

STAtOR

Het dieetprobleem

Het voorspellen van de leeftijd van spermasporen

Locatie-optimalisatie van vangsystemen voor plastic afval uit het water

Unprecedented Speed

Een Bayesiaanse ervaring

In memoriam Chris Klaassen

Interview met Yori Zwols

Spreeken de Nederlanders echt beter Engels dan de Zweden?



STATOR

Jaargang 26, nummer 2, juli 2025

STATOR is een uitgave van de Vereniging voor Statistiek en Operations Research (VVSOR). STATOR wil leden, bedrijven en overige geïnteresseerden op de hoogte houden van ontwikkelingen en nieuws over toepassingen van statistiek en operations research. Verschijnt 3 of 4 keer per jaar.

Redactie

Joaquim Gromicho (hoofdredacteur), Caroline Jagtenberg, Alex Kuiper, Miriam Loois, Guus Luijben (eindredacteur), Kerry Malone, Gerard Sierksma, Richard Starmans, Gerrit Stermerdink (eindredacteur), Vanessa Torres van Grinsven, en Inez Zwetsloot. Vaste medewerkers: Jelke Bethlehem, John Poppelaars en Henk Tijms.

Kopij en reacties richten aan

Prof. dr. J.A.S. Gromicho (hoofdredacteur), Universiteit van Amsterdam Faculteit Economie en Bedrijfskunde, Sectie Operations Management | Amsterdam Business School, Plantage Muidergracht 12, 1018 TV Amsterdam, stator@vvsor.nl

Bestuur van de VVSOR

Voorzitter: prof. dr. Casper Albers, db@vvsor.nl; Secretaris: secretaris@vvsor.nl; Penningmeester: dr. Rebecca Kuiper, penningmeester@vvsor.nl; Algemeen bestuurslid: dr. Marianne Jonker, db@vvsor.nl.

Voorzitters van de secties: dr. Marianne Jonker (Biometrical Section); dr. ir. Marjan van den Akker (Section for Operations Research); prof. dr. ir. Frank van der Meulen (Section Mathematical Statistics); dr. Rebecca Kuiper (Social Sciences Section); dr. Michel van de Velden (Economics Section); dr. Iris Yocarini (Section Data Science); Luise Gummi, MSc (Young Statisticians); dr. Sanne Willems (Section Statistics Communication); dr. Stéphanie van den Berg (Section Statistics Education).

Leden- en abonnementenadministratie van de VVSOR

VVSOR, Maarsbergseweg 20, 3956 KW Leersum, admin@vvsor.nl. Raadpleeg onze website www.vvsor.nl over hoe u lid kunt worden van de VVSOR of een abonnement kunt nemen op STATOR.

Voor advertenties

Prof. dr. J.A.S. Gromicho, stator@vvsor.nl STATOR verschijnt in maart, juli en december.

Uitgever

© Vereniging voor Statistiek en Operations Research
ISSN 1567-3383

Wetenschap en samenleving

Voor de meeste van onze lezers zal het vanzelfsprekend zijn dat ons werk erop gericht moet zijn om de samenleving te verbeteren. Zelfs de allervroegste oorsprong van de statistiek, die zich bezighield met het registreren van voorraden of bevolkingsaantallen, had als doel meer inzicht te verkrijgen en daarmee beter bestuur mogelijk te maken.

Maar soms kan dat schuren en moet men zich afvragen of men wel wil meewerken aan een bepaalde ontwikkeling. Het bekendste voorbeeld is de ontwikkeling van de atombom, waar de latere Nobelprijswinnaar Richard Feynman als jonge wetenschapper bij betrokken was. Na het bijwonen van de eerste proefexplosie vroegen hij en enkele collega's zich af "what a terrible thing we have made".

Dikwijls is echter niet op voorhand duidelijk dat ons werk discutabele gevolgen kan hebben. Zo is er het 'Dietprobleem' dat tot de standaardtoepassingen van lineaire programmering behoort. Het is ontwikkeld om, met minimale kosten, een verantwoord dieet voor gevangenen vast te stellen. Dat was lovenswaardig, maar Sophie Huiberts laat in haar column zien dat deze toepassing ook raakt aan de ellendige voedselsituatie in Gaza.

De rest van dit nummer gaat over onderwerpen die op dit ogenblik niet beladen lijken te zijn, maar we kunnen niet weten of dit in de toekomst mogelijk anders zal zijn. Nadenken over waarmee we bezig zijn kan daarom nooit kwaad.

Bart Kapiteijn et al vertellen hoe het voorspellen van de leeftijd van spermasporen kan worden verbeterd met Machine Learning en Anne-Fleur Dijkhorst kijkt naar het probleem van het wegvangen van plastic afval uit oppervlaktewater. John Einmahl en Yi He hebben bestudeerd welke er grenzen zijn aan menselijke sprintrecords, kortom drie zeer afwisselende onderwerpen.

Een bijzondere bijdrage is het interview dat Nicky van Foreest en Gerard Sierksma hadden met Yori Zwols, een in Groningen opgeleide OR-wetenschapper die betrokken is bij de AI-ontwikkeling van Google Deep Mind.

Verder treft u een tweetal columns aan: Gerrit Stermerdink over een misverstand dat mogelijk een Bayesiaanse oorsprong had en Jelke Bethlehem die op de voor hem kenmerkende grondige manier bekijkt of de opvatting dat de Nederlanders beter Engels spreken dan de Zweden wel juist is.

Als verenigingsnieuws een In Memoriam voor Chris Klaassen, hij heeft veel voor de statistiek en de VVSOR betekend, en mededelingen over de Young Statisticians en een NGB Young OR Day.

Wij wensen u een ontspannen vakantie en veel leesplezier!

De STATOR-redactie



INHOUD

2 Wetenschap en samenleving

4 **Het dieetprobleem** – column | Sophie Huiberts

10 **Het voorspellen van de leeftijd van spermasporen met behulp van machine learning voor forensisch onderzoek** | Bart Kapiteijn, Elise van Swol, Laila Abdelbaey en Lindy Schermer

14 **Locatie-optimalisatie van vangsystemen voor plastic afval uit het water** | Anne-Fleur Dijkhorst

18 **Unprecedented Speed: How Data Heterogeneity Improves Ultimate Sprint Record Estimates** | John Einmahl, Yi He

20 NGB Young OR Day

21 **Een Bayesiaanse ervaring** – column | Gerrit Stermerdink

22 **In memoriam Chris Klaassen (1949 - 2025)** | Ronald Does, Mathisca de Gunst, Aad van der Vaart

24 **"Ik ben van nature een optimalisatiepersoon"** | Nicky van Foreest en Gerard Sierksma

28 **Spreek de Nederlanders echt beter Engels dan de Zweden?** – column | Jelke Bethlehem

31 Young Statisticians



Het dieetprobleem

Het dieetprobleem is populair in studieboeken over lineair programmeren. Maar deze berekening heeft kwalijke toepassingen die lange tijd zijn genegeerd.

Lineair programmeren (LP) kent ontelbare toepassingen overal in de maatschappij. Dat maakt het makkelijk om gevarieerde en realistische voorbeelden te vinden die zowel gepast als verantwoord zijn. Desondanks beginnen de meeste studieboeken over lineair programmeren met hetzelfde voorbeeld. Gegeven is een tabel met drie tot vijf voedingsmiddelen, inclusief prijs en twee of drie voedingswaarden. Een arme student, restaurantmanager of kok is benieuwd naar het goedkoopste voedingspatroon zodat een standaard man zijn aanbevolen dagelijkse hoeveelheden van alle voedingsstoffen binnenkrijgt. Elk voedingsmiddel krijgt een niet-negatieve variabele toegewezen. De prijsvector geeft de doelfunctie en voor elke voedingsstof is er een lineaire randvoorwaarde.

De populariteit van het dieetprobleem is begrijpelijk. In 1948 werd de eerste benchmark van de simplexmethode verricht (zie figuur 1) en het dieet-

probleem was hiervoor de invoer. Dit feit geeft het dieetprobleem een historische waarde. Didactisch is dit LP probleem ook prettig, want de verschillende voedingsmiddelen en voedingsstoffen geven houvast voor dimensieanalyse.

Ondanks deze voordelen zal ik beargumenteren dat het gebruik van het dieetprobleem in de collegezaal moet veranderen. Hiervoor gebruik ik een aantal historische toepassingen, elk met een hedendaagse relevantie. Uit deze toepassingen zal blijken dat berekeningen van deze vorm vaak moreel verontrustend, methodologisch onverantwoord en juridisch onhoudbaar zijn.

Armoede

Voor de benchmarkberekening in 1948 werd data gebruikt uit een paper van George Stigler, een invloedrijke conservatieve Amerikaanse econoom van de Chicago school. Stigler wilde bepalen wat de kosten van levensonderhoud waren, en hij modelleerde een toereikend dieet als een dieet dat voldoet aan alle aanbevolen dagelijkse hoeveelheden voedingsstoffen (Stigler, 1945). In 1945 wist men 9 voedingsstoffen te benoemen: calorieën, eiwitten, calcium,

ijzer en vitamines A, B1, B2, B3 en C. Door experimenten met voedingssupplementen wist men dat enkel deze voedingsstoffen niet toereikend zijn om gezond te blijven, en dus dat er nog meer (onbekende) voedingsstoffen bestaan. Voor zijn eigen berekening ging Stigler er van uit dat elk dieet van natuurlijke voedingsmiddelen de overgebleven voedingsstoffen in toereikende mate bevat. Hij gebruikte een tabel met in totaal 77 voedingsmiddelen.

In 1945 bestond er nog geen algoritme om LPs op te lossen. Toch wist Stigler te berekenen dat het goedkoopste geldige dieet in 1939 tussen de \$39.40 en \$39.93 per jaar zou kosten. Onder zijn voorgestelde dieet koopt iemand in een jaar 168 kilo bloem, 57 blikken gecondenseerde melk, 50 kilo kool, 130 kilo gedroogde witte bonen en 10 kilo spinazie. Het Stiglerdieet is onsmakelijk, ongezonder en eentonig. Ondanks dat hij erkent dat zijn berekende dieet niet aan te raden is, suggereert hij dat een jaarlijkse voedseluitgave rond dit prijspunt voldoende is voor een toereikende levensstijl.

Stigler was niet de eerste academicus die een goedkoop en gezond dieet probeerde te formuleren. Voedingsdeskundigen en diëtisten schatten dat een gezond dieet in 1939 ongeveer \$94 - \$115 per jaar zou kosten (Stigler, 1945). Dat prijspunt is vergelijkbaar met wat arme stedelijke huishoudens daadwerkelijk uitgaven in die tijd (Williams en Hanson, 1942). Dat Stigler's dieet zo veel goedkoper is, daar geeft hij twee verklaringen voor. Ten eerste, zegt hij dat diëtisten rekening houden met smakelijkheid, afwisseling en andere culturele aspecten. Stigler is van mening dat deze aspecten geen plaats hebben in de bepaling van een wetenschappelijk minimaal dieet. Ten tweede, schrijft hij dat "belasting-gefinancierde bureaucraten en professoren ook een ander motief zouden kunnen hebben voor hun gedrag."



Figuur 1: Computers van het Mathematical Tables Project (1938-1948) waar de simplexmethode voor het eerst werd getest (Grier, 2013).

Stigler was van mening dat economen zich moesten uitspreken over regelgeving en beleid. Om een indruk te krijgen van de politieke context van zijn LP-berekening, kunnen we kijken naar een artikel dat hij een jaar later schreef (Stigler, 1946). Het federale minimumloon in de VS was destijds zo laag dat het effectief niet bestond, en er gingen stemmen op om het te verhogen tot 60-75 cent per uur. Hoewel Stigler vond dat extreme armoede moet worden uitgeroeid, dacht hij dat een minimumloon niet de juiste aanpak was. Hij had een theoretisch argument dat een minimumloon slecht zou zijn voor de economie en dat een minimumloon niet effectief zou zijn om de armste huishoudens te helpen.

In de laatste paragraaf krijgen we een blik in Stigler's waarden en motivaties wanneer hij stelt dat "we armoede willen uitroeien, grotendeels omdat het tot ondervoeding leidt." In de contrapositie: als er geen ondervoeding zou zijn, dan is er veel minder noodzaak om armoede te bestrijden. Stigler merkt op dat ondervoeding voorkomt in iedere inkomensklasse, waaruit hij concludeert dat het onbetaalbaar zou zijn om ondervoeding uit te roeien door enkel inkomens te verhogen. Hij stelt twee mogelijke wegen voor: ofwel de inkomensverhoging moet worden ondersteund met een educatieprogramma over budgetteren, ofwel de staat moet voor mensen hun voedsel uitkiezen.

We hebben hier twee papers van Stigler bekeken, uit 1945 en 1946. In elkaars context is het duidelijk dat beide werken dezelfde functie hebben: in plaats van inkomens te verhogen, wilde Stigler liever dat de armste Amerikanen hun voedingsuitgaven tot wel 50% zouden verlagen. Zijn artikelen waren een poging om deze uitkomst te bewerkstelligen. Wat begon als een berekening van het minimum noodzakelijke voedselbudget werd daarin getransformeerd tot het maximum voorgeschreven budget. De oorsprong van het dieetprobleem was niet om mensen te helpen met hun huishoudboekje, maar vooral om economische herverdeling te voorkomen.

Opsluiting

Een studieboek voor managementwetenschappen (F. S. Hillier en M. S. Hillier, 2011) bevat een opgave over een keukenmanager in Sing Sing gevangenis. In de opgave wil ze een combinatie van melk, bonen en sinaasappels vinden die zo min mogelijk kost, maar toch voldoet aan het wettelijk minimum voor vitamines B1, B3 en C. De taak aan de student is om te berekenen wat ze de gevangenen moet voeren.

Sing Sing is een echte gevangenis, 20 kilometer

ten noorden van New York. Mensen die in Sing Sing zijn opgesloten krijgen nauwelijks verse groente of fruit, en ze mogen geen pakketten met voedsel ontvangen van vrienden of familie buiten (Slingerland en Fullilove, 2023).

Lineair programmeren van het dieetprobleem gebeurt ook in echte gevangenissen. In 1975 schreef Sitansu Mittra van Penn State University over zijn ervaringen met lineair programmeren in de staatsgevangenissen van Pennsylvania en New York (Mittra, 1975). Mittra begon dit project in 1971 en omschreef het LP als het "huisvrouwenprobleem", met het doel om een dieet met toereikende voedingsstoffen te vinden voor opgesloten mensen. Het dieetprobleem voor gevangenissen werd ook in andere staten opgepakt. In 1984 schrijven Hy, Feig en Regoli een artikel over LP voor studenten in justitie administratie (Hy, Feig en Regoli, 1984). In dit lesmateriaal zijn alle voorbeelden variaties op het dieetprobleem.

De introductie van operations research in de gevangeniskeuken had snel gevolgen voor de beschikbaarheid en kwaliteit. In een recent rapport (Soble, Stroud en Weinstein, 2020) beschrijft de organisatie Impact Justice een tijdlijn van gevangenissen. In het midden van de vorige eeuw hadden gevangenissen veel vrijheid over wat ze kochten en kookten; het was waarschijnlijker dat ze te veel voedsel aanboden dan te weinig. Volgens Impact Justice slaat dit in de jaren zeventig om. Enerzijds door een mentaliteitsverschuiving richting een straffende rol voor gevangenissen, anderzijds doordat de staten hun menu's centraal gingen plannen om kosten te besparen en te voldoen aan voedingsrichtlijnen.

Impact Justice schrijft over hedendaagse gevangenismaaltijden dat ze "ontworpen zijn om de benodigde calorieën en voedingsstoffen te voorzien, maar niks meer dan dat". Ondanks dat de maaltijden op papier toereikend zijn qua voedingswaarden, is de praktijk anders. Op een enquête antwoorden 94% van de respondenten dat ze niet genoeg eten kregen om zich vol te voelen. Het rapport beschrijft dat sommige mensen zo hongerig zijn dat ze aan gevangenisbendes deelnemen of seks ruilen voor extra eten.

De invoering van LP-menuplanning was ongeveer gelijktijdig met het moment dat Amerikaanse gevangenissen minder voedsel begonnen te voorzien. Hoewel we nog niet met zekerheid kunnen aantonen of dit een oorzakelijk verband is, weten we een ding zeker: het LP-model heeft er niet voor heeft gezorgd dat iedereen toereikende toegang tot voedsel heeft.



Figuur 2: Markt in al-Maghazi vluchtelingenkamp in de Gazastrook. Foto: Suhair Karam / IRIN 2010

Bezetting

In 2007 sloot Israël alle grenzen van de Gazastrook voor mensen en goederen. De minister van defensie besloot dat er nog slechts 106 vrachtwagens met humanitaire goederen per werkdag naar binnen mochten, waaronder 77 basis-voedingsmiddelen. Een aantal maanden later deden de Israëlische ministeries van defensie en gezondheid een stel berekeningen om te bepalen of dit genoeg voedsel was om ondervoeding te voorkomen. In een tweetal presentaties, gedateerd 1 en 27 januari 2008, lezen we over hun werk om een minimum bestaanspakket te berekenen. Dit minimumpakket laat een hoeveelheid "voeding toe die voldoende is voor levensonderhoud zonder dat er ondervoeding optreedt" (Coordinator of Government Activities in the Territories, 2008; Gisha, 2012).

In de eerste presentatie staat een tabel die voorschrijft, voor 12 verschillende leeftijds/seksegroepen, tot op 0,01 gram nauwkeurig hoe veel bloem, rijst, groente, fruit, melk, vlees, olie en suiker iedereen dient te eten. Niet alleen zijn deze getallen uiterst nauwkeurig, maar ook staan er getallen in deze tabel die gelijk zijn aan nul. Deze nullen zijn de hoeveelheid olie voor kinderen van 2 tot 3 jaar oud en meisjes, suiker voor kinderen tussen 4 en 6 jaar oud en olie voor meisjes tussen 11 en 24 jaar oud. Het onderschrift van deze tabel geeft aan dat ze deels zijn gebaseerd op Israëlische consumptie-standaarden en niet op minimum levensonderhoud, maar wat dit betekent wordt niet specifiek uitgelegd.

Verder concludeert de eerste presentatie dat 104 vrachtwagens vol eten per werkdag nodig zijn. Dit zou veel te veel zou zijn om met 106 vrachtwagens totaal in alle levensbehoeftes te voorzien, waaronder ook schoonmaakmiddelen, medische

noodzakelijkheden, huishoudelijk textiel en apparatuur, en zo veel meer. Daarom worden in de tweede presentatie nieuwe getallen bepaald waarin enkel standaarden voor minimum levensonderhoud zijn meegenomen. De nieuwe berekening concludeert dat 88,8 vrachtwagens per werkdag toereikend zijn voor het levensonderhoud van de Palestijnen in Gaza. Dit getal ligt heel dicht bij het blokkadeplan, dat 89 vrachtwagens met voedsel per dag voorstelde.

Waar kwamen deze cijfers vandaan? Het is onwaarschijnlijk dat de tabel consumptiestatistiek bevat, die zijn nooit meer dan 2 cijfers significant. De getallen in de presentatie zijn 3 tot 5 cijfers significant. Deze onrealistische nauwkeurigheid suggereert dat de voedselpakketten in de presentatie prescriptief zijn, niet descriptief. Verder zijn metingen van voedingswaarden ook vaak onnauwkeurig, ze hebben regelmatig foutmarges van meer dan 10%. Dit geeft een indicatie dat de voorschriften zijn opgesteld door een algoritme, niet door een deskundige met relevante vakkenis.

In het berekende voedselpakket zien we meerdere getallen die gelijk zijn aan hun ondergrens. Ook in de tweede presentatie komen de hierboven genoemde nullen terug in het becijferde voedselpakket. Aangezien olie nodig is om te koken, zouden we niet verwachten dat dit getal ooit gelijk is aan nul in een realistisch voedselpakket. Naast de woordelijke omschrijving van "het minimum bestaanspakket" zijn deze nullen de sterkste aanleiding die we hebben dat het voorgeschreven dieet mogelijk is verkregen als hoekpunt in een lineair programmeringsprobleem. En we hebben gezien dat 88,8 heel dicht in de buurt van 89 is, dus het is mogelijk dat de blokkade en de 106 limiet is bepaald aan de hand van een soortgelijke berekening.

De Israëlische mensenrechtenorganisatie Gisha heeft de blokkade omschreven als collectieve bestraffing van de Palestijnse bevolking. Deze omschrijving is te herkennen in een toespraak van de Israëlische premier Ehud Olmert:

We zullen geen humanitaire crisis laten ontstaan, maar we zijn niet van plan hun leven makkelijker te maken. En hoe moeilijker hun levens, humanitaire schade daargelaten, we zullen ze niet toestaan om een aangenaam leven te leiden (Strand, 2014).

De blokkade is erin geslaagd om mensen hun leven moeilijk te maken. De directeur van UNRWA omschreef de voedselsituatie van de burgerbevolking: "veel families eten eens per twee dagen. Iedereen heeft honger; niemand verhongert" (El-Haddad en Schmitt, 2012). Physicians for Human Rights schrijft

dat tijdens de blokkade het tunnelnetwerk tussen de Gazastrook en Egypte sterk is gegroeid, en dat de import van goederen door deze tunnels een belangrijk reden is waarom de meeste mensen niet verhongerden (Wright en Robinson, 2010).

Hoogleraar Israëlisch recht Ayal Gross en advocaat Tamar Feldman beschreven in een paper in 2015 de juridische achtergrond van de blokkade (Gross en Feldman, 2015). Volgens hen is het onderscheid tussen "honger" en "verhongering" belangrijk voor Israël's interpretatie van het oorlogsrecht. Zo stelt Aanvullend Protocol I van de Geneefse Conventies dat vechtende partijen een plicht hebben om burgers te beschermen, inclusief toegang tot toereikend voedsel. Israël koos om deze plicht beperkt te interpreteren: zolang er genoeg goederen over de grens komen zodat het Palestijnse volk niet hoeft te verhongeren, is dat toereikend en is er geen schending van het oorlogsrecht.

De functie van de berekening in de presentatie is daarmee duidelijk: om een zo streng mogelijke blokkade in te stellen zonder dat de rechtbank dit een oorlogsmisdaad zou noemen. Wat ogenschijnlijk was geformuleerd als een berekening van de minimale noodzakelijke hoeveelheid voedsel, werd door de blokkade getransformeerd tot de maximale hoeveelheid beschikbaar voedsel.

Er is nog veel onduidelijk over deze blokkade. Het kleine beetje informatie dat naar buiten is gekomen is het resultaat van jarenlange inspanning door mensenrechtenactivisten. En hoewel we geen sluitend bewijs hebben dat de blokkade was opgesteld door middel van het dieetprobleem, is het duidelijk dat de twee berekeningen vergelijkbaar zijn in doel en opzet.

Ook vandaag zijn berekeningen aan de grens van uithongering belangrijk. In de huidige crisis, die door gerenommeerde onderzoekers wordt bestempeld als genocide (Laarhoven, Peek en Walters, 2025), speelt voedsel een grote rol. Israël laat al geruime tijd geen voedsel of andere humanitaire hulpgoederen meer binnen in de Gazastrook. Israël verdedigde zijn acties maandenlang door te stellen dat er voldoende voedsel aanwezig is, terwijl de VN deze bewering "belachelijk" noemde (Nichols, 2025).

Op het moment van schrijven (19 mei) heeft Netanyahu bekend gemaakt dat er weer een minimale hoeveelheid voedsel naar binnen zal worden gelaten. Dit gebeurt wegens druk vanuit bondgenoten rond afbeeldingen van zichtbare verhongering (Gritten, 2025). Hoe veel hulpgoederen precies naar binnen zal worden gelaten is nog onduidelijk, en het Israëlische leger stelt dat politici nog moeten

besluiten (Angel en Al-Mughrabi, 2025). Noodhulpcoördinator Tom Fletcher van de VN pleit juist voor het weghalen van alle quotas voor humanitaire goederen.

Het fundamentele recht op voedsel

Mensenrechten zoals opgesteld in het Internationaal Verdrag inzake Economische, Sociale en Culturele Rechten gelden voor iedereen: rijk of arm, man of vrouw, misdadiger of onschuldige, levend onder bezetting of in een vrij land. Twee belangrijke mensenrechten uit dit verdrag zijn het recht van iedereen op een toereikende levensstandaard, inclusief toereikend voedsel, en het fundamentele recht om vrij te zijn van honger (Verenigde Naties, 1967).



Figuur 3: Staten die het Internationaal Verdrag inzake Economische, Sociale en Culturele Rechten hebben ondertekend (lichtgroen), hebben ondertekend en geratificeerd (donkergroen) en niet hebben ondertekend (grijs). Afbeelding: Wikipedia gebruiker Louperivois

In 1999 publiceerde het Comité voor Economische, Sociale en Culturele Rechten hun algemene aanbeveling inzake het recht op toereikend voedsel. In deze uitwerking beschrijft het comité onder andere dat het recht op toereikend voedsel een breed recht is, en dat het begrip van "toereikend voedsel" ook sociale, economische, culturele, klimatologische, ecologische en andere omstandigheden omvat. Voedsel moet beschikbaar, bereikbaar, vrij van schadelijke stoffen en cultureel acceptabel zijn. In het bijzonder schrijft het comité:

Het recht op toereikend voedsel zal daarom niet worden geïnterpreteerd in een nauwe of beperkte zin die het gelijkstelt aan een minimumpakket calorieën, eiwitten en andere specifieke voedingsstoffen (VN Comité inzake Economische, Sociale en Culturele Rechten, 1999).

Deze zin spreekt ons direct aan. In het dieetprobleem doen we precies datgene wat het comité afraadt: we stellen dat een vector voedingsstoffen een "toereikend dieet" is zolang het aan een aantal eenvoudige lineaire randvoorwaarden voldoet, en

daarmee interpreteren we het recht op voedsel in de beperktst mogelijke zin.

Conclusie

In 1951 wilde George Dantzig een aantal kilo afvalen. Toen hij probeerde om het dieetprobleem op te lossen voor zijn eigen menuplanning, reageerde Ray Fulkerson met

Ben je helemaal gestoord? We lossen modellen op om optimale activiteitenschema's te vinden voor anderen om op te volgen, niet voor onszelf (Dantzig, 1990)

Zoals we in de voorgaande paragrafen hebben gezien was Fulkerson's uitspraak profetisch: het dieetprobleem wordt nog steeds voornamelijk gebruikt voor de onderdrukking van "de ander."

Berekeningen van een "minimaal dieet onder voedingswaarde randvoorwaarden" blijken regelmatig een functie te hebben in het bepalen van een bovengrens op beschikbaar voedsel. Deze bovengrenzen hebben slachtoffers, en als operations researchers en academici moeten we ons tot deze context verhouden. Ik geloof dat dit ethische vraagstukken oproept waarvan de invulling specifiek is aan OR als vakgebied. Ik beschrijf er twee.

Ten eerste dienen we respectvol te zijn richting de slachtoffers van het dieetprobleem, ongeacht of ze aanwezig zijn in onze collegezaal of niet. Als we over het dieetprobleem praten zonder deze slachtoffers te erkennen, geven we dan wel voldoende rekenschap aan onze wetenschappelijke en maatschappelijke verantwoordelijkheden?

Ten tweede hebben we een verantwoordelijkheid richting onze studenten. We onderwijzen methodes die ze kunnen gebruiken in hun werkzame leven. Wanneer een methode slechte gevolgen kan hebben, dan moeten we overwegen of deze wel in de studie thuishoort. En als we toch besluiten om een bepaalde methode te onderwijzen, dan hebben we misschien wel een plicht om de studenten te waarschuwen voor de risico's die daarmee verbonden zijn.

Het doel van dit essay is niet om het dieetprobleem uit de collegezaal te verbannen. Wel wil ik dat we de functie van dit LP heroverwegen. In plaats van te dienen als voorbeeld van het nut en gebruik van lineair programmeren, kunnen we het dieetprobleem misschien beter gebruiken om een gesprek te starten over ethische en morele aspecten van deze wiskunde.

Literatuur

- M. Angel en N. Al-Mughrabi. „Netanyahu says Israel will control Gaza as pressure mounts on aid“. In: *Reuters* (mei 2025). URL: <https://www.reuters.com/world/middle-east/israel-kills-20-palestinians-airstrikes-after-agreeing-let-some-aid-into-gaza-2025-05-19/>.
- Coordinator of Government Activities in the Territories. *Food Consumption in the Gaza Strip - Red Lines*. Jan 2008. URL: <http://www.gisha.org/UserFiles/File/publications/redlines/red-lines-presentation-eng.pdf>.
- G. B. Dantzig. „The Diet Problem“. In: *Interfaces* 20.4 (aug 1990), p. 43–47. DOI: 10.1287/inte.20.4.43.
- Gisha. Reader: „Food Consumption in the Gaza Strip - Red Lines“. 2012. URL: <https://www.gisha.org/UserFiles/File/publications/redlines/redlines-position-paper-eng.pdf>.
- D. A. Grier. *When Computers Were Human*. Princeton University Press, dec 2013. ISBN: 9781400849369. DOI: 10.1515/9781400849369.
- D. Gritten. „Israel lets aid into Gaza after 11-week blockade but UN calls it 'drop in ocean'“. In: *BBC* (mei 2025). URL: <https://www.bbc.com/news/articles/c3d4kz8p00eo>.
- A. Gross en T. Feldman. „We Didn't Want to Hear the Word Calories: Rethinking Food Security, Food Power, and Food Sovereignty-Lessons from the Gaza Closure“. In: *Berkeley Journal of International Law* 33 (2015), p. 379–441. DOI: 10.15779/Z38JW1T. URL: <https://lawcat.berkeley.edu/record/1126816>.
- L. El-Haddad en M. Schmitt. *The Gaza Kitchen: a Palestinian cultural journey*. Just World Books, 2012. ISBN: 9781682570968.
- F. S. Hillier en M. S. Hillier. *Introduction to management science*. 4de ed. The McGraw-Hill/Irwin series. Operations and decision sciences. McGraw Hill Higher Education, 2011. ISBN: 978-0-07-809660-0.
- R. J. Hy, D. Feig en R. M. Regoli. „Linear programming in criminal justice administration“. In: *American Journal of Criminal Justice* 8.2 (jun 1984), p. 195–213. DOI: 10.1007/bf02885865.
- K. van Laarhoven, E. Peek en D. Walters. „Zeven gerenommeerde wetenschappers vrijwel eensgezind: Israël pleegt in Gaza genocide“. In: *NRC* (mei 2025). URL: <https://www.nrc.nl/nieuws/2025/05/14/zeven-gerenommeerde-wetenschappers-vrijwel-eensgezind-israel-pleegt-in-gaza-genocide-a4893293>.
- S. S. Mittra. „Application of Operations Research Methods To Correctional Problems“. In: *Correctional Psychologist* 2.2 (jun 1975), p. 169–179. DOI: 10.1177/009385487500200206.

- M. Nichols. „Israel says plenty of food in Gaza, UN says that's ridiculous“. In: *Reuters* (apr 2025). URL: <https://www.reuters.com/world/middle-east/israel-says-planty-food-gaza-un-says-thats-ridiculous-2025-04-01/>.
- I. Slingerland en R. Fullilove. „Opinion: Bringing Fresh Produce to New York's Correctional Facilities“. In: *City Limits* (mrt 2023). URL: <https://citylimits.org/opinion-bringing-fresh-produce-to-new-yorks-correctional-facilities/>.
- L. Soble, K. Stroud en M. Weinstein. *Eating Behind Bars: Ending the Hidden Punishment of Food in Prison*. 2020. URL: <https://impactjustice.org/wp-content/uploads/IJ-Eating-Behind-Bars.pdf>.
- G. J. Stigler. „The cost of subsistence“. In: *Journal of Farm Economics* 27.2 (1945), p. 303–314.
- G. J. Stigler. „The economics of minimum wage legislation“. In: *The American Economic Review* 36.3 (1946), p. 358–365.
- T. Strand. „Tightening the Noose: The Institutionalized Impoverishment of Gaza, 2005-2010“. In: *Journal of Palestine Studies* 43.2 (feb 2014), p. 6–23. DOI: 10.1525/jps.2014.43.2.6.
- Verenigde Naties. *Internationaal Verdrag inzake Economische, Sociale en Culturele Rechten*. 1967. URL: https://treaties.un.org/doc/Treaties/1976/01/197601032009-57%20PM/Ch_IV_03.pdf.
- VN Comité inzake Economische, Sociale en Culturele Rechten. *E/C.12/1999/5: General Comment No. 12 on the right to adequate food: Substantive issues arising in the implementation of the international covenant on economic, social and cultural rights*. Mei 1999. URL: <https://www.ohchr.org/en/documents/general-comments-and-recommendations/ec1219995-general-comment-no-12-right-adequate-food>.
- F. M. Williams en A. C. Hanson. *Money Disbursements of Wage Earners and Clerical Workers in Five Cities in the Pacific Region, 1934-36*. Bulletin of the United States Bureau of Labor Statistics 639. US Government Printing Office, 1942. URL: <https://fraser.stlouisfed.org/title/4152>.
- F. Wright en S. Robinson. „Humanitarian Minimum: Israel's Role in Creating Food and Water Insecurity in the Gaza Strip“. In: *Physicians for Human Rights-Israel* (dec 2010). URL: <https://reliefweb.int/report/occupied-palestinian-territory/%E2%80%9C9Chumanitarian-minimum%E2%80%9D-israel%E2%80%99s-role-creating-food-and-water>.

Sophie Huiberts is onderzoeker bij CNRS in het LIMOS lab aan de Université Clermont Auvergne. Haar proefschrift over lineair programmeren won de Gijs de Leve prijs en de Stieltjesprijs. E-mail: sophie@huiberts.me



Het voorspellen van de leeftijd van spermasporen met behulp van machine learning voor forensisch onderzoek

Bart Kapiteijn, Elise van Swol, Laila Abdelbakey en Lindy Schermer

Het weten van de leeftijd van biologische sporen, zoals vingerafdrukken, bloed- of spermasporen, is cruciaal in forensisch onderzoek. Deze leeftijd is echter vaak moeilijk te bepalen. In opdracht van het Amsterdam Universitair Medisch Centrum (Amsterdam UMC) is er onderzoek gedaan naar een methode voor het voorspellen van de leeftijd van spermasporen, aan de hand van machine learning. Zie ook (Abdelbakey e.a., 2024).

Fysiek bewijs van biologische sporen, zoals spermasporen, is van belang voor het vaststellen of een delict heeft plaatsgevonden, het identificeren van de dader of het vrijpleiten van een verdachte. Biologische sporen kunnen helpen om het slachtoffer en de verdachte met elkaar te verbinden, het slachtoffer aan de plaats delict te koppelen, of de verdachte aan de plaats delict te relateren.

Het veroordelen van daders van seksueel geweld is vaak een uitdaging door een gebrek aan bewijs en ooggetuigen. De leeftijdsbepaling van sporen is daarom cruciaal in forensisch onderzoek voor het reconstrueren van een tijdlijn van de gebeurtenissen. Ook kunnen hiermee de verklaringen van slachtoffers worden versterkt. De leeftijd van een spermaspoor is echter moeilijk te bepalen door de

complexe variatie in de chemische samenstelling.

Fluorescentiespectroscopie is een onderzoeksmethode die kan worden gebruikt voor kwantitatieve metingen van chemische producten. Deze methode is gebaseerd op de fluorescentie-eigenschappen van een monster (*sample*) en analyseert de fluorescentie van een molecuul. Fluorescentiespectroscopie zorgt voor de excitatie van elektronen in moleculen van bepaalde verbindingen. Hierdoor stralen de moleculen licht uit. Deze stralen worden opgevangen, waardoor metingen kunnen worden verricht. Een fluorometer is een apparaat dat wordt gebruikt om fluorescentie te meten. Het apparaat geeft de intensiteit en distributie van golflengtes van een emissiespectrum na excitatie bij een specifieke golflengte van licht.

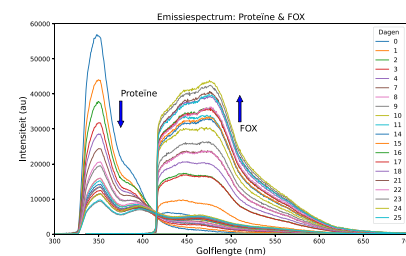
Bij fluorescentie van spermasporen wordt er gekeken naar proteïne en fluorescerende oxidatieproducten (FOX). FOX ontstaan bij het oxidatieproces van bepaalde vetten en eiwitten. Met de tijd vindt er een afname van de fluorescentie-intensiteit van proteïne plaats en een toename van de intensiteit van FOX. Er zijn inmiddels meerdere onderzoeken uitgevoerd door het Amsterdam UMC naar de leeftijdsbepaling van biologische sporen aan de hand van fluorescentiespectroscopie. Metingen zijn gedaan met behulp van een draagbaar apparaat, ontwikkeld door het Amsterdam UMC. De betrokken onderzoekers hebben een verouderingsfunctie ontwikkeld

die de gedragsverhouding van proteïne-tot-FOX in de tijd in kaart brengt. Hierdoor kon de leeftijd worden voorspeld voor de trainset tot vijftien dagen met een *Mean Absolute Error* (MAE) van 1,9 dagen en voor de testset tot elf dagen met een MAE van 1,0 dagen (Achetib e.a., 2023). De MAE-score is een maat die de gemiddelde afwijking tussen de voorspelde en werkelijke leeftijd aangeeft. In dit geval betekent een MAE van 1,0 dagen dat het model een fout heeft van gemiddeld één dag per voorspelling. Hoe lager de MAE, hoe preciezer de voorspellingen. Bij de verouderingsfunctie zijn zes opeenvolgende metingen nodig voor een voorspelling, waarbij elke dag één meting wordt uitgevoerd.

In dit onderzoek staat de volgende vraag centraal: is er een beter data-analysemodel – dan het model van het Amsterdam UMC – voor het maken van een leeftijdsschatting van spermasporen voor forensische toepassingen? Hierbij houdt 'een beter model' in: voorspellen tot meer dagen en/of een nauwkeurigere schatting en/of het nodig hebben van minder meetmomenten voor het maken van een voorspelling.

De data

In dit onderzoek wordt gebruik gemaakt van dezelfde dataset als in het eerdere onderzoek van het Amsterdam UMC. Hierdoor kunnen de resultaten van beide onderzoeken met elkaar vergeleken worden. De data bestaan uit de emissiespectra van FOX en proteïne van tien proefpersonen waarvan drie samples per persoon, voor maximaal 27 meetdagen na plaatsing van een spoor. In Figuur 1 is het emissiespectrum van een sample voor alle meetdagen te zien. In deze figuur is de daling van de intensiteit van proteïne en toename van de intensiteit van FOX over de dagen goed te zien.



Figuur 1: Fluorescentie-emissiespectrum van een sample over tijd.

Data-analyse voor feature engineering

Een groot onderdeel van dit onderzoek is data-analyse. Data-analyse wordt uitgevoerd vanwege variaties in de data. Zo bevat elke sample een andere hoeveelheid proteïne en FOX en zijn er onzekerheden door de manier van meten. Het emissiespectrum van alle samples per dag volgt een vast patroon. Naarmate de tijd vordert, neemt de intensiteit van proteïne af en die van FOX toe. Door data-analyse kunnen eigenschappen worden geïdentificeerd, die kunnen bijdragen aan het voorspellen van de leeftijd van spermasporen. Het ontwikkelen van variabelen (*features*) maakt het mogelijk om belangrijke informatie in de data te onderscheiden van minder belangrijke informatie.

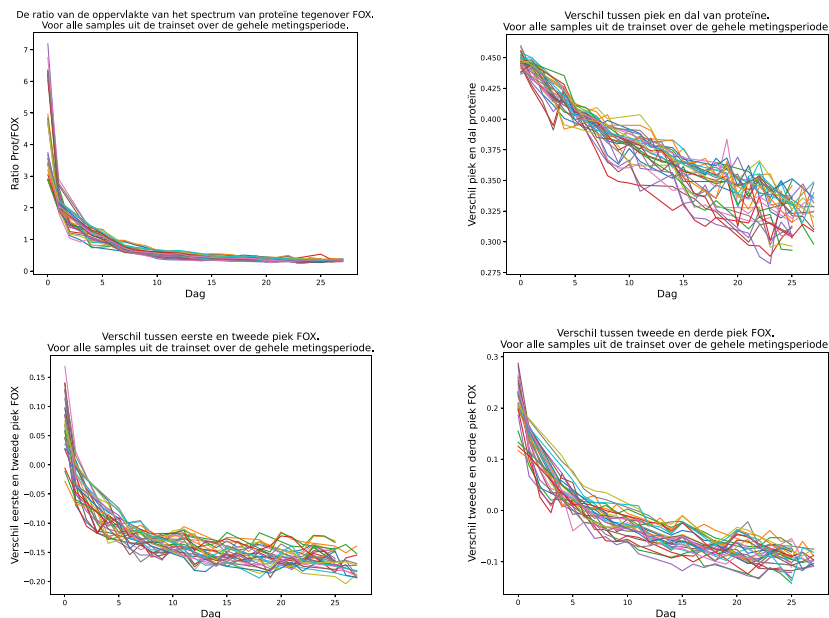
Uit de data-analyse zijn de volgende voorspellingen de features naar voren gekomen:

1. De ratio van de oppervlakte van het emissiespectrum van proteïne en FOX (dit staat centraal in het onderzoek van het Amsterdam UMC);
2. Het verschil tussen piek en dal van het genormaliseerde emissiespectrum van proteïne;
3. Het verschil tussen eerste (bij 424,8 nm) en tweede (bij 446,7 nm) pieken van het genormaliseerde emissiespectrum van FOX;
4. Het verschil tussen tweede en derde (bij 475,6 nm) pieken van het genormaliseerde emissiespectrum van FOX.

Bovenstaande features zijn gebruikt bij de modellering. De features zijn visueel weergegeven in Figuur 2. Het is te zien dat alle features dalen in de tijd. De feature van het verschil tussen de piek en het dal van de proteïne lijkt lineair af te nemen (rechts-boven), terwijl de andere features een exponentiële afname lijken te vertonen.

Voorspellen aan de hand van een Support Vector Machine (SVM)

Voor het voorspellen van de leeftijd van spermasporen zijn Support Vector Machine-modellen getraind. Een SVM kan namelijk relatief goed omgaan met een kleine als complexe dataset. De modellen zijn gevalideerd aan de hand van de MAE-score. Zo kunnen de prestaties van de modellen direct vergeleken worden met die van het model van het Amsterdam UMC.



Figuur 2: De gebruikte features bij de modellering. Iedere kleur vertegenwoordigt een ander sample.

Er zijn vier modellen getraind:

- Model 1 (alle features van spermatozoën van 0 t/m 27 dagen oud met één benodigd meetmoment). MAE-score op de testset: 2,28 dagen.
- Model 2 (alle features van spermatozoën van 0 t/m 15 dagen oud met één benodigd meetmoment) presteert het best op de testset: MAE-score: 1,27.
- Model 3 (alle features van spermatozoën van 0 t/m 27 dagen oud met twee benodigde meetmomenten). MAE-score op de testset: 2,25.
- Model 4 (één feature van spermatozoën van 0 t/m 27 dagen oud met één benodigd meetmoment), het eenvoudigste model met één variabele (het verschil tussen piek en dal van het eiwit spectrum van proteïne). MAE-score op de testset: 2,11.

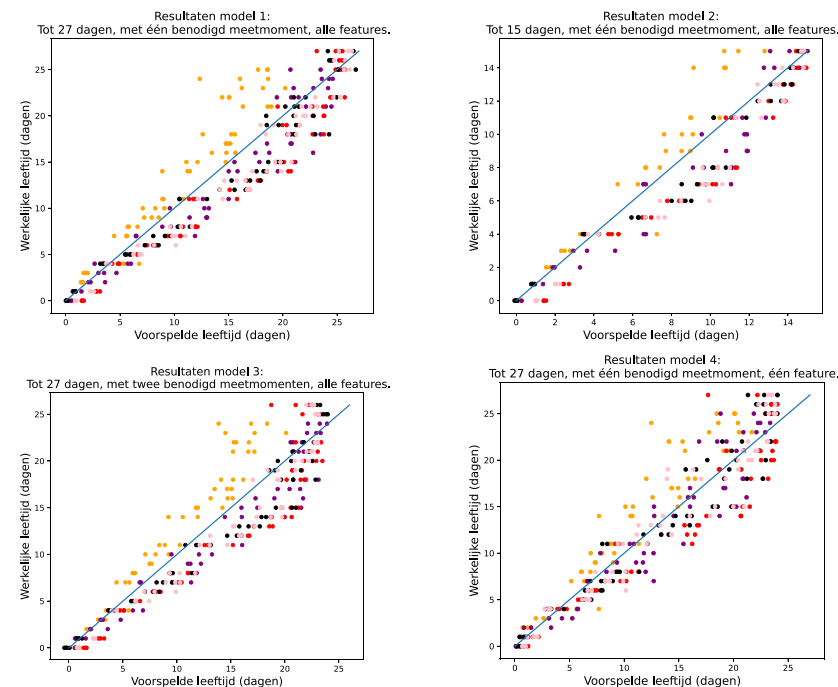
De prestaties van de getrainde modellen liggen dicht bij elkaar, behalve die van model 2. Dit is te zien in Figuur 3 doordat de voorspelde en werkelijke waarde bij de vier modellen vergelijkbaar rond de diagonale lijn liggen. Model 3 suggereert dat het wijzigen van het aantal dagen input (twee dagen in plaats van één dag) nauwelijks een verbetering oplevert ten opzichte van model 1. Model 2 heeft daarentegen door het verkleinen van het aantal voorspelde da-

gen (15 dagen in plaats van 27) wel een positieve invloed op de score in vergelijking met model 1. Bij de meeste features is er tot 15 dagen nog een duidelijk afnemend patroon zichtbaar naarmate de dagen toenemen. Na dag 15 stagneert deze afname. Het model kan daarom oudere samples minder goed voorspellen. Tot slot onderscheidt model 4 zich door zijn eenvoud. Het maakt namelijk gebruik van slechts één variabele en is daarmee het minst complexe model.

Over het algemeen zijn de MAE-scores op de testdata hoger dan die op de trainingdata. Dit kan duiden op overfitting of op een te beperkte hoeveelheid testdata. Als de trainingdata erg beperkt zijn, kan het model niet goed generaliseren, wat resulteert in slechtere prestaties op de testdata. Bij een klein aantal datapunten kan elke waarneming een grote invloed hebben op de score, wat de nauwkeurigheid van het model verder kan verminderen.

Conclusie en vergelijking met het model van het Amsterdam UMC

Modellen 2 en 4 worden aanbevolen voor gebruik. Model 2 voorspelt tot en met 15 dagen en behaalde de beste score van alle onderzochte modellen.



Figuur 3: De werkelijke en de voorspelde leeftijd van de spermatozoa in de testset voor de vier modellen. De kleuren vertegenwoordigen de samples van de donoren uit de testset.

Model 4 maakt gebruik van slechts één feature om voorspellingen te doen van 0 tot en met 27 dagen en vereist slechts één meetmoment. Dit is het eenvoudigste model vanwege het gebruik van slechts één variabele. De huidige foutmarge van de voorspelling is helaas nog te groot om als juridisch bewijs in de rechtszaal te gebruiken. De modellen kunnen echter wel als hulpmiddel dienen om een schatting te maken van de leeftijd van spermatozoa.

De modellen voorspellen tot meer dagen dan het model van het Amsterdam UMC. Bovendien hebben de modellen slechts één meetmoment nodig, in plaats van zes achtereenvolgende meetmomenten. De nauwkeurigheid van de voorspelling van model 2 is vergelijkbaar met die van het model van het Amsterdam UMC.

Toekomstig onderzoek

In een toekomstig onderzoek kan een eenvoudige lineaire regressie worden toegepast met één verklarende variabele op model 4, aangezien de gebruikte

feature een lineair verband vertoont met de leeftijd van een sample. Daarnaast wordt aanbevolen om meer data te verzamelen, omdat dit de kans op overfitting verlaagt. Het verzamelen van meer data biedt tevens de mogelijkheid om complexere modellen te gebruiken, waardoor de leeftijd van spermatozoa nauwkeuriger zou kunnen worden geschat.

Literatuur

- L. Abdelbakey e.a. *Leeftijdsschatting van spermatozoa (Hbo-scriptie)*. Hogeschool van Amsterdam. 2024.
- N. Achetib e.a. *Towards Onsite Age Estimation of Semen Stains Using Fluorescence Spectroscopy*. 2023. DOI: <https://doi.org/10.3390/s23136148>.

Bart Kapiteijn, Elise van Swol, Laila Abdelbakey en Lindy Schermer zijn studenten Toegepaste Wiskunde aan de Hogeschool van Amsterdam. Tijdens hun bedrijfsfase onderzochten zij de leeftijdsvoorspelling van spermatozoa. Dit artikel is een samenvatting van dat onderzoek (Abdelbakey et al., 2024). E-mail: bart.kapiteijn.kapiteijn@hva.nl, elise.van.swol@hva.nl, laila.abdelbakey@hva.nl, lindy.schermer@hva.nl



Locatie-optimalisatie van vangsystemen voor plastic afval uit het water

Anne-Fleur Dijkhorst

Optimalisatie kan worden gebruikt om de wereld een stukje schoner te maken als we de beste locaties kiezen om zo veel mogelijk plastic afval met vangsystemen uit een waternetwerk op te vangen. Hiervoor omschrijven we de stroom van plastic afval door een waternetwerk als een Markov Decision Process en hebben we een "Plastic Waste Flow Capturing Location Model" opgesteld om de beste oplossing te vinden als we de omgevingsfactoren als input geven.

Hoe vangen we zo veel mogelijk plastic afval voor het in de zee terechtkomt?

Vervuiling door plastic afval is een groot probleem, dat steeds meer op de agenda komt. Zodra plastic afval het land verlaat en in het water terechtkomt, wordt het steeds moeilijker om het op te ruimen. Het brengt schade toe aan omliggende ecosystemen, breekt af tot microplastics en komt vervolgens vaak via drinkwater en voedsel in ons lichaam terecht. Het is aangetoond dat een aanzienlijk deel van het plastic via kanalen en rivieren in de oceaan terechtkomt (Meijer e.a., 2021). Daarom is het belangrijk om het afval dichterbij de bron in deze kanalen en rivieren op te vangen.

Noria is een tech scale-up uit Delft die afrekenet met urban plastic soup. Bij Noria houden ze van de schoonheid van de natuur in het algemeen en water in het bijzonder. Die schoonheid willen ze behouden voor toekomstige generaties, met behulp van hun passie voor het ontwikkelen van duurzame oplossingen. Niet alleen pakken ze het plastic aan, ze richten zich op al het zwerfafval. Dit doen ze door middel van onderzoek naar de locatie en de stroom van plastic afval in lokale waterwegen en door het plastic afval te verwijderen met behulp van twee verschillende soorten vangsystemen die zij zelf hebben ontwikkeld.

Het actieve vangstelsysteem, de "CirCleaner" is een bewegend rad met scheppen die het plastic afval uit het water optillen en in een container laten vallen. De beweging van het rad wordt aangestuurd via stroomtoevoer vanaf de kant of door een klein zonnepaneel aan te sluiten.

Het passieve vangstelsysteem, de "CanalCleaner" is een drijvende kooi met een opening aan één zijde, bedekt met een houten vlot. Het plastic afval drijft dus vanzelf naar binnen als de stroming van het water in de richting van de opening van het vangstelsysteem staat. Binnen in het passieve vangstelsysteem zijn schotten geplaatst zodat dat het plastic naar binnen toe wordt geleid maar zodat de terugstroom van plastic naar de buitenkant van het systeem wordt geblokkeerd.

Om deze vangsystemen optimaal te benutten, is het belangrijk om te bepalen welke locaties het meest geschikt zijn om ze te plaatsen. Het doel is om zoveel mogelijk plastic afval uit het water te verwijderen; dat betekent dat we moeten proberen de locaties te vinden waar de systemen het meeste plastic kunnen opvangen. Ons doel is dan ook om een optimalisatiemodel te ontwerpen dat helpt bij de besluitvorming over de locaties van opvangsystemen voor plastic afval, zodat er met een gegeven budget voor de vangsystemen zoveel mogelijk plastic afval kan worden opgevangen.

Flow capturing location problem

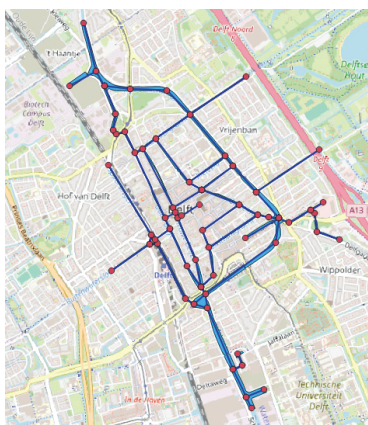
Maar hoe kunnen we dan weten waar de vangsystemen het meeste plastic afval kunnen opvangen? Noria heeft al een plastic-hotspot-voorspellingsmodel ontwikkeld waarbij voorspeld wordt op welke plekken veel plastic afval vast zou komen te zitten. Hierin werd gekeken naar de oorsprong van plastic afval, de windrichting, de geometrie van de waterweg (bijvoorbeeld scherpe hoeken) en vegetatie langs de kust of in het midden van het

water. Soms werd ook naar de stroming van het water gekeken als er bijvoorbeeld een pompsysteem in de buurt was, maar op basis van tests en gesprekken met lokale experts leek de beweging van het plastic afval in grachten en kanalen vooral afhankelijk van de wind.

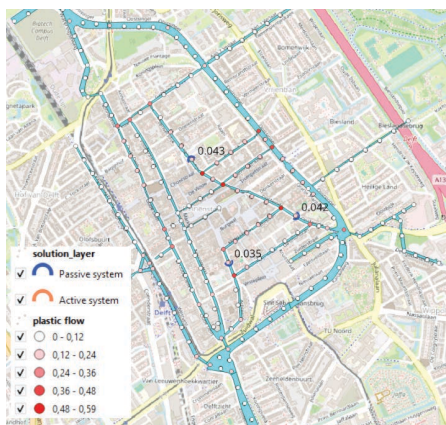
Alleen weten waar plastic-hotspots verwacht worden is helaas nog niet genoeg om te kunnen weten waar het meeste plastic afval kan worden afgevangen. Doordat de vangsystemen van Noria op hun vaste plek blijven liggen en niet door het water heen en weer bewegen moet het plastic namelijk eigenlijk juist vóórdat het is komen vast te zitten worden opgevangen. Dit is hetzelfde probleem als het "Flow Capturing Location Problem" (FCLP) (Berman, Krass en Wei Xu (1995a); Berman, Krass en Wei Xu (1995b)). Dit probleem komt al voor in de literatuur en werd vooral gebruikt om de optimale locatie van tankstations te vinden. Daarbij is het namelijk ook het geval dat een tankstation op een vaste plek is en dat je de meeste voertuigen 'opvangt' als ze dichtbij de stroom van voertuigen liggen. Een groot verschil met die setting is wel dat voertuigen vaak een kleine afwijking van hun route te maken omdat ze graag willen tanken. Het plastic afval zal helaas niet van de route afwijken om te worden opgevangen en moet dus echt precies langs de gekozen locatie stromen. Dat komt meer overeen met een situatie zoals het vinden van de optimale locaties voor het monitoren van voertuigen (snelheidscontroles of gewichtscontroles van vrachtwagens).

Er zijn twee manieren om het FCLP te formuleren, namelijk door de volledige paden van de stroom te beschrijven (het begin, de route ertussen, en het einde), of door het probleem op te schrijven als een Markov Decision Process (MDP). Bij een MDP hoeven niet alle paden bekend te zijn, maar is het genoeg als er initiële kansen en overgangskansen worden gedefinieerd. We hebben geen data van de precieze paden van plastic afval van begin tot eind. Gelukkig kunnen we wel een inschatting maken van de kans dat er op een bepaalde locatie plastic in het water terecht komt en waar plastic vanaf een bepaalde plek naartoe zal gaan drijven. Dit doen we op basis van omgevingsfactoren zoals plekken waar veel mensen tijd doorbrengen langs het water (op bankjes en in parken) en op basis van de windrichting.

Het bestaande model uit de literatuur hebben we aangepast door nog twee aspecten toe te voegen: de mogelijkheid om een "gevoelig gebied" aan te geven, en de mogelijkheid om meerdere soorten locaties (vangsystemen) te gebruiken. In deze context bete-



Figuur 1: Waterwegen van Delft als netwerk.



Figuur 2: Voorbeeld van een oplossing van het model. De plastic flow is aangegeven als een fractie van de totale plastic flow in het gebied. De getallen bij de drie vangsystemen geven aan hoe veel plastic flow er in de vangsystemen wordt verwacht.

kent een gevoelig gebied dat het extra belangrijk is om in dit gebied plastic afval uit het water op te vangen. Denk bijvoorbeeld aan een natuurgebied, of een gebied waar het water (en dus het plastic afval) steeds dichterbij de zee zal komen. De verschillende eigenschappen van de twee soorten vangsystemen worden ook meegenomen in het model, zoals de aanschafkosten en eventuele restricties in waar ze geplaatst mogen worden, bijvoorbeeld door hun formaat of door gebrek aan stroomtoevoer voor het actieve systeem.

Het model weegt vervolgens door middel van gewichten drie verschillende doelstellingen af: als eerste doel om zo veel mogelijk plastic af te vangen, als tweede doel om dit zo veel mogelijk weg te halen uit eventuele gevoelige gebieden en als derde om het budget dat wordt besteed zo goed mogelijk af te wegen met de hoeveelheid extra plastic die per extra vangstelsysteem wordt afgevangen.

Omgevingsfactoren en data

Het model heeft data van een aantal omgevingsfactoren nodig om het op een bepaald gebied te kunnen toepassen. Zo is er natuurlijk een netwerk

van waterwegen nodig. Er is een aantal handige trucjes om in het geografische informatiesysteem QGIS een netwerk te krijgen van lijnen van de ligging van de waterwegen. Als eerste heeft dit netwerk alleen knopen op een kruispunt waar het water in meerdere richtingen beweegt, zoals te zien is in Figuur 1. Vervolgens worden er extra knopen toegevoegd op de rechte stukken in het water door middel van een Python-script. De maximale afstand tussen deze knopen moet als input worden gegeven en bepaalt de precisie van de oplossing van het model.

Verder moet de plastic stroom van het model beschreven worden. Voor de initiële kans dat plastic in het water terecht komt, gebruiken we data over de mogelijke bronnen van plastic afval dichtbij de kust. Afhankelijk van hoeveel plastic afval bronnen er in de buurt van een locatie zijn, wijzen we een bepaalde kans toe aan een knoop, ingeschat op basis van veldwerk. Voor de overgangskansen tussen de aangrenzende knopen is de windrichting per dag van een jaar gebruikt in de buurt van het gebied waar het model wordt toegepast. Verder wordt data over de geometrie van het waternetwerk (scherpe hoeken), woonboten en vegetatie gebruikt om op basis van veldwerk de kans in te schatten dat

het plastic vast komt te zitten en niet meer verder stroomt. Als het model voor een nieuw gebied wordt gebruikt kan deze open source geografische informatie worden geïmporteerd in QGIS en wordt het automatisch door onze geschreven Python scripts verwerkt in de implementatie van het model.

Op deze manier hebben we twee case studies gedaan waar we het model hebben toegepast op de waterwegen in Delft en in Groningen. Zo konden we kijken hoe het model werkt onder verschillende input data door de verschillende omgevingsfactoren in Delft en Groningen.

Hoe werkt het model?

In Figuur 2 is een voorbeeld te zien van hoe het model er voor de gebruiker uit zou zien. Nadat alle data over omgevingsfactoren is ingeladen moet een budget voor de vangsystemen worden gekozen. Vervolgens zal het model de optimale locatie(s) vinden. Die worden dan op de kaart weergegeven met een aparte kleur voor beide vangsystemen. De richting van de opening van de vangsystemen is ook weergegeven op de kaart, zodat de gebruiker weet hoe het vangstelsysteem moet worden neergelegd. De verwachte plastic afval stroom langs elk punt is ook weergegeven met een kleur. De punten waar het meeste plastic lijkt te stromen zijn op de kruispunten van het water. Daar kan helaas geen vangstelsysteem worden geplaatst omdat er boten langs moeten kunnen varen en omdat de richting van het vangstelsysteem op een kruispunt onduidelijk is, wat als voorwaarde in het model wordt meegenomen.

Discussie

In het onderzoek hebben we een uitgebreide gevoeligheidsanalyse gedaan, zowel voor de case study in Delft als in Groningen. Daarmee kunnen we zien welke input data de oplossing van het model het sterkst lijkt te beïnvloeden. Hieruit blijkt dat het vooral belangrijk is om nog eens te kijken naar de kans dat het plastic vast blijft zitten op een plek door vegetatie in het water. Het plastic verplaatst namelijk een stuk minder ver als deze kans hoog is, wat in de zomer het geval lijkt te zijn als het plastic nauwelijks beweegt tussen kroos en waterlelies in het centrum van Delft. Hiervoor kunnen tests gedaan worden met veldwerk. Het is dan ook een goed idee om nog extra tests te doen om te verifiëren of het plastic afval aan het oppervlak van het water inderdaad voornamelijk met de windrichting meegaat.

Het is lastig om een optimale oplossing te vinden als het model wordt gebruikt voor een groot

gebied met veel mogelijke locaties voor de vangsystemen en wanneer er genoeg budget is om veel vangsystemen te plaatsen. Dan zijn er namelijk veel variabelen en veel verschillende combinaties van locaties mogelijk om de vangsystemen te plaatsen. Dat zou gebeuren als het model wordt toegepast op bijvoorbeeld het niveau van een hele provincie in plaats van enkel een stad. Voor zulke gevallen hebben we ook een heuristiek geschreven die goed werkt op een aantal test instanties. Uiteindelijk zou het natuurlijk mooi zijn als het model gebruikt zou kunnen worden voor zulke grote gebieden, om de optimale locaties te vinden om zo veel mogelijk plastic afval in Nederland op te vangen uit het water vóórdat het in de zee en in de oceaan terecht komt.

Literatuur

- O. Berman, D. Krass en C. Wei Xu. „Locating Discretionary Service Facilities Based on Probabilistic Customer Flows“. In: *Transportation Science* 29.3 (1995), p. 276–290. URL: <http://www.jstor.org/stable/25768694> (bezoekt op 13-10-2023).
- O. Berman, D. Krass en C. Wei Xu. „Locating flow-intercepting facilities: New approaches and results“. In: *Annals of Operations Research* 60 (1995), p. 121–143. URL: <https://link.springer.com/article/10.1007/BF02031943>.
- L. J. J. Meijer e.a. „More than 1000 rivers account for 80% of global riverine plastic emissions into the ocean“. In: *Science Advances* 7.18 (2021), eaaz5803. DOI: 10.1126/sciadv.aaz5803. URL: <https://www.science.org/doi/abs/10.1126/sciadv.aaz5803>.

Dankwoord Dit artikel is een samenvatting van mijn masterthesis, die geschreven is in samenwerking met het bedrijf Noria onder begeleiding van Jur van Wijk van Noria en Dr. Theresia van Essen van de TU Delft. Ik wil hen graag nogmaals bedanken voor de fijne begeleiding en samenwerking en ik wil het team van Noria nogmaals bedanken voor de ontzettend gezellige en fijne tijd binnen het bedrijf tijdens mijn afstudeertraject. De volledige versie van de thesis is te lezen op <https://repository.tudelft.nl/record/uuid:28744e34-9bbc-40dc-9039-d52ed2f771e5>.

Anne-Fleur Dijkhorst Anne-Fleur heeft in 2024 de master Applied Mathematics van de TU Delft afgerond, met als afstudeer richting Discrete Mathematics and Optimization. Anne-Fleur heeft met de thesis waarop dit artikel gebaseerd is de prijs “Best Master Thesis in Applied Mathematics 2024” gewonnen, uitgereikt door het Centrum voor Wiskunde en Informatica. Ze is nu begonnen aan een PhD traject bij Erasmus Universiteit Rotterdam, gericht op operationele uitdagingen in ontwikkelingslanden en humanitaire hulp. <https://www.linkedin.com/in/anne-fleur-dijkhorst-430703192/>



Unprecedented Speed: How Data Heterogeneity Improves Ultimate Sprint Record Estimates

John Einmahl, Yi He

How fast can humans ultimately run? Using extreme value theory on a large dataset of top performances, our analysis suggests that it is highly unlikely that a man will run 100 meters under 9.49 seconds or a woman under 10.20 seconds anytime soon. Our novel method accounts for substantial individual heterogeneity, allowing us to significantly increase the effective sample size by including a wide range of athletes. By further quantifying these individual differences using multiple records for each athlete, our approach reduces statistical uncertainty by up to 88 percent compared to existing analyses.

Tomorrow's Ultimate Record

How fast can humans run? This longstanding question has intrigued not only the general public and biologists, but also data scientists. A natural approach might be to forecast future records by extrapolating historical trends. However, such ad hoc methods often rely on limited data and tend to produce

unreliable estimates. These forecasts aim to estimate a maximum possible speed, or equivalently, a minimum time, which may depend on the evolving human population.

Now, imagine a hard limit that could be achieved 'tomorrow', rather than in the distant future. We define this as the ultimate record: a performance so extreme that it would remain unbroken regardless of population growth, unless there were fundamental changes in human physiology or sport technology. By definition, this ultimate record must surpass all past achievements, assuming it has not already been reached.

For the 100-meter sprint, any such ultimate record must at least match the current world records: 9.58 seconds for men (set by Usain Bolt in 2009) and 10.49 seconds for women (set by Florence Griffith-Joyner in 1988).

Beyond the Range: Extreme Value Theory

How can we estimate the fastest possible speed that has never been observed? The answer lies in a powerful mathematical framework known as extreme value theory (de Haan and Ferreira, 2006),

which models not only the single best performance but also those approaching the ultimate limit. The underlying idea is simple: if there is any consistent pattern in the gap between the most extreme observations and the absolute record, it must follow a specific form.

By expressing sprint performance in terms of speed rather than time, and fitting this structure to data from 1991 to 2023 (including up to five performances per athlete), we find that the gap between top performances and the ultimate speed follows a predictable pattern:

$$\text{Men: Speed} \approx -0.52 \cdot (\text{Ranking})^{0.20} + 37.72,$$

$$\text{Women: Speed} \approx -0.86 \cdot (\text{Ranking})^{0.17} + 35.00,$$

where the speed is in km/h, and ranking equals to 1 for the fastest observed time, 2 for the second fastest, and so on. Figure 1 illustrates this well-fitted relationship. Each circle represents a record, while the red line indicates the fitted model. It is worth noting that the deviation of a few top records (for men) due to randomness aligns with the theoretical result, and that our linear extrapolation remains valid.

An interesting question follows: what is the ranking of the ultimate record? According to the model, it is rank 0, which represents the theoretical limit. Therefore, the intercepts of these lines estimate the fastest possible speeds, which correspond to time limits of 9.54 seconds for men and 10.29 seconds for women.

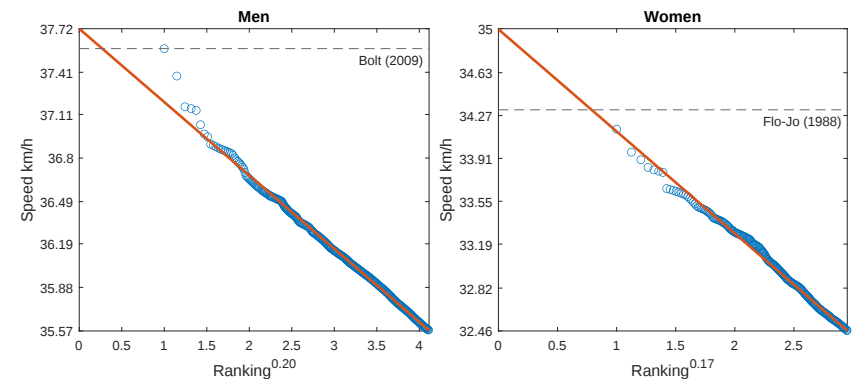


Figure 1: Linear extrapolation using extreme value theory. The x-axis represents the power-transformed record ranking. Solid lines show the fitted models, with intercepts estimating the speed limits. Dashed lines indicate current world records. Note that Florence Griffith-Joyner's 1988 record falls outside our sample period.

The Power of Heterogeneity

A key goal of statistical inference is to quantify the uncertainty in our estimates. In general, this uncertainty decreases as the amount of data increases. One of the main challenges in extreme value analysis is that only a small fraction of the most extreme data points, typically about 5%, can be used effectively. For example, Einmahl and Magnus, 2008 and Einmahl and Smeets, 2011 relied on hundreds of personal best performances, yet the effective sample size remained in the two-digit range.

To increase the sample size, our recent work (Einmahl and He, 2025) expanded the number of athletes to nearly five times, reaching a total of 5,618 male athletes and 2,528 female athletes. However, as the population size grows, variation among individuals also increases. It becomes difficult to assume that all athletes come from a single common superpopulation, if such a population exists.

Instead, we argue that rankings may reflect a combination of two components: chance, meaning the probability of achieving an exceptionally high record, and differences in skill among athletes. Since skill is not considered random, it should not contribute to statistical uncertainty. To assess how much skill varies across athletes, we use multiple records per athlete (best annual records from different years, up to five times in total) and apply a method developed in our previous work, He and Einmahl, 2024. The central idea is straightforward: if an athlete's various records tend to occupy similar

ranking positions, this consistency is more likely to reflect underlying skill rather than random luck.

Using this approach, we identify a significant degree of heterogeneity that appears similar for both men and women. This leads to an estimated 35% reduction in statistical uncertainty compared to models that treat all athletes as identical.

By combining all improvements – the larger number of athletes, the use of multiple records per athlete, and the reduction in uncertainty due to heterogeneity – we achieve an overall reduction in statistical uncertainty between 67% and 88% compared to earlier studies. As a result, we conclude with 95% confidence that the ultimate 100-meter sprint records are at best:

- 9.49 seconds for men;
- 10.20 seconds for women.

These estimates vary slightly depending on the choice of the fraction of extreme observations used for fitting. The values reported above represent the median results across a range of reasonable choices.

Looking Ahead

Our results confirm that current world records are exceptional and that the margins for improvement are small, yet improvement remains possible. These findings carry important implications for the world of athletics. The closeness of current records to the estimated performance limits underscores the remarkable achievements of athletes such as Usain Bolt and Florence Griffith Joyner. Furthermore, this research highlights the potential of advanced statistical methods for analyzing the growing volume of data in sports science. By accounting for individual heterogeneity, data scientists can better understand the factors contributing to athletic excellence and provide more accurate estimates of ultimate sport records.

References

- L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. New York: Springer, 2006.
- J. H. J. Einmahl and Y. He. *Accurate estimates of ultimate 100-meter records*. <https://doi.org/10.48550/arXiv.2502.04085>. Preprint. 2025.
- J. H. J. Einmahl and J. R. Magnus. "Records in athletics through extreme-value theory." In: *Journal of the American Statistical Association* 103.484 (2008), pp. 1382–1391.

J. H. J. Einmahl and S. G. W. R. Smeets. "Ultimate 100-m world records through extreme-value theory." In: *Statistica Neerlandica* 65.1 (2011), pp. 32–42.

Y. He and J. H. J. Einmahl. *Extreme value inference for general heterogeneous data*. <http://dx.doi.org/10.2139/ssrn.4816397>. Preprint. 2024.

John Einmahl is a research scholar and emeritus professor of Statistics in the Department of Econometrics & Operations Research at Tilburg University.
Email: j.h.j.einmahl@tilburguniversity.edu

Yi He is an Associate Professor of Data Science at the Amsterdam School of Economics, University of Amsterdam.
Email: yi.he@uva.nl



NGB Young OR Day

This year, the Young OR Day will be organized at ASML in Autumn 2025!

ASML

Learn how operations research is shaping the world's most complex supply chain, as well as optimizing the ASML machines itself. During the day, you will get a tour of the ASML Experience Center to learn more. And above all, it is an opportunity to meet fellow Young OR members!



Stay tuned for the date...

Around the summer the date of the Young OR Day will be announced to all NGB members and registration opens. While we prioritize our members, given sufficient capacity, non-members are also welcome to attend. If you would like to stay up to date, follow us on LinkedIn or keep an eye on our event website: <https://www.vvsnr.nl/operations-research/events/>.

Gerrit Stermerdink



Een Bayesiaanse ervaring

De Bayesiaanse aanpak is vrij eenvoudig uit te leggen als het bijstellen van initiële veronderstellingen aan de hand van extra binnenkomende informatie.

Je veronderstelt bijvoorbeeld dat een willekeurige volwassen persoon een lengte van 1,70 meter heeft. Maar dan kom je te weten wat het geslacht van die persoon is en vervolgens stel je die verwachting bij naar boven of beneden, mannen zijn nu eenmaal gemiddeld iets langer dan vrouwen.

Zonder dat we het expliciet in de gaten hebben passen we dat principe regelmatig toe in het dagelijks leven. Uit kleine aanvullende waarnemingen denken we dikwijls gevolgtrekkingen te kunnen maken. Dat levert bijvoorbeeld bij de politie al snel het gevaar van etnische profilering op.

Zelf had ik lang geleden een collega die wist dat ik vier kinderen had. In de jaren '70 kwam zo'n uitbundige kinderschaar hoofdzakelijk voor bij gereformeerde of katholieke gezinnen, uitzonderingen daargelaten natuurlijk.

Hij zal bij een eerste kennismaking met het feit

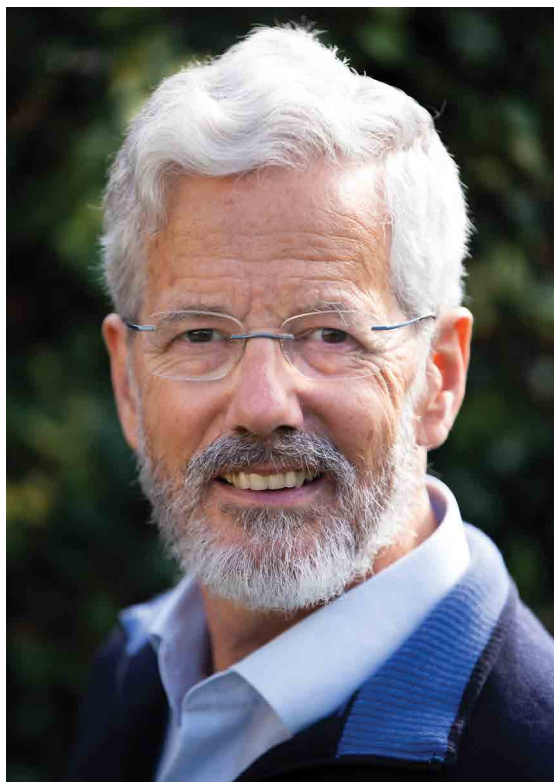
dat ik vader was ongetwijfeld het idee hebben gehad dat ik een of twee kinderen had, het meest voorkomende aantal in die tijd. Maar toen ik vertelde dat ik er vier had sloeg bij hem kennelijk het Bayesiaans denken toe, en dacht hij dat ik gereformeerd of katholiek zou moeten zijn. Een niet zo onlogisch idee, we waren tenslotte op een statistisch congres en dan ga je automatisch een beetje statistisch denken.

Maar naast die extra informatie over mijn aantal kinderen beschikte hij óók nog eens over de extra informatie dat ik bij de Katholieke Universiteit Nijmegen werkte. Zijn op grond van al die informatie bijgestelde verwachting was dan ook dat ik een mede-katholiek was. Helaas voor hem was dat niet zo, ik was eind jaren '60 een van de zeer weinige niet katholieke werknemers van de KUN.

Maar gelachen hebben we wel, vooral toen we ons realiseerden hoe vaak we in de praktijk onbewust dit soort Bayesiaanse principes toepassen.

Gerrit Stermerdink is eindredacteur van STATOR.
E-mail: gjstermerdink@hotmail.com

Column



In memoriam Chris Klaassen (1949 - 2025)

Ronald Does, Mathisca de Gunst, Aad van der Vaart

Christianus Antonius Johannes Klaassen, emeritus hoogleraar Mathematische Statistiek aan de Universiteit van Amsterdam, is op 1 april jongstleden overleden in de leeftijd van 76 jaar.

Chris was tot het eind toe actief als mathematisch en toegepast statisticus. In de laatste paar jaar werd hij daarbij gehinderd door een voortschrijdende kanker, en al eerder moest hij omgaan met het overlijden van zijn geliefde echtgenote Angeli, maar hij bleef zeer positief. Chris was en bleef een trouw deelnemer aan nationale bijeenkomsten in de statistiek, en werkte door aan zijn onderzoek, ook toen hij tot zijn huis werd beperkt. Sommige projecten

wist hij af te ronden en voor andere zocht hij samenwerkingen, zodat zijn inzichten niet verloren zouden gaan.

Chris werd geboren op 3 februari 1949 te Nijmegen, en bezocht daar het Canisius college (Gymnasium bèta, 1967) en de Radboud Universiteit (doctoraal Wiskunde, 1974). Chris vervolgde zijn opleiding met een proefschrift onder leiding van Willem van Zwet, dat hij verdedigde aan de Universiteit Leiden (1980), maar schreef in zijn functie als Wetenschappelijk Medewerker aan het Mathematisch Centrum in Amsterdam (thans Centrum voor Wiskunde en Informatica), toentertijd het centrum voor de statistiek in Nederland, nog onder leiding van Jan Hemelrijk. Tijdens zijn promotietijd kwam hij via Van Zwet in contact met Peter Bickel, wat leid-

de tot een verblijf aan de University of California, Berkeley, in 1983, en een latere samenwerking aan een boek. Na zijn promotie was Chris Universitair (Hoofd)Docent aan de Universiteit Leiden, totdat hij in 1989 werd benoemd als hoogleraar Mathematische Statistiek aan de Universiteit van Amsterdam. Daar was hij promotor voor 13 promovendi, na een eerder co-promotorschap in Leiden. Voorts was hij Wetenschappelijk Directeur van het Korteweg-De Vries Instituut en enige tijd Vice-Decaan. Chris bevorderde de vestiging van leerstoelen in de Industriële Statistiek (verbonden aan het Instituut voor Bedrijfs- en Industriële Statistiek (IBIS UvA)) en de Forensische Statistiek, en was nauw betrokken bij het opzetten van de masterprogramma's in "Stochastics and Financial Mathematics" en Forensische Wetenschappen. Op nationaal niveau was Chris onder meer programmeur van het "Statistics for Biology" project van EURANDOM (2003-2009) in Eindhoven, en lid van het eerste bestuur van het STAR-cluster (vanaf 2009). In 2014 ging hij met emeritaat, maar bleef sindsdien altijd actief in onze gemeenschap.

Het speerpunt in het onderzoek van Chris was de semi-parametrische statistiek, gericht op statistische modellen gekenmerkt door een eindig-dimensionale "parameter of interest", gecombineerd met een oneindig-dimensionale "nuisance parameter". De twee belangrijkste voorbeelden van deze nieuwe manier van statistisch modelleren waren het Cox model in de levensduur analyse (1975) en het symmetrisch locatie model, dat intensief bestudeerd was na de baanbrekende informatieberekeningen van Stein (1956). In de jaren 1980-1990-2000 kwam de semi-parametrische statistiek tot volle ontwikkeling. Chris zijn idee van "sample-splitting" (1987) voor het construeren van semi-parametrisch-efficiënte schatters wordt ook vandaag veelvuldig toegepast, bijvoorbeeld in de "double-robust machine learning", vaak zonder bronvermelding. Karakteristiek voor de fundamentele aanpak van Chris is dat hij niet alleen het constructie principe uiteenzette, maar ook bewees dat deze constructie altijd mogelijk is wanneer semi-parametrisch-efficiënte schatters bestaan. Ander veel geciteerd werk betreft semi-parametrie in uiteenlopende onderwerpen als copulas, GARCH en andere tijdreeksen, en differentiaalvergelijkingen. Chris zijn "spread inequality", welke een scherper alternatief geeft voor de Cramér-Rao ongelijkheid, was al aanwezig in Chris zijn proefschrift over locatie modellen, en werd door hem verder ontwikkeld gedurende zijn carrière. Chris was één van de vier auteurs (met Peter Bickel, Ya'acov Ritov en Jon Wellner) van het standaardwerk over semi-parametrische statistiek "Efficient and adap-

tive estimation for semiparametric models" (560 pagina's), dat in 1993 verscheen bij Johns Hopkins University Press en later door een andere uitgever opnieuw werd uitgebracht. Chris deed ook een aantal vondsten in de kansrekening, waaronder een generalisatie van Chernoff's ongelijkheid voor de variantie van een functie van een normale variabele, en Edgeworth ontwikkelingen, in de lijn van zijn promotor. Ook aan sommige kleinere projecten beleefde hij veel genoegen. Zo bewees hij dat het kwadraat van de scheefheid van een uni-modale verdeling niet groter kan zijn dan de kurtosis plus 186/125, een verbetering van een vondst van Karl Pearson uit 1916.

Chris zijn wiskundige invalshoek leverde originele oplossingen voor praktische, statistiek-gerelateerde vragen. Hij was betrokken in verschillende projecten van IBIS UvA. Zo bedacht hij een nieuw steekproef-systeem voor een bedrijf dat edelmetalen voorwerpen van een keurmerk voorziet, dat origineel genoeg was voor een publicatie in Technometrics, en een schattingsmethode voor het aantal vrije parkeerplaatsen in Rotterdam, dat eveneens een publicatie opleverde maar te ingewikkeld werd bevonden voor implementatie. De meeste impact had een project bij de Rechtbank Maastricht, dat later ook goed toepasbaar bleek om spreekuren in de polikliniek in te plannen, en verschillende promoties voor IBIS medewerkers opleverde. Als hij een wiskundig sluitende oplossing had gevonden, dan ging Chris eventuele controverse niet uit de weg, zoals in zijn onderzoek naar wetenschappelijke fraude van een hoogleraar in de psychologie aan de Universiteit van Amsterdam, waarvoor Chris nieuwe door hem ontwikkelde statistische methoden toepaste. De conclusie dat fraude was gepleegd werd niet op prijs gesteld en leidde tot een aanval op de methode.

Ondergetekenden werkten samen met Chris in verschillende rollen en omgevingen: kamergenoot, docent, co-promotor, collega, bestuurder, wandelpartner tijdens de jaarlijkse Lunteren bijeenkomst. Wij herinneren Chris in al deze rollen als hartelijk, bescheiden, oplossingsgericht en altijd positief. Hij was een fijne collega en goede vriend, voor ons, en vele andere collega's in de nationale en internationale wereld van de statistiek. Chris zelf vertelde graag, maar met een bezwaard gemoed, dat hij in zijn bestuurswerk, toch vaak een bron van wrijving, één keer ruiz had gehad. Dat was nota bene met één van ons drie, en die onenigheid is naderhand met volle overgave bijgelegd. Chris was niet alleen scherp in de wiskunde, maar ook wijs en warm in de omgang. We missen hem.



"Ik ben van nature een optimalisatiepersoon"

Nicky van Foreest en Gerard Sierksma

Het interview met Yori Zwols, in 1982 geboren in Groningen, vindt plaats online aan het begin van donderdagmiddag 28 november. Yori heeft zich gedurende ruim twee uur kunnen onttrekken aan de wereldwijde AI-race waarin ook Google DeepMind, zijn werkgever, is verwickeld. Hij zit tijdens ons interview in een rustig zaaltje bij DeepMind op Kings Cross in Londen. Door een glazen deur zijn schimmen van haastige passanten zichtbaar.

"Ik heb bij DeepMind verschillende functies gehad. Tot twee jaar geleden leidde ik een van de grootste teams in ons bedrijf. Het doel van dat team was om met behulp van software engineering expertise DeepMind's meest veelbelovende onderzoeksprojecten te versnellen. Daarna heb ik een nieuwe rol opgepakt als technisch directeur van de afdeling die Google's prominente generatieve AI-modellen voor afbeeldingen, video en audio (muziek/spraak) ontwikkelt. Voor de kenners: Veo en Imagen." Hij vervolgt: "Bij DeepMind ben ik bijna vanaf het begin manager geweest, en heb me daarnaast ook altijd deels met onderzoek bezig kunnen houden."

Met deze oud-Groningse student hadden wij het interview! De sfeer was 'onder bekenden', ontspannen dus, vrijwel steeds inhoudelijk en soms ook 'uit de losse pols'.

Yori's eerste algoritme

Al tijdens zijn gymnasiumjaren in Groningen (1993-99) spelen optimalisatieproblemen door zijn hoofd.

"Wiskunde en natuurkunde waren toen mijn favoriete vakken. Ik was 15 jaar toen ik in de bus naar huis een roosterprobleem heb opgelost voor een gymleraar. Toen ik later het begrip 'optimaal' leerde, ontdekte ik dat dat algoritme –spijtig genoeg niet noodzakelijkerwijs de 'optimale oplossing' vond. Niet zo verrassend, want toen begreep ik ook dat het probleem 'NP-hard' was."

In Groningen direct top

2004 is het afstudeerjaar aan de Rijksuniversiteit Groningen. Samen met zijn begeleider Gerard Sierksma werkt hij zijn scriptie om tot een artikel dat vijf jaar later verschijnt in het INFORMS-toptijdschrift Operations Research. Deze eerste publicatie, met de titel 'Training Optimization for the Decathlon' weerspiegelt Yori's passie voor wiskunde en sport. Ondertussen werkt hij parttime bij ORTEC-Sports (voorheen Team Support Systems) aan de ontwikkeling van software voor voetbalteam-analyses.

Waarom ben je in Groningen eigenlijk geen wiskunde gaan studeren?

"Goeie vraag. Ik heb daar een aantal open dagen bezocht: van wiskunde, natuurkunde, en econometrie. De open dag van econometrie stak er wat mij betreft bovenuit omdat die over wiskunde en statistiek met daarbij interessante toepassingen ging."

Die ervaring had ik ook. Bij econometrie gaat het over sociale, menselijke dingen en niet over ineenslopende sterrenstelsels.

"Inderdaad. Bij Econometrie en Operations Research klikte er gewoon iets. Daarom ben ik die kant opgegaan bij de Groningse universiteit."

En die klik is gebleven, niet?

"Zeker, die is gebleven. Ik hou van optimaliseren en paralleliseren. Vroeger, tijdens het lopen door de stad, dacht ik tegelijkertijd al na over het optimaal afstellen van de verkeerslichten."

Van PhD in New York naar Postdoc in Montreal

Yori vertrekt in 2004 naar Providence, Rhode Island, en studeert daar economie. In 2006 verhuist hij naar New York en haalt aldaar in 2010 zijn PhD. Daarna werkt hij twee jaar in Montreal als postdoc aan McGill University en later aan Concordia University onder de welbekende wetenschapper Vašek Chvátal.

Na je studie in Groningen koos je ervoor te promoveren. Wat bracht jou tot de wetenschap?

"Nieuwe dingen leren heb ik altijd leuk gevonden en ik wilde dat het liefst in een nieuwe omgeving voortzetten. En dat werd eerst Brown University in Providence en daarna Columbia University in New York City. Aan Columbia ontmoette ik een hele goeie professor, Maria Chudnovsky, één van de oplossters van de 'strong perfect graph conjecture'."

En bij haar in New York ben je gepromoveerd?

"Bij professor Maria Chudnovsky ben ik inderdaad gepromoveerd, met een proefschrift over 'geïnduceerde subgrafen'. Dat was een ontzettend leuke en boeiende tijd."

Was je na je PhD- en postdoc-tijd niet verzadigd van het tamelijk eenzame puur wiskundig-inhoudelijk werk?

"Tijdens mijn postdoc in Montreal was het (professionele) leven inderdaad weleens eenzaam. Ik heb mensen nodig en een whiteboard en dan samen dingen uitzoeken. Ik reisde geregeld af naar andere universiteiten om te werken met andere onderzoekers. Maar tijdens mijn postdoc zat ik soms ook wel de hele dag in m'n eentje naar zo'n whiteboard te staren. Dat gevoel kennen jullie ook vast wel."

Ja ja, dat snap ik goed. Samen aan een whiteboard, dat inspireert enorm.

Naar Londen, Google

In mei 2013 begint Yori's Google-loopbaan in Londen. Eerst als programmeur, werkend aan een intern rapportageplatform, en al snel krijgt hij de uitdagende opdracht om de vele tabellen achter dat Google-platform 'aan elkaar te knopen' en database-opdrachten te optimaliseren.

"Na twee weken in m'n eentje nadenken heb ik met grafentheorie en dynamisch programmeren de klus geklaard. Een cool probleem. Het was wat ongebruikelijk om zo lang te puzzelen in plaats van te programmeren, maar het was die investering waard: hoewel de implementatie van het algoritme wat meer tijd kostte, 'draaide' het resultaat nog lang nadat ik bij die afdeling vertrok."

DeepMind

In januari 2015 begint Yori aan zijn succesvolle tijd bij DeepMind, op dat moment net overgenomen door Google.

Eerst even dit. Je bent dan getrouwd.

"Ja. Rouba en ik leerden elkaar kennen tijdens onze PhD-tijd in New York. Zij werkte daar aan haar PhD over wachtrijtheorie. Ze is geboren in Libanon. Fijn is dat Rouba ook een baan heeft in Londen. Ze is hoogleraar Operations and Technology aan de School of Management van het University College London."

Jullie hebben samen sinds 2019 een zoontje met de naam Milo. 'MILO' het acroniem van Mixed Integer Linear Optimization?

"Haha, ja dat klopt, dat was wel een grappig toeval. Milo is ondertussen inderdaad vijf jaar."

Te jong dus om het grappje te snappen. En bevalt het wonen in het drukke Londen, of een beetje wel en een beetje niet? Ga je op de fiets naar je werk?

"Ja, het bevalt goed in Londen. We wonen ten zuiden van de Theems met een mooi park in de buurt. Fietsen in Londen van huis naar Kings Cross is zo'n 11 kilometer maar ook gevaarlijk en met veel verkeerslichten en dus veel stoppen en weer 'gas geven'. Met dat fietsen in Londen ben ik snel gestopt en ik ga nu, samen met vele andere Londenaren, met de metro."

DeepMind, het begin en het geweten

Yori vertelt over zijn beginperiode bij DeepMind, waarin hij eerst hoofdzakelijk bezig is met het programmeren en draaien van Machine Learning (ML) experimenten.

Van grafentheorie naar AI lijkt me een grote stap. Of is er nog wel een directe relatie met econometrie?

CV Yori Zwols

1993 – 1999	Prædinius Gymnasium Groningen.
1999 – 2004	Master in Econometrie en Operations Research aan de Rijksuniversiteit Groningen. Scriptie <i>Training Optimization for the Decathlon</i> onder begeleiding van Gerard Sierksma.
2001 – 2004	Part-time onderzoeker en programmeur bij ORTEC-Sports.
2004 – 2006	Master in Economie aan Brown University in Providence.
2006 – 2010	PhD in Industrial Engineering and Operations Research aan Columbia University in New York. Proefschrift <i>Forbidden Induced Subgraphs</i> onder begeleiding van Maria Chudnovsky.
2010 – 2012	Postdocs in Montréal aan McGill University (Bruce Reed) en Concordia University (Vašek Chvátal).
2013 – 2015	Software ingenieur bij Google.
2015	Publicatie boek <i>Linear and Integer Optimization: Theory and Practice</i> met Gerard Sierksma.
2015 – nu	Onderzoekingenieur bij Google DeepMind. Nu technisch directeur.

"ML is eigenlijk statistiek op grote schaal met heel veel parameters. Zoveel parameters dat het eigenlijk weer 'niet-parametrisch' is. Die grote schaal betreft grote, diepe neurale netwerken, onder meer getraind met behulp van Markovketens. In zekere zin dus niet heel anders dan econometrie en statistiek. Alleen de taal is heel anders; je denkt anders na over de concepten."

Maar wat is dat 'anders nadenken'?

"Als econometrist formuleer je een model waarvan je vermoedt dat het je data beschrijft, en dan schat je de parameters. Het model wordt zo simpel mogelijk gehouden (maar niet simpeler) en is gebaseerd op economische theorie. Je gebruikt vervolgens standaard software om het model te schatten door bijvoorbeeld een 'maximum likelihood' probleem op te lossen. Die software vindt doorgaans optimale parameterwaarden. In machine learning gebruik je modellen die zo algemeen mogelijk zijn: diepe neurale netwerken, tegenwoordig met miljarden parameters. Het optimaliseringsproces is daarbij juist één van de uitdagingen: afgezien van de enorme schaal, kun je in de praktijk ook nooit optimale parameterwaarden vinden. Sterker nog: vaak is dat niet eens wenselijk, want 'lokale optima' zijn in de praktijk robuuster dan 'globale optima'."

De robuustheid en dat contrast tussen simpele, en dus interpreteerbare, modellen in de econometrie tegenover algemene modellen in ML is belangrijk. Stel je bent computer-scientist en doet ook aan performance-analyse van algoritmen. Hoe verstandig is het als zo

iemand dergelijke algoritmen echt gaat toepassen? Dan moeten die modellen toch, laat ik zeggen, gewetensvol zijn. Daarmee bedoel ik, ML kan alles voorspellen. Dan heb je toch het geweten nodig en niet dat van de fysicus of de econometrist? Maar wel dat van de social engineer, die zegt: ik kan niet alles uitrekenen en dat moet ik ook niet willen.

"Daar ben ik het wel mee eens. ML is in feite een grote correlatiemachine. Als statisticus weten we dat correlatie iets anders is dan causaliteit, en dat je goed moet opletten correlatie niet voor causaliteit aan te zien. En dat weten ML-onderzoekers natuurlijk ook."

Om het wat duidelijker te maken. Een parabool kun je goed benaderen met een cirkel, althans het top- en daldeel. Iedereen die werkt in de approximatie-theorie weet: pas op met generaliseren.

"Inderdaad. De vraag is altijd: zijn de modellen nog betrouwbaar als je ver buiten de trainingsverdeling zit? Dit in-de-praktijk-testen is een belangrijk onderdeel van het inzetten van ML-modellen."

Zou je zo'n algoritme zo kunnen trainen dat wordt aangegeven: nou begin ik te bazelen, ik weet het niet zo goed meer, ik raak hier buiten mijn traindomein en kan maar beter nu mijn mond houden? Kun je zoiets bouwen?

"Absoluut, en dat is deels al ingebouwd in chatbots met behulp van andere losstaande ML-modellen, en dit blijft ook een interessant onderzoeksgebied. Je kunt bijvoorbeeld fact checking laten doen met een losstaand ML-model. Bij lastige vragen kan de chatbot zeggen: 'Hier kan ik niks over zeggen.'"

Een volgende carrièrestap

Interviewer vindt het 'gewetensprobleem' voorlopig wel voldoende uitgediept en keert terug naar Yori's carrière bij DeepMind. Dat Yori een creatieve wetenschapper is blijkt vooral uit de diepgang van zijn publicaties. De tijdschriften Operations Research en Nature staan reeds in zijn lijst. Daarnaast is Yori een fervent hardloper, te volgen op Strava (voor de kenner) en daarnaast nog een enthousiast pianospeler.

"Na vijf jaar onderzoekingenieur te zijn geweest, kreeg ik meer managementverantwoordelijkheid toen ik hoofd werd van het team "Core Research Engineering". Bij bijna elke belangrijke DeepMind publicatie waren wel één of meer ingenieurs van mijn team betrokken."

Die stap paste wel mooi bij jouw passie voor 'het onbekende' en 'nieuwe dingen leren'. Maar waar bleef en blijft 'de optimalisatie'?

"Ik heb een natuurlijke drijfveer om echt alles beter te willen maken, te optimaliseren. Ik probeer altijd wat groter te denken en heb zeker ook een natuurlijke neiging om organisaties te verbeteren. Was het eerst dus wiskundige optimalisatie en nu, misschien nog wel veel boeiender, 'human optimization'."

Je zou ook jouw liefde voor management en jouw liefde voor wiskunde en algoritmen kunnen, zoals je dat zelf noemt, paralleliseren, te behouden en te verenigen. Wat zou je eraan kunnen doen om beide uitdagingen te behouden?

"Dat is een kwestie van timemanagement. Niet te veel met de korte termijn bezig zijn. Proberen de dopaminehit van de snelle, makkelijke taken afronden te omzeilen, en juist tijd wegzetten voor wat je zelf belangrijk vindt. En elk half jaar eens terug- en vooruitkijken of je je tijd goed besteedt."

En zou je nog wel terug willen naar de wetenschap, de universiteit? Je kunt een point of no return passeren door te zeer het contact met de wetenschap kwijt te raken.

"Ik ben me daar inderdaad van bewust. Het kan best zijn dat ik ooit weer de wetenschap inga, maar dat is allemaal voor later. Ik zou daarvoor wel gericht in mijzelf moeten investeren. Maar nu AI zo'n hot topic is, is er voor mij genoeg interessant werk bij Google DeepMind. Als technisch directeur bij 'Generative Media' houd ik mij nu bijvoorbeeld veel bezig met de software-infrastructuur voor het trainen van grote modellen en het opleveren van die modellen op productie-infrastructuur."

Ik kan me voorstellen dat sommige lezers denken: ik zou ook wel bij Google DeepMind willen werken. Waar kunnen lezers meer informatie vinden over vacatures?

"We zijn altijd op zoek naar getalenteerde en gemotiveerde mensen! Vacatures en meer informatie over Google DeepMind is te vinden op www.deepmind.google."

Nicky van Foreest is universitair hoofddocent operations research aan de Rijksuniversiteit Groningen.
E-mail: n.d.van.foreest@rug.nl

Gerard Sierksma is emeritus hoogleraar operations research aan de Rijksuniversiteit Groningen.
E-mail: g.sierksma@rug.nl

PEILINGPRAKTIJEN



Education First

Bron: EF Education First

Sprekende Nederlanders echt beter Engels dan de Zweden?

Er zijn veel peilingen in Nederland. Dat merk je vooral als er verkiezingen zijn. Maar ook buiten de verkiezingen om vragen peilingen steeds vaker onze mening over allerlei onderwerpen.

Het principe van een peiling is dat je aan een *steekproef* van personen uit een doelgroep een *vragenlijst* voorlegt. Die vragen kunnen gaan over feitelijke zaken, gedrag en meningen. Met de verkregen antwoorden probeer je dan uitspraken te doen over de doelgroep als geheel. We noemen dat *generaliseren*. Dat kan, maar dan moet je het onderzoek wel op methodologisch verantwoorde wijze hebben opgezet en uitgevoerd. Met de *checklist voor peilingen* kun je vaststellen of dit inderdaad het geval is. Je kunt hem hier downloaden: met de url aan het einde van

deze column.

Er zijn goede en slechte peilingen. En het is niet eenvoudig om het kaf van het koren te scheiden. Deze checklist kan je daarbij helpen. Door de vragen in de checklist één voor één af te lopen en te beantwoorden, krijg je een goede eerste indruk van de kwaliteit van de peiling. Lijkt de kwaliteit goed te zijn, dan verdienen de uitkomsten van de peiling nadere aandacht. Roept het doorlopen van de checklist veel vragen op over de kwaliteit van het onderzoek en de uitkomsten, dan is het verstandig er verder geen of weinig aandacht aan te besteden.

Het is dus belangrijk om publicaties over peilingen in de media zorgvuldig te bekijken alvorens je de uitkomsten gaat gebruiken. We geven een voorbeeld van een niet zo goede peiling die ten onrechte toch veel aandacht kreeg. Dat is de *English Proficiency Index* van het taalinstituut *Education First*.

De English Proficiency Index

Elke jaar weer maakt taalinstituut *Education First* de *English Proficiency Index* (EPI) bekend. De EPI meet hoe goed de inwoners van een land de Engelse taal beheersen. Het gaat hierbij om landen waarin Engels niet de moedertaal is. Nederland en de Scandinavische landen scoren altijd erg hoog. Ze komen meestal op de eerste of tweede plaats in de ranglijst terecht. In 2020 stond Nederland weer bovenaan, gevolgd door Denemarken en Finland. De vraag is echter hoe dit onderzoek naar Engelse taalvaardigheid in elkaar zit en of de uitkomsten wel representatief zijn voor de deelnemende landen. Is de aandacht die dit onderzoek van Education First in de media krijgt, wel terecht? Tijd voor een nadere analyse!

In 2015 spraken wij Nederlanders volgens Education First (bijna) het beste Engels van alle niet-Engelstalige landen. Alleen de Zweden deden het nog beter. Dat stond in een bericht in het AD van 3 november 2015. En de NOS kopte 'Nederlander is wereldtop in Engels'.

In 2016 beheersten volgens Education First de Nederlanders de Engelse taal het best. Daarover werd uitgebreid bericht door de NOS (op 15 november 2016). De kop boven het artikel was nogal uitbundig: 'We rock! Nederlanders beheersen de Engelse taal het best'.

Is Education First een marktonderzoeksbureau?

Uit een bezoek aan de website van Education First wordt onmiddellijk duidelijk dat dit geen marktonderzoeksbureau is, maar een taalinstituut dat wereldwijd cursussen aanbiedt. Op de website kun je testen hoe goed je Engels is. Dit betekent dat

het instituut belang heeft bij de uitkomsten van de peiling: als je kennis van de Engelse taal te wensen overlaat, kun je bij hen een cursus volgen. Het is dus oppassen met de objectiviteit van de onderzoekers.

Op 17 november 2020 publiceerde Education First de uitkomsten van de English Proficiency Index (EPI) voor 2020. Het onderzoek was gebaseerd op de gegevens van 2,2 miljoen mensen in 100 verschillende landen en regio's. Dit is een hele grote groep, maar helaas is hij niet representatief. Die 2,2 miljoen mensen vormen namelijk geen aselechte steekproef. Het zijn alleen maar mensen die spontaan een Engelse taaltest hebben gedaan op de website van Education First.

Education First heeft voor al die landen de English Proficiency Index (EPI) berekend. Dat is de gemiddelde vaardigheid met het Engels voor alle volwassenen die de taaltest hebben gedaan op de website. Tot en met 2019 was de EPI een getal tussen de 0 en 100. In 2020 veranderde EF de berekening van de index. Vanaf dat jaar is het een score tussen de 0 en 800. Dat is erg onhandig als de trends door de jaren heen wilt bestuderen. Tabel 1 bevat de top drie van landen met de hoogste score voor de jaren 2015 t/m 2020.

Is er sprake van een aselechte steekproef?

De waarde van de index varieert in de periode van 2015 t/m 2019 tussen de 67,93 en de 72,16. De scores liggen dus dicht bij elkaar. In de ranglijst is er sprake van stuivertje wisselen tussen Nederland, Zweden en Denemarken.

De vraag is nu of je uitkomsten van dit onderzoek kunt generaliseren naar de gehele bevolking van de landen. Kun je de gegevens zien als een aselechte

Tabel 1: De English Efficiency Index (EPI) van de top 3 landen voor de periode 2015-2020.

Jaar	Eerste plaats	Tweede plaats	Derde plaats
2015	Zweden 70,94	Nederland 70,58	Denemarken 70,10
2016	Nederland 72,16	Denemarken 71,15	Zweden 70,81
2017	Nederland 71,45	Zweden 70,40	Denemarken 69,93
2018	Zweden 70,72	Nederland 70,31	Singapore 68,63
2019	Nederland 70,27	Zweden 68,74	Noorwegen 67,93
2020	Nederland 652	Denemarken 632	Finland 631

steekproef uit de bevolking? Helaas niet. Education First geeft ook zelf toe dat de steekproef niet representatief is. In het onderzoeksrapport staat:

'The test-taking population represented in this Index is self-selected and not guaranteed to be representative. Only those who want to learn English or are curious about their English skills will participate in one of these tests. This could skew scores lower or higher than those of the general population'.

De steekproef bestond dus uit personen die Engels wilden leren, wilden weten hoe goed hun Engels was, de website van EF hebben bezocht en daar een taaltest hebben gedaan. Dat is duidelijk geen goede afspiegeling van de bevolking van de landen.

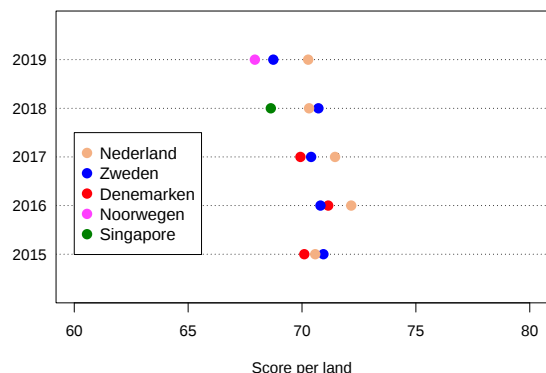
Er is nog een tweede probleem dat de representativiteit aantast. De gegevens zijn namelijk verzameld via een test op het internet. In Nederland is internet niet zo'n probleem. Volgens cijfers van Eurostat (het statistische bureau van de EU) heeft 98% van de Nederlandse huishoudens internet (cijfers van 2019). Dat is bijna iedereen. En ook in Scandinavië hebben veel huishoudens internet: in Zweden 96% en in Denemarken 95%.

In veel andere landen was de dekking van internet echter veel lager. Volgens Eurostat had bijvoorbeeld in Bulgarije maar 75% internet, in Griekenland was het 79% en in Portugal (71%) was het al niet veel beter (cijfers van Eurostat over 2019). In dit soort landen konden dus heel wat mensen de taaltest niet doen. En de mensen die hem wel deden, waren vast beter opgeleid dan de mensen die geen internet hebben. Er was dus sprake van een te rooskleurig beeld. Dat geeft Education First ook toe in het onderzoeksrapport:

'People without Internet access would be automatically excluded. In parts of the world where Internet usage is low, we would expect the impact of an online format to be strong. This sampling bias would tend to pull scores upward by excluding poorer and less educated people'.

Merk ook op dat de drie landen die meestal in de top drie staan (Nederland, Zweden en Denemarken) ook de landen zijn met de hoogste internetdekking (95% tot 98%).

De English Proficiency Index van Education First (2015–2019)



Figuur 1: De English Proficiency Indices van de top 5 landen liggen dicht bij elkaar.

Nederland, Zweden of Denemarken?

Welke land beheerst de Engelse taal nu het best? Is het Nederland, of Zweden, of misschien Denemarken? Om daar wat meer inzicht in te krijgen, hebben we een grafiek gemaakt van de gegevens in de tabel 1. Het zijn de scores van de beste drie landen in de jaren van 2015 t/m 2019. Door de verandering in de definitie van de EPI in 2020 konden we die scores niet meenemen in de analyse. Zie de dotplot in figuur 1.

Het eerste dat opvalt is dat de scores van Nederland, Zweden en Denemarken erg dicht bij elkaar in de buurt liggen, zo rond de 70. Nederland heeft drie keer de hoogste score (in 2016, 2017 en 2019). Zweden scoort twee keer het beste (in 2015 en 2018). Denemarken scoort niet het beste, maar ligt wel heel dicht bij Nederland en Zweden in de buurt.

Je kunt je afvragen waar de jaarlijkse schommeling in scores vandaan komt. Het lijkt onwaarschijnlijk dat de taalvaardigheid van een land van jaar op jaar ineens verandert. Het zal eerder liggen aan een andere samenstelling van de (niet-representatieve) steekproef.

Je kunt je ook afvragen of de verschillen per land echte (significante) verschillen zijn of dat het slechts gaat om ruis van de steekproef is. Om hier achter te komen, zou je onzekerheidsmarges moeten berekenen. Dat hebben de onderzoekers helaas niet gedaan.

Bedenk hierbij dat de omvang van de totale steekproef (2,2 miljoen personen) weliswaar heel groot is, maar er doen ook veel landen (100) mee, zodat de steekproef per land aanzienlijk kleiner is. Er zouden dus best nog behoorlijke onzekerheidsmarges kunnen zijn.

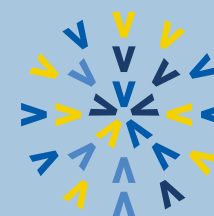
Conclusie

Kortom, dit onderzoek laat veel te wensen over. Het is eerder een reclamecampagne voor taalcursussen. Jammer dat media als het AD en de NOS er aandacht aan hebben besteed zonder het bericht eerst te controleren op correcte toepassing van de methodologie. EF doet dit onderzoek jaarlijks met dezelfde onderzoeksopzet. In 2015 berichtten het AD en de NOS al over dit onderzoek. En ook toen was er al kritiek. Daarbij maakte AD ook nog de fout te melden dat in 2015 Denemarken bovenaan stond. Dit was niet juist. Het was Zweden.

Checklist voor peilingen:

<https://peilingpraktijken.nl/wp-content/uploads/2021/05/Checklist-voor-peilingen-versie-4.pdf>

Jelke Bethlehem is expert op het gebied van steekproeven, vragenlijsten en weergave van onderzoeksresultaten.
E-mail: mail@jelkebethlehem.nl



Young Statisticians

After the success of our Statistics Café on February 21st, we want to thank everyone who joined us and warmly invite you to our future events. As we plan for organising more fun and engaging events in the future, we would like to invite you to follow us on our Instagram page or reach out to us with any suggestions for events you'd be interested in seeing from us in the future!

@youngstatisticians

youngstatisticians@vwsor.nl

vwsor.nl/young-statisticians

