DRAFT: A2.1 Activity rate Loppersum

Rakesh Paleja, Matthew Jones, David Randell, Stijn Bierman

April 30, 2015

1 Summary

On 1st Jan 2014, the rate of production in the area of Loppersum was reduced. Here we seek to understand if reduction in production has changed the characteristics of inter-event times for seismic events in the vicinity of Loppersum. This reports shows that if only events of magnitude >= 1.5 are considered, there is little or no evidence for change in production having influenced inter-event time. If all events (i.e. events of all magnitudes) are considered (as reported in the KNMI catalogue), there is considerable evidence that the so-called '**catalogue** inter-event time' after production change is different to the earthquake inter-event time immediately prior to the production change. Specifically, we find that events of small magnitude were more frequent before production change than after. This observation should be viewed with considerable caution, since (1) recording of events of magnitude < 1.5 is uncertain both in spatial and time domain (according to the information provided by KNMI), and (2) the period of observation since production change is short.
It is further recommended that a further study of inter-event time with magnitudes >= 1.0 should be conducted to understand the influence of change in production more fully, for carefully-specified spatial subdomains. These spatial subdomains should include locations within the Loppersum boundary where trusted geophones are located, partially reducing uncertainty in our analysis for events of magnitude < 1.5 (since we do not know the exact location of the geophones).

2 Methodology

2.1 Inter-event time

Let there be n number of events after the production was reduced on 1st Jan 2014 up to the current date (18th Jan 2014, the last date in the event catalog) in the area defined by the production clusters covering Loppersum. 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 parameter and one assigns λ a prior $p(\lambda)$. Bayesian posterior density of λ is given by the following equation

$$p(\lambda|t_1, t_2...t_n) = \frac{f(t_1, t_2, ..t_n|\lambda)p(\lambda)}{\int_0^\infty f(t_1, t_2, ..t_n|\lambda)p(\lambda)d\lambda} \tag{1}$$

The above equation can be simplified to proportionality so that:

$$p(\lambda|t_1, t_2...t_n) \propto f(t_1, t_2, ..t_n|\lambda)p(\lambda) \tag{2}$$

If activity rate follows a Poisson process (ref Bourne and Oates Activity rate model), t_1, t_2 .. t_n are from Exponential distribution.

$$f(t_1, t_2, ..t_n|\lambda) = \prod_{i=1}^{n} f(t_i|\lambda) \propto \frac{1}{\lambda^n} e^{\frac{-\sum t_i}{\lambda}} \tag{3}$$

This can be written as:

$$f(t_1, t_2, ..t_n|\lambda) \propto \lambda^{-n} e^{\frac{-s}{\lambda}} \quad (4)$$

where s is the summation of the n inter-event time.

Substituting 4 in 2 gives the posterior density of λ as:

$$p(\lambda|t_1, t_2...t_n) \propto \lambda^{-n} e^{\frac{-s}{\lambda}} p(\lambda) \quad (5)$$

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

$$p(\lambda) \propto \frac{1}{\lambda} \quad (6)$$

which when substituted in 5 leads to the posterior density

$$p(\lambda|t_1, t_2...t_n) \propto \frac{\lambda^{-n} e^{\frac{-s}{\lambda}}}{\lambda} \quad (7)$$

or

$$p(\lambda|t_1, t_2...t_n) \propto \lambda^{-n-1} e^{\frac{-s}{\lambda}} \quad (8)$$

It is often beneficial to deal with log-likelihood for the posterior density in order to deal with small numbers which leads to:

$$log(p(\lambda|t_1, t_2...t_n)) \propto -(n+1) log\lambda - \frac{s}{\lambda} \quad (9)$$

It is also possible to assume that the prior density $p(\lambda)$ is the non-informative Jeffrys' prior which would lead to the posterior density to be slightly different.

$$log(p(\lambda|t_1, t_2...t_n)) \propto -(n+\frac{1}{2}) log\lambda - \frac{s}{\lambda} \quad (10)$$

Also, one can assume the prior density to be uniformly distributed which leads to

$$log(p(\lambda|t_1, t_2...t_n)) \propto -nlog\lambda - \frac{s}{\lambda} \quad (11)$$

It should be clear from equations 9, 10 and 11 that the assumption about the prior can influence the posterior density to some extent. Also, for an exponential variate, an informative prior can come from gamma distribution with two known non-zero shape parameters α and β. While we can estimate them for the period prior to production change, it will be hard to estimate them for the period after the production was changed due to the limited sample size. For the purpose of the this report, we will therefore use equation 9 while posteriors based on other non informative or informative prior can be computed at a later stage if required. Also, for the sake of simplicity, we henceforth refer to the posterior density $p(\lambda|t_1, t_2...t_n)$ simply as $p(\lambda)$.

In order to estimate the posterior density, we follow Markov Chain Monte Carlo (MCMC) approach. To do so, we initiate a sequence of λ with an initial value λ^0 which could be a random draw from a uniform distribution. In the next step, we simulate a candidate value λ^* as the 1^{st} proposal in the sequence from a normal proposal density $g(\lambda^*|\lambda^0)$ and compute the difference

$$R = log(p(\lambda^*)) - log(p(\lambda^0)) \quad (12)$$

If $R > log(u)$ where u is a simulated random uniform random variate [0, 1], then the proposal λ^* is accepted in the sequence and λ^1 is λ^* otherwise λ^1 is λ^0. We now repeat the steps i.e. generate λ^*, from normal proposal density $g(\lambda^*|\lambda^1)$ and compute the difference :

$$R = log(p(\lambda^*)) - log(p(\lambda^1)) \tag{13}$$

λ^2 is λ^* if $R > log(u)$ otherwise λ^2 is λ^1. In this way, λ^3, ... λ^k where k is some large number are simulated in the sequence. The end of the chain is stable and can be used to compute the kernel density of λ.

The above discussion for estimating the posterior density for λ after the production change occurred can also be used to estimate the posterior density of λ before the production change between the period 2003 to Dec 2013. In this case, if N is the total number of events for the entire period (2003 to 2013) but in the same area that defines n, $N >>> n$ simply because of the time period prior to the production change has allowed large number of events to occur over time. So, to estimate λ prior to the change in production in such a way so that it can be compared with λ after the change in production, we simply divide the N events into k contiguous blocks, each of size n and use equation 9 for each block. There will be k λ's and λ_{k-2} to λ_k should be informative for comparison with λ after the production was changed on 1st Jan 2014. This is because λ for blocks $k-2$ to k which refer to the time period before the production change are closest in time to λ for the block after the production was reduced and the influence of other covariates on the seismic activity rate can be ruled out. If the density functions of λ before and after the production change are overlapping, we can conclude that the weight of evidence suggests that production change has resulted in a small change in the activity rate up to the current date. On the other hand if the density functions do not overlap, we can infer that the weight of evidence for a change in activity rate due to change in production rate is large.

2.2 Defining the Influence Area

We now need a systematic way to define the area whose seismic activity can be influenced by production change in Loppersum as this defines n , N and s used to estimate λ. In all there are 22 production clusters in the Groningen field. Five of these production clusters namely Leermens, Overschild, Ten Post, t Zandt and De Pauwen belong to Loppersum. They are north west of Groningen field and their production was cut on 1st January 2014 as shown in figure 1. Raw production is on a daily basis and smooth production is a 30 day moving average with zero phase lag. Notice that the production post Jan 14 does not remain at zero at all times and it seems some volume is produced during high demand, winter season. Also, in pre-2014 period, the production for each cluster is highly cyclical with peaks near December/January and troughs in a summer month.

There are three possible ways to understand the seismic area influenced by these five clusters qualitatively. This is shown in figure 2. In the first scenario, the area of seismic influence is assumed to not only cover the interior area drawn by the polygon that defines the five clusters but also extends in the north west direction. This is a fair assumption as there are no production clusters north west of Ten Post and t Zandt. It is therefore safe to say that only production from these two clusters can cause seismic events in north west direction if we assume a causal relationship between production and events. In the second scenario, the area of seismic influence is assumed to be moderate and extends only few kilometers to the north west. In the third scenario, the area within the polygon and a small area outside the polygon is assumed to be influenced. Notice that in all the three scenarios, the area influenced is not overlapping with other production clusters.

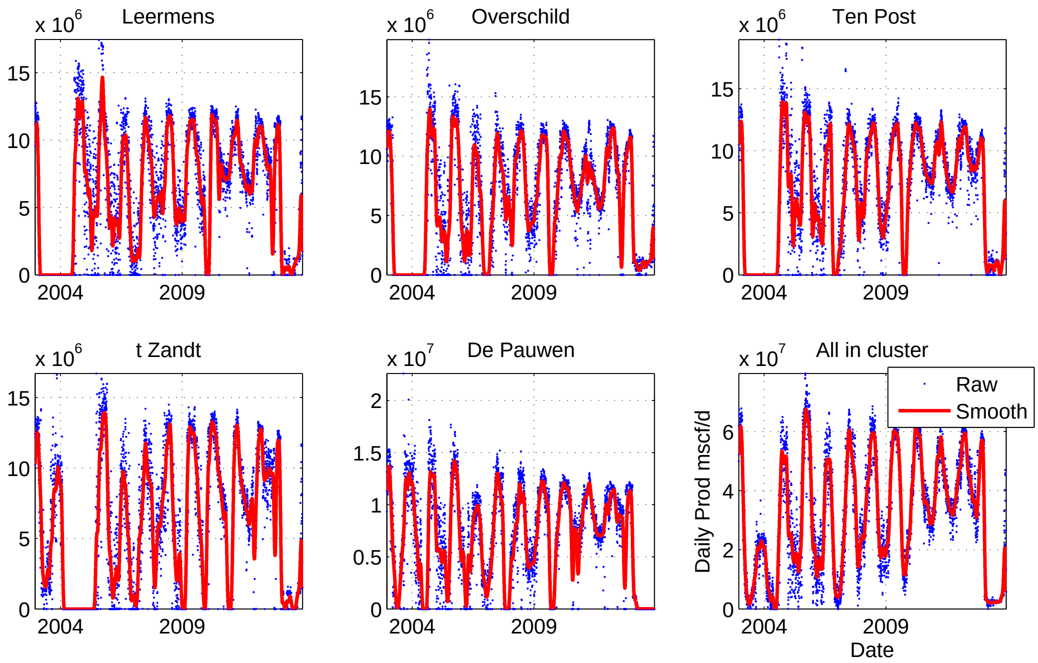

Figure 1: Raw and smoothed production by clusters belonging to Loppersum and their sum (bottom right) for the period 2003 to Jan 2015.

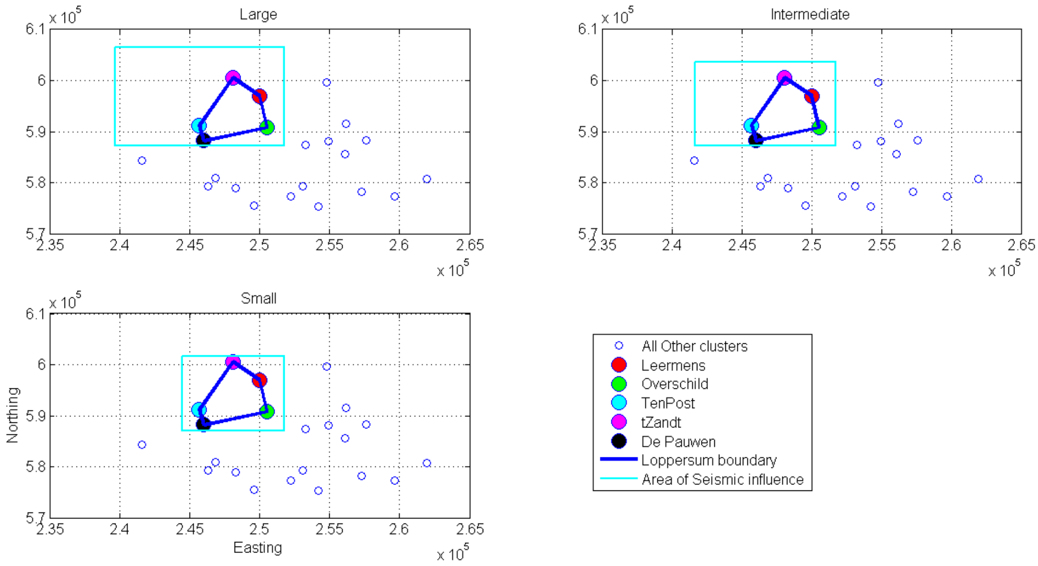

Figure 2: Clusters belonging to Loppersum coloured circles. Production changes to these clusters is assumed to influence seismic activity in the area covered by cyan rectangles of different sizes. Three scenarios are considered. Top left: Large influence area Top Right : Moderate influence area. Bottom left: Narrow influence area. Notice influence areas are not overlapping with other clusters.

2.3 Treating for aftershocks

A seismic event can also trigger another seismic event. This is often called as an aftershock. It is important to understand that an aftershock is not a true event but just a consequence of another event. An aftershock event has to be properly identified and removed from our analysis as it can give a biased view of the the underlying inter-event time. There are various methods suggested in the literature to identify an aftershock but for the purpose of our work, we have treated any event which occurs within three days of the preceding event as an aftershock. Thus, by this methodology, the minimum inter-event time in any block has to be 3 days.

3 Events of Magnitude >=1.5

We now try to understand if the change in production has an influence on seismicity using inter-event time (λ) as a metric described in the preceding section on methodology under three possible. In other words, does inter-event time λ before and after production change when the area influenced by the production change is assumed to be broad, moderate and narrow?

3.1 Broad influence area

When the seismic area influenced by the change in production in Loppersum clusters is assumed to be broad, the spatial distribution of seismic events is shown in figure 3. There are ($n =$) 8 events of magnitude $>= 1.5$ after the production was changed in Jan 2014. The events that occurred prior to Dec 2013 can be arranged into ($k =$) 11 blocks of 8 events each such that the last three blocks (blocks 9, 10 and 11) are closest to 2014. Figure 4 shows the temporal distribution along with the sum of production for the five clusters while figure 5 shows inter-event time, which is also shown in a tabular form in table 1. Posterior λ's, given the information, can now be estimated for all the blocks using Markov Chain Monte Carlo (MCMC) technique and their smooth kernel density is shown in figure 6. λ density for blocks 9, 10 and 11 corresponds to inter-event time before the production change was effected while block 12 corresponds to λ for the post production change period. Figure 7 shows that all density curves of λ overlap suggesting that null hypothesis cannot be rejected and the available data suggest that change in production has not impacted the activity rate. Figure 8 shows the 95 percent confidence interval of λ for all the blocks. We can potentially argue that the trend for median λ is declining with time (activity rate is potentially increasing). However, the spread is large and the median for the last block (refers to period after production change) is not any different to the median from all the other blocks. Effect of production change, if any, is not evident in this limited time frame.

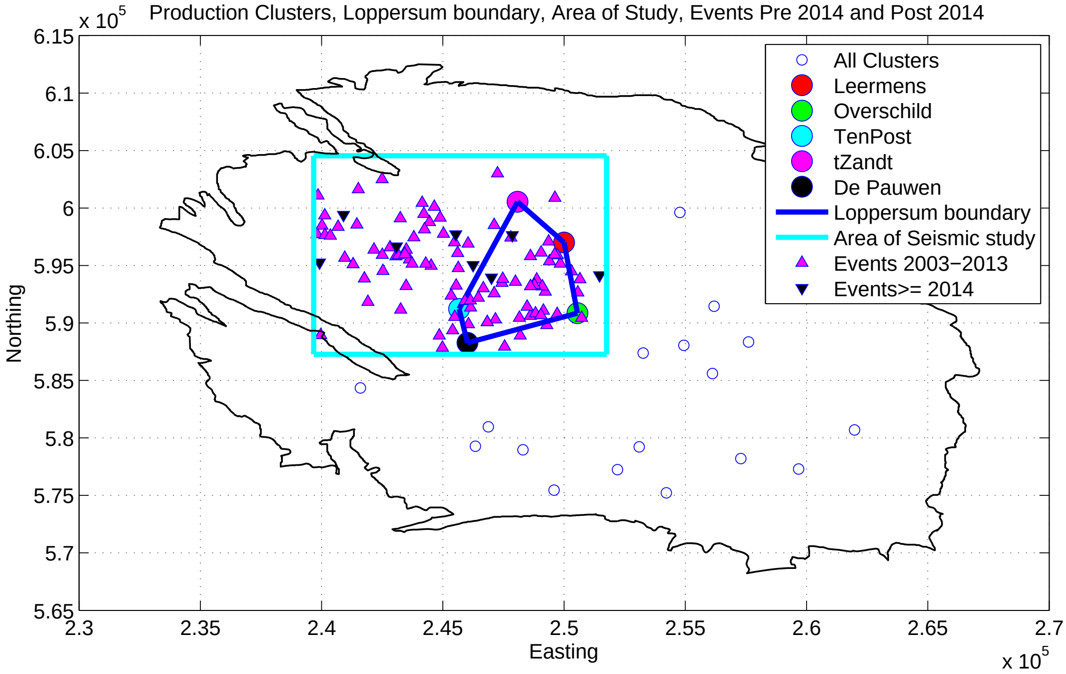

Figure 3: Clusters with reduced production. Seismic events before and after production cut. Broad influence area.

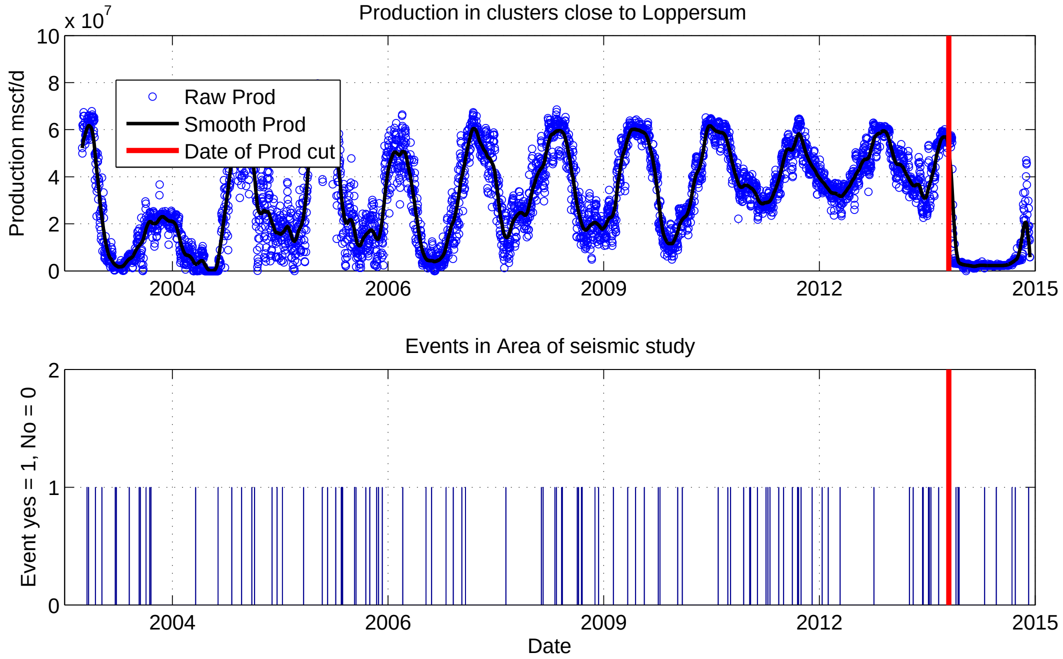

Figure 4: Production(top) and Activity(bottom) with broad influence area.

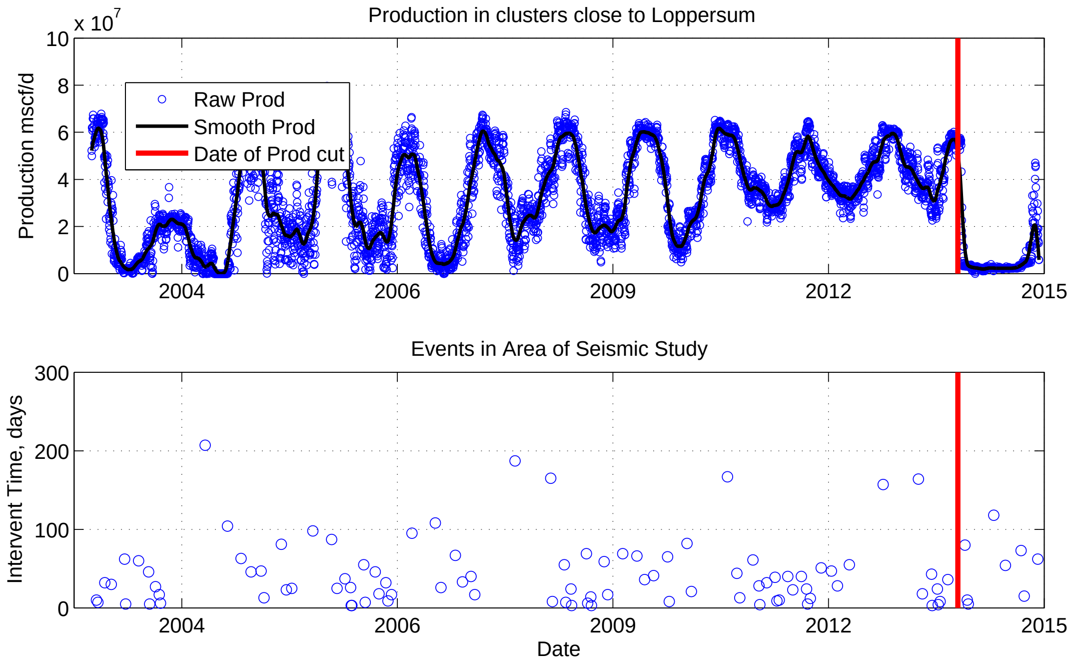

Figure 5: Production(top) and earthquake inter-event time (bottom) before and after production cut with broad influence area.

Table 1: Table of earthquake inter-event time: Broad influence area. Block 12 refers to inter-event time after production was changed. Blocks 1 to 11 refer to inter-event time before production was changed.

Evnt No	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Block 11	Block 12
1	30	6	23	7	26	55	69	21	39	5	4	80
2	62	207	25	46	67	7	66	167	9	12	137	10
3	5	104	98	18	33	23	19	44	10	50	18	5
4	60	63	87	32	40	69	17	13	14	47	43	118
5	46	44	25	9	17	6	41	61	26	28	24	54
6	5	47	37	17	187	14	65	28	23	55	4	73
7	27	13	24	95	164	58	8	4	40	156	8	15
8	17	81	55	108	8	17	80	32	24	19	36	62

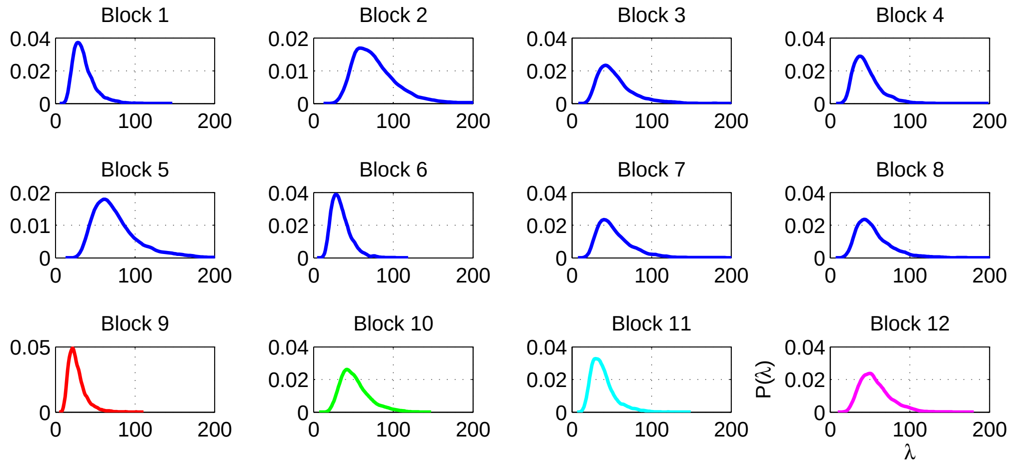

Figure 6: Kernel density for λ for each block. Magenta (bottom right curve) is for earthquake events after production was changed. The three density curve in Red, Green and Cyan are for three periods prior to production change and are in period 2012-2013.

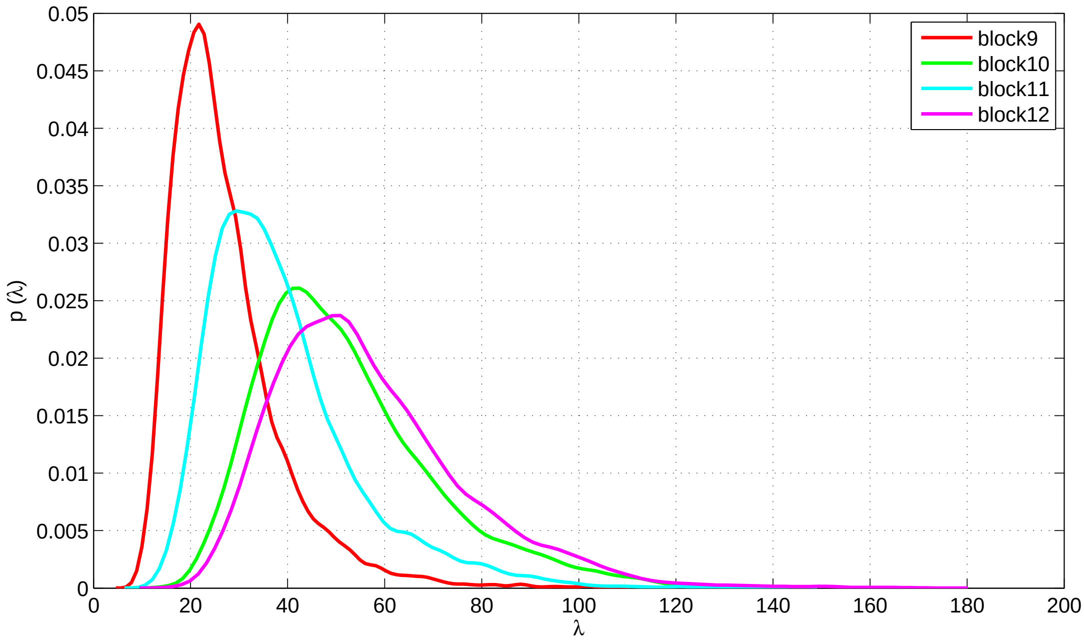

Figure 7: Kernel density for last three blocks prior to production change in red, cyan and green. Density curve in magenta is for period after the production was changed. All density curves overlap.

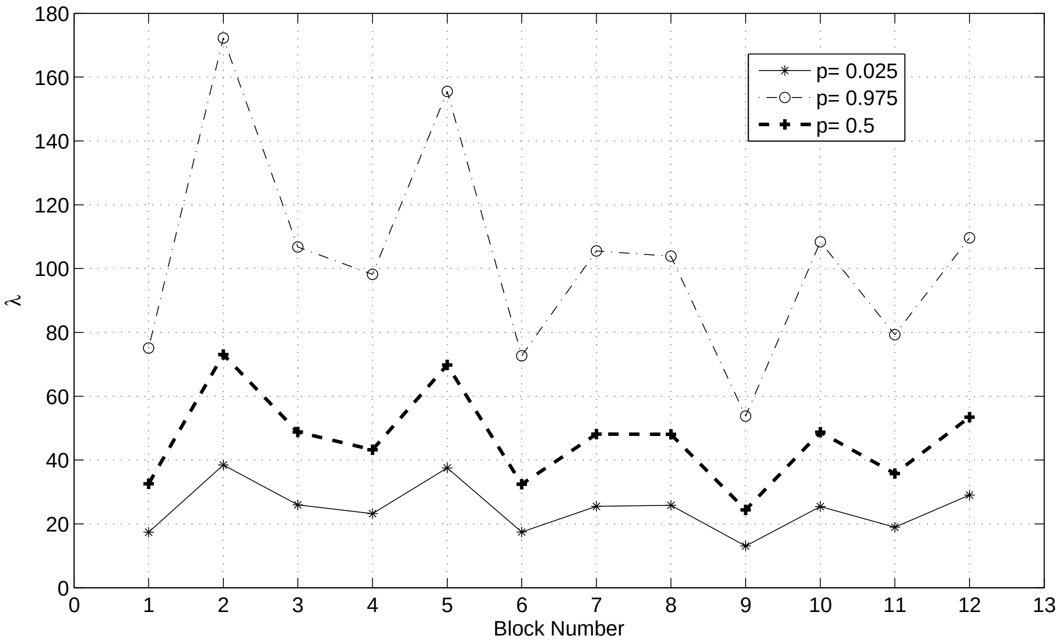

Figure 8: 95 percent confidence interval λ for all blocks.

3.2 Moderate influence area

When the seismic area influenced by the change in production is assumed to be moderate, the spatial distribution of seismic events is shown in figure 9. There are $(n =)$ 6 events after the production was changed. The events that occurred prior to Dec 2013 can be arranged in$(k =)$ 13 blocks of 6 each in a similar way as discussed in the previous chapter. Figure 10 shows the temporal distribution, figure 11 shows inter-event time which is also tabulated in table 2. Posterior λ's are in figures 12 and 13 which shows that the densities of λ before and after production change overlap. This again suggests that the null hypothesis cannot be rejected and the available data suggest that change in production has not impacted the activity rate. Figure 14 shows the trend in median λ and 95 percent confidence interval. Again, it is difficult to claim that median λ had changed significantly after production change.

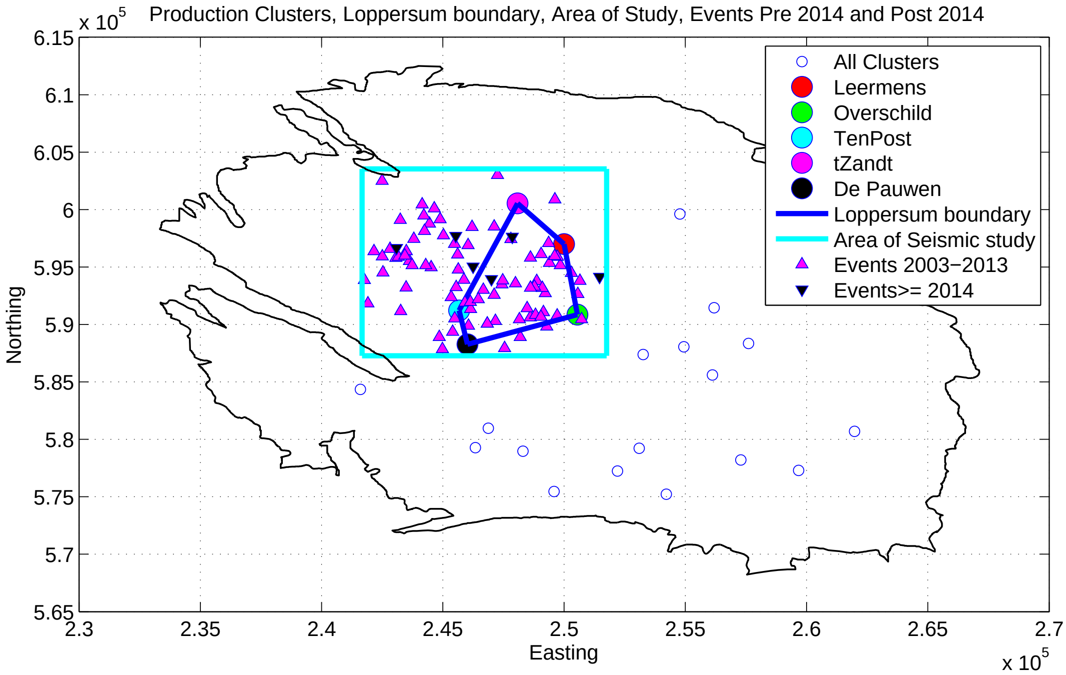

Figure 9: Clusters with reduced production, Seismic events before and after production cut. Intermediate influence area.

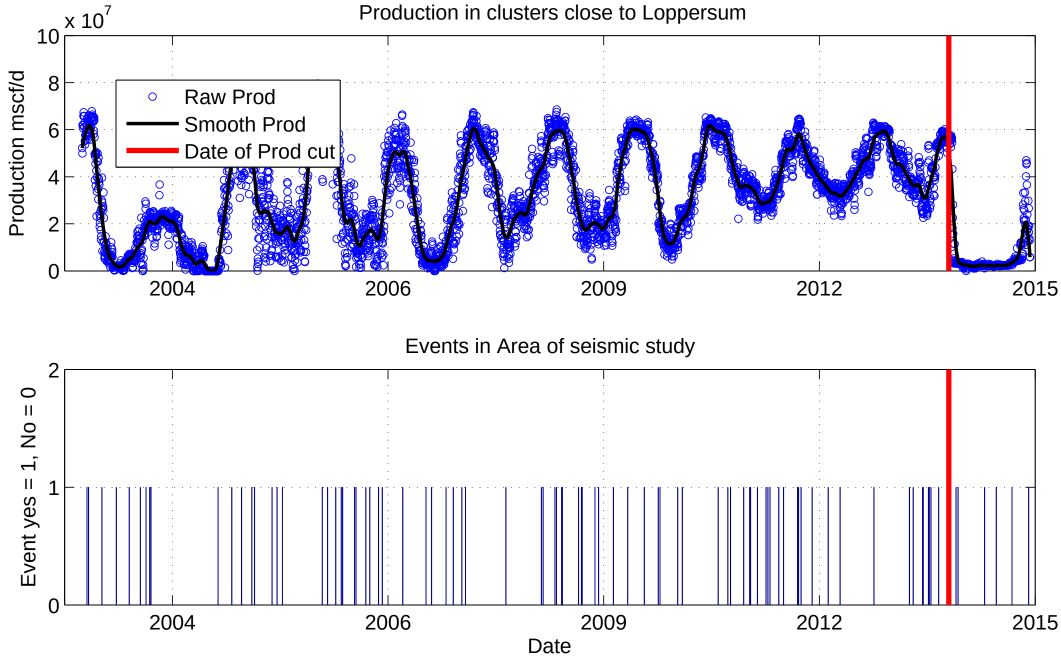

Figure 10: Production(top) and Activity(bottom) with intermediate influence area.

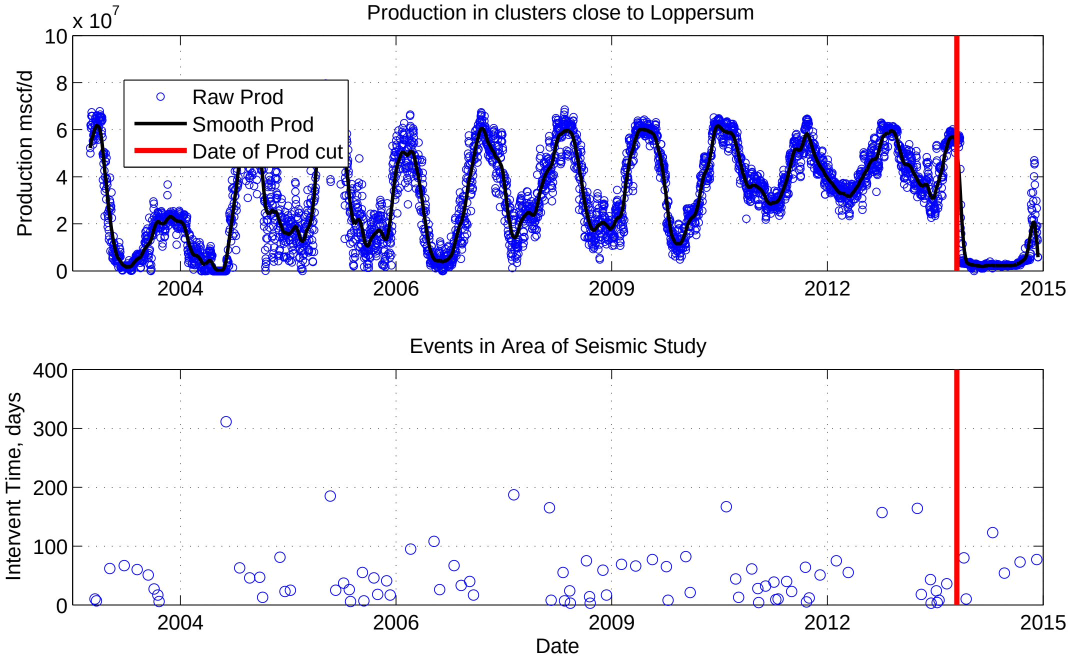

Figure 11: Production(top) and earth quake inter-event time (bottom) before and after production change with narrow influence area.

Table 2: Table of earthquake inter-event time: Intermediate Influence Area. Block 14 refers to period after production was changed. Block 1 to 3 refer to period before production was changed.

Evnt No	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Block 11	Block 12	Block 13	Block 14
1	67	311	23	6	17	40	7	59	8	61	10	51	43	80
2	60	63	25	55	95	17	24	17	82	28	40	75	3	10
3	51	46	185	7	108	187	3	69	21	4	23	55	24	123
4	27	47	25	46	26	165	75	66	167	32	64	157	4	54
5	17	13	37	18	67	8	14	77	44	39	5	164	8	73
6	6	81	26	41	33	55	3	65	13	9	12	18	36	77

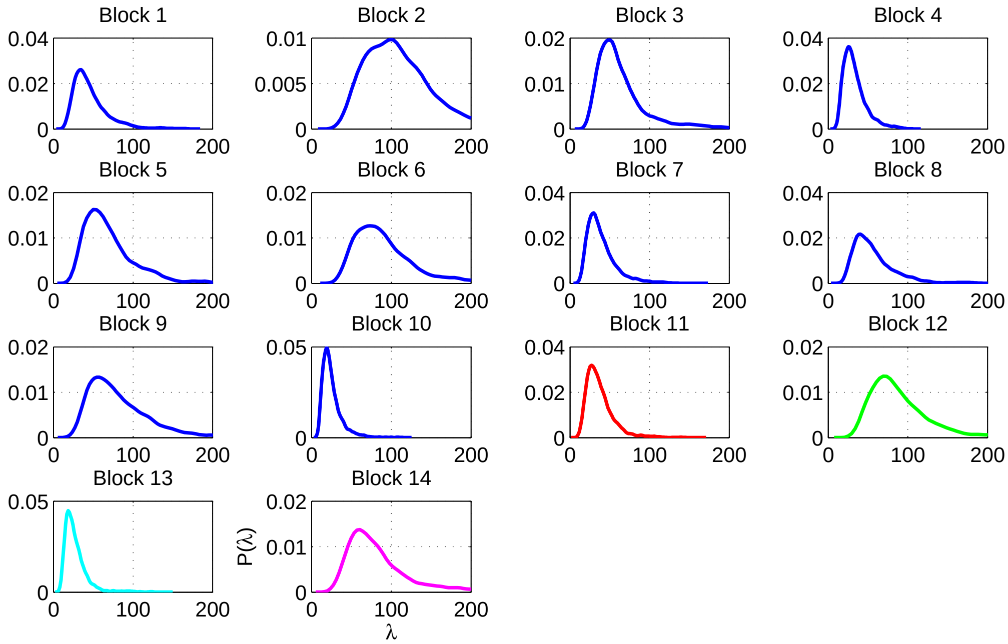

Figure 12: Kernel density for λ for each block. Magenta (bottom right curve) is for earthquake events after production was changed. The three density curve in Red, Green and Cyan are for three periods prior to production change and are in period 2012-2013.

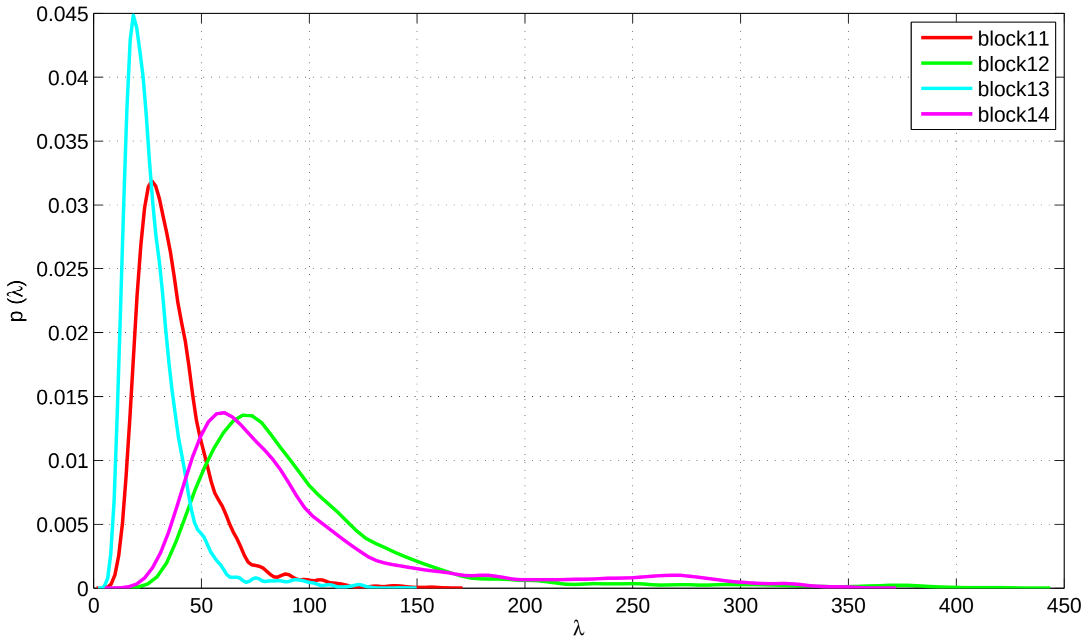

Figure 13: Kernel density for last three blocks prior to production change in red, cyan and green. Density curve in magenta is for period after the production was changed. All density curves overlap.

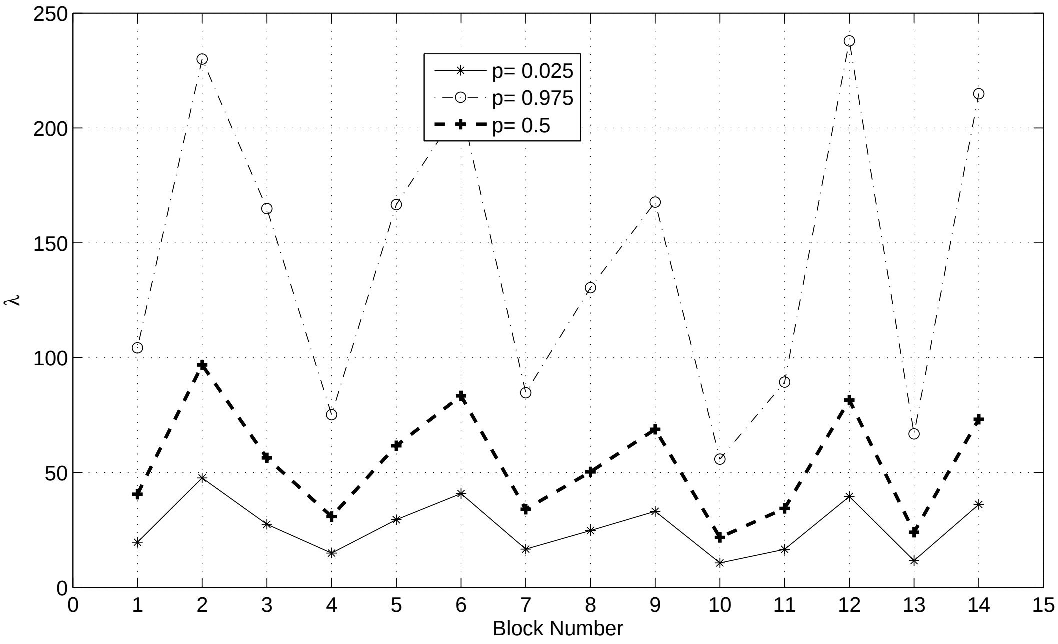

Figure 14: 95 percent confidence interval λ for all blocks.

4 Narrow influence area

When the seismic area influenced by the change in production is assumed to be restricted to the area of the polygon and its vicinity, there are only 5 seismic events in the period after the production change. Figure 15 shows spatial distribution of events, figure 16 shows temporal distribution, figure 17 shows distribution of inter-event time while figures 18 and 19 show posterior densities for λ. Table 3 shows inter-event time. We again see overlap between density curves in cyan (post production change) and pre-production change period suggesting the available data cannot be used to suggest that changes to production have influenced the seismic activity.

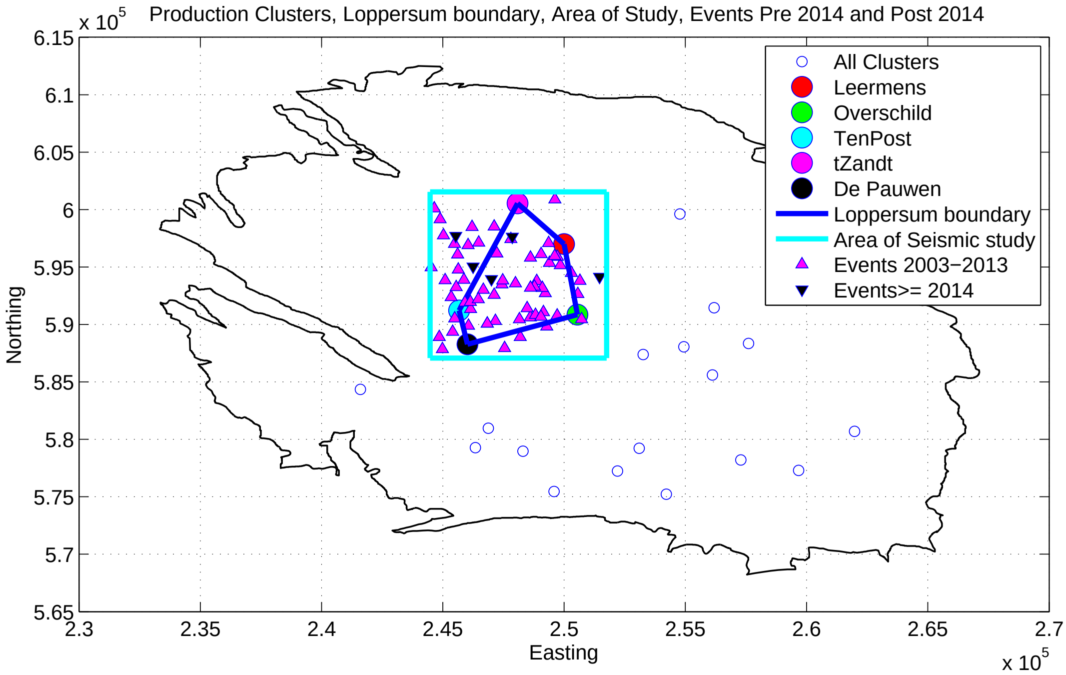

Figure 15: Clusters with reduced production, Seismic events before and after production change.

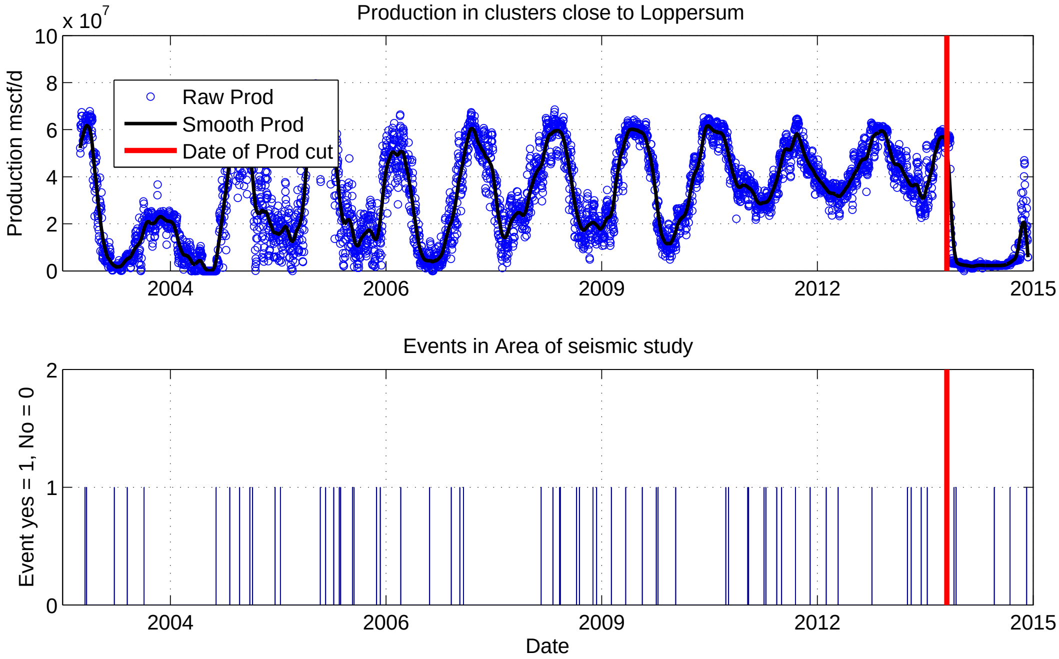

Figure 16: Production(top) and Activity(bottom), narrow influence area.

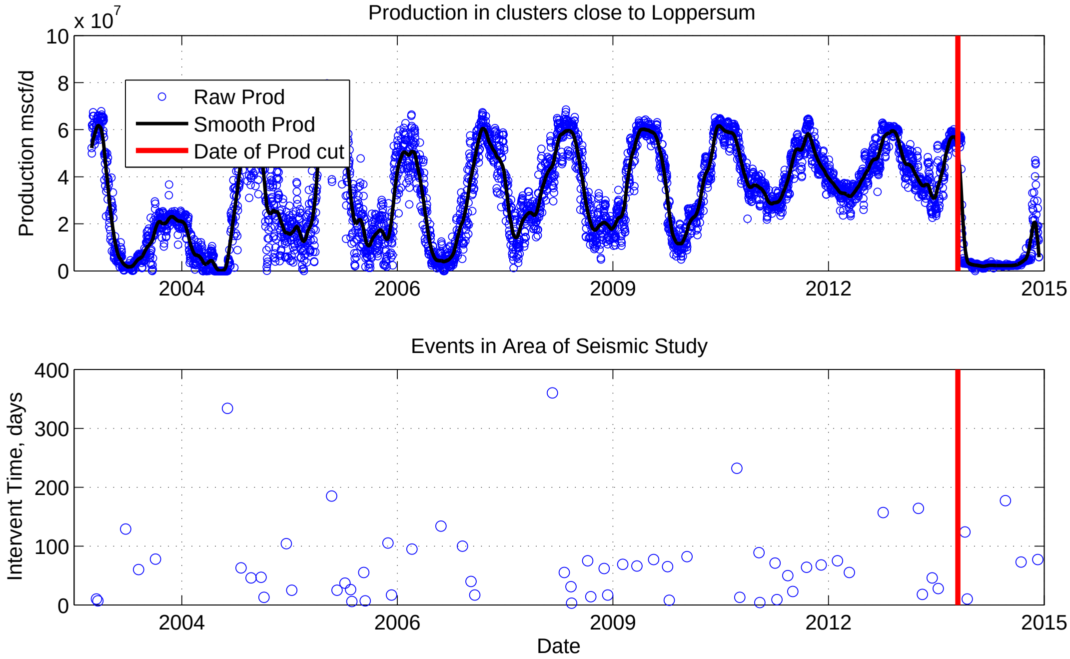

Figure 17: Production(top) and earthquake inter-event time (bottom) before and after production change, narrow influence area.

Table 3: Table of earthquake inter-event time: Narrow Influence Area. Block 13 refers to period after production was changed. Blocks 1-12 refer to period before production was changed.

Evnt No	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10	Block 11	Block 12	Block 13
1	7	63	25	6	95	188	14	66	82	71	68	4	124
2	129	44	185	55	134	172	4	19	232	9	75	137	10
3	60	47	25	7	100	55	58	58	13	50	55	18	177
4	78	13	37	105	40	31	17	65	89	23	157	46	73
5	334	104	24	17	17	75	69	8	4	64	21	28	77

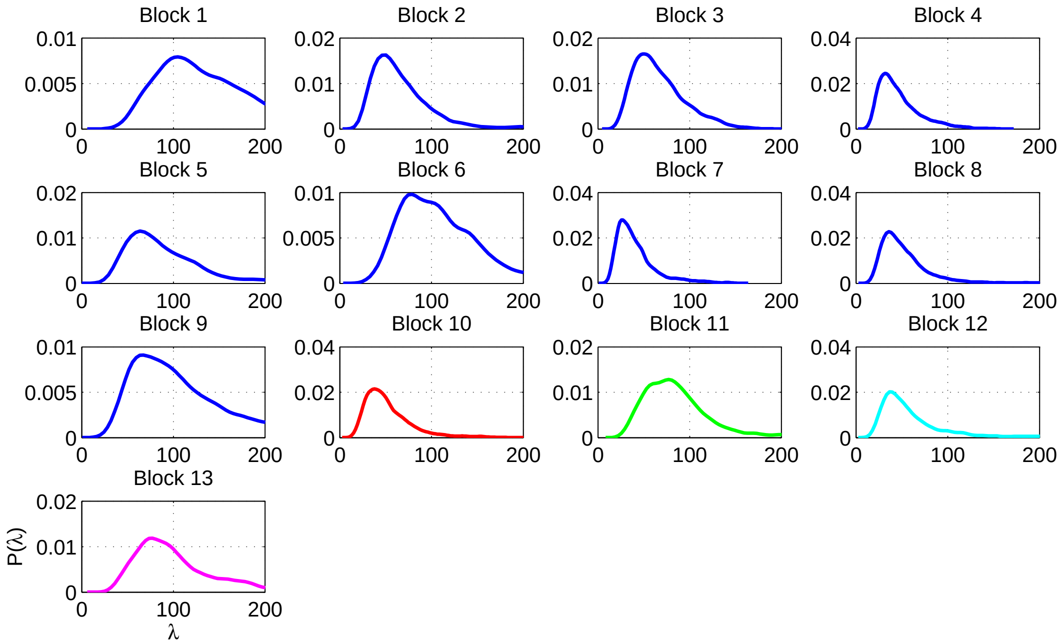

Figure 18: Kernel density for λ for each block. Magenta (bottom right curve) is for events after production was changed. The three density curve in Red, Green and Cyan are for three periods prior to production change and are in period 2012-2013.

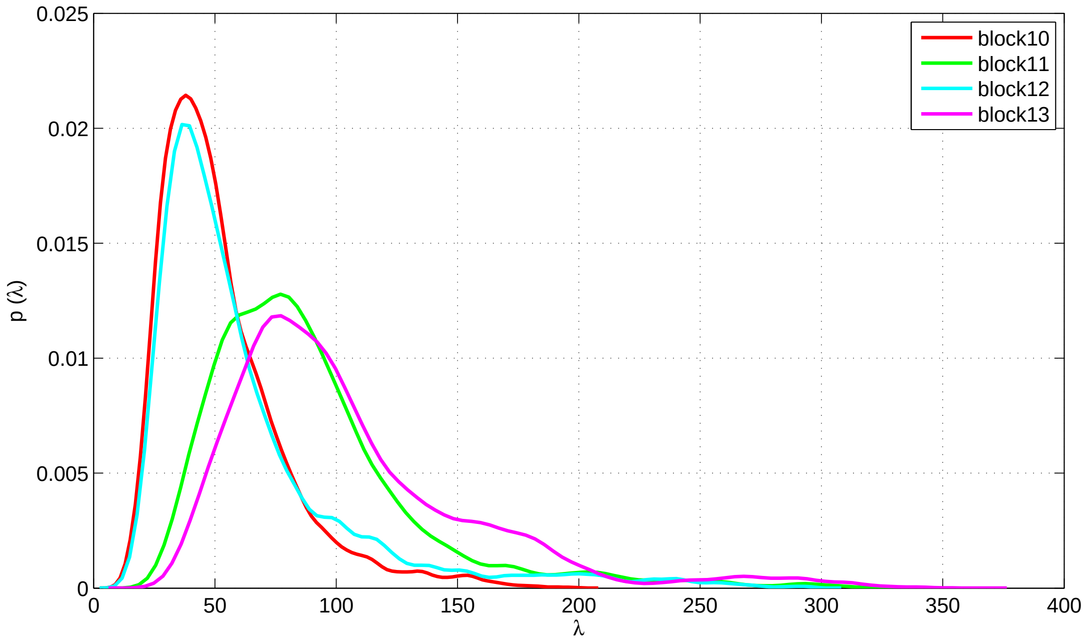

Figure 19: Kernel density for last three blocks prior to production change in red, cyan and green. Density curve in magenta is for period after the production was changed. All density curves overlap.

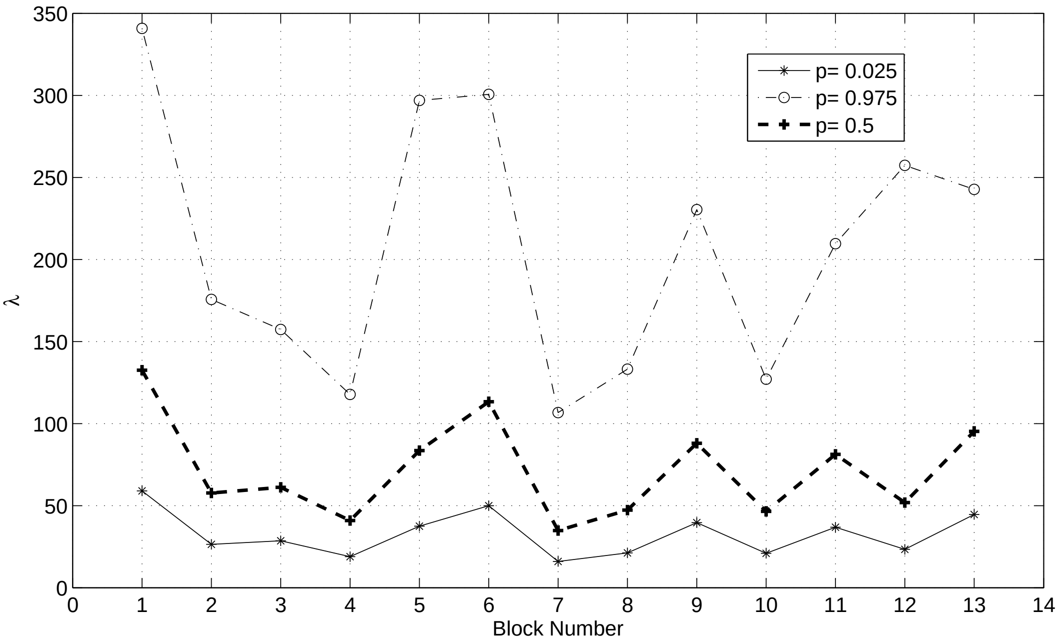

Figure 20: 95 percent confidence interval λ for all blocks.

5 All events

In the previous section, we used events of magnitude >=1.5 to understand the relation between production and seismicity. In this section, we do a similar analysis but include all the events in the catalog that exists on the KNMI website. It should be mentioned that though this analysis is statistically robust, the results cannot be treated as reliable since events of magnitude < 1.5 cannot be recorded accurately. The inter-event time should therefore be read as **Catalogue inter-event time throughout the text**

5.1 Broad Influence Area

Figure 21 shows the spatial distribution of all events that occurred prior to production change and also after production change in the area which we believe can be influenced by production. The number of events is much greater than when only events of magnitude 1.5 and above are considered. After the production was changed, 22 events of small and large magnitude occurred in the defined area. The temporal distribution and inter-event time is shown in figures 22 and 23 respectively. We can observe that the events get denser near 2011 and continue to remain so immediately after the production change was made . The inter-event time for block 11 which falls in post production change period is trending upwards as seen in Figure 24 which implies that production change has perhaps reduced the number of seismic events of smaller magnitude after production change. However, we must wait for a longer period of time in order to confirm this trend.

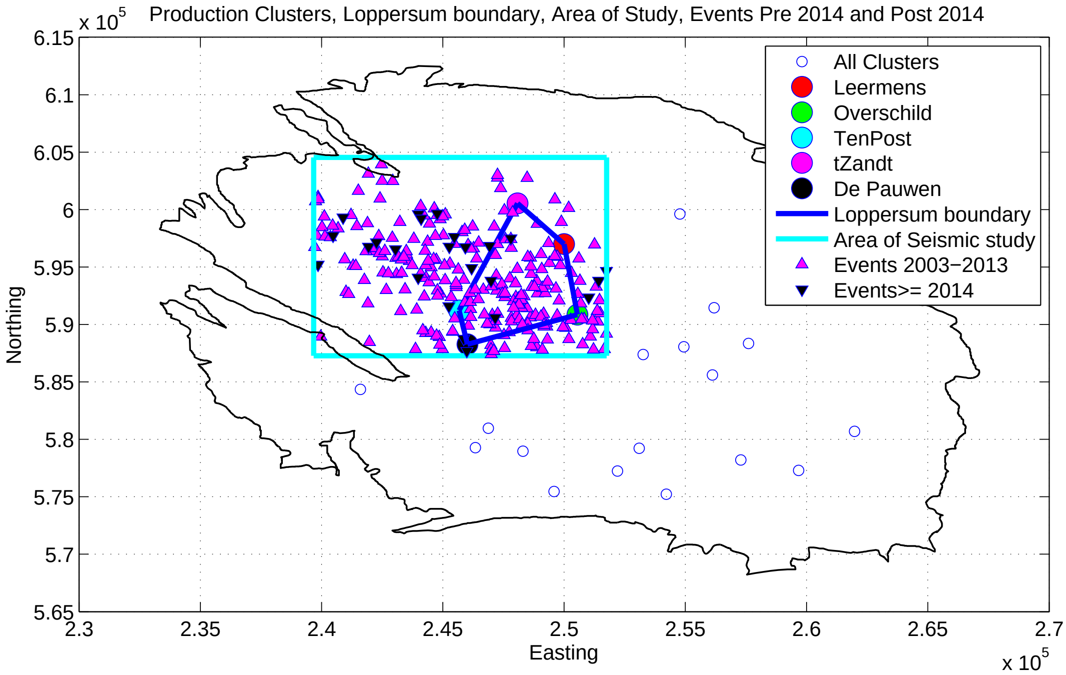

Figure 21: Clusters with reduced production, catalogue events before and after production change.

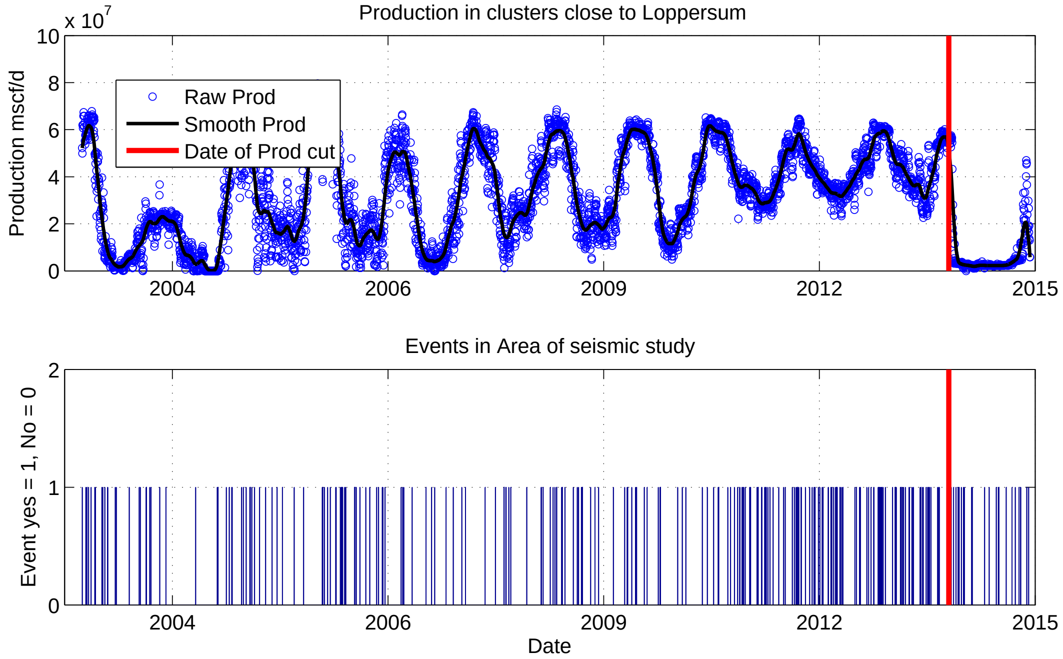

Figure 22: Production and catalogue events (Yes = 1, No = 0).

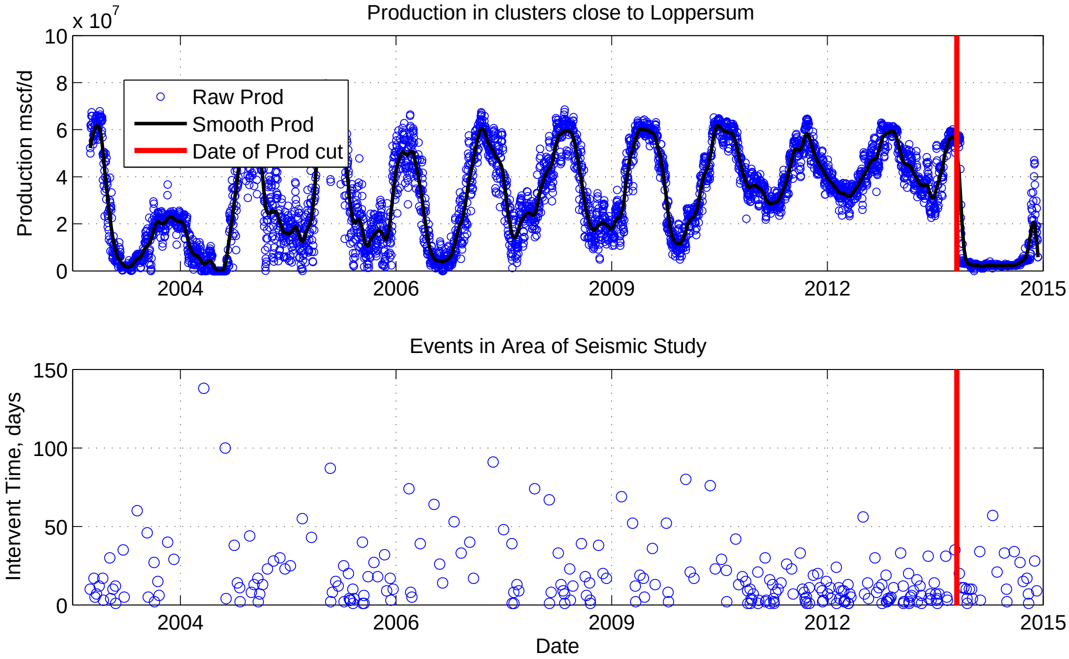

Figure 23: Production and catalogue Inter event time.

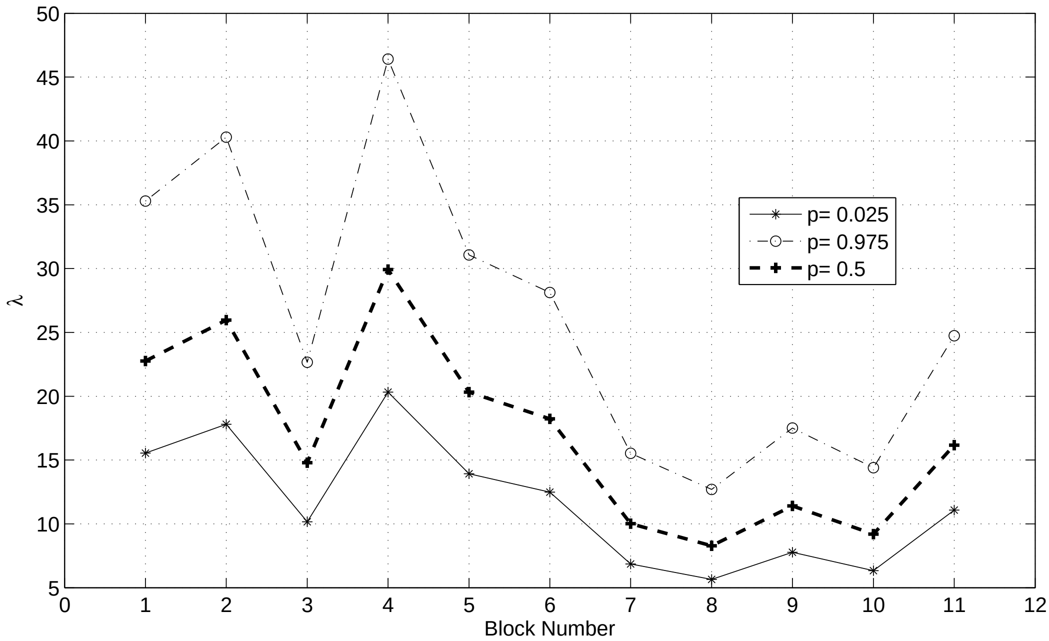

Figure 24: 95 percent confidence interval of λ.

5.2 Moderate Influence Area

Figure 25 shows the spatial distribution of the seismic events that have occurred in the area of influence before and after production change was made. Figure 26 shows the temporal distribution of events. We again observe that the events occurred more frequently from 2010 to end of 2013, the date of production change. After the production change, they were relatively less frequent as seen in figure 27. Figure 28 shows that there is a good evidence to believe that the inter-event time has increased after production change but again, we must wait for a longer period before confirming this observation.

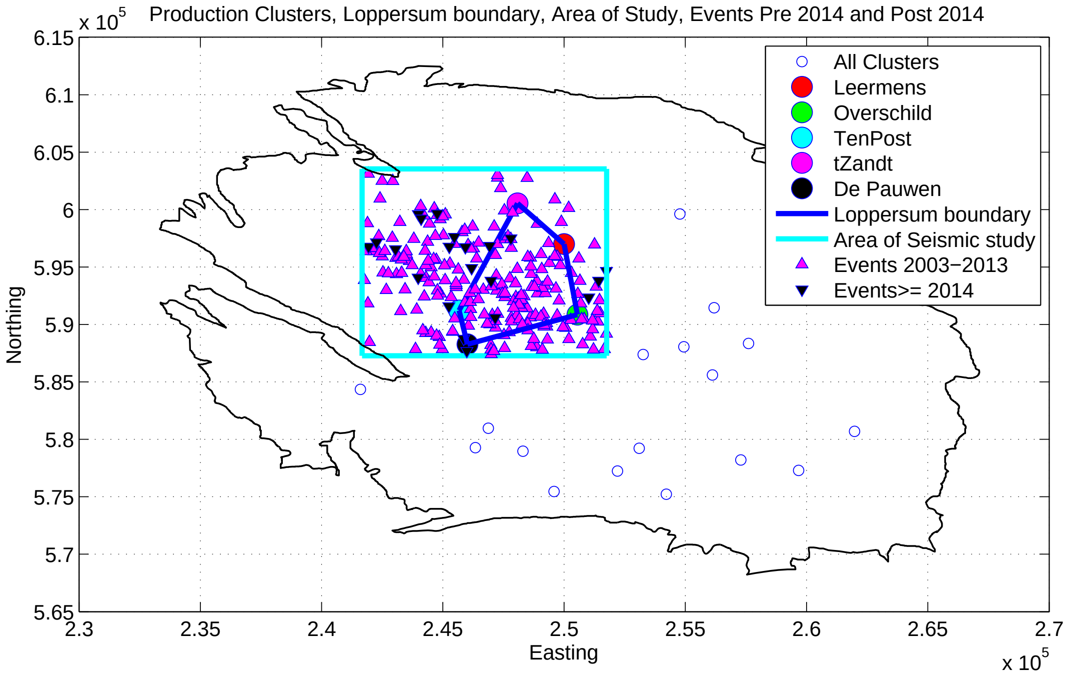

Figure 25: Clusters with reduced production, catalogue events before and after production change.

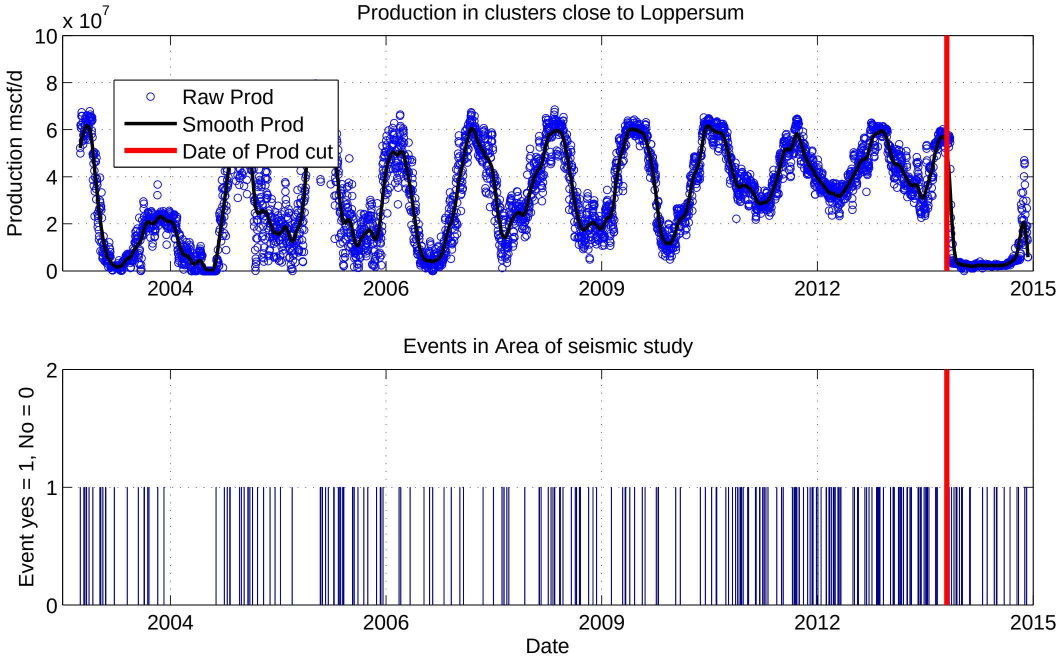

Figure 26: Production and catalogue events (Yes = 1, No = 0).

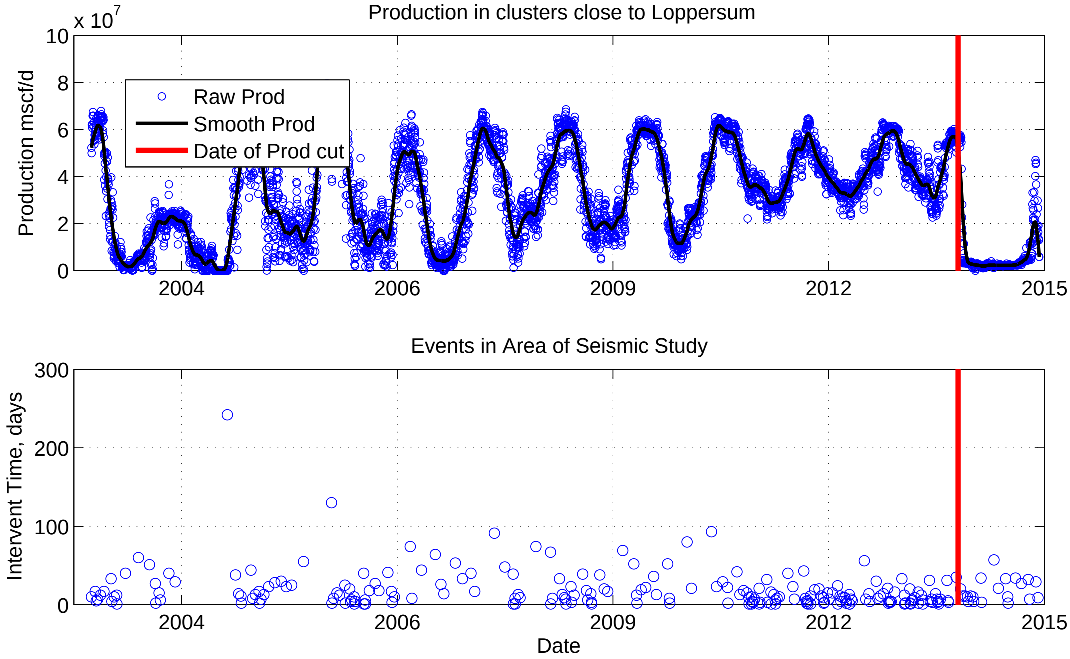

Figure 27: Production and catalogue inter-event time.

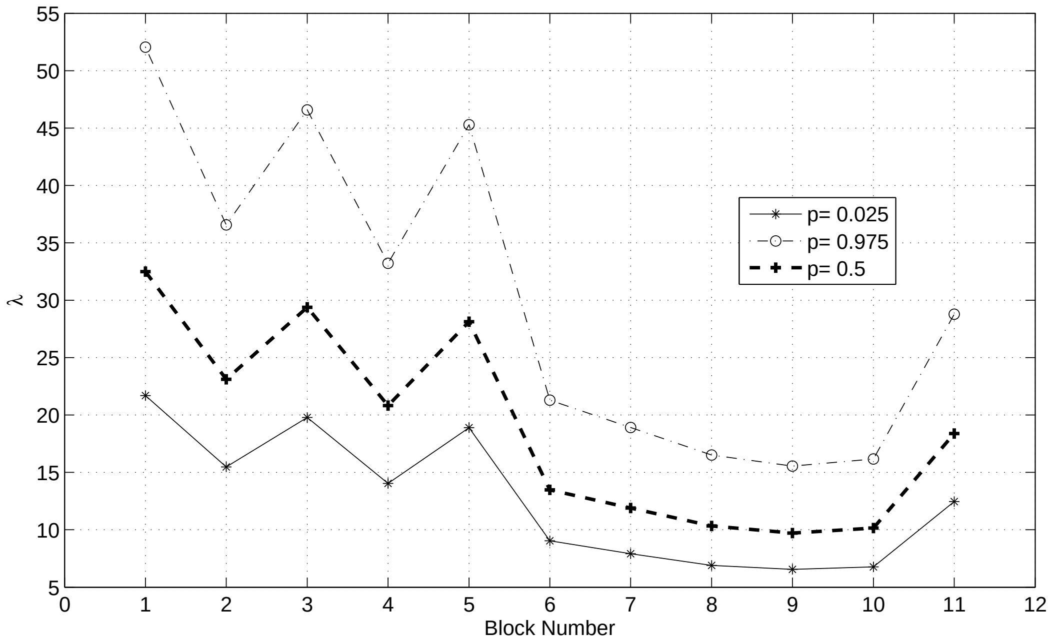

Figure 28: 95 percent confidence interval of λ.

5.3 Narrow Influence Area

Figure 29 shows the spatial distribution of all the events in what we believe to be a narrow influence area after production change was made. We again see that the temporal distribution of events became dense since 2011 (Figure 30) in this area and continued to remain so just after production change was made. Later during the year, events seem to get less dense which is also supported by Figure 31 and Figure 32.

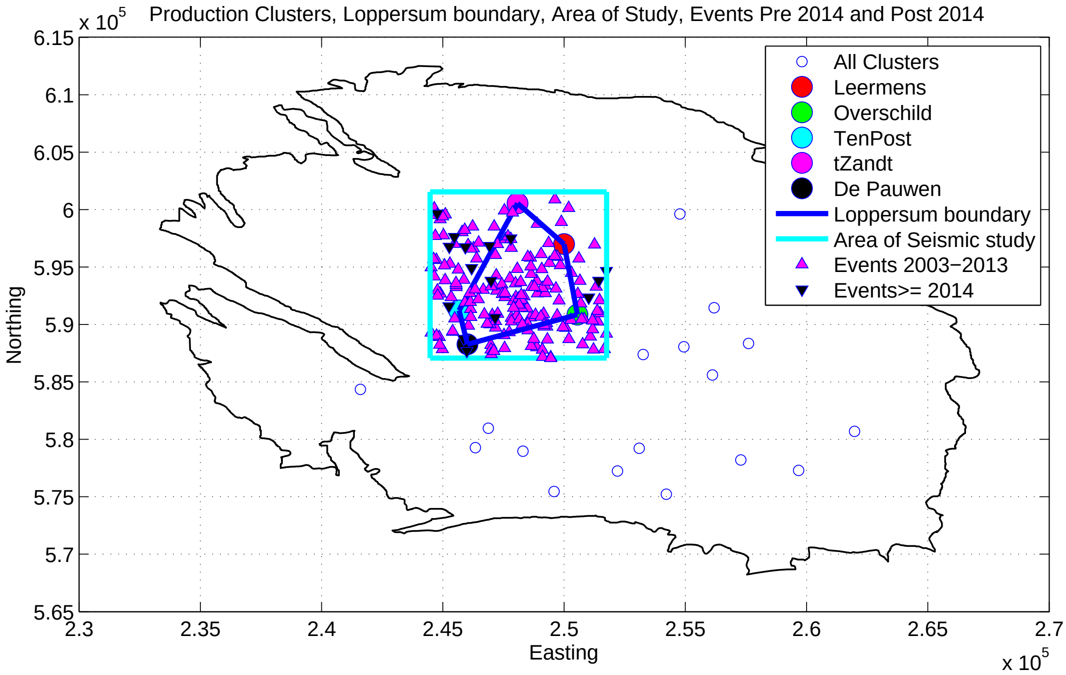

Figure 29: Clusters with reduced production, catalogue events before and after production change.

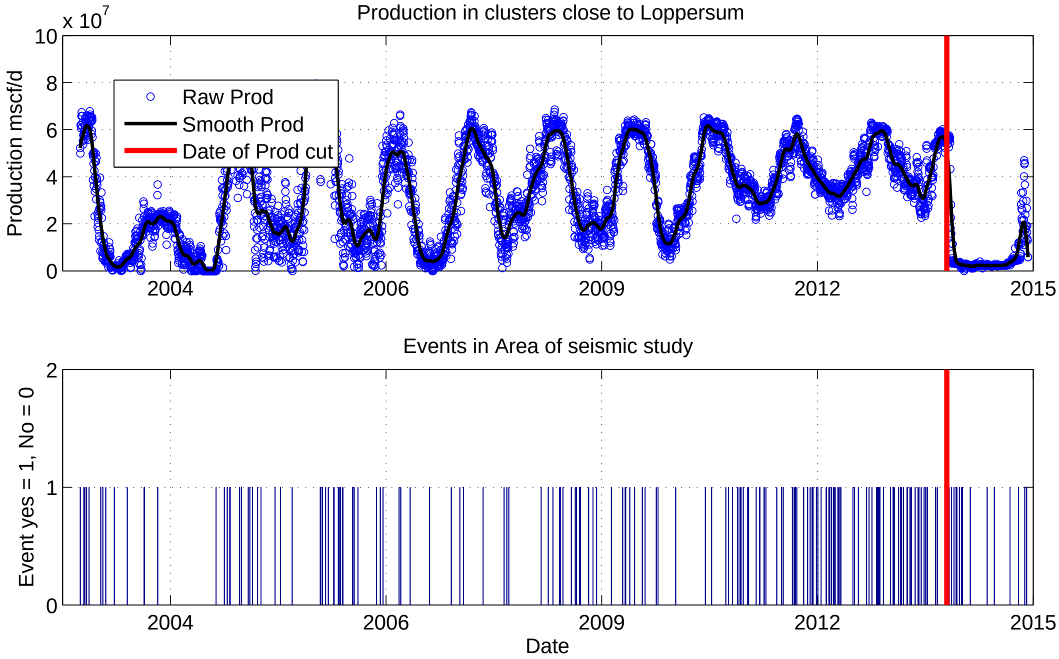

Figure 30: Production and catalogue events (Yes = 1, No = 0).

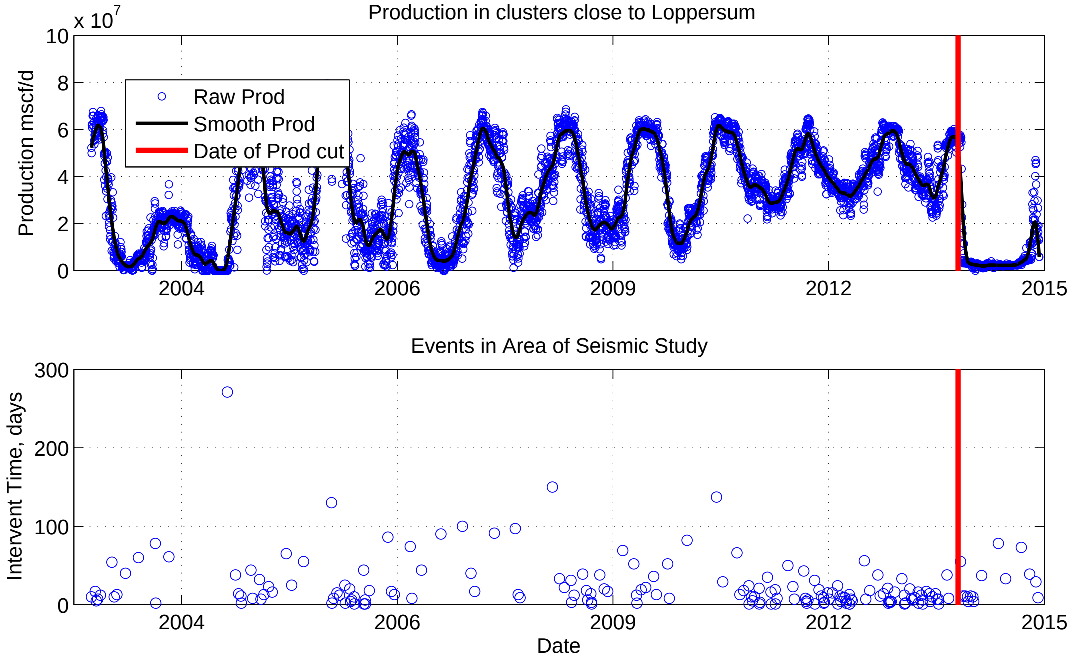

Figure 31: Production and catalogue inter-event time.

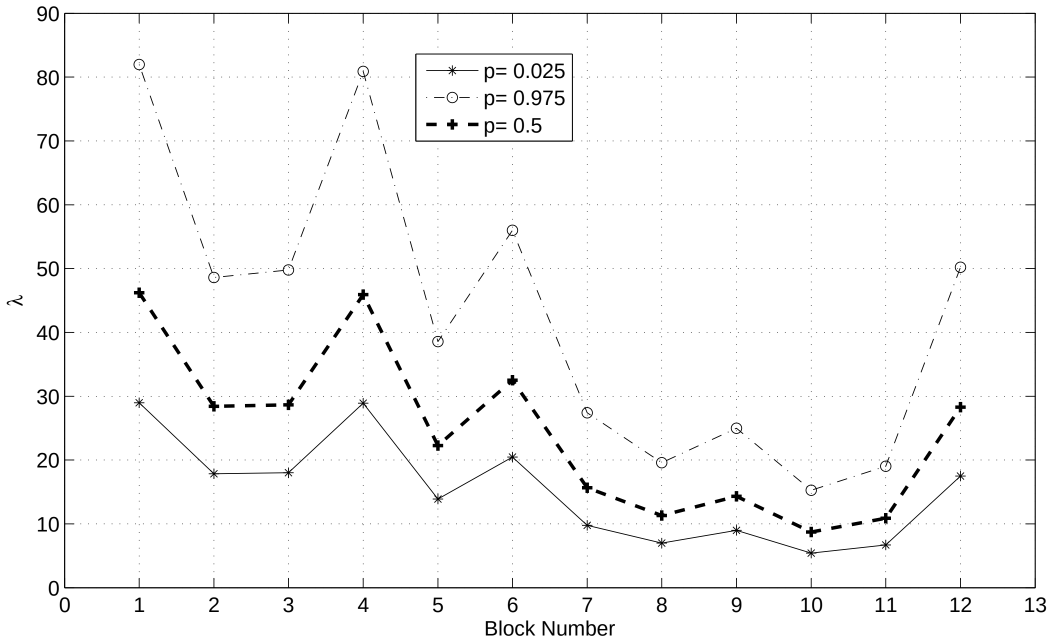

Figure 32: 95 percent confidence interval of λ.

6 Remarks and Notes

The current study shows that the inter-event time of seismic events of magnitude $>= 1.5$ after the production change was made has not changed but it appears that if all the events are considered, the seismic activity is reduced under all possible scenarios of seismic area influence influenced by production. We should, however, caution that though the statistical approach and results are robust, the study is covering a limited time period (January 14 to Jan 15) with reduced production. The sample size, in terms of number of events in this period is small which can influence the outcome. Also, it is uncertain if events of magnitude < 1.5 can be recorded accurately and there can be spatial and temporal uncertainty in the catalogue provided by KNMI website. The inter-event time when all events are considered should therefore be read as **Catalogue inter-event time**.

Production for the clusters belonging to Loppersum (and for other clusters as well) is highly cyclical. Each year, the peak in production occurs in one of the winter months and trough in summer. Despite the troughs in summer, it is not the case that summer months are devoid of any activity. Also, while Dec and Jan months have peak production, it should be kept in mind that December is the month of lowest seismic activity as only 5 seismic events have occurred in this month over 2003 to 2013 period. On the other hand, January shows high seismic activity and at-least one event has occurred in this month for all the years from 2003 to 2013. This strange observation shows that other covariates such as seasonality in seismicity besides production are (also) involved and should be given due importance.

In the present work, we have assumed that the posterior λ is an Exponential variate with non-informative prior that is inversely proportional to λ. For exponential distribution, an informative prior can be based on

Gamma distribution with two hyper-parameters. If we have prior belief about the hyper-parameters, the analysis can be re-run although we believe there will be a minimum impact on the analysis and conclusions.

Inter-event time in just one metric by which we can understand the influence of production change on seismicity. Other metrics are total seismic moment, frequency of large magnitude events and correlation between production and events. These need to be considered to have a proper view on the relationship between production changes and seismicity.

It is recommended that a further study of inter-event time of a spatial subset of events of magnitude $>= 1.0$ should be conducted to understand the influence of production. This spatial subset should come from locations within the Loppersum boundary where the geophones are present. This will overcome the uncertainty in our analysis for events of magnitude < 1.5 as we do not know the exact location of the geophones used to record seismic events.