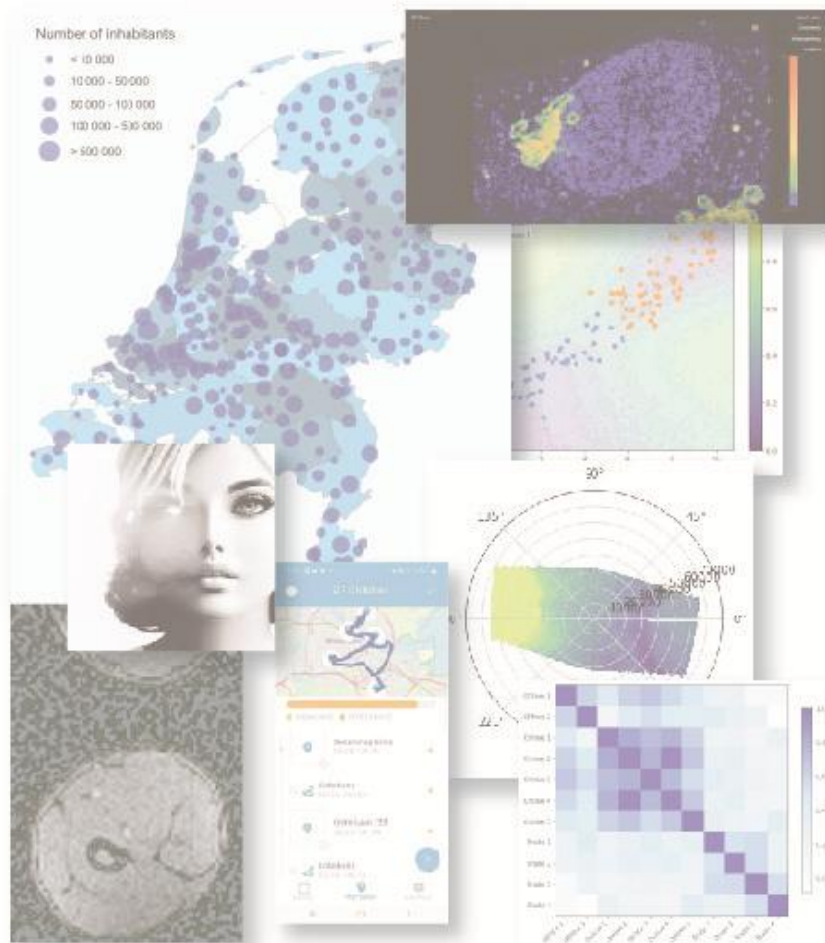


Studiegroep Wiskunde met de Industrie 2025

Publiekssamenvatting



Studiegroep Wiskunde met de Industrie 2025

Publiekssamenvatting

Colofon

De [Studiegroep Wiskunde met de Industrie 2025](#) vond plaats aan de Universiteit Utrecht, van 27 tot en met 31 januari 2025.

De [Studiegroep Wiskunde met de Industrie 2025](#) werd financieel gesteund door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), NWO-ENW en NWO-TTW, en de deelnemende bedrijven.

Deze [Public Proceedings](#) van de [Studiegroep Wiskunde met de Industrie 2025](#) werden geschreven en opgemaakt door wetenschapsschrijver Gieljan de Vries.

Voorwoord

De 183e Studiegroep Wiskunde met Industrie (SWI 2025, in het Engels, de *European Study Group Mathematics with Industry*) vond plaats van 27 tot en met 31 januari 2025 aan de Universiteit Utrecht in Nederland. SWI 2025 werd bijgewoond door 52 wiskundigen uit België, Denemarken, Engeland, Georgië, Duitsland, Italië, Noorwegen, Schotland, Servië en Nederland.

Problemen zijn ingebracht door zes organisaties: het Centraal Bureau voor de Statistiek (CBS), Jasco Applied Sciences, het RIVM, Opella, SVI (Scientific Volume Imaging) en het UMC Utrecht.

Volgens traditie werden de problemen maandagochtend gepresenteerd door de organisaties, waarna de wiskundigen zich in teams verdeelden om gedurende de rest van de week aan de vraagstukken te werken. Op vrijdag presenteerden de onderzoeksgroepen hun voorlopige bevindingen, waarvan de uiteindelijke uitwerkingen verschenen in de Scientific Proceedings.

In deze bundel kijkt de lezer in de keuken van de wiskunde. De problemen worden beschreven, de wiskundigen leggen uit hoe ze te werk zijn gegaan, en de bedrijven vertellen wat ze aan de oplossingen denken te hebben.

De Studiegroepen Wiskunde met Industrie in Nederland worden mogelijk gemaakt door de voortdurende en genereuze steun van NWO, in het bijzonder NWO-ENW en NWO-TTW, waar wij als organisatie zeer dankbaar zijn. Ook bedanken wij de deelnemende bedrijven voor hun financiële bijdragen.

Het organisatieteam van SWI 2024:

Martin Bootsma
Sjoerd Dirksen
Chiheb Ben Hammouda
Jason Frank (voorzitter)
Artem Kaznatcheev
Ivan Kryven
Cristian Spitoni

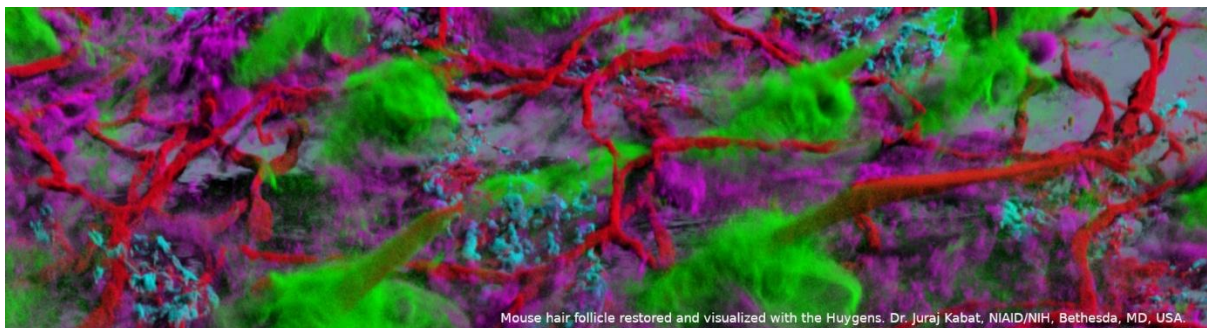
Inhoud

Voorwoord	4
Inhoud	5
Identiek licht uit een fluorescentiemicroscoop sorteren	6
Kaarten van de oceaانبodem koppelen	9
De computer die leert wat je vandaag deed	13
Woelen in de MRI.....	15
Corona uit de riooldata filteren.....	17
Het beste kanaal voor een advertentie	21

Identiek licht uit een fluorescentiemicroscop sorteren

Met fluorescentiemicroscopie breng je verschillende onderdelen van levende cellen in beeld, door ze aan eiwitten met verschillende kleurstoffen te laten binden. Het bedrijf Scientific Volume Imaging ontwikkelt de analysesoftware voor zulke microscoopbeelden en vraagt: kun je die eiwitten ook onderscheiden als ze bijna dezelfde kleur uitzenden? De deelnemers aan de Studiegroep Wiskunde met de Industrie vinden een verrassend krachtige oplossing.

Wil je als levenswetenschapper tot binnenin een levende cel kijken, dan is *fluorescentiemicroscopie* een uitgelezen techniek. Door fluorescerende 'labels' te koppelen aan eiwitten die zich aan specifieke celonderdelen binden, kun je precies zien waar bijvoorbeeld de celkern of -wand zit, of waar veel energie wordt verbruikt. Na belichting met de juiste kleur licht (*excitatie*) zenden de kleurstoffen namelijk fluorescerend licht uit met een kenmerkende golflengte.



Fluorescentiemicroscopie-opname van een muizenfollikel, gereconstrueerd en gevisualiseerd met het softwarepakket Huygens van SVI. Bron: Dr. Juraj Kabat, NIAID / NIH, Bethesda, MD, VS.

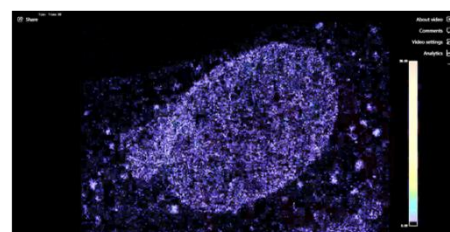
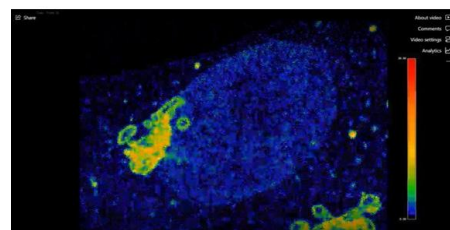
Onderscheiden

Het Nederlandse bedrijf *Scientific Volume Imaging* ontwikkelt software die data uit fluorescentiemicroscopen pixel voor pixel analyseert. "Als je meerdere celonderdelen tegelijkertijd onderzoekt, is het zaak om verschillende kleurstoffen te gebruiken. Anders zijn ze niet makkelijk te onderscheiden", legt wetenschappelijk ontwikkelaar Michel Ram van SVI uit.

Soms is de keuze van kleurstoffen echter beperkt, bijvoorbeeld door de bindingsmechanismen aan de biologische componenten. Om de verschillende eiwitten dan toch te kunnen onderscheiden, gebruikt SVI de levensduur van het fluorescentieproces. Ram: "Na excitatie duurt het typisch 1 tot 5 nanoseconden voor de kleurstof een lichtdeeltje uitzendt. Die vervaltijd verschilt per kleurstof."

Vervaltijden

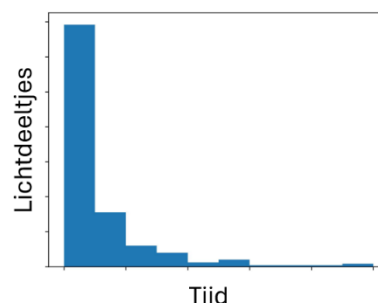
Een nieuw type sensor kan de aankomsttijd van lichtdeeltjes in de fluorescentiemicroscopie meten. Is daaruit de vervaltijd te bepalen om zo kleurstoffen te onderscheiden? Dat zou interessant zijn, want die vervaltijden zijn deels afhankelijk van de temperatuur of zuurtegraad. Afwijkingen van de standaard geven dus extra informatie over het binnenste van de cel. En als je die vervaltijden eenmaal weet, kun je dan per pixel bepalen wat de verhouding is van de twee kleurstoffen? Michel Ram en zijn collega Frans van der Have leggen het vraagstuk voor aan SVI.



Fluorescentie van getagde eiwitten op verschillende tijdstippen. Bron: SVI

De case van SVI lijkt op maat gesneden voor wiskunde-promovendus Koen Keijzer: aan de Universiteit Leiden onderzoekt hij namelijk de wiskunde van biologische systemen zoals de groei van bloedvaten. "Dit leek me een mooie manier om de andere kant van mijn dagelijkse samenwerking met biologen te bekijken, namelijk hoe de microscoop-opnames van mijn collega's worden omgezet in bruikbare data."

Simpelweg de aankomsttijd van lichtdeeltjes per pixel plotten en exponentieel vervallende functies door de datawolk trekken is niet genoeg, geeft SVI de wiskundigen alvast mee. Per pixel komen maar zo'n 10 lichtdeeltjes binnen. Door meet-onzekerheden zijn daarom altijd meerdere combinaties van vervaltijden te vinden die hetzelfde gedrag geven.

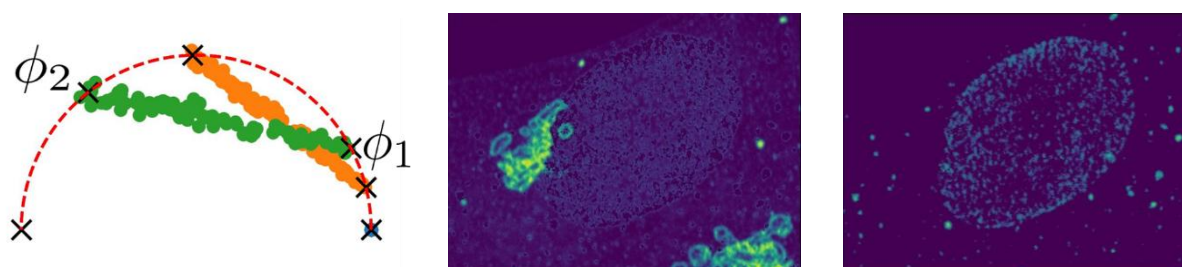


Gemeten lichtdeeltjes in de tijd.
Bron: SVI

Robuust

De wiskundigen bouwen daarom in een kleine week een tool die kunstmatige fluorescentiedata maakt om hun analyses op uit te voeren. Ze gebruiken eerst de *phasor*-analyse die SVI zelf ook heeft getest. Deze analyse onderscheidt trage en snelle veranderingen in de uitdovende kleurstoffen met een wiskundige *Fouriertransformatie*.

"De phasor-aanpak was heel populair in de recente literatuur, maar wiskundig was op voorhand niet duidelijk waarom dit de beste keus zou zijn", stelt Keijzer. Toch blijkt de methode onverwacht 'robuust'. Na wat analyse snaptten de onderzoekers waarom: "Deze aanpak is heel effectief in het onderdrukken van meetruis." Beter zelfs dan standaardmethoden uit de statistiek die de wiskundigen uitproberen, waaronder de 'method of moments'; een techniek om eigenschappen van een verzameling metingen te analyseren.



Van links naar rechts: phasor-berekening en reconstructies van twee verschillende kleurstoffen. Bron: SWI

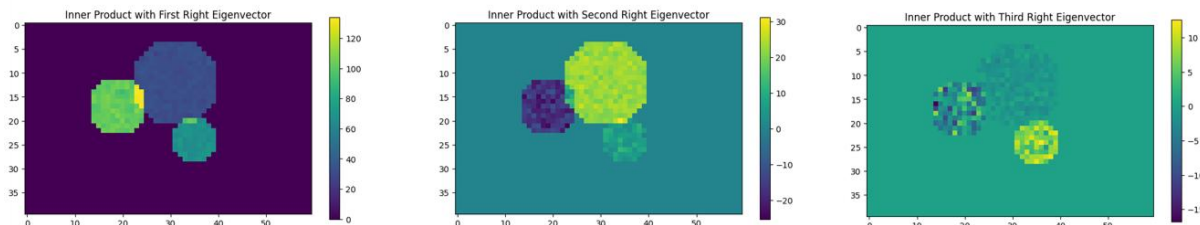
Onverwacht

Door op de ingeslagen weg dus, voor SVI? Niet helemaal. In een eindsprint werkt een van de wiskundigen een compleet nieuwe manier uit om de vervaltijden van fluorescerende eiwitten te bepalen.

De nieuwe aanpak werkt niet per afzonderlijke pixel, maar kijkt juist naar de relatie tussen naburige pixels. Die is te vangen in een tijdsafhankelijke matrix, en door die op de juiste manier uiteen te rafelen met *singular value decomposition* scoort het team een onverwacht resultaat: de analyse blijkt te kunnen voorspellen hoeveel verschillende kleurstoffen er in een meting aanwezig zijn. Die informatie moet je in andere methodes handmatig toevoegen.

"We waren onder de indruk dat de groep in zo'n korte tijd verschillende, vernieuwende oplossingsrichtingen heeft getest", kijkt Michel Ram terug. Zo is er ook een model opgezet dat het vervalgedrag van kleurstoffen voorspelt met exacte formules. Dat kost wel extra rekentijd, omdat je de uitkomst steeds verder verfijnt met herhaalde rekenstappen. Ram: "Dit *Discrete Time Multi-Rate Decay Model* is mogelijk te combineren met de *Maximum Likelihood*-methode die wij zelf ook al toepassen en verdient zeker een eervolle vermelding!"

Ook de singular value decomposition valt in de smaak: "Deze methode lijkt goed geschikt als screening vooraf, om te zien hoeveel verschillende stoffen er in de data zitten", denkt Michel Ram. "Op basis van deze eerste schatting kunnen we de phasor-methode vervolgens gericht uitvoeren. We zijn zeker van plan deze bevindingen verder uit te diepen."



Singular value decomposition-aanpak, die door analyse van naburige pixels het aantal verschillende kleurstoffen in een opname kan bepalen. Bron: SWI

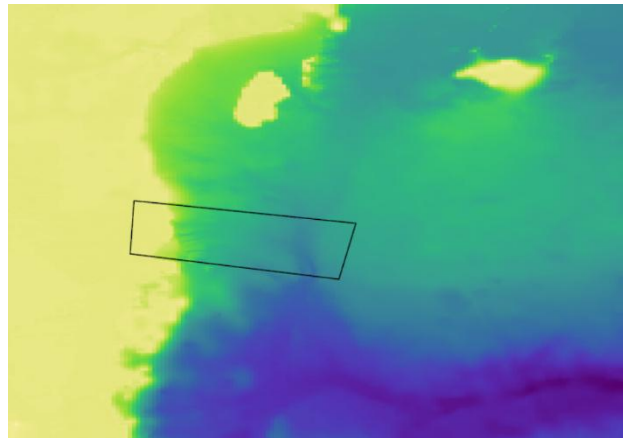
Bij de hoorns

Michel Ram is positief over de deelname aan SWI: "De deelnemers stelden doelgerichte vragen en vatten het probleem meteen bij de hoorns. We hebben ook genoten van hun bezoek aan ons bedrijf om te praten over het probleem en de mogelijke oplossingsrichtingen, ondanks de tijdsdruk voor dit project. Dat klikte zo goed dat we zijn blijven zitten voor een gezellige borrel."

Ook Koen Keijzers is tevreden met het resultaat. "Ik was geïnteresseerd in het soort problemen dat je als wiskundige tegenkomt buiten de universiteit en onderzoeksinstituten. Dat bleek erg leuk. Deze workshop was anders dan de universiteit, waar je soms op een eilandje alleen werkt. Hier kreeg ik veel energie van werken aan een gezamenlijk doel met een diverse club mensen. Nu terug naar het afronden van mijn proefschrift!"

In elkaar passen

"Jasco vroeg ons hoe je dieptekaarten kunt samenvoegen, waarbij ruwe kaarten worden verbeterd op basis van de fijne." Aan het woord is SWI-deelnemer Bas van 't Hof van het bedrijf Vortech, een terugkerend deelnemer aan SWI. Niet om cases in te brengen of opdrachten op te doen, vertelt hij: "Ik ben hier voor mijn eigen inspiratie. Bij SWI werk je een week lang samen met nieuwe mensen van uiteenlopende achtergronden. Echt een kans om te leren."



Een kleine en gedetailleerde kaart van de zeebodem ingepast in het gebied van een grotere, ruwere kaart van de omliggende regio. Bron: JASCO

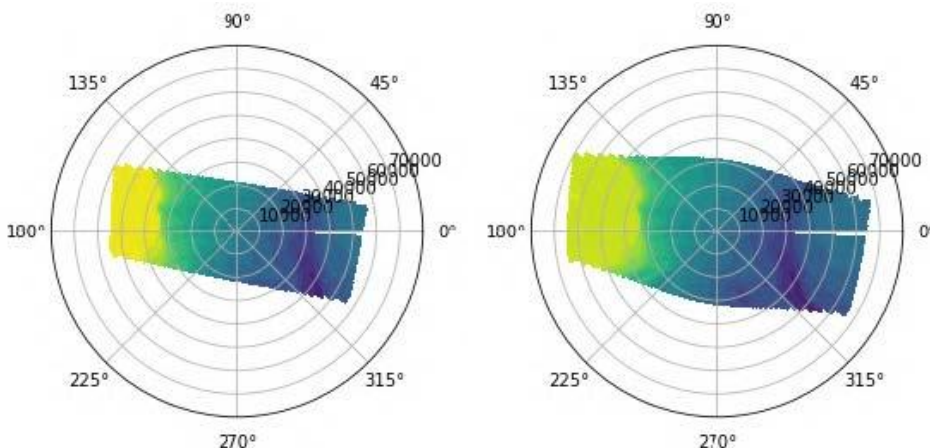
Van 't Hof en zijn medewiskundigen krijgen van JASCO bodemkaarten met botsende meetwaardes. Dat gaf de vraag: probeer je vanaf de randen van de hoge-resolutiekaart gladjes naar de omliggende lage-resolutie data te bewegen, of begin je al binnen het hoge-resolutiegebied met middelen?

"Het bedrijf had daar duidelijke voorkeuren in", herinnert Van 't Hof zich: "Wij opperden dat je bijvoorbeeld een deel van de hoge-resolutie data op zou kunnen offeren om met een zacht verloop te middelen tussen de twee kaarten. Maar dat zou waardevolle, betrouwbare datapunten kosten, wat voor JASCO geen optie was." Met die marsorders gaan de SWI-deelnemers in groepjes uiteen, om drie verschillende aanpakken uit te werken.

Cylindrische dronkenmanswandeling

De meerderheid van de onderzoekers duikt op een techniek die ze 'cylindrisch' noemen, vanwege de start vanuit een centraal punt. Vanuit het midden van de hoge-resolutiekaart worden telkens lijnen naar buiten getrokken. Langs zo'n lijn analyseren de wiskundigen de hoogtevariatie en proberen dat gedrag door te trekken tot in het lage-resolutiegebied.

Hoe dieper in de lage-resolutiekaart, hoe zwaarder die meetelt, zodat er uiteindelijk een vloeiende overgang is. De onderzoekers gebruiken een zogeheten *stochastische* aanpak; een dronkenmanswandeling waarbij de voorspelling nu eens omhoog, dan weer omlaag gaat, binnen de variatie die al is gezien in de hoge-resolutiekaart. Zo weten ze de kaarten te verweven. Toch blijven er oneffenheden achter, zoals een vingerafdruk van de rekenmethode: straalvormige patronen in de nieuwe kaart.



Originele fijnmazige kaart (links) en de uitbreiding (rechts) met de cilindrische methode. Bron: SWI

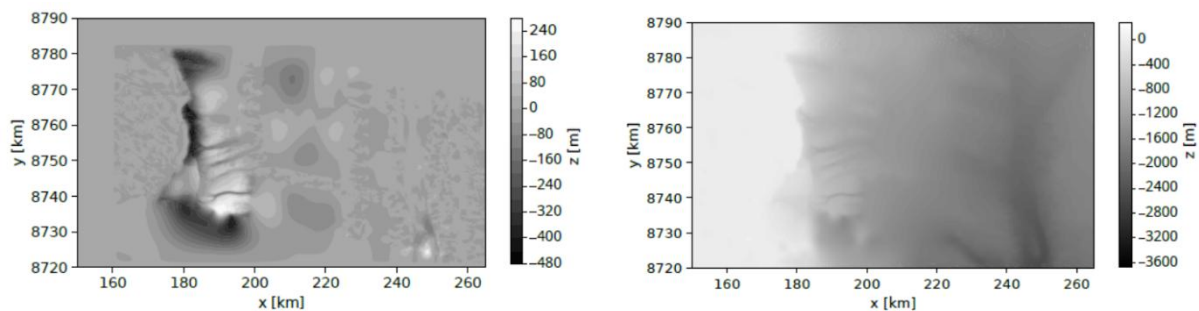
Biharmonisch lakens spannen

Bas van 't Hof werkte zelf aan een aanpak gebaseerd op differentiaalvergelijkingen. Die vermijdt scherpe randen tussen de twee kaarten gegarandeerd: "We zoeken een nette functie die begint op de rand van het fijne domein en glad naar de grofmazige data toebeweegt."

Alsof je een dun, elastisch laken met wasknijpers vastpint op alle randen van de fijnmazige kaart, neer laat dalen op het ruwe gebied, en dan de plooien en richels volgt. De vorm van zo'n laken wordt voorgeschreven door de *Laplace-vergelijking*, en geeft voor elke coördinaat in de ruwe kaart een nieuwe hoogtewaarde.

"Gebruik je niet de Laplace-vergelijking maar de *Biharmonische* of *Dubbele Laplace*, dan kun je naast de hoogtewaarde ook de helling van het terrein meenemen", vertelt Van 't Hof, die zo ook knikken tussen twee aansluitende gebieden kan voorkomen.

"Een veelbelovende aanpak, al is dit wel rekenintensief", denkt opdrachtgever Camillo Schenone. "De grote uitdaging wordt om dit uit te breiden naar kaarten die niet rechthoekig zijn of op een net grid liggen. Misschien een uitdaging voor volgend jaar?"

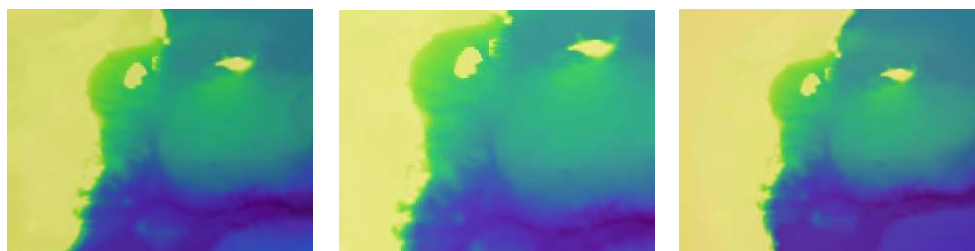


Verskil tussen de fijne en ruwe kaart (links) en de samengevoegde versie op basis van de biharmonische functie (rechts). Bron: SWI

AI: van fjord tot fotogallerij

De derde aanvliegroute van het JASCO-vraagstuk is een AI-aanpak. Die verwerkt hoogtekaarten in de vorm van afbeeldingen, in plaats van als hoogtewaardes op een coördinatenstelsel. De onderzoekers gebruiken zogeheten *Generative Adversarial Networks*, waarbij één AI een nieuwe afbeelding verzint en een andere AI beoordeelt of die door de beugel kan. In die touwtrekwedstrijd scherpen beide AI's hun voorspellingen aan.

Met de techniek van *uncropping* voert het team hun AI-tool de fijne en ruwe kaart, met de vraag om te voorspellen wat er buiten de randen van het fijnmazige gebied gebeurt. Alsof je Photoshop vraagt om het landschap achter de Mona Lisa door te tekenen buiten de randen van het schilderij.



Uncropping van een gedetailleerde zeebodemkaart, waarbij telkens 20% extra oppervlak wordt toegevoegd. Bron: SWI

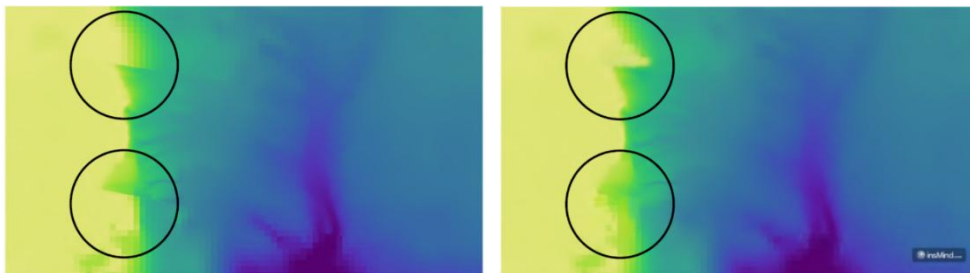
Uncropping blijkt overgangen tussen kaarten te verzachten, maar heeft ook zo zijn risico's, zoals de neiging om terug te vallen op het beeldmateriaal waarop de tool getraind is. Gezichten, zo blijkt als een zeebodemkaart op griezelige wijze overgaat in een vrouwengezicht. In het wetenschappelijke verslag staat droogjes: "Bij opeenvolgende stappen van telkens 20% groter waren de resultaten realistisch. (...) Te ambitieus *uncroppen* (200% in één keer) gaf een pathologische uitkomst". Waarvan akte.



*Te ambitieus gebruik van uncropping zorgt dat de AI-tool terugvalt op zijn trainingsdata; in dit geval portretten.
Bron: SWI*

Naast *uncropping* blijkt ook de *inpainting*-aanpak in staat om een fijne kaart in een ruwe te passen. Met uncropping kun je een moedervlek, krasjes, of andere beeldelementen aanpassen. De software vergelijkt omliggende structuren en kan zo harde overgangen verzachten. Ook daar is het nog zoeken naar de beste techniek. De originele kaarten zijn aangeleverd in kleur, maar de tool weet niet wat met elke kleur bedoeld wordt. Alles omzetten naar grijs tinten blijkt verwarring te voorkomen en nettere resultaten te geven.

"In deze tests hebben we publiek beschikbare tools gebruikt", legt Bas van 't Hof uit. "Dat ging snel, maar heeft zo zijn valkuilen, zoals de neiging om gezichten te tekenen. Als je een AI op maat traint op zeebodemkaarten, voorkom je dat probleem."



Overlap van ruwe en fijnmazige kaart, met in de twee cirkels de botsing tussen de twee datasets. Boven: originele data. Onder: Aangepast met infilling. Bron: JASCO / SWI

Geluidstest gezocht

Aan het eind van de SWI-week liggen er drie verschillende aanpakken om kaarten netter aan te laten sluiten, maar zijn de wiskundigen niet 100% zeker of ze ook de onderliggende vraag hebben opgelost. De echte test is namelijk niet of je kaarten kunt samenvoegen, maar of de simulaties van onderzees geluid niet hikken op de gladgestreken plooiën en bijvoorbeeld toch scherpe reflecties geven die er in de echte wereld niet zouden zijn.

"We hebben de deelnemers bewust niet gevraagd om de oplossingen te evalueren", reflecteert Schenone: "Een week tijd is simpelweg te weinig tijd om de hele machinerie van akoestisch modelleren door te werken. Maar ik begrijp de frustratie. Als wiskundige wil je weten hoe goed je resultaat is."

Zijn de deelnemers en opdrachtgever tevreden? Schenone, die tijdens zijn wiskundestudie al eens deelnam aan een vergelijkbaar programma van wiskunde en industrie, is vol lof. "Het team was ontzettend actief, elke keer dat we elkaar spraken hadden ze weer nieuwe en toegespitste vragen. Telkens bleek hoeveel werk ze hadden verzet en hoe ze met hun oplossingen onze vragen en randvoorwaarden probeerden te balanceren. Aan ons nu de uitdaging om hun technieken te testen in onze workflows."

Ook wat Van 't Hof betreft was de SWI-week zondermeer een succes: "Ik leer altijd enorm van de mensen waarmee ik samenwerk. En het is fantastisch om mezelf te testen tegen al die andere experts."

De computer die leert wat je vandaag deed

Informatie over de dagbesteding van Nederlanders is razend interessant voor het Centraal Bureau voor de Statistiek (CBS). Deelnemers krijgen daarbij hulp van een app die hun lokatiegegevens bijhoudt. Ze moeten die nog wel zelf 'labelen' met het doel van elk dagdeel. Kan een computermodel leren om die 'labels' te voorspellen?

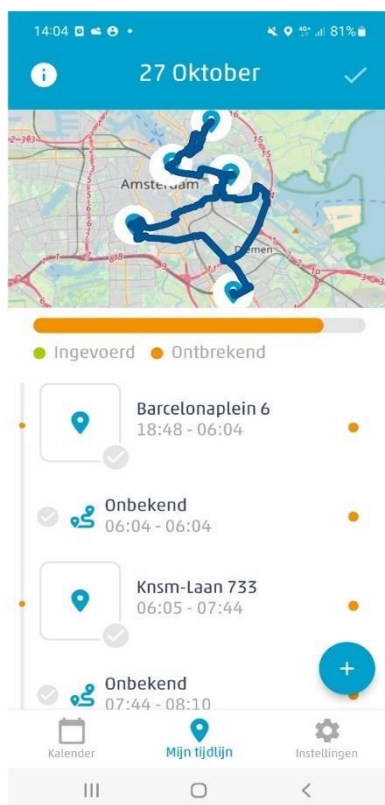
Hoeveel tijd besteedt de gemiddelde Nederlander aan werken, winkelen, onderwijs of sport? En hoe lang duurt eigenlijk het reizen tussen al die locaties? Het is interessante en nuttige informatie die inzicht geeft in hoe mensen hun leven organiseren.

CBS-methodoloog Barry Schouten en zijn collega's Chris Lam en Jonas Klingwort zoeken naar nieuwe manieren om de dagbesteding in kaart te brengen. Vroeger ging dat via dagboeken en interviews afnemen; foutgevoelig, want wie een paar dagen oversloeg, moest op het geheugen de afgelopen week reconstrueren. Handmatig registreren is ook tijdsintensief. Mensen kunnen daardoor de motivatie verliezen om mee te doen, of minder secuur worden.



Bron: Pexels / Andrea Piacquadio

Tijd dus voor digitale ondersteuning: een app, die via positiegegevens weet waar iemand zoal is, of via ingescande bonnetjes inzicht krijgt in het bestedingspatroon. De gebruiker hoeft dan alleen nog maar labels toe te voegen volgens een internationale standaard. 'Thuis', 'werk' en 'sporten' bijvoorbeeld voor locatiegegevens, of 'boodschappen' en 'ontspanning' voor uitgaves.



Schouten: "Dit zijn echte praktijkvoorbeelden. Onderzoek naar mobiliteit doen we bijvoorbeeld in opdracht van het Ministerie van Infrastructuur en Waterstaat. Die zijn geïnteresseerd in hoe infrastructuur wordt gebruikt. En voor de EU en het Ministerie van Economische Zaken verzamelen we via gescande bonnetjes informatie over uitgaves. Dat is belangrijk voor inflatiecijfers en voor inzicht in hoe huishoudens moeten en kunnen rondkomen."

Automatisch labelen

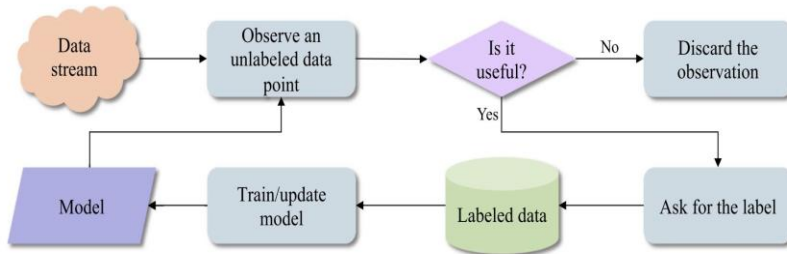
Ook met de app kost het een deelnemer redelijk wat tijd om bijvoorbeeld van elke reisbeweging te registreren wat het doel was. Kan een computermodel werk uit handen nemen? Met die vraag benaderen Schouten en collega's de Studiegroep Wiskunde met de Industrie (SWI).

De acht deelnemers aan de CBS-case zoomen al snel in op een aanpak die *active learning* heet. Relinde Jurrius, hoofddocent in combinatoriek en discrete wiskunde bij de Nederlandse Defensie Akademie: "Het idee is dat het computermodel alle nog ongelabelde locatiegegevens bekijkt en de gebruiker alleen vraagt om informatie te geven aan locaties waar dat echt wat toevoegt."

Reisbewegingen bijhouden in een app. Bron: CBS

Overeenkomsten

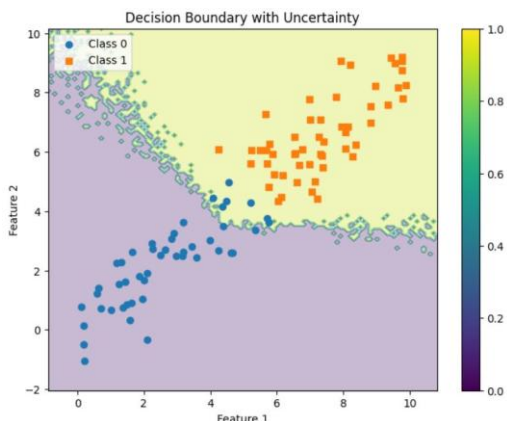
Hoe leert de computer wat wel en geen relevante informatie is? "Wij keken naar data over beweging en de reden van stops", legt Jurrius uit: "ga je met de auto of met de fiets, en stopte je voor school of voor boodschappen?" Een app kan meten wanneer je beweegt of niet, en kan ook wel leren dat die plaats waar je van 's avonds tot 's ochtends bent je huis is. "Maar het verschil tussen werk en school is een stuk lastiger, zeker als je verder niets van de gebruiker weet, dus dat wil je dan juist wel vragen. *Active learning* kan op een formele manier regelen hoe dit proces in zijn werk gaat."



Stappenplan om te bepalen of live binnenkomende data om annotatie vragen.

Bron: Cacciarelli, D., & Kuhlaci, M. (2024). Active learning for data streams: a survey. Machine Learning, 113(1), 185-239.

De wiskundigen trainen hun model door het steeds meer verplaatsingsgegevens te voeren. Die waren aangeleverd door het CBS. Twee verschillende algoritmes zoeken overeenkomsten en verschillen tussen datapunten en bepalen in welke categorie een tijdsbesteding thuishoort. Daarna testen ze hoe goed het model de rest van de data voorspelt.



Datapunten catagoriseren. Bron: SWI

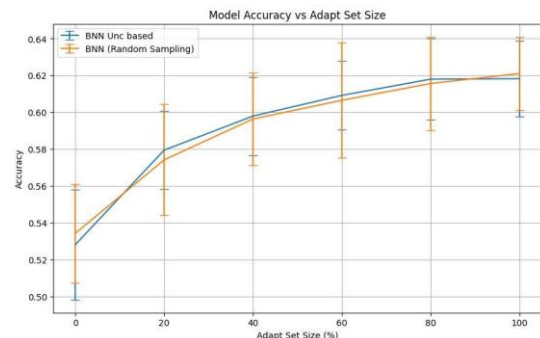
Opmerkelijk: dat lukt al goed na trainen op maar 30% van de data. Ook blijkt het mogelijk om al te gaan voorspellen voordat de volledige dataset van een hele dag binnen is — handig, want dan kan een app-gebruiker al direct na een activiteit de vraag krijgen om die te labelen.

Een computermodel dat steeds bijleert is onmisbaar, onderstreept Schouten: "In het dagelijks leven is niks bevroren. Winkels passen doorlopend hun assortiment aan en kiezen regelmatig nieuwe namen voor producten. Ook de infrastructuur verandert voortdurend."

Erachter komen wat er precies speelt

Terugkijkend op de SWI-week is opdrachtgever Schouten tevreden met het resultaat: "Een redelijk losse organisatie, precies de goede toon. En: een gedreven team. Ze hebben in korte tijd een rudimentair algoritme opgeleverd dat zeer de moeite waard is om te gaan simuleren, en recente theorie gevonden die nog niet in deze context is toegepast. Voor herhaling vatbaar."

Deelnemer Relinde Jurrius: "Ik wist al jaren dat SWI bestond, maar het paste telkens niet in de planning. Nu ik me meer met toegepaste wiskunde bezighoud, leek het me extra interessant eens deel te nemen. A priori was niet echt duidelijk wat precies het probleem was. Daar ligt volgens mij ook één van de grootste voordelen die wiskundigen hebben bij het aanpakken van toegepaste problemen: heel goed kunnen kijken en vragen om erachter te komen wat er nu precies speelt."



Nauwkeurigheid van voorspellingen bij toenemende trainingsdata. Bron: SWI

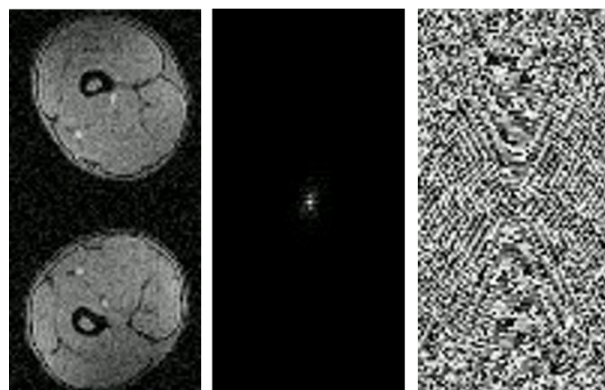
Woelen in de MRI

In de MRI-scan moet je stil liggen voor een zo scherp mogelijk beeld. Bewegende organen zoals longen, het hart en zelfs een langzaam samengedrukte beenspier zijn lastiger te onderzoeken. Zijn er slimme analysemethodes te vinden die toch bruikbare data opleveren over veranderingen en krachten in het lichaam?

MRI-scans zijn gemeengoed in de medische wereld, maar de techniek is nog lang niet uitontwikkeld. Zo zoekt het Universitair Medisch Centrum Utrecht (UMCU) naar manieren om naast stilstaande plaatjes ook bewegend weefsel zoals een kloppend hart te onderzoeken, of een orgaan dat zachtjes vervormt tijdens het liggen. "Dan zouden we allerlei informatie krijgen over trek- en duwkrachten in het lichaam", vertelt UMCU-onderzoeker Ray Sheombarsing.

De uitdaging: de meeste wiskundige technieken om beweging te analyseren vergelijken ruimtelijke plaatjes van verschillende momenten, maar de ruwe "*k-space*" data die een MRI meet is in de vorm van frequenties, waar die technieken niet voor werken. Pas na een wiskundige transformatie ontstaat de bekende dwarsdoorsnede van het lichaam.

Verschillende snapshots nemen van bewegende lichaamsdelen zou veel extra meettijd kosten, legt Sheombarsing uit. "Een andere oplossing is om direct met de *k-space* data te werken; en met minder, en dus incomplete metingen zodat de acquisitietijden aanzienlijk verkort kunnen worden. De methode die we nu gebruiken vraagt veel rekenkracht, dus we zijn benieuwd of er alternatieven zijn." Met die vraag benaderen Sheombarsing en zijn collega Alessandro Sbrizzi de Studiegroep Wiskunde met de Industrie (SWI).



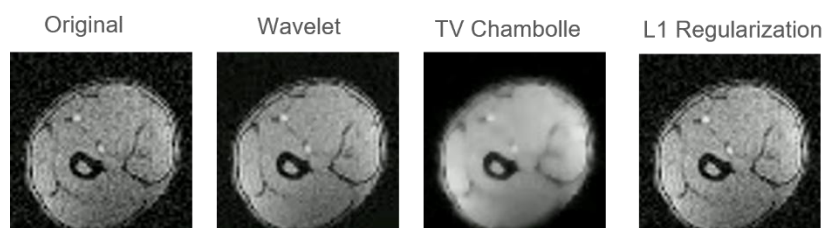
Video data Magnitude kspace Phase kspace

MRI-beeld (links) en k-space data. Bron: UMCU

Meetruiis

Wiskunde-promovenda Jolien Kamphuis van de Universiteit Leiden is pittige vraagstukken op het gebied van dynamica en partiële differentiaalvergelijkingen wel gewend. Met haar team ontdekt ze al snel dat het vraagstuk veel verschillende onderdelen heeft. "We zijn daarom opgesplitst in subgroepen met allemaal een eigen onderwerp en aanpak."

Eén groep richt zich op een bekend dataprobleem: meetruiis. Bij het zoeken naar verschillen tussen twee opnames wil je namelijk zeker weten dat een pixel van waarde verandert omdat het lichaam bewoog en niet door ruis in de opname. Kamphuis: "We hebben daar drie verschillende technieken gebruikt, waarvan er één, 'L1 regularisatie', het beste presteerde." Om goed te kunnen vergelijken, passen de onderzoekers in het vervolg toch de huis-aanpak van het UMCU toe; een analyse op basis van een neurale netwerk.

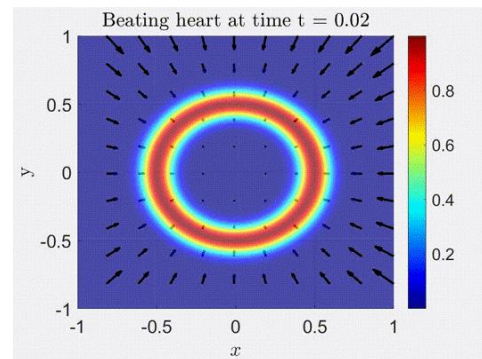


Verschillende technieken om uit kspace data dwarsdoorsneden van het lichaam af te leiden. Bron: SWI

Transformatie

Naast het zo zuiver mogelijk krijgen van de meetdata willen de wiskundigen ook beter begrip krijgen over de tijdsevolutie van de ruwe meetdata in k-space. Zo proberen ze de beweging van gescande lichaamsdelen direct uit de ruwe k-space data te halen.

Het team stelt verschillende aanpakken voor. Een daarvan is een model voor *optical flow*: formules die beschrijven hoe de pixels van een voorwerp over het beeld schuiven als het voorwerp langzaam van vorm of plaats verandert. "De manier waarop het team uitwerkt hoe je verschillende rekenmodellen kunt vergelijken is heel inzichtelijk", reageert opdrachtgever Sheombarsing.

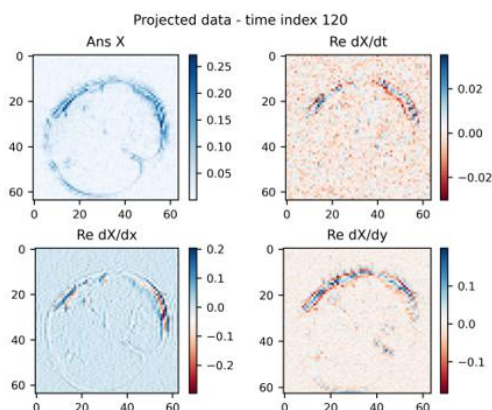


Numeriek model van een kloppend hart. Bron: SWI

Kloppend hart en samengedrukt been

Naast analytisch werk, het geduldig afleiden van relaties tussen meting en afbeelding door formules te manipuleren, zien de SWI-deelnemers ook een andere aanpak: een versimpeld model van een bewegend hart maken in de vorm van een numerieke benadering. Door dat kunsthart door de wiskundige transformaties van k-space naar ruimtelijke afbeelding te halen, willen ze inzicht krijgen in de invloed van ruis op beweging.

Om te toetsen of ze ook met echte MRI-data kunnen werken, analyseren de wiskundigen tenslotte de beweging van een been dat wordt samengedrukt door een bloeddrukmeter. In de MRI-data die het UMCU aanlevert, zoeken ze de kenmerkende tekenen van beweging in ruimte en tijd met een zogeheten principal component analysis. Die aanpak stuit op alle al bekende problemen: van het belang



Reconstructie van een been. Bron: SWI

van meetruis filteren, tot het probleem dat de data niet compleet genoeg zijn om de beweging perfect te reconstrueren. De analyse blijkt onderdelen van het kunsthart die wel in het originele model zitten, niet te kunnen reconstrueren.

Opdrachtgever Sheombarsing is desondanks tevreden: "De wiskundigen hebben heel interessante resultaten laten zien. Ook de discussies waren heerlijk intensief en leerzaam: ze dwongen ons echt om onze problematiek goed te definiëren. Bovendien kregen we door dat we onderschat hadden hoe goed onze eigen methode is in meetruis wegfilteren. Dat is absoluut niet triviaal, zoals de deelnemers hebben ontdekt. Wij hebben nu het vertrouwen dat we geen methode hebben gemist in de vakliteratuur en dat het de moeite waard is om onze eigen aanpak verder te ontwikkelen."

Eerste aanzetten

Aan het eind van de SWI-week ligt er kortom geen kant en klare oplossing voor het vraagstuk van beweging in de MRI-scanner. Jolien Kamphuis: "We hadden veel ideeën, maar 2,5 dag aan een probleem werken is uiteindelijk best kort! Toch hebben we leuke eerste resultaten gevonden die verder uitgewerkt kunnen worden."

Kamphuis: "Ik kan meedoen met SWI zeker aanraden, vooral als er een bedrijf tussen zit wat je interessant lijkt. Ik vond de samenwerkingen erg prettig en het was leuk om een inkijkje te krijgen in de problemen van de bedrijven die meededen. Bij zulke praktische problemen werkt je ideale wiskundige aanpak of favoriete aanname niet zondermeer, je hebt altijd praktische beperkingen. De kunst is dan om de essentie van het vraagstuk zo te formuleren dat je er wél iets mee kan. Het was genieten om te zien hoe mensen de verschillende onderwerpen hebben aangepakt."

Corona uit de riooldata filteren

De coronapandemie is voorbij, maar het Nederlandse RIVM houdt nog steeds het aantal deeltjes coronavirus in ons rioolwater bij. Door natuurlijke variaties pieken die waarden soms. Kan de Studiegroep Wiskunde met de Industrie bepalen of dat toevallige meetafwijking zijn, of het teken van een nieuwe uitbraak?

Sinds 5 mei 2023 beschouwt de wereldgezondheidsorganisatie WHO COVID-19 niet meer als pandemie. Toch worden er nog steeds mensen geïnfecteerd, soms met de jarenlang slepende aandoening *long covid* als gevolg. Ook nieuwe varianten van het virus zijn de moeite om in de gaten te houden. Daarom houdt het Nederlandse RIVM (Rijksinstituut voor Volksgezondheid en Milieu) de vinger aan de pols van corona.

"Een van de manieren waarop je de verspreiding van het coronavirus kunt meten is via het aantal virusdeeltjes in het rioolwater", vertelt Wouter Hetebrij. Na zijn promotie in de wiskunde werd hij modelleur bij het RIVM. "Even omschakelen; je moet toch een hoop achtergrond opdoen. Maar daarna kon ik mijn vaardigheden prima toepassen."

Bijvoorbeeld op de analyse van coronavirusdeeltjes in het rioolwater. Op diezelfde manier wordt al gemonitord op de ziekte polio en naar bacteriën die ongevoelig zijn voor antibiotica. Het levert een interessant vraagstuk op voor de Studiegroep Wiskunde met de Industrie (SWI), waarvan op voorhand niet duidelijk is welke techniek de beste oplossing zal geven.

Metingen teruggeschroefd

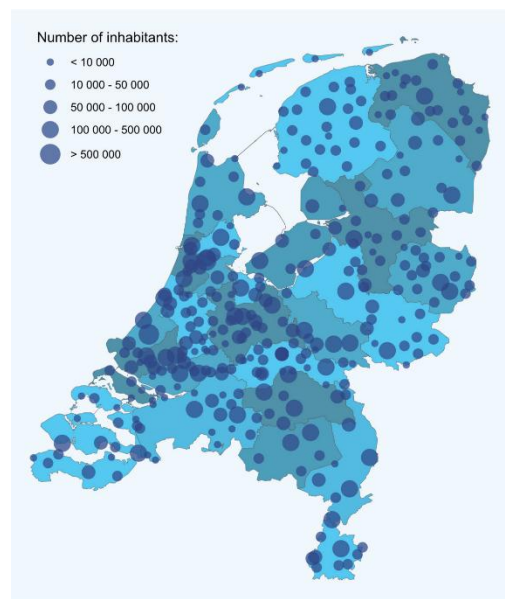
Op het hoogtepunt van de coronapandemie namen de Nederlandse waterschappen vier keer per week monsters bij de ruim 300 rioolwaterzuiveringsinstallaties (RWZI's) in ons land. Het RIVM analyseerde die monsters om per RWZI-gebied het aantal virusdeeltjes per 100.000 inwoners te achterhalen. Steeg het coronagehalte in het rioolwater ineens, dan was dat een aanwijzing voor een virusuitbraak.

Twee jaar na het eind van de pandemie is het aantal rioolwatermetingen teruggeschroefd. Niet alleen is de hoge nood voorbij, ook de kosten van het meten en analyseren tellen mee. Veertig procent van de RWZI's analyseert twee keer per week op SARS-CoV-2, de rest nog maar wekelijks.

Betrouwbaarheid

Met nog maar 1 à 2 metingen per week groeit de invloed van meetonzekerheden. Hetebrij: "Je kunt altijd een meting krijgen die eruit springt, waarbij het aantal virusdeeltjes ineens piekt of duikt." Daarbij is het lastig dat er geen data van eergisteren meer is om mee te vergelijken. Om te zorgen dat er geen onrealistische gegevens in officiële rapportages komen, doet het RIVM aan trendanalyse. Daarbij worden nieuwe metingen vergeleken met voorgaande waarden van dezelfde RWZI; een flinke rekenklus die enkele uren kost.

Kan dat sneller en betrouwbaarder, door cijfers van nabijgelegen RWZI's met elkaar te vergelijken? Voor naburige grote steden zou je vergelijkbaar gedrag verwachten. Hetebrij, die tijdens zijn eigen



Nederlandse rioolwaterzuiveringsinstallaties in 2022. Bron: van Boven, Hetebrij, et al. Euro Surveill (2023)

promotie al eens deelnam aan de Studiegroep Wiskunde met de Industrie (SWI), legt het vraagstuk voor aan de organisatie.

Populaties

Voor Niek Mooij van de Universiteit Utrecht is de RIVM-case een natuurlijke keuze. Zijn promotie gaat over hoe de vorm van netwerken met meerdere locaties complexe processen zoals ecologie of economie beïnvloeden. Daar komen veel differentiaalvergelijkingen bij kijken. "Die techniek leek ons ook kansrijk om de coronadeeltjes in het rioolwater te beschrijven."

In differentiaalvergelijkingen hangt de snelheid waarmee een meetwaarde verandert af van zijn huidige waarde. Zo kun je bijvoorbeeld de populaties van roof- en prooidieren modelleren, maar ook de verspreiding van virussen. Tijdens de SWI-week bouwt een deel van de wiskundigen een model van het aantal coronainfecties in het gebied rond een RWZI.

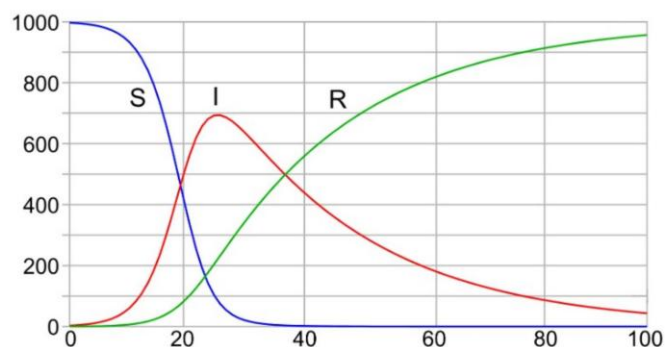


Het rekenmodel SIR. Bron: SWI

Het zogeheten SIR-rekenmodel (voor *Susceptible*, *Infectious* en *Recovered*) simuleert drie verschillende groepen mensen: mensen die nog *vatbaar* zijn voor infectie, mensen met de ziekte die *besmettelijk* zijn en dus virusdeeltjes verspreiden, en zij die *hersteld* zijn. Elke tijdstap van het rekenmodel kunnen mensen Corona oplopen of daarvan genezen.

Hoeveel mensen in het model besmet raken of genezen, hangt telkens af van de grootte van hun groep, vermenigvuldigd met een voorfactor: veel vatbaren betekent veel besmettingen. De uitdaging bij het SIR-model is om de juiste voorfactoren te vinden die de werkelijkheid goed beschrijven. Dat doet het team door hun formules te vergelijken met historische coronadata.

Met een betrouwbaarheidsinterval is te bepalen of een nieuw datapunt binnen de normale spreiding rond de functie ligt, of onwaarschijnlijk is. De SWI-week blijkt te kort om die analyse in te bouwen; ook is er geen tijd om de SIR-modellen van naburige steden met elkaar te vergelijken.



Een mogelijke oplossing van het SIR-model, met horizontaal de tijd en verticaal de grootte van de drie populaties. Bron: SWI

Voor Wouter Hetebrij is het SIR-model interessant, maar nog geen directe oplossing voor zijn vraagstuk. En ook Niek Mooij kiest uiteindelijk voor een andere aanpak: "Met differentiaalvergelijkingen moet je heel zeker zijn dat je het juiste gedrag modelleert. In het geval van corona heb je veel factoren die je hier niet mee kunt nemen." Herbesmettingen en nieuwe varianten bijvoorbeeld, maar ook vaccinaties. Mooij sluit zich daarom aan bij een team dat zich baseert op methodes uit de kansrekening.

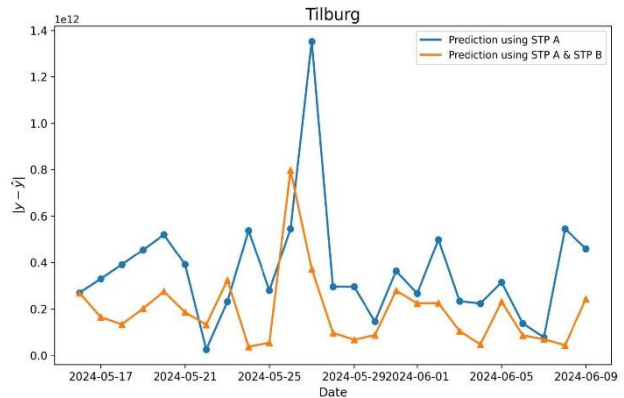
Autoregressie

In plaats van de werkelijkheid na te bootsen in een model, kiezen de statistici ervoor om wiskundige patronen af te leiden uit de werkelijkheid. Ze gebruiken de techniek van *autoregressie*, die populair is

om patronen te vinden in tijdreeksen van meetdata. Ook de relatie tussen verschillende rioolwaterzuiveringsinstallaties die het RIVM meet, is er goed in te vangen.

"Bij autoregressie bouw je een functie op basis van de historische data", legt Mooij uit. "Als je honderd datapunten, hebt, voer je een computer de eerste vijf en vraag je om een functie die het zesde datapunt voorspelt. Die verfijn je door met datapunt twee tot en met zes de zevende waarde te voorspellen. En zo door tot je bij het heden bent."

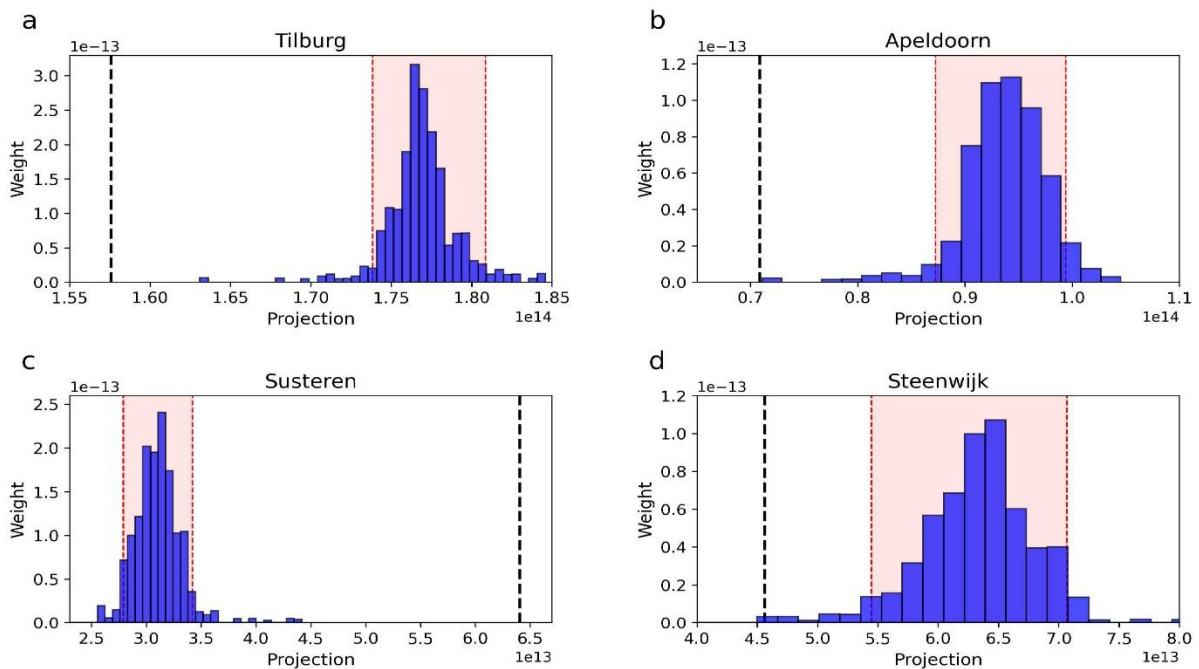
De uitkomst is serie coëfficiënten: per tijdverschil een coëfficiënt die het precieze gedrag bepaalt. En een bijbehorende onzekerheid in de voorspelling, want de data bevatten immers meetruis. De cruciale truc, vertelt Mooij: "vervolgens doe je de hele exercitie nog een keer met de data van een naburige RWZI erbij. Als dat een betere voorspelling levert, weten we hoe sterk die twee verbonden zijn."



Autoregressie van de Coronawaardes in het rioolwater van Tilburg. In de oranje uitkomst is ook een naburige RWZI meegenomen, in de blauwe alleen Tilburg. Bron: SWI

Door alle 300 ruim RWZI's met elkaar te vergelijken, krijgen de wiskundigen een netwerk van sterk en minder sterk op elkaar lijkende RWZI's. Na het doorploegen van de historische data gaat de berekening snel: elke nieuwe week hoef je per RWZI maar één a twee nieuwe rekenstappen te doen. Rekening: anderhalve minuut.

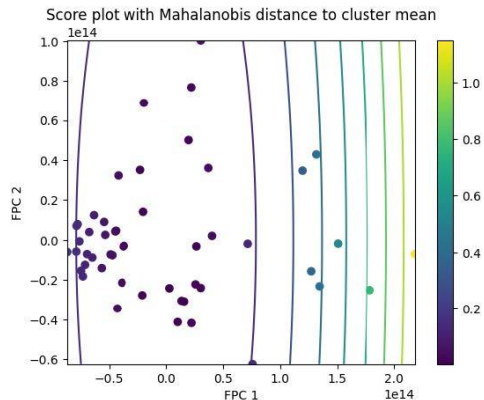
"Valt een nieuwe meetwaarde dan buiten de bulk van wat het netwerk zou verwachten, dan weet je dat die écht onverwacht is", schetst Mooij. Omgekeerd: als meerdere sterk verbonden gebieden tegelijkertijd pieken, is de uitkomst betrouwbaarder.



Betrouwbaarheidsintervallen in de autoregressieaanpak. Bron: SWI

Onderliggende patronen

Ook een derde groep onderzoekers zoekt naar onderliggende patronen in de coronadata. Maar waar autoregressie per RWZI één beschrijvende functie oplevert, maakt *Functional Principal Component Analysis* (FPCA) er zoveel je maar wil. Het achterliggende idee is om de gemeten data te ontbinden in een serie basisfuncties die samen het waargenomen gedrag voorspellen.



De meeste coronawaardes bij de RWZI Holten blijken te voorspellen met vergelijkbare combinaties van FPCA-functies. Bron: SWI

Alsof je losse tonen bij elkaar optelt tot je de juiste klank krijgt: de mix van de componenten is allesbepalend en de precieze verhoudingen werken zelfs als een vingerafdruk. Als een nieuw meetpunt een heel andere mix vraagt om goed te voorspellen, is dat weer een teken dat de meting buiten de verwachtingswaarde valt.

Ook voor de FPCA-team blijkt een week te kort om het RIVM-vraagstuk volledig op te lossen. Hun data-analyse werkt, maar naburige meetlocaties vergelijken is voor de toekomst. Dat kan bijvoorbeeld door grafieken van rioolwaterzuiveringspunten die op verschillende momenten beginnen te pieken te verschuiven tot ze zoveel mogelijk over elkaar vallen.

De kracht van wiskunde

Ook al ligt er geen kant-en-klare oplossing voor zijn vraagstuk, opdrachtgever Wouter Hetebrij is enthousiast over de opbrengst van SWI: "Het is mooi hoeveel verschillende methodes er zijn geprobeerd. Toch de kracht van wiskunde: je hebt uiteenlopende gereedschappen zoals modelleren of statistiek die allemaal toepasbaar zijn. Vooral de autoregressie lijkt goed bruikbaar. We gaan nu bespreken welke van de drie methodes van SWI we een tijd uittesten naast onze standaardaanpak, zodat we ze over een tijdje kunnen vergelijken."

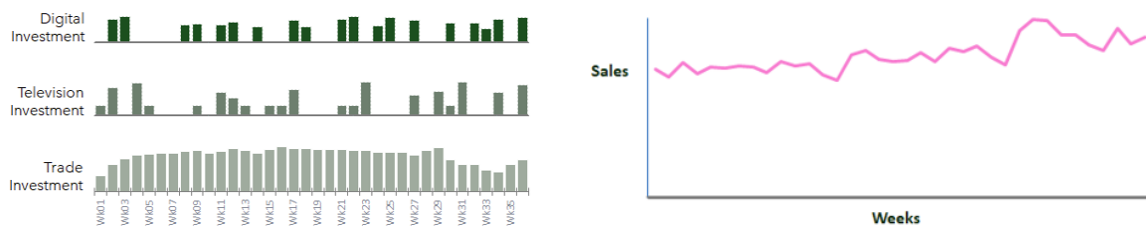
Ook promovendus Niek Mooij is tevreden over de SWI-week: "In een team werken aan een probleem uit de praktijk was heel interessant, erg anders dan ieder aan zijn eigen project. Het was extra leuk om samen te werken met mensen uit andere gebieden van de wiskunde, die je normaal gesproken niet 1-2-3 tegenkomt. Samen konden we heel wat meters maken!"

Opella

Het beste kanaal voor een advertentie

Opella is een van 's werelds grootste geneeskunde-firma's en maakt onder andere zelfzorgmiddelen zoals voedingssupplementen. Aan de Studiegroep Wiskunde met de Industrie (SWI) stellen ze de vraag: kunnen we met statistiek achterhalen wat het effect is van advertenciacampagnes via verschillende kanalen op onze verkoop, ook als die tegelijkertijd lopen met een vergelijkbare investering?

Een beetje marketingafdeling werkt met advertenties op verschillende sociale media, in magazines en kranten, op radio en tv. En elke marketeer en communicator kent de vraag: welk van onze kanalen presteert het beste? Die vraag legt Opella voor aan de deelnemers van SWI 2025 in Utrecht.

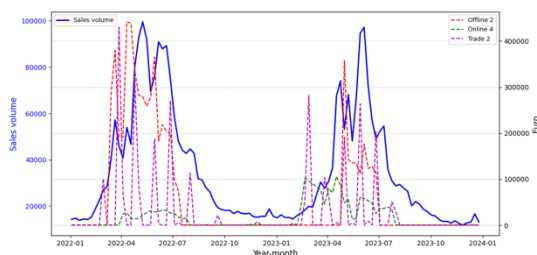


Investerings in verschillende mediakanalen en het bijbehorende salesvolume. Bron: SWI / Opella

In de pas

Om de relatie tussen advertentie-investeringen en uiteindelijke sales te achterhalen, gebruikt Opella Marketing Mix Modeling, een onderwerp binnen de econometrie. Een belangrijk ingrediënt daarvan is statistisch modelleren, om de waarschijnlijke relaties tussen investeringen en sales bloot te leggen. En dat is een uitdaging, legt het bedrijf de SWI-deelnemers uit.

In theorie is het vraagstuk namelijk op te lossen: varieer per reclamecampagne telkens de investering in één kanaal en kijk naar het effect op de sales. In de praktijk is dat niet te doen. Niet alleen gebruiken marketeers vrijwel altijd een mix van kanalen, de investeringen daarin lopen ook in de pas: rond de feestdagen krijgen bijvoorbeeld alle kanalen een extra impuls.



Multicollineariteit tussen verschillende mediakanalen. Bron: Opella / SWI

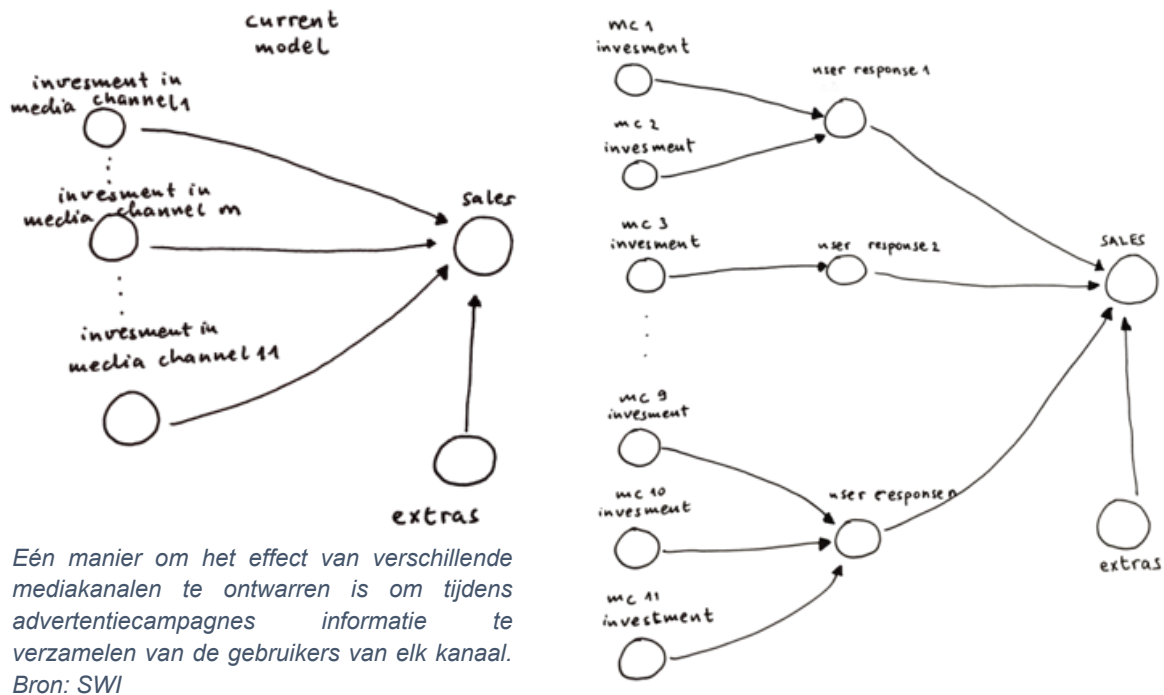
Inschatting vooraf

Promovendus Tijn Jacobs is een van de wiskundigen die de tanden in de Opella-case zet. Aan de Vrije Universiteit Amsterdam specialiseert hij zich in de Bayesiaanse statistiek die ook Opella gebruikt in zijn analyses. Anders dan bij frequentistische statistiek stellen Bayesianen vooraf een uitdrukking op van waarschijnlijke parameters, om die zogeheten prior daarna te confronteren met de data.

"De kern van dit probleem is wat we multi-collineariteit noemen", legt Jacobs uit. "Je probeert dan het individuele gedrag van een reeks parameters te voorspellen die lastig van elkaar te onderscheiden zijn." Het vraagstuk moet dan worden versimpeld, óf er is extra informatie nodig.

Een van de adviezen die de wiskundigen geven is dan ook om tijdens een marketingcampagne onderzoek te doen bij de mensen die een reclameuiting onder ogen krijgen. Die tussenstap kan extra informatie geven over hun koopbereidheid, of over hoe snel een advertentie tot verkoop leidt; alles om

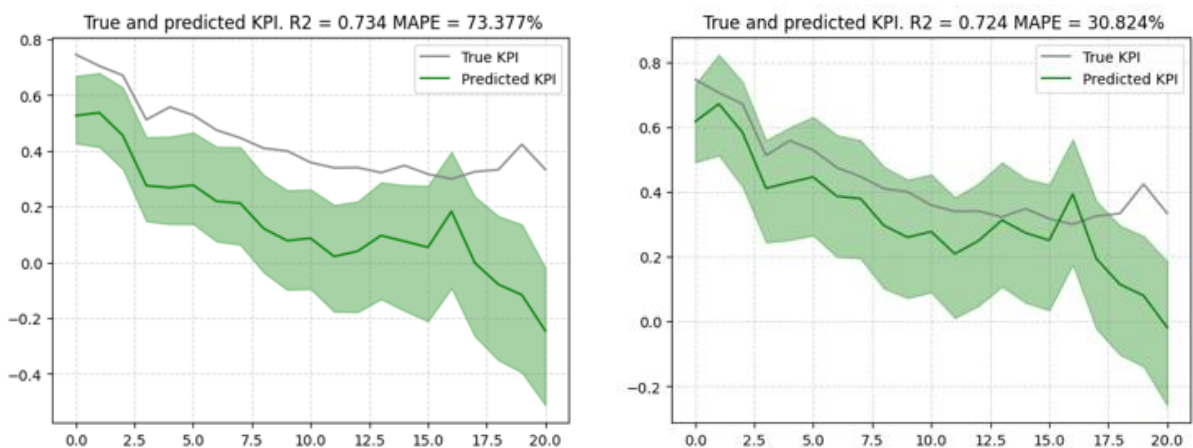
de verschillende kanalen te kunnen onderscheiden. Tijdens de SWI-week is die informatie er niet, maar de onderzoekers werken wel een methodiek uit die zulke data kan verwerken.



Clusteren

Ook zonder extra informatie zien de SWI-deelnemers kansen op meer inzicht in de Opella-sales. Jacobs: "Een standaardaanpak bij multi-collineariteit is dimensie-reductie, waarbij je bijvoorbeeld alle social kanalen samenvoegt." In een paar uur zetten de onderzoekers een data-analyse op waaruit blijkt dat de voorspellende waarde van het model daardoor zou verbeteren. "Maar: dat wilde het bedrijf expliciet niet, ze waren echt op zoek naar de unieke respons van ieder apart kanaal."

Met een slimme list weten de onderzoekers toch voortgang te boeken. "Wij dachten: eigenlijk klopt die aanname dat alle kanalen uniek en onafhankelijk zijn niet", vertelt Jacobs. "Als je een druk Instagram-gebruiker bent, zit je waarschijnlijk ook regelmatig op Facebook of TikTok. Daar zit dus een gekoppeld effect."



Data-analyse waarin alle mediakanalen onafhankelijk worden behandeld (links), vergeleken met dimensie-reductie waarbij kanalen worden geclusterd. Die laatste heeft een grotere voorspellende waarde. Bron: SWI

Filosofisch

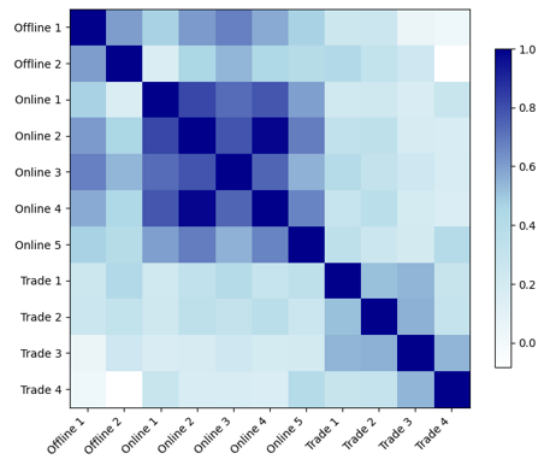
De onderzoekers besluiten verschillende kanalen te clusteren op zo'n manier dat investeringen in die kanalen wel allemaal hetzelfde effect hebben op sales, maar waarbij er toch variatie binnen een cluster mogelijk is. "Met zulke regularisatie (*shrinking*) introduceer je wel extra variabelen, maar op zo'n manier dat je structuur aanbrengt in je probleem en de wiskunde kunt versimpelen", stelt Jacobs.

Een data-analyse laat drie grote clusters zien in de aangeleverde sales-data: sociale media, drukwerk zoals kranten en magazines, en uitzendingen op radio en tv. Toch kiezen de wiskundigen een net andere mix in hun analyse. Jacobs noemt het "bijna een filosofische discussie. Als Bayesiaan vind ik dat je je prior op basis van je eigen overtuigingen moet opstellen. Als je je teveel laat leiden door de data, kun je zelfversterkende effecten krijgen."

Het resultaat van de exercitie: na een kleine week werken liggen er scherpere voorspellingen voor de respons van sales op investeringen in de verschillende mediakanalen. Toch is er nog werk aan de winkel, denkt Jacobs. Bijvoorbeeld door extra data te verzamelen over hoe snel verschillende kanalen sales genereren.

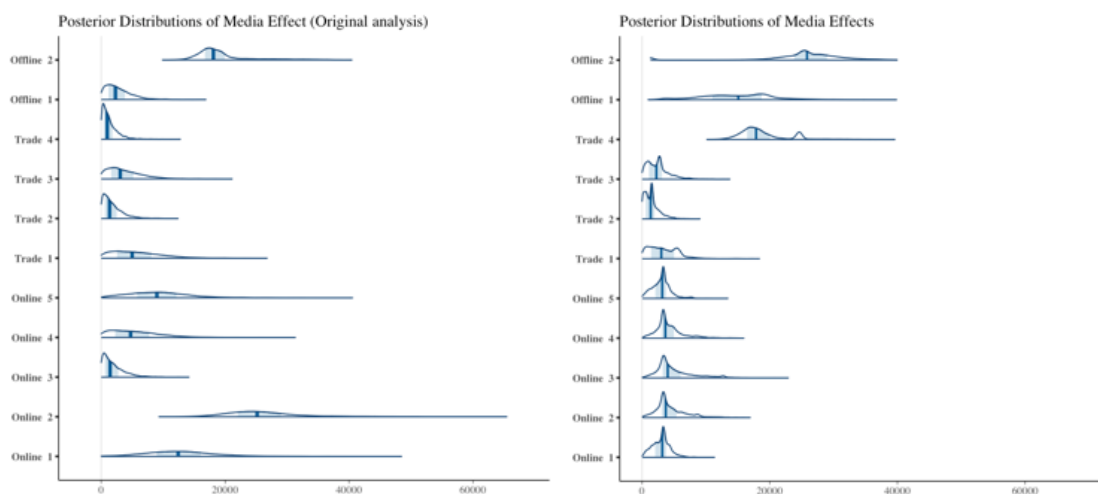
Scherp krijgen

Terugkijkend op zijn eerste SWI-workshop is Tijn Jacobs enthousiast: "Het waren vijf ongelooflijk intense en lange dagen, maar ik had het voor geen cent willen missen. Ik was eigenlijk de enige die vertrouwd was met de Bayesiaanse aanpak. Anderen hadden achtergronden van theoretische optimalisatietechnieken en numerieke wiskunde tot meer toegepast georiënteerd programmeren."



Cluster-analyse van mediakanalen. Bron: SWI

De wiskundigen besloten daarom om eerst de vraagstelling en aanpak te definiëren. "In discussie met het bedrijf merkten we dat zij zelf hun methodes en knelpunten ook nog niet helemaal scherp voor ogen hadden. Dat vroeg constant schakelen tussen enerzijds zelf de diepte ingaan en anderzijds coördineren, uitzoeken en gericht vragen zodat iedereen naar hetzelfde doel werkte. Een uitdagende samenwerking, maar ontzettend leerzaam."



Links: originele posteriors per mediakanaal. De profielen zijn relatief breed, wat een teken is van een lage voorspellende waarde. In de nieuwe analyse (rechts) zijn de profielen scherper gepiekt, wat een sterkere voorspelling betekent. Bron: SWI