

NEDU steekproef allocatie

Auteur	Jan Telman
Datum	April 2021
Referentie	RAP/20xxx/Tmn
Project nummer	CQM1965

Samenvatting en conclusies

Dit rapport beschrijft de resultaten van een onderzoek naar de opzet van een steekproefmethodiek ten behoeve van de allocatie van de totale netuitwisseling naar te onderscheiden categorieën van aansluitingen. Het onderzoek is uitgevoerd in opdracht van en nauwe samenwerking met de NEDU-werkgroep B1.

Doelstelling

De doelstelling is gekwantificeerd in termen van de volgende eis: de steekproefgrootte moet zodanig worden gekozen dat de nauwkeurigheid van de voorspelde totale afname (c.q. Invoeding) in een categorie binnen 10% ligt met een betrouwbaarheid van 98%. Dit is vertaald naar de eis dat de relatieve standaardafwijking maximaal 4,3% mag bedragen. Het opleveren van een tabel met steekproefgroottes is de ultieme deliverable van dit project.

Data en methodiek

Hoofdstuk 2 gaat in op de beschikbare gegevens. De focus ligt op 3 categorieën, E1B, E1C en E2B, waarbij onderscheid wordt gemaakt naar aansluitingen zonder en met invoeding. Verder wordt er onderscheid gemaakt in de 17 netgebieden. De analyses zijn uitgevoerd op basis van de steekproefdata van 2020 en de C-AR data over die periode.

Hoofdstuk 3 beschrijft op hoofdlijnen de aanpak bij de berekening van de allocatiewaarden en de betrouwbaarheidsmarge daarop. Meer details zijn nog opgenomen in de Appendix. Belangrijkste resultaat van deze aanpak is de formulematige route van steekproefdata (aantal, gemiddelde, spreiding) en populatiedata (totaal aantal aansluitingen, gemiddelde SJA) naar een relatieve standaardafwijking (RSD), zowel per netgebied per categorie als over categorieën heen.

Resultaten - steekproeftabel

Hoofdstuk 4 beschrijft de resultaten bij de huidige steekproefgroottes. Overall is in 2020 de steekproefgrootte 3.584 aansluitingen over de verschillende categorieën en netgebieden. Dit is ongeveer 0,05% van het totaal aantal aansluitingen. Het blijkt dat de huidige steekproef voor vrijwel alle netgebied en categorie combinaties een RSD oplevert die niet voldoet aan de gestelde eis van een RSD van 4,3%.

Daarom zijn berekeningen uitgevoerd om te bepalen wat de steekproef (per netgebied en per categorie) minimaal moet zijn om wél aan de eis van $RSD = 4,3\%$ te kunnen voldoen. Dit is in 2 stappen uitgevoerd. In een eerste stap (in hoofdstuk 5) zijn de berekeningen voor alle combinaties uitgevoerd volgens de daartoe geschikte methodiek. En in een tweede stap (hoofdstuk 6) zijn de steekproefgroottes herberekend voor de netgebied-categorie combinaties waarvoor in de oorspronkelijke steekproef zo weinig data zitten dat een rekenkundige extrapolatie niet betrouwbaar genoeg is. In die situaties zijn de steekproefgroottes afgeleid uit die van naburige gebieden.

Overall is een steekproefgrootte van 27.784 berekend (0,36% van het totaal aantal aansluitingen), ongeveer 8 maal de oorspronkelijke steekproef. Meer details staan in de steekproeftabel in hoofdstuk 6.

Onderzoek naar combinatiemogelijkheden

Onderzocht is nog of het mogelijk is om bepaalde gebieden of bepaalde categorieën te combineren. Dan zou een gezamenlijke steekproef over zo'n combinatie kunnen worden overwogen. Hoofdstuk 7 beschrijft een methodiek om verschillen tussen de profielen te berekenen en toont de resultaten. De voorlopige conclusie is niet hoopvol. Tussen netgebieden zijn de verschillen in profielen in het algemeen groot. Tussen categorieën zijn verschillen soms kleiner, maar hier is het beeld niet consistent over bijvoorbeeld winter- en zomermaanden.

Aanbevelingen

Aanbevolen wordt om voorlopig te werken met de steekproeftabel, zoals weergegeven in hoofdstuk 6. Mogelijk moeten de steekproeven met $n=100$ nog wat worden bijgesteld bij heel kleine categorie-netgebied combinaties.

Voer over enige maanden een validatie uit om na te gaan of de spreiding tussen de aansluitingen binnen categorie-netgebied combinaties zodanig is dat met de nieuwe steekproefgrootte voldaan wordt aan de 4,3% eis voor de RSD. Deze validatie is tweeledig:

- Toetsing van de 4,3% eis voor de totale steekproef
- Vergelijking van de deelverzameling van de oude steekproef tussen 2020 en 2021. Dit biedt inzicht in hoeverre 2020 afwijkend is ten opzichte van 2021.

Verder onderzoek:

- In de huidige data is de spreiding in patronen tussen aansluitingen soms extreem groot, leidend tot een grote steekproef. Die verschillen waren niet verklaarbaar uit postcodes, maar worden veroorzaakt door onbekende specifieke kenmerken van die aansluitingen. Meer intelligentie in het steekproefmodel brengen is raadzaam. Vanuit de praktische kant door na te denken welke kenmerken van aansluitingen (zowel in de populatie als in de steekproef) bekend (kunnen) zijn en die kenmerken gebruiken om de steekproef nog een slag verder te segmenteren in meer homogene delen. Vanuit een data-science kant kan worden nagegaan welke aansluitingen 'extreem' zijn vervolgens kan worden nagegaan wat die extreme aansluiting aan gemeenschappelijke kenmerken hebben.
- Het combineren van netgebieden of categorieën om de steekproeflast te drukken blijft aantrekkelijk. In dit onderzoek is hierbij een eerste verkenning gedaan, met als conclusie in de meeste gevallen de verschillen in profielen te groot zijn. Maar deze verkenning was beperkt en een snelle slag over totalen van 2 perioden in het jaar. Aan te bevelen is:
 - De berekeningen wellicht te beperken tot kritische tijdvakken (bijvoorbeeld piekuren op werkdagen).
 - Voor die momenten te onderzoeken wat de relatie is tussen een berekend verschil in profiel en de consequenties voor de allocatiemethodiek.
 - Te bepalen welke netgebieden of welke categorieën een voldoende klein verschil in profiel hebben.
 - Deze exercitie uit te voeren voor elk van de categorie / netgebied combinaties.

1. Doelstelling

Doelstelling van dit project is het opstellen van een representatieve steekproef, waarmee de uitwisseling op het net – in termen van het profiel per 15 minuten – kan worden gealloceerd naar de te onderscheiden categorieën. Deze allocatie gebeurt dagelijks voor de voorgaande dag, afzonderlijk per netgebied.

Representativiteit van de steekproef kan worden bereikt door – per categorie en per netgebied – een random set van aansluitingen te kiezen waarvoor het verbruikspatroon wordt geregistreerd. De betrouwbaarheid van de allocatie wordt aan de ene kant beïnvloed door de variatie in de kwartierwaarden tussen de verschillende aansluitingen, en aan de andere kant door de keuze van de steekproefomvang.

Als uitgangspunt is door NEDU gesteld dat totaal over alle aansluitingen binnen een categorie en een netgebied de gesommeerde kwartierwaarde moeten worden vastgesteld met een marge van +/- 10% bij 98% betrouwbaarheid.

2. Beschikbare data

2.1 Onderscheid in categorieën, met en zonder invoeding

Het onderzoek beperkt zich tot de volgende categorieën:

- E1B: aansluitwaarde $\leq 3 \times 25A$, dubbel tarief nachtstroom (schakeling om 23:00u)
- E1C: aansluitwaarde $\leq 3 \times 25A$, dubbel tarief avondstroom (schakeling om 21:00u)
- E2B: aansluitwaarde $> 3 \times 25A$ en $\leq 3 \times 80A$, dubbel tarief.

Binnen elke categorie wordt onderscheid gemaakt tussen aansluitingen met en zonder invoeding (eigen productie), aangeduid met de afkortingen AMI en AZI.

2.2 Onderscheid in netgebieden

Bij de berekeningen is onderscheid gemaakt tussen 17 netgebieden. Deze zijn weergegeven in onderstaande Tabel 1. In de rest van dit rapport wordt gewerkt met de nummers en afgekorte namen.

gebiednr	netgebied	netbeheerder	netbeh afkorting
1	871685900000056162	Liander NWN	Liander NWN
2	871687120000052782	Liander	Liander
3	871687400000002254	Stedin Netbeheer Utrecht	Stedin Utr
4	871687800090000015	Westland Infra Netbeheer B.V.	Westland
5	871687910000219120	Enexis Brabant	Enexis NB
6	871688520000076884	Enexis Limburg	Enexis L
7	871688600000002202	Stedin Netbeheer Delfland	Stedin Delfland
8	871689200000010161	Stedin Netbeheer	Stedin
9	871690200000000007	Enduris Elektriciteit	Enduris
10	871690499910000003	Enexis Maastricht	Enexis Maastr
11	871690910000025589	Liander EWR	Liander EWR
12	871691280000000008	Rendo Netbeheer	Rendo
13	871691600019188908	Coteq Netbeheer B.V.	Coteq
14	871692100000010038	Stedin Netbeheer Midden-Holland	Stedin Mid-H
15	871692510000000005	Stedin Netbeheer Schiedam	Stedin Schiedam
16	871694600000002173	Stedin Netbeheer Zuid-Kennemerland (ELK)	Stedin Z-Kenneml
17	871694830000000309	Enexis Noord	Enexis Noord

Tabel 1. Overzicht van de netgebieden.

2.3 Meetgegevens in de huidige steekproef

Voor elke aansluiting in de steekproef zijn de volgende gegevens beschikbaar:

- Kwartierwaarden (in Wh) voor 2020, dus een meetreeks van $366 \times 24 \times 4 = 35.136$ waarden.
- Het standaard jaarverbruik SJV, bij de aansluiting met invoeding (AMI) nog opgesplitst naar SJA en SJI
- De postcode in 4 cijfers.

Voor de kwartierwaarden moet onderscheid gemaakt worden tussen aansluitingen *zonder* invoeding (AZI's) en aansluitingen *met* invoeding (AMI's). De AZI's hebben alleen een verbruikswaarde V. De AMI's hebben zowel een verbruik V als een productie P, momenteel vrijwel geheel via zonnepanelen. In het onderzoek wordt steeds gesproken over een afnamepatroon A en een (eventueel) invoedingspatroon I

- Voor de AZI's geldt dat er één meetreeks beschikbaar is, namelijk alleen afname A: hierbij geldt $A = V$.
- Voor de AMI's geldt dat er twee meetreeksen beschikbaar zijn, namelijk afname A en invoeding I. Hierbij geldt dat er vóór de meterkast wordt gesaldeerd, dat wil zeggen de A is het overschot aan verbruik (na verrekening van eventuele productie) en de I is het overschot aan productie (na verrekening van eventueel verbruik). Er zijn dus 2 situaties denkbaar:
 - Overschot aan verbruik, dus $V \geq P$. Dan is $A = V - P$ en $I = 0$.
 - Overschot aan productie, dus $V < P$. Dan is $A = 0$ en $I = P - V$.

Dit betekent dat er in de meetreeksen A en I voor de AMI's een substantieel aantal nullen zullen voorkomen.

De in 2020 gehanteerde steekproefgroottes zijn weergegeven in onderstaande Tabel 2. Vermeld zijn de bemeten aantallen aansluitingen per categorie per netgebied, met onderscheid naar AZI en AMI.

gebied	netbeheerder	AZI			AMI			totaal
		E1B	E1C	E2B	E1B	E1C	E2B	
1	Liander NWN	72		34	109		51	266
2	Liander	77	3	38	222	17	124	481
3	Stedin Utr	79		41	80		27	227
4	Westland	8		3	4		3	18
5	Enexis NB		395	111		459	115	1080
6	Enexis L		202	63		288	59	612
7	Stedin Delfland	18		7	19		14	58
8	Stedin	124		42	88		40	294
9	Enduris	12		12	33		21	78
10	Enexis Maastr		17	5		8	2	32
11	Liander EWR	1	65	11		102	12	191
12	Rendo	3		1	4		2	10
13	Coteq	5		3	6		2	16
14	Stedin Mid-H	14		3	15		3	35
15	Stedin Schiedam				3			3
16	Stedin Z-Kenneml	2					2	4
17	Enexis Noord	75		71	124		105	375
	TOTAAL	490	682	445	707	874	582	3780

Tabel 2. Steekproefgroottes per categorie per netgebied.

2.4 Data in de populatie

Voor alle aansluitingen in de populatie zijn gegevens bekend uit de zogenaamde C-AR data over 2020. Deze database geeft voor alle postcodes het aantal aansluitingen, en de bijbehorende categorie en netgebied. Verder is voor ongeveer 88% van de aansluitingen bekend wat de waarde van de SJA was en wat de leveringsrichting is (LVR,TLV,CMB). Hieruit kunnen de volgende gegevens worden berekend:

- Het totaal aantal aansluitingen per categorie en per netgebied.
- De gemiddelde SJA in kWh (voor ca 88% van de aansluitingen is deze waarde bekend)
- Het percentage van de aansluitingen AZI en AMI, op basis van de aantallen per leveringsrichting: LVR = AZI; CMB of TLV = AMI.
- De gemiddelde SJA in kWh, afzonderlijk voor aansluitingen AZI en AMI.

Het resultaat staat in Tabel 3. De grijze velden markeren enkele bijzonderheden.

- Voor netgebied 2, categorie E1C is geen informatie beschikbaar over de SJA en leveringsrichting.
- Voor netgebied 16, categorie E2B werden sterk afwijkende resultaten gevonden. Deze zijn niet betrouwbaar, omdat ze gebaseerd zijn op slechts 13 aansluitingen met bekende waarden voor de SJA en de leveringsrichting.

gebiednr	netbeh afkorting	categorie E1B					categorie E1C					categorie E2B				
		N	SJA	%AZI	SJA_AZI	SJA_AMI	N	SJA	%AZI	SJA_AZI	SJA_AMI	N	SJA	%AZI	SJA_AZI	SJA_AMI
1	Liander NWN	1.177.074	2.772	87,5%	2.750	2.936	4.289	3.136	84,5%	3.075	3.473	45.929	25.371	92,8%	25.792	19.917
2	Liander	1.266.148	2.981	80,0%	2.981	2.982	4.487					63.268	24.987	89,0%	25.353	22.011
3	Stedin Utr	536.106	2.996	85,0%	3.001	2.967	0					23.108	24.143	90,4%	24.473	21.044
4	Westland	55.815	3.160	83,9%	3.139	3.274	0					3.322	26.108	87,1%	26.313	24.724
5	Enexis NB	0					1.027.649	3.186	82,9%	3.162	3.301	54.894	26.250	90,6%	27.059	18.451
6	Enexis L	0					471.398	3.133	78,2%	3.117	3.193	24.605	27.323	92,4%	27.840	21.067
7	Stedin Delfland	147.805	2.975	86,1%	2.960	3.072	0					5.569	23.273	93,9%	23.632	17.738
8	Stedin	1.051.582	2.753	91,1%	2.725	3.071	0					41.479	22.043	94,0%	22.330	17.533
9	Enduris	177.118	2.635	80,3%	2.787	2.156	0					8.728	23.459	88,9%	24.155	17.902
10	Enexis Maastr	0					41.925	2.740	91,0%	2.730	2.838	2.101	28.347	93,2%	29.451	13.325
11	Liander EWR	973	3.428	76,0%	3.491	3.242	260.726	2.952	85,2%	2.939	3.023	10.640	23.593	89,0%	24.488	16.366
12	Rendo	30.122	2.729	77,9%	2.772	2.590	0					1.209	27.404	87,4%	27.384	27.546
13	Coteq	50.238	2.833	85,6%	2.832	2.843	0					2.007	26.784	91,3%	27.550	18.704
14	Stedin Mid-H	91.700	2.991	85,6%	2.974	3.091	0					4.487	20.321	86,9%	20.737	17.553
15	Stedin Schiedam	34.008	2.653	93,3%	2.616	3.303	0					1.558	21.577	99,1%	21.579	21.381
16	Stedin Z-Kenneml	8.785	3.429	88,0%	3.436	3.375	0					393	31.995	92,3%	34.660	13
17	Enexis Noord	884.575	2.941	78,9%	2.930	2.979	1					43.780	24.443	88,3%	25.339	17.698
	TOTAAL	5.512.049	2.872	84,5%	2.859	2.946	1.810.475	3.127	82,2%	3.106	3.227	337.077	24.660	91,1%	25.170	19.430

Tabel 3. Overzicht van aantallen aansluitingen en gemiddelde SJA per categorie per gebied.

3. Globale aanpak

Per netgebied, per kwartier, is op hoofdlijnen de volgende werkwijze gehanteerd. Een meer gedetailleerde beschrijving is opgenomen in de Appendix 1.

3.1 Berekeningen per categorie

Onderscheid is gemaakt in 6 verschillende categorieën, met de volgende codering voor de totale afname en invoeding:

- E1B zonder invoeding: A1 = afname E1B_AZI_A
- E1C zonder invoeding: A2 = afname E1C_AZI_A
- E2B zonder invoeding: A3 = afname E2B_AZI_A
- E1B met invoeding: A4 = afname E1B_AMI_A en I4 = invoeding E1B_AMI_I
- E1C met invoeding: A5 = afname E1C_AMI_A en I5 = invoeding E1C_AMI_I
- E2B met invoeding: A6 = afname E2B_AMI_A en I6 = invoeding E2B_AMI_I

Schematisch is dit weergegeven in onderstaande tabel.

	aansluitingen zonder invoeding			aansluitingen met invoeding			aansluitingen met invoeding		
	afname (= verbruik)			afname (=verbruik minus productie)			invoeding (=productie minus verbruik)		
categorie	E1B	E1C	E2B	E1B	E1C	E2B	E1B	E1C	E2B
categorienr	1	2	3	4	5	6	4	5	6
totaal per cat	E1B_AZI_A	E1C_AZI_A	E2B_AZI_A	E1B_AMI_A	E1C_AMI_A	E2B_AMI_A	E1B_AMI_I	E1C_AMI_I	E2B_AMI_I
code	A1	A2	A3	A4	A5	A6	I4	I5	I6

Tabel 4. Overzicht categorieën.

Per categorie is over alle aansluitingen de totale afname A_j ($j = 1 \dots 6$) en de totale invoeding I_j ($j = 4 \dots 6$) berekend, inclusief de betrouwbaarheidsmarge.

Dit gebeurt door eerst per categorie het profiel te berekenen op basis van de meetreeksen in de steekproef, inclusief de onzekerheidsmarge op dit profiel aangeduid met een standaardafwijking.

De profielwaarden per kwartier zijn berekend als: $p_A = \text{mean}A / \text{meanSJA}_{\text{sample}}$,

waarbij meanA het gemiddelde is van de afname (of invoeding) in een bepaald kwartier en $\text{meanSJA}_{\text{sample}}$ is het gemiddelde van de standaard jaarafname (of invoeding) van de aansluitingen in de steekproef.

Deze profielwaarden tellen over alle kwartieren in het jaar op tot 1.

De onzekerheid in de profielwaarden worden gekarakteriseerd door de standaardafwijking

$s_p \approx \frac{\text{std}A / \sqrt{n}}{\text{meanSJA}_{\text{sample}}}$. Hierin is stdA de standaardafwijking van de afnames (of invoeding) in de steekproef en n de steekproefgrootte. Het $\approx$ teken geeft aan dat dit een benadering is. In de appendix wordt ingegaan op de exacte berekening.

Vervolgens is deze profielwaarde omgerekend naar de afname (of invoeding) per kwartier totaal voor alle aansluitingen in betreffende categorie door de profielwaarde te vermenigvuldigen met de SJA of SJI van een aansluiting en vervolgens op te tellen over alle aansluitingen. In de appendix wordt beschreven dat de volgende relatie gelden.

Per categorie wordt (per kwartier) de totale afname over alle aansluitingen in het netgebied berekend als:

$$A = p_A \times N \times F \times \text{meanSJA}_{\text{sample}}$$

De onzekerheid op die berekende waarde wordt berekend als de standaardafwijking

$$s_A = N \times F \times \frac{\text{std}A}{\sqrt{n}}$$

De relatieve standaardafwijking wordt dan berekend als

$$RSD = \frac{s_A}{A} = \frac{1}{p_A \times \text{meanSJA}_{\text{sample}}} \times \frac{\text{std}A}{\sqrt{n}} \times 100 \text{ (\%)}$$

Hierin geldt:

- N is het totaal aantal aansluitingen binnen betreffende categorie in het netgebied
- n is het aantal aansluitingen in de steekproef, binnen betreffende categorie en netgebied
- F is een factor, namelijk de verhouding van de gemiddelde SJA in de populatie en de steekproef, dus $F = \text{meanSJA}_{\text{pop}} / \text{meanSJA}_{\text{sample}}$.

De factor F is bedoeld om te corrigeren voor een eventueel verschil in meanSJA's in de totale populatie van aansluitingen en die in de steekproef. De meanSJA in de steekproef is te berekenen uit de bekende jaartotalen in de data. De meanSJA in de populatie kan in principe worden berekend uit de C-AR data en staan in Tabel 3 onder de kolommen SJA_AZI en SJA_AMI.

Hier trad echter een probleem op. De gemiddelde SJI's zijn niet bekend in de C-AR data en de gemiddelde SJA's gaven een grote afwijking ten opzichte van de SJA's in de steekproeven. Om deze reden is in de berekening noodzakelijkerwijs verder gewerkt met een F-factor van 1. Dit betekent dat impliciet is aangenomen dat de steekproeven qua SJA en SJI representatief zijn voor de populatie.

Merk op dat in de relatieve standaardafwijking RSD de factor F wegvalt en dus geen rol meer speelt bij de berekeningen per categorie afzonderlijk.

De berekeningen van het totaal, de standaardafwijking en de relatieve standaardafwijking gebeurt op eenzelfde wijze en afzonderlijk voor zowel afnames als eventuele invoedingen per categorie. In bovenstaande uitdrukkingen kan voor de A dus worden ingevuld A1, A2, A3, A4, A5, A6, I4, I5 of I6 voor zover van toepassing in betreffend netgebied.

3.2 Eis aan de betrouwbaarheid per categorie

Als uitgangspunt bij de berekeningen is gehanteerd dat de onzekerheidsmarge op de totale afname of invoeding maximaal +/-10% mag zijn, met 98% betrouwbaarheid. Uitgaande van de normale verdeling (Gauss-verdeling) ligt het 98%-betrouwbaarheidsinterval tussen -2,326 en +2,326 maal de standaardafwijking ten opzichte van het gemiddelde.

Praktisch vertaald betekent een marge van +/- 10% dus dat de relatieve standaardafwijking op een geschatte totale afname of invoeding maximaal $10\% / 2,326 = 4,3\%$ mag bedragen.

De aanname van de normale verdeling geldt in het algemeen niet voor de verdeling van de afnames (of invoedingen) tussen de aansluitingen binnen een kwartier. Maar vanuit de statistiek is bekend dat steekproefgemiddelden van dergelijke meetwaarden wel naar een normale verdeling tenderen.

In overleg is afgesteld dat deze eis van 4,3% voor de relatieve standaardafwijking, ofwel +/-10% onzekerheidsmarge, moet gelden:

- Voor de totale afname of invoeding, afzonderlijk per categorie en per netgebied
- Volume gewogen over alle kwartieren van het jaar.

3.3 Berekeningen van totalen over categorieën heen

Op basis van de totale afname (of invoeding) per categorie is in tweede instantie over de categorieën heen de totale afname en invoeding per kwartier berekend als:

- Totale afname $A = A1 + A2 + A3 + A4 + A5 + A6$
- Totale invoeding $I = I4 + I5 + I6$
- Totale netuitwisseling $T = A + I$
- Totale saldo $S = A - I$

Hierbij is ook de standaardafwijking voor de onzekerheid per categorie vertaald naar de standaardafwijkingen voor totalen. Daarbij worden de standaardafwijkingen per categorie op een gewogen wijze bij elkaar opgeteld, met gewichten evenredig aan de totale SJA (of SJI) per categorie. In de appendix staan meer details over de berekeningswijze.

De resultaten in termen van volume gewogen relatieve standaardafwijkingen RSD staan in de volgende hoofdstukken. Eerst is wat de resulterende RSD's zijn bij de huidige steekproefgrootte. Vervolgens is aangegeven hoe groot de steekproef moet worden gekozen om te voldoen aan de gestelde eis van $RSD = 4,3\%$.

4. Resultaten bij de huidige steekproef

De resultaten bij de huidige steekproef zijn samengevat in onderstaande Tabel 5. De kolommen onderscheiden de 17 netgebieden.

In de eerste 9 regels met percentages geven de volume gewogen RSD's voor de afname en invoeding voor elk van de onderscheiden categorieën. Met een groene markering is aangegeven waar de eis van 4,3% wordt gehaald met de huidige steekproef. Voor de afzonderlijke categorieën is dit vrijwel nooit het geval. Voor de totalen (A, I, A+I en A-I) zijn de RSD's altijd kleiner dan die voor de afzonderlijke categorieën, omdat er uitmiddeling van onzekerheden plaatsvindt.

In de rijen achter “steekproef” staan de huidige aantallen aansluiting in de steekproef. Hier kunnen kleine afwijkingen optreden ten opzichte van de eerder genoemde aantallen (zie bovenstaande Tabel 2), omdat niet altijd meetgegevens beschikbaar bleken. Belangrijk zijn de oranje velden, die aanduiden dat de steekproef < 20 was. In die situaties zijn de resultaten te onbetrouwbaar om er gefundeerde conclusies aan te verbinden. Verder zijn er velden geel gemarkeerd, waar in de huidige steekproef geen aansluitingen zitten, maar volgens de C-AR data wel aansluitingen bestaan.

In de rijen achter “populatie” staan aantallen aansluitingen per categorie totaal in betreffend netgebied, berekend op basis van de C-AR data over 2020.

De onderste rij laat zien hoe groot de verhouding is in aantallen aansluitingen in de steekproef (totaal over categorieën) en de totale populatie van aansluitingen. Overall dus 0,05% (ca 3600 op 7,7 miljoen aansluitingen).

		Liander NWN	Liander	Stedin Utr	Westland	Enexis NB	Enexis L	Stedin Delfland	Stedin	Enduris	Enexis Maastr	Liander EWR	Rendo	Coteq	Stedin Mid-H	Stedin Schiedam	Stedin Z- Kenneml	Enexis Noord	
	netgebied	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
afname	E1B_AZI	11,6%	13,2%	12,2%	27,9%			19,3%	9,7%	25,2%			34,1%	33,3%	21,9%			11,6%	
	E1C_AZI		34,6%			5,5%	7,5%				22,9%	12,5%							
	E2B_AZI	17,9%	17,1%	15,7%	26,4%	8,4%	7,9%	28,0%	14,5%	18,8%	31,4%	23,7%		42,9%	31,0%			9,4%	
	E1B_AMI	11,6%	9,2%	17,4%	34,1%			24,5%	12,8%	20,2%			40,4%	36,7%	25,4%	42,1%		12,9%	
	E1C_AMI		34,2%			6,7%	7,6%				31,4%	12,0%							
invoed	E2B_AMI	18,7%	11,8%	18,3%	15,7%	12,2%	14,5%	26,0%	17,5%	29,0%	40,7%	29,6%	13,8%	55,0%	40,9%		32,0%	12,2%	
	E1B_AMI	8,2%	5,5%	8,9%	25,9%			16,9%	9,1%	10,3%			23,6%	20,3%	13,8%	18,1%		7,0%	
	E1C_AMI		17,5%			3,6%	5,1%				23,4%	6,8%							
totalen	E2B_AMI	12,8%	8,3%	14,5%	24,4%	7,8%	12,0%	21,7%	13,4%	13,4%	29,4%	29,9%	5,5%	22,1%	23,9%		0,3%	9,2%	
	totaal A	9,3%	9,2%	9,7%	19,4%	4,1%	5,0%	16,8%	8,0%	16,0%	19,1%	10,6%	29,0%	27,6%	17,8%	42,1%	32,0%	7,6%	
	totaal I	7,8%	5,0%	8,0%	23,8%	3,3%	4,9%	16,2%	8,6%	9,0%	20,5%	7,0%	18,7%	19,9%	12,6%	18,1%	0,3%	6,3%	
	totaal A + I	5,1%	5,5%	4,7%	8,3%	2,6%	2,8%	7,5%	3,8%	8,3%	8,5%	5,0%	6,7%	11,0%	7,9%	8,5%	4,8%	4,7%	
steekproef	saldo A - I	5,7%	6,7%	5,4%	9,9%	3,1%	3,5%	8,3%	4,1%	10,5%	9,1%	5,6%	9,5%	13,9%	9,8%	15,7%	5,1%	5,8%	totaal
	E1B_AZI	70	72	73	8	.	.	16	116	12	.	1	3	5	13	.	1	71	461
	E1C_AZI	.	3	.	.	.	370	186	.	.	15	64	.	.	.	.	.	.	638
	E2B_AZI	32	36	40	3	98	55	7	41	11	5	11	1	3	3	.	.	62	408
	E1B_AMI	107	211	75	4	.	.	17	85	33	.	.	4	6	14	3	.	118	677
	E1C_AMI	.	17	.	.	444	280	.	.	.	8	98	.	.	.	.	.	.	847
	E2B_AMI	47	117	25	3	112	56	13	39	20	2	11	2	2	3	.	2	99	553
populatie	totaal	256	456	213	18	1.024	577	53	281	76	30	185	10	16	33	3	3	350	3.584
	E1B_AZI	1.029.375	1.012.819	455.803	46.833	0	0	127.227	957.903	142.168	0	739	23.456	42.990	78.457	31.721	7.735	697.862	4.655.088
	E1C_AZI	3.624	3.688	0	0	852.425	368.671	0	0	38.132	222.080	0	0	0	0	0	0	0	1.488.620
	E2B_AZI	42.633	56.339	20.882	2.895	49.734	22.728	5.230	38.996	7.757	1.957	9.468	1.056	1.833	3.901	1.544	363	38.646	305.962
	E1B_AMI	147.699	253.329	80.303	8.982	0	0	20.578	93.679	34.950	0	234	6.666	7.248	13.243	2.287	1.050	186.713	856.961
	E1C_AMI	665	799	0	0	175.224	102.727	0	0	3.793	38.646	0	0	0	0	0	0	0	321.854
	E2B_AMI	3.296	6.929	2.226	427	5.160	1.877	339	2.483	971	144	1.172	153	174	586	14	30	5.134	31.115
steekproef%	totaal	1.227.292	1.333.903	559.214	59.137	1.082.543	496.003	153.374	1.093.061	185.846	44.026	272.339	31.331	52.245	96.187	35.566	9.178	928.355	7.659.600
	steekproef%	0,02%	0,03%	0,04%	0,03%	0,09%	0,12%	0,03%	0,03%	0,04%	0,07%	0,03%	0,03%	0,03%	0,01%	0,03%	0,04%	0,05%	

Tabel 5. Resultaten bij de huidige steekproefomvang per categorie en netgebied.

5. Benodigde steekproef bij RSD < 4,3%

Op basis van de resultaten bij de huidige steekproef is vervolgens berekend hoe groot de steekproef per categorie moet worden om wel te voldoen aan de eis van 4,3% voor elke categorie afzonderlijk.

De resultaten staan in onderstaande Tabel 6. Uiteraard zijn nu alle velden voor de RSD's groen geworden. Sommige percentages zijn zelfs kleiner dan 4,3%. Dit wordt veroorzaakt door het feit dat bij de aansluitingen met invoeding (AMI) de steekproef voor de afname uiteraard dezelfde is als voor de invoeding en de steekproefgrootte is bepaald door de grootte (afname of invoeding) met de meeste spreiding.

Voorlopig is de conclusie dat de steekproef moet worden vergroot om de eis te halen. In totaal van 3.584 naar 32.045.

Deze conclusie is voorlopig, omdat niet te veel waarde moet worden gehecht aan de oranje velden. Dit betreft de situaties waarin de originele steekproef te klein was (< 20). Rekentechnisch is de berekening wel mogelijk (behalve de 3 lege oranje velden waar de oorspronkelijke steekproef 1 was), maar de extrapolatie naar de eis van 4,3% is te onbetrouwbaar. Bovendien moeten de gele velden nog worden ingevuld.

Als voorlopige oplossing kan in betreffende netgebieden beter de steekproef worden overgenomen van de naburige netgebieden. Een voorstel hiervoor staat in het volgende hoofdstuk.

		Liander NWN	Liander	Stedin Utr	Westland	Enexis NB	Enexis L	Stedin Delfland	Stedin	Enduris	Enexis Maastr	Liander EWR	Rendo	Coteq	Stedin Mid-H	Stedin Schiedam	Stedin Z- Kenneml	Enexis Noord	
	netgebied	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
afname	E1B_AZI	4,3%	4,3%	4,3%	4,3%			4,3%	4,3%	4,3%			4,3%	4,3%	4,3%			4,3%	
	E1C_AZI		4,3%			4,3%	4,3%				4,3%	4,3%							
	E2B_AZI	4,3%	4,3%	4,3%	4,3%	4,3%	4,3%	4,3%	4,3%	4,3%	4,3%	4,3%		4,3%	4,3%			4,3%	
	E1B_AMI	4,3%	4,3%	4,3%	4,3%			4,3%	4,3%	4,3%			4,3%	4,3%	4,3%	4,3%		4,3%	
	E1C_AMI		4,3%			4,3%	4,3%				4,3%	4,3%							
	E2B_AMI	4,3%	4,3%	4,3%	2,8%	4,3%	4,3%	4,3%	4,3%	4,3%	4,3%	4,1%	4,2%	4,3%	4,3%		4,3%	4,3%	
invoed	E1B_AMI	3,1%	2,6%	2,3%	3,3%			3,1%	3,1%	2,2%			2,5%	2,4%	2,4%	1,8%		2,4%	
	E1C_AMI		2,2%			2,3%	3,0%				3,2%	2,5%							
	E2B_AMI	3,0%	3,1%	3,5%	4,3%	2,8%	3,6%	3,7%	3,4%	2,1%	3,1%	4,3%	1,7%	1,7%	2,5%		0,0%	3,3%	
totalen	totaal A	3,4%	2,9%	3,3%	3,0%	2,9%	2,8%	3,7%	3,5%	2,9%	3,4%	3,0%	3,6%	3,5%	3,4%	4,3%	4,3%	2,9%	
	totaal I	2,9%	2,4%	2,0%	3,0%	2,1%	2,8%	3,0%	2,9%	1,9%	2,8%	2,3%	2,0%	2,3%	2,2%	1,8%	0,0%	2,2%	
	totaal A + I	1,9%	1,8%	1,6%	1,3%	1,8%	1,6%	1,6%	1,6%	1,5%	1,5%	1,4%	0,8%	1,4%	1,5%	0,9%	0,6%	1,8%	
	saldo A - I	2,1%	2,2%	1,9%	1,5%	2,2%	2,0%	1,8%	1,8%	1,9%	1,6%	1,6%	1,1%	1,8%	1,9%	1,6%	0,7%	2,2%	totaal
steekproef	E1B_AZI	497	685	589	338			330	585	413			188	300	350			525	4.800
	E1C_AZI		195			602	574				434	534						2.339	
	E2B_AZI	555	581	533	113	381	189	298	469	198	259	324		298	155			308	4.661
	E1B_AMI	765	967	1.226	252			564	749	720			353	436	489	288		1.055	7.864
	E1C_AMI		1.076			1.082	872				426	768							4.224
	E2B_AMI	896	874	466	97	899	638	476	636	890	180	579	21	328	271		111	795	8.157
	totaal	2.713	4.378	2.814	800	2.964	2.273	1.668	2.439	2.221	1.299	2.205	562	1.362	1.265	288	111	2.683	32.045
populatie	E1B_AZI	1.029.375	1.012.819	455.803	46.833	0	0	127.227	957.903	142.168	0	739	23.456	42.990	78.457	31.721	7.735	697.862	4.655.088
	E1C_AZI	3.624	3.688	0	0	852.425	368.671	0	0	0	38.132	222.080	0	0	0	0	0	0	1.488.620
	E2B_AZI	42.633	56.339	20.882	2.895	49.734	22.728	5.230	38.996	7.757	1.957	9.468	1.056	1.833	3.901	1.544	363	38.646	305.962
	E1B_AMI	147.699	253.329	80.303	8.982	0	0	20.578	93.679	34.950	0	234	6.666	7.248	13.243	2.287	1.050	186.713	856.961
	E1C_AMI	665	799	0	0	175.224	102.727	0	0	0	3.793	38.646	0	0	0	0	0	0	321.854
	E2B_AMI	3.296	6.929	2.226	427	5.160	1.877	339	2.483	971	144	1.172	153	174	586	14	30	5.134	31.115
	totaal	1.227.292	1.333.903	559.214	59.137	1.082.543	496.003	153.374	1.093.061	185.846	44.026	272.339	31.331	52.245	96.187	35.566	9.178	928.355	7.659.600
	steekproef%	0,22%	0,33%	0,50%	1,35%	0,27%	0,46%	1,09%	0,22%	1,20%	2,95%	0,81%	1,79%	2,61%	1,32%	0,81%	1,21%	0,29%	0,42%

Tabel 6. Resultaten bij de huidige steekproefomvang per categorie en netgebied.

6. Steekproefgrootte gecorrigeerd voor onbetrouwbare input

In de voorgaande tabellen is met een oranje markering aangegeven voor welke categorieën en netgebieden de berekeningen te onzeker zijn, omdat in de oorspronkelijke steekproef te weinig (< 20) aansluitingen zaten. Daarnaast waren er enkele situaties geel gemarkeerd, waarin geen aansluiting in de steekproef zitten maar wel in de populatie.

Om voorlopig toch een steekproefgrootte te kunnen bepalen, is voor die situaties de steekproefgrootte ingeschat op basis van nabijgelegen netgebieden. Dat zijn netgebieden met nagenoeg dezelfde postcode en een overeenkomend aansluitingenbestand (d.w.z. beide of E1B of E1C).

Hierbij zijn de volgende koppelingen gemaakt:

- Netgebied 10 met netgebied 6
- Netgebieden 4, 6, 9, 11, 14, 15 en 16 met netgebied 8
- Netgebieden 12 en 13 met netgebied 17
- Voor netgebieden 1 en 2 was een dergelijk koppeling niet duidelijk te maken en is de steekproefgrootte voor E1C_AZI berekend naar verhouding met E1B_AZI en die voor E1C_AMI uit die van E1B_AMI.

Achtereenvolgens de volgende rekenregels toegepast:

- Bepaal de steekproefgrootte in de oranje of gele velden naar evenredigheid van de populatiegrootte; bijvoorbeeld voor E1B_AZI in netgebied 7 is de steekproefgrootte berekend uit die in netgebied 8 via de berekening: $(127227/957903) \times 585 = 0,133 \times 585 = 78$.
- Als deze berekening op een waarde < 100 uitkomt, bereken dan de steekproef als 0,5% van de populatie; deze 0,5% is ongeveer de gemiddelde relatieve steekproefgrootte die eerder is berekend (0,42%, zie Tabel 6).
- Als zelfs deze berekening een waarde < 100 geeft, wordt als minimum waarde 100 aangehouden. Dit minimum is mogelijk te laag om een RSD van 4,3% te kunnen halen, maar aan de andere kant zijn in betreffende situaties (oranje en gele velden) de populaties van het totaal aantal aansluitingen te klein om een grotere steekproef te rechtvaardigen. Voor de E2B_AMI is voor een aantal netgebieden de populatie zo klein dat overwogen kan worden deze met een veel kleinere steekproef te bemonsteren, of zelfs helemaal niet.

Het resultaat staat in onderstaande Tabel 7. Overall wordt de steekproefgrootte nu 27.784 (0,36% van het totaal aan aansluitingen).

		Lander NWN	Lander	Stedin Utr	Westland	Enexis NB	Enexis L	Stedin Delfland	Stedin	Enduris	Enexis Maastr	Lander EWR	Rendo	Coteq	Stedin Mid-H	Stedin Schiedam	Stedin Z-Kenneml	Enexis Noord	
	netgebied	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
steekproef	E1B_AZI	497	685	589	234	.	.	636	585	711	.	100	117	215	392	159	100	525	5.545
	E1C_AZI	100	100	.	.	602	574	.	.	.	191	534	.	.	.	.	.	.	2.101
	E2B_AZI	555	581	533	100	381	189	100	469	100	100	114	100	100	100	100	100	308	4.030
	E1B_AMI	765	967	1.226	100	.	.	165	749	720	.	100	100	100	106	100	100	1.055	6.352
	E1C_AMI	100	100	.	.	1.082	872	.	.	.	100	768	.	.	.	.	.	.	3.022
	E2B_AMI	896	874	466	109	899	638	100	636	249	100	300	100	100	271	100	100	795	6.733
	totaal	2.913	3.307	2.814	544	2.964	2.273	1.001	2.439	1.780	491	1.916	417	515	869	459	400	2.683	27.784
populatie	E1B_AZI	1.029.375	1.012.819	455.803	46.833	0	0	127.227	957.903	142.168	0	739	23.456	42.990	78.457	31.721	7.735	697.862	4.655.088
	E1C_AZI	3.624	3.688	0	0	852.425	368.671	0	0	0	38.132	222.080	0	0	0	0	0	0	1.488.620
	E2B_AZI	42.633	56.339	20.882	2.895	49.734	22.728	5.230	38.996	7.757	1.957	9.468	1.056	1.833	3.901	1.544	363	38.646	305.962
	E1B_AMI	147.699	253.329	80.303	8.982	0	0	20.578	93.679	34.950	0	234	6.666	7.248	13.243	2.287	1.050	186.713	856.961
	E1C_AMI	665	799	0	0	175.224	102.727	0	0	0	3.793	38.646	0	0	0	0	0	0	321.854
	E2B_AMI	3.296	6.929	2.226	427	5.160	1.877	339	2.483	971	144	1.172	153	174	586	14	30	5.134	31.115
	totaal	1.227.292	1.333.903	559.214	59.137	1.082.543	496.003	153.374	1.093.061	185.846	44.026	272.339	31.331	52.245	96.187	35.566	9.178	928.355	7.659.600
	steekproef%	0,24%	0,25%	0,50%	0,92%	0,27%	0,46%	0,65%	0,22%	0,96%	1,11%	0,70%	1,33%	0,99%	0,90%	1,29%	4,36%	0,29%	0,36%

Tabel 7. Tabel met steekproefgroottes, met invulling van onbetrouwbare velden op basis van nabijgelegen gebieden.

7. Onderzoek naar verschillen tussen profielen

7.1 Berekening van verschil tussen profielen

Onderzocht is hoe groot de verschillen zijn tussen tweetallen profielen. Hierbij is onderscheid gemaakt tussen:

- Vergelijking tussen netgebieden voor eenzelfde categorie
- Vergelijking tussen categorieën voor eenzelfde netgebied

Doelstelling van dit onderzoek is om na te gaan of er mogelijkheden zijn om de steekproeven over netgebieden, of over categorieën te combineren.

Het verschil tussen twee profielen is berekend als de wortel van het gemiddelde kwadratische verschil over een bepaalde periode.

Dus voor 2 profielen p_1 en p_2 is de afstand: $D = \sqrt{\frac{1}{K} \sum (p_{1i} - p_{2i})^2}$ waarbij wordt gesommeerd over kwartieren $i = 1 \dots K$.

De profielwaarden zijn heel kleine getallen, omdat ze over een jaar sommeren tot 1. De gemiddelde waarde van een profielwaarde per kwartier is dus $1/(366 \times 96) = 2,85 \times 10^{-5}$. Het gevolg is dat ook de afstanden D zeer kleine getallen opleveren. Voor een eenvoudiger leeswijze zijn de afstanden daarom gedeeld door die $2,85 \times 10^{-5}$, dus vermenigvuldigd met 366×96 .

Uiteindelijk is dus berekend en gerapporteerd: $D = 366 \times 96 \times \sqrt{\frac{1}{K} \sum (p_{1i} - p_{2i})^2}$

Bij de vergelijking van twee profielen zijn veel keuzes mogelijk met betrekking tot het tijdvak waarbinnen die vergelijking plaatsvindt. In overleg is besloten om de vergelijking uit te voeren over alle kwartieren van het jaar, opgedeeld in 2 tijdvakken: namelijk de wintermaanden (jan-mrt en okt-dec) en de zomermaanden (apr-sep).

7.2 Verschillen tussen netgebieden per categorie

In onderstaande tabellen zijn de onderlinge afstanden tussen netgebieden in een matrix tegen elkaar uitgezet.

De bovenste rij in elke tabel toont hoe groot de steekproef in betreffend netgebied voor betreffende categorie was. Steekproeven kleiner dan 20 zijn met een rood cijfer weergegeven. In die gevallen is een vergelijking niet erg betrouwbaar.

Alle tabellen geven met een kleurovergang aan of de afstanden relatief klein zijn (blauw) of groot (rood).

Er zijn 18 vergelijkingen mogelijk, namelijk voor de afname in 6 categorieën en de invoeding in 3 categorieën, en dan nog voor winter- en zomermaanden. Ingezoomd is op de categorie waar de grootste verschillen mogen worden verwacht, namelijk de E2B aansluitingen met invoeding.

Uit onderstaande Tabel 8 en Tabel 9 blijkt dat voor de afnameprofielen voor E2B_AMI de verschillen relatief klein zijn voor netgebieden 2, 5 en 17.

In de zomer zijn de onderlinge verschillen in afnameprofielen kleiner dan in de winter, maar het beeld van de onderlinge verschillen lijkt niet sterk te veranderen.

Opvallend is ook dat kleinere steekproeven blijkbaar leiden tot grilligere patronen, want de onderlinge verschillen zijn vooral groot bij netgebieden met kleine steekproeven.

	E2B AMI afname winter																	
	steekpr	47	117	25	3	112	56	13	39	20	2	11	2	2	3	0	2	99
	net	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Liander NWN	1		0,30	0,38	0,78	0,27	0,35	0,43	0,32	0,66	0,82	0,49	0,69	1,37	0,84		0,71	0,31
Liander	2	0,30		0,31	0,64	0,20	0,28	0,36	0,30	0,71	0,77	0,35	0,55	1,36	0,86		0,70	0,18
Stedin Utr	3	0,38	0,31		0,83	0,35	0,38	0,53	0,42	0,75	0,80	0,48	0,69	1,37	0,85		0,74	0,31
Westland	4	0,78	0,64	0,83		0,68	0,74	0,59	0,70	1,04	1,03	0,63	0,67	1,55	1,19		0,90	0,70
Enexis NB	5	0,27	0,20	0,35	0,68		0,29	0,36	0,28	0,68	0,79	0,42	0,61	1,36	0,85		0,70	0,23
Enexis L	6	0,35	0,28	0,38	0,74	0,29		0,43	0,33	0,69	0,74	0,44	0,62	1,34	0,83		0,72	0,33
Stedin Delfland	7	0,43	0,36	0,53	0,59	0,36	0,43		0,38	0,73	0,85	0,47	0,62	1,43	0,93		0,76	0,41
Stedin	8	0,32	0,30	0,42	0,70	0,28	0,33	0,38		0,67	0,77	0,43	0,63	1,36	0,83		0,69	0,35
Enduris	9	0,66	0,71	0,75	1,04	0,68	0,69	0,73	0,67		1,01	0,81	0,81	1,55	1,04		0,92	0,70
Enexis Maastr	10	0,82	0,77	0,80	1,03	0,79	0,74	0,85	0,77	1,01		0,80	0,93	1,47	1,06		0,98	0,80
Liander EWR	11	0,49	0,35	0,48	0,63	0,42	0,44	0,47	0,43	0,81	0,80		0,59	1,38	0,94		0,75	0,41
Rendo	12	0,69	0,55	0,69	0,67	0,61	0,62	0,62	0,63	0,81	0,93	0,59		1,51	1,06		0,87	0,59
Coteq	13	1,37	1,36	1,37	1,55	1,36	1,34	1,43	1,36	1,55	1,47	1,38	1,51		1,55		1,49	1,38
Stedin Mid-H	14	0,84	0,86	0,85	1,19	0,85	0,83	0,93	0,83	1,04	1,06	0,94	1,06	1,55			1,06	0,87
Stedin Schiedam	15																	
Stedin Z-Kenneml	16	0,71	0,70	0,74	0,90	0,70	0,72	0,76	0,69	0,92	0,98	0,75	0,87	1,49	1,06			0,73
Enexis Noord	17	0,31	0,18	0,31	0,70	0,23	0,33	0,41	0,35	0,70	0,80	0,41	0,59	1,38	0,87		0,73	

Tabel 8. Verschillen tussen profielen E2B AMI, afname in de winter.

	E2B AMI afname zomer																	
	steekpr	47	117	25	3	112	56	13	39	20	2	11	2	2	3	0	2	99
	net	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Liander NWN	1		0,25	0,31	0,70	0,23	0,30	0,36	0,25	0,36	0,71	0,32	0,88	0,97	0,53		0,59	0,27
Liander	2	0,25		0,22	0,66	0,19	0,25	0,37	0,26	0,40	0,67	0,32	0,80	0,97	0,55		0,56	0,17
Stedin Utr	3	0,31	0,22		0,72	0,27	0,30	0,41	0,31	0,44	0,69	0,36	0,79	0,99	0,56		0,59	0,23
Westland	4	0,70	0,66	0,72		0,63	0,73	0,70	0,67	0,82	0,95	0,73	1,07	1,15	0,90		0,81	0,71
Enexis NB	5	0,23	0,19	0,27	0,63		0,26	0,35	0,23	0,38	0,67	0,33	0,82	0,96	0,54		0,56	0,22
Enexis L	6	0,30	0,25	0,30	0,73	0,26		0,44	0,27	0,41	0,64	0,34	0,76	0,98	0,51		0,60	0,28
Stedin Delfland	7	0,36	0,37	0,41	0,70	0,35	0,44		0,38	0,45	0,80	0,46	0,94	1,02	0,64		0,65	0,37
Stedin	8	0,25	0,26	0,31	0,67	0,23	0,27	0,38		0,39	0,67	0,32	0,82	0,96	0,52		0,57	0,29
Enduris	9	0,36	0,40	0,44	0,82	0,38	0,41	0,45	0,39		0,78	0,47	0,92	1,02	0,61		0,70	0,40
Enexis Maastr	10	0,71	0,67	0,69	0,95	0,67	0,64	0,80	0,67	0,78		0,70	0,92	1,16	0,78		0,82	0,67
Liander EWR	11	0,32	0,32	0,36	0,73	0,33	0,34	0,46	0,32	0,47	0,70		0,87	0,99	0,55		0,61	0,35
Rendo	12	0,88	0,80	0,79	1,07	0,82	0,76	0,94	0,82	0,92	0,92	0,87		1,28	0,93		0,98	0,79
Coteq	13	0,97	0,97	0,99	1,15	0,96	0,98	1,02	0,96	1,02	1,16	0,99	1,28		1,09		1,08	0,98
Stedin Mid-H	14	0,53	0,55	0,56	0,90	0,54	0,51	0,64	0,52	0,61	0,78	0,55	0,93	1,09			0,75	0,56
Stedin Schiedam	15																	
Stedin Z-Kenneml	16	0,59	0,56	0,59	0,81	0,56	0,60	0,65	0,57	0,70	0,82	0,61	0,98	1,08	0,75			0,57
Enexis Noord	17	0,27	0,17	0,23	0,71	0,22	0,28	0,37	0,29	0,40	0,67	0,35	0,79	0,98	0,56		0,57	

Tabel 9. Verschillen tussen profielen E2B AMI, afname in de zomer.

Onderstaande Tabel 10 en Tabel 11 tonen de onderlinge verschillen in de invoedingsprofielen voor E2B_AMI. In de wintermaanden is de onderlinge afstand alleen tussen netgebieden 2 en 17 relatief klein. In de zomermaanden zijn alle onderlinge verschillen tussen de netgebieden groot.

	E2B AMI invoeding winter																
	steekpr	51	124	27	3	115	59	14	40	21	2	12	2	2	3	0	2
	net	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Liander NWN	1		0,35	0,32	0,51	0,48	0,62	0,56	0,38	0,45	0,74	0,50	0,87	0,92	0,52		0,76
Liander	2	0,35		0,33	0,55	0,35	0,46	0,58	0,41	0,45	0,66	0,54	0,76	0,79	0,56		0,82
Stedin Utr	3	0,32	0,33		0,52	0,38	0,53	0,54	0,36	0,43	0,69	0,53	0,86	0,89	0,52		0,78
Westland	4	0,51	0,55	0,52		0,56	0,68	0,54	0,42	0,53	0,80	0,55	1,01	0,98	0,59		0,85
Enexis NB	5	0,48	0,35	0,38	0,56		0,34	0,59	0,41	0,40	0,61	0,61	0,90	0,87	0,60		0,85
Enexis L	6	0,62	0,46	0,53	0,68	0,34		0,71	0,56	0,53	0,53	0,71	0,93	0,92	0,70		0,95
Stedin Delfland	7	0,56	0,58	0,54	0,54	0,59	0,71		0,52	0,63	0,86	0,65	1,01	1,00	0,69		0,88
Stedin	8	0,38	0,41	0,36	0,42	0,41	0,56	0,52		0,36	0,71	0,48	0,95	0,94	0,43		0,84
Enduris	9	0,45	0,45	0,43	0,53	0,40	0,53	0,63	0,36		0,69	0,59	0,95	0,98	0,53		0,85
Enexis Maastr	10	0,74	0,66	0,69	0,80	0,61	0,53	0,86	0,71	0,69		0,83	1,07	1,04	0,77		0,99
Liander EWR	11	0,50	0,54	0,53	0,55	0,61	0,71	0,65	0,48	0,59	0,83		0,98	0,96	0,55		0,96
Rendo	12	0,87	0,76	0,86	1,01	0,90	0,93	1,01	0,95	0,95	1,07	0,98		0,98	1,07		1,24
Coteq	13	0,92	0,79	0,89	0,98	0,87	0,92	1,00	0,94	0,98	1,04	0,96	0,98		1,04		1,15
Stedin Mid-H	14	0,52	0,56	0,52	0,59	0,60	0,70	0,69	0,43	0,53	0,77	0,55	1,07	1,04			0,92
Stedin Schiedam	15																
Stedin Z-Kenneml	16	0,76	0,82	0,78	0,85	0,85	0,95	0,88	0,84	0,85	0,99	0,96	1,24	1,15	0,92		0,84
Enexis Noord	17	0,44	0,28	0,44	0,63	0,49	0,58	0,67	0,54	0,55	0,74	0,63	0,66	0,75	0,66		0,84

Tabel 10. Verschillen tussen profielen E2B AMI, invoeding in de winter.

	E2B AMI invoeding zomer																
	steekpr	51	124	27	3	115	59	14	40	21	2	12	2	2	3	0	2
	net	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Liander NWN	1		0,67	0,71	1,03	0,84	1,06	1,24	0,72	0,91	1,52	0,96	1,40	1,54	1,07		1,80
Liander	2	0,67		0,62	1,11	0,58	0,80	1,29	0,78	0,87	1,41	1,08	1,19	1,32	1,19		2,03
Stedin Utr	3	0,71	0,62		1,10	0,68	0,95	1,20	0,76	0,95	1,54	1,15	1,32	1,41	1,19		2,03
Westland	4	1,03	1,11	1,10		1,14	1,30	1,29	0,95	1,09	1,74	1,19	1,66	1,80	1,32		2,12
Enexis NB	5	0,84	0,58	0,68	1,14		0,67	1,33	0,77	0,84	1,38	1,14	1,31	1,39	1,26		2,10
Enexis L	6	1,06	0,80	0,95	1,30	0,67		1,54	1,02	1,00	1,19	1,25	1,37	1,50	1,42		2,24
Stedin Delfland	7	1,24	1,29	1,20	1,29	1,33	1,54		1,26	1,44	1,97	1,55	1,86	1,84	1,63		2,15
Stedin	8	0,72	0,78	0,76	0,95	0,77	1,02	1,26		0,77	1,49	0,89	1,45	1,58	0,97		1,96
Enduris	9	0,91	0,87	0,95	1,09	0,84	1,00	1,44	0,77		1,49	1,10	1,46	1,67	1,18		2,13
Enexis Maastr	10	1,52	1,41	1,54	1,74	1,38	1,19	1,97	1,49	1,49		1,66	1,93	1,96	1,68		2,26
Liander EWR	11	0,96	1,08	1,15	1,19	1,14	1,25	1,55	0,89	1,10	1,66		1,60	1,80	1,22		2,10
Rendo	12	1,40	1,19	1,32	1,66	1,31	1,37	1,86	1,45	1,46	1,93	1,60		1,60	1,81		2,56
Coteq	13	1,54	1,32	1,41	1,80	1,39	1,50	1,84	1,58	1,67	1,96	1,80	1,60		1,93		2,47
Stedin Mid-H	14	1,07	1,19	1,19	1,32	1,26	1,42	1,63	0,97	1,18	1,68	1,22	1,81	1,93			2,00
Stedin Schiedam	15																
Stedin Z-Kenneml	16	1,80	2,03	2,03	2,12	2,10	2,24	2,15	1,96	2,13	2,26	2,10	2,56	2,47	2,00		2,17
Enexis Noord	17	0,88	0,51	0,79	1,26	0,78	0,95	1,40	1,00	1,08	1,55	1,24	1,05	1,23	1,40		2,17

Tabel 11. Verschillen tussen profielen E2B AMI, invoeding in de zomer.

Een voorlopige conclusie is dat in de meeste gevallen de onderlinge verschillen in patronen te groot zijn om de steekproeven van netgebieden samen te voegen.

Getoonde berekeningen geven een eerste verkenning. Het is aan te bevelen om

- De berekeningen wellicht te beperken tot kritische tijdvakken (bijvoorbeeld piekuren op werkdagen).
- Voor die momenten te onderzoeken wat de relatie is tussen een berekend verschil in profiel en de consequenties voor de allocatiemethodiek.
- Te bepalen welke netgebieden voor welke categorieën een voldoende klein verschil in profiel hebben.
- Deze exercitie uit te voeren voor elk van de categorieën.

7.3 Verschillen tussen categorieën per netgebied

In onderstaande tabellen staan de verschillen in profiel tussen categorieën weergegeven, met onderscheid naar het netgebied (nummer linksboven in elk van de 6 deeltabellen). Alleen de netgebieden met een steekproef groter dan 20 zijn meegenomen.

De categorieën horen 3 aan 3 bij elkaar (nl. afname bij AZI, afname in AMI en invoeding in AMI). Dit is ondersteund met dikkere gridlijnen.

Uit de tabellen blijkt dat in de wintermaanden (Tabel 12):

- De onderlinge afstand is vaak klein voor de invoedingsprofielen I4, I5 en I6 in de AMI-categorieën.
- Maar tegelijk is de afstand groot voor de afnameprofielen A4, A5 en A6 in diezelfde categorieën.
- Voor de AZI-categorieën zijn de onderlinge afstanden klein in netgebieden 5 en 17, maar groter in de overige netgebieden

In de zomermaanden (Tabel 13) is het beeld anders dan in de wintermaanden:

- De onderlinge afstand voor de invoedingsprofielen is nu juist groot (met uitzondering van netgebied 5).
- De onderlinge afstand voor de afnameprofielen is nogal wisselend over de netgebieden.

	steekpr	70	0	32	107	0	47	109	0	51
1	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1			0,36	0,40		0,48			
E1C AZI	A2									
E2B AZI	A3	0,36			0,48		0,41			
E1B AMI	A4	0,40		0,48			0,35			
E1C AMI	A5									
E2B AMI	A6	0,48		0,41	0,35					
E1B AMI	I4									0,19
E1C AMI	I5									
E2B AMI	I6							0,19		

	steekpr	0	186	55	0	280	56	0	288	59
6	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1									
E1C AZI	A2			0,29		0,47	0,54			
E2B AZI	A3		0,29			0,61	0,58			
E1B AMI	A4									
E1C AMI	A5		0,47	0,61			0,34			
E2B AMI	A6		0,54	0,58		0,34				
E1B AMI	I4									
E1C AMI	I5									0,14
E2B AMI	I6								0,14	

	steekpr	73	0	40	75	0	25	80	0	27
3	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1			0,33	0,46		0,58			
E1C AZI	A2									
E2B AZI	A3	0,33			0,49		0,58			
E1B AMI	A4	0,46		0,49			0,39			
E1C AMI	A5									
E2B AMI	A6	0,58		0,58	0,39					
E1B AMI	I4									0,22
E1C AMI	I5									
E2B AMI	I6							0,22		

	steekpr	116	0	41	85	0	39	88	0	40
8	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1			0,36	0,44		0,45			
E1C AZI	A2									
E2B AZI	A3	0,36			0,64		0,54			
E1B AMI	A4	0,44		0,64			0,35			
E1C AMI	A5									
E2B AMI	A6	0,45		0,54	0,35					
E1B AMI	I4									0,26
E1C AMI	I5									
E2B AMI	I6							0,26		

	steekpr	0	370	98	0	444	112	0	459	115
5	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1									
E1C AZI	A2			0,27		0,43	0,42			
E2B AZI	A3		0,27			0,58	0,45			
E1B AMI	A4									
E1C AMI	A5		0,43	0,58			0,30			
E2B AMI	A6		0,42	0,45		0,30				
E1B AMI	I4									
E1C AMI	I5									0,13
E2B AMI	I6								0,13	

	steekpr	71	0	62	118	0	99	124	0	105
17	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1			0,26	0,46		0,49			
E1C AZI	A2									
E2B AZI	A3	0,26			0,51		0,40			
E1B AMI	A4	0,46		0,51			0,36			
E1C AMI	A5									
E2B AMI	A6	0,49		0,40	0,36					
E1B AMI	I4									0,18
E1C AMI	I5									
E2B AMI	I6							0,18		

Tabel 12. Verschillen tussen profielen in wintermaanden.

	steekpr	70	0	32	107	0	47	109	0	51
1	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1			0,29	0,46		0,50			
E1C AZI	A2									
E2B AZI	A3	0,29			0,45		0,42			
E1B AMI	A4	0,46		0,45			0,25			
E1C AMI	A5									
E2B AMI	A6	0,50		0,42	0,25					
E1B AMI	I4									0,42
E1C AMI	I5									
E2B AMI	I6							0,42		

	steekpr	73	0	40	75	0	25	80	0	27
3	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1			0,24	0,54		0,60			
E1C AZI	A2									
E2B AZI	A3	0,24			0,47		0,52			
E1B AMI	A4	0,54		0,47			0,33			
E1C AMI	A5									
E2B AMI	A6	0,60		0,52	0,33					
E1B AMI	I4									0,44
E1C AMI	I5									
E2B AMI	I6							0,44		

	steekpr	0	370	98	0	444	112	0	459	115
5	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1									
E1C AZI	A2			0,23		0,48	0,46			
E2B AZI	A3		0,23			0,57	0,49			
E1B AMI	A4									
E1C AMI	A5		0,48	0,57			0,22			
E2B AMI	A6		0,46	0,49		0,22				
E1B AMI	I4									
E1C AMI	I5									0,23
E2B AMI	I6								0,23	

	steekpr	0	186	55	0	280	56	0	288	59
6	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1									
E1C AZI	A2			0,26		0,52	0,61			
E2B AZI	A3		0,26			0,65	0,69			
E1B AMI	A4									
E1C AMI	A5		0,52	0,65			0,23			
E2B AMI	A6		0,61	0,69		0,23				
E1B AMI	I4									
E1C AMI	I5									0,29
E2B AMI	I6								0,29	

	steekpr	116	0	41	85	0	39	88	0	40
8	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1			0,29	0,48		0,49			
E1C AZI	A2									
E2B AZI	A3	0,29			0,57		0,53			
E1B AMI	A4	0,48		0,57			0,24			
E1C AMI	A5									
E2B AMI	A6	0,49		0,53	0,24					
E1B AMI	I4									0,43
E1C AMI	I5									
E2B AMI	I6							0,43		

	steekpr	71	0	62	118	0	99	124	0	105
17	cat	A1	A2	A3	A4	A5	A6	I4	I5	I6
E1B AZI	A1			0,21	0,49		0,54			
E1C AZI	A2									
E2B AZI	A3	0,21			0,52		0,50			
E1B AMI	A4	0,49		0,52			0,27			
E1C AMI	A5									
E2B AMI	A6	0,54		0,50	0,27					
E1B AMI	I4									0,33
E1C AMI	I5									
E2B AMI	I6							0,33		

Tabel 13. Verschillen tussen profielen in zomermaanden.

De conclusie is voorlopig dat in de meeste netgebieden de profielen tussen categorieën te weinig op elkaar lijken om eventuele samenvoegingen aan te brengen. Het blijkt dat er te weinig consistentie in de verschillen tussen categorieën zit:

- De verschillen in invoeding zijn klein terwijl de verschillen in afname groot zijn, of andersom.
- De verschillen in de wintermaanden liggen anders dan die in de zomermaanden.

Getoonde berekeningen geven een eerste verkenning. Het is aan te bevelen om

- De berekeningen wellicht te beperken tot kritische tijdvakken (bijvoorbeeld piekuren op werkdagen).
- Voor die momenten te onderzoeken wat de relatie is tussen een berekend verschil in profiel en de consequenties voor de allocatiemethodiek.
- Te bepalen welke categorieën een voldoende klein verschil in profiel hebben.

Appendix.

8. Aanpak in meer detail

8.1 Data cleaning

In overleg is besloten geen statistische toetsing toe te passen op uitbijters. Het is een gegeven dat de meetreeksen per aansluiting sterk kunnen variëren, zowel in de steekproef als in het totaal van een netgebied. Er zijn alleen enkele meetreeksen uitgesloten op technische gronden.

In een aantal gevallen werden kwartierwaarden gevonden die boven de technische doorlaatwaarden liggen. Omgerekend naar kwartierwaarden zijn die grenswaarden $3 \times 25 \times 230 \times 0,95 / 4 = 4097$ Wh voor E1B en E1C, en $3 \times 80 \times 380 \times 0,95 / 4 = 21660$ Wh voor E2B. In die gevallen zijn de data van betreffend etmaal weggelaten uit de analyses. Onderstaande tabel geeft een overzicht van deze gevallen, met onderscheid naar categorie en netgebied. Het betreft met name 2 aansluitingen (geel gemarkeerd).

		E1B		E1C	
		AMI_A	AZI_A	AMI_A	AMI_I
gebied	aansluiting				
2	433		1		
	836	118			
	880			3	
	935	12			
6	321			1	
	1267				639
17	990	1			
	eindtotaal	131	1	4	639

Tabel 14. Overzicht van aantal kwartierwaarden groter dan de technische doorlaatwaarde.

Daarnaast werden enkele aansluitingen aangetroffen met een telwerk in 1000-tallen. Dit leidt tot een onrealistische reeks kwartierwaarden met veel 0-waarden en af en toe een waarde 1 als een nieuw 1000-tal wordt bereikt. Deze aansluitingen zijn eveneens weggelaten.

Hierbij ging het om 4 aansluitingen in de categorie AZI_A in netgebieden 7 (1x) en 8 (3x).

Berekening profiel per categorie per netgebied

De beschrijving van de rekenstappen is beschreven voor de afname, maar geldt op eenzelfde manier voor de invoeding.

8.1.1 Kentallen voor de aansluitingen in de steekproef

In een bepaalde categorie in een bepaald netgebied zijn meetreeksen (kwartierwaarden) beschikbaar in een steekproef van n aansluitingen.

Van deze n aansluitingen in de steekproef is berekend:

- Van de jaartotalen (SJA's):
 - De gemiddelde waarde in Wh, genoteerd als $\text{meanSJA}_{\text{sample}}$ (in Wh)

- De spreiding van die jaartotalen, genoteerd als een standaardafwijking $\text{stdSJA}_{\text{sample}}$ (in Wh)
- Voor ieder kwartier:
 - De gemiddelde waarde van de afname, genoteerd als meanA (in Wh)
 - De standaardafwijking van de afname, genoteerd als stdA (in Wh)

8.1.2 Berekening profielwaarden

Uit bovenstaande gegevens is per kwartier de profielwaarde berekend als de gemiddelde kwartierwaarde gedeeld door gemiddelde SJA (beide gemiddelden berekend over de aansluitingen in de steekproef).

In formule: $p_A = \text{meanA} / \text{meanSJA}_{\text{sample}}$.

Deze profielwaarden tellen over alle kwartieren in het jaar op tot 1.

8.1.3 Berekening onzekerheid in profielwaarden

De profielwaarde heeft een onzekerheid die kan worden beschreven met een standaardafwijking s_p . Hierbij geldt dat $s_p \approx \frac{\text{stdA} / \sqrt{n}}{\text{meanSJA}_{\text{sample}}}$.

In principe wordt de s_p dus groter evenredig met de standaardafwijking stdA tussen aansluitingen en kleiner evenredig aan de wortel van de steekproefgrootte n .

Het $\approx$ teken geeft aan dat deze uitdrukking voor s_p een benadering is, omdat hierin alleen de onzekerheid in meanA is verdisconteerd. In de uitdrukking $p_A = \text{meanA} / \text{meanSJA}_{\text{sample}}$ heeft niet alleen meanA een onzekerheid, maar ook $\text{meanSJA}_{\text{sample}}$. Bovendien zijn die onzekerheden positief gecorreleerd: een aansluiting in de steekproef met een hogere kwartierwaarde dan de rest heeft meestal ook een hogere SJA. Het effect hiervan is meegenomen in de berekeningen.

8.2 Berekening kwartierwaarden totaal per categorie per netgebied, huidige steekproefgrootte

De profielwaarden p_A – en de onzekerheid in de vorm van de standaardafwijking s_p – zijn vervolgens omgerekend naar alle aansluitingen in betreffende categorie en netgebied.

8.2.1 Berekening totale afname per categorie en netgebied

De totale afname per kwartier is gelijk aan de profielwaarde maal de SJA, en dan gesommeerd over alle aansluitingen in het netgebied (populatie).

Noem N het totaal aantal aansluitingen in een netgebied (binnen een categorie) en de gemiddelde SJA over die aansluitingen $\text{meanSJA}_{\text{pop}}$.

De totale afname in betreffende categorie en netgebied is dan (per kwartier):

$$A = \sum p_A \times SJA_i = p_A \times \sum SJA_i = p_A \times N \times \text{meanSJA}_{\text{pop}}$$

De factor F wordt gedefinieerd als de verhouding van de gemiddelde SJA in de populatie en de steekproef, dus $F = \text{meanSJA}_{\text{pop}} / \text{meanSJA}_{\text{sample}}$.

Daarmee kan de afname worden herschreven als:

$$A = p_A \times N \times F \times \text{meanSJA}_{\text{sample}}$$

Het voordeel van deze schrijfwijze is dat het totale afnamepatroon altijd berekend kan worden uit de data, ook in situaties waarin de totale SJA (of SJI) in de populatie niet bekend is. Dit is het geval voor de SJI en ook voor de SJA is er een onzekerheid in de correctheid van de data uit CA-R (zie 2.4) in verband met mogelijke saldering in de data.

De aanpak is dan:

- Als de SJA of SJI in de populatie wel betrouwbaar bekend is wordt de factor F berekend als $\text{meanSJA}_{\text{pop}} / \text{meanSJA}_{\text{sample}}$. Het effect is dan als volgt: als de steekproef een representatief beeld geeft van de SJA's dan ligt de waarde van de factor F dichtbij 1, en als de steekproef een relatief lage (c.q. hoge) schatting van de SJA geeft, wordt F groter (c.q. kleiner) dan 1 en de berekende afname hiervoor dus gecorrigeerd.
- Als de SJA of SJI niet of niet betrouwbaar bekend is, is er geen andere oplossing dan aan te nemen dat de steekproef representatief is en wordt gerekend met een factor $F = 1$.

8.2.2 Berekening onzekerheid in de totale afname per categorie per netgebied

In 8.1.3 is aangegeven dat de onzekerheid in het profiel p_A wordt beschreven door de

standaardafwijking s_p , waarbij $s_p \approx \frac{\text{std}A / \sqrt{n}}{\text{meanSJA}_{\text{sample}}}$.

De onzekerheid in de totale afname $A = p_A \times N \times F \times \text{meanSJA}_{\text{sample}}$ wordt dan dus beschreven door een standaardafwijking s_A die een factor $N \times F \times \text{meanSJA}_{\text{sample}}$ maal die s_p is.

Dus $s_A = N \times F \times \frac{\text{std}A}{\sqrt{n}}$

8.2.3 Samengevat: totale afname en onzekerheid per categorie per netgebied

Per categorie wordt (per kwartier) de totale afname over alle aansluitingen in het netgebied berekend als:

$$A = p_A \times N \times F \times \text{meanSJA}_{\text{sample}}$$

De onzekerheid op die berekende waarde wordt berekend als de standaardafwijking

$$s_A = N \times F \times \frac{\text{std}A}{\sqrt{n}}$$

De relatieve standaardafwijking wordt dan berekend als

$$RSD = \frac{s_A}{A} = \frac{1}{p_A \times \text{meanSJA}_{\text{sample}}} \times \frac{\text{std}A}{\sqrt{n}} \times 100 \text{ (\%)}$$

De berekeningen van het totaal, de standaardafwijking en de relatieve standaardafwijking gebeurt afzonderlijk voor zowel afnames als eventuele invoedingen per categorie. Dus in termen van

hoofdstuk 3 afzonderlijk voor alle totalen A1, A2, A3, A4, A5, A6, I4, I5 en I6 voor zover van toepassing in betreffend netgebied.

8.2.4 Volumegewogen RSD per categorie bij huidige steekproefomvang

Bovenstaande berekeningen leveren voor ieder kwartier een relatieve standaardafwijking RSD voor de totale afname of invoeding in betreffende categorie. Omdat de spreiding stdA tussen de aansluitingen in de steekproef in elk kwartier anders kan zijn, is voor het overall beeld een volumegewogen RSD_{tot} berekend.

Dus $RSD_{tot} = \frac{\sum(A \times RSD)}{\sum A}$, waarbij de sommatie Σ loopt over alle kwartieren in een periode.

Ter illustratie zijn voor een willekeurig netgebied de RSD's – volumegewogen per dag – weergegeven in onderstaande figuren. De resultaten gelden voor de huidige steekproefgroottes per categorie. Dit betrof:

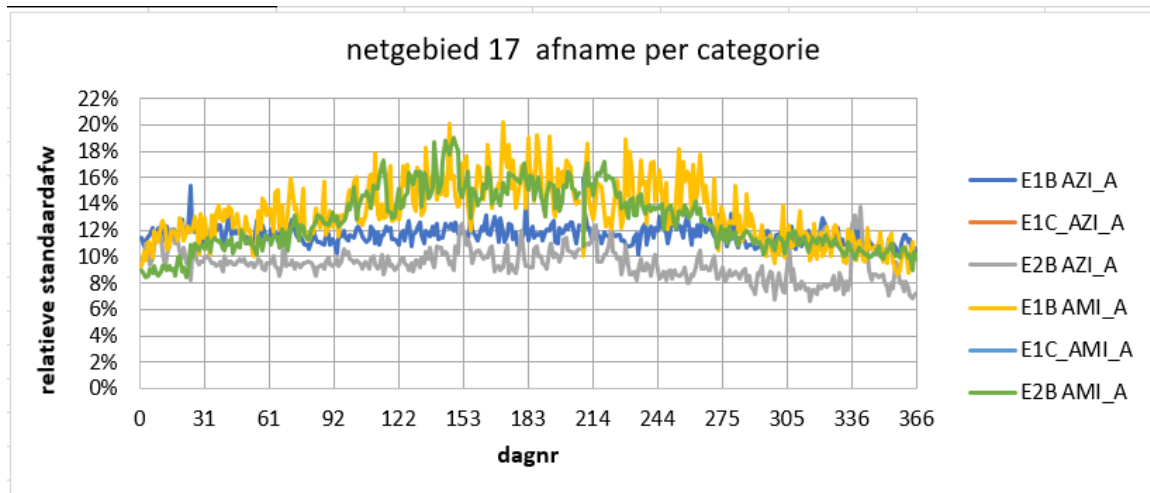
- n = 75 voor E1B zonder invoeding, dus afname E1B_AZI_A (=A1)
- n = 71 voor E2B zonder invoeding, dus afname E2B_AZI_A (=A3)
- n = 124 voor E1B met invoeding, dus afname E1B_AMI_A (=A4), invoeding E1B_AMI_I (=I4)
- n = 105 voor E2B met invoeding, dus afname E2B_AMI_A (=A6), invoeding E2B_AMI_I (=I6)

Voor dit netgebied zijn er dus geen aansluitingen categorie E1C. In de grafieken ontbreken dus lijnen voor E1C_AZI_A (=A2), E1C_AMI_A (=A5) en E1C_AMI_I (=I5).

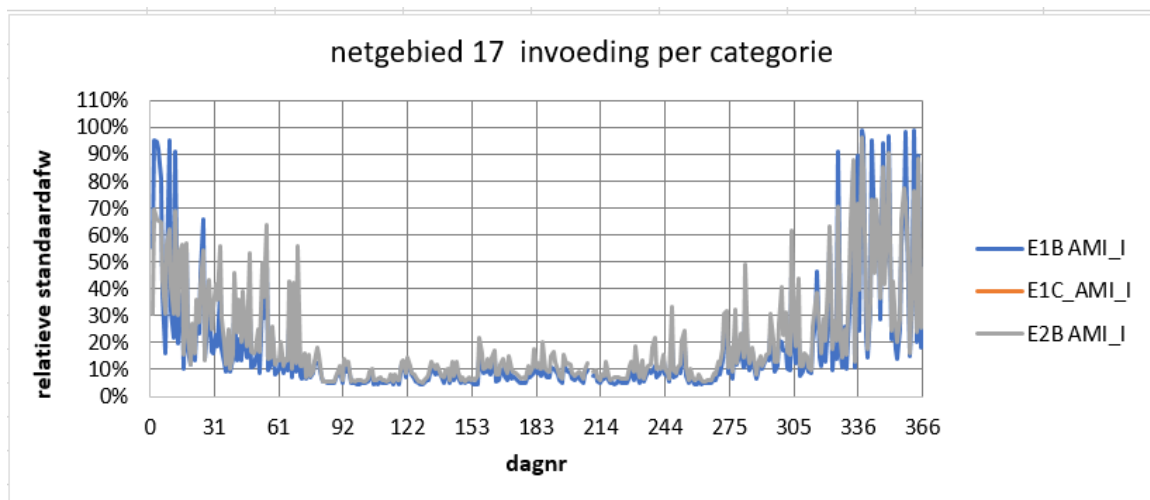
De grafieken tonen een fluctuerend beeld over het jaar heen (horizontale as):

- Voor de afnames zonder invoeding E1B_AZI_A en E2B_AZI_A is er voor de RSD nauwelijks sprake van een effect van seizoenen. Over het hele jaar is de volumegewogen RSD respectievelijk 11,6% en 9,4%
- Voor de afnames met invoeding E1B_AMI_A en E2B_AMI_A is de RSD in de zomermaanden hoger. Maar in die maanden is het volume in deze categorieën lager. Over het hele jaar is de volumegewogen RSD respectievelijk 12,9% en 12,2%.
- Voor de invoedingen E1B_AMI_I en E2B_AMI_I lijkt het beeld verontrusten met hoge RSD's in de wintermaanden. Dit is echter een direct gevolg van het feit dat in die periode nauwelijks invoeding plaatsvindt. Overall over het hele jaar is de RSD respectievelijk 7,0% en 9,2%.

De genoemde RSD's laten zien dat bij de huidige steekproef (in dit voorbeeld in netgebied 17) de eis van 4,3% niet wordt gehaald.



Figuur 1. Relatieve standaardafwijking voor de totale afname per categorie, volume gewogen per dag. Netgebied 17, huidige steekproef.



Figuur 2. Relatieve standaardafwijking voor de totale invoeding per categorie, volume gewogen per dag. Netgebied 17, huidige steekproef.

8.3 Benodigde steekproefgrootte per categorie voor een RSD van 4,3%

Bij de huidige steekproefgrootte wordt in vrijwel alle categorieën en netgebieden de vereiste nauwkeurigheid niet gehaald. Deze eis is een RSD van 4,3%. Dit komt overeen met een marge van +/- 10% bij 98% betrouwbaarheid.

In paragraaf 8.1.3 is afgeleid dat de relatieve standaardafwijking wordt berekend als:

$$RSD = \frac{s_A}{A} = \frac{1}{p_A \times \text{meanSJA}_{\text{sample}}} \times \frac{\text{stdA}}{\sqrt{n}} \times 100 \text{ (%)}$$

Hieruit blijkt dat de RSD – en daarmee ook de volume gewogen RSD – een relatie $1/\sqrt{n}$ heeft met de steekproefgrootte n . Dus om bijvoorbeeld de RSD te halveren moet de steekproef 4 maal zo groot worden. Meer algemeen geldt dat de optimale steekproefgrootte n_{opt} kan worden berekend als:

$$n_{opt} = n \times \left(\frac{RSD}{4,3} \right)^2, \text{ waarbij } n \text{ en } RSD \text{ de waarden in de huidige steekproef voorstellen.}$$

Als voorbeeld is hieronder voor netgebied 17 weergegeven hoe de berekeningen verlopen. Bij de afnames in de verschillende categorieën was de RSD in de oorspronkelijke situatie tussen 9,4% en 12,9%. Dit is een factor 2,5 tot 3 maal de vereiste waarde van 4,3%. De steekproeven moeten dan dus via het kwadratische verband ongeveer 6 tot 9 maal zo groot worden gekozen om op de 4,3% uit te komen.

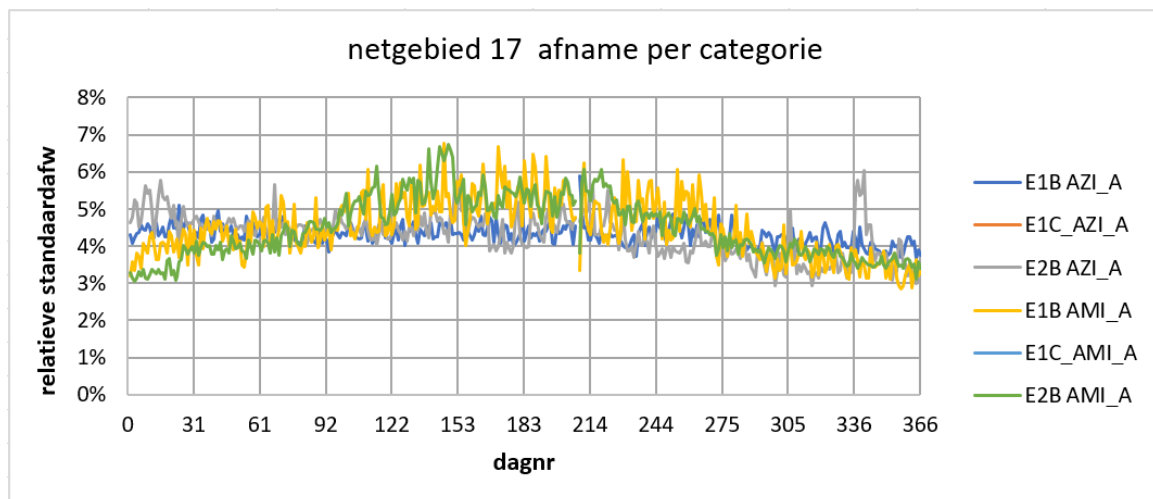
Het resultaat van deze berekeningen staat in Tabel 15. Te zien is dat voor elk van de categorieën de RSD op 4,3% staat en welke steekproefgrootte daarvoor nodig is.

Voor de invoeding is het resultaat wat beter dan de 4,3%. Dit komt omdat de steekproefgrootte voor de AMI's noodzakelijk op 1.055 en 795 moest komen in verband met de spreiding in afnames.

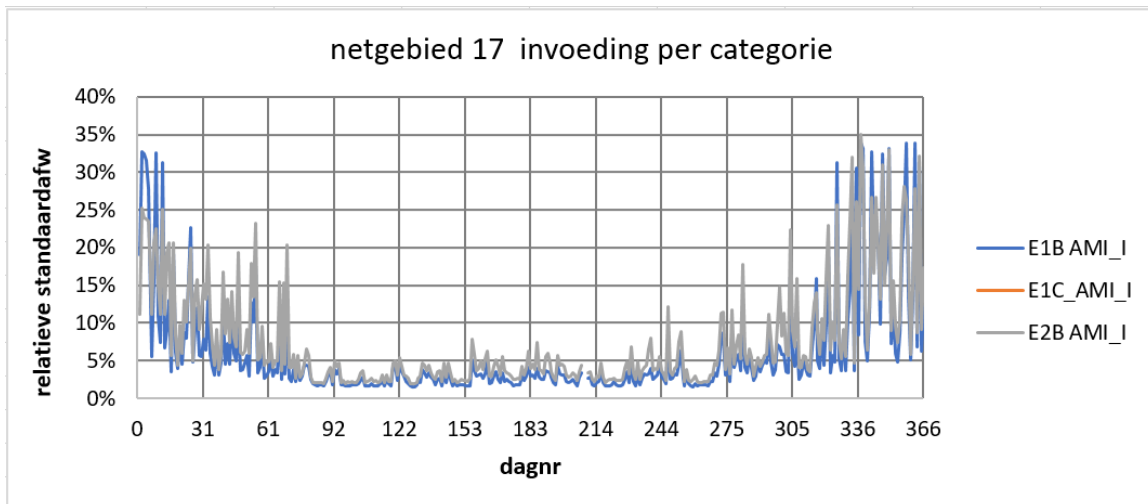
17	volume gewogen RSD		steekproef	
	afname	invoeding	afname	invoeding
E1B_AZI	4,3%		525	
E1C_AZI			0	
E2B_AZI	4,3%		308	
E1B_AMI	4,3%	2,4%	1.055	1.055
E1C_AMI			0	0
E2B_AMI	4,3%	3,3%	795	795
		totaal	2.683	

Tabel 15. Vereiste steekproefgroottes bij RSD = 4,3%. Netgebied 17.

Het verloop van de RSD, volume gewogen per dag, staat in onderstaande figuren. Ten opzichte van de oorspronkelijke steekproef (zie bovenstaande Figuur 1 en Figuur 2) liggen de waarden nu lager, want rond de 4,3%. Bovendien zijn de fluctuaties in RSD van dag tot dag kleiner.



Figuur 3. Relatieve standaardafwijking voor de totale afname per categorie, volume gewogen per dag. Netgebied 17, optimale steekproef.



Figuur 4. Relatieve standaardafwijking voor de totale afname per categorie, volume gewogen per dag. Netgebied 17, optimale steekproef.

8.4 Berekening totalen over categorieën per netgebied

Voor de verschillende categorieën zijn de volgende waarden voor de totale afname en invoeding bekend:

- E1B zonder invoeding: A1 = afname E1B_AZI_A
- E1C zonder invoeding: A2 = afname E1C_AZI_A
- E2B zonder invoeding: A3 = afname E2B_AZI_A
- E1B met invoeding: A4 = afname E1B_AMI_A en I4 = invoeding E1B_AMI_I
- E1C met invoeding: A5 = afname E1C_AMI_A en I5 = invoeding E1C_AMI_I
- E2B met invoeding: A6 = afname E2B_AMI_A en I6 = invoeding E2B_AMI_I

Voor elk van deze waarden A1 ... A6 en I4 ... I6 wordt het totaal over alle aansluitingen in het netgebied berekend volgens de formules van paragraaf 8.1.3.

Dus voor het totaal:

$$A1 = p_{A1} \times N \times F \times meanSJA1_{sample}$$

En voor de standaardafwijking op dit totaal

$$s_{A1} = N \times F \times \frac{stdA1}{\sqrt{n}}$$

Hierin is p_{A1} het uit de steekproef geschatte profiel, N het totaal aantal aansluitingen in de categorie, n de steekproefgrootte F de verhouding in gemiddelde SJA (of SJI) in de populatie en de steekproef en stdA1 de standaardafwijking voor spreiding in de steekproef.

Waar A1 staat kan voor de andere totalen ook A2...A6 en I4...I6 worden ingevuld.

Hieruit kunnen nu eenvoudig de totalen worden berekend over de categorieën:

- Totale afname $A = A1 + A2 + A3 + A4 + A5 + A6$
- Totale invoeding $I = I4 + I5 + I6$
- Totale netuitwisseling $T = A + I$
- Totaal saldo $S = A - I$

Bovendien kunnen de standaardafwijkingen voor die totalen worden berekend. Voor A en I geldt dat de standaardafwijking wordt berekend als de wortel van de kwadratische optelling van de s_{A1} t/m s_{A6} of s_{I4} t/m s_{I6} . Voor T en S geldt iets soortgelijks, maar hier moet rekening worden gehouden met het feit dat de opgetelde delen niet allemaal statistisch onafhankelijk zijn. Er geldt namelijk een

negatieve correlatie tussen A4 en I4, tussen A5 en I5 en tussen A6 en I6. Immers, door de saldering zal een grotere invoeding vaak samenvallen met een lagere afname en omgekeerd.

Op dezelfde wijze als voor de waarden per categorie kan ook voor de totalen een relatieve standaardafwijking RSD worden berekend per kwartier. En vervolgens ook volume gewogen gemiddeld over een tijdsperiode van een dag of een jaar.

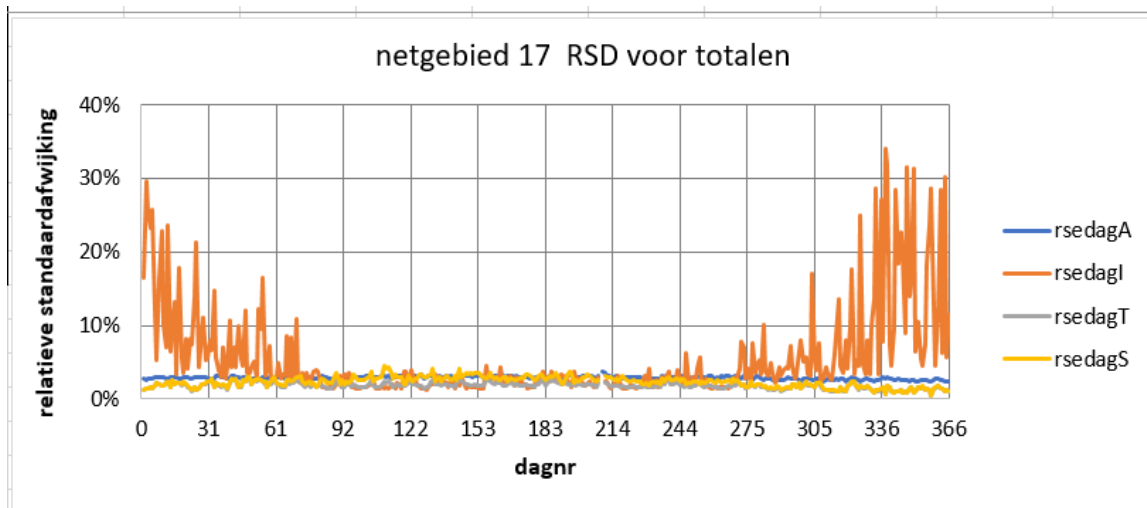
Als we het voorbeeld van netgebied 17 voortzetten, blijkt dat voor de totalen het jaargewogen gemiddelde voor de RSD waarden worden gevonden zoals vermeld in onderstaande tabel. De waarden gelden voor de geoptimaliseerde steekproefgroottes, waarbij voor elke categorie afzonderlijk de jaargewogen RSD 4,3% bedraagt.

Te zien is dat voor de totalen de jaargewogen RSD's lager liggen dan die 4,3%. Dit komt omdat er bij de optelling altijd een vorm van uitmiddeling van onzekerheden optreedt.

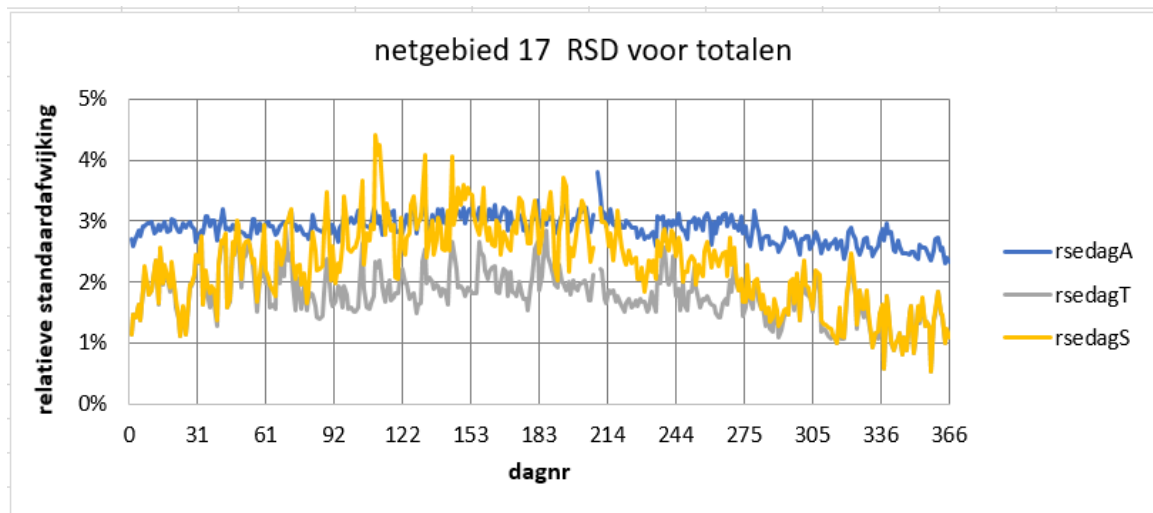
totaal A	2,9%
totaal I	2,2%
totaal A + I	1,8%
saldo A - I	2,2%

Tabel 16. Jaargewogen gemiddelde RSD voor netgebied 17, optimale steekproef.

Onderstaande figuren tonen het verloop van de volume gewogen RSD per dag. In Figuur 5 wordt het beeld bepaald door de totale invoeding I. Deze heeft in de wintermaanden een grote RSD. In Figuur 6 is deze I weggelaten. De verticale schaalverdeling laat nu beter het verloop van A, T en S zien.



Figuur 5. Relatieve standaardafwijking voor de totalen A, I, T en S over categorieën, volume gewogen per dag. Netgebied 17, optimale steekproef.



Figuur 6. Relatieve standaardafwijking voor de totalen A, T en S over categorieën, volume gewogen per dag. Netgebied 17, optimale steekproef.

CONSULTANTS IN QUANTITATIVE METHODS

Tallose variabelen. Tallose x'jes. Zijn ze incidenteel of structureel? Wat is hun samenhang, hun invloed, wat is echt van belang en wat niet?

CQM helpt organisaties om complexe processen inzichtelijk te maken. Met kwantitatieve modellen creëren wij de structuur om fact based beslissingen te nemen en processen te analyseren. Zo optimaliseren wij planning en logistiek en verbeteren wij product- en procesinnovatie. Intelligentie, die organisaties structureel op een hoger niveau brengt. CQM analyseert en lost op, met een groot inlevingsvermogen in uw specifieke problematiek.

[Zo werken wij.](#) [Van X naar U.](#)

CQM B.V.

T +31 40 750 23 23

E info@cqm.nl

I www.cqm.nl

Vonderweg 16
5616 RM Eindhoven
P.O. Box 414
5600 AK Eindhoven
The Netherlands

Trade register 17076484
IBAN NL61RABO0359340598
VAT NL801228505B01