

STATOR

periodiek van de VVS jaargang 10 nummer 4, december 2009

**De wereld geoptimaliseerd!
Waarom Operations Research essentieel
is bij voedselhulp**

Structurele Vergelijkingen Modellen

**OR in de energiesector: een praktische
toepassing voor elektriciteitsnetbeheerders**

**Optimale onderzoeksdesigns:
meer efficiëntie voor minder geld**

STATOR

Jaargang 10, nummer 4, december 2009

STATOR is een uitgave van de Vereniging voor Statistiek en Operationele Research (VVS). STATOR wil leden, bedrijven en overige geïnteresseerden op de hoogte houden van ontwikkelingen en nieuws over toepassingen van statistiek en operationele research. Verschijnt 4 keer per jaar.

Redactie

Goos Kant (hoofdredacteur), Ana Isabel Barros, Mirjam Moerbeek, Gerrit Stemerding (eindredacteur), Hilde Tobi, Marnix Zoutenbier. Vaste medewerkers: Johan van Leeuwen, Fred Steutel

Kopij en reacties richten aan

Prof. dr. G. Kant (hoofdredacteur), Faculteit der Economische Wetenschappen van de Universiteit van Tilburg, Postbus 90153, 5000 LE Tilburg, telefoon 013 - 4668234, mobiel 06-11045089, <G.Kant@uvt.nl>.

Bestuur van de VVS

Voorzitter: prof. dr. R. Gill <voorzitter@vvs-or.nl>
Secretaris: dr. C.G.H. Diks <c.g.h.diks@uva.nl>
Penningmeester: prof. dr. ir. C.A.G.M. van Montfort <kvmontfort@feweb.vu.nl>
Statistische dag: prof. dr. A.W. van der Vaart <aad@cs.vu.nl>
Namens de Bedrijfssectie (BDS):
prof. dr. R.J.M.M. Does <R.J.M.M.Does@uva.nl>
Namens de Biometrische Sectie (BMS):
prof. dr. A.H. Zwinderman <a.h.zwinderman@amc.uva.nl>
Namens de Economische Sectie (ECS):
dr. P.H.F.M. van Casteren <casteren@fee.uva.nl>
Namens het Ned. Genootschap voor Besliskunde (NGB):
prof. dr. ir. C.P.M. van Hoesel <s.vanhoesel@ke.unimaas.nl>
Namens de Sectie Mathematische Statistiek (SMS):
dr. P.J.C. Spreij <spreij@science.uva.nl>
Namens de Sociaal Wetenschappelijke Sectie (SWS):
prof. dr. J.K. Vermunt <j.k.vermunt@uvt.nl>

Leden- en abonnementenadministratie van de VVS

VVS, Postbus 244, 6700 AE Wageningen, telefoon 0317 - 419572, fax 0317 - 421364, <admin@vvs-or.nl>. Raadpleeg onze website over hoe u lid kunt worden van de VVS of een abonnement kunt nemen op STATOR of op een van de andere periodieken.

VVS-website

<http://www.vvs-or.nl>

Advertentieacquisitie

Marieke Klein, p/a Vrije Universiteit, afdeling Econometrie & Operationele Research, De Boelelaan 1105, 1085 HV Amsterdam, <adverteren.stator@vvs-or.nl>. STATOR verschijnt in maart, juni, september en december.

Ontwerp en opmaak

Pharos | M. van Hootegem, Nijmegen

Uitgever

© Vereniging voor Statistiek en Operationele Research
ISSN 1567-3383

Inhoud

3 Redactioneel

4 De wereld geoptimaliseerd! Waarom Operations Research essentieel is bij voedselhulp

Tom Plat, Timon van Dijk en John Poppelaars

9 Structurele Vergelijkingen Modellen

Remco Feskens

13 Agenda & VVS-OR-nieuws

14 OR in de energie sector: een praktische toepassing voor elektriciteitsnetbeheerders

Sonja Bouwman en Gabriël Bloemhof

20 Dromen van zomerscholen – column

Johan van Leeuwen

26 Optimale onderzoeksdesigns: meer efficiëntie voor minder geld

Martijn P.F. Berger

27 Voertaal: gebroken Engels – column

Fred Steutel

29 Emotionele hokjes. Mental Accounting: liever overzichtelijk dan kostenefficiënt

30 IM Jaap Wessels



WERELDPROBLEMATIEK

In de donkere dagen tussen Sinterklaas en Kerst schrijf ik deze inleiding. Buiten is het crisis, in Kopenhagen proberen meer dan 190 landen afspraken te maken over het klimaat, in Nederland waart de Q-koorts en de Mexicaanse griep rond en verschijnen er diverse toplijstjes als terugblik op het eerste decennium van de 21ste eeuw. Voldoende redenen dus om weg te duiken in *STAtOR*, sinds dit decennium elk kwartaal de vaste rots in de branding.

Bij de klimaattop gaat het vooral om het verminderen van de CO₂-uitstoot, ofwel hoe we slimmer met energie kunnen omgaan. Bouwman en Bloemhof hebben onderzocht hoe we in Nederland slimmer de energie-apparatuur, zoals transformatoren, kunnen onderhouden. Ze sluiten af met de conclusie dat OR erg bruikbaar is in de energiesector. Dit wisten wij en u natuurlijk al. Maar mocht dit bericht de klimaatdeskundigen te laat bereiken, dan hebben we nog altijd het artikel uit *STAtOR* 2007-4 hoe we onze dijken moeten gaan ophogen.

Met de huidige crisis is een artikel als 'Meer efficiëntie voor minder geld' van Berger natuurlijk altijd welkom. Maar de wereldproblematiek is groter. Iemand van het wereldvoedselprogramma van de VN vroeg mij eens waarom OR niet simpel de voedselverdeling in de wereld kon oplossen:

zet het aanbod links, de vraag rechts, inventariseer de onderlinge kosten, en los het toewijzingsprobleem op. Tsja. In deze *STAtOR* alvast een aanzet over het gebruik van OR bij voedselverdeling in ontwikkelingslanden.

Hergebruik is ook altijd goed in tijden van crisis. Hoewel, het artikel over *Structural Equation Models* mag amper onder deze noemer vallen. Feskens presenteert hierin een hedendaagse versie van een techniek die al zo'n 40 jaar succesvol wordt toegepast, maar in de 'oude tijden' enigszins gehinderd werd door de complexiteit van de toenmalige software.

Kortom, wij schuwen de grotere vragen niet. Maar er blijven nog vele uitdagingen over voor het komende decennium. Wat te denken van het optimaliseren van bankiersbonussen of het verbeteren van verdraagzaamheid in de wereld? Johan van Leeuwen droomt hartje winter al van prachtige *Summer Schools* om met OR en statistische technieken hiermee aan de slag te gaan. En als het aan onze Fred Steutel ligt graag niet in gebroken Engels, dus heet het daarom in dit nummer voor de zekerheid ook zomerscholen.

Veel leesplezier toegewenst.

De redactie



DE WERELD GEOPTIMALISEERD!

Waarom Operations Research essentieel is bij voedselhulp

TOM PLAT, TIMON VAN DIJK EN JOHN POPPELAARS

Het thema van het vorige nummer van *STATOR* was voedsel. Voor ons is voedsel meestal niet een thema waar we veel over nadenken. We lopen de supermarkt binnen en kiezen een van de vele varianten van hetzelfde product. Tel het aantal broodsoorten maar eens! Als we al bewust aan voedsel denken is dat vanwege veiligheid of gezondheid. Voedsel wordt door ons vooral geassocieerd met gezelligheid, zoals tijdens het kerstdiner. Voor velen op de wereld is het echter anders, voedsel is daar prioriteit 1. Voedsel staat daar gelijk aan overleven. Ook dit jaar spreken krantenkoppen voor zich: 'Droogte heeft een voedselcrisis veroorzaakt in Ethiopië. Zes miljoen mensen dreigen in hongersnood te geraken.' Dit jaar groeit het aantal

mensen dat chronisch honger lijdt naar 1 miljard, ofwel 1 op de 6 wereldburgers.¹ Een efficiëntere distributie van noodhulp en voedsel betekent dat deze mensen een betere kans krijgen om te overleven.

Het Wereldvoedselprogramma & Optimising the World

De voedselhulp in de derde wereld wordt voor een belangrijk deel uitgevoerd door het Wereldvoedselprogramma (WFP)². Het is de grootste humanitaire organisatie ter wereld en onderdeel van de Verenigde Naties. Het is vertegenwoor-

digd in 78 landen en voorziet jaarlijks ongeveer 87 miljoen mensen van voedsel, waarvan 58 miljoen kinderen. In 2008 is 2,8 miljoen ton voedsel gedistribueerd met een totale waarde van US\$ 1,4 miljard. Voor het overgrote deel betreft het steun bij conflicten en in rampgebieden. Daarnaast is WFP actief in de distributie van voedsel aan de allerarmsten. Hierbij is de doelstelling vooral om kinderen en jonge moeders van de noodzakelijke hulp te voorzien. Door diverse programma's worden mensen gestimuleerd om iets te doen voor de samenleving in ruil voor voedsel.

Als onderdeel van haar sponsorprogramma *Optimising the World* heeft ORTEC zich tot doel gesteld haar kennis, kunde en werknemers in te zetten voor het WFP. Met de oplossingen voor bijvoorbeeld routing van transportmiddelen kan een betere distributie van voedsel en andere hulpgoederen worden bereikt. Deze betrokkenheid bij het Wereldvoedselprogramma dateert uit 2006 toen er voor het eerst, in samenwerking met TNT's Moving the World programma, een optimalisatieproject in Liberia is uitgevoerd. Doel van het project was de verbetering van de distributie van voedsel aan scholen. Door die verbetering kon de voedselhulp, ondanks forse bezuinigingen, op hetzelfde niveau worden gehouden.

De noodhulp supply chain

Noodhulp organiseren stelt hoge eisen aan de betrouwbaarheid en robuustheid van de keten. Vaak begint die keten in de haven of op de luchthaven in de nabijheid van het rampgebied. Van daaruit wordt de hele distributieketen opgebouwd. Het is immers niet zo dat de noodgebieden altijd op dezelfde plek liggen, er is dus maatwerk vereist, steeds opnieuw. Dit vertoont enige gelijkenis met een militaire invasie.

Een distributieketen voor noodgoederen bestaat vaak uit meerdere tussenopslagplaatsen

(ofwel hubs) verbonden door meerdere routes met soms ook verschillende modaliteiten. Dit om er zeker van te zijn dat, als wegen bijvoorbeeld vanwege de regen of plaatselijke onlusten niet toegankelijk zijn, een tijdje vooruit te kunnen met de voorraad hulpmiddelen of een alternatieve route te hebben.

Soms is er geen tijd om een hele keten op te zetten en wordt met helikopters direct eerste hulp geboden, zoals tijdens de overstromingen in Uganda. Als de keten eenmaal ontworpen is, dan volgt de voorraadstrategie en het ontwerp van de uiteindelijke distributie van de goederen. Zo op het eerste gezicht is deze uitdaging niet anders dan die van producenten in de eerste wereld. Een topdown aanpak vanuit infrastructurele planning naar voorraadstrategie en distributie.

De praktijk

Vorig jaar hebben Timon van Dijk en Tom Plat projecten uitgevoerd voor het WFP. De missie in Ethiopië had tot doel het bestaande distributienetwerk in het grensgebied met Somalië te evalueren en te verbeteren. Doel van de missie in Mali was de optimalisatie van de distributie van noodhulpgoederen in de Mopti en Tombouctou regio.

Ethiopië

Voor het bestrijden van honger heeft het WFP een aantal Targeted Supplementary Feeding Programmes, ofwel TSFP's, opgezet. Zo ook in Ethiopië, meer specifiek in het grensgebied met Somalië. Ieder kwartaal wordt 2.931 ton voedsel gedistribueerd. Het voedsel wordt vanaf de haven van het nabijgelegen Djibouti of vanuit de hoofdstad Addis Abeba via een drietal hubs (Nazret, Dire Dawa en Jijiga) naar de uiteindelijke voedseldistributiepunten (FDP's) gebracht. In totaal zijn er 209 FDP's, die elk één keer per kwartaal worden bevoorrad. In het regenseizoen, dat vier

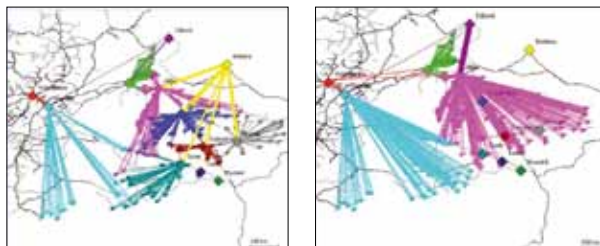


maanden duurt, kan een aantal FDPs niet worden bereikt met grote trucks. Doel van de missie is om de bereikbaarheid van die locaties te verbeteren, bijvoorbeeld door het aantal hubs uit te breiden of een alternatieve verbinding tot stand te brengen.

De kwartaalkosten voor de huidige infrastructuur zijn US\$ 270.000. Per kwartaal werden er 104 trucks ingezet voor het vervoer van in totaal 2.931 ton aan voedsel. Met behulp van een *supply chain* optimalisatie model is de huidige infrastructuur gemodelleerd. Ook zijn alternatieve locaties in het model opgenomen. Het vinden van data voor parameters als rijnsnelheid, capaciteit van de voertuigen en locatie, kosten voor vervoer en opslag en het wegennet is een verhaal op zich. Veel van die gegevens ontbreken en moeten dus met veel tijd en moeite boven tafel worden gehaald. In het

vinden van de beste oplossing kwamen ook een aantal niet te modelleren condities naar voren. Bijvoorbeeld: niet alle chauffeurs zijn bereid de betreffende gebieden waar de FDP's zich bevinden in te rijden, vanwege onlusten tussen verschillende stammen. Soms is die bereidheid er wel, maar alleen met hoge toeslagen.

Met behulp van het model is de meest optimale infrastructuur per seizoen vastgesteld. Belangrijke toevoeging in de nieuwe infrastructuur is het vanuit de haven van Berbera bevoorraden van een aantal hubs, waardoor de FDP's in het uiterste zuiden van Ethiopië ook gedurende het regenseizoen bereikt kunnen worden. Ook is het aantal hubs van 3 naar 7 uitgebreid. Het aantal benodigde trucks daalt. In het droge seizoen zijn er nog 82 nodig, in het regenseizoen stijgt dit naar 131. De nieuwe infrastructuur zorgt ervoor dat de kosten per kwartaal dalen naar US\$ 215.000, een besparing van 21%. Ook is de verdeling van de goederen beter; er is een minder grote afhankelijkheid van de bevoorrading vanuit Addis Abeba. Als gevolg van de lagere kosten kan meer voedsel worden ingekocht en meer mensen van voedsel worden voorzien.



Huidige situatie en (rechts) geoptimaliseerde *supply chain*



Markt aan de Niger in Mopti

Mali

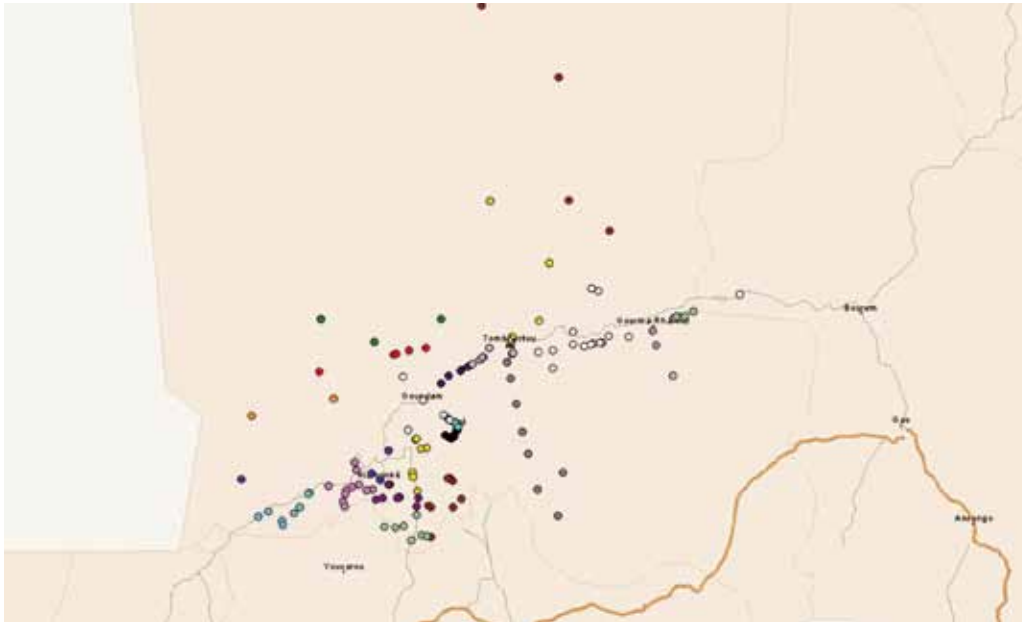
In Mali had de missie de doelstelling de huidige distributie van voedsel naar de FDP's te analyseren en te verbeteren. De missie concentreerde zich op Mopti en Tombouctou, de twee grootste regio's van Mali. Per jaar wordt er 6.300 ton voedsel verdeeld over 6.289 afzonderlijke leveringen. De woestijn en de rivieren zorgen ervoor dat het transport bemoeilijkt wordt in beide regio's, sommige locaties zijn moeilijk bereikbaar.

Naast bereikbaarheid speelt ook het seizoen in Mali een rol. In het regenseizoen worden locaties veel moeilijker bereikbaar. Het WFP in Mali heeft geen eigen wagenpark en huurt lokale transportbedrijven in, Ieder jaar worden de groepen locaties geveild. In iedere regio doen de transportbedrijven biedingen op groepen locaties waarop WFP besluit met welke transporteur ze in zee gaat. Omdat deze prijsbiedingen bepaald worden door allerlei moeilijk te kwantificeren factoren, niet altijd eerlijke handelspraktijken en kartelvorming (wat niet ongebruikelijk is op het platteland van Mali) lijkt het gebruik van Operations Research technieken op het eerste gezicht niet zinvol. De

prijs van de biedingen is echter direct gerelateerd aan het aantal kilometers dat gereden moet worden voor een bepaalde groep locaties. Hierdoor is het minimaliseren van de kilometers over alle groepen locaties een manier om de te verwachten kosten van WFP te minimaliseren. Doelstelling van de missie werd dus het vinden van 'mooie' groepen locaties in de Mopti en Tombouctou regio.

Om mooie groepen locaties te bepalen kun je wel met OR technieken aan de slag. 'Mooie' groepen locaties – ook wel clusters genoemd – moeten voldoen aan de volgende bepalingen:

- De grootste afstand tussen willekeurig twee locaties in een cluster is zo klein mogelijk, voor alle clusters. Hierdoor worden de groepen compact.
- De hoeveelheid locaties in een cluster mag niet teveel verschillen van het gemiddelde aantal locaties in een cluster. Daarbij moet ook het volume dat naar locatie vervoerd worden meegewogen. Hiermee blijft de grootte van clusters in evenwicht, zodat er niet een heleboel volume



Clusters in de regio Tombouctou.

in één cluster komt, wat het biedingproces kan verstoren.

- Goede bereikbaarheid tussen locaties in een cluster en vanuit aanwezige depots zodat mooie 'rondjes' gemaakt kunnen worden.

Gebaseerd op cijfers van 2008 zijn de kosten van de huidige clustering van WFP doorgerekend met een systeem voor de optimalisatie van distributievraagstukken. Voor beide regio's en gegeven het aantal clusters zijn nieuwe indelingen gemaakt. Deze leiden tot geschatte besparingen van 1% (Mopti) tot 3% (Tombouctou) ten opzichte van de huidige kosten.

Operations Research in de derde wereld

De tot nu toe uitgevoerde missies zijn voor ORTEC de eerste op het gebied van de toepassing van Operations Research in ontwikkelingslanden. Zoals uit de beschreven missies blijkt is een ontwikkelingsland een heel andere omgeving dan waarin we normaal werken en dat brengt nogal

wat interessante factoren met zich mee. De vaak toch al betrekkelijke betrouwbaarheid van data (zelfs in ons deel van de wereld) brengt in deze setting een nog grotere mate van onzekerheid. Gebrek aan data is echter niet de enige beperking. In Ethiopië en Mali kwamen een aantal niet te modelleren condities naar voren. Beide missies laten echter zien dat ondanks deze bijzondere factoren een structurele methode zorgt voor een beter inzicht, afwegingen concreet maakt en meetbare resultaten oplevert.

TOM PLAT, TIMON VAN DIJK en JOHN POPPELAARS zijn werkzaam bij ORTEC Logistiek, Divisie Consulting. TOM PLAT is consultant. Hij vertrekt in januari 2010 naar Malawi voor zijn volgende missie: het ontwerpen van een Jatropha supply chain. E-mail <Tom.Plat@ortec.com>. TIMON VAN DIJK is consultant. JOHN POPPELAARS is principal consultant. E-mail: <john.poppelaars@ortec.com>.

NOTEN

1. Zie <www.wfp.org/hunger>.
2. Zie <www.wfp.org> voor meer informatie over het Wereldvoedselprogramma.
3. Zie <<http://john-poppelaars.blogspot.com/2007/10/uganda-flood-response.html>>.

STRUCTURELE VERGELIJKINGEN MODELLEN

In dit artikel worden Structural Equations Models (SEM) besproken. SEM wordt al decennialang toegepast, met in de jaren zeventig van de vorige eeuw een voortrekkersrol voor het softwareprogramma LISREL. Ze bieden steeds meer mogelijkheden tot verschillende analyses. Een structureel vergelijkingen model is een model gebaseerd op een achterliggende theorie, waarbij met name de mogelijkheid om een complexe theorie in een model te analyseren SEM aantrekkelijk maakt.

REMCO FESKENS

SEM (Structural Equations Models oftewel structurele vergelijkingen modellen) is een soort combinatie van traditionele regressie of pad-analyse en factor-analyse. Een SEM-analyse volgt gewoonlijk de volgende stappen: 1. op basis van een theorie wordt een model gemaakt; 2. vervolgens wordt de benodigde data verzameld; 3. de verzamelde data wordt als ruwe data ingevoerd, die door de programmasoftware geconverteerd wordt naar covariantie matrices en gemiddelden; 4. de onderzoeker beoordeelt de model fit en; 5. past het model eventueel aan; 6. daarna dienen de schat-

tingen geëvalueerd te worden; 7. tenslotte dient de onderzoeker een terugkoppeling te maken van de empirische resultaten naar de theorie waar het allemaal mee begonnen is.

Deze stappen zullen later in dit artikel uitgebreider worden besproken en geïllustreerd aan de hand van een praktijkvoorbeeld. Eerst zal ik het specifieke SEM-jargon toelichten. Daarna zal ik, aan de hand van een voorbeeld, de SEM-theorie bespreken. Ten slotte geef ik een overzicht van de software waarmee SEM geanalyseerd kunnen worden.

De taal van SEM

SEM gebruikt een eigen taal en symbolen voor het aanduiden van variabelen en relaties. Onafhankelijke variabelen worden *exogene* variabelen genoemd; afhankelijke variabelen worden *endogene* variabelen genoemd. Variabelen kunnen in tegenstelling tot veel andere statistische procedures gelijktijdig zowel afhankelijk als onafhankelijk zijn. Direct geobserveerde variabelen worden *manifeste* variabelen genoemd en worden in een diagram van het model weergegeven door vierkanten. Niet direct observeerbare inhoudelijke concepten kunnen worden gemeten door een set van manifeste variabelen en worden *latente* variabelen genoemd. Latente variabelen zijn dus factoren en als zodanig vergelijkbaar met de factoren die worden gevonden in een exploratieve factor analyse. Latente variabelen worden in een model weergegeven door ovals of cirkels. Tenslotte, een structureel vergelijkingen model kan bestaan uit een *meet* gedeelte, een *structureel* gedeelte of een combinatie van beide. Het meetgedeelte heeft betrekking op het meten

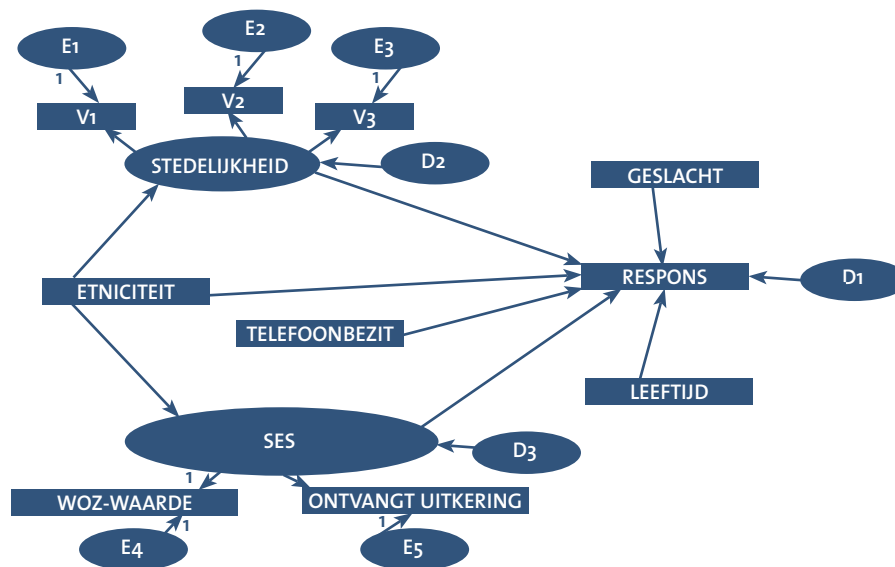
van latente variabelen, het structurele gedeelte heeft betrekking op de voorspelde relaties tussen variabelen.

SEM aan de hand van een voorbeeld

Het voorbeeld is afkomstig uit een eerdere studie naar de non-respons bij een enquêteonderzoek (Feskens et al, 2007). In deze studie wordt nader onderzocht of bovengemiddelde non-responscijfers onder allochtonen toegeschreven kunnen worden aan hun etniciteit of aan andere factoren, zoals stedelijkheid en sociale economische status (SES).

1. De onderzoeker specificeert een model op basis van bestaande theorie.

Dit heeft geresulteerd in het model zoals in figuur 1 is afgebeeld. Het meetgedeelte van dit model bestaat uit de latente variabelen stedelijkheid en SES, die door 3 respectievelijk 2 geobserveerde variabelen worden gemeten. De pijlen van de factoren naar de variabelen zijn de factorladin-



Figuur 1. Structureel vergelijkingen model

gen. De factoren verklaren niet alle variantie in de geobserveerde variabelen, dit wordt gerepresenteerd door een 'residuele error' term en wordt ook wel meetfout genoemd. Deze term wordt ook weergegeven door een cirkel. De paden van de andere geobserveerde variabelen en latente variabelen op de dichtome afhankelijke variabele respons vormen het structurele gedeelte van het model. Om te onderzoeken of de lage respons onder allochtonen toegeschreven kan worden aan etniciteit of andere factoren zijn er ook regressiepaden van etniciteit naar de latente variabelen gemodelleerd. Nu kan dus naast het directe effect van etniciteit naar respons ook een indirect effect van etniciteit naar respons via stedelijkheid en SES gemeten worden. Dit wordt ook wel aangeduid als een mediator effect. Het gedeelte van de variantie in de (afhankelijke) endogene variabelen dat niet verklaard wordt door de (onafhankelijke) exogene variabelen wordt weergegeven door de *disturbances*. Deze worden grafisch ook weergegeven door ovalen.

2. Dataverzameling en 3. Invoer van de data

Een basisvoorwaarde voor SEM is een voldoende grote steekproefomvang. Hierbij wordt uitgegaan van tenminste 15 cases per variabele (Stevens, 1996), waarbij dit aantal groter zou moeten zijn wanneer de data bijvoorbeeld incompleet of scheefverdeeld is. De cijfers die in dit artikel gegeven worden, zijn afkomstig van het Permanent Onderzoek Leefsituatie (POLS) 1998, uitgevoerd door het Centraal Bureau voor de Statistiek (CBS). De steekproefomvang is 39.572.

4. Beoordeling van de model fit

Een voordeel van SEM is dat het veel meer maten voor model fit levert dan bijvoorbeeld traditionele regressieanalyse. Deze fitmaten kunnen worden ingedeeld in tests voor absolute fitmaten en relatieve fitmaten. Een veelgebruikte absolute fitmaat is de Chi-kwadraat test. In dit voorbeeld is de

Chi-kwadraat waarde 125,432 met 4 vrijheidsgraden (zie tabel 1). Dit betekent dat de nulhypothese voor deze test (het model fit aan de data) verworpen wordt. De Chi-kwadraat test is echter gevoelig voor de steekproefomvang. In dit geval zijn ruim 39.000 cases geïncludeerd, dus dit probleem kan hier een serieuze rol spelen. Daarom zijn er voor SEM analyses alternatieve fitmaten ontwikkeld, de zogenaamde relatieve fitmaten. Relatieve fitmaten maken gebruik van het gekozen model en het onafhankelijkheidsmodel, een model waar geen relaties tussen de variabelen zijn gespecificeerd. Er bestaan twee families van relatieve fitmaten die meestal naast elkaar gebruikt worden. Een fitindex geeft de proportie van de verbetering in de totale fit van het gekozen model ten opzichte van het onafhankelijkheidsmodel aan, deze moeten naderen tot 1 om een goede fit aan te tonen. In dit voorbeeld is Bentlers comparative fit index (CFI) gebruikt. De CFI was 0,971 en dus een prima fit. Een maat die het model vergelijkt met het 'ware' model is de root mean square error of approximation (RMSEA). Een lage waarde duidt op een goede fit. In dit geval is de RMSEA = 0,046, dus wederom een goede fit.

5. Eventuele aanpassing van het model

Vaak dient een model aangepast worden, omdat het geen goede fit heeft. SEM geeft hier ook zogenaamde *modificatie indices* (of Lagrange multiplier test) voor. Hiermee wordt aangegeven welke toegevoegde of weggehaalde paden een significante verbetering van de modelfit zullen opleveren.

6. Evaluatie van de schattingen

In tabel 1 zijn de ongestandaardiseerde schattingen en de gestandaardiseerde schattingen (tussen haakjes) weergegeven. De verandering in de afhankelijke variabele door een verandering in de onafhankelijke variabele kan worden afgelezen uit de ongestandaardiseerde regressiecoëfficiën-

	Schattingen op respons	S.E.	Schattingen op stedelijkheid	S.E.	Schattingen op SES	S.E.
Grootte van de gemeente (V1)			1,000 gefixeerd (0,874)	0,000		
Stedelijkheid van gemeente (V2)			0,825** (0,914)	0,011		
Stedelijkheid op postcode niveau (V3)			0,870** (0,887)	0,015		
WOZ waarde					1,000 gefixeerd (0,648)	0,000
Ontvangt uitkering?					-0,370** (-0,491)	0,020
Geslacht	0,075** (0,037)	0,021				
Leeftijd	-0,002** (-0,033)	0,001				
Stedelijkheid	-0,097** (-0,142)	0,011				
SES	0,047** (0,063)	0,018				
Vaste telefoon?	0,341** (0,139)	0,031				
Etniciteit	-0,285** (-0,072)	0,048	1,453** (0,250)	0,054	-1,805** (-0,340)	0,073

Noot: respons, geslacht en etniciteit zijn respectievelijk als volgt gecodeerd 1 = respons, 0 = non-respons; 1 = vrouw, 0 = man; 1= niet-westerse allochtonen 0 = autochtonen & westerse allochtonen.

**p < 0,01

Chi square = 125,432 (df = 4); RMSEA = 0,046; CFI = 0,972, R square = 0,072.

Tabel 1. Resultaten structureel vergelijkings model

ten. De gestandaardiseerde regressiecoëfficiënten kunnen gebruikt worden om de relatieve bijdrage van de onafhankelijke variabelen op de afhankelijke variabele te beoordelen. Uit tabel 1 blijkt dat de relatie tussen etniciteit en respons vrijwel in zijn geheel gemedieerd wordt door met name stedelijkheid. Het directe gestandaardiseerde pad van etniciteit naar respons is -0,072. Het effect van stedelijkheid op respons is -0,142. SES beïnvloedt daarentegen nauwelijks de respons waarschijnlijkheid (0,063).

7. Terugkoppeling naar de theorie

De laatste stap in een SEM-analyse is de terugkoppeling van de empirische onderzoeksresultaten naar de theorie waar het allemaal mee begonnen is. In dit geval is de conclusie: onder allochtonen worden bovengemiddelde non-responscijfers gevonden. Deze hoge non-respons cijfers zijn echter meer het resultaat van andere factoren dan

etniciteit zelf. In het bijzonder de stedelijke woon-situatie van allochtonen verklaart de hoge non-respons onder deze groep. Dit is overigens ook het geval voor de autochtone bevolking: onafhankelijk van de etniciteit, inwoners van stedelijke gebieden responderen slechter dan (autochtone en allochtone) inwoners van niet-stedelijke gebieden.

Software

Er zijn veel software programma's waarin SEM analyses gedraaid kunnen worden. De bekendste zijn LISREL, AMOS en EQS, waarbij AMOS wellicht de meest gebruikersvriendelijke software is. Mplus staat bekend om haar grote flexibiliteit (met name in het gebruik van dichotome afhankelijke variabelen). Mx is een software die gratis gebruikt kan worden.

AGENDA

Conclusie en samenvatting

SEM heeft een aantal voordelen die het de moeite waard maken het te gebruiken. Samengevat zijn dat:

- een variabele kan gelijktijdig gemodelleerd worden als een onafhankelijke en als een afhankelijke variabele;
- latente variabelen kunnen in SEM geïncorporeerd worden;
- meetfouten worden expliciet gemodelleerd;
- SEM geven veel fit indices;
- ten slotte biedt SEM de mogelijkheid tot het flexibel modelleren van meer complexe designs, zoals longitudinale analyses, multiple group analyses en multilevel analyses en kan het goed omgaan met data die niet normaal is verdeeld of missende waarden bevat.

Mocht u meer willen weten over de toepassingen van SEM dan is *Principles and Practice of Structural Equation Modeling* geschreven door Kline (2005) een toegankelijke inleiding. Dit geldt ook voor het al wat oudere artikel van Blok en Saris uit 1980. Er bestaat ook een levendige discussielijst op internet <<http://alabamamaps.ua.edu/cgi-bin/wa?Ao=SEMNET>>, waarbij vele SEM-onderwerpen besproken worden.

LITERATUUR

- Blok, H. en Saris, W.E. (1980). Relevante variabelen bij het doorverwijzen na de lagere school; een structureel model. *Tijdschrift voor Onderwijs Research*, 25,3, 231-248.
- Feskens, R.C.W., Hox, J.J., Lensvelt-Mulders, G.J.L.M., en Schmeets, J.J.G. (2007). Nonresponse among ethnic minorities: A multivariate analysis. *Journal of Official Statistics*, 23, 3, 387-408.
- Kline, R.B. (2005). *Principles and practice of Structural Equation Modeling*. New York: Guilford Press.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Lawrence Erlbaum Publishers.

REMCO FESKENS is senior research fellow bij IQ Healthcare aan het Universitair Medisch Centrum St. Radboud Nijmegen. E-mail: <r.feskens@iq.umcn.nl>

12-14 januari 2010

CONFERENTIE MATHEMATISCHE BESLISKUNDE

De 35-ste Lunteren Conferentie Mathematische Besliskunde, georganiseerd door het LNMB, vindt plaats van 12-14 januari 2010.

De plenaire sprekers zijn: Dimitris Bertsimas (MIT, Cambridge, USA), Jose Blanchet (Columbia University, New York, USA), Anton Kleywegt (Georgia Tech, Atlanta, USA) en Andrea Lodi (University of Bologna, Bologna, Italy).

Tevens zal tijdens deze conferentie een speciale sessie worden georganiseerd ter gelegenheid van het afscheid van de directeur van het LNMB, Lodewijk Kallenberg, die met emeritaat gaat.

Daarnaast is er op 14 januari een seminar over het onderwerp 'Operations Research and Health Care', georganiseerd in samenwerking met het NGB. Voor meer informatie en inschrijving zie <www.lnmb.nl/conferences/lunteren2010>

VVS-OR NIEUWS

Prof. Richard D. Gill, voorzitter van de VVS-OR, is verkozen tot Distinguished Lorentz Fellow (DLF) van het academisch jaar 2010-2011! Meer informatie over het fellowship is te vinden op de website <www.vvs-or.nl>.

Statistische en OR-dag 2010

Op 1 april 2010 zal er weer een Statistische en OR-dag plaatsvinden. De eerste plenaire spreker zal B. Efron zijn. Naar een slotspreker voor de dag wordt nog gezocht. Suggesties aan de voorzitter zijn welkom (gill@math.leidenuniv.nl).

Tevens zal op deze dag de van Dantzigprijs 2010 worden uitgereikt. Het bestuur is bezig een jury te benoemen. Ondertussen kunnen suggesties voor kandidaten voor de prijs ook aan de voorzitter van de VVS-OR worden doorgestuurd.



MEER ENERGIE DOOR OR

OR IN DE ENERGIE SECTOR

een praktische toepassing voor elektriciteitsnetbeheerders

SONJA BOUWMAN, GABRIËL BLOEMHOF

Het toepassen van OR technieken in de energiesector wint steeds meer terrein. Tijdens het NGB seminar in Lunteren in 2008 was het thema dan ook 'Operations Research and Energy'. Tijdens dit seminar hebben wij enkele OR technieken laten zien die KEMA bij haar klanten toepast. KEMA is een wereldspeler op het gebied van 'strategic & technical energy consultancy, operational support, measurements & inspection, and testing & certification'. Het leveren van op OR-technieken gebaseerde diensten is onderdeel van onze innovatieve dienstenpakket.

Dit artikel gaat over het toepassen van OR technieken voor elektriciteitsnetbeheerders (verder afgekort tot netbeheerders). Het beheren van netten wordt steeds complexer. Dit komt enerzijds doordat de toezichthouder stringenter eisen stelt aan netbeheer en anderzijds doordat de omgeving van de netbeheerders sterk aan verandering onderhevig is. Bij het laatste moet onder andere worden gedacht aan:

- de toename van decentrale en duurzame energiebronnen;
- de komst van elektrische auto's en van andere opslagsystemen;
- de toename van im- en export van elektrische energie als gevolg van de marktkoppelingen.

Als gevolg van de dynamiek en onzekerheid die kenmerkend is voor dit soort ontwikkelingen kan de netbeheerder steeds minder vraag en aanbod van elektriciteit voorspellen. Van de netbeheerder wordt nu verwacht dat hij met deze nieuwe onzekerheden moet kunnen omgaan. Het gebruik van slimme (wiskundige) technieken die hem hierbij van nut kunnen zijn, is dan ook hard nodig, zowel bij lange termijn planning als bij onderhoud en korte termijn beheer. We zullen eerst één project uitgebreid toelichten, waarna een overzicht van andere, in de energiesector gebruikte, OR-gerelateerde technieken volgt.

Vervangings- en onderhoudsmodel voor transformatoren

Het begon met een vraag van één van onze klanten, een netbeheerder, over één van hun

belangrijkste assets: 'We hebben 220 transformatoren in verschillende stations in het door ons beheerde middenspanningsnet. Deze transformatoren verschillen in onderhoudstaat en leeftijd. Volgens onze schatting moet een aantal hiervan mogelijk worden vervangen. 42 transformatoren hiervan hebben een relatief grote kans om binnenkort te falen. Het falen van een transformator kan grote gevolgen hebben voor de energievoorziening, bijvoorbeeld onderbreking van vitale werkzaamheden bij fabrieken en kantoren, huishoudens in het donker, etc. We hebben niet genoeg technisch personeel (zowel in de eigen organisatie als bij toeleveranciers) en niet voldoende budget om alles tegelijk te vervangen. Hoe kunnen we een optimaal vervangings- en onderhoudsplan maken voor deze transformatoren?'

Voor niet-ingewijden: een gemiddelde transformator in het middenspanningsnet meet 4 x 2 x 2,5 meter, weegt 50 ton en kost ruim € 300.000. Het gaat dus om significante apparatuur en investeringen (zie de foto).

Om bovenstaande vraag van de netbeheerder te kunnen beantwoorden is het REMA-T (**RE**placement and **MA**intenance for **T**ransformers) model ontwikkeld. Met het model kunnen vragen worden beantwoord, zoals: wanneer (in welk jaar) moet deze transformator worden vervangen, wat is het optimale tijdstip van onderhoud, welk onderhoud moet worden uitgevoerd en wat voor effect heeft deze vervanging op de betrouwbaarheid van het hele net? Dit zijn dagelijkse vragen voor een *asset manager* (functionaris belast met het opstellen

van investerings- en onderhoudsplannen). Voor één transformator zijn deze vragen al lastig te beantwoorden, laat staan voor een heel transformatorpark tegelijk, zeker wanneer de budgetten en mankracht beperkt zijn en over alle stations moeten worden verdeeld.

Omdat de transformatoren niet op zichzelf staan, maar deel uitmaken van het elektriciteitsnet, is ook de betrouwbaarheid van stroomlevering aan de aangesloten klanten van groot belang bij de prioritering van de vervangingen.

Opzet van het model

Het model berekent de optimale strategie op basis van minimale kosten (de doelfunctie). Deze kosten bestaan uit onderhouds- en investeringskosten, kosten van de betrouwbaarheid (uitgedrukt door een waardering voor de CML (Customer Minutes Lost)) en netverliezen. In het model zijn de consequenties van onderhoud en vervanging voor de faalkansen (en dus indirect

op de restlevensduur) meegenomen.

De asset manager (gebruiker) bepaalt wat optimaal is door zijn keuze voor de weging van CML en de andere parameters en zijn keuze voor de randvoorwaarden. Mogelijke randvoorwaarden zijn:

- aantal te vervangen, te reviseren en te onderhouden transformatoren per jaar
- budget restricties voor het hele transformatorpark
- restricties wat betreft mankracht

Het model bestaat uit verschillende geneste modules (zie figuur 1): De kern wordt gevormd door een module om de betrouwbaarheid te berekenen en een tweede voor de kosten. Met deze beide samen kan één strategie worden doorgerekend. Met behulp van een solver kan de beste strategie worden bepaald. Het model is gekoppeld aan MS Excel om de resultaten weer te geven en om een gevoeligheidsanalyse uit te voeren.

Het model bekijkt verschillende alternatieven per transformator per jaar: Niets doen, Onderhoud,



Figuur 1. Opbouw van het model in verschillende modules

Revisie, Vervangen. Niets doen betekent dat er aan die specifieke transformator in dat jaar geen onderhoud of revisie wordt uitgevoerd en dat deze niet wordt vervangen. Dit zal een bepaald effect hebben op de faalkans van desbetreffende transformator in het volgende jaar. De optimale strategie voor de asset-manager van de netbeheerder omvat alle transformatoren simultaan.

Output

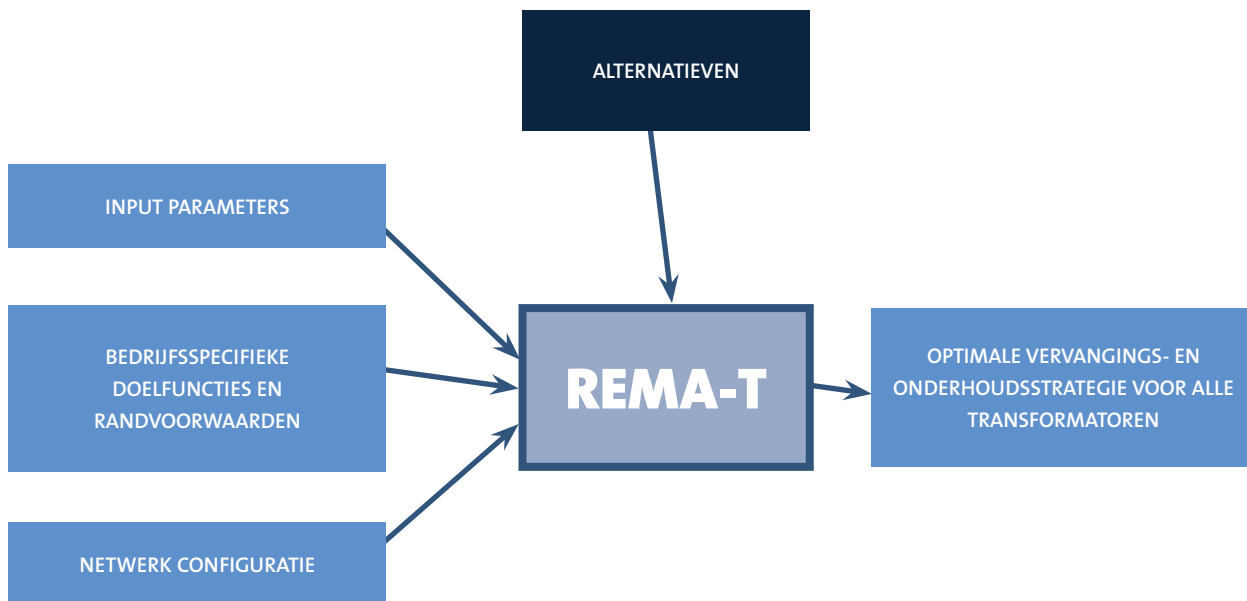
Als alle parameters in REMA-T model zijn ingevuld (figuur 2), verschijnt met één druk op de knop de optimale strategie. Om een idee te krijgen van de uitkomsten van het model geeft tabel 1 een voorbeeld van het resultaat voor 16 transformatoren in 5 stations.

De aanbevolen acties worden per station per jaar weergegeven. In 2011 wordt bijvoorbeeld transformator 2 in station 1 vervangen. De latere jaren zijn minder nauwkeurig, maar ook minder relevant voor de planner. Op dat moment is het

model namelijk allang weer een keer doorgerekend met de nieuwste inputgegevens. Immers, de asset manager actualiseert elk jaar, op basis van nieuwe gegevens, het onderhouds- en vervangingsprogramma.

Het model berekent de beste strategie. Handmatig kunnen hierin nog wijzigingen worden aangebracht. Daarmee kan worden gekeken wat het verschil is tussen de eigen strategie en de modelstrategie. Indien randvoorwaarden worden overschreden geeft het model dit aan. De gebruiker krijgt zo een goed gevoel over het model en behoudt enige flexibiliteit omtrent zijn eigen keuzes. Twee voorbeelden:

- Stel dat de vervanging van transformator 1 in station 5 drie jaar wordt uitgesteld?
Dan zullen de totale kosten stijgen 3 % (kosten investeringen dalen met 1%, kosten betrouwbaarheid stijgen met 12%.
- Stel dat de revisie van transformator 2 in station 2 één jaar eerder wordt uitgevoerd?
Dan is er niet genoeg mankracht om deze oplossing uit te voeren (maximum is 500 uur / jaar).



Figuur 2. Input / output overzicht van het REMA-T model

JAAR		2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
STATION 1	trafo 1	--	--	--	--	--	--	--	--	B	--	--	--	--	--	--	--
	trafo 2	--	V	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	trafo 3																
	trafo 4																
	trafo 5																
STATION 2	trafo 1	--	--	--	B	--	--	--	--	--	--	--	--	--	--	--	--
	trafo 2	--	B	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	trafo 3	--	--	--	--	--	--	--	B	--	--	--	--	--	--	--	--
	trafo 4																
	trafo 5																
STATION 3	trafo 1	--	--	B	--	--	--	--	O	--	--	O	--	--	O	--	--
	trafo 2	--	--	--	--	--	O	--	--	O	--	--	O	--	--	--	--
	trafo 3	--	V	--	--	--	--	--	--	O	--	--	O	--	--	O	--
	trafo 4	B	--	--	--	--	--	O	--	--	O	--	--	--	O	--	--
	trafo 5																
STATION 4	trafo 1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
	trafo 2	--	--	--	--	--	--	--	--	--	B	--	--	--	--	--	--
	trafo 3																
	trafo 4																
	trafo 5																
STATION 5	trafo 1	V	--	--	--	--	--	--	O	--	--	--	O	--	--	O	--
	trafo 2	V	--	--	--	--	--	O	--	--	O	--	--	--	O	--	--
	trafo 3	--	--	--	--	--	--	--	--	--	O	--	--	O	--	--	--
	trafo 4	--	O	--	--	O	--	--	O	--	--	O	--	--	O	--	--
	trafo 5	O	--	--	O	--	--	O	--	--	O	--	--	O	--	--	O

Tabel 1. Optimaal instandhoudingsplan voor 16 transformatoren in 5 stations voor 15 jaar (R=revisie, O=onderhoud, -- = niets doen, V=vervanging)

Conclusie REMA-T

De netbeheerder is erg tevreden met het model en gebruikt het in zijn dagelijkse praktijk met name voor een betere inschatting van de budgetten voor vervanging en onderhoud, het beter balanceren van hulpmiddelen en om een bijdrage te leveren aan hun risk management systeem. Het model heeft zijn eigen kosten en de kosten van inspecties (om de faalkansen beter in te kunnen schatten) binnen een jaar al meer dan 5 keer terugverdiend.

Andere OR gerelateerde projecten

Netbeheerders zitten natuurlijk nog met vele andere vragen omtrent het plannen en beheren van hun elektriciteitsnetten, zoals:

- Productie vertrekt naar het buitenland; om inzicht te krijgen in de relatie tussen vermindering van het opgestelde productievermogen in een bepaald net en de extra investeringen die nodig zijn om het net binnen de gestelde grenzen van kwaliteit en leveringszekerheid te kunnen beheren is grafentheorie gebruikt en

een branch & bound methode toegepast. (KEMA 2004).

- Optimale locatie voor beveiligingen; om de betrouwbaarheid te verhogen van een elektrisch netwerk kunnen extra beveiligingen (beveiligingsautomaten) worden geplaatst. Met behulp van genetische algorithmes hebben we de optimale locatie en optimale aantallen kunnen bepalen. (CIRED 2009).
- Waarin moeten we als eerste investeren? Planning, ontwerp en projecteren van elektriciteits- en gasnetten is de laatste jaren sterk van karakter veranderd. Waar voorheen werd gestreefd naar een technisch optimaal net, wordt tegenwoordig een keuze gemaakt uit een aantal alternatieven, die qua prestatie, risico en kosten met elkaar zijn vergeleken. De onderbouwing van de keuze dient helder en eenduidig te zijn. Daarom hebben we een methode ontwikkeld om beslissingen te nemen onder onzekerheid voor netbeheerders (CIRED 2003).
- Wat is de interactie tussen de verschillende marktpartijen? FLEXNET is een business game waarbij de elektriciteitssector kan worden nagebootst. Op het spelbord wordt een elektriciteitsnet opgebouwd compleet met productie en consumptie. De pc wordt gebruikt om het net technisch door te rekenen en om congesties en problemen inzichtelijk te maken. Met het spel wordt inzicht verkregen in hoe de sector werkt. Ook kunnen wijzigingen op het gebied van regelgeving worden gesimuleerd (ISAGA 2007).
- Hoe gaan we om met alle nieuwe onzekerheden door decentrale productie en duurzame energie? Met onzekerheden kun je op twee manieren omgaan. Je kunt ze verkleinen (bijvoorbeeld door voorspellingen te maken) of je kunt ze accepteren en inzichtelijk maken. Door de traditionele *loadflows* (netwerkberekeningen) te koppelen aan Monte Carlo simulaties

is het mogelijk om probabilistische *loadflow*-berekeningen uit te voeren. Hiermee is het eenvoudiger om alle nieuwe onzekerheden mee te nemen in het netbeheer.

Conclusie

OR technieken zijn erg bruikbaar voor de energie sector, niet alleen voor netbeheerders, maar ook voor producenten en andere partijen. Er spelen veel verschillende problemen waar OR technieken kunnen worden ingezet.

LITERATUUR

CIRED 2003, *Better Decisions Using Forecasting And Probabilistic Risk*, S. Bouwman, G.A. Bloemhof, J. Knijp, A. Jongepier, R. Trekop ANALYSIS, Cired 2003, may 2003, session 5.

ISAGA 2007, *Simulation in the energy-infrastructure; FLEXNET, a flexible approach*, Gabriël A. Bloemhof, Femmy Combrink, Peter T.M. Vaessen; in: Igor Mayer & Hanneke Mastik (eds.) 2007, *Organizing and Learning through Gaming and Simulation*, Proceedings of Isaga 2007, Eburon, Delft, ISBN 978 90 5972 2316.

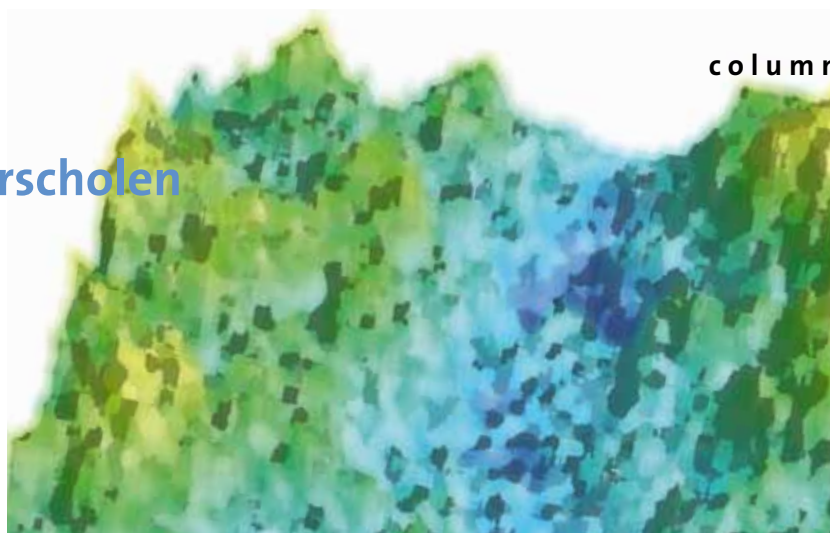
KEMA (2004). *Hoeveel investeren? Kostenbepaling voor netbeheerders bij verminderende productie*. PREGO 16 rapport, downloaden via www.kema.com/nl/consulting_services/transmission_and_distribution/publications/prego/measurements/.

CIRED 2009, *Increasing Quality of Supply of EDP through optimal and strategic Distribution Automation design*, Rui Oliveira, Gabriël Bloemhof, Aurélio Blanquet, Cired 2009, Praag, June 2009, paper 0459.

SONJA BOUWMAN is afgestudeerd aan de Universiteit Twente in 1999. Sinds 2001 is ze werkzaam bij KEMA Consulting op het gebied van beslistheorie, wiskundig modelleren, voorspellen, statistiek en data-analyse. E-mail: <Sonja.Bouwman@kema.com>.

GABRIËL BLOEMHOF is afgestudeerd aan de Universiteit Twente in 1982. Sinds 1986 is hij werkzaam bij KEMA Consulting op het gebied van netwerkplanning, risicomangement, storingsstatistiek en innovatie. E-mail: <Gabriel.Bloemhof@kema.com>.

Dromen van zomerscholen



Het is herfst en het semester staat in volle bloei. Ik loop over de parkeerplaats, op weg naar de collegezaal. Het miezert, wat tot optimisme stemt, maar helaas zie ik de plas achter de bruine Volvo te laat. De leren zool van mijn gehavende voorjaarschoen zuigt het regenwater gretig op. Even later betreed ik de nog lege zaal op vloer elf. Ik zet mijn koffie op de lessenaar en hang mijn klamme jas op de haak. De klok zegt 8:37. Door de gekneusde lamellen staar ik naar de contouren van de stad. De motregen maakt mijn uitzicht tot een pointillistische ansichtkaart. Ik denk aan verre oorden en zomerscholen...

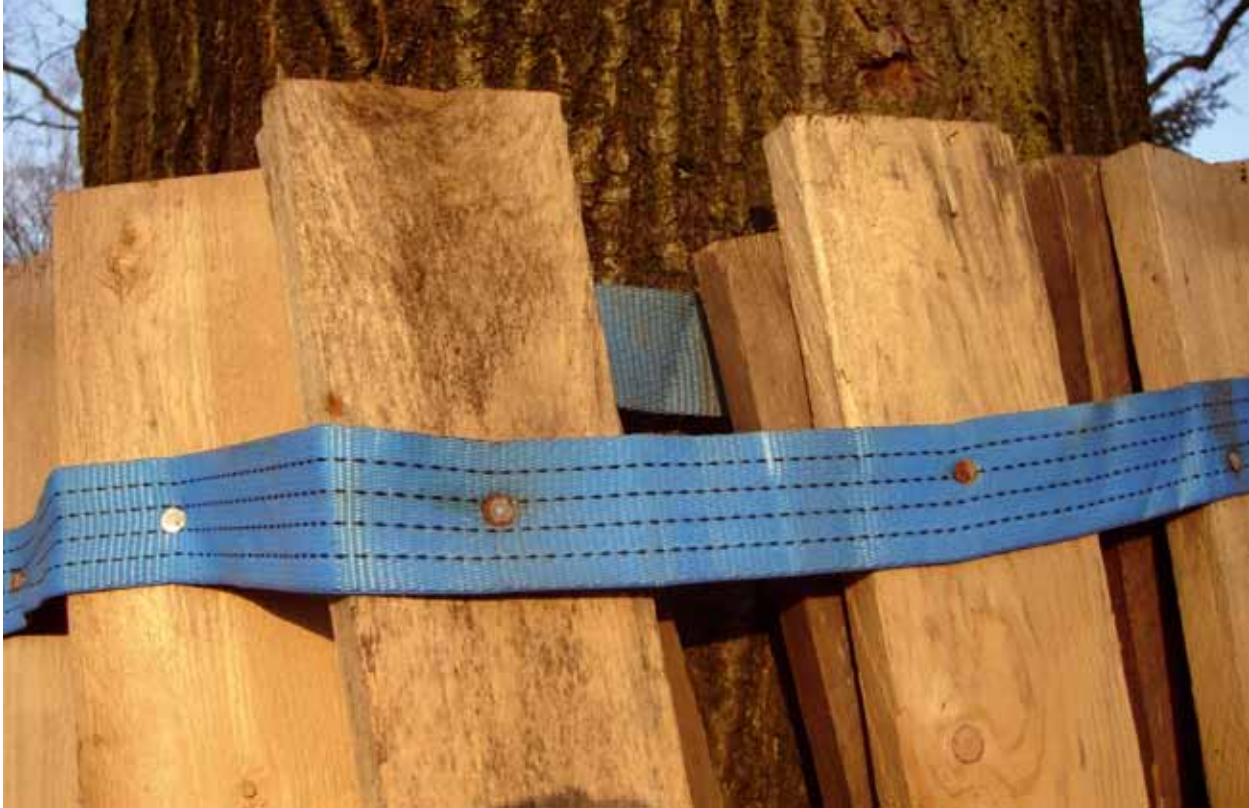
Zomerscholen. Verdieping zoeken in de zomer door het volgen van lessen, op een klein gebied, gegeven door grote namen. Lessen door experts voor experts in spe, voor studenten en onderzoekers in een bepaalde discipline. Zomerscholen zijn het talrijkst in de muziek, maar ze lenen zich ook voor de schilderkunst en wiskunde. De fijne kneepjes van het vak als voorbereiding op het echte werk. Vaak bij mooi weer in een rustige omgeving, ver weg van de afleiding. Oude kloosters met harde bedden vormen geschikte locaties.

In de kansrekening staan ieder jaar verscheidene zomerscholen op het programma: Cornell Probability Summer School (Ithaca), Ecole d'Été de Probabilités (Saint-Flour), Escola Brasileira

de Probabilidade (São Paulo) en PIMS Summer School in Probability (Vancouver en Washington). Daarentegen kennen de operationele wetenschappen eigenlijk geen zomerscholen. Zelfs de machtige Amerikaanse vakorganisatie INFORMS heeft geen jaarlijkse zomerschool, terwijl de honger van het aanstormend talent toch groot is. Misschien ligt er wel een kans voor Nederland. We doen al decennialang mee op het hoogste niveau van de OR. Nederland is te klein voor de Olympische Spelen, maar een jaarlijkse OR zomerschool laat onze infrastructuur wel toe. Zou het niet prachtig zijn als de LNMB (Landelijk Netwerk Mathematische Besliskunde) bijeenkomst in Lunteren er een broertje bij krijgt in de zomer?

Ik schrik op van de krakende deur. De eerste student arriveert. Ik sta geleund tegen het bord, met het krijt in de hand. De tekens op het bord lijken een wirwar van strepen. Maar bij het uitvegen herken ik in de linkerbovenhoek een zon, daarnaast 14 juli 2035, en daaronder in fragiele letters: frisse lucht en fantastische lessen, groeten van de Veluwe, Johan jr.

*JOHAN VAN LEEUWAARDEN is werkzaam in de groep Stochastische Besliskunde bij de faculteit Wiskunde en Informatica van de TU Eindhoven. Tevens is hij research fellow bij EURANDOM.
E-mail: <j.s.h.v.leeuwaarden@tue.nl>.*



OPTIMALE ONDERZOEKSDESIGNS MEER EFFICIËNTIE VOOR MINDER GELD

Tot nu toe werden optimale designmethoden niet echt veel toegepast in sociaalwetenschappelijk en biomedisch onderzoek om onderzoeksdesigns te kiezen. Aangezien het gebruik van optimale of hoog-efficiënte onderzoeksdesigns de kosten voor het uitvoeren van onderzoek kunnen reduceren, is het in toenemende mate belangrijk dat sociaalwetenschappelijke en biomedische onderzoekers deze methoden toepassen voor hun designkeuze. In dit artikel worden enkele voorbeelden uit de praktijk van het onderzoek gegeven die laten zien hoe met gelijkblijvende kosten de efficiëntie van de parameterschatters verhoogd kan worden.

MARTIJN P.F. BERGER

Het onderzoeksdesign in sociaalwetenschappelijk en biomedisch onderzoek beschrijft hoe de data verzameld worden. De keuze van een design is cruciaal en wel om twee redenen. Ten eerste

omdat het design aan bepaalde voorwaarden moet voldoen om betrouwbare en valide conclusies te trekken. Deze voorwaarden zijn uitgebreid in de methodologische literatuur behandeld door

o.a. Cox and Reid (2000) en Shadish, Cook and Campbell (2002).

Een tweede reden is dat met een designkeuze getracht wordt om zoveel mogelijk informatie voor het schatten van de modelparameters en het toetsen van hypothesen te verkrijgen. Aan de vraag welk design leidt tot de meest efficiënte parameterschatters is tot nu toe in sociaalwetenschappelijk en biomedisch onderzoek weinig aandacht besteed. Met efficiëntie wordt de nauwkeurigheid van de schatter in termen van variantie bedoeld. Hoe kleiner de variantie van de parameterschatter, hoe nauwkeuriger deze schatter is en hoe meer vermogen de toets heeft om echte effecten te vinden.

De vraag of het onderzoekdesign tot de meest efficiënte schatters voor de modelparameters leidt staat centraal in *optimal design* onderzoek. Sinds het werk van Fisher (1935) hebben statistici gewerkt aan problemen rond de keuze van een optimaal experimenteel design. Zie Atkinson, Donev and Tobias (2007) voor een overzicht. Hoewel dit onderzoek aantoonde dat het gebruik van optimale designs tot grote efficiëntieverbeteringen en kostenbesparingen kan leiden, zijn deze resultaten nog betrekkelijk weinig toegepast. Hiervoor is een aantal redenen te noemen.

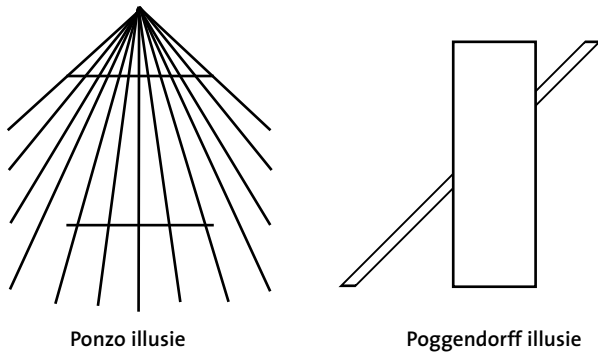
De nadruk lag in de begintijd vooral op het onder bepaalde assumpties analytisch afleiden van het optimale design voor eenvoudige (regressie)modellen. Deze assumpties gingen in de praktijk vaak niet op en het optimale design was dan niet bruikbaar. Een andere reden waarom de resultaten van optimaal designonderzoek weinig werden toegepast, is dat hoewel de meeste sociaal-wetenschappelijke en biomedische onderzoekers getraind zijn in statistische methoden, zij niet bekend zijn met de concepten uit optimale designtheorie.

De laatste jaren is er echter een toename van onderzoek naar optimale designs te constateren, waarbij de nadruk meer ligt op het zoeken naar

hoog-efficiënte designs, die niet per se optimaal zijn, maar toch een hogere efficiëntie hebben dan de traditionele designs. Door verbetering van rekenmethoden werd het gemakkelijker om de robuustheid van een optimaal design tegen schendingen van de modelassumpties te onderzoeken, zodat de toepassingsmogelijkheden in de praktijk verruimd werden. Uitbreiding van de rekenmogelijkheden maakte het ook gemakkelijker om optimale en hoog efficiënte designs te vinden voor multilevel en random effect modellen die vooral in sociaalwetenschappelijk en biomedisch onderzoek gebruikt worden (zie o.a. Raudenbush, 1997, Moerbeek, van Breukelen & Berger, 2000 en Moerbeek, 2001). In Berger and Wong (2005) wordt een overzicht gegeven van toepassingen van optimaal designonderzoek in verschillende wetenschapsgebieden.

Met de stijgende kosten voor het uitvoeren van onderzoek worden de resultaten van optimale design onderzoek voor onderzoekers en subsidiegevers steeds belangrijker: kleinere steekproeven om efficiënter te schatten met een reductie van de kosten. Optimale designs kunnen vergeleken met traditionele designs leiden tot 20-40% reductie van de steekproefomvang. Bij gelijkblijvende steekproefgrootte zal het vermogen om echte effecten te vinden ook stijgen.

Het optimale design probleem is te omschrijven als het zoeken naar die kenmerken van een design welke de variantie van de parameterschatter minimaliseren. Als er meer dan een parameter in het model is, dan wordt een functie van de variantie-covariantie matrix van de parameterschatters geminimaliseerd. In de literatuur zijn verschillende functies of criteria voorgesteld. De meest bekende zijn de alfabetische optimaliteitscriteria, zoals het A- of het spoorcriterium, het D- of determinantcriterium en het E- of grootste eigenwaarde criterium. Het D-criterium heeft prettige eigenschappen en wordt daarom vaak gebruikt. Dit criterium minimaliseert het volu-



Figuur 1. Ponzo en Poggendorff illusies

me van een betrouwbaarheidselipsoïde van de modelparameters en is invariant onder lineaire transformatie van de onafhankelijke variabele (Zie Atkinson, Donev and Tobias, 2005; Berger and Wong, 2009) In de volgende secties wordt een tweetal toepassingen van optimale designtheorie gegeven.

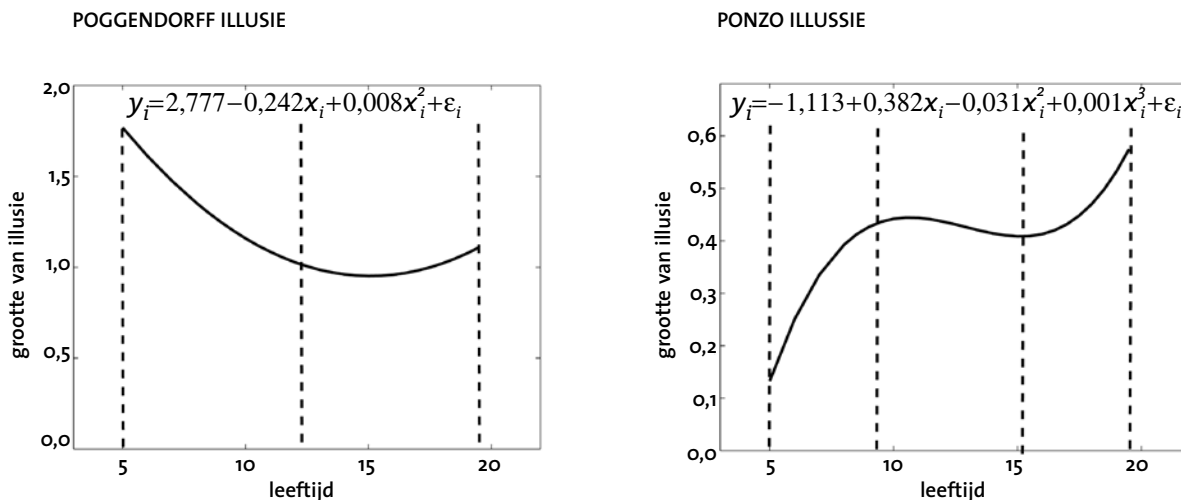
Poggendorff en Ponzo illusies

Twee bekende visuele illusies zijn het Poggendorff en het Ponzo effect. Deze illusies staan afgebeeld in Figuur 1.

Leibowitz en Judisch (1967) en Leibowitz en Gwozdicki (1967) hebben bij 144 kinderen de relatie van deze illusies met de leeftijden onderzocht. In Figuur 2 wordt de relatie met leeftijd voor het Poggendorff effect beschreven door een kwadratische polynoom en voor het Ponzo effect door een derde machtspolynoom.

Het experimentele design bestond uit negen verschillende leeftijdsgroepen variërend van 5 tot 19,5 jaar met ieder 16 kinderen. De vraag is nu of deze negen leeftijdsgroepen allemaal nodig zijn om efficiënt de polynomische parameters van de kwadratische en de derde machtspolynoom te schatten. Hoeveel leeftijdsgroepen en hoeveel kinderen in iedere leeftijdsgroep zijn nodig om deze parameters zo efficiënt mogelijk te schatten? Het antwoord op deze vraag hangt af van het model en van het gekozen optimaliteitscriterium.

Optimale designs voor polynome modellen zijn bekend (Pukelsheim, 1993, hoofdstuk 9). Voor een leeftijdsspreiding van 5-19,5 jaar zijn de D-optimale leeftijdscategorieën respectievelijk [5, 9, 15,5 en 19,5] (Ponzo illusie) en [5, 12,25 en 19,5] (Poggendorff illusie). Deze D-optimale leeftijdscategorieën zijn als verticale stippellijnen in Figuur 2 weergegeven en staan ook in Tabel 1. Het kan zijn dat de leeftijdscategorieën 12,25 en 15,5 niet



Figuur 2. De grootte van de Ponzo en Poggendorff illusies als functie van leeftijd

Ponzo	D-optimal design				Sub-optimal design				RE _{sub-optimal}	RE _{oorsp. design}
Leeftijden	5	9	15,5	19,5	5	8,5	16,5	19,5	0,9627	0,8020
# kinderen	36	36	36	36	36	36	36	36		
Poggendorff										
Leeftijden	5	12,25	19,5		5	12,5	19,5		0,9992	0,7243
# kinderen	48	48	48		48	48	48			

Tabel 1. Relatieve efficiencies van oorspronkelijk en suboptimale designs voor de Ponzo en Poggendorff experimenten

beschikbaar zijn voor het onderzoek. In dat geval kunnen deze categorieën wat aangepast worden. In Tabel 1 staan ook sub-optimale designs met enkele aangepaste categorieën. De relatieve efficiënties (RE) voor deze designs ten opzichte van de optimale designs zijn erg hoog. In Tabel 1 staat ook de relatieve efficiëntie van het oorspronkelijke design met 9 leeftijdscategorieën en 16 kinderen per categorie. Deze zijn respectievelijk $RE = 0,8020$ en $RE = 0,7243$. Dit betekent dat bijna $(1/RE - 1)100\% = 25\%$ respectievelijk 38% meer observaties voor het oorspronkelijke experimentele design nodig zijn om even efficiënt te worden als het D-optimale design voor het Ponzo en Poggendorff effect.

Tot slot, de volgende vuistregel voor het kiezen van een hoog efficiënt design kan worden gebruikt. Kies zoveel leeftijdscategorieën als er parameters in het model zijn die simultaan geschat worden. Voor een tweede machtsfunctie of derde machtsfunctie zijn dit respectievelijk 3 en 4 parameters. Voor de zekerheid kan men dan nog één of twee extra leeftijdscategorieën opnemen om modelfit mogelijk te maken.

Grootschalige cohort onderzoeken

Grootschalige cohort onderzoeken worden in de gezondheidswetenschappen vaak uitgevoerd om

effecten van behandelingen over de tijd te meten (Calle et al 2002, van den Akker et al 2008) Deze designs op zijn gebaseerd op heel grote steekproeven met veel herhaalde metingen. Zij zijn daarom erg kostbaar en hebben lange tijd nodig om te voltooien. Dit heeft tot gevolg dat de kans op uitval (drop-out) groot is.

Een cohort design wordt gekenmerkt door het aantal cohorten dat over de tijd gemeten wordt en de tijdstippen en het aantal tijdstippen waarop de metingen plaatsvinden. Het aantal patiënten en het aantal herhaalde metingen in een cohort zijn bepalend voor de efficiëntie. Hoe meer personen en herhaalde metingen hoe efficiënter de parameterschattingen en hoe meer vermogen er is om vast te stellen dat de effecten niet toevallig zijn. Dit heeft natuurlijk een prijs, namelijk dat de kosten omhoog gaan.

In Figuur 3 staan drie verschillende designs. Het bovenste design is een longitudinaal design met slechts 1 cohort gemeten over het gehele tijdsinterval. Het middelste design is een typisch gemengd cohort design waarbij drie cohorten gemeten worden over drie opeenvolgende tijdsintervallen en het onderste design is een cross-sectioneel design waarbij drie cohorten gemeten worden op drie tijdstippen. De personen in het longitudinale design worden tijdens het gehele tijdsinterval gemeten, terwijl de personen van het

LONGITUDINAAL DESIGN

1 cohort			
	Tijdstippen		

COHORT DESIGN

Cohort 1			
Cohort 2			
Cohort 3			
	Tijdstippen		

CROSS-SECTIONEEL DESIGN

Cohort 1				
Cohort 2				
Cohort 3				
	Tijdstippen			

Figuur 3. Drie voorbeelden van cohort designs

cross-sectionele design slechts op een van de drie tijdstippen gemeten worden.

Het optimale cohort design probleem bestaat uit een drietal vragen:

- Wat is het optimale aantal cohorten ?
- Wat is het optimaal aantal tijdstippen?
- Wat is de optimale locatie van deze tijdstippen?

Om deze vragen te kunnen beantwoorden moet een aantal zaken worden vastgelegd. Ten eerst moet het statistische model met parameters

gespecificeerd worden. Een veel gebruikt model voor cohort onderzoek is het lineaire mixed effect model (Verbeke and Molenberghs, 2000), waarbij personen gekarakteriseerd worden door ‘random’ parameters en effecten door *fixed* parameters. De efficiëntie kan weergegeven worden door de determinant van de variantie-covariantie matrix van de *fixed* parameterschatters. De kosten C van een cohort design met S cohorten kunnen door de volgende functie weergegeven worden:

$$C = \sum_{s=1}^S n_s (c_1 + c_2 (q_s - 1))$$

De kosten C bestaan uit de kosten voor het opnemen en meten van een nieuwe persoon c_1 en de kosten voor het verkrijgen van een (herhaalde) meting per persoon c_2 . Het aantal personen en het aantal herhaalde metingen per cohort is n_s en q_s .

Het vaststellen van een optimaal cohort design kan vertaald worden in een optimaliseringsprobleem waarbij het determinant criterium wordt geminimaliseerd onder bepaalde voorwaarden. Een voorwaarde kan zijn dat de totale kosten C gelijk blijven en/of dat de kosten voor het selecteren van een persoon altijd hoger zijn dan de kosten voor het herhaald meten van die persoon, dat wil zeggen $c_1 > c_2$. Tekle et al (2008) hebben een aantal vragen van het optimale cohort design probleem onderzocht. Voor de door hen onderzochte condities vonden zij numeriek dat het optimale aantal herhaalde metingen voor een cohort design met S cohorten en een mixed effectenmodel met p parameters gelijk is aan $q^* = S + (p-1)$. Hun vergelijking tussen verschillende cohort designs toonde dat het design met slechts $S = 1$ cohort het meest efficiënt was.

Conclusie

Het gebruiken van resultaten uit optimaal design onderzoek kan leiden tot efficiëntere keuzes voor onderzoeksdesigns met meer vermogen om echte effecten te vinden tegen lagere kosten voor uitvoering van onderzoek. De auteur van dit artikel spreekt de wens uit dat sociaalwetenschappelijke en biomedische onderzoekers in de toekomst de resultaten uit optimaal design onderzoek meer zullen gebruiken voor hun design keuze.

LITERATUUR

- Atkinson, A.C., Donev, A.N., and Tobias, R.D. (2007). Optimum experimental designs, with SAS. Oxford University Press.
- Berger, M.P.F. and Wong, W.K. (Eds.) (2005). *Applied Optimal Designs*. John Wiley & Sons.
- Berger, M.P.F. and Wong, W.K. (2009) *An Introduction to Optimal Designs for Social and Biomedical Research*. John Wiley & Sons.
- Calle, E.E., Rodriguez, C., Jacobs, E.J., Almon, M.L., Chao, A., McCullough, M.L., Feigelson, H.S., and Thun, M.J. (2002). The American Cancer Society cancer prevention study II: Nutrition cohort, *Cancer*, 94, 500-511.
- Cox, D.R. and Reid, N. (2000). *The Theory of the Design of Experiments*. Boca Raton: Chapman & Hall/CRC.
- Fisher, R.A. (1935). *The Design of Experiments*. Oliver & Boyd, Edinburgh.
- Leibowitz, H.W. and Gwozdicki, J. (1967). The Magnitude of the Poggendorff illusion as a function of age. *Child Development*, 38, 573-580.
- Leibowitz, H.W. and Judisch, J.M. (1967). The relation between age and the magnitude of the Ponzo illusion. *American Journal of Psychology*, 80, 105-109.
- Moerbeek, M., van Breukelen, G.J.P. and Berger, M.P.F. (2000). Design issues for experiments in multilevel populations. *Journal of Educational and Behavioural Statistics*, 25, 2, 271-284.
- Moerbeek, M. (2001). Het ontwerpen van experimenten voor genestelde datastructuren. *STATOR*, 2, 2, 16-19.
- Pukelsheim, F. (1993). *Optimal Design of Experiments*. New-York: John Wiley & Sons
- Raudenbush, S.W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods*, 2, 2, 173-185.
- Shadish, W.R., Cook, T.D. and Campbell, D.T. (2002). *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston: Houghton-Mifflin.
- Tekle, F., Tan, F.E.S. and Berger, M.P.F. (2008). D-optimal cohort designs for linear mixed-effects models. *Statistics in Medicine*, 27, 2586-2600.
- Van den Akker, M., Spigt, M.G., De Raeve, L., Van Steenkoste, B., Metsemakers, J.F.M., Van Voorst, E.J., De Vries, H. (2008). The SMILE study: A study of medical information and lifestyles in endhoven. The rationale and contents of a large prospective dynamic cohort study. *BMC Public health*, 8, 19.
- Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*, Springer Verlag, New York.
- MARTIJN P.F. BERGER is hoogleraar aan de Universiteit Maastricht, vakgroep Methodologie en Statistiek. Zijn onderzoek richt zich vooral op optimale designproblemen voor medisch en gezondheidswetenschappelijk onderzoek. E-mail: <Martijn.Berger@STAT.unimaas.nl>.

VOERTAAL: GEBROKEN ENGELS

FRED STEUTEL

Engels heeft de wereld veroverd, nou ja, de halve wereld, en dat nog zeer onvolledig. Buiten de zogenaamde *native speakers* spreekt bijna iedereen een vorm van gebroken Engels: de woordkeus is klein en de uitdrukkingsmogelijkheden zijn beperkt. Hoe zeg je in het Engels (of in een andere vreemde taal) ‘het gras voor de voeten wegmaaien’ of ‘de wind uit de zeilen nemen’. Wat we geleerd hebben is als een tuin zonder bloemen, als een bos met maar één soort bomen. Hoe blij, of hoe ongelukkig, moeten we zijn met de hegemonie van het Engels in de gesproken en geschreven taal en dat speciaal in de statistiek?

Statistiek is een internationaal bedrijf. Vroeger vond je, enigszins paradoxaal, die internationaliteit terug in de nationale tijdschriften. Elk land had minstens één statistiektijdschrift in zijn

eigen taal. Zo hadden wij *Statistica (Neerlandica)*, in het Nederlands, zij het dat er wel eens een stukje in het Engels instond, zoals in 1957 ‘*Statistical Priesthood*’ van Van Dantzig. In 1982 werd de voertaal – uitsluitend – Engels. Vroeger mocht er, meen ik, in het Nederlands, Frans, Duits en Engels in worden geschreven.

In Duitsland was er het *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, door lezers liefdevol afgekort tot *Zeitschrift*, zelfs Amerikanen konden dat woord zeggen. Sinds 1986 heet het *Probability Theory and Related Fields*; je mag er nog wel in het Duits en het Frans in schrijven, maar bijna niemand doet dat.

In Scandinavië hadden we het *Skandinavisk aktuarietidskrift*, inmiddels omgedoopt tot *Scandinavian actuarial journal*. Er mag worden geschreven in Engels, Duits en Frans, en in de Scandinavische talen – volgens mijn woordenboek zijn dat Zweeds, Deens, Noors, IJslands en Faerøers – de Scandinaviërs hebben meer respect voor de eigen taal dan Nederlanders – maar de meeste artikelen zijn toch in het Engels.

De Fransen hebben nog steeds de *Annales de l’Institut Henri Poincaré, Probabilités et Statistique*, zij het dat de toevoeging is veranderd in *Probability and Statistics*. Er wordt weinig Frans in geschre-

ven, maar het mag wel. De enige Franse titel die ik aantrof in het nummer dat ik inkeek, luidde ‘*Le poisson n’a pas d’arêtes*’ – de vis heeft geen graten. Zo leer je nog eens wat: ‘graten’ heten in het Frans ‘*arêtes*’ – toch handig in een Frans visrestaurant. De ‘*poisson*’ heeft niets met kansrekening van doen; het is een visvormige figuur – zonder graten.

In Spanje had je de *Trabajos de Estadística*. ‘*Estadística*’ betekent natuurlijk ‘statistiek’, en ‘*trabajos*’ is Spaans voor ‘*travaux*’, een woord dat we allemaal kennen van de wegomleidingen in Frankrijk. Het blad heet nu *Trabajos de Estadística y de Investigación Operativa*, sedert 1985 ‘*no longer indexed*’ door *Mathematical Reviews*. Toch leren we weer iets: het ‘Nederlandse’ *Operations Research* is in het Spaans *Investigación Operativa*.

De Russen hebben *Teorija Verojatnostej i ee Primenenija (Theory of Probability and its Applications)*. Ik ken mensen die de Russische titel zonder haperen konden uitspreken en die niet konden wachten op de Engelse vertaling en daarom Russisch leerden. Zelf heb ik in dit verband nog een boekje met Russische wiskundewoorden aangeschaft. De voertaal is nu Russisch of Engels, maar soms wordt een vertaling in het Russisch geëist.

Het wonderlijkste tijdschrift in een ‘vreemde’ taal gaat niet over statistiek, het heet *Funkcialaj Ekvacioj*. Dat is – u dacht het misschien al – Esperanto voor ‘functionaalvergelijkingen’. Het wordt uitgegeven door een aantal Japanse faculteiten op het gebied van taal en wiskunde. Hoewel Esperanto een Europese taal is, was het heel populair in Azië. Maar ook hier heeft het Engels toege-

slagen: je kunt in *Funkcialaj Ekvacioj* niet meer, zoals vroeger, in het Esperanto of het Frans publiceren – alleen in het Engels. Ook vroeger werd Esperanto weinig gebruikt, en, als dat al gebeurde, werd het door MathSciNet niet vertaald: Masuo Hukuhara. (Esperanto) *Funkcialaj Ekvacioj* 9 1966 ii—iii; verder niets.

Op internet lees ik een discussie over Esperanto in de wetenschap. Er zijn een paar opmerkingen van de statisticus Herbert Rubin, die naar zijn zeggen met behulp van Frans, Italiaans en Spaans wel eens een artikel in het Esperanto heeft gelezen. Hij spreekt en passant schande van een Nederlands tijdschrift: ‘The Dutch journal, *Journal of Nuclear Materials*, accepts articles in English, French and German only – not Dutch.’ Hij was mogelijk niet bekend met *Statistica Neerlandica*.

Onlangs las ik een in memoriam voor een Nederlandse wiskundige. De auteur was Duits en de taal was – gebroken – Engels. Er was van alles mis mee; verkeerde woorden, verkeerde uitdrukkingen en verkeerde volgorde. Je bent dan al blij dat het niet wordt uitgesproken. Ik lees vele malen liever goed Duits dan slecht Engels.

Samenvattend stel ik vast dat we met de ‘anglicatie’ van de Europese taal meer verloren hebben dan gewonnen, en niet alleen in de statistiek.

*FRED STEUTEL is emeritus hoogleraar kansrekening aan de TU Eindhoven.
E-mail: <f.w.steutel@tue.nl>.*



de auto gedurende vier jaar hetzelfde strakke aflossingsschema wordt gevolgd – inclusief rentebetalingen – zoals banken dat eisen, komt het doel van het spaargeld niet in gevaar en is men aanzienlijk goedkoper uit. Mensen hebben bepaalde zaken kennelijk graag ‘goed, gescheiden, geregeld’ en accepteren daarvoor een nadelige *spread* in de rentepercentages.

EMOTIONELE HOKJES

Mental Accounting: liever overzichtelijk dan kostenefficiënt

Het blijkt dat iedereen in zijn hoofd een compleet bankbedrijf meevoert. Met rekeningen die eigen waarderingen genereren voor bij- en afschrijvingen. En die op bepaalde momenten geopend of gesloten kunnen zijn. Zo kan € 5 in het ene geval ongemerkt worden uitgegeven, waar zo'n zelfde bedrag – afgeschreven van een andere rekening – misschien als een zware molensteen wordt ervaren. Het fenomeen staat bekend als *mental accounting*. Twee voorbeelden.

Auto van de kinderen?

De irrationele dwang tot het opdelen van het beschikbare geld in ‘potjes’ gaat ver. Veel huishoudens hebben leningen lopen voor de aanschaf van een auto, een wasmachine of een koelkast. Daartegenover staan dan vaak spaarrekeningen met een bepaald doel: de komende studies van opgroeiende kinderen, pensioenaanvulling of een *sabbatical* over 10 jaar. Het verschil tussen de te betalen en te ontvangen rentepercentages is groot. Toch is het financieren van een auto uit de studiepots van jonge kinderen meestal volstrekt onbespreekbaar. Vreemd, want als voor

Rad voor ogen

Gebruikers van creditcards draaien zichzelf met het grootste gemak een rad voor ogen. De betrokken maatschappijen doen daar graag aan mee. Want het gaat om veel geld. De uitgaven die met een creditcard worden gedaan worden in het algemeen niet als ‘schuld’ in het geheugen gegrift. Zolang ze maandelijks na de eerste factuur worden afgerekend is daar niets mis mee. Creditcardmaatschappijen bieden echter de mogelijkheid het openstaande bedrag gespreid te betalen. Dan is de schuld snel geboren, zo blijkt.

De meeste creditkaarthouders nemen zich voor de schuld op de kaart niet te laten oplopen. Tijdens een onderzoek van de University of Michigan onder een grote groep Amerikaanse creditcardhouders zegt slechts 27% van de kaarthouders bewust wél van de kredietfaciliteiten gebruik te maken, de rest zegt nooit te willen lenen en alles direct af te rekenen. Maar de cijfers wijzen uit dat 75% van de kaarthouders daarop regelmatig rentebetalingen doet. Het waardevolle ‘plastic’ leidt dus tot schulden die niet als zodanig worden ervaren. En tegen extreem hoge kosten. Want de maatschappijen rekenen forse rentepercentages die in geen verhouding staan tot de rente die de kaarthouder betaalt op een normale rekening.

Overgenomen met toestemming uit *Beslissen is Menselijk*, een uitgave van Cardano Risk Management, Rotterdam (2003).

IN MEMORIAM JAAP WESSELS (1939 - 2009)

Op 30 juli 2009 is prof.dr. Jaap Wessels, emeritus hoogleraar van de Technische Universiteit Eindhoven, op 70-jarige leeftijd in Veldhoven overleden.

Jaap Wessels werd geboren op 19 januari 1939 in Amsterdam. Zijn vader was fabrieksarbeider bij Fokker, en had in de fabriek en op een avondschool het vak van bank- en plaatwerker geleerd. Na de lagere school ging Jaap, sterk gestimuleerd door het schoolhoofd (zijn ouders vonden dat aanvankelijk niet zo nodig), naar de 3-jarige en later de 5-jarige HBS. Zijn wiskundeleraar bracht hem er toe daarna niet naar de HTS te gaan, maar wiskunde te gaan studeren. In 1963 behaalde hij het doctoraal examen Wis- en Natuurkunde aan de Universiteit van Amsterdam. Vervolgens ging hij naar Eindhoven en startte hij aan de TU/e zijn promotie-onderzoek in het jonge vakgebied Markov beslissingsprocessen, onder supervisie van Professor Benders. Dat mondde in 1968 uit in het proefschrift *Decision rules in Markovian decision problems with incompletely known transition probabilities*. In hetzelfde jaar werd hij benoemd tot lector en in 1973 tot hoogleraar in de stochastische beslissonde aan de faculteit Wiskunde en Informatica van de TU/e.

Door de impulsen van Wessels in Eindhoven, en van De Leve en zijn leerlingen in Amsterdam, werd Nederland in de zeventiger jaren in feite het

internationale centrum van het onderzoek op het gebied van de Markov beslissingsprocessen. Jaap Wessels vormde een belangrijke school, en veel van zijn toenmalige promovendi werden hoogleraar: Wijngaard, Van Nunen, Van Hee, Van der Wal, Zijm en Van Damme. In die tijd begon hij ook veel van zijn onderzoek te baseren op samenwerkingsprojecten met het bedrijfsleven en overheidsinstanties. Bijvoorbeeld op de gebieden voorraadplanning en personeelsplanning. Wat later ging Jaap zich op andere onderwerpen richten, zoals netwerken van wachtrijen (veelal in het kader van computer performance), neurale netwerken en gestructureerde Markovprocessen. Ook in elk van deze onderwerpen zette hij Eindhoven op de internationale kaart, en wederom werden enige van zijn promovendi hoogleraar: Van der Aalst, Adan, Van Houtum.

Zelf heeft hij zes boeken gepubliceerd en zo'n 150 artikelen, een indrukwekkende wetenschappelijke nalatenschap. Bij al zijn theoretische onderzoek hield Jaap dus ook een scherp oog voor de toepassingen. Hij bouwde aan een netwerk van contacten met de industrie, en ook met hoogleraren van de faculteit Bedrijfskunde. Hij speelde een grote rol bij het totstandkomen van de ontwerpersopleidingen Wiskunde voor de Industrie en Logistics Management Systems, die inmiddels ruim 20 jaar bestaan. Ook bij de totstandkoming van onderzoekschool Beta, een samenwerkingsverband van Twente en Eindhoven, was de rol van Wessels groot. Beta eerde Jaap enige jaren geleden door hem tot Fellow te benoemen.



Jaap ontplooië ook landelijk tal van activiteiten en initiatieven. Hij was mede-oprichter van het Landelijk Netwerk Mathematische Besliskunde (LNMB). Ook was hij voorzitter van de VVS, de Vereniging voor Statistiek. De VVS heeft hem later tot erelid benoemd.

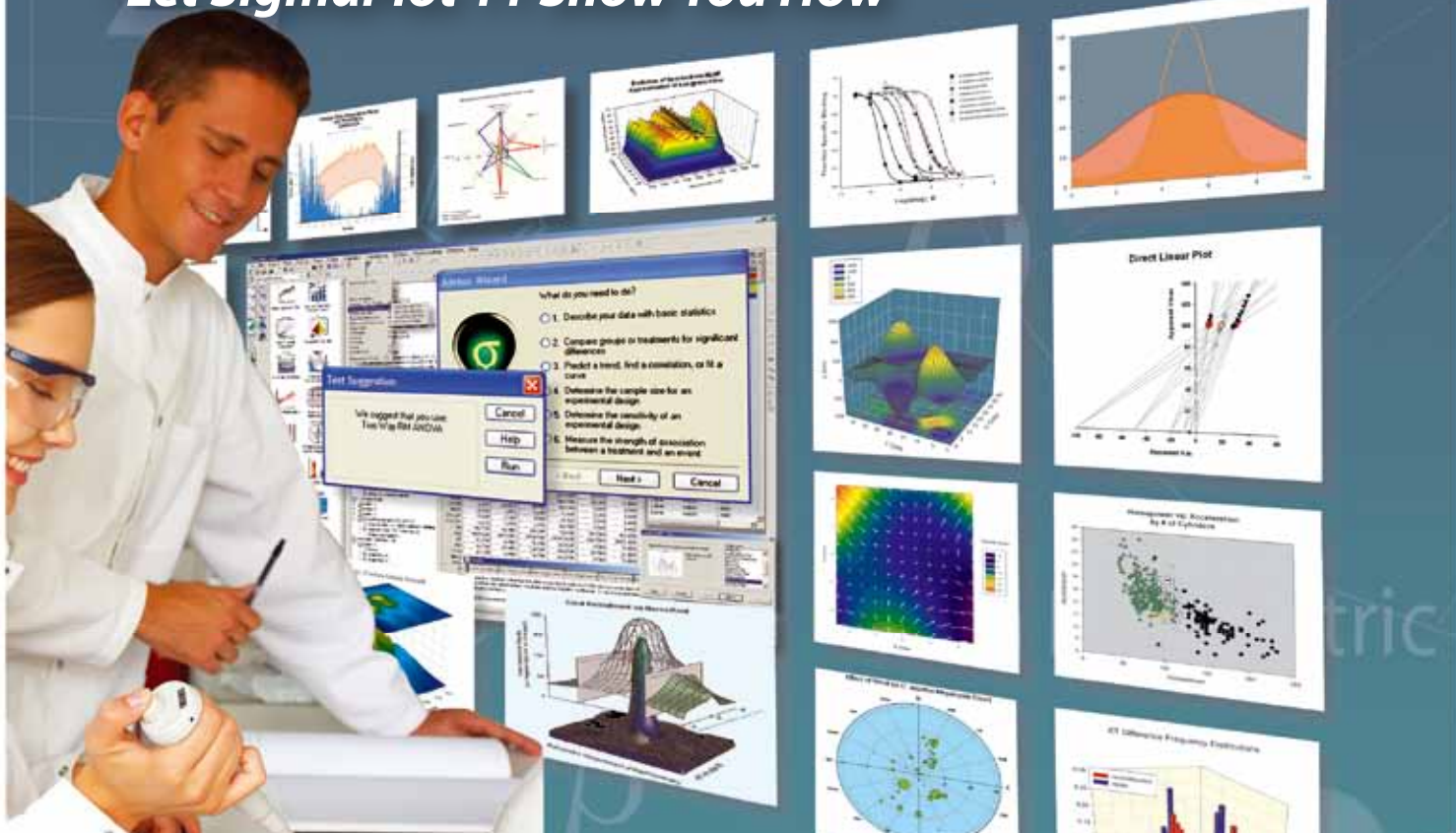
Bij Jaaps internationale activiteiten dienen naast zijn redactionele werk vooral IASA en EURANDOM genoemd te worden. Jaap bracht geruime tijd in Wenen door om het internationale project *Methodology of Decision Analysis* van IASA te leiden. Hij was daarnaast een van de founding fathers van EURANDOM.

In juni 2000 ging Jaap met emeritaat. Bij zijn afscheidsrede werd hem de legpenning van de universiteit overhandigd, als blijk van waardering voor zijn grote verdiensten voor de TU/e. Tevens werd hem een special issue aangeboden van *Statistica Neerlandica* (Volume 54, nr.2, July 2000) als blijvend monument voor zijn werk. Hij bleef actief in EURANDOM, waar hij een belangrijke rol achter de schermen speelde als adviseur van de opeenvolgende directies. Ook functioneerde hij nog een aantal jaren als onderzoeker en co-leider van EURANDOM's wachtrijprogramma, tot ziekte hem noopte deze bezigheden te beëindigen.

Jaap Wessels is geruime tijd ziek geweest. De wijze waarop hij de afgelopen jaren zijn ziekte en andere persoonlijke tegenslagen heeft gedragen heeft grote indruk gemaakt op allen die met hem te maken hadden. Wij zullen Jaap niet alleen als leermeester maar vooral ook als vriend zeer missen.

Namens zijn leerlingen,
Onno Boxma en Jan van der Wal

Statistical Analysis is No Longer a Daunting Task - Let SigmaPlot 11 Show You How



The Simplest and Most Effective Way to Analyze and Graph Data !

Join more than 250,000 researchers worldwide who use SigmaPlot to easily customize every graphic detail and create compelling publication-quality graphs that clearly present their results for technical publications, presentations or the web. Gain deeper insight into your data with easy-to-use data analysis tools - from sophisticated curve fitting to advanced mathematical calculations. SigmaPlot 11 has 50 of the most frequently used statistical tests to analyze scientific data. Take advantage of the Advisor Wizard that walks the non-statistician through the analysis of their data.

With SigmaPlot 11 you can:

- > Perform over 50 of the most frequently used statistical tests in scientific research with step-by-step guidance that includes suggestion of the most appropriate statistical test, checking for violations of data assumptions and generation of an easy-to-interpret report that has minimal statistical jargon.
- > Fit your data easily and accurately to solve simple and advanced curve fitting problems including Global Curve Fitting.
- > Choose from over 100 easily customizable 2D and 3D graph types to create publication-quality graphs that communicate exactly what you want. Use pre-formatted worksheets to easily arrange your data according to the required graph type.
- > Take advantage of an updated look-and-feel across the entire application and customize your workspace according to your usage.
- > Automate repetitive tasks to save your time and effort.

"I find SigmaPlot to be the most user-friendly and advanced... It really is quite exceptional, especially for someone who doesn't have a huge statistical background."

- Carmen Cuffari, MD
Associate Professor
Department of Pediatrics



For information and availability of SigmaPlot 11 you can contact :

Cosinus Computing BV, Groenewoud 27, 5151 RM Drunen, NL
e-mail : info@cosinus.nl Tel. : +(31) 416 378 125 www.cosinus.nl

Download a free 30-day trial of SigmaPlot from www.sigmaplot.com

Systat Software Inc., A Cranes Software International Limited Company