ULTUURTECHNIEK EN WATERHUISHOUDING

NOTA 395, d.d. 7 juni 1967

Enkele beschouwingen bij het samenbrengen van onderdelen van cultuurtechnische onderzoekingen

Ph. Th. Stol

Nota's van het Instituut zijn in principe interne communicatiemiddelen, dus geen officiële publikaties.
Hun inhoud varieert sterk en kan zowel betrekking hebben op een eenvoudige weergave van cijferreeksen, als op een concluderende discussie van onderzoeksresultaten. In de meeste gevallen zullen de conclusies echter van voorlopige aard zijn omdat het onderzoek nog niet is afgesloten.
Bepaalde nota's komen niet voor verspreiding buiten het Instituut in aanmerking.

Inhoud | Pag.

Inleiding

Ten behoeve van het onderzoek in de Gelderse Achterhoek worden verschillende aspecten van de neerslag bestudeerd. Bij deze onderzoekingen doet zich de vraag voor in hoeverre de resultaten in een samenvattend onderzoek kunnen worden ingepast. Zowel de toepasbaarheid als de betrouwbaarheid van de bestudeerde grootheid dienen onderzocht te worden, wanneer het er om gaat deze in het samenvattende onderzoek op te nemen.

De toepasbaarheid is overigens een aspect dat in het samenvattende onderzoek aan de orde komt. Het bestuderen van een of andere grootheid kan wetenschappelijke waarde hebben, bij het invoegen van deelresultaten in een samenvattend onderzoek dient de te bestuderen grootheid wel overwogen gekozen te worden.

In veel gevallen zal nog blijken dat niet de gekozen grootheid zelf doch een functie daarvan bepalend is in het verdere onderzoek. Om deze reden zou beter van een parameter gesproken kunnen worden en dient het onderzoek te worden gericht op de functie die de bestudeerde parameter transformeert in de toe te passen grootheid.

Een onderzoek naar de betrouwbaarheid van de uitkomsten met de gekozen parameter verkregen kan aan het licht brengen hoe vaak, onder identieke omstandigheden, eenzelfde resultaat verwacht kan worden. Is dit aantal hoog, dan pleit dit voor de methode en de werkwijze, waarmee de parameter werd bepaald, terwijl een criterium ontstaat volgens hetwelk alternatieve werkwijzen tegen elkaar kunnen worden afgewogen. Een hoge betrouwbaarheid behoeft overigens nog niet een goede toepasbaarheid in te houden, slechts de mate van reproduceerbaarheid wordt met een objectieve maatstaf vastgelegd.

De boven weergegeven gedachtengang is nader uitgewerkt en bood mogelijkheden om los van het speciale geval een beschouwing te wijden aan enkele aspecten van onderzoek-technische aard.

De gegeven formuleringen zijn zonder nadere specificatie niet steeds numeriek te evalueren. De volgende beschouwingen hebben echter voornamelijk het doel na te gaan tot hoever een mathematische formulering kan dienen voor het vastleggen van de vorderingen bij een onderzoek bereikt en het mogelijk te maken deze onderling tegen elkaar af te wegen.

De toepasbaarheid van deelonderzoekingen

De uitkomsten van deelonderzoekingen welke in het samenvattende onderzoek moeten worden ingebracht zullen toepasbaar moeten zijn in deze zin dat de uitkomsten dienst moeten doen als relevante factoren (parameters) in het samenvattende onderzoek. De toepasbaarheid heeft dus een fysische betekenis.

Hoewel de volgende beschouwingen meer algemeen geldig zijn zal een voorbeeld gehanteerd worden om de ontwikkelde gedachten een practische interpretatie te geven. Gebruik zal worden gemaakt van de bekende vergelijking van de waterbalans.

$$A = N + K - V - O - B \qquad (1)$$

waarin A = afvoer van een bepaald gebied over een tijdsinterval $t_{i+1} - t_i$
N = neerslag over hetzelfde gebied en hetzelfde tijdsinterval
K = kwel, idem
V = verdamping, idem
O = overige onttrekkingen, idem
B = verschil in geborgen hoeveelheid in het gegeven gebied over hetzelfde tijdsinterval.

Alle grootheden N tot en met O kunnen door hun aard slechts waarden aannemen ≥ 0, terwijl B slechts positief is als er in het genoemde tijdsinterval water in het gebied is achtergebleven.

Dat een dergelijke balans moet bestaan is uit elementaire fysische overwegingen onmiddellijk duidelijk. De moeilijkheid is echter om bij elke hydrologische- en meteorologische situatie, aangeduid met S, van elk van de bovenvermelde grootheden de bijbehorende waarde vast te stellen. Door substitutie van gemeten waarden in (1) wordt vervolgens de berekende afvoer vastgesteld die met de gemeten afvoer kan worden vergeleken.

Het is noodzakelijk dat relevante parameters in deze betrekking worden ingebracht, dat wil zeggen dat de algemene form zal luiden

$$A = p(N) + p(K) - p(V) - p(O) - p(B) \qquad (2)$$

waarin p het symbool van een functie voorstelt en bedoelt aan te geven dat er een voorschrift bestaat waarmee uit metingen of onderzoek van de variabelen in het argument de relevante waarden voor de gevraagde grootheden worden verkregen.

De functies kunnen van geheel verschillende structuur zijn. Zo is het voor de neerslag voor de hand liggend als parameter een functie van de gemeten neerslag zelf te nemen, zodat

$$p(N) = f_n(N) \tag{3}$$

waarvoor dan veelal de identiteit

$$f_n(N) = N \tag{4}$$

genomen wordt,met N de in een regenmeter opgevangen hoeveelheid neerslag. De problemen die zich nu gaan voordoen houden verband met de vraag in hoeverre met $f_n(N)$ het bedrag gegeven wordt dat in de waterbalans (1) moet worden ingebracht (DE ZEEUW, 1963). De vraag kan gesteld worden of bijvoorbeeld een vertraging in rekening gebracht moet worden (STOL, 1962), of een middeling over grote oppervlakten moet worden toegepast dan wel met isohyeten een integratie over het stroomgebied moet plaatsvinden (STOL, 1967).

Dergelijke en andere overwegingen gelden ook voor de andere functies. Zo zal voor de kwel

$$p(K) = f_k(\text{hydrologie, geologie}) \tag{5}$$

een meer gecompliceerde formulering gelden. Met (5) wordt dus bedoeld dat de in de waterbalans benodigde waarden voor de kwel p(K) verkregen worden uit een onderzoek naar de hydrologie en de geologie van een gebied (DE RIDDER en WIT, 1967). Hiermee is aangegeven dat een maat p(K) voor de kwel verkregen wordt uit een onderzoek naar het geologisch profiel, de waterstaatkundige en de topografische indeling van een gebied. De functie f_k is zeker geen eenvoudige. Naast fysische formuleringen bevat ze integraties over oppervlakten en profielen die in sommige gevallen met behulp van kaarten wordt uitgevoerd.

Meer analoog aan het eerder beschreven neerslagonderzoek is de directe meting van de kwel (VAN DER WEERD 1966). In dit geval gaat (5) over in

$$p(K) = g_k(K)$$

zodat als relevante parameter voor de kwel een functie g_k van de rechtstreeks gemeten kwel kan worden gebruikt. De variaties van de metingen naar tijd en plaats maken het in dit geval duidelijk dat de identiteit

$$g_k(K) = K$$

geen relevante parameter oplevert aangezien geen enkel meetpunt het gehele gebied naar oppervlakte en tijd volledig representeert.

Van geheel andere structuur is vervolgens de parameter voor de verdamping p(V). Rechtstreekse metingen van de verdamping door middel van evaporimeters of pannen leveren geen waarden die eenvoudig ingebracht kunnen worden in een functie die de parameters voor de verdamping levert. De fysiche benadering (RIJTEMA, 1965), kan symbolisch worden weergegeven met bijvoorbeeld

$$p(V) = f_v(R,W,E_o,B,G) \qquad (6)$$

waarin R = straling
W = windsnelheid
E_o = verdamping van vrij wateroppervlak
B = bodemfysische factoren
G = gewasfactoren

zodat een met (5) vergelijkbare situatie ontstaat. Uiteindelijk zal over oppervlakte, tijd, gewas en bouwplan geïntegreerd moeten worden zodat voor het bepalen van f_v tevens een bedrijfsanalyse moet worden uitgevoerd.

Het bovenstaande mag als voorbeeld dienen voor overeenkomstige beschouwingen voor de posten 'overige onttrekkingen' (drinkwater) en de berging.

Verschillende parameters bevatten impliciet identieke elementen, terwijl voor het vaststellen van de waterbalans steeds integratie over tijd en stroomgebied moet plaatsvinden.

De onzekerheden in de vaststelling van de grenzen van een stroomgebied doen zich in elk van de parameters gevoelen. Met een poging tot rubricering

zou, met toenemende gevoeligheid voor de juiste topografische ligging van de grenzen van het stroomgebied, kunnen gelden:

$$f_v < f_n < f_b < f_k < f_o$$

Daarnaast komt correlatie in verschillende vorm tussen de parameters voor. Zo hangt de verdamping onder andere af van meteorologische factoren evenals het geval is voor de neerslag; de berging wordt mede bepaald door de neerslagintensiteit, de verdamping en overige onttrekkingen. Het is niet ondenkbeeldig dat deze correlatie nog verder gaat, doordat bijvoorbeeld de berging bepaald wordt volgens

$$p(B) = f_b(p(N), p(V), \text{hydrologie}, \text{geologie}) \qquad (7)$$

wat weer van belang is bij overwegingen aangaande de voortplanting van de fouten inhaerent aan het meten en bepalen van de benodigde parameters.

Tenslotte verdient de uiteindelijke uitkomst, de afvoer A, nog enige aandacht. Bij een rechtstreekse meting wordt de afvoer als intensiteit maar geïntegreerd over het stroomgebied, vastgesteld. De grenzen van het stroomgebied zijn exact in rekening gebracht doch niet vastgesteld kan worden waar deze liggen ten tijde van de meting en welke variaties hier nog een rol kunnen spelen (DE JAGER, 1965; DE ZEEUW, 1966).

In de betrekking (1) wordt deze regulatie van stroomgebied-grootte eveneens automatisch aanwezig geacht aangezien met deze vergelijking de werkelijke situatie is weergegeven; het ideale geval dat slechts als denkmodel bestaat.

De werkelijkheid wordt voorgesteld door (2) waarin de parameters voor de vereiste grootheden voorkomen. De uitkomst levert een waarde voor de afvoer die een maat zal zijn voor de werkelijke afvoer, zodat in het algemeen slechts een parameter voor de afvoer gevonden wordt. De feitelijke situatie is nu dat (2) geschreven moet worden als een som van relevante parameters die als uitkomst een maat voor de werkelijke afvoer oplevert volgens:

$$p(N) + p(K) - p(V) - p(O) - p(B) \rightarrow p(A) \qquad (8)$$

waarbij tenslotte het samenvattende onderzoek zal bestaan uit het bestuderen

van de relatie

$$A = f_a\{p(A)\} \tag{9}$$

Het is hiermede duidelijk dat in het samenvattende onderzoek zowel de onzekerheden van de deelonderzoekingen de relatie vertroebelen - zich cumulatief uitend in p(A) - als de onzekerheden van het gehele systeem van werken - zich uitend in de functie f_a. Aangezien in (9) substitutie moet plaatsvinden van (8) en daarin weer vergelijkingen van het type (7), (6) enz., zal (9) uiteindelijk resulteren in (10):

$$A = F_a(N,R,W,E_o, \ldots, \text{geologie, hydrologie}, \ldots, \ldots) \tag{10}$$

waarin geen splitsing meer gemaakt kan worden in de mogelijke bronnen die de onzekerheden in uitkomst veroorzaken, zodat er zonder meer geen voor de hand liggende methode is aan te geven om het relevant zijn van de parameters vast te stellen. Hiertoe zal dan gebruik gemaakt moeten worden van fysisch triviale uitkomsten zoals langdurige droge perioden waarin $N \equiv 0$ ongeacht de vorm van f_n; perioden met nagenoeg gelijke waterstand aan begin en eind waardoor $B = 0$, wat al weer moeilijker te constateren valt daar waterstanden slechts steekproefsgewijs gemeten kunnen worden, enz.

De invloed van onzekerheden in deelonderzoekingen

Stel dat over een zeker tijdsverloop de hydrologische en meteorologische situaties dusdanig geweest zijn dat van een situatie S_o sprake geweest is. Dit houdt in dat over dit tijdsverloop de werkelijke waterbalans als volgt kan worden weergegeven (zie (1)):

$$A_o = N_o + K_o - V_o - O_o - B_o \tag{11}$$

De over hetzelfde tijdsverloop vastgestelde parameterwaarden leveren dan de betrekking

$$p(N)_o + p(K)_o - p(V)_o - p(O)_o - p(B)_o \rightarrow p(A)_o \tag{12}$$

Stel vervolgens dat over een ander (i) evenlang tijdverloop de bovenomschreven situatie S_o zich herhaalt[1]. Het is duidelijk dat ook (11) zich herhaalt als zijnde een representatie van de werkelijkheid. Evenzeer is het duidelijk dat (12) niet weer opnieuw dezelfde waarden oplevert, doch een die hiervan afwijkt, bijvoorbeeld $p(A)'_o$. De deelonderzoekingen hebben andere, wat afwijkende, resultaten opgeleverd, en de uitkomst is dan ook in het algemeen

$$p(A)_i = A_o + \Delta(A)_i, \quad i = 1,2, \ldots \tag{13}$$

wat wil zeggen dat de gevonden waarde uit het samenvattende onderzoek een bedrag $\Delta(A)_i$ van de werkelijke waarde A_o zal afwijken op de i-de keer dat de hydrologische situatie S_o optreedt.

Voor de overige parameters kunnen analoge betrekkingen opgesteld worden. Een verdere beschouwing kan nu plaatsvinden door aan de afwijkingen Δ bepaalde voorwaarden op te leggen. Voor een deel <u>moeten</u> aan de Δ's voorwaarden opgelegd worden opdat het onderzoek in de gewenste resultaten zal uitmonden, anders gezegd: het onderzoek moet op systematische wijze plaatsvinden waardoor de Δ's aan zekere wetmatigheden gaan voldoen.

Zowel grote als kleine waarden van Δ zullen optreden, al zal niet bekend zijn wanneer deze zich zullen voordoen. De afwijkingen Δ krijgen hiermede een stochastisch karakter. Geëist kan nu worden dat bij onderzoek van vele situaties S_o grote waarden van Δ slechts een kleine fractie mogen uitmaken, en dus met een kleine frequentie van voorkomen mogen optreden, wat - onder geschikte voorwaarden geëxtrapoleerd - inhoudt dat in volgende situaties S_o grote waarden Δ met een kleine kans op voorkomen zullen optreden.

Voor het in rekening brengen van nieuw verworven inzichten in het onderwerp van studie kan de volgende gedachtengang gevolgd worden. Stel dat de verantwoording van de post 'neerslag' op de waterbalans het gevolg is van een inbreng $p_k(N)$ waarbij de index k aanduidt dat de inbreng afkomstig is van een neerslagonderzoek dat zich in het stadium k namelijk het huidige

---

1) Met de index i worden de achtereenvolgens optredende situaties S_o aangeduid, waarbij deze niet chronologisch behoeven aaneen te sluiten. Afgezien wordt verder van de vraag of gelijkheid van S_o wel ooit geconstateerd kan worden zo het zich al voordoet

kennisniveau bevindt. We stellen ons hierbij voor dat naarmate verdere onderzoekingen worden uitgevoerd (k + 1, k + 2, ...) het inzicht in het gestelde probleem niet afneemt zodat in de reeks p_k, p_{k+1}, p_{k+2}, ... de parameterwaarden steeds beter de relevante factoren gaan representeren. Vanzelfsprekend behoeven k, k + 1, k + 2, ... niet aequidistant in de tijd te zijn.

Beschouwen we thans het verschil, analoog aan (13), voor de neerslagterm dan ontstaat (14) met

$$N_o - p_k(N)_i = \Delta_k(N)_i, \; k = 1, 2, \ldots, n \qquad (14)$$

Er kunnen nu twee principieel verschillende beschouwingen gevolgd worden.

De eerste is de <u>deterministische</u>. Hierbij treedt het eenvoudigste model op want gesteld wordt, dat voor $n \to \infty$ het inzicht in het gestelde probleem - - het facet 'waterbalans' voor de neerslag - toeneemt en tot steeds vollediger kennis voert. Dit houdt in dat in (14)

$$\lim_{k \to \infty} |\Delta_k(N)_i| \to 0 \qquad (15)$$

voor elke i, dus voor elke keer dat de hydrologische- en meteorologische situatie S_o optreedt. Na elk onderzoek wordt het totale proces beter beschreven en begrepen en de absolute waarden van de afwijkingen zullen na elk onderzoek kleiner worden. Doch niet alleen voor de situatie die symbolisch met S_o is aangegeven geldt deze betrekking, naarmate k toeneemt zal (15) ook voor andere situaties S_j geldigheid bezitten. Dit model is echter naar de huidige opvattingen te simplistisch en kan worden vervangen door één dat dichter bij de werkelijkheid aansluit.

De tweede beschouwing is de <u>indeterministische</u>. Weer wordt gesteld dat voor $n \to \infty$ het inzicht toeneemt en tot steeds vollediger kennis voert. Beschouw weer (14). De ervaring leert dat, hoewel het inzicht is toegenomen en een betere beschrijving van het proces is verkregen, in het stadium (k-1) toch gevallen voorkomen met een kleinere afwijking Δ dan later wordt gevonden, zodat - omgekeerd - ook in de stadia na k grote afwijkingen niet tot de onmogelijkheden behoren.

Wordt een, naar de huidige situatie van kennis, kriterium d_k gesteld waarboven de afwijkingen Δ_k liefst niet uit mogen komen dan zal er toch een kleine kans w_k bestaan dat dit bij een van de situaties S_o gebeurt. In formule geldt voor deze overschrijdingskans, uitgeschreven voor de neerslagterm:

$$P\{|\Delta_k(N)_i| \geqq d_k\} \stackrel{\text{def}}{=} w_k, \; 0 < d_k < \infty \tag{15}$$

In woorden: de kans P dat bij een kennisniveau k een van de afwijkingen van de werkelijke situatie een positief bedrag d_k overtreffen is gelijk aan w_k.

Een toename van het kennisniveau wordt nu tot uitdrukking gebracht met

$$P\{|\Delta_{k+1}(N)_i| \geqq d_k\} = w_{k+1} \quad \text{met} \quad w_{k+1} < w_k \tag{16}$$

en tenslotte kan de eis gesteld worden dat de overschrijdingskans voldoen moet aan

$$P\{|\Delta_j(N)_i| \geqq d_k\} \leqq \alpha_k \qquad \begin{array}{l} d_k > 0 \\ 0 \leqq \alpha_k < w_k \\ j > k \end{array} \tag{17}$$

(KENDALL and STUART, 1961 Pt II:3; KENNY and KEEPING, 1959:85) wat betekent dat grote afwijkingen niet tot de onmogelijkheden behoren, maar met toenemend inzicht minder frequent dan met een vooraf gestelde kleine overschrijdingskans α_k zullen optreden.

Bij steeds verder voortschrijdend kennisniveau, symbolisch aan te duiden met $j \to \infty$ kan nog gesteld worden dat

$$\lim_{j \to \infty} P\{|\Delta_j(N)_i| \geqq d_k\} = 0, \; d_k > 0$$

De parameter voor de neerslag krijgt hiermede het predicaat <u>bruikbaar</u> aangezien het bij voldoende inzicht in het gestelde probleem op den duur met kans P = 0 onmogelijk is dat afwijkingen groter dan d_k zullen optreden.

Een formulering voor het afwegen van bereikte resultaten

Het resultaat van de vorige beschouwing is dat duidelijk tot uiting komt dat twee kriteria gesteld dienen te worden:

Bij het huidige kennisniveau k wordt geëist dat geen grotere overschrijdingen van de werkelijkheid dan die ter grootte d_k mogen voorkomen met dien verstande dat overschrijdingen die niet frequenter dan met een kleine kans α_k zullen voorkomen, worden toegestaan. Volgens (16) kan met toenemende k (nieuwe ontwikkelingen in het onderzoek, verfijnde meetmethoden e.d.) op den duur aan de eis worden voldaan en men zal dan gaan overwegen d_k lager te stellen, dat wil zeggen met eenzelfde kleine kans α_k mogen nog slechts afwijkingen $d_{k+m} < d_k$ voorkomen, of men kan de afwijking zelf voldoende klein achten doch eisen dat de overschrijdingskans ervan kleiner wordt zodat voor meer situaties S_o aan het kriterium d_k voldaan wordt. De noodzaak tot het lager stellen van deze waarden kan berusten op ontwikkelingen op verschillend gebied waarvoor het huidige kennisniveau niet meer toereikend is. Het hoger stellen van de minimum eis levert de impuls om nieuw onderzoek op te zetten. De eisen aan d_k en α_k te stellen zijn niet onafhankelijk.

Zo zal volgens de ongelijkheid van TCHEBYCHEFF (KENNY and KEEPING, 1959:85) voor elke verdeling met eindige variantie gelden:

$$P\{|\Delta_k| \geq C\} \leq \frac{\sigma^2_{\Delta_k}}{C^2}$$

(STAM, 1964:283) waardoor nu dus met de definitie voor w_k in (15) verkregen wordt:

$$d_k = C$$

en

$$w_k \leq \frac{\sigma^2_{\Delta_k}}{C^2}$$

waaruit volgt dat steeds voldaan kan zijn aan

$$\alpha_k < w_k \leq \frac{\sigma^2_{\Delta_k}}{d_k^2}$$

Hoeveel w_k kleiner zal zijn hangt van de verdeling af die hier onbekend is, maar tot uiting is gebracht dat w_k en d_k niet onafhankelijk zijn.

Bij toenemend kennisniveau zal $\sigma^2_{\Delta_j}$ afnemen en bij gevolg op den duur ook w_j wanneer d_k constant gehouden wordt, zodat vanaf zekere n geldt

$$\alpha_k \geq \frac{\sigma^2_{\Delta_{k+n}}}{d_k^2} \geq w_{k+n} > \alpha_{k+n}$$

vanaf welk moment de eis α_k respectievelijk d_k^2 kleiner gesteld kan worden bijvoorbeeld α_{k+n} enz.

In de regel worden alleen overwegingen aangaande de afwijkingen d_k gegeven aangezien over de overschrijdingskans w_k en de hieruit voortvloeiende eis α_k in de regel weinig bekend is en het vaststellen ervan een type onderzoek vereist dat op het kansbegrip gebaseerd is terwijl d_k in de oorspronkelijke fysische eenheden kan worden uitgedrukt. In het eerste geval dient de kansverdeling van de Δ's bepaald te worden. Deze kan nog van velerlei aard zijn, alleen de eis (16) werd gesteld.

Verder zou nog nagegaan dienen te worden volgens welke overwegingen het kriterium d_k op den duur lager gesteld wordt. Veelal dienen hier economische motieven aan ten grondslag te liggen terwijl het verder geenszins uitgesloten is te achten dat dit de vertaling inhoudt van de eis dat w_{k+m} een lagere waarde moet aannemen.zonder over de kansen zelf een uitspraak te hoeven doen.

Vooruitgang verkregen bij het onderzoek kan nu uitgedrukt worden in de ongelijkheden

$$w_k > w_{k+1} > w_{k+2} > \ldots, \; d_k \text{ constant}$$

Men kan nog aannemen dat elke volgende w een fractie van de voorgaande

is, zodat

$$w_{k+1} = \gamma w_k, \; 0 < \gamma < 1 \tag{18}$$

De oplossing van deze differentie-vergelijking luidt (LEVY and LESSMAN, 1959):

$$w_k = w_o \gamma^k = w_o e^{k \ln \gamma}, \; 0 < \gamma < 1$$

waarin w_o een constante (beginniveau) is en waarin $\gamma < 1$ en bijgevolg $\ln \gamma < 0$, zodat

$$w_k = w_o e^{-\beta k}, \; -\beta = \ln \gamma \quad \text{en} \quad \beta > 0 \tag{18a}$$

waardoor een exponentiële afname ontstaat. Deze opzet vindt zijn motivering in de aanname dat een onderzoek niet eerder afgesloten zal worden dan nadat een bepaald resultaat bereikt is. Dit resultaat wordt hier dus gedefinieerd als een verdere procentuele afname van de overschrijdingskans w, waarmee dan het volgend kennisniveau (k+1) bereikt is.

Om nu te voldoen aan de eis α_k moet na zekere tijd (van nieuwe onderzoekingen) de overschrijdingskans w tot beneden deze waarde gedaald zijn, zodat

$$w_k = w_o \gamma^k > \alpha_k > w_o \gamma^{k+m} = w_{k+m}$$

in woorden: huidig niveau > eis > toekomstig niveau.

Kan de overschrijdingskans slechts afgeschat worden met behulp van de ongelijkheid van TCHEBYCHEFF, dan wordt dit

$$\frac{\sigma^2_{\Delta_k}}{d_k^2} \geqq P\{|\Delta_k| \geqq d_k\} \overset{\text{def}}{=} w_k > \alpha_k \geqq w_{k+m} = \gamma^m w_k \geqq \frac{\sigma^2_{\Delta_{k+n}}}{d_k^2} \geqq$$

$$\geqq P\{|\Delta_{k+n}| \geqq d_k\} \overset{\text{def}}{=} w_{k+n} = \gamma^n w_k$$

waarin nu $n \geq m$, en de gelijk tekens rechts van w_{k+m} gelden als $n = m$. Deze uitdrukking geldt dus als wel steeds de mogelijkheid bestaat om σ_Δ^2 te berekenen, maar de verdeling van Δ niet bekend is.

Het praktische gevolg van deze onbekendheid is dat meer en intensiever onderzoek zal moeten worden verricht om aan de eis van de kleine overschrijdingskans α_k te voldoen.

Om tot een juist afwegen van de noodzaak tot verder onderzoek te komen zou het de moeite kunnen lonen vanaf een bepaald moment meer aandacht te gaan besteden aan de werkelijke aard van de verdeling van de sluitfouten bij het huidige kennisniveau dan met een globale schatting hiervan te trachten door verder onderzoek verfijningen in reeds bereikte resultaten aan te brengen.

Het samenspel van eisen aan nauwkeurigheid

Vergelijking (15) kan, voor de neerslag N, nog als volgt als cumulatieve kans geschreven worden:

$$P\{|\Delta_k(N)_i| < d_k^{(N)}\} = (1 - w_k^{(N)})$$

zodat analoog voor de andere posten op de waterbalans verkregen wordt onder andere:

$$P\{|\Delta_k(K)_i| < d_k^{(K)}\} = (1 - w_k^{(K)})$$

waarmede ook de andere cumulatieve kansen gedefinieerd zijn als

$$(1 - w_k^{(V)}),\ (1 - w_k^{(O)}) \quad \text{en} \quad (1 - w_k^{(B)})$$

Moet aan alle voorwaarden tegelijk voldaan worden namelijk dat alle absolute waarden van afwijkingen kleiner zijn dan de criteria, dan is de kans $P_<$ dat deze gebeurtenis optreedt gelijk aan het product van de kansen:

$$P_< = \prod_{x=N}^{B}(1 - w_k^{(x)}),\ x = N, K, V, O, B$$

Voor kleine waarden van $w^{(x)}$ zou hiervoor nog geschreven kunnen worden

$$P_< = 1 - (w_k^{(N)} + w_k^{(K)} + w_k^{(V)} + w_k^{(O)} + w_k^{(B)})$$

en de kans $P_>$ op een overschrijding van enig van de kriteria is nu dus

$$P_> = w_k^{(N)} + w_k^{(K)} + w_k^{(V)} + w_k^{(O)} + w_k^{(B)}$$

Wil men nu dus bereiken dat in de toekomst, aangeduid als (k+1), de waarde van $P_>$ in ieder geval kleiner is dan een vooraf aangegeven waarde P_{k+1} dan luidt de eis

$$w_{k+1}^{(N)} + w_{k+1}^{(K)} + w_{k+1}^{(V)} + w_{k+1}^{(O)} + w_{k+1}^{(B)} < P_{k+1} \qquad (19)$$

Aangezien alle $w^{(x)} > 0$ volgt hieruit dat in ieder geval voldaan moet zijn aan

$$w_{k+1}^{(x)} < P_{k+1}, \text{ (alle x)} \qquad (20)$$

wat betekent dat indien als eis gesteld wordt dat de overschrijdingskans van alle afwijkingen van het type d_k niet groter dan P_{k+1} mag zijn een inzicht zal moeten zijn verkregen in de grootte van de overschrijdingskansen w_{k+1} opdat uitgemaakt kan worden welk van de grootheden in de deelonderzoekingen het eerst voor verder en intensiever onderzoek in aanmerking komt. Overigens volgt uit (19) nog dat indien (20) metterdaad geldt, de grootheden niet los van elkaar gezien behoeven te worden en het voldoende is van een aantal grootheden de overschrijdingskans w enigszins te verlagen om aan (19) te voldoen, wat met minder intensief onderzoek gepaard zal gaan, zodat op meer economische wijze eenzelfde resultaat verkregen wordt.

Een verdere beslissing aangaande de grootheden die onderzocht moeten worden kan plaatsvinden op grond van vergelijking (18a) die, in combinatie met (19) levert:

$$\sum_{x=N}^{B} w_o^{(x)} e^{-\beta_x^{(k+1)}} < P_{k+1} \qquad (21)$$

Het meest efficiënt is het nu die grootheid het eerst in het onderzoek te betrekken waarvan β_x groot is en die dus een snelle afname van de overschrijdingskans garandeert.

Het is duidelijk dat de evaluatie van een en ander op moeilijkheden zal stuiten. Een gelukkige greep bij het onderzoek veroorzaakt een β die groot is. Dit garandeert niet dat een volgende sprong met evenveel gemak genomen kan worden. Met de eerste gelukkige greep kunnen alle verdere mogelijkheden reeds uitgeput zijn. Verder zal het zo zijn dat de parameter β niet zo zeer gebonden is aan het onderwerp van studie, dan wel aan de onderzoeker, of het team van onderzoekers, mogelijk aan een combinatie van onderzoeker, onderwerp en faciliteiten die bij het onderzoek ondervonden worden.

- o - o -

Van veel belang is echter het verkrijgen van inzicht in de vraag of men een onderzoek afgesloten acht als een bepaalde, overigens niet bewust geëvalueerde, winst aan nauwkeurigheid is bereikt, zodat in (18) γ een constante is, dan wel of men een onderzoek na een optimaal aantal jaren afsluit waardoor voor opeenvolgende kennisniveaus k, k+1, ..., zou gelden dat γ afhankelijk is van k.

Vermeld wordt nog dat in dat geval de oplossing van (18) wordt

$$w_k = \gamma_o \gamma_1 \gamma_2 \cdots \gamma_{k-1} w_o$$

en het aantal verstreken jaren mogelijk optimaal geacht wordt wanneer γ_k, hoewel nog niet gelijk aan γ_{k-1} geworden, deze toch op een bepaald percentage is benaderd. Hiermee wordt aangegeven dat het steeds moeilijker wordt een bepaalde afname van w_k te realiseren, wanneer het kennisniveau voortschrijdt. Uiteindelijk wordt hiermede een ander model verkregen en wel

$$w_k = w_o \gamma_o^k g^{\sum_1^{k-1} n} \quad , \quad 0 < \gamma_o < 1$$
$$0 < g < 1$$

wat weer in de volgende vorm gebracht kan worden die aangeeft hoe w_k afneemt,

$$w_k = w_o \exp(- \beta_o k - g_o \sum_1^{k-1} n), \; - g_o = \ln g$$

Is men bij zijn onderzoek aan tijd gebonden dan hangt de afname van w bij opeenvolgende kennisniveaus dus mede af van g_o de constante die de procentuele afname van γ vertegenwoordigt. De eerste oplossing (18a) bestond hierin dat $g = 1$ of $g_o = 0$, de laatste wordt dus verondersteld aequidistante punten in de tijd te bezitten, dat wil zeggen dat na elke stap de kennisniveaus langzamer toenemen doch regelmatiger in de tijd.

- o - o -

De gevolgde gedachtengang mondt nu uit in hanteerbare termen, namelijk dat bij een samengesteld onderzoek

1) de meest onnauwkeurige onderdelen die afzonderlijk nog niet aan de gestelde eis voldoen het eerst onderzocht moeten worden;
2) het onderzoek te laten uitvoeren door die medewerkers of groep van medewerkers waarvan een grote efficiëncy β verwacht wordt.

Voor zover deze opmerkingen niet nieuw zijn moet wel bedacht worden dat hier een basis van een exacte wijze van afwegen van de noodzaak tot studie is gegeven en dat een formulering is verkregen waarin de beslissingskriteria elk hun plaats hebben.

De voortplanting van fouten

Behalve de eis dat voor de afwijkingen Δ de kans steeds kleiner zal zijn dat een minimum afwijking d_k overschreden wordt is de eis gerechtvaardigd dat bij elk kennisniveau k de grote en de kleine, positieve en negatieve afwijkingen tegen elkaar opwegen zodat, gemiddeld over velerlei situaties i de

met de kans van voorkomen gewogen gemiddelde Δ gelijk is aan 0. Dit betekent dat de verwachtingswaarde gelijk is aan

$$E(\Delta) = 0$$

Wordt ook hieraan voldaan, dan heet de werkwijze zuiver in die zin dat zich geen systematische afwijkingen zullen voordoen. Namelijk met (14) geldt voor de verwachting:

$$E\{p_k(N)\} = E\{N_o - \Delta_k(N)\} = N_o, \text{ alle } k$$

De verwachtingswaarde van de gebruikte parameter is de werkelijke te gebruiken neerslag.

Met deze eis mondt het resultaat uit in de goede waarde voor de afvoer, namelijk nu zonder indices:

$$E\{p(A)\} = E\{p(N) + p(K) - p(V) - p(O) - p(B)\} = N_o + K_o - V_o - O_o - B_o = A_o$$

Hiermede is echter nog niets gezegd over de fout van de uitkomsten. Hoewel het gemiddelde de gevraagde uitkomst levert, is voor elk geval afzonderlijk een fout aanwezig die nog groot kan zijn. Bovendien volgt de fout in het resultaat uit de voortplanting van de fouten uit de deeluitkomsten.

Wanneer een systematische afwijking waarvoor $E(\Delta x) = \varepsilon_x$, optreedt ontstaat de volgende situatie:

$$\begin{aligned} E\{p(A)\} &= E\{p(N) + p(K) - p(V) - p(O) - p(B)\} = \\ &= N_o + \ldots - B_o + E(\Delta N + \Delta K - \Delta V - \Delta O - \Delta B) = \\ &= A_o + \varepsilon_N + \varepsilon_K - \varepsilon_V - \varepsilon_O - \varepsilon_B \end{aligned}$$

wat geconstateerd wordt als

$$E\{p(A)\} = A_o + \varepsilon_A \qquad (22)$$

zodat

$$\varepsilon_A = \varepsilon_N + \varepsilon_K - \varepsilon_V - \varepsilon_O - \varepsilon_B \tag{23}$$

wat een enkele vergelijking oplevert in 5 onbekenden. Deze zijn dus niet te scheiden.

Pogingen hiertoe berusten op de volgende onderzoekstechniek. Uitgezet worden de berekende tegen de gemeten afvoer A zodat verkregen wordt uit (22), algemeen:

$$p(A) = \alpha A_o + \Delta(A) \tag{23a}$$

waarin $\alpha = 1$ en $E\{\Delta(A)\} = 0$ de gewenste ideale oplossing levert. Een procentuele en constante afwijking kunnen hiermee geconstateerd worden wanneer de curve niet door de oorsprong gaat en op een lineaire samenhang duidt. De verwachting van het linkerlid levert dan (22) en bij gevolg geldt (23).

Een splitsing van de termen in (23) is slechts mogelijk door weer gebruik te maken van fysisch triviale uitkomsten zoals reeds eerder werd beschreven. Voor perioden met langdurige droogte bijvoorbeeld is de hoeveelheid neerslag exact gelijk aan 0 en dus $E\{\Delta(N)\} = \varepsilon_N = 0$. In de winter zal de verdamping gering zijn en zal bij benadering gelden $\varepsilon_V = 0$. Combinaties van deze situaties maken het mogelijk (23a) te bestuderen onder verschillende omstandigheden, met andere woorden (23) als één vergelijking met ≤ 5 onbekenden.

Door deze wijze van werken kan een steeds verder inzicht in het gedrag van (23) worden verkregen, en ontstaat de mogelijkheid systematische fouten te elimineren.

- o - o -

De afwijkingen Δ_i komen elk dus met een zekere kans voor zodat naast de verwachtingswaarde ook de variantie van de Δ's een indicatie omtrent hun gedrag geven. Deze variantie is gedefinieerd als

$$Var(\Delta) = E(\Delta - E(\Delta))^2 = E(\Delta)^2 = \sigma_\Delta^2$$

Dit betekent tevens dat bij kennisniveau k en bij identieke situaties j van S_o zal gelden, weer voor de neerslag als voorbeeld:

$$Var\{\Delta_k(N)\} = E\{N_o - p_k(N)\}^2 = E\{p_k(N)\}^2 = Var\{p_k(N)\}^2$$

zodat de variantie van de afwijkingen gelijk is aan de variantie van de parameterwaarden zelf bij gelijke situaties S_o.

Op overeenkomstige wijze zijn de covarianties gedefinieerd en er geldt

$$Cov\{\Delta_k(X), \Delta_k(Y)\} = E\{\Delta_k(X)\}\ \{\Delta_k(Y)\} = \rho_{\Delta x \Delta y} \cdot \sigma_{\Delta x} \cdot \sigma_{\Delta y} \qquad (24)$$

waarmee de covariantie uitgedrukt is in de (wortel van de) varianties en de correlatie ρ tussen de fouten of afwijkingen $\Delta(X)$ en $\Delta(Y)$.

Met behulp van deze formules kan nu nagegaan worden wat de voortplanting van fouten in het deelonderzoek voor een effect heeft op het samenvattende onderzoek.

In het voorgaande is weliswaar nog aangenomen dat de fout additief is ten opzichte van de inbreng, dit behoeft echter niet altijd het geval te zijn. De verdamping bijvoorbeeld wordt in de regel met formules bepaald. Hierin vindt eveneens een voortplanting van fouten plaats. Stel dat de functie van de verdamping die toegepast wordt, luidt: f(V) en dat een afwijking ΔV optreedt, dan geldt bij benadering

$$\frac{f(V_o + \Delta V)}{\Delta V} \simeq \left[\frac{df}{dV}\right]_o$$

dus bij benadering gelijk aan de afgeleide ter plaatse met index 0. Zodat

$$f(V_o + \Delta V) \simeq \left[\frac{df}{dV}\right] \Delta V$$

en de afwijking ΔV dus met een factor vermenigvuldigd wordt die zelf weer van V afhankelijk kan zijn. De verandering van de functie ter plaatse 0 is nu tevens van belang. Er geldt overigens

$$E\{\left[\frac{df}{dV}\right]_o \Delta V\} = \left[\frac{df}{dV}\right] E\{\Delta V\} = 0$$

zodat met toevallige afwijkingen met verwachting 0 de formulering (1) bruikbaar blijft.

Toepassing van de algemene foutenvoortplantingsregels op de waterbalansvergelijking

Toegepast op de waterbalansvergelijking is het duidelijk dat alle partiële afgeleiden de waarde + 1 of - 1 hebben zodat geschreven kan worden

$$\Delta A = \Delta N + \Delta K - \Delta V - \Delta O - \Delta B \qquad (25)$$

waarin de Δ's zoals uiteengezet, een stochastisch karakter hebben.

Wordt van het rechterlid de verwachtingswaarde van de fouten Δ genomen dan blijkt deze gelijk aan nul zodat ook

$$\Delta A = 0$$

en, bij afwezigheid van systematische fouten zal, over eindig veel situaties gemiddeld, de fout in de einduitkomst gering zijn.

Voor het berekenen van het kwadraat van de fout (variantie) in A wordt (25) gekwadrateerd waardoor er ontstaat:

$$(\Delta A)^2 = (\Delta N)^2 + (\Delta K)^2 + (\Delta V)^2 + (\Delta O)^2 + (\Delta B)^2 + 2(\Delta N)\ (\Delta K) + \ldots + 2(\Delta O)\ (\Delta B)$$

Gemiddeld over velerlei situaties wordt dit

$$E(\Delta A)^2 = E(\Delta N)^2 + \ldots + 2E(\Delta N)\ (\Delta K) + \ldots$$

en tenslotte, in verband met (24)

$$\sigma_A^2 = \sigma_N^2 + \ldots + 2\rho_{NK}\sigma_N\sigma_K - 2\rho_{NV}\sigma_N\sigma_V - \ldots \qquad (26)$$

De voorwaarde waaronder de som van de varianties in de deelresultaten gelijk is aan de variantie in het eindresultaat volgt onmiddellijk uit (26) namelijk wanneer alle fouten ongecorreleerd zijn dus $\rho_{NK} = \rho_{NV} = \ldots = 0$.

Wordt aangenomen dat de deelonderzoekingen elk op eigen wijze tot een resultaat komen en van andere reeksen basisgegevens gebruik maken, dan zal aan deze voorwaarde voldaan zijn en zal gelden

$$\sigma_A^2 = \sigma_N^2 + \sigma_K^2 + \sigma_V^2 + \sigma_O^2 + \sigma_B^2 \qquad (27)$$

Hoewel het veelal lastig zal zijn een goede kwantitatieve maat voor deze grootheden te vinden kan deze uitkomst toch dienen voor enkele nadere beschouwingen.

Er worden twee gevallen onderscheiden:

1. Voor de hand liggend lijkt het veelal te eisen dat σ_A^2 zo klein mogelijk gemaakt moet worden, dat wil zeggen men wil bij eenzelfde situatie S_o op dezelfde waarde voor A uitkomen. Uit (27) die uit een som van kwadraten bestaat volgt dat σ_A^2 zo klein mogelijk wordt als elk van de termen in het rechterlid zo klein mogelijk wordt. Dit houdt in dat elke opvoering van de nauwkeurigheid in de deel-resultaten steeds zal leiden tot opvoering van de nauwkeurigheid in het eindresultaat, en men kan steeds doorgaan met σ_A^2 zo klein mogelijk te maken. Het blijkt dus dat de eis niet praktisch is. Er wordt geen kwantitatieve maat gesteld, terwijl ook niet afgewogen kan worden welke onderdelen van het onderzoek meer aandacht verdienen te krijgen en in welke mate.
2. Een meer praktische voorwaarde wordt verkregen wanneer inzicht in het gehele probleem is verworven. Dan kan de eis gesteld worden dat σ_A^2 niet groter dan een bepaalde waarde mag zijn, bijvoorbeeld $\sigma_{A_{max}}^2$, maar ook niet lager behoeft te wezen.
   Uit de kansverdeling, waarvan σ_A^2 een parameter is, kan opgemaakt worden welk risico gelopen wordt door het stellen van deze eis wat op economische gronden moet worden gebaseerd. Het bewust kiezen van een maat voor de variantie betekent dat men zekere afwijkingen van het gemiddelde toestaat en overgaat op een interval-schatting. Dit past beter bij de realiteit daar nu geen niet realiseerbare nauwkeurigheid wordt nagestreefd.

Ook nu wordt weer gevonden dat die grootheden in ieder geval onderzocht dienen te worden waarvoor

$$\sigma_X^2 > \sigma_A^2, \; X = N, \ldots, B$$

Is het voor enige grootheid niet mogelijk te voldoen aan

$$\sigma_X^2 < \sigma_A^2$$

dan is de eis met betrekking tot de nauwkeurigheid waarmee de afvoer gevraagd wordt te streng. Ook hier geldt weer dat enige bekendheid met de variaties van de deelresultaten tot duidelijker inzicht voert in de eisen die aan het eindresultaat gesteld mogen en kunnen worden.

Het is hierbij goed denkbaar dat een economische beschouwing pleit voor het genoegen nemen met een minder nauwkeurige eindoplossing gezien het feit dat dit voor het gestelde doel voordeliger kan zijn dan het onderzoek te intensiveren, wat tijd vraagt en kosten met zich brengt.

-

Samenvatting

Uitgaande van de vergelijking voor de waterbalans werd een beschouwing gegeven die tot doel had een mathematische formulering toe te passen op gedachten omtrent het inpassen van deelonderzoekingen in een samenvattend onderzoek. Een formulering werd gegeven voor het voortschrijden van het onderzoek waarin verschillende van de hierbij van belang zijnde facetten in een model werden opgenomen.

Literatuur

JAGER, A.W. DE, 1965. Hoge afvoeren van enige nederlandse stroomgebieden. Wageningen.

KENDALL, M.G. and A. STUART, 1961. The advanced theory of statistics PE II, Griffin, London.

KENNY, J.F. and E.S. KEEPING, 1959. Mathematics of Statistics Pt II; Nostrand Company, New York.

LEVY, H. and F. LESSMAN, 1959. Finite difference equations,Pitman, London.

RIDDER, N.A. DE and K.E. WIT, 1967. Seepage flow analyses of a small polder in the south-western part of the Netherlands. Techn. Bulletin 49, I.C.W. Wageningen.

RIJTEMA, P.E. 1965. An analysis of actual evapotranspiration. Wageningen.

STAM, A.J. 1964. Inleiding tot de waarschijnlijkheidsrekening. Stam, Haarlem.

STOL, PH.TH. 1962. Een oriënterend onderzoek naar de invloed van de voorperiode op de betrekking tussen neerslag en afvoer. I.C.W.-NOTA 166.

_____, 1967. Verschillen in dagelijkse neerslaghoeveelheden tussen regenstations in de Gelderse Achterhoek. Een kwalitatieve vergelijking. NOTA 398, I.C.W.

WEERD, B. VAN DER, 1966. Apparatuur voor het meten van slootkwel. Med. I.C.W. Nr. 95.

ZEEUW, J.W. DE, 1963. Over de werkelijkheidsbenadering van gemeten neerslagen. Landbk.T. 75:815.

_____, 1966. Analyse van het afvoerverloop van gebieden met hoofdzakelijk grondwaterafvoer. Wageningen.