Statistics Netherlands

2016 Scientific Paper

Trend changes in tremor rates in Groningen

update Nov 2016

The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.

Frank P. Pijpers

Contents

Nederlands

Deze rapportage behelst een voortzetting van onderzoek dat is uitgevoerd sinds midden 2014 in het kader een onderzoeksproject door het CBS in opdracht van Staatstoezicht op de Mijnen (SodM). Dit onderzoek is ten behoeve van een statistische onderbouwing van het meet- en regelprotocol voor gasexploitatie in de provincie Groningen.

In dit rapport ligt de aandacht op een analyse van de tijden en locaties van aardbevingen die worden gerapporteerd door het Koninklijk Nederlands Meteorologisch Instituut (KNMI) gebaseerd op hun analyses van de gegevens verzameld door het netwerk van seismometers dat het KNMI beheert. Deze analyse is een uitbreiding van het onderzoek eerst gerapporteerd in 2014, met actualisaties in 2015, in de zin dat aardbevingen tot eind augustus 2016 in de catalogus van het KNMI zijn meegenomen in de analyse. In deze actualisatie ligt de focus op de verschillen in aardbevingsfrequentie voor en na de datum van 1 maart 2015. Deze datum is gekozen omdat de totale maandelijkse productie door gaswinning vanaf begin 2015 minder in de tijd is gevarieerd dan in de periode daarvoor. Voor een aantal gaswinningsclusters is de productie al langer vlak en laag. Met behulp van een Monte Carlo analyse kan worden bepaald dat het aantal aardbevingen na 1 maart 2015 statistisch significant lager is dan het zou zijn geweest wanneer de trend van de periode daarvoor zou zijn voortgezet.

Het uitgangspunt voor deze analyse is om zoveel mogelijk data gedreven te zijn en onafhankelijk van modellen. In combinatie met de eerder gepubliceerde analyses is een direct causaal verband tussen productievariaties en frequentie van bevingen voldoende plausibel als werkhypothese, maar de analysestappen worden uitgevoerd zonder gebruikmaking van deze hypothese. In sectie 5 van dit rapport wordt wel nader ingegaan op een mogelijk verband tussen bevingsfrequenties en de winning van gas en is er dus sprake van empirische modellen.

English

This report is a continuation of research, commenced in 2014, which is part of a research project being carried out by Statistics Netherlands and commissioned by State Supervision of Mines (SodM). This research is part of the underpinning of the statistical methods employed to support the protocol for measurement and regulation of the production of natural gas in the province of Groningen.

In this report, the focus is on an analysis of the times and locations of earthquakes as reported by the Royal Netherlands Meteorological Institute (KNMI) based on their processing of the network of seismometers that they manage. This analysis is an update of the reports of December 2014 and 2015 in the sense that it uses earthquake data recorded by the KNMI up to the end of August 2016. In this update the focus lies on differences in tremor frequencies before and after March 1 2015. This date is chosen because the total monthly gas production in the epoch starting in 2015 has varied much less in time than in the previous epoch. For a few specific clusters the production has been flat and low for longer. A Monte Carlo analysis is employed to demonstrate that the rate at which earthquakes occur after March 1 2015 is significantly lower than would be expected under a null hypothesis that the rate follows the same trend as before that date.

This analysis was purposely set up to be data driven and as much as feasible to remain model-independent. In combination with previously published analyses, a direct causal connection between production variations and tremor frequencies is sufficiently plausible that

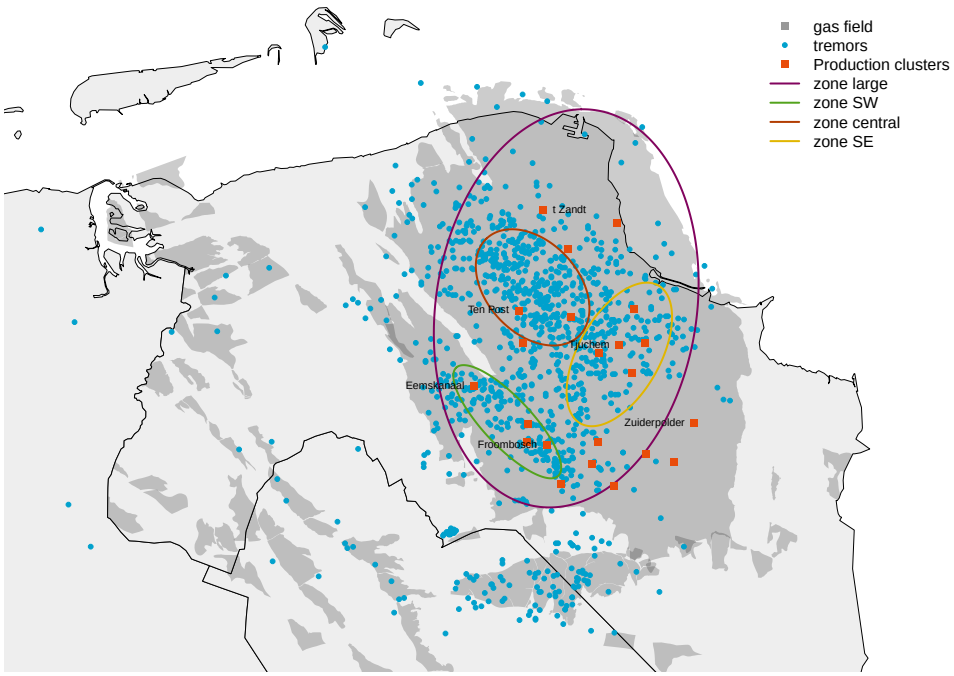
it may be used as a working hypothesis. However, the various steps in the analysis reported here do not require making use of this hypothesis. In section 5 of this report a relationship between tremor rates and the extraction volumes of gas is explored and so there empirical models are involved.

1 Introduction

For some decades earthquakes of modest magnitudes have occurred in the Groningen gas field. It is recognized that these events are induced by the production of gas from the field. Following an $M_L = 3.6$ event near Huizinge, and the public concern that this raised, an extensive study program has started into the understanding of the hazard and risk due to gas production-induced earthquakes.

A protocol needs to be established with the aim of mitigating these hazards and risks by adjusting the production strategy in time and space. In order to implement this regulation protocol and adaptively control production it is necessary also to measure the effects on subsidence and earthquakes in order to provide the necessary feedback.

Figure 1.1 The locations of earthquakes as reported by the KNMI. The red squares are locations of the production clusters, some of which are identified by name. The purple ellipse `zone large' demarks the reference area for earthquake rates. The red and green smaller ellipses (central and SW respectively) mark the two regions of interest also reported on in previous reports. The yellow ellipse (SE) is an additional region first considered in the report of Nov 2015. The production field is also shown in dark gray, overplotted on a map of the region



The causality of the earthquakes induced by gas production is likely to be through the interaction of compaction of the reservoir rock with existing faults and differentiated geology of the subsurface layers. The ground subsidence occurs because with the extraction of gas, pressure

support decreases in the layer from which the gas is extracted. The weight of overlying layers then compacts that extraction layer until a new pressure equilibrium can be established, cf. Dake (1978); Doornhof et al. (2006). The technical addendum to the winningsplan Groningen 2013 "Subsidence, Induced Earthquakes and Seismic Hazard Analysis in the Groningen Field" (Nederlandse Aardolie Maatschappij BV, 2013) discusses all of these aspects in the context of the Groningen reservoir in much more detail.

The seismic network of the KNMI has been in operation for some decades, and detailed reporting on and (complete) data for earthquakes in the Groningen region are available from 1991 onwards. The locations of all earthquakes in the region are shown in fig. 1.1, together with the locations of the gas production clusters. Also indicated are the boundaries of the regions for which the earthquake rates are determined in this report, which are the same as in the previous semi-annual updates (cf. Pijpers (2014, 2015a,c, 2016b))

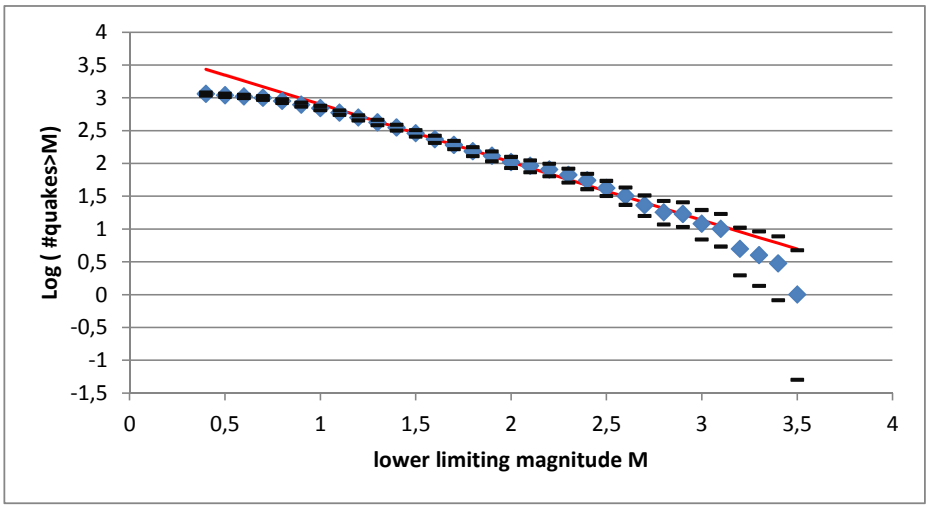
In this technical report, the available earthquake data are examined for a signature of changes in rates. The analysis procedure is unchanged from previous reports, cf. Pijpers (2014, 2015a,c, 2016b), and is presented as well as the conclusions one can draw from this phase of the research project. In a sense this is therefore a classical approach, as opposed to for instance the Bayesian approach reported in Nepveu et al. (2016).

2 Background

2.1 The earthquake data

The available earthquake dataset contains in total 1293 events recorded after 1 Jan. 1991 up to 1 Sep. 2016. Of these, there are 896 that are located within the zone indicated as 'zone large' in fig. 1.1. An earthquake magnitude and time of event as well as the KNMI's present best estimate of the longitude-latitude position is available for each of the earthquakes. The KNMI has indicated that the network of seismometers was designed to be complete in terms of both detection and localisation of earthquakes in the Groningen region above magnitudes of 1.5. Above magnitude 1.0 the detection is fairly secure, but the localisation may be more problematic. Starting towards the end of 2014 an upgrade to the network of seismometers has been implemented which has pushed down these limiting magnitudes for completeness and localisation. This may also have consequences for analyses such as these, because a better detection rate will imply that more will be recorded in the catalog which is only an apparent increase in the rate of tremors. In previous analyses carried out by CBS cf. Pijpers (2014, 2015a,c, 2016b), all data collected since 1995 has been used, but there might be some issues with completeness for the earlier years. While it is unlikely that such issues, even if present, materially affect those analyses, for the analysis reported here all data from before Sep. 1 2004 are excluded. This leaves a total of 934 tremors of which 719 occurred within the ellipse 'zone large' that can be included in the analysis. Furthermore, comparisons between epochs, and also the tests of the various hypotheses discussed here, will be done taking only tremors with magnitudes greater than $M = 1$ into account. While this is therefore slightly lower than the completeness limit quoted, the data itself suggest that incompleteness becomes severe only at lower magnitudes than $M = 1$. With this restriction in time and magnitude, there are data on 605 $M \geq 1$ earthquakes available. Of these, there are 466 within the large ellipse shown in fig. 1.1.

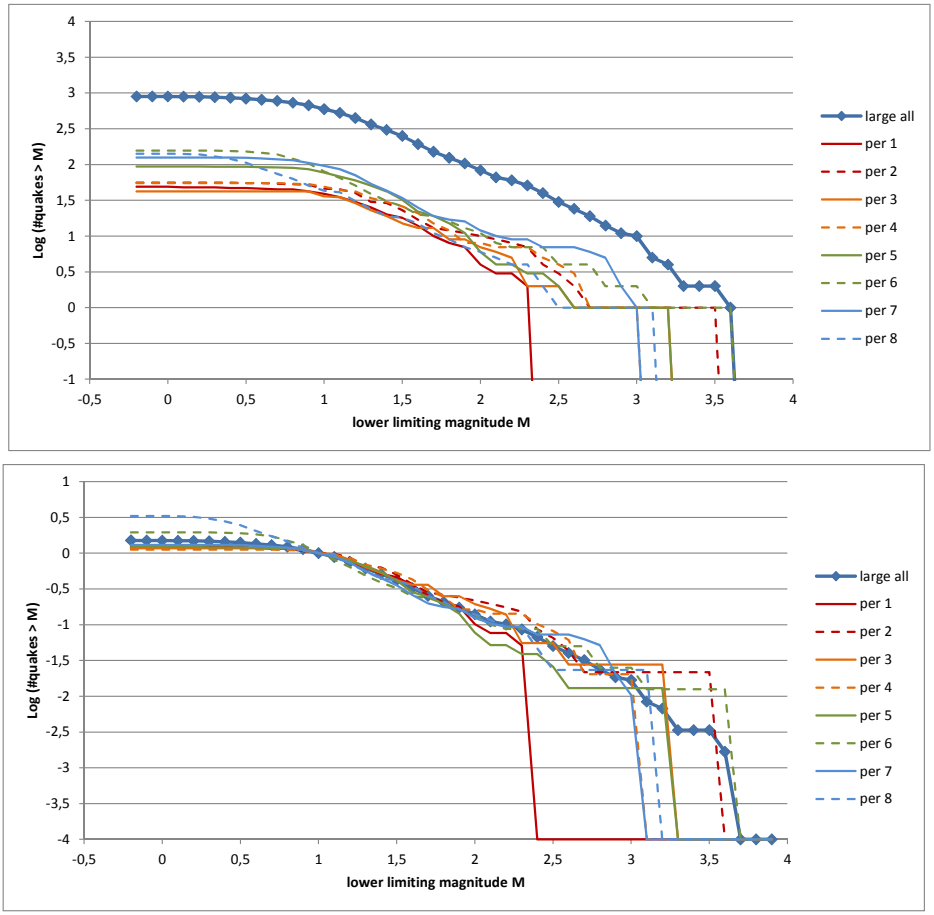
Figure 2.1 The logarithmic cumulative magnitude distribution of earthquakes for all earthquakes in the set. The red line is a linear function with a slope of -0.9 similar to values reported elsewhere (Dost et al., 2012). 95% confidence intervals are indicated under the assumption that the underlying process obeys Poisson statistics.



It is evident from fig. 1.1 that the distribution of events is not uniform over the area under consideration. It is also known that the distribution function of earthquakes is not uniform as a function of magnitude. For all 896 quakes in the catalog since 1995 that occurred within zone large the distribution is shown in fig. 2.1. The way in which this is plotted is in a cumulative form: all earthquakes with a magnitude above a lower limit are counted and the base-10 logarithm of that count is shown as a function of the lower limiting magnitude. As this limiting magnitude increases there are fewer and fewer earthquakes with magnitudes above that limit, so this is a cumulative distribution function (or cdf) when reading the figure from right to left. This is a commonly used way to represent earthquakes in the field, known as frequency-magnitude or Gutenberg-Richter plot. The horizontal lines indicate margins of 95% confidence under the assumption that within each interval of the cumulative distribution in quake magnitude the value obeys Poisson statistics (e.g. Garwood (1936)). The statistics of induced earthquakes is not well known, which implies that using margins of confidence from a particular probability distribution function such as the Poisson distribution may well be inappropriate. Towards higher values of the lower limiting magnitude, the margins of uncertainty become larger because there are fewer events on which to build the statistics.

Also shown in fig. 2.1 is a linear function with a slope of -0.9, i.e. very close in value to the results of Dost et al. (2012) and an offset selected to match the range $1.1 < M_L < 3.1$. This shows that the slope of the distribution function appears to be constant over this range. For lower limiting magnitudes the distribution function is systematically lower than the straight line. The apparent 'deficit' of earthquakes with very low magnitudes is known to be indicative of the limitations of the sensitivity of the seismometer network. If tremors of such small magnitudes occur too far away from any of the seismometers in the network the signal becomes indistinguishable from noise or cannot be located with sufficient accuracy. For tremors with magnitudes below about 1.0 the 'missing' smaller earthquakes or tremors probably do occur but the detection of such events is no longer complete. The KNMI may use a higher value than this lower limit, such as 1.5, when taking into account not only the magnitude as is done here but also an accurate localisation of the events, which requires that a positive detection is available from at least 3 seismic wells in order to carry out the triangulation.

Figure 2.2 The logarithmic cdf-s of earthquakes for the 8 consecutive equal length time intervals, covering the period from 1 Sep. 2004 to 1 Sep. 2016. Upper panel: before normalization, lower panel: after normalization at $M = 1$.

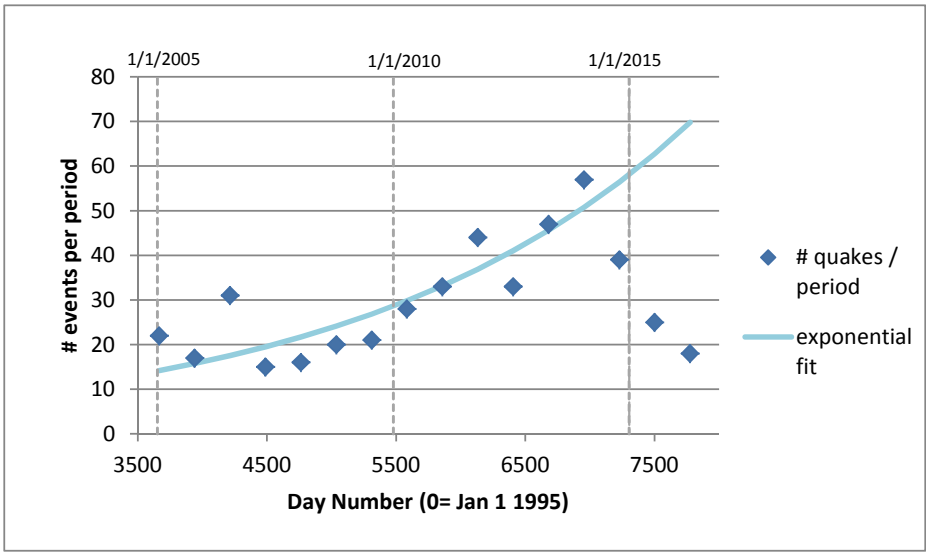


The catalog of quake events is likely also to contain events that are aftershocks. This means that some fraction of events has not occurred completely independently from preceding ones, which implies that it is inappropriate to assume Poissonian statistics. This is considered in more detail in section 4.

Further, in the strict sense a Poissonian process is homogeneous (i.e. must not sense absolute time), which implies that it should be stationary in time. A process can be inhomogeneous and quasi-Poissonian if the change in the time parameter of the process is very slow, and a separation of time scales may allow modelling as a combination of Poisson processes with different time parameters. If the quakes were due to a stationary stochastic process this implies that if the dataset were to be divided into subsets that are equal in length of time, the number of quakes in each subset should not show any significant difference. To test the assumption, the dataset is divided into eight equal sections, and the logarithmic cdf is determined separately for each subset, numbered consecutively from 1 to 8. Subdividing the dataset into a larger or smaller numbers of subsets has also been tested, which confirms that this behaviour does not depend on the precise boundaries between the divisions or the number of divisions used. Note that the 8 divisions used here are different epochs from those of the previous CBS reports, although there is of course overlap.

From fig. 2.2 it is clear that the number of earthquakes increases with time, and therefore the underlying process cannot be a stationary stochastic Poissonian process. In principle this might

Figure 2.3 The total number of earthquakes for 16 consecutive subsets of equal length, obtained by splitting in two each of the time intervals used above. Also shown is an exponential fit to these points, omitting the third point and the final three.



have been due to a detection effect: if over time the seismometer network has expanded, or the individual seismometers have become more sensitive due to upgrades, or the processing has improved to reduce noise levels, the dataset would contain more earthquakes towards later times, because more of the smaller earthquakes are being detected. However, if this were the case, the magnitude at which the cdf bends over should have moved progressively to lower magnitudes. In the most recent years, an upgrade of the network of seismometers has been realised, which is also reflected in the cdf for period 8 in fig. 2.2. It is clear that in this epoch, covering March 2015 up to Sep. 1 2016, the flattening of the slope from higher to lower magnitudes occurs at lower magnitude than in all other epochs. This is exactly what would be expected from the increased sensitivity of the network. At magnitudes well above the completeness limit of around 0.8 the cdf-s should not show any trends with time if the process were stationary.

It is evident from fig. 2.2 that the shape of the distribution functions is very similar between the subsets, and that they appear to simply shift upwards from every period to the next, except for the most recent period, which has as its initial date March 1 2015. Here it is evident that the network sensitivity has increased so that in particular below magnitude $M = 0.8$, more tremors are detected. If the distribution functions are normalized, by dividing by the total number of quakes with $M \geq 1$ in each period, fig. 2.2 is obtained. From this it can be seen that there is very little change in the shape of the distribution functions. At the high end, with lower limiting magnitudes of around 2.5 and higher, the number of events is so small that there may simply be no such events in a given period. Hence the distribution functions for the different periods have different cut-offs at the high end. The near invariance of the shape of the distribution functions since Sep 2004 means that the increased rate is unlikely to be due to improved sensitivity of the network apart from the period after March 1 2015, and more likely to be due to a genuinely increasing rate of quake events.

Fig. 2.3 shows the total number of events, where each of the 8 periods considered is further split into two. Also a fit to these points is shown of the form $A\exp(t/\tau)$. Two sets of fit parameters are determined : once using all points except the final two, covering the most recent period. For the other fit, two further points are omitted (i.e. given 0 weight) which are the third and third

from last points. The characteristic timescale τ that is determined from the former fit of the function to these data, indicates that the rate of quake events doubles roughly every 5.5 years. The latter fit, which is the one shown, implies a shorter doubling time of 4.3 years. Both a least-squares and a maximum likelihood fitting has been performed, with the same result, within the uncertainty of 0.2 years, of the doubling time.

A straightforward method to analyse the behaviour of rate changes of tremor events would be to divide the time axis into sections of several hundred days (e.g. half a year or less), and for each section to count the number of events, with magnitudes above a fixed threshold. This is similar to what is done in fig. 2.3 but more fine-grained. Such a time series would have more sampling points than fig. 2.3 allowing applying standard time series analysis techniques. However, when this is done it becomes clear that the number of events per section is no higher than a few tens at best. This has the consequence that assessing the statistical significance of trend changes becomes so sensitive to the unknown properties of the underlying distribution function, produced by the process that generates the tremors, that no meaningful conclusions can be drawn. For this reason such a straightforward approach was abandoned, and a Monte Carlo technique was adopted.

2.2 Monte Carlo simulations

The data indicate that the process by which the earthquakes arise is neither stationary in time, nor homogeneous in spatial distribution over the area. This prevents applying the statistics of Poissonian processes to asses whether in particular subregions the rate of earthquakes has altered, following the reduction in production. However, it is possible to use the dataset itself to test various hypotheses. This is done by means of a technique referred to in the literature as bootstrapping or Monte Carlo simulation. Extensive descriptions and applications of this technique can be found e.g. in textbooks by Robert and Casella (2004), Tarantola (2004).

Since in each simulation all the 466 earthquakes with $M \geq 1$ are assigned, the same limitations apply to the simulations as apply to the real data. A close similarity in this aspect, between the artificial and real data, is an essential requirement for the method to function. In the present case the technique is applied in order to test several hypotheses. The way one proceeds is to use the magnitude of the 466 events as recorded and reported by the KNMI. For the simulations, the location and timing of each event are not used. Instead locations and timings are assigned stochastically, using a random number generator and a pre-set likelihood for an event to belong to a certain group. In the present case there are eight relevant groupings, constructed by a subdivision in time and subdivisions in space :

1. A grouping in time : the event *either* occurs in the period from Sep. 1 2004 up to March 1 2015, *or* it occurs in the period from March 1 2015 to Sep. 1 2016.
2. A grouping in space : the event occurs either within the contours of the area marked ‘zone SW’ in fig. 1.1, or within the area marked ‘zone central’, or within ‘zone SE’, or not in any of these regions, but within the area marked as ‘zone large’ in fig. 1.1.

The eight groups are obtained by events within each zone occurring in either the first or the second time range. There are several null hypotheses that are tested within the scope of this research. The most simple hypothesis is that, despite appearances, the probability for an event to occur is constant over the entire domain ‘zone large’ and also constant in time. Under this null

hypothesis the likelihood for an event to occur within each of the three spatial groups is simply proportional to the area of each zone. Also, the likelihood for an event to occur in the first or the second of the two time ranges is proportional to the length of each range. The combined likelihoods are obtained assuming independence i.e. by straightfoward multiplication of the likelihoods for the spatial divisions and for the division in time.

The next step is to assign each event (quake magnitude) to one of the eight groups using a random number generator twice: once to decide which of the spatial groups to assign the event to, and once to decide which period. After all 466 events are assigned, a cdf can be constructed for each group. This assigment process is repeated a large number of times, for the present case 10000 repetitions was considered sufficient, since there does not appear to be a need determine the simulated number of quakes N_{sim} and the standard deviation σ to more than 3 significant digits. Using these 10000 simulations an average distribution function for quake magnitudes can be constructed for each group, as well as 95% and 99% confidence limits, because each of the 10000 simulations will produce a different realisation from the stochastic assignment.

Other likelihoods than the ones described above can be assigned as well, giving rise to different null-hypotheses for testing. The measured / true distribution in space and in time of all 466 events can then be used in each case to test wether the null-hypothesis can be rejected or not. The total number of events for each group is shown in table 2.1. The following sections present the results for 5 separate null-hypotheses.

Note that by proceeding in this way, the only assumption that is made about the stochastic properties of the physical processes underlying the generation of earthquakes, is that the events are independent. The presence of considerable numbers of aftershocks in the catalog would violate this assumption. The consequences of that are explored in section 4. By using the bootstrapping technique it is possible to circumvent the necessity of having a spatiotemporal model for the generation of tremors and aftershocks. In particular, by using the earthquake magnitudes of the 466 actual events the distribution functions for magnitudes can be simulated. The detected total number of tremors in each zone can thus be compared directly with the percentiles of the Monte Carlo distributions for the total numbers which directly translates to whether given percentile confidence limits are exceeded, for each group, without requiring a model for the rate at which quakes with magnitudes of any particular strength will be produced. In all cases the proper limits (percentiles) for the probability distribution function as determined from the Monte Carlo simulations are used as (non-)rejection criterion.

2.3 Null-hypothesis I: homogeneous and stationary process

From section 2.1 it does not appear very probable a-priori that quake events are spread uniformly over the area of interest and that there is no time dependence in the rate at which quakes occur. Nevertheless it is useful to present these results as a measure of the capability of the Monte Carlo approach to test hypotheses. Also, the relevant probabilities are a useful reference to asses by how much quake rates are enhanced or lowered in the other models. Under this null hypothesis the likelihood for an event to occur within each of the three spatial groups is simply proportional to the area of each zone. Also, the likelihood for an event to occur in the first or the second of the two time ranges is proportional to the length of each range. The combined likelihoods are obtained assuming independence i.e. by straightfoward multiplication of the likelihoods for the spatial divisions and for the division in time.

Table 2.1 The measured total number of quake events since Sep. 1 2004, for each group and the probabilities for assignment to each group for homogeneous and stationary test case.

region	period	Number of events	probability
zone SW	before 01-03-2015	60	0.0596
	after 01-03-2015	8	0.0086
zone central	before 01-03-2015	149	0.0985
	after 01-03-2015	9	0.0141
zone SE	before 01-03-2015	65	0.1080
	after 01-03-2015	12	0.0155
zone large	before 01-03-2015	148	0.6084
(not SW, SE or central)	after 01-03-2015	15	0.0873

Table 2.2 Simulated number of quake events for each group for homogeneous and stationary test case, and standardised difference. Entries 'bef' and 'aft' in the column period refer to before and after the date of March 1, 2015. The columns 99%l and 99%u refer to the lower and upper 99% confidence levels, rounded to the nearest integer, and the column 95%l and 95%u refer to the same for the 95% confidence levels.

region	period	N_{sim}	σ	$\frac{(N_{true}-N_{sim})}{\sigma}$	99% l	95% l	95% u	99% u
zone SW	bef	27.7	5.10	6.3	15	18	38	41
	aft	4.0	1.99	2.0	0	1	8	10
zone central	bef	45.9	6.43	16.0	30	34	59	63
	aft	6.6	2.52	1.0	1	2	12	14
zone SE	bef	50.4	6.74	2.2	33	38	64	68
	aft	7.3	2.64	1.8	2	3	13	15
zone large	bef	283.5	10.55	−12.8	255	262	304	310
(not SW, SE or central)	aft	40.7	6.13	−4.2	26	29	53	57

Using the probabilities shown in table 2.1, the cdf-s of quake magnitudes are determined. The total number of events in each group can be compared directly, and tested for significance, with the true numbers shown in table 2.1. From the simulations a mean value and a standard deviation can be determined, and also the 1%, 5%, 95%, and 99% percentiles of the distributions of total numbers of events.

The mean and standard deviation for the 10000 simulations are shown in table 2.2, as well as the standardised difference between the measured and simulated total number of quake events for each group. While the distribution of the simulated data does not conform exactly to a normal distribution, a value larger than ~ 2 for the standardised difference implies a statistically significant deviation at the 95% level at least in most cases, except for very small total counts. The standardised differences are a good enough indicator here to see directly that, apart from the result for the groups ‘zone central ’ and ‘zone SE’ after March 1 2015, this null hypothesis is strongly rejected. The standardised differences are shown for illustrative purposes; proper limits (percentiles) for the probability distribution function as determined from the Monte Carlo simulations are used as (non-)rejection criterion. The null hypothesis is rejected at a confidence level of 99%.

2.4 Null-hypothesis II: non-homogeneous and stationary process

More of interest for the problem at hand is to test the null-hypothesis that the rate at which quakes occur has not changed with time, but that the spatial distribution of that rate is not homogeneous: there is an enhanced likelihood in the various regions of interest. Geophysical modelling of the subsurface and the response of existing fractures to pressure changes might in future enable predicting a rate, but at present the true probability is not known with high precision. For this operational reason in the Monte Carlo simulation the probability is assigned according to the proportions of the true total number of events in each region, combined for both before and after March 1 2015.

Table 2.3 Probabilities for assignment to each group for non-homogeneous and stationary test case. For convenience the numbers of true events are repeated.

region	period	probability	Number of events
zone SW	before 01-03-2015	0.1276	60
	after 01-03-2015	0.0183	8
zone central	before 01-03-2015	0.2965	149
	after 01-03-2015	0.0425	9
zone SE	before 01-03-2015	0.1445	65
	after 01-03-2015	0.0207	12
zone large	before 01-03-2015	0.3059	148
(not SW, SE or central)	after 01-03-2015	0.0439	15

Table 2.4 Simulated number of quake events for each group for non-homogeneous and stationary test case, and standardised difference. The columns are as in table 2.3

region	period	N_{sim}	σ	$\frac{(N_{true}-N_{sim})}{\sigma}$	99% l	95% l	95% u	99% u
zone SW	bef	59.4	7.16	0.1	41	45	74	78
	aft	8.5	2.88	−0.2	2	3	15	17
zone central	bef	138.1	9.89	1.1	113	119	157	164
	aft	19.9	4.27	−2.5	10	12	29	32
zone SE	bef	67.4	7.65	−0.3	48	53	83	88
	aft	9.6	3.07	0.8	3	4	16	18
zone large	bef	142.5	9.93	0.5	118	124	162	168
(not SW, SE or central)	aft	20.5	4.46	−1.2	10	12	30	33

Comparing table 2.3 with table 2.1, the probability for quakes to occur within zone SW is now enhanced by a factor of ~ 2.1 over the homogeneous value, and for the zone central the probability is enhanced by a factor of roughly ~ 3.0. For zone SE there is a more modest enhancement of a factor of 1.3. Using these probabilities, shown in table 2.3, the cdf-s of quake magnitudes are again determined, following the same procedures as in section 2.3. The total number of events in each group which can be compared directly, and tested for significance, with the true numbers also shown in table 2.3. From the simulations the mean value and the standard deviation is shown in table 2.4.

The mean and standard deviation for the 10000 simulations are shown, as well as the standardised difference between the measured and simulated total number of tremor events for each group. As one would expect this null-hypothesis is better in the sense that it is not rejected for more groups. However, there is still strong rejection of this hypothesis for the central zone for the period after March 1 2015. The overall number of tremors for all zones combined after March 1 2015 is also lower than the number of simulated events, and therefore also too low before March 1 2015. Effectively this means that the tremor rate since March 1 2015 for tremors

with $M \geq 1$, is lower than the average rate between Sep. 1 2004 and March 1 2015, and this result is statistically significant.

2.5 Null-hypothesis III: non-homogeneous and exponentially increasing process A

From the discussion in section 2.1 it is clear that the stationary null hypothesis also does not appear very realistic. Using the number of earthquakes recorded in each of the 16 successive periods discussed in section 2.1 (fig. 2.3) , one can re-assess the likelihood for earthquakes to occur after March 1 2015 by extending the trend over the past years. Using the fit discussed, with the doubling time $\tau = 5.5$ years, the likelihoods can be re-determined for each of the 8 groups and Monte Carlo simulations produced to test whether this time dependence, together with the same enhanced likelihoods in the regions of interest is consistent with the data. Comparing the probabilities for a quake to occur after March 1 2015 from table 2.5 with the probabilities of the previous section (table 2.3), shows that this probability is now higher by a factor of roughly 2.

Table 2.5 Probabilities for assignment to each group for non-homogeneous and exponentially increasing test case with $\tau = 5.5$ years. For convenience the numbers of true events are repeated.

region	period	probability	Number of events
zone SW	before 01-03-2015	0.1103	60
	after 01-03-2015	0.0357	8
zone central	before 01-03-2015	0.2562	149
	after 01-03-2015	0.0828	9
zone SE	before 01-03-2015	0.1249	65
	after 01-03-2015	0.0404	12
zone large	before 01-03-2015	0.2643	148
(not SW, SE or central)	after 01-03-2015	0.0855	15

Table 2.6 Simulated number of quake events for each group for non-homogeneous and exponentially increasing test case, and standardised difference. The columns are as in table 2.3

region	period	N_{sim}	σ	$\frac{(N_{true}-N_{sim})}{\sigma}$	99% l	95% l	95% u	99% u
zone SW	bef	48.6	6.54	1.7	32	36	61	66
	aft	19.4	4.28	−2.7	9	11	28	31
zone central	bef	112.9	9.28	3.9	89	95	131	137
	aft	45.1	6.32	−5.7	30	33	58	62
zone SE	bef	55.1	7.06	1.4	37	41	69	74
	aft	22.0	4.58	−2.2	11	13	31	35
zone large	bef	116.5	9.37	3.4	92	98	135	141
(not SW, SE or central)	aft	46.5	6.44	−4.9	30	34	60	64

From the final five columns in table 2.6 it can be seen that for most regions this null hypothesis is rejected at 99% confidence, with one exception. The zone SE, in the period before March 1 2015 had a number of tremors that does not lead to rejection of the hypothesis. The total number of quakes over all regions, after March 1 2015, is lower by a statistically significant amount compared to the general increasing trend. The null hypothesis of continuation of this trend is rejected at a confidence level of 99%.

2.6 Null-hypothesis IV: non-homogeneous and exponentially increasing process B

In fig. 2.3 a fit is shown with a doubling time of 4.3 years which is even shorter than the 5.5 years of the previous subsection, and which was also found in the previous reports, which included data as far back as 1995. From the previous subsection it is evident that this null hypothesis is also rejected, but it is illustrative to nevertheless present the results in the same format.

Table 2.7 Probabilities for assignment to each group for non-homogeneous and exponentially increasing test case with $\tau = 4.3$ years. . For convenience the numbers of true events are repeated.

region	period	probability	Number of events
zone SW	before 01-03-2015	0.1043	60
	after 01-03-2015	0.0416	8
zone central	before 01-03-2015	0.2424	149
	after 01-03-2015	0.0967	9
zone SE	before 01-03-2015	0.1181	65
	after 01-03-2015	0.0471	12
zone large	before 01-03-2015	0.2500	148
(not SW, SE or central)	after 01-03-2015	0.0997	15

The probabilities and simulated events are shown in tables 2.7 and 2.8. Comparing with the probabilities from the previous sections it is clear that the probability for an event to fall after March 1 2015 is enhanced by a factor of roughly 2.3 over the stationary case, and of course slightly higher than is predicted when using the doubling time of 5.5 years of the previous subsection.

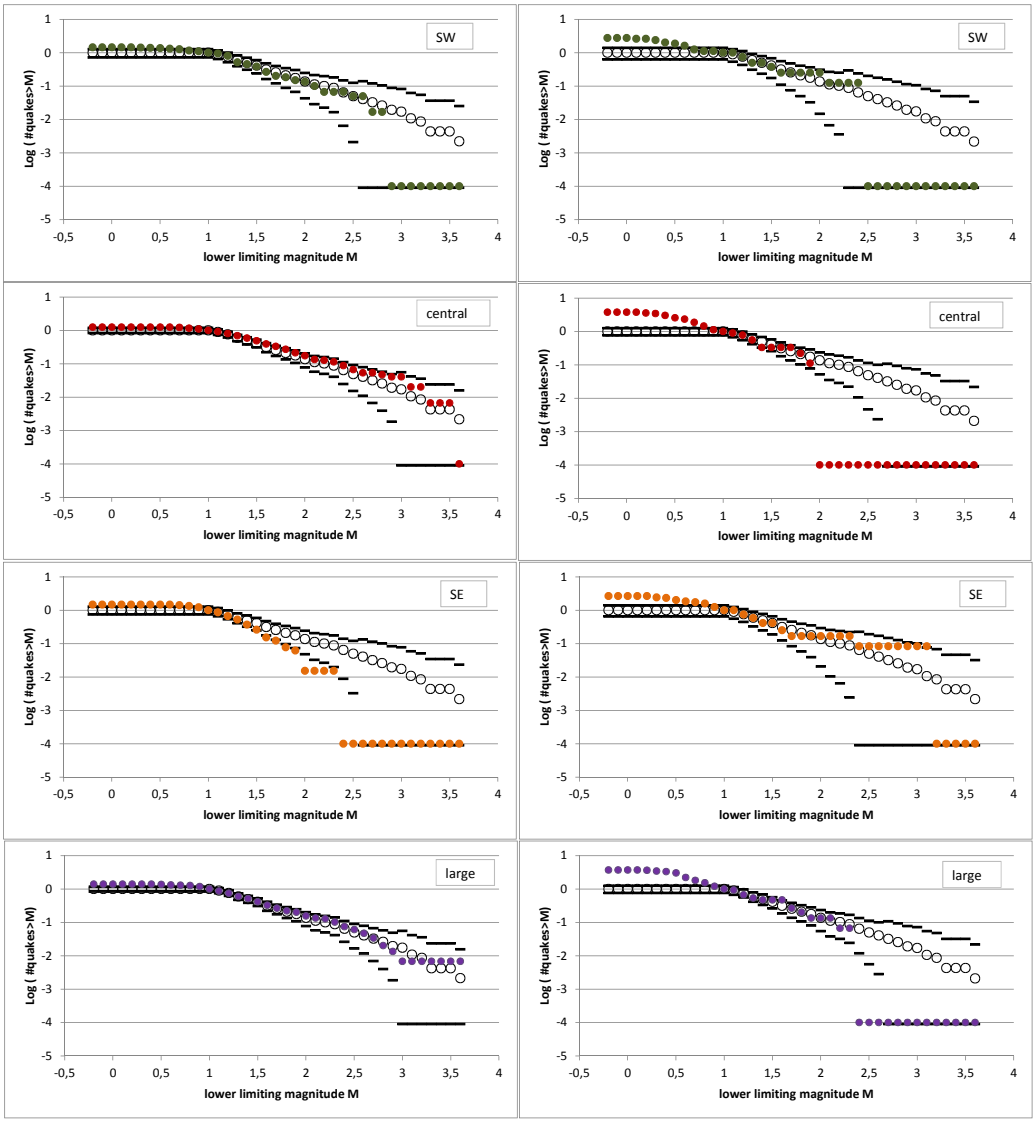
Table 2.8 Simulated number of quake events for each group for non-homogeneous and exponentially increasing test case, and standardised difference. The columns are as in table 2.3

region	period	N_{sim}	σ	$\frac{(N_{true}-N_{sim})}{\sigma}$	99% l	95% l	95% u	99% u
zone SW	bef	44.3	6.29	2.5	29	32	57	61
	aft	23.8	4.78	−3.3	12	15	34	37
zone central	bef	102.8	8.95	5.2	80	86	120	126
	aft	55.2	7.03	−6.6	38	42	69	74
zone SE	bef	50.0	6.66	2.2	33	37	63	68
	aft	26.9	5.05	−3.0	15	18	32	37
zone large	bef	106.0	9.02	4.7	83	88	124	129
(not SW, SE or central)	aft	57.0	7.07	−5.9	40	44	71	76

While this exponentially increasing trend with the short doubling time of 4.3 years appears to fit quite well to the rates shown in fig. 2.3, the predicted numbers of tremors are now rejected at well over 99% confidence for all regions, except zone SE where there is rejection at 95%..

The simulated cdfs for the 4 regions, before and after March 1 2015, are shown in fig. 2.4. These are of additional interest because the shape of the cdfs of tremor magnitudes of the real data, also shown in each panel, depart from those of the simulations in some cases. Most notable is that compared to the simulations the zone central appears to have a distinct deficit in tremors with $M > 2$. On the other hand, the zone SE which had a slight relative deficit of such tremors before March 1 2015, now appears to have 'caught up'. The cdfs are normalized at magnitudes $M = 1$ with the result that the improved sensitivity of the network after March 2015 shows as an

Figure 2.4 The cdfs of earthquakes for the 8 spatio-temporal groups, from the bootstrapping procedure for this hypothesis, as well as the real data, normalised at $M = 1$. Panels in left column are before March 1 2015, right column panels are after March 1 2015. Open/ black symbols are the simulation results and the coloured filled dots are the data.



excess at lower magnitudes. This illustrates why in this paper, in the comparisons of data with simulations, only tremors with $M > 1$ are counted.

2.7 Null-hypothesis V: non-homogeneous process with reverted rates except for zone central

Of all the null-hypotheses considered thus far in this report, the best performing one is null-hypothesis II. In the light of the time dependence shown in fig. 2.3 this must be interpreted in the sense that the (exponentially) increasing trend is not continuing, and that after March 1 2015 the tremor rate has dropped back down to a value, roughly equal to its long term (ten year) average value. However, for the central zone, where gas production was already reduced a year earlier and thus has remained flat and low for longer, the rate appears to have dropped lower

still. One may for instance suppose a tremor rate that is half that ten year average. The resulting probabilities (table 2.9) and simulation results (table 2.10) are shown.

Table 2.9 Probabilities for assignment to each group for non-homogeneous test case with rate reverting to long term average after March 1 2015, except for zone central where the rate is half this average rate. For convenience the numbers of true events are repeated.

region	period	probability	Number of events
zone SW	before 01-03-2015	0.1276	60
	after 01-03-2015	0.0183	8
zone central	before 01-03-2015	0.3178	149
	after 01-03-2015	0.0213	9
zone SE	before 01-03-2015	0.1445	65
	after 01-03-2015	0.0207	12
zone large	before 01-03-2015	0.3059	148
(not SW, SE or central)	after 01-03-2015	0.0439	15

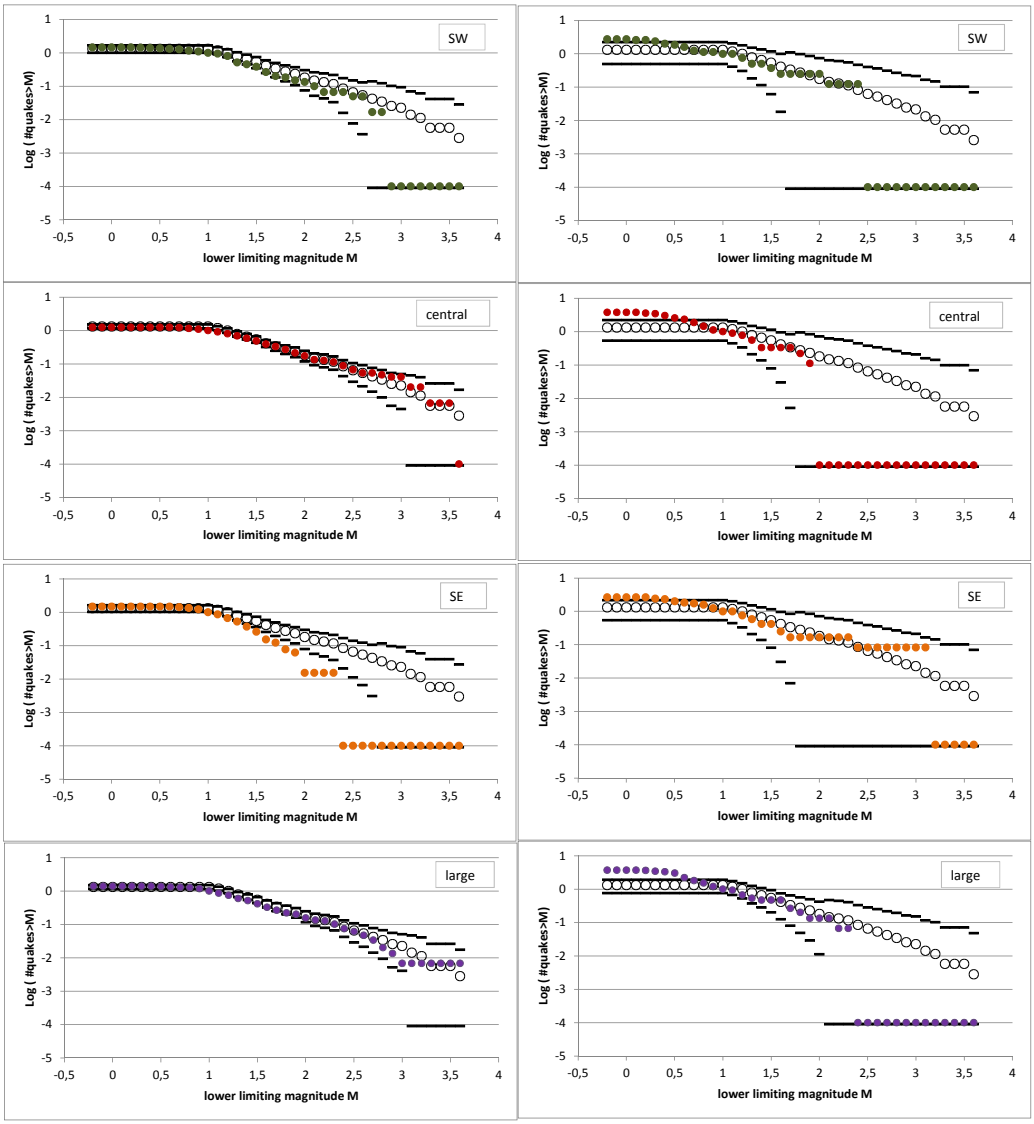
The mean value of the simulated total number of events after March 1 2015 is slightly larger than the number of actually recorded events. However, the numbers in each zone, before and after March 1 2015 are now consistent with the actually recorded numbers of events, hence this hypothesis is not rejected.

Table 2.10 Simulated number of quake events for each group for non-homogeneous and exponentially increasing test case, except for zone central where the rate stabilises after March 1 2015, and standardised difference. The columns are as in table 2.3

region	period	N_{sim}	σ	$\frac{(N_{true}-N_{sim})}{\sigma}$	99% l	95% l	95% u	99% u
zone SW	bef	59.5	7.22	0.1	42	46	74	79
	aft	8.5	2.88	−0.2	2	3	14	17
zone central	bef	148.1	10.08	0.1	122	128	168	174
	aft	9.9	3.17	−0.3	3	4	17	19
zone SE	bef	67.3	7.57	−0.3	49	53	82	87
	aft	9.7	3.08	0.8	3	4	16	18
zone large	bef	142.5	9.85	0.6	118	124	162	168
(not SW, SE or central)	aft	20.4	4.49	−1.2	10	12	30	33

Under this hypothesis the number of tremors in each region after March 1 2015 is rather lower than under hypothesis IV, which means that the margins of uncertainty around the shape of the cdfs is wider. This is shown in fig. 2.5. From this one can see that in zone central, the deficit of $M > 2$ tremors is not statistically significant, as it was under hypothesis IV. In other words, with the reduced overall likelihood of tremors occurring, the absence of stronger tremors is not beyond what would be expected. However, for zone SE the same is seen as under hypothesis IV: where there appears to have been a somewhat lower than expected number of tremors with $M > 2$ in the ten years before March 1 2015, there is now no such deficit. The reduction in the steepness of the cdf may not yet be statistically significant but certainly warrants monitoring.

Figure 2.5 The cdfs of earthquakes for the 8 spatio-temporal groups, from the bootstrapping procedure for this hypothesis, as well as the real data, normalised at $M = 1$. Panels in left column are before March 1 2015, right column panels are after March 1 2015. Open/ black symbols are the simulation results and the coloured filled dots are the data.



3 The influence of incompleteness

3.1 The exclusion of tremors with magnitudes below 1

In the previous reports, cf. Pijpers (2014, 2015a,c, 2016b), the full catalog of events was used, including a range of low magnitudes where it is likely that not all events have been detected. From comparison of the shape of the cdfs before and after March 1 2015, it appears that this is likely to play a role in particular at tremor magnitudes $M < 0.8$. For this reason, in all analyses for this report the simulations have been done for the magnitude range $M \geq 1$ and compared with counts of real tremors in the same range. In the previous reports, attention was paid to this limited range as well as to the full range, so it is straightforward to make comparisons of the previous results with what is reported here.

In the previous report Pijpers (2016b) the doubling time for tremors with $M \geq 1$ was deduced to be around 6.2 years. This is in contrast with the shorter time scales deduced in this report, but the previous reports included data from 1995 onwards. It appears likely that the shorter timescales deduced in this report reflect that in the period from 1995 to 2004 the time scales of increase were much longer overall, and have become shorter around or after 2004.

3.2 Excluding tremors with magnitudes below 1.5

While fig. 2.1 appears to indicate that incompleteness becomes a serious issue only below magnitudes of 1, it is known that tremors with magnitudes between 1 and 1.5, although they are detected, are often difficult to localise because the signal exceeds the noise at only 1 or 2 seismic wells which means standard triangulation is impossible. The lower resulting spatial accuracy of the catalog at these magnitudes might also influence the statistics. For this reason the analysis is repeated, excluding all tremors with magnitudes below 1.5. In total there are then 192 tremors left in the catalog since Sep. 1 2004.

Table 3.1 Probabilities for assignment to each group for hypothesis V test case. Only tremors with magnitudes 1.5 .

region	period	probability	Number of events
zone SW	before 01-03-2015	0.1184	23
	after 01-03-2015	0.0170	3
zone central	before 01-03-2015	0.3710	73
	after 01-03-2015	0.0248	3
zone SE	before 01-03-2015	0.1002	17
	after 01-03-2015	0.0144	5
zone large	before 01-03-2015	0.3097	61
(not SW, SE or central)	after 01-03-2015	0.0444	7

Table 3.2 Simulated number of quake events, with magnitudes > 1.5 , for each group for the hypothesis V test case, with standardised difference. The columns are as in table 2.3

region	period	N_{sim}	σ	$\frac{(N_{true}-N_{sim})}{\sigma}$	99% l	95% l	95% u	99% u
zone SW	bef	22.8	4.46	0.1	12	14	32	35
	aft	3.3	1.79	−0.2	0	0	7	9
zone central	bef	71.2	6.63	0.3	54	58	84	88
	aft	4.8	2.19	−0.8	0	1	9	11
zone SE	bef	19.2	4.15	−0.5	10	12	28	31
	aft	2.8	1.66	1.3	0	0	6	8
zone large	bef	59.5	6.40	0.2	43	47	72	77
(not SW, SE or central)	aft	8.5	2.89	−0.5	2	3	15	17

If only the best-localised tremors with magnitudes above 1.5 are taken into account, the spatial enhancements in the zones SW and central become 2.0 and 3.5 respectively, but for zone SE it is now 0.9 (i.e. a lowering rather than an enhancement). The equivalent of hypothesis II can be rejected at a 99% confidence level for zone central, since there are very few tremors after March 1 2015. Hypothesis V, (see tables 3.1 and 3.2), is not rejected.

4 The influence of aftershocks

It is possible that some of the tremors in the catalog, even at magnitudes higher than 1 or 1.5, are events that are triggered by preceding tremors. This means that there is some finite correlation, both in time and in space, in the likelihood for a tremor to occur. This likelihood for a tremor to occur close in time and space to a previous tremor is then slightly in excess of what it would be if each event occurred completely independently from all previous events. This would mean that the fluctuations around a mean trend or inhomogeneities in spatial distribution are somewhat higher than a random assignment simulation produces. Conversely, the confidence limits used to determine whether a particular deviation is statistically significant must then be appropriately enlarged, from what is obtained from simulations that do not take correlations into account.

Table 4.1 Probabilities for assignment to each group for hypothesis V test case. Only tremors with magnitudes > 1, and excluding all events within 10 days and 5 km of a previous event recorded in the catalog. The relevant recorded number of events are given in the final column.

region	period	probability	Number of events
zone SW	before 01-03-2015	0.1304	42
	after 01-03-2015	0.0187	6
zone central	before 01-03-2015	0.2794	90
	after 01-03-2015	0.0187	6
zone SE	before 01-03-2015	0.1358	42
	after 01-03-2015	0.0195	8
zone large	before 01-03-2015	0.3476	117
(not SW, SE or central)	after 01-03-2015	0.0499	11

Table 4.2 Simulated number of quake events, with magnitudes > 1 and excluding all events within 10 days and 5 km of a previous event recorded in the catalog, for each group for non-homogeneous and exponentially increasing test case and standardised difference. The columns are as in table 2.3

region	period	N_{sim}	σ	$\frac{(N_{true}-N_{sim})}{\sigma}$	99% l	95% l	95% u	99% u
zone SW	bef	41.9	5.94	0.0	27	30	54	58
	aft	6.0	2.41	−0.0	1	2	11	13
zone central	bef	90.0	8.00	0.0	70	74	106	111
	aft	6.0	2.48	−0.0	1	2	11	13
zone SE	bef	43.7	6.20	−0.3	28	32	56	60
	aft	6.3	2.48	0.7	1	2	11	13
zone large	bef	112.0	8.49	0.6	90	96	129	134
(not SW, SE or central)	aft	16.1	3.92	−1.3	7	9	24	27

The Monte Carlo simulations used for this paper do not have such an excess of correlation. In principle it would be possible to introduce this, for instance through adding a Markov chain process to the simulations, with a finite probability for a tremor to be flagged as an aftershock in the simulations, and then assigned an appropriate location and time relatively close to the preceding tremor rather than completely at random. However, this would require a knowledge of the likelihood for an earthquake of a given strength to produce an aftershock, and distribution functions for the distances and times between progenitor and aftershocks. Relevant methods of analysis reported in the literature are Huc and Main (2003) and Naylor et al. (2009), or a modelling approach for aftershock generation (Kumazawa and Ogata, 2014) to simulate data. Progress on the analysis using these methods is to be reported at a later stage.

An alternative approach is to exclude from the catalog any event that is sufficiently close in space

and in time to a preceding event, so that one might reasonably suppose that it could be an aftershock. There are a number of windowing methods, the simplest of which is described in Gardner and Knopoff (1974). There is some arbitrariness in the possible choice of parameters but to operationalise this exclusion criterion, shocks are designated as aftershocks, and therefore excluded from the analysis, if they occur within 10 days of an earlier entry in the catalog, and are located within a distance of 5 km of that entry. The numbers of events remaining after applying the exclusion criteria are shown in table 4.1. There are other methods for identifying aftershocks, e.g. as described in Baiesi and Paczuski (2004). Using that method as a check, a set of events are identified as aftershocks which to a large degree overlaps with the set using the criteria described here, and which therefore would produce comparable results.

The probabilities for tremor rates in accordance with hypothesis V are shown in table 4.1, and the simulation results in table 4.2. This model is similar to the one discussed in section 2.7, but excluding potential aftershocks. The spatial enhancement factors in the zones, SW, central, and SE are 2.2, 2.6, and 1.3 respectively. This model is not rejected, whereas the equivalent of hypothesis II is rejected for the central zone at a 99% confidence level..

In sum there is no strong evidence that the existence of aftershocks, and the correlation between events that this produces, affects the data to such a large extent that the confidence limits produced by the Monte Carlo simulations are a severe underestimate. Thus the comparison of the results from the various hypotheses appear to point to a genuine reduction in the rate of generation of tremors.

5 Gas production as explanatory variable

5.1 Linear frame rate effect

There is a possibility that the tremor rates have been reduced, merely because the overall production rate is reduced, but that the total number of tremors will eventually reach the same values if a particular total cumulative production is to be reached. This is what is commonly referred to as a 'frame rate' effect. If $N(t \geq t_0)$ is defined as the cumulative number of tremors satisfying the selection criteria in terms of magnitudes and not being identified as aftershocks, where the first such shock is catalogued at time t_0, and $Q(t \geq t_0)$ is the cumulative amount of gas produced, then:

$$N(t \geq t_0) = \alpha Q(t \geq t_0) \tag{1}$$

In such a scenario, the expectation value for the number of quakes generated per unit volume of extracted gas is constant and equal to α in Eq. (1). There are several ways in which this hypothesis can be tested. One possibility is to assign probabilities for tremors to be generated in each region, proportional to the volumes of gas extracted at those production clusters which are connected to the various tremor regions by the flow lines in the field. In order to explore this hypothesis, the production clusters are grouped as follows:

1. Eemskanaal, Froombosch, Kooipolder, Sappemeer, Slochteren, Spitsbergen, Tusschenklappen, Zuiderveen

2. Leermens, Overschild, de Pauwen, ten Post
3. Amsweer, Schaapbulten, Siddeburen, Tjuchem

There are other production clusters ('t Zandt, Bierum, De Eeker, Scheemderzwaag, and Zuiderpolder), but the flow lines in the field (cf Nederlandse Aardolie Maatschappij BV (2016)) suggest that these primarily drain different regions of the reservoir. The three groups above are associated with the region SW, central and SE respectively. It should be noted however, that gas within the reservoir can flow into or out of each of the spatial regions, even if all the relevant clusters for that region are not producing. The grouping of clusters should therefore reflect the dominant influences of the gas production but may nevertheless give an incomplete picture.

Table 5.1 Probabilities for assignment to each group for production related test case. Production in normalized gas volumes is shown in column 3, where for zone large the total Groningen production is used. Only tremors with magnitudes > 1, and excluding all events within 10 days and 5 km of a previous event recorded in the catalog. The relevant recorded number of events are given in the final column.

region	period	gas prod. $[10^9 m^3]$	probability	Number of events
zone SW	before 01-03-2015	106.8	0.1307	42
	after 01-03-2015	15.03	0.0184	6
zone central	before 01-03-2015	105.4	0.2922	90
	after 01-03-2015	2.151	0.0060	6
zone SE	before 01-03-2015	111.7	0.1360	42
	after 01-03-2015	15.81	0.0193	8
zone large	before 01-03-2015	412.3	0.3577	117
(not SW, SE or central)	after 01-03-2015	45.91	0.0398	11

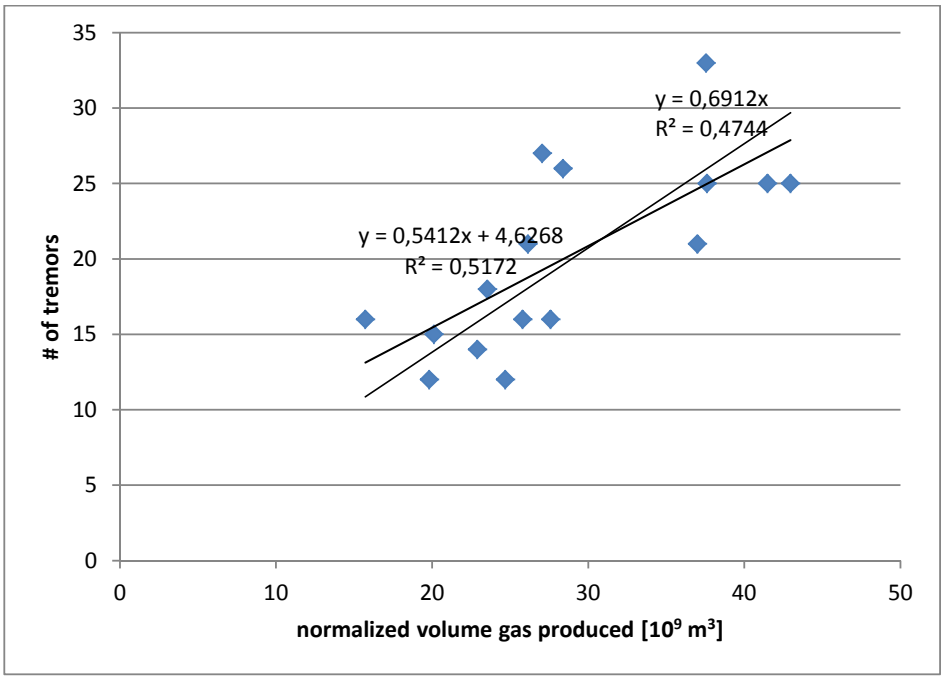
From the previous reports, cf. Pijpers (2014, 2015a,c, 2016b), a delay of roughly 3 months between the 2014 production changes and the tremor rate changes was deduced. Nepveu et al. (2016) conclude on the basis of the seasonality patterns of seismicity and production variations that perhaps a somewhat wider range of delays is valid, depending on which areas and which clusters are involved. Taking this into account, the probabilities are determined by using for each group of clusters the ratio of the production totals for the period from Dec. 1 2014 to Jun. 1 2016, and the cumulative totals since Jul. 1 2004. For this analysis in particular it appears important to exclude potential aftershocks, so the same selection criteria are applied as in section 4. These probabilities are shown in table 5.1, which are clearly very similar to those of section 4. The largest shift occurs for the zone central. The simulation results to test this hypothesis are shown in table 5.2. From this it appears that for almost all regions and epochs, this hypothesis is not rejected. There is one exception, which is the central zone. This appears to have experienced more tremors than is predicted by this hypothesis, leading to a rejection at a confidence level of 99%.

While it appears from this result that a linear 'frame rate' effect is an incomplete explanation for the tremor counts, it could nevertheless be a partial explanation. An alternative way to investigate this is to use the counts similar to those shown in fig. 2.3, but again excluding potential aftershocks in the same way as described in section 4, and relate these to the total gas production over successive 9 month windows, again displaced by the 3 month time delay. From taking the time derivative, or finite differences, of both left- and right-hand side of eq. (1) it can be seen that the same relationship should hold between tremor rate and production rate. In fig. 5.1 the data are plotted. A pure frame rate effect should have no offset at production rates of 0, so such a fit is shown, as well as a standard linear fit with a 0-point offset. The fit without a 0-point offset has a goodness-of-fit R^2 measure of 0.47, whereas the fit with a 0-point offset

Table 5.2 Simulated number of quake events, with magnitudes > 1 and excluding all events within 10 days and 5 km of a previous event recorded in the catalog, for each group for production related tremor rates. The columns are as in table 2.3

region	period	N_{sim}	σ	$\frac{(N_{true}-N_{sim})}{\sigma}$	99% l	95% l	95% u	99% u
zone SW	bef	42.1	6.08	-0.0	27	31	55	58
	aft	5.9	2.40	0.0	1	2	11	13
zone central	bef	94.0	8.09	-0.5	74	78	110	116
	aft	1.9	1.40	2.9	0	0	5	6
zone SE	bef	43.9	6.12	-0.3	29	32	56	60
	aft	6.2	2.45	0.7	1	2	11	13
zone large	bef	115.2	8.64	0.2	93	98	132	137
(not SW, SE or central)	aft	12.8	3.55	-0.5	5	6	20	23

Figure 5.1 The total number of earthquakes with $M > 1$ and excluding potential aftershocks, for 16 consecutive subsets of equal length, plotted against the gas production. Linear fits to the data are also plotted, with and without a 0-point offset.



included is slightly better with an $R^2 = 0.52$. The 0-point offset of 4.6 would suggest that even in the absence of gas extraction, a 'background level' of some 4 or 5 tremors every 9 months should be expected. This is somewhat counter-intuitive. Both this offset and the relatively low R^2 measures of either fit appear to suggest that the tremor rates are not purely due to a frame rate effect of the type described by eq. (1). Additionally, delays of 2 and 4 months between gas production and tremor rates have also been assessed, but these lead to lower R^2 measures.

5.2 More complex relationships

An alternative scenario to the linear frame rate could be that eq. (1) is merely the lowest order approximation, and that in fact the relationship is:

$$N(t \geq t_0) = f\,(Q(t \geq t_0)) \tag{2}$$

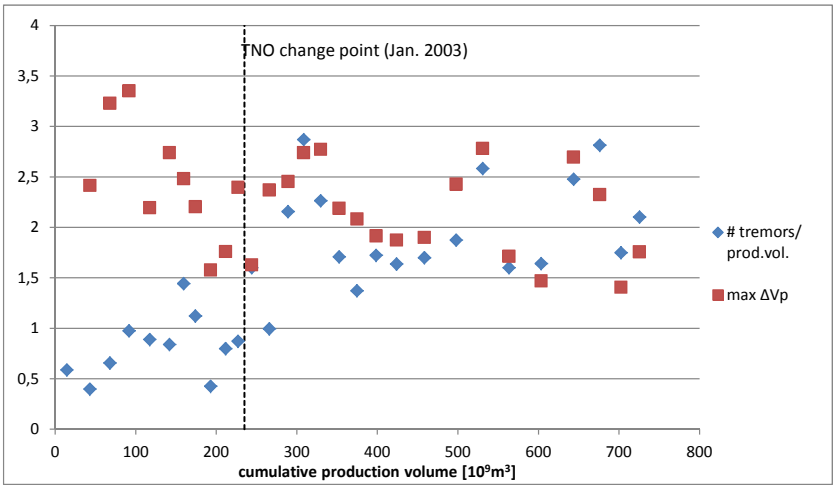
where f is a monotonically increasing function and $f(0) = 0$. In this case the tremor and production rates, N_t and V_p respectively, are related by:

$$\frac{\partial N(t \geq t_0)}{\partial t} \equiv N_t = \frac{df(Q)}{dQ}\frac{\partial Q}{\partial t} \equiv \frac{df(Q)}{dQ}V_p \tag{3}$$

which can be rearranged as:

$$\frac{N_t}{V_p} = \frac{df(Q)}{dQ} \tag{4}$$

Figure 5.2 Blue diamonds: The number N_t of earthquakes with $M > 1$ (excluding potential aftershocks), per produced volume of gas in the same time window, for 28 consecutive subsets of equal length, plotted against the cumulative gas production. Red squares: the maximum month-to-month increase in V_p within each of the windows.

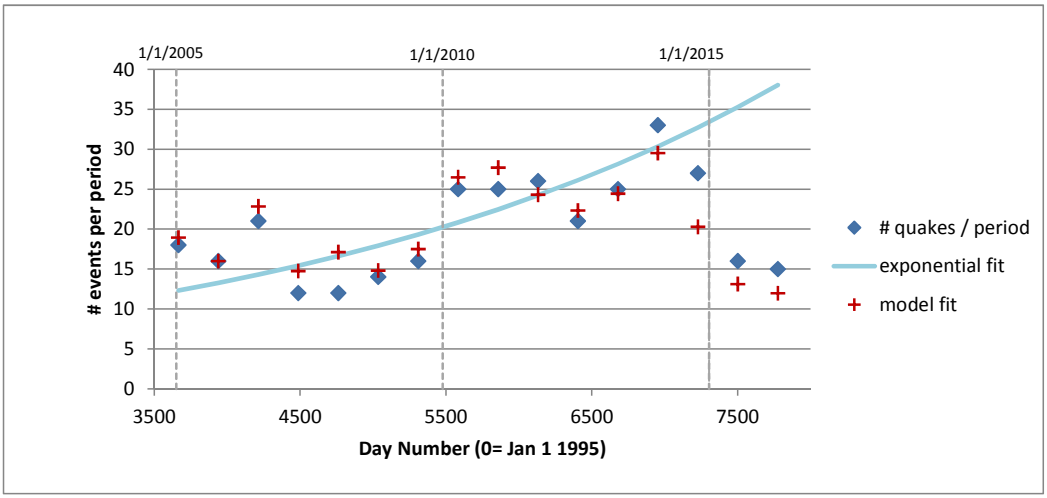


One can plot this ratio N_t/V_p as a function of Q which is done in fig. 5.2 for the same data as in the previous section. In essence the variable Q plays a similar role to time, where the 'clock' runs slightly faster or slower as gas production is increased or decreased. From the blue diamonds in fig. 5.2 it can be seen that there is rather a lot of variation, and a linear frame rate, i.e. $df/dQ = cst.$ does not appear likely, nor do simple extensions such as an f that is quadratic in Q.

In previous reports, it is suggested that a variable production rate can lead to strong gradients in gas pressure inside the reservoir; cf. Pijpers (2015b, 2016a), which would lead to an additional stimulating effect for tremors. In order to explore this, from the monthly gas production time series per cluster a time series of first order time differences is calculated for each production cluster: $\Delta V_{p\,m} = V_{p\,m} - V_{p\,m-1}$ where m is the index for the month. For the set of production clusters listed above, the maximum value of this ΔV_p is determined within 9 month windows. In fig. 5.2 the 'jumps' in production rate ΔV_p are shown as red squares, i.e. for each window the maximum in month-to-month increase in V_p in units of $[10^8\ m^3]$, shifted by one window-width. For cumulative production values higher than about $235\ 10^9\ m^3$, corresponding to a date of January 2003, there is a remarkably close correspondence between the series. That date of Jan. 2003 has been shown in van Thienen-Visser et al. (2015), to be a change point for the tremor rates. This would be consistent with a 'regime change' for the relationship between tremor rates and production.

Using the maximum of $\Delta V_{p\,m}$ within the window as an explanatory variable, in addition to the production rates V_p, a bilinear relationship is fitted. The time delay between a maximum ΔV_p and

Figure 5.3 The total number of earthquakes with $M > 1$ and excluding potential aftershocks, for 16 consecutive subsets of equal length, plotted against the prediction using the bilinear fit of eq. (5) (red crosses) as well as an exponential trend line.



an increase in tremors is not necessarily the same as the 3 months used for the production levels, so a range of delays is tested. From this the best fitting relationship is:

$$N_t = -13,0 + 0.6\frac{V_p}{10^9} + 7.0\frac{\Delta V_{p\ max}}{10^8} \tag{5}$$

In which N_t is the number of tremors in each 9 month window, V_p is the production volume in normalized m^3, and $\Delta V_{p\ max}$ is the maximal first order difference in normalized $m^3/month$. The R^2 of this fit is 0.77 if a time delay of 12 months is used between $\Delta V_{p\ max}$ and N_t. This value for R^2 is clearly a substantial improvement over 0.47 for a relationship using just V_p without an offset. Other time delays with values between 3 and 13 months have been attempted, but the best result is for a 12 months delay. In addition, a much better result would be quite surprising, since that would imply a near-deterministic generation of tremors. The remaining unexplained variation is consistent with what would be expected for a Poisson process at the count levels seen here. This implies that while the point process of the tremor occurrence is not a stationary Poisson process, the fluctuations remaining in this point process after removing a deterministic part may conform to a stationary Poisson process. In fig. 5.3 the number of tremors, with $M > 1$ and excluding potential aftershocks, is shown for each of the 16 successive intervals, with an exponential fit of the same type as is shown in fig. 2.3, and also with the numbers that the bilinear relationship of Eq. (5) produces. Note that while the data and fit are shown as a function of time, time is not used in the fitting procedure, and no time series analysis is applied.

In sum it appears unlikely that the variations in tremor rates can be explained purely in terms of a 'frame rate effect' where the gas production levels fully determine the expectation level of the tremor rate. A better phenomenological model is obtained with a bilinear relationship, using both the gas production rate and maxima of first order time differences of the monthly gas production as explanatory variables with time delays of 3 months and 12 months respectively. While it is clear that adding degrees of freedom in the modelling will never produce a worse model in terms of R^2, if the added degree of freedom has no explanatory power, there will not be an improvement in R^2. In this sense adding $\Delta V_{p\ max}$ produces a genuinely better model. With the parameters determined from the best fit eq. (5) one could surmise that in the complete absence of variations, i.e. $\Delta V_{p\ max} = 0$, levels of production below about $22\ 10^9 m^3$ per 9-month window, i.e. annual production levels of below $29\ 10^9 m^3$ should produce very few tremors. Even very modest, and perhaps operationally unavoidable, month-to-month variations

of 0.1 $10^9 m^3$ reduce this threshold to 13 $10^9 m^3$ per annum however. Perhaps the relationship (5) indicates that a non-linear effect such as a threshold in reservoir flow plays a role in the generation of tremors. However, in the absence of a properly physically motivated model for the effect of gas extraction on the dynamics of the gas reservoir and fault lines within it, phenomenological relationships such as (5) should only be used as working hypotheses, requiring monitoring.

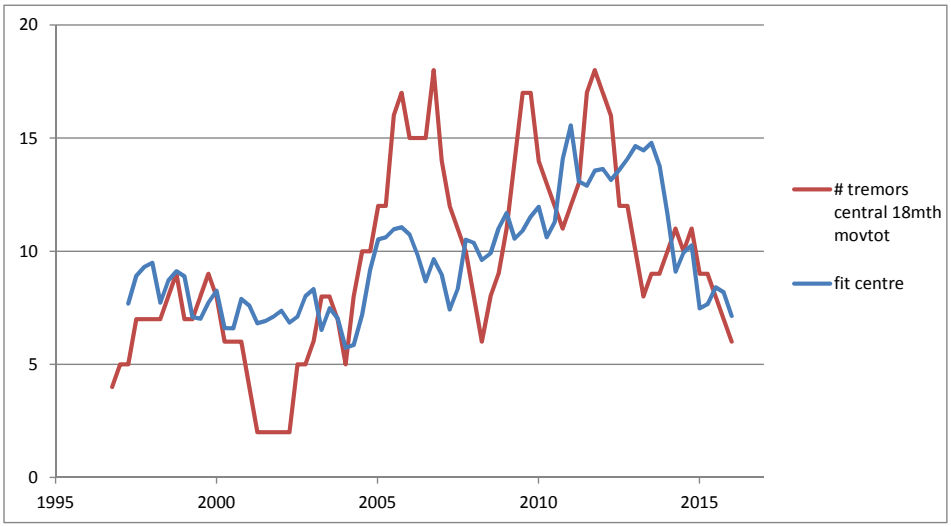
5.3 Reservoir gas pressure

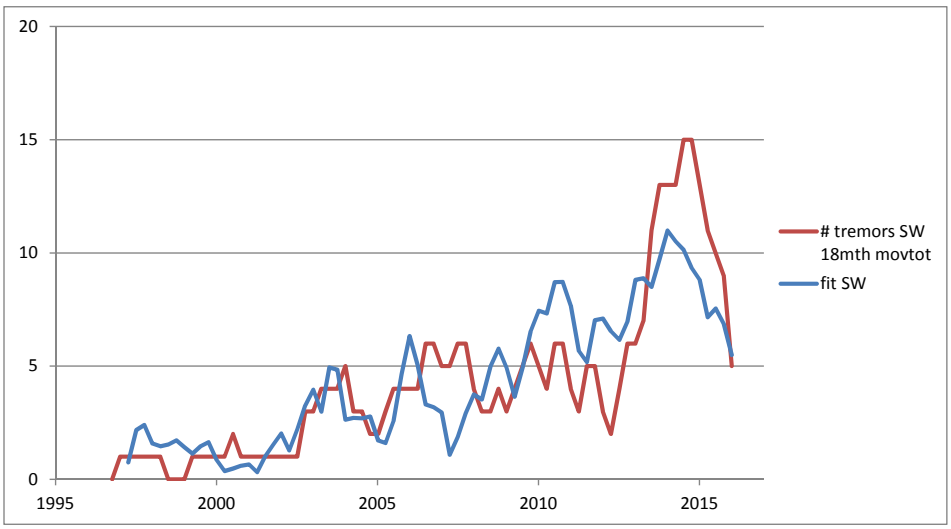
The effects of the gas production on the dynamical balance of the reservoir layers, and hence the generation mechanism for tremors, is mediated by the gas pressure inside the reservoir. While this is one step removed from the direct gas production which to some extent affects the ease with which direct intervention can be implemented, the reservoir gas pressure might provide a closer connection to the statistical properties of the tremors. The gas pressures can be obtained using a reservoir model which simulates gas flow in a porous media. If a history match is obtained for the measured production, pressures, water flow and subsidence, gas pressures are likely to be (on average) a good representation of the actual gas pressures in the reservoir. One of these modelling tools is called MoRes: for a more detailed description see e.g. Nederlandse Aardolie Maatschappij BV (2016). NAM has provided an input file of MoRes to TNO. For this application MoReS output pressures are obtained from TNO at the request of State Supervision of Mines. Volume weighted averages over the entire depth interval of specific regions of the MoReS output reservoir gas pressure as a function of time are used, over the three tremor zones of interest: central, SW, and SE (see fig. 1.1).

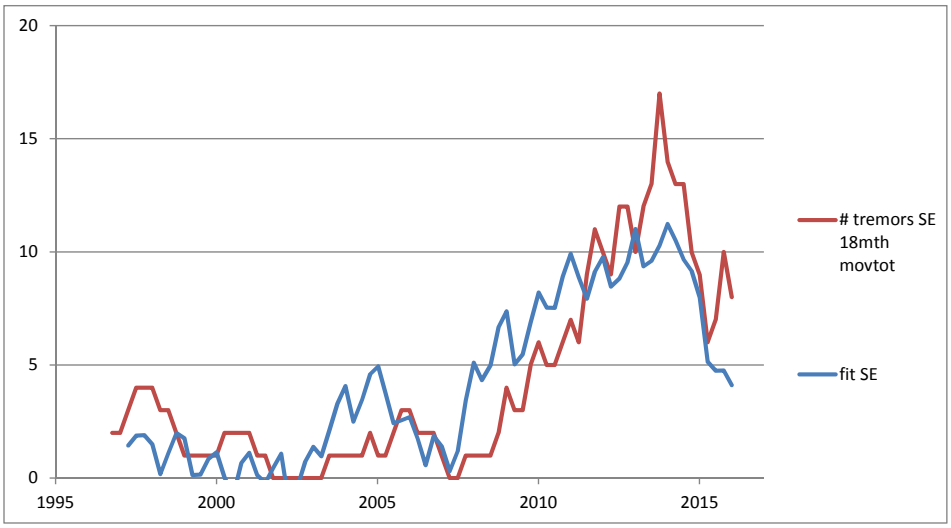
Since the aim is in part to assess regional variations, the intention is to compare the tremor rates in each of the three regions with the appropriate reservoir gas pressure. For the purposes at hand it is useful to have regularly sampled time series which is difficult to construct for a point process of tremor events without some form of processing. Moving averages in time are used, both for the gas pressures and for the tremor counts. The use of moving averages does have its disadvantages and pitfalls in terms of statistical measures of significance, and frequency content of time series. In the discussion of results, some care is necessary to avoid overinterpretation. For the width of the window used to construct the averaged time series, window lengths of 6 or 9 months proved to be so short that the event statistics became quite poor. A window length of 12 months would also be undesirable, because this would strongly suppress seasonal variations. Given these considerations a window length of 18 months is used. The time series are oversampled: a sampling time of once every 3 months is used. The moving averages or totals are taken symmetrically with respect to the sampling time.

Gas production in a semi-steady state, i.e. an amount of gas extracted per unit of time, induces a certain reservoir gas pressure reduction per unit time. The relationship between the two does have some dependence on the reservoir gas pressure itself. Without such a dependence the equivalent of the production quantity V_p, when using (functions of) the reservoir gas pressure is dP_{res}/dt, or in terms of finite differences $\Delta P_{res}/\Delta t$. Similarly, the role of the quantity $\Delta V_{p\ m}$ of the previous section could be played by a second derivative d^2P_{res}/dt^2 or $\Delta\Delta P_{res}/\Delta t^2$. In addition to using the (moving averages of) first and second finite differences of the reservoir gas pressure time series, it is also of interest to consider relative pressure first and second differences, i.e. the same two quantities but divided by the (moving average of) the reservoir pressure. The reasoning here is that if a pressure drop of 1 bar has a stronger

Figure 5.4 Moving total number of earthquakes with $M > 1$ and excluding potential aftershocks, 18 month window sampled every 3 months, plotted together with the fit prediction using first differences and maxima in second differences in reservoir gas pressure as explanatory variables.







geomechanical/dynamical impact on a layer where the average gas pressure is some tens of bar, than on a layer where that pressure is a few hundred bar, the relative pressure change would have more explanatory power than the pressure change itself in a regression model.

Contrary to the gas production quantity V_p, one would expect little time delay between tremor rates and the averaged reservoir gas pressure in the same region, since the diffusive 'travel time' for the production signal from the cluster to be communicated to the entire region is already accounted for in the reservoir simulation. In practice several time delays were tested but the best fits were indeed obtained when no delay was applied between reservoir gas pressure reduction rates and tremor rates. Downward jumps in gas pressure can be measured by e.g. maxima in the second differences $\Delta\Delta P_{res} \equiv P_{res\ m+1} - 2P_{res\ m} + P_{res\ m-1}$. Contrary to the gas pressure reduction rate over a region, downward jumps in gas pressure are likely to in fact arise at rather more localized subregions within a zone. The full magnitude of a rapid gas pressure drop is probably suppressed by the averaging over the zone. If these are to have an effect on spatially localised regions sensitive to generating tremors, such as fault lines, it can be argued that a finite delay time between tremors and second differences is plausible. Rather than using a window with a width of 18 months a width of 6 months is used instead for the second differences, in part to mitigate the suppression of temporal variation in this quantity and in part to improve the precision of the determination of any time delay.

Best fits for all three regions to the regional tremor rates were found using the relative first and second differences, ie:

$$\widehat{P'} \equiv \frac{1}{P_{res}} \frac{\Delta P_{res}}{\Delta t} \quad without\ any\ time\ delay \tag{6}$$

$$\widehat{P''}_9 \equiv max \frac{1}{P_{res}} \frac{\Delta\Delta P_{res}}{\Delta t^2} \quad with\ a\ time\ delay\ of\ 9\ months \tag{7}$$

where $\widehat{P'}$ and $\widehat{P''}_9$ are in units of $[yr^{-1}]$ and $[yr^{-2}]$ respectively. Omitting one or the other variable leads to distinctly poorer quality fits. The relationships between the number of quakes per 18 month window and the fitting variables, and their resp. R^2 measures are:

$$\begin{aligned} N_t(central) &= 2.3 + 6.2\,\widehat{P'} + 3.1\,\widehat{P''}_9 \quad R^2 = 0.36 \\ N_t(SW) &= -1.8 + 4.6\,\widehat{P'} + 1.1\,\widehat{P''}_9 \quad R^2 = 0.65 \\ N_t(SE) &= -4.3 + 5.7\,\widehat{P'} + 1.3\,\widehat{P''}_9 \quad R^2 = 0.70 \end{aligned} \tag{8}$$

In assessing the R^2 for these relationships, care should be taken here. Because of the oversampling of the time series, only every 6^{th} point is statistically independent. There are 6 ways to choose every 6^{th} point of the series to calculate R^2 measures with from a set of of genuinely statistically independent points. Therefore in fact for each fit there is a small range of applicable R^2 values, where the value quoted above is an appropriate mean. The ranges of R^2 for each region are: central $[0.28, 0.47]$, SW $[0.55, 0.79]$, and SE $[0.64, 0.78]$.

The moving averages time series of tremor rates and the fits are shown in fig. 5.4 as a function of time for the three regions. Note that time is not used as a variable in the fitting process and no time series analysis techniques are applied. Inspecting the relationships (8) for the three regions between tremor counts and the fitting variables, regions SW and SE have similar fitted proportionality constants, but different offsets. The zone central has a lower R^2 and different fitted constants. The zone central is also the largest of the three, and perhaps is more heterogeneous in terms of the properties of the reservoir layer within its boundary than the other two, or has two quite distinct regions within its contour. The remaining variance, of event counts minus fit, for this zone is 3.3 which is larger by roughly a factor of $\sqrt{2}$ than for the other two zones which is also consistent with such an interpretation. For all zones the relationship (8)

has a slightly lower R^2 than relationship (5). In part this is a reflection of the lower counts per region and hence higher relative contribution of noise than if the entire region (i.e. within the purple ellipse of fig. 1.1) is fitted as one, as in (5). Another issue is that for relationship (5) the $max\Delta V_p$ per cluster is used, whereas in this section $\widehat{P''}_9$ is determined only **after** averaging over the entire zone in question. This suppresses variations in this variable, and hence is detrimental to the quality of the fits.

6 Conclusions

From the analysis presented in this report, it can be concluded that, averaged over the period from Sep. 1 2004 to Sep. 1 2016, there is a statistically significant spatial enhancement of the earthquake rate in all three zones, SW, central and SE, where gas production takes place, compared to the surrounding region, by factors of around 2.2, 2.8 and 1.3 respectively. If only the best-localised tremors with magnitudes above 1.5 are taken into account the enhancements for SW and central are 1.8 and 3.8 respectively. The region SE in this case does not show enhancement.

For all regions SW, SE, and central, as well as the area directly surrounding these regions of particular interest, there is an increasing trend in the earthquake rate with time since Sep. 1 2004, which can be fit with an exponential increase with a doubling time of ~ 5.4 years or even a shorter time of 4.3 years if two somewhat anomalous points are excluded from the fitting procedure. In the most recent 18 months, since March 1 2015, the data are consistent with a reversal in this trend, reducing the tremor rates to a rate consistent with the twelve-year average from Sep.1 2004 to Sep. 1 2016. An exception is the central zone, where production was reduced earlier, from the beginning of January 2014. For this region the tremor rate is consistent with the hypothesis that the rate has reduced to about half of the twelve-year average from Sep.1 2004 to Sep. 1 2016.

There is sufficient evidence, from all previous work, to assume as a working hypothesis that there is a causal relationship between gas production and tremors, with a relatively short reaction time cf. Pijpers (2015b, 2016a). At present the data discussed in this report are consistent with the hypothesis that the reduced and flat production has had the effect of reducing the earthquake rate to well below the rates found just previous to this epoch. It appears unlikely that the tremor rates are determined purely by a 'frame rate effect' i.e. due to gas production rates alone. In addition, in the central zone the cdf might be steepening which would imply a stronger reduction in the tremor rate for tremors with $M > 2$ than for the smaller tremors. The absence of tremors with $M > 2$ in the central zone, after March 1 2015, might also still be due to the general reduction in tremor rate. It may take another year of data collection in order to be able to make a statistically significant distinction between the two possibilities.

The region SE, while conforming to these conclusions, does give cause for a recommendation for close monitoring, since there are some tentative indications that the slope of the cdf is reducing, which might imply that tremors with magnitudes above about 2 have become more prevalent compared to tremors with lower magnitudes than was the case previously in this region.

The relationships presented in section 5 may suggest that there is some form of control possible over the generation of tremors, through appropriate production adjustments. However, before

such relationships are used in this way, this requires at a minimum the validation of the relationships using data not used in determining the fits, i.e. monitoring in the year(s) to come. Until such validation has become possible, the relationships of section 5 must be regarded with due caution, and treated as working hypotheses only.

References

Baiesi, M. and M. Paczuski (2004). Scale-free networks of earthquakes and aftershocks. *Phys. Rev. E 69(6)*, 066106--066114.

Dake, L. (1978). *fundamentals of reservoir engineering*. Elsevier. (17^{th} impr. 1998, chapter 3).

Doornhof, D., T. Kristiansen, N. Nagel, P. Pattillo, and C. Sayers (2006). Compaction and Subsidence. Technical report, Schlumberger.

Dost, B., F. Goutbeek, T. van Eck, and D. Kraaijpoel (2012). Monitoring induced seismicity in the North of the Netherlands: status report 2010; wr 2012-03. Technical report, KNMI.

Gardner, J. and L. Knopoff (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bull. Seis. Soc. Am. 64(5)*, 1363--1367.

Garwood, F. (1936). Fiducial limits for the Poisson distribution. *Biometrika 28*, 437--442.

Huc, M. and I. Main (2003). Anomalous stress diffusion in earthquake triggering: correlation length, time dependence, and directionality. *J. Geophys. Res. 108 (B7)*, 2324.

Kumazawa, T. and Y. Ogata (2014). Nonstationary ETAS models for nonstandard earthquakes. *Ann. Appl. Stat. 8*, 1825--1852.

Naylor, M., I. Main, and S. Touati (2009). Quantifying uncertainty on mean earthquake inter-event times for a finite sample. *J. Geophys. Res. 114 (B0)*, 1316.

Nederlandse Aardolie Maatschappij BV (2013). A technical addendum to the winningsplan Groningen 2013 subsidence, induced earthquakes and seismic hazard analysis in the Groningen field. Technical report.

Nederlandse Aardolie Maatschappij BV (2016). A technical addendum to the winningsplan Groningen 2016 production, subsidence, induced earthquakes and seismic hazard and risk assessment in the Groningen field. Technical report.

Nepveu, M., K. van Thienen-Visser, and D. Sijacic (2016). Statistics of seismic events at the Groningen field. *Bull. Earthquake Eng*, 1--20.

Pijpers, F. (2014). Phase 0 report 2 : significance of trend changes in tremor rates in Groningen. Technical report, Statistics Netherlands.

Pijpers, F. (2015a). Phase 1 update may 2015 : significance of trend changes in tremor rates in Groningen. Technical report, Statistics Netherlands.

Pijpers, F. (2015b). A phenomenological relationship between gas production variations and tremor rates in Groningen. Technical report, Statistics Netherlands.

Pijpers, F. (2015c). Trend changes in tremor rates in Groningen : update november 2015. Technical report, Statistics Netherlands.

Pijpers, F. (2016a). A phenomenological relationship between reservoir pressure and tremor rates in Groningen. Technical report, Statistics Netherlands.

Pijpers, F. (2016b). Trend changes in tremor rates in Groningen : update may 2016. Technical report, Statistics Netherlands.

Robert, C. and G. Casella (2004). *Monte Carlo Statistical Methods*. Springer.

Tarantola, A. (2004). *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM.

van Thienen-Visser, K., P. Fokker, M. Nepveu, D. Sijacic, J. Hettelaar, and B. van Kempen (2015). Recent developments on the seismicity of the Groningen field in 2015. Technical report, TNO.

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress: Statistics Netherlands, Grafimedia
Design: Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contact form: www.cbs.nl/information

Where to order
verkoop@cbs.nl
Fax +31 45 570 62 68
ISSN 1572-0314