

3 okt. 1963

KONINKLIJK NEDERLANDS
METEOROLOGISCH INSTITUUT

Verslagen V-128
(R III-279-1963)

Kansbeschouwingen bij door regen belemmerde
openluchtwerkzaamheden.

.....

(tevens verslag van een colloquium,
op 28 mei 1963 gehouden).

door

Dr. C. Levert

551.577:
519.2

.....

INHOUD

1. Inleiding.
2. Het artikel van J.D. Mc Quigg en W.L. Decker.
 - 2.1 citaten
 - 2.2 formules
3. De gemiddelde wachttijd en de gemiddelde "klaardatum".
4. Numerieke voorbeelden.
5. Slotopmerking (doel van de studie).
6. Literatuur.

.-.-.-.

1. INLEIDING

In [1] wordt op bladzijde 37 een aktuele praktijkvraag geformuleerd, die het onderwerp van het colloquium vormde. De vraag, niet letterlijk geciteerd, luidt: "Er moet zeker werk verricht worden. Indien het niet door regen onderbroken wordt komt het klaar in W, zegge 7, dagen. Het werk is van dien aard, dat het bij h of meer, zegge 3 of meer, mm per etmaal onderbroken moet worden. Bij hervatting moet men (vrijwel) opnieuw beginnen. Hervatting kan pas geschieden nadat er, na de laatste regendag D, zegge 2, dagen zonder neerslag gepasseerd zijn. Men wil dit werk op dag T (target day; begindag B), zegge 10 juli, beginnen, d.w.z. men wil het klaar hebben op dag $B + W - 1$.

Gevraagd wordt de kans $F(E)$, dat dit werk uiterlijk op dag E, zegge 1 augustus, klaar zal zijn. De berekening levert bijv. $F = 0.60$. Wellicht vindt men het risico $R = 1 - F$, in casu $1 - 0.60 = 0.40$, om niet op of vóór 1 augustus klaar te zijn groot. De F kan worden vergroot (en R verkleind) o.a. door W te verkleinen (7 → 5 dagen). Zulks impliceert meer arbeiders, meer werktuigen, meer kosten. Misschien kunnen ook zodanige maatregelen genomen worden, dat het werk eerst bij regenhoeveelheden van 5 of meer mm per etmaal onderbroken moet worden. Ook deze maatregelen verhogen de kosten. Kortom, de vraag rijst: hoe hangt F, voor gegeven B en E, samen met h, D, W, en hoe met de keuze van B en E?

Dit probleem laat zich niet met de gegevens in de frequentieboeken [2] oplossen, daar hierin de jaarlijkse gang van een dag met tenminste h mm (0.1; 0.2; etc.) binnen de kalendermaand verwaarloosd wordt. Bovendien verschaffen de boeken slechts de (onvoorwaardelijke) "overall-kansen" p op dagen met h of meer mm, d.w.z. de kans, dat op een dag h of meer mm valt, niet lettende op wat er tijdens de vorige dagen viel; anders gezegd, de persistentie wordt niet beschouwd. Wil men derhalve én de jaarlijkse gang én de persistentie in rekening brengen (en dat is zeker in genoemd probleem gewenst), dan kunnen de frequentieboeken niet helpen, doch dienen de initiaalkaarten (de ponskaarten met de dagsommen) geraadpleegd te worden.

De oplossing van het bovenbeschreven probleem vonden wij in een artikel van J.D. Mc Quigg en W.L. Decker [3] . Aan dit artikel is hoofdstuk 2 gewijd. Hier en daar zijn kritische opmerkingen toegevoegd.

Aangezien de praktijk voor een eerste oriëntatie ook naar de gemiddelde "wachtijd" en de gemiddelde "klaardatum" vraagt en genoemde auteurs deze niet behandelen, leek het ons nuttig in hoofdstuk 3 daaraan enkele beschouwingen te wijden.

In hoofdstuk 4 lichten wij een en ander numeriek toe.

2. HET ARTIKEL van J.D. Mc Quigg en W.L. Decker.

2.1 Wij citeren vrij uit de inleiding van het Amerikaanse artikel:

"Vele landbouwkundige en commerciële werkzaamheden moeten bij regen onderbroken worden. Het komt voor, dat er gedurende lange tijdvakken weinig regen valt, maar er komen ook perioden voor met zo weinig dagen met "werkbaar weer", dat zeer veel materieel en arbeiders nodig zijn om het werk op tijd klaar te krijgen. Voor sommige activiteiten in de buitenlucht moeten werktuigen en arbeiders weken of maanden van te voren "besproken" worden." Verder:

"Bij de definitie van de "target date" T wordt verondersteld, dat er een goede reden is waarom de uitvoering niet kan beginnen vóór deze T-dag. Er kunnen contracten lopen (bouwwerken). In de landbouw kan de T door het weer bepaald worden."

Voorts:

"De "droogtijd" D kan uiteenlopen, al naar de hoeveelheid regen, die er in totaal vóór de "droogtijd" viel. De vraag hoe lang het duren zal alvorens de bodem voldoende droog is voor een hervatting van het werk en de vraag of deze duur redelijk constant is dan wel van het seizoen en de grondsoort afhangt, verdient verdere studies."

Op bladzijde 181 van het artikel: "Bij de definitie van de operatie-duur werd gesteld, dat de run van droge dagen ononderbroken is. Echter is het bij sommige werkzaamheden redelijk te stellen, dat de uitvoeringstijd gedeeltelijk onderbroken mag

worden door natte dagen, zodat wèl de totale operatie-duur een gegeven aantal dagen omvat. Inderdaad is dikwijls deze situatie aanwezig, doch het waarschijnlijkheidsmodel, op basis waarvan dán de berekeningen zouden moeten geschieden, zou mathematisch onhandelbaar worden. De eis, dat de droge en operationele periodes continue reeksen van droge dagen moeten zijn, stelt dus een limietgeval voor. Als de operationele tijd zou kunnen worden onderbroken en men, zodra de natte dagen voorbij zijn, het werk kan voortzetten, dan zou dit werk nooit later klaarkomen dan wanneer de droge en operationele tijden wèl continue runs van droge dagen zijn. Daarom zullen de kansschattingen, voortvloeiende uit de huidige mathematische analyse, nooit kleiner zijn dan de ware kansen om het werk klaar te krijgen op een gegeven dag." Tot zover onze citaten.

2.2 Formules

Definities:

droge of min-dag = etmaal met minder dan h mm

natte of plus-dag = etmaal met h of meer mm.

B = begindag = vroegste dag, waarop de uitvoering beginnen kan.

Worden de dagen in bijvoorbeeld maart t/m september met 1, 2, enz. aangeduid, dan stelt B = 77 de dag 16 mei voor.

E = klaardatum = einddag. E = 92 zou 31 mei voorstellen.

W = werktijd = aantal achtereenvolgende droge dagen, direct na D achtereenvolgende droge dagen, waarin het werk achterelkaar voltooid kan worden.

D = droogtijd = aantal achtereenvolgende droge dagen, nodig na de laatste natte dag, alvorens het werk opnieuw begonnen kan worden.

$$\beta = D + W$$

$f(i) \equiv f_i$ = kans het werk te voltooien op dag i.

$F(i) \equiv F_i$ = kans het werk te voltooien op dag i of eerder.

$p(i)$ = kans, dat dag i nat is.

$q(i, i+1, \dots, i+m)$ = kans, dat elk der $m+1$ achtereenvolgende dagen i, i+1, . . . , i+m droog is.

het werk niet klaar zijn, ook niet eerder (anders zou het zinloos zijn naar $f(B + W + x)$ te vragen). De kans, dat het werk niet af is op of voor dag $B - D + x - 1$, wordt geschreven als

(4) $1 - F(B - D + x - 1)$.

Mc Quigg en Decker schrijven $f(B + W + \underline{x})$ als het product der drie kansen (2), (3) en (4), q.e.d. *

Per definitie is $F(B + W + \underline{x}) = \sum_{i=-1}^{\underline{x}} f(B + W + i)$

Als nu nog $\underline{y} = W + \underline{x}$ wordt ingevoerd ($x \geq -1$; $y \geq W - 1$), dan laten zich drie situaties onderscheiden:

(5)
$$\begin{cases} y < W - 1, \text{ dan } f(B + y) = 0; F(B + y) = 0. \\ y = W - 1, \text{ dan } f(B + W - 1) = q(B - D, \dots B + W - 1). \\ y > W - 1, \text{ dan } f(B + y) = [1 - F(B - \beta - 1 + y)] p(B - \beta + y) [q(B - \beta + y + 1; \dots \\ \dots B + y)], \\ \text{met } F(B + y) = \sum_{i=W-1}^y f(B + i) \end{cases}$$

In bijgaande schets, fig. 2, is $B=1$, $E=15$, $D=2$, $W=3$

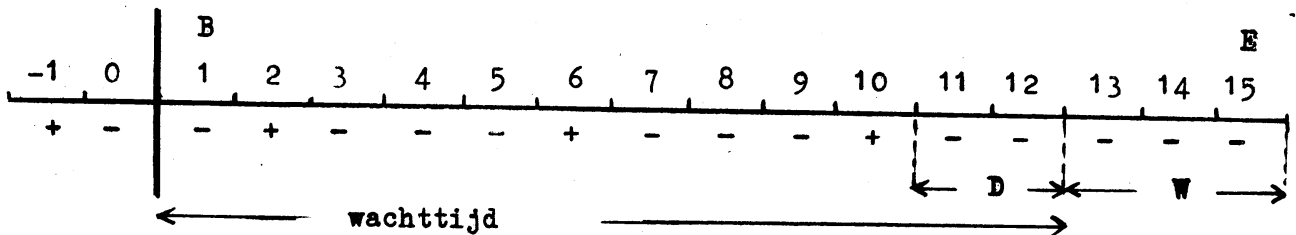


fig.2

Twee omstandigheden maken de berekening van de f 's en F 's via (5) zo langdradig, vooral bij grote W en D .

- a) er is een jaarlijkse gang in de kans op een plus-dag
- b) er is persistentie, d.w.z. de kans, dat een aangewezen dag een plus-dag zal zijn, hangt samen met wat de gepasseerde dagen waren.

* Is dit correct? Het is correct alleen als de drie "gebeurtenissen" onafhankelijk zijn. Zijn zij dat? Indien er persistentie is zal de q uit (2) afhangen van het "karakter" der vorige dagen (droog, nat). Misschien is deze afhankelijkheid van minder betekenis naarmate β en h groter zijn. Hoe dan ook: ze is er, zeker in Nederland, wellicht niet of zo goed als niet in het door de auteurs bestudeerde materiaal. Kortom: (1) geldt exact als er geen persistentie is.

De schrijvers achten b) afwezig. Omdat zij weten, dat vele andere onderzoekers het niet met hen eens zullen zijn, drukken zij zich voorzichtig uit: "For many years climatologists have contended that the relationship between days will cause a greater number of runs of dry days than expected from random theory. In this analysis, this characteristic is not apparent and the need for further study is indicated".

Als jaarlijkse gang en persistentie afwezig zijn, gaat (5) over in:

$$(6) \begin{cases} y < W - 1; f(B; y) \equiv 0; F(B + y) \equiv 0 \\ y = W - 1; f(B + W - 1) = q^\beta \\ y > W - 1; f(B + y) = [1 - F(B - \beta - 1 + y)] pq^\beta \end{cases}$$

De schrijvers delen mede, dat voor de berekeningen via (5) en (6) een elektronische rekenmachine gebruikt werd. Er werd een groot aantal grafieken getekend, bedoeld voor de praktijk, waarvan bijgaande fig.3 er een is.

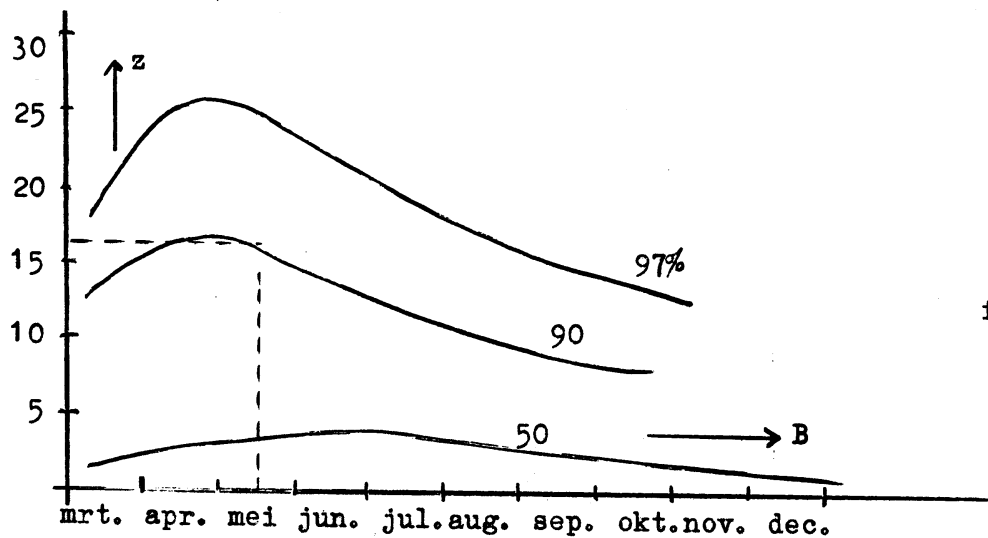


fig.3

Bijvoorbeeld: $B = 77$ (16 mei); z is gedefinieerd door $i = B+W+z-1$. De loodlijn op $B = 77$ ontmoet de F -curve van 90% in de ordinaat $z = 16$. Bij gevolg is $i = 77 + 5 + 16 - 1 = 97$, d.i. 5 juni. Interpretatie: voor $B = 16$ mei, $\beta = 6$, $W = 5$ dagen is er $F = 90\%$ kans, dat het werk uiterlijk op 5 juni klaar is. Mocht men 0.90 te klein vinden, dan kan overwogen worden W te verkleinen (meer materiaal; meer arbeiders; meer kosten).

Misschien komt er dan (via een andere grafiek) $F = 99\%$. Ook kan onderzocht worden of vervroeging van B de F groter maken kan. Misschien ook kan men de h-drempel verhogen. Kortom, de praktijk wenst over vele van deze grafieken (vele h-waarden; vele β -waarden; vele D, W combinaties; krommen van bijvoorbeeld 99, 95, 90, 75, 50%) te kunnen beschikken.

3. DE GEMIDDELDE WACHTTIJD EN DE GEMIDDELDE "KLAARDATUM"

Zie fig.2. In dit geval heeft het 12 dagen, te beginnen met de geplande startdag, geduurd alvorens men het werk achterelkaar (te beginnen op dag 13) kon afmaken. Laten wij deze tijd de "wachttijd" \underline{t} noemen. De naam is niet correct; men houde in het oog, dat men niet werkelijk al die tijd heeft moeten wachten; in het schetsje ziet men, dat men aan het werk reeds werkte op de dagen 5 en 9 (niet op de welis-waar droge dagen 1, 3, 4, 7, 8, 11, 23 en wel omdat er eerst 2 droge na een natte moeten volgen alvorens met het werk opnieuw begonnen kan worden). Deze wachttijd \underline{t} is, evenals de "klaardatum" \underline{E} , een stochastische grootheid; $\underline{t} = \underline{E} - W$.

De praktijk zal zich zeker ook interesseren voor de gemiddelde wachttijd $\underline{\mathcal{E}t} = \underline{\mathcal{E}E} - W$.

Hierover zeggen de auteurs niets. Daarom leiden wij in het volgende de formule voor $\underline{\mathcal{E}t}$ af, met gebruikmaking van de formule voor $f(\underline{E})$ van de auteurs.

Wij stellen vast, dat de kans $g(\underline{t})$ op een wachttijd \underline{t} dezelfde is als de kans $f(\underline{E})$ op een klaardatum $\underline{E} = \underline{t} + W$.

De kleinste wachttijd \underline{t} is 0; het werk begint dan op dag B en komt klaar op dag $B + W - 1$. Bij gevolg is

$$(7) \quad \underline{\mathcal{E}t} = \sum_0^{\infty} \underline{t} g(\underline{t}) = \sum_0^{\infty} \underline{t} f(\underline{t} + W)$$

Deze uitdrukking laat zich in het algemeen niet verder vereenvoudigen. Onderstellen we echter én de jaarlijkse gang én de persistentie afwezig, zodat we ook $B = 1$ mogen stellen (waardoor de berekening vereenvoudigd wordt), dan komt er (zie (6) en fig.2):

$$(8) \quad \begin{cases} y < W - 1, \text{ dan } f_{y+1} = 0 \text{ en } F_{y+1} = 0; \\ y = W - 1, \text{ dan } f_w = q^\beta ; F_w = f_w = q^\beta \\ y > W - 1, \text{ dan } f_{y+1} = (1 - F_{y-\beta}) pq^\beta ; \beta = D + W \end{cases}$$

$$(14) \quad \underline{\epsilon}_t = \left(\frac{1}{pq^W} - 1 \right) - \left(W-1 + \frac{1}{p} \right) = M - N$$

Hierin stelt pq^W de kans voor, bij willekeurig aanwijzen van een dag, een plus-dag te steken, gevolgd door tenminste W min-dagen. Het gemiddelde aantal dagen tussen dit soort plus-dagen in een random reeks van plus- en min-dagen (kansen p en $q = 1-p$), is daarom $(pq^W)^{-1} - 1$. Voorts is de gemiddelde lengte van een min-run, die uit W of meer min-dagen bestaat, gelijk aan $W-1 + \frac{1}{1-q}$ (hier niet bewezen). Zie de volgende schets, fig. 4, waarin $D = 0$, $W = 2$.

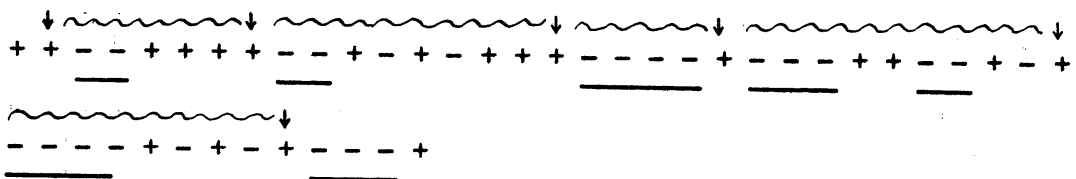


Fig. 4

Met de pijltjes zijn die plus-dagen gemarkeerd, die door tenminste $W = 2$ min-dagen gevolgd worden. De tussenafstanden (in de schets 5, 8, 4, 9, 8 dagen) zijn met \sim aangegeven. De min-runs van lengten ≥ 2 , zijn --- onderstreept (in de schets lengten 2, 2, 4, 3, 2, 4, 3). Als in deze random reeks van plus- en min-dagen $p = 0.4$ en $q = 0.6$, dan zouden (voor een oneindig lange reeks) de \sim segmenten een gemiddelde lengte $M = (0.4 \cdot 0.6^2)^{-1} - 1 = 5.95$ dagen en de --- segmenten een gemiddelde lengte van $N = 2 - 1 + \frac{1}{1-0.6} = 3.50$ dagen hebben, zodat $\underline{\epsilon}_t = 4.45$ dagen.

Wij menen (maar wij kunnen het niet bewijzen), dat dezelfde uitdrukking voor $\underline{\epsilon}_t$, geschreven in de vorm $M - N$, óók voor reeksen, die wel persistentie bezitten, zal gelden. Let wel: wij beweren niet, dat dan voor M en N dezelfde formules zouden gelden als in (14), maar wel dezelfde interpretaties: $M =$ gemiddelde afstand tussen die plus-dagen, die door W of meer min-dagen gevolgd worden en $N =$ gemiddelde lengte van een min-run, bestaande uit tenminste W min-dagen. In een nog niet gepubliceerd rapport laten wij zien, dat voor de gemiddelde lengte van een run van tenminste W droge dagen, geschreven $\bar{L}(dr \geq W)$, geldt

$$\bar{L}(dr \geq W) = W - 1 + \frac{1}{1-q^W}$$

waarin $q^* = q (\geq 1 | \geq W)$ = kans op één of meer droge dagen ná W of meer droge dagen. Zo is er ook een $q^{**} = q (\geq 1 | W)$ = kans op één of meer droge dagen nadat er precies W gepasseerd zijn. Persistentie impliceert : $q (\geq 1 | 1) < q (\geq 1 | 2) < q (\geq 1 | 3)$ etc.

4. NUMERIEKE VOORBEELDEN

4.1 Berekening van f en F bij p=0.2, q=0.8, D=0, W=5.

Wij denken de jaarlijkse gang en de persistentie afwezig, waardoor de uitdrukkingen voor f en F zo eenvoudig mogelijk worden. Zij p=0.2 de kans op een natte (d.i. plus-dag) en q=1-p=0.8 die op een droge (d.i. min-dag). Wij mogen B=1 stellen. Aldus levert (8):

$$f_{y+1} = 0 \text{ en } F_y = 0 \text{ als } y < 4$$

$$f_5 = q^5 \text{ en } F_5 = f_5 \text{ als } y = 4$$

$$f_{y+1} = (1 - F_{y-5})pq^5 \text{ als } y > 4 \text{ met } F_{y-5} = 0 \text{ voor } y < 10$$

$$\text{en } F_{y-5} \neq 0 \text{ voor } y \geq 10$$

$$\text{en } F_{y-5} = \sum_{j=5}^{y-5} f_j, \text{ zie verder tabel 1.}$$

Tabel 1.

y	f_{y+1}	F_{y+1}	Opmerking
4	$0.8^5 = 0.328$	$0.8^5 = 0.328$	De kansverdeling van y en dus ook van $E = y+1$, is min of meer J-vormig, d.w.z. met toenemende y neemt de kans op deze y niet toe (d.w.z. blijft gelijk of neemt af). De grootste kans van optreden heeft hier de kleinste klaardatum, t.w. $E=5$ en wel is deze kans 0.33.
5	$0.2 \cdot 0.8^5 = 0.066$	0.393	
6	"	0.459	
7	"	0.524	
8	"	0.590	
9	"	0.655	
10	0.044	0.699	
11	0.040	0.739	
enz.	enz.	enz.	
34	0.001	0.991	

Een gevolg van de scheefheid van de kansverdeling van de y en dus van de $E = y + 1$ en dus ook van de $t = E - 5$ is, dat de gemiddelde waarde niet met het 50% fractiel coïncideert

$$\epsilon_t = (pq^w)^{-1} - (W + \frac{1}{p}) = (0.2 \cdot 0.8^5)^{-1} - (5 + \frac{1}{0.2}) = 5.3 \text{ dag};$$

$$\epsilon_E = 10.3 \text{ dag}$$

$$\epsilon_y = 11.3 \text{ dag}$$

Zie tabel 2.

Tabel 2.

t	E	y	F%
2.6	7.6	6.6	50
$\epsilon = 5.3$	10.3	9.3	
8.8	13.8	12.8	80
13.5	18.5	17.5	90
18.5	23.5	22.5	95
30	35	34	99

Interpretatie:

Wanneer men als vroegste begin-datum B ("target day") voor zeker werk, dat $W=5$ successieve dagen droog weer vergt, bijv. 15 juni projecteert en de kans op een droge dag $q=0.8$ is (op een natte dag dus 0.2) en indien niet alleen de persistentie mag worden verwaar-

loosd doch ook de jaarlijkse gang binnen bijvoorbeeld 30 dagen, dan zal dit werk, gemiddeld genomen, klaar zijn uiterlijk 24 à 25 juni (omdat $B + \epsilon_E - 1 = 15 + 10.3 - 1 = 24.3$). Er is verder 50% kans, dat het werk op of voor 21 à 22 juni (uit $15 + 7.6 - 1$) klaar zal zijn en 99% kans op of voor 19 juli (uit $15 + 34 = 49$; $49 - 30 = 19$).

4.2 Berekening van de gemiddelde wachttijd voor een aantal p, W-combinaties.

In tabel 3 hebben wij voor $p=0.1, 0.2, 0.3, 0.4, 0.5$ (waarvoor te De Bilt in de maand juli blijkens het frequentieboek No.4 dagsommen van resp. tenminste 8.0, 3.5, 1.5, 0.4 en 0.1 mm behoren) en $W=1, 2, 4, 5$ dagen ($D=0$ onderstellende) de waarden van ϵ_t in dagen bijeengebracht. Binnen het dikomlijnde gedeelte stellen de 4 getallen onder de ϵ_t -waarde de 50, 80, 95 en 99% fractielen van de t -verdeling voor. Opnieuw werd de persistentie en de jaarlijkse gang verwaarloosd.

Andere onderzoekingen leerden ons, dat zulk een verwaarlozing des te meer geoorloofd is naarmate h hoger ligt; m.a.w. de tabel levert ϵt -waarden, die in de rechterbeneden hoek beter de werkelijkheid zullen benaderen dan in de hoek, links boven. Men zou zulk een tabel a.v. kunnen lezen: voor $h=1.5$ mm en $W=5$ dagen, bedraagt $\epsilon t = 11.5$ dag (bij een geplande B-dag 1 juli, is het werk gemiddeld op 11 à 12 juli klaar) en er is 5% kans op een wachttijd van meer dan 31.9 dagen (klaardatum 2 augustus). Het kan zijn, dat 2 augustus te laat gevonden wordt. Men zou W van 5 tot 4 dagen kunnen bekorten (meer arbeiders; meer werktuigen; hogere kosten) dan gaan 11.5 (11 à 12 juli) resp 31.9 (2 augustus) over in 6.6 (6 à 7 juli) en 19.2 (19 à 20 juli). Zou men h kunnen verhogen van 1.5 tot 3.5 mm/dag (p verlagen van 0.3 tot 0.2), dan gaan 11.5 en 31.9 over in resp. 5.3 (5 à 6 juli) en 17.5 (17 à 18 juli). Tenslotte kan men ook overwegen én h van 1.5 tot 3.5 mm/dag te verhogen én W van 5 tot 4 dagen te bekorten. Dan $11.5 \rightarrow 3.3$ en $31.9 \rightarrow 11.5$ dagen.

5. SLOTOPMERKING (doel van de studie)

Deze studie had tot doel enig inzicht te krijgen in de aard en omvang van die ponsmechanische berekeningen, op basis waarvan voor enige in ons land gelegen regenstations (natuurlijk denken wij aan de in de frequentieboeken behandelde 24 stations) verzamelingen grafieken zouden kunnen worden getekend, waarmede snel vragen van het type als in de inleiding geformuleerd werd zouden kunnen worden beantwoord. Naar onze mening zou men daarmee de praktijk een goede dienst bewijzen.

Tabel 3. De Bilt; juli.
Gemiddelde wachttijd en 50, 80, 90, 95% fractielen.

p h Wdg	0.5	0.4	0.3	0.2	0.1
	0.1	0.4	1.5	3.5	8.0 mm
1	1	0.67	0.43	0.25	0.11
2	4	2.4	1.5	0.81	0.35
4	30	12.8	6.6 6.4 14.0 19.2 24.6	3.3 4.1 8.3 11.5 14.5	1.2
5	57	25	11.5 10.9 22.9 31.9 41.0	5.3 6.6 12.8 17.5 22.3	1.9

6. LITERATUUR.

- [1] K.N.M.I.; "Frequenties van k-daagse neerslagsommen op Nederlandse stations", No.25 "Verklaring en toelichting", deel B.
"Tabellarische en grafische, parameter vrije behandeling van enige aktuele neerslagvragen uit de praktijk", 1963.
- [2] K.N.M.I.: "Frequenties van k-daagse neerslagsommen op Nederlandse stations" Nummers 1 t/m 24.
- [3] Mc QUIGG, J.D. and DECKER, W.L.: "The probability of completion of outdoor work"; Journ.of applied Meteor. 1 (1962), 178.