

A network diagram with various sized light blue circles connected by thin white lines, set against a solid blue background. The circles are of different diameters, and the lines connect them in a complex, non-hierarchical pattern.

KWR 2025.007 | April 2025

**Beoordeling van
disaggregatiemethoden
op synthetische digitale
watermeterdata**

Colofon

Beoordeling van disaggregatiemethoden op synthetische digitale watermeterdata

KWR 2025.007 | April 2025

Opdrachtnummer

404559/001

Projectmanager

B. (Bram) Hillebrand, MSc

Opdrachtgever

Collectief Nederlandse drinkwaterbedrijven

Auteur(s)

dr.ir. J. (Joeri) Willet, B. (Bram) Hillebrand, MSc, dr.ir. E.J.M. (Mirjam) Blokker

Kwaliteitsborger(s)

dr.ir. E.J.M. (Mirjam) Blokker

Verzonden naar

Projectgroep en de deelnemende commerciële partijen.

Openbaarheid

Dit rapport is openbaar.

Keywords

Disaggregatie, Digitale Watermeter

Jaar van publicatie
2025

Meer informatie

dr. ir. J. Willet
T 030-6069752
E joeri.willet@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

T +31 (0)30 60 69 511
E info@kwrwater.nl
I www.kwrwater.nl

KWR

April 2025 ©

Alle rechten voorbehouden aan KWR. Niets uit deze uitgave mag - zonder voorafgaande schriftelijke toestemming van KWR - worden veeelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier.

Samenvatting

Ontwikkelingen op het gebied van digitale watermeters en data-disaggregatie methodes doen vermoeden dat deze een waardevolle toevoeging kunnen zijn voor het Watergebruik Thuis (WT) onderzoek dat wordt uitgevoerd aan de hand van enquêtes. Het was echter onbekend in hoeverre data-disaggregatie het waterverbruik, zoals gemeten door digitale watermeters, nauwkeurig kan toekennen aan de individuele eindverbruiken (wc, douche, wasmachine, etc.). Een belangrijke reden hiervoor is een gebrek aan meetgegevens waarmee de uitkomsten vergeleken kunnen worden; er zijn niet genoeg locaties waar het waterverbruik op elk tappunt binnen een huis gemeten wordt.

Binnen dit onderzoek is gebruikgemaakt van synthetische waterverbruiksdata, gegenereerd met een aangepaste versie van pySIMDEUM (een model om waterverbruikspatronen te genereren, Blokker (2010)), als alternatief voor fysieke metingen. Externe partijen zijn benaderd om deze data te disaggregeren. De resultaten zijn beoordeeld op een viertal criteria: (1) de aanwezige apparaten/eindgebruiken, (2) waterverbruik naar volumen, (3) labels van de dataset, en (4) verbruiksgebeurtenissen onderscheiden.

In totaal zijn de resultaten van vijf disaggregaties, uitgevoerd door drie partijen, beoordeeld. Uit de resultaten blijkt dat de methodes/algoritmes over het algemeen ver gevorderd zijn en dat deze met een hoge mate van nauwkeurigheid de synthetische data kunnen disaggregeren. De aanwezigheid apparaten/eindverbruiken wordt tussen de 84% en 92% van de simulaties correct geïdentificeerd. Het waterverbruik naar volume wordt tussen de 82% en 91% correct toegewezen aan elk van de eindverbruiken. Het labels van de dataset gebeurt met een nauwkeurigheid tussen de 65% en 77%. Minimaal 58% en maximaal 89% van de individuele waterverbruiksgebeurtenissen worden met een marge van 1 seconde correct geïdentificeerd.

Deze resultaten laten zien dat disaggregatiemethodes een bijdrage kunnen leveren aan het WT onderzoek. Hiervoor gelden wel een aantal aanbevelingen/vervolgstappen: (1) herhaling van disaggregatie van synthetische data na verdere ontwikkeling van pySIMDEUM (voor een realistischere weergave van huishoudelijk waterverbruik), (2) aanvullend onderzoek met nieuwe synthetische datasets onder nieuwe/deelnemende softwareleveranciers met een andere nauwkeurigheid van de data.

Inhoud

Colofon	2
Samenvatting	3
Inhoud	4
1 Inleiding	6
1.1 Aanleiding en doel	6
1.2 Leeswijzer	7
2 Deelname externe partijen	8
3 SIMDEUM	10
4 Aanpassingen aan pySIMDEUM	11
4.1 Opeenvolgende dagen	11
4.2 Stroomsnelheid van het toilet	11
4.3 Patronen voor vaatwasser en wasmachine	11
4.4 Ruis	12
4.5 Blokform van waterverbruik	13
4.6 Verlaging van volumestroom bij simultaan verbruik	13
4.7 Koppelingen in waterverbruik	14
5 Genereren van testdata	15
5.1 Conversie naar digitale watermeterdata	15
5.2 Dataset voor disaggregatie	16
5.3 Test data, Voorkennis en trainingsdata	17
6 Beoordelingscriteria	18
7 Resultaten	19
7.1 Aanwezige apparaten/eindverbruiken	20
7.2 Waterverbruik per huishouden	23
7.3 Labelen van de dataset	28
7.4 Gebeurtenissen onderscheiden	36
8 Conclusies en aanbevelingen	38
9 Referenties	40
I Begeleidende brief en vragenlijst voor commerciële partijen	41
II Detail huishoudsamenstelling	45
III Instructies voor deelnemende partijen	46
IV Begeleidende brief oplossingen partij A	54

V	Prestaties van de disaggregatiemethoden naar huishoudsamenstelling	55
VI	Volumetrisch waterverbruik per uur en per dag vergeleken met het totale waterverbruik	56
VII	Vervolgstappen disaggregatieonderzoek	61

1 Inleiding

1.1 Aanleiding en doel

Vewin laat sinds 1992 iedere drie jaar onderzoek doen naar de opbouw van het huishoudelijk drinkwaterverbruik in Nederland, het “Watergebruik Thuis” onderzoek (WT met aanduiding van het jaar). De data zijn jarenlang middels enquêtes en het invullen van een dagboek vergaard, eerst door Kantar (en voorgangers), en de laatste keer door het CBS. Het voorlaatste onderzoek (WT-2017) vertoonde te grote afwijkingen met de afzetdata die waterbedrijven zelf vergaren. In WT-2017 zijn daarom correctiefactoren toegepast op de data. Onderzocht wordt dus in hoeverre een alternatieve methode gebruikt kan worden.

Er zijn digitale watermeters op de markt die met een hoge resolutie kunnen meten en de data kunnen ook hoogfrequent worden gelogd op een externe of ingebouwde logger. Er zijn verschillende aanbieders van de hardware. Er zijn ook aanbieders die de data kunnen analyseren; soms zijn dit partijen die ook hardware aanbieden. Enkele partijen hebben een methode ontwikkeld om de metingen van de watermeter om te zetten naar individuele tappingen (disaggregatie), en deze te classificeren volgens de componenten van huishoudelijk verbruik, zoals douche, toiletspoeling, wasmachine. Het blijkt echter dat niet gekwantificeerd is hoe goed de verschillende algoritmes zijn in disaggregatie en het classificeren van het verbruik. Bunt Planet gaf bijvoorbeeld aan dat ze geen beschikking hebben over gelabelde data waarmee ze hun resultaten kunnen valideren. KWR heeft in het BTO ook geprobeerd om methodes uit de literatuur toe te passen voor disaggregatie en classificatie. De betrouwbaarheid bleek beperkt¹.

Voordat metingen van digitale watermeters in combinatie met analysesoftware het dagboekonderzoek kunnen vervangen t.b.v. het WT-onderzoek, is het belangrijk om te bepalen hoe goed deze software werkt. Met behulp van SIMDEUM kan KWR realistische verbruikspatronen genereren waarbij precies bekend is of het verbruik van een douche is, wc, wasmachine etc. De volgende stap is deze “perfecte” data om te zetten naar de vorm die een digitale watermeter zou meten (som van alle verbruiken, een bepaalde puls grootte en logfrequentie) en om (eventueel) nog enige ruis toe te voegen. Het is eenvoudig om lange tijdreeksen te maken (tot meerdere maanden) voor veel (meer dan 100) verschillende huishoudens, met elk hun eigen gezinssamenstelling (éénpersoonshuishouden, tweepersoonshuishouden en gezinnen) met verschillende waterverbruikende apparaten (1 of 2 wc’s, douche met en zonder waterbesparende douchekop, wel of geen bad, wel of geen vaatwasser, wel of geen tuin) en verschillend gedrag (week of weekend, zuinig gedrag of niet). Zo kan een goede set aan testdata worden gecreëerd aan de hand waarvan disaggregatiesoftware kan worden beoordeeld.

In dit project is een objectieve beoordeling uitgevoerd naar de effectiviteit van de verschillende disaggregatie-oplossingen welke ontwikkeld zijn door marktpartijen. De software-aanbieders ontvangen de gesimuleerde watermeterdata, zonder de labels van de verbruikscomponenten. Zij leveren de resultaten van hun analyse: een indeling van waterverbruik met labels. KWR vergelijkt de resultaten met de echte labels en beoordeelt de nauwkeurigheid van de methode. De beoordeling zal worden gedaan in het licht van toepasbaarheid in het WT-onderzoek.

¹ (70%, afhankelijk van de component, betere match voor douche dan voor vaatwasser)

1.2 Leeswijzer

Dit rapport is opgebouwd uit de volgende hoofdstukken:

- Externe partijen zijn benaderd met de vraag of ze hun disaggregatie software wilden testen op de synthetische digitale watermeterdata (Hoofdstuk 0).
- Een korte beschrijving van de werking en structuur van het SIMDEUM model waarmee de synthetische watermeterdata gegenereerd is (Hoofdstuk 1)
- De aanpassingen aan SIMDEUM voor de doeleinden van dit project (Hoofdstuk 4).
- De conversiemethode van SIMDEUM output naar digitale watermeterdata en het genereren van de testdata (Hoofdstuk 0).
- Het ontwerp van de beoordelingscriteria (Hoofdstuk 0).
- De resultaten van de disaggregatie door de externe partijen (Hoofdstuk 7).
- Conclusies aanbevelingen (Hoofdstuk 0)

Omdat de deelnemende partijen niet-Nederlandstalig zijn, is de communicatie met hen in het Engels gevoerd. De meeste tabellen en figuren in dit rapport zijn daarom in het Engels uitgevoerd, en wordt een decimale punt gebruikt voor de cijfers.

De termen *watergebruik* en *waterverbruik* worden in de literatuur vaak door elkaar gebruikt. In dit rapport is gekozen voor de term *verbruik*. In de Engelse tabellen en figuren gebruiken we 'end uses' voor de verschillen *eindverbruiken*.

Voor de leesbaarheid van sommige tabellen en figuren worden de eindverbruiken afgekort zoals in Tabel 1.

Tabel 1 Afkortingen voor eindverbruiken in tabellen en figuren.

End use	Abbreviation
BathroomTap	BTap
Bathtub	BTub
Dishwasher	Dish
KitchenTap	KTap
OutsideTap	OTap
Shower	Show
WashingMachine	Wash
Wc	Wc

2 Deelname externe partijen

Uit eerder contact met een aantal commerciële partijen was bekend dat er aan de disaggregatie van digitale watermeterdata gewerkt wordt. Deelname aan deze disaggregatie-opdracht zou deze partijen de mogelijkheid bieden om hun software/algoritmes te testen en te valideren. Daarnaast kan de evaluatie ontwikkeld binnen dit project uitgroeien tot een soort benchmark dat gebruikt wordt door de drinkwaterbedrijven bij aanbestedingen. In totaal zijn er 16 partijen aangeschreven om mee te doen met dit project. Het merendeel van deze partijen was ofwel bekend bij de leden van de begeleidingsgroep of bij de onderzoekers van KWR. Een aantal partijen zijn online gevonden. Aangezien niet alle algoritmes op dezelfde manier werken en de motivatie voor deelname verschillend kan zijn hebben de deelnemende partijen een vragenlijst ontvangen om te bepalen wat de eigenschappen van de dataset moesten worden. Een kopie van de begeleidende brief en de vragenlijst is te vinden in Bijlage I.

Van de 16 partijen waaraan deze brief toegestuurd is hebben 6 partijen toegezegd mee te willen doen aan het huidige project. De huidige opzet maakt het mogelijk om deelname in vervolgstappen makkelijk uit te breiden. In samenspraak met de deelnemende partijen is er gekozen voor een dataset met een resolutie van 0.1 L per puls en een tijdstap van 1 s (Tabel 2). Niet elke deelnemende partij heeft aangegeven dezelfde data nodig te hebben, zoals de huishoudsamenstelling of de aanwezige eindverbruiken, voor het uitvoeren van de disaggregatie. Onderdeel van de beoordeling is welke data gebruikt is.

Uiteindelijk hebben drie partijen daadwerkelijk de disaggregatie uitgevoerd en hun resultaten teruggestuurd ter evaluatie. Twee van de partijen wilden anoniem blijven in deze rapportage. We refereren naar de partijen als partij A, B, en C. Partij B wilde wel bij naam genoemd worden, dit is Procter & Gamble (P&G).

Tabel 2 Overzicht van de antwoorden op de vragenlijst naar de uiteindelijke deelnemers (anoniem). De dikgedrukte en onderstreepte waarden zijn gebruikt in het genereren van de dataset.

Question	Participant A	Participant B	Participant C
1 Temporal resolution	<u>1s</u>	<u>1s</u> , 10s, 5s	<u>1s</u> , (preferably 0.01 seconds)
2 Flow resolution	<u>0.1 L/pulse</u>	<u>0.1l/pulse</u> , 1l/pulse	<u>0.1 L/pulse</u>
3 Household information	Number of residents, Water using appliances present, Potential labels to choose from	-	Number of residents, Water using appliances present, Potential labels to choose from
4 Time information	Day of the week	Day of the week/weekend, timestamp with hour and minute	Day of the week
5 Training dataset	Test dataset, Training dataset (with labels), Start up timeseries with labels (fingerprint for appliances)	Training dataset (with labels)	Start up timeseries with labels (fingerprint for appliances)
6 Aggregation level	Aggregated tap level - does not really matter	End use level, Tap level	The end-use level
7 Assessment criteria*	Overall score, score per label, score per water use at home, percentage of volume labelled correctly, percentage of events labelled correctly, Insight into false negatives/false positives	All	Percentage of volume labelled correctly, Percentage of events labelled correctly
8 Algorithm	Post-processing - hours	Post processing: ~5 min. to extract features, ~30 sec. for training, and ~5 sec. for inference, Accuracy: ~70% balanced accuracy on end-use labels.	-
9 Processing time**	6 weeks	4 weeks	4 weeks
10 Restrictions	Decision is made after reading the draft report	Publicly available	Must be checked with relevant department

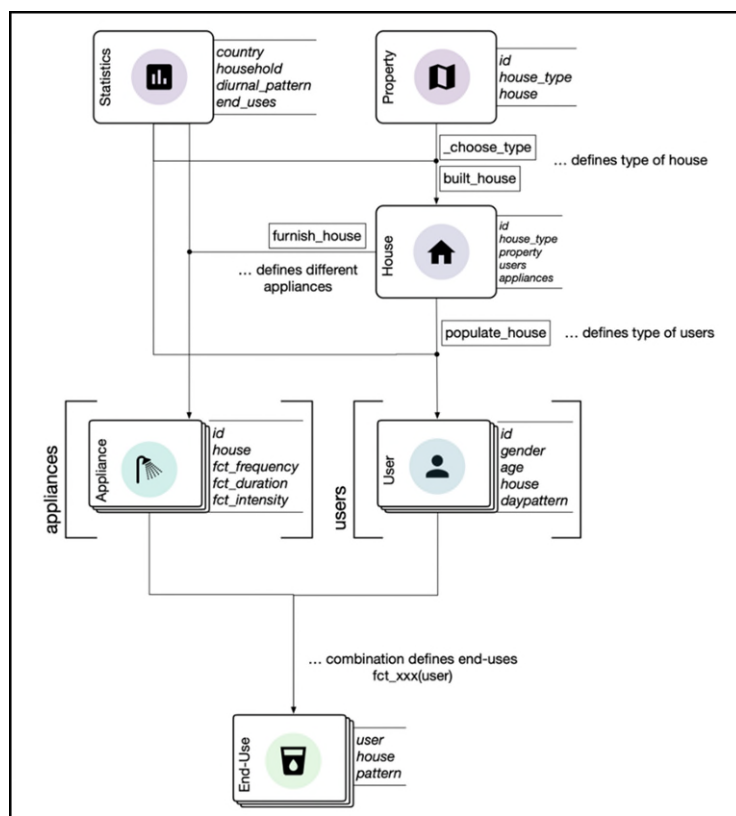
*The final assessment criteria were the same for all participants. The final set of criteria was established in consultation with the Dutch drinking water companies and the participants.

**Initial estimate at the time of project start, actual processing time is likely to differ in case of actual implementation and optimization.

3 SIMDEUM

Voor het genereren van de synthetische waterverbruikspatronen is het SIMDEUM (Blokker, 2010) model gebruikt. SIMDEUM is een stochastisch model waarin de verschillende toepassingen voor waterverbruik gesimuleerd worden met Poisson Rechthoekige Pulsen. Per eindverbruik is een kansverdeling bepaald welke het begin, de intensiteit, en de duur van een puls bepaalt. De parameters voor de kansverdelingen worden bepaald op basis van het Watergebruik Thuis onderzoek door de Nederlandse drinkwaterbedrijven. SIMDEUM telt acht typen eindverbruik: WC, douche, wasmachine, vaatwasser, keukenkraan, badkamerkraan, bad, en buitenkraan. Waterverbruik door deze eindverbruikers wordt toegekend aan de gebruikers (bewoners) van een huishouden. Binnen SIMDEUM zijn er drie huishoudsamenstellingen gedefinieerd: eenpersoons, tweepersoons (zonder kinderen), en familie (twee mensen, met kinderen). Het waterverbruik van de gebruikers wordt beïnvloed door hun leeftijd, gender, en beroep (werkend, niet werkend, school). De samenstelling van huishoudens is gebaseerd op Nederlandse data. In eerste instantie is SIMDEUM in MATLAB® geïmplementeerd. Een python versie, pySIMDEUM, is in 2022 beschikbaar gemaakt als een open-source package (Steffelbauer et al., 2024) en is te vinden op git-hub (<https://github.com/KWR-Water/pysimdeum>).

De modelstructuur van pySIMDEUM is weergegeven in Figuur 1. In eerste instantie wordt een huis gegenereerd op basis van de drie mogelijke soorten huishouden (built_house). Hierna worden de eindverbruiken welke aanwezig zijn bepaald aan de hand van de penetratiegraad (furnish_house), en worden de eigenschappen van de gebruikers bepaald op basis van de landelijke statistieken (populate_house). Op basis van dit huishouden worden de waterverbruikspatronen gegenereerd.



Figuur 1 Overzicht van pySIMDEUM, de python implementatie van SIMDEUM.

4 Aanpassingen aan pySIMDEUM

Bij aanvang van het project zijn er een aantal aspecten binnen pySIMDEUM geïdentificeerd waarvoor aanpassingen nodig waren voor dit project, of een meerwaarde zouden hebben.

4.1 Opeenvolgende dagen

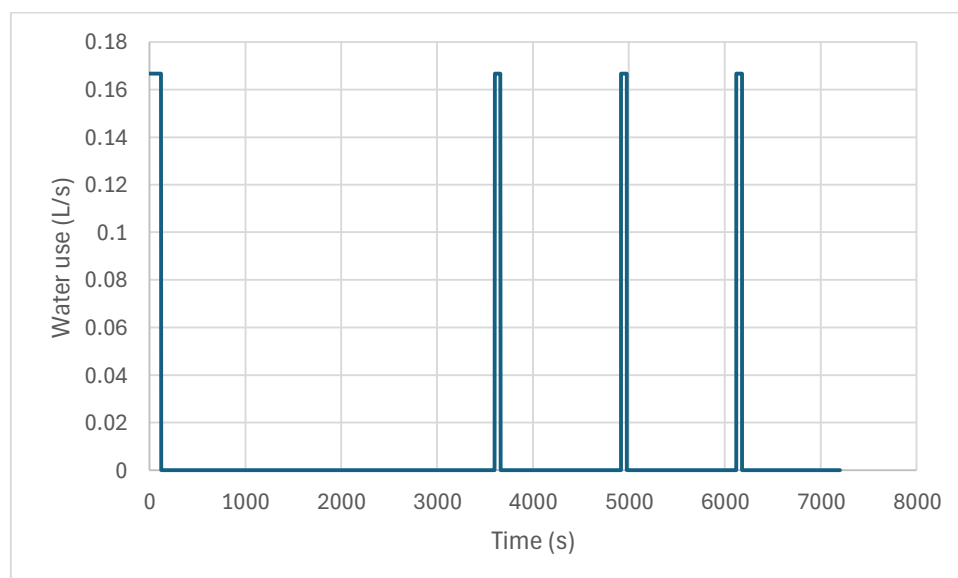
Het originele doel van SIMDEUM was om het gemiddelde waterverbruik van een huishouden, of een groep van huishoudens, te simuleren en op dagbasis te rapporteren. In pySIMDEUM is de simulatieperiode in eerste instantie een dag, waarna de tijdsreeks wordt afgebroken. In dit onderzoek is het echter van belang om meerdere dagen achter elkaar te kunnen simuleren en ervoor te zorgen dat waterverbruik aan het eind van een dag doorloopt naar de volgende dag, bijvoorbeeld wanneer iemand net voor middernacht onder de douche stapt moet deze douchebeurt doorlopen in de eerste minuten van de volgende dag.

4.2 Stroomsnelheid van het toilet

In de originele versie van pySIMDEUM was de volumestroom voor het vullen van het toiletreservoir $0.042 \text{ L}^3/\text{s}$. Hiervan was bekend dat dit aan de lage kant was. Een nieuw geïnstalleerd toilet, waarvan het model bekend was, is meerdere malen doorgespoeld om een nieuw debiet van 0.125 L/s vast te stellen en te implementeren.

4.3 Patronen voor vaatwasser en wasmachine

In de originele versie van pySIMDEUM is er maar een mogelijk patroon voor de vaatwasser en voor de wasmachine geïmplementeerd. Beide apparaten hebben vier innamecycli. Bij de wasmachine is het totale waterverbruik 50.2 L , duurt de eerste innamecyclus 121 seconden en de rest 60 seconden (Figuur 2). Bij de vaatwasser is het totale waterverbruik 14 L en duren alle innamecycli 24 seconden.



Figuur 2 Voorbeeld van de originele innamecycli van een wasmachine in SIMDEUM

In werkelijkheid hebben deze apparaten meerdere programma's die met verschillende frequenties gebruikt worden. Voor de wasmachine zijn er drie programma's geïmplementeerd, voor de vaatwasser twee. Het verschil in totaal waterverbruik wordt behaald door de duur van de innamecycli aan te passen (Tabel 3).

Tabel 3 Eigenschappen van verschillende programma's voor de wasmachine en de vaatwasser

Property	Washing Machine			Dishwasher	
	Normal	Eco	Short	Normal	Eco
Duration (s)	7200	9000	1800	7200	12600
Water use (L)	51.2	51.2	40.8	14.4	11.4
Intake duration	301	301	240	96	76
Flow (L/s)	0.17	0.17	0.17	0.15	0.15
Usage probability	45%	46%	9%	50%	50%

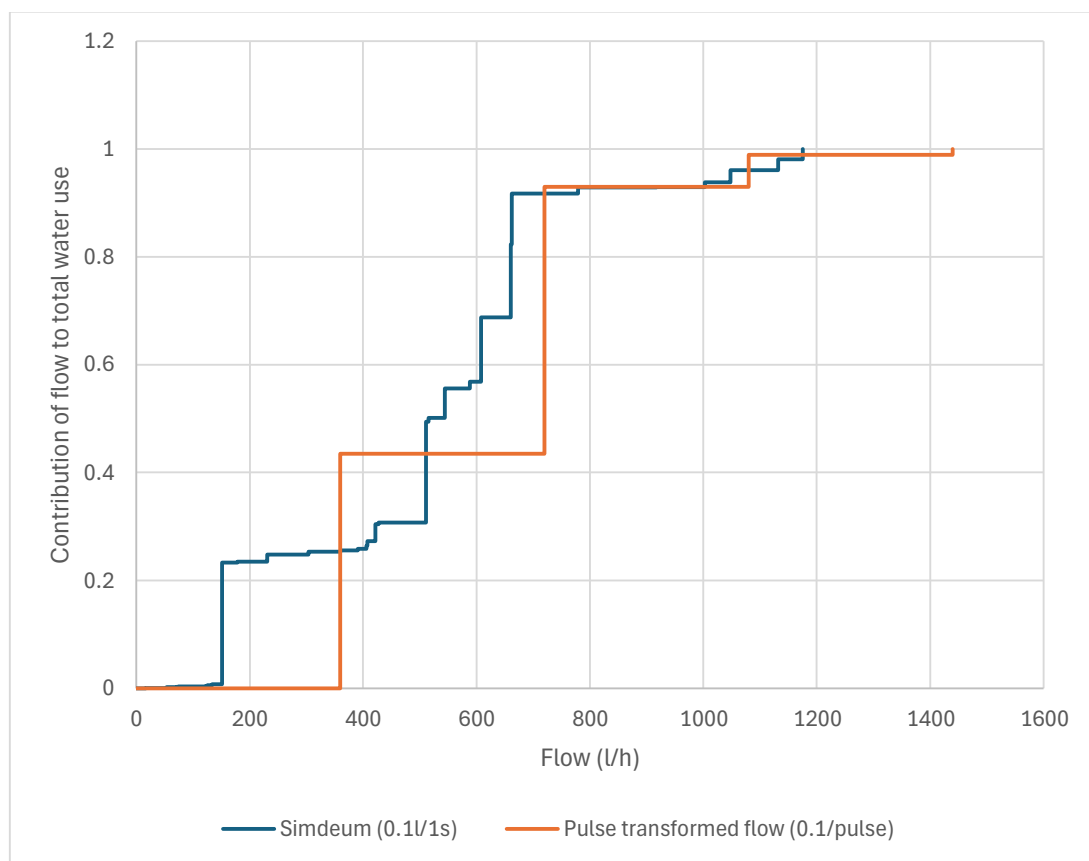
4.4 Ruis

Bij aanvang van het project was het onduidelijk of het nodig was om synthetische ruis toe te voegen aan de SIMDEUM-data om de disaggregatie-opdracht realistisch genoeg te maken. Daarnaast was er ook de hypothese dat het toevoegen van ruis nodig zou zijn om de opdracht niet te makkelijk te maken. Om dit te onderzoeken moest de SIMDEUM-data eerst omgezet worden naar het signaal van een digitale watermeter om te kunnen evalueren of ruis toevoegen nodig was. De hoeveelheid pulsen die simultaan geregistreerd kunnen worden zijn discreet verdeeld, waardoor de mogelijke stroomsnelheden van de originele simulatie teruggebracht worden tot een gelimiteerde set (Tabel 4).

Tabel 4 Voorbeeld van mogelijke volumestromen op basis van simultane pulsen uit een digitale watermeter. In de derde kolom is een SIMDEUM-simulatie omgezet naar pulsen ($0.1 \text{ L puls}^{-1} \text{ s}^{-1}$) en geëvalueerd hoeveel van de totale volumestroom hieruit voortkomt.

# Simultaneous pulses	Flow (L/h)	Percentage of total water use
0	0	0%
1	360	43%
2	720	50%
3	1080	6%
4	1440	1%

Wanneer de stroomsnelheden van de originele simulatie vergeleken worden met de stroomsnelheden na de pulstransformatie dan blijkt dat een groot deel van de originele simulatie tussen de 0 tot 2 pulsen valt (Figuur 3). De grote hoeveelheid aan mogelijke stroomsnelheden, die gereduceerd worden tot maximaal 4 stroomsnelheden, en het feit dat veel van deze tussen de mogelijke discrete waarden invallen, kan gezien worden als een soort ruis. Er is daarom besloten om geen extra ruis aan de data toe te voegen. In andere woorden, het omzetten van de stroomsnelheden uit SIMDEUM naar pulldata is voorlopig voldoende bevonden om het toevoegen van extra ruis overbodig te maken.



Figuur 3 Contributie van verschillende stroomsnelheden aan het totale watergebruik van een huishouden binnen een SIMDEUM simulatie en de stroomsnelheden na de pulstransformatie

4.5 Blokform van waterverbruik

In SIMDEUM worden de verschillende waterverbruiken gesimuleerd met een blokfunctie waarbij de waarde ofwel de volumestroom van het eindverbruik of nul is (Blokker, 2010). In werkelijkheid kan het voorkomen dat waterverbruik zich niet altijd aan deze blokform houdt. Dit kan gebeuren door het gedrag van de bewoners of omdat er schommelingen zijn in de druk van het systeem waardoor er variatie ontstaat in de volumestroom bij het eindverbruik. Dit aspect van SIMDEUM is bekend maar toepassing zou teveel tijd kosten en de toegevoegde waarde ervan is te gering bevonden.

4.6 Verlaging van volumestroom bij simultaan verbruik

Het is bekend dat wanneer meerdere apparaten in een huishouden tegelijk gebruikt worden de volumestroom per apparaat kan afnemen vergeleken met enkel gebruik. Bijvoorbeeld, de volumestroom uit de keukenkraan wordt (tijdelijk) lager wanneer de WC doorgespoeld wordt of de douche gebruikt wordt. Dit effect is niet aanwezig in SIMDEUM of pySIMDEUM. Voor implementatie van dit fenomeen zijn twee mogelijkheden geïdentificeerd:

1. Een hydraulisch model van een huishouden koppelen aan de watervraag gegenereerd vanuit SIMDEUM. Een volledige implementatie hiervan zou ook een terugkoppeling van het hydraulisch model naar SIMDEUM vergen; een lagere volumestroom vanwege simultaan gebruik zal er voor zorgen dat de keukenkraan langer open staat bij het vullen van een pan water.
2. Metingen bij meerdere huishoudens uitvoeren om de volumestroom van eindverbruiken te bepalen tijdens simultaan gebruik. De verhoudingen van enkel gebruik ten opzichte van simultaan gebruik kunnen aan SIMDEUM toegevoegd worden binnen de simulatie zelf, of als postprocessing van de simulatie.

Beide van bovenstaande opties hebben op dit moment onvoldoende meerwaarde buiten de context van dit project om deze aanpassingen in pySIMDEUM te implementeren.

4.7 Koppelingen in waterverbruik

Wanneer mensen het toilet doorspoelen dan zullen ze over het algemeen vlak daarna een kraan gebruiken om hun handen te wassen. Dit soort koppelingen van waterverbruik is niet geïmplementeerd in SIMDEUM. SIMDEUM is in eerste instantie ontworpen om gemiddelde verbruikspatronen te generen van een huishouden over meerdere dagen, van meerdere huishoudens in een straat/wijk, of van grotere gebouwen. Bij het middelen/optellen van meerdere dagen of huishoudens vervalt het belang van details over de volgordelijkheid van waterverbruik. In SIMDEUM wordt op dit moment een enkele set van kansverdelingen gebruikt om te bepalen op welke momenten water verbruikt wordt. De implementatie van koppelingen in waterverbruik vergt een tweede set kansverdelingen waarmee eindverbruik geactiveerd kan worden op basis van voorgaand waterverbruik. Het bepalen van deze kansverdelingen, en de benodigde aanpassingen aan de modelstructuur van pySIMDEUM, vallen buiten de scope van dit project. Voor mogelijke toekomstige implementatie van deze koppelingen is het van belang om het effect hiervan in relatie te zien tot de eindverbruiken met het grootste verbruik, en op basis hiervan te prioriteren hoe belangrijk deze koppelingen zijn.

5 Genereren van testdata

5.1 Conversie naar digitale watermeterdata

SIMDEUM produceert een tabel met waterverbruik per eindverbruik voor elke tijdstap in de simulatieperiode (Tabel 5).

Tabel 5 Voorbeeld van een SIMDEUM output. Waardes voor de eindverbruiken zijn in liter (L)

Time	WcNew	Bathtub	BathroomTap	Dishwasher	KitchenTap	FancyShower	WashingMachine
00:03:28	0	0	0	0	0	0.0123	0
00:03:29	0	0	0	0	0	0.0123	0
00:03:30	0	0	0	0	0	0.0123	0
00:03:31	0.125	0	0.0045	0	0	0.0123	0
00:03:32	0.125	0	0.0045	0	0	0.0123	0
00:03:33	0.125	0	0.0045	0	0	0.0123	0
00:03:34	0.125	0	0.0045	0	0	0.0123	0
00:03:35	0.125	0	0.0045	0	0	0.0123	0
00:03:36	0.125	0	0.0045	0	0	0.0123	0
00:03:37	0.125	0	0	0	0	0.0123	0
00:03:38	0.125	0	0	0	0	0.0123	0
00:03:39	0.125	0	0	0	0	0.0123	0

Deze data is omgezet naar een vergelijkbaar format waarmee digitale watermeters waterverbruik loggen. Op deze manier genereren we een gelabelde dataset welke gebruikt kan worden voor training en evaluatie. Tot noch toe werden voor dit soort onderzoeken startup protocollen in combinatie met dagboeken gebruikt. Binnen dit project is er gekozen voor een meetmethode waarbij pulsen een bepaalde precisie hebben (volume) en een bepaalde frequentie waarmee deze geregistreerd worden. Daarnaast wordt uitgegaan van een 100% nauwkeurigheid van de digitale watermeter; wanneer de precisie 0.1 L is dan zal elke 0.1 L daadwerkelijk gelogd worden. Een voorbeeld van deze pulstransformatie is weergegeven in (Tabel 6). In dit voorbeeld is de eerste keer dat een puls van 0.1 L behaald wordt om 00:18:28 ($3 \times 0.042 = 0.126$), de tweede keer dat een puls geregistreerd wordt is twee seconden later om 00:18:30 ($0.126 + 0.042 + 0.042 = 0.21$).

Tabel 6 Voorbeeld van de puls transformatie van SIMDEUM naar het signaal van een digitale watermeter

Time	SIMDEUM (L)	Cumulative water use (L)	Logger pulse (0.1L/1s)
2024-03-04 00:18:23	0	0	0
2024-03-04 00:18:24	0	0	0
2024-03-04 00:18:25	0	0	0
2024-03-04 00:18:26	0.042	0.042	0
2024-03-04 00:18:27	0.042	0.084	0
2024-03-04 00:18:28	0.042	0.126	1
2024-03-04 00:18:29	0.042	0.168	0
2024-03-04 00:18:30	0.042	0.21	1
2024-03-04 00:18:31	0.042	0.252	0
2024-03-04 00:18:32	0.042	0.294	0

2024-03-04 00:18:33	0.042	0.336	1
2024-03-04 00:18:34	0.042	0.378	0

5.2 Dataset voor disaggregatie

Binnen dit project is het waterverbruik van 100 huizen gesimuleerd met verschillende samenstellingen en aanwezigheid van eindverbruiken. De aanwezigheid van de verschillende eindverbruiken is gebaseerd op de penetratiegraad in Nederland. Binnen SIMDEUM bestaan de mogelijkheden voor eenpersoonshuishoudens, tweepersoonshuishoudens of families (3 of meer personen). De compositie van de gesimuleerde huishoudens is weergegeven in Tabel 7. Elk van de samenstellingen had een 34% (eenpersoons), 30% (tweepersoons), en 36% (familie) kans om gegenereerd te worden elke keer dat een huis gesimuleerd is.

Tabel 7 Huishoudsamenstelling in de gesimuleerde huishoudens

Inhabitants	Number of households
1	36
2	34
3	13
4	14
5	3

De aanwezigheid van de apparaten in elk van de huishoudens is weergegeven in Tabel 8. Voor een detailweergave van de eindverbruiken aanwezig in de simulatie zie Bijlage II.

Tabel 8 Overzicht van de verdeling van de eindgebruiken aanwezig in de honderd gesimuleerde huishoudens naar de huishoudsamenstelling.

Users	BTap		BTub		Dish		KTap		OTap		Show		Wash		Wc	
	Nr	%	Nr	%	Nr	%	Nr	%	Nr	%	Nr	%	Nr	%	Nr	%
1	36	100%	7	19%	14	39%	36	100%	20	56%	36	100%	32	89%	36	100%
2	34	100%	16	47%	24	71%	34	100%	20	59%	34	100%	33	97%	34	100%
3	13	100%	8	62%	12	92%	13	100%	11	85%	13	100%	12	92%	13	100%
4	14	100%	8	57%	14	100%	14	100%	8	57%	14	100%	14	100%	14	100%
5	3	100%	2	67%	2	67%	3	100%	3	100%	3	100%	3	100%	3	100%
Total	100		41		66		100		62		100		94		100	

De deelnemers werd mogelijk gemaakt om hun algoritme/methode te trainen op een deel van de data. Hiervoor is voor elk huishouden een simulatie van 14 dagen uitgevoerd. De eerste 7 dagen konden gebruikt worden voor training. Dagen 8 tot en met 14 waren de testdata voor de disaggregatie. Een voorbeeld van de test data die de deelnemers ontvingen is te zien in Tabel 9.

Tabel 9 Voorbeeld van de testdata format

Time	Day	Weekday	Pulses	Flow (L)
06:12:49	8	Weekday	0	0
06:12:50	8	Weekday	0	0
06:12:51	8	Weekday	0	0
06:12:52	8	Weekday	1	0.1
06:12:53	8	Weekday	1	0.1
06:12:54	8	Weekday	2	0.2

06:12:55	8	Weekday	1	0.1
06:12:56	8	Weekday	1	0.1
06:12:57	8	Weekday	1	0.1

5.3 Test data, Voorkennis en trainingsdata

De deelnemende partijen ontvingen allemaal dezelfde data om te disaggregeren, de testdata. Daarnaast kregen ze de mogelijkheid om voorkennis over het huisouden en/of trainingsdata te ontvangen om hun algoritmes op te zetten/trainen. De data die ze konden ontvangen, is weergegeven in Tabel 10.

Tabel 10 Overzicht van de data de deelnemende partijen konden ontvangen

Name	Information type	Description
Test data	Data for disaggregation	Pulse data for days 8-14 of the water use patterns generated with pySIMDEUM for 100 households.
Composition data	Household info	List with the composition of each household in terms of the number of users.
Appliance data	Household info	List with the appliances present in each household.
Training tap log	Training data	A list of water using events of the first seven simulation days (other than the test data) for each household. For each event the start time, the end use and the duration is reported.
Startup protocol	Training data	An example of the pulse fingerprint generated by each appliance in a household.
Unlabeled pulse data	Training data	The pulse data for the first seven simulation days of 25 out of the 100 households.

6 Beoordelingscriteria

De beoordelingscriteria voor de disaggregatie zijn bedoeld om een zo groot mogelijke meerwaarde te bieden voor het Watergebruik Thuis onderzoek, en zijn tot stand gekomen in een iteratief proces tussen de onderzoekers van KWR en de begeleidingsgroep. Input van de deelnemende partijen door middel van de vragenlijst is meegenomen, maar was niet doorslaggevend.

De beoordelingscriteria zijn opgebouwd uit vier categorieën:

1. Aanwezige apparaten/eindverbruiken: een evaluatie van de mate waarin de disaggregatie een correcte identificatie maakt van de eindgebruiken binnen het huishouden.
2. Totaalverbruik per eindverbruik: het totale waterverbruik van de disaggregatie in relatie tot de testdata op basis van verschillende tijdsvlakken binnen de simulatie – de totale dataset, per dag, en per uur. Hierin wordt onderscheid gemaakt tussen het relatieve verschil binnen de eindverbruikscategorie, en in relatie tot het totale waterverbruik.
3. Labelen van de dataset: aangeven welke eindverbruiken (mede)verantwoordelijk zijn voor elk van de pulsen in de test data. Het percentage correcte labels (het label uit de disaggregatie komt overeen met het label uit de dataset), vals-positieven (volgens de disaggregatie wel waterverbruik maar niet in de testdata), en vals-negatieven (volgens de disaggregatie geen waterverbruik maar wel in de testdata) wordt gerapporteerd. Momenten zonder puls in de disaggregatie en ook niet in de testdata worden buiten beschouwing gelaten. Naast deze percentages rapporteren we de precisie en de recall van de disaggregatie methodes. Deze scoregegevens worden vaak gebruikt om de classificatie van binaire datasets te beoordelen.
 - Precisie: maat voor de positief-voorspellende waarde van een model. Binnen het disaggregatieproject betekent precisie de waarschijnlijkheid dat de disaggregatiemethode het verbruik van water goed voorspelt op het moment dat waterverbruik voorspeld wordt. Precisie wordt berekend volgens: $\text{correcte labels} / (\text{correcte labels} + \text{vals-positieven})$
 - Recall: de maat waarin waterverbruik correct voorspeld wordt door de disaggregatiemethode op het moment dat water in de testdata daadwerkelijk gebruikt wordt. Recall wordt berekend volgens: $\text{correcte labels} / (\text{correcte labels} + \text{vals-negatieven})$
4. Gebeurtenissen onderscheiden: de hoeveelheid gebeurtenissen per eindverbruik over de hele simulatieperiode, per dag, of per uur. Daarnaast is er gekeken naar het moment waarop gebeurtenissen plaatsvinden. Hiervoor is een tijds marge in de evaluatie ingebouwd. Bijvoorbeeld een Wc die om 16:31:25 is doorgespoeld, maar om 16:31:34 uit de disaggregatie komt, zou nog steeds goed gerekend worden met een tijds marge van 10 seconden.

Er is gekozen om alle criteria apart weer te geven, en dus geen weegfactoren toe te kennen aan elk van deze categorieën om tot een totaalscore te komen. Voor meer detail en de manier waarop deze criteria gecommuniceerd zijn met de deelnemende partijen zie Bijlage II.

7 Resultaten

Een overzicht van de disaggregatie resultaten is te zien in Tabel 11. Twee partijen, A en C, hebben op eigen initiatief twee verschillende versies opgestuurd. Binnen A_1 is geen gebruik gemaakt van de compositie van de huishoudens of het startprotocol van de eindverbruiken. In A_2 is deze voorkennis wel gebruikt. Daarnaast verschillen de twee oplossingen in de manier waarop de algoritmes werken (zie Bijlage IV). Binnen C_1 is geen rekening gehouden met mogelijkheid dat de WC-spoeling onderbroken wordt, bij C_2 wordt hier wel rekening mee gehouden. De resultaten van disaggregatie C_1 dienen beschouwd te worden als een voorbeeld van het effect dat zulke mechanismen kunnen hebben op de uitkomsten. Het negeren van de onderbreking van de WC-spoeling zal er namelijk toe leiden watergebruik aan andere eindgebruiken wordt toegewezen.

Tabel 11 Overzicht van de resultaten naar voorkennis en gebruik van trainingsdata

Participant	Disaggregatie	Household info	Training data
A	A_1	-	Unlabeled pulse data, Training tap log
	A_2	Composition data, Appliance data	Unlabeled pulse data, Training tap log, Startup protocol
B	B	-	Training tap log
C	C_1*	Composition data, Appliance data	Startup protocol
	C_2	Composition data, Appliance data	Startup protocol

*C_1 does not use flush interruption for the toilet and illustrates the possible effect of this mechanism on the disaggregation results.

De deelnemers is gevraagd om drie bestanden aan te leveren:

1. appliance_data.csv: een bestand met een lijst van de verschillende eindverbruiken voor elk van de huishoudens in de dataset (Tabel 12). Dit bestand hoefde alleen aangeleverd te worden wanneer deelnemers geen voorkennis over de aanwezigheid van de eindverbruiken ontvingen.

Tabel 12 Voorbeeld van appliance_data.csv, een lijst van eindverbruiken voor elk huishouden bepaald door de disaggregatiemethode

House ID	Appliances
11926619-0816-46f7-b475-d52f807b469a	['Wc', 'BathroomTap', 'WashingMachine', 'KitchenTap', 'Shower', 'Dishwasher', 'OutsideTap']
37b6b8ed-4b6d-41fc-8d58-7ab4d47e3ae1	['BathroomTap', 'Wc', 'KitchenTap', 'OutsideTap', 'Shower', 'WashingMachine', 'Dishwasher', 'Bathtub']
52d569a7-fca6-43d5-b88c-2d8932e2e758	['BathroomTap', 'Wc', 'KitchenTap', 'OutsideTap', 'Shower', 'WashingMachine', 'Dishwasher', 'Bathtub']
5457ca6f-d8b3-494f-9740-5bacc783cbcb	['Wc', 'KitchenTap', 'BathroomTap', 'Dishwasher', 'Shower', 'WashingMachine']
6fca3589-1769-48e8-a650-5458df9cf102	['Wc', 'KitchenTap', 'BathroomTap', 'Shower', 'Dishwasher', 'WashingMachine', 'OutsideTap', 'Bathtub']
7d812407-51f3-4ad0-b384-17b1b16ae0b1	['Wc', 'BathroomTap', 'KitchenTap', 'Shower', 'Dishwasher', 'WashingMachine', 'OutsideTap', 'Bathtub']

2. labeled_pulses.csv: een tabel waarin met '1' aangegeven is welke eindverbruiken verantwoordelijk zijn voor de pulsen in de test data.

Tabel 13 Voorbeeld van labeled_pulses.csv, een '1' geeft aan dat het eindverbruik verantwoordelijk is voor de puls in de testdata

time	day	Wc	BathroomTap	Shower	KitchenTap	WashingMachine	Dishwasher	OutsideTap
06:12:50	8	0	0	0	0	0	0	0
06:12:51	8	0	0	0	0	0	0	0
06:12:52	8	1	0	0	0	0	0	0
06:12:53	8	1	0	0	0	0	0	0
06:12:54	8	1	0	0	0	0	0	0
06:12:55	8	1	0	0	0	0	0	0
06:12:56	8	1	0	0	0	0	0	0

3. tap_log_processed.csv: Een lijst met momenten van waterverbruik door de verschillende eindverbruiken.

Tabel 14 Voorbeeld van tap_log_processed.csv

time	Day	weekday	end use	duration(s)	water use (L)
06:12:52	8	Weekday	Wc	48	6
06:27:32	8	Weekday	BathroomTap	112	5.2
07:25:56	8	Weekday	Wc	49	6.4
07:26:45	8	Weekday	BathroomTap	9	0.1
07:34:38	8	Weekday	Wc	48	6
07:45:41	8	Weekday	Wc	48	6

7.1 Aanwezige apparaten/eindverbruiken

A en C hebben informatie over de samenstelling van de huishoudens en de aanwezige eindverbruiken ontvangen. A heeft aangegeven dat hun eerste oplossing (A_1) deze voorkennis niet gebruikt. Partij B heeft ervoor gekozen om deze data niet te ontvangen. Partij C heeft aangegeven de samenstelling van het huishouden niet te gebruiken in hun algoritme. Met of zonder deze informatie is het van belang om te evalueren in hoeverre de eindverbruiken in de disaggregatie overeenkomen met die van de testdata. Het kan namelijk voorkomen dat een huishouden wel een wasmachine heeft, maar dat deze binnen de simulatieperiode niet gebruikt wordt. Wanneer er een exacte overeenkomst is dan is dit een 'Matches'. Daarnaast kan het voorkomen dat een eindverbruik wel in de originele data voorkomt maar niet in de disaggregatieresultaten, d.w.z. 'Under identifications' of andersom 'Overidentifications'. Een overzicht van de prestatie van de disaggregatiemethodes bij het identificeren van eindverbruiken is weergegeven in Tabel 15.

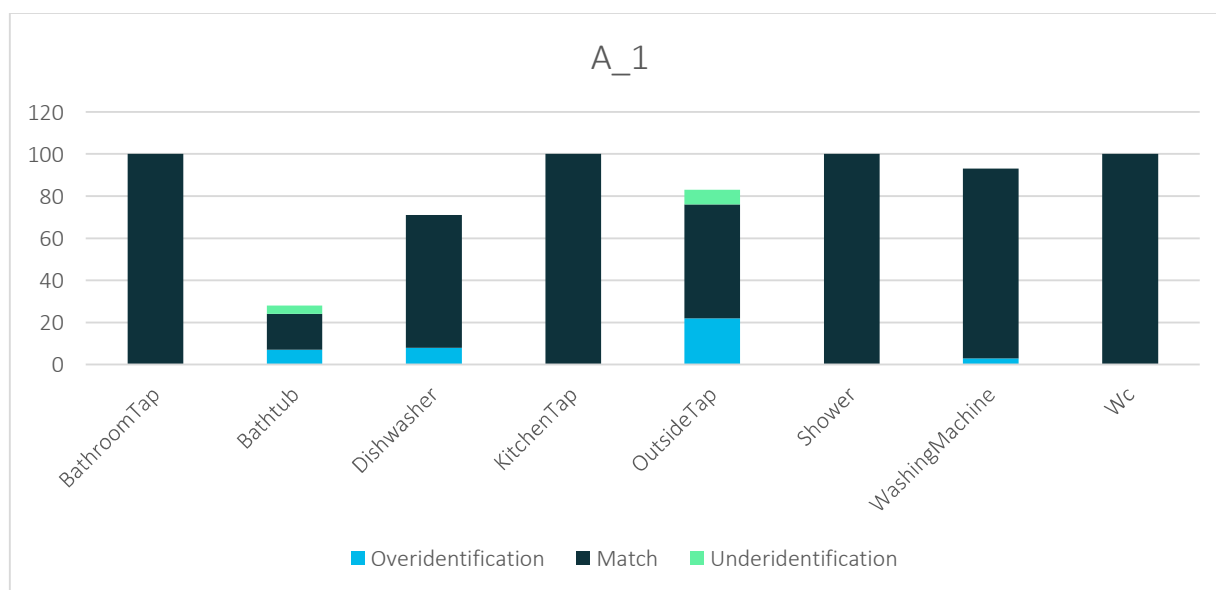
Tabel 15 Overzicht van de matches, over- en onderidentificatie van de disaggregatiemethodes.

Methode	Matches	Overidentifications	Underidentifications	Match percentage*
A_1	624	40	11	92%
A_2	615	2	20	96%
B	632	80	3	87%
C_1**	553	4	82	84%
C_2	553	4	82	84%

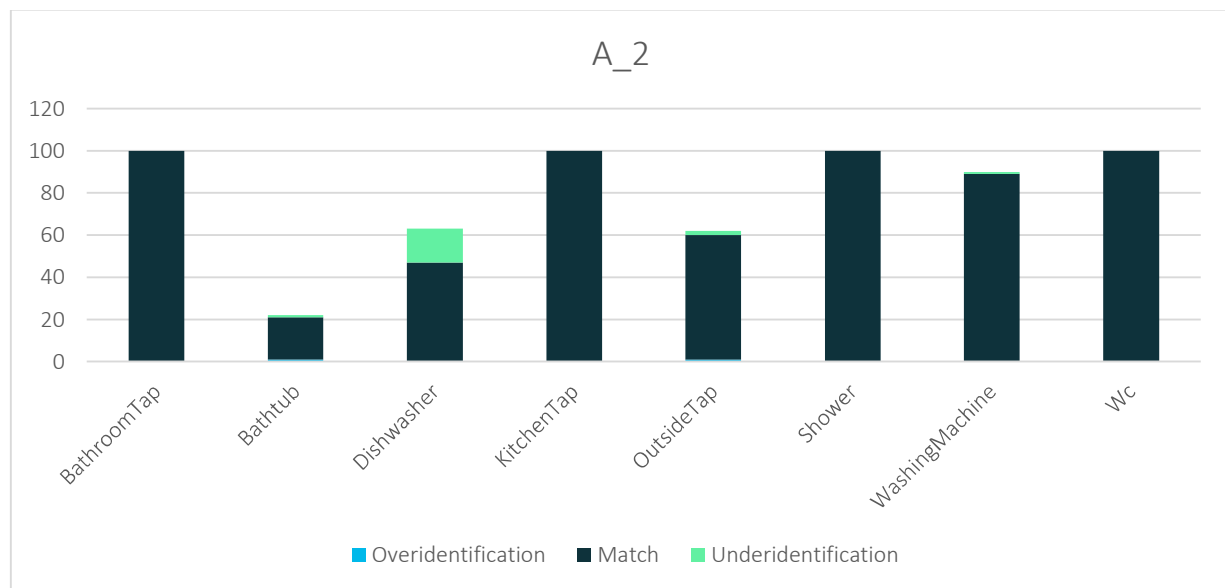
*Matchpercentage is calculated with: $1 - (\text{overidentifications} + \text{underidentifications}) / \text{matches}$

**C_1 does not use flush interruption for the toilet and illustrates the possible effect of this mechanism on the disaggregation results.

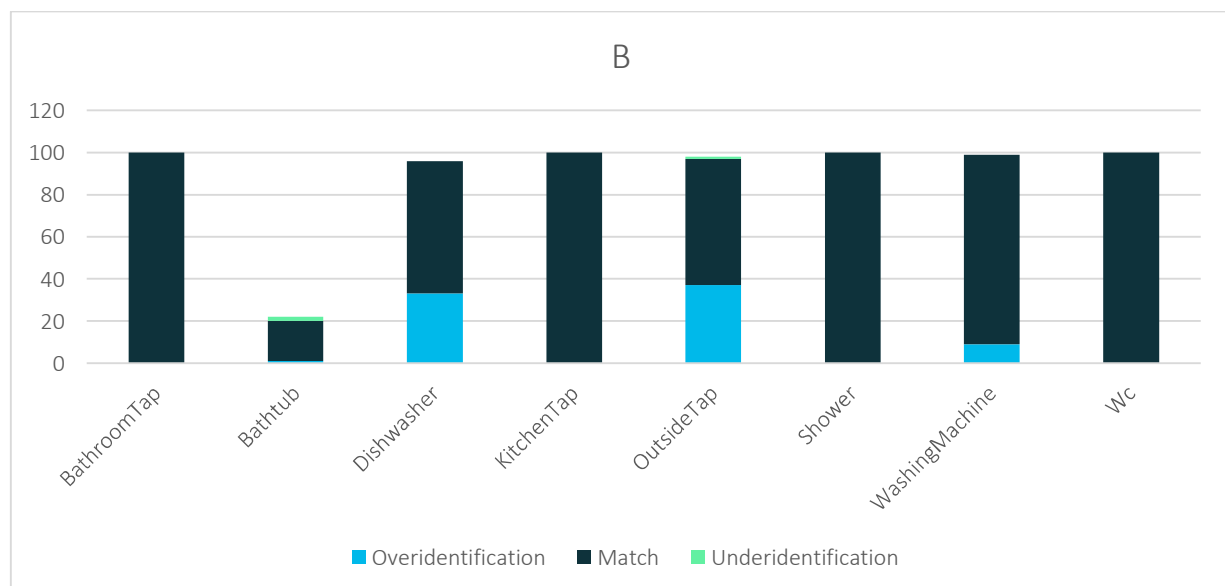
In onderstaande figuren (Figuur 4 - Figuur 8) is te zien hoe vaak de resultaten van de verschillende disaggregatiemethoden wel of niet overeenkomen met de originele data voor elk van de eindverbruiken. Voor de eindverbruiken waarbij de penetratiegraad 100% is (BathroomTap, KitchenTap, Shower, Wc) worden geen fouten gemaakt door de disaggregatiemethoden. Wanneer er geen voorkennis gebruikt wordt, komt het geregeld voor dat er in de resultaten van de disaggregatie eindverbruiken gerapporteerd worden die niet in de originele data aanwezig waren (Overidentification), zie A_1 (Figuur 4) en B (Figuur 6). Met voorkennis gebeurt dit niet (A_2) of zelden (C_1 en C_2). Met of zonder voorkennis komt het tegenovergestelde voor. In het geval van C_1 en C_2 worden 'Bathtub' en 'OutsideTap' helemaal niet meegenomen in de resultaten, de voorkennis over hun aanwezigheid was wel aangeleverd maar is niet gebruikt tijdens de disaggregatie.



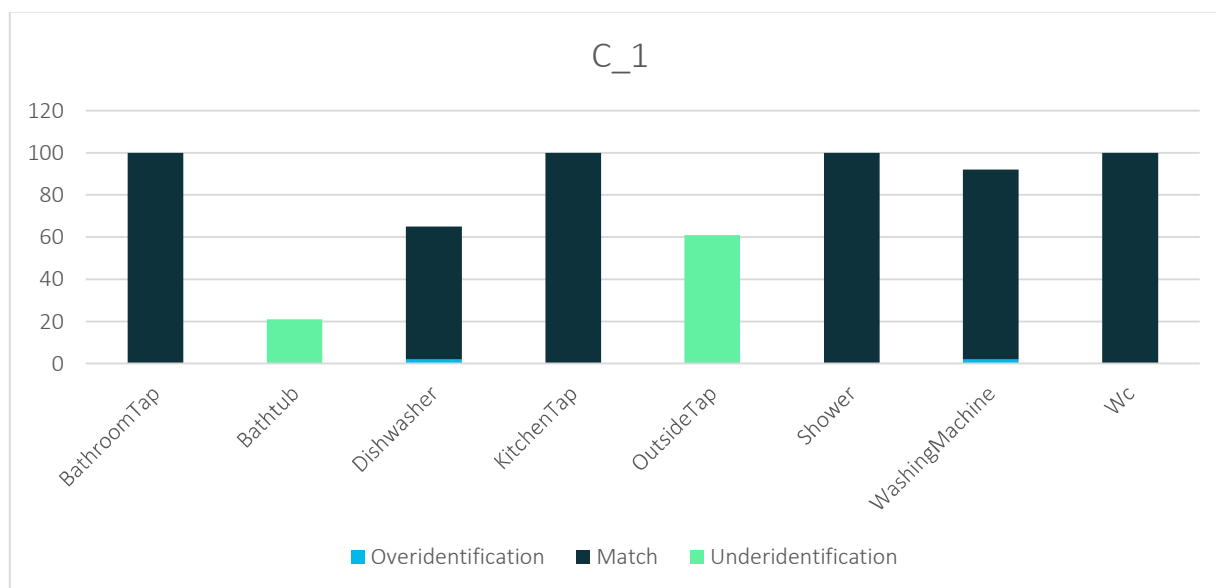
Figuur 4 Aanwezigheid van eindverbruiken ten opzichte van de originele data voor A_1



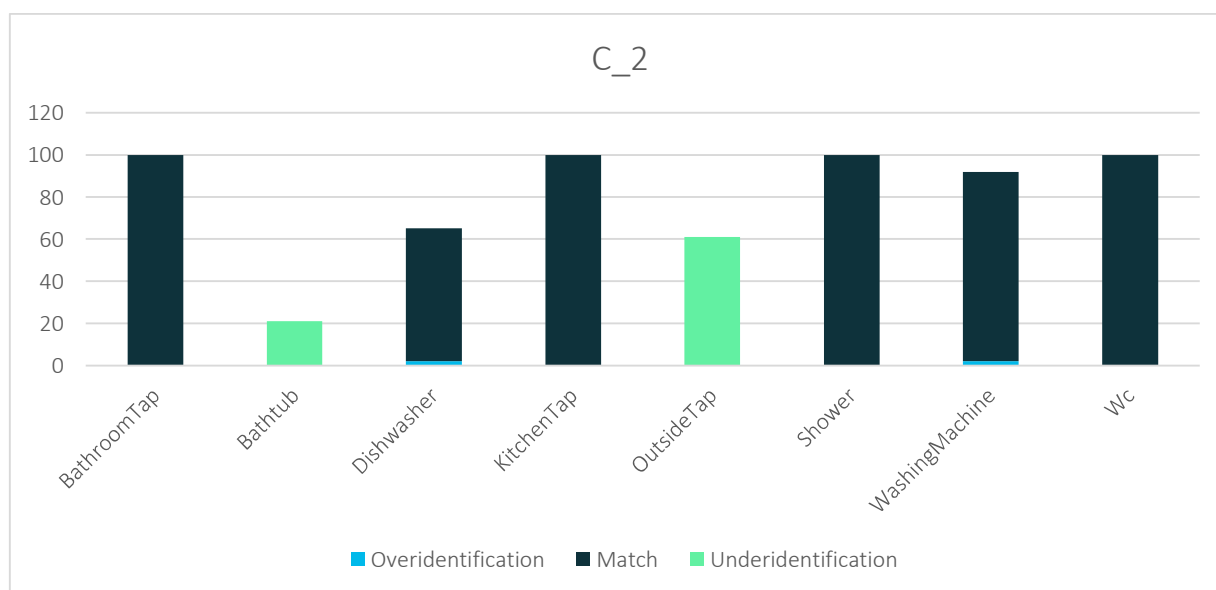
Figuur 5 Aanwezigheid van eindverbruiken ten opzichte van de originele data voor A_2



Figuur 6 Aanwezigheid van eindverbruiken ten opzichte van de originele data voor B



Figuur 7 Aanwezigheid van eindverbruiken ten opzichte van de originele data voor C_1, C_1 maakt geen gebruik van spoelonderbreking voor de WC en illustreert het mogelijke effect van dit mechanisme op de disaggregatieresultaten.



Figuur 8 Aanwezigheid van eindverbruiken ten opzichte van de originele data voor C_2

7.2 Waterverbruik per huishouden

Binnen elk huishouden is bepaald hoeveel water correct (c) is toegewezen aan elk eindverbruik. Het toegewezen water kan meer zijn dan het origineel (above original, ao) of onder het origineel (below original, bo). De hoeveelheid correct toegewezen water wordt bepaald volgens: $c = 1 - (ao + bo)/tw$.

Bijvoorbeeld: voor een hypothetisch huis met een totaal waterverbruik van 200 L (WC = 150 L, en Douche = 50 L) waarvan een disaggregatie de WC 100 L en de douche 100 L toewijst scoort $1 - \frac{50+50}{200} = 0.5 \rightarrow 50\%$. Dat wil zeggen, 50% van het water is correct toegewezen.

Het totale waterverbruik (total water use, tw) van de 100 huishoudens over de simulatieperiode van 7 dagen is 184675 liter. Alle disaggregatiemethodes zijn in staat om meer dan 70% van het totale waterverbruik correct toe te

wijzen op basis van $\sum ao$ en $\sum bo$ voor alle huishoudens (Tabel 16). Het valt daarnaast op dat alle disaggregatiemethodes afwijken in hun tw waarde ten opzichte van de ontvangen data, wat aangeeft dat de waterbalans in de disaggregatiemethoden niet 100% sluitend is. Methodes die in totaal meer water identificeren dan in het origineel (A_1) karakteriseren meer water boven hun werkelijke waarde. Hetzelfde geldt andersom voor methodes die in totaal minder water identificeren.

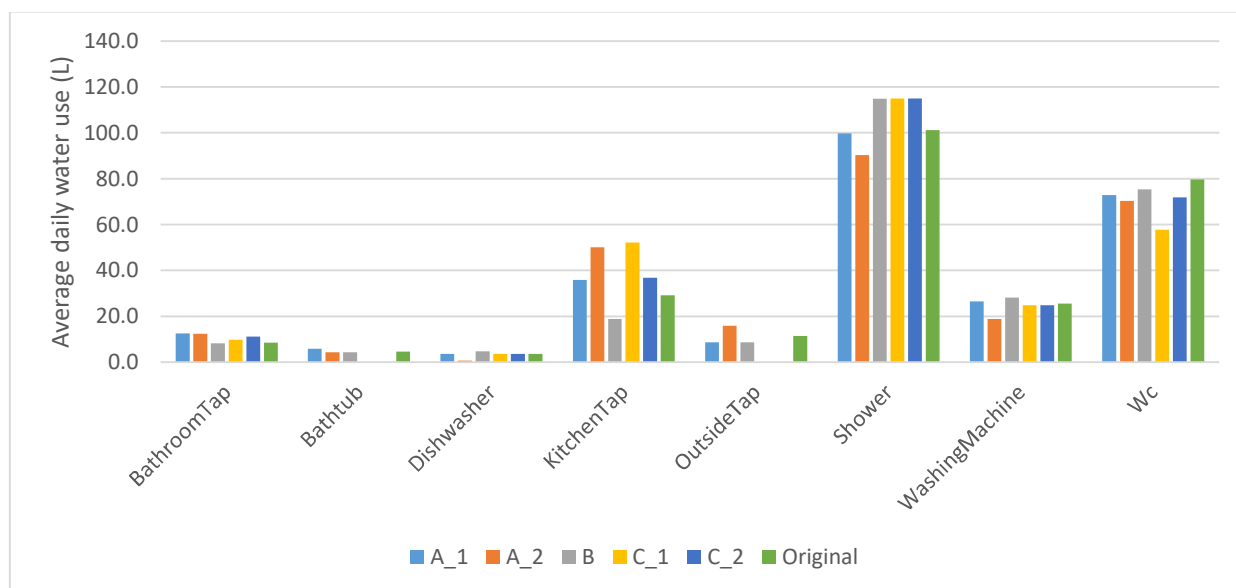
Tabel 16 Overzicht van de resultaten uit de disaggregatiemethoden in relatie tot de originele data. De waardes zijn de som van het totale waterverbruik voor de 100 huishoudens voor de 7 dagen in de simulatieperiode.

Method	Disaggregation total water use (L)	Disaggregation water use above original $\rightarrow \sum ao$ (L)	Disaggregation water use below original $\rightarrow \sum bo$ (L)	Correct $\rightarrow c$ (%)
A_1	185937	9055	7794	91%
A_2	183970	20495	21201	77%
B	184248	12132	12559	87%
C_1*	184227	26625	27074	71%
C_2	184227	16794	17243	82%

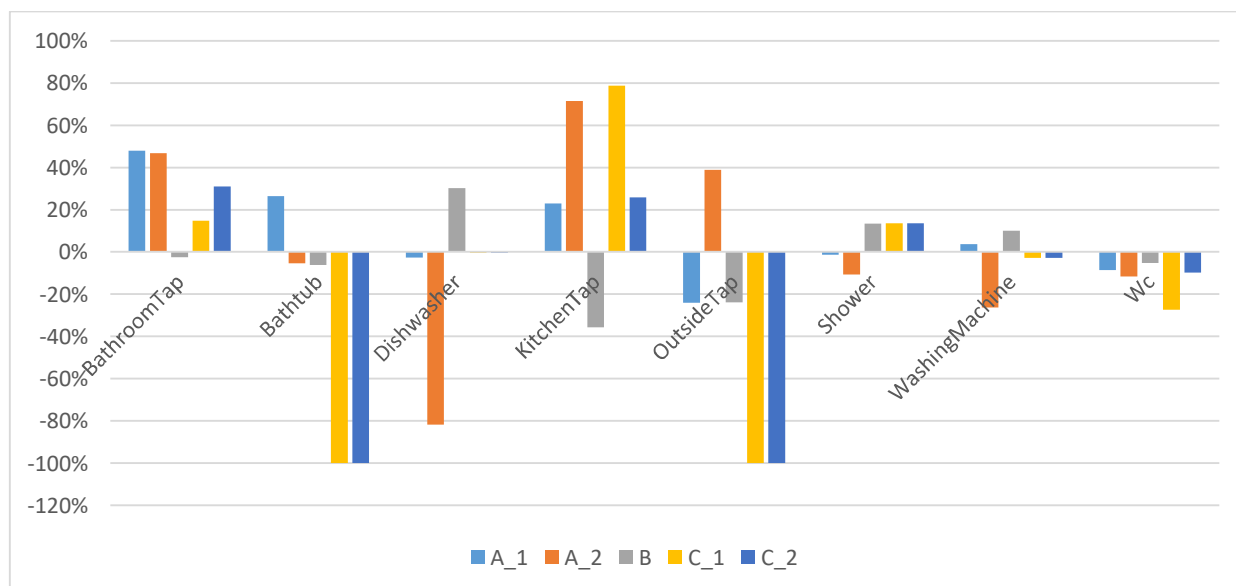
*C_1 does not use flush interruption for the toilet and illustrates the possible effect of this mechanism on the disaggregation results.

Figuur 9 tot Figuur 11 geven een overzicht van de prestaties van de verschillende disaggregatiemethoden in relatie tot het volume waterverbruik in de originele data. Het totale waterverbruik voor de disaggregaties en het origineel is te zien in Figuur 9. Het verschilpercentage tussen de disaggregatie en de originele data in relatie tot elk eindverbruik is te zien in Figuur 10. Bijvoorbeeld: A_1 rapporteert een gemiddeld waterverbruik voor de BathroomTap van 12.6 L per huishouden, terwijl in dit in de originele data 8.5 is, een verschil van 48%. De disaggregatie van watermeterdata is ook belangrijk in relatie tot het totale waterverbruik (Figuur 11). Bijvoorbeeld: A_1 rapporteert een gemiddeld waterverbruik voor de BathroomTap van 12.6 L per huishouden, terwijl in dit in de originele data 8.5 is en het totale waterverbruik 264 L per dag is dan is de foutmarge 1.6% ten opzichte van het totaal verbruik. Een foutmarge van 10% bij de Wc is veel meer water dan dezelfde foutmarge bij de buitenkraan.

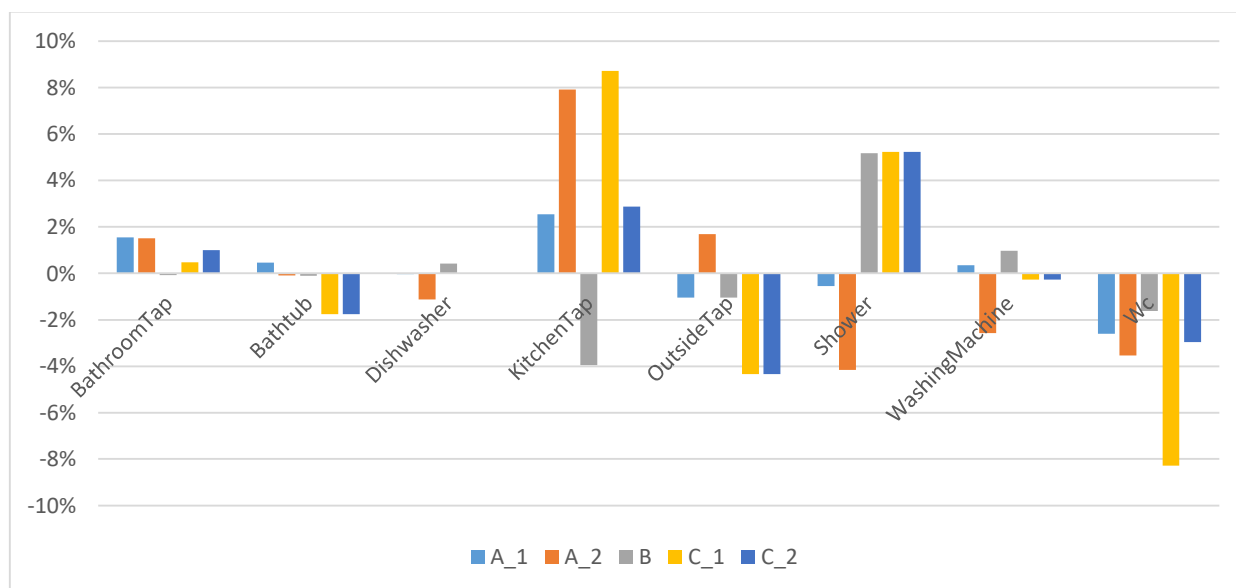
In onderstaande figuren is te zien dat de verschillende disaggregatiemethodes soms het verbruik onderschatten en soms overschatten. Tussen de resultaten van A valt het volume van de kranen op. A_1 kent kranen over het algemeen een lager volume toe vergeleken met A_2. Om de waterbalans kloppend te houden wordt er in A_1 een hoger volume toegekend aan de douche, wasmachine en wc. Daarnaast is te zien dat de toevoeging van de mogelijkheid om de Wc-spoelbeurt te onderbreken (de kleine knop/stopknop) in C_2 zorgt voor een aanzienlijk betere prestatie binnen de Wc zelf. Deze toevoeging zorgt er ook voor dat de prestatie van de keukenkraan verbetert.



Figuur 9 Overzicht van het dagelijks waterverbruik per huishouden op basis van de disaggregatiemethoden en uit de originele dataset

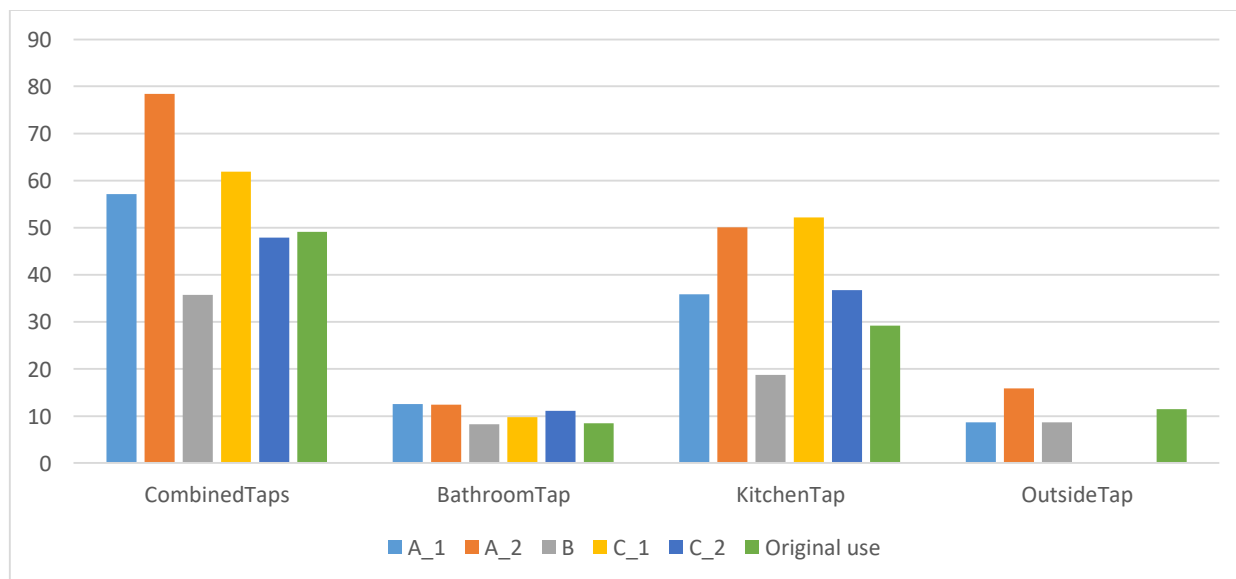


Figuur 10 Verschil in waterverbruik tussen de disaggregatieresultaten en de originele data per eindverbruik



Figuur 11 Verskil in waterverbruik tussen de disaggregatieresultaten en de originele data in relatie tot het totale waterverbruik

Wanneer we naar de resultaten kijken als de verschillende kranen (keukenkraan, badkamerkraan, buitenkraan) samengenomen worden dan komt de disaggregatie methode C_2 het dichtst bij het originele waterverbruik (97%), zie Figuur 12. Het onderscheid tussen verschillende kranen is waarschijnlijk moeilijker te maken in een dataset van daadwerkelijk waterverbruik ten opzichte van de synthetische data in dit onderzoek.



Figuur 12 Waterverbruik van de disaggregatiemethoden en de originele data wanneer kranen gecombineerd worden

Op basis van de volumestroom is er geen duidelijk trend te zien tussen hoe goed de disaggregatiemethoden werken en het aantal mensen dat een huis bewonen, zie Bijlage V. Er kunnen hoogstens een aantal kanttekeningen geplaatst worden over de prestaties van de disaggregatiemethoden in relatie tot de hoeveelheid bewoners:

- De disaggregatiemethoden presteren constant voor de vaatwasser en de wasmachine onafhankelijk van de hoeveelheid bewoners en hebben over het algemeen een kleine afwijking ten opzichte van de testdata.
- Het bad, de buitenkraan, en de WC zijn de eindverbruiken waarbij het verschil met de testdata meer onderschat wordt naarmate er meer bewoners zijn.
- De douche en de keukenkraan worden meer overschat naarmate er meer bewoners zijn.

Naast het algehele waterverbruik is ook het waterverbruik per dag en per uur vergeleken met de originele data, voorbeelden van A_1 in Tabel 17 en Tabel 18. Een compleet overzicht van deze resultaten is te vinden in Bijlage V.

Tabel 17 Verschil tussen de disaggregatieresultaten en de originele data in het dagelijkse waterverbruik voor A_1. Hoe geler de achtergrond hoe hoger de disaggregatiewaarde boven de originele waarde valt; blauw onder de originele waarde.

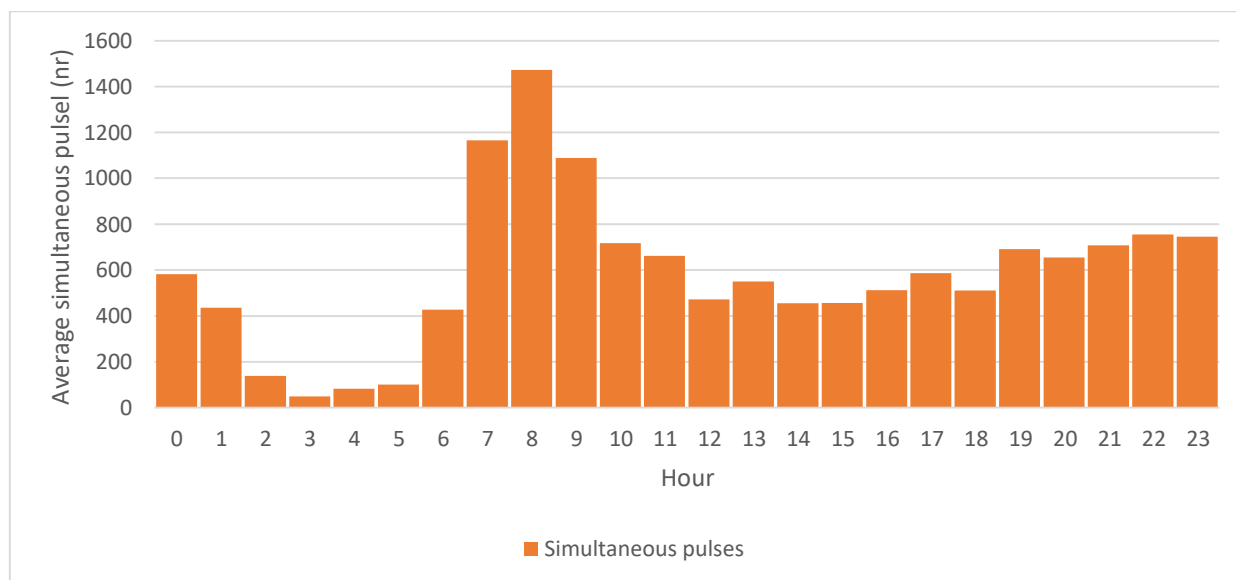
day	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
8	0.27%	0.02%	0.00%	0.37%	-0.23%	-0.08%	0.06%	-0.31%
9	0.26%	0.00%	0.00%	0.47%	-0.05%	-0.12%	0.00%	-0.46%
10	0.18%	0.25%	0.00%	0.13%	-0.42%	0.08%	0.13%	-0.29%
11	0.24%	0.14%	-0.01%	0.28%	-0.13%	-0.01%	-0.02%	-0.43%
12	0.21%	-0.04%	-0.01%	0.30%	-0.02%	-0.18%	0.10%	-0.34%
13	0.18%	-0.06%	0.00%	0.47%	-0.10%	-0.09%	0.03%	-0.33%
14	0.21%	0.17%	-0.02%	0.52%	-0.09%	-0.15%	0.05%	-0.44%

Tabel 18 Verschil tussen de disaggregatieresultaten en de originele data op uur basis voor A_1. Hoe geler de achtergrond hoe hoger de disaggregatiewaarde boven de originele waarde valt; blauw onder de originele waarde.

hour	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
0	-0.03%	0.00%	0.00%	0.18%	0.01%	-0.03%	0.02%	-0.14%
1	0.01%	0.00%	0.00%	0.18%	-0.06%	0.01%	-0.01%	-0.12%
2	0.00%	0.00%	0.00%	0.08%	-0.03%	0.00%	0.00%	-0.05%
3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
4	0.00%	0.00%	0.00%	0.02%	-0.01%	0.01%	0.00%	-0.01%
5	0.00%	-0.06%	0.00%	0.03%	0.03%	0.02%	0.00%	-0.01%
6	0.05%	0.00%	0.00%	0.04%	-0.02%	-0.01%	0.02%	-0.09%
7	0.18%	0.09%	0.01%	0.19%	0.03%	-0.08%	-0.02%	-0.28%
8	0.21%	0.23%	0.01%	0.29%	-0.04%	-0.34%	0.12%	-0.32%
9	0.09%	0.18%	0.00%	0.20%	-0.05%	-0.07%	0.00%	-0.25%
10	0.10%	0.13%	0.00%	0.17%	-0.08%	-0.05%	-0.04%	-0.11%
11	0.16%	-0.06%	0.00%	-0.03%	-0.06%	0.06%	0.03%	-0.08%
12	0.12%	-0.06%	0.00%	-0.08%	0.04%	0.00%	-0.01%	0.02%
13	0.07%	0.00%	0.01%	0.03%	-0.25%	0.23%	0.05%	-0.11%
14	0.04%	0.00%	0.00%	0.08%	-0.08%	-0.01%	0.04%	-0.07%
15	0.05%	0.00%	0.00%	0.18%	-0.11%	-0.04%	0.00%	-0.08%
16	0.13%	0.00%	0.00%	0.05%	-0.02%	-0.04%	0.02%	-0.09%
17	0.10%	0.00%	0.01%	-0.04%	0.06%	-0.04%	0.03%	-0.10%
18	0.12%	0.00%	-0.01%	-0.03%	-0.01%	-0.06%	0.02%	-0.08%
19	0.06%	-0.06%	-0.01%	0.33%	-0.04%	-0.10%	0.02%	-0.15%
20	0.02%	-0.04%	-0.01%	0.20%	0.01%	-0.04%	-0.03%	-0.08%
21	0.03%	-0.03%	-0.01%	0.14%	-0.07%	0.03%	0.02%	-0.13%
22	0.04%	0.10%	-0.02%	0.06%	-0.12%	0.06%	-0.01%	-0.10%
23	0.00%	0.06%	-0.02%	0.28%	-0.18%	-0.04%	0.08%	-0.18%

De verschillende dagen of uren kunnen met elkaar vergeleken worden door eerst de data te normaliseren (z-score normalisatie, (waarde – gemiddelde van dataset)/standard deviatie van dataset) en daarna de absolute afstand van het gemiddelde te nemen. Op deze manier kunnen de dagen/uren geïdentificeerd worden met de grootste afwijking van het gemiddelde. In de meeste disaggregatieresultaten is dag 10 het verst verwijderd van het gemiddelde, maar is niet duidelijk te achterhalen wat de achterliggende oorzaak hiervoor is. Over het algemeen is

het moment met de grootste afwijking 8 uur in de ochtend. Alleen voor C_2 verschuift dit naar 7 uur in de ochtend. Dit is ook de periode van de dag wanneer simultaan gebruik door meerdere eindverbruiken het meest voorkomt (Figuur 13). Een direct causaal verband tussen de prestatie van de disaggregaties en de hoeveelheid simultaan waterverbruik is echter moeilijk te bewijzen. De hoeveelheid simultaan gebruik neemt namelijk ook toe naarmate de hoeveelheid bewoners toeneemt. De toename in de hoeveelheid bewoners veroorzaakte geen duidelijke trend in de prestaties van de disaggregaties.



Figuur 13 Gemiddelde hoeveelheid pulsen per uur met simultaan waterverbruik in de originele dataset

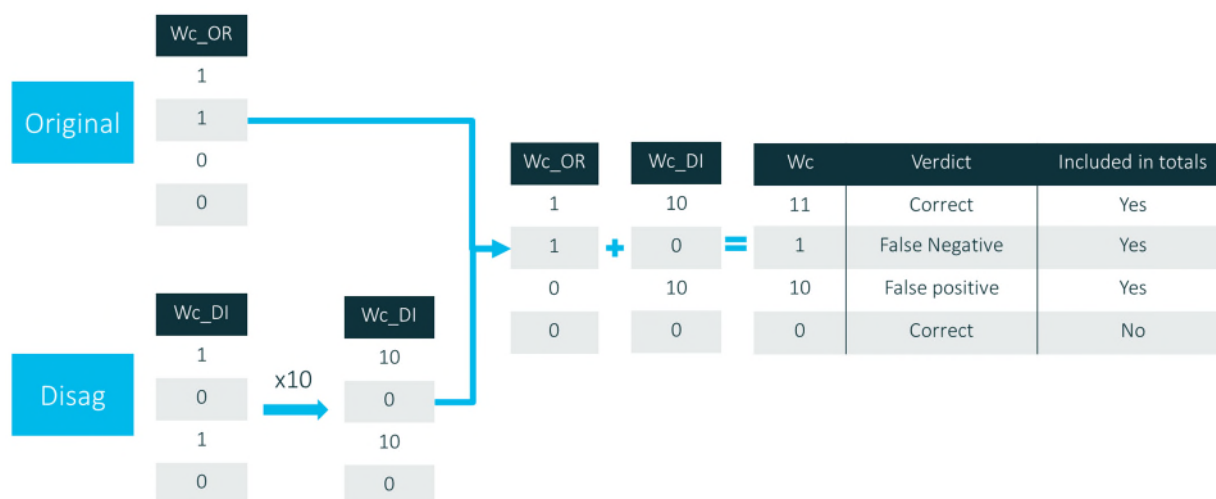
7.3 Labelen van de dataset

Bij het labelen van de dataset is het de bedoeling dat een of meerdere eindverbruiken worden toegewezen aan elk van de pulsen in de testdata. De labels zijn binair (0) geen waterverbruik of (1) wel waterverbruik door desbetreffend eindverbruik. De uitkomst van het labelen kan zijn:

- Correct: het label uit de disaggregatie komt overeen met het label uit de dataset – beiden labelen met 1 of met 0
- Vals negatief: de disaggregatie geeft aan dat een bepaald eindverbruik geen water verbruikt terwijl dat in de dataset wel zo is – disaggregatie (0), dataset (1)
- Vals positief: de disaggregatie geeft aan dat een bepaald eindverbruik wel water verbruikt terwijl dat in de dataset niet zo is – disaggregatie (1), dataset (0)

De momenten waarop er geen puls toegekend wordt door de disaggregatie (0) noch door de dataset (0) zijn in principe wel correct maar zijn uitgesloten van de totale score. Reden hiervoor is dat anders 99% van de pulsen correct gelabeld zijn omdat er op de meeste momenten geen waterverbruik plaatsvindt.

De evaluatiemethode voor het labelen is uitgebeeld in Figuur 14. De labels uit de disaggregatie worden vermenigvuldigd met 10 en daarna opgeteld bij de originele data. Afhankelijk van deze som kan bepaald worden of het label correct, vals positief, of vals negatief is.



Figuur 14 Visualisatie van de evaluatiemethode voor het labelen van de pulsen

Een overzicht van de labelprestatie is weergegeven in Tabel 19. De originele dataset telt in totaal 1401608 labels met waarde '1'. Alle disaggregatiemethodes kennen meer '1' labels toe dan in het origineel. De hoeveelheid correcte labels ligt tussen de 64% en 77%. De gemiddelde hoeveelheid vals negatieven en vals positieven is 15% en 16% respectievelijk.

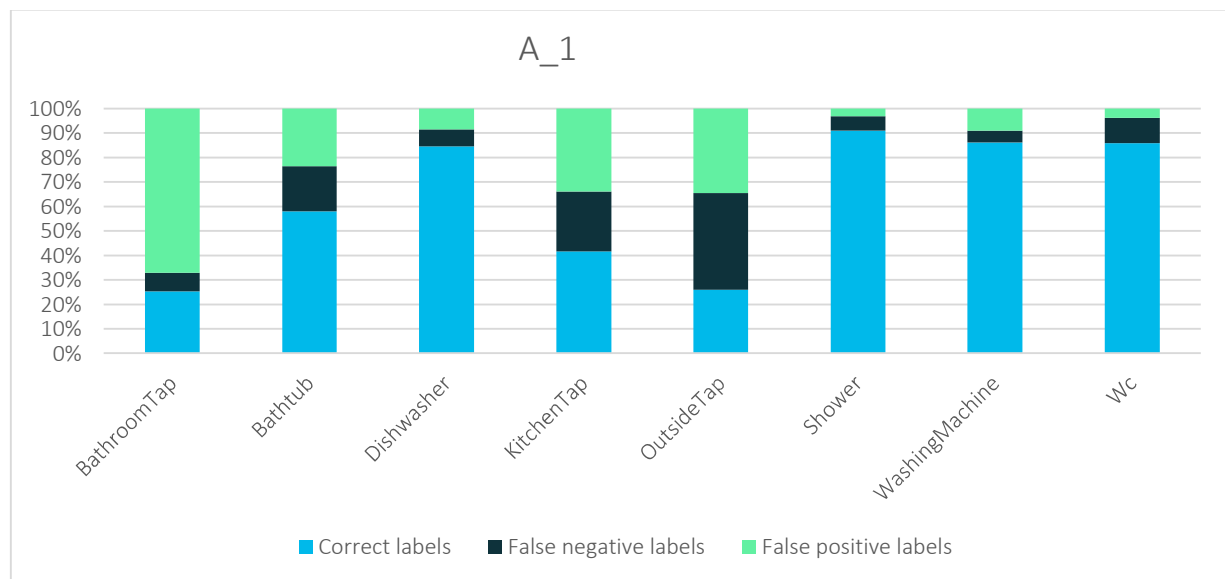
Tabel 19 Overzicht van de label prestatie van de verschillende disaggregatie methodes

Method	Total labels	Label count (% of original)	Correct label		False negatives		False positives	
A_1	1699896	121%	1200723	71%	200885	12%	298288	18%
A_2	1784246	127%	1156415	65%	245193	14%	382638	21%
B	1552571	111%	1198151	77%	203457	13%	150963	10%
C_1*	1688663	120%	1078945	64%	322663	19%	287055	17%
C_2	1613269	115%	1155158	72%	246450	15%	211661	13%

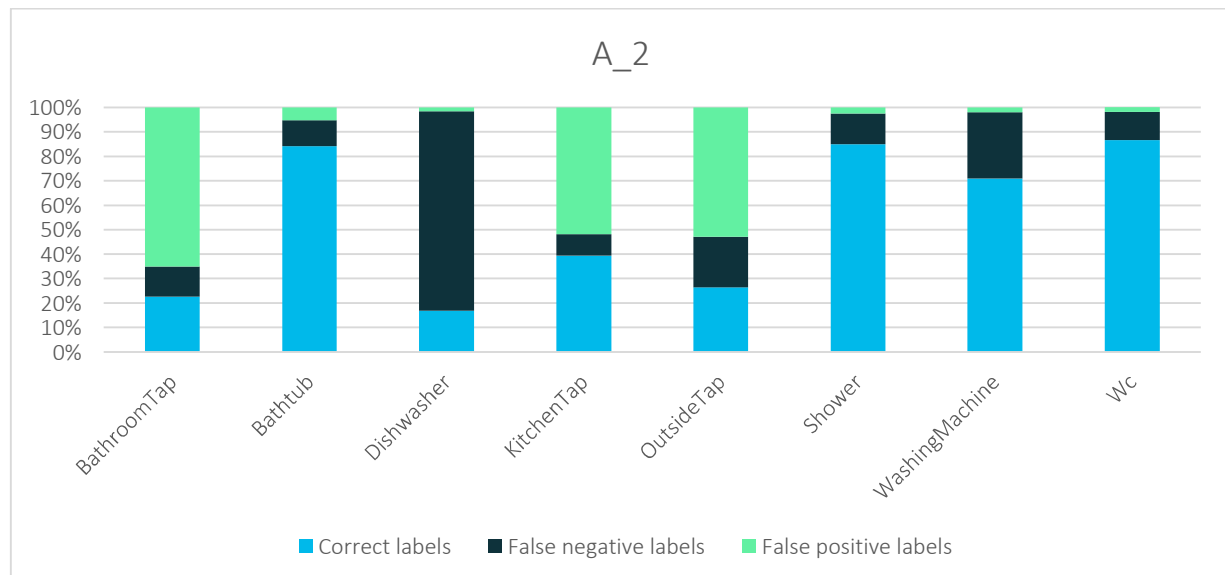
*C_1 does not use flush interruption for the toilet and illustrates the possible effect of this mechanism on the disaggregation results.

De gedetailleerde uitkomsten voor het labelen zijn weergegeven in Figuur 15 tot en met Figuur 19. Uit de gemiddelde scores voor elk van de disaggregaties blijkt:

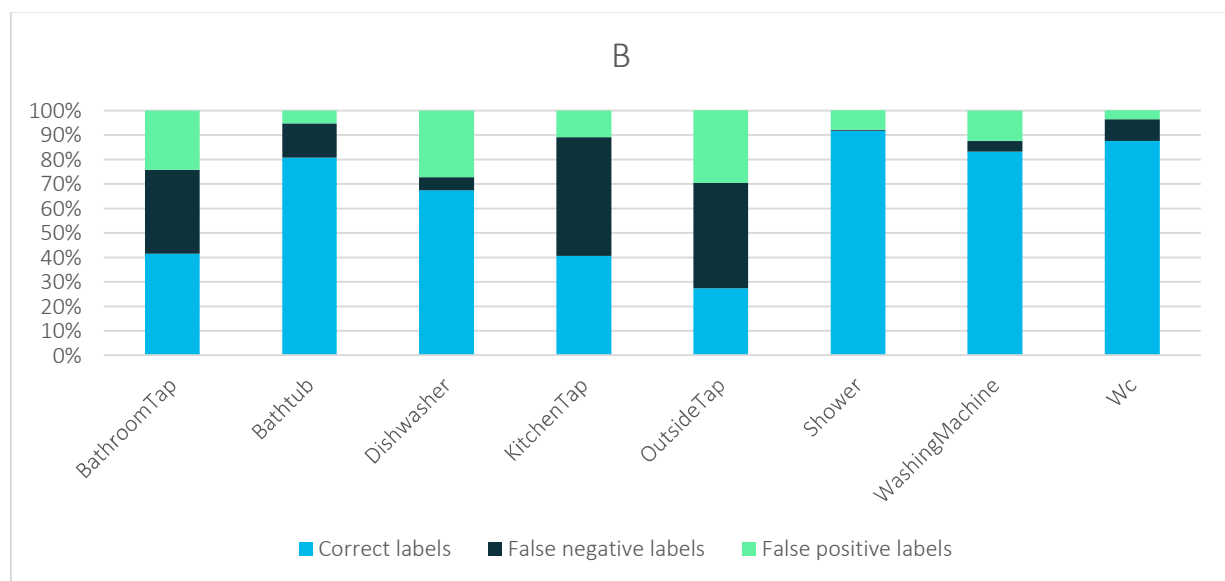
- Over het algemeen worden de douche (87%), vaatwasser (66%), en wasmachine (85%) het vaakst correct gelabeld.
- Vals negatieven komen het vaakst voor bij de buitenkraan (61%) en het bad (49%). Wanneer C_1 en C_2 uit deze analyse weggelaten worden (omdat deze de buitenkraan niet meenemen) dan is de buitenkraan (34%) nog steeds het eindverbruik met de meeste vals negatieven, gevolgd door de vaatwasser (31%) en de keukenkraan (27%). De gemiddelde prestatie voor de vaatwasser wordt sterk beïnvloed door disaggregatie A_2, met een vals negatief score van 82%.
- Vals positieven komen het vaakst voor bij de badkamerkraan (45%) en de keukenkraan (37%).



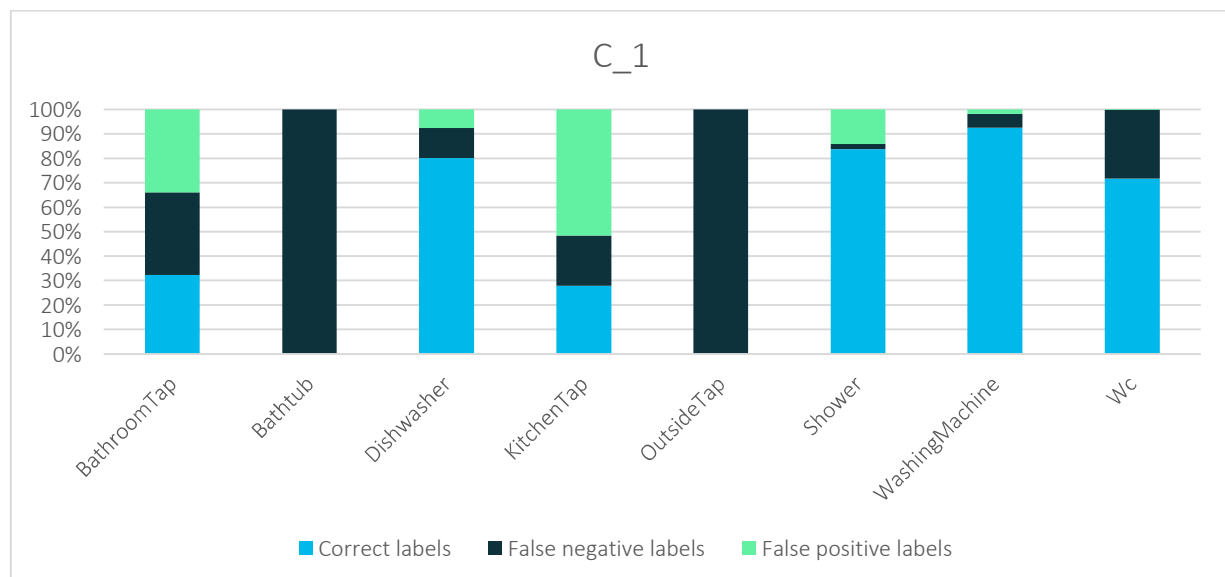
Figuur 15 Label prestatie van A_1



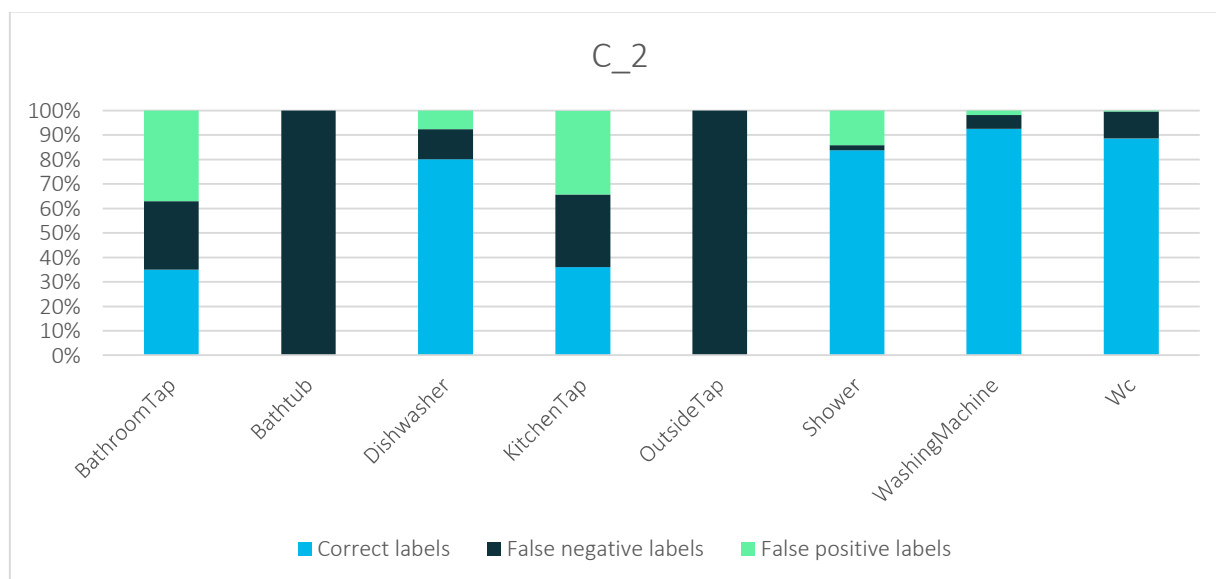
Figuur 16 Label prestatie van A_2



Figuur 17 Label prestatie van B

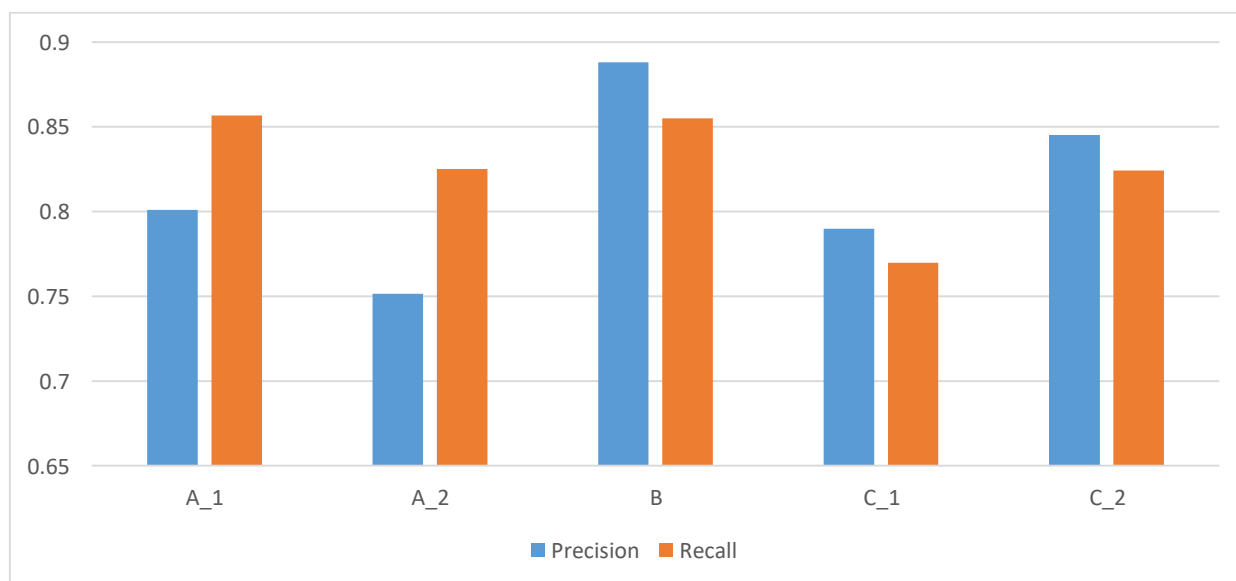


Figuur 18 Label prestatie van C_1, C_1 maakt geen gebruik van spoelonderbreking voor de WC en illustreert het mogelijke effect van dit mechanisme op de disaggregatieresultaten.



Figuur 19 Label prestatie van C_2

Uit de correcte labels, vals positieven en vals negatieven worden de precisie en de recall berekend. Over het algemeen scoort de methode van partij B hoog op beide aspecten (Figuur 20), maar zijn er significante verschillen tussen de precisie en recall voor elk van de eindverbruiken.



Figuur 20 Algemene precisie en recall score voor de disaggregatie methodes.

- Precisie: De hoogste precisie wordt behaald door C_1 en C_2 bij de WC en de wasmachine (Tabel 20). Dat wil zeggen, wanneer men enkel over de resultaten van de disaggregatiemethode zou beschikken, en deze geeft aan dat de WC in gebruik is, dan is er een 99.6% kans dat dit correct is.

Tabel 20 Precisie voor elk eindverbruik per disaggregatiemethode

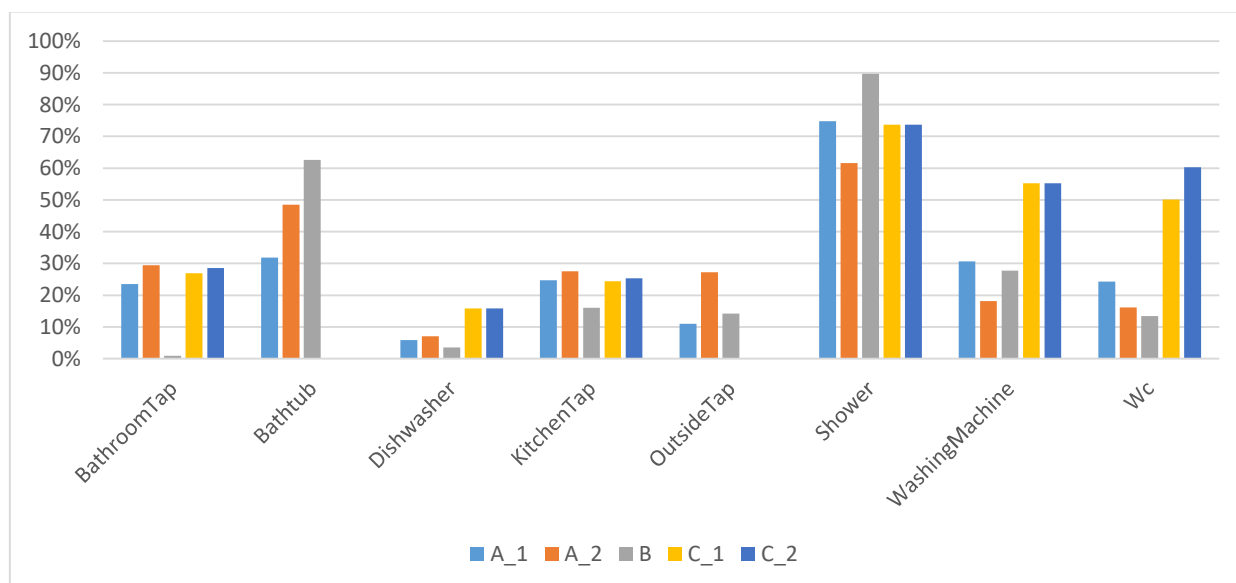
end_use	A_1	A_2	B	C_1	C_2
BathroomTap	27.4%	25.7%	63.1%	48.8%	48.6%
Bathtub	71.2%	94.1%	94.0%	0.0%	0.0%
Dishwasher	90.9%	91.1%	71.3%	91.3%	91.3%
KitchenTap	55.1%	43.2%	78.9%	35.1%	51.3%
OutsideTap	42.9%	33.2%	48.0%	0.0%	0.0%
Shower	96.6%	97.1%	92.1%	85.5%	85.5%
WashingMachine	90.5%	97.2%	87.0%	98.1%	98.1%
Wc	95.7%	97.8%	96.1%	99.6%	99.6%

- Recall: De hoogste recall waarde wordt behaald binnen de douche door disaggregatiemethode B (Tabel 21). Dat wil zeggen, wanneer men alleen beschikt over de testdata, en deze geeft aan dat de douche in gebruik is, dan zou disaggregatiemethode B in 99.5% van de gevallen hetzelfde voorspellen.

Tabel 21 Recall voor elk eindverbruik per disaggregatiemethode

end_use	A_1	A_2	B	C_1	C_2
BathroomTap	76.9%	64.4%	54.7%	48.9%	55.6%
Bathtub	75.9%	88.9%	85.2%	0.0%	0.0%
Dishwasher	92.5%	17.1%	92.6%	86.7%	86.7%
KitchenTap	63.0%	81.5%	45.6%	57.3%	54.9%
OutsideTap	39.6%	56.1%	38.9%	0.0%	0.0%
Shower	93.9%	87.1%	99.5%	97.7%	97.7%
WashingMachine	94.6%	72.5%	95.0%	94.2%	94.2%
Wc	89.2%	88.4%	90.7%	72.0%	89.0%

Binnen het waterverbruik in een huishouden kan er globaal gezien onderscheid gemaakt worden tussen twee situaties: waterverbruik van een enkel eindverbruik en simultaan gebruik van meerdere apparaten. In Figuur 21 is te zien hoe de verschillende methodes presteren wanneer er gekeken wordt naar de subset van momenten met simultaan waterverbruik. De douche presteert binnen alle disaggregaties het best tijdens simultaan gebruik. C_1 en C_2 presteren aanzienlijk beter bij de wasmachine en de WC op momenten van simultaan gebruik dan de andere methodes.



Figuur 21 Percentage correcte labels van de disaggregatie methoden voor de subset van momenten met simultaan waterverbruik

Voor het WT-onderzoek is het van belang om een inschatting te krijgen over welke eindverbruiken door elkaar gehaald worden op het moment dat waterverbruik gelabeld wordt, zie Tabel 22. In deze tabellen is het percentage momenten te zien waarop een puls wordt gelabeld met hetzelfde of een ander eindverbruik, vergeleken met de originele data. Bijvoorbeeld, voor A_1 worden 14.3% van de momenten met een puls in de buitenkraan aangezien voor een badkamerkraan. Over het algemeen blijkt uit deze analyse:

- Vooral de kranen worden door elkaar gehaald.
- De meeste methodes wijzen veel labels foutief toe aan de keukenkraan.
- Veel water van de buitenkraan wordt aangezien als douche. Waarschijnlijk wanneer de buitenkraan voor een langere periode aanstaat.
- Voor C_1 en C_2 kunnen er geen conclusies getrokken worden over het bad en de buitenkraan omdat deze niet meegenomen zijn als mogelijke eindverbruiken/labels.

Tabel 22 Percentage momenten met een puls dat toegewezen is per eindverbruik vergeleken met origineel

ber 22 Percentage momenten met een paal dat toegevoegd is per eindverbruik vergeleken met origineel

A_1

		Identified as								
		BathroomTap	Dishwasher	KitchenTap	Outside Shower	WashingMachine	Wc	Bathtub	not identified	
Appliance	BathroomTap	77.1	0.1	10.3	3.0	0.2	0.6	0.6	0.2	8.0
	Dishwasher	0.2	92.7	4.5	0.0	0.0	0.1	0.4	0.0	2.1
	KitchenTap	19.3	0.5	63.0	5.0	2.1	2.7	4.7	0.6	2.1
	OutsideTap	14.3	0.0	19.1	40.0	15.0	6.0	0.6	2.2	2.7
	Shower	0.0	0.0	3.4	1.0	94.1	0.2	0.3	0.4	0.6
	WashingMachi	0.2	0.0	2.9	0.3	0.5	94.9	0.1	0.0	1.1
	Wc	0.4	0.0	5.9	1.3	0.7	0.1	89.3	0.1	2.3
	Bathtub	0.0	0.0	13.9	9.8	0.0	0.0	0.0	76.1	0.3

A_2

		Identified as								
		BathroomTap	Dishwasher	KitchenTap	Outside Shower	WashingMachine	Wc	Bathtub	not identified	
Appliance	BathroomTap	65.3	0.0	23.6	5.0	0.0	0.0	0.0	0.0	6.2
	Dishwasher	0.0	17.7	79.1	3.0	0.0	0.0	0.0	0.0	0.2
	KitchenTap	11.8	0.0	81.9	3.1	1.1	0.0	0.3	0.4	1.3
	OutsideTap	7.8	0.0	26.9	56.5	6.2	0.9	0.0	0.0	1.5
	Shower	0.7	0.0	3.0	6.0	90.0	0.0	0.0	0.0	0.4
	WashingMachi	0.3	0.1	21.8	1.7	0.1	75.3	0.0	0.0	0.7
	Wc	1.0	0.0	6.5	0.7	1.0	0.0	90.2	0.0	0.7
	Bathtub	0.0	0.0	0.0	0.3	0.0	0.0	0.0	99.6	0.0

B

		Identified as								
		BathroomTap	Dishwasher	KitchenTap	Outside Shower	WashingMachine	Wc	Bathtub	not identified	
Appliance	BathroomTap	54.7	0.4	9.4	13.7	3.7	0.6	2.6	0.0	14.8
	Dishwasher	0.0	92.6	2.0	0.0	0.7	0.4	1.8	0.0	2.5
	KitchenTap	13.3	3.8	45.6	10.9	6.6	5.2	8.3	0.6	5.7
	OutsideTap	2.2	0.3	6.8	38.9	36.2	9.2	1.6	0.0	4.8
	Shower	0.0	0.0	0.3	0.0	99.5	0.0	0.1	0.0	0.0
	WashingMachi	0.0	0.0	1.4	0.4	0.4	95.0	0.2	0.1	2.4
	Wc	0.0	0.1	1.1	0.4	2.9	0.3	90.7	0.0	4.5
	Bathtub	0.0	0.0	11.1	0.0	1.6	0.0	0.0	85.2	2.1

C_1

		Identified as								
		BathroomTap	Dishwasher	KitchenTap	Outside Shower	WashingMachine	Wc	Bathtub	not identified	
Appliance	BathroomTap	49.0	0.0	30.3	0.0	6.4	0.1	0.0	0.0	14.2
	Dishwasher	0.7	87.4	6.5	0.0	0.4	0.2	0.0	0.0	4.9
	KitchenTap	14.5	0.9	58.1	0.0	19.1	1.1	0.2	0.0	6.1
	OutsideTap	8.3	0.0	48.8	0.0	41.0	0.2	0.0	0.0	1.6
	Shower	0.0	0.0	1.9	0.0	97.7	0.0	0.0	0.0	0.3
	WashingMachi	0.1	0.0	2.9	0.0	0.5	95.0	0.0	0.0	1.6
	Wc	0.3	0.0	18.4	0.0	6.0	0.0	72.5	0.0	2.7
	Bathtub	0.0	0.0	99.5	0.0	0.0	0.0	0.0	0.0	0.5

C_2

		Identified as								
		BathroomTap	Dishwasher	KitchenTap	Outside Shower	WashingMachine	Wc	Bathtub	not identified	
Appliance	BathroomTap	55.6	0.0	23.6	0.0	6.4	0.1	0.0	0.0	14.2
	Dishwasher	0.9	87.5	5.9	0.0	0.4	0.2	0.2	0.0	4.9
	KitchenTap	16.7	0.9	55.7	0.0	19.2	1.1	0.4	0.0	6.1
	OutsideTap	10.0	0.0	47.2	0.0	41.0	0.2	0.0	0.0	1.5
	Shower	0.0	0.0	1.9	0.0	97.7	0.0	0.0	0.0	0.3
	WashingMachi	0.1	0.0	2.9	0.0	0.5	95.0	0.0	0.0	1.6
	Wc	0.3	0.0	1.4	0.0	6.0	0.0	89.6	0.0	2.7
	Bathtub	0.0	0.0	99.5	0.0	0.0	0.0	0.0	0.0	0.5

7.4 Gebeurtenissen onderscheiden

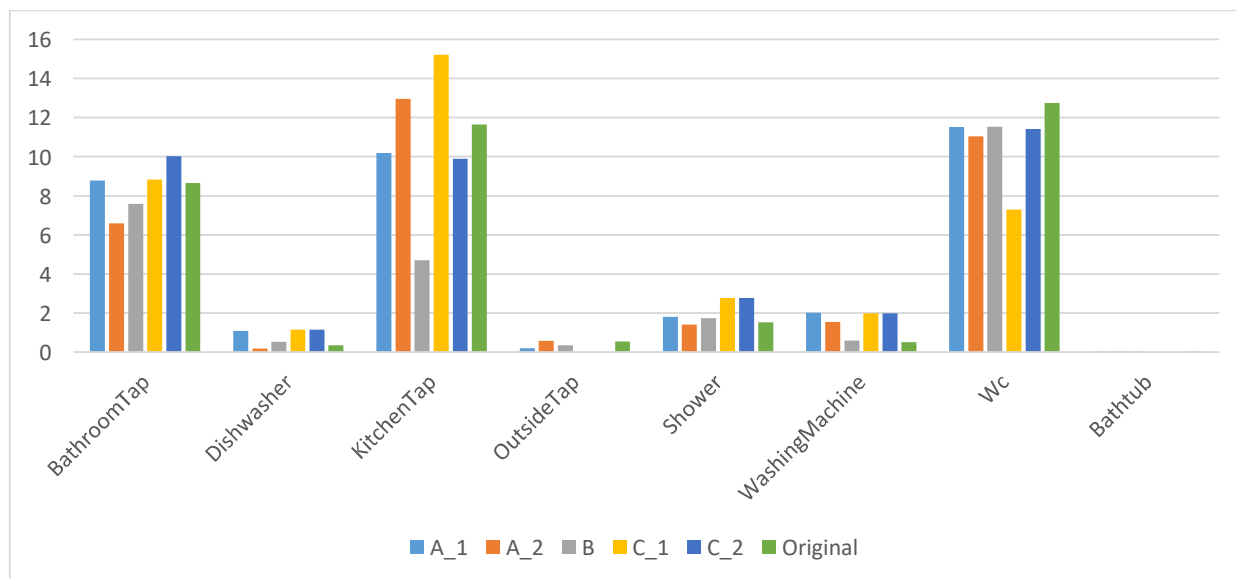
Voor het WT-onderzoek is het van belang om inzicht te krijgen in het aantal keer dat elk van de eindverbruiken ingeschakeld is. De totale hoeveelheid gebeurtenissen per disaggregatie en in de originele dataset is te zien in Tabel 23, deze zijn direct verkregen uit de pySIMDEUM output zonder de puls transformatie. A en B identificeren over het algemeen minder individuele gebeurtenissen dan in de originele dataset, waarbij B aanzienlijk minder (25%) gebeurtenissen identificeert.

Tabel 23 Geïdentificeerde gebeurtenissen per disaggregatiemethode en in de originele dataset

Disaggregation method	Events
A_1	24936
A_2	24045
B	18956
C_1*	26067
C_2	26067
Original	25183

*C_1 does not use flush interruption for the toilet and illustrates the possible effect of this mechanism on the disaggregation results.

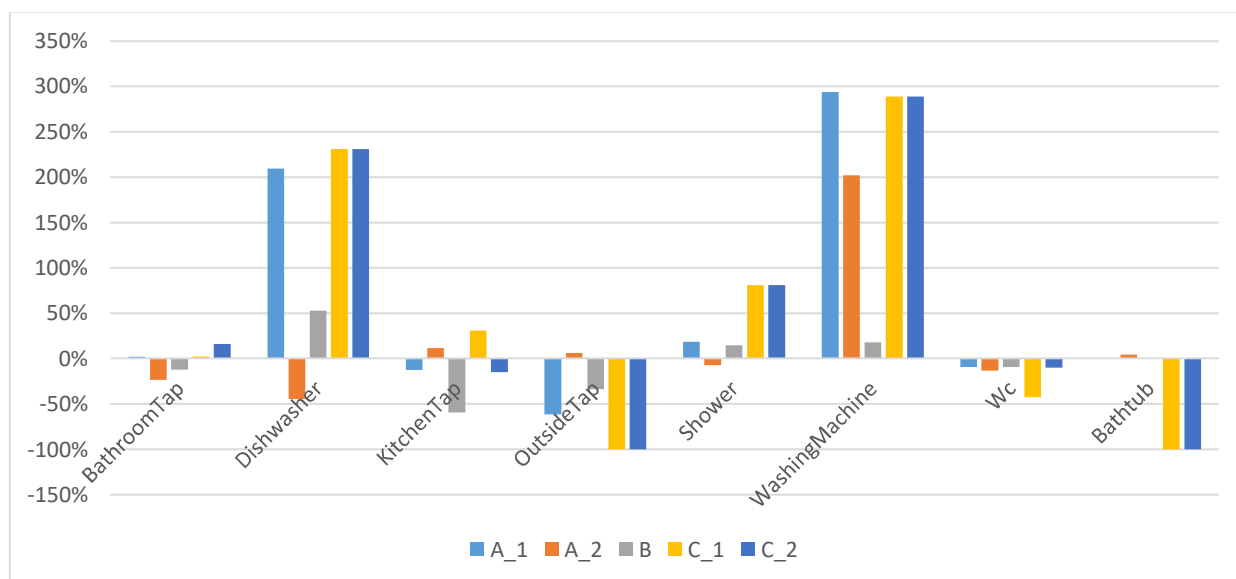
De gemiddelde hoeveelheid gebeurtenissen per huishouden per dag per eindverbruik is te zien in Figuur 22. Hier wordt geen onderscheid gemaakt tussen huishoudens met meer of minder bewoners of de penetratiegraad van de eindverbruiken. Tussen C_1 en C_2 is de totale hoeveelheid gebeurtenissen hetzelfde maar is er een verschuiving tussen de keukenkraan en de WC.



Figuur 22 Gemiddelde hoeveelheid gebeurtenissen per dag per eindverbruik per woning

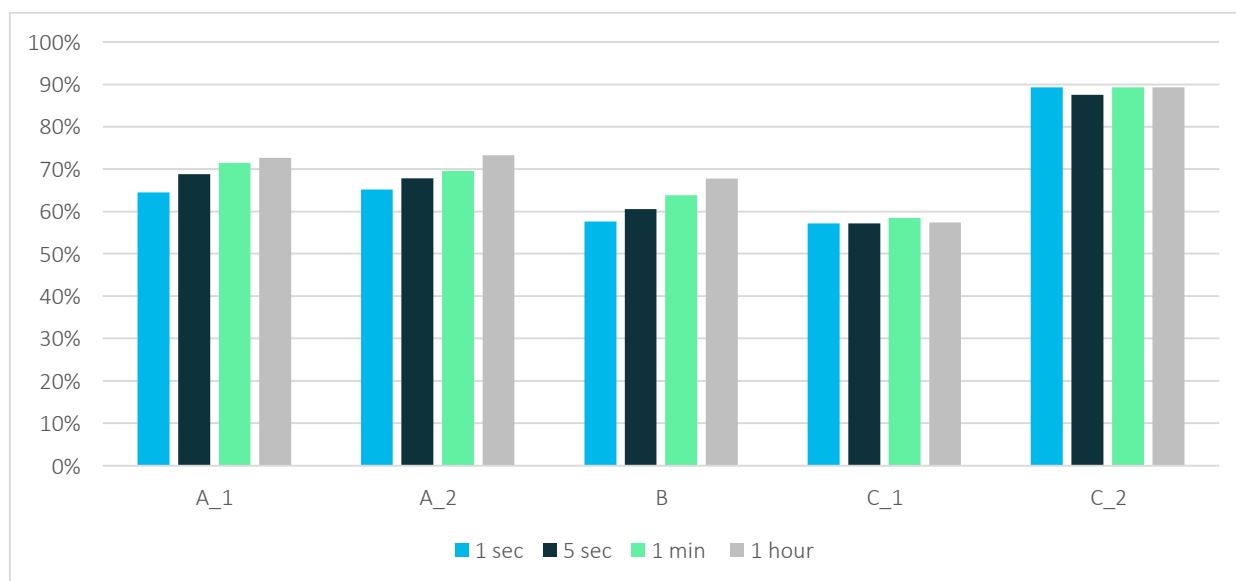
De hoeveelheid gebeurtenissen per eindverbruik in relatie tot het origineel is te zien in Figuur 23. In de meeste disaggregatiemethodes is de hoeveelheid gebeurtenissen van de vaatwasser en de wasmachine significant hoger dan in werkelijkheid. Reden hiervoor is dat elk van de inname-momenten gerapporteerd is als een individuele gebeurtenis. Gezien beide eindverbruiken vier inname-momenten hebben zou te verwachten zijn dat er volgens deze methode ongeveer 300% meer gebeurtenissen gerapporteerd worden. Voor A_1, C_1 en C_2 is dit het geval voor de wasmachine. Voor de vaatwasser vallen de waardes over het algemeen lager uit, maar dit wordt

veroorzaakt door een fout tijdens het genereren van de dataset, algemene verbeteringen aan pySIMDEUM op dit onderdeel worden besproken in de aanbevelingen voor vervolgstappen.



Figuur 23 Relatieve prestatie van de disaggregatie methode in de het identificeren van gebeurtenissen in relatie tot de werkelijke hoeveelheid gebeurtenissen

Naast de totale hoeveelheid gebeurtenissen is het van belang wanneer deze gebeurtenissen plaatsvinden. Door simultaan waterverbruik kan het gebeuren dat een gebeurtenis niet op het exacte moment herkend wordt als in de testdata. Disaggregatiemethodes A en B presteren beter naarmate het tijdsinterval tussen de voorspelde gebeurtenis en de werkelijke gebeurtenis groter wordt, maar methode C lijkt hier ongevoelig voor te zijn (Figuur 24). Over het algemeen presteert methode C_2 het beste wanneer een gebeurtenis uit de disaggregatie teruggezocht wordt binnen de originele dataset, voor elke onderzochte tijdsrange. Vooralsnog is onduidelijk waarom dit verschil tussen de methodes optreedt.



Figuur 24 Percentage correct geïdentificeerde gebeurtenissen naar tijdsrange waarin gezocht wordt

8 Conclusies en aanbevelingen

In dit project is een disaggregatie-opdracht opgezet om te evalueren in hoeverre algoritmes de individuele waterverbruikstoepassingen kunnen onderscheiden uit de pulldata van digitale watermeters. De evaluatiecriteria voor de disaggregatie zijn opgesteld in samenspraak met de Nederlandse drinkwaterbedrijven en de synthetische testdata is genereerd met pySIMDEUM. Naast het hoofddoel van dit onderzoek, uitspraak te doen over de potentiële waarde van disaggregatiemethodes binnen het Watergebruik Thuis onderzoek, is gebleken dat pySIMDEUM mogelijkheden biedt om synthetische datasets te generen waarmee een inschatting gemaakt kan worden over de prestaties van disaggregatiemethodes.

Het belang van een mechanisme als de onderbreking van de spoelknop van de WC wordt duidelijk geïllustreerd door de verschillen tussen de resultaten van C_1 en C_2. De uitkomsten van C_1 zijn daarom weergegeven ter illustratie maar worden niet meegenomen in verdere conclusies.

De eindverbruiken in de huishoudens worden met een nauwkeurigheid van 84% of hoger geïdentificeerd door de disaggregatiemethodes. Bij het identificeren van eindverbruiken kan er onderscheid gemaakt worden tussen methodes met en zonder voorkennis. Met voorkennis kan het alsnog voorkomen dat er waterverbruik wordt toegewezen aan eindverbruiken welke in de simulatieperiode niet gebruikt zijn, een wasmachine kan bij toeval een week ongebruikt blijven. Met of zonder voorkennis worden vooral de buitenkraan, het bad, en de vaatwasser over- of ondergerapporteerd in de resultaten. Deze eindverbruiken waren dus: (1) niet aanwezig in het huishouden of werden niet gebruikt tijdens de simulatie periode maar werden wel als waterverbruikers bestempeld door de disaggregatie (overraportage) of (2) waren wel aanwezig en werden ook gebruikt maar werden niet gedetecteerd door de disaggregatiemethode (onderraportage).

De disaggregatiemethodes karakteriseren tussen de 82% en 91% van het verbruikte water naar het correcte eindverbruik. De methodes presteren over het algemeen het best op de douche, wasmachine en WC wanneer het van belang is om het waterverbruik naar volume van de individuele eindverbruiken te bepalen. Gezien het totale waterverbruik van de douche en de WC betekent de foutmarge binnen deze eindverbruiken alsnog relatief de hoogste foutmarge in relatie tot het totale waterverbruik. De meeste afwijking tussen de disaggregatiemethodes en de originele data gebeurt tussen 7:00 en 8:00. Dit is ook het moment waarop het vaakst door meerdere eindverbruiken simultaan water wordt verbruikt.

Bij het toekennen van binaire labels ligt het percentage correcte labels tussen de 72% en 77%, en presteren de disaggregatiemethodes het beste op de douche, vaatwasser en wasmachine. De kranen en het bad zijn de eindverbruiken waar de meeste vals positieven en vals negatieven aan worden toegekend. De disaggregatiemethodes presteren verschillend op precisie en recall per eindverbruik. Over het algemeen presteert methode B het beste op beide vlakken terwijl C_1 en C_2 een zeer hoge precisie voor de wasmachine en WC vertonen. Dit betekent dat het van belang is om een specifieke disaggregatiemethode te kiezen wanneer men geïnteresseerd is in een specifiek eindverbruik. Tijdens simultaan waterverbruik is de douche het best te onderscheiden van de andere eindverbruiken. Over het algemeen worden labels onderling door elkaar gehaald bij de kranen, en is de keukenkraan het eindverbruik waar labels aan toegewezen worden wanneer onduidelijk is welk ander eindverbruik het zou kunnen zijn.

Bij het identificeren en onderscheiden van individuele gebeurtenissen van waterverbruik valt op dat de disaggregatiemethodes de totale hoeveelheid kunnen overschatten maar ook onderschatten. Over het algemeen wordt de hoeveelheid keer dat de WC verbruikt wordt onderschat; voor de andere eindverbruiken dit afhankelijk is

van de disaggregatiemethode. Op dit moment kunnen er nog geen conclusies getrokken worden over de individuele gebeurtenissen van de vaatwasser; tijdens het genereren van de dataset werd deze voortijdig opnieuw geactiveerd en afgekapt aan het eind van de dag. Methode C_2 presteert het beste wanneer gebeurtenissen geïdentificeerd moeten worden binnen een uur van het moment waarop ze daadwerkelijk voorkomen.

Op basis van bovenstaande resultaten kan geconcludeerd worden dat de huidige staat van disaggregatiemethodes voor (synthetische) digitale watermeterdata redelijk ver gevorderd is. Gezien de verschillen tussen de methodes kan er geen uitsluitend gegeven worden over de meest 'geschikte' methode, en zal dit bepaald worden afhankelijk van het gewenste doel. In het geval van het Watergebruik Thuis onderzoek is het vooral van belang dat het totale waterverbruik, het verbruik per keer, en de hoeveelheid gebeurtenissen correct geïdentificeerd wordt.

Verdere ontwikkeling aan pySIMDEUM, gericht op een reëlere weerspiegeling van huishoudelijk waterverbruik, zouden een beter beeld geven van de daadwerkelijke mogelijkheden voor disaggregatie en wordt aanbevolen als vervolgstap. De belangrijkste aspecten welke aandacht verdienen zijn:

- Her-evalueren en dubbelchecken van de modelstructuur van pySIMDEUM om ervoor te zorgen dat veranderingen aan eindverbruiken (vooral de vaatwasser en wasmachine) makkelijk doorgevoerd worden in een simulatie met meerdere huishoudens.
- Verlagingen van volumestromen en aanpassingen aan blokfuncties in geval van simultaan waterverbruik. In een vervolgstap zou onderzocht moeten worden of dit via een mechanistisch sub-model, of op basis van correlaties op basis van meetgegevens, toegevoegd kan worden aan pySIMDEUM.

Gezien de prestaties van de huidige disaggregatiemethodes is een vervolgstap op dit onderzoek aan te bevelen. Vooralsnog is niet bekend hoe de methodes presteren wanneer de nauwkeurigheid van de digitale puls data afneemt. Kennis hierover kan waardevol zijn om de gewenste specificaties van digitale watermeters te bepalen, in combinatie met de mogelijkheden om data effectief te verzenden/op te slaan. Vanuit de huidige opzet kunnen deze nieuwe datasets makkelijk gegenereerd worden en zouden de disaggregatiemethodes van de huidige deelnemers met weinig aanpassingen de evaluatie opnieuw kunnen uitvoeren. Op basis van de resultaten van dit onderzoek is de betrokken drinkwaterbedrijven/Vewin gevraagd welke vervolgstappen hun voorkeur heeft (zie Bijlage VII). Verder onderzoek zal (onder andere) moeten toewerken naar: (1) het achterhalen of disaggregatie dezelfde betrouwbaarheid als het Watergebruik Thuis onderzoek kan behalen, (2) bepalen wat het effect van datanauwkeurigheid is op de disaggregatieresultaten, (3) bepalen van de vereisten voor digitale watermeters voor een uiteindelijke pilot studie.

9 Referenties

- Blokker, E. J. Maria. (2010). *Stochastic water demand modelling for a better understanding of hydraulics in water distribution networks*. Water Management Academic Press.
- Steffelbauer, D. B., Hillebrand, B., & Blokker, M. (2024, March 1). *pySIMDEUM - An open-source stochastic water demand end-use model*. <https://doi.org/10.4995/wdsa-ccwi2022.2022.14774>

I Begeleidende brief en vragenlijst voor commerciële partijen

Dear Potential Disaggregation Participant,

My name is Mirjam Blokker, and I am a principal scientist at KWR Water Research Institute. We are doing a project for eight Dutch drinking water utilities, who are orientating themselves on the use of smart water meters in households. I would like to invite your company to participate in our project, as I consider you one of the front runners on disaggregation of (water) use patterns.

The project “Evaluating disaggregation of water consumption data to individual components” is meant to quantify how well (commercially) available algorithms perform for disaggregating and labelling water consumption data from high-resolution smart water meters. The goal is to advise the drinking water utilities on the use cases of high-resolution smart water meter data in combination with smart algorithms: are these accurate enough to replace the four yearly diary-based inquiries into “water use at home”[1]? Are they accurate enough to quantify the effect of water reduction measures? Are they accurate enough to be helpful in reducing water demand? And other uses. In order to do the quantification, KWR will provide a dataset (a timeseries of flow data of up to 100 households), of which we know the labels (such as toilet flushing, shower, washing machine, dishwasher, kitchen tap, etc.). We expect the participants to use their algorithms and return the dataset with labels. KWR will then evaluate the performance. As the accuracy of the algorithms depends partly on the temporal resolution (1 s, 10 s, 1 min, 5 min, 15 min) and the accuracy of the flow measurements (0.1 l/pulse, 1 l/pulse) we would like to test several variations of these resolutions.

KWR will create these datasets with SIMDEUM [2]. SIMDEUM simulates (residential) water demand at the end-use (and end-user) level, and with a temporal resolution of 1 s. It was shown that SIMDEUM provides realistic water demand patterns at aggregated levels, e.g. at the level of a few households and 10 min [3], or per household per 10 s [4]. For the purposes of this specific project, we will enhance SIMDEUM with some features to make the demand patterns at the level of the end-use and per 1 s more realistic (non-perfect blocks of on/off flow) and with a little more variability within a home (patterns of washing machines e.g.), and make sure that consecutive days (instead of random daily patterns) are being simulated. There will be no explicit seasonal influence (holidays, temperature and precipitation, school vacations) on the generated water demand pattern. Leakage is not being introduced at this point. A post processing step will mimic the logging of a smart water meter.

Creating a synthetic dataset has the advantage of being relatively cheap to achieve a large dataset. Another advantage is that all labels are exactly known, whereas creating a fully labelled dataset from measurements is very challenging [5]. The downside is that computer algorithms will work better on computer generated (virtual) data than on actual human behavior data, and thus a too optimistic result may be found. Given earlier attempts of disaggregation and labelling [5, 6] and the difficulty to get a good labelled dataset, we take this disadvantage for granted at this stage. In a later stage we may try to make the dataset even more realistic, e.g. based on high-resolution measured data and expert judgement, by adding even more variability within a household. The dataset and evaluation criteria that we create will serve as a standardize protocol for the evaluation of available disaggregation and labelling tools. They will be available to our project partners for future reference, e.g. when they are tendering for smart water meters.

We would be glad if you join our project. If you confirm your participation before 10 July 2024, we can include you in the project. We will provide the dataset to the participants by the end of September. If you are interested in joining our project, please answer the questions on the following pages and sign the document. If you require any further clarifications don't hesitate to contact us.

Best regards,

Mirjam Blokker

Requirements for participation

If you are interested in joining our project, please answer the following questions. Sometimes you can select multiple options. There is also room to indicate “other” requirements. We may not always be able to grant your requests.

1. Which temporal resolution will you require for the dataset? You can choose one or more of the following. If you choose multiple, be aware that those datasets will be different, i.e. not an aggregated version of each other.

- ☐ 1 s
- ☐ 10 s
- ☐ 1 min
- ☐ 5 min
- ☐ Other:

2. Which resolution in flow will you require for the dataset? You can choose one or more of the following. If you choose multiple, be aware that those datasets will be different, i.e. not an aggregated version of each other.

- ☐ 0.1 L/pulse
- ☐ 1 L/pulse
- ☐ Other:

3. KWR will provide a dataset of 100 individual households. Which extra information do you require for each household? The performance score of your disaggregation method may be influenced by the amount of extra information you require.

- ☐ Number of residents
- ☐ Water using appliances that are present
- ☐ The potential labels to choose from
- ☐ Other:

4. KWR will provide a pattern of 14 consecutive days for each household. Which extra information do you require for the pattern?

- ☐ Indication of day of the week (week day/weekend)
- ☐ Other:

5. Do you require a training dataset with labels (supervised learning) or only a test dataset without labels?

- ☐ Test dataset (no label)
- ☐ Training dataset (with labels, see also the next question)
- ☐ For each home a start-up timeseries (few hours), with labels. We can introduce a protocol where each tap is used, so that a fingerprint of all taps (w.r.t. volume or flow rate) can easily be discerned (no overlapping measurements).

6. At which aggregation level will your output be? KWR will prepare for those labels.

- ☐ The end-use level (e.g. showering, clothes washing, washing hands, etc.)
- ☐ The tap level (shower, bath, washing machine, dishwasher, toilet nr 1, toilet nr 2, kitchen tap, bathroom tap, outside tap)
- ☐ The aggregated tap level (shower and bath, washing machine, dishwasher, toilet, tap)
- ☐ Other:

7. We expect that your results will be a timeseries of water using events with a label indicating the end-use, the flow rate and the duration of the event. KWR will evaluate the results per label, and according to the division

that is of value for the “water use at home” study (see appendix 1). The exact assessment criteria are still to be determined. We will inform you about the criteria before you receive the dataset. Do you have any preference on the assessment criteria?

- ☐ Overall score (for all labels together)
- ☐ Score per label
- ☐ Score per water use at home division
- ☐ Percentage of volume labelled correctly
- ☐ Percentage of events labelled correctly
- ☐ Insight into false negatives, false positives

8. We would like to get some insight into the algorithm that you use. Are you willing to share some information on that with us? This question is optional.

- ☐ The method that is used. Name: Reference:
- ☐ The method works real-time
- ☐ The method works as a post-processing and requires some time. Indicate number of minutes / hours.
- ☐ Please indicate how reliable your results are expected to be:

9. KWR will provide the dataset at a specified date. This date will be confirmed on time with the participants. How much time do you require before you can return the results?

- ☐ 2 weeks
- ☐ 4 weeks
- ☐ 6 weeks
- ☐ Other:

10. We will report the results. Do you have any restrictions with respect to the report?

- ☐ Our company name can only be mentioned if the report is restricted to the project clients (not publicly available)
- ☐ If the report is publicly available, our company's results can be reported with the company's name. This will allow the company to refer to the report to inform (potential) customers
- ☐ If the report is publicly available, our company's results can only be reported anonymously
- ☐ Our company wants to participate, but we need a non-disclosure agreement before proceeding (we will work with the KWR NDA format for consistency)
- ☐ Other:

11. Is there anything else you would like us to know?

Please sign below if you wish to participate in this project:

Name

Company

Signature

Date

References

- [1] Bakker, J., F. van der Mooren, and H.J. Boonstra, Watergebruik thuis 2021. 2022: CBS.
- [2] Blokker, E.J.M., Stochastic water demand modelling for a better understanding of hydraulics in water distribution networks, 2010, Delft University of Technology. p. 212.
- [3] Blokker, E.J.M., C. Agudelo-Vera, A. Moerman, P. van Thienen, and E.J. Pieterse-Quirijns, Review of applications for SIMDEUM, a stochastic drinking water demand model with a small temporal and spatial scale. Drinking Water Engineering and Science, 2017. 10(1): p. 1-12.
- [4] Zeidan, M. and B. Hillebrand, A golden standard of representative flow rate distributions. Inventory for the new inspection system for water meters. 2023, Nieuwegein: KWR. KWR 2023.044
- [5] Levasseur, G., Human activity recognition in water meter data. 2023.
- [6] Mazzoni, F., S. Alvisi, M. Franchini, and M. Blokker, Exploiting high-resolution data to investigate the characteristics of water consumption at the end-use level: A Dutch case study. Water Resources and Industry, 2022: p. 100198.

Appendix 1: Water use at home table

	P (%)	F (# per day)	D (min)	Q (l/min)	V (l per event)	Volume (l/day)
Bath total						5.3
Bath full (V > 50 L)	42	0.09	10.3	13.9	142.4	5.2
Bath partial (V < 50 L)	5	0.12			20.2	0.1
Shower total	99	0.82	7.4	7.7	56.9	46.2
Flow rate A: 6-7.5 l/min	39	0.75	7.1	6.3	44.8	13.1
Flow rate B: 7.5-9.0 l/min	61	0.79	5.7	8.0	45.9	22.1
Flow rate C: 9-10.5 l/min	24	1.04	4.1	10.3	42.1	10.5
Flow rate D: 10.5-15 l/min	2	0.3	6.1	13.6	83.3	0.5
Toilet	100					30.2
Full	100	2.93			6.7	19.7
Partial	78	3.65			3.7	10.5
Dishwasher	76	0.33			11.7	2.9
Washing machine	98	0.35			48.9	16.9
Outdoor tap	75	0.00	49.5	5.9	289.3	0.4
Washbasin		1.89			4.6	8.7
Kitchen tap						17.5
Very large volume (V > 10 L)		4.6			0.2	0.9
Large volume (4 < V < 10 L)		1.0			1.2	1.2
Medium volume (2 < V < 4 L)		4.8			2.2	10.4
Small volume (0.5 < V < 2 L)		0.9			4.5	3.9
Very small volume (V < 0.5 L)		0.1			19.2	1.1

N.B. At this time, leaking taps and toilets will not be included in the dataset. Also, the residents in a home are the only users; visitors and long-term absence of residents is not being considered.

II Detail huishoudsamenstelling

End use composition	Residents (#)					Total
	1	2	3	4	5	
['Wc', 'BathroomTap', 'Dishwasher', 'KitchenTap', 'OutsideTap', 'Shower', 'WashingMachine']	5	7	3	4		19
['Wc', 'BathroomTap', 'Dishwasher', 'KitchenTap', 'OutsideTap', 'Shower']	1	1				2
['Wc', 'BathroomTap', 'Dishwasher', 'KitchenTap', 'Shower', 'WashingMachine']	7	4	2	2		15
['Wc', 'BathroomTap', 'KitchenTap', 'OutsideTap', 'Shower', 'WashingMachine']	8	1			1	10
['Wc', 'BathroomTap', 'KitchenTap', 'OutsideTap', 'Shower']	1					1
['Wc', 'BathroomTap', 'KitchenTap', 'Shower', 'WashingMachine']	7	5				12
['Wc', 'Bathtub', 'BathroomTap', 'Dishwasher', 'KitchenTap', 'OutsideTap', 'Shower', 'WashingMachine']		9	6	4	2	21
['Wc', 'Bathtub', 'BathroomTap', 'Dishwasher', 'KitchenTap', 'OutsideTap', 'Shower']			1			1
['Wc', 'Bathtub', 'BathroomTap', 'Dishwasher', 'KitchenTap', 'Shower', 'WashingMachine']		3		4		7
['Wc', 'Bathtub', 'BathroomTap', 'Dishwasher', 'KitchenTap', 'Shower']	1					1
['Wc', 'Bathtub', 'BathroomTap', 'KitchenTap', 'OutsideTap', 'Shower', 'WashingMachine']	4	2	1			7
['Wc', 'Bathtub', 'BathroomTap', 'KitchenTap', 'OutsideTap', 'Shower']	1					1
['Wc', 'Bathtub', 'BathroomTap', 'KitchenTap', 'Shower', 'WashingMachine']	1	2				3
Total	36	34	13	14	3	100

III Instructies voor deelnemende partijen

Information about disaggregation dataset

Hello Disaggregation participants,

With this document we inform you about the next steps of the disaggregation project. This document covers three points 1) the format of the dataset you will receive for the disaggregation of smart water meter data 2) the format in which we expect you to return the data to us 3) the evaluation method we will use on your disaggregation results 4) the timing for the rest of the project. We hope that this covers the complete procedure and enables you to perform the disaggregation effectively and in such a way that we can evaluate the results.

Format of Dataset

The dataset you receive will be a collection of comma-separated values (CSV) files and will have the overall structure shown in *Figure 1*. The separator used in the CSV files is ‘;’. The dataset contains a collection of folders named with a unique ‘id’. Each of these folders represents the simulated dataset for a single house. Depending on your response to question 3 of the questionnaire you will also receive the ‘appliance_data.csv’ and ‘composition_data.csv’ files. The ‘appliance_data.csv’ file contains an overview of the appliances present in each of the simulated houses. If you chose not to receive this file then generating it becomes part of the disaggregation and the evaluation (see Section 0). This list of appliances corresponds to the potential labels you can choose from when labeling the data. The ‘composition_data.csv’ file contains an overview of the number of residents/users in the house. In both these files you will find the ‘id’ which matches the folder names of the simulations.

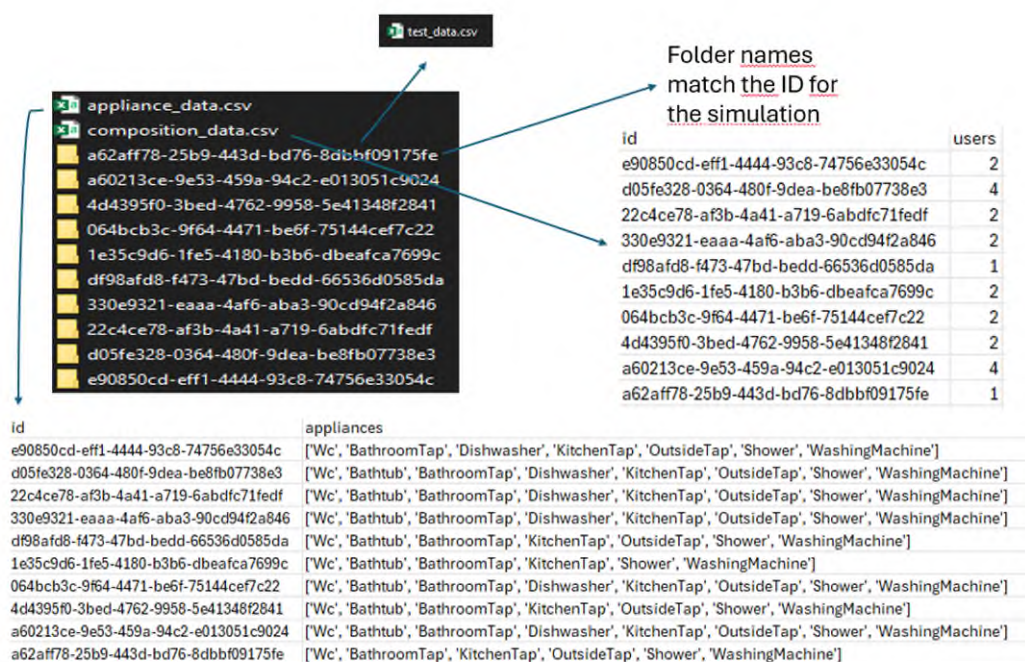


Figure 1 Overall dataset structure

Within each of the folders you will find a file named ‘test_data.csv’ (*Figure 1* **Error! Reference source not found.**). This is the household water consumption in the format that would be produced by a smart

water meter in the form of a timeseries. The generated consumption patterns were 14 days long, where the first 7 days serve as potential training data. For this reason the test_data.csv file on which you perform the disaggregation starts on day 8 (see **Error! Reference source not found.**). The temporal resolution and the precision of the pulses in this file corresponds to your answer for questions 1 and 2 of the questionnaire. The file has 5 columns:

- time: the time "hh:mm:ss" of the datapoint
- day: the day of the datapoint
- weekday: whether the datapoint is on a 'weekday' or in the 'weekend'
- Pulses: the number of pulses during the timestep, this can also be '0'
- Flow_L: the flow during the timestep in liter. This corresponds to the precision of the meter multiplied by the number of pulses. For example 4 pulses * 0.1 L/pulse = 0.4 L

The 'time', 'day', and 'weekday' columns combined serve as the index for all the datasets. Each file will have the same index. When you return the data to us, please use the same index format to make evaluation possible.

Table 1 Example of the test_data.csv file

time	day	weekday	Pulses	Flow_L
07:48:53		8 Weekday	0	0
07:48:54		8 Weekday	0	0
07:48:55		8 Weekday	0	0
07:48:56		8 Weekday	2	0.2
07:48:57		8 Weekday	1	0.1
07:48:58		8 Weekday	1	0.1
07:48:59		8 Weekday	1	0.1
07:49:00		8 Weekday	1	0.1
07:49:01		8 Weekday	2	0.2

Depending on your response to question 5 of the questionnaire you will receive the following additional training data:

A folder containing unlabeled test data. The folder contains a collection of 'unlabeled_pulse_data.csv' files – this is additional pulse data as generated by a smart water meter which can be used to train your model/algorithms. These are dataset you can use to train your algorithm, but which you do not need to return to us after processing

'training_tap_log.csv' – this is a labeled output of the water use events generated with SIMDEUM. You will find the starting time of the event, the end use, and the duration of the event (Table 2). This file serves as the training dataset for your algorithms. The training dataset has a length of 7 days, while the total dataset has a length of 14 days. As such you can use the first 7 days to train your model and disaggregate the next 7 days.

Table 2 Example of the training_tap_log.csv file

time	day	weekday	end use	duration(s)
00:04:35	1	Weekday	Shower	540
00:07:40	1	Weekday	Wc	48
00:18:58	1	Weekday	Wc	48
03:15:43	1	Weekday	Wc	48
05:03:26	1	Weekday	KitchenTap	24
08:01:52	1	Weekday	KitchenTap	15

‘startup_protocol.csv’ – this is a startup protocol for the house to create a fingerprint for the different end uses in the house. If the end use is present in the house the startup protocol shown in *Table 3* is followed to generate the fingerprint for each of the appliances. The order may be different, but each appliance fingerprint should be present. The startup protocol works with the same smart water meter pulse precision as the rest of the dataset and is represented as a matrix with a specific number of pulses per timestep (*Table 4*).

Table 3 Startup protocol for end uses

Kitchen tap max	Open the kitchen tap for 30 seconds at maximum intensity
Kitchen tap avg	Open the kitchen tap for 30 seconds at average intensity (maximum intensity/2)
WC	Full flush (i.e. no flush interruption / half flush) and wait until the reservoir is full again
Bathroom tap high	Open the bathroom tap for 30 seconds at high intensity. In SIMDEUM the intensity for the bathroom tap has a normal distribution so this is not necessarily the maximum intensity
Bathroom tap avg	Open the bathroom tap for 30 seconds at average intensity (high intensity/2)
Shower	Open the shower for 30 seconds at maximum intensity
Outside tap max	Open the outside tap for 30 seconds at maximum intensity
Dishwasher	Turn on the normal program of the dishwasher and do not use any other taps
Washing machine	Turn on the short program of the washing machine and do not use any other taps
Bathtub	Open the bath tap (the bathtub) for 30 seconds (the bath is not filled fully)

Table 4 Example of the startup_protocol.csv which serves as a fingerprint for each of the appliances. The values are the number of pulses for that timestep.

time	day	weekday	Wc	Bathtub	BathroomTap	Dishwasher	KitchenTap	Shower	WashingMachine
00:00:00	startup_protocol	N.A.	1	0	0	0	0	0	0
00:00:01	startup_protocol	N.A.	1	0	0	0	0	0	0
00:00:02	startup_protocol	N.A.	1	0	0	0	0	0	0
00:00:03	startup_protocol	N.A.	1	0	0	0	0	0	0
00:00:04	startup_protocol	N.A.	2	0	0	0	0	0	0
00:00:05	startup_protocol	N.A.	1	0	0	0	0	0	0
00:00:06	startup_protocol	N.A.	1	0	0	0	0	0	0
00:00:07	startup_protocol	N.A.	1	0	0	0	0	0	0

Format for the returning disaggregated data

We request the returned data to be in the same format as it was supplied, in CSV. Please return a folder with the same structure as the folder you received, where the subfolder names correspond to the id's of the houses we sent you.

Within the main folder please return ‘appliance_data_processed.csv’ (see *Figure 1*). This is a list of the water using appliances you identified with your disaggregation method. Make sure that you link them to the correct simulation with the ‘id’. You only need to include this data for the test data where this information was not provided through the original ‘appliance_data.csv’, ‘training_tap_log.csv’, or ‘startup_protocol.csv’. The possible labels you can assign are: Wc, Bathtub, BathroomTap, Dishwasher, KitchenTap, Shower, WashingMachine, OutsideTap. Please use these exact strings to make evaluation possible.

Within each sub-folder please return the following CSV files:

'tap_log_processed.csv' – a list of water using events in the same format as training_tap_log.csv described above (Table 5). This file contains a list of water using events with the starting time (please keep the time, day, weekday index), the end use your disaggregation assigned to the event, the duration of the event in seconds, and the amount of water used in the event.

Table 5 Example of tap_log_processed.csv

time	day	weekday	end use	duration(s)	water use (L)
06:53:53		2 Weekday	KitchenTap	6	...
08:06:05		2 Weekday	KitchenTap	10	...
08:07:50		2 Weekday	...	...	...
08:09:25		2 Weekday	...	...	...
08:17:18		2 Weekday	...	...	...
08:20:39		2 Weekday	...	...	...
08:28:47		2 Weekday	...	...	...

'labeled_pulses.csv' – a matrix with the end uses as columns and a binary indication when a specific end use is responsible for a pulse trigger of the smart water meter (1 = pulse(s) present, 0/No Data = no pulse). See **Error! Reference source not found.** for an example. In this matrix we do not expect you to assign a specific number of pulses to a specific end use, only whether the end use consumed water during any of the recorded pulses.

Figure 2 Example of 'labeled_pulses.csv'. On the left, the data of 'test_data.csv' is shown (you will receive this data). On the right, the expected binary representation of the end uses responsible for the pulses based on your disaggregation method is shown.

time	day	weekday	Pulses		time	day	weekday	Wc	Bathtub	BathroomTap	Dishwasher	KitchenTap	Shower	WashingMachine
00:04:30	1	Weekday	0		00:04:30		1 Weekday							
00:04:31	1	Weekday	0		00:04:31		1 Weekday							
00:04:32	1	Weekday	0		00:04:32		1 Weekday							
00:04:33	1	Weekday	0		00:04:33		1 Weekday							
00:04:34	1	Weekday	0		00:04:34		1 Weekday							
00:04:35	1	Weekday	1		00:04:35		1 Weekday						1	
00:04:36	1	Weekday	1		00:04:36		1 Weekday						1	
00:04:37	1	Weekday	1		00:04:37		1 Weekday						1	

Evaluation method

In this section the evaluation method for the disaggregated data is explained. For each step of the evaluation method the level of detail increases. In the final report the results for each of the disaggregation procedures is reported according to the steps described below.

NOTE: The generated datasets were 14 days long, the first 7 days were made available for training (based on your response to question 5 of the questionnaire), the last 7 days are for disaggregation and evaluation. This means that the disaggregated data you return to us should always start at 00:00 of day 8. Make sure that the original time, day, weekday index is used when returning the data.

1. Identified end-uses: the correctly identified end-uses is the first evaluation metric. This metric will only be evaluated for the test data where 'training_tap_log.csv', 'startup_protocol.csv' or 'appliance_data.csv' were not provided. This information should be placed within 'appliance_data_processed.csv'. We will report on the overall accuracy across all appliances, and the accuracy within each appliance category.
2. Total water user per end-use: based on the returned 'tap_log_processed.csv' files the correctly identified water use over different time periods (in terms of water volume) will be determined. We will evaluate the performance over the complete simulation period of 7 days, per day, and

per hour per day. A margin of error will be used around the day (around 5 min) and hour (around 1 min) boundary. E.g. if a disaggregation procedure returns a water consumption of 5 L for a WC at 00:03 of day 2, while this water use actually occurred at the end of day 1 this would still be considered correct. A distinction will be made between the percentage difference (above or below) and the absolute percentage difference of the water volume (Table 6). Also the difference of the disaggregation and the test data in relation to the total water use is evaluated.

Table 6 Example of the evaluation table based on total water use over 7 days

End-Use	SIMDEUM (L/7days)	Disaggregation (L/7days)	Difference (%)	Absolute difference (%)	Difference compared with total water use (%)
BathroomTap	87.3	90.8	4%	4%	0.1%
Bathtub	0.0	0.0	-	-	-
Dishwasher	14.0	15.8	13%	13%	0.1%
Shower	1439.1	1358.2	-6%	13%	-2.7%
KitchenTap	132.2	115.2	-13%	6%	-0.6%
OutsideTap	118.2	102.8	-13%	13%	-0.5%
WashingMachine	297.4	303.3	2%	2%	0.2%
Wc	916.7	1018.6	11%	14%	3.4%
Total	3004.7	3004.7	-	-	-
Average	375.6	375.6	0%	9%	0.0%

Labeling of pulses: the returned 'labeled_pulses.csv' will be compared with the simulated data (Table 7) to identify the percentage of correctly labeled pulses, the percentage of false positives, and the percentage of false negatives (Table 8 **Error! Reference source not found.**). In this evaluation no temporal margin is allowed, since the pulse data in 'test_data.csv' has an exact timestamp should not be altered but only assigned to an end-use. An additional distinction will be made for events occurring during simultaneous use of different end-uses (Table 9).

Table 7 Example of the binary pulse labels from SIMDEUM and a disaggregation procedure

SIMDEUM pulse labels (binary)							Disaggregation pulse labels (binary)						
time	day	Weekday	Wc	Bathtub	BathroomTap	Dishwasher	time	day	Weekday	Wc	Bathtub	BathroomTap	Dishwasher
07:48:54	8	Weekday	1	0	1	0	07:48:54	8	Weekday	1	1	1	0
07:48:55	8	Weekday	1	0	0	0	07:48:55	8	Weekday	1	1	1	1
07:48:56	8	Weekday	0	0	1	0	07:48:56	8	Weekday	1	1	1	0
07:48:57	8	Weekday	1	0	1	0	07:48:57	8	Weekday	1	1	0	1
07:48:58	8	Weekday	1	0	0	0	07:48:58	8	Weekday	0	0	1	0
07:48:59	8	Weekday	1	0	0	0	07:48:59	8	Weekday	0	0	1	1
07:49:00	8	Weekday	1	0	1	0	07:49:00	8	Weekday	1	0	0	1
07:49:01	8	Weekday	1	0	1	0	07:49:01	8	Weekday	0	0	0	0

Table 8 Example of evaluation of the labels

time	day	Weekday	Wc	Bathtub	BathroomTap	Dishwasher
07:48:54	8	Weekday	Correct label	False Positive	Correct label	Correct label
07:48:55	8	Weekday	Correct label	False Positive	False Positive	False Positive
07:48:56	8	Weekday	False Positive	False Positive	Correct label	Correct label
07:48:57	8	Weekday	Correct label	False Positive	False Negative	False Positive
07:48:58	8	Weekday	False Negative	Correct label	False Positive	Correct label
07:48:59	8	Weekday	False Negative	Correct label	False Positive	False Positive
07:49:00	8	Weekday	Correct label	Correct label	False Negative	False Positive
07:49:01	8	Weekday	False Negative	Correct label	False Negative	Correct label
Correct label			50%	50%	25%	50%
False positives			13%	0%	38%	50%
False negatives			38%	50%	38%	0%

Table 9 Evaluation of correct pulse labels during single end-use vs simultaneous end-use

	% Correct labels	
Simultaneous use	Yes	No
Yes	30%	35%
No	15%	20%

- Evaluation of the individual events: the individual events in the 'tap_log_processed.csv' will be compared with the original 'tap_log.csv'.
The number of events for each end use will be compared to the number of events of the original data (Table 10 and Table 11). We report on the absolute difference in the number of event (the absolute sum of the number of event above and below the true value), the average difference, and absolute average different (Table 12).

NOTE: the WashingMachine and the Dishwasher have multiple water intake moments when they are used. All intake moments for a single program count as a single 'event'.

Table 10 Example of event numbers from the original SIMDEUM data

SIMDEUM event data (number of events)								
day	BathroomTap	Bathtub	Dishwasher	KitchenTap	Shower	WashingMachine	Wc	Total
8	23	0	0	22	3	2	17	67
9	25	0	1	28	4	2	27	87
10	20	0	2	7	4	3	24	60
11	23	0	0	23	4	1	14	65
12	20	0	0	3	4	0	25	52
13	23	0	1	13	3	1	24	65
14	24	0	0	10	5	1	23	63
Total	158	0	4	106	27	10	154	459

Table 11 Example of event numbers generated by a disaggregation method

Disaggregation event data (number of events)								
day	BathroomTap	Bathtub	Dishwasher	KitchenTap	Shower	WashingMachine	Wc	Total
8	25	2	1	10	4	2	22	66
9	27	0	0	8	3	0	24	62
10	21	0	1	21	2	1	26	72
11	13	0	1	39	2	0	20	75
12	20	0	1	3	4	1	24	53
13	21	0	1	21	3	3	28	77
14	18	0	3	29	3	1	24	78
Total	145	2	8	131	21	8	168	483

Table 12 Example of the comparison between the event data of Table 10 and Table 11

Comparison between SIMDEUM events and Disaggregation events (number of events)										
day	BathroomTap	Bathtub	Dishwasher	KitchenTap	Shower	WashingMachine	Wc	ABS total difference	Average difference	ABS average difference
8	2	2	1	-12	1	0	5	23	0	3
9	2	0	-1	-20	-1	-2	-3	29	-4	4
10	1	0	-1	14	-2	-2	2	22	2	3
11	-10	0	1	16	-2	-1	6	36	1	5
12	0	0	1	0	0	1	-1	3	0	0
13	-2	0	0	8	0	2	4	16	2	2
14	-6	0	3	19	-2	0	1	31	2	4
ABS total difference	23	2	8	89	8	8	22			
Average difference	-2	0	1	4	-1	0	2			
ABS average difference	1.6	0.1	0.6	6.4	0.6	0.6	1.6			

- a. The start and end time for each event of the disaggregation will be compared to the original data (Table 13). We expect a margin of error will be needed for this, but the magnitude remains to be decided.

Table 13 Example of the comparison of SIMDEUM and disaggregation results for the timing of individual events

SIMDEUM event data				Disaggregation event data				Within error margin (5min)
time	day	weekday	end use	time	day	weekday	end use	
00:48:07	1	Weekday	BathroomTap	00:40:07	1	Weekday	BathroomTap	No
04:21:48	1	Weekday	BathroomTap	04:23:48	1	Weekday	BathroomTap	Yes
05:04:09	1	Weekday	Wc	05:24:09	1	Weekday	Wc	No
06:02:58	1	Weekday	Wc	06:07:48	1	Weekday	Wc	Yes
06:13:04	1	Weekday	Wc	06:12:04	1	Weekday	Wc	Yes
06:27:16	1	Weekday	WashingMachine	07:10:16	1	Weekday	WashingMachine	No
06:43:53	1	Weekday	BathroomTap	06:41:53	1	Weekday	BathroomTap	Yes
							Percentage within margin	57%

- b. The water volume per event of the disaggregation will be compared to the original data. This comparison only takes place for events that fall within the temporal error margin defined in the previous step (Table 14). These figures will be aggregated per end-use to calculate the total volume of water which was disaggregated within the temporal error margin. There is the possibility to perform this evaluation with several temporal error margins. For each error margin we will report on the total water volume characterized correctly/incorrectly on event level.

Table 14 Example of event based water consumption by volume comparison

SIMDEUM event data					Disaggregation event data					Difference in water use by volume (%)
time	day	weekday	end use	Water use (L)	time	day	weekday	end use	Water use (L)	
04:21:48	1	Weekday	BathroomTap	0.34	04:23:48	1	Weekday	BathroomTap	0.46	35%
06:02:58	1	Weekday	Wc	3.6	06:07:48	1	Weekday	Wc	3.5	-3%
06:13:04	1	Weekday	Wc	6	06:12:04	1	Weekday	Wc	5.7	-5%
06:43:53	1	Weekday	BathroomTap	1.35	06:41:53	1	Weekday	BathroomTap	0.7	-48%

Project timeline

The overall planning for the rest of the project is shown in *Table 15*. Please review this timeline and let us know if it is feasible for you.

Table 15 Project timeline

Period/moment	Activities	Responsible
1 st of September	Dataset is sent to participants	KWR
1 st September – 15 th October	Disaggregation of dataset by participants	Participants
15 th October – 10 th November	Evaluation of disaggregation results – Preliminary outcomes are shared with participants to discuss final publication restrictions (Question 10 of questionnaire)	KWR & Participants
10 th November – 30 th November	Report writing	KWR
1 st December – 24 th December	Quality assurance and project completion	KWR

IV Begeleidende brief oplossingen partij A

Cover letter from [REDACTED]

First of all, thank you very much for letting us participate in this project, it has been great fun. Our results are delivered in the requested structure using .csv format. If something is missing or you have questions to the delivered, please do not hesitate to reach out. Below we have some remarks regarding our results.

Our solutions are both based on targeting single events. When a consumption event takes place a feature set is generated. Based on the training sets provided the most probable label can be selected using an ML enhanced selection process. Multi usage events are trickier. These are tried to be split into single events and afterwards handled as single events.

We deliver two solutions. The first (solution [REDACTED]) do not use the information contained in the composition_data and startup_procol files. Therefore, the appliance_data file including all 100 homes is also contained in this solution. The other one (solution [REDACTED]) have used the composition_data and the appliance_data file is not included. The two solutions also differ in the sense of how multi events are handled. The first solution is rule based whereas the other is a fitting approach. In general, we find that the most uncertain label is the tab label since some of the tap consumptions are quite similar.

In the answer files (tap_log_proceeded and labeled_pulses) we only label the data where there is consumption i.e., the non-consuming periods of the washing machine and dishwasher are not labelled. In the case of the washing machine, this results in 4 label periods for a single event. The reason for this approach is that a meter-based prediction takes place immediately after the end of an event (defined by end of water usage) to simplify things.

During our time spent with the provided data we have collected some observations that we would like to share with you.

- The constant consumption devices such as Wc, Dishwasher and Washing machine could have benefited from noise. Since system pressure varies over the day, the filling time of these devices will also vary.
- Maybe we are not smart enough, but we have not been able to find a correspondence between the use of BathroomTab and Wc. We would have expected this.
- There are some funny things going on in the data, when looking at the provided labels. An example is a KitchenTap beginning within a present KitchenTab event.

16:45:17;1;Weekday;KitchenTap;49.0

16:45:48;1;Weekday;KitchenTap;8.0

During this 8s event the consumption is lower than it was before and is after. We are not saying that this is not right, but we found it challenging to interpretating what was meant to be going on. We see the same for Wc where a new Wc begins within an existing Wc event, without changing the flow rate.

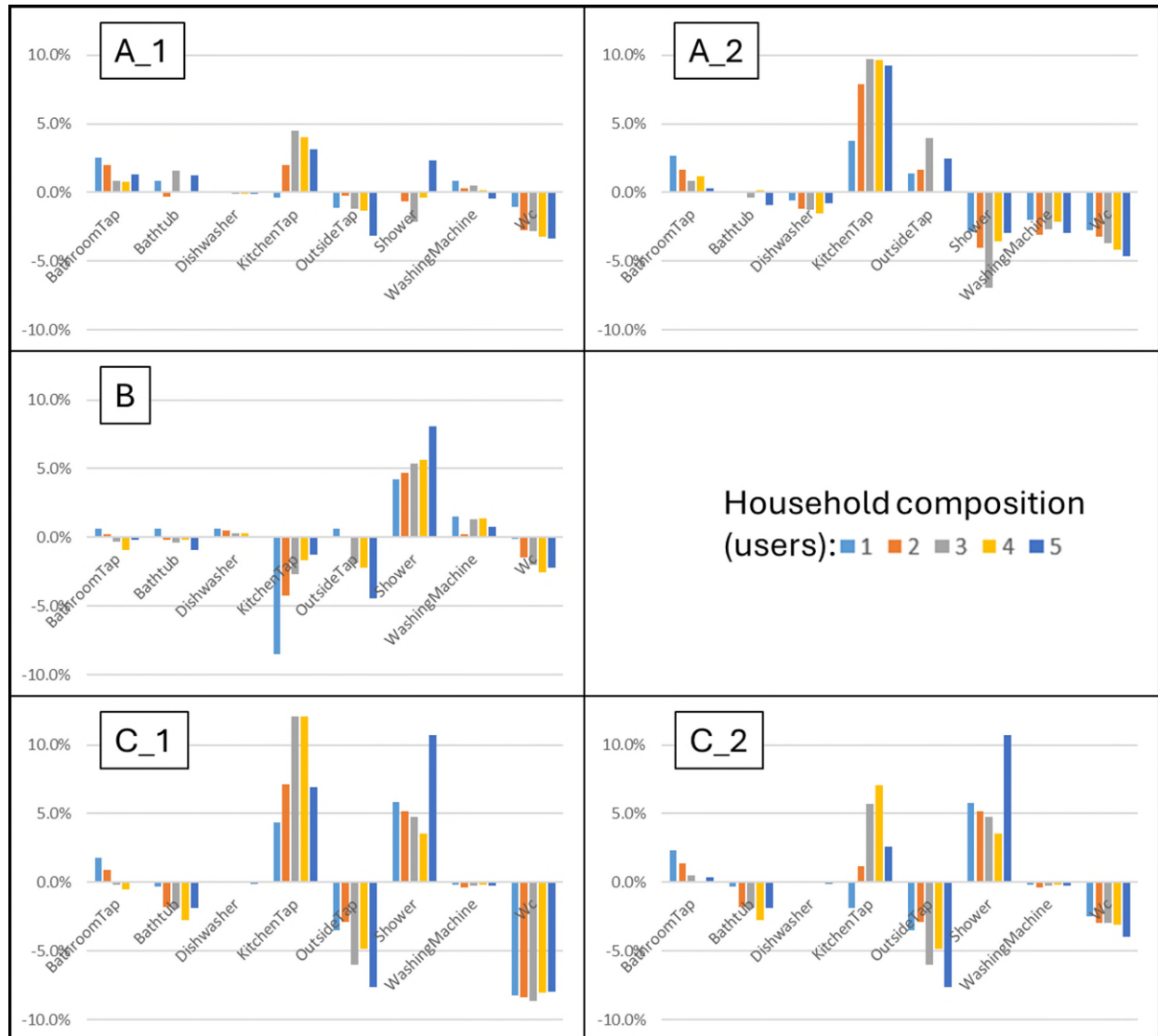
- Leakages like running toilets would have been interesting additional aspect to include. We know that many of these exists and complicates the analysis.

Once again this has been great fun, and we are looking forward to your analysis of the data.

Kind regards

[REDACTED]

V Prestaties van de disaggregatiemethoden naar huishoudsamenstelling



VI Volumetrisch waterverbruik per uur en per dag vergeleken met het totale waterverbruik

VI.I A_1

day	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
8	0.27%	0.02%	0.00%	0.37%	-0.23%	-0.08%	0.06%	-0.31%
9	0.26%	0.00%	0.00%	0.47%	-0.05%	-0.12%	0.00%	-0.46%
10	0.18%	0.25%	0.00%	0.13%	-0.42%	0.08%	0.13%	-0.29%
11	0.24%	0.14%	-0.01%	0.28%	-0.13%	-0.01%	-0.02%	-0.43%
12	0.21%	-0.04%	-0.01%	0.30%	-0.02%	-0.18%	0.10%	-0.34%
13	0.18%	-0.06%	0.00%	0.47%	-0.10%	-0.09%	0.03%	-0.33%
14	0.21%	0.17%	-0.02%	0.52%	-0.09%	-0.15%	0.05%	-0.44%

hour	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
0	-0.03%	0.00%	0.00%	0.18%	0.01%	-0.03%	0.02%	-0.14%
1	0.01%	0.00%	0.00%	0.18%	-0.06%	0.01%	-0.01%	-0.12%
2	0.00%	0.00%	0.00%	0.08%	-0.03%	0.00%	0.00%	-0.05%
3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
4	0.00%	0.00%	0.00%	0.02%	-0.01%	0.01%	0.00%	-0.01%
5	0.00%	-0.06%	0.00%	0.03%	0.03%	0.02%	0.00%	-0.01%
6	0.05%	0.00%	0.00%	0.04%	-0.02%	-0.01%	0.02%	-0.09%
7	0.18%	0.09%	0.01%	0.19%	0.03%	-0.08%	-0.02%	-0.28%
8	0.21%	0.23%	0.01%	0.29%	-0.04%	-0.34%	0.12%	-0.32%
9	0.09%	0.18%	0.00%	0.20%	-0.05%	-0.07%	0.00%	-0.25%
10	0.10%	0.13%	0.00%	0.17%	-0.08%	-0.05%	-0.04%	-0.11%
11	0.16%	-0.06%	0.00%	-0.03%	-0.06%	0.06%	0.03%	-0.08%
12	0.12%	-0.06%	0.00%	-0.08%	0.04%	0.00%	-0.01%	0.02%
13	0.07%	0.00%	0.01%	0.03%	-0.25%	0.23%	0.05%	-0.11%
14	0.04%	0.00%	0.00%	0.08%	-0.08%	-0.01%	0.04%	-0.07%
15	0.05%	0.00%	0.00%	0.18%	-0.11%	-0.04%	0.00%	-0.08%
16	0.13%	0.00%	0.00%	0.05%	-0.02%	-0.04%	0.02%	-0.09%
17	0.10%	0.00%	0.01%	-0.04%	0.06%	-0.04%	0.03%	-0.10%
18	0.12%	0.00%	-0.01%	-0.03%	-0.01%	-0.06%	0.02%	-0.08%
19	0.06%	-0.06%	-0.01%	0.33%	-0.04%	-0.10%	0.02%	-0.15%
20	0.02%	-0.04%	-0.01%	0.20%	0.01%	-0.04%	-0.03%	-0.08%
21	0.03%	-0.03%	-0.01%	0.14%	-0.07%	0.03%	0.02%	-0.13%
22	0.04%	0.10%	-0.02%	0.06%	-0.12%	0.06%	-0.01%	-0.10%
23	0.00%	0.06%	-0.02%	0.28%	-0.18%	-0.04%	0.08%	-0.18%

VI.II A_2

day	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
8	0.38%	-0.06%	-0.09%	1.18%	0.13%	-0.66%	-0.43%	-0.50%
9	0.23%	0.00%	-0.17%	1.08%	0.18%	-0.47%	-0.37%	-0.55%
10	0.26%	0.10%	-0.17%	1.31%	-0.21%	-0.44%	-0.42%	-0.48%
11	0.27%	0.00%	-0.15%	0.92%	0.35%	-0.59%	-0.37%	-0.47%
12	0.18%	0.00%	-0.13%	1.12%	0.50%	-0.87%	-0.33%	-0.54%
13	0.05%	-0.06%	-0.23%	1.04%	0.36%	-0.41%	-0.29%	-0.49%
14	0.13%	-0.06%	-0.19%	1.26%	0.37%	-0.69%	-0.35%	-0.51%

hour	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
0	-0.17%	0.14%	0.05%	-0.01%	-0.08%	-1.68%	1.83%	-1.50%
1	-0.10%	0.16%	-0.03%	0.03%	-0.13%	-1.16%	1.28%	-1.11%
2	-0.04%	0.03%	-0.04%	0.02%	-0.02%	-0.39%	0.43%	-0.36%
3	-0.01%	0.01%	-0.02%	0.02%	0.01%	-0.19%	0.18%	-0.10%
4	-0.01%	0.03%	-0.01%	0.01%	-0.11%	-0.15%	0.25%	-0.15%
5	-0.04%	-0.01%	0.06%	-0.01%	-0.01%	-0.19%	0.20%	-0.30%
6	-0.12%	0.11%	0.06%	-0.37%	0.45%	-1.18%	1.08%	-0.92%
7	-0.26%	0.41%	0.17%	-1.56%	2.22%	-2.58%	1.87%	-2.54%
8	-0.30%	0.52%	0.09%	-1.72%	2.65%	-3.64%	2.34%	-2.60%
9	-0.19%	0.23%	-0.01%	-0.83%	1.12%	-2.25%	1.29%	-1.20%
10	-0.12%	0.08%	0.00%	-0.38%	0.59%	-1.24%	0.03%	-0.12%
11	-0.11%	0.01%	0.06%	-0.85%	1.12%	-1.22%	0.30%	-0.26%
12	-0.06%	0.01%	0.01%	-0.79%	1.02%	-0.78%	0.07%	-0.25%
13	-0.09%	-0.01%	0.09%	-0.42%	0.28%	-0.73%	0.57%	-0.64%
14	-0.12%	0.22%	-0.01%	-0.18%	0.27%	-1.18%	0.62%	-0.48%
15	-0.13%	0.14%	0.04%	-0.26%	0.26%	-0.83%	0.52%	-0.63%
16	-0.10%	0.15%	-0.01%	-0.79%	0.92%	-1.37%	1.11%	-0.79%
17	-0.15%	0.14%	0.00%	-1.06%	1.46%	-1.60%	1.27%	-0.84%
18	-0.12%	-0.01%	-0.02%	-0.82%	1.28%	-1.15%	0.81%	-0.97%
19	-0.15%	0.18%	-0.07%	-0.33%	0.84%	-2.22%	1.73%	-1.19%
20	-0.14%	0.06%	0.00%	-0.24%	0.41%	-1.61%	1.30%	-1.24%
21	-0.15%	0.05%	-0.03%	-0.15%	0.15%	-1.63%	1.69%	-1.57%
22	-0.19%	0.16%	0.03%	-0.22%	0.06%	-1.60%	1.74%	-1.67%
23	-0.20%	0.16%	-0.13%	0.09%	-0.12%	-1.76%	2.01%	-1.63%

VI.III

day	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
8	0.03%	-0.01%	0.08%	-0.51%	-0.19%	0.77%	0.00%	-0.19%
9	0.02%	0.00%	0.06%	-0.65%	0.14%	0.66%	0.09%	-0.29%
10	-0.05%	0.00%	0.06%	-0.69%	-0.58%	1.28%	0.21%	-0.25%
11	0.02%	0.00%	0.06%	-0.64%	-0.06%	0.69%	0.10%	-0.20%
12	-0.04%	0.02%	0.05%	-0.67%	-0.07%	0.52%	0.30%	-0.17%
13	-0.06%	-0.06%	0.04%	-0.19%	-0.14%	0.51%	0.10%	-0.23%
14	0.01%	-0.06%	0.07%	-0.61%	-0.13%	0.75%	0.18%	-0.27%

hour	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
0	-0.05%	0.00%	0.00%	-0.02%	-0.03%	-0.06%	0.00%	-0.12%
1	-0.02%	0.00%	0.00%	0.02%	-0.04%	0.17%	-0.01%	-0.08%
2	-0.01%	0.00%	0.00%	0.00%	0.01%	0.05%	-0.01%	-0.02%
3	0.00%	0.00%	0.00%	0.01%	0.00%	0.03%	0.00%	-0.01%
4	0.00%	0.00%	0.00%	0.01%	-0.01%	0.04%	0.00%	0.00%
5	-0.02%	0.00%	0.00%	0.00%	-0.02%	0.06%	0.00%	0.00%
6	0.01%	0.00%	0.02%	-0.14%	0.05%	0.11%	-0.03%	-0.04%
7	0.02%	-0.05%	0.06%	-0.56%	0.09%	0.60%	0.00%	-0.23%
8	0.04%	0.00%	0.07%	-0.74%	-0.14%	0.87%	0.01%	-0.22%
9	-0.03%	0.00%	0.03%	-0.28%	0.08%	0.42%	0.16%	-0.23%
10	0.00%	-0.06%	0.02%	-0.16%	-0.07%	0.35%	0.01%	-0.03%
11	0.06%	0.00%	0.04%	-0.41%	0.02%	0.16%	0.15%	0.04%
12	0.05%	0.00%	0.03%	-0.32%	0.00%	0.14%	0.06%	0.06%
13	0.02%	0.00%	0.02%	-0.16%	-0.30%	0.49%	-0.06%	-0.06%
14	-0.05%	0.00%	0.01%	-0.01%	-0.06%	0.13%	0.08%	-0.05%
15	-0.02%	0.00%	0.01%	0.00%	-0.34%	0.20%	0.26%	-0.06%
16	0.03%	0.06%	0.01%	-0.31%	0.03%	0.14%	0.03%	0.01%
17	0.00%	0.06%	0.05%	-0.50%	0.15%	0.24%	0.09%	-0.06%
18	0.02%	0.00%	0.04%	-0.32%	0.06%	0.12%	0.02%	0.02%
19	-0.03%	0.00%	0.00%	-0.10%	-0.02%	0.23%	0.07%	-0.09%
20	-0.02%	-0.06%	-0.01%	0.01%	-0.06%	0.20%	-0.01%	-0.05%
21	0.00%	-0.06%	0.01%	0.06%	-0.08%	0.14%	0.16%	-0.14%
22	-0.04%	0.00%	0.01%	-0.08%	-0.22%	0.46%	-0.03%	-0.09%
23	-0.05%	0.00%	-0.01%	0.07%	-0.15%	-0.10%	0.03%	-0.15%

VI.IV C_1

day	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
8	0.07%	-0.32%	-0.01%	1.09%	-0.61%	0.95%	-0.05%	-1.15%
9	0.08%	-0.13%	-0.01%	1.24%	-0.53%	0.60%	-0.04%	-1.22%
10	0.06%	-0.13%	0.01%	1.22%	-1.06%	1.00%	-0.03%	-1.08%
11	0.12%	-0.19%	0.00%	1.12%	-0.56%	0.78%	-0.01%	-1.28%
12	0.05%	-0.39%	0.00%	1.34%	-0.40%	0.71%	-0.05%	-1.28%
13	0.01%	-0.26%	0.01%	1.38%	-0.67%	0.64%	-0.06%	-1.07%
14	0.09%	-0.32%	0.01%	1.33%	-0.52%	0.56%	-0.03%	-1.20%

hour	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
0	0.05%	-0.06%	0.00%	0.43%	-0.12%	0.26%	0.00%	-0.28%
1	0.07%	0.00%	0.00%	0.26%	-0.14%	-0.01%	0.01%	-0.21%
2	0.02%	0.00%	0.00%	0.10%	-0.04%	-0.02%	0.01%	-0.09%
3	0.00%	0.00%	0.00%	0.02%	0.00%	-0.02%	0.00%	-0.03%
4	0.00%	0.00%	0.00%	0.14%	-0.10%	-0.02%	0.00%	-0.04%
5	0.02%	-0.07%	0.00%	0.14%	0.00%	-0.01%	0.00%	-0.10%
6	0.11%	-0.06%	-0.02%	0.26%	-0.15%	0.10%	0.03%	-0.25%
7	0.40%	-0.14%	-0.04%	0.87%	-0.34%	0.00%	-0.06%	-0.64%
8	0.39%	-0.13%	-0.06%	1.05%	-0.28%	-0.17%	-0.01%	-0.71%
9	0.09%	-0.07%	-0.03%	0.82%	-0.38%	0.03%	-0.18%	-0.44%
10	-0.03%	-0.06%	-0.02%	0.79%	-0.19%	-0.16%	-0.07%	-0.32%
11	-0.07%	-0.13%	-0.04%	0.70%	-0.19%	0.15%	-0.16%	-0.31%
12	-0.06%	-0.07%	-0.03%	0.55%	-0.07%	-0.02%	-0.08%	-0.27%
13	-0.04%	-0.12%	-0.01%	0.73%	-0.12%	-0.20%	0.04%	-0.24%
14	0.01%	-0.01%	-0.01%	0.36%	-0.10%	-0.03%	-0.09%	-0.19%
15	-0.05%	-0.06%	0.00%	0.65%	-0.13%	-0.01%	-0.28%	-0.17%
16	-0.05%	-0.07%	-0.01%	0.52%	-0.08%	-0.01%	-0.05%	-0.25%
17	-0.01%	-0.11%	-0.03%	0.58%	-0.26%	0.09%	-0.08%	-0.21%
18	-0.05%	-0.08%	-0.04%	0.45%	-0.09%	0.16%	-0.02%	-0.29%
19	0.01%	-0.12%	-0.01%	0.58%	-0.10%	-0.07%	-0.08%	-0.27%
20	-0.05%	-0.06%	0.01%	0.52%	-0.09%	0.00%	-0.01%	-0.32%
21	-0.08%	-0.07%	-0.01%	0.72%	-0.10%	-0.08%	-0.17%	-0.31%
22	-0.05%	-0.07%	-0.03%	0.68%	-0.12%	-0.11%	0.02%	-0.34%
23	-0.06%	-0.07%	-0.03%	0.75%	-0.11%	0.22%	-0.04%	-0.39%

VI.V C_2

day	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
8	0.14%	-0.32%	-0.01%	0.26%	-0.61%	0.95%	-0.05%	-0.39%
9	0.20%	-0.13%	-0.01%	0.36%	-0.53%	0.60%	-0.04%	-0.45%
10	0.12%	-0.13%	0.01%	0.41%	-1.06%	1.00%	-0.03%	-0.33%
11	0.18%	-0.19%	0.00%	0.25%	-0.56%	0.78%	-0.01%	-0.48%
12	0.12%	-0.39%	0.00%	0.51%	-0.40%	0.71%	-0.05%	-0.52%
13	0.11%	-0.26%	0.01%	0.55%	-0.67%	0.63%	-0.06%	-0.34%
14	0.14%	-0.32%	0.01%	0.53%	-0.52%	0.56%	-0.03%	-0.45%

hour	BTap	BTub	Dish	Ktap	OTap	Show	Wash	Wc
0	0.00%	-0.06%	0.00%	0.13%	-0.14%	0.19%	0.00%	-0.12%
1	0.04%	0.00%	0.00%	0.09%	-0.18%	0.15%	0.00%	-0.11%
2	0.01%	0.00%	0.00%	0.03%	-0.03%	0.03%	0.00%	-0.04%
3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
4	0.00%	0.00%	0.00%	0.10%	-0.12%	0.01%	0.00%	0.00%
5	0.00%	-0.06%	0.00%	0.06%	-0.02%	0.05%	0.00%	-0.02%
6	0.12%	-0.06%	0.00%	-0.08%	-0.10%	0.21%	0.01%	-0.10%
7	0.42%	-0.19%	0.02%	-0.18%	-0.25%	0.60%	-0.06%	-0.39%
8	0.43%	-0.13%	0.01%	-0.24%	-0.42%	0.70%	0.00%	-0.38%
9	0.09%	-0.06%	0.00%	0.11%	-0.30%	0.45%	-0.02%	-0.27%
10	0.01%	-0.13%	0.00%	0.33%	-0.26%	0.20%	-0.06%	-0.09%
11	0.02%	-0.13%	0.00%	0.08%	-0.17%	0.31%	-0.01%	-0.10%
12	0.02%	-0.06%	0.00%	0.05%	-0.07%	0.12%	-0.02%	-0.05%
13	0.01%	-0.12%	0.00%	0.37%	-0.42%	0.29%	-0.02%	-0.13%
14	-0.02%	-0.01%	0.00%	0.17%	-0.16%	0.10%	-0.01%	-0.08%
15	-0.05%	-0.06%	0.01%	0.47%	-0.47%	0.19%	-0.02%	-0.08%
16	0.02%	-0.02%	0.00%	0.01%	-0.05%	0.13%	-0.02%	-0.08%
17	0.02%	-0.05%	0.02%	-0.10%	-0.11%	0.33%	0.01%	-0.12%
18	-0.01%	-0.08%	0.00%	-0.06%	-0.03%	0.28%	0.00%	-0.11%
19	0.01%	-0.12%	-0.01%	0.21%	-0.12%	0.15%	-0.01%	-0.13%
20	-0.02%	-0.12%	0.00%	0.22%	-0.15%	0.19%	-0.02%	-0.10%
21	-0.03%	-0.14%	-0.01%	0.42%	-0.17%	0.06%	-0.01%	-0.14%
22	-0.03%	-0.06%	-0.02%	0.23%	-0.35%	0.35%	0.00%	-0.12%
23	-0.06%	-0.06%	-0.04%	0.43%	-0.26%	0.12%	0.00%	-0.19%

VII Vervolgstappen disaggregatieonderzoek

Op basis van de resultaten is betrokken drinkwaterbedrijven en Vewin gevraagd naar hun wensen en voorkeuren bij vervolg van onderzoek naar disaggregatie methodes voor digitale watermeter data.

Vewin, Waternet & Vitens

Disaggregatie digitale watermeterdata naar gebruikstoepassing, hoe verder?

De resultaten van het onderzoek naar de betrouwbaarheid van software voor disaggregatie zijn hoopgevend. Om disaggregatie te kunnen inzetten ter vervanging van het periodiek onderzoek Watergebruik Thuis is nog wel vervolgonderzoek nodig, gericht op de doelen van Watergebruik Thuis: basis voor prognoses en informeren maatschappij.

Basisinformatie voor prognoses

Voor de drinkwatersector zijn de uitkomsten Watergebruik Thuis belangrijk als basis voor het opstellen van betrouwbare drinkwaterprognoses. In het landelijk prognosemodel wordt het toekomstig gebruik per toepassing (douche, toilet etc.) voorspeld op basis van 1) geprognostiseerd hoofdelijk gebruik (liter per persoon per dag) per leeftijdscategorie in een jaar x 2) aantal inwoners per leeftijdscategorie in het jaar zoals voorspeld door het CBS. Onderdeel 1 wordt geprognostiseerd op basis van het product van de verwachte ontwikkeling van a) de penetratie in Nederland van de toepassing (bijv. vaatwasser) op persoonsniveau x b) de frequentie van het gebruik van de toepassing per persoon x c) de hoeveelheid drinkwater die per keer door de toepassing wordt gebruikt.

De digitale watermeter maakt helaas geen onderscheid van het gebruik naar gezinslid, maar zal in ieder geval op huishoudensniveau de factoren penetratie x frequentie x gebruik per keer betrouwbaar moeten kunnen vaststellen. Voor de omrekening naar persoonsniveau zal het nodig blijven om extra informatie op te vragen.

Het CBS rapport Watergebruik Thuis 2021 en (bijbehorende Excel tabellen set) bevat per component informatie over de betrouwbaarheid van de geschatte penetraties, frequenties en gebruiken per keer. Hiertoe heeft het CBS bij elk van deze factoren een bijbehorende standaard afwijking gerapporteerd. Als we Watergebruik Thuis opstellen d.m.v. disaggregatie software willen we minimaal een gelijke betrouwbaarheid bereiken.

Informeren maatschappij

Wat betreft de doelstelling om de maatschappij te informeren over soorten watergebruik krijgen we, indien we alleen gebruik maken van disaggregatie, te maken met informatieverlies. Er worden minder gebruikstoepassingen onderscheiden dan in de vroegere dagboekonderzoeken. Er zal een grotere post “overig gebruik” zijn, omdat de disaggregatiesoftware niet alle componenten kan onderscheiden (denk aan handen wassen, wastafel, consumptie). Als we deze componenten in de toekomst willen blijven weten, dan zou de post ‘overig’ via een aanvullende vragenlijst kunnen worden uitgevraagd.

Vervolgonderzoek

- Wat is de betrouwbaarheid van elk van de componenten penetratie, frequentie en gebruik per keer zoals vastgesteld door de disaggregatiesoftware van de leveranciers?

- Kan de betrouwbaarheid van de software van de leveranciers vergroot worden, bijvoorbeeld doordat zij hun algoritmes trainen met gebruikscurven van soorten gebruik?
- Kan een analyse gemaakt worden van gerapporteerde gebruikstoepassingen in de vroegere rapportage Watergebruik Thuis vs. de technische mogelijkheden van disaggregatie ten aanzien van onderscheid in gebruikstoepassingen? Welke componenten kunnen we niet meer apart weergeven? Neem bij de analyse ook het kraantje in het toilet mee. Het is onduidelijk of deze component in de gerapporteerde disaggregaties wordt meegenomen bij het toiletgebruik of bij andere kranen.
- Kunnen de bestanden voor disaggregatie die voor onderzoek worden geleverd aan de software leveranciers meer praktijkgericht worden gemaakt (Simdeum gaat bijvoorbeeld bij de douche uit van vnl. douchen in de ochtend, terwijl in de avond en in de middag ook vaak gedoucht wordt)?
- En/of kunnen digitale watermeter data uit de praktijk worden gebruikt?
- Kan disaggregatie ook gedaan worden met minder gevoelige data dan op 0,1 liter nauwkeurig en/of met langer tijdsinterval dan op secondebasis?
 - o Wat zijn de maximaal toelaatbare tijdsintervallen per toepassing Watergebruik Thuis?
 - o Wat is de bijbehorende minimale gevoeligheid van de slimme watermeter per toepassing?
 - o Kan op basis van deze kennis een analyse gemaakt worden t.a.v. welke componenten er wel/niet nog separaat moeten worden vastgesteld, afgewogen tegen de noodzaak om die component apart te weten en de kosten die dat met zich meebrengt vanwege benodigde gevoeligheid en meetfrequentie van techniek / apparatuur.
- Kunnen we alles afwegende de vereisten definiëren van de aan te schaffen digitale watermeters nodig voor disaggregatie?
- Met die meters zou in een volgende stap een pilot uitgevoerd kunnen worden, waarin bij een aantal mensen een digitale watermeters wordt geplaatst en, ter vergelijking, tegelijkertijd de dagboekmethode wordt toegepast.

Evides

Evides denk dat het goed zou zijn om het vervolgonderzoek naar de haalbaarheid en het nut van disaggregatie van watermeterdata te koppelen aan een duidelijke use case. Hiervoor zou bijvoorbeeld het in kaart brengen van het watergebruik thuis met als doel waterbesparing een interessante keuze kunnen zijn. Het vervolgonderzoek heeft als doel om aan te tonen dat, disaggregatie van watermeterdata verkregen van fysieke watermeters met verbruikerspatronen gegenereerd door huishoudens, mogelijk is. Evides zelf heeft komende tijd nog geen concrete ambities om te vervolgonderzoek/experimenten uit te voeren met de disaggregatie van watermeterdata, dit kan uiteraard wijzigen op basis van de uitkomsten van het vervolgonderzoek.