

Enkele gedachten over de kosten van proefvelden

G. J. VERVELDE,

Instituut voor Biologisch en Scheikundig Onderzoek van Landbouwgewassen, Wageningen

1 INLEIDING

De kosten van proefvelden kunnen om twee soorten redenen hoger zijn dan strikt noodzakelijk is.

Allereerst kan de uitvoering duurder worden dan nodig is. Verbetering van de experimenteervaardigheid bij de voor de proeven verantwoordelijke onderzoekers en ontwikkeling van hun kosteninzicht zullen dan tot besparingen leiden.

Het kan ook zijn, dat de kosten „hoog” blijken te zijn ten gevolge van de lage gebruikswaarde der verkregen uitkomsten, bijv. wegens gebrekkige betrouwbaarheid blijkend uit de grootte der toevalspreiding. Voor dit punt roept men gewoonlijk statistici als doelmatigheidsexperts te hulp. Kennis van doeltreffende proefopzetten en doelmatigheidsgevoel gaan bij hen hand in hand. Op grond van dit laatste zullen zij zich af en toe ook wel eens een opmerking veroorloven over de uitvoering van de proef. De onderzoeker zal dan met zijn technisch inzicht moeten vaststellen of een dergelijke aanbeveling al of niet deugdelijk is.

Beide soorten kostenfactoren, de eerste van technische en de tweede van statistische aard, zullen in het navolgende iets nader worden gezien.

¹ Voordracht voor de Studiek. v. Stat. Techn. op 18-2-'60.

2 KOSTENFACTOREN VAN TECHNISCHE AARD

Besparingen moet men bij de uitvoering van proefveldenwerk om te beginnen nagenoeg geheel in dezelfde richting zoeken als bij elke andere bezigheid. Het gaat dan om zaken als:

- een weldoordachte werkplanning
- doelmatiger werkwijzen
- inzet van machines voor dure handenarbeid
- verhoging van de vaardigheid der uitvoerders, enz.

Voor proefvelden kornen daar nog wel enkele bijzondere factoren bij, zoals:

- lengte van de reisroute van personeel en materiaal
- verspreidingspatroon van proefvelden, die gelijktijdig zorg vereisen
- mogelijkheid om een deel van het werk door plaatselijke krachten te doen verrichten.

Heel veel verbeteringen zijn in de loop der jaren al werkende aangebracht. Zij zijn verwerkt in de diverse handleidingen, die in verschillende talen bestaan voor de inrichting en verzorging van proefvelden, bijv. Handleiding voor veldproeven, 1951 (1). Toch zijn er nog steeds verdere aanpassingen nodig, in het bijzonder

doordat de arbeid steeds duurder wordt. Zo werd enige jaren geleden een „Werkgroep Rationalisatie Proefveldwerkzaamheden” aan het werk gezet om nog weer nieuwe verbeteringen op te sporen. Daaruit ontstond een rapport „Rationalisatie van Proefveldwerkzaamheden” (2). Het beschrijft diverse nieuwe of vernieuwde methoden. Er wordt vooral gewezen op het nut van machinaal werken, met aanpassing van de veldjesafmetingen aan de werkbreedte der machines. Het blijkt gewenst, er naar te streven, dat de proeven zoveel mogelijk met de in de praktijk gebruikte machines bewerkt kunnen worden. Men behoeft dan geen bijzondere werktuigen aan te schaffen.

Technische kostenfactoren, die zich nauwelijks lenen voor beïnvloeding maar die toch de moeite van kennisneming ten volle waard zijn, zijn: de soort proef (het onderwerp van de proef dus) en de aard van het verbouwde gewas. Een in het jaar 1954 door de Inspectie Landbouwkundig

Onderzoek ondernomen analyse van de materiële kosten van ruim 1500 door de Rijkslandbouwconsulenten beheerde proefnemingen leert hierover meer. Enkele cijfers als voorbeeld van de strekking der verkregen gegevens zijn opgenomen in tabel 1.

Het was bij de bedoelde studie helaas niet mogelijk ook de loonkosten te achterhalen. Globaal kan men zeggen, dat deze wel 75 procent van de totale kosten uitmaken. Een hectare proefvelden met landbouwgewassen komt thans gemiddeld op f 2500 kosten met afwijkingen naar boven of naar beneden, al naar het betreffende gewas meer of minder teeltzorg vereist en al naar het duurdere of minder dure typen onderzoek betreft. Grotere oppervlakte van een proef, bij gelijkblijvend aantal vakjes, verhoogt de kosten. Hetzelfde geldt voor een groter aantal vakjes bij gelijkblijvende oppervlakte.

Een soortgelijke kostenanalyse wordt opnieuw ondernomen voor het jaar 1959.

Tabel 1 Gemiddelde materiële kosten van consulentenproeven in 1954

	Aantal proeven	Gem. opp. per proef in ares	Gem. kosten per proef in guldens
<i>Gemiddelde van alle proeven</i>	1570	17	95
<i>Proefgewas</i>			
Granen	405	11	76
Vezelgewassen	30	13	124
Grasland	338	30	129
<i>Onderwerp</i>			
Grondbewerking	36	68	136
Gewas- of rasvergelijking	427	13	67
Cultuurtechniek	21	125	228
Ziekten, plagen en onkruiden enz.	124	7	32

3 KOSTENFACTOREN VAN TECHNISCHE EN TEVENS VAN STATISTISCHE AARD

Onder dit hoofdje horen factoren thuis zoals de gelijkmatigheid van het proefterrein en van de verdere proefomstandigheden, de grootte der proefvakjes en de vorm dezer vakjes. Dat ze de kosten allereerst rechtstreeks beïnvloeden, zal bij enig nadenken gemakkelijk duidelijk zijn. Doorgaans is het niet raadzaam om deze factoren zodanig te regelen, dat de proef in geld uitgedrukt zo goedkoop mogelijk wordt. Het kan dan namelijk gebeuren, dat de toevalscomponent van de totale variantie der verkregen waarnemingen dusdanig ongunstig komt te liggen, dat het proefveld toch ondoelmatig en daardoor te kostbaar geacht moet worden. Anderzijds is het ook weer niet wenselijk de bedoelde factoren zodanig te reguleren, dat de toevalsvariantie maximaal gunstig komt te liggen. Gelijkmatigheid van het proefterrein is bijv. slechts met zeer veel kosten en dan nog onvolledig te bewerkstelligen.

Helaas was het bij de studies van de reeds genoemde Werkgroep voor de Rationalisatie van Proefveldwerkzaamheden niet mogelijk aan deze factoren gedegen aandacht te geven. Hiervoor zou o.a. een uitvoeriger studie nodig zijn van de gegevens van blanco proeven of van de restspreiding van grotere veldproeven. Daarvoor bleek in de betreffende korte werktijd van de Werkgroep geen gelegenheid te zijn. Het blijft echter een punt, dat vooral in de aandacht van de proefveldstatistici aanbevolen mag blijven.

Elke proeftechnisch onderlegde statisticus is wel vertrouwd met de syste-

matische proefschema's, waarmee men de gevolgen van de ongelijkmatigheid der proefomstandigheden zo klein mogelijk maakt. Het lukt dan min of meer om de systematische componenten van de ongelijkmatigheidsvariantie onschadelijk te maken. Het zou zeker de moeite lonen nog eens een nadere studie te maken van de aard van de ongelijkmatigheid, zoals die op onze proefvelden wordt aangetroffen.

Heeft deze bijv. een gestreept patroon, al of niet in twee richtingen, dan zijn schema's met eliminatie van rij- en kolomeffecten doelmatig en legt men liefst gerekte proefvelden met de lengtes loodrecht op de hoofdrichting van de streping aan. Is een gelijkmatig golvend patroon aanwezig, dan kan men wellicht werken met vereffening volgens een vloeiend vruchtbaarheidsvlak, terwijl de vorm der veldjes dan in hoofdzaak op het gemak van werken gericht mag zijn. Is het patroon daarentegen bont vlekkelig en grillig, dan is vereffening volgens zeer kleine onvolledige blokjes wellicht juist.

Dat de grootte der veldjes de kosten beïnvloedt is reeds vermeld. Meestal gaat men er van uit, dat grote veldjes, omdat ze doorgaans ook tot grote proefvelden leiden, onvoordelig zijn voor de grootte van de restspreiding. Met de grotere afstand zou de kans op toevalsafwijkingen toenemen. Bovendien zijn grote proefvelden niet gemakkelijk in kort tijdsbestek te verzorgen of af te werken, hetgeen ook de ongelijkmatigheid pleegt te versterken. Een en ander zou dan wijzen in de richting van kleine veldjes.

Aan de verkleining van proefvakjes zijn intussen ook grenzen, al kan men vrij ver gaan. De ondoelmatige peuterachtigheid bij de bewerking (extra arbeid!), de grote kwetsbaarheid, en

vooral ook de individuele plantverschillen bij kleine aantallen planten per vakje bepalen de ondergrens. Ongelijkmatigheid van plant tot plant kan op de erfelijke constitutie van het gewas (bieten, asperges) berusten, maar ook op verschillen in voorgeschiedenis (fruit- en bosbomen). Een andere ervaring is, dat verkleining van de proefvakjes beneden een zeker niveau ook nauwelijks meer tot opvallende verkleining van de ongelijkmatigheid van vakje tot vakje leidt. Blijkbaar is het micropatroon van de ongelijkmatigheid al even grillig als het macropatroon.

4 KOSTENFACTOREN VAN STATISTISCHE AARD

a *De totale omvang van proefnemingen als kostenfactor*

In het voorgaande is al wel duidelijk geworden, dat ook bij de beste technische inrichting van de proef toch altijd een niet uit te schakelen invloed van ongelijkmatigheid in proefomstandigheden overblijft. Statistisch inzicht leert, dat het middelen van door herhaalde waarnemingen verkregen cijfers tot betrouwbaarder uitkomsten leidt. Parallellen brengen echter stijging van kosten mee, hetgeen alleen aanvaardbaar is zolang de extra kosten overtroffen worden door de extra doelmatigheid ten gevolge van grotere betrouwbaarheid. Men moet dus bij het vaststellen van het aantal parallellen niet verder gaan dan tot de optimale doelmatigheids-/kostenverhouding bereikt is.

Het aangeven van het minimum aantal herhalingen voor een proef was in de ontwikkeling van de statistische proeftechniek een der eerste glans-

punten bij het rationaliseren van proefnemingen. De basisgegevens die men bij het ontwerpen van één enkele proef nodig heeft, zijn: de grootte van de te verwachten toevalsvariantie, de minimumwaarde van het te achterhalen gegeven en het overschrijdingsrisico, d.i. de nog aanvaardbaar geachte kans dat men de bedoelde minimumwaarde ten onrechte onderscheidt van nul of van enigerlei andere waarde.

Bij heel veel onderzoeken ligt de zaak echter moeilijker en ondoorzichtiger. Bijna altijd heeft men nl. met series van proefnemingen te maken, waarbij een veel groter aantal overwegingen te pas komt. Er zijn in ieder geval twee soorten serieproeven te onderscheiden.

α Men wenst de uitkomsten der afzonderlijke proeven op zich zelf te interpreteren. Zo kan men bijv. het plan hebben deze uitkomsten in verband te brengen met de bij elke proef geldende milieu-omstandigheden, om ten slotte een totaalbeeld te verkrijgen van het reactiepatroon van de gemeeten grootheid of grootheden. In feite heeft men dan het geval dat elke proef op zichzelf gegevens moet opleveren met een voldoende betrouwbaarheid, juist zoals in het geval van één enkele proef. Alleen mag men verwachten, dat de afzonderlijke proeven elkaar wat betreft de bestudering van kwantitatieve milieufactoren wel ondersteunen, zodat in zo'n geval wellicht op de betrouwbaarheid der uitkomsten van de afzonderlijke proeven kan worden toegegeven.

β De serie heeft geen verdergaande bedoeling dan om een gemiddeld

eindgegeven voor een verscheidenheid van milieu's op te leveren; ze tracht daartoe een zo representatief mogelijk monster van de verschillende milieu's te zijn. Men wil dan de uitkomst gebruiken als aanwijzing voor nog te verwachten toekomstige uitkomsten, ten behoeve van voorspellingen dus.

Deze werkwijze is in de landbouw niet ongebruikelijk en af en toe zelfs onvermijdelijk. Zo zal men bijv. het gedrag van planterrassen niet voor de velerlei grondsoorten en bemestingstoestanden afzonderlijk kunnen vaststellen. Evenmin gaat dat voor elke weersgesteldheid, die men binnen het heersende klimaat zou kunnen verwachten, nog afgezien van het feit dat men deze kennis toch niet in praktijk kan brengen doordat men onkundig is van het weer van het komende seizoen. Evenzo zal men ook geen kans zien om het voedereffect van een voedermiddel bij dieren vast te stellen voor elke erfelijke constitutie, die binnen de gebruikte vee-rassen kan voorkomen, noch voor alle typen rantsoenen, die op onze landbouwbedrijven in gebruik zijn.

Om nu voor dergelijke verzamelingen van milieu's toch met een behoorlijke trefkans uitspraken te kunnen doen, moeten ook weer de waarnemingen in zodanige aantallen worden verricht, dat de onzekerheid binnen aanvaardbare grenzen komt. Nu bestaat de spreiding bij dergelijke proevenseries gewoonlijk uit diverse systematische componenten naast „toevallige” componenten. Zo zal bij series rassenproeven een deel van de spreiding ontstaan door de verschillende oogstjaren, een ander deel ten gevolge van grondsoortenverschil, enz. Hamming (3) spreekt in dit verband van „heterogene universa”. De onzekerheid, die het gevolg is van deze verschillende spreidingsoorzaken, kan in dit geval worden verkleind door de proeven-

serie te verdelen over meer dan één jaar en over de verschillende grondsoorten, terwijl de toevalspreiding door parallellen binnen de proeven moet worden achterhaald. Men kan wel aanvoelen, dat de herhalingen evenwichtig over de verschillende spreidingsoorzaken verdeeld moeten zijn. Belegt men nl. één dezer spreidingsoorzaken niet of onvoldoende met herhalingen, bijv. de jaarinvloed niet, dan zal deze onzekerheidscomponent volledig blijven bestaan, hoe krachtig men de andere componenten ook door herhalingen gaat terugdringen.

Het is om die reden dat Hamming in zijn reeds genoemde artikel voor samenvatting van rassenproeven een systeem van gewichten heeft ontworpen.

Een dergelijk gewicht wordt toegekend aan bepaalde sub-gemiddelden op basis van het aantal waarnemingen waarop dat gemiddelde berust. Het blijkt dat het gewicht ten slotte nauwelijks meer toeneemt, ook al blijft het aantal herhalingen, dat voor het sub-gemiddelde is gebruikt, groeien. Men komt tot het werken met gewichten, als men niet in de hand heeft, hoe de waarnemingen verdeeld zullen zijn over de verschillende spreidingscomponenten.

Het is goed mogelijk (zoals een kleine niet gepubliceerde studie van Hoveyn, verbonden aan het I.B.S., dezer dagen nog weer bevestigde en waarop Hamming ook al wees) om reeds van de aanvang af op een evenwichtige verdeling van de herhalingen aan te sturen.

Voor beide bedoelde werkwijzen (gewichtentoekenning, c.q. het tevoren ontwerpen van een voordelig proevenprogramma) is het nodig dat men cij-

fers in handen heeft van de verwachte grootte der variantiecomponenten. Deze ontbreken veelal. Het zou niet ondienstig zijn voor dit doel een aantal reeds voltooide proevenseries nog eens te analyseren.

Wellicht (maar dit is een hypothese) kan men verder nog iets doen met de veronderstelling, dat zgn. oogstzekere gewassen de neiging hebben tamelijk standvastig te reageren van jaar tot jaar (dus een kleine jaarcomponent in de totaal-variantie op te leveren), terwijl de wispelturiger gewassen omgekeerd reageren. Eenzelfde gedachten-gang kan worden gevolgd voor de reactie op de verschillende groeiplaatsen: er zijn gewassen, die zich op een verscheidenheid van grondsoorten goed redden, die een brede zgn. „ecologische spreiding” hebben, en andere die erg kieskeurig zijn.

Er is één kostenaspect, dat men bij het samenstellen van de voordeligste proevenserie niet over het hoofd mag zien. Men zou nl. zijn proefvelden-dienst zeer ongelukkig maken, wanneer men met een programma zou komen, waarin de proeven voorkomen met bijv. slechts één parallel. Daardoor zouden er immers zeer veel van dergelijke proeven nodig zijn. Elke proefveldpracticus beseft nl. uit ervaring dat met elke nieuwe proef een groot bedrag aan aanloopkosten gemoeid is, nagenoeg onafhankelijk van de grootte van de proef. Men kan het ook zo zeggen, dat a proeven met b parallellen lang niet zo duur zijn als ab proeven met 1 parallel.

De bij iedere proef onvermijdelijke vaste kosten omvatten bijv. het ontwerpen en op papier vastleggen van de proef, het opsporen van en het overleg met een proefveldhouder, de aan- en afvoerkosten van personeel en materiaal, verliezen ten gevolge van randen e.d. Het blijkt echter, volgens

een door Hoveyn uitgewerkt geval, ook dan geen bezwaar de bovengenoemde beschouwingen door te voeren, wanneer maar een schatting voorhanden is van de verhouding der vaste kosten van een proefveld tot die van een extra parallel. Men moet dan zijn programma afstemmen op de voordeligste uitvoering en niet op de puur statistisch berekende doeltreffendste verdeling der herhalingen.

Het bestaan van de zojuist bedoelde heterogene universa is ongetwijfeld dikwijls een groot struikelblok voor het snel verkrijgen van trefzekere voorspellingen. Er bestaat dan ook altijd de neiging om deze universa homogener te maken ter wille van de experimenten. Zo iets zou bijv. mogelijk zijn door rassenproeven steeds op één of slechts op een paar naverwante grondsoorten te leggen. Men maakt op overeenkomstige wijze bij dierproeven wel gebruik van eeneïge tweelingen om bepaalde effecten scherp vast te stellen. Men moet dan echter wél beseffen dat de generalisatie van de uitkomsten, het gebruiken ervan voor voorspellingen dus, een matige of zelfs slechte aanpassing bij de werkelijkheid oplevert. Het vernauwen van het universum is daarom vooral bruikbaar wanneer men in een vroeg stadium van een onderzoek al wil weten of een bepaalde werking of invloed bestaat. De gewetensvolle onderzoeker zal daarna zich vergewissen of de uitkomst onder de gebruikelijke praktijkomstandigheden eenzelfde geldigheid bezit. De neiging om de uitkomsten van zeer smalle universa nadrukkelijk naar voren te brengen zonder op de beperkingen te wijzen, heeft af en toe een commerciële reden, maar behoort toch zoveel mogelijk beteugeld te worden.

b *De gewenste zekerheid als kostenfactor*

Betrouwbaarder uitspraken over grootheden, die onderhevig zijn aan toevallige spreiding, kunnen in het algemeen slechts verkregen worden door rijkere waarnemingen en dus door besteding van meer kosten. Hoe lager de betrouwbaarheidsdrempel is, waarmee men genoegen kan nemen, hoe voordeliger in het algemeen de proevenseries kunnen zijn. De ligging van die drempel is al van oudsher een punt van levendige discussie in kringen van statistici geweest, waarop thans niet nader hoeft te worden ingegaan. Eén punt slechts zij genoemd, dat soms een puur statistisch ingestelde werker wel wat verwonderd doet staan kijken. Het blijkt nl., dat de praktijk van de landbouw soms met erg onzeker schijnende gegevens al tot een duidelijk oordeel komt. De praktijk heeft nl. steeds te beslissen voor slechts een klein deel van het heterogene universum, een sterk vernauwd universum dus, in tegenstelling tot de meer centraal geplaatste onderzoeker die een uitspraak voor het hele universum nastreeft. Bovendien zal de praktijk zijn oordeel baseren op een complex van overwegingen, niet op een aantal los van elkaar staande overwegingen. Een voorbeeld uit het rassenonderzoek, geheel hypothetisch samengesteld overigens, kan dit duidelijk maken.

Stel dat bij vergelijking van twee graanrassen voor een vrij breed traject van omstandigheden na enige jaren beproeven de uitspraak bereikt wordt, dat ras A t.o.v. ras B wat eerder legt, meer neiging tot schot geeft en min-

der stro oplevert, waarbij deze drie uitspraken echter slechts berusten op resp. 75 %, 80 % en 80 % positieve waarnemingen, terwijl de overige waarnemingen in andere richting wijzen. De zaadopbrengst van ras A lijkt wèl hoger te zullen zijn dan van B; dit is nl. gevonden in 90 % der waarnemingen. De heterogeniteit der uitkomsten kan ten dele aan het toeval, ten dele aan de heterogeniteit van het universum worden toegeschreven. Een onderzoeker die op grond van dergelijke uitkomsten een generaliserende uitspraak zou moeten doen over de landbouwkundige voorkeur voor het ene resp. het andere ras, zal willen wachten op aanvullende cijfers. De praktijk weet in veel gevallen echter voor zijn eigen omstandigheden al voldoende. In een streek waar legering weinig waarschijnlijk is en het stro niet veel waard is, zal de kans op meer opbrengst de doorslag geven in de richting van ras A. Heeft men veel last van legering en is de stroopbrengst waardevol (het stro dus duur), dan is er gemiddeld slechts $0,25 \times 0,20 = 0,05 = 5\%$ kans dat ras A in beide opzichten toch nog zal meevallen. Deze kans is te klein, om daarmee op de grotere zaadopbrengst van ras A te gokken.

Het is met deze inzichten voor ogen niet langer onbegrijpelijk, hoe sommige landbouwconsulenten jaarlijks voor hun gebied een rassenbeoordeling in druk gereed kunnen hebben op het ogenblik waarop de opbrengstgegevens nog niet vaststaan. Men heeft zijn oordeel nl. geveld op grond van de andere, reeds verrichte waarnemingen en hoeft dat oordeel slechts zelden op het laatst te wijzigen, wanneer totaal onverwachte opbrengstcijfers daartoe dwingen. Een landelijke berichtgever, die noodzakelijkerwijs elk

produktiekenmerk der rassen afzonderlijk wil karakteriseren, zit hiermee echter veel minder gemakkelijk!

5 SLOTPMERKING

Voor degenen, die betrokken zijn of raken bij het uitstippelen van het brede onderzoekbeleid, nog enkele opmerkingen over kostenaspecten. Er is een neiging van onderzoekers en ook van onderzoekinstanties, om goedkoop onderzoek voorrang te geven. Hierboven is reeds iets gezegd over kostenverschillen tussen verschillende typen onderzoek. Het is geen toeval dat ons landbouwkundig onderzoek het eerst begonnen is met rassen- en met ziektebestrijdingsonderzoek, ter-

wijl grondbewerkings- en cultuurtechnisch onderzoek lang op zich heeft laten wachten. Ter voorkoming van onevenwichtigheden zal de hogere leiding van het onderzoek dan af en toe speciale stappen moeten doen om de ontstane hiaten, tegen de neiging van de rechtstreeks betrokkenen in, op te vullen.

Een tweede punt, dat nog weer eens herhaald mag worden, is het in kringen van statistici reeds overbekende gezichtspunt, dat de kosten besteed aan een goede voorbereiding van de proefnemingen en aan een zodanige opzet, dat wiskundige bewerking vlot mogelijk is, altijd verantwoord zijn.

LITERATUUR

- 1 Handleiding voor veldproeven. *Mededeling nr. 59 van de Landbouwvoorlichtingsdienst*. 1951. (Een herdruk is in voorbereiding).
- 2 Rapport „Rationalisatie van Proefveldwerkzaamheden” (1958). Nat. Raad v.h. Landbouwkundig Onderzoek T.N.O., Afd. Akker- en Weidebouw, Den Haag.
- 3 HAMMING, G.: Interactie problemen. *Landbouwkundig Tijdschrift* 64 (1952) 162—173.