9-6-2011 12:00 AM

The Analysis of Extreme Synoptic Winds

David A. Gatey, *The University of Western Ontario*

Supervisor: Dr. Craig Miller, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Civil and Environmental Engineering

THE ANALYSIS OF EXTREME SYNOPTIC WINDS
(Thesis format: Monograph)

by

David Allan Gatey

Graduate Program in Civil and Environmental Engineering

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor:

.....
Dr. C.A. Miller

Supervisory Committee:

.....
Dr. H. Hangan

.....
Dr. E. Savory

Examiners:

.....
Dr. W.J. Braun

.....
Dr. H.P. Hong

.....
Dr. M. Kasperski

.....
Dr. G.A. Kopp

The thesis by

David Allan Gatey

entitled:

The Analysis of Extreme Synoptic Winds

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

.....
Date

.....
Chair of the Thesis Examination Board

Abstract

Time histories of wind speed and direction from 394 surface observation stations were obtained to calculate synoptic 50-year return period wind speeds for 11 countries in Europe. Preliminary investigation indicated wind speed differences along national borders were successfully reduced by application of a simple consistent methodology to wind speed data. This study considers the ideal methodology for calculating synoptic 50-year return period wind speeds.

Wind speed data requires standardisation through quality control measures, exposure correction and adjustment for disjunct sampling. A quality control algorithm was successfully applied to identify shifts of monthly mean wind speeds and data conversion issues. Three exposure correction models were evaluated and two-layer models were found to perform better than internal boundary layer models. The differences arise as a result of how the models adapt to an upstream change of roughness. Furthermore, an empirical model was formed to correct observations at stations which were not recording measurements hourly.

Extreme value analyses were carried out using a robust estimator to fit the extreme value distribution type I to storm and yearly maxima. The latter was found to provide more consistent results. Comparison of the resulting 50-year return period wind speeds to existing literature found that several regions were in good agreement, while other regions exhibited similar spatial variation but greater magnitudes. The differences in magnitude were partially related to exposure correction methods, thus lending support to the importance of a single consistent methodology. Directional factors were calculated and subsequently grouped into six regions exhibiting similar directional characteristics.

Background wind fields were calculated from mean sea-level pressure data using the geostrophic approximation and consideration of other improved approximations, however, variations in

the pressure field led to a breakdown of the methodology. A background 50-year return period wind field calculated from upper-level wind fields was significantly lower than surface wind speed estimates due to spatial and temporal smoothing. Finally, assimilation of the 50-year return period wind speeds from surface observations and the background wind field was explored using the Bratseth scheme for statistical interpolation. The Bratseth scheme provided an overall 50-year return period wind speed map.

Keywords: 50-year return period wind speeds, homogenising wind data, exposure correction, boundary layer, extreme value analysis, outliers, Bratseth, synoptic winds, European wind map

Acknowledgements

Work on a thesis quite often finishes as it starts, with many long hours and sleepless nights spent in the office. I could not have completed this thesis without the support of the technical and support staff of the Boundary Layer Wind Tunnel Laboratory, my friends and, of course, family. Special thanks needs to be given to particular individuals who have been there on a daily basis since the beginning.

I thank my supervisor, Dr. Craig Miller, who took me on as a undergraduate research student over 7 years ago. He has always had an open door and the time to provide guidance and support throughout the course of this thesis. In addition, he provided me with the opportunity and means to speak at various international conferences and to get involved with, and occasionally distracted by, several interesting projects. I truly believe I am a more well-rounded and stronger researcher because of his guidance.

I would like to thank Tom Mara and Zach Taylor for the many interesting discussions and late evenings bouncing ideas off one another, or simply unwinding, over a good Islay.

To Laurie, thank you for your love and firm determination to ensure I finished this thesis on time. Your support, particularly over these final months, has been remarkable.

To both of my grandfathers, Cor Dorssers and Allan Gatey, who taught me the value of good, hard work and the importance of a strong education. I know that they would have been truly proud of what I have accomplished and, with both of their passings occurring in these final months, I dedicate this thesis to them.

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

1.1 Overview

One of the primary concerns in the field of wind engineering is the design and response of structures subjected to strong winds in the atmospheric boundary layer (ABL). The ABL is the region of the atmosphere in contact with, and influenced by, the surface of the earth. Due to the interaction of the atmosphere with the surface, structures contained within the ABL are subjected to both mean and fluctuating wind effects. In engineering design codes for structures, the pressure and associated loading applied to these structures is inherently derived from a 50-year return period wind speed at 10 m height in open-country exposure (e.g. National Building Code of Canada, NBCC; Eurocode). Moderate differences in wind speed can result in greatly varying wind loads due to the squared relationship between wind speed and pressure. A logical conclusion is that accurate estimation of the 50-year return period wind speed is crucial to all wind susceptible structures and structural elements. To appropriately consider the effects of wind action on structures, not only the 50-year return period wind speed, but also the wind climate, requires consideration. The wind climate

provides additional information about important factors such as direction, duration, spatial variation and storm mechanisms. Arguably the wind climate is still as significant today as it was when defined as the 'critical link' in the wind loading chain by Davenport (1983, 1999).

50-year return period wind speeds are typically calculated from historical surface wind speed records which may have been corrected for the effects of non-standard anemometer heights and upstream changes of surface roughness and terrain, then statistically analysed using an appropriate extreme value distribution. Since the early years of wind engineering, various methodologies have existed to calculate wind speeds for design purposes. At the inaugural Wind Effects on Structures and Buildings conference in the United Kingdom (UK), which would later become the International Conference on Wind Engineering (ICWE), Davenport (1963) presented a gradient level British wind speed map and Shellard (1963) presented a second British wind speed map based on surface-level gust wind speed measurements. The complexity of the problem has grown over the last 50 years as alternative methods now exist for each step in the process of calculating 50-year return period wind speeds. The consequences are clearly illustrated through the attempt to create a unified 50-year return period wind speed map of Europe for the original Eurocode. Along national borders, severe discontinuities exist between 50-year return period wind speeds for neighbouring countries. Although recent work has indicated a possible reduction of the largest differences by modifying the underlying methodology (Sacré *et al.*, 2007), accurate estimation of wind speeds used for design has been shown to be critical. The discontinuities can be significant and they explicitly define the underlying problem in wind engineering design: the various techniques used by different nations can often result in significantly different 50-year return period wind speeds.

In this work, the theory that a consistent methodology will reduce observed differences is considered and the ideal process of determining 50-year return period wind speeds resulting

from synoptic-scale events, events on the order of hundreds to thousands of kilometres such as pressure systems, is explored. The former can be examined by the use of a simplified technique provided it is consistently applied. The latter, however, requires the investigation of multiple techniques, several of which remain largely unaddressed within the wind engineering community. A number of the issues discussed in this work include the type and quality of data, surface corrections, disjunct sampling, directionality, extreme value analysis, outliers and data assimilation. The methods available for considering each of these aspects contribute differently to the final prediction of 50-year return period wind speeds. In instances where multiple accepted options have been established within the wind engineering or meteorological communities, the provided analyses compare the feasibility and performance of each approach. Alternatively, where existing techniques are lacking, new methods are proposed by means of empirical models using current data or by the extension of existing models. The purpose is to establish a consistent and ideal methodology for analysing extreme synoptic winds, and to apply this methodology to generate a unified synoptic 50-year return period wind speed map of Europe. The remainder of the current chapter identifies the objectives of the study based on a review of the methodologies currently employed to calculate synoptic 50-year return period wind speeds. The review provides a necessary framework through which discontinuities between current and suggested practices will be identified and examined.

1.2 Objectives

Many current design codes throughout the world remain based on analyses carried out in the 1990's, prompting researchers in recent years to explore various improved methods of developing 50-year return period wind speeds. The United States, while greatly improving the methods for mixed climates, notably the estimation of wind speeds for design in hurri-

cane prone regions, have left the remainder of the country relatively unchanged in ASCE 7-10. Concerns remain regarding the methodology in which stations were amalgamated when forming superstations, and a lack of a proper representation of the varied extreme wind climate throughout the central regions of the country (Simiu *et al.*, 2003). Despite these concerns, the majority of the country is governed by a single 3-second gust wind speed of 40 m/s in ASCE 7-10 (Vickery *et al.*, 2010) which originates from Peterka and Shahid (1998) for ASCE 7-98. Similarly, the NBCC has had no substantial review of the process for estimating 50-year return period wind speeds since 1995 (Yip and Auld, 1993; Yip *et al.*, 1995). Recently An and Pandey (2007) examined 50-year return period wind speeds in the province of Ontario and have recommended improved statistical methods for updating the 50-year return period wind speed maps within the NBCC. However, the 50-year return period wind speeds published in NBCC 1995 still exist in original form in NBCC 2005.

Europe presents a unique opportunity to further this research, as recognition of the discontinuities between national borders in Europe has spurred a renewal of interest towards the improvement of existing 50-year return period wind speed maps for both Europe and its individual nations. The extreme wind climate in central and northwestern Europe is dominated by the passage of extratropical cyclones (ETCs), or depressions. Depressions typically originate near the Icelandic Low, the northern pole of the North Atlantic Oscillation, and track northeast across central Europe. In the summer months, thunderstorms occur throughout Europe with activity typically peaking for central and northern Europe in July (Boucher, 2005). Gomes and Vickery (1978) recommended separating extreme wind climates for individual analysis, which is still considered an essential requirement for calculating wind speeds for design (Holmes *et al.*, 2005; Kasperski, 2009). As the focus of the current study is on synoptic 50-year return period wind speeds, methods for detecting and extracting thunderstorms are discussed in Section 3.1.4. With the exception of Kasperski

(2002), separation of wind climates has rarely been carried out in national wind mapping studies in Europe.

Individual stations within an observation network are often subject to potentially erroneous measurements, varying temporal frequency of measurements and are influenced by physical surroundings such as land cover. These differences must be corrected for in order to create a consistent wind speed map; this process is herein referred to as the standardisation of surface wind speed data. A crucial step in the process of estimating 50-year return period wind speeds is the initial removal of spurious observations which often result in greatly overestimated, and potentially unrealistic, design requirements. Surprisingly, despite the obvious importance of detecting such records, little to no discussion is provided by the majority of researchers on whether such observations were detected or even sought. Two exceptions are Sacré *et al.* (2007) and Burton and Allsop (2009b), where the former implement a detection technique used by Météo-France for climatic parameters and the latter provide details of an identification process. Available quality control methods are further addressed in Section 3.1.

In contrast to the lack of documentation of quality control methods, the correction of wind speed measurements for exposure to an open-country equivalent is quite common. The difficulty of selecting an appropriate exposure correction model can arise as there are several different methods available. Miller *et al.* (2001) and Burton and Allsop (2009a) use an internal boundary layer model, while Kristensen *et al.* (2000) and Sacré *et al.* (2007) use the commercially available software Wind Atlas Analysis and Application Program (WAsP). Others such as George (2006) use the geostrophic drag law, typically within a two-layer boundary layer model. As several options are presently available to appropriately correct wind speed measurements for both site and upstream surface roughness, each approach will provide different corrections for a single location depending on the surrounding roughness and fetch. Thus, the available models require comparison in order to identify the most ap-

propriate method. Section 3.3 provides these necessary comparisons to improve upon the existing models.

Throughout the operational lifetime of an observation station it is not uncommon for a change to occur in the temporal frequency of measurements, most notably with the switch from manual to automated observation systems. Disjunct sampling has only been accounted for by Frank (2001) and Larsén and Mann (2009), while remaining unaddressed by the majority of the wind engineering community. A new empirical model is derived in Section 3.4 and is subsequently compared to existing alternatives, despite its absence from discussion in the literature. General characteristics regarding the duration of wind storms and the relative intensity of hours adjacent the peak is inferred from the results of the proposed disjunct sampling model.

The methodology for the statistical analysis of extreme wind observations is covered in Chapter 4. A summary of the types of datasets and extreme value techniques currently used in calculating 50-year return period wind speeds is provided in Table 1.1. It is clear that both the type of sampling and extreme value distribution vary among studies. Holmes

Author	Region	Sampling	Distribution	Directional
Yip *et al.* (1995)	Canada	Annual	Gumbel	No
Żurański and Jaśpińska (1996)	Poland	Annual	Gumbel	Yes
Peterka and Shahid (1998)	US	Annual	Gumbel	No
Kristensen *et al.* (2000)	Denmark	Two Months	Gumbel	Yes
Frank (2001)	Denmark	Annual	Gumbel	No
Miller *et al.* (2001)	UK	Storm	Gumbel	Yes
Kasperski (2002)	Germany	Storm	GEV(III)	Yes
Sacré (2002)	France	Annual	Gumbel	No
Miller (2003)	Northern Europe	Storm	GPD	No
George (2006)	UK	Annual	Gumbel	No
An and Pandey (2007)	Canada	Storm, r-LOSS	Gumbel	No
Sacré *et al.* (2007)	France	Storm	Gumbel, GPD	No
Burton and Allsop (2009a)	Ireland	Annual, Storm	Gumbel	Yes
Larsén and Mann (2009)	Multiple	Annual	Gumbel	Yes

Table 1.1: Statistical methods used in wind engineering to analyse extreme winds

et al. (2005) do not recommend the use of annual extremes only, as other significant wind events are often not represented. When considering annual maxima, the Gumbel extreme value distribution, a special form of the generalised extreme value distribution (GEVD), has typically been selected as the appropriate statistical model. To increase the number of wind events considered for statistical analysis, the GEVD has also been applied to other block maxima such as independent storms. Alternative methods which researchers have utilised in an attempt to increase the number of events considered for statistical analysis are r-largest ordered statistics (r-LOSS) and the generalised Pareto distribution (GPD). In conjunction with increased extreme sampling, Kasperski and Geurts (2005) and Kasperski (2009) suggest the focus should shift to the consideration of storm duration and the magnitude of wind speeds in hours adjacent the peak. The associated statistical field of study is titled 'dependant extremes' and its full analysis is beyond the scope of the current work. (e.g. Fawcett and Walshaw, 2008).

Synoptic 50-year return period wind speeds have historically been calculated using a single source of data, such as, time histories of surface wind speed observations, upper-air wind speed measurements obtained from radiosonde, or wind speeds calculated from mean sea-level pressure (MSLP) fields. Most analyses utilise surface wind measurements or MSLP fields as radiosonde are not typically released in severe extreme wind conditions. Wind speeds calculated from MSLP fields are typically based on the assumption of geostrophic balance, thus frictionless flow occurs between straight, parallel isobars and is assumed to be representative of a wind field sufficiently far from the effects of the surface. The geostrophic drag law is utilised to calculate an estimate of the associated surface wind speed from the geostrophic wind components (Miller, 2003; Larsén and Mann, 2009). A review and analysis of the various methods for deriving upper-level wind fields from available re-analysis data is covered in Chapter 5.

Data assimilation techniques allow data obtained from multiple sources to be appropriately

merged. As previous studies have considered 50-year return period wind speeds calculated exclusively from surface observations or MSLP fields, a truly unique approach is considered in Chapter 6. The possibility of utilising the 50-year return period wind field derived from upper-air measurements in Chapter 5 to supplement the 50-year return period wind speeds calculated from surface measurements in Chapters 3 and 4 is explored in Chapter 6. Lastly, the conclusions of the study are presented in Chapter 7.

Chapter 2

Background

2.1 Preliminary Study

A preliminary study was carried out to investigate whether observed differences between 50-year return period wind speeds along national borders in Europe could be reduced by applying a simple, yet consistent, methodology and was published by Gatey and Miller (2007). Five examples of regions in Europe where differences exist between national borders are identified in Figure 2.1, with each region shown in detail in Figure 2.2. The plot of each region has two portions; the left panel represents peer-reviewed 50-year return period wind speeds which have been published in conjunction with the methodology, and the right panel contains a comparison of the latest national building codes or national annexes (NAs) to Eurocode. Sources of the various 50-year return period wind speed values shown in Figure 2.2 are summarised in Table 2.1. All 50-year return period wind speeds are 10-minute mean wind speeds at 10 m height for a roughness length of 0.05 m with the exception of the values for the UK. Both Miller *et al.* (2001) and BS6399-2 (based on Cook and Prior, 1987) provide hourly-mean wind speeds at 10 m height for a roughness length of 0.03 m,

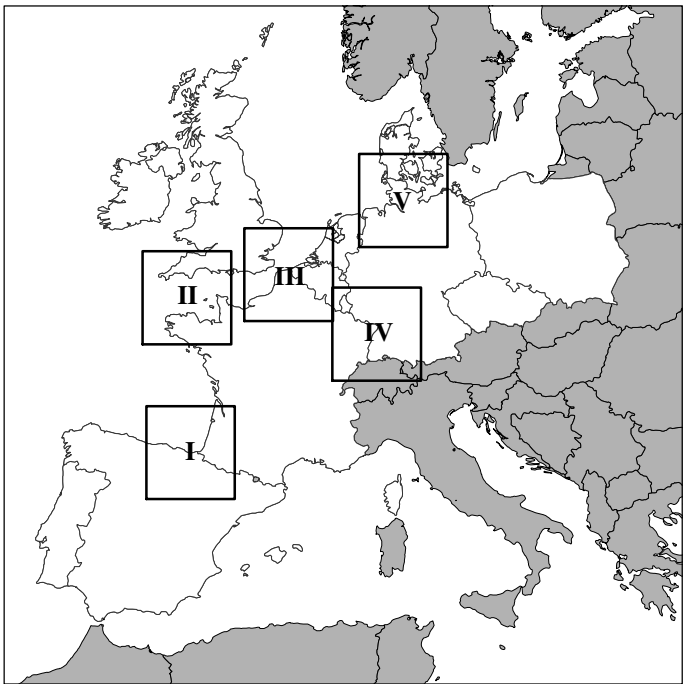
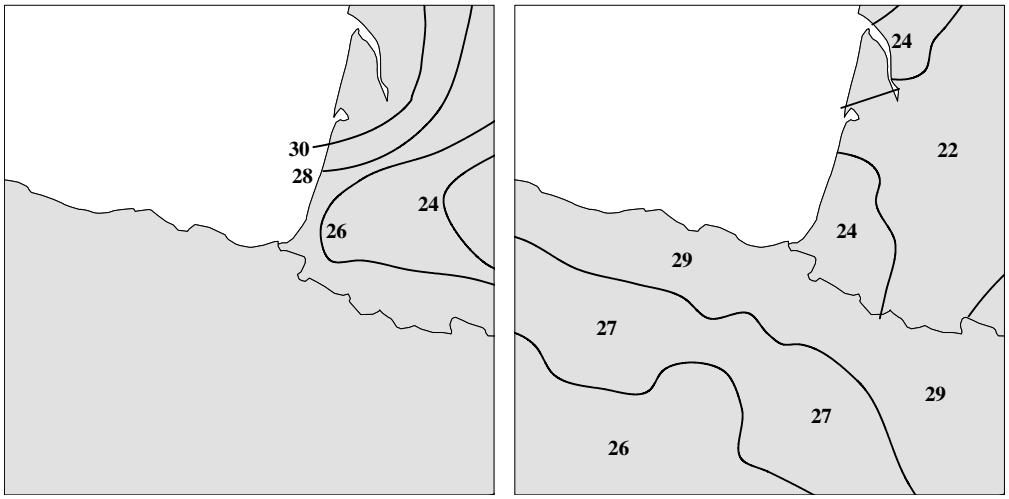


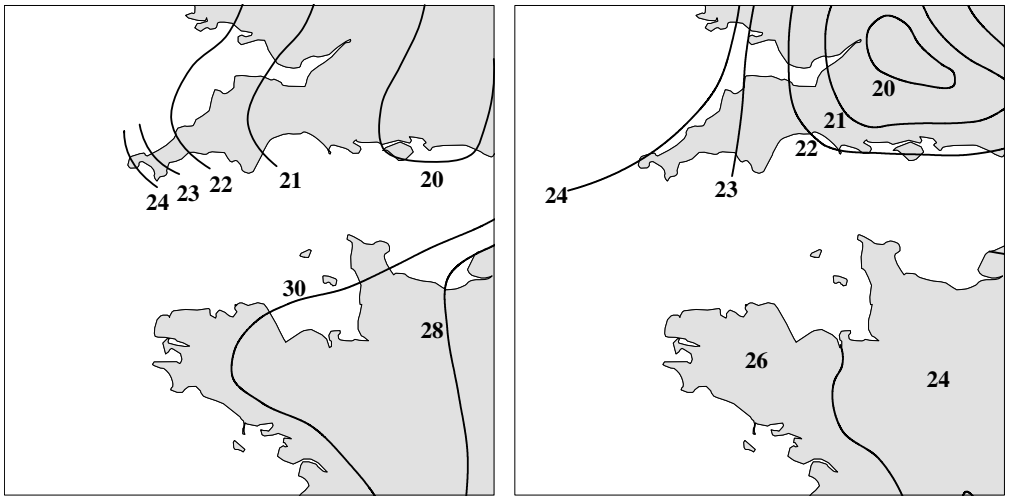
Figure 2.1: Regions of interest in Europe for the preliminary study

however, the combined correction from hourly-mean to 10-minute mean wind speed and a roughness length of 0.03 m to 0.05 m is generally taken as unity. Visual inspection suggests the German NA is based on Kasperski (2002) and the values for the French NA are possibly based on the methodology described by Sacré *et al.* (2007) who cite a reduced value (26 m/s) for a station on the French coast near the Belgian border which matches the French NA. Several discrepancies are noted here:

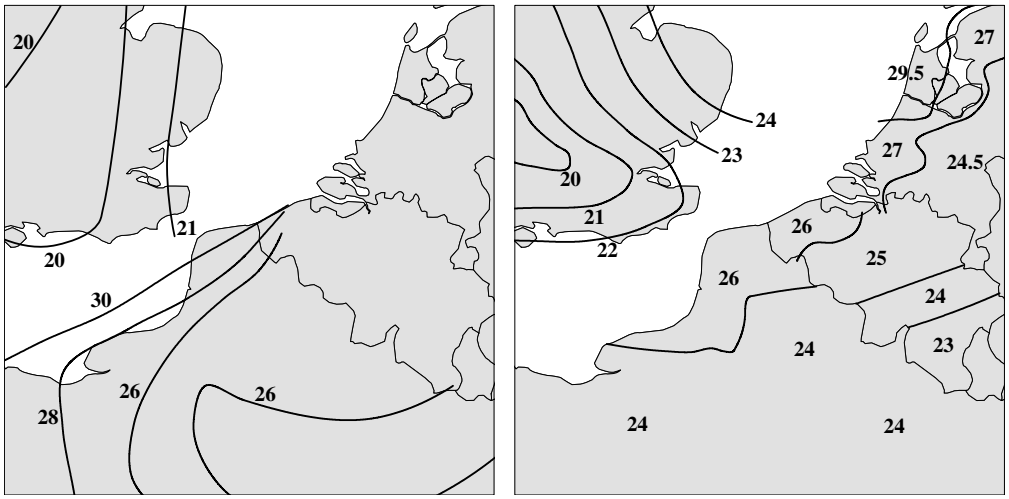
- 50-year return period wind speeds in France have been considerably reduced.
- The existing French values provided a better match to the Spanish code along the France-Spain border. Since the border between France and Spain follows the Pyrenées, a true difference between wind speeds may exist.
- Differences have been reduced between the UK and France, however, there is still a considerable difference between the UK and both France and Belgium.
- Differences have been reduced along the border between France and Germany, however the values lack continuity which is likely a result of differences in contouring.



(a) Section I: France-Spain border



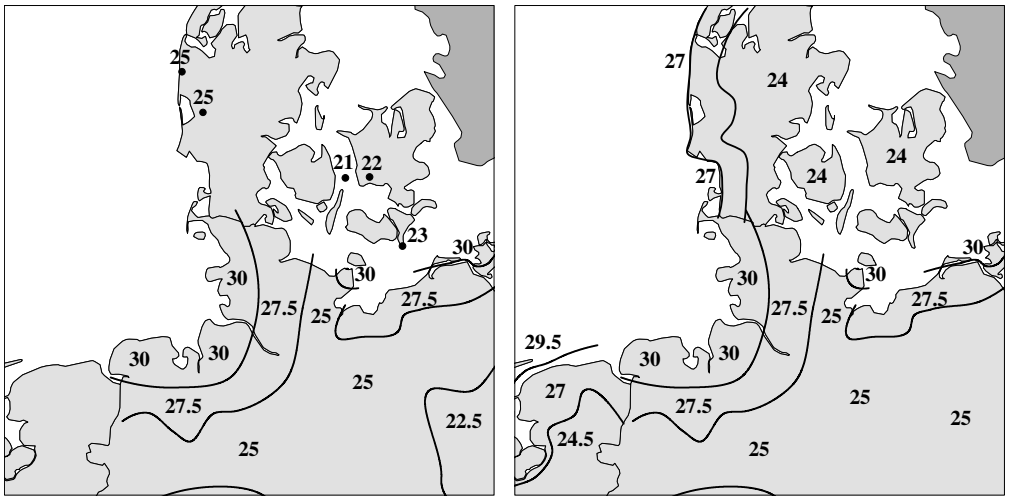
(b) Section II: English Channel (West)



(c) Section III: English Channel (East)

Figure 2.2: Comparison of European 50-year return period wind speed maps, published (*left*) and National Annexes (*right*), *continued on next page*

(d) Section IV: France-Germany border



(e) Section V: Denmark-Germany border

Figure 2.2: Comparison of European 50-year return period wind speed maps, published (*left*) and National Annexes (*right*)

Country	Published	Code
Spain	–	DB SE-AE 2009
France	Sacré (2002)	NF EN 1991-1-4 NA
UK	Miller *et al.* (2001)	BS6399-2
Belgium	–	NBN EN 1991-1-4 NA
Netherlands	–	NEN EN 1991-1-4 NA
Germany	Kasperski (2002)	DIN EN 1991-1-4 NA
Denmark	Kristensen *et al.* (2000)	DK EN 1991-1-4 NA

Table 2.1: Sources of 50-year return period wind speeds by country

- Large differences continue to exist between Denmark and Germany for 50-year return period wind speeds.

The majority of the NAs were unavailable or incomplete at the onset of the preliminary study, thus the study sought to reduce the discrepancies between the published values in Sections II through V. The NA values will provide additional comparison for the current study despite a lack of documentation regarding the underlying methodology for several nations.

Global surface summary of the day data was obtained from the National Climatic Data Center (NCDC), a division of the National Oceanic and Atmospheric Administration (NOAA). Basic quality control checks found that multiple years of data, at several stations, had to be omitted as a result of errors stemming from improper unit conversion. The data was processed using a basic traditional methodology for calculating 50-year return period wind speeds. Corrections were applied for anemometer height and the surface roughness representative of the site, thereby neglecting upstream effects, using the Deaves and Harris model (further discussed in Section 3.3.3). Annual maxima were extracted for each station and estimates with probability of exceedance of 0.02 were calculated using the Gumbel distribution (defined in Section 4.1). One of the major recommendations arising from the extreme value analysis in the preliminary study was for future investigations to consider methods for statistically identifying outliers, either spurious or relating to a longer return period. A method for statistically identifying outliers appearing in a dataset is presented in Section 4.3.

The conclusions for the study included improved correlation across the English Channel and along the France-Germany border. Differences along the border between Denmark and Germany were less than those resulting from Kasperski (2002) and Kristensen *et al.* (2000), but still displayed notable discrepancies. Overall, the unified process found better

correlation across borders than various complex procedures being used individually by each nation. By identifying an ideal methodology for calculating 50-year return period wind speeds, estimates from the current study can be directly compared to Figure 2.2 to evaluate whether discrepancies arise from over- or under-estimation by a single nation or if a compromise can be established between existing values.

2.2 Surface Data

The dataset obtained for the current work consists of global hourly and synoptic observations from the Integrated Surface Database (ISD), digital dataset DS-3505, managed by the NCDC. The ISD contains two fixed length and three variable length sections. The former two are the control and mandatory data sections and the latter three are the additional data, remarks data and element quality data sections. The observations of interest are the mean wind speed and wind direction (mandatory data section), present weather identifiers and supplementary wind observations (additional data section) and observer comments (remarks section). Full ISD documentation can be found in NCDC (2010).

The present weather identifiers and supplementary wind observations are recorded with varying temporal frequency. For example, the latter contains the recorded gust wind speed and/or gust wind direction, which are typically only recorded when the velocity exceeds a predetermined threshold (which may also vary temporally). In some instances the gust wind speeds may not be recorded at all. The remarks section occasionally contains additional mean wind speed or gust wind speed measurements, as well as comments regarding thunderstorms or other relevant meteorological observations. Many of the remarks contained within the section follow the practices outlined in NOAA (2005).

Mean wind speeds may also vary in sampling duration; common measurements include

hourly, continuous 10-minute mean, 10-minute mean before the hour and a 2-minute mean. Similarly, gust wind speeds may be block or continuous measurements. The directive of the World Meteorological Organization (WMO) mandates that anemometers should be located in open-country exposure and standard averaging times of 10-minutes and 3-seconds should be used for the mean and gust wind speeds respectively (WMO, 2008). As such, the mean and gust wind speeds documented in the ISD have been assumed to be a 10-minute mean wind speed recorded during the 10-minute period prior to the hour and a nominal 3-second gust wind speed observed throughout the hour. Measuring the 10-minute mean exclusively on the hour neglects 50-minutes of available wind observations. Continuous 10-minute mean wind speeds are thereby preferable, however, 10-minute mean wind speeds measured during the 10-minute period prior to the hour are available for the longest periods. Observation networks such as the Automated Surface Observing System (ASOS) have now been recording continuously since 1998 and in the next decade will provide enough data to improve estimates of the true continuous 10-minute mean wind speeds in the US. Many stations currently report two times per hour, in these instances the observation nearest the hour is selected to maintain consistency throughout the record. When analysing a wind event it is important to consider the averaging time which will best represent the type of system or storm. Giving consideration to the characteristics of the European synoptic wind climate, the 10-minute mean wind speed is selected as the data type which best represents the synoptic events of the region. The mean wind speed is also the most commonly available wind measurement thereby ensuring sufficient data for analysis.

In addition to the errors identified by Gatey and Miller (2007), Burton and Allsop (2009b) found that annual extremes extracted from a NCDC dataset do not directly match values obtained from records provided by local authorities. The problem may arise from mixing of recorded mean and gust wind data, the data from the local authorities may be of a different averaging time, or an observation may have simply been incorrectly transcribed. Identifi-

cation of these observations, and treating them in a consistent manner, is an important step in the overall methodology and is discussed further in Section 3.1.

2.3 Station Selection

The countries of interest located within Europe are Portugal, Spain, France, Ireland, UK, Belgium, Netherlands, Germany, Denmark, Czech Republic and Poland. For each WMO station present within the ISD, an inventory is available which indicates the available number of observations per month on a annual basis. As a preliminary classification, all WMO stations within the countries of interest are queried and classified as primary, secondary or tertiary based on the number of complete years and number of observations per month. For the latter criteria, thresholds of 500 and 200 observations per month are selected as the minimum number of observations as they correspond to approximately a 28 day month containing 18 and 8 observations per day respectively. The criteria for the three classifications are identified in Table 2.2. The inventory was originally parsed for observations commencing January 1970, however, few stations were found to have records in the period 1970-1972. As such, a consistent start date of 1973 was selected. Stations were mapped and hand-selected to ensure adequate spatial coverage where available, with preference given to stations of higher classification. The resulting 394 selected stations are shown in Figure 2.3 and a listing is provided in Appendix A. The entire data record is obtained from the ISD for each selected station and the relevant observations discussed in Section 2.2 are extracted for analysis.

Classification	Criteria
Primary (I)	Minimum 25 years of data and 500 observations/month
Secondary (II)	Minimum 25 years of data and 200 observations/month
Tertiary (III)	Minimum 15 years of data and 200 observations/month

Table 2.2: Classifications and associated criteria for station selection

Figure 2.3: Map of selected WMO stations

Chapter 3

Standardisation of Wind Speed Data

50-year return period wind speeds are to be representative of wind speeds recorded at 10 m height in open-country exposure. Standardising wind speed data provides a consistent framework for engineers to adjust standardised values to better suit the conditions at a specific location. In the current work, standardisation is carried out in two steps. First, wind speed data is assessed by quality control algorithms, from which erroneous measurements and discontinuities in time histories are identified in a consistent manner. Second, wind speed measurements are modified to allow observations at different locations to be directly compared irrespective of site characteristics or sampling frequency. The latter process is commonly referred to as homogenisation. The following sections address quality control algorithms, atmospheric boundary layer models, site exposure corrections and disjunct sampling.

3.1 Quality Control Measures

3.1.1 Background

When carrying out a statistical analysis of extremes, if left undetected spurious observations can greatly affect 50-year return period wind speed estimates. Methods for identifying such observations require attention, particularly for the current work where a subset of maxima is sought and errors are known to exist within the ISD. Quality control measures can also aid in determining if annual mean wind speeds are consistent over the entire data record or if considerable discontinuities exist. Identification of a shift may indicate changes of instrument location, height or local surroundings.

Throughout the operational lifetime of a synoptic weather station, it is not uncommon for the height or location of the anemometer to change, or for instrumentation to be upgraded. Quite often a meteorological agency will upgrade all instrumentation for a given date, although in practice it may be several months before the upgrades are operational at every location. Each of the possible changes will have a specific effect on the measured wind speed. A change of anemometer height will be most apparent from a shift of the mean wind speed for all directions, however, a change of anemometer location can be much more complex. If the old and new site have very similar exposures for all directions, a change may not be detectable unless otherwise documented. Conversely, a new location where the exposure differs directionally from the prior location will experience changes of mean wind speed and gustiness in the affected directions. The gustiness at a site can be represented by the gust factor, the ratio of the gust wind speed to mean wind speed. Lastly, a change of the anemometer or chart recorder should not be apparent from the mean wind speed records, as any changes to the response length or gust averaging time should be filtered out over a sufficient averaging period. The change will most likely affect measurements of the gustiness.

Documentation is occasionally available from meteorological agencies identifying dates of location or instrumentation changes, if unavailable, it is important to identify these changes to at least be aware that they exist.

At present, only two 50-year return period wind speed studies identify the methods chosen to pre-process meteorological data. Sacré *et al.* (2007) implement a detection method called PRODIGE which is described by Caussinus and Mestre (2004) and used by Météo-France. The PRODIGE algorithm is applied to annual mean data from stations assumed to be influenced by the same climatic conditions. Each series is assumed to be a combination of a climate effect, station effect and random white noise (Caussinus and Mestre, 2004). In performing the analysis across multiple stations, the climatic effect should be spatially redundant, thus allowing differences due to station effects to be identified. A penalised log-likelihood procedure of Caussinus and Lyazrhi (1997) is used to detect change-points and outliers, and least-square estimates of the climate and station effects are used to correct the data. To vastly reduce the number of hypotheses and computational time, a preliminary stage consisting of pairwise comparisons of the station record with those from neighbouring locations is required. These difference series, in conjunction with the penalised log-likelihood procedure and manual synthesis, are used for pre-selection of change-points and outliers in monthly or annual mean data. A procedure which can be automated without requiring a pre-selection stage is preferred for the number of stations considered here. In addition, the resulting 'corrected' data may not be appropriate for wind observations particularly those exhibiting significant directional variation. For an anemometer sited in relatively open terrain, a change of height will likely have an isotropic influence on the wind speeds, in this situation a single station effect will be appropriate for all wind speeds measured at the location. However, if the location of an anemometer has changed, then differences in surface roughness may only occur for certain azimuths, therefore, the true station effect may exhibit anisotropy. Caussinus and Mestre (2004) note the PRODIGE

model requires better detection of gradual changes and of breaks when the shift of the mean is less than the standard deviation.

Burton and Allsop (2009a) pre-process wind speed data in an attempt to identify individual observations for removal. Mean wind speeds greater than 20 m/s and three times greater than both adjacent mean hourly observations are classified as errors or thunderstorms, both of which are excluded from a synoptic climate analysis. For a number of regions in Europe of interest in the current work, the 50-year return period wind speed is less than 27 m/s, as was shown in Figure 2.2. A representative set of annual maxima will likely contain a subset of extremes which are less than 20 m/s, therefore, the maxima contained within the subset are not necessarily validated e.g. an annual maximum of 18 m/s is not considered by the pre-processing scheme. Such a situation is likely to arise, particularly when evaluating directional extremes where maxima occurring from non-dominant wind directions are, in general, substantially lower than dominant wind directions. A lower threshold of 15 m/s suggested by Burton and Allsop (2009b) is likely more appropriate. Ideally, a method which can be applied to ensure the quality of every hourly wind observation is desired. The data can then be used to accurately derive the parent distribution if desired and, more importantly, ensures the validation of maxima regardless of the strength of the wind climate.

A combination of the aforementioned quality control measures are required. Quality control measures are divided into two levels for the current work, global or high-level quality control measures and localised or low-level quality control measures. The global quality control measures include physical limits checks based on DeGaetano (1997) and a homogenisation algorithm by Domonkos (2011). Localised quality control measures are based on the wind speed variability checks established by DeGaetano (1997) for hourly surface measurements and are expanded to consider additional information relevant to the current analysis. Additional quality control measures for a range of meteorological parameters are discussed by Graybeal *et al.* (2004).

3.1.2 Global Quality Control Measures

Several global quality control measures of ranging complexity are considered. The most basic checks are for observations flagged as suspect or failing the ISD quality control described by Lott (2004). Other minor tests include physical limit checks identified by DeGaetano (1997). The checks ensure the mean wind speed is less than the gust wind speed, the wind direction is a multiple of 10 and that measurements obtained during calm periods are properly transcribed.

The majority of observations within the ISD are reported on the hour, 10-minutes prior or 10-minutes after. Observations are prioritised in this order and, where multiple records exist, the highest ranking observation is selected resulting in a single observation for each hour. After culling the redundant observations, each station is tested against the tertiary classification outlined in Table 2.2 to ensure all stations meet the stated basic requirements. If a year of observations fail to meet this criterion, the year is removed from the record to ensure sufficient temporal resolution of observations throughout the year. Insufficient observations may be due to downtime associated with anemometer damage, measurement system replacement, freezing, or a site change. For several measurement stations in Germany, there are years where no records are reported at the expected reporting times, instead reporting was performed at 44 minutes past the hour. Thus, if a year is to be omitted due to insufficient measurements at the expected reporting times, a procedure is implemented to scan the previously parsed observations to evaluate whether there is a specific reporting minute which satisfies the minimum observation criterion.

A method for identifying shifts in the annual mean wind speeds by Caussinus and Mestre (2004) was discussed in Section 3.1.1. Although the resulting corrected time histories may not be appropriate for wind speed data, the detection of change-points and outliers is of interest. An automatic homogenisation procedure based on the PRODIGE method is given

by Domonkos (2011). The method, an adapted Caussinus-Mestre algorithm for networks of temperature series, is herein referred to as ACMANT. The PRODIGE and ACMANT methods are both recommended based on standardised benchmark tests carried out by Venema *et al.* (2011). The ACMANT method contains two detection schemes, the main detection is based on annual means and summer-winter differences and secondary detection is used to identify short-term inhomogeneities. Domonkos (2011) notes that the radiation intensity affects temperature measurement and as a result, anomalies between time series during the summer naturally exhibit larger inhomogeneities than during the winter. Thus, the secondary detection scheme is based on monthly mean values and includes a harmonic annual cycle to account for the seasonal variation of inhomogeneity size. Theoretically, a similar cycle potentially exists for mean wind speed observations at mid- and upper-latitude locations as a result of the seasonal variation of surface roughness. In the winter months, deciduous plant species will shed their foliage and the surface is typically covered by snow. Under such conditions, wind speeds likely exhibit greater spatial correlation, reducing anomalies between time series, than during summer months when anemometers are affected by varying types and degrees of local vegetation. The correlation between monthly mean wind speeds as a function of month is shown for Bournemouth Airport, Hurn, UK and Caen-Carpiquet Airport, Carpiquet, FR in Figure 3.1 and indicates the assumption is appropriate.

The ACMANT method was carried out on six overlapping regions of approximately 100 stations as shown in Figure 3.2. For stations in overlapping regions, the detected change-points were found to be consistent between runs since the ACMANT method bases inhomogeneity detection on differences between stations whose series are well correlated. Figure 3.3 shows a typical time series and the identification of detected shifts by year.

To evaluate the impact of the previous assumption regarding the variation of inhomogeneity size by month, the monthly mean wind speeds for stations within the zone encompassing

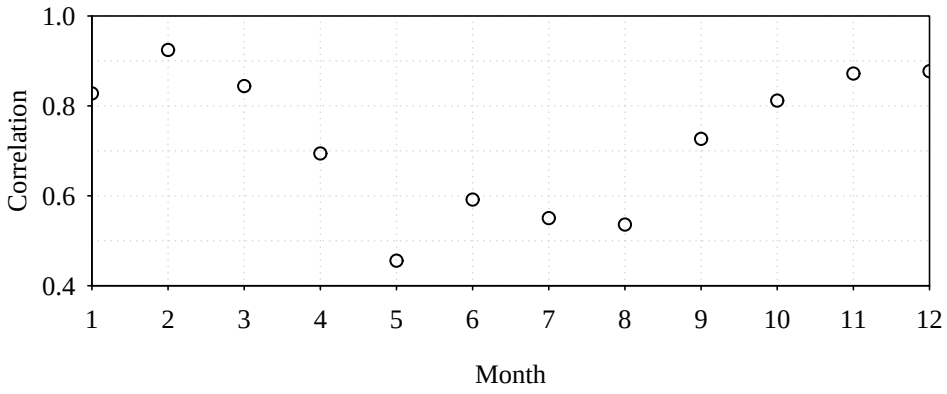


Figure 3.1: Correlation of monthly mean wind speeds between Bournemouth Airport, UK (WMO 03862) and Caen-Carpiquet Airport, FR (WMO 07027)

the UK and Ireland was shifted by six months and the ACMANT method was carried out a second time. Between the original and modified time series, approximately 75 percent of the change-points were common within a couple months between the two datasets. Analysis of the change-points detected for winter and summer months between the two sets indicates a difference of only six percent, thus, even in a severe case where the annual cycle is assumed out of phase, the algorithm does not greatly impact the detected change-points for the current data.

Detected change-points were compared to what limited documentation on location, height and instrumentation changes could be found from the websites of various meteorological agencies, particularly Met Éireann, Koninklijk Nederlands Meteorologisch Instituut (KNMI) and Deutscher Wetterdienst (DWD); publications, Cappelan and Jorgensen (1999), Traup and Kruse (1996) and Verkaik (2001); and other available resources. Shifts verified to be related to a change of location or height are further considered in Section 3.3 for exposure correction.

Gatey and Miller (2007) showed that there were data conversion issues in the ISD for several stations based on mean wind speed data and fluctuations of the associated gust factor. In examining the detected change-points by country it was found that for several

Country	Number of Stations	Period
Belgium	12 of 13	1996-1997
Germany	25 of 81	1999-2000
Ireland	10 of 12	1996-1998
Netherlands	16 of 16	1996-1997
Portugal	5 of 6	1997-2000

Table 3.1: Summary of conversion errors by country

exhibiting possible conversion related errors are listed in Table 3.1.

3.1.3 Localised Quality Control Measures

The localised quality control algorithm is based on the excessive wind speed variability checks described by DeGaetano (1997). The general procedure is to extract a subset of data centred about an observation of interest and compare the maximum two hour wind speed difference in the subset, to the difference between the current observed wind speed and all other observations in the subset. Several criteria are to be met to identify a wind speed as suspect, including:

- The difference between the current observation and all observations in the subset must be greater than the maximum two hour difference in the subset.
- The difference between the current observation and all observations in the subset, neglecting the hour prior and after, must be greater than 7.7 m/s (15 kt).
- The current wind speed must be at least 3.1 m/s (6 kt) greater than the neighbouring hours.

In addition, if a strong shower or thunderstorm is present at the time of measurement, or occurred in the previous or following hour, then the wind speed is accepted as a valid measurement.

PW Identifier	Type	Localised Quality Control	Thunderstorm
AU[1 – 9]	Automated	2, 3	2
AW[1 – 4]	Automated	18, 26, 42, 44, 46, 48, 58, 63, 66, 68, 73, 76, 80 : 97	26, 90 : 97
AY[1 – 2]	Manual	8, 9	9
AZ[1 – 2]	Automated	8, 9	9
MV[1 – 7]	Manual	1, 2	1
MW[1 – 7]	Manual	17, 18, 19, 25, 26, 27, 29, 59, 64, 65, 67, 69, 74, 75, 80 : 98	17, 19, 29, 91 : 98

Table 3.2: Present weather and thunderstorm identifiers

Several identifiers exist within the additional data section of the ISD which summarise the weather at the time of observation. The relevant present weather (PW) indicators for values related to strong showers and thunderstorms are shown in Table 3.2. The remarks section of the ISD often contains additional comments indicating the presence of thunderstorms. NOAA (2005) indicates a standard reporting style of TSB*bb*E*ee* where *b* and *e* represent the hour relating to the start and end of a thunderstorm. Further investigation found that it is more common for a shorthand form of 'thunderstorm' to be reported. In general, the following word segments are capable of identifying the shorthand entries: STORM, THUN, T/ST, TSTO, TSTR. If an observation meets all rejection criteria, and one of the mentioned weather phenomena did not occur, the observation is removed.

The remarks section of the ISD often contains manual entries identifying the wind direction, wind speed and, occasionally, gust wind speed measurements. Several of the entry formats follow those outlined in NOAA (2005), while other formats have been identified manually. Overall, eight different entries have been identified. Representing the wind direction, wind speed and gust wind speed as *d*, *w* and *g* respectively, the formats are: *dddww*/*gg*KT, *dddww*G*gg*KT, *dddwww*KT, *dddww*KT, *ddww*KT, MAX *gg*KT, MAX*gg*KT and *ww*KT. If an observation is rejected by the localised quality control algorithm, and an alternative value exists in the remarks section in one of the above formats, the alternative value is selected and the quality control check is repeated for the associated time step.

The quality control algorithms by DeGaetano (1997) were intended for complete hourly wind records, however, recommendations for application to a three-hour sampling interval were provided. In the current work, three assumptions regarding the average number of observations per day, calculated by month, have been made. Months where a median of 18 observations per day or greater exist, are treated in the same manner as those having 24 hourly observations. A median of $9-17$ observations per day typically indicated observations were being recorded during the daily operational hours of a site. Lastly, a three hour sampling interval was assumed if the median number of observations per day was between 6.5 and 9. Table 3.3 contains the temporal interval considered in the local quality control subset, the corresponding subset size and the minimum number of observations in the subset required for the localised quality control check to be applied to the current observation. The subset interval for hourly and three-hour sampling intervals are provided by DeGaetano (1997). The minimum subset size is considered here to ensure sufficient measurements are present to adequately evaluate the current observation. The criteria for measurements occurring throughout operational hours is defined in relation to the criteria for the hourly and three-hour sampling intervals.

The current procedure relies on accurate and complete records of PW identifiers. In the instance where the PW identifiers are incomplete, a short duration high-intensity wind (HIW) event may be rejected if it was not recorded that an associated incident, such as a thunderstorm, was present. Given the focus on synoptic winds, the potential rejection of a measurement associated with a HIW event is not of great concern as the events are typically driven by convective mechanisms. In addition, it is possible for two closely spaced

Median Obs./Day	Subset Interval	Max. Subset Size	Min. Subset Size
$18 \leq$	-11.5 to $+12.5$ hours	24	16
$9-17$	-18.5 to $+21.5$ hours	14-39	14
$6.5-9$	-18.5 to $+21.5$ hours	14	11

Table 3.3: Localised quality control criteria

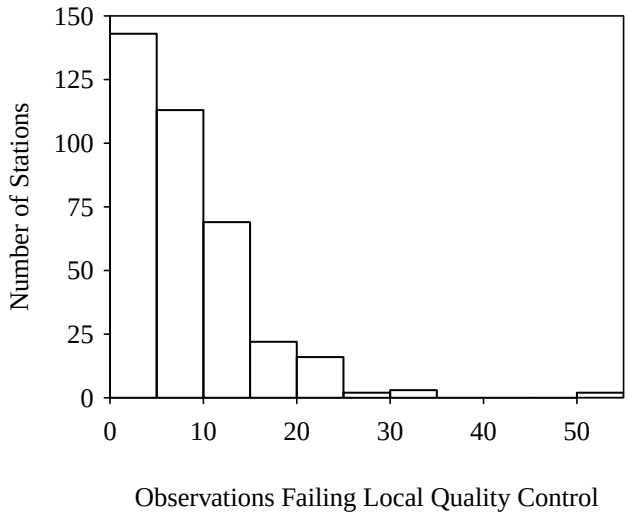


Figure 3.4: Distribution of observations failing local quality control checks

erroneous measurements to shelter one another from identification as shown by DeGaetano (1997). Overall, it was found that the average number of rejected observations per station was 9 with a maximum of 52 for Sniezka, Poland. The distribution of rejected observations per station is shown in Figure 3.4. The localised quality control method described here to validate individual wind speed measurements is an important analysis used to identify spurious observations. The method can be applied to an entire time history or to a set of extracted maxima, provided the required temporally adjacent observations are available.

3.1.4 Thunderstorm Identification

The quality control algorithms discussed in Sections 3.1.2 and 3.1.3 focus on validating the obtained wind records. Since the focus of the current work is on synoptic winds, thunderstorm observations require removal from the dataset. Convective storms are typically short duration events with single cell thunderstorms lasting approximately 30 minutes (Holton, 2004). Lombardo et al. (2009) show that approximately 91 percent of the recorded thunderstorms at a location in the US had a duration of 2 hours or less. Depending on when

a thunderstorm ended during an observation hour, a one hour event could affect the local wind climate, and associated measurements, for up to two or three hours prior or after the recorded observation. Thus, the adjacent two hours, both before and after a reported thunderstorm hour are extracted as contaminated synoptic observations and archived for future investigation. The weather indicators relevant to strictly thunderstorms are listed in Table 3.2.

The criteria for reporting a thunderstorm will often vary between national meteorological organisations. A thunderstorm classification may be based on hearing thunder or seeing lightning, however, there is no guarantee that the wind speed recorded is representative of a thunderstorm wind. Future algorithms to identify thunderstorms could benefit by giving consideration to the temperature and the ratio of gust to mean wind speed at a location, provided all three measurements are available with sufficient temporal resolution. Such a scheme would allow thunderstorms to be detected in the absence of present weather identifiers and to verify the reverse.

3.2 Atmospheric Boundary Layer

3.2.1 Background

In extreme synoptic wind events, convection is negligible and mechanical turbulence production due to wind shear and surface roughness governs. Under such conditions, the assumption of a neutral or near-neutral atmosphere is acceptable. Modelling of the wind profile of the ABL for engineering application, where wind loading is a concern, is conveniently simplified by the assumption of atmospheric neutrality. The lowest portion of the ABL is known as the surface layer, which is defined by the law of the wall. Within the

surface layer, the law of the wall indicates that the velocity scales as a function of height and surface roughness, which holds for approximately the bottom 10 percent of the ABL (Simiu and Scanlan, 1996). The upper or outer layer of the ABL is defined by the velocity defect law which is a function of the velocity at the top of the boundary layer and the height of the boundary layer. An intermediate layer between the surface and outer layers is assumed to exist in which both layers overlap. Blackadar and Tennekes (1968) use asymptotic similarity theory (AST) to equate these two layers from which the log-law and geostrophic drag law equations are derived. The log-law for neutrally stable conditions is commonly expressed as

$$u(z) = \frac{u_*}{\kappa} \ln\left(\frac{z}{z_0}\right) \tag{3.1}$$

near the surface, where $u(z)$ is the mean velocity at height z, u_* is the friction velocity, κ is the von Karman constant and z_0 is the roughness length. The geostrophic wind speed (G) is calculated from the geostrophic drag law as

$$G = \frac{u_*}{\kappa} \sqrt{\left[\ln\left(\frac{u_*}{f z_0}\right) - A\right]^2 + B^2} \tag{3.2}$$

where f is the Coriolis parameter, and, A and B are generally treated as dimensionless parameters, although they have been identified as functions of stability and boundary layer height (Zilitinkevich and Esau, 2002). A summary of values selected by researchers to represent the two parameters in Equation 3.2 is provided by Zilitinkevich (1989)

In the field of wind engineering, the power-law was originally used to model the mean wind profile within the ABL due to its simplicity and improved estimates away from the surface. The power-law is expressed as

$$u(z) = u_{ref}\left(\frac{z}{z_{ref}}\right)^{1/\alpha} \tag{3.3}$$

where u_{ref} is a wind speed at reference height z_{ref} and α is an empirically derived exponent dependant upon exposure. The Engineering Sciences Data Unit (ESDU) standard for over 30 years is a semi-empirical boundary layer model proposed by Deaves and Harris (1978). The model is based on the assumption of neutral steady-state conditions and AST. Empirical estimates are used to determine four theoretically derived constants which yield a parabolic profile for a majority of the boundary layer. The mean wind profile of the Deaves and Harris model is expressed as

$$u(z) = \frac{u_*}{\kappa}\left[\ln\left(\frac{z}{z_0}\right) + 5.75\left(\frac{z}{z_h}\right) - 1.878\left(\frac{z}{z_h}\right)^2 - 1.333\left(\frac{z}{z_h}\right)^3 + 0.25\left(\frac{z}{z_h}\right)^4\right] \tag{3.4}$$

where z_h is the height of the boundary layer. Despite its widespread acceptance in the wind engineering community, the model has never gained popularity in other fields due to a lack of publishing in peer-reviewed journals outside of the community.

Gryning *et al.* (2007) recently proposed a boundary layer model which has been validated using data obtained from several tall towers. The model contains three separate wind profiles corresponding to the neutral, stable or unstable conditions. In addition, the model considers length scales appropriate for the surface, middle and upper layers of the boundary layer. The model was validated against 160 m and 250 m towers by Gryning *et al.* (2007) and, a 300 m tower and the Leipzig, Germany wind profile up to 900 m by Peña *et al.* (2010).

Observations from the Leipzig wind profile as re-examined by Lettau (1950), along with wind profile fits from the aforementioned models, are shown in Figure 3.5. The fits are calculated using $u_* = 0.65$ m/s from Lettau (1950) and $z_0 = 0.1$ m as determined by Peña *et al.* (2010). The exponent for the power-law is estimated from

$$\alpha = \frac{1}{\ln(10/z_0)} \tag{3.5}$$

where the wind speed at 850 m is used as the reference value in Equation 3.3. Considering the uncertainty associated with the measurements obtained from the 28 pilot balloons used to form the Leipzig wind profile, the majority of ABL models perform quite well with the exception of the log-law which, as given by Equation 3.1, is only valid near the surface. Due to the more extensive and transparent validation techniques performed by Gryning *et al.* (2007) and Peña *et al.* (2010), and the flexibility of the model to allow for potential consideration of stable and unstable boundary layers, the model proposed by Gryning *et al.* (2007) is selected for modelling the mean wind profile in the current work.

3.2.2 Gryning ABL Model

The ABL profile model proposed by Gryning *et al.* (2007), herein referred to as the Gryning ABL model, is based on the assumption that there are three components which contribute to the length scale, one each for the surface, middle and upper layers of the ABL. Length scales in the surface layer appropriately scale with height, while those in the middle layer are assumed to be dependant on stability. The influence of the length scale in the upper layer is thought to be relatively unknown and as a result, the length scale is assumed to decrease to zero as a function of the distance from the top of the boundary layer, similar to scaling in the surface layer. The mean wind profile under neutral conditions is given by Gryning *et al.* (2007) as

$$u(z) = \frac{u_{*0}}{\kappa}\left[\ln\left(\frac{z}{z_0}\right) + \frac{z}{L_M} - \frac{z}{z_h}\left(\frac{z}{2L_M}\right)\right] \tag{3.6}$$

where L_M is the length scale in the middle layer of the ABL. Length scales in the middle layer were parametrised by equating Equations 3.2 and 3.6. The approximation for the length scale in the middle layer under neutral conditions was determined empirically based on the dependence between u_{*0}/fz_0 and u_{*0}/fL_M from the data obtained at the two towers

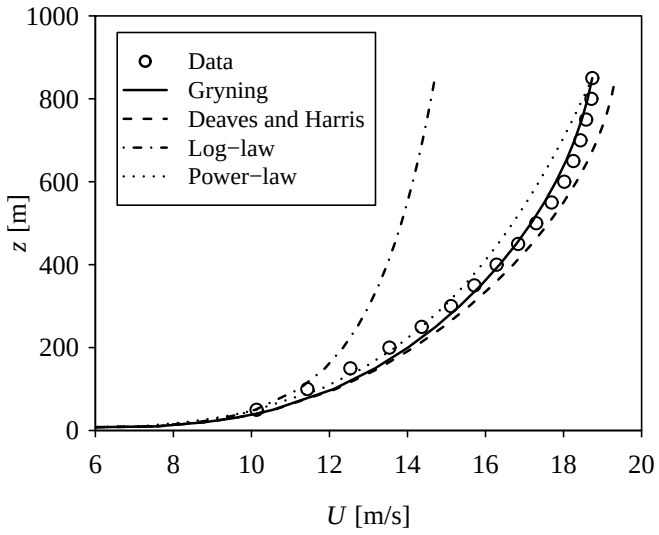


Figure 3.5: Fits of various wind profiles to the Leipzig wind profile

analysed by Gryning *et al.* (2007) and is expressed as

$$\frac{u_{*0}}{fL_M} = -2\ln\left(\frac{u_{*0}}{fz_0}\right) + 55. \tag{3.7}$$

The boundary layer height can be approximated from the Rossby-Montgomery formula, given as

$$z_h \approx cu_{*0}/f \tag{3.8}$$

where c is a constant (Rossby and Montgomery, 1935). Gryning *et al.* (2007) suggest a value of $c = 0.1$, however, Peña *et al.* (2010) found a constant value of 0.15 provided the best fit to the Leipzig wind profile. A similar value was derived by Peña *et al.* (2010) assuming the dimensionless parameter A was equal to 1.7.

3.3 Heterogeneous Exposure Correction

3.3.1 Background

The majority of stations providing the greatest temporal duration of measurements are typically located on airfields. Anemometers located on airfields are generally sited in near open-country exposure, however, after initial placement local disturbances may arise due to the expansion of airport facilities or hangers. In addition, airfields are typically placed sufficiently far from urban or suburban centres, although urban sprawl may result in reduced fetch between the two regions. Correction of observed wind speeds (u) to standardised values (u_B) allows for the direct comparison of final predicted wind speeds. The ratio of the observed to corrected value is referred to as the correction factor. For the current work, observed values are standardised to the WMO standard height (10 m) and open-country exposure ($z_0 = 0.05$ m). The effects of heterogeneous exposure can be corrected to standard open-country exposure by several different approaches of ranging complexity.

The combined effects of upstream heterogeneous exposure can be summarised in terms of an effective roughness length at the site. The effective roughness length can be calculated from the wind profile, turbulence intensity of the wind, or gustiness at a site. Barthelmie *et al.* (1993) compared several of these methods, including determination of the roughness length from aerial photographs, and concluded that the roughness lengths derived from aerial images gave acceptable results while gustiness- and turbulence-derived roughness lengths resulted in underestimated wind speeds. The standard deviation method produced the greatest errors, however, Vega (2008) notes that the standard deviation method was applied to observations further from the surface ($\sim$ 34 m) for a threshold wind speed intended for measurements at 10 m. The measurements used for the standard deviation method are likely a result of combined mechanical and thermal mixing. In addition, a single factor was

selected to represent the ratio of the standard deviation of the wind to the friction velocity which likely requires additional consideration of the impact of the transfer functions of the measurement instruments on the wind spectrum.

By deriving roughness lengths from wind measurements, one can easily evaluate changes of roughness over time or identify periods potentially affected by local sheltering. Wind profile-derived roughness lengths are ideal, however, in practice the availability of such information is limited to locations where towers are instrumented at multiple heights. Roughness lengths derived from the standard deviation of the wind, or turbulence intensity, require measurements obtained with a sampling frequency which is much greater than the hourly measurements obtained for the current work. Stations with sufficient data to perform such an analysis are available from ASOS for the US, whereas a Europe-wide equivalent is currently unavailable. Roughness lengths can be derived from gustiness if sufficient gust wind speed data is available for a specified location using methods proposed by Wieringa (1973, 1976) and Beljaars (1987). Application of the latter to ASOS data is discussed by Masters *et al.* (2010).

The gustiness-derived roughness length method is preferred for the current work as it does not require knowledge of the exact anemometer location throughout the operational lifetime of the station. The method requires knowledge of the gust factor which, when calculated for each wind azimuth, requires the assumption that the gust wind speed is from the same, or very nearly same, direction as the mean wind speed. Thus, the resulting gust factor is representative of exposure and terrain conditions in the mean wind azimuth. Verkaik (2000) compared the two methods and found Wieringa's gustiness model to produce larger correction factors than Beljaars' model. For a mean averaging time of 10-minutes, Wieringa's model produced correction factors approximately 3 to 10 percent larger than Beljaar's model. Neither of the models were able to fully account for a change in the measuring chain. Overall, the study identified the differences between the two models but was unable

to evaluate the accuracy.

Preliminary investigation found that the majority of stations in the current work provide insufficient gust wind speed observations at high mean wind speeds in the non-dominant wind directions. In addition, documentation defining the characteristics of existing measuring chains is largely unavailable. Thus, to maintain the ideology that a consistent methodology is an important factor in the estimation of 50-year return period wind speeds, the only alternative for the current work is to calculate a correction factor based on land use land cover (LULC) information. A simple approach to correcting exposure at a site to the reference height and roughness length considers only the height of the anemometer and roughness length at the site. However, it is known that changes in upstream roughness can have a significant impact on the wind profile. Letchford *et al.* (2001) have shown that upstream roughness effects can be significant, thus, correcting by both direction and distance is desirable. More sophisticated models exist which calculate the effects of non-uniform surface roughness on the boundary layer by both distance and direction.

Two approaches which model the effects of non-uniform surface roughness are: internal boundary layer (IBL) models and two-layer (TL) models. IBL models assume that following a change of roughness, a new internal boundary layer will form that is in local equilibrium with the new exposure. Above this layer the wind profile remains in equilibrium with the upstream roughness. Between these two profiles exists a transition region where the wind profiles from above and below are assumed to blend smoothly. With knowledge of the model equations, surrounding roughness and anemometer height, it is possible to calculate the effect of these changes of roughness on the wind profile at a particular location. A model accounting for upstream roughness effects was proposed by Deaves (1981) based on the Deaves and Harris ABL model described in Section 3.2. Cook (1985) proposed a series of correction factors which account for various aspects influencing a site and is here referred to as the Deaves and Harris IBL model. The Deaves and Harris IBL model was

shown by Cook (1997) to greatly reduce the directional variance in comparison to assuming uniform exposure; reductions in variance were not as significant for sites affected by topography.

TL models focus on two regions, a lower surface layer and an upper layer. The height at which these two regions meet is identified as the mesolevel or blending height (Wieringa, 1976, 1986). TL models consider the effects of the local roughness within the surface layer and the mesoscale roughness, which is representative of a larger region, within the upper or macrolevel. Through consideration of these roughness effects, model equations, and a given height, the resulting wind speed at a particular location can be determined.

The TL model of Wieringa (1986) can be used to predict the wind speed profile over multiple changes of roughness. An approximation to the area within which the roughness length contributes to the surface flux was incorporated into the model by Verkaik (2003). The combination of the TL model proposed by Wieringa (1986) and the footprint approximation of Verkaik (2003) is here referred to as the Hydra TL model. Verkaik (2003) reported relative errors in wind speed predictions of 10 to 15 percent but expected better results after revision of the model. Many of the stations for which measured and estimated wind speeds were compared, were located on the coast and some large distance inland. The model was validated by using observations from multiple locations to calculate a macrolevel windfield, then interpolating the wind speed to a separate site and comparing the estimate to recorded data.

Over large changes of roughness, Verkaik (2003) found IBL models tend to predict more abrupt adjustments to wind speeds within short distances, less than two kilometres, than TL models. In a separate investigation by Letchford *et al.* (2001), wind records from two hurricanes were compared for two anemometers on the same airfield. The records showed similar readings until changes in the wind direction occurred. Discrepancies in the records,

due to the change in wind direction, were related to the effects of nearby woodland. Simulating this change of roughness using the Deaves and Harris model, Letchford *et al.* (2001) found that the model typically overestimated the speed up associated with a change from a rough to smooth surface, predominantly with mean wind speeds as opposed to gusts.

In this section, the Deaves and Harris IBL and the Hydra TL models as proposed by Cook (1985) and Verkaik (2003), respectively, are summarised briefly. An alternative model is proposed which combines the concept of the TL model with the Gryning ABL wind profile. For several stations in the Netherlands, correction factors are calculated for 30-degree sectors from gustiness- and LULC-derived effective roughness length estimates using each of the three correction methods. The differences between the gustiness- and LULC-derived correction factors are calculated and compared across the three models.

3.3.2 Methodology

Correction factors are calculated for the three exposure correction models based on gustiness-derived effective roughness lengths and compared to the respective correction factors calculated by the three models based on LULC-derived roughness lengths for a subset of stations in the Netherlands. For stations in the KNMI observation network, documentation of the individual station histories, including site, anemometer and chart recorder changes are provided by Verkaik (2001). Beljaars' gustiness model is selected to estimate the effective roughness lengths for several reasons. Since the method allows for selection of the input wind spectrum, the transfer functions for the anemometer and chart recorder are easily applied directly to the wind spectrum prior to application of the gustiness model. Accounting for the transfer functions in this way allows continuous or discretely sampled observations to be considered through application of the appropriate transfer function. Conversely, Wieringa's gustiness model is not applicable for discretely sampled data, which comprises

the majority of the KNMI observation network since converting to automatic weather stations in the 1990's (Verkaik, 2000). Through application of Beljaars' gustiness model to the observations at 13 stations, effective roughness lengths are calculated by direction. The effective roughness lengths are calculated for 30-degree sectors having a minimum of 30 gust factor values with a minimum mean wind speed of 10 m/s. For each wind direction satisfying this criterion, the log-average of the estimated effective roughness lengths for the sector is calculated. This method provides a single effective roughness length for each direction and is applied for each measurement chain identified at the location. If different measurement chains were consecutively employed at a consistent mast location, then the weighted average of the effective roughness length determined from the measurement chains was calculated. This final step ensures a single effective roughness length is evaluated for each direction and mast location at the station.

To calculate the LULC-derived correction factors, a geographic information systems (GIS) tool has been developed to sample a LULC database by distance and direction from a site of interest. The Coordinate Information on the Environment (CORINE) LULC database is a Europe-wide 44 class LULC raster database developed by the European Environment Agency (EEA) and is selected for the current work. The database was created by compiling LULC databases from individual countries using a common framework. In theory, the common framework should provide fairly consistent results. The version of the CORINE database used in this study has a pixel resolution of 100 m by 100 m where each pixel contains an integer value representing one of the 44 LULC classes. Taking into consideration both the various land cover nomenclature set out by the CORINE documentation (Bossard *et al.*, 2000) and a review of roughness lengths by Wieringa (1993), roughness lengths are assigned to the 44 LULC classes as shown in Table 3.4. Gatey and Miller (2007) and Sacré *et al.* (2007) also use the CORINE LULC database to assign roughness lengths. A few variations exist in the assignment of roughness lengths between the studies, however, the

z_0 [m]	CORINE LULC Classes
0.003	331, 332, 335, 511, 512, 521, 522, 523
0.005	333, 422, 423
0.010	412, 421
0.034	142, 231
0.05	124, 213, 321, 322, 334, 411
0.06	211, 212
0.10	121, 132
0.15	211, 222, 223, 241, 242, 243
0.25	131, 133, 323
0.3	244
0.5	122, 123, 141, 324
0.6	112
1.0	111,331,312,313

Table 3.4: LULC roughness assignments

values overall show good agreement. The selected CORINE database is for the year 2000, which is the only version to include the UK. Given the database is representative of a single year, the current analysis will not account for any changes of roughness over time.

Due to the coarseness of the CORINE LULC pixels, radial bands of the sampling grid have a minimum thickness of 0.2 km. Bands nearest the site are smaller, while outer bands are much thicker. The sampling grid extends to a distance of 55 km and bands range in thickness from 0.2 km to 2.5 km. Twelve 30-degree sectors are considered for which the logarithms of the roughness lengths are area-averaged for each segment of the band within a sector. The process provides a single effective roughness length representing the cell, which has been shown to be an adequate estimate of the effective roughness length by Taylor (1987). An example of the sampling grid and underlying CORINE LULC raster for Den Helder Airport, De Kooy, NL is shown in Figure 3.6. The following sections provide brief summaries of the equations for each of the models discussed here.

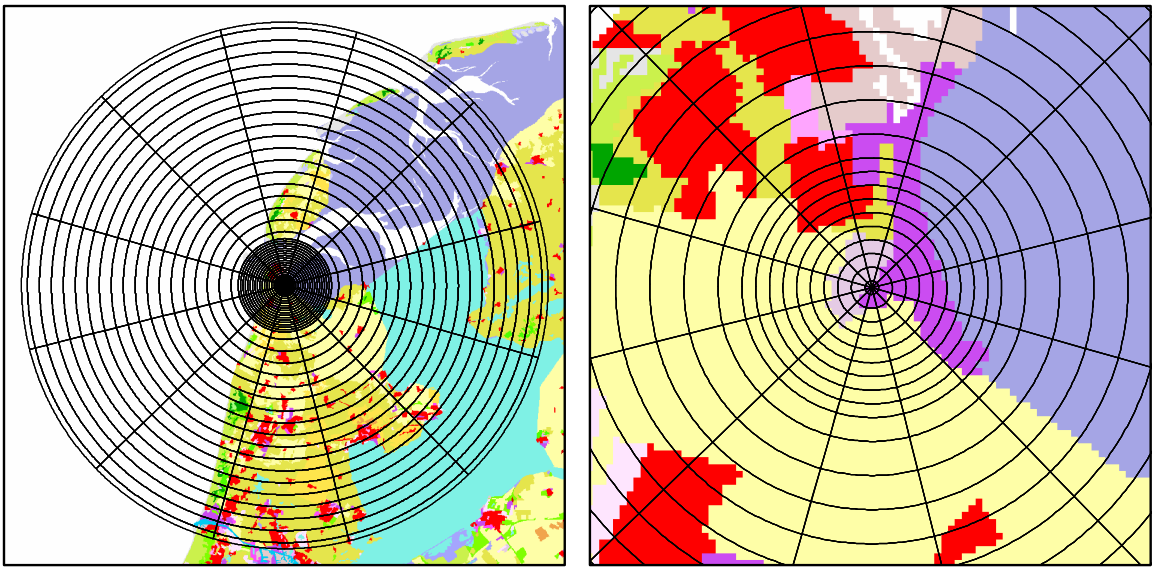


Figure 3.6: CORINE LULC and sampling grid (meso- and local-scale) for Den Helger Airport, De Kooy, NL (WMO 06235)

3.3.3 Deaves and Harris IBL Model

Parameters known as S-factors are utilized to account for several aspects including height (S_Z), site exposure (S_E) and fetch (S_X). It is the net effect of these factors which allows for manipulation of wind speed data. The mean wind speed as calculated using the Deaves and Harris S-Factors is given by the equation

$$u = S_{X\{n-m\}}...S_{X\{c-b\}}S_{X\{b-a\}}S_{Z\{a\}}S_{E\{a\}}u_B \tag{3.9}$$

where u is the mean wind speed at anemometer height z, u_B is the basic wind speed at height z_B for roughness length $z_{0,B}$, n denotes the furthest upstream roughness, and a is the site roughness. The S-factors as given by Cook (1985) are determined from the following expressions

Height Factor

$$S_Z = \frac{\ln(z/z_0) + 5.75(z/z_{h,0}) - 1.875(z/z_{h,0})^2 - 4(z/z_{h,0})^3/3 + (z/z_{h,0})^4/4}{\ln(10/z_0)} \tag{3.10}$$

where $z_{h,0}$ is the height of the ABL for roughness length z_0.

Exposure Factor

$$S_E = \frac{[\ln(z_{h,B}/z_{0,B} + 2.79]\ln(10/z_0)}{[\ln(z_{h,0}/z_0 + 2.79]\ln(10/z_{0,B})} \tag{3.11}$$

where $z_{h,B}$ is the height of ABL for the basic roughness.

Fetch Factor

$$S_{X\{j\to i\}} = \left[1 - \frac{\ln(z_{0,j}/z_{0,i})}{0.42 + \ln m_0}\right] \frac{\ln(10/z_{0,i})}{\ln(10/z_{0,j})} \frac{S_{E,j}}{S_{E,i}} \tag{3.12}$$

where j and i denote the upstream and downstream profiles respectively, and m_0 is calculated as

$$m_0 = 0.32X/\left[z_{0,i}(\ln m_0 - 1)\right] \tag{3.13}$$

where X refers to the distance of the site downstream of the change of roughness.

The height of the ABL is calculated through iteration of the friction velocity and the mean velocity profile. One final calculation is necessary to determine which IBL is controlling the wind speed. A simple approach assumes that the transition region between IBLs has zero thickness. For smooth to rough transitions, the wind speed is determined from the smallest value of the profiles as calculated by Equation 3.9, and for a rough to smooth transition, the largest value.

3.3.4 Hydra TL Model

Prior to the calculation of wind speeds, the Hydra model first approximates a local and mesoscale footprint. For each cell, the drag coefficient is calculated from

$$C_d = \frac{\kappa}{\ln(z_{bh}/z_{o,eff})} \tag{3.14}$$

where z_{bh} is the blending height chosen by Wieringa (1986) to be 60 m and $z_{0,eff}$ is the log-averaged effective roughness length of the cell. The average drag coefficient for each 30-degree sector is then expressed as

$$C'_d = \frac{\Sigma W(x/D) C_d}{\Sigma W(x/D)} \tag{3.15}$$

where weighted averages are calculated based on the distance of a cell from the site as

$$W(x, D) = \exp(-x/D) \tag{3.16}$$

where x is the distance from the site and D is given by Verkaik (2003) to be 600 m and 3000 m for local and meso-scale footprints respectively.

Two differences exist between the calculation of the effective roughness lengths in the current work and the Hydra model:

- Verkaik (2003) proposed 5-degree wide sectors for which C'_d is smoothed using a weighted moving average considering three sectors on either side of the centre sector. In the current work, all sectors are 30-degrees and are not smoothed as the resolution of the LULC grid is significantly more coarse.
- Verkaik (2003) used the Charnock relation to account for the drag relation for water, whereas a single fixed value has been assumed here.

Once an effective drag coefficient has been calculated for the local and mesoscale footprints, the effective roughness lengths are solved for. From the local effective surface roughness, the mesolevel wind is calculated as

$$u_{meso} = u \times \left[\frac{\ln(z_{bh}/z'_{o,s})}{\ln(z/z'_{o,s})} \right] \tag{3.17}$$

where u is the surface wind speed at anemometer height z and $z'_{o,s}$ is the local effective roughness length. Incorporating the mesoscale effective roughness length ($z'_{o,m}$) the friction velocity is given by

$$u_{*,m} = \kappa u_{meso} / \ln(z_{bh}/z'_{o,m}) \tag{3.18}$$

Once the friction velocity is known, the macroscale wind can be calculated as

$$U_{macro} = (u_*/\kappa) \left[\ln\left(\frac{u_*}{f z_{o,m}} \right) - A \right] \tag{3.19}$$

where $A = 1.9$ and $B = 4.5$ for neutral conditions. By reversing this process and assuming that $z'_{o,m} = z'_{o,s} = 0.05$ m and $z = 10$ m, the basic wind speed can be calculated and an appropriate correction factor determined.

3.3.5 TL Model: Gryning ABL

The effective local and meso-scale roughness lengths are calculated in the same manner as in the Hydra model. In place of extrapolating the measured wind speed using the log-law and geostrophic drag law in Equations 3.17 and 3.19, the Gryning ABL wind profile is utilised. The boundary layer height is calculated iteratively based on the observed wind speed and the effective local roughness length using the equations outlined in Section 3.2.2. The velocity at the blending height can then be calculated and the boundary layer height is

once again determined iteratively using the mesoscale roughness length. Once the boundary layer height is known, the velocity at the top of the boundary layer is calculated, and is assumed to the be representative of the gradient wind speed. The boundary layer height for the base roughness length is then calculated iteratively assuming an equivalent gradient wind speed. Once the boundary layer height has been determined, the equivalent 10 m wind speed and related correction factor are calculated

3.3.6 Beljaar's Gustiness Model

Given mean wind speed $\overline{u}$, the gust wind speed $\hat{u}$ can be written as

$$\hat{u} = \overline{u} + (\hat{u} - \overline{u}) \tag{3.20}$$

from which the gust factor is defined by

$$\frac{\hat{u}}{\overline{u}} = 1 + \frac{\hat{u} - \overline{u}}{\overline{u}} \tag{3.21}$$

$$= 1 + \frac{\sigma_u}{\overline{u}} \frac{\hat{u} - \overline{u}}{\sigma_u} \tag{3.22}$$

$$= 1 + g\frac{\sigma_u}{\overline{u}} \tag{3.23}$$

where g is the peak factor or normalised gust and $\sigma_u/\overline{u}$ is the turbulence intensity. The mean value of the peak factor is equal to

$$g = (2 \ln \nu T)^{1/2} + 0.5772(2 \ln \nu T)^{-1/2} \tag{3.24}$$

where T is the averaging period of the mean wind speed and ν is the zero-crossing or cycling rate defined as

$$v^2 = \frac{\int_0^\infty f^2 S(f)\,\mathrm{d}f}{\int_0^\infty S(f)\,\mathrm{d}f} \tag{3.25}$$

where $S(f)$ is the wind spectrum and f is the frequency. For the current work, the Kaimal (1978) wind spectrum is selected since Beljaars (1987) showed that at a height of 10 m, the Kaimal spectrum fit the available data from the Cabauw tower better than the wind spectrum of Højstrup (1982). The Kaimal spectrum contains high frequency, transition and low frequency regions. The high frequency region scales with height, while the low frequency region instead scales with the height of the boundary layer (Verkaik, 2000). The Kaimal wind spectrum is a function of normalised frequency and is given as

$$\frac{fS_u(f)}{u_*^2} = \begin{cases} \left(12 + 0.5\left|\dfrac{z_h}{L}\right|\right)^{2/3} \dfrac{n_b}{1 + 3.1n_b^{5/3}} & n < \dfrac{3z}{2z_h} \\ \left(1 + 0.75\left|\dfrac{z}{L}\right|^{2/3}\right) 0.48(2n)^{-p} & \dfrac{3z}{2z_h} \le n < \dfrac{1}{2} \\ \left(1 + 0.75\left|\dfrac{z}{L}\right|^{2/3}\right) 0.3n^{-2/3} & n \ge \dfrac{1}{2} \end{cases} \tag{3.26}$$

where $n = fz/\overline{u}$, $n_b = fz_h/\overline{u}$ and

$$p = \ln\left[0.44\frac{(12 + 0.5|z_h/L|)^{2/3}}{1 + 0.75|z/L|^{2/3}}\right] \Big/ \ln\left(\frac{n_b}{3z}\right) \tag{3.27}$$

The dimensionless Obukhov length scale (z/L) is equal to zero for the neutral condition. The resulting wind spectrum is characterised by the observed wind spectrum and the net effect of the transfer functions of the anemometer, chart recorder and averaging time. The transfer functions of the anemometer and chart recorder are

$$T_{anem}(f) = [1 + (2\pi f d/\overline{u})^2]^{-1} \tag{3.28}$$

$$T_{rec}(f) = [1 + (2\pi f t_{rec})^2]^{-1} \tag{3.29}$$

where d is the response length of the anemometer and t_{rec} is the response time of the chart recorder. For an analogue running-average filter, the filter operates on a continuous signal and averages of a t-second gust duration (t_{gust}), the transfer function of which is

$$T_{avg}(f) = \left(\frac{\sin \pi f t_{gust}}{\pi f t_{gust}}\right)^2 . \tag{3.30}$$

A discrete running-average filter averages over N consecutive samples with a sampling period of Δ giving

$$T_{avg}(f) = \frac{1}{N^2}\left(\frac{\sin \pi f \Delta N}{\sin \pi f \Delta}\right)^2 . \tag{3.31}$$

The final wind spectrum can be written as a combination of the Kaimal wind spectrum and the net effects of the transfer functions giving

$$S(f) = T_{anem}(f) T_{rec}(f) T_{avg}(f) S_u(f) \tag{3.32}$$

The properties of the anemometers and the chart recorders used by the KNMI can be found in Verkaik (2000) and Verkaik (2001). The turbulence intensity in Equation 3.23 can be rewritten as a function of the low-law and the integral of the Kaimal wind spectrum which equals

$$\frac{\sigma_u}{\overline{u}} = \frac{\sigma_u}{u_*}\frac{u_*}{\overline{u}} \tag{3.33}$$

$$= \left[\int_0^\infty \frac{1}{f}\frac{fS(f)}{u_*^2}\,\mathrm{d}f\right]^{1/2}\left[\frac{\kappa}{\ln(z/z_0)}\right] \tag{3.34}$$

By substituting Equations 3.24 and 3.34 into Equation 3.23 and considering the ratio of the observed gust to mean wind speeds, the effective roughness length can be solved directly.

3.3.7 Results

Correction factors are calculated for 13 stations using three exposure correction methods based on gustiness- and LULC-derived roughness lengths to evaluate the performance of each model. The percent error for each model is calculated as

$$e_r = \frac{|CF_{lulc} - CF_{gustiness}|}{CF_{gustiness}} \tag{3.35}$$

where CF_{lulc} and $CF_{gustiness}$ are the LULC- and gustiness-derived correction factors respectively. The mean percent error for the tested models along with the 10^{th} and 90^{th} percentiles are summarised in Table 3.5. The TL models provide lower mean percent error than the Deaves and Harris IBL model and do not exhibit nearly as large relative errors at the 90^{th} percentile.

A radial plot of the correction factors as calculated by the three models compared to the correction factors determined from the gustiness model for Den Helder Airport, De Kooy, NL is shown in Figure 3.7. The correction factors from the gustiness-derived roughness lengths for the three exposure models were nearly identical, the largest difference between the maximum and minimum correction for a sector being 0.02. The average of the three gustiness-derived correction factors is shown and considered the observed correction factor. The Deaves and Harris IBL model quite clearly over-predicts the correction factors for the sectors of 210 to 270-degrees by approximately 0.1. The upstream exposure in the corresponding wind directions is open-country for the majority of a 5 km fetch before

Model	Mean Error (%)	10^{th} Percentile (%)	90^{th} Percentile (%)
Deaves and Harris IBL	9.5	1.0	21.3
Hydra TL	6.5	0.9	13.7
Gryning TL	7.5	1.2	14.9

Table 3.5: Comparison of the relative error from IBL and TL model correction factors

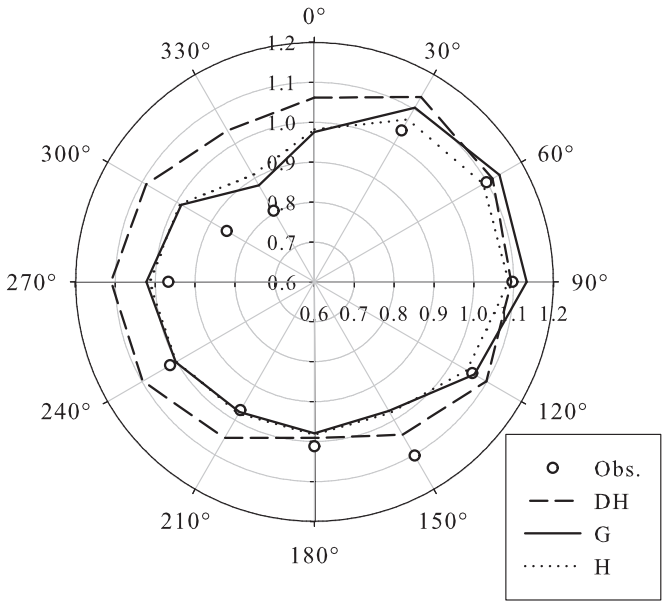


Figure 3.7: Directional exposure correction factors for Den Helger Airport, De Kooy, NL (WMO 06235)

reaching the North Sea. Thus, the over-prediction of the correction factor corresponds to an under-prediction of the slowdown associated with a smooth to rough transition. The gustiness-derived correction factors from all three models and the TL model correction factors calculated from LULC, indicate that at 10 m height the wind speed in the boundary layer is nearly in equilibrium with the local open-country exposure. A correction factor of 1.1 calculated from the Deaves and Harris IBL model based on LULC indicates the model has not achieved equilibrium with the local exposure and is still affected by the North Sea. The model indicates the wind speed is 10 percent greater than the wind speed observed once equilibrium with the local exposure is achieved. The difference between correction factors from the TL and IBL models with increasing fetch for a smooth to rough transition is shown in Figure 3.8. The correction factors are calculated for an idealised smooth to rough transition where the smooth exposure is open-water (z_0 = 0.003 m) and the rough exposure is open-country (z_0 = 0.05 m). The TL model clearly experiences rapid adjustment to the downstream exposure as the correction factor reaches unity within 10 km for wind observations at 10 m height. The IBL model experiences much slower adjustment

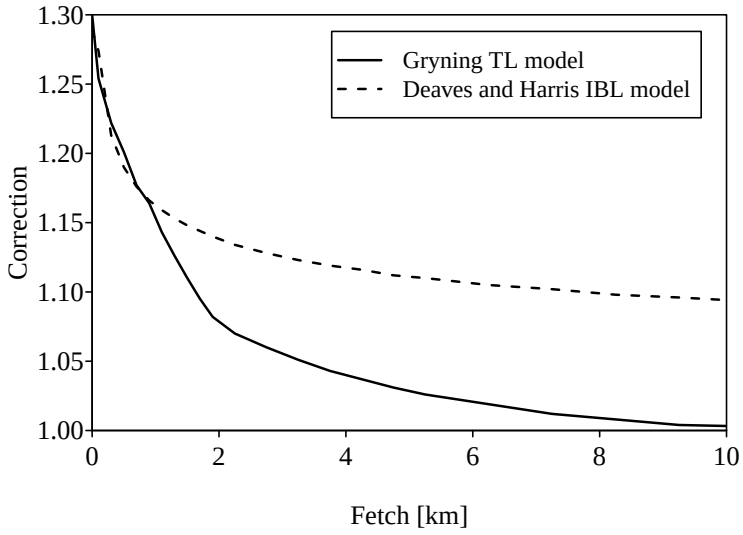


Figure 3.8: Comparison of correction factors for a smooth to rough transition

to the downstream roughness after 1 km, in fact at 55 km the IBL model has reached a correction factor of 1.04. The Deaves and Harris IBL model has been shown by Cook (1985) to require fetches in excess of 100 km to achieve equilibrium with a downstream homogeneous exposure. Correction factors for sectors of 30 to 120-degrees show much better agreement as the upstream open-water exposure is approximately 1.0 km, which falls in the region where the models exhibit similar corrections.

Due to similarities between the Deaves and Harris ABL and Gryning ABL models shown in Section 3.2 for wind profile fits to the Leipzig wind data, the differences between the IBL and TL exposure correction models arise entirely from the way the models consider the upstream roughness and fetch. Based on the similarities between the two ABL models, application of the TL model framework incorporating the Deaves and Harris ABL model would likely produce similar results to the Gryning TL model.

Between the two TL models, the Hydra TL model performs slightly better than the Gryning TL model, however, the Hydra TL model exhibits a much smaller range of correction factors than the Gryning TL and Deaves and Harris IBL models. To balance both the range of correction factors and relative error, the Gryning TL model is selected to correct for

heterogeneous exposure. The Gryning TL model exhibits relative error statistics similar to those of the Hydra TL model and a range of correction factors similar to the Deaves and Harris IBL model.

For the current work, correction factors are calculated for all stations by direction using the Gryning TL model. Correction factors are calculated and applied accordingly for additional measuring heights and mast locations throughout the operational lifetime of each station. Where no information was known regarding the measuring height at a location, the anemometer height was assumed to be 10 m. Corrections are based on the periods and changes identified by the ACMANT algorithm of Section 3.1.2 in conjunction with the available documentation. As previously mentioned, the version of the CORINE LULC database is for the year 2000, thus, the assumption is required that corrections based on the LULC at existing mast locations are representative of the site conditions at the time. The distribution of correction factors for all stations and directions is shown in Figure 3.9. The majority of stations have a correction factor near unity which is consistent with the assumption that anemometers should be sited in open-country exposure.

For both the LULC- and gustiness-derived correction factors there are considerable sources of uncertainty. The gustiness-derived corrections are based on a gust factor which is variable and wind observations which are susceptible to local effects from neighbouring structures. The LULC-derived corrections are based on generalised LULC classifications and a single roughness length for each class. Many of the anemometers are located on airports which are assigned a single roughness length despite variations between airports concerning the length of grass and proportion of runways and structures. Lastly, the 100 m grid resolution provides insufficient spatial resolution to identify the localised effects which will contribute to the gustiness-derived corrections. Considering the numerous sources of uncertainty, relative mean error estimates of 6 to 10 percent for both TL and IBL models are not unreasonable. However, it has been shown that fundamental differences do exist

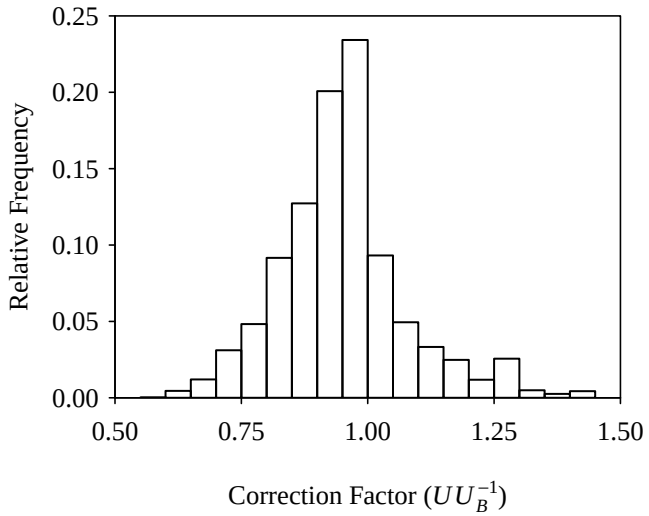


Figure 3.9: Distribution of correction factors

between TL and IBL models at moderate fetches for large roughness changes. The inadequacy of performing a simple correction, which only considers the site exposure, is clearly shown by the variation of the correction factors with direction in Figure 3.7.

3.4 Disjunct Sampling Correction

3.4.1 Background

Disjunct sampling concerns observation stations which are not sampling hourly. Depending on the sampling frequency, a gap of several hours may exist between consecutive measurements. When carrying out statistical predictions of extreme wind events, the maxima of representative subsets are required. The extraction of maxima from a time history of hourly observations may yield larger maxima than observations recorded every three hours since peak wind speeds may have occurred during interim hours. A correction is therefore required to adjust maxima from stations which are not recording hourly. A similar argument is valid for adjusting a 10-minute mean in the 10-minute period prior to the hour to

a continuous 10-minute mean, however, given the resolution of the data available for the current work, the additional correction is unattainable.

A simple approach to accounting for disjunct sampling was explored by Frank (2001). Three months of continuous data were extracted and 10-minute mean wind speeds were calculated from measurements at heights of 44, 77 and 125 m. A subset of observations taken at six-hour intervals was extracted and the maximum wind speed from the subset was compared to the true maximum wind speed. The maximum of the six-hour interval measurements was found to be approximately 89 percent of the true maximum.

Two methods which account for disjunct sampling, one theoretical and one empirical were suggested by Larsén and Mann (2006). Both methods are capable of being applied omni-directionally or sector-wise. The methods provide an estimate of the ratio of the mean annual extreme extracted from a t_d-hour sampling interval to the true mean annual extreme. For the current work, fitting and estimation of the integral time scale from the measured spectra for every station is not ideal. Since the empirical method was shown to provide good estimates of the theoretical method for the omni-directional case, the method is described here. The ratio of the mean annual extremes is calculated as

$$\frac{\bar{u}_{d,max}}{\bar{u}_{max}} = b - a(\log(s))^2 \tag{3.36}$$

where s is the sectoral frequency. The empirically derived values are functions of the sampling interval and equal

$$a = 0.0209(t_d - 1/6)^{0.4627} \tag{3.37a}$$

$$b = 1 - 0.0342(t_d - 1/6)^{0.5346} \tag{3.37b}$$

Practical consideration of disjunct sampling and a synoptic wind climate suggests an ap-

propriate correction factor is dependant upon sampling interval and will likely exhibit wind speed dependency. High wind speeds associated with strong depressions are generally sustained for prolonged periods depending on the severity of the storm. In this situation, it is reasonable to assume that for increasing wind speed there is greater correlation between recorded maximums regardless of sampling interval.

The correction models of Frank (2001) and Larsén and Mann (2006) do not allow changes in sampling interval, which occur throughout the year, to be accounted for. In the situation where years exhibit mixed sampling intervals, corrections should be based on the largest t_d-hour sampling period. Alternatively, a method which considers monthly maxima recorded at a specific sampling interval and separated by wind speed is preferred. Application of the process to each month allows the annual maximum to be determined from the adjusted monthly maximums. Carrying out the process by month allows all years to be utilised regardless of changes to the sampling interval throughout the year. To investigate the influence of disjunct sampling on extracted maxima, an empirical model is derived in the following section.

3.4.2 Methodology

Monthly maxima are extracted from all stations for months which have greater than 22 observations per day. The extracted monthly maxima are not true continuous maxima, thus the corrections presented here are in relation to 10-minute mean wind speeds recorded during the 10-minute period prior the hour. For each of the resulting 74273 months, maxima are extracted for two-, three-, four-, and six-hour sampling intervals and daytime observations only. The extracted disjunct maxima are sorted by 5 m/s wind speed increments and sampling interval. The mean ratio of the disjunct maximum to the ‘true’ monthly maximum are calculated across each wind speed bin and sampling interval. Third-order polynomials

are fit to the ratios across all wind speed bins to form correction factors which are a function of wind speed.

3.4.3 Results

The resulting correction factors from the analysis are shown in Figure 3.10 and the corresponding equations are presented in Table 3.6. For comparison, the empirical model provided by Larsén and Mann (2006), given by Equation 3.36 for the irrespective of direction case, yields correction factors of 0.953, 0.940, 0.930 and 0.912 for two-, three-, four-, and six-hour sampling intervals. These values correspond well with the six-hour sampling interval correction from Frank (2001) and the current corrections at 12.5 m/s. Direct comparison of the correction factors should be performed with caution as the values provided by the model of Larsén and Mann (2006) are meant for application to the mean annual maximum and the corrections provided here are applied to monthly maxima. Since the correction factors determined by Larsén and Mann (2006) match with the lower wind speed estimates considered here, application of the correction factors from Equation 3.36

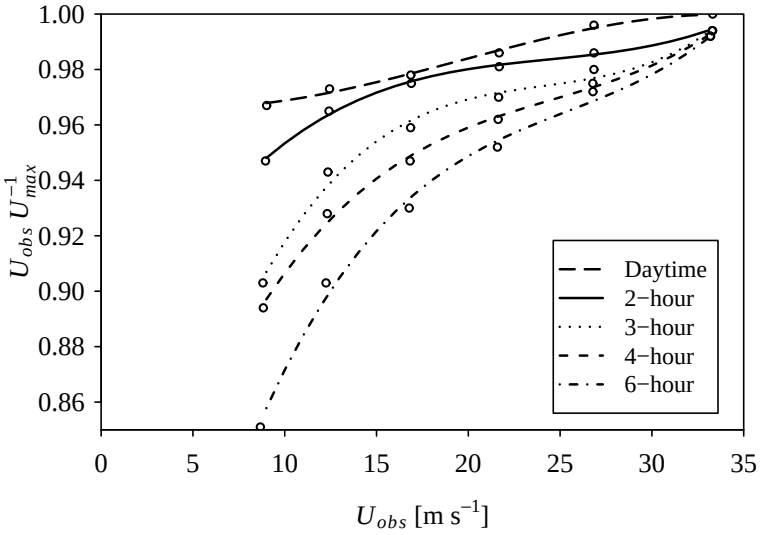


Figure 3.10: Disjunct sampling correction factors

Sampling Interval	Correction Factor
2	$7.02 \cdot 10^{-6} u_{obs}^3 - 5.13 \cdot 10^{-4} u_{obs}^2 + 1.31 \cdot 10^{-2} u_{obs} + 8.66 \cdot 10^{-1}$
3	$1.55 \cdot 10^{-5} u_{obs}^3 - 1.12 \cdot 10^{-3} u_{obs}^2 + 2.79 \cdot 10^{-2} u_{obs} + 7.35 \cdot 10^{-1}$
4	$1.08 \cdot 10^{-5} u_{obs}^3 - 8.01 \cdot 10^{-4} u_{obs}^2 + 2.17 \cdot 10^{-2} u_{obs} + 7.59 \cdot 10^{-1}$
6	$1.48 \cdot 10^{-5} u_{obs}^3 - 1.12 \cdot 10^{-3} u_{obs}^2 + 3.11 \cdot 10^{-2} u_{obs} + 6.58 \cdot 10^{-1}$
Daytime	$-2.93 \cdot 10^{-6} u_{obs}^3 + 1.71 \cdot 10^{-4} u_{obs}^2 - 1.56 \cdot 10^{-3} u_{obs} + 9.70 \cdot 10^{-1}$

Table 3.6: Disjunct sampling correction factor equations

appear to be conservative for strong winds. The differences may not be as large if an additional factor is applied to the current results to account for the adjustment from hourly to continuous sampling.

The correction factors shown in Figure 3.10 approach unity with increasing wind speed across all sampling intervals. Given the correction factor approaches unity for high wind speeds, the results indicate that wind observations during these events are representative of the true maximum regardless of sampling interval. Thus, stronger wind events likely exhibit wind speeds sustained for increased durations. For example, when sampling at 6-hour intervals the greatest possible duration between a maximum wind speed and an observation is 3 hours. Kasperski (2007) similarly found that strong frontal depressions lasted on average three hours with the second and third event hours exhibiting wind speed ratios of approximately 0.97 and 0.93 the maximum event wind speed. The results here suggest that for up to 3 hours adjacent an event maximum, the observed wind speeds are nearly as strong as the maximum wind speed when the maximum is greater than 25 m/s. High wind speed events which exhibit sustained wind speeds are linked to several ETCs affecting Europe including the Great Storm of ’87 (October, 1987), the Burns’ Day Storm (January, 1990) and Anatol (December, 1999).

Chapter 4

Statistical Methods for the Estimation of Extreme Winds

50-year return period wind speeds are representative of a standardised wind speed with annual probability of exceedance of 0.02. Sufficient observations are rarely available to determine wind speeds at high probabilities of exceedance, requiring statistical methods to extrapolate the relationship between exceedance probability and wind speed. The traditional approach to extreme value analysis in the wind engineering community is to base predictions on annual maxima and extrapolate using the Gumbel extreme value distribution, a case of the generalised extreme value distribution (GEVD). Current methodologies have utilised larger datasets by considering storm maxima or peak-over threshold methods (Kasperski, 2002; Miller, 2003; Sacré *et al.*, 2007). The following chapter provides an overview of the statistical methods traditionally used by the wind engineering community, and proposes the use of advancements available from the field of extreme value statistics.

4.1 Classical Extreme Value Theory

Given a sequence of observed wind data $\boldsymbol{x} = \{x_1, ..., x_n\}$, measured with consistent temporal frequency, the series is traditionally assumed to be identically distributed random observations of the Weibull distribution, a generalisation of the Rayleigh and Exponential distributions. The series may also satisfy independence depending on the sampling interval of the observations and the scale of local meteorological processes. Classical extreme value theory aims to estimate the behaviour of the upper tail of the parent distribution. In general, maxima of a representative subset are assumed to be independent and identically distributed random observations, although a site subject to a mixture of meteorological processes may violate the latter. By defining and fitting a limiting distribution to extracted maxima, estimates for various probabilities of exceedance can be calculated. In the field of engineering, the probability of exceedance is often considered in terms of the return period.

The most common statistical distributions used to model extreme values of synoptic winds by the wind engineering community are the Gumbel distribution, or extreme value distribution (EVD) type I, and to lesser extents, the EVD type III and generalised Pareto distribution (GPD). The EVD types I and III are part of the GEVD family which is traditionally applied to annual maxima, although the distribution has been applied to other block maxima. Application of the GEVD to annual maxima discards a large amount of potentially useful wind data for a single value each year. Alternative approaches have been proposed which utilise a much greater percentage of the data by considering storm maxima (Cook, 1982; Harris, 1999, 2009) or r-largest order statistics (An and Pandey, 2007).

The GPD is a threshold method which makes use of all data over a chosen threshold, allowing more data to be utilised than a block-maxima approach. The selection of an appropriate threshold can often be challenging. Mean residual life plots of extreme wind data rarely exhibit clear linear behaviour, resulting in the selection of a threshold which

proves to be intractable in many cases. The proper selection of a threshold for the GPD requires a somewhat more extensive process. Stability plots of each parameter over a range of thresholds with 95% confidence limits should be produced. Beginning with the largest potential threshold on the stability plot, fit a straight line with zero slope through as many consecutive thresholds as possible. The last value which successfully fits the line should provide an ideal threshold. Alternatively, more sophisticated methods exist. One such method developed by Dupuis (1998), incorporates the use of robust estimators for assessing threshold selection. Unfortunately few, if any, of the threshold selection methods are easily automated to facilitate the number of locations considered in the current work. In addition, the use of the GPD is still widely debated in the wind engineering community, see Holmes and Moriarty (1999), Cook and Harris (2001), and Harris (2005). For these reasons the GEVD is selected for modelling extreme synoptic winds.

4.1.1 Generalised Extreme Value Distribution

The basis of the GEVD was established by Dodd (1923) who identified a relationship between the asymptotic growth of the maximum of a set of independent and identically distributed random variables and the rate at which the tail of the parent probability density distribution approaches zero. Fisher and Tippet (1928) concluded that the behaviour of a series of identically distributed extreme values approaches that of one of three families, EVD types I, II and III, referred to as Gumbel, Fréchet and Weibull respectively. The three distributions were later formed into a parametric model by von Mises (1936) where the cumulative distribution function, given parameter vector $\boldsymbol{\theta}$ in parameter space $\boldsymbol{\Theta}$ ($\boldsymbol{\theta} \in \Theta$), is written as

$$F_{\boldsymbol{\theta}}(x) = \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \qquad \text{for } x : 1+\xi(x-\mu)/\sigma > 0 \text{ and } \sigma > 0 \quad (4.1)$$

where $\boldsymbol{\theta} = [\mu\ \sigma\ \xi]^T$ and the parameters are the location, scale, and shape respectively. The value of the shape parameter determines the associated family such that EVD types I, II, and III correspond to the cases $\xi = 0$, $\xi > 0$ and $\xi < 0$ respectively. The EVD type II is typically considered inappropriate for predicting extreme winds as it results in a lower bound and rapid growth with decreasing exceedance probability. Harris (2004) argues the EVD type III should be abandoned as significance tests indicate that a null hypothesis of the shape parameter equalling zero is accepted and the appearance of a EVD type III is the result of a poorly converged EVD type I due to insufficient data.

The Gumbel distribution (EVD type I) is a special case of the GEVD and is defined as the limiting form of Equation 4.1 as the shape parameter approaches zero ($\xi \to 0$) such that

$$F_{\theta}(x) = \exp\left\{-\exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\} \tag{4.2}$$

is defined for all x and $\boldsymbol{\theta} = [\mu\ \sigma]^T$. Several methods and estimators exist for fitting the unknown parameters to a given dataset and are covered in Section 4.2. The equations defining additional statistical properties such as the probability density function of the GEVD are provided in Appendix B.1.

4.2 Estimators

Several classical estimators exist which appropriately estimate unknown parameters of a distribution, either graphically or numerically, for a specific dataset. An estimator is a function of the observed data which is utilised to estimate the unknown parameters or estimands. Graphical or least-square methods have been historically preferred in the field of wind engineering over numerical methods. One of the debates when using these methods is the selection of an appropriate plotting position. The original plotting position given by

Gumbel (1958) as

$$F(x_i) = \frac{i}{N+1} \tag{4.3}$$

has been considered biased and is often replaced by the plotting position given by Gringorten (1963) as

$$F(x_i) = \frac{i - 0.44}{N + 0.12}. \tag{4.4}$$

The debate of the appropriate plotting position has been recently revisited and is ongoing, see Makkonen (2006), Makkonen (2008), and Cook (2011). Overall, the methods typically applied by the wind engineering community to calculate the fitted parameters of the extreme value distribution are outdated when one considers the efficient numerical estimators utilised by statisticians. The plotting position debate is extraneous considering the statistical techniques which are available to directly solve estimands (de Haan, 2007).

In the field of extreme value statistics, the method of moments (MoM), maximum likelihood estimators (MLE), and probability weighted moments (PWM) are established estimators. Despite the general disregard for such methods by the wind engineering community, in the field of statistical modelling of extremes, the methods are considered classic when one considers current research. Updated methods include optimal bias-robust estimators (OBRE) and Bayesian methods.

Parameter estimation using MoM is carried out by solving the population moments (e.g. mean, variance) using the sample moments. The estimator is easily biased as calculation of the sample mean can be sensitive to outliers for small sample sizes. An alternative to MoM which is less sensitive to outliers is PWM. PWM belong to the family of L-estimates introduced by Greenwood *et al.* (1979) and further developed by Landwehr *et al.* (1979) and Hosking *et al.* (1985). L-estimators tend to be less sensitive to outliers than other

classical estimators as they are calculated from linear functions of the data, rather than the individual values (Hosking, 1990). The estimator differs from conventional moments since the estimates are calculated from linear combinations of ordered data. Hosking *et al.* (1985) show that PWM have reduced bias, which often provides a better fit to observed data than MLE. Alternatively, Dupuis and Field (1998b) found PWM can be biased by a single large event, thus, the authors suggest the use of OBRE. OBRE are a robust extension of MLE which produce similar parameter estimates as PWM and provide additional information describing the quality of the fit to each observation.

4.2.1 Maximum Likelihood Estimators

MLE were introduced by Fisher (1912, 1922) and were applied to the GEVD by Jenkinson (1969). Huber (1964) proposed a generalisation of the MLE by a class of estimators called M-estimators which provides the basis of OBRE discussed in Section 4.2.2. The formulation of the MLE is summarised here in the context of M-estimators, the general form of which is given by Huber (1964) as

$$\min \sum_{i=1}^{n} \rho(x_i, \boldsymbol{\theta}) \tag{4.5}$$

where ρ is an appropriate function. An estimate of the parameters minimising Equation 4.5 are calculated by setting the derivative of ρ, expressed as

$$\boldsymbol{\psi}(x; \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \rho(x, \boldsymbol{\theta}), \tag{4.6}$$

equal to zero and solving the resulting implicit equation

$$\sum_{i=1}^{n} \boldsymbol{\psi}(x_i; \theta) = 0. \tag{4.7}$$

In the case of the MLE, the general form of the M-estimators is rewritten as a maximum by taking the negative value of the function ρ giving

$$\max \sum_{i=1}^{n} -\rho(x_i, \boldsymbol{\theta}). \tag{4.8}$$

The parameters which will maximise the likelihood function, defined as

$$L(\boldsymbol{\theta}; x) = f_{\boldsymbol{\theta}}(x_1, ..., x_n|\boldsymbol{\theta}) \tag{4.9}$$

$$= \prod_{i=1}^{n} f_{\boldsymbol{\theta}}(x_i) \tag{4.10}$$

are then sought. Equation 4.9 can be written in the form given by Equation 4.10 provided $\boldsymbol{x}$ satisfies independence. By taking the logarithm, a monotonic transformation, of Equation 4.10 the log-likelihood is written as

$$\log L(\boldsymbol{\theta}; x) = \sum_{i=1}^{n} \log f_{\boldsymbol{\theta}}(x_i). \tag{4.11}$$

Thus, the function described in the general form of the M-estimators is equal to

$$\rho(x; \theta) = -\log f_{\boldsymbol{\theta}}(x). \tag{4.12}$$

and its derivative, defined by Equation 4.6, equals

$$\boldsymbol{\psi}(x_i; \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log f_{\boldsymbol{\theta}}(x; \boldsymbol{\theta}) \tag{4.13}$$

which is commonly referred to as the maximum likelihood scores function, $\boldsymbol{s}(x_i; \boldsymbol{\theta})$ (Hampel *et al.*, 1986). The score functions for the GEVD are derived in Appendix B.1.

4.2.2 Optimal Bias-Robust Estimators

The most significant shortcoming of the classical estimators is the lack of robustness. Depending on the number of observations, small deviations from the underlying model can greatly affect the estimands if the influence function (IF) is unbounded (Dupuis and Field, 1998b). By bounding the IF of an estimator, small contaminations in the data will not largely affect the outcome of the estimator. The influence function in its general form is provided in Appendix B.2.1. To mitigate the influence of deviations from the assumed model, robust estimators are based on the data that are well fit by the model. Observations not well fit by the assumed model are therefore weighted lower than those fit well by the model. OBRE have been successfully applied to environmental extremes such as temperature (Dupuis and Field, 1998b) and wind measurements obtained from buoys moored in the Pacific Ocean (Dupuis and Field, 2004).

The M-estimators discussed in Section 4.2.1 form a starting point for OBRE. The IF of MLE is unbounded as a result of the score function, given by Equation 4.13, being unbounded in x (Dupuis and Field, 1998b). An overview of the IF for MLE is provided in Appendix B.2.2. To construct a bounded influence function for the MLE, a bounded version of Equation 4.6 is required which is as similar to Equation 4.13, the maximum likelihood scores function, as possible. To bound the influence of observations not well fit by the model, the Huber function forms the basis of a weighting function. The Huber function maps values of function z which are outside the bounds of $h_c(z)$ to the nearest value on $h_c(z)$ ($z \mapsto h_c(z)$), thus reducing the influence of the furthest values (Hampel *et al.*, 1986). The multidimensional Huber function is given by

$$h_c(z) = zW_c(z) = z \min\left(1, \frac{c}{\| z \|}\right) \tag{4.14}$$

where W_c is the weighting function, c is the robustness constant and $\| \cdot \|$ denotes the

Euclidean norm. When the robustness constant in Equation 4.14 equals infinity, the MLE is achieved since $W_c(x, \theta) = 1$ for all observations and parameters. The complete derivation of the estimator is provided by Hampel *et al.* (1986) and Dupuis and Field (1998b), while the resulting bounded estimator and associated algorithm for the OBRE procedure are provided in Appendix B.3. The OBRE algorithm provides estimates of the fitted parameters and the weight applied to each observation.

4.3 Outlier Identification

A relatively simple extension of the OBRE algorithm was developed by Dupuis and Field (2004) to identify observations which are not well fit by the assumed distribution for wind data recorded at moored buoys in the Pacific Ocean. The OBRE algorithm will inherently downweight poorly fitted observations, however, a process which can identify and evaluate the likelihood of such events has additional benefits. In the wind engineering community, there has been little discussion regarding how to handle outliers. Designers may choose to leave a questionable observation in the dataset, while others may arbitrarily remove such an observation. Gatey and Miller (2007) identified four stations where sets of annual maxima were contaminated by a potential outlier. The removal of the potential outlier, and inclusion of the subsequent maximum wind speed observed for the year, resulted in an average reduction of the 50-year return period wind speed of 5.5 m/s which changed from the EVD type II to EVD type III and 1.2 m/s for the EVD type I. The outliers were not necessarily errors. In several instances observations were linked to significant depressions, such as the Burns' Day Storm which influenced northwestern Europe on January 25, 1990. Wind measurements from strong depressions may be associated with longer return periods, such as 100- or 200-year return period events.

The weight applied to an observation by the OBRE algorithm is a measure of how well

the observation is fit by the model. Conceptually, the procedure identifies potential outliers based on the weights of the same ordered point in a series of simulations of the fitted model. The procedure identified by Dupuis and Field (2004) is outlined below:

1. Fit the observed data using the OBRE method to the selected EVD, obtaining estimates of the fitted parameters and weights applied to each observation.
2. Using the fitted parameters, simulate a dataset equal in size to the observed dataset from the model distribution.
3. Fit the simulated dataset using the OBRE method, and compute weights for each observation.
4. Repeat Steps 2 and 3 for 1000 simulations.

For a potential outlier in the observed dataset, the weight applied to the observation is compared to the weights of the simulated data for the corresponding ordered data point. That is, the weight of the ith highest observation is compared to the weight of the ith highest data point within each simulation. By forming an empirical distribution of the weights from the corresponding ordered point, an approximate p-value for the observed weight can be calculated. If the p-value of the weight is smaller than the fifth percentile, then there is statistical evidence that the observation is an outlier.

Once an outlier is identified, caution is required when drawing conclusions from the results. The detection scheme identifies points that are not fit well by the modelled distribution and provides statistical evidence that an observation is in fact an outlier. The analysis does not identify, or omit the possibility, that the data point is a result of measurement error, alternative meteorological mechanism, from a longer return period or incorrect choice of model. Additional sources of information such as neighbouring locations or historical records may require examination to aid in identifying the underlying cause of an outlier.

4.4 Results

4.4.1 Annual Maxima

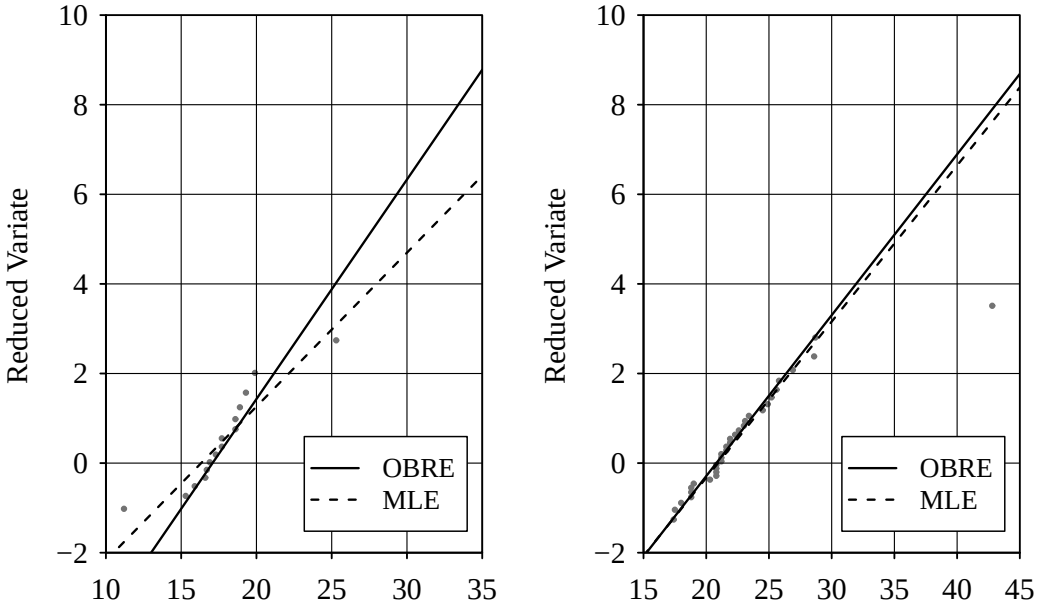
The OBRE procedure discussed in this chapter is applied to fit the EVD type I to the surface observation data obtained from the ISD. For each measurement station, the distribution is fit to the annual maximum wind speeds. The maxima are extracted yearly based on a full seasonal cycle, from 1 May to 30 April, to ensure a single strong wind season (winter) is not divided between two annual blocks. In addition, a minimum of 72 hours is required between maxima to ensure independence. The fitted EVD type I parameters are used to extrapolate wind speeds at various probability of exceedances, calculate confidence intervals and check for outliers.

The effect of the bounded influence function of the OBRE is shown by comparison to MLE fits in Figure 4.1. Wind speeds are plotted versus reduced variate (Y) equal to

$$Y = -\log(-\log(F_{\theta}(x))) \tag{4.15}$$

$$= \frac{x-\mu}{\sigma} \tag{4.16}$$

for the EVD type I. Figure 4.1(a) exemplifies the difference in predictions from OBRE and MLE, and the ability of the OBRE algorithm to identify and mitigate the impact of annual maxima which may be much larger or smaller than the majority of the data. For the second case shown in Figure 4.1(b), the outlier has little impact on the MLE fit. The single outlier, although exhibiting large magnitude, is unable to skew the statistical fit as a result of the 32 years of data considered. Although use of OBRE resulted in negligible differences to the statistical fit of the dataset, the outlier algorithm identified a wind speed was present which was extremely unlikely given the fitted distribution. The weight assigned by the



(a) Angermuende, DE (WMO 09291) (b) San Sebastian, SP (WMO 08029)

Figure 4.1: EVD type I fit to annual maxima with potential outlier

OBRE process to the potential outlier is 0.62. From the simulation procedure, comparison of the weight to the respective weights of the corresponding ordered data point across all simulations results in a p-value of 0.013. Thus, only 1.3 percent of the simulations drew a wind speed which was assigned such a low weight by the estimator, indicating that the wind speed is likely an outlier. Manual investigation of the outlier indicates the observation is an error which could not be removed by the localised quality control algorithm due to insufficient neighbouring observations to form a proper assessment. In this situation the outlier is removed and replaced with the appropriate annual maximum since the outlier is considered an erroneous observation. In the case where the outlier is deemed to be a real observation, the decision is required whether to remove the outlier and replace the observation with the second largest wind speed in the year, allow the OBRE algorithm to mitigate the influence of the observation, or assign full weight to the observation for complete consideration. A typical good quality fit to annual maxima is shown in Figure 4.2 for Valley, UK. A comparison of bootstrap and asymptotic intervals for the GEVD by

Dupuis and Field (1998a) found bootstrap intervals did not perform as well as asymptotic intervals. Thus, asymptotic confidence intervals are constructed for the quantiles of the EVD type I.

4.4.2 Storm Maxima

The station selection criteria described in Section 2.3 ensures that the temporal resolution of measurements at the selected stations is a minimum of 8 observations per day, however, the majority of selected locations measure 24 observations per day. The large quantity of data allows a comparison of predictions based on storm and annual maxima to be carried out. Independent storms are defined as the periods occurring between lulls, identified here in a similar method as the one described by Cook (1982).

The storm threshold is selected using the Beaufort wind force scale. Beaufort wind force number 9 corresponds to a strong gale and is classified by a minimum wind speed of 20.8 m/s in open-water conditions. Applying a correction factor for open-water to open-country exposure of approximately 1.3, based on the Gryning TL model, to the minimum wind speed, provides an open-country threshold of 16 m/s. The maximum wind speed observed within the temporal bounds of the storm is extracted, and carried out for each storm identified in the time history. A minimum of 72 hours is required between consecutive maxima to satisfy independence. Selection of an appropriate universal storm threshold can be difficult. A threshold set too low may not reflect the behaviour of the upper tail of the observations as shown in Figure 4.3, and if set too high may result in an insufficient number of observations for analysis.

The use of storm maxima requires the probability of exceedance to be adjusted to account

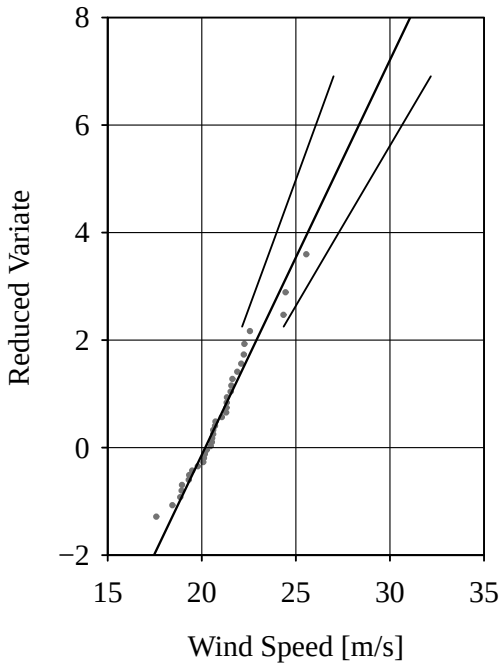
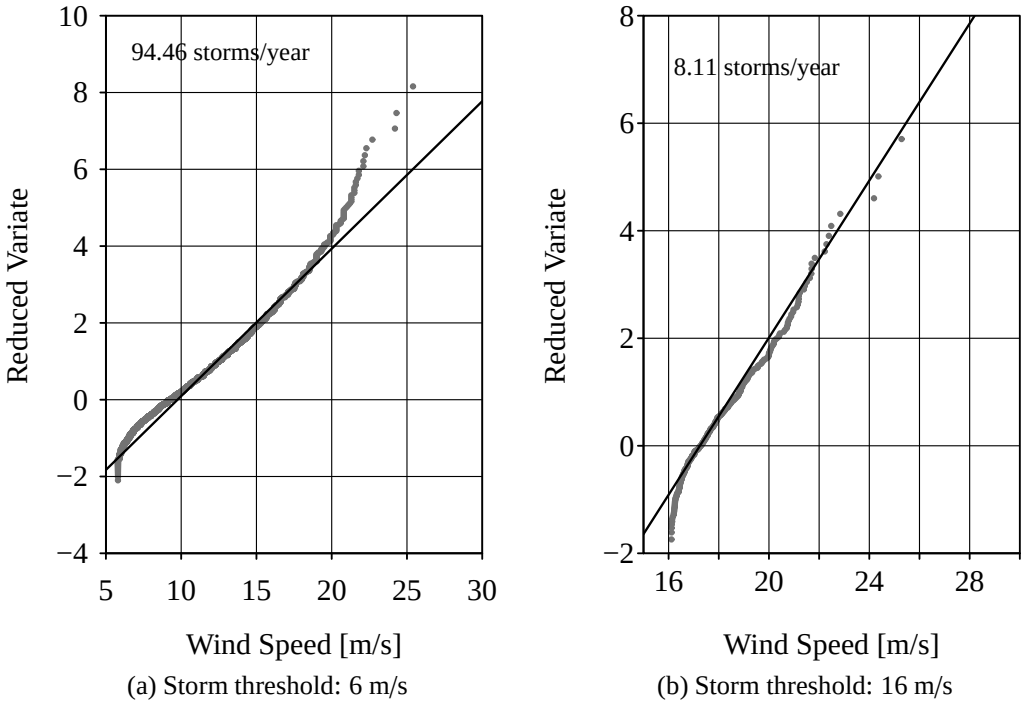


Figure 4.2: Typical good-quality EVD type I fit to annual maxima - Valley, UK (WMO 03302)



(a) Storm threshold: 6 m/s

(b) Storm threshold: 16 m/s

Figure 4.3: Influence of storm threshold selection - Valley, UK (WMO 03302)

for the mean annual frequency of storms (s) such that

$$1 - F_{\theta}(x < x_{ref}) = \frac{1}{50s} \tag{4.17}$$

for the equivalent 50-year return period estimate. Due to the increasing density of observations with decreasing wind speed, fits based on storm maxima are biased by the low wind speed values. To avoid the impact of the low wind speed values, Kasperski (2002) fits over the range of probability of non-exceedances greater than 0.3 for the EVD type III. Similarly, Harris (2009) shows the minimum usable reduced variate is approximately -1.8 for measurements in the UK. Fitting over a selected range is more easily achieved when graphical or least-squares techniques are utilised to estimate the parameters, which is likely a significant reason why the methods are still common in engineering. The application of such modifications to classical numerical estimators such as MLE and PWM is analytically intractable. However, since the OBRE procedure assigns a weight to each observation, the weighting function is modified such that observations in the lower region are assigned a negligible weight of 0.01. Storm maxima less than the Beaufort classification for a storm, wind force number 10, are assigned a negligible weight. The associated wind speed for a Beaufort classification of storm in open-country exposure is 18.8 m/s. The influence of the downweighting is shown in Figure 4.4.

The influence of a single outlier when greater than 30 years of annual observations are available has been shown to be negligible. Thus, due to the large number of storm maxima, the influence of a single observation was found to have minimum impact on the associated EVD type I fit. A good quality fit of storm maxima is shown for Valley, UK in Figure 4.5. The strength of the extreme wind climate will vary spatially, making it difficult to successfully apply a single criterion to the number of stations considered here. For stations located in regions where the synoptic extreme wind climate is relatively weak, defined here as locations having less than 16 observations greater than or equal to the downweighting threshold

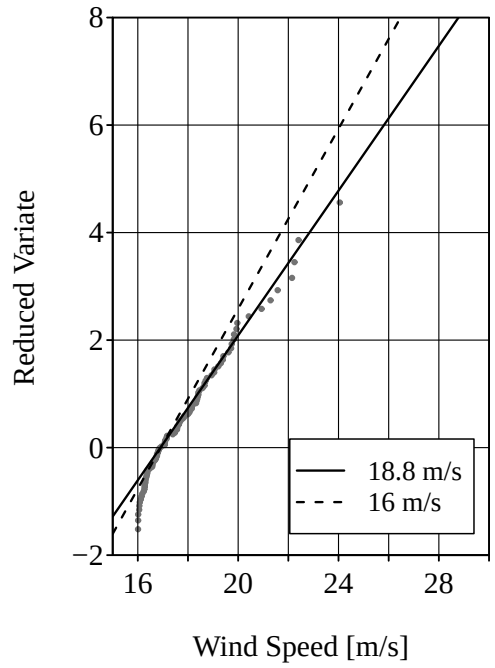


Figure 4.4: Influence of OBRE downweighting - Portalegre, PO (WMO 08571)

(18.8 m/s), a reduced storm threshold was applied. The storm threshold, previously 16 m/s, was reduced to the maximum of either the lowest annual maximum wind speed or 13.3 m/s, the open-country exposure wind speed of the Beaufort wind force number 8, classified as a gale. If the storm threshold criterion was reduced, the downweighting threshold was adjusted to assign negligible weights to observations less than a Beaufort wind force number of 9.

50-year return period wind speed estimates extrapolated from statistical fits to storm and annual maxima exhibit strong correlation. Comparison of the wind speed estimates indicates 76 percent of 50-year return period wind speeds calculated from the two types of maxima differ by less than 2 m/s. Of the remaining stations differing by greater than 2 m/s, 34 percent (8 percent overall) differ by greater than 3 m/s and less than 6 percent (1.5 percent overall) differ by more than 5 m/s. The distribution of the difference between the 50-year return period wind speeds predicted from yearly and storm maxima ($V_{50,yearly} - V_{50,storm}$) is shown in Figure 4.6. The resulting differences indicate that the yearly maxima tend to

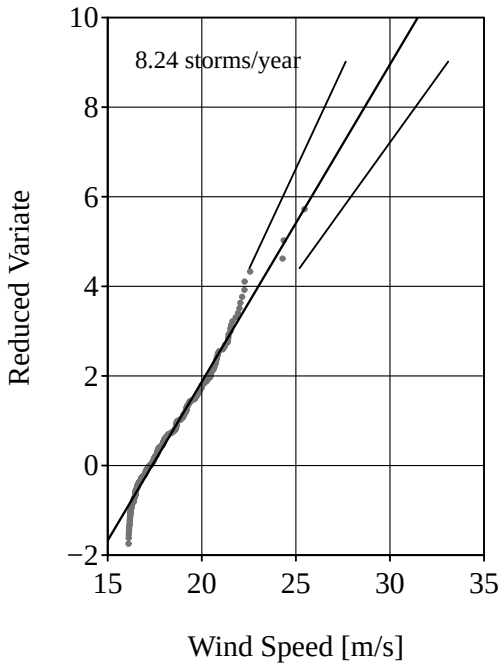


Figure 4.5: Typical good-quality EVD type I fit to storm maxima - Valley, UK (WMO 03302)

produce estimates which are on average 1 m/s greater than the storm estimates.

In cases where appreciable differences exist between 50-year return period wind speed estimates from storm and annual maxima, fits based on annual maxima were consistently better than those based on storm maxima, provided sufficient years of data were available. The differences are not necessarily due to an inadequacy of using storm maxima to estimate 50-year return period wind speeds. The higher number of poorly fit datasets is rather a shortcoming of the attempt to use a set of predefined criteria to select storm and down-weighting thresholds for a significant number of stations in a large spatial region. For this reason the 50-year return period wind speeds computed from annual maxima are selected for mapping.

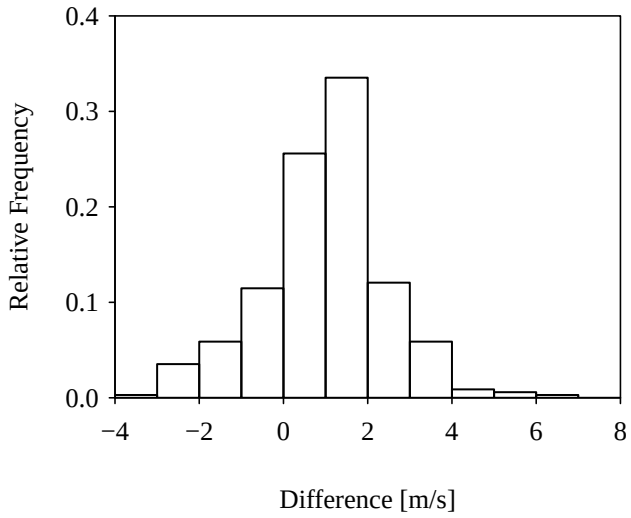


Figure 4.6: Distribution of 50-year return period wind speed differences

4.4.3 Mapping

Prior to mapping the 50-year return period wind speeds derived from annual maxima, the goodness-of-fit of the model to the observed maxima was evaluated using the the Anderson-Darling (A-D) test statistic. The A-D test statistic provides a measure of whether a dataset is from a specified probability distribution (Anderson and Darling, 1952). The simple test statistic, as given by Anderson and Darling (1954), is calculated as

$$W^2 = -n - \sum_{i=1}^{n} \left\langle \frac{2i-1}{n} \left[\log F_{\theta}(x_i) + \log(1 - F_{\theta}(x_{n+1-i})) \right] \right\rangle \tag{4.18}$$

where x_i are the observations in the dataset, n is the length of the dataset and $F_{\theta}(x_i)$ is the cumulative distribution given in Equation 4.2. A modified test statistic which accounts for the weight assigned to each observation by the OBRE is provided by Dupuis and Field

(2004) as

$$nW^2 = \sum_{i=1}^{n-1} \Big(w(x_i) \left[-F_{\theta}(x_{i+1}) + F_{\theta}(x_i) \right] + w(x_i)(i/n)^2 \left[\log F_{\theta}(x_{i+1}) - \log F_{\theta}(x_i) \right] - w(x_i)(i/n - 1)^2 \left\langle \log \left[1 - F_{\theta}(x_{i+1}) \right] - \log \left[1 - F_{\theta}(x_i) \right] \right\rangle \Big) + w(x_n) \left(-1 + F_{\theta}(x_n) - \log F_{\theta}(x_n) \right). \quad (4.19)$$

where $w(x_i)$ is the corresponding weight of observations x_i.

Asymptotic significance points have been given for A-D tests for the GEVD, however, due to the downweighting of observations, a quantitative approach is taken here. The A-D test statistic is computed for all parameter estimates and observations. Based on the empirical distribution of test statistics shown in Figure 4.7, fits to the data whose A-D test statistic was above the 95^{th} percentile (≈ 0.04) were examined manually. Further investigation revealed that fits to the data were found to be quite poor for fits exhibiting an A-D test statistic greater than 0.05, or a percentile of 98.5. Thus, an A-D test statistic threshold of 0.05 was applied to the 50-year return period wind speed estimates, resulting in the removal of six values.

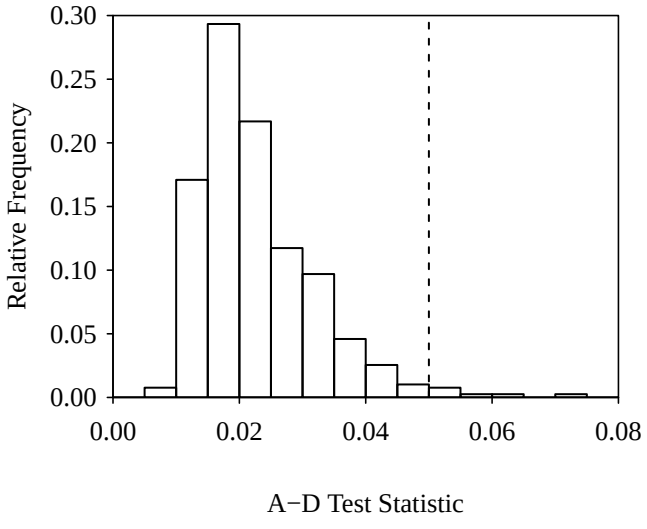


Figure 4.7: Distribution of the A-D test statistic

The 50-year return period wind speeds derived from statistical fits of independent annual maxima for Europe were converted to gridded estimates using the simple inverse distance weighted interpolation described by Shepard (1968), and subsequently low-pass filtered. The gridded estimates are mapped into 2 m/s zones shown in Figure 4.8. For codification purposes, the boundary of each zone would ideally be adjusted to follow municipal, state or national divisions (Holmes *et al.*, 2005). Since wind speeds are designated into zones, the specified 50-year return period wind speed for each zone indicates the upper limit of the individual 50-year return period wind speeds within the zone.

For proper comparison to the summary of published and codified 50-year return period wind speeds for the five regions of Europe discussed in Section 2.1, the interpolated wind speeds are recalculated into zones or contours which correspond to the intervals presented by the codified 50-year return period wind speeds as shown within the right panels of Figure 2.2. Use of the original mapping intervals to compare the 50-year return period wind speeds

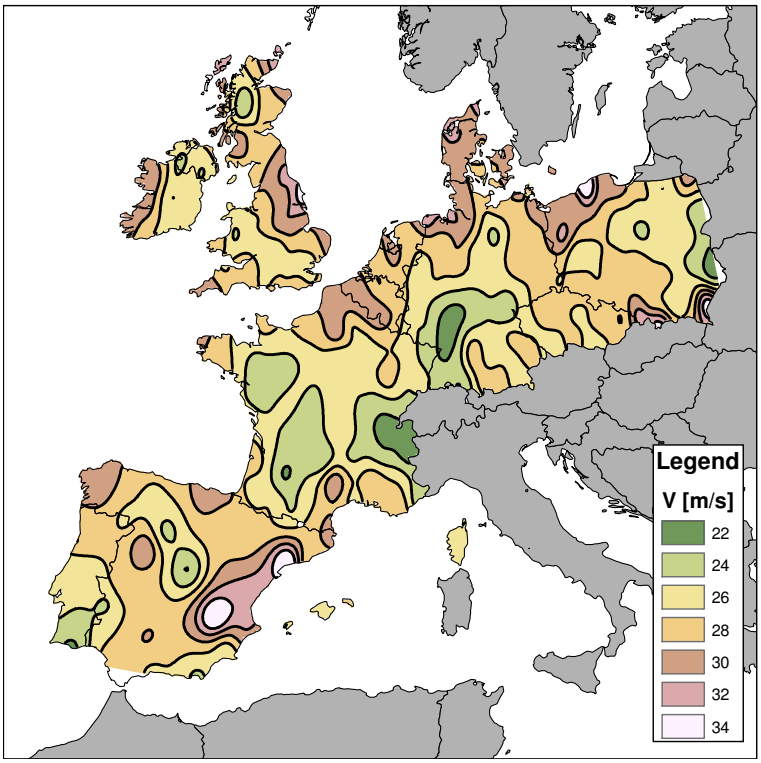
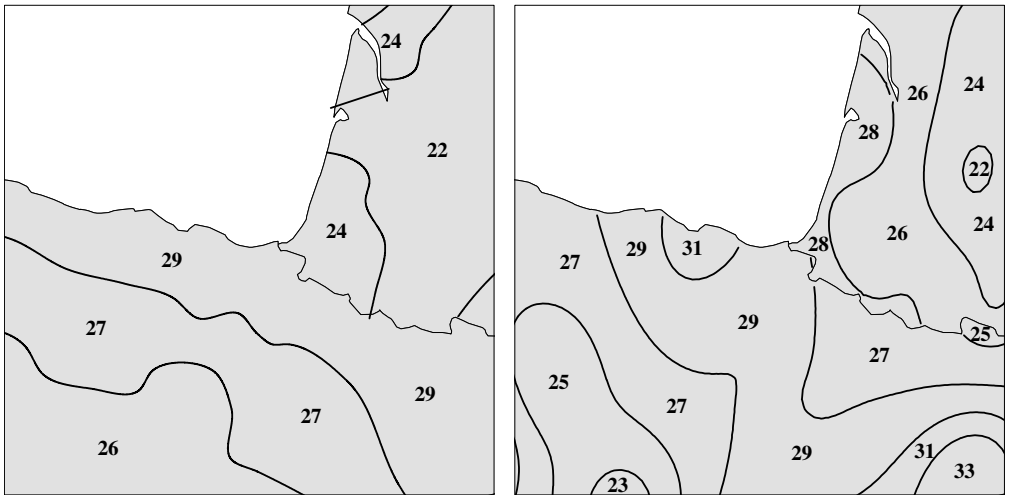


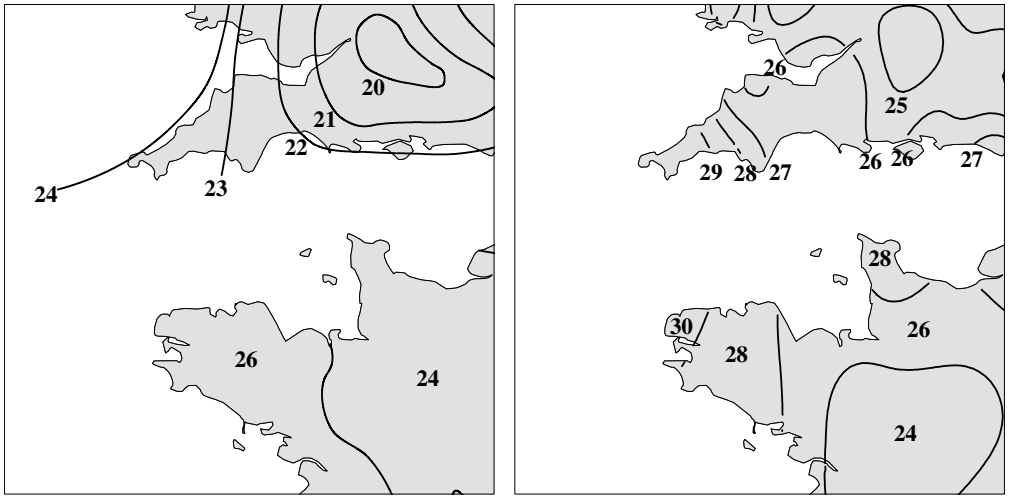
Figure 4.8: 50-year return period wind speed zones (annual maxima)

for the five regions in Europe allows for better examination of differences which arise due to mapping and those which arise due to differences in wind speed predictions. A comparison of the codified 50-year return period wind speeds to the 50-year return period wind speeds derived in the current work using mapping techniques corresponding to the codified values is shown in Figure 4.9. Several observations can be made when comparing the current results to the codified values:

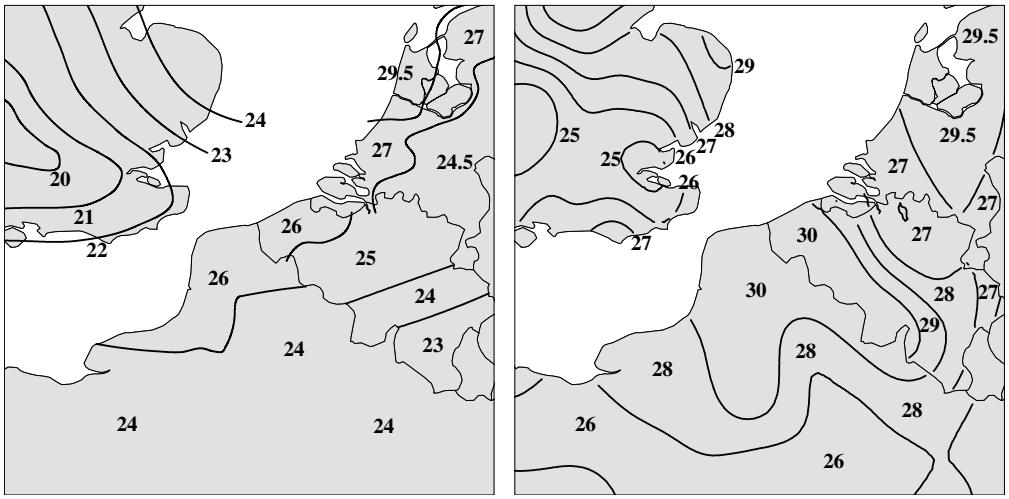
- Section I: France-Spain border
  - Wind speeds for France are greater than the codified estimates. The presented 50-year return period wind speeds correlate better with the original findings of Sacré (2002), shown in Figure 2.2(a, left panel).
  - Wind speeds exhibit similar magnitudes in Spain, however, there are differences in the contouring. Peaks and pits which appear in the contouring are the result of the low spatial resolution of high-quality stations in Spain.
  - Differences along the border through the Pyrenées mountain range, identifiable in the topographic map shown in Figure 4.10, are greatly reduced.
- Section II: English Channel (West)
  - Predictions for the UK exhibit a similar spatial pattern to codified values and those from Miller *et al.* (2001). However, the presented 50-year return period wind speeds are 5 m/s greater throughout the region.
  - Values in France along the English Channel are approximately 2-4 m/s greater which falls between the current codified values and those from Sacré (2002).
  - A difference of up to 1-2 m/s exists between proposed 50-year return period wind speeds across the English Channel.
- Section III: English Channel (East)
  - 50-year return period wind speeds are approximately 4 m/s greater than those



(a) Section I: France-Spain border

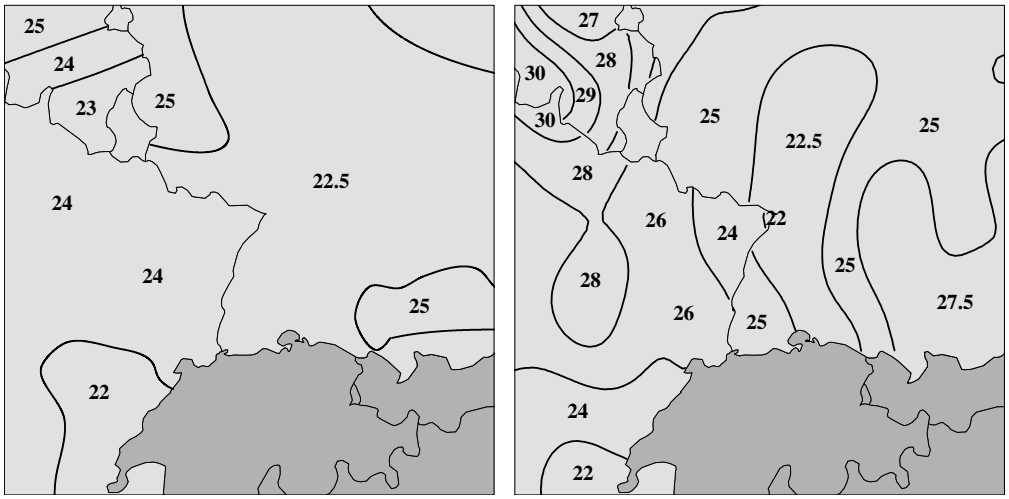


(b) Section II: English Channel (West)



(c) Section III: English Channel (East)

Figure 4.9: Comparisons of European 50-year return period wind speed maps, National Annexes (*left*) and current work (*right*), *continued on next page*



(d) Section IV: France-Germany border

(e) Section V: Denmark-Germany border

Figure 4.9: Comparisons of European 50-year return period wind speed maps, National Annexes (*left*) and current work (*right*)

cited in the Belgian NA due to an unusual 30 m/s wind speed band penetrating inland along the France-Belgium border.

– Coastal wind speeds for the Netherlands are similar to codified values, however, the wind speed zones penetrate further inland.

– Similar to Section II, France and the UK both exhibit greater 50-year return period wind speeds. Differences across the English Channel, vary from 1-3 m/s.

- Section IV: France-Germany border
  - Estimates through the Rhine valley along the southern borders of France and Germany exhibit low values and match well with the respective code estimates. The stations in the region are likely sheltered from strong winds from the west as a result of the north/south orientation of the valley as shown in Figure 4.10.
  - 50-year return period wind speeds have similar magnitudes across the border in both cases, thus, the lack of continuity is due to differing zone intervals, as suggested in Section 2.1.
- Section V: Denmark-Germany border
  - 50-year return period wind speeds are consistent and similar in magnitude in both maps across the Netherlands-Germany border.
  - Coastal wind speed estimates for northern Germany are correlated relatively well with the codified values.

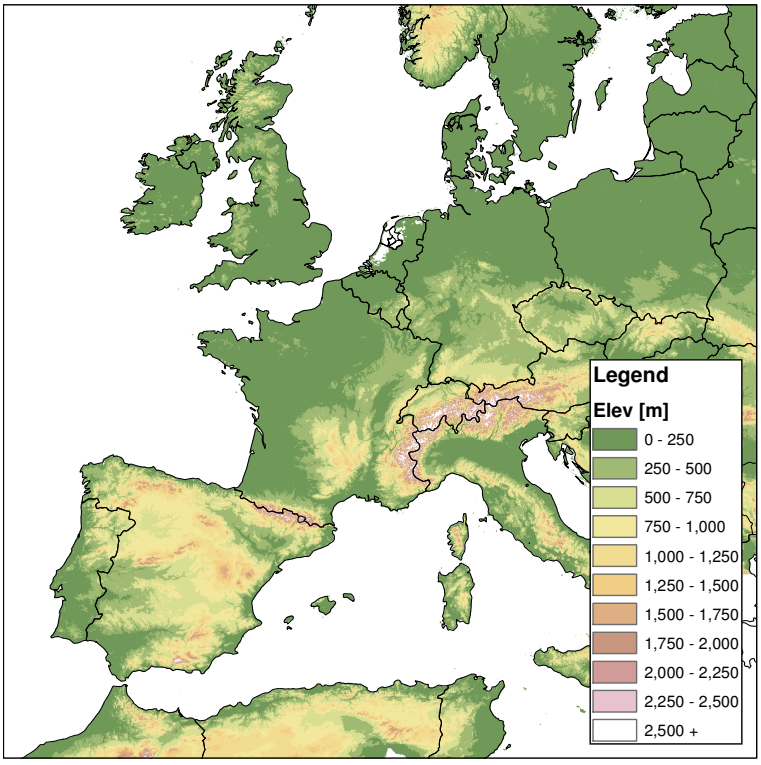
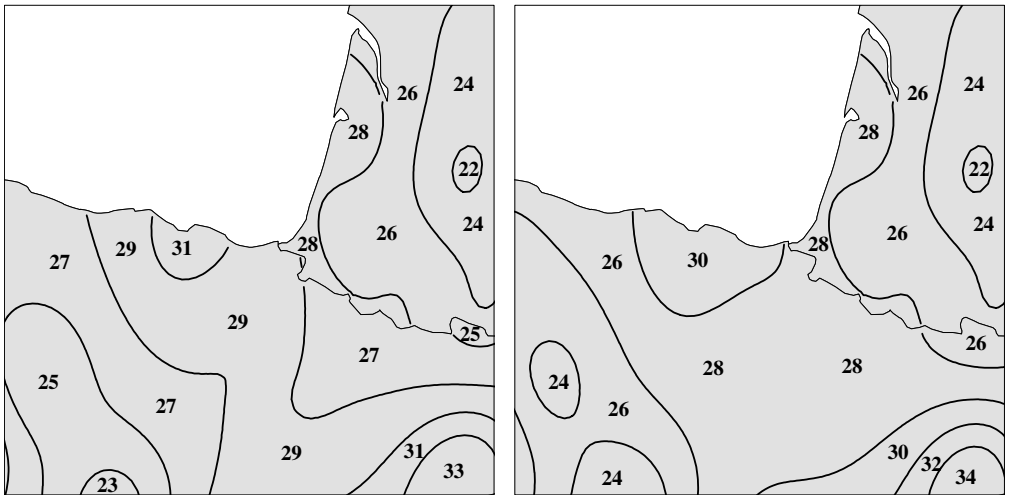


Figure 4.10: Elevation: USGS GTOPO30

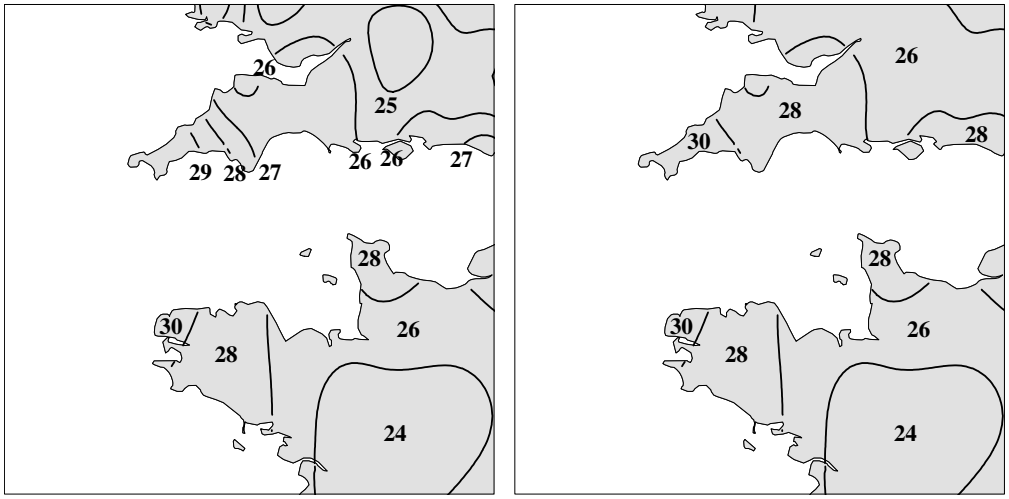
  - Danish NA 50-year return period wind speeds are 3 to 4 m/s lower than estimates from the current work and the codified values, which transition smoothly from northern Germany to Denmark.

- General observations from Figure 4.8
  - The 50-year return period wind speeds match the results of Burton and Allsop (2009a) for eastern Ireland if the contours are grouped into 2 m/s increments, while estimates for western Ireland are 1-2 m/s larger.
  - Despite differences in magnitude for the UK, the current work and Cook and Prior (1987) similarly indicate a region of reduced 50-year return period wind speeds in northern UK.
  - The 30 m/s 50-year return period wind speed band following the northeastern border of Germany penetrates much further inland than indicated by the German NA and Kasperski (2002).

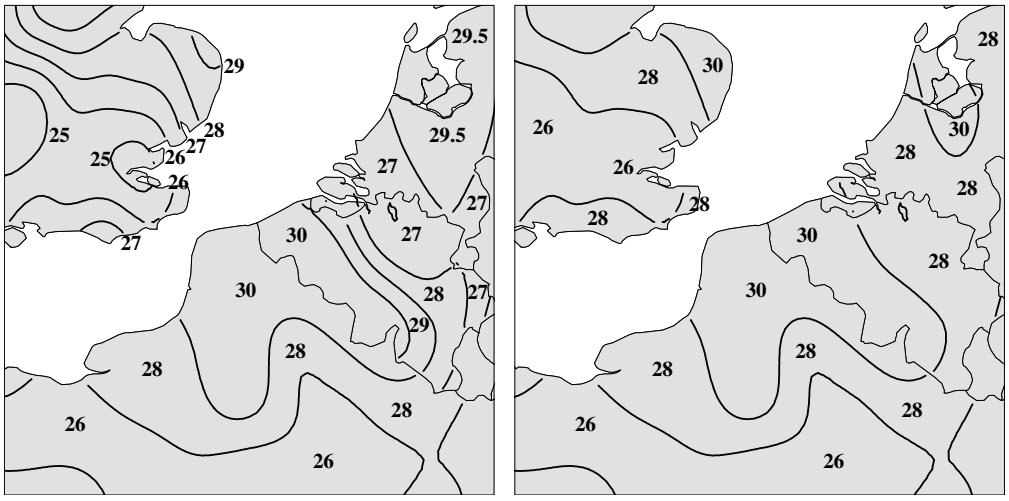
Differences between the 50-year return period wind speeds calculated in the current work and those from national and published 50-year return period wind speed maps are not unexpected. Discrepancies will arise based on the number of steps considered in standardising the observed wind speeds and the methods undertaken to calculate such corrections, each of which will vary from study to study. By using a consistent methodology, differences were largely reduced in Sections I, II, and V due to changes in magnitude of the 50-year return period wind speeds, while Sections III and IV are thought to display additional differences primarily due to mapping. By considering a common zoning system, we seek to further reduce differences in Sections I, II, and V and show that the differences in Sections III and IV are related to zoning. The five regions above are re-examined by comparing the computed 50-year return period wind speeds mapped with zones corresponding to each national code, to the consistent 2 m/s zoning system presented in Figure 4.8. The mapping differences are shown for each of the five regions in Figure 4.11. For regions II-V an increase in continuity



(a) Section I: France-Spain border

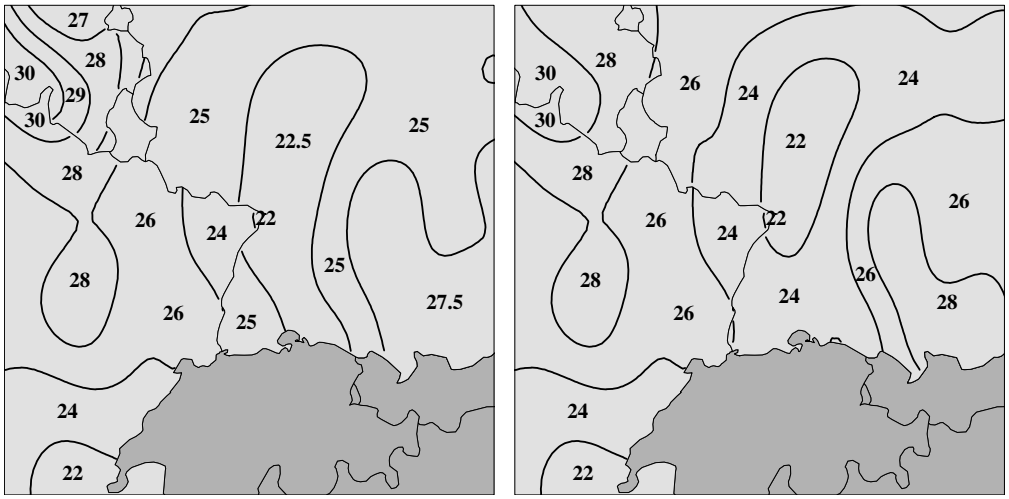


(b) Section II: English Channel (West)

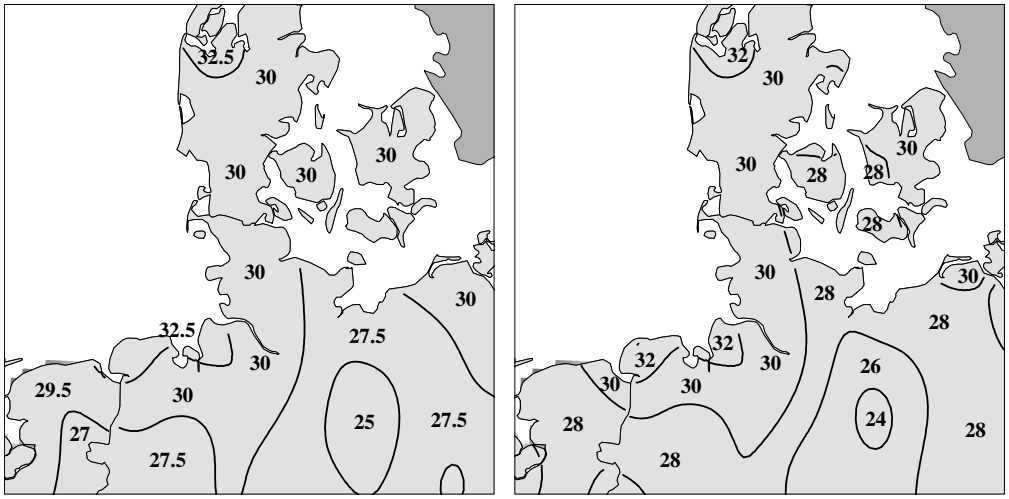


(c) Section III: English Channel (East)

Figure 4.11: Comparisons of European 50-year return period wind speed maps based on the current work, code-based mapping (*left*) and consistent zoning (*right*), *continued on next page*



(d) Section IV: France-Germany border



(e) Section V: Denmark-Germany border

Figure 4.11: Comparisons of European 50-year return period wind speed maps based on the current work, code-based mapping (*left*) and consistent zoning (*right*)

of 50-year return period wind speeds across borders is apparent and shows that differences up to 2 m/s can appear strictly as a result of contouring or zoning. The only exception is region I which indicates a natural shift may occur across the Pyrenées.

The greatest difference between the 50-year return period wind speeds computed in the current work and previously published or codified values occurs for the UK. The appropriateness of the 5 m/s increase in 50-year return period wind speeds requires additional discussion. In the preliminary study by Gatey and Miller (2007) discussed in Section 2.1,

50-year return period wind speeds along the English Channel were not corrected for upstream exposure. The wind speeds for each site were corrected for two different irrespective of direction, or circular, cases. The first case was based on the assumption that the upstream exposure was equal to the roughness length at the site, typically airport exposure. The second case was based on the assumption that, depending on the fetch from the station to the English Channel, the observations may be representative of open-water exposure. In theory, the true exposure correction factor accounting for upstream effects likely falls within these limits, depending on the distance from the coast. For coastal stations along the southern UK, 50-year return period wind speed ranges for Plymouth and St. Catherine's Point were found to be 23.4 - 33.3 m/s and 27.9 - 35.1 m/s respectively. The estimates along the northern coast of France were 22.1 - 31.4 m/s, 23.5 - 26.4 m/s and 21.4 - 24.0 m/s for Brehat, Guernsey and Jersey respectively. Based on the current zoning system, Figure 4.8 indicates a 50-year return period wind speed of 28 m/s with a nearby region of 26 m/s. The zonal wind speed of 28 m/s would suggest that the individual calculated wind speed predictions in the region are approximately 26 - 28 m/s, which correlates well with the limits identified by Gatey and Miller (2007).

The 50-year return period wind speeds of Miller *et al.* (2001) for the UK were corrected for exposure using the Deaves and Harris IBL model. The resulting 50-year return period wind speeds along the English Channel range from $20 - 24$ m/s as shown in Figure 2.2(b,c). Recalling the comparison of correction factors calculated from three different exposure correction models for Den Helder Airport, De Kooy, NL shown in Figure 3.7, the Deaves and Harris IBL model was found to over-predict the slowdown associated with a transition from smooth to rough exposure. Based on exposure correction alone, the 50-year return period wind speeds calculated by the existing study could be 10-percent greater if one of the TL models discussed in Sections 3.3.4 and 3.3.5 were alternatively selected.

Overall, differences between national 50-year return period wind speeds were reduced

through a consistent methodology for calculating 50-year return period wind speeds and a consistent zoning system. Based on the preliminary study of Gatey and Miller (2007) and differences in exposure correction, a 5 m/s increase in 50-year return period wind speeds for the UK is plausible. Lastly, Figure 4.8 indicates there are still several regions which are locally affected by a single station. Sharp changes in these regions arise due to an inability to correct for every factor influencing a site and, in certain regions, a lack of spatial resolution of stations. These issues are addressed in Chapters 5 and 6.

4.4.4 Directionality

A climatic or directional factor in building codes allows designers to reduce the 50-year return period wind speeds for non-dominant wind directions. When considering directionality, it is important to avoid directional masking (Moriarty and Templeton, 1983). Directional masking occurs when a maximum wind speed for a sector is not recorded as a result of a greater wind speed being observed in a different sector. Cook (1982, 1983) calculated a climatic factor based on the ratio of the extrapolated sectoral wind speed prediction to the irrespective of direction prediction. These values were then adjusted to account for correlation by considering a series of k-factors. An alternative approach proposed by Melbourne (1984) extrapolates sectoral wind speed estimates to smaller exceedance probabilities than the irrespective of direction prediction. The adjusted exceedance probabilities are a function of the number of sectors over which storms are expected to participate. The two methods were compared by Vega (2008) and it was found that the final directional factors were similar between the two methods when normalised by the largest directional factor.

Coles and Walshaw (1994) suggest resolving wind speed and direction observations into components and calculating the magnitude of the wind speed for each azimuth. The method alleviates the concern of directional masking as a time history is constructed for each sector.

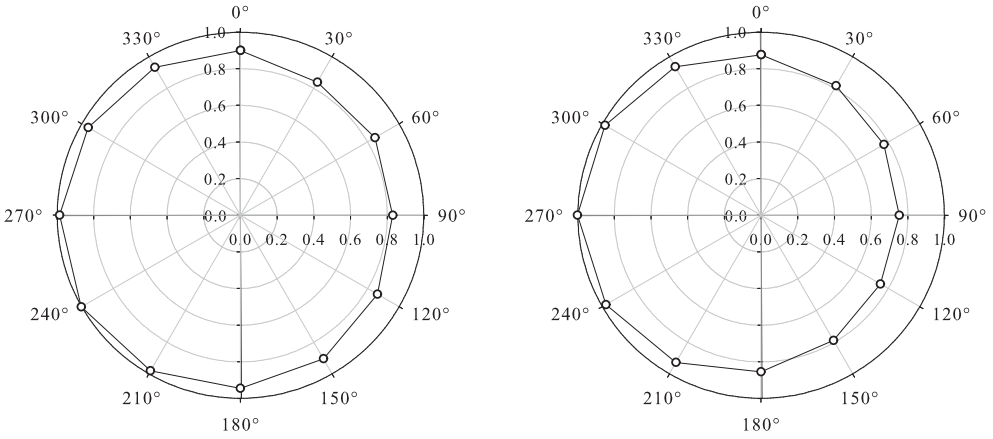
Correlation across wind directions will exist as a result of calculating the wind components for adjacent directions and secondly as a result of fluctuations in direction during a wind event. After calculation of individual time histories for each azimuth, if independence is assumed across directions then separate distributions can be fitted for each directional sector (Coles and Walshaw, 1994). The resulting distributions cannot be used to directly calculate the joint probability across multiple sectors without further consideration of the correlation across wind directions (Coles and Walshaw, 1994; Palutikof *et al.*, 1999). Independent predictions for each wind sector will suffice for the current work since the joint probability across multiple sectors is not required.

For each station, the time histories of wind speed and direction are used to resolve the observations into wind components at each azimuth. The analysis forms 37 time histories for each station, one for the original irrespective of direction case and one for each 10-degree sector. The directional time histories are combined into time histories for 30-degree sectors by selecting the largest annual maximum from the three contributing directions, e.g. for each storm, the 30-degree sector centred at 0-degrees consists of the greatest annual maxima from the 350-, 0- and 10-degree azimuths. 50-year return period wind speed estimates are calculated for each 30-degree azimuth and normalised by the sector containing the largest estimate. For each nation, the normalised values for each sector are averaged across all stations to provide a directional factor representative of the region. Nations with greater spatial coverage, such as France and Germany, are subdivided into regions which are thought to exhibit different directionality based on typical storm tracks of northeast travelling ETCs. The resulting directional factors are shown in Figure 4.12. Further investigation suggests that the directional factors can be combined into six climates based on similarities between neighbouring nations:

- I - Portugal, Spain
- II - France (Atlantic), France (Interior)

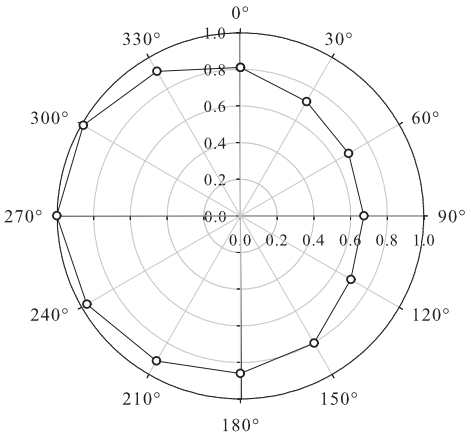
- III - Ireland, United Kingdom, France (English Channel)
- IV - Germany (South), Belgium, Netherlands
- V - Germany (North), Denmark
- VI - Czech Republic, Poland

The directional factors for the UK shown in Figure 4.12(e) match very well with the results of Cook and Prior (1987) and Miller *et al.* (2001). Burton and Allsop (2009a) recommend the use of the directional factors applied in the UK for Ireland. A comparison of the directional plots of Ireland and the UK shown in Figures 4.12(d,e) respectively, suggest that Ireland is susceptible to stronger winds from the south than the UK. The directional variations indicate that ETCs approaching from the south and making landfall in the UK are weaker after crossing France than those which make landfall in the south of Ireland. The use of the UK factors in Ireland may not be appropriate. If the directional factors of Ireland and the UK are to be combined as suggested here, the largest directional factor for each direction should be adopted. The distributions of climatic factors for the six regions support the underlying assumption of grouping nations based on the location relative to the typical storm tracks of ETCs.

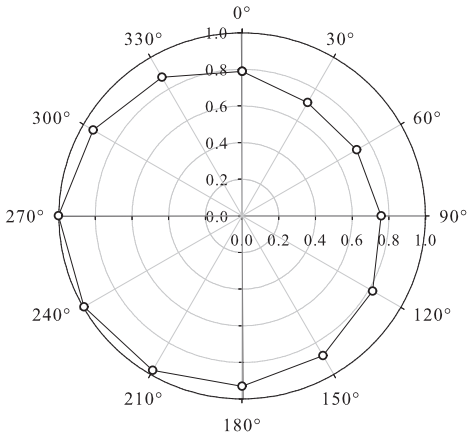


(a) Portugal (b) Spain

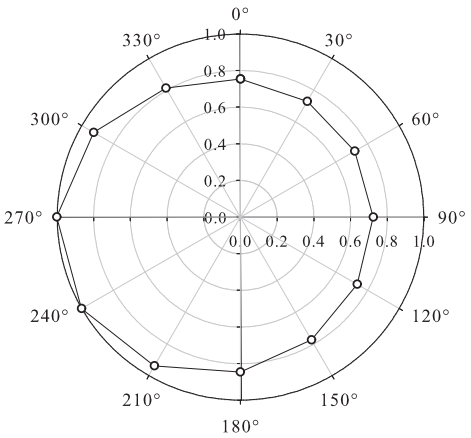
Figure 4.12: Directional factors by nation, *continued on next page*



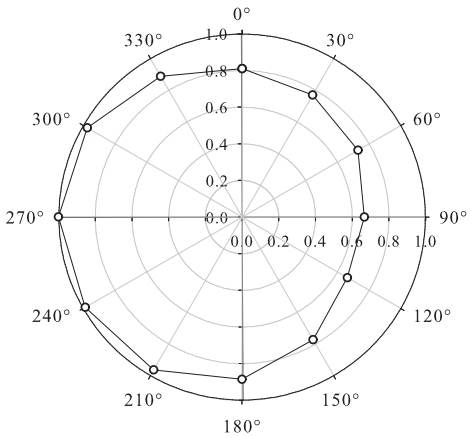
(c) France (Atlantic)



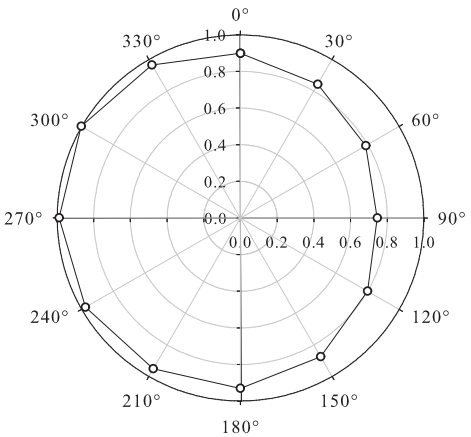
(d) Ireland



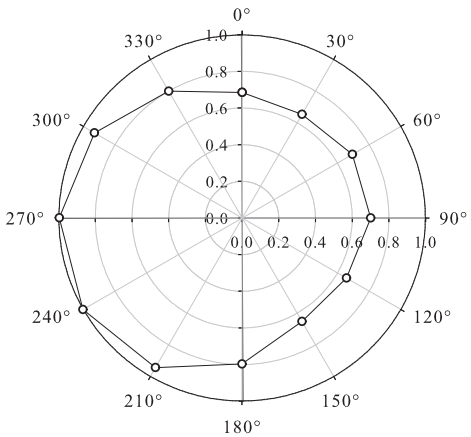
(e) United Kingdom



(f) France (English Channel)

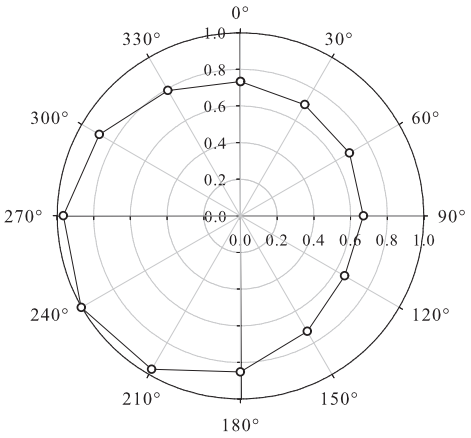


(g) France (Interior)

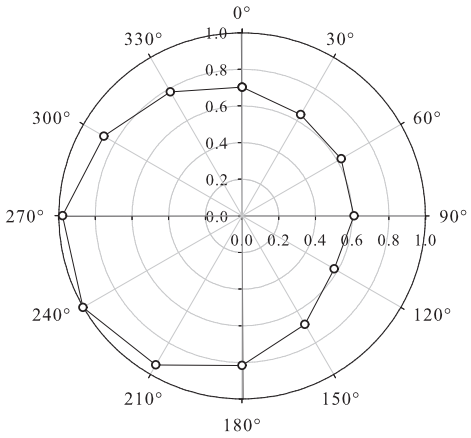


(h) Germany (South)

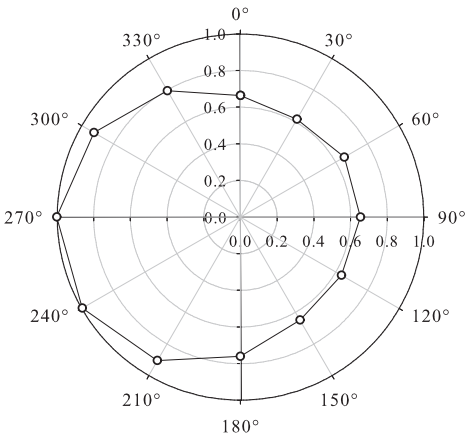
Figure 4.12: Directional factors by nation, *continued on next page*



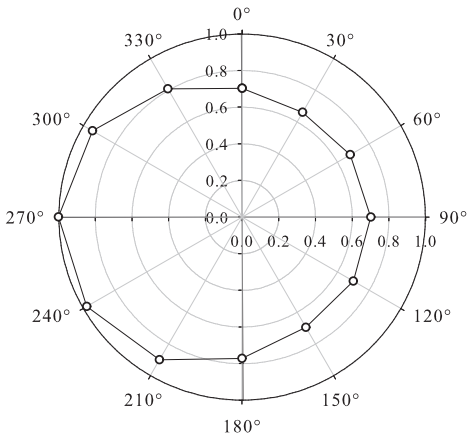
(i) Belgium



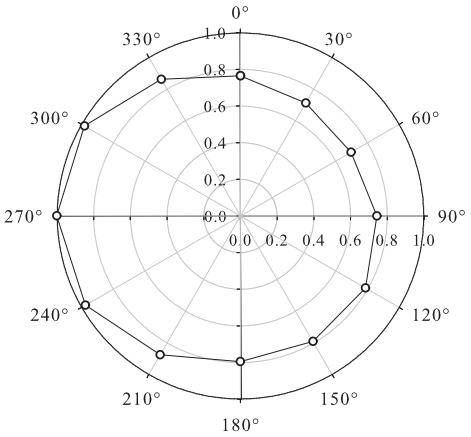
(j) Netherlands



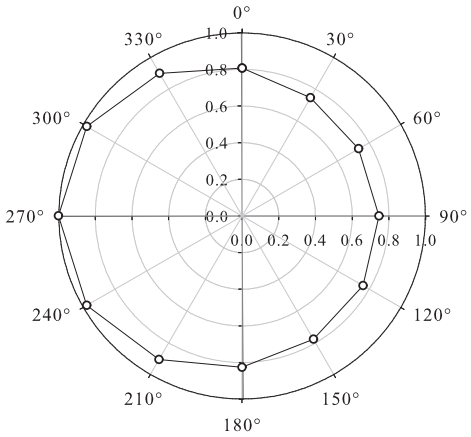
(k) Germany (North)



(l) Denmark



(m) Czech Republic



(n) Poland

Figure 4.12: Directional factors by nation

Chapter 5

Background Wind Field

50-year return period wind speeds have historically been calculated from a time history of surface wind speed observations, upper-air wind speed measurements obtained from radiosondes, or wind speeds calculated from pressure fields. Most analyses utilise surface wind measurements, as in Chapter 4, or wind fields calculated from pressure fields. Radiosonde data is not ideal for the current work since the instruments are not typically released in severe wind conditions. Wind speeds calculated from pressure fields are typically based on the assumption of geostrophic balance, thus, frictionless flow occurs between straight, parallel isobars. The geostrophic drag law is utilised to calculate the associated surface wind speed from the geostrophic wind components (Miller, 2003; Larsén and Mann, 2009). Wind fields which are assumed to be representative of upper-level wind fields, such as geostrophic wind fields, will ideally be less influenced by surface roughness and topography. Therefore, an extreme value analysis of wind fields derived from the upper-level wind fields should neglect localised surface effects and depict the overall spatial variability of extreme wind speeds. The following sections provide analyses of various methods used for deriving upper-level wind fields to form a background 50-year return period wind speed map.

5.1 Background

The assumption of geostrophic balance allows calculation of wind components, assumed to be representative of the upper-level wind field, directly from the balance of the Coriolis force and pressure gradient. Watson *et al.* (2001) used mean sea-level pressure (MSLP) data obtained from the US National Centers for Environmental Prediction (NCEP) for a 13-year period (1985-1997) at 6-hour intervals. After interpolation of the MSLP data from a 2.5-degree to 0.5-degree grid, the geostrophic wind field was calculated for each pressure field. The geostrophic balance is based on the underlying assumption of straight parallel isobars, thus, for strong ETCs experienced in northern Europe, the assumption will not likely hold. Kristensen and Jensen (1999) found that the gradient wind speed better represented upper-air wind speeds, particularly in instances where isobars exhibit strong curvature.

Watson *et al.* (2001) validate the assumption of geostrophic balance by comparing the calculated geostrophic wind speed and direction obtained from MSLP data to radiosonde observations, reporting overall agreement between the two datasets. Further inspection of the comparison reveals that good correlation was achieved between the two datasets when comparing 'normal' winter and summer days. However, the two datasets did not match for ETCs, including the Burns' Day Storm (January, 1990). Despite differences between predicted and observed wind speeds during ETCs, the geostrophic wind speed estimates were concluded to be an 'excellent representation of frictionless flow'. Direct comparison of geostrophic wind speeds to radiosonde data for the Burns' Day Storm showed geostrophic estimates of approximately 55 m/s were much greater than radiosonde estimates of 30 m/s. Although the directional data from the radiosonde showed good agreement, the wind speed data was deemed unreliable and a wind speed cut off for radiosonde of 30 m/s was suggested. Depending on the location of the radiosonde relative to the storm maxima, the true value likely resides between the two values, as the inclusion of the centrifugal force

in the calculation of the gradient wind speed acts to reduce the magnitude of wind speeds calculated from the geostrophic balance for cyclones in the Northern Hemisphere.

Comparison of monthly mean wind speeds from radiosonde measurements and geostrophic estimates matched well. The largest differences were found to occur in months where the wind speeds were highest (Watson *et al.*, 2001). Assuming the geostrohpic drag law has been applied consistently, if the largest discrepancies occur in months where wind speeds are highest, the results indicate the geostrophic wind speeds are overestimating the true upper-air wind speed in strong wind situations. The result is likely due to the inappropriate application of the geostrophic approximation for strong depressions, as will be shown in Section 5.3.1

Based on the conclusion that the geostrophic winds and radiosonde exhibit good agreement by Watson *et al.* (2001), both Miller (2003) and Larsén and Mann (2009) validate use of the geostrophic balance to calculate wind fields from MSLP fields. Miller (2003) digitised MSLP maps to calculate geostrophic wind fields every 6 hours for the period of 1953-1995. Extreme value analyses were performed at each grid point using the GPD to extrapolate estimates to 50-year return period levels. Larsén and Mann (2009) calculated geostrophic wind fields from MSLP data acquired from the NCEP/NCAR re-analysis. Annual extremes were fitted with the EVD type I using PWM. Both reports utilised the geostrophic drag law to calculate 50-year return period wind speeds at the surface. 50-year return period wind speeds for Europe from both reports exhibit similar trends as contours tend to follow coastlines and increase in magnitude from south to north. The results of Larsén and Mann (2009) tend to be 1-2 m/s higher than those of Miller (2003), however, the former adjusted estimates for disjunct sampling using the model of Larsén and Mann (2006).

5.2 ECMWF Re-analysis

The European Centre for Medium-Range Weather Forecasts (ECMWF) has produced several re-analyses through a joint effort with meteorological and environmental institutions. The ECMWF 40-years re-analysis (ERA-40) is a re-analysis of collected meteorological observations for the 45-year period from September, 1957 to August, 2002 (Uppala *et al.*, 2005). The basic ERA-40 datasets are interpolated to a latitude-longitude grid having a spatial resolution of 2.5 by 2.5 degrees. The most recent re-analysis is the ECMWF Interim re-analysis (ERA-Interim) which is available for the period of January, 1989 to present on a latitude-longitude grid of 1.5 by 1.5 degrees (Berrisford *et al.*, 2009). Surface and pressure level datasets are available at 6-hour intervals for both re-analyses and are available from the ECMWF data server for research use. The ERA-Interim is selected for the current work since the re-analysis contains updated models and better spatial resolution than the ERA-40. The re-analysis data is processed in a uniform manner which should provide consistency between statistical predictions across Europe. Such consistency may not be present for individual station observations acquired from multiple European meteorological agencies.

5.3 Wind Fields from Mean Sea Level Pressure Data

The surface analysis dataset of the ERA-Interim contains records of MSLP data which are used in the current section to calculate wind fields representative of upper-level wind conditions. MSLP fields are obtained from the ERA-Interim and fit using bi-cubic splines to interpolate to a 0.5-degree grid. The equations of the bi-cubic splines allow the derivatives of pressure to be calculated directly. Two examples of the resulting MSLP fields are shown in Figures 5.1 and 5.2 for the Burns' Day Storm (January, 1990) and Anatol (December,

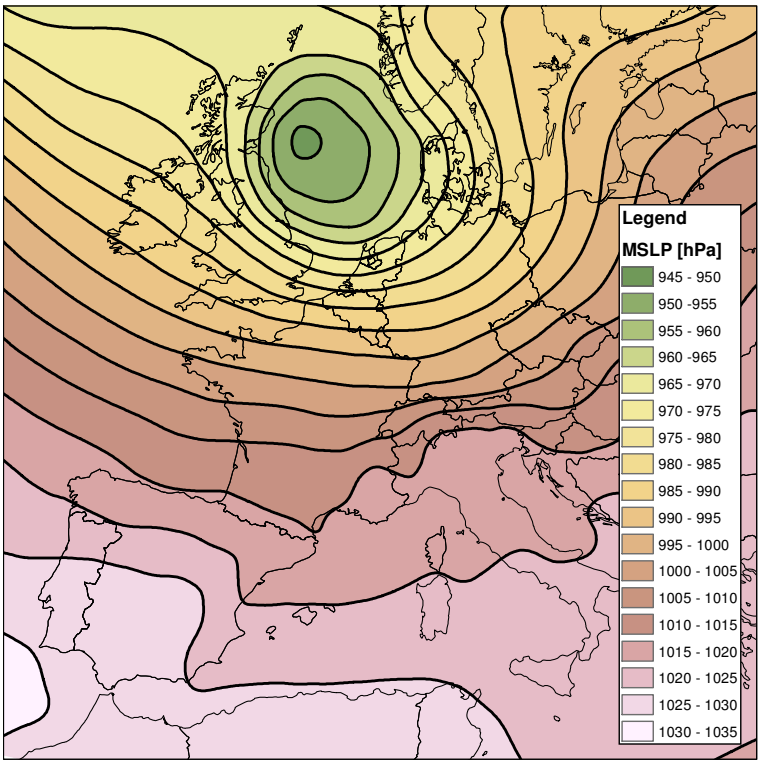


Figure 5.1: MSLP field: Burns' Day Storm, January 26, 1990 at 0000 hr

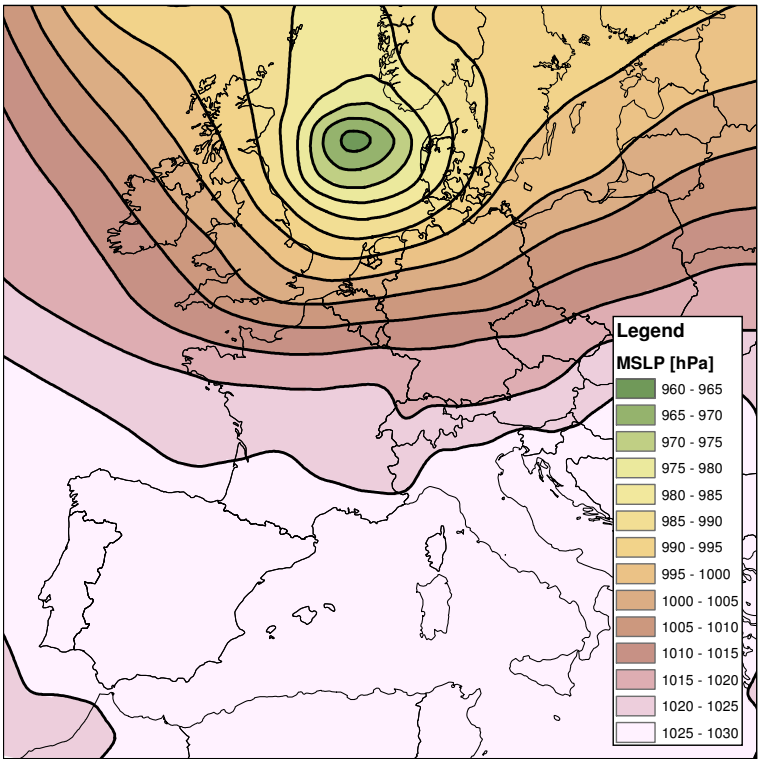


Figure 5.2: MSLP field: Anatol, December 3, 1999 at 1800 hr

1999) respectively. Calculation of wind fields from the re-analysis MSLP data can be carried out with varying degrees of complexity. Based on a scale analysis of the Navier-Stokes equation for an incompressible fluid with constant viscosity, the viscous terms can be neglected for synoptic-scale events (Holton, 2004). The resulting equations considered here for the x- and y-directions are

$$\frac{Du}{Dt} = -\frac{1}{\rho}\frac{\partial P}{\partial x} + fv \tag{5.1}$$

$$\frac{Dv}{Dt} = -\frac{1}{\rho}\frac{\partial P}{\partial y} - fu \tag{5.2}$$

where u and v are the wind components in directions x and y respectively, D/Dt represents the material derivative, ρ is the air density, P is pressure, and f is the Coriolis parameter. Various approximations to the Equations 5.1 and 5.2 are explored in the subsequent sections, including the geostrophic, quasi-geostrophic and semi-geostrophic approximations.

5.3.1 Geostrophic Approximation

When considering a force balance along straight, parallel isobars there is approximate balance between the pressure gradient and Coriolis forces, known as the geostrophic balance. The geostrophic approximation is calculated from MSLP data which is an idealised pressure field at sea-level estimated from surface-level measurements. The associated wind field represents frictionless flow at sea-level, which is assumed to describe the upper-level wind field sufficiently far from the influence of the surface. The simple force balance reduces Equations 5.1 and 5.2 to the two geostrophic wind components, given as

$$u_g = -\frac{1}{\rho f}\frac{\partial P}{\partial y} \tag{5.3}$$

$$v_g = \frac{1}{\rho f}\frac{\partial P}{\partial x}. \tag{5.4}$$

The appropriateness of calculating wind fields using the geostrophic approximation was discussed in Section 5.1 based on the results of Watson *et al.* (2001). The geostrophic approximation was concluded to provide poor estimates of wind speeds near strong depressions. Since depressions dominate the European synoptic wind climate, and the extremes are of interest here, the geostrophic approximation is inadequate. The geostrophic wind fields calculated from the MSLP fields in Figures 5.1 and 5.2 are shown in Figures 5.3 and 5.4, respectively, for later comparison.

5.3.2 Quasi-geostrophic Approximation

The quasi-geostrophic approximation improves upon the geostrophic approximation through consideration of the advection terms in the Navier-Stokes equation. The quasi-geostrophic approximation assumes that the geostrophic wind components are much larger than the ageostrophic wind components. The ageostrophic wind components are defined as the difference between the true wind components and the geostrophic wind components, such that

$$u_a = u - u_g \tag{5.5}$$

$$v_a = v - v_g. \tag{5.6}$$

The approximation implies that advection is largely governed by geostrophic advection and the velocity gradients may be well represented by the geostrophic velocity gradients. The quasi-geostrophic approximation can then be written as

$$\frac{D_g u_g}{D_g t} - f v_a = 0 \tag{5.7}$$

$$\frac{D_g v_g}{D_g t} + f u_a = 0 \tag{5.8}$$

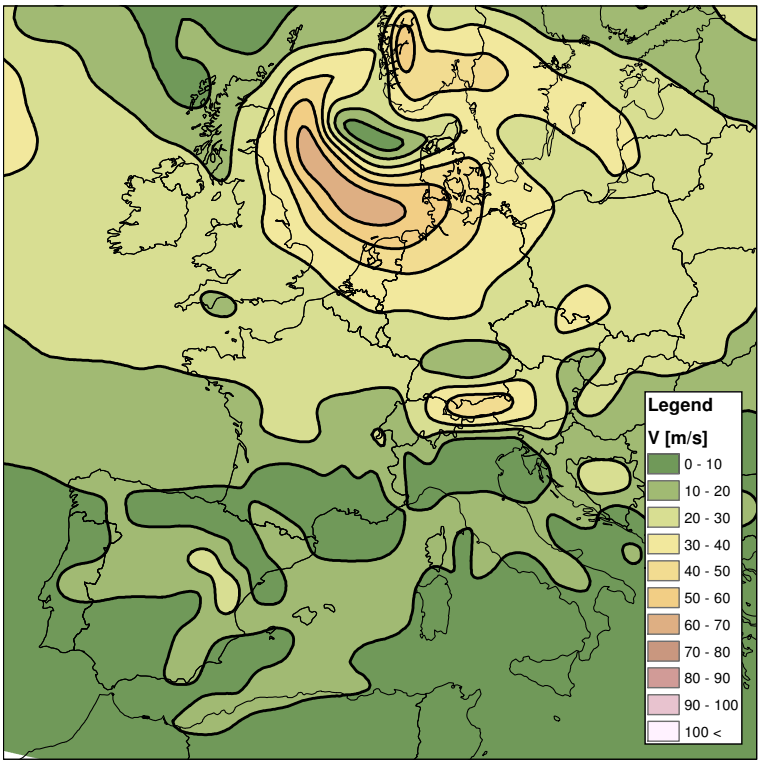


Figure 5.3: Geostrophic wind field: Burns' Day Storm, January 26, 1990 at 0000 hr

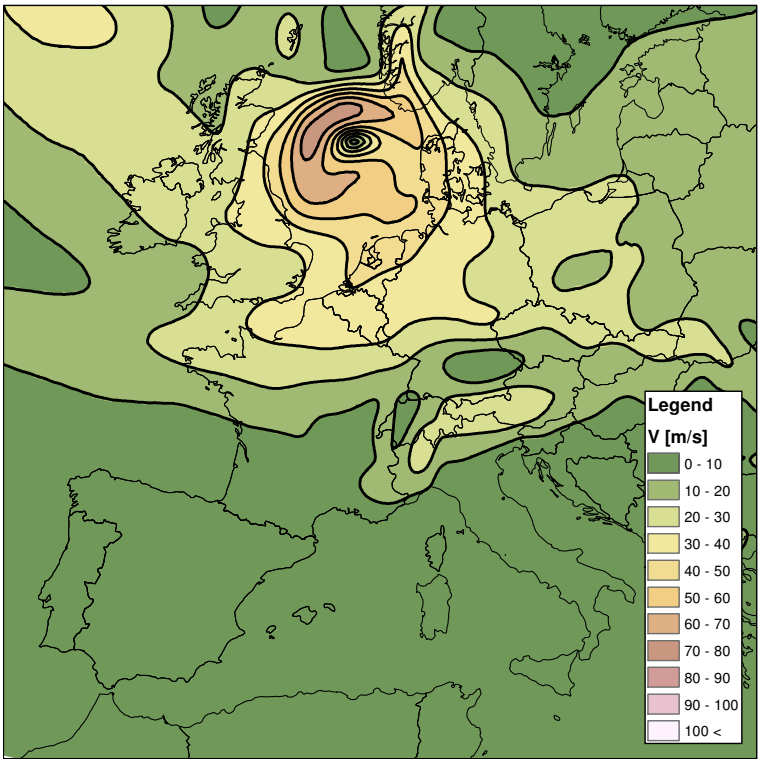


Figure 5.4: Geostrophic wind field: Anatol, December 3, 1999 at 1800 hr

where expansion of the left-hand side leads to

$$\frac{\partial u_g}{\partial t} + u_g\frac{\partial u_g}{\partial x} + v_g\frac{\partial u_g}{\partial y} - fv_a = 0 \tag{5.9}$$

$$\frac{\partial v_g}{\partial t} + u_g\frac{\partial v_g}{\partial x} + v_g\frac{\partial v_g}{\partial y} + fu_a = 0. \tag{5.10}$$

Substitution of the geostrophic wind components defined by Equations 5.3 and 5.4, and rearranging, gives

$$v_a = -\frac{1}{\rho f^2}\frac{\partial^2 P}{\partial y \partial t} + \frac{1}{\rho^2 f^3}\frac{\partial P}{\partial y}\frac{\partial^2 P}{\partial x \partial y} - \frac{1}{\rho^2 f^3}\frac{\partial P}{\partial x}\frac{\partial^2 P}{\partial y^2} \tag{5.11}$$

$$u_a = -\frac{1}{\rho f^2}\frac{\partial^2 P}{\partial x \partial t} + \frac{1}{\rho^2 f^3}\frac{\partial P}{\partial y}\frac{\partial^2 P}{\partial x^2} - \frac{1}{\rho^2 f^3}\frac{\partial P}{\partial x}\frac{\partial^2 P}{\partial x \partial y}. \tag{5.12}$$

The quasi-geostrophic wind field can be calculated by substituting Equations 5.3, 5.4, 5.11 and 5.12 into Equations 5.5 and 5.6 and solving for u and v. The quasi-geostrophic wind fields for the Burns' Day Storm and Anatol are shown in Figures 5.5 and 5.6 respectively. From both figures it is obvious that the quasi-geostrophic approximation has failed in regions near the storm centre by producing unrealistic estimates. Sharp pockets of wind speeds are located throughout Europe, particularly in regions exhibiting large topographic variation such as Sweden, and in the Alps through Switzerland and northern Italy. The contributing factors to these regions are explored further in the following section.

5.3.3 Semi-geostrophic Approximation

The semi-geostrophic approximation improves upon the quasi-geostrophic approximation by including the effects of ageostrophic advection. Strong depressions and fronts will be better described by the semi-geostrophic equations (Hoskins, 1975). In the semi-geostrophic approximation, advection is carried out by the combination of both geostrophic and ageostrophic velocities, however, the velocity gradients are still approximated by the geostrophic

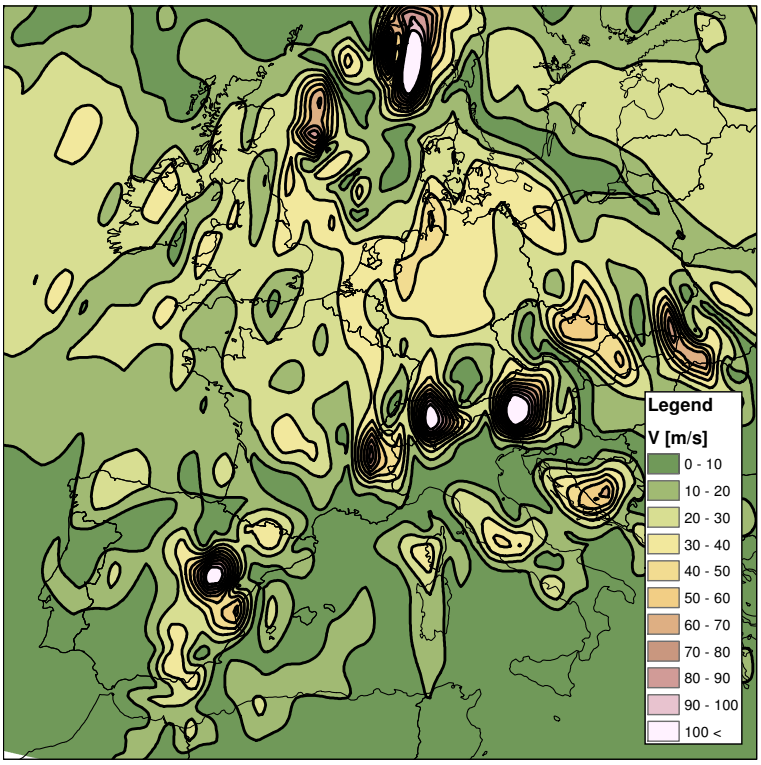


Figure 5.5: Quasi-geostrophic wind field: Burns' Day Storm, January 26, 1990 at 0000 hr

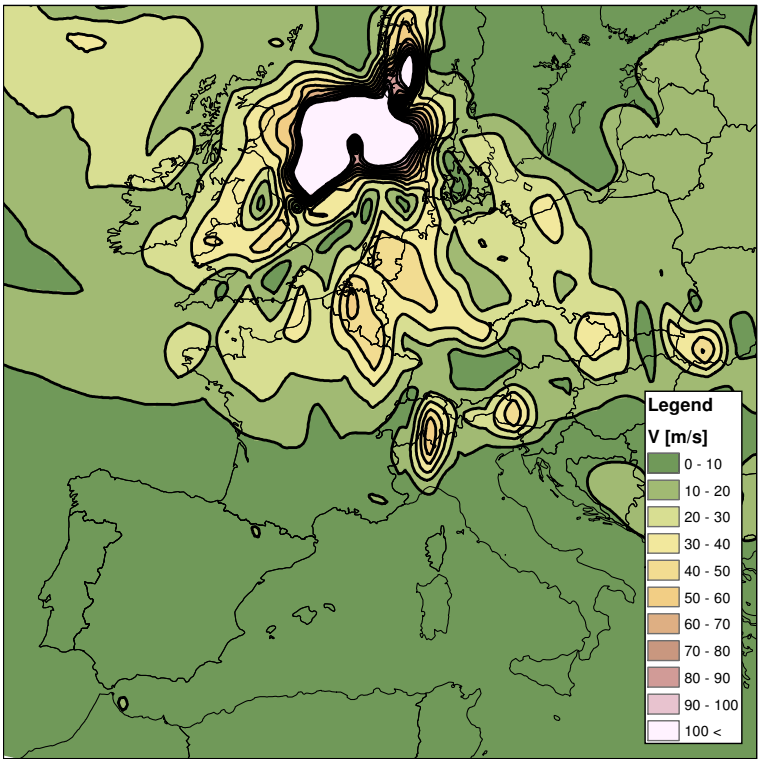


Figure 5.6: Quasi-geostrophic: Anatol, December 3, 1999 at 1800 hr

velocity gradients such that

$$\frac{Du_g}{Dt} - fv_a = 0 \tag{5.13}$$

$$\frac{Dv_g}{Dt} + fu_a = 0 \tag{5.14}$$

and expansion of the left-hand side leads to

$$\frac{\partial u_g}{\partial t} + u\frac{\partial u_g}{\partial x} + v\frac{\partial u_g}{\partial y} - fv_a = 0 \tag{5.15}$$

$$\frac{\partial v_g}{\partial t} + u\frac{\partial v_g}{\partial x} + v\frac{\partial v_g}{\partial y} + fu_a = 0 \tag{5.16}$$

where substitution of the geostrophic wind components defined by Equations 5.3 and 5.4, and rearranging, gives

$$-\frac{1}{\rho f}\frac{\partial^2 P}{\partial y \partial t} - \frac{u}{\rho f}\frac{\partial^2 P}{\partial x \partial y} - \frac{v}{\rho f}\frac{\partial^2 P}{\partial y^2} - fv + \frac{1}{\rho}\frac{\partial P}{\partial x} = 0 \tag{5.17}$$

$$\frac{1}{\rho f}\frac{\partial^2 P}{\partial x \partial t} + \frac{u}{\rho f}\frac{\partial^2 P}{\partial x^2} + \frac{v}{\rho f}\frac{\partial^2 P}{\partial y \partial x} + fu + \frac{1}{\rho}\frac{\partial P}{\partial y} = 0. \tag{5.18}$$

Equations 5.17 and 5.18 can be rearranged for direct calculation of the wind components. The semi-geostrophic wind fields for the Burns' Day Storm and Anatol are shown in Figures 5.7 and 5.8 respectively. The approximation breaks down in regions near the storm centre and additional regions of failed estimates are found throughout southern Europe and Sweden. These regions correlate strongly to regions of significant elevation shown in Figure 4.10. Thus, the underlying equations used in the re-analysis to reduce surface pressure observations to MSLP values appear to perform poorly in these regions. Since the second derivatives of the pressure field are required, small although abrupt changes in MSLP appear to greatly affect estimates of the wind components. To reduce the influence of these errors, gaussian and mean smoothing filters are applied to the MSLP fields. Smoothing of the MSLP fields greatly reduces the number of erroneous estimates as shown in Figure 5.9.

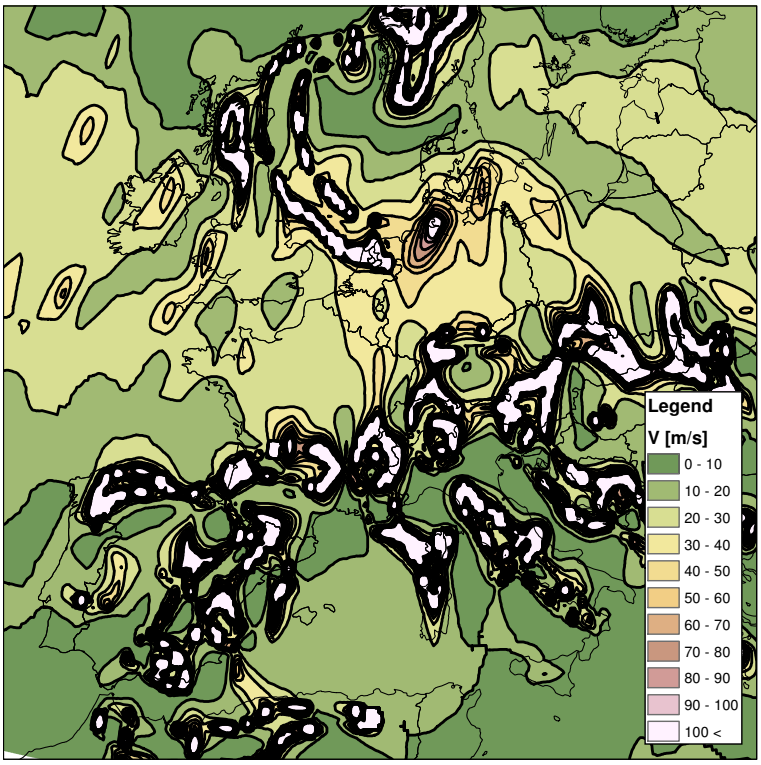


Figure 5.7: Semi-geostrophic wind field: Burns' Day Storm, January 26, 1990 at 0000 hr

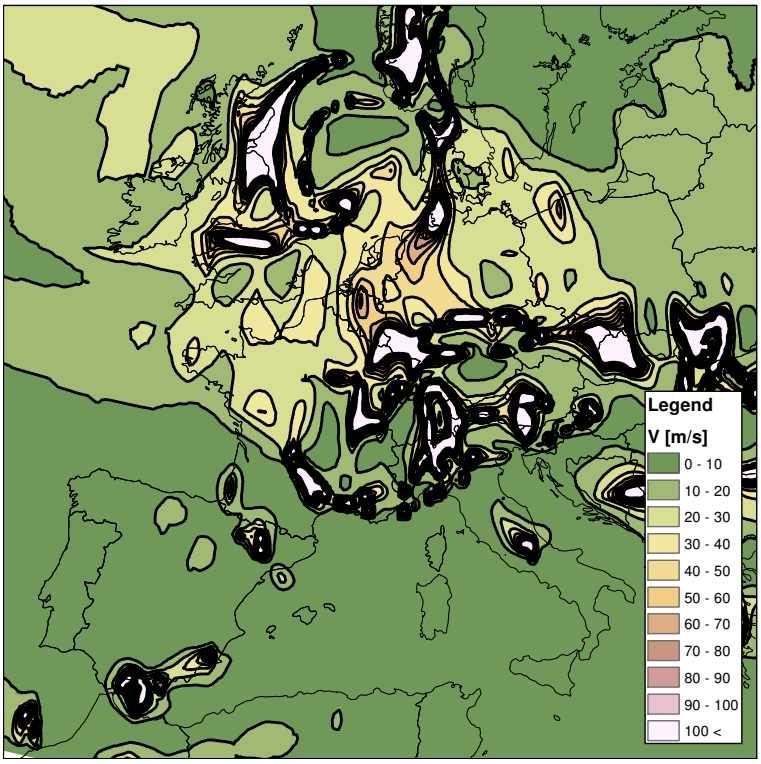


Figure 5.8: Semi-geostrophic wind field: Anatol, December 3, 1999 at 1800 hr

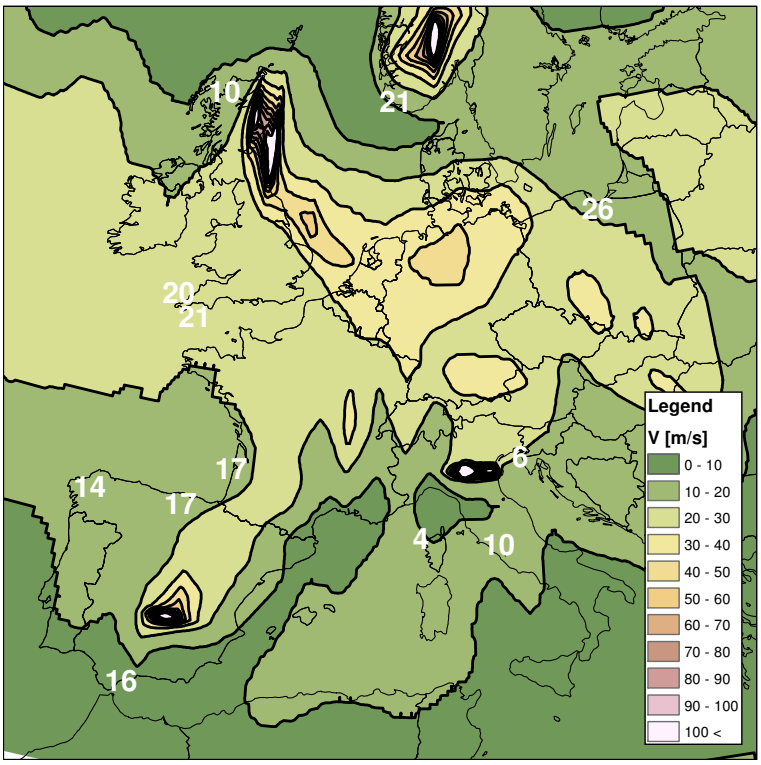


Figure 5.9: Semi-geostrophic wind field (Smoothed MSLP field): Burns' Day Storm, January 26, 1990 at 0000 hr:

The values overlaid in Figure 5.9 are radiosonde measurements quoted by Watson *et al.* (2001) which correlate well, despite failure of the approximation in several regions. Smoothing of the MSLP fields is not preferable as the algorithms will also decrease the pressure gradient near the centre of a depression, further reducing wind speeds in the region. Overall, the methods considered here to calculate wind fields representative of upper-level wind conditions from MSLP fields, fail to provide accurate results for the full spatial extents of the two ETCs considered.

5.4 Wind Fields from Pressure-level Data

Methods of calculating upper-level wind fields from MSLP data have been shown to be analytically intractable for the methods considered in the previous section. Alternatively,

formulation of a background wind field can be calculated directly from wind fields provided by the ERA-Interim. Wind components at 10 m height are available from the ERA-Interim, however, a background field which is relatively unaffected by surface friction and topography is desired. For wind fields selected at pressure-levels sufficiently far from the surface, wind speed estimates should not be greatly influenced by either condition. The ERA-Interim provides *u*- and *v*-wind components at various pressure-levels which can be utilised to calculate upper-air wind fields. Since the height of a constant pressure-level will vary spatially, the wind direction and magnitude at 1000 m are linearly interpolated from wind components at two pressure surfaces. Given the typical synoptic boundary layer profile, shown in Figure 3.5, linear interpolation between wind speeds at heights away from the surface will be adequate. The 875 and 925 mb pressure levels are selected for the current analysis. Calculation of the wind speeds at a specific height above the surface requires not only consideration of the geopotential of the pressure-level, but the geopotential of the

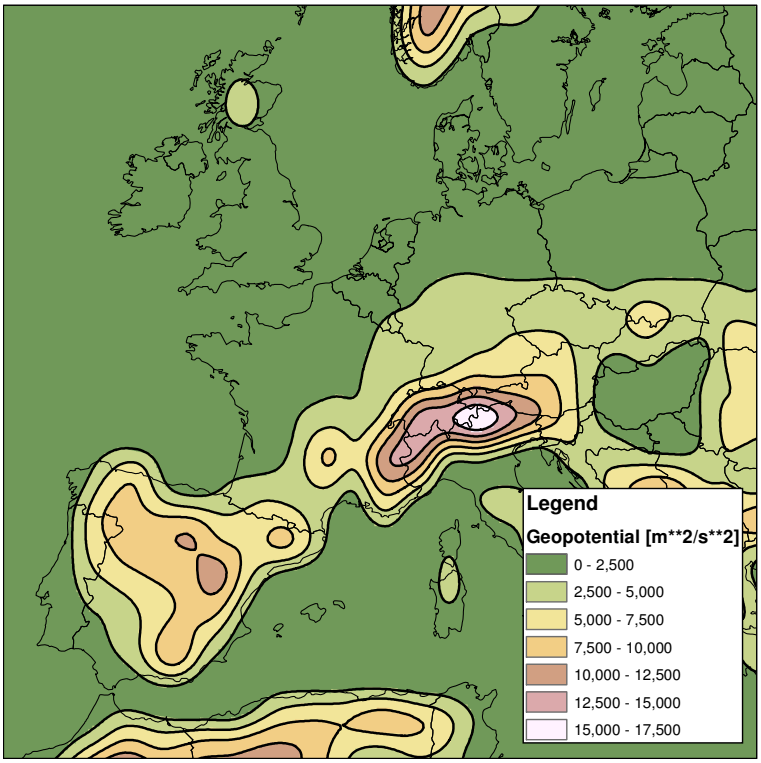


Figure 5.10: Surface geopotential of the ERA-Interim (invariant)

surface. The geopotential of the surface used in the ERA-Interim is shown in Figure 5.10. The resulting 1000 m wind fields for the Burns' Day Storm and Anatol are shown in Figures 5.11 and 5.12. The Burns' Day Storm wind field matches well with the radiosonde data cited by Watson *et al.* (2001) and wind fields in both figures tend to display the typical placement of the high wind speed region on the right-hand side of a cyclone in the Northern Hemisphere. Maximum wind speeds are lower than the corresponding geostrophic estimates in Figures 5.3 and 5.4, which is expected if the curvature of the isobars is appropriately considered. Due to the spatial resolution of the re-analysis dataset, the features of the wind field are not as well resolved as the wind fields calculated from interpolated pressure fields.

5.5 Results

The upper-level wind fields at 1000 m calculated in Section 5.4 are used to estimate the wind speed at 10 m height for a roughness length of 0.05 m using the Gryning ABL profile from Section 3.2.2. Wind speeds at individual points are further corrected for disjunct sampling intervals of 6-hours before OBRE are applied to fit the EVD type I to annual maxima wind speeds. The fitted parameters are used to calculate a background 50-year return period wind field for Europe shown in Figure 5.13. To evaluate the potential effects of terrain on the estimates, the above procedure was repeated for wind fields based on 1000 m wind components which were not corrected for the geopotential of the surface. The resulting 50-year return period wind speeds are shown in Figure 5.14. Comparison of Figures 5.13 and 5.14 shows estimates through Switzerland and northern Italy exhibit the greatest variation in wind speed as a result of the largest surface geopotential. For countries of interest in the current work, appreciable differences are observed for Spain and southern Germany when surface geopotential is considered. The final wind speeds over Spain and

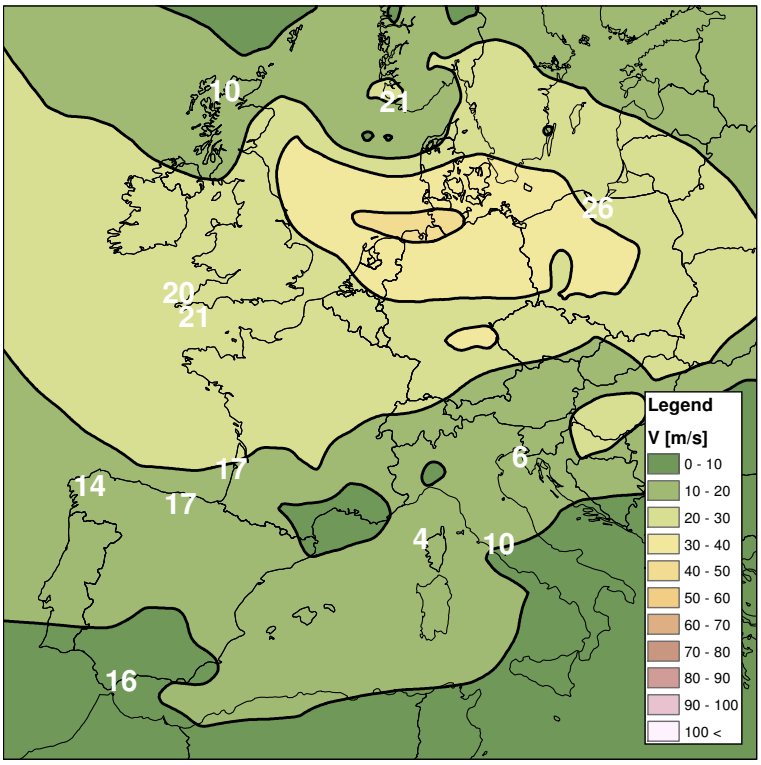


Figure 5.11: Wind field at 1000 m: Burns' Day Storm, January 26, 1990 at 0000 hr

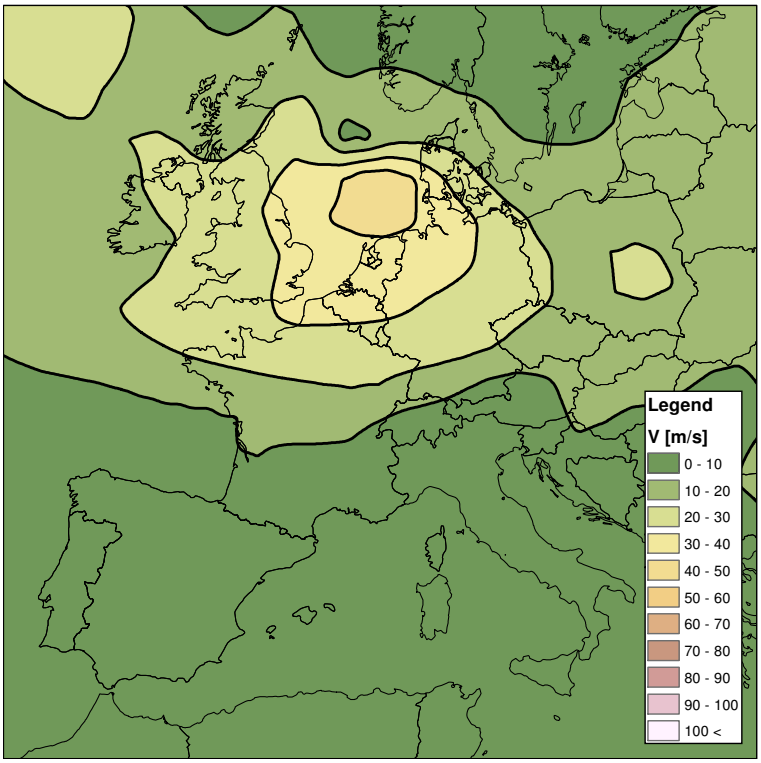


Figure 5.12: Wind field at 1000 m: Anatol, December 3, 1999 at 1800 hr

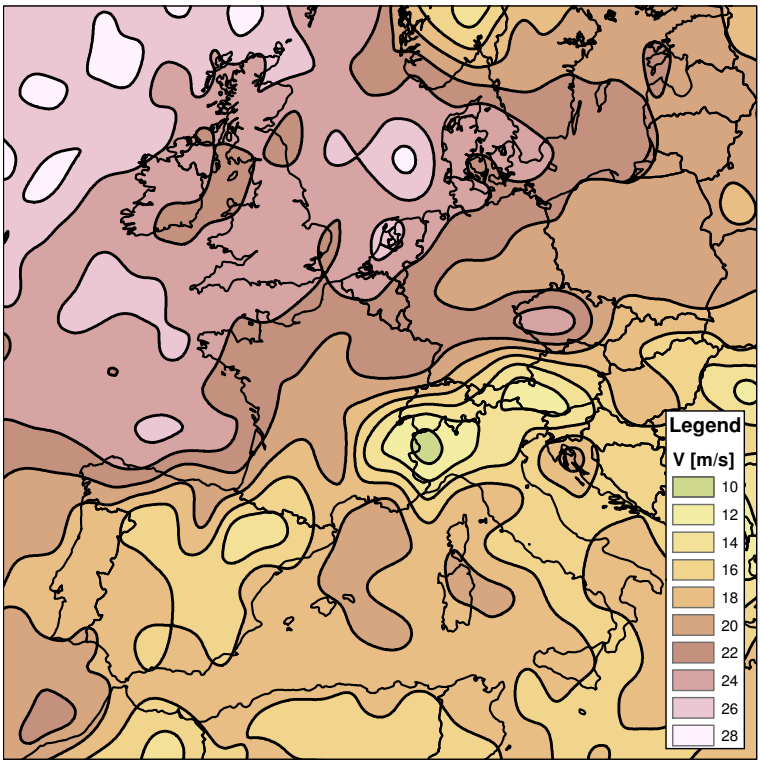


Figure 5.13: 50-year return period wind speeds adjusted for surface geopotential

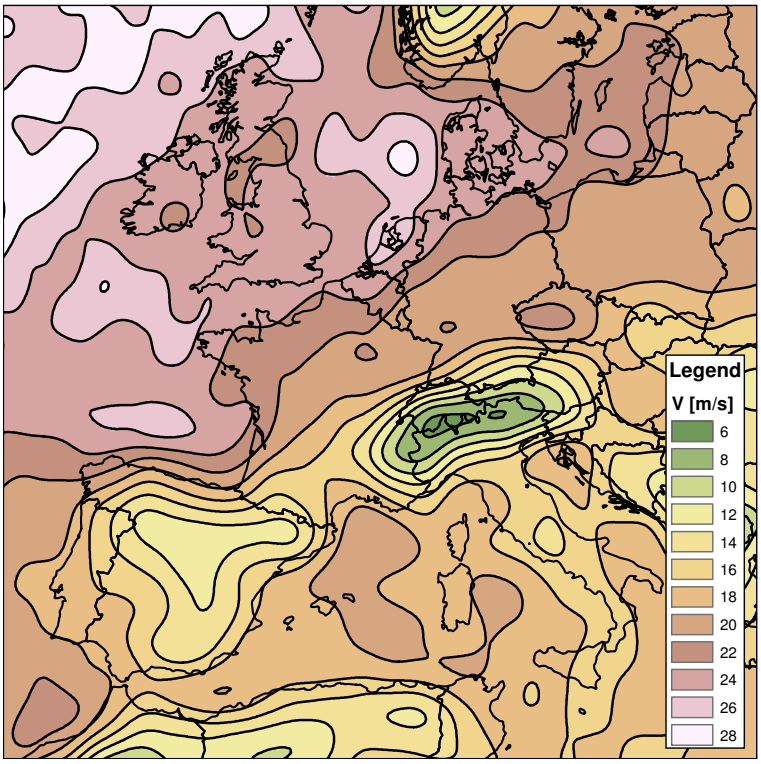


Figure 5.14: 50-year return period wind speeds unadjusted for surface geopotential

much of southern Europe still appear to follow surface geopotential contours despite being appropriately corrected.

Since the procedure carried out here to calculate 50-year return period wind speeds from annual maxima is identical to the procedure in Section 4.4, the empirically derived A-D test statistic threshold should apply. Comparison of the A-D test statistics indicates the A-D test statistic from the previous chapter corresponds to an A-D test statistic percentile of 99.5 based on the statistical fits of the background wind speeds. The improved statistical fits to the background wind data may be a result of smoothing of the observations by the re-analysis model. The impacts of smoothing are explored later in this section. Due to the A-D test statistic threshold, six grid points contributing to the background 50-year return period wind field are removed.

The 50-year return period wind speeds shown in Figure 5.13 are compared to the results of Miller (2003) and Larsén and Mann (2009), from which the following conclusions can be made:

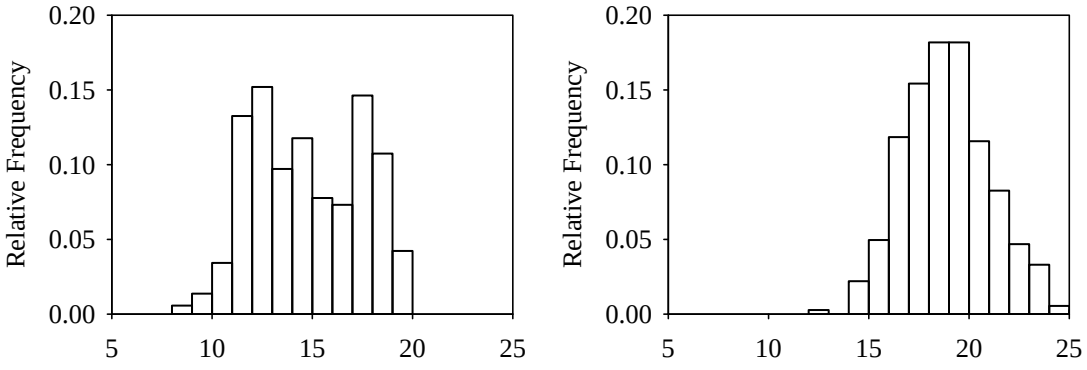
- Miller (2003) calculated 50-year return period wind speeds 2 m/s less through France, Germany and the UK and 3-4 m/s less over Denmark.
- Accounting for contouring differences, Larsén and Mann (2009) found similar wind speeds for Ireland, UK and nations bordering the North Sea and English Channel.
- The spatial distribution of wind speeds over Spain is different for Larsén and Mann (2009).
- Larsén and Mann (2009) have a 50-year return period wind speed low of 15 m/s in the Czech Republic, while the current analysis found a high region of 24 m/s.

In comparison to the results of Chapter 4, the 50-year return period wind speeds from the current chapter are:

- 2-6 m/s lower for the UK, Denmark, Netherlands, and Belium.
- 6-8 m/s lower for France and Poland.
- 10-15 m/s lower for Spain.

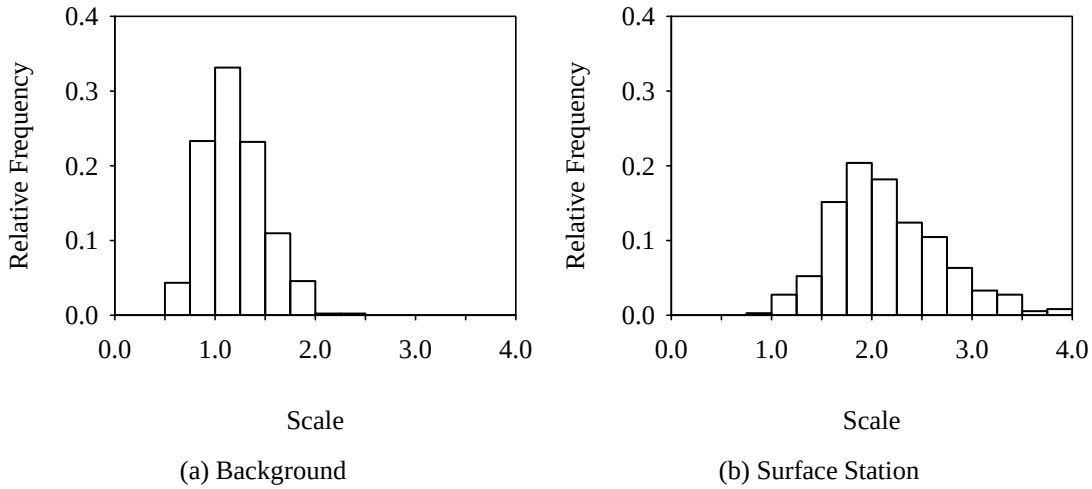
The underestimation of 50-year return period wind speeds by the background wind field is not unexpected as observations are spatially and temporally smoothed when assimilated for the re-analysis. Application of the disjunct sampling correction derived in Section 3.4 should somewhat alleviate the impact of the latter, however, spatial smoothing remains uncorrected as shown by the differences noted above. Smoothing of observed data will likely reduce the magnitude of the extreme wind speeds measured at point locations. In these instances, fits of the EVD type I are expected to exhibit reduced location and scale parameters due to a general reduction of the magnitude of annual maxima. Histograms of the fitted parameters computed here, and those calculated in Chapter 4 for EVD type I fits of annual maxima from surface stations, are shown in Figures 5.15 and 5.16 for the location and scale, respectively. A significant shift in the parameters is clearly exhibited, verifying the expected reductions for both the location and scale parameters due to smoothing.

The location parameters calculated from the background wind field exhibit a bi-modal be-



(a) Background (b) Surface Station

Figure 5.15: Distribution of the location parameters



(a) Background (b) Surface Station

Figure 5.16: Distribution of the scale parameters

haviour as shown in Figure 5.15(a). Further investigation finds the second peak is a reflection of larger location parameters for grid points over-water, particularly the North Atlantic. The shifted peak supports the aforementioned conclusion that regions of terrain impact the re-analysis upper-level wind fields despite consideration of surface geopotential and the calculation of wind speeds at heights significantly far above the surface.

The annual maxima of greatest magnitude may exhibit larger reductions due to spatial smoothing if the events are relatively localised compared to the grid resolution. In theory, if all wind observations are equally affected by smoothing, a constant factor equivalent to the ratio of the background to surface location parameter should be applied to all wind speeds. As a result, the scale parameter should also reflect an equivalent reduction. The mean values of the location parameters for the background and surface based 50-year return period wind speeds are 14.7 and 19.0 m/s respectively. The background field is restricted to only grid points within the countries of interest in the current work. Thus, the ratio of the background to surface location parameter is approximately 0.78. If all wind speeds are assumed to be equivalently affected by smoothing, the ratio of the scale parameters should be similar. Carrying out the same analysis for the scale parameters, the mean values for the background and surface scale parameters are 1.2 and 2.2 respectively, resulting in a ratio

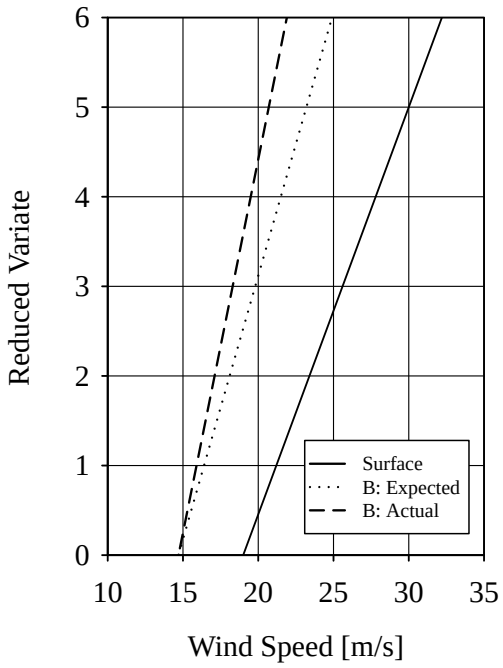


Figure 5.17: Comparison of mean background and surface EVD type I fits

of 0.55. The difference between the two ratios indicates the largest annual maxima are reduced more significantly than the lower annual maxima. Figure 5.17 shows the influence of smoothing on the mean parameter values discussed here.

Chapter 6

Data Assimilation

Methodologies for deriving 50-year return period wind speed maps from surface observations and upper-level wind fields have been evaluated and established independently in Chapters 4 and 5 respectively. The corrections made to surface observations to calculate 50-year return period wind speeds in Chapters 3 and 4 will never fully account for all conditions influencing a specific location. The background 50-year return period wind field is representative of the broader climatological scales and was shown in Chapter 5 to provide estimates in which maxima are smoothed both spatially and temporally. A combination of these two techniques could provide an ideal 50-year return period wind speed map. In practice, 50-year return period wind speed maps are calculated from a single source of data. The possibility of utilising the 50-year return period wind field derived from upper-air measurements in Chapter 5 to supplement the surface based 50-year return period wind speeds analysed in Chapters 3 and 4 has never been considered. Assimilation techniques allow data obtained from multiple sources to be appropriately merged, and the method based on Bratseth (1986) is considered in the current chapter. The algorithm is selected due to its widespread application in numerical weather forecasting. In numerical weather forecasting, observations are merged with the output from an appropriate numerical weather

prediction (NWP) model through consideration of the error or uncertainty of the two data sources. The resulting analysis field is used to represent the current conditions and initialise an NWP model for the following forecast. The Bratseth method involves iteratively adjusting the background field through use of point observations and is investigated here to create a 50-year return period wind speed map which considers the results of both Chapters 4 and 5.

6.1 Bratseth Scheme

Bratseth (1986) proposed a successive correction method which converges to statistical interpolation, thereby allowing the benefits of statistical interpolation to be achieved in a more computationally efficient manner. Statistical optimal interpolation allows data from multiple sources to be considered jointly, based on the uncertainty associated with each source. The successive correction method proposed by Bratseth (1986) for calculating analysis values, at both grid and observation points, is determined for iteration k from

$$F_{a,i}^{(k+1)} = F_{a,i}^{(k)} + \sum_{j=1}^{n} \alpha_{ij} \left(F_{o,j} - F_{a,j}^{(k)} \right) \tag{6.1}$$

where $F_{a,i}$ is the analysis point, $F_{o,j}$ is a point observation and, for the optimal solution,

$$\alpha_{ij} = \frac{\rho_{ij}}{M_j} \tag{6.2}$$

$$\alpha_{ij} = \frac{\rho_{ij} + \delta_{ij}(\sigma_i^2/\sigma_b^2)}{M_j} \tag{6.3}$$

for the analysis of the grid points and observations respectively (Bratseth, 1986). In Equation 6.3, δ_{ij} represents the Kronecker delta and σ_i^2/σ_b^2 is the ratio of the point variance to

the background variance. The assumption of Gaussian spatial correlations leads to

$$\rho_{ij} = \exp\left(-\frac{d_{ij}^2}{D^2}\right) \tag{6.4}$$

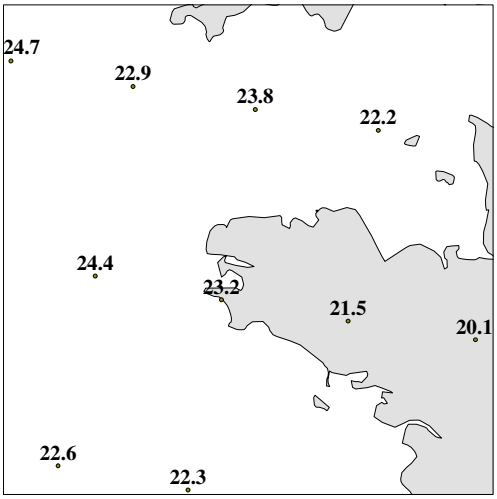
where d_{ij} is the distance between two locations and D is the radius of influence. In Equations 6.2 and 6.3, the spatial correlations are normalised by

$$M_j = (\sigma_j^2/\sigma_b^2) + \sum_k^n \rho_{jk} \tag{6.5}$$

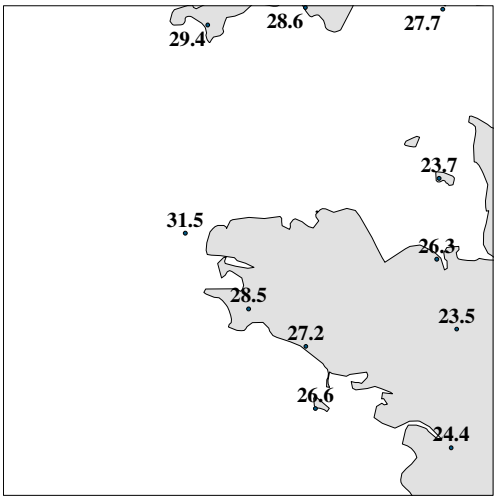
which is representative of the density of weights around the observation. For calculation of the distance between two points, the algorithm given by Vincenty (1975) is selected. The algorithm provided by Vincenty (1975) provides accurate distance estimates due to consideration of the Earth's surface as an ellipsoid rather than a sphere. A sample of the input background 50-year return period wind field, surface point 50-year return period wind speeds and final analysis 50-year return period wind field with corresponding wind zones is shown in Figure 6.1 for the western region of the English Channel and northwestern France. The complete results and discussion are provided in Section 6.3.

6.2 Methodology

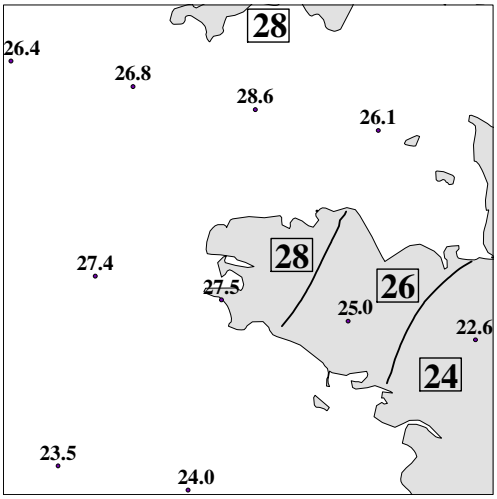
A subset of the results from Chapters 4 and 5 is considered here to examine the suitability of the Bratseth scheme to assimilate surface observations with re-analysis data. The spatial extents of the subset are limited to northern Europe due to the large discrepancies experienced through southern Europe as a result of topography.



(a) Background wind field



(b) Surface based wind speeds



(c) Analysis wind field and zones

Figure 6.1: Example of the Bratseth Scheme

6.2.1 Bratseth Scheme

The background 50-year return period wind field derived in Chapter 5 is bi-linearly interpolated to calculate the 'analysis' 50-year return period wind speed at each of the surface station locations. Based on the calculated asymptotic confidence interval for each fit, the standard deviation of each 50-year return period wind speed estimate in the background field is calculated. The mean of the standard deviations is considered the representative standard deviation of the background error. Thus, the error term reflects the uncertainty of the statistical model at the 50-year return period and does not consider potential measurement errors and the associated impact on the wind speed predictions.

The 50-year return period wind speed estimates calculated in Chapter 4, which were subsequently filtered based on the A-D test statistic, are selected as point estimates. The standard deviation of the 50-year return period wind speed at each surface station is calculated for consideration by the Bratseth scheme. The background 50-year return period wind field and the mean standard deviation of the estimates are merged with the surface based 50-year return period wind speed point estimates and respective error statistics using the equations defined in Section 6.1. The Bratseth scheme is carried out for two radii of influence, 125 and 250 km, to assess the impact of the Gaussian spatial correlation function.

6.3 Results

50-year return period wind speeds calculated from surface and background datasets in Chapters 4 and 5, are assimilated based on statistical interpolation and subsequently mapped into 2 m/s zones. The Bratseth scheme is considered for two radii of influence, 125 km and 250 km, and the results are shown in Figures 6.2 and 6.3 respectively. The Bratseth analysis based on a radius of influence of 125 km exhibits spatial variations similar to the 50-

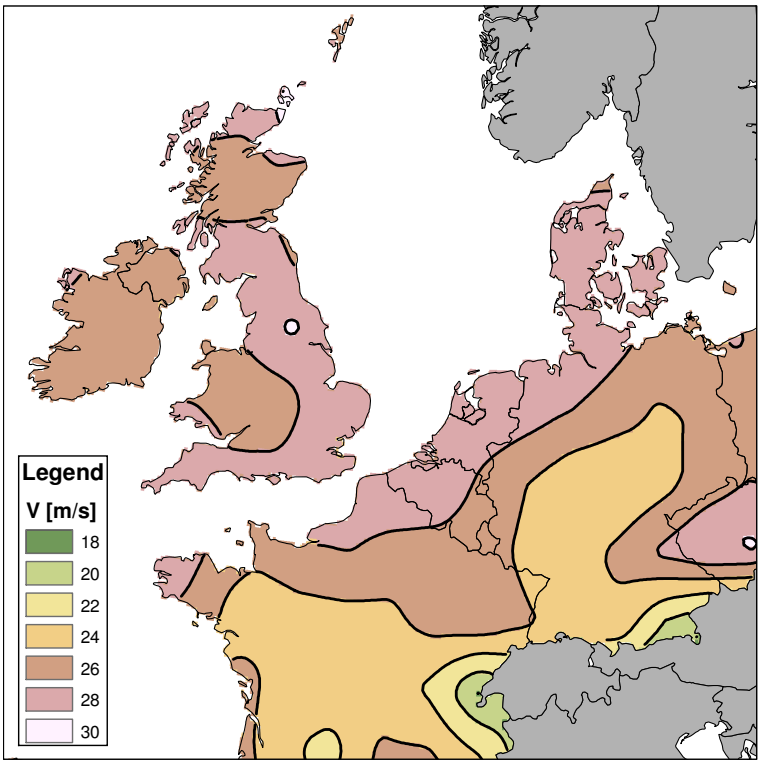


Figure 6.2: 50-year return period wind field - Bratseth scheme (D = 125 km)

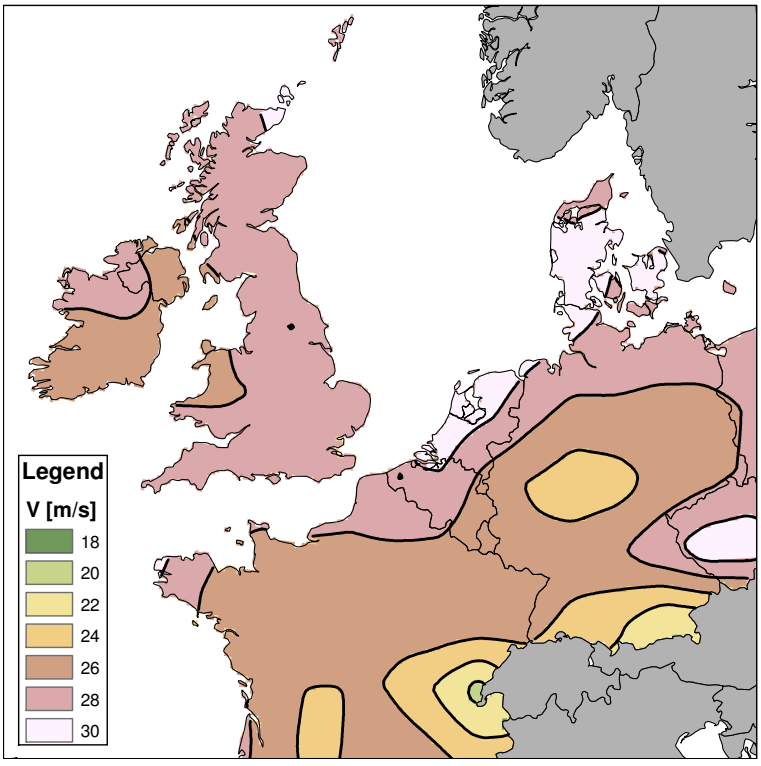


Figure 6.3: 50-year return period wind field - Bratseth scheme (D = 250 km)

year return period wind speed map shown in Figure 4.8.

The 50-year return period wind speed map based on the Bratseth analysis using a radius of influence of 250 km exhibits less variation through the interior of France and Germany. However, a 30 m/s wind speed zone is present along the coast of the North Sea through the Netherlands, Germany and Denmark which is consistent with the Dutch and German NAs. Further inspection of the 50-year return period wind speeds along the northwestern coast of Germany for the 125 km radius of influence case, reveals the values are very close to 28 m/s. The design wind speed zone provided by Kasperski (2002) and the German NA along the coast is 30 m/s and the remaining zones are spaced at 2.5 m/s increments, shown in Figure 2.2(e). Thus, if similar zone intervals are adopted, the coastal region would similarly be considered a 30 m/s wind speed zone since the wind speeds are greater than 27.5 m/s.

The Bratseth scheme is favourable since both datasets, background and surface, are derived independently, allowing their goodness-of-fit to be assessed prior to inclusion in the interpolation scheme. In addition, the spatial influence of the observations, and uncertainty of the estimates can be controlled. Given the spatial resolution of the surface stations considered in the current work, a 125 km radius of influence is more appropriate than a 250 km radius of influence. The former exhibits a smoothed field similar to the results of Chapter 4, while maintaining suitable spatial variability.

Chapter 7

Conclusions

7.1 Overview

Recent years have seen researchers explore various improved methods of calculating 50-year return period wind speeds. Previous studies attempt to calculate 50-year return period wind speeds using various methodologies of ranging complexity which often lack consistency between studies. The impact of inconsistent methodology is apparent in national design codes, as discontinuities exist between specified wind speeds across borders for neighbouring countries in Europe. The current work mitigated these differences through use of a consistent methodology and identified ideal techniques for calculating synoptic 50-year return period wind speeds.

A preliminary study tested the hypothesis that a consistent methodology, regardless of complexity, will reduce observed differences between 50-year return period wind speeds along national borders. At each site, simple corrections for the height of the anemometer and surface roughness were applied, and an extreme value analysis of annual maxima using the

EVD type I distribution was carried out. Several regions exhibiting differences between 50-year return period wind speeds across national borders were explored and the impact of a consistent methodology was considered.

Based on the results of the preliminary study, the focus of the thesis was directed at establishing the a methodology to calculate synoptic 50-year return period wind speeds. The stages considered for surface based wind speeds included quality control, exposure correction, disjunct sampling and extreme value analysis. At each stage, consideration was given to the traditional models used in the wind engineering community. Where available, existing models were compared to proposed empirical models derived from the wind speed data obtained from the ISD, or the application of recent scholarly research available from the fields of meteorology and statistics. The overall methodology was applied to 394 surface observation stations in Europe to calculate a map of synoptic 50-year return period wind speeds.

A background 50-year return period wind field was examined through consideration of wind fields calculated from MSLP and upper-level wind fields available from the ERA-Interim. Several methods for calculating wind fields from MSLP fields were explored and found to perform poorly for two wind storms affecting northern Europe. Upper-level wind fields were then determined by interpolation between wind speeds at the 875 and 925 mb pressure levels. The interpolated wind fields were adjusted to 10 m height in open-country terrain and corrected for disjunct sampling prior to carrying out the extreme value analysis. The resulting background 50-year return period wind field was examined and discrepancies between the results and the surface observations were discussed. The possibility of utilising the 50-year return period wind field derived from upper-air measurements to supplement the surface based 50-year return period wind speeds using data assimilation techniques was considered for the first time in the field of wind engineering. The use of statistical interpolation was selected to assimilate the surface based 50-year return period wind speeds

with the background design wind field.

The various 50-year return period wind speed maps produced at different stages throughout the current work were compared with those from the literature and NAs for a number of countries in Europe. The results include a consistent 50-year return period wind speed map of northern Europe and identification of methodologies for carrying out an analysis of extreme synoptic wind speeds for surface and upper-level wind speed estimates.

7.2 Conclusions

7.2.1 Standardisation and Homogenisation of Wind Speed Data

- Global and localised quality control is an essential component in pre-processing surface wind speed data. Identification of periods containing data conversion errors in the ISD requires the ability to detect short-term departures or outliers from homogeneous subsets of annual mean winds. The ACMANT algorithm provided a method to detect shifts based on monthly mean wind speeds and was used to identify years potentially subject to data conversion problems.
- Localised quality control measures ensured the validity of individual wind speed observations. When considering annual maxima for modelling extremes, a small number of erroneous wind speeds can greatly affect return period estimates. The excessive wind speed variability checks described by DeGaetano (1997) provided a framework for identifying such observations and the method was extended to consider additional sampling intervals. A mean of 9 observations were rejected per station with a maximum of 52 (Sniezka, Poland).
- Wind speed measurements identified as exhibiting excessive variability were compared to entries of thunderstorms and wind speed indicators contained within the

remarks section of the ISD. For the stations considered, additional non-standard formats not outlined in NOAA (2005) were identified manually to extract observer remarks for thunderstorms and wind speed.

- Three exposure correction models were examined to correct wind speeds for upstream exposure by direction. For a subset of KNMI measurement stations, correction factors were derived by direction using effective roughness lengths calculated from measurements of the gust factor using Beljaars' gustiness model and the three exposure correction models. The correction factors were compared to those calculated by direction using the CORINE LULC database and the three correction models. Roughness lengths were assigned to each of the LULC classes and a sampling tool was created to sample the LULC database by both distance and direction. Correction factors derived from the CORINE LULC database provided reasonable estimates of gustiness-derived correction factors, when knowledge of the anemometer and chart recorder characteristics was available.
- The two TL exposure correction models performed better than the IBL exposure correction model. Relative error estimates were 6 to 7 percent for the former and 10 percent for the latter. The Hydra TL model performed slightly better than the Gryning TL model, however, the Hydra TL model had a smaller range of correction factors. The Gryning TL model was selected because it provided a range of corrections similar to the Deaves and Harris IBL model, while exhibiting relative error statistics similar to the Hydra TL model.
- The Deaves and Harris IBL model over-predicted the correction for smooth to rough exposure which translates to an underestimation of the true wind speed. TL models were found to converge to local equilibrium at 10 m height over much shorter fetches, approximately 10 km for open-water to open-country exposure, as opposed to the Deaves and Harris IBL model which requires greater than 100 km.
- Exposure correction was identified as being subject to many sources of uncertainty

including the assignment of a single roughness length to LULC classes, spatial resolution of the LULC database, and differences between the type of exposure correction model selected, TL or IBL.

- An empirical model for correcting observed wind speeds recorded from sampling intervals of 2-, 3-, 4-, and 6-hours and daytime observations to equivalent hourly measurements was formed. Corrections for disjunct sampling were found to approach unity with increasing wind speed, regardless of sampling interval. The convergence to unity across sampling intervals indicated the strong winds were sustained for prolonged periods. Thus, wind speeds were nearly as great during the 2 to 3 hours adjacent the hour exhibiting the maximum wind speed. These events were related to significant depressions affecting Europe, including the Great Storm of '87 (October, 1987), the Burns' Day Storm (January, 1990) and Anatol (December, 1999).

7.2.2 Extreme Value Analysis

- The use of OBRE are recommended to estimate the parameters of the EVD type I to observed data. The estimators provided robust fits to observed data and a framework to detect potential outliers. The OBRE fitting algorithm was applied to storm and annual maxima data for all surface stations. Identification of independent storms was carried out in a similar method to that described by Cook (1982). The Beaufort wind force scale was used as a guideline to select suitable thresholds for defining storms and downweighting maxima.
- Improved fits to storm maxima were found when wind speeds in the lowest Beaufort wind force number were downweighted by the OBRE algorithm to a negligible weight of 0.01. Such modifications are analytically intractable when considering MLE and PWM. When datasets of sufficient length were available, statistical fits to storm and annual maxima provided similar estimates of design wind speeds. Where

differences existed, fits to annual maxima were found to be consistently better than fits to storm maxima.

- Poor statistical fits to storm maxima resulted from attempting to use a predefined criteria to select storm and downweighting thresholds for a significant number of stations in a large spatial region.
- Annual maxima tended to be more robust to a single outlier when the sample size was greater than 30. Similar results were found for storm maxima due to the inherently larger size of the sample sets. The OBRE outlier algorithm was shown to identify observations which were not extracted by the localised quality control measure due to insufficient neighbouring observations for proper evaluation. Although the impact of a single outlier was found to be mitigated for large sample sets, the outlier algorithm provided a means to identify and further assess such observations.
- Surface based 50-year return period wind speeds were converted to gridded estimates using a simple inverse distance weighted technique. The resulting gridded estimates were grouped into 2 m/s zones. Filtering of locations by evaluation of the A-D test statistic led to to the elimination of several locations with poor fits of the selected model to the observed data. The 50-year return period wind speeds calculated in the current work showed good agreement with wind speeds for several nations, while other regions exhibited similar spatial variation but were greater in magnitude than existing reported values. 50-year return period wind speeds along the English Channel were found to be consistent with the range of wind speeds inferred from the preliminary investigation. The impact of topography was apparent from the reductions of design wind speed through the Rhine valley. Topographic effects, in conjunction with poor spatial resolution of data, resulted in increased spatial variability of 50-year return period wind speeds in Spain.
- Differences of up to 2 m/s were found to occur as a result of using various types of contouring and zoning intervals for neighbouring nations.

- 50-year return period wind speeds in the current work exhibit similar spatial variation as previously reported 50-year return period wind maps for the UK. Miller *et al.* (2001) corrected wind speeds for the UK using the Deaves and Harris IBL model. Due to the differences between the adjustment of IBL and TL models to upstream changes of roughness, the difference in magnitude for values over the UK could be reduced by application of the same exposure correction model. The differences exemplify the importance of a consistent methodology to calculate 50-year return period wind speeds.
- Maxima were resolved into components to form time histories for each 30-degree azimuth. 50-year return period wind speeds were calculated for each sector and normalised by the largest 50-year return period wind speed estimate to calculate a set of direction factors. The factors were averaged over all stations within a country or region, and subsequently combined into six climates based on similar directional effects related to typical storm tracks. The six regions exhibiting similar directional characteristics are:
  - I: Portugal, Spain
  - II: France (Atlantic), France (Interior)
  - III: Ireland, United Kingdom, France (English Channel)
  - IV: Germany (South), Belgium, Netherlands
  - V: Germany (North), Denmark
  - VI: Czech Republic, Poland

7.2.3 Background 50-year Return Period Wind Field

- The use of re-analysis data from the ERA-Interim was explored to calculate a background 50-year return period wind field which is free from the influence of the surface. Several approximations to the Navier-Stokes equation were explored using

MSLP fields. The geostrophic approximation allows wind fields to be calculated from MSLP fields based on the balance between the pressure gradient and Coriolis forces. The geostrophic wind fields represent frictionless flow at the sea-level which are assumed to describe the upper-level wind fields sufficiently far from the influence of the surface. For regions affected by strong, tightly-wound ETCs the geostrophic approximation was shown to overestimate upper-level wind speeds.

- Quasi- and semi-geostrophic approximations which improve upon the geostrophic approximation through consideration of advection, broke down in regions near the storm centre and in regions of large topography. MSLP fields were found to exhibit 'pocketing' in regions of large topography, indicating shortcomings of the underlying equations for reducing surface pressures to MSLP values in mountainous areas. Smoothing of the pressure fields improved the results of the semi-geostrophic approximation, however, smoothing of MSLP fields was identified as unfavourable since it would lead to decreased estimates of wind speed near the storm centre. Neither approximation was found to provide adequate wind speed estimates for the full spatial extents of MSLP fields from the Burns' Day Storm (January, 1990) and Anatol (December, 1999).
- Due to the inability to satisfactorily calculate wind fields at upper-levels from pressure data, upper-level wind fields were alternatively derived by interpolating wind speed and direction fields from the ERA-Interim at the 875 and 925 mb pressure levels to a height of 1000 m. Wind fields calculated in this manner showed good agreement with radiosonde estimates for the Burns' Day storm. The wind fields were not as resolved as those calculated from pressure measurements.
- Wind fields at 1000 m were used to estimate the wind speed at 10 m height in open-country exposure by application of the Gryning ABL model. A correction was applied for disjunct sampling, and the EVD type I was fit to annual maxima using OBRE. 50-year return period wind speeds were found to be less than those calcu-

lated from surface observations in Chapter 4. The results were found to be similar to other 50-year return period wind fields calculated from pressure and re-analysis data by Miller (2003) and Larsén and Mann (2009) respectively.

- The overall reduction of 50-year return period wind speeds was a result of the spatial and temporal smoothing of point observations by the re-analysis model. The smoothing inherent to re-analysis data resulted in significant reductions of the location and scale parameters. The distribution of the location parameter exhibited two peaks. The additional peak was found to be related to location parameters of grid points over the North Atlantic. The results indicate that, despite utilising upper-level wind data which should be relatively free from surface effects, the upper-level wind fields of the ERA-Interim are influenced by the surface.

7.2.4 Data Assimilation

- Data assimilation techniques allow data from multiple sources to be jointly considered. The possibility of utilising the background 50-year return period wind field derived from upper-air measurements to supplement the surface based 50-year return period wind speeds was considered for the first time. The selected method was the Bratseth scheme for statistical interpolation.
- The Bratseth scheme with a radius of influence of 125 km was found to provide a compromise between capturing the spatial variation of 50-year return period wind speeds and smoothing of the wind field. This method produced the overall 50-year return period wind speed map.

7.3 Future work

There are four recommendations for extension of the current work. First, the localised quality control method relies on the completeness of indicators representing the weather at the time of observation, at a site, to identify thunderstorms. It was discussed that a reported thunderstorm may only involve hearing thunder or observing lightning, not necessarily observing a 'thunderstorm wind'. Thus, a method is required to better identify thunderstorm events from time histories of related parameters such as wind speed, wind direction, and temperature. Second, further validation of the IBL and TL exposure correction models is required to properly assess which models are appropriately capturing the impact of fetch following a change of roughness. This could have a significant impact on all corrected wind speeds. Third, the relationship between the 10-minute mean recorded in the 10-minute period prior to the hour and a continuous 10-minute mean is required to fully account for disjunct sampling by the proposed empirical model. Lastly, additional examination of the uncertainty associated with the final 50-year return period wind speeds is required. The largest sources of uncertainty arise from measurement error, exposure correction and analysis of the extremes, however, only the uncertainty associated with the extreme value analysis is typically considered. Although the predicted 50-year return period wind speeds would likely remain unchanged, increased uncertainty will influence data assimilation methods such as the Bratseth scheme.

References

An, Y. and Pandey, M. (2007) The r largest order statistics model for extreme wind speed estimation. *J. Wind Eng. Ind. Aerod.*, **95(3)**, 165–182.

Anderson, T. and Darling, D. (1952) Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes. *Ann. Math. Stat.*, **23(2)**, 193–212.

Anderson, T. and Darling, D. (1954) A test of goodness of fit. *J. Am. Stat. Assoc.*, **49(268)**, 765–769.

Barthelmie, R., Palutikof, J. and Davies, T. (1993) Estimation of sector roughness lengths and the effect on prediction of the vertical wind speed profile. *Bound.-Lay. Meteorol.*, **66**, 19–47.

Beljaars, A. (1987) The influence of sampling and filtering on measured wind gusts. *J. Atmos. Ocean. Tech.*, **4**, 613–626.

Berrisford, P., Dee, D., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S. and Uppala, S. (2009) The ERA-Interim archive. Technical report, European Centre for Medium-Range Weather Forecases.

Blackadar, A. and Tennekes, H. (1968) On the extension of the wind profile over homogeneous terrain beyond the surface boundary layer. *J. Atmos. Sci.*, **25**, 1015–1020.

Bossard, M., Feranec, J. and Otahel, J. (2000) The revised and supplemented CORINE land cover nomenclature. Technical Report 38, European Environmental Agency.

Boucher, K. (2005) Climate of Europe. In *Encyclopedia of World climatology*. Springer Dordrecht.

Bratseth, A. (1986) Statistical interpolation by means of successive corrections. *Tellus*, **38A**, 439–447.

Burton, M. and Allsop, A. (2009a) Irish national annex to the wind Eurocode (EN1991-1-4): Derivation of the wind map. Technical Report 119646, Ove Arup & Partners Ltd.

Burton, M. and Allsop, A. (2009b) Predicting design wind speeds from anemometer records: Some interesting findings. In 11^{th} *Americas Conference on Wind Engineering*. San Juan, Puerto Rico.

Cappelan, J. and Jorgensen, B. (1999) Observed wind speed and direction in Denmark - with climatological standard normals, 1961-90. Technical Report DMI TR-99-13, Danish Meteorological Institute.

Caussinus, H. and Lyazrhi, F. (1997) Choosing a linear model with a random number of change-points and outliers. *Ann. Inst. Statist. Math.*, **49(4)**, 761–775.

Caussinus, H. and Mestre, O. (2004) Detection and correction of artificial shifts in climate series. *Appl. Stat.-J. Roy. St. C*, **53(3)**, 405–425.

Coles, S. and Walshaw, D. (1994) Directional modelling of extreme wind speeds. *J. Roy. Stat. Soc. C-App.*, **43(1)**, 139–157.

Cook, N. (1982) Towards better estimation of extreme winds. *J. Wind Eng. Ind. Aerod.*, **9**, 295–323.

Cook, N. (1983) Note on directional and seasonal assessment of extreme winds for design. *J. Wind Eng. Ind. Aerod.*, **12(3)**, 365–372.

Cook, N. (1985) *The designer's guide to wind loading of building structures: Part I.* Butterworths, London, UK.

Cook, N. (1997) The Deaves and Harris ABL model applied to heterogeneous terrain. *J. Wind Eng. Ind. Aerod.*, **66**, 197–214.

Cook, N. (2011) Comments on 'Plotting positions in extreme value analysis'. *J. Appl. Meteorol. Clim.*, **50(1)**, 255–266.

Cook, N. and Harris, R. (2001) Discussion on application of the generalized pareto distribution to extreme value analysis in wind engineering by J.D. Holmes, W.W. Moriarty. *J. Wind Eng. Ind. Aerod.*, **89(2)**, 215–224.

Cook, N. and Prior, M. (1987) Extreme wind climate of the United Kindom. *J. Wind Eng. Ind. Aerod.*, **26**, 371–389.

Davenport, A. (1963) The relationship of wind structure to wind loading. In *Proceedings of the National Physical Laboratory Symposium 16*. Wind Effects on Structures and Buildings, Teddington, UK.

Davenport, A. (1983) The relationship of reliability to wind loading. *J. Wind Eng. Ind. Aerod.*, **13(1-3)**, 3–27.

Davenport, A. (1999) The missing links. In *Proceedings of the* 10^{th} *International Conference on Wind Engineering*. International Associations for Wind Engineering, Rotterdam.

Deaves, D. (1981) Computations of wind flow over changes in surface roughness. *J. Wind Eng. Ind. Aerod.*, **7**, 65–91.

Deaves, D. and Harris, R. (1978) A mathematical model of the structure of strong winds. Technical Report CIRIA76, Construction Industry Research and Information Association.

DeGaetano, A. (1997) A quality control procedure for hourly wind data. *J. Atmos. Ocean. Tech.*, **14**, 308–317.

Dodd, E. (1923) The greatest and the least variate under general laws of error. *T. Am Math. Soc.*, **25(4)**, 525–539.

Domonkos, P. (2011) Adapted Caussinus-Mestre algorithm for networks of temperature series ACMANT. *Int. J. Geosci.*, **2**, 239–309.

Dupuis, D. (1998) Exceedances over high thresholds: A guide to threshold selection. *Extremes*, **1(3)**, 251–261.

Dupuis, D. and Field, C. (1998a) A comparison of confidence intervals for generalized extreme-value distributions. *J. Stat. Comput. Sim.*, **61(4)**, 341–360.

Dupuis, D. and Field, C. (1998b) Robust estimation of extremes. *Can. J. Stat.*, **26(2)**, 199–215.

Dupuis, D. and Field, C. (2004) Large wind speeds: Modeling and outlier detection. *J. Agr. Biol. Envir. St.*, **9(1)**, 1–17.

Efron, B. and Hinkley, D. (1978) Assessing the accuracy of the maximum likelihood estimator: observed versus expected fisher information. *Biometrika*, **65(3)**, 457–483.

Fawcett, L. and Walshaw, D. (2008) Bayesian inference for clustered extremes. *Extremes*, **11**, 217–233.

Fisher, R. (1912) On an absolute criterion for fitting frequency curves. *Messeng. Math.*, **41**, 155–160.

Fisher, R. (1922) On the mathematical foundations of theoretical statistics. *Philos. Trans. Roy. Soc. London Ser. A*, **222**, 309–368.

Fisher, R. and Tippet, L. (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Cambridge*, **24(2)**, 180–190.

Frank, H. (2001) Extreme winds over Denmark from the NCEP/NCAR reanalysis. Technical Report Risø-R-1238(EN), RisøNational Laboratory.

Gatey, D. and Miller, C. (2007) An investigation into 50-year return period wind speed differences for Europe. *J. Wind Eng. Ind. Aerod.*, **95(9-11)**, 1040–1052.

George, S. (2006) *United Kingdom Windspeed*. Ph.D. thesis, University College London.

Gomes, L. and Vickery, B. (1978) Extreme wind speeds in mixed wind climates. *J. Wind Eng. Ind. Aerod.*, **2(4)**, 331–344.

Graybeal, D., DeGaetano, A. and Eggleston, K. (2004) Complex quality assurance of historical hourly surface airways meteorological data. *J. Atmos. Ocean. Tech.*, **21**, 1156–1169.

Greenwood, J., Landwehr, J., Matalas, N. and Wallis, J. (1979) Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resour. Res.*, **15(5)**, 1049–1054.

Gringorten, I. (1963) A plotting rule for extreme value probability paper. *J. Geophys. Res.*, **68**, 813–814.

Gryning, S., Batchvarova, E., Brümmer, B., Jørgensen, H. and Larsen, S. (2007) On the extension of the wind profile over homogeneous terrain beyond the surface boundary layer. *Bound.-Lay. Meteorol.*, **124**, 251–268.

Gumbel, E. (1958) *Statistics of Extremes*. Columbia University Press, New York, USA.

de Haan, L. (2007) Comments on 'Plotting positions in extreme value analysis'. *J. Appl. Meteorol. Clim.*, **46(3)**, 396.

Hampel, F. (1968) *Contributions to the Theory of Robust Estimation*. Ph.D. thesis, University of California.

Hampel, F. (1974) The influence curve and its role in robust estimation. *J. Am. Stat. Assoc.*, **69(346)**, 383–393.

Hampel, F., Ronchetti, E., Rousseeuw, P. and Stahel, W. (1986) *Robust Statistics: The Approach Based On Influence Functions*. John Wiley, New York, NY.

Harris, R. (1999) Improvements to the 'Method of independent storms'. *J. Wind Eng. Ind. Aerod.*, **80(1-2)**, 1–30.

Harris, R. (2004) Extreme value analysis of epoch maxima - convergence, and choice of asymptote. *J. Wind Eng. Ind. Aerod.*, **92(11)**, 897–918.

Harris, R. (2005) Generalised Pareto methods for wind extremes. Useful tool or mathematical mirage. *J. Wind Eng. Ind. Aerod.*, **93(5)**, 341–360.

Harris, R. (2009) XIMIS, a penultimate extreme value method suitable for all types of wind climate. *J. Wind Eng. Ind. Aerod.*, **97(5-6)**, 271–286.

Højstrup, J. (1982) Velocity spectra in the unstable planetary boundary layer. *J. Atmos. Sci.*, **39**, 2239–2248.

Holmes, J., Kasperski, M., Miller, C., Żurański, J. and Choi, E. (2005) Extreme wind prediction and zoning. *Wind Struct.*, **8(4)**, 269–281.

Holmes, J. and Moriarty, W. (1999) Application of the generalized Pareto distribution to extreme value analysis in wind engineering. *J. Wind Eng. Ind. Aerod.*, **83(1-3)**, 1–10.

Holton, J. (2004) *An Introduction to Dynamic Meteorology*. Elsevier Academic Press, London, UK.

Hosking, J. (1990) L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. Roy. Stat. Soc. B Met.*, **52(1)**, 105–124.

Hosking, J., Wallis, J. and Wood, E. (1985) Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, **27(3)**, 251–261.

Hoskins, B. (1975) The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.*, **32(2)**, 233–242.

Huber, P. (1964) Robust estimation of a location parameter. *Ann. Math. Stat.*, **35(1)**, 73–101.

Jenkinson, A. (1969) Statistics of extremes in, Estimation of maximum floods. Technical Report 98, World Meteorological Office.

Kaimal, J. (1978) Horizontal velocity spectra in an unstable surface layer. *J. Atmos. Sci.*, **35**, 18–24.

Kasperski, M. (2002) A new wind zone map of Germany. *J. Wind Eng. Ind. Aerod.*, **90**, 1271–1287.

Kasperski, M. (2007) Design wind loads for a low-rise building taking into account directional effects. *J. Wind Eng. Ind. Aerod.*, **95**, 1125–1144.

Kasperski, M. (2009) Specification of the design wind load - a critical review of code concepts. *J. Wind Eng. Ind. Aerod.*, **97(7-8)**, 335–357.

Kasperski, M. and Geurts, C. (2005) Reliability and code level. *Wind Struct.*, **8(4)**, 295–307.

Kristensen, L. and Jensen, G. (1999) Geostrophic winds in Denmark: a preliminary study. Technical Report Risø-R-1145(EN), RisøNational Laboratory.

Kristensen, L., Rathmann, O. and Hansen, S. (2000) Extreme winds in Denmark. *J. Wind Eng. Ind. Aerod.*, **87(2-3)**, 147–166.

Landwehr, J., Matalas, N. and Wallis, J. (1979) Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resour. Res.*, **15(5)**, 1055–1064.

Larsén, X. and Mann, J. (2006) The effects of disjunct sampling and averaging time on maximum mean wind speeds. *J. Wind Eng. Ind. Aerod.*, **94(8)**, 581–602.

Larsén, X. and Mann, J. (2009) Extreme winds from the NCEP/NCAR reanalysis data. *Wind Energy*, **12(6)**, 556–573.

Letchford, C., Gardner, A., Howard, R. and Schroeder, J. (2001) A comparison of wind prediction models for transitional flow regimes using full-scale hurricane data. *J. Wind Eng. Ind. Aerod.*, **89**, 925–945.

Lettau, H. (1950) A re-examination of the Leipzig wind profile considering some relations between wind and turbulence in the frictional layer. *Tellus*, **2**, 125–129.

Lott, N. (2004) The quality control of the integrated surface hourly database. In *84th American Meteorological Society Annual Meeting*. American Meteorological Society, Boston, MA.

Makkonen, L. (2006) Plotting positions in extreme value analysis. *J. Appl. Meteorol. Clim.*, **45(2)**, 334–340.

Makkonen, L. (2008) Bringing closure to the plotting position controversy. *Commun. Stat. A-Theor.*, **37(3)**, 460–467.

Masters, F., Vickery, P., Bacon, P. and Rappaport, E. (2010) Toward objective, standardized intensity estimates from surface wind speed observations. *B. Am. Meteorol. Soc.*, **91(12)**, 1665–1681.

Melbourne, W. (1984) Designing for directionality. In 1^{st} *Workshop on Wind Engineering and Industrial Aerodynamics*. Highett, Australia.

Miller, C. (2003) A once in 50-year wind speed map for Europe derived from mean sea level pressure measurements. *J. Wind Eng. Ind. Aerod.*, **91**, 1813–1826.

Miller, C., Cook, N. and Barnard, R. (2001) Towards a revised base wind speed map for the United Kingdom. *Wind Struct.*, **4(3)**, 197–212.

von Mises, R. (1936) La distribution de la plus grande de n valeurs. *Revue Math. de l'Union Interbalkanique*, **1**.

Moriarty, W. and Templeton, J. (1983) On the estimation of extreme wind gusts by direction sector. *J. Wind Eng. Ind. Aerod.*, **13**, 127–138.

NCDC (2010) Federal climate complex data documentation for Integrated Surface Data. Technical Report TD-3505, National Climatic Data Center.

NOAA (2005) Federal meteorological handbook No. 1: Surface weather observations and reports. Technical Report FCM-H1-2005, National Oceanic and Atmospheric Administration.

Palutikof, J., Brabson, B., Lister, D. and Adcock, S. (1999) A review of methods to calculate extreme wind speeds. *Meteorol. Appl.*, **6**, 119–132.

Peña, A., Gryning, S., , Mann, J. and Hasager, C. (2010) Length scales of the neutral wind profile over homogeneous terrain. *J. Appl. Meteorol. Clim.*, **49**, 792–806.

Peterka, J. and Shahid, S. (1998) Design gust wind speeds in the United States. *J. Struct. Eng.-ASCE*, **124**, 207–214.

Prescott, P. and Walden, A. (1983) Maximum likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution from censored samples. *J. Stat. Comput. Sim.*, **16**, 241–250.

Rossby, C. and Montgomery, R. (1935) The layers of frictional influence in wind and ocean currents. *Pap. Phys. Oceanogr. Meteorl.*, **3(3)**.

Sacré, C. (2002) Extreme wind speed in France: the '99 storms and their consequences. *J. Wind Eng. Ind. Aerod.*, **90**, 1163–1171.

Sacré, C., Moisselin, J., Sabre, M., Flori, J. and Dubuisson, B. (2007) A new statistical approach to extreme wind speeds in France. *J. Wind Eng. Ind. Aerod.*, **95(9-11)**, 1415–1423.

Shellard, H. (1963) The estimation of design wind speeds. In *Proceedings of the National Physical Laboratory Symposium 16*. Wind Effects on Structures and Buildings, Teddington, UK.

Shepard, D. (1968) A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 ACM National Conference*. Association for Computing Machinery, New York, USA.

Simiu, E. and Scanlan, R. (1996) *Wind effects on structures: Fundamentals and applications to design* , 3^{rd} *edition*. John Wiley and Sons, Inc., New York, US.

Simiu, E., Wilcox, R., Sadek, F. and Filliben, J. (2003) Wind speeds in ASCE 7 standard peak-gust map: Assessment. *J. Struct. Eng.-ASCE*, **129**, 427–439.

Taylor, P. (1987) Comments and further analysis on effective roughness lengths for use in numerical three-dimensional models. *Bound.-Lay. Meteorol.*, **39**, 403–418.

Traup, S. and Kruse, B. (1996) *Winddaten für Windenergienutzer*. Deutscher Wetterdienst, Offenbach am Main.

Uppala, S., Kallberg, P., Simmons, A., Andrae, U., Costa, V. D., Fiorino, M. and Gibson, J. (2005) The ERA-40 re-analysis. *Q. J. Roy. Meteor. Soc.*, **131(612)**, 2961–3012.

Vega, R. (2008) *Wind directionality: A reliability-based approach*. Ph.D. thesis, Texas Tech University.

Venema, V., Mestre, O., Aguilar, E., Auer, I., Guijarro, J., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P., Viarre, J., Muller-Westermeier, G., Lakatos, M., Williams, C. N., Menne, M., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andresen, L., Acquaotta, F., Fratianni, S., Cheval, S., Klancar, M., Brunetti, M., Gruber, C., Prohom-Duran, M., Likso, T., Esteban, P. and Brandsma, T. (2011) Benchmarking monthly homogenization algorithms. *Clim. Past Discuss.*, **7**, 2655–2718.

Verkaik, J. (2000) Evaluation of two gustiness models for exposure correction calculations. *J. Appl. Meteorol.*, **39**, 1613–1626.

Verkaik, J. (2001) Documentatie windmetingen in Nederland. Technical Report KNMI-2001, Koninklijk Nederlands Meteorologisch Instituut.

Verkaik, J. (2003) A method for the geographical interpolation of wind speed over heterogeneous terrain. Technical Report KNMI-HYDRA 11-12, Koninklijk Nederlands Meteorologisch Instituut.

Vickery, P., Wadhera, D., Galsworthy, J., Peterka, J., Irwin, P. and Griffis, L. (2010) Ultimate wind load design gust wind speeds in the United States for use in ASCE-7. *J. Struct. Eng.-ASCE*, **136**, 613–625.

Vincenty, T. (1975) Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Surv. Rev.*, **176**, 88–93.

Watson, G., Halliday, J., Palutikof, J., Holt, T., Barthelmie, R., Coelingh, J., van Zuylen, E. and Cleijne, J. (2001) Predicting offshore wind energy resources (POWER). Technical report, POWER Consortium.

Wieringa, J. (1973) Gust factors over open water and built-up country. *Bound.-Lay. Meteorol.*, **3**, 424–441.

Wieringa, J. (1976) An objective exposure correction method for average wind speeds measured at a sheltered location. *Q. J. Roy. Meteor. Soc.*, **102**, 241–253.

Wieringa, J. (1986) Roughness-dependent geographical interpolation of surface wind speed averages. *Q. J. Roy. Meteor. Soc.*, **112(473)**, 867–889.

Wieringa, J. (1993) Representative roughness parameters for homogeneous terrain. *Bound.-Lay. Meteorol.*, **63**, 323–363.

WMO (2008) Guide to meteorological instruments and methods of observation. Technical Report WMO-No. 8, World Meteorological Organization.

Yip, T. and Auld, H. (1993) Updating the 1995 National building code of Canada wind pressures. In *Proceedings of the Electricity '93 Engineering and Operating Division Conference*. Canadian Electrical Assocation, Montréal, Canada.

Yip, T., Auld, H. and Dnes, W. (1995) Recommendations for updating the 1995 National building code of Canada wind pressures. In *Proceedings of the 9th International Conference on Wind Engineering*. International Associations for Wind Engineering, New Delhi, India.

Zilitinkevich, S. (1989) Velocity profiles, the resistance law and the dissipation rate of mean flow kinetic energy in a neutrally and stably stratified planetary boundary layer. *Bound.-Lay. Meteorol.*, **46**, 367–387.

Zilitinkevich, S. and Esau, I. (2002) On integral measures of the neutral barotropic planetary boundary layer. *Bound.-Lay. Meteorol.*, **104**, 371–379.

Żurański, J. and Jaśpińska, B. (1996) Directional analysis of extreme wind speeds in Poland. *J. Wind Eng. Ind. Aerod.*, **65(1-3)**, 13–20.

Appendix A

Station Listing

A.1 Selected Stations

WMO	Station Name	Country	Latitude	Longitude
030030	Sumburgh	UK	59.87871	-1.29970
030170	Kirkwall	UK	58.95252	-2.90528
030220	Benbecula Island	UK	57.47576	-7.36661
030260	Stornoway Airport	UK	58.21454	-6.32726
030340	Aultbea No 2	UK	57.85958	-5.63168
030470	Tulloch Bridge	UK	56.86716	-4.70763
030660	Kinloss	UK	57.64499	-3.56386
030750	Wick Airport	UK	58.45566	-3.09241
030910	Dyce/Aberdeen	UK	57.20807	-2.19780
031000	Tiree	UK	56.49795	-6.87759
031110	Machrihanish	UK	55.44002	-5.69546
031320	West Freugh	UK	54.85922	-4.93447
031350	Prestwick No 2	UK	55.50184	-4.58302
031530	Dundrennan	UK	54.80349	-4.00764
031580	Charterhall	UK	55.70909	-2.38430
031600	Turnhouse/Edinburgh	UK	55.94941	-3.34748
031710	Leuchars	UK	56.37513	-2.86376
032040	Ronaldsway	UK	54.08503	-4.62721
032080	Point of Ayre	UK	54.41492	-4.36929
032140	Walney Island	UK	54.12512	-3.25851
032260	Warcop Range	UK	54.57226	-2.41374
032300	Redesdale Camp	UK	55.28616	-2.27786
032400	Boulmer	UK	55.42241	-1.60264
032433	Newcastle	UK	55.03356	-1.68472

Continued on Next Page. . .

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
032570	Leeming	UK	54.29706	-1.53364
032660	Linton on ouse	UK	54.04577	-1.24934
032920	Bridlington Mrsc	UK	54.09450	-0.17503
033020	Valley	UK	53.25523	-4.54157
033160	Crosby	UK	53.49697	-3.05742
033180	Blackpool/Squires Gate	UK	53.77668	-3.03865
033210	Hawarden Airport	UK	53.17486	-2.98664
033340	Ringway/Manchester	UK	53.35661	-2.27716
033463	Leeds/Bradford	UK	53.92881	-1.59867
033600	Finningley	UK	53.48315	-1.00764
033735	Humberside	UK	53.63043	-0.28132
033770	Waddington	UK	53.16358	-0.52889
033910	Coningsby	UK	53.09392	-0.17284
034140	Shawbury	UK	52.80134	-2.66519
034185	East Midlands	UK	52.88372	-1.27848
034530	Cottesmore	UK	52.72906	-0.65463
034620	Wittering	UK	52.61261	-0.46767
034820	Marham	UK	52.64689	0.56535
034950	Coltishall	UK	52.75642	1.36331
035020	Aberporth	UK	52.13862	-4.57394
035030	Trawsgoed	UK	52.34486	-3.94724
035070	Sennybridge No 2	UK	52.06396	-3.61405
035290	Pershore	UK	52.14873	-2.04161
035340	Elmdon/Birmingham	UK	52.45323	-1.74318
035440	Church Lawford	UK	52.45505	-1.75494
035580	Bedford Airport	UK	52.35902	-1.33110
035600	Bedford	UK	52.22686	-0.46488
035660	Wyton	UK	52.35397	-0.11541
035773	Mildenhall	UK	52.37282	0.47157
035860	Honington	UK	52.34142	0.79135
035900	Wattisham	UK	52.12244	0.95735
036030	Brawdy	UK	51.88071	-5.12301
036040	Milford Haven	UK	51.70810	-5.05174
036440	Fairford	UK	51.68279	-1.77214
036490	Brize Norton	UK	51.74521	-1.58209
036553	Upper Heyford	UK	51.97448	-1.22620
036580	Benson	UK	51.62111	-1.09789
036830	Stansted	UK	51.88590	0.22242
036930	Shoeburyness	UK	51.55542	0.82807
036960	Walton-on-the-naze	UK	51.85487	1.27963
037070	Chivenor	UK	51.08519	-4.14374

Continued on Next Page. . .

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
037150	Rhoose	UK	51.40002	-3.35623
037243	Bristol/Lulsgate	UK	51.42009	-2.71253
037400	Lyneham	UK	51.50318	-1.99203
037460	Boscombe Down	UK	51.16214	-1.75469
037610	Odiham	UK	51.23947	-0.94493
037630	Bracknell/Beaufort Park	UK	51.39079	-0.78479
037760	Gatwick	UK	51.14492	-0.19796
037970	Manston	UK	51.35372	1.34856
038090	Culdrose	UK	50.08617	-5.25503
038170	St Mawgan	UK	50.44472	-5.00207
038270	Plymouth/Mount Batten	UK	50.35531	-4.12051
038390	Exeter	UK	50.73116	-3.41617
038530	Yeovilton	UK	51.00688	-2.64062
038550	Isle of Portland	UK	50.51404	-2.45621
038620	Bournemouth/Hurn	UK	50.77832	-1.84039
038660	Wight: St Catherines Point	UK	50.57577	-1.29736
038800	Newhaven	UK	50.78276	0.05762
038940	Guernsey Airport	UK	49.43408	2.59520
038950	Jersey Airport	UK	49.20955	-2.19428
039030	St Angelo	UK	54.39555	-7.64508
039170	Aldergrove	UK	54.65519	-6.22908
039230	Glenanne No 2	UK	54.23300	-6.50000
039520	Roches Point	IE	51.79311	-8.25412
039530	Valentia	IE	51.93972	-10.24444
039570	Rosslare	IE	52.24972	-6.33442
039600	Kilkenny	IE	52.66528	-7.26944
039620	Shannon Airport	IE	52.70151	-8.92120
039650	Birr	IE	53.09028	-7.89028
039690	Dublin Airport	IE	53.43404	-6.26196
039700	Claremorris	IE	53.71074	-8.99220
039710	Mullingar	IE	53.53704	-7.36194
039740	Clones	IE	54.18333	-7.23333
039760	Belmullet	IE	54.22778	-10.00694
039800	Malin Head	IE	55.37222	-7.33889
060240	Thisted Lufthavn	DN	57.06691	8.71335
060300	FSN lborg	DN	57.09507	9.85660
060410	Skagen Fyr	DN	57.73657	10.63190
060520	Thyborn	DN	56.70040	8.22171
060600	FSN Karup	DN	56.29632	9.12081
060700	Tirstrup	DN	56.30022	10.63623
060800	Esbjerg Lufthavn	DN	55.52610	8.57337

Continued on Next Page. . .

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
061040	Billund Lufthavn	DN	55.73919	9.15452
061100	FSN Skrydstrup	DN	55.21933	9.28643
061180	Soenderborg	DN	54.96914	9.78688
061190	Kegns Fyr	DN	52.60911	-0.47518
061200	Odense Lufthavn	DN	55.47679	10.33626
061500	Avno	DN	55.08300	11.78300
061700	Roskilde Lufthavn	DN	55.58696	12.13521
061790	Mn Fyr	DN	54.94659	12.53984
061800	Kbenhavns Lufthavn	DN	55.62011	12.66418
062250	Ijmuiden	NL	52.46290	4.55544
062350	De Kooy	NL	52.92875	4.78454
062420	Vlieland	NL	53.24144	4.92025
062500	Terschelling	NL	53.35283	5.18413
062650	Soesterberg	NL	52.12982	5.27444
062700	Leeuwarden	NL	53.22492	5.74658
062750	Deelen	NL	52.05583	5.87366
062800	Groningen	NL	53.12402	6.58586
062900	Twenthe	NL	52.27306	6.89667
063100	Vlissingen	NL	51.44236	3.59620
063300	Hoek Van Holland	NL	51.98760	4.08522
063400	Woendrecht	NL	51.68496	4.44938
063440	Rotterdam	NL	51.69218	4.45321
063500	Gilze Rijen	NL	51.56733	4.93659
063700	Eindhoven	NL	51.44555	5.41353
063750	Volkel	NL	51.65682	5.70561
063800	Beek/Maastrict	NL	50.91900	5.77519
064000	Koksijde	BX	51.08803	2.65235
064070	Middlekerke/Oostende	BX	51.20046	2.88723
064320	Chievres	BX	50.57396	3.83281
064510	Melsbroek/Bruxelles	BX	50.89636	4.52697
064520	Brasschaat	BX	51.33733	4.50453
064560	Florennes	BX	50.23661	4.65340
064580	Beauvechain	BX	50.75684	4.76872
064700	St. Truiden	BX	50.79014	5.19087
064760	St-Hubert	BX	50.03385	5.40562
064780	Bierset	BX	50.64594	5.45567
064790	Kleine Brogel	BX	51.17008	5.46595
064900	La Sauveniere/Spa	BX	50.48173	5.91084
065900	Luxembourg	BX	49.62105	6.20135
070030	Le Touquet	FR	50.51459	1.62275
070050	Abbeville	FR	50.14333	1.82574

Continued on Next Page...

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
070100	Dunkerque	FR	51.05000	2.33333
070150	Lille Lesqui	FR	50.56700	3.10000
070240	Maupertus	FR	49.64749	-1.47406
070270	Caen Carpiquet	FR	49.18004	-0.45626
070370	Rouen	FR	49.38930	1.17944
070380	Evreux	FR	49.02870	1.21990
070610	Saint-Quentin	FR	49.81807	3.20568
070700	Reims	FR	49.31176	4.04824
071000	Ouessant	FR	48.46550	-5.05738
071100	Brest	FR	48.45160	-4.40664
071210	Brehat Island	FR	48.85000	-3.00000
071250	Dinard	FR	48.58611	-2.08239
071300	Rennes	FR	48.06913	-1.73335
071530	Melun	FR	48.60921	2.67834
071570	Paris	FR	49.02217	2.51737
071680	Troyes Barberey	FR	48.32708	4.01404
071690	Saint-Dizier	FR	48.63654	4.90490
071790	Toul Rosiere	FR	48.78131	5.98366
071810	Nancy-Ochey	FR	48.58310	5.95500
071900	Strasbourg-Entzheim	FR	48.54251	7.63801
071970	Meyenheim-Colmar	FR	47.92110	7.40079
072010	Quimper	FR	47.97370	-4.17199
072050	Lorient Lan Bihoue	FR	47.76060	-3.44000
072070	Belle Ie Le Talut	FR	47.29475	-3.21864
072220	Nantes	FR	47.15006	-1.60883
072300	Angers	FR	47.49843	-0.57491
072350	Le Mans	FR	47.94909	0.19893
072400	Tours	FR	47.44035	0.72922
072490	Orleans	FR	47.99079	1.77790
072570	Avord	FR	47.05192	2.63096
072650	Auxerre	FR	47.80000	3.55000
072800	Dijon	FR	47.26890	5.09000
072830	Langres	FR	47.85000	5.33333
072920	Luxeuil	FR	47.78820	6.34994
072990	Bale-Mulhouse	FR	47.59207	7.52157
073350	Poitiers	FR	46.58475	0.30903
073540	Chateauroux/Deols	FR	46.86220	1.73070
074120	Cognac	FR	45.65830	-0.31750
074340	Limoges	FR	45.86276	1.18020
074600	Clermont-Ferrand	FR	45.78897	3.16623
074710	Le Puy	FR	45.07774	3.76686

Continued on Next Page...

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
074810	Lyon-Satolas	FR	45.73250	5.09410
074820	Amberieu	FR	45.98859	5.33119
074860	Grenoble-St-Geoirs	FR	45.36461	5.32880
074910	Chambery	FR	45.63669	5.88310
074970	Bourg St-Maurice	FR	45.61444	6.77856
075020	Cazaux	FR	44.53333	-1.13333
075240	Agen	FR	44.17450	0.59210
075490	Aurillac	FR	44.89447	2.41753
075580	Millau	FR	44.11860	3.02029
075790	Orange	FR	44.14031	4.86359
075880	St Auban Sur Duranc	FR	44.06220	5.99328
076030	Dax	FR	43.68962	-1.07022
076100	Pau	FR	43.38494	-0.41632
076210	Tarbes-Ossun	FR	43.18869	0.00029
076270	Saint Girons	FR	43.00780	1.10320
076300	Toulouse Blagnac	FR	43.62186	1.37144
076350	Carcassonne	FR	43.21771	2.29884
076430	Montpellier	FR	43.58163	3.96950
076460	Nimes	FR	43.75751	4.41531
076500	Marignane	FR	43.43818	5.20519
076670	Hyeres	FR	43.09730	6.14600
076750	Le Luc	FR	43.38322	6.38631
076900	Nice	FR	43.64887	7.20610
077470	Perpignan	FR	42.73947	2.87576
080020	La Coruna/Alvedro	SP	43.30194	-8.37722
080110	Asturias/Aviles	SP	43.56361	-6.03472
080230	Santander	SP	43.48495	-3.78305
080250	Bilbao/Sondica	SP	43.30111	-2.91060
080290	San Sebastian	SP	43.35650	-1.79060
080420	Santiago	SP	42.89639	-8.41528
080550	Leon	SP	42.58889	-5.65556
080750	Burgos	SP	42.35750	-3.62476
080800	Vitoria	SP	42.88280	-2.72444
080840	Logrono/Agoncillo	SP	42.46025	-2.32579
080850	Pamplona/Noain	SP	42.77000	-1.64639
081410	Valladolid	SP	41.65000	-4.76667
081605	Zaragoza	SP	41.66343	-1.02294
081750	Reus	SP	41.14740	1.16720
081810	Barcelona	SP	41.28416	2.07293
081840	Gerona/Costa Brava	SP	41.90100	2.76050
082020	Salamanca/Matacan	SP	40.95210	-5.50200

Continued on Next Page...

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
082270	Madrid	SP	40.48333	-3.45000
082610	Caceres	SP	39.46733	-6.33721
082800	Albacete/Los Llanos	SP	38.94853	-1.85699
082840	Valencia	SP	39.48754	-0.48194
083060	Palma de Mallorca	SP	39.54117	2.73113
083140	Menorca	SP	39.86983	4.22351
083300	Badajoz	SP	38.88674	-6.82938
083600	Alicante	SP	38.28289	-0.57131
083730	Ibiza	SP	38.87290	1.37310
083910	Sevilla	SP	37.41678	-5.90858
083970	Moron de la Fronter	SP	37.17490	-5.61590
084100	Cordoba	SP	37.84290	-4.84697
084190	Granada	SP	37.18861	-3.77722
084290	Murcia/Alcantarilla	SP	37.95046	-1.23291
084330	Murcia/San Javier	SP	37.78336	-0.80711
084510	Jerez de la Fronter	SP	36.75089	-6.05560
084820	Malaga	SP	36.66653	-4.48944
084870	Almeria	SP	36.84576	-2.35687
085360	Lisboa	PO	38.76677	-9.13336
085450	Porto	PO	41.24810	-8.68140
085540	Faro	PO	37.01578	-7.97298
085620	Beja	PO	38.01667	-7.86667
085710	Portalegre	PO	39.28330	-7.41670
085750	Braganca	PO	41.85610	-6.70600
090910	Arkona	DD	54.68167	13.43667
091610	Boltenhagen	DD	54.00255	11.19039
091700	Warnemuende	DD	54.18022	12.08054
091770	Teterow	DD	53.76333	12.62000
092910	Angermuende	DD	53.03194	13.99278
093850	Berlin/Schonefeld	DD	52.37507	13.52270
093930	Lindnberg	DD	52.20944	14.12222
094600	Artern	DD	51.37611	11.29333
094690	Leipzig	DD	51.41806	12.23028
094880	Dresden	DD	51.12789	13.75448
094990	Goerlitz	DD	51.16417	14.95861
095540	Erfurt	DD	50.97722	10.96222
095670	Gera	DD	50.88167	12.13000
100200	List/Sylt	DL	55.01333	8.41306
100220	Leck	DL	54.79111	8.95222
100260	Husum	DL	54.51716	9.14749
100370	Schleswig-Jagel	DL	54.46341	9.51659

Continued on Next Page...

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
100460	Kiel	DL	54.37917	10.14778
101130	Norderney Island	DL	53.71249	7.15184
101260	Wittmundhaven	DL	53.54806	7.66778
101290	Bremerhaven	DL	53.53639	8.57639
101360	Nordholz	DL	53.76583	8.65528
101420	Itzehoe	DL	53.98976	9.56980
101470	Hamburg	DL	53.63333	9.98750
101560	Luebeck	DL	53.80485	10.70945
102180	Ahlhorn	DL	52.88917	8.24000
102240	Bremen	DL	53.04667	8.79694
102340	Rotenburg	DL	53.12861	9.35333
102460	Gassberg	DL	52.91972	10.19083
102530	Luchow	DL	52.97389	11.13861
103140	Hopsten	DL	52.33847	7.54337
103200	Guetersloh	DL	51.92728	8.30915
103210	Diepholz	DL	52.58944	8.34528
103280	Detmold	DL	51.94142	8.90236
103350	Bueckeburg	DL	52.27871	9.09157
103380	Hannover	DL	52.46500	9.68833
103480	Braunschweig	DL	52.29472	10.44611
103680	Wiesenburg	DL	52.12111	12.46028
103840	Berlin/Tempelhof	DL	52.47417	13.41556
104000	Dusseldorf	DL	51.29556	6.77528
104020	Wildenrath	DL	51.11428	6.20876
104040	Kalkar	DL	51.73333	6.26667
104050	Laarbruch	DL	51.60200	6.14280
104160	Dortmund	DL	51.51845	7.60918
104260	Paderborn	DL	51.61182	8.61221
104390	Fritzlar	DL	51.13694	9.29556
105020	Noervenich	DL	50.83444	6.67500
105130	Koln/Bonn	DL	51.73667	7.19444
105140	Mendig	DL	50.37028	7.32444
105150	Bendorf	DL	50.41609	7.58333
105360	Fulda	DL	50.54824	9.65410
106100	Bitburg	DL	49.95034	6.57260
106130	Buechel	DL	50.17556	7.06083
106140	Ramstein	DL	49.43463	7.59146
106160	Hahn	DL	49.94623	7.26446
106330	Wiesbaden	DL	50.04905	8.32820
106370	Frankfurt Main Airport	DL	50.04639	8.59861
106420	Hanau	DL	50.16830	8.96170

Continued on Next Page...

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
106530	Geibelstadt	DL	49.64830	9.96670
106850	Hof	DL	50.31389	11.87778
107140	Zweibruecken	DL	49.21600	7.40815
107220	Karlsruhe	DL	48.78708	8.08584
107280	Coleman	DL	49.56644	8.46677
107380	Stuttgart	DL	48.68889	9.22139
107430	Niederstetten	DL	49.39180	9.95820
107610	Weissenburg	DL	49.02028	10.96167
107630	Nurnberk	DL	49.49722	11.07806
107710	Kuemmersbruck	DL	49.41665	11.90000
107880	Straubing	DL	48.82917	12.56056
108050	Lahr	DL	48.36930	7.82770
108360	Stoetten	DL	48.66639	9.86583
108370	Laupheim	DL	48.21834	9.91138
108520	Augsberg	DL	48.42519	10.93773
108530	Neuburg	DL	48.71252	11.21099
108580	Fuerstenfeldbruck	DL	48.20000	11.26667
108690	Erding	DL	48.31667	11.95000
109000	Bremgarten	DL	47.90444	7.61861
109210	Neuhassen Ob Eck	DL	47.97722	8.91250
109350	Frierichshafen	DL	47.64944	9.48361
109470	Memmingen	DL	47.99000	10.23694
109540	Altenstadt	DL	47.83611	10.86778
114060	Cheb	CZ	50.06862	12.39110
114140	Karlovy Vary	CZ	50.20192	12.91060
114230	Primda	CZ	49.66942	12.67810
114380	Tusimice	CZ	50.37672	13.32810
114870	Kocelovice	CZ	49.46692	13.84080
115180	Praha	CZ	50.10082	14.25780
115410	Ceske Budejovice	CZ	48.94578	14.43071
116030	Liberec	CZ	50.77002	15.02420
116360	Kostelni Myslova	CZ	49.16002	15.43920
116480	Hradec Kralove	CZ	50.25000	15.85000
116590	Pribyslav	CZ	49.58282	15.76250
116790	Usti Nad Orlici	CZ	49.98032	16.42220
116980	Kucharovice	CZ	48.88252	16.08640
117100	Luka	CZ	49.65222	16.95330
117230	Brno/Turany	CZ	49.15972	16.69560
117740	Holesov	CZ	49.31862	17.57330
117820	Ostrava	CZ	49.69832	18.12170
121000	Kolobrzeg	PL	54.18000	15.58000

Continued on Next Page...

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
121050	Koszalin	PL	54.20480	16.15430
121150	Ustka	PL	54.58300	16.86700
121200	Leba	PL	54.75000	17.53306
121350	Hel	PL	54.60000	18.81667
121500	Gdansk-Rebiechowo	PL	54.37760	18.46620
121850	Ketrzyn	PL	54.06670	21.36670
121950	Suwalkl	PL	54.13000	22.95000
122000	Swinoujscie	PL	53.92290	14.24100
122050	Szczecin	PL	53.40000	14.61667
122350	Chojnice	PL	53.70000	17.55000
122500	Torun	PL	53.02920	18.54590
122700	Mlawa	PL	53.10000	20.35000
122720	Olsztyn	PL	53.77301	20.41397
122800	Mikolajki	PL	53.78330	21.58330
122950	Bialystok	PL	53.10140	23.17060
123000	Gorzow	PL	52.73330	15.26670
123300	Poznan	PL	52.42100	16.82630
123450	Kolo	PL	52.20000	18.66700
123600	Plock	PL	52.58000	19.73000
123750	Warszawa	PL	52.16580	20.96710
123850	Siedlce	PL	52.25000	22.25000
124000	Zielona	PL	52.14020	15.79700
124150	Legnica	PL	51.20000	16.20000
124240	Wroclaw II	PL	51.10270	16.88580
124350	Kalisz	PL	51.76666	18.06666
124550	Wielun	PL	51.20000	18.55000
124650	Lodz	PL	51.72190	19.39810
124950	Lublin Radaweic	PL	51.22190	22.39470
124970	Wlodawa	PL	51.55000	23.53306
125000	Jelenia Gora	PL	50.89890	15.78560
125100	Sniezka	PL	50.73630	15.73970
125200	Klodzko	PL	50.43330	16.61000
125300	Opole	PL	50.80000	17.96666
125600	Katowice	PL	50.24058	19.03274
125660	Krakow	PL	50.07775	19.79482
125700	Kielce	PL	50.81655	20.70256
125750	Tarnow	PL	50.03000	20.98000
125800	Rzeszow-Jasionka	PL	50.11104	22.03154
125950	Zamosc	PL	50.70000	23.20000
126000	Bielsko-Biala	PL	49.80500	19.00188
126500	Kasprowy Wierch	PL	49.23306	19.98306

Continued on Next Page...

Table A.1 - *Continued*

WMO	Station Name	Country	Latitude	Longitude
126950	Przemysl	PL	49.80000	22.76670

Table A.1: Listing of Selected Stations

Appendix B

Statistical Methods

B.1 Generalised Extreme Value Distribution: Statistical Properties

Probability density function (p.d.f.)

$$f_\theta(x) = \begin{cases} \dfrac{1}{\sigma}\left[1+\xi\left(\dfrac{x-\mu}{\sigma}\right)\right]^{-1/\xi-1} \exp\left\{-\left[1+\xi\left(\dfrac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} & \xi \neq 0 \\ \dfrac{1}{\sigma}\exp\left[-\left(\dfrac{x-\mu}{\sigma}\right)\right]\exp\left\{-\exp\left[-\left(\dfrac{x-\mu}{\sigma}\right)\right]\right\} & \xi = 0 \end{cases} \tag{B.1}$$

Maximum likelihood score function

$$\boldsymbol{s}(x;\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log f_{\boldsymbol{\theta}}(x) \tag{B.2}$$

$$= \left[s_1(x;\boldsymbol{\theta}) \ldots s_p(x;\boldsymbol{\theta}) \right], \text{for } \boldsymbol{\theta} \text{ length } p \tag{B.3}$$

for the GEVD when $\xi \neq 0$

$$s_1(x;\boldsymbol{\theta}) = \frac{\partial}{\partial \mu} \log f_{\boldsymbol{\theta}}(x) \tag{B.4}$$

$$= -\frac{\xi}{\sigma}\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1} - \frac{1}{\sigma}\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi-1} \tag{B.5}$$

$$s_2(x;\boldsymbol{\theta}) = \frac{\partial}{\partial \sigma} \log f_{\boldsymbol{\theta}}(x) \tag{B.6}$$

$$= -\frac{1}{\sigma} + (1+\xi)\left(\frac{x-\mu}{\sigma^2}\right)\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1} - \left(\frac{x-\mu}{\sigma^2}\right)\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi-1} \tag{B.7}$$

$$s_3(x;\boldsymbol{\theta}) = \frac{\partial}{\partial \xi}\log f_{\boldsymbol{\theta}}(x) \tag{B.8}$$

$$= -\frac{1}{\xi^2}\log\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]\left\{1-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} - \left(\frac{x-\mu}{\xi\sigma}\right)$$

$$\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1}\left\{1+\xi-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{B.9}$$

for the GEVD when $\xi = 0$

$$s_1(x;\boldsymbol{\theta}) = \frac{\partial}{\partial \mu}\log f_{\boldsymbol{\theta}}(x) \tag{B.10}$$

$$= \frac{1}{\sigma}\left\{1-\exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\} \tag{B.11}$$

$$s_2(x;\boldsymbol{\theta}) = \frac{\partial}{\partial \sigma}\log f_{\boldsymbol{\theta}}(x) \tag{B.12}$$

$$= \sigma + \left(\frac{x-\mu}{\sigma^2}\right)\left\{1-\exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\} \tag{B.13}$$

B.2 Influence Function

B.2.1 Overview

The influence function (IF) was originally developed by Hampel (1968, 1974) as the influence curve and forms the basis of robustness theory. Conceptually, the IF measures the asymptotic bias resulting from an induced perturbation ($\epsilon\delta_x$) at a point in the observations.

Hampel (1974) defines the IF of the estimator T at the probability distribution F by

$$IF(x;T,F) = \lim_{\epsilon \to 0} \left\{ \frac{T[(1-\epsilon)F + \epsilon\delta_x] - T(F)}{\epsilon} \right\} \tag{B.14}$$

which represents the directional derivative in the direction of a point mass, δ_x, at x . The asymptotic variance is then given by

$$V(T,F) = \int IF(x;T,F)IF(x;T,F)^T dF(x). \tag{B.15}$$

The IF allows the robustness of an estimator to be evaluated. Ideally the IF will be bounded, thus ensuring that any small contamination in the data does not largely affect the outcome of the estimator.

B.2.2 Derivation: Maximum Likelihood Estimators

Hampel *et al.* (1986) derive the influence function for the M-estimators from Equation B.14, giving

$$IF(x;\psi,F_{\boldsymbol{\theta}}) = \boldsymbol{M}(\boldsymbol{\psi},F_{\boldsymbol{\theta}})^{-1}\boldsymbol{\psi}(x,\boldsymbol{\theta}) \tag{B.16}$$

where

$$\boldsymbol{M}(\boldsymbol{\psi},F_{\theta}) = -\int \frac{\partial}{\partial \boldsymbol{\theta}}\boldsymbol{\psi}(x,\boldsymbol{\theta})dF_{\boldsymbol{\theta}}(x) \tag{B.17}$$

and the asymptotic covariance is given by

$$\boldsymbol{V}(\boldsymbol{\psi},F_{\boldsymbol{\theta}}) = \boldsymbol{M}(\boldsymbol{\psi},F_{\boldsymbol{\theta}})^{-1}\boldsymbol{Q}(\boldsymbol{\psi},F_{\boldsymbol{\theta}})\boldsymbol{M}(\boldsymbol{\psi},F_{\boldsymbol{\theta}})^{-T} \tag{B.18}$$

where

$$\boldsymbol{Q}(\boldsymbol{\psi}, F_\theta) = \int \boldsymbol{\psi}(x, \boldsymbol{\theta})\boldsymbol{\psi}(x, \boldsymbol{\theta})^T dF_\theta(x) \tag{B.19}$$

Substituting the score function (from Equation 4.13)

$$\boldsymbol{s}(x_i; \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log f_\theta(x; \boldsymbol{\theta}). \tag{B.20}$$

into Equation B.17 yields the Fisher Information matrix $\boldsymbol{J}(\boldsymbol{\theta})$. The Fisher Information is the variance of the score and for a given probability density function represents a measure of the amount of information that an observation carries about the estimated parameters. The Fisher Information matrix may be formulated as

$$J(\boldsymbol{\theta}) = -\int \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \boldsymbol{s}(x, \boldsymbol{\theta}) dF_\theta(x) \tag{B.21}$$

$$= \int \boldsymbol{s}(x, \boldsymbol{\theta})\boldsymbol{s}(x, \boldsymbol{\theta})^T dF_\theta \tag{B.22}$$

The former is often referred to as the observed Fisher Information and the latter as the expected Fisher Information. The expected Fisher Information is generally favoured since the second derivatives of the score functions do not require calculation, although several authors (Prescott and Walden, 1983; Efron and Hinkley, 1978) have shown that the observed information matrix can provide better estimates. By substitution of the Fisher Information matrix into Equation B.16, the influence function becomes

$$IF(x; \boldsymbol{\psi}, F_\theta) = \boldsymbol{J}(\boldsymbol{\theta})^{-1}\boldsymbol{\psi}(x, \boldsymbol{\theta}) \tag{B.23}$$

where for the MLE, $\boldsymbol{\psi}(x, \boldsymbol{\theta}) = \boldsymbol{s}(x, \boldsymbol{\theta})$. The equation is unbounded as a result of the score function given by Equation B.20 being unbounded in x (Dupuis and Field, 1998b).

B.3 Optimal Bias-Robust Estimators

B.3.1 Estimator

A detailed derivation of the estimator is provided by Hampel *et al.* (1986) for both single and multidimensional estimators. A bounded function for $\boldsymbol{\psi}$ can be derived as

$$\boldsymbol{\psi}_c^{A,a}(x) = h_c\{\boldsymbol{A}[\boldsymbol{s}(x,\boldsymbol{\theta}) - \boldsymbol{a}]\} \tag{B.24}$$

where $\boldsymbol{A}$ and $\boldsymbol{a}$ arise from Fisher consistency and ensuring the function coincides with the influence function of the M-estimator given by Equation B.16. The matrix $\boldsymbol{A}$ is lower triangular and is written as

$$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{M}_2^{-1} \tag{B.25}$$

and the vector $\boldsymbol{a}$ is solved implicitly by

$$\boldsymbol{a}(\boldsymbol{\theta}) = \frac{\int \boldsymbol{s}(x,\boldsymbol{\theta}) w_c(x,\boldsymbol{\theta}) dF_{\boldsymbol{\theta}}(x)}{\int w_c(x,\boldsymbol{\theta}) dF_{\boldsymbol{\theta}}(x)} \tag{B.26}$$

Thus, considering the estimator given by Equation B.24, the matrix $\boldsymbol{M}$ constructed in Equation B.17, which is analogous to the expected Fisher Information defined by Equation B.22, can be written as

$$\boldsymbol{M}_k = \int \{\boldsymbol{s}(x,\boldsymbol{\theta}) - \boldsymbol{a}\}\{\boldsymbol{s}(x,\boldsymbol{\theta}) - \boldsymbol{a}\}^T W_c(x)^k dF(x) \tag{B.27}$$

where the weighting function as defined in Equation 4.14 is

$$W_c(x) = \min\{1, c/\parallel \boldsymbol{A}[\boldsymbol{s}(x,\boldsymbol{\theta}) - \boldsymbol{a}] \parallel\}. \tag{B.28}$$

If the robustness constant c equals infinity the MLE is achieved since $W_c(x, \boldsymbol{\theta}) = 1$ for all observations and parameters. For the standardised OBRE case given here $c > \sqrt{p}$, where p is the length of the parameter vector $\boldsymbol{\theta}$, and the most robust case is achieved as $c \rightarrow \sqrt{p}$ (Dupuis and Field, 1998b).

B.3.2 Algorithm

The algorithm for the OBRE procedure closely follows the procedure outlined by Dupuis and Field (1998b) and the standardised OBRE algorithm of Hampel *et al.* (1986). For a sequence $\boldsymbol{x} = \{x_1, ..., x_n\}$ of observed wind data:

1. Develop an initial estimate $\hat{\boldsymbol{\theta}}$ for the parameter vector $\boldsymbol{\theta}$ from MLE or PWM and set the robustness coefficient c.
2. Calculate $\boldsymbol{A} = \boldsymbol{J}^{1/2}(\boldsymbol{\theta})^{-T}$ where $J(\boldsymbol{\theta})$ is given by Equation B.22 and set $\boldsymbol{a} = 0$.
3. Solve for $\boldsymbol{A}$ and $\boldsymbol{a}$ using Equations B.25 and B.26 respectively, and iterate until convergence.
4. Calculate M_1 from Equation B.27.
5. Solve for $\Delta\boldsymbol{\theta}$ where

$$\Delta\boldsymbol{\theta} = M_{ave}\{h_c\{\boldsymbol{A}[\boldsymbol{s}(x,t) - \boldsymbol{a}]\}\} \tag{B.29}$$

$$= M_1^{-1}\frac{1}{n}\sum\{[\boldsymbol{s}(x,t) - \boldsymbol{a}]W_c(x,\boldsymbol{\theta})\} \tag{B.30}$$

6. If $\Delta\boldsymbol{\theta}$ is less than the desired precision the answer has converged. Otherwise $\hat{\boldsymbol{\theta}} + \Delta\boldsymbol{\theta}$ and return to step 3.

The integrals in Equations B.22 and B.27 of Steps 2 and 4 can be calculated using the empirical distribution function, however, numerical integration must be used in Equation

B.26 for the calculation of $\boldsymbol{a}$ in Step 3 otherwise every estimate will satisfy $\sum_{i=1}^{n} \boldsymbol{\psi}(x_i; \theta) = 0$ (Dupuis and Field, 1998b).

Curriculum Vitae

Name:	David Allan Gatey	
Post-Secondary Education and Degrees:	Ph.D., Civil & Environmental Engineering The University of Western Ontario London, Ontario, Canada	2006 - 2011
	B.E.Sc.(Hons.), Civil & Environmental Engineering The University of Western Ontario London, Ontario, Canada	2001 - 2006
Honours and Awards:	NSERC Postgraduate Scholarship Doctoral (National)	2008-2011
	Ontario Graduate Scholarship Doctoral - Declined (Provincial)	2008-2009
	Ontario Graduate Scholarship Masters (Provincial)	2006-2008
	Civil Undergraduate Engineering Thesis First Place (Institutional)	2006
	Dr. James A. Vance Gold Medal in Civil Engineering Top of Graduating Class (Institutional)	2006
	NSERC Undergraduate Student Research Award Undergraduate (National)	2005
Related Work Experience:	Teaching Assistant The University of Western Ontario	2006 - 2011

Publications:

Refereed Journal Articles:

Gatey, D.A. and Miller, C.A. (2007) An investigation into 50-year return period wind speed differences for Europe, *Journal of Wind Engineering and Industrial Aerodynamics*, 95: 1040-1052, doi:10.1016/j.jweia.2007.01.016.

Conference and Technical Reports:

Gatey, D.A., Case, P., Ho, T.C.E., Isyumov, N. and Miller, C.A. (2011) Design wind climates for very tall structures, *13th International Conference on Wind Engineering*, Amsterdam, Netherlands, July 10-15.

Gatey, D.A., Miller, C.A., Mitchell, S.J., Sagi, P. and Newson, T. (2011) Tree interaction with wind - a scaled experimental system, *Wind and Trees*, International Union of Forest Research Organizations Research Conference.

Sagi, P., Newson, T., Mitchell, S.J., **Gatey, D.A.** and Miller, C.A. (2011) Scaled tree-root system behaviour in a sand substrate under wind loading, *Wind and Trees*, International Union of Forest Research Organizations Research Conference.

Gatey, D.A. and Ho, T.C.E. (2011) Summary of the wind climate analysis for Blythe, CA, *BLWTL-SSXX-2011*, Boundary Layer Wind Tunnel Laboratory, London, Ontario, Canada.

Ho, T.C.E., **Gatey, D.A.** and Garnham, D. (2011) Site-specific wind climate and topographic studies for Mekka, Saudi Arabia, *BLWTL-SSX-2011*, Boundary Layer Wind Tunnel Laboratory, London, Ontario, Canada.

Gatey, D.A. and Miller, C.A. (2007) Assessment of models simulating wind effects over non-uniform terrain, *Proceedings of the 12th International Conference on Wind Engineering*, Cairns, Australia, July 1-6.

Gatey, D.A. and Miller, C.A. (2007) A H*WIND derived pressure-wind speed relationship for the North Atlantic, University of Western Ontario Graduate Engineering Research Day (Poster).

Gatey, D.A. (2006) Assessment of models simulating wind effects over non-uniform terrain, *Bachelor of Engineering Science Thesis*, The University of Western Ontario.

Gatey, D.A. and Miller, C.A. (2005) An investigation into European design wind speed differences, *Proceedings of the 4th European and African Conference on Wind Engineering*, Prague, Czech Republic, July 11-15.