

STATOR

Een Ig Nobelprijs voor het opgooien van muntjes

Ik erger me!

Citizen Science in Transient Astronomy: The Black Hole Finder App

Dynamische afsprakenroosters in de zorg

In memoriam: Prof. dr. Willem Schaafsma

In memoriam: dr. Alessandro Di Bucchianico

**Programme of the Annual Meeting of the VVSOR 2025:
80 years of statistics and operations research in the
Netherlands: past, present and beyond**

Statistische prijzen en de erfenis van Stella Cunliffe

**Optimalisatie voor de sturing van de grootste pompen
van Nederland**

Explainable Decision Making

De Amerikaanse presidentsverkiezingen beter in beeld



STATOR

Jaargang 26, nummer 1, maart 2025

STATOR is een uitgave van de Vereniging voor Statistiek en Operations Research (VWSOR). 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 Stemerding (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@vwsor.nl

Bestuur van de VWSOR

Voorzitter: prof. dr. Casper Albers, db@vwsor.nl; Secretaris: secretaris@vwsor.nl; Penningmeester: dr. Rebecca Kuiper, penningmeester@vwsor.nl; Algemeen bestuurslid: dr. Tsega Kahsay Gebretekla en dr. Marianne Jonker, db@vwsor.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); Elena Petridou, BSc (Young Statisticians); dr. Sanne Willems (Section Statistics Communication); dr. Stéphanie van den Berg (Section Statistics Education).

Leden- en abonnementenadministratie van de VWSOR

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

Voor advertenties

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

Uitgever

© Vereniging voor Statistiek en Operations Research
ISSN 1567-3383

De volgende 80 jaar

Het zal in augustus van dit jaar tachtig jaar geleden dat de VWSOR werd opgericht. Een belangrijk initiatiefnemer was David van Dantzig, die tijdens zijn onderduiktijd lang had nagedacht over de mogelijkheden om wiskunde dienstbaar te kunnen laten zijn voor de samenleving. Naast de VVS, zoals de vereniging toen nog heette, kwam vanuit die gedachte ook het Mathematisch Centrum, nu CWI, tot stand. Het is dan ook geen toeval dat de belangrijkste prijs van de VWSOR naar hem is genoemd. Mensen die de tachtig jaar bereiken worden gewoonlijk gerekend tot de leeftijd der zeer sterken. Maar in de statistiek ligt dat iets genuanceerder: de Royal Statistical Society is van 1834, de American Statistical Association van 1839 en het International Statistical Institute werd 1885 opgericht na een reeks voorafgaande International Statistical Congresses die in 1853 startten. Wij moeten dus nog even om een dergelijke respectabele leeftijd te bereiken.

Op onze Annual Meeting zal op tachtig jaar VWSOR worden teruggeblikt, maar belangrijker nog zal ook worden gekeken naar de toekomst. Het programma vindt u in dit nummer, naast verder verenigingsnieuws en een tweetal IM's voor overleden prominente leden.

Dit nummer heeft, zoals u van ons gewend bent, een afwisselende inhoud. Van een Ig Nobelprijs voor het opgooien van muntjes, door Eric-Jan Wagenmakers en František Bartoš, tot een App waarmee iedereen kan meewerken aan het detecteren van Zwarte Gat, door Fiorenzo Stoppa. Passend bij het toekennen van drie Prijzen tijdens de aanstaande AM is een artikel van Richard Starmans over enkele internationale statistische prijzen. Essentieel voor de Nederlandse waterhuishouding, en dus onze droge voeten, zijn grote pompen; Kieran van Gaalen et al. beschrijven hoe de sturing daarvan geoptimaliseerd kan worden. Iedereen heeft wel eens meegemaakt dat het soms grote moeite kost om prachtige optimaliseringen geaccepteerd te krijgen. Het goed uitleggen daarvan kan makkelijker worden door hier rekening mee te houden bij het model: Jannis Kurtz gaat uitgebreid op in dit Explainable Decision Making. Afsprakenroosters in de zorg geven vaak irritatie, regelmatig moet men lang wachten doordat voorgaande afspraken zijn uitgelopen. Roshan Mahes bekijkt de mogelijkheden om deze roosters dynamisch aan te passen. Tot slot de columns: Jelke Bethlehem kijkt terug op de Amerikaanse presidentsverkiezingen en Gerrit Stemerding ventileert zijn ergernis over vragenlijsten. Henk Tijms zult u vanaf nu missen, hij heeft aangegeven na 34(!) columns te stoppen als vaste columnist. Wel is hij van plan incidenteel nog bijdragen te leveren. Heel veel dank Henk!

Laten wij als redactie ook naar de toekomst kijken: we hebben het voornemen op geregelde tijden een interview op te nemen met een bekende Nederlandse OR- of Statistiek-wetenschapper die in het buitenland werkzaam is. Uw suggesties zijn van harte welkom. Wij hopen u op 20 maart in Utrecht te ontmoeten en wensen u tot dan veel leesplezier!

De STATOR-redactie



INHOUD

2 De volgende 80 jaar

4 Een Ig Nobelprijs voor het 350.757 keer opgooien van muntjes: achtergrond en sfeerimpressie | Eric-Jan Wagenmakers and František Bartoš

8 Ik erger me! – column | Gerrit Stemerding

10 Citizen Science in Transient Astronomy: The Black Hole Finder App | Fiorenzo Stoppa

13 Young Statisticians

14 Dynamische afsprakenroosters in de zorg | Roshan Mahes

20 In memoriam: Prof. dr. Willem Schaafsma (1937 – 2024) | Casper Albers

21 In memoriam: dr. Alessandro Di Bucchianico (1963 – 2024) | Edwin van den Heuvel

23 Vergroot je kennis tijdens het ISI World Statistics Congress 2025 in Den Haag

24 Letter from the Upcoming President

25 Annual Meeting of the Netherlands Society for Statistics and Operations Research (VWSOR)

30 Statistische prijzen en de erfenis van Stella Cunliffe | Richard Starmans

35 Optimalisatie voor de sturing van de grootste pompen van Nederland | Kieran van Gaalen, Klaudia Horváth en Rens de Heer

38 Explainable Decision Making: How to explain optimization models in Operations Research | Jannis Kurtz

41 De Amerikaanse presidentsverkiezingen beter in beeld – column | Jelke Bethlehem



Een Ig Nobelprijs voor het 350.757 keer opgooien van muntjes: achtergrond en sfeerimpressie

Eric-Jan Wagenmakers and František Bartoš

De Ig Nobel prijs is een ludieke erkenning voor onderzoek dat eerst op de lachspieren werkt maar daarna tot nadenken stemt. In 2024 mochten wij de Ig Nobel prijs in ontvangst nemen in de categorie 'Probability', voor het 350.757 keer opgooien van muntjes. Hier beschrijven we de voorgeschiedenis van dit project, de bewust chaotische Ig Nobel prijsuitreiking en de nasleep.

Voorgeschiedenis

In het onderwijs over de kansrekening spelen dobbelstenen, urnen en muntjes van oudsher een belangrijke rol. Zo was het dat een van ons [EJW] tijdens de voorbereiding van de cursus *Bayesian inference for psychological science* bij toeval in aanraking kwam met het artikel "Dynamical bias in the coin toss" van Persi Diaconis, Susan Holmes and Richard Montgomery. Deze statistici ontwikkelden een natuurkundig model over hoe mensen muntjes opgooien, en deden vervolgens een opmerkelijke voorspelling:

"We analyseren het natuurlijke proces van het opgooien van een munt die in de hand wordt opgevangen. We laten zien dat krachtig opgegooide munten de neiging hebben om op dezelfde zijde te landen als die waarop ze begonnen. (...) Metingen (...) op basis van hogesnelheidsfotografie worden gerapporteerd. Voor natuurlijke worpen is de kans dat het muntstuk landt zoals het begon ongeveer .51." (Diaconis, Holmes en Montgomery, 2007, p. 211)¹

De natuurkundige verklaring van deze neiging is niet intuïtief en berust op de 'wiebel' (precessie) waarmee mensen een munt opgooien. Hoe wiebeliger de worp, hoe groter de proportie van de tijd die een munt in de lucht op de aanvangszijde doorbrengt. Kan dit nu echt waar zijn? Het lijkt eenvoudig om de voorspelling empirisch te toetsen. Maar het verschil tussen 0,50 en 0,51 is klein, en om het eventuele effect nauwkeurig te kunnen meten is een verrassend groot aantal worpen nodig:

"Onze schatting van de bias voor opgegooide munten is $p = 0,51$. Om p nabij $1/2$ te schatten met een standaardfout van $1/1000$ vereist $\frac{1}{2\sqrt{n}} = 1/1000$ of $n = 250,000$ worpen. Hoewel dit in de praktijk niet onhaalbaar is, zeker niet als er een nationaal muntwerp-evenement wordt georganiseerd, maakt dit het wel minder verrassend dat het huidige onderzoek niet empirisch is getest." (Diaconis, Holmes en Montgomery, 2007, p. 219)

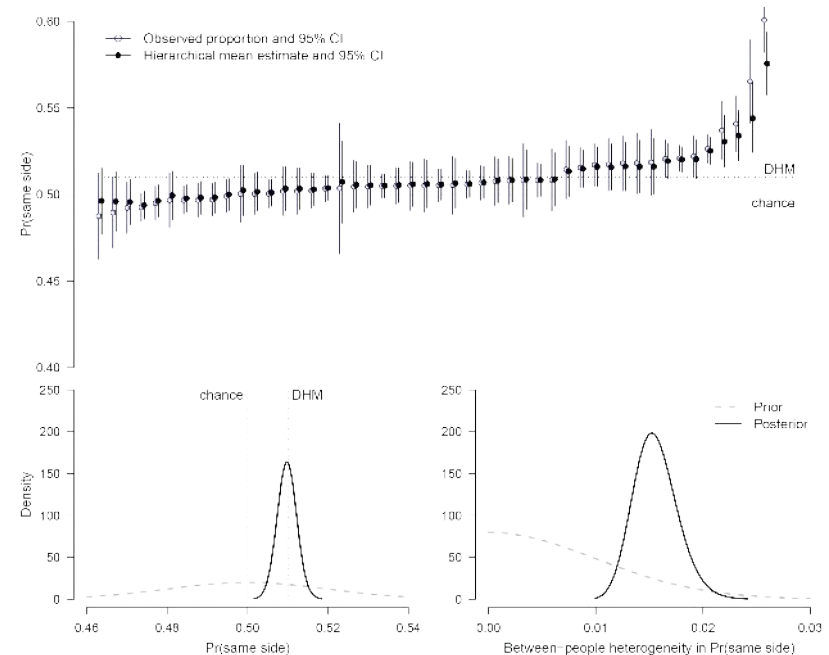
Gefascineerd door deze voorspelling heeft de eerste auteur jarenlang studenten (tevergeefs) proberen te enthousiasmeren voor dit onderzoek. In het kader van de Bayesiaanse cursus werd gedurende vier jaar 'slechts' 6927 keer geworpen, hetgeen resulteerde in 3334 landingen op de aanvangszijde (i.e., 51.9%; zie het hoofdstuk "Diaconis's Wobbly Coin" in Wagenmakers en Matzke, 2024). In 2022 ging de tweede auteur [FB] uiteindelijk de logistieke uitda-

ging ten volle aan. Na preregistratie van design en analyseplan kon de dataverzameling beginnen.

Muntwerp-marathon

In totaal hebben 49 werpers (studenten, docenten, en vrienden) samen 44 verschillende type munten opgegoooid, hetgeen ongeveer 650 uur (= 81 werkdagen) in beslag nam. Het leeuwendeel van de worpen werd verricht tijdens gemeenschappelijke 'muntwerp-marathons', waarvan de langste 12 uur duurde. Dit hobbyproject werd uitgevoerd in de avonden en de weekenden, en is niet financieel ondersteund door een subsidieverstrekker.

Het was natuurlijk van groot belang dat de munten netjes opgegoooid werden. De werpers kregen een korte instructie en moesten ter controle iedere worp opnemen met hun laptopcamera. Foutief uitgevoerde worpen (bv. muntjes die niet omdraaiden, of niet werden gevangen) zijn ter plekke vervangen.²



Figuur 1: Munten hebben de neiging om te landen op hun aanvangszijde, hetgeen ondersteuning biedt voor het model van Diaconis, Holmes en Montgomery (DHM). Bovenste paneel: werper-specifieke resultaten, geordend op effect-grootte. Paneel links onder: prior en posterior verdeling voor de overall kans dat muntjes landen op de aanvangszijde. Paneel rechts onder: prior en posterior verdeling voor de heterogeniteit tussen werpers in de kans op een 'zelfde zijde' uitkomst. Figuur uit Bartoš e.a. (2024).

Resultaten

De belangrijkste bevindingen zijn te zien in Figuur 1 (zie Bartoš e.a., 2024 voor een gedetailleerde omschrijving van de Bayesiaanse modellering en hypothese toetsing).³ Het bovenste paneel toont voor iedere werper apart de kans dat het muntje landt op de aanvangszijde. Een aantal werpers laat geen effect zien, een aantal werpers laat een effect zien nabij 0,51 (zoals voorspeld door Diaconis c.s.), en een aantal werpers laat een effect zien dat nog aanzienlijk groter is. Het paneel linksonder toont de prior en posterior verdeling voor de overall kans dat muntjes landen op de aanvangszijde. De posterior verdeling is geconcentreerd rond de waarde voorspeld door Diaconis c.s. en kent weinig massa toe aan waarden nabij 0,50. Het paneel rechtsonder toont de prior en posterior verdeling voor de heterogeniteit tussen werpers. Deze posterior verdeling kent weinig massa toe aan waarden nabij de nul, en suggereert derhalve dat de grootte van het effect afhankelijk van de werper; dit is consistent met de visuele impressie van de resultaten in het bovenste paneel. Binnen de context van het DHM model kunnen we speculeren dat wiebelige werpers een relatief groot effect laten zien.

Aanvankelijke aandacht

Direct nadat we de preprint online beschikbaar hadden gemaakt werden de resultaten vermeld op de radio, televisie, en in de krant, zowel binnen Neder-

land als daarbuiten. Het was toen al duidelijk dat de resultaten tot de verbeelding spraken – waarschijnlijk niet alleen vanwege het eindresultaat en de theoretische voorspelling, maar ook vanwege het feit dat er 350.757 maal was geworpen, iets wat op eerste gezicht een oefening lijkt in nuttelosheid.

In 2023 ontving het muntjes-onderzoek de *pineapple science award*, de Chinese tegenhanger van de Ig Nobel prijs. Deze award voor “nieuwsgierigheid” werd uitgereikt in het Olympisch stadion van Wenzhou, tijdens de World Young Scientists Summit. In het Westen is de pineapple science award relatief onbekend, maar zoals Figuur 2 suggereert was het evenement groots opgezet, professioneel georganiseerd, en strak geregisseerd. De moderator was een van China’s meest bekende presentatoren.⁴

Een ander hoogtepunt was het feit dat Stephen Colbert in 2023 de draak stak met het onderzoek (of beter gezegd, met František in *The Late Show*. In het fragment ‘Coin flip math’ concludeert Colbert:

“Je wilt *niet* verstrikt raken in een gesprek met deze jongen tijdens een feestje: “...en toen, op worp tweehonderdvierenvijftigduizend zeshonderdenachtentwintig, was het *kop!*! ...En daarna, op worp tweehonderdvierenvijftigduizend zeshonderdenegenentwintig, was het...opnieuw *kop!*!”

Al met al kwam het dus niet als een grote verrassing toen we te horen kregen dat het muntjes-onderzoek in 2024 een Ig Nobel prijs had gewonnen.

Ig Nobel prijs 2024

De Ig Nobel prijs wordt toegekend aan onderzoek dat je eerst laat lachen en vervolgens doet nadenken. De uitreikingsceremonie vond plaats in een collegezaal van MIT in Boston. Zoals de foto bovenaan dit artikel doet vermoeden verliep de ceremonie bijzonder chaotisch, dat wil zeggen, precies zoals gewenst. Het thema van 2024 was *Murphy’s law* (“alles wat fout kan gaan, gaat ook fout”). Er werd gegooid met papieren vliegtuigjes, er werd eindeloos handen geschud met echte Nobelprijswinnaars, er werd gezongen, en er werden flauwe grappen gemaakt.

Het leukste aan de ceremonie (die in zijn geheel kan worden teruggezien op YouTube⁵) waren de acceptatiespeeches. De winnaars hadden precies één minuut om hun onderzoek uit te leggen. In vlot tempo ging het onder andere over kamikazeduiven die hadden geleerd om raketten naar hun doel te leiden (dit onderzoek kreeg overigens de Ig Nobel prijs voor de vrede), het zwemgedrag van een dode vis (het blijkt dat vissen vanwege hun anatomie weinig moeite hoeven te doen om te zwemmen), en de bewegingen van wormen die eerst dronken waren gevoerd.⁶ Dat laatste onderzoek was uitgevoerd door een team van natuurkundigen die toevallig net als wij werkzaam zijn aan de Universiteit van Amsterdam. Geheel in de geest van het evenement spendeerden ze de meeste tijd met het nabootsen van een race tussen een sobere en een dronken worm (de sobere worm boekte een overtuigende overwinning).

Zelf waren we behoorlijk zenuwachtig, want als je maar één minuut hebt wil je niet over je woorden struikelen. Gelukkig gebeurde dat niet, en hoefde de gevreesde *Miss Sweetie Poo* ook niet in actie te komen. *Miss Sweetie Poo* is een achtjarig meisje dat bij overschrijding van de spreektijd naar het podium loopt en vervolgens op luide stem constant dezelfde zin herhaalt: “Please stop. I’m bored! Please stop.”

Nasleep

De uitreiking van de Ig Nobel prijs was voor radio, televisie, en kranten een aanleiding om zich nog een tweede keer op het muntjes-onderzoek te storten. Een van ons [EJW] deed een optreden bij het wetenschappelijk jaaroverzicht van het televisieprogramma *Focus*. Daar werd door een ter plekke geronselde assistent (radiohost Ruben Rosen Jacobson) 1000 maal geworpen, waarvan de munt 578 keer landde op de aanvangszijde – een relatief groot effect dat apart beschouwd al statistisch overtuigend is.

Over de uiteindelijke bestemming van de preprint kan op dit moment nog geen definitieve uitspraak worden gedaan. Het modelleren en reviseren heeft uiteindelijk meer tijd in beslag genomen dan de dataverzameling. Een reviewer suggereerde dat het effect misschien verandert over de tijd, en dat lijkt ook zo te zijn: het effect is het grootst bij aanvang, waarna het langzaam verwatert. Dit zou kunnen komen doordat wiebelige werpers gaandeweg geoefend raken en minder wiebelig gaan gooien.

Lange tijd hebben we gedacht dat dit hobbyproject geen enkel praktisch nut zou hebben. Daar zijn we inmiddels van teruggekomen. De data zijn van nut voor het onderwijs en het populariseren van de statistiek.

Voetnoten

- 1) Alle oorspronkelijk Engelse teksten zijn vrij vertaald naar het Nederlands, soms met assistentie van DeepL.
- 2) Data en opnames zijn publiekelijk beschikbaar op <https://osf.io/pxu6r/>.
- 3) Een flyer die de resultaten samenvat is beschikbaar op <https://osf.io/d6mv9/>.
- 4) Een uitgebreider verslag is te vinden op <https://www.bayesianspectacles.org/coin-tossing-paper-wins-chinese-ignobel-pineapple-award/>.
- 5) Zie ook <https://improbable.com/ig/archive/2024-ceremony/>; het specifieke gedeelte over de muntjes is te vinden op <https://skepsis.nl/ig-nobelprijs-voor-eric-jan-wagenmakers/>.
- 6) Zie ook het verslag in *Science* op <https://www.science.org/content/article/scienceadviser-what-do-you-do-wit-h-a-drunken-worm>.

Literatuur

- F. Bartoš e.a. „Fair Coins Tend to Land on the Same Side They Started: Evidence From 350,757 Flips”. In: *ArXiv Preprint* (2024). URL: <https://doi.org/10.48550/arXiv.2310.04153>.
- P. Diaconis, S. Holmes en R. Montgomery. „Dynamical Bias in the Coin Toss”. In: *SIAM Review* 49 (2007), p. 211–235.
- E. Wagenmakers en D. Matzke. *Bayesian Inference from the Ground Up: The Theory of Common Sense*. 2024. URL: <https://www.bayesianspectacles.org/free-course-book/>.

Eric-Jan Wagenmakers is hoogleraar Bayesiaanse Methodologie bij de afdeling Psychologische Methodenleer aan de Universiteit van Amsterdam.
Email: EJ.Wagenmakers@gmail.com

František Bartoš is promovendus bij de afdeling Psychologische Methodenleer aan de Universiteit van Amsterdam.
Email: f.bartos96@gmail.com



Figuur 2: Alexandra Sarafoglou presenteert de resultaten voorafgaand aan de uitreiking van de *pineapple science award* op de 2023 World Young Scientists Summit in Wenzhou, China.



Ik erger me!

Zo'n zestig jaar geleden heb ik een tijdlang bij een markt-onderzoeksbureau gewerkt. Daarom weet ik uit ervaring hoe moeilijk het is een goede groep respondenten te krijgen. Als ik weer eens een mail krijg met het verzoek mee te doen aan een onderzoek kan ik het dan ook niet over mijn hart verkrijgen dit te weigeren. Maar het zal wel aan mijn inmiddels fiks oplopende leeftijd

liggen dat ik mij steeds vaker erger aan de vragenlijst die ik dan krijg. Is er bij die bureaus dan niemand die hier toezicht op houdt? Een soort eindredacteur is hard nodig, vooral bij vragenlijsten die lijken te zijn samengesteld door het gedachteloos achter elkaar plakken van vragen voor verschillende opdrachtgevers.

Survey Form

Name: _____	Age: _____		
Date: _____	Address: _____		
Place: _____			
	Yes	Maybe	No
Would you losem ipsum dolor sit amet, consectetur adipiscing elit?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Would you losem ipsum dolor sit amet, consectetur adipiscing elit?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Would you losem ipsum dolor sit amet, consectetur adipiscing elit?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Would you losem ipsum dolor sit amet, consectetur adipiscing elit?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Would you losem ipsum dolor sit amet, consectetur adipiscing elit?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Would you losem ipsum dolor sit amet, consectetur adipiscing elit?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Would you losem ipsum dolor sit amet, consectetur adipiscing elit?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Thank you for your valuable feedback! Lorem ipsum dolor sit amet, consectetur adipiscing elit.			
<input type="button" value="next"/>			

Zo maar wat ergernissen, in een willekeurige volgorde:

- Veelvuldige taalfouten, vaak tegen de d/dt regels, of verkeerd gespelde woorden, of onjuist gebruikte vaste uitdrukkingen etc.
- Binnen één-en-dezelfde lijst afwisselend worden aangesproken met u of jij.
- Het nagenoeg stelselmatig ontbreken van antwoordcategorieën als 'weet niet', 'niet van toepassing' etc. Kennelijk heeft men liever een volstrekt irrelevant antwoord dan géén antwoord.
- Lange lijsten van bijvoorbeeld steden die men bezocht kan hebben. Pas na eindeloos scrollen kan men aan het slot aangegeven geen tripjes te hebben ondernomen.
- Inconsequente manieren van antwoord geven op vergelijkbare vragen: soms een hokje aanvinken, soms een cijfer van 1-10 geven, 5- en 7-puntsschalen, en dat alles binnen een-en-dezelfde vragenlijst etc.

De laatste jaren is er vermoedelijk een afhankelijkheid ontstaan van een softwareprogramma om vragenlijsten te componeren. Dat programma, dat ik overigens niet ken, zal best prima werken, maar het ontbeert noodzakelijke opties voor het zogenaamde routeren: het automatisch vragen overslaan die op basis van eerder gegeven antwoorden niet meer relevant zijn. Maar misschien heb ik het fout en zijn die opties er wel, en is het gewoon luiheid of onwetendheid van de samensteller.

Toch is er één eigenschap van dat programma waarbij een systematicus als ik zich steeds weer afvraagt

waarom dat zo knullig moet: het al dan niet automatisch doorspringen naar een volgende vraag.

Vaak springt men na het geven van een antwoord automatisch door naar de volgende vraag. Maar het komt ook voor dat men na het beantwoorden van een vraag zélf moet aangeven door te willen gaan. Beide systemen zijn goed, als tenminste voor alle vragen voor dezelfde manier wordt gekozen. Maar het gaat mis als men beide gebruikt, dat is behoorlijk irritant. Soms bestaat een vraag uit een opsomming van gelijkvormige deelvragen, zoals het geven van een mening over een reeks eigenschappen van een persoon, biermerk, krant, of wat dan ook. Daarbij moet voor iedere eigenschap een punt op een schaal worden ingevuld. Na het aanklikken van het antwoord voor één eigenschap springt het programma dan automatisch door naar de volgende eigenschap. Maar niet na de laatste eigenschap op de lijst! Om dan naar de volgende vraag te gaan moet men weer zelf op een 'ga door' knop klikken. Waarom kan dat automatisme niet ook voor die laatste deelvraag gelden? Het antwoord zal wel zijn dat het programma dat nu eenmaal zo regelt. Maar dat is een té gemakkelijk antwoord, je zou ook kunnen zoeken naar een ander programma of er bij de producent op aandringen dit te veranderen.

Als je kostbare respondenten tevreden wilt houden zul je toch zo nu en dan zélf met de ogen van een kritische respondent naar je eigen vragenlijsten moeten kijken. En zoals ik al zei is dat typisch een klus voor een goede eindredacteur (PS: dit is geen verkapte sollicitatie).

Gerrit Stemerink is eindredacteur van STATOR.
E-mail: gstemerink@hotmail.com

Citizen Science in Transient Astronomy: The Black Hole Finder App

Fiorenzo Stoppa

The Black Hole Finder App is a Citizen Science Project that invites the public to help identify astronomical transients, such as objects that suddenly appear or change brightness in the night sky. Users are given a set of three images: a recent image captured by the BlackGEM telescope array in Chile, a reference image of the same sky region, and a difference image that makes it easier to spot changes between the first two. Their task is to compare these images and decide whether the observed change is a real celestial event or simply an artifact caused by the imaging process or human-

made interference, like satellites or airplanes passing through the field of view. This article explains how the app gathers and uses public input to support scientific research. It describes how user classifications are used to train machine learning models, which help improve the accuracy of detecting transient celestial events. By combining public participation with data analysis, the Black Hole Finder App helps us better understand these brief cosmic events and promotes a collaborative approach to astronomical discovery.

Introduction

Astronomy always captivates the imagination of both scientists and public. With advancements in technology, the opportunities for public involvement in scientific research have significantly expanded. The Black Hole Finder App is an example of this trend, enabling citizen scientists to contribute to the identification of astronomical transients.

The app was inspired by groundbreaking events like the 2017 kilonova, where the merger of two neutron stars was observed through both gravitational waves and electromagnetic signals. This rare event highlighted the need for rapid and precise identification of short-lived cosmic phenomena. The BlackGEM telescope array, located in Chile, was developed to respond to such signals quickly, capturing new images of the sky whenever a gravitational wave is detected. However, distinguishing between genuine astronomical events and false signals, such as spurious signals from various sources, presents a significant challenge.

This is where the Black Hole Finder App comes into play. The app invites the users to help analyse the vast amounts of data generated by BlackGEM, such as to separate real cosmic events from artifacts. While the app's name might suggest a direct search for black holes—elusive objects that are invisible in optical light—it actually focuses on detecting cosmic events that may indicate their presence, such as merging neutron stars or exploding stars.

The Black Hole Finder App: How It Works

Every night, the BlackGEM telescope array in Chile captures vast amounts of data, continuously surveying large areas of the sky for any signs of change. This influx of data is immense, with each observation contributing to the potential discovery of new celestial events.

Once the images are captured, they undergo a rapid processing phase. Within just 15 minutes, new images are compared against reference images—composites of earlier observations of the same sky region. Any significant changes detected in this comparison are flagged as candidate transients. These candidates are then sent to a central database for further analysis and public engagement.

To make this data accessible to the public, the Black Hole Finder App was developed in collaboration with the Dutch company DDQ Pocket Science. DDQ Pocket Science manages the infrastructure

that receives and distributes these images, ensuring that they are efficiently delivered to users through the app.

When users open the app, they are provided with three key images:

- **New image:** The freshly captured image from the BlackGEM telescope, highlighting the current state of a particular sky region.
- **Reference image:** A composite image created from earlier observations of the same region, serving as a stable baseline for comparison.
- **Difference image:** A processed output showing the differences between the new and reference images, clearly highlighting areas where changes have occurred, such as the appearance of a transient object.

After reviewing these images, users are prompted to classify the central object by selecting one of three options: Real, Bogus, or Unknown. Real refers to a genuine celestial event, such as a supernova or a variable star. A supernova, as shown in the first three images of the first row in Figure 1, appears as a circular source at the center of the New image, often located close to a nearby galaxy. In the Reference image, only the galaxy is visible, while in the Difference image, the transient stands out clearly as a new source.

A variable star, illustrated by the last three images of the first row in Figure 1, appears as a bright circular object in both the New and Reference images. In the Difference image, the variable star shows as a positive source if the star became brighter, as in this example, or as a negative source if the star had dimmed, reflecting its changing brightness over time.

Bogus refers to spurious objects caused by artifacts such as imaging errors, diffraction spikes, cosmic rays, or human-made interference. For example, a diffraction spike is shown in the first three images of the second row in Figure 1. Here, the “change” is due to optical effects near a bright star rather than a real transient. Similarly, the last three images of the second row depict a cosmic ray, which appears as a sharp, pixelated signal in the New and Difference images but has no corresponding feature in the Reference image.

If users cannot confidently determine whether the observed change is a genuine celestial event or an artifact, they can select “Unknown.” This option ensures that ambiguous cases are flagged for further review by experts, allowing the system to capture potentially complex scenarios that require additional analysis.

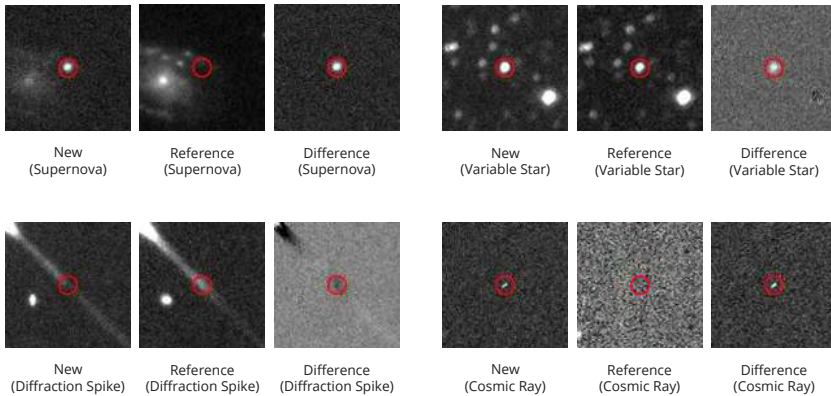


Figure 1: Cutouts used in the Black Hole Finder App: The first row shows real transients (supernova and variable star), while the second row shows bogus sources (diffraction spike and cosmic ray). Each set of three images includes the New, Reference, and Difference images.

Additionally, users who download the app are provided with a tutorial that guides them through distinguishing between Real and Bogus classifications.

Engaging Citizen Scientists

Citizen science relies on public participation to contribute to scientific research, and the Black Hole Finder App exemplifies this approach. By enabling users to assist in the identification of transient celestial events, the app gathers valuable data while providing an accessible way for the public to participate in astronomy. The task's simplicity, paired with the possibility of contributing to scientific discoveries, offers an engaging experience for a wide range of participants. As of January 2025, the Black Hole Finder App has attracted 9,250 registered users who have collectively classified 2,218,476 transients across 412,236 unique events.

To further motivate users, the app features a scoreboard that highlights the top 10 classifiers based on the number of transients they have reviewed. While the app currently does not track accuracy metrics—since the transients identified are unlabelled and awaiting professional confirmation—the scoreboard serves as a fun and competitive way to recognize active contributors. This element of gamification helps maintain user engagement and encourages continued participation.

In addition to identifying transients, selected citizen scientists are now able to request follow-up observations of particularly interesting events. These follow-up observations are carried out using the telescopes from the Las Cumbres Observatory (LCO)

network, a global array of twenty-five telescopes at seven sites around the world. The LCO network allows astronomers to quickly and efficiently capture additional data on transient events, enabling more detailed analysis. This functionality empowers app participants to play an even more active role in scientific discovery by prioritizing observations for events they find intriguing.

A notable example of a successful follow-up is the recently discovered Iax Supernova, AT 2024vj¹. Within minutes to hours of its discovery by BlackGEM, 12 users requested follow-up observations of this event from the app. Observing time on the LCO network was quickly allocated, resulting in multiple follow-up images being captured in a wide range of filters. One of these images is featured as the main image in this article.

From Citizen Classifications to Machine Learning

A key goal of the Black Hole Finder App is to generate a vast dataset of human classifications, forming the foundation for training machine learning models to automatically identify transient celestial events.

Unlike typical machine learning projects, where ground-truth labels are readily available, transient events in this context lack definitive labels. As a result, the classifications provided by app users become the core of our labeled training set. To address the absence of true labels, we aggregate user classifications to determine the most likely label for each transient. This typically involves analyzing the consensus among multiple users: if a majority classify

a transient as “Real”, it’s more likely to be labeled as such, and similarly for “Bogus.” While consensus-based labeling serves as the foundational method for creating a reliable training set, it has limitations, as user accuracy can vary. To address these challenges and improve label quality, we are exploring statistical methods like Latent Class Analysis (LCA). LCA enhances the labeling process by analyzing individual user classification patterns, accounting for variations in reliability and expertise. This approach allows us to probabilistically assign labels to transients, even in cases of disagreement, by identifying which users are more likely to provide accurate classifications over time.

The generated labels serve as the foundation for training a Convolutional Neural Network (CNN). CNNs are powerful machine learning models designed to process image data by automatically extracting patterns, such as shapes and textures. By comparing the New, Reference, and Difference images, the CNN learns to recognize patterns associated with genuine transients and distinguish them from spurious artifacts, such as imaging errors or cosmic rays.

However, the process does not end with the initial training. To ensure the network remains accurate and adapts to new data, we will rely on continuously updated labels from users. This allows us to periodically retrain the model, improving its performance and ensuring it reflects current observations. Additionally, the trained CNN can serve as a tool to cross-check human classifications: when the model’s predictions deviate from human consensus, these discrepancies highlight cases that may need further human review and relabeling. This iterative process creates a feedback loop, where human input and machine learning work together to improve both the model and the dataset.

As the app continues to be developed and more users contribute, its outputs may provide new opportunities for scientific discovery and deepen our understanding of transient celestial phenomena.

Notes

- 1) <https://www.wis-tns.org/object/2024vjm>

Florenzo Stoppa specialises in the intersection of advanced statistical methods, machine learning techniques, and astronomy. His research is focused on developing innovative solutions and tools that not only streamline data analysis within the astronomical community but also contribute to the advancement of scientific understanding in the field. <https://sites.google.com/view/fstoppa/home>.



Young Statisticians

The Young Statisticians have a new board in 2025. We are MSc Students in Statistics and Data Science at the University of Leiden. We are planning our first event for this year on 21st February 2025: A Statistics Cafe on Sports Data Science.



Next event

If you would like to attend or learn more details, sign up for our newsletter, follow us on LinkedIn and follow our Instagram: @youngstatisticians, or keep an eye on our website: [vvsor.nl/young-statisticians](https://www.vvsor.nl/young-statisticians).



Photo by Amsterdam UMC

Dynamische afsprakenroosters in de zorg

Roshan Mahes

Artsen en medische apparatuur zijn duur en moeten daarom zoveel mogelijk worden benut. Om dit te bereiken worden er veelal afsprakenroosters gebruikt die ervoor zorgen dat patiënten op tijd aanwezig zijn. De meeste afsprakenroosters zijn vaak statisch en houden geen rekening met de onvoorspelbaarheid gedurende de dag, behalve door die van tevoren mee te wegen. Met andere woorden: er kunnen situaties optreden met lange wachttijden of veel uitloop.

In dit artikel bekijken we daarom een alternatieve dynamische aanpak voor afspraakplanning, waarbij real-

time updates naar patiënten kunnen worden gestuurd. Door gebruik te maken van dynamisch programmeren kunnen we de effectiviteit van zo'n aanpak doorrekenen.

Drie aanpakken worden uitgelegd: bijwerken bij aankomst van de patiënt, bijwerken op vaste tijden gedurende de dag en bijwerken op basis van het aantal wachtenden in de wachtkamer. Resultaten tonen aan dat het geven van updates door de dag de potentie heeft om niet alleen wachttijden te reduceren, maar ook de efficiëntie te verhogen.

Afspraakplanningen

In de gezondheidszorg worden patiënten vaak geconfronteerd met lange wachttijden, of het nu gaat om een bezoek aan de huisarts, een afspraak met een specialist, of het gebruik van dure medische apparatuur zoals MRI-scanners. Denk bijvoorbeeld aan een patiënt die om 10:00 uur een afspraak heeft voor een MRI-scan. Onvoorzien uitloop bij eerdere patiënten kunnen ertoe leiden dat de wachttijd snel oploopt, waardoor deze patiënt pas een uur later aan de beurt is. Dit kan leiden tot frustratie en ook de benutting van schaarse middelen negatief beïnvloeden.

Traditionele afsprakenroosters, geïntroduceerd in Bailey, 1952, hebben als doel de vraag van patiënten (behandelingen) af te stemmen op de beschikbaarheid van middelen (zoals artsen, of medische apparatuur). Doorgaans zijn deze roosters gebaseerd op de gemiddelde tijd.

Onvoorspelbare Behandelduren

In de praktijk zijn behandelduren echter niet deterministisch. Soms duurt een afspraak nou eenmaal langer dan anders. Door deze onvoorspelbaarheid van de behandeltijden kan *ieder* schema leiden tot een kettingreactie van vertragingen. Dit resulteert in langere wachttijden en een afname van de kwaliteit van zorg. Het wordt nog erger als er strak geroosterd is; een afsprakenrooster gebaseerd op gemiddeldes

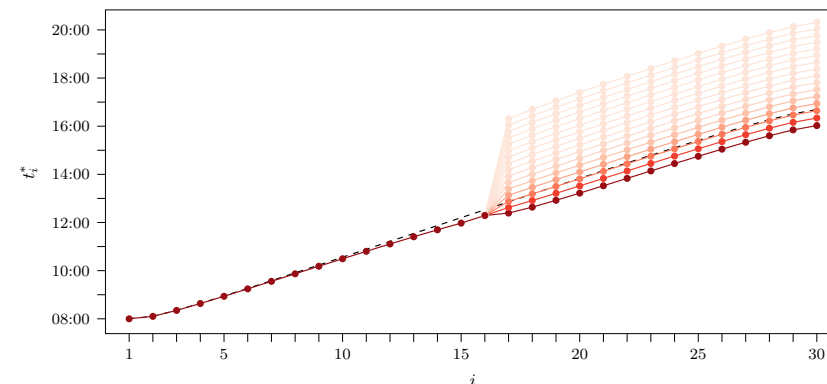
is zo'n strak schema.

Stel je voor dat voor iedere klant i de gemiddelde behandelduur 15 minuten is, ofwel $EB_i = 15$ (min.). Als we dit modelleren met een exponentiële verdeling, dan is $B_i \sim \text{Exp}(\lambda = \frac{1}{15})$. De cumulatieve verdelingsfunctie is gelijk aan $1 - e^{-\lambda x}$, ofwel bij roosteren op het gemiddelde is 63% van de afspraken korter dan 15 minuten, terwijl 37% langer duurt. Bovendien, wanneer een afspraak is uitgelopen, heeft dit directe impact op de eerstvolgende klanten. Je haalt die tijd niet meer terug omdat er op de gemiddelde tijd is geroosterd. Het kan dus verstandiger zijn om iets ruimer te roosteren. Te ruim roosteren kan echter leiden tot onnodige wachttijd aan de kant van de dokter.

Om afsprakenroosters te optimaliseren wordt er daarom naar een kostenfunctie gekeken die zowel de wachttijd voor patiënten als de inactiviteit van apparatuur of dokter in overweging neemt:

$$C = \omega \times \text{inactiviteit} + (1 - \omega) \times \text{wachttijd}.$$

In deze formule geeft ω aan hoe belangrijk het is om inactiviteit te beperken ten opzichte van wachttijden. Als je deze functie minimaliseert over alle mogelijke afsprakenroosters krijg je het optimale rooster. Zo'n schema wordt weergegeven in Figuur 1 door de zwarte gestreepte lijn. Hieronder leggen we uit wat de andere strepen betekenen, en dat het alles te maken heeft met het versturen van updates tijdens de sessie.



Figuur 1: Een voorbeeld van een dynamisch rooster. De stippen geven de aankomsttijden per patiënt aan. Herroostering vindt plaats bij de aankomst van patiënt 16, waarna de kleurintensiteit van de vervolgroosters bepaald wordt door de waarschijnlijkheid.

Het versturen van updates

Een mogelijke oplossing om inactiviteit en wachttijd te verminderen is een meer dynamische aanpak van afspraakplanning te kiezen. Met dynamisch bedoelen we een systeem dat zich aanpast aan veranderingen gedurende de dag. Hierin stellen wij een strategie voor waarbij patiënten gedurende de sessie één of meerdere updates ontvangen over hun afspraak. Door gebruik te maken van moderne communicatiemiddelen, zoals sms en apps, kunnen zorgverleners de patiënten makkelijk bereiken en daarmee in real time informeren over vertragingen of wijzigingen.

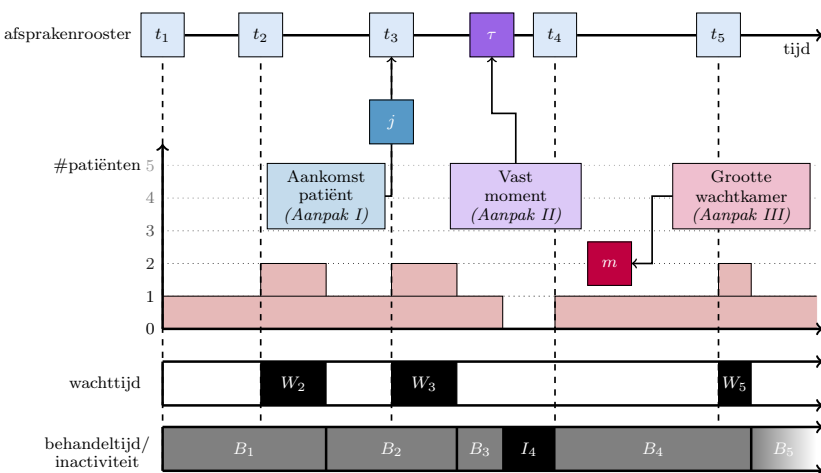
Als een MRI-scanner bijvoorbeeld achterloopt op schema, kunnen patiënten tijdig een melding krijgen om hun aankomsttijd uit te stellen, zodat ze minder lang in de kliniek hoeven te wachten. Dit maakt de ervaring voor patiënten een stuk prettiger en blijkt tevens ervoor te zorgen dat de medische apparatuur efficiënter wordt gebruikt. De traditionele manier van het optimaliseren van een afsprakenrooster moet wel herzien worden. Aangezien je tijdens de sessie ingrijpt, wil je dat deze informatie op een of andere manier zijn weg terugvindt in de formulering. Technieken van dynamisch programmeren liggen voor de hand.

Dynamisch programmeren

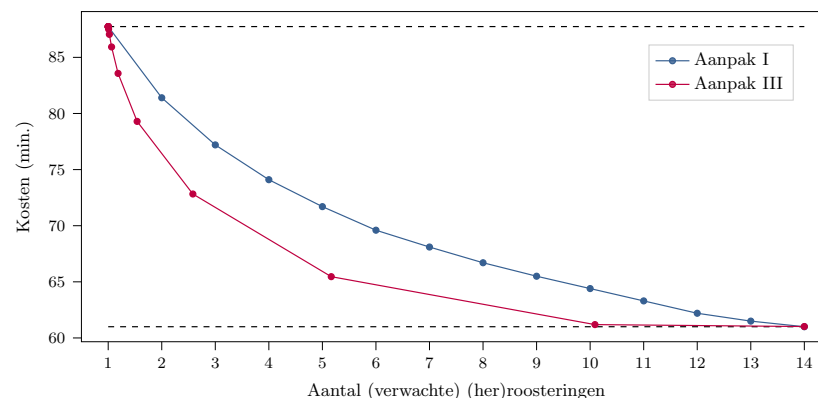
Dynamisch programmeren is een wiskundige techniek die helpt bij het oplossen van een groter probleem door het op te splitsen in kleinere subproblemen. Stel dat we halverwege de dag een update willen doorvoeren. Op dat moment kunnen we met behulp van dynamisch programmeren de huidige staat van het systeem (bijvoorbeeld het aantal wachtende patiënten) evalueren om de resterende afspraken te optimaliseren.

Voor afsprakenroosters werkt dynamisch programmeren als volgt. Gezien een sessie eindig is, is het handig om juist eerst de laatste beslissingen in het schema (de afspraken aan het einde van de dag) te optimaliseren, en vervolgens stap voor stap terug te werken naar het begin van de dag.

Door dynamisch programmeren toe te passen en te herroosteren bij *iedere* aankomst kan wel een kostenbesparing van 30% gerealiseerd worden, zie Mahes, 2024. Toch lijkt het niet altijd praktisch om bij elke aankomst het schema volledig te herzien. Niet alleen vanwege het eventuele werk wat bij het implementeren en versturen van zulke updates komt kijken, maar ook omdat de volgende patiënt dus pas erg laat weet hoe laat deze definitief moet komen. Dit kan onhandig en frustrerend zijn wanneer de pa-



Figuur 2: Drie herroosterparadigma's: op aankomsten (I), op vaste momenten (II) en op het aantal wachtende patiënten (III).



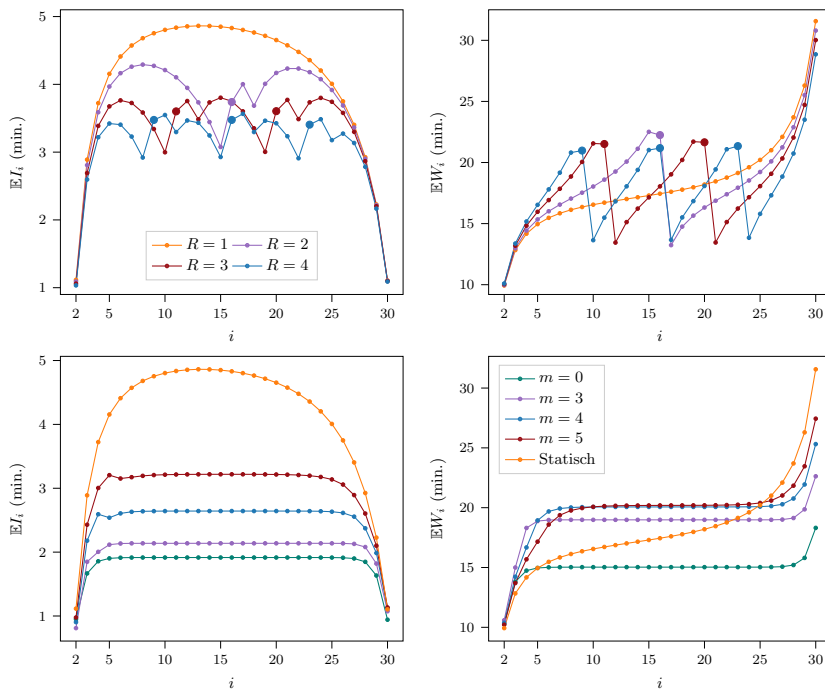
Figuur 3: Paradigma's I en III: kosten als functie van het aantal keer dat we ingrijpen. De bovenste gestreepte lijn weergeeft de kosten wanneer er niet ingegrepen wordt, terwijl de onderste lijn het extreme scenario waarbij bij elke aankomst herroosterd wordt illustreert.

tiënt bijvoorbeeld al onderweg is. Of de patiënt zou misschien veel eerder moeten komen dan verwacht, wat niet haalbaar is voor de patiënt. We kijken nu of we een balans kunnen vinden tussen het aantal keer updates en de kostenbesparing.

Drie paradigma's voor updates

Er zijn drie hoofdanpakken (of paradigma's) voor het toepassen van dynamisch programmeren bij het bijwerken van afsprakenroosters, zie ook Figuur 2:

- I. Bijwerken bij **specifieke aankomsten**. In deze benadering wordt het schema bijgewerkt op het moment dat een nieuwe patiënt j arriveert, bijvoorbeeld patiënt 16 als in Figuur 1. Dynamisch programmeren wordt hier gebruikt om op dat moment een beslissing te nemen op basis van de huidige situatie. Als er bijvoorbeeld veel patiënten in de wachtkamer zijn, kan de kostenfunctie C opnieuw worden geëvalueerd om te bepalen of toekomstige afspraken moeten worden aangepast om wachttijden te verminderen. Elke keer dat een patiënt arriveert, zou het schema aangepast kunnen worden om een nieuw, optimaal schema te berekenen dat past bij de huidige omstandigheden; we bepalen van tevoren bij welke aankomsten we het afsprakenrooster gaan veranderen.
- II. Bijwerken op **vaste momenten**. In deze aanpak worden updates op vooraf bepaalde tijdstippen gedurende de dag (bijvoorbeeld 10:00 en 13:00) uitgevoerd, ongeacht de aankomst van patiënten. Met dynamisch programmeren kunnen we op elk vast tijdstip $\tau_1, \tau_2, \dots, \tau_R$ het optimale schema herberekenen op basis van de huidige toestand van het systeem. Dit helpt ons om vroegtijdig in te grijpen en aanpassingen door te voeren die toekomstige vertragingen kunnen voorkomen. Hier wordt de hele dag als een reeks intervallen beschouwd, waarbij elk interval opnieuw wordt geoptimaliseerd op basis van de situatie op dat moment.
- III. Bijwerken op basis van het aantal patiënten in de **wachtkamer**. Hier wordt het schema aangepast wanneer het aantal wachtende patiënten een bepaald maximum m bereikt. Dynamisch programmeren helpt hier door continu te monitoren hoeveel patiënten er wachten. Merk op dat het aantal wachtende patiënten enkel kan toenemen bij een nieuwe aankomst. Op het moment dat er minimaal m wachtende patiënten zijn, wordt het schema opnieuw berekend om de resterende afspraken zó aan te passen dat hun wachttijden worden verkort. Deze methode is reactiever en grijpt dus alleen in wanneer de situatie in de wachtkamer 'kritiek' wordt.



Figuur 4: Paradigma's I (boven) en III (onder): de inactieve E_i (links) en wachttijden EW_i (rechts) per geroosterd patiënt i . Het aantal herroostermomenten in de bovenste panelen varieert van 0 ($R = 1$) tot en met 3 ($R = 4$). De dik gedrukte stippen geven de optimale herroostermomenten aan. In de onderste panelen geeft m het niveau aan wanneer er herroosterd gaat worden.

Wat is de beste aanpak?

Hoewel dynamisch programmeren het mogelijk maakt om het afsprakenrooster op elk moment aan te passen, is het niet altijd nodig om dit te vaak te doen. Het aantal updates moet worden afgewogen tegen het afnemende rendement; de eerste paar updates hebben vaak het grootste effect op het verminderen van wachttijden en inactiviteit, terwijl latere updates steeds minder verbetering brengen, zie ook Figuur 3. Daarom is het kiezen van de juiste momenten om te updaten essentieel. Uit onze analyse blijkt dat het strategisch kiezen van één of twee updatemomenten gedurende de dag, zoals halverwege, cf. patiënt 16 in Figuur 1, of op één derde en

tweederde van de dag, vaak al voldoende is om de efficiëntie aanzienlijk te verbeteren.

Uit de numerieke resultaten blijkt dat de aanpak op basis van de wachtkamer (III) tot de grootste kostenbesparing kan leiden, zie Figuur 3. In essentie grijpt deze aanpak alleen in wanneer het echt nodig is, namelijk bij een overvolle wachtkamer. Hierdoor wordt een optimale balans gevonden tussen het verminderen van wachttijden en het efficiënt gebruiken van schaarse middelen.

Echter hebben andere aanpakken, zoals herroosteren bij aankomsten (I) of op vaste momenten (II), ook hun voordelen. De roosteraar hoeft namelijk niet bij iedere aankomst de wachtkamer in de gaten te houden; de specifieke updatemomenten worden

van tevoren gekozen. Maar ook dan is het de kunst om de juiste momenten voor deze updates te vinden voor de hoogst mogelijke kostenbesparing.

Een ander voordeel van de aanpak op basis van de wachtkamer (III) is dat deze resulteert in stabielere en daarmee eerlijkere wachttijden. Zo laat Figuur 4 een vergelijking zien van de te verwachten inactieve tijden E_i en wachttijden EW_i per klant i voor paradigma's I en III. Voor paradigma I zijn tevens de (optimale) herroostermomenten aangegeven. Uit het paneel rechtsboven blijkt interessant genoeg dat juist de patiënten na een herroostermoment profiteren van de update, terwijl degenen net ervoor er op achteruitgaan. Deze effecten zijn onder paradigma III niet aanwezig.

Extra uitdagingen

Naast het bepalen van de paradigma's is het ook belangrijk factoren mee te nemen die de prestatie van de aanpak zouden kunnen beïnvloeden. We bespreken vier verschillende richtingen.

Allereerst kan men denken aan het loslaten van de exponentiële verdeling. Een belangrijke motivatie voor het overwegen van andere verdelingen is dat de exponentiële verdeling vrij variabel is, terwijl behandelduren in de zorg doorgaans dichter bij het gemiddelde liggen dan de exponentiële verdeling het modelleert. Ten tweede is een overweging het typische verschijnsel van zogenaamde no-shows. No-shows zijn patiënten die niet komen opdagen voor hun afspraak, in Nederland typisch niet meer dan 5%—in Amerika zijn deze percentages doorgaans veel hoger—wat leidt tot inactiviteit van de medische hulpmiddelen. Deze kunnen eenvoudig meegenomen worden door de verdeling van de behandelduren aan te passen: bij een no-show percentage van 5%, zet je met 5% kans de behandelduur gelijk aan nul, terwijl je met 95% kans de originele verdeling van de behandeltijd neemt.

Daarnaast kun je denken aan overtijd O_T als het gewenst is dat de sessie niet langer dan een bepaald tijdstip T (sluittingstijd) duurt, bijvoorbeeld 18:00 in Figuur 1. De verwachte overtijd, i.e., het aantal minuten dat de laatste klant vertrekt na sluitingstijd, kan toegevoegd worden aan de doelfunctie. Tot slot is het nuttig om een bevroingsperiode toe te voegen, d.w.z. om het schema niet aan te passen voor de patiënten die net na een herroostermoment zullen

komen. Je kunt je voorstellen dat het best wel onwenselijk kan zijn dat in Figuur 1 patiënt 17 een update krijgt; waarschijnlijk is deze namelijk al onderweg.

Zowel overtijd als de toevoeging van een bevroingsperiode maken het dynamisch programma aanzienlijk gecompliceerder, gezien het gehele probleem niet meer opgedeeld kan worden in losse stukken. Bijvoorbeeld bij één keer herroosteren, is de informatie in het eerste deel bepalend hoe de bevroingsperiode in de subproblemen geëvalueerd moet worden. Oftewel: je kunt niet meer gewoonweg beginnen met het laatste deel (d.w.z. vanaf het herroostermoment) evalueren. Ook wanneer je naar overtijd kijkt, moet je weten hoeveel tijd van de sessie al 'verbruikt' is bij het herroostermoment, terwijl bij bevroingsperiodes de patiënten die al geroosterd zijn maar nog aan moeten komen meegenomen moeten worden.

Kortom, door dynamisch programmeren toe te passen op afspraakplanning kunnen zorgverleners effectief inspelen op de veranderlijke omstandigheden van de dag. Dit leidt tot een betere patiëntervaring en een efficiënter gebruik van waardevolle middelen.

Dankwoord

Dit artikel is gebaseerd op mijn proefschrift (hoofdstuk 3) geschreven onder begeleiding van Michel Mandjes, Marko Boon en Alex Kuiper. Een deel van het werk is ook door Alex Kuiper gepresenteerd op de VVSOR Annual Meeting (2024).

Literatuur

- N. Bailey. „A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times”. In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 14.2 (1952), p. 185–199.
- R. Mahes. „Appointment and Delivery Rescheduling”. PhD thesis. Universiteit van Amsterdam, nov 2024. ISBN: 978-94-6506-567-0.

Roshan Mahes is gepromoveerd (2024) aan de Amsterdam Business School en het Korteweg-de Vries Instituut van de Universiteit van Amsterdam.
E-mail: a.v.mahes@uva.nl



Promovendi van Willem Schaafsma
(bron: Mathematics Genealogy Project)

Promovendus	Jaar promotie
Tom Snijders	1979
Paul de Jong	1985
Vaclav Fidler	1986
Otto Kardaun	1986
Ton Steerberman	1987
Johan Akkerboom	1988
Wilhelmus Gerver	1988
Antonius Ambergen	1989
Klaas Poortema	1989
Anne van Streun	1989
Rienk Tamminga	1990
Bert van der Meulen	1992
Robert de Bruin	1993
Bert Kroese	1994
Diemer Salomé	1998
Bojan Basrak	2000
Casper Albers	2003

In memoriam: Prof. dr. Willem Schaafsma (1937 – 2024)

Casper Albers

Op 23 december 2024 is professor Willem Schaafsma overleden. Hij was erelid van de VVSOR en is zeer actief voor de vereniging geweest, onder meer in het bestuur en als editor van *Statistica Neerlandica*.

Schaafsma werd in 1937 geboren in Leeuwarden en promoveerde in 1966 bij prof. L.J. Smid aan het Mathematisch Instituut te Groningen op de monografie *“Hypothesis testing problems with the alternative restricted by a number of inequalities”*. Bij dit onderzoek maakte hij veel gebruik van de ZEBRA (“zeer eenvoudige binaire rekenautomaat”), de eerste computer in Groningen.

Hierna bezocht hij in Amerika de universiteiten van Berkeley, Atlanta en Pittsburg, waar hij samengewerkt heeft met statistische grootheden als Jerzy Neyman, David Blackwell, Erich Lehmann en C.R. Rao. Na de Amerikaanse periode keerde hij terug aan de Rijksuniversiteit Groningen, waaraan hij tot aan zijn emeritaat verbonden bleef.

Hij richtte zich in zijn onderzoek op de fundamente van het toetsen van nulhypotheseën en andere vormen van statistische inferentie, Walds decision theory en classificatiemethoden. Schaafsma publiceerde hierover met Groningse collega's en grote namen als C.R. Rao en David Cox.

Zijn passie lag bij onderwijs en onderzoek, maar bo-

venal bij het geven van statistisch advies aan derden. De pragmatische focus op de toepassing leverde hem publicaties op in een breed scala van wetenschapsgebieden. Hij heeft talloze jonge Groningse onderzoekers, tot aan Wubbo Ockels toe, voorzien van uitgebreidere adviezen dan waarom gevraagd werd. Ook voor zijn eigen studenten en promovendi maakte hij altijd tijd om zeer kritisch en gedetailleerd te lezen. De eindversie van mijn afstudeerscriptie staat vol met handgeschreven commentaren van Willem, maar ook de eindconclusie “je hoeft er niks meer aan te doen” en een prima cijfer.

Schaafsma was promotor van 17 promovendi (zie Tabel) waarvan er vijf (Tom Snijders, Ton Steerberman, Anne van Streun, Bojan Basrak en ikzelf) later zelf ook hoogleraar geworden zijn, vier ervan in Groningen. Hij is het statistisch geweten van meerdere generaties Groningse statistici en was de drijvende kracht achter de oprichting van de Groningse doctoraalstudie statistiek, die helaas na de introductie van het Bachelor/Master-systeem ter ziele gegaan is.

In de jaren negentig begeleidde Willem meerdere promovendi op het gebied van “distributional inference”. Na zijn pensionering stortte hij zich, samen met Otto Kardaun, volledig op dit thema met als doel een gezamenlijk boek hierover te schrijven. De totstandkoming van de definitieve versie van dit boek zal Willem helaas niet meer meemaken.



In memoriam: dr. Alessandro Di Bucchianico (1963 – 2024)

Edwin van den Heuvel

Alessandro Di Bucchianico, associate professor in industrial statistics and program director of the Bachelor and Master programs in applied mathematics at Eindhoven University of Technology (TU/e), passed away far too early on July 8, 2024, from cancer, after a very brief illness at the Catharina Hospital in Eindhoven.

Dit In Memoriam is ook geplaatst in het Nieuw Archief voor de Wiskunde.

Alessandro studied mathematics at the University of Amsterdam and graduated in 1987. Initially, he specialized in the field of mathematics known as “analysis”, and he began a PhD in this area at the University of Groningen. His PhD thesis, titled “Polynomials

of Convolution Type,” earned him his doctorate in 1991. He remained in Groningen as a postdoc for two years but decided to take a different professional path. He began focusing on applied statistics, becoming an assistant professor at TU/e in 1993, where he spent the remainder of his professional career.

In 2001, Alessandro was promoted to associate professor, marking a significant milestone in recognizing his research contributions in industrial statistics. In 2007, he joined the Laboratory for Industrial Mathematics Eindhoven (LIME) as the indispensable right hand man of its founder, Prof. Dr. Bob Mattheij. LIME was established as a research institute at TU/e to support industry through innovation in processes and products. Alessandro was instrumental in securing funding through statistical projects in industry. LIME achieved significant success as a business and joined Sioux in 2011, but Alessandro chose to re-

main at the TU/e to continue his academic career. In 2018, Alessandro became a program director for the education in mathematics and in 2022, the master program on industrial and applied mathematics received the George Smith INFORMS prize as the first university outside the USA.

Alessandro was a pioneer because he was one of the first academic researchers in the mathematics department of a Dutch university who focused on statistical science for real-world applications. This was rather unconventional in the 1990s, because statistics at mathematics departments in the Netherlands was dominated by mathematical statistics. I believe that researchers like John Tukey and George Box inspired him. Thus, Alessandro wanted to make a change, and being confident in his strong mathematical foundation, he sought to apply mathematical rigor to practical problems. The phrase “nothing is more practical than a good theory”, often attributed to Kurt Lewin, who was considered a pioneer in social psychology, resonated deeply with Alessandro.

When an academic researcher focuses on developing theory for practical applications across many areas, there is a risk that the work or results may lack concentrated impact in any single discipline or application domain because it is scattered across many different areas. Although Alessandro also published on many different topics, both in statistics and different types of applications, he gained strong recognition. He was primarily recognized for his expertise and research in reliability and monitoring and, to a somewhat lesser extent, on exact distributions and predictive maintenance. He successfully advanced these areas by linking theory to practical applications, solidifying his reputation as a researcher in industrial statistics.

Although it is impossible to fully cover Alessandro's contributions to statistics, I will highlight a few achievements in reliability and statistical monitoring, demonstrating his dedication to both theory and applications. He also made significant contributions to topics like nonparametrics (e.g., tolerance regions and null distributions) and density estimation, as well as providing entries in the Encyclopedia of Statistics in Quality and Reliability (e.g., Coefficient of Determination, Leverage, Measures of Scale, Outliers, Pareto Chart, and Sum of Squares), but his most prominent contributions were in reliability and monitoring.

Alessandro's early contributions to reliability were on software reliability, where he developed testing procedures to estimate and improve soft-

ware quality and worked on estimating reliability models for software systems and failure probabilities. He also created statistical tools for software reliability analysis that were useful for industry and society. Later, Alessandro contributed to the field of end-of-life management, where loss of economic value would occur before technical failure, and performance monitoring and prediction of component failures are more important. He also contributed to the field of lifetime analysis, where degradation data and accelerated tests are used to estimate products' (functional) lifetime, and where uncertainty quantification is necessary, in particular when machine learning methods are used for prediction. I believe that he was often motivated to work in this area to make sure that more mathematical rigor was applied in the field of reliability because he could become annoyed when confronted with sloppy and imprecise statistical practices.

In statistical monitoring, Alessandro's contributions are much more distinguished by “out-of-the-box” thinking, because he sought to revolutionize the field of Statistical Process Control (SPC) and go beyond the standard thinking in SPC. While the field was well-developed when he began his work, Alessandro believed that the statistical methods often relied on unrealistic or simplistic assumptions. Some of his notable contributions that demonstrate this out-of-the-box thinking include:

- Integrating SPC (focused on monitoring processes) with Automatic Process Control (focused on adjusting processes),
- Evaluating out-of-control situations beyond standard shift and trend assumptions,
- Monitoring high-volume processes, sparse events, and biological reference standards,
- Developing optimal sampling plans and generalized likelihood ratio control charts for monitoring impurities,
- Monitoring personal health for early disease detection,
- Linking predictability concepts with Shewhart's in-control monitoring concepts,
- Using copulas for multivariate monitoring of non-normally distributed data.

Besides his academic achievements and the societal impact of his work, Alessandro was deeply involved in the industrial statistics and mathematics communities. From 2007 to 2019, he served as director and webmaster of the European Network for Business and Industrial Statistics (ENBIS), a network founded in 1999 by Prof. Dr. Ronald Does to help connect

academia and industry on topics of business and industrial statistics. Professionals working closely with Alessandro during his period at ENBIS as director spoke highly of his professionalism and dedication to the network. Furthermore, he also contributed to nENBIS, the Dutch section of ENBIS, as part of the Dutch Society for Statistics and Operations Research (VSOR).

In addition to ENBIS, Alessandro played pivotal roles in several Dutch communities. He was instrumental in the first ten years of Eurandom, a national research platform on stochastics in the Netherlands established in 1998, which later evolved into a national workshop platform on the subject. Alessandro supervised many young researchers at Eurandom during its formative years. He also was a driving force behind Study Group Mathematics with Industry (SWI) and the innovation committee of the Dutch Mathematics Platform (PWN), connecting Dutch mathematicians with industry. His extensive knowledge, commitment, and professional network made him an effective and respected bridge builder.

Alessandro shared his passion for statistics—both its theory and applications—with students. His numerous courses combined deep statistical knowledge with real-world industrial problems to illustrate the societal relevance of mathematics. Many of his master's students completed their thesis on practical problems, delivering empirical proof and mathematical rigor to the solutions for these problems and often spending time in industry to fully grasp the topics. Alessandro was incredibly proud of this achievement, but also for his role in establishing and contributing to the “modeling week”, a mandatory element in the Applied Mathematics Master's program at TU/e. During this week, student groups tackled real challenges from the industry under the guidance of academic mentors and industry specialists. Alessandro was a dedicated supervisor during these weeks and responsible for attracting industry participants and ensuring meaningful collaborations. Many industry problems have been (partly) solved through the modeling week.

Besides his brilliance and numerous contributions to industry and academia, Alessandro was known for his kindness, integrity, and warmth as a person. He was always ready to assist others and was highly professional with a keen eye for detail. He was a great mentor, friend, and guiding light to many. His untimely passing is an immense loss to the broader statistics community in the Netherlands and beyond.



Vergroot je kennis tijdens het ISI World Statistics Congress 2025 in Den Haag

Ontdek het veelzijdige wetenschappelijk programma van dit toonaangevende congres, en breid je kennis uit tijdens de waardevolle Short Courses voorafgaand aan de congresdagen.

De 65e editie van het ISI World Statistics Congress zal van 5 tot 9 oktober 2025 plaatsvinden in Den Haag. Voor Nederlandse statistici en datawetenschappers is het congres een unieke kans om hun kennis te vergroten en te netwerken met collega's van over de hele wereld.

Boeiend wetenschappelijk programma

Met meer dan 160 Invited Paper Sessions en 90 Contributed Paper Sessions biedt het wetenschappelijk programma een schat aan statistische kennis en inzichten. Verken de nieuwste trends, uitdagingen en innovaties in de wereld van statistiek en datawetenschap. Het congres biedt een breed scala aan onderwerpen die aansluiten bij alle statistische vakgebieden. Bezoek de website <https://isi-next.org> en klik op Congress en vervolgens op Scientific Programme om het volledige wetenschappelijk programma te bekijken.

Kennis opdoen tijdens 1- en 2-daagse cursussen

Ook voorafgaand aan het congres zijn er mogelijkheden om je statistische kennis uit te breiden! Het Short Courses Programme vindt plaats als side-event in de dagen voorafgaand aan het congres. Dit programma omvat diverse 1- en 2-daagse cursussen, gegeven op 3 en 4 oktober door toonaangevende experts uit verschillende statistische vakgebieden. Deze cursussen bieden deelnemers de kans om zich te verdiepen in specifieke onderwerpen. Bezoek de website isi-next.org en klik op Side Events en vervolgens op Short Courses voor een overzicht van alle cursussen.

Aanmelden

Maak deel uit van dit vermaarde wereldwijde statistiek evenement waarvan de wortels teruggaan tot 1853! Vergroot je kennis, en bouw waardevolle internationale connecties op. Meld je aan vóór 12 mei via <https://isi-next.org> om te profiteren van de early bird tarieven. Na deze datum gelden reguliere tarieven.



Letter from the Upcoming President

Last summer, I was honoured to be asked by the VVSOR-board to consider succeeding Casper Albers as its president. Having been active for the society in different roles and at varying intensity during the past 35 years, but not so much during the past 10, this request came somewhat as a surprise. After discussions with several people, especially also on the current situation of the society and its plans, I gladly and confidently agreed to make myself available for the task and look forward to collaborating with the many active members of the society.

The VVSOR is a society with a rich history. It was founded 80 years ago, a period in which many developments were made in statistics, operations research and its intersections. At our forthcoming Annual Meeting, our former president, Gerrit Timmer, will share lessons learned during his long career simultaneously working in industry and academia with us.

The VVSOR is not only an association with a rich history, but definitely also with a future. To realise this, we only need to observe the scale at which data are currently being used to build computer models, which are combined with optimization techniques to automatize many things and support decisions. Activities of this kind go under many fancy names, but the (combined) role of statistics and operations research is evident. Still many developments are going on and needed, developments that ask for active involvement of people of our profession. At the Annual Meeting, Marjolein Fokkema will introduce particular challenges associated to the use of machine learning models in the behavioural sciences and how these can be addressed.

Past, present and future of the VVSOR come together in the prizes to be awarded at the Annual Meeting. Their names refer to three prominent figures in the history of the VVSOR: Willem van Zwet,

his promotor Jan Hemelrijk and his promotor David van Dantzig. History shows that many of the past awardees grew into influential and visible representatives of our community, something we also wish for the prize winners of this year.

I think we all feel that the theme "80 years of statistics and operations research in the Netherlands: past, present and beyond" of our forthcoming Annual Meeting is natural. It is a pleasure to announce this here. The meeting will be held on March 20, at familiar location "In De Driehoek" in Utrecht. The programme committee has managed to compose an interesting programme. Besides the two keynote contributions by dr. Marjolein Fokkema and prof. dr. Gerrit Timmer and talks by the various award winners, presentations will be held on subjects ranging from teaching statistics to causal inference and data integration to age-based maintenance in reliability theory. In addition to attending the talks, I warmly recommend participation in the social parts of the program, including lunch, drinks and dinner. Registration can be done via the website of the VVSOR: vvsor.nl, where also more information on the program, venue, costs etc. can be found. I am looking forward to meeting many of you at our meeting in Utrecht!

Geurt Jongbloed received his PhD from TU Delft in 1995, under supervision of Piet Groeneboom. He was appointed full professor of statistics at TU Delft in 2007. His research focuses on shape constrained statistical inference. He also works on applications of statistics in various fields, including materials science, health and climate. Within VVSOR, Geurt served as chair of the section Mathematical Statistics for several years and as member of the editorial board of *Statistica Neerlandica*. Outside the VVSOR, he acted as president of the Royal Dutch Mathematical Society, board member of PWN and chair of the department of Applied Mathematics at TU Delft. He currently serves as director of the Faculty Graduate School within the Faculty of Electrical Engineering, Mathematics and Computer Science at TU Delft.

80 Years of
Statistics and Operations Research
in the Netherlands

Past, Present and Beyond



VVSOR

Annual meeting
March 20, 2025

Annual Meeting of the Netherlands Society for Statistics and Operations Research (VVSOR)

Thursday March 20, 2025

10:15 – 17:45

In de Driehoek

Willemsplantsoen 1 C, 3511 LA Utrecht

- prof.dr. Gerrit Timmer (ORTEC)
- dr. Marjolein Fokkema (Leiden University)

Just like last year, this year's Annual Meeting will be at In de Driehoek in Utrecht. We will have a general assembly for members, followed by the actual event with two keynote speakers, three parallel sessions and two award presentations. The AM 2025 will be in English.

Attending the meeting at In de Driehoek (including drinks and lunch) costs 65 euro. Reduced price for students: 30 euro. Additional registration is required for dinner and pubquiz at De Rechtbank.

Please register on the [vvsor](http://www.vvsor.nl)-website
<https://www.vvsor.nl/vvsor-annual-meeting>

DATE

Thursday, March 20, 2025

VENUE

In de Driehoek, Willemsplantsoen 1C, 3511 LA Utrecht

REGISTRATION

Registration for the conference is mandatory at <https://www.vvsor.nl/vvsor-annual-meeting>. Detailed information can be found on our website.

LANGUAGE

The talks at the annual meeting will be in English.

ALGEMENE LEDENVERGADERING (ALV)

The Annual General Meeting of members (ALV) takes place on March 20, 10:15 – 11:15. The relevant documents will be e-mailed two weeks before the meeting.

SNACKS AND DRINKS

Lunch and drinks during the breaks will be provided.

DINNER WITH PUBQUIZ

Dinner at De Rechtbank, Utrecht. The pubquiz will be organized by the Young Statisticians.

ORGANIZING COMMITTEE

The annual meeting is organized by a special committee in cooperation with the board of the VVSOR. For questions, contact the organizers by email at annualmeeting@vvsor.nl.

PLEASE REGISTER BEFORE MARCH 17

- 9:45 - 10:15 **Registration + coffee & tea**
- 10:15 - 11:00 **ALV, General Assembly (members only)**
- 11:00 - 11:15 **Break with coffee & tea**
- 11:15 - 11:30 **Prof. dr. Casper Albers | Welcome & Opening of the AM 2025**
- 11:30 - 12:15 **Making an impact with OR and Statistics**
prof.dr. Gerrit Timmer
ORTEC
- 12:15 - 12:45 **Short presentations (session 1)**
- 12:45 - 13:30 **Lunch at In de Driehoek**
- 13:30 - 14:20 **Ceremony of the Willem R. van Zwet Award and the Jan Hemelrijk Award**
Prize winners will be presented by the juries, followed by a short presentation by the laureates
- 14:20 - 14:50 **Short presentations (session 2)**
- 14:50 - 15:00 **Short break**
- 15:00 - 15:40 **Ceremony of the Van Dantzig Award**
Prize winner will be presented by the jury, followed by a short presentation by the laureate
- 15:40 - 16:25 **Machine Learning in Behavioural Science**
dr. Marjolein Fokkema
Leiden University
- 16:25 - 16:30 **Wrap up & Finish**
- 16:30 - 17:30 **Drinks at In de Driehoek**
- 17:45 - 21:00 **Dinner + Pubquiz at De Rechtbank (extra registration required, walk-in from 17.45)**

Key note speaker 1

11:30 - 12:15

Making an impact with OR and Statistics

Gerrit Timmer
ORTEC

In the past 44 years, I have been in the position to observe hundreds of projects in various industries and application areas. Clearly, a lot has changed in computer capabilities, data availability, and available techniques. But independent of these changes, I noticed how subtle differences in circumstances and approach led to the impact varying from huge to none at all. I will summarize this experience in a number of lessons learned.

Gerrit Timmer is one of the founders of ORTEC in 1981 and is currently non-executive board member. He was chairman of the VVSOR from 1998-2003 and (part-time) professor of Business Econometrics at the Free University of Amsterdam for more than 25 years.

Short presentations (session 1)

12:15 - 12:45

Teaching Statistical Reasoning Using Quantitative Replication

Lucie Zicha
Leiden University

The challenge of fostering statistical literacy among high school and university students lies in the fact that it extends beyond merely applying statistical tools or interpreting results. Instead, it requires a broader framework that integrates rigorous reasoning skills, numeracy, and scientific literacy—specifically, an understanding of the scientific method. The close relationship between statistical literacy and scientific reasoning presents pedagogical challenges: to effectively cultivate statistical literacy, students must also develop a solid foundation in scientific reasoning. The replicatED project makes the case that teaching statistical literacy through quantitative replication addresses this issue directly as it bridges the disconnect that students often manifest between the mechanical understanding of the statistical technique (the so called “cookbook” approach) and struggling to relate it confidently to the substance of the question and the scientific research design. Replication work challenges students to critically assess methodology, measurement, and interpretation of findings. Students learn to scrutinize research claims, and identifying potential sources of error and bias. Lastly, replication teaches students about “researcher degrees of freedom”—the numerous points where subjective choices can impact results. In our paper, we introduce the replicatED project developed at Leiden University College that integrates replication as a pedagogical tool into statistics education. The project features an online database of replication datasets, searchable by the method used in the original study, and offers a range of pedagogical resources tailored to different levels of statistical knowledge among students.

Lucie Zicha is an Assistant Professor of Quantitative Research Methods at Leiden University College (LUC). Since 2015, she has served as the convenor of the quantitative methods track in the BSc program at LUC, and in 2024, she became the Major Team Lead for the BSc Governance, Economics, and Development program. She co-developed LUC's first-year Introduction to Statistics course and has played a pivotal role in advancing quantitative education within the curriculum. Lucie holds a PhD in Comparative Politics and Quantitative Methods from Binghamton University, U.S.

Short presentations (session 1)

12:15 - 12:45

Causal Inference in Finite Samples: The Potential of Invalid Adjustments

Nadja Rutsch
VU Amsterdam

Traditional covariate selection methods for causal inference focus on achieving unbiasedness and asymptotic efficiency. In many practical scenarios, researchers must estimate causal effects from observational data with limited sample sizes or when covariates are difficult or costly to measure. Their needs might be better met by selecting adjustment sets that are finite sample-optimal in terms of Mean Squared Error (MSE). We aim to find the adjustment set that minimizes the MSE of the causal effect estimator, taking into account the joint distribution of the variables and the sample size. We present examples where the MSE-optimal adjustment set differs from the asymptotically optimal adjustment set. To identify the MSE-optimal adjustment set, we introduce a sample size criterion for comparing adjustment sets in linear Gaussian models. We develop graphical criteria to reduce the search space for this adjustment set based on the causal graph. In experiments with simulated data, we show that the MSE-optimal adjustment set can outperform the asymptotically optimal adjustment set in finite sample size settings, making causal inference more practical in such scenarios.

Nadja Rutsch is a PhD student in the BayCause project. Her research focuses on Bayesian causal inference in high-dimensional data, under the supervision of Dr. Stéphanie van der Pas (Amsterdam UMC) and Dr. Sara Magliacane (UvA). Her work involves using directed acyclic graphs (DAGs) to improve our understanding of causal relationships.

Her academic background includes a Master's in Artificial Intelligence from the University of Amsterdam, where she graduated cum laude and completed a thesis on joint intervention detection and causal discovery.

Short presentations (session 2)

14:20 - 14:50

Bayesian Survey Data Integration: Can We Enhance Inference and Reduce Costs?

Camilla Salvatore
Utrecht University

Traditional probability sample surveys, the gold standard for population inference, are becoming increasingly expensive and suffer from declining response rates. As a result, researchers are increasingly turning to less costly but potentially biased non-probability sample surveys. This research addresses a crucial question: Can integrating a small probability sample with a larger non-probability sample improve analytic inference while reducing survey costs? We answer this question through an application involving logistic regression, utilizing both simulated and real-world data (e.g., volunteering, voting behaviour). Our findings reveal that integrating data sources significantly lowers the Mean Squared Errors of regression coefficients, compared to scenarios without integration, and can result in cost savings of up to 70 percent. A Shiny web app allows for deeper exploration of the methods and results. We also discuss the potential of extending this approach to studies that incorporate both self-reported and objective measurements (e.g., medical data, digital trace data).

Camilla Salvatore Dr. Camilla Salvatore works as an assistant professor at the department of Methodology and Statistics at Utrecht University, where she specializes in survey research. Her interests include survey data integration, inference with non-probability samples, survey weighting, nonresponse, the use of digital trace data and their integration with surveys.

Short presentations (session 2)

14:20 - 14:50

Combining age-based maintenance with an imperfect fault detection model

Ingeborg de Pater
Delft University of Technology

Age-based maintenance is a popular maintenance strategy where components are always maintained at a certain age or upon failure. There is, however, an increasing interest in predictive (also called condition-based) maintenance instead. In predictive maintenance, the measurements of the sensors installed around a component are used to develop prognostic models. A prognostic fault detection model detects when a component becomes unhealthy, but has not failed yet. Unfortunately, prognostic models are usually imperfect, with false positives and false negatives. By solely planning maintenance based on the outcomes of an imperfect prognostic model, the number of failures and maintenance tasks might actually increase, compared to age-based maintenance.

However, also an imperfect prognostic model might still provide useful information on the potential failures of the components. In this presentation, we will therefore combine an age-based maintenance strategy with an imperfect prognostic fault detection model. With Bayes theorem, we derive a formula for the probability of a false positive based on the age of the component. We subsequently use the classical renewal reward theory to optimize the age at which to preventively replace components, if we also replace components based on the alarms of the imperfect fault detection model. We analyse this approach with a small case study, where we assume a Weibull lifetime distribution for a component. We show how even prognostic models with a very high false negative rate can still contribute to lowering the maintenance costs.

Ingeborg de Pater started this summer as an assistant professor at the TU Delft on predictive maintenance and maintenance optimization for aircraft. Her research focuses on developing health estimation and RUL prediction models, the validation and uncertainty quantification of these models, and finding the optimal maintenance moment under uncertainty. Before this, she completed her PhD at the TU Delft on the same topic.

Key note speaker 2

15:40 - 16:25

Machine Learning in Behavioural Science

Marjolein Fokkema
Leiden University

Flexible algorithms from machine learning can provide highly accurate predictions, but they are a black box to users. This is problematic for behavioural scientists, who not only want to predict human behaviour, but also want to understand how it works. We are developing effect sizes that quantify the shape and magnitude of predictor variables' effects in ML models. Furthermore, we develop uncertainty quantification, to allow for hypothesis tests and meta-analysis with ML. This will allow behavioural scientists to not only profit from the highly accurate predictions of ML, but to also use it for testing and improving scientific theories.

Marjolein Fokkema is associate professor at Leiden University's Institute of Psychology



Statistische prijzen en de erfenis van Stella Cunliffe

Richard Starmans

Het instellen van internationale wetenschappelijke prijzen is belangrijk voor de status en zichtbaarheid van een vak. Lange tijd kwam de statistiek er wat dit betreft nogal bekaaid af, maar daarin is de laatste jaren enige verandering gekomen. Aan de vooravond van het ISI

World Statistics Congres, dat van 5 t/m 9 oktober 2025 in Den Haag plaatsvindt, laten we enkele ontwikkelingen kort de revue passeren.

Prijzen en controversen

De instelling en toekenning van wetenschappelijke prijzen blijkt niet zelden een hachelijke onderneming, die gepaard gaat met debat, verwijten van willekeur, persoonlijke voorkeuren, verstoorde politieke verhoudingen of andere onverwikkelijke zaken. Indien we de Amerikaanse historicus van de statistiek Stephen Stigler (1941) mogen geloven, lijkt dit alles onvermijdelijk. Zijn "Wet van de Eponymie" stelt dat alle resultaten in de exacte wetenschappen getooid gaan met namen van onderzoekers die nauwelijks of in het geheel niet bij de ontdekking of uitvinding ervan betrokken waren (Stigler, 1980). Dat hij vervolgens met een knipoog en een kwinkslag deze bewering aanduidt als "Stiglers Wet" toont zijn

talent voor zelfspot, maar de ondertoon is ernstig. Zegt deze verkapte "onmogelijkheidsstelling" niet iets over de grenzen van de historiografie? Indien we er niet eens in slagen wetenschappelijke prestaties juist te adresseren, hoe kunnen we ze dan rechtvaardig belonen?

Ook de Nobelprijs kent een roemruchte geschiedenis, vol mysteries en controversen. De Engelse letterkundige Burton Feldman (1926-2003) laat zich in zijn boek *The Nobel Prize: A History of Genius, Controversy, and Prestige* uit 2000 daarover dan ook niet onbetuigd. Een en ander begon al met de instelling ervan in de 19e eeuw. Waarom introduceerde Alfred Nobel (1833-1896) geen Nobelprijs voor wiskunde en wel een voor literatuur en voor de vrede? Over deze vraag wordt tot op heden gespeculeerd. Sommigen wijzen op de slechte verstandhouding tussen Nobel en de vooraanstaande Zweedse wiskundige Gösta Mittag-Leffler (1846-1927). Deze zou een relatie met Nobels vriendin Sofie Hess zijn begonnen. Uit angst dat deze man de prijs wel eens in de wacht zou kunnen slepen, zou Nobel van de weeromstuit hebben afgezien van de instelling van een Nobelprijs voor wiskunde. Saillant is dat de historica Erica Rummel in 2017 een uitvoerige studie aan de kwestie wijdde, getiteld *A Nobel Affair; The Correspondence between Alfred Nobel and Sofie Hess*,

maar ook daarmee werd de zaak niet geheel duidelijk. Aannemelijker is dat Nobel als scheikundige en praktiserend ingenieur niet overtuigd was van de fundamentele betekenis van de wiskunde en haar directe bijdrage aan vooruitgang en het heil van de mensheid. Dat hij de voorkeur gaf aan een op voorhand betwistbare Nobelprijs voor de vrede had volgens sommigen bovendien alles te maken met het oppoetsen van zijn blazoen. Nobel besefte dat dit laatste hard nodig was, vooral toen hij in 1888 slachtoffer werd van een persoonsverwisseling en in de krant zijn eigen necrologie las, die abusievelijk door een slordige journalist was gepubliceerd. Deze meldde dat de "merchant of death" was overleden; een man die lange tijd werkzaam was geweest in de wapenindustrie, dynamiet had uitgevonden, daarmee onheil over de mensheid had gebracht en daaraan ook nog eens flink had verdiend.

Hoe dan ook, wiskundigen moesten lang wachten op een "substituutprijs", de Abel Prize, die pas in 2003 voor het eerst werd toegekend en sterk overeenkomt met de Nobelprijs, wat betreft prijzengeld, frequentie, uitreikingsceremonie, selectiecriteria en ook het ontbreken van een leeftijdsgrens voor de kandidaten. Mede daardoor lijkt de prestigieuze Fields Medal, die een leeftijdsgrens van 41 hanteert, enigszins naar de achtergrond te zijn verdrongen. Dat mag zo zijn, maar hoe staat het met de internationale erkenning van en waardering voor fundamenteel of toegepast statistisch onderzoek met aantoonbaar grote maatschappelijke impact? Lange tijd kwam de statistiek er wat dit betreft nogal bekaaid af, maar daarin is de laatste jaren enige verandering gekomen. Aan de vooravond van het tweejaarlijkse ISI World Statistics Congres, dat van 5 t/m 9 oktober 2025 in Den Haag plaatsvindt, laten we enkele recente ontwikkelingen kort de revue passeren.

Wickham en de R-revolutie

Sedert 1981 kent de Committee of Presidents of Statistical Societies (COPSS) jaarlijks de *COPSS Presidents' Award* toe aan een jonge statisticus die zich heeft onderscheiden wegens "outstanding contributions to the profession of statistics". De eerste laureaat was Peter Bickel (1940) en inmiddels is de prijs 45 maal toegekend, steeds ongedeeld, dat wil zeggen aan één onderzoeker, en bovendien vrijwel altijd aan statistici werkzaam aan een Amerikaanse of Canadese universiteit. In 2019 viel de eer ten deel aan de Australische statisticus en datascientist Hadley Wickham (1979). Wickham was en is direc-

teur van RStudio (sedert 2022 hernoemd tot Posit PBC), een geïntegreerde ontwikkelomgeving met grafische gebruikersinterface, die inmiddels 15 jaar bestaat en zeer populair is in de open-source R community. Wickham zelf bouwde in deze gemeenschap een cult-status op, mede dankzij door hem geschreven R-packages als *dplyr* (voor data wrangling) en *ggplot2* (visualisatie), ontwikkeld volgens een strikte "filosofie" en ondergebracht in het onvolprezen *tidyverse*. Daarmee werden generaties datascientists die statistiek met programmeren trachtten te combineren en een gedegen achtergrond in informatica ontbeerden op weg geholpen in R, bij uitstek de taal ontwikkeld "door en voor statistici". Dit alles geldt *a fortiori* voor onderzoekers in de sociale en biomedische wetenschappen, die lange tijd vertrouwden op standaardpakketten als SAS en SPSS, maar nu de beschikking kregen over een eigen experimenteer- en ontwikkelomgeving, een virtueel laboratorium met een uitgebreide toolkit aan recente statistische en machine learning technieken. Zijn "methodologie" wordt onder meer ondersteund door het toegankelijke en inmiddels in vele talen vertaalde boek *R for Data Science* dat Wickham samen met Garrett Grolemund in 2017 publiceerde (Salsbury, 2001). Daarmee speelde hij een belangrijke rol in de R-revolutie, zoals die in voornoemde disciplines gestalte kreeg.

De jury benoemde in haar rapport een viertal specifieke redenen voor de toekenning: "for influential work in statistical computing, visualization, graphics, and data analysis; for developing and implementing an impressively comprehensive computational infrastructure for data analysis through R software; for making statistical thinking and computing accessible to a large audience; and for enhancing an appreciation for the important role of statistics among data scientists".

Duidelijk is dat Wickham niet zozeer zelf grensverleggende theoretische bijdragen leverde of nieuwe statistische methoden ontwikkelde, maar als bruggebouwer onderzoekers aanmoedigde het tijdperk van big data te betreden en zelfs hun methodologie te hervormen. Daarmee oefent hij daadwerkelijk invloed uit op de onderzoeks- en beroepspraktijk en vergroot hij de reikwijdte van en waardering voor de statistiek.

De jury van de COPSS-Award lijkt steeds meer te onderkennen dat vooruitgang vooral op het snijvlak van verschillende disciplines plaatsvindt en in de "trading zones" om met Peter Galison te spreken (Starmans, 2024). Dat blijkt ook uit meer recente toekenningen van deze prijs, waaraan overigens een bescheiden prijzengeld van \$2000,- is verbonden.



De poëzie van de wetenschap

Dit jaar vindt in Den Haag de 65e editie van het ISI World Statistical Congress plaats. Dit tweeverjaars event wordt georganiseerd onder auspiciën van het aldaar gevestigde International Statistical Institute (ISI) en tijdens de bijeenkomst wordt ook de *International Prize in Statistics* uitgereikt. Deze prijs, waaraan een bedrag van 80.000 EURO is verbonden, bestaat pas sedert 2016, maar kent al enkele illustere laureaten. In 2017 was dat Sir David Cox (1924-2022) voor zijn bijdragen aan survival-analyse binnen geneeskunde, wetenschap en engineering. Zijn parametrische proportional hazards model is al decennialang voor velen een begrip. Vervolgens viel in 2019 de eer ten deel aan Bradley Efron (1938) voor zijn werk op het gebied van de bootstrap, dat eveneens weinig toelichting behoeft. Sedert het eind van de vorige eeuw krijgt de inferentiële statistiek ook volop gestalte via computer-intensieve statistische methoden. Denk aan de aloude exact-tests, permutatie-tests en randomisatie-tests, aan de populaire Monte Carlo methoden, maar zeker ook aan bootstrap sampling distributions, die in pakketten zoals SPSS al vele jaren beschikbaar zijn. In 2021 ging de prijs naar de Amerikaanse statistica Nan Laird (1943) met name voor haar onderzoek op het terrein van longitudinale studies. Ook het expectation-maximization (EM) algorithm voor ML-estimation, dat zij in 1977 samen met Arthur Dempster en Donald Rubin introduceerde, vindt in brede kring toepassing. Tot slot werd de International Prize in Statistics in 2023 toegekend aan de Indische statisticus C.R. Rao (1920-2023). Elke statisticus kent de Cramér-Rao bound en de Rao-Blackwell stelling, die ruim driekwart eeuw eerder reeds waren gepubliceerd. Maar de jury roemde ook zijn werk op het interdisciplinaire terrein van de "information geometry", dat volgens het rapport een belangrijke rol speelt bij het begrijpen en optimaliseren van metingen van/aan het Higgsboson. Uiteraard laten de vier toekenningen van deze jonge prijs nog geen verregaande conclusies toe. Vooralsnog gaat het ook hier om ongedeelde prijzen, al is dat volgens de ISI-website geen noodzakelijkheid. De verschillen met de COPSS-award zijn duidelijk en betreffen niet alleen het prijzengeld, maar vooral

het ontbreken van leeftijdsgrenzen. Door dat laatste ligt het voor de hand dat vooral emeriti-hoogleraren de prijs ontvangen, soms als een soort oeuvreprijs, maar bovenal voor resultaten die de tand des tijds hebben doorstaan, de reikwijdte en (h)erkenning van statistiek vergroten en ook een rol spelen bij canonicvorming.

Feit is dat ISI met deze prijs een belangrijke apologie van de statistiek beoogt en zij kwijt zich met verve van deze taak. De website kent dan ook vele ronkende teksten en bloemrijk taalgebruik wordt niet geschuwd. Zo wordt de Amerikaanse schrijver Frank Emerson Andrews ten tonele gevoerd, die statistiek typeert als "the poetry of science" en ISI doet daar nog een schepje bovenop door het vak te kenschetsen als "the soul of scientific enquiry". Het instituut benadrukt dat de lijst met bijdragen aan "the advancement of humankind" schier eindeloos is, en laat vele toepassingsgebieden één voor één de revue passeren. Vaak is die statistische bijdrage gemaskeerd of impliciet. Verschillende Nobelprijzen voor de Economie betreffen puur statistisch onderzoek; samenwerking met astronomen heeft pas recentelijk tot de subdiscipline van "astrostatistics" geleid; computationele biologie betreft voor een belangrijk deel "statistical modeling of large databases of genomic and proteomic measurements". Dit alles geldt *a fortiori* voor vele resultaten binnen machine learning, aldus het ISI, dat tegelijkertijd stelt dat statistics helaas nog steeds een "invisible science" is "unknown to the general public and, perhaps more troubling, not valued as a key contributor to research by most scientists in other fields".

Een eigentijdse mecenas

Ook de Belgische statisticus Peter Rousseeuw (1956) was ontevreden met de status quo en zag de positie van statistiek als louter hulpdiscipline weerspiegeld in het landschap van prijzen en awards. Zelf zou hij een vermogen verdienen bij het hedgefonds Renaissance Technologies, opgericht door de wiskundige en investeerder James Simons (1938-2024). Hier werken niet zozeer traditionele beursgoeroes of "wolves" van Wall Street, maar vooral wiskundigen,

statistici en informatici en het bedrijf behaalde jarenlang hoge rendementen. Met het verdiende geld besloot Rousseeuw een eigen internationale prijs voor statistiek in te stellen, ruwweg opgesteld naar het model van de Nobelprijs. De *Rousseeuw Prize for Statistics* behelst één miljoen dollar en wordt tweeverjaars uitgereikt onder auspiciën van de Koning Boudewijnstichting. In 2022 werd de prijs toegekend aan de Amerikaanse biostatisticus James Robins en aan zijn (oud-)collega's Miguel Hernán, Thomas Richardson, Andrea Rotnitzky en Eric Tchetgen voor hun onderzoek op het gebied van Causal Inference, toegepast in geneeskunde en volksgezondheid. Het geld werd volgens een door de jury bepaalde verdeelsleutel aan de laureaten uitgekeerd. Dat Causal Inference werd gehonoreerd is om twee redenen saillant. Causaliteit was ooit een obscuur begrip, dat volgens velen uit de taal der wetenschap diende te worden verbannen, maar is nu volop in de mode. Na de toekenning van de Turing-award aan AI-wetenschapper Judea Pearl in 2011 en na de toekenning van de Nobelprijs voor economie aan econometrist Guido Imbens in 2021 wordt opnieuw onderzoek naar causaliteit met een belangrijke prijs gehonoreerd. Bovendien maakt de toekenning aan Robins c.s. duidelijk dat ook statistici een belangrijke rol spelen in het causaliteitsonderzoek en dat dit niet het exclusieve domein is van informatici en AI-ers, zoals statistiek-"basher" Pearl, die in zijn *The Book of Why* uit 2018 in "Jaccusee stijl" tegen de statistiek van leer trekt (Starmans, 2018).

De tweede toekenning is nog relevanter. In 2024 ging de prijs naar de Israëlische statistici Yoav Benjamini, Daniel Yekutieli en Ruth Heller, allen werkzaam aan de Universiteit van Tel Aviv. De prijs werd toegekend vanwege hun pionierswerk op het terrein van "False Positivity Rate (FDR) and methods to control it", op een moment dat "schijnontdekkingen" volop in de belangstelling staan en velen een "replicatiecrisis" ontwaren. Het begon allemaal in 1990 toen Benjamini een belangrijke paper schreef samen met Yosef Hochberg, die van de jury postuum een eervolle vermelding ontving. Beiden hadden kennisgenomen van Branco Soric's werk, die in 1988 zijn paper "Statistical 'Discoveries' and Effect-Size Estimation" had gepubliceerd en waarschuwde dat het negeren van issues rond multiple testing tot vele foute resultaten in de wetenschappen zou leiden. Dat was ruim 15 jaar voordat John Ioannides in 2005 zijn gerucht makende artikel "Why Most Published Research Findings Are False" deed verschijnen. Benjamini en Hochberg realiseerden zich dat de proportie foute ontdekkingen niet louter een signaal of waar-

schuwing was, maar kon fungeren als selectie criterium, dat zij vervolgens formeel definieerden als de "proportion of false discoveries" en benoemden als "False Discovery Rate". In hetzelfde artikel stelden zij een procedure voor om de FDR te controleren "using marginal p-values", die bekend zou worden werd als de Benjamini-Hochberg procedure (BH).

De receptie was ronduit slecht. Deel 1 werd door verschillende journals afgewezen en verscheen pas in 1995, deel 2 pas in 2000. Door Yekutieli en Heller werden belangrijke uitbreidingen en verbeteringen aangebracht en de relevantie van dit alles nam volgens de jury toe toen begin deze eeuw het wetenschappelijk onderzoek een "industrialisation process" onderging. Schaalvergroting en automatische analyse en verwerking van big data maakten nieuwe methoden voor multiple comparisons / multiple testing nodig "to cater to the emerging needs of the scientific community". De door de laureaten ontwikkelde concepten en technieken spelen hierbij een belangrijke rol, aldus de jury. Hoe dan ook deze prijs, die verreweg het hoogst is gedoteerd, laat na twee toekenningen vanzelfsprekend nog minder dan de International Prize in Statistics verregaande conclusies toe. Bij de toekenning spelen diverse factoren een rol, maar zeker ook reikwijdte en toepasbaarheid van statistiek in grootschalig (industriële) empirisch onderzoek. Een traditie, die in de geschiedenis van de statistiek al een fameuze voorloper kende. De Amerikaanse statisticus William Edwards Deming (1900-1993) vond in eigen land geen gehoor, maar zou in het naoorlogse Japan de industrie (met name staalindustrie, auto-industrie en elektronica) moderniseren op basis van zijn statistische methodologie voor systematische proces- en kwaliteitscontrole. David Salsburg riep hem daarom ooit uit in zijn onvolprezen essaybundel *The Lady Tasting Tea* tot "the man who remade industry" (Salsburg, 2001).

De erfenis van Stella Cunliffe

Het is dit jaar precies een halve eeuw geleden dat de Britse statistica Stella Cunliffe (1917-2012) als eerste vrouw werd benoemd tot President van de Britse Royal Statistical Society. In haar inauguratie-rede, die zij in 1976 publiceerde, nam zij geen blad voor de mond en schroomde niet om haar collega's op een luchtige toon flink de les te lezen. Zo vond zij dat statistici zich te veel in een ivoren toren ophielden, aldaar "with elegant brilliance" theoretische problemen oplossen met een overmaat aan wiskundig vernuft, maar geen belangstelling aan de dag



legden voor de statistische problemen van “research workers” in de alledaagse praktijk. Of het nu ging om universitaire wetenschappers, biochemici in laboratoria, industriële onderzoekers of ambtenaren die worstelden met “official statistics”. Statistiek in het algemeen en de Society in het bijzonder kunnen alleen hun waarde voor de samenleving bewijzen door actief de dialoog te zoeken met andere disciplines en beroepsvelden, aldus Cunliffe (Cunliffe, 1976). Juist wanneer deze wellicht “less scientific than we may like” blijken te zijn, moet de statisticus binnentreden “in what I can only describe as the airy-fairy world of the sociologist, and if I dare say it, sometimes of the psychologist”. Ofschoon zij soms enigszins geringschattend spreekt over de “less elevated” en hun “slapdash methods” legt zij de schuld toch grotendeels bij haar collega’s, die verzuimen “to proselytize our ware to social scientists” en andere wetenschappers. Zij moeten hun waarde gedurende het gehele onderzoekstraject bewijzen, een solide methodologie aanbieden, maar dan wel de taal van de onderzoeker leren en vooral niet gebieden als Official Statistics met dedain terzijde schuiven.

Bij het grote publiek is Cunliffe vandaag de dag grotendeels onbekend en ook in menig historisch overzichtswerk ontbreekt zij. Het is dan ook niet meer dan terecht dat Salburg in zijn voornoemde boek een apart essay wijdt aan haar visie en invloed (Salsbury, 2001). Natuurlijk is er inmiddels veel veranderd. Toen Cunliffe haar rede uitsprak bestond geen van de drie besproken statistische prijzen en van een computationele / statistische omwenteling in de voornoemde domeinen was meestal evenmin sprake. Toch dragen alle prijzen een deel van de boodschap uit, wat betreft interactie, het streven naar zichtbaarheid bij wetenschappers en het grote publiek en de betekenis voor de samenleving. Indien we ISI en Rousseeuw mogen geloven is er tegelijkertijd -ondanks of wellicht deels door deze nadruk op dienstverlening- nog veel werk te doen om de status van statistiek als “invisible science” te

verbeteren. In 1925, precies honderd jaar geleden publiceerde Ronald Fisher zijn inmiddels klassieke *Statistical Methods for Research Workers*. Maar de “research workers” van de moderne tijd spreken de taal van machine learning en AI en zijn niet altijd overtuigd van het belang van een statistische methodologie. In het AI-data-science debat spelen statistici een bescheiden rol. En sterker nog, wanneer AI-sceptici doemscenario’s schetsen door de opmars van deep learning en generatieve AI benadrukken zij dikwijls dat het slechts om statistiek gaat, correlaties en kansen, geen echte kennis, et cetera. Die pejoratieve duiding beschreef historicus van de statistiek Ian Hacking al in zijn *The Taming of Chance* uit 1990, waarin hij schetst hoe leidende intellectuelen in de 19e eeuw statistiek in het beklagdenbankje plaatsten. Tegen dit alles bieden de hier besproken organisaties en prijzen dan ook een welkom tegenwicht.

Literatuur

- S. Cunliffe. „Interaction”. In: *Journal of the Royal Statistical Society* 139.Part 1 (1976).
- D. Salsbury. *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*. New York: Henry Holt en Company, 2001.
- R. J. C. M. Starmans. „Statistiek en Causaliteit; voortschrijdende liaison of moeizame samenspraak”. In: *STATOR* 19.4 (2018).
- R. J. C. M. Starmans. „Statistiek, astronomie en de constructie van data”. In: *STATOR* 25.3 (2024).
- S. M. Stigler. „Stigler’s Law of Eponymy”. In: *Transactions of the New York Academy of Sciences* 39 (1980). Red. door F. Gieryn.

Richard Starmans is verbonden aan de Faculteit Bèta-wetenschappen (Department of Information and Computing Sciences) van de Universiteit Utrecht en aan Tilburg University. Hij doet onderzoek op het snijvlak van filosofie, statistiek, informatica en AI.
E-mail: starmans@cs.uu.nl



source: Staartjes Fotografie Drenthe, Ewout Staartjes

Optimalisatie voor de sturing van de grootste pompen van Nederland

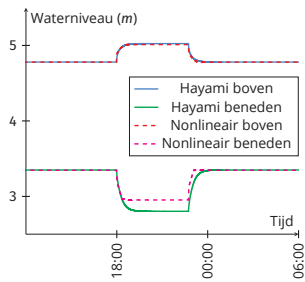
Kieran van Gaalen, Klaudia Horváth en Rens de Heer

Hoe helpt optimalisatie om droge voeten te houden? Om onze voeten droog te houden, moeten de waterpeilen in de binnenwateren tussen bepaalde grenzen blijven. Dat betekent dat het meeste water dat via rivieren en regen Nederland binnenkomt, ook weer afgevoerd moet worden. Een groot deel van dit water stroomt naar de Noordzee via pompen en sluizen.

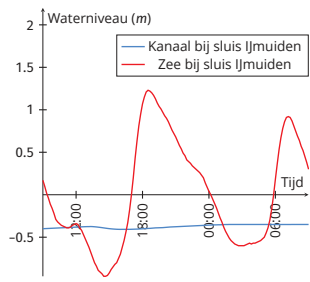
Maar hoe zorgen we ervoor dat dit proces zo veilig en milieuvriendelijk mogelijk verloopt? Optimalisatie kan daarbij een belangrijke rol spelen.

Wanneer en hoeveel water moet gepompt worden?

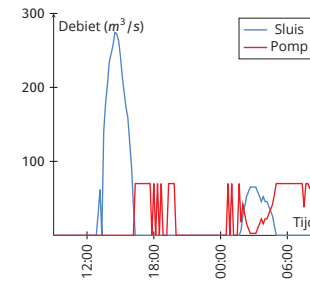
Elk jaar wordt 3 miljard m³ water afgevoerd vanuit Nederland via het Spui- en Gemaalcomplex IJmuiden, dit is 2000 keer de Grote Plas bij Maarseveen. Het water stroomt door het Amsterdam-Rijnkanaal dat een open verbinding met het Noordzeekanaal heeft, waarna door de sluizen en pompen bij IJmuiden het water naar de zee gevoerd wordt. Het waterniveau van het Noordzeekanaal - Amsterdam-Rijnkanaal moet op peil gehouden worden ten behoeve van veiligheid en scheepvaart, terwijl voor ecologische redenen ook de afvoer en de waterkwaliteit binnen bepaalde grenzen moeten blijven.



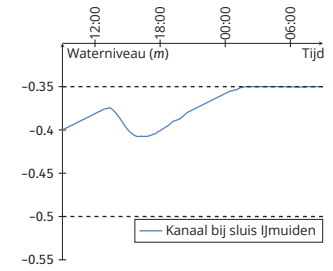
Figuur 1: Waterniveaus van Hayami en niet-lineaire model.



Figuur 2: Waterniveau aan beide kanten van de sluis bij IJmuiden.



Figuur 3: Debiet van de sluis en pomp bij IJmuiden.



Figuur 4: Waterniveau aan kanaalzijde IJmuiden.

De grootste stuurknoppen van dit systeem zijn de pompen en sluisen bij IJmuiden. Hoe kunnen we het waterniveau op peil houden en de sluisen en pompen zo efficiënt mogelijk inzetten, rekening houdend met alle bedieningsregels van deze objecten (bijvoorbeeld minimum aan en uit tijden) en overige eisen (bijvoorbeeld zo visvriendelijk mogelijk pompen) die aan het watersysteem gesteld worden? Wiskundige optimalisatie geeft een antwoord.

Oplossing

Om het waterniveau in het Noordzeekanaal onder controle te houden moet er een pomp- en sluischema gevonden worden voor de komende dag, waarbij gedurende elke uur bekend moet zijn hoeveel water door de verschillende pompen en sluisen afgevoerd moet worden. We lossen deze vraag op als een MILP (mixed-integer linear optimization).

Eerst moet de beweging van water in kanalen gemodelleerd worden. Waterniveau en afvoer (ook wel debiet genoemd) in een kanaal staan in verhouding met elkaar. Als in een stuk kanaal meer water instroomt dan uitstroomt, zal het waterniveau stijgen. Deze verhoudingen, samen met aspecten als zwaartekracht en wrijvingskracht worden beschreven door de Saint-Venant vergelijkingen. Dit zijn twee partiële differentiaalvergelijkingen, waarvan er een niet-lineair is. Ook de formules die het spuidebiet beschrijven zijn niet-lineair. Om snelle lineaire optimalisatie technieken te kunnen gebruiken, gaan we voor deze vergelijkingen lineaire benaderingen gebruiken.

Een van deze benaderingen voor de Saint-Venant formules is de lineaire Hayami vergelijking (Hayami, 1951). Deze kan in de plaats van de niet-lineaire Saint-Venant vergelijking gebruikt worden. Dit model is gebaseerd op een equilibrium van waterhoogte. Hoe meer het model daarvan afwijkt, hoe groter de foutmarge van de benadering zal zijn. Om dit

model te kunnen gebruiken moet er nog een discretisatie op worden toegepast, in de vorm van de eindige-differentiemethode. Elk tijd- en ruimtepunt wordt hierin benaderd door het verschil van de twee punten ernaast te berekenen en te delen door de afstand tussen die twee punten. Dit geeft ons lineaire vergelijkingen. Als die vergelijkingen zijn gebruikt samen met randvoorwaarden in de vorm van debiet of waterniveau aan de uiteinden van het kanaal, kan het waterniveau en debiet op elk punt van het kanaal berekend worden als het waterniveau en debiet van de vorige tijdstap bekend is. Om voor elke tijdstap dit te kunnen berekenen moet er een beginsituatie gegeven worden. Met al deze informatie kan voor elke tijdstap op elk discrete punt van het kanaal het waterniveau berekend worden.

Het debiet van een sluis is afhankelijk van de wortel van het waterhoogte verschil (Te Chow, 1959). Om dit te lineariseren pakken we twee punten waartussen de meest voorkomende waterhoogte verschillen zitten, en trekken we een lijn daartussen. Dit zal voor de meeste situaties een redelijke benadering zijn, maar zal, net zoals Hayami, meer gaan afwijken hoe verder het model van het equilibrium gaat afwijken.

Door het vinden van lineaire benaderingen kunnen veel snellere optimalisatietechnieken toegepast worden dan bij niet-lineaire vergelijkingen. Eerst gaan we de gevonden formules toevoegen als voorwaarden van een lineair programmeerprobleem (LP). De randvoorwaarde voor een kanaalsegment naast een pomp of sluis zal het debiet van de pomp of sluis zijn. Hiermee kan dan het waterniveau berekend worden, wat zelf ook niet hoger of lager dan een aangegeven waarde mag zijn. In het huidige model is het echter nog mogelijk om de sluis te openen wanneer de buitenwaterstand op zee hoger is dan het kanaalpeil. Om dit te voorkomen, dient een binaire variabele te worden toegevoegd

die de waarde nul aanneemt zodra de buitenwaterstand op zee hoger is. Dit brengt echter een aanzienlijke vertraging met zich mee, aangezien een lineair programmeringsprobleem in polynomiale tijd oplosbaar is, terwijl een mixed-integer lineair programmeringsprobleem (MILP) dat niet is.

Resultaten

Als voorbeeld gebruiken we het model van het Noordzeekanaal waarbij het debiet bovenstrooms constant is, met een kleine toename van debiet rond 18:00. Voor de randvoorwaarde benedenstrooms gebruiken we data van het zeeniveau bij IJmuiden gedurende een dag. Door het veranderende zeeniveau zal het hoogteverschil van de sluis gedurende de dag verschillen. Wanneer het zeeniveau hoger is dan het kanaalpeil, zal de sluis dicht moeten.

In figuur 2 is het waterniveau aan beide kanten van de sluis te zien. Wanneer het waterniveau van de zee hoger is dan het kanaalpeil, mag de sluis niet open. Andersom mag de sluis wel open als het kanaalpeil hoger is dan het zeeniveau. Dit gebeurt rond 16:00 en 04:00, wat ook terug te zien is in figuur 3. Zowel rond 16:00 als rond 04:00 is er een piek te zien in het debiet van de sluis. Wanneer de sluis dus open mag, is het optimaal om dat te doen. Wanneer de sluis niet open mag, moet alsnog het waterniveau in stand gehouden worden. Dit wordt gedaan door de pomp te gebruiken. In figuur 3 is goed te zien dat sluisdebiet en pompdebiet elkaar afwisselen.

Figuur 4 laat nogmaals het kanaalpeil bij IJmuiden zien, maar dan met meer detail inclusief het minimum en maximum waterniveau. Te zien is dat het kanaalpeil gestaag stijgt zolang de sluis niet open kan. Zodra de sluis open mag, daalt het kanaalpeil weer. Hieruit blijkt dat de optimale strategie is om de sluisafvoer te maximaliseren, om hiermee de benodigde pompafvoer te minimaliseren.

Verdere ontwikkelingen

De belangrijkste verbetering van deze optimalisatie methode ten opzichte van bestaande simulatie methodes die ditzelfde probleem oplossen, is de mogelijkheid om het probleem op te lossen tijdens de optimalisatie routine zelf waarbij alle tijdstappen in de simulatieperiode worden meegenomen. In simulatie methodes kan pas bepaald worden wanneer de sluis open en dicht moet als het gesimuleerde peil van de vorige berekening afgerond is. Gebruik van optimalisatie technieken is een van de eerste stappen die nodig is voor optimaal beheer van de sluisen en pompen. De onderzochte lineaire methode kan de bestaande niet-lineaire stap vervangen in het proces, zodat er veel beter en efficiënter rekening gehouden kan worden met overige beheerdoelen zoals visvriendelijker pompen. Ook is het mogelijk om nog meer elementen van de optimalisatie, die voorheen door hoge rekentijden noodgedwongen achteraf werden berekend, te verplaatsen naar de optimalisatie zelf. Dit zou betekenen dat de minder praktische resultaten van de pomp in figuur 3 aanzienlijk verbeterd kunnen worden.

Literatuur

- S. Hayami. „On the propagation of flood waves”. In: *Bulletins-Disaster Prevention Research Institute, Kyoto University* 1 (1951), p. 1–16.
- V. Te Chow. *Open channel hydraulics*. 1959, p. 508.

Kieran van Gaalen is masterstudent computer science bij de Universiteit Utrecht. Voor zijn scriptie is hij bezig met optimalisatie van watersystemen bij Deltares.
E-mail: k.vangaalen@students.uu.nl

Klaudia Horváth werkt als onderzoeker en adviseur bij Deltares, afdeling Operatie Waterbeheer aan optimalisatie van watersystemen.
E-mail: klaudia.horvath@deltares.nl

Rens de Heer is promovendus bij de Universiteit Utrecht, waar hij onderzoek doet naar service locatie planning bij de NS.
E-mail: c.e.deheer@uu.nl

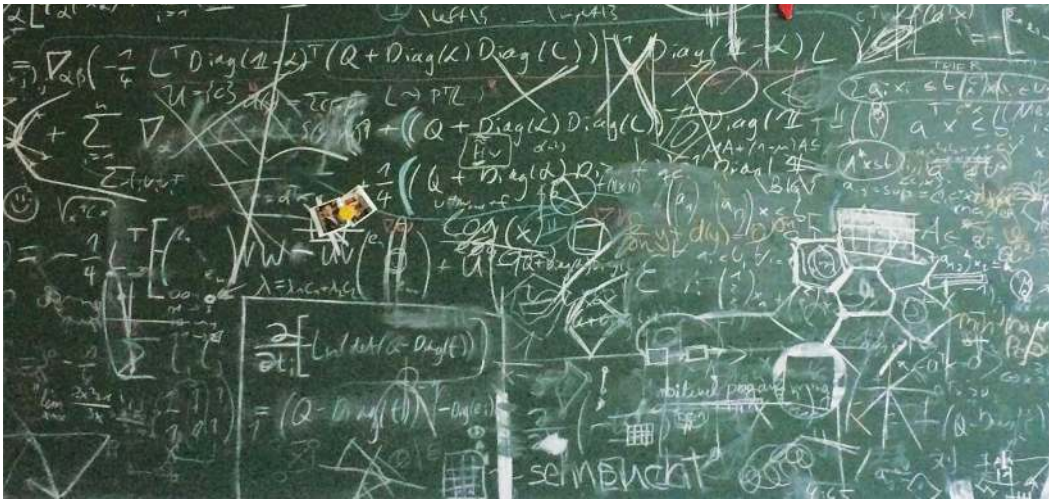


Photo by Jannis Kurtz

Explainable Decision Making: How to explain optimization models in Operations Research

Jannis Kurtz

Automated systems have become an integral part of our daily life. These systems, driven by advanced algorithms and technologies like artificial intelligence (AI) or Operations Research (OR), nowadays play a pivotal role in different domains ranging from transportation and healthcare to communication and disaster management. Due to its overarching presence the work and free time of most individuals is affected by decisions derived from optimization models. This situation gives rise to an important question: how do we explain and/or justify the decisions made by an optimization model? In this article we discuss why explanations are needed in OR and how explanations can be derived by calculating so-called counterfactual explanations.

Why do we need explanations?

One of the narratives frequently repeated in the field of Operations Research is that our goal as an OR expert is to model the “real world”. However, as most experts admit, the “real world” is too complex and we can only provide models which are an abstraction of it. But this real-world-narrative is misleading and even problematic since it gives a false authority for the decisions made by OR algorithms. While the goal in physics (or other natural sciences) is indeed modeling the natural laws of the real world, in OR we are doing nothing like this. Operations in a society are not like natural laws since they depend highly on the system and ideology society is embedded in. In fact, what we are doing instead is modeling the interests of a certain group of people in an organization and often do not coincide with the interests of other individuals who are part of the organization (e.g. workers or users) which was already debated in the 80s in Rosenhead and Thun-

hurst (1982). Hence, if your industry partner decides the solution of your model is not implementable, this does not mean that you failed to model the “real world”, it just means you failed to model the interests of this group of people.

What does this have to do with explanations? The real-world-narrative leads to the assumption that there is no need to explain or justify our optimization models, as long as these models (kind of) accurately model what “reality” demands from us. If we calculate solutions which are optimal for our model, there is no further discussion allowed. However, there are multiple examples where the choice of the optimization model had a severe effect on individuals. For example, in December 2006, Netherlands Railways introduced a completely new timetable by using sophisticated OR techniques (Kroon et al., 2009). The first versions of the schedule led to much social unrest for the employees. The schedules were considered as “boring,” since for each employee it was more or less the same every day. After these aspects were included in the approach, and the resulting schedule was better explained, the employees finally accepted the new schedule. A similar situation appeared when optimization methods were developed for Boston Public Schools to optimize the bus routes for picking up the students (Bertsimas et al., 2020). The calculated changes in the schools’ start times led to strong vocal opposition from some school communities that would have been affected negatively. These examples show that the decision maker has the responsibility to explain and justify the decisions derived by a mathematical model since these decisions are usually not comprehensible for non-expert individuals affected by them. This need for comprehensive explanations has been legally underpinned by the legislative initiatives such as the General Data Protection Regulation, the European Union AI Act, and the US Blueprint for an AI Bill of Rights.

Besides the previously mentioned motives, explanations can also provide a sophisticated tool for the OR expert, to get insights into the behavior of an optimization model and the properties of its decisions. While huge progress in this regard was made by a vast amount of works on Sensitivity Analysis, this methodology is limited in its capabilities to provide explanations due to the focus on local behaviors of the model. Understanding why a model chooses a certain solution and not a vastly different one, requires a more global analysis of the model and a different methodology is needed.

What is a counterfactual explanation?

A counterfactual explanation (CE) is a causal statement of the form “If A had not happened, B would not have happened.”. This concept is frequently applied in the realm of Explainable AI to provide explanations for the behaviour of black-box AI models. Consider for example a bank which trained a machine learning model (e.g. a decision tree) on the personal data of its customers to predict if a customer should get granted a loan or not. If this model denies a loan to a new customer, this decision has to be explained. A counterfactual explanation could be “If your annual salary would be 2 500 EUR higher, we would have granted you a loan”.

Translated to the optimization world, a CE is of the form “If the problem parameters (e.g. the constraint parameters) would be changed in this way, the optimal solution of the model would have the favored properties.”. To make the concept more clear let us consider the diet problem, a linear optimization model which is used in the food supply chain optimization for the United Nations’ World Food Programme (WFP) or to optimize the ingredients for cattle feed implemented in the mobile application Feed Calculator; see Figure 1.

The diet problem calculates which food commodities should be included in the food basket such that all daily nutritional requirements are satisfied, while the costs are minimized. An optimal solution of the diet problem applied to the data from the WFP is shown in Figure 2. This decision made by the WFP has a significant effect on the business of the local suppliers who are providing the food. Assume the WFP has for every type of food a fixed supplier who can offer the product in case the WFP wants to purchase it. Imagine one of the suppliers who is offering Lentils, Rice and Chickpeas wants to know why no product of her assortment is bought by the WFP. She can ask the counterfactual question “How much would the prices of the products have to change, such that the WFP includes at least 200 gram of my food into the daily food basket?”.

To answer the question the WFP has to determine, what is the smallest change in the price parameters of the diet problem, such that an optimal solution includes at least 200g of Lentils, Rice or Chickpeas. It turns out that providing an answer to this questions, i.e., calculating a CE, can be done by solving a bilinear optimization problem which we derive in our work (Kurtz, Birbil, and Den Hertog, 2024). The calculated counterfactual explanation for the question is “If your price for Lentils decreases



Figure 1: The mobile application Feed Calculator. Picture taken from <https://www.feedcalculator.org/theapp>.

by 50% and the price for Wheat increases by 25.2%, then we would buy at least 200 gram of your food.". The intuition behind the latter explanation is that the prize of Wheat has to increase such that it is not part of the optimal food basket any more but is replaced by the now cheaper Lentils.

Counterfactual explanation variants

When analyzing the concept of CEs for optimization problems it turns out that the situation is usually more tricky. The explanation the WFP is providing above only takes into account, that for the given price changes there *exists* an optimal solution which fulfills the properties the supplier is demanding. We call such an explanation a *weak counterfactual explanation*. However, since an optimization problem can have multiple optimal solutions this explanation does not guarantee that the WFP chooses the one which is favored by the supplier. To overcome this issue we developed the concept of so called *strong counterfactual explanations* which calculates a change in prices such that all optimal solutions of the corresponding problem are fulfilling the requirements of the supplier.

The models of the latter two concepts turn out to have very complicated mathematical structures, making it intractable to calculate the weak or strong CEs for large-dimensional problems at the current state of algorithmic and hardware development. Hence, we propose a third variant to provide explanations which we call *relative counterfactual explanations*. Applied to the diet problem above, in the relative CE model the WFP provides the price changes the supplier has to perform such that there is a solution, which buys at least 200g of the three products of the supplier, which has costs not larger than the current costs the WFP has. In this case the WFP could implement a solution fulfilling the demands of the supplier without increasing the costs. This model ignores the optimality condition used



Figure 2: Optimal food basket for the WFP diet problem.

in the weak and strong CE models which leads to a tractable model which can be used to calculate relative CEs in the same time as the underlying linear optimization problems, i.e., usually in seconds.

Outlook

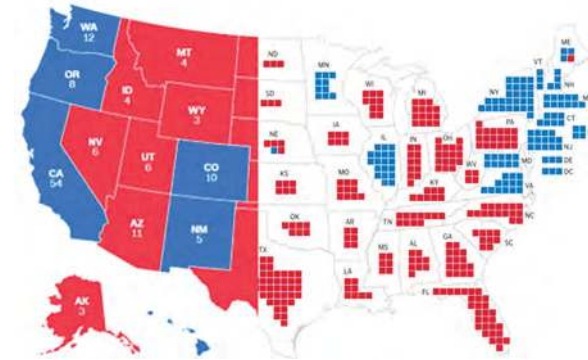
While our developed concepts can be applied to a variety of problems appearing in practice, our solution methods rely on using optimality conditions for linear optimization problems. Hence, they can only be applied to linear optimization problems with continuous decision variables. It would be important to develop efficient methods to calculate CEs also for problems involving integer decision variables or non-linear problem structures. Furthermore, future work should also take into account the expertise of psychologists or law experts to develop explanations which are useful for the human user and aligned with the law. In general an active and diverse research field is needed to discuss and analyze different viewpoints on this topic.

References

- D. Bertsimas et al. "Bus Routing Optimization Helps Boston Public Schools Design Better Policies." In: *INFORMS Journal on Applied Analytics* 50.1 (2020), pp. 37–49.
- L. Kroon et al. "The New Dutch Timetable: The OR Revolution." In: *Interfaces* 39.1 (2009), pp. 6–17.
- J. Kurtz, Ş. İ. Birbil, and D. Den Hertog. "Counterfactual Explanations for Linear Optimization." In: *arXiv preprint arXiv:2405.15431* (2024).
- J. Rosenhead and C. Thunhurst. "A materialist analysis of operational research." In: *Journal of the Operational Research Society* 33.2 (1982), pp. 111–122.

Jannis Kurtz is Assistant Professor at the Amsterdam Business School of the University of Amsterdam.
Email: j.kurtz@uva.nl

presidentsverkiezingen



Bron: eigen compositie

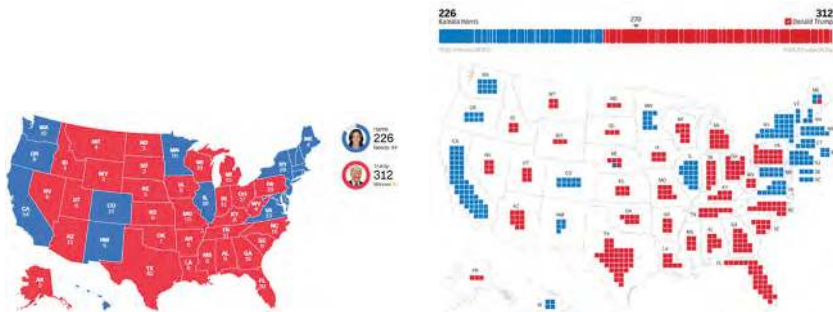
De Amerikaanse presidentsverkiezingen beter in beeld

Op 5 november 2024 waren er weer presidentsverkiezingen in de Verenigde Staten. De strijd ging tussen de Democraat Kamala Harris en de Republikein Donald Trump. En het was een spannende campagne. Afgaande op de peilingen was er sprake van een nek-aan-nek race. Maar uiteindelijk werd het een overwinning voor Trump.

Tijdens de verkiezingscampagne waren er veel peilingen. Voor het weergeven van de uitkomsten van die peilingen, en ook voor het in beeld brengen van de werkelijke uitslag, werd door de media heel veel gebruikgemaakt van thematische kaarten. Een thematische kaart is een speciaal soort grafiek waarmee je de verdeling van een variabele over een geografisch gebied kunt weergeven.

Thematische kaarten kunnen heel informatief zijn. Het in beeld brengen van de variabele moet je echter wel zorgvuldig doen, want het gevaar van een verkeerde interpretatie ligt op de loer. Een specifiek probleem is de vertekening die wordt veroorzaakt door het feit dat niet alle regio's in een geografisch gebied een even grote oppervlakte hebben. In het Engels duiden we dit probleem aan met *area bias*. Veel kaarten over de verkiezingen in de VS lijden aan *area bias*. Je moet dus voorzichtig zijn met dit soort kaarten.

De verkiezingen op 5 november 2024 werden dus gewonnen door Donald Trump. Hij kreeg 51% van de stemmen. Kamala Harris kreeg 48% van de stemmen. Dit wordt de *popular vote* genoemd. Het verschil tussen beide kandidaten was dus niet zo groot (3 procentpunten). Die uitslag zou je moeten kunnen terugvinden op de thematische kaarten. Maar was dat ook zo?



(a) Choropleet met de uitslag van de presidentsverkiezingen op 5 november 2024. Bron: CNN.

(b) Cartogram met de uitslag van de presidentsverkiezingen op 5 november 2024. Bron: The Associated Press.

Figuur 1: Vergelijking van de verkiezingsuitslagen in twee visualisaties.

In een *STATOR*-column van juli 2024 is al uitgebreid ingegaan op de verschillende manieren om thematische kaarten te maken. Al die manieren hebben hun voor- en nadelen, zie Bethlehem (2024). Hieronder passen we enkele van die manieren toe op de uitslag van de Amerikaanse verkiezingen van 25 november 2024.

Choropleten

Voor het weergeven van de resultaten van de Amerikaanse presidentsverkiezingen, en de bijbehorende peilingen, maken de media vaak gebruik van een speciaal soort thematische kaart, en dat is de *choropleet*. Figuur 1a bevat een voorbeeld van zo'n kaart. Deze kaart werd gemaakt door de Amerikaanse nieuwswijzer CNN (CNN, 2024).

Een choropleet is een kaart van een geografisch gebied dat is opgedeeld in een aantal regio's. De regio's zijn gekleurd, waarbij de kleur van elke regio de waarde van de te tonen variabele weergeeft. In figuur 1a zijn de regio's de staten van de Verenigde Staten. De Democratische staten zijn blauw en de Republikeinse staten rood. Bij nadere beschouwing van de grafiek valt het op dat de oppervlakte van de rode staten samen heel veel groter lijkt dan die van de blauwe staten. Op grond van deze kaart zou je kunnen concluderen dat Trump de verkiezingen met grote overmacht heeft gewonnen. Een *landslide* dus!

De Amerikaanse president wordt gekozen door het *Electoral College*. Dit college bestaat uit 538 kiesmannen. De winnaar van de verkiezingen is de kan-

didat met minstens 270 kiesmannen. Elke staat heeft een aantal kiesmannen. Het aantal per staat is gebaseerd op het aantal inwoners. Californië (CA) heeft de meeste kiesmannen (54). Verschillende kleine staten, zoals bijvoorbeeld Montana (MT), hebben slechts drie of vier kiesmannen. De kaart in figuur 1a bevat voor elke staat de afkorting van de naam (twee letters) en het aantal kiesmannen. De kaart van een staat geeft aan welke partij de staat heeft gewonnen: blauw voor de Democraten en rood voor de Republikeinen.

In 2024 won Donald Trump dus de verkiezingen. Hij kreeg 312 kiesmannen achter zich. Kamala Harris kreeg slechts 226 kiesmannen. Dus Trump had een meerderheid van 86 kiesmannen. De uitslag van deze verkiezingen is door CNN weergegeven in figuur 1a. Helaas is deze kaart misleidend. Hij lijdt aan *area bias*. Dit probleem ontstaat doordat de geografische oppervlaktes van de staten niet overeenkomen met de waarden van de variabele (hier: de aantallen kiesmannen).

Aangezien de verschillen tussen de Democraten en Republikeinen behoorlijk groot waren, zou je meer rood dan blauw verwachten. De kaart bevat echter wel heel veel rood en maar weinig blauw. Klopt de kaart wel? Als je de aantallen blauwe en rode pixels op de kaart in figuur 1a telt, dan blijkt slechts 28% van de pixels blauw te zijn en maar liefst 72% rood. Bij de kiesmannen liggen de verhoudingen heel anders. Van de kiesmannen gaat 58% naar Trump en slechts 42% naar Harris. Zie ook tabel 1. We kunnen dus concluderen dat de kaart een be-

	Rood (Trump)	Blauw (Harris)
Gekleurd oppervlak	72%	28%
Percentage kiesmannen	58%	42%
Popular vote	51%	48%

Tabel 1: Vergelijking van de uitkomsten voor Trump en Harris.

hoorlijk vertekend beeld geeft van de uitslag van de verkiezingen.

Als we de choropleet in figuur 1a nader bekijken, dan wordt duidelijk wat er mis is. We geven een voorbeeld. Kijk daarvoor eens naar de staat Montana (MT) in het noordwesten van de VS. De geografische omvang van Montana is behoorlijk groot, maar deze staat heeft toch slechts vier kiesmannen. De reden is dat er maar weinig mensen wonen in deze staat. Het is een dunbevolkt gebied. Kijk vervolgens eens naar de staat Pennsylvania (PA) in het noordoosten van het land. De geografische omvang (oppervlakte) van deze staat is veel kleiner dan die van Montana. Toch heeft de staat 19 kiesmannen, want het is een dichtbevolkt gebied. Verder is het hele blauwe noordoosten ondervetegenwoordigd. Er zijn hier verschillende democratische staten met een grote bevolkingsdichtheid en een kleine geografische omvang. We moeten wel tot de conclusie komen dat de kaart te lijden heeft van *area bias*.

Een cartogram

Je moet dus oppassen met het gebruik van choropleten. Het gevaar van *area bias* ligt altijd op de loer. Gelukkig zijn er nog andere manieren om thematische kaarten te maken. Ze gaan allemaal uit van het principe dat je de verdeling van de variabele in ieder geval altijd correct in beeld moet brengen. Dat mag eventueel ten koste gaan van de geografische nauwkeurigheid (de kaart is verwrongen). Zulke kaarten noemen we ook wel *cartogrammen*. Diverse voorbeelden van cartogrammen kun je vinden in, bijvoorbeeld, Bethlehem (2022).

We geven een voorbeeld van een cartogram. Hij stond op de website van het persbureau The Associated Press, Press, 2024. Deze grafiek is gereproduceerd in figuur 1b. De staten zijn niet gekleurd. In plaats daarvan wordt in elke staat een aantal 'tegeltjes' gezet. Het aantal tegeltjes correspondeert

met de waarde van de variabele (hier het aantal kiesmannen). Door deze aanpak is er relatief meer blauw te zien dan in de choropleet in figuur 1. Een staat, en dus ook de corresponderende waarde van de variabele, hangen nu niet meer af van de oppervlakte van de staat. Deze storende invloed is verdwenen. Daarom geeft figuur 1b een beter beeld van de verkiezingsuitslag.

Conclusie

De conclusie is dat je voorzichtig moet zijn met het gebruik van thematische kaarten, en in het bijzonder choropleten. Het gevaar bestaat dat je als gevolg van *area bias* een verkeerd beeld krijgt van de informatie in de grafiek, en dus een verkeerde conclusie trekt. Meer over thematische kaarten kun je bijvoorbeeld vinden in Bethlehem (2022) en Bethlehem (2024).

Dankwoord

Met dank aan Hayo Bethlehem die de berekeningen uitvoerde voor de bepaling van de percentages rode en blauwe oppervlakten.

Literatuur

- J. Bethlehem. „De Amerikaanse verkiezingen in kaart gebracht“. In: *Stator* 25.2 (2024), p. 13–17.
- J. Bethlehem. *Het Grafiekenboek*. Amsterdam University Press, 2022. URL: <https://www.aup.nl/en/book/9789463720984/het-grafiekenboek>.
- CNN. *2024 Electoral College Map*. 2024. URL: <https://edition.cnn.com/election/2024/electoral-college-map>.
- T. A. Press. *Election Results 2024*. 2024. URL: <https://apnews.com/projects/election-results-2024/>.

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