

Rapport  
aan  
Dienst der Kustwegen  
betreffende

## **ANALYSE VAN WINDSNELHEIDGEGEVENS METEOPAAL ZEEBRUGGE**

### **1. INLEIDING**

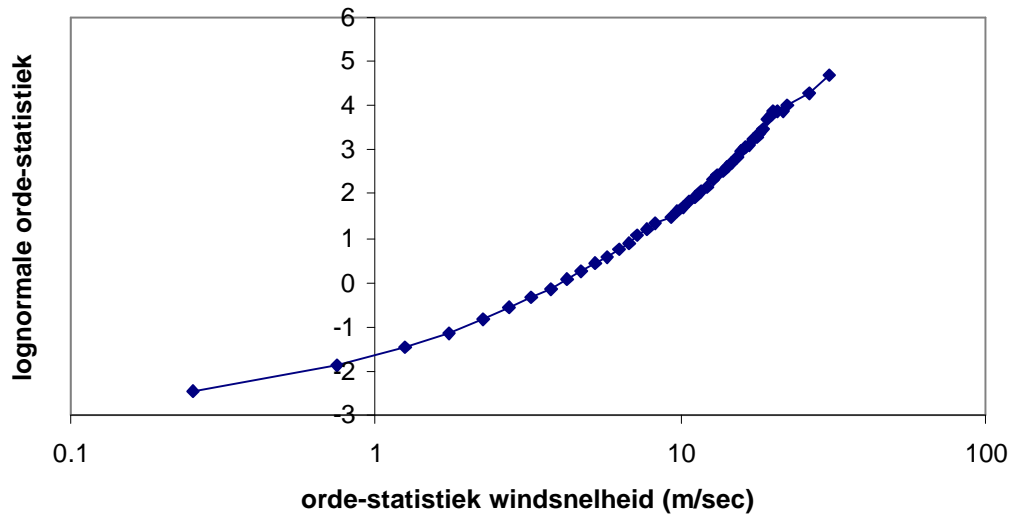
Dit rapport beschrijft de analyse resultaten van de windsnelheidgegevens voor de meteopaal Zeebrugge en de overeenkomstige voorstellingswijzen in de Meteopaal. Volgende punten zijn onderzocht:

1. de keuze van de verdelingsvorm (zie Hoofdstuk 2);
2. de schatting van de marginale verdelingsvorm (zie Hoofdstuk 3);
3. de gemiddelde overschrijdingsduur in functie van de drempelwaarde (zie Hoofdstuk 4);
4. een aanpassing van de ontwerpwaarden rekening houdend met de overschrijdingsduur (zie Hoofdstuk 5).

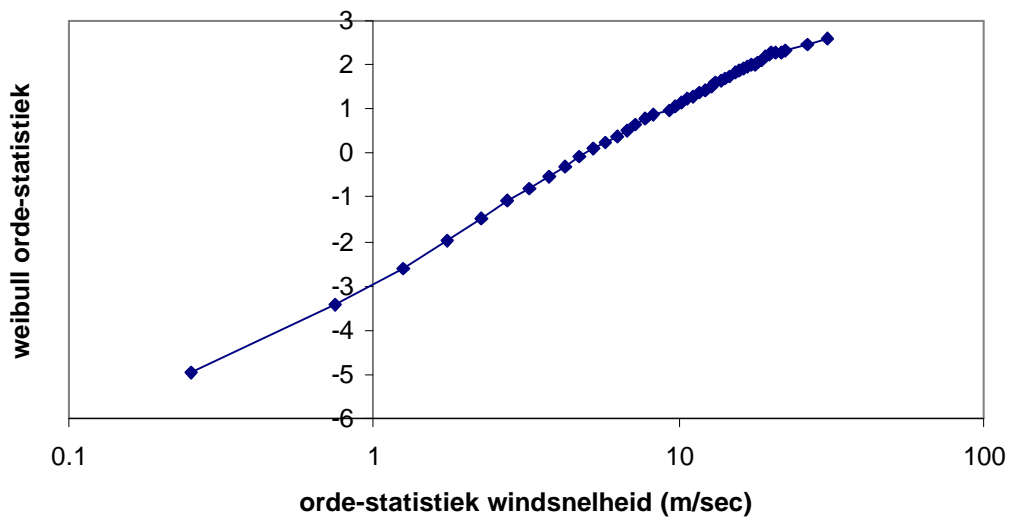
### **2. KEUZE VAN DE VERDELINGSVORM**

Voor windsnelheidsmetingen wordt veelal hetzij een lognormale, hetzij een Weibull verdeling verondersteld. Om een keuze te maken van de meest gepaste verdelingsvorm is de empirische verdeling van de windsnelheidsmetingen voor de periode van 1977 tot en met 1993 (meetwaarden elke 15 minuten) uitgezet zowel op lognormaal als op Weibull waarschijnlijkheidspapier. Deze voorstellingswijzen gaan uit van de empirische kwantielplot waarbij voor iedere orde-statistiek  $x_i$  (de  $i$ -de hoogste waarde) de proportie wordt uitgezet waarmee deze waarde wordt onderschreden, zoals bepaald door de Blom-score  $b=(i-0.5)/n$ , waarbij  $n$  verwijst naar het totaal aantal gegevens. Zowel de orde-statistiek in de  $x$ -as, als de Blom-score in de  $y$ -as wordt echter getransformeerd zodanig dat, indien de veronderstelde verdeling van toepassing is, de uitgezette curve naar een rechte lijn dient te convergeren. Volgende figuren tonen het resultaat van deze operatie, respectievelijk voor de veronderstelde lognormaal verdeling als voor de veronderstelde Weibull verdeling.

### Lognormale Q-Q plot



### Weibull Q-Q plot



Omdat het aantal gegevens bijzonder groot is (324713 metingen zijn uitgevoerd over deze periode) zijn de orde-statistieken gediscrèteerd in intervallen van 0.25 m/sec en de blom-score is bepaald voor de midpunten van deze klassen. Dit heeft echter geen verdere invloed op de interpretatie van het resultaat.

Het is zeer duidelijk dat de empirische Weibull Q-Q plot veel beter een rechte lijn benadert dan de lognormale Q-Q plot. Daarom wordt in volgend hoofdstuk een Weibull fit gebruikt om de marginale verdeling te benaderen.

### 3. FIT VAN DE MARGINALE VERDELING

Voor de windsnelheidsmetingen zijn metingen tot en met 1993 uitgevoerd met een 15 minuten interval, terwijl de metingen na 1993 zijn uitgevoerd met een 10 minuten interval. De verandering van het tijdsinterval waarover de metingen zijn bepaald kan de verdelingsvorm beïnvloeden en daardoor stelt het zich het combinatieprobleem van de resultaten bekomen door afzonderlijke analyse van de datasets. We bespreken eerst de resultaten voor de afzonderlijke analyses en vervolgens hoe men deze resultaten kan combineren om meer nauwkeurige schattingen te bekomen van de ontwerpwaarden.

#### 3.1. Schatting van de marginale verdeling voor één enkele dataset

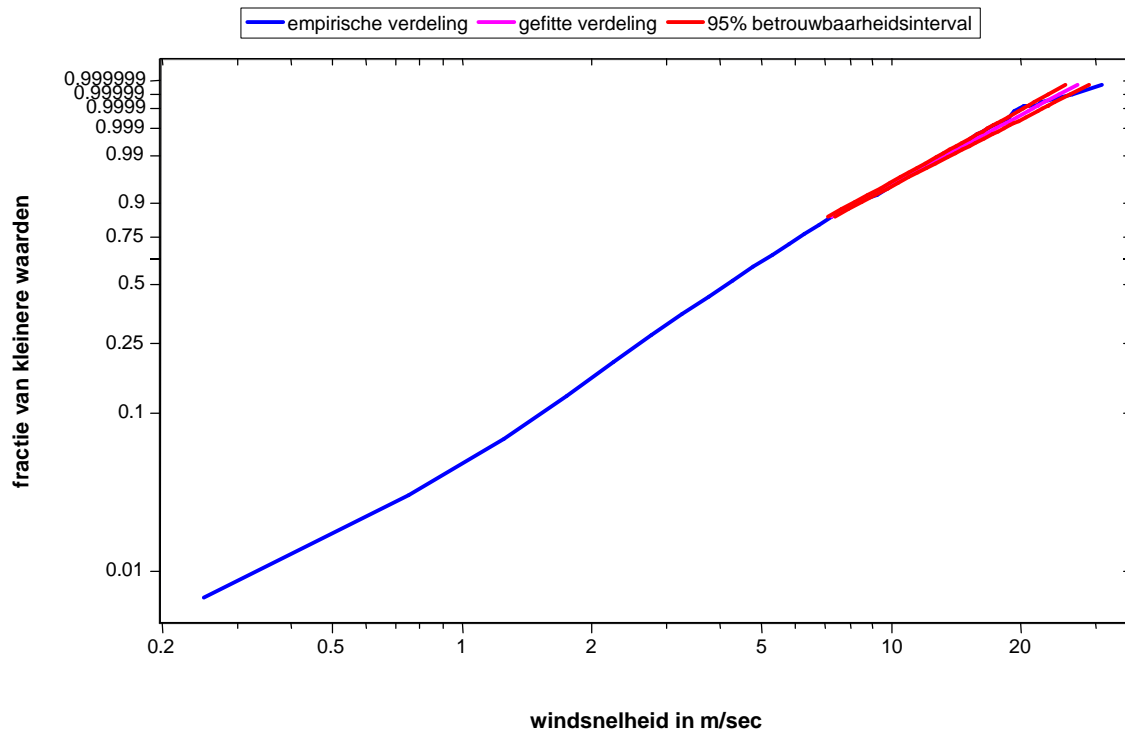
Een procedure voor de schatting van de Weibull verdeling op basis van de meetgegevens is eerder uitgewerkt voor de verwerking van significante golfhoogte. De methode die hiertoe is ontwikkeld wordt in detail uiteengezet in "Schatting van de kansverdeling van de significante golfhoogte". Specifieke karakteristieken van de methode die ook hier van toepassing zijn betreffen volgende aspecten:

- de Weibull verdeling wordt gefit voor een zo groot mogelijk bereik van de hogere orde-statistieken in zoverre een voldoende "goodness-of-fit" wordt bereikt. Voorgaande Weibull Q-Q plot toont dat ook voor de windgegevens het lineaire gedrag afwijkt bij de lagere orde-statistieken van dat bij de hogere orde-statistieken;
- de fitting gebeurt op basis van de ML-methode, waarbij correcties worden gemaakt voor de dataomvang in de verschillende seizoenen, zodat deze conform is met de werkelijke kalendertijd van deze seizoenen. Deze correctie is ook in dit geval van toepassing;
- de onzekerheid op de geschatte parameters en de overeenkomstige ontwerpwaarden wordt geschat door middel van een parametrische "bootstrap" (d.w.z. de schattingsmethode wordt herhaald toegepast op basis van gesimuleerde steekproeven uit een Weibull verdeling met de geschatte parameters). De steekproefomvang die wordt gebruikt bij deze bootstrap wordt echter voorafgaand bepaald op basis van de variatie van de schattingen die wordt vastgesteld bij de toepassing van een niet-parametrische "geblokte" bootstrap. Bij deze methode worden uit de originele meetreeks lukraak blokken van metingen (rond de 500 opeenvolgende metingen) geselecteerd, waarna de schattingsmethode wordt herhaald. Door de selectie van tijdsblokken van metingen eerder dan individuele metingen wordt de correlatie van de metingen doorbroken en krijgt men een minder vertekend beeld van de werkelijke onzekerheid. Vermits opeenvolgende metingen van de windsnelheid sterk gecorreleerd zijn, is deze techniek ook hier van toepassing.

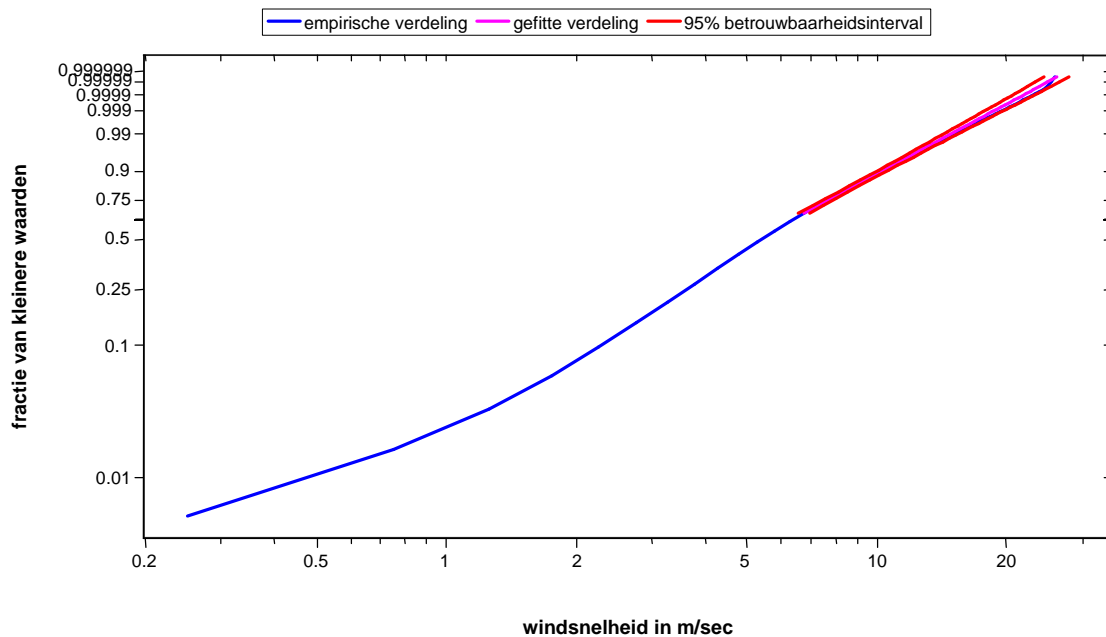
De schattingsmethode is toegepast zowel voor de dataset van windgegevens tot en met 1993 als na 1993. Selectieve resultaten worden hierna bondig besproken

### 3.1.1. Marginale kansverdeling

cumulatieve verdeling van de windsnelheid (gegevens tot en met 1993)



cumulatieve verdeling van de windsnelheid (gegevens na 1993)



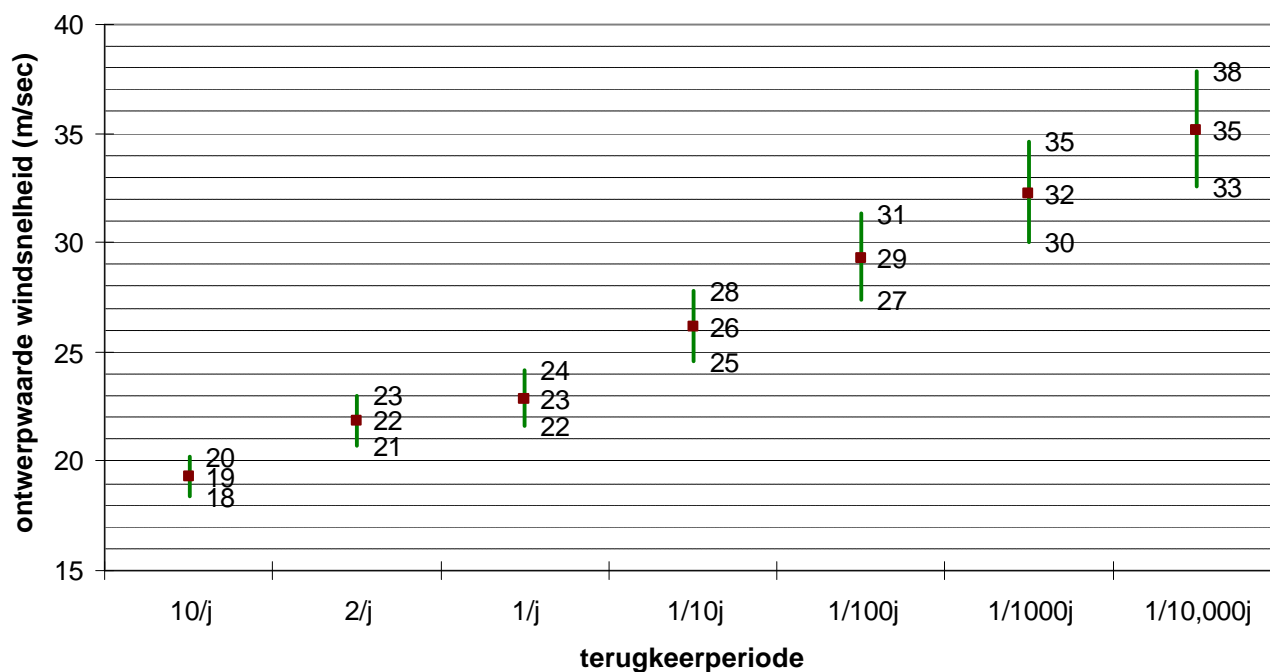
Plots van de empirische en gefitte Weibull-verdeling toont dat voor beide datasets een goede fit wordt bekomen vanaf een drempelwaarde van respectievelijk 7 m/sec (voor data tot en met 1993) en 6.5 m/sec (voor data na 1993).

### 3.1.2. Ontwerpwaarden

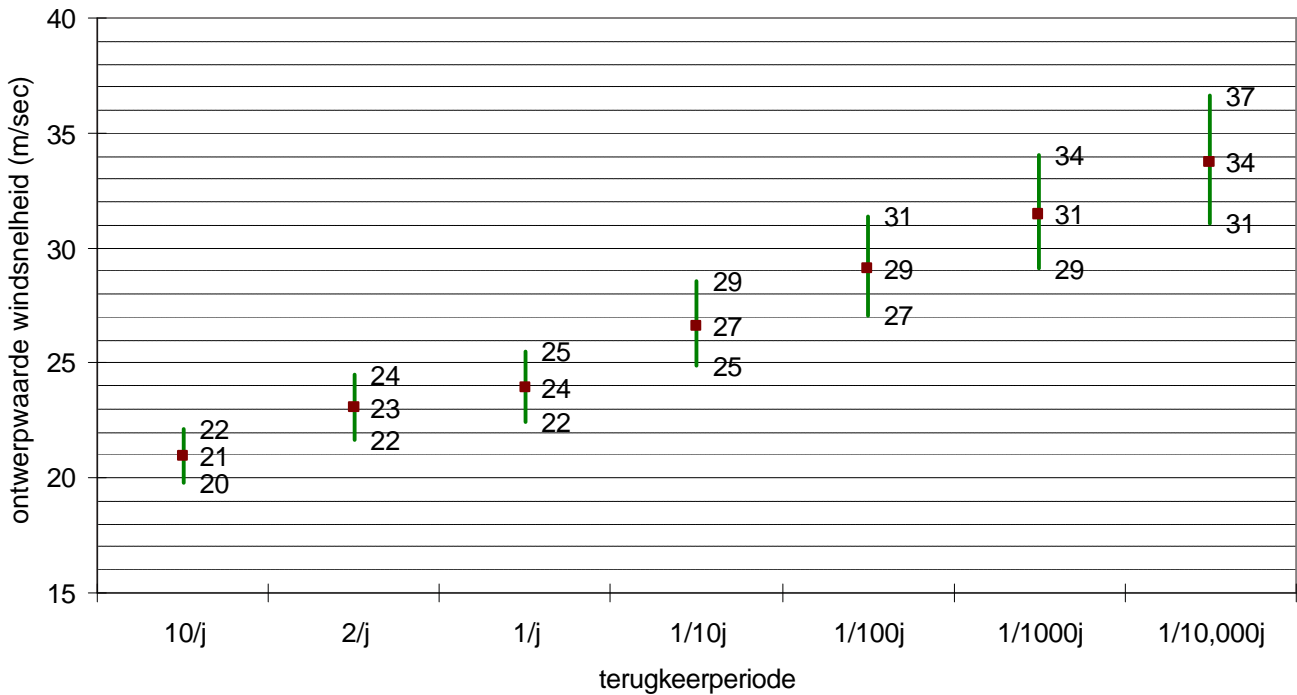
Op basis van de marginale kansverdelingen kunnen de windsnelheden berekend worden die voor een bepaalde “terugkeerperiode” gemiddeld 1 maal worden overschreden. Bij de interpretatie van dit resultaat dient men echter rekening te houden met de tijdsduur van de overschrijding. Voor zeer hoge waarden komt dit gemiddeld overeen met het tijdsinterval waarover wordt gemeten; voor lage waarden doen overschrijdingen zich echter meestal in groepen voor. Dit probleem wordt verder besproken in Hoofdstuk 4 bij de analyse van de overschrijdingsduren.

Volgende figuren tonen het resultaat van de ontwerpwaarde schatting op basis van de twee datasets.

**meetwaarden tot en met 1993**



meetwaarden na 1993



Bij de vergelijking van deze resultaten dient men echter rekening te houden met het verschil in tijdsinterval. Zo vindt men bijvoorbeeld op basis van de analyse van de meetwaarden tot en met 1993 dat de windsnelheid van 26 m/sec gemiddeld 15 minuten wordt overschreden per 10 jaar, terwijl men op basis van de analyse van de meetwaarden na 1993 vindt dat de windsnelheid van 27 m/sec gemiddeld 10 minuten wordt overschreden per 10 jaar.

Hoe men tot een rechtstreekse vergelijking van de resultaten kan komen en de resultaten ook kan combineren tot een meer nauwkeurige schatting die gebruik maakt van beide datasets wordt besproken in Sectie 3.2

### 3.1.3. Enkele numerische resultaten

Volgende tabel vergelijkt enkele numerische resultaten voor de beide datasets.

RESULTAAT	TOT EN MET 1993	NA 1993
totaal aantal meetpunten	324712	278602
equivalent aantal onafhankelijke meetgegevens	6390	2269
drempelwaarde	7	6.5
aantal meetpunten gelijk aan of lager dan de drempelwaarde	273519.3	181437.5
geschatte waarde van u	4.6	6.3
geschatte standaardfout van u	0.1	0.1
geschatte waarde van k	1.47	1.79
geschatte standaardfout van k	0.037	0.054
geschatte correlatie tussen u en k	0.81	0.64
G-statistiek voor de fit	4.65	5.28
significantiëniveau van de G-statistiek	0.03	0.63
gemiddelde van log(u) in de bootstrap methode	1.53	1.84
standaarddeviatie van log(u) in de bootstrap methode	0.016	0.017
gemiddelde van 1/k in de bootstrap methode	0.68	0.56
standaarddeviatie van 1/k in de bootstrap methode	0.017	0.017
correlatie van log(u) en 1/k in de bootstrap methode	-0.81	-0.64
TIJDSCATEGORIE		
correctie voor januari-maart	1.04	1.13
correctie voor april-juni	1.02	1.07
correctie voor juli-september	0.93	0.92
correctie voor oktober-december	1.00	0.91

Hoewel de observatieperiode van 1977 tot en met 1993 aanzienlijk hoger is dan die na 1993 is het aantal meetpunten toch van dezelfde orde grootte omdat na 1993 een kleiner tijdsinterval werd gebruikt. Het equivalent aantal onafhankelijke meetgegevens wordt echter door de schattingsmethode op correcte wijze als aanzienlijk hoger ingeschat voor de eerste dataset dan voor de tweede dataset. De drempelwaarde die automatisch wordt gekozen door de schattingsmethode is, zoals eerder vermeld, in beide gevallen vergelijkbaar. Het significantieniveau van de G-statistiek toont echter dat voor de eerste dataset een minder goede fit wordt bekomen dan voor de tweede dataset: de waarde 0,03 duidt aan dat een waarde van de G-statistiek kleiner of gelijk aan de genoteerde waarde van 4,65 zich slechts in 3% van de onafhankelijke schattingen zou voordoen indien de schattingsmethode werd herhaald op basis van onafhankelijke steekproeven uit een Weibull verdeling met de geschatte parameters (met een steekproefomvang gelijk aan het equivalent aantal meetpunten). Onderaan de tabel worden de correcties aangeduid voor de verschillende tijdscategoriën. Meetpunten in de zomerperiode zijn oververtegenwoordigd en worden in beide datasets verminderd. Hoewel de correcties niet bijzonder groot zijn, zijn ze toch niet te verwaarlozen voor het eindresultaat.

De schattingen van de parameters u en k voor de twee datasets blijken opvallend verschillend te zijn. Bij deze beoordeling dient men echter rekening te houden met de

correlatie tussen de waarden u en k die bijzonder hoog is. Overigens is het duidelijk uit voorgaande plots van de ontwerpwaarden dat, hoewel de schattingen van u en k significant verschillen, de schattingen van de ontwerpwaarden slechts geringe verschillen vertonen.

### 3.2. Vergelijking en combinatie van de ontwerpwaarde resultaten

Om de resultaten bekomen op basis van de datasets dienen de resultaten eerst en vooral te worden herberekend op basis van dezelfde tijdstap. Dit kan eenvoudig gebeuren door de kans van overschrijding die overeenkomt de N-jaarlijkse ontwerpwaarde te herberekenen als:

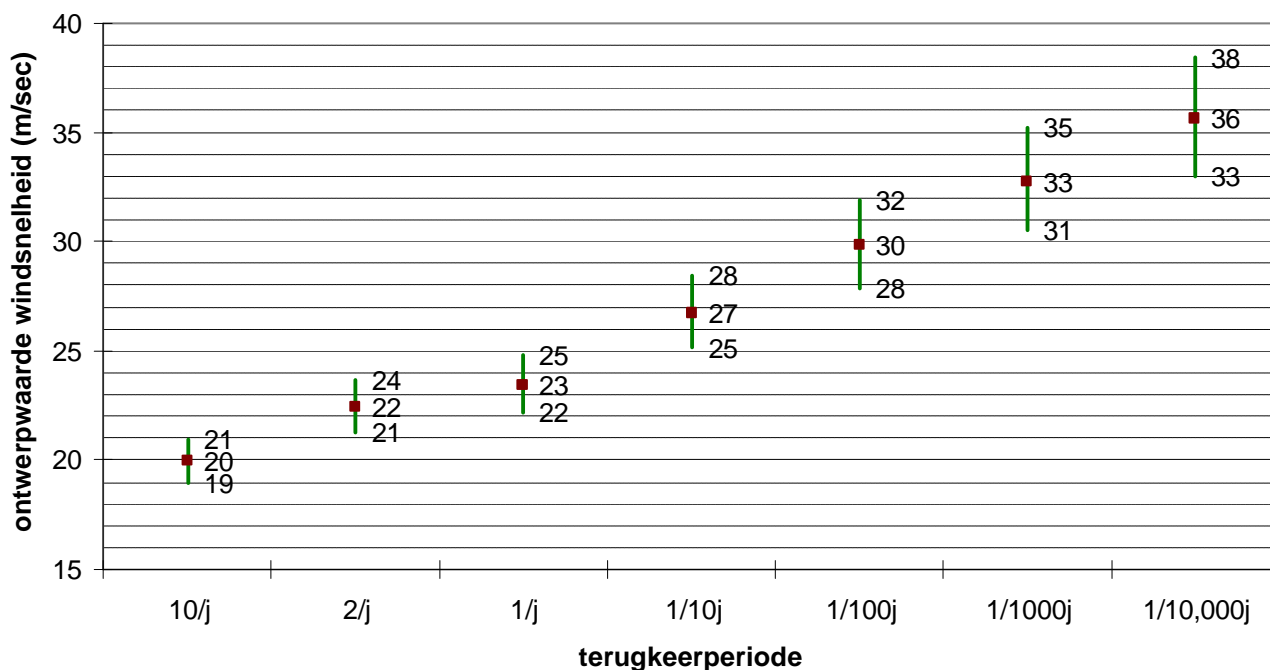
$$p = \frac{\Delta t}{N} \quad (1)$$

waarbij  $\Delta t$  de tijdsduur van de overschrijding voorstelt uitgedrukt in jaren.

Voor deze overschrijdingskansen kunnen dan de gebruikelijke formules worden toegepast om de overeenkomstige ontwerpwaarde en zijn onzekerheid af te leiden (zie rapport "Schatting van HS").

De volgende figuur toont het resultaat van deze nieuwe berekening wanneer men  $\Delta t$  gelijkstelt aan 10 minuten voor de gegevens gemeten voor 1993 (met een oorspronkelijke tijdstap van 15 minuten).

#### meetwaarden tot en met 1993 (voor 10 minuten)





De resultaten zijn lichtjes hoger dan degene die eerder werden gevonden voor een overschrijdingsduur van 15 minuten.

De resultaten van de twee analyses kunnen nu gecombineerd worden door de onafhankelijke schattingen te combineren rekening houdend met hun respectieve nauwkeurigheid. Dit gebeurt door een gewogen gemiddelde te maken van het logaritm van de respectieve schatters. De gewichten zijn invers proportioneel met de varianties van dit logaritm. Mathematisch betekent dit dat:

$$\log \hat{h}_N = w_1 \log \hat{h}_{N1} + w_2 \log \hat{h}_{N2} \quad (2)$$

waarbij  $\hat{h}_N$  de nieuwe gecombineerde schatting voorstelt en  $\hat{h}_{N1}, \hat{h}_{N2}$  de respectieve schattingen voorstellen voor de afzonderlijke datasets.  $w_1$  en  $w_2$  worden bepaald als:

$$w_1 = \frac{1}{\sigma_{\log \hat{h}_{N1}}^2} \quad \text{en} \quad w_2 = \frac{1}{\sigma_{\log \hat{h}_{N2}}^2} \quad (3)$$

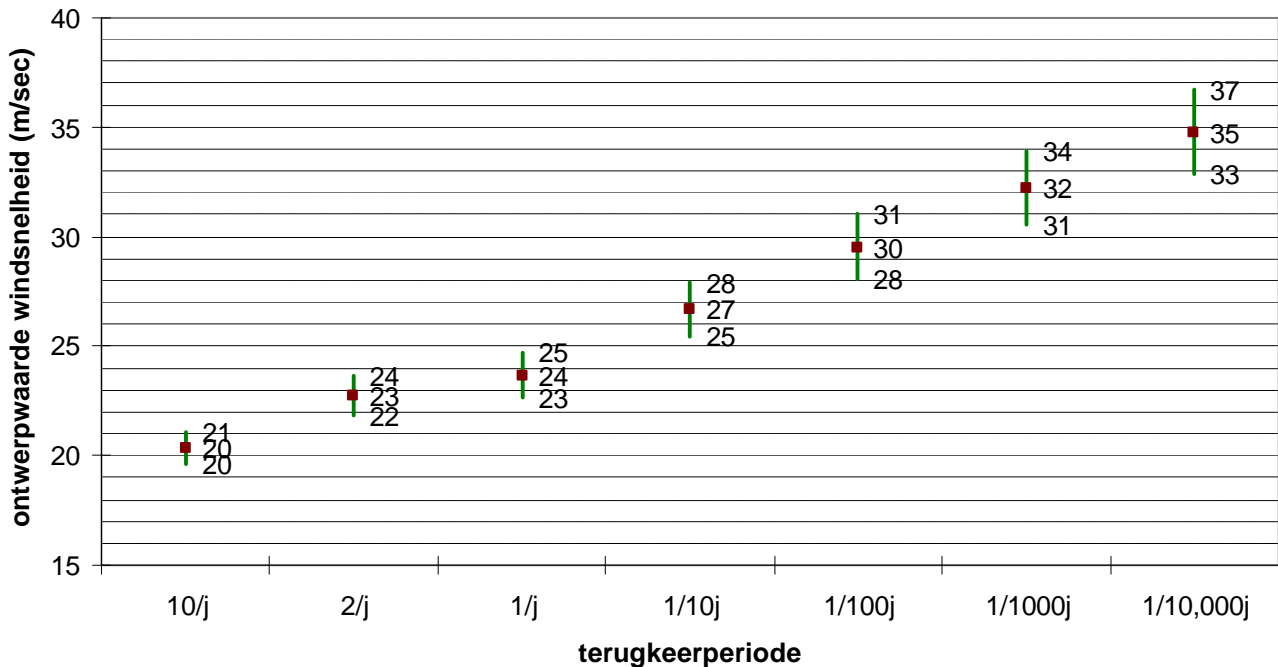
waarbij  $\sigma_{\log \hat{h}_{N1}}^2$  en  $\sigma_{\log \hat{h}_{N2}}^2$  verwijzen naar de variantie van de afzonderlijke schattingen na logtransformatie.

De variantie van het logaritm van de gecombineerde schatting komt dan overeen met:

$$\sigma_{\log \hat{h}_N}^2 = \frac{1}{w_1 + w_2} \quad (4)$$

De volgende figuur toont het eindresultaat van de combinatie van de twee analyses: de ontwerpwaarden liggen nu tussen beide schattingen in, waarbij de eerste dataset een iets groter gewicht wordt toegekend, vermits deze resultaten nauwkeuriger zijn.

### alle meetwaarden (voor 10 minuten)



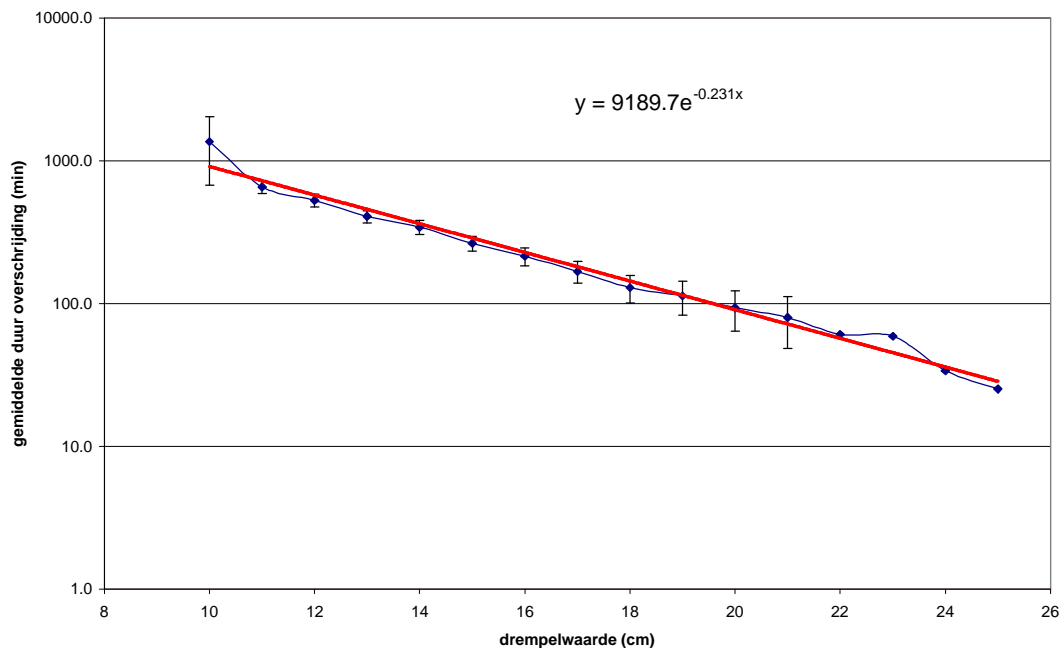
Merken we nogmaals op dat deze “ontwerpwaarden” betrekking hebben op de windsnelheid die gemiddeld 10 minuten per N jaar worden overschreden. In volgend hoofdstuk wordt de gemiddelde overschrijdingsduur onderzocht voor een gegeven drempelwaarde van de windsnelheid. Deze analyse laat toe om ontwerpwaarden af te leiden die betrekking hebben op onafhankelijke overschrijdingen die voor een gegeven jaar met kans  $1/N$  voorkomen..

#### 4. OVERSCHRIJDINGSDUUR VAN DE WINDSNELHEID

De overschrijdingsduur van de windsnelheid wordt op dezelfde wijze bepaald als voor de golfhoogte (zie rapport "Overschrijdingsduur van significante golfhoogte"). Voor verschillende drempelwaarden worden eerst "storm" periodes bepaald die overeenkomen met een continue overschrijding van de windsnelheid. Vervolgens worden deze stormen samengevoegd, indien ze binnen een korte tijdspanne van elkaar gebeuren. Tenslotte wordt voor iedere drempelwaarde de gemiddelde duur van de overschrijding van het drempelniveau binnen een "storm" periode berekend.

Voor de uitvoering van deze analyses kunnen de "storm" periodes van de beide observatieperiodes worden samengevoegd. Het resultaat voor de gemiddelde overschrijdingsduur voor deze samengevoegde resultaten wordt getoond in de volgende figuur. Voor drempelwaarden waarbij een voldoende aantal stormen wordt gedetecteerd is eveneens de standaardfout van de schatting van het gemiddelde

Meteostation Zeebrugge, W15+W11, Windmetingen (15 min), WVS (1977/01-1998/12)



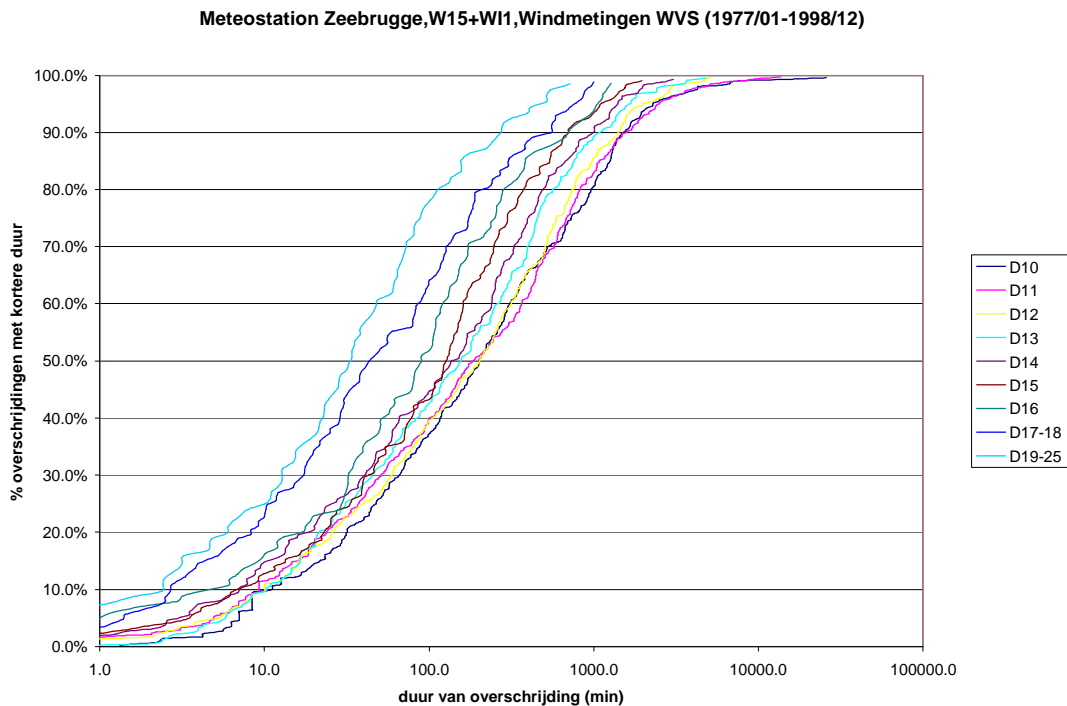
aangeduid. De rechte lijn komt overeen met een kleinste kwadratenfit van de waarden.

Het is onmiddellijk duidelijk dat het verband tussen de gemiddelde overschrijdingsduur en de drempelwaarde zeer goed benaderd wordt door een exponentieel dalende curve van de vorm:

$$E(T) = 9187.9 \times e^{-0.231V} \quad (5)$$

waarbij T verwijst naar de overschrijdingsduur en V verwijst naar de beschouwde drempelwaarde.

Men dient echter op te merken dat de variatie van de overschrijdingsduur voor een gegeven drempelwaarde bijzonder hoog kan zijn. Dit wordt geïllustreerd door de volgende figuur, die voor de verschillende drempelwaarden de cumulatieve frequentieverdeling van de stormduren toont (voor drempelwaarden waarbij slechts enkele stormen worden gedecteerd, zijn verschillende klassen samengevoegd).



Een formele analyse van deze gegevens is niet uitgevoerd binnen het bestek van deze studie. Voorgaande plot suggereert echter een lognormale verdeling voor de overschrijdingsduur bij een gegeven drempelwaarde. Verder blijkt dat de verdelingsvorm quasi-identiek is voor de verschillende drempelwaarden en enkel de mediaanwaarde verandert. De spreiding is echter bijzonder hoog: typisch vallen 90% van de waarnemingen tussen 1/10 en 10x de mediaanwaarde.

## 5. AANPASSING VAN DE ONTWERPWAARDEN VOOR DE VERWACHTTE OVERSCHRIJDINGSDUUR

Voorgaande analyse van de overschrijdingsduur toont dat de gemiddelde overschrijdingsduur sterk varieert in functie van de overschreden drempelwaarde. Voor een windsnelheid van 20 m/sec vindt men bijvoorbeeld een gemiddelde van 1,5 uur, terwijl men voor een windsnelheid van 29,5 m/sec slechts een gemiddelde duur van 10 minuten zou verwachten.

Het is duidelijk dat voor de lagere waarden de aanname van onafhankelijke overschrijdingen niet opgaat (gemiddeld zijn er 9 opeenvolgende metingen die de waarde van 20 m/sec overschrijden). De proportie van marginale overschrijdingen van  $\Delta t$  meetwaarden tijdens een periode van N jaar tijdens onafhankelijke storm overschrijdingen komt nu overeen met:

$$p = \frac{E(T(V))}{N} \quad (6)$$

waarbij  $E(T(V))$  verwijst naar de verwachtingswaarde van de overschrijdingsduur.

Vergelijking (6) kan gebruikt worden om voor iedere N-jaar terugkeerperiode terugkeerperiode de waarde van V te zoeken waarbij p voldoet aan vergelijking (6). Alternatief kan men stellen dat men in Vergelijking (1) de waarde van  $\Delta t$  dient te zoeken, die een ontwerpwaarde V oplevert waarbij  $\Delta t = E(T(V))$ . Merk op dat in deze berekening gebruikt gemaakt wordt van de eerder besproken gecombineerde oplossing voor de ontwerpwaarden.

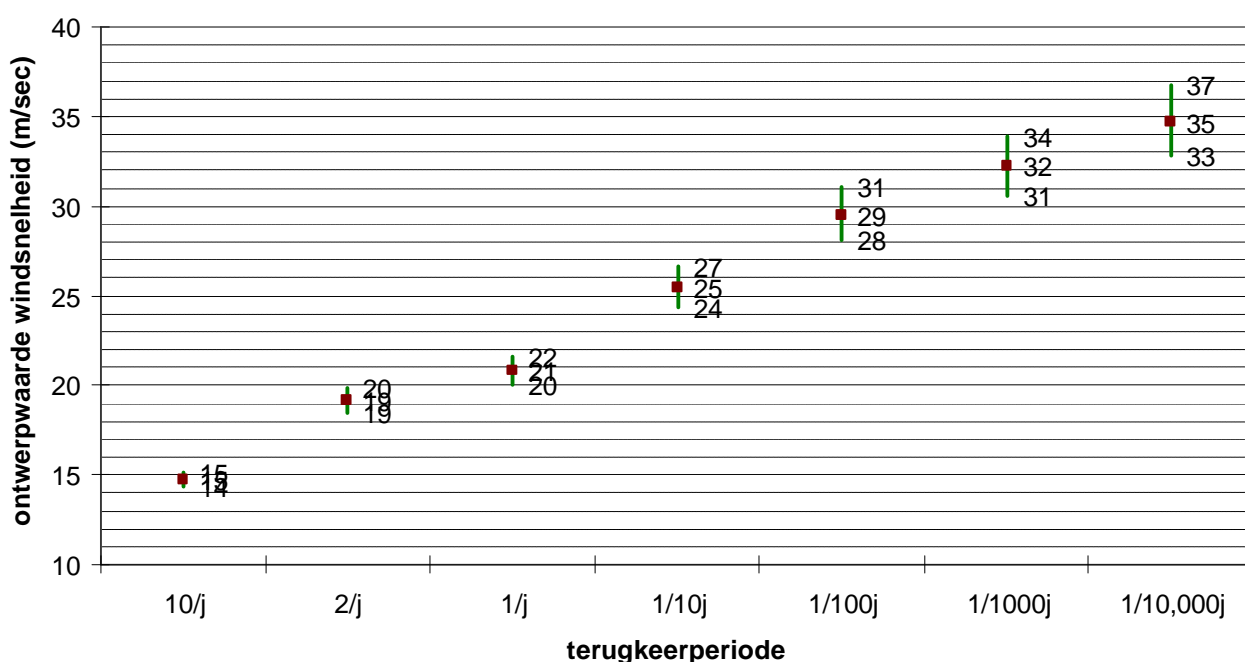
De numerische oplossing gebeurt door gebruik te maken van de "SOLVER" optie in EXCEL. Volgende tabel toont het resultaat:

terugkeer- periode	ontwerpwaarde (m/sec)			gemiddelde overschrijdingsduur (min)
	95% beneden- grens	beste schatting	95% boven- grens	
10/j	14	15	15	306
2/j	19	19	20	110
1/j	20	21	22	76
1/10j	24	25	27	26
1/100j	28	29	31	10
1/1000j	31	32	34	10 (5)
1/10,000j	33	35	37	10 (3)

Voor de kleinere terugkeerperiodes is er een substantiële verlaging van de ontwerpwaarden wanneer men rekening houdt met de overschrijdingsduur. Voor de hogere terugkeerperiodes (1/1000j en 1/10,000j), waarbij de verwachte overschrijdingsduur kleiner is dan 10 minuten, worden de vorige resultaten behouden omdat het weinig zinvol is om over overschrijdingsduren kleiner dan het meetinterval te spreken.

Volgende figuur toont visueel het eindresultaat.

### ontwerpwaarden gecorrigeerd voor overschrijdingsduur



Zoals reeds eerder is uiteengezet in het rapport betreffende de schatting van de golfhoogte dient men deze ontwerpwaarden op een andere wijze te interpreteren dan de ontwerpwaarden gebaseerd op basis van de marginale kansverdelingen. De nieuwe waarden verwijzen naar onafhankelijke overschrijdingen en zijn dan ook van toepassing voor de bepaling van een kans van overschrijding binnen één jaar, waarbij men vervolgens rekening dient te houden met het feit dat de overschrijding een bepaalde duur heeft. De ontwerpwaarden van de marginale kansverdeling daarentegen verwijzen naar een gemiddelde periode van overschrijding ( $\Delta t/N$ ) per jaar, maar houden geen rekening met de typische groepering van deze overschrijdingen.