

Robuustheid regressiemodel voor kapitaalkosten gebaseerd op aansluitdichtheid

Dr.ir. P.W. Heijnen
Faculteit Techniek, Bestuur en Management
Technische Universiteit Delft

22 april 2010

1 Introductie

De Energiekamer heeft in 2009 onderzoek laten doen door Cosentec naar de relatie tussen kabel- en lijngerelateerde *kapitaalkosten* en de *aansluitdichtheid* in het Nederlandse elektriciteits- en gasnetwerk. Uit dit onderzoek is voor het elektriciteitsnetwerk een verband geschat tussen deze *kapitaalkosten* en de *aansluitdichtheid* binnen de gebieden van de 9 Nederlandse netbeheerders, m.b.v. gegevens over het jaar 2006. Met het gevonden verband wil de Energiekamer voortaan per netbeheerder een schatting maken van de verwachte *kapitaalkosten* per jaar. Het verschil tussen de geschatte *kapitaalkosten* en de werkelijke *kapitaalkosten* zal daarna worden verrekend met de individuele netbeheerders. De Energiekamer heeft besloten om in 2010 het regressiemodel opnieuw te bepalen m.b.v. de meest recente data uit 2009 en het resulterende model voor 3 jaar vast te leggen en te gebruiken.

De methode die gebruikt is om de relatie tussen de *kapitaalkosten* en de *aansluitdichtheid* te bepalen, lijkt vrij arbitrair. Er is een aantal aannamen gedaan om de bovengenoemde variabelen te construeren. Om de betrouwbaarheid van de resultaten te beoordelen moet onderzocht worden in hoeverre de resultaten gevoelig zijn voor deze aannamen.

Bovendien is het regressiemodel uit 2006 slechts gebaseerd op 9 meetpunten en het nieuwe regressiemodel in 2009 zal zelfs maar gebaseerd zijn op 8 meetpunten. Ieder meetpunt komt overeen met één van de Nederlandse netbeheerders. Voor een betrouwbare regressieanalyse is dit een zeer klein aantal. De invloed van individuele meetpunten kan daardoor zeer groot zijn. Ook dit moet nader bekeken worden.

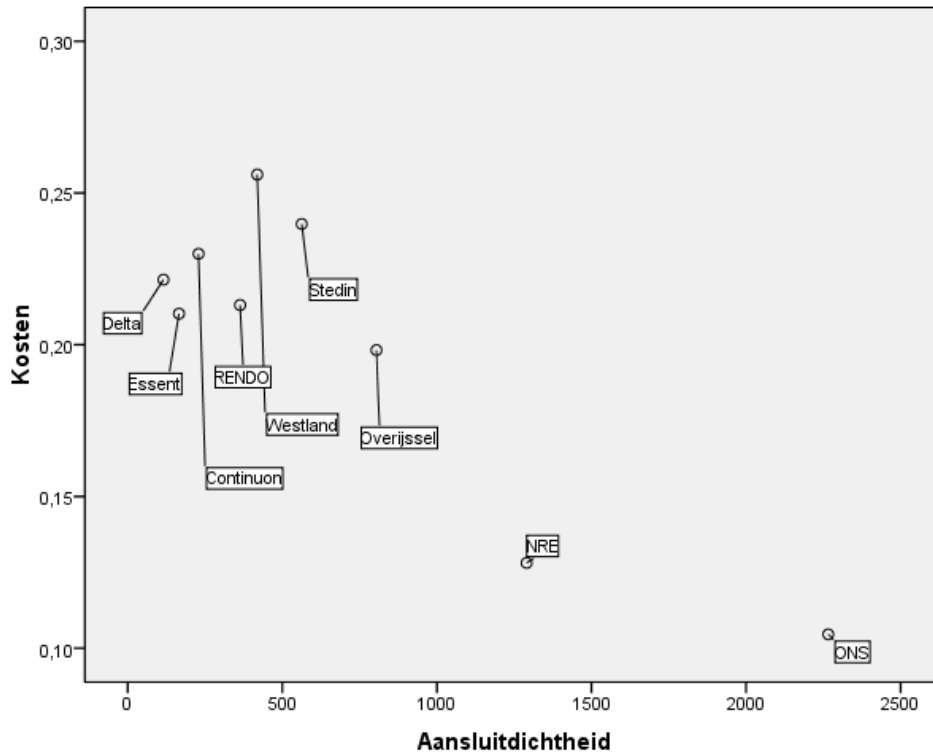
Stedin is van oordeel dat beide punten terdege moeten worden onderzocht voordat kan worden geconcludeerd dat *aansluitdichtheid* een objectief regionaal verschil vormt. In dit rapport is in dit verband nader onderzoek uitgevoerd naar het tweede punt: de invloed van individuele meetpunten.

2 Regressiemodel voor kapitaalkosten op basis van aansluitdichtheid

Een simpel regressiemodel, zoals door de Energiekamer voorgesteld, geeft een model van het lineaire verband tussen een afhankelijke en een onafhankelijke variabele. Het regressiemodel kan voor twee doeleinden gebruikt worden. Ten eerste om m.b.v. het regressiemodel een schatting te maken van de verwachte waarde van de afhankelijke variabele en ten tweede om de bekende waarde van een afhankelijke variabele te verklaren d.m.v. de waarde van een onafhankelijke variabele.

Het regressiemodel dat de Energiekamer wil gaan gebruiken geeft het verband weer tussen de afhankelijke variabele *kapitaalkosten* en de onafhankelijke variabele *aansluitdichtheid*. Op basis van de *aansluitdichtheid*, zoals deze is bepaald voor iedere netbeheerder afzonderlijk, wil men m.b.v. het regressiemodel een schatting geven voor de *kapitaalkosten* van de netbeheerder. De werkelijke *kapitaalkosten* van de netbeheerder zullen aan deze schatting worden gerelateerd en afwijkingen hiervan zullen in positieve of negatieve zin worden verrekend. Om ervoor te zorgen dat het totaal aan geschatte waarden gelijk is aan het totaal van de daadwerkelijke *kapitaalkosten*, worden de te verrekenen bedragen geschaald met de quotiënt van de totale werkelijke kosten en de totale geschatte kosten.

Voor de netbeheerders, en in het bijzonder voor die netbeheerders waarvoor de schatting lager uitvalt dan hun werkelijke *kapitaalkosten*, is het dus van groot belang te weten dat er een duidelijk verband bestaat tussen de *aansluitdichtheid* en de *kapitaalkosten*, dat de schatting gemaakt met het regressiemodel betrouwbaar is en dat het daadwerkelijk een goede indicatie geeft van de *kapitaalkosten* die bij de voor hen specifieke *aansluitdichtheid* haalbaar zouden moeten zijn. Dit rapport zal zich dan ook voornamelijk richten op deze aspecten.

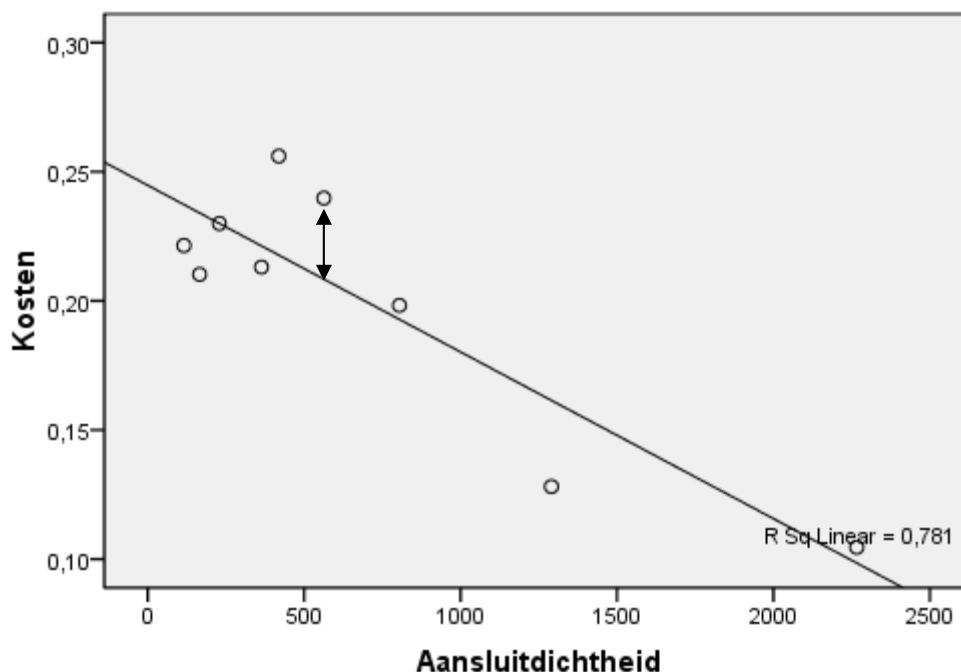


Figuur 1: Spreidingsdiagram van de *kapitaalkosten* versus de *aansluitdichtheid*

Figuur 1 toont de gebruikte meetpunten. Het onderzoek is gedaan op data uit het jaar 2006 [p258, ontwerp-methodebesluit]. In 2006 waren er 9 netbeheerders. Van alle 9 netbeheerders zijn de data bekend verondersteld.

De correlatiecoëfficiënt van de *kapitaalkosten* en de *aansluitdichtheid* bedraagt -0.884. Dit duidt op een sterke negatieve lineaire relatie tussen de twee variabelen. Hiervan is dan ook gebruik gemaakt voor het schatten van het regressiemodel. Een belangrijke kanttekening hierbij is echter dat de correlatiecoëfficiënt in sterke mate beïnvloed kan zijn door individuele meetpunten die afwijken van de rest van de meetdata. Hier wordt later op teruggekomen.

Het regressiemodel is bepaald met een *ordinary least squares* methode [p258, ontwerp-methodebesluit]. Dit is de standaardmethode voor het bepalen van een regressiemodel. Iedere netbeheerder heeft hetzelfde gewicht gekregen [p266, ontwerp-methodebesluit] en er wordt een lineair verband tussen de *kapitaalkosten* en de *aansluitdichtheid* bepaald zodanig dat het verschil tussen de meetpunten en de regressielijn zo klein mogelijk is.



Figuur 2: Spreidingsdiagram en regressielijn

De Energiekamer gebruikt het Schwartz Bayesian criterium om een keuze te maken tussen een lineair, kwadratisch of loglineair verband [p260, ontwerp-methodebesluit]. Op basis van dit criterium is gekozen voor een lineair verband. Figuur 2 toont de meetpunten en de lineaire regressielijn zoals hij bepaald is door Energiekamer voor de data over 2006 [Formule 35, bijlage 1, ontwerp-methodebesluit].

Voor de accuraatheid van een regressiemodel wordt over het algemeen de determinatiecoëfficiënt, ook wel de R^2 gebruikt. Dit is een maat die aangeeft welk percentage van de variantie van de afhankelijke variabele wordt verklaard door de onafhankelijke variabele. Voor een simpel regressiemodel met slechts één onafhankelijke variabele is de R^2 gelijk aan het kwadraat van de correlatiecoëfficiënt. In dit geval, verklaart het regressiemodel op basis van de *aansluitdichtheid* 78.1% van de variantie gevonden in de *kapitaalkosten*.

De Energiekamer concludeert hieruit, dat de netbeheerders met een meetpunt boven de regressielijn te hoge *kapitaalkosten* hebben gegeven hun *aansluitdichtheid*.

Voordat dit echter geconcludeerd mag worden, moet er eerst gedegen worden aangetoond dat de *aansluitdichtheid* inderdaad een goede voorspeller is voor de *kapitaalkosten* die verwacht kunnen worden.

In het onderzoek van Cosentec lijkt met onderstaande zaken echter onvoldoende tot geen rekening gehouden te zijn:

- De dataset is compleet. Dat betekent dat er geen sprake is van een steekproeffout of van onzekerheid in de verkregen relaties. De significantiewaarde die volgt uit de statistische analyses kan niet gebruikt worden om aan te tonen dat de resultaten praktische relevantie hebben. Hiervoor is een ander criterium nodig.

- De dataset is erg klein, waardoor de invloed van individuele netbeheerders onaanvaardbaar groot kan worden en het verdwijnen of ontstaan van een nieuwe netbeheerder het regressiemodel structureel kan veranderen.
- De dataset is erg klein, waardoor niet is aan te tonen dat aan een aantal belangrijke regressievoorwaarden is voldaan. De gebruikte regressiemethode is echter niet geldig als niet aan deze voorwaarden is voldaan. In dat geval kan het regressiemodel niet worden gebruikt om betrouwbare schattingen te maken.
- De keuze om alle netbeheerders hetzelfde gewicht te geven in de regressieanalyse kan erg bepalend zijn voor het uiteindelijke resultaat. Het zorgt ervoor dat een zeer kleine netbeheerder dezelfde invloed op het regressiemodel heeft als een hele grote netbeheerder. De keus voor een gewogen regressiemodel kan de uitkomsten structureel veranderen.

Bovengenoemde punten zullen hieronder één voor één besproken worden.

3 Onderzoek op de complete populatie van netbeheerders

In het onderzoek zijn de gegevens van alle 9 netbeheerders bekend verondersteld. Er is hier dus geen sprake van een steekproef, maar van een onderzoek op de totale populatie. Dat betekent dat er geen sprake is van een steekproeffout of van een statistische significantie van bijvoorbeeld een correlatiecoëfficiënt of een regressiecoëfficiënt. Bij een steekproefonderzoek geeft de statistische significantie de kans aan om deze waarde (of groter) van de correlatiecoëfficiënt in de steekproef te vinden, terwijl hij in werkelijkheid voor de gehele populatie gelijk is aan 0. Als deze kans heel klein is, kleiner dan 5% in het algemeen, wordt ervan uitgegaan dat de correlatiecoëfficiënt significant is, wat betekent dat dezelfde *relatie* ook voor de gehele populatie zal gelden, maar dit betekent niet dat exact dezelfde *waarde* van de correlatiecoëfficiënt ook voor de populatie geldt.

In het geval van een onderzoek op de complete populatie kent de waarde van een correlatiecoëfficiënt of regressiecoëfficiënt m.b.t. de gebruikte data geen onzekerheid. Wel is er sprake van onzekerheid als het er over gaat of de resultaten kunnen worden toegepast om de *kapitaalkosten* voor komende jaren te schatten op basis van de *aansluitdichtheid*. Hiernaar wordt echter geen onderzoek gedaan.

In het ontwerp-methodebesluit wordt meerdere malen de statistische significantie van resultaten gebruikt, om de praktische relevantie van de resultaten aan te tonen [p258, p260, ontwerp-methodebesluit]. Dit is niet terecht¹. Als het om het verband gaat tussen de *aansluitdichtheid* en de *kapitaalkosten* is alleen de hoogte van de correlatiecoëfficiënt en het daarmee samenhangend percentage verklaarde variantie (de R^2) van belang. De Energiekamer moet zelf een eenduidige ondergrens vastleggen voor de hoogte van de correlatiecoëfficiënt die zij nog relevant acht. Dit kan niet gebeuren op basis van de statistische significantie.

4 Invloed van individuele netbeheerders op het model

Zoals in Figuur 1 en Figuur 2 duidelijk te zien is, wordt de regressielijn in sterke mate bepaald door de netbeheerder ONS en in iets mindere mate door de netbeheerder NRE. Zonder deze twee netbeheerders zou er nauwelijks sprake zijn van een lineair verband tussen de *kapitaalkosten* en de

¹ In het ontwerp-methodebesluit wordt zelfs gesproken over “als de t-waarde niet hoger is dan 5%” [p258], hier wordt waarschijnlijk bedoeld, “als de significantie van de t-waarde niet hoger is dan 5%”.

aansluitdichtheid. De correlatiecoëfficiënt bedraagt in dat geval nog slechts -0.137 en een regressiemodel zonder deze 2 netbeheerders zal nog slechts 1.9% van de variantie kunnen verklaren van de verschillen in *kapitaalkosten* zoals gevonden voor de verschillende netbeheerders. Hieraan kan geen enkele praktische relevantie meer worden toegekend.

Hieronder zal m.b.v. een aantal statistische parameters worden aangetoond dat de invloed van de netbeheerders ONS en NRE op de regressieresultaten inderdaad erg groot is.

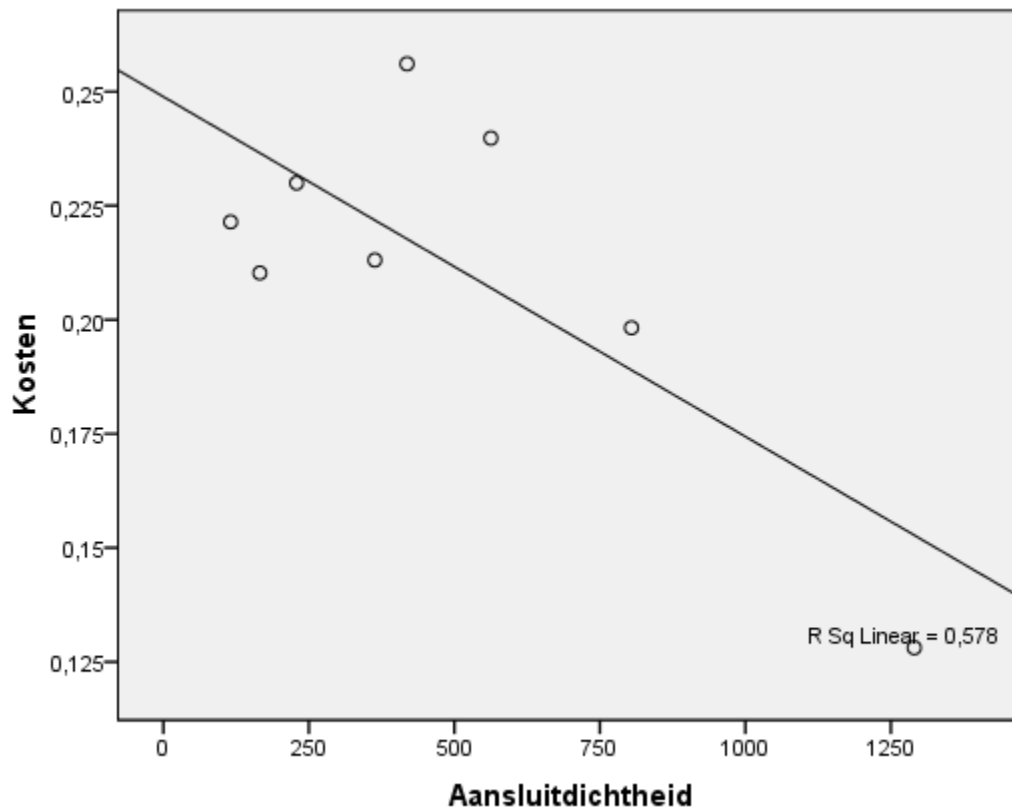
De invloed van de individuele meetpunten op een regressiemodel kan op een aantal manieren worden gemeten:

- **Grootte van het residu:** Dit is het verschil tussen gemeten waarde en geschatte waarde. Meetpunten met een groot residu vertonen afwijkend gedrag van de andere meetpunten en hun waarde wordt door het model niet goed verklaard. De residuen worden over het algemeen gestandaardiseerd, zodat ze een gemiddelde van 0 en een variantie van 1 hebben. Een vuistregel is dat residuen groter dan 3 of kleiner dan -3 verwijderd zouden moeten worden uit de analyse.
- **Cook's distance:** Dit is een maat voor de verandering van de regressiecoëfficiënt als het meetpunt uit de analyse wordt verwijderd. De Cook's distance moet laag zijn (dichtbij 0), anders wordt het regressiemodel in hoge mate bepaald door dat ene meetpunt, wat voor de betrouwbaarheid van het model niet gewenst is. Als vuistregel geldt dat meetpunten boven $\frac{4}{(n-k-1)}$ verwijderd moeten worden. Hierin is n de grootte van de dataset en k het aantal onafhankelijke variabelen in het regressiemodel. In dit geval zouden meetpunten met een Cook's distance groter dan 0.57 ($n = 9$ en $k = 1$) verwijderd moeten worden.
- **Leverage value:** Dit is een maat voor de afstand van het individuele meetpunt tot alle andere meetpunten. Ook de Leverage value moet laag zijn (dichtbij 0), anders vertoont het meetpunt teveel afwijkend gedrag van de andere meetpunten en kan daardoor het regressiemodel teveel beïnvloeden. Als vuistregel geldt dat meetpunten met een Leverage value boven de $\frac{2(k+1)}{n}$ teveel invloed hebben en verwijderd moeten worden. In dit geval zouden meetpunten met een Leverage value groter dan 0.44 verwijderd moeten worden.

Netbeheerder	Residu (<3)	Cook's distance (<0.57)	Leverage value (<0.44)
Continuon	0,0000	0,0000	0,0550
Overijssel	0,2130	0,0033	0,0033
Delta	-0,6234	0,0591	0,0854
Stedin	1,2345	0,1123	0,0042
Essent	-0,9361	0,1195	0,0711
NRE	-1,3135	0,2777	0,0928
ONS	0,2387	0,3492	0,6411
REDO	-0,3229	0,0098	0,0276
Westland	1,5097	0,1962	0,0191

Figuur 3: Analyse van invloedrijke punten

Figuur 3 toont aan dat de Netbeheerder ONS een te grote leverage value en de grootste Cook's distance heeft van alle netbeheerders in de dataset. Daarmee heeft deze netbeheerder dus erg veel invloed op de dataset en zou verwijderd moeten worden uit de analyse.



Figuur 4: Spreidingsdiagram en regressielijn (zonder ONS)

Figuur 4 toont opnieuw het spreidingsdiagram en de regressielijn, zonder de netbeheerder ONS. Dit nieuwe regressiemodel verklaart nog 57,8% (R^2) van de variantie in de *kapitaalkosten*.

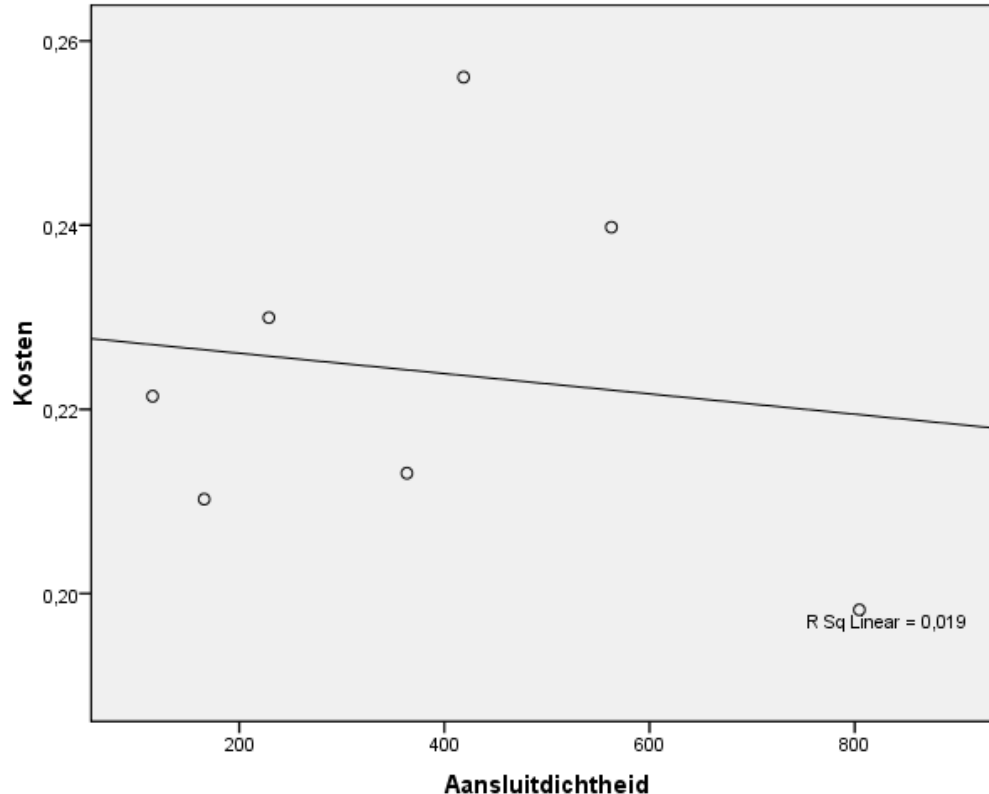
Opnieuw moet echter weer gecontroleerd worden wat voor dit nieuwe model de invloed is van de individuele meetpunten. De grenzen voor Cook's distance en de leverage value zijn aangepast aan de nieuwe situatie met $n = 8$ meetpunten en $k = 1$.

Netbeheerder	Residu (<3)	Cook's distance (<0.67)	Leverage value (<0.5)
Continuon	-0.0697	0.0007	0.0650
Overijssel	0.3434	0.0204	0.0895
Delta	-0.6984	0.1140	0.1326
Stedin	1.2157	0.1262	0.0044
Essent	-0.9743	0.1774	0.0997
NRE	-0.9144	3.6165	0.5879
ONS			
RENDO	-0.3240	0.0100	0.0157
Westland	1.4217	0.1739	0.0052

Figuur 5: Analyse van invloedrijke punten (zonder ONS)

Uit Figuur 5 blijkt dat zonder netbeheerder ONS, de netbeheerder NRE onaanvaardbaar veel invloed heeft op het gevonden regressiemodel.

Zoals we hierboven echter al hebben gezien, verklaart een regressiemodel zonder de netbeheerders ONS en NRE nog slechts 1.9% van de variantie in de *kapitaalkosten*.



Figuur 6: Spreidingsdiagram en regressielijn (zonder ONS en NRE)

Uit Figuur 6 blijkt duidelijk dat er geen sprake meer is van een eenduidig lineair verband. Verder onderzoek naar de invloed van de individuele netbeheerders heeft dan ook geen zin.

Op basis van de analyses hierboven blijkt dus dat het gevonden regressiemodel niet robuust is omdat het in sterke mate wordt bepaald door 2 van de 9 netbeheerders. Deze 2 netbeheerders hebben samen slechts 2% van alle aansluitingen in hun beheer.

Het regressiemodel op deze manier gebruiken betekent in feite dat alle andere netbeheerders (die samen 98% van alle aansluitingen beheren) worden vergeleken met deze 2 netbeheerders en dat hun verrekening aan de hand van de geschatte kosten afhangt van de prestaties van deze 2 kleine netbeheerders.

5 Regressievoorwaarden waaraan moet zijn voldaan

Bij het gebruik van een ordinary least squares regressiemethode wordt verondersteld dat aan een aantal voorwaarden is voldaan. Zo niet, dan zijn de schattingen met het gevonden regressiemodel niet betrouwbaar.

De voorwaarden waaraan moet worden voldaan luiden:

1. De meetpunten zijn onafhankelijk.
2. De variabelen zijn van minimaal ordinaal meetniveau.
3. Er is sprake van een lineair verband tussen de afhankelijke en de onafhankelijke variabele.
4. Voor een vaste waarde van de onafhankelijke variabele moet de afhankelijke variabele normaal verdeeld zijn.
5. De afhankelijke variabele moet een constante variantie hebben voor de verschillende waarden van de onafhankelijke variabele.

Hoewel in het vorige hoofdstuk is aangetoond dat het regressiemodel in sterke mate bepaald is door twee individuele meetpunten, wordt hier toch uitgegaan van het oorspronkelijke model. De voorwaarden zullen hieronder één voor één besproken worden.

Voorwaarde 1: De meetpunten zijn onafhankelijk

Aangezien het hier alle 9 onafhankelijke netbeheerders betreft, mag er van worden uitgegaan dat aan deze voorwaarde is voldaan.

Voorwaarde 2: De variabelen zijn van minimal ordinaal meetniveau

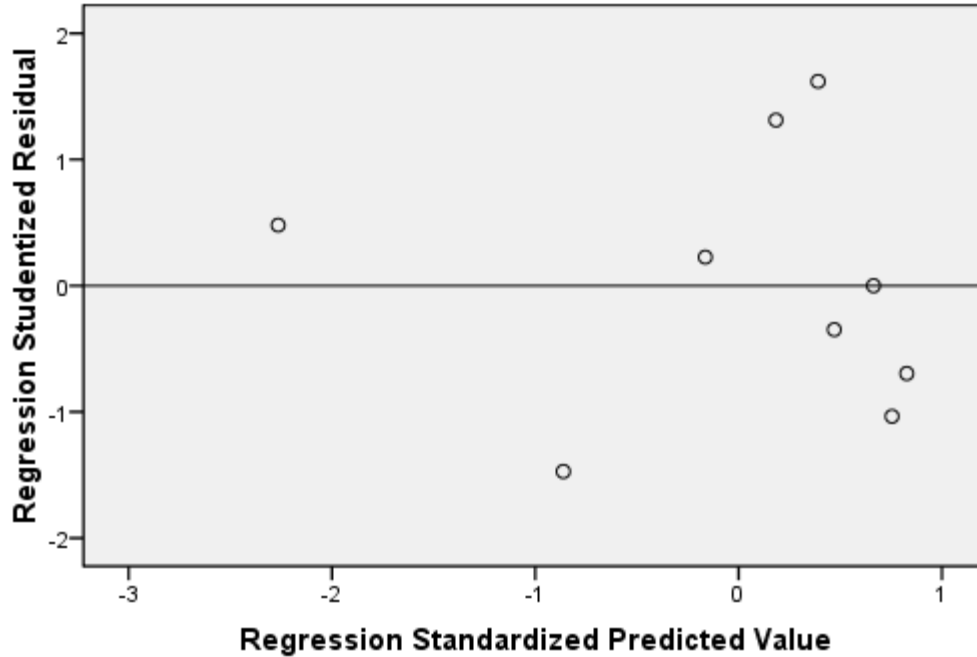
Dit betekent dat er minstens een ordening tussen de waarden van de variabelen aanwezig moet zijn. In dit geval is hier zeker aan voldaan, want beide variabelen *aansluitdichtheid* en *kapitaalkosten* zijn op een numerieke schaal gemeten.

Voorwaarde 3: Er moet een lineair verband aanwezig zijn

Aan deze voorwaarde wordt niet in voldoende mate voldaan. Het spreidingsdiagram en ook de analyses uit het vorige hoofdstuk tonen duidelijk aan dat het statistisch gevonden lineaire verband vooral wordt bepaald door de aanwezigheid van de netbeheerders ONS en NRE.

Voorwaarde 4 en 5: Normaalverdeling en constante variantie van de afhankelijke variabele voor vaste waarden van de onafhankelijke variabele

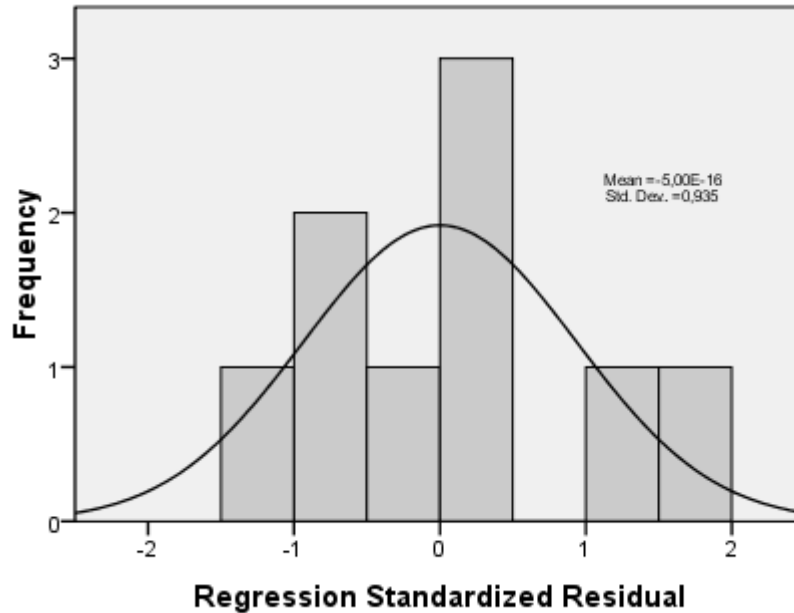
Voor grotere datasets wordt dit aangetoond door naar een aantal plots van de residuen te kijken. In de eerste plot, worden de residuen uitgezet tegen de geschatte waarden van de afhankelijke variabele. Als deze plot een mooie goed verdeeld puntenwolk laat zien, zijn er geen afwijkingen in de verdeling van de afhankelijke variabele te verwachten.



Figuur 7: Spreidingsdiagram van de residuen t.o.v. de geschatte waarden

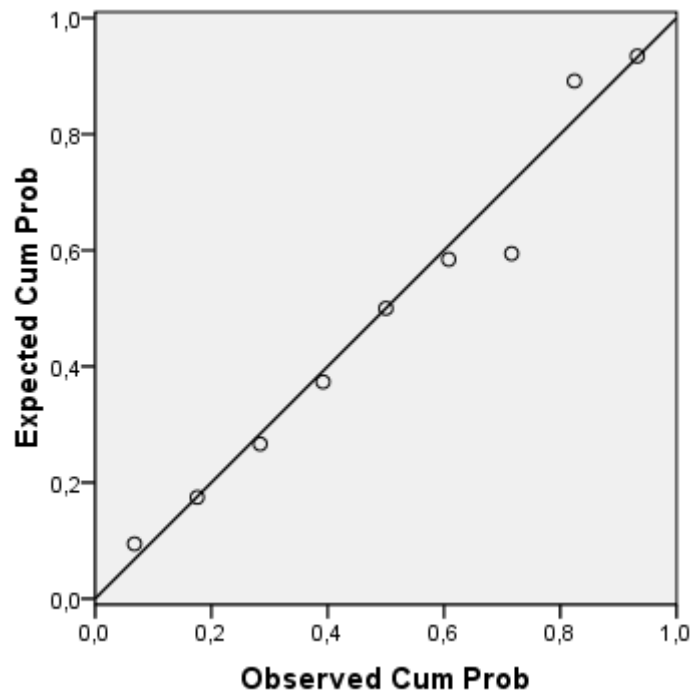
Figuur 7 toont deze plot van de residuen. Het is duidelijk dat op basis van zo weinig meetpunten niets zinnigs gezegd kan worden over de verdeling van de residuen.

Een histogram van de residuen moet aantonen dat de residuen normaal verdeeld zijn.



Figuur 8: Histogram van de residuen

Figuur 8 toont dit histogram van de residuen en de bijbehorende (zelfde gemiddelde, zelfde variantie) normaalverdeling. Ook hierbij is duidelijk dat op basis van zo weinig meetpunten niets zinnigs valt af te leiden over de verdeling van de residuen. Om die reden is het onbegrijpelijk dat de Energiekamer in de bijlage 1, formule 35, concludeert “Deze term kent een statistische normale verdeling”, waarbij “Deze term” verwijst naar de residuen. Dit wordt nergens bewezen en kan ook niet worden bewezen.



Figuur 9: Normalwaarschijnlijkheidsplot

Figuur 9 toont de normaalwaarschijnlijkheidsplot van de residuen. Deze plot vergelijkt het gevonden percentage meetpunten kleiner dan een zekere waarde met het verwachte aantal meetpunten kleiner dan die waarde als de normaalverdeling was gevolgd. De normaalwaarschijnlijkheidsplot toont een redelijke overeenkomst met de normaalverdeling op een paar punten na. Maar ook hier geldt weer dat de normaalverdeling voor zo'n kleine dataset niet is aan te tonen.

6 Gewogen versus ongewogen regressiemodel

In het regressiemodel zoals gepresenteerd door de Energiekamer is uitgegaan van een ongewogen regressiemodel [p266, ontwerp-methodebesluit]. Zoals in Hoofdstuk 3 al is aangetoond, is het daardoor mogelijk dat het lineaire verband, waarop het regressiemodel is gebaseerd, wordt veroorzaakt door twee netbeheerders ONS en NRE, die samen slechts 2% van alle aansluitingen beheren.

De keuze om een ongewogen regressiemodel te schatten, is slechts één van de vele keuzes die gemaakt hadden kunnen worden, maar één die wel zeer bepalend is voor het uiteindelijke resultaat.

Een betere keuze zou zijn om de invloed van een netbeheerder uit te drukken in het aantal aansluitingen dat de netbeheerder beheert. In dat geval wordt een gewogen regressieanalyse uitgevoerd. Het gewogen regressiemodel luidt:

$$\text{kosten} = 0,228872 - 0,000016 * \text{aansluitdichtheid}$$

Dit regressiemodel verklaart echter nog maar 4.4% van de variantie in *kapitaalkosten* van de verschillende netbeheerders.

7 Conclusies

Op basis van het in dit rapport besproken onderzoek, kan de volgende conclusie worden getrokken.

Het regressiemodel zoals het is gebaseerd op de dataset met gegevens over de *aansluitdichtheid* en *kapitaalkosten* van 9 netbeheerders uit 2006 toont niet aan dat er een sterke relatie is tussen de *aansluitdichtheid* en de *kapitaalkosten* en dient dus ook niet als basis te dienen voor een vergelijking met de werkelijke *kapitaalkosten* en de daarbij behorende verrekening.

Deze conclusie is gebaseerd op de volgende observaties:

- Het onderzoek is uitgevoerd op een complete dataset. Er is dus geen sprake van onzekerheid in de gevonden relatie. De significantiewaarde die volgt uit de statistische analyses kan zeker niet gebruikt worden om aan te tonen dat de resultaten enige praktische relevantie hebben. Alleen de hoogte van de correlatiecoëfficiënt en de daarmee samenhangende verklaarde variantie door het geschatte regressiemodel is hiervoor bepalend. Er bestaat echter geen eenduidige ondergrens waarboven sprake is van een sterke relatie tussen de twee variabelen *aansluitdichtheid* en *kapitaalkosten*.
- De dataset is compleet maar met 9 netbeheerders erg klein. De invloed van twee individuele netbeheerders is hierdoor onaanvaardbaar groot en zij bepalen voor het grootste gedeelte de relatie tussen *aansluitdichtheid* en *kapitaalkosten*. Na verwijdering van deze 2 netbeheerders, die samen slechts verantwoordelijk zijn voor 2% van alle aansluitingen,

wordt nog nauwelijks enige variantie (1.9%) in *kapitaalkosten* van de overige netbeheerders verklaard.

- De dataset is erg klein, hierdoor is het niet aan te tonen dat aan een aantal belangrijke regressievoorwaarden is voldaan. De gebruikte regressiemethode is echter niet geldig als niet aan deze voorwaarden is voldaan. In dat geval kan het regressiemodel niet worden gebruikt om betrouwbare schattingen te maken.
- De keuze om alle netbeheerders hetzelfde gewicht te geven in de regressieanalyse is zeer bepalend voor het uiteindelijke resultaat. Als een gewogen regressieanalyse wordt uitgevoerd, waarbij het gewicht van een netbeheerder wordt bepaald door het aantal aansluitingen dat zij beheren, dan blijft er van het gevonden verband tussen *aansluitdichtheid* en *kapitaalkosten* nauwelijks meer wat over en de verklaarde variantie (4.4%) is ook in dat geval te laag om nog enige praktische relevantie te bezitten.

De verwachting is dat dezelfde problemen zullen optreden als data wordt gebruikt uit enig ander jaar, maar dit moet nog verder onderzocht worden.