

Rapport

aan

Minsterie van de Vlaamse Gemeenschap  
Departement Leefmilieu en Infrastructuur  
Aministratie Waterwegen en Zeewezen

AFDELING WATERWEGEN KUST

**POT ANALYSE VAN SIGNIFICANTE GOLFHOOGTE  
EN SCHATTING VAN ONTWERPWAARDEN**

- 14 Mei, 2000 -

# **POT ANALYSE VAN SIGNIFICANTE GOLFHOOGTE EN SCHATTING VAN ONTWERPWAARDEN**

## **1. INLEIDING**

De bepaling van de frequentie van voorkomen van extreme significante golfhoogten kan gebeuren hetzij door een kansverdeling te schatten voor alle gemeten waarden, hetzij door een kansverdeling te schatten enkel voor de meest extreme gemeten waarden. In dit rapport wordt deze laatste methode beschreven, waarbij enkel de piekwaarden van de significante golfhoogte die een bepaalde drempelwaarde overschrijden worden weerhouden. Het voordeel van de analyse van zulke piekwaarden (een zogenaamde POT of “peak over treshold” analyse) is dat de verschillende meetwaarden als onafhankelijke waarnemingen kunnen worden beschouwd en dat, op basis van resultaten van extreme waarden theorie, een welbepaalde vorm van de kansverdeling wordt verwacht.

Dit rapport beschrijft hoe de POT analyse wordt uitgevoerd. Het rapport is gestructureerd als volgt:

- Hoofdstuk 2 detailleert hoe de piekwaarden worden bepaald op basis van de oorspronkelijke tijdreeks;
- Hoofdstuk 3 bespreekt de schattingsmethode die is toegepast om ontwerpwaarden voor de significante golfhoogte te bepalen;
- Hoofdstuk 4 toont resultaten van de schattingsmethode voor verschillende meetstations.

## 2. EXTRACTIE VAN PIEKWAARDEN

Extractie van de piekwaarden van de significante golfhoogte HS gebeurt in volgende 3 stappen:

1. de tijdreeks van de gemeten waarden van HS wordt samengevat in overschrijdingsperiodes ("stormen") waarbij de hoogste waarde tijdens die overschrijding wordt genoteerd. Dit gebeurt voor verschillende overschrijdingsniveaus;
2. voor ieder overschrijdingsniveau worden de stormen gegroepeerd, om correlatie te vermijden tussen de verschillende piekwaarden. Enkel volledig gemeten stormen en hun overeenkomstige piekwaarden worden weerhouden;
3. voor het overschrijdingsniveau gekozen in de analyse worden de piekwaarden weerhouden en wordt een correctie gemaakt van de observatieperiode die rekening houdt met onvolledig gemeten stormen.

In de volgende paragrafen wordt ieder van deze stappen verder uiteengezet.

### 2.1. Extractie van Stormen en Piekwaarden

De extractie van stormen gebeurt op dezelfde wijze als in de studie van de overschrijdingsduur van stormen. De procedure wordt hier nogmaals samengevat.

Een storm wordt in deze studie beschouwd als een periode waarbij HS continu in de tijd een bepaalde waarde overschrijdt. Voor een gekozen overschrijdingsniveau  $x$ , wordt de tijdreeks HS doorlopen in functie van de tijd  $t$ . Wanneer HS op tijdstip  $t$  de waarde  $x$  overschrijdt terwijl de voorgaande waarde van HS lager is dan  $x$ , dan wordt een nieuwe storm genoteerd. Indien de voorgaande meting meer dan 3 uur voorafgaat, dan wordt de storm beschouwd als onvolledig gemeten, omdat er mogelijk andere waarden van HS die tot dezelfde storm behoren voorafgaan. Na de aanvang van een nieuwe storm worden de daaropvolgende waarden van HS verder doorlopen en wordt de maximum waarde van HS genoteerd. Het einde van de storm valt samen met het tijdstip waarop HS groter is dan  $x$ , terwijl de daaropvolgende waarde kleiner is dan  $x$ . Indien de volgende meting slechts na meer dan 3 uur gebeurt wordt de storm opnieuw als onvolledig genoteerd.

Deze procedure wordt herhaaldelijk toegepast op de meetreeks gebruikmakend van verschillende overschrijdingsniveaus  $x$ . Voor zeer lage waarde van  $x$  vindt men slechts enkele stormen vermits HS steeds de waarde  $x$  overschrijdt. Naarmate  $x$  stijgt, neemt het aantal stormen toe om vervolgens terug te dalen voor hogere waarden van  $x$ . Het is enkel in deze tweede fase dat men werkelijk van "stormen" kan

spreken. In de toepassing voor deze studie zijn overschrijdingsniveaus beschouwd startend vanaf 2 meter met stappen van 10 cm. Typisch bevindt men zich dan net in de dalende fase (afname van het aantal stormen met de waarde x).

## 2.2. Groeperen van Stormen

De lijst van stormen die op voorgaande wijze wordt bepaald leidt tot overschrijdingen die vaak vlak na elkaar gebeuren, onderbroken door onderschrijdingen van korte duur. Daarom worden de individuele overschrijdingen sequentieel doorlopen (voor ieder overschrijdingsniveau x) en verder gegroepeerd als volgt: indien de tijd tussen het einde van een overschrijding en het begin van de volgende overschrijding minder bedraagt dan 1 dag of de duur van de (tot dan toe) samengevoegde storm, dan wordt die overschrijding beschouwd als deel van de voorgaande storm. In zulk geval wordt de overschrijding toegevoegd aan de voorgaande storm en het maximum van de piekwaarden in voorgaande storm en de huidige overschrijding weerhouden als de POT waarde.

## 2.3. Keuze van POT Waarden en Observatieperiode

Voorgaande procedure leidt voor ieder overschrijdingsniveau x tot een lijst van “stormen” met bijhorende piekwaarden. Om te vermijden dat de piekwaarden vertekend zijn door onvolledig gemeten stormen, worden enkel stormen waarbij de aanvang en einde gekend is weerhouden. De observatieperiode van de tijdreeks dient dan echter te worden aangepast als volgt: noemen we T de effectieve tijd waarover HS is gemeten<sup>1</sup>, NS het totaal aantal gegroepeerde stormen over deze periode en NV het totaal aantal gegroepeerde “volledige” stormen. De NV piekwaarden die worden gebruikt in deze analyse zijn dan geassocieerd met een equivalente observatieperiode:

$$TE = \frac{NV}{NS} T \quad (1)$$

De uiteindelijke keuze van het overschrijdingsniveau x waarvoor de analyse wordt uitgevoerd is van minder belang, omdat in de verdere analyse een tweede drempelwaarde wordt ingevoerd waarboven de extreme-waarden theorie van toepassing is. In deze studie werd x per meetstation gekozen, zodanig dat enerzijds een zo hoog mogelijk aantal pieken worden weerhouden terwijl anderzijds de dalende trend van het aantal stormen met x zich duidelijk inzet. Dit laatste is nodig is om te vermijden dat twee onafhankelijke piekwaarden worden samengevoegd.

---

<sup>1</sup> Deze tijd wordt bepaald als het aantal meetpunten vermenigvuldigd met het gebruikte meetinterval en komt dus niet noodzakelijkerwijze overeen met de kalendertijd.

### 3. POT ANALYSE

De POT analyse die in deze studie wordt toegepast is een verdelingsvrije methode zoals wordt uiteengezet in de volgende paragraaf. Hoe de methode precies wordt toegepast wordt verder verduidelijkt in een tweede paragraaf. Resultaten van de toepassing van de methode worden getoond in het volgend hoofdstuk.

#### 3.1. Verdelingsvrije Schattingsmethode van Ontwerpwaarden

Voor de verdelingsvrije schatting van de ontwerpwaarde wordt gebruik gemaakt van de extreme-waarden theorie. Deze theorie veronderstelt dat voor geschikt gekozen constanten  $a_n$  en  $b_n$  het genormaliseerde maximum  $M_n$  van  $n$  onafhankelijke toevalsvARIABLEN  $X_1, X_2, \dots, X_n$  met identieke verdelingsfunctie  $F$  convergeert voor  $n \rightarrow \infty$  naar een extreme-waarden verdeling  $G_\gamma(x)$ . Indien deze veronderstelling geldt dan kan men aantonen dat de extreme-waardenverdeling van de volgende vorm dient te zijn:

$$P\left(\frac{M_n - b_n}{a_n} < x\right) \approx G_\gamma(x) = \exp^{-(1+\gamma x)^{-1/\gamma}} \quad \text{voor } 1 + \gamma x > 0 \quad (2)$$

$\gamma$  noemt men de extreme-waarden index. Voor positieve waarden van  $\gamma$ , is de maximum waarde onbegrensd; voor negatieve waarden is de maximum waarde begrensd.  $\gamma=0$  komt overeen met het speciale geval van de Gumbel-klasse van extreme-waardenverdelingen. Vermelden we nog dat voorgaande convergentie veronderstelling van toepassing is op een zeer ruime klasse van functies  $F_X$ .

Noemen we nu  $x_p$  de waarde  $x$  die met kans  $p$  wordt overschreden door de originele toevalsvARIABLE:

$$F_X(x_p) = P(X \leq x_p) = p \quad (3)$$

Wanneer de convergentie naar de extreme-waarden verdeling geldt, dan kan men aantonen (zie bijvoorbeeld Beirlant et al., 1996) dat voor hoge kwantielwaarden (d.w.z. lage waarden van  $p$ ) de waarde van  $x_p$  kan benaderd worden als volgt:

$$x_p \approx b\left(\frac{n}{k}\right) + \frac{\left(\frac{k}{np}\right)^\gamma - 1}{\gamma} a\left(\frac{n}{k}\right) \quad (4)$$

waarbij  $k$  een functie is van  $n$  waarvoor geldt dat  $k/n \rightarrow 0$  wanneer  $n \rightarrow \infty$ , terwijl  $a$  en  $b$  verwijzen naar de constanten  $a$  en  $b$  in Vergelijking 4, ditmaal voor de gereduceerde steekproefomvang  $n/k$ .

Vergelijking (4) beschrijft het verloop van de rechterstaart van de verdeling  $F_x$ . In essentie kan men vergelijking (4) interpreteren als volgt:

- voor  $\gamma > 0$ , daalt de overschrijdingskans proportioneel met de macht  $1/\gamma$  van  $1/(1+x/\alpha)$  waarbij  $\alpha$  een normaliseringsconstante voorstelt;
- voor  $\gamma = 0$ , daalt de overschrijdingskans exponentieel met  $x/\alpha$ ;
- voor  $\gamma < 0$ , daalt de overschrijdingskans met de macht  $1/\gamma$  van de afstand  $(1-x/\alpha)$  waarbij  $\alpha$  een bovengrens voorstelt voor  $x$  en  $x$  steeds kleiner is dan  $\alpha$ .

Voor de praktische toepassing van Vergelijking (4) is het nodig om de normalisatiecoëfficiënten  $a$  en  $b$  en de extreme-waarden index  $\gamma$  te schatten op basis van een beperkt aantal steekproefgegevens  $x$ . Voor  $b$  gebruiken we de benadering  $x_{n-k}^*$  voorgesteld door Dillingh et al. (1993) in een rapport van de Rijkswaterstaat. De normalisatiecoëfficiënt  $a$  wordt geschat door een rechte lijn te fitten (met behulp van

de kleinste-kwadraten techniek) door de punten  $\left( \frac{x_{n-i}^* - x_{n-k}^*}{x_{n-k}^*}, \frac{\left(\frac{k}{i}\right)^{\hat{\gamma}} - 1}{\hat{\gamma}} \right)$  voor  $i=1$  tot  $k$ .

Dan geldt:

$$\hat{x}_p \approx x_{n-k}^* + \frac{\left(\frac{k}{np}\right)^{\hat{\gamma}} - 1}{\hat{\gamma}} x_{n-k}^* \times \hat{a} \quad (5)$$

waarbij:

- $x_i^*$  de  $i$ 'de orde-statistiek voorstelt van de gegevens (de  $i$ 'de hoogste waarde);
- $\hat{\gamma}$  een schatter is van de extreme-waarden index waarvoor zo dadelijk verscheidene schattingsformules worden voorgesteld;
- de schatter van  $a$  overeenkomt met:

$$\hat{a} = \frac{\sum_{i=1}^k \left( \frac{\left(\frac{k}{i}\right)^{\hat{\gamma}} - 1}{\hat{\gamma}} \right) \left( \frac{x_{n-i}^* - x_{n-k}^*}{x_{n-k}^*} \right)}{\sum_{i=1}^k \left( \frac{x_{n-i}^* - x_{n-k}^*}{x_{n-k}^*} \right)^2} \quad (6)$$

Verschillende schatters kunnen worden voorgesteld voor de extreme-waarden index  $\gamma$ .

Een eerste klassieke schatter komt overeen met  $H_{k,n}$  in Vergelijking (6) en wordt de Hill schatter genoemd. Deze schatter is eenvoudig te interpreteren en komt, na logtransformatie van de ordestatistieken, overeen met het gemiddelde van het verschil van de ordestatistieken hoger dan orde-statistiek  $k$  en diezelfde orde-statistiek. De Hill-schatter:

$$\hat{\gamma}_H = H_{k,n} \quad (7)$$

is echter steeds positief en dus duidelijk vertekend wanneer  $\gamma < 0$ .

Een alternatieve en betere schatter gebruikt door Rijkswaterstaat (Dillingh et al., 1993) is de momentenschatter ontwikkeld door Dekkers et al. (1989). Deze schatter gebruikt naast het gemiddeld verschil ook de dispersie van de ordestatistieken boven de orde-statistiek  $x_{n-k}^*$  en is van de vorm:

$$\hat{\gamma}_M = \hat{\gamma}_H + 1 - \frac{1}{2} \left( \frac{M_{k,n}^{(2)}}{M_{k,n}^{(2)} - H_{k,n}^2} \right) \quad (8)$$

waarbij:

$$M_{k,n}^{(2)} = \frac{1}{k} \sum_{i=0}^{k-1} \left( \ln x_{n-i}^* - \ln x_{n-k}^* \right)^2 \quad (9)$$

Beirlant et al. (1996) stellen nog een andere schatter voor, die in de meeste gevallen een kleinere variantie vertoont dan de momentenschatter. Deze schatter, hierna UH-schatter genoemd, is van de vorm:

$$\hat{\gamma}_{UH} = \frac{1}{k} \sum_{i=0}^{k-1} \left( \ln(x_{n-i}^* H_{i,n}) - \ln(x_{n-k}^* H_{k,n}) \right) \quad (10)$$

Het is deze laatste schatter die in deze studie zal worden toegepast.

Blijft nog de keuze van het aantal ordestatistieken  $k$  voor de schatting van  $\gamma$  en de toepassing van Vergelijking (5). Dit is een bijzonder moeilijk probleem waarvoor verschillende oplossingen zijn voorgesteld. Een klassieke methode bestaat erin om de variatie van de schattingen of een andere statistiek te tekenen in functie van de gekozen waarde van  $k$ . In dit geval zoekt men de waarde van  $k$  waarbij de schatting stabiel wordt. Andere methodes maken gebruik van asymptotische vergelijkingen voor de variantie en vertekening van de schatter (zie bijvoorbeeld Beirlant et al., 1996). Tenslotte kan men bootstrapping trachten toe te passen om een optimale waarde te bepalen (zie bijvoorbeeld, Caers en Van Dyck, 1999).

In deze studie kiezen we voor een eerder pragmatische benadering: door het lukraak herbemonsteren van de gegevens (met teruglegging) is het mogelijk om de variatie van de geschatte waarden te bepalen. Men verwijst naar deze techniek als een niet-parametrische bootstrap. Hoewel deze techniek niet toelaat om op betrouwbare wijze de eventuele vertekening van de schatter te bepalen (tenzij men speciale technieken toepast zoals in Caers en Van Dyck, 1999) is het mogelijk om op deze wijze een statistiek te genereren die een aanduiding geven van zowel de toevallige variatie van de fit (die daalt naarmate meer gegevens worden gebruikt en dus daalt met een hogere waarde van  $k$ ) als de afwijking van de fit ten opzichte van de geobserveerde gegevens (die stijgt naarmate men tracht meer gegevens te modelleren en dus stijgt met hogere waarde van  $k$ ). Hoe dit juist gebeurt wordt verder uiteengezet in volgende paragraaf. Het is op basis van de variatie van deze statistiek met de waarde van  $k$  samen met een visuele beoordeling van de "goodness-of-fit" dat een optimale keuze wordt gemaakt.

### 3.2. Toepassing van de Methode

De ontwerpwaarde wordt gevonden door toepassing van Vergelijking 6 waarbij de kans  $p$  dient te voldoen aan de voorwaarde:

$$p \frac{\text{aantal overschrijdingen}}{\text{meetperiode}} = \frac{1}{\text{terugkeerperiode}} \quad (11)$$

of mathematisch

$$p \frac{n}{T} = \frac{1}{T_R} \quad (12)$$



waarbij  $n$  overeenkomt met het aantal POT waarden weerhouden in de analyse (de variabele  $NV$  in vergelijking 1) en  $T$  overeenkomt met de equivalente observatieperiode (de variabele  $TE$  in vergelijking 1).

Deze voorwaarde volgt uit het feit dat de inverse van de terugkeerperiode de jaarlijkse frequentie aanduidt waarmee de ontwerpwaarde kan voorkomen terwijl  $n/T$  de jaarlijkse frequentie van piekwaarden boven de gekozen drempel aanduidt.  $p$  dient dan overeen te komen met de verhouding van deze twee frequenties. Bijgevolg is de term  $k/(np)$  in Vergelijking (5) gelijk aan:

$$\frac{k}{np} = k \frac{T_R}{T} \quad (13)$$

De schattingsmethode omvat dan volgende stappen:

1. extractie van de POT waarden voor een gegeven drempelwaarde, zoals eerder beschreven in Hoofdstuk 2;
2. een bootstrap simulatie (50 maal) waarbij de gegevens zonder teruglegging worden gekozen uit de POT waarden en voor iedere simulatie:
  - a. de waarde  $\hat{\gamma}_{UH}$  wordt geschat (Vergelijking 10) voor alle waarden van  $k$ ;
  - b. de waarde  $H_{k,n}$  (de Hill-schatter beschreven in Vergelijking 6) wordt geschat voor alle waarden van  $k$ ;
3. berekening van het gemiddelde en de standaarddeviatie van de schatter  $\hat{\gamma}_{UH}$ ;
4. berekening voor elke waarde van  $k$  van een “gemiddelde totale fout” of MSE waarde van de gefitte curve voor de eerste 10 ordestatistieken. Dit gebeurt door:
  - a. voor de eerste 10 ordestatistieken en voor iedere waarde  $k$  een voorspelde waarde  $\hat{x}_p$  (met  $p=k'/n$ ) te berekenen voor elke simulatie en hiervan het gemiddelde en de standaarddeviatie te berekenen (hiertoe dient ook de coëfficiënt  $a$  in Vergelijking 6 geschat te worden);
  - b. de totale variantie van deze voorspellingen (inclusief de afwijking van het simulatiegemiddelde ten overstaan van de schatting) samen te voegen met de som van de gekwadrateerde afwijkingen van de schatting ten overstaan van de geobserveerde waarden:

$$MSE = \sum_{k'=1}^{10} \text{var}(\hat{x}_{p=k'/n}^B) + \left(\bar{\hat{x}}_{p=k'/n}^B - \hat{x}_{p=k'/n}\right)^2 + \left(\hat{x}_{p=k'/n} - x_{k'}\right)^2 \quad (14)$$

waarbij  $\text{var}(\hat{x}_{p=k'/n}^B)$  verwijst naar de variantie van de voorspelde waarden voor de verschillende bootstraps,  $\overline{\hat{x}_{p=k'/n}^B}$  verwijst naar de gemiddelde waarde van de voorspelde waarde voor de verschillende bootstrap simulaties, en  $\hat{x}_{p=k'/n}$  verwijst naar de beste schatting van de ordestatistiek  $x_k$ ;

5. bepaling van de waarde van  $k$  waarvoor MSE minimaal is en berekening van de ontwerpwaarden voor verschillende waarden van  $p$  gebruikmakend van de schatter  $\hat{\gamma}_{UH}$  voor deze drempelwaarde en voor de originele POT waarden. Deze berekening wordt herhaald voor iedere simulatie om de standaardfout op de voorspelde ontwerpwaarden te berekenen.

Resultaten van de toepassing van deze methode worden besproken in het volgend hoofdstuk. De automatische keuze van de waarde  $k$  op basis van Vergelijking (14) leidt echter niet altijd tot een redelijk resultaat en daarom is, in enkele gevallen,  $k$  handmatig gekozen.

## 4. RESULTATEN

In dit hoofdstuk worden resultaten van de POT analyse uitgelegd en de keuze (hetzij automatisch, hetzij handmatig) van de drempelwaarde  $k$  besproken. Het overschrijdingsniveau voor de bepaling van stormen is in alle gevallen gelijkgesteld aan 2.3 meter. Vanaf deze waarde vindt men voor alle meetstations een duidelijke afname van het aantal stormen met hogere drempelwaarden, zodat men er van kan uitgaan dat het typisch stormen betreft met enkelvoudige maxima. De gemiddelde stormduur voor deze waarden varieert van meetstation tot meetstation, maar is typisch gelijk aan ongeveer 10 uur.

In de volgende beschrijving worden enkel de resultaten voor de A2-boei (H33 metingen) uitvoerig gecommentarieerd. Voor de overige meetstations wordt enkel de handmatige correctie van de keuze van de drempel  $k$ , in zoverre die nodig bleek, beschreven.

## 4.1. A2-boei, H33 metingen

Figuur 1 toont een samenvatting van de resultaten van de POT analyse voor dit meetstation.

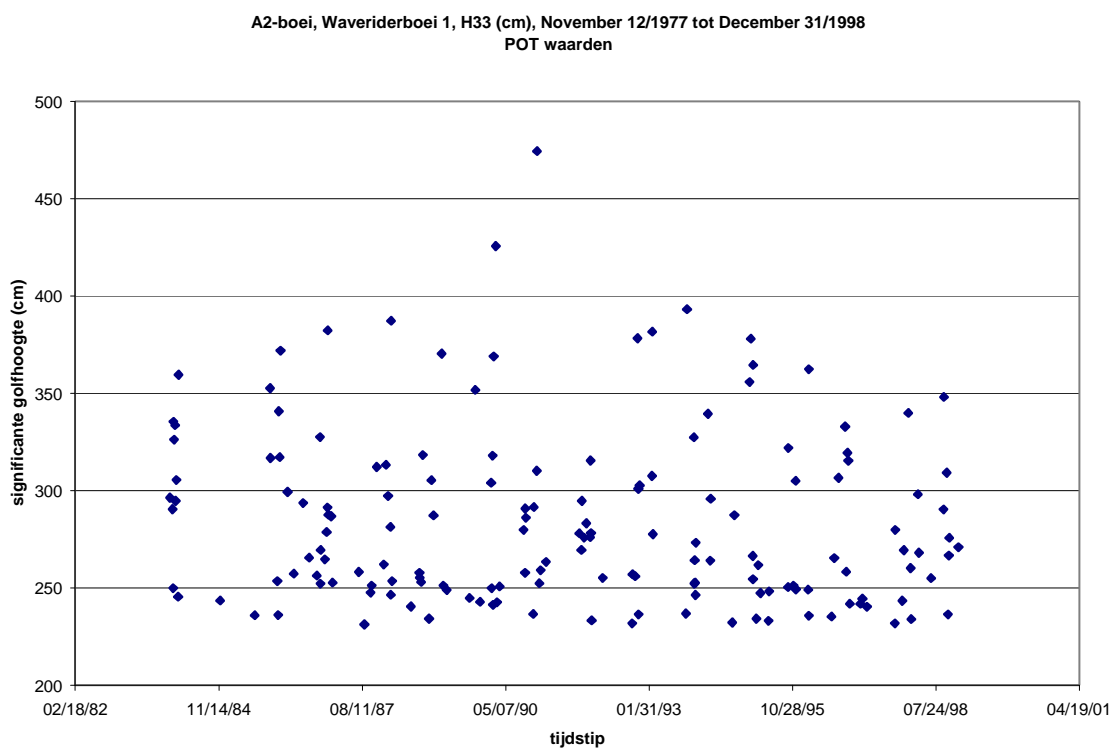
### RESULTATEN VAN DE POT ANALYSE

Beschrijving		A2-boei, Waveriderboei 1, H33 (cm), November 12/1977 tot December 31/1998			
Drempelwaarde (cm)		230			
Aantal POT waarden		158			
Observatieperiode (jaren)		10.45			
Analyse methode		UH			
Aantal stormen		185			
Equivalentente observatieper. (jaren)		8.93			
Aantal orde-statistieken voor fit		59			
Orde-statistiek (cm)		290.5			
Extreme-waarde index		-0.089			
Standaardfout extreme-waarde index		0.163			
Hill extreme-waarde index		0.129			
a-coëfficiënt voor terugkeerperiode		0.161			
Terugkeerperiode (jaar)	waarde (cm)	standaardfout	95% 2-zijdig betrouwbaarheidsinterval		
0.2	303	6	291	315	
0.5	344	9	325	362	
1	372	12	349	395	
2	399	15	370	427	
5	431	21	391	472	
10	455	26	403	506	
20	476	33	412	541	
50	503	43	420	586	
100	522	51	423	621	
200	539	59	424	655	
500	561	71	423	700	
1000	577	80	419	734	
2000	591	91	414	769	
5000	609	105	402	815	
10000	621	117	391	851	

**Figuur 1 - Resultaten van de POT analyse voor de A2 boei**

Voor de A2 boei worden de metingen gebruikt vanaf november 12, 1977 (het tijdstip van de eerste “storm”) tot december 31, 1998 (het tijdstip van de laatste gegevens beschouwd in deze analyse). Over deze kalenderperiode van 21.15 jaar werden 366,531 metingen uitgevoerd representatief voor een tijdsinterval van 15 minuten. De overeenkomstige effectieve observatieperiode is 10.45 jaar. Over deze periode worden 185 stormen gedetecteerd waarbij de drempelwaarde van 2.3 meter wordt overschreden. 158 van deze stormen worden weerhouden als “volledig” gemeten en het is voor deze stormen dat de piekwaarden worden gebruikt in de analyse. De overeenkomstige equivalentente observatieperiode bedraagt 8.93 jaren.

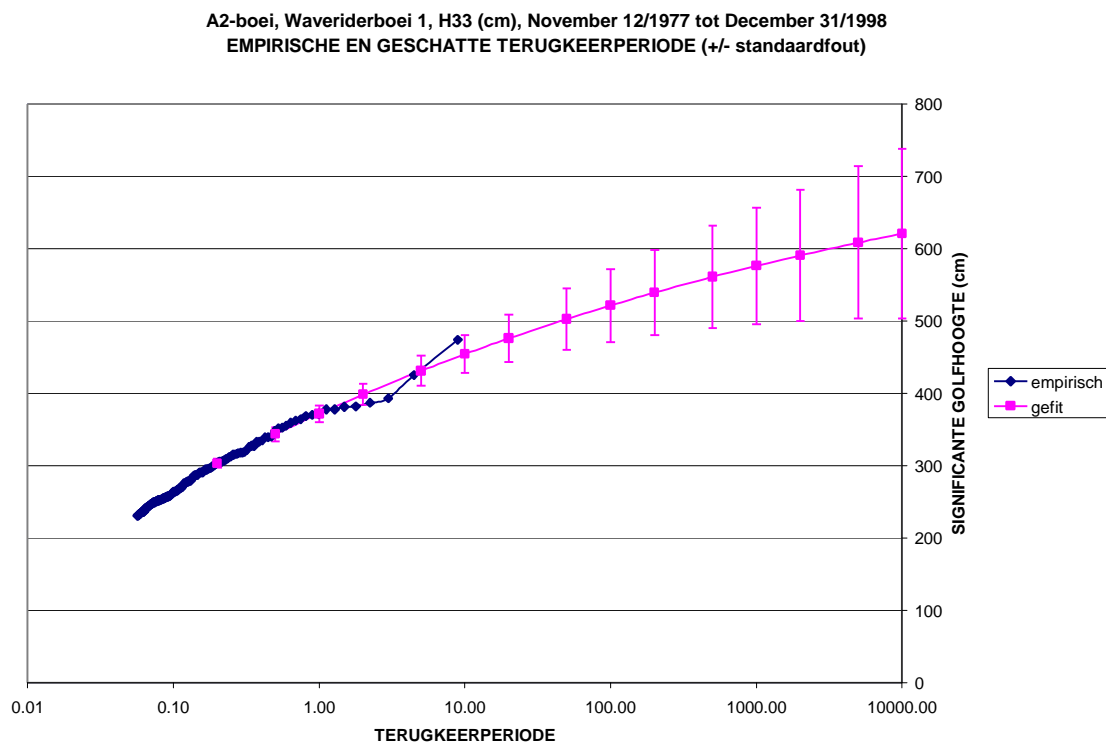
Figuur 2 toont de POT waarden in functie van de tijd.



**Figuur 2 - Plot van de POT waarden in functie van het tijdstip van de waarneming**

De hoogste waarde (4.75 meter) werd gemeten op 12 december 1990.

Figuur 3 toont voor de ordestatistieken van deze POT waarden de overeenkomstige terugkeerperiode, berekend als  $(i)/TE$  waarbij  $i$  overeenkomt met de orde (bv.  $i=1$  verwijst naar de maximum waarde) en  $TE$  overeenkomt met de equivalente observatieperiode. In deze figuur wordt eveneens de fit van deze waarde getoond wanneer men Vergelijking (5) toepast vanaf de 59<sup>ste</sup> ordestatistiek, die overeenkomt met een waarde van 290.5 cm. Voor deze gefitte waarden wordt eveneens de standaardfout die volgt uit de bootstrapping getoond (de verticale lijn komt overeen met +/- 1 standaardfout). Zoals aangeduid in de resultaten van Figuur 1 is de geschatte extreme-waarden index gelijk aan  $-0.089$  wanneer men gebruik maakt van de UH methode en de standaardfout komt overeen met 0.163. De schatter die men bekomt door middel van de Hill-methode komt overeen met 0.129. De schatting van de coëfficiënt  $a$  gebruikt in Vergelijking (5) komt overeen met 0.161.



**Figuur 3 - Empirische en gefitte terugkeerperiode**

De fit in Figuur 3 benadert de empirische curve op redelijke wijze. Numerische waarden voor de ontwerpwaarden van HS die overeenkomen met een bepaalde terugkeerperiode worden getoond in Figuur 1 samen met een 95% tweezijdig betrouwbaarheidsinterval.

Vergelijking van deze waarden met de schatting van de HS waarde<sup>2</sup> die met een frequentie  $1/(365 \cdot 24 \cdot 4 \cdot T_R)$  voorkomen wanneer men alle 15-minuten metingen in rekening brengt (in plaats van enkel piekwaarden) leidt tot het resultaat in Tabel 1, waar de benedengrens van het 95% tweezijdig betrouwbaarheidsinterval, de beste schatting en de bovengrens van het 95% tweezijdig betrouwbaarheidsinterval worden getoond voor verschillende terugkeerperiodes.

terugkeerperiode	POT analyse	marginale kansverdeling
0.5	(325, 344, 362)	(368, 392, 419)
1	(349, 372, 395)	(385, 412, 441)
10	(403, 455, 506)	(442, 477, 513)
100	(423, 522, 621)	(496, 538, 582)
1000	(419, 577, 734)	(547, 596, 649)
10,000	(391, 621, 851)	(597, 652, 713)

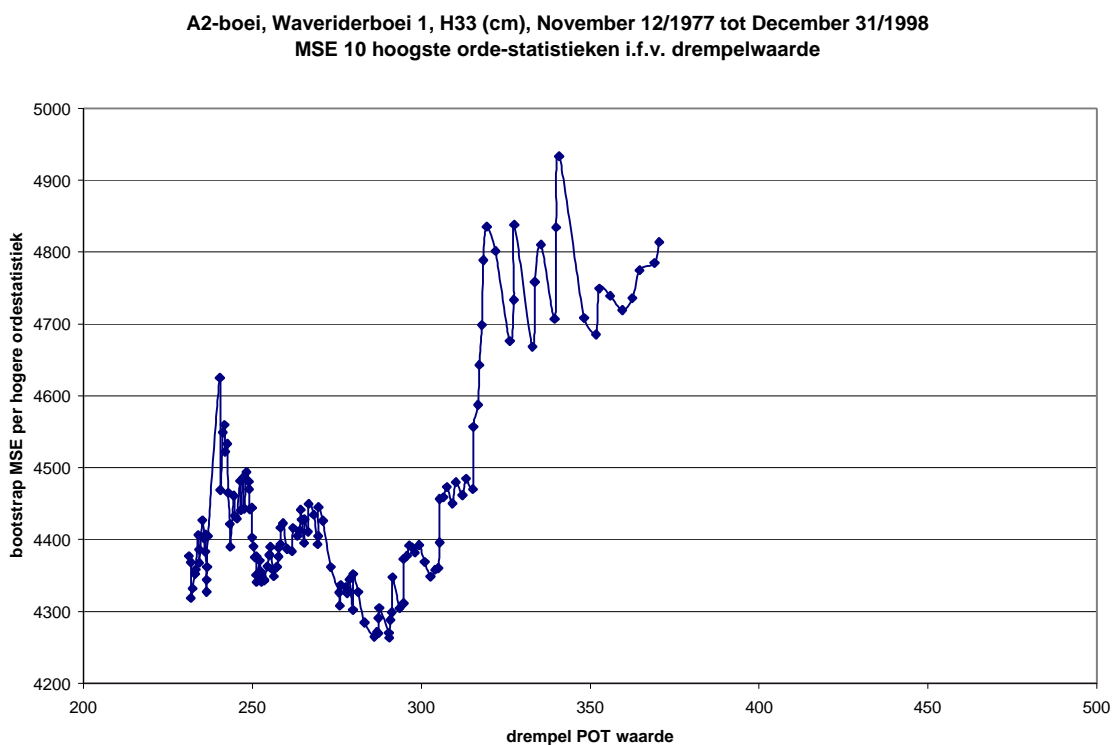
**Tabel 1 - Beste schatting (in cm) en 95% tweezijdig betrouwbaarheidsinterval voor de ontwerpwaarden bepaald op basis van de POT analyse en op basis van de marginale kansverdeling**

De waarden in deze tabel tonen duidelijk dat voor lage terugkeerperiodes  $T_R$  de POT analyse tot kleinere waarden leidt. Dit is zo omdat voor deze waarden, een overschrijding van de POT waarde typisch overeenkomt met meervoudige overschrijdingen binnen een "storm" van deze waarde. Merk ook op dat vanaf een terugkeerperiode van 100 jaar, de resultaten van beide analyses vergelijkbaar zijn en dat, althans voor dit meetstation, de breedte van het betrouwbaarheidsinterval in het geval de POT analyse merkkelijk groter is dan de breedte bepaald op basis van de marginale kansverdeling.

---

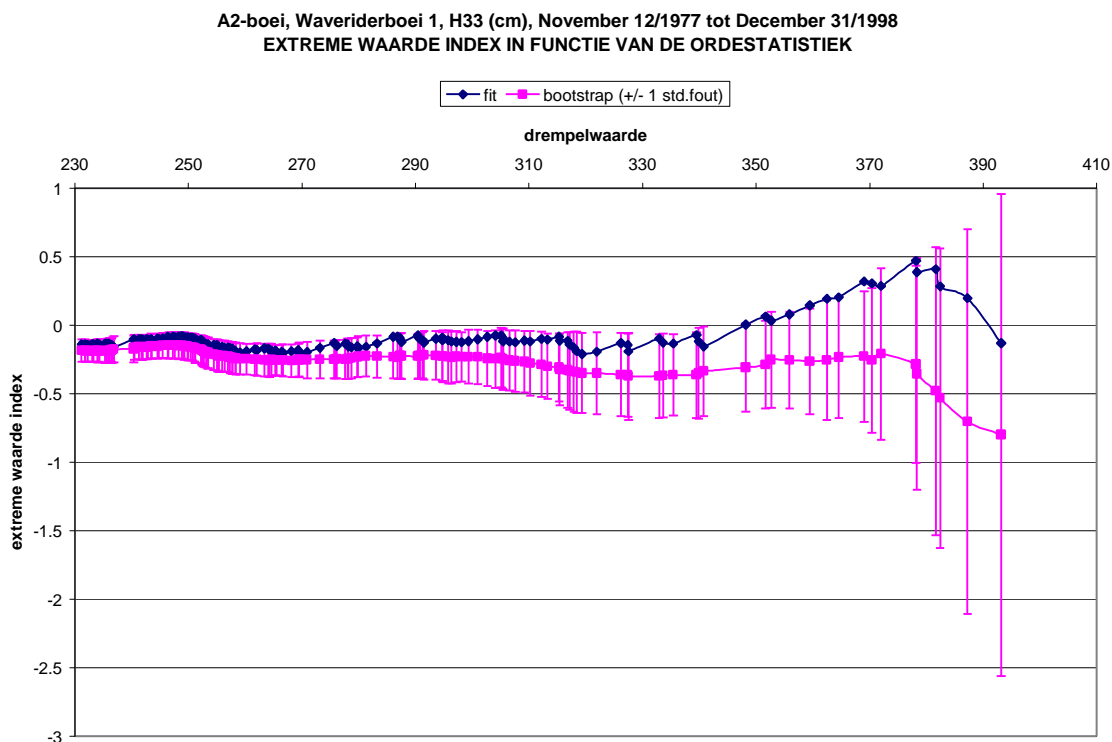
<sup>2</sup> zulke schattingen vindt men onder de sectie "Weibull: N-jaarlijkse ontwerpwaarden" van de Meteoatlas en zijn eerder besproken in het theorie gedeelte "Schatting van de Kansverdeling van Significante Golfhoogte".

De k-waarde voor de POT analyse hiervoor getoond werd automatisch gekozen. Figuur 4 toont de evolutie van de MSE waarde in functie van de drempelwaarde. De minimum waarde voor MSE vindt men bij de 59'ste ordestatistiek (drempelwaarde 290.5). Een tweede lokaal optmimum lijkt zich voor te doen bij een iets lagere drempelwaarde kiezen (rond 250 cm) en alternatief zou men deze drempelwaarde kunnen gebruiken.



**Figuur 4 - Plot van de MSE waarde (zie Vergelijking 14) in functie van de waarde van de k'de ordestatistiek**

Figuur 5 toont hoe de schatting van de extreme-waarden index fluctueert in functie van de gekozen drempelwaarde.



**Figuur 5 - Geschatte extreme-waarden index in functie van de drempel ordestatistiek. Voor de bootstrap simulaties wordt het gemiddelde +/- 1 standaarddeviatie getoond.**

De schatting is relatief stabiel in functie van de gekozen drempelwaarde en men zou ook de lagere drempelwaarde rond de 250 cm kunnen verkiezen. Zulk een keuze leidt tot een vergelijkbaar resultaat met echter kleinere betrouwbaarheidsintervallen..



## 4.2. Overige meetstations

Dezelfde procedure is toegepast voor de overige meetstations en we verwijzen naar de Meteoatlas voor detailresultaten.

Zoals in voorgaand voorbeeld leidt de automatische keuze van de drempelwaarde tot redelijke fits, maar ook alternatieve waarden (overeenkomstig met lokale minima in de MSE plot) zouden tot een redelijk resultaat leiden en de betrouwbaarheid kunnen verhogen. Voor de presentatie in de Meteoatlas werd gekozen voor de automatische selectie omdat een alternatief selectiecriteria (dat ook objectief kan gefundeerd worden) voorlopig ontbreekt.

Uitzondering dient gemaakt voor de studie van de POT waarden ter plaatse van Westhinder (H33 metingen). De automatische selectie weerhoudt in dit geval enkel de 43 hoogste orde-statistieken en de extreme waarde index wordt geschat als 0.293 hetgeen bijzonder hoge waarden oplevert voor de ontwerpwaarden. Dit is echter volledig te wijten aan de extreme meetwaarde van 815 cm in november 93. Deze waarde kan enkel goed gefit worden door een dergelijk hoge extreme-waarde index te gebruiken die dan enkel van toepassing is voor de allerhoogste waarden. Om de invloed van deze waarneming te beperken werd daarom in de meteoatlas gekozen voor een handmatige selectie van de drempelwaarde (161 orde-statistieken en een drempelwaarde van 258.5 cm). Dit levert een meer redelijk resultaat, hoewel de resultaten nog steeds duidelijk tonen dat de waarneming van 815 cm duidelijk ongewoon is. Deze waarneming veroorzaakt verder een zeer breed betrouwbaarheidsinterval.

Voor een vergelijkende studie en discussie van de ontwerpwaarden die worden bekomen door deze POT analyse en de waarden die volgen door studie van de marginale kansverdeling, mits eventuele aanpassing voor de duur van overschrijding, verwijzen we naar een afzonderlijk rapport in deze Meteoatlas.

## REFERENTIES

Beirlant, J., J.L. Teugels, P. Vynckier, 1996, Practical Analysis of Extreme Values, Leuven University Press

Caers, J. en J. Van Dyck, 1999, "Nonparametric Tail Estimation Using a Double Bootstrap Method", *Computational Statistics & Data Analysis* 29 (1999), pp. 191-211

Dillingh, D., L. de Haan, R. Helmers, G.P. Können, J. van Malde, 1993, "De Basispeilen langs de Nederlandse Kust: Statistisch Onderzoek", Rapport DGW-93.023 Deel 1 – Tekst, Dienst Getijdewateren, Rijksinstituut voor Kust en Zee, Directoraat-Generaal Rijkswaterstaat

de Ronde, J.G., J.G.A. van Marle, A.P. Roskam, J.H. Andorka Gal, 1995, "Golfrandvoorwaarden langs de Nederlandse Kust op Relatief Diep Water", Rapport RIKZ-95.024, Rijksinstituut voor Kust en Zee, Directoraat-Generaal Rijkswaterstaat

Hydro Meteo Atlas, Meetnet Vlaamse Banken, Prototype versie juli 1999, CD-ROM, Afdeling Waterwegen Kust, administratie Waterwegen en Zeewegen, Departement Leefmilieu en Infrastructuur, Ministerie van de Vlaamse Gemeenschap

Pickands, J., 1975, "Statistical Inference Using Extreme Order Statistics", *The Annals of Statistics*, 3, pp. 119-131