Analysis of the Rijkoort-Weibull model

A. Smits (Ilja)
KNMI, WM/KD
P.O. Box 201, 3730 AE De Bilt
Tel: +31 30 2206874, Fax: +31 30 2210407
Email: Ilja.Smits@knmi.nl
Date: 2 February, 2001

Index:

Abstract

This report makes part of the KNMI HYDRA project and it gives an overview of the reproduction and analysis of and adjustments to the so called “A Compound Weibull Model For The Description Of Surface Wind Velocity Distributions” (Rijkoort, 1983). This model is used to calculate return levels of hourly mean wind speeds with high return periods (up to 10^4 years) at several stations in the Netherlands.
It turned out that the persistence correction used in the model causes unreliable results with respect to the return levels of interest, especially when the correction is used in combination with the occurrence of relatively many low wind speed values.
In order to use the model properly, adjustments have to be made. It makes sense to use a threshold to prevent the impact of low wind speed values and to determine the persistence correction based on physical grounds.
Alternative extreme wind models also have been examined. The combined GEV distribution results in unreliable return levels due to the split up into seasons and wind direction groups. The one-step Markov chain model has to be extended substantially before results of interest can be obtained. Finally, the analysis of separate storms instead of the parent distribution of wind speed or the distribution of annual maxima of wind speed is discussed briefly.

1 Introduction

The National Institute for Coastal and Marine Management (RIKZ) and the Institute for Inland Water Management and Waste Water Treatment (RIZA) in the Netherlands are legally obliged to redo their risk assessment of the Dutch dike systems regularly. On the IJssellake, the Waddensea and parts of the Zeeland waters wind is the main source of waves. This urged the need for an updated wind climate assessment of the Netherlands, including the water-land (and vice versa) transition zones. This is the contribution of the Royal Netherlands Meteorological Institute (KNMI) to the HYDRA (**HYD**raulische **RA**ndvoorwaarden) project.

One of the main goals of the project is to provide return levels of hourly mean wind speeds corresponding with return periods up to 10^4 years. At this moment, the Directorate-General of Public Works and Water Management (RWS) uses the outcomes from a model described in detail in “A Compound Weibull Model For The Description Of Surface Wind Velocity Distributions” (Rijkoort, 1983). Because of the doubt about the quality of this model (henceforth the Rijkoort-Weibull model (RW-model)) and because of the extension of available time series up to now, it seemed worthwhile to review and to update the model and to adjust it if necessary. The results of this are set out in this report.

First of all, the available time series and stations that are used during the research are discussed in Chapter 2. Chapter 3 describes the RW-model in rough lines. In this Chapter, a number of phrases and formulas are literally taken from the paper of Rijkoort, but often more explanation about the interpretation of the theory turned out to be necessary which has been added. In Chapter 4, the reproduction of the model is discussed. This reproduction is based on the same observational series Rijkoort used. Chapter 5 evaluates the model and Chapter 6 reviews some possible adjustments to it. Also alternative wind extreme models have been examined. These models are discussed in Chapter 7. Chapter 8 summarises the main conclusions and recommendations are given in Chapter 9. Finally, several numerical and graphical outcomes can be found in the Appendices.

2 Available stations and time series

In this report wind data is used from several stations of the KNMI-network in the Netherlands. At the stage of reproduction of the RW-model, a selection of twelve stations is made and used. These stations are equal to the stations Rijkoort used in the application of his model. Furthermore, also the time period used is equal to the time period used by Rijkoort (1962-1976). Both the same stations and the same time period makes comparison between the results of Rijkoort and the reproduction valid.
More stations and longer time periods have become available for analysis. These data also are used in this report. Figure 2.1 shows the geographical locations of all the stations used and Figure 2.2 shows the tabulated representation of the time series that have been used.

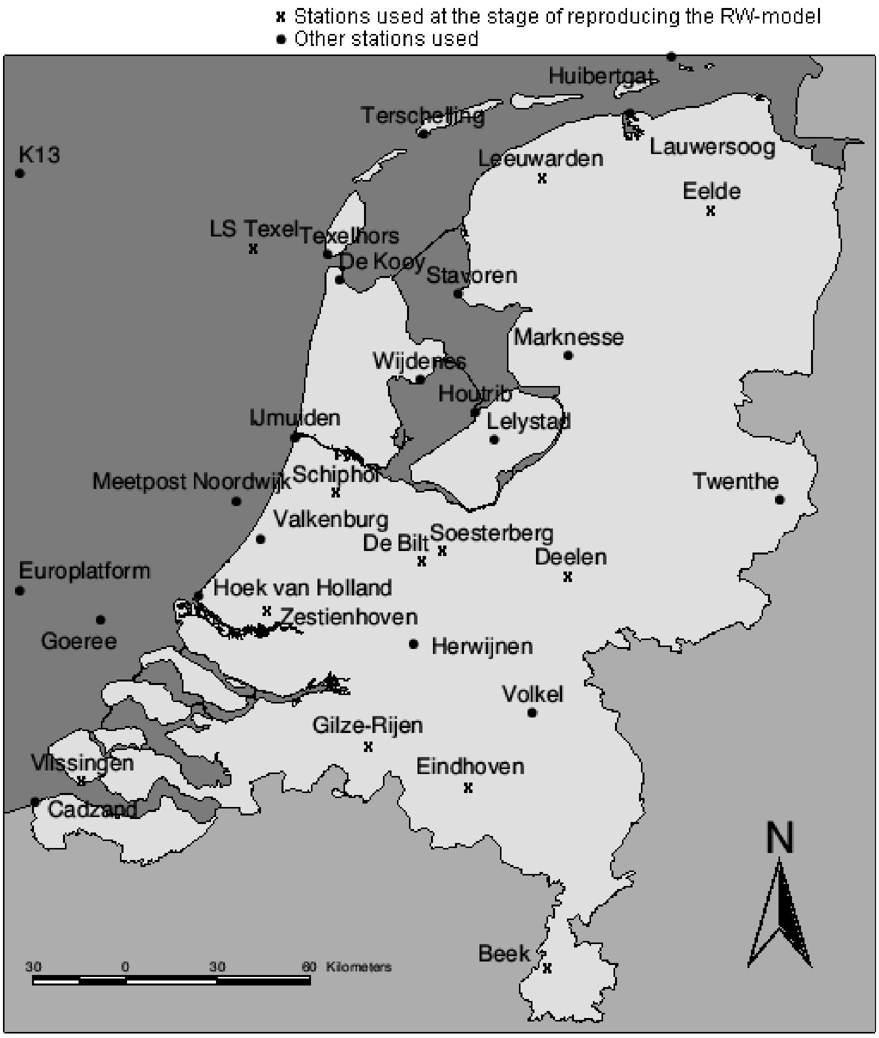

Figure 2.1: Stations used.

The wind speed data that are used are from a data set compiled for this project (Verkaik, 2000a). The time series that are included by this data set consist of hourly mean wind speed values. These values are corrected for local terrain roughness in a way that the measured wind speeds are translated to a height of 10 meters over open terrain (roughness length z_0 0.03 meter). These translated measured wind speeds are called potential wind speeds. The exposure correction method was developed by Wieringa (1976, 1986). It includes the estimation of direction dependent roughness lengths at the anemometer locations by gustiness analysis. It was applied in the assessment of the Dutch wind climate by Wieringa & Rijkoort (1983). The original parameters of the RW-model are based on these observations only. The reproduction of the RW-model also uses these observations. More recent observations are corrected using the same method although a different gustiness model has been used (Beljaars 1987, Verkaik 2000b). Also these data have been used in this report.

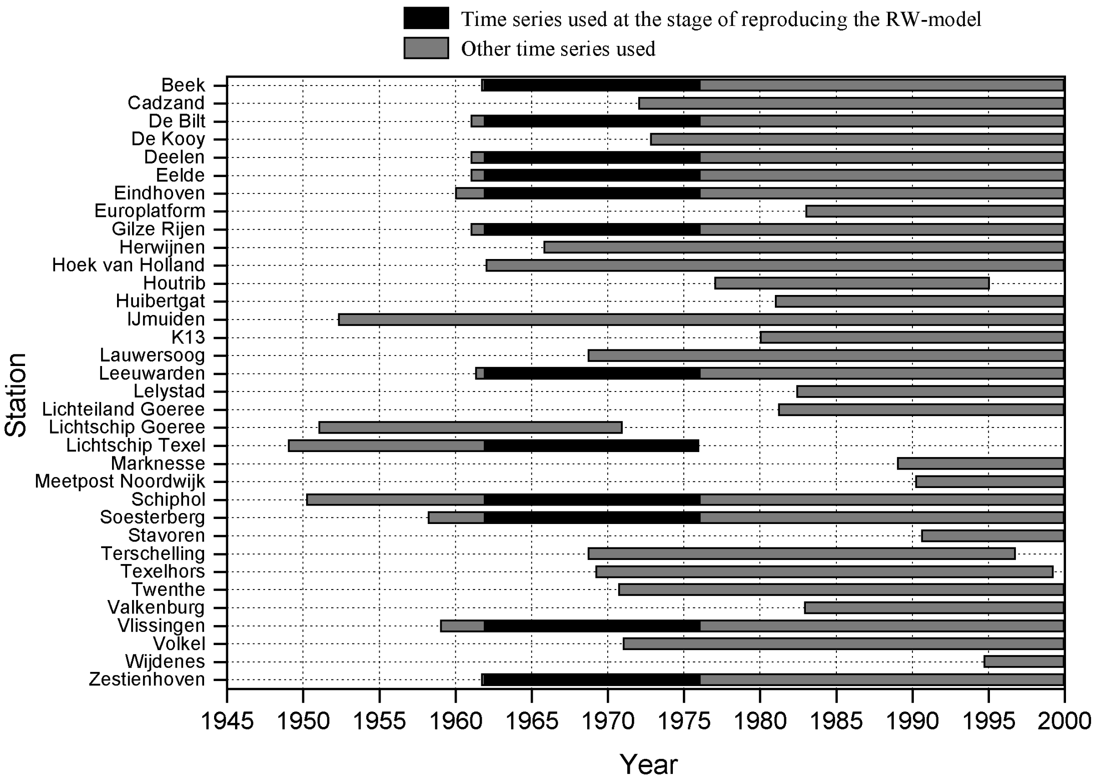

Figure 2.2: Time series used.

3 Description of the Rijkoort-Weibull model

3.1 The Compound model

The RW-model is based on the Weibull distribution. This distribution is used frequently for wind speeds and is commonly accepted. Its distribution function has the form

$$F(u) = 1 - \exp\left[-\left(\frac{u}{\alpha}\right)^{\kappa}\right], \quad \textbf{(3.1)}$$

where u is the wind speed, κ the shape parameter and α the scale parameter. In this report, the wind speed concerns only hourly mean values and the unit of wind speed used equals meters per second (m/s). The shape parameter κ determines the degree of peakedness of the distribution function around the mean and the width of the tails of the distribution function. Because our interest concerns extreme wind speeds the value of κ is very essential in the model. The scale parameter α can be seen as a multiplication factor and therefore as a good indication for the mean wind speed.

Although the Weibull distribution fits wind speed reasonably well, Rijkoort introduced a modification of the Weibull distribution for nighttime situations (Rijkoort, 1980). This modification approaches (3.1) at high wind speeds. The modification results in two separate distributions for nighttime ($F_n(u)$) and daytime ($F_d(u)$):

$$F_n(u) = 1 - \exp\left[-\left(\frac{u}{\alpha}\left\{1 + \gamma \exp\left(-\frac{u}{5}\right)\right\}\right)^{\kappa}\right], \quad \textbf{(3.2)}$$

$$F_d(u) = 1 - \exp\left[-\left(\frac{u}{\alpha}\right)^{\kappa}\right], \quad \textbf{(3.3)}$$

where γ is the extra parameter in the nighttime distribution and also called the stability parameter. The factor 5 in (3.2) has been derived empirically by Rijkoort.
$F_n(u)$ indicates the probability that on an arbitrary moment during nighttime the wind speed is less than u.
$F_d(u)$ is defined likewise but nighttime is replaced by daytime.
When using these distributions it's necessary to define the days and nights. Table 3.1 is used for this purpose:

Month	Hours
January	10-14
February	09-15
March	08-16
April	07-17
May	06-18
June	05-18

Month	Hours
July	06-18
August	06-18
September	07-17
October	08-16
November	09-14
December	10-13

Table 3.1: Hours assigned to daytime per month (UTC). The tabulated hours refer to the previous hours. For example, 10 means 09:00 – 10:00 UTC.

A next step in the RW-model is to split up the data into six seasons and twelve wind direction groups. This is done because the overall (without dividing the data into seasons and wind directions) Weibull distribution seems to fit over a limited range (at most between 3 and 15 m/s). When using a subdivision into seasons and wind directions the separate Weibull fits seem to fit in a better way. The subdivision into seasons is listed in Table 3.2 and the subdivision into wind directions is listed in Table 3.3.

	Season	Months
1	Midwinter	January – February
2	Spring	March – April
3	Presummer	May – June
4	Midsummer	July – August
5	Autumn	September – October
6	Prewinter	November – December

Table 3.2: Subdivision into seasons.

Wind direction		
1	30°	015°–044°
2	60°	045°–074°
3	90°	075°–104°
4	120°	105°–134°
5	150°	135°–164°
6	180°	165°–194°
7	210°	195°–224°
8	240°	225°–254°
9	270°	255°–284°
10	300°	285°–314°
11	330°	315°–344°
12	360°	345°–014°

Table 3.3: Subdivision into wind directions.

This split up into season and wind direction results in:

$$F_n(u \mid i,j) = 1 - \exp\left[-\left(\frac{u}{\alpha_{ij}}\left\{1+\gamma_{ij}\exp\left(-\frac{u}{5}\right)\right\}\right)^{\kappa_{ij}}\right], \tag{3.4}$$

$$F_d(u \mid i,j) = 1 - \exp\left[-\left(\frac{u}{\alpha_{ij}}\right)^{\kappa_{ij}}\right], \tag{3.5}$$

where i (1, 2, . . . ,12) indicates the azimuth group and j (1, 2, . . . ,6) the season group. $F_n(u \mid i, j)$ indicates the probability that on a arbitrary moment during nighttime the wind speed is less than u given that this moment lies in season j and that the wind direction is i. $F_d(u \mid i, j)$ is defined likewise but nighttime is replaced by daytime.

After the split up we have a great number of Weibull parameters (12·6·3 = 216 per station):

$$\alpha_{ij}, \kappa_{ij} \text{ and } \gamma_{ij}.$$

These parameters are fitted by the method of maximum likelihood.

When we combine day- and nighttime, the wind distribution for an individual season-azimuth group is now

$$F(u \mid i,j) = \frac{\delta_{ij}F_d(u \mid i,j) + \nu_{ij}F_n(u \mid i,j)}{\delta_{ij} + \nu_{ij}}, \tag{3.6}$$

where the parameters δ_{ij} and ν_{ij} indicate the numbers of day- and nighttime observations in a year during season j with wind direction is i. The parameters δ_{ij} and ν_{ij} increase the total number of parameters to 360

(=(12·6·3)+(12·6·2)). Estimates of δ_{ij} and ν_{ij} can be made by averaging the observations over the years that are used for analysis.

Furthermore, when we combine these individual azimuth-season distributions we get a distribution for each wind direction

$$F(u \mid i) = \frac{\sum_{j=1}^{6}\left(\delta_{ij} F_d(u \mid i,j) + \nu_{ij} F_n(u \mid i,j)\right)}{\sum_{j=1}^{6}\left(\delta_{ij} + \nu_{ij}\right)}, \tag{3.7}$$

and for each season

$$F(u \mid j) = \frac{\sum_{i=1}^{12}\left(\delta_{ij} F_d(u \mid i,j) + \nu_{ij} F_n(u \mid i,j)\right)}{\sum_{i=1}^{12}\left(\delta_{ij} + \nu_{ij}\right)}. \tag{3.8}$$

When we combine all season-azimuth groups we get the distribution for any arbitrary moment:

$$F(u) = \frac{\sum_{i=1}^{12}\sum_{j=1}^{6}\left(\delta_{ij} F_d(u \mid i,j) + \nu_{ij} F_n(u \mid i,j)\right)}{\sum_{i=1}^{12}\sum_{j=1}^{6}\left(\delta_{ij} + \nu_{ij}\right)}. \tag{3.9}$$

$F(u \mid i, j)$ indicates the probability that on a arbitrary moment the wind speed is less than u given that this moment lies in season j and that the wind direction is i. $F(u \mid i)$ indicates the probability that on a arbitrary moment the wind speed is less than u given that the wind direction at this moment is i. $F(u \mid j)$ indicates the probability that on a arbitrary moment the wind speed is less than u given that this moment lies in season j. Finally, $F(u)$ indicates the probability that on a arbitrary moment the wind speed is less than u without any given situation.

3.2 Smoothing of the model-parameters

In order to make the 360 distribution parameters per station more manageable, Rijkoort has chosen to "smooth" these parameters. Because of the often bimodal structure of a wind rose he subjected the 12 azimuthal parameters to harmonic analysis. This resulted is the following sine forms:

$$\alpha_{ij} = a_{j0} + a_{j1}\sin\left(30i + a_{j2}\right) + a_{j3}\sin\left(60i + a_{j4}\right), \tag{3.10}$$

$$\kappa_{ij} = k_{j0} + k_{j1}\sin\left(30i + k_{j2}\right) + k_{j3}\sin\left(60i + k_{j4}\right), \tag{3.11}$$

$$\gamma_{ij} = g_{j0} + g_{j1}\sin\left(30i + g_{j2}\right) + g_{j3}\sin\left(60i + g_{j4}\right), \tag{3.12}$$

$$\delta_{ij} = d_{j0} + d_{j1}\sin\left(30i + d_{j2}\right) + d_{j3}\sin\left(60i + d_{j4}\right), \tag{3.13}$$

$$\nu_{ij} = n_{j0} + n_{j1}\sin\left(30i + n_{j2}\right) + n_{j3}\sin\left(60i + n_{j4}\right). \tag{3.14}$$

The first harmonic parameter (respectively a_{j0}, k_{j0} , g_{j0} , d_{j0} and n_{j0}, $j = 1, . . ,6$) is equal to the mean of the parameter in question combined over all azimuth sectors. The parameters d_{j0} and n_{j0} are in fact no parameters but constants, because their means combined over all azimuth sectors are equal to the numbers of hours in season j divided by the total numbers of azimuth sectors (=12). d_{j0} and n_{j0} are therefore fixed numbers and the same for all stations. The values of d_{j0} and n_{j0} are given in Table 3.4.

In this fashion the 360 parameters per station are reduced to 150 (=5·6·5) parameters per station. For geographically interpolating purposes though, 150 parameters is still an huge amount of parameters. For this reason, more steps of simplification and smoothing are performed which are described in detail in the paper of Rijkoort (1983). Some of these steps will be discussed in Chapter 4.

	Jan-Feb	Mar-Apr	May-Jun	Jul-Aug	Sep-Oct	Nov-Dec
d_{j0}	29.40	50.75	68.58	67.17	50.75	25.33
n_{j0}	89.10	71.25	53.42	56.83	71.25	96.67

Table 3.4: Values of the constants d_{j0} and n_{j0}.

3.3 Model extension for the calculation of extreme values

The main goal of the RW-model is to calculate return levels of wind speed with high return periods (T_p). So, our interest doesn't concern the whole range of the wind speed distribution, but only the tail of it: the extremes.

In theory, a possible method to calculate extreme return levels of wind speed is to derive these values directly from the wind speed distributions determined in the preceding Paragraphs 3.1 and 3.2. For example, when using the combined distribution $F(u)$ based on the several seasons and wind directions, the return level u corresponding with return period $T(u)$ expressed in hours can be calculated by

$$u = F^{-1}(1 - \frac{1}{T(u)}), \tag{3.15}$$

where F^{-1} represents the inverse of the distribution function F.
However, doing this will not result in return levels of interest. Because time series of wind speed are strongly autocorrelated, many extreme values are clustered within a relative short time, which is the case during a storm. Such a storm is only one event we are interested in, while a certain wind speed can be exceeded more than once during that same storm.
As an example the hourly mean wind speed time series of station Schiphol in the period 1951-1999 can be taken. In this period the value of 22.8 m/s is exceeded 51 times (on average 1.04 times a year). Furthermore, each time the value of 22.8 is exceeded on average 2.2 more wind speeds exceed this value, so that an exceedance lasts on average 3.2 hours.
The mean return period of a single exceedance is equal to 1/1.04 = 0.96 years, but the mean return period of each event (the one of interest) equals 3.03 years, because the exceedance frequency of each event is not 1.04 times a year but 1.04/3.2 = 0.33 times a year. Clearly, calculation by ways of (3.15) does not take this into account.
Based on this example the relation between the exact definition of the return period and the exceedance frequency is given by

$$T(u) = \frac{1}{f(u)}, \tag{3.16}$$

where $T(u)$ represents the mean duration in years between exceedances of the level u and $f(u)$ represents the average number of exceedances per year of the level u.

The method Rijkoort applies to solve the problem of the autocorrelation between successive values of wind speed is the determination of the distribution of the maximum value in a predefined not too short period (for example, the RW-model uses one year), because these maximum values are rather independent of each other. The maxima are not fully independent, because a storm can occur at the end of a period while continuing in the following period. This doesn't occur often, so this dependency is not that severe. More important is the dependency between successive years. Experience learns that years with many stormy periods often are followed by years with many high winds also. Nevertheless, the

dependency between the maxima of successive years will be much less than the dependency between successive hours.
Using the distribution of the maximum, though, yields a different interpretation of the resulting return periods. When the period is defined as a year, the resulting return period corresponding with a certain wind speed can not be less than one year, while exceedances for the considered wind speed can occur more frequently within the year. So, a change in definition about the return period has been introduced.
So theoretically we have, expressed in formula,

$$P(U_{\max} > u) \le f(u)\,, \tag{3.17}$$

where $U_{\max}$ represents the annual maximum wind speed.
But if it can be assumed that an exceedance of u never occurs together with an other exceedance of u within the same year, then the equal sign holds (Geerse, 1999). In other words, if the probability of two or more exceedances of u in the same year can be neglected, then the probability that $U_{\max}$ exceeds u equals the exceedance frequency $f(u)$, and from (3.16) then

$$P(U_{\max} > u) = f(u) = \frac{1}{T(u)}\,. \tag{3.18}$$

Moreover, it's easy to show that the probability of two or more exceedances of u within the same year can be neglected if $f(u)$ is smaller than say 1/10 or 1/20 per year and if exceedances of u can be approximately assumed to be independent.
In the RW-model it is assumed that (3.18) holds to a sufficiently good approximation. This means that the two definitions of return periods, the one in terms of exceedance frequencies and the other in terms of the annual maxima, coincide.

Rijkoort uses the results from the preceding paragraphs to extend the model for the calculation of extreme wind speed values, because the exact general formula for maximum values is based on the basic distribution model

$$G_N(x) = \{F(x)\}^N\,, \tag{3.19}$$

where $F(x)$ is the probability that an outcome of the stochastic variable X is smaller than x, whereby (3.19) means the probability that all outcomes of X from a sample of size N are smaller than x, so that also the largest element $x_N < x$. Then the probability that this maximum x_N exceeds x is:

$$\overline{G}_N(x) := 1 - G_N(x)\,. \tag{3.20}$$

In the case of the compound model we get for an individual azimuth-season group (i,j):

$$G_{ij}(u) = \{F_d(u \mid i,j)\}^{\delta_{ij}} \{F_n(u \mid i,j)\}^{\nu_{ij}}\,. \tag{3.21}$$

(3.21) indicates the probability that in season j the maximum of all wind speeds with wind direction i is smaller than u, given that $\delta_{ij}+\nu_{ij}$ hours occur in season j with wind direction i. Because the number of hours in a certain season with a certain wind direction fluctuates over years and normally doesn't equal the average value over years $(\delta_{ij}+\nu_{ij})$, (3.21) is an approximation of the probability that in season j the maximum of all wind speeds with wind direction i is smaller than u. For example, imagine a year where season j includes a high number of hours with wind direction i. In this year, the probability that in season j the maximum of all wind speeds with wind direction i is smaller than u will be less compared with a year where season j includes a low number of hours with wind direction i. (3.21) does not take account of this fluctuation in the number of hours over years. This means the following:

$$P(U_{\max,ij} > u) \approx 1 - G_{ij}(u)\,, \tag{3.22}$$

where $U_{\max,ij}$ represents the annual wind speed maxima per azimuth sector and season.
Formula (3.21) only can be used when the hourly mean values of wind speed are mutually independent. Because this is evidently not the case for wind speeds (like stated earlier), Rijkoort introduces a persistence correction factor $q_{ij}(u)$ in order to reduce the δ_{ij} and ν_{ij} to numbers of seemingly independent elements, as follows:

$$G_{ij}(u) = \{F_d(u \mid i,j)\}^{\frac{\delta_{ij}}{q_{ij}(u)}} \{F_n(u \mid i,j)\}^{\frac{\nu_{ij}}{q_{ij}(u)}} . \tag{3.23}$$

Determination of the persistence correction factor $q_{ij}(u)$ is described in detail in Paragraph 3.4.
For the extreme values, belonging respectively to azimuth sectors, seasons and the full year, the distributions have the following basic forms:

$$G_i(u) = \prod_{j=1}^{6} G_{ij}(u) , \tag{3.24}$$

$$G_j(u) = \prod_{i=1}^{12} G_{ij}(u) , \tag{3.25}$$

$$G(u) = \prod_{i=1}^{12}\prod_{j=1}^{6} G_{ij}(u) . \tag{3.26}$$

Again, these formulas only can be used when the maximum hourly mean values of wind speed per season and azimuth sector are independent of each other. However, it is possible that some dependence exists between neighbouring azimuth sectors (during a storm multiple wind directions can occur) or between subsequent seasons (a storm can occur at the end of a season while continuing in the following season). In that case additional corrections for persistence might be required. Determination of these factors is also described in Paragraph 3.4.

Finally, return periods in years for individual season-azimuth sectors, seasons, azimuth sectors and for the whole year can be calculated based on (3.23), (3.24), (3.25) and (3.26):

$$T_{ij}(u) = \frac{1}{1 - G_{ij}(u)} , \tag{3.27}$$

$$T_j(u) = \frac{1}{1 - G_j(u)} , \tag{3.28}$$

$$T_i(u) = \frac{1}{1 - G_i(u)} , \tag{3.29}$$

$$T(u) = \frac{1}{1 - G(u)} . \tag{3.30}$$

3.4 Determination of persistence correction factors

Rijkoort determines the persistence correction factors, that are needed as a result of the dependency between successive hourly wind speed values, with the help of the observed annual maxima of wind speed.
First, these extremes are plotted on Gumbel graph paper following the plotting position formula of Benard and Bos-Levenbach (1953), which often is used for wind extremes. This position is given by

$$H(u_r) = \frac{r - 0.3}{N + 0.4} , \tag{3.31}$$

where r equals the ranked position of the annual extreme (from low to high) and N the total number of years in the time series. These positions can be interpolated linearly, so that non-exceedance probabilities of integer values of the wind speed (u_I) between the lowest and highest annual extreme can be calculated. These probabilities are denoted as $H(u_I)$.
Next, Rijkoort calculates the non-exceedance probabilities $P(U_{\max,ij} \leq u_I)$ by ways of (3.21) and (3.26) as follows:

$$P(U_{\max,ij} \leq u_I) = \prod_{i=1}^{12}\prod_{j=1}^{6}\left[\{F_d(u_I \mid i,j)\}^{\delta_{ij}}\{F_n(u_I \mid i,j)\}^{\nu_{ij}}\right] = \prod_{i=1}^{12}\prod_{j=1}^{6} G_{ij}(u_I). \quad \textbf{(3.32)}$$

Now the $q(u_I)$-values can be obtained by

$$\left(\prod_{i=1}^{12}\prod_{j=1}^{6} G_{ij}(u_I)\right)^{\frac{1}{q(u_I)}} = H(u_I), \text{ or } q(u_I) = \frac{\ln\left(\prod_{i=1}^{12}\prod_{j=1}^{6} G_{ij}(u_I)\right)}{\ln H(u_I)}. \quad \textbf{(3.33)}$$

The resulting $q(u_I)$-values can be plotted against u_I. Full correctness of the model would require $q(u_I) \rightarrow 1$ (asymptotic independency between successive hourly wind speed values) for increasing u_I ($u_I \rightarrow \infty$). If the $q(u_I)$-values become significantly smaller than 1 or stay significantly above 1, besides the persistence correction also a tail correction is introduced. It appeared that the $q(u_I)$-values become significantly smaller than 1 what results in higher non-exceedance probabilities and thereby in higher return levels of wind speed.
Values of the persistence correction and tail correction have been determined for the various individual azimuth-season sectors. This results in:

$$G_{ij}(u_I)^{\frac{1}{q_{ij}(u_I)}} = H_{ij}(u_I), \text{ or } q_{ij}(u_I) = \frac{\ln G_{ij}(u_I)}{\ln H_{ij}(u_I)}. \quad \textbf{(3.34)}$$

Furthermore, there appears to be quite some variation in the position of the $q_{ij}(u_I)$-lines between stations: The $q_{ij}(u_I)$-lines of coastal or offshore stations seem to lie on a higher level than inland stations. This is not surprising because one can expect that inland stations reach earlier the level of independence (q=1) than offshore stations as a result of the difference in the average wind speed. This made Rijkoort to take the station mean wind speed into account. Therefore the $q_{ij}(u_I)$-values are plotted against u_I / α_{ij}, where α_{ij} is the scale parameter of the Weibull distribution for the wind speeds in season j with wind direction i. The result from this is that now the position of the $q_{ij}(u_I)$-lines seems to lie on the same level reasonably well.
By trial and error Rijkoort found the following relation between $q_{ij}(u_I)$ and u_I / α_{ij}:

$$\ln q_{ij}(u_I) = A_{ij}\left(\frac{u_I}{\alpha_{ij}}\right)^2 + B_{ij}, \quad \textbf{(3.35)}$$

with $A_{ij} < 0$.
Because Rijkoort found no definite systematic variation between the values of A_{ij} and B_{ij} for different stations, these parameters were averaged over the stations. From these averages and the station values of α_{ij} the persistence/tail factors $q_{ij}(u_I)$ can be calculated, and from there the extreme value distribution for any arbitrary azimuth-season sector (i,j) by (3.23).

Next, the distributions for separate seasons, for separate azimuth sectors, and for the overall (year) maxima can be determined. It proved necessary to introduce additional persistence correction factors that take into account the dependence between neighbouring azimuth sectors and subsequent seasons. For seasonal calculations the mutual dependency between azimuth groups was accounted for by a persistence correction factor 2.0. For azimuthal calculations the mutual dependency of seasons required a persistence correction factor 1.2. For yearly calculations the mutual dependency between azimuth groups and between seasons required a persistence correction factor of again 2.0. Based on these correction factors, (3.24), (3.25) and (3.26) are replaced by

$$G_j(u) = \left(\prod_{i=1}^{12} G_{ij}(u) \right)^{\frac{1}{2.0}}, \qquad \textbf{(3.36)}$$

$$G_i(u) = \left(\prod_{j=1}^{6} G_{ij}(u) \right)^{\frac{1}{1.2}}, \qquad \textbf{(3.37)}$$

$$G(u) = \left(\prod_{i=1}^{12} \prod_{j=1}^{6} G_{ij}(u) \right)^{\frac{1}{2.0}}. \qquad \textbf{(3.38)}$$

4 Reproduction of the Rijkoort-Weibull model

In this chapter step by step the reproduction of the RW-model will be explained (Paragraph 4.1-4.3). This reproduction makes use of the same theory (Chapter 3) and the same station wind observations (Chapter 2) as Rijkoort (1983). Problems that arose during the reproduction are discussed in detail. Finally, the reproduction results (Table B.1, B.2, B.3 and B.4 in Appendix B) are compared with the results of Rijkoort (Table A.1, A.2, A.3 and A.4 in Appendix A) to see what and why certain deviations occurred (Paragraph 4.4).

4.1 The Compound model

In Paragraph 3.1 it was shown that the RW-model requires a split up of the wind speed data according to three criteria: season, azimuth sectors and day- and nighttime for the purpose of fitting a reasonable Weibull distribution. When subdividing between azimuth sectors Rijkoort encountered a problem. This is due to the fact that wind speeds can have a variable wind direction and therefore they can not be placed into one of the twelve azimuth sectors. Furthermore, anemometers have a finite starting speed, so they often can not register weak wind speeds below approximately 2 m/s and therefore they are not able to measure the wind direction reliably. In such circumstances though, these wind speeds have to be distributed over the azimuth sectors like the wind speeds above 2 m/s with a certain wind direction. Rijkoort applied a method which distributes this class over the azimuth sectors in proportion to the amount of data in each azimuth sector.
Rijkoort did not describe how to handle wind speeds above 2 m/s with a variable wind direction. During the reproduction these wind speeds are subdivided by the same method as applied for the wind speed values below 2 m/s. Because the frequency of occurrence of this class is relative small (station average 0.36%) compared with the frequency of occurrence of the class with wind speeds below 2 m/s (station average 14.57%) the method of subdividing of this class won't influence the results in a severe way.

For each azimuth-season sector the wind speed data has been fitted to the Weibull distribution for daytime hours and to the modified Weibull distribution for nighttime hours. The maximum likelihood method was used to estimate the parameters α_{ij}, κ_{ij} and γ_{ij} in (3.4) and (3.5). In this, the parameters α_{ij} and κ_{ij} are set equal for day- and nighttime hours, so the optimal fit has been obtained by combining the Weibull distribution for daytime hours and the modified Weibull distribution for nighttime hours.
Before doing this the data is grouped in classes. The widths of these classes are not defined in the report of Rijkoort but it can be assumed that they each have width one meter per second, except for the lowest class which width is twice as wide because of the reason explained above. Therefore, the classes are of the following form:

$$[0.0,1.9]\ ,\ [2.0,2.9]\ ,\ [3.0,3.9]\ ,\ .\ .\ . \qquad \textbf{(4.1)}$$

Some attention has to go to the fitted values of γ_{ij}. This parameter is called the stability parameter and therefore it indicates the measure of difference between night and day. When there is no difference the parameter is equal to zero and the distributions for day and night are the same. At inland stations, one can expect a larger difference between day and night wind speeds than at coastal or sea stations, because of more stable conditions over sea. This indeed seemed to be the case for example for L.S. Texel, a station about 30 kilometres off the coast. About 50% of the azimuth-season sectors produced a γ-value that was a little smaller than zero, while the remaining 50% of the azimuth-season sectors produced a γ-value that was a little above zero. That outcome is not surprising, because in cases where a parameter equals zero and nevertheless is being estimated, a 50% chance exists to get an parameter estimate that is smaller than zero when this parameter is not bounded theoretically. Therefore, in cases with no difference between day- and nighttime, it might be wise to set the γ-parameter equal to zero.
In his report, Rijkoort did not take notice of this problem. Therefore it is assumed that he did not set any parameter equal to zero, but that he estimated each parameter. For this reason, each parameter is

estimated during the reproduction.
However, whether Rijkoort did set a number of γ-parameters equal to zero or not, the deviation between the results following his procedure and the results following the reproduction will not be of great importance because of the mostly near zero values for the estimates of γ.

4.2 Smoothing of the model parameters

Paragraph 3.2 showed which function is used to subject the wind direction to harmonic analysis. This function is a mixed function of two sine functions with each an amplitude and phase shift which has to be estimated. Furthermore these sine functions have period lengths which are set equal to 1 and ½.
In Rijkoort's report (Rijkoort, 1983), it is not mentioned how this mixed function has been fitted, but it can be assumed that the method described in Rijkoort (1980) is used where for instance function (3.10) is rewritten as

$$\alpha_{ij} = a_{j0} + b_1 \sin 30i + b_2 \cos 30i + b_3 \sin 60i + b_4 \cos 60i\,, \qquad \textbf{(4.2)}$$

where

$$b_1 = a_{j1} \cos a_{j2}\,,$$
$$b_2 = a_{j1} \sin a_{j2}\,,$$
$$b_3 = a_{j3} \cos a_{j4}\,,$$
$$b_4 = a_{j3} \sin a_{j4}\,,$$
$$i = 1,\ldots,12\,.$$

(4.2) can be fitted using the method of multiple least squares after which the estimates of b_1, b_2, b_3 and b_4 can be used to calculate the estimates of a_{j1}, a_{j2}, a_{j3} and a_{j4} in the following way:

$$a_{j2} = \arctan\left(\frac{b_2}{b_1}\right), \qquad \textbf{(4.3)}$$

$$a_{j1} = \frac{b_1}{\cos(a_{j2})}, \qquad \textbf{(4.4)}$$

$$a_{j4} = \arctan\left(\frac{b_4}{b_3}\right), \qquad \textbf{(4.5)}$$

$$a_{j3} = \frac{b_3}{\cos(a_{j4})}. \qquad \textbf{(4.6)}$$

This harmonic analysis reduced the number of parameters from 360 to 150 parameters per station.
A further reduction has also been made. Some points of this reduction are discussed below.

The separate seasonal values of the first harmonic parameter of d_{j0} and n_{j0} are listed in Table 3.4. These values, however, are not equal to the values that are listed in the report of Rijkoort. The values written down in his report are equal to the total number of hours in season j divided by the total number of azimuth sectors (=12), but cumulated over the years that are used. Rijkoort used in his analyses the time period 1962-1976, so his values are a factor 15 times the values in Table 3.4.
Although Rijkoort listed values that are a factor 15 too high, he probably used the correct values in his calculations. This because of using the wrong values would result in unrealistic high return levels of the wind speed, which evidently is not the case is his report.

Another parameter requires some explanation. The RW-model calculates this parameter to reduce the station noise in the seasonal variation of the stability parameter g_{j0} and is defined as follows:

$$g_{j0}^{*} = \frac{1}{S}\sum_{s=1}^{S}\left(\frac{g_{j0}}{g}\right)_{s}, \qquad \textbf{(4.7)}$$

where S equals the number of stations used in the analysis and

$$g = \frac{1}{6}\sum_{j=1}^{6} g_{j0}.$$

(4.7) stands for the station average value of the ratio between the seasonal mean of the stability parameter and the annual mean of the stability parameter. After analysis, it turned out that some stations could not be used for this purpose. L.S. Texel, for instance, has a calculated value for g that is close to zero. This value blows up the fraction in (4.7), so it's not possible to obtain a realistic value for the parameter. The only way to get a value that is realistic, is to average over all stations minus the stations with a g-value close to zero. Rijkoort followed this procedure, what can be deduced from Figure 8 from Rijkoort's report, where the line belonging to station L.S. Texel is missing.
A way to deal with these low values of g could be using linear regression to fit the relation between g_{j0} and g.

In an earlier stage (equation (3.10) to (3.14)) of the reproduction, harmonic analysis was used to reduce the wind direction dependent parameters. For one parameter (a_{j0}) this analysis has been performed for the different seasons also:

$$\frac{a_{j0}}{a} = 1 + c\sin(60j + \chi), \qquad \textbf{(4.8)}$$

where

$$a = \frac{1}{6}\sum_{j=1}^{6} a_{j0} \text{ and } j = 1,\ldots,6.$$

The difference with the earlier harmonic analysis is that the sine structure is unimodal instead of bimodal. Rijkoort stated that the parameter c in (4.8) can be set to a constant (0.11) because of the little variation in the separate c-values of the different stations. The values calculated for the stations are in the range of 0.09 up to 0.13, which indeed is rather constant, with the exception of L.S. Texel where c equals 0.19.
Because of this and because of the earlier demonstrated special behaviour of station L.S. Texel, it can be argued whether offshore stations have to be treated the same as coastal and inland stations. In this stage however, during the reproduction the specified constant is used for each station, also for station L.S. Texel.

The parameters d^* and n^* (see Table B.1 in Appendix B) are calculated by projection on a regression line. These parameters represent the seasonal mean of the normalised amplitudes of the first harmonic of respectively δ_{ij} and ν_{ij}. It is assumed that Rijkoort used normal projection, though this is not mentioned explicitly in his report.

The table with seasonal dependent parameter values in Rijkoort's report (Table 2 on page 19) contains three errors (probably typing errors). The corrected values of this table (and those of the other tables) are copied into Table B.2 in Appendix B).
The first error concerns the parameter D_{11}. The original value of this parameter is 0.10, while it should be the value –0.10. This conclusion is based on two reasons:
First of all an almost perfect correlation exists between the parameter D_{j1} calculated by Rijkoort and calculated by the reproduction, except for D_{11}. When changing this value into –0.10 the correlation coefficient (R^2) increases from 0.7262 to 0.9981.
Furthermore, in Figure 6.4 from Rijkoort's report can be seen that almost all different station values of d_{11}/d_{10} (top of the figure) lie beneath the annual mean, what should have been resulted in a negative value for D_{11}.

The remaining two errors concerns the parameters n_{22}^* and n_{42}^*. The original values of these parameters are respectively 232° and 255°, while it's almost for sure that these values have to be replaced by 132° and 235°.
This is based on Figure 6.5 from Rijkoort's report which shows clearly that the different station values of the March/April component of n_{j2} (second figure from above) lie around 132° instead of 232° and those of the November/December component of n_{j2} around 235° instead of 255°. In this case, when the values are changed into 132° and 235° the correlation coefficient R^2 between the values of n_{j2} calculated by Rijkoort and calculated by the reproduction increases from 0.4702 to 0.9993.

4.3 Determination of persistence correction factors

Like described in Paragraph 3.4, the parameter q is determined with the help of the observed annual wind speed maxima. These values are plotted on Gumbel graph paper after which the observations are interpolated to obtain probabilities for integer values.
During the reproduction, this interpolation was firstly done using the method of linear least-squares regression. This method corresponds with fitting the values to a Gumbel distribution. Doing this however, the variation in the observations is reduced at this stage, while this has to be done at a stage somewhat further following the RW-model.
For this reason it's more likely that Rijkoort interpolated linearly between two successive points instead of using linear regression. In this way, the variation in the observation can be maintained at this stage.

After calculating the q-values and with the help of the parameter estimates calculated before, return periods can be calculated for each station.
Rijkoort noticed a slightly worse fit for some of the stations (Eelde, Deelen and Leeuwarden). For this reason he adjusted the k-values for these stations. However, he did not mention in what manner these adjustments took place so it was not possible during the reproduction to adjust these values in the same way. Therefore, the k-values in Table A.1 and Table B.1 are the original, not adjusted values to keep the comparison between the calculations of Rijkoort and the reproduction objective.

4.4 Comparison between results calculated by Rijkoort and resulting from the reproduction

Because of the points made in this chapter the results of Rijkoort do not need to correspond exactly with the outcomes resulting from the reproduction. Indeed this was not the case. Relatively small differences in the parameters resulted in severe differences for the return levels of wind speed at certain stations.
The deviations between the station dependent parameters are graphically listed in Figure C.1 of Appendix C. Clearly can be seen in this figure that the correlation in the parameters k, d^* and n^* is quite bad. The correlation in the parameters a, χ and g on the contrary is quite good.
Parameter k is the most important parameter of the (modified) Weibull distribution. This parameter handles the shape of the distribution and therefore it is mainly responsible for the width of the tail. The way to obtain an estimate for parameter k is to calculate the mean of the separate smoothed season/azimuth maximum likelihood estimates of κ (κ_{ij}). Because no serious problems arose by doing so, it's not clear how the deviations could become as severe as they are now.
The parameters d^* and n^* don't have much influence in the outcomes of the RW-model, so deviations in the estimates of these parameters can't do much harm. Something can be said though about one station of which the estimates deviate more than those of the other stations. In the lower two plots of Figure C.1 (the ones about d^* and n^*) can be seen that one point (right at the top) shows abnormal behaviour. This point corresponds with station Beek which is the most inland station Rijkoort used in his analysis. Both from the results from Rijkoort and the results from the reproduction can be concluded that the d^* and n^* value of this station is rather high. This implies that some azimuth-season sectors at this station obtain relatively few observations (and other azimuth-season sectors relatively many) compared with other stations. After smoothing, even negative frequencies were calculated for some of the 72 (=12·6) azimuth-season sectors (respectively 8 and 7 for day and night) which evidently is not realistic. To avoid negative frequencies, Rijkoort probably lowered the d^* and n^* values such that negative values didn't occur anymore. These

levels occurred at values of 0.52 for both d^* and n^*: When the values are set on 0.53 one azimuth-season sector gives a negative frequency, but when the values are set on 0.52 the frequency becomes slightly positive. The reproduction of the RW-model does not lower the d^* and n^* values but replaces negative values by 1, because Rijkoort did not describe the procedure to lower d^* and n^*. This explains the deviation between the results from Rijkoort and the results from the reproduction in the case of station Beek.

Next, the estimated seasonal dependent parameters are compared. Graphically, these deviations are listed in Figure C.2 of Appendix C. Two parameters (d_{j0} and n_{j0}) are not included in this figure, because they (like mentioned earlier) are constants and don't have to be estimated. Consequently, no deviations exist between d_{j0} and n_{j0} calculated by Rijkoort and resulting from the reproduction.
Most of the other season dependent parameters show a good (g_{jo}^*, D_{j1}, d_{j2}^*, N_{j1}, n_{j2}^*) or reasonable (a_{j1}, a_{j2}, k_{j0}^*, g_{j1}, d_{j3}, d_{j4}^*, n_{j4}^*) correlation, but quite bad correlations exist in the case of the parameters a_{j3}, a_{j4}, k_{j1}, k_{j2} and g_{j2}. Again, it's not obvious what the causes are of the relative large deviations between the parameters calculated by Rijkoort and resulting from the reproduction.

The deviations in the persistence/tail correction are shown in Figure C.3 of Appendix C. Again, these deviations are quite large. This is not strange, because the values of the persistence/tail parameters are partly based upon the values of the station and season dependent parameters. So, certain deviations in these parameters will result in deviations in the persistence/tail correction parameters too.

Return levels of wind speed can be calculated using the station dependent parameters, the season dependent parameters and the persistence/tail correction parameters. Deviations in these parameters will therefore result in deviations in the return levels. This can be seen in Figure C.4, where the return levels calculated by Rijkoort and resulting from the reproduction are set out against each other. Apparently, the deviations are relatively small in the case of short return periods (2, 5 and 10 years), but they are extremely large for some stations in the case of longer return periods (500 or 1000 year). For example, for station Deelen the reproduction gives a return level of 51.4 m/s for a return period of 1000 years, while Rijkoort calculates a return level of 38.0 m/s. Both values are not realistic (especially the value of 51.4 m/s). The question is of course how these return levels became so high, which will be discussed in Paragraph 5.1. This paragraph evaluates model weaknesses at different steps in the model.

An explanation for the severe deviations in the return levels of wind speeds for high return periods can be the separate deviations in the three parameter tables (Table B.1, B.2 and B.3 in comparison with Table A.1, A.2 and A.3). Together, these deviations have an cumulative and strengthening effect on the return levels. That is why it is sensible to look at the model with at certain points different choices compared with the original RW-model, which makes it possible to isolate certain effects of parameter estimates. This and more properties of the RW-model will be evaluated in the following chapter.

5 Evaluation of the Rijkoort-Weibull model

The results of the reproduction of the RW-model (Table B.1, B.2, B.3 and B.4 in Appendix B) can be used to examine the quality of the model. The reproduction of the RW-model produces intermediate results what makes it possible to examine step by step what effects certain decisions have on the results. In Paragraph 5.1 each of these steps will be discussed. Paragraph 5.2 evaluates the persistence corrections used between neighbouring azimuth sectors and subsequent seasons and the consequences of class definition of wind direction and season will be examined in Paragraph 5.3. Finally, the effect of the length and quality of time series on the results of the RW-model will be discussed in Paragraph 5.4.

5.1 The Rijkoort-Weibull model step by step

The RW-model can be split up into two rough parts: The first part concerns the estimation of the azimuth-season (modified) Weibull parameters while the second part concerns the determination of the persistence correction between successive wind speed values (where this persistence factor turned out to function as a tail correction also). Return levels of wind speed can be estimated without the second part (so without application of the persistence/tail correction), which makes it possible to concentrate on the effects of parameter estimation. The mutual azimuthal and seasonal persistence correction factors (like used in (3.36), (3.37) and (3.38)) will be applied in both cases (without and with the persistence/tail correction) though.

The parameter estimation itself can be split up into a number of parts also. First, each azimuth-season sector is fitted to the modified Weibull distribution, which results in a number of individual Weibull parameters. Next, the individual Weibull parameters are subjected to harmonic analysis, which results in a number of harmonic Weibull parameters. Finally, the harmonic parameters are subjected to further smoothing, which results in a number of smoothed Weibull parameters.

Each step of the parameter estimation can be ran with or without the persistence/tail correction to produce return levels of wind speed, which makes it possible to analyse the effect of each of these extensions of the RW-model.

The persistence/tail correction itself can also be split up into two parts: the first without averaging over the several stations (station dependent persistence/tail correction), the second with averaging over the several stations (station independent persistence/tail correction). The effect of this difference will be analysed as well.

An overview of the separate steps with corresponding model names is given in Table 5.1. In this paragraph will be referred to the names in this table.

parameters	persistence/tail correction		
	no persistence/tail correction	station dependent persistence/tail correction	station independent persistence/tail correction
individual Weibull parameters	model A	model B	model C
harmonic Weibull parameters	model D	model E	model F
smoothed Weibull parameters	model G	model H	RW-model

Table 5.1: Listing of several steps in the RW-model.

5.1.1 Individual Weibull parameters without persistence/tail correction

The most simple model with minimal extensions is the model which makes use of individual (modified) Weibull parameters without implementation of the persistence/tail correction (model A). The resulting return levels of wind speed of this model are listed in Table 5.2.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	28.1	29.2	30.6	31.7	32.7	34.1	35.1	36.1	37.5	38.5
Schiphol	25.5	26.7	28.1	29.2	30.2	31.5	32.5	33.5	34.7	35.6
De Bilt	19.5	20.3	21.4	22.1	22.8	23.8	24.4	25.1	26.0	26.6
Soesterberg	21.6	22.6	23.9	24.9	25.8	27.0	27.9	28.8	30.0	30.8
Leeuwarden	22.2	23.1	24.3	25.1	26.0	27.1	27.9	28.7	29.8	30.6
Deelen	24.4	25.6	27.0	28.1	29.2	30.6	31.6	32.7	34.0	35.0
Eelde	21.6	22.7	24.0	25.0	25.9	27.2	28.1	29.0	30.2	31.0
Vlissingen	23.4	24.4	25.7	26.7	27.6	28.8	29.7	30.6	31.7	32.6
Zestienhoven	23.9	24.9	26.1	26.9	27.8	28.9	29.7	30.4	31.5	32.2
Gilze-Rijen	21.3	22.3	23.5	24.4	25.3	26.4	27.3	28.1	29.1	30.0
Eindhoven	21.7	22.7	23.9	24.8	25.6	26.7	27.6	28.4	29.4	30.2
Beek	20.4	21.3	22.4	23.1	23.9	24.9	25.6	26.3	27.2	27.8

Table 5.2: Return levels in m/s (model A).

At first the results seem reasonably realistic, so a closer look at the results is necessary to obtain a better idea of the quality of the model. A method to validate the outcomes is to compare the return levels with measured wind speeds. The observed annual maximum values of hourly mean wind speeds can be plotted at Gumbel graph paper using the plotting position formula of Benard and Bos-Levenbach (1953), described in Paragraph 3.4. The values in Table 5.2 can also be plotted in this figure. The deviations between the annual maxima and the model output can be interpreted as a measure of quality of the model. An example of this presentation is Figure 5.1. A complete list of this presentation for all models and all stations is placed into Appendix D.

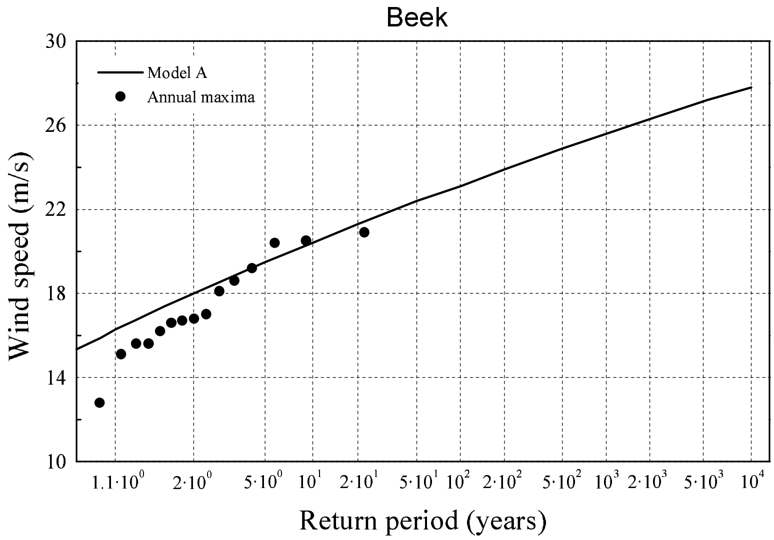

Figure 5.1: Comparison between model A and annual maxima for station Beek.

Figure 5.1 shows clearly the relative high levels of wind speeds for low return periods following model A. This deviation seems to disappear for higher return periods. A similar pattern occurs in the case of the other stations (Figure D.1 in Appendix D), but this is not surprising because there is no correction for the persistence in the time series. Each exceedance of a certain wind speed level is now considered as one event while a number of exceedances within a storm has to be seen as one event (Paragraph 3.3). Because the persistence will decrease gradually at higher wind speed levels, especially for relative low levels of wind speed model A will produce return periods that are too short.

5.1.2 Individual Weibull parameters with station dependent persistence/tail correction

In this model (model B) a list of persistence/tail parameters like Table A.3 in Appendix A and Table B.3 in Appendix B has been determined for each station, so without averaging these values over the several stations (like done in the RW-model).
The results from this model are listed in Table 5.3.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	27.7	29.3	31.5	33.1	34.8	37.2	39.0	40.8	43.2	44.9
Schiphol	25.2	26.7	28.7	30.2	31.7	33.7	35.2	36.6	38.5	39.9
De Bilt	21.0	22.3	24.0	25.2	26.4	27.9	29.0	30.1	31.5	32.6
Soesterberg	22.1	23.7	25.9	27.5	29.0	31.1	32.6	34.1	36.1	37.7
Leeuwarden	24.4	26.0	27.9	29.4	30.8	32.7	34.0	35.4	37.2	38.5
Deelen	24.8	26.5	28.7	30.3	31.8	33.9	35.4	37.0	39.0	40.5
Eelde	22.9	24.7	27.0	28.7	30.3	32.5	34.1	35.6	37.7	39.2
Vlissingen	23.2	24.6	26.5	27.9	29.2	31.1	32.4	33.8	35.5	36.8
Zestienhoven	24.0	25.6	27.5	28.9	30.2	31.9	33.2	34.4	35.9	37.0
Gilze-Rijen	22.5	24.0	25.9	27.4	28.8	30.7	32.2	33.7	35.6	37.1
Eindhoven	22.5	23.9	25.8	27.2	28.7	30.6	32.1	33.6	35.7	37.2
Beek	20.8	22.1	23.7	24.9	26.0	27.5	28.5	29.5	30.9	31.9

Table 5.3: Return levels in m/s (model B).

To evaluate the values in Table 5.3, again the results are presented graphically like in Subparagraph 5.1.1 (Figure D.2 in Appendix D). The example below for station Beek (Figure 5.2) demonstrates the effect of the persistence/tail correction.

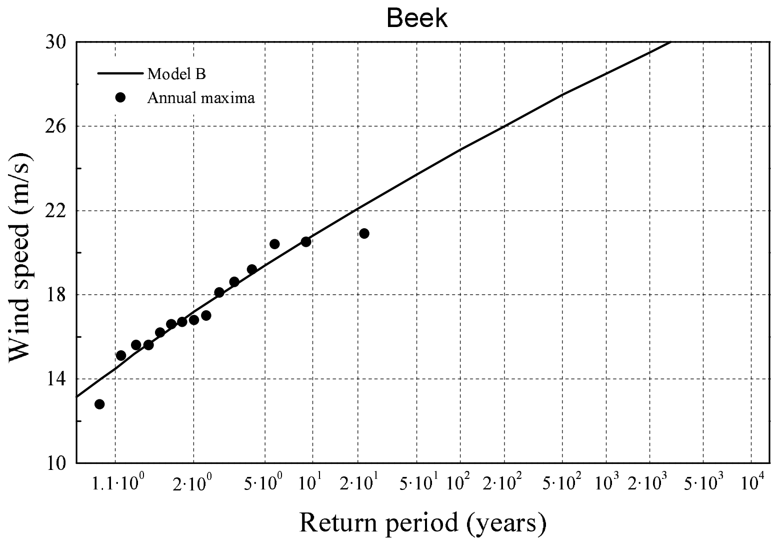

Figure 5.2: Comparison between model B and annual maxima for station Beek.

The return periods seem no longer to be too short for relative low return levels. This effect appears also for the several other stations (Figure D.2 in Appendix D). But when Figure D.2 is compared to Figure D.1 more can be noticed. Like stated earlier, the persistence correction also seems to function as a tail correction. All stations show the same effect of higher return levels of wind speed in the tail of the distribution (at high return periods).
The return level corresponding with a return period of 10,000 years increases at all stations rather strong (from 4.1 m/s at station Beek up to 8.2 m/s at station Eelde). This behaviour can be explained by taking a closer look at the q-factors described in Paragraph 3.4. From all the individual azimuth-season q-factors that could be calculated (12 stations · 12 wind direction · 6 seasons = 864) 47.2% of them were smaller than 1 at the highest azimuth-season wind speed at each station. Because the q-factor is defined in a way

that the factor will decrease at higher wind speeds the percentage of them lower than 1 will increase at higher wind speeds. At the level of wind speed corresponding with a return period of 10,000 year the percentage is increased to 79.3. This percentage is rather constant for all stations considered.
Based on this it's not surprising that the return levels resulting from model B are rather high for all stations in comparison with the outcomes of model A.

5.1.3 Individual Weibull parameters with station independent persistence/tail correction

Model B can be extended to the model where the persistence/tail correction parameters are averaged over the stations (like done in the RW-model). The results from this model (model C) are listed in Table 5.4.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	29.3	31.2	33.8	35.8	37.8	40.5	42.4	44.2	46.7	48.4
Schiphol	25.8	27.5	29.6	31.2	32.7	34.7	36.2	37.6	39.4	40.8
De Bilt	20.2	21.6	23.2	24.4	25.6	27.1	28.2	29.2	30.6	31.6
Soesterberg	28.0	32.0	39.3	52.6	*	*	*	*	*	*
Leeuwarden	23.8	25.5	28.0	30.0	32.1	35.1	37.4	39.7	42.7	45.0
Deelen	29.0	31.5	35.0	37.8	40.8	45.1	48.6	52.5	*	*
Eelde	23.0	24.9	27.2	29.0	30.7	33.0	34.7	36.4	38.5	40.1
Vlissingen	23.5	25.1	27.4	29.1	30.9	33.2	34.9	36.5	38.6	40.1
Zestienhoven	25.1	26.5	28.4	29.7	31.0	32.7	33.9	35.1	36.7	37.8
Gilze-Rijen	22.8	24.5	26.7	28.3	29.9	31.9	33.4	34.8	36.7	38.1
Eindhoven	23.4	25.1	27.4	29.1	30.8	33.1	34.9	36.6	39.0	40.8
Beek	21.3	22.8	24.8	26.2	27.6	29.4	30.6	31.9	33.5	34.6

Table 5.4: Return levels in m/s (model C).

It's evident that the values in Table 5.4 are generally higher than the values in Table 5.2 and Table 5.3. The return levels for the stations Deelen and Soesterberg are even that high at certain return periods that it was not possible to calculate these values in a proper way. These values have been noted with a star (*). With the help of Figure D.3 in Appendix D it becomes clear that also the return levels of the stations L.S. Texel (Figure 5.3) and Leeuwarden have become much higher. Only the calculated 10,000 year return levels for the stations De Bilt and Gilze-Rijen are lower in comparison with model B.
The return levels concerning Soesterberg and Deelen will be discussed later in this chapter, for now we will concentrate on the return levels of wind speed calculated for station L.S. Texel. For this station, observed annual maxima are compared with model values in Figure 5.3.

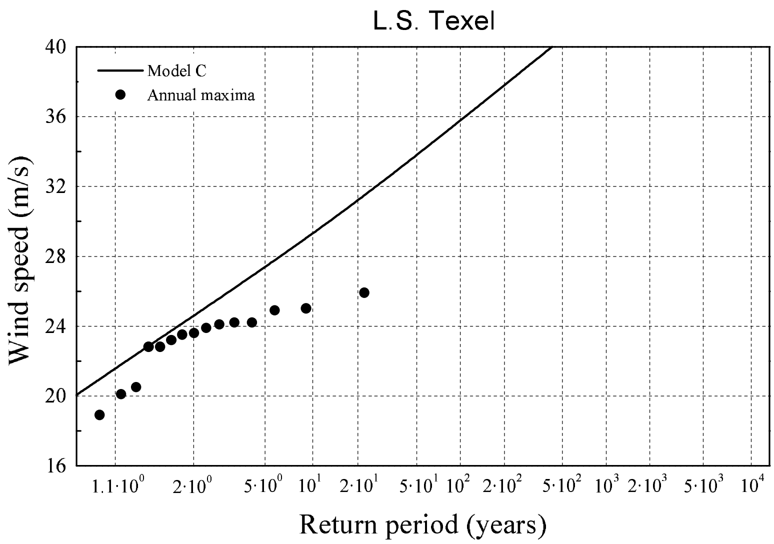

Figure 5.3: Comparison between model C and annual maxima for station L.S. Texel.

The return levels listed are based on the several (12·6=72) individual azimuth-season sectors (see (3.38)). These sectors are combined in a way that the resulting return levels are mostly higher than the maximum return levels of each individual azimuth-season sector ("mostly" because of the constant factors used for correcting dependency between neighbouring azimuth sectors and subsequent seasons). So when one certain azimuth-season sector produces unrealistic high return levels, the final return level will probably be unrealistically high as well. It's therefore sensible to look at separate return levels for the 72 azimuth-season sectors. 10,000 year return levels for each azimuth-season sector resulting from model C are listed in Table 5.5 for L.S. Texel.

Azimuth	Season					
	1	2	3	4	5	6
1	27.6	23.7	18.0	17.3	19.0	30.4
2	26.7	27.5	17.1	18.0	19.8	28.1
3	30.3	28.7	15.1	15.0	18.2	29.5
4	25.7	**47.8**	14.4	16.7	18.6	23.2
5	19.1	18.4	13.7	22.6	25.4	23.1
6	28.6	25.4	20.5	23.0	26.5	28.5
7	32.1	26.2	24.6	27.0	33.4	30.6
8	33.0	24.7	24.4	23.7	32.0	31.6
9	36.6	31.2	23.4	27.1	30.2	32.4
10	42.3	31.9	26.4	28.7	34.1	38.3
11	40.5	33.0	26.4	27.8	34.2	34.5
12	29.4	25.9	19.3	21.8	34.2	32.3

Table 5.5: 10,000 year return levels for L.S. Texel following model C.

As can be seen from Table 5.5, only one of the azimuth-season sectors produces an unrealistic return level of wind speed. This concerns sector (4,2) (the first number represents the azimuth, the second number represents the season). Azimuth 4 (between east and south) in combination with season 2 (March/April) is not a sector with a high frequency of strong winds, so a resulting 10,000 year return level of 47.8 is very surprising. If we look at the list of κ-values, we notice that this azimuth-season sector produces the lowest Weibull κ-value (1.90) as can be seen in Table 5.6.

Azimuth	Season					
	1	2	3	4	5	6
1	2.24	2.36	2.84	2.80	2.70	2.19
2	2.48	2.37	2.84	2.75	2.61	2.30
3	2.33	2.25	3.02	3.01	2.71	2.34
4	2.57	**1.90**	2.84	2.71	2.71	2.68
5	2.95	2.65	2.78	2.36	2.23	2.61
6	2.35	2.22	2.44	2.38	2.30	2.48
7	2.32	2.35	2.40	2.36	2.07	2.58
8	2.19	2.49	2.36	2.49	2.26	2.49
9	2.18	2.29	2.37	2.35	2.33	2.55
10	2.02	2.31	2.30	2.33	2.12	2.32
11	2.08	2.31	2.33	2.43	2.14	2.39
12	2.29	2.44	2.65	2.52	2.12	2.58

Table 5.6: Individual Weibull κ-parameter following model A, B and C.

Partly this is not surprising because the value of κ determines the length of the tail. Normally, when this κ-value is low the tail of the distribution is long which results in relative high return levels and when this κ-value is high the tail of the distribution is short which results in relative low return levels. Figure 5.4 shows clearly the difference in wind speed distribution between azimuth-season sectors with low κ-values and azimuth-season sectors with relative high κ-values.

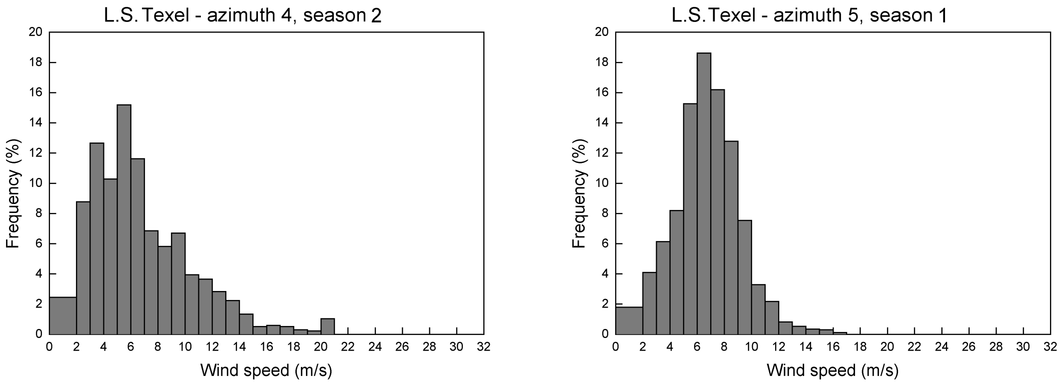

Figure 5.4: Histograms of Texel. The left histogram corresponds with azimuth-season sector (4,2) with a low κ-value (1.90). The right histogram corresponds with azimuth-season sector (5,1) with a κ-value that is relatively high (2.95).

However, this correlation between the value of κ and the values of the return values can't explain the extreme value for sector (4,2) in a direct way. Therefore it's possible that this low κ-value is indirectly responsible for the unrealistic high return value. What we do know at this point, is that the station independent persistence/tail correction plays an important role in it, because after applying the station dependent persistence/tail correction (model B) no azimuth-season sectors produced unrealistic return levels of wind speed (azimuth-season sector (4,2) produces a 10,000 year return level of 31.2 m/s following model B).
For a better understanding why the impact of the station independent persistence/tail correction is so strong, it's necessary to take a better look at its definition. This persistence/tail correction consists of three parts:

- The first part of the correction determines the relation between the annual extremes and the model without persistence/tail correction factor (in this case model A) per azimuth-season sector:

$$G_{ij}(u_I)^{\frac{1}{q_{ij}(u_I)}} = H_{ij}(u_I), \text{ or } q_{ij}(u_I) = \frac{\ln G_{ij}(u_I)}{\ln H_{ij}(u_I)}. \qquad \textbf{(3.34)}$$

- The second part of the persistence/tail correction draws a smoothed curve through these $q_{ij}(u_I)$-lines by least squares regression:

$$\ln q_{ij}(u_I) = A_{ij}\left(\frac{u_I}{\alpha}\right)_{ij}^{2} + B_{ij}, \qquad \textbf{(3.35)}$$

  where B_{ij} corresponds with the persistence parameter and A_{ij} with the tail parameter.

  (3.35) tries to bring the separate $q_{ij}(u_I)$-lines for each station to the same level by dividing by the location parameter (α_{ij}) of the (modified) Weibull distributions. The purpose of this division by the location parameter (α_{ij}) of the (modified) Weibull distribution is to make it possible to average over the separate station values of A_{ij} and B_{ij}, what should reduce the noise produced by the several stations.

- The final part is the averaging over the separate station values of A_{ij} and B_{ij} for the reasons described above. If this averaging is not applied the resulting model equals model B, so the difference between model B and model C is the averaging over the station values of A_{ij} and B_{ij}.

To do the final part of the the persistence/tail correction properly, Rijkoort divided in the second part the wind speed values by the location parameter (α_{ij}) of the (modified) Weibull distribution. Without this

division, the separate station values of A_{ij} and B_{ij} are not comparable, because stations off shore (with in general high wind speeds) produce q-factors that converge at a higher level of wind speed than inland stations where the wind speed is less strong generally. So, when no division is applied into the second part of the definition one can expect to get structural differences between the separate station values of A_{ij} and B_{ij}.
Rijkoort "solved" this problem by introducing the division by the location parameter (α_{ij}) of the (modified) Weibull distribution. However, after doing this the problem has not been solved in a proper way. For low return periods the division by α_{ij} may be sufficient but for higher return periods the impact of the k-value of the (modified) Weibull distribution becomes stronger and stronger. Figure 5.5, where 10,000 return levels are presented as a function of α en κ, makes the above more clear.

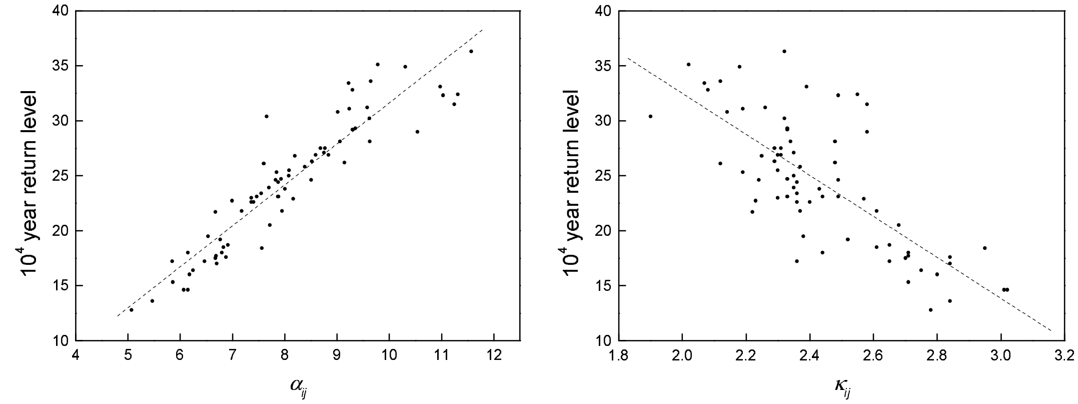

Figure 5.5: Scatterplots of α (left) and κ (right) against corresponding 10,000 year return levels for each azimuth-season sector for station L.S. Texel resulting from model A The dashed lines represent the regression lines.

Figure 5.5 shows that besides the α_{ij}-parameters also the κ_{ij}-parameters seem to play an important role in determining 10,000 year return levels following model A. This is not surprising, but the figure can be misleading though. If the α_{ij}-parameters and the κ_{ij}-parameters are strongly correlated, the additional explanation power on the 10,000 year return levels of the κ_{ij}-parameters will be less than could be interpreted from the right part of Figure 5.5. The correlation between these two parameters in the case of station L.S. Texel is not very strong though ($R^2 = 0.21$), so the level of the κ_{ij}-parameters indeed plays an important role in determining the 10,000 year return levels following model A.

The above explains that structural deviations between the separate station values of A_{ij} and B_{ij} won't disappear after dividing by the location parameters α_{ij}. As a result of this azimuth-season sectors with a relative low κ-value in comparison with it's α-value will produce return levels that are too high and azimuth-season sectors with a relative high κ-value in comparison with it's α-value will produce return levels that are too low.
An overview of what this means for azimuth-season sector (4,2) at station L.S. Texel is listed in Table 5.7.

	model A	model B	model C
A_{42}	0.000	-0.247	-0.555
B_{42}	0.00	3.39	2.95
α_{42}	7.655	7.655	7.655
$q_{42}(30.4)$	1.0000	0.6033	0.0031
$T_{42}(30.4)$	10000	6033	31

Table 5.7: Example of impact of station averaging of persistence/tail correction on return level.

Table 5.7 makes clear that the change in values of A_{42} and B_{42} in model C in comparison with model B has far-reaching consequences. If we calculate the 10,000 year return level following model C, the return level increases a lot, because the corresponding q-value will be even less than 0.0031.

5.1.4 Harmonic Weibull parameters without persistence/tail correction

We return to the model without persistence/tail correction (model A) but subject the azimuth dependent Weibull parameters to harmonic analysis (model D). By this, the individual Weibull parameters change to harmonic Weibull parameters. The results of this model are listed in Table 5.8.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	28.0	29.0	30.3	31.3	32.2	33.4	34.3	35.2	36.4	37.3
Schiphol	25.7	26.8	28.2	29.2	30.2	31.5	32.4	33.3	34.4	35.3
De Bilt	19.8	20.6	21.7	22.5	23.3	24.2	25.0	25.7	26.6	27.2
Soesterberg	22.3	23.4	24.9	25.9	27.0	28.3	29.3	30.2	31.5	32.4
Leeuwarden	22.8	23.9	25.1	26.1	27.0	28.2	29.1	29.9	31.1	31.9
Deelen	25.0	26.2	27.7	28.8	29.8	31.2	32.2	33.2	34.5	35.5
Eelde	21.8	22.8	24.1	25.0	25.9	27.1	27.9	28.8	29.8	30.6
Vlissingen	23.0	23.9	25.1	25.9	26.8	27.8	28.6	29.4	30.4	31.1
Zestienhoven	24.1	25.1	26.4	27.3	28.2	29.3	30.1	30.9	32.0	32.8
Gilze-Rijen	22.4	23.4	24.7	25.6	26.5	27.7	28.6	29.4	30.5	31.3
Eindhoven	22.6	23.7	24.9	25.9	26.8	27.9	28.8	29.6	30.7	31.5
Beek	20.6	21.6	22.7	23.6	24.4	25.4	26.2	27.0	28.0	28.7

Table 5.8: Return levels in m/s (model D).

The results have been set out graphically in Figure D.4 in Appendix D. As example, in Figure 5.6 again the results for station Beek are compared to observed annual maxima.

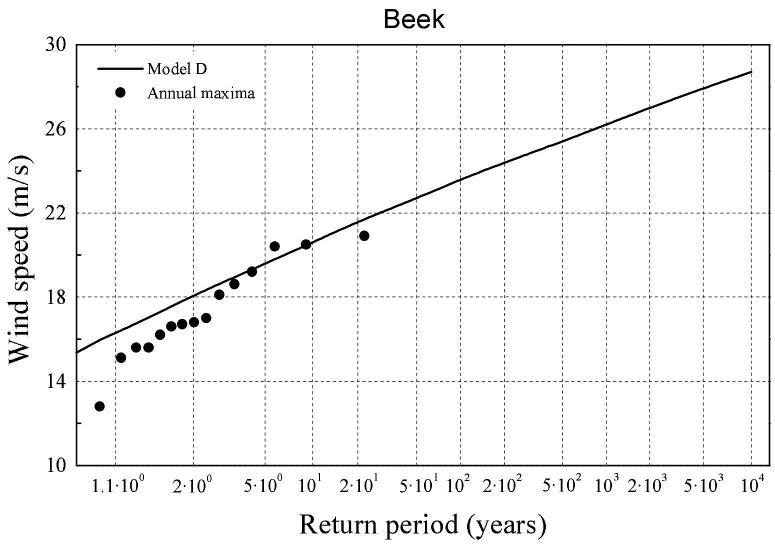

Figure 5.6: Comparison between model D and annual maxima for station Beek.

As can be seen in Figure D.4 in Appendix D and Figure 5.6 the results are very similar to those resulting from model A. The harmonic analysis has been performed to reduce the noise produced by the wind directions. However, this reduction doesn't seem to improve the results. This can be explained by the fact that each Weibull parameter has been corrected independently from the other parameters. This disadvantage could neutralise the advantage of reducing the noise between the wind directions.

5.1.5 Harmonic Weibull parameters with station dependent persistence/tail correction

Model D has been extended with the station dependent persistence/tail correction (model E). The results of this model are listed in Table 5.9 and in Figure D.5 in Appendix D. For station Beek the comparison to observed annual maxima is presented in Figure 5.7.

Based on Figure 5.7 and Figure D.5 in Appendix D it becomes clear that even after applying the station dependent persistence/tail correction the changes in return levels compared with the model with individual Weibull parameters (model B) are rather small. So, at this stage of the RW-model, harmonic analysis of the wind direction doesn't seem to improve the results.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	27.7	29.3	31.3	32.9	34.5	36.5	38.0	39.5	41.4	42.8
Schiphol	25.2	26.7	28.7	30.2	31.7	33.6	35.0	36.3	38.1	39.4
De Bilt	20.9	22.3	24.0	25.2	26.4	27.9	29.1	30.2	31.6	32.7
Soesterberg	22.1	23.7	25.9	27.5	29.1	31.2	32.7	34.2	36.3	37.8
Leeuwarden	24.3	26.0	28.0	29.5	30.9	32.8	34.2	35.6	37.3	38.6
Deelen	24.8	26.5	28.7	30.3	31.9	34.0	35.6	37.2	39.3	40.8
Eelde	23.0	24.8	27.1	28.8	30.5	32.7	34.3	35.8	37.8	39.3
Vlissingen	23.2	24.7	26.5	27.9	29.2	31.0	32.3	33.6	35.3	36.5
Zestienhoven	24.0	25.5	27.4	28.8	30.1	31.7	32.9	34.1	35.6	36.7
Gilze-Rijen	22.4	23.9	25.8	27.2	28.6	30.3	31.7	33.0	34.7	35.9
Eindhoven	22.4	23.8	25.6	27.0	28.3	30.0	31.2	32.5	34.1	35.3
Beek	20.9	22.2	23.9	25.1	26.3	27.8	28.9	30.1	31.5	32.6

Table 5.9: Return levels in m/s (model E).

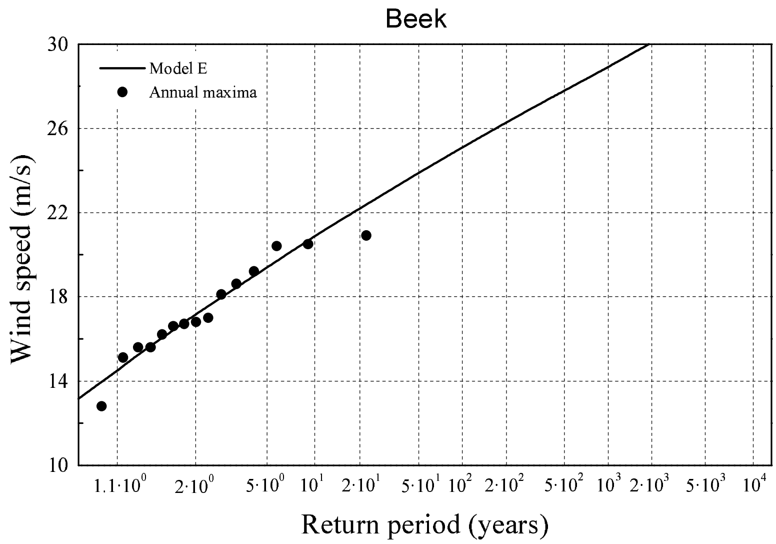

Figure 5.7: Comparison between model E and annual maxima for station Beek.

5.1.6 Harmonic Weibull parameters with station independent persistence/tail correction

In this subparagraph the station dependent persistence/tail corrections (model E) have been replaced by the station independent persistence/tail corrections (model F). The results of this model are given in Table 5.10 and Figure D.6 in Appendix D.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	28.4	29.9	31.8	33.2	34.6	36.4	37.8	39.1	40.8	42.0
Schiphol	25.2	26.7	28.6	29.9	31.3	32.9	34.1	35.3	36.8	37.9
De Bilt	20.2	21.4	23.0	24.2	25.3	26.8	27.8	28.9	30.2	31.2
Soesterberg	26.3	29.1	32.8	35.7	38.8	43.1	46.8	51.0	*	*
Leeuwarden	23.4	24.9	27.0	28.5	30.0	31.9	33.3	34.7	36.6	37.9
Deelen	28.4	30.5	33.3	35.4	37.5	40.4	42.6	44.8	47.7	49.8
Eelde	22.7	24.4	26.6	28.2	29.7	31.7	33.1	34.5	36.3	37.7
Vlissingen	22.4	23.7	25.4	26.7	27.9	29.5	30.6	31.8	33.3	34.5
Zestienhoven	24.7	26.0	27.8	29.0	30.2	31.7	32.8	33.9	35.3	36.4
Gilze-Rijen	22.5	24.1	26.2	27.8	29.3	31.3	32.7	34.2	36.0	37.4
Eindhoven	23.1	24.7	26.7	28.2	29.8	31.7	33.2	34.6	36.4	37.8
Beek	20.6	22.0	23.8	25.2	26.5	28.3	29.6	30.9	32.6	33.9

Table 5.10: Return levels in m/s (model F).

As with model C here an example is given for station L.S. Texel (Figure 5.8).

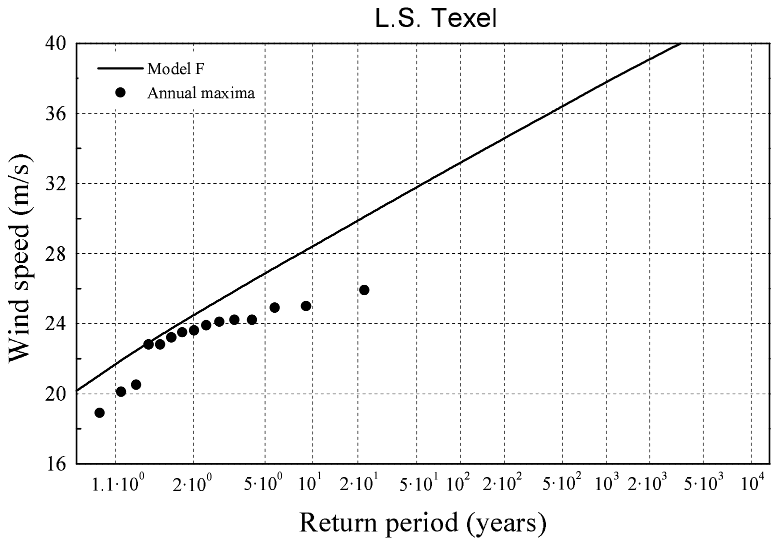

Figure 5.8: Comparison between model F and annual maxima for station L.S. Texel.

When we compare Figure 5.8 with Figure 5.3 it becomes clear that model F produces better results than model C in the case of station L.S. Texel. This can be explained as follows:
As discussed in Subparagraph 5.1.3, model C performs very badly as a consequence of the occurrence of some relatively low *k*-values. For station L.S. Texel this concerned sector (4,2) with a *k*-value of 1.90. But as can be seen in Table 5.6 this value is an outlier compared with the values of the other wind directions in season 2. As show in Table 5.11 and Figure 5.9, with the help of harmonic analysis this value could be smoothed.

Azimuth	Season					
	1	2	3	4	5	6
1	2.28	2.40	2.78	2.75	2.49	2.30
2	2.37	2.30	2.94	2.91	2.73	2.31
3	2.49	2.20	2.98	2.88	2.77	2.41
4	2.61	**2.19**	2.87	2.68	2.59	2.54
5	2.65	2.27	2.69	2.46	2.33	2.61
6	2.56	2.39	2.52	2.35	2.18	2.60
7	2.36	2.43	2.40	2.37	2.18	2.53
8	2.15	2.38	2.34	2.42	2.25	2.47
9	2.03	2.31	2.32	2.41	2.25	2.46
10	2.04	2.30	2.33	2.36	2.17	2.46
11	2.12	2.35	2.41	2.38	2.12	2.44
12	2.20	2.41	2.58	2.52	2.24	2.37

Table 5.11: Harmonic Weibull κ-parameter following model D, E and F.

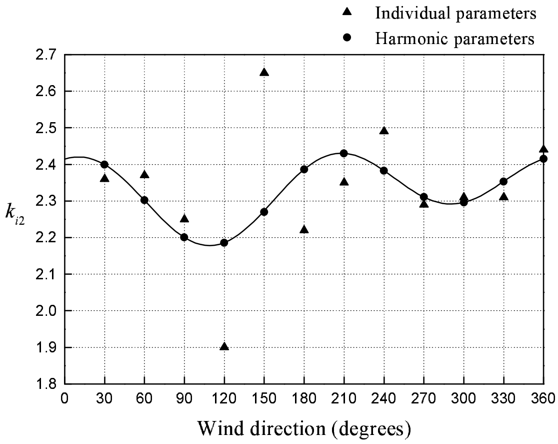

Figure 5.9: Comparison between individual (model A, B and C) and harmonic (model D, E and F) parameters for station L.S. Texel in season 2 (March/April) .

But if we take a close look at Figure D.6 in Appendix D we see that the results of Soesterberg and Deelen are still unsatisfactory. Apparently the harmonic analysis doesn't smooth the low k-values for these stations. If we look at the seasons that are responsible for the unrealistic high return levels (season 1 for both stations) then we see that there are no wind directions that are outliers, but that for these station the low k-values are structural for several azimuth sectors (Figure 5.10).

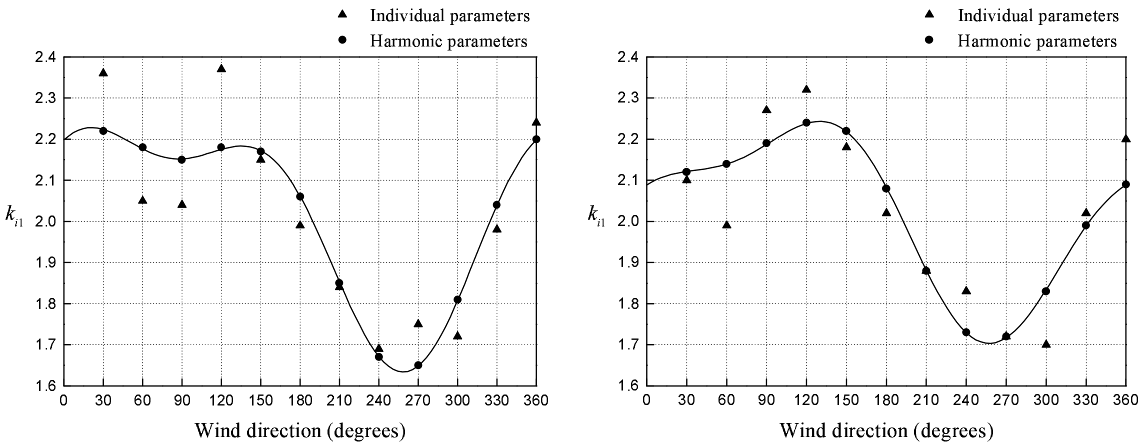

Figure 5.10: Comparison between individual (model A, B and C) and harmonic (model D, E and F) parameters for the stations Soesterberg (left) and Deelen (right) in season 1 (January/February).

The reason for this structural behaviour of low k-values is the presence of too many zero values of wind speed for these stations as a consequence of the finite starting speed of anemometers at some stations, especially in the past. As a result of this the hourly mean values of wind speed are often too low. If we look at the percentage of wind speed values between 0 and 1 m/s ([0,1)) for each station (Table 5.12), then we see that this percentage for the stations Soesterberg and Deelen is rather high.

Station	% [0,1)
L.S. Texel	2.7
Schiphol	4.1
De Bilt	6.0
Soesterberg	12.4
Leeuwarden	2.6
Deelen	10.7
Eelde	3.0
Vlissingen	2.5
Zestienhoven	6.4
Gilze-Rijen	6.6
Eindhoven	7.6
Beek	3.2

Table 5.12: Percentages of wind speed between 0 and 1 m/s.

The effect of these percentages on the wind speed histograms is visualised in Figure 5.11, where Soesterberg and Deelen are compared to Beek.

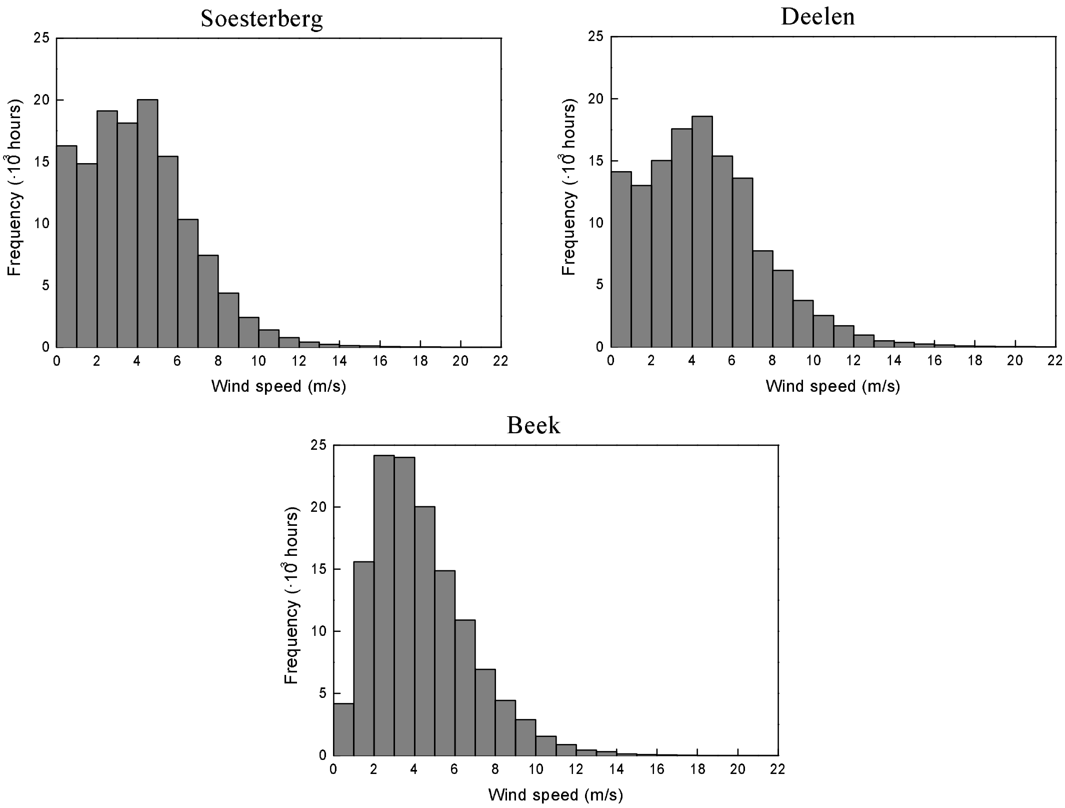

Figure 5.11: Histograms of stations with a relative high frequency of low wind speeds (left above and right above) and a histogram of a station with a normal frequency of low wind speeds (bottom).

A relative high frequency of low wind speeds results in low κ-values of the Weibull distribution. Because a low κ-value implies a heavy tail this seems a bit strange. How can too many low values result in a tail that is too heavy? The answer to this question is the increasing standard deviation of the distribution. The

mean distance to the mean of the distribution will increase and with this the standard deviation. As a result the κ-value will decrease and the tail-width will increase.
If this behaviour is structural, the harmonic analysis will not correct for it and the model will produce unrealistic return levels of wind speed. This is the case for the stations Soesterberg and Deelen.
The conclusion of this is again that relative low κ-values in combination with the station independent persistence/tail correction will result in unrealistic return levels of wind speed.

5.1.7 Smoothed Weibull parameters without persistence/tail correction

The harmonic analysis of the parameters of the Weibull distribution (model F) can be subjected to further smoothing. Without applying the persistence/tail correction (model G) this will result in the return levels of wind speed listed in Table 5.13.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	28.9	30.0	31.4	32.5	33.5	34.7	35.7	36.6	37.7	38.6
Schiphol	23.5	24.4	25.7	26.6	27.4	28.5	29.3	30.1	31.1	31.9
De Bilt	20.2	21.1	22.2	23.0	23.8	24.8	25.5	26.2	27.1	27.8
Soesterberg	21.8	22.8	24.1	25.0	25.9	27.1	27.9	28.8	29.9	30.7
Leeuwarden	22.4	23.3	24.4	25.2	26.0	27.0	27.8	28.5	29.4	30.1
Deelen	25.3	26.5	28.0	29.1	30.2	31.6	32.6	33.7	35.0	35.9
Eelde	20.5	21.3	22.4	23.2	23.9	24.9	25.6	26.2	27.1	27.8
Vlissingen	22.7	23.6	24.8	25.6	26.4	27.4	28.2	28.9	29.9	30.6
Zestienhoven	24.4	25.5	26.8	27.8	28.7	29.9	30.8	31.7	32.8	33.6
Gilze-Rijen	21.2	22.1	23.2	24.0	24.8	25.8	26.5	27.2	28.2	28.8
Eindhoven	22.0	22.9	24.1	25.0	25.8	26.9	27.7	28.5	29.5	30.3
Beek	18.8	19.6	20.5	21.2	21.8	22.7	23.3	23.9	24.7	25.3

Table 5.13: Return levels in m/s (model G).

Comparison with observed annual maxima has been performed again which can be found in Figure D.7 in Appendix D. One station from this figure has been picked out in Figure 5.12.

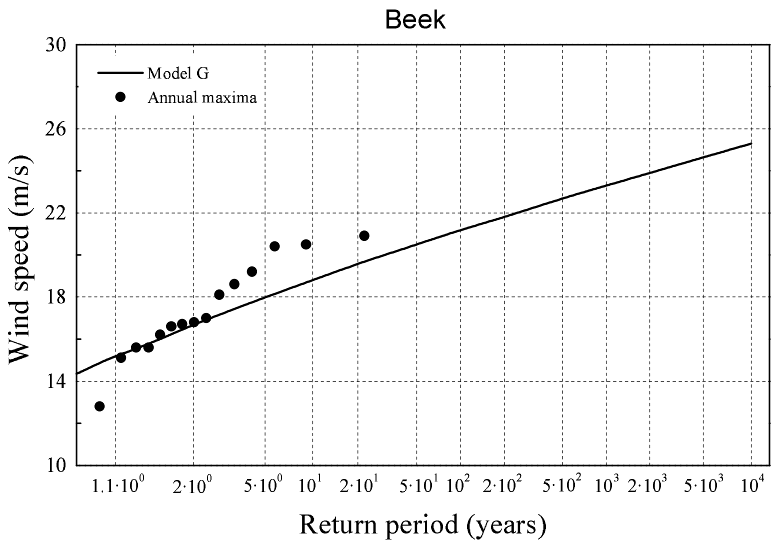

Figure 5.12: Comparison between model G and annual maxima for station Beek.

The return levels for station Beek are significantly smaller following model G than following model A or D which are both without persistence/tail correction. A possible reason for this can be the problem encountered in Paragraph 4.4. In this paragraph was shown that Rijkoort probably lowered the d^* and n^* parameter values for station Beek such that negative frequencies didn't occur anymore.

The other stations show no strongly deviations with model A or D.

5.1.8 Smoothed Weibull parameters with station dependent persistence/tail correction

Model G can be extended with the station dependent persistence/tail correction (model H). The results of this model are listed in Table 5.14.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	28.1	29.8	31.8	33.4	34.9	36.8	38.3	39.7	41.5	42.8
Schiphol	24.9	26.4	28.3	29.8	31.2	33.0	34.4	35.7	37.4	38.7
De Bilt	21.0	22.3	24.0	25.2	26.5	28.0	29.2	30.3	31.8	33.0
Soesterberg	21.9	23.5	25.5	26.9	28.4	30.2	31.6	33.0	34.7	36.0
Leeuwarden	24.0	25.4	27.2	28.5	29.7	31.3	32.4	33.5	34.9	35.9
Deelen	24.7	26.5	28.7	30.4	32.0	34.2	35.8	37.4	39.6	41.2
Eelde	22.6	24.1	26.1	27.5	28.9	30.6	31.9	33.2	34.7	35.9
Vlissingen	23.2	24.6	26.5	27.8	29.1	30.8	32.0	33.2	34.7	35.8
Zestienhoven	24.0	25.5	27.4	28.8	30.1	31.9	33.2	34.5	36.1	37.3
Gilze-Rijen	22.1	23.4	25.1	26.3	27.5	29.1	30.2	31.4	32.8	33.9
Eindhoven	22.4	23.8	25.6	26.9	28.2	29.9	31.1	32.4	34.0	35.2
Beek	20.5	21.7	23.2	24.2	25.2	26.5	27.5	28.4	29.5	30.4

Table 5.14: Return levels in m/s (model H).

Again, Figure D.8 in Appendix D compares the model results with annual maxima. For station Beek this yields Figure 5.13.

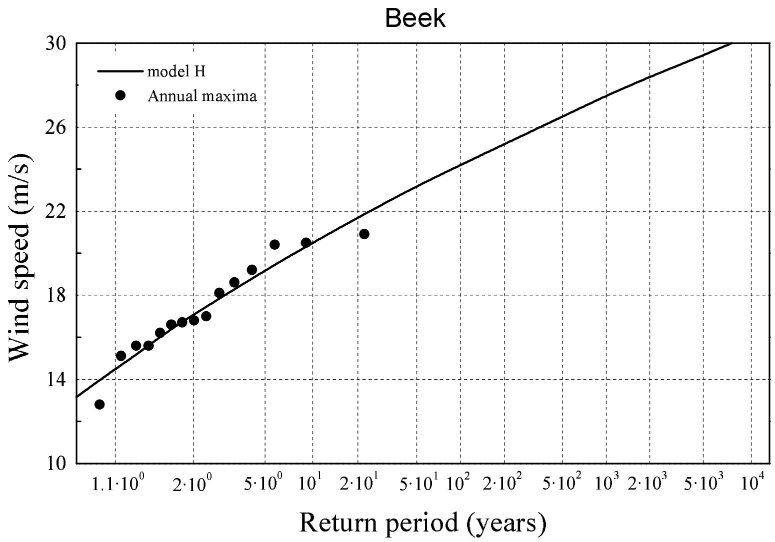

Figure 5.13: Comparison between model H and annual maxima for station Beek.

The results for the several stations following model H show no severe deviations compared with model E, so there's no indication that smoothing the parameters improves the model substantially.

5.1.9 The Rijkoort-Weibull model

Model H can be extended to the model where the persistence/tail correction parameters are averaged over the stations. This model equals the final RW-model. The outcomes of this model are listed in Table 5.15 which are equal to those listed in Table B.4 in Appendix B.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	28.7	30.3	32.3	33.7	35.1	36.8	38.1	39.4	40.9	42.1
Schiphol	23.4	24.8	26.6	27.9	29.1	30.7	31.9	33.0	34.5	35.5
De Bilt	20.7	22.0	23.8	25.0	26.3	27.9	29.0	30.2	31.7	32.7
Soesterberg	24.7	26.6	29.1	31.0	32.9	35.4	37.2	39.0	41.4	43.0
Leeuwarden	21.8	23.0	24.6	25.7	26.7	28.1	29.1	30.1	31.3	32.2
Deelen	31.2	34.2	38.3	41.5	44.8	49.4	51.4	54.5	55.9	57.5
Eelde	20.2	21.4	22.9	24.0	25.0	26.4	27.4	28.4	29.6	30.5
Vlissingen	22.6	23.9	25.6	26.8	28.0	29.5	30.6	31.7	33.1	34.1
Zestienhoven	25.4	27.1	29.2	30.8	32.3	34.2	35.7	37.1	38.9	40.2
Gilze-Rijen	21.2	22.5	24.1	25.3	26.5	28.0	29.1	30.1	31.5	32.5
Eindhoven	22.7	24.2	26.1	27.5	28.8	30.6	31.9	33.1	34.7	35.9
Beek	18.4	19.5	20.9	21.9	22.8	24.0	24.9	25.8	26.9	27.7

Table 5.15: Return levels in m/s (RW-model).

The results of the RW-model have been set out graphically in Figure D.9 in Appendix D. For station Deelen and Soesterberg this yields Figure 5.14.

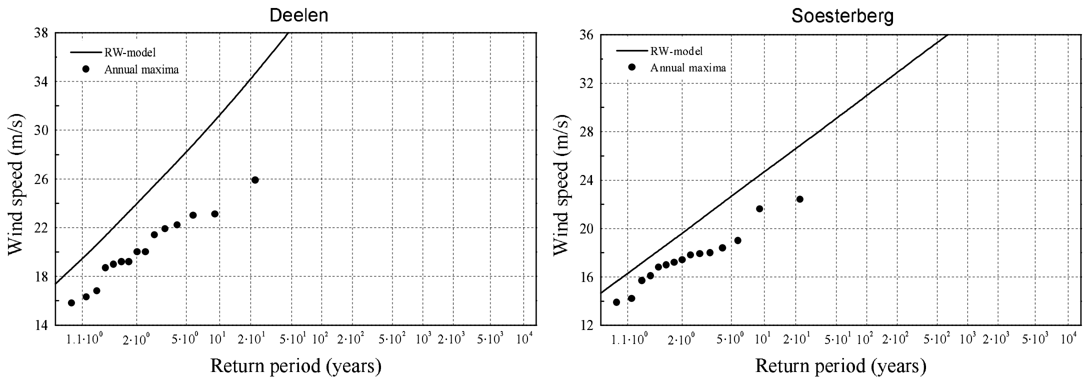

Figure 5.14: Comparison between the RW-model and annual maxima for station Deelen and Soesterberg.

Figure 5.14 together with Figure D.9 in Appendix D shows that also the further smoothing of the Weibull parameters does not correct the low *k*-values satisfactorily; the results for Soesterberg have been improved, but the results for Deelen have become even worse compared with the results from model F where the harmonic parameters were used in combination with the station independent persistence/tail correction.

5.2 Persistence corrections between neighbouring azimuth sectors and subsequent seasons

Paragraph 3.4 describes that constant factors were used for the mutual dependency between azimuth sectors and between subsequent seasons. Furthermore, the persistence factor for the dependency between hourly mean values of wind speed was determined as a declining function of the wind speed level. The idea was that the persistence would converge to 1 (no persistence) for high wind speeds. The same you would expect for the dependency between azimuth sectors and seasons, so it's a bit strange that Rijkoort has chosen to use constant factors for these dependencies. He probably determined these constant factors on empirical basis. Theoretically however, constant factors for these dependencies can not function properly when these factors are above 1. Then there will be return levels of wind speed for a certain return period corresponding with a certain wind direction that are higher than the overall (all wind directions) return level. In most cases, this behaviour only occurs for high return levels, but when the differences between wind directions and season are substantial, the behaviour will occur for lower wind speeds as well. Imagine for example the following values for $G_{ij}(u)$ in equation (3.23):

$$G_{ij}(u) = 0.80 \text{ for } i = 1,$$

$$G_{ij}(u) = 0.99 \text{ for } i = 2,..,12.$$

Then using (3.25):

$$G_j(u) \approx 0.846 .$$

Finally, using (3.27) and (3.28) we obtain:

$$T_{ij}(u) = 5 \text{ for } i = 1,$$
$$T_{ij}(u) = 100 \text{ for } i = 2,..,12.$$
$$T_j(u) \approx 6.5 .$$

Let's take for instance j equals 1 and u equals 20.
The above then says that the return period for a wind speed value of 20 m/s with wind direction 1 in season 1 equals 5 year and that the return period for a wind speed value of 20 m/s independent of the wind direction in season 1 equals approximately 6.5 year. This is of course contradictive.
The RW-model has more than one such cases.

5.3 Effect of class definition

The RW-model divides the wind data into several seasons and wind direction sectors. This is based on theoretical grounds and application grounds as well.
Theoretically (like stated in Paragraph 3.1) the division has been done because the (modified) Weibull distribution does not fit very well for the whole wind speed range, but only for the wind speeds between 3 and 15 m/s (Rijkoort, 1983). After split up into several seasons and wind direction sectors the modified Weibull distribution produces substantially better fits.
Besides this theoretical aspect, the division also suits from the point of view of application purposes. Dikes for example are sensitive for the direction the wind is coming from. For this reason it's better to obtain return periods that correspond with a certain wind direction than return periods that correspond with the wind directions combined. The same can be stated in the case of division into seasons.
The intervals for wind directions and seasons are chosen so that the (modified) Weibull distribution fit reasonably well. For this goal a division into as many seasons and wind directions as possible is preferred. Too many sectors though results in noise between the seasons and wind directions caused by the loss of information per azimuth-season sector. After examination Rijkoort has chosen to divide the wind data into six seasons and twelve wind direction groups.
What remains is how to define these seasons and wind direction groups. Table 3.2 and Table 3.3 in Paragraph 3.1 show the final class definition in the RW-model.
Subparagraph 5.3.1 and Subparagraph 5.3.2 discuss the results of the RW-model when the wind directions and seasons are defined differently.

5.3.1 Wind direction

From the point of view of application purposes it's important that with the help of the RW-model return levels can be derived for several azimuth sectors. In the Netherlands, differences in wind speed between the several wind directions are rather large due to the position of the country between water (North Sea and Atlantic Ocean) in the west and land in the east (Europe). A split up into enough wind directions is necessary to prevent smoothing out the effect of this between land-sea position of the Netherlands. For this reason a subdivision into less than 12 azimuth sectors is not wise.
An increase of the number of wind directions is another option but because a combination is made with a number of seasons, one has to be careful with the results due to lack of data in the several azimuth-season sectors. This lack of data can result in too much noise between the wind directions.
The above has been examined by replacing the split up into 12 azimuth sectors following the original RW-model by a split up into 18 azimuth sectors (Table 5.16). With this new definition of azimuth sectors the RW-model keeps its properties. Only the constants 30 and 60 in (3.10)-(3.14) have to be replaced by 20 and 40.

Wind direction		
1	15°	005°–024°
2	35°	025°–044°
3	55°	045°–064°
4	75°	065°–084°
5	95°	085°–104°
6	115°	105°–124°
7	135°	125°–144°
8	155°	145°–164°
9	175°	165°–184°
10	195°	185°–204°
11	215°	205°–224°
12	235°	225°–244°
13	255°	245°–264°
14	275°	265°–284°
15	295°	285°–304°
16	315°	305°–324°
17	335°	315°–344°
18	355°	335°–364°

Table 5.16: Subdivision into wind directions by increasing the number of azimuth sectors.

When the RW-model was reproduced with this new definition of wind direction sectors, the variation between the return levels corresponding with a certain azimuth sector didn't increase substantially. This is not surprising because of the harmonic analysis of wind direction. Furthermore, there were no significant differences in comparison with the original definition of wind direction splitting up into 12 azimuth sectors. Of course, a number of azimuth sectors produced different return values than in the original RW-model, but these deviations didn't seem to be structural, but more arbitrary.
So, the increasing variation between the azimuth sectors has been accounted for not by significant differences between azimuth sectors but by increasing noise between the azimuth sectors as a result of lack of data in the several azimuth-season sectors. Based on this, it's not wise to increase the numbers of azimuth sectors.

Also the effect of horizontal shifting of the original azimuth sectors has been examined. Shifts of 10 degrees leftwards and rightwards were performed (Table 5.17).

Wind direction		
RW-model	Shift to left	Shift to right
015°–044°	005°–034°	025°–054°
045°–074°	035°–064°	055°–084°
075°–104°	065°–094°	085°–114°
105°–134°	095°–124°	115°–144°
135°–164°	125°–154°	145°–174°
165°–194°	155°–184°	175°–204°
195°–224°	185°–214°	205°–234°
225°–254°	215°–244°	235°–264°
255°–284°	245°–274°	265°–294°
285°–314°	275°–304°	295°–324°
315°–344°	305°–334°	325°–354°
345°–014°	335°–004°	355°–024°

Table 5.17: Subdivisions into wind directions with different class definitions.

An overview of the results of these shiftings has been set out graphically in Figure E.1 in Appendix E, where the 10,000 year return periods (for Deelen 1,000 year as a results of the earlier stated problem of calculating some extreme values) have been set out following the separate class definitions of wind

directions.
As can be seen in this figure the greatest part of the differences between the azimuth sectors has been accounted for by the harmonic analysis of wind direction explained in Paragraph 3.2. As a result, the effect of different class definitions are rather small. Based on this there is no apparent reason to use another class definition than the one that has been defined in the RW-model.

5.3.2 Season

Because in the Netherlands extreme hourly mean values of wind speed mostly occur during winter rather than during summer, also the effect of seasonal class definition has been examined.

Firstly, the number of seasons has been increased from the original 6 seasons to 12 seasons (corresponding with months). As with the split up into more azimuth sectors, the RW-model keeps most of its properties when increasing the number of seasons to 12. Only the constant 60 in (4.8) has to be replaced by 30. Again the increasing variance didn't seem to be structural. No seasons produced significantly different return levels compared with the original definition of splitting up the seasons.

Also the effect of a horizontal shift of 1 month has been examined (Table 5.18).

Season	
RW-model	Shift by one month
January – February	December – January
March – April	February – March
May – June	April – May
July – August	June – July
September – October	August – September
November – December	October – November

Table 5.18: Subdivision into seasons with different class definitions.

An overview of the results of this shifting has been set out graphically in Figure E.2 in Appendix E, where the 10,000 year return periods (for Deelen again 1,000 year) have been set out following the separate class definitions of seasons. This figure shows that the structural differences between the class definitions are greater than in the case of the azimuth sectors, but they are still rather small. Based on this, again no much reason exists to use another class definition of season than the one used in the original RW-model.

5.4 Effect of length and quality of time series

In his analysis, Rijkoort used the time period 1962 till 1976 for twelve stations. Most of these stations started measuring before 1962 and except for one they kept measuring after 1976 (Figure 2.2). One can consider using all the data that are available for analysis. The more data you have the more exact will be the estimation of the return levels, one could argue. But because the wind speed in a certain year is always different from the preceding year, estimates of return levels will be different when using that certain year compared with the estimated values of return levels derived without using that same year. So, when after a certain year with many high winds the estimated return levels of wind speed are updated, the estimates probably will increase. The degree of this increase depends on the used extreme model in combination with the length of the series used in that model.
If one is interested in the return levels for only one station (this is mostly not the case), it's sensible to use as many data as possible. If, on the other hand, one wants to calculate the return levels of multiple stations (like in this paper), it's necessary to use the same time period for each station to make the comparison between the stations valid. By this the use of the same period for all the stations by Rijkoort has been justified and it means that the extension of the series with data after 1976 used to estimate e.g. 10^4 year return levels should be similar for all stations.

As stated earlier, the RW-model produced some very unrealistic return values for some stations. Partly, this was a consequence of the finite starting speed of some anemometers resulting in sometimes relatively high frequencies of low wind speed values. It shows that an important condition of the RW-model is that the quality of the wind speed data is optimal, especially for low wind speeds. A remarkable condition, because we are interested in return periods of extreme wind speeds and certainly not in those of low wind speeds. In general this asks for relatively recent periods of analysis because the quality of wind measurement has improved considerably over the last decades in the Netherlands (Table 5.19).

Station	% [0,1)	
	1962-1976	1981-1995
L.S. Texel	2.7	*
Schiphol	4.1	1.4
De Bilt	6.0	6.2
Soesterberg	12.4	6.2
Leeuwarden	2.6	2.5
Deelen	10.7	4.3
Eelde	3.0	2.3
Vlissingen	2.5	1.7
Zestienhoven	6.4	3.5
Gilze-Rijen	6.6	2.1
Eindhoven	7.6	4.2
Beek	3.2	1.2

Table 5.19: Percentages of wind speed between 0 and 1 m/s in two different 15 year time periods.

Another condition that has to be met in the RW-model is that reasonably long time series should be used. Because in the persistence/tail correction the annual maxima are used, sample sequences must be sufficiently long (see also Paragraph 7.1).
Both conditions (optimal quality of the wind speed data and long time series) are not in favour of the RW-model.

6 Adjustments to the Rijkoort-Weibull model

Several points in the RW-model appeared to be weak points. This has been discussed in the preceding chapter (Chapter 5). To adjust the RW-model one has to be careful to change these points without weakening relatively strong points in the model, because the preceding chapter indicated that a combination of separate points in the model can have a strengthening effect on the results of the RW-model.

The following list gives an overview of the weak points in the RW-model.

- The Weibull fits make use of the whole range of wind speed values, including low wind speed values that are not of interest to us.
- The determination of the persistence correction between successive hourly mean wind speed values, which also functions as a tail correction.
- The constant factors for the persistence correction between neighbouring azimuth sectors and between subsequent seasons.

Possible improvements of these points are discussed in Paragraph 6.1, 6.2 and 6.3. Finally, a different approach to the problem of persistence is discussed in Paragraph 6.4.

6.1 Improving the fit of the Weibull distribution

The RW-model uses all wind speed values to fit the Weibull parameters. We saw that relatively high frequencies of low wind speed values reduce the k-value which has a clear effect on the return levels of wind speed obtained by the RW-model. A method to prevent this is to impose a threshold value of wind speed and to fit the wind speed values above this threshold by the conditional (modified) Weibull distribution. This threshold however can not be put too high, because of the use of annual maxima during the determination of the persistence correction. Because some azimuth-season sectors contain only low values of wind speed, the derived annual maximum will be low as well. If this maximum is lower than the threshold value, no value can be calculated for $G_{ij}(u)$. For this reason the imposed threshold has to be as low as possible but at that level where the finite starting speed behaviour of the anemometer is not important any more (~4 m/s).
The general formula of the conditional distribution is defined as follows:

$$P(U < u \mid U > \omega) = \frac{P(U < u, U > \omega)}{P(U > \omega)} = \frac{F(u) - F(\omega)}{1 - F(\omega)}, \quad u > \omega \tag{6.1}$$

where ω represents the threshold value.
In the case of the RW-model the conditional Weibull distribution can be imposed as follows to obtain the parent distribution of wind speed for a certain azimuth-season sector:

$$F_{\text{con}}(u \mid i, j) = \left(1 - P_{ij}(\omega)\right) + P_{ij}(\omega)\frac{F(u \mid i, j) - F(\omega \mid i, j)}{1 - F(\omega \mid i, j)}, \quad u > \omega \tag{6.2}$$

where $P_{ij}(\omega)$ represents the percentage of hourly mean wind speed values in season j with wind direction i that exceed the threshold value ω. Furthermore, $F(u \mid i, j)$ equals (3.6).

6.2 Improving the persistence correction between successive hourly mean wind speed values

It was shown that the persistence correction factor also functions as a tail correction and that by averaging over the several stations a substantial error was created with severe effect on the return levels of wind speed. This can be prevented in two ways:

Firstly, one can take the station dependent persistence/tail corrections, so without calculating the average over the stations. By this, the extreme behaviour like in Subparagraph 5.1.3, 5.1.6 and 5.1.9 will not occur anymore, because near zero values of the q-values won't occur anymore. If no average over the stations is calculated it's also not necessary any more to divide the wind speed by the individual azimuth-season scale-parameter α_{ij}. Then the smoothing function (3.35) can be replaced by

$$\ln q_{ij}(u_I) = A_{ij}u_I^2 + B_{ij}\,. \qquad \textbf{(6.3)}$$

Doing this, most values of q will still converge to values below 1, so the persistence factor will still function as a tail correction also. To prevent this, the calculated smoothed values of q (by (3.35)) can be increased by adding the constant 1. This may not be the most elegant way, but a structural increase of the return levels won't appear anymore when this is done.
The best thing to do is to use both methods. Because when the averaging over the stations is skipped, the noise between stations increases and with this the probability of strange patterns in the q-values.

6.3 Improving the persistence correction between neighbouring azimuth sectors and subsequent seasons

In the RW-model, the dependency between neighbouring azimuth sectors and subsequent seasons has been accounted for by using constant values of persistence correction factors. It was shown that this can result in contradictive return levels when comparing individual azimuth sectors with the overall distribution. The same can be stated when comparing individual seasons with the whole year distribution.
Like with the dependency between successive hourly mean values of wind speed it's more reasonable that this factor follows a decreasing function with respect to the wind speed and converges for high wind speed values to one (at the level of no persistence). This can be done by treating the dependency between azimuth sectors and seasons in the same way as the successive hourly mean values of wind speed.

For azimuth sectors, this results in the following persistence correction:

$$G_i(u_I)^{\frac{1}{q_i(u_I)}} = H_i(u_I)\,, \text{ or } q_i(u_I) = \frac{\ln G_i(u_I)}{\ln H_i(u_I)}\,. \qquad \textbf{(6.4)}$$

These q-values can be smoothed the same way as has been done for the successive hourly mean values of wind speed:

$$\ln q_i(u_I) = A_i u_I^2 + B_i\,. \qquad \textbf{(6.5)}$$

The same procedure for seasons yields:

$$G_j(u_I)^{\frac{1}{q_j(u_I)}} = H_j(u_I)\,, \text{ or } q_j(u_I) = \frac{\ln G_j(u_I)}{\ln H_j(u_I)}\,, \qquad \textbf{(6.6)}$$

$$\ln q_j(u_I) = A_j u_I^2 + B_j\,. \qquad \textbf{(6.7)}$$

Like with the successive hourly mean values of wind speed the final values of q can be obtained without averaging the A- and B-values over the stations and by adding the constant value of 1.

6.4 A different approach to the problem of persistence

As discussed in Paragraph 3.3 the parent distribution $F(u)$ of wind speed can not be used directly to derive return periods of interest, because of the clustering of wind speed values. If one uses the parent distribution in a indirect way, one has to correct for the clustering in successive wind speed values. Rijkoort has followed this approach by deriving the distribution of the maximum in a year $G(u)$ and imposing a persistence/tail correction.
Another way to impose the persistence correction factor is to use the kind of information given in the example in Paragraph 3.3. In this example, the mean exceedance frequency of the wind speed value 22.8 per storm equals 3.2 for station Schiphol. For this reason, the return period of interest (the return period of a storm with at least one hourly mean wind speed value of 22.8 or higher) is a factor 3.2 larger than the return period of a single exceedance. This mean exceedance frequency is most probably a decreasing function of the wind speed. So

$$\widetilde{T}(u) = T(u) \cdot q(u) \,, \qquad \textbf{(6.8)}$$

where $\widetilde{T}(u)$ represents the return period of interest in hours, $T(u)$ the return period of a single hourly mean exceedance in hours and $q(u)$ the mean exceedance frequency of the hourly mean wind speed u per storm. When we rewrite (6.8) the following can be derived:

$$\widetilde{T}(u) = T(u) \cdot q(u) = \frac{1}{1 - F(u)} \cdot q(u) = \frac{1}{\frac{1}{q(u)}\left(1 - F(u)\right)} = \frac{1}{1 - \widetilde{F}(u)}, \qquad \textbf{(6.9)}$$

where

$$\widetilde{F}(u) = 1 + \frac{F(u) - 1}{q(u)} \,. \qquad \textbf{(6.10)}$$

In (6.9) $F(u)$ represents the parent distribution of u and $\widetilde{F}(u)$ can be interpreted as the parent distribution of u corrected for clustering. For low wind speeds $q(u)$ will be greater than 1 and therefore (6.9) won't work for wind speed values below ~2 m/s. Because wind speed values of interest are at a much higher level, this is not an important restriction.

(6.9) gives the return period for any event based on the parent distribution of the wind speed. But as stated earlier it's necessary to divide the wind data into several season and wind direction sectors based on theoretical and application grounds as well. For clarity, in the following the difference between the day- and nighttime distribution will be ignored. Then, return periods for separate azimuth-season sectors can be calculated as follows:

$$\widetilde{T}_{ij}(u) = T_{ij}(u) \cdot q_{ij}(u) = \frac{1}{1 - F(u \mid i,j)} \cdot \frac{1}{p_{ij}} \cdot q_{ij}(u) = \frac{1}{\frac{1}{q_{ij}(u)}\left(1 - F(u \mid i,j)\right)} \cdot \frac{1}{p_{ij}} = \frac{1}{1 - \widetilde{F}(u \mid i,j)} \cdot \frac{1}{p_{ij}}, \qquad \textbf{(6.11)}$$

where

$$F(u \mid i,j) = 1 - \exp\left[-\left(\frac{u}{\alpha_{ij}}\right)^{\kappa_{ij}}\right], \qquad \textbf{(6.12)}$$

$$\widetilde{F}(u \mid i,j) = 1 + \frac{F(u \mid i,j) - 1}{q_{ij}(u)}, \qquad \textbf{(6.13)}$$

$q_{ij}(u)$ represents the mean number of hourly mean values of wind speed with wind direction i that lie in season j per storm that exceed the value u and p_{ij} represents the proportion of hours with wind direction i and season j.

Now it can be shown that $q_{ij}(u)$ in (6.11) has the same meaning as the $q_{ij}(u)$ in the RW-model. If we return to this RW-model and ignore the difference between the day- and nighttime distribution again, the distribution of the maximum in a certain azimuth-season sector is given by

$$G_{ij}(u) = F(u \mid i, j)^{\frac{\delta_{ij}+\nu_{ij}}{q_{ij}(u)}} \quad \textbf{(6.14)}$$

If $q_{ij}(u)$ is introduced before the distribution of the maximum is derived we get

$$\bar{F}(u \mid i, j) = F(u \mid i, j)^{\frac{1}{q_{ij}(u)}} \quad \textbf{(6.15)}$$

It is not difficult to prove that the differences in outcomes between (6.13) and (6.15) are not that great and that they converge rather quickly to each other for higher wind speeds where the $q_{ij}(u)$ values converge to 1.

The difference between $q_{ij}(u)$ in the RW-model and in the new approach is the way of estimating these values. In the RW-model this has been done by comparing the RW-model without correction with observed annual maximum values of wind speed. In the new approach $q_{ij}(u)$ can be estimated by the mean exceedance frequencies of the wind speed per storm. The only thing what has to be done in the new approach is to smooth the separate $q_{ij}(u)$-values with the purpose of extrapolation. This smoothing though, won't be as difficult as in the case in the RW-model, because no separate $q_{ij}(u)$-values of below one will occur in the new approach.
More advantages of the new approach compared with that of Rijkoort are the physical meaning of the $q_{ij}(u)$-values and the quite easy extension to separate seasons and azimuth sectors which will be discussed below.

The estimation of the persistence correction can be improved following the new approach. The new approach however, calculates only return periods for separate azimuth-season sectors so far.
Following the new approach, return periods for a certain azimuth sector, for a certain season or for an arbitrary moment can be easily estimated, but again a correction has to be made for the dependency between neighbouring azimuth sectors and subsequent seasons.
Imagine the following extension of the example with wind speed data from station Schiphol (used in Paragraph 3.3): The wind data is split up into two azimuth sectors and these sectors are on average evenly represented per storm. We saw that on average the exceedance frequency of the wind speed 22.8 m/s was equal to 3.2 per storm. Then the exceedance frequency per azimuth sector equals 1.6. If we now derive the return period of the return level 22.8 m/s for an arbitrary moment the resulting return period will be a factor 2 too small if this return period has been estimated by combining the azimuth sectors without correction factor. This can be accounted for by using the mean number of wind directions during a storm. Theoretically this yields the return period for season j (ignoring differences between day and night):

$$\tilde{T}_j(u) = T_j(u) \cdot q_j(u) = \frac{1}{1 - F(u \mid j)} \cdot \frac{1}{p_j} \cdot q_j(u) = \frac{1}{\frac{1}{q_j(u)}\left(1 - F(u \mid j)\right)} \cdot \frac{1}{p_j} = \frac{1}{1 - \tilde{F}(u \mid j)} \cdot \frac{1}{p_j}, \quad \textbf{(6.16)}$$

where

$$F(u \mid j) = \frac{\sum_{i=1}^{12}\left(\tilde{F}(u \mid i, j) \cdot (\delta_{ij} + \nu_{ij})\right)}{\sum_{i=1}^{12}\left(\delta_{ij} + \nu_{ij}\right)}, \quad \textbf{(6.17)}$$

$$\tilde{F}(u \mid j) = 1 + \frac{F(u \mid j) - 1}{q_j(u)}, \quad \textbf{(6.18)}$$

$q_j(u)$ represents the mean number of azimuth sectors per storm above the threshold u in season j and p_j represents the proportion of hours in season j.

In analogy, for azimuth sectors this yields:

$$\widetilde{T}_i(u) = T_i(u)\cdot q_i(u) = \frac{1}{1-F(u\,|\,i)}\cdot\frac{1}{p_i}\cdot q_i(u) = \frac{1}{\frac{1}{q_i(u)}\left(1-F(u\,|\,i)\right)}\cdot\frac{1}{p_i} = \frac{1}{1-\widetilde{F}(u\,|\,i)}\cdot\frac{1}{p_i}, \qquad \textbf{(6.19)}$$

where

$$F(u\,|\,i) = \frac{\sum_{j=1}^{6}\left(\widetilde{F}(u\,|\,i,j)\cdot(\delta_{ij}+\nu_{ij})\right)}{\sum_{j=1}^{6}\left(\delta_{ij}+\nu_{ij}\right)}, \qquad \textbf{(6.20)}$$

$$\widetilde{F}(u\,|\,i) = 1 + \frac{F(u\,|\,i)-1}{q_i(u)}, \qquad \textbf{(6.21)}$$

$q_i(u)$ represents the mean number of seasons per storm above the threshold u for each azimuth sector i and p_i represents the proportion of hours with wind direction i.

It can be argued that $q_i(u)$ won't deviate strongly from 1, because switches between subsequent seasons don't occur often.

Overall return periods can now be calculated based on the combined seasonal distributions $\widetilde{F}(u\,|\,j)$:

$$\widetilde{T}(u) = T(u)\cdot q(u) = \frac{1}{1-F(u)}\cdot q(u) = \frac{1}{\frac{1}{q(u)}\left(1-F(u)\right)} = \frac{1}{1-\widetilde{F}(u)}, \qquad \textbf{(6.22)}$$

where

$$F(u) = \frac{\sum_{i=1}^{12}\sum_{j=1}^{6}\left(\widetilde{F}(u\,|\,j)\cdot(\delta_{ij}+\nu_{ij})\right)}{\sum_{i=1}^{12}\sum_{j=1}^{6}(\delta_{ij}+\nu_{ij})}, \qquad \textbf{(6.23)}$$

$$\widetilde{F}(u) = 1 + \frac{F(u)-1}{q(u)} \qquad \textbf{(6.24)}$$

and $q(u)$ represents the mean number of seasons per storm above the threshold u.

Overall return periods can also be calculated based on the combined azimuthal distributions $\widetilde{F}(u\,|\,i)$. In that case $q(u)$ represents the mean number of azimuth sectors per storm above the threshold u.

7 *Alternative extreme wind models*

Besides the adjustments within the RW-model and the new approach for the persistence problem in the RW-model also a number of other extreme wind models have been examined. This chapter discusses the pros and cons of these models in relation to the RW-model.

7.1 Combined Generalized Extreme Value model

In the RW-model, one important weak point is the determination of the persistence correction between successive wind speed values. Classical extreme value estimation though, is often used to calculate long return periods, whereby it is not needed any longer to correct for persistence.
The classical extreme value estimation describes how the maxima of samples of size N from sufficiently long sequences of independent and identically distributed random variables with distribution function $F(x)$ can be fitted asymptotically (for large N) to one of three basic families derived by Fisher and Tippett (1928). These families were combined into a single distribution by Von Mises (1936, in French; see Jenkinson (1955) for an explanation in English), known as the Generalized Extreme Value (GEV) distribution with distribution function:

$$F(x) = \exp\left[-(1-\theta y)^{\frac{1}{\theta}}\right], \qquad \theta \neq 0 \qquad \textbf{(7.1a)}$$

$$F(x) = \exp[-\exp(-y)], \qquad \theta = 0 \qquad \textbf{(7.1b)}$$

where θ is the shape parameter and y, the reduced variate, is given by

$$y = \frac{x-\lambda}{\beta}, \qquad \textbf{(7.2)}$$

where β is the scale parameter and λ the location parameter.
The shape parameter θ determines the type of distribution. (7.1b) represents the Fisher-Tippett Type I distribution, known as the Gumbel distribution. (7.1a) with a negative value of θ represents the Fisher-Tippett Type II, known as the Frechet distribution. (7.1a) with a positive value of θ represents the Fisher-Tippett Type III, known as the Reverse Weibull distribution.
The Gumbel distribution and the Frechet distribution are both unbounded at the right end of the distribution, but the Frechet distribution has a heavier tail than the Gumbel distribution and is bounded at the left end of the distribution. Furthermore, the Reverse Weibull distribution is bounded at the right end of the distribution.

Often the GEV is fitted to annual maxima, because they are often considered to be independent of each other (Paragraph 3.3). If we want to use the GEV distribution with respect to our point of view of application purposes, it's necessary to fit the distribution to each azimuth-season sector defined in Table 3.2 and Table 3.3. Then, the GEV distribution of the seasonal maxima of hourly mean wind speed values for each azimuth sector is given by:

$$G_{ij}(u) = \exp\left[(1-\theta_{ij} y_{ij})^{\frac{1}{\theta_{ij}}}\right], \qquad \theta_{ij} \neq 0 \qquad \textbf{(7.3a)}$$

$$G_{ij}(u) = \exp[-\exp(-y_{ij})], \qquad \theta_{ij} = 0 \qquad \textbf{(7.3b)}$$

where

$$y_{ij} = \frac{u-\lambda_{ij}}{\beta_{ij}}. \qquad \textbf{(7.4)}$$

In (7.3a) and (7.3b) $G_{ij}(u)$ indicates the probability that in season j the maximum of all wind speeds with wind direction i is smaller than u. The parameters β_{ij}, λ_{ij} and θ_{ij} have been estimated using the probability weighted moments (Palutikof, Brabson, Lister and Adcock, 1999).
Return periods corresponding for each azimuth-season sector can now be calculated using (3.27).
Next step in the model is to combine the separate azimuth-season $G_{ij}(u)$ distributions to derive probabilities for seasons, azimuth sectors and the whole year. Like in the RW-model, one has to take into account the dependency between neighbouring azimuth sectors and subsequent seasons. We have seen that several methods can be used to estimate these persistence correction factors. For practical reasons, the constant factors estimated in the original RW-model are used for this purpose, although better ways exist (discussed in Paragraph 6.4). Probabilities for seasons, azimuth sectors and the whole year are then derived using (3.24), (3.25) and (3.26). Finally, return periods can be calculated using (3.28), (3.29) and (3.30).
The results of this model (the combined GEV model) for the whole year are listed in Table 7.1.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	26.3	28.0	30.8	32.9	35.3	39.5	43.5	48.3	56.5	64.1
Schiphol	24.3	25.6	29.1	32.9	38.2	47.6	57.0	68.9	89.3	109.2
De Bilt	20.7	22.9	26.2	29.2	33.0	39.3	45.4	52.7	64.8	76.2
Soesterberg	21.3	22.9	25.0	27.0	29.4	33.4	37.4	42.2	50.1	57.6
Leeuwarden	23.6	25.8	29.3	32.8	36.9	43.4	49.3	56.3	67.5	77.7
Deelen	23.9	25.9	29.1	32.7	37.3	46.5	57.3	72.7	102.3	133.8
Eelde	21.7	23.7	26.3	28.4	30.9	35.9	40.9	47.0	57.5	68.1
Vlissingen	22.5	23.7	25.5	27.5	30.0	34.3	38.6	44.3	54.0	62.9
Zestienhoven	22.8	24.5	27.6	31.0	35.7	43.8	51.8	62.0	79.9	97.5
Gilze-Rijen	22.2	23.9	26.4	28.6	31.2	36.1	41.1	47.8	60.2	72.9
Eindhoven	22.2	24.0	26.6	28.7	31.0	35.0	38.9	43.6	51.5	58.6
Beek	20.3	21.8	24.2	26.3	28.9	33.2	37.6	43.3	53.5	63.6

Table 7.1: Return levels in m/s (combined GEV model) based on Rijkoort stations and time period 1962-1976.

When comparing the values listed in Table 7.1 to observed annual maxima it's very clear that the 10,000 year return levels are overestimated very much.

To find out why the combined GEV model performs unsatisfactory, we return to the individual azimuth-season distributions of the maximum ($G_{ij}(u)$), because they form the basis of the model.
In the beginning of this paragraph we saw that a condition of the GEV distribution is the independence between successive samples of random variables with distribution function $F(x)$ for asymptotic purposes. If this condition has been met, the observed annual maxima usually fit the asymptotic extreme value theory very well when N exceeds 100 (Tabony, 1983). Because it's evident that successive wind speed values are not independent at all, the real number of independent elements per azimuth-season sector decreases substantially. So, it's the question whether the split up into the several azimuth-seasons sectors doesn't reduce the numbers of values in each sample too much for meeting the above described condition of the GEV distribution.
This can be examined by looking whether the observed annual maxima in each azimuth-season sector make part of the tail of its parent distribution. When N (for certain azimuth-season sectors) is so small that some of the maxima are not being drawn from the tail then this is an indication that the assumption of extreme value theory is not met. The tail of the parent distribution can be loosely defined as the top 10-15% (Tabony, 1983). Table 7.2 lists the average number (relative) of annual maxima per azimuth-season sector for each station that is not drawn from the top 15% and the top 10% of the corresponding empirical parent distributions. It shows that the percentages of the annual maxima not being drawn from the top of the parent distribution are rather large. This indicates that the number of hourly mean values of wind speed per azimuth-season sector is too small to meet the assumptions of extreme value theory.
Based on this, we can doubt whether the combined GEV model is applicable with our azimuth-season classification.

Station	% of maxima below 85%	% of maxima below 90%
L.S. Texel	12.9	22.7
Schiphol	9.1	15.6
De Bilt	9.7	15.8
Soesterberg	11.4	19.7
Leeuwarden	10.7	18.5
Deelen	10.3	17.4
Eelde	8.2	15.6
Vlissingen	9.3	16.2
Zestienhoven	10.4	16.9
Gilze Rijen	10.0	17.2
Eindhoven	9.8	17.6
Beek	7.9	14.2

Table 7.2: Percentage of yearly maxima per azimuth sector and season not belonging to the tail.

Furthermore, Cook (1985) suggests that at least 20 years of data should be used for reliable results, and that the method must not be applied with fewer than 15 years of data. Because the return periods in Table 7.1 are based on 15 years of data, this is again an indication that the method will not work very properly for the Rijkoort data set.
For this reason, the method has been applied to all available stations and time series (Figure 2.2). The results are listed in Table 7.3.

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
Valkenburg	24.6	26.2	28.5	30.6	32.8	36.5	39.9	44.0	50.9	57.4
L.S. Texel	25.9	27.0	28.5	29.9	31.5	34.5	37.4	40.5	45.0	48.6
IJmuiden	23.1	23.9	24.9	25.6	26.4	27.5	28.5	29.7	31.4	32.9
Texelhors	25.4	26.9	29.0	30.7	32.4	34.7	36.4	38.1	40.4	42.3
De Kooy	24.2	25.4	27.0	28.3	29.5	31.1	32.5	34.5	37.4	39.8
Schiphol	23.8	24.9	26.4	27.6	28.8	30.3	31.3	32.3	33.7	35.0
Wijdenes	22.7	23.6	27.2	32.0	39.1	54.3	72.1	98.3	154.8	226.3
Terschelling	24.4	25.3	26.6	27.6	28.8	30.8	32.4	34.2	36.8	39.1
K13	27.0	28.5	31.2	33.8	36.6	40.8	44.4	48.9	56.8	64.5
M. Noordwijk	24.8	27.2	31.5	36.4	43.4	57.2	72.2	93.4	135.4	182.8
De Bilt	19.4	20.4	21.7	22.6	23.5	24.7	25.7	26.7	28.0	29.0
Soesterberg	20.6	22.0	23.8	25.3	27.0	29.8	32.2	34.7	38.1	40.8
Stavoren	22.3	23.9	27.0	29.7	33.0	38.5	43.7	49.9	59.9	69.3
Houtrib	24.3	26.5	29.8	32.6	35.7	40.4	44.8	50.2	59.7	69.5
Lelystad	25.0	27.5	31.0	34.2	38.2	45.8	53.7	63.3	79.2	94.2
Leeuwarden	22.8	24.0	25.6	27.0	28.4	30.1	31.3	32.4	33.8	34.8
Marknesse	21.0	22.4	24.3	26.0	28.4	32.7	37.3	43.3	54.1	64.6
Deelen	21.7	23.2	25.2	26.7	28.4	30.9	33.3	36.3	41.5	46.3
Lauwersoog	23.5	24.8	26.5	28.1	30.0	32.6	34.7	36.7	39.4	41.4
Eelde	20.9	22.1	23.8	25.5	27.2	29.2	30.7	32.0	33.8	35.0
Huibertgat	24.7	25.5	26.7	27.9	29.6	32.3	34.7	37.3	41.2	44.5
Twenthe	20.0	21.8	24.4	26.8	29.3	32.9	35.9	39.0	43.6	47.5
Cadzand	23.9	25.1	26.7	27.8	28.9	30.3	31.3	32.4	34.0	35.7
Vlissingen	22.0	22.9	24.0	24.8	25.6	26.9	28.1	29.7	33.2	36.6
Goeree	24.9	26.2	28.1	29.5	30.9	32.9	34.6	36.5	39.3	41.7
Europlatform	27.1	29.2	32.5	35.2	38.1	42.5	46.2	50.3	56.5	61.9
H. van Holland	22.8	23.7	24.7	25.4	26.1	27.0	27.8	28.9	31.4	33.6
Zestienhoven	21.9	22.9	24.1	25.0	26.0	27.6	28.8	30.0	31.7	33.0
Gilze-Rijen	20.6	21.8	23.5	24.8	26.0	27.7	28.9	30.3	32.3	34.0
Herwijnen	22.3	23.5	25.1	26.4	27.8	30.2	32.1	34.0	36.6	38.5
Eindhoven	21.1	22.5	24.4	25.8	27.2	29.1	30.6	32.2	34.6	36.8
Volkel	20.2	21.6	23.6	25.4	27.7	31.6	34.9	38.6	44.1	48.8
Beek	20.2	21.2	22.5	23.5	24.4	25.7	26.7	27.7	29.1	30.3

Table 7.3: Return levels in m/s (combined GEV model)) based on all stations and time periods available in the Netherlands.

Applied on all data available (from 5 years at Wijdenes to 47 years at IJmuiden), the overestimation of wind speed levels at high return periods still exists for a number of stations, but it seems that the average overestimation has been reduced. If we look more closely to the results we see a strong correlation between the length of the time series and the degree of overestimation. But even for some stations with a reasonable long history of measurements (± 35 years), the overestimation is quite large (Deelen, Twenthe and Volkel). The reason for this behaviour lies in the combination of the several azimuth-seasons sectors. Because there are 72 azimuth-season sectors, the probability is high that at least one of them produces unrealistic return levels. When the sectors are combined, this leads to unrealistic overall return levels as well. An example is given in Figure 7.1, where it can be seen that the azimuth-season sector (7,1) produces a 10,000 year return level (49.7) that is rather unrealistic, although the curve (Frechet) follows reasonably well the annual maxima. Without having any knowledge about other azimuth-season sectors the overall return level should be at least of this level. In this case the overall 10,000 year return level for Volkel is less than 49.7 (48.8). This contradictive behaviour is the result of using constant values for the dependency between neighbouring azimuth sectors and subsequent seasons.

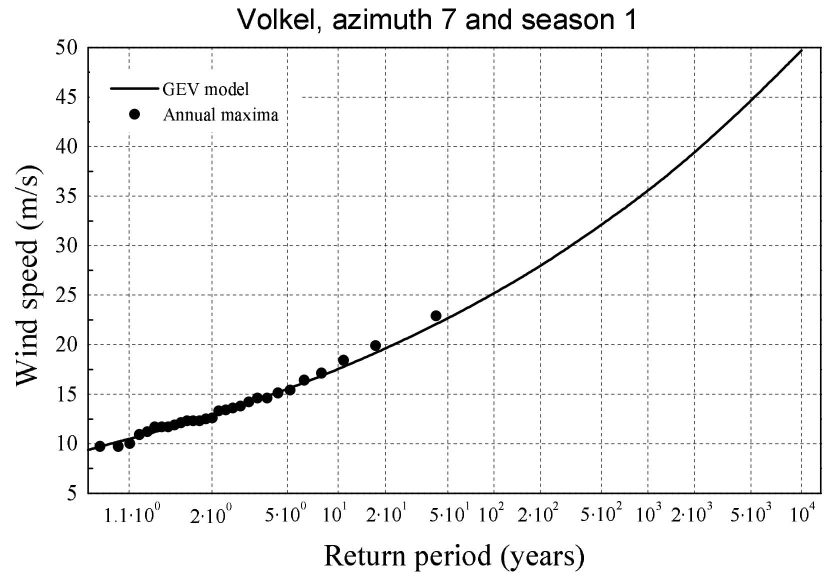

Figure 7.1: GEV fit for azimuth-season sector 7-1 at Volkel.

Figure 7.1 shows that even if the GEV distribution seems to fit reasonably well, the result is not always what one would expect. There are also cases some extreme points in an azimuth-season sector indicate bad performance. A good example of this can be seen in Figure 7.2, where the maxima of a severe storm on April 2, 1973 has a substantially effect on the fit of the GEV distribution.
Without this extreme of 26.9 m/s the fit would be a Reverse Weibull with a 10,000 year return level of about 20 m/s, but with the extreme the fit changes into a Frechet with a 10,000 year return level of 37.5 m/s. So, individual wind speed values can have a substantial effect on the way of fitting.

Based on what we have seen for station Volkel, we can conclude that with a split up into 72 azimuth-season sectors we need substantially more than 20 years of data for a good performance of the model. In combination with what we saw before (not meeting the asymptotic condition of the GEV distribution) the combined GEV model is not an improvement in comparison with the (adjusted) RW-model.

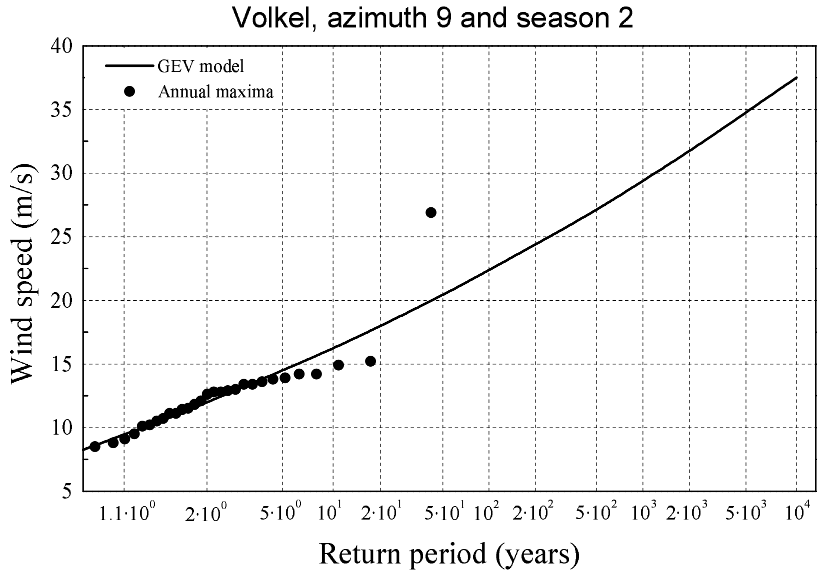

Figure 7.2: GEV fit for a azimuth-season sector 9-2 at Volkel.

7.2 One-step Markov chain model

Another extreme wind model has been examined. This method simulates long time series of hourly mean wind speed values and is described in detail by Dukes and Palutikof (1995). Based on the simulated time series estimates have been derived for return periods of interest.

The procedure is as follows:
Firstly, a transitional and probability matrix (TPM) is created. For this purpose, the wind speed values in our data sets were transformed into wind speed classes. Each wind speed class includes wind speeds between certain values. A class width of 1 m/s is applied, except for the lowest class where a width of 2 m/s is applied. For the top class an upper limit has to be set. The top class includes all wind speeds above this limit, which has been set on the nearest integer of the 99.5% percentile. Now the TPM can be created. This matrix shows the probabilities p_{ij} of a wind speed u in class i in hour n changing into class j in the following hour n+1. An example of a TPM for our data set is given in Table 7.4.

class	class												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.774	0.204	0.018	0.003	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.174	0.579	0.215	0.027	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.019	0.206	0.528	0.215	0.027	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000
4	0.003	0.035	0.228	0.495	0.201	0.033	0.005	0.001	0.000	0.000	0.000	0.000	0.000
5	0.001	0.006	0.045	0.242	0.475	0.196	0.030	0.004	0.001	0.000	0.000	0.000	0.000
6	0.000	0.001	0.010	0.055	0.253	0.460	0.185	0.030	0.005	0.001	0.000	0.000	0.000
7	0.000	0.000	0.003	0.012	0.055	0.273	0.444	0.177	0.029	0.005	0.001	0.001	0.000
8	0.000	0.000	0.000	0.002	0.016	0.069	0.267	0.423	0.185	0.031	0.005	0.001	0.001
9	0.000	0.000	0.000	0.001	0.003	0.020	0.081	0.272	0.421	0.170	0.029	0.005	0.001
10	0.000	0.000	0.000	0.001	0.003	0.008	0.018	0.090	0.285	0.375	0.176	0.037	0.007
11	0.000	0.000	0.000	0.001	0.000	0.006	0.008	0.028	0.094	0.279	0.374	0.169	0.041
12	0.000	0.000	0.001	0.000	0.000	0.003	0.003	0.010	0.022	0.104	0.264	0.370	0.223
13	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.005	0.007	0.021	0.065	0.165	0.733

Table 7.4: TPM for Eelde, 1962-1976. Listed are the probabilities p_{ij} that a wind speed in class i (rows) at hour n changes into class j (columns) at hour $n+1$.

Using this TPM simulated time series of wind speed classes were created using a random number generator. To transform the wind speed classes into wind speed values an uniform random generator can

be used for the classes except for the lowest and highest wind speed class. For these classes a shifted exponential distribution random number generator has been used.
Estimations of return levels have been calculated by simulating a number (here 100) of time series (each with different random number initiator) with length equal to the requested return period and then by averaging over the maximum wind speed values in each time serie.

The results of this model (one-step Markov chain model) are listed in Table 7.5. For return periods below 20 years the results have not been listed, because of the different definition of return periods. In the RW-model and other models, the return period of the annual maximum has been derived, while in this model the return period of an arbitrary wind speed is derived. As stated in Paragraph 3.3, for return periods above about 20 year the difference becomes negligible. Because of the different definition of return periods, the results are not graphically displayed and compared with the annual maxima (like before); the graph would begin at return periods where the observed annual maxima would end.

Station	Return period in years								
	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	31.4	32.7	33.5	35.1	37.0	38.0	39.4	40.9	42.1
Schiphol	28.7	30.6	31.9	33.0	34.8	35.8	37.0	38.4	39.4
De Bilt	22.2	23.4	24.2	25.2	26.5	27.4	28.3	29.6	30.6
Soesterberg	23.9	25.4	26.9	27.6	29.0	30.2	31.1	32.4	33.5
Leeuwarden	27.2	28.7	29.7	30.8	32.2	33.3	34.4	35.8	36.8
Deelen	26.1	27.5	29.0	30.0	31.4	32.4	33.6	34.8	35.7
Eelde	25.3	26.5	27.5	28.6	30.3	31.2	32.5	33.7	34.5
Vlissingen	25.2	26.6	27.7	28.6	30.1	30.9	31.7	33.0	34.0
Zestienhoven	25.4	27.2	27.7	29.0	30.1	31.5	32.4	33.8	34.8
Gilze-Rijen	24.4	25.0	26.3	27.1	28.5	29.5	30.6	31.9	32.9
Eindhoven	26.3	28.0	28.7	30.2	31.8	33.0	34.1	35.6	36.6
Beek	22.9	24.1	25.1	26.2	27.4	28.2	29.2	30.5	31.4

Table 7.5: Return levels in m/s (one-step Markov chain model) based on Rijkoort stations and time period 1962-1976.

Some remarks have to be made about the values listed in Table 7.5 and about the applicability of the model in the HYDRA project:

Firstly, the maximum derived from a time series of, let's say, 1000 year, is normally not a good estimation of the 1000 year return level of wind speed. The maximum (or with this method the average of the maxima) probably corresponds with a return period that is higher than the length of the time series. Consequently, return periods calculated by the one-step Markov chain model are underestimated and return levels overestimated.
The reason is that the maximum of the time series is a good estimate for the median return period but not for the mean return period (Tabony, 1983). The degree of difference between the median and mean depends on the underlying distribution, which has not been estimated by the method described in this paragraph. It's therefore hard to say something about the degree of underestimating the return periods. For annual maxima that are Gumbel distributed, the underestimation of the return period when the estimation is based on the highest observed annual maximum is about 44%, which can be derived from

$$H(u_m) = \frac{r - 0.44}{N + 0.12} \tag{7.5}$$

(Palutikov et al., 1999) where r represents the ranked number of the annual maxima and N represents the total years included in the observation series. For the highest wind speed u_N in the time period with length N =100 year, $H(u_N) \approx 0.9944$ and the corresponding return period is then $1/(1\text{-}0.9944) \approx 179$ years. For very long time series the relative difference between the length of these time series and the return periods corresponding with the highest values in these time series tends to ~ 44%.

The highest wind speed class is fitted by the shifted exponential distribution. It can be questioned whether this is the most appropriate distribution to use. The conditional Weibull distribution is the most obvious distribution to use with respect to wind speed values.

An important disadvantage of the values in Table 7.5 is that they are derived without splitting up the data into seasons and azimuth sectors, which is necessary from the point of view of application purposes. So if return periods that correspond with certain azimuth-seasons sectors have to be derived, the model has to be extended. If only extra information about separate seasons is required, the model can be extended with a number of transitional probability matrices for each season, which is easily done. However, if also information with respect to the wind direction is required, the wind direction has to be simulated besides the wind speed. This extension makes the model much more complex. In literature not that much is found on this subject. Cheng (1991) and Cheng & Chiu (1990; 1992) have accounted for some of the features described above. They analysed the simulated time series for its extreme value properties instead of deriving the maximum from it to estimate return levels.

7.3 Analysis of separate storms

So far, the basic extreme wind models discussed in this report are based on all data (RW-model, one-step Markov chain model (Paragraph 7.2)) or on annual maxima (combined GEV model (Paragraph 7.1)). We saw that methods based on all data can result in unreliable results due to the relative strong effect of uninteresting low wind speed values in the tail of the wind speed distribution (Chapter 5). Furthermore, when using annual maxima one has to use relative long periods of measurements to prevent overestimating return levels of wind speed if the GEV distribution is fitted on separate azimuth-season sectors (Paragraph 7.1). Using annual maxima also means losing important information. During a year, for example, multiple severe storms can occur, while there are years with no severe storms. The consequence of this is that when only the maxima per year are used, sometimes an interesting storm won't be used while in other cases wind speeds are used that are not that interesting.

In summary, one can say that using all data is not optimal because of the use of too many uninteresting wind speeds and using annual maxima is not optimal because of the use of too few interesting wind speeds. A possible way in between is using separate storms and fitting a distribution (exact or asymptotic) to the maximum wind speed per storm.
Models that use separate storms are widely used and accepted and much can be found on this subject in literature. It seems thereby worthwhile to examine these kinds of models more in detail. Recently, RIKZ has developed a method based on this theory but applied on water levels, wave heights and wave periods. This method could be examined for application of wind speeds.

8 Conclusions

It turns out that the Rijkoort-Weibull model has a number of shortcomings with the consequence that the return levels for hourly mean values of wind speed derived from the model are very unreliable.
Especially, the determination of the persistence correction for successive wind speed values in combination with the fit of the Weibull distribution where all the data are used sometimes results in unrealistic high return levels. The performance of the persistence correction is unsatisfactory because it is defined in a way that it also functions as a (rather poor) tail correction. Furthermore, the correction is calculated by averaging over the stations used in the analysis while it turns out that systematic deviations exist between these stations.
For some stations harmonic analysis of wind direction dependent parameters neutralises this effect, but generally taken, the harmonic analysis and further smoothing of parameters do not improve the results substantially.
Another weak point in the Rijkoort-Weibull model is the usage of constant values for the factors to correct for the dependency between neighbouring azimuth sectors and subsequent seasons. This results in some contradictive return levels with respect to certain azimuth sectors or seasons in comparison with the whole year (e.g. return levels corresponding to individual azimuth sectors that are higher than the overall return level).

A number of adjustments within the basic concept of the Rijkoort-Weibull model have been examined.
To prevent low wind speed values from having too much weight in the fit procedures of the Weibull distribution, a threshold is imposed and values above this threshold are fitted with the conditional Weibull distribution. Doing this, the tail estimation is less influenced by wind speed levels that are not interesting.
Furthermore, the determination of the persistence correction factor is modified by calculating the factor without averaging over stations and by forcing the factor to go to the level of no persistence at high wind speed values. By performing a persistence correction in this way, an unwanted tail correction is avoided.
Finally, another approach to the problem of persistence is introduced. This approach determines the duration of a storm and therefore has a physical meaning. An advantage of the new approach is that it produces much more robust persistence correction factors in comparison with those of the Rijkoort-Weibull model. This approach also can be used to correct for the dependency between neighbouring azimuth sectors and subsequent seasons properly.

Alternative extreme wind models also have been examined.
The first one, the combined GEV distribution, results in unrealistic high return levels because of the split up into many seasons and wind direction sectors. For a good performance of this model, the analysis should include longer time series and maxima derived per two years or more.
The second model, the one-step Markov chain model, has been performed without splitting the wind data into wind directions and seasons. The resulting return levels are rather robust. To be able to use this model in the HYDRA project, it needs to be extended substantially to obtain seasonal and azimuthal return levels.
Finally, the advantages of the analysis of separate storms have been described. There are many examples of models for the analysis of storms in literature (e.g. the RIKZ approach applied on water levels, wave heights and wave periods). In part, it is from these models that the idea of the new approach for determining the persistence correction arose.

9 Recommendations

- The original Rijkoort-Weibull model should not be used anymore in the HYDRA project, nor its results because of the shortcomings identified in this report. These shortcomings have a strong impact on the wind speed levels at high return periods.

- Two important adjustments are recommended within the basic concept of the Rijkoort-Weibull model that will substantially improve the results of the original model. They are:

  1) imposing a threshold and
  2) following a new approach to the problem of persistence.

  Applied in the HYDRA project, this will lead to a revised wind statistics in the near future.

- The results of the (adjusted) Rijkoort-Weibull model that focuses on the parent distribution of the wind speed should be compared with alternative extreme wind models that focus on the extremes. Two of the three alternatives considered in this report are promising because they do have interesting properties with respect to return levels of wind speed. One of these models (the One-step Markov chain), has to be extended substantially before it can be used within the HYDRA project. Because this extension is not well covered in literature, it is advised to explore the other one (that analyses separate storms) in more detail. A method based on this theory, used by RIKZ to analyse water levels, wave heights and wave periods, could be examined for application on wind speeds. Compared with the above adjustments to the Rijkoort-Weibull model, a larger effort will be needed before it contributes to an improved wind statistics.

Acknowledgements

First of all I would thank Chris Geerse (RIZA) who introduced me to the world of the Rijkoort-Weibull model. Our discussions about this subject and his earlier research after interpretation aspects of the model convinced me that the theory can not be interpreted in a simple clear way. For this reason, Chapter 3 has been performed with his contribution substantially.
Furthermore, Albert Klein Tank is thanked for his many remarks after reading earlier versions of this report and also is thanked Job Verkaik who gave me lots of helpful suggestions during the writing.
Finally, Kees Kok and Adri Buishand are thanked for their useful remarks after reading the final draft of this report.

References

Beljaars, A. C. M., 1987. The influence of sampling and filtering on measured wind gusts. *Journal of Atmospheric and Oceanic Technology*, **4**, 613-626.

Benard, A. and Bos-Levenbach, E. C., 1953. The plotting of observations on probability paper. *Statistica*, **7**, pp.163-173.

Cheng, E. D. H., 1991. Wind data generator: a knowledge-based expert system. *Journal of Wind Engineering & Industrial Aerodynamics*, **38**, 101-108.

Cheng, E. D. H. and Chiu, A. N. L., 1990. An expert system for extreme wind simulation. *Journal of Wind Engineering & Industrial Aerodynamics*, **36**, 1235-1243.

Cheng, E.D.H. and Chiu, A.N.L., 1992. A knowledge-based extreme wind simulation technique. *Journal of Wind Engineering & Industrial Aerodynamics*, 41-44, 543.

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

Dukes, M. D. G. and , Palutikov, J. P., 1995. Estimation of extreme wind speeds with very long return periods. *Journal of Applied Meteorology*, **34**, 1950-1961.

Fisher, R. A. and Tippett, L. H. C., 1928. Limiting forms of the frequency distribution of the largest or smallest members of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, **24**, 180-190.

Geerse, C. P. M., 1999. De interpretatie van het Rijkoort Weibull model. Concept, Institute for Inland Water Management and Waste Water Treatment (RIZA).

Jenkinson, A. F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, **81**, 158-171.

Palutikof, J. P., Brabson, B.B., Lister, D.H. and Adcock, S.T., 1999. A review of methods to calculate extreme wind speeds. *Meteorol. Appl.*, **6**, 119-132

Rijkoort, P. J., 1980. Het gebruik van de Weibull-verdeling voor windsnelheidsfrequenties gesplitst naar dag of nacht. **VSB 80-3**, Royal Netherlands Meteorological Institute (KNMI).

Rijkoort, P. J., 1980. Toepassing van de Weibull-verdeling op windsnelheidsgegevens van een aantal kuststations. **VSB 80-4**, Royal Netherlands Meteorological Institute (KNMI).

Rijkoort, P. J., 1983. A compound Weibull model for the description of surface wind velocity distributions. Scientific Report, **WR 83-13**, Royal Netherlands Meteorological Institute (KNMI).

Tabony, R. C., 1983. Extreme value analysis in meteorology. *The Meteorological Magazine*, **112**, 77-98.

Verkaik, J. W., 2000a. Documentatie windmetingen in Nederland. Royal Netherlands Meteorological Institute (KNMI), Climatological Services.

Verkaik, J. W., 2000b. Evaluation of two gustiness models for exposure correction calculations. *Journal of Applied Meteorology*, **39**, 1613-1626.

Von Mises, R., 1936. La Distribution de la Plus Grande de *n* Valeurs. Reprinted in *Selected Papers*, **II**, Amer. Math. Soc., Providence, RI (1954), 271-294.

Wieringa, J., 1976. An objective exposure correction method for average wind speeds measured at a sheltered location. *Quarterly Journal of the Royal Meteorological Society*, **102**, 241-253.

Wieringa, J., 1986. Roughness-dependent geographical interpolation of surface wind speed averages. *Quarterly Journal of the Royal Meteorological Society*, **112**, 867-889.

Wieringa, J. and Rijkoort, P. J., 1983. *Windklimaat van Nederland*. Staatsuitgeverij, Den Haag, 263 p.

List of notation

a	mean Weibull scale parameter
$a_{j0},.., a_{j4}$	harmonic Weibull scale parameters
$b_1,..,b_4$	harmonic regression parameters
c	amplitude parameter in seasonal harmonic analysis
con	conditional
d	daytime
$d_{j0},..,d_{j4}$	harmonic daytime observations parameters
d^*, n^*	seasonal means of the normalised amplitudes of the first harmonic of δ and ν
$d^*_{j2}, d^*_{j4}, g^*_{j0}, k^*_{j0}, n^*_{j2}, n^*_{j4}$	parameters derived from d_{j2}, d_{j4}, g_{j0}, k_{j0}, n_{j2} and n_{j4}
f	exceedance frequency
g	mean Weibull stability parameter
$g_{j0},..,g_{j4}$	harmonic Weibull stability parameters
h	proportion of hours
i	azimuth group
j	season group
k	mean Weibull shape parameter
$k_{j0},..,k_{j4}$	harmonic Weibull shape parameters
max	maximum
n	nighttime
$n_{j0},..,n_{j4}$	harmonic nighttime observation parameters
p	transition probability
q	persistence/tail correction factor
r	rank number
u	wind speed
x	variable
y	reduced variate
A	persistence correction factor
B	tail correction factor
$D_{j1}, D_{j3}, N_{j1}, N_{j3}$	parameters derived from d_{j1},d_{j3},n_{j1} and n_{j3}
F	distribution function
$\widetilde{F}$	adjusted distribution function
G	distribution function of the maximum
$\overline{G}$	complement of G
H	plotting position
N	sample size
P	probability
R^2	correlation coefficient
T	return period
$\widetilde{T}$	adjusted return period
α	Weibull scale parameter
β	GEV scale parameter
χ	shifting parameter in seasonal harmonic analysis
δ	daytime observations
γ	Weibull stability parameter
κ	Weibull shape parameter
λ	GEV location parameter
ν	nighttime observations
θ	GEV shape parameter

Appendices

Several outcomes which resulted from the different models have been set out.
The parameters and return levels calculated by Rijkoort are shown in Appendix A (Table A.1, A.2, A.3 and A.4). The values in Table A.1, A.2 and A.4, are not identical to the values listed in Table 1, Table 2 and Figures 22.2-22.8 in Rijkoort (1983). From Table 1, the original not corrected values are copied into Table A.1. From Table 2, most of the values are copied directly into Table A.2, except for the parameters D_{11}, n_{22}^* and n_{42}^* which have been subjected to typing errors. Finally, the values in Table A.4 are partly based on the values in Table A.1 and Table A.2.
Appendix B sets out the parameters and return levels calculated by the reproduction (Table B.1, B.2, B.3 and B.4).
Appendix C shows scatterplots in which the results from Rijkoort and the reproduction have been set out (Figure C.1, C.2, C.3 and C.4).
Appendix D shows the resulting return periods for the several steps in the RW-model. The plotting positions of the annual maximum wind speeds are calculated following the formula of Benard and Bos-Levenbach (1953).
Appendix E shows the effect of class definition of wind direction and season in a graphic way.

A Parameters and return levels calculated by Rijkoort

Station	a	χ	k	g	d^*	n^*
L.S. Texel	7.99	129	2.47	0.00	0.36	0.27
Schiphol	6.36	74	2.35	0.76	0.43	0.30
De Bilt	5.05	58	2.28	0.79	0.42	0.34
Soesterberg	5.24	38	2.26	0.87	0.44	0.36
Leeuwarden	6.38	68	2.48	0.70	0.42	0.30
Deelen	5.77	66	2.18	0.82	0.44	0.35
Eelde	5.52	58	2.37	0.74	0.43	0.39
Vlissingen	5.86	101	2.33	0.30	0.45	0.41
Zestienhoven	6.21	78	2.35	0.75	0.44	0.41
Gilze-Rijen	5.58	60	2.35	0.85	0.49	0.51
Eindhoven	5.60	58	2.32	0.79	0.48	0.42
Beek	4.69	75	2.25	0.45	0.52	0.52

Table A.1: Station dependent parameters (Rijkoort).

	a_{j1}	a_{j2}	a_{j3}	a_{j4}	
January - February	0.98	198	0.60	312	
March - April	0.90	163	0.65	317	
May - June	0.77	165	0.60	316	
July - August	1.04	179	0.47	318	
September - October	1.28	200	0.47	313	
November - December	1.39	195	0.51	298	
	k_{j0}^*	k_{j1}	k_{j2}		
January - February	0.941	0.25	-19		
March - April	1.034	0.13	59		
May - June	1.066	0.20	67		
July - August	1.089	0.20	54		
September - October	0.956	0.20	8		
November - December	0.915	0.13	-40		
	g_{j0}^*	g_{j1}	g_{j2}		
January - February	0.566	0.12	122		
March - April	1.191	0.26	138		
May - June	1.316	0.20	173		
July - August	1.304	0.32	169		
September - October	1.117	0.26	143		
November - December	0.503	0.10	149		
	d_{j0}	D_{j1}	d_{j2}^*	D_{j3}	d_{j4}^*
January - February	29.4	-0.10	249	-0.05	327
March - April	50.7	-0.21	151	-0.03	335
May - June	68.6	-0.05	170	-0.04	333
July - August	67.2	0.06	174	0.04	324
September - October	50.7	0.05	233	0.01	350
November - December	25.3	0.25	232	0.08	354
	n_{j0}	N_{j1}	n_{j2}^*	N_{j3}	n_{j4}^*
January - February	89.0	0.05	258	0.00	-27
March - April	71.3	-0.23	132	0.00	-11
May - June	53.4	-0.18	156	0.00	1
July - August	56.8	-0.06	194	0.00	-2
September - October	71.3	0.16	257	0.00	10
November - December	96.7	0.30	235	0.00	-5

Table A.2: Season dependent parameters (Rijkoort).

Wind Direction	January - February		March – April		May – June	
	A	B	A	B	A	B
30º	-0.412	2.86	-0.571	2.83	-0.630	3.16
60º	-0.587	3.20	-0.695	3.24	-0.602	3.28
90º	-0.613	3.49	-0.834	3.76	-0.604	3.05
120º	-0.638	3.50	-0.842	3.27	-0.698	2.80
150º	-0.511	3.08	-0.560	2.30	-0.518	2.32
180º	-0.585	3.19	-0.532	2.49	-0.422	2.56
210º	-0.536	3.45	-0.446	2.60	-0.475	2.90
240º	-0.374	3.25	-0.480	2.92	-0.584	3.18
270º	-0.381	3.28	-0.761	3.65	-0.590	3.20
300º	-0.423	2.88	-0.925	3.75	-0.824	3.48
330º	-0.466	2.28	-1.131	3.77	-0.952	3.40
360º	-0.368	2.55	-0.635	2.72	-0.677	2.87
Wind Direction	July - August		September – October		November - December	
	A	B	A	B	A	B
30º	-0.956	3.76	-0.557	2.80	-0.434	2.78
60º	-0.799	3.56	-0.500	2.82	-0.583	2.93
90º	-0.687	2.97	-0.525	2.52	-0.647	3.18
120º	-0.815	2.65	-0.470	2.41	-0.548	3.04
150º	-0.746	2.65	-0.519	2.88	-0.590	2.70
180º	-0.749	3.13	-0.508	2.88	-0.572	2.76
210º	-0.612	3.23	-0.470	2.78	-0.485	2.82
240º	-0.501	3.10	-0.448	3.31	-0.337	2.91
270º	-0.550	3.32	-0.365	2.74	-0.340	2.56
300º	-0.835	3.45	-0.418	2.29	-0.480	2.97
330º	-1.118	3.69	-0.521	2.16	-0.398	1.55
360º	-0.835	3.35	-0.679	2.79	-0.494	2.27

Table A.3: Persistence/tail correction parameters (Rijkoort).

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	27.8	29.2	30.9	32.2	33.4	34.9	36.0	37.1	38.5	39.5
Schiphol	24.1	25.5	27.2	28.5	29.7	31.3	32.5	33.6	35.1	36.2
De Bilt	20.8	22.1	23.8	24.9	26.1	27.6	28.7	29.8	31.2	32.3
Soesterberg	21.5	22.9	24.6	25.8	27.0	28.6	29.8	31.0	32.5	33.6
Leeuwarden	21.9	23.0	24.5	25.5	26.5	27.8	28.8	29.7	30.9	31.7
Deelen	26.4	28.2	30.5	32.2	34.0	36.3	38.0	39.7	41.9	43.6
Eelde	20.8	22.0	23.4	24.5	25.6	26.9	27.9	28.9	30.2	31.1
Vlissingen	23.4	24.8	26.4	27.6	28.8	30.3	31.5	32.5	33.9	35.0
Zestienhoven	23.7	25.1	26.8	28.0	29.2	30.8	31.9	33.1	34.5	35.6
Gilze-Rijen	21.4	22.6	24.2	25.3	26.4	27.8	28.9	29.9	31.2	32.2
Eindhoven	21.9	23.2	24.9	26.1	27.2	28.7	29.8	30.9	32.3	33.4
Beek	20.7	22.0	23.6	24.8	26.0	27.5	28.7	29.8	31.2	32.3

Table A.4: Return levels in m/s (Rijkoort).

B Parameters and return levels resulting from the reproduction

Station	a	χ	k	g	d^*	n^*
L.S. Texel	8.06	128	2.43	-0.02	0.41	0.28
Schiphol	6.44	74	2.40	0.67	0.44	0.32
De Bilt	5.16	59	2.31	0.79	0.43	0.41
Soesterberg	5.28	34	2.17	0.87	0.46	0.42
Leeuwarden	6.47	73	2.50	0.70	0.43	0.31
Deelen	5.80	65	2.11	0.83	0.47	0.40
Eelde	5.64	61	2.43	0.71	0.45	0.43
Vlissingen	5.98	100	2.39	0.32	0.48	0.44
Zestienhoven	6.26	76	2.29	0.77	0.46	0.43
Gilze-Rijen	5.67	60	2.38	0.83	0.53	0.54
Eindhoven	5.65	59	2.30	0.78	0.52	0.45
Beek	4.90	83	2.43	0.44	0.67	0.64

Table B.1: Station dependent parameters (Reproduction).

	a_{j1}	a_{j2}	a_{j3}	a_{j4}	
January - February	0.88	193	0.48	314	
March - April	0.77	166	0.56	313	
May - June	0.63	170	0.51	313	
July - August	0.88	183	0.36	316	
September - October	1.08	200	0.31	311	
November - December	1.18	191	0.34	308	
	k_{j0}^*	k_{j1}	k_{j2}		
January - February	0.937	0.27	355		
March - April	1.021	0.11	24		
May - June	1.061	0.18	12		
July - August	1.090	0.25	25		
September - October	0.960	0.34	16		
November - December	0.931	0.22	2		
	g_{j0}^*	g_{j1}	g_{j2}		
January - February	0.549	0.11	122		
March - April	1.195	0.22	153		
May - June	1.347	0.17	177		
July - August	1.320	0.28	178		
September - October	1.088	0.20	180		
November - December	0.501	0.10	146		
	d_{j0}	D_{j1}	d_{j2}^*	D_{j3}	d_{j4}^*
January - February	29.41	-0.11	248	-0.06	322
March - April	50.75	-0.22	151	-0.04	336
May - June	68.58	-0.05	169	-0.03	335
July - August	67.17	0.07	175	0.04	324
September - October	50.75	0.06	230	0.02	346
November - December	25.33	0.25	231	0.07	348
	n_{j0}	N_{j1}	n_{j2}^*	N_{j3}	n_{j4}^*
January - February	89.13	0.02	256	0.00	328
March - April	71.25	-0.26	136	0.00	347
May - June	53.42	-0.18	156	0.00	0
July - August	56.83	-0.06	193	0.00	355
September - October	71.25	0.19	252	0.00	4
November - December	96.67	0.30	233	0.00	348

Table B.2: Season dependent parameters (Reproduction).

Wind Direction	January - February		March - April		May - June	
	A	B	A	B	A	B
30°	-0.408	3.11	-0.442	2.74	-0.538	3.11
60°	-0.641	3.65	-0.692	3.34	-0.632	3.36
90°	-0.658	3.57	-0.827	3.80	-0.700	3.19
120°	-0.576	3.62	-0.748	3.12	-0.612	2.40
150°	-0.424	3.04	-0.407	1.81	-0.457	2.02
180°	-0.558	3.25	-0.562	2.53	-0.529	2.67
210°	-0.515	3.59	-0.475	2.64	-0.539	2.94
240°	-0.419	3.66	-0.463	2.81	-0.597	3.34
270°	-0.361	3.25	-0.720	3.69	-0.327	2.85
300°	-0.435	2.95	-0.783	3.61	-0.611	3.21
330°	-0.492	2.10	-0.952	3.55	-0.623	2.85
360°	-0.330	2.31	-0.376	2.34	-0.130	1.95
Wind Direction	July - August		September - October		November - December	
	A	B	A	B	A	B
30°	-0.909	3.79	-0.726	3.14	-0.416	2.90
60°	-0.935	3.84	-0.689	3.15	-0.805	3.55
90°	-0.728	3.01	-0.540	2.68	-0.922	3.65
120°	-0.781	2.34	-0.243	1.94	-0.595	3.08
150°	-0.809	2.53	-0.430	2.80	-0.550	2.92
180°	-0.794	3.11	-0.434	2.77	-0.499	2.82
210°	-0.642	3.32	-0.270	2.64	-0.353	2.67
240°	-0.435	3.01	-0.290	3.12	-0.225	2.82
270°	-0.477	3.19	-0.139	2.60	-0.301	2.90
300°	-0.734	3.36	-0.241	2.05	-0.481	2.67
330°	-0.948	3.38	-0.475	1.87	-0.444	1.38
360°	-0.660	2.99	-0.685	2.81	-0.615	2.37

Table B.3: Persistence/tail correction parameters (Reproduction).

Station	Return period in years									
	10	20	50	100	200	500	1000	2000	5000	10000
L.S. Texel	28.7	30.3	32.3	33.7	35.1	36.8	38.1	39.4	40.9	42.1
Schiphol	23.4	24.8	26.6	27.9	29.1	30.7	31.9	33.0	34.5	35.5
De Bilt	20.7	22.0	23.8	25.0	26.3	27.9	29.0	30.2	31.7	32.7
Soesterberg	24.7	26.6	29.1	31.0	32.9	35.4	37.2	39.0	41.4	43.0
Leeuwarden	21.8	23.0	24.6	25.7	26.7	28.1	29.1	30.1	31.3	32.2
Deelen	31.2	34.2	38.3	41.5	44.8	49.4	51.4	54.5	55.9	57.5
Eelde	20.2	21.4	22.9	24.0	25.0	26.4	27.4	28.4	29.6	30.5
Vlissingen	22.6	23.9	25.6	26.8	28.0	29.5	30.6	31.7	33.1	34.1
Zestienhoven	25.4	27.1	29.2	30.8	32.3	34.2	35.7	37.1	38.9	40.2
Gilze-Rijen	21.2	22.5	24.1	25.3	26.5	28.0	29.1	30.1	31.5	32.5
Eindhoven	22.7	24.2	26.1	27.5	28.8	30.6	31.9	33.1	34.7	35.9
Beek	18.4	19.5	20.9	21.9	22.8	24.0	24.9	25.8	26.9	27.7

Table B.4: Return levels in m/s (Reproduction).

C Scatterplots of results calculated by Rijkoort and resulting from the reproduction

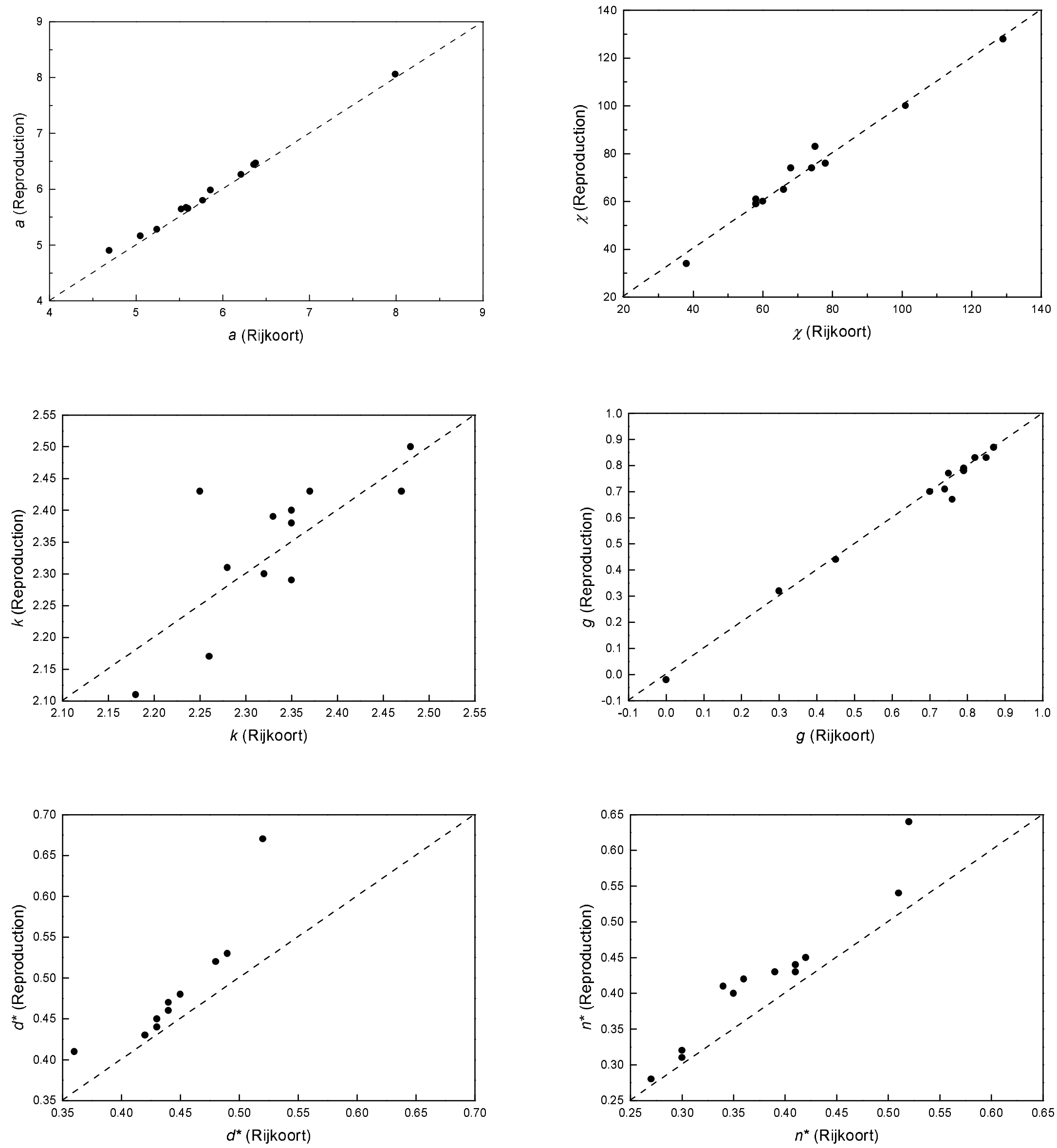

Figure C.1: Scatterplots of station dependent parameters calculated by Rijkoort and resulting from the reproduction. The diagonal dashed lines in these figures represent the one to one relations between the results.

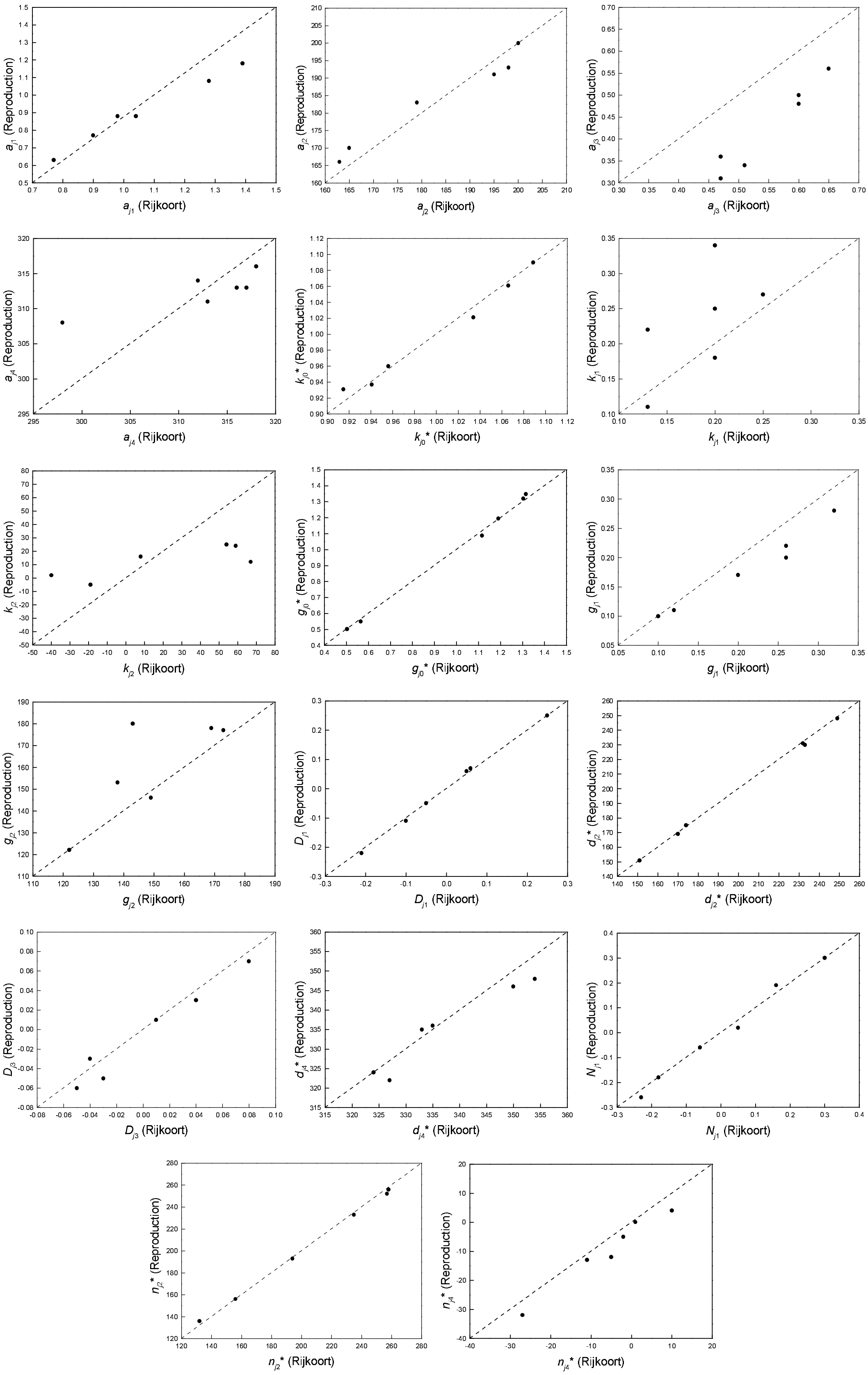

Figure C.2: Scatterplots of season dependent parameters calculated by Rijkoort and resulting from the reproduction. The diagonal dashed lines in these figures represent the one to one relations between the results.

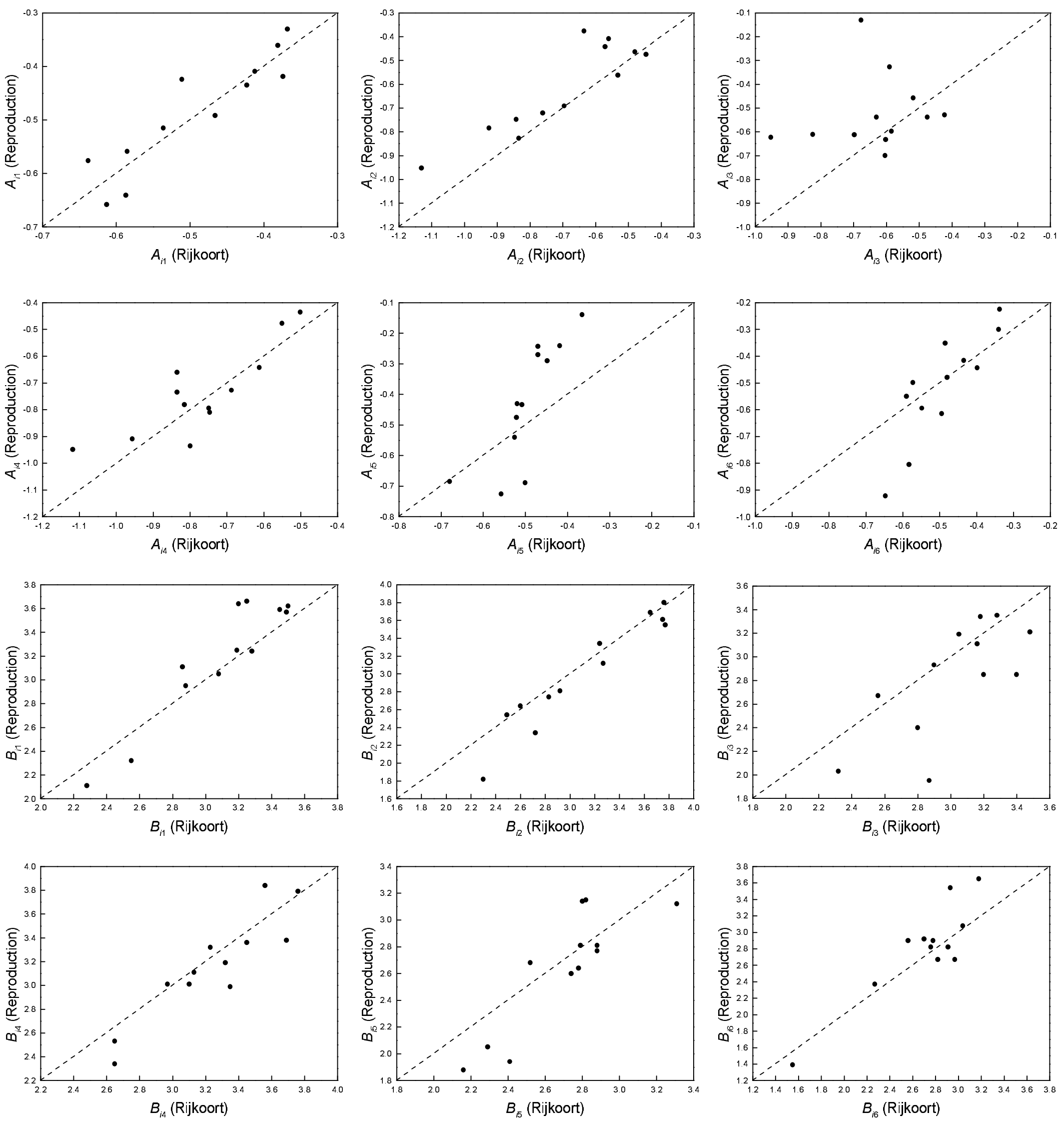

Figure C.3: Scatterplots of persistence/tail correction parameters calculated by Rijkoort and resulting from the reproduction. The diagonal dashed lines in these figures represent the one to one relations between the results.

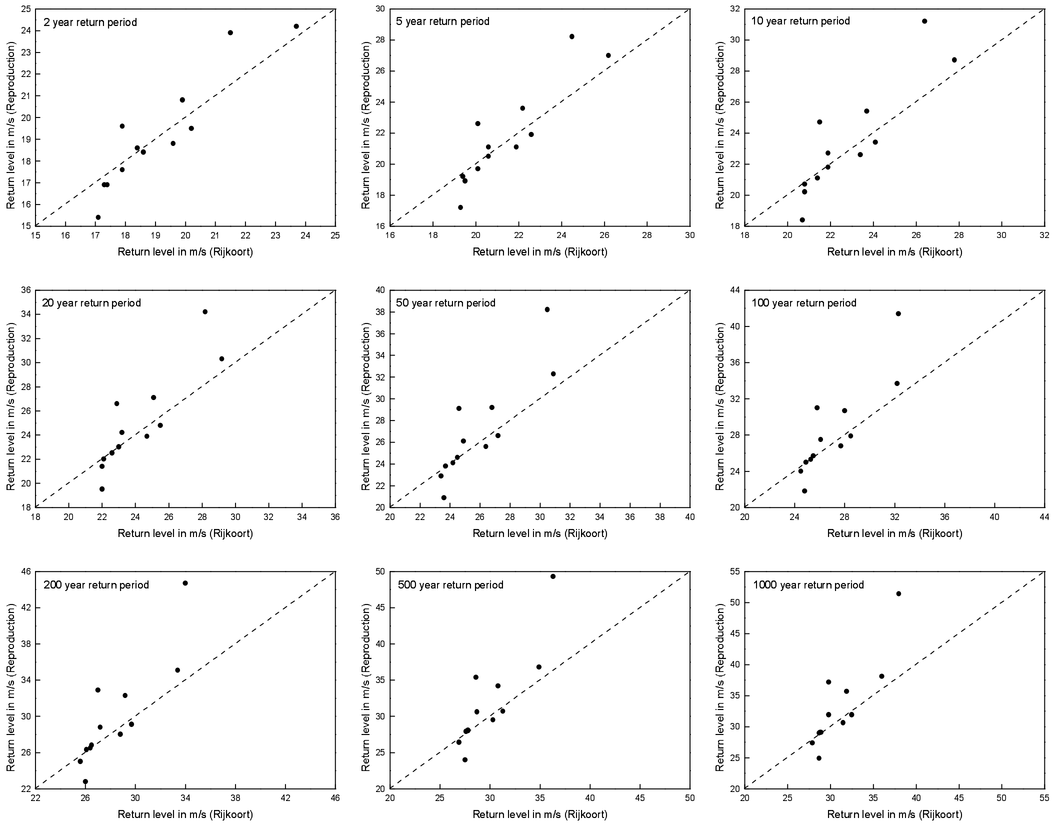

Figure C.4: Scatterplot of return levels calculated by Rijkoort and resulting from the reproduction. The diagonal dashed lines in these figures represent the one to one relations between the results.

D Resulting return periods in the RW step by step models

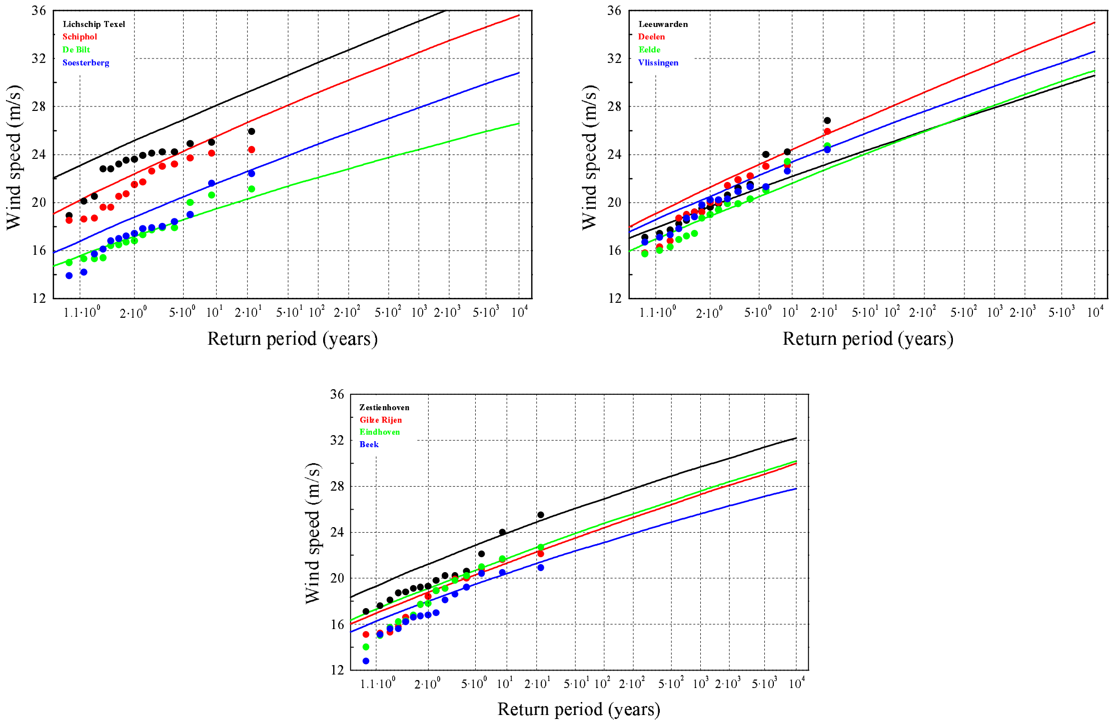

Figure D.1: Return periods and return levels following model A. The closed circles represent the annual maximum wind speed.

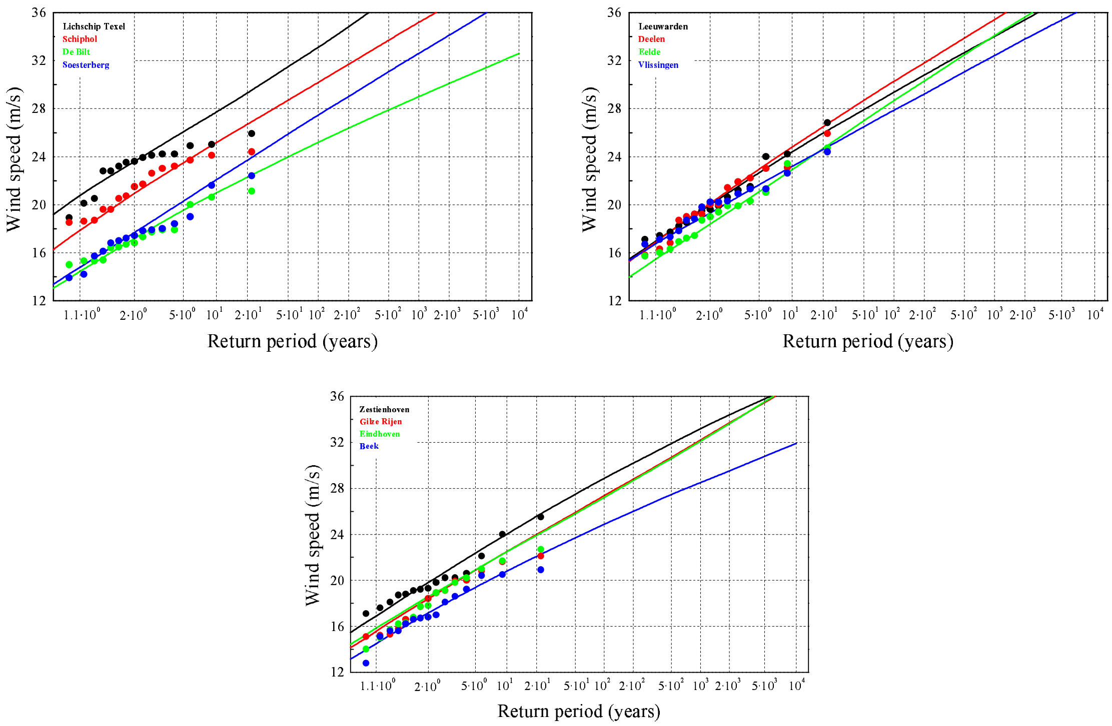

Figure D.2: Return periods and return levels following model B. The closed circles represent the annual maximum wind speed.

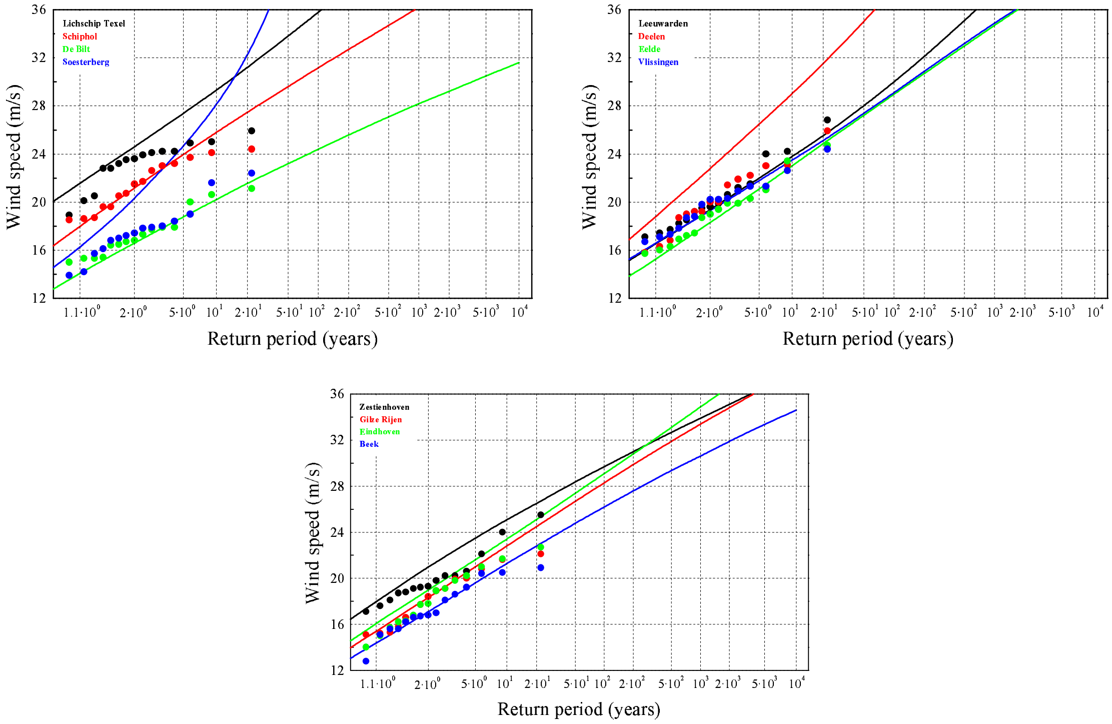

Figure D.3: Return periods and return levels following model C. The closed circles represent the annual maximum wind speed.

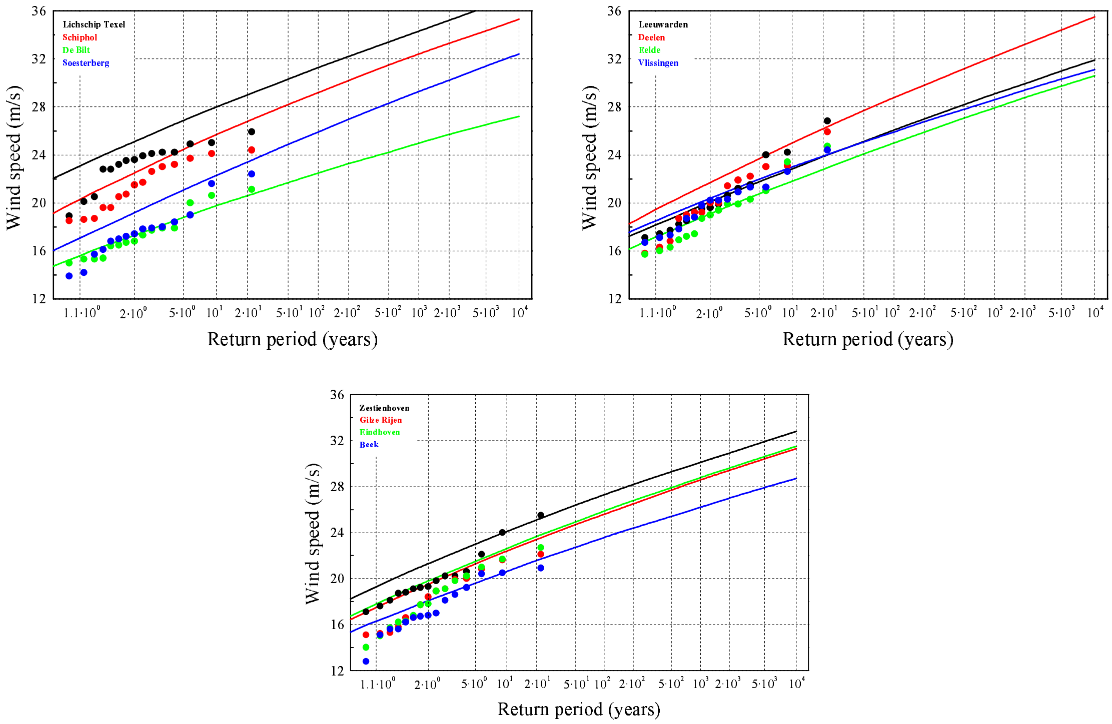

Figure D.4: Return periods and return levels following model D. The closed circles represent the annual maximum wind speed.

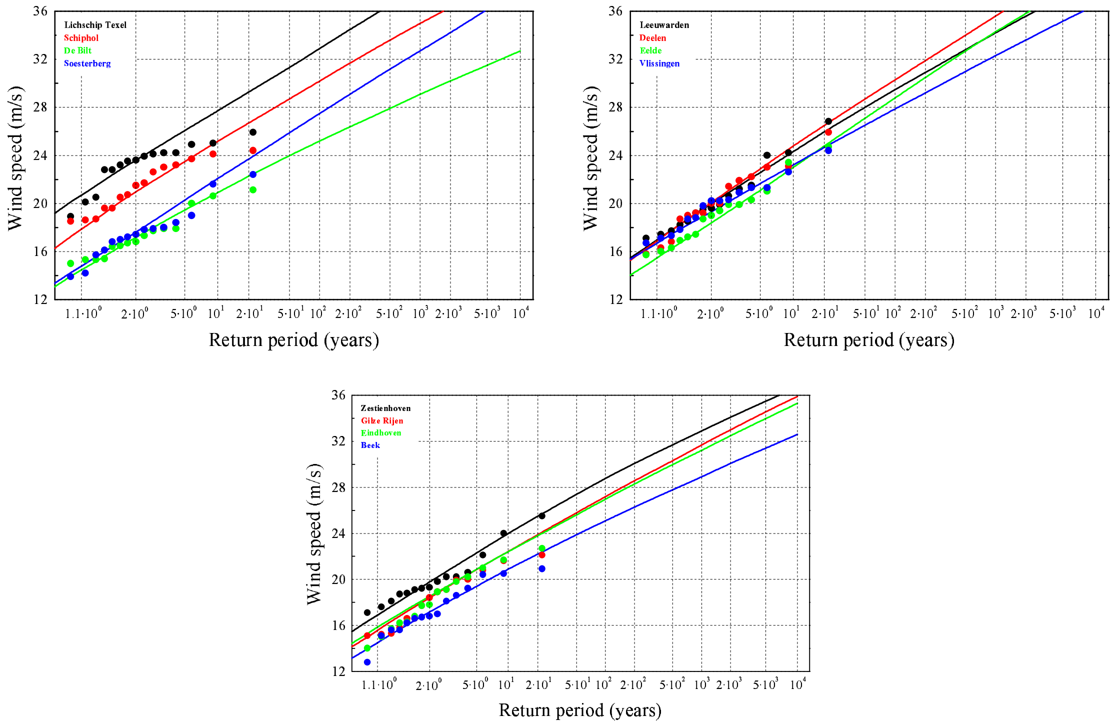

Figure D.5: Return periods and return levels following model E. The closed circles represent the annual maximum wind speed.

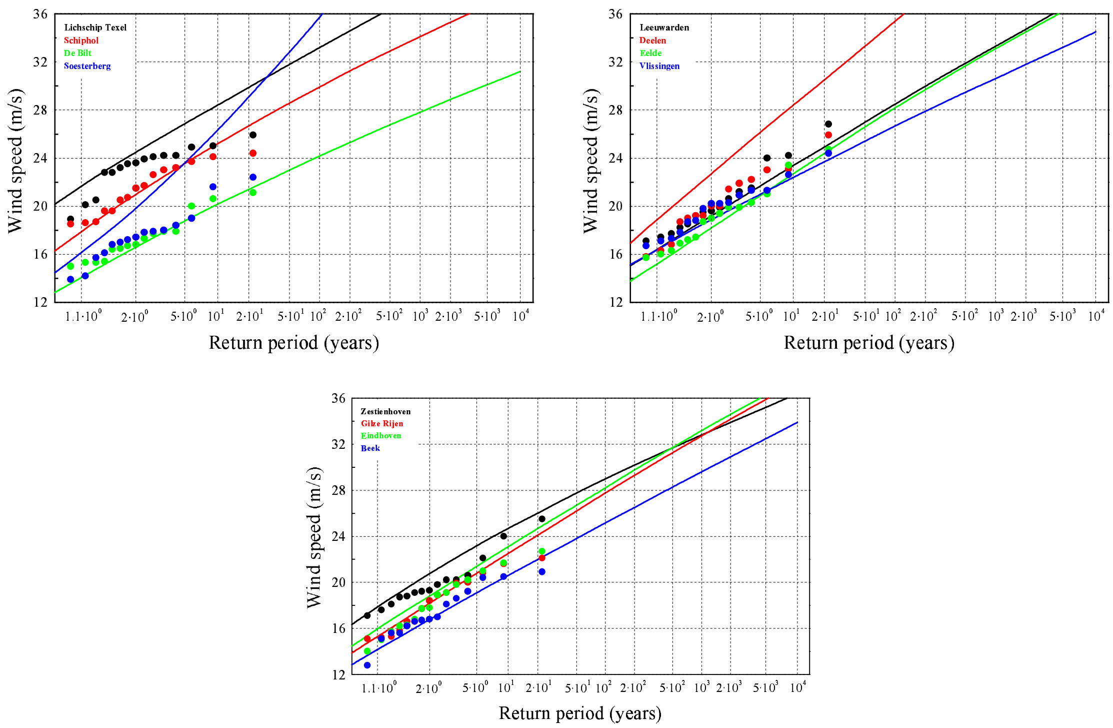

Figure D.6: Return periods and return levels following model F. The closed circles represent the annual maximum wind speed.

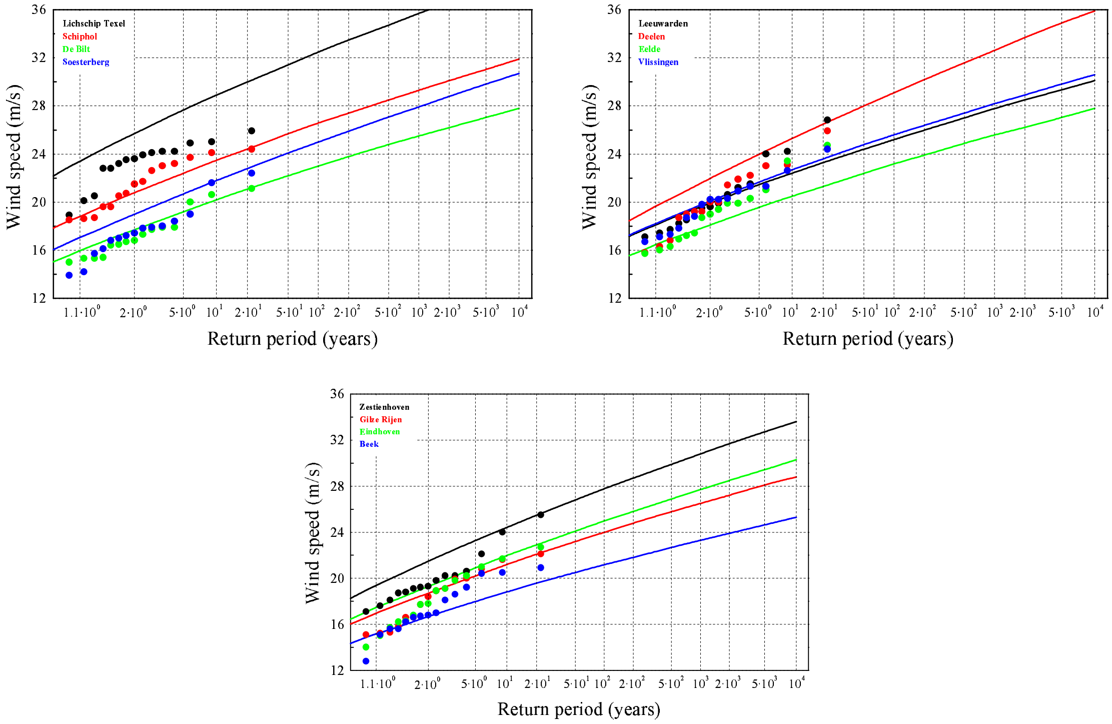

Figure D.7: Return periods and return levels following model G. The closed circles represent the annual maximum wind speed.

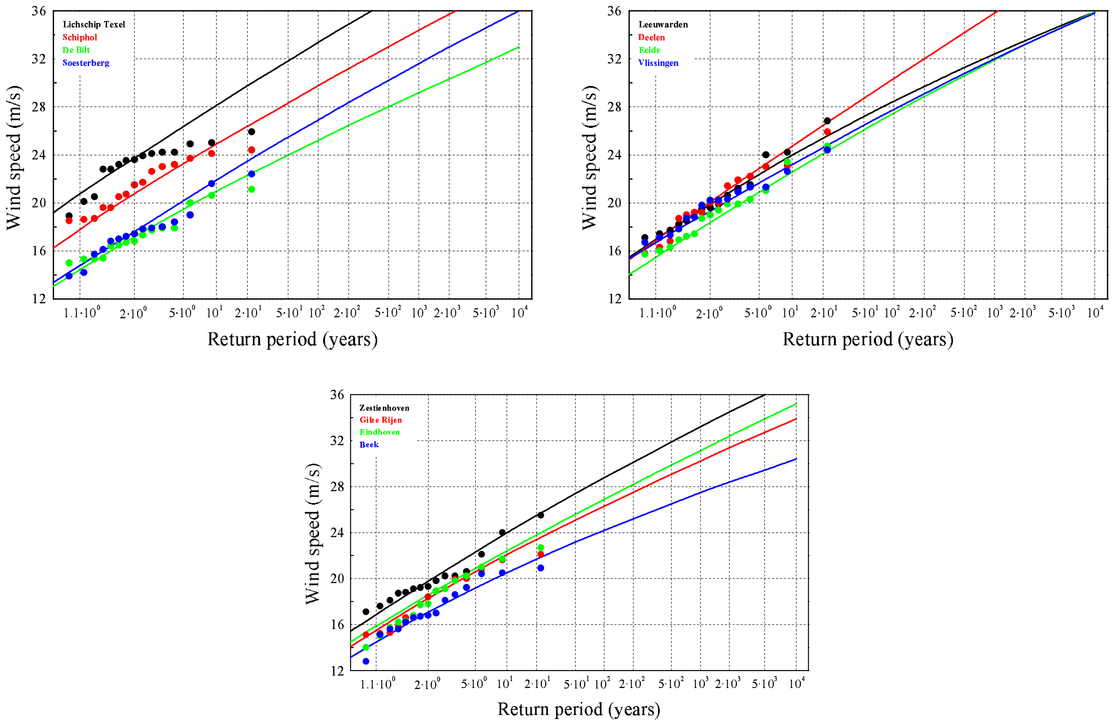

Figure D.8: Return periods and return levels following model H. The closed circles represent the annual maximum wind speed.

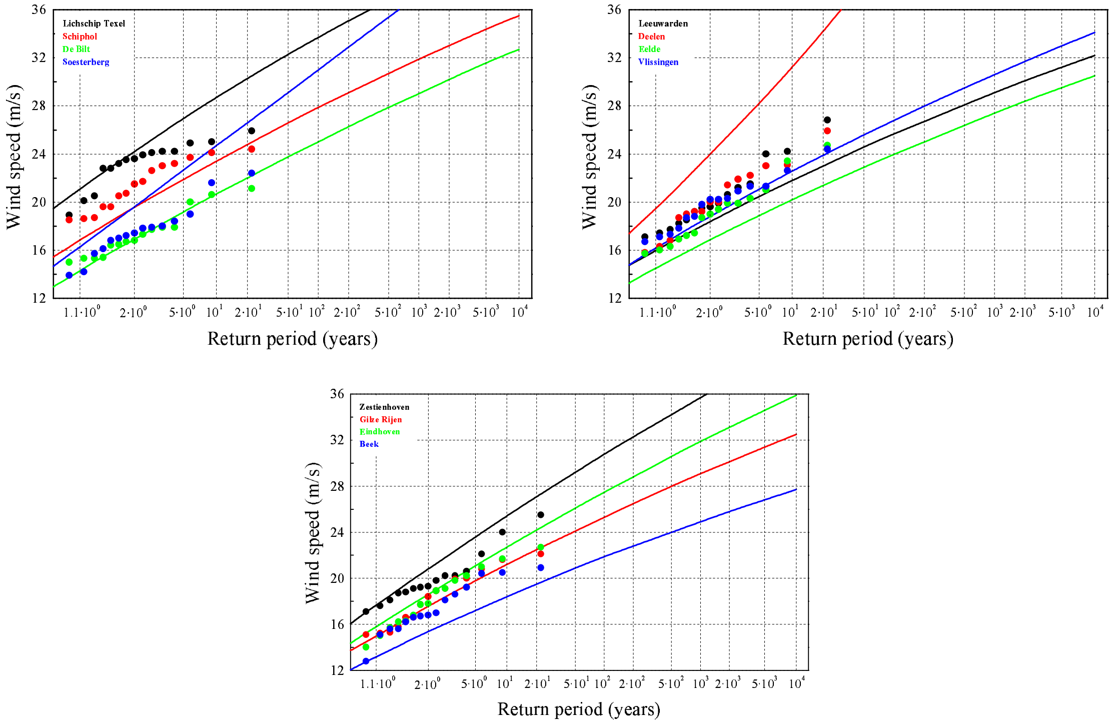

Figure D.9: Return periods and return levels following model I (the RW-model). The closed circles represent the annual maximum wind speed.

E Effect of class definition

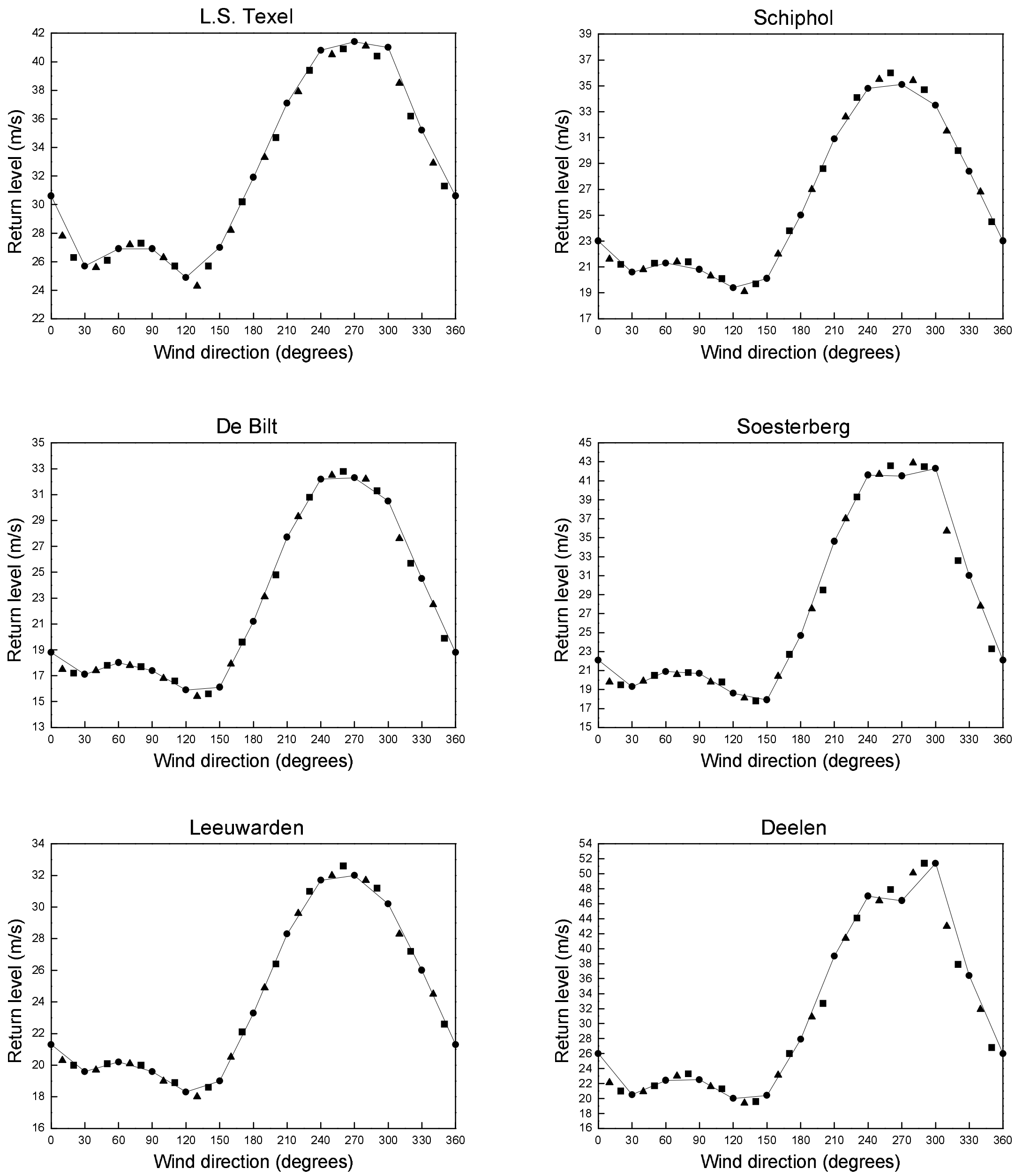

Figure E.1a: Effect of class definition of wind direction. The interpolated circles represent the 10,000 year return periods (Deelen: 1,000 year) calculated for each wind direction class used in the original RW-model (see Table 3.3 in Paragraph 3.1). The return periods represented by squares are calculated after shifting of 10 degrees to the left and the return periods represented by triangles are calculated after shifting of 10 degrees to the right.

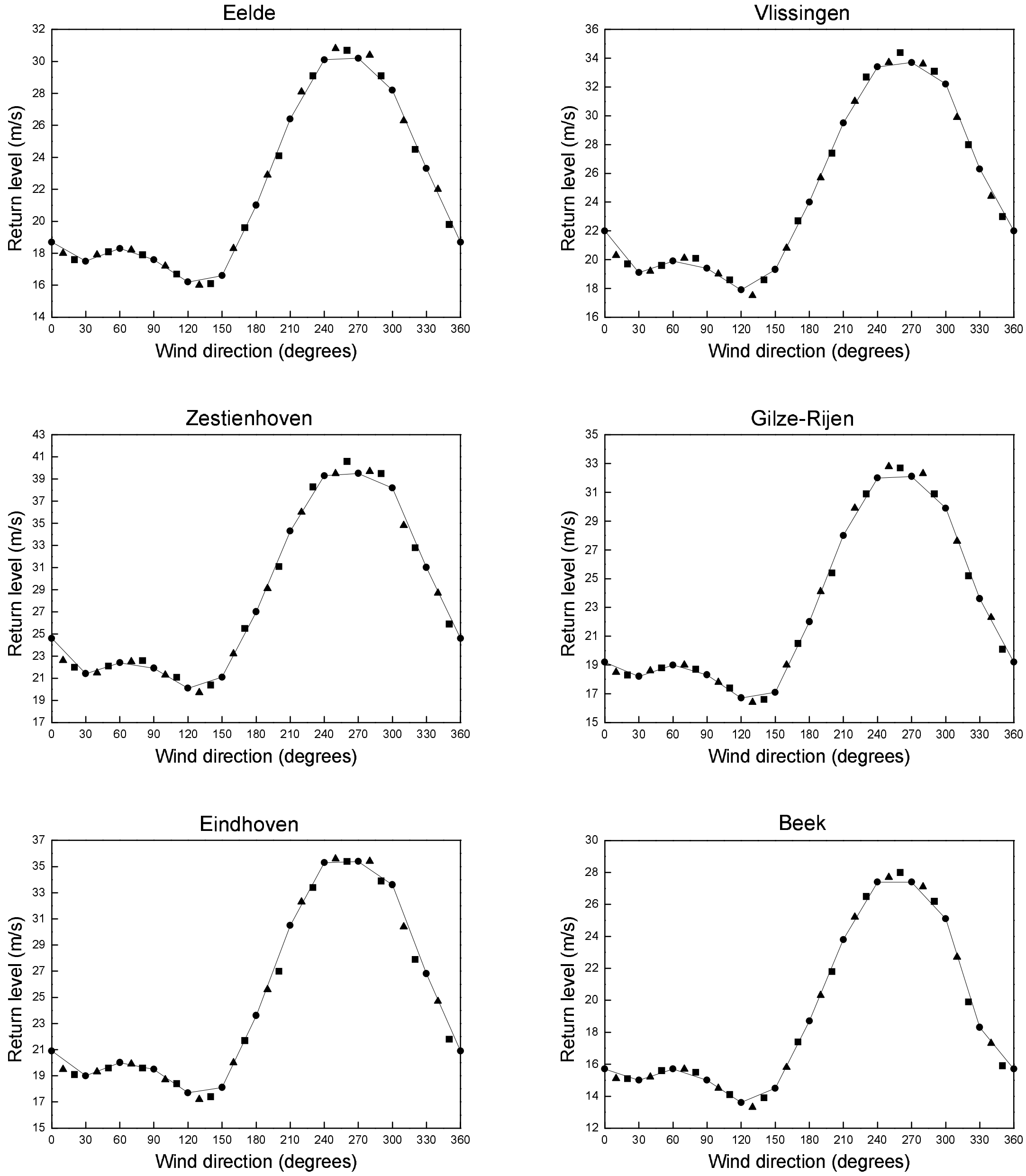

Figure E.1b: Continuation of Figure E.1a.

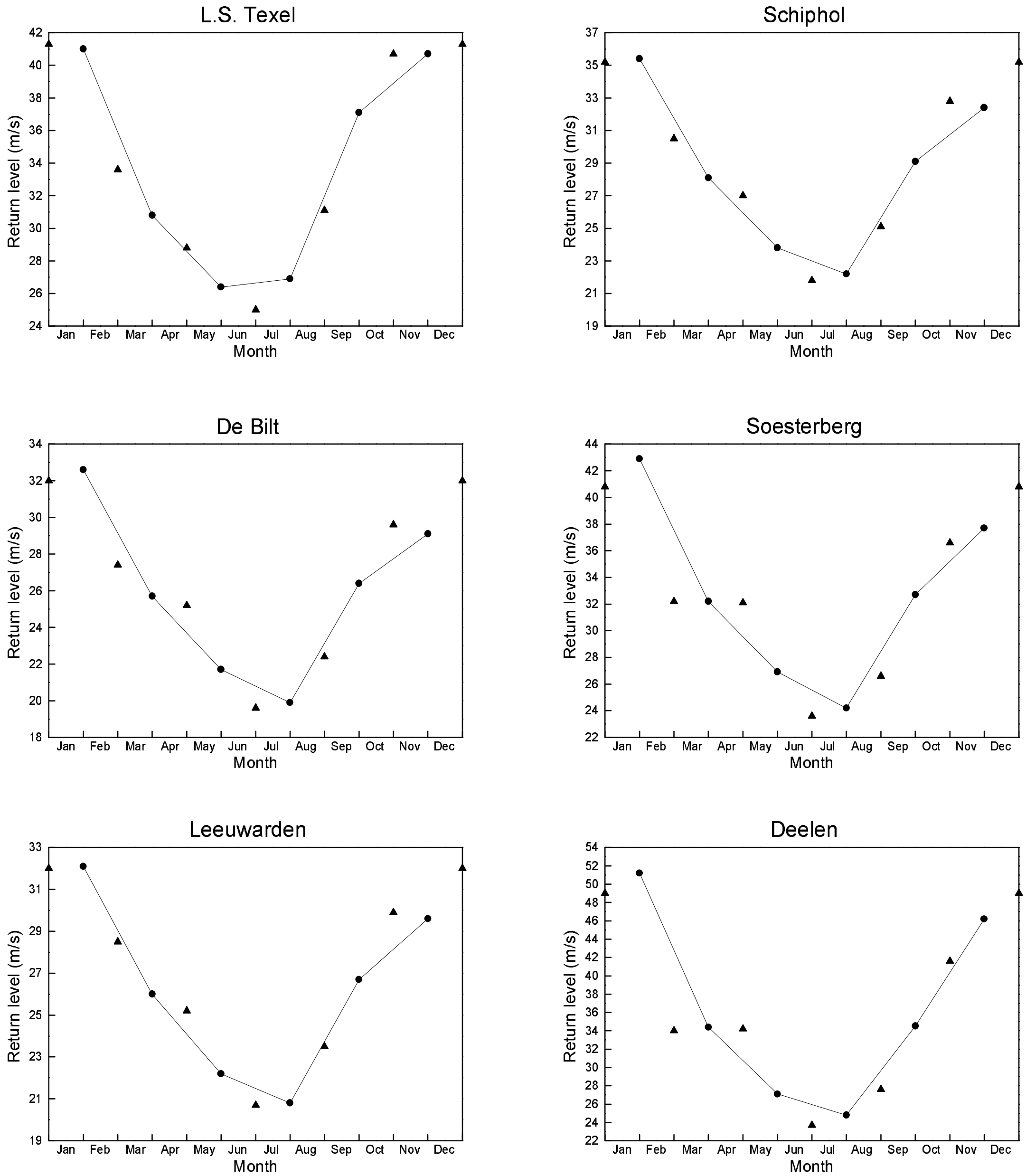

Figure E.2a: Effect of class definition of season. The interpolated circles represent the 10,000 year return periods (Deelen: 1,000 year) calculated for each season class used in the original RW-model (see Table 3.2 in Paragraph 3.1). The return periods represented by triangles are calculated after shifting of one month.

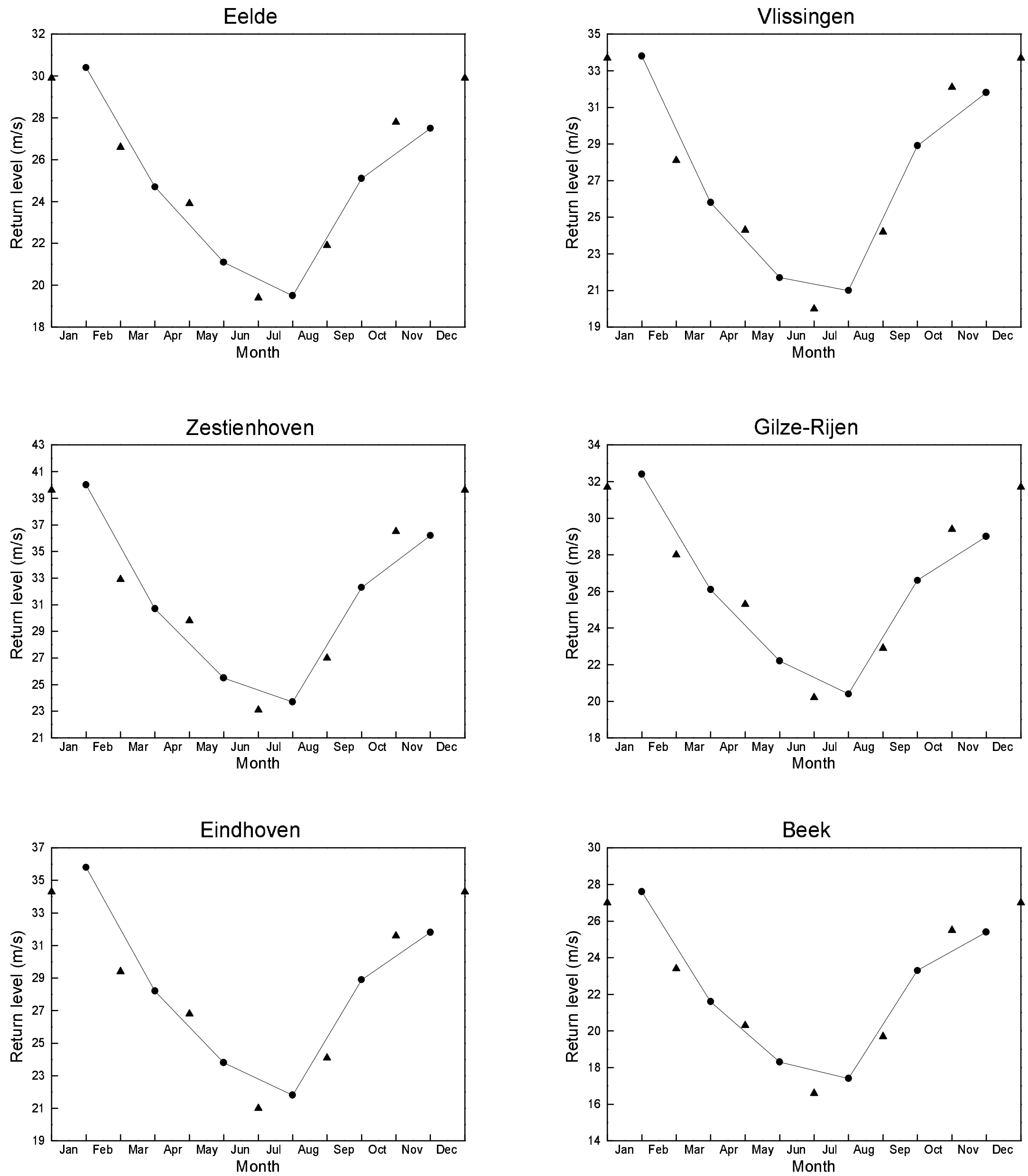

Figure E.2b: Continuation of Figure E.2a.