

University of Groningen

## 20 Jaar statistisch rekenen

Stokman, F.N.

*Published in:*  
EPRINTS-BOOK-TITLE

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
1983

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
Stokman, F. N. (1983). 20 Jaar statistisch rekenen. In *EPRINTS-BOOK-TITLE* s.n..

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

## 20 jaar statistisch rekenen

F.N. Stokman, Rijksuniversiteit te Groningen

"It is undesirable for a statistical package to be the product of a single person. What should individuals do who want to develop statistical software? One should try to become part of a group such as those who participated in NAG, GLIM and GENSTAT. Some packages, such as SAS and RS/1, have taken effort to make it easy for sophisticated users to add their own modules. Many individuals have also integrated their work with SPSS and P-STAT. SAS, SPSS and RS/1 have sizable user groups that facilitate integrating (at least at a local level) of user developed software with the major package. When developing software, one should use components from standard subroutine libraries such as those of IMSL and NAG as much as possible."

James W. Frane, 1983, 71.

### 1. Ontwikkelingen in de jaren '60 en '70

Onder de wel zeer brede titel "20 jaar statistisch rekenen" die de organisatoren van het Symposium Statistical Software 1983 aan mijn lezing hebben meegegeven, wil ik een schets geven van de hoofdlijnen van de ontwikkelingen op het gebied van statistisch rekenen, zoals die zich binnen de Nederlandse universiteiten op het terrein van de sociale wetenschappen hebben voorgedaan, om vandaar een aantal aanbevelingen te doen voor de toekomst. Ik wil u daarbij deelgenoot maken van mijn eigen "ommezien in verwondering" over de snelheid van deze ontwikkelingen: met het huidige voorzieningenniveau realiseert men zich onvoldoende hoe kort deze voorzieningen nog maar bestaan. De volgende anecdote moge dat verhelderen. In maart 1968, dus slechts 15 jaar geleden, heb ik mijn doctoraalscriptie politicologie over politieke opinieleiders voltooid. De scriptie is gebaseerd op een enquête onder 269 mannelijke kiesgerechtigden in Amsterdam, die gehouden was in het kader van een werkcollege onder leiding van drs. F. Bergsma en (toendertijd) drs. R.J. Mokken. In het onderzoek werden een aantal indices, zoals die voor politieke kennis, politiek zelfvertrouwen en politiek opinieleiderschap, geconstrueerd met behulp van cumulatieve schaalanalyse volgens het Mokken-model. Met uitzondering van de schaalanalyses zelf, die door het Mathematisch Centrum werden verricht, werden alle overige bewerkingen van het materiaal door mij verricht met behulp van een tel- en sorteermachine, de Bull D3, en de PRD kaartlezer en -ponser. Voor elke kruistabel moesten

de ponskaarten eerst op de eerste variabele worden uitgesorteerd en vervolgens wederom doorgevoerd om per categorie van de eerste variabele de frequenties op de tweede variabele vast te stellen. Ook de sommatiescores van de Mokken-schalen moesten via sorteerprocedures worden vastgesteld en als extra variabele in de ponskaarten worden gestanst, waarna zij via kruistabellen op andere variabelen konden worden betrokken.

Toch moet op dat moment de stand van zaken aan de Universiteit van Amsterdam als uitzonderlijk worden beschouwd. De Universiteit van Amsterdam had geen eigen rekencentrum, zoals dat buiten Amsterdam gebruikelijk was. Op instellingsniveau was met de Stichting Mathematisch Centrum een contract afgesloten ten behoeve van consultatie en rekenwerk. Hiermee lag de gehele verwerkingsfase van onderzoek in handen van een externe organisatie. De hiervoor benodigde programmatuur werd als intern product van het Mathematisch Centrum en niet als eindproduct van de consultatie beschouwd. Interessante numerieke en statistische procedures werden in een procedurebibliotheek opgenomen en eventueel in MC-publicaties of elders gepubliceerd. Deze overdracht van de verwerkingsfase naar een externe instantie werd in toenemende mate als een belemmerende factor voor de opbouw en overdracht van eigen kennis en ervaring op dit gebied ervaren, temeer daar de oriëntering van het Mathematisch Centrum op de door haarzelf ontwikkelde computers (ARRA, EL-X1 en EL-X8) en compilers (Algol-60) een aansluiting op elders ontwikkelde statistische en verwerkingsprogrammatuur belemmerde. Vanuit deze onvrede werd dan ook op 28 juni 1966 het "Technisch Centrum voor de Faculteit der Sociale Wetenschappen" door de Faculteit der Sociale Wetenschappen van de Universiteit van Amsterdam ingesteld. Feitelijk begon het Technisch Centrum pas in 1968 te functioneren (Van der Wijngaart, 1978).

Ontwikkeling van (standaard)programmatuur werd de belangrijkste taak van het Technisch Centrum in de eerste jaren van haar bestaan. Ook elders vindt inmiddels veel ontwikkeling van programmatuur plaats. Omdat in die tijd ook elders de verwerking van onderzoeksgegevens veelal gelegd werd in handen van speciaal daarvoor aangetrokken personeel en nog niet werd overgelaten aan de onderzoeker zelf, lag het accent op de ad-hoc programmering. Zo ontstonden aan de Katholieke Universiteit van Nijmegen reeds vanaf 1965 aparte kernen met programmeurs binnen de disciplines psychologie, sociologie en pedagogie, die nauw geliëerd waren met methodologen. Samenwerking tussen de disciplines startte pas in 1968, zodat daarna althans enige afstemming en wederzijds gebruik van programmatuur gerealiseerd werd. In Groningen kwamen de eerste aanzetten tot standaardisatie wat eerder tot stand in het hercoderings- en rechte-tellingen-programma ZAEHL en een reeks pro-

gramma's vanuit het rekencentrum en psychologie (AHP en GRON PSYCH series).

Een eerste poging tot landelijke afstemming en bundeling van menskracht werd door drs. A.J. van Reeken (Katholieke Hogeschool in Tilburg) ondernomen die eerst op de Statistische Dag in 1965 en vervolgens op die in 1966 een oproep plaatste om te komen tot de oprichting van een sectie "Statistiek en Computer" binnen de VVS. Dit resulteerde in de Werkgroep Statistiek en Computer (WSC), die voor het eerst in april 1966 bijeenkwam en in februari 1967 door het VVS-bestuur de status van werkgroep binnende VVS verkreeg. Op basis van een enquête onder haar leden werd als doelstelling van de werkgroep geformuleerd: "het vergemakkelijken van de programmering van statistische analysetechnieken alsmede het gebruik van statistische standaardprogramma's, ten dienste van statistici en rekencentra". Naast de organisatie van voordrachten, discussie-bijeenkomsten en een centrale programmatheek in de vorm van een inventarislijst was een belangrijke taak van de WSC de oprichting van werkcommissies. Deze moesten zich bezighouden "met de vastlegging van eisen waaraan een standaardprogramma moet voldoen, de vervaardiging van dit programma (inclusief de wiskundige hulpmodulen), de vervaardiging van een programmabeschrijving, en de opsomming van de toepassingvoorwaarden". (Van Meurs en Van der Weele, 1982) Met deze initiatieven en ideeën is Van Reeken zijn tijd ver vooruit. In feite anticiperen zij reeds op de situatie waarin de verwerking van gegevens en gebruik van programmatuur bij de onderzoeker zelf wordt gelegd, met alle problemen van ondeskundig gebruik vandien. Van Reeken werkt dit uit in een artikel in *Statistical Computing*, waarin hij als uitgangspunt neemt dat geen enkel rekencentrum in staat is genoeg statistici te hebben "- to screen the statistical computations carried out on their machine for statistical misapplications. -- I cannot imagine that statisticians are the only ones using these programs. The only possible solution I see for this problem is to have 'foolproof' programs". (Van Reeken, 1970, 74).

Dergelijke "gekbestendige" programma's moeten ook nog "algemeen" zijn: "This means that a program performing a particular analysis should include 'all kinds of everything' relevant to this analysis, and a program testing a particular hypothesis should do this, with regard to the underlying basic assumptions of the data. Of course, there is no need to have this big programs constantly in core, and, of course they will be programmed as modules. Both types should take data from a data-base. For output we need a report generator, not merely a memory dump". (Van Reeken, 1970, 75)

Omdat de werkcommissies van de WSC niet voldoende van de grond kwamen, werd in 1969 de taakstelling van de WSC tot een minder ambitieus niveau teruggebracht ("inventariseren en uitwisselen van programma's"), maar ook dat

was een onvoldoende stimulans, zodat de WSC in 1970 de geest gaf.

"Gekbestendige" programma's zijn nooit gerealiseerd, maar "algemene" programma's kwamen eerder dan toen werd voorzien, niet door eigen ontwikkeling, maar door import van programmapakketten vanuit de Verenigde Staten. Was SPSS het eerst in Leiden beschikbaar, zowel vanuit Nijmegen als Amsterdam was het baken eerst gericht op OSIRIS dat in Michigan was ontwikkeld (1971). Spoedig daarna echter werd ook daar het baken verzet naar SPSS, mede wegens de conversieproblemen van OSIRIS naar CDC computers, terwijl de conversie van SPSS naar CDC juist voltooid was toen besloten was tot aanschaf van CDC-apparatuur in Amsterdam, Groningen en Utrecht (1972). In het overleg over universitaire statistische programmatuur, dat uit de gemeenschappelijke keuze voor CDC tussen Amsterdam, Groningen en Utrecht voortvloeide (OUSP-overleg), ontstond een scherp verschil van inzicht over de vraag welk algemeen pakket het meest wenselijk was: OSIRIS, SPSS of een eigen te ontwikkelen pakket, gebaseerd op een door CDC aangekondigd, maar nooit aangeboden pakket. Bovendien presenteerde Groningen zijn ideeën om voortbouwend op ZAEHL en de AHP/PSYCH series een Gronings Statistisch Pakket te ontwikkelen (GRONSTA). Nadat dit bruusk door Amsterdam en Utrecht werd bekritiseerd en afgewezen, is Groningen op deze weg doorgegaan onder de naam WESP ofwel Waarlijk Eenvoudig Statistisch Pakket. Dit betekende het voorlopige einde van elk overleg tussen de drie instellingen (1973). WESP werd langzaam maar zeker verder uitgebouwd. In 1974 volgde een interactieve versie (TTWESP) en in 1975 de conversie van Algol-60 naar FORTRAN (Van der Weele et.al., 1978). Buiten Groningen is WESP feitelijk nooit gebruikt. In Groningen is WESP lange tijd aanzienlijk vaker gebruikt dan SPSS (in 1975 in de verhouding 4 : 1), maar na 1981 is dit drastisch gewijzigd (in 1982 1 : 4) ten gevolge van de overschakeling van psychologen en sociologen op het statistiekboek van Nijdam en Van Buuren (1980) met de daarbij behorende SPSS-oefeningen. Hiermee is het lot van het enige Nederlandse algemene programmapakket definitief bezegeld en mogen op dit vlak geen eigen Nederlandse ontwikkelingen meer worden verwacht.

Vanzelfsprekend kon SPSS niet voldoen aan alle statistische behoeften. Afgezien van het beperkte scala van statistische technieken dat in de beginfase in SPSS was opgenomen, is het inherent aan een algemeen programmapakket dat het slechts veel gebruikte en beproefde statistische technieken bevat. Voor alle universitaire rekencentra bleef het daarom noodzakelijk een systeem te ontwikkelen waardoor de overige statistische programmatuur kon worden aangeboden, gedocumenteerd en gescreend. Hiervoor is in alle gevallen gekozen voor een programmabibliotheek. In eerste instantie werd daar-

bij vaak gestreefd naar een standaardisatie van de invoer, zodat de gebruiker zich slechts één stuurtaal eigen hoefde te maken. Zo combineerde het Technisch Centrum de conversie van haar programma's van de X1/X8 naar de CDC 6500 met uniformering van de stuurtaal, naar het model van OSIRIS. Vervolgens ging het Technisch Centrum nog één stap verder toen zij het plan opvatte om de belangrijkste TC programma's als los aanhangsel op SPSS aan te bieden, zodat er met het belangrijkste algemene programmapakket een eenheid in datastructuur en stuurtaal zou ontstaan. Eind 1973 werd met de ontwikkeling van deze Statistical Appendix op SPSS (STAP) begonnen. De ontwikkeling daarvan zou veel langer duren dan gepland, zodat STAP pas in 1981 beschikbaar werd op een moment dat SPSS zelf reeds was uitgebreid met een aantal in STAP opgenomen technieken. Als een tweede voorbeeld van een poging tot standaardisatie zou ik de in 1977 gestarte omzetting van in Nijmegen ontwikkelde software in KUNST (Katholieke Universiteit Nijmegen Statistische Toepassingen) willen noemen. Ook daarvan liep de ontwikkeling trager dan gepland en werd de noodzaak van KUNST grotendeels achterhaald door een toenevende accentverschuiving van locale programmatuur naar elders ontwikkelde programmatuur. Tussen de eigen ontwikkelde programmatuur en de algemene pakketten schoof zich namelijk een hele reeks van min of meer op een grote groep van gebruikers gerichte programma's rond nieuwe technieken en modellen. Deze programma's werden veelal ontwikkeld en gedistribueerd door teams van onderzoekers en programmeurs rond de ontwikkelaar van de techniek of werden na een periode van interne ontwikkeling overgedragen aan commerciële bureaus, omdat onderhoud, aanpassing, conversie en distributie een te grote aanslag op de tijd van de onderzoekers ging leggen. Voorbeelden van deze "speciale" programma(pakketten) zijn LISREL, ECTA, GLIM, CLUSTAN, CADA, PML, MULTIVARIANCE en het MDS pakket van Coxon. Ook in Nederland zijn voorbeelden van dergelijke teams en programmaontwikkeling te vinden. Zo wordt een hele reeks van programma's ontwikkeld en gedistribueerd door een team van onderzoekers rond Prof.dr. J. de Leeuw (Leiden) die gebaseerd zijn op een door De Leeuw ontwikkeld algoritme op het gebied van meerdimensionale schaalanalyse (Alternative Least Squares): HOMALS, PRINCALS, CANALS etc. Het meest gevraagde en unieke onderdeel van STAP is het door Mokken ontwikkelde cumulatieve schaalmodel met de daarbij behorende schaalprocedures (MOKKEN SCALE en MOKKEN TEST). Vanwege de conversieproblemen van STAP naar andere typen computers worden deze programma's ook afzonderlijk gedistribueerd. Een derde voorbeeld is het programmapakket GRADAP (Graph Definition and Analysis Package) dat door een interuniversitair team van onderzoekers en programmeurs van de universiteiten van Amsterdam, Groningen en Nijmegen

is ontwikkeld (Sprenger en Stokman, 1982) ten behoeve van sociale netwerkanalyse. Enerzijds werd daarbij voortgebouwd op een programmapakket dat eerder door Jac.M. Anthonisse van het Mathematisch Centrum was ontwikkeld (Graphlib), terwijl anderzijds via een data interface en stuurtaal nauwe aansluiting met SPSS werd gezocht. Een vierde voorbeeld is het programma FISHER dat door Prof.dr. A. Verbeek is ontwikkeld, waarin voor kruistabellen met kleine aantallen de exacte verdeling wordt berekend in plaats van de gebruikelijke benaderingen, die voor grote aantallen gebruikt mogen worden. Uit het wat verdere verleden kan de inbreng van Prof.dr. E.E.Ch.I. Roskam op het gebied van meerdimensionale schaalanalyse worden genoemd (MINISSA, MINIRSA, MINICPA) en de contrastgroepenanalyse vanuit Tilburg, later uit het gezichtsveld verdwenen door de ontwikkeling van AID (Automatic Interaction Detection). Ook op dit moment zijn een aantal projecten aan de gang die naar verwachting tot op distributie gerichte programmaontwikkeling t.b.v. specifieke methoden kunnen leiden. Ik denk hierbij met name aan een aantal onderzoeksgroepen die vanuit de NOSMO (Nederlandse Organisatie voor Sociaal-Methodologisch Onderzoek) en de Stichting Psychometrie zijn opgericht op het gebied van de Raschanalyse, preferentieanalyse, magnitude estimation en inhoudsanalyse.

Vanuit deze ontwikkelingen, waarin voor academische centra voor data-analyse aanschaf van programmatuur een steeds belangrijkere plaats in gaat nemen boven ontwikkeling van eigen programmatuur, is het niet verwonderlijk dat in de tweede helft van de jaren '70 een aantal nieuwe samenwerkingsverbanden ontstaan die op evaluatie en coördinatie van aanschaf van programmatuur zijn gericht. Zo wordt op 8 april 1976 de Contactgroep Statistische Programmatuur (CSP) opgericht als opvolger van de in 1970 overleden Werkgroep Statistiek en Computer. In tegenstelling tot de WSC heeft de CSP het vervaardigen van programmatuur niet in haar vaandel staan, maar slechts inventarisatie, uitwisseling van evaluatie (het opstellen van kwaliteitsrapporten). Het gelukt nu wel een aantal evaluatierapporten af te ronden op het gebied van clusteranalyse (Van Meurs, 1978), meerdimensionale schaalanalyse (Brouwer *et.al.* 1980), defaults (Van Nes, 1978) en behandeling van missing data (Van der Sluis, 1981), maar ook nu is een structurele aanpak van de evaluatie van beschikbare programmatuur te weinig van de grond gekomen (Van Meurs en Van der Weele, 1982). De CSP noch haar opvolger, de SSP (Sectie Statistisch Programmatuur), lijken in de verste verte op zoiets als een Nederlandse consumentenbond voor statistische programmatuur. Toch vereist het brede gebruik van programmatuur door een groot aantal onderzoekers, die niet in staat zijn de technische aspecten van de programmatuur te evalueren en de juistheid van de uitkomsten te controleren, een dergelijk

forum, zoals Van der Vorst terecht bepleit en met voorbeelden adstrueert (Van der Vorst, 1982). Naast de CSP (en vervolgens SSP) wordt de samenwerking geïntensiveerd tussen universiteiten die dezelfde configuratie bezitten. Het meest intensief is dat tussen de vier universiteiten met de CDC-configuratie, die hun aanschaf en conversie vanaf 5 april 1978 coördineren in de VUGA\* (VU, RUU, RUG en UvA). Deze samenwerking is aanzienlijk intensiever dan die tussen Leiden en Nijmegen die beide een IBM-configuratie hebben. Door de ICL-configuratie in Tilburg en de Burroughs in Eindhoven zijn deze hogescholen veel sterker dan de andere universiteiten een eigen weg ingeslagen en nemen zij in landelijke overleggroepen een marginale plaats in.

## 2. Stand van zaken anno 1983

De bovenstaande ontwikkelingen samenvattend en complementierend ontstaat het volgende beeld met betrekking tot statistical software in de sociale wetenschappen:

1. Voor documentatie en manipulatie van data en verrichting van standaardanalyses bestaan een beperkt aantal algemene programmapakketten die door commerciële bureaus worden ontworpen en gedistribueerd. Naast SPSS zijn gelukkig een aantal alternatieven ontstaan, zoals BMDP, GENSTAT, P-STAT, SAS en MINITAB. (Voor IBM-configuraties is SAS een zo aantrekkelijk alternatief voor SPSS dat zij haar monopoliepositie wel eens zou kunnen gaan verliezen.) Met name rond SPSS is een familie van programma's ontstaan die via data interfaces en gemeenschappelijke stuurtaal daar nauw mee verbonden zijn, zoals SIR (t.b.v. archivering en hiërarchische datastructuren), STAP, GRADAP en RIQS (waarvan de support helaas is gestopt). Op het gebied van algemene programmapakketten is geen Nederlandse inbreng te verwachten.
2. Rond specifieke methoden zijn speciale programmapakketten ontstaan, ontwikkeld door teams van onderzoekers en programmeurs, waarvan de distributie na verloop van tijd veelal wordt overgedragen aan commerciële instanties. Op dit terrein is er in Nederland wel degelijk een eigen inbreng geweest en nog te verwachten.
3. Eigen ontwikkelde, niet op distributie gerichte programmatuur, veelal na enige screening in een programmabibliotheek opgenomen.
4. Voor de eigen ontwikkelingen sub 2 en 3 is de beschikbaarheid van numerieke en statistische procedures van belang. Ook in deze heeft Nederland een eigen bijdrage geleverd, namelijk via het Mathematisch Centrum, met name voor wat de numerieke programmatheken betreft (NUMAL en NUMPAS); daarnaast heeft het Mathematisch Centrum eigen statistische programmathe-

VUGA\*moet zijn FUGA (Federatie Utrecht Groningen Amsterdam). De FUGA is een samenwerkingsverband tussen het Centrum voor Data-Analyse, de LIS-TOR-groep van de RUG en het Technisch Centrum; twee medewerkers van de VU en één medewerker van IWIS/TNO wonen de vergaderingen bij.



ken STATAL (Algol 60) en STAR (Fortran) met een groot aantal statistische procedures.

5. Door deze ontwikkelingen is de taak van centra voor data-analyse verschoven van zelfstandige programmering en verwerking naar acquisitie, documentatie en consultatie. Naast consultatie aan een zeer kleine groep van sterk gespecialiseerde sociale wetenschappers en statistici die zelf aan modelontwikkeling en programmering doen, richt de consultatie zich vooral op de grote groep van gebruikers. De grootste groep daarvan bestaat ongetwijfeld uit personen die uitsluitend gebruik maken van een algemeen pakket, vaak in het kader van verplicht statistiek- of methodologieonderwijs. Daarnaast bestaat er eveneens een kleine groep van methodologisch geschoolde gebruikers die voldoende van het aanbod van 100 tot 250 programma's afweet om daar een gepast gebruik van te maken. Het meest problematisch is de groep van gebruikers die zonder adequate kennis van het aanbod toch behoefte heeft aan of sterk geholpen zou zijn met meer specifieke analyses. In Groningen bestaan goede ervaringen met opvang via een door gevorderde studenten geleide methodologiewinkel onder verantwoordelijkheid van de facultaire vakgroep Statistiek en Meettheorie, die zelf weer terug kunnen vallen op specialisten voor elk der technieken. Goede en betrouwbare verwijskanalen tussen eerste lijnsconsultatie en meer gespecialiseerde consultatie is van wezenlijk belang om onoordeelkundig gebruik of in het geheel geen gebruik van beschikbare programmatuur te voorkomen. Dit is des te meer van belang omdat een aantal technieken via schatting en toetsing van integrale modellen een veel direktere verbinding tussen inhoudelijke theorieën en empirisch onderzoek legt dan tot nu toe mogelijk was (Stokman, 1983). Zonder dergelijke betrouwbare verwijskanalen is interactieve consultatie, zoals TCINFO die beoogt, alleen maar gevaarlijk; met dergelijke verwijskanalen kan hij een nuttig hulpmiddel zijn (Bakker et.al., 1982; zie ook Brouwer, 1980).
6. Naast individuele consultatie is kennisoverdracht via cursorisch onderwijs een belangrijk hulpmiddel om een verantwoord gebruik van statistische programmatuur te verkrijgen. Buiten de colleges methodologie/methoden en technieken/statistiek worden vanuit de centra voor data-analyse c.q. het rekencentrum veelal korte cursussen verzorgd die primair gericht zijn op het leren omgaan met de computer en aanwezige programmatuur. In toenemende mate is de behoefte ontstaan dit onderwijs uit te bouwen in de richting van en te onderbouwen vanuit Sociaal-Wetenschappelijke Informatica (SWI). Met name in Amsterdam en Groningen, maar ook elders (Leiden en Utrecht) wordt hard gewerkt aan uitbouw van Sociaal-Wetenschappelijke Informatica en de

inbedding van Statistical Computing daarin (Brouwer et.al., 1982). Uitbouw van en inbedding in SWI lijkt mij van wezenlijk belang voor ontwikkeling en evaluatie van statistische programmatuur alsmede voor een kritisch gebruik daarvan door een breed scala van onderzoekers.

### 3. Beleidsaanbevelingen

Uit het bovenstaande overzicht blijkt dat de Nederlandse inbreng in de ontwikkeling van speciale, op een brede groep van gebruikers gerichte software pakketten rond eigen ontwikkelde technieken en modellen niet onbeduidend is geweest. Gezien de kracht van een aantal methodologische kernen in Nederland mag verwacht worden dat deze inbreng, althans vanuit de methodologieontwikkeling, niet tot het verleden behoort. Toch moet gevreesd worden dat deze Nederlandse inbreng verloren zal gaan, wanneer op dit terrein geen landelijke infrastructuur en speerpuntenbeleid wordt ontwikkeld, waarbij de ontwikkeling van statistische programmatuur uit de sfeer van hobbyisme wordt gehaald en erkend wordt als wetenschappelijke output. Stelde Van Reeken in 1970 nog:

"The experience is that statistical programmers live under constant pressure of their work. It is a vicious circle; because they are busy, they are not able to co-operate, and because they are not able to co-operate, they will always be busy" (Van Reeken, 1970, 77).

Op dit moment geldt dat zonder landelijke samenwerking statistische programmaontwikkelaars spoedig geen ontwikkelingswerk meer zullen doen, omdat dat het einde van zelfstandige, op distributie gerichte programmaontwikkeling in Nederland zal betekenen. Ik zal daarom nu een programma van eisen formuleren die vervuld moeten worden om de Nederlandse inbreng in deze te continueren. Realisering van die infrastructuur schept tevens de gelegenheid tot een kritischer aankoopbeleid en meer systematische evaluatie van statistische software. Ook maakt zij een betere afstemming tussen technische mogelijkheden en software ontwikkeling mogelijk. Terecht constateert Van de Wijngaart (1981) dat de ontwikkeling van programmatuur meer en meer achterloopt bij die van apparatuur en dat nog nauwelijks gebruik wordt gemaakt van een aantal technische mogelijkheden als interactief gebruik en gebruik van micro's, hetzij stand-alone hetzij als intelligente terminal op een centrale computer.

Het programma van eisen luidt:

1. Vorming van interdisciplinaire teams en oprichting van een landelijk centrum voor ontwikkeling van statistische software. Ontwikkeling van statistische software, niet voor zichzelf maar voor een grote groep van gebruikers, is slechts mogelijk vanuit een team, waarin toegepaste informatici, statistici en sociale wetenschappers samenwerken. De ontwikkeling

van GRADAP is aan een dergelijke hechte interuniversitaire samenwerking tussen methodologen, wiskundigen, statistici, toegepaste informatici en programmeurs te danken geweest, waarbij de verantwoordelijkheden (inclusief die van de projectleiding) duidelijk geregeld waren. Dit laatste vereist dat de uitvoering, i.c. de programmering, zoveel mogelijk geconcentreerd moet worden op één plaats. De locale centra voor data-analyse hebben te weinig formatie en know-how om een dergelijke taak op zich te nemen. Wij zien dan ook in vrijwel alle universitaire centra een toenemend accent op aankoop, onderhoud, aanpassing en consultatie, hetgeen onder druk van de toenemende bezuinigingen alleen maar zal toenemen. Daarom ligt een landelijk centrum voor de hand, dat wel organisatorisch nauw kan worden verbonden met een der lokale centra. Anderzijds zal de methodologische inbreng van verschillende universiteiten moeten komen, omdat geen universitaire kern van methodologen in staat is zoveel methodologische innovatie te realiseren dat alleen van haar uit een konstante stroom van ontwikkelingsprojecten wordt gewaarborgd. Voor een actief en wervend beleid van het landelijke centrum is het van belang dat er een interuniversitair samengestelde projectcommissie wordt gevormd die zowel projektaanvragen en vorming van teams entameert als projecten op haalbaarheid en belang toetst. Nagegaan moet worden in hoeverre het landelijk centrum ook als transferpunt voor het bedrijfsleven kan fungeren, waarbij ook projecten, vandaaruit geformuleerd, kunnen worden aangenomen.

2. Vanuit Sociaal-Wetenschappelijke Informatica opstelling van een pakket van eisen waaraan dergelijke statistische pakketten dienen te voldoen.

Hierboven is reedsesignaleerd dat de ontwikkeling van software meer en meer achterloopt bij die van hardware. Vanuit informatica en Sociaal-Wetenschappelijke Informatica zijn algemene eisen te formuleren waaraan dergelijke pakketten moeten voldoen en waarbij ook mogelijkheden voor ontwikkeling op decentrale computers uitdrukkelijk moet worden nagegaan. Als meest belangrijke eisen zou ik willen noemen:

- a. portabiliteit. Hoewel GRADAP op het gebied van sociale netwerkanalyse één uniek pakket is, wordt de distributie, m.n. naar de Verenigde Staten, ernstig belemmerd doordat bij de ontwikkeling optimalisering voor de CDC-configuratie een hogere prioriteit heeft gekregen dan portabiliteit. Hierdoor wordt voor ander configuraties GRADAP tenminste twee jaar achterop gezet, waardoor GRADAP wellicht tegen die tijd wel een concurrerend pakket zal ontmoeten. Voor portabiliteit van statistische software dient een algemene oplossing gevonden te worden, bijvoorbeeld door constructie van een aparte bibliotheek van subroutines die sy-

steemafhankelijk zijn en in vrijwel alle software pakketten gebruikt worden.

- b. aansluiting bij een algemeen statistisch pakket via data interfaces en applicatietaal. Hierdoor worden aaneensluitende analyses met verschillende technieken en modellen vergemakkelijkt, wordt voorkomen dat data-manipulatiemogelijkheden, die via algemene pakketten beschikbaar zijn, opnieuw geprogrammeerd moeten worden en wordt documentatie van ingevoerde data voor gebruikers aanzienlijk vereenvoudigd. Bij GRADAP is, ondanks de afwijkende datastructuur van netwerkgegevens, gekozen voor aansluiting bij SPSS, waarbij via een data-interface met SPSS zowel informatie van de punten als van de lijnen van een netwerk naar een SPSS systeem file geschreven als daarvan ingevoerd kan worden. Bovendien sluit de applicatietaal van GRADAP nauw aan bij die van SPSS.
- c. interactief. Statistische software is tot nu toe te sterk op batchverwerking gericht. Interactieve sessies van gebruikers worden meestal gebruikt om batch-jobs op te bouwen, te verzenden en eventueel uitvoer te inspecteren alvorens deze naar de printer te sturen. Interactieve algemene pakketten beginnen daadwerkelijk geleverd te worden (SCSS, SAS, MINITAB). Interactieve specifieke pakketten zijn er vrijwel niet, op een enkele uitzondering, zoals CADA, na, maar kunnen wel extra analyse-mogelijkheden voor de gebruikers opleveren, vooral in explorerende analyses. Zo zou bijvoorbeeld een echt interactief programma voor één-dimensionale schaalanalyse, waarin zowel het Rasch- als Mokkenmodel zou zijn opgenomen met de daarbij behorende procedures voor itemselectie, de vaststelling van de definitieve schalen aanzienlijk kunnen versnellen en optimaliseren.
- d. gericht op gebruikers op verschillende niveaus van kennis en ervaring. Vooral bij interactieve programma's is het mogelijk verschillende niveaus van kennis en ervaring van gebruikers in te bouwen door informatie over de technieken en de betekenis van resultaten aan te passen aan een opgegeven kennisniveau van de gebruiker. Vele fouten, soms zeer wezenlijke, ontstaan doordat:
  - \* pakketten onvoldoende informatie verschaffen over het uiteindelijke objekt van analyse. Ongetwijfeld zijn honderden resultaten van SPSS verkeerd geïnterpreteerd en gerapporteerd omdat gebruikers bijvoorbeeld via selecties het objekt van analyse veranderen zonder de daaruit resulterende datafile van een andere naam te hoeven voorzien dan de data-base waaruit deze selecties zijn geschied. Soortgelijke onduidelijkheden bestaan er ten aanzien van in- c.q. uitsluiting van

missende gegevens (Van der Sluis, 1981 en Van der Vorst, 1982).

- \* pakketten onvoldoende informatie verschaffen over de uiteindelijk gekozen analytische procedures. De applicatietaal van pakketten laat vaak onduidelijkheid bestaan over de status van daarin opgenomen begrippen (specificeren zij het object van analyse, de procedure of hebben zij een andere functie zoals documentatie of commentaar?). Door defaults en gebrekkige informatie over de op basis daarvan geselecteerde procedure is het onduidelijk welke procedure uiteindelijk is gekozen. In combinatie met het voorgaande wordt het aan de gebruiker overgelaten duidelijke conclusies te trekken uit een niet volledig gespecificeerde procedure toegepast op een niet volledig gespecificeerd object.
  - \* pakketten geen vooronderzoek verrichten of de data aan de assumpties van het model voldoen en op grond daarvan waarschuwingen aan de gebruiker geven over mogelijke verkeerde interpretaties. Zo zou op basis van het proefschrift van Boomsma (1983) zeer goed waarschuwingen in LISREL in te bouwen zijn bij gebruik van scheef verdeelde variabelen en kleine steekproeven. Op deze wijze zou toch nog, zij het 15 jaar na dato, iets van de ideeën van Van Reeken over gekbestendige programma's worden gerealiseerd.
  - \* pakketten veel niet relevante informatie verschaffen en relevante informatie onoverzichtelijk weergeven, zodat een snel inzicht in de resultaten van de analyse ernstig wordt bemoeilijkt, met name voor niet ervaren gebruikers. Vanuit de ergonomie en Sociaal-Wetenschappelijke Informatica ontstaan nieuwe inzichten die niet of nauwelijks bij ontwikkelaars van programmatuur doordringen (zie bv. Ehrenberg, 1975; Joiner, 1982; Molenaar, 1983).
- e. Optimale mogelijkheden tot locale uitbreidingen en, waar mogelijk en zinvol, macro's voor eigen modelspecificaties door deskundige gebruikers. Op deze wijze voorkomt men rigiditeit en behoeft men als ontwerper niet het onbereikbare doel na te streven om alle mogelijk en zinvolle functies, modellen etc. uit te denken en voor te programmeren.
3. Erkenning van programmatuurontwikkeling als onderzoeksoutput.
- Uit het vorige punt is naar voren gekomen hoezeer programmatuurontwikkeling het resultaat is van wetenschappelijk onderzoek en daarop gebaseerd dient te zijn. Programmaontwikkeling wordt echter niet of nauwelijks als zodanig erkend. Wanneer ik de regels voor het Wetenschappelijk Verslag van de RUG strikt zou hebben geïnterpreteerd, dan zou de 350 pagina's tellende manual van GRADAP (Stokman en Van Veen, 1981) daarin niet opgenomen mogen worden; het pakket GRADAP is nooit in enig Wetenschappelijk Verslag

vermeld. Zoals psychologen regels hebben opgesteld, wanneer testontwikkeling als wetenschappelijk product mag worden meegeteld, zo zouden ook regels moeten worden ontwikkeld wanneer programmaontwikkeling als wetenschappelijk product mag worden verantwoord. Als een sprekend voorbeeld van de onderwaardering voor distributie van software zou ik de opvatting van de beheerscommissie van het Technisch Centrum willen vermelden volgens welke aan het portabel maken van STAP en GRADAP een lage prioriteit moet worden toegekend, omdat dat voor de dienstverlening binnen Amsterdam irrelevant is. Distributie van eigen statistische pakketten zou in mijn ogen een even hoge prioriteit moeten krijgen als het publiceren van artikelen in nationale en internationale tijdschriften.

4. Transfer van statistische software ten behoeve van commerciële distributie. Zelden zal een statistische techniek zulk een verspreiding krijgen dat ontwikkeling van de statistische programmatuur aan de afnemers doorberekend kan worden. Zo een stringent commercieel criterium zou ook belemmerend werken op ontwikkeling van belangrijke doch sterk gespecialiseerde technieken. Wel dient de distributie van gereedgekomen programmatuur op een professionele wijze te worden georganiseerd, waarbij de kosten van onderhoud en distributie kunnen worden doorberekend. Dit zou kunnen worden ingepast in het ministeriële beleid ten aanzien van programmatuurontwikkeling zoals dat recentelijk in het Eindrapport van de Werkgroep Programmatuurontwikkeling van het Ministerie van Onderwijs en Wetenschappen (Programmatuur naar een hoger plan, 1982) is neergelegd.
5. Intensivering van landelijke samenwerking ten behoeve van evaluatie, aanschaf en verkenning van ontwikkelingen op het gebied van statistische software, ondersteund vanuit het landelijke centrum en kernen op het gebied van Sociaal-Wetenschappelijke Informatica.

Reeds eerder is gesteld dat ontwikkeling van software achterblijft bij die van hardware. Adekwate screening van nieuwe ontwikkelingen, met name ook die rond micro-computers, is niet alleen voor de eigen ontwikkeling van software, maar ook voor een doelmatige aanschaf en evaluatie van software van het grootste belang. Met name rond micro-computers dreigt zich dezelfde situatie te herhalen die hierboven voor de zestiger jaren is weergegeven voor de centrale computers, waarbij ieder voor zich ad-hoc programmatuur zit te ontwikkelen die op korte termijn door ontwikkelingen van buitenaf is achterhaald.

Men kan zich afvragen hoe iemand het in zijn hoofd haalt anno 1983 dergelijke voorstellen te doen als ik hierboven heb gedaan. Ik ben echter van

mening dat bovengenoemde voorstellen een voldoende kader scheppen voor een zinvolle besteding van één van de zes miljoen gulden, die de Minister van Onderwijs en Wetenschappen heeft uitgetrokken ten behoeve van fundamenteel onderzoek in de sociale wetenschappen. Zonder deze infrastructuur zal er wederom een terrein bijkomen waarop Nederland geen eigen inbreng meer heeft.

#### Dankwoord

Het overzicht van de ontwikkelingen op het gebied van statistische software kon mede tot stand komen door een aantal gesprekken die ik hierover heb gevoerd en de documentatie die mij daarbij werd overhandigd. Ik dank hiervoor met name drs. E. Brouwer en C. van de Wijngaart (TC-UvA), drs. A. de Jong en Drs. G.P. van der Vorst (Centrum voor data-analyse-RUU), drs. B.F. Schriever en R.D. Gill (MC), L. van der Weele (Rekencentrum RUG) en Prof.dr. A.J.A. Felling (KUN).

#### Referenties

Bakker, D., U. Brouwer en P. Debets, 1982

Automatisering van de voorlichting inzake programmatuur, Kwantitatieve Methoden, 3, nr. 5, 92-115.

Boomsma, Anne, 1983

On the robustness of LISREL (maximum likelihood estimation) against small sample size and non-normality, Proefschrift Rijksuniversiteit te Groningen.

Brouwer, Ulbe, 1980

Consultatie m.b.t. statistische programmatuur, TC-Rapport 80-2, 3-9.

Brouwer, U., P. Debets en W. van Schuur, 1980

Multidimensional Scaling: Some evaluations concerning the distance model in MDS programs, using symmetric input matrices, Amsterdam: Technisch Centrum.

Brouwer, Ulbe, Gerrit van der Vorst en Leo van der Weele, 1982

Onderwijs in "Statistical Computatjon" binnen de sociale wetenschappen, TC-Rapport 82-1, 32-55.

Ehrenberg, A.S.C., 1975

Data reduction. Analysing and Interpreting Statistical Data, New York: Wiley.

Frane, James W., 1983

Book review of STATCAT, Computational Statistics and Data Analysis, 1, no. 1, 70-71.

Joiner, B.L., 1982

The Frontiers of Statistical Analysis. H. Caussinus, P. Ettinger en R. Tomassone (eds.), COMPSTAT 1982, Wien: Physica-Verlag, 278-281.

Meurs, A. van (red.), 1978.

Clusteranalyse, Utrecht: Geografisch Instituut.

Meurs, Ton van en Leo van der Weele, 1982

Naar een Sectie Statistische Programmatuur. De geschiedenis van de WSC en de CSP, Kwantitatieve Methoden, 3, nr. 5, 82-91.

Molenaar, I.W., 1983

Gedragswetenschap en softwarebeleid, Symposium Statistische Software 1983, Utrecht.

Nes, Floor van (red.), 1978

Defaults, een zegen of een ramp? Methoden en Data Nieuwsbrief van de sociaal-wetenschappelijke sectie van de VVS, 3, nr. 3, 63-76.

Nijdam, B.M. en H. van Buuren, 1980

Statistiek voor de sociale wetenschappen, Alphen a/d Rijn: Samsom.

Programmatuur naar een hoger plan, 1982

Eindrapport van de werkgroep programmatuurontwikkeling, Ministerie van Onderwijs en Wetenschappen.

Reeken, A.J. van, 1970

Report on the Work of the Dutch Working Party on Statistical Computing, Statistical Computing, 73-79.

Sluis, D.M. van der, 1981

Statistical Programs and Missing Data, Kwantitatieve Methoden, 2, nr. 2, 56.

Sprenger, C.J.A. en Stokman, F.N., 1982

Applied Graph Analysis in the Social Sciences: The Software Project GRADAP. H. Caussinus, P. Ettlinger en R. Tomassone (eds.), COMPSTAT 1982, Wien: Physica-Verlag, 113-120.

Stokman, F.N., 1983

Het meten en analyseren met modellen in sociologisch onderzoek. S. Lindenberg en F.N. Stokman (eds.), Modellen in de Sociologie, Van Loghum Slaterus, Deventer, 177-194.

Stokman, F.N. en Van Veen, F., 1981

GRADAP GRAPh Definition and Analysis Package User's Manual, Inter-university project-group GRADAP, Universiteit van Amsterdam.

Vorst, G.P. van der, 1982

De computer als black box, Intermediair, 18, nr. 51, 1-5.

Wijgaart, C. van de, 1978

Tien jaar TC, TC-rapport 78-3, 3-16.

Wijgaart, C. van de, 1981

Statistische programmatuur en computers: Enkele ontwikkelingen, Kwantitatieve Methoden, 2, nr. 3, 90

WESP, Waarlijk Eenvoudig Statistisch Pakket, Methoden en Nieuwsbrief van de sociaal-wetenschappelijke sectie van de VVS, nr. 3, 77-81.