

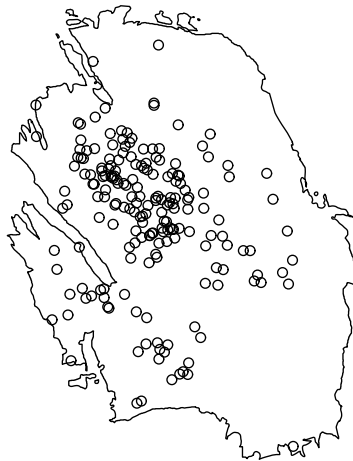
Geïnduceerde aardbevingen in gasveld Groningen

Een statistische analyse

Bacheloropdracht Technische Wiskunde
Begeleid door: Prof. dr. Marie-Colette van Lieshout
Universiteit Twente

Eline van Hove, s1314858
Rico van Lingen, s1381083
Sem Riemens, s1344064

17 juni 2015



Samenvatting

Sinds 1963 wordt er in het noorden van Nederland aardgas gewonnen. De aardbevingen die hierdoor ontstaan worden 'geïnduceerde aardbevingen' genoemd. De hoofdvraag van dit onderzoek luidt: hoe kun je in wiskundig opzicht de locaties en tijden van geïnduceerde aardbevingen in het gasveld Groningen modelleren en voorspellen? Uit de exploratieve analyse blijkt dat er een sterk positief verband is tussen het aantal aardbevingen en de hoeveelheid gewonnen gas. Een toename in magnitudes is niet aantoonbaar. In dit verslag wordt een eenvoudig wiskundig model geschetst dat de jaarlijkse bevingensintensiteit modelleert. Dit model is gebaseerd op een inhomogeen Poissonproces waarbij de relevante data zoals breuklijnen en gaswinning wordt meegenomen als covariaten in de vorm van pixelbeelden. Uit voorspellingen met dit model blijkt dat het jaarlijks aantal aardbevingen convergeert wanneer de jaarlijkse gasproductie constant is. Wanneer de gaskraan wordt dichtgedraaid convergeert het jaarlijks aantal bevingen binnen enkele jaren naar dertien.

Voorwoord

Dit onderzoek is gedaan als eindopdracht van de bachelor Technische Wiskunde aan de Universiteit Twente. Het is uitgevoerd door drie studenten onder begeleiding van Prof. dr. Marie-Colette van Lieshout. Zij is tevens de opdrachtgever. Deze bachelor eindopdracht nam per persoon 10 EC in beslag in het tweede semester van het collegejaar 2014-2015. Een deel van deze tijd is besteed aan het wegwijs worden binnen de ruimtelijke statistiek en bijbehorende software.

Inhoudsopgave

1	Inleiding	4
2	Modellen en technieken in de ruimtelijke statistiek	6
2.1	Ruimtelijke puntpatronen	6
2.2	Toetsen puntpatronen	7
2.3	Ondersteunende technieken	11
2.4	Drie stochastische processen: Poisson-cluster, Cox en Hawkes	13
3	Geologische achtergrond	16
3.1	Oorzaak van seismiciteit	17
3.2	Model van de bodem	17
3.3	Gaswinning en druk in het reservoir	18
3.4	Veranderingen in het reservoir: compactie	19
3.5	Gevolgen van compactie: bodemdaling en seismiciteit	20
3.6	Naschokken	20
4	Exploratieve analyse	22
4.1	Locaties aardbevingen	22
4.2	Het aantal aardbevingen en gaswinning	26
4.3	Magnitude aardbevingen en gaswinning	32
5	Model	34
5.1	Modelopbouw	34
5.2	Schatten van de parameters	34
5.3	Verklarende variabelen	35
6	Parameterschatting en modelvalidatie	38
6.1	De geschatte parameters β_i	38
6.2	Validatie van het model voor 2014	38
7	Voorspellingen	40
7.1	Verwachting voor 2015	40
7.2	De gaskraan gaat dicht, en dan?	41
7.3	Andere toekomstscenario's	42
8	Conclusie	43
9	Discussie	45
	Referenties	46
10	Bijlagen	47
10.1	Data gaswinning	47
10.2	Data bodemdaling	48
10.3	Data breuklijnen	49
10.4	R codes	50
10.5	Toetsen met F, G, en J-functie	56

1 Inleiding

De aardbevingen in Groningen zijn geregeld onderwerp van het nieuws. In kranten, op de radio en op televisie wordt er, vooral de laatste maanden, vaak aandacht besteed aan de situatie in het noorden van ons land. De emoties lopen soms hoog op. Prikkelende koppen prijken boven krantenartikelen: ‘Grote kans op zware aardbevingen in Groningen’¹, ‘Weer aardbeving Groningen door aardgaswinning’², ‘Onderzoeksraad: risico’s gaswinning Groningen genegeerd’³, ‘Nog minstens 1100 aardbevingen in Groningen’⁴. De vraag is in hoeverre deze uitspraken gegrond zijn.

Sinds 1963 wordt er in het noorden van Nederland aardgas gewonnen. Als gevolg van deze aardgaswinning daalt de interne gasdruk in het reservoir. De gashoudende lagen in het gesteente worden daardoor samengedrukt door bovenliggende gesteentelagen. Compactie veroorzaakt enerzijds bodemdaling en anderzijds seismiciteit, oftewel: aardbevingen (zie TNO en Dost 2014). Deze aardbevingen als gevolg van aardgaswinning noemen we geïnduceerde aardbevingen. Geïnduceerde aardbevingen zullen het onderwerp zijn van ons onderzoek. Wij zullen ons niet richten op de geologische achtergrond van aardbevingen, maar er vanuit de statistiek naar kijken. Er zal in dit verslag worden gesproken over ‘aardbevingen’, daarmee worden vanaf nu de geïnduceerde aardbevingen bedoeld.

De hoofdvraag van ons onderzoek luidt: hoe kun je in wiskundig opzicht de locaties en tijden van geïnduceerde aardbevingen in het gasveld Groningen modelleren en voorspellen? De deelvragen die we daarbij stellen zijn:

1. Is er een toename in het aantal aardbevingen?
2. Is er een toename in de magnitudes van de aardbevingen?
3. Is er een aantoonbaar verband tussen de hoeveelheid gewonnen gas en het aantal aardbevingen?
4. Is er een aantoonbaar verband tussen de hoeveelheid gewonnen gas en de magnitudes van aardbevingen?
5. Is er een aantoonbaar verband tussen de bodemdaling en het aantal aardbevingen?
6. Hoe zijn de aardbevingen verdeeld over de ruimte?
7. Na hoeveel tijd vinden er geen aardbevingen meer plaats, wanneer de gaswinning wordt stopgezet?
8. Wat gebeurt er met het aantal aardbevingen wanneer er een constante hoeveelheid gas wordt gewonnen?

Er werd en wordt veel onderzoek gedaan naar de situatie in Groningen met betrekking tot de geïnduceerde aardbevingen. De afgelopen jaren heeft het kabinet veertien onderzoeken laten uitvoeren rondom de aardbevingen. De reden hiervoor was om een goed besluit te kunnen nemen over de toekomst van gaswinning in Groningen. Grotendeels zijn deze onderzoeken niet relevant voor ons onderzoek. Zo is er bijvoorbeeld onderzocht wat de effecten zijn van sterke bevingen op dijken en gebouwen en hoe de leefbaarheid kan worden verbeterd in Groningen. Er zitten echter ook onderzoeken tussen die aansluiten bij ons onderzoek.

TNO heeft in 2014 onderzoek gedaan naar de effecten van productiebeperking (zie TNO en Dost 2014). Ze hebben de trend in het jaarlijks aantal optredende bevingen aan de hand van een statistische methode geanalyseerd. Voor de periode 2003 tot en met 2013 geeft deze analyse aan dat de waargenomen toename van het aantal bevingen in het veld in de tijd significant is. Op 17 januari 2014 werd de productie van de vijf productieputten in de omgeving van Loppersum met 80% verminderd. Vanaf deze datum laat de statistische analyse van TNO een lichte voorkeur zien voor een model dat een afname in het aantal bevingen in 2014 voorspelt. Het aantal waargenomen bevingen over 2014 is echter niet voldoende om statistisch sterke uitspraken te kunnen doen.

¹NRC, 9 oktober 2013, Wubby Luyendijk

²AD.nl, 30 december 2014

³De Volkskrant, 18 februari 2015, Jurre van den Berg

⁴RTL nieuws, 4 april 2015

Ook het CBS heeft in 2014 onderzoek gepubliceerd over de aardbevingen in Groningen. Dit onderzoek werd uitgevoerd in opdracht van Staatstoezicht op de Mijnen (SodM) (zie Pijpers 2014). Uit dit onderzoek bleek dat er twee zones kunnen worden aangewezen waarin de intensiteit van het aantal aardbevingen significant hoger is dan in de omgeving. Dit zijn zones waarin gasproductie plaatsvindt.

In de gedane onderzoeken zijn modellen gebouwd om de situatie in Groningen te beschrijven. Deze modellen zijn echter zeer gecompliceerd. Ze zijn opgebouwd uit meerdere lagen en maken gebruik van veel parameters met grote onzekerheden. Voor tectonische aardbevingen bestaan simpelere modellen. Omdat in Groningen de aardbevingen plaatsvinden als gevolg van gaswinning, en niet als gevolg van tectonische bewegingen, zijn deze modellen hier echter niet toepasbaar. Wij zullen relatief simpele wiskundige technieken gebruiken om de situatie zo goed mogelijk te modelleren. We zullen bij dit onderzoek het statistisch softwarepakket R gebruiken. Binnen dit programma zullen we met name gebruik maken van het pakket SPATSTAT (zie Baddeley 2014) dat is geschreven voor ruimtelijke statistiek.

Bij ons onderzoek wordt data van het KNMI (KNMI 2015) gebruikt over de locaties, tijdstippen en magnitudes van geïnduceerde aardbevingen. Het KNMI geeft aan dat het netwerk van seismometers ontworpen is om aardbevingen te detecteren vanaf een magnitude van 1,5 (Pijpers 2014 pag.5). Wij zullen dan ook alleen de bevingen met een magnitude van 1,5 of groter beschouwen. Wanneer in dit verslag wordt gepraat over ‘aantal aardbevingen’, gaat het dus alleen om het aantal aardbevingen met een magnitude van 1,5 of hoger.

Vanaf het jaar 1986 zijn er aardbevingen in Noord-Nederland gemeten. We gebruiken echter alleen de data vanaf 1 januari 1995. Tussen 1986 en 1995 was de detectieapparatuur minder goed, waardoor niet alle aardbevingen werden gemeten of goed werden gelokaliseerd. Deze data is dus niet betrouwbaar (Pijpers 2014 pag.5).

Wat is er in dit verslag te lezen? In het hoofdstuk Modellen en technieken in de ruimtelijke statistiek wordt een inleiding gegeven in de theorie van de ruimtelijke statistiek, waarbij ruimtelijke puntpatronen een belangrijke rol spelen. Er worden een aantal stochastische processen beschreven die de basis vormen voor modellen in de ruimtelijke statistiek: Poisson-processen, Cox-processen en Hawkes-processen.

Hoewel we in ons model niet ingaan op de geologische achtergrond van seismiciteit, zullen we hier wel een hoofdstuk aan wijden om een compleet beeld te geven van de situatie (zie hoofdstuk Geologische achtergrond). In dit hoofdstuk zal worden uitgelegd hoe de aardbevingen ontstaan. Bovendien geeft dit hoofdstuk een beeld van de plaats van ons onderzoek binnen de bestaande onderzoeken. Hierna zal de data over aardbevingen en gaswinningen worden beschouwd in het hoofdstuk Exploratieve analyse. Er zal in dit hoofdstuk worden onderzocht welke verbanden en patronen er te ontdekken zijn en vooral ook welke verbanden niet aantoonbaar zijn.

De kennis die we hebben opgedaan in de exploratieve analyse gebruiken we bij het opstellen van een model voor de aardbevingen in gasveld Groningen. Dit model staat beschreven in hoofdstuk 5. In het daaropvolgende hoofdstuk, Parameterschatting en modelvalidatie, zal worden geverifieerd in hoeverre dit model correspondeert met de werkelijkheid. Met behulp van het geverifieerde model zullen we verschillende toekomstscenario's doorrekenen. De resultaten van deze berekeningen zijn te vinden in hoofdstuk Voorspellingen. Tot slot staan in de hoofdstukken 8 en 9 de conclusies en discussie. Een aantal belangrijke codes en data zijn bijgevoegd in de bijlagen (zie hoofdstuk 10).

Er wordt bij de lezer enige voorkennis van statistiek verwacht. Begrippen uit de ruimtelijke statistiek zullen worden toegelicht. Niet alle theorie wordt in dit verslag uitgebreid uitgediept. Er zal geregeld worden verwezen naar de literatuur.

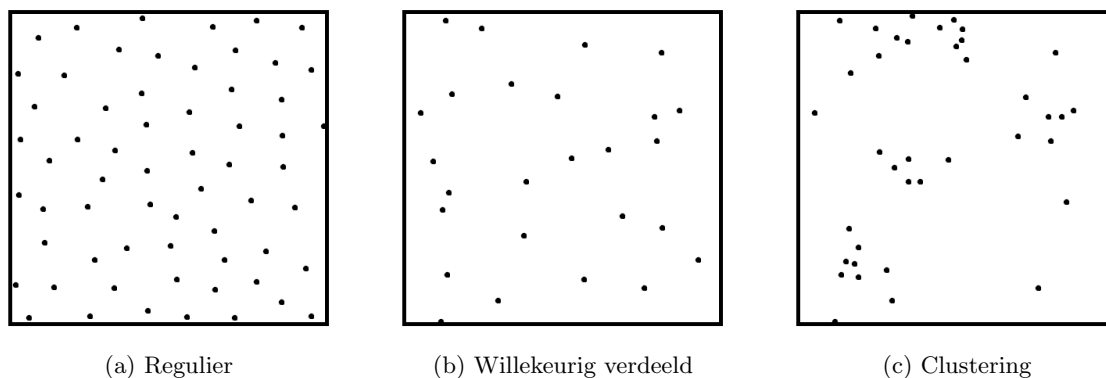
2 Modellen en technieken in de ruimtelijke statistiek

We zullen in dit hoofdstuk beginnen met een inleiding in de theorie van de ruimtelijke statistiek. Ruimtelijke puntpatronen spelen daarbij een belangrijke rol. In paragraaf 2.1 zal worden uitgelegd wat ruimtelijke puntpatronen zijn en welke drie typen we kunnen onderscheiden. De manieren waarop we kunnen toetsen met welk type puntpatroon we te maken hebben, staan beschreven in paragraaf 2.2. In de daaropvolgende paragraaf worden een aantal ondersteunende technieken uit de ruimtelijke statistiek uitgelegd. Deze technieken hebben we toegepast tijdens ons onderzoek. In paragraaf 2.4 zullen drie soorten stochastische processen worden beschreven die de basis vormen voor modellen in de ruimtelijke statistiek: Poisson-clusterprocessen, Cox-processen en Hawkes-processen.

2.1 Ruimtelijke puntpatronen

In ons onderzoek naar de geïnduceerde aardbevingen in gasveld Groningen zullen puntpatronen een belangrijke rol spelen. In een puntpatroon zijn observaties in een gedefinieerde ruimte weergegeven, bijvoorbeeld als punten in een tweedimensionaal vlak. Een stochastisch model bij deze puntpatronen heet een puntproces. Een voorbeeld van een toepassing van puntpatronen is het weergeven van bomen in een bepaald bosgebied. Niet alleen het aantal punten in het puntpatroon geeft informatie over hoe het bos eruit ziet, ook de verdeling van de punten over het vlak speelt een belangrijke rol. Staan de punten geclusterd of zijn de punten gelijkmatig verspreid over het oppervlak? Of zien we dat er in een bepaald gebied totaal geen punten staan? Dit kan duiden op bijvoorbeeld een waterplas in het gebied. In Diggle 1983 worden een aantal processen uit de biologie gemodelleerd als puntprocessen. In dit verslag zullen we locaties van aardbevingen op een tweedimensionaal vlak bestuderen met behulp van puntprocessen.

Grofweg kunnen we drie typen verdelingen bij puntpatronen onderscheiden: reguliere verdeling, clustering en willekeurige verdeling. Wanneer er sprake is van een reguliere verdeling zal elk punt zo ver mogelijk van al zijn burens vandaan zijn. Clustering is daarvan de tegenhanger: punten in groepjes geconcentreerd waardoor zowel gebieden met een relatief hoge als gebieden met een relatief lage punten-dichtheid ontstaan.



Figuur 1: Drie verschillende typen puntpatronen.

Bij een willekeurige verdeling over het oppervlak zal een punt overal in het gebied gelijke kans hebben van voorkomen. Bovendien beïnvloeden de posities van punten elkaar niet. Bij een willekeurige verdeling over het oppervlak wordt uitgegaan van een homogeen Poissonproces binnen het bestudeerde gebied. De puntendichtheid wordt meestal aangeduid met λ . Dit wordt ook wel de intensiteit genoemd. Wanneer je een willekeurig deelgebied bekijkt, dan is de verdeling van het aantal punten Poisson verdeeld. Bovendien is de verdeling in het deelgebied onafhankelijk van de puntenverdeling in andere disjuncte deelgebieden.

In een gebied R met een willekeurige verdeling en intensiteit λ is de kans om precies k punten te vinden

$$P(k|\lambda, R) = \frac{(|R|\lambda)^k}{k!} \cdot e^{-(|R|\lambda)} \quad (1)$$

met $|R|$ de maat van het gebied.

In veel toepassingen zullen de gebeurtenissen zich niet volgens een homogeen Poissonproces verdelen. Het kan zijn dat de kans om een aantal gebeurtenissen in een gebied te vinden verschilt, in welk geval er sprake is van een inhomogeen Poissonproces. De intensiteit hangt dan af van de locatie $\mathbf{x} \in R$:

$$P(k|\lambda, R) = \frac{(\int_R \lambda(\mathbf{x}) d\mathbf{S})^k}{k!} \cdot e^{-\int_R \lambda(\mathbf{x}) d\mathbf{S}}. \quad (2)$$

Aan de andere kant kunnen de gebeurtenissen ook invloed op elkaar hebben, bijvoorbeeld doordat ze elkaar aantrekken of afstoten. Dit heet interactie. Het kansproces dat de punten genereert wordt dan beschreven door de intensiteit geconditioneerd op de al aanwezige punten. Formule (2) geeft dan de kans om nog k punten te vinden in een gebied R , gegeven de uitkomst van het proces “tot nu toe”. De intensiteit λ mag afhangen van de locaties van andere punten. In dit geval dienen de punten wel geordend te zijn. Die ordening wordt bijna altijd door de tijd bepaald. De “reeds aanwezige punten” zijn dan simpelweg de gebeurtenissen die al hebben plaatsgevonden. Verschillende modellen, met en zonder interactie(-componenten), worden beschreven vanaf paragraaf 2.4.1.

In de volgende paragraaf zullen methoden worden besproken waarmee kan worden getoetst of er sprake is van een willekeurige verdeling over de ruimte. Merk op dat deze methoden geen verschil kunnen maken tussen interactie en inhomogeniteit. Meer informatie over verschillende verdelingen van puntprocessen is te vinden in Diggle 1983 hoofdstuk 4: Spatial Point Processes.

2.2 Toetsen puntpatronen

We hebben beschreven welke drie typen puntpatronen we kunnen onderscheiden: reguliere verdeling, clustering en willekeurige verdeling. In deze paragraaf zullen we een aantal methoden bespreken om te toetsen of een puntpatroon willekeurig is verdeeld over de ruimte. We beginnen met de twee relatief eenvoudigste methoden Kwadrantmethode en Naaste-Buur-methode. Hierna zullen we Ripley’s K -functie bespreken. We maken onderscheid tussen de homogene en inhomogene variant van deze functie.

2.2.1 Kwadrantmethode

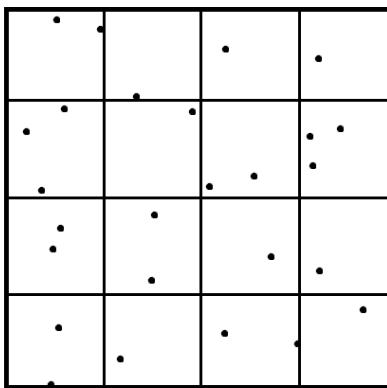
Een relatief eenvoudige methode om te toetsen of een puntpatroon willekeurig is verdeeld over de ruimte, is de kwadrantmethode. Bij deze methode wordt het gebied opgedeeld in gelijke deelgebieden, cellen genaamd. ‘Gelijk’ houdt niet alleen in dat de cellen een gelijke oppervlakte hebben, maar ook dat ze gelijkvormig zijn. Meestal wordt het gebied met behulp van een raster verdeeld in cellen. Vervolgens wordt het aantal punten per cel geteld en de puntendichtheid per cel berekend.

Stel, we beschouwen een rechthoekig gebied R met daarin n punten. We verdelen R in m congruente rechthoeken C_1, \dots, C_m met ieder een oppervlakte $|C_i|$, $i = 1, 2, \dots, m$. Omdat we alle cellen zo hebben gekozen dat de oppervlakken gelijk zijn, zal $|C_i|$ constant zijn voor alle waarden van $i = 1, 2, \dots, m$. De stochast voor het totaal aantal punten in cel C_i noemen we N_i . De geschatte dichtheid in gebied R is $\hat{\lambda}$ met

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^m |C_i|} = \frac{n}{m|C_i|} = \frac{n}{|R|}.$$

Onder de hypothese van willekeurige verdeling moet dan gelden:

$$P(N_i = k|\hat{\lambda}, C_i) = \frac{(\hat{\lambda}|C_i|)^k}{k!} \cdot e^{-(\hat{\lambda}|C_i|)}, \quad k = 0, 1, 2, \dots$$



Figuur 2: Kwadrantmethode met $m = 16$

Met de Pearson χ^2 -toets kunnen we nu toetsen of er daadwerkelijk sprake is van een willekeurige verdeling over de ruimte. Onder de nulhypothese is het verwachte aantal punten in elke cel

$$E(N_i|\hat{\lambda}) = |C_i| \cdot \hat{\lambda} = |C_i| \cdot \frac{n}{m|C_i|} = \frac{n}{m}.$$

De toetsingsgrootte die wordt gebruikt is

$$T = \sum_{i=1}^m \frac{n_i - n/m}{n/m}$$

met n_i de waargenomen waarde van stochast N_i . Onder de nulhypothese van willekeurige verdeling heeft toetsingsgrootte T een χ^2 -verdeling met $m - 1$ vrijheidsgraden. Kleine waarden van T duiden op regulariteit en grote waarden van T op clustering. Meer over deze χ^2 -toets is te lezen in Diggle 1983 in paragraaf 3.2 Quadrat Counts.

Een nadeel van de kwadrant-methode is dat er geen eenduidige ‘beste’ manier is om het gebied te verdelen in cellen. De uitkomst van de toets kan afhankelijk zijn van de gekozen rastergrootte (zie ook Smith 2014, hoofdstuk 3.1 Quadrat Method). Wanneer de cellen te klein worden gekozen, dan zullen verschillen tussen deelgebieden minder zichtbaar worden. Wanneer daarentegen de cellen te groot zijn, dan zullen mogelijke verschillen in dichtheid binnen een cel niet worden opgemerkt. Het verdelen van R in cellen kan ook worden bemoeilijkt door de vorm van het gebied, bijvoorbeeld wanneer we te maken hebben met grillige landgrenzen.

Een ander nadeel van de kwadrantmethode is dat mogelijke patronen niet worden ontdekt, omdat deze methode niet is gericht op de relaties van punten onderling maar vooral op de punt dichtheid. Bovendien zullen variaties binnen cellen C_i niet worden opgemerkt, omdat deze methode alleen iets zegt over de totale cel.

2.2.2 Naaste-Buur-methode

Een tweede methode om te toetsen of een puntpatroon willekeurig is verdeeld over de ruimte, is de Naaste-Buur-methode. In tegenstelling tot de kwadrantmethode, gaat deze methode juist uit van onderlinge relaties van punten. Deze methode is dus een goede aanvulling op de kwadrantmethode.

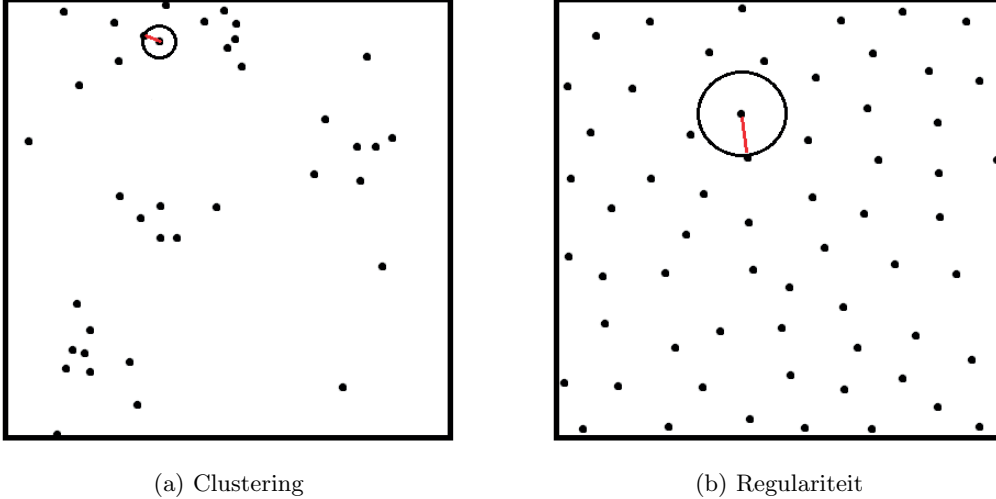
Bij de Naaste-Buur-methode wordt voor ieder punt uitgerekend wat de afstand is tot het dichtstbijzijnde andere punt in R . Het berekenen van deze afstand wordt gedaan met de Euclidische afstand. Gegeven punten (x_i, y_i) en (x_j, y_j) in \mathbb{R}^2 , dan is de afstand tussen deze twee punten

$$d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \quad (3)$$

De minimale afstand van een punt (x_i, y_i) tot een willekeurig ander punt in R noteren we als

$$d_i = d_i(S_R) = \min\{d_{i,j} | (x_j, y_j) \in S_R, (x_j, y_j) \neq (x_i, y_i)\}$$

met S_R , de verzameling punten in $R \subseteq \mathbb{R}^2$.



Figuur 3: Naaste-Buur-methode bekijkt de minimale afstand van een punt i tot een willekeurig ander punt in de ruimte.

De grootte van de kortste afstand is een directe maat voor regulariteit (bij reguliere verdeling) of clustering in het puntpatroon. Wanneer er sprake is van regulariteit, zal de kortste afstand groot zijn en bij clustering zal deze klein zijn. Maar wat is de verdeling van de kortste afstand wanneer er sprake is van een willekeurige verdeling over de ruimte? Een uitgebreide uitleg over het toetsen met de Naaste-Buur-methode is te lezen in Smith 2014 paragraaf 3.2.1 Nearest-Neighbor Distribution under CSR. Het blijkt dat de geschaalde som S_m ,

$$S_m = 2\lambda\pi \sum_{i=1}^m D_i^2$$

met (D_1, \dots, D_m) een aselechte steekproef van m Naaste-Buur-afstanden, onder de nulhypothese χ^2 is verdeeld met $2m$ vrijheidsgraden. Dit geeft een methode om te toetsen, gebaseerd op de Naaste-Buur-methode.

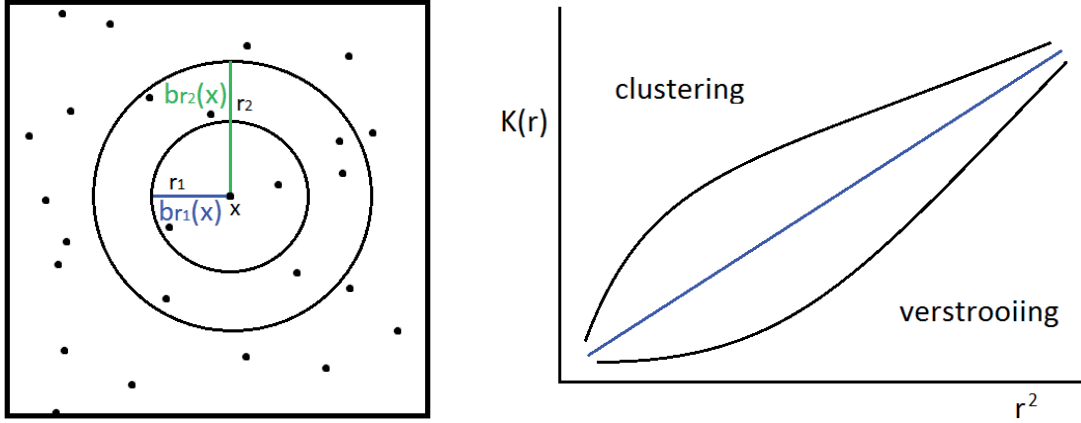
2.2.3 Homogene K-functie

Een andere methode om te toetsen of een puntpatroon totaal willekeurig is verdeeld over de ruimte is met Ripley's K-functie, ook wel de homogene K-functie. Een aanname die we doen bij deze methode is dat er sprake is van isotropie ⁵. Ripley's K-functie neemt de afstanden tussen punten mee op verschillende schaalniveaus. We bekijken bij deze functie het aantal punten in een omgeving met straal r rondom een willekeurig gekozen punt $x \in R$. De gesloten cirkel rondom punt x met straal r noteren we met $b_r(x)$. De K-functie gaat uit van het verwachte aantal andere punten in cirkel $b_r(x)$. We definiëren een functie K als volgt:

$$K(r) = \frac{1}{\lambda} E(\#y \in b_r(x) | x, y \in R, x \neq y, x \text{ willekeurig gekozen})$$

met λ wederom de intensiteit. De K-functie kan nu worden geplot tegen de straal r . De vorm van deze functie zegt iets over de verdeling van het puntpatroon. In figuur 4 is meer te zien over de vorm van functie $K(r)$ en de bijbehorende verdeling.

⁵Bij isotropie ga je ervan uit dat de verdeling van het puntpatroon rotatie-invariant is.



Figuur 4: Ripley's K-functie. De blauwe lijn geeft het verloop van $K(r)$ onder de nulhypothese dat het puntpatroon totaal willekeurig is verdeeld over het vlak. Wanneer er sprake is van clustering, zal $K(r)$ bij kleine r relatief groot zijn. Wanneer er sprake is van verstrooiing, zal $K(r)$ bij kleine r relatief klein zijn.

We kunnen een schatter $\hat{K}(r)$ voor $K(r)$ beschouwen. Deze schatter berekent voor ieder punt (i) hoeveel andere punten (j) in de cirkel $b_r(i)$ liggen. Daarna wordt er over deze n aantallen gesommeerd en gedeeld door n . Zo wordt het gemiddelde aantal punten per cirkel berekend. Tot slot wordt dit gemiddelde vermenigvuldigd met een factor $\frac{1}{\lambda}$. We vinden nu de volgende schatter $\hat{K}(r)$:

$$\hat{K}(r) = \frac{1}{\hat{\lambda}n} \sum_{i=1}^n \sum_{j \neq i} I(d_{i,j} \leq r)$$

met $d_{i,j}$ zoals gedefinieerd in functie (3), n het totale aantal punten in R , $\hat{\lambda} = \frac{n}{|R|}$ en

$$I(d_{i,j} \leq r) = \begin{cases} 1 & \text{if } d_{i,j} \leq r \\ 0 & \text{if } d_{i,j} > r. \end{cases}$$

Het probleem met deze schatter is echter dat er geen rekening wordt gehouden met de randen van het bestudeerde gebied R . Bij cirkels $b_r(x)$ met grote straal r zal een deel van de cirkel buiten gebied R vallen. Om ook de invloed van de randen mee te nemen in de schatter $\hat{K}(r)$ kan er een factor $w_{i,j}^{-1}$ worden toegevoegd aan de functie. Deze factor kan op verschillende manieren worden gedefinieerd. Ripley (zie Ripley 1984) definieerde $w_{i,j}$ als volgt:

$$w_{i,j} = \frac{a_{i,j}}{2\pi d_{i,j}},$$

met $a_{i,j}$ de lengte van de straal die binnen gebied R valt. De nieuwe schatter wordt nu gedefinieerd als

$$\hat{K}(r) = \frac{1}{\hat{\lambda}n} \sum_{i=1}^n \sum_{j \neq i} w_{i,j}^{-1} I(d_{i,j} \leq r).$$

Hoe ziet de K-functie er uit onder de nulhypothese dat het puntpatroon willekeurig verdeeld is over de ruimte? Onder de nulhypothese geldt $K(r) = \pi r^2$. Dit is eenvoudig af te leiden uit het feit dat het verwachte aantal punten y ($y \neq x$) in een bolletje $b_r(x)$ onder de nulhypothese gelijk is aan $\lambda \cdot |b_r(x)| = \lambda \cdot \pi r^2$. Door het vergelijken van de bovengenoemde schatter $\hat{K}(r)$ en de functie $K(r)$ onder de nulhypothese, kan een uitspraak worden gedaan over het mogelijk verwerpen van de nulhypothese. Intuïtief is al wel te zeggen dat de nulhypothese zal worden verworpen wanneer de geplote schatter $\hat{K}(r)$ en $K(r) = \pi r^2$ veel verschillen. Met behulp van betrouwbaarheidsintervallen zijn hier uitspraken over te doen. In Baddeley 2014 is meer te lezen over het construeren van betrouwbaarheidsintervallen en het toetsen van de nulhypothese.

We gingen tot nu toe uit van een homogene intensiteit λ : overal in gebied R is de verwachte punten-dichtheid gelijk. Er bestaat echter ook de mogelijkheid om de K-functie te gebruiken bij een inhomogene intensiteit. In de volgende paragraaf zullen we deze methode verder toelichten.

2.2.4 Inhomogene K-functie

Tot nu toe gingen we uit van een homogene intensiteit λ . Wanneer is vastgesteld dat het puntpatroon niet totaal willekeurig is verdeeld, is het mogelijk om meer onderzoek te doen naar de (inhomogene) intensiteit. ‘Inhomogeen’ duidt erop dat de intensiteit niet constant is over de ruimte. Een belangrijke vraag is of de inhomogene intensiteit van het puntproces zich hetzelfde gedraagt gedurende de tijd. Oftewel: is de inhomogene intensiteit stationair? Om deze vraag te kunnen beantwoorden kan een nieuwe versie van de K-functie worden gebruikt: de inhomogene K-functie.

Met deze K-functie kan een inhomogene intensiteit geschat worden op basis van een gegeven puntpatroon. We gebruiken hierbij een intensiteitsschatter. De meest voorkomende intensiteitsschatter is de Gaussische kernschatter waarbij klokvormige kernen over de punten in het puntpatroon worden gelegd. Hoe deze kernschatter werkt staat beschreven in paragraaf 2.3.2. De enige te kiezen parameter van deze kernschatter is de bandbreedte σ . Wanneer deze parameter te groot gekozen wordt, zullen de klokken breed worden en daarmee zal de intensiteit te vlak zijn. Wordt σ echter te klein gekozen, dan zal de schatting van de intensiteit te veel sterke pieken bevatten en erg lijken op het puntpatroon zelf. Het is zaak om een goede balans hiertussen te vinden. Er bestaat geen eenduidige ‘beste’ manier om de waarde σ te kiezen. Diggle en Berman doen een voorstel voor de keuze van σ , dit is te lezen in Berman en Diggle 1989. Zie Baddeley, Møller en Waagepetersen 2000 voor meer details over de inhomogene K-functie.

De definitie van inhomogene K-functie is dezelfde als in het homogene geval, met de aanvulling dat de intensiteit nu afhangt van de locatie van het gekozen punt. Deze inhomogene K-functie geeft de mogelijkheid om na te gaan of de (inhomogene) intensiteit stationair is. Dit kan door de intensiteit op basis van een zekere tijdsperiode te schatten, om vervolgens te toetsen of een ander gekozen tijdsperiode zich volgens diezelfde intensiteit gedraagt. Bijvoorbeeld: er is data beschikbaar van de afgelopen vijftien jaar. We schatten de intensiteit op basis van de data van de eerste vijf jaar. Hierna toetsen we of deze intensiteit ook geldt voor de data van de overige tien jaar.

2.3 Ondersteunende technieken

In deze paragraaf zullen een aantal veelgebruikte technieken uit de ruimtelijke statistiek worden besproken: het maken van pixelbeelden, het smoothen met Gaussische kernen en het opdelen van een gebied met behulp van een Voronoi tessellatie. Deze technieken zitten ook ingebouwd in het pakket SPATSTAT van softwarepakket R.

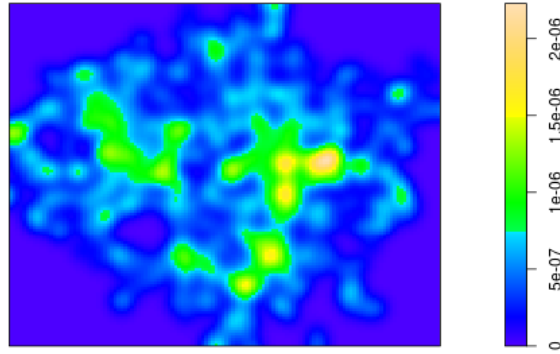
2.3.1 Pixelbeelden

In de ruimtelijke statistiek zijn pixelbeelden veelgebruikte objecten. Ze kunnen worden gebruikt om te discretiseren over een tweedimensionale ruimte. Wat zijn pixelbeelden?

Een pixelbeeld wordt gemaakt door een tweedimensionale ruimte op te delen in kleine deelgebieden met behulp van een fijn rooster. Deze deelgebieden noemen we pixels. We kunnen nu aan elke pixel een waarde - en een bijbehorende kleur - koppelen. Zo kunnen we een gediscrètiseerde functie over een tweedimensionaal oppervlak met een pixelbeeld inzichtelijk maken. Wanneer het rooster zeer fijn wordt gekozen kan de pixelbeeld er met het blote oog uit zien als een tweedimensionale gladde functie. In figuur 5 is een voorbeeld te zien van een pixelbeeld in SPATSTAT.

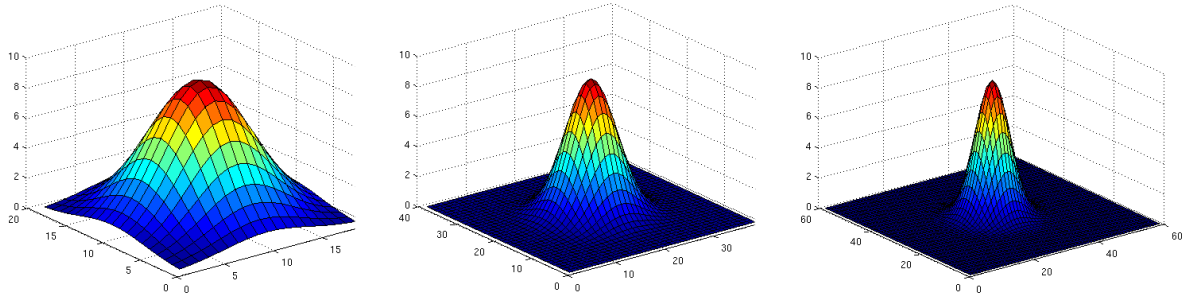
2.3.2 Smoothen met Gaussische kernen

Om een tweedimensionale intensiteitsfunctie (zoals in figuur 7) te maken met behulp van een gegeven puntpatroon kunnen Gaussische kernen worden gebruikt. Grofweg gaat deze methode als volgt: over alle punten in het puntpatroon wordt een klokvormige kern gelegd. De bandbreedte van deze kernen is variabel en te bepalen met de parameter σ . Wanneer σ groot wordt gekozen zal de functie erg glad worden.



Figuur 5: Voorbeeld van een pixelbeeld in SPATSTAT.

Er zal bovendien weinig reliëf te zien zijn. Wanneer echter de bandbreedte σ klein wordt gekozen, zullen in het oppervlak scherpe pieken te zien zijn. In figuur 6 zijn drie Gaussische kernen te zien met verschillende bandbreedtes. In deze afbeelding lijken de kernen dezelfde hoogte te hebben. In werkelijkheid wordt de hoogte van de kernen bepaald door de bandbreedte: de integraal over het oppervlak onder kern is namelijk altijd gelijk aan 1. Bij een grote bandbreedte is de kern lager dan bij een kleine bandbreedte.

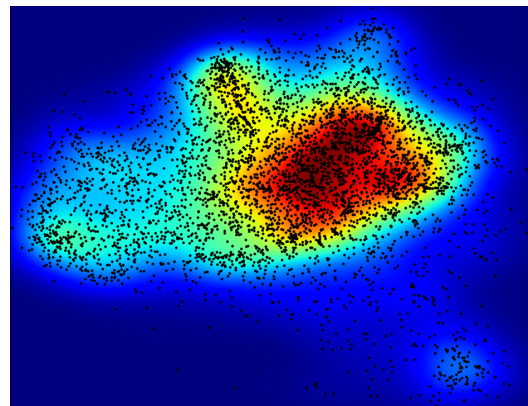


Figuur 6: Voorbeelden van Gaussische kernen met verschillende bandbreedtes.

Het is mogelijk om de kernen verschillend te schalen, zodat de kernen in verschillende mate worden meegenomen in het model. De coëfficiënten die dit bepalen worden gewichten (w_i) genoemd. De gewichten die aan de kernen worden meegegeven kunnen bijvoorbeeld een functie zijn van de marks in het puntpatroon. De functie die ontstaat heeft de vorm $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. De functiewaarde in een punt (x, y) in het vlak is gelijk aan de som van alle geschaalde kernen in dat punt. Oftewel:

$$f(x, y) = w_1 k_1(x, y) + \dots + w_n k_n(x, y)$$

met (x, y) een punt in het tweedimensionale vlak en k_i de functie van de kern over punt i in het gegeven puntpatroon. We gaan uit van een puntpatroon met n punten, dus $i = 1, \dots, n$. Er geldt dat $\sum_{i=1}^n w_i = 1$.



Figuur 7: Voorbeeld van een dichtheidsfunctie gemaakt met Gaussische kernen.

Wat is een Gaussische kern precies? De Gaussische kern kan worden afgeleid uit de bivariate normale verdeling. Deze tweedimensionale normale verdeling wordt bepaald door vijf parameters: μ, σ, v, τ en ρ , respectievelijk: de verwachting en standaardafwijking in de eerste en tweede dimensie en de correlatiecoëfficiënt tussen beide dimensies. Omdat wij uitgaan van isotropie geldt in ons geval dat $\tau = \sigma$,

$\mu = v$ en $\rho = 1$. De kansdichtheid in een punt (x, y) is als volgt:

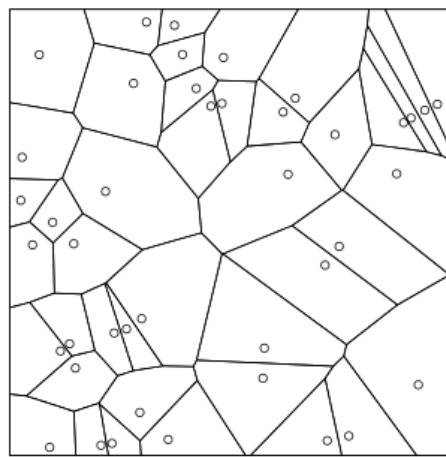
$$k(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\left(\left(\frac{x-\mu}{\sigma}\right)^2 + \left(\frac{y-\mu}{\sigma}\right)^2\right)\right).$$

Smoothen met Gaussische kernen is een manier om van een gegeven puntpatroon en een gladde functie te maken. Dit kan bijvoorbeeld worden gebruikt, wanneer de punten in het puntpatroon meetlocaties zijn. Door het smoothen wordt ondervangen dat het aantal meetpunten beperkt is. Een andere manier om dit beperkte aantal meetpunten te ondervangen is met behulp van Voronoi tessellatie. In de volgende paragraaf wordt daar verder op in gegaan.

2.3.3 Voronoi tessellatie

Naast Gaussische kernen kan ook een Voronoi tessellatie worden gebruikt om met een gegeven puntpatroon een functie te maken over een tweedimensionale ruimte. Bij deze methode wordt echter geen gladde functie verkregen, maar een diagram zoals te zien is in figuur 8. Aan elke cel in het diagram kan een kleur worden gekoppeld. Een andere benaming voor Voronoi tessellatie is Voronoi-betegeling of Dirichlet-betegeling.

Bij Voronoi tessellatie wordt een ruimte opgedeeld in deelruimten (cellen) op basis van de Euclidische afstand tot gegeven punten. Hier beschouwen we weer een tweedimensionaal geval. Een cel in een Voronoi-diagram wordt bepaald door de verzameling coördinaten in het vlak dat het dichtst bij een punt in het gegeven puntpatroon ligt. Zo worden met Voronoi tessellatie de punten in een puntpatroon centra van cellen.



Figuur 8: Voorbeeld van een Voronoi-diagram.

2.4 Drie stochastische processen: Poisson-cluster, Cox en Hawkes

In deze paragraaf zullen drie soorten stochastische processen worden beschreven die de basis vormen voor modellen in de ruimtelijke statistiek: Poisson-clusterprocessen, Cox-processen en Hawkes-processen.

2.4.1 Poisson-clusterprocessen

Poisson-clusterprocessen vormen de basis voor veel modellen in de ruimtelijke statistiek. Een belangrijk aspect van Poisson-clusterprocessen is dat er gebeurtenissen zijn die ook weer gebeurtenissen kunnen genereren. Die genererende gebeurtenissen worden ouder-gebeurtenissen genoemd. Het Poisson-clusterproces wordt gedefinieerd aan de hand van de volgende drie postulaten:

1. De ouder-gebeurtenissen vormen een Poissonproces met intensiteit ρ .
2. Iedere ouder-gebeurtenis genereert een stochastisch aantal S nakomelingen, onafhankelijk en identiek gerealiseerd door een kansverdeling met als uitkomstenruimte de gehele niet-negatieve getallen.
3. De posities van de nakomelingen ten opzichte van hun ouder-gebeurtenis zijn identiek en onafhankelijk verdeeld volgens een kansdichtheidsfunctie.

Het is gebruikelijk om als realisatie van het hele Poisson-clusterproces alleen te kijken naar nakomelingen. Echter is het ten behoeve van de toepassing toegestaan om ook de ouder-gebeurtenissen deel van de

realisatie te laten zijn. Voor de kansverdeling van S wordt in veel toepassingen gebruikt gemaakt van een Poisson-verdeling. Voor de positie van de nakomelingen ten opzichte van hun ouder-gebeurtenis wordt veelal gebruikt gemaakt van de normale verdeling. Dat juist deze twee verdelingen veel voorkomen heeft als reden dat deze verdelingen vaak het beste de situatie beschrijven. Natuurlijk zijn er veel meer mogelijkheden, aangezien de definitie van Poisson-clusterprocessen vrij algemeen is. In Diggle 1983 en Daley en Vere-Jones 2007 is meer te vinden over Poisson-clusterprocessen.

2.4.2 Cox-processen

Bij het Poisson-clusterproces wordt clustering veroorzaakt door interactie tussen gebeurtenissen. Het kan echter ook zijn dat de clustering in het puntpatroon te wijten is aan inhomogeniteit van de intensiteit. In een puntpatroon zullen dan clusters ontstaan op die plaatsen waar de intensiteit pieken vertoont. Het kan natuurlijk zijn dat de inhomogeniteit niet van te voren exact te beschrijven is, maar zelf stochastisch is. In dat geval spreken we van een Cox-proces. Een Cox-proces wordt gedefinieerd door de volgende postulaten:

1. $\{\Lambda(x) : x \in \mathbb{R}^2\}$ is een stochastisch proces dat geen negatieve waarden aanneemt.
2. Geconditioneerd op $\{\Lambda(x) : x \in \mathbb{R}^2\}$ vormen de gebeurtenissen een inhomogeen Poissonproces met intensiteit $\lambda(x)$.

Het puntproces is stationair dan en slechts dan als het proces $\Lambda(x)$ stationair is. Hetzelfde geldt voor isotropie. Men zegt ook wel dat het Cox-proces gedreven is door het intensiteitsproces $\Lambda(x)$. Het Cox-proces is een veralgemenisering van het inhomogeen Poissonproces - een niet-stochastische $\Lambda(x)$ geeft een proces met intensiteit $\lambda(x) = \Lambda(x)$. In Diggle 1983 en Daley en Vere-Jones 2007 staat het Cox-proces uitgebreid beschreven.

We willen niet onvermeld laten dat er Poisson-clusterprocessen bestaan die ook kunnen worden beschreven door een Cox-proces en vice versa. Daarnaast willen we ingaan op het onderscheid tussen clustering en inhomogeniteit, zoals te lezen is in Diggle 1983. Het is namelijk niet altijd even duidelijk om vast te stellen wat de oorzaak is van een aantal dicht bij elkaar gelegen punten in een puntpatroon. Hiervoor heb je extra informatie nodig over de totstandkoming van het puntpatroon, bijvoorbeeld in de vorm van covariaten.

2.4.3 Hawkes-processen

Het Hawkes-proces is een veelgebruikt model in de ruimtelijke statistiek. Het wordt ook wel een ‘self-exciting process’ genoemd, waarmee een belangrijke eigenschap van het Hawkes-proces naar voren komt: de eigenschap dat een gebeurtenis een positieve bijdrage levert aan de kans dat er later weer een gebeurtenis plaatsvindt. Nemen wij als voorbeeld aardbevingen, dan is het erg aannemelijk dat een aardbeving ook naschokken genereert. Dat wil dus zeggen dat de oorspronkelijke beving de kans vergroot dat er even later ook weer een beving plaatsvindt. Bij het Hawkes-proces is dus de intensiteit mede bepaald door het voorkomen van, of het gebrek aan, gebeurtenissen in het verleden. De intensiteit van een Hawkes-proces kan daarom opgedeeld worden in twee delen, te weten de achtergrondintensiteit en een interactieterm. De achtergrondintensiteit geeft weer wat de intensiteit zou zijn geweest als er geen gebeurtenissen waren in het verleden. De interactieterm beschrijft de bijdrage van de gebeurtenissen in het verleden aan de intensiteit. Zo kunnen we de voorwaardelijke intensiteit van een Hawkes-proces in ruimte en tijd als volgt uitdrukken:

$$\lambda(x, t | H_t) = \mu(x, t) + \sum_{i: t_i < t} \nu(|x - x_i|, t - t_i), \quad (4)$$

met $\mu(x, t)$ de achtergrondintensiteit en $\nu(|x - x_i|, t - t_i)$ de functie die de bijdrage weergeeft van de gebeurtenis op (x_i, t_i) aan de intensiteit op (x, t) . De bruikbare informatie over de gebeurtenissen uit het verleden wordt wiskundig beschreven door de filtratie H_t . Een filtratie is een familie van σ -algebra's, welke algebra's van verzamelingen zijn die gesloten zijn onder aftelbare vereniging. In Coculescu en Nikeghbali 2007 is meer te lezen over filtraties. Het Hawkes-proces is ook weer te geven als een Poisson-clusterproces maar vanwege de vele toepassingen worden deze processen als aparte klasse beschouwd.

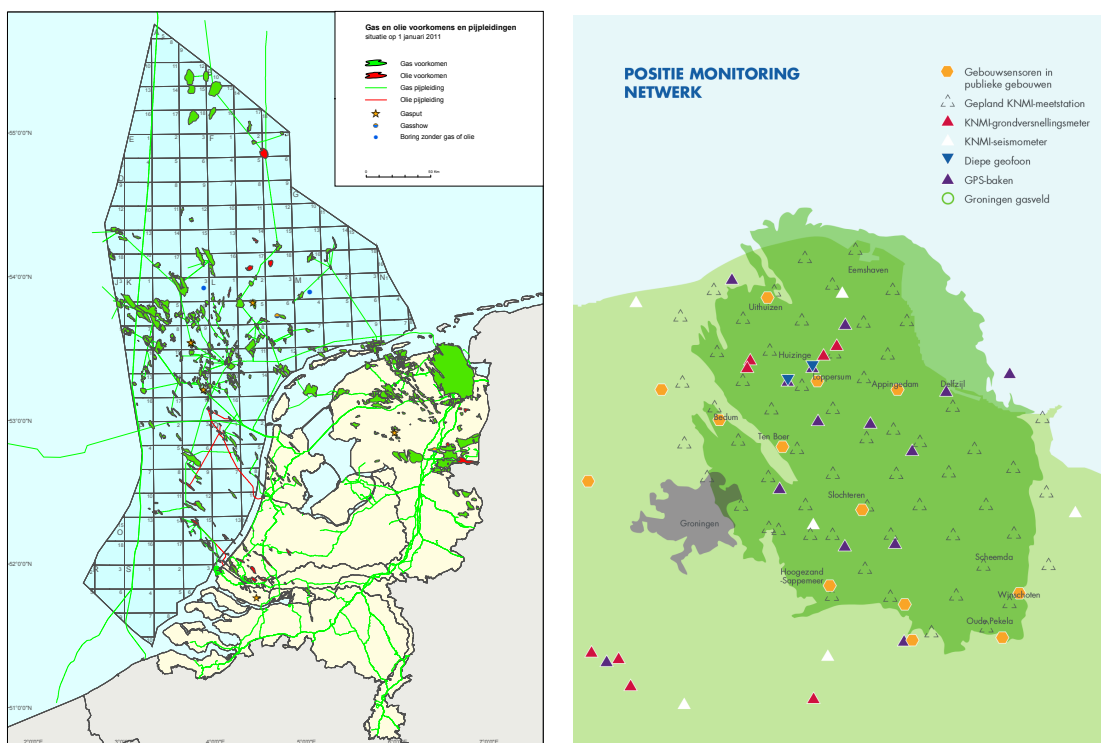
Het Hawkes-proces is oorspronkelijk een een-dimensionaal puntproces in de positieve tijd. Het werd echter al vrij snel uitgebreid naar een multivariaat proces, waardoor het Hawkes-proces vele toepassingen kent. Een belangrijke toepassing van Hawkes-processen is het modelleren van aardbevingen, waarbij het ETAS-model een bekend Hawkes-proces beschrijft (Ogata 2013). In Daley en Vere-Jones 2008 wordt het Hawkes-proces uitgebreid beschreven.

3 Geologische achtergrond

Om een compleet beeld te geven van de situatie, zullen we in dit hoofdstuk de geologische achtergrond bespreken van de geïnduceerde aardbevingen. Daarnaast geeft dit hoofdstuk een beeld van het onderzoek dat werd en wordt gedaan rondom de aardbevingen in Noord-Nederland. In paragraaf 3.1 wordt in het kort uitgelegd hoe seismiciteit ontstaat. In de hieropvolgende paragrafen zullen de tussenstappen verder worden uitgediept.

De meeste geïnduceerde bevingen in Nederland vinden plaats in Groningen, onder het grootste onshore gasveld van Europa. De NAM voert daarom een “kleine-veldenbeleid”. Dit houdt in dat uit het Groninger gasveld relatief weinig gas wordt gewonnen, terwijl de productie uit andere, kleinere, gasvelden wordt gemaximaliseerd. Ondanks dit beleid is de gasproductie in het Groninger gasveld vele malen hoger dan in de andere gasvelden. Ook in Drenthe, Friesland en Noord-Holland zijn er plaatsen waar seismiciteit wordt geïnduceerd, maar delfstofwinning leidt zeker niet in alle gevallen tot aardshokken (zie figuur 9 voor een totaaloverzicht van onder andere gasvoorkomens in Nederland).

De aardbevingen in Groningen zijn uniek omdat er vrij veel schade wordt veroorzaakt door bevingen met een relatief lage magnitude. Ook in andere landen worden er bevingen geïnduceerd door delfstofwinning. Het verschil met de situatie in Groningen is dat deze bijna altijd plaatsvinden in afgelegen gebieden, waardoor de schade beperkt blijft. Onderzoek naar geïnduceerde seismiciteit in Groningen is in volle gang. De belangrijkste onderzoeksgroepen zijn verbonden aan het Centraal Bureau voor de Statistiek (CBS), het Koninklijk Nederlands Meteorologisch Instituut (KNMI), de Nederlandse Aardolie Maatschappij (NAM), TNO-AGE en het Staatstoezicht op de Mijnen (SodM). Zij hebben een meetnetwerk opgetuigd om de oorzaken en gevolgen van bevingen te analyseren. In figuur 9 is een kaart van de belangrijkste sensoren getekend; data van aardbevingen zelf komt voornamelijk van de KNMI-meetstations.

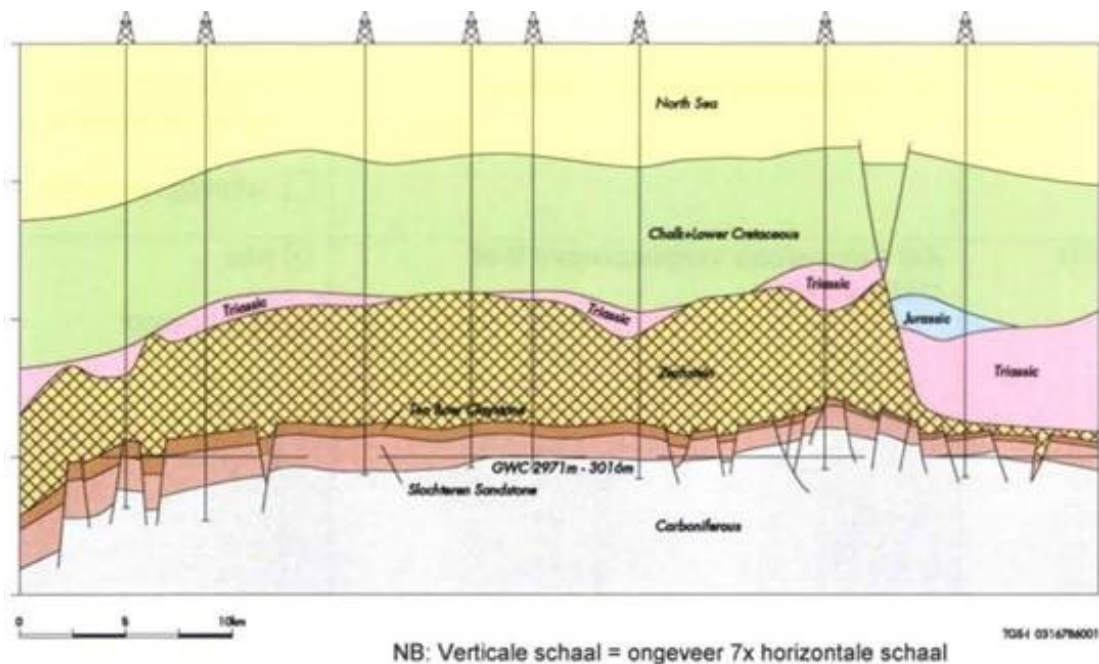


Figuur 9: Links: kaart van olie- en gasinfrastructuur van Nederland (uit 2012). Niet alle winning is in handen van de NAM - per gasveld is er een aanbestedingsprocedure geweest. Bron: www.aardgas-in-nederland.nl. Rechts: trilling- en hoogtemeetnetwerk met op de achtergrond de locatie van het Groningen gasveld. Bron: NAM platform (NAM 2015).

3.1 Oorzaak van seismiciteit

Het precieze ontstaan van aardbevingen door gaswinning is een vrij ingewikkeld proces. In eerste instantie zit het gas opgesloten in een zandsteenlaag (“het reservoir”) tussen niet-poreuze gesteentelagen. Na het slaan van een put kan er gas worden gewonnen uit de poriën van het reservoirgesteente. Hierdoor daalt de druk in het veld en ondervinden de gesteentelagen erboven minder weerstand. Dit heeft tot gevolg dat het reservoir inzakt - een proces dat bekend staat als compactie. In eerste instantie wordt de laag van reservoirgesteente enigszins samengedrukt. Bovendien zal aan de bovengrond (“het maaiveld”) de bodem in een breed gebied gaan dalen. Hierdoor kunnen er spanningen in de grond ontstaan, vooral daar waar er al breuken in de aardkorst bestonden. De spanning kan vrijkomen als seismische energie in een paar grote of een reeks kleine bevingen.

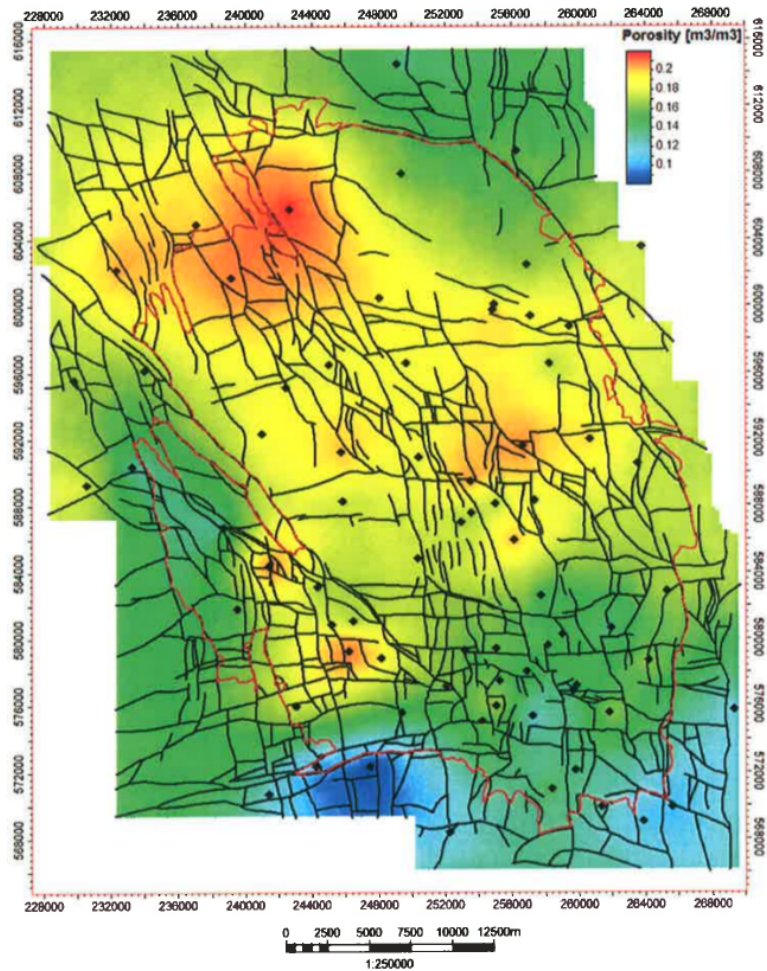
TNO en NAM modelleren al deze stapjes als losse onderdelen (zie o.a. TNO 2013) en gebruiken een eindige-elementenmethode om de gevolgen van bodemdaling te simuleren. Dat wil zeggen dat de bodem wordt opgedeeld in een aantal roostervlakjes, waarop de modellen één voor één worden toegepast. We zullen in de komende paragrafen de verschillende tussenstappen in het proces van gaswinning tot aardbeving uitgebreider behandelen.



Figuur 10: Dwarsdoorsnede van de bodem in Groningen. Onderaan de poreuze gaslaag (“Carboniferous”) en daarboven verschillende gesteentelagen. Bron: NAM winningsplan 2007

3.2 Model van de bodem

Bij de exploitatie van gasvelden in Nederland en het buitenland is gebleken dat elk reservoir anders reageert op gaswinning. De structuur van de ondergrond is dan ook allesbepalend voor de ontwikkeling van seismiciteit. Met behulp van een combinatie van verschillende meettechnieken en interpolatie heeft de NAM een volledig model van het Groningen gasveld gemaakt en ingevoerd in het computerprogramma Petrel. Dit driedimensionale model beschrijft onder andere de eigenschappen van de verschillende gesteentelagen die samen het reservoir vormen. Belangrijke steenlagen zijn het zandsteen waar gas en water in opgesloten zitten (dit heet de “Slochteren-formatie”). De dichtheid of porositeit van dit gesteente bepaalt hoe vrijelijk het gas en water zich kunnen verspreiden door de bodem. In de gesteentelagen zijn breuken aanwezig, waar spanning staat tussen losse gesteenteformaties.



Figuur 11: Geïnterpreteerde breuken en geschatte porositeit van de Slochteren (gas) gesteentelaag. Bron: TNO en Dost 2014, fig 2.19

Het Petrel model is gebaseerd op een groot aantal metingen, deels aan putten en deels met seismische meetapparatuur. In het Groningen gasveld zijn 365 putten geslagen. Het grootste gedeelte bevindt zich op één van 28 productielocaties, maar daarnaast zijn er ook nog een fors aantal observatieputten die puur bedoeld zijn voor het verkrijgen van meetdata. Hier kan met sondes in de putschacht en door analyse van opgegraven materiaal lokaal een goed beeld van de bodem verkregen worden. Om hierbuiten ook informatie te vergaren wordt gebruik gemaakt van seismiek. Dat is de analyse van trillingen in de bodem. Met bijvoorbeeld een hei-installatie worden schokgolven de grond ingestuurd en deze worden door verschillende steensoorten anders doorgegeven. Door met een seismograaf de reflectie van de golven te analyseren, kan er veel informatie over de bodem worden verkregen. Met deze techniek zijn de grenzen van het gasveld bepaald en ook de locatie van zo'n 1800 breuken, welke uiteindelijk door de NAM tot 707 breuklijnen ("geïnterpreteerde breuken") gereduceerd zijn (zie figuur 11). Omdat elke aardstok onstaat in een breuk zijn hiermee de mogelijke epicentra vastgelegd (hoewel sommige breuken te klein zijn om met seismiek te detecteren).

3.3 Gaswinning en druk in het reservoir

Het gas in de Slochteren-steenlaag staat onder druk doordat de bovenliggende grond het gesteente naar beneden drukt. De druk wordt voornamelijk bepaald door de hoeveelheid gas per volume, welke in beginsel alleen afhangt van de porositeit van het gesteente. Rond productieputten daalt de reservoirdruk echter door gaswinning. Door de winning uit verschillende productieclusters te variëren, kan de exploi-

tant van een gasveld de druk over grote delen van het reservoir vrij goed regelen. Meestal wordt een constante drukverlaging door het hele veld heen nagestreefd. Wanneer uiteindelijk overal de einddruk is bereikt, is het veld uitgeproduceerd.

Een verandering in de gaswinning heeft ook veranderingen in de mate van drukverlaging tot gevolg. Deze veranderingen verspreiden zich als een schokgolf door de gaslaag. Het golffront heeft een lage snelheid: in 12 maanden zal, afhankelijk van de porositeit van het gesteente, maximaal 7 kilometer zijn afgelegd. Verminderde bodemdaling, als effect van een productieverlaging, is waarschijnlijk in een straal van 3 kilometer rondom een boorput meetbaar. (Resultaten uit paragraaf 2.2 en figuur 2.2 van TNO en Dost 2014 - een uitgebreide afleiding is te vinden in appendix A van dat rapport).

3.4 Veranderingen in het reservoir: compactie

Als gevolg van drukverlaging wordt het reservoirgesteente samengedrukt. De mate van compactie door een bepaalde drukdaling kan per plaats verschillen, afhankelijk van de porositeit van het reservoirgesteente. Ook de aanwezigheid van waterhoudende lagen (“aquifers”) onder de gaslaag beïnvloedt de samendrukbaarheid. In boorputten kan men de compactie meten en met laboratoriumexperimenten is de porositeit van het gesteente te bepalen. De relatie tussen die grootheden is uitgebreid onderzocht en er zijn een aantal compactiemodellen in omloop, welke in TNO 2013 worden beschreven en vergeleken. Een simpel bilineair model werd gebruikt door de NAM in NAM 2010. Daarin is de compactie lineair in de drukdaling, met een omslagpunt na het bereiken van een bepaalde transitiedruk:

$$\Delta H = \begin{cases} C_{m,init} H_0 \Delta P & P \leq P_{trans} \\ C_{m,end} H_0 \Delta P & P > P_{trans} \end{cases} \quad (5)$$

Hierin is P_{trans} de transitiedruk, ΔH het verschil in hoogte, H_0 de initiële hoogte en C de compactiecoëfficiënt. In woorden zegt dit model: de verandering van hoogte van het reservoir is gelijk aan de beginhoogte maal de drukverlaging (sinds het begin) maal een compactiecoëfficiënt C . Deze C is alleen afhankelijk van de porositeit van het reservoirgesteente.

Het blijkt echter dat de bodem niet direct reageert op drukverschillen en er een doorgaande daling gaat ontstaan die minder afhankelijk is van de huidige druk. De vertraagde, doorgaande compactie die hierdoor ontstaat wordt “kruip” genoemd.

De NAM heeft een nieuw model ontwikkeld om deze kruip te beschrijven: het zogenaamde Time Decay model. Er wordt nu weer uitgegaan van een lineair verband tussen compactie en drukverlaging, maar de kruip is meegenomen door middel van een tijdsconvolutie:

$$\Delta H = (C_m \Delta P) *_t \left(\frac{1}{\tau} e^{-\frac{t}{\tau}} \right) \quad (6)$$

waarbij τ de tijdsvertragingconstante is waar het model zijn naam aan dankt. Deze is bepaald op ongeveer 7 jaar, waardoor een verandering van gasproductie wordt uitgesmeerd over een periode van 5 tot 15 jaar (voor meer informatie zie TNO 2013).

TNO gebruikt een totaal ander compactiemodel, genaamd Rate Type Compaction Model (RTCM), beschreven in TNO 2013. Het RTCM onderscheidt aparte componenten van compactie: de directe daling, puur afhankelijk van de drukverschillen, en de seculiere daling, afhankelijk van drukverschil en totale daling tot nu toe. De seculiere component reageert vertraagd op drukverschillen en modelleert zo de kruip. We leggen hier niet het totale model uit. Het belangrijkste verschil tussen het RTCM en het Time Decay model is dat het Time Decay model veel minder gevoelig is voor vermindering van gaswinning.

3.5 Gevolgen van compactie: bodemdaling en seismiciteit

Door compactie wordt er energie opgeslagen in de bodem in de vorm van spanning. Het grootste gedeelte wordt omgezet in instantane compactie en daarmee bodemdaling. Een deel kan echter vrijkomen in de vorm van seismiciteit bij breuken in de aardkorst. De hoeveelheid seismische energie die vrijkomt als deel van de totale energie ondergronds kan empirisch worden bepaald. Deze hoeveelheid wordt weergegeven met de partiticoëfficiënt α'_f , welke de grootste bron is van onzekerheid in huidige onderzoeken (zie onder andere de paragrafen “partiticoëfficiënt” en “onzekerheden” in TNO en Thienen-Visser 2014). Het lijkt duidelijk dat de partiticoëfficiënt afhangt van het aantal breuken in de aardkorst. Immers: hoe meer breuken, hoe groter de kans dat er één gaat trillen. Bovendien blijkt dat de partiticoëfficiënt afhankelijk is van de compactie: hoe meer spanning er op breuken komt te staan, hoe groter het gedeelte van die spanning dat in aardbevingen wordt omgezet. Verder moet gelden dat $\alpha'_f \geq 0$ en dus kiest de NAM voor een exponentiële relatie:

$$\alpha'_f = e^{f+g\Delta H}. \quad (7)$$

waarbij f en g onbekende parameters zijn, welke kunnen worden bepaald door metingen van α'_f . In het centrum van het Groninger gasveld is α'_f ongeveer 10^{-3} . TNO heeft in TNO 2013 laten zien dat dit een redelijk model is, maar stelt uit fysische overwegingen dat de partiticoëfficiënt begrensd zou moeten zijn tot maximaal 1%. In het exponentiële model is de partiticoëfficiënt echter onbegrensd. Deze zou zelfs groter dan 100% kunnen worden.

Seismische energie kan op meerdere manieren uit de bodem vrijkomen: in een paar hele grote bevingen of in meerdere kleine bevingen. Seismisch Hazard Analyse richt zich op het berekenen van de kans dat bepaalde grondversnellingsniveaus overschreden worden. Dit gebeurt met een Ground Motion Prediction Equation. Op basis hiervan kan ook het seismisch risico worden bepaald (kans op bepaalde schadeniveaus). De trillingsenergie van een beving wordt meestal uitgedrukt in een magnitude op de schaal van Richter. Met behulp van de Gutenberg-Richter vergelijking kan de verdeling van bevingsmagnitudes ook worden bepaald. Het blijkt namelijk dat de magnitudes zich ongeveer als volgt verdelen: het aantal bevingen N met magnitude tenminste M is:

$$N(M) = 10^{a-bM}. \quad (8)$$

Hierbij zijn a en b parameters - b is in de meeste gevallen (en ook in het geval van het Groninger gasveld) ongeveer 1 en a correspondeert met de totale seismiciteit van de beschouwde regio.

Op de Gutenberg-Richter vergelijking valt nog het nodige aan te merken. Ten eerste moet worden aangenomen dat de seismiciteit stationair is, anders is a een functie van de tijd en verdwijnt het analytisch bepaalbare gedrag van de magnitudeverdeling. Ook neemt de vergelijking bevingen mee die niet meetbaar zijn vanwege een te lage magnitude. Om dit op te lossen moet de minimaal detecteerbare magnitude M_0 bepaald en ingevuld worden volgens $N(M) = 10^{a-b(M-M_0)}$. Daarnaast geeft de formule een positieve kans op bevingen van alle magnitudes, ook zeer zware bevingen die fysisch onmogelijk zijn. Een aangepaste methode (Bounded Gutenberg-Richter) begint bij een vooraf bepaalde magnitude M_0 en vlakt af bij M_{max} :

$$N(M) = e^{a-bM_0} \frac{e^{b(M-M_0)} - e^{b(M_{max}-M_0)}}{1 - e^{b(M_{max}-M_0)}}. \quad (9)$$

In aangepaste vorm heet deze vergelijking ook wel de Cornell-Vandermarke verdeling. Het blijkt moeilijk te zijn om voor Groningen een goede waarde van M_{max} te vinden: de rapporten TNO 2013 en Pijpers 2014 verschillen daarin van mening.

3.6 Naschokken

De formule voor de partiticoëfficiënt in (7) laat het feit dat aardbevingen naschokken kunnen veroorzaken totaal buiten beschouwing. Het kan echter zo zijn dat de spanning in het veld alleen maar verplaatst of zelfs vergroot is door een beving, waardoor er meer bevingen volgen. Dit kan zelfs een hele keten van naschokken tot gevolg hebben. Deze hoeven niet direct plaats te vinden. In Pijpers 2014) wordt de aanwezigheid van naschokken onderkend maar geen van de huidige onderzoeken neemt deze mee in hun model. Het effect van naschokken is namelijk niet zo groot, omdat deze vaak óf een hele kleine

magnitude hebben óf net zo goed gezien kunnen worden als gewone bevingen. Er bestaat wel een breed gedragen model voor naschokken, dat bekend staat als ETAS. Dit brengt ons weer terug naar de theorie van puntpatronen.

Het ETAS-model, dat staat voor Epidemic Type Aftershock Sequence-model, is een bijzondere vorm van het Hawkes-proces (zie paragraaf 2.4.3). Het werd geïntroduceerd door Ogata in 1988 (Zhuang, Ogata en Vere-Jones 2002) en is gebaseerd op de wetten van Gutenberg-Richter en Omori-Utsu. Dit model is vooral geschreven voor tectonische seismiteit, maar zou ook voor geïnduceerde bevingen kunnen werken.

ETAS modelleert hoe bevingen naschokken veroorzaken, door elke aardshok als mogelijke “ouder” van een aantal “kinder”-bevingen te zien in het Hawkes proces. Het verloop van het proces is dus afhankelijk van eerdere gebeurtenissen, ofwel het verleden $H_t = \{(x_i, t_i, m_i) | \forall i : t_i \leq t\}$.

Dit is uit te drukken als een intensiteitsfunctie op tijdstip t , afhankelijk van H_t . H_t is eigenlijk de realisatie tot dan toe van het kansproces gedreven door de intensiteitsfunctie. De intensiteit is dus in zekere zin afhankelijk van zichzelf en is niet direct uit te drukken. De intensiteit is alleen geconditioneerd op het verleden van het proces. De voorwaardelijke intensiteit drukt dan de invloed van het verleden op de intensiteit in het heden uit. Tijdstip, locatie en magnitude van een ouder-beving worden gebruikt om de kans verschillende aantallen nakomelingen in de buurt (in ruimte en tijd) te bepalen. De uiteindelijke intensiteitsfunctie, geconditioneerd op de observaties tot tijdstip t , wordt gegeven door:

$$\lambda(x, t | H_t) = \mu(x) + \sum_{i: t < t_i} k(m_i) \cdot g(t - t_i) \cdot f(|x - x_i| | m_i). \quad (10)$$

De functie $\mu(x)$ geeft de achtergrondintensiteit weer, welke in het ETAS-model stationair verondersteld wordt. Voor tectonische bevingen is dit redelijk, maar voor geïnduceerde is dit niet zomaar gemakkelijk aan te nemen. De functie $k(m_i)$ geeft aan hoeveel naschokken een beving naar verwachting produceert en hiervoor wordt de (Bounded) Gutenberg-Richter vergelijking uit (8) (of (9)) gebruikt. Daar komt ook de stationariteitsaannname vandaan. De Omori-Utsu wet komt terug in de functie g en geeft een kansverdeling voor de tijd tussen ouderbeving en naschok. Meestal is dit een inverse machtsdistributie:

$$g(t) = (p - 1)c^{p-1}(t + c)^{-p} \quad \text{als } t > 0. \quad (11)$$

Hierbij moeten de parameters c en p worden gefit. Als laatste geeft de functie f een kansverdeling voor de afstand tussen beving en naschok (ofwel de spreiding van naschokken) en hiervoor bestaan verschillende uitdrukkingen. In zekere zin is het een kwestie van smoothen over de ruimte. Er wordt dan ook vaak een Gaussische kern gebruikt (zo ook in Zhuang, Ogata en Vere-Jones 2002):

$$f(x, y) = \frac{1}{2\pi d \frac{k(m_i)}{k(M_0)}} \exp\left(-\frac{x^2 + y^2}{2d \frac{k(m_i)}{k(M_0)}}\right). \quad (12)$$

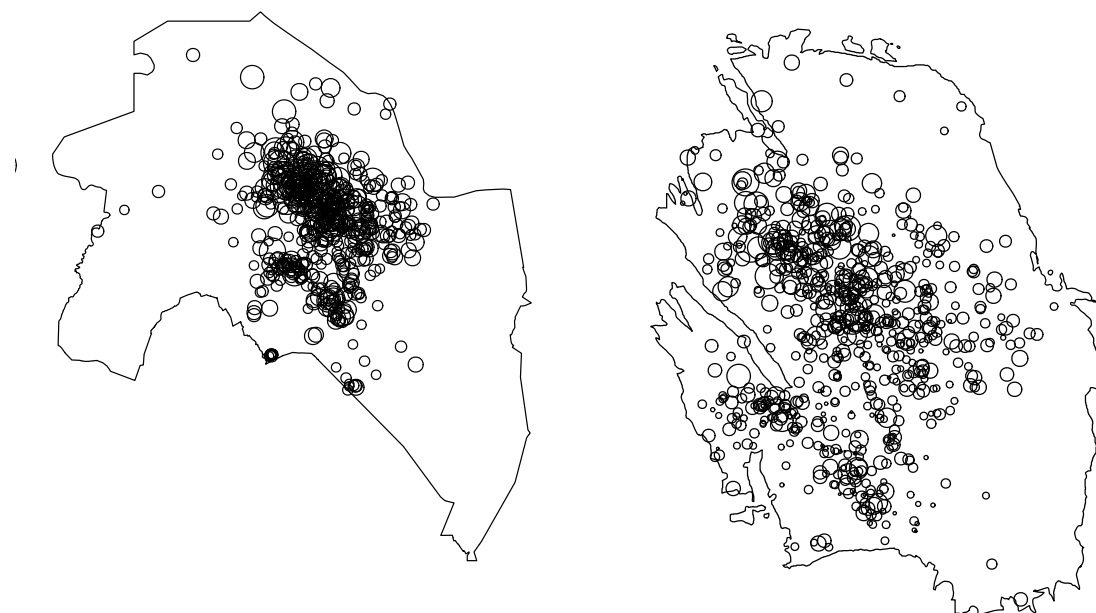
In deze formulering heeft het model acht parameters die de verdeling van naschokken beschrijven. Wat overblijft is een onbekende functie μ , de achtergrondintensiteit. In Zhuang, Ogata en Vere-Jones 2002 wordt een methode genaamd “Stochastic Declustering” geïntroduceerd die deze parameters en de achtergrondintensiteit kan schatten. Voor elke beving wordt iteratief de kans bepaald dat het een achtergrondbeving of naschok is. Zo kan uiteindelijk een meest aannemelijke schatting van de parameters bepaald worden en daarmee een achtergrond- en clusteringintensiteit per punt (in tijd en ruimte).

4 Exploratieve analyse

In dit hoofdstuk beschouwen we data over de aardbevingen en gaswinning. We onderzoeken welke verbanden en patronen er te ontdekken zijn. Daarnaast zullen we ook kijken naar welke verbanden niet aantoonbaar zijn. In paragraaf 4.1 zullen we de data van het KNMI⁶ beschouwen die beschrijft waar en wanneer de aardbevingen in Noord-Nederland hebben plaatsgevonden. De locaties van aardbevingen in het vlak beschouwen we als puntpatronen. In paragraaf 2.1 hebben we de drie typen puntpatronen beschreven: regulier verdeeld, willekeurig verdeeld en geclusterd. Met behulp van K-functies, beschreven in paragrafen 2.2.3 en 2.2.4, zullen we onderzoeken welk type puntpatroon van toepassing is op de aardbevingen in Noord-Nederland.

In paragraaf 4.2 zullen we zowel de data over het aantal jaarlijkse aardbevingen als de data over de gaswinning analyseren. Is er sprake van een toenemend aantal aardbevingen? We zullen onderzoek doen naar eventuele verbanden tussen gaswinning en aantal gemeten bevingen. In paragraaf 4.3 zullen we ook de magnitudes bestuderen. Zien we een stijging in de gemiddelde of grootst gemeten magnitude? We zullen verbanden tussen gaswinning en de magnitudes van de bevingen analyseren.

Aan de hand van de exploratieve analyse kunnen we beslissen hoe het model vorm gaat krijgen.



Figuur 12: Gemeten aardbevingen vanaf 1995. De grootte van de cirkels geeft de grootte van de magnitude weer. Links: provincie, rechts: gasveld.

4.1 Locaties aardbevingen

In deze paragraaf zullen we de data van het KNMI beschouwen die beschrijft wanneer in het verleden op welke locaties aardbevingen hebben plaatsgevonden. We zullen in deze paragraaf de aardbevingen in de ruimte weergeven als een puntpatroon (zie paragraaf 2.1). In figuur 12 zijn de aardbevingen vanaf 1995 als cirkels in de ruimte weergegeven. We zien dat vooral in het centrum van de provincie Groningen, in de omgeving van Loppersum, veel bevingen gemeten zijn. Het overgrote deel van deze bevingen valt binnen het gasveld Groningen. Vandaar dat we in deze paragraaf enkel de bevingen in het gasveld Groningen beschouwen. Een blik op figuur 12 doet vermoeden dat er ook binnen het gasveld sprake is van clustering. Maar kan dit ook onderbouwd worden?

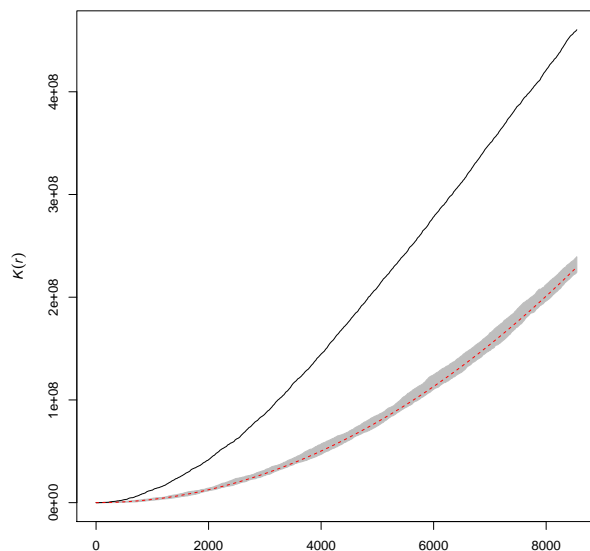
⁶Bron: www.KNMI.nl, KNMI 2015

In paragraaf 2.2 zijn methoden beschreven om te toetsen of er sprake is van een willekeurige verdeling over het vlak. In paragraaf 2.2.1 wordt de kwadrantmethode beschreven. Bij deze methode wordt het totale gebied R opgedeeld in m congruente deelgebieden. De toetsingsgrootte die wordt gebruikt is

$$T = \sum_{i=1}^m \frac{n_i - n/m}{n/m}$$

met n_i de waargenomen waarde van stochast N_i , het aantal punten in deelgebied C_i . Onder de nulhypothese van willekeurige verdeling heeft toetsingsgrootte T een χ^2 -verdeling met $m - 1$ vrijheidsgraden. Kleine waarden van T duiden op een reguliere verdeling en grote waarden van T op clustering. Wanneer we deze toets toepassen op het gasveld Groningen met $m = 86$ (we delen het gebied op in 86 deelgebieden) vinden we een toetsingsgrootte $T = 4,6 \cdot 10^2$. Dit levert een p-waarde op van $1 \cdot 10^{-3}$. De uitkomst van de toets correspondeert met onze verwachtingen: we kunnen de nulhypothese van de willekeurige verdeling verwerpen voor het gasveld Groningen.

In paragraaf 2.2.1 werd het al beschreven: de kwadrantmethode heeft ook nadelen. Er is geen eenduidige ‘beste’ manier om het totale gebied, in ons geval de provincie of het gasveld, te verdelen in subgebieden. Omdat wij te maken hebben met grillige gasveldgrenzen is het niet mogelijk om enkel gelijkvormige deelgebieden te construeren. Daarnaast worden de relaties tussen punten niet meegenomen met deze methode. Daarom zullen we ook Ripley’s K-functie gebruiken om het puntpatroon te analyseren. Kort gezegd: de K-functie is een functie van straal r en geeft het aantal punten aan binnen een afstand r van een willekeurig punt. In paragrafen 2.2.3 en 2.2.4 is meer over deze functie te lezen.



Figuur 13: Homogene K-functie, aardbevingen gasveld Groningen 1995-2014.

We beginnen met het toepassen van de homogene K-functie, hierna zullen we ook de inhomogene K-functie toepassen. De intensiteitsschattingen zullen we op verschillende manieren bepalen.

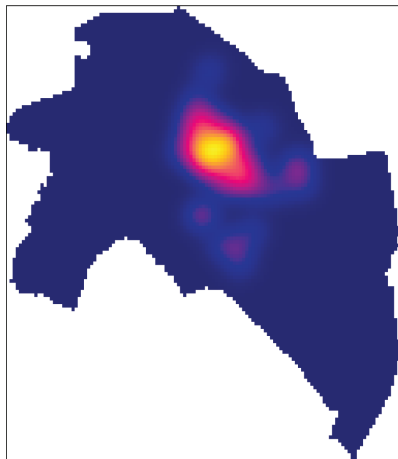
De komende figuren zijn allemaal in een vorm vergelijkbaar met figuur 13. Wat is er te zien in de figuur? We zien een grijze band met daarin een rode lijn. De rode lijn geeft het verwachte verloop van de K-functie aan, onder de nulhypothese dat het puntpatroon beschreven kan worden als een binomiaal proces met de gekozen intensiteit. Deze gekozen intensiteit zullen we variëren. De bandbreedte van de grijze band is bepaald door het doen van simulaties onder de nulhypothese. Er zijn steeds 99 simulaties uitgevoerd waarbij een vast aantal punten wordt verdeeld over de ruimte. Dit aantal punten is precies het aantal punten dat te vinden is in het puntpatroon. Na discretisatie is voor elke waarde van r de minimale en de maximale waarde bepaald van de functie $K(r)$. Deze waarden vormen de bovenkant

en de onderkant van de grijze band. De zwarte lijn in het figuur laat het waargenomen verloop van de functie $K(r)$ zien. De assen zijn gebaseerd op afstanden in meters.

Om de nulhypothese te toetsen wordt van te voren een waarde van r gekozen. Wanneer voor deze waarde van r de zwarte lijn buiten het grijze gebied ligt, wordt de nulhypothese verworpen. Dit gebeurt dan met significantie $\alpha = 0,02$.

In figuur 13 wordt uitgegaan van de homogene K-functie, oftewel: de intensiteit onder de nulhypothese is constant over de ruimte. We zien dat het verloop van de K-functie duidelijk buiten de grijze band valt. Uit deze toets blijkt dus dat we de nulhypothese dat de bevingen willekeurig zijn verdeeld over de ruimte kunnen verwerpen. Dit correspondeert met het resultaat uit de kwadrantmethode.

Omdat uit voorgaande toetsen bleek dat een Poissonverdeling met homogene intensiteit geen goed model is van de situatie in gasveld Groningen, zullen we nu een inhomogene intensiteit gebruiken zoals uitgelegd in paragraaf 2.2.4. Op deze manier onderzoeken we of schijnbare clustering te wijten is aan inhomogeniteit. Op basis van de data over de bevingen die plaatsvonden tussen 1995 en 2004 bepalen we een schatting voor de intensiteit. Hierbij wordt een Gaussische kern gebruikt met een gekozen bandbreedte σ . Hier is σ gekozen zoals Diggle en Berman dat doen in Berman en Diggle 1989. Ter illustratie is in figuur 14 de intensiteitsschatting weergegeven voor de provincie Groningen.

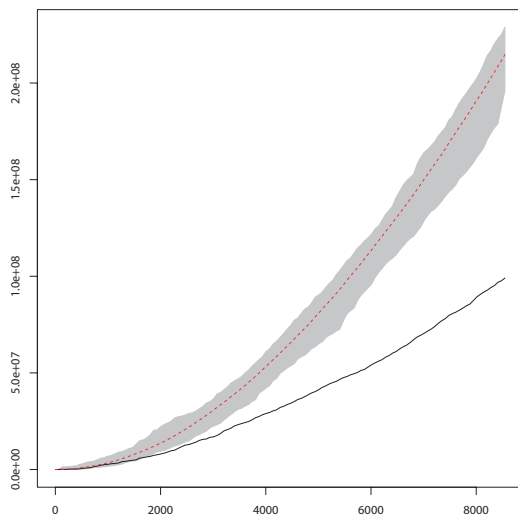


Figuur 14: Voorbeeld van een intensiteitsschatting, geschat op basis van periode 1995-2004. Hoe lichter de kleur, hoe hoger de intensiteit.

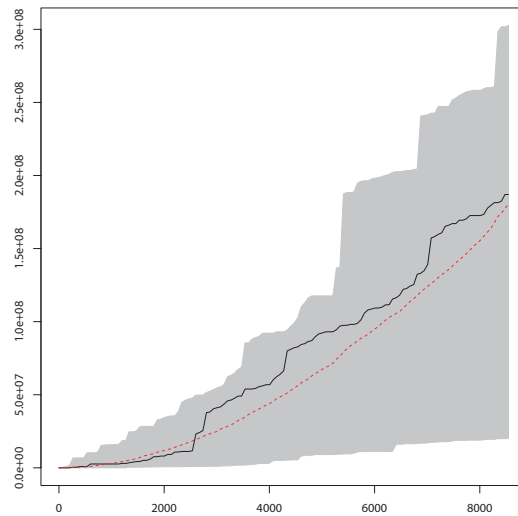
We zien in 15a dat de K-functie totaal buiten de grijze band valt. Dit kan twee verschillende oorzaken hebben: de intensiteit is niet stationair of er is sprake van clustering bovenop de inhomogene intensiteit. Om dit verder te onderzoeken zullen we de inhomogene intensiteit bepalen op basis van een kortere periode, namelijk 1995 tot en met 1999. Met deze intensiteitsfunctie voeren we steeds een toets uit over een periode van 5 jaar. Dit helpt ons om te onderzoeken hoe het resultaat van figuur 15a te verklaren is. Achtereenvolgens toetsen we over de periodes: 2000-2004, 2005-2009 en 2010-2014. In figuur 15 is het resultaat van deze toetsen te zien.

Opvallend aan de uitkomst van deze toetsen is de uitkomst uit figuur 15d waar de K-functie ver buiten de grijze band valt. Hoe is dit te verklaren? Is de periode 2010-2014 inderdaad erg afwijkend van de periode 1995-2000? Om te controleren of dit het geval is voeren we nog een aantal andere toetsen uit. We gebruiken de inhomogene F-functie, inhomogene G-functie en inhomogene J-functie. Deze functies zullen wij niet nader toelichten, omdat de strekking lijkt op die van de inhomogene Ripley's K-functie. In het overzichtsartikel van Szmyt 2014 is meer te lezen over deze functies. De uitkomsten van het toetsen met de inhomogene F-, G- en J-functie staan in figuur 37 in bijlage 10.5. In alle drie de gevallen zien we dat de functie buiten de grijze band valt. Hierdoor vermoeden we dat er inderdaad sprake is van non-stationariteit. Om dit te valideren hebben we de periodes 1995-1999 en 2010-2014 getoetst met een intensiteit die gebaseerd is op basis van de gemeten bevingen in diezelfde periodes. Deze resultaten zijn

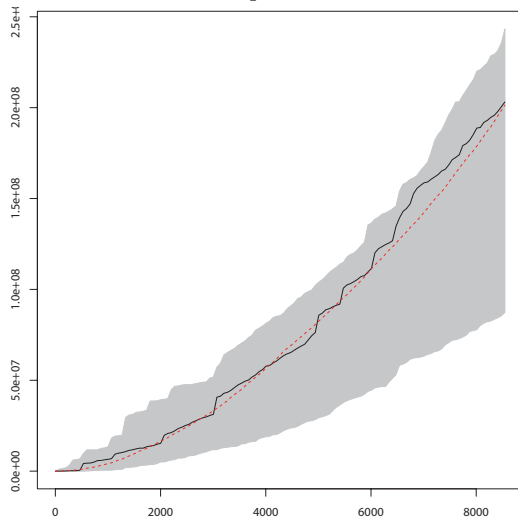
te vinden in figuur 38 in de bijlage 10.5. We zien dat de observaties ruim binnen de grijze banden liggen. We concluderen dat het opvallende resultaat te wijten is aan de non-stationairiteit van de intensiteit over de tijd.



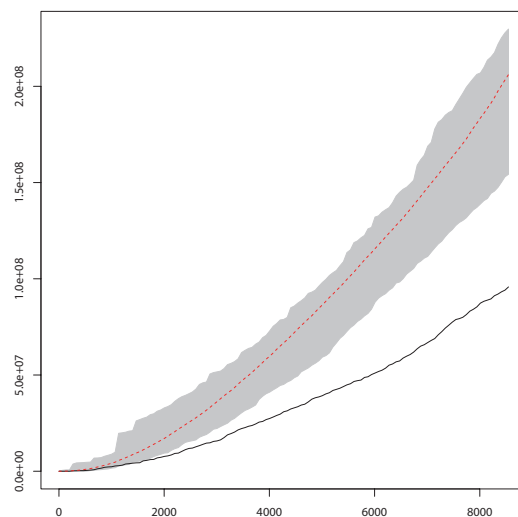
(a) Inhomogene K-functie, gasveld Groningen. Intensiteit geschat op basis van periode 1995-2004. Getoetst over periode 2005-2014.



(b) Gasveld Groningen, getoetst over periode 2000-2004.



(c) Gasveld Groningen, getoetst over periode 2005-2009.



(d) Gasveld Groningen, getoetst over periode 2010-2014.

Figuur 15: Inhomogene K-functie. Intensiteit geschat op basis van periode 1995-1999.

4.2 Het aantal aardbevingen en gaswinning

In deze paragraaf zullen we de data van het KNMI⁷ over de aardbevingen in Noord-Nederland verder onderzoeken. We zullen kijken naar het jaarlijkse aantal aardbevingen. Dit brengen we in verband met data over de gaswinning in Noord-Nederland afkomstig van TNO⁸. Zijn er verbanden te ontdekken tussen de gaswinning en het aantal aardbevingen?

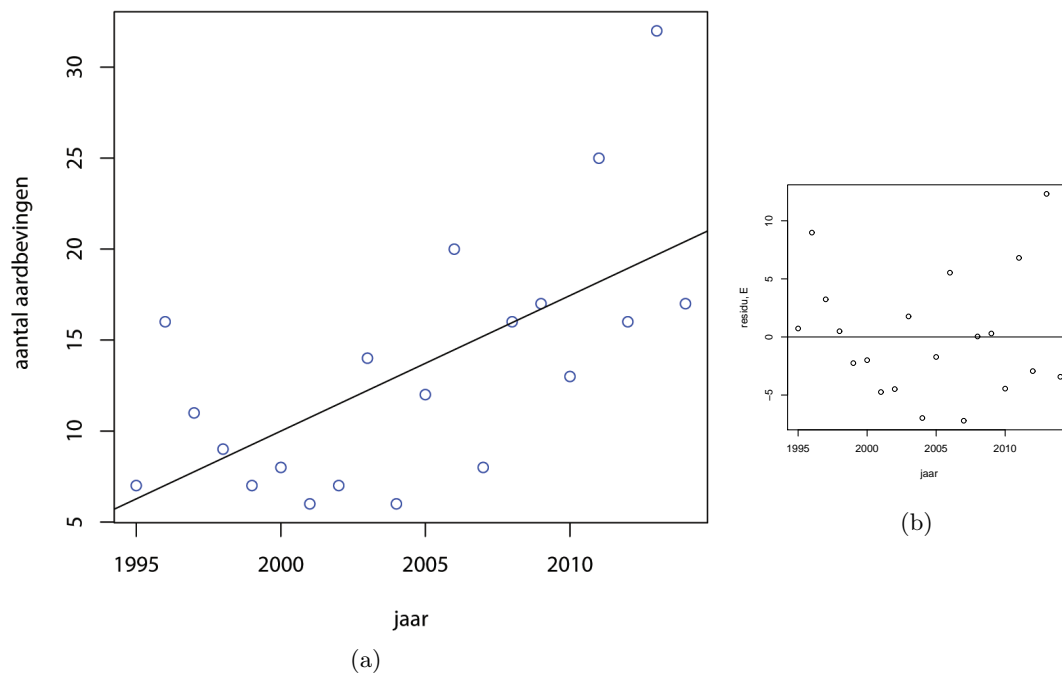
In deze paragraaf introduceren we de data over de gaswinning. Deze data zullen we ook in de volgende paragraaf, over het verband tussen gaswinning en de magnitudes, gebruiken.

4.2.1 Jaarlijkse aantal aardbevingen

We bekijken eerst het jaarlijkse aantal waargenomen aardbevingen. In figuur 16a zien we deze aantallen geplot in de tijd. Met behulp van enkelvoudige lineaire regressie (zie Barnett en Dobson 2008) is er een trendlijn doorheen geplot. Het model is van de vorm:

$$Y_i = \beta_0 + \beta_1 x_{i,1} + \epsilon_i$$

met Y_i het aantal aardbevingen, $x_{i,1}$ het jaartal en ϵ_i de fout met een verdeling $\epsilon_i \sim N(0, \sigma^2)$, onderling onafhankelijk. We vinden de volgende schatters: $\hat{\beta}_0 = -1480,23$ en $\hat{\beta}_1 = 0,7451$. De F-toets om te toetsen of $\beta_1 = 0$, geeft een p-waarde van $2,154 \cdot 10^{-3}$. Dit resultaat en het feit dat $\hat{\beta}_1 > 0$ duidt op een stijging van het aantal aardbevingen in de tijd.



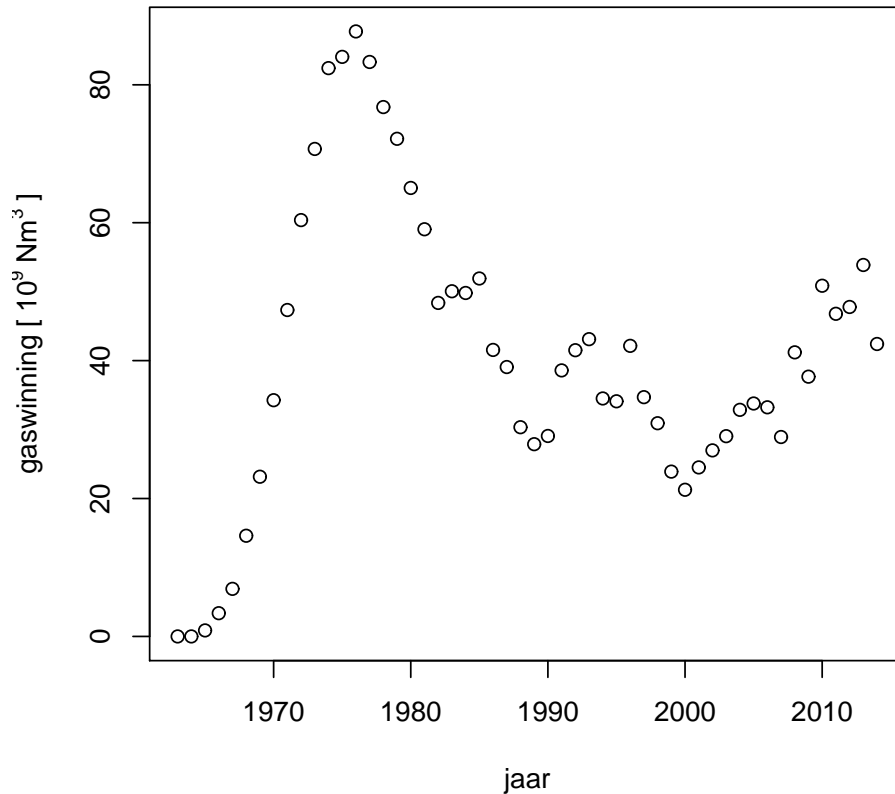
Figuur 16: Blauw: jaarlijkse aantal gemeten aardbevingen, zwart: trendlijn, 1995-2014. Figuur (b) geeft de residuenplot.

4.2.2 Jaarlijkse gaswinning

We bekijken de hoeveelheid gaswinning door de jaren heen. Deze hoeveelheid gaswinning is uitgedrukt in Nm^3 , normaal kubieke meter. We spreken hier in eenheden van $10^9 Nm^3$, miljarden normaal kubieke meter. Vanaf 1963 wordt er gas gewonnen in Groningen. De jaarlijkse cijfers zijn te zien in figuur 17 (bron: TNO). We zien dat er vooral eind jaren '70 veel gas is gewonnen.

⁷bron: www.KNMI.nl

⁸bron: www.nlog.nl



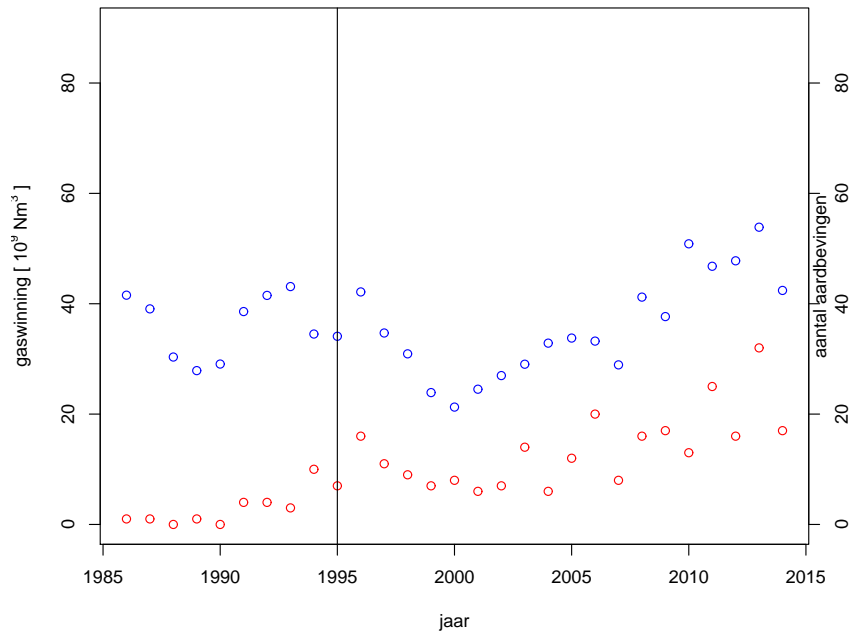
Figuur 17: Totale jaarlijkse gaswinning Groningen 1963-2014.

We zijn geïnteresseerd in het verband tussen de jaarlijkse gaswinning en het aantal aardbevingen per jaar. Vanaf 1986 zijn er gegevens beschikbaar over de gemeten aardbevingen. Echter, pas vanaf 1995 zijn deze gegevens betrouwbaar. Voor 1995 was er namelijk geen goede apparatuur, waardoor niet alle bevingen werden gedetecteerd of niet juist werden gelokaliseerd (Pijpers 2014). Wanneer we het aantal aardbevingen en de hoeveelheid gaswinning per jaar in één figuur plotten (zie figuur 18), lijken we een duidelijke samenhang te zien.

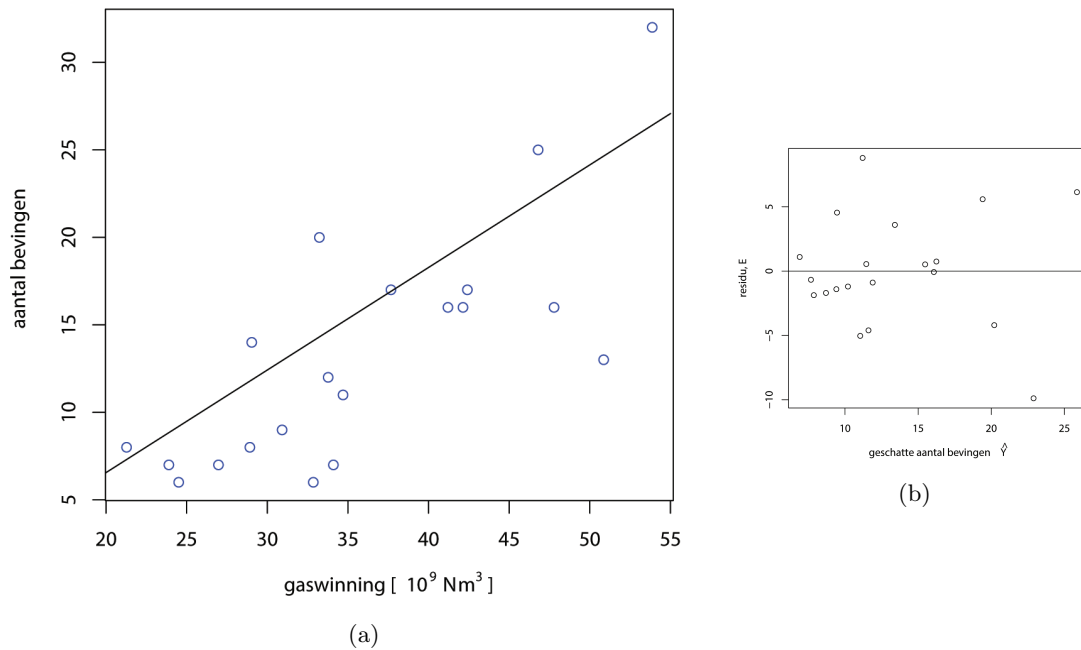
Kunnen we deze samenhang ook onderbouwen? In figuur 19a is het jaarlijkse aantal gemeten bevingen geplotted tegen de jaarlijkse gaswinning. Met behulp van enkelvoudige regressie is er een lijn doorheen geplotted die het verband tussen het aantal aardbevingen en de gaswinning schat. Het model is van de vorm (zie ook Barnett en Dobson 2008):

$$\log(Y_i) = \beta_0 + \beta_1 x_{i,1} + \epsilon_i$$

met Y_i het aantal aardbevingen, $x_{i,1}$ de jaarlijkse gaswinning en ϵ_i de fout met een verdeling $\epsilon_i \sim N(0, \sigma^2)$, onderling onafhankelijk. Aangezien we ervan uitgaan dat de variabelen Y_i een Poisson-verdeling hebben, maken we gebruik van Poisson-regressie, ook wel log-lineaire regressie genoemd. We vinden de volgende schatters: $\hat{\beta}_0 = 1,070$ en $\hat{\beta}_1 = 0,0405$. De F-toets om te toetsen of er sprake is van een verband, geeft een p-waarde van $7,43 \cdot 10^{-10}$. Dit duidt op een aantoonbaar verband tussen de gaswinning en het aantal aardbevingen. Ter controle van het model bekijken we de residuenplot (figuur 19b).



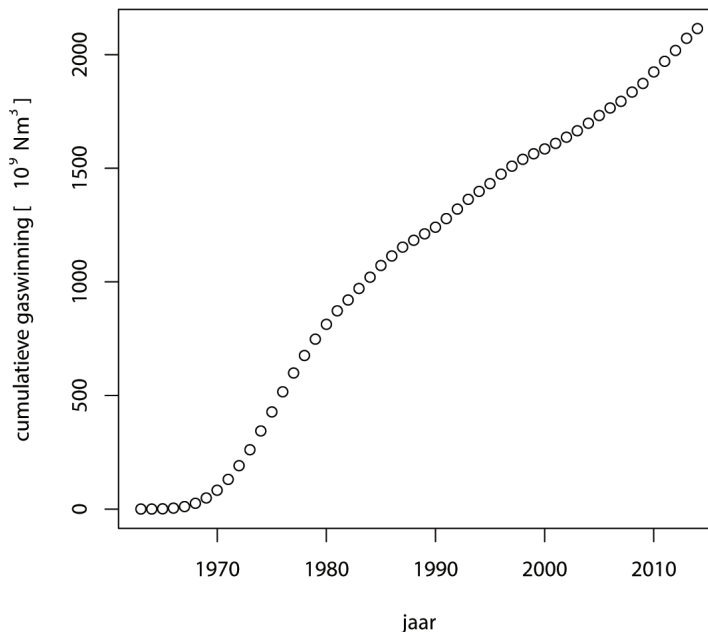
Figuur 18: Blauw: totale jaarlijkse gaswinning Groningen, rood: aantal bevingen. Voor 1995 is de data over het gemeten aantal aardbevingen niet betrouwbaar.



Figuur 19: Verband tussen het jaarlijkse aantal gemeten bevingen en de jaarlijkse gaswinning. Data aardbevingen uit 1995-2014. Figuur (a) geeft de meetwaarden in blauw en regressielijn in zwart. Figuur (b) geeft de residuenplot met $E = Y - \hat{Y}$, $Y :=$ aantal aardbevingen per jaar.

4.2.3 Cumulatieve gaswinning

Zou er behalve het verband met de jaarlijkse gaswinning, ook een verband zijn met de cumulatieve gaswinning? Onder de cumulatieve gaswinning verstaan we het totaal aantal kubieke meter gas dat gewonnen is sinds 1963. In figuur 20 is de cumulatieve gaswinning te zien.

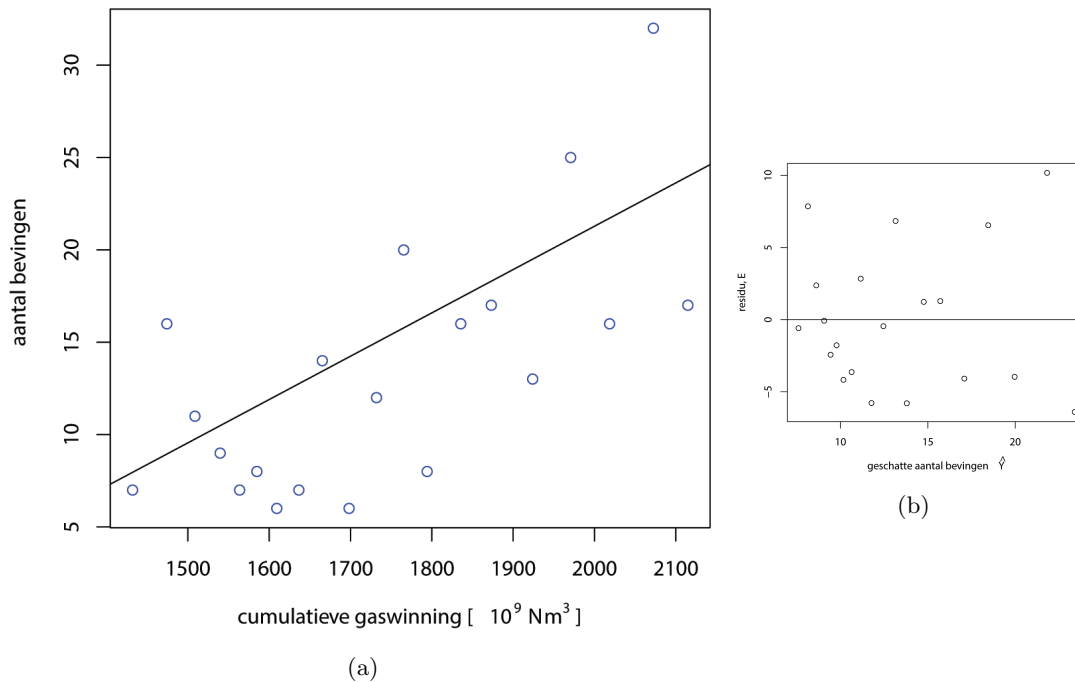


Figuur 20: Cumulatieve gaswinning Groningen 1963-2014.

We plotten nu het jaarlijkse aantal gemeten aardbevingen tegen de cumulatieve gaswinning (zie figuur 21a). Met behulp van enkelvoudige regressie plotten we weer een regressielijn. Het model is van de vorm (zie ook Barnett en Dobson 2008):

$$\log(Y_i) = \beta_0 + \beta_2 x_{i,2} + \epsilon_i$$

met Y_i het aantal aardbevingen, $x_{i,2}$ de cumulatieve gaswinning en ϵ_i de fout met een verdeling $\epsilon_i \sim N(0, \sigma^2)$, onderling onafhankelijk. We vinden de volgende schatters: $\hat{\beta}_0 = -0,333$ en $\hat{\beta}_2 = 1,65 \cdot 10^{-3}$. De F-toets om te toetsen of er sprake is van een verband, geeft een p-waarde van $4,3 \cdot 10^{-8}$. We kunnen dus concluderen dat er ook een significant verband bestaat tussen de cumulatieve gaswinning en het aantal aardbevingen. Ter controle van het model bekijken we de residuenplot (figuur 21b).



Figuur 21: Verband tussen het jaarlijkse aantal gemeten bevingen en de cumulatieve gaswinning, gemeten vanaf 1963. Data aardbevingen uit 1995-2014. Figuur (a) geeft de meetwaarden in blauw en regressielijn in zwart. Figuur (b) geeft de residuenplot met $E = Y - \hat{Y}$, $Y :=$ aantal aardbevingen per jaar.

4.2.4 Meervoudige regressie

Nu we in paragrafen 4.2.2 en 4.2.3 hebben gezien dat er een verband is tussen het aantal aardbevingen en zowel de jaarlijkse gaswinning als de cumulatieve gaswinning, gaan we meervoudige regressie toepassen. Het model krijgt de volgende vorm (zie ook Barnett en Dobson 2008):

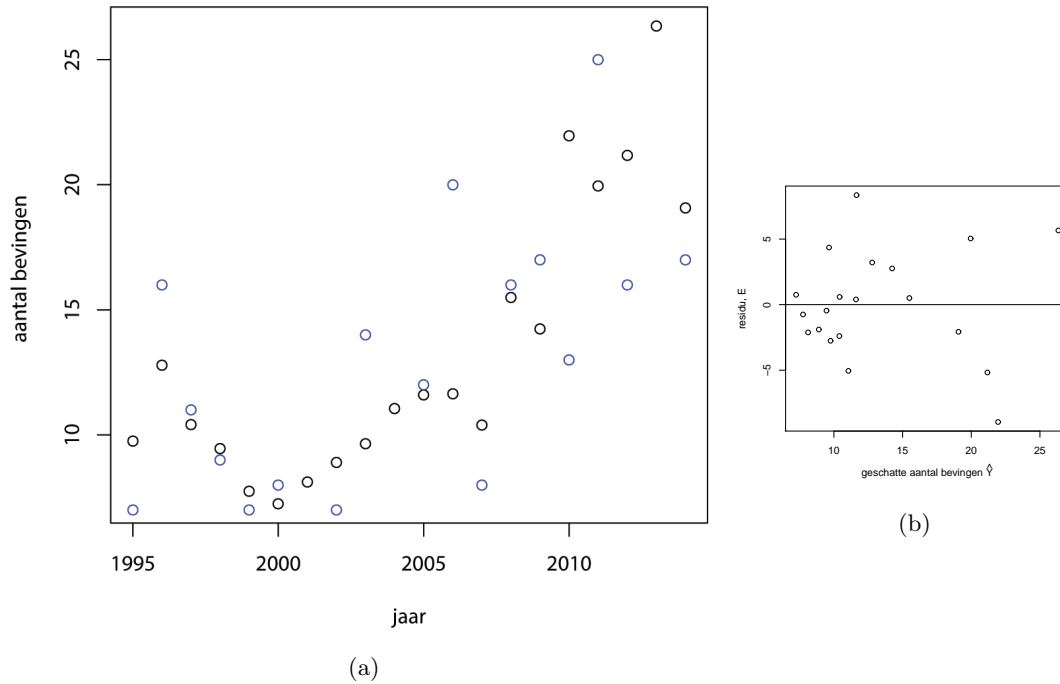
$$\log(Y_i) = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i$$

met Y_i het aantal aardbevingen, $x_{i,1}$ de jaarlijkse gaswinning, $x_{i,2}$ de cumulatieve gaswinning en ϵ_i de fout met een verdeling $\epsilon_i \sim N(0, \sigma^2)$, onderling onafhankelijk.

We vinden de volgende schatters: $\hat{\beta}_0 = 0,3644$, $\hat{\beta}_1 = 3,045 \cdot 10^{-2}$ en $\hat{\beta}_2 = 6,110 \cdot 10^{-4}$. We zien dat de cumulatieve gaswinning minder zwaar weegt in het model dan de jaarlijkse gaswinning. In figuur 22 is het resultaat te zien van meervoudige regressie.

Tabel 1: Output bij meervoudige regressie van de vorm: $Y_i = \beta + \beta_0 x_i + \beta_1 x_{i-1} + \beta_2 x_{i-2} + \beta_3 x_{i-3} + \beta_4 x_{i-4}$ met Y_i het aantal aardbevingen in jaar i en x_i de hoeveelheid gaswinning in $10^9 Nm^3$ in jaar i . In de laatste kolom staat de p-waarde van de F-toets.

	$\hat{\beta}$	p-waarde
β_0	0,034	0,00247
β_1	0,019	0,106
β_2	-0,022	0,128
β_3	0,019	0,133
β_4	-0,013	0,294



Figuur 22: Predictieplot voor het beschreven model waarbij zowel jaarlijkse als cumulatieve gaswinning wordt meegenomen als verklarende variabele. Data aardbevingen uit 1995-2014. Figuur (a) geeft de meetwaarden in blauw en de voorspellingen in zwart. Figuur (b) geeft de residuenplot met $E = Y - \hat{Y}$, $Y :=$ aantal aardbevingen per jaar.

In TNO en Dost 2014 wordt beweerd dat wijzigingen in de gasproductie invloed hebben op het aantal bevingen in de toekomst. We hebben daarom ook meervoudige regressie toegepast met als verklarende variabele de gaswinning van 1, 2, 3, of 4 jaar geleden. Het model heeft de volgende vorm (zie ook Barnett en Dobson 2008):

$$Y_i = \beta + \beta_0 x_i + \beta_1 x_{i-1} + \beta_2 x_{i-2} + \beta_3 x_{i-3} + \beta_4 x_{i-4} + \epsilon_i$$

met Y_i het aantal aardbevingen in jaar i en x_i de hoeveelheid gaswinning in $10^9 Nm^3$ en ϵ_i de fout met een verdeling $\epsilon_i \sim N(0, \sigma^2)$, onderling onafhankelijk. De resultaten van meervoudige regressie staan in tabel 1. We zien dat de verbanden met gaswinning in voorgaande jaren niet significant zijn wanneer ook de gaswinning van het huidige jaar als verklarende variabele mee wordt genomen in het model.

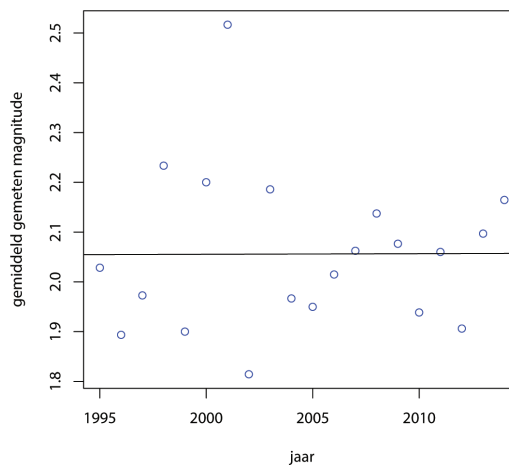
4.3 Magnitude aardbevingen en gaswinning

In deze paragraaf zullen we verder gaan met de data die is geïntroduceerd in de vorige paragraaf. Behalve de locaties en de tijdstippen, zijn ook de magnitudes van de aardbevingen bekend. We zullen in deze paragraaf onderzoeken of er een verband is tussen de gaswinning en magnitudes van de aardbevingen.

4.3.1 Gemiddelde magnitude



(a) Gemiddeld gemeten magnitude per jaar.



(b) Blauw: gemiddeld gemeten magnitude per jaar, zwart: trendlijn, 1995-2014.

We beschouwen eerst de gemiddelde magnitude gemeten per jaar. In figuur 23a is de gemiddelde magnitude in de tijd geplot. We zien dat de gemiddelde magnitude rond de 2,0 schommelt. In jaren 1988 en 1990 zijn geen bevingen gemeten, waardoor de gemiddelde magnitude in deze jaren gelijk is aan nul. Er is geen duidelijke trend te ontdekken. Voor 1995 lijkt de gemiddelde magnitude hoger te liggen. Dit kan komen door onbetrouwbare meetgegevens: niet alle bevingen konden worden gedetecteerd, met name de lichte bevingen.

In figuur 23b zien we de gemiddelde magnitude geplot in de tijd (vanaf 1995). Met behulp van enkelvoudige lineaire regressie is er een trendlijn doorheen geplot. Het model is van de vorm (zie ook Barnett en Dobson 2008):

$$Y_i = \beta_0 + \beta_1 x_{i,1} + \epsilon_i$$

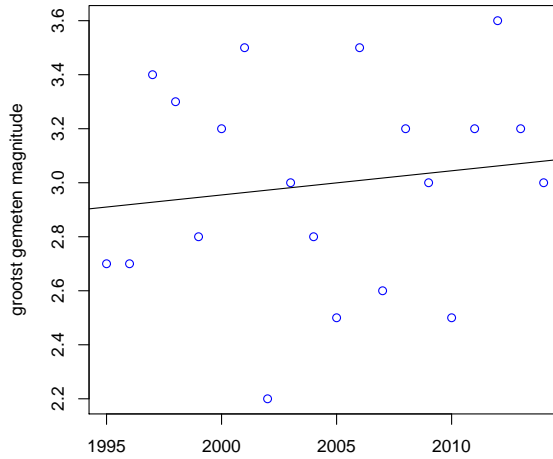
met Y_i het aantal aardbevingen, $x_{i,1}$ het jaartal en ϵ_i de fout met een verdeling $\epsilon_i \sim N(0, \sigma^2)$, onderling onafhankelijk. We vinden de volgende schatters: $\hat{\beta}_0 = 1,817$ en $\hat{\beta}_1 = 1,192 \cdot 10^{-4}$. De F-toets om te toetsen of $\beta_1 = 0$, geeft een p-waarde van 0,985. Dit resultaat duidt erop dat we geen duidelijke stijging kunnen aantonen.

4.3.2 Grootst gemeten magnitude

We bekijken nu ook de grootst gemeten magnitude per jaar. In figuur 23 zien we dat er meer fluctuatie is in deze maximale magnitude, vergeleken met de gemiddelde magnitude. Er lijkt echter nog steeds geen duidelijke trend te ontdekken. Met behulp van enkelvoudige lineaire regressie is er een trendlijn doorheen geplot. Het model is van de vorm (zie ook Barnett en Dobson 2008):

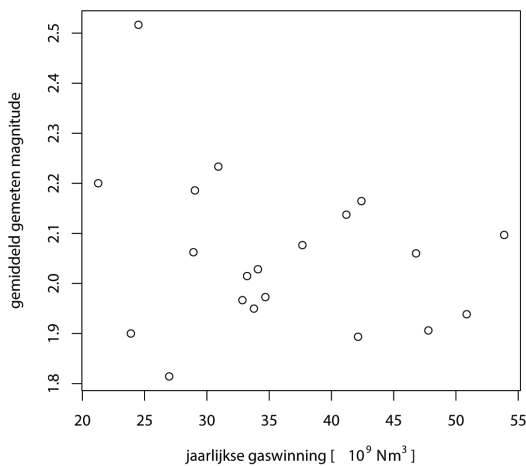
$$Y_i = \beta_0 + \beta_1 x_{i,1} + \epsilon_i$$

met Y_i het aantal aardbevingen, $x_{i,1}$ het jaartal en ϵ_i de fout met een verdeling $\epsilon_i \sim N(0, \sigma^2)$, onderling onafhankelijk. We vinden de volgende schatters: $\hat{\beta}_0 = -14,94$ en $\hat{\beta}_1 = 8,947 \cdot 10^{-3}$. De F-toets om te toetsen of $\beta_1 = 0$, geeft een p-waarde van 0,564. Dit resultaat duidt erop dat we geen duidelijke stijging kunnen aantonen in de maximale magnitude.

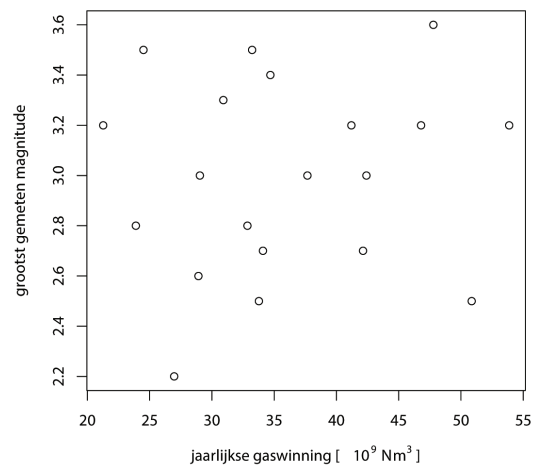


Figuur 23: Blauw: grootst gemeten magnitude per jaar, zwart: trendlijn, 1995-2014.

4.3.3 Lineaire regressie



(a) Verband tussen de gemiddelde magnitude van de aardbevingen per jaar en de hoeveelheid gewonnen gas van dat jaar.



(b) Verband tussen de grootste magnitude van de aardbevingen per jaar en de hoeveelheid gewonnen gas van dat jaar.

Figuur 24: Er lijkt in beide figuren geen verband te zijn.

Wanneer we de jaarlijkse gaswinning plotten tegen de gemiddelde magnitude (zie figuur 24a) zien we geen duidelijk verband. We stellen het volgende model op (zie ook Barnett en Dobson 2008):

$$\log(M_i) = \beta_0 + \beta_1 x_i + \epsilon_i$$

met M_i het de gemiddelde magnitude, x_i de jaarlijkse gaswinning en ϵ_i de fout met een verdeling $\epsilon_i \sim N(0, \sigma^2)$, onderling onafhankelijk. We toetsen de nulhypothese $H_0 : \beta_1 = 0$. Met een p-waarde van 0,904 kunnen we H_0 niet verwerpen.

Is er dan wel een duidelijk verband tussen de gaswinning en de grootst gemeten magnitude per jaar? Ook nu plotten we de jaarlijkse gaswinning tegen de grootst gemeten magnitude (zie figuur 24b). De puntenwolk is nog chaotischer dan de vorige wolk. Bij het toetsen naar een verband vinden we dan ook een p-waarde van 0,9179. We kunnen concluderen dat er ook geen verband is tussen de gaswinning en de grootst gemeten magnitude.

5 Model

5.1 Modelopbouw

We beschouwen een puntproces waarbij elk punt een aardbeving voorstelt, met coördinaten (x, y) . We discretiseren in ons model over de tijd. In ons geval betekent dat dat we het aantal aardbevingen in een bepaald jaar beschouwen als mark j . Deze keuze voor tijdstappen van een jaar is gebaseerd op de beschikbare data over de gaswinning en het aantal bevingen. Vanaf 2010 is de gaswinning per maand per productielocatie bekend. Door het kleine aantal bevingen is het moeilijk om iets over het maandelijkse aantal aardbevingen te zeggen, vandaar de keuze voor een jaar als tijdseenheid.

Op basis van de exploratieve analyse concluderen we dat de intensiteit van de bevingen niet homogeen is. We hebben ervoor gekozen het proces te modelleren als een inhomogeen Poissonproces. Wij willen met ons model de bevingsintensiteit $\lambda^j(x, y)$ modelleren. Hieronder verstaan we het aantal aardbevingen per oppervlakte-eenheid in het jaar j als functie van de coördinaten (x, y) . Deze intensiteit beperkt zich tot het gasveld Groningen en doet dus geen enkele uitspraak over bevingen buiten het gasveld. Deze keuze hebben we gemaakt aan de hand van de beschikbare data en het feit dat veruit de meeste aardbevingen plaatsvinden in het gasveld.

We willen de bevingsintensiteit als continue functie over de ruimte modelleren. Dit zal een functie zijn van een aantal verklarende variabelen, zeg X_1, \dots, X_n . Om de intensiteit te bepalen aan de hand van onze verklarende variabelen maken we gebruik van een log-lineair model. Dit wil zeggen dat de logaritme van de intensiteit een lineaire functie is van de variabelen X_1, \dots, X_n . Ons model is dus van de vorm:

$$\log(\lambda^j(x, y)) = \beta_0 + \beta_1 X_1(x, y) + \beta_2 X_2(x, y) + \dots + \beta_k X_k(x, y) + \beta_{k+1} X_{k+1}^j(x, y) + \dots + \beta_n X_n^j(x, y)$$

met $\beta_i \in \mathbb{R}$ bepaalde coëfficiënten en X_i een pixelbeeld ($i = 1 \dots n$) die de waarde van de verklarende variabele in het punt (x, y) weergeeft. Zie paragraaf 2.3.1 voor de uitleg over pixelbeelden. Een aantal variabelen is hetzelfde voor elk jaar van voorspellen. Dit zijn de variabelen X_1, \dots, X_k . Daarentegen, de variabelen X_{k+1}, \dots, X_n zijn wel afhankelijk van het jaar dat we beschouwen. Daarom dragen deze variabelen ook index j . De output van ons model is een pixelbeeld die de verwachte intensiteit weergeeft in het gasveld Groningen. Voor de details over de verklarende variabelen die wij meenemen verwijzen wij u naar paragraaf 5.3.

5.2 Schatten van de parameters

In het beschreven model zijn de parameters β_i nog niet bepaald. We zullen deze parameters bepalen aan de hand van de data over de jaren 2010 tot en met 2014. Dit doen we met behulp van de functie `ppm` in `SPATSTAT`. Aan de hand van het puntpatroon en pixelbeelden van de verklarende variabelen worden meest aannemelijke waarden voor de β 's bepaald. Dit betekent dat de β 's zo worden gekozen, dat het uiteindelijke puntpatroon bij die intensiteitsfunctie een zo hoog mogelijke “kans” heeft om plaats te vinden. Natuurlijk heeft elke individuele realisatie kans 0, omdat de coördinaten continu zijn. Daarom wordt de hoogte van de kansdichtheidsfunctie gebruikt om de combinatie van parameters te vinden met een zo hoog mogelijke aannemelijkheid voor het puntpatroon. Bij het berekenen van de aannemelijkheid van een intensiteitsschatting moet er eigenlijk een integraal over het oppervlak worden uitgerekend. Dit kan efficiënt worden uitgerekend door een discretisatie: een som van de waarde op een aantal locaties (waaronder de punten van het puntpatroon, aangevuld met dummy punten). Deze methode heet *Berman-Turner computational approximation*, geïntroduceerd in Berman en Turner 1992. Het maximaliseren van aannemelijkheid kan op die manier omgeschreven worden naar een ‘standaard’ probleem, waar de ingebouwde optimalisatie-functie van R wel raad mee weet. Wat er nog meer achter de functie `ppm` schuilgaat, staat verder uitgelegd in Baddeley 2014. In de bijgevoegde code in bijlage 10.4 is aan het begin van het bestand `model.R` te zien hoe het fitten uiteindelijk in zijn werk gaat.

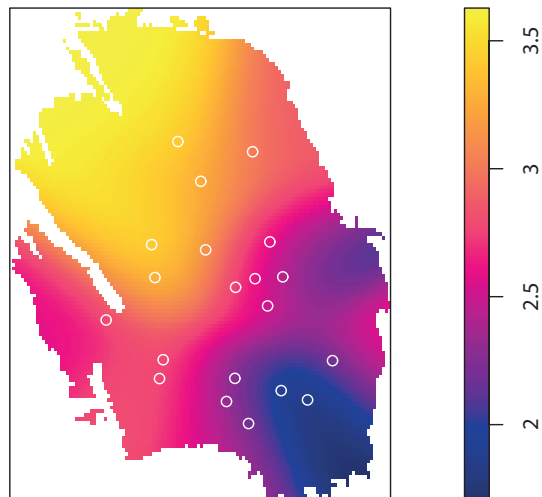
Het maximaliseren van de aannemelijkheid wordt gedaan voor alle jaren 2010-2014 tegelijk. Dit betekent dat we het puntpatroon van de bevingen in de jaren 2010-2014 ingeven. Daarnaast geven we de pixelbeelden van de variabelen voor elk jaar in. Door te werken met marks zorgen we ervoor dat alleen

de juiste variabelen worden gebruikt voor ieder jaar. Op deze manier krijgen we voor elke β_i de meest aannemelijke waarde over de jaren 2010-2014, in plaats van de meest aannemelijke waarde voor een enkel jaar. Hiermee verkomen we dat we waarden vinden voor de parameters die slechts voor één jaar goed passen.

5.3 Verklarende variabelen

In het model worden een aantal verklarende variabelen meegenomen (X_i). Deze verklarende variabelen hebben de vorm van een pixelbeeld. Wat een pixelbeeld is, staat beschreven in paragraaf 2.3.1. In deze paragraaf zullen alle meegenomen pixelbeelden worden beschreven.

In de gebruikte data speelt afstand vaak een rol: onder andere in de grenzen van de gebieden, de locaties van de bevingen en de locaties van de boorputten. Als standaardeenheid voor de afstanden worden in het model kilometers gebruikt in plaats van meters. Op de schaal die we beschouwen is dit een logischere keuze. Bovendien heeft het werken met kilometers numeriek voordelen boven het werken met meters: de getallen uitgedrukt in meters kunnen erg groot worden, bijvoorbeeld bij het berekenen van oppervlakken.



Figuur 25: Pixelbeeld behorende bij de gaswinning in 2013. De punten geven de locaties aan van de boorputten.

5.3.1 Gaswinning

Allereerst wordt als verklarende variabele de gaswinning meegenomen in het model. Dit doen we aan de hand van productielocaties. In het gasveld Groningen zijn putten geboord waaruit het gas wordt gewonnen. Deze boorputten liggen in kleine groepjes geclusterd in het gasveld. Per productielocatie bepalen we één centrum. Deze centra zijn met witte rondjes weergegeven in figuur 25.

Van deze centra zijn de jaarlijkse gaswinningscijfers beschikbaar. De gaswinningscijfers zijn uitgedrukt in de eenheid $10^9 Nm^3$, miljard normaal kubieke meter. We gebruiken de techniek van smoothen met Gaussische kernen om een gladde functie te krijgen over het gasveld. In paragraaf 2.3.2 is meer over deze techniek te lezen. De gekozen bandbreedte is $\sigma = 3,5$ km. Deze waarde voor σ is gebaseerd op onderzoek van TNO waaruit blijkt dat de effecten van gaswinning tot 3,5 km meetbaar zijn (zie TNO en Dost 2014).

De Gaussische kernen worden geschaald naar de hoeveelheid gaswinning per productielocatie per jaar: hoe meer gaswinning, hoe hoger de gewichten w_i . Zo ontstaat er voor elk jaar een pixelbeeld zoals te zien in figuur 25. Dit pixelbeeld hoort bij het jaar 2013. Bij het doen van voorspellingen is de gaswinning uit de toekomst uiteraard nog niet bekend. Mogelijke gaswinningscijfers kunnen als input worden meegegeven aan het model.

De data over de gaswinning wordt door de NAM gepubliceerd. De gegevens zijn te vinden op het online NAM platform (NAM 2015). In de bijlage in paragraaf 10.1 is de gebruikte data weergegeven.

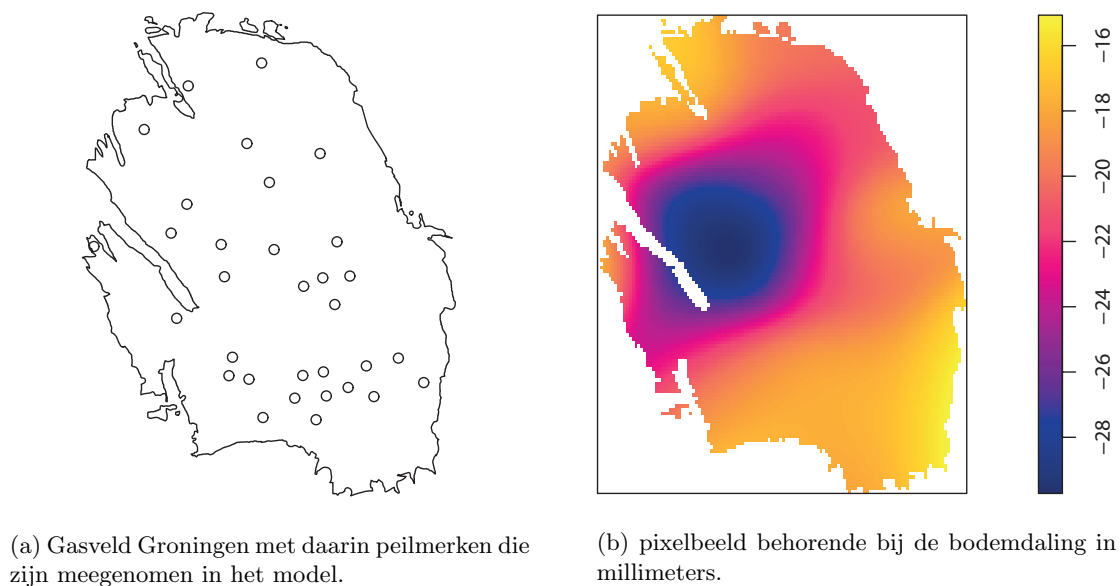
5.3.2 Bodemdaling

Als tweede verklarende variabele nemen we de bodemdaling mee in ons model. Doordat het gasveld steeds leger wordt, treedt er compactie op waardoor de bodem daalt. In het gasveld Groningen is een groot aantal peilmerken waar de bodemdaling wordt gemeten. Wij hebben 32 peilmerken in de buurt

van productielocaties geselecteerd en daarvan de bodemdalingcijfers tussen 2003 en 2008 gebruikt. In figuur 26a zijn de meegenomen peilmerken in het gasveld Groningen weergegeven.

De marks behorende bij deze peilmerken zijn de bodemdalingcijfers in millimeters in de periode 2003 tot en met 2008. Door smoothen met Gaussische kernen (zie paragraaf Smoothen met Gaussische kernen) wordt nu een gladde functie gemaakt over het gasveld. De hoogtes van de Gaussische kernen zijn hier gelijk aan de bijbehorende marks. Als bandbreedte wordt $\sigma = 3,75$ km gebruikt. Deze sigma is iets hoger gekozen dan de bandbreedte van de gaswinning, omdat bodemdaling minder lokaal effect heeft (zie paragraaf 3.1). In figuur 26b is het pixelbeeld te zien dat we meenemen als verklarende variabele in het model.

Omdat ons model gebruikt zal worden om te voorspellen over enkele jaren, kunnen we de bodemdaling beschouwen als stationair. De data over de bodemdaling is afkomstig van TNO, van de website www.nlog.nl. In de bijlage 10.2 is een tabel opgenomen met daarin de peilmerken en de bijbehorende bodemdaling.



Figuur 26: Bodemdaling als verklarende variabele.

5.3.3 Breuklijnen

De aardbevingen ontstaan in breuklijnen in de bodem (zie paragraaf 3.2). We nemen dan ook de locaties van de grootste breuklijnen in ons model mee. In het rapport van TNO (TNO 2013) zijn figuren opgenomen met daarin de breuklijnen in kaart gebracht (zie ook figuur 11). De grootste breuklijnen hebben wij benaderd met rechte lijnstukken en opgenomen in ons model. We hebben geen toegang tot de data van seismiek waarin de precieze breuklijnen beschreven staan. Daarom hebben we voor deze oplossing gekozen. In figuur ?? zijn deze breuklijnen in kaart gebracht. De precieze locaties van de lijnen in ons model zijn te vinden in de bijlage in paragraaf 10.3. Door smoothen over de lijnstukken wordt een gladde functie gemaakt over het gasveld. De gekozen bandbreedte is in dit geval $\sigma = 3,5$ km. In figuur ?? is het pixelbeeld te zien dat mee wordt genomen in het model. Net als de bodemdaling heeft deze variabele geen jaarindex.

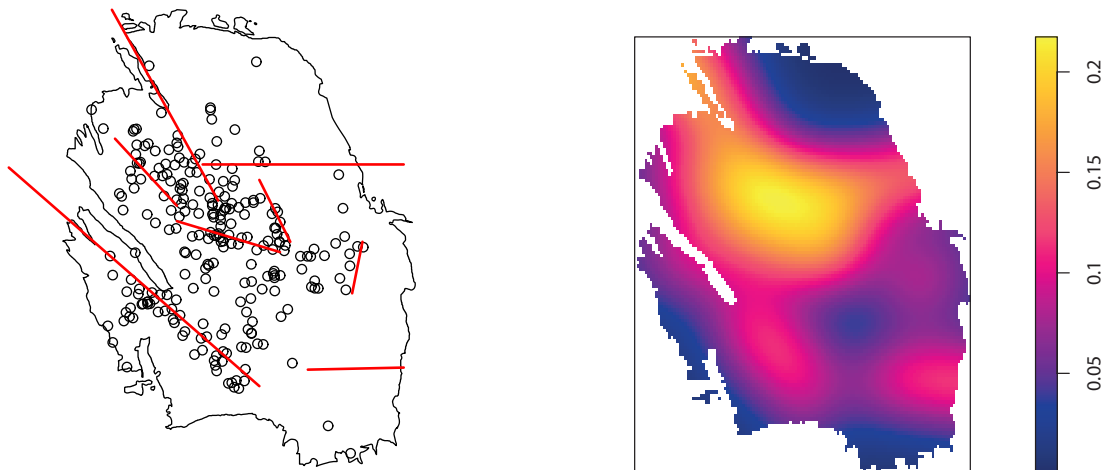
5.3.4 Bevingen in het verleden

Tenslotte wordt ook het verleden van de aardbevingen meegenomen in het model: we nemen de (geschatte) bevingensintensiteit mee van het afgelopen jaar. Net als bij de andere verklarende variabelen wordt ook hier gebruik gemaakt van smoothen met Gaussische kernen. In dit geval smoothen we over de aardbevingen uit het verleden. Als bandbreedte wordt $\sigma = 3,75$ km gebruikt. Vanwege de implementatie

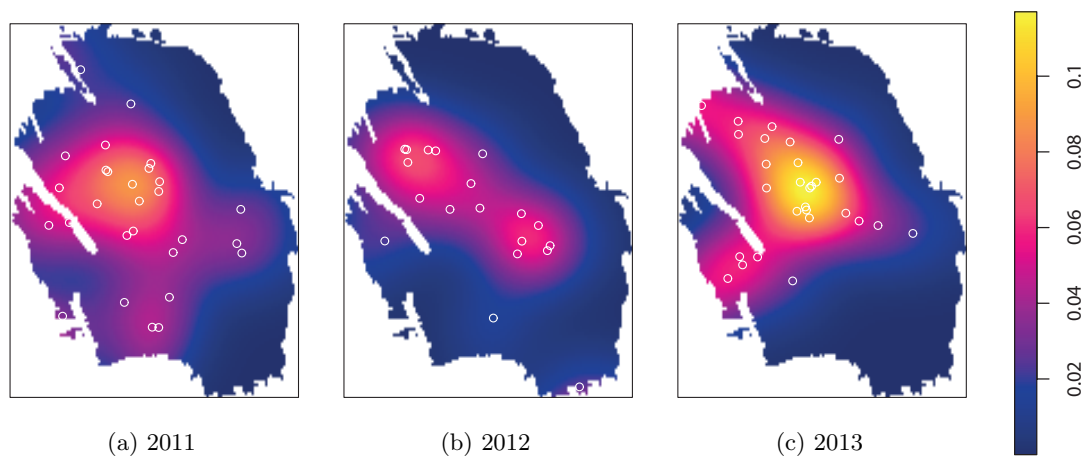
was het helaas niet mogelijk om gewichten w_i aan de verschillende kernen mee te geven, waardoor alle bevingen gelijk zijn geschaald. Het zou wenselijk zijn om bevingen met een grotere magnitude zwaarder mee te nemen in het model dan bevingen met een kleine magnitude. In de discussie gaan we hier verder op in.

In figuren 28a, 28b en 28c zijn pixelbeelden weergegeven die op deze manier zijn gemaakt. Het gaat hier om de aardbevingen uit 2011, 2012 en 2013. Wanneer we gaan voorspellen over jaar k en de gemiddelde bevingintensiteit van het jaar $k - 1$ niet bestaat (aangezien jaar $k - 1$ in de toekomst ligt) gebruiken we onze voorspelde intensiteit van jaar $k - 1$. Met behulp van onze voorspellingen doen we dan weer nieuwe voorspellingen.

De data over de locaties, tijden en magnitudes van aardbevingen worden gepubliceerd op www.KNMI.nl (KNMI 2015).



Figuur 27: Breuklijnen als verklarende variabele. Links: aardbevingen in gasveld Groningen met de breuklijnen die zijn meegenomen. Rechts: pixelbeeld behorende bij de breuklijnen.



Figuur 28: In figuren (a), (b) en (c) zijn de dichtheidsfuncties te zien van de jaren 2011, 2012 en 2013. De witte punten geven de aardbevingen aan die in dat jaar hebben plaatsgevonden.

6 Parameterschatting en modelvalidatie

In dit hoofdstuk zullen we de parameters β_i bepalen aan de hand van de beschikbare data van 2010 tot en met 2014. Tevens zullen we het model valideren voor het jaar 2014.

6.1 De geschatte parameters β_i

Wanneer we parameterwaarden voor β_i bepalen met `ppm` geeft dit de waarden vermeld in tabel 2. Het model is gefit zoals aangegeven in paragraaf 5.2, waarbij uiteindelijk de volgende modelvergelijking is gebruikt:

$$\begin{aligned}
 \log(\lambda^{2010}(x, y)) & & X_3^{2010}(x, y) & & X_4^{2010}(x, y) \\
 \log(\lambda^{2011}(x, y)) & & X_3^{2011}(x, y) & & X_4^{2011}(x, y) \\
 \log(\lambda^{2012}(x, y)) & = \beta_0 + \beta_1 X_1(x, y) + \beta_2 X_2(x, y) + \beta_3 & X_3^{2012}(x, y) + \beta_4 & X_4^{2012}(x, y) \\
 \log(\lambda^{2013}(x, y)) & & X_3^{2013}(x, y) & & X_4^{2013}(x, y) \\
 \log(\lambda^{2014}(x, y)) & & X_3^{2014}(x, y) & & X_4^{2014}(x, y)
 \end{aligned} \tag{13}$$

waarbij X_1 het pixelbeeld behorend bij de bodemdaling voorstelt, X_2 de breuklijnen, X_3^j gaswinning in jaar j en X_4^j de geschatte bevingintensiteit in jaar $j - 1$.

Tabel 2: Uitkomsten van het schatten van de parameters.

Variabele	Schatting	Standaarddeviatie	Ondergrens 95% BI	Bovengrens 95% BI	Z-toets
Intercept	-7,9100056	0,7094296	-9,3004622	-6,5195491	***
Bodemdaling	0,1248741	0,0364693	0,0534957	0,1964526	***
Breuklijnen	7,6533184	2,4129051	2,9239767	12,3823909	**
Gaswinning	0,1795713	0,1555108	-0,1252243	0,4843669	
Bevingen voorgaande jaar	-1,2457946	5,5499119	-12,1234220	9,6318327	

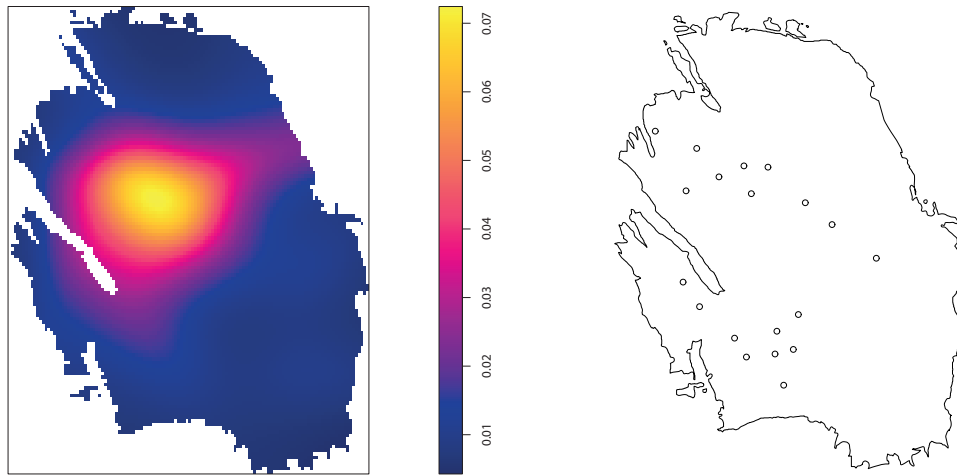
We zien in tabel 2 dat de coëfficiënten voor de gaswinning en de bevingen in het voorgaande jaar niet significant zijn. We willen deze variabelen echter wel meenemen. Het is immers gebleken dat de gaswinning wel van invloed is op het aantal aardbevingen, zie paragraaf 4.2.2. De bevingen uit het voorgaande jaar zijn in ons model van belang om doorgaande tijdseffecten te modelleren.

6.2 Validatie van het model voor 2014

De gaswinning in het jaar 2014 wijkt sterk af van de gaswinning in de jaren 2012-2013, zie figuur 31. De jaren 2010-2011 lijken qua gaswinning erg op 2012-2013. Daarom zullen we de data uit het jaar 2014 gebruiken om het model te valideren. We schatten de bevingintensiteit aan de hand van ons model. We geven dus de pixelimages van de juiste variabelen in en we gebruiken de in paragraaf 6.1 gevonden parameters. In figuur 29a is de schatting voor de bevingintensiteit voor het jaar 2014 gegeven en in figuur 29b zijn de bevingen in het gasveld Groningen in het jaar 2014 weergegeven.

Om te bepalen of de geschatte bevingintensiteit een plausibele schatting is, maken we gebruik van de inhomogene K-functie en de J-functie, zie paragraaf 4.1. De resultaten hiervan zijn weergegeven in figuur 30.

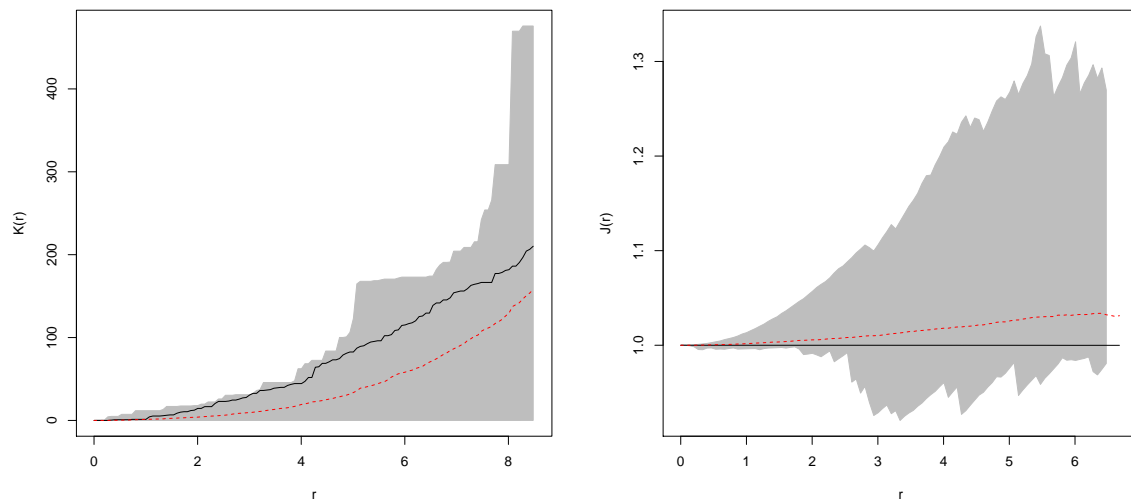
We zien dat zowel de K-functie als de J-functie binnen de grijze band vallen. Wanneer we kijken naar het aantal bevingen in 2014, is dit gelijk aan negentien. Wanneer we integreren over de geschatte intensiteit, waarmee we het aantal verwachte bevingen in 2014 benaderen, geeft dit de waarde 17,58976. Hieruit concluderen we dat ons model een plausibele schatting geeft voor de bevingintensiteit.



(a) De geschatte bevingintensiteit

(b) De bevingen in het jaar 2014

Figuur 29: De geschatte bevingintensiteit voor het jaar 2014 en de bevingen in het jaar 2014.



(a) De K-functie

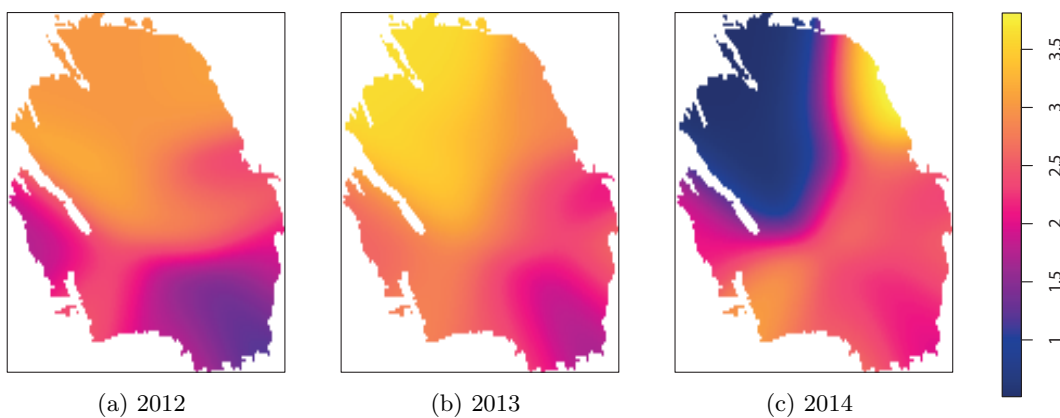
(b) De J-functie

Figuur 30: De K-functie en de J-functie voor de bevingen in 2014 op basis van de geschatte bevingintensiteit.

7 Voorspellingen

In dit hoofdstuk zullen we met ons model voorspellingen doen. Aangezien voorspellingen afhankelijk zijn van de gaswinning, is de toekomst niet eenduidig te voorspellen. Hoeveel gas er in de toekomst zal worden gewonnen is afhankelijk van keuzes van de politiek. We kiezen ervoor om verschillende toekomstscenario's door te rekenen. We zullen in paragraaf 7.1 een voorspelling doen voor het huidige jaar. De gaswinningscijfers van het eerste deel van 2015 zullen we hierbij doortrekken. In de daaropvolgende paragraaf zullen we kijken naar de extreme situatie dat de gaskraan totaal wordt dichtgedraaid. In paragraaf 7.3 worden nog twee andere toekomstscenario's doorgerekend.

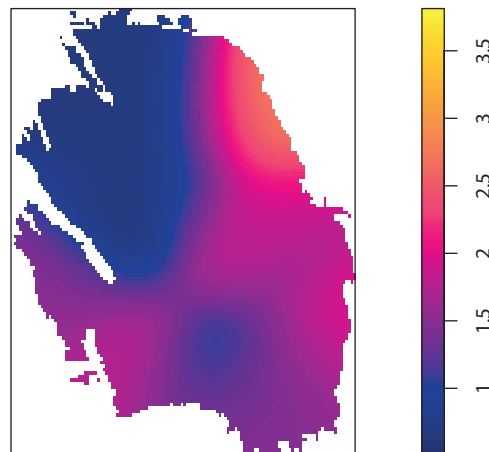
In figuur 31 zijn de pixelbeelden weergegeven behorend bij de gaswinning van de jaren 2012 tot en met 2014. We zien dat er in 2014 een enorme verandering heeft plaatsgevonden: de gaswinning op locaties rondom Loppersum zijn tot 80 procent teruggedrongen. Hieruit blijkt dat het niet eenvoudig is om te voorspellen hoeveel gas er in de toekomst zal worden gewonnen. Er wordt veel gediscussieerd hoeveel en waar gas moet worden gewonnen.



Figuur 31: De pixelbeelden behorende bij de gaswinning in de jaren 2012 tot en met 2014 [$10^9 Nm^3$].

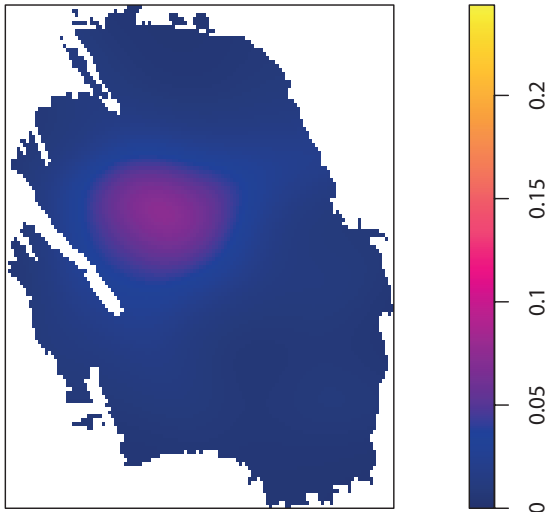
7.1 Verwachting voor 2015

Allereerst zullen we met ons model doorrekenen wat het verwachte aantal aardbevingen is in 2015. De gaswinningscijfers van januari 2015 tot en met april 2015 zijn door het NAM beschikbaar gesteld (NAM 2015). In de bijlage (10.1) zijn deze gegevens weergegeven. Wanneer we ervan uitgaan dat de gaswinning in een jaar constant is per maand, kunnen we de verwachte gaswinning voor heel 2015 berekenen. De gegeven data beslaat vier maanden, een derde jaar. Wanneer we deze cijfers vermenigvuldigen met drie krijgen we de verwachte gaswinning in heel dit lopend jaar. Dit zullen we gebruiken als input van onze voorspelling. In figuur 32 is het bijbehorende pixelbeeld te zien. Als we dit vergelijken met de gaswinning in eerdere jaren, zien we dat de gaswinning flink is terugeschroefd. Zou dit invloed hebben op het verwachte aantal bevingen?

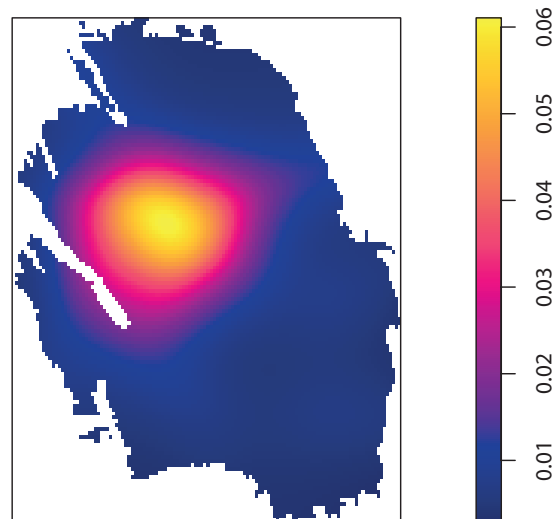


Figuur 32: pixelbeeld behorend bij verwachte gaswinning 2015 [$10^9 Nm^3$]

In figuur 33 is de verwachte bevingsintensiteit voor 2015 weergegeven. Deze verwachte bevingsintensiteit lijkt laag te zijn. Het verwachte aantal bevingen in gasveld Groningen in 2015 kan worden berekend met de integraal over het oppervlak. Hieruit komt een verwacht aantal bevingen van $18,11061 \approx 18$. Lijkt dit een reële schatting? In de maanden januari tot en met april 2015 hebben in totaal 7 aardbevingen plaatsgevonden met een magnitude van 1,5 of groter⁹. Wanneer we ervan uitgaan dat de aardbevingen uniform over het jaar zijn verdeeld, zullen er naar verwachting drie maal zoveel aardbevingen plaatsvinden in heel 2015. Dit komt uit op $7 \cdot 3 = 21$ aardbevingen. Onze schatting van 18 aardbevingen lijkt dus een reële schatting.



Figuur 33: Verwachte bevingsintensiteit 2015 [bevingen per km²].



Figuur 34: Verwachte bevingsintensiteit 2018 wanneer de gaskraan vanaf mei 2015 wordt dichtgedraaid [bevingen per km²]. Let op de schaal!

7.2 De gaskraan gaat dicht, en dan?

Het tweede toekomstscenario dat we doorrekenen is wanneer de gaskraan vanaf mei 2015 totaal is dichtgedraaid. Als input geven we de gaswinningscijfers die bekend zijn (tot en met april 2014). Hierna zetten we de gaswinningscijfers op nul. Uit voorspellingen komen de getallen weergegeven in tabel 3. We zien dat het verwachte aantal aardbevingen convergeert naar 13. In de pixelbeelden die als output worden verkregen is te zien dat de spreiding van de intensiteit ook convergeert. In figuur 34 is te zien hoe deze spreiding eruit ziet. Let op: de kleurverdeling in deze afbeelding suggereert een hoge intensiteit. Aan de schaal is echter te zien dat de intensiteiten niet boven de 0,06 bevingen per km² komen.

Tabel 3: Verwachte aantal aardbevingen wanneer er vanaf mei 2015 geen gas meer wordt gewonnen.

jaar	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026
bevingen	15,29	14,04	13,37	13,83	13,51	12,85	13,52	13,46	13,10	13,49	12,66	12,87
afgerond	15	14	13	14	14	13	14	13	13	13	13	13

⁹Bron: www.KNMI.nl

7.3 Andere toekomstscenario's

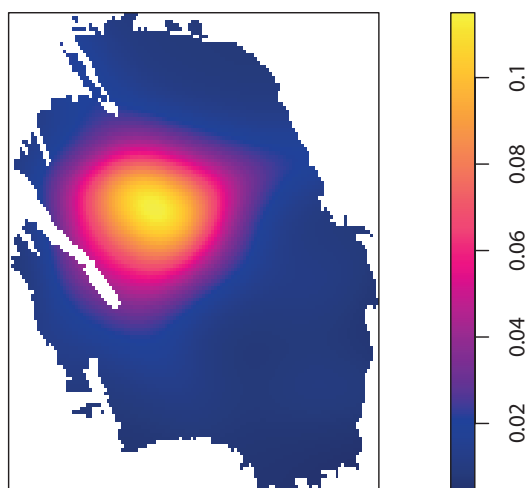
We bekijken nu het geval dat de gaswinning vanaf 2016 constant is en gelijk is aan de gaswinning in 2013. We gaan ervan uit dat de gaswinning in 2015 is zoals in paragraaf 7.1. In de eerste rijen van tabel 4 is te zien, hoe het verwachte aantal aardbevingen zich ontwikkelt in de komende jaren in dit scenario. We zien dat het jaarlijks aantal aardbevingen convergeert naar 23 à 24. De spreiding van de intensiteit over het gasveld is in de voorspelde jaren redelijk constant. In figuur 35 is te zien hoe deze spreiding eruit ziet.

We doen nu hetzelfde als hiervoor, alleen vervangen we de gaswinning van 2013 door de gaswinning zoals in 2014. In figuur 31 is te zien dat de gaswinning in deze twee jaren sterk verschilt. Niet alleen werd er in 2014 minder gas gewonnen dan in 2013, ook de spreiding van de gaswinning over het gasveld is flink veranderd. In tabel 4 is het verwachte aantal aardbevingen voor de komende jaren weergegeven in dit scenario. We zien dat het jaarlijks aantal aardbevingen convergeert naar achttien.

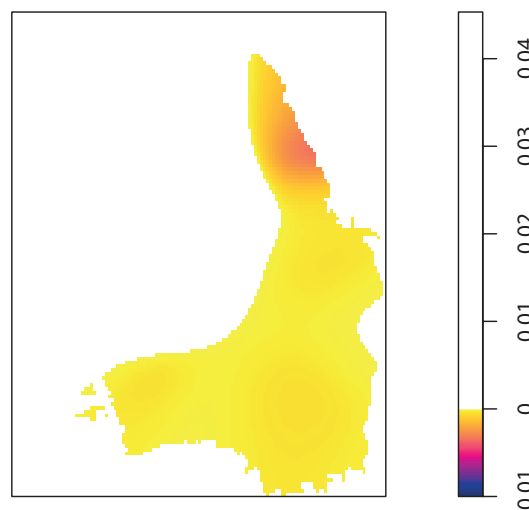
We zijn nu benieuwd naar de spreiding van de bevingen in vergelijking met het scenario waarin er jaarlijks evenveel gas werd gewonnen als in 2013. In figuur 36 is dit verschil in een pixelbeeld geplot. De kleurenschaal is bewust gekozen. Hierdoor is zichtbaar dat er een gebied bestaat waar de bevingsintensiteit toe is genomen in 2014. Dit komt overeen met de uitkomst van het onderzoek uit TNO en Dost 2014, waarin wordt geconcludeerd dat aan de randen van het gasveld de bevingsintensiteit toeneemt bij dit gaswinningsscenario.

Tabel 4: Verwachte aantal aardbevingen wanneer er vanaf 2016 jaarlijks evenveel gas wordt gewonnen als in 2013 (boven) respectievelijk 2014 (onder).

jaar	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026
bevingen	18,11	24,62	23,82	24,43	24,05	23,41	24,07	24,36	23,48	24,18	23,40	23,82
afgerond	18	25	24	24	24	23	24	24	23	24	23	24
bevingen	18,11	18,58	18,08	18,46	18,20	17,78	18,21	18,42	17,84	18,26	17,73	18,03
afgerond	18	19	18	18	18	18	18	18	18	18	18	18



Figuur 35: Verwachte bevingsintensiteit 2018 wanneer er vanaf 2016 jaarlijks evenveel gas wordt gewonnen als in 2013 [bevingen per km²].



Figuur 36: Verschil tussen twee verwachte bevingsintensiteiten voor 2018: wanneer er jaarlijks op dezelfde manier gas wordt gewonnen als in 2013 respectievelijk 2014 [bevingen per km²]. Het blijkt dat de bevingsintensiteit in het oranje gebied toeneemt.

8 Conclusie

In dit hoofdstuk komen we terug op de hoofdvraag en deelvragen die we hebben gesteld in de inleiding. De hoofdvraag luidt: ‘Hoe kun je in wiskundig opzicht de locaties en tijden van geïnduceerde aardbevingen in het gasveld Groningen modelleren en voorspellen?’ Voordat we terugkomen op de hoofdvraag, beantwoorden we eerst onze deelvragen.

1. Is er een toename in het aantal aardbevingen?

Ja, we kunnen spreken van een toename in het aantal aardbevingen in de tijd. In paragraaf 4.2.1 is beschreven hoe dit is getoetst.

2. Is er een toename in de magnitudes van de aardbevingen?

Voor zowel de gemiddelde magnitude per jaar als de grootst gemeten magnitude per jaar geldt dat een stijging niet aantoonbaar is. Dit blijkt uit toetsen beschreven in paragraaf 4.3. We concluderen daarom dat er geen sprake is van toename in magnitudes in de tijd.

3. Is er een aantoonbaar verband tussen de hoeveelheid gewonnen gas en het aantal aardbevingen?

Ja, er is een aantoonbaar positief verband tussen de gaswinning en het aantal aardbevingen. Hoe dit is getoetst, staat beschreven in paragraaf 4.2.2. Ook tussen de cumulatieve gaswinning (gemeten vanaf 1963) en het aantal aardbevingen is een positief verband aantoonbaar (zie paragraaf 4.2.3). We zien echter dat in een model waar zowel de jaarlijkse gaswinning als de cumulatieve gaswinning wordt meegenomen, de invloed van de cumulatieve gaswinning gering is.

We hebben ook onderzocht of er een verband is tussen het aantal aardbevingen en gaswinning uit voorgaande jaren. Wanneer ook de gaswinning van het huidige jaar als verklarende variabele wordt meegenomen in het model, blijkt de gaswinning in de voorgaande jaren niet significant (zie paragraaf 4.2.4).

4. Is er een aantoonbaar verband tussen de hoeveelheid gewonnen gas en de magnitudes van aardbevingen?

Uit de toets beschreven in paragraaf 4.3.1 blijkt dat er geen aantoonbaar verband bestaat tussen de gaswinning en de gemiddelde magnitude per jaar. Ook tussen de gaswinning en de grootst gemeten magnitude per jaar blijkt geen aantoonbaar verband te bestaan (zie paragraaf 4.3.2).

5. Is er een aantoonbaar verband tussen de bodemdaling en het aantal aardbevingen?

Door gebrek aan data was het niet mogelijk om de jaarlijkse bodemdaling te vergelijken met het jaarlijks aantal gemeten aardbevingen. Wel bleek uit ons model dat bodemdaling significante invloed heeft op de bevingsintensiteit (zie tabel 2).

6. Hoe zijn de aardbevingen verdeeld over de ruimte?

In paragraaf 4.1 zijn verschillende toetsen gedaan om te toetsen of er sprake is van homogeen Poissonproces. Uit deze toetsen kunnen we concluderen dat hier geen sprake van is. We zijn in het model daarom uitgegaan van een inhomogeen Poissonproces. Onze gemodelleerde intensiteit is hoog in het gebied rondom Loppersum (centraal gelegen in het gasveld) en laag aan de randen van het gasveld.

7. Na hoeveel tijd vinden er geen aardbevingen meer plaats, wanneer de gaswinning wordt stopgezet?

Wanneer de gaswinning wordt stopgezet volgt uit ons model geen totale uitdoving van het aantal aardbevingen. Het blijkt dat het jaarlijks aantal aardbevingen binnen enkele jaren convergeert naar dertien. Ons model is enkel geschikt voor voorspellingen op de korte termijn. In de discussie wordt hier verder op ingegaan.

8. Wat gebeurt er met het aantal aardbevingen wanneer er een constante hoeveelheid gas wordt gewonnen?

Wanneer er een constante hoeveelheid gas wordt gewonnen, zien we in ons model dat het jaarlijks aantal aardbevingen convergeert. De vorm van de uiteindelijke bevingsintensiteit is afhankelijk van de manier waarop gas wordt gewonnen. Hiermee bedoelen we zowel hoeveelheid als locatie.

Wij hebben een model gemaakt dat de jaarlijkse bevingsintensiteit in gasveld Groningen modelleert. Dit model is gebaseerd op een inhomogeen Poissonproces. Hierbij hebben we relevante data meegenomen als covariaten in de vorm van pixelbeelden (2.3.1). Dit model is minder waarheidsgetrouw dan de bestaande modellen van bijvoorbeeld TNO en NAM. Het model onderscheidt zich door zijn eenvoud. Bovendien vergt het weinig rekenkracht om met het model te voorspellen, te simuleren en te analyseren.

9 Discussie

Helaas hebben wij met de beschikbare tijd en mankracht de modellen van professionele onderzoeksteams bij CBS, NAM, SodM en TNO niet kunnen evenaren. Ons model heeft dan ook minder verklarende en voorspellende waarde dan de modellen die deze organisaties voorstellen. Dit blijkt vooral uit de gevoeligheid van ons model bij het voorspellen van intensiteiten in verschillende scenario's van gaswinning: hier zien we betrekkelijk weinig verschil in de spreiding van bevingsintensiteit, terwijl bijvoorbeeld in TNO en Dost 2014 duidelijkere veranderingen voorspeld worden.

Ons log-lineaire model steunt op een aantal aannames die een behoorlijke versimpeling inhouden. Vooral van grote invloed is het feit dat we bodemdaling als constant (stationair) in ons model houden. Natuurlijk is bodemdaling sterk gecorreleerd met zowel gaswinning als verspreiding van aardbevingen en zeker niet constant in de tijd. Hierdoor zijn onze voorspellingen minder nauwkeurig naarmate we verder vooruit voorspellen. We zien dit ook aan onze regressiecoëfficiënten voor gaswinning en eerdere bevingen, die (waarschijnlijk vanwege deze correlatie) niet significant zijn. Een beter model zou voorspelde bodemdaling mee kunnen nemen. In het rapport van NAM 2010 is de bodemdaling voorspeld tot het jaar 2070.

Als tweede grote versimpeling hebben we aangenomen dat aardbevingen binnen een jaar plaatsvinden via een Poisson proces met inhomogene intensiteit. Dan moeten de bevingen wel onafhankelijk zijn van elkaar, een aanname die vanwege de aanwezigheid van naschokken niet direct te onderbouwen valt. Hierin volgen we de onderzoekers van CBS en TNO. Seismologen gebruiken daarentegen vaak modellen zoals die gepresenteerd in paragraaf 3.6, waar naschokken expliciet mee gemodelleerd worden.

Een duidelijke tekortkoming in ons model is het feit dat we alle bevingen even zwaar meenemen. Dit terwijl er gewoon data (van het KNMI) en theorie (zoals de Gutenberg-Richter vergelijking in (8) beschikbaar is om de kracht van aardbevingen te vergelijken. Helaas hebben we door een tekortkoming in het door ons gebruikte softwarepakket SPATSTAT deze invloed niet mee kunnen nemen. We hadden de magnitudes van bevingen mee willen nemen als mark, maar helaas kan SPATSTAT geen modellen fitten met continue (numerieke) markwaardes. Bij het smoothen van bevingen om de intensiteit te schatten is het wel mogelijk om de magnitude als gewicht mee te nemen. Dit zou voordelig zijn omdat je dan meer data mee kunt nemen.

Toevoegingen op ons model, en daarmee ook mogelijkheden voor vervolgonderzoek, hangen vooral op de beschikbaarheid van geschikte data. Onze predicties zijn erg gelimiteerd doordat er slechts vanaf 2010 ruimtelijke informatie is over de gaswinning in Groningen; daarnaast waren er maar tot 2008 makkelijk bruikbare gegevens van bodemdaling beschikbaar. Vooral de pixelbeelden die we voor breuklijnen gemaakt hebben, zijn wat twijfelachtig doordat we geen toegang hadden tot het Petrel-bestand van de NAM. Gegevens van seismiek zijn alleen tegen betaling beschikbaar en dan nog in een zeer ingewikkeld formaat. Was deze data wel beschikbaar, dan had de ruimtelijke component van ons model veel waarheidsgetrouwer kunnen zijn. We hebben via verschillende instanties geprobeerd om meer data te verkrijgen. Dit is echter niet gelukt.

Wat nog een waardevolle toevoeging zou kunnen zijn op ons model is een onzekerheidsanalyse. Met behulp van betrouwbaarheidsintervallen van onze regressiecoëfficiënten zouden we verschillende mogelijke scenario's beter kunnen analyseren. Dit hebben we niet gedaan omdat de betrouwbaarheidsintervallen bij ons te groot zijn, waarbij de uitersten zeer onwaarschijnlijk of zelfs fysiek onmogelijk zijn. De oorzaak hiervan is het feit dat er sprake is van "overfitting", doordat we niet genoeg data hebben en doordat deze te sterk gecorreleerd zijn.

Een uitgebreider model zou de intensiteit van aardbevingen continu in de tijd kunnen modelleren, in plaats van per jaar gediscrèteerd. Dit zou bijvoorbeeld kunnen met de R-pakketten PTPROCESS en STPP. Als de ruimtelijke spreiding van bevingen verwaarloosd wordt, dan zou er in tijdszin beter kunnen worden voorspeld met de beschikbare data. Dit is voor ons een te grote stap.

Referenties

- Baddeley, A. et al. (2014). *On Tests of Spatial Pattern Based in Simulation Envelopes*.
- Baddeley, A.J., J. Møller en R. Waagepetersen (2000). “Non- and Semi-Parametric Estimation of Interaction in Inhomogeneous Point Patterns”. In: *Statistica Neerlandica* 54, pp. 329-350.
- Barnett, A.G. en A.J. Dobson (2008). *An Introduction to Generalized Linear Models, third edition*. CRC Press.
- Berman, M. en P. Diggle (1989). “Estimating Weighted Integrals of the second-order Intensity of a Spatial Point Process”. In: *Journal of the Royal Statistical Society* 51.
- Berman, M. en T.R. Turner (1992). “Approximating Point Process Likelihoods with GLIM”. In: *Applied Statistics* 41, pp. 31-38.
- Coculescu, D. en A. Nikeghbali (2007). *Filtrations*. <http://arxiv.org/abs/0712.0622v1>. [Online; bezocht op 24-04-2015].
- Daley, D.J. en D. Vere-Jones (2007). *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods, second edition*. Springer.
- Daley, D.J. en D. Vere-Jones (2008). *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure, second edition*. Springer.
- Diggle, Peter J. (1983). *Statistical Analysis of Spatial Point Patterns, first edition*. Academic Press.
- KNMI (2015). *Seismologie, Aardbevingen in Nederland*. <http://www.knmi.nl/seismologie/aardbevingen-nederland.html>. [Online; bezocht op 01-05-2015].
- NAM (2010). “Bodemdaling door Aardgaswinning, Statusrapport 2010 and Prognose tot het Jaar 2070”. In: EP201006302236.
- NAM (2015). *NAM platform, feiten and cijfers*. <http://feitenencijfers.namplatform.nl/>. [Online; bezocht op 03-04-2015].
- Ogata, Y. (2013). “A Prospect of Earthquake Prediction Research”. In: *Statistical Science* 28.
- Pijpers, Frank P. (2014). “CBS Phase 0 Report 2: Significance of Trend Changes in Tremor Rates in Groningen”. In: *Statistics Netherlands*.
- Ripley, B.D. (1984). “Edge Corrections for Spatial Processes”. In: *Stochastic Geometry, Geometrical Statistics, Stereology*.
- Smith, Tony E. (2014). *Notebook for Spatial Data Analysis, Part 1: Spatial Point Pattern Analysis*.
- Szmyt, J. (2014). “Spatial Statistics in Ecological Analysis: From Indices to Functions.” In: *Silva Fennica* 48.1.
- TNO (2013). “Toetsing van de Bodemdalingsprognoses and Seismische Hazard ten gevolge van Gaswinning van het Groningen Veld”. In: TNO-2013 R11953.
- TNO (2015). *NL Olie- and Gasportaal*. <http://www.nlog.nl/>. [Online; bezocht op 01-05-2015].
- TNO Wassing, B.B.T. en B. Dost (2014). “Recent Developments of the Groningen Field in 2014 and, Specifically, the Southwest Periphery of the Field”. In: TNO-2014 R11703.
- TNO van Breunese, J. N. en K. Thienen-Visser (2014). “Effecten Verschillende Productiescenario’s op de Verdeling van Compactie in het Groningen Veld in de Periode 2014 tot and met 2016”. In: TNO-2014 R10426.
- Zhuang, J., Y. Ogata en D. Vere-Jones (2002). “Stochastic Declustering of Space-Time Earthquake Occurrences”. In: *Journal of the American Statistical Association* 97.458, pp. 369-380. DOI: <http://dx.doi.org/10.1198/016214502760046925>.

10 Bijlagen

In deze bijlagen is een deel van de gebruikte data opgenomen. De data van het KNMI met de gegevens over alle aardbevingen vanaf 1986 hebben wij hierin niet opgenomen. Deze informatie is vrij beschikbaar op www.KNMI.nl en wordt geregeld bijgewerkt met nieuwe bevingen. De precieze data over seismologie in Nederland is te vinden via KNMI 2015.

10.1 Data gaswinning

Hieronder staan de in het model gebruikte cijfers van de gaswinning per productielocatie. De cijfers worden gepubliceerd door NAM op www.feitenencijfers.namplatform.nl/gaswinning (NAM 2015). Zoals te zien in deze tabel zijn er een aantal locaties waar geen gas is gewonnen de afgelopen jaren. De cijfers voor 2015 zijn de winningscijfers tot en met april.

Tabel 5: Gaswinningscijfers uitgedrukt in $10^9 Nm^3$.

Put	X_RD	Y_RD	2010	2011	2012	2013	2014	2015 (t/m april)
Amsweer	256285	591519.7	2.57	2.36	1.85	2.98	3.03	0.9
Barnheem	241070	592308	0	0	0	0	0	0
Bedum	234000.2	591093	0	0	0	0	0	0
Bierum	254734.9	599603.7	3.26	3.31	3.08	2.86	4.02	1.17
De Eeker	259669.6	577338	1.29	1.05	0.97	1.5	1.75	0.60
De Paauwen	245963.3	588312.5	2.77	3.03	3.25	3.34	0.28	0.08
Eemskanaal	241580.8	584511.3	2.39	1.62	1.69	2.55	2.09	0.54
Froombosch	248204.5	578911.5	0	0	0	0	0	0
Kooipolder	246704	580929.5	2.74	2.06	2.25	2.83	3.17	0.80
Leermens	250083.5	596956.3	2.49	3.16	2.99	3.19	0.57	0.2 2
Midwolda	264223.3	578612.8	0	0	0	0	0	0
Nieuw-Scheemda	258975.8	580143.9	0	0	0	0	0	0
Noordbroek	255029.7	579596.1	0	0	0	0	0	0
Oldorp	238605	601801	0	0	0	0	0	0
Oudeweg	256082.8	585760.1	2.95	2.74	2.89	2.92	2.72	0.52
Overschild	250508.5	590804.5	2.64	3.09	2.84	3.33	0.62	0.27
Sappemeer	249482.1	575413.7	0	0	0	0	0	0
Schaapbulten	257449.5	588371.4	2.84	2.73	2.84	1.26	1.7	0.72
Scheemderzwaag	257306.5	578164.2	1.51	1.09	1.11	1.34	2.12	0.58
Siddeburen	253211.9	587434.3	2.9	2.85	3.24	3.05	3.18	0.90
Slochteren	246381.4	579250.2	3.06	2.53	2.66	2.84	3.14	0.77
Spitsbergen	252398.5	577196.9	2.53	1.77	1.86	2.81	2.29	0.61
Stedum	242518	594969	0	0	0	0	0	0
Ten Post	245664	591275.4	2.88	3.44	3.28	3.65	0.6	0.32
Tjuchem	254962.5	588214.2	2.33	2.69	2.78	2.62	2.98	0.86
Tusschenklappen	254366.1	575209.1	2.76	1.99	2.01	2.49	2.94	0.66
't Zandt	248027.1	600526	2.97	2.59	3.02	3.63	0.51	0.22
Uiterburen	255307.5	577395.6	0	0	0	0	0	0
Uithuizen	242633	605803	0	0	0	0	0	0
Uithuizermeden	249363	607901	0	0	0	0	0	0
Zuiderpolder	261906.8	580858.8	2.52	1.67	2.05	2.77	2.65	0.77
Zuiderveen	253140.7	579261.6	1.45	1.03	1.1	1.91	2.04	0.05

10.2 Data bodemdaling

Hieronder staan de in het model gebruikte data van de bodemdaling tussen 2003 en 2008 in gasveld Groningen. De data over bodemdaling is gepubliceerd door TNO en is te vinden op www.nlog.nl (TNO 2015).

Tabel 6: Peilmerken meegenomen in het model.

Peilmerknr	X_RD	Y_RD	daling [mm]
7F0261	256285.0	591519.7	14.6
7E0023	241070.0	592308.0	30.0
7B0111	234000.2	591093.0	16.2
3H0047	254734.9	599603.7	20.9
7H0255	259669.6	577338.0	16.6
7E0007	245963.3	588312.5	33.4
7G0045	241580.8	584511.3	24.3
7G0173	248204.5	578911.5	18.7
7G0165	246704.0	580929.5	21
7E0200	250083.5	596956.3	26.4
8C0195	264223.3	578612.8	12.4
7H0036	258975.8	580143.9	21.2
7H0271	255029.7	579596.1	17
7B0216	238605.0	601801.0	18
7H0052	256082.8	585760.1	21.6
7E0203	250508.5	590804.5	32.1
7G0221	249482.1	575413.7	15
7F0240	257449.5	588371.4	20.6
7H0256	257306.5	578164.2	16.1
7H0050	253211.9	587434.3	22.3
7G0238	246381.4	579250.2	19.6
7H0130	252398.5	577196.9	18
07E0031	242518	594969	29
7E0020	245664.0	591275.4	29
7F0132	254962.5	588214.2	20.8
12F0080	254366.1	575209.1	20.1
7E0062	248027.1	600526.0	23.1
7H0236	255307.5	577395.6	16.5
3G0012	242633	605803	16
3G0012	249363	607901	20
8C0224	261906.8	580858.8	17.3
7H0106	253140.7	579261.6	17.9

10.3 Data breuklijnen

Hieronder zijn de in het modelgebruikte data van de breuklijnen in gasveld Groningen weergegeven. In het rapport van TNO (TNO 2013) zijn figuren opgenomen met daarin de breuklijnen in kaart gebracht. De grootste breuklijnen hebben wij benaderd met rechte lijnstukken uit onderstaande tabel. De lijnsegmenten lopen van punt (x_1, y_1) naar (x_2, y_2) .

Tabel 7: Coördinaten van de begin- en eindpunten van de lijnstukken die de grootste breuklijnen modelleren.

x_1	y_1	x_2	y_2
239,5	609,5	248,0	594,0
246,5	597,5	266,0	597,5
227,7	601,2	252,0	576,0
256,7	577,6	266,0	577,8
255,0	590,0	252,0	596,0
243,0	583,5	253,0	583,0
244,0	593,5	238,0	600,0
244,0	592,0	254,0	589,0
261,0	585,0	262,0	590,0

10.4 R codes

In deze paragraaf is een selectie van de geschreven en gebruikte R codes weergegeven.

hulpfuncties.R

```
##### Aardbevingen inlezen

laad_grenzen_gasveld <- function() {
  library(maptools)
  gasveld_shape <- readShapeSpatial("gasvelden.shp")
  gasveld_shape <- subset(gasveld_shape, gasveld_shape@data$Field_na_1=="Groningen")
  gasveld_shape <- as(gasveld_shape, "SpatialPolygons")
  return(gasveld_shape)
}

laad_data_spatstat <- function(min_mag = 1.5) {
  data_bevingen <- read.csv("KNMI - Genduceerde bevingen vanaf 1986.csv", sep=";", quote="")
  data_bevingen <- subset(data_bevingen, data_bevingen$MAG>=min_mag &
    data_bevingen$YYMMDD>19950000 & data_bevingen$YYMMDD<20150000)

  data_bevingen$jaar <- as.factor(floor(data_bevingen$YYMMDD/10000))

  library(spatstat)
  observatiegrens <- as.owin(laad_grenzen_gasveld())

  # Herschalen m -> km
  data_bevingen$X_RD <- data_bevingen$X_RD/1000
  data_bevingen$Y_RD <- data_bevingen$Y_RD/1000
  observatiegrens <- rescale(observatiegrens, 1000, unitname="km")

  aardbevingen <- ppp(data_bevingen$X_RD, data_bevingen$Y_RD, window=observatiegrens, marks =
    factor(data_bevingen$jaar, levels=1995:2014))

  names(marks(aardbevingen)) <- "jaar"
  attr(aardbevingen, "rejects") <- NULL # verwijder punten buiten het gasveld
  return(aardbevingen)
}

# Filter alle bevingen niet in (jaar) uit de data (jaar kan een vector zijn).
bevingen_in_jaar <- function(ppp, jaar) {
  return(subset(ppp, ppp$marks %in% jaar))
}

aardbevingen_1995_2014 <- laad_data_spatstat()

##### Validatie functies

k_functie_test <- function(geschatte_intensiteit, realisatie) {
  KinhomEnvelop<-envelope(realisatie, Kinhom,
    nsim=99,simulate=expression(rpoispp(geschatte_intensiteit)),
    verbose=FALSE,clipdata=TRUE,global=FALSE,saveall=FALSE)
  plot(KinhomEnvelop,main="Ripley's K functie test", sub="", xlab="r", ylab="K(r)")
}

j_functie_test <- function(geschatte_intensiteit, realisatie) {
  JinhomEnvelop<-envelope(realisatie, Jinhom,
    nsim=99,simulate=expression(rpoispp(geschatte_intensiteit)),
    verbose=FALSE,clipdata=TRUE,global=FALSE,saveall=FALSE)
}
```

```

    plot(JinhomEnvelop,main="J functie test", sub="", xlab="r", ylab="J(r)")
  }

##### Exploratieve functies
aantal_bevingen_per_jaar <- function(data_bevingen) {
  return(table(data_bevingen$marks))
}

# Aantal bevingen per magnitude (vergelijk Gutenberg-Richter negatief exponentieel)
aantal_bevingen_per_magnitude <- function(data_bevingen) {
  return(aggregate(data.frame(count = data_bevingen$MAG), list(value = data_bevingen$MAG),
    length))
}

# Gutenberg-richter plot om magnitude < 1.5 te analyseren
plot_gr <- function() {
  betas <- coefficients(lm(log(aantal_bevingen_per_magnitude(bevingen1.5)$count) ~
    aantal_bevingen_per_magnitude(bevingen1.5)$value))
  plot(seq(0,3.6,0.1),exp(6.245 - 1.789 * seq(0,3.6,0.1)), col="red", xlab="magnitude",
    ylab="aantal bevingen in Groninger gasveld")
  points(aantal_bevingen_per_magnitude(bevingen))
  segments(1.5, 0, 1.5, 80, col="green")
} # we zien dat magnitude 1.1 echt het uiterste is.

```

alle_data.R

```

source("hulpfuncties.R")

observatiegrens <- as.owin(laad_grenzen_gasveld())
observatiegrens <- rescale(observatiegrens, 1000, unitname="km")

##### Bevingen verleden

#Lijst met pixel images aardbevingen 2001 - 2014
dens<-list()
for(i in 2001:2014)
{
  dens[[i-2000]]<-density(bevingen_in_jaar(aardbevingen_1995_2014,i), sigma=3.750)
}

# Hoge magnitudes zwaarder meerekenen. Werkt helaas niet in de huidige versie van spatstat.
bevingen_weights <- function(i){return(exp(1.788796*(marks(i)$MAG-1.5)))}

##### Bodemdaling

peilmerken <- read.csv("bodemdaling.csv", sep=";", quote="", header=TRUE)
peilmerken <- subset(peilmerken,peilmerken[,4]!="")
dalingen <- peilmerken[,4]

peilmerken <- ppp(peilmerken$X_RD/1000, peilmerken$Y_RD/1000,window=observatiegrens,
  marks=dalingen)

BD <- Smooth(peilmerken,sigma=3.750)
plot(BD)

##### Gaswinning

winning <- read.csv("NAM - gaswinning per locatie per jaar.csv", sep=";", quote="")

```

```
winning <- subset(winning, winning$X2010 > 0)

G14 <- Smooth(ppp(winning$X_RD/1000, winning$Y_RD/1000, window=observatiegrens,
marks=winning$X2014), sigma=3.500)
G13 <- Smooth(ppp(winning$X_RD/1000, winning$Y_RD/1000, window=observatiegrens,
marks=winning$X2013), sigma=3.500)
G12 <- Smooth(ppp(winning$X_RD/1000, winning$Y_RD/1000, window=observatiegrens,
marks=winning$X2012), sigma=3.500)
G11 <- Smooth(ppp(winning$X_RD/1000, winning$Y_RD/1000, window=observatiegrens,
marks=winning$X2011), sigma=3.500)
G10 <- Smooth(ppp(winning$X_RD/1000, winning$Y_RD/1000, window=observatiegrens,
marks=winning$X2010), sigma=3.500)
```

```
##### Breuklijnen
```

```
maak_breuklijn <- function(x1, y1, x2, y2) {
  return( psp(x1, y1, x2, y2, observatiegrens, check=FALSE)[observatiegrens])
}
```

```
breuklijn1 <- maak_breuklijn(239.500, 609.500, 248.000, 594.000)
breuklijn2 <- maak_breuklijn(246.500, 597.500, 266.000, 597.500)
breuklijn3 <- maak_breuklijn(227.700, 601.200, 252.000, 576.000)
breuklijn4 <- maak_breuklijn(256.700, 577.600, 266.000, 577.800)
breuklijn5 <- maak_breuklijn(255.000, 590.000, 252.000, 596.000)
breuklijn6 <- maak_breuklijn(261.000, 585.000, 262.000, 590.000)
breuklijn7 <- maak_breuklijn(244.000, 593.500, 238.000, 600.000)
breuklijn8 <- maak_breuklijn(244.000, 592.000, 254.000, 589.000)
```

```
breuklijnen <- superimpose.psp(breuklijn1, breuklijn2, breuklijn3, breuklijn4,
breuklijn5, breuklijn6, breuklijn7, breuklijn8, W=observatiegrens)
BL <- density(breuklijnen, sigma=3.500)
plot(BL)
```

CSR_en_Stationariteit.R

```
source("hulpfuncties.R")
```

```
ppp_bevingen <- laad_data_spatstat()
sigma <- 3.750 # km
```

```
##### Complete Spatial Randomness tests op het hele gasveld
```

```
# Monte Carlo (MonteCarlo) of Chi kwadraat (Chisq)
```

```
MC<-quadrat.test(ppp_bevingen,10,10,alternative=c("two.sided", "regular", "clustered"),
method="MonteCarlo",)
plot(MC,main="Monte Carlo, Gasveld Groningen", sub="Aardbevingen Groningen 1995-2014 met
magnitude >1.5")
```

```
# Ripley's K functions Test
```

```
K<-Kest(aardbevingen_1995_2014, correction="best")
envelop<-envelope(ppp_bevingen, fun=Kest, nsim=19, verbose=FALSE, clipdata=TRUE, global=TRUE, ginterval=c(0,9000), sav
plot(envelop,main="Ripley's K functie, Gasveld Groningen", sub="Aardbevingen Groningen
1995-2014 met magnitude >1.5", xlab="r", ylab="K(r)")
```

```
##### Stationariteit/Clustering tests
```

```
# Vergelijk de ruimtelijke spreiding van een serie jaren (bijv 1995-2000) met die in een
andere serie (bijv 2000-2005)
```

```

# De intensiteit wordt geschat op basis van de eerste jaren, geschaald om tijdseffecten te
  scheiden en dan vergeleken met de uitkomst

bevingen_1995_2000 <- bevingen_in_jaar(bevingen, 1995:2000)
bevingen_2000_2005 <- bevingen_in_jaar(bevingen, 2000:2005)

dens_est <- density(bevingen_1995_2000, sigma)
dens_est_2000_2005 <- dens_est * bevingen_1995_2000$n / bevingen_2000_2005$n

k_functie_test(dens_est_2000_2005, bevingen_2000_2005)
plot(KinhomEnvelop,main="Ripley's inhomogene K functie, Gasveld Groningen", sub="Aardbevingen
  Groningen 2000-2005 geschat o.b.v. 1995-2000", xlab="r", ylab="K(r)")

```

Exploratief_gaswinning.R

```

##### Inladen gaswinning
#C1: jaartal, C2: totale gaswinning, C3-C14: gaswinning per maand

GBB <- read.csv("GBB - Maandelijkse Cijfers Gaswinning.csv", sep=";", dec=",", header=T)
Gas <- GBB[1:52,1:14] #gegevens t/m 2014

gas0 <- Gas[33:52,2] #x0-waarden xi=gaswinning i jaar geleden
gas1 <- Gas[32:51,2] #x1-waarden
gas2 <- Gas[31:50,2] #x2-waarden
gas3 <- Gas[30:49,2] #x3-waarden
gas4 <- Gas[29:48,2] #x4-waarden

##### Aantallen aardbevingen
source("hulpfuncties.R")
aant_bevingen <- aantal_bevingen_per_jaar(aardbevingen_1995_2014)

##### Multiple Linear Regression: bevingen ~ gaswinning nu, gaswinning vorig jaar, etc..
quake.lm <- lm(aant_bevingen ~ gas0 + gas1 + gas2 + gas3 + gas4)
summary(quake.lm)

quake_est <- predict(quake.lm)

##### Plotjes

# Jaarlijkse gaswinning
plot(Gas[,1], Gas[,2],main="Jaarlijkse gaswinning Groningen", xlab="jaar", ylab="gaswinning")

# Aantal bevingen per jaar
plot(aant_bevingen, sub="Aardbevingen Gasveld Groningen 1995-2015 met magnitude>1.5",
  xlab="jaar", ylab="#bevingen")

# Gaswinning tegen aantal bevingen
plot(gas0,aant_bevingen, main="Gaswinning / #bevingen", sub="Aardbevingen Noord-Nederland
  1995-2015 met magnitude>1.5", xlab="Gaswinning", ylab="# bevingen")

# Regressie diagnostic plots
plot(quake.lm)

```

model.R

```

##### Bepalen coefficienten
# uitgaande van stationair in de tijd

```

```

# covariate: list met images

source("hulpfuncties.R")
source("alle_data.R")

# ppm herkent pixel images in lijsten niet. We moeten ze dus even kopiëren.
dens09 <- dens[[9]]
dens10 <- dens[[10]]
dens11 <- dens[[11]]
dens12 <- dens[[12]]
dens13 <- dens[[13]]

Q_10_14 <- bevingen_in_jaar(aardbevingen_1995_2014, 2010:2014)

regressie <- ppm(Q_10_14, trend=~ BD + BL
  + ifelse(marks==2010, G10,
    ifelse(marks==2011, G11,
      ifelse(marks==2012, G12,
        ifelse(marks==2013, G13,
          G14))))
  + ifelse(marks==2010, dens09,
    ifelse(marks==2011, dens10,
      ifelse(marks==2012, dens11,
        ifelse(marks==2013, dens12,
          dens13))))))
show(summary(regressie))

# De 'regressie' ppm is niet echt bruikbaar voor voorspellen, dat is me ten minste niet gelukt.
# Maak dus een nieuwe ('dummy') ppm. Hier gaan we bij het voorspellen de coëfficiënten van
overschrijven.
dummy <- ppm(bevingen_in_jaar(aardbevingen_1995_2014, 2010), trend=~ BD+BL+G+D,
  covariates=list(BD=BD, BL=BL,G=G10,D=dens09))

# voorspelde intensiteit (in dit geval voor 2014)
schatting_2014 <- predict.ppm(dummy, covariates=list(BD=BD,BL=BL, G=G14,D=dens13), new.coef =
  coef(regressie))

##### Validatie

k_functie_test(schatting_2014, bevingen_in_jaar(aardbevingen_1995_2014, 2014))
j_functie_test(schatting_2014, bevingen_in_jaar(aardbevingen_1995_2014, 2014))

##### Voorspellen

forecast<-function(gaswinning_voorspelling) {
  fore <-dens #fore wordt de lijst met pixelimages van 2001 tot te voorspellen jaar
  n <- length(dens)
  for(i in 1:length(gaswinning_voorspelling)) { #in deze loop vullen we lijst 'fore' aan
    fore[[n+i]] = predict.ppm(dummy, covariates=list(BD=-BD,BL=BL,
      G=gaswinning_voorspelling[[i]], D=dens[[i]]), new.coef=coef(regressie))
  }

  return(fore)
}

gaswinning_zoals_2013 <- forecast(list(G13, G13, G13))
gaswinning_zoals_2014 <- forecast(list(G14, G14, G14))
geen_gaswinning <- forecast(list(0, 0, 0))

plot(gaswinning_zoals_2014[[17]])

```

```

##### Animatie

# Bepaal n schaal voor een lijst met afbeeldingen
bepaal_z_range <- function(density_lijst) {
  return(range(unlist(lapply(density_lijst, function(x){summary(x)$range}))))
}

# Plot alle afbeeldingen als animatie
animatie <- function(voorspelling_lijst) {

  library(animation)
  for (i in 1:length(voorspelling_lijst)) {
    plot(voorspelling_lijst[[i]], main=bquote("(Voorspelde) intensiteit in jaar" ~
      .(2000+i)), xlim = bepaal_z_range(voorspelling_lijst))
    ani.pause()
  }
  ani.pause()
  ani.options(oopt)
}

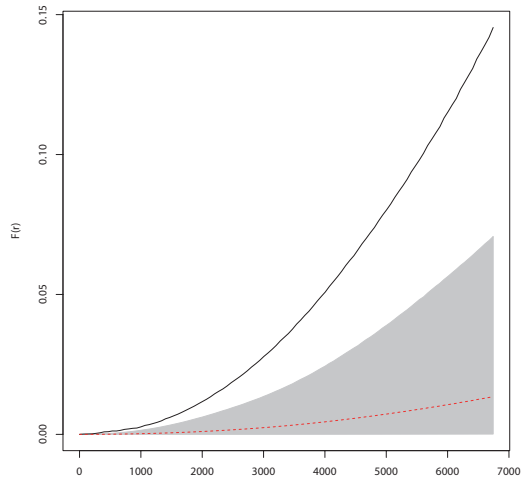
animatie(gaswinning_zoals_2014)

# Animatie opslaan als .Latex
# Zie pagina 78: http://cran.r-project.org/web/packages/animation/animation.pdf
saveLatex({animatie()}, img.name="voorspelling_zoals_2014")

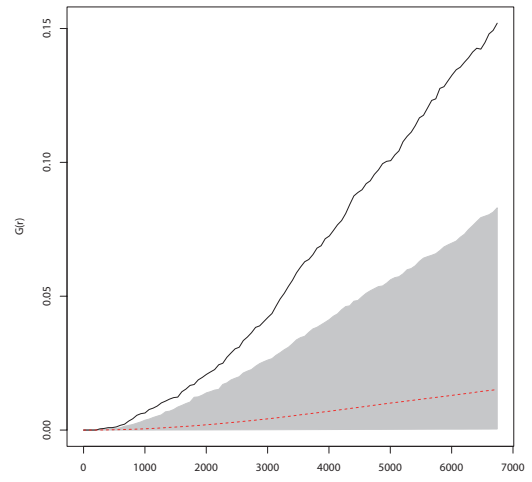
```

10.5 Toetsen met F, G, en J-functie

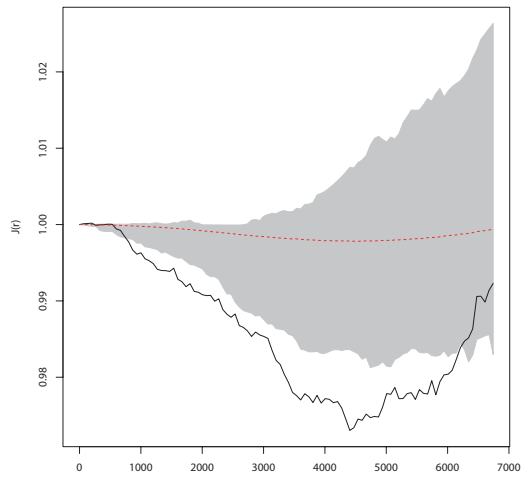
In paragraaf 4.1 worden een aantal toetsen genoemd. Hier zijn de bijbehorende figuren afgebeeld.



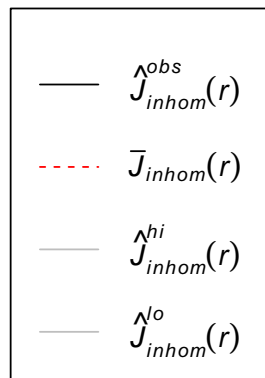
(a) Inhomogene F-functie



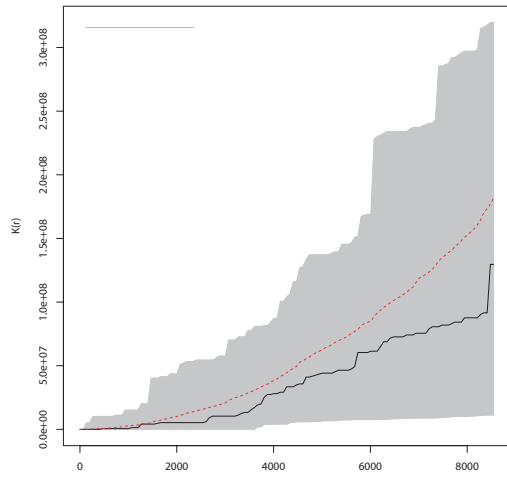
(b) Inhomogene G-functie



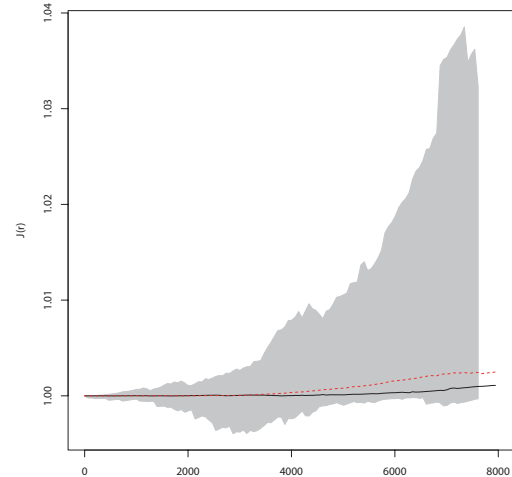
(c) Inhomogene J-functie



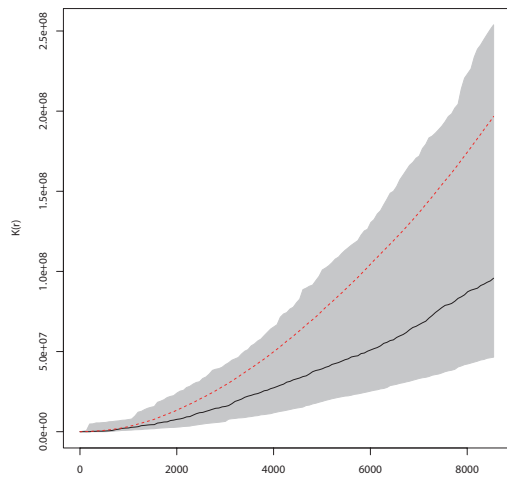
Figuur 37: Gasveld Groningen. Intensiteit geschat op basis van periode 1995-1999, getoetst over periode 2010-2014.



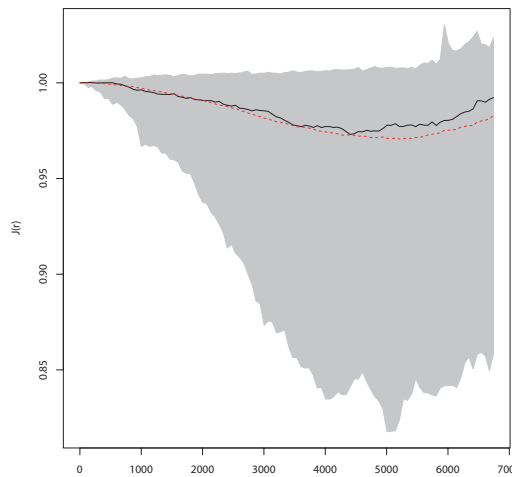
(a) Inhomogene K-functie in de periode 1995-1999



(b) Inhomogene J-functie in de periode 1995-1999



(c) Inhomogene K-functie in de periode 2010-2014



(d) Inhomogene J-functie in de periode 2010-2014

Figuur 38: Gasveld Groningen. Intensiteit geschat op basis van perioden 1995-1999 en 2010-2014, getoetst over dezelfde periodes.