

Verlagen van de

Studiegroep Wiskunde met de Industrie 2024



Colofon

De [Studiegroep Wiskunde met de Industrie 2024](#) vond plaats aan de Vrije Universiteit Amsterdam, van 29 januari tot en met 2 februari 2024.

De [Studiegroep Wiskunde met de Industrie 2024](#) werd financieel gesteund door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), Doing the Math, Groendus, IKEA, KLM, The Driving Force en VSL, met ondersteuning van Math4NL en de Vrije Universiteit Amsterdam

Deze [Public Proceedings](#) van de [Studiegroep Wiskunde met de Industrie 2024](#) werden geschreven en opgemaakt door wetenschapsschrijver Gieljan de Vries.

Voorwoord

Voor u ligt het verslag van de Studiegroep Wiskunde met de Industrie (SWI) 2024, gehouden van 29 januari tot 2 februari aan de Vrije Universiteit Amsterdam. Deze editie bracht een groep van ongeveer veertig gedreven wiskundigen samen om zes uitdagende probleemstellingen, aangedragen door diverse bedrijven, aan te pakken.

De SWI 2024 zag bijdragen van bedrijven uit verschillende sectoren, elk met unieke en complexe vraagstukken. IKEA, Doing the Math, KLM, Groendus, VSL, en The Driving Force presenteerden opdrachten die uiteenliepen van het optimaliseren van energie- en logistieke netwerken tot het verbeteren van algoritmen voor boekingscurves en het plannen van voedselproductie.

De intensieve samenwerking tussen wiskundigen en industriële partners leverde innovatieve oplossingen op die vele gebieden van toegepaste wiskunde omvatten, variërend van statistiek en operations research tot machine learning. Deze samenwerking onderstreepte de kracht van wiskundige modellering en probleemoplossing in de praktijk.

De succesvolle organisatie van de SWI 2024 was mede te danken aan de financiële steun van NWO en de ondersteuning van Math4NL en de Vrije Universiteit Amsterdam. Wij willen hen, evenals alle deelnemers en industriële partners, hartelijk bedanken voor hun bijdrage en inzet.

De gedrevenheid en het enthousiasme van de deelnemers waren inspirerend. Binnen slechts een week zijn er veelbelovende resultaten behaald, die niet alleen bijdragen aan de oplossing van de specifieke problemen, maar ook aan de verdere ontwikkeling van toegepaste wiskunde.

We hopen dat de traditie van de Studiegroep Wiskunde met de Industrie nog vele jaren voortgezet zal worden, en dat de samenwerking tussen academici en de industrie zal blijven groeien en bloeien.

Het organisatieteam van SWI 2024,

Sandjai Bhulai (voorzitter)
Rob van der Mei
Joost Hulshof
Angela Evenhuis (afdelingsmanager)
Marielle van der Aa (management assistent)

Inhoudsopgave

Voorwoord.....	3
Inhoudsopgave.....	4
De knooppunten van een nieuw energienetwerk.....	5
Laat dat pakje maar van verder komen	9
Statistiek voor het waterstofnet.....	13
Meer vlaai dan verwacht	17
Schuiven met laadcapaciteit	21
"Wij laten de data zichzelf indelen"	25

Doing the Math

De knooppunten van een nieuw energienetwerk

De duurzame stroom komt het elektriciteitsnetwerk niet op of af. Lastig voor wie zonnepanelen op het dak heeft, maar ook voor bedrijven die willen uitbreiden en verduurzamen. Adviesbureau Doing the Math vroeg de deelnemers aan SWI om wiskundige tools die bedrijventerreinen kunnen helpen zo zelfvoorzienend mogelijk te worden.

Wanneer een ander deelnemend bedrijf een week voor de start van de Studiegroep Wiskunde met de Industrie (SWI) uitvalt, hoeft econometrist John Poppelaars niet lang te denken: hij stapt graag in met een uitdaging aan de wiskundigen. Poppelaars' adviesbureau *Doing the Math* helpt bedrijven met wiskundige tools allerlei duurzaamheidsvragen door te denken. Daar is wel een pittig vraagstuk voor SWI uit te halen.

Kunnen de deelnemers bijvoorbeeld modelleren hoe bedrijventerreinen het beste hun duurzame energieopwekking en energieopslag kunnen delen? Zo'n tool moet niet alleen doorrekenen hoeveel zonnepanelen en batterijcapaciteit bedrijven nu nodig hebben, maar ook hoe ze verstandig kunnen groeien en investeren om verder te vergroenen.

Aan visie geen gebrek: Poppelaars houdt de SWI-ers een toekomst voor waarin zulke bedrijventerreinen knooppunten worden in een groter energienetwerk. Grootverbruikers die meer op eigen groene benen staan zijn niet alleen goed voor die bedrijven en voor het milieu, denkt Poppelaars, maar ook voor netbeheerders die kampen met een overvol energienet.

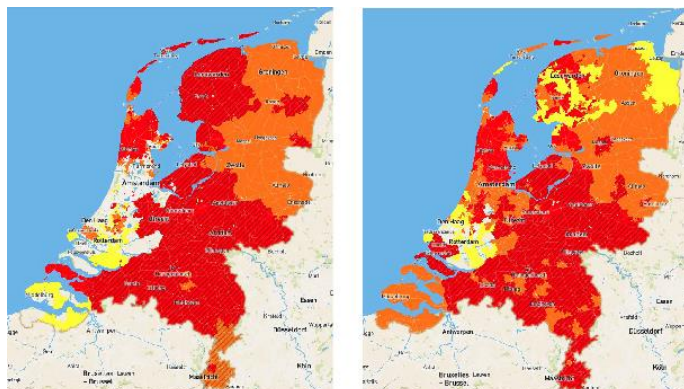


Hoogspanningsmast. Bron: Windpowerengineering.com

Tot de kern

Een alternatief energienetwerk doordenken in de krap vier werkdagen van SWI is zacht gezegd een uitdaging. Naast verschillende technologieën spelen namelijk ook wijd uiteenlopende tijdschalen mee, van dagelijks inspelen op de veranderende weersomstandigheden, tot plannen van meerjarige investeringscycli bij het uitbreiden van de duurzame energieinstallaties.

Een interessant vraagstuk, maar te groot en complex om in een week te kraken, oordelen de SWI-ers dan ook. Deelnemer Martijn Onderwater, gepromoveerd in de wiskunde aan de Vrije Universiteit Amsterdam, neemt na een uitstap naar het bedrijfsleven deel aan SWI om weer aan de wiskunde te snuffelen. "We zeiden al heel snel tegen elkaar: dit is veel te groots, dat krijgen we niet uitgewerkt in een week tijd."



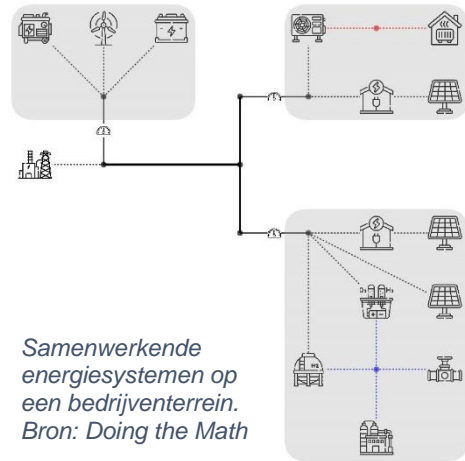
Die complexiteit was paradoxaal ook wat Onderwater aantrok in de opdracht. "De originele opdracht was erg vaag en breed, maar als wiskundige leer je om door te vragen en snel tot de kern van een probleem te komen. We vroegen elkaar: 'Waar gaat dit nou eigenlijk over? En kunnen we in een paar dagen iets opzetten, dat later kan uitgroeien tot een algemenere tool?' "

Het Nederlandse stroomnet heeft weinig overcapaciteit om stroom in of uit het netwerk te sturen.

Bron: <https://capaciteitskaart.netbeheernederland.nl>

Experts

Samen met John Poppelaars werken de vijf wiskundigen een behapbare deelopdracht uit, over de beste manier om van dag tot dag de gedeelde energieproductie en -opslag op een bedrijventerrein te optimaliseren. Die berekening moet rekening houden met het eigen verbruik en de opbrengst van de zonnepanelen op het terrein, maar ook de capaciteit van de gedeelde batterij en hoeveel het kost en opbrengt om stroom in te kopen of te verkopen.



Samenwerkende energiesystemen op een bedrijventerrein.
Bron: Doing the Math

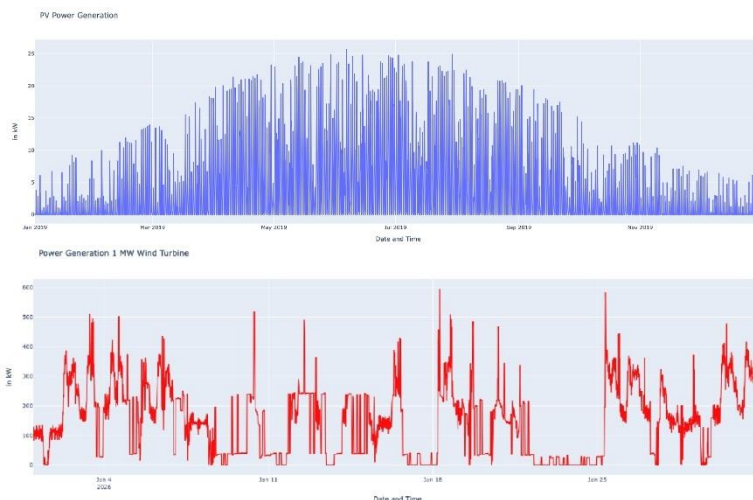
Na een gemeenschappelijke notatie in te voeren om elkaars resultaten te kunnen vergelijken, gaan de wiskundigen in twee teams uiteen. Dat is heel natuurlijk, denkt Onderwater: " Als je een probleem eenmaal goed hebt geformuleerd, zijn er in de wiskunde allerlei methodes om vraagstukken op te lossen. We hadden twee teamleden die elk expert waren in een andere techniek; het lag voor de hand om die in twee teams uit te proberen."

Kansspel

Hoe modelleren je zoiets onvoorspelbaars als de hoeveelheid zon per uur en de prijs van netstroom? In hun eerste aanpak gebruiken de SWI-ers een *stochastische* aanpak die zulke getallen van uur tot uur varieert. Tegen die achtergrond draait een algoritme dat tegen zo laag mogelijke kosten het energieverbruik van het bedrijf moet dekken. De batterij mag daarbij niet voller raken dan de maximum capaciteit, of onder een minimale noodcapaciteit komen.

Een gesimuleerde dag begint met een volle batterij, om te compenseren dat de zonnepanelen door de lage zonnestand nog weinig energie leveren. In opeenvolgende tijdstappen kiest het algoritme of het bedrijf de elektriciteit krijgt uit de zonnepanelen, de batterij of het elektriciteitsnetwerk.

Het model blijkt complex genoeg om verschillende tactieken mee uit te testen, zoals continu stroom kopen of verkopen, of juist voornamelijk interen op de batterij tot de zonnepanelen genoeg stroom leveren om de eigen vraag te dekken of zelfs stroom te verkopen aan het elektriciteitsnetwerk.



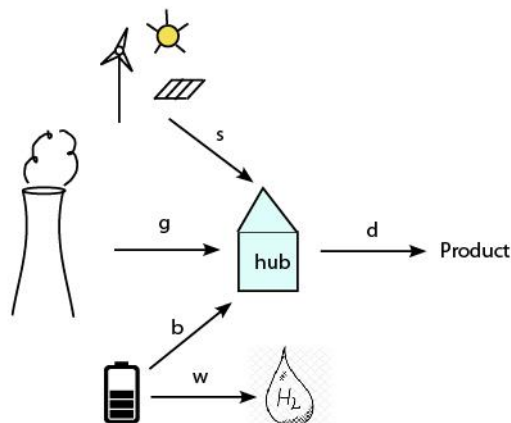
Duurzame energie is sterk variabel op een schaal van uren, dagen en maanden. In blauw de variatie in de opbrengst van een zonnepaneel gedurende een jaar, in rood die van een windmolen gedurende de maand januari.
Bron: Doing the Math

Slim spelen met die in- en verkoop-techniek blijkt fictieve energiekosten terug te kunnen brengen van 18 geldeenheden per dag tot 13, of zelfs 6. In dat laatste geval weigert het algoritme om tijdens de dure piekuren van 12u tot 18u stroom in te kopen. Sterker nog; met dynamisch inkoopgedrag komt het er zelf achter dat het slim is om vlak voor die piek de batterij af te topen met netstroom, om die tegen hogere prijs terug te verkopen.

Holle bolle functies

In de energietransitie en het modelleren daarvan zijn er altijd meerdere wegen naar Rome. Zo werken de SWI-deelnemers naast hun stochastische model een andere formulering uit die de energiehuishouding van hun bedrijventerrein beschrijft in ongelijkheden.

Martijn Onderwater: "Dit is een iets andere aanpak dan het stochastische model. Met deze tweede aanpak krijg je uiteindelijk een hol of bol lopende grafiek (in vaktermen: convex of concaaf) die je probleem beschrijft. Door de manier waarop we het probleem formuleren zit de beste oplossing altijd bij een extreem. De techniek om dat optimum te vinden werkt heel snel, dat is een voordeel boven de uitgebreide simulatie van energiestromen in het andere model."



Cartoon van een energieknooppunt.
Bron: SWI-team Doing the Math

Twee rekenmodellen op verschillende leest, hoe vergelijk je die fatsoenlijk? De wiskundigen gingen weliswaar in gespecialiseerde clubjes uiteen, maar hielden dagelijks contact om elkaar bij te praten. "We hadden het geluk dat ons teamlid, nu bezig aan haar promotie in Oxford, beide wiskundige technieken in de vingers had", vertelt Onderwater. "Daardoor kon ze niet alleen sparringpartner zijn voor onze twee experts, ze werd als vanzelf de brug die ons als team bij elkaar hield. Dat was indrukwekkend om te zien."

Ondersteunend

Vergeleken met het originele verzoek om de samenstelling en uitbreiding van kleine energiemaatschappijen te modelleren, lijkt een rekentool voor de energiehuishouding van het zonnepark misschien een wat bescheiden opbrengst. Maar opdrachtgever John Poppelaars ziet groeipotentieel.

"Dit klinkt als een klein probleem, maar het is een handige tool voor bedrijven om mee uit te werken wat voor capaciteit aan zonnepanelen en batterijen ze nodig hebben", denkt Poppelaars. Het tekent de stijl van Doing the Math, dat wiskundige tools ontwikkelt om managers te ondersteunen in hun besluitvorming. Bijvoorbeeld als er bedrijven bij komen op het terrein, of als er naast elektriciteit ook duurzame warmte of waterstof wordt geproduceerd en opgeslagen.

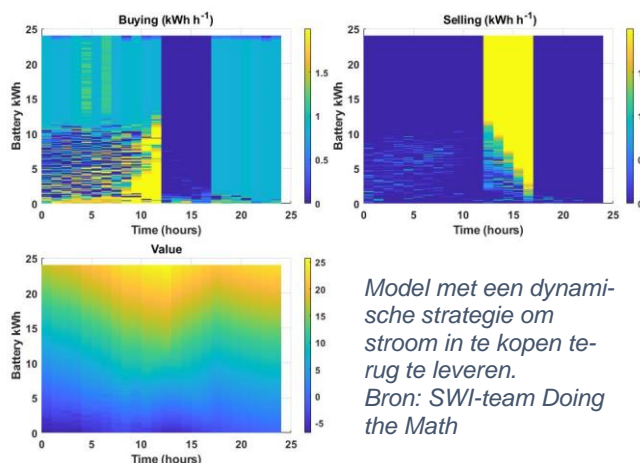
Uitbreiden

"Waterstof maken uit duurzame stroom als langdurige energieopslag naast batterijen hebben we uiteindelijk niet uitgebreid doorgerekend", vertelt wiskundige Martijn Onderwater. Een gevolg van tijdsgerebrek: een extra opslagmedium is zo in formules gezet, maar maakt het vraagstuk meteen complexer

om op te lossen. Dat moet namelijk gaan kiezen tussen verschillende routes om energie heen te sturen.

In de drie werkdagen van SWI was waterstofopslag teveel gevraagd, denkt ook opdrachtgever John Poppelaars. "Die uitbreiding kan later komen. Het is goed om nu te zien hoe goed de basis al werkt."

Gevraagd of zijn originele vraagstelling niet te ambitieus was: "Ik heb me geen moment zorgen gemaakt dat ik deze case 'te groot'



Model met een dynamische strategie om stroom in te kopen terug te leveren.
Bron: SWI-team Doing the Math

heb gemaakt. Wiskunde is veel meer dan sommetjes oplossen, het helpt je gedachten te structureren, complexe problemen te doorzien en daar behapbare vragen van maken. Dat is ook hier perfect gelukt."

Gaan de rekenmodellen van de SWI-ers ook echt gebruikt worden? Poppelaars ziet kansen. Zijn vraagstuk over lokale industrieterreinen die hun stroom en warmte delen is namelijk ook interessant voor netbeheerders, als manier om het hoofdnet te ontlasten. "Die organisaties inventariseren nu al waar de komende jaren nieuwe knelpunten kunnen ontstaan. Als bedrijven door slim samenwerken minder op het hoofdnetwerk hoeven te leunen en misschien zelfs als energieopslag kunnen dienen voor de omgeving, is dat winst voor iedereen."

IKEA

Laat dat pakje maar van verder komen

IKEA is al lang niet meer door de winkel dolen op Tweede Paasdag; steeds meer klanten bestellen online. Een uitdaging voor de webshop, die live moet bepalen welke winkel of distributiecentrum de bestelling het beste kan leveren. Vijf SWI-ers bouwen er vijf uiteenlopende modellen voor, mét Zweedse knipoog.

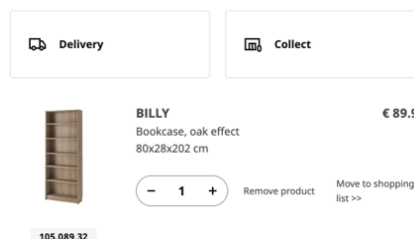
Een hanglamp, een bureau en (Billy) boekenkast voor een studentenkamer, een complete keuken: de webshop van IKEA heeft het allemaal. Klik op 'bezorgopties tonen', en de website geeft je een lijstje van mogelijke bezorgmomenten waarop jouw nieuwe huisraad voor de deur kan staan.

Om die opties te bepalen moet de IKEA-server in een oogwenk uitzoeken welke winkel of distributiecentrum de bestelling het beste kan versturen. Geen simpel sommetje, stelt data-expert Jan de Munck van IKEA. Naast transportkosten is er ook het aspect van *order picking*: hoeveel tijd het medewerkers kost om artikelen in het magazijn op een karretje te laden. Dat varieert per locatie vanwege de beschikbare capaciteit en de indeling van de winkels.

"Ons huidige algoritme berekent telkens wat de goedkoopste oplossing is voor die ene order", legt De Munck uit. Meestal een prima oplossing, maar op drukke dagen kan de capaciteit van een kleine locatie door die toewijzing snel vollopen. Een grotere order later op de dag past dan niet meer in de planning. Die moet dan tegen hogere kosten van verder weg komen.

Shopping cart

How would you like to receive your order?



Order summary

Products price € 89.95

Delivery Will be visible in the next step

Subtotal € 89.95

Have a discount code?

Continue to checkout

Screenshot van de IKEA-webshop. Bron: SWI-team IKEA

Kan het slimmer? De Munck heeft vanuit zijn vorige werkgever PostNL al twee keer deelgenomen aan SWI en stelt voor om die vraag in te brengen namens IKEA. Een ideaal vraagstuk voor SWI, vindt ook De Munck's collega Arturo Pérez Rivera, hoofd van de afdeling orderallocatie.

"In feite vragen we een manier om in de toekomst te kijken en bij de orderverdeling rekening te houden met toekomstige bestellingen", vat De Munck samen. Door opdrachten beter te verdelen over de winkels moet er ook op drukke dagen capaciteit overblijven. "We vermoeden dat dat opweegt tegen de iets hogere transportkosten."



Girig



Optimål



Heuristik



Linjär



Netvörk

Ontwikkelde algoritmes tijdens de workshop

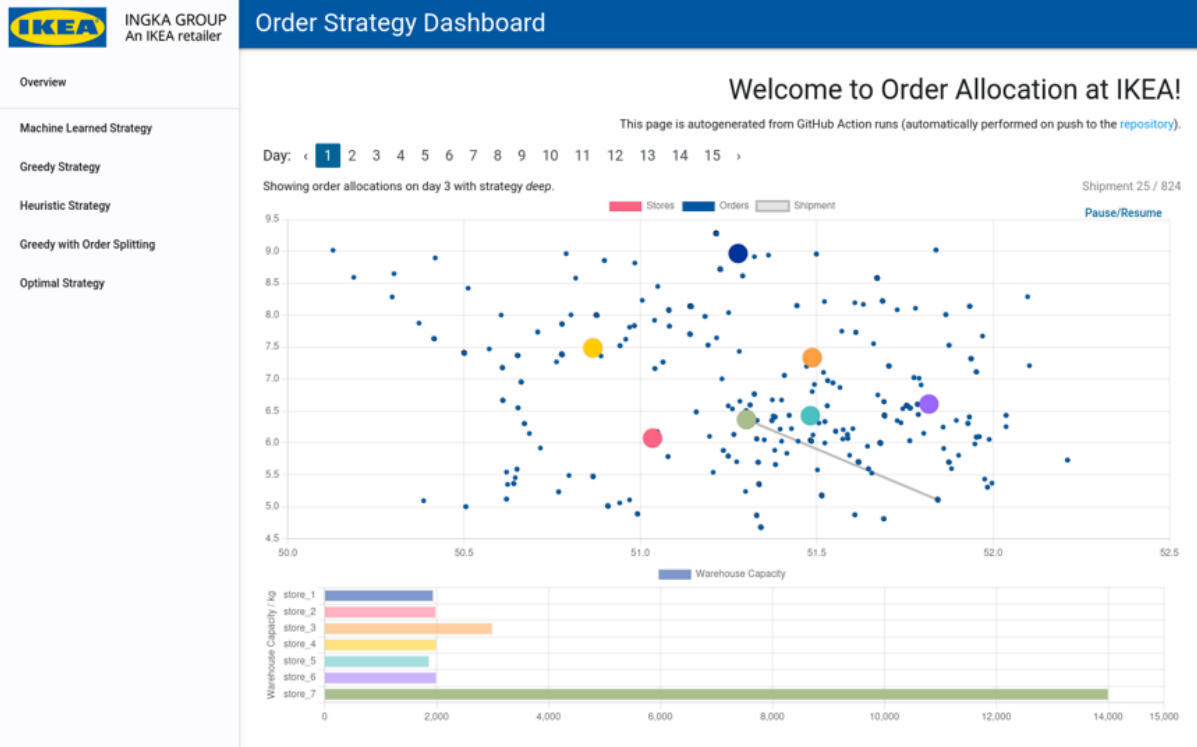
Creatieve wiskunde

Vraag SWI-ers één glazen bol en je krijgt er vier. Aan het eind van de workshop presenteren vijf deelnemers een gelijke dashboardomgeving die vier nieuwe rekenmodellen in detail vergelijkt met IKEA's bestaande algoritme. De sfeer zit erin: een verloren uurtje is besteed aan het verzinnen van Zweeds klinkende namen voor de algoritmes.

"We hadden een ideaal team voor deze opdracht", verklaart teamlid Yasmin Roshandel de opbrengst. "We hadden experts in data-analyse, optimalisatieproble-

men en *deep learning*, en iemand die thuis is codebases beheren zodat we alle modellen in één dashboard konden vergelijken."

In het dagelijks leven promoveert Roshandel aan de Vrije Universiteit Amsterdam op het optimaliseren van logistieke ketens voor het bedrijf Pon. "Optimalisatie is creatieve wiskunde: je bedenkt telkens nieuwe oplossingen om zo snel en goed mogelijk je doel te bereiken. Die oplossingen direct toepassen voor een bedrijf is echt heel leuk."



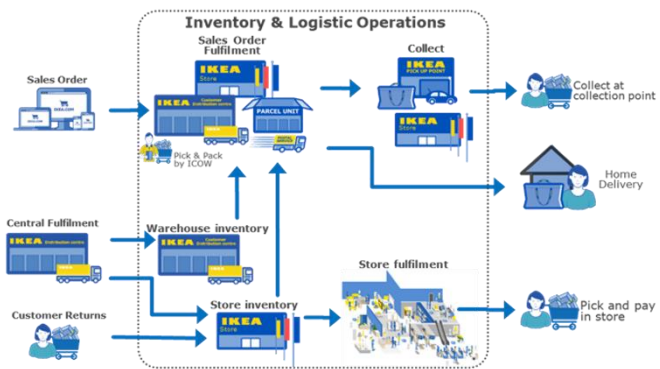
Bekijk het dashboard met verschillende algoritmes dat SWI ontwikkelde voor IKEA: <https://ikea-order-strategy-dashboard.netlify.app/>

Beste oplossing: Optimål

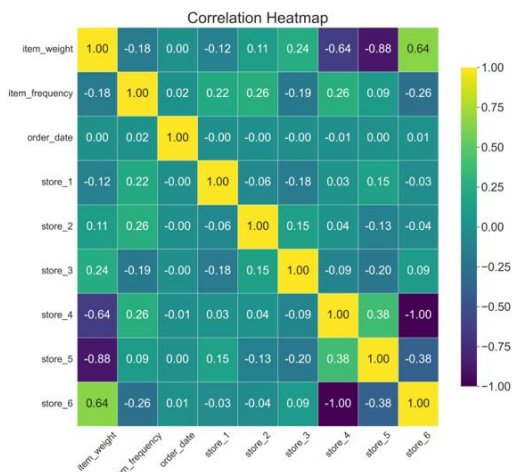
Aan het begin van hun werkweek krijgen de wiskundigen van Jan de Munck een uitgebreid startpakket. Daarin zit een sterk versimpelde dataset met producten en hun gewicht, winkels en hun voorraad, een maand aan fictieve orders én een implementatie van IKEA's bestaande 'greedy' algoritme, prompt omgedoopt tot *Girig*. Kunnen de SWI-ers het beter dan dit algoritme, dat consequent de goedkoopste optie kiest zonder rekening te houden met de capaciteit van het netwerk?

Als eerste vingeroefening zochten Roshandel en haar collega's onder de vlag *Optimål* uit of er een ideale manier is om de dagorders tegen de laagste totale kosten te verdelen over de winkels. Daarvoor gebruikten ze de krachtige maar rekenintensieve techniek van *Integer Linear Programming (ILP)*.

"Bij ILP moet je je complete vraagstuk formuleren in wiskundige relaties tussen grootheden", legt Roshandel uit. Ook aspecten die voor een mens vanzelfsprekend zijn, zoals dat een artikel moet worden verzonden vanuit dezelfde winkel waar het op voorraad is. Een geautomatiseerde ILP-oplosser kan dan de best



Inventaris en logistiek binnen IKEA. Bron: IKEA



Analyse van de inventaris van distributiecentra in de trainingsdata. Bron: SWI-team IKEA

mogelijke oplossing vinden. Daarvoor moet het systeem wel alle te verdelen orders kennen. Roshandel: "We doen dus alsof we aan het eind van een winkel-dag terugkijken en de orders toewijzen."

In het dashboard is meteen te zien hoeveel beter Optimál op drukke dagen presteert dan Girig: de totale kosten zijn de helft lager, en maar 1% van geplaatste orders kan niet worden vervuld, tegen Girig's 4%. Nadeel: deze optimale oplossing is totaal ongeschikt voor de webshop vanwege de forse rekentijd en de eis dat toekomstige bestellingen al bekend zijn. Is daar een mouw aan te passen?

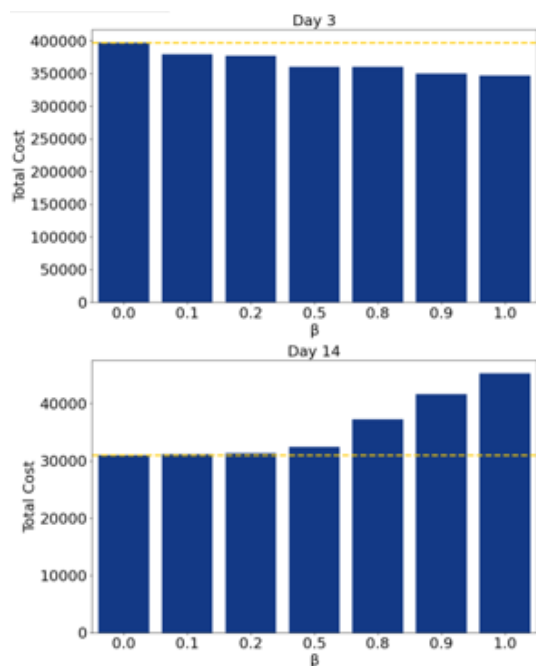
Linjär algoritme

In de toekomst kijken kan niet, in het verleden maar al te goed. Het Linjär-algoritme maakt daar dankbaar ge-

bruik van. De basis is weer een ILP-berekening die informatie nodig heeft over toekomstige opdrachten. Die faken de SWI-ers slim door willekeurige bestellingen uit de database van historische orders te plukken. Linjär doet alsof ook die gesimuleerde bestellingen ingepland moeten worden en bepaalt waar de bestelling van de huidige klant vandaan moet komen.

Met maar 600 gesimuleerde orders haalt dit model al 75% van de besparingen van het optimale algoritme, vertellen de wiskundigen in hun presentatie. Dat is netjes, maar met een minuut rekentijd te traag voor de webshop. Dat kan in theorie sneller, denken de wiskundigen. Linjär houdt rekening met de steeds dalende ordercapaciteit van winkels om opdrachten over alle winkels te spreiden. Een sneller alternatief is om de ordercapaciteit te formuleren als

schaduwrijzen. Zo mijdt het algoritme vanzelf 'dure' locaties, oftewel de winkels waar de ordercapaciteit bijna op is. Die schaduwrijzen kan buiten het ILP-algoritme worden berekend en aangepast, wat de tool in de webshop rekentijd zou schelen.



Invloed van de drukke-dagenfactor op de totale bezorgkosten in Heuristik. Bron: SWI-team IKEA

Heuristik

Linjär en Optimál vragen veel rekenkracht. Daarom proberen andere teamleden het snelle Girig-algoritme te verbeteren. Vergelijkbaar met het schaduwrijzenconcept houdt Heuristik de restcapaciteit per locatie bij. Het algoritme kiest niet simpelweg de goedkoopste winkel net als in Girig, maar weegt de resterende capaciteit van elke winkel mee. Zo worden overbelaste winkels minder aantrekkelijk voor het algoritme.

De Heuristik-aanpak heeft geen informatie over toekomstige bestellingen nodig, omdat het informatie over de restcapaciteit gebruikt. Een drukke-dagen-factor geeft aan hoe zwaar die capaciteit moet wegen en die drukke-dagen-factor rust weer op een voorspelling van

hoeveel bestellingen er worden verwacht die dag. Resultaat: zo'n 15% lagere kosten dan het Girig-algoritme, met dezelfde rekentijd.

Netvörk

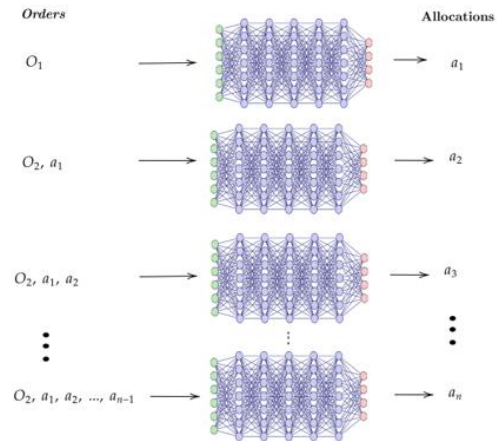
Zowat het enige SWI-algoritme dat niet beter presteerde dan IKEA's standaardberekening is een *deep learning*-aanpak. Hier moest een *neuraal netwerk* patronen leren herkennen in de orderverdeling tijdens een winkeldag. Zo'n patroon ('drukke dag, veel zware bestellingen') wordt dan gekoppeld aan passende orderverdelingen, geleerd van oplossingen van het Optimäl-model.

"Helaas hadden we hier te weinig tijd en te weinig data voor", haalt Roshandel de schouders op. Een maand aan trainingsdata lijkt veel, maar blijkt te weinig voor het grote aantal variabelen zoals winkels, ordersamenstellingen, artikelprijzen en -gewichten. Netvörk is razendsnel, maar presteert -nu nog- minder dan Girig. Met meer trainingsdata is dat waarschijnlijk te verbeteren.

Aan de slag

IKEA's Jan de Munck steekt zijn waardering voor het SWI-resultaat niet onder stoelen en banken. "Wat de SWI-ers hebben geleverd is in een woord geweldig. Ze hebben meerdere algoritmes bedacht die echt kosten kunnen besparen."

"Het idee van schaduwrijzen voor winkels die tegen hun capaciteit aanlopen is inspirerend! En het model dat toekomstige orders simuleert is alsof je een kristallen bol in handen hebt. Zelf vind ik dit de mooiste oplossing, maar ik weet nog niet of we die ook binnen ons bestaande systeem kunnen implementeren."



Neuraal netwerk. Bron: SWI-team IKEA

Ook Yasmin Roshandel en haar groepsgenoten zijn enthousiast; ze overwegen zelfs een publicatie in een vakblad over hun oplossingen. Hierbij hebben ze een hoop ideeën om machine learning te gebruiken om de algoritmes te voeden met informatie uit het verleden. Roshandel: "IKEA had hun probleem heel goed geformuleerd: ze wisten wat ze hieruit wilden halen en we kregen alle ingrediënten voor een vliegende start. Daardoor konden we veel voor elkaar krijgen."

Waardevol

De SWI-algoritmes zijn volgens De Munck mooie startpunten. Toch is er nog veel werk nodig om ze klaar te maken voor de veel complexere werkelijkheid van IKEA. Om de SWI-ers een vliegende start te geven, heeft De Munck zijn casus namelijk net als bij voorgaande edities "gigantisch versimpeld. Ik houd ervan om het probleem al zover mogelijk in te kaderen. Als opdrachtgever krijg ik meer uit de week als de wiskundigen hun kwaliteiten helemaal kunnen richten op slimme oplossingen verzinnen."

Producten hebben in de SWI-casus bijvoorbeeld een gewicht, maar geen afmetingen; magazijnen hebben een product óf onbeperkt op voorraad of helemaal niet. Ook het aspect tijd is platgeslagen. Orders moeten op een dag worden geleverd; ze worden niet doorgeschoven naar rustiger tijden. Dat lijkt weinig realistisch, maar laat wel de essentie van het capaciteitsprobleem zien.

Ook oplossingen voor versimpelde problemen zijn waardevol, vindt De Munck. Hij prijst de grote nieuwsgierigheid van de deelnemers: "Ze pellen je vraagstuk helemaal uit en komen met oplossingen die je zelf nooit had bedacht. Dankzij het optimale algoritme kunnen wij bijvoorbeeld zien hoe ver onze eigen webshop boven de laagst mogelijke kosten zit. Heel handig."

Roshandel: "Deze week was heel intensief, maar ik heb enorm genoten! Ik zou andere promovendi zeker aanraden om mee te doen aan SWI. Je leert mensen kennen, je ontdekt hoe het is om samen te werken met bedrijven, en je komt allerlei wiskundige technieken tegen die je nog niet kent. Leuk én waardevol."

VSL

Statistiek voor het waterstofnet

Willen we in de toekomst de duurzame energiedrager waterstof door het aardgas mengen, dan moeten we voortdurend de kwaliteit van de gasmix meten. Daar blijkt pittige statistiek achter te zitten. Zes SWI-wiskundigen buigen zich over het vraagstuk.

Nederland, aardgasland. Volgens het Centraal Bureau voor de Statistiek is nog maar 8,1 % van de huizen volledig aardgasvrij; de rest gebruikt aardgas om te koken of de woning te verwarmen. Daar moet verandering in komen, want aardgas verbranden—methaan, voor de chemici—produceert veel CO₂ en draagt bij aan klimaatopwarming. Tegelijkertijd hebben we decennialang geïnvesteerd in het aardgasnet.

Zijn er manieren om het aardgasnet te verduurzamen, bijvoorbeeld door duurzaam geproduceerd waterstofgas door het methaan te mengen? Die schone energiedrager stoot geen CO₂ uit bij het verbranden. Ook de productie kan schoon, door water te splitsen (elektrolyse) met schone stroom uit wind en zon. Bestaande installaties kunnen al omgaan met een klein percentage waterstof in de aardgasmix. Toekomstige installaties kunnen zelfs volledig op waterstof draaien.



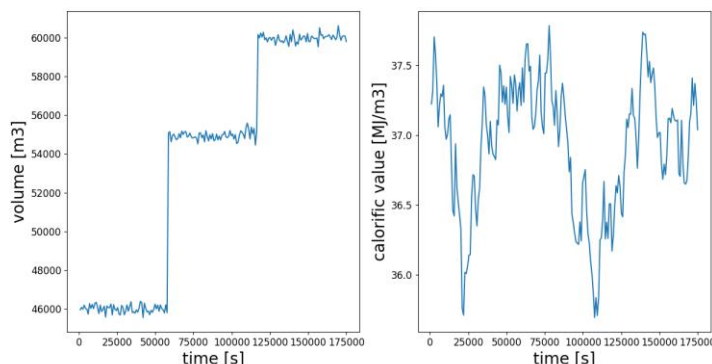
Aardgasnetwerk. Bron: Gasunie

Een mooi vooruitzicht, maar er zijn nog genoeg hordes te nemen voor we waterstof door het aardgas mengen. Hoe meet en bewaak je bijvoorbeeld de kwaliteit van de gasmix? Dat is de achtergrond van de SWI-casus van het nationaal metrologisch instituut VSL, de Nederlandse experts in meetstandaarden en nauwkeurige meettechnieken. Zij onderzoeken de lastige statistiek van zorgvuldig meten van energietransport door het toekomstige gasnet.

Meetpunten

Fysisch chemicus Adriaan van der Veen van Data Science & Modelling bij VSL vertelt over de haken en ogen van metingen aan een uitgebreide gasinfrastructuur. Vanuit VSL coördineert hij het Europese onderzoeksproject [MET4H2](#) (*Metrology for the hydrogen supply chain*) over waterstof bijmengen in het gasnet en waterstofgasnetten.

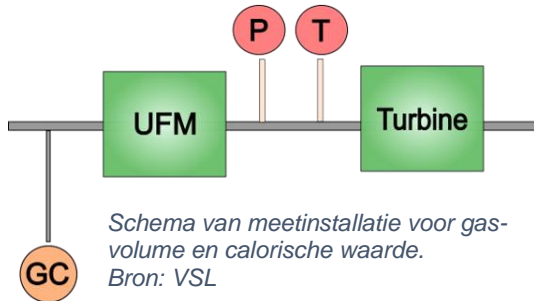
De gasmix door het hele netwerk kennen is lastiger dan aan de bron bijhouden hoeveel methaan en waterstof je erin pompt, legt Van der Veen uit. En belangrijk. Klanten willen bijvoorbeeld de juiste verhouding van gassen krijgen voor hun apparatuur en willen weten hoeveel energie ze geleverd krijgen. Ook voor leveranciers is het belangrijk om te weten welke mix er aankomt bij de klant, en of er geen ontmenging plaatsvindt in delen van het netwerk.



Meting van passerend gasvolume en de calorische waarde.
Bron: VSL

Leveranciers hebben dan ook meetpunten door hun hele netwerk. Bij de bron, bij verdeelstations en bij eindstations in steden meten stromings-, druk- en tem-

peratuursensoren en analyseapparatuur hoeveel gas er met welke kwaliteit langs stroomt. In de toekomst moeten die meetstations ook het waterstofgehalte bepalen: hoeveel waterstof zit er door het methaan gemengd? Die verhouding bepaalt de hoeveelheid energie per kuub gas. Vermenigvuldigd die *energiedichtheid* met het debiet (het langs gestroomde volume) en je weet hoeveel energie er per tijds-eenheid is getransporteerd.



Onzekerheidsmarge

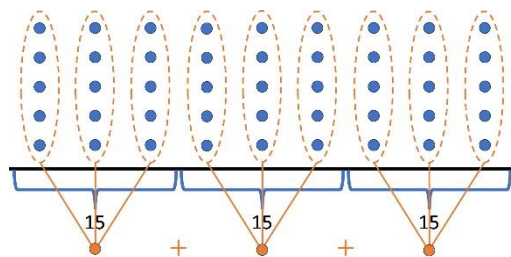
Bij het meten van debiet en samenstelling van het gasmengsel worstelen Van der Veen en zijn collega's nog met het bepalen van de bijbehorende onzekerheidsmarge. Die geeft aan hoe goed je de uitkomst van de metingen kunt vertrouwen; cruciaal om te weten of er de juiste hoeveelheid gas of energie wordt geleverd.

"Voor het debiet en de energiedichtheid kun je aparte onzekerheden bepalen", legt de VSL-onderzoeker uit.

Die hangen samen met de gebruikte sensoren en zijn voor elke sensor te bepalen. "De vraag is wat er gebeurt als je de datapunten met elkaar vermenigvuldigt om de totale energie te berekenen."

In de handboeken statistiek staan allerlei methodes om de onzekerheidsmarge van zulke berekeningen te bepalen. Tenminste: als de gebruikte datareeksen geen interne samenhang hebben in de tijd. In werkelijkheid is die er natuurlijk wel, omdat de meetwaarden van nu sterk gerelateerd zijn aan die uit het verleden.

Van der Veen: "De handboeken en normen waarschuwen dat de standaardmethodes te kleine onzekerheidsmarges voorspellen als de meetreeksen een interne samenhang hebben, maar de methode om die onzekerheid in dat geval wél goed te berekenen is er nog niet. Hoe ga je daarmee om? Dat wilden we weten van SWI."



Tijdreeksen combineren en middelen.
Bron: SWI-team VSL

Volgorde

Tijdens de workshop buigen zes wiskundigen zich over de statistiek van VSL. Onder hen is postdoctoraal onderzoeker Erwin Luesink van de Universiteit Twente. Zijn vakgebied: stochastische geometrische mechanica, ofwel hoe je omgaat met natuurlijke variaties en onzekerheden in meetdata. Bijvoorbeeld als je satellietmetingen van de oceaantemperatuur wil koppelen aan klimaatmodellen.

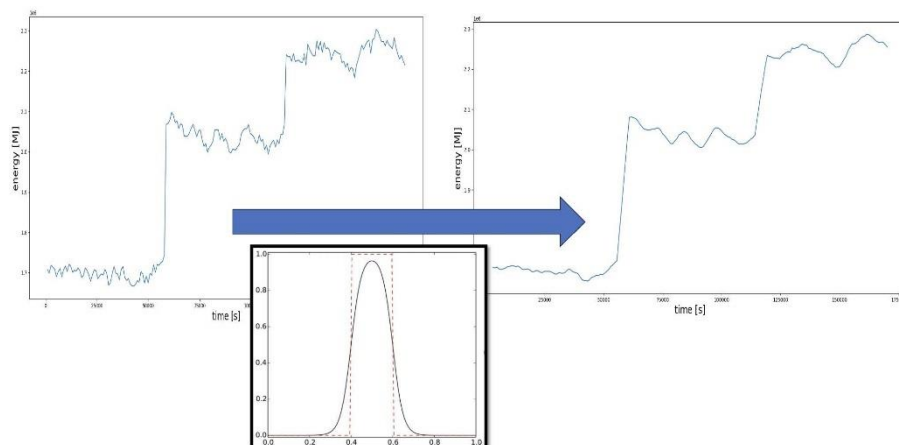
"Deze case van VSL ligt wel op mijn pad, maar is ook weer niet helemaal mijn eigen specialisme. Ik wilde een project waar ik iets nieuws tegen zou komen. Het interessante aan dit vraagstuk is die gekoppelde onzekerheid doordat metingen op verschillende tijdstippen niet onafhankelijk zijn. En ook de volgorde van bewerkingen is belangrijk, ontdekten we tijdens deze week."

Neem je bijvoorbeeld eerst het gemiddelde van een stel debietmetingen tijdens een stel meetmomenten, en vermenigvuldigd je dat met het gemiddelde van de bijbehorende energiedichtheden? Dan krijg je een andere uitkomst dan wanneer je eerst de twee grootheden per meetmoment vermenigvuldigt en daarna pas een gemiddelde neemt. Zo ook voor de gekoppelde meetonzekerheid.

Gladstrijken

Luesink en zijn vijf teamgenoten hebben niet allemaal een achtergrond in de statistiek, maar weten wel al snel hoe ze het VSL-vraagstuk aan willen vliegen. "We hebben te maken met een signaal met daarbovenop ruis, we weten dat we geïnteresseerd zijn in statistische eigenschappen zoals variantie en covariantie; dan ligt het erg voor de hand om filtertheorie toe te passen."

Filteren is een manier om grillige meetseries glad te strijken en alleen de essentiële eigenschappen over te houden, legt de onderzoeker uit. Door elk meetpunt te middelen met zijn burens, verdwijnt chaotische ruis en blijft er een gladder signaal over. Dat kan eenvoudig - simpelweg het gemiddelde van de drie metingen - of complexer, waarbij naburige waardes gewogen meetellen.



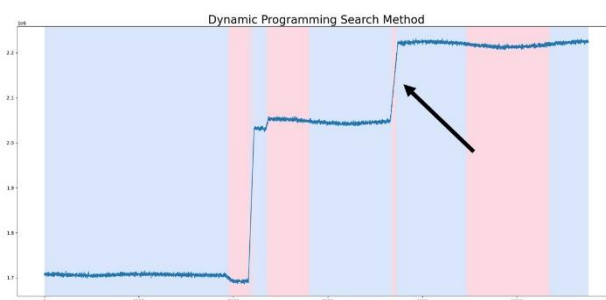
*De filtermethode maakt gewogen gemiddeldes van naburige meetpunten.
Bron: SWI-team VSL*

Sprongherkenning

Zo gezegd, zo gedaan. Maar dan blijken er addertjes onder het gras te zitten. Zo zitten er scherpe sprongen in de meetgegevens, waarbij volume of samenstelling ineens omhoog of omlaag schieten en dan stabiliseren op een nieuw niveau. Volledig verklaarbaar: op koude winterochtenden springen de verwarmingen van huizen bijvoorbeeld op vergelijkbare momenten aan en stijgt het verbruik. Maar zulke sprongen gooien lelijk roet in het eten van de filtermethode.

Om de data te verwerken testen de wiskundigen algoritmes die automatisch sprongsgewijs gedrag herkennen in de meetseries. Data tijdens zo'n sprong worden apart geanalyseerd, zodat ze de omliggende data niet vervuilen. Aandachtspunt blijft het goed identificeren van de precieze momenten waarop een sprong begint en eindigt, zodat de verandering wordt uitgesmeerd over het juiste tijdsinterval.

Opricht afgerond? Niet bepaald. Als de SWI-ers eenmaal hun automatische sprongdetectie aan de gang hebben, testen ze hun filteraanpak op de aangeleverde data. Wanneer ze daar statistische grootheden als variantie en covariantie voor uitrekenen, slaan de uitkomsten nergens op. Een verrassing, maar wel een die mooie bijvangst gaat leveren.



Sprongen in de data herkennen. Bron: SWI-team VSL

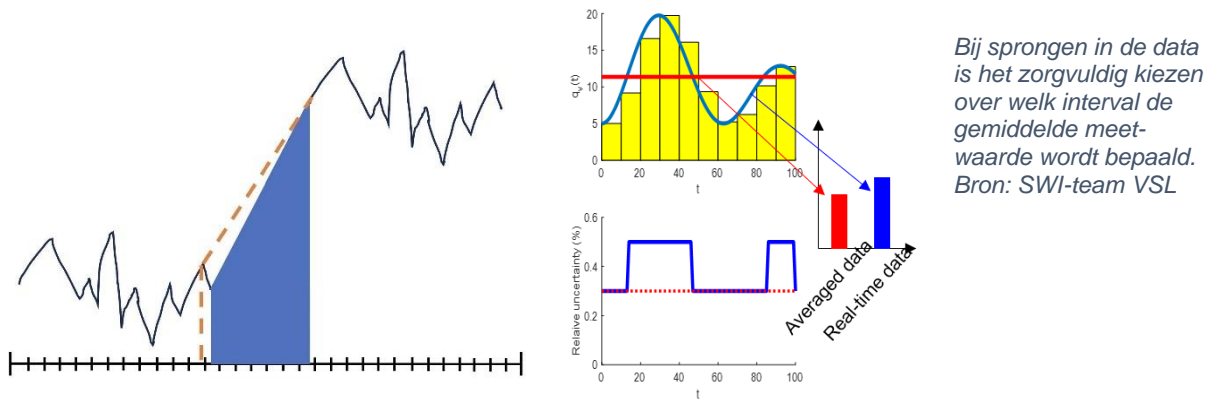
Data error

Luesink: "VSL gebruikt zelf meetgegevens van een commerciële energiepartij, die ze niet zomaar aan ons mogen geven. Daarom hebben ze zelf kunstmatige datareeksen gegenereerd die geen bedrijfsgeheimen weggeven."

Begrijpelijk, maar in die gesimuleerde meetreeksen blijken metingen op opeenvolgende momenten niet met elkaar samen te hangen. Precies die afhankelijkheid zorgt dat de standaardstatistiek vastloopt op de echte datareeksen waar VSL mee worstelt.

Luesink: "We ontdekten al snel dat die data niet representatief waren toen we ze door een paar statistische tests haalden. Daarom hebben we zelf synthetische data geproduceerd waar die relaties tussen datapunten wel in zitten. Met het blote oog zie je het onderscheid niet, ga je dieper kijken dan is er een wereld van verschil."

De datagenerator maakt meteen een nieuwe onderzoeksstrategie mogelijk, denkt opdrachtgever Van der Veen: een geijkte test van de analysemethodes. Door data te analyseren waar de relevante statistische eigenschappen van vaststaan, kun je namelijk testen hoe goed je analysealgoritme die er weet uit te halen.



Europees

Hoe goed is de originele vraag van VSL beantwoord? Adriaan van der Veen ziet veel handvatten voor verder onderzoek. "Er is mooi werk geleverd in de vorm van die automatische filtermethode en sprongherkenning." Hij verwacht interesse van Europese gasbedrijven, die voor de uitdaging staan om hun infrastructuur zo goed (en goedkoop) mogelijk te bemeten. "Uiteindelijk willen die weten op welke punten in hun netwerk ze uitgebreide meetpunten moeten plaatsen die elk kwartier uitlezen, en waar het voldoende is om een sensor te plaatsen die eens per dag rapporteert. Dit onderzoek over de betrouwbaarheid van de data is daar een mooi startpunt voor."

Tevreden opdrachtgever, tevreden wiskundigen? Erwin Luesink, ondertussen voor de tweede keer deelnemer aan SWI, is enthousiast. "Het was een verfrissende uitdaging om eens een week helemaal anders te werken dan normaal. Als wiskundige zit je niet zo vaak in groepsverband aan hetzelfde project. En dan ook nog een nieuw onderwerp, met mensen en technieken die je nog niet kent - bijna alsof je even vakantie hebt.

"Als een bedrijf zijn opdracht goed voorbereidt, kom je als wiskundige in een prachtige speeltuin waar je met allerlei vrijheid onderzoek kunt doen en aan een bruikbaar resultaat werkt."

The Driving Force

Meer vlaai dan verwacht

Hoeveel vlaaien moet een bakker elke dag maken om geen klanten teleur te stellen? En verandert dat antwoord als je weet dat klanten in plaats van hun eerste keus ook genoeg nemen met andere producten? Het SWI-team 'Vlaai Substitutieprobleem' schotelt drie antwoorden voor op dit smakelijke probleem.

Of ze tevreden zijn met de opbrengst? Wiskunde-promovendus Luke Visser aan de Technische Universiteit Eindhoven: "Onze analyses doen het iets beter dan de standaardaanpak, dat is gewoon goed. Revolutionair is het ook weer niet, maar we hebben mooie wiskundige oplossingen geformuleerd."

Opdrachtgever Frank Ottenhof van adviesbureau *The Driving Force* is meer uitgesproken positief. "Ik ben juist behoorlijk verbaasd over wat de wiskundigen allemaal voor elkaar hebben gekregen! In een week tijd hebben ze goed bruikbare voorspellende tools gebouwd. En door hun werk zijn wij ervan overtuigd dat we terug moeten keren op een eerdere keuze. Zo kunnen we onze klanten beter helpen. Een fantastisch resultaat dus."



Hoeveel stuks van elke soort gebak moet de bakker elke dag voorbereiden? Bron: the Driving Force

moet je zorgen dat je op precies x procent van de dagen niet uitverkocht raakt. Dus als je een kleine marge hebt van 20%, moet je 80% van de tijd uitverkocht raken."

Voor de Studiegroep Wiskunde met de Industrie 2024 duiken Visser en vijf collega-onderzoekers in de vraag: hoe verandert de klassieke Newsvendor-oplossing als klanten van een Limburgse bakkerij ook genoeg nemen met een tweede keus wanneer hun favoriete vlaai niet meer op voorraad is? En hoe zit dat als de klant een product alleen koopt als er ook een ander op voorraad is - koffie en een vlaai-punt - of wanneer de verkoper de specialiteit van het huis altijd op voorraad wil hebben?

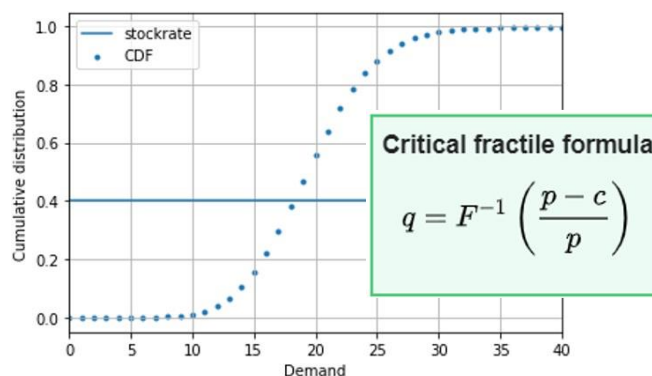
Opdrachtgever *The Driving Force* bakt zelf geen vlaaien, maar levert advies en tools op maat voor midden- en kleinbedrijven. Bijvoorbeeld met software die de optimale dagvoorraad voor versbedrijven uitrekent.

*Newsvendor-probleem.
Bron: SWI-team the Driving Force*

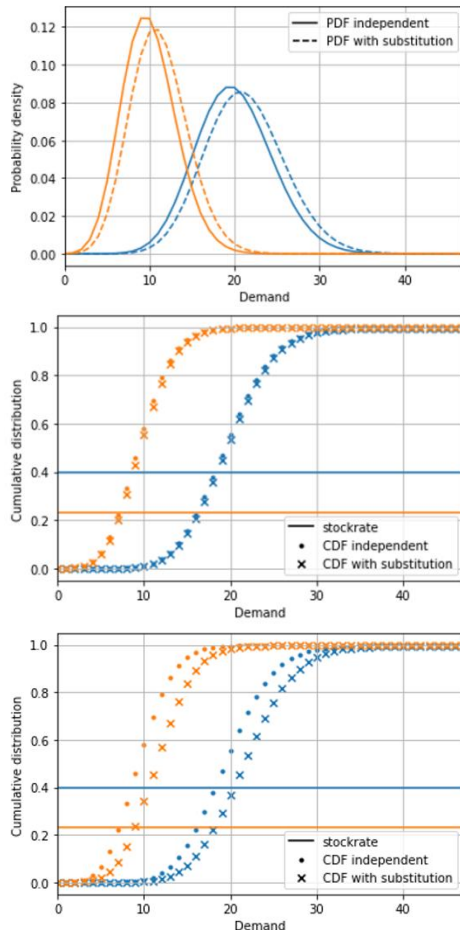
Tweede keus

We hebben het over de wiskunde van verse, bederfelijke waar. Hoeveel voorraad moet je als klein bedrijf maken of inslaan zodat je je klanten niet teleurstelt, maar ook niet met onverkoopbare producten blijft zitten? Dat *Newsvendor Problem* (krantenverkoper-probleem) lijkt belegen - de eerste formulering stamt tenslotte uit 1888 en de moderne oplossing uit de jaren 1950. Maar *Newsvendor* is onverminderd actueel.

Luke Visser: "Volgens Newsvendor hangt de hoeveelheid producten die je in moet slaan af van je marges. Als x procent van je verkoopprijs winst is,



Voor SWI heeft The Driving Force als inspiratie een Limburgse bakker genomen waar ze al mee samenwerken. Maar een uitspraak over verse voorraden is breder toepasbaar, denkt Frank Ottenhof: "Kleine bedrijven hebben anders dan grote concerns geen gespecialiseerde afdelingen die zulke tools zelf kunnen maken. Terwijl de sector versproducten enorm is. Met een algemeen inzetbare tool waar die bedrijven allemaal iets aan hebben, kunnen we een flink verschil maken."



*Uitgebreide versie van de Newsvendor-aanpak (boven). Afhankelijk van hoe bereid klanten zijn om een ander product te accepteren, adviseert de berekening om meer of minder voorraad in te slaan.
Bron: SWI-team the Driving Force*

Kansenspel

Elke bakkersklant is anders; de eerste neemt alleen genoeg met aardbeinvlaai, de volgende wil voor een verjaardagsfeest met drie verschillende smaken naar huis, voor ieder wat wils. Dat is met de exacte formules van het Newsvendor-probleem maar lastig te beschrijven, stelt Visser. Een statistische aanpak biedt meer mogelijkheden.

In hun kansenspel werken de onderzoekers alle mogelijke klantgedrag uit, en geven ze verhoudingen voor hoe vaak elke soort klant langskomt in de winkel. Dat palet aan klanten laten ze vervolgens los op een klein aantal verschillende vlaaien. Door alle mogelijke bestellingen af te gaan, ontstaat een 'hittekaart' van welke producten het snelst opraken. De aanpak blijkt krachtig, maar vraagt wel redelijk wat reken capaciteit. Luke Visser: "Voor elke vlaai die je toevoegt, krijg je een extra dimensie in de berekening, oftewel extra opties en combinaties om af te gaan." Daardoor zijn er ook meer rekenstappen nodig om alle mogelijkheden af te gaan.

Een complete winkel met tientallen verschillende producten simuleren, zit er volgens de promovendus dan ook niet in met deze aanpak. Wél kun je zo inzicht krijgen in het effect van verschillende soorten bestellingen op je voorraad. "In feite bouw je intuïtie op, die je kunt inpassen in complexere modellen."

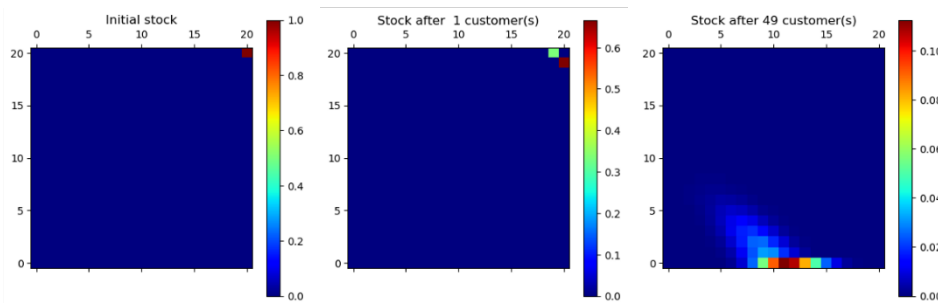
Uitgebreide Newsvendor-aanpak

Zoals zoveel SWI-groepen probeert ook het vlaaienteam meerdere aanvalsroutes uit. Als eerste aanpak proberen ze de standaardoplossing van het Newsvendor-probleem aan te passen, zodat die meerdere producten aankan.

Het vraagstuk van klanten die bereid zijn om over te stappen op een alternatief product heet in de vakliteratuur het *substitutieprobleem*. TU/e-promovendus Luke Visser: "We vonden een publicatie die uitwerkt hoe je dat toepast voor het Newsvendor-probleem, maar die aanpak leverde een vrij ruwe schatting op. Die hebben we zelf weten te verbeteren."

De wiskundigen weten het substitutieprobleem zo in formules te gieten, dat ze vlaaien waar klanten best tussen willen wisselen kunnen beschrijven als onafhankelijke producten. Die nieuwe methode voorspelt dat je best een paar procent meer voorraad in kunt slaan, als je weet dat klanten niet vastzitten aan hun eerste keus.

Het substitutieprobleem is daarmee deels opgelost, maar de andere delen van het vraagstuk "zijn te taai om op deze manier aan te pakken", vertellen de wiskundigen in hun eindpresentatie. Om algemenere voorspellingen te doen over vlaai-verkoop, grijpen ze naar simulaties.



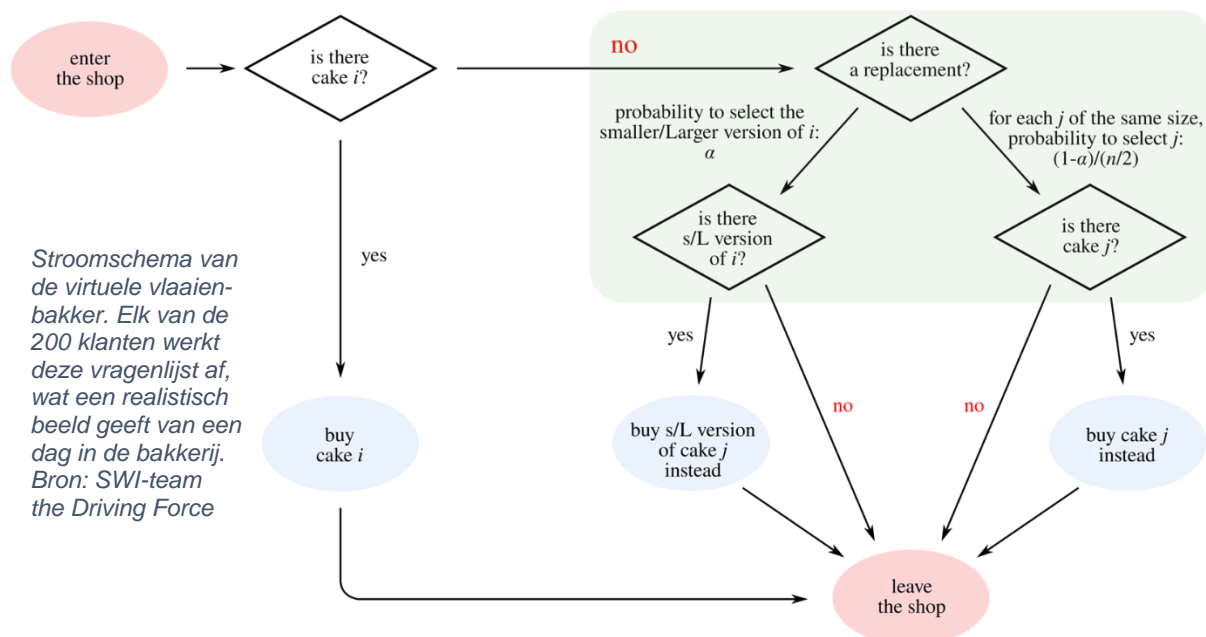
Statistische aanpak. Elke rekenstap bepaalt de kans dat een klant de voorraad van een bepaalde soort vlaai opmaakt. Bron: SWI-team the Driving Force

Virtuele vlaaienbakker

Kan een computermodel een realistische vlaaienbakker nabootsen, die in een dag honderden klanten langs krijgt die kunnen kiezen uit een waaier van producten? Niet als dat model alle opties af wil gaan - dat vraagt teveel rekentijd. Een minder nauwkeurige aanpak is wel mogelijk, besluiten de onderzoekers.

De wiskundigen bouwen een simulatie van een winkel met vijf verschillende soorten vlaaien, elk in een grote en kleine uitvoering. Voor elk van de 200 virtuele klanten die langs zullen komen wordt een willekeurig verlanglijstje gegenereerd aan de hand van een keuzediagram. Wil de klant 1, 2 of 3 vlaaien? Van welk formaat? Wat is de kans dat de klant voor een tweede keus gaat als de eerste op is, en wat is het favoriete alternatief?

De virtuele vlaaienbakker levert een gedetailleerd overzicht van hoe gevarieerd bestellingen kunnen worden. Om de simulatie nog realistischer te maken, zouden de wiskundigen de simulatie uit willen breiden met informatie over welke soorten vlaaien het populairst zijn. In de SWI-week is daar helaas te weinig tijd voor. Wel kunnen ze al berekenen wat het effect is op de totale winst als de bakker 's ochtends meer vlaaien maakt.

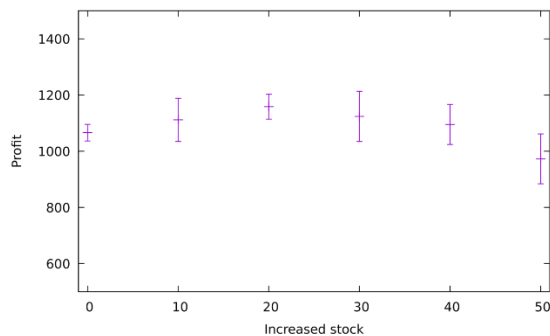


Paar procent

Namens The Driving Force reageert Joris Slootweg enthousiast op het resultaat van de SWI-week. Slootweg is zelf wiskundige en promoveert bij het CWI in Amsterdam op het werk dat hij voor The Driving Force doet, aan slimme algoritmes om infrastructuur beter te gebruiken. Dat loopt uiteen van optimaliseren van de wachtrijen op Schiphol tot toerisme spreiden over attracties en wegwerkzaamheden plannen voor Rijkswaterstaat.

Slootweg: "We hebben van de SWI-ers geen wereldschokkende nieuwe verkoopstrategieën gekregen, maar die hadden we ook niet verwacht." De onderzoeker en ondernemer ziet alsnog mooie uitkomsten. "De hittekaart-aanpak had ik bijvoorbeeld nog niet eerder gezien." Dat de modellen voorspellen dat je meer in kunt slaan dan de klassieke Newsvendor-oplossing, ziet Slootweg als winst: "Dat is niet heel intuïtief, maar het is heel mooi om te zien dat ondernemers zelfs nog iets ruimer kunnen zitten!"

Ook directeur Frank Ottenhof van The Driving Force is in zijn nopjes over het resultaat van de SWI-casus. "Je ziet hier hoe zelfs 'eenvoudige' vragen al snel te ingewikkeld worden voor onze intuïtie. Dat is wat mij betreft de kracht van wiskunde, dat die je helpt om situaties nauwgezet door te werken."



Analyse van de invloed van de voorraad vlaai op de totale winst. Te weinig voorraad, en de bakker laat verkoopkansen liggen; teveel, en er blijft te vaak vlaai over.

Bron: SWI-team the Driving Force

Van gedachten veranderd

De weg vooruit is volgens Ottenhof om de rekenmodellen van SWI verder te ontwikkelen tot tools waar het MKB zijn voorraden mee door kan rekenen. Hij geeft zelfs toe dat hij door de SWI-ers heeft besloten om een eerder advies aan zijn klanten om te gooien.

"Onze bedrijven en de wiskundigen lopen er tegenaan dat ze niet in het hoofd van klanten konden kijken. Grote concerns komen aan die informatie door hun kassasysteem. Dat houdt de aankopen van miljoenen klanten bij. In kleinere winkels kun je dat niet; iemand achter de toonbank hoort natuurlijk wel dat de klant liever een vlaai met gemengd fruit had gewild, maar dat komt niet terecht op de kassabon."

De oplossing: een aparte kassa-app bouwen die op een tablet draait, zodat de gegevens later automatisch geanalyseerd kunnen worden. Ottenhof: "Daar wilden we oorspronkelijk niet aan, want kassa-apps zijn er al zoveel. Maar door SWI zien we dat we met die apps betere voorspellingen kunnen maken van wat klanten willen. Aan de slag dus!"

Groendus

Schuiven met laadcapaciteit

Adviesbureau Groendus helpt bedrijven om hun energiehuishouding te verduurzamen. Bijvoorbeeld met laadpalen voor elektrisch rijdende medewerkers en klanten, gevoed door de eigen zonnepanelen van het bedrijf. Maar hoe plan je eigenlijk het optimale schema om al die accu's te laden?

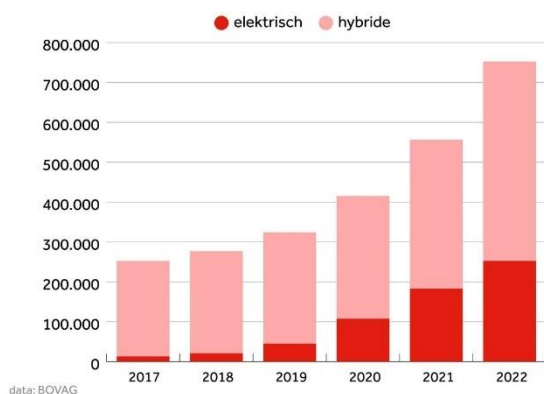
Over op duurzame energie om klimaatverandering af te remmen? Dan moeten we over op elektrisch rijden. Europese regeringen spraken daarom af dat er vanaf 2035 geen nieuwe fossiele brandstofauto's meer mogen worden verkocht. Volgens het Centraal Bureau voor de Statistiek telde ons land begin 2023 al 511.000 elektrische auto's; 6% van het totaal. In het bedrijfsleven ligt dat percentage fors hoger. Een kwart van de zakelijke personenauto's is elektrisch, de helft van het totale aantal stekkerauto's.



Elektrische laadpaal aan huis. Bron: Vattenfall

Bedrijven zoeken daarom naar manieren om hun elektrisch rijdende klanten en medewerkers van stroom te voorzien. Zonnepanelen op het dak, laadpalen op de parkeerplaats, en iedereen kan aan het eind van de werkdag zorgeloos met een volle accu naar huis.

Elektrisch laden op het parkeerterrein van je bedrijf klinkt ideaal voor klanten en medewerkers, maar blijkt een lastige puzzel. Want wanneer begin je met een auto opladen - zo gauw die inpluigt, of pas als de zonnepanelen veel vermogen leveren? Vroeg laden betekent dure netstroom inkopen, maar met te lang wachten op zon loop je het risico dat auto's niet genoeg lading krijgen om naar huis te komen.



*Verkoop van elektrische en hybride auto's.
Bron: NOS. Data: BOVAG*

Energievraag balanceren

Om zijn laadschema te bepalen, werkt adviesbureau Groendus momenteel met een eenvoudig algoritme dat de prijs van netstroom en zonne-productie vergelijkt. Uitgangspunt is dat auto's zo vol worden geladen, dat de bestuurder weer thuis kan komen. Het algoritme is slim genoeg om stroom terug te leveren aan het net als er meer wordt opgewekt dan verbruikt, maar kan de energievraag gedurende de dag niet overzien en balanceren; soms wordt daarom dure netstroom gebruikt om auto's te laden, terwijl later op de dag zonnestroom over is.

De vraag aan SWI: ontwerp een beter algoritme, dat aan de hand van voorspellingen over de zonne-productie en energieverbruik van de volgende dag een optimaal laadplan uitwerkt.

Nogal een vraagstuk, vonden wiskunde-promovenda Evie Nielen (Technische Universiteit Eindhoven) en haar teamgenoten aan het begin van de SWI-week: "Ik heb best wel eens gedacht: 'dit gaat niet werken'". Aan het eind van de week is die mening gedraaid: "Toen hadden we een best solide aanpak, waar Groendus ook direct van zei: 'dit is nuttig!' Een mooi resultaat dus."

Nielen promoveert in het dagelijks leven op het onderwerp *model order reduction*, oftewel "goedkoop afschatten van dure vergelijkingen". Door slimme versimpelingen kunnen vraagstukken die veel reken-

kracht kosten sneller en goedkoper worden opgelost. SWI leek haar een mooie kans om een week aan heel andere onderwerpen te werken, en te ontdekken wat voor vraagstukken je als wiskundige in het bedrijfsleven tegen kunt komen.

Splitsen

"We kwamen er al snel achter dat dit vraagstuk te complex is om in een week tijd op te lossen", herinnert Nielen zich. Daar maken de onderzoekers snel korte metten mee: als het vraagstuk te groot is, moet je het splitsen in onderdelen die wél behapbaar zijn.

De wiskundigen besluiten een knip te maken tussen het voorspellen van de toekomstige energievraag en -verbruik, en de optimalisatieslag om die voorspelde auto's met de beschikbare energie zo goed en goedkoop mogelijk te laden. Dat laatste vraagstuk pakken ze in twee stappen aan: eerste de ideale oplossing, dan een realistische.

$$\min_{E^+, E^- \in \mathbb{R}_{\geq 0}^T} \sum_{t=1}^T \underbrace{E_t^+ p_t^+}_{\text{buy energy}} - \underbrace{E_t^- p_t^-}_{\text{sell energy}} + \text{"penalty term"}$$

Formule om het laden van elektrische auto's te optimaliseren. De strafterm stimuleert het algoritme om duurzaam opgewekte stroom zoveel mogelijk te gebruiken om auto's op te laden, in plaats van die terug te verkopen aan het stroomnetwerk. Bron: SWI-team Groendus

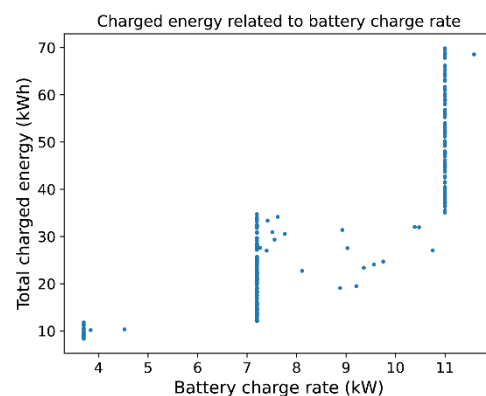
Die 'ideale aanpak voor het stekkerautoprobleem waar Nielen en collega's op uitkomen, stoelt op de techniek van *Mixed Integer Linear Programming* (MILP). Dat is een bekende optimalisatietechniek in de wiskunde. Door alle aspecten van het vraagstuk in vergelijkingen te gieten, kan een automatische *solver* de optimale oplossing uitrekenen waarbij auto's zo vol mogelijk worden geladen tegen de laagste kosten.

MILP is een rekenkundig dure techniek. Er gaat relatief veel rekentijd in zitten, omdat allerlei verschillende opties in details worden nagegaan. Dat vraagt een bijna pietepouterig exacte formulering van het vraagstuk in de vorm van formules, legt Nielen uit. Hoeveel opladers er zijn, het aantal tijdstappen en de hoeveelheid zonnestroom per tijdstip, de kosten en winst om stroom uit of in het stroomnet te sturen, zelfs de maximale laadsnelheden van accu's gaan in het model.

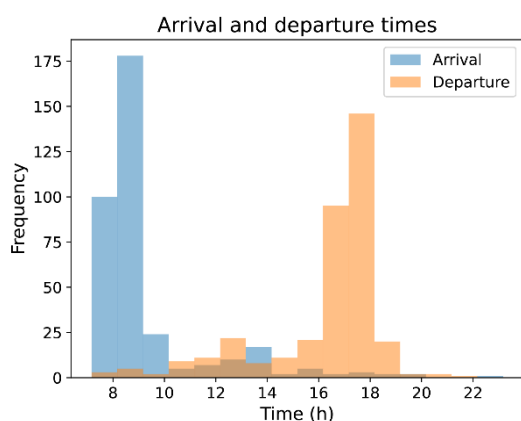
De MILP-techniek vraagt ook zo compleet mogelijke informatie: alsof je aan het begin van de dag exact weet wanneer welke auto binnenkomt en hoeveel stroom de zonnepanelen elk moment opleveren. Weinig realistisch, maar een goed startpunt, zo blijkt.

Met hun MILP-aanpak weten de wiskundigen laad-schema's uit te rekenen die het allerlaatste uit de beschikbare zonnestroom halen. Dat doen ze door hun algoritme strafpunten te geven als het zonnestroom verkoopt terwijl het op een ander moment op de dag energie uit het stroomnet haalt.

Door die strafterm aan te passen, blijken de wiskundigen te kunnen kiezen hoeveel auto's er aan het eind van de dag genoeg lading hebben om naar huis te komen. De grote vraag blijft: werkt deze aanpak ook als je geen perfecte voorkennis hebt over wat de dag gaat brengen?



De MILP-aanpak vraagt veel informatie, zoals hoeveel capaciteit de batterij van elke laadauto heeft en hoe vol die aankomt op het parkeerterrein. Uit een literatuurstudie komen drie groottes en bijbehorende lading naar voren. Bron: SWI-team Groendus



Aankomst- en vertrektijden van elektrische auto's. Bron: SWI-team Groendus

"In ons vraagstuk zijn er twee soorten stroomgebruikers", legt Evie Nielen uit: "allereerst de medewerkers van het bedrijf, die aan het begin van de dag aankomen en inpluggen en aan het eind van de werkdag weer vertrekken. Daarnaast zijn er gasten en klanten die in principe de hele dag door aan kunnen komen."

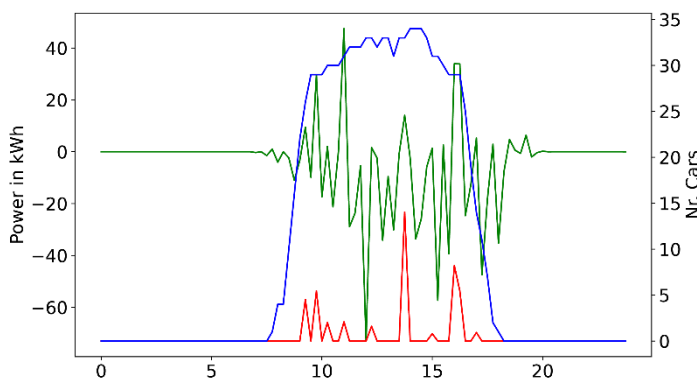
Om grip te krijgen op de aankomsttijden van stekkerauto's grijpen de wiskundigen naar statistiek. In hun aanpak zijn werknemers er niet exact van negen tot vijf, maar verspreiden ze zich in een klokvormige *normaalverdeling* rondom die tijdstippen. Voor klanten is dat anders: die kunnen in principe de hele dag door aanwaaien.

Zon en ziekmelden

Kunnen de voorspellingen scherper, en kan het MILP-model omgaan met beperkte voorkennis? In hun tweede onderzoeksfase bekijken Nielen en haar collega's een variant op hun MILP-aanpak. Die wordt niet eens per dag doorgerekend, maar krijgt bij elke tijdsstap aangescherpte informatie over bijvoorbeeld de verwachte zonnestroom en de hoeveelheid stekkerauto's op de parkeerplaats.

"Een van de aspecten waar we naar kijken is de kans dat werknemers niet naar hun werk komen", vertelt Nielen. Dat kan allerlei oorzaken hebben; ziekte bijvoorbeeld, een thuiswerkdag, of een afspraak buiten de deur. "We hebben een model met kansberekening opgezet op basis van de gemiddelde aankomsttijd van een werknemer. Als die een uur later nog niet binnen is, kun je daar een kans uit halen dat die persoon helemaal niet meer komt; dat scheelt laadcapaciteit die dan vrijkomt voor andere doeleinden."

Of die informatie niet op een andere manier te vinden is? "Je zou dit natuurlijk uit het ziekmeldsysteem kunnen halen", lacht Nielen, "maar of dat zo legaal is?" Een realistischer alternatief is een ticketsysteem waarmee iemand laadcapaciteit reserveert. Dat vraagt wel een dashboardfunctie waarmee het MILP-model makkelijk kan worden bediend door bedrijven, geeft de wiskundige toe: "die mooie strik zit er nu nog niet om, we hebben deze week echt nodig gehad om de wiskundige functionaliteit op te zetten. Maar in de toekomst is dat zeker een optie."



Uitkomst van het rekenmodel. Blauw: aanwezige elektrische auto's. Groen: energieoverschot. Rood: actieve laadpalen. Bron: SWI-team Groendus

Dicht op de toepassing

Hoe kijkt Nielen terug op de drukke SWI-week? "Het was aanpoten, dat zeker! We hebben veel, hard en lang gewerkt met ons groepje, echt een hogedrukpan. Dat was ook meteen een mooi verschil met de normale manier van werken als onderzoeker. Soms wel wat chaotisch - iedereen wil de diepte in, alles begrijpen, maar je moet ook overleggen om samen verder te komen.

"Tijdens de week werd duidelijk dat ons project op zijn pootjes ging landen. Erg mooi om zo snel resultaat te boeken en zo een bijdrage te leveren aan een probleem uit de praktijk. Als wiskundig onderzoeker zit je soms erg theoretisch te werken. In dit project hadden we een vraagstuk waarmee we juist direct op de toepassing zaten. Ook dat was een interessante ervaring. SWI is echt een aanrader om aan mee te doen."

KLM

"Wij laten de data zichzelf indelen"

Wanneer boeken passagiers hun vliegticket? Luchtvaartmaatschappij KLM gebruikt gedetailleerde boekingscurves om dat te voorspellen. Die curves zijn cruciale informatie om de ticketprijs goed te bepalen. Aan SWI de uitdaging om de boekingscurves nog beter te berekenen.

Als passagiers vinden we wat ruimte om ons heen wel prettig, maar luchtvaartmaatschappijen verkopen het liefst alle beschikbare stoelen op een vlucht. Om die verkoop te voorspellen gebruikt KLM zogeheten boekingscurves, net zoals het hotelwezen. Die geven aan hoe ver voor een vlucht tickets typisch geboekt worden: stapsgewijs bijvoorbeeld, of in één sprint vlak voor de vlucht.

Boekingscurves zijn cruciale informatie voor KLM om de ticketprijs aan te passen aan vraag en aanbod, en zoveel mogelijk te verdienen aan elke reis. Ken je klanten en hoe ze boeken, en je kunt er je ticketprijs op afstemmen. De vraag aan de Studiegroep Wiskunde en de Industrie (SWI): kunnen de wiskundigen die boekingscurves beter bepalen dan KLM zelf?

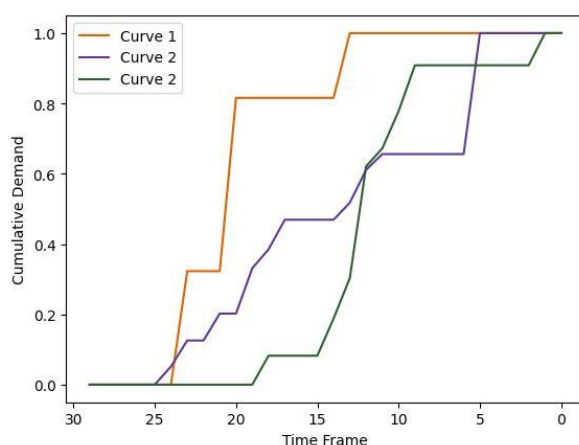
Bijdragen

"Van alle cases in de workshop sprong deze er voor me uit. Hier kon ik echt betekenisvol iets aan bijdragen", vertelt derdejaars promovendus statistiek Tyron Lardy van het Centrum voor Wiskunde en Informatica en de Universiteit Leiden.



KLM Cityhopper. Bron: Wayne Jackson / Pexels

De statisticus kwam terecht in een divers team: "Iedereen had een iets andere achtergrond en opleiding, en dus een eigen manier om over dingen na te denken. Dat was heel nuttig, want daardoor begreep iedereen net een ander deel van het vraagstuk."



Voorbeeld van boekingscurves. Bron: KLM

boomdiagram op van steeds verder toegespitste resultaten. De uiteindelijke boekingscurve van een nieuwe vlucht wordt dus bepaald door historische vluchten met dezelfde kenmerken.

Beslissingsboom

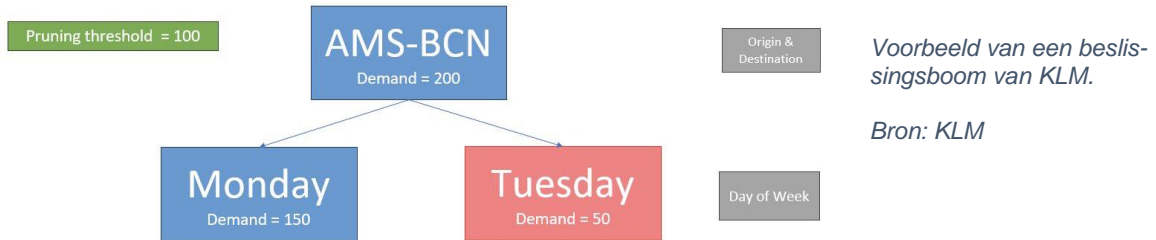
Tickets kun je bij KLM tot een jaar voor de vlucht boeken. De kunst voor het bedrijf is om te voorspellen wanneer klanten hun tickets zullen boeken. Dat gebeurt met een 'beslissingsboom' (*decision tree*), een serie opeenvolgende ja/nee-vragen over de vlucht.

In de beslissingsboom spelen tientallen factoren mee. Het begin- en eindpunt van de reis bijvoorbeeld; het tijdstip, de dag in de week en het jaargetijde; en of we zakelijk of privé reizen. De truc van een goede beslissingsboom is de juiste vragen in de juiste volgorde stellen. Dat levert een

Resultaten uit het verleden

De SWI-ers krijgen van KLM een database met twee jaar aan boekingen en de bijbehorende informatie zoals routes, data en tijden. Alle achttien factoren in de database beïnvloeden in principe de vorm van de boekingcurve. De vraag is: hoe precies? En blijft die relatie altijd zo, of verandert die in de loop der tijd?

Lardy: "Beslissingsbomen zijn een krachtige methode om data op te splitsen in categorieën, maar KLM gebruikt nu een vaststaande serie vragen, die zijn bedacht door een menselijke specialist. In de statistiek weten we dat resultaten uit het verleden geen garantie bieden voor de toekomst. We vliegen nu tenslotte niet meer zoals twintig jaar geleden." Tijd dus voor een andere aanpak dan een vaststaande vragenlijst.



Beslissingsbomen en regressie

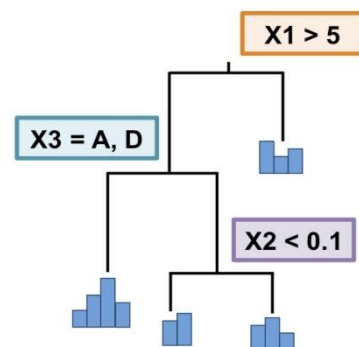
In Lardy's team liepen de meningen uiteen over de beste aanpak. Daarom zijn twee verschillende methodes gebruikt. Eén team dook in de lineaire regressie: een manier om lineaire relaties te vinden in immense databases. In principe kun je daarmee opsporen welke factoren de vorm van een boekingcurve bepalen.

"In de praktijk betekent dat, dat je enorme tabellen (*matrices*) opstelt met alle mogelijke relaties tussen de eigenschappen van de reis en wanneer er geboekt wordt", legt Lardy uit. "Die analyseer je tot je een lineaire afhankelijkheid hebt die de historische data het beste verklaart."

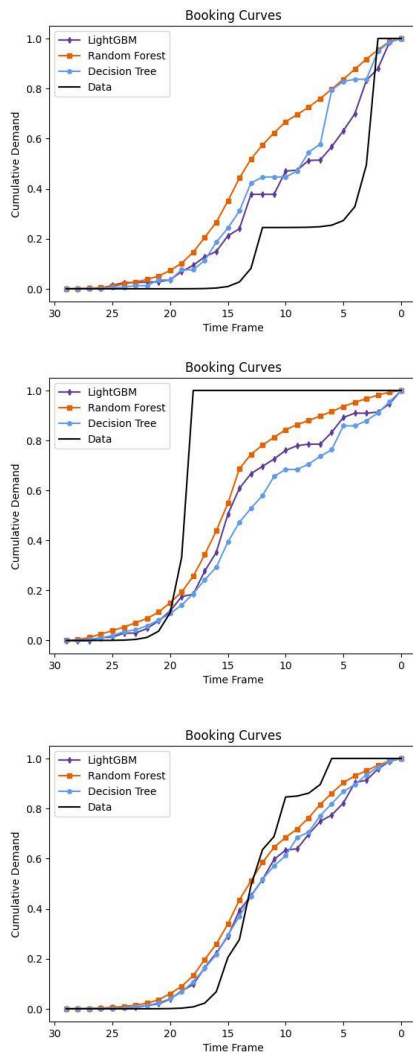
Lineaire regressie is een krachtige en snelle techniek, maar niet zomaar toe te passen op de KLM-casus. Zo is lineaire regressie bijvoorbeeld toepasselijk als de factoren binair zijn: business class of economy. Extra opties zoals economy+ in dat schema stoppen vraagt meer rekentijd om alle combinaties af te gaan. Daar komt bij dat het team een maatwerk algoritme schrijft, dat niet zo snel is als kant en klare varianten uit de literatuur. Uiteindelijk weten ze daarmee in de beperkte tijd drie van de parameters achter de boekingcurves door te rekenen. Te weinig om de vorm van de curves betrouwbaar te voorspellen. Tijd voor een snellere methode.

Zelflerend systeem

KLM heeft met hun beslissingsboom een geschikte aanpak gekozen, denkt Lardy. Maar in plaats van de onveranderlijke categorieën van de luchtvaartmaatschappij, willen de wiskundigen een dynamische aanpak, "waarbij we de data zichzelf in laten delen". Met hun database vol boekingen uit het verleden laten ze een computer steeds andere mogelijke onderverdelingen bedenken. Door automatisch te testen of zulke vragen ook boekingcurves opleveren die op elkaar lijken, komen ze met relatief weinig rekenwerk achter een handige manier om de boekingcurve van toekomstige vluchten te voorspellen.



Zelflerende beslissingsboom.
Bron: SWI-team KLM



Prestaties van drie verschillende technieken om beslissingsbomen te genereren voor de bookingcurves. Bron: SWI-team KLM

getraind op een andere subset van de data. Zo voorkom je dat je de bomen te veel toespitst op historische data en daarmee voorspellingskracht verliest."

Bedrijfsgeheim

In hun werkweek plotten de wetenschappers ook de prestaties van hun algoritmes, en vergelijken ze de berekende bookingcurves via verschillende foutenmarges met de historische reisgegevens. *Gradient boosting* haalt de laagste foutenmarges, al geeft het team ruitertlijk toe dat de drie methodes die ze hebben gekozen om de prestaties van de algoritmes te bepalen nog beter moeten worden toegespitst op de case van KLM. Dat werpt de vraag op: zijn de beslissingsbomen van SWI ook echt beter dan die van de luchtvaartmaatschappij?

De gradient-boosted analyse wijst de maand en dag van de week aan als belangrijkste factoren die de vorm van beslissingsbomen bepalen. Bron: SWI-team KLM

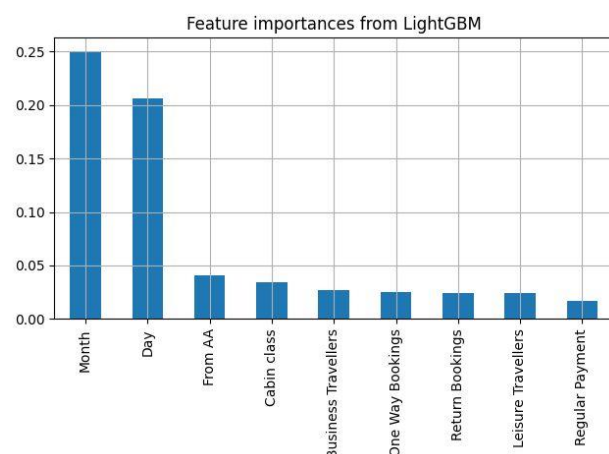
"Van de historische data leren we welke vragen je moet stellen om tot zinnige categorieën te komen", vertelt Lardy. "Als zo'n ja/nee-vraag een verzameling reizen oplevert die duidelijk bij elkaar horen omdat ze dezelfde vorm booking-curve hebben, ben je klaar. Zo niet, dan probeer je de reizen verder op te splitsen. Vervolgens kun je nieuwe routes door dezelfde vragenlijst halen. Het idee is dat ze dan ook dezelfde soort bookingcurve hebben als de reizen uit de trainingsdata."

Woud van beslissingsbomen

Om hun beslissingsbomen zo goed mogelijk te krijgen, halen de SWI-ers moderne statistiek uit de kast. Het 'standaardexemplaar' trainen ze op op twee jaar aan reisdata, met 18 parameters per booking. Dat levert een aardig eerste resultaat, al vallen de voorspelde bookingcurves nog lang niet perfect samen met de historische resultaten. Tijd voor een verfijnde aanpak.

De wiskundigen gebruiken twee technieken om hun beslissingsbomen aan te scherpen: *gradient boosting* en *random forests*. "Als je al je trainingsdata gebruikt om één beslissingsboom te genereren, geeft die niet per se de beste oplossing", legt Lardy uit. "Met *gradient boosting* zet je alternatieve bomen op, andere vragenseries, die je optimaliseert om goed te presteren waar je hoofdboom zwakke plekken heeft."

Dat samenspel van meerdere modellen levert betere resultaten, net als een *random forest*, de andere techniek die het team uitprobeert: "Dan zet je meerdere bomen op, elk



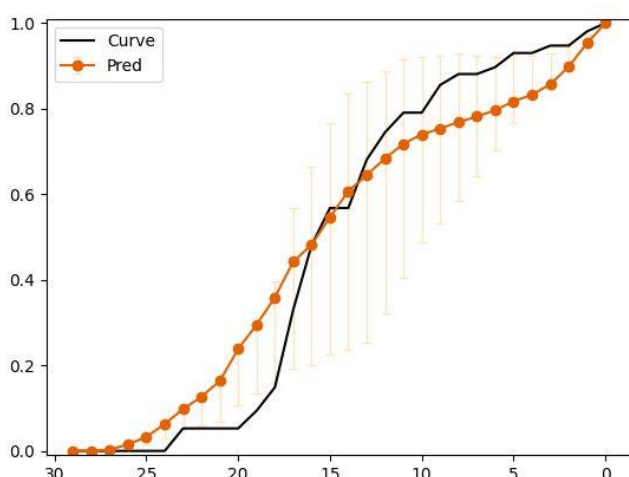
"Gek genoeg weten we niet precies hoe goed we het nou echt hebben gedaan", vertelt Tyron Lardy. Om dat te bepalen zou het team namelijk naar de bedrijfsalgoritmes en -data van KLM moeten kijken, en die zijn tijdens de workshop niet beschikbaar gemaakt. Wel geven de vertegenwoordigers van het bedrijf aan dat ze brood zien in de verschillende methodes, al is het bijvoorbeeld nog de vraag of hun ICT- en veiligheidsteams zomaar toestaan dat er gebruik wordt gemaakt van de open source rekenpakketten die de SWI-ers hebben ingezet.

"Die willen vast iets dat nét niet kan"

Hoe kijkt Tyron Lardy terug op de SWI-week? "Ik nam deel aan SWI om erachter te komen wat voor uitdagingen bedrijven hebben voor wiskundigen. Tegen het eind van je promotie vraag je je toch af: wat wil ik hierna met mijn leven - door in het onderzoek, of het bedrijfsleven in? SWI was een mooie manier om wat ervaring op te doen met werken in het bedrijfsleven."

Lardy vond de SWI-week een prettige verrassing: "Eerst dacht ik dat die grote bedrijven wel extreem geoliede machines zouden zijn, met enorm gespecialiseerde tools en vraagstukken - die vragen vast iets aan ons dat met bestaande technieken nét niet kan. Maar ze bleken met dezelfde tools te werken als wij. Dat is het mooie van wiskunde: of ik nou een boekingcurve wil voorspellen of iemands lengte, uiteindelijk zijn daar allerlei standaardtechnieken voor. De kunst is hoe je ze toepast.

"Ik zou zeker nog een keer meedoen aan SWI, als er opdrachten bij zitten die goed bij mijn achtergrond passen. Dan kun je echt een fantastische ervaring hebben."



Bepalen van de onzekerheidsmarge van een boekingcurve.

Bron: Zhang, Haozhe, et al. "Random forest prediction intervals." The American Statistician (2019).

