K O N I N K L IJ K N E D E R L A N D S
M E T E O R O L O G I S C H I N S T I T U U T

Wetenschappelijk Rapport W.R. 60-2 (III-248)

Dr. C. Levert

Onderzoek naar de interdiurne variabiliteit van enkele meteorologische grootheden

De Bilt - 1960

KONINKLIJK NEDERLANDS
METEOROLOGISCH INSTITUUT

Wetenschappelijk Rapport
W.R. 60-2 (III-248)
U.D.C.: 551.5: 519.2

Dr. C. Levert

Onderzoek naar de interdiurne variabiliteit van enkele meteorologische grootheden

INHOUD

0. Inleiding

Het initiatief tot het hier beschreven onderzoek is uitgegaan van een in 1955 door de Adviescommissie voor het Tuinbouwvestigingsplan opgerichte werkgroep, die tot taak had na te gaan in hoeverre verschillende gebieden in Nederland in klimatologisch opzicht geschikt voor tuinbouwvestiging zouden kunnen zijn. Daarbij kwam naar voren, dat voor verschillende takken van tuinbouw grote waarde gehecht moest worden aan een zekere "gelijkmatigheid" van het klimaat.

Het bleek zeer moeilijk om met behulp van de onmiddellijk beschikbare klimatologische gegevens deze "gelijkmatigheid" op bevredigende wijze in cijfers uit te drukken. De moeilijkheid is niet slechts gelegen in het ontbreken van desbetreffende gegevens, doch ook in de omstandigheid, dat het begrip "gelijkmatigheid" niet gemakkelijk te definiëren is. Het heeft, voelt men wel aan, zeker te doen met de veranderingen van dag op dag, maar met deze niet alleen. Voor een eerste oriëntatie achtte men een onderzoek van deze interdiurne variabiliteit (verderop nader gepreciseerd) het meest op zijn plaats.

Deze interdiurne variabiliteit werd reeds eerder bestudeerd in Nederland door Hartman. Aangezien zijn onderzoek, gepubliceerd in Med. en Verh. K.N.M.I., No. 24, 1918,

a) op slechts één element betrekking heeft; de temperatuur (dagelijkse ware-, dagelijkse maximum- en dagelijkse minimumwaarde)
b) slechts één station betreft: De Bilt
c) niet op recente gegevens berust (1897-1916)
d) voor ons doel niet voldoende uitvoerig is

werd echter besloten een nieuwe studie te ondernemen

Het volgende drietal stations werd in beschouwing genomen; Naaldwijk, Gemert en Groningen:

Naaldwijk kan als representatief voor het Westland (glascultuur) beschouwd worden. Niet allen zijn het daarmee eens. Sommigen achten de mogelijkheid aanwezig, dat het station onderhevig was aan iets als een "stadseffect", niet zo zeer vanwege een dichte behouwing met huizen, als wel gezien de grote oppervlakte aan glas en gezien het feit, dat de kassen voor een groot deel van het jaar verwarmd worden. Het effect zal zeker minder groot zijn dan dat te Groningen.

Gemert verkeert in omstandigheden, die niet veel zullen verschillen van die in het glascentrum rondom Venlo.

De keuze van Groningen is misschien minder gelukkig. Er is hierbij gedacht aan vestiging van nieuwe glastuinbouwgebieden in het noorden (noordoosten) des lands. Het voor dit gebied representatieve station moest in of rondom Groningen gezocht worden. Bij gebrek aan beter werd dat te Groningen gekozen. Het is zeer waarschijnlijk, dat het de invloed van de nabij gelegen Waddenzee ondervindt en het is zeker dat er een "stadseffect" zal zijn. Wat dit "stadseffect" betreft: In doorsnee genomen zullen in een stad de maximumtemperaturen hoger en de minimumtemperaturen minder laag liggen. Dit is allang bekend. Minder bekend zijn de consequenties op de tussendaagse schommelingen. Wij vermoeden, dat deze, weer is doorsnee genomen, groter, resp. kleiner zullen zijn, dan op een in de onmiddellijke nabijheid gelegen "vrij" station. Misschien is dit later eens te verifiëren, zodra het station Eelde over voldoende gegevens beschikt.

Het bleek uitermate gewenst de daggegevens op ponskaarten te brengen, zodat talrijke bewerkingen via hulpkaarten volledig machinaal konden worden uitgevoerd. Tegelijk werd aandacht geschonken aan het verschijnsel der autocorrelatie (d.i. persistentie). Het is wellicht op het eerste gezicht niet duidelijk wat dit verschijnsel met de I.V. (afkorting voor "interdiurne variabiliteit) te doen heeft, maar als men in statistische finesses treedt (waartoe men wordt genoodzaakt bij het toepassen van statistische toetsen, bij kwesties van betrouwbaarheden, bij vragen naar de realiteit van verschillen tussen stations enz.), dan blijken de I.V. en de persistentie onmogelijk los van elkaar beschouwd te kunnen worden. Zo groeide dit onderzoek uit tot een studie, die veel meer dan de I.V. alleen omvat.

Het onderhavige rapport beschrijft zo goed als volledig de gevolgde werkwijze, de statistische overwegingen, de toelaatbaarheid van benaderingen en tenslotte de constructie van een nomogram, waarmede vele met de I.V. samenhangende vragen eenvoudig en snel kunnen worden beantwoord. Tevens worden enkele typen van vragen geformuleerd, die met dit nomogram niet of slechts gebrekkig (en onbetrouwbaar) kunnen worden beantwoord, maar met het beschikbare ponskaartenmateriaal wel aangepakt zouden kunnen worden. Het rapport beschouwt niet de tuinbouwtechnische aspecten der resultaten. Zie hiervoor het blad "Mededelingen van de Directeur van de Tuinbouw", augustus 1960.

Voor Naaldwijk en Gemert is de beschouwde periode 1931-1956 (26 jaren), voor (de stad) Groningen 1931-1950 (20 jaren). Er worden 3 deelperioden onderscheiden: 1931-1940 (10 jaren), 1941-1950 (10 jaren) en 1950-1956 (6 jaren), welke later zou kunnen worden verlengd tot en met 1959.

De bestudeerde elementen zijn de dagelijkse minimumtemperaturen t, de dagelijkse maximumtemperatuur T en de 14h-relatieve vochtigheid (eenheden ^{o}C en %).

Verduidelijking: De minimumtemperatuur tussen 19.40 uur van bijv. 3 en 19.40 uur (M.E.T.) van 4 januari te Naaldwijk heet de minimumtemperatuur t_4 van 4 januari. De interdiurne verandering (van de eerste orde) ${}_1\Delta_4$ (korter: Δ_4) is dan gedefinieerd als ${}_1\Delta_4 = t_4 - t_3$ en is met 4 januari gedateerd. De differentie heeft een teken. Algemener: aan dag no. i zijn 6 differenties, orden k = 1, 2, 3, 4, 5 en 6, gekoppeld (natuurlijk zijn er nog meer, doch gemeend werd met 6 orden te kunnen volstaan), geschreven ${}_k\Delta_i$, gedateerd op i (in een dekade: i = 1, 2, ... 10) definitie is: ${}_k\Delta_i = t_i - t_{i-k}$.

Al het rekenwerk geschiedde per dekade. Er waren grote machinetechnische voordelen als iedere dekade precies 10 dagen zou tellen (het jaar derhalve 360 dagen lang zou zijn). Daarom werden de dekaden gedefinieerd als volgt:

jan. I : 1-10 II : 11-20 III: 21-30	feb. I : 31 jan. - 9 feb. II : 10 " - 19 " III: 20 " - 1 mrt. (c.q. 29 feb.)
mrt. I : 2-11 II : 12-21 III: 22-31	overige 9 maanden I : 1-10 II : 11-20 III: 21-30

De dagen 31 mei, 31 juli, 31 augustus, 31 oktober en 31 december (1 mrt. in een schrikkeljaar) werden weggelaten. In een niet-schrikkeljaar is feb. III: 20 feb. - 1 mrt., in een schrikkeljaar feb. III: 20 feb. - 29 feb.

Ondanks deze definitie bleven alle Δ's "eerlijk"; ter verduidelijking: ${}_1\Delta_1$ (bijv. 1 juni) heeft <u>niet</u> betrekking op het verschil in minimumtemperatuur van 1 juni en 30 mei, doch van 1 juni en 31 mei. De dekaden werden alle even lang gemaakt (door weglating van enige data), eerst nadat alle Δ's naar behoren berekend waren.

1. Karakterisering van de interdiurne veranderlijkheid

Hoe zullen wij een bewering, dat op het station S de tussendaagse veranderlijkheid (interdiurne variabiliteit I.V.) van de minimumtemperatuur in de maand april groter is dan in de maand juni in een getal (getallen) uitdrukken?

Als ${}_1\Delta_i = t_i - t_{i-1}$ de eerste orde differentie voorstelt, waarbij t_i de minimumtemperatuur op dag i (i = 1, 2, ... 30) is, zou men er mee kunnen bedoelen

$$\overline{|{}_1\Delta_i|}_{apr.} > \overline{|{}_1\Delta_i|}_{jun.},$$

waarbij $e \equiv \overline{|{}_1\Delta_i|}$ het veeljarig gemiddelde van ${}_1\Delta_i$ is, zonder op het teken van Δ te letten (26 jaren: 26 x 30 = 780 getallen). Deze e laat zich gemakkelijk machinaal berekenen. Wil men echter een bewering verifiëren, dat grote zowel als kleine (of misschien alleen de grote) Δ-waarden in april waarschijnlijker zijn dan in juni, dan is kennis van de volledige kansverdelingen (of een "staart" ervan) $f(\Delta)\Delta$ in de universa der in de maanden april en juni gelegen Δ-waarden nodig. Deze kansverdeling is zeker niet gekarakteriseerd door $\varepsilon \equiv \mathcal{E}|\Delta|$ = universumgemiddelde van Δ alleen (onder het april-universum kan worden verstaan het aantal Δ's in oneindig veel april-maanden), zelfs niet als mocht blijken, zoals inderdaad het geval is, dat het universumgemiddelde, geschreven $\mathcal{E}\Delta$, nul is. Ter verduidelijking: ook al zou de meting leveren, dat $\overline{|\Delta|}_{apr.} > \overline{|\Delta|}_{jun.}$ en wel over een zodanig groot aantal jaren, dat het zeer waarschijnlijk is, dat ook voor de twee Δ-universa de ongelijkheid geldt (d.i. $\mathcal{E}_{apr.} > \mathcal{E}_{jun.}$), dan behoeft daaruit nog niet noodzakelijk te volgen, dat voor iedere positieve Δ^* de kans op een $\Delta \geqslant \Delta^*$ in april groter is dan in juni. De gevolgtrekking is wèl juist als de twee kansverdelingen $f_{apr.}(\Delta)d\Delta$ en $f_{jun.}(\Delta)d\Delta$ symmetrisch rondom nul zijn. Doch ook zulk een kansverdeling is niet volledig door de waarde van ε, geschat door $e = \overline{|\Delta|}$, bepaald. De enige symmetrische, rondom nul gelegen, kansverdeling, waarvoor dit wel geldt, is de <u>normale</u>. "Alles" weten van de I.V. eist daarom een studie van de volledige kansverdeling van Δ. In het volgende wordt duidelijk gemaakt, dat beter dekaden dan maanden beschouwd konden worden. Mocht nu blijken, dat al deze 36 Δ-kansverdelingen (één voor jan. I, één voor jan. II, ... één voor dec. III) normaal zijn rondom nul, dan behoeven wij er nog slechts één grootheid van te kennen, waarna ze volkomen vastliggen. Het ligt voor de hand hiervoor de standaarddeviatie $\sigma_\Delta = [\int_{-\infty}^{\infty}\Delta^2 f(\Delta)d\Delta]^{\frac{1}{2}}$ te kiezen (het kwadraat heet variantie). Maar men kan ook de gemiddelde absolute Δ-waarde nemen, gedefinieerd door $\varepsilon_\Delta = \int_{-\infty}^{\infty}|\Delta|\, f(\Delta)d\Delta$, doch dit is betrekkelijk ongebruikelijk (in de moderne statistica). Tussen σ_Δ en ε_Δ bestaat de (slechts voor normale verdelingen, die een gemiddelde nul hebben, geldende) relatie van Cornu:

$$\varepsilon_\Delta = \sqrt{\frac{2}{\Pi}}\sigma_\Delta \approx \frac{4}{5}\sigma_\Delta$$

Sommige auteurs kiezen ter karakterisering nog een andere grootheid en wel de zgn. kwartiel-waarde of "probable error" β, gedefinieerd door:

$$\int_{-\beta}^{+\beta} f(\Delta)d\Delta = \tfrac{1}{2}$$

Men verwarre deze β niet met de mediaan τ, die gedefinieerd wordt door:

$$\int_{-\infty}^{\tau} f(\Delta)d\Delta = \int_{\tau}^{+\infty} f(\Delta)d\Delta = \tfrac{1}{2}$$

In iedere symmetrische verdeling zijn μ en τ aan elkaar gelijk; in ons geval zijn beide bovendien nul.

Voor een normale verdeling geldt: $\beta = 0.6745\,\sigma_\Delta$. Bijgevolg geldt:

$$\beta = 0.84\varepsilon_\Delta;\quad \varepsilon_\Delta = \sqrt{\frac{2}{\Pi}}\sigma_\Delta \approx \frac{4}{5}\sigma_\Delta \quad \text{en } \beta = 0.6745\,\sigma_\Delta$$

Voor een normale verdeling van Δ met gemiddelde nul is de kans op een $|\Delta| \leqslant \sigma_\Delta$ ongeveer 2/3, en op een $|\Delta| \leqslant 2\sigma_\Delta$ ongeveer 0.95 (zie overigens de tabellen voor de standaardnormale verdeling).

Wij kozen als karakteriserende grootheid voor I.V. de standaarddeviatie σ_Δ, geschat door s_Δ. De ε_Δ, geschat door e_Δ[1], is een hulpgrootheid, die aanvankelijk berekend werd met het oog op toepassing van de normalisatietoets van Geary. Overigens wordt de e zo goed als nooit bij statistische toetsen gebruikt, de s daarentegen zeer veel (zie verder hoofdstuk 9).

---

1) Sommige auteurs vermelden $\bar{\Delta}_-$ der negatieve Δ's en $\bar{\Delta}_+$ der positieve Δ's afzonderlijk en karakteriseren aldus de I.V. door twee grootheden. In onze tabellen noemen wij volledigheidshalve ook $\bar{\Delta}_+$ en $\bar{\Delta}_-$, in het bijzonder om te doen uitkomen, hoe weinig doorgaans deze twee gemiddelden verschillen (gevolg van de symmetrie-om-nul der Δ-verdelingen).

2. Inhoud van de machinaal geleverde lijsten van statistische rekengrootheden, samen met een korte toelichting

Wij noemen hier de lijsten betreffende de minimumtemperatuur t, welke voor elk der stations Gemert, Naaldwijk en Groningen werden samengesteld. Overeenkomstige lijsten werden voor de maximumtemperatuur en de 14-uurs relatieve vochtigheid geleverd.

A. De eerste lijst bevat voor elk der 12 x 3 = 36 dekaden van elk der 26 jaren van elke dag de t en de ${}_k\Delta_i$-waarden (k = 1, 2, ... 6; i = 1, 2, ... 10). Deze Δ's (alle in nieuwe ponskaarten geponst) werden afgedrukt in twee kolommen, resp. voor $\Delta < 0$ en voor $\Delta \geqslant 0$. Per individuele dekade werden de sommen $\Sigma^{10}_{1} t_i$, $\sum_{i=1}^{10} ({}_k\Delta_i < 0)$, $\sum_{i=1}^{10} ({}_k\Delta_i \geqslant 0)$, $\sum_{i=1}^{10} |{}_k\Delta_i|$ machinaal berekend en afgedrukt. Voorts werden de totaalsommen $S_{jk} = \sum_{1}^{10} \left\{ \sum_{i=1}^{10} |{}_k\Delta_i| \right\}_j$, k = 1, ... 6 berekend en afgedrukt voor elk der drie deelperioden, die met j = 1, 2, 3 zijn aangeduid en voor elk der 36 dekaden. Deling van deze S_{jk} door 100 (of 60) geeft $\overline{|{}_k\Delta|}$, d.i. de gemiddelde waarde van de absolute k-de orde differentie voor een dekade, gemiddeld over de gehele deelperiode. Evenzo werden de totaalsommen $M_j = \sum_{1}^{10} \left\{ \sum_{i=1}^{10} t_i \right\}_j$, j = 1, 2, 3 afgedrukt. Deling van deze sommen door 100 (60) geeft de gemiddelde waarde $\bar{t}$ van de dagelijkse minimumtemperatuur over de gehele periode.

B. De tweede lijst bevat voor elk der 36 dekaden, gesommeerd over de gehele 26-jarige periode, de 6 totale sommen $\Sigma({}_k\Delta < 0)$ en de 6 totale sommen $\Sigma({}_k\Delta \geqslant 0)$; k = 1, ... 6. Aldus kon men $\bar{\Delta} = \frac{1}{260} \Sigma\, {}_k\Delta$ berekenen, d.i. de gemiddelde waarde van de k-de orde differentie in een dekade in kwestie, gemiddeld over alle 26 jaren. Voor de eerste orde k = 1 (de eigenlijke interdiurne verandering) is $\bar{\Delta}$ ongveer nul. Het is gemakkelijk in te zien, dat ongeveer geldt:
${}_k\bar{\Delta} = k\,{}_1\bar{\Delta}$.

C. De derde lijst bevat voor iedere dag van iedere dekade van elk der 26 jaren het verschil $d_i = t_i - \bar{t}$. Het is duidelijk, dat de machine daartoe eerst $\bar{t}$ moest berekenen (zie onder A). De d's werden in twee kolommen afgedrukt, nl. voor $d < 0$ en voor $d \geqslant 0$. Voor iedere dekade leverde de machine de som $N_j = \Sigma^{10} |d_i|$; j = 1, 2, ... 26. Als $N = \sum_{1}^{26} N_j$, dan stelt $\frac{1}{260} N$ de ${}^1\overline{|d|}$ voor, d.i. het 26-jarige gemiddelde van de absolute d. Dit gemiddelde is natuurlijk positief, terwijl $\overline{d}$ zelf nul is, louter vanwege de definitie van $\overline{t}$. Wij lieten deze $\overline{|d|}$ berekenen om er, via de formule $\sigma = \varepsilon\sqrt{\frac{\pi}{2}}$, een schatting mee te maken van de standaarddeviatie in de kansverdeling der t-waarden. Deze schatting is slechts correct als deze verdeling normaal is. Blijkens door anderen uitgevoerde onderzoekingen is dit in goede benadering het geval. Aldus kwam voor elk der 36 dekaden een schatting van σ tot stand.

D. De <u>vierde lijst</u> bevat de distributieve en cumulatieve frequentieverdeling van de $|\Delta|$-waarden (alleen eerste orde), waarbij $|\Delta|$ in klassen 0.0-0.9; 1.0-1.9 .... ^{o}C werd ingedeeld en wel voor

de 100 dagen der dekaden jan. I					uit deelperiode			j = 1
de 100	"	"	"		"	"	"	j = 2
de 60	"	"	"		"	"	"	j = 3
de 100	"	"	"	jan. II	"		"	j = 1

enz.

t/m de 60 dagen der dekaden dec. III uit deelperiode j = 3; derhalve 12 x 3 frequentieverdelingen (berustende op 100 of 60 getallen) der $|\Delta|$-waarden.

E. De <u>vijfde lijst</u> bevat de waarden van enkele hulpgrootheden, die nodig zijn ter berekening van de steekproefvarianties $s^2_{|\Delta|}$ der $|\Delta|$-waarden in elk der onder D bedoelde frequentieverdelingen. Uit deze werden weer $s^2_{|\Delta|}$-waarden voor de gehele 26-jarige periode berekend (36 stuks) en uit deze weer werden afgeleid de s^2_{Δ}-waarden der Δ's zelf. (let wel: $s^2_{|\Delta|} \neq s^2_{\Delta}$!). Voor de relatieve vochtigheid werden de frequentieverdelingen der Δ's zelf, alsmede de s^2_{Δ}-waarden daarvan machinaal berekend en afgedrukt.

3. Een onderzoek naar de normaliteit van de kansverdeling der differenties van de eerste orde

De constructie van een nomogram als waarvan in hoofdstuk 10 van dit rapport sprake zal zijn, is slechts mogelijk (althans zonder bijzondere transformaties) als de kansverdelingen der differenties van de eerste orde normaal zijn en wel rondom nul, of als zij toelaatbaar weinig van normale verschillen. Het was dus zaak aan de hand van de frequentieverdelingen deze voorwaarde te toetsen.

Voor de dagelijkse maximum- en de dagelijkse minimumtemperatuur werd voor elk der drie genoemde stations voor elk der 36 dekaden een frequentieverdeling der $|\Delta|$-waarden machinaal gemaakt. Deze verdelingen berusten voor Naaldwijk en Gemert elk op 260 $|\Delta|$-waarden, voor Groningen op 200 waarden. Uit deze 72 frequentieverdelingen werden de frequentieverdelingen der Δ-waarden zelf (dus met inachtneming van de tekens) gemaakt, hetgeen slechts geoorloofd is als aangenomen mag worden, dat elke kansverdeling symmetrisch rondom nul gelegen is. Enkele steekproeven leerden, dat aan deze voorwaarde voldoende goed voldaan is. De aldus verkregen frequentieverdelingen werden cumulatief op lineair waarschijnlijkheidspapier uitgezet. De punten bleken doorgaans vrij goed "lineair te liggen"; dit duidt erop, dat de kansverdelingen zeer waarschijnlijk normaal zijn. Bovendien werd geheel machinaal, voor elk der genoemde 72 frequentieverdelingen de variantie $s^2_{|\Delta|}$ der $|\Delta|$-waarden berekend, waaruit weer de variantie s^2_Δ der Δ-waarden kon worden afgeleid (§9). In elke grafische voorstelling werd op basis van deze s_Δ de "gaussische" rechte getekend. In het algemeen blijkt deze rechte fraai tussen de uitgezette punten door te lopen, hetgeen weer wijst op het normaalverdeeld-zijn van de kansverdelingen.

In het geval van de 14-uurs relatieve vochtigheid U werd, weer machinaal, de frequentieverdeling der Δ-waarden zelf en rechtstreeks de s^2_Δ berekend. In totaal gaat het dus om 3 x 3 x 36 = 324 frequentieverdelingen (en 324 grafische voorstellingen), nl. vanwege 3 stations, 3 elementen en 36 dekaden. Om tijd te sparen werd ervan afgezien op al deze frequentieverdelingen de χ^2-toets toe te passen om te beoordelen in hoeverre de kansverdelingen normaal kunnen zijn. Aanvankelijk werd de Geary-toets, die veel minder rekenwerk vergt, toegepast. Deze toets beoordeelt het quotiënt $q \equiv \overline{|\Delta - \bar{\Delta}|} / s_\Delta$ of nog eenvoudiger $\overline{|\Delta|}/s_\Delta$. In een normale verdeling is $\varepsilon/\sigma \cong \frac{4}{5}$ (exact: $\sqrt{2/\Pi}$). De $\overline{|\Delta|}$ is een schatting van ε; de s_Δ is een schatting van σ. Door toeval (steekproefeffecten) kan q van $\frac{4}{5}$ verschillen. Indien de gemeten q toelaatbaar weinig van $\frac{4}{5}$ verschilt (dit te beoordelen m.b.v. bepaalde tabellen), dan mag de kansverdeling als normaal beschouwd worden (natuurlijk behoudens een kleine, bekende kans, dat toch de uitspraak fout is). Spoedig bleek, dat deze (overigens toch weinig door statistici gebruikte) toets te zwak was. De beoordeling of Δ normaal verdeeld is, geschiedde in het vervolg slechts grafisch-visueel. Aldus bleken alle Δ-kansverdelingen voldoende goed normaal.[1]

Hier moge de volgende opmerking toegevoegd worden. Het ligt in de bedoeling met behulp van deze drie frequentieverdelingen (denken wij, als voorbeeld aan jan. I, drie deelperioden) de homogeniteit van de gehele reeks te onderzoeken (trend? klimaatsverandering? toenemende s_Δ?). Dit onderzoek kon nog niet geschieden. Voorlopig werden voor elk der 36 dekaden de bedoelde drie frequentieverdelingen samengevoegd. Aldus ontstonden er 36 frequentieverdelingen voor elk element en elk station. Toch behoort elk dezer frequentieverdelingen bij een universum, dat in zekere zin principieel statistisch inhomogeen is of kan zijn. Ter verduidelijking: beschouw de t_1, t_2 ..... t_{10} der dagen in de eerste dekade van april. Er is sprake van een universum van t_1-waarden, met gemiddelde $\mu_1 \equiv \mathcal{E} t_1$. Zo is er ook een μ_2, ... μ_{10}. Gesteld: $\mu_1 < \mu_2$ ..... μ_{10} (men zou dit kunnen toetsen). Er is ook een universum van Δ_2-waarden, met $\Delta_2 = t_2 - t_1$, met gemiddelde $\beta_2 = \mu_2 - \mu_1 > 0$. Zo ook een $\beta_3 = \mu_3 - \mu_2 > 0$ ..... $\beta_{10} = \mu_{10} - \mu_9 > 0$. Het kan zijn (maar noodzakelijk is het niet), dat $\beta_1 = \beta_2 = \ldots\ldots \beta_{10}$, hetgeen betekent, dat de μ's een rekenkundige reeks vormen (de jaarlijkse gang der μ's is dan lineair). Zelfs als dit zo zou zijn dan zijn deze 10 universa niet per se identiek, want ze kunnen nog in hun varianties verschillen. Een samenvoeging heeft eigenlijk eerst dan zin als de universa identieke kansverdelingen hebben. Als men daarvan niet volkomen zeker is, laat men uitspraken, die gebaseerd zijn op de frequentieverdeling van alle 260 Δ's in de 26 dekaden april I gemakshalve maar gelden voor elk der dagen in deze dekade afzonderlijk. Hoe het met deze details gesteld is, werd niet in extenso onderzocht. Wij willen met het bovenstaande doen uitkomen, dat $\beta_i = 0$ geen noodzakelijkheid is. Echter bleek deze onderstelling toelaatbaar, het verdere rekenwerk (in het bijzonder de constructie van een nomogram) werd er door vergemakkelijkt.

---

1) Dit mag wellicht iets verbazen t.a.v. de 14-uurs relatieve vochtigheid. Een grootheid, die aan tenminste één zijde begrensd is bij een eindige waarde (de U is dit zelfs tweezijdig: $0 \leqslant U \leqslant 100$) kan, exact beschouwd, nooit normaal verdeeld zijn. In onze studie gaat het echter niet om U zelf, maar om de eendaagse veranderingen $_1\Delta$. Deze zijn al gauw wel in goede benadering normaal verdeeld, zoals in ons onderzoek, ook al is de initiaalvariabele het niet.

4. Over de noodzakelijkheid van kennis van de persistentie bij onderzoekingen van de jaarlijkse gang van de interdiurne variabiliteit en van de verschillen tussen stations

4.0 Inleiding

Als men wil onderzoeken of het verschil tussen s_Δ^2 bijv. voor april II en s_Δ^2 bijv. voor december III te Gemert voor de minimum temperatuur statistisch significant is (d.i. of de twee universa der Δ's verschillende varianties hebben) moeten wij wel bedenken, dat de toets kennis eist van de ware "aantallen graden van vrijheid", waarop elk der steekproefvarianties berust. De nominale aantallen zijn weliswaar 260 (= 26 x 10), doch de statistisch effectieve aantallen zijn groter. Hoeveel groter? Het antwoord kan eerst gegeven worden als de persistentie- of autocorrelatie-coëfficiënten der orden 1, 2, ... in de Δ-reeks bekend zijn en deze weer kunnen pas berekend worden als die in de t-reeks bekend zijn. Wij hebben hier met een klimatologisch betrekkelijk zeldzaam verschijnsel te doen, t.w. met een <u>negatieve</u> persistentie, die, voor zover mij bekend is, alleen in tijdreeksen van afgeleide grootheden optreedt, en dus o.a. bij de eerste orde differentie Δ, die een afgeleide grootheid is: $\Delta_1 = t_1 - t_0$; $\Delta_2 = t_2 - t_1$ etc. (men ziet hoe Δ_1 en Δ_2 de "t_1 gemeen hebben"). Reden te meer om wat dieper op de statistische details in te gaan.

4.1 De persistentie-coëfficiënten

Weer zullen t_0, t_1, t_2 .... de dagelijkse minima der dagen 0, 1, 2 ..... zijn. Wij beginnen direct met de onderstelling, dat de t-reeks een positieve persistentie bezit, die beschreven kan worden met de wet $\rho_k = \rho^k$. Hierbij is ρ_k de k^{de}-orde autocorrelatie-coëfficiënt tussen t_i en t_{i+k} en ρ de autocorrelatie-coëfficiënt van de orde 1. Uit de t-reeks volgt de reeks der differenties van de eerste orde ${}_1\Delta_i$: $\Delta_1 = t_1 - t_0$; $\Delta_2 = t_2 - t_1$ ..... Ook deze reeks bezit persistentie. Laten de autocorrelatie-coëfficiënten der orden 1, 2 ..... voor deze reeks γ_1, γ_2 ..... zijn. Wij gaan γ_k in ρ uitdrukken. Ter vereenvoudiging van de berekening voeren we de volgende onderstellingen in:

a) over een dekade verandert ρ niet of niet noemenswaardig.

b) binnen een dekade mogen wij $\sigma^2(t_i)$ (i = 1, 2 ..... 10) onafhankelijk van i onderstellen, d.i. $\sigma(t_1) = \sigma(t_2) = \dots (t_{10}) = \sigma$.

Dit betekent: de minimumtemperatuur op 1 jan. spreidt evenveel als die op 2 of 3 of 4 .... of 10 januari. Deze onderstellingen zullen voor een dekade zeker nog wel geoorloofd zijn.

Dan is per definitie:

$$\gamma_k = \frac{\varepsilon(\Delta_i.\Delta_{i+k}) - \varepsilon(\Delta_i).\varepsilon(\Delta_{i+k})}{\sqrt{\{\varepsilon(\Delta_i^2) - \varepsilon^2(\Delta_i)\}.\{\varepsilon(\Delta_{i+k}^2) - \varepsilon^2(\Delta_{i+k})\}}}$$

Met gebruikmaking van

$${}_1\Delta_j = t_j - t_{j-1}$$

herleiden we als volgt:

$$\gamma_k = \frac{\{\varepsilon(t_i t_{i+k}) - \varepsilon(t_i).\varepsilon(t_{i+k})\} + \{\varepsilon(t_{i-1} t_{i+k-1}) - \varepsilon(t_{i-1})\,\varepsilon(t_{i+k-1})\} - \{\varepsilon(t_{i-1}.t_{i+k}) - \varepsilon(t_{i-1})\,\varepsilon(t_{i+k})\} - \{\varepsilon(t_i.t_{i+k-1}) - \varepsilon(t_i)\,\varepsilon(t_{i+k-1})\}}{\sqrt{[\{\varepsilon(t_i^2) - \varepsilon^2(t_i)\} - 2\{\varepsilon(t_i t_{i-1}) - \varepsilon(t_i)\,\varepsilon(t_{i-1})\} + \{\varepsilon(t_{i-1}^2) - \varepsilon^2(t_{i-1})\}].\,[\;]_{i+k}}} =$$

$$= \frac{2\rho^k\sigma^2 - \rho^{k+1}\sigma^2 - \rho^{k-1}\sigma^2}{\sqrt{(\rho^0\sigma^2 - 2\rho\sigma^2 + \rho^0\sigma^2)\,(\rho^0\sigma^2 - 2\rho\sigma^2 + \rho^0\sigma^2)}} = \frac{-(1-\rho)^2\,\rho^{k-1}\sigma^2}{2(1-\rho)\,\sigma^2} =$$

$$= -\frac{1}{2}\,(1-\rho)\,\rho^{k-1}$$

Algemeen geldt: $\boxed{\gamma_k = \dfrac{-\rho_{k-1} + 2\rho_k - \rho_{k+1}}{2(1-\rho)}}$

De formule is zo simpel alleen indien de persistentie der t-waarden met de eenvoudige wet $\rho_k = \rho^k$ beschreven kan worden. Het is dus wel zaak te toetsen of deze wet geldt (zie paragraaf 6). Ter controle: als de t-reeks correlatie-vrij is ($\rho = 0$), zijn toch de Δ's nog negatief gepersisteerd en wel volgens $\gamma_1^* = -\frac{1}{2}$; $\gamma_2^* = \gamma_3^* = \ldots. = 0$. Nu de t's positief gepersisteerd zijn wordt γ_1 minder negatief, want $-\frac{1}{2} < -\frac{1}{2}(1-\rho)$ en breekt de negatieve persistentie eigenlijk nergens af, want $\gamma_k = -\frac{1}{2}\rho^{k-1}(1-\rho)$ gaat met toenemende k asymptotisch naar nul. Wel dient bedacht te worden, dat wij k niet onbeperkt kunnen laten toenemen zonder aan de vermelde onderstellingen geweld te doen.

Hier volgen enige getallen ter illustratie

als $\rho_k = \rho^k$ geldt, dan is						
	γ_1	γ_2	γ_3	γ_4	γ_5	enz.
bij ρ = 0	$-\frac{1}{2}$	0	0	0	0	0
= 0.1	- 0.450	- 0.045	- 0.004	- 0.0004	- 0.00004	afnemend met factor ρ
= 0.5	- 0.250	- 0.125	- 0.062	- 0.0312	- 0.01562	" " " "
= 0.9	- 0.050	- 0.045	- 0.040	- 0.0364	- 0.03280	" " " "

N.B.: Het gaat in het onderzoek, waarvan dit rapport verslag geeft, over <u>interdiurne</u> veranderingen. Wij kunnen ons voorstellen, dat in andere onderzoekingen ook twee-, driedaagse veranderingen van belang zouden kunnen zijn. Beschouwen wij bijv. de tussen-twee-daagse veranderingen: ${}_2\Delta_j = t_j - t_{j-2}$; $j = 1, 2 \ldots$ Ook deze reeks bezit autocorrelatie, gekarakteriseerd door autocorrelatie-coëfficiënten δ_k van de orden k (1, 2 ...). Ook deze laten zich uitdrukken in de ρ_k's van de t-reeks. Wij geven direct de uitkomst

$$\delta_k = \frac{-\rho_{k-2} + 2\rho_k - \rho_{k+2}}{2(1-\rho_2)}$$

Substitueren we $\rho_k = \rho^k$ (zuivere persistentie), dan komt er $\delta_k = -\frac{1}{2}(1-\rho^2)\rho^{k-2}$ als $k \geqslant 2$ en $\delta_1 = +\frac{1}{2}\rho$. In deze reeks zijn dus alle autocorrelatie-coëfficiënten negatief, m.u.v. de eerste.

Vergelijkende met γ_k, blijkt, voor elke $k \geqslant 2$, de $\delta_k < \gamma_k$ te zijn. De ${}_2\Delta$-reeks bezit derhalve minder persistentie dan de ${}_1\Delta$-reeks.

Algemeen: de reeks der differenties der m^de^ orde, p.d. ${}_m\Delta_j = t_j - t_{j-m}$, $j = 1, 2 \ldots.$, heeft een k-de orde autocorrelatie-coëfficiënt $\boxed{\rho_k({}_m\Delta) = -\frac{1}{2}(1-\rho^m)\rho^{k-m},}$ althans als $k \geqslant m$. Verdere mathematische details gaan wij voorbij.

4.2 <u>De effectieve getallen</u>

Beschouw n successieve Δ's : $\Delta_1, \Delta_2 \ldots. \Delta_n$, met gemiddelde $g_n = \frac{1}{n}(\Delta_1 + \ldots.. \Delta_n)$.*) Als deze Δ's onafhankelijk zijn (d.i. als $\gamma_1 = \gamma_2 = \ldots.. = 0$) dan heeft deze g_n een standaarddeviatie $\tau/\sqrt{n}$, wanneer iedere Δ_i dezelfde variantie τ^2 heeft (de bekende $\sqrt{n}$ - wet).

*) $g_n = \frac{1}{n}(t_n - t_0) = \frac{1}{n}\,{}_n\Delta_n$

Echter zijn de Δ's zeker wèl afhankelijk, d.i. er bestaat persistentie, t.w. volgens $\gamma_k = -\frac{1}{2}\rho^{k-1}(1-\rho)$, indien in de initiaalreeks zelf de autocorrelatie gehoorzaamt aan de relatie $\rho_k = \rho^k$. De variantie van g_n wordt dan

$$\sigma^2(g_n) = \frac{\tau^2}{n}\left[1 + 2\left\{\frac{n-1}{n}\gamma_1 + \frac{n-2}{n}\gamma_2 + \dots \frac{1}{n}\gamma_{n-1}\right\}\right],$$ geschreven $\frac{\tau^2}{n_e}$ [1], met n_e = effectief aantal = n/w als $w = 1 + 2\left\{\frac{n-1}{n}\gamma_1 + \frac{n-2}{n}\gamma_2 + \dots \frac{1}{n}\gamma_{n-1}\right\}$

dus $n_e < n$

Meestal is $w > 1$, omdat men doorgaans, in het geval van autocorrelatie, met <u>positieve</u> autocorrelatie te doen heeft. In ons geval (successieve differenties) echter is de persistentie negatief, d.w.z. $\gamma_k < 0$ en dus $0 < w < 1$ en dus $\underline{n_e > n}$.
Wij vinden bij uitwerking

$$w = 1 - \frac{1-\rho}{n}[(n-1) + (n-2)\rho + \dots \rho^{n-2}] = \frac{1}{n}\frac{1-\rho^{\cancel{2}\,n}}{1-\rho}$$

en zien: voor een correlatie-vrije initiaal-reeks is $\rho = 0$; $w = \frac{1}{n}$ en $n_e = n^2$. Is er positieve correlatie, dan $\rho_k = \rho^k$ en

$$w = \frac{1}{n}\frac{1-\rho^n}{1-\rho}; \quad n_e = n^2\frac{1-\rho}{1-\rho^n}; \quad \boxed{n \leqslant \frac{1-\rho}{1-\rho^n}n^2 < n^2}$$

Hoe kleiner de $\rho(\to 0)$, hoe dichter ligt n_e bij n^2; altijd $n_e < n^2$; hoe groter de $\rho(\to 1)$, hoe dichter ligt n_e bij n; altijd $n_e > n$.

In ons geval beschouwen wij bijv. alle dekaden jan. I, met in totaal 26 x 10 differenties Δ. Dit aantal is effectief equivalent met

$$26\left[10^2\frac{1-\rho}{1-\rho^{10}}\right] \quad \text{en niet met} \quad (260)^2\frac{1-\rho}{1-\rho^{260}}$$ [2]

---

1) Zie o.a. bij Bartels: Gesetz und Zufall in der Geophysik; Naturwissenschaften <u>31</u> 421, 1943.

2) Om het verschil te doen uitkomen volgen hier enkele getallen:

a) Het totale aantal, 260 stuks, interdiurne differenties, die aanwezig zijn in 26 dekaden jan. I, is, als $\rho_k = \rho^k$ geldt in de initiaal-reeks zelf, met $\rho = 0.73$, effectief equivalent met $26\left[10^2\frac{1-\rho}{1-\rho^{10}}\right] = 730$, zegge 2 jaren.

b) Het totale aantal, 260 stuks, interdiurne differenties, die aanwezig zijn in een chronologische ononderbroken reeks van 260 dagen, is, als weer $\rho_k = \rho^k$ met $\rho = 0.73$ zou gelden, effectief equivalent met $(260)^2\frac{1-\rho}{1-\rho^{260}} \approx (260)^2(1-\rho)$ d.i. met ong. 50 jaren.

4.3 Nog enige toelichting bij het begrip "effectief aantal"

Als wij zeggen, dat het nominale aantal n der Δ's equivalent is met effectief n_e bedoelen wij daarmee per definitie, dat zich een groep van n chronologisch successieve Δ's "gedraagt" als een groep van n_e Δ's die onderling zo ver in de tijdas van elkaar verwijderd zijn, dat ze onafhankelijk zijn (of mogen heten). Met "gedraagt zich" wordt dan bedoeld: de gemiddelden van beide groepen hebben gelijke varianties. Wij weten, dat het gemiddelde van de tweede groep een variantie σ^2/n_e heeft. In het algemeen genomen leidt in een tijdreeks positieve autocorrelatie tot $1 < n_e < n$ en negatieve tot $n_e > n$ (negatieve in de betekenis van: iedere persistentie-coëfficiënt is negatief). In ons geval: $n < n_e < n^2$.

Dit effectieve aantal moeten wij kennen, bijv. bij een onderzoek van de statistische realiteit van het verschil tussen twee steekproefvarianties s_1^2 en s_2^2. De bedoeling is om te onderzoeken of, indien $s_1^2 < s_2^2$, ook tot $\sigma_1^2 < \sigma_2^2$ (in de universa) mag worden besloten. Hoe groter nu de effectieve aantallen n_{e1} en n_{e2}, hoe minder waarschijnlijk is het, dat een gegeven verschil tussen s_1^2 en s_2^2 louter toevallig is, d.i. met de mogelijkheid $\sigma_1^2 = \sigma_2^2$ niet in strijd is.

Enkele getallen-voorbeelden

ρ	$\frac{1-\rho}{1-\rho^{10}}$	Het nominale aantal 10 is equivalent met het effectieve aantal n_e
1.00	0.100	10.0
0.95	0.124	12.4
0.90	0.154	15.4
0.70	0.309	30.9
0.50	0.50	50
0.30	0.70	70
0.10	0.90	90
0.05	0.95	95
0.00	1.00	100

5. De berekening van de autocorrelatie-coëfficiënten r_k

In het vorige hoofdstuk is uiteengezet, dat het noodzakelijk is de persistentie-coëfficiënten γ der chronologisch successieve Δ-waarden te kennen. Dit eist kennis van de autocorrelatie-coëfficiënten in de t-reeks (als wij, bij wijze van voorbeeld, even aan de minimumtemperaturen denken). Vervolgens zou het van groot voordeel zijn als mocht blijken dat de laatste zouden gehoorzamen aan de wet $\rho_k = \rho^k$.

Wij hebben in het hierna volgende een procedure uitgewerkt, die via de differenties ${}_k\Delta_i$ (i = 1, 2 .... 10; k = 1, 2 .... 6), zonder met kwadraten te werken (dit heeft machine-technische voordelen), toch redelijke schattingen r_k van de ρ_k's verschaft. Er wordt vervolgens nagegaan of de metingen r_1, r_2, r_3, r_4, r_5, r_6 "redelijk goed" aan de formule $\rho_k = \rho^k$ gehoorzamen.

De volgende stelling geldt: als x en y kansverdelingen hebben met varianties σ_1^2 en σ_2^2 en gecorreleerd zijn met een correlatie-coëfficiënt ρ, dan volgt het verschil $\Delta = x - y$ een kansverdeling met variantie $\sigma^2(\Delta) = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1.\sigma_2$. De stelling is onafhankelijk van de typen kansverdelingen. In ons geval komen de x-waarden uit de tijdreeks der temperaturen t_1, t_2 ....; de y-waarden uit dezelfde, doch één element verschoven, reeks: t_2, t_3 ..... Dan is $\Delta = t_2 - t_1$; $t_3 - t_2$ ... en $\sigma^2(\Delta) = 2\sigma^2(1 - \rho)$, als de t_i-waarden een variantie σ^2 hebben. Bij gevolg is

$$\boxed{\rho = 1 - \frac{1}{2}\frac{\sigma^2(\Delta)}{\sigma^2}} \qquad (1)$$

Deze ρ is de autocorrelatie-coëfficiënt van de eerste orde in de t-reeks.

Als wij $\sigma(\Delta)$ en σ kunnen berekenen, is aldus ρ bekend. Voor een normale verdeling geldt (zie hoofdstuk 1):

$$\varepsilon = \sqrt{\frac{2}{\pi}}\,\sigma \approx \frac{4}{5}\sigma$$

Indien dus zowel t als Δ een normale verdeling volgt (als t het doet, doet Δ het automatisch), dan mag $\sigma(\Delta)/\sigma$ vervangen worden door $\varepsilon(\Delta)/\varepsilon$. Deze ε's zijn machine-technisch veel eenvoudiger te berekenen dan de varianties $\sigma^2(\Delta)$ en σ^2.

Wij moeten wel bedenken, dat wij met steekproeven te doen hebben. Zo vormt jan. I 1931 de eerste steekproef; jan. I 1932 de tweede ... jan. I 1956 de 26ste. Alle steekproeven bestaan uit 10 elementen. Zij leveren stuk voor stuk (af te lezen van door de machine samengestelde lijsten) $\bar{t}$, $\bar{\Delta}$, $\overline{|t - \bar{t}|}$ en $\overline{|\Delta - \bar{\Delta}|}$. Aldus is uit te rekenen

$$r^{(1)} = 1 - \frac{1}{2}\left[\frac{\overline{|\Delta - \bar{\Delta}|}}{\overline{|t - \bar{t}|}}\right]^2 = 1 - \frac{1}{2}\left[\frac{\Sigma|\Delta - \bar{\Delta}|}{\Sigma|t - \bar{t}|}\right]^2 \qquad (2)$$

/

als een eerste schatting van de eerste-orde-correlatie-coëfficiënt ρ. De dekade jan. I 1932 levert evenzo een tweede schatting, $r^{(2)}$, en aldus $r^{(3)}$ ...., $r^{(26)}$. Nu behoeft niet in elke kleine steekproef (en 10 elementen per steekproef is inderdaad weinig, ook al is deze genomen uit een normale verdeling) de verhouding van e (schatting van ε) tot s (schatting van σ) precies $\sqrt{\frac{2}{\Pi}}$ te zijn. De verhouding zelf volgt (vanwege het steekproef-effect) een kansverdeling, waarin de spreiding met de grootte van de steekproef samenhangt. Bij gevolg behoeft niet per steekproef de verhouding $s(\Delta)/s(t)$ precies gelijk te zijn aan $e(\Delta)/e(t)$. Het gevolg daarvan is weer, dat de schatting (2) zeer grof, zelfs ontoelaatbaar grof (nl. $r < -1$ of $r > +1$) kan zijn. Inderdaad bleek dit zo dikwijls te gebeuren, dat wij besloten de berekening van $r^{(i)}$ per individuele dekade achterwege te laten. Daarmede verviel ook de mogelijkheid om te onderzoeken hoe sterk de autocorrelatie van maand tot maand of van jaar tot jaar veranderen kan. De kleinheid van de steekproef was dus de oorzaak van deze tegenslag; met een maand of met een seizoen, i.p.v. met een dekade, was een en ander misschien wel gelukt, maar er was nu eenmaal besloten alle bewerkingen per dekade te laten verrichten (natuurlijk is er nog de andere oplossing: de r volgens de juiste methode, via de s-waarden, te berekenen, maar dit zou machine-technisch moeilijker geweest zijn).

Zo werd dus besloten niet voor elk der 36 dekaden de $r^{(1)}$, $r^{(2)}$ ... $r^{(36)}$ te berekenen, gevolgd door middelen: $r = \frac{1}{36}\Sigma_1^{36} r^{(i)}$, maar om te werken met

$$r = 1 - \frac{1}{2}\left[\frac{\Sigma|\Delta|}{\Sigma|t - \bar{t}|}\right]^2 \qquad (3)$$

De sommatie heeft hier betrekking op bijv. alle 26 dekaden jan. I tezamen. De $\bar{t}$ is thans bedoeld als het gemiddelde van alle minimumtemperaturen op alle 26 x 10 dagen in deze dekaden. In de teller zou eigenlijk $\Sigma|\Delta - \bar{\Delta}|$ hebben moeten staan (zoals in (2)), waarin $\bar{\Delta}$ de gemiddelde interdiurne verandering is over alle 26 x 10 in de 26 dekaden jan. I aanwezige dagen. Aangezien echter bleek, dat deze $\bar{\Delta}$ meestal zeer dichtbij nul gelegen is (negatief of positief; hooguit 0.3°C) en proefberekeningen leerden, dat het weinig verschil maakte t.a.v. r of wij de ene of de andere som in de teller gebruiken, geven wij de voorkeur aan formule 3, die machine-technisch eenvoudiger is. [1]

[1] Natuurlijk kunnen nooit alle 36 $\bar{\Delta}$-waarden tegelijk nul zijn; dit zou nl. zeer moeilijk te rijmen zijn met de toch zeker reële gemiddelde jaarlijkse gang der dagelijkse minima.

Een getallen-voorbeeld: Gemert, dagelijkse minimumtemperatuur; 26 dekaden jan. I; $\bar{t} = -0.6^{o}C$; $\Sigma|t - \bar{t}| = 1007.3$; $\frac{1}{260}\Sigma|t - \bar{t}| = 3,88$; $\bar{\Delta} = -0.1^{o}C$ (de drie deelperioden afzonderlijk: -0.2; -0.2; $+0.1$); $\Sigma|\Delta| = 592.9$; $\frac{1}{260}\Sigma|\Delta| = 2.28$; dus $r = 1 - \frac{1}{2}\left(\frac{592.9}{1007.3}\right)^2 = 0.83$, aangeduid met r_1, omdat het hier de eerste-orde correlatie-coëfficiënt betreft (gebaseerd op de eerste orde differenties ${}_1\Delta$).

Als in (3) met de tweede orde differenties gewerkt wordt, moet er $\Sigma|{}_2\Delta|$ gesubstitueerd worden. Dan $r_2 = 0.38$ (jan. I). Evenzo $r_3 = 0.50$; $r_4 = 0.38$; $r_5 = 0.27$ en $r_6 = 0.19$.

De fout die gemaakt wordt door met $\Sigma|{}_j\Delta|$ en niet met $\Sigma|{}_j\Delta - {}_j\bar{\Delta}|$ te werken, wordt groter met grotere j, o.a. doordat dan ${}_j\bar{\Delta}$ meer van nul verschillen gaat, maar de fout leek ons steeds toelaatbaar.

Het bovenstaande zal duidelijk gemaakt hebben, dat het voor deze r-berekeningen (3 stations; 36 dekaden; 4 elementen; 6 orden) nodig was, dat de machine direct de waarden van de grootheden $\Sigma|\Delta|$ en $\Sigma|t - \bar{t}|$ afdrukte (zie hoofdstuk 2).

6. De aanpassing met $\rho_k = \rho^k$ aan de zes correlatie-coëfficiënten $r_1, \dots r_6$

In hoofdstuk 5 is uiteengezet hoe voor iedere dekade voor elk der drie stations en voor elk der drie weerselementen de autocorrelatie-coëfficiënten van de orden 1 t/m 6 r_1, r_2 .... r_6 berekend werden ($36 \times 3 \times 3 = 324$ stellen van zes stuks).

Dat wij juist bij de orde 6 afbraken heeft hoofdzakelijk een machinetechnische reden; het bleek mogelijk "in één slag" de 6 differenties te bepalen, op de lijst af te drukken en op ponskaarten te brengen. Bovendien meenden wij dat de persistentie wel niet veel verder zou grijpen, d.w.z. dat het bij een niet bijzonder groot materiaal toch spoedig moeilijk zou worden om hoge-orde autocorrelatie-coëfficiënten significant van nul te onderkennen. Wij wisten, dat de dagsom neerslag significante positieve r_1, r_2, r_3-waarden gaf, en wij konden wel vermoeden, dat de persistentie in de temperatuur èn groter èn uitgestrekter zou zijn. Bij het huidige materiaal is het zeker de moeite waard de correlatie-coëfficiënten met een eenvoudige formule $\rho_i = f(i)$ te beschrijven. Het verdere rekenwerk wordt dan allicht eenvoudiger. Bij pogingen daartoe moeten wij met de betrekkelijk grote onnauwkeurigheid der r_i-waarden (hoe groter de i, hoe minder de nauwkeurigheid) rekening houden. Zo moeten wij er ons niet over verbazen, dat, ook al zou $\rho_1 > \rho_2 > \rho_3 \dots$ (in de universa) zijn, de steekproef toch bijv. $r_1 > r_2 > r_3$; $r_3 < r_4 > r_5$ etc. leveren kan. Ook kunnen negatieve r-waarden gevonden worden, terwijl toch elke $\rho_i > 0$ is. Wij hebben toen direct de relatie $\rho_k = \rho^k$ aangepast.

De "beste" ρ, passend bij een gegeven stel $r_1, \ldots. r_6$, zou gedefinieerd kunnen worden als die ρ, welke de som $\Sigma_1^6(r_k - \rho^k)^2$ minimaal maakt. Daartoe zou ρ uit $\partial\Sigma/\partial\rho = 0$ moeten worden opgelost, een elfdegraadsvergelijking. Het werk daaraan verbonden ($3 \times 3 \times 36 = 324$ keer) is zodanig groot, dat wij de volgende minder correcte weg kozen. Noem $y_i = \lg r_i$; $\eta_i = \lg \rho_i = i\eta$. Maak niet $\Sigma_1^6(r_i - \rho^i)^2$ minimaal, doch $T = \Sigma(y_i - \eta_i)^2$. Los dus op $\partial T/\partial\eta = 0$. Er komt:

$$\eta = \frac{y_1 + 2y_2 + 3y_3 + \ldots. 6y_6}{1^2 + 2^2 + \ldots. 6^2}$$

Hieraan is weinig en zeer eenvoudig rekenwerk verbonden. Natuurlijk kunnen alleen de positieve r-waarden gebruikt worden. Aldus is ook ρ bekend, via $\eta = \lg\rho$.

Men mag aannemen, dat $\rho_i = 0$ (behoudens de kans 0.05 op een foutieve conclusie) als de $|r_i| < 0.125$ (wanneer r op 260 paren gebaseerd is) of als $|r_i| < 0.142$ (200 paren). Immers, als r op een groot aantal paren gebaseerd is, heeft deze r, indien althans het bi-universum normaal verdeeld is en $\rho = 0$, een standaarddeviatie $\sigma(r) \sim 1/\sqrt{N}$. Voor $N = 260$ is $\sigma(r) \sim 0.062$; voor $N = 200$ is $\sigma(r) \sim 0.071$. Louter door toeval ligt dan r met 0.95 kans tussen -2σ en $+2\sigma$.

Nadat aldus de ρ berekend is, moet nog de goedheid van de aanpassing beoordeeld worden. Hoe dit geschiedde lichten we toe aan een getallen-voorbeeld.

Gemert; dagelijkse minimum; jan. I; $r_1 = 0.83$; $r_2 = 0.62$; $r_3 = 0.50$; $r_4 = 0.38$; $r_5 = 0.27$; $r_6 = 0.19$. Er komt $\rho = 0.77$, derhalve $\rho_2 = \rho^2 = 0.60$; $\rho_3 = \rho^3 = 0.46$; $\rho_4 = \rho^4 = 0.35$; $\rho_5 = \rho^5 = 0.27$ en $\rho_6 = \rho^6 = 0.21$. De vraag is of elke r_i toelaatbaar van ρ_i verschilt. Hiertoe lezen wij uit het nomogram van F.N. Davids ("Tables of the correlation coefficient" 1938) af tussen welke grenzen met 95 % zekerheid de r zal liggen als $\rho = 0.77$. Nodig is, dat het bi-universum dubbelnormaal verdeeld is. Aangezien het hier om autocorrelatie gaat, moeten wij zeker weten, dat de dagelijkse minimumtemperaturen normaal of in goede benadering normaal verdeeld zijn. Wij hebben dit aangenomen (de tijd voor verificatie ontbrak). Voorts dient met het effectieve aantal gewerkt te worden.[1] Hoewel het om 260 paren gaat (nl. 26 keren een 10-tal), is dit aantal door de autocorrelatie effectief kleiner, afhankelijk van de grootte van de autocorrelatie. In hoofdstuk 7 wordt beredeneerd, dat, in goede **benadering** 10 door $10\,\frac{1-\rho}{1+\rho}$ vervangen moet worden, als althans $\rho_i = \rho^i$. Dan is hier 260 equivalent met 34. De bovenbedoelde grenzen voor r worden 0.59 en 0.89 (en **zouden bij 260 geweest zijn** 0.72 en 0.82). **De meting** 0.83 is er juist tussen gelegen. Voor de overige r-waarden werden ook zulke marges opgezocht. Aldus voor alle 324 aanpassingen. Het totale resultaat werd, dat doorgaans de r_i binnen de bij ρ_i behorende r-marge gelegen is. Aangezien ρ vrij hoog ligt (zegge > 0.70) is het effectieve aantal n_e (vrij) veel kleiner dan 260 (zegge < 50) en hierdoor zijn de bedoelde r-marges vrij breed, zodat de gemeten r-waarden een "goede kans maken" erin te liggen. Het valt derhalve niet mee de wet $\rho_i = \rho^i$ af te wijzen; als wij dus besluiten, dat de wet $\rho_i = \rho^i$ geldt, is niet bewezen, dat hiermede de juiste beschrijving van $\rho = \rho(i)$ gevonden zou zijn.[2] De wet is echter zo eenvoudig, dat we haar zelfs prefereren, ook al zouden de eindresultaten iets minder betrouwbaar zijn in het geval de ware wet iets anders zou luiden.

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

In hoofdstuk 6 werd beweerd, dat 10 paren $t_0, t_1; t_1, t_2 \ldots t_9, t_{10}$ equivalent zijn met effectief ongeveer $10\,\frac{1-\rho}{1+\rho}$ stuks.

---

1) Toelichting: in vele gevallen is in elk der reeksen $x_1, x_2 \ldots, x_N$ en $y_1, y_2 \ldots {}_1y_N$ waarbij het gaat om een onderzoek naar de correlatie tussen de paren x_i, y_i, geen autocorrelatie aanwezig, zodat wij met N paren (nominaal en effectief N) te doen hebben. Welke consequenties de aanwezigheid van autocorrelatie heeft (als in ons geval) behandelt o.a. V.d. Bijl in zijn dissertatie "Toepassing van statistische methoden in de klimatologie" pg. 137.

2) Deze opmerking maken wij met het oog op andere statistische modellen van persistentie, die ook recht van bestaan hebben, zolang wij ze nog niet met behulp van statistische toetsen of op fysische gronden alleen hebben kunnen afwijzen. Zo zou veel te zeggen zijn voor die modellen, bij welke de persistentie begrensd is (d.w.z. $\rho_k \equiv 0$ voor alle $k > k_0$). Het zou te ver voeren hierop in te gaan (zie W.R. 60-4; R III 251, 1960).

Hier volgt het bewijs:

Beschouw n successieve x-termen $x_1, x_2 \dots x_n$ uit een tijdreeks met persistentie-coëfficiënten $\rho_1, \rho_2 \dots$ (orden 1, 2 ...), die de wet $\rho_k = \rho^k$ volgen. Onderstel verder dat iedere x_i eenzelfde variantie σ^2 heeft. In het algemeen geldt, dat het groepsgemiddelde $g = \frac{1}{n}\Sigma_1^n x_i$ een variantie $\sigma^2(g) = \sigma^2/n_e$, met $n_e = n/w$, heeft waarin

$$w = 1 + 2\left\{\frac{n-1}{n}\rho_1 + \frac{n-2}{n}\rho_2 + \dots \frac{1}{n}\rho_{n-1}\right\}$$

w = effectieve herhalingsgetal van Bartels. Als elke $\rho_k = 0$ (autocorrelatie afwezig) is $w = 1$ en $n_e \equiv n$ (voor de standaarddeviatie $\sigma(g)$ geldt dan de "$\sqrt{n}$-wet "). Substitueer $\rho_k = \rho^k$ en er komt

$$w = 1 + 2\left[\frac{n-1}{n}\rho + \frac{n-2}{n}\rho^2 + \dots \frac{1}{n}\rho^{n-1}\right] = 1 + 2\frac{\rho^{n-1}}{n}\left[1 + 2\rho^{-1} + 3\rho^{-2} + \dots (n-1)\rho^{-(n-2)}\right]$$

Een tussenstelling luidt: $1 + 2m + 3m^2 + \dots km^{k-1} = \dfrac{(1-m^k) - km^k(1-m)}{(1-m)^2}$

Hier toegepast, komt er:

$$w = 1 + 2\frac{\rho^{n-1}}{n}\left[\frac{1-\rho^{-(n-1)} - (n-1)\rho^{-(n-1)}(1-\rho^{-1})}{(1-\rho^{-1})^2}\right] =$$

$$\boxed{\frac{1+\rho}{1-\rho} - \frac{2\rho}{n}\,\frac{1-\rho^n}{(1-\rho)^2}}$$

Er staat dus $w \cong \frac{1+\rho}{1-\rho}$ als althans de tweede term mag worden verwaarloosd. Mag dit inderdaad als n = 10 ?

De volgende tabel geeft enige numerieke waarden. De exacte waarde van w blijkt kleiner dan de benadering $(1+\rho)/(1-\rho)$; bij gevolg is de exacte waarde van $n_e = 10/w$ altijd groter dan de benadering $10\,\frac{1-\rho}{1+\rho}$, maar de verschillen zijn voor bijv. $\rho < 0.80$ niet groot (met toenemende n wordt het verschil kleiner; wij namen echter alles per dekade, n = 10). Toch werd eenvoudigheidshalve steeds met de benadering gewerkt. Bij gevolg zijn de in hoofdstuk 6 bedoelde r-marges wat nauwer dan ze werden berekend. Belangrijke consequenties heeft deze iets minder correcte handelwijze zeker niet.

ρ	A = benadering van w $\frac{1+\rho}{1-\rho}$	B $0.2\rho\,\frac{1-\rho^{10}}{(1-\rho)^2}$	exacte w $w = A - B$	exacte $n_e/10$ $= \frac{1}{w}$	benadering van $n_e/10$ $\frac{1-\rho}{1+\rho}$
1.0	∞	∞	10.0	0.100	0.000
0.9	19	11.7	7.3	0.143	0.053
0.8	9	3.57	5.43	0.185	0.111
0.5	3	0.40	2.60	0.340	0.333
0.2	1.5	0.0625	1.44	0.695	0.666
0.1	1.22	0.0246	1.195	0.834	0.820
0.0	1.00	0	1.000	1.000	1.000

8. De jaarlijkse gang van de eerste orde autocorrelatie-coëfficiënt

Nadat, zoals uiteengezet werd in hoofdstuk 6, voor elk der 3 stations, voor elk der 3 elementen, voor elke dekade afzonderlijk, de ρ uit de wet $\rho_k = \rho^k$ gevonden was, bestudeerden wij de jaarlijkse gang. Deze is voor ieder element zeer markant en statistisch significant. Zo wisselt voor Naaldwijk de ρ der dagelijkse maximumtemperaturen tussen ongeveer 0.90 (zeer hoog!) in jan. III en ongeveer 0.62 in jun. I. De jaarcurve vertoont twee maxima (0.90 jan. III en 0.80 aug. III) en twee minima (0.63 jun.II en 0.64 nov. I). Men kan de ligging en de hoogten van deze maxima en minima onmogelijk exact aangeven, daarvoor spreiden de ρ-waarden te sterk, terwijl ze bovendien elk zekere onbetrouwbaarheid hebben. Wij hebben ervan afgezien zekere analytische functies (die vermoedelijk verre van eenvoudig geweest zouden zijn) voor de jaarlijkse gang aan te passen. Het viel op, dat de jaarcurven der drie stations voor eenzelfde element zoveel gelijkenis vertonen, dat het voor de hand lag de drie curven te middelen. De nieuwe 36 ρ-waarden spreiden dan minder. Vervolgens werd een vloeiende kromme uit de hand getrokken. Van deze werden de ρ-waarden afgelezen. Daarmede konden de effectieve aantallen (zie 4.3), die bij de s^2-waarden van belang zijn, worden berekend.

9. Berekening van de $s(\Delta)$ uit de $s(|\Delta|)$

In hoofdstuk 3 werd uiteengezet (zie ook lijst E in hoofdstuk 2), dat ten aanzien van de maximum- resp. minimumtemperatuur, de frequentieverdelingen der $|\Delta|$-waarden machinaal vervaardigd werden, terwijl ook de $s^2(|\Delta|)$ berekend werd. Het gaat ons echter om de frequentieverdelingen van Δ en om $s^2(\Delta)$.

Laat x een kansverdeling $\varphi(x)dx$ volgen met gemiddelde μ en variantie σ^2, bijv. $-\infty < x < \infty$. Zij $\hat{x} = |x|$, met $0 \leqslant \hat{x} < \infty$. De kansverdeling $\Psi(\hat{x})d\hat{x}$, waaraan $\hat{x}$ gehoorzaamt, is zeker niet gelijk aan $\varphi(x)dx$. Wij willen $\sigma^2(\hat{x})$ in μ en σ^2 uitdrukken.

(a) Allereerst: $\Psi(\hat{x}) = \varphi(x) + \varphi(-x)$, omdat uit $\hat{x} = a$ volgt, dat òf $x = a$ òf $x = -a$.

(b) Zij p.d. $\gamma \equiv \mathcal{E}\hat{x} \equiv \int_0^\infty \hat{x}\varphi(\hat{x})d\hat{x}$ en $\sigma^2(\hat{x}) \equiv \mathcal{E}(\hat{x} - \gamma)^2 = \mathcal{E}\hat{x}^2 - \gamma^2$

Daarin is

(c) $\gamma = \int_0^\infty x\varphi(x)dx + \int_0^\infty x\varphi(-x)dx$ en $\mathcal{E}\hat{x}^2 = \int_0^\infty x^2\varphi(x)dx + \int_0^\infty x^2\varphi(-x)dx$

Onverschillig de vorm van $\varphi(x)$, geldt altijd $\mathcal{E}\hat{x}^2 = \mathcal{E}x^2$, echter is $\mathcal{E}\hat{x} \neq \mathcal{E}x$. Het statistisch probleem luidt dus: gegeven de $\bar{\hat{x}} = \frac{1}{n}\Sigma_1^n \hat{x}_i$ en $s^2(\hat{x}) = \frac{1}{n-1}\Sigma_1^n(\hat{x}_i - \bar{\hat{x}})^2$, gezocht de $\mu(x)$ en $\sigma^2(x)$, d.w.z. schattingen daarvan, nu wij weten dat geldt

$$\sigma^2(\hat{x}) \equiv \int_0^\infty x^2\varphi(x)dx - \left[\int_0^\infty x\{\varphi(x) + \varphi(-x)\}dx\right]^2, \text{ terwijl}$$

$$\mu(x) \equiv \int_{-\infty}^\infty x\varphi(x)dx \text{ en } \sigma^2(x) \equiv \int_{-\infty}^\infty x^2\varphi(x)dx - \left[\int_{-\infty}^\infty x\varphi(x)dx\right]^2.$$

De $\varphi(x)dx$ is de onbekende kansverdeling van x.

Dit probleem schijnt onoplosbaar, tenzij wij daarbij mogen onderstellen, dat $\varphi(x)$ een normale verdeling $N(\mu; \sigma)$ volgt, dus dat

$$\varphi(x)dx = (\sigma\sqrt{2\pi})^{-1} \exp. \left\{-\tfrac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} dx.$$

Dan laat zich berekenen, dat

$$\gamma = 2\sigma\left(A + \frac{C}{\sqrt{2\pi}}\right), \text{ met } A = p\int_0^p \phi(t)dt \text{ en } C = e^{-\frac{1}{2}p^2},$$

waarbij $p = \mu/\sigma$ en $\phi(t)dt$ voor de standaardnormale verdeling $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$, staat

(d) Bijgevolg is $\sigma^2(\hat{x}) = \sigma^2 + \mu^2 - 4\sigma^2\left(A + \frac{C}{\sqrt{2\pi}}\right)^2$

(e) Naarmate μ kleiner is t.o.v. σ, d.i. als $p \to 0$, ligt A dichter bij nul en C dichter bij 1 en $\sigma^2(\hat{x}) \to \sigma^2 - \mathcal{E}^2 = \sigma_x^2(1 - \frac{2}{\pi})$ of $\sigma_x^2 \cong 1.47\,\sigma$

$\sigma^2(\hat{x})$ of $\sigma = 1.21\ \sigma(\hat{x})$

De bij het hier gestelde probleem te volgen weg is: substitueer de eendaagse verandering Δ voor x. Als de machine levert ~~$s^2(\hat{x})$, $\bar{\Delta}$~~ dan hebben wij in

$$s^2(|\Delta|) = \sigma^2 + \mu^2 - 4\sigma^2 (A + \frac{C}{\sqrt{2\pi}})^2$$

$$\text{(f)} \quad \overline{|\Delta|} = 2\ (A + \frac{C}{\sqrt{2\pi}})$$

$$\mu = p\sigma$$

3 vergelijkingen met 3 onbekenden: σ, μ, p, waarbij μ en σ het gemiddelde en de standaarddeviatie in de normaal gedachte kansverdeling van Δ voorstellen. Toen echter enkele keren gebleken was, dat een exacte (overigens toch grafische) oplossing weinig andere waarden voor σ en μ opleverde, dan de approximatieve oplossing $\sigma_{\Delta}^2 = s^2(|\Delta|) + \{\overline{|\Delta|}\}^2$, volgende uit (e), werd voor verdere bewerkingen met de laatste gewerkt. Deze handelwijze is des te minder ongeoorloofd naarmate de Δ-verdeling "dichter bij de normale rondom nul gelegen is".

10. Nomogrammen der interdiurne variabiliteit

Nadat bewezen was, dat alle Δ-populaties in goede benadering normaal verdeeld waren, derhalve slechts in de varianties σ^2 verschilden, lag het voor de hand op gaussisch-waarschijnlijkheidspapier een nomogram te construeren, zodat op eenvoudige en snelle wijze vele typen vragen kunnen worden beantwoord. Een normale verdeling wordt dan immers door een rechte lijn weergegeven. Men zou hiertoe op één vel lineair-gaussisch-waarschijnlijkheidspapier, bijv. voor de I.V. van de maximumtemperatuur, 3 x 12 = 36 rechte lijnen kunnen trekken; zij gaan alle door het punt $\Delta = 0$, $P = 0.50$ en hellen onder verschillende hoeken. Wanneer men op hetzelfde vel 3 x 36 rechten (3 stations) trekken wil, wordt het beeld zo vol en druk, dat men er niet meer mee werken kan. Dan ligt het voor de hand om de rechten niet te trekken, doch om een constructie uit te voeren, waarbij men een doorzichtige loper in een bepaalde stand moet leggen, opdat de rechte lijn op deze loper bijv. de cumulatieve kansverdeling der eerste orde differenties van de dagelijkse maximumtemperatuur te Naaldwijk voorstellen zal. Deze gedachtengang leidde tot het hierbij gevoegde nomogram. Het spreekt vanzelf, dat de s_Δ-waarden vereffend moesten worden ten einde de jaarlijkse gang voor een niet te grillige kromme te kunnen voorstellen. Hiertoe werd in de regel gladgestreken door twee keren de formule $b^1 = \frac{1}{3}(a + b + c)$

toe te passen.[1] Tussen de daarna verkregen punten werd op het oog een vloeiende kromme gelegd (die dus per station, per element, de waarde van $\sigma(\Delta)$ als functie van de dekade aangeeft). Wij hebben nog getracht diegene der s_Δ-waarden samen te nemen waarvoor dit statistisch toelaatbaar is, maar het resultaat kwam aan de overzichtelijkheid niet ten goede. Dit al of niet samennemen ("poolen") is in grote mate afhankelijk van de effectieve aantallen. Hoe groter zij zijn (en zij zijn hier "zeer groot", veel groter dan de nominale aantallen 260 en 200) hoe "minder vlug" zal men mogen samennemen bij eenzelfde verschil tussen de steekproefvarianties. Het leek ons daarom beter het oordeel over dit samennemen geheel aan de gebruiker over te laten. In de toelichting bij het nomogram wordt uitvoerig uiteengezet, hoe men daartoe handelen moet.

Voor het overige moge verwezen worden naar de bij dit rapport gevoegde nota, getiteld:

"Numerieke voorbeelden als toelichting bij het gebruik van de nomogrammen der Interdiurne Variabiliteit (I.V.)": V-62; R III 243-1960.

Hier volgt een opsomming der 8 nomogrammen

(1) $\Delta(t)$-nomogram voor de I.V. der minimumtemperatuur, te Gemert, Naaldwijk en Groningen.

(2) $\Delta(T)$-nomogram voor de I.V. der maximumtemperatuur, te Gemert, Naaldwijk en Groningen.

(3) $\Delta(U)$-nomogram voor de I.V. der 14-uurs relatieve vochtigheid, zelfde stations.

(4) H-nomogram, dat ik construeerde bij de Hartley-toets.

(5) $B(\alpha = 0.005)$-nomogram, dat ik construeerde voor de aflezing van betrouwbaarheidsmarges, bij een kans 0.010 op een foutieve conclusie.

---

1) Er zijn vele andere vereffeningsformules in gebruik. De vraag of de ene "beter" is dan de andere, is niet gemakkelijk te beantwoorden. Het antwoord hangt niet slechts af van de statistische definitie van het begrip "beter", doch bovendien van de vorm van de curve, waarop de ware onbekende waarden, die men via metingen benadert, gelegen zijn. Wij meenden, dat het voldoende was die vereffening te kiezen, die het minste rekenwerk vergde.

(6) B(α = 0.025)-nomogram, idem.

(7) B(α = 0.050)-nomogram, idem.

Deze drie B-nomogrammen besprak ik in Statistica Neerlandia 13 1, 1959.

(8) F-nomogram, dat ik construeerde speciaal voor grote aantallen vrijheidsgraden. Het wordt gebruikt om uit te maken of twee varianties s_1^2 en s_2^2 (berustende op ν_1 en ν_2 aantallen graden van vrijheid) statistisch verschillen.

11. Typen van vragen, die niet met de nomogrammen te beantwoorden zijn

11.1 Vraag:

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 1°C verandert? Zulk een reeks bezit een bepaalde "standvastigheid" in de maximumtemperatuur.

Antwoord:

Met het Δ-nomogram vinden wij, dat de kans op een $|\Delta| \geqslant 1°C$ is $P = 1 - 2 \times 0.34 = 0.32$. Hierbij wordt echter niets gegeven met betrekking tot de grootten of tekens der voorafgegane Δ's. De gestelde vraag zou kunnen worden beantwoord als men ook voorwaardelijke kansen zou kennen van het soort als: welke kans is er op een $|\Delta| \leqslant 1°C$ als ook elk der twee direct voorafgaande Δ's in absolute zin kleiner dan 1°C is? Deze kansen werden niet berekend, doch zouden met het beschikbare materiaal wel berekend kunnen worden. De min of meer voor de hand liggende schatting $P(1 - P)^5 P = 0.015$ van de gevraagde kans kan ontoelaatbaar grof zijn.

11.2 Vraag:

De minimumtemperatuur op zekere dag zij gegeven: 4°C, jan. II, Naaldwijk. Welke kans is er op vorst (dus $\Delta \leqslant - 4°$ C)?

Antwoord:

Dit is weer een voorwaardelijke kans, nu echter van een andere soort dan de bovenvermelde, waarvoor het Δ-nomogram evenmin bedoeld is. Het nomogram levert: $P(\Delta \leqslant - 4°C) = 0.047$, maar daarbij wordt de minimumtemperatuur zelf niet gegeven. Het ligt voor de hand, dat de kans op een $\Delta \leqslant - 4°C$ afhangen zal van het niveau van de temperatuur, maar wij weten niet numeriek hoe. Het is zeer waarschijnlijk, dat bij een lagere minimumtemperatuur eenzelfde negatieve Δ minder waarschijnlijk is. In dit geval wordt geëist, dat twee gebeurtenissen gelijktijdig plaatsvinden; de ene gebeurtenis is, dat de minimumtemperatuur ligt tussen bijv. $3\frac{1}{2}$ en $4\frac{1}{2}^{o}$C en de tweede is, dat $\Delta \leqslant -4^{o}$C. Derhalve bestaan er twee kansen, k en P. De kans, dat de twee gebeurtenissen plaatsvinden is echter niet $k \times P$ en wel doordat de gebeurtenissen niet onafhankelijk zijn (de sterkte van de afhankelijkheid zou een punt van studie kunnen zijn).

11.3 <u>Vraag</u>:

Welke kans is er, dat de gehele dekade jan. II te Naaldwijk tenminste één $\Delta \geqslant 4^{o}$ telt?

<u>Antwoord</u>:

Een gegeven dekade bevat 10 Δ's, t.w. $\Delta_1 = t_1 - t_0$; $\Delta_2 = t_2 - t_1$ ... $\Delta_{10} = t_{10} - t_9$. Het nomogram leert, dat de kans op een $\Delta \geqslant 4^{o}$C ongeveer $P = 0.047$ is. De kans K, dat van 10 Δ's er tenminste één $\Delta \geqslant 4^{o}$ is, zou $1 - (1 - 0.047)^{10} = 0.38$ geweest zijn als deze 10 Δ's ongecorreleerd zijn. Echter zijn ze wel gecorreleerd (d.i. gepersisteerd). De exponent 10 moet daarom worden vervangen door een getal, waarmede de 10 effectief equivalent is. Dit effectieve aantal is: $10_e = 10^2 \frac{1-\rho}{1-\rho^{10}}$, als althans de persistentie in de t-reeks zelf met $\rho_i = \rho^i$ beschreven kan worden. Van het nomogram lezen wij af: $\rho = 0.87$; dus $10_e = 17$ en K gaat over in $K^1 = 1 - (1 - 0.047)^{17} = 0.56 > 0.38$. De gezochte kans is dus aanmerkelijk vergroot door de persistentie. Ook al zijn de t's niet gepersisteerd, de Δ's zijn het altijd en dan zelfs zo zwaar mogelijk, want 10_e is dan 10^2. Met deze 10_e zou de kans geweest zijn $K'' = 1 - (1 - 0.047)^{100} = 0.99$.

De beantwoording van zulk een vraag op de boven geschetste wijze blijft, statistisch gesproken, ietwat riskant, maar er valt weinig anders te doen, als er geen betere gegevens ter beschikking staan. Deze betere gegevens zouden kunnen zijn: kennis van de kans Q_i op een $\Delta < 4^{o}$C <u>als er tenminste i stuks Δ's $< 4^{o}$C direct voorafgaan</u> (dus weer van het ook in 11.1 bedoelde type). Dan is $Q_0 = 1 - P$. De kans op een dekade, waarin zich tenminste één $\Delta \geqslant 4^{o}$C bevindt is dan 1 - kans, dat elk der 10 Δ's kleiner dan 4^{o}C is, d.i. $1 - Q_0 . Q_1 . Q_2 \ldots . Q_9$. Zoals reeds gezegd, werden deze Q's niet berekend.

Men zou echter elke Q_i door 1 - P kunnen vervangen en tegelijkertijd het nominale aantal door het effectieve aantal. Als wij $1-(1-P)^{10} = 1-(1-0.047)^{10} = 1-0.62 = 0.38$ vergelijken met $1-(1-P)\,Q_1.Q_2 \ldots\ldots Q_9$ en deze uitdrukking weer vervangen door $1-(1-P)^{17} = 1-0.44 = 0.36$ moet de conclusie luiden, dat $Q_1 < 1-P$; $Q_2 < Q_1$; $Q_3 < Q_2$ .... M.a.w. hoe meer dagen er al direct voorafgaan met een $\Delta < 4^{\circ}C$, hoe kleiner is de kans, dat ook de volgende dag een $\Delta < 4^{\circ}C$ zal brengen en dit alles louter ten gevolge van de negatieve persistentie Δ's.

N.B.: Negatieve persistentie in niet afgeleide, gebruikelijke, klimatologische reeksen is zeldzaam.

11.4 Vraag:

Welke kans is er te Gemert in mei II, dat de maximumtemperatuur in één dag met tenminste $6^{\circ}C$ toeneemt en tegelijkertijd de 14-uurs relatieve vochtigheid met tenminste 25 % stijgt?

Antwoord:

Als zulke gebeurtenissen onafhankelijk zouden zijn, zou men met het product der kans op de gebeurtenissen afzonderlijk kunnen rekenen, waarbij weer de nomogrammen geraadpleegd zouden kunnen worden. Nu zij dit niet zijn (althans dit vermoeden wij) is het antwoord op de vraag niet te geven.

11.5 Vraag:

Welke is de gemiddelde lengte van een onafgebroken reeks van dagen met toenemende maximumtemperatuur (d.i. positieve Δ's) in oktober II, te Naaldwijk?

Antwoord:

Vergelijken wij deze vraag met de in 11.1 behandelde vraag, dan is de daar bedoelde drempel thans $0^{\circ}C$.
Ook nu kan het nomogram niet helpen. Het zegt slechts, dat de "overall-kans" op een positieve Δ gelijk $\frac{1}{2}$ is, doch de conclusie, dat derhalve de kans op een reeks van precies a positieve Δ's gelijk $(\frac{1}{2})(\frac{1}{2})^a(\frac{1}{2}) = 2^{-(a+2)}$ zou zijn, is fout, omdat de reeks der Δ's niet persistentie-vrij is (te denken aan de γ-waarden, die principieel **ongelijk** nul zijn).
De volgende stelling geldt algemeen: in een reeks met elementen A en B, die elkaar uitsluiten en elkaar qua optreden niet beïnvloeden

(persistentie-vrij), waarbij de kansen zijn p en $q = 1 - p$, is de kans op een successie (iteratie of run) van a A's gelijk aan q^2p^a. De gemiddelde lengte $\overline{L}_A$ van een A-iteratie blijkt $\frac{1}{1-p} = \frac{1}{q}$. De formule voor $\overline{L}_A$ verandert niet voor het geval de reeks wèl persistentie bezit, alleen moeten wij de p vervangen door p_{11}, betekenende: de kans op tenminste 1 of meer A's, als er tenminste 1 of meer A's gepasseerd zijn. De overall-kans op A wordt dan geschreven p_{10}, betekenende de kans op een iteratie van 1 of meer A's als er wel of niet A's gepasseerd zijn. In een reeks met positieve persistentie is $p_{11} > p_{10}$ en bij gevolg is de $\overline{L}_A$ door de persistentie vergroot. In een reeks met negatieve persistentie (anti-persistentie), zoals in ons geval met de Δ-reeks, is $p_{11} < p_{10}$ en is $\overline{L}_A$ verkleind. In ons geval is $p_{10} = \frac{1}{2}$ en $p_{11} < \frac{1}{2}$ en $\overline{L}_A < 2$, waarbij het alternatief A voor een $\Delta > 0$ en het alternatief B voor een $\Delta < 0$ staat. Zonder positieve of negatieve persistentie zou $\overline{L}_A = \overline{L}_B = 2$ geweest zijn. De moeilijkheid is om p_{11} in p_{10} uit te drukken d.m.v. de autocorrelatie-coëfficiënten. Ons zijn geen publicaties bekend, waarin dit geschiedt.

N.B.: De gemiddelde lengte van een iteratie van dagen, waarbij op elke dag de maximumtemperatuur boven normaal is, zou, als er geen autocorrelatie zou zijn, precies 2 geweest zijn. Nu er wèl positieve correlatie is (elke $\rho_k > 0$), is de gemiddelde lengte groter dan 2. De gemiddelde lengte van een iteratie van dagen, waarbij van dag tot dag de maximumtemperatuur stijgt, is kleiner dan 2, ook als de initiaal-reeks zonder persistentie is, omdat de differentiereeks altijd negatieve persistentie bezit. Nu deze initiaal-reeks positieve autocorrelatie vertoont, is de bedoelde gemiddelde lengte nog kleiner.

12. Addendum

Indien in steekproeven autocorrelatie aanwezig is, verliezen vele overigens zeer gebruikelijke uitdrukkingen als gemiddelde $\overline{x} = \frac{1}{n}\Sigma_1^n x_i$ en $s^2 = \frac{1}{n-1}\Sigma_1^n (x_i - \overline{x})^2$, hun "geldigheid", d.w.z. zij zijn dan geen zuivere schattingen meer van de onbekende parameters μ (gemiddelde) en σ^2 (variantie) van het universum. De statistische terminologie luidt: de schattingen hebben in zo'n geval "bias"; ze zijn "biased", "onzuiver". Zo is de uitdrukking $\overline{x} = \frac{1}{n}\Sigma_1^n x_i$ een zuivere schatting van μ

(autocorrelatie in de steekproef afwezig gedacht) - en wel hoe ook de verdeling in de populatie moge zijn - omdat de kansverdeling van $\bar{x}$ een gemiddelde μ heeft. Anders geformuleerd: berekent men m keren dit gemiddelde (n vast), $\bar{x}_1$, $\bar{x}_2$ ... $\bar{x}_m$, dan is $\lim\limits_{m\to\infty} \bar{\bar{x}} \equiv \lim\limits_{m\to\infty} \frac{1}{m} \Sigma_{j=1}^{m} \bar{x}_j = \mu$. Analoog is $\lim\limits_{m\to\infty} \bar{s}^2 \equiv \lim\limits_{m\to\infty} \frac{1}{m} \Sigma_{j=1}^{m} s_j^2 = \sigma^2$. Wij schrijven ook: $\mathcal{E}\underline{\bar{x}} = \mu$ en $\mathcal{E}\underline{s}^2 = \sigma^2$. In het geval van autocorrelatie is $\mathcal{E}\underline{\bar{x}} = f\mu$ en $\mathcal{E}\underline{s}^2 = gu$, waarin f en g functies zijn van de autocorrelatie-coëfficiënten ρ_1, ρ_2 ... in de reeks. Als $f > 1$ dan heet de $\bar{x}$-uitdrukking een "schatter" met positieve "bias". Ditzelfde geldt voor de s^2-uitdrukking. Bovendien worden, ten gevolge van de autocorrelatie-coëfficiënten, de nauwkeurigheden van $\bar{x}$ en s^2 anders (beter gezegd: de kansverdelingen van $\underline{\bar{x}}$ en $\underline{s}^2$ krijgen standaarddeviaties, die met de ρ's samenhangen). Het is nu een mathematisch belangrijk probleem te zoeken naar die wèl zuivere "schatters" van μ en σ^2, die bovendien zo nauwkeurig mogelijk zijn. Het zou te ver voeren deze hier te behandelen. Zij zijn zeer ingewikkeld en machine-technisch onhandelbaar. Daarbij komt, dat de uitdrukkingen nog alleen bekend zijn voor het geval de ρ's bekend zijn; in ons onderzoek echter moeten deze uit hetzelfde materiaal berekend worden, dat $\bar{x}$, s^2 etc. leverde. Bovendien werd in ons onderzoek, ter beperking van de omvang daarvan, de s afgeleid uit de gemiddelde absolute afwijking (mean deviation) e zodat er een tweetal nieuwe problemen rijst, namelijk (a) in hoeverre is $e = \frac{1}{n} \Sigma_1^n |x|$ een zuivere schatting van $\mathcal{E}$ en (b) hoe is het met de verhouding $\mathcal{E} : \sigma$ (bij de steekproef e:s) in het geval van autocorrelatie?

Ik ben Prof. N.H. Kuiper en zijn assistenten (Afd. Wiskunde, Landbouwhogeschool, Wageningen) veel dank verschuldigd voor hun bereidheid een studie te maken van de statistische aspecten van parameterschattingen met behulp van steekproeven met autocorrelatie. Het bovenstaande is ontleend aan deze nog niet afgesloten studie.

Voor de "practicus" is het natuurlijk van belang te weten, of hij veel "wint" door zo exact mogelijk te werk te gaan. Het is nog niet mogelijk deze vraag te beantwoorden. Wij krijgen echter de indruk, dat de in het rapport beschreven procedure voldoende goed is voor het doel.

13. Tabellen

De volgende tabellen zijn gemaakt:

(1) Voor elk der drie stations, voor elk der 3 elementen, voor elke der 36 dekaden, de 6 r_k-waarden ($k = 1, 2 \ldots 6$) alsmede de 6 ρ_k ($= \rho^k$)-waarden.

(2) Voor elk der 3 stations, voor elk der 3 elementen, voor elke der 36 dekaden, voor elk der 26 (20) jaren 1931 t/m 1956 ('50), de gemiddelde waarde.

(3) Voor elk der drie stations, voor elk der 3 elementen, voor elke der 36 dekaden, gemiddelde over 26 (20) jaren, de gemiddelde eendaagse differentie, geacht en ongeacht het teken.

(4) Voor elk der 3 stations, voor elk der 3 elementen, voor elke der 36 dekaden, over 26 (20) jaren, de standaarddeviatie der dagelijkse waarde.

(5) Voor elk der 3 stations, voor elk der 3 elementen, voor elke der 36 dekaden, over 26 (20) jaren, de variantie en de standaarddeviatie van de eendaagse differentie.

(6) Voor elk der 3 stations, voor elk der 3 elementen, voor elke der 36 dekaden, over 26 (20) jaren, de grootste positieve en de grootste negatieve eendaagse differentie, met datum.

De tabellen (3) en (6) werden bij dit rapport ingesloten.

Summary

0. Introduction

Several questions from practice, expecially with regard to the rating of new regions in the Netherlands on their suitability for horticulture, necessitated a far more thorough study of some meteorological elements than was published by Hartman in Med. en Verh. K.N.M.I. No. 24, 1918.

The following stations were taken into consideration:

Naaldwijk (west), Gemert (south-east), Groningen (north-east).

The elements studied are: the daily minimumtemperature t, the daily maximumtemperature T and the relative humidity U of 14.40 M.E.T. (all measured at a height of 2.20 m).

The basic periods are: Na: 1931-1956, Ge: 1931-1956 and Gr: 1931-1950.

The temperatures t and T refer to a 24 hours period between 19.40 and 19.40 M.E.T.

The one day, two days, etc.-differences are defined as ${}_k\Delta_i = t_i - t_{i-k}$ for day no. i and order k (= 1, 2, 3 ...) and element t.

The well-known interdiurnal variation is the one day difference or difference of order one: ${}_1\Delta_i = t_i - t_{i-1}$ (shorter: Δ_i).

1. Characterizing the interdiurnal variability

Some statistical considerations are given as to characterizing the interdiurnal variability.

The population of Δ_i-values of t situated in, for instance, all decades I of January at Naaldwijk (a sample of 26 x 10 = 260 values) proves to be normal around zero, and the same holds for all other decades, each of the three stations and each of the three elemts. Hence we do well to use the standard deviation σ and to compare different decades or different stations as to the interdiurnal variability on the ground of this standard deviation.

Most authors on the subject of interdiurnal variability only mention the mean absolute value $\overline{|\Delta|}$; this is the sample value of ε in the population, just like it is s for σ. For normal populations (and only for normal populations) the relation $\varepsilon = \sigma\sqrt{2/\Pi} = 0.8\sigma$ holds.

2. The lists

This chapter describes all the lists which were produced wholly mechanically by the tabulator and the sorter with the aid of the basic-punched cards, containing the ${}_k\Delta_i$-values ($k = 1, 2 \dots 6$; $i = 1, 2 \dots 365$).

3. The normality of the Δ-populations

It turned out to be necessary to investigate whether the probability distributions of the populations of the one day differences (Δ_i) are normal. Since this had to be done for 324 frequency distributions (3 stations; 3 elements; 36 decades) the graphical way was chosen: each cumulative distribution was plotted on arithmatic probability paper. Nearly all graphs proved to be linear. Consequently all distributions were (approximately) gaussian. This enabled us to construct nomographs (chapter 10).

4. On the necessity of the knowledge of persistence in investigations about the annual course of the interdiurnal variability or of the differences between stations

4.1 The persistence coefficient

In this chapter the problem is considered how to examine the statistical reality of the difference between, for instance, two measured sample standard deviations s_1 and s_3, for the decades 1 and 3 of april, for t and for Naaldwijk.

We should take into account the fact that neither the t series (t_1, t_2 ....) nor the Δ series ($\Delta_1 = t_1 - t_0$; $\Delta_2 = t_2 - t_1$ ....) is random. It is easily proved that the t-series posesses a positive autocorrelation described by the so called autocorrelation coefficients ρ_1, ρ_2 ... (orders $k = 1, 2 \dots$); ρ_k represents the correlation coefficient between t_i and t_{i+k} ($i = 1, 2, 3 \dots$); each $\rho_k < \rho_{k-1}$ and each $\rho_k > 0$. The well-known law $\rho_k = \rho^k$ could be affirmed for all decades, for all stations and all elements. Even if the t series would not be autocorrelated, the Δ series would be so. The autocorrelation coefficients in the Δ series would then be $\gamma_1 = -\frac{1}{2}$; $\gamma_2 = \gamma_3 = \dots = 0$. Since the t-series does possess positive autocorrelation ($\rho_k = \rho^k$) it can be shown that $\gamma_k = -\frac{1}{2}\rho^{k-1}(1 - \rho)$ (the index 1 in ρ_1 is omitted).

4.2 The effective numbers

A group of n successive Δ-values has a mean value $g_n = \frac{1}{n} \Sigma_1^n \Delta_i$ which is a stochastic variable with a variance

$$\sigma^2(g_n) = \frac{\tau^2}{n}\left[1 + 2\left\{\frac{n-1}{n}\gamma_1 + \frac{n-2}{n}\gamma_2 + \dots \frac{1}{n}\gamma_{n-1}\right\}\right] \equiv \frac{\tau^2}{n_e}$$

with τ^2 = variance of Δ_i; n_e is called "effective number". The γ_k's are the autocorrelation coefficients in the Δ series. From $\gamma_k = -\frac{1}{2}\rho^{k-1}(1-\rho)$ follows

$$n_e = n^2\frac{1-\rho}{1-\rho^n} \text{ and } n \leqslant \frac{1-\rho}{1-\rho^n} n^2 < n^2 \text{ for } 1 > \rho > 0$$

Stress is laid on the surprising result $n_e > n$ (generally $n_e < n$) which is only a consequence of the fact that the autocorrelation is negative; negative autocorrelation is rare in climatological series and seems to occur only in series of derived climatological elements (the primary series is t_1, t_2 ...; the derived series is Δ_1, Δ_2 ...). When examining the statistical significance of the difference between two sample variances $s^2(\Delta)$ we decided to enter, for instance, the F-table with the mentioned effective numbers n_e as numbers of degrees of freedom.

5. The evaluation of the autocorrelation coefficients r_1, r_2 ... r_6

For the sake of speed we developed an approximative method to compute the 6 correlation coefficients r_1, ... r_6 for each of the 36 decades, for each of the 3 stations and for each of the 3 elements. This was done with the help of the differences $_k\Delta$ of orders k = 1, 2 ... 6. The reasoning was: from the definiton $_k\Delta_i = t_i - t_{i-k}$ follows

$$\sigma^2(_k\Delta) = 2(1-\rho_k)\,\sigma^2(t) \text{ or } \rho_k = 1 - \frac{1}{2}\left[\frac{\sigma(_k\Delta)}{\sigma(t)}\right]^2 = 1 - \frac{1}{2}\left[\frac{\varepsilon(_k\Delta)}{\varepsilon(t)}\right]^2$$

with $\varepsilon = \sigma\sqrt{2/\Pi}$ for normal populations.
Since ε is estimated in the sample as $e = \overline{|\Delta|}$, the ρ_K is estimated as

$$r_k = 1 - \frac{1}{2}\frac{\overline{|_k\Delta|}}{\overline{|t-\bar{t}|}} \quad \text{, with } \overline{_k\Delta} \cong 0.$$

6. The fitting of $\rho_k = \rho^k$ to the computed set $r_1, r_2 \dots r_6$

The "law" ("pure persistence") $\rho_k = \rho^k$ is tested and affirmed. In consequence of this geometrical relation all further formulas become extremely simple.

7. The formula $n_e = n \frac{1 - \rho}{1 + \rho}$ in time series with autocorrelation coefficients $\rho_k = \rho^k$

In this chapter the statistical consequence is treated as to the law $\rho_k = \rho^k$ on a group of n successive elements t_i. Then the mean g of the group has a standard deviation $\sigma(g) = \sigma/\sqrt{n_e}$ with

$$n_e = n : w \quad \text{and} \quad w = \frac{1 + \rho}{1 - \rho} - 2\frac{\rho}{n}\,\frac{1 - \rho^n}{(1 - \rho)^n}$$

For large n and small ρ the formula simplifies to $w = \frac{1 + \rho}{1 - \rho}$

8. The annual course in the first order autocorrelation coefficient ρ

There is, for the same element, no difference between the 3 stations as to the value of ρ. However, there is a marked annual course, in particular for T (two maxima, two minima). The ρ's are situated around a high level 0.6 à 0.9. The daily amounts of rainfall show no annual course; $\rho_k = \rho^k$, with $\rho = 0.25$.

9. The evaluation of s_Δ from $s_{|\Delta|}$

In this connection the following particular mathematical-statistical problem is treated.

Let x satisfy a probability distribution $\varphi(x)dx$ with a mean μ and a variance σ^2 (for instance $-\infty < x < \infty$). Let be $\hat{x} \equiv |x|$, with $0 \leqslant \hat{x} < \infty$. This $\hat{x}$ follows a new probability distribution $\Psi(\hat{x})d\hat{x}$. We want to express the variance $\sigma^2(\hat{x})$ of $\hat{x}$ in terms of μ and σ^2.

The problem seems to be solvable only in case x follows a normal distribution. Then the result is

$$\sigma^2(\hat{x}) = \sigma^2 + \mu^2 - 4\sigma^2\left(A + \frac{C}{\sqrt{2\Pi}}\right)^2 \quad \text{with } A = p\int_0^p \phi(t)dt;\; C = e^{-\frac{1}{2}p^2}$$

while $p = \mu/\sigma$ and $\phi(t)dt$ is the standardized normal distribution ($\mu = 0$; $\sigma = 1$).

10. Nomographs of the interdiurnal variability

Since all 324 Δ-populations (3 x 3 x 36 = 324; 3 elements, 3 stations, 36 decades) are normal around zero, it is possible to construct three simple analoguous nomographs, to wit for each of the elements seperately. Then, for instance, the nomograph of the interdiurnal variability (on arithmetic probability paper) would contain 36 straight lines for Naaldwijk, 36 for Gemert and 36 for Groningen, in total 108. Such a graph could not be read off easily. Therefore in the right upper corner of the nomograph three curves were drawn (each presenting s_{Δ} against the decade). A transparant piece of paper, on which is drawn a sharp straight line, can easily be posed in such a way that the straight line represents, for instance, the probability distribution of Δ or t for Naaldwijk, April decade. Then it is easy to read off, for instance, the probability of a Δ-value $\geqslant 3^{o}$C. Of course several different questions can be answered with these nomographs. A separate report gives numerical examples, V-62; R III 243, 1960.

Questions of reliability of the answers given by means of nomographs are also considered. For this purpose a set of other nomographs has been drawn.

11. Questions that cannot be answered with the aid of the nomographs

Attention is paid to many practical questions which can be important, but cannot be answered with the aid of the nomographs. However, the basic data on punched cards enable us to solve also these problems, if desired. We may ask, for instance, what is the mean length of an uninterrupted series of days in spring with daily maximumtemperatures which do not change more than 2^{o}C from day to day.

12. Addendum

This deals with a mathematical statistical question in connection with the fact that in autocorrelated series the usual expressions for the mean value and the standard deviation do not strictly hold. These expressions give biased estimates of the population parameters. The unbiased and most efficient estimates are very complicated and certainly they are untreatable in a mechanical way. Moreover, the statistical computations being still in progress when this report was written, we prefered to use the usual expressions and to work with the effective numbers (instead of with the nominal ones), although this is statistically not the best method.

Tabel 3a

Naaldwijk: $\Delta^{o}C$ van dagelijkse minimumtemperatuur t

	I				II				III				$\overline{t}^{o}C$		
	$\overline{\Delta}$	$\overline{\lvert\Delta\rvert}$	$\frac{1}{2}\overline{\Delta}_+$	$\frac{1}{2}\overline{\Delta}_-$	$\overline{\Delta}$	$\overline{\lvert\Delta\rvert}$	$\frac{1}{2}\overline{\Delta}_+$	$\frac{1}{2}\overline{\Delta}_-$	$\overline{\Delta}$	$\overline{\lvert\Delta\rvert}$	$\frac{1}{2}\overline{\Delta}_+$	$\frac{1}{2}\overline{\Delta}_-$	I	II	III
jan	-0.14	2.06	0.96	1.10	+0.07	2.05	1.06	0.99	-0.14	2.07	0.97	1.10	0.3	0.8	-1.1
feb	+0.16	2.02	1.09	0.93	+0.01	2.02	1.01	1.01	-0.01	2.04	1.02	1.02	-0.0	-0.1	-0.2
mrt	+0.09	1.85	0.97	0.88	+0.26	1.97	1.14	0.86	-0.09	2.06	0.99	1.08	0.6	2.0	3.2
apr	+0.20	1.80	1.00	0.80	+0.00	2.04	1.04	1.00	+0.08	1.79	0.94	0.86	4.5	5.2	5.6
mei	+0.21	1.81	1.01	0.80	+0.09	1.99	1.04	0.95	+0.08	2.07	1.08	1.00	7.0	8.3	9.3
jun	+0.11	1.84	0.97	0.87	+0.07	2.03	1.05	0.98	+0.09	2.06	1.08	0.99	10.7	11.1	12.4
jul	+0.10	1.96	1.03	0.93	+0.06	1.63	0.84	0.78	-0.00	1.69	0.84	0.85	13.1	13.7	14.0
aug	+0.03	1.78	0.90	0.88	-0.03	1.91	0.94	0.97	-0.11	1.85	0.87	0.98	14.0	13.7	13.1
sep	-0.14	2.15	1.00	1.14	-0.09	2.12	1.01	1.10	-0.08	2.28	1.10	1.18	12.3	11.7	10.6
okt	-0.15	2.24	1.04	1.20	-0.05	2.40	1.18	1.22	-0.37	2.21	0.92	1.29	8.9	8.2	6.1
nov	+0.05	2.11	1.08	1.03	-0.19	1.94	0.88	1.07	-0.01	2.14	1.06	1.07	5.5	4.3	3.4
dec	-0.16	1.98	0.91	1.07	-0.14	2.35	1.10	1.25	+0.08	2.14	1.11	1.03	2.6	1.4	0.7
	$\Delta^{o}C$ van dagelijkse maximumtemperatuur T												$\overline{T}^{o}C$		
jan	-0.05	1.75	0.85	0.90	+0.04	1.68	0.86	0.82	-0.11	1.57	0.73	0.84	4.5	5.3	3.8
feb	+0.15	1.67	0.91	0.76	-0.02	1.62	0.80	0.82	+0.10	1.56	0.83	0.73	4.7	4.7	5.4
mrt	+0.17	1.70	0.93	0.77	+0.27	2.11	1.19	0.92	-0.04	2.15	1.06	1.09	6.4	9.2	10.7
apr	+0.15	2.14	1.14	1.00	+0.20	2.34	1.27	1.07	+0.04	2.42	1.23	1.19	11.0	12.9	12.8
mei	+0.30	2.62	1.46	1.16	-0.04	2.85	1.41	1.44	+0.10	2.84	1.47	1.37	15.3	16.5	17.7
jun	+0.16	2.83	1.49	1.33	+0.13	2.46	1.30	1.17	+0.01	2.39	1.20	1.20	19.1	18.6	20.1
jul	+0.08	2.45	1.27	1.18	-0.05	2.25	1.10	1.15	+0.04	2.16	1.10	1.06	20.9	20.9	21.3
aug	-0.00	2.13	1.06	1.06	+0.06	1.85	0.96	0.90	-0.10	1.81	0.86	0.96	21.3	21.1	21.2
sep	-0.11	1.97	0.93	1.04	-0.10	1.92	0.91	1.01	-0.18	1.75	0.78	0.96	20.1	19.0	17.6
okt	-0.13	1.45	0.66	0.79	-0.16	1.41	0.63	0.78	-0.35	1.74	0.70	1.04	15.6	14.7	12.1
nov	-0.00	1.44	0.72	0.72	-0.16	1.60	0.72	0.88	-0.14	1.65	0.76	0.90	10.8	8.9	8.0
dec	-0.16	1.59	0.72	0.87	-0.11	1.91	0.90	1.01	+0.03	1.77	0.90	0.87	7.0	5.8	5.2
	Δ % van dagelijkse 14h-relatieve vochtigheid U												$\overline{U}$%		
jan	-0.2	10.6	5.2	5.4	+0.2	10.6	5.4	5.2	0.0	9.0	4.5	4.5	85	85	84
feb	-0.5	10.9	5.2	5.7	+0.2	15.4	5.8	5.6	-0.7	11.5	5.4	6.1	83	82	78
mrt	-0.1	12.7	6.3	6.4	-0.1	13.5	6.7	6.8	-0.4	13.8	6.7	7.1	76	72	69
apr	0.0	14.6	7.3	7.3	-0.8	14.4	6.8	7.6	+0.3	13.8	7.1	6.7	73	65	66
mei	-1.1	14.8	6.9	7.9	+1.0	14.2	7.6	6.6	+0.1	14.5	7.3	7.2	64	65	66
jun	-0.2	13.1	6.5	6.6	0.0	13.5	6.7	6.8	-0.2	12.6	6.2	6.4	65	68	67
jul	+0.6	13.2	6.9	6.3	+0.1	12.7	6.4	6.3	-0.4	11.6	5.6	6.0	69	70	67
aug	-0.1	11.1	5.5	5.6	-0.6	13.1	6.2	6.9	+0.3	11.5	5.9	5.6	69	69	66
sep	-0.1	12.2	6.1	6.1	+0.3	11.9	6.1	5.8	+0.5	12.5	6.5	6.0	68	69	72
okt	-0.2	12.0	5.9	6.1	+0.2	11.4	5.8	5.6	+0.5	11.9	6.2	5.7	72	75	77
nov	+0.3	12.4	6.4	6.0	+0.6	10.5	5.5	5.0	0.0	10.4	5.2	5.2	81	84	85
dec	+0.1	10.7	5.4	5.3	-0.2	9.5	4.7	4.8	0.0	8.6	4.3	4.3	85	87	88

Toelichting: I = 1e dekade; II = 2e dekade; III = 3e dekade

Δ = interdiurne verandering; $\Delta_+ \geqslant 0$ en $\Delta_- < 0$

N = aantal Δ's; N = 26 x 10 (Na en Ge) of 20 x 10 (Gr); N_+ = aantal Δ_+'s; N_- = aantal Δ_-'s

$N = N_+ + N_-$; vanwege de symmetrie der Δ-verdeling is $N_+ = N_- = \frac{1}{2}N$

$\overline{\Delta} = \frac{1}{N}\Sigma\,\Delta$; $\overline{|\Delta|} = \frac{1}{N}\,\Sigma\,\overline{|\Delta|}$; $\overline{\Delta}_+ = \frac{1}{N_+}\,\Sigma\,\Delta_+$; $\overline{\Delta}_- = \frac{1}{N_-}\,\Sigma\,\Delta_-$

Tabel 3b

Gemert; Δ van dagelijkse minimumtemperatuur t

	I				II				III				$\bar{t}$ °C		
	$\bar{\Delta}$	$\overline{\lvert\Delta\rvert}$	$\frac{1}{2}\bar{\Delta}_+$	$\frac{1}{2}\bar{\Delta}_-$	$\bar{\Delta}$	$\overline{\lvert\Delta\rvert}$	$\frac{1}{2}\bar{\Delta}_+$	$\frac{1}{2}\bar{\Delta}_-$	$\bar{\Delta}_-$	$\overline{\lvert\Delta\rvert}$	$\frac{1}{2}\bar{\Delta}$	$\frac{1}{2}\bar{\Delta}$	I	II	III
jan	-0.10	2.28	1.09	1.19	+0.05	2.32	1.19	1.13	-0.09	2.23	1.07	1.16	-0.6	-0.4	-1.8
feb	+0.17	2.47	1.32	1.15	-0.06	2.48	1.21	1.27	+0.06	2.48	1.27	1.21	-0.7	-1.0	-1.0
mrt	+0.07	1.98	1.02	0.96	+0.02	2.37	1.30	1.07	-0.03	2.41	1.19	1.22	-0.1	1.5	2.8
apr	+0.16	2.42	1.29	1.13	-0.05	2.45	1.20	1.25	+0.12	2.41	1.27	1.14	3.9	4.4	4.6
mei	+0.19	2.31	1.25	1.06	+0.06	2.52	1.29	1.23	+0.20	2.42	1.31	1.11	6.2	7.5	9.0
jun	+0.07	2.48	1.27	1.21	+0.02	2.70	1.36	1.34	+0.06	2.36	1.21	1.15	10.0	10.2	11.8
jul	+0.10	2.39	1.24	1.15	+0.04	2.28	1.16	1.12	+0.01	2.31	1.16	1.15	12.5	12.8	12.9
aug	+0.03	2.15	1.06	1.09	+0.04	2.20	1.12	1.08	-0.12	2.13	1.01	1.12	13.0	12.8	12.1
sep	-0.15	2.29	1.07	1.22	-0.07	2.47	1.20	1.27	-0.16	2.51	1.18	1.33	11.1	10.5	9.2
okt	-0.16	2.59	1.22	1.37	-0.03	2.69	1.33	1.36	-0.31	2.50	1.09	1.41	7.4	6.9	5.0
nov	+0.07	2.49	1.28	1.21	-0.10	2.16	1.03	1.13	-0.08	2.40	1.16	1.24	4.2	3.2	2.1
dec	-0.19	2.17	0.99	1.18	-0.09	2.48	1.20	1.28	-0.01	2.43	1.21	1.22	1.5	0.4	-0.5
	Δ van dagelijkse maximumtemperatuur T												$\bar{T}$		
jan	+0.04	1.90	0.97	0.93	+0.06	2.20	1.13	1.07	-0.04	2.01	0.98	1.02	4.1	5.1	3.9
feb	+0.16	1.92	1.04	0.88	-0.05	1.96	0.96	1.01	+0.12	1.91	1.02	0.89	5.2	4.9	5.8
mrt	+0.19	2.27	1.23	1.04	+0.33	2.44	1.38	1.06	-0.02	2.40	1.19	1.21	7.6	10.6	12.4
apr	+0.24	2.28	1.26	1.02	+0.18	2.49	1.33	1.15	-0.04	2.80	1.38	1.42	12.8	15.1	15.0
mei	+0.40	2.87	1.64	1.23	+0.09	2.88	1.48	1.39	-0.07	2.60	1.27	1.33	17.5	19.1	20.0
jun	+0.20	2.68	1.44	1.24	+0.11	2.54	1.32	1.22	+0.09	2.59	1.34	1.25	21.5	21.2	22.7
jul	-0.01	2.82	1.40	1.41	-0.06	2.43	1.19	1.24	-0.01	2.37	1.18	1.19	23.3	22.7	23.6
aug	-0.03	2.43	1.20	1.23	+0.03	2.20	1.12	1.09	-0.09	2.18	1.05	1.13	23.4	23.0	22.9
sep	-0.11	2.18	1.04	1.14	-0.15	2.20	1.02	1.18	-0.18	2.08	0.95	1.13	21.6	19.9	18.4
okt	-0.20	1.97	0.88	1.09	-0.14	1.97	0.94	1.03	-0.45	2.03	0.97	1.24	16.3	15.1	12.3
nov	-0.02	1.99	0.99	1.01	-0.15	1.91	0.90	1.01	-0.17	1.92	0.88	1.05	10.5	8.8	7.5
dec	-0.16	1.86	0.85	1.01	-0.12	2.04	0.96	1.08	-0.05	2.00	0.97	1.03	6.5	5.3	4.5
	Δ van dagelijkse 14h-relatieve vochtigheid U												$\bar{U}$ %		
jan	-0.6	9.6	4.5	5.1	+0.2	9.9	5.1	4.8	-0.3	9.8	4.7	5.1	83	83	78
feb	-0.2	11.9	5.9	6.0	-0.1	12.5	6.2	6.3	-0.5	11.7	5.6	6.1	78	77	73
mrt	-0.1	14.7	7.3	7.4	-0.6	14.5	6.9	7.6	-0.2	15.3	7.6	7.7	69	64	61
apr	-0.3	15.1	7.4	7.7	-0.7	15.4	7.4	8.0	+0.9	14.4	7.7	6.7	64	58	58
mei	-0.9	15.2	7.1	8.1	-0.0	13.0	6.5	6.5	+1.0	13.8	7.4	6.4	58	56	61
jun	-0.6	13.2	6.3	6.9	-0.4	14.8	7.2	7.6	+0.2	13.8	7.0	6.8	59	61	60
jul	+0.6	13.2	6.9	6.3	+0.4	14.7	7.5	7.2	-0.5	12.1	5.8	6.3	62	66	61
aug	+0.1	12.9	6.5	6.4	-0.1	12.5	6.2	6.3	+0.3	12.9	6.6	6.3	62	63	62
sep	+0.2	12.8	6.5	6.3	+0.6	13.0	6.8	6.2	+0.1	13.2	6.7	6.5	65	69	71
okt	-0.2	12.2	6.0	6.2	+0.3	11.9	6.1	5.8	+0.4	11.6	6.0	5.6	70	73	76
nov	-0.1	10.1	5.0	5.1	+0.6	9.5	5.1	4.4	-0.2	9.6	4.7	4.9	79	82	82
dec	+0.6	8.8	4.7	4.1	-0.1	8.7	4.3	4.4	+0.2	8.1	4.2	3.9	84	83	85

Zie legenda bij Naaldwijk

Tabel 3c

Groningen: Δ van dagelijkse minimumtemperatuur t

	I				II				III				$\bar{t}$ °C		
	$\bar{\Delta}$	$\overline{\lvert\Delta\rvert}$	$\frac{1}{2}\bar{\Delta}_+$	$\frac{1}{2}\bar{\Delta}_-$	$\bar{\Delta}$	$\overline{\lvert\Delta\rvert}$	$\frac{1}{2}\bar{\Delta}_+$	$\frac{1}{2}\bar{\Delta}_-$	Δ	$\overline{\lvert\Delta\rvert}$	$\frac{1}{2}\bar{\Delta}_+$	$\frac{1}{2}\bar{\Delta}_-$	I	II	III
jan	-0.22	2.16	0.97	1.19	+0.07	1.98	1.03	0.96	-0.09	2.03	0.97	1.06	-0.4	-0.6	-2.2
feb	+0.21	2.03	1.12	0.91	+0.06	1.82	0.94	0.88	-0.02	1.93	0.96	0.97	-0.6	0.2	-0.4
mrt	+0.05	1.99	1.02	0.97	+0.27	1.98	1.13	0.85	-0.03	1.66	0.81	0.84	-0.3	1.6	2.8
apr	+0.15	2.00	1.07	0.93	-0.01	1.93	0.96	0.97	+0.11	1.88	1.00	0.89	3.8	5.0	5.3
mei	+0.17	1.90	1.04	0.86	+0.08	1.86	0.97	0.89	+0.21	2.00	1.11	0.89	6.4	8.1	9.5
jun	+0.12	1.94	1.03	0.91	-0.02	2.06	1.02	1.04	+0.17	1.89	1.03	0.86	10.5	10.7	12.3
jul	+0.08	1.76	0.92	0.84	-0.05	1.67	0.86	0.81	-0.00	1.68	0.84	0.84	12.9	13.4	13.5
aug	-0.04	1.81	0.89	0.92	-0.01	1.83	0.91	0.92	-0.08	1.80	0.86	0.94	13.8	13.1	13.1
sep	-0.18	2.04	0.93	1.11	+0.01	1.71	0.86	0.85	-0.14	1.90	0.88	1.02	11.9	11.3	9.8
okt	-0.15	2.09	0.97	1.12	-0.11	1.93	0.91	1.02	-0.34	2.13	0.90	1.23	8.5	7.9	5.2
nov	+0.10	2.05	1.08	0.98	-0.14	1.64	0.75	0.89	-0.10	1.97	0.93	1.04	4.7	3.4	2.9
dec	-0.17	1.71	0.77	0.94	-0.20	2.01	0.91	1.11	+0.11	2.27	1.19	1.08	1.5	0.3	-0.4
	Δ van dagelijkse maximumtemperatuur T												$\bar{T}$ °C		
jan	-0.11	1.89	0.89	1.00	+0.05	1.99	1.02	0.97	-0.03	1.77	0.87	0.90	3.7	3.7	2.5
feb	+0.16	1.62	0.89	0.73	+0.12	1.65	0.88	0.77	+0.03	1.98	1.01	0.98	4.0	4.9	5.1
mrt	+0.06	1.83	0.94	0.89	+0.42	2.24	1.33	0.91	+0.02	2.16	1.09	1.07	5.2	8.8	10.6
apr	+0.13	2.33	1.23	1.10	+0.22	2.48	1.35	1.13	-0.18	2.62	1.22	1.40	11.1	13.2	13.1
mei	+0.47	2.80	1.63	1.17	+0.16	3.11	1.64	1.47	+0.03	2.94	1.49	1.46	15.2	17.2	18.2
jun	+0.19	2.83	1.51	1.32	+0.01	2.60	1.31	1.30	+0.02	2.74	1.38	1.36	19.7	19.2	20.7
jul	+0.11	2.83	1.47	1.36	-0.07	2.32	1.13	1.19	+0.05	2.38	1.21	1.16	21.5	21.1	21.8
aug	-0.08	2.31	1.11	1.20	+0.12	2.00	1.06	0.94	-0.17	2.17	1.00	1.17	21.9	21.4	21.6
sep	-0.10	1.97	0.94	1.04	-0.15	2.14	0.99	1.15	-0.09	1.86	0.88	0.98	19.9	18.8	17.3
okt	-0.22	1.54	0.66	0.88	-0.15	1.69	0.77	0.92	-0.37	1.70	0.66	1.03	15.1	14.2	11.2
nov	-0.00	1.89	0.94	0.95	-0.15	1.69	0.77	0.92	-0.21	1.78	0.78	1.00	9.7	7.8	7.4
dec	-0.14	1.66	0.76	0.90	-0.07	1.77	0.85	0.92	+0.01	1.87	0.94	0.93	5.7	4.6	3.9
	Δ van dagelijkse 14h-relatieve vochtigheid U												$\bar{U}$ %		
jan	-0.6	9.6	4.5	5.1	+0.3	10.3	5.3	5.0	-0.4	9.6	4.6	5.0	84	84	80
feb	+0.3	11.1	5.7	5.4	-0.7	11.8	5.6	6.2	-0.4	14.8	7.2	7.6	81	77	73
mrt	+0.1	13.2	6.6	6.6	-0.6	14.5	7.0	7.5	-0.2	14.2	7.0	7.2	72	67	63
apr	-0.1	15.1	7.5	7.6	-0.8	14.8	7.0	7.8	+1.2	16.2	8.7	7.5	64	61	61
mei	-1.9	13.8	5.9	7.9	+0.7	15.9	8.3	7.6	+0.5	13.7	7.1	6.6	56	58	61
jun	+0.1	13.9	7.0	6.9	-0.1	15.5	7.7	7.8	+0.1	14.9	7.5	7.4	58	62	61
jul	+0.2	14.3	7.3	7.0	+0.4	14.1	7.3	6.8	-0.3	12.8	6.2	6.6	62	66	63
aug	+0.2	13.0	6.6	6.4	-0.8	12.9	6.1	6.8	+0.6	13.2	6.9	6.3	64	63	62
sep	-0.6	14.0	6.7	7.3	+1.3?	15.2	8.3	6.9	-0.1	13.1	6.5	6.6	64	68	68
okt	0.0	13.6	6.8	6.8	+0.1	12.3	6.2	6.1	+0.2	13.5	6.9	6.6	71	74	74
nov	0.0	11.2	5.6	5.6	+0.5	10.8	5.7	5.1	+0.1	9.9	5.0	4.9	80	84	84
dec	0.0	9.4	4.7	4.7	-0.4	8.9	4.3	4.6	+0.5	8.3	4.4	3.9	84	83	87

Zie legenda bij Naaldwijk

Tabel 6a

Grootste veranderingen van dag tot dag van de minimum temperatuur t

		Naaldwijk (1931 t/m 1956				Gemert (1931 t/m 1956)				Groningen (1931 t/m 1950)			
		+		−		+		−		+		−	
jan	I	8.2	10- 1-'53	5.4	10- 1-'42	10.5	3- 1-'32	7.9	10- 1-'42	9.7	3- 1-'32	7.9	2- 1-'40
	II	8.3	15- 1-'40	9.6	20- 1-'32	9.3	15- 1-'40	8.0	13- 1 '46	7.9	14- 1-'40	9.3	16- 1-'40
	III	13.2	28- 1-'42	8.1	22- 1-'40	14.8	28- 1-'42	7.3	27- 1-'42	9.3	28- 1-'45	7.7	21- 1-'45
feb	I	9.3	5- 2-'40	7.8	31- 1-'56	13.0	8- 2-'41	9.5	31- 1-'56	9.4	8- 2-'41	6.7	4- 2-'44
	II	11.7	15- 2-'40	6.4	16- 2-'48	10.1	15- 2-'40	12.5	14- 2-'55	9.2	15- 2-'40	7.6	13- 2-'40
	III	10.3	27- 2-'56	6.7	28- 2-'32	8.9	25- 2-'34	8.1	23- 2-'56	9.6	21- 2-'40	7.3	28- 2-'32
mrt	I	8.4	8- 3-'48	6.4	11- 3-'39	10.7	8- 3-'48	7.8	7- 3-'50	8.5	8- 3-'48	6.7	8- 3-'47
	II	7.4	17- 3-'35	6.7	14- 3-'52	9.7	13- 3-'49	10.5	14- 3-'40	7.5	12- 3-'31	7.4	14- 3-'40
	III	8.3	25- 3-'43	7.4	24- 3-'48	9.2	29- 3-'53	9.6	25- 3-'50	6.6	28- 3-'31	5.1	31- 3-'31
apr	I	7.7	9- 4-'44	5.6	9- 4-'47	8.4	9- 4-'44	8.3	7- 4-'49	8.6	1- 4-'40	5.1	5- 4-'32
	II	8.8	12- 4-'53	5.9	19- 4-'46	8.7	14- 4-'46	8.3	19- 4-'46	6.9	16- 4-'33	7.6	13- 4-'33
	III	10.0	22- 4-'43	6.1	21- 4-'49	8.6	22- 4-'43	8.5	21- 4-'49	8.2	29- 4-'32	5.5	25- 4-'50
mei	I	7.6	5- 5-'48	7.7	3- 5-'41	8.6	6- 5-'42	9.5	3- 5-'41	6.4	6- 5-'42	6.0	3- 5-'50
	II	11.3	20-5-'43	6.1	11- 5-'53	12.0	17- 5-'41	7.8	18- 5-'41	5.8	12- 5-'44	6.8	14- 5-'45
	III	6.9	24- 5-'31	6.2	27- 5-'39 21- 5-'43	9.1	24- 5-'55	10.0	22- 5-'37	7.9	24- 5-'31	7.0	28- 5-'47
jun	I	7.1	7- 6-'39	7.2	8- 6-'39	7.8	3- 6-'35	10.3	1- 6-'37	6.2	4- 6-'37	5.7	1- 6-'37
	II	6.0	16- 6-'46	6.8	16- 6-'55	7.1	19- 6-'43	8.2	16- 6-'55	5.5	18- 6-'36	5.7	16- 6-'50
	III	13.1	28- 6-'52	6.9	30- 6-'31	7.6	27- 6-'32	10.2	30- 6-'47	7.1	21- 6-'35	5.7	23- 6-'31 30- 6-'31
jul	I	8.5	2- 7-'54	6.4	2- 7-'42 3- 7-'52	7.3	5- 7-'43	9.0	2- 7-'49	5.8	8- 7-'38	5.3	5- 7-'41
	II	6.5	18- 7-'43	7.1	17- 7-'43	8.5	18- 7-'43 12- 7-'49	8.9	15- 7-'49	5.2	18- 7-'36	6.6	17- 7-'45
	III	7.0	28- 7-'35	6.1	24- 7-'34	8.4	25- 7-'53	8.0	29- 7-'49	6.8	22- 7-'49	6.0	28- 7-'46
aug	I	6.2	5- 8-'53	6.0	1- 8-'40	10.0	5- 8-'53	7.5	4- 8-'37	6.4	7- 8-'32	5.4	7- 8-'45
	II	7.0	20- 8-'49	5.7	14- 8-'31	8.1	20- 8-'49	8.7	17- 8-'53	4.9	16- 8-'41	6.7	13- 8-'44
	III	5.9	27- 8-'32	7.0	28- 8-'32	8.6	27- 8-'32	9.4	29- 8-'52	7.1	26- 8-'41	5.6	21- 8-'42
sep	I	7.9	8- 9-'34	7.7	6- 9-'34	7.7	10- 9-'43	8.6	9- 9-'37	7.0	4- 9-'34	5.8	7- 9-'43
	II	9.1	13- 9-'35	5.9	15- 9-'38	8.8	13- 9-'35	9.1	17- 9-'44	6.8	16- 9-'37	5.9	17- 9-'44
	III	8.7	25- 9-'31	7.7	27- 9-'31	9.9	25- 9-'31	7.5	25- 9-'39	8.2	29- 9-'39	5.4	21- 9-'35

Vervolg tabel 6a

		+		-		+		-		+		-	
okt	I	10.5	6-10-'45	8.7	5-10-'45	9.9	6-10-'45	8.2	2-10-'42	8.1	1-10-'43	5.8	5-10-'36 9-10-'50
	II	10.2	18-10-'51	7.6	12-10-'41	9.5	15-10-'35	8.7	16-10-'55	9.2	18-10-'43	6.6	13-10-'40
	III	11.3	29-10-'31	6.7	24-10-'46	12.3	29-10-'31	7.3	27-10-'48	7.0	29-10-'31	6.9	26-10-'48
nov	I	10.1	1-11-'40	7.6	10-11-'33	9.8	1-11-'40	6.8	8-11-'37	8.6	1-11-'40	5.7	5-11-'42
	II	10.4	17-11-'41	6.3	14-11-'43	11.8	17-11-'41	7.1	13-11-'54	9.1	17-11-'41	4.6	14-11-'41
	III	10.7	26-11-'56	8.1	22-11-'39	15.0	21-11-'47	8.2	24-11-'47	13.0	21-11-'47	6.8	24-11-'47
dec	I	7.5	4-12-'31 7-12-'33 8-12-'43	6.6	5-12-'48	10.1	4-12-'41	6.7	4-12-'43	7.4	7-12-'33	5.7	3-12-'33
	II	10.1	20-12-'47	8.2	12-12-'55	10.6	19-12-'33	7.7	18-12-'55	9.3	19-12-'43	8.9	14-12-'46
	III	9.7	29-12-'48	5.9	29-12-'31	11.9	25-12-'46	8.0	29-12-'50	10.9	21-12-'50	7.1	28-12-'41

Tabel 6b

Grootste veranderingen van dag tot dag
van de maximum temperatuur T

		Naaldwijk (1931 t/m 1956)				Gemert (1931 t/m 1956)				Groningen (1931 t/m 1950)			
		+		−		+		−		+		−	
jan	I	8.2	2- 1-'32	7.6	2- 1-'40	10.6	2- 1-'32	7.6	1- 1-'41	11.3	9- 1-'47	6.9	4- 1-'40
	II	9.0	15- 1-'39	9.2	17- 1-'40	9.2	20- 1-'36	11.2	12- 1-'42	8.0	13- 1-'44	10.4	17- 1-'40
	III	10.9	28- 1-'42	10.1	27- 1-'40	13.1	23- 1-'40	9.1	27- 1-'42	12.1	28- 1-'42	6.1	26- 1-'37
feb	I	8.3	3- 2-'56	7.8	1- 2-'56	9.3	31- 1-'37	9.8	1- 2-'56	6.9	7- 2-'41	5.0	9- 2-'40
	II	7.9	11- 2-'35	5.7	16- 2-'56	6.8	13- 2-'45	7.2	11- 2-'33	9.2	11- 2-'35	4.8	10- 2-'40
	III	8.1	28- 2-'41	4.9	25- 2-'37	9.2	28- 2-'41	5.7	21- 2-'43	9.2	25- 2-'38 28- 2-'41	6.3	26- 2-'47
mrt	I	8.1	9- 3-'54	7.1	11- 3-'36	8.3	2- 3-'52	8.5	11- 3-'41	11.0	6- 3-'48	6.3	11- 3-'41
	II	9.2	14- 3-'42	6.6	21- 3-'52	8.8	18- 3-'40	8.9	16- 3-'45	8.7	14- 3-'42	6.3	16- 3-'45
	III	7.8	25- 3-'53	7.6	30- 3-'33	7.1	31- 3-'56	7.5	28- 3-'49	7.5	28- 3-'49	7.3	27- 3-'49
apr	I	9.8	2- 4-'49	9.5	5- 4-'46	6.9	8- 4-'44	7.4	9- 4-'47	7.4	2- 4-'46	10.9	6- 4-'46
	II	9.8	11- 4-'33	8.5	17- 4-'46	8.2	11- 4-'34 16- 4-'40	9.9	17- 4-'46	10.6	20- 4-'40	10.4	19- 4-'49
	III	11.2	30- 4-'40	8.8	25- 4-'40	8.5	26- 4-'46	9.8	21- 4-'52	9.9	26- 4-'46	9.7	23- 4-'48
mei	I	7.2	3- 5-'47	9.1	5- 5-'37	12.4	10- 5-'52	14.6	9- 5-'52	11.1	3- 5-'33	9.5	5- 5-'37
	II	10.7	15- 5-'32	13.7	14- 5-'45	11.1	14- 5-'40	11.9	14- 5-'45	10.2	14- 5-'40	11.5	14- 5-'45
	III	9.8	28- 5-'56	9.9	27- 5-'47	10.0	28- 5-'56	11.1	30- 5-'54	11.5	28- 5-'31	9.5	29- 5-'31
jun	I	9.8	8- 6-'48	12.0	5- 6-'47	8.5	5- 6-'36	9.9	6- 6-'41	10.1	10- 6-'35	11.4	8- 6-'39
	II	10.2	17- 6-'36 20- 6-'43	10.5	14- 6-'36	9.9	20- 6-'43	9.1	12- 6-'35	10.2	20- 6-'43	6.9	12- 6-'37
	III	8.9	22- 6-'35	9.5	25- 6-'42	7.9	22- 6-'35	8.9	25- 6-'42	7.9	24- 6-'47	9.5	25- 6-'42
jul	I	8.9	6- 7-'33	10.5	5- 7-'46	7.5	6- 7-'33 7- 7-'38	10.9	10- 7-'56	9.3	6- 7-'33	9.2	8- 7-'44
	II	7.3	15- 7-'43	11.7	19- 7-'34	9.6	13- 7-'56	8.4	16- 7-'37	7.0	13- 7-'50	8.5	17- 7-'37
	III	9.8	22- 7-'51	10.6	29- 7-'33	8.2	30- 7-'32	9.2	29- 7-'33	8.8	24- 7-'46	9.6	26- 7-'31
aug	I	7.3	6- 8-'32	7.8	1- 8-'51	6.7	5- 8-'49	7.7	9- 8-'37	8.3	4- 8-'38	9.5	9- 8-'37
	II	7.1	19- 8-'32	8.1	20- 8-'42	7.2	12- 8-'53	8.0	20- 8-'42	7.1	19- 8-'32	6.3	13- 8-'35
	III	5.8	23- 8-'44	8.5	22- 8-'39	7.1	29- 8-'36	7.3	29- 8-'34	7.7	22- 8-'44	10.9	22- 8-'43
sep	I	9.0	7- 9-'34	11.2	6- 9-'49	7.7	7- 9-'34	8.3	9- 9-'34	7.5	3- 9-'42	7.5	6- 9-'49
	II	6.4	15- 9-'46	7.6	17- 9-'47	6.9	16- 9-'32	8.4	17- 9-'47	7.8	13- 9-'43	9.3	17- 9-'47
	III	8.8	28- 9-'34	6.8	24- 9-'55	7.7	28- 9-'35 29- 9-'43	8.3	25- 9-'43 24- 9-'44	10.0	25- 9-'41	6.4	29- 9-'33 23- 9-'35

Vervolg tabel 6b

		+		-		+		-		+		-	
okt	I	6.2	7-10-'42	6.8	7-10-'31	7.6	7-10-'42	8.7	8-10-'31	7.0	6-10-'31	5.5	7-10-'31
	II	8.2	17-10-'43	6.8	16-10-'43	11.6	17-10-'43	8.6	13-10-'55	7.3	20-10-'45	6.8	1-10-'42 13-10-'43
	III	7.3	22-10-'33	6.7	27-10-'49	7.1	22-10-'33	8.4	27-10-'49	6.4	26-10-'37	6.3	27-10-'49
nov	I	6.9	3-11-'55	6.3	4-11-'37	6.5	6-11-'41	8.1	4-11-'37	6.5	3-11-'48	5.7	6-11-'43
	II	8.7	19-11-'37	7.8	17-11-'37	10.0	20-11-'47	7.9	15-11-'41	8.7	19-11-'37	5.4	17-11-'37
	III	10.2	25-11-'56	5.2	22-11-'37	7.9	28-11-'50	7.5	29-11-'48	10.2	28-11-'50	5.3	26-11-'44
dec	I	8.4	2-12-'40	5.9	8-12-'52	10.1	4-12-'31	6.7	5-12-'43	10.0	4-12-'31	5.7	5-12-'37 9-12-'49
	II	7.0	18-12-'43	7.7	16-12-'54	8.4	11-12-'45	6.9	16-12-'54	8.3	14-12-'43	7.0	18-12-'39
	III	10.1	27-12-'38	6.9	21-12-'40	9.4	27-12-'38	6.8	25-12-'42	8.0	27-12-'38	7.1	26-12-'41

Tabel 6c

Grootste veranderingen van dag tot dag van de relatieve vochtigheid U te 14 uur

		Naaldwijk (1931 t/m 1956)				Gemert (1931 t/m 1956)				Groningen (1931 t/m 1950)			
		+		−		+		−		+		−	
jan	I	36	1- 1-'39	47	2- 1-'56	30	2- 1-'50	45	2- 1-'56	42	5- 1-'50	41	4- 1-'50
	II	32	18- 1-'42 11- 1-'48	36	17- 1-'40 16- 1-'52	43	19- 1-'33	40	17- 1-'40 19- 1-'54	51	14- 1-'40	36	14- 1-'31
	III	34	30- 1-'54	35	22- 1-'47	53	22- 1-'50	41	24- 1-'54	41	28- 1-'33	32	24- 1-'43
feb	I	39	31- 1-'50	58	31- 1-'56	46	31- 1-'50	55	31- 1-'56	58	8- 2-'39	31	2- 2-'34
	II	55	15- 2-'34	39	14- 2-'34	39	18- 2-'50	50	10- 2-'32	50	15- 2-'34	57	14- 2-'34
	III	33	27- 2-'38 25- 2-'44	43	1- 3-'54	47	1- 3-'37	39	24- 2-'55	37	20- 2-'50	40	25- 2-'38
mrt	I	41	4- 3-'40	46	7- 3-'56	53	3- 3-'32	40	6- 3-'41	47	8- 3-'48	41	2- 3-'40
	II	52	11- 3-'43 21- 3-'31	46	18- 3-'50	54	18- 3-'53	53	17- 3-'56	48	21- 3-'46	49	14- 3-'38
	III	55	30- 3-'33	54	22- 3-'42	49	30- 3-'40	56	26- 3-'31	47	25- 3-'38	57	27- 3-'48
apr	I	68	7- 4-'34	64	6- 4-'34	51	3- 4-'46	59	5- 4-'54	57	4- 4-'31	48	6- 4-'40
	II	55	12- 4-'49	53	16- 4-'45	68	12- 4-'49	52	18- 4-'44	61	17- 4-'31	61	13- 4-'40
	III	50	28- 4-'46	50	30- 4-'51	58	25- 4-'49	39	21- 4-'34	58	28- 4-'49	42	23- 4-'35
mei	I	57	4- 5-'37	44	9- 5-'35	44	7- 5-'34 5- 5-'49	45	6- 5-'50	53	5- 5-'37	50	2- 5-'41
	II	48	15- 5-'54	50	15- 5-'42	50	17- 5-'34	49	19- 5-'50	67	11- 5-'42	53	13- 5-'32
	III	59	29- 5-'41	52	22- 5-'47	66	25- 5-'56	37	25- 5-'50 26- 5-'51	53	30- 5-'46	62	23- 5-'31
jun	I	58	4- 6-'55	45	9- 6-'41	53	7- 6-'55	46	1- 6-'42	50	9- 6-'38	54	4- 6-'31
	II	52	16- 6-'54	42	12- 6-'39	50	19- 6-'47	52	16- 6-'52	56	15- 6-'48	47	13- 6-'43
	III	43	25- 6-'54	42	24- 6-'50	55	29- 6-'55	57	28- 6-'53	52	22- 6-'36	50	25- 6-'39
jul	I	59	3- 7-'50	38	5- 7-'51	50	1- 7-'32	45	7- 7-'38	47	6- 7-'32	43	7- 7-'32
	II	40	13- 7-'53	35	18- 7-'46	52	14- 7-'50	49	17- 7-'50	58	14- 7-'34	45	13- 7-'50
	III	41	25- 7-'31	47	25- 7-'55	50	27- 7-'53	47	28- 7-'53	44	27- 7-'41	38	26- 7-'42
aug	I	42	8- 8-'37	39	1- 8-'53	49	6- 8-'52	37	3- 8-'34 2- 8-'51 1- 8-'53	49	8- 8-'45	37	5- 8-'34
	II	46	13- 8-'37	37	20- 8-'31	50	16- 8-'50	49	17- 8-'50	43	16- 8-'33	43	17- 8-'42 17- 8-'49
	III	48	30- 8-'33 27- 8-'35	37	25- 8-'39	50	24- 8-'49	42	25- 8-'50	53	21- 8-'32	50	30- 8-'34

Vervolg tabel 6c

		+		-		+		-		+		-	
sep	I	49	6- 9-'49	37	6- 9-'35	52	9- 9-'34	39	5- 9-'42 1- 9-'54	41	7- 9-'44	44	4- 9-'32
	II	36	13- 9-'35 14- 9-'50	41	15- 9-'38	46	12- 9-'55	47	17- 9-'55	50	20- 9-'43	47	12- 9-'45
	III	53	30- 9-'54	47	27- 9-'41	45	25- 9-'43	52	21- 9-'48	43	24- 9-'31 22- 9-'35 25- 9-'37	43	21- 9-'35 29- 9-'42
okt	I	40	3-10-'40	37	8-10-'38	44	5-10-'54	40	3-10-'50	47	7-10-'31	50	6-10-'31
	II	51	17-10-'40	33	20-10-'41 18-10-'48 12-10-'45 15-10-'56	46	14-10-'50	35	20-10-'46	37	11-10-'39	37	14-10-'48 13-10-'50
	III	40	27-10-'40 27-10-'46	42	27-10-'53	36	29-10-'52	43	26-10-'49	45	27-10-'46	43	26-10-'49
nov	I	44	2-11-'36	40	1-11-'40	54	5-11-'31	27	8-11-'48 7-11-'52	36	6-11-'40	46	3-11-'34
	II	34	16-11-'54	39	17-11-'45	34	20-11-'44	30	13-11-'50	36	20-11-'47	33	16-11-'48
	III	35	22-11-'44	36	27-11-'32	31	29-22-'39	36	30-11-'48	38	29-11-'48	28	28-11-'40 23-11-'42
dec	I	37	5-12-'49	46	9-12-'45	36	5-12-'52	36	3-12-'33	33	7-12-'43	30	5-12-'40
	II	42	13-12-'49	38	15-12-'46	35	12-12-'47	30	13-12-'33	47	15-12-'33	61	13-12-'33
	III	35	28-12-'44	43	27-12-'44	37	30-12-'44	33	27-12-'42 29-12-'44	51	29-12-'41	41	28-12-'41