12 APR. 1960

KONINKLIJK NEDERLANDS
METEOROLOGISCH INSTITUUT

Verslagen V-64
(R III-245-1960)

Over de interdiurne variabiliteit van enkele klimatologische grootheden

door

Dr. C. Levert

Over de interdiurne variabiliteit van enkele klimatologische grootheden

0. Inleiding

In dit rapport wordt een beknopt overzicht gegeven van de berekeningen inzake de interdiurne variabiliteit (afgekort I.V.), d.w.z. de veranderingen van dag op dag, in de maximum temperatuur T ^{o}C, in de minimum temperatuur t ^{o}C (beide geldig voor het etmaal tussen 19.40 en 19.40 h M.E.T.) en in de relatieve vochtigheid U%, gemeten om 14.40 M.E.T. De berekeningen werden uitgevoerd voor 3 stations: Naaldwijk (26 jaren: 1931 - '56), Gemert (26 jaren: 1931 - '56) en Groningen (20 jaren: 1926 - '50), voor elke dekade van elke maand.

Een volledig verslag, beschrijvende de gevolgde werkwijze, de statistische overwegingen, de toelaatbaarheid van benaderingen, de constructie van een nomogram, dat kan worden gebruikt om snel vele vragen uit de praktijk te beantwoorden, is opgenomen in het K.N.M.I.-rapport W.R. 60-2 (R III-248-1960).

Vooraf mogen enkele opmerkingen gemaakt worden.

De minimumtemperatuur, gemeten bijv. in het etmaal tussen 19.40 van 3 januari en 19.40 van 4 januari, heet t_4, kortweg de "minimumtemperatuur van 4 januari". De interdiurne verandering (de verandering van de ene op de direkt volgende dag) is gedefinieerd als $\Delta_4 = t_4 - t_3$, gedateerd op 4 januari. Deze differentie Δ_i heeft een teken. Als de minimumtemperatuur van dag i-1 op dag i toeneemt, is Δ_i positief. Analoge opmerkingen gelden voor de dagelijkse maximumtemperatuur en de relatieve vochtigheid.

In het boven aangehaalde rapport worden ook tussen-twee-daagse, tussen-drie-daagse, enz., veranderingen (differenties van de orden twee, drie, enz.) besproken, geschreven ${}_k\Delta_i = t_i - t_{i-k}$; i = datum; k = orde.

Zolang wij over de "gewone" interdiurne veranderingen (eerste-orde-differenties) spreken, wordt de vooraangehangen index 1 in ${}_1\Delta_i$ weggelaten.

De dekaden zijn gedefinieerd a.v.

jan.I	:	1 t/m 10 jan.	feb.I	:	31 jan. t/m 9 feb.
" II	:	11 t/m 20 "	" II	:	10 feb. t/m 19 feb.
" III	:	21 t/m 30 "	" III	:	20 " t/m 1 mrt., c.q.29 feb.

mrt. I	:	2 t/m 11 mrt.
" II	:	12 t/m 21 "
" III	:	22 t/m 31 "

Overige maanden: I: 1 t/m 10; II: 11 t/m 20; III: 21 t/m 30. De dagen 31 mei, 31 juli, 31 augustus, 31 oktober, 31 december en in een schrikkeljaar ook 1 maart werden weggelaten. Wel werd er voor gezorgd, dat elke Δ_i op de verandering van één dag betrekking heeft; ter verduidelijking: de 10 Δ's der derde dekade van juli zijn Δ_{21} = t(21 juli) - t(20 juli) .... Δ_{30} = t(30 juli) - t(29 juli), direkt gevolgd door Δ_1 = t(1aug.)-t(31 juli)....... Δ_{10} = t(10 aug.)-t(9 aug.) enz.

Elk van de 36 frekwentieverdelingen der Δ-waarden, t.w. voor elke dekade één, is gebaseerd, althans in het geval dat de basisperiode 26 jaren lang is, op 260 Δ-waarden; immers er liggen, om een voorbeeld te noemen, in de 26 jan.I-dekaden der 26 jaren tezamen 260 dagen.

In het volgende noemen wij de titels der hoofdstukken uit het genoemde W.R.60-2; R III 248 , terwijl slechts een beknopte samenvatting van de inhoud van elk hoofdstuk gegeven wordt.

1. Karakterisering van de interdiurne variabiliteit

Dikwijls ziet men de I.V. gekarakteriseerd door het gemiddelde der absolute Δ-waarden, geschreven $\overline{|\Delta|}$. De praktijkvragen zijn echter zodanig, dat men de gehele kansverdeling der Δ-waarden, de tekens in acht genomen, moet kennen. Daardoor komen vele statistische details ter sprake. Is de kansverdeling symmetrisch (anders gezegd komen pluswaarden gemiddeld even dikwijls voor als minwaarden) ? Is de kansverdeling normaal? Welke statistische grootheden (parameters) (μ, σ, ε en β) zullen ter karakterisering van de I.V. gebruikt kunnen worden? Ter toelichting: σ heet standaarddeviatie; σ^2 heet variantie; ε heet gemiddelde absolute waarde; β heet waarschijnlijke waarde.

Indien f(Δ) dΔ de waarschijnlijkheidsverdeling van Δ (teken in acht genomen) voorstelt, dan is $\sigma^2 = \int_{-\infty}^{\infty} (\Delta - \mu)^2 f(\Delta) d\Delta$, waarin

$\mu = \int_{-\infty}^{\infty} \Delta f(\Delta) d\Delta$; $\varepsilon = \int |\Delta| f(\Delta) d\Delta$; $\int_{-\beta}^{\beta} f(\Delta) d\Delta = \frac{1}{2}$

Deze parameters hebben betrekking op het universum der Δ -waarden. Gekozen werd σ ter karakterisering van de I.V., geschat in de steekproeven als s . In de tabellen wordt echter ook $e = \overline{|\Delta|}$, schatting van ε , genoemd.

2. De machinaal geleverde lijsten van statistische rekengrootheden

Als voorbeeld wordt hier de dagelijkse minimumtemperatuur t te Gemert gekozen.

Voor elk der dagen uit het 26 jarige tijdvak 1931 t/m 1956 werd de dagelijkse minimumtemperatuur (t_i op dag i) geponst. Het tijdvak werd verdeeld in drie deeltijdvakken, t.w. twee van 10 jaren (1931-'40 en 1941-'50) en één van 6 jaren (1951-'56), dat later tot 10 jaren uitgebreid wordt. De bedoeling is aldus een "trend" te onderzoeken.
De volgende bewerkingen werden machinaal verricht.
Voor iedere dag i van elke dekade werden 6 differenties ${}_k\Delta_i$ machinaal berekend, voor k = 1, 2, 3, 4, 5, 6 , gedefinieerd a.v.:

$k = 1$	${}_1\Delta_1 = t_1 - t_0$	${}_1\Delta_2 = t_2 - t_1$	......${}_1\Delta_{10} = t_{10} - t_9$
$k = 2$	${}_2\Delta_1 = t_1 - t_{-1}$	${}_2\Delta_1 = t_2 - t_0$	......${}_2\Delta_{10} = t_{10} - t_8$
.			
.			
$k = 6$	${}_6\Delta_1 = t_1 - t_{-5}$	${}_6\Delta_1 = t_2 - t_{-4}$	......${}_6\Delta_{10} = t_{10} - t_4$

Deze ${}_k\Delta_i$'s zijn positief, nul of negatief.
Nadat al deze differenties berekend waren, werden ze voor elk der drie genoemde deelperioden gegroepeerd naar de 36 dekaden per jaar, waarbij, zoals gezegd, elke dekade tien dagen lang gemaakt werd.
Deze ${}_k\Delta$-waarden, in het bijzonder hun sommen S_k en N (verderop genoemd), zijn nodig voor de berekening van de autocorrelatiecoëfficiënten van de orden 1, 2, 3.... (aangeduid met r_1, r_2....) in de tijdreeks der t-waarden, bij welke berekening een vereenvoudigde procedure (§ 5) gevolgd werd.

Thans volgt een korte opsomming van de lijsten A, B, C, D, E.

A. De eerste grote lijst bevat voor elk der dagen van de 36 dekaden, voor elk der 26 jaren afzonderlijk, de t en de ${}_k\Delta$, k = 1,2,3,4,5,6.
De Δ's werden in twee kolommen afgestempeld: $\Delta < 0$ en $\Delta \geqslant 0$. Bovendien werden berekend en gestempeld per afzonderlijke dekade de sommen $\sum_{i=1}^{10} t$, $\sum_{i=1}^{10} ({}_k\Delta_i < 0)$, $\sum_{i=1}^{10} ({}_k\Delta_i \geqslant 0)$ en $\sum_{i=1}^{10} |{}_k\Delta_i|$. De deelperioden, aan-

geduid met $j = 1,2,3$, werden ook voor elke dekade afzonderlijk gesommeerd over elk der drie deelperioden; zo werden verkregen de 3 x 6 = 18 sommen $S_{jk} = \sum_j \left\{ \sum_{i=1}^{10} |_k\Delta_i| \right\}$, $j = 1,2,3$; $k = 1,2 \ldots 6$. Zij $S_k = \sum_j S_{jk}$, dan stelt $\frac{1}{260} S_k$ de gemiddelde absolute Δ, d.i. $\overline{|\Delta|}$, voor. Eveneens afgestempeld werden voor elk der dekaden en gesommeerd over de deelperioden, de sommen $M_j = \sum_j \left\{ {}^{10}_{1} t_i \right\}$, $j = 1,2,3$. Zij $M = \sum_1^3 M_j$, dan stelt $M : 260$ de gemiddelde minimumtemperatuur $\bar{t}$ voor, d.i. het 26 jarige gemiddelde voor één bepaalde dekade.

B. Een tweede lijst bevat voor elke dekade afzonderlijk en gesommeerd over de gehele 26 jarige periode, de zes totale sommen $\sum ({}_k\Delta < 0)$ en de zes totale sommen $\sum ({}_k\Delta \geqslant 0)$, $k = 1,2 \ldots 6$. Hiermede werd $\bar{\Delta} = \frac{1}{260} \sum \Delta$ berekend, d.i. de gemiddelde waarde van de 260 ééndaagse differenties uit één dekade (over 26 jaren). Dit gemiddelde ligt doorgaans vlak bij nul, waarin een aanwijzing gezien mag worden van de symmetrie der Δ - verdeling.

C. Een derde lijst bevat voor elke dag van bijv. jan.I van elk der 26 jaren, het verschil $d_i = t_i - \bar{t}$, waarbij $\bar{t}$ het veeljarig gemiddelde voorstelt (genoemd in A). Deze d's werden afgestempeld over twee kolommen, $d < 0$ en $d \geqslant 0$. Tevens werd berekend en gestempeld voor elke individuele dekade de som $N_v = \sum_1^{10} |d_i|$; $v = 1,2 \ldots 26$. Dit levert in totaal voor één station 26 x 3 = 78 sommen. Als $N = \sum_1^{26} N_v$, stelt $N : 260$ de $\overline{|d|}$ voor, d.i. het gemiddelde van de absolute d voor één bepaalde dekade. Vanzelfsprekend is $\overline{|d|} > 0$, terwijl, louter door de definitie van $\bar{t}$, echter $\bar{d} = 0$ is. Aldus $\overline{|d|}$ bekend zijnde, vormt $\overline{|d|} \sqrt{\tfrac{1}{2}\pi}$ een schatting van de standaarddeviatie der t-waarden in het universum. Deze schatting is slechts correct als de t-waarden een normale verdeling volgen, hetgeen wij gemakshalve als juist onderstelden. Aldus werden op een machinaal zeer eenvoudige wijze de 36 standaarddeviaties s_t der 36 t-universa berekend.

D. Een vierde lijst bevat de distributieve en cumulatieve frekwentieverdelingen der $|\Delta|$-waarden (alleen ${}_1\Delta$), ondergebracht in de klassen 0.0-0.9 1.0-1.9.....°C. Dit afzonderlijk voor alle dagen uit de dekaden jan.I uit deelperiode 1; idem dekaden jan.I, deelperiode 2; idem dekaden jan.I, deelperiode 3 enz. enz. t/m dekaden december III, deelperiode 3. Zo ontstonden 12x3x3 = 108 frekwentieverdelingen der $|\Delta|$'s (elk gebaseerd op òf 100 òf 60 getallen), immers men heeft 12 maanden;

elke maand 3 dekaden; 3 perioden. Dit werd uitgevoerd voor t, T, U, tezamen 324 frekwentieverdelingen.
Het ligt in de bedoeling om aan de hand van deze frekwentieverdelingen een "trend" gedurende de jaren 1931 t/m 1956 te onderzoeken. Aannemende, dat een "trend" afwezig is, werden de frekwentieverdelingen met de hand samengevoegd, zodat er 36 nieuwe frekwentieverdelingen der $|\Delta|$ -waarden ontstonden (althans voor één station en één element). Deze werden omgezet in frekwentieverdelingen der Δ's zelf (plus, nul en minwaarden), hetgeen kon geschieden slechts als de symmetrie der Δ-verdelingen ondersteld mocht worden.
Voor verschillende statistische notities bij de tegenoverelkaarstelling van de frekwentieverdeling der t-waarden enkel op bijv. 3 januari (26 getallen) en die der t-waarden van 1,2,....10 januari (260 getallen) en idem voor de -waarden, verwijzen wij naar het W.R. 60-2

E. Een vijfde lijst bevat voor elk der 108 onder D genoemde frekwentieverdelingen der $|\Delta|$ -waarden zekere hulpgrootheden, met behulp waarvan het mogelijk is om de variantie $s^2_{|\Delta|}$ der $|\Delta|$ -waarden te berekenen. Uit de $s^2_{|\Delta|}$ kan weer de s^2_{Δ} berekend worden.

3. <u>Een onderzoek naar de normaliteit van de kansverdeling der eerste-orde-differenties</u>

Een eenvoudig nomogram bleek slechts dan geconstrueerd te kunnen worden als de Δ-verdelingen alle normaal (volgens Gausz) zouden zijn. Het onderzoek daaromtrent zou hebben kunnen geschieden bijv. met de χ^2-toets of met de toets van Geary. Aangezien het hier 324 frekwentieverdelingen betreft kozen wij, om werk te sparen, de visuele weg. Alle frekwentieverdelingen werden op lineair waarschijnlijkheidspapier uitgezet en op het oog werd beoordeeld of de puntenreeksen vrijwel lineair lagen.
Conclusie: zeer waarschijnlijk zijn alle Δ-verdelingen zo goed als normaal, zowel wat T, t als U betreft.

4. <u>Over de noodzakelijkheid van kennis van de persistentie bij onderzoekingen van de jaarlijkse gang van de interdiurne variabiliteit en van de verschillen tussen stations onderling.</u>

Hoewel iedere variantie s^2_{Δ} op 10 x 26 = 260 Δ-waarden berust, is het statistisch effectieve aantal (in het geciteerde W.R. nader gedefinieerd) anders; in het onderhavige geval, waarin de tijdreeks der Δ-waarden een negatieve autocorrelatie bezit (veel meer komt positieve voor) is het effectieve aantal groter dan 260 en wel groter in een van de autocorrelatiecoëfficiënten afhankelijke mate. Kennis van deze effectieve aantallen is onontbeerlijk bij een onderzoek van de statistische realiteit van de verschillen tussen twee of meer δ_Δ^2 -waarden, welke voor één enkel station op verschillende dekaden uit het jaar betrekking kunnen hebben of voor één enkele dekade op verschillende stations. De autocorrelatie in de Δ-reeks is in de plaats het gevolg van het feit, dat Δ_i "een term gemeen" heeft met Δ_{i-1}, immers $\Delta_i = t_i - t_{i-1}$ en $\Delta_{i-1} = t_{i-1} - t_{i-2}$. Ze wordt verkleind qua grootte en vergroot qua doorwerking door het feit, dat de tijdreeks der t-waarden een positieve autocorrelatie bezit (men zegt: de t-reeks bezit persistentie). Er wordt aangetoond, dat de autocorrelatiecoëfficiënten van de orden 1,2,3.. (d.i. $\rho_1, \rho_2, \ldots$) gehoorzamen aan de wet $\rho_k = \rho^k$. Een groep van n successieve Δ's bevat, uitgaande van deze wet, effectief $n^2(1-\rho):(1-\rho^n)$ elementen.

5. De berekening van de autocorrelatiecoëfficiënten r_k.

In 4 werd uiteengezet waarom kennis der r_k's ($k = 1,2,3\ldots$) nodig bleek. Er werd een vereenvoudigde - voor machinale berekening geschikte - procedure uitgewerkt, volgens welke de r_k-waarden werden berekend ($k = 1$ t/m 6). De procedure is zeker niet exact, doch voldoende goed. Zij maakt gebruik van de $|{}_k\Delta|$ waarden, redenen, waarom de differenties der orden 1 t/m 6 berekend en, absoluut, gemiddeld werden, gelijk in 2 uiteengezet werd.

6. De aanpassing van de wet $\rho_k = \rho^k$ aan de berekende zes autocorrelatiecoëfficiënten $r_1, \ldots r_6$.

Voor elk der 36 dekaden over de gehele 26 (20) jarige reeks, voor elk der 3 elementen, voor elk der 3 stations, werden $r_1, r_2 \ldots r_6$ berekend. Vervolgens was nodig een formule op te stellen voor de relatie tussen r_k en k (r_k neemt af met k). De in de literatuur veel genoemde - echter slechts onder bepaalde voorwaarden geldende - wet $\rho_k = \rho^k$ werd getoetst. Zij bleek voldoende goed te gelden. Slechts indien deze simpele relatie geldt worden verdere berekeningen zo eenvoudig mogelijk. In dit hoofdstuk wordt de aanpassing en de beoordeling van de juistheid daarvan uitvoerig besproken.

7. De formule $n_e = n\frac{1-\rho}{1+\rho}$ in tijdreeksen met autocorrelatiecoëfficiënten $\rho_k = \rho^k$

In dit hoofdstuk wordt aangetoond dat, indien $\rho_k = \rho^k$ geldt, de exacte uitdrukking voor het met het nominale aantal n statistisch equivalente effectieve aantal n_e te schrijven is, mits n voldoende groot en ρ voldoende klein is, als $n\frac{1-\rho}{1+\rho}$. Natuurlijk wordt ook het begrip "statistisch equivalentie" gedefinieerd.

8. De jaarlijkse gang in ρ

Aangezien $\rho_k = \rho^k$ geldt, behoeven wij slechts de jaarlijkse gang in ρ_1 (de index 1 weglatend: ρ), d.i. de autocorrelatiecoëfficiënt van de eerste orde te onderzoeken. Het bleek, dat de ρ's voor de verschillende stations, mits op eenzelfde element en eenzelfde dekade betrekking hebbend weinig of niet verschillen, zodat wij de drie stations in dit opzicht samen mochten nemen. De ρ is daarentegen wel duidelijk van het element afhankelijk, terwijl de jaarlijkse gang markant is.

9. Berekening van σ_Δ uit $\sigma_{|\Delta|}$

In hoofdstuk 1 werd uiteengezet, dat alle Δ -kansverdelingen (234 stuks; 3 elementen; 26 dekaden; 3 stations) normale verdelingen rondom nul zijn, die alleen in de standaarddeviaties σ_Δ verschillen. De berekening van de 234 σ_Δ-waarden geschiedde weer voor een groot deel machinaal, gelijktijdig met de constructie der frekwentieverdelingen. In het geval van de Δ -waarden bij de minimum- en de maximumtemperatuur (t resp. T) geschiedde dit via de standaarddeviaties $\sigma_{|\Delta|}$ der $|\Delta|$ -verdelingen. De σ_Δ laat zich uit de $\sigma_{|\Delta|}$ berekenen, alleen als voor de Δ -verdeling aan zekere voorwaarden voldaan is, o.m. als zij normaal is. Een en ander is puur statistisch zeer interessant, redenen, waarom er in dit hoofdstuk vrij uitvoerig over gesproken wordt.

10. Het nomogram bij de interdiurne variabiliteit.

Na alle genoemde noodzakelijke berekeningen bleek het mogelijk per element één enkel nomogram te maken, i.p.v. met 3 x 36 frekwentieverdelingen (3 stations; 36 dekaden) te werken. Zulk een nomogram geldt dan bijv. voor t; het bevat drie krommen, t.w. één voor Gemert, één voor Naaldwijk en één voor Groningen.

Bij dit hoofdstuk behoort een afzonderlijke nota: V-62; RIII-243-1960: "Numerieke voorbeelden als toelichting bij het gebruik van de nomogrammen der interdiurne variabiliteit".

11. Typen vragen, die niet met het nomogram te beantwoorden zijn.

11.1 Welke kans is er te Naaldwijk in jan.II op een onafgebroken reeks van precies 5 dagen, gedurende welke van dag tot dag de maximumtemperatuur niet meer dan 4 °C verandert?
Uiteengezet wordt, hoe deze vraag zou kunnen worden beantwoord, mits er meer berekeningen uitgevoerd worden.

11.2 Welke is de gemiddelde lengte van een onafgebroken reeks van dagen met toenemende maximumtemperatuur in jan.II te Naaldwijk?

11.3 Gegeven de minimumtemperatuur te Naaldwijk is op zekere dag in jan.II 4 °C. Welke kans is er op vorst de volgende dag (d.i. $\Delta \leqslant -4$ °C)?

11.4 Welke kans is er, dat de gehele jan.II te Naaldwijk tenminste één $\Delta \geqslant 4$ °C telt?

11.5 Welke kans is er te Gemert in jan.II dat in een enkele dag de maximumtemperatuur tenminste 6 °C stijgt, terwijl de relatieve vochtigheid met tenminste 25% daalt?

24 maart 1960.