

# het sneeuwdek in Nederland

T. A. Buishand

wetenschappelijke rapporten WR 86-6

met medewerking van

A. Denkema en

P.J.M. van der Veer

T.A. Buishand (1986): "Snow cover in The Netherlands" - Royal Netherlands Meteorological Institute (K.N.M.I.), Sci. Rep. W.R. 86- 6

### Abstract

This report deals with the following climatological properties of the snow cover in The Netherlands:

1. The number of days with snow cover.
2. The distribution of the annual maximum snow depths.
3. Run lengths and the distribution of snow depths.

Snow depth data from 13 synoptic stations were studied for the 28 winters 1955/56, ... , 1982/83. The main body of these data is given in the form of the synoptic code S. This code can take the following values:

S = 0	no snow cover
1	depth < 2 cm
2	depth 2-4 cm
3	depth 5-9 cm
4	depth 10-14 cm
5	depth 15-24 cm
6	depth 25-49 cm.

Snow depths over 49 cm have not been observed at the stations considered in this study.

Three long-term series of snow depth observations in the neighbourhood of The Netherlands were taken into account to find a suitable distribution for the annual maxima, namely Diepholz (58 yr.) and Kleve (47 yr.) in West Germany, and Uccle-Brussels (95 yr.) in Belgium. Further, some attention was paid to data from Borkum and Köln (Cologne) in West Germany over the same period as the data for The Netherlands. The station locations are given in Fig. 1.1.

### The number of days with snow cover

In this study a day is said to have a snow cover if  $S > 0$  for the morning observation (0800 GMT or, since 1981, 0600 GMT). The mean annual number of days with snow cover ranges from 12 in the south-west to 31 in the north-east (Fig. 2.1). In about 30% of these days less than half of the ground is snow covered. For the stations considered in this study there is a strong relation

between the mean annual number of days with snow cover and the 24-hour mean temperature  $\bar{w}$  for the 3-month period 1 December - 28/29 February (Fig. 2.2). The mean annual number of days with snow cover decreases with increasing mean winter temperature  $\bar{w}$  by about 10 per 1°C. This relation holds for lowland regions; in hilly regions (Ardennes) there is an extra increase in the annual mean number of days with snow cover due to orographic enhancement of precipitation.

The distribution of the number of days with snow cover in a winter can be described by the 3-parameter lognormal distribution with probability density:

$$f(x) = \frac{1}{(x-a) \sigma \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left\{ \frac{\ln(x-a) - \mu}{\sigma} \right\}^2 \right]. \quad (1)$$

A suitable choice of the parameter  $a$  is -5 for all stations; the values of the parameters  $\mu$  and  $\sigma$  depend on the mean winter temperature  $\bar{w}$  (Fig. 2.7). The validity of the lognormal distribution is doubtful for  $x < 3$  days. Therefore values of  $x$  less than 3 days were treated as left-censored for estimating the parameters  $\mu$  and  $\sigma$ . The information from the sample then consists of the number of winters with  $x < 3$  days and the actual values of  $x$  for winters with  $x \geq 3$  days.

The fitted lognormal distribution was used to derive quantiles of the distribution of the number of days with snow cover in a winter. These quantiles depend on the mean winter temperature  $\bar{w}$  (Fig. 2.8).

#### The distribution of the annual maximum snow depths

The following theoretical distributions were fitted to the annual maximum snow depths of the long-term records from Diepholz, Kleve and Uccle-Brussels:

1. The 2-parameter lognormal distribution. For this distribution the parameter  $a$  in eq. (1) equals 0.
2. The Gumbel distribution. This is a 2-parameter distribution with distribution function:

$$F(x) = \Pr(X \leq x) = \exp \left\{ -e^{-(x-\mu)/\sigma} \right\}. \quad (2)$$

For both distributions estimates of  $\mu$  and  $\sigma$  were obtained from a censored annual maximum series, in which snow depths less than 5 cm were considered as left-censored values.

For the 13 synoptic stations in The Netherlands the Gumbel distribution was

fitted to the frequency distribution of the annual maxima of the synoptic code S. Also in this case the estimates of the parameters  $\mu$  and  $\sigma$  depend on the mean winter temperature  $\bar{w}$  (Fig. 3.5). These relations were used to derive maps of the 10-year and 50-year snow depths (Figs. 3.6 and 3.7). The standard errors of these estimates are large (3 and 5 cm, respectively), because they are based on a relatively short period of 28 years. For Diepholz, Kleve and Uccle-Brussels the estimates obtained from the data for this 28-year period correspond nicely with the estimates from the entire record. Incomplete knowledge about the right tail of the distribution leads to an extra uncertainty in the 50-year snow depths, especially in the northern coastal region.

#### Run lengths and the distribution of snow depths

The distribution of run lengths of days with snow cover was studied for the five synoptic stations Den Helder/De Kooy, Eelde, De Bilt, Vlissingen and Beek. The mean run length is the same for these stations (slightly over 4 days) and also the distribution of the run lengths does not vary over the country. The mean annual number of runs ranges, however, from about 3 for the coastal site Vlissingen to nearly 7 for the inland site Eelde. The longest run of days with snow cover (about 70 days for all sites) occurred in the exceptional winter 1962/63.

The distribution of snow depths on days with snow cover can be described by a 2-parameter exponential distribution with distribution function:

$$F(x) = \Pr(X \leq x) = 1 - Ce^{-\lambda x}, \quad x > 0. \quad (3)$$

The parameter C is needed to describe a departure from the classical 1-parameter exponential distribution near  $x=0$ . From the fitted exponential distribution it follows that the mean snow depth on days with snow cover ranges from 2 cm at coastal sites to slightly over 4 cm at inland sites. Some attention was also paid to the distribution of run lengths of days with snow depths above a certain threshold (5, 10 or 15 cm). In contrast with runs of days with snow cover the distribution of these run lengths depends strongly on the geographical location.

## INHOUD

1. Inleiding	6
2. Het aantal dagen met een sneeuwdek	11
2.1 Inleiding	11
2.2 Het gemiddelde aantal dagen met een sneeuwdek	12
2.3 De homogeniteit van het materiaal	17
2.4 Nadere bijzonderheden over het gemiddelde aantal dagen met een sneeuwdek in aangrenzende landen	19
2.5 De kansverdeling van het aantal sneeuwdekdagen	22
3. De kansverdeling van de jaarmaxima van de hoogte van het sneeuwdek	33
3.1 Inleiding	33
3.2 Toepassing van de Gumbel-verdeling op jaarmaxima van langjarige reeksen van Diepholz, Kleef en Ukkel-Brussel	35
3.3 Toepassing van de Gumbel-verdeling op frequentieverdelingen van synoptische codecijfers	40
3.4 Opmerkingen over de nauwkeurigheid van parameterschattingen	47
4. Runlengten en de kansverdeling van de hoogte van het sneeuwdek	53
4.1 Inleiding	53
4.2 Runlengten van sneeuwdekdagen	53
4.3 De kansverdeling van de hoogte van het sneeuwdek	55
4.4 Runlengten van dagen met een sneeuwdek boven een bepaald niveau	60
5. Samenvatting	64
Dankwoord	67

Appendices

A. Homogeniteitsanalyse met betrekking tot het aantal sneeuwdekdagen	68
B. Het schatten van de parameters van een kansverdeling bij geschorste steekproeven	71
B1. Toepassing van de methode van de grootste aannemelijkheid bij de lognormale verdeling	72
B2. Toepassing van de methode van de grootste aannemelijkheid bij de Gumbel-verdeling	75
C. Aanpassingstoetsen	77
D. Het schatten van de parameters van de lognormale verdeling bij het aantal sneeuwdekdagen	83
E. Het schatten van de parameters van de Gumbel-verdeling bij waarnemingen in klassen	87
F. Het schatten van de parameters van de exponentiële verdeling bij sneeuwdekhogten	90
Literatuur	92

## 1. INLEIDING

Regelmatig worden vanuit de praktijk vragen gesteld over het aantal dagen dat de grond met sneeuw bedekt is en over de hoogte van het sneeuwdek. Om dit soort vragen te beantwoorden zijn waarnemingsreeksen nodig die bovendien nog vrij lang moeten zijn daar het voorkomen van een sneeuwdek een verschijnsel is dat van jaar tot jaar sterk varieert. In België en Duitsland vinden sneeuwdekhogtemetingen plaats vanaf de tweede helft van de vorige eeuw. In vergelijking daarmee zijn regelmatige metingen van het sneeuwdek in Nederland van vrij recente datum. Vanaf december 1939 wordt in Wageningen de hoogte van het sneeuwdek gemeten door het Laboratorium van Natuur- en Weerkunde van de Landbouwhogeschool. In De Bilt vingen de sneeuwdekhogtemetingen aan in februari 1942 toen de sneeuw zich opstapelde tot de zeer uitzonderlijke hoogte van 42,5 cm. Sinds december 1945 wordt het sneeuwdek tevens gemeten op de hoofdstations Den Helder, Eelde, Vlissingen en Beek (L.). Van een uitgebreid waarnemingsnet is pas sprake in 1954 wanneer op ruim 300 regenstations sneeuwdekhogtemetingen plaatsvinden.

Een probleem bij de bestudering van eigenschappen van het sneeuwdek in Nederland was dat tot voor kort de waarnemingsgegevens vrij ontoegankelijk en slecht gecontroleerd waren. Hierin is recentelijk een verbetering gekomen dankzij de activiteiten van een tweetal weeramateurs. Door Piekema (1984) werd een reeks van dagelijkse sneeuwdekmetingen van Midden Nederland gepubliceerd voor het tijdvak 1950-1984. Deze reeks is samengesteld uit metingen van De Bilt en Amersfoort. Daarnaast werd door Van der Horst (1984) een overzicht gepubliceerd van sneeuwdekhogtegegevens van negentien synoptische\* stations voor het tijdvak 1956-1977.

De meeste sneeuwdekhogtemetingen in Nederland zijn omgewerkt tot een codecijfer. Bij het netwerk van synoptische stations is van 1955 tot 1981 het zogenaamde S-cijfer in gebruik geweest, zie tabel 1.1.

Het S-cijfer heeft betrekking op de sneeuwdekhogtemeting van 8 uur UT (Universal Time, vroeger Middelbare Tijd van Greenwich GMT) en werd vanaf 1956 opgenomen in het synoptische codebericht van 9 uur UT. De waarnemingstijd van

\* Synoptische stations zijn meteorologische stations die elk uur een beschrijving van de weersgesteldheid ter plaatse geven. Deze informatie wordt o.a. gebruikt bij het vervaardigen van weerkaarten.

8 uur UT was landelijk vatgesteld en week af van de internationaal gebruikelijke waarnemingstijd van 6 uur UT. Slechts De Bilt had naast een S-cijfer van 8 uur UT ook een S-cijfer van 6 uur UT daar dit station deel uitmaakte van een internationaal net van waarnemingsstations.

Vanaf januari 1981 worden op de meeste synoptische stations de hoogten van het sneeuwdek gemeten om 0, 6, 12 en 18 uur UT. De getallen van 6 uur UT zijn vanaf januari 1984 gepubliceerd in het dagoverzicht van het weer in Nederland.

Naast een codecijfer S voor de hoogte van het sneeuwdek hebben synoptische stations ook nog een codecijfer E voor de toestand van de grond. Uit dit laatste codecijfer, dat in 1949 werd ingevoerd, kan men opmaken of de grond droog, nat, bevroren of met sneeuw bedekt was. Er was geen sneeuw indien  $E=0,1,2,3$  of 4 (het S-cijfer is dan gelijk aan nul). Indien  $E=5$  dan was de bodem wel met sneeuw bedekt, echter voor minder dan de helft van de omgeving van het waarnemingspunt. Het E-cijfer bedraagt 6,7,8 of 9 op tijdstippen waarop meer dan de helft van de bodem met sneeuw bedekt was. Vóór 1949 was een ander codecijfer in gebruik om de toestand van de grond te karakteriseren. Voor De Bilt is vanaf oktober 1930 een voorloper van het E-cijfer in de synoptische waarnemingsboeken opgetekend; voor een aantal andere synoptische stations is een dergelijk codecijfer vanaf het einde van de jaren dertig beschikbaar.

Behalve de synoptische codecijfers voor het sneeuwdek is er ook een klimatologisch codecijfer dat in gebruik is bij de regenstations. Het klimatologisch codecijfer is speciaal voor de Nederlandse situatie ontworpen. Voor stations die òn als synoptisch station òn als regenstation fungeren zijn reeksen van beide codecijfers beschikbaar. Daar het bestand van klimatologische codecijfers nog vrij ontoegankelijk is, zal dit hier buiten beschouwing worden gelaten.

De basis voor deze studie over het sneeuwdek in Nederland vormt het materiaal in de publikatie van Van der Horst (1984), echter aangevuld met gegevens over 1955 en 1978-1983\*. Voor het tijdvak 1981-1983 is steeds de ochtendwaarneming

\* De gegevens over 1955 en 1978-1980 zijn ontleend aan lijsten van magneetbandbestanden; de gegevens over 1981-1983 zijn afkomstig uit microfiches van het KNMI. De getallen over deze jaren verschillen soms iets van die in de juist verschenen publikatie van Van der Horst (1986). Bij de berekeningen in hoofdstuk 4 is uitsluitend gebruik gemaakt van computerbestanden van het KNMI.



van 6 uur UT aangehouden. Van de negentien synoptische stations die Van der Horst beschouwde, zijn er dertien die een volledige waarnemingsreeks hebben over de 28 winterhalfjaren 1955/56,.....,1982/1983. Voor de namen en de ligging van de betreffende stations wordt verwezen naar tabel 1.2 en fig. 1.1. De meeste sneeuwdekhoogtewaarnemingen van deze stations zijn gegeven in de vorm van het S-cijfer. Dit codecijfer heeft als nadeel dat het slechts ruwe informatie geeft over de hoogte van het sneeuwdek. Zo zijn er voor hoogten tussen de 15 en 50 cm, die juist voor Nederland zeer interessant zijn, slechts twee klassen gereserveerd. Om een beter inzicht te krijgen in de kansverdeling van de sneeuwdekhoogten werden drie langjarige waarnemingsreeksen met werkelijke sneeuwdekhoogten geanalyseerd, namelijk Diepholz (40 km noordnoordoostelijk van Osnabrück), Kleef en Ukkel-Brussel. Daar deze reeksen aanzienlijk langer zijn dan de synoptische waarnemingsreeksen in Nederland bieden zij tevens de mogelijkheid iets te zeggen over de representativiteit van het tijdvak 1955/56-1982/83 voor een langer tijdvak. Naast deze drie langjarige reeksen werden ook de sneeuwdekhoogtegegevens van Midden Nederland, Borkum en Keulen geanalyseerd.

Hoofdstuk 2 gaat in op het aantal dagen in een jaar dat de bodem met sneeuw bedekt is. De kansverdeling van de maximale hoogte van het sneeuwdek binnen een winterhalfjaar wordt behandeld in hoofdstuk 3. Hoofdstuk 4 geeft nadere informatie over de kansverdeling van de hoogte van het sneeuwdek en over runlengten van sneeuwdekdagen. Onder een run van sneeuwdekdagen wordt een reeks van opeenvolgende dagen verstaan waarop de bodem met sneeuw bedekt is, voorafgegaan en gevolgd door één of meer dagen zonder sneeuwbedekking. Behalve aan runs van sneeuwdekdagen wordt ook enige aandacht geschonken aan reeksen van opeenvolgende dagen waarop het sneeuwdek een bepaald niveau heeft overschreden. Statistische details zijn opgenomen in een zestal appendices.

Tabel 1.1. Het synoptisch codecijfer S voor de hoogte van het sneeuwdek. De codecijfers 7, 8 en 9 zijn nooit waargenomen op de Nederlandse stations.

S = 0: geen sneeuwdek	S = 5: 15-24 cm
1: < 2 cm	6: 25-49 cm
2: 2-4 cm	7: 50-99 cm
3: 5-9 cm	8: 100-199 cm
4: 10-14 cm	9: > 199 cm

Tabel 1.2. Sneeuwdekreeksen die in deze studie onderzocht zijn.

Synoptische stations met codecijfer S		Reeksen met werkelijke sneeuwdekhogten
Den Helder/De Kooy	Soesterberg	Midden Nederland
Leeuwarden	Vlissingen	Borkum
Eelde	Gilze-Rijen	Diepholz
Twente	Volkel	Kleef
Deelen	Eindhoven	Keulen
Valkenburg (Z.H.)	Beek (L.)	Ukkel-Brussel
De Bilt		



Fig. 1.1. Ligging stations waarvan sneeuwdekreeksen onderzocht zijn.

Geographical location of stations.

## 2. HET AANTAL DAGEN MET EEN SNEEUWDEK

### 2.1. Inleiding

De aanwezigheid van een sneeuwdek heeft een sterke verandering van de fysische eigenschappen van het oppervlak tot gevolg en leidt tot allerlei maatschappelijke ongemakken (slechte berijdbaarheid van wegen; afgelastingen veldsporten). Vandaar dat er in de praktijk veel vragen leven over het gemiddelde aantal dagen dat de grond met sneeuw bedekt is en hoeveel dit aantal sneeuwdekdagen van winter tot winter kan variëren.

Er bestaat geen algemeen aanvaarde definitie van een sneeuwdek dag. In deze studie wordt een dag als sneeuwdek dag aangeduid indien het S-cijfer in tabel 1.1 van nul verschilt. Deze definitie houdt in dat ook dagen worden meegeteld waarop slechts een gedeelte van de grond met sneeuw bedekt is. Daar het S-cijfer op een ochtendwaarneming betrekking heeft, zal zo nu en dan de betere benaming sneeuwdekochtend gebruikt worden.

De hier gehanteerde definitie van een sneeuwdek dag vloeit voort uit het feit dat het bestand van S-cijfers het enige uitvoerige bestand is van goed gecontroleerde gegevens in Nederland. Verschillen met andere publikaties ontstaan doordat velen pas van een sneeuwdek dag spreken als een bepaald gedeelte van de bodem (gewoonlijk de helft) met sneeuw bedekt is. Een andere bron van verschillen is dat soms met meer dan één waarneming op een dag gewerkt wordt. Met dergelijke verschillen in de definitie van een sneeuwdek dag moet rekening worden gehouden indien men de getallen voor Nederland wil vergelijken met die in aangrenzende landen.

De paragrafen 2.2, 2.3 en 2.4 handelen over het gemiddelde aantal dagen in een jaar dat de bodem met sneeuw bedekt is. In de paragrafen 2.2 en 2.3 valt de nadruk op de situatie in Nederland; par. 2.4 gaat daarentegen in op het aantal sneeuwdek dagen in laaggelegen gebieden van de ons omringende landen. In par. 2.5 wordt de kansverdeling van het aantal sneeuwdek dagen besproken. Aangegeven wordt in deze paragraaf hoe men voor een willekeurige plaats in Nederland het aantal sneeuwdek dagen kan bepalen dat met een bepaalde kans (bijv. gemiddeld eens in de 10 jaar) wordt overschreden.

## 2.2 Het gemiddelde aantal dagen met een sneeuwdek

Fig. 2.1 geeft een kaartje van Nederland met het gemiddelde aantal sneeuwdekdagen in een winterhalfjaar. De figuur laat zien dat er vrij grote verschillen binnen Nederland zijn. De getallen lopen uiteen van ongeveer 12 dagen in Zeeland tot ruim 30 dagen in het noordoosten van ons land. Naast een kust-effect is er ook een duidelijk noord-zuid effect (bijv. in Drente meer ochtenden met een sneeuwdek dan in Limburg).

Het sneeuwdekklimaat van een gebied wordt bepaald door de hoeveelheid sneeuw die er valt en door de tijd gedurende welke de sneeuw blijft liggen. Het gemiddelde aantal sneeuwdekdagen op een bepaalde plaats hangt daardoor in sterke mate af van de gemiddelde temperatuur, de hoogteligging en de lokale omgeving van het waarnemingspunt (begroeiing, beschutting tegen de wind e.d.). Voor de in deze studie beschouwde Nederlandse stations is er een sterk verband tussen het gemiddelde aantal sneeuwdekoctenden en de gemiddelde etmaaltemperatuur\*  $\bar{w}$  in de wintermaanden december, januari en februari, zie fig. 2.2. De figuur laat zien dat bij een toename van  $\bar{w}$  met 1°C het gemiddelde aantal sneeuwdekoctenden met ongeveer 10 afneemt. De hoogteligging speelt in fig. 2.2 nauwelijks een rol. Het hoogst gelegen station Beek (L.), dat zich op 114 m boven de zeespiegel bevindt, sluit nog vrij goed bij het gegeven verband aan.

In fig. 2.1 zijn ook de gemiddelden gegeven van een aantal stations in België. Deze gemiddelden zijn verkregen met behulp van gegevens uit De Vogelaere (1976). De getallen in België sluiten vrij matig aan bij de getallen in Nederland. Vermoedelijk spelen lokale terreinomstandigheden en subjectieve criteria van waarnemers hier een belangrijke rol. De invloed van deze twee factoren zal in de volgende paragraaf nader worden gedemonstreerd.

Bij het trekken van de isolijnen in fig. 2.1 is gebruik gemaakt van de gemiddelde wintertemperaturen van een vijftigtal klimatologische stations. In het zuiden van Limburg is het echter niet goed mogelijk om isolijnen te trekken. In heuvelachtige en bergachtige gebieden neemt het gemiddelde aantal sneeuwdekdagen sterk toe met de hoogte. Voor de Ardennen vond De Vogelaere (1976) de volgende waarden voor het gemiddelde aantal sneeuwdekdagen in het tijdvak

\* De gemiddelde etmaaltemperaturen hebben betrekking op het normaal-tijdvak 1951-1980 en zijn ontleend aan KNMI (1982).

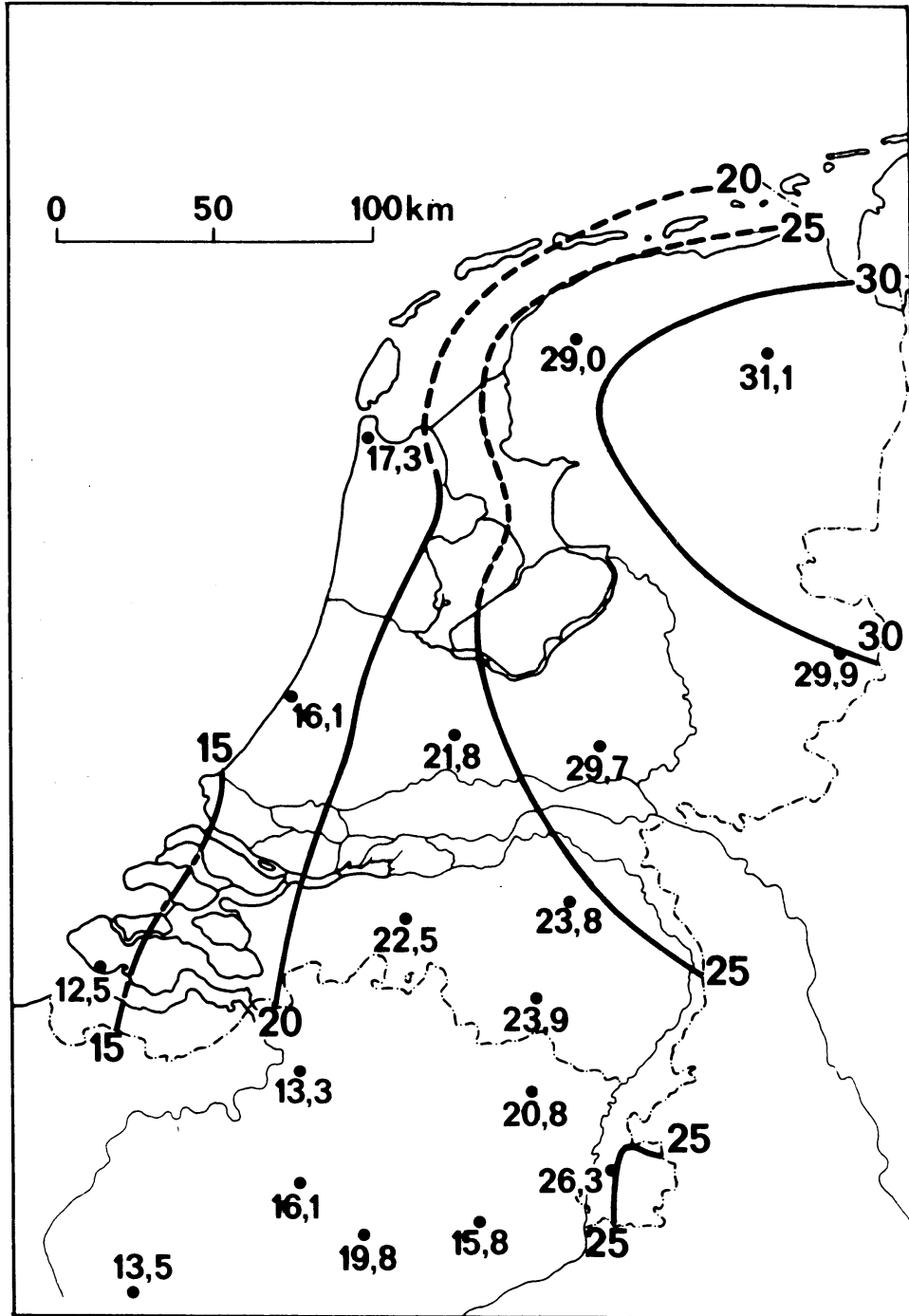


Fig. 2.1. Het gemiddelde aantal sneeuwdekochtenden per winterhalfjaar voor het tijdvak 1955/56-1982/83. Een ochtend wordt een sneeuwdekochtend genoemd als het S-cijfer in tabel 1.1 van nul verschilt.

Mean annual number of days with snow cover (1955/56-1982/83).

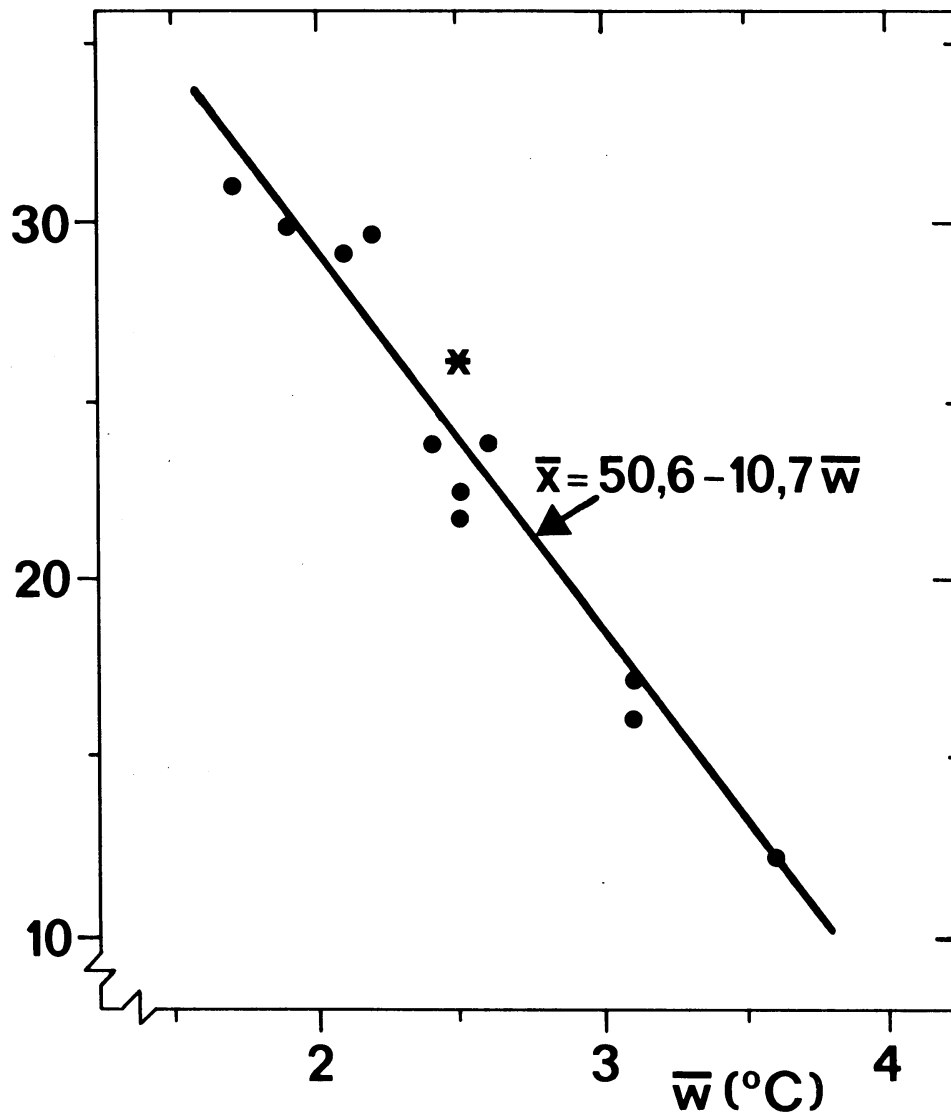
$\bar{x}$  (dagen)

Fig. 2.2. Het verband tussen het gemiddelde aantal sneeuwdekochten per winterhalfjaar  $\bar{x}$  en de gemiddelde wintertemperatuur  $\bar{w}$ . Het gemiddelde aantal sneeuwdekochten van het relatief hooggelegen station Beek (L.) is aangegeven met een sterretje. De getrokken lijn is verkregen met behulp van de kleinste kwadratenmethode. Voor de correlatiecoëfficiënt tussen  $\bar{x}$  en  $\bar{w}$  wordt gevonden  $\text{corr}(\bar{x}, \bar{w}) = -0,97$ .

Relation between the mean annual number of days with snow cover  $\bar{x}$  and the mean winter temperature  $\bar{w}$ . The asterisk indicates the station with the highest elevation (114 m above mean sea level).

1950/51-1974/75:

Bierset (195 m)	21,0 dagen
Florennes (284 m)	32,6 dagen
Luxemburg (378 m)	38,0 dagen
Spa (485 m)	52,8 dagen
St. Hubert (560 m)	61,6 dagen.

Het gemiddelde aantal sneeuwdekdagen neemt in dit geval met ongeveer 11 dagen per 100 m toe. Voor een deel is dit toe te schrijven aan de temperatuurafname met de hoogte (ongeveer  $0,6^{\circ}\text{C}$  per 100 m). Volgens het verband in fig. 2.2 zou deze temperatuurafname een toename van ongeveer 6 sneeuwdekdagen per 100 m opleveren. Het feit dat er een veel sterkere toename van het aantal sneeuwdekdagen met de hoogte is, wordt veroorzaakt door een toename van de gemiddelde neerslaghoeveelheid met de hoogte.

Voor hoogten kleiner dan 200 m is er bij de gegevens van De Vogelaere geen duidelijke toename van het aantal sneeuwdekdagen met de hoogte (het gemiddelde aantal sneeuwdekdagen van het station Bierset ligt in dezelfde orde van grootte als dat van Eindhoven). Het gemiddelde aantal sneeuwdekdagen in Florennes, dat op 284 m ligt, bedraagt 32,6. Vermoedelijk moeten we ook in de hoogste delen van Limburg rekening houden met een gemiddelde van ruim 30 sneeuwdekdagen per jaar.

Of de grond al of niet geheel met sneeuw bedekt is, kan worden opgemaakt uit het synoptische E-cijfer. Voor een aantal laaggelegen stations in België en Nederland blijkt dat in ongeveer 30% van de sneeuwdekochtenden de grond voor minder dan de helft met sneeuw bedekt is, zie tabel 2.1. Bij kuststations is in ruim 40% van het aantal sneeuwdekochtenden de grond geheel met sneeuw bedekt; bij stations in het binnenland bedraagt dit percentage ongeveer 55. De gegevens van de Belgische stations in tabel 2.1 zijn ontleend aan De Vogelaere (1976).

Tot nu toe hebben we ons slechts beperkt tot het gemiddelde jaarlijks aantal sneeuwdekochtenden. Tabel 2.2 geeft voor de stations Vlissingen, De Bilt en Eelde een opsplitsing naar maandgemiddelden. Op deze stations is een sneeuwdek uitsluitend voorgekomen in de periode november t/m april. Het hoogste gemiddelde aantal sneeuwdekochtenden wordt in de maanden januari en februari aangetroffen. Het gemiddelde over deze twee maanden bedraagt ongeveer  $2/3$  van het jaargemiddelde. De verschillen in het gemiddelde aantal sneeuwdekochtenden



Tabel 2.1. Indeling van sneeuwdekochtenden naar de mate van sneeuwbedekking van de grond met percentages van voorkomen bij enkele Belgische en Nederlandse stations.

	Station(s)	Minder dan de helft E=5	Meer dan de helft, maar niet geheel E=6,8	Geheel E=7,9
Kust:	Den Helder/De Kooy	29%	28%	43%
	Vlissingen	36	27	36
	Oostende	38	20	42
	Koksijde	33	25	42
	Gemiddeld	34	25	41
Binnenland:	Eelde	23	19	58
	De Bilt	24	19	56
	Beek (L.)	27	22	51
	Melsbroek	29	18	53
	Kleine Brogel	30	13	57
	Gemiddeld	27	18	55

Tabel 2.2. Gemiddeld aantal sneeuwdekochtenden per maand en per jaar (tijdvak 1955/56-1982/83).

Station	Nov.	Dec.	Jan.	Feb.	Mrt.	April	Jaar
Vlissingen	0,3	2,3	4,8	3,9	1,1	0,1	12,5
De Bilt	0,9	3,9	7,0	7,3	2,3	0,4	21,8
Eelde	1,7	6,3	10,0	10,0	2,6	0,4	31,1

tussen de kust en het binnenland zijn relatief het grootste in het begin van het winterseizoen (november, december). In deze periode is ook het gemiddelde temperatuurverschil tussen de kust en het binnenland groter dan in het verdere verloop van het winterseizoen.

### 2.3. De homogeniteit van het materiaal

In het beoordelen van de toestand van de grond zit een sterk subjectief element. Daarnaast hangt het al of niet voorkomen van een sneeuwdekdag enigszins af van de lokale omstandigheden van het meetterrein. Verandering van waarnemer en/of verplaatsing van een meetstation kunnen daardoor leiden tot een sterke inhomogeniteit in de meetreeks.

Voor de dertien Nederlandse synoptische stations in tabel 1.1 werd nagegaan in hoeverre de reeksen homogeen waren met betrekking tot het aantal sneeuwdekdagen. Bij dit onderzoek werden tevens de reeksen van Borkum, Diepholz, Kleef, Keulen en Ukkel-Brussel meegenomen. In totaal waren dus achttien stations bij de homogeniteitsanalyse betrokken.

Laat nu  $x_1, x_2, \dots, x_N$  de reeks van sneeuwdekdagen zijn in  $N$  opeenvolgende winterhalfjaren voor een bepaald station. We kijken nu naar de volgende verschillenreeks:

$$v_i = x_i - g_i, \quad i = 1, \dots, N \quad (2.1)$$

waarbij  $g_i$  het gemiddelde aantal sneeuwdekdagen is in winterhalfjaar  $i$  van de zeventien overige stations. In het algemeen zal een inhomogeniteit in de reeks  $\{x_i\}$  tot eenzelfde inhomogeniteit leiden in de verschillenreeks  $\{v_i\}$ . Doordat  $x_i$  en  $g_i$  zwaar positief gecorreleerd zijn heeft de verschillenreeks  $\{v_i\}$  een veel kleinere variantie dan de oorspronkelijke reeks  $\{x_i\}$ . De verschillenreeks verdient daardoor de voorkeur voor het toetsen van inhomogeniteiten.

Grafisch kan een inhomogeniteit het beste zichtbaar worden gemaakt door de cumulatieve verschillen:

$$S_k = \sum_{i=1}^k v_i, \quad k=1, \dots, N \quad (2.2)$$

tegen de tijd uit te zetten. Dit is in fig. 2.3 gedaan voor de stations De Bilt en Soesterberg.

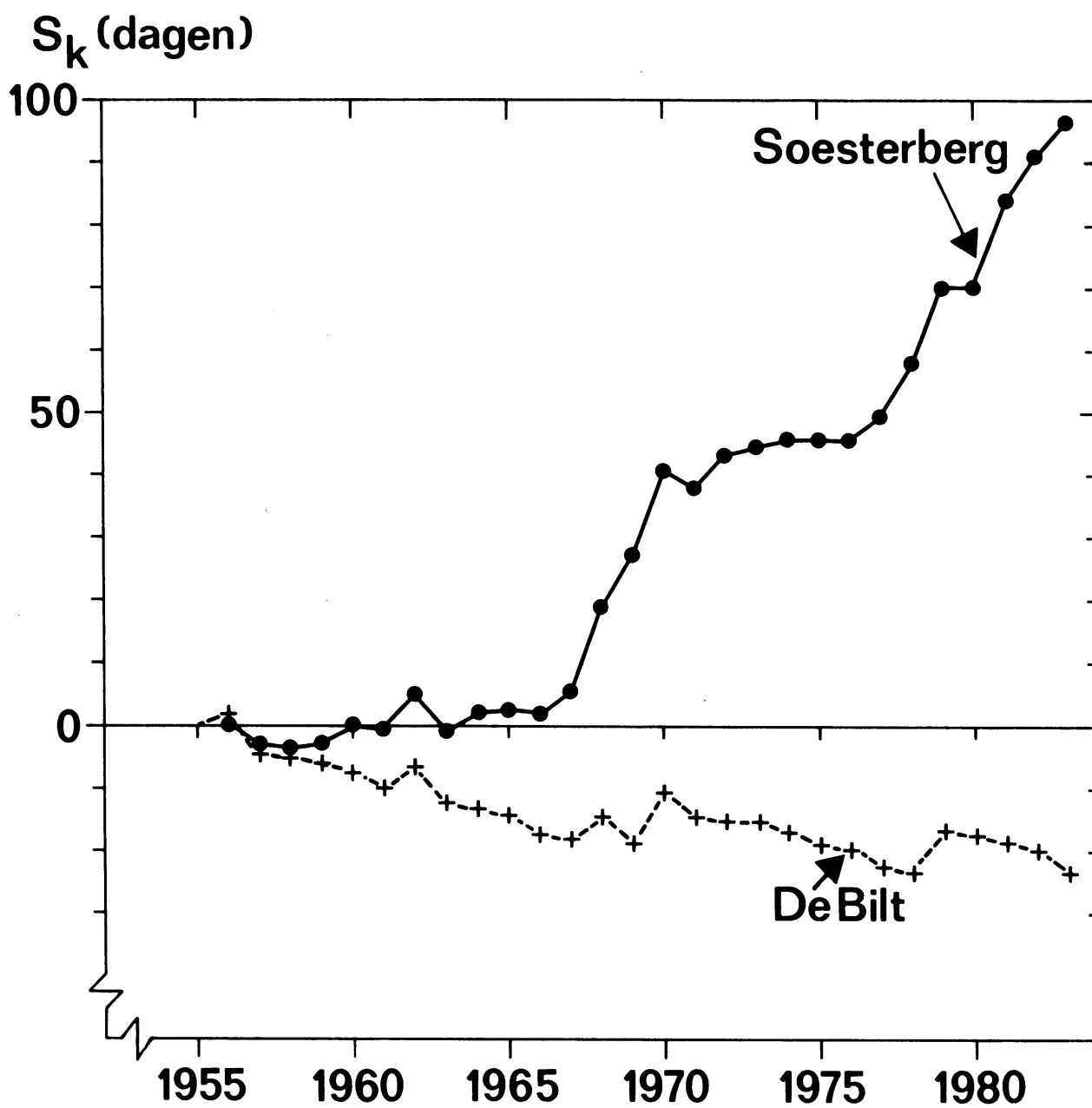


Fig. 2.3. Cumulatieve verschillen  $S_k$  voor de reksen van De Bilt en Soesterberg. De  $S_k$ 's worden gedefinieerd door de vergn. (2.1) en (2.2) in de tekst.

Tot omstreeks 1966 loopt de curve van  $S_k$  voor Soesterberg vrijwel horizontaal, daarna stijgt hij van ongeveer 0 dagen tot 100 dagen in 1983. Dit laatste komt neer op een relatieve toename van ongeveer zes dagen per winterhalfjaar ten opzichte van de overige stations. De curve van De Bilt vertoont niet een dergelijke plotseling verandering.

Een statistische toets op inhomogeniteiten wordt besproken in appendix A. Het blijkt dat de inhomogeniteit bij Soesterberg zeer significant is. Dit station is dan ook weggelaten in de figuren 2.1 en 2.2. Bij de andere Nederlandse synoptische stations konden geen inhomogeniteiten worden aangetoond. Wel was er sterke statistische evidentie voor inhomogeniteiten bij een aantal stations in West-Duitsland.

#### 2.4. Nadere bijzonderheden over het gemiddelde aantal dagen met een sneeuwdek in aangrenzenden landen

Deze paragraaf geeft een literatuuroverzicht over het aantal sneeuwdekdagen in België, West-Duitsland en Engeland. Tevens wordt nagegaan in hoeverre het verband met de gemiddelde wintertemperatuur in fig. 2.2 geldig is voor vlakke gebieden in aangrenzende landen.

##### België

Informatie over het aantal sneeuwdekdagen in België wordt gegeven door Sneyers (1964, 1967) en De Vogelaere (1976). De laatstgenoemde publikatie heeft het voordeel dat eenzelfde definitie van een sneeuwdekdag wordt aangehouden als in deze studie. Het werk van de Vogelaere heeft echter betrekking op het tijdvak 1950/51-1974/75. Voor de stations Vlissingen, Ukkel-Brussel en Beek (L.) blijkt dat er in dit tijdvak gemiddeld ongeveer twee sneeuwdekdagen per jaar meer zijn dan in het tijdvak 1955/56-1982/83. In fig. 2.1 zijn de getallen van De Vogelaere herleid tot gemiddelden voor het tijdvak 1955/56-1982/83.

De publikaties van Sneyers (1964, 1967) gaan uit van meer dan één waarneming op een dag. In zijn publikatie van 1964 over de sneeuwbedekking van de grond te Ukkel-Brussel wordt een dag een sneeuwdekdag genoemd als tijdens minstens één van de drie waarnemingstijden een sneeuwdek aanwezig was ( $S$ -cijfer  $\geq 1$  of  $E$ -cijfer  $\geq 5$ ). In zijn verhandeling over het sneeuwdeklimaat van België gaat Sneyers (1967) noodgedwongen over op een andere definitie. Sneeuwdekdagen zijn hierin dagen waarop tijdens minstens één van de waarnemingsuren meer dan de

helft van de grond met sneeuw bedekt was. Op deze publikatie zal worden teruggekomen in par. 2.5.

#### West-Duitsland

In West-Duitsland gaat men uit van de toestand van de grond om 7 uur 's ochtends (middelbare plaatselijke tijd). Bij geen of weinig sneeuw kende men tot 1965 de volgende notatie:

- . geen sneeuwdek aanwezig ( $S=0$ ).
- 0 de hoogte van het sneeuwdek was  $< \frac{1}{2}$  cm of minder dan de helft van de omgeving van het waarnemingspunt was met sneeuw bedekt.
- F1 "Schneeflecken". Deze term werd gebruikt als er slechts enige niet samenhangende resten waren. Er werd dan geen hoogte opgegeven.

Met ingang van 1965 wordt echter de volgende notatie aangehouden:

- . geen sneeuwdek aanwezig ( $S=0$ ).
- 0 de hoogte van het sneeuwdek was  $< \frac{1}{2}$  cm.
- F1 "Schneeflecken". Deze term betekent nu dat minder dan de helft van de omgeving van het waarnemingspunt met sneeuw bedekt was. Sneeuwresten worden echter buiten beschouwing gelaten. Soms wordt tevens een hoogte opgegeven, die niet gelijk aan nul hoeft te zijn.
- R "Schneereste". Bij sneeuwresten wordt geen hoogte opgegeven.

Als officiële sneeuwdekdag houdt men een dag aan waarvoor een hoogte van het sneeuwdek is opgegeven (zgn. "Schneedecke  $\geq 0$  cm"), zie Caspar (1962), Deutscher Wetterdienst (1957,1960,1964,1967) en Sieland (1979). Deze definitie van een sneeuwdekdag geeft voor de vier Duitse stations in tabel 1.1 de volgende gemiddelden voor het tijdvak 1955/56-1982/83:

Borkum	15,8 dagen
Diepholz	23,7 dagen
Kleef	15,8 dagen
Keulen	15,8 dagen.

Om een goede aansluiting te krijgen met de gemiddelden van de Nederlandse stations in fig. 2.1, zouden we de getallen met ongeveer tien dagen moeten vermeerderen. Dit komt ruwweg overeen met het aantal ochtenden waarop de grond voor minder dan de helft met sneeuw bedekt is, zie tabel 2.1. Tellen we alle ochtenden met "Schneeflecken" en "Schneereste" mee, dan wordt weliswaar een groter gemiddeld aantal sneeuwdekdagen verkregen, maar dit soort reeksen is voor de meeste stations sterk inhomogeen (appendix A) en daardoor voor onder-

linge vergelijking niet geschikt.

Voor laaggelegen gebieden in West-Duitsland kan men het verband met de gemiddelde wintertemperatuur in fig. 2.2 gebruiken. Voor Berlijn ( $\bar{w} \approx 0^\circ\text{C}$ ) geeft Sieland (1979) op grond van een 85-jarige reeks een gemiddelde van 42,1 sneeuwdekdagen. Dit gemiddelde heeft betrekking op "Schneedecke  $\geq 0$  cm". Het gemiddelde aantal dagen met een S-cijfer  $\geq 1$  is ongeveer 10 dagen hoger en bedraagt dus ongeveer 52 voor Berlijn. Dit getal komt goed overeen met de waarde van 50,6 dagen, die uit fig. 2.2 volgt voor  $\bar{w} = 0$ .

### Engeland

In het Verenigd Koninkrijk is het gebruikelijk het aantal ochtenden te tellen waarop om 9 uur UT meer dan de helft van de omgeving met sneeuw bedekt is (Jackson, 1978; Stirling, 1982). Ochtenden met een E-cijfer 5 worden dus niet meegeteld. Uit de gegevens van tabel 2.1 blijkt dat het aantal sneeuwdekdagen met ongeveer een factor 1,5 toeneemt indien ochtenden met een E-cijfer 5 meegeteld worden. Passen we deze correctiefactor toe dan krijgen we voor Londen, Cambridge en York\* de volgende getallen voor het gemiddelde aantal dagen met een S-cijfer  $\geq 1$ :

Londen-Kew ( $\bar{w} = 4,8^\circ\text{C}$ )	13,8 dagen
Cambridge ( $\bar{w} = 4,0^\circ\text{C}$ )	21,8 dagen
York ( $\bar{w} = 4,0^\circ\text{C}$ )	26,3 dagen.

Het gemiddelde aantal sneeuwdekdagen in Londen komt ruwweg overeen met dat in Zeeland; de plaatsen Cambridge en York zijn daarentegen vergelijkbaar met het midden van Nederland. De gemiddelde wintertemperaturen van de Engelse stations zijn echter aanzienlijk hoger dan die in Nederland. Het station met de hoogste gemiddelde wintertemperatuur in fig. 2.2 is Vlissingen waarvoor  $\bar{w} = 3,6^\circ\text{C}$ , wat nog altijd lager is dan de opgegeven waarden voor de Engelse stations. Het verband in fig. 2.2 geeft voor deze stations veel te lage waarden voor het gemiddelde aantal ochtenden met een sneeuwdek.

Dit verschijnsel kan slechts verklaard worden door een hogere sneeuwval in Engeland. Deze hogere sneeuwval moet vermoedelijk voor een belangrijk deel toegeschreven worden aan sneeuwbuien die bij een oostelijk stromingspatroon boven een relatief warm Noordzee-oppervlak ontstaan. Ongeveer 40% van het aantal dagen met sneeuwval in zuidoost Engeland is toe te schrijven aan dit soort sneeuwbuien (Clarke, 1969).

\* Het gemiddelde aantal sneeuwdekdagen is ontleend aan Stirling (1982); de gemiddelde wintertemperaturen  $\bar{w}$  zijn gebaseerd op normaalwaarden voor het tijdvak 1931-1960 in Kalb und Noll (1980).

## 2.5 De kansverdeling van het aantal sneeuwdekdagen

Het gemiddelde aantal sneeuwdekdagen in een bepaald gebied geeft slechts beperkte informatie over het sneeuwdekklimaat. Het aantal sneeuwdekdagen in een winterhalfjaar is immers een kansvariabele\*  $\underline{x}$ . Deze variabele kan in Nederland waarden aannemen van vrijwel nul in zachte winters tot enkele malen het gemiddelde in strenge winters.

### De lognormale verdeling

Fig. 2.4 laat de frequentieverdeling zien van het jaarlijks aantal sneeuwdekdagen in De Bilt. De bovenste grafiek heeft betrekking op het oorspronkelijke aantal sneeuwdekdagen  $\underline{x}$ ; bij de onderste grafiek is de transformatie  $\underline{z} = \ln(\underline{x} + 5)$  toegepast. De verdeling van de getransformeerde aantallen kan beschreven worden door de normale verdeling (verdeling van Gauss). Deze kansverdeling kan men karakteriseren door zijn kansdichtheid:

$$f_z(z) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{1}{2} \frac{(z-\mu)^2}{\sigma^2}\right] . \quad (2.3)$$

De functie  $f_z(z)$  geeft aan hoe de totale kansmassa over de verschillende klassen is verdeeld. De kans dat  $\underline{z}$  in het interval  $(z, z+dz)$  ligt wordt gegeven door:

$$\Pr(z < \underline{z} < z + dz) = f_z(z) dz \quad (2.4)$$

en het verwachte aantal waarnemingen in dit interval is gelijk aan  $N f_z(z) dz$ . De vloeiende curve in de onderste grafiek is op deze wiskundige uitdrukking gebaseerd (met  $dz = 0,5$  en  $N = 28$ ). De curve sluit goed aan bij de gegeven frequentieverdeling van het aantal sneeuwdekdagen in De Bilt.

De parameters  $\mu$  en  $\sigma$  geven resp. de verwachtingswaarde (populatiegemiddelde) en de standaardafwijking van  $\underline{z}$  aan. Deze onbekende parameters worden zodanig gekozen dat de theoretische verdeling zo goed mogelijk bij de waarnemingen past, zie de appendices B en D.

De verdeling van de getransformeerde aantallen sneeuwdekdagen  $\underline{z}$  is symmetrisch t.o.v. de verwachtingswaarde  $\mu$ . Daarentegen is de verdeling van de oorspronke-

\* Ter onderscheiding van gewone variabelen als  $x$  en  $y$ , worden kansvariabelen in dit rapport onderstreept:  $\underline{x}$  en  $\underline{y}$ .

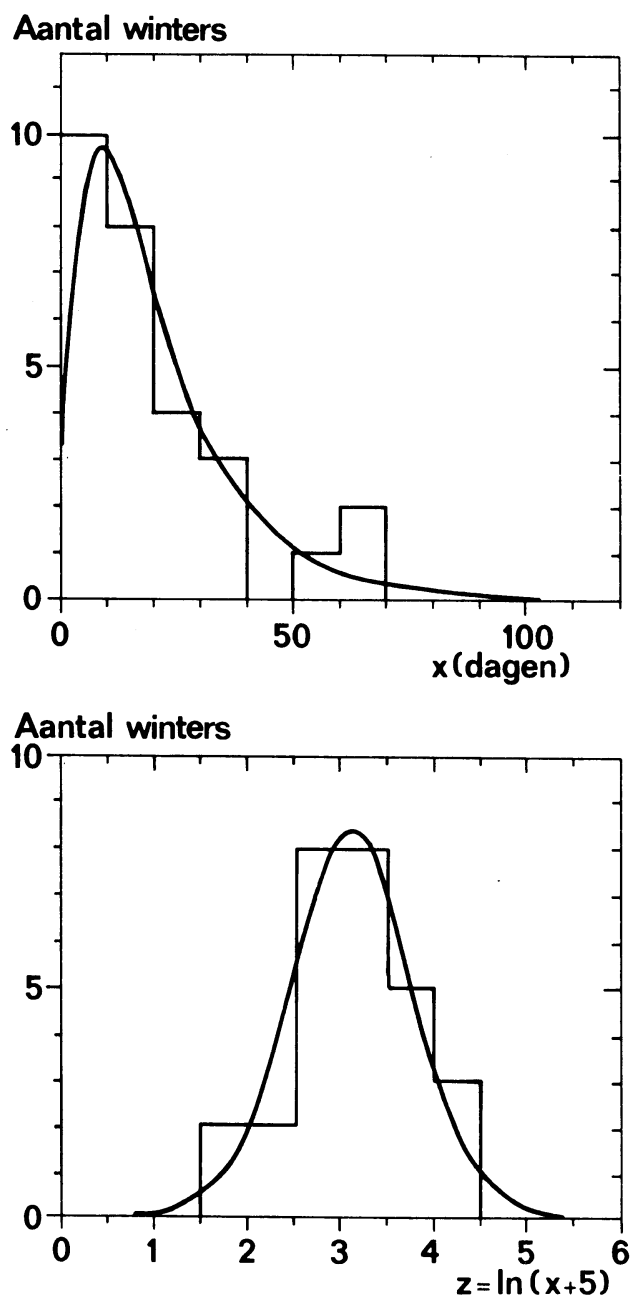


Fig. 2.4. Frequentieverdeling van het aantal sneeuwdekochtenden  $x$  in een winterhalfjaar te De Bilt (1955/56-1982/83). De bovenste frequentieverdeling heeft betrekking op de oorspronkelijke aantallen sneeuwdekochtenden, bij de onderste frequentieverdeling is een transformatie toegepast en wel zodanig dat de kansverdeling van de getransformeerde aantallen op een normale verdeling lijkt.



lijke aantallen  $x$  niet symmetrisch of scheef. Uit fig. 2.4 blijkt dat de kansdichtheid van  $x$  een maximum heeft voor  $x \approx 10$  dagen. Deze zogenaamde modus verschilt aanzienlijk van het gemiddelde, dat ongeveer 22 dagen bedraagt. Links van de modus liggen ongeveer 35% van de waarnemingen; rechts van de modus komen niet alleen meer waarnemingen voor, maar deze kunnen ook zeer sterk van de modus (of het gemiddelde) afwijken. Men spreekt dan in de statistiek van een kansverdeling met een lange rechter staart (positieve scheefheid).

Als  $z = \ln(x-a)$  normaal verdeeld is, dan zegt men dat  $x$  een lognormale verdeling heeft. Hierin is  $a$  een constante. Als  $a=0$  dan spreekt men van de 2-parameter lognormale verdeling. De lognormale verdeling is veel gebruikt om de verdeling van het jaarlijks aantal sneeuwdekdagen te beschrijven (Sneyers, 1967; Jackson, 1978). Voor de constante  $a$  houdt Sneyers een waarde aan van -10 dagen; in deze studie is echter de voorkeur gegeven aan  $a = -5$ , zie appendix D.

De variabele  $x$  is slechts gedefinieerd op het interval  $(a, \infty)$ . Voor de kansdichtheid van  $x$  geldt op dit interval (Johnson and Kotz, 1970, par. 14.1):

$$\begin{aligned} f_x(x) &= \frac{1}{x-a} f_z \{ \ln(x-a) \} \\ &= \frac{1}{(x-a)\sigma \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left\{ \frac{\ln(x-a) - \mu}{\sigma} \right\}^2 \right]. \end{aligned} \quad (2.5)$$

De vloeiende lijn in de bovenste grafiek van fig. 2.4 is op deze vergelijking gebaseerd. Een klein probleem kan zijn dat voor  $a = -5$  er een kleine kans is dat de lognormale variabele een negatieve waarde aanneemt. Op dit probleem keren we later in deze paragraaf nog terug.

#### De cumulatieve verdeling

Naast de kansdichtheid kan men een kansvariabele ook karakteriseren door zijn verdelingsfunctie  $F(x) = \Pr(\underline{x} \leq x)$ . Deze functie geeft dus voor alle niveaus  $x$  de onderschrijdingskans aan en kan geschreven worden als integraal van de kansdichtheid:

$$F(x) = \int_{-\infty}^x f_x(y) dy. \quad (2.6)$$

Indien  $x$  een lognormale verdeling heeft, dan is er geen expliciete uitdrukking voor  $F(x)$ . Voor de bepaling van  $F(x)$  moet men dan gebruik maken van tabellen

of van numerieke benaderingen.

Voor  $x \rightarrow \infty$  gaat (2.6) over in:

$$F(\infty) = \int_{-\infty}^{\infty} f_x(y) dy = \Pr(\underline{x} \leq \infty) = 1 \quad (2.7)$$

d.w.z. de totale oppervlakte onder de kansdichtheidscurve is gelijk aan 1.

Empirisch kan men  $F(x)$  beschrijven door de aantallen waarnemingen in opeenvolgende klassen van een frequentieverdeling op te tellen. Men spreekt dan ook meestal van een cumulatieve verdeling. Het indelen in klassen kan vermeden worden door gebruik te maken van de zogenaamde empirische verdelingsfunctie:

$$F_N(x) = \frac{\text{aantal waarnemingen} \leq x}{\text{totaal aantal waarnemingen } N} \quad (2.8)$$

Merk op dat  $F_N(x) = 0$  als  $x$  kleiner is dan de kleinste waarneming en dat  $F_N(x) = 1$  als  $x$  groter is dan (of gelijk is aan) de grootste waarneming.

Fig. 2.5 geeft de cumulatieve verdeling van het jaarlijks aantal sneeuwdekochtenden in De Bilt. De trapjeslijn van de empirische verdeling sluit goed aan bij de vloeiende curve van  $F(x)$ . Uit de figuur leest men bijvoorbeeld af dat de kans op minder dan 45 sneeuwdekdagen ongeveer gelijk is aan 0,9. Gemiddeld eens in de 10 winters mag men in De Bilt dus minstens 45 sneeuwdekochtenden verwachten. Tussen de gemiddelde herhalingsperiode  $T$  van overschrijding van een niveau  $x$  en de verdelingsfunctie  $F(x)$  bestaat de volgende relatie:

$$T = 1/\{1-F(x)\} \quad (2.9)$$

#### Het gebruik van waarschijnlijkheidspapier

In de praktijk wordt vaak van waarschijnlijkheidspapier gebruik gemaakt om de cumulatieve verdeling weer te geven. Op dit grafiekenpapier zet men de waarnemingen langs de verticale as en de onderschrijdingskansen langs de horizontale as. De laatste as wordt zodanig getransformeerd dat het verband tussen  $x$  en  $F(x)$  door een rechte lijn wordt weergegeven.

In plaats van de empirische verdelingsfunctie zet men de geordende steekproef uit. Deze verkrijgt men door de  $N$  waarnemingen te rangschikken in niet afnemende volgorde

## Onderschrijdingskans $F(x)$

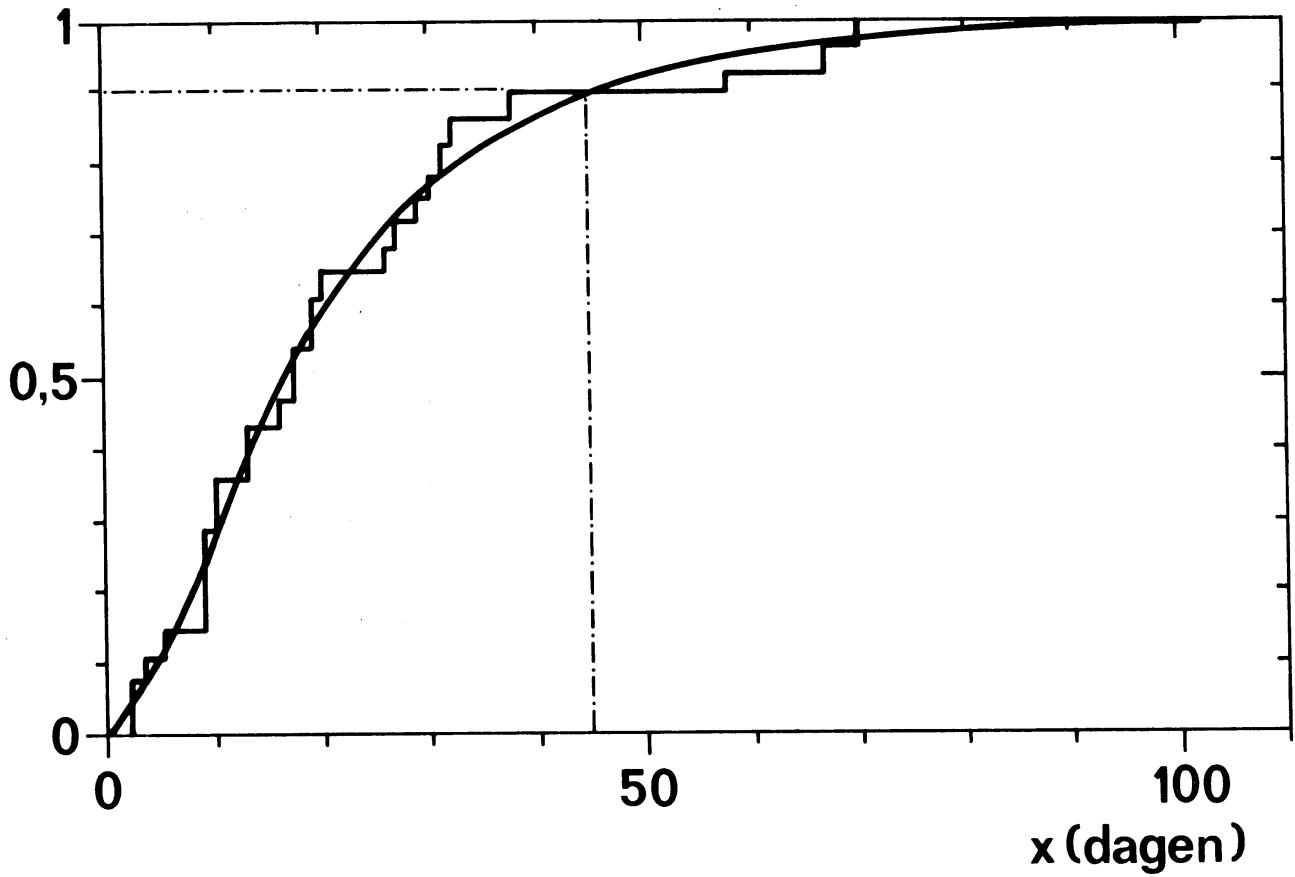


Fig. 2.5. Cumulatieve verdeling van het aantal sneeuwdekochtenden  $x$  in een winterhalfjaar te De Bilt (1955/56-1982/83). De trapjeslijn geeft de empirische verdelingsfunctie  $F_N(x)$ ; de vloeiende curve is een theoretische verdelingsfunctie die gebaseerd is op de lognormale verdeling.

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(N)} \cdot$$

Bij elke waarde van  $x_{(i)}$  wordt een plotkans  $F_i$  bepaald. Voor de normale verdeling wordt de plotkans gegeven door de formule van Blom  $F_i = (i-3/8)/(N+1/4)$ . De plotkans  $F_i$  geeft ruwweg de fractie van het aantal waarnemingen  $\leq x_i$  aan\*. De punten  $(F_i, x_{(i)})$  zullen nagenoeg op een rechte lijn liggen indien de waarnemingen afkomstig zijn uit de verdeling waarvoor het waarschijnlijkheidspapier is geconstrueerd.

Fig. 2.6 geeft voor een drietal stations de verdeling van het aantal sneeuwdekdagen  $x$  op normaal waarschijnlijkheidspapier. In dit geval zijn de getransformeerde aantallen  $z = \ln(x+5)$  normaal verdeeld. Deze hebben een lineaire schaal in fig. 2.6. Uit de figuur lezen we voor De Bilt af dat met 90% kans het aantal sneeuwdekdagen in een bepaald winterhalfjaar minder dan 45 zal bedragen, een resultaat dat ook reeds in fig. 2.5 was aangegeven. Voor Vlissingen en Eelde bedragen deze aantallen resp. 30 en 60.

De punten van de geordende steekproef liggen nagenoeg op een rechte lijn. Een uitzondering vormt echter het station Vlissingen waar voor winters met minder dan drie sneeuwdekdagen de punten sterk van het rechte lijnig verband afwijken, zodat we moeten concluderen dat voor deze waarden van  $x$  de lognormale verdeling niet meer toepasbaar is. Er is echter geen bezwaar om de lognormale verdeling toe te passen voor winters met drie of meer sneeuwdekdagen, d.w.z. voor  $x \geq 2,5$ . Het feit dat de lognormale verdeling niet geldig is bij zeer lage waarden van  $x$  brengt met zich mee dat de onbekende parameters  $\mu$  en  $\sigma$  geschat moeten worden op basis van een geschorste steekproef (Eng.: "censored sample"):

$$x_1, x_2, \dots, x_{N-m}$$

waarbij  $x_1, x_2, \dots, x_{N-m}$  de waarnemingen zijn die de drempelwaarde van 2,5 dag

\* Er zijn kleine verschillen tussen de empirische verdelingsfunctie  $F_N(x)$  en de plotkansen  $F_i$ . De functie  $F_N(x)$  is bedoeld om een schatting voor de onderschrijdingskans  $F(x)$  te krijgen, zonder dat aannamen worden gemaakt omtrent het type kansverdeling. Voor alle  $x$  geldt dat de verwachting (gemiddelde over een oneindig aantal steekproeven) van  $F_N(x)$  gelijk is aan  $F(x)$ . De plotkansen  $F_i$  worden echter zodanig gespecificeerd dat voor de veronderstelde verdeling de verwachtingswaarden van de  $x_{(i)}$ 's nagenoeg op een rechte lijn liggen. Voor nadere bijzonderheden, zie Cunnane (1978).

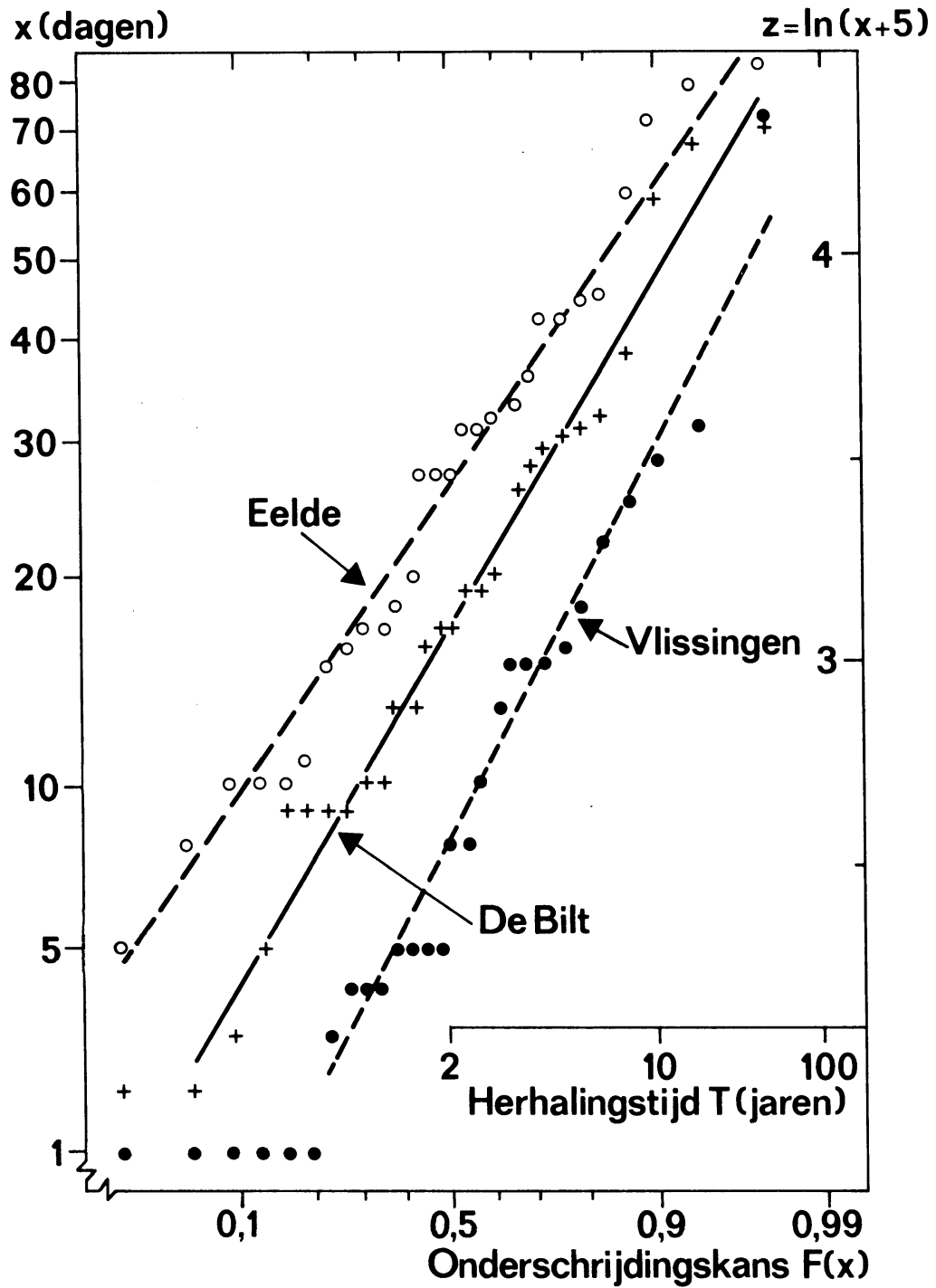


Fig. 2.6. Het aantal sneeuwdekoctenden  $x$  voor de winterhalfjaren 1955/56-1982/83 te Eelde, De Bilt en Vlissingen uitgezet op normaal waarschijnlijkheidspapier.

overschrijden. Het aantal waarnemingen  $m$  dat de drempelwaarde niet overschrijdt wordt tevens in beschouwing genomen. Hierop zal in appendix B nader worden ingegaan. Door ervan uit te gaan dat de lognormale verdeling pas geldig is voor  $x = 2,5$  dag is ook het probleem verdwenen dat bij deze verdeling negatieve waarden mogelijk zijn als  $a < 0$  in (2.5).

#### De relatie tussen de kansverdeling van het aantal sneeuwdekdagen en de gemiddelde wintertemperatuur

Om kansuitspraken te doen over het aantal sneeuwdekdagen voor een bepaalde plaats is slechts kennis van de parameters  $\mu$  en  $\sigma$  voor die plaats nodig. Evenals bij het gemiddelde aantal sneeuwdekdagen het geval was, kan de grootte van deze parameters gerelateerd worden aan de gemiddelde wintertemperatuur  $\bar{w}$ , zie fig. 2.7. Vooral voor de parameter  $\mu$  is er een sterke relatie met  $\bar{w}$ .

De parameter  $\sigma$  neemt toe met  $\bar{w}$ . Een hoge waarde van  $\sigma$  wil zeggen dat de verdeling van  $x$  zeer asymmetrisch (scheef) is. De hoogste waarden van  $\sigma$  worden aangetroffen bij kuststations, die gekenmerkt worden door een laag gemiddeld aantal sneeuwdekdagen, terwijl in lange strenge winters (1962/63!) het aantal sneeuwdekdagen vaak weinig verschilt van dat van stations in het binnenland.

Met behulp van de relaties in fig. 2.7 kunnen we voor een willekeurige plaats in Nederland de grootte van de parameters  $\mu$  en  $\sigma$  afleiden en daarmee tevens het aantal sneeuwdekdagen dat met een bepaalde kans wordt overschreden. Fig. 2.8 geeft de aantallen sneeuwdekdagen die gemiddeld eens in de 2,5,10 en 20 jaar worden overschreden als functie van de gemiddelde wintertemperatuur  $\bar{w}$ . Langs de horizontale as is naast  $\bar{w}$  tevens het gemiddelde aantal sneeuwdekdagen  $\bar{x}$  uitgezet, waarbij gebruik is gemaakt van de relatie in fig. 2.2.

Voor Amsterdam lezen we uit fig. 2.1 af dat er gemiddeld 20 sneeuwdekdagen per jaar zijn. Uit fig. 2.8 lezen we nu voor deze plaats af dat er gemiddeld eens in de 10 jaar een winter is met minstens 40 sneeuwdekdagen.

De lijn bij  $T = 2$  jaar in fig. 2.8 geeft de zogenaamde mediaanwaarde aan. In het Verenigd Koninkrijk wordt aan deze mediaanwaarde de voorkeur gegeven boven het gemiddelde (Jackson, 1978). Uit fig. 2.8 volgt dat de mediaan lager is dan het gemiddelde (positieve scheefheid). Voor de meeste plaatsen in Nederland bedraagt het verschil ongeveer vijf dagen.

De lijnen in fig. 2.8 gelden in principe alleen voor laaggelegen gebieden.

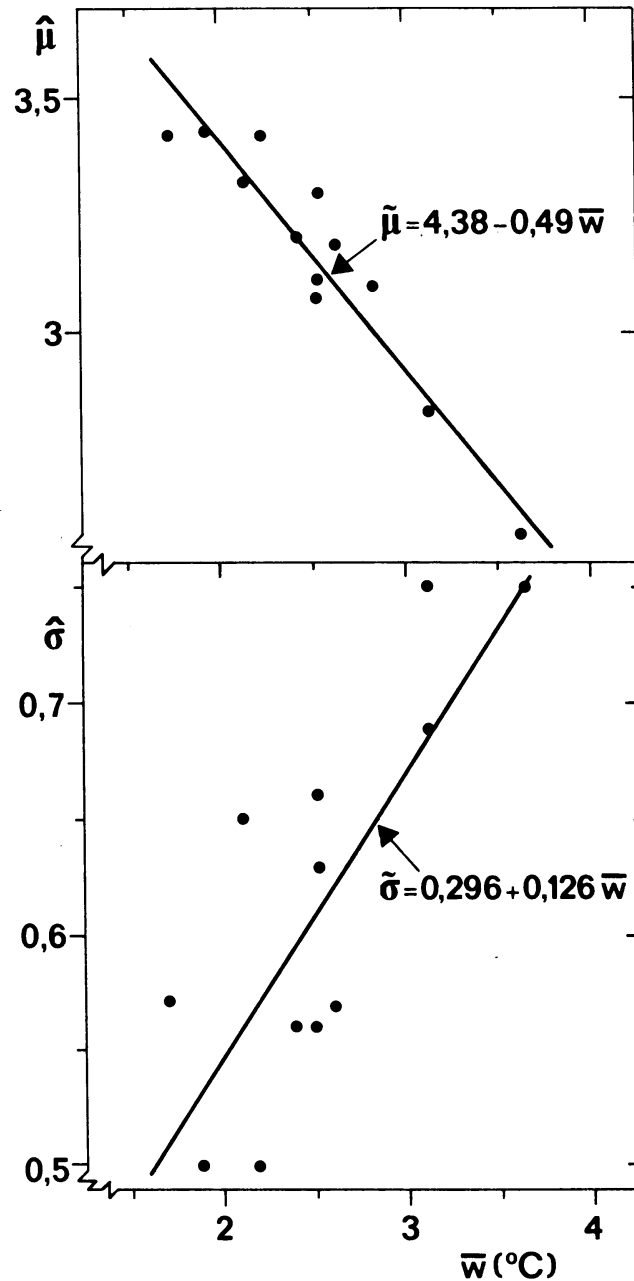


Fig. 2.7. Schattingen  $\hat{\mu}$  en  $\hat{\sigma}$  voor de parameters  $\mu$  en  $\sigma$  van de lognormale verdeling, uitgezet tegen de gemiddelde wintertemperatuur  $\bar{w}$ . Voor de correlatiecoëfficiënten wordt gevonden  $\text{corr}(\hat{\mu}, \bar{w}) = 0,96$  en  $\text{corr}(\hat{\sigma}, \bar{w}) = 0,76$ . De getrokken lijnen zijn verkregen met behulp van de kleinste kwadratenmethode. De waarden voor  $\mu$  en  $\sigma$  die uit deze lijnen volgen zijn aangegeven als  $\tilde{\mu}$  en  $\tilde{\sigma}$ .

Relation between the estimates of the parameters  $\mu$  and  $\sigma$  of the lognormal distribution and the mean winter temperature  $\bar{w}$ .

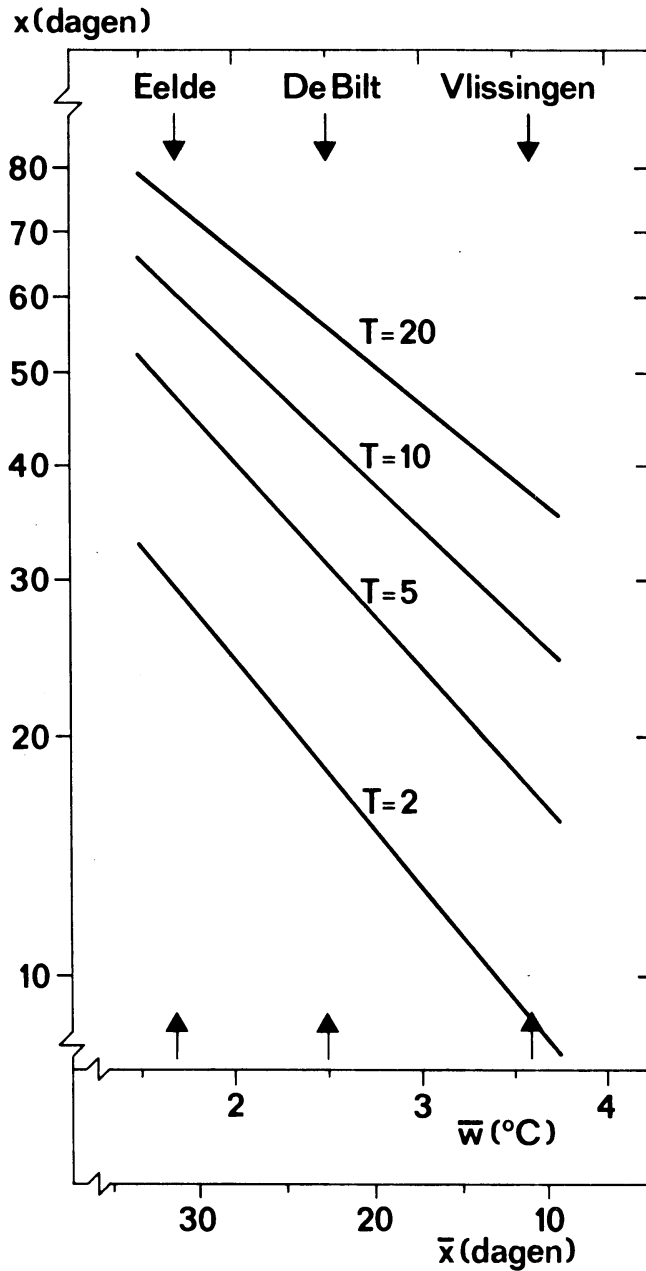


Fig. 2.8. Het aantal sneeuwdekochtenden  $x$  dat gemiddeld eens in de  $T$  jaar in een winterhalfjaar wordt overschreden als functie van de gemiddelde wintertemperatuur  $\bar{w}$  (of het gemiddeld aantal sneeuwdekochtenden  $\bar{x}$ ).

Quantiles of the distribution of the number of days with snow cover in a winter  $x$  as a function of mean winter temperature  $\bar{w}$  (or mean annual number of days with snow cover  $\bar{x}$ ).  $T$  denotes the mean return period of exceedance (years).



Voor plaatsen in België tussen de 200 en 500 m blijkt uit de studie van Sneyers (1967) dat de mediaanwaarde met ongeveer 10 dagen per 100 m toeneemt. Ondanks het feit dat hier van een andere definitie van een sneeuwdek dag wordt uitgegaan dan in de publikatie van De Vogelaere (1976) komt dit getal goed overeen met de in par. 2.2 vermelde toename van 11 dagen per 100 m voor het gemiddelde. Eenzelfde toename wordt ook gevonden voor de aantallen sneeuwdek dagen die gemiddeld eens per 5,10 of 20 jaar worden overschreden.

### 3. DE KANSVERDELING VAN DE JAARMAXIMA VAN DE HOOGTE VAN HET SNEEUWDEK

#### 3.1. Inleiding

Naarmate de hoogte van het sneeuwdek toeneemt zal het meer ongerief met zich meebrengen. Bij daken van gebouwen moet men rekening houden met sneeuwbelastingen die in sneeuwrijke winters kunnen voorkomen. Een eerste indicatie voor de mogelijke waarden van de sneeuwbelasting kan worden verkregen met behulp van een kansverdeling van extreme waarden van sneeuwdekhogten.

In fig. 3.1 is het verloop van de hoogte van het sneeuwdek te De Bilt weergegeven gedurende het tijdvak december 1968 - februari 1969. De figuur geeft twee perioden met een sneeuwdek te zien, namelijk december 1968 en februari 1969. De maximale hoogte van het sneeuwdek in deze winter bedraagt 18 cm. Deze hoogte werd gemeten op 18 februari en is in fig. 3.1 aangegeven met een M. De grootte M wordt meestal kortweg het jaarmaximum genoemd.

Van praktische betekenis zijn waarden van M die gemiddeld eens in de 10 à 50 jaar kunnen voorkomen. Het is daarom soms nodig een kleine extrapolatie toe te passen daar de meeste waarnemingsreeksen niet langer zijn dan ongeveer 30 jaar. Dergelijke extrapolaties zijn slechts toepasbaar indien we iets weten over de vorm van de rechter staart van de kansverdeling van M. Hoe de rechter staart van deze verdeling eruit ziet kan worden nagegaan met behulp van langjarige reeksen van sneeuwdekhogtemetingen in België en Duitsland. Hierop zal worden ingegaan in par. 3.2.

Voor stations in Nederland geeft het bestand van S-cijfers informatie over de kansverdeling van M. In par. 3.3 wordt aangegeven hoe men met dit bestand een kaartje van Nederland kan construeren met sneeuwdekhogten die gemiddeld eens in de T jaar worden overschreden. Als voorbeeld worden kaartjes gegeven voor  $T = 10$  jaar en  $T = 50$  jaar.

In par. 3.4 worden enige opmerkingen gemaakt over de nauwkeurigheid van de resultaten in par. 3.3. Nagegaan wordt in hoeverre het tijdvak 1955/56-1982/83 representatief is voor een langer tijdvak. Tevens wordt ingegaan op het verlies aan informatie dat optreedt bij het vervangen van sneeuwdekhogten door S-cijfers.

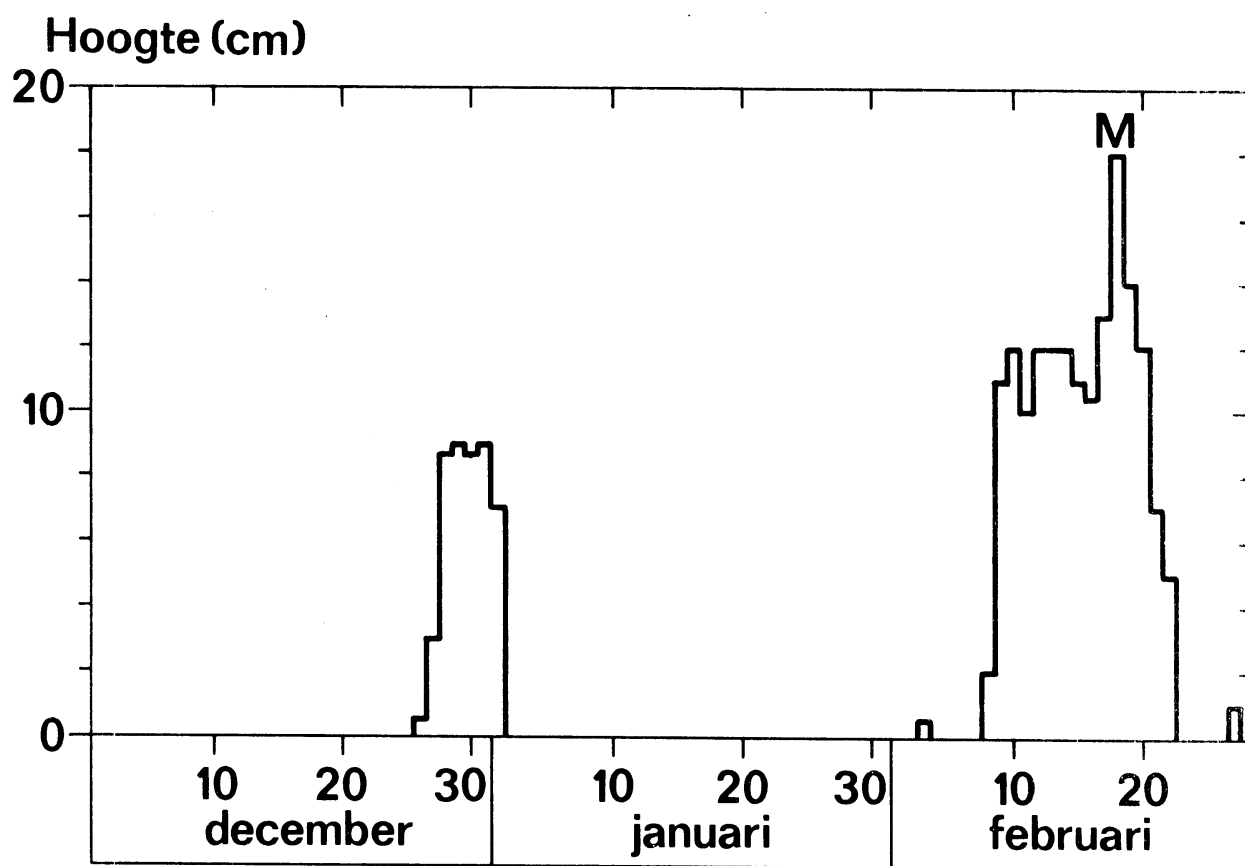


Fig. 3.1. De hoogte van het sneeuwdek (8 uur middelbare plaatselijke tijd) te De Bilt in het tijdvak december 1968 - februari 1969. M geeft het jaarmaximum aan.

### 3.2. Toepassing van de Gumbel-verdeling op jaarmaxima van langjarige reeksen van Diepholz, Kleef en Ukkel-Brussel

Om de verdeling van jaarmaxima te beschrijven wordt vaak gebruik gemaakt van de Gumbel-verdeling. Dit is een kansverdeling met verdelingsfunctie:

$$F(x) = \Pr(\underline{M} \leq x) = \exp\{-e^{-(x-\mu)/\sigma}\} \quad (3.1)$$

en kansdichtheid:

$$f(x) = F'(x) = \frac{1}{\sigma} e^{-(x-\mu)/\sigma} \exp\{-e^{-(x-\mu)/\sigma}\} . \quad (3.2)$$

De verdeling heeft twee onbekende parameters  $\mu$  en  $\sigma$  die op grond van waarnemingen geschat moeten worden\*. Uit (3.1) volgt dat  $F(\mu) = e^{-1} \approx 0,37$ , d.w.z. de parameter  $\mu$  is het niveau dat met een kans van 37% niet overschreden wordt. Dit is tevens de modus van  $\underline{M}$ . De parameter  $\sigma$  is een maat voor de spreiding van de jaarmaxima. Men kan aantonen dat voor de variantie van  $\underline{M}$  geldt  $\text{var } \underline{M} = \pi^2 \sigma^2 / 6$  (Gumbel, 1958, par. 5.2.4). De vorm van de Gumbel-verdeling vertoont veel overeenkomst met die van de lognormale verdeling.

Voor jaarmaxima van sneeuwdekhoogten blijkt de Gumbel-verdeling vaak goed te voldoen. Succesvolle toepassingen zijn te vinden in Sneyers (1964, 1967) voor België en in Jackson (1977) voor het Verenigd Koninkrijk. De laatste auteur geeft tevens aan dat de Gumbel-verdeling goed past bij de gegevens in Schlaak (1968) voor Berlijn en bij de gegevens in Fröhlich (1970) voor Stiermarken (Oostenrijk). Daarnaast is deze kansverdeling veel toegepast bij de jaarmaxima van de waterinhoud van het sneeuwdek (Zerche, 1974; Martinec, 1977; Kuusisto, 1984). Hierbij moet opgemerkt worden dat uitschieters door de eerste twee auteurs weggelaten werden. In de Verenigde Staten geeft men daarentegen vaak de voorkeur aan de 2-parameter lognormale verdeling boven de Gumbel-verdeling (Thom, 1966; Ellingwood and Redfield, 1983, 1984).

De verschillen tussen de Gumbel-verdeling en de 2-parameter lognormale verdeling treden hoofdzakelijk op in de rechterstaart en daar we geïnteresseerd

\* Hier is dezelfde notatie aangehouden als bij de lognormale verdeling. Door Gumbel (1958) worden de parameters van de verdeling aangeduid als  $u(=\mu)$  en  $\alpha(=1/\sigma)$ .

zijn in sneeuwdekhogten, die met een geringe kans worden overschreden, is het belangrijk hierop wat nader in te gaan. Het beste kunnen de verschillen zichtbaar gemaakt worden met behulp van waarschijnlijkheidspapier daar dit grafiekenpapier de rechter staart vergroot weergeeft. In de figuren 3.2, 3.3 en 3.4 zijn de jaarmaxima van resp. Diepholz (58 jaar), Kleef (47 jaar) en Ukkel-Brussel (95 jaar) op Gumbel-papier uitgezet. De Gumbel-verdeling wordt op dit grafiekenpapier door een rechte lijn weergegeven; de gebogen lijn stelt de lognormale verdeling voor.

De verschillen tussen de Gumbel-verdeling en de lognormale verdeling zijn gering voor herhalingstijden  $T$  kleiner dan 10 jaar. De sneeuwdekhogte die gemiddeld eens in de 10 jaar wordt overschreden bedraagt voor alle drie stations ongeveer 20 cm ongeacht welk type verdeling gekozen wordt. Voor grotere herhalingstijden geeft de lognormale verdeling aanzienlijk hogere waarden dan de Gumbel-verdeling. Bij een herhalingstijd van 50 jaar bedraagt het verschil gemiddeld ongeveer 7 cm (25%).

Het verschil tussen de Gumbel-verdeling en de lognormale verdeling is moeilijk te toetsen. Klassieke toetsen als de chi-kwadraat toets en de Kolmogorov-Smirnov toets zijn niet gevoelig voor afwijkingen in de rechter staart (appendix C). Het feit dat in alle drie grafieken de hoogste waarde beneden de curve van de lognormale verdeling ligt wijst erop dat de rechter staart van deze verdeling te lang is. Met name bij het station Diepholz (fig. 3.2) past de curve van de lognormale verdeling slecht bij de grootste gemeten sneeuwdekhogten. Hierop zal nader worden ingegaan in appendix C.

Het hoogste jaarmaximum in de reeks van Kleef bedraagt 42 cm (fig. 3.3). Deze waarde werd gemeten in februari 1942. In het midden van Nederland werden in deze maand ook hoogten van ongeveer 40 cm gemeten (Piekema, 1984). Bij vrijwel alle langjarige reeksen van Duitse stations\* in het grensgebied met Nederland komen sneeuwdekken van 40 cm en meer niet voor, zodat we een dergelijke hoogte zeer uitzonderlijk mogen noemen. Toch ligt de sneeuwdekhogte van 42 cm onder de curve van de lognormale verdeling, wat er weer op wijst dat de rechter staart van de verdeling te lang is.

\* De wintermaxima t/m 1967/68 van alle betrouwbare langjarige reeksen in West-Duitsland zijn gepubliceerd in een rapport van Caspar und Krebs (1974).

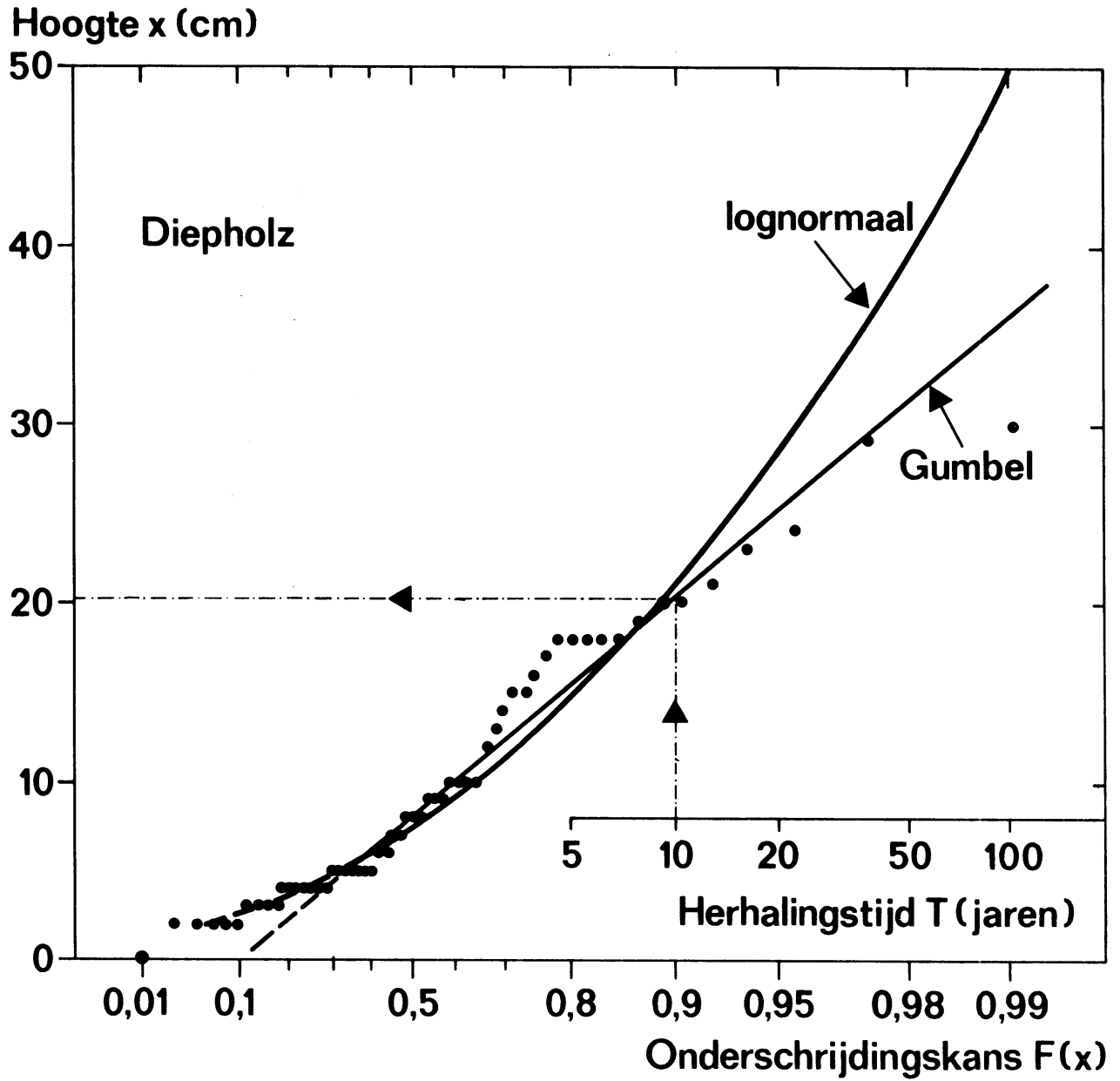


Fig. 3.2. Jaarmaxima van de hoogte van het sneeuwdek in Diepholz (1925/26-1982/83) uitgezet op Gumbel-papier.

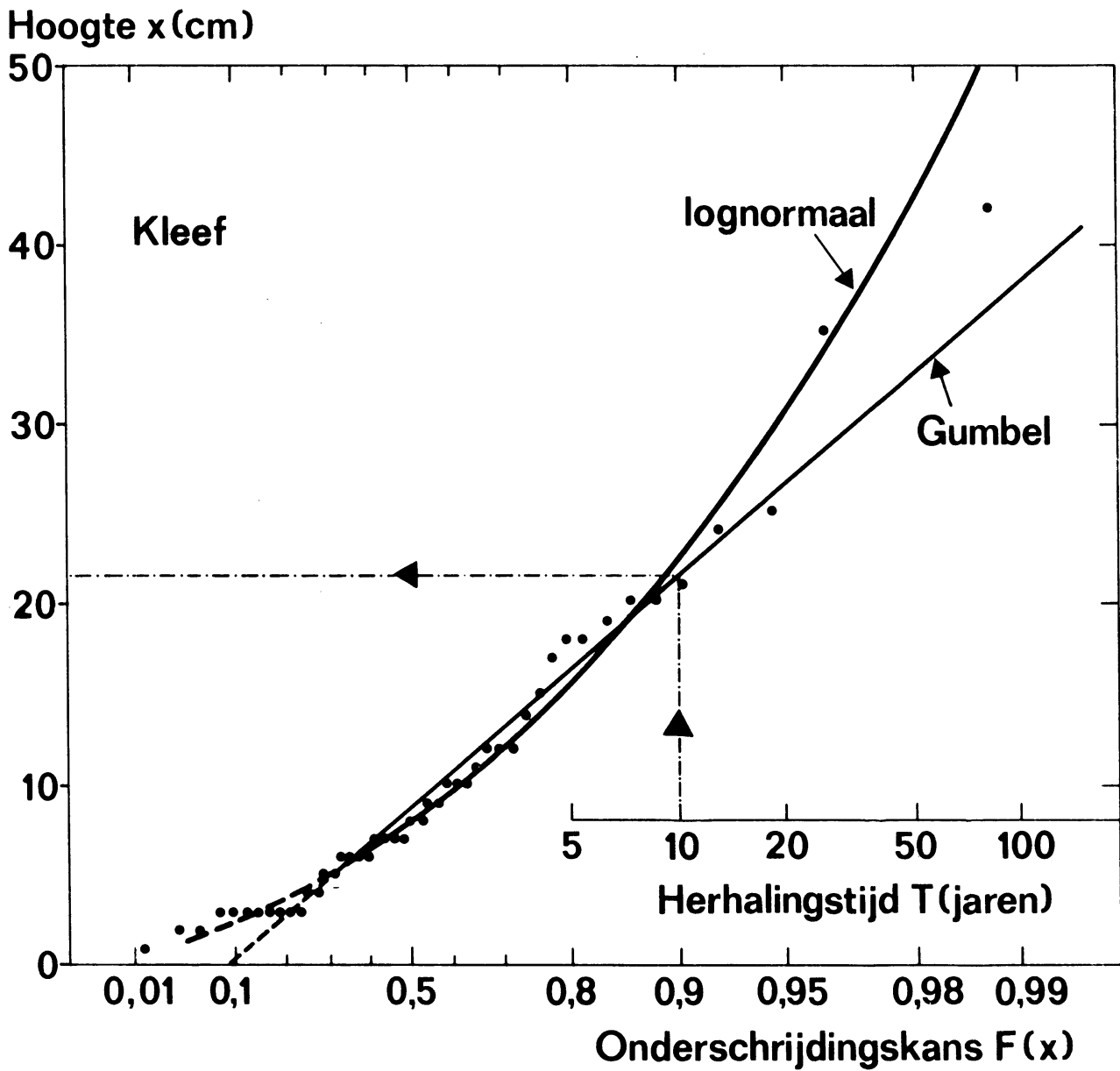


Fig. 3.3. Jaarmaxima van de hoogte van het sneeuwdek in Kleef (1934/35-1982/83 met uitzondering van 1944/45 en 1945/46) uitgezet op Gumbel-papier.

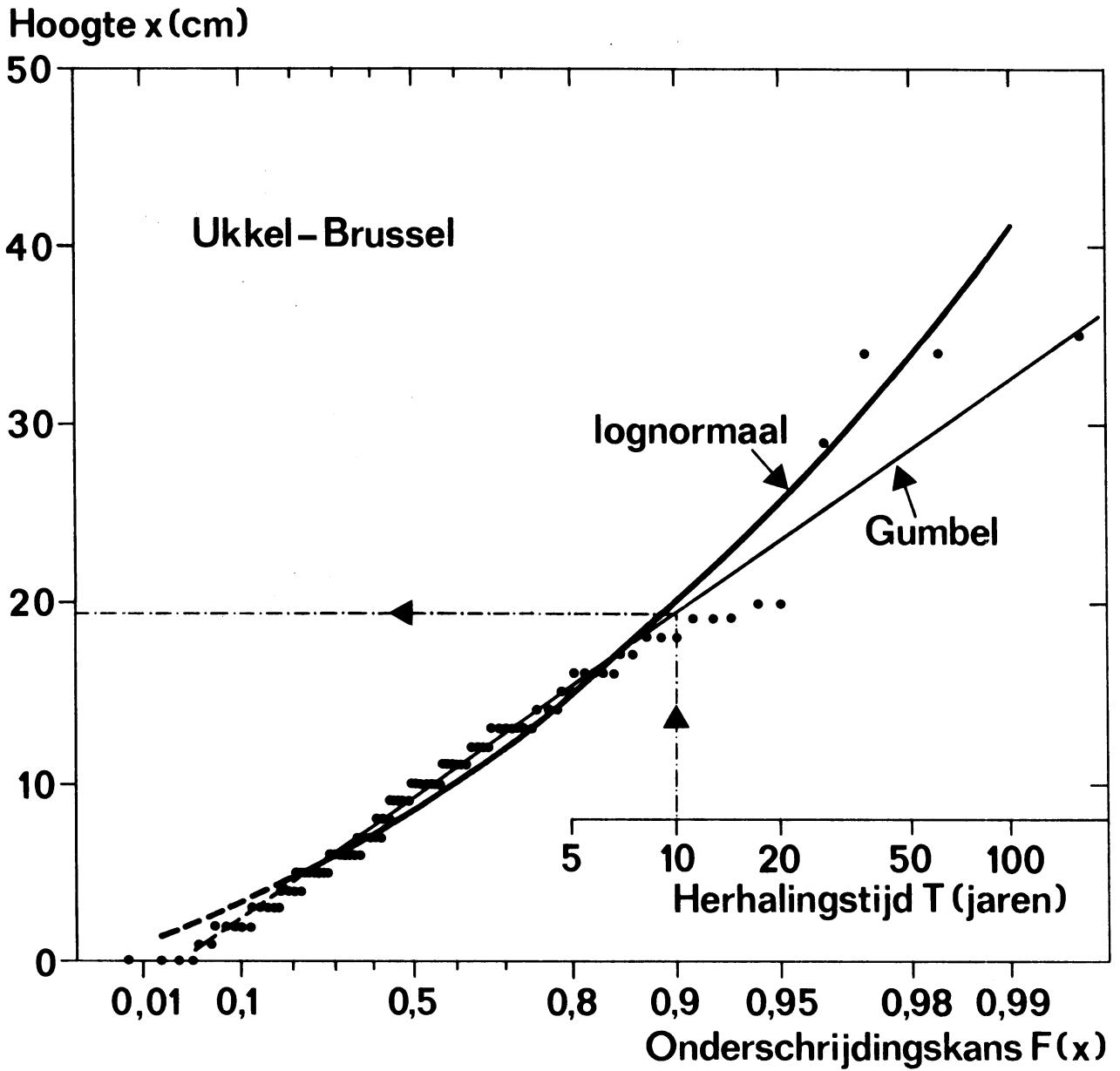


Fig. 3.4. Jaarmaxima van de hoogte van het sneeuwdek te Ukkel-Brussel (1888/89-1982/83) uitgezet op Gumbel-papier.



Op grond van het voorgaande blijkt er enige voorkeur voor de Gumbel-verdeling boven de lognormale verdeling. In de volgende paragrafen zullen we ons daarom uitsluitend concentreren op de Gumbel-verdeling.

Hoewel het praktische belang in de rechter staart van de verdeling ligt is het toch nodig enige woorden te besteden aan de linker staart. Een Gumbel-variabele  $x$  kan waarden aannemen op het interval  $(-\infty, \infty)$ . De Gumbel-rechten in de figuren 3.2, 3.3 en 3.4 geven bij een hoogte van 0 cm een onderschrijdingskans  $F(x)$  van 5 à 10% aan. Evenals bij de lognormale verdeling het geval was (par. 2.5) moeten we er daarom vanuit gaan dat de Gumbel-verdeling pas geldig is bij een bepaalde waarde van  $x$  en de parameters schatten op basis van een geschorste steekproef. Details over de schattingsprocedure worden gegeven in appendix B. Bij de sneeuwdekhogten in de figuren 3.2, 3.3 en 3.4 is ervan uitgegaan dat de Gumbel-verdeling en de lognormale verdeling pas geldig zijn bij een sneeuwdekhogte van 4,5 cm (meet-waarden van 5 cm en meer). Beneden deze drempelwaarde zijn de theoretische verdelingen in de betreffende figuren door een streeplijn weergegeven.

### 3.3 Toepassing van de Gumbel-verdeling op frequentieverdelingen van synoptische codecijfers

In tabel 1.2 werden dertien synoptische stations genoemd waarvoor vanaf 1955 waarnemingsreeksen van het S-cijfer beschikbaar zijn. In plaats van op het jaarmaximum van de gemeten sneeuwdekhogte moeten we ons bij deze stations baseren op het hoogste S-cijfer in het winterhalfjaar\*. Voor elk station kunnen we voor dit jaarmaximum een frequentieverdeling maken, zie tabel 3.1. In de tabel zijn ook de frequentieverdelingen van een vijftal buitenlandse stations opgenomen waarvoor wel gemeten hoogten van het sneeuwdek beschikbaar waren.

De S-cijfers 0,1 en 2 zijn in tabel 3.1 in één klasse samengevoegd. Deze klasse bevat alle jaarmaxima die kleiner zijn dan de in de vorige paragraaf gehanteerde drempelwaarde van 4,5 cm. Het aantal winterhalfjaren waarvoor het S-cijfer altijd  $\leq 2$  is varieert van ongeveer 20% in het binnenland tot 50 à 60% bij de kuststations Den Helder/De Kooy, Valkenburg (Z.H.) en Vlissingen.

\* Vanaf 1981 zijn voor de meeste Nederlandse synoptische stations de werkelijke sneeuwdekhogten beschikbaar. De jaarmaxima van deze sneeuwdekhogten werden herleid tot het oude S-cijfer. Daar het hier nog een vrij kort tijdvak betreft, treedt door deze herleiding nauwelijks verlies aan informatie op.

Bij deze kuststations is een S-cijfer groter dan 5 (sneeuwdekhoogte  $\geq 25$  cm) nooit waargenomen. Dit is ook het geval bij de stations Gilze-Rijen en Eindhoven.

Tabel 3.1. Frequentieverdeling van jaarmaxima van S-cijfers voor het tijdvak 1955/56-1982/83.

Station(s)	Aantal winterhalfjaren met maximale S-waarde				
	$\leq 2$	3	4	5	6
Den Helder/De Kooy	16	5	4	3	-
Leeuwarden	7	7	5	6	3
Eelde	7	7	5	7	2
Twente	5	9	5	6	3
Deelen	6	10	5	4	3
Valkenburg (Z.H.)	14	8	4	2	-
De Bilt	7	7	9	4	1
Soesterberg	8	7	4	7	2
Vlissingen	16	6	4	2	-
Gilze-Rijen	5	11	5	7	-
Volkel	7	10	5	4	2
Eindhoven	7	9	8	4	-
Beek (L.)	6	10	9	2	1
Borkum	8	10	5	3	2
Diepholz	7	8	5	6	2
Kleef	7	11	3	5	2
Keulen	8	8	5	4	3
Ukkel-Brussel	4	8	9	5	2

In par. 3.2 werd aangegeven dat de kansverdeling van de jaarmaxima redelijk door een Gumbel-verdeling kan worden benaderd. Voor elk station kunnen we de parameters  $\mu$  en  $\sigma$  van deze theoretische verdeling zodanig kiezen dat deze zo goed mogelijk aansluit bij de frequentieverdeling in tabel 3.1. In de statistiek staat dit bekend als het aanpassen van een kansverdeling aan tellingen (Eng.: "grouped data"). Bijzonderheden hierover worden gegeven in appendix E. Voor de Nederlandse stations zijn in fig. 3.5 de verkregen geschatte waarden van  $\mu$  en  $\sigma$  uitgezet tegen de gemiddelde wintertemperatuur  $\bar{w}$ . Doordat het gemiddelde jaarmaximum groter wordt naarmate de gemiddelde wintertemperatuur lager is neemt  $\mu$  af met  $\bar{w}$ . De parameter  $\sigma$  neemt ook af met  $\bar{w}$ , echter in veel minder sterke mate. Het quotiënt  $\sigma/\mu$  dat een maat is voor de relatieve variatie (of variatie-coëfficiënt) neemt daardoor toe met  $\bar{w}$ . De hoogste waarden voor dit quotiënt treffen we daarom aan in het kustgebied.

In de figuren 3.6 en 3.7 zijn geschatte waarden van de sneeuwdekhoogten gegeven die gemiddeld eens in de 10 jaar ( $\hat{M}_{10}$ ), resp. eens in de 50 jaar ( $\hat{M}_{50}$ ) worden overschreden. De waarde voor  $\hat{M}_{10}$  loopt uiteen van 13 cm in het zuidwesten tot ongeveer 25 cm in het noordoosten. De waarde voor  $\hat{M}_{50}$  is ongeveer 1,5 maal zo hoog als die voor  $\hat{M}_{10}$ . Relatief zijn de plaatselijke verschillen met betrekking tot  $\hat{M}_{10}$  en  $\hat{M}_{50}$  wat kleiner dan de plaatselijke verschillen bij het gemiddeld jaarlijks aantal sneeuwdekdagen in fig. 2.1.

De getallen bij de stations in de figuren 3.6 en 3.7 berusten op de Gumbel-verdeling die het beste past bij de frequentieverdeling in tabel 3.1. Deze getallen vertonen een vrij onregelmatig verloop over Nederland. Zo zijn bijvoorbeeld de waarden voor Eindhoven en Beek (L.) relatief laag ten opzichte van die van Keulen en Ukkel-Brussel. De oorzaak hiervan is dat de geschatte waarde voor  $\sigma$  vrij onnauwkeurig is. Op de nauwkeurigheid van parameterschattingen zal in de volgende paragraaf uitvoerig worden ingegaan.

De getallen van Eindhoven en Beek (L.) zijn laag ten opzichte van hetgeen op grond van hun gemiddelde wintertemperatuur verwacht mag worden. Op basis van de relaties in fig. 3.5 worden de volgende waarden verkregen:

Eindhoven	$\tilde{M}_{10} = 19$ cm;	$\tilde{M}_{50} = 28$ cm
Beek (L.)	$\tilde{M}_{10} = 19$ cm;	$\tilde{M}_{50} = 29$ cm.

Deze waarden sluiten beter aan bij de waarden van Keulen en Ukkel-Brussel dan de getallen in de figuren 3.6 en 3.7. De isolijnen in de figuren zijn gebaseerd op het verloop van de gemiddelde wintertemperatuur. Evenals in fig. 2.1 werd bij het trekken van de lijnen gebruik gemaakt van temperatuurgegevens van

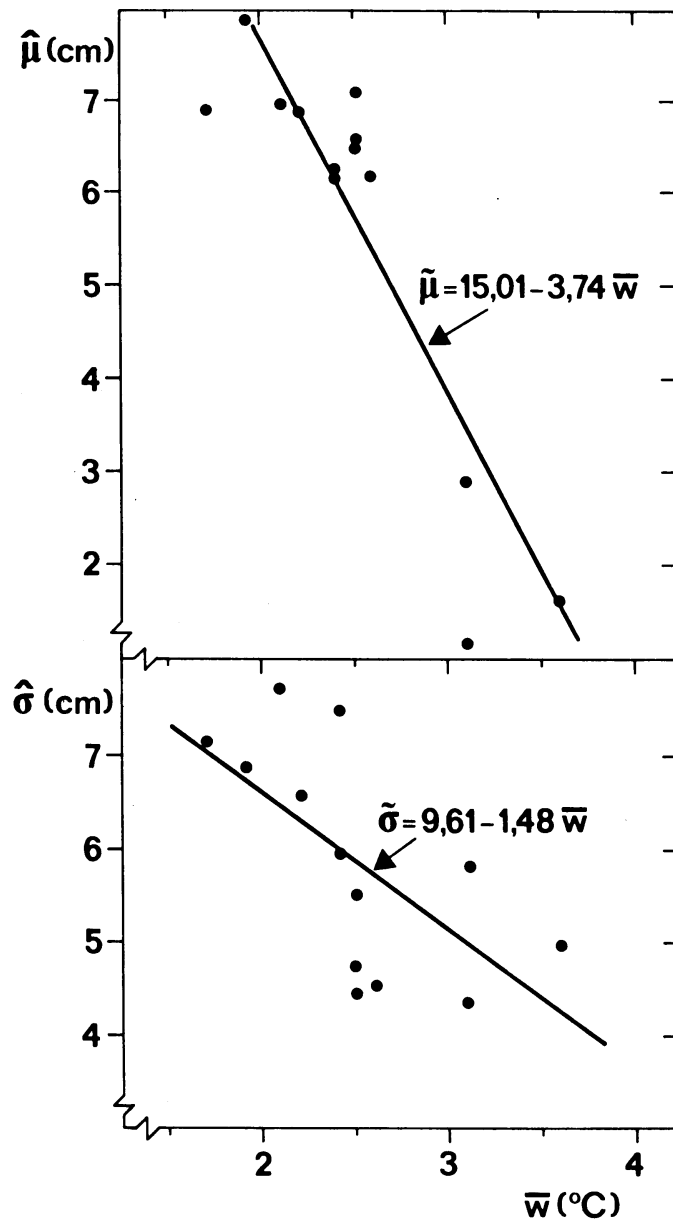


Fig. 3.5. Schattingen  $\hat{\mu}$  en  $\hat{\sigma}$  voor de parameters  $\mu$  en  $\sigma$  van de Gumbel-verdeling, uitgezet tegen de gemiddelde wintertemperatuur  $\bar{w}$ . Voor de correlatiecoëfficiënten wordt gevonden  $\text{corr}(\hat{\mu}, \bar{w}) = -0,88$  en  $\text{corr}(\hat{\sigma}, \bar{w}) = -0,64$ . De getrokken lijnen zijn verkregen met behulp van de kleinste kwadratenmethode. De waarden voor  $\mu$  en  $\sigma$  die uit deze lijnen volgen zijn aangegeven als  $\tilde{\mu}$  en  $\tilde{\sigma}$ .

Relation between the estimates of the parameters  $\mu$  and  $\sigma$  of the Gumbel distribution and the mean winter temperature  $\bar{w}$ .

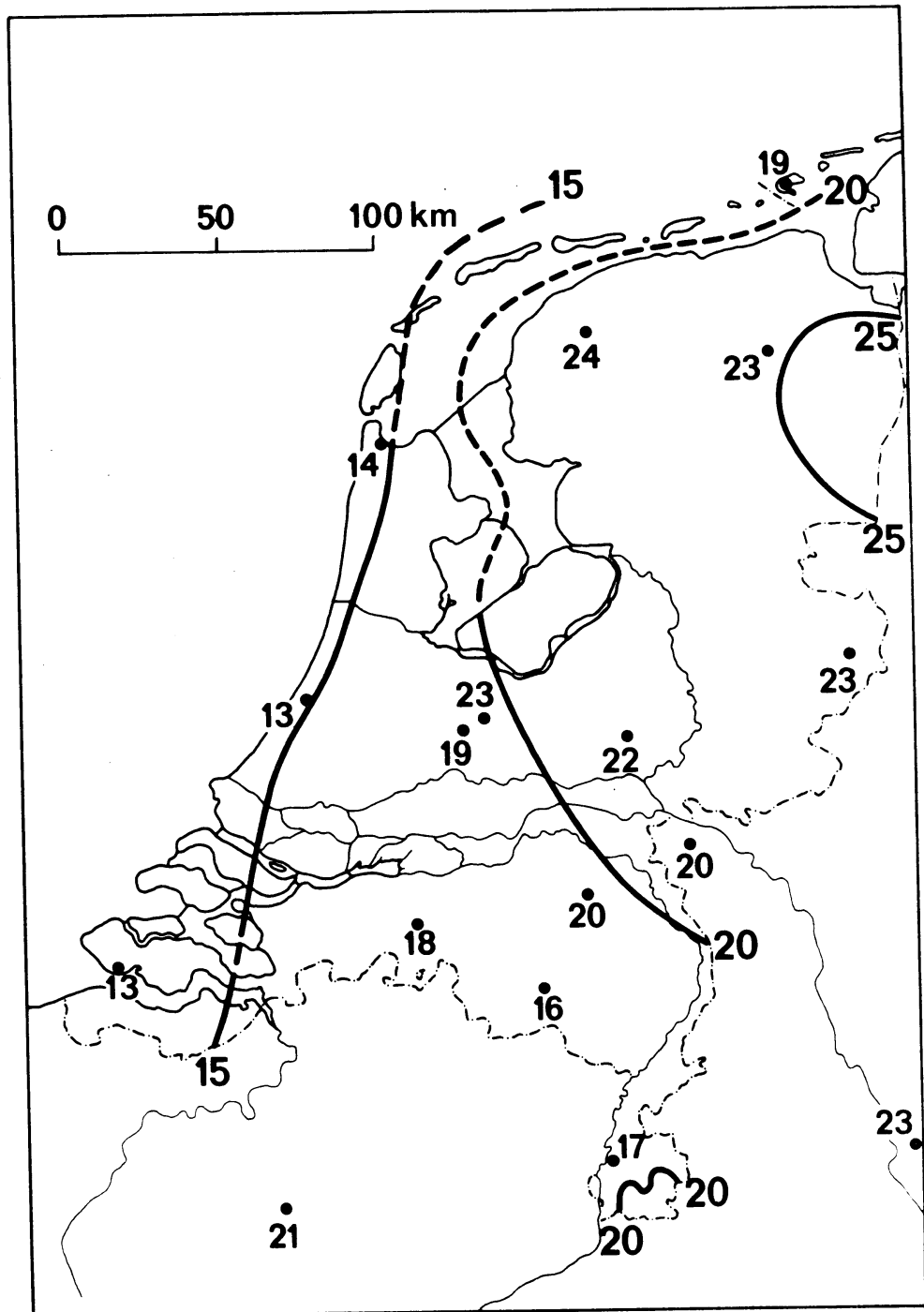


Fig. 3.6. Sneeuwdekhoogten in cm die gemiddeld eens in de 10 jaar ( $\hat{M}_{10}$ ) worden overschreden. De getallen bij de stations zijn schattingen gebaseerd op de frequentieverdelingen in tabel 3.1.

Estimates of 10-year snow depths (cm).

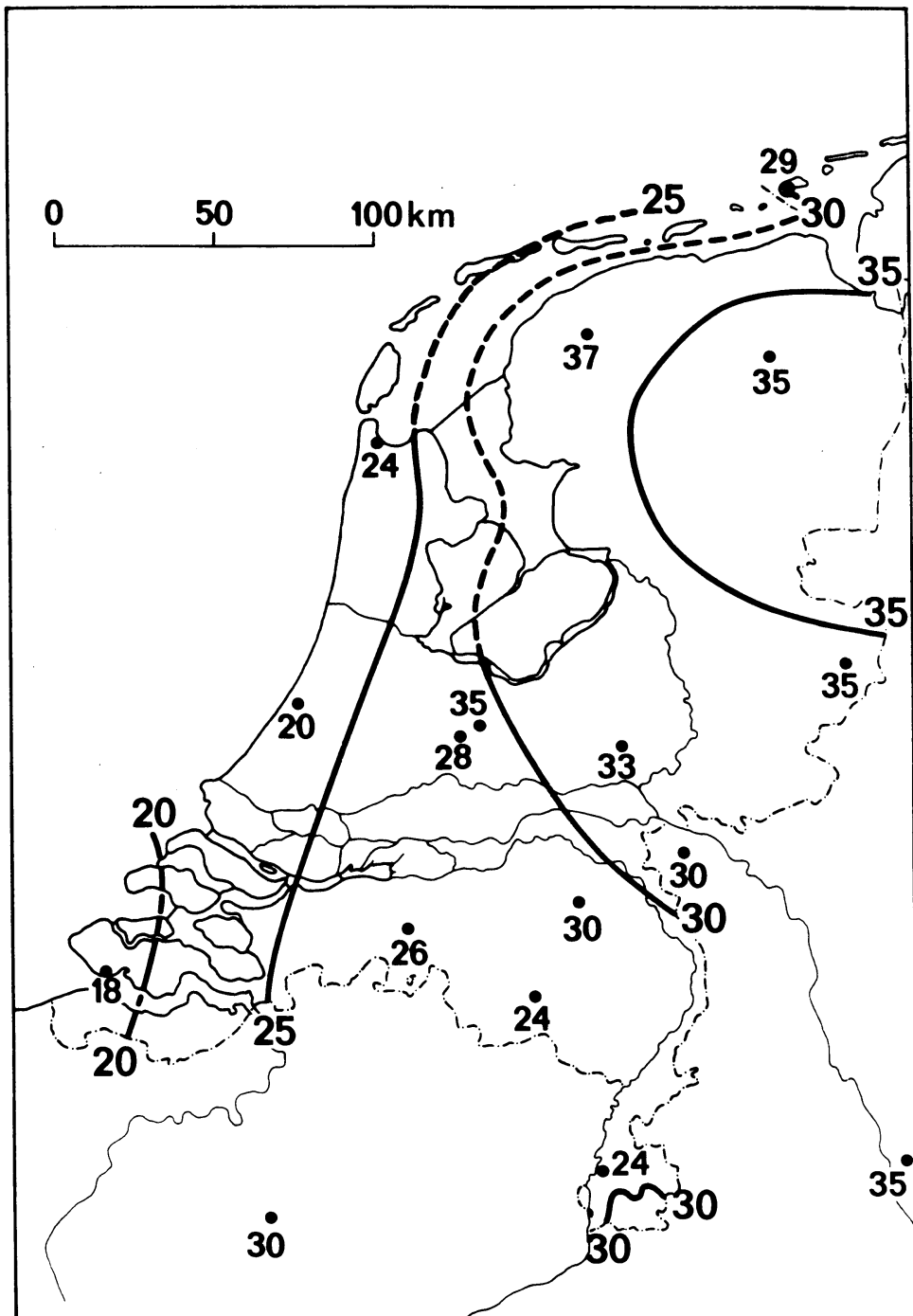


Fig. 3.7. Sneeuwdekhoogten in cm die gemiddeld eens in de 50 jaar ( $\hat{M}_{50}$ ) worden overschreden. De getallen bij de stations zijn schattingen gebaseerd op de frequentieverdelingen in tabel 3.1

Estimates of 50-year snow depths (cm).

een vijftigtal klimatologische stations.

Bij een gemiddelde temperatuurafname met de hoogte van  $0,6^{\circ}\text{C}$  per 100 m zou men op grond van de relaties in fig. 3.5 verwachten dat  $M_{10}$  ongeveer met 4 cm per 100 m toeneemt en  $M_{50}$  met 6 cm per 100 m. Voor België vindt Sneyers (1967) echter iets hogere waarden namelijk 6 en 8 cm voor resp.  $M_{10}$  en  $M_{50}$ , wat weer moet worden toegeschreven aan het orografisch neerslag-effect. Vanwege de toename van  $M_{10}$  en  $M_{50}$  met de hoogte zijn in het zuiden van Limburg extra isolijnen getekend in de figuren 3.6 en 3.7. Deze isolijnen vallen ongeveer samen met de hoogtelijn van 150 m.

Zelfs voor laaggelegen gebieden is het niet toegestaan de temperatuurrelaties in fig. 3.5 te extrapoleren. Voor Berlijn geven deze relaties te hoge waarden voor  $M_{10}$  en  $M_{50}$ . Daarentegen krijgt men voor Engeland, evenals bij het aantal sneeuwdekdagen, te lage waarden voor  $M_{10}$  en  $M_{50}$ .

Voor andere herhalingsstijden dan 10 en 50 jaar kan men  $M_T$  uit de figuren 3.6 en 3.7 schatten met behulp van de relatie:

$$M_T = M_{10} + 0,61 (M_{50} - M_{10}) (y_T - 2,25) \quad (3.3)$$

waarbij:

$$y_T = -\ln\{-\ln(1-1/T)\} . \quad (3.4)$$

Deze relatie berust op de geldigheid van de Gumbel-verdeling. Voor  $T \geq 10$  jaar kan  $y_T$  benaderd worden door  $y_T = \ln(T-1/2)$ .

Met behulp van de figuren 3.6 en 3.7 kan snel beoordeeld worden hoe extreem de hoogte van het sneeuwdek was in de recente strenge winters van 1984/85 en 1985/86. Hiervan was de winter van 1984/85 de meest sneeuwrijke. Toch werd zelfs in deze winter in een groot deel van Nederland de waarde van  $M_{10}$  in fig. 3.6 niet overschreden. Zo was de maximale hoogte van het sneeuwdek in Leeuwarden 23 cm, in Eelde 16 cm, in Deelen 19 cm, in De Bilt 11 cm en in Eindhoven 15 cm. Uitzonderlijk grote sneeuwdekhogten werden in de winter van 1984/85 slechts gemeten in het zuiden van Limburg en op de Waddeneilanden. De maximale hoogte van het sneeuwdek in Beek (L.) bedroeg 30 cm wat ongeveer overeenkomt met de waarde van  $M_{50}$  in fig. 3.7. Op 5 januari 1985 meldde de

regenwaarnemer in Hoorn (Terschelling) een sneeuwdekhogte van 38 cm terwijl de regenwaarnemer in Oost-Vlieland voor 6 januari 1985 een sneeuwdekhogte noteerde van 44 cm. Soortgelijke hoogten van het sneeuwdek op Vlieland en Terschelling zijn ook opgegeven door weeramateurs (Muizelaar en Ybema, 1985). Op grond van de Gumbel-verdeling heeft een hogte van 40 cm in dit gebied een gemiddelde herhalingstijd van ongeveer 1000 jaar. Deze herhalingstijd geldt voor waarnemingen op een punt. De kans dat in een bepaalde winter ergens in Nederland een sneeuwdek met een dergelijke herhalingstijd wordt gemeten is echter veel groter dan 1/1000.

Men dient ook rekening te houden met de mogelijkheid dat door het frequenter voorkomen van sneeuwbuien de rechter staart van de verdeling van de sneeuwdekhogten in het Waddengebied minder goed beschreven wordt door de Gumbel-verdeling. Om dit uit te maken is een uitgebreider waarnemingsmateriaal van dit gebied nodig dan de hier beschouwde reeksen van Den Helder/De Kooy en Borkum.

#### 3.4. Opmerkingen over de nauwkeurigheid van parameterschattingen

In de vorige paragraaf werd gebruik gemaakt van meetreeksen van 28 jaren om de parameters  $\mu$  en  $\sigma$  te schatten. Zou men een ander 28-jarig tijdvak nemen dan krijgt men andere schattingen die vrij sterk kunnen afwijken van de hier gegeven waarden omdat de maximale hogte van het sneeuwdek van jaar tot jaar sterk varieert. In deze paragraaf zullen we de geschatte waarde op basis van een eindige meetreeks als een kansvariabele opvatten. Van deze kansvariabele kunnen we een standaardafwijking berekenen (details hierover worden gegeven in de appendices B2 en E). Is de standaardafwijking groot dan zegt men dat de geschatte waarde onnauwkeurig is.

Tabel 3.2 geeft voor de stations Eelde, De Bilt en Vlissingen de grootte van de standaardafwijking van  $\hat{\mu}$ ,  $\hat{\sigma}$ ,  $\hat{M}_{10}$  en  $\hat{M}_{50}$  bij een reeks van 28 jaren. Hierbij is uitgegaan van de veronderstelling dat de jaarmaxima van de hoogten van het sneeuwdek een Gumbel-verdeling hebben waarvan de parameters  $\mu$  en  $\sigma$  gegeven worden door de relaties met de gemiddelde wintertemperatuur in fig. 3.5. Uit de tabel lezen we bijvoorbeeld af dat de standaardafwijking van  $\hat{M}_{50}$  voor Eelde ruim 5 cm bedraagt. Voor dit station kan men daardoor voor het ene 28-jarige tijdvak een waarde voor  $\hat{M}_{50}$  van 30 cm vinden, terwijl voor een ander sneeuwrijker tijdvak van dezelfde lengte de geschatte waarde voor  $M_{50}$  mogelijk



Tabel 3.2. Correlatiecoëfficiënten en standaardafwijkingen van de schatters  $\hat{\mu}$  en  $\hat{\sigma}$  voor de parameters  $\mu$  en  $\sigma$  van de Gumbel-verdeling en standaardafwijkingen van de waarden, die gemiddeld eens in de 10 en 50 jaar worden overschreden ( $\hat{M}_{10}$  resp.  $\hat{M}_{50}$ ). De getallen gelden voor een steekproefomvang  $N=28$  (aantal jaarmaxima in het tijdvak 1955/56-1982/83). Bij gemeten sneeuwdekhogten moet men van de bovenste rij getallen uitgaan; bij synoptische codecijfers van de onderste rij.

Station	$\tilde{\mu}$ (cm)	$\tilde{\sigma}$ (cm)	corr ( $\hat{\mu}, \hat{\sigma}$ )	Standaardafwijking (cm)			
				$\hat{\mu}$	$\hat{\sigma}$	$\hat{M}_{10}$	$\hat{M}_{50}$
Eelde	8,7	7,1	0,07	1,48	1,24	3,25	5,17
			0,08	1,50	1,33	3,44	5,50
De Bilt	5,7	5,9	-0,15	1,37	1,16	2,76	4,52
			-0,15	1,39	1,24	2,93	4,83
Vlissingen	1,5	4,3	-0,65	1,80	1,17	2,00	3,65
			-0,67	1,86	1,24	2,08	3,86

Tabel 3.3. Geschatte waarden van sneeuwdekhogten die gemiddeld eens in de 10 en 50 jaar worden overschreden ( $\hat{M}_{10}$  resp.  $\hat{M}_{50}$ ) zowel op basis van de langjarige reeksen in de figuren 3.2, 3.3 en 3.4 als op basis van de jaarmaxima in het tijdvak 1955/56-1982/83.

Station	Langjarige reeks		Tijdvak 1955/56-1982/83	
	$\hat{M}_{10}$ (cm)	$\hat{M}_{50}$ (cm)	$\hat{M}_{10}$ (cm)	$\hat{M}_{50}$ (cm)
Diepholz	20	31	22	33
Kleef	22	33	20	30
Ukkel-Brussel	19	29	21	30

Tabel 3.4. Schattingen  $\hat{\mu}$  en  $\hat{\sigma}$  voor de parameters  $\mu$  en  $\sigma$  van de Gumbel-verdeling zowel op basis van de gemeten hoogten als op basis van de jaarmaxima van S-cijfers in het tijdvak 1955/56-1982/83.

Sneeuwdekreeks	Gemeten hoogten		Synoptisch codecijfer S	
	$\hat{\mu}$ (cm)	$\hat{\sigma}$ (cm)	$\hat{\mu}$ (cm)	$\hat{\sigma}$ (cm)
Midden Nederland	6,0	5,7	6,0	5,7
Borkum	5,3	5,6	5,6	5,9
Diepholz	6,3	6,8	6,7	6,8
Kleef	6,2	6,1	6,1	6,2
Keulen	6,2	6,6	5,9	7,4
Ukkel-Brussel	8,2	5,6	8,3	5,7

40 cm zal bedragen. Gezien de vrij grote spreiding in mogelijke uitkomsten van  $\hat{M}_{10}$  en  $\hat{M}_{50}$  doet zich de vraag voor of het tijdvak 1955/56-1982/83 al of niet een uitzonderlijk tijdvak is geweest.

Uit tabel 3.3 blijkt dat voor de stations Diepholz, Kleef en Ukkel-Brussel de waarden van  $\hat{M}_{10}$  en  $\hat{M}_{50}$  voor de gehele langjarige reeks weinig verschillen van de waarden op grond van gegevens over het tijdvak 1955/56-1982/83. Hieruit kan geconcludeerd worden dat dit tijdvak redelijk representatief is voor een langer tijdvak.

Naast een onzekerheid tengevolge van de lengte van de meetreeks is er in de waarde van  $M_{50}$  nog een extra onzekerheid tengevolge van het feit dat de Gumbel-verdeling deze waarde mogelijk systematisch onderschat of overschat.

Doordat bij het indelen in klassen verlies aan informatie optreedt krijgt men met synoptische codecijfers wat hogere waarden voor de standaardafwijkingen van  $\hat{\mu}$ ,  $\hat{\sigma}$ ,  $\hat{M}_{10}$  en  $\hat{M}_{50}$  dan wanneer men van de werkelijk gemeten hoogten uitgaat. De getallen in tabel 3.2 laten echter zien dat de verschillen gering zijn, namelijk 1 à 3% voor de standaardafwijking van  $\hat{\mu}$  en ruim 5% voor de standaardafwijkingen van  $\hat{\sigma}$ ,  $\hat{M}_{10}$  en  $\hat{M}_{50}$ .

Voor een zestal reeksen worden in tabel 3.4 de geschatte waarden van  $\mu$  en  $\sigma$  op basis van gemeten hoogten en synoptische codecijfers onderling vergeleken. De grootste verschillen worden aangetroffen bij het station Keulen. De geschatte Gumbel-verdelingen voor dit station zijn in fig. 3.8 op waarschijnlijkheidspapier uitgezet. Beide rechten sluiten op het oog zeer goed aan bij de geplotte punten van de geordende steekproef.

Aan het slot van deze paragraaf keren we nogmaals terug naar de getallen in tabel 3.2. Opvallend is dat de standaardafwijking van  $\hat{\mu}$  het grootste is bij het kuststation Vlissingen, terwijl de standaardafwijkingen van  $\hat{\sigma}$ ,  $\hat{M}_{10}$  en  $\hat{M}_{50}$  juist relatief hoog zijn bij het station Eelde. Bij deze plaatselijke verschillen spelen de volgende factoren een rol:

- (i) De grootte van de parameter  $\sigma$ . De standaardafwijkingen van parameterschattingen zijn evenredig aan  $\sigma$ . De hoogste waarde van  $\sigma$  wordt bij stations in het binnenland aangetroffen.
- (ii) De fractie geschorste waarnemingen. Het schorsen van waarnemingen werkt verhogend op de standaardafwijking van parameterschattingen. De fractie geschorste waarnemingen is het hoogste voor reeksen van kuststations. Bij het aanpassen van de Gumbel-verdeling aan synoptische codecijfers

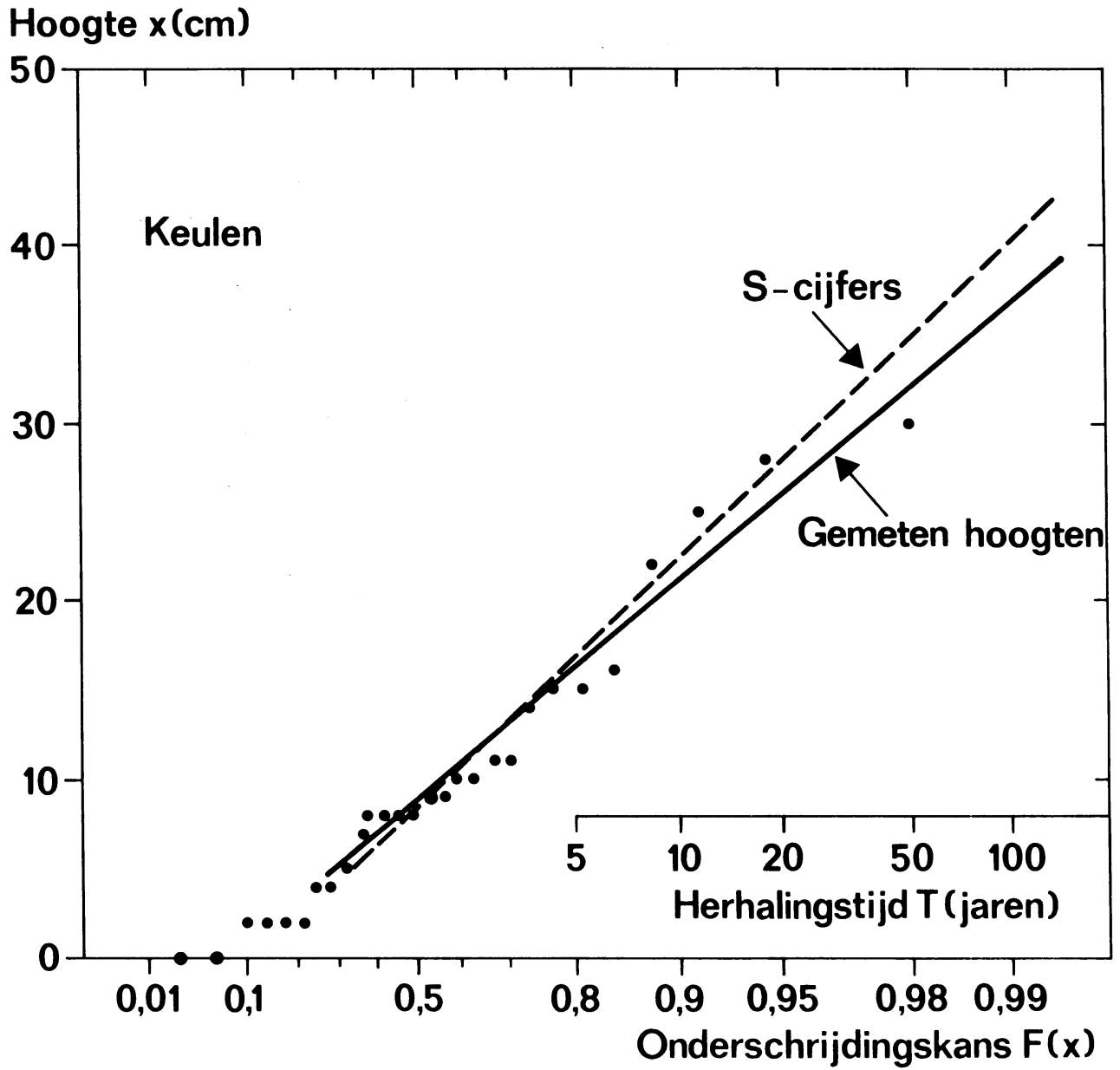


Fig. 3.8. Jaarmaxima van de hoogte van het sneeuwdek in Keulen (1955/56-1982/83) uitgezet op Gumbel-papier. De lijn op basis van de gemeten hoogten wijkt iets af van die op basis van de S-cijfers.

is ook de situatie bij kuststations ongunstig omdat ruwweg de helft van het aantal waarnemingen in de klasse  $S \leq 2$  ligt.

- (iii) De correlatie tussen  $\hat{\mu}$  en  $\hat{\sigma}$ . Een negatieve correlatie werkt verlagend op de standaardafwijking van  $\hat{M}_{10}$  en  $\hat{M}_{50}$ . Wat dit punt betreft is de situatie ongunstig bij stations in het binnenland.

Bij de plaatselijke verschillen in de grootte van de standaardafwijking van  $\hat{\mu}$  overheerst het effect van (ii); bij de standaardafwijking van  $\hat{\sigma}$  wordt dit effect echter volledig gecompenseerd door (i). Punt (iii) speelt een dominerende rol bij plaatselijke verschillen in de standaardafwijkingen van  $\hat{M}_{10}$  en  $\hat{M}_{50}$ .

#### 4. RUNLENGTEN EN DE KANSVERDELING VAN DE HOOGTE VAN HET SNEEUWDEK

##### 4.1. Inleiding

Dagen met een sneeuwdek liggen min of meer in groepjes bij elkaar. Zo laat fig. 3.1 zien dat het tijdvak december 1968-februari 1969 in De Bilt naast twee solitaire dagen met een sneeuwdek ook twee langere perioden kende waarop de bodem met sneeuw bedekt was. Een periode van opeenvolgende sneeuwdekdagen die aan beide zijden begrensd wordt door een dag zonder sneeuwdek wordt een run van sneeuwdekdagen genoemd. In fig. 3.1 kunnen we vier runs van sneeuwdek-  
dagen onderscheiden waarvan er twee een lengte van één dag hebben. De twee andere runs hebben lengten van respectievelijk 7 en 15 dagen. De kansverdeling van deze runlengten wordt in par. 4.2 besproken.

Behalve naar het al of niet voorkomen van een sneeuwdek kunnen we ook kijken naar dagen waarop het sneeuwdek een bepaalde hoogte bereikt of overschreden heeft. Uit fig. 3.1 kan afgelezen worden dat het tijdvak december 1968-februari 1969 in De Bilt 19 ochtenden telde met een sneeuwdek van minstens 5 cm. In deze winter waren er twee runs van dagen met een sneeuwdek van minstens 5 cm waarvan de eerste een lengte had van 5 dagen en de tweede een lengte van 14 dagen.

Paragraaf 4.3 geeft nadere informatie over het gemiddelde aantal dagen dat het sneeuwdek een bepaalde hoogte overschrijdt. In par. 4.4 wordt ingegaan op de runs van dagen met een sneeuwdek boven een bepaald niveau.

Om de lengte van de diverse tabellen te beperken worden in dit hoofdstuk alleen de sneeuwdekgegevens van de vijf klassieke hoofdstations Den Helder/De Kooy, De Bilt, Eelde, Vlissingen en Beek (L.) behandeld. Soms zal echter een vergelijking worden gemaakt met de uitkomsten voor sneeuwdekgegevens van Belgische of Duitse stations.

##### 4.2. Runlengten van sneeuwdekdagen

In tabel 4.1 is een aantal eigenschappen van runs van sneeuwdekdagen samengevat. Uit de tabel blijkt dat er relatief veel korte runs zijn met een lengte van één dag of enkele dagen. De langste run heeft voor alle stations een lengte van ongeveer 70 dagen. Deze zeer zeldzame lange sneeuwdekperiode

Tabel 4.1. Eigenschappen van runs van sneeuwdekochtenden. Een ochtend wordt een sneeuwdekochtend genoemd als het S-cijfer in tabel 1.1 van nul verschilt. De aantallen hebben betrekking op de 28 winterhalfjaren 1955/56-1982/83.

Station(s)	Totaal aantal runs	Aantal runs met lengte (dagen)					Langste run (dagen)	Gemiddel- de lengte (dagen)
		1	2-4	5-9	10-19	$\geq 20$		
Den Helder/De Kooy	108	42	39	17	7	3	71	4,5
Eelde	190	88	57	26	11	8	76	4,6
De Bilt	145	62	48	22	9	4	69	4,3
Vlissingen	85	41	23	14	5	2	70	4,2
Beek (L.)	177	78	55	26	14	4	71	4,2

Tabel 4.2. Procentuele verdeling van runlengten van sneeuwdekochtenden voor enkele Belgische en Nederlandse stations.

Station(s)	Percentage per lengteklasse				
	1	2-4	5-9	10-19	$\geq 20$
<b>Kust:</b>					
Den Helder/De Kooy	39%	36%	16%	6%	3%
Vlissingen	48	27	16	6	2
Oostende	46	34	15	4	1
Koksijde	43	39	12	4	2
Gemiddeld	44	34	15	5	2
<b>Binnenland:</b>					
Eelde	46	30	14	6	4
De Bilt	43	33	15	6	3
Beek (L.)	44	31	15	8	2
Melsbroek	43	34	17	3	3
Kleine Brogel	40	38	14	5	3
Gemiddeld	43	33	15	6	3

kwam voor in de winter van 1962/63. Een opvallend feit is dat de gemiddelde runlengte voor alle stations ongeveer gelijk is en wel ruim vier dagen.

Behalve het gemiddelde loopt ook de vorm van de kansverdeling van de runlengten van plaats tot plaats weinig uiteen. Dit blijkt uit tabel 4.2 waarin voor een aantal stations in België en Nederland een procentuele verdeling van de runlengten is gegeven. De gegevens van de Belgische stations in deze tabel zijn ontleend aan De Vogelaere (1976).

Het gemiddelde aantal sneeuwdekdagen in fig. 2.1 is het produkt van de gemiddelde runlengte en het gemiddelde aantal runs per jaar. Alleen de laatste grootte varieert van plaats tot plaats en wel van ongeveer 3 in Vlissingen tot bijna 7 in Eelde.

#### 4.3 De kansverdeling van de hoogte van het sneeuwdek

Voor dagen waarop het synoptische codecijfer  $S$  in tabel 1.1 van nul verschilt kunnen we een kansvariabele  $H$  definiëren die de hoogte van het sneeuwdek aangeeft. In tabel 4.3 wordt het gemiddelde aantal dagen per jaar gegeven dat  $H$  een waarde van 2, 5, 10 of 15 cm bereikt of overschrijdt. Deze niveaus zijn steeds de ondergrens van een bepaald codecijfer  $S$ . Evenals bij de frequentieverdeling van de jaarmaxima van het codecijfer  $S$  kunnen we op de gegevens in tabel 4.3 een continue functie aanpassen. Van zo'n aangepaste functie wordt in dit hoofdstuk gebruik gemaakt om de gemiddelde hoogte van het sneeuwdek te berekenen.

Om een geschikte kansverdeling voor  $H$  te vinden beschouwen we eerst een gemiddelde frequentieverdeling van sneeuwdekhogten van een groot aantal Duitse stations nabij de Nederlandse grens en wel stations ten westen van de  $7^{\circ}30'$  meridiaan en ten noorden van de  $50^{\circ}30'$  parallel. Door Caspar (1962) wordt voor verschillende hoogten  $x$  het gemiddelde aantal dagen per jaar  $A_x$  gegeven dat  $H \geq x$ . Van 21 stations met een hoogteligging van 50 m of minder werd voor elke  $x$  de mediaanwaarde van  $A_x$  bepaald. In fig. 4.1 is deze mediaanwaarde van  $A_x$  uitgezet tegen  $x$ . De verticale as in deze figuur heeft een logaritmische schaal, zodat we tussen  $A_x$  en  $x$  de volgende relatie krijgen:

$$A_x = b e^{-\lambda x} \quad (4.1)$$



Tabel 4.3. Gemiddeld aantal ochtenden per jaar dat de bodem bedekt is met een sneeuwdek van een bepaalde hoogte (tijdvak 1955/56-1982/83).

Station(s)	Ochtenden met een sneeuwdek	Ochtenden met een sneeuwdek van tenminste			
		2 cm	5 cm	10 cm	15 cm
Den Helder/De Kooy	17,3	7,1	2,5	0,7	0,11
Eelde	31,1	16,9	10,9	4,7	2,04
De Bilt	21,8	12,9	8,1	3,1	0,86
Vlissingen	12,5	4,3	2,1	0,7	0,07
Beek (L.)	26,3	12,3	6,4	2,7	0,64

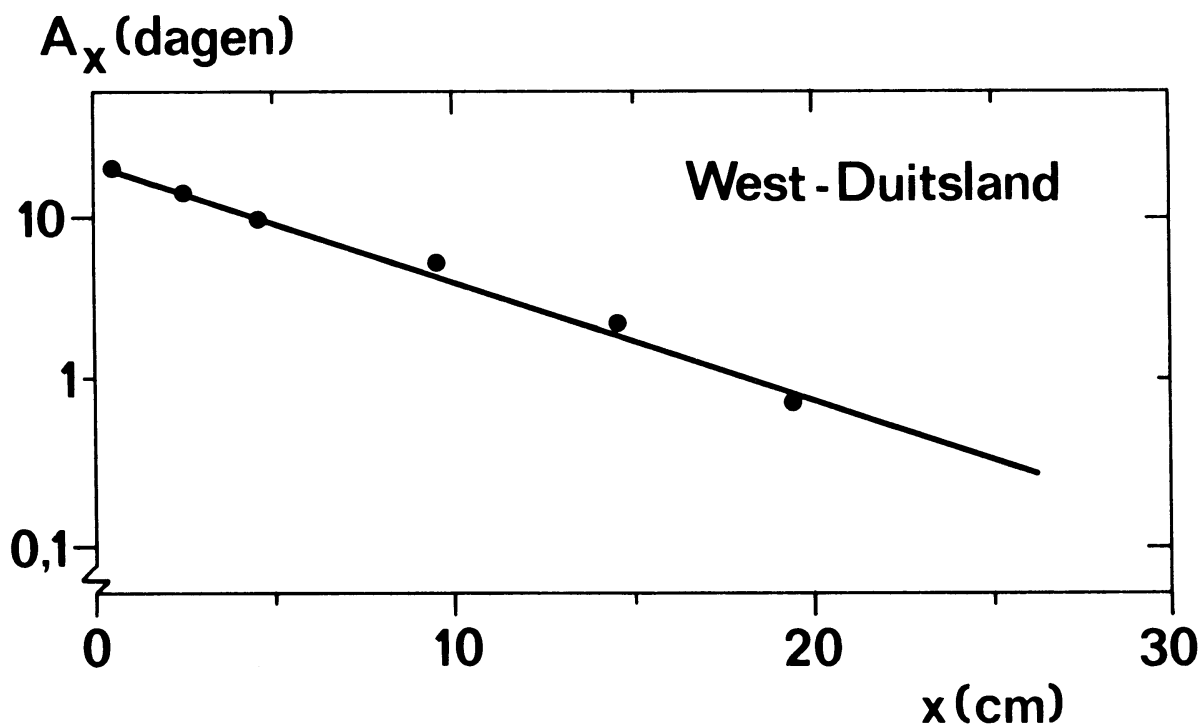


Fig. 4.1. Het gemiddelde aantal dagen  $A_x$  per winterhalfjaar dat het sneeuwdek een hoogte  $x$  bereikt of overschrijdt voor laaggelegen Duitse stations in het grensgebied met Nederland (tijdvak 1936/37-1957/58 met uitzondering van 1944/45 en 1945/46).

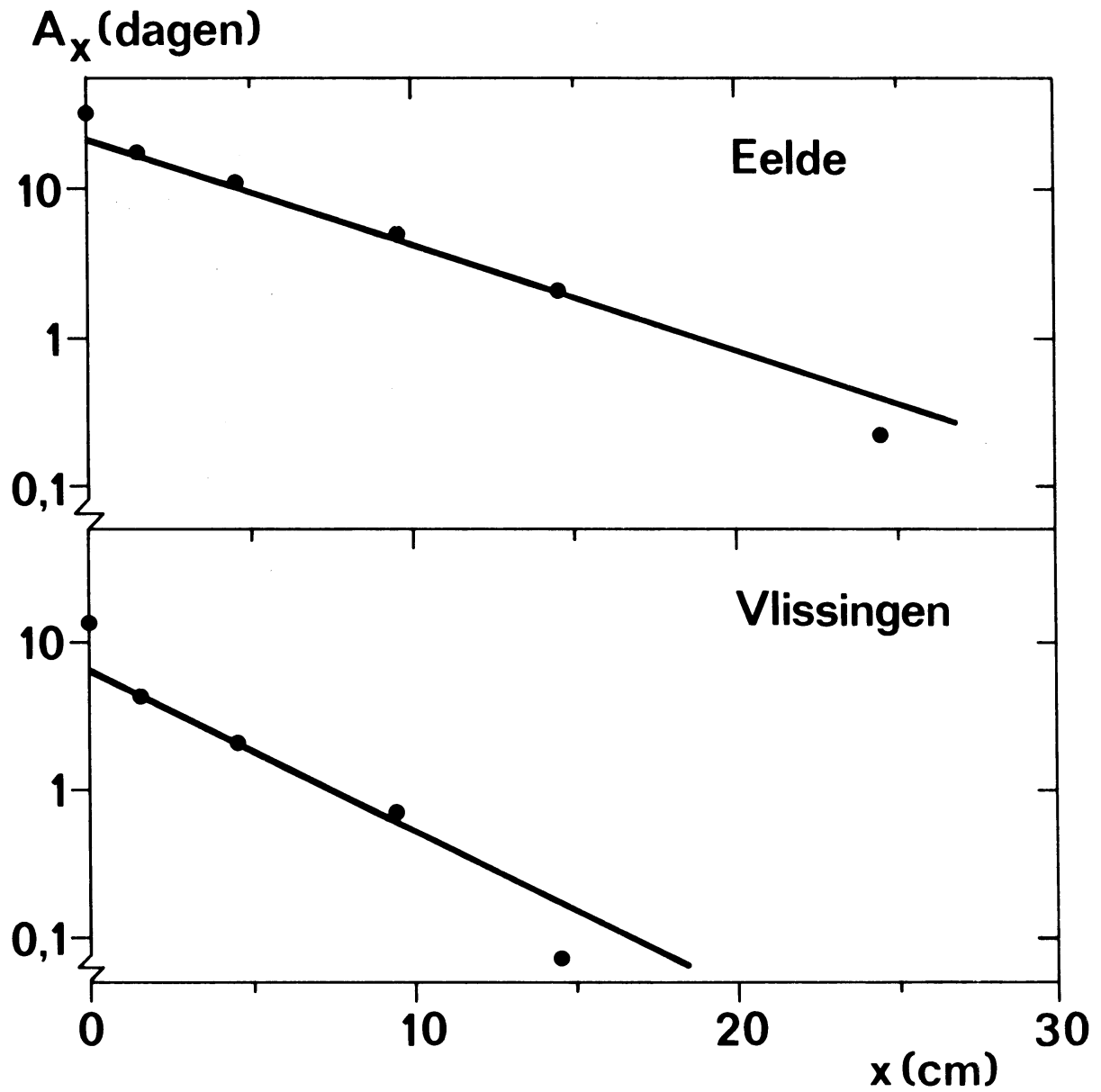


Fig. 4.2. Als fig. 4.1 maar nu voor Eelde en Vlissingen (tijdvak 1955/56-1982/83).

waarbij  $b$  en  $\lambda$  constanten zijn. De grootheid  $A_x$  neemt dus exponentieel af met  $x$ .

Fig. 4.2 geeft voor Eelde en Vlissingen het verband tussen  $A_x$  en  $x$ . Ook in deze figuur heeft de verticale as een logaritmische schaal. De figuur laat zien dat er een discontinuïteit is bij  $x=0$ . Bij de Duitse stations in fig. 4.1 is deze discontinuïteit niet te zien omdat dagen met een sneeuwdekhoogte  $< \frac{1}{2}$  cm buiten beschouwing gelaten zijn daar dit soort dagen in Duitsland vaak niet als sneeuwdekdag wordt opgevat, zie par. 2.4.

Het verloop van  $A_x$  met  $x$  in fig. 4.2 kan beschreven worden door een 2-parameter exponentiële verdeling. De verdelingsfunctie hiervan luidt:

$$F(x) = \Pr(X \leq x) = 1 - C e^{-\lambda x}, \quad x > 0. \quad (4.2)$$

De parameter  $C$  is geïntroduceerd om de discontinuïteit bij  $x=0$  te kunnen beschrijven. Als  $C=1$ , dan is  $F(x)$  continu in de omgeving van  $x=0$ . Is  $C < 1$ , dan heeft  $F(x)$  een discontinuïteit ter grootte van  $1-C$  bij  $x=0$ , zie fig. 4.3. Als  $x=0$  dan is  $F(x)$  altijd gelijk aan nul omdat  $\underline{H}$  per definitie altijd positief is; voor kleine positieve  $x$  neemt  $F(x)$  echter de waarde  $1-C$  aan.

Verg. (4.2) leidt tot de volgende relatie tussen  $A_x$  en  $x$ :

$$A_x = A_0 C e^{-\lambda x}, \quad x > 0 \quad (4.3)$$

waarbij  $A_0$  het gemiddelde aantal dagen per jaar met een sneeuwdek aangeeft. De gemiddelde hoogte  $\mu_H$  van het sneeuwdek volgt uit:

$$\mu_H = C/\lambda. \quad (4.4)$$

De onbekende parameters  $C$  en  $\lambda$  in verg. (4.2) worden zodanig gekozen dat de theoretische verdeling zo goed mogelijk aansluit bij de frequentieverdeling van het codecijfer  $S$  voor alle dagen met een sneeuwdek. Bijzonderheden hierover worden gegeven in appendix F. De parameterschattingen voor de vijf klassieke hoofdstations worden gegeven in tabel 4.4.

### Onderschrijdingskans $F(x)$

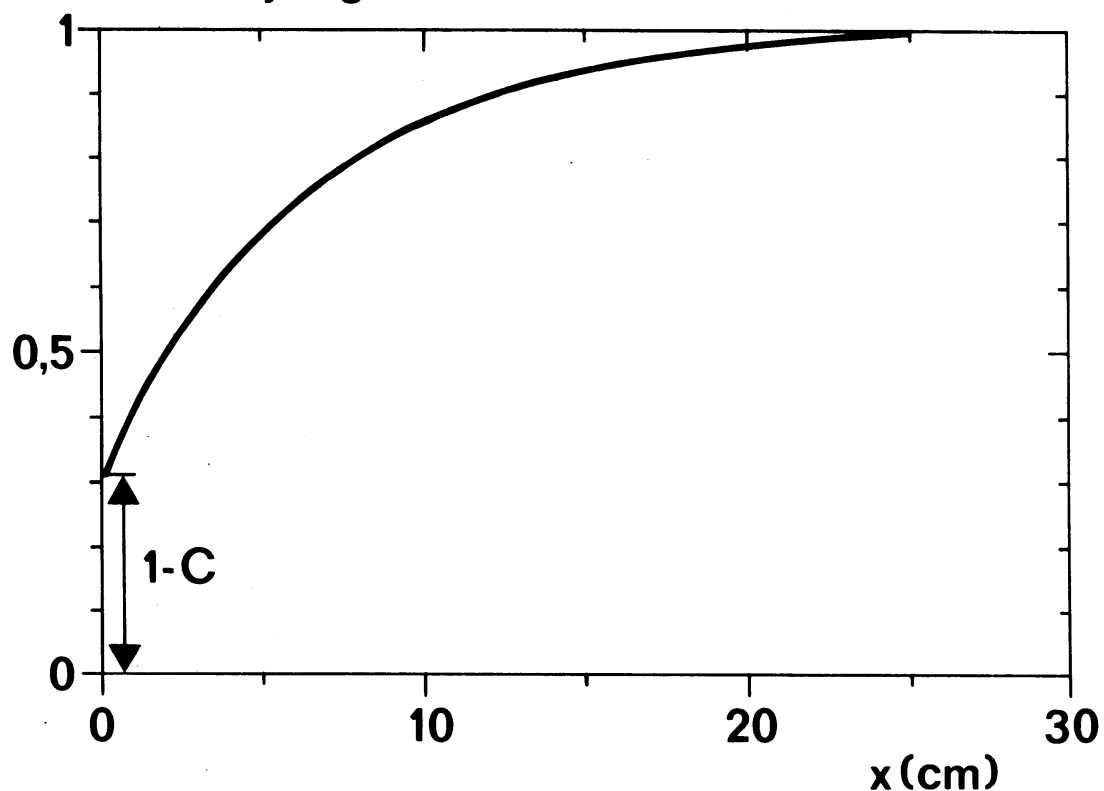


Fig. 4.3 Cumulatieve verdeling van de hoogte van het sneeuwdek in Eelde volgens de exponentiële verdeling gegeven door verg. (4.2).

Tabel 4.4. Schattingen  $\hat{C}$  en  $\hat{\lambda}$  van de parameters  $C$  en  $\lambda$  van de exponentiële verdeling alsmede de geschatte gemiddelde hoogte  $\hat{\mu}_H$  van het sneeuwdek.

Station(s)	$\hat{C}$	$\hat{\lambda}(\text{cm}^{-1})$	$1/\hat{\lambda}(\text{cm})$	$\hat{\mu}_H(\text{cm})$
Den Helder/De Kooy	0,67	0,32	3,1	2,1
Eelde	0,69	0,16	6,1	4,2
De Bilt	0,76	0,19	5,4	4,1
Vlissingen	0,48	0,25	4,0	2,0
Beek (L.)	0,64	0,21	4,9	3,1

De waarde van  $\hat{\lambda}$  is relatief hoog bij kuststations. De geschatte waarde van  $C$  is het hoogste voor De Bilt en het laagste voor Vlissingen. De waarde van deze laatste parameter is sterk afhankelijk van de fractie van het aantal sneeuwdek-dagen met een S-cijfer 1 (dagen met  $H < 2$  cm). Deze fractie wordt zwaar beïnvloed door beoordelingscriteria van waarnemers en mogelijk ook door de lokale situatie van het meetterrein.

Uit tabel 4.3 volgt dat er in De Bilt gemiddeld ongeveer 9 sneeuwdekochtenden zijn met een S-cijfer van 1; in Beek (L.), dat qua sneeuwdekklimaat vergelijkbaar is met De Bilt, zijn er echter gemiddeld 14 van dergelijke sneeuwdekochtenden. Dit verschil in het gemiddelde aantal ochtenden met een S-cijfer 1 leidt tot een vrij groot verschil in het gemiddelde aantal sneeuwdekochtenden tussen beide stations. Doordat Beek (L.) relatief veel ochtenden telt met een sneeuwdek van geringe hoogte is  $\hat{\mu}_H$  van dit station vrij laag. Voor De Bilt wordt daarentegen een vrij hoge waarde voor  $\hat{\mu}_H$  gevonden.

De geschatte hoogte  $\hat{\mu}_H$  voor De Bilt kunnen we verifiëren door deze te vergelijken met de gemiddelde hoogte van de metingen van de reeks van Midden Nederland. Deze reeks bestaat voor een belangrijk deel uit sneeuwdekhogtemetingen te De Bilt. Voor het 28-jarige tijdvak, waarop tabel 4.4 berust, bedraagt de gemiddelde hoogte van de ochtendwaarneming in de reeks van Midden Nederland 4,3 cm. Dit getal is in goede overeenstemming met de 4,1 cm die in de tabel voor De Bilt is aangegeven.

#### 4.4. Runlengten van dagen met een sneeuwdek boven een bepaald niveau

Tabel 4.5 geeft een overzicht van runs van dagen waarop het sneeuwdek een hoogte van minstens 5, 10 of 15 cm heeft bereikt. Voor het hoogste niveau worden alleen de stations in het binnenland beschouwd daar bij kuststations een sneeuwdekhogte boven dit niveau zeer zelden voorkomt. In tegenstelling tot de runs van sneeuwdekdagen loopt bij de runs in tabel 4.5 de gemiddelde lengte sterk van plaats tot plaats uiteen. De langste runs treft men aan bij de stations in het binnenland. Voor alle niveaus is bij Eelde de gemiddelde runlengte groter dan de gemiddelde lengte van de runs van sneeuwdekdagen in tabel 4.1. Daarentegen is bij de kuststations Den Helder/De Kooy en Vlissingen het tegenovergestelde het geval. De gemiddelde runlengte in tabel 4.5 neemt in het algemeen af met de hoogte van het niveau.

Tabel 4.5. Eigenschappen van runs met een sneeuwdek boven een bepaald niveau.  
De aantallen hebben betrekking op de 28 winterhalfjaren 1955/56-  
1982/83.

	Totaal aantal runs	Aantal runs met lengte (dagen)					Gemiddelde	
		1	2-4	5-9	10-19	$\geq 20$	lengte (dagen)	hoogte (cm)
<u>Hoogte <math>\geq 5</math> cm</u>								
Den Helder/De Kooy	21	8	8	4	1	-	3,3	7,6
Eelde	43	12	12	8	8	3	7,1	10,6
De Bilt	44	13	14	11	4	2	5,1	9,9
Vlissingen	16	7	5	2	2	-	3,7	8,5
Beek (L.)	41	13	18	6	3	1	4,4	9,4
<u>Hoogte <math>\geq 10</math> cm</u>								
Den Helder/De Kooy	8	5	2	-	1	-	2,5	12,6
Eelde	22	3	9	4	6	-	5,9	15,6
De Bilt	18	4	10	2	2	-	4,7	14,9
Vlissingen	9	3	6	-	-	-	2,1	13,5
Beek (L.)	20	7	9	3	-	1	3,9	14,4
<u>Hoogte <math>\geq 15</math> cm</u>								
Eelde	11	3	2	4	2	-	5,1	20,6
De Bilt	7	4	2	-	1	-	3,3	19,9
Beek (L.)	4	1	1	1	1	-	4,5	19,4

In de laatste kolom van tabel 4.5 wordt de gemiddelde hoogte van het sneeuwdek voor de betreffende runs gegeven. Bij de bepaling van deze gemiddelde hoogte is gebruik gemaakt van de 2-parameter exponentiële verdeling in verg. (4.2) en wel als volgt.

Voor de sneeuwdekhogten boven een bepaald basisniveau gaan we uit van een voorwaardelijke of conditionele kansverdeling met verdelingsfunctie:

$$F_{H|t}(x,t) = \Pr(\underline{H} \leq x \mid \underline{H} > t), \quad x \geq t. \quad (4.5)$$

Verg. (4.5) geeft de kansverdeling van de sneeuwdekhogten onder de voorwaarde dat het sneeuwdek een hoogte van minstens  $t$  heeft. Tussen deze voorwaardelijke verdelingsfunctie en de onvoorwaardelijke verdelingsfunctie  $F_H(x)$  bestaat het volgende verband:

$$F_{H|t}(x,t) = \frac{F_H(x) - F_H(t)}{1 - F_H(t)}, \quad x \geq t. \quad (4.6)$$

Merk op dat  $F_{H|t}(x,t) = 0$  als  $x=t$ . Substitutie van verg. (4.2) voor  $F_H(x)$  geeft:

$$F_{H|t}(x,t) = 1 - e^{-\lambda(x-t)}, \quad x \geq t. \quad (4.7)$$

De voorwaardelijke verdeling hangt niet af van de parameter  $C$  en wordt dus volledig bepaald door de parameter  $\lambda$ . De verdeling in verg. (4.7) is de klassieke vorm van de exponentiële verdeling (Johnson and Kotz, 1970, par. 18). De gemiddelde hoogte volgt nu uit:

$$\mu_{H|t} = t + 1/\lambda. \quad (4.8)$$

Het gemiddelde krijgt men dus door bij het basisniveau\* een constante  $1/\lambda$  op te tellen. Deze constante is voor kuststations lager dan voor stations in het binnenland (tabel 4.4). De grootste gemiddelde hoogten van het sneeuwdek worden dan ook aan bij stations in het binnenland aangetroffen.

\* De drie niveaus in tabel 4.5 hebben betrekking op meetwaarden die op hele centimeters zijn afgerond. Voor  $t$  moet men dan een waarde aanhouden die 0,5 cm lager is dan het aangegeven niveau in deze tabel.

Runs van dagen met een sneeuwdek boven een bepaald niveau liggen in groepjes bij elkaar. In een run van sneeuwdekdagen zijn er gemiddeld ongeveer 1,2 runs van dagen waarop de niveaus van 5, 10 of 15 cm worden bereikt of overschreden.



## 5. SAMENVATTING

In dit rapport werd een aantal klimatologische eigenschappen van het sneeuwdek in Nederland behandeld. Dit betrof enerzijds de duur dat de grond in Nederland met sneeuw bedekt kan zijn en anderzijds een tweetal kansverdelingen van de hoogte van het sneeuwdek. Als uitgangsmateriaal dienden de gegevens van dertien Nederlandse synoptische stations voor de 28 winterhalfjaren 1955/56, ..., 1982/83 (28 winters).

De gemeten sneeuwdekhogten van deze stations zijn meestal herleid tot een codecijfer, het zogenaamde S-cijfer. Naast reeksen van codecijfers werd ook een beperkt aantal reeksen van werkelijk gemeten hoogten geanalyseerd. Hiervan had één reeks betrekking op metingen in Midden Nederland; de andere reeksen waren afkomstig van stations in België en West-Duitsland.

Het gemiddelde aantal dagen per jaar dat de grond met sneeuw bedekt is loopt uiteen van ongeveer 12 dagen in Zeeland tot ruim 30 dagen in het noordoosten van ons land (fig. 2.1). In ongeveer 30% van de ochtenden met een sneeuwdek is echter de grond voor minder dan de helft met sneeuw bedekt. Voor de in deze studie betrokken stations is er een sterke relatie tussen het gemiddelde aantal dagen met een sneeuwdek en de gemiddelde etmaaltemperatuur  $\bar{w}$  over de maanden december, januari en februari (fig. 2.2). Het gemiddelde aantal dagen per jaar dat de grond met sneeuw bedekt is neemt met ongeveer 10 af als de gemiddelde wintertemperatuur  $\bar{w}$  met 1°C toeneemt. Deze relatie kan ook toegepast worden in laaggelegen delen van België en Duitsland; echter niet in Engeland. In heuvelachtige gebieden (Ardennen) neemt niet alleen het gemiddelde aantal sneeuwdekdagen toe door een afname van de gemiddelde wintertemperatuur met de hoogte, maar is er ook een toename in het gemiddelde aantal sneeuwdekdagen vanwege het orografisch neerslag-effect.

De kansverdeling van het aantal dagen met een sneeuwdek in een winter kan met behulp van een lognormale verdeling beschreven worden (fig. 2.4 t/m 2.6). Daar in de omgeving van  $x=0$  dagen (sneeuwarde winters) de geldigheid van de verdeling zeer discutabel is werden de parameters  $\mu$  en  $\sigma$  geschat op basis van een geschorste steekproef. Evenals bij het gemiddelde aantal sneeuwdekdagen kunnen de waarden van  $\mu$  en  $\sigma$  gerelateerd worden aan de gemiddelde wintertemperatuur  $\bar{w}$  (fig. 2.7). Met behulp van deze relaties kan men voor een willekeurige plaats in Nederland de kansverdeling van het jaarlijks aantal sneeuwdekdagen uit de gemiddelde wintertemperatuur afleiden. Voor de aantallen sneeuwdekdagen, die

gemiddeld eens in de 2, 5, 10 en 20 jaar worden overschreden is dit resultaat in een figuur weergegeven (fig. 2.8).

De kansverdeling van het jaarmaximum van de hoogte van het sneeuwdek kan goed beschreven worden met de Gumbel-verdeling. Een andere kandidaat, de lognormale verdeling, gaf systematisch te hoge waarden in de rechter staart (fig. 3.2 t/m 3.4). Bij reeksen van de gemeten hoogten werden de parameters  $\mu$  en  $\sigma$  van de Gumbel-verdeling geschat met behulp van de methode van de grootste aannemelijkheid op basis van een geschorste steekproef van de jaarmaxima (vanwege problemen in sneeuwarme winters); bij reeksen van synoptische codecijfers werd de methode van de grootste aannemelijkheid toegepast op een frequentieverdeling van de jaarmaxima van het S-cijfer. Evenals bij het jaarlijks aantal sneeuwdekdagen konden de parameters van de verdeling gerelateerd worden aan de gemiddelde wintertemperatuur  $\bar{w}$  (fig. 3.5). Hierdoor is het weer mogelijk voor een willekeurige plaats in Nederland de verdeling van het jaarmaximum van de hoogte van het sneeuwdek af te leiden uit de gemiddelde wintertemperatuur  $\bar{w}$ . Kaartjes worden gegeven van de hoogten die gemiddeld eens in de 10 resp. 50 jaar worden overschreden (fig. 3.6 en 3.7).

De grote jaarlijkse variabiliteit van de hoogte van het sneeuwdek resulteert in vrij grote standaardafwijkingen van de parameterschattingen. Voor de hoogte die gemiddeld eens in de 50 jaar wordt overschreden, is er nog een extra onzekerheid omdat de vorm van de rechter staart van de verdeling niet volledig bekend is. Uit een vergelijking met de waarden van een drietal langjarige reeksen bleek dat het tijdvak 1955/56-1982/83 noch extreem sneeuwarm, noch extreem sneeuwrijk is. Voorts bleek dat bij het schatten van de parameters van de Gumbel-verdeling er slechts een gering verlies aan informatie optrad bij overgang van gemeten waarden naar synoptische codecijfers.

De kansverdeling van runlengten van sneeuwdekdagen varieert nauwelijks van plaats tot plaats (tabel 4.2). Overal in Nederland bedraagt de gemiddelde runlengte ruim vier dagen (tabel 4.1). In de uitzonderlijke winter van 1962/63 was overal in Nederland de grond met sneeuw bedekt gedurende een aaneengesloten periode van ongeveer 70 dagen. Het gemiddelde aantal runs van sneeuwdekdagen loopt uiteen van ongeveer 3 per jaar in Vlissingen tot bijna 7 per jaar in Eelde. De kansverdeling van runlengten van dagen met een sneeuwdek boven een niveau van 5,10 of 15 cm loopt wel sterk van plaats tot plaats uiteen (tabel 4.5). De langste runs worden aangetroffen bij stations in het binnenland.

De verdeling van de hoogte van het sneeuwdek op dagen dat de grond met sneeuw bedekt is, kan goed beschreven worden met een exponentiële verdeling. De hier gekozen exponentiële verdeling bevat een extra parameter vanwege een discontinuïteit in de verdeling van sneeuwdekhogtes bij een hoogte van 0 cm (fig. 4.2 en 4.3). Met behulp van deze exponentiële verdeling kon worden afgeleid dat de gemiddelde hoogte van het sneeuwdek ongeveer 2 cm is langs de kust en ruim 4 cm in het binnenland.

## DANKWOORD

In de eerste fase van het onderzoek werd medewerking verleend door de heer A. Denkema bij het bijeenbrengen van de gegevens van de stations in West-Duitsland en bij het uitvoeren van enkele bewerkingen. De heer P. van der Veer nam het programmeerwerk voor de tabellen in hoofdstuk 4 voor zijn rekening. Tevens droeg hij bij aan het opzoeken van sneeuwdekgegevens uit de jaarboeken van de "Deutscher Wetterdienst".

De auteur is dank verschuldigd aan drs. A.J. Coops (Instituut voor Meteorologie en Oceanografie, Rijksuniversiteit Utrecht) en dr. R. Sneyers (Koninklijk Meteorologisch Instituut van België) voor het kritisch doorlezen van een eerdere versie van de hoofdstukken 1 t/m 3 en de appendices A t/m E. Dr. Sneyers wordt tevens dank gebracht voor het toesturen van de sneeuwdekgegevens van Ukkel-Brussel. De maximale sneeuwdekhogten van Kleef over 1934/35 en 1938/39-1952/53 werden voor dit onderzoek beschikbaar gesteld door de "Deutscher Wetterdienst".

Het tekenwerk van dit rapport werd vakkundig verzorgd door de heer C. van Stralen. Anja de Bree en Marleen Kaltoven wordt dank gebracht voor het uitstekende typewerk.

## APPENDIX A - HOMOGENITEITSANALYSE MET BETREKKING TOT HET AANTAL SNEEUWDEKDAGEN

Er zijn vele methoden om de homogeniteit van een tijdreeks te toetsen (Buishand, 1982). In deze studie werd gebruik gemaakt van de volgende toetsingsgrootte:

$$U = \frac{1}{N(N+1)} \sum_{k=1}^{N-1} \left\{ \sum_{i=1}^k (v_i - \bar{v}) \right\}^2 / s_v^2 \quad (A1)$$

Hierin staat N voor het aantal waarnemingen,  $\bar{v}$  voor het gemiddelde van de  $v_i$ 's en  $s_v^2$  voor de steekproefvariantie:

$$s_v^2 = \frac{1}{N} \sum_{i=1}^N (v_i - \bar{v})^2 \quad (A2)$$

Hoge waarden van de toetsingsgrootte U wijzen op inhomogeniteiten. Een tabel met kritieke waarden wordt gegeven door Buishand (1984).

Tabel A1 geeft voor de stations in par. 2.3 de steekproefstandaardafwijking  $s_v$  en de waarde van de toetsingsgrootte U. Bij de Duitse stations is de homogeniteitsanalyse uitgevoerd met reeksen van hun officiële sneeuwdekking ("Schneedecke  $\geq 0$  cm") als ook met reeksen waarbij alle dagen met "Schneeflecken" en "Schneereste" als sneeuwdekking zijn opgevat\*. In het laatste geval is er bij de reeksen van Borkum, Diepholz en Kleef zeer sterke statistische evidentie voor inhomogeniteiten. Bij Borkum en Diepholz hangt de inhomogeniteit vermoedelijk samen met het gebruik van andere notaties sinds 1965. Vanaf dat jaartal neemt het gemiddelde aantal sneeuwdekkingen op deze stations sterk toe.

Bij de Nederlandse synoptische stations is er slechts statistische evidentie voor een inhomogeniteit bij de reeks van Soesterberg (zie par. 2.3).

Voor de meeste stations langs de rand van het onderzoeksgebied (Eelde,

\* In het gemiddelde  $g_i$  in verg. (2.1) zijn alleen de dagen met "Schneedecke  $\geq 0$  cm" opgenomen bij het toetsen van de homogeniteit van de reeksen in Nederland en België.

Tabel Al. Toets op homogeniteit van reeksen van het aantal sneeuwdekdagen voor het tijdvak 1955/56-1982/83 (N=28).

Station(s)	$s_v$ (dagen/jr)	U
<u>Nederland, België</u>		
Den Helder/De Kooy	5,6	0,165
Leeuwarden	7,9	200
Eelde	7,0	052
Twente	5,2	244
Deelen	5,8	125
Valkenburg (Z.H.)	5,6	255
De Bilt	3,1	056
Soesterberg	5,2	568*
Vlissingen	6,6	125
Gilze-Rijen	4,1	034
Volkel	3,7	388
Eindhoven	4,0	309
Beek (L.)	4,0	046
Ukkel-Brussel	4,6	243
<u>West-Duitsland ("Schneedecke <math>\geq</math> 0 cm")</u>		
Borkum	6,6	0,070
Diepholz	7,5	156
Kleef	7,0	285
Keulen	6,7	271
<u>West-Duitsland (inclusief "Schneeflecken" en "Schneereste")</u>		
Borkum	12,4	0,727**
Diepholz	8,4	551*
Kleef	4,2	181
Keulen	6,2	917**

\* Significant bij een onbetrouwbaarheidsdrempel van 5%, maar niet bij een onbetrouwbaarheidsdrempel van 1%.

\*\* Significant bij een onbetrouwbaarheidsdrempel van 1%.

Leeuwarden, stations West-Duitsland) is de steekproefstandaardafwijking  $s_v$  relatief hoog. Dit komt doordat voor deze stations de correlatie tussen  $x_i$  en  $g_i$  in het algemeen wat lager is dan bij stations die meer in het midden gelegen zijn. Een hoge standaardafwijking heeft tot gevolg dat de kans vrij gering is dat een inhomogeniteit opgespoord wordt.

Vervanging van het gemiddelde  $g_i$  door een gemiddelde van stations die niet meer dan 150 km van het beschouwde station verwijderd liggen levert in een aantal gevallen een verschillenreeks op met een veel kleinere standaardafwijking. Toepassing van de toetsingsgrootte  $U$  op dit soort verschillenreeksen bracht echter geen nieuwe inhomogeniteiten aan het licht.

Behalve op verschillenreeksen van de oorspronkelijke aantallen sneeuwdekdagen  $x_i$  werd de toets ook toegepast op verschillenreeksen van getransformeerde aantallen  $z_i = \ln(x_i+5)$ . In tegenstelling tot de  $x_i$ 's zijn de  $z_i$ 's voor de meeste stations bij benadering normaal verdeeld (par. 2.5), waardoor beter voldaan wordt aan de veronderstellingen die aan de toets ten grondslag liggen. Het toetsen op verschillenreeksen van getransformeerde aantallen gaf vrijwel dezelfde resultaten als in tabel A1.

De overgang van het station Den Helder naar De Kooy in augustus 1972 heeft geleid tot een relatieve toename van bijna twee sneeuwdekkochtenden per winterhalfjaar ten opzichte van de overige stations. Deze toename is echter statistisch niet significant. De gegevens over de drie winterhalfjaren (1955/56 t/m 1957/58) dat de waarnemingen van het station Vlissingen in Souburg werden verricht, laten geen afwijkingen zien ten opzichte van de andere sneeuwdekgegevens in de reeks van Vlissingen.

APPENDIX B - HET SCHATTEN VAN DE PARAMETERS VAN EEN KANSVERDELING BIJ  
GESCHORSTE STEEKPROEVEN

In de paragrafen 2.5 en 3.2 kwam het schatten van de parameters van een kansverdeling op basis van een geschorste steekproef ter sprake. Het betrof hier steekproeven die aan de linkerzijde geschorst waren (Eng.: "left censoring"). Dit houdt in dat van waarnemingen onder een bepaalde drempelwaarde  $t$  de werkelijke waarde buiten beschouwing gelaten wordt; wel wordt het aantal  $m$  van deze geschorste waarnemingen bijgehouden. De reden dat gewerkt werd met geschorste steekproeven was dat getwijfeld werd aan de geldigheid van de verdeling voor sneeuwarme winters. Daarnaast kan het schatten van de parameters op basis van een geschorste steekproef noodzakelijk zijn doordat de relatief grote onnauwkeurigheid van lage waarden de kwaliteit van de schattingen zeer ongunstig beïnvloedt.

In deze appendix zal het schatten van de parameters  $\mu$  en  $\sigma$  van zowel de lognormale verdeling als de Gumbel-verdeling worden besproken. Bij de lognormale verdeling zal de parameter  $a$  bekend verondersteld worden. Het vinden van een geschikte waarde voor deze parameter zal in appendix D worden beschreven.

Schattingen voor  $\mu$  en  $\sigma$  kunnen in principe met behulp van waarschijnlijkheidspapier worden verkregen. De parameter  $\mu$  bepaalt het niveau van de lijn, terwijl  $\sigma$  de helling weergeeft. Bij het aanpassen van een rechte lijn moet men de geschorste waarnemingen buiten beschouwing laten. Men moet echter wel met deze waarnemingen rekening houden bij de bepaling van de plotkansen  $F_i$ . Objectief kan men  $\hat{\mu}$  en  $\hat{\sigma}$  bepalen met behulp van de kleinste kwadratenmethode. Hoewel dit bij de normale en lognormale verdeling een goede methode is, heeft zij bij de Gumbel-verdeling het bezwaar dat te veel gewicht toegekend wordt aan de grootste waarneming. Het gevolg hiervan is dat men met name een vrij onnauwkeurige schatting voor de parameter  $\sigma$  krijgt, wat er weer toe leidt dat de geschatte waarde van een grootheid als  $M_{50}$  sterk uiteen kan lopen tussen stations met vrijwel hetzelfde sneeuwklimaat. Dit kan worden geïllustreerd aan de hand van de jaarmaxima van de stations Diepholz en Kleef in de figuren 3.2 en 3.3. Volgens de Gumbel-rechte in deze figuren bedragen de waarden voor  $\hat{M}_{50}$  resp. 31 en 33 cm. De iets hogere waarde voor het station Kleef moet hoofdzakelijk worden toegeschreven aan het effect van de twee grootste waarnemingen



in de steekproef. Klimatologische verschillen en het feit dat bij Diepholz met een iets langere waarnemingsreeks wordt gewerkt zijn van minder belang. Een rechte lijn volgens de kleinste kwadratenmethode geeft voor  $\hat{M}_{50}$  een waarde van 30 cm bij Diepholz en 36 cm bij Kleef. Het verschil tussen deze stations is dan driemaal zo groot als eerder vermeld werd.

De lijnen in de figuren 2.6, 3.2, 3.3 en 3.4 zijn verkregen met behulp van de methode van de grootste aannemelijkheid (Eng.: "maximum likelihood"). Hierbij worden de parameters  $\mu$  en  $\sigma$  zodanig gekozen dat de aannemelijkheidsfunctie gemaximaliseerd wordt. Deze functie kan bij een geschorste steekproef als volgt worden gedefinieerd:

$$L(\mu, \sigma; x_1, x_2, \dots, x_n) = F^m(t) \prod_{i=1}^n f(x_i) \quad (B1)$$

met  $m$  het aantal waarnemingen dat de drempelwaarde  $t$  niet overschrijdt; de  $x_i$ 's geven de waarnemingen boven de drempelwaarde aan.

Het maximaliseren van  $L$  naar  $\mu$  en  $\sigma$  is gelijkwaardig aan het maximaliseren van  $\ln L$  (de "log likelihood") naar  $\mu$  en  $\sigma$ . Voor  $\ln L$  kunnen we schrijven:

$$\ln L = m \ln F(t) + \sum_{i=1}^n \ln f(x_i). \quad (B2)$$

Eerst zullen we het maximaliseren van  $\ln L$  bespreken voor de lognormale verdeling; daarna komt de Gumbel-verdeling aan de orde.

#### B1. Toepassing van de methode van de grootste aannemelijkheid bij de lognormale verdeling

Bij de lognormale verdeling kunnen we overgaan van de aannemelijkheidsfunctie van de  $x_i$ 's naar de aannemelijkheidsfunctie  $L^*$  van de  $z_i$ 's waarbij  $z_i = \ln(x_i - a)$ . Uit (2.3) en (2.5) kan het volgende verband tussen  $\ln L^*$  en  $\ln L$  worden afgeleid:

$$\ln L = \ln L^* - \sum_{i=1}^n z_i. \quad (B3)$$

Daar de laatste term in deze vergelijking niet van  $\mu$  en  $\sigma$  afhangt bereiken  $\ln L$  en  $\ln L^*$  hun maximum voor dezelfde waarden van  $\mu$  en  $\sigma$ .

De  $z_i$ 's vormen een geschorste steekproef uit een normale verdeling waarbij de drempelwaarde gelijk is aan  $g = \ln(t-a)$ . De kans dat deze drempelwaarde niet overschreden wordt is dan:

$$F_z(g) = \Pr(\underline{z} \leq g) = \Phi(y_g) \quad (\text{B4})$$

met  $y_g = (g-\mu)/\sigma$  en  $\Phi(\cdot)$  de verdelingsfunctie van een normale variabele waarvoor  $\mu=0$  en  $\sigma=1$  (de standaardnormale variabele).

Voor  $\ln L^*$  geldt volgens verg. (B2):

$$\begin{aligned} \ln L^* &= m \ln F_z(g) + \sum_{i=1}^n \ln f_z(z_i) \\ &= m \ln \Phi(y_g) - n \ln \sqrt{2\pi} - n \ln \sigma - \frac{1}{2} \sum_{i=1}^n \frac{(z_i - \mu)^2}{\sigma^2}. \end{aligned} \quad (\text{B5})$$

Differentiëren naar  $\mu$  en  $\sigma$  geeft:

$$\frac{\partial \ln L^*}{\partial \mu} = -\frac{m}{\sigma} \frac{\phi(y_g)}{\Phi(y_g)} + \frac{1}{\sigma^2} \sum_{i=1}^n (z_i - \mu) \quad (\text{B6})$$

$$\frac{\partial \ln L^*}{\partial \sigma} = -\frac{m y_g}{\sigma} \frac{\phi(y_g)}{\Phi(y_g)} - \frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (z_i - \mu)^2 \quad (\text{B7})$$

waarbij  $\phi(\cdot)$  de kansdichtheid is van een standaardnormale variabele:

$$\phi(y) = \phi'(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} y^2} \quad (\text{B8})$$

De waarden van  $\mu$  en  $\sigma$  waarvoor  $\ln L^*$  maximaal is krijgt men door de partiële afgeleiden in (B6) en (B7) gelijk aan nul te stellen. Dit leidt tot de volgende vergelijkingen:

$$\bar{z} - \mu = \sigma B(y_g) \quad (\text{B9})$$

$$s^2 + (\bar{z} - \mu)^2 = \sigma^2 \{1 + y_g B(y_g)\} \quad (\text{B10})$$

waarbij

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i; \quad s^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2; \quad B(y_g) = \frac{m}{n} \frac{\phi(y_g)}{\Phi(y_g)}. \quad (\text{B11})$$

De oplossing van het stelsel (B9), (B10) wordt behandeld door Cohen (1959, 1961) en David (1981, par. 6.3). Eerst lost men  $y_g$  op uit de vergelijking:

$$\frac{1 - B(y_g) \{B(y_g) - y_g\}}{\{B(y_g) - y_g\}^2} = \frac{s^2}{(\bar{z} - g)^2}. \quad (\text{B12})$$

Dit is een niet-lineaire vergelijking, waarvan de oplossing door iteratie moet worden bepaald (bijv. de Newton-Raphson methode). Een geschikte beginwaarde is  $\Phi^{-1}(m/N)$  met  $N = m + n$ .

Heeft men de oplossing  $\hat{y}_g$  van (B12) bepaald, dan volgen de "maximum  
likeliheidschattingen"  $\hat{\mu}$  en  $\hat{\sigma}$  uit:

$$\hat{\mu} = \bar{z} - \lambda (\bar{z} - g) \quad (\text{B13})$$

$$\hat{\sigma}^2 = s^2 + \lambda (\bar{z} - g)^2 \quad (\text{B14})$$

waarbij:

$$\lambda = B(\hat{y}_g) / \{B(\hat{y}_g) - \hat{y}_g\}. \quad (\text{B15})$$

Door Cohen (1961) en David (1981, pag. 142-143) wordt een tabel gegeven van  $\lambda$  als functie van  $m/N$  en  $\gamma = s^2 / (\bar{z} - g)^2$ . Deze tabel kan men gebruiken als alternatief voor numerieke oplossing van (B12).

Als alle waarnemingen boven de drempelwaarde liggen ( $m=0$ ), dan zijn  $\lambda$  en  $B(y_g)$  gelijk aan nul en krijgt men  $\bar{z}$  en  $s$  als schattingen voor  $\mu$  en  $\sigma$ .

B2. Toepassing van de methode van de grootste aannemelijkheid bij de Gumbel-verdeling

De methode van de grootste aannemelijkheid bij de Gumbel-verdeling is analoog aan die bij de Weibull-verdeling. Dit komt doordat de negatieve logaritme van een Weibull-variabele een Gumbel-verdeling heeft, waardoor tussen de aannemelijkheidsfuncties van beide kansvariabelen een soortgelijk verband bestaat als (B3). De methode van de grootste aannemelijkheid bij de Weibull-verdeling wordt behandeld door Cohen (1965), zowel voor volledige als voor geschorste steekproeven.

Uit substitutie van (3.1) en (3.2) in (B2) volgt:

$$\ln L = -m e^{-(t-\mu)/\sigma} - n \ln \sigma - \sum_{i=1}^n (y_i + e^{-y_i}) \quad (\text{B16})$$

met  $y_i = (x_i - \mu)/\sigma$ . Differentiëren naar  $\mu$  en  $\sigma$  geeft:

$$\frac{\partial \ln L}{\partial \mu} = -\frac{m}{\sigma} e^{-(t-\mu)/\sigma} + \frac{1}{\sigma} \sum_{i=1}^n (1 - e^{-y_i}) \quad (\text{B17})$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{m}{\sigma^2} (t-\mu) e^{-(t-\mu)/\sigma} - \frac{n}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^n (y_i - y_i e^{-y_i}). \quad (\text{B18})$$

Uit  $\partial \ln L / \partial \mu = 0$  volgt:

$$\mu = \sigma \ln(n/S_0) \quad (\text{B19})$$

waarbij  $S_0$  wordt gegeven door:

$$S_0 = m e^{-t/\sigma} + \sum_{i=1}^n e^{-x_i/\sigma}. \quad (\text{B20})$$

Substitutie van (B20) in (B18) geeft na enig rekenwerk:

$$\frac{\partial \ln L}{\partial \sigma} = -n(\sigma - \bar{x} + S_1/S_0)/\sigma^2 = 0 \quad (\text{B21})$$

waarin

$$S_1 = mt e^{-t/\sigma} + \sum_{i=1}^n (x_i e^{-x_i/\sigma}) \quad (\text{B22})$$

Vergelijking (B21) kan worden herschreven als:

$$\sigma - \bar{x} + S_1/S_0 = 0. \quad (\text{B23})$$

Uit deze vergelijking kan  $\sigma$  door iteratie worden opgelost. Substitutie van de verkregen oplossing  $\hat{\sigma}$  in het rechterlid van (B19) geeft de "maximum likelihood" schatting  $\hat{\mu}$ .

De "maximum likelihood" vergelijkingen (B19) en (B23) zijn identiek aan die bij volledige steekproeven. Het enige verschil is dat bij geschorste steekproeven de werkelijke waarden van de geschorste waarnemingen in de sommaties  $S_0$  en  $S_1$  vervangen worden door de drempelwaarde  $t$ .

In par. 3.4 werden standaardafwijkingen van  $\hat{\mu}$ ,  $\hat{\sigma}$ ,  $\hat{M}_{10}$  en  $\hat{M}_{50}$  gegeven. De getallen in tabel 3.2 werden verkregen door met behulp van een computer 4000 reeksen van 28 onafhankelijke Gumbel-variabelen te genereren. De gegeven standaardafwijkingen zijn de steekproefstandaardafwijkingen van de "maximum likelihood" schattingen van deze gegenereerde reeksen. Daarnaast kan men uit de verwachtingswaarden van de tweede afgeleiden van  $\ln L$  asymptotische uitdrukkingen voor de varianties krijgen. Deze uitdrukkingen zijn te vinden in een artikel van Leese (1973). Als alternatief kan men gebruik maken van een tabel in Meeker en Nelson (1977) voor de asymptotische covariantiematrix van "maximum likelihood" schattingen. Deze asymptotische formules geven de beste resultaten voor stations in het binnenland. Voor Eelde bedragen de verschillen met de gegeven standaardafwijkingen in tabel 3.2 ongeveer 1%; bij Vlissingen zijn de verschillen met de getallen in deze tabel echter bijna 10%.

## APPENDIX C - AANPASSINGSTOETSEN

In de statistiek zijn verschillende toetsen ontwikkeld om te beoordelen of een kansverdeling bij de waarnemingen past. Een vrij populaire toets is die van Kolmogorov-Smirnov. Voor een geschorste steekproef wordt de toetsingsgrootheid gegeven door:

$$D = \max_{x > t} | F(x) - F_N(x) | \quad (C1)$$

waarbij  $t$  de drempelwaarde aangeeft. De grootheid  $D$  geeft het maximale verschil aan tussen de theoretische verdelingsfunctie  $F(x)$  en de empirische verdelingsfunctie  $F_N(x)$ . In fig. 2.5 is  $D$  de grootste verticale afstand tussen de vloeiende lijn en de trapjeslijn.

Bij hoge waarden van  $D$  moeten we  $F(x)$  verwerpen. Men mag hiervoor echter niet een klassieke tabel van kritieke waarden voor de Kolmogorov-Smirnov toets gebruiken, om de volgende twee redenen:

- (i) Bij de oorspronkelijke Kolmogorov-Smirnov toets moeten de parameters  $\mu$  en  $\sigma$  van te voren gespecificeerd zijn. In deze studie zijn  $\mu$  en  $\sigma$  echter geschat op grond van de waarnemingen. Hierdoor zal  $F(x)$  beter aansluiten bij  $F_N(x)$  dan wanneer men  $\mu$  en  $\sigma$  onafhankelijk van de waarnemingen kiest. Men moet daarom bij geschatte parameters lagere kritieke waarden gebruiken dan die bij de oorspronkelijke Kolmogorov-Smirnov toets.
- (ii) Bij de oorspronkelijke Kolmogorov-Smirnov toets wordt uitgegaan van volledige steekproeven. Het maximum wordt dan genomen over alle  $x$ -waarden, terwijl bij een geschorste steekproef slechts gemaximaliseerd wordt over de  $x$ -waarden boven het drempelniveau  $t$ , wat in een lagere waarde voor het maximum kan resulteren. Men moet daarom bij een geschorste steekproef lagere kritieke waarden aanhouden dan bij een volledige steekproef, zie ook Barr and Davidson (1973). De hoogte van de kritieke waarde neemt af met de fractie geschorste waarnemingen  $h$ .

Een derde punt is nog dat de kritieke waarden bij een Kolmogorov-Smirnov toets slechts gelden voor continue kansvariabelen. Daar sneeuwdekhoogten op hele centimeters worden afgerond hebben we in feite te maken met een discrete

variabele. In tegenstelling tot de twee eerder genoemde punten is het effect van afrondingen te verwaarlozen.

In tabel C1 worden kritieke waarden van de grootheid  $\underline{D}/N$  gegeven. De getallen in deze tabel gelden voor een toets op een normale verdeling, waarbij de parameters  $\mu$  en  $\sigma$  geschat zijn met de methode van de grootste aannemelijkheid (appendix B1). Voor volledige steekproeven ( $h=0$ ) zijn de waarden in tabel C1 iets hoger dan die in Pearson and Hartley (1976, tabel 54), hoofdzakelijk vanwege het feit dat in Pearson and Hartley de parameter  $\sigma$  op een iets andere wijze wordt geschat.

De getallen in tabel C1 werden verkregen met behulp van een Monte-Carlo methode. Voor elke waarde van  $N$  werden 4000 onafhankelijke steekproeven uit een standaardnormale verdeling gegenereerd; de drempelwaarde  $t$  werd steeds gelijk gesteld aan het 100  $h$ -procentpunt van de verdeling.

Voor volledige steekproeven bedragen de kritieke waarden in tabel C1 ongeveer 65% van die van de Kolmogorov-Smirnov toets waarbij de parameters  $\mu$  en  $\sigma$  van tevoren gespecificeerd zijn; voor geschorste steekproeven zijn de relatieve verschillen zelfs nog iets groter.

De getallen in tabel C1 zijn ook geldig voor de 2-parameter lognormale verdeling, mits de parameters geschat worden met de methode van de grootste aannemelijkheid.

Procentpunten van  $\underline{D}/N$  voor volledige steekproeven van een Gumbel-variabele zijn gepubliceerd door Chandra et al. (1981). Deze procentpunten zijn iets lager dan die voor de normale verdeling in tabel C1. Uit een Monte Carlo studie bleek dat ook voor geschorste steekproeven de kritieke waarden voor een toets op een Gumbel-verdeling weinig verschillen van die voor een toets op een normale verdeling. We kunnen daarom tabel C1 ook gebruiken om te toetsen of de waarnemingen uit een Gumbel-verdeling afkomstig zijn. Voorwaarde is echter wel dat de parameters geschat worden met de methode van de grootste aannemelijkheid.

De Kolmogorov-Smirnov toets werd toegepast om te zien of de verdeling van de jaarmaxima van de sneeuwdekhoogten van de stations Diepholz, Kleef en Ukkel-Brussel (fig. 3.2, 3.3 en 3.4) beschreven kan worden door de Gumbel-verdeling en de 2-parameter lognormale verdeling. De waarden van de toetsingsgrootheid zijn gegeven in tabel C2. De lengte van de waarnemingsreeks is bij Diepholz en Kleef ongeveer 50 jaar en bij Ukkel-Brussel bijna 100 jaar. Bij alle drie

Tabel C1. Procentpunten van de grootheid  $D/N$  (Kolmogorov-Smirnov) voor een toets op een normale of lognormale verdeling. Het p- procentpunt geeft de waarde aan die met een kans van p-procent niet overschreden wordt.

		Fractie geschorste waarnemingen h			
		0	0,25	0,50	0,75
90%	10	0,78	0,72	0,61	0,47
	20	80	73	63	47
	50	82	75	64	47
95%	10	0,84	0,79	0,68	0,52
	20	87	80	69	53
	50	88	82	71	52
99%	10	0,96	0,91	0,81	0,65
	20	1,02	96	79	64
	50	1,03	96	85	63

Tabel C2. Waarden van de toetsingsgrootheid van Kolmogorov-Smirnov voor de jaarmaxima van de sneeuwdekhogten in de figuren 3.2, 3.3 en 3.4.

Station	Lognormaal		Gumbel	
	D	D/N	D	D/N
Diepholz	0,097	0,74	0,086	0,66
Kleef	0,064	0,44	0,069	0,47
Ukkel-Brussel	0,060	0,59	0,054	0,53



stations zijn bij het schatten van de parameters ongeveer 25% van de waarnemingen geschorst. Alle getallen in tabel C2 zijn lager dan het 90%-punt van  $\underline{D}/N$  bij  $h=0,25$  zodat de toets geen aanleiding geeft om zowel de lognormale als de Gumbel-verdeling te verwerpen.

Hoewel de lognormale verdeling en de Gumbel-verdeling grote verschillen in de rechter staart te zien gaven, zegt de Kolmogorov-Smirnov toets dat beide verdelingen aanvaardbaar zijn. Dit wordt veroorzaakt door het feit dat in de staarten van de verdeling de verschillen in de onderschrijdingskansen gering zijn. Bij de jaarmaxima van de sneeuwdekhogten in Kleef (fig. 3.3) geldt voor  $x = 40$  cm:

$$F_N(x) = 0,979$$

$$F(x) = 0,976 \text{ bij de lognormale verdeling}$$

$$F(x) = 0,993 \text{ bij de Gumbel-verdeling.}$$

De verschillen in onderschrijdingskans bedragen derhalve nog geen 2%, terwijl de toetsingsgrootte  $D$  pas significant wordt bij verschillen in de orde van 10%. De Kolmogorov-Smirnov toets is alleen maar in staat verschillen tussen  $F(x)$  en  $F_N(x)$  in het midden van de verdeling te detecteren; voor afwijkingen in de staarten is de toets ongevoelig. Hetzelfde geldt ook voor de chi-kwadrat toets. Deze toets is ongevoelig voor afwijkingen in de rechter staart doordat de hoogste waarnemingen in één klasse worden samengevoegd.

Om te beoordelen of een bepaalde verdeling bij de waarnemingen past kan men ook wat meer gebruik maken van de ligging van de geplote punten van de geordende steekproef op het betreffende waarschijnlijkheidspapier. Als maat voor de rechtlijnigheid kan men de correlatiecoëfficiënt  $r$  gebruiken. Indien de verdeling bruikbaar is dan moet deze plot-correlatiecoëfficiënt vrijwel gelijk aan 1 zijn; lagere waarden van  $r$  wijzen op afwijkingen.

Bij jaarmaxima van de waterinhoud van het sneeuwdek bevelen Ellingwood and Redfield (1983, 1984) het gebruik van de plot-correlatiecoëfficiënt aan om te kiezen tussen verschillende mogelijke verdelingen. Voor de normale verdeling levert de plot-correlatiecoëfficiënt een goede toets op, zie Filliben (1975). Deze auteur geeft ook kritieke waarden van de toetsingsgrootte voor volledige steekproeven. Bij geschorste steekproeven moet men lagere kritieke waarden aanhouden, omdat het weglaten van waarnemingen onder een bepaald niveau verlagend werkt op de hoogte van de plot-correlatiecoëfficiënt.

Tabel C3 geeft kritieke waarden van de plot-correlatiecoëfficiënt  $r$  als functie van de steekproefomvang  $N$  en de fractie geschorste waarnemingen  $h$ . De tabel werd verkregen door middel van een Monte Carlo methode (5000 onafhankelijke reeksen bij  $N=10$  en  $N=20$ ; 2000 reeksen bij  $N=50$  en  $N=100$ ). De waarde van een bepaald procentpunt neemt in het algemeen af met  $h$ . Een uitzondering hierop treft men aan bij het 10%-punt en het 50%-punt voor  $h=0,75$  en  $N=10$ . Bij deze combinatie van  $h$  en  $N$  komt het vaak voor dat  $r$  gelijk is aan 1 doordat er slechts twee waarnemingen boven de drempelwaarde liggen.

Voor de logaritmen van de jaarmaxima van de reeksen in par. 3.2 op normaal waarschijnlijkheidspapier wordt gevonden:

$r = 0,963$  bij Diepholz,

$r = 0,993$  bij Kleef,

$r = 0,983$  bij Ukkel-Brussel.

De waarde van  $r$  voor Diepholz is significant bij een onbetrouwbaarheidsdrempel van 5%. Dit wordt veroorzaakt door afwijkingen in de rechter staart, zie fig. 3.2.

Tabel C3. Procentpunten van de plot-correlatiecoëfficiënt  $r$  bij een normale verdeling. Het p-procentpunt geeft de waarde aan die met een kans van p procent niet overschreden wordt.

		Fractie geschorste waarnemingen h			
N		0	0,25	0,50	0,75
1%	10	0,879	0,852	0,815	0,807
	20	923	902	865	804
	50	968	955	930	879
	100	982	973	958	923
5%	10	0,918	0,895	0,878	0,875
	20	948	937	909	877
	50	977	970	953	921
	100	987	982	972	949
10%	10	0,933	0,918	0,903	0,913
	20	960	949	928	904
	50	981	977	963	938
	100	989	986	979	960
50%	10	0,970	0,965	0,963	0,990
	20	981	977	970	965
	50	990	989	984	973
	100	994	994	990	983

APPENDIX D - HET SCHATTEN VAN DE PARAMETERS VAN DE LOGNORMALE VERDELING BIJ  
HET AANTAL SNEEUWDEKDAGEN

In par. 2.5 werd de 3-parameter lognormale verdeling geïntroduceerd. De variabele  $x$  wordt lognormaal genoemd indien  $z = \ln(x-a)$  normaal verdeeld is. De parameters  $\mu$  en  $\sigma$  zijn afhankelijk van de ligging van het station. In deze appendix wordt daarom aan deze parameters een stationsindex  $j$  toegevoegd. In par. 2.5 werden twaalf Nederlandse synoptische stations beschouwd. De logaritme van de aannemelijkheidsfunctie van het  $j$ -de station duiden we aan als  $L(\mu_j, \sigma_j, a)$ ,  $j=1, \dots, 12$ . Voor een aantal vaste waarden van  $a$  werd voor ieder station afzonderlijk de aannemelijkheidsfunctie gemaximaliseerd naar  $\mu_j$  en  $\sigma_j$ . De waarden van  $\mu_j$  en  $\sigma_j$  waarvoor het maximum bereikt wordt duiden we aan als  $\hat{\mu}_j$  en  $\hat{\sigma}_j$  (de "maximum likelihood" schattingen). We kijken nu naar de gemiddelde waarde van het maximum:

$$\bar{L}(a) = \frac{1}{12} \sum_{j=1}^{12} L(\hat{\mu}_j, \hat{\sigma}_j, a) \quad (D1)$$

als functie van  $a$ . Uit tabel D1 blijkt dat  $\bar{L}(a)$  maximaal is voor  $a=-5$ . Voor deze waarde van  $a$  is ook de gemiddelde plot-correlatiecoëfficiënt  $\bar{r}(a)$  maximaal. Beide functies veranderen echter slechts zeer geleidelijk in de omgeving van  $a=-5$ .

Een waarde voor  $a$  groter dan  $-5$  leidt er toe dat de rechter staart te lang wordt bij stations in het binnenland. Een kleinere waarde van  $a$  dan  $-5$  kan het grote aantal sneeuwdekochtenden bij kuststations in zeer strenge winters niet verklaren. Bovendien blijkt dat de waarde van de toetsingsgrootte van Kolmogorov-Smirnov dan toeneemt, zie tabel D2. Op grond hiervan is in deze studie voor  $a$  de waarde  $-5$  aangehouden, zulks in tegenstelling tot Sneyers (1967) die de voorkeur gaf aan  $a=-10$ .

Bij geen enkel station is de waarde van de toetsingsgrootte van Kolmogorov-Smirnov in tabel D1 significant bij een onbetrouwbaarheidsdrempel van 5%. Voor  $a=-5$  blijkt dit ook het geval bij de plot-correlatiecoëfficiënt  $r$ . Hieruit kan geconcludeerd worden dat de lognormale verdeling goed bij de waarnemingen past.

In par. 2.5 was gebleken dat schattingen voor  $\mu_j$  en  $\sigma_j$  konden worden verkregen uit de gemiddelde wintertemperatuur  $\bar{w}$ . Deze schattingen zullen we hier aanduiden als  $\tilde{\mu}_j$  en  $\tilde{\sigma}_j$ . De overgang van de "maximum likelihood" schattingen  $\hat{\mu}_j$  en  $\hat{\sigma}_j$  naar  $\tilde{\mu}_j$  en  $\tilde{\sigma}_j$  leidt tot een afname in  $\ln L$  of een toename in  $-\ln L$ . Tabel D3 geeft de toename in  $-2 \ln L$ .

Enig idee van de bruikbaarheid van de schattingen  $\tilde{\mu}_j$  en  $\tilde{\sigma}_j$  kunnen we krijgen door de getallen in tabel D3 te vergelijken met de procentpunten van een  $\chi^2$ -variabele ("likelihood ratio" toets). Voor alle stations is de toename kleiner dan het 90%-punt (=4,6) van deze variabele, zodat we mogen concluderen dat de parameters  $\mu_j$  en  $\sigma_j$  redelijk geschat kunnen worden uit de gemiddelde wintertemperatuur  $\bar{w}$ .

De "likelihood ratio" toets geeft echter alleen maar een zeer ruwe indicatie in dit geval. De waarden van  $\tilde{\mu}_j$  en  $\tilde{\sigma}_j$  zijn niet geheel onafhankelijk van de waarnemingen van het j-de station, waardoor men in feite lagere kritieke waarden moet aanhouden. Daarnaast is de asymptotische  $\chi^2$ -benadering enigszins twijfelachtig omdat hier gewerkt wordt met reeksen van 28 getallen.

Tabel D1. Gemiddelden (over 12 stations) van de logaritme van het maximum van de aannemelijkheidsfunctie  $\bar{L}(a)$  en van de plot-correlatiecoëfficiënt  $\bar{r}(a)$  voor verschillende waarden van de ondergrens  $a$  van de lognormale verdeling bij reeksen van aantallen sneeuwdekochtenden.

$a(\text{dagen})$	$\bar{L}(a)$	$\bar{r}(a)$
0	-111,6	0,987
-5	-111,4	988
-10	-111,7	985
-15	-112,1	981
-25	-113,0	975
-35	-113,3	969
-45	-113,5	965

Tabel D2. Waarden van de toetsingsgrootheid  $D/\sqrt{N}$  (Kolmogorov-Smirnov) als functie van de ondergrens  $a$  van de lognormale verdeling bij reeksen van aantallen sneeuwdekochtenden.

Station(s)	$a=0$ dagen	$a=-5$ dagen	$a=-10$ dagen
Den Helder/De Kooy	0,25	0,41	0,49
Leeuwarden	29	31	39
Eelde	44	46	51
Twente	48	53	55
Deelen	40	51	58
Valkenburg (Z.H.)	45	61	64
De Bilt	37	35	41
Vlissingen	56	71	73
Gilze-Rijen	38	59	69
Volkel	58	74	81
Eindhoven	54	71	79
Beek (L.)	69	70	69

Tabel D3. Toename in  $-2 \ln L$  bij overgang van "maximum likelihood" schattingen van de parameters van de lognormale verdeling naar schattingen gebaseerd op de gemiddelde wintertemperatuur  $\bar{w}$  (fig. 2.7).

Station(s)	Toename $-2 \ln L$	Station	Toename $-2 \ln L$
Den Helder/De Kooy	0,3	De Bilt	0,7
Leeuwarden	1,4	Vlissingen	0,1
Eelde	2,4	Gilze-Rijen	0,1
Twente	0,6	Volkel	0,2
Deelen	2,2	Eindhoven	1,0
Valkenburg (Z.H.)	0,1	Beek (L.)	3,1

APPENDIX E - HET SCHATTEN VAN DE PARAMETERS VAN DE GUMBEL-VERDELING BIJ  
WAARNEMINGEN IN KLASSEN

De sneeuwdekhoogtemetingen van de synoptische stations in Nederland waren beschikbaar in de vorm van S-cijfers. In par. 3.3 kwam het aanpassen van de Gumbel-verdeling aan een frequentieverdeling van de jaarmaxima van deze S-cijfers ter sprake. De parameters van deze verdeling kunnen geschat worden met de methode van de grootste aannemelijkheid.

Het S-cijfer kan tien verschillende waarden aannemen (tabel 1.1). De eerste drie S-cijfers waren echter in één klasse ondergebracht, zodat we acht verschillende klassen krijgen. Voor de k-de klasse ( $k=1, \dots, 8$ ) definiëren we:

- $g_{k-1}$  : de ondergrens,
- $g_k$  : de bovengrens,
- $n_k$  : het aantal waarnemingen.

Het totaal aantal waarnemingen kunnen we schrijven als:

$$N = \sum_{k=1}^8 n_k. \quad (E1)$$

We stellen verder dat  $g_0 = -\infty$  en  $g_8 = \infty$ . De uitbreiding van de laagste klasse met het interval  $(-\infty, 0)$  is nodig omdat de theoretische Gumbel-verdeling waarden kan aannemen op dat interval.

Voor de aannemelijkheidsfunctie geldt nu:

$$L(\mu, \sigma; n_1, \dots, n_8) = \frac{N!}{n_1! n_2! \dots n_8!} \prod_{k=1}^8 \{F(g_k) - F(g_{k-1})\}^{n_k}. \quad (E2)$$

Hierin worden  $F(g_{k-1})$  en  $F(g_k)$  gegeven door verg. (3.1). Dit zijn de enige grootheden in het rechterlid van (E2), die afhankelijk zijn van de parameters  $\mu$  en  $\sigma$ . Het maximaliseren van (E2) naar  $\mu$  en  $\sigma$  is daardoor equivalent aan het maximaliseren van:

$$\ln \tilde{L} = \sum_{k=1}^8 n_k \ln \{F(g_k) - F(g_{k-1})\} \quad (E3)$$



waarin  $k_{\max}$  de hoogste waarde van  $k$  is waarvoor  $n_k > 0$ . Voor de jaarmaxima van het S-cijfer in tabel 3.1 is  $k_{\max}$  gelijk aan 4 of 5.

Differentiëren naar  $\mu$  en  $\sigma$  geeft:

$$\frac{\partial \ln \tilde{L}}{\partial \mu} = \sum_{k=1}^{k_{\max}} n_k \frac{F_{\mu}(g_k) - F_{\mu}(g_{k-1})}{F(g_k) - F(g_{k-1})} \quad (E4)$$

$$\frac{\partial \ln \tilde{L}}{\partial \sigma} = \sum_{k=1}^{k_{\max}} n_k \frac{F_{\sigma}(g_k) - F_{\sigma}(g_{k-1})}{F(g_k) - F(g_{k-1})} \quad (E5)$$

waarin:

$$F_{\mu}(g) = -\frac{1}{\sigma} e^{-(g-\mu)/\sigma} F(g) \quad (E6)$$

$$F_{\sigma}(g) = -\frac{(g-\mu)}{\sigma^2} e^{-(g-\mu)/\sigma} F(g) \quad (E7)$$

Gelijkstellen van  $\partial \ln \tilde{L} / \partial \mu$  en  $\partial \ln \tilde{L} / \partial \sigma$  aan nul levert twee niet-lineaire vergelijkingen in  $\mu$  en  $\sigma$  op. De numerieke oplossing van deze vergelijkingen wordt behandeld door Van Montfort (1981).

Asymptotische uitdrukkingen voor de standaardafwijkingen van de "maximum likelihood" schattingen volgen uit de verwachtingswaarden van de tweede afgeleiden van  $\ln L$  (of  $\ln \tilde{L}$ ). Evenals bij geschorste steekproeven (appendix B2) geven deze asymptotische uitdrukkingen voor kleine  $N$  de beste resultaten bij stations in het binnenland.

Overgang van "maximum likelihood" schattingen naar schattingen gebaseerd op de gemiddelde wintertemperatuur  $\bar{w}$  leidt tot een toename van  $-\ln L$  (appendix D). Tabel E1 geeft de toename in  $-2 \ln L$ . Alle getallen in deze tabel zijn kleiner dan het 90%-punt van de  $\chi^2_2$  variabele, zodat we mogen concluderen dat uit de gemiddelde wintertemperatuur  $\bar{w}$  redelijke schattingen voor  $\mu$  en  $\sigma$  worden verkregen.

Tabel E1. Toename in  $-2 \ln L$  bij overgang van "maximum likelihood" schattingen van de parameters van de Gumbel-verdeling naar schattingen gebaseerd op de gemiddelde wintertemperatuur  $\bar{w}$  (fig. 3.5).

Station(s)	Toename $-2 \ln L$	Station	Toename $-2 \ln L$
Den Helder/De Kooy	1,7	De Bilt	0,5
Leeuwarden	0,8	Soesterberg	1,2
Eelde	1,3	Vlissingen	0,6
Twente	0,0	Gilze-Rijen	2,5
Deelen	0,0	Volkel	0,0
Valkenburg (Z.H.)	0,0	Eindhoven	1,8
		Beek (L.)	2,2

APPENDIX F - HET SCHATTEN VAN DE PARAMETERS VAN DE EXPONENTIËLE VERDELING BIJ  
SNEEUWDEKHOOGTEN

In par 4.3 kwam voor een vijftal klimatologische hoofdstations in Nederland het aanpassen van een 2-parameter exponentiële verdeling aan een frequentieverdeling van het synopische codecijfer S ter sprake. Hiertoe kunnen we gebruik maken van de in appendix E besproken methode van de grootste aannemelijkheid. De eerste klasse bevat nu het aantal ochtenden met een S-cijfer 1; de tweede klasse het aantal ochtenden met een S-cijfer 2, enz. Daar een exponentiële variabele geen negatieve waarden kan aannemen mogen we de ondergrens  $g_0$  van de eerste klasse gelijk aan 0 stellen.

Voor  $\ln \tilde{L}$  in verg. (E3) geldt nu\*:

$$\begin{aligned} \ln \tilde{L} &= \sum_{k=1}^{k_{\max}} n_k \ln \{F(g_k) - F(g_{k-1})\} \\ &= n_1 \ln (1 - Ce^{-\lambda g_1}) + \sum_{k=2}^{k_{\max}} n_k \ln \{Ce^{-\lambda g_{k-1}} - Ce^{-\lambda g_k}\} \end{aligned} \quad (F1)$$

daar  $F(g_0)=0$ , terwijl voor  $k>0$  de waarde van  $F(g_k)$  gegeven wordt door verg. (4.2). De bovengrens in de sommatie  $k_{\max}$  is gelijk aan 5 (Den Helder/De Kooy, Vlissingen) of 6 (Eelde, De Bilt, Beek). Verg. (F1) kan worden herschreven als:

$$\begin{aligned} \ln \tilde{L} &= n_1 \ln (1 - Ce^{-\lambda g_1}) + \ln C \sum_{k=2}^{k_{\max}} n_k \\ &+ \sum_{k=2}^{k_{\max}} n_k \ln \{e^{-\lambda g_{k-1}} - e^{-\lambda g_k}\}. \end{aligned} \quad (F2)$$

Differentiëren naar C en  $\lambda$  geeft:

$$\frac{\partial \ln \tilde{L}}{\partial C} = \frac{-n_1 e^{-\lambda g_1}}{1 - C e^{-\lambda g_1}} + \sum_{k=2}^{k_{\max}} n_k / C \quad (F3)$$

$$\frac{\partial \ln \tilde{L}}{\partial \lambda} = n_1 \frac{g_1 C e^{-\lambda g_1}}{1 - C e^{-\lambda g_1}} + \sum_{k=2}^{k_{\max}} n_k \frac{-g_{k-1} e^{-\lambda g_{k-1}} + g_k e^{-\lambda g_k}}{e^{-\lambda g_{k-1}} - e^{-\lambda g_k}}. \quad (F4)$$

\* Strikt genomen is L hier geen echte aannemelijkheidsfunctie omdat er afhankelijkheid is tussen de sneeuwdekhogten op opeenvolgende dagen.

Uit  $\partial \ln \tilde{L} / \partial C = 0$  volgt:

$$C = (1 - n_1/N) e^{\lambda g_1} \quad (F5)$$

waarbij  $N$  het totale aantal waarnemingen is. Vergelijking (F5) geeft aan dat de theoretische overschrijdingskans  $C e^{-\lambda g_1}$  van de bovengrens  $g_1$  van de eerste klasse gelijk is aan de empirische overschrijdingskans  $1 - n_1/N$ . Substitutie van (F5) in (F4) en gelijkstellen aan nul geeft:

$$\frac{\partial \ln \tilde{L}}{\partial \lambda} = g_1 \sum_{k=2}^{k_{\max}} n_k + \sum_{k=2}^{k_{\max}} n_k \frac{g_k e^{-\lambda g_k} - g_{k-1} e^{-\lambda g_{k-1}}}{e^{-\lambda g_{k-1}} - e^{-\lambda g_k}} = 0. \quad (F6)$$

Uit deze vergelijking kan  $\lambda$  door iteratie worden opgelost. Opgemerkt dient te worden dat de oplossing  $\hat{\lambda}$  niet afhangt van het aantal waarnemingen  $n_1$  in de eerste klasse. Substitutie van  $\hat{\lambda}$  in (F5) geeft de "maximum likelihood" schatting  $\hat{C}$ .

Bij de gegevens van de Duitse stations in fig. 4.1 moet uitgegaan worden van de voorwaardelijke verdeling (4.7) met  $t = \frac{1}{2}$ . Voor de logaritme van de aannemelijkheidsfunctie in verg. (E3) krijgen we nu:

$$\ln \tilde{L} = \sum_{k=1}^k n_k \ln \{ e^{-\lambda(g_{k-1} - \frac{1}{2})} - e^{-\lambda(g_k - \frac{1}{2})} \}. \quad (F7)$$

Voor de ondergrens van de eerste klasse geldt nu  $g_0 = \frac{1}{2}$ ; de bovengrens  $k_{\max}$  in de sommatie is bij de gegevens van de Duitse stations gelijk aan 6.

Vergelijking (F7) kan omgewerkt worden tot:

$$\ln \tilde{L} = \frac{\lambda N}{2} + \sum_{k=1}^k n_k \ln \{ e^{-\lambda g_{k-1}} - e^{-\lambda g_k} \} \quad (F8)$$

waarbij  $N$  weer het totale aantal waarnemingen aangeeft. De "maximum likelihood" schatting van  $\lambda$  wordt nu verkregen uit de vergelijking:

$$\frac{\partial \ln \tilde{L}}{\partial \lambda} = \frac{N}{2} + \sum_{k=1}^k n_k \frac{g_k e^{-\lambda g_k} - g_{k-1} e^{-\lambda g_{k-1}}}{e^{-\lambda g_{k-1}} - e^{-\lambda g_k}} = 0. \quad (F9)$$

Deze vergelijking heeft dezelfde gedaante als vergelijking (F6).

## LITERATUUR

- Barr, D.R. and Davidson, T. (1973): A Kolmogorov-Smirnov test for censored samples. *Technometrics*, 15, 739-757.
- Buishand, T.A. (1982): Some methods for testing the homogeneity of rainfall records. *J. Hydrol.*, 58, 11-27.
- Buishand, T.A. (1984): Tests for detecting a shift in the mean of hydrological time series. *J. Hydrol.*, 73, 51-69.
- Caspar, W. (1962): Die Schneedecke in der Bundesrepublik Deutschland (Tabellen). Deutscher Wetterdienst, Offenbach am Main.
- Caspar, J.W. und Krebs, M. (1974): Auswertung langjähriger Beobachtungen über Schneehöhen und Schneelasten. Forschungsbericht Deutscher Wetterdienst, Offenbach am Main.
- Chandra, M., Singpurwalla, N.D. and Stephens, M.A. (1981): Kolmogorov statistics for tests of fit for the extreme-value and Weibull distributions. *J. Amer. Statist. Assoc.*, 76, 729-731.
- Clarke, P.C. (1969): Snowfalls over south-east England, 1954-1969. *Weather*, 24, 438-447.
- Cohen, A.C., Jr. (1959): Simplified estimators for the normal distribution when samples are singly censored or truncated. *Technometrics*, 1, 217-237.
- Cohen, A.C., Jr. (1961): Tables for maximum likelihood estimates: singly truncated and censored samples. *Technometrics*, 3, 534-541.
- Cohen, A.C. (1965): Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. *Technometrics*, 7, 579-588.
- Cunnane, C. (1978): Unbiased plotting positions - a review. *J. Hydrol.*, 37, 205-222.
- David, H.A. (1981): *Order Statistics*, 2nd edition. John Wiley, New York.
- Deutscher Wetterdienst (1957): *Klima-atlas von Rheinland-Pfalz*. Bad Kissingen.
- Deutscher Wetterdienst (1960): *Klima-atlas von Nordrhein-Westfalen*. Offenbach am Main.
- Deutscher Wetterdienst (1964): *Klima-atlas von Niedersachsen*. Offenbach am Main.
- Deutscher Wetterdienst (1967): *Klima-atlas von Schleswig-Holstein, Hamburg und Bremen*. Offenbach am Main.
- Ellingwood, B. and Redfield R. (1983): Ground snow loads for structural design. *J. Struct. Div. Amer. Soc. Civ. Eng.*, 109, 950-964.

- Ellingwood, B. and Redfield, R.K. (1984): Probability models for annual extreme water-equivalent ground snow. *Mon. Weather Rev.*, 112, 1153-1159.
- Filliben, J.J. (1975): The probability plot correlation coefficient test for normality. *Technometrics*, 17, 111-117.
- Fröhlich, W. (1970): Die n-Jährlichkeit grösster winterlicher Schneehöhen in der Steiermark. *Wetter und Leben*, 22, 105-107.
- Gumbel, E.J. (1958): *Statistics of Extremes*. Columbia University Press, New York.
- Horst, J. van der (1984): Sneeuwdekhoogten in Nederland naar metingen van de synoptische stations 1956-1977. *Techn. Rapp. TR-51*, KNMI, De Bilt.
- Horst, J. van der (1986): Sneeuwdekhoogten in Nederland naar metingen van de synoptische stations in 1955 en 1978-1984. *Techn. Rapp. TR-51a*, KNMI, De Bilt.
- Jackson, M.C. (1977): A method of estimating the probability of occurrence of snow water equivalents in the United Kingdom. *Hydrological Sciences Bulletin*, XXII, 127-142.
- Jackson, M.C. (1978): Snow cover in Great Britain. *Weather*, 33, 298-308.
- Johnson, N.L. and Kotz, S. (1970): *Distributions in Statistics. Continuous Univariate Distributions-1*. John Wiley, New York.
- Kalb, M. und Noll, H. (1980): *Klimadaten von Europa. Teil I: Nord-, West- und Mitteleuropa*. Deutscher Wetterdienst, Offenbach am Main.
- KNMI (1972): *Klimaatatlas van Nederland*. Staatsuitgeverij, 's-Gravenhage.
- KNMI (1982): *Normalen en standaardafwijkingen voor het tijdvak 1951-1980. Klimatologische gegevens van Nederlandse stations, No. 10*, KNMI, De Bilt.
- Kuusisto, E. (1984): Snow accumulation and snowmelt in Finland. *Publications of the Water Research Institute, No. 55*, Helsinki.
- Leese, M.N. (1973): Use of censored data in the estimation of Gumbel distribution parameters for annual maximum flood series. *Water Resources Research*, 9, 1534-1542.
- Martinec, J. (1977): Expected snow loads on structures from incomplete hydrological data. *J. Glaciology*, 81, 185-197.
- Meeker, W.Q., Jr. and Nelson, W. (1977): Weibull variances and confidence limits by maximum likelihood for singly censored data. *Technometrics*, 19, 473-476.
- Montfort, M.A.J. van (1981): Comparison of distributions. *Technical Note, No. 81-1*, Dept. of Mathematics, Agricultural University, Wageningen.

- Muizelaar, P. en Ybema, K. (1985): Sneeuw-overzicht, januari 1985. Weerspiegel, 12, 87-109.
- Pearson, E.S. and Hartley, H.O. (1976): Biometrika Tables for Statisticians, Vol. 2. Biometrika Trust, London.
- Piekema, J.J. (1984): Het sneeuwdek in Midden Nederland: dagelijkse metingen van de sneeuwhoogte 1950-1984. Techn. Rapp. TR-50, KNMI, De Bilt.
- Schlaak, P. und R. (1968): Tägliche Schneehöhen in Berlin 1892-1967 (75 Jahre). Meteorologische Abhandlungen, Bd. 92, Hft. 1, Beilage 10/68. Institut für Meteorologie und Geophysik der Freien Universität Berlin.
- Sieland, K. (1979): Tägliche Schneehöhen in Berlin seit 1892. Berliner Wetterkarte, Neue Folge, Serie B, Bd. 23, Heft 12, Beilage. Institut für Meteorologie der Freien Universität Berlin.
- Sneyers, R. (1964): La statistique de l'enneigement du sol en Belgique. Arch. Met. Geoph. Biokl., Serie B, 13, 503-520.
- Sneyers, R. (1967): Les propriétés statistiques de l'enneigement du sol en Belgique. Pub. Série A, No. 63, Inst. R. Mét. Belg., Uccle-Bruxelles.
- Stirling, R. (1982): The Weather of Britain. Faber and Faber, London.
- Thom, H.C.S. (1966): Distribution of maximum annual water equivalent of snow on the ground. Mon. Weather Rev., 94, 265-271.
- Vogelaere, D. de (1976): Sneeuwklimatologie van België. Licentiaatsverhandeling Katholieke Universiteit Leuven.
- Zerche, M. (1974): Die jährlichen Höchstwerte des Wassergehalts der Schneedecke als Grundlage der potentiellen Schneelastgefährdung in der DDR. Zeitschrift für Meteorologie, 24, 155-160.