An Adapted Båth's Law Framework for Assessing Trailing Seismicity in Dutch Gas Fields

Bubu Schmal
To obtain the degree Bachelor of Science
At Delft University of Technology
To be defended on 9th of July 2026

Student Number: 6060919
Project Duration: 18th of May - 6th of July
Thesis committee: Annemarie Muntendam-Bos (Supervisor 1)
David Naranjo Hernandez (Supervisor 2)
Anne Pluymakers (external committee member)

TU Delft

Table of Contents

Table of Contents 2
1. Abstract 3
2. Introduction 4
3. Mechanisms of induced and trailing seismicity 5
4. Methodology 6
5. Code Description 11
- 5.1 TrailDistance 11
- 5.2 dMtrail 11

6. Selection of Gas Fields 12
7. Parameters Input 14
- 7.1 Groningen 14
- 7.2 Annerveen 16
- 7.3 Eleveld 16
- 7.4 Roswinkel 16

8. Results 17
- 8.1 Beta Distribution fit 17
- 8.2 Groningen 18
  - 8.2.1 Present 18
  - 8.2.2 Adjusted Groningen 20
- 8.3 Annerveen 22
- 8.4 Eleveld 23
- 8.5 Roswinkel 25

9. Discussion 27
- 9.1 Gas Field Robustness 27
- 9.2 Justification; Groningen 27
- 9.3 Justification: Small Dutch Gasfields 28
- 9.4 Limitations 29
- 9.5 Future 30

10. Conclusion 31
11. Reference List 32
Appendix 34
- Appendix A 34
- Appendix B 36
- Appendix C 37

1. Abstract

Induced seismicity related to gas production remains an important scientific and societal issue in the Netherlands, especially because earthquakes may continue after production has stopped. This post-production activity, known as trailing seismicity, creates uncertainty in seismic hazard assessment, since the largest earthquake does not necessarily occur during the production phase. This thesis investigates whether the largest observed trailing earthquakes in Dutch gas fields, with specific focus on the 2025 Zeerijp earthquake in Groningen, can be justified using an adapted framework of Ryan Schultz combined with Gutenberg-Richter magnitude-frequency statistics.

Earthquake data from the Koninklijk Nederlands Meteorologisch Instituut (KNMI) catalogue were divided into production and post-production events for selected Dutch gas fields. The main parameters used in the analysis were the magnitude of completeness, the number of production and trailing earthquakes, the production-seismicity ratio, the Gutenberg-Richter b-value, and the maximum magnitude parameter. Groningen was used as a control case because of its larger earthquake catalogue, while Annerveen, Eleveld and Roswinkel were analysed as smaller gas field examples. A Monte Carlo approach and logic-tree structure were used to include uncertainty in the input parameters and to estimate distributions of the largest expected trailing magnitude.

The results show that the observed Zeerijp earthquake with magnitude 3.4 falls within the predicted Groningen magnitude distribution. For the smaller gas fields, their observed largest trailing magnitudes are justified as well; however, uncertainty increases. Overall, the adapted framework can justify the largest observed trailing events within the selected Dutch gas fields, but the reliability of the prediction depends on the assumed b-value, completeness magnitude, seismicity ratio and maximum-magnitude constraint. Therefore, the method is useful as a probabilistic tool but should be interpreted together with field-specific geological and catalogue uncertainties.

2. Introduction

Induced seismicity has become an important scientific, political, and societal issue in the Netherlands. For several decades, natural gas production contributed to the Dutch energy supply and economy. The Groningen gas field, discovered in 1959 and brought into production in 1963, became one of the most important gas fields in Europe. Revenues from Groningen gas production contributed greatly to Dutch prosperity, with the Dutch state earning more than €363 billion from the field over approximately sixty years (Parliamentary Committee of Inquiry, 2023). However, the same production also caused reservoir pressure depletion, subsidence, and induced earthquakes, which led to damage to structures, safety concerns, and loss of trust among regional residents.

Groningen shows that induced seismicity is not only a geological or technical problem but also a societal problem. Even relatively small earthquakes can have serious consequences when they occur at near-surface depths and close to houses and infrastructure. In the Netherlands, induced earthquakes related to gas production are generally of a low magnitude ($M < 3$), but because they are shallow, they can still cause damage and concern among the local population (Dost et al., 2004). In Groningen, repeated earthquakes resulted in large numbers of damage claims and a long-term reinforcement operation for buildings. According to the parliamentary inquiry into Groningen gas extraction, more than 267,000 damage claims had been reported by early 2023, and many buildings still required assessment or reinforcement (Parliamentary Committee of Inquiry, 2023). Eventually, the risks and damage related to induced earthquakes became too significant to ignore, and the Dutch government therefore decided to end gas extraction from the Groningen field, and the field was permanently closed in April 2024 (Government of the Netherlands, 2024).

A further risk is that induced seismicity does not always stop immediately when the operation stops. Schultz et al. (2022) state that induced earthquakes can continue after injection or stimulation has ended. This is called trailing seismicity. In some cases, the largest earthquake in a sequence may occur after shut-in, which makes operational risk management more difficult (Schultz et al., 2022). A notable trailing event is the Zeerijp event of Groningen, with a magnitude of 3.4, one of the highest magnitudes recorded in Groningen. Trailing seismicity is not only relevant for Groningen but also for small Dutch gas fields (Grigoratos et al., 2023).

This study focuses on estimating and interpreting induced seismicity parameters for Dutch gas fields. By analysing production-period events, trailing events, event ratios, and expected maximum magnitudes, the study contributes to a better understanding of how seismicity behaves after gas production. Furthermore, it aims to see if the Zeerijp event of magnitude 3.4 can be justified and, similarly, the top events of other Dutch gas fields within current observations.

3. Mechanisms of induced and trailing seismicity

Induced seismicity refers to earthquakes caused by anthropogenic activities in the earth's subsurface. These activities disturb the natural stress state of the earth by changing pore or reservoir pressure, temperature or load of rocks and faults. Examples are gas injection, gas storage, geothermal energy production, impoundment of reservoirs, salt solution mining and mining in general (Ellsworth, 2013; Foulger et al., 2018; Muntendam-Bos et al., 2022). In the Netherlands, induced seismicity is relevant because of their involvement in gas production and increase of subsurface use, such as energy storage and geothermal heat extraction (Muntendam-Bos et al., 2022).

Because several related terms are used throughout this thesis, the most important concepts are defined below:

- Induced seismicity: earthquakes that are mainly caused by human activity. In this thesis, this mainly refers to earthquakes related to gas production and the resulting pressure and stress changes in the reservoir.
- Triggered seismicity: earthquakes that occur on faults that may already be close to failure, but where sudden activity changes the stress conditions enough to bring the earthquake forward in time. In this case, the human activity does not necessarily create the fault instability completely, but it can trigger slip on an already critically stressed fault.
- Trailing seismicity: earthquakes that occur after the operation has stopped. In this thesis, this refers to earthquakes that occur after gas production has ended. Trailing seismicity is important because the largest earthquake in a sequence does not always occur during the production phase.

Physically, the cause of induced seismicity is the reactivation of faults. Faults may already exist and could already be close to failure due to natural stresses. Human activity can then alter the existing stress state enough to make the fault slip. For injection, the most important factor is the increase of pore pressure, reducing the normal stress acting, leading to easier slippage of the fault (Ellsworth, 2013; Keranen et al., 2014). A different mechanism occurs for gas extraction: during production reservoir pressure decreases, changing the effective stress in and around the reservoir, which could lead to activation of faults (Muntendam-Bos et al., 2022; Richter et al., 2020).

Induced earthquakes in the Netherlands are generally of low magnitude (roughly $M \leq 3.6$, the largest recorded) but can still be felt because they occur at shallow depths where weak topsoils amplify ground motion. Events of magnitude 1.5-2.0 are already perceptible to the public, and larger events can cause non-structural damage to houses and infrastructure (Muntendam-Bos et al., 2022).

4. Methodology

The aim of the methodology is to test whether the largest observed post-production earthquakes in Dutch gas fields are statistically plausible within a trailing-seismicity framework. To do this, each earthquake sequence must first be defined in a consistent way. This requires three steps: selecting earthquakes that can reasonably be related to gas production, separating the selected events into production and post-production phases and applying a magnitude threshold so that only reliably recorded events are used.

In induced seismicity studies, both the location and timing of an earthquake are important. An earthquake can only be interpreted as induced if it has a plausible spatial and temporal relation to a subsurface operation. Therefore, earthquakes were not automatically assigned to the nearest gas field. Instead, the selection followed the Dutch induced seismicity classification approach described by Muntendam-Bos et al. (2022). In this approach, events are assessed using their spatial relation to gas fields and subsurface operations, while also considering location uncertainty, hypocentre depth, and nearby geological structures. This is consistent with the general induced seismicity reasoning of Davis and Frohlich (1993), where spatial relation, timing, depth, and regional seismic history are used to assess whether earthquakes may be induced.

The earthquake data used in this study were taken from the KNMI earthquake catalogue (Koninklijk Nederlands Meteorologisch Instituut, n.d.). This catalogue provides the origin time, epicentre location, depth, and local magnitude M_L of recorded earthquakes. These variables are needed to determine when each earthquake occurred, where it occurred, and how large it was. After the spatial attribution step, the timing of each selected earthquake was compared with the production history of the associated gas field.

This timing step is important because induced seismicity can occur both during and after the operational period. Earthquakes that occur while gas extraction is still ongoing are considered production-phase events. However, seismicity does not always stop immediately after production ends. Stress changes caused by reservoir pressure depletion and compaction can continue to redistribute after shut-in, and faults may still respond during this post-production period, thus causing trailing seismicity. Schultz et al. (2022) therefore distinguish between an operational phase and a trailing phase when analysing induced seismicity sequences.

In this thesis, the same distinction is applied to Dutch gas fields. The operational phase is defined as the period between the start and end of gas production. The trailing phase is defined as the period after production has stopped. Each selected earthquake is therefore assigned to either the production phase or the trailing phase based on its occurrence time relative to the field-specific production dates.

Before the event numbers and maximum magnitudes are calculated, the catalogue is filtered using the magnitude of completeness, M_C. The magnitude of completeness is the lowest magnitude above which the earthquake catalogue is assumed to be complete. Only earthquakes with magnitudes equal to or larger than this threshold are included:

$$M_i \geq M_C$$

This filtering step is necessary because smaller earthquakes may be missing from the catalogue, especially in older periods or in areas with fewer seismic stations. If incomplete low-magnitude data were included, the calculated production and trailing event counts could be biased. By applying M_C, the comparison between production-phase and trailing seismicity is based only on events that are expected to be reliably recorded. After event attribution, phase separation, and completeness filtering, the following variables are defined for each gas field:

N_P = number of earthquakes during production
N_T = number of earthquakes after production
$M_{P,max}$ = largest earthquake during production
$M_{T,Max}$ = largest earthquake after production

Resulting into a maximum observed difference of:

$$\Delta M_{obs} = M_{T,max} - M_{P,max}$$

Schultz et al. (2022) base their maximum-magnitude framework on the unbounded Gutenberg-Richter magnitude-frequency distribution (GR-MFD):

$$N = 10^a 10^{-bM}$$

or equivalently:

$$log_{10}(N) = a - bM$$

Where N is the number of earthquakes with magnitude greater than or equal to M, the a-value describes earthquake productivity and the b-value describes the relative proportion of small and large earthquakes (Schulz et al., 2022).

The confidence-based maximum-magnitude estimate is derived from the sample size hypothesis. This hypothesis states that the expected largest earthquake in a catalogue depends on how many earthquakes are sampled from the Gutenberg-Richter distribution. A larger catalogue has a higher probability of containing a larger event. Assuming an unbounded Gutenberg-Richter distribution, the largest observed event scales with the logarithm of the total number of induced earthquakes (Van der Elst et al., 2016):

$$M_{max} \approx M_C + \frac{1}{b} log_{10}(N)$$

To include uncertainty in the predicted largest earthquake, a confidence parameter $\boldsymbol{u}$ from the sample-size approach used by van der Elst et al. (2016) and Schultz et al. (2022) is included.

$$M_{max}(u) \approx M_C + \frac{1}{b}log_{10}(N) - \frac{1}{b}log_{10}(-ln(u))$$

Where $M_{max}(u)$ is the maximum magnitude at confidence level $\boldsymbol{u}$, and $\boldsymbol{u}$ is a confidence parameter between 0 and 1. Schultz et al. (2022) use this sample-size expression as the basis for their trailing seismicity model.

Now here we use the same logic as Schulz et al. (2022) to derive the formulation for the relation between gas production earthquakes and its trailing earthquakes:

$$\Delta M = M_{max} - M_{P,max} \approx \frac{1}{b}log_{10}\left(1 + \frac{N_T}{N_P}\right) + \frac{1}{b}log_{10}\left(\frac{ln(u_P)}{ln(u)}\right)$$

This follows directly from:

$$M_{P,max}(u_P) \approx M_C + \frac{1}{b}log_{10}\left(N_P\right) - \frac{1}{b}log_{10}\left(-ln(u_P)\right)$$
$$M_{max}(u) \approx M_C + \frac{1}{b}log_{10}\left(N_P + N_T\right) - \frac{1}{b}log_{10}(-ln(u))$$

Now the difference in magnitude between the production and the combined phase is

$$\Delta M = M_{max}(u) - M_{P,max}(u_P)$$
$$\Delta M = \frac{1}{b}log_{10}\left(\frac{N_P+N_T}{N_P}\right) + \frac{1}{b}log_{10}\left(\frac{ln(u_P)}{ln(u)}\right)$$

Here:

$$\frac{N_P+N_T}{N_P} = 1 + \frac{N_T}{N_P}$$

Thus:

$$\Delta M \approx \frac{1}{b}log_{10}\left(1 + \frac{N_T}{N_P}\right) + \frac{1}{b}log_{10}\left(\frac{ln(u_P)}{ln(u)}\right)$$

Where:

$$\frac{1}{R_P} = \frac{N}{N_P} = 1 + \frac{N_T}{N_P} = 1 + R_{TP}$$
$$R_{TP} = \frac{N_T}{N_P}$$
$$R_P = \frac{N_P}{N_P+N_T}$$

In addition to the unbounded Gutenberg-Richter magnitude-frequency distribution, a bounded Gutenberg-Richter distribution can also be considered. In the unbounded formulation, there is no physical upper limit on the possible earthquake magnitude. This assumption is used by Van der Elst et al. (2016), who show that the largest observed induced earthquakes can be explained by sampling statistics from an unbounded Gutenberg-Richter distribution. Schultz et al. (2022) use this same sample-size logic in their statistical framework for trailing seismicity. This is useful for statistical sampling, but it may be less realistic for gas fields where the maximum possible rupture size is limited by reservoir dimensions, fault length, reservoir thickness, and stress conditions. Therefore, a bounded Gutenberg-Richter distribution introduces an upper magnitude limit. We introduce this as the M_U;

$$M_C < M < M_U$$

In the bounded Gutenberg-Richter model, the cumulative distribution function can be written as:

$$F_B(M) = \frac{1-10^{-b(M-M_C)}}{1-10^{-b(M_U-M_C)}}$$

Here, $F_B(M)$ is the probability that an earthquake magnitude is smaller than or equal to M. The denominator normalises the distribution between M_C and M_U (Beirlant et al., 2019). This means that:

$$F_B(M_C) = 0$$

$$F_B(M_U) = 1$$

The corresponding exceedance probability is the following:

$$P(M_i \geq M) = \frac{10^{-b(M-M_C)} - 10^{-b(M_U-M_C)}}{1-10^{-b(M_U-M_C)}}$$

This formula follows directly from the same bounded cumulative distribution, because exceedance probability is the probability of observing a magnitude equal to or larger than M within the bounded interval.

To estimate the largest magnitude in a catalogue of N earthquakes, the sample-size approach is used. Van der Elst et al. (2016) show that induced earthquake maximum magnitudes can be interpreted using Gutenberg-Richter sampling statistics. For a bounded distribution, if $M_{lrg,B}$ is the largest earthquake in the sample, then all N earthquakes must be smaller than or equal to this value:

$$P(M_{lrg,B} \geq M) = F_B(M)^N$$

Using a confidence level u, this becomes:

$$F_B(M)^N = u$$

Taking the N-th root gives:

$$F_B(M) = u^{\frac{1}{N}}$$

Substituting the bounded Gutenberg-Richter cumulative distribution gives:

$$\frac{1-10^{-b(M-M_C)}}{1-10^{-b(M_U-M_C)}} = u^{\frac{1}{N}}$$

Rearranging gives:

$$1 - 10^{-b(M-M_C)} = u^{\frac{1}{N}}\left(1 - 10^{-b(M_U-M_C)}\right)$$

Therefore:

$$10^{-b(M-M_C)} = 1 - u^{\frac{1}{N}}\left(1 - 10^{-b(M_U-M_C)}\right)$$

Taking the $\log_{10}$ and solving for M gives the following:

$$M_{lrg,B}(u) = M_C - \frac{1}{b} log_{10}\left(1 - u^{\frac{1}{N}}\left(1 - 10^{-b(M_U-M_C)}\right)\right)$$

This is the bounded largest-magnitude equation. It follows the same sample-size reasoning used by Van der Elst et al. (2016), but it includes the upper-bound term from the truncated Gutenberg-Richter distribution described by Beirlant et al. (2019).

For this thesis the bounded Gutenberg-Richter approach has been used.

5. Code Description

For this project two MATLAB scripts were used, script_MakeTrailDist_main.m and script_dMtrail.m. These were originally created by Ryan Schultz, after minimally altered by Annemarie Muntendam-Bos and finally adjusted by Bubu Schmal to fit for this project.
Full code descriptions are readable in Appendix A and B respectively.

5.1 TrailDistance

The TrailDistance script analyses the production-seismicity ratio R_p and fits the observed data to a beta distribution. The ratio R_p describes the fraction of earthquakes that occur during the operation compared with the total number of earthquakes. Schultz et al. (2022) showed that that the distribution for injection/stimulation induced seismicity follows a beta distribution. TrailDistance uses a bootstrap methodology to derive the optimum parameter for the beta distribution and a Kolmogorov-Smirnov test to determine the significance of this beta distribution compared to a random distribution.

5.2 dMtrail

The second script used was the script_dMtrail.m. This script estimates the distribution of the largest expected trailing earthquake magnitude, M_{lrg}, after production has stopped or after a red-light/shut-in situation.

The main purpose of the script is to estimate the possible magnitude range of trailing seismicity. Instead of using one fixed value, the script samples possible combinations of input parameters. These include the Gutenberg-Richter b-value, the seismicity ratio R_p, and the maximum magnitude parameter M_{max}. By sampling these parameters, the script creates a probabilistic distribution of possible largest trailing magnitudes.

6. Selection of Gas Fields

The Dutch gas fields considered for this research are shown in Table 1, alongside the dates of production start and end, their statistical relations according to chapter 4 'Methodology', and the largest observed magnitude. These include the total number of events (N_{total}, M ≥ M_C), observed events during production (N_P, M ≥ M_C), observed trailing events (N_T, M ≥ M_C) and the ratios N_T/N_P and R_P (N_P/N_{total}).

Considered Dutch Gas Fields									
Field	Production start	Production end	Mc	Np	Nt	Ntotal	Nt/Np	Rp = Np/Ntotal	Mlrg, observed
Annerveen	01-01-1973	01-07-2021	0.9	45	3	48	0.067	0.938	2.43
Appelscha	01-01-1999	01-09-2011	1.5	4	0	4	0.000	1.000	2.5
Bergen	13-10-1978	01-12-2024	0.	3	0	3	0.000	1.000	2.7
Boerakker	01-07-1998	01-03-2022	0.9	0 ± 1	0	0 ± 1	0.000	1.000	0.94
Dalen	01-01-1974	01-01-2018	1.2	4	1	5	0.250	0.800	2.24
Eleveld	01-01-1975	01-01-2026	1.2	42 ± 2	2	44 ± 2	0.047 ± 0.002	0.955 ± 0.002	2.99
Emmen	01-01-1969	01-02-2018	1.2	15 ± 1	0	15 ± 1	0.000	1.000	2.26
Emmen-Nieuw Amsterdam	01-01-1975	01-01-2019	1.2	0 ± 1	0	0 ± 1	0.000	1.000	1.7
Groningen	01-01-1963	01-10-2023	1.2	596	22 ± 1	618 ± 1	0.038 ± 0.001	0.964 ± 0.001	3.6
Houwerszijl	01-06-2000	01-10-2010	0.9	0	0 ± 1	0 ± 1	no value	0.000	1.44
Metslawier	01-12-1997	01-07-2016	0.9	1	0	1	0.000	1.000	1.43
Pasop	01-08-1997	01-03-2015	0.9	0 ± 1	0	0 ± 1	0.000	1.000	0.94
Roden	01-01-1976	01-02-2003	0.9	5 ± 2	2 ± 2	7 ± 4	0.400 ± 0.400	0.714 ± 0.286	2.1
Roswinkel	01-09-1980	01-01-2006	1.2	35	2	37	0.057	0.946	3.35
Saaksum-West	01-04-1999	01-01-2024	0.9	4 ± 1	0	4 ± 1	0.000	1.000	1.7
Ureterp	01-01-1978	01-03-2024	0.9	1 ± 2	0	1 ± 2	0.000	1.000	1.8
Vries-Centraal	01-01-1976	01-01-2026	0.9	4	0	4	0.000	1.000	1.2
Vries-Noord	01-01-1976	01-04-2025	0.9	3 ± 4	0	3 ± 4	0.000	1.000	1.94
Witterdiep	01-06-2007	01-12-2013	0.9	1 ± 2	0	1 ± 2	0.000	1.000	2.5
Zuidwending-Oost	01-07-2006	01-01-2024	0.9	0 ± 1	0	0 ± 1	0.000	1.000	1.6

Table 1: Overview of the Dutch gas fields considered, including their production dates and statistical relations

The first notable thing is that many Dutch gas fields have a low number of earthquakes. Several fields have no trailing earthquakes at all, while others have only one or a few production-related events. Although these fields are still considered, their R_p values are less reliable, because a single event can strongly influence the calculated ratio. For example, Roden has 5+/-2 events observed, with 2+/-2 trailing events, which results in an R_p-value of 0.714+/-0.266. In contrast, Eleveld has 42+/-2 events and 2 trailing events, which results in an R_p-value of 0.955+/-0.002. Therefore, the R_p-values of fields with a low number of observed events contain a large uncertainty. If all fields are included in the beta-distribution fit with equal weight, these uncertainties are not taken into account and introduce a bias in the fit.

For this reason, the selection was narrowed down to the fields with >20 earthquakes in the minimum scenario, thus reducing the uncertainty in R_P to < 0.05. Groningen, Annerveen, Eleveld, and Roswinkel clearly stand out in the table because they have a relatively high number of earthquakes without an R_P of 1. Groningen clearly dominates the total seismicity; however, in these smaller gas fields enough seismicity has occurred to create meaningful data.

Groningen will also be used here as a control field because its abundant seismicity provides the statistical basis needed to validate the method. Its catalogue contains 1,396 detected events (December 1991–November 2021) across a field with over 1,100 mapped faults (Muntendam-Bos & Grobbe, 2022).

7. Parameters Input

As the methodology is based on a Monte Carlo simulation, different input parameters need to be stated. The gas field-specific constant parameters include the M_C and N_p. The logic tree parameters include b-value, R_P and M_{max}.

7.1 Groningen

Groningen is the most extensively researched Dutch gas field. For that reason, besides parameters, a decision tree has been made. Besides the parameters in the Groningen logic tree, there is one more:

N_P	596

Mmax is one of the key but also uncertain inputs for the Monte Carlo simulation. Therefore, the Groningen Mmax value is taken from the logic tree of the Groningen Mmax Panel of Nederlandse Aardolie Maatschappij B.V. (2022). The Groningen logic tree, visible in figure 2, represents knowledge-based uncertainty by the Groningen Mmax Panel in the maximum possible earthquake magnitude, Mmax, for seismic hazard and risk analysis of the Groningen gas field. In this context, Mmax is not treated as one fixed value but as a distribution of defensible upper-bound magnitudes. This is important because the Groningen Mmax Panel explains that Mmax is the largest magnitude that a seismic source is expected to produce. This means that Mmax sets the highest possible magnitude used in the hazard model (Nederlandse Aardolie Maatschappij B.V., 2022). The report also emphasises that the Groningen Mmax assessment applies specifically to seismicity caused by gas extraction from the Groningen field and not to natural tectonic earthquakes in the region.

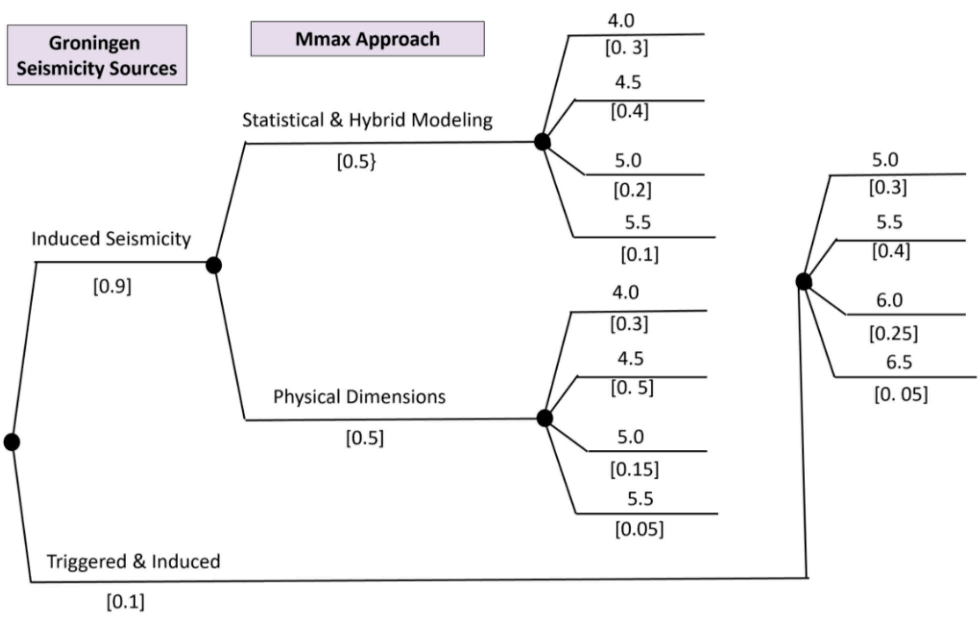

Figure 2: The Groningen logic tree from the Nederlandse Aardolie Maatschappij B.V. (2022)

The first split in the Groningen logic tree separates two possible seismicity-source models: induced seismicity only and triggered-and-induced seismicity. The induced seismicity branch receives a dominant weight of 0.9. This branch assumes that future Groningen earthquakes are mainly caused by gas production. The report explains that this branch is supported by evidence that observed Groningen seismicity is strongly related to gas extraction processes and reservoir compaction (Nederlandse Aardolie Maatschappij B.V. 2022).

The second main branch represents a combined triggered-and-induced seismicity scenario and receives a lower weight of 0.1. This branch is included because the possibility of triggered seismicity cannot be fully ruled out, even though the report states that triggered seismicity has not been clearly observed so far in Groningen. Triggered events may involve ruptures extending beyond the capabilities of the gas production-related seismicity, which means that this branch can include larger possible M_{max} values than the induced-only branch (Nederlandse Aardolie Maatschappij B.V. 2022).

Within the induced branch, the logic tree is divided into two equally weighted approaches for estimating Mmax: the statistical and hybrid modelling approach and the physical-dimensions approach. Both receive a weight of 0.5. The statistical and hybrid modelling branch uses the observed Groningen earthquake catalogue together with physical modelling approaches, such as stress changes, failure criteria, and simulations of spatial and temporal seismicity patterns. The physical-dimensions branch estimates possible maximum magnitudes from the dimensions of plausible ruptures within the reservoir, using information on mapped faults, reservoir structure, rupture length, rupture width, and empirical magnitude-scaling relationships. The panel gives equal weight to these two approaches because they represent different but defensible ways of estimating Mmax: one based mainly on observed seismicity and modelling and the other based on possible rupture within the reservoir (Nederlandse Aardolie Maatschappij B.V., 2022).

The final Groningen Mmax logic tree, therefore, gives most probability to lower maximum magnitudes in the induced-seismicity branch while retaining a small probability of larger magnitudes through the triggered-and-induced branch. In the 2022 assessment, the Mmax distribution ranges from M=4.0 to M=6.5, with most of the probability mass between M=4.0 and M=5.0. The discrete probability weights are visible in Figure 2 and Figure 3. This shows that the model is dominated by induced seismicity but still includes a conservative low-probability tail for larger triggered scenarios (Nederlandse Aardolie Maatschappij B.V. 2022).

In this project, the Groningen logic tree was adapted for use in the Monte Carlo model. The literature-based Mmax structure was combined with a normal distribution of Groningen-specific seismicity parameters, a b-value of 0.89 (Nederlandse Aardolie Maatschappij B.V. 2022) and an R_p value of 0.95. This adapted decision tree is presented in figure 3.

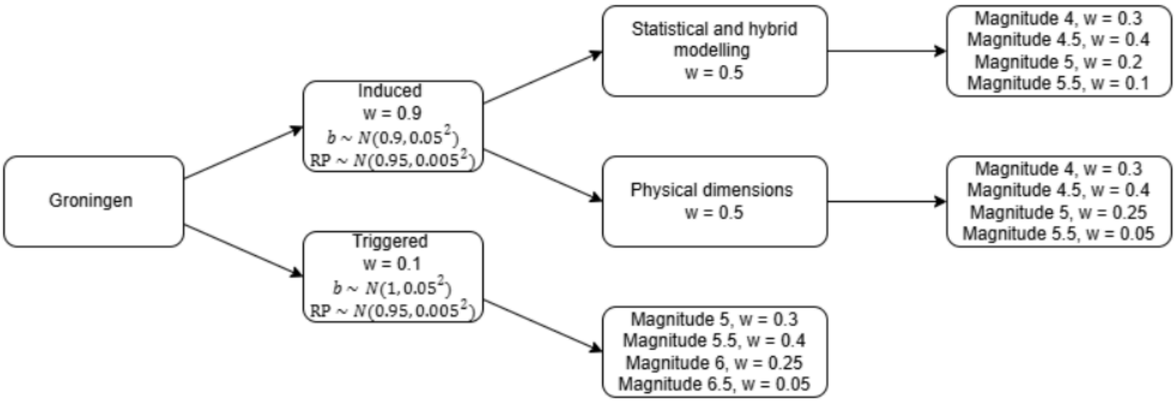

Figure 3: Adapted Groningen logic tree

7.2 Annerveen

The input parameters for Anneveen are the following:

N_p	45
b-value	A normal distribution with the tectonically standard distribution of $N(1.0, 0.05^2)$
R_p	A normal distribution of $(0.9375, 0.005^2)$
M_{max}	4.0 (TNO, 2022). From this a normal distribution has been made of $N(4.0, 0.1^2)$

7.3 Eleveld

The input parameters for Eleveld are the following:

N_p	44
b-value	A normal distribution with the tectonically standard distribution of $N(1.0, 0.05^2)$
R_p	A normal distribution of $N(0.9550, 0.005^2)$
M_{max}	3.6 (Nederlandse Aardolie Maatschappij B.V. [NAM], 2016). From this a normal distribution has been made of $N(3.6, 0.1^2)$

7.4 Roswinkel

The input parameters for Roswinkel are the following:

N_p	35
b-value	A normal distribution with the tectonically standard distribution of $N(1.0, 0.05^2)$
R_p	A normal distribution of $N(0.9459, 0.005^2)$
M_{max}	3.8 (Nederlandse Aardolie Maatschappij B.V. [NAM], 2007). From this, a normal distribution has been made of $N(3.8, 0.1^2)$

8. Results

Groningen is presented first as a control field, followed by Annerveen, Eleveld, and Roswinkel. For each field, the results report the input parameters used, the modelled M_{lrg} output and the R_p value or adjustment where applicable.

8.1 Beta Distribution fit

The four selected fields have R_P values close to 0.95 (Table 1). Figure 1 shows the fitted beta distribution for the fraction of earthquakes that occur during production, R_P. The fitted beta distribution shows a strong peak near high R_P values. The scatter plot of the bootstrapped alpha and beta parameters shows a clear positive relationship, while the histograms show that most bootstrap results cluster around the mean parameter values. The fitted mean parameters are α = 495 and β = 26, whereas the background data of Schultz et al. (2022) is α = 0.73 and β = 0.11.

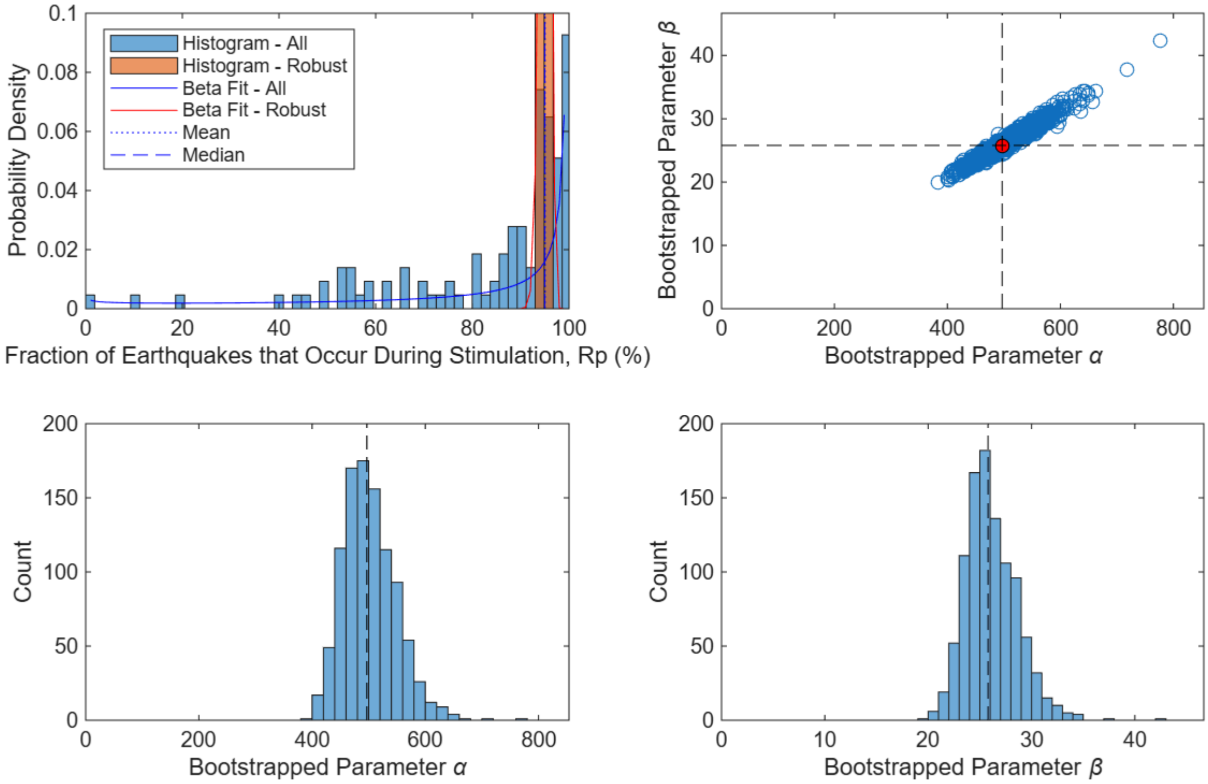

Figure 1: Fitted beta-distribution of the selected gas fields on top of a database from Schulz et al. (2022)

8.2 Groningen

The Groningen results are presented in two steps. First, the present model uses the observed production and trailing seismicity available at the time of analysis, together with the original R_p value of approximately 0.95. This model is used to test whether the observed M = 3.4 Zeerijp earthquake falls within the predicted M_{lrg} distribution. Second, an adjusted Groningen model is introduced to account for the fact that Groningen is still seismically active after production stopped. This adjusted model uses a lower R_p value to include projected future trailing seismicity.

8.2.1 Present

The input parameters for Groningen are described in Section 7.1. The production earthquake count used in the model is NP = 596. The initial Groningen logic tree used a b-value of 0.89, an Rp value of approximately 0.95, and the Mmax structure shown in Figure 3.

The Zeerijp earthquake occurred on November 14th, 2025, with a magnitude of 3.4. This is the largest trailing earthquake that occurred in the Groningen gas field. Figure 4 models the distribution of the largest expected magnitude, M_{lrg}, and shows that values centre around magnitude 4, with a range extending from approximately 3 to above 5.5. The observed Zeerijp earthquake with magnitude 3.4 falls within the lower part of the predicted distribution.

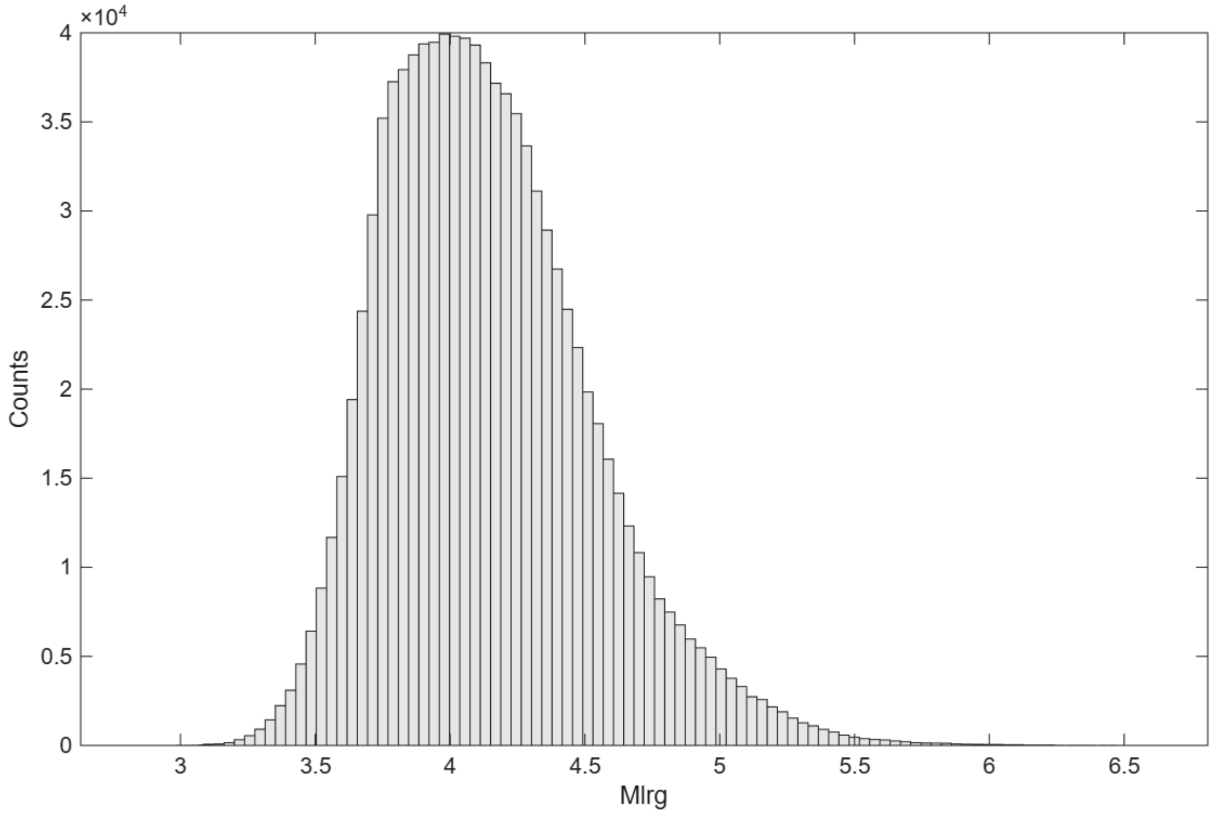

Figure 4: M_{lrg} distribution for Groningen based on the Figure 3 logic tree

Figure 5 displays how the input parameters distribute for Figure 4 and shows the predicted total number of events as per the April 2026 dataset. The predicted number of events peaks at 627 and ranges from approximately 612 to 642.

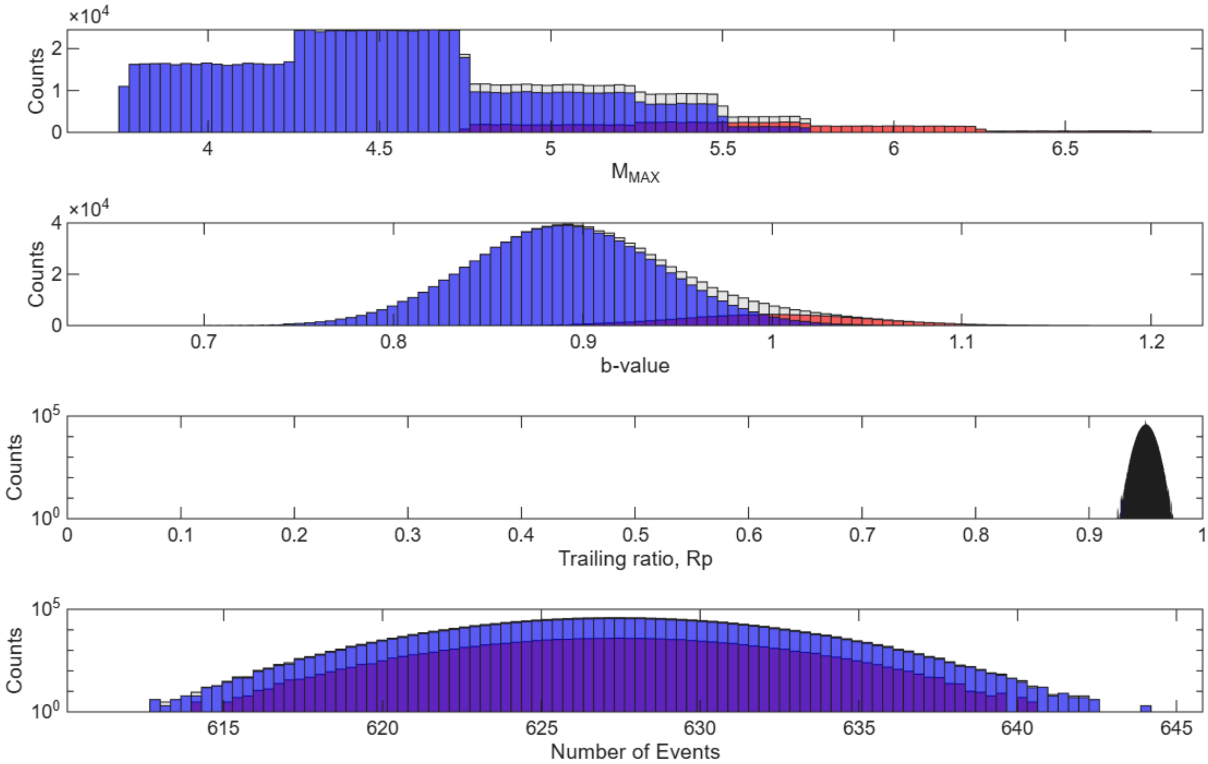

Figure 5: The input parameters for Figure 4 and the predicted total amount of events

8.2.2 Adjusted Groningen

The Groningen gas field is still seismically active and is predicted to remain active for a while. In the previous results, R_p was based on the number of trailing events observed from 2024 until April 2026: 17 earthquakes above Mc = 1.2. In order to make a prediction of the distribution of Mlrg, Figure 4-9 (Appendix C) of Nederlandse Aardolie Maatschappij (2025) was used. Based on this, Table 2 was constructed. From Table 2, the expected number of trailing earthquakes predicted for 2026-2050 is 92. Together with the 17 observed trailing events from 2024 until April 2026, this results in a predicted R_p of 0.838. Furthermore, to obtain the increased range the R_p normal distribution has been changed to $N(0.833, 0.03^2)$ The resulting final decision tree is shown in Figure 6.

Projected Seismicity per Year / Period			
Year	**Minimal (#/Year)**	**Expected (#/Year)**	**Maximum (#/Year)**
2026	2	8	20
2027	0	8	18
2028	0	6	16
2029	0	6	16
2030	0	6	14
2031	0	4	14
2032-2036	0	4	12
2037-2039	0	4	10
2040-2045	0	2	10
2046-2050	0	2	8
Total	**2**	**92**	**288**

Table 2: Projected seismicity according Ministerie van Klimaat en Groene Groei (2025), Appendix I

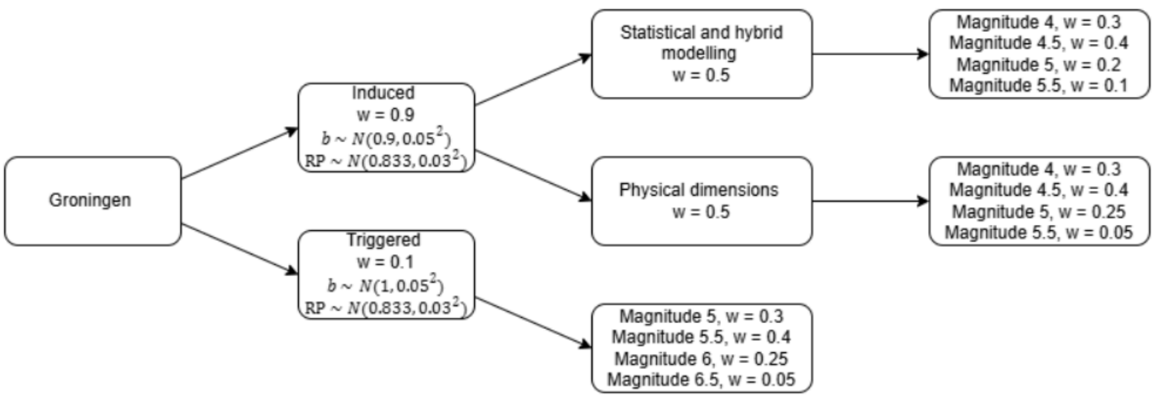

Figure 6: Updated Groningen logic tree as per newly introduced R_p

Using the predicted R_p, the newly produced distribution of M_{lrg} and input parameter figures are visible in Figures 7 and 8, respectively. With the adjusted R_p the expected number of total earthquakes lies between ~620 and ~850 above M_c = 1.2 and peaks around ~710. Meaning, as per April 2026, the expected number of earthquakes still to occur is between 0 and ~230. Furthermore, figure 7 does not show a major change in the M_{lrg} distribution compared to figure 5.

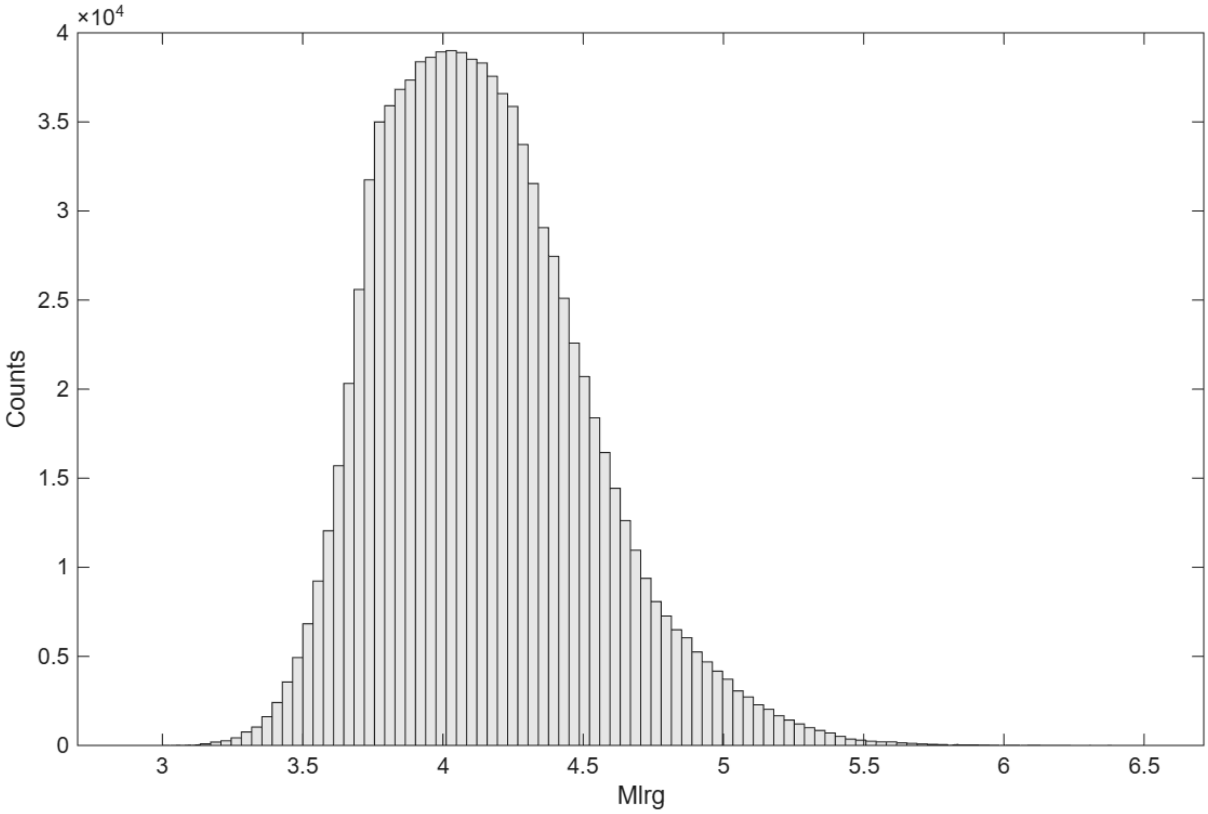

Figure 7: Updated M_{lrg} distribution based on the updated R_p value

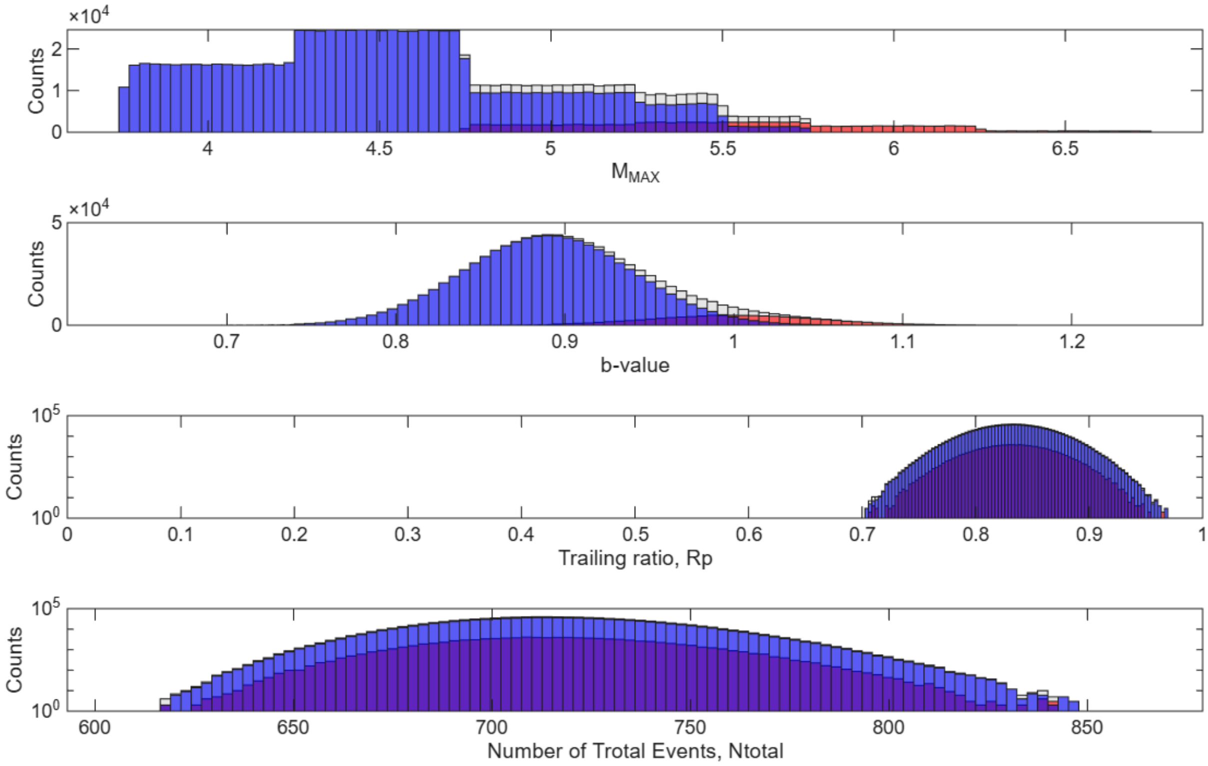

Figure 8: The updated input parameters as per updated R_p value

8.3 Annerveen

Annerveen has NP = 45 production events, NT = 3 trailing events, Ntotal = 48, Rp = 0.938, and an observed largest magnitude of Mlrg = 2.43 (Table 1). The input parameters used in the model, as described in section 7, are b = N(1.0, 0.05^2), Rp = N(0.9375, 0.005^2), and Mmax = N(4.0, 0.1^2).

The largest magnitude observed for Annerveen is 2.43. Figure 9 models the M_{lrg} distribution for the Annerveen gas field. The distribution peaks around 2.6 and ranges from approximately 1.5 to 4.3. Therefore, the largest observed magnitude of 2.43 lies just left of the distribution peak.

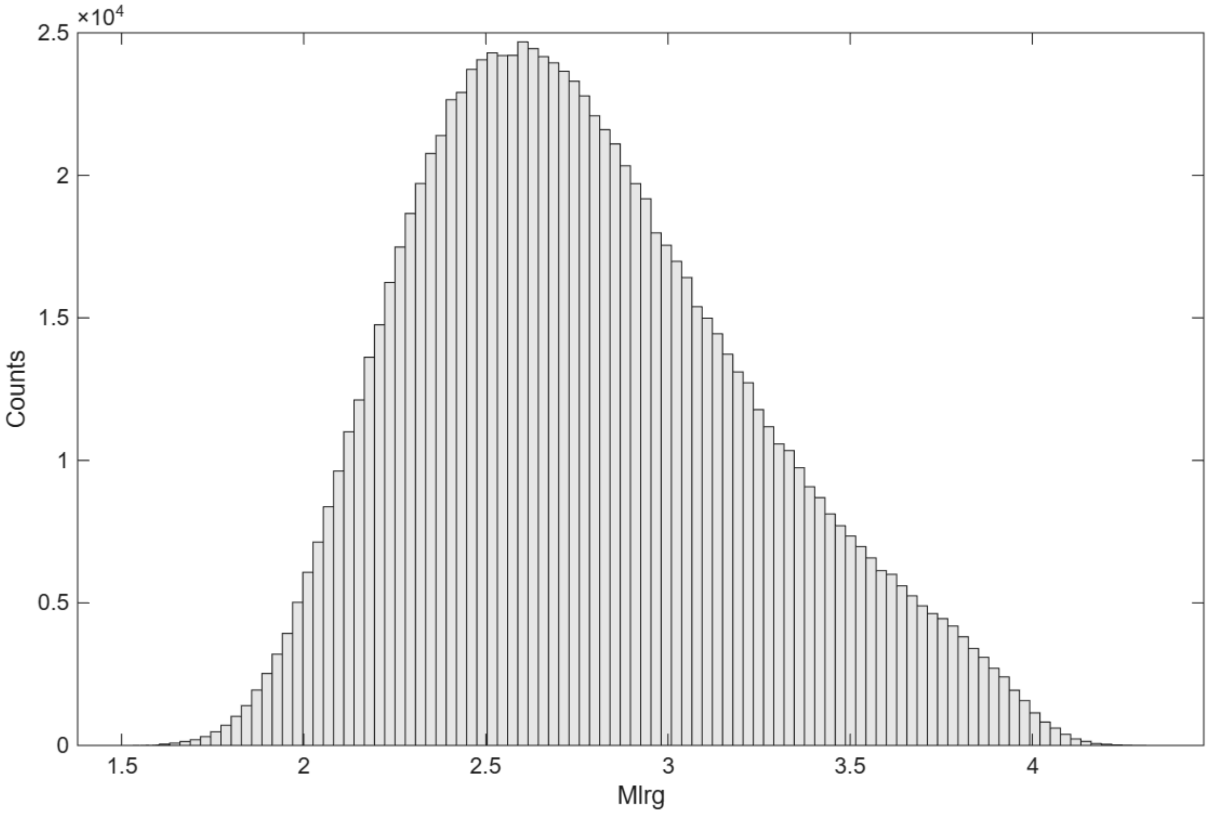

Figure 9: M_{lrg} distribution for Annerveen

Figure 10 models how the input parameters distribute for Figure 9. Furthermore, it models the total number of events to be between approximately 47 and 49.

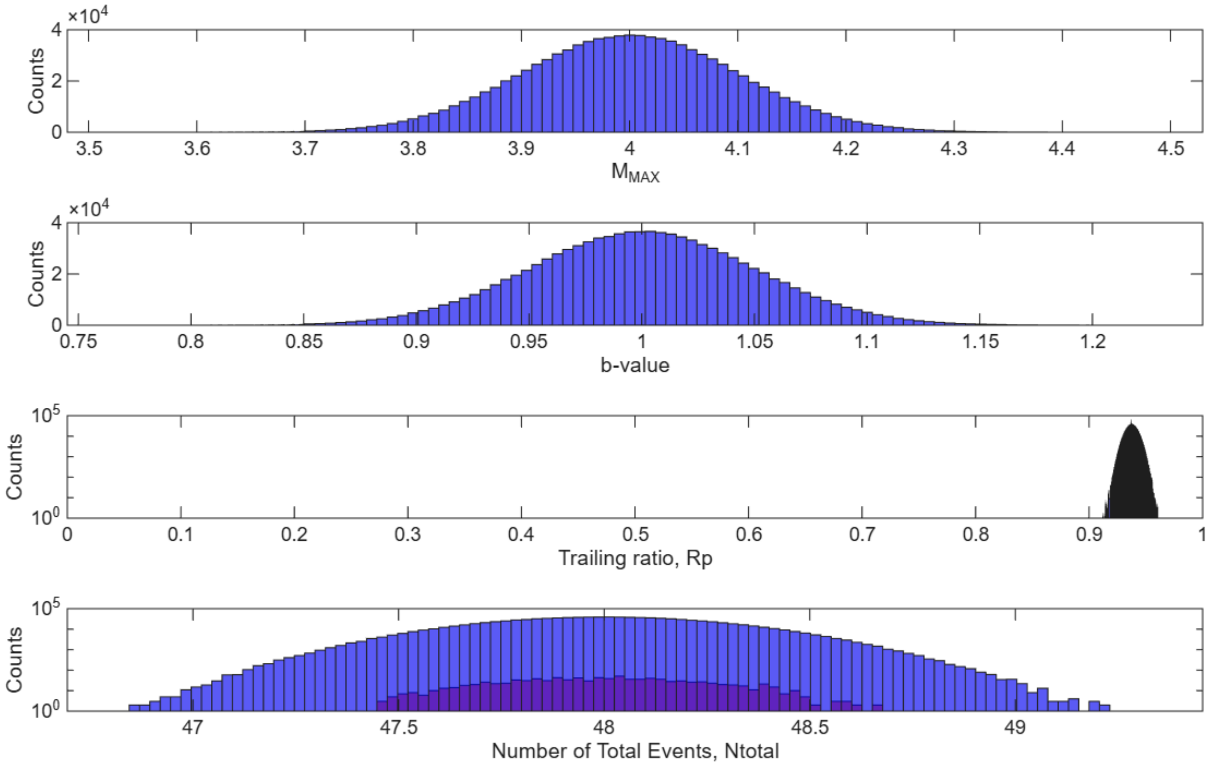

Figure 10: The input parameters for Figure 9 and the predicted total amount of events

8.4 Eleveld

Eleveld has NP = 42 +/- 2 production events, NT = 2 trailing events, Ntotal = 44 +/- 2, Rp = 0.955 +/- 0.002, and an observed largest magnitude of Mlrg = 2.99 (Table 1). The input parameters used in the model, as described in section 7, are b = N(1.0, 0.05^2), Rp = N(0.9550, 0.005^2), and Mmax = N(3.6, 0.1^2).

The largest magnitude observed for Eleveld is 2.99. Figure 11 models the M_{lrg} distribution for the Eleveld gas field. The distribution peaks around 2.6 and ranges from approximately 1.5 to 3.8. Therefore, the largest observed magnitude of 2.99 lies to the right of the distribution peak.

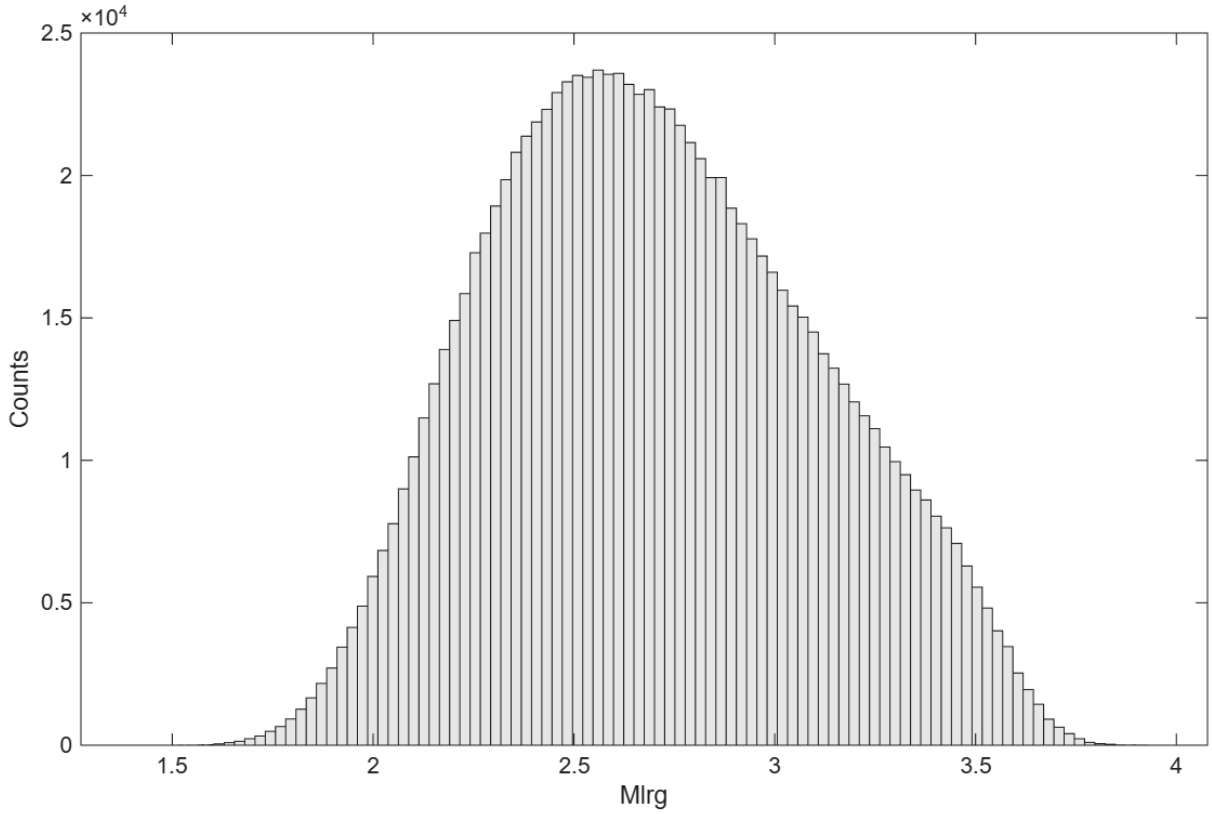

Figure 11: M_{lrg} distribution for Eleveld

Figure 12 models how the input parameters distribute for Figure 11. Furthermore, it models the total number of events to be between approximately 45 and 47.

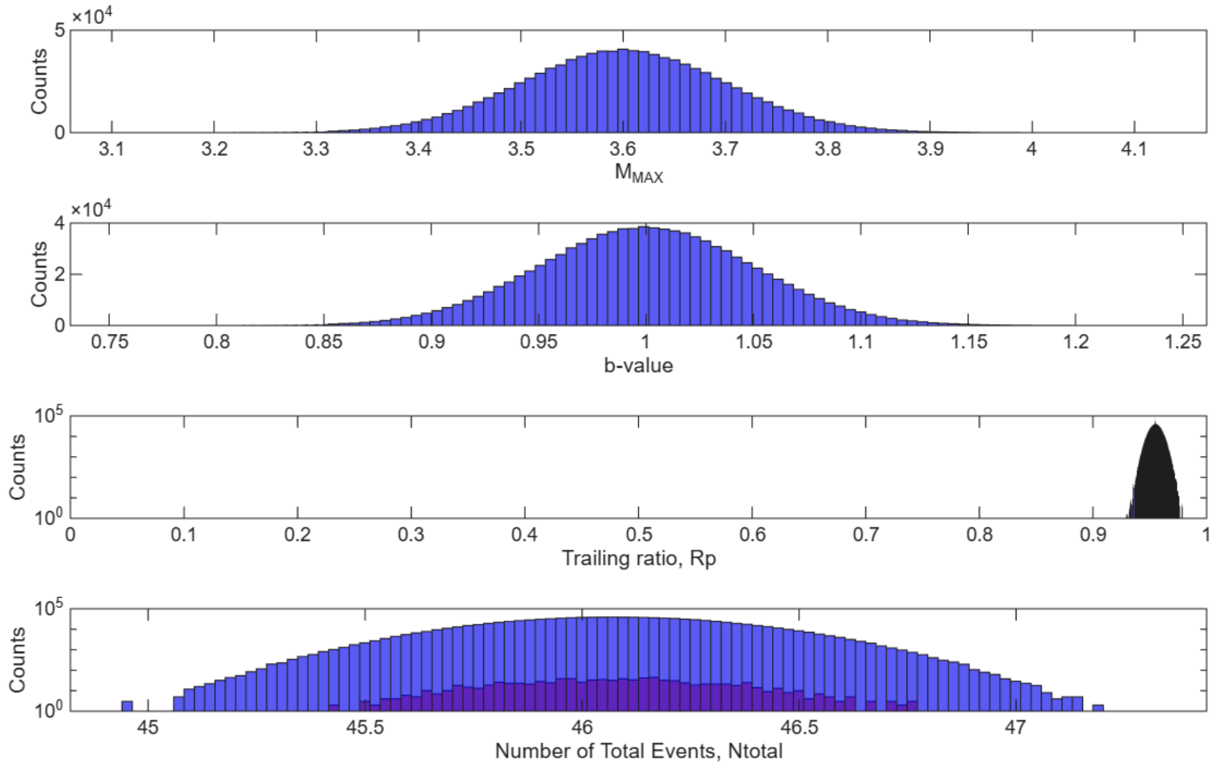

Figure 12: The input parameters for Figure 11 and the predicted total amount of events

8.5 Roswinkel

Roswinkel has NP = 35 production events, NT = 2 trailing events, Ntotal = 37, Rp = 0.946, and an observed largest magnitude of Mlrg = 3.35 (Table 1). The input parameters used in the model are b = N(1.0, 0.05^2), Rp = N(0.9459, 0.005^2), and Mmax = N(3.8, 0.1^2).

The largest magnitude observed for Roswinkel is 3.35. Figure 13 models the M_{lrg} distribution for the Roswinkel gas field. The distribution peaks around 2.5 and ranges from approximately 1.5 to 4.1. Therefore, the largest observed magnitude of 3.35 lies to the right of the distribution peak.

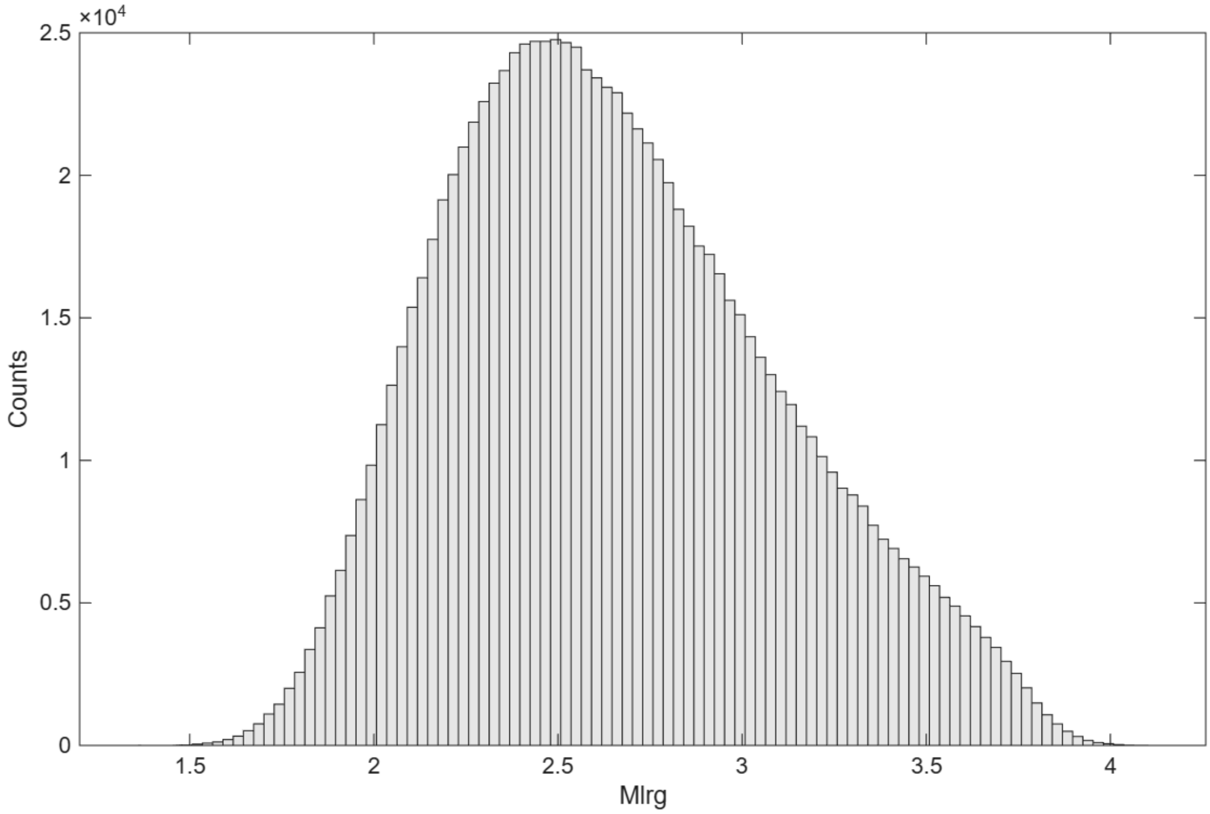

Figure 13: M_{lrg} distribution for Roswinkel

Figure 14 models how the input parameters distribute for Figure 13. Furthermore, it models the total number of events to be between approximately 36 and 38.

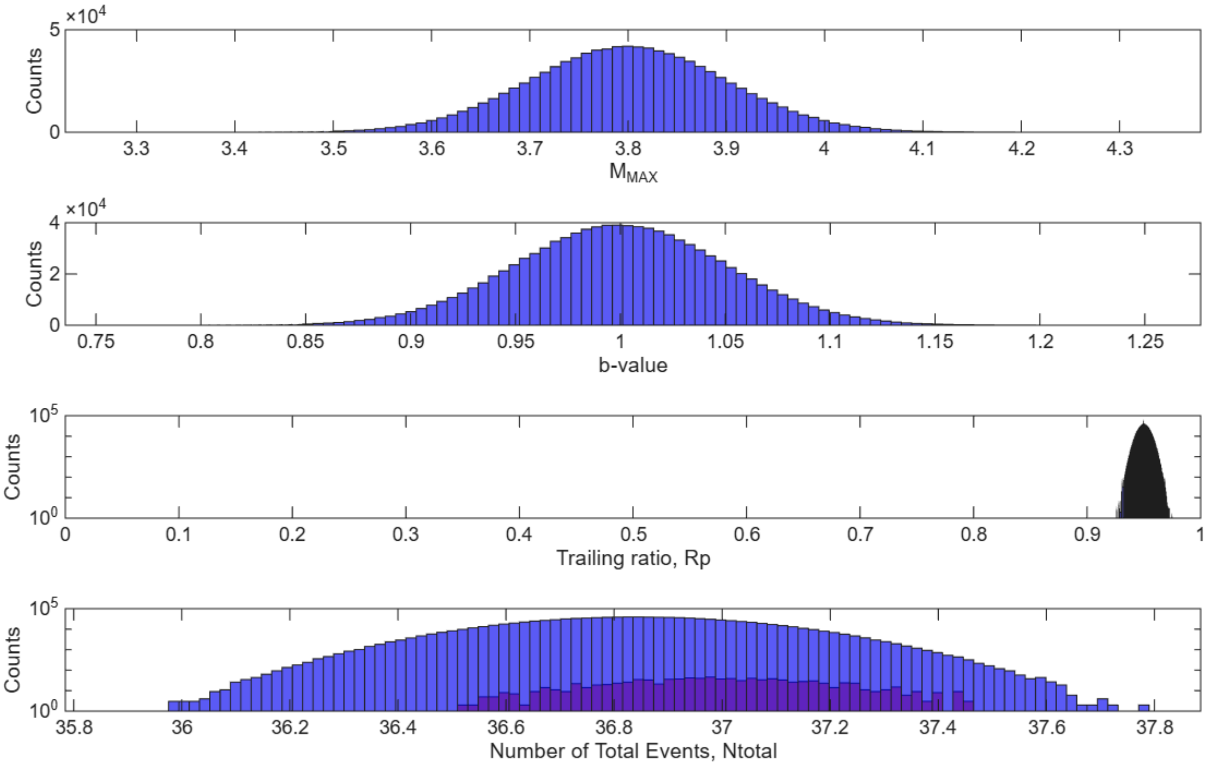

Figure 14: The input parameters for Figure 13 and the predicted total amount of events

9. Discussion

The aim of this study was to test whether the largest observed earthquakes in selected Dutch gas fields can be explained using a bounded Gutenberg-Richter-based adaptation of the Schultz et al. (2022) trailing-seismicity framework.

9.1 Gas Field Robustness

The beta-distribution fit shows that the selected Dutch gas fields are strongly concentrated near high R_p values. This means that, for these fields, most earthquakes occurred during the production period and only a small part occurred during the trailing period. The clustering of the bootstrapped alpha and beta parameters around the mean suggests that the fitted beta distribution is stable for the selected dataset.

However, the fitted parameters for the selected Dutch gas fields, α=495 and β=26, differ strongly from the background dataset of Schultz et al. (2022), where α=0.73 and β=0.11. This difference is important because it shows that the selected Dutch gas fields do not follow the same broad R_p distribution as the full Schultz et al. dataset. The Dutch fields form a much narrower subset, with R_p values concentrated close to 0.95.

This difference may be caused by the field selection used in this thesis, which is limited. Fields with too few events or R_p = 1 were excluded, which reduces the spread of R_p values. Therefore, the fitted beta distribution should not be interpreted as a general distribution for all induced-seismicity cases. Instead, it should be interpreted as a field-specific fit for the selected Dutch gas fields.

The high R_p values suggest that trailing seismicity forms only a small part of the total seismicity in these selected fields. However, this does not mean that trailing seismicity is unimportant. Even a small number of trailing events can include the large observed events, as shown by the Zeerijp earthquake in Groningen. Therefore, the R_p distribution is useful for estimating trailing-event occurrence, but it must be interpreted together with the modelled M_{lrg} distribution.

9.2 Justification Groningen

Groningen is used as the control field because it has by far the largest earthquake catalogue of the selected fields. With a N_p of 596 the Groningen model is better constrained than the models for Annerveen, Eleveld, and Roswinkel. Therefore, Groningen is used to test whether the method can reproduce a realistic largest observed magnitude for the field with the strongest data basis.

The largest observed trailing earthquake in Groningen is the Zeerijp earthquake, with a magnitude of 3.4. In the first Groningen model, the M_{lrg} distribution peaks around magnitude 4 and ranges from approximately magnitude 3 to above 5.5. The Zeerijp earthquake falls in the lower part of this predicted distribution. This means that the observed largest trailing magnitude is justified by the model, since it lies within the expected range.

However, the Zeerijp event is smaller than the most likely model outcome. The peak of the distribution is around magnitude 4, while Zeerijp has magnitude 3.4. This means that the observed maximum magnitude is not unusually large according to the model. Instead, it is a relatively low value within the predicted M_{lrg} range.

The adjusted Groningen model uses a lower R_p value because Groningen is still seismically active after production stopped. The original R_p value of approximately 0.95 was adjusted using 17 observed post-2024 earthquakes above M_c = 1.2 and 92 projected trailing earthquakes for 2026-2050. This resulted in a predicted R_p of approximately 0.838, with the R_p distribution changed to a normal distribution of $N(0.833, 0.03^2)$. This adjustment mainly increases the expected number of total earthquakes, which changes from a peak around 627 in the first model to a peak around 710 in the adjusted model.

Even after this R_p adjustment, the Mlrg distribution does not change drastically compared with the first Groningen model. However, the peak of the distribution shifts upward by approximately 0.2 magnitude units. This shows that the R_p adjustment has some influence on the predicted largest magnitude, but the overall shape and range of the distribution remain relatively similar. Therefore, for Groningen, the largest expected magnitude is still mainly controlled by the bounded Mmax logic-tree structure and the b-value, rather than only by the exact R_p value. The R_p adjustment mainly affects the expected number of future events, while only moderately shifting the position of the M_{lrg} peak.

Therefore, the Groningen result supports the method as a control case. The observed Zeerijp earthquake is within the predicted distribution and can be justified by the model. At the same time, the model predicts that a larger event than Zeerijp would also still be possible within the bounded distribution. This does not mean that such an event will occur, but it shows that the observed M = 3.4 event is not close to the upper limit of the modelled Groningen range.

9.3 Justification: Small Dutch Gasfields

For the smaller fields, the main question is whether the maximum observed magnitude can be justified by the modelled M_{lrg} distribution. As Annerveen, Eleveld, and Roswinkel contain far fewer events than Groningen, the model should not be interpreted as a precise prediction. Instead, it is used to test whether the observed largest event is located within the expected magnitude range.

For Annerveen, the largest observed magnitude is 2.43. The modelled M_{lrg} distribution peaks around magnitude 2.6 and ranges from approximately 1.5 to 4.3. The observed maximum magnitude lies just left of the distribution peak. This means that the Annerveen maximum observed magnitude is well justified by the model, because it falls close to the most likely part of the predicted distribution.

For Eleveld, the largest observed magnitude is 2.99. The modelled M_{lrg} distribution peaks around magnitude 2.6 and ranges from approximately 1.5 to 3.8. The observed maximum magnitude lies to the right of the peak, but still within the predicted range. This means that the Eleveld maximum observed magnitude is higher than the most likely model outcome, but it is still justified by the modelled distribution.

For Roswinkel, the largest observed magnitude is 3.35. The modelled M_{lrg} distribution peaks around magnitude 2.5 and ranges from approximately 1.5 to 4.1. The observed maximum magnitude lies clearly to the right of the distribution peak but remains inside the predicted range. This means that Roswinkel is the least central of the three small-field results, but the maximum observed magnitude is still justified by the bounded model.

Together, these results show that the maximum observed magnitudes of Annerveen, Eleveld, and Roswinkel can be justified by their modelled M_{lrg} distributions. None of the observed maximum magnitudes falls outside the predicted range. However, the position of the observed event differs between fields. Annerveen lies close to the peak, Eleveld lies moderately above the peak, and Roswinkel lies clearly above the peak. Therefore, the modelled results support the observed maxima.

9.4 Limitations

The main limitation of this study is the small number of suitable Dutch gas fields. Many fields contain only a few recorded earthquakes, and several fields have no observed trailing events. Because of this, their R_p values are sensitive to single events. One additional trailing earthquake can strongly change the R_p value for a small field, while the same change would have only a minor effect for Groningen. This is why the selection was limited to fields with more than 20 events in the minimum scenario, but even within this selection, Annerveen, Eleveld and Roswinkel are still much less statistically robust than Groningen. This means that the smaller field results should be interpreted with more caution than the Groningen result.

A second limitation is the uncertainty in the magnitude of completeness, M_C. The model only uses earthquakes above M_C, because events below this threshold may not be completely recorded. However, the choice of M_C directly affects the number of production and trailing events used in the calculation. If M_C is chosen too low, small events may bias the event counts. If M_C is chosen too high, useful data is removed and the catalogue becomes smaller. Therefore, the final R_P ranges partly depend on the selected completeness threshold.

A third limitation is the use of simplified b-values for the smaller gas fields. Groningen has enough seismicity to support a field-specific b-value, but for the smaller fields, a tectonically standard b-value was used. This is practical, but it also introduces uncertainty, because the b-value controls the relative probability of small and large earthquakes. A change in b-value can shift the predicted largest-magnitude distribution.

Another important limitation is the choice of the upper magnitude bound, Mmax. The bounded Gutenberg-Richter approach is more physically realistic than an unbounded model because it introduces an upper limit related to reservoir and fault dimensions. However, Mmax is still uncertain and must be based on literature values and logic-tree choices. Therefore, the model does not independently determine the true physical maximum magnitude. It tests whether observed events are reasonable within the chosen bounded framework. This is especially important for Groningen, where the M_{lrg} distribution is strongly influenced by the Mmax logic-tree structure.

A final limitation is that the method is mainly statistical. It uses event counts, b-values, R_P values and Mmax distributions, but it does not directly model the physical stress changes in the reservoir or on individual faults. These changes are therefore only indirectly represented. This means that the model can show whether an observed largest event is statistically plausible, but it cannot explain exactly why that event occurred at a specific time and location.

9.5 Future

Future work could first improve the method by updating the earthquake catalogues as new trailing events occur. This is relevant especially for Groningen, since its projected remaining events are very widespread. Adding new events would make it possible to update R_P through the new number of trailing events and the predicted M_{lrg} distribution over time. Furthermore, with more events occurring, more small fields could be considered.

A second improvement would be to estimate field-specific b-values for more gas fields. In this study, the smaller fields used a standard b-value distribution. If gas field-specific b-values were calculated and used, the M_{lrg} distributions would become more accurate.

Future research could also test different choices for Mmax. Since the bounded model is sensitive to the selected upper magnitude limit, it would be useful to compare multiple Mmax scenarios for each field. This would show how much of the final M_{lrg} distribution is caused by the Mmax assumption itself.

Finally, the method could be tested on other induced-seismicity gas-related datasets outside the Netherlands. The Dutch gas fields form a narrow subset with high R_P values, while the broader Schultz et al. (2022) dataset has a much wider R_P distribution. Applying the same bounded method to other gas production cases would show whether the approach is generally applicable or mainly useful for Dutch gas field seismicity.

10. Conclusion

This study adapted the Schultz et al. (2022) trailing-seismicity framework to selected Dutch gas fields by combining R_P, b-value and a bounded Gutenberg-Richter Mmax approach. The aim was to test whether the largest observed earthquakes in Groningen, Annerveen, Eleveld and Roswinkel can be justified within the modelled M_{lrg} distributions.

The results show that the selected Dutch gas fields have high R_p values, meaning that most earthquakes occurred during production and only a smaller part occurred after production stopped. However, trailing seismicity remains important, because a relatively small number of trailing events can still include a significant earthquake, such as the Zeerijp event in Groningen.

For Groningen, the observed Zeerijp earthquake of M = 3.4 falls within the predicted M_{lrg} distribution and can therefore be justified by the model. However, it lies below the distribution peak, which is around magnitude 4. This means that Zeerijp event is not close to the upper limit of the modelled Groningen range. After adjusting R_P for expected future trailing events, the expected number of earthquakes increases, and the M_{lrg} peak shifts upward by about 0.2 magnitude units. This shows that R_P has some influence, but that the Groningen result is mainly controlled by the bounded Mmax logic tree and the b-value.

For Annerveen, Eleveld and Roswinkel, the largest observed magnitudes also fall within the modelled distributions. Annerveen lies close to the distribution peak, while Eleveld and Roswinkel lie to the right of the peak but still within the predicted range. Therefore, the observed largest magnitudes of all selected fields can be justified by the bounded model.

Overall, the bounded Gutenberg-Richter adaptation is useful for testing whether observed largest earthquakes are statistically plausible. However, the results should be interpreted with caution because they depend on M_C, R_P, b-value, catalogue size and the chosen Mmax assumptions. The method should therefore be seen as a probabilistic justification tool, not as an exact prediction of future earthquakes.

11. Reference List

Beirlant, J., Kijko, A., Reynkens, T., & Einmahl, J. H. J. (2019). Estimating the maximum possible earthquake magnitude using extreme value methodology: The Groningen case. *Natural Hazards, 98*(3), 1091–1113. https://doi.org/10.1007/s11069-017-3162-2

Davis, S. D., & Frohlich, C. (1993). *Did (or will) fluid injection cause earthquakes? Criteria for a rational assessment*. *Seismological Research Letters, 64*(3–4), 207–224. https://doi.org/10.1785/gssrl.64.3-4.207

Dost, B., van Eck, T., & Haak, H. W. (2004). *Seismic hazard due to small-magnitude, shallow-source, induced earthquakes in the Netherlands*. *Netherlands Journal of Geosciences, 83*(2), 113–121.

Ellsworth, W. L. (2013). *Injection-induced earthquakes*. *Science, 341*(6142), 1225942. https://doi.org/10.1126/science.1225942

Foulger, G. R., Wilson, M. P., Gluyas, J. G., Julian, B. R., & Davies, R. J. (2018). *Global review of human-induced earthquakes*. *Earth-Science Reviews, 178*, 438–514. https://doi.org/10.1016/j.earscirev.2017.07.008

Government of the Netherlands. (2024). *Herstel na gaswinning Groningen en Noord-Drenthe*. Rijksoverheid.

Grigoratos, I., Schultz, R., van Ginkel, J., Gunatilake, T., & Wiemer, S. (2024). *Review of the Seismic Hazard and Risk Protocol for induced seismicity related to gas production from small gas fields in the Netherlands* (Version 4.2) [Technical report]. Swiss Seismological Service, ETH Zurich. https://doi.org/10.13140/RG.2.2.22667.50729

Keranen, K. M., Weingarten, M., Abers, G. A., Bekins, B. A., & Ge, S. (2014). *Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection*. *Science, 345*(6195), 448–451. https://doi.org/10.1126/science.1255802

Koninklijk Nederlands Meteorologisch Instituut. (z.d.). *Aardbevingscatalogus*. KNMI. Retrieved June 29, 2026, from https://www.knmi.nl/kennis-en-datacentrum/dataset/aardbevingscatalogus

Ministerie van Klimaat en Groene Groei. (2025). *Periodieke rapportage seismiciteit Groningen*. Open Overheid. https://open.overheid.nl/documenten/1154f06a-d172-4557-afe1-db18945849c5/file

Muntendam-Bos, A. G., Hoedeman, G., Polychronopoulou, K., Draganov, D., Weemstra, C., van der Zee, W., Bakker, R. R., & Roest, H. (2022). *An overview of induced seismicity in the Netherlands*. *Netherlands Journal of Geosciences, 101*, e1. https://doi.org/10.1017/njg.2021.14

Muntendam-Bos, A.G., Grobbe, N. Data-driven spatiotemporal assessment of the event-size distribution of the Groningen extraction-induced seismicity catalogue. Sci Rep 12, 10119 (2022). https://doi.org/10.1038/s41598-022-14451-z

Nederlandse Aardolie Maatschappij B.V. (NAM). (2007). *Winningsplan Zuidoost Drenthe LoCal*.

Nederlandse Aardolie Maatschappij B.V. (NAM). (2016). *Aanvraag instemming winningsplan Westerveld: Actualisering September 2016*.

Nederlandse Aardolie Maatschappij B.V. (NAM). (2022). Report on the Second Workshop on Mmax for Seismic Hazard and Risk Analysis in the Groningen Gas Field, MainRept., The Infinity Building, Amsterdam, The Netherlands, 30pp., available at https://nam-onderzoeksrapporten.data-app.nl/reports/download/groningen/en/77951661-552a-46bc-9f2e-f1580cd6abc3 Full report available at https://nam-onderzoeksrapporten.data-app.nl/reports/download/groningen/en/fe2d5640-7fbd-477e-8be4-125f66d2260d, 1222 pp

Nederlandse Aardolie Maatschappij. (2025). Rapportage seismiciteit Groningen gasveld: Kalenderjaar 2024. Open Overheid. Available at https://open.overheid.nl/documenten/1154f06a-d172-4557-afe1-db18945849c5/file

Parliamentary Committee of Inquiry. (2023). *Groningers before gas: Conclusions and recommendations*. House of Representatives of the States General.

Richter, G., Hainzl, S., Dahm, T., & Zöller, G. (2020). *Stress-based statistical modeling of the induced seismicity at the Groningen gas field, the Netherlands*. Environmental Earth *Sciences, 79*, 252.

Schultz, R., Ellsworth, W. L., & Beroza, G. C. (2022). *Statistical bounds on how induced seismicity stops*. *Scientific Reports, 12*, 1184. https://doi.org/10.1038/s41598-022-05216-9

TNO. (2022). *Deterministische hazard analyse voor geïnduceerde seismiciteit (DHAIS): Actualisatie 2021* (TNO-rapport 2021 R10977).

Van der Elst, N. J., Page, M. T., Weiser, D. A., Goebel, T. H. W., & Hosseini, S. M. (2016). *Induced earthquake magnitudes are as large as (statistically) expected*. *Journal of Geophysical Research: Solid Earth, 121*(6), 4575–4590. https://doi.org/10.1002/2016JB012818

Appendix

Appendix A

First, the script defines several constants. The number of random samples used to create a reference distribution is set to N_r the number of bootstrap simulations N_{mc}. The parameter f that states how much information the bootstrap of the available information keeps.
The script then reads the relevant columns from the dataset. The case ID, the ratio (R_s), and the largest magnitude (M_{lrg}). Any cases without a valid (R_s) value are removed from the analysis.

The next step is to assign weights to the data. Several weighting options are included in the code, but only one should be used at a time. Equal weighting gives every case the same importance. Alternatively, the data can be weighted by the number of events, by how close the largest magnitude is to a target value, or by selecting only specific cases such as "Seismically active gas fields". Furthermore there is an option to let certain ID's weight more than others. These weights influence which cases are considered more important in the final distribution.

The script then creates a random reference distribution. This is done by generating two random variables and calculating. This produces random values between 0 and 1, which are later used for comparison with the observed R_s values using a Kolmogorov-Smirnov test.

After this, the function `get_Rstats` is used to calculate the main statistics of the observed R_s data. It calculates the mean (R_s), the median (R_s), and the fitted beta distribution parameters. The beta distribution has two shape parameters, called alpha and beta. These parameters control the shape of the probability distribution, R_a.close to 1 means a higher alpha and a R_s close to 0 means a higher beta. A fit is made once using the chosen weights and once using equal weights for all data.

The most important part of the script is the bootstrap loop. In each bootstrap iteration, the function 'bootstrap_Decimate' selects a random weighted subset of approximately f. The mean, median, and beta distribution parameters are then recalculated for this subset. This is repeated N_{mc} times. The result is not just one estimate of the beta parameters, but a distribution of possible estimates. This gives an indication of the uncertainty and robustness of the fitted (R_s) distribution.

The script also performs a two-sample Kolmogorov-Smirnov test in each bootstrap run. This test compares the observed R_s values from the bootstrap sample with the random reference distribution. The resulting p-values are stored and later plotted. Here, low p-values indicate that the observed (R_s) distribution differs significantly from the random reference distribution.

Finally, the script creates several figures. The first figure shows a histogram of the observed R_s values together with the fitted beta distributions. It also shows the bootstrapped mean and median R_s. The second part of the same figure shows the bootstrapped alpha and beta parameters, including their scatter and separate histograms. The second figure shows the uncertainty in the mean and median R_s values based on the bootstrap simulations. The third figure shows the distribution of the Kolmogorov-Smirnov p-values.

Appendix B

First, the script defines several constants. The red-light magnitude, M_{red} and the completeness magnitude, M_c. The value of M_c is used as the lower magnitude threshold above which the earthquake catalogue is considered complete. Furthermore the input parameters are the number of Monte Carlo samples N, the magnitude bin size dM, and the N_s, representing the number of production-related earthquakes used in the calculation. Finally, the flag relating to which period of earthquakes is calculated: 'trail' for trailing seismicity, 'full' for all seismicity and 'red+trail' for the red-light and trailing seismicity.

The next step is to create the logic tree structure using the function 'MakeLogicTreeStructure'. The logic tree contains different branches that represent different possible interpretations of the seismicity. In this script, two main branches are used: a triggered branch and a driven branch. Within the driven branch, two sub-branches are included. These sub-branches represent different assumptions for the maximum magnitude distribution.

After the logic tree is created, the function 'LogicTreeSample' is used to randomly sample the input parameters from gas field specific values. For each of the simulations, a value is drawn for the b-value, R_p, and M_{max}. These values are sampled from parameter distributions related to respective fields.

The function 'dM_distribution' then uses the sampled b-values, R_p values, M_{max} values, and N_s to generate the expected distribution of trailing magnitudes. The output includes the simulated largest magnitudes M_{lrg}, the effective seismicity ratio R_p, and the number of trailing events $N_{t.}$

The script separates the results according to the logic-tree branches. Histogram counts are calculated for the triggered and driven branches separately. This makes it possible to see how much each branch contributes to the final magnitude distribution. In the figures, the total distribution is shown together with the branch-specific contributions, allowing the influence of the logic-tree assumptions to be inspected visually.

Finally, the script creates several figures. The first figure shows the normalised Gutenberg-Richter magnitude-frequency distribution. It includes the simulated magnitude distribution and the fitted Gutenberg-Richter trend. This figure is used to interpret the likelihood of different values of M_{lrg} . The second figure shows the sampled input distributions for M_{max}, the b-value, R_p, and the number of trailing events N_t. These histograms show how the logic tree and parameter uncertainty influence the model input. The third figure shows the final distribution of the simulated trailing magnitudes.

Appendix C

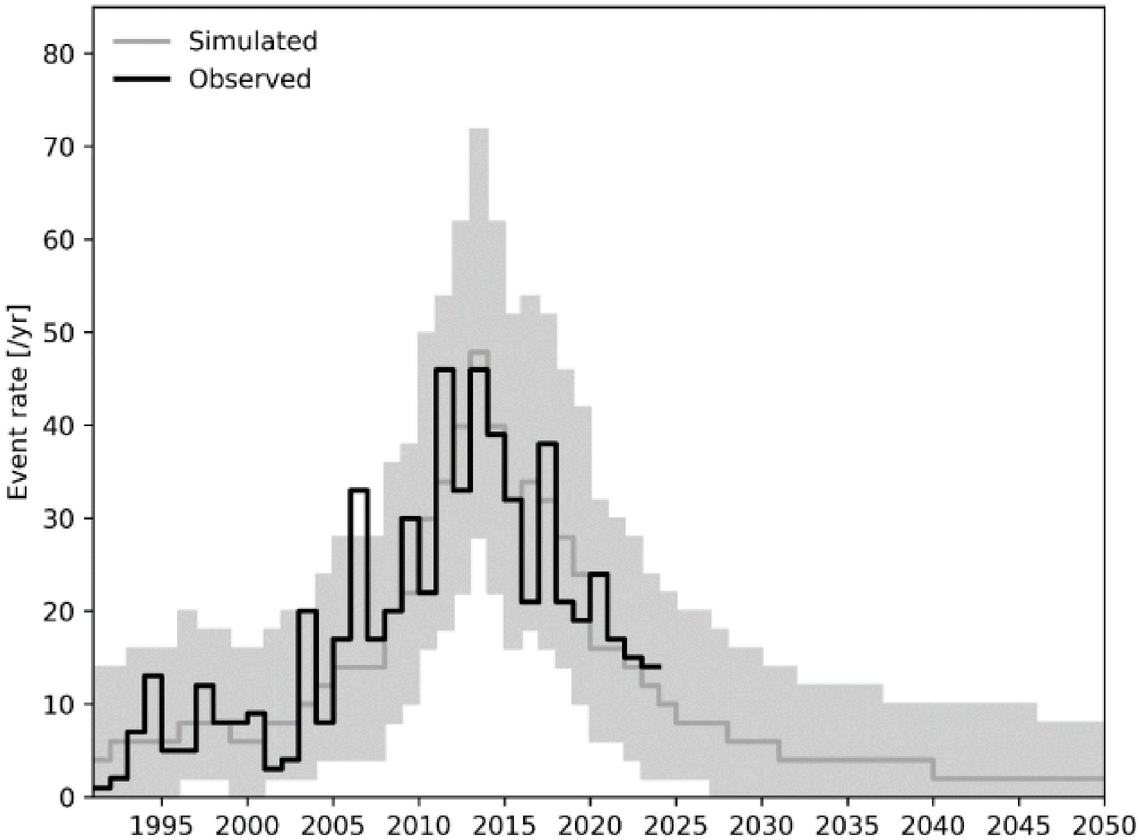

Figure 4-9 from the Nederlandse Aardolie Maatschappij (2025), simulations of the yearly number of earthquakes per year.