Impact of Production Shut-in on Inter-Event time in Groningen.

A statistical perspective.

Restricted
SR.15.13107
October 2015

Restricted **SR.15.13107**

Impact of Production Shut-in on Inter-Event time in Groningen.

A statistical perspective.

by

Rakesh Paleja (PTD/ TASE); Stijn Bierman(PTD/TASE); Matthew Jones (PTD/TASE)

Shell Global Solutions (UK) (a division of Shell Research Ltd.)

Further electronic copies can be obtained from the Global Information Centre.

Executive Summary

Gas production from the Groningen Field located in the north of the Netherlands induces earthquakes that are causing a concern with the local population. To address the issue, Nederlandse Aardolie Maatschappij BV (NAM), the ministry of Economic Affairs, and regulator SodM have designed a protocol to measure and control the rate at which induced earthquakes occur. It is thought that if production is reduced, the number of induced earthquakes, called activity rate, will reduce. With this assumption, production in Loppersum area, which is part of the larger Groningen field was stopped on 1st Jan 2014.

This report is motivated to quantitatively understand the influence of production shut-in in Loppersum on activity rate. The report is continuation of the research note that was published in May 2015 which used the event catalogue up to 1st Jan 2015. Since more events have occurred since 1st Jan 2015, the inferences drawn in this report are more robust and quantitative in nature. The work has also been extended to other production areas besides Loppersum and covers Oost, Zuidwest and Eemskanal.

In order to draw conclusions, we have downloaded a catalogue of events on 9th September 2015. The catalogue is divided into two parts. The part that contains events before the production was shut, i.e up to 31st December 2013 is call pre shut-in period. The part that contains events from 1st Jan 2014 up to 9th Sept 2015 is called post shut-in period. Using this catalogue we first estimate the parameter inter-event time, λ, the expected time difference between two successive events in the post shut-in period (called λ_{post}) and pre shut-in period (called λ_{pre}). The difference between λ_{post} and λ_{pre} , called $\Delta\lambda$ and its 95% Confidence Interval (CI) is derived using statistical techniques and forms our test statistic. If the 95% CI for $\Delta\lambda$ contains only positive values, we conclude that the test statistic is significant. In such a case we infer that activity rate has reduced and there are fewer events post shut-in than pre shut-in. On the other hand if the 95% CI for $\Delta\lambda$ contains all negative values, the test statistic is significant and we conclude that activity rate has increased and there are more events post shut-in than pre shut-in. For all other cases, the test statistic is not significant and we conclude that activity rate post shut-in is no different to pre shut-in.

The method allows partitioning the event catalogue in the pre shut-in period into various equal sized, non overlapping blocks in time. This permits us to temporally compare activity rate post shut-in to various pre shut-in periods for each of the four regions without introducing a bias caused by different sample sizes (i.e. the number of events) in post shut-in and pre shut-in periods. By taking a subset of events of only certain magnitudes such as M⩾0.5, M⩾1.0 and M⩾1.5, we can also make quantitative inference on the influence of production shut-in on a subset of events.

Using the methodology we derive the following statistically significant changes in activity rate between post shut-in and the pre shut-in period that is closest in time to the post shut-in period but with the same number of events as in the post shut-in period:

- The activity rate in Loppersum area post shut-in is lower than the activity rate in Loppersum in the pre shut-in period from Jan 2013 up to Dec 2013. This is true for events of all magnitudes.
- In Loppersum, post shut-in, we can expect inter-event time, the time between two consecutive events, to have increased by 20 days for M⩾1.5. The expected increase is 7 days for M⩾1.0 and 5 days M⩾0.5 compared to the period from Jan 2013 up to Dec 2013.
- The activity rate in Zuidwest post shut-in has increased for events of M⩾1.0 and M⩾0.5 compared to any pre shut-in period. The increase in activity rate is despite the fact that production in Zuidwest post shut-in is at a historic low, along with production in the other

three production areas. There is no noticeable change in activity rate yet when M⩾1.5 are considered.

- In Zuidwest, post shut-in, we can expect inter-event time between two consecutive events to have decreased by 35 days for M⩾1.0 compared to pre shut-in spanning from Dec 2009 to Dec 2013. The expected decrease is 50 days for M⩾0.5 when post shut-in period is compared to pre shut-in period from Dec 2010 to Dec 2013.

We do not see a statistically significant difference in activity rate in Oost and Eemskanaal between the post shut-in period and pre shut-in period that is most adjacent to post shut-in period with the same number of events.

We conclude this summary by informing the reader that there are temporal trends in activity rate for each region and each magnitude. The inferences on the difference in activity rate post shut-in and the most adjacent pre shut-in drawn above should not be generalized and extended beyond the time frame described. Thus, while the activity rate in Loppersum post shut-in is lower than in the period from Jan 2013 to Dec 2013 for all magnitudes, activity rate in Loppersum post shut-in is similar to activity rates:

- from 1998 to 2012 for M⩾1.5,
- from 2003 to 2012 for M⩾1.0 and
- from 2009 to 2012 for M⩾0.5

The temporal trends are described in the main body of the report. Further, we note that the report also discusses an alternative method and test statistic which allows us to draw a similar conclusion suggesting robustness of our approach.

Manchester, October 2015.

Table of Contents

List of Figures

List of Tables

1. Introduction

The people living in Groningen have been confronted with an increasing number of induced earthquakes which has caused anxiety and distress to the population. To address the concern, Nederlandse Aardolie Maatschappij BV (NAM), the ministry of Economic Affairs, and regulator SodM have designed a protocol to measure and control the rate at which induced earthquakes, called activity rate, occur in Groningen [1]. It is thought that activity rate is positively correlated to production from the field. Thus, by reducing the production, we can control the number of earthquake events that are observed. The Loppersum region, which is a part of the larger Groningen field, has historically shown the highest activity rate. In order to test the influence of production on activity rate, it was decided to shut the production in Loppersum on 1st Jan 2014 and measure activity rate post this shut-in period.

In the document 'DRAFT: A2.1 Activity rate Loppersum'[2] (old report), we shared preliminary results on the impact of production shut-in in the Loppersum area on activity rate. We showed that for events with $M \geqslant 1.5$, activity rate has not been impacted by shut-in as the inter-event time, the expected time between two consecutive events, before and after shut-in were largely similar. On the other hand, when all events (i.e. of all magnitudes are considered), the inter-event time seemed to have increased after production shut-in. We must note that in the old report, the event catalogue up to 1st Jan 2015 was used.

Since issuing the old report, the event catalogue has been enriched with more events in the period from 1st January 2015 to 9th Sept 2015. Analysis of this new data allows much more robust and meaningful inferences to be drawn as the sample size has increased. The objective of this report is to share our views on the impact of shut-in on activity rate using this supplemented data.

The old report was reviewed by SodM [3], NAM and other stakeholders. Many valuable comments were received and they have been incorporated in this report. Notably, we have addressed the following comments:

- In the old report, our inferences on the impact of production shut-in on activity rate were largely qualitative and subjective. In this report, quantitative inferences are drawn.
- The cumbersome process of Markov Chain Monte Carlo (MCMC) sampling has been solved by an analytical solution.
- No distinction is made between 'aftershock' and 'true' events and all the events that are in the catalogue are taken into account.
- The current work extends the geographical boundaries to other production areas besides Loppersum.

The current report also shares the statistical framework, assumptions and limitations of the methodologies that have been proposed and used in this work. Possible improvements and future work are suggested towards the end.

2. Statistical framework, methodology and assumptions

In our current work, we have divided the Groningen field into 4 sub regions/production areas referred to as Loppersum, Oost, Zuidwest and Eemskanaal (see Figure 5.1 for example). The boundaries defining the regions are based on expert judgment provided by NAM and the objective of such sub-division is to assign all the events that have occurred in the catalogue of events to one of the four production areas. This way no event is ignored in our assessment.
A catalogue of events was downloaded on 9th Sept 2015. As described in the old report, we divided the catalogue of events into pre shut-in and post shut-in periods. The pre shut-in period contains events from 1st January 1991 to 31st December 2013 while the post shut-in period contains events from 1st January 2014 to the date the event catalogue was generated i.e. 9th September 2015. Thus the post shut-in period contains events from 1st Jan 2014 to 9th September 2015. Events in each period are assigned to the four production areas described above. The assignment is done based on the events' coordinates and the boundary defining each region. Once the assignment is done, we infer whether activity rate post shut-in is different to pre shut-in using the two methodologies described below.

2.1. Method 1: Posterior inter-event time using Bayesian inference

This method has been described in the old report. Subsequent to the publication of this, we have adapted the methodology to improve the computational efficiency and accuracy of the statistical parameters that we estimate. The description is as follows:
Let there be n events in the post shut-in period up to 9th Sept 2015 in any production area. Let t_1, t_2, ... t_n be the inter-event time from a sampling density $f(t_1, t_2, ..t_n|\lambda)$ where λ is the expected time between two consecutive events. If we assign the parameter λ a prior density $g(\lambda)$ the posterior conditional density of λ is given by the following equation:

$$p(\lambda|t_1, t_2...t_n) = \frac{f(t_1, t_2, ..t_n|\lambda)g(\lambda)}{\int_0^\infty f(t_1, t_2, ..t_n|\lambda)g(\lambda)d\lambda} \tag{2.1}$$

The above equation can be simplified to proportionality so that:

$$p(\lambda|t_1, t_2...t_n) \propto f(t_1, t_2, ..t_n|\lambda)g(\lambda) \tag{2.2}$$

If the activity rate follows a Poisson process [4], t_1, t_2 .. t_n are from Exponential distribution, the probability distribution function for which is well known [5]

$$f(t_1, t_2, ..t_n|\lambda) = \prod_{i=1}^{n} f(t_i|\lambda) \propto \frac{1}{\lambda^n} e^{\frac{-\sum t_i}{\lambda}} \tag{2.3}$$

This can be written as:

$$f(t_1, t_2, ..t_n|\lambda) \propto \lambda^{-n} e^{\frac{-S}{\lambda}} \tag{2.4}$$

where S is the summation of the n inter-event times $t_1...t_n$.
Substituting 2.4 in 2.2 gives the posterior density of λ as:

$$p(\lambda|t_1, t_2...t_n) \propto \lambda^{-n} e^{\frac{-S}{\lambda}} g(\lambda) \tag{2.5}$$

The prior density function for λ is unknown and one can assume a non informative prior up to proportionality as:

$$g(\lambda) \propto \frac{1}{\lambda} \tag{2.6}$$

which when substituted in 2.5 leads to the posterior density

$$p(\lambda|t_1, t_2...t_n) \propto \frac{\lambda^{-n} e^{\frac{-S}{\lambda}}}{\lambda} \tag{2.7}$$

or

$$p(\lambda|t_1, t_2...t_n) \propto \lambda^{-n-1} e^{\frac{-S}{\lambda}} \tag{2.8}$$

This is very similar to inverse gamma distribution [6] and the posterior density of λ can be estimated as follows:

$$p(\lambda|t_1, t_2...t_n) = \frac{S^n}{\Gamma(n)} \lambda^{-n-1} e^{\frac{-S}{\lambda}} \tag{2.9}$$

In the above equation, S defines the scale, n defines the shape and $\Gamma(.)$ denotes gamma function. Posterior λ is bound between $(0, \infty)$ but realistically, we expect λ between [0 ,1000] to be more realistic and it is prudent to evaluate it on a very fine grid in this range. This avoids sampling from MCMC sampler and improves the efficiency and accuracy of λ estimate. An example of the density function of λ is shared in figure A.2.
The above discussion for estimating the posterior λ in post shut-in prior can also be used to estimate λ in pre shut-in period. In this case, if N is the total number of events in the entire pre shut-in period in the same production area that defines n, N will be >>> than n. This is simply because the pre shut-in period spans from 1st Jan 1991 to 31st Dec 2013 which allows a large number of events to accumulate over time. So, to estimate λ in the pre shut-in period with same number of events as in post shut-in period, we simply divide the N events into k contiguous blocks, each of size n and use equation 2.9 for each block. There will be k λ's for pre shut-in and one in post shut-in period. We simply see the temporal trend in λ to qualitatively suggest how activity rate has changed over the entire history of each region. This was described in our old report. We went to the extent of suggesting that if the densities of λ post and pre shut-in period are not overlapping, the two λs are different. But this remained qualitative and we have taken steps to make more quantitative assessment since the old report.
To quantitatively compare if activity rate has changed in the post shut-in period compared to pre shut-in period, we draw samples of λs in the pre shut-in period and the block for post shut-in period. Notice that λs is from inverse gamma distribution and since both the shape and scale parameters are known it is trivial to draw a very large number of samples representing the population. Next we take the difference between the λ post shut-in and λ pre shut-in and determine the density $p(\lambda_{post}$- $\lambda_{pre})$ from the drawn samples. After trimming the tails, we retain ($\Delta\lambda = \lambda_{post}$- λ_{pre}) in the 2.5% to 97.5% percentile band. This is our test statistic. If all the retained values are positive we conclude that activity has decreased in the post shut-in period compared to pre shut-in period and that this is statistically significant. On the other hand if all the values are negative, we conclude that seismic activity has increased and that this is statistically significant. If the values contain both positive and negative values we cannot quantitatively comment on the change in activity rate.

2.2. Method 2: Difference in proportion

If there are n events in the post shut-in period with inter-event time t_1, t_2, ... t_n , the sample proportion of days an event has been observed over S days is given by

$$\hat{\alpha} = \frac{n}{S} \tag{2.10}$$

Following the discussions in the previous section, in the pre shut-in period, all the N events are divided into k blocks each of size n and compute $\hat{\alpha}$ for each block. If $\widehat{\alpha_{post}}$ is the sample proportion in the post shut-in period and $\widehat{\alpha_{pre}}$ is the sample proportion in a block of pre shut-in period, the null hypothesis is that the population proportions, α_{post} and α_{pre}, are equal.

$$H_0 : \alpha_{post} = \alpha_{pre} \tag{2.11}$$

The alternative hypothesis is that the population proportions are not equal.

$$H_1 : \alpha_{post} \neq \alpha_{pre} \tag{2.12}$$

Before we can actually conduct the hypothesis test, we'll have to derive an appropriate test statistic. The test statistic for testing the difference in two population proportions, that is, for testing the null hypothesis H_0: α_{post}- α_{pre}=0 is [7]

$$Z = \frac{(\widehat{\alpha_{post}} - \widehat{\alpha_{pre}}) - 0}{\sqrt{\hat{A}(1 - \hat{A})(\frac{1}{S_{pre}} + \frac{1}{S_{post}})}} \tag{2.13}$$

where

$$\hat{A} = \frac{2n}{S_{pre} + S_{post}} \tag{2.14}$$

The Z statistic is normally distributed with mean 0 and variance 1. This means that if Z is $<$-1.96, we can infer that there is less activity in post shut-in period than in pre shut-in period. If Z is $>$ 1.96, we can infer that activity rate has increased. On the other hand -1.96⩽Z⩽1.96 implies we cannot draw any quantitative conclusions.

2.3. Assumptions, Advantages and Disadvantages of the two method

Both methods assume that inter-event times $t_1., t_2, ...t_n$ are from the exponential distribution. This is a reasonable assumption and is empirically supported by the catalogue data. Figure 2.1 shows the inter-event time for all M⩾1.5 in Loppersum in the pre shut-in period. The top panel of the chart is the time series plot of inter-event time. Up to event number 45, labeled as Group 1, there is a marginal temporal decline in inter-event time. After event number 45, labeled as Group 2, there is little temporal decline. The bottom panel is the frequency plot of inter-event time of stationary, Group 2, events (sample size 105). We observe that the histogram resembles exponential distribution, thus justifying our choice of exponential likelihood function. We note that our inference about the sample distribution is drawn using a reasonably large sample size during a period when activity rate was near-stationary. It is not advisable to derive such inference if the sample size is small and/or when the process in not stationary.
The first method assumes a non-informative prior for λ. This is because we do not have any subjective evidence so far that can support an alternative prior. It should be apparent that with a non-informative prior, we let the data speak for itself. If, on the other hand, we have some prior information about λ, the first method becomes really powerful and posterior is more reliable given the

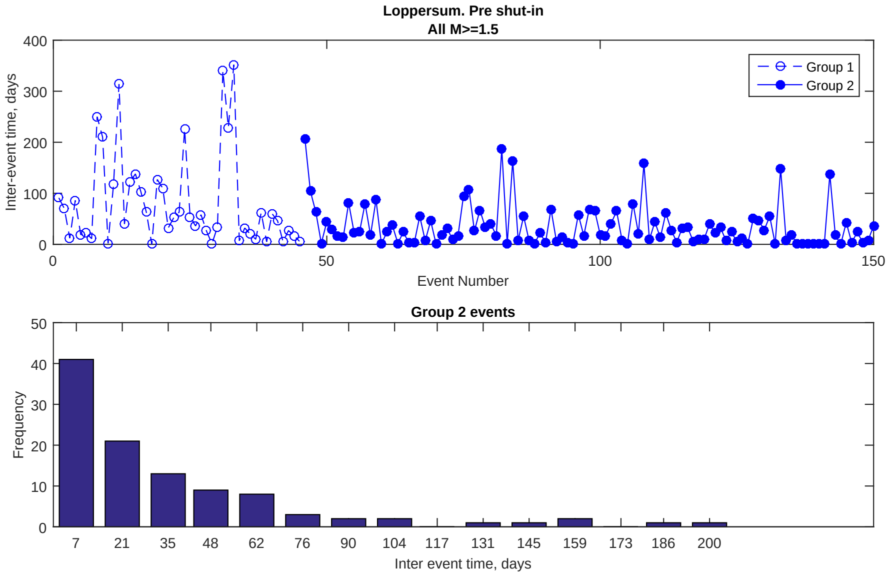

Figure 2.1.: Top: Inter-event times in Loppersum, M ⩾1.5 in the pre shut-in period from 1991 to end 2013. Group 1 (open circles) refers to period when inter-event time shows some temporal trend. Group 2(closed circles) refers to period when inter-event time does not show strong temporal trend. Bottom: Histogram of Group 2 inter-event times.

prior information. We should bear in mind that the shape of the likelihood function is strongly influenced by n particularly when n is small. This can introduce bias in the posterior λ in such cases: In a hypothetical case where only two events were observed, the posterior λ can change dramatically if for same S, a couple of more events were observed. An additional disadvantage of the first method is that it is novel and may be difficult to understand which may discourage its use.
The second method offers the advantage in that it is widely used in clinical trials. It is easy to understand and well accepted. The second method assumes that the Z statistic is from standard normal distribution. This is potentially a limitation of the second method as it is not always likely to be the case with real data. S will be a Poisson variate which means $\hat{\alpha}$ is also a Poisson variate. When n is large, the Poisson distribution approaches normality. The Z statistic which is based on the difference in $\hat{\alpha}$, will therefore be normally distributed. However when n is small, the second method encounters the same problem as the difference between two Poisson variates is not necessarily normal. Thus, like the first method, our conclusions may be biased. This is a general problem for which there is no simple solution and we recommend accumulating the events as was suggested in the old report and during meetings with various stakeholders. The current catalogue is definitely richer than the one we used for old report and the conclusions we draw now are therefore more reliable with both the methods.
The biggest disadvantage with the second method is that we can accept or reject the null hypothesis but we cannot directly estimate how the inter-event time has changed without further computation. This is not the case with the first method as the test statistic $\Delta\lambda$, when significant, itself is informative about the change in the inter-event time.
We must note that the two methods are directionally opposite. Thus when activity rate increases, λ and $\Delta\lambda$ decrease but α and Z statistics increase.
We conclude this section highlighting the point that both the methods have their own advantages and disadvantages. In our current report, we base our conclusion based on evidence supported by both the methods.

3. Production Time Series

Figure 3.1 shows the overall production from Groningen. Both raw and smoothed production are shown in the plots. It is evident from the figures there are seasonal trends in production as more gas is produced during winter than in summer. The long term trend shows that the overall production was stationary during 2004 to 2007 period, while it showed an upward trend post 2007 up to 2014. The overall production post shut-in is significantly lower than any previous years. We note that the production data is available from 2003 onwards. For 2015, data is available only up to Sept 2015.

Figure 3.2 shows the production by the four production areas. Production area Oost has the highest production followed by Loppersum, Zuidwest and Eemskanaal. All the four production areas show a similar long term trend as observed with the overall production; stationary production during 2004 to 2007, an upward trend post 2007 which declines significantly post shut-in. The production in all the four areas is at a historic low in the post shut-in period.

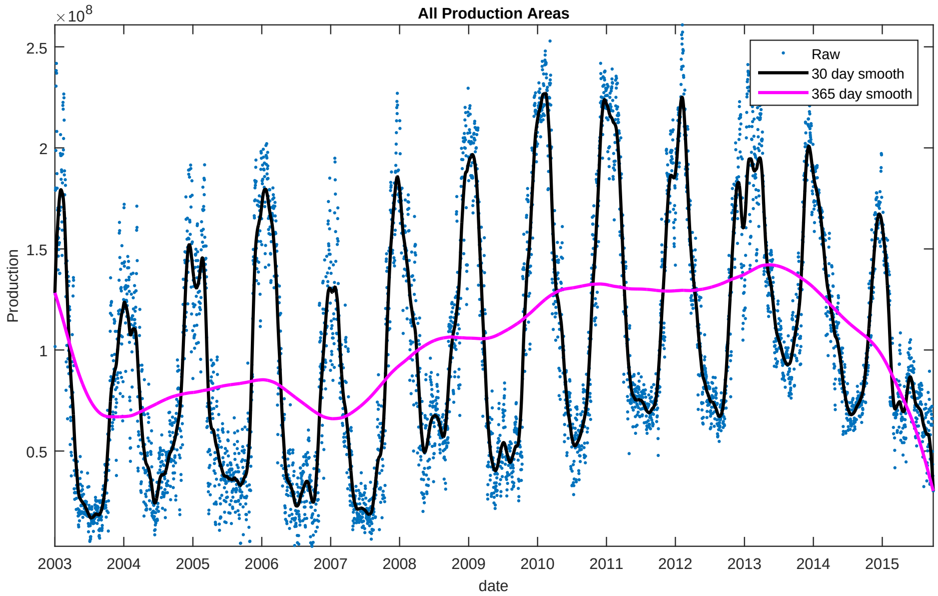

Figure 3.1.: Raw daily production (blue), 30 day smooth (black) and 365 day smooth (magenta) production from all four production areas.

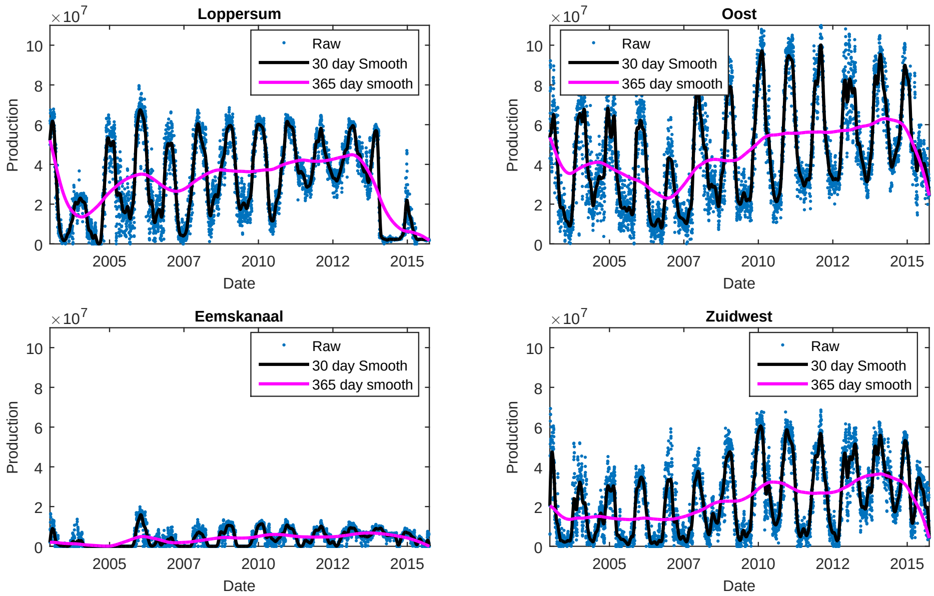

Figure 3.2.: Raw daily production (blue), 30 day smooth (black) and 365 day smooth (magenta) production in the four production areas.

4. Activity rate Time Series

As in the previous chapter, which showed temporal variations in production, we can study the temporal variation in activity rate. In order to do so, the catalogue data is split by regions and by the magnitude of events and yearly event count is derived. Figure 4.1 shows time series plots of yearly event count with M⩾1.5. For the entire Groningen field, activity rate was stationary up to 2003. It increased from 2003 and shows a peak at 2013. Following this time, there seems to be a decline. At the regional level, the trend is similar. When events with M⩾1.0 or M⩾0.5 are considered, we see a similar trend (figure 4.2 and figure 4.3). We note that seismic activity in Zuidwest shows a monotonic increase from 2010 for M⩾0.5 . We also note that besides this monotonic increase, which is an exception, there are large yearly variations for all other cases. Thus, for instance, with M⩾ 1.0, Eemskanaal had 4 events in 2008 but there was no observed event in 2010.

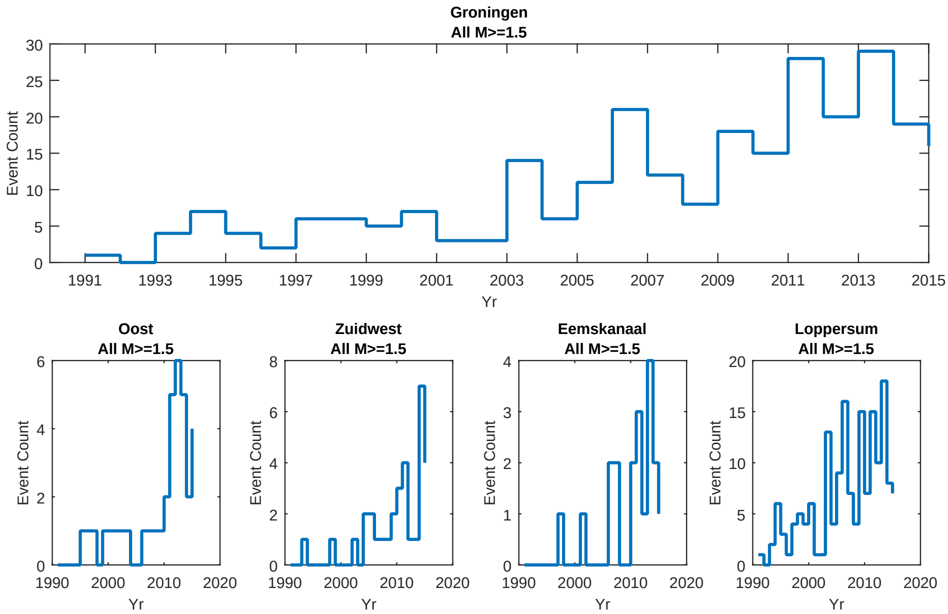

Figure 4.1.: Event count time series for Groningen (top) and four production areas (bottom). M⩾1.5

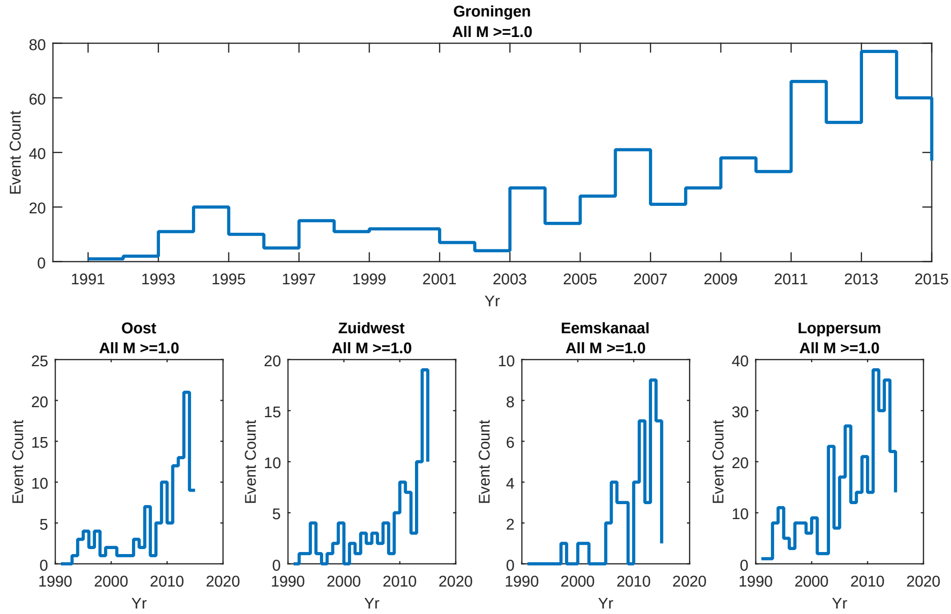

Figure 4.2.: Event count time series for Groningen (top) and four production areas (bottom). M⩾1.0

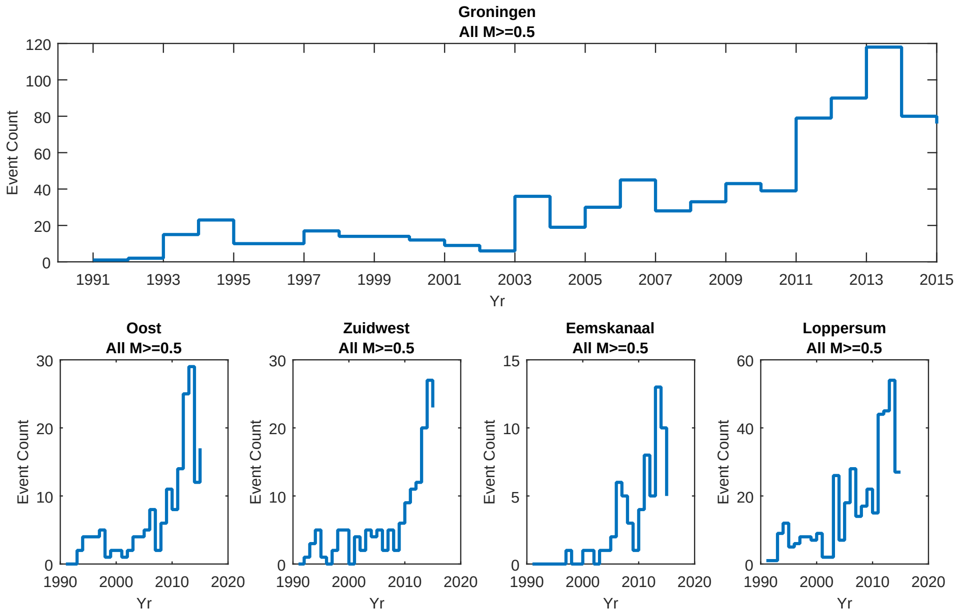

Figure 4.3.: Event count time series for Groningen (top) and four production areas (bottom). M⩾0.5

5. Statistical inference on change in activity rate M⩾1.5

We now use the two methods described in section 2 to infer if activity rate post shut-in is different to activity rate in the pre shut-in period for all the four production areas. The spatial distribution of all the events of M⩾1.5 for the two time periods is shown in figure 5.1. Table 5.1 shows the number of events that have occurred in each production area over the two periods. The number of events in each production area post shut-in is used to partition the events in the pre shut-in period into equal sized, non-overlapping blocks. Thus, for example, in Loppersum, there are 15 events post shut-in and 152 events in the pre shut-in period. The last 150 of the 152 events are partitioned into 10 non overlapping blocks each of 15 events and the very first two events left out of the analysis. This gives us 11 blocks in all; Blocks 1-10 belonging to the pre shut-in period and block 11 belonging to the post shut-in period, each containing 15 events.
The start date, end date and the time difference between the two dates for each block is shown in table 5.2. For other production areas, please refer to Appendix in section A. The date difference between two consecutive events in a block is called inter-event time. For Loppersum, this is shown in figure 5.2. We note that in each block the expected inter-event time and the variance are similar. This would suggest that the underlying activity rate is Poisson with little or no over dispersion. Similar partitioning of events into blocks is done for other regions.
The inter-event time in each block is used to estimate the posterior density of the expected inter-event time, λ, and its 95% CI using method 1. This is shown in figure 5.3. The last block in each sub plot refers to the post shut-in period while all other blocks refer to the pre shut-in period. We can make the following **qualitative** inferences:

- λ generally declines and we progress from left to right in each sub-plot. Since the first block dates back in time and the last block is towards 2014/15, we infer that activity rate increases with time.
- The decline is not monotonic: For instance for Loppersum, block 5 (time 12/02/2006 to 26/01/2007) has a lower λ than block 6(time 14/05/2007 to 04/20/2009).
- For Loppersum and Eemskanaal, λ post shut-in has shown an upward tick indicating that activity rate has reduced in these production areas. Oost does not show such an upward trend. For Zuidwest, λ post shut-in has declined indicating activity rate has increased in this region.

Since there is uncertainty in λ due to variance in inter-event times (see Figure 5.2), it is prudent to take into account the difference in λ between the post shut-in block and all the pre shut-in blocks to infer if the λ post shut-in has changed with respect to other blocks pre shut-in. The results are shown in figure 5.4. We can derive the following **quantitative** conclusions:

- For Loppersum, post shut-in activity rate is lower than the period defined by block 10, pre shut-in, which spans from 07/02/2013 to 15/11/2013. This is statistically significant as the test statistic, $\Delta\lambda$, has all positive values. The interval defined by trimmed $\Delta\lambda$ is the change in inter-event time. This happens to be between +5 days to +55 days with a mean of +20 days. In other words, post shut-in, the average time between two consecutive events has increased by 20 days than in the period from 07/02/2013 to 15/11/2013.
- On the other hand, activity rate for Loppersum is higher than that for the period defined by block 1 which spans from 22/12/1993 to 04/11/1997. $\Delta\lambda$ in now -5 days to -140 days with a mean of -60 days. In other words, the time between two consecutive events now is 60 days shorter than during the 1993 to 1997 period.
- Activity rate in Oost post shut-in is higher than activity rate in Oost during the periods defined by blocks 1 and 2 for this region. The dates covered by these two periods are from

15/05/1995 to 16/04/2009. The change in inter-event time is very large. We can now expect time difference between two subsequent events to be shorter by 600 days compared to 1995-2009 period.

- For all other pre shut-in periods and region combination, activity rate post shut-in is not very different to activity rate in the past.

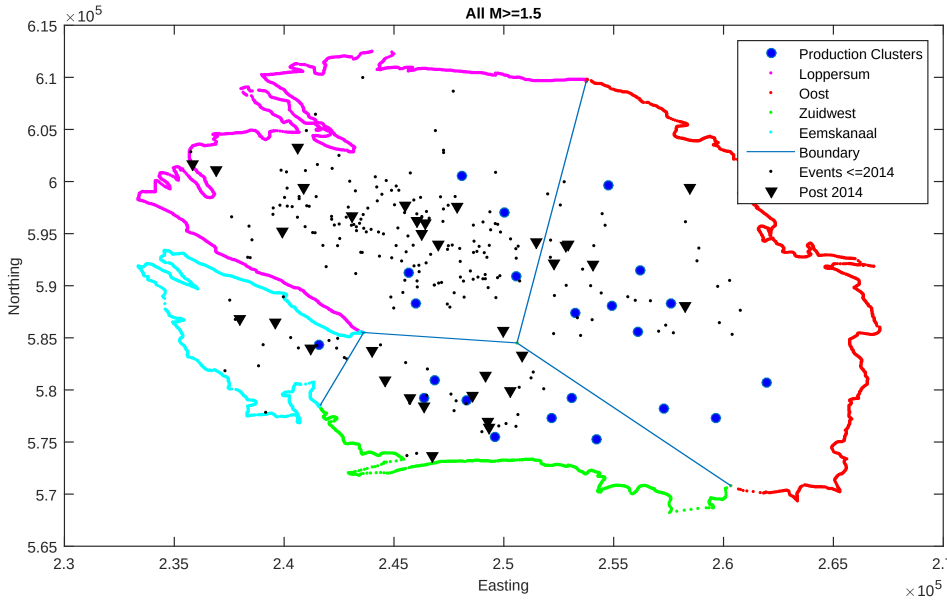

Figure 5.1.: Spatial distribution of all events M⩾1.5

Table 5.1.: Count and proportion of all events M⩾1.5 in each production area pre shut-in and post shut-in

Region	Count pre shut-in	Count post shut-in
Loppersum	152	15
Oost	30	6
Zuidwest	21	11
Eemskanaal	16	3
All	219	35

Table 5.2.: Blocks, start date, end date and days between start and end dates. Loppersum production area. M⩾1.5. Blocks 1-10 refer to pre shut-in period. Block 11 refers to post shut-in period.

Block Number	Start Date	End Date	No. of days
Block 1	22/12/1993	04/11/1997	1413
Block 2	15/02/1998	16/07/2000	882
Block 3	21/06/2001	16/11/2003	878
Block 4	10/06/2004	18/01/2006	587
Block 5	12/02/2006	26/01/2007	348
Block 6	14/05/2007	04/02/2009	632
Block 7	14/04/2009	05/05/2010	386
Block 8	24/07/2010	04/11/2011	468
Block 9	27/11/2011	07/02/2013	438
Block 10	07/02/2013	15/11/2013	281
Block 11	03/02/2014	09/09/2015	583

Block/Event No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean	Sigma
Block 1	91	70	12	86	19	22	11	250	210	2	117	314	40	122	138	100	95
Block 2	103	63	0	127	110	32	52	63	226	54	36	58	26	1	34	66	58
Block 3	340	229	352	7	32	20	10	62	5	60	46	5	27	17	6	81	121
Block 4	207	104	63	2	44	30	17	13	81	23	25	80	18	87	0	53	54
Block 5	25	37	2	24	3	3	55	7	46	0	18	32	9	17	95	25	26
Block 6	108	26	67	33	40	17	187	1	164	8	55	7	1	23	3	49	59
Block 7	69	6	14	3	1	58	17	69	66	19	17	41	65	8	2	30	28
Block 8	80	21	158	9	44	13	61	28	4	32	33	6	9	10	40	37	40
Block 9	23	33	7	24	5	12	1	50	47	28	55	1	148	8	19	31	37
Block 10	0	2	2	0	2	2	137	18	0	43	3	24	4	8	36	19	36
Block 11	80	10	5	28	90	54	73	15	62	135	6	10	4	73	18	44	40

Figure 5.2.: Inter-event times Loppersum M⩾1.5. Blocks 1-10 refer to pre shut-in period. Block 11 refers to post shut-in period.

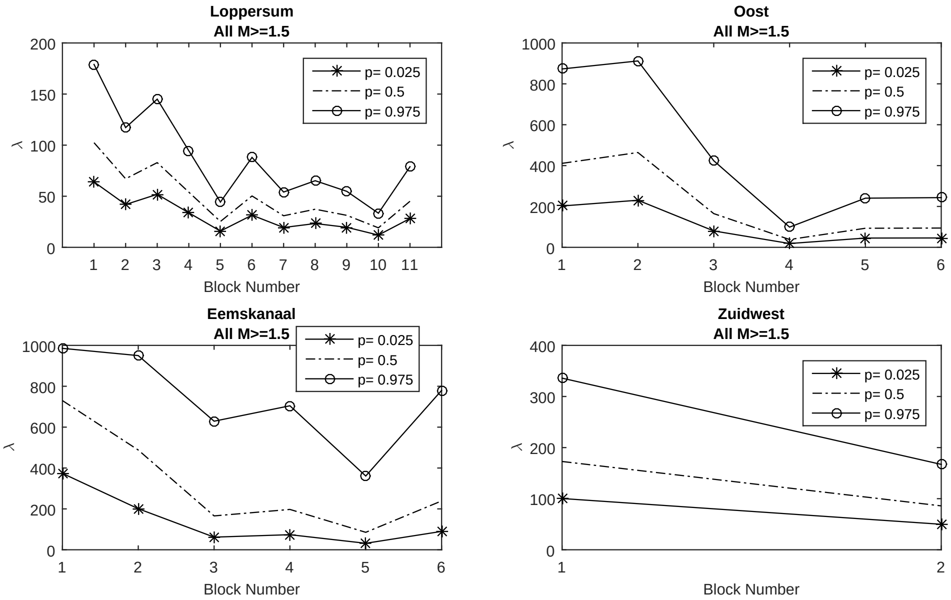

Figure 5.3.: λ for all the blocks in pre shut-in period and post shut-in (the last block in each plot) for the four production area & M⩾1.5.

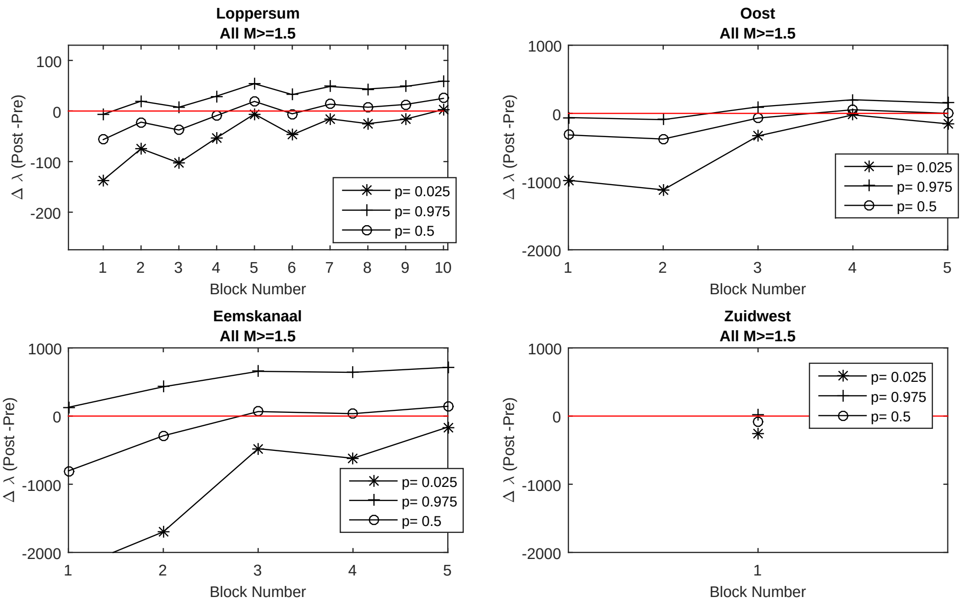

Figure 5.4.: $\Delta\lambda$ post shut-in minus all pre shut-in periods indicated by block number on x-axis for M⩾1.5. The horizontal red line is draw at y = 0. If all the points for a block are above or below the red line, $\Delta\lambda$ for that block is statistically significant.

We now use the second method to understand the change in activity rate. Figure 5.5 shows the temporal trend in the proportion of days $\hat{\alpha}$ an event has been observed from the sample data. We can derive the following **qualitative** conclusions:

- $\hat{\alpha}$ generally increases as we move from left to right in each sub-plot suggesting that activity rate increases with time.

- The increase is not monotonic.
- For Lopersum and Eemskanaal, $\hat{\alpha}$ post shut-in has shown a downward tick. For Oost, there is no directional change in $\hat{\alpha}$ post shut-in. Zuidwest shows an increase in $\hat{\alpha}$ which suggests an increase in activity rate.

In order to **quantitatively** assess changes in activity rate, we use the Z test statistic and the results are shown in figure 5.6. We can conclude that:

- The Z test statistic in Loppersum for block 10 is statistically significant. This means that post shut-in, activity rate in Loppersum is lower than in the period defined by block 10 (07/02/2013 to 15/11/2013).
- The Z test statistic in Loppersum for block 1 is also statistically significant but the statistic suggest that activity rate now is higher than activity rate defined by block 1 period (22/12/1993 to 04/11/1997).
- However, when the post shut-in period is compared to other pre shut-in periods, the Z test statistic is not statistically significantly. Thus for instance, in the post shut-in period, activity in Loppersum is similar to activity rate in the period 27/11/2011 to 07/02/2013 (block 9); 24/07/2010 to 04/11/2011 (block 8) or even further back in time.
- In Oost, activity rate in post shut-in period is higher than the time defined by blocks 1 and 2 in Oost. This covers time period from 15/05/1995 to 16/04/2009.
- In Eemskanaal, activity rate post shut-in period is higher than activity rate defined by the time period for block 1 which spans from 18/03/2001 to 13/04/2006.
- There is no difference in activity rate post shut-in period for other regions & period combinations.

We note that the two methods provide similar evidence on significant shifts in activity rate. We also note that because there are temporal changes in activity rate, it is prudent to search for such evidence using blocks such that each block contains identical number of events. It is not advisable to bundle the entire pre shut-in together and assume temporal homogeneity across regions.

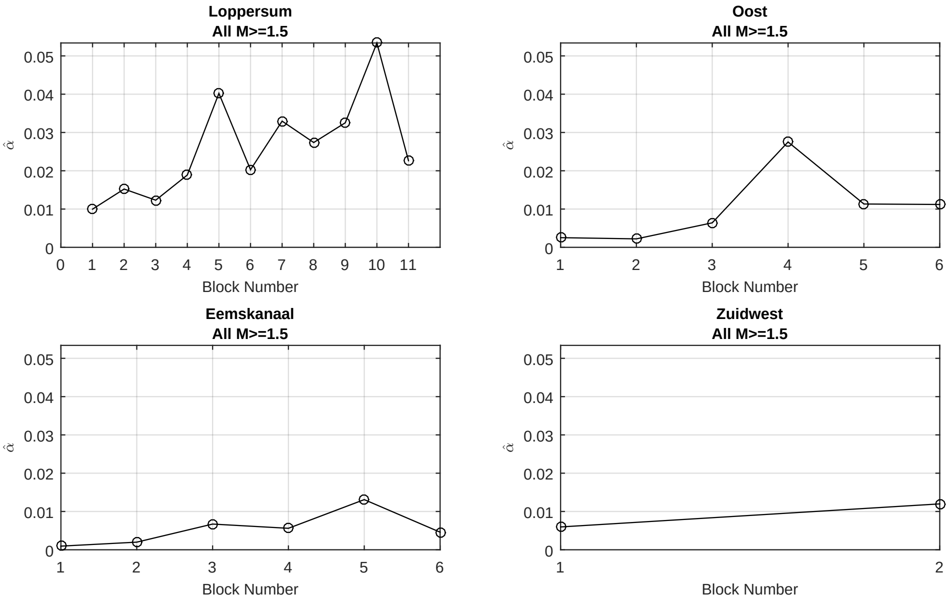

Figure 5.5.: α, the proportion of days on which an event occurred in each block and each production area. M⩾1.5. The last block in each sub-plot refers to post shut-in period. Other blocks refer to pre shut-in periods.

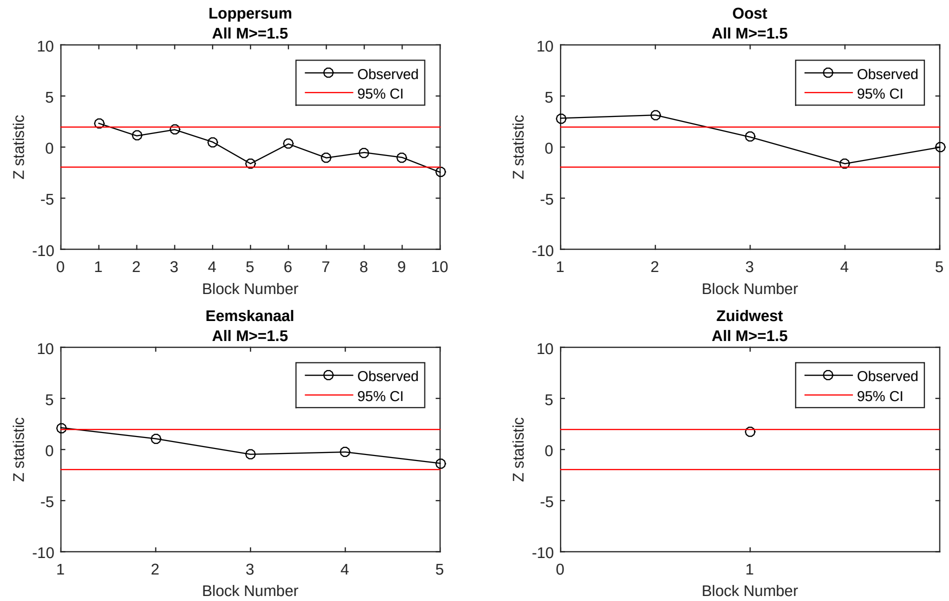

Figure 5.6.: Z statistic, the difference in proportion between post shut-in period and all pre shut-in periods. Red lines show 95% CI. M⩾1.5. If the black line for any block is outside the 95% CI, Z statistic for that block is statistically significant.

6. Statistical inference of change in activity rate M⩾1.0

Figure 6.1 shows the spatial distribution of events of M⩾1.0 in the two time periods and table 6.1 shows the number of events that have occurred in each production area. As has been described before the number of events post shut-in in each production area is used to partition the events pre shut-in into equal sized blocks. This gives 9 blocks for Loppersum such that blocks 1 to 8 belong to the pre shut-in period and block 9 belongs to post shut-in period. The start and end dates are shown in table 6.2. Please refer to appendix B for other three regions.
Posterior λ for the four regions is shown in figure 6.2. We again see a temporal decline in λ as seen for M⩾1.5. However we see that the decline is much smoother. This is because each block now has more events than in the case of M⩾1.5. Thus for instance, for Loppersum, each block now has 36 events as against 15 when data set with M⩾ 1.5 was used. We note again that for Loppersum and Eemskanaal, there is an upward trend in λ post shut-in as indicated by the last block for the respective regions. This **qualitatively** suggest that activity rate post shut-in is lower than pre shut-in. There is some evidence of an upward trend in Oost but this is not very clear. Zuidwest, on the other hand, shows a decline suggesting that activity rate has increased.
$\Delta\lambda$ provides **quantitative** evidence on change in seimicity. This is shown in figure 6.3. The main inferences are:

- For Loppersum, post shut-in has lower activity rate that the period defined by block 8 (11/01/2013 to 23/12/2013). The expected time difference between two successive events in post shut-in has increased by 7 days.
- On the other hand, post shut-in activity rate is higher than in the periods defined by blocks 1 and 2 which span from 08/06/1994 to 26/12/2003. The expected time difference has decreased by 40 days and 30 days respectively.
- In Zuidwest, events are occurring more frequently than ever before. Post shut-in inter-event time in this region has decreased by 150 days when compared to block 1 which spans from 20/11/1997 to 18/11/2009 or 35 days when compared to block 2 which spans from 21/12/2009 to 22/12/2013.
- Other pre shut-in and region combinations do not show statistically significant differences in activity rate compared to the post shut-in period.

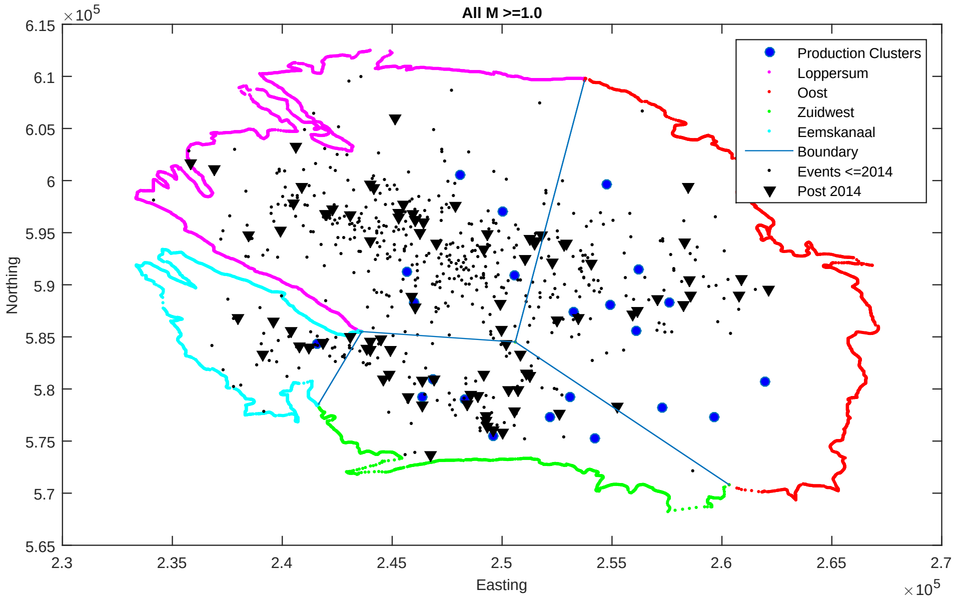

Figure 6.1.: Spatial distribution of all events M⩾1.0

Table 6.1.: Count and proportion of all events M⩾1.0 in each production area pre shut-in and post shut-in

Region	Count pre shut-in	Count post shut-in
Loppersum	303	36
Oost	101	18
Zuidwest	65	29
Eemskanaal	38	8
All	507	91

Table 6.2.: Blocks, start date, end date and days between start and end dates. Loppersum production area. M⩾1.0. Blocks 1-8 refer to pre shut-in period. Block 9 refers to post shut-in period.

Block Number	Start Date	End Date	No. of days
Block 1	08/06/1994	20/12/1999	2021
Block 2	12/02/2000	26/12/2003	1413
Block 3	24/01/2004	22/04/2006	819
Block 4	02/05/2006	26/10/2008	908
Block 5	29/10/2008	24/07/2010	633
Block 6	14/08/2010	15/09/2011	397
Block 7	25/09/2011	25/12/2012	457
Block 8	11/01/2013	23/12/2013	346
Block 9	03/02/2014	09/09/2015	583

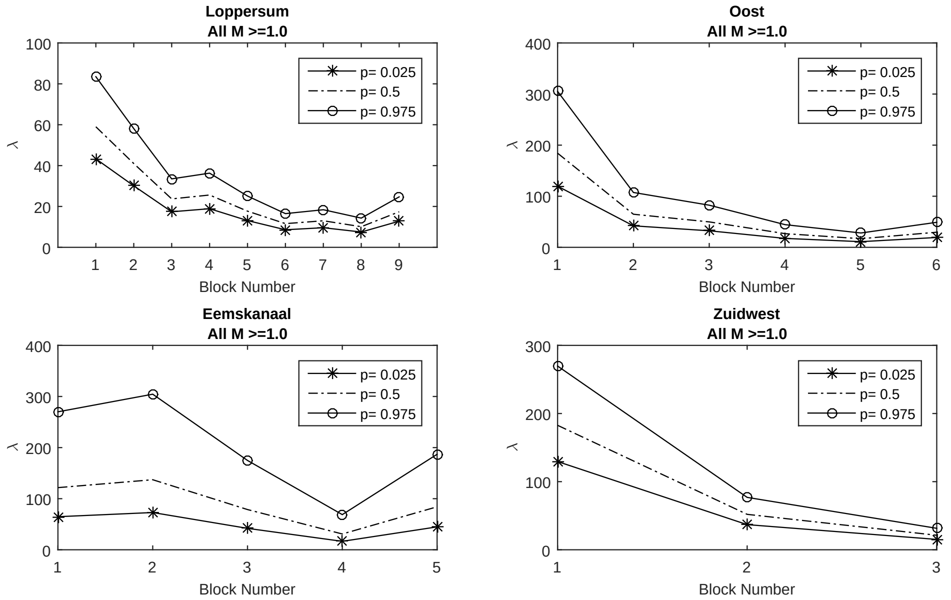

Figure 6.2.: λ for all the blocks in pre shut-in period and post shut-in (the last block in each plot) for the four production area & M⩾1.0

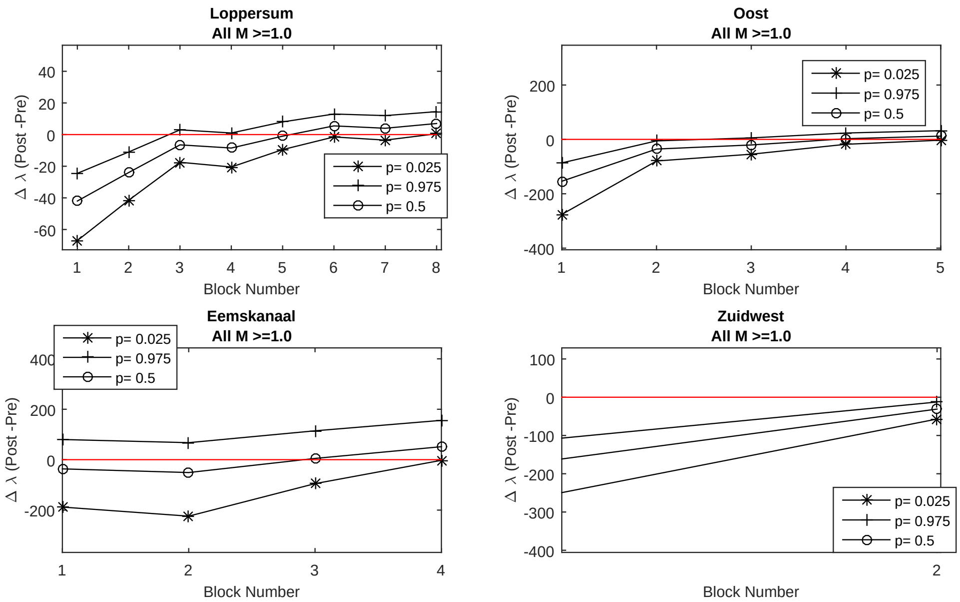

Figure 6.3.: $\Delta\lambda$ post shut-in minus all pre shut-in periods indicated by block number on x-axis for M⩾1.0. The horizontal red line is draw at y = 0. If all the points for a block are above or below the red line, $\Delta\lambda$ is statistically significant.

Figure 6.4 shows temporal behaviour of $\hat{\alpha}$ for all the blocks. The last block in each sub-plot shows $\hat{\alpha}$ post shut-in while the other blocks are for pre shut-in periods. Figure 6.5 shows Z statistic for post shut-in block when compared with pre shut-in blocks. Quantitative inferences one can draw using the the second method are similar to those drawn using the first method.

- The Z test statistic of block 8 for Loppersum is significant. Since the test statistic is negative, there are less events post shut-in compared to the period 11/01/2013 to 23/12/2013 defined by this block.

- The Z statistic is also significant for blocks 1 and 2 which span from 08/06/1994 to 26/12/2003.
- The Z statistic for Zuidwest is significant. There are more events now than ever before.
- Other pre shut-in and region combinations do not show statistically significant difference in activity rate compared to the post shut-in period.

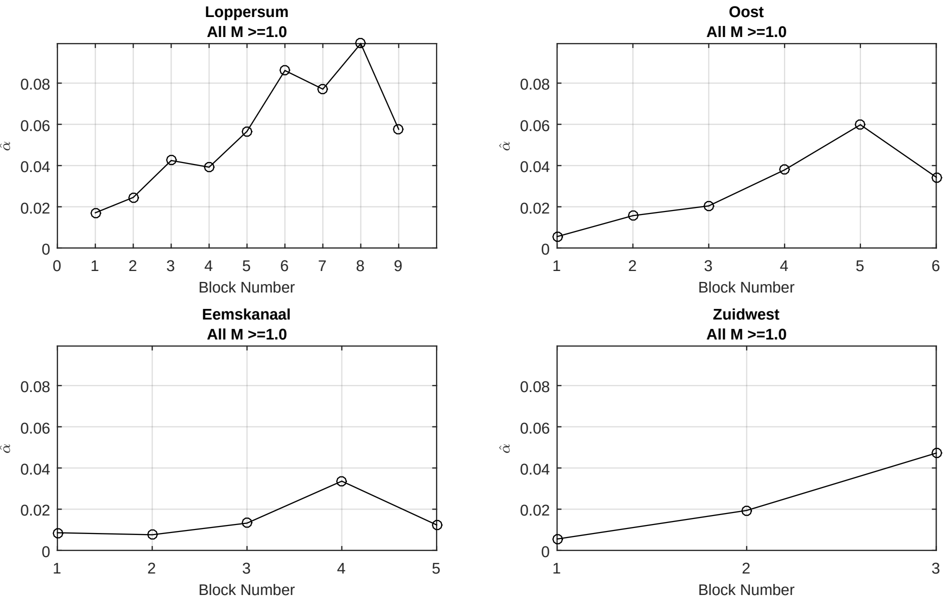

Figure 6.4.: α, the proportion of days on which an event occurred in each block and each production area. M⩾1.0. The last block in each sub-plot refers to post shut-in period. Othe blocks refer to pre-shut-in periods.

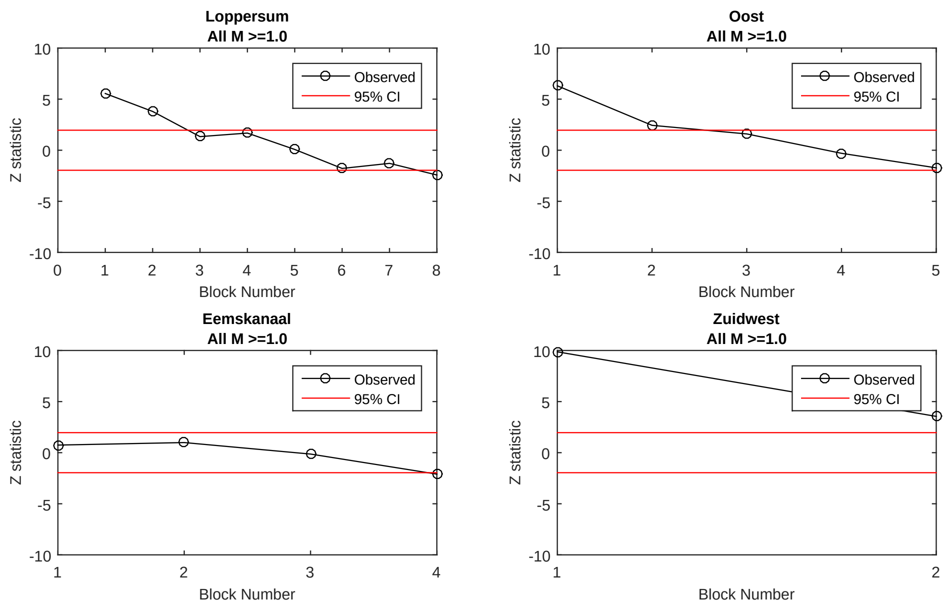

Figure 6.5.: Z statistic , the difference in proportion between post shut-in period and all pre shut-in periods. Red lines show 95% CI. M⩾1.0. If the black line for any block is outside the 95% CI, Z statistic is statistically significant.

7. Statistical inference of change in activity rate M⩾0.5

Figure 7.1 shows the spatial distribution of events of M⩾0.5 in the two time periods and table 7.1 shows the number of events that have occurred in each production area. The increase in activity rate in Zuidwest is apparent. In a short time span post shut-in period, 50 events have occurred in Zuidwest where as it took more than 20 years to accumulate 109 events in the entire pre shut-in period.
Following the established procedure, events in the pre shut-in period were partitioned into blocks. The start and end dates for Loppersum are shown in table 7.2. Similar tables for other regions are shown in Appendix C. Posterior λ, $\Delta\lambda$, $\hat{\alpha}$ and Z statistic are shown in figures 7.2, 7.3, 7.4 and 7.5. In brief, we can quantitatively conclude the following:

- Activity in Loppersum post shut-in is lower than activity rate in Loppersum during the period 05/01/2013 to 23/12/2013 (block 6) and we can expect the time difference between two successive events to be prolonged by 5 days.
- Activity rate in Loppersum post shut-in is more pronounced than during the time interval 29/04/1997 to 22/07/2009 (blocks 1 to 3). The difference in time between two consecutive events has decreased by 10 to 30 days.
- Zuidwest is showing highest seismic activity now than in anytime before. The time interval between successive events has shortened by 50 days.
- Oost is more active now than during time period 14/05/2003 to 10/07/2008 (block 1). The time interval between successive events has shortened by 50 days.
- Eemskanaal does not show a change in activity rate over time.
- For other combinations of time periods pre shut-in and regions, there is no change in activity rate.

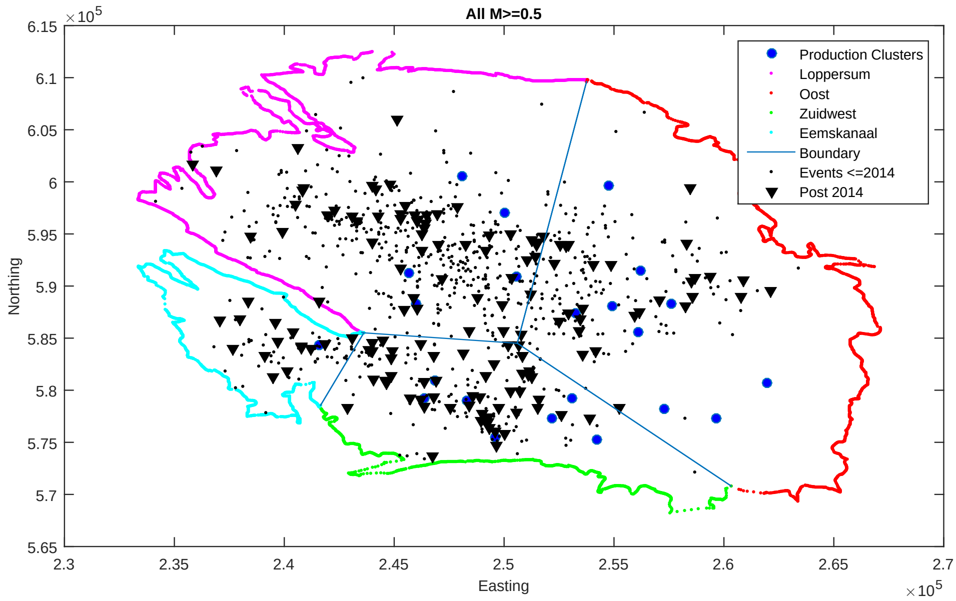

Figure 7.1.: Spatial distribution of all events M⩾0.5

Table 7.1.: Count and proportion of all events M⩾0.5 in each production area pre shut-in and post shut-in

Region	Count pre shut-in	Count post shut-in
Loppersum	360	54
Oost	143	29
Zuidwest	109	50
Eemskanaal	52	15
All	664	148

Table 7.2.: Blocks, start date, end date and days between start and end dates. Loppersum production area. M⩾0.5. Blocks 1 to 6 refer to pre shut-in period. Block 7 refers to post shut-in period.

Block Number	Start Date	End Date	No. of days
Block 1	29/04/1997	27/09/2003	2342
Block 2	24/10/2003	26/08/2006	1037
Block 3	27/09/2006	22/07/2009	1029
Block 4	29/09/2009	04/09/2011	705
Block 5	06/09/2011	25/12/2012	476
Block 6	05/01/2013	23/12/2013	352
Block 7	12/01/2014	09/09/2015	605

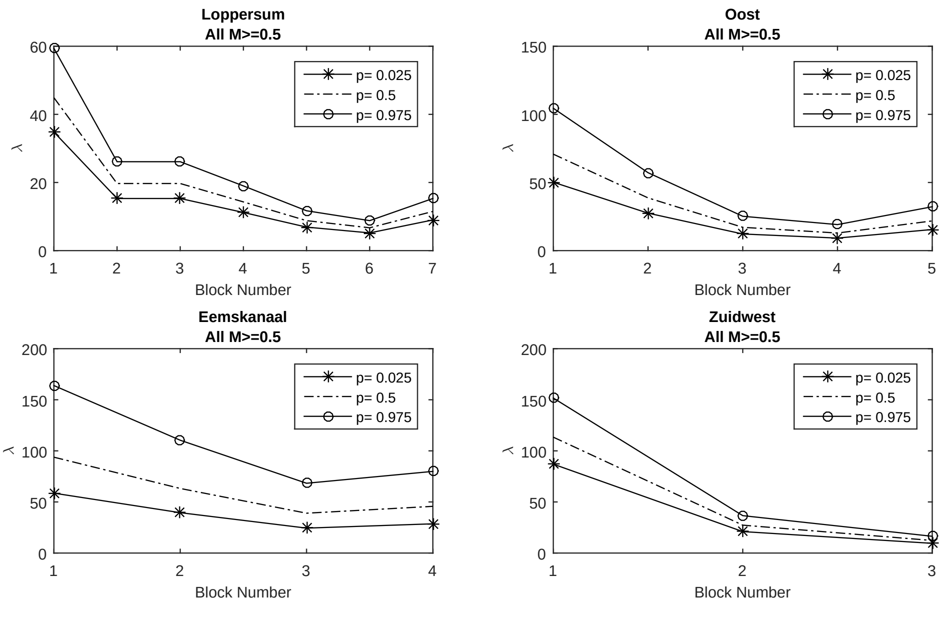

Figure 7.2.: λ for all the blocks in pre shut-in period and post shut-in (the last block in each plot) for the four production area & M⩾0.5

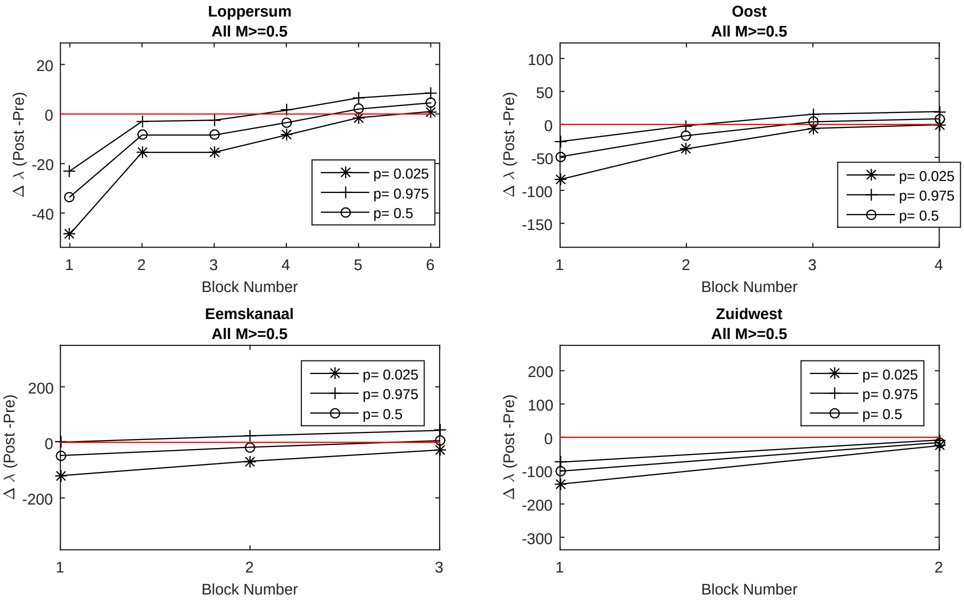

Figure 7.3.: $\Delta\lambda$ post shut-in minus all pre shut-in periods indicated by block number on x-axis for M⩾0.5. The horizontal red line is draw at y = 0. If all the points for a block are above or below the red line $\Delta\lambda$ for that block is statistically significant.

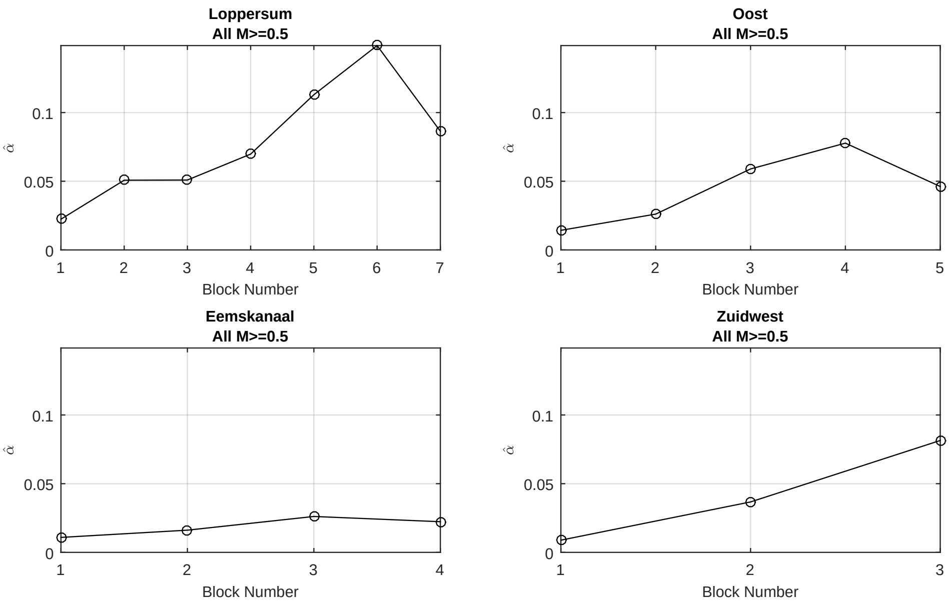

Figure 7.4.: α, the proportion of days on which an event occurred in each block and each production area. M⩾0.5. The last block in each sub-plot refers to post shut-in period. Other blocks refer to pre shut-in periods.

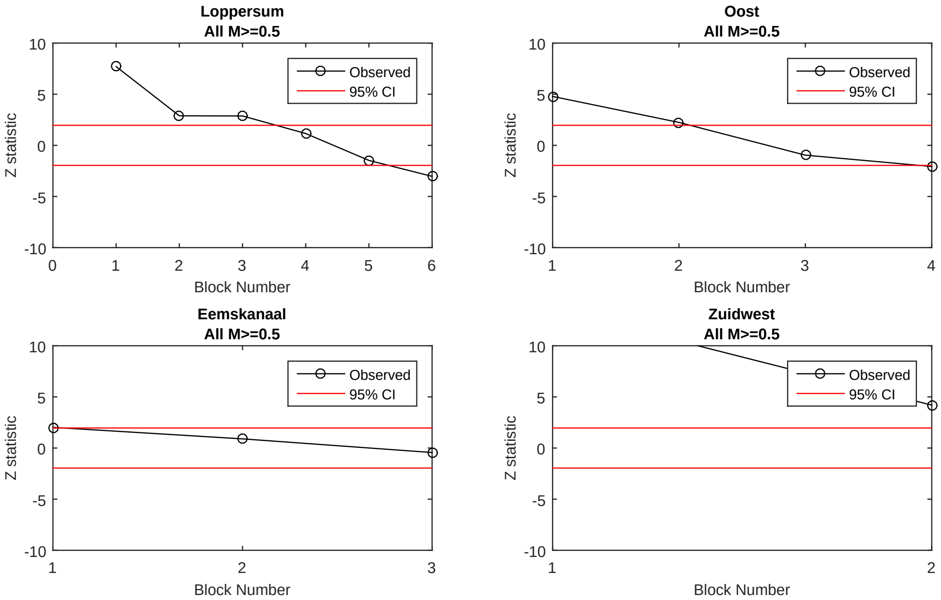

Figure 7.5.: Z statistic, the difference in proportion between post shut-in period and all pre shut-in periods. Red lines show 95% CI. M⩾0.5. If the black line for any block is outside the 95% CI, Z statistic for that block is statistically significant.

8. Conclusions

The report describes the statistical methods that can be used to understand the change in activity rate before the production shut-in in Loppersum and after production shut-in in the four production areas of the Groningen field. We have used λ, inter-event time, as the parameter to qualitatively understand the temporal trends in activity rate while $\Delta\lambda$ is use to quantitatively assess the significant differences in activity rate post shut-in compared to various time periods pre shut-in. The work shows that there is indeed a temporal change in activity rate for all the four regions and for all the magnitudes. When the post shut-in period is compared to the most adjacent pre shut-in period such that both of them has the same number of events, we derive the following conclusions from chapters 5, 6 and 7;

- The activity rate in Loppersum area post shut-in is lower than the activity rate in Loppersum in the pre shut-in period from Jan 2013 up to Dec 2013. This is true for events of all magnitudes.
- In Loppersum, post shut-in, we now expect inter-event time, the time between two consecutive events, to have increased by 20 days for $M \geqslant 1.5$. The expected increase is 7 days for $M \geqslant 1.0$ and 5 days $M \geqslant 0.5$ compared to the period from Jan 2013 up to Dec 2013.
- The activity rate in Zuidwest post shut-in has increased for events of $M \geqslant 1.0$ and $M \geqslant 0.5$ compared to any pre shut-in period. The increase in activity rate is despite the fact that production in Zuidwest post shut-in is at a historic low, along with production in the other three production areas. There is no noticeable change in activity rate yet when $M \geqslant 1.5$ are considered.
- In Zuidwest, post shut-in, we can expect inter-event time between to consecutive events to have decreased by 35 days for $M \geqslant 1.0$ compared to pre shut-in spanning from Dec 2009 to Dec 2013. The expected decrease is 50 days for $M \geqslant 0.5$ when post shut-in period is compared to pre shut-in period from Dec 2010 to Dec 2013.

We do not see a statistically significant difference in activity rate in Oost and Eemskanaal between the post shut-in period and pre shut-in period that is most adjacent to post shut-in period with the same number of events.

We conclude that, due to temporal trends in activity rate for each region and each magnitude, the inferences on the difference in activity rate post shut-in and the closest pre shut-in period drawn above should not be generalized and extended beyond the time frame described. Thus, while the activity rate in Loppersum post shut-in is *lower* than in the period from Jan 2013 to Dec 2013 for all magnitudes, activity rate post shut-in is *similar* to activity rates in Loppersum in the following periods:

- from 1998 to 2012 for $M \geqslant 1.5$,
- from 2003 to 2012 for $M \geqslant 1.0$ and
- from 2009 to 2012 for $M \geqslant 0.5$

We have applied two methods to derive quantitative inferences and both of them provide similar results. We believe the work done is therefore robust and conclusions are reliable.

9. Future work

In the current work, two methods are used. Both the methods assume a certain type of distribution from which samples are drawn. The assumption though reasonable needs to be thoroughly tested by fitting various distributions to a sample of events when the activity rate is stationary. Further, when a suitable distribution that describes the observed events is found to be most appropriate, the likelihood function in equation 2.3 should be changed and the posterior density should be re-evaluated based on such information. We also intend to estimate λ and $\Delta\lambda$ purely by bootstrapping the catalogue data. This avoids any distributional assumptions and so the inferences will be more robust and reliable. So far, we have not taken such an approach since the sample size of the number of events post shut-in has been small. We believe a sample size of 20 events is necessary for the bootstrapping method to be reliable.

The methodology and the results will be discussed with various stakeholders. Based on the outcome of the discussions, we feel that there is a need to deploy a simple application in the 'cloud' such that various stakeholders can interact with the application and update themselves with changes in activity rate using the most recent catalogue. This however requires consent of all the stakeholders and risk/benefits should be properly understood.

Appendix A.

Tables & Figures M⩾1.5.

Table A.1.: Blocks, start date, end date and days between start and end dates. Oost production area. M⩾1.5. Last block refers to post shut-in period. Other blocks refer to pre shut-in period.

Block Number	Start Date	End Date	Date difference
Block 1	15/05/1995	12/11/2001	2373
Block 2	10/05/2002	16/04/2009	2533
Block 3	09/05/2010	09/11/2011	549
Block 4	15/11/2011	14/06/2012	212
Block 5	15/06/2012	26/11/2013	529
Block 6	15/03/2014	16/05/2015	427

Table A.2.: Blocks, start date, end date and days between start and end dates. Zuidwest production area. M⩾1.5. Last block refers to post shut-in period. Other blocks refer to pre shut-in periods.

Block Number	Start Date	End Date	Date difference
Block 1	18/11/2009	11/02/2013	1181
Block 2	11/03/2014	18/08/2015	525

Table A.3.: Blocks, start date, end date and days between start and end dates. Eemskanaal production area. M⩾1.5. Last block refers to post shut-in period. Other blocks refer to pre shut-in periods.

Block Number	Start Date	End Date	Date difference
Block 1	18/03/2001	13/04/2006	1852
Block 2	16/02/2007	30/05/2010	1199
Block 3	21/06/2010	20/08/2011	425
Block 4	09/10/2011	05/02/2013	485
Block 5	16/08/2013	22/09/2013	37
Block 6	26/01/2014	07/07/2015	527

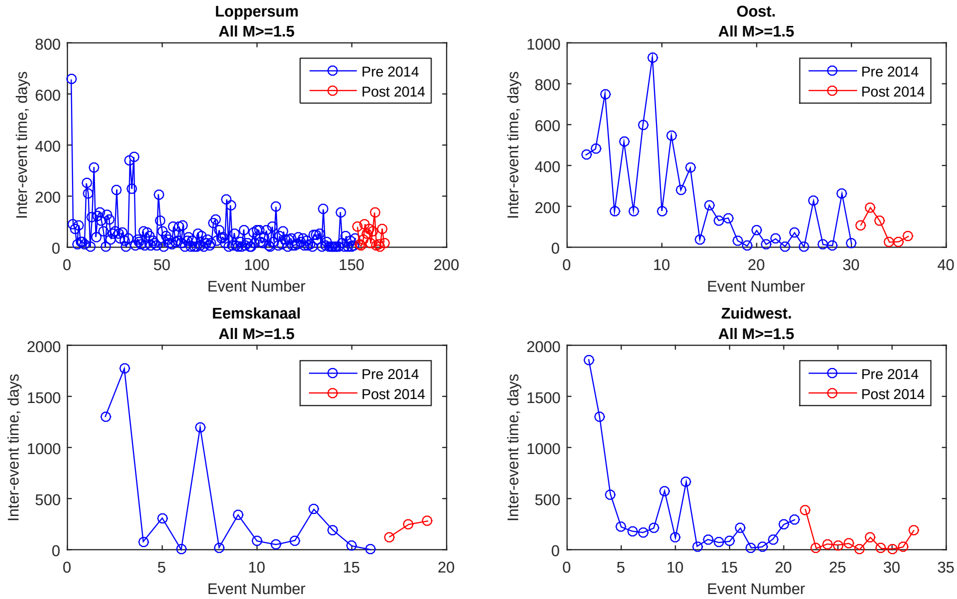

Figure A.1.: inter-event time pre shut-in (blue) and post shut-in(red) for the four regions & M⩾1.5

.

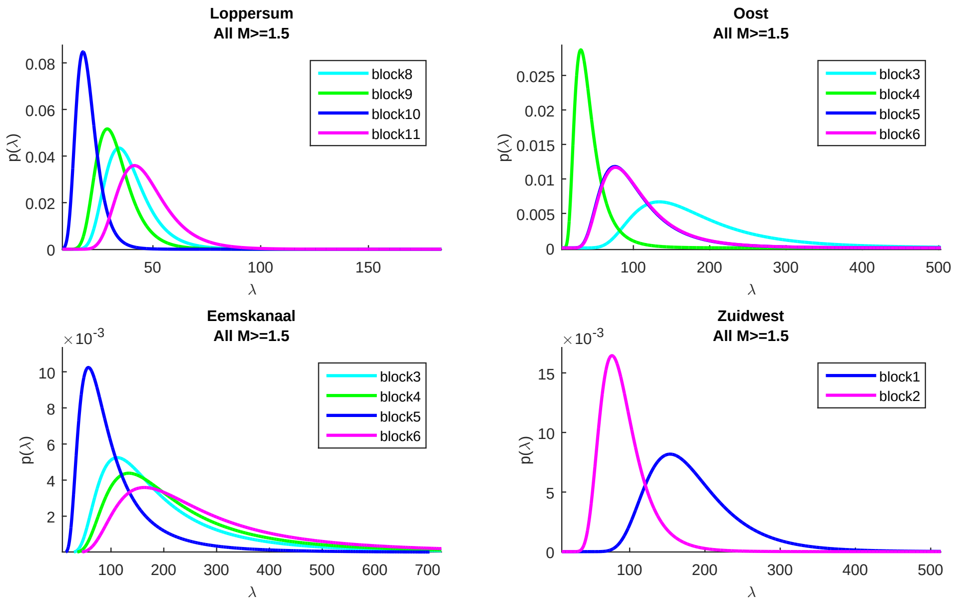

Figure A.2.: λ for block in post shut-in(magenta) and for three blocks in pre shut-in (blue, green and red) closest to 2014 for the four production area & M⩾1.5

.

Appendix B.

Tables M⩾1.0.

Table B.1.: Blocks, start date, end date and days between start and end dates. Oost production area. M⩾1.0. Last block refers to post shut-in period. Other blocks refer to pre shut-in periods.

Block Number	Start Date	End Date	Date difference
Block 1	06/06/1997	05/04/2006	3225
Block 2	12/04/2006	26/05/2009	1140
Block 3	26/05/2009	24/10/2011	881
Block 4	09/11/2011	10/02/2013	459
Block 5	10/02/2013	08/12/2013	301
Block 6	16/01/2014	16/05/2015	485

Table B.2.: Blocks, start date, end date and days between start and end dates. Zuidwest production area. M⩾1.0. Last block refers to post shut-in period. Other blocks refer to pre shut-in periods.

Block Number	Start Date	End Date	Date difference
Block 1	20/11/1997	18/11/2009	4381
Block 2	21/12/2009	22/12/2013	1462
Block 3	02/01/2014	28/08/2015	603

Table B.3.: Blocks, start date, end date and days between start and end dates. Eemskanaal production area. M⩾1.0. Last block refers to post shut-in period. Other blocks refer to pre shut-in periods.

Block Number	Start Date	End Date	Date difference
Block 1	23/01/2006	23/07/2008	912
Block 2	20/09/2008	10/06/2011	993
Block 3	20/08/2011	03/02/2013	533
Block 4	05/02/2013	29/09/2013	236
Block 5	04/01/2014	07/07/2015	549

Appendix C.

Tables M⩾0.5.

Table C.1.: Blocks, start date, end date and days between start and end dates. Oost production area. M⩾0.5.Last block refers to post shut-in period. Other blocks refer to pre shut-in periods.

Block Number	Start Date	End Date	Date difference
Block 1	14/05/2003	10/07/2008	1884
Block 2	01/01/2009	27/07/2011	937
Block 3	08/10/2011	30/11/2012	419
Block 4	11/01/2013	08/12/2013	331
Block 5	16/01/2014	28/08/2015	589

Table C.2.: Blocks, start date, end date and days between start and end dates. Zuidwest production area. M⩾0.5. Last block refers to post shut-in period. Other blocks refer to pre shut-in periods.

Block Number	Start Date	End Date	Date difference
Block 1	15/07/1995	31/03/2010	5373
Block 2	25/04/2010	22/12/2013	1337
Block 3	02/01/2014	28/08/2015	603

Table C.3.: Blocks, start date, end date and days between start and end dates. Eemskanaal production area. M⩾0.5. Last block refers to post shut-in period. Other blocks refer to pre shut-in periods.

Block Number	Start Date	End Date	Date difference
Block 1	02/01/2006	19/08/2009	1325
Block 2	30/05/2010	04/03/2012	644
Block 3	08/08/2012	29/09/2013	417
Block 4	04/01/2014	02/08/2015	575

References

[1] NAM. Technische bijlage behorend bij het Meet- en Regelprotocol Groningen . *NAM Policy Document*, 2015.

[2] R.Paleja, M.Jones, and S.Bierman. DRAFT: A2.1 Activity rate Loppersum. *Internal Shell Report*, 2015.

[3] Dirk Kraaijpoel, Jesper Spetzler, and Bernard Dost. Review of 3 reports by NAM . *KNMI*, 2015.

[4] S.J.Bourne and S.J.Oates. An Activity rate model of induced seismicity withing the Groningen Field. *Internal Shell Report*, Sept 2014.

[5] Wikipedia. Exponential Distribution. 2013. [Online; accessed 19-Oct-2015].

[6] Wikipedia. Inverse Gamma Distribution. 2013. [Online; accessed 19-Oct-2015].

[7] John.I.Gallin and Frederic.P.Ognibene. Principles and Practice of Clinical Research. *Elsevier*, 2012.

Bibliographic information

Classification	Restricted
Report Number	SR.15.13107
Title	Impact of Production Shut-in on Inter-Event time in Groningen.
Subtitle	A statistical perspective.
Authors	Rakesh Paleja (PTD/ TASE); Stijn Bierman(PTD/TASE); Matthew Jones (PTD/TASE)
Keywords	Groningen, Activity rate, inter-event time, Loppersum, Shut-in, Production
Issue Date	October 2015
Period of work	Jan 2014 Oct 2015
US Export Control	Non US — Public Domain
WBSE Code	ZZPT/015656/0101125
Reviewed by	Maggie Wenham (PTD/TASE)
Approved by	Phil Jonathan (PTD/TASE)
Sponsoring Company/ Customer	Shell Research B.V
Spons./Cust. Address	
Issuing Company	Shell Global Solutions (UK) (a division of Shell Research Ltd.) Brabazon House Concord Business Park Threapwood Road Manchester M22 0RR England

Report distribution

Electronic distribution (PDF)

Name, Company, Ref. Ind.	
PT Information Services, PTT/TIKE, PT-Information-Services@Shell.com	PDF
Electronic distribution (PDF) Name, Company, Ref. Ind. PT Information Services, PTT/TIKE, PT-Information-Services@Shell.com	PDF
Jonathan Philip GSUK-PTD/TASE	
Gibbs Rick, GSUSI-PTD/TASE	PDF
Bierman Stijn, GSUK-PTD/TASE	PDF
Greenall Carol, GSUK, PTD/FTCT	PDF
Maron Karel, P NAM-UIO/T/DGR	PDF
van Elk Jan, NAM-UIO/T/DGR	PDF
Bourne Stephen, GSNL-PTI/RC	PDF
Harris Christopher, GSNL-PTI/RC	PDF
Wentinck Rick, GSNL-PTI/RC	PDF
Paleja Rakesh, GSUK-PTD/TASE	PDF
Kraaijeveld Famke, GSNL-PTI/RC	PDF
Park Timothy, GSUK-PTD/TASE	PDF
Randell David, GSUK-PTD/TASE	PDF
van den Bogert Peter, AJ GSNL-PTU/E/Q	PDF
Beckers Janwillem, NAM-UIO/T/D	PDF
Wenham Maggie, SEPL PTD/TASE	PDF
Das Saptarshi, PTIN-PTI/CT	PDF
van Beest Bertwim, PTIN-PTI/CT	PDF
Reppart William, GSUSI-PTD/TASE	PDF

Paper copy distribution

Name, Company, Ref. Ind.	*No. of copies*

Principal typesetting performed with LaTeX system (MiKTeX) using the PTreport2015.cls class based on the KOMA-Script with the necessary modifications to match the Shell Projects and Technology house style.

The PTreport2015.cls class file is written and maintained by IDLTeX (www.idltex.com). All rights reserved.