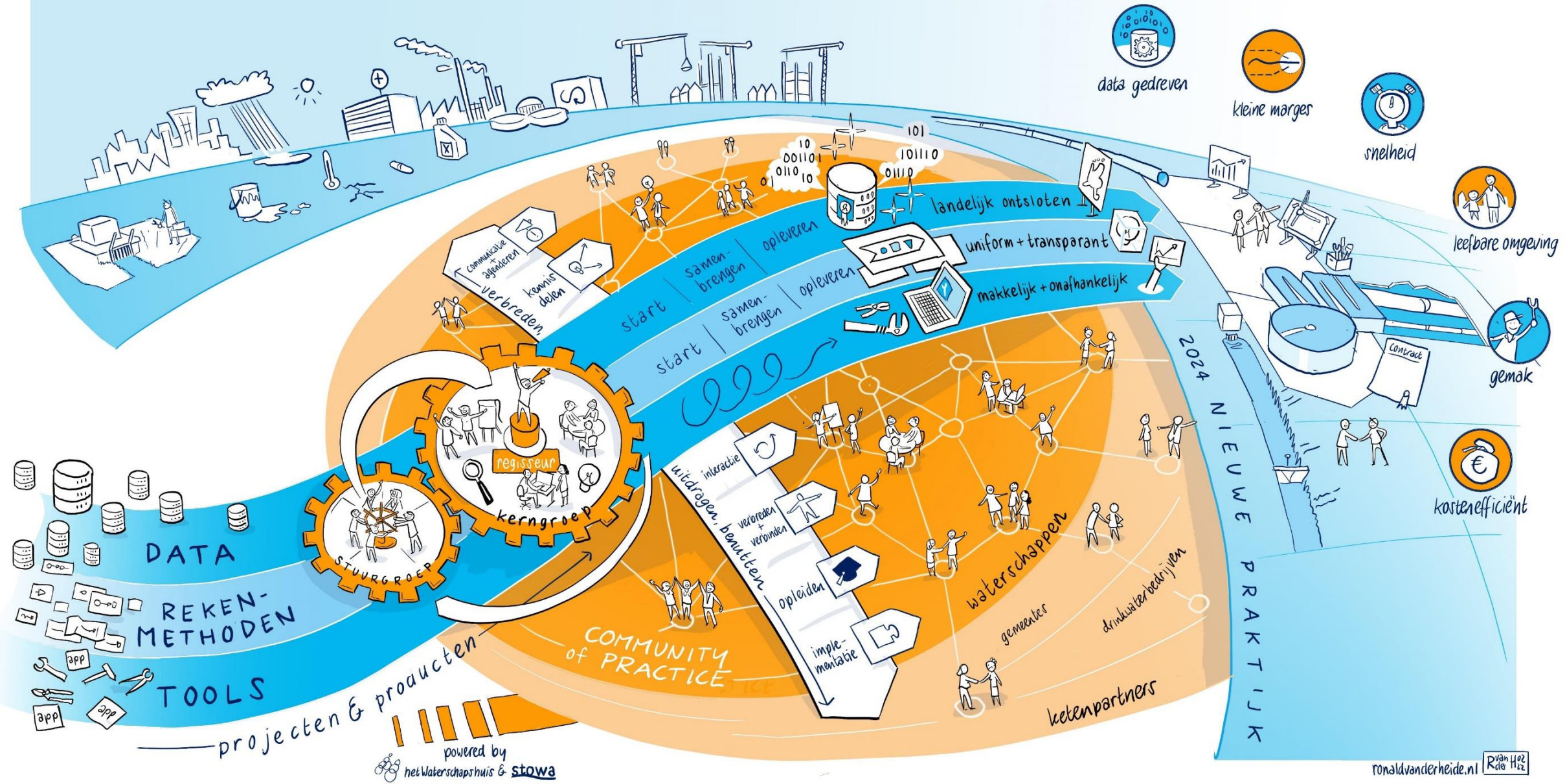


afvalwaterprognoses NAAR EEN NIEUWE PRAKTIJK

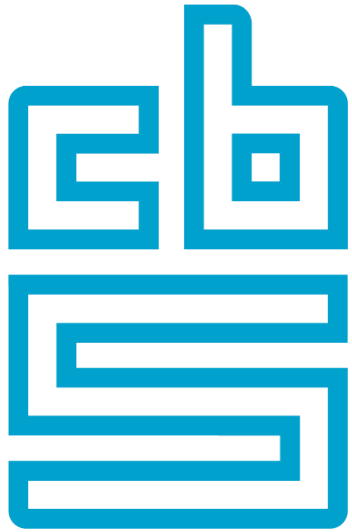


Webinar 3

'Datascience en verbeterde kengetallen'

In samenwerking met het CBS is er verkend wat de mogelijkheden zijn om afvalwatervoorspellingen te maken op perceelsniveau. De eerste resultaten worden gepresenteerd en er wordt een doorkijk gegeven naar de vervolgstappen.

Presentator: Jeldrik Bakker (CBS) – 14 juni 2023 11:00 - 11:30



WATERVERBRUIK PER ECONOMISCHE SECTOR

Jeldrik Bakker, Rob Willems, Yvonne
Gootzen, Marc Ponsen
2023-06-14

Pilot project CBS / Hollands Noorderkwartier / Waterschapshuis

Inhoud van de presentatie

- Introductie
- Probleemstelling en onderzoeksvragen
 - Per onderzoeksvraag:
 - Data (koppeling)
 - Methode
 - Resultaten en Conclusies
- Conclusies
- Toekomstig werk

Introductie

- Inzicht in de waterketen is cruciaal voor beleidsbepalers en de waterschappen. Huidige waterverbruik statistieken (van CBS en waterschappen) worden gemaakt op basis van vragenlijsten en domein expertise

Waterusers	Periods	Tap water			Groundwater			Surface water			Type bijzondere bebouwing	Maatgevende belasting (indicatief)	
		Total Tap water	Use of drinking water	Use of industrial water	Total use of groundwater	Use of groundwater for cooling	Other use of groundwater	Total use of surface water	Use of salt surface water	Use of fresh surface water			
mln m3													
A Agriculture, forestry and fishing	2017 ^o	48.4	48.4	0.0	97.0	0.0	97.0	22.7	22.7	0.0	Hotels	per bed	10 – 40 l/h
	2018 ^o	48.9	48.9	0.0	225.3	0.0	225.3	76.8	76.8	0.0	Restaurants	per werknemer	50 l/h
	2019 ^o	43.8	43.8	0.0	198.8	0.0	198.8	54.6	54.6	0.0	Cafés	per werknemer	25 l/h
B Mining and quarrying	2017 ^o	2.6	2.6	0.0	0.0	0.0	0.0	35.3	0.7	34.6	Laboratoria	per werknemer	25 l/h
	2018 ^o	2.9	2.9	0.0	0.1	0.0	0.0	36.2	0.8	35.4	Internaten	per bed	15 l/h
	2019 ^o	2.2	2.2	0.0	0.1	0.0	0.1	27.7	0.0	27.7	Ziekenhuizen	per patiënt	30 l/h
C Manufacturing	2017 ^o	205.9	136.2	69.7	130.2	60.1	70.1	2,999.1	2,390.7	608.4	Bejaardencentrum	per bewoners + personeel	15 l/h
	2018 ^o	202.4	132.0	70.4	132.3	62.7	69.5	2,980.0	2,326.1	653.9	Kazernes en gevangenissen	per bewoner	15 l/h
	2019 ^o	213.0	144.6	68.4	127.6	61.3	66.4	2,906.3	2,285.2	621.1	Scholen	per leerling	2 – 3 l/h
											'Droge' bedrijven en industrieën	per werknemer	6 l/h
											Recreatieparken of vakantiebungalows	per bewoner	10 l/h
											Campings, jachthavens etc.	per kampeerder	5 l/h
											Melk/(rundvee)houderij	per bedrijf	72 – 100 l/h
											Landbouwbedrijven	per bedrijf gelimiteerd op	3000 – 5000 l/h
											Glastuinbouw	per hectare	200 – 1200 l/h

Introductie

- De wens is om data-gedreven schattingen te maken van waterverbruik per economische sector op basis van o.a. sensorgegevens uit het Z-Info systeem. Z-Info is een database beheerd door waterschappen met daarin onder andere dag debieten (doorstromend water) gemeten bij rioolwaterzuiveringsinstallaties
- Vergeleken met statistieken op basis van experten of vragenlijsten, zijn statistieken op basis van sensorgegevens **in potentie**....
 - ... *Kosten-efficiënter*
 - ... *Actueler/gedetailleerder*
 - ... *Nauwkeuriger*

Introductie

Het doel van dit pilot project is dus een op Z-Info gebaseerde data-gedreven schatting van (delen van) onderstaande Statline tabel:

Waterusers <input type="text"/>		Tap water			Groundwater			Surface water		
Periods <input type="text"/>		Total Tap water	Use of drinking water	Use of industrial water	Total use of groundwater	Use of groundwater for cooling	Other use of groundwater	Total use of surface water	Use of salt surface water	Use of fresh surface water
		mln m ³								
A Agriculture, forestry and fishing	2017*	48.4	48.4	0.0	97.0	0.0	97.0	22.7	22.7	0.0
	2018*	48.9	48.9	0.0	225.3	0.0	225.3	76.8	76.8	0.0
	2019*	43.8	43.8	0.0	198.8	0.0	198.8	54.6	54.6	0.0
B Mining and quarrying	2017*	2.6	2.6	0.0	0.0	0.0	0.0	35.3	0.7	34.6
	2018*	2.9	2.9	0.0	0.1	0.0	0.0	36.2	0.8	35.4
	2019*	2.2	2.2	0.0	0.1	0.0	0.1	27.7	0.0	27.7
C Manufacturing	2017*	205.9	136.2	69.7	130.2	60.1	70.1	2,999.1	2,390.7	608.4
	2018*	202.4	132.0	70.4	132.3	62.7	69.5	2,980.0	2,326.1	653.9
	2019*	213.0	144.6	68.4	127.6	61.3	66.4	2,906.3	2,285.2	621.1

Probleemstelling en onderzoeksvragen

Probleemstelling:

Huidige waterverbruik statistieken (van CBS en waterschappen) zijn gebaseerd op vragenlijsten en domein expertise en zijn tijdrovend (en dus duur) om te maken. Via twee onderzoeksvragen zoeken we naar een oplossing voor dit probleem.

Onderzoeksvraag 1:

Is het mogelijk waterverbruik door personen en bedrijven te schatten op basis van Z-Info meet gegevens?

Onderzoeksvraag 2:

Is het mogelijk waterverbruik toe te kennen aan economische activiteiten op basis van CBS bedrijven- en persoonsregisters?

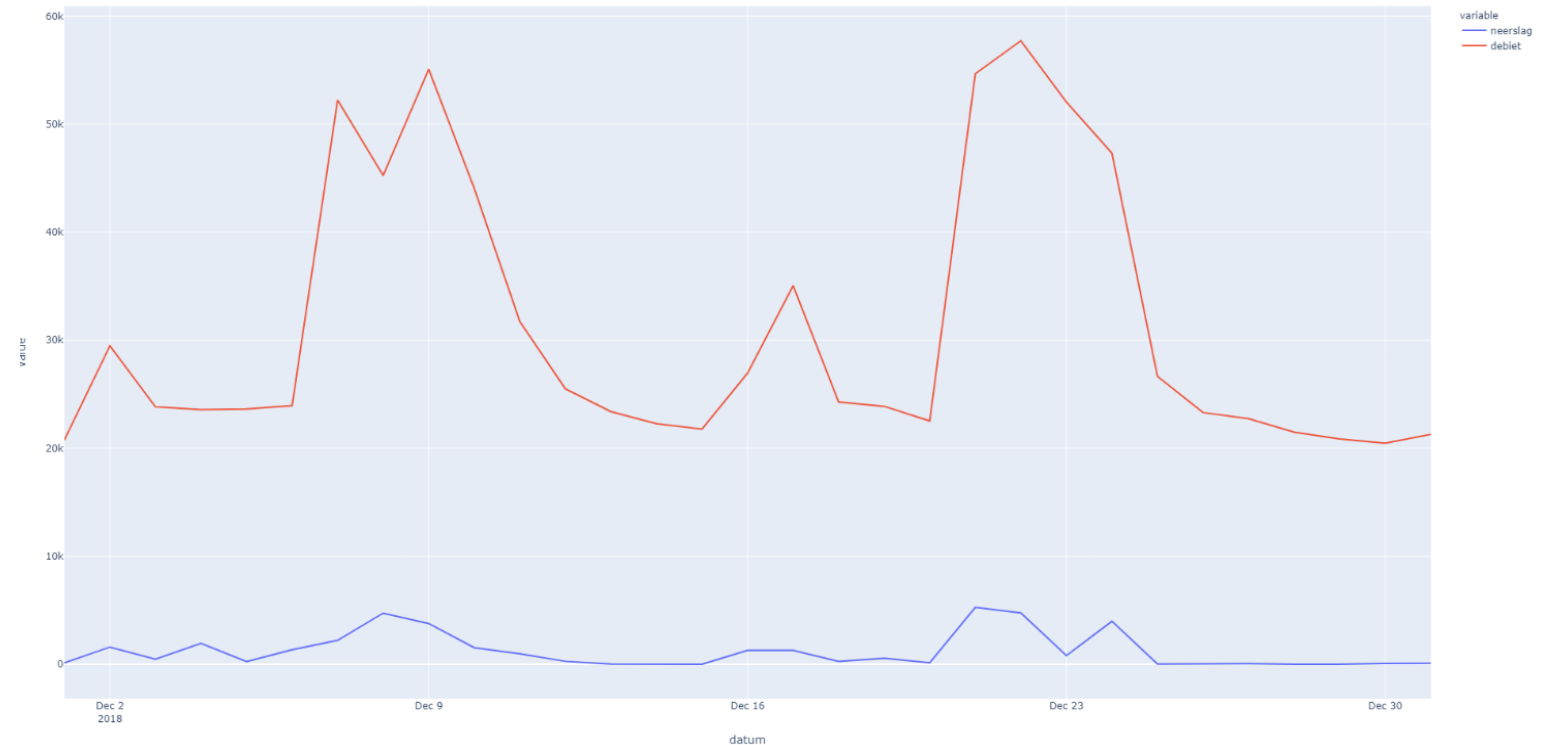
Onderzoeksvraag 1

Is het mogelijk waterverbruik (door personen en bedrijven) te schatten op basis van Z-Info meet gegevens?

Bedrijven en huishoudens zijn aangesloten op een zogenaamde rioolwaterzuiveringsinstallaties (RWZI), maar niet alleen diens waterverbruik komt terecht in de riolen, maar ook bijvoorbeeld regenwater. We schatten een gecorrigeerd dag debiet op basis van neerslaggegevens van het KNMI



z-info Amstelveen december 2018



Data | Z-Info

- Centraal systeem voor informatie van de afvalwaterzuivering van negentien van de eenentwintig waterschappen in Nederland ([Z-info | Het Waterschapshuis](#))
- Data beheerd door de waterschappen
- Data via API toegankelijk voor CBS

Datum	Installatie	Omschrijving intallatie	Deelproces	Omschrijving deelproces	Deelproces 2	Omschrijving deelproces 2	Product	Omschrijving product	Grootheid	Omschrijving grootheid	Parameter	Hoedanigheid	Eenheid	Compartment	Dagdebiet
1-1-2018	NL.07.ATN	RWZI Aalten	tITA	Transportlijn (influent)	wLON	Ontvangen	INFET	Influent	VOLME	Volume	NVT	NVT	m3	AW	16566
1-1-2018	NL.07.DPL	RWZI Dinxperlo	tITA	Transportlijn (influent)	wLON	Ontvangen wI	INFET	Influent	VOLME	Volume	NVT	NVT	m3	AW	10122
1-1-2018	NL.07.ETN	RWZI Etten	tITA	Transportlijn (influent)	wLON	Ontvangen wI	INFET	Influent	VOLME	Volume	NVT	NVT	m3	AW	78179
1-1-2018	NL.07.HTN	RWZI Holten	tITA	Transportlijn (influent)	wLON	Ontvangen	INFET	Influent	VOLME	Volume	NVT	NVT	m3	AW	21908
1-1-2018	NL.07.LTV	RWZI Lichtenvoorde	tITA	Transportlijn (influent)	wLON	Ontvangen	INFET	Influent	VOLME	Volume	NVT	NVT	m3	AW	15769
1-1-2018	NL.07.NGF	RWZI Nieuwgraaf	tITA	Transportlijn (influent)	wLON	Ontvangen	INFET	Influent	VOLME	Volume	NVT	NVT	m3	AW	165700
1-1-2018	NL.07.OBG	RWZI Olburgen	tITA	Transportlijn (influent)	wLON	Ontvangen	INFET	Influent	VOLME	Volume	NVT	NVT	m3	AW	64207
1-1-2018	NL.07.RLO	RWZI Ruurlo	tITA	Transportlijn (influent)	wLON	Ontvangen	INFET	Influent	VOLME	Volume	NVT	NVT	m3	AW	5231

Data | KNMI

- Het dataplatform van het Koninklijk Nederlands Meteorologisch Instituut (KNMI) biedt toegang tot data over weer, klimaat en seismologie (<https://www.daggegevens.knmi.nl/klimatologie/monv/reeksen>)
- Open data

Station nr	Station naam	Neerslag	Datum	Starttijd	Eindtijd	Hoogte	Positie X	Positie Y	Positie NB	Positie OL
235	Castricum	77	1-1-2018	1-7-1927	31-12-9999	3,74	104,9	506,5	52.5	4.6
217	Heiloo	94	1-1-2018	1-10-1943	31-12-9999	-0,29	110,6	511,6	52.5	4.7
264	Krommenie	35	1-2-2020	31-1-2020	31-12-9999	-1,00	112,4	503,3	52.5	4.7
234	Bergen NH	68	1-1-2018	18-6-1927	31-12-9999	9,20	107,5	521,5	52.6	4.6
226	Wijk aan Zee	103	1-1-2018	1-1-1910	31-12-9999	5,60	101,3	498,6	52.4	4.5
263	Assendelft	92	1-1-2018	1-3-2006	31-12-9999	-1,30	110,0	498,3	52.4	4.7
16	Petten	58	1-1-2018	1-4-1891	31-12-9999	0,40	105,4	530,6	52.7	4.6
435	Heemstede	133	1-1-2018	1-12-1866	31-12-9999	-4,50	105,0	482,9	52.3	4.6
233	Zaandam (Hembrug)	118	1-1-2018	13-10-1949	31-12-9999	-1,40	115,2	494,3	52.4	4.8
255	Obdam (nieuw)	73	1-1-2018	1-1-1951	31-12-9999	-2,00	120,3	520,9	52.6	4.8

Onderzoeksvraag 1 | Methode

Is het mogelijk waterverbruik door personen en bedrijven te schatten op basis van Z-Info meet gegevens?

- Generalized lineair model (GLM) die gecorrigeerde dag debieten schat voor een rwz_i^j op basis van gesommeerde neerslag over heel Nederland (op die dag en deels dag ervoor)
- Deze gecorrigeerde dag debieten gelden dan als input voor de volgende onderzoeksvraag

$$Dagdebiet_i \text{ in } m^3 \sim \beta_{0,i} + \beta_{1,i} \cdot Neerslag + \beta_{2,i} \cdot Neerslag_vertraagd + \varepsilon_i$$

$$Dagdebiet_i \text{ gecorrigeerd} = Dagdebiet_i - \beta_{1,i} \cdot Neerslag - \beta_{2,i} \cdot Neerslag_vertraagd$$

Onderzoeksvraag 1 | Resultaten en Conclusies

Is het mogelijk waterverbruik door personen en bedrijven te schatten op basis van Z-Info meet gegevens?



Onderzoeksvraag 2

Is het mogelijk waterverbruik toe te kennen aan economische activiteiten op basis van CBS bedrijven- en persoonsregisters?

In onderzoeksvraag 1 berekenden we een gecorrigeerd dag debiet op basis van neerslaggegevens van het KNMI, dit als schatter voor “waterverbruik door bedrijven en personen”. Maar welk aandeel hebben de verschillende economische activiteiten in dat waterverbruik?

- CBS heeft onlangs intern een koppeling gemaakt tussen het Algemeen Bedrijvenregister (ABR) en Basisregistratie Adressen en Gebouwen (BAG). Deze nieuw gekoppelde dataset (ABR + BAG) maakt het mogelijk economische activiteiten (geografisch) te koppelen aan de berekende waterverbruik schattingen
- In dit pilot project kijken we naar de meest grove SBI classificaties, omdat die overeenkomen met die in Statline.

Data | BAG

- De Basisregistratie Adressen en Gebouwen (BAG) is onderdeel van het overheidsstelsel van basisregistraties. Gemeenten zijn bronhouders van de BAG. Zij zijn verantwoordelijk voor het opnemen van de gegevens in de BAG
- Kadaster beheert de BAG en stelt gegevens beschikbaar aan diverse afnemers, waaronder het CBS.



Bag ID	Postcode	Huisnr	Huisnr Toevoeging	Woonfunctie	Kantoorfunctie	Logiesfunctie	Oppervlakte	bouwjaar	status	WOZ waarde object
B0983010000036751	5914XN	49		0	0	0	159	1938	in gebruik	345000
B1709010000008849	4791BH	92		0	1	0	345	1997	in gebruik	687000
B0166010000025769	8271RW	10	8	0	0	0	2244	1986	in gebruik	2187000
B0307010000443807	3811ND	41		0	0	1	40	2012	in gebruik	490000
B0736010000005195	3645AR	19 A 81A		0	0	0	121	2007	in gebruik	873000

Data | ABR + BAG

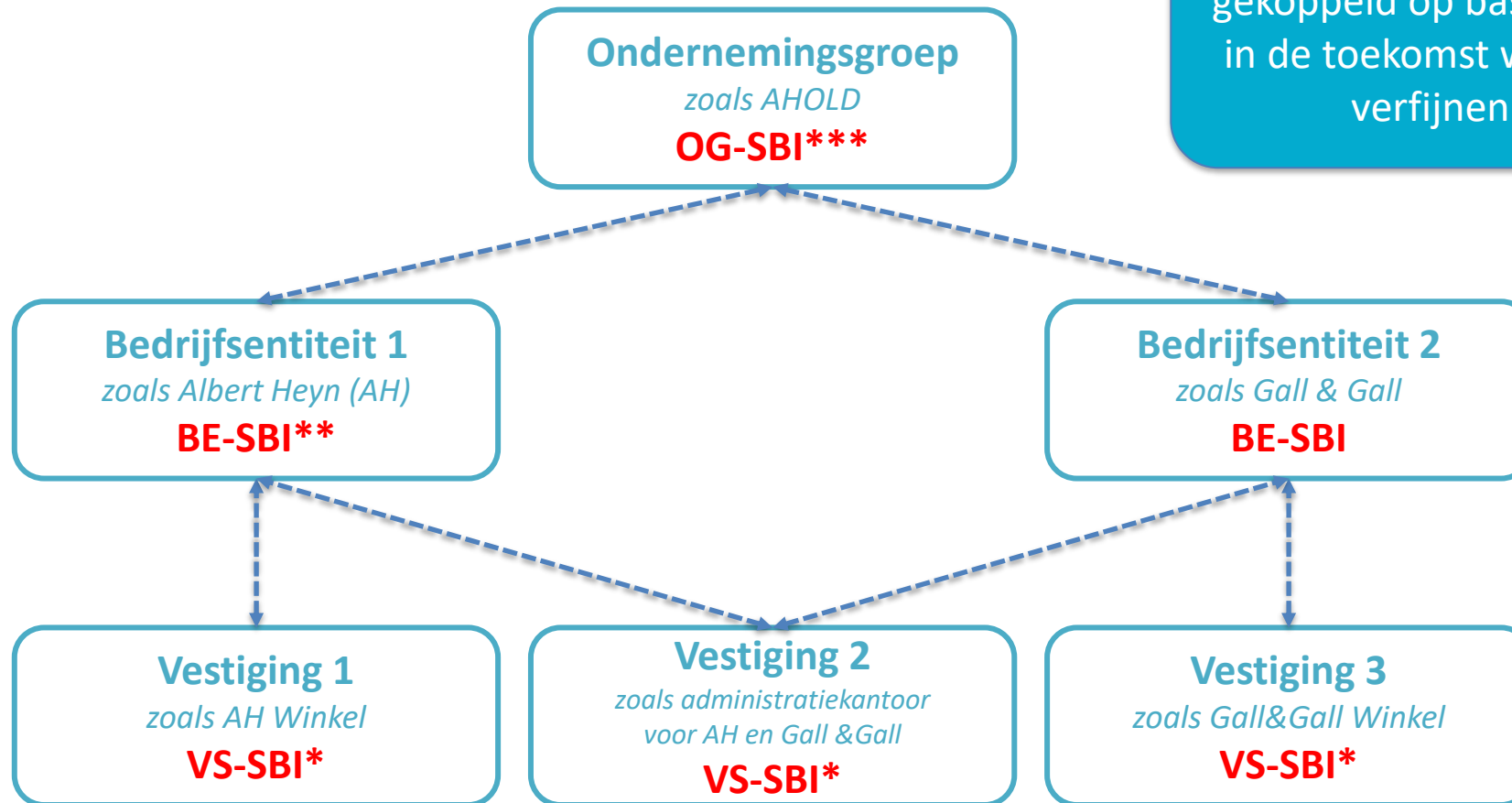
- In het Algemeen Bedrijvenregister (ABR) worden bedrijven en instellingen, met hun identificatie- en structuurgegevens, vastgesteld en geregistreerd in voor statistisch onderzoek geschikte eenheden. De gegevens worden uitgesplitst naar economische activiteit, gebaseerd op de Standaard Bedrijfsindeling 2008 (SBI)

[SBI codes · Centraal Bureau voor de Statistiek \(cbs.nl\)](#)

StandId	Bedrijfseenheid ID	SBI	Vestigings nummer	Postcode	Huisnr	Huisnr Toevoeging	BAG ID
1801	67350844	1512	32543255	5914XN	49		B0983010000036751
1801	55823904	53202	2383152	4791BH	92		B1709010000008849
1801	73876399	70222	34179828	8271RW	10	8	B0166010000025769
1801	53717139	6201	33530157	3811ND	41		B0307010000443807
1801	57678235	9609	5775337	3645AR	19	A 81A	B0736010000005195

Data | **ABR + BAG**

Op dit moment is de ABR + BAG data gekoppeld op basis van BE-SBIs, maar in de toekomst willen we de analyse verfijnen naar VS-SBIs



- * Op basis van kamer van koophandel inschrijving waarbij hoofd bedrijfsactiviteit gegeven moet worden
- ** BE-SBI bepaald op basis van VS-SBIs (op basis van bepaling van economisch zwaartepunt)
- *** OG-SBI bepaald op basis van BE-SBIs (op basis van bepaling van economisch zwaartepunt)

Data | BRP + BAG

- De Basisregistratie Personen (BRP) bevat persoonsgegevens van inwoners van Nederland (ingezetenen) en van personen die Nederland hebben verlaten (niet ingezetenen).

BAG ID (versleuteld)	Persoons ID	Drinkwaterbedrijf	Inkomen	Type woning	Gemeente	Stedelijkheid	Provincie	Opleidingsniveau	Aantal bewoners
000003cb8bb8066bfb2ed8079db7eaab	788705595	Dunea	87e perce	Huurwoning zonder huurtoeslag	855	Zeer sterk (>=2500 omgevingsadressen/km2)	Noord-Brabant	3	2
0000047333d4499957b11c30858a2c7f	488516883	Waternet	73e perce	Eigen woning	363	Zeer sterk (>=2500 omgevingsadressen/km2)	Noord-Holland		2
00001673d41cf7751dd40169cf95a979	54099568	Evides	69e perce	Eigen woning	867	Matig (1000 tot 1500 omgevingsadressen/km2)	Noord-Brabant	3	3
00001721c93bddd0134cd58d4b3a193f	213235124	Vitens Overijssel	68e perce	Eigen woning	1896	Weinig (500 tot 1000 omgevingsadressen/km2)	Overijssel		2
0000193c39581396705dffab444e35a7	236644306	Waternet	26e perce	Huurwoning zonder huurtoeslag	363	Zeer sterk (>=2500 omgevingsadressen/km2)	Noord-Holland		2
0000199a2fa71b69825ebfe4e621f24f	123475383	Dunea	93e perce	Eigen woning	1926	Sterk (1500 tot 2500 omgevingsadressen/km2)	Zuid-Holland	3	1

Data Koppeling | **ABR** | **BAG** | **BRP** | **Z-INFO (KNMI)**

Op basis van de RWZI indelingen uit 2020, heeft CBS bij alle unieke BAG ID's (sinds het begin der tijden) het desbetreffende RWZI gebied gekoppeld. Dit maakt verdere koppeling met CBS registers mogelijk

Z-INFO (KNMI)

Datum	Installatie	Omschrijving intallatie	Dagdebiet
1-1-2018	NL.07.ATN	RWZI Aalten	16566
1-1-2018	NL.07.DPL	RWZI Dinxperlo	10122
1-1-2018	NL.07.ETN	RWZI Etten	78179
1-1-2018	NL.07.HTN	RWZI Holten	21908
1-1-2018	NL.07.LTV	RWZI Lichtenvoorde	15769
1-1-2018	NL.07.NGF	RWZI Nieuwgraaf	165700
1-1-2018	NL.07.OBG	RWZI Olburgen	64207
1-1-2018	NL.07.BLO	RWZI Ruurlo	5231

BAG ID	x	y	RWZI code CBS	Naam	WS_zinfo
000003cb8bb8066bfb2ed8079db7eaab	134654	397036	32002	Schiphol(derden)	cbs.1
0000047333d4499957b11c30858a2c7f	128314	478937	11014	Bruggehof (particulier)	cbs.2
0000134a3c75547a6350f0c5434a7b36	126918	478781	11014	Stadskanaal	NL.33.SK
000015e051700259b7046d20ca738da6	201247	356176	30024	Ter Apel	NL.33.TA
00001673d41cf7751dd40169cf95a979	132898	410583	25003	Uithuizermeeden	NL.34.UM
00001721c93bdd0134cd58d4b3a193f	203539	511888	4005	Ulrum	NL.34.UR

Bag ID	Oppervlakte
B0983010000036751	159
B1709010000008849	345
B0166010000025769	2244
B0307010000443807	40
B0736010000005195	121

BAG

SBI	BAG ID
1512	B0983010000036751
53202	B1709010000008849
70222	B0166010000025769
6201	B0307010000443807
9609	B0736010000005195

ABR-BAG

Aantal bewoners	BAG ID
2	000003cb8bb8066bfb2ed8079db7eaab
2	0000047333d4499957b11c30858a2c7f
3	00001673d41cf7751dd40169cf95a979
2	00001721c93bdd0134cd58d4b3a193f
2	0000193c39581396705dffab444e35a7
1	0000199a2fa71b69825ebfe4e621f24f

BRP-BAG

Onderzoeksvraag 2 | Methode

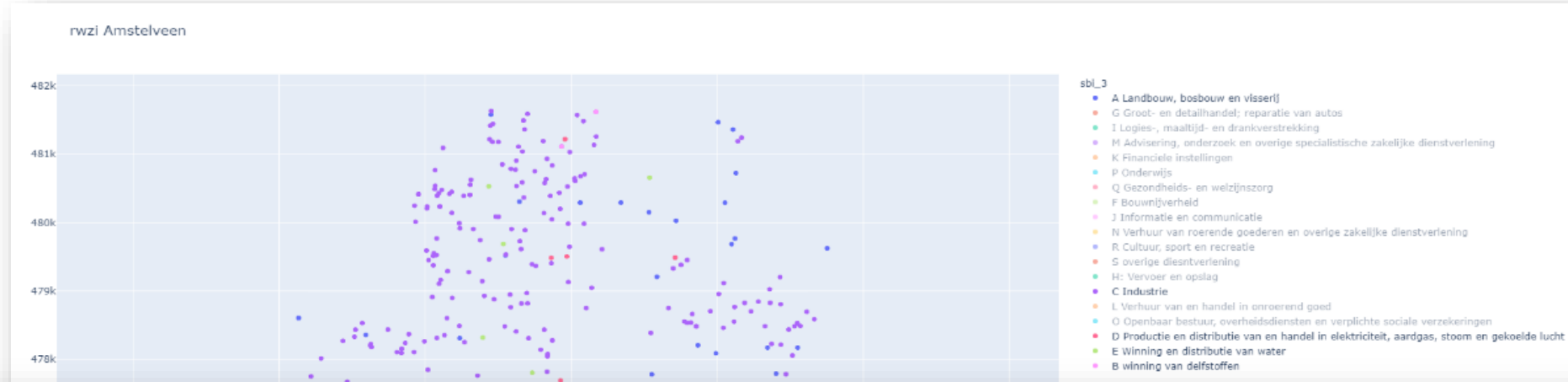
Is het mogelijk waterverbruik toe te kennen aan economische activiteiten?

- Markov Chain Monte Carlo Small Area Estimation (multi-level model) dat per rwzi maar ook landelijk verbanden leert in de data, eventueel ook voorzien van hulpinformatie (via zogenaamde *priors* uit Statline tabellen, per SBI hoofdklasse werd *het gemiddelde* en *standaarddeviatie* berekend op basis van de jaartallen 2018-2021)
- Verbanden worden gezocht tussen de doelvariabele “gecorrigeerd dag debiet” en contextuele variabelen zoals “aantallen bedrijven per SBI hoofdklasse” voor een specifieke rwzi

$$\text{Dagdebiet in } m^3 \text{ RWZI}_i \sim \beta_{1,i} \cdot \text{Aantal}_{pers.,i} + \beta_{2,i} \cdot \text{Aantal}_{bedr.,i} + \beta_{3_A,i} \cdot \text{Aantal SBI}_{A,i} + \dots + \beta_{3_U,i} \cdot \text{Aantal SBI}_{U,i} + \varepsilon_i$$

Onderzoeksvraag 2 | Methode

Is het mogelijk waterverbruik toe te kennen aan economische activiteiten?



A Landbouw, bosbouw en visserij	B winning van delfstoffen	C Industrie	D Productie en distributie van en handel in elektriciteit, E Winning en distributie van water	aantal_bedrijven	aantal_bedrijven_2	rwzi_naam_zinfo	datum	neerslag	neerslag_cor	debiet	debiet_cor
154	1	281	11	6	10891	19662 rwzi Amstelveen	1-1-2018	36023	36023	69690.5	31242
154	1	281	11	6	10891	19662 rwzi Amstelveen	2-1-2018	15730	29259	42555.1	21035
154	1	281	11	6	10891	19662 rwzi Amstelveen	3-1-2018	42498	24653	75772.1	36652
154	1	281	11	6	10891	19662 rwzi Amstelveen	4-1-2018	14678	33225	43788.7	21637
154	1	281	11	6	10891	19662 rwzi Amstelveen	5-1-2018	10542	13299	37370.4	25154
154	1	281	11	6	10891	19662 rwzi Amstelveen	6-1-2018	9376	10153	29429.8	19151
154	1	281	11	6	10891	19662 rwzi Amstelveen	7-1-2018	455	6402	25795	23230
154	1	281	11	6	10891	19662 rwzi Amstelveen	8-1-2018	10	307	24722.6	24608
154	1	281	11	6	10891	19662 rwzi Amstelveen	9-1-2018	2	7	23754.1	23750
154	1	281	11	6	10891	19662 rwzi Amstelveen	10-1-2018	2152	719	8628.5	6833
154	1	281	11	6	10891	19662 rwzi Amstelveen	11-1-2018	5054	3119	22905.9	18188
154	1	281	11	6	10891	19662 rwzi Amstelveen	12-1-2018	1020	3709	21742.6	19714
154	1	281	11	6	10891	19662 rwzi Amstelveen	13-1-2018	477	839	21828.5	21193
154	1	281	11	6	10891	19662 rwzi Amstelveen	14-1-2018	19	324	21854.4	21727

$$\text{Dagdebiet in } m^3 \text{ RWZI}_i \sim \beta_{1,i} \cdot \text{Aantal}_{pers.,i} + \beta_{2,i} \cdot \text{Aantal}_{bedr.,i} + \beta_{3_A,i} \cdot \text{Aantal SBI}_{A,i} + \dots + \beta_{3_U,i} \cdot \text{Aantal SBI}_{U,i} + \varepsilon_i$$

Onderzoeksvraag 2 | Resultaten en Conclusies

Is het mogelijk waterverbruik toe te kennen aan economische activiteiten?

– Het model leert de volgende parameters:

ZONDER STATLINE PRIOR	Coëfficiënt (β)	Standaard deviatie	Ondergrens (5%)	Bovengrens (95%)
Personen	0,12	0	0,12	0,12
Bedrijven (algemeen)	0,03	0,02	0,02	0,08
A Landbouw, bosbouw en visserij	0,79	0,57	0,7	1,92
B Delfstoffenwinning	37,51	26,3	34,06	83,74
C Industrie	0,52	0,46	0,39	1,38
D Energievoorziening	8,92	6,72	7,52	21,3
E Waterbedrijven en afvalbeheer	95,87	24,61	97,17	133,57
F Bouwnijverheid	0,28	0,23	0,22	0,75
G Handel	0,19	0,17	0,14	0,55
H Vervoer en opslag	1,13	0,95	0,87	3,06
I Horeca	0,68	0,6	0,52	1,89
J Informatie en communicatie	0,31	0,28	0,22	0,89
K Financiële dienstverlening	0,4	0,36	0,3	1,12
L Verhuur en handel van onroerend goed	1,45	1,29	1,09	4,18
M Specialistische zakelijke diensten	0,1	0,09	0,07	0,28
N Verhuur en overige zakelijke diensten	0,48	0,47	0,34	1,4
O Openbaar bestuur en overheidsdiensten	31,71	18,04	31,71	60,73
P Onderwijs	0,36	0,33	0,27	1
Q Gezondheids- en welzijnszorg	0,22	0,21	0,16	0,65

MET STATLINE PRIOR	Coëfficiënt (β)	Standaard deviatie	Ondergrens (5%)	Bovengrens (95%)
Personen	0,12	0,00	0,12	0,13
Bedrijven (algemeen)	0,06	0,04	0,01	0,14
A Landbouw, bosbouw en visserij	1,63	0,53	0,96	2,71
B Delfstoffenwinning	13,24	4,64	6,10	20,49
C Industrie	4,31	0,69	3,47	5,75
D Energievoorziening	19,46	6,54	10,11	30,96
E Waterbedrijven en afvalbeheer	6,96	2,24	3,14	10,19
F Bouwnijverheid	0,05	0,02	0,02	0,09
G Handel	0,16	0,06	0,08	0,25
H Vervoer en opslag	0,33	0,11	0,15	0,50
I Horeca	1,21	0,43	0,57	1,89
J Informatie en communicatie	0,04	0,01	0,02	0,07
K Financiële dienstverlening	0,07	0,02	0,03	0,10
L Verhuur en handel van onroerend goed	0,40	0,14	0,18	0,62
M Specialistische zakelijke diensten	0,04	0,01	0,02	0,06
N Verhuur en overige zakelijke diensten	0,05	0,01	0,02	0,07
O Openbaar bestuur en overheidsdiensten	4,11	1,36	1,87	6,10
P Onderwijs	0,27	0,10	0,13	0,43
Q Gezondheids- en welzijnszorg	0,37	0,14	0,18	0,62

$$\text{Dagdebiet in } m^3 \text{ RWZI}_i \sim \beta_{1,i} \cdot \text{Aantal}_{pers.,i} + \beta_{2,i} \cdot \text{Aantal}_{bedr.,i} + \beta_{3A,i} \cdot \text{Aantal SBI}_{A,i} + \dots + \beta_{3U,i} \cdot \text{Aantal SBI}_{U,i} + \varepsilon_i$$

Onderzoeksvraag 2 | Resultaten en Conclusies

Is het mogelijk waterverbruik toe te kennen aan economische activiteiten?

– Toepassen model op waterverbruik per SBI:

Debieten (miljoen m3) voor 2018	Model (geen prior)	Model (met prior)	Statline (leidingwater totaal)
Personen	665,3	673,7	837,2
A Landbouw, bosbouw en visserij	20,4	42,4	48,9
B Delfstoffenwinning	8,9	3,1	2,9
C Industrie	16,3	130,6	202,4
D Energievoorziening	6,2	13,6	13,3
E Waterbedrijven en afvalbeheer	99,5	7,3	5,5
F Bouwnijverheid	18,1	6,6	2,8
G Handel	21,6	22,2	13,5
H Vervoer en opslag	19,1	6,4	4,7
I Horeca	14,5	26	20,3
J Informatie en communicatie	9,4	2,9	1
K Financiële dienstverlening	14,5	4,1	1,9
L Verhuur en handel van onroerend goed	17,5	5,4	4,1
M Specialistische zakelijke diensten	13	10,1	3,5
N Verhuur en overige zakelijke diensten	13,5	2,7	1
O Openbaar bestuur en overheidsdiensten	47,3	6,2	4,7
P Onderwijs	10,3	8,6	6,6
Q Gezondheids- en welzijnszorg	13,8	23,6	19

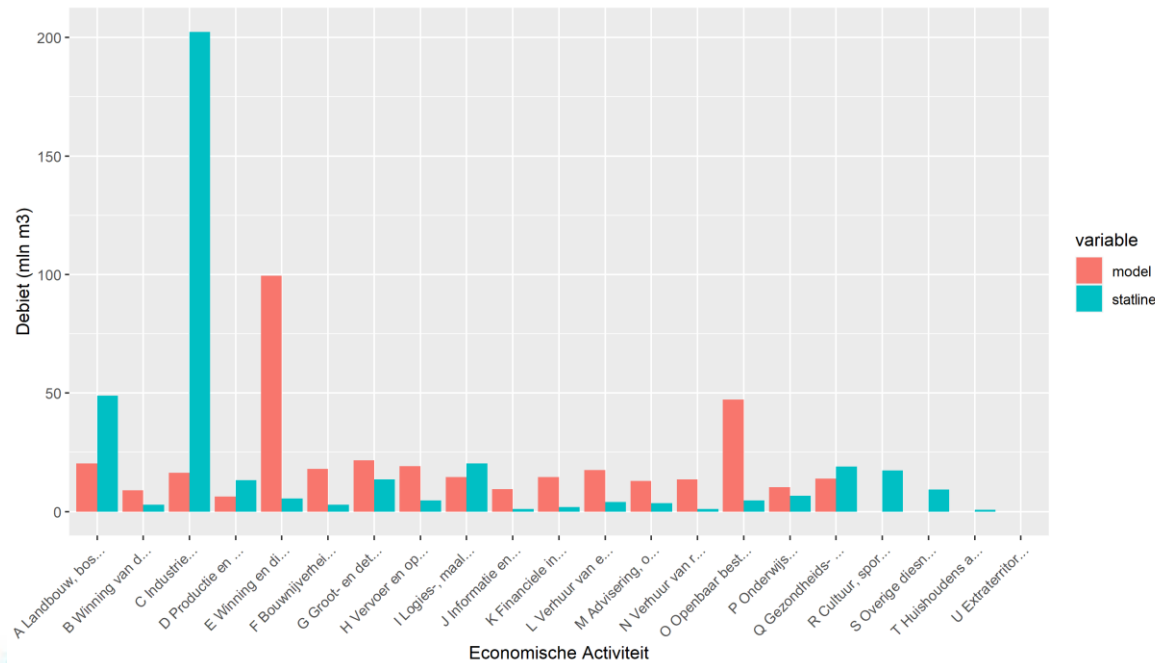
$$Gebruik\ per\ SBI_X = \beta_{2,i} \cdot Aantal_{bedr.,i} + \sum_{RWZI\ i} \beta_{3_X,i} \cdot Aantal\ SBI_{X,i}$$

Onderzoeksvraag 2 | Resultaten en Conclusies

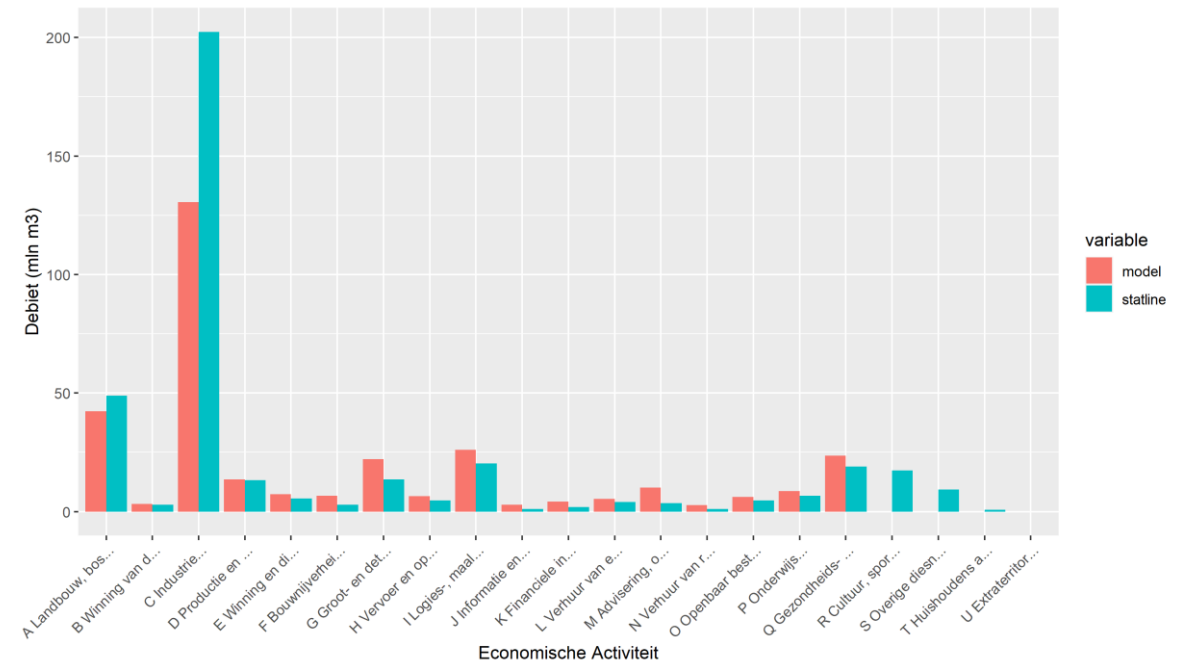
Is het mogelijk waterverbruik toe te kennen aan economische activiteiten?

- Het model zonder hulpinformatie is verre van goed (maakt grove misschattingen t.o.v. Statline). Priors “duwen” schattingen de goede kant op, maar het onderliggende model (data + algoritmie) blijft daarmee problematisch.

Vergelijking van modelschattingen (Geen_Priors) en statline (Leidingwater.Totaal, 2018*)



Vergelijking van modelschattingen (Priors) en statline (Leidingwater.Totaal, 2018*)



Conclusies

- Wat is bereikt tijdens dit pilot project?
 - Data verzameling en koppeling met zowel interne register data en externe data van de waterschappen en KNMI
 - Kennis opgedaan over de data
 - Een workflow is neergezet die jaarlijkse waterverbruik schattingen (statistieken) geeft per SBI en die vervolgens valideert met Statline
 - Potentie inzichtelijk gemaakt
- Wat is niet bereikt dit pilot project?
 - Op dit moment is een accurate data-gedreven schatting van de Statline tabel nog niet mogelijk.
 - Enkele verklaringen zijn
 - Tijdens data-verzameling zijn pragmatische keuzes gemaakt en data-bronnen sloten niet altijd op elkaar aan wat betreft de jaartallen, maar werden desalniettemin gekoppeld
 - Het project heeft besloten eerst z.s.m. toe te werken naar het eindproduct, i.t.t. focus te leggen op het verbeteren van de afzonderlijke onderdelen. Beide modellen toegepast (lees: voor schatten waterverbruik en vervolgens dat toekennen aan economische activiteiten) zijn relatief simplistisch en niet afzonderlijk verbeterd
 - Het bepalen van economische activiteit per gebouw gebeurt op dit moment niet op het meest gedetailleerde niveau, de economische activiteit kan afwijken van wat nu gemodelleerd wordt

Toekomstig werk

- Onderzoeksverslag voorleggen ter beoordeling aan divisie (milieurekeningen)
 - Is de geschetste output nuttig?
 - Vergelijken we niet appels met peren (tussen Z-Info en Statline)?
- Data verzameling en koppeling rechtekken over de jaartallen heen (in huidige analyse mismatch tussen beschikbare data per jaartallen)
- Data bovendien structureel koppelen aan BAG en op een veilige manier beschikbaar stellen aan statistische onderzoekers (o.a., voor verdere verrijking met externe data)
- Modellen:
 - M.b.t. schatten waterverbruik.
 - Verbetering *regionale* neerslagcorrectie (dus per rwzi of zelfs per overnamepuntgebied)
 - M.b.t. schatten waterverbruik per economische activiteit
 - Modelleren op maand- of zelfs jaarniveau (i.t.t., dagelijks in huidige analyse)
 - Modelleren op uurniveau (i.t.t., dagelijks in huidige analyse)
 - SBI toekenning op gebouw niveau (eerste inschatting is dat tot 12% van SBIs verkeerd geclassificeerd waren in huidige analyse)
 - Modelleren op gedetailleerder SBI niveau (i.t.t. meest grove indeling in huidige analyse)
 - Contextuele informatie verbreden, zoals oppervlakte van bedrijven (BAG), pand/bedrijf nog actief (BAG), aantal werknemers bedrijven (ABR)
 - Toepassen van andere modellen

afvalwaterprognoses NAAR EEN NIEUWE PRAKTIJK

