K O N I N K L I J 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. 66-2 R III 297-1966

Dr. C. Levert

Statistische beschouwingen bij de vraag hoe metingen van eenzelfde fysische grootheid, verricht met verschillende typen instrumenten en volgens verschillende meettechnieken tot één gemeenschappelijke basis herleid kunnen worden.

De Bilt, 1966

U.D.C.:
519.2
53.08

INHOUD

.-.-.-.-.

Tevens is dit een verslag van het K.N.M.I. colloquium d.d. 20 oktober 1964 en van een voordracht, d.d. 18 mei 1965, gehouden voor de Studiekring voor Statistische Techniek van het K.G.L. en de Landbouwkundige Sectie van de V.V.S.

1. INLEIDING

Tijdens de Derde Zitting van de C.I.M.O. (Commission for Instruments and Methods of Observation), 29 januari - 16 februari 1962, werd besloten een "Working Group on Radiation Measurements and Observations for General Use" op te richten [Res. 7 (CIMO III)] . Uit de "terms of reference" lichten wij hier, in verband met het onderhavige rapport, slechts de volgende twee taken: "to provide guidance for Members on comparisons of sunshine recorders and to study the results of comparisons reported by Members".

Als "Interim Reference Sunshine Recorder", (in ons rapport afgekort tot I.S.) werd gekozen de Campbell Stokes- zonneschijnautograaf, vervaardigd in overeenstemming met de door het Britse "Meteorological Office" gestelde eisen, en voorzien van diagrammen (stroken), die voldoen aan de door de Franse "Météorologie Nationale" geformuleerde voorschriften.

In het besluit van het Uitvoerend Comité wordt voorts aan alle Leden, die met andere instrumenten*) en volgens andere technieken de zonneschijnduur meten, gevraagd simultane metingen te willen verrichten met de nationale zonneschijnmeter (voortaan aangeduid met N.S.) en de genoemde I.S., teneinde tot herleidingscoëfficiënten (reductiefactoren) te komen, die toegepast zouden moeten worden op de reeds vroeger gepubliceerde en in de toekomst alsnog te publiceren, met de N.S. gemeten, maandtotalen, waardoor deze beter vergelijkbaar zullen zijn. De procedure (waarop uitvoerig in hoofdstuk 4 wordt ingegaan) is in het kort deze. Op elk station wordt een I.S. opgesteld naast de N.S. Iedere maand moeten dan door de nationale routine-dienst én de stroken van de N.S. én die van de I.S. uitgetrokken worden,enwel die van de N.S. volgens de ter plaatse gebruikelijke voorschriften en die van de I.S. volgens een internationaal voorschrift, dat als aanhangsel in Res. 23 (EC XIV) is opgenomen.

Een "Experts Group on International Comparison of Sunshine Recorders" werd, in overeenstemming met de boven onder a en b genoemde taken, belast met:

a) het geven van voorlichting ten aanzien van de door de Leden uit te voeren vergelijkingen;

b) de studie der door de Leden gerapporteerde metingen.

Van deze subgroep maken deel uit: de voorzitter R. Lamboley (F), de

---

*) De zonneschijnduur wordt gemeten in hoofdzaak met vier instrumenten: 1. Campbell- Stokes zonneschijnautograaf; 2. Jordan-zonneschijnautograaf; 3. Marvin-type en 4. Foster-type.

leden R.H. Collingbourne (U.K.), F.H. Mac Donald (in de loop van 1964 vervangen door A.N. Hill; beiden U.S.A.) en C. Levert (N). De voorzitter schreef in april 1964 een ontwerp-handleiding, waaraan Levert een "Beschouwing over de statistische aspecten met betrekking tot de berekening van herleidingscoëfficiënten" toevoegde. De hierin ontwikkelde aanpak voor de behandeling van de met zonneschijnmeters verkregen resultaten kan algemeen van nut zijn bij de bewerkingen van soortgelijke metingen met instrumenten, waarmede men eenzelfde ware, weliswaar onbekende, welgedefiniëerde waarde beoogt te vinden, doch waarbij zowel toevallige als systematische, van instrument en persoon afhankelijke, fouten onvermijdelijk zijn.

Gedacht wordt vooral aan instrumenten, die "registraties" (diagrammen) afleveren, die door een of meer personen, een of meer keren, "gelezen", "uitgemeten", "uitgetrokken", "geanalyseerd" moeten worden, teneinde getallen te verwerven, waarmede "verder gewerkt" kan worden.

2. TOEVALLIGE EN SYSTEMATISCHE FOUTEN

Gesteld persoon A trekt de N.S.-stroken van januari 1964, De Bilt, uit volgens het geldende K.N.M.I. voorschrift en vindt x uren zonneschijn in totaal. Een tweede uittrekking (onbeïnvloed door de eerste) levert zeker een ander getal; een derde uittrekking weer een ander getal (zulke "experimenten" zijn op het K.N.M.I. op uitgebreide schaal verricht), etc. Kortom: $\underline{x}$ is een stochastische of statistische variabele[1]. Anders gezegd: bij deze persoon, bij deze maand, bij dit voorschrift, behoort één bepaalde $\underline{x}$-populatie. De kansverdeling van $\underline{x}$ is in goede benadering normaal en dus gekenmerkt door slechts twee parameters, te weten het gemiddelde μ en de standaarddeviatie σ (variantie σ^2). Deze σ is omgekeerd evenredig met de nauwkeurigheid, waarmede A deze stroken uittrekt (velen noemen de σ kortweg de onnauwkeurigheid van de meting). Het stochastische verschil $\underline{\delta} = \underline{x} - \mu$, dat de toevallige fout heet, volgt dan eveneens een vrijwel normale kansverdeling. Er geldt $\mathcal{E}(\underline{\delta}) = 0$ en $\sigma(\underline{\delta}) = \sigma(\underline{x}) \equiv \sigma$. De bedoeling van de "meting" ("meting" $\equiv$ inbranding in strook + uitmeting van brandspoor) is een schatting te verkrijgen van de onbekende ware waarde $\mathring{\mu}$ van de zonneschijnduur te De Bilt in januari 1964. Men zou deze ware waarde kunnen definiëren als de totale duur, gedurende welke de zon in deze maand

1) Een stochastische of statistische variabele is een variabele, die een kansverdeling volgt. De letter wordt onderstreept: $\underline{x}$. Een bepaalde, numerieke waarde van $\underline{x}$ heet x. Het gemiddelde van $\underline{x}$ heet $\mathcal{E}(\underline{x})$, ook wel $\mu(\underline{x})$ of μ_x of, als geen verwarring mogelijk is, μ. De standaarddeviatie heet $\sigma(\underline{x})$ of σ_x of σ. In de steekproef duiden wij het gemiddelde en de standaarddeviatie aan met $\bar{x}$ en s_x (de variantie is het kwadraat van de standaarddeviatie).

de plaats, waar het instrument opgesteld is, bescheen met een intensiteit (gemeten loodrecht op de stralingsrichting) van bijv. tenminste 0,3 cal/min.cm^2. Essentiëel in ons betoog is, dat ondersteld wordt dat er aldus een welgedefiniëerde ware waarde is, ofschoon zij numeriek onbekend is.

Het in de inleiding ruw geschetste Project beoogt niet uit te maken of de I.S. een "goed" instrument of misschien wel "het beste" is, maar in de eerste plaats om de uitkomsten afkomstig van de N.S. met die van de I.S. als referentie-instrument te vergelijken.

Gewoonlijk zullen μ en $\mathring{\mu}$ verschillen. Dit constant[1] gedachte verschil, geschreven als $\varepsilon = \mu - \mathring{\mu}$, heet systematische fout.

Bij een tweede persoon B behoort bij dezelfde stroken (januari 1964, De Bilt) en dezelfde instructie een tweede x-populatie. Wij hangen derhalve aan de $\underline{x}$, μ, $\underline{\delta}$, ε, σ de index A of de index B en noteren:

$\underline{x}_A = \mu_A + \underline{\delta}_A = \mathring{\mu} + \varepsilon_A + \underline{\delta}_A$, met een σ_A

en $\underline{x}_B = \mu_B + \underline{\delta}_B = \mathring{\mu} + \varepsilon_B + \underline{\delta}_B$, met een σ_B

Als $\sigma_A = \sigma_B$ en $\varepsilon_A = \varepsilon_B$ (d.i. $\mu_A = \mu_B$) heten A en B volledig gelijk; als $\sigma_A \neq \sigma_B$ en $\varepsilon_A \neq \varepsilon_B$ (d.i. $\mu_A \neq \mu_B$) heten A en B volledig ongelijk. Van het "persoonseffect" kunnen wij spreken als $\mu_A \neq \mu_B$ (daarbij kan $\sigma_A = \sigma_B$ of $\sigma_A \neq \sigma_B$ zijn); van het herhalings (repetitie) effect voor A kunnen wij spreken als $\sigma_A \neq 0$.

Aldus behoort er bij ieder stel stroken, bij iedere persoon, bij ieder voorschrift, één bepaalde populatie, met kenmerkende parameters μ, σ en ε. Bij een "populatie" van personen - daarbij gedacht aan de potentiële aanwezigheid van een zeer groot (oneindig) aantal personen - behoort dan een populatie van μ's, een populatie van σ's en een populatie van ε's. De μ-populatie heeft een gemiddelde μ_μ en een standaarddeviatie σ_μ; bij de σ-populatie behoort een tweetal μ_σ en σ_σ en bij de ε-populatie een tweetal μ_ε en σ_ε.

Indien een andere instructie en (of) een andere assistenten-opleiding en (of) een ander instrument kan leiden tot een kleinere σ_μ en (of) een kleinere σ_σ, dan zouden de uitkomsten van de personen A, B, C, D etc. onderling "beter vergelijkbaar" genoemd kunnen worden. Dit is het geval in

1) De systematische fout zou een functie kunnen zijn van o.a. het niveau $\mathring{\mu}$ (zie ook in hoofdstuk 11). "Echt constant" behoeft de systematische fout dus niet te zijn.

fig. 1b, waarin de μ-waarden dichter bij elkaar liggen en de σ-waarden minder verschillen dan in fig. 1a. In fig. b ligt tevens de μ_μ dichter bij $\mathring{\mu}$ dan in fig. a, maar dat lijkt ons niet per se nodig om van "meer vergelijkbaar" te mogen spreken. In het Project in kwestie is het namelijk, zoals reeds gezegd, niet in de eerste plaats de bedoeling de ware waarde (die trouwens op meer dan één wijze gedefinieerd kan worden) dichter te benaderen, maar om een vaststelling van de huidige "<u>graad van vergelijkbaarheid</u>".

fig. 1 a

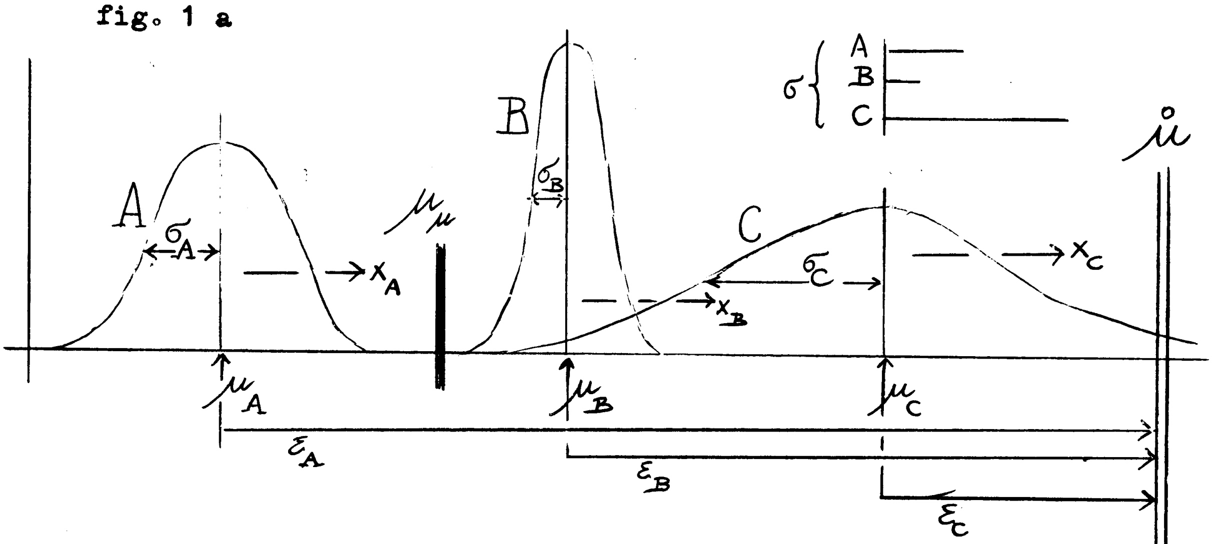
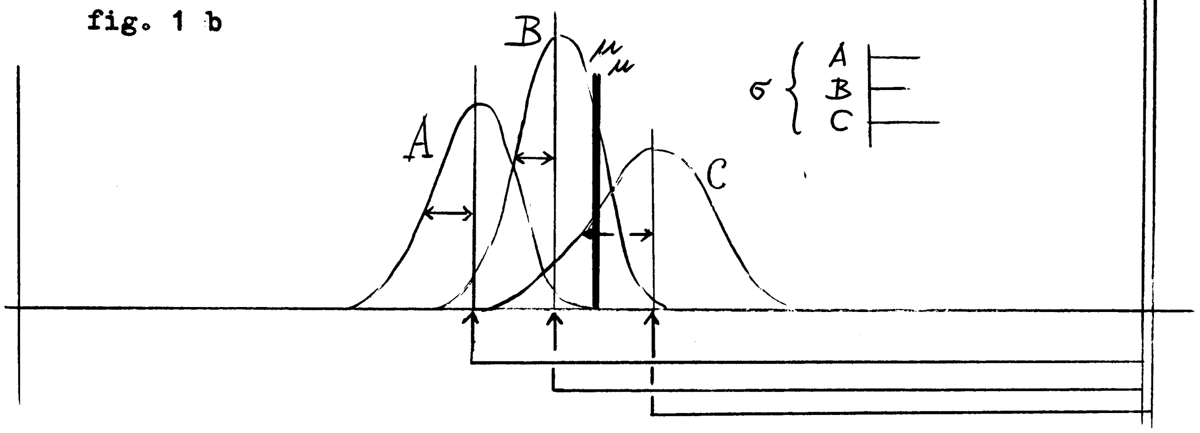


fig. 1 b



N.B. <u>Persoonseffecten</u> kunnen <u>afwezig</u> zijn ($\mu_A = \mu_B = \mu_C = \ldots$; zegge μ), terwijl <u>systematische fouten</u> gelijktijdig <u>aanwezig</u> zijn ($\varepsilon \neq 0$, d.i. $\mu \neq \mathring{\mu}$, zodat alle analisten dezelfde systematische fout hebben).

3. VERGELIJKBAARHEID

Het is de bedoeling van het Vergelijkingsproject coëfficiënten te vinden, waarmee de, met de N.S. gemeten, maandelijkse zonneschijnduren moeten worden vermenigvuldigd "voor een herleiding tot het I.R.S.R.-niveau", opdat zij hierdoor "beter vergelijkbaar" zullen worden. Wij willen nu het begrip "vergelijkbaar" preciseren.

Gesteld een onbekende, ware waarde $\overset{o}{\mu}_I$ wordt gemeten met zeker instrument, volgens zekere techniek, terwijl een andere daarvan verschillende, eveneens onbekende, ware waarde $\overset{o}{\mu}_{II}$ met eenzelfde of een ander instrument, volgens eenzelfde of een andere techniek gemeten wordt. Een voorbeeld: gesteld $\overset{o}{\mu}_I$ stelt de ware, totale zonneschijnduur in januari 1964 te Londen voor; gemeten wordt er met de Campbell-Stokes, waarbij de diagrammen volgens de engelse instructie uitgetrokken worden. Gesteld $\overset{o}{\mu}_{II}$ is de ware, totale zonneschijnduur in januari 1964 te De Bilt; hier wordt gemeten bijv. met de Campbell-Stokes, waarbij de stroken worden uitgemeten naar nederlandse voorschrift of bijv. met de Jordan-heliograaf. Meestal $\overset{o}{\mu}_I \neq \overset{o}{\mu}_{II}$. De metingen zullen heten $\underline{x}_I$, zijnde een stochastische schatting (behept met een toevallige en een systematische fout) van $\overset{o}{\mu}_I$ en $\underline{x}_{II}$, zijnde een stochastische schatting (eveneens behept met een toevallige en een systematische fout) van $\overset{o}{\mu}_{II}$. Noem $\Delta \equiv \overset{o}{\mu}_I - \overset{o}{\mu}_{II}$ en zij $\Delta < 0$. Noem $\underline{d} = \underline{x}_I - \underline{x}_{II}$; deze $\underline{d}$ is een stochastische schatting van Δ. Ofschoon $\Delta < 0$, kan toch $d > 0$ zijn (zie fig. 2a), m.a.w. een $d > 0$ duidt niet per se op een $\Delta > 0$. Wij zouden willen, dat de kans $P[\underline{d} > 0 | \Delta < 0]$ op een $d > 0$ bij een gegeven $\Delta < 0$, nul zou zijn. Dán zou van "ideale vergelijkbaarheid" sprake zijn. Indien nu, door het tweede instrument of de tweede techniek "anders te kiezen" of door de meting met dit tweede instrument of volgens deze tweede techniek op de een of andere wijze te "herleiden" ("reduceren", "corrigeren"), kan worden bereikt, dat deze P (meestal wél ongelijk aan nul zijnde) bij dezelfde Δ kleiner wordt, dan zeggen wij met deze "manipulatie" de twee metingen "meer vergelijkbaar" gemaakt te hebben*).

---

*) Men kan vragen, waarom wordt gewerkt met herleidingsfactoren (dus multiplicatief) en niet met herleidingstermen (additief) ? Wat beter is, valt moeilijk vooruit te zeggen. Het is bekend, dat verschillende klimatologische elementen (neerslag, temperatuur, luchtdruk, enz.) zich in dit opzicht verschillend gedragen. C.I.M.O. verzuimde o.i. onmiddellijk de raad van statistici in te winnen, vóór het Project begon. Wij vermoeden, dat men gemakshalve (omdat in vele andere projecten, waarin het doel is op standaardinstrumenten te herleiden, op procentuele correcties aangestuurd wordt) voor herleidingsfactoren gekozen heeft. Wij hebben ons aan dit besluit moeten houden.

4. DE PROCEDURE VAN DE INTERNATIONALE VERGELIJKING VAN ZONNESCHIJNMETERS

Wij gaan nu dieper op het begrip "vergelijkbaar maken" in. Zij weer $x_I > x_{II}$ (zie in vorig hoofdstuk). Toch mag uit dit enkele meetresultaat niet worden geconcludeerd dat ook $\mathring{\mu}_I > \mathring{\mu}_{II}$ is en wel omdat er toevallige én systematische fouten zijn ($\underline{\delta}$ en ε), die misschien te Londen sterk verschillen van die te De Bilt (zie fig. 2a). Zo rijst de vraag: welke totale zonneschijnduur zou er met de engelse N.S. in januari 1964 gemeten zijn, indien hij niet te Londen, maar te De Bilt, naast de nederlandse N.S., gestaan zou hebben (fig. 2b) ? Dan was zeker $\mathring{\mu}_I = \mathring{\mu}_{II}$ geweest. Aldus ligt de volgende procedure voor de hand: breng de engelse N.S. naar De Bilt (of de nederlandse N.S. naar Londen, of beide N.S.'s naar eenzelfde station*), elders en verricht hier maanden lang vergelijkende metingen, waarbij de stroken van de nederlandse N.S. worden uitgetrokken door de nederlandse meteorologische dienst volgens de nederlandse instructie en die van de engelse N.S. door de engelse meteorologische dienst in overeenstemming met het engelse voorschrift. Zie fig. 2b. Weer is het onwaarschijnlijk, dat x_I gelijk aan x_{II} zou zijn, ofschoon nù wel $\mathring{\mu}_I = \mathring{\mu}_{II}$, zegge $\mathring{\mu}$. De ongelijkheid van x_I en x_{II} kan thans met een eventuele ongelijkheid van de systematische fouten ε_I en ε_{II} en die van de toevallige fouten $\underline{\delta}_I$ en $\underline{\delta}_{II}$ begrepen worden; in fig. 2a daarentegen moest ook de mogelijkheid $\mathring{\mu}_I \neq \mathring{\mu}_{II}$ overwogen worden. Hierbij werd de ε_I uit fig. 2b gelijk aan die uit fig. 2a genomen, evenzo bij ε_{II}, hoewel dit geen noodzakelijkheid is, daar de ε nog van de plaats afhankelijk zou kunnen zijn (hetgeen ook van de $\sigma(\underline{x})$ gedacht zou kunnen worden).

De eenmalige metingen x_I en x_{II} van fig. 2b leveren één quotient $c = x_{II}/x_I$, dat een schatting is van het onbekende, theoretische, ware quotiënt $\gamma = \mu_{II}/\mu_I$, waarom het ons te doen is. Een n-voudige heruittrekking der stroken zou een gemiddeld quotiënt $\bar{c} = \sum_1^n c_i/n$ leveren, eveneens een schatting, en wel een meer nauwkeurige van γ . Wij wensen door zulk een herleiding te bereiken, dat de kans, dat de meting x_I (te Londen, zekere maand) een gegeven bedrag (zegge a) groter is dan de meting x_{II} (te De Bilt, zelfde maand), terwijl toch de ware waarde $\mathring{\mu}_I$ een gegeven bedrag (zegge b) kleiner is dan de ware waarde $\mathring{\mu}_{II}$, kleiner wordt, overgaande op de kans dat de

---

*) Misschien is de keuze niet helemaal onverschillig, zie in de discussies rond de procedures A en B.

meting x_I hetzelfde bedrag a groter is dan de herleide meting $x^*_{II} = \bar{c}.x_{II}$, terwijl weer de ware waarde $\overset{o}{\mu}_I$ hetzelfde bedrag b beneden $\overset{o}{\mu}_{II}$ gelegen is. In overeenstemming met het vorige hoofdstuk kunnen wij dit schrijven als:

$$\overbrace{P\left[\underline{d}^* = \underline{x}_I - \underline{x}^*_{II} = a > 0 \,\middle|\, \Delta = \overset{o}{\mu}_I - \overset{o}{\mu}_{II} = b < 0\right]}^{\text{ná herleiding}} < \overbrace{P\left[\underline{d} = \underline{x}_I - \underline{x}_{II} = a \,\middle|\, \Delta = b\right]}^{\text{vóór herleiding}}$$

Generaliserende: deze ongelijkheid moet gelden, voor alle waarden van a en b, ongeacht hun tekens.

.-.-.-.-.

Wij zien twee procedures A en B, volgens welke men de internationale vergelijking zou kunnen doen plaatsvinden.

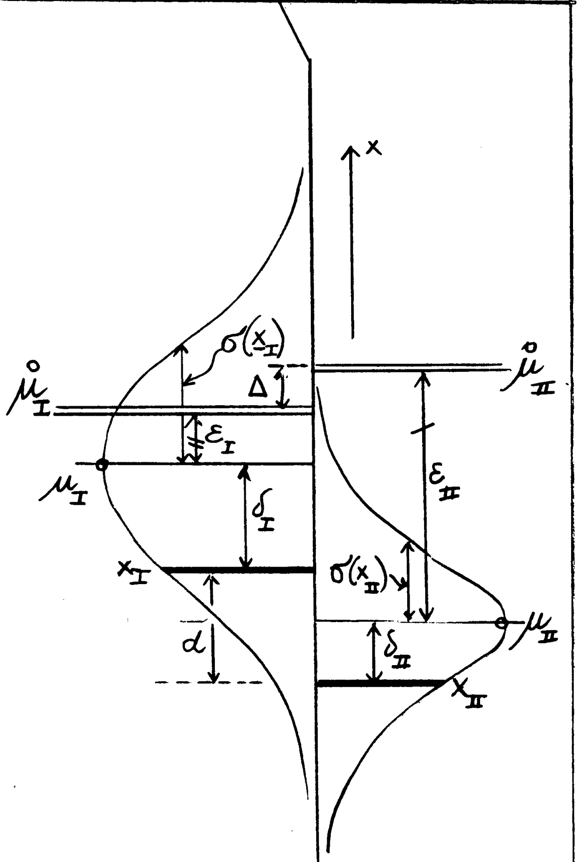
A

Elk der m deelnemende landen L_1, L_2,... L_m schaft zich een I.S. aan; noem deze I.S.$_1$, I.S.$_2$ ... I.S.$_m$. Het land L_i stelt I.S.$_i$ (i=1, ...m) naast zijn N.S. op. Vervolgens worden vele maanden achtereen vergelijkende metingen verricht, waaruit herleidingscoëfficiënten zullen resulteren (we gaan hier niet in op vragen als: "bezit de herleidingscoëfficiënt een jaarlijkse gang ?" enz.) De N.S.-stroken worden uitgemeten door de meteorologische dienst van het land L_i, volgens de daar geldende regels. De I.S.$_i$-stroken daarentegen worden uitgetrokken door één vaste persoon, die dit doet voor alle m landen, waarbij hij één enkele, internationale instructie volgt.

B

Elk der m deelnemende landen L_1 ...L_m zendt zijn N.S. naar één station, waar de m N.S.'s naast elkaar opgesteld worden. Men verricht er vele maanden lang vergelijkende metingen, en kiest òf een der m N.S.$_i$'s (noem deze N.S.*) òf het groepsgemiddelde (noem dit N.S.**) als "referentie", teneinde daarmede elk der N.S.$_j$'s te vergelijken. Aldus resulteren er òf m-1 verhoudingscoëfficiënten q_i òf m stuks. De stroken van de N.S.$_i$ worden uitgemeten door de dienst van land L_i volgens de voor dat land geldende voorschriften. Het is dan deze q_i, waarmede de maandtotalen, gemeten met de N.S.$_i$ in het land L_i, voortaan moeten worden vermenigvuldigd om ze te "herleiden tot het N.S.*-, dan wel het N.S.**-niveau" ("common basis").

jan. 1964	
I	II
engelse N.S. te Londen, engelse instructie	ned. N.S. te De Bilt, ned. instructie



$\underline{x}_I = \overset{\circ}{\mu}_I + \varepsilon_I + \underline{\delta}_I = \mu_I + \underline{\delta}_I$

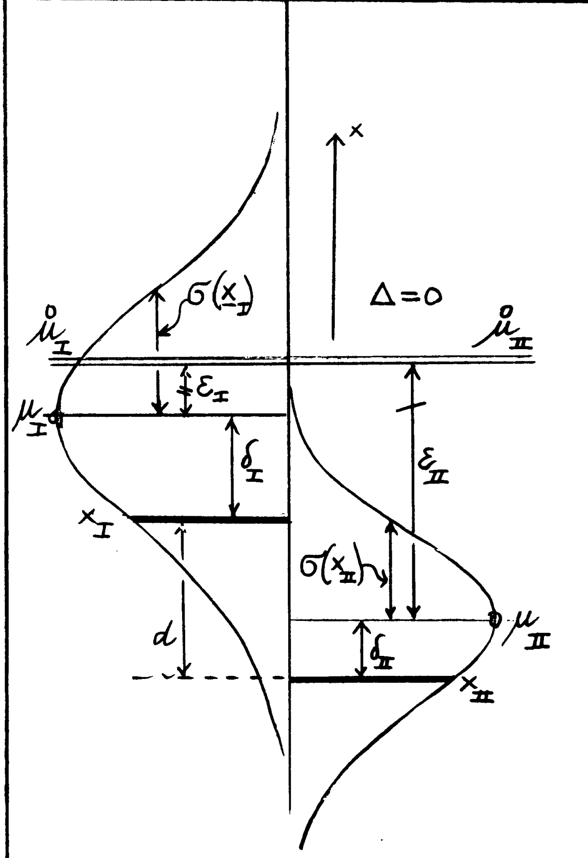
$\underline{x}_{II} = \overset{\circ}{\mu}_{II} + \varepsilon_{II} + \underline{\delta}_{II} = \mu_{II} + \underline{\delta}_{II}$

b.v. $x_I > x_{II}$ en

$\overset{\circ}{\mu}_I < \overset{\circ}{\mu}_{II}$ met $\mu_I > \mu_{II}$

Als $x^{*}_{II} = \bar{c}\, x_{II}$ dan is

jan. 1964	
I	II
engelse N.S. te De Bilt, engelse instructie	ned. N.S. te De Bilt, ned. instructie



$\underline{x}_I = \overset{\circ}{\mu} + \varepsilon_I + \underline{\delta}_I = \mu_I + \underline{\delta}_I$

$\underline{x}_{II} = \overset{\circ}{\mu} + \varepsilon_{II} + \underline{\delta}_{II} = \mu_{II} + \underline{\delta}_{II}$

b.v. $x_I > x_{II}$ en

$\overset{\circ}{\mu}_I \equiv \overset{\circ}{\mu}_{II} \equiv \overset{\circ}{\mu}$ met $\mu_I > \mu_{II}$

Bereken $c = x_I / x_{II}$ en $\bar{c} = \frac{1}{n}\sum c$

$$\boxed{P\left[x_I - x^{*}_{II} > 0 \;\middle|\; \overset{\circ}{\mu}_I - \overset{\circ}{\mu}_{II} < 0\right] < P\left[x_I - x_{II} > 0 \;\middle|\; \overset{\circ}{\mu}_I - \overset{\circ}{\mu}_{II} < 0\right]}$$

fig. 2a

fig. 2b

Aldus wordt voor het land L_i een verhoudingsfactor (factoren eventueel) c_i gevonden, waarmede de maandtotalen, hier gemeten, moeten worden vermenigvuldigd om ze te "herleiden tot het I.S.-niveau". ("common basis").

De procedures A en B zijn zeker niet gelijk. Hier volgt enig commentaar.

1) In procedure A moeten de $I.S._1$, $I.S._2$ . . . $I.S._m$ (let wel: m=86= aantal deelnemende stations (meer dan één station per land) "identiek" zijn; zijn zij dat ? Wat is "identiek" ? En als m apparaten op het station S_1 identiek zouden zijn, zijn zij het dan ook op een ander station S_2 ?
2) Bij procedure B is het zeker, dat met alle instrumenten gelijktijdig dezelfde ware, onbekende waarde gemeten wordt; in procedure A zijn de ware waarden in groepjes van 2 gelijk, d.i. op station S_1, meten $N.S._1$ en $I.S._1$ dezelfde ware waarde, zegge $\overset{o}{\mu}_1$, terwijl op station S_2 de $N.S._2$ en $I.S._2$ dezelfde, ware waarde, zegge $\overset{o}{\mu}_2$, meten, waarbij meestal $\overset{o}{\mu}_1 \neq \overset{o}{\mu}_2$ zal zijn, etc.
3) Dat de volgens procedure B te vinden herleidingscoëfficiënten q_1, q_2 . . . q_m meestal niet statistisch gelijk zullen zijn aan de volgens procedure A te vinden c's, is niet belangrijk, omdat het eerste doel is door toepassing der herleidingsfactoren de vergelijkbaarheid te verbeteren. Wel zou het kunnen zijn, dat de q's samenhangen met de ligging van het voor de uitvoering der vergelijking gekozen station.
4) Procedure B lijkt ons goedkoper uit te voeren, doch moeilijker te organiseren, dan procedure A.

Het is ons niet bekend of door C.I.M.O. aan het bestaan van deze procedure gedacht is en op welke gronden procedure A gekozen werd, zij het in een wat afwijkende vorm. Als interim-referentie-instrument werd gekozen de Campbell-Stokes zonneschijnmeter, geleverd door de firma Casella (te Londen) en gebouwd in overeenstemming met de door het Britse Meteorological Office gestelde eisen, terwijl diagrammen, in Frankrijk vervaardigd volgens het door de Franse Météorologie Nationale gegeven voorschrift, gebruikt werden. Over de afwijkende vorm het volgende. Door het Uitvoerend Comité werd besloten, dat op <u>ieder</u> station <u>zowel</u> de stroken van de N.S. als die van de ernaast geplaatste I.S. door de nationale routinedienst zouden worden uitgetrokken en wel die van de N.S. volgens de geldende nationale instructie en die van de I.S. volgens het gegeven internationale voorschrift. Naar onze mening wordt hiermede een essentiëel element over het hoofd gezien, te weten het (in Nederland kwantitatief vastgestelde) feit, dat verschillende personen bij het uitmeten van eenzelfde stel stroken ongelijke uitkomsten zullen krijgen, ook al doen ze dit zeer vele keren en ook al volgen zij daarbij eenzelfde voorschrift. Daar de internationale instructie bovendien niet bijzonder scherp geformuleerd*) werd (kortheid en eenvoud waren de eerste eisen), zodat er ruimte is voor persoonlijke interpretatie, vrezen wij, dat de door de W.M.O. vermoedelijk eenvoudigsheidshalve aanvaarde procedure, afwijkend van de op pagina 8 beschreven procedure A, statistische gevolgen zal hebben. (Wij komen erop terug in de laatste alinea van 8.2). Misschien zijn deze voor de praktijk van geen belang, dit valt moeilijk vooruit te zeggen. Een afzonderlijk na-onderzoek zou wellicht gewenst zijn, om te voorkomen, dat men zou besluiten herleidingscoëfficiënten toe te passen, die misschien de toets der statistische kritiek niet zullen kunnen doorstaan.

5. DE HERLEIDINGSCOËFFICIËNT ALS STATISTISCHE VARIABELE

Persoon A vindt voor januari 1964 uit de N.S.-stroken, uitgetrokken volgens de nationale instructie, een totale zonneschijnduur x uren en uit de I.S.-stroken (I.S. opgesteld naat de N.S.), uitgetrokkenvolgens de internationale instructie, y uren. Zowel $\underline{x}$ als $\underline{y}$ is een statistische

*) Nog een bijzonderheid:
De Internationale Instructie (I.I.) verschilt niet slechts van de meeste Nationale Instructies (N.I.'s) wat details inzake het uitmeten van de brandsporen aangaat (hoe te handelen met "brandpunten", "insnoeringen", "flauwe inzetten en uitlopen", etc.), maar ook in de wijze, waarop de dagelijkse zonneschijnduur bepaald wordt. Vrijwel iedere N.I. eist, dat iedere strook voor elk der afzonderlijke uren wordt uitgemeten; de som dezer uursommen is dan de dagsom (de maandsom is de som dezer dagsommen). De I.I. ordonneert nadrukkelijk, dat <u>niet</u> per uur wordt uitgetrokken (wat dat betreft hadden de uurstrepen dus weggelaten kunnen worden), maar over de gehele dag, direct(de maandsom is weer de som der dagsommen). De "regelrechte" dagsom is meestal groter dan de som der uursommen als gevolg van afrondingsfouten, die weer samenhangen met de manipulaties met de duren der "insnoeringen" en "brandpunten" (wij gaan er hier niet verder op in).

variabele, d.w.z. als A dezelfde stroken talloos vele keren zou uitmeten, zou er een, zo goed als normaal verdeelde, x-populatie ontstaan; evenzo een y-populatie en evenzo een c-populatie, waarbij $\underline{c} = \underline{y}/\underline{x}$. De vraag rijst: hoe zullen het gemiddelde μ_c en de standaarddeviatie σ_c van de $\underline{c}$-populatie samenhangen met het gemiddelde μ_x en de standaarddeviatie σ_x van de $\underline{x}$-populatie en het gemiddelde μ_y en de standaarddeviatie van de y-populatie ?

Het antwoord luidt (hier niet bewezen):

(1a) $\mu_c \approx \gamma(1 + \kappa_x^2)$ en $\sigma_c^2 \approx \gamma^2(\kappa_x^2 + \kappa_y^2)$

Hierin zijn per definitie $\gamma \equiv \mu_y/\mu_x$; $\kappa_x \equiv \sigma_x/\mu_x$; $\kappa_y \equiv \sigma_y/\mu_y$.

Als ook nog $\kappa_x \ll 1$ is en per definitie $\kappa_c \equiv \sigma_c/\mu_c$ gesteld

(1b) wordt, dan gelden $\kappa_c^2 \approx \kappa_x^2 + \kappa_y^2$ ("Stelling van Pythagoras" voor relatieve varianties) en

(1c) $\mu_c \approx \gamma$

De formules 1a gelden slechts als $\underline{x}$ en $\underline{y}$ ongecorreleerd zijn. Deze onderstelling is zeker vervuld, daar de bij de meting x gemaakte meetfout niets te doen heeft met de bij de meting y gemaakte meetfout.

De $\underline{c}$ volgt een zo goed als normale verdeling (hier niet bewezen); de benadering is beter, hoe meer aan $\kappa_x \ll 1$ en $\kappa_y \ll 1$ voldaan is (en dus wel bij maandtotalen, doch niet bij dagtotalen).

Men mag niet a priori $\kappa_x = \kappa_y$ stellen. De instructie zelf heeft ook invloed op de κ , d.i. een minder scherpe instructie leidt tot een grotere κ (hetgeen in Nederland experimenteel bleek), en er is hier sprake van twee instructies. Het bovenstaande gold voor b.v. persoon A. Voor persoon B gelden dezelfde beschouwingen. Iedere persoon heeft, in het algemeen gesproken, voor iedere individuele maand, zijn eigen μ_x, σ_x, μ_y, σ_y en μ_c, σ_c.

6. MEERVOUDIGE ANALYSEN IN HET LABORATORIUM

Bij laboratoriumwerk (b.v. chemische analysen) is het een bekend statistisch probleem: "Waarvan hangt de nauwkeurigheid van het gemiddelde van meervoudige metingen, verricht door een aantal analisten, af?" Wij besteden hier enige aandacht aan omdat het mogelijk is de statistische beschouwing te analogiseren (in hoofdstuk 7) ten behoeve van het herleidingsvraagstuk, dat onderwerp van het onderhavige rapport is. In hoofdstuk 8 volgt een toepassing teneinde een antwoord te geven op de vraag hoe lang het Vergelijkingsproject minimaal duren moet.

Denk aan de potentiële aanwezigheid van een oneindig aantal analisten (de "populatie"). Zij $\mathring{\gamma}$ de onbekende, ware waarde (b.v. van de p_H van een vloeistof). De door analist A (behorende tot deze populatie) verrichte meting $\underline{c}_A$ volgt een kansverdeling met een gemiddelde $\mu_A(\underline{c})$, zegge γ_A en een standaarddeviatie $\sigma_A(\underline{c})$, zegge σ_A. Voor een andere analist B: γ_B en σ_B, enz. Bij een populatie van analisten behoort aldus een populatie van γ-waarden (gemiddelde μ_γ ; standaarddeviatie σ_γ), eveneens een populatie van σ-waarden (gemiddelde μ_σ ; standaarddeviatie σ_σ). De σ_γ is een maat voor de "spreiding tussen de analisten"; de σ_A is een maat voor de spreiding "binnen de analist A". Laat ieder van k <u>aselect gekozen</u>*) analisten m bepalingen van $\mathring{\gamma}$ verrichten. Het rekenkundige gemiddelde $\underline{\bar{c}}$ van deze km bepalingen is een stochastische variabele. Deze $\underline{\bar{c}}$ volgt (hier niet bewezen), een verdeling, met een gemiddelde $\mu(\underline{\bar{c}})$ en een standaarddeviatie $\sigma(\underline{\bar{c}})$. De verdeling is, bij voldoend grote km, zo goed als normaal. Er geldt:

(2) $$\mu(\underline{\bar{c}}) = \mu_\gamma$$

(3) $$\sigma(\underline{\bar{c}}) = [\sigma_\gamma^2/k + \mu_\sigma^2/km]^{\frac{1}{2}}$$

In het algemeen verschilt μ_γ van $\mathring{\gamma}$. Persoon A maakt de systematische fout $\varepsilon_A = \mathring{\gamma} - \mu_A$; B: $\varepsilon_B = \mathring{\gamma} - \mu_B$, etc. Er is een populatie van ε-waarden, met een gemiddelde $\mu_\varepsilon = \mathring{\gamma} - \mu_\mu$ (dat men de systematische fout van de personen-populatie zou kunnen noemen), en een standaarddeviatie σ_ε.

Ook als de k personen <u>steeds dezelfde</u>*) zouden zijn, zou $\underline{\bar{c}}$ een statistische variabele zijn. Dan is

---

*) Toelichting:
<u>Geval I</u>: kies k = 3, m = 2. De drie personen A, B, C doen elk twee metingen; het gemiddelde der 6 metingen zij $\bar{c}_1$. Daarna doen weer drie personen, nu b.v. A, C, E elk twee metingen; $\bar{c}_2$. Daarna b.v. C, F, H elk 2 metingen; $\bar{c}_3$. Daarna b.v. P, Q, R elk 2 metingen; $\bar{c}_4$, enz. Deze $\bar{c}$'s volgen een kansverdeling met een door (2) en (3) gegeven gemiddelde en standaarddeviatie.
<u>Geval II</u>: kies weer k = 3, m = 2. Van drie personen A, B, C doet nu elk twee metingen; het gemiddelde zij $\bar{c}_1$. Vervolgens doen dezelfde personen A, B, C weer allen twee metingen; $\bar{c}_2$. Daarna doet weer ieder twee metingen; $\bar{c}_3$, etc. Ook deze $\bar{c}$'s volgen een kansverdeling. Gemiddelde en standaarddeviatie daarvan zijn nu door (4) en (5) gegeven.

(4) $\mu(\underline{\bar{c}}) = \bar{\gamma}$

(5) $\sigma(\underline{\bar{c}}) = \sqrt{(\bar{\sigma})^2/km}$

Daarin staat $\bar{\gamma}$ voor het gemiddelde van de bij deze k personen behorende $\gamma_{\bullet}$-waarden en $\bar{\sigma}$ voor het gemiddelde bij dezelfde personen behorende $\sigma_{\bullet}$-waarden. Meestal: $\bar{\gamma} \neq \mu_\gamma$ en $\bar{\sigma} \neq \mu_\sigma$.

Hier volgt enig commentaar.

A. bij (2) en (3):

a) altijd $\sigma(\underline{\bar{c}}) \geqslant \sigma_\gamma / \sqrt{k}$ ($\neq 0$)

b) $\sigma(\underline{\bar{c}})$ is kleiner, d.i. $\bar{c}$ is nauwkeuriger,
als σ_γ kleiner is, d.i. als de personen onderling minder verschillen;
als μ_σ kleiner is, d.i. als iedere analist nauwkeuriger meet;
als k groter is (meer personen);
als m groter is, d.i. meer metingen per persoon.

c) $\sigma(\underline{\bar{c}})$ bij k=3 en m=2 is kleiner dan $\sigma(\underline{\bar{c}})$ bij k=2 en m=3, hoewel in beide gevallen km = 6.

d) als $\mu_\sigma \gg \sigma_\gamma$, d.i. als de personen onderling weinig verschillen, doch elk met een grote onnauwkeurigheid meet, dan is $\sigma(\underline{\bar{c}}) \approx \mu_\sigma / \sqrt{km}$

e) als $\mu_\sigma \ll \sigma_\gamma$ (tegengestelde situatie), dan $\sigma(\underline{\bar{c}}) \approx \sigma_\gamma / \sqrt{k}$

f) $\mu(\underline{\bar{c}})$ hangt niet van m en k af

g) als k=1; m=1, dan $\sigma(\underline{c}) = \sqrt{\sigma_\gamma^2 + \mu_\sigma^2}$

B. bij (4) en (5)

a) zowel voor grotere k als voor grotere m neemt $\sigma(\underline{\bar{c}})$ af naar nul.

b) $\sigma(\underline{\bar{c}})$ heeft voor een vaste waarde van het product km niet per se dezelfde waarde, daar $\bar{\sigma}$ een andere waarde kan aannemen én bij dezelfde k (weer een k-tal personen, maar andere) én bij andere k.

c) $\mu(\underline{\bar{c}})$ hangt nu wel met k samen, zij het onoverzichtelijk.
Of $\bar{\gamma}$ bij k_1 groter dan, gelijk aan, kleiner dan $\bar{\gamma}$ bij k_2 zal zijn, indien $k_2 \neq k_1$, danwel $k_2 = k_1$, is niet vooruit te zeggen.

d) als k=m=1, dan $\sigma(\underline{c}) = \sigma_A$ voor analist A.

Wij gaan deze formules gebruiken in de hoofdstukken 7 en 8.

7. MEERVOUDIGE UITTREKKINGEN VAN DEZELFDE ZONNESCHIJNSTROKEN

In hoofdstuk 6 spraken wij van een analist A, die m keren een meting van de onbekende, ware waarde $\overset{\circ}{\gamma}$ deed. Nu denken wij aan de assistent A, die m keren de onbekende, ware waarde $\overset{\circ}{\gamma}$ van de herleidingscoëfficiënt "meet" (berekent). Gesteld **hij trekt de** stroken én van N.S. én van I.S. bij een gegeven maand m keren uit, zodat er m onafhankelijke quotiënten*) $c = y/x$ resulteren; x = maandsom volgens N.S.; y volgens I.S. Echter ook als hij m maanden elk één keer uittrekt, resulteren er m quotiënten c. Daarbij wordt ondersteld, dat de ware waarde van de herleidingscoëfficiënt $\overset{o}{\gamma}$ voor iedere maand dezelfde is, welke onderstelling getoetst moet worden. Een tweede assistent B doet hetzelfde, een derde C dito, etc. Zo zou januari 1964 kunnen worden uitgetrokken én 2 keren door A, én 2 keren door B én twee keren door C. Het gemiddelde der 6 c's hete $\bar{c}_I$. Vervolgens wordt februari 1964 behandeld: 2 keren door B, 2 keren door D en 2 keren door F. Aldus ontstaat een $\bar{c}_{II}$. Daarna maart dito door A, F, H ($\bar{c}_{III}$). Daarna april door B, F, G ($\bar{c}_{IV}$), enz. Kortom: steeds doet een ieder van drie aselect uit de populatie gekozen personen een duplometing; d.i. k=3, m=2. Aldus leveren a=5 maanden n=kma=30 c-waarden en a=5 $\bar{c}_.$-waarden. Het gemiddelde $\bar{\bar{c}}$ dezer n c-waarden, dat is tevens het gemiddelde der a $\bar{c}_.$-waarden, is een schatting van de $\overset{o}{\gamma}$. De verwachtingswaarde van deze $\bar{\bar{c}}$ is niet $\overset{o}{\gamma}$, doch μ_γ , zie formule (2), terwijl de nauwkeurigheid van de $\bar{\bar{c}}$ gegeven wordt door (3). Wij hebben dit in fig.3 geïllustreerd. Persoon A heeft zijn eigen c-populatie (gekarakteriseerd door gemiddelde γ_A en standaarddeviatie σ_A); ondersteld wordt, dat deze c-populatie voor iedere maand dezelfde is. A heeft dus ook zijn eigen systematische fout $\varepsilon_A = \overset{o}{\gamma} - \gamma_A$. Een populatie van personen bepaalt een populatie van $\gamma_.$-waarden, een populatie van $\sigma_.$-waarden en een populatie van $\varepsilon_.$-waarden, met gemiddelden resp. μ_γ , μ_σ en μ_ε . In het algemeen is $\mu_\gamma \neq \overset{o}{\gamma}$ en dus $\mu_\varepsilon \neq 0$. Daar iedere afzonderlijke $\sigma_.$ (b.v. σ_A) én van σ_x én van σ_y afhangt (zie formule 1a, 1b en 1c), hangt ook μ_σ samen met de σ_x- en σ_y-waarden (of met de κ_x- en κ_y-waarden) aller personen.

---

*) Er resulteren eigenlijk m^2 quotiënten $c_{ij} = y_i/x_j$, met i, j = 1, 2 ... m, doch deze zijn niet afhankelijk. Zo zijn $c_{23} = y_2/x_3$ en $c_{33} = y_3/x_3$ afhankelijk, doordat de breuken de noemer gemeenschappelijk hebben. Deze afhankelijkheid levert statistische complicaties, die wij wilden vermijden. Wij beschouwen dus slechts $c_i = y_i/x_i$, i = 1, 2 .... m.

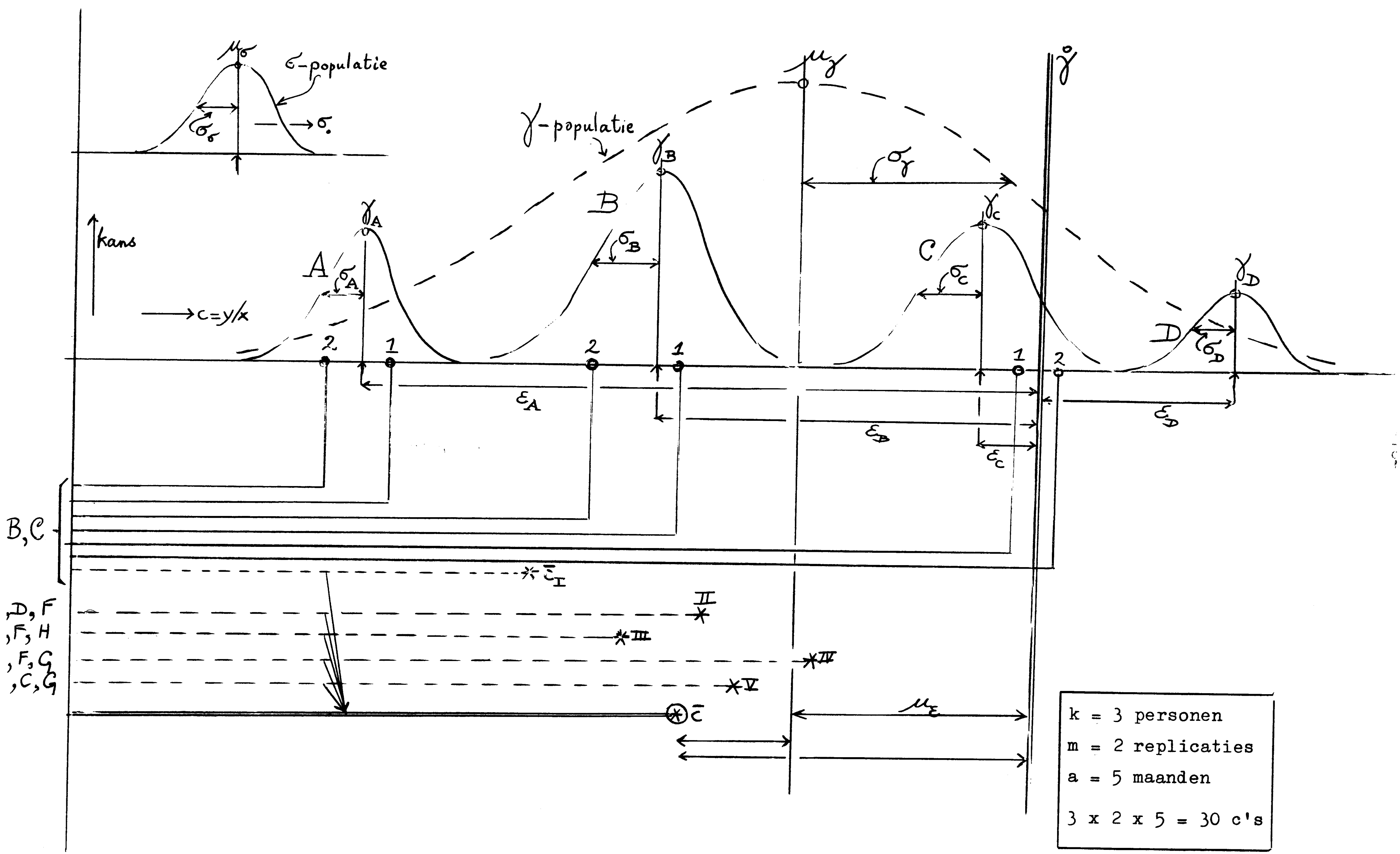

σ-populatie
γ-populatie
kans
c=y/x
A
B
C
D
B, C
,D, F
,F, H
,F, G
,C, G
k = 3 personen
m = 2 replicaties
a = 5 maanden
3 x 2 x 5 = 30 c's

fig. 3

8. DE MINIMUM DUUR VAN HET VERGELIJKINGSPROJECT

8.1 Het toetsen van hypothesen

Gesteld men wil weten of de onbekende, ware waarde γ van een "object" al of niet een voorgeschreven bedrag Δ van een gegeven waarde γ_0 verschilt. Het is duidelijk, dat men zich hierover nooit volledige zekerheid verschaffen kan, hoevele metingen men ook zal willen verrichten. Volledige zekerheid is er namelijk slechts als γ bekend zal zijn. Zo rijst de vraag: Hoevele metingen (met bekende meetnauwkeurigheid σ) moeten tenminste verricht worden om dan toch een "redelijke zekerheid" te verkrijgen ? De redenering is als volgt.

De nulhypothese H_0 luidt, dat de ware waarde γ_0 is. De alternative hypothese H_a luidt, dat de ware waarde $\gamma_a (\neq \gamma_0)$ is. Men doet n metingen $\underline{c}_1, \underline{c}_2, \ldots \underline{c}_n$ en berekent aldus een gemiddelde $\bar{\underline{c}} = \Sigma\, \underline{c}_i/n$. Dit stochastische gemiddelde volgt een kansverdeling, welke, onder H_0, een gemiddelde γ_0 en een standaarddeviatie $\sigma(\bar{\underline{c}}) = \sigma/\sqrt{n}$ bezit. Hoe groter n, hoe meer nadert deze verdeling tot een normale (volgt $\underline{c}$ zelf een normale verdeling, dan is $\bar{\underline{c}}$ normaal verdeeld voor elke n). Zou H_a en niet H_0 waar zijn, dan zou $\bar{\underline{c}}$ weer aan een, eveneens vrijwel normale, kansverdeling gehoorzamen, nu met gemiddelde γ_a en dezelfde standaarddeviatie $\sigma/\sqrt{n}$. In het algemeen zal $\bar{c} \neq \gamma_0$ zijn; toch houdt dit niet in, dat dus H_0 verworpen moet worden. Er zijn twee soorten fouten:

a) de fout van de <u>eerste soort</u>, d.i. de fout, gemaakt bij het verwerpen van H_0 ofschoon H_0 geldt. De kans op deze fout duiden wij aan met α (zodat $1 - \alpha$ de kans is op het <u>terecht aanvaarden van H_0</u> is);

b) de fout van de <u>tweede soort</u>: d.i. de fout, gemaakt bij het aanvaarden van H_0, hoewel H_0 niet geldt. De kans op deze fout duidt men aan met β (zodat $1 - \beta$ de kans is op het <u>terecht verwerpen van H_0</u>). Zie de figuren 4 en 5.

Nadat α gekozen is (b.v. 0,05) bepaalt men rondom γ_0 een "gebied van aanvaarding van H_0", zodanig breed, dat $\bar{\underline{c}}$, als H_0 geldt, met een kans $1 - \alpha$ daarin liggen zal. Het ligt, als $\bar{\underline{c}}$ symmetrisch verdeeld is, symmetrisch rondom γ_0 en heeft, als $\bar{\underline{c}}$ bovendien normaal verdeeld is, een halve breedte $\tau\,.\,\sigma(\bar{\underline{c}}) = \tau\,.\,\sigma/\sqrt{n}$. Daarbij staat τ voor die t-waarde, die in de t- of Student-verdeling op n-1 graden van vrijheid met kans $\frac{1}{2}\alpha$ overschreden wordt. Voorbeelden: als $\alpha = 0{,}05$ is, dan bij $n = 5$ een $\tau = 2{,}45$; 10 met 2,23; 25 met 2,06; ∞ met $1{,}96 \approx 2$. Het gebied buiten dit aanvaardingsinterval

	H_o geldt H_a geldt niet	H_o geldt niet H_a geldt
H_o verwerpen	α	$1-\beta$
H_o aanvaarden	$1-\alpha$	β

fig. 4

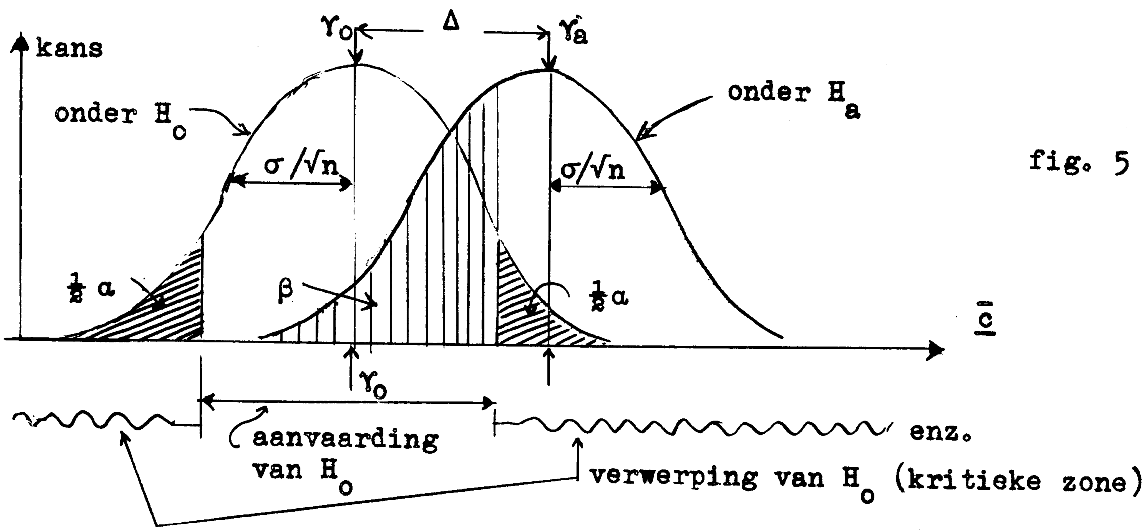


fig. 5

heet het "gebied van verwerping van H_o" of de "kritieke zone". Indien de gemeten $\bar{c}$ daarin valt (hierop is, onder H_o, een kans α), dan wordt H_o verworpen (elk der twee gearceerde "staarten" van de linker klokkromme in fig.5 representeert $\frac{1}{2}\alpha$). Nu kan, ook als niet H_o, doch H_a geldt, toch $\bar{c}$ in het aanvaardingsgebied vallen. De kans β , dat dit geschiedt, is een functie van α , n, γ_o en γ_a. (Het verticaal gearceerde deel in de linkerhelft van de rechter klokkromme representeert β). Omgekeerd: geeft men β , dan is $\Delta \equiv |\gamma_a - \gamma_o|$ een functie van α , n, β . Hier volgt (zonder bewijs)

(6) $\quad \Delta = (\tau+\nu)\sigma(\bar{\bar{c}}) = R\sigma(\bar{c}) = R\sigma/\sqrt{n}.$

Daarin is $R = \tau + \nu$; τ werd reeds gedefiniëerd; ν stelt die t-waarde voor, die in de t-verdeling met n-1 graden van vrijheid met een kans β overschreden wordt.

De volgende drie 2x2 tabellen (n=5; 25; ∞) geven de R-waarden bij de vier combinaties α , β = 0, 05 0,05; 0,05 0,01; 0,01 0,05 en 0,01 0,01 te zien.

n = 5

β \ α	0,05	0,01
0,05	4,9	6,7
0,01	6,5	8,4

n = 25

4,1	4,8
5,2	5,9

n = ∞

3,6	4,2
4,3	4,9

Uit (6) volgt

(7) $$n = R^2 \; (\sigma / \Delta)^2$$

Interpretatie: indien men met "redelijke zekerheid" wil te weten komen, of de onbekende, ware waarde al of niet een voorgeschreven bedrag Δ van een hypothetisch gestelde waarde verschilt en indien men daarbij een kans α op een fout van de eerste soort en een kans β op een fout van de tweede soort wil toelaten, dan moeten tenminste zovele metingen verricht worden als (7) aangeeft. Kiest men b.v. β = 0,01, dan is die "redelijke zekerheid" $1 - \beta = 0{,}99$. Men ziet n groeien met grotere R (d.i. kleinere α en kleinere β), grotere σ en kleinere Δ .

8.2 Toepassing van de "theorie van het toetsen van hypothesen" op het herleidingsprobleem

Daar bij identieke meetinstrumenten of -technieken geen herleiding op elkaar nodig is, ligt het voor de hand als nulhypothese H_0 te formuleren, dat $\gamma_0 = 1$ is. Wij moeten ons goed bewust zijn wat deze γ_0 betekent. Gesteld, in de meteorologische dienst zou het de gewoonte zijn, dat alle stroken van iedere maand altijd door drie personen (maar niet steeds dezelfde drie) worden uitgetrokken, en door ieder van hen in enkelvoud. Worden de N.S.-stroken van b.v. jan. 1964 uitgemeten door A, K, Z, dan zouden de I.S.-stroken van dezelfde maand door dezelfde personen moeten worden uitgetrokken. Aldus zou de maand januari 1964 bij een éénmalige uitmeting door A, K, Z één $\bar{c}_{AKZ}$ (uit 3 c's) geven en bij een oneindigmalige uitmeting een populatie van $\bar{c}_{AKZ}$-waarden, met een gemiddelde $\gamma_{AKZ} = \frac{1}{3}\,(\gamma_A + \gamma_K + \gamma_Z)$ en een variantie $\sigma^2_{AKZ} = \frac{1}{9}\,(\sigma^2_A + \sigma^2_K + \sigma^2_Z)$; γ_A, σ_A stellen gemiddelde, resp. standaarddeviatie der bij A behorende $\underline{c}$-populatie voor. Indien B, D en V de maand februari 1964 zouden behandelen, dan één $\bar{c}_{BDV}$ en een γ_{BDV} en een σ^2_{BDV}.

Indien A, B, C de maand maart 1965 voor hun rekening zouden nemen, dan één $\bar{c}_{ABC}$ en een γ_{ABC} en een σ^2_{ABC}, etc. Daar het (alles in theorie !) op oneindig vele manieren mogelijk is uit een populatie van personen een drietal te kiezen, is er ook een populatie van $\gamma_{,,,}$-waarden, rondom een gemiddelde μ_γ (dezelfde μ_γ als in fig. 3), en een populatie van $\sigma^2_{,,,}$-waarden, rondom een gemiddelde $\frac{1}{3}\mu^2_\sigma$ (dezelfde μ^2_σ als in fig. 3). De in de aanvang van 8.1 genoemde σ moet dan door deze $\frac{1}{3}\mu_\sigma\sqrt{3}$ vervangen worden en de in de aanvang van 8.2 in de nulhypothese genoemde γ_o zou dan μ_γ zijn.

Zouden de drie personen steeds dezelfde zijn (b.v. A, B, C), dan zou γ_o vervangen moeten worden door $\gamma_{\overline{ABC}} = \frac{1}{3}(\gamma_A + \gamma_B + \gamma_C)$ en σ door $\sigma_{\overline{ABC}} = \frac{1}{3}(\sigma_A + \sigma_B + \sigma_C)$ of, beter misschien, door $\sqrt{\frac{1}{3}(\sigma^2_A + \sigma^2_B + \sigma^2_C)}$.
De werkelijkheid is dikwijls nòg anders; dan worden misschien én de stroken van N.S. én die van I.S. van 15 dagen van januari 1964 door A, die van 10 andere dagen door K en die van de resterende 6 dagen door Z, alle één keer, uitgetrokken. Er resulteert één c. In februari trekt misschien B 20 dagen uit en D de resterende 9. In maart 1964 trekt A 4 dagen uit, B 3 andere, C 9 weer andere, D 2 weer andere en E 3 weer andere dagen, etc. Deze situatie is zo ingewikkeld, dat ze "onmogelijk" volkomen statistisch verantwoord aan te pakken is.

De alternative hypothese luidt $\gamma_a \neq 1$. Men zou het verschil $|\gamma_a - 1|$ pas "interessant" (voor de praktijk) kunnen noemen als het zeker, nader overeen te komen, bedrag Δ^* overtreft. De, nu zo algemeen mogelijk gestelde, vraag luidt: gedurende hoevele jaren (n jaren) moeten er vergelijkende metingen verricht worden, indien iedere maand voor ieder van k personen (vaste k, maar niet maand voor maand per se dezelfde personen) alle N.S.- en alle I.S.-stroken m keren uitgetrokken worden ? Er resulteren dan in totaal n = 12 Nmk c-waarden, met een gemiddelde $\bar{c}$. Al deze c-waarden zijn schattingen van eenzelfde onbekende $\overset{o}{\gamma}$, als althans iedere maand, hoe lang en op welke wijze ook de zon scheen, dezelfde γ bezit.

Substitueer de $\sigma(\bar{c})$ uit (3), na km door de genoemde n vervangen te hebben, in (6) en vervang Δ door Δ^*. Er komt, opgelost naar n,

(8) $$\boxed{n = 12\ Nmk = R^2 \frac{(\mu_\sigma / \Delta^*)^2}{1 - \frac{R^2}{k}\left(\frac{\sigma_\gamma}{\Delta^*}\right)^2}}$$ maanden, waarin, strikt beschouwd, de R een functie van n is.

Interpretatie: als de werkwijze van de routinedienst zodanig zou zijn, dat van iedere maand alle stroken door een vast aantal k, maar niet steeds dezelfde, personen uitgemeten worden, dan zouden in het Vergelijkingsproject, teneinde met een "zekerheid" 1 - β te weten te kunnen komen of de onbekende herleidingscoëfficiënt al of niet een gegeven bedrag Δ^* van 1 verschilt (daarbij een kans α op een fout van de eerste soort en een kans β op een fout van de tweede soort toelatend) tenminste zovele vergelijkende metingen moeten worden verricht als formule (8) aangeeft. De α en β bepalen tezamen de R (zie (6)). Bij deze n kunnen N, m en k in allerlei combinaties gekozen worden. Zo leveren de vier combinaties:

k, m, N: 5, 1, 4 5, 8, $\frac{1}{2}$ 1, 1, 20 en 1, 8, $2\frac{1}{2}$

alle eenzelfde product n = 12 kmN = 240, doch de duren van het Project zouden daarbij uiteenlopen tussen 6 maanden en 20 jaren, al naar gelang één of vijf personen de stroken uittrekken en al naar gelang een ieder het in enkelvoud of achtvoud doet. Bestaat de routinedienst evenwel uit k, steeds dezelfde personen, dan substituere men $\sigma(\underline{c})$ uit (5), na opnieuw km door n = 12 Nmk vervangen te hebben, in (6), zodat er komt

(9) $$n = 12\ Nmk = R^2 \left(\frac{\bar{\sigma}}{\Delta^*}\right)^2 \text{ maanden.}$$

Bij (8) behoort een $\mu(\underline{c}) = \mu_\gamma$, bij (9) een $\mu(\underline{c}) = \bar{\gamma}$

N (<u>minimum duur van het Project</u>) is dus een vrij ingewikkelde functie van α, β, μ_σ, Δ^*. De N is kleiner bij kleinere μ_σ, kleinere σ_γ, grotere k, kleinere R (d.i. grotere α en β) en grotere Δ^*. Als men k, m, Δ^* en R geeft, laat zich N berekenen, als althans in (8) μ_σ én σ_γ bekend zijn en in (9) $\bar{\sigma}$ bekend is.

De breuk (8) geeft aanleiding tot de volgende opmerking. De noemer mag niet negatief zijn. Dit impliceert dat men alleen over die Δ^*-waarden "wat te weten komen kan" (d.i. na een eindig aantal maanden n), die groter dan $R\sigma_\gamma/\sqrt{k}$ zijn. Is b.v. k=1 en zijn α = 0,05 en β = 0,01 (d.i. R $\geq$ 4,3), dan zijn alleen de Δ^*'s, waarvoor $\Delta^* > 4{,}3\ \sigma_\gamma$, aantoonbaar. Hoe meer de personen onderling verschillen, hoe groter de σ_γ.

Mocht $\Delta^* \gg R\sigma_\gamma/\sqrt{k}$ (hoe groter k en hoe kleiner σ_γ, hoe eerder deze ongelijkheid geldt), dan gaat (9) over in (10).

(10) $$n = 12\ Nmk \approx R^2\left(\frac{\mu_\sigma}{\Delta^*}\right)^2 \text{ maanden,}$$

waardoor, behoudens het verschil tussen $\bar{\sigma}$ en μ_σ, (9) wordt verkregen.

De formule wordt fraaier als men, bij $\alpha = 0{,}05$ (de waarde, die de meeste statistici kiezen), de β op 0,07 (de $1 - \beta$ op 0,93) stelt, waardoor $R^2 \gtrsim 12$ en (10) overgaat in (10*)

$$(10^*) \qquad \boxed{N \approx \frac{(\mu_\sigma / \Delta^*)^2}{mk}} \quad \text{jaren}$$

Wanneer $k = 1$ en $m = 1$ (in de meeste landen worden de stroken maar éénmaal uitgetrokken en in sommige landen geschiedt dit iedere maand door slechts één persoon, die niet maand voor maand dezelfde is, dan gaat (10*) over in (10**)

$$(10^{**}) \qquad N \approx (\mu_\sigma / \Delta^*)^2 = 1 : (\Delta^*/\mu_\sigma)^2 \quad \text{jaren}$$

Voorbeeld: wil men met een "zekerheid" 0,93 (d.i. met een kans 0,07 op een fout van de tweede soort) en een kans 0,05 op een fout van de eerste soort, te weten komen of de onbekende ware herleidingscoëfficiënt al of niet een bedrag $\Delta^* = \frac{1}{3}\mu_\sigma$ van 1 verschilt, dan moeten er tenminste 9 jaren lang vergelijkende metingen verricht worden, welke metingen dan $9 \times 12 = 108$ c-waarden met een gemiddelde $\bar{c}$, leveren. Kiest men Δ^* wat groter, b.v. $\frac{1}{2}\mu_\sigma$, dan $N = 4$; bij $\Delta^* = 2\mu_\sigma$, dan $N = \frac{1}{4}$ jaar. Als een duur van 9 jaren te lang gevonden wordt, zou men kunnen kiezen b.v. $k = 2$, $m = 2$, waardoor $N = 2\frac{1}{4}$ jaar wordt. Wel wordt ondersteld, dat (10*) toegepast mág worden, d.w.z. dat $\Delta^* \gg R\sigma_\gamma / \sqrt{k}$, en om dat zeker te weten moet men σ_γ kennen (d.w.z. moet men weten hoe "sterk de personen uiteenlopen"). Echter moet óók μ_σ bekend zijn, want anders heeft het weinig zin om te zeggen, dat het verschil $|\gamma_a - 1|$ voor de praktijk "interessant" is, eerst als het b.v. $\frac{1}{3}\mu_\sigma$ bedraagt. Kent men μ_σ? In de meeste landen niet. In ons land is κ_x, "deel uitmakend" van σ_c (zie in hoofdstuk 5), vrij goed bekend uit twee- en meervoudige metingen (zie ook hoofdstuk 10). Het is jammer, dat men (voor zover wij weten) nergens elders duplometingen verrichtte, al was het maar gedurende enige tijd, en toch geven alleen deze een indruk van de onnauwkeurigheden en de intrinsieke waarden der maandtotalen der zonneschijnduur. Het gevolg is, dat de formules (9) t/m (10**) ons van weinig nut zijn, ténzij men de μ_σ durft te schatten zonder werkelijk over duplometingen te beschikken.

Er is nog een moeilijkheid, die boven reeds gesignaleerd werd, te weten het feit, dat iedere maand noch dezelfde persoon, noch dezelfde personen, noch hetzelfde aantal personen de stroken uittrekken. Op het K.N.M.I. b.v. was in de maanden dec. 1964, jan. 1965 en april 1965 de situatie als volgt:

personen	A	B	C	D	E	F	G	H	som	aantal personen
dec. '64	13	12	4	1	1	-	-	-	31	5
jan. '65	7	16	3	1	1	1	2	-	31	7
apr. '65	-	24	2	1	-	1	-	2	30	5

Zo trok A 13 dagen van december uit, zowel de stroken van N.S. als I.S., alle in enkelvoud. B nam 12 andere dagen voor zijn rekening (ook alle stroken één keer). C weer 4 andere dagen, enz. Met zulk een "model" is het zeer moeilijk, zonder speciale onderstellingen, statistisch werken.

Het bovenstaande heeft duidelijk gemaakt, dat het zeker zin hebben zou $m > 1$ te nemen, d.w.z. de routinedienst in elk land de stroken in meervoud uit te doen trekken. Wij kunnen dan én een indruk krijgen van de nauwkeurigheid van het uitmeten van de N.S.- zowel als van de I.S.-stroken én de totale duur van het Project bekorten. Helaas was de Voorzitter van de "Experts Group" van mening, dat zulks niet te organiseren zou zijn.

Formule (10^*) laat o.a. zien, dat Nmk kleiner is bij kleinere μ_σ. Kunnen wij μ_σ klein maken ? De μ_σ stelt, zoals gezegd, voor het gemiddelde aller σ-waarden (gemiddelde van σ_A, σ_B, σ_C etc., waarin σ_A de standaarddeviatie der bij A behorende c-populatie is,etc.) Iedere $\sigma^2 = \gamma^2(\kappa_x^2 + \kappa_y^2)$. Op de κ_x-waarden (N.S.-stroken) hebben wij weinig invloed; zij zijn groot of klein al naar de omstandigheden, die op de verschillende stations in de wereld zeer verschillend kunnen zijn (kwaliteit van personen; scherpten der nationale instructies).Vermoedelijk is de κ_x het kleinst in Nederland, waar (lees [5]) de instructie het meest gedetailleerd is. De κ_y-waarden kunnen we wèl beïnvloeden, althans voor zover het de internationale instructie aangaat. Hoe scherper deze geformuleerd wordt, hoe kleiner de κ_y en hoe geringer dan de bijdrage tot σ_c en via alle σ_c's tot de μ_σ. Daarom adviseerden wij de Voorzitter de internationale instructie "super scherp" te stellen en zeker scherper dan te New Delhi (1962) werd voorgesteld. Helaas was het niet meer mogelijk veranderingen in de Internationale Instructie aan te brengen.

9. HET SCHEMA VOOR DE BEREKENING VAN DE HERLEIDINGSCOËFFICIËNT

Aangezien op elk der aan het Internationale Project deelnemende stations voor iedere maand zowel de stroken van de nationale zonneschijnmeter (Campbell Stokes, Jordan) als die van de internationale referentie zonneschijnmeter (Campbell Stokes) door de nationale meteorologische dienst slechts eenmaal zouden worden uitgetrokken (alleen in De Bilt vonden de uittrekkingen bovendien door twee personen plaats, en wel door ieder van hen in duplo), waarbij bij de eerste stroken de nationale en bij de laatste de internationale instructie gevolgd werd, konden wij volstaan met het maken van één schema voor de berekening van de herleidingscoëfficiënt, (eventueel) toe te passen op de maandtotalen, gemeten met de N.S., ter herleiding op het I.S.-niveau.

Het schema moest zodanig "eenvoudig"*) zijn, dat, indien niet de Experts Groep de berekeningen (86 stations; basismateriaal tenminste 2 jaren) zou uitvoeren, deze door de diensten zelf zouden kunnen plaatsvinden.

Vanzelfsprekend moet in dit schema aandacht geschonken worden onder meer aan die onderstellingen, die noodzakelijkerwijze gemaakt moeten worden, nu er geen duplometingen zijn. Twee belangrijke vragen rijzen

a) heeft de theoretische herleidingscoëfficiënt een jaarlijkse gang ?

b) bezit de nauwkeurigheid van de meting van c een jaarlijkse gang ?

Een volkomen verantwoord statistisch onderzoek van zulk een jaarlijkse gang, doch ook de beantwoording van andere vragen, zou de beschikking eisen over een vele jaren bestrijkend vergelijkingsmateriaal (liever 10 dan 2 jaren). Daar men echter het Project twee, misschien drie, jaren wilde laten duren, moesten wij het probleem anders stellen en hebben wij ons afgevraagd of de uitkomsten ons aanleiding geven de hypothese, dat de theoretische herleidingsfactor niet met het niveau (de totale maandsom) zou samenhangen, al of niet af te wijzen.

Wij geven thans een beschrijving van het "schema van de statistische berekeningen" in fig. 6.

Gegeven: n paren x_i, y_i ($i = 1, 2, \ldots n$); x_i = totale zonneschijnduur, gemeten met de N.S., volgens de nationale instructie, in zekere maand i; y_i = totale zonneschijnduur, gemeten met de I.S., volgens de internationale instructie, in dezelfde maand; in beide gevallen zijn de stroken uitgetrokken (altijd in enkelvoud) door dezelfde plaatselijke meteorologische dienst .

*) Daarom ook kozen wij parameter-vrije toetsen.

Gevraagd:

1) Is het wenselijk een herleidingscoëfficiënt (misschien meer dan één) toe te passen op de maandsom*), gevonden met N.S., ter herleiding op het I.S.-niveau ?

2) Zo ja, welke zijn de puntschatting en de 95%-intervalschatting van de theoretische herleidingsfactor γ ?

Oplossing:

Stap A. Wij beginnen met als nulhypothese H_A te stellen, dat x en y in elk der n maanden alleen door toevallige fouten verschillen, d.w.z. dat $(\mu_x)_i = (\mu_y)_i$ voor iedere i, en toetsen deze met de Rang-Teken-Toets van Wilcoxon [1]. Het antwoord op de vraag of (op de 5% dempel van significatie) deze H_A verworpen moet worden kan zijn:

1. Neen. Stop. Conclusie: het heeft geen zin een herleidingscoëfficiënt toe te passen.

2. Ja. Doe nu stap B.

Stap B. De nulhypothese H_B is thans: de theoretische herleidingscoëfficiënt hangt niet met het niveau μ_x samen. Pas de Toets van Spearman toe [2]. Daartoe wordt de Spearmanse correlatiecoëfficiënt tussen c en x tegen nul getoetst; er zijn n paren c_i, x_i met $c_i = y_i/x_i$. Op de vraag of deze H_B verworpen moet worden zijn weer twee antwoorden mogelijk:

1. Neen. Bereken $\bar{c} = \frac{1}{n} \Sigma c$. Deze $\bar{c}$ is een zuivere schatting van de gezochte herleidingscoëfficiënt. Als mocht blijken, dat het toepassen van een herleidingscoëfficiënt statistisch zinvol is (dat hangt af van wat de stappen D en E leveren), dan met deze $\bar{c}$. Ga naar C.

2. Ja. Verdeel de n c-waarden over zovele (zegge p), niet te smalle, klassen, dat H_B in geen dezer klassen verworpen behoeft te worden. Doe daarna in elk der klassen (in fig. 6 is p = 2) stap C.

Stap C. Terwijl γ niet met μ_x samenhangt, kan $\sigma(c)$ het wel doen. Stel daarom de nulhypothese H_C: $\sigma(c)$ hangt niet met μ_x samen en toets deze met de Toets van Hartley. Verdeel daartoe de n c-waarden over q klassen van

---

*) Men kan vragen: waarom niet voor decade-sommen ? Dan is het basismateriaal 3 keren zo groot. Echter rijzen dan (in sommige landen meer dan in andere) moeilijkheden met gevallen als x = 0, y = 0, c = ? en x = 0, y ≠ 0, c = ? Bovendien: is voor kleine μ_x wel aan $\kappa_x \ll 1$ voldaan ? Maandsommen en zeker seizoensommen zijn zelden nul. Natuurlijk kunnen wij de genoemde moeilijkheden omzeilen door niet op procentuele maar op additieve correcties aan te sturen. Zie echter de noot op pg. 24 .

(niet te on)gelijke grootte. Bereken in elke klasse de variantie s_j^2 ($j = 1, 2, \ldots q$). De toets verifieert of deze q varianties "statistisch gelijk" zijn. Op de vraag of H_C verworpen moet worden, zijn weer twee antwoorden mogelijk:

1. <u>Neen</u>. Ga verder met stap D.
2. <u>Ja</u>. Voer stap D uit voor elk der genoemde q groepen afzonderlijk (in fig. 6 is $q = 2$). Eventueel zullen er q $\bar{c}$-waarden komen, die in hun $\sigma(\bar{c})$'s statistisch verschillen.

<u>Stap D</u>. Bereken nu ook de s(c), d.i. de standaarddeviatie der n c-waarden; (of voor zovele c-waarden als er in de ondergroep zijn).

$$s^2(c) = \{\Sigma c^2 - (\Sigma c)^2/n\} : (n - 1).$$

Dan stelt $s(\bar{c}) = s(c)/\sqrt{n}$ de nauwkeurigheid van $\bar{c}$ voor. Terwijl $\bar{c}$ de "puntschatting" van γ is, wordt de "intervalschatting van γ " gegeven door het zg. 95% betrouwbaarheidsinterval

$$\bar{c} - 2\, s(\bar{c}) < \gamma < \bar{c} + 2\, s(\bar{c}).$$

Mocht n "zeer klein" zijn (dus vooral in de ondergroepen), dan moet men de factor 2 vervangen door de uit de Student-tabel bij n-1 graden van vrijheid en 5% af te lezen "kritieke waarde τ ". Zo is bij $n = 5$ de $\tau = 2{,}8$; bij $n = 20$, $\tau = 2{,}1$; bij $n \geq 50$, $\tau \approx 2{,}0$. Op de vraag of het betrouwbaarheidsinterval de waarde 1 bevat, zijn twee antwoorden mogelijk:

1. <u>Ja</u>. Het is dan o.i. niet nodig een herleidingscoëfficiënt toe te passen (ook al is, natuurlijk, $\bar{c} \neq 1$).
2. <u>Neen</u>. Nu ligt het betrouwbaarheidsinterval òf geheel beneden òf geheel boven 1. Zulks impliceert o.i. toch nog niet, dat het zinvol is een herleidingscoëfficiënt toe te passen. Men zou twee <u>beslissingscriteria</u> kunnen opstellen. Bij beide is de grondgedachte, dat het slechts dan zin heeft een herleidingscoëfficiënt toe te passen als deze voldoende van 1 verschilt, d.w.z. voor een bedrag, dat "veel groter" is dan de variatie tussen de individuele c-waarden. Zo komen wij er toe twee criteria te formuleren (welk gekozen moet worden, kan de statisticus niet zeggen):

a) <u>in relatieve zin</u>. Kies een f^*, b.v. 1 of 2 en pas $\bar{c}$ eerst toe als $f = |1 - \bar{c}| : s(c) > f^*$. De keuze van de numerieke waarde van de constante f^* is geen statistische zaak, maar een van international overleg.

b) <u>in absolute zin</u>. Kies een Δ^*, b.v. 0,15 of 0,20 en pas $\bar{c}$ eerst toe als $\Delta \equiv | 1 - \bar{c} | > \Delta^*$. Ook de keuze van Δ^* is geen statistische zaak, doch een van internationaal overleg.

<u>Stap E</u>. Bereken f (of Δ). Op de vraag of $f > f^*$ (of $\Delta > \Delta^*$) is, zijn twee antwoorden mogelijk:

1. <u>Neen</u>. Pas geen herleiding toe.
2. <u>Ja</u>. Pas wel herleiding toe.

In fig. 6 is met de gestippelde weg het eenvoudigste geval van een herleiding aangegeven (met $f^* = 1$). Er resulteert één herleidingscoëfficiënt, op welk niveau de maandsom ook gelegen moge zijn.

<u>Enige opmerkingen</u>.

a. Als men besloten heeft een herleidingscoëfficiënt toe te passen, maar om een of andere reden de $\bar{c}$ nog niet nauwkeurig genoeg acht, d.w.z. als men de breuk $g_n \equiv s(\bar{c})/\bar{c}$ nog niet voldoende klein vindt, zette men de vergelijkende metingen voort tot er $n'(> n)$ paren x en y zijn. Geeft men g^* vooruit (b.v. 0,01; 0,10) en wenst men een $g_{n'}$ (op basis van n' paren x, y), kleiner dan g^*, dan is nodig, dat

$$n' \geqslant n(g_n/g^*)^2$$

b. De volgende tabel illustreert waar in het gehele statistische betoog <u>de min of meer subjectieve keuzen</u> gelegen zijn.

	vraag	antwoord eist keuze van	zie
duur Project	minimum duur	$\alpha, \beta, \mu_\sigma / \Delta$, k, m	form.8
γ-interval ligt buiten 1	herleiding uitvoeren?	α; f^* of Δ^*	$f \equiv \frac{\lvert 1-\bar{c}\rvert}{s(\bar{c})} > f^*$? of $\Delta \equiv \lvert 1-\bar{c}\rvert > \Delta^*$?
gevonden herleidings-coëfficiënt $\bar{c}$	$\bar{c}$ voldoende nauwkeurig ?	g^*	$n' > n(g/g^*)^2$ met $g \equiv [s(\bar{c})/\bar{c}]_n$

Schema der Statistische berekeningen

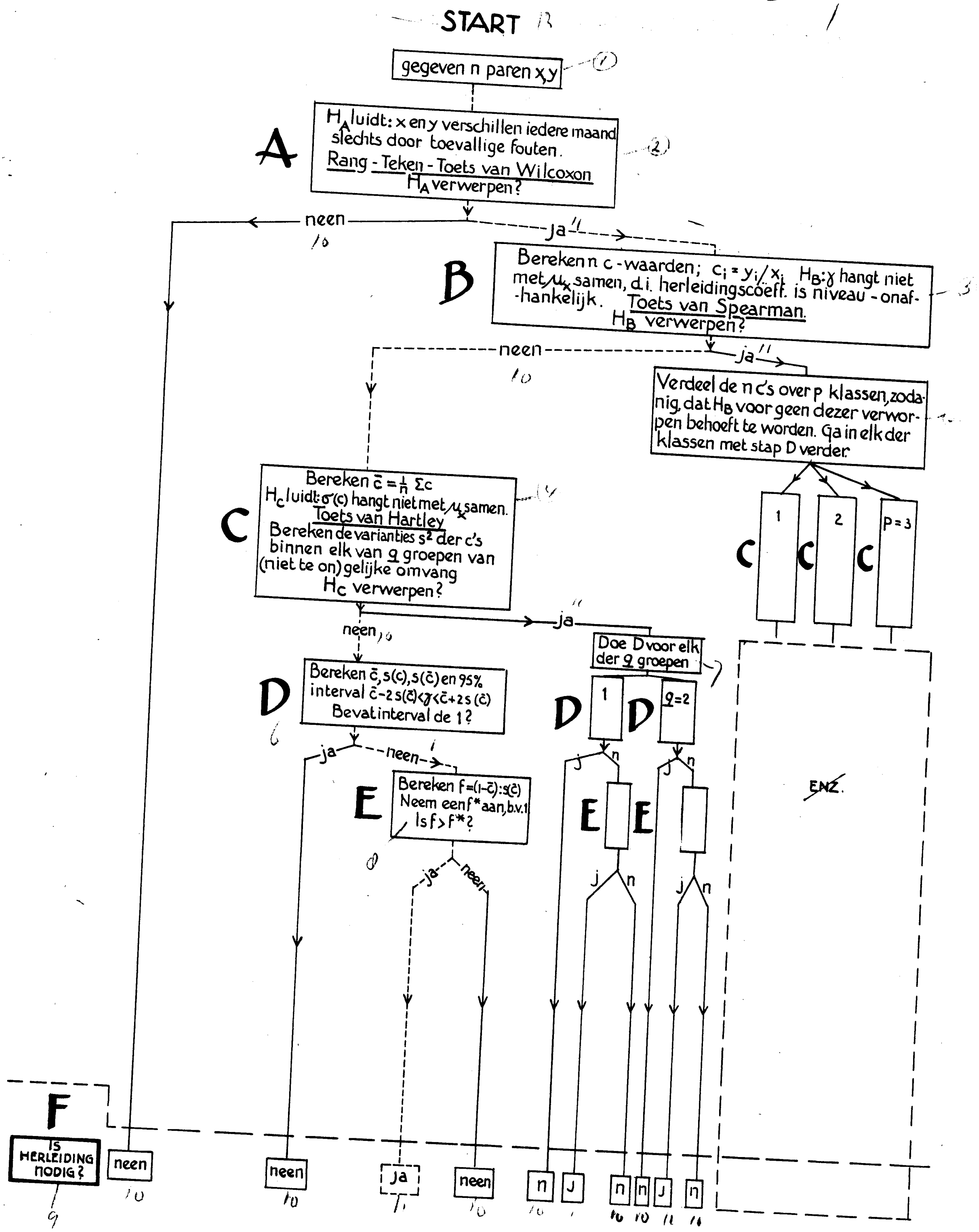


Fig. 6

10. ENKELE NUMERIEKE RESULTATEN.

Wij beëindigen het rapport men numerieke uitkomsten. Het zijn er slechts enkele om twee redenen:

1) Zoals reeds gezegd, is het de bedoeling in dit rapport meer aandacht te schenken aan de statistische aspecten van het vergelijkingsprobleem dan aan concrete kwantitatieve resultaten.

2) Op het moment van afsluiten van het onderhavige verslag (mei 1965) staan ons slechts van 7 buitenlandse stations gegevens ter beschikking, die betrekking hebben op tijdvakken van slechts 5 tot 10 maanden. In De Bilt werd het onderzoek beëindigd met de maand mei 1965, omdat de berekeningen leerden, dat een herleiding tot het I.S.-niveau niet nodig is. Gedurende het tijdvak november 1963 tot en met december 1964 werden te De Bilt door twee op het uittrekken van zonneschijnstroken geoefende, van de Klimatologische Dienst deel uitmakende, personen P en Q zowel de N.S.-stroken (volgens de nationale instructie) als de I.S.-stroken (volgens de internationale instructie) uitgetrokken en wel in duplo, terwijl de klimatologische routine dienst (variërende én van grootte, én van samenstelling) de N.S.- en de I.S.-stroken in enkelvoud uitmat, gedurende het tijdvak juni 1964 tot en met mei 1965.

Wij volgden voor Nederland het in fig. 6 opgenomen schema, maar hebben daarin, op nadrukkelijk verzoek van Dr. Lamboley, een sterke vereenvoudiging aangebracht bij toepassing op de gegevens van stations in het buitenland. De vereenvoudiging bestaat hierin, dat wij in stap B de Toets van Spearman vervingen door de Twee-Steekproeven - Toets van Wilcoxon, waartoe de x-waarden in twee groepen van gelijke omvang ondergebracht werden, te weten $x <$ mediaan en $x >$ mediaan. Verder werd de Toets van Hartley (stap C) achterwege gelaten. Zo resulteren er enige tabellen, waarvan wij er hier drie publiceren:

Tabel A bevat o.m. de waarden van κ , d.i. de relatieve meetonnauwkeurigheid (hoofdstuk 5), die juist via duplometingen berekend kan worden.

Tabel B vermeldt de antwoorden op de vragen of herleidingen op het I.S.-niveau wenselijk zijn, als óf alleen P óf alleen Q òf de routinedienst alle stroken zou uittrekken.

Tabel C geeft een overzicht van al die rekengrootheden, die tot het antwoord op de twee vragen "al of niet herleiden ?" en "zo ja, waarmee ?" leidden, voor De Bilt (routine dienst) en voor 7 buitenlandse stations. Beneden elke tabel staat enig commentaar.

TABEL A

De Bilt, november 1963 - december 1964;
14 maanden. Nauwkeurigheidsresultaten ,
verkregen uit duplometingen

	uren		in %	
	$\overline{N.S.}$	$\overline{I.S.}$	N.S.	I.S.
P	118,6	118,2	1,4	1,2
Q	118,0	113,4	0,4	3,1

Toelichting:

Ieder van twee ervaren personen P en Q trok zowel de N.S.-stroken (de nationale instructie volgende) als de I.S.-stroken (de internationale instructie volgende) twee keren uit. Aldus verkreeg P voor b.v. november 1963 twee maandsommen voor N.S. en twee maandsommen I.S. en 2 onafhankelijke c-waarden I.S./N.S. Zo ook voor de 13 andere maanden en zo ook voor Q. Met behulp van deze duplometingen kan zowel voor P als Q de procentuele standaarddeviatie κ (deze onafhankelijk van het niveau gedacht*)) berekend worden.

De gemiddelde waarden van de maandsommen zijn $\overline{N.S.}$ en $\overline{I.S.}$

Conclusies:

a) Q trok de N.S.-diagrammen nauwkeuriger uit dan P, doch de I.S.-diagrammen minder nauwkeurig. Immers $0{,}4 < 1{,}4$ en $3{,}1 > 1{,}2$.

b) P trok de N.S.- en de I.S.-stroken ongeveer even nauwkeurig uit (immers $1{,}4 \approx 1{,}2$), maar Q trok de N.S.-stroken veel nauwkeuriger uit dan de I.S.-stroken (immers $0{,}4 < 3{,}1$).

---

*) Wij hebben de hypothese van onafhankelijkheid getoetst; ze behoefde niet verworpen te worden, noch voor P, noch voor Q.

TABEL B

De Bilt, november 1963 - december 1964;
14 maanden. Herleidingscoëfficiënten voor maandelijkse zonneschijnduren

tijdvak		n	$\bar{c}$	s(c)	s($\bar{c}$)	95% betrouwbaarheids-interval voor γ	bevat het de 1 ?	f= | 1-$\bar{c}$ | : s(c)	f > 1 ?	herleiding gewenst ?
nov.1963 - dec.1964	P	28	1,006	0.030	0,006	0,995 tot 1,017	ja	0,2	neen	neen
	Q	28	0,970	0,049	0,009	0,951 tot 0,989	neen	0,7	neen	neen
juli 1963 - dec.1964	routine-dienst	6	0,999	0,033	0,014	0,971 tot 1,027	ja	0,0	neen	neen
de maandsommen liepen uiteen van 36 tot 240 uren.										

Toelichting: De stroken van de N.S. en die van de I.S. werden uitgemeten door twee ervaren personen P en Q en door de "routine dienst", allen behorende tot Afd. 3 van het K.N.M.I. Daarbij deden P en Q het ieder in duplo, doch de routine dienst in enkelvoud. Aldus verkregen P en Q ieder n = 2 x 14 onafhankelijke c-waarden, met c = I.S./N.S. De 28 c-waarden bij P lagen om een gemiddelde $\bar{c}$ = 1,006, met een standaarddeviatie s(c) = 0,030. Dan stelt $s(\bar{c}) = s(c)/\sqrt{n}$ = 0,006 de nauwkeurigheid van $\bar{c}$ voor. Het 95% betrouwbaarheidsinterval voor de onbekende herleidingscoëfficiënt γ ligt dan tussen $\bar{c} - 2s(\bar{c})$ en $\bar{c} + 2s(\bar{c})$, d.i. tussen 0,995 en 1,017, zodat de eenheid 1 erin valt. Bijgevolg is een herleiding van N.S. op I.S. niet gewenst. Om dezelfde reden evenmin voor de routine dienst; nu n = 6 (geen duplo's), etc. Ofschoon voor Q het 95% betrouwbaarheidsinterval voor γ_Q wel buiten 1 (er onder) valt, is echter het "beslissingsquotiënt" f = 0,7 en dus $f < f^*$, indien, arbitrair, f^* op 1 gesteld wordt, zodat toch geen herleiding gewenst is.

Conclusie: Noch voor P, noch voor Q, noch voor de routine dienst schijnt toepassing van een herleidingscoëfficiënt statistisch zinvol.

TABEL C

Resultaten van de Vergelijking tussen simultane metingen van de maandelijkse zonneschijnduur, gemeten met de N.S. en de I.S., op 8 stations in de wereld

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
tijdvak	station	land	N.S. is Campbell S. of Jordan	N mnd.	R.S. T.W.	W	$\bar{c}$	s(c)	s($\bar{c}$)	95% betr. interval voor γ	bevat het 1 ?	$f=\frac{\lvert 1-\bar{c}\rvert}{s(c)}$	f> 1?	is herleiding noodzakelijk ?	voeg toe aan N.S. %
nov.63-feb.65	De Bilt	Nederland	C	16	ja	ja								neen	-
jan.64-jun.65	Jokionen	Finland	C	5	ja	ja								neen	-
jul.64-feb.65	Narey	Frankrijk	J	8	neen	ja	0,968	0,023	0,008	0,951-0,985	neen	1,4	ja	ja	-3,2
mei 64-feb.65	Rennes	Frankrijk	J	10	neen	ja	0,963	0,030	0,009	0,964-1,002	ja	0,6	neen	neen	-
apr.64-okt.64	Tokyo	Japan	J	7	neen	ja	0,862	0,051	0,019	0,824-0,900	neen	2,7	ja	ja	-13,8
apr.64-okt.64	Tokyo	Japan	J	7	neen	ja	0,918	0,058	0,022	0,874-0,962	neen	1,4	ja	ja	-8,2
apr.64-sep.65	Nairobi	Kenya	C	6	neen	ja	0,854	0,030	0,012	0,830-0,874	neen	4,9	ja	ja	-14,6
mrt.64-okt.64	Kelburn	New Zealand	C	8	neen	neen	0,937	0,010	0,005	0,927-0,947	neen	6,5	ja	ja	-6,3[1)]
						neen	0,955	0,008	0,004	0,947-0,963	neen	5,8	ja	ja	-4,5[2)]

Toelichting: Voor het beslissingsquotiënt f^* werd weer 1 gekozen.
Kolom 6 geeft het antwoord op de vraag:"Verschillen de maandsommen, gevonden met N.S. en I.S., iedere maand louter vanwege toevallige fouten ?"Getoetst werd met de"Rank-Sign-Test of Wilcoxon"(R.S.T.W.). Kolom 7 geeft het antwoord op de vraag:"Is de theoretische herleidingscoëfficiënt γ onafhankelijk van het niveau ?" Getoetst werd met de"Two-Sample-Test of Wilcoxon"(W).

Noten:
1) De gehele regel geldt voor N.S.-maandsommen < 143 h.
2) De gehele regel geldt voor N.S. > 143 h.; hierbij is 143 de mediaan der 8 x-waarden.

11. SLOTOPMERKINGEN

11.1 De Campbell-Stokes zonneschijnautograaf (C.S.) in zijn kwaliteit*) van meter van de zonneschijnduur is hier niet in discussie geweest, hoe belangrijk dit aspect ook moge zijn. Dat was ook niet de bedoeling van hen, die tot het Project het initiatief namen. Een feit is, dat men moeilijkheden ondervond en ondervindt b.v. bij het tekenen van isoheliën-patronen op "wereldschaal", welke moeilijkheden men geneigd is toe te schrijven aan systematische verschillen. Het staat vast, dat men de C.S. in vele landen nog vele jaren gebruiken zal, ook al zijn er (ernstige) bezwaren tegen het instrument. Ongetwijfeld is het dan nuttig een kwantitatieve voorstelling te verkrijgen van de (betrekkelijke) waarde van de metingen met de C.S., opdat men weten zal, in welke onderzoekingen men het instrument wel, en in welke niet gebruiken mag. Een statistisch onderzoek als het bovenstaande zou kunnen helpen zulk een kennis te verschaffen.

11.2 Natuurlijk rijzen vragen als: hoe willen wij fysisch verklaren, dat eenzelfde persoon bij herhaalde uitmetingen van de stroken van eenzelfde, gegeven, maand wisselende maandtotalen krijgen zal ?; hoe komt het, dat zulks bij de ene persoon markanter het geval is dan bij een andere ?;

---

*) 1. Zo gebeurt het, dat men de zon ziet (door dunne wolken heen; bij opkomst en ondergang), ofschoon er op de strook niet het flauwste brandspoor ontstaat.

2. Zo gebeurt het, dat men visueel onderbreking in de zonneschijn constateert (voorbijtrekken van wolken), hoewel in het brandspoor niets of nauwelijks iets van insnoeringen terug te vinden is.

3. Zo gebeurt het, dat de zon heel even schijnt (enkele seconden) en daarbij brandpunten in de strook veroorzaakt, die gemakkelijk, bij het uittrekken van de strook, qua zonneschijnduur overschat worden. Enz.

Aan sommige van deze bezwaren (waarvan de betekenis nog van de ligging van het station - gematigde streken, tropen - kan afhangen) is wel tegemoet te komen door de voorschriften voor het uittrekken van de stroken zeer zorgvuldig te formuleren (gelijk in Nederland geschiedde), maar de instructie wordt dan al gauw zeer gecompliceerd. Hier rijzen ook vragen als: hoe definieert men de zonneschijnduur (E: "bright sunshine") ? De tijd, dat men de zon met ongewapend oog ziet ? De tijd, zolang de zon zichtbaar schaduwen werpt? De tijd, dat de zon met een intensiteit $\geqslant$ i cal/cm^{2}min schijnt? Welke i ? 0,03; 0,3; 3,0 ? Bij het ene soort papier ligt de i-drempel, waarboven een eerste indicatie van een brandpunt te zien is, een orde van grootte anders dan bij een ander soort papier. Al deze, zeker interessante, zaken, die reeds jaren een punt van studie vormen (niet te De Bilt), vallen buiten het kader van het onderhavige rapport.

hoe komt het, dat ook de scherpte, waarmede de uittrekregels geformuleerd zijn, van invloed kan zijn ? Men voelt wel aan, dat ook (het niet gemakkelijk te definiëren begrip) "de aard" van de maand er toe doet. Als op iedere dag de zon ononderbroken geschenen zou hebben, dan zouden de bij herhaling van metingen verkregen maandsommen weinig of niet spreiden, noch voor persoon A, noch voor persoon B. Maar als de zon met zeer vele onderbrekingen geschenen zou hebben, dan zou de bedoelde spreiding wel eens groot kunnen zijn. Het is het "karakter" van de brandsporen (flauwe inzetten; interrupties of afwisselingen; insnoeringen; brandpunten), die de grootte van de standaarddeviatie σ der randomfouten bij de uittrekkingen veroorzaakt. De ene maand is "veel moeilijker" dan de andere. Uit dien hoofde zullen er vele maanden nodig zijn voor en aleer er een betrouwbare σ resulteert en metingen in tweevoud, liefst in b.v. tienvoud, door ieder van vele assistenten, zullen niet gemist kunnen worden. Men zou dus de herleidingsfactor ook nog van het "karakter" van de maand kunnen laten afhangen ! Doch is dit in de praktijk uitvoerbaar ? Wij vertelden reeds, dat het ons niet gelukte, anders dan in Nederland, duplo-uittrekkingen (en uittrekkingen tot in vijfvoud) door verschillende personen te doen verrichten, terwijl C.I.M.O. het Project het liefst niet langer dan twee jaren wilde doen duren.

11.3 Wij zijn ons bewust, dat er in de praktijk vele effecten zijn, waarmede helaas zeer moeilijk in het herleidingsprobleem rekening gehouden kan worden. Hier volgen er enige:

a) Na regen, zware dauw, dikke mist, gevolgd door plotseling doorbrekende zonneschijn kan een kwaliteitsverandering van de strook nadelige gevolgen op de scherpte van het brandspoor hebben.
b) De stroken, ofschoon van dezelfde kleur en dikte gebleven zijnde, kunnen van zekere dag af van andere kwaliteit (samenstelling) zijn, zonder dat de leverancier dit mededeelt of zelfs weet.
c) De helderheid van de bol kan geleidelijk minder worden.
d) Het is te begrijpen, dat niet bij iedere C.S. zonneschijnmeter de drie mantels "volmaakt correct" rondom de door het bolmiddelpunt gedachte hemelas (// aardas) gecentreerd zullen zijn. Wij zullen dit als een technische tolerantie moeten aanvaarden, doch welke is de invloed van deze tolerantie op de nauwkeurigheid van de meetresultaten ?

e) Er kunnen zeer geleidelijk veranderingen in de wijze van uittrekken van de stroken ontstaan, hoewel men zich dit niet bewust is. Hoe scherper de formulering van de instructie, hoe kleiner de kans op zulke veranderingen.

f) Nergens zal de routinedienst van grootte, kwaliteit en samenstelling constant zijn.

g) In sommige landen worden de stroken door de dienst van het station, waar de N.S. staat, uitgemeten. Er zijn dan zovele routinediensten als er stations zijn. Is het dan zinvol om op alle stations van zulk een land met éénzelfde herleidingscoëfficiënt te werken ?

Al deze opmerkingen bedoelen duidelijk te maken, dat de "werkelijke waarde" van een herleidingscoëfficiënt vaak gering, althans twijfelachtig, is hoe verantwoord hij ook berekend moge zijn. De herleidingscoëfficiënt werd gevonden tijdens het Project, maar men gaat hem toepassen op voorbije en toekomstige metingen, hetgeen, strikt beschouwd, slechts mag als letterlijk alles (maar wat is dit precies ?) volmaakt hetzelfde was en blijft als gedurende het Project. Velen menen: ook al is de gevonden reductiefactor niet "de beste", "enigszins reduceren" zal vaak "beter zijn" dan helemaal niet.

11.4 Een fraaie statistische beschouwing van toevallige en systematische fouten van instrumentele en persoonlijke aard, vindt men in [4] .

○-○-○-○-○

12. LITERATUUR

[1]	Rangtekentoets van Wilcoxon	Stat. Tabellen en Nomogr. (onder redactie van Ver. v. Stat.), no. 19.1 en 19.2.
[2]	Rangcorrelatiecoëfficiënt van Spearman	P.Rijkoort; K.N.M.I. R III 120 1953; toets III, 4.
[3]	Toets van Hartley	Stat. Tabellen en Nomogr. (onder redactie van Ver. v. Stat.), no.22.1*
[4]	Etude statistique des erreurs de mesure. P.J.Delaporte	Bull.de l'Inst. Internationale de Stat. 38 387 (1961).
[5]	Über die Auswertung der Registrierungen des Sonnenscheinautographen Campbell-Stokes. C. Levert	Arch. für Meteor., Geoph. und Bioklim., Serie B 11 135 (1961).
[6]	Über die Auswertung der Registrierungen des Sonnenscheinautographen Campbell-Stokes. M. Bider	Arch. für Meteor., Geoph. und Bioklim., Serie B 9 199 (1958).

.-.-.-.-.

13. SUMMARY

Statistical considerations with respect to the question how to reduce the measurements with different types of instruments and (or) techniques to a common basis.

Chapter 1

An examination of published values of duration of bright sunshine will show that differences of up to even 20 per cent in monthly totals can be explained by the use of different types of instruments and recording paper and by the use of different methods of measuring the cards. In order to reduce these monthly totals to a common basis, C.I.M.O., at its Third Session,(at New Delphi in 1962 January), proposed in its Recommandation 6, incorporated by the Executive Committee in its Resolution 23 (E.C.-XIV), to adopt as the Interim Reference Sunshine Recorder (I.S.) the Campbell-Stokes sunshine recorder made in compliance with the specifications of the British Meteorological Office, and used with the record cards which meet the detailed requirements,made by the French Meteorological Service. The reduction coefficient should be determined by a careful sufficiently long comparison between the national sunshine recorder (N.S.)*) and I.S. The Experts Group, which should provide guidance for Members on the comparison of sunshine recorders and study the results of comparisons reported by Members, consisted of R. Lamboley (chairman), R.H. Collingbourne, T.H. Mac Donald (in the course of 1964 replaced by A.N. Hill) and C. Levert. The chairman wrote a "Draft Guide for the International Comparison of Sunshine Recorders", which chiefly refers to the installation of the recorders and the completion of special forms, whereas Levert made 1) a study on the statistical aspects with regard to this comparison of instruments and techniques and 2) constructed a diagram, showing how to calculate the reduction coefficient on the basis of the monthly pairs of total duration as measured by the N.S. and the I.S. during the period of comparison.

The present report deals with these statistical considerations, which probably could also be useful for all instruments, furnishing records (diagrams, cards, etc.), which must be read off, measured or analyzed in

---

*) Four main types of instrument are available: a) Campbell-Stokes type, b) Jordan-type, c) Marvin-type, d) Foster-type.

more or less detail, before obtaining figures to be used in further work. Then errors of systematic and random nature, caused by instruments as well as by persons, will arise.

Chapter 2

In this chapter definitions are given of the systematic error (ε) and the random error ($\underline{\delta}$). If person A would measure the same monthly set of cards again and again, following always the same rules of evaluation, a population of monthly durations (mean value μ) would arise with standard deviation σ. It is explained that ε, σ, μ could depend on the given set of cards, the person and the measuring rules.

Chapter 3

In this chapter a statistical definition of the concept "comparable" is given. This is necessary because the purpose of applying reduction coefficients to the N.S.-monthly totals, in order to bring them to a common basis, is to make these totals "more comparable".

Chapter 4

In this chapter two possible procedures for the International Comparison between the N.S. and the I.S. are dealt with and discussed. The procedure which was recommended by C.I.M.O. is as follows:
"At each of all participating stations (to the number of 86) an I.S. must be installed near to the N.S. Simultaneous measurements of the monthly sunshine durations have to be made during several months. The N.S.-cards as well as the I.S.-cards have to be read off by the local service, but for the N.S.-cards the national instruction and for the I.S.-cards the international one has to be followed. In this way at each station for each individual month a ratio of the simultaneous monthly totals (which is an estimate of the theoretical reduction coefficient) has to be computed. The decision whether the mean value of this ratio over several months should be applied to passed and future monthly sums, given by the N.S., could depend on special statistical criteria."

Stress is laid upon two statistically regrettable points:

a) the International Rules for Evaluation are so little precise that personal interpretation certainly will occur;

b) on the ground of the fact, that two persons will always obtain different monthly totals of duration of sunshine when measuring, mutually independently, the same monthly set of cards, even if they both follow the same rules of measuring the cards (the less precise these rules, the more marked this difference), we think, it would have been better that C.I.M.O. had recommended that at each station, wherever situated in the world, the I.S.-cards have to be measured by only one and the same person or by only one and the same national service.

Chapter 5

In this chapter the fact is stressed that, in a sense, the reduction coefficient is a statistical variable. Since both N.S.-measurement of the total sunshine (x) of e.g. January 1964 and the I.S.-measurement of the total duration (y) of the same month is inaccurate, also the ratio $c = y/x$ is inaccurate. This inaccuracy could depend on the person, who reads the cards,as well as on the monthly sum itself or on the number of interruptions of sunshine in this particular month, etc.

Chapter 6

In this chapter reference is made to an interesting, statistical problem in the field of (chemical) analyses in the laboratory. This problem is: If each of k analysts, chosen at random from a hypothetical population of analysts, makes m measurements of the same unknown true value $\overset{o}{\gamma}$ (e.g. the P_H of a solution), then the overall mean $\bar{\underline{c}}$ of the k.m measurements c_i will be a statistical estimate of $\overset{o}{\gamma}$. The theoretical mean value and the standarddeviation (accuracy) of this estimate are given and discussed.

Chapter 7

In this chapter the analogy of the statistical reasoning and the formulae given in the preceding chapter are applied to the reduction problem in question. Now each of k persons, again chosen at random from a hypothetical population of persons, who all could measure the cards, evaluates m times a couple of monthly sets of cards (N.S.; I.S.), each time obtaining a reduction coefficient c. The mean $\bar{\underline{c}}$ of these k.m c-values represents a stochastical estimate of the theoretical reduction coefficient $\overset{o}{\gamma}$. But now mind the difference with chapter 6: m = 3 could imply that e.g.

person A would measure both the N.S.-cards and the I.S.-cards of three consecutive months, but m = 3 could also mean that this person would remeasure the same sets of cards three times. The statistical aspects and the underlying hypotheses are treated fully in this chapter.

Chapter 8

In this chapter an estimate is given of the minimum duration of the Project of Comparison. Therefore first of all in section 8.1 some words are said on the "Theory of testing hypotheses" and the "errors of the first and second kind". The central problem sounds: "How many measurements (of given accuracy) must be made in order to know, "reasonably well", whether the difference between the unknown true value and a hypothetical value is of a prescribed size, whereas the chance of an error of the first kind will be α and the chance of an error of the second kind will be β ?"

The theoretical results are applied in section 8.2. Simplifying the formulae, the minimum duration (N years) of the Project can be calculated approximately, The value of N will depend on m, k, μ_σ and Δ^*. Here k stands for the number of persons; each of them evaluates each month both the N.S.- and I.S.-cards. The symbol Δ^* represents the difference between the reduction coefficient and the unit 1, which is felt as sufficiently interesting. The μ_σ stands for the theoretical mean value of the σ_0-values of all persons of the population, where, for instance, σ_A represents the accuracy with which person A calculates the c-value, which accuracy itself depends both on the national and the international instruction. Unfortunately in all countries but one duplo-measurements could not be made during the Project. Only in the Netherlands measurements in duplo by two persons took place, whereas measurements in quintuple have been performed by six persons over a period of some months.
Because of the general lack of figures on the accuracies of evaluation it is hardly possible to estimate the minimum duration of the International Project of Comparison even in an approximate way. For that reason it was decided to start with a period of comparison of two years and then to study the results.

Chapter 9

In this chapter a diagram is given for the calculation of the reduction coefficients in a statistically systematical way. Hypotheses must be verified wherever they must be made. Distribution-free tests have been chosen, in order to keep the calculations simple. The principle is: n couples x_i, y_i give n c_i-values, where $c_i = y_i/x_i$, with a mean $\bar{c} = \Sigma c/n$; the standard deviation s(c) of these c-values is found as $\sqrt{\{ \Sigma c^2 - (\Sigma c)^2/n \} : (n-1)}$ and the accuracy of the mean $\bar{c}$ is given by $s(\bar{c}) = s(c)/\sqrt{n}$. The point-estimate of the theoretical reduction coefficient γ is given by $\bar{c}$, and the 95% reliability-interval of γ is expressed by $\bar{c} \pm 2\, s(\bar{c})$. If this interval would contain the unit 1, then, in our opinion, it seems not necessary to apply a reduction coefficient. If it would lie outside the unit, then it seems necessary to apply a reduction only if the following "decision criterion" $|1-\bar{c}| : s(c) > f^*$ is fulfilled. The underlying idea is that it is not senseful to apply a reduction if the difference between $\bar{c}$ and 1 is of the same order of size as the variation between the n individual c-values. The numerical value of the constant f^* (e.g. 1, 2, 3) should be agreed internationally.

Chapter 10

In this chapter some numerical results are given. See the tables A, B, C. These results chiefly refer to the questions: "should we reduce?" and "if so, how is the reduction coefficient?" For De Bilt also accuracy results are shown, relating to two persons and based on duplo-measurements. Calculations of the reduction coefficients have been made for De Bilt, and for 7 stations in Japan, France, Norway, Finland, New Zealand and Africa.

Chapter 11

In this chapter some aspects are mentioned, which show that the intrinsic value of the reduction coefficient, how well it may be computed, can never be very great. The coefficient is found during the Project, but it is meant to apply it to passed and future monthly totals; this is fully correct only if all circumstances during these passed and these future measurements have been and will remain quite the same as those during the Period of Comparison.

Notwithstanding this fact it could be important to obtain any quantitative idea of what can be done with the Campbell-Stokes sunshine-recorder and in what investigations the instrument could be reasonably useful and in what investigations quite a different apparatus should be preferred.

o-o-o-o-.

Interim Reference Sunshine Recorder

↘ denotes I.R.S.R.

↘ denotes N.S.