Maximum Magnitude of Induced Earthquakes in the Groningen Gas Field

ExxonMobil Upstream Research Company,

Nora DeDontney, Christine Gans, William A Burnett, Damian Burch, Jorge Garzon, Grant Gist, Sheng-Yuan Hsu, Suvrat Lele, Darren Pais, Pablo Sanz Rehermann, Kevin Searles, Bill Symington, Jelena Tomic, Martin Terrell and Adel Younan

Date July 2016

General Introduction

Early March 2016, 36 independent experts from around the world gathered in Amsterdam for a workshop to discuss the maximum magnitude of earthquakes in Groningen. The workshop resulted in a probability distribution of the maximum magnitudes of induced and triggered tectonic earthquakes in Groningen. The results of the workshop are recorded in the comprehensive Report on Mmax Expert Workshop (Ref. 1), which includes both the main conclusions of the expert panel and all individual contributions of the workshop participants to the discussion.

The current report provides background to one of these presentations held on the afternoon of the second day of the workshop. This report was prepared after the workshop and not available to the panel.

References

1. Report on Mmax Expert Workshop, Mmax panel chairman Kevin Coppersmith, June 2016

NAM

Title	**Maximum Magnitude of Induced Earthquakes in the Groningen Gas Field**		**Date**	July 2016
			Initiator	NAM
Autor(s)	**Nora DeDontney, Christine Gans, William A Burnett, Damian Burch, Jorge Garzon, Grant Gist, Sheng-Yuan Hsu, Suvrat Lele, Darren Pais, Pablo Sanz Rehermann, Kevin Searles, Bill Symington, Jelena Tomic, Martin Terrell and Adel Younan**	**Editors**		
Organisation	**ExxonMobil Upstream Research Company**	**Organisation**	NAM	
Place in the Study and Data Acquisition Plan	Study Theme: Seismological Model Comment: Early March 2016, 36 independent experts from around the world gathered in Amsterdam for a workshop to discuss the maximum magnitude of earthquakes in Groningen. The workshop resulted in a probability distribution of possible maximum magnitudes of induced and triggered tectonic earthquakes in Groningen. The results of the workshop are recorded in the comprehensive Report on Mmax Expert Workshop (Ref. 1), which includes both the main conclusions of the expert panel and all individual contributions of the workshop participants to the discussion. The current report provides background to one of these presentations held on the afternoon of the second day of the workshop. This report was not available to the panel.			
Directliy linked research	(1) Geology of the Carboniferous (2) Geomechanics of fault rupture (3) Sesimological Model			
Used data				
Associated organisation				
Assurance	ExxonMobil Upstream Research Company			

Maximum Magnitude of Induced Earthquakes in the Groningen Field

ExxonMobil Upstream Research Company

June 22, 2016

Nora DeDontney
Christine Gans

William A Burnett
Damian Burch
Jorge Garzon
Grant Gist
Sheng-Yuan Hsu
Suvrat Lele
Darren Pais
Pablo Sanz Rehermann
Kevin Searles
Bill Symington
Jelena Tomic
Martin Terrell
Adel Younan

Contents

0.0 Executive Summary 1

1.0 Introduction 3

1.1 Overview of Groningen seismicity 3

1.2 Geology of the Groningen field and the structure of the fault network 7

1.3 Definition of maximum magnitude 10

1.4 Framework of this study 12

2.0 Maximum magnitude constraints using the existing earthquake catalog 13

2.1 Statistical analysis of the earthquake catalog 14

2.1.1 Uncertainty in the statistical analysis 16

2.2 Extreme-value analysis 17

2.3 Conclusions drawn from the observed catalog 17

3.0 Confinement of rupture to the vicinity of the reservoir 18

3.1 Methods of Confinement 18

3.2 Criteria for determining if stress will inhibit rupture propagation 20

3.2.1 Fault friction during sliding 20

3.2.2 Rupture propagation into areas of low fault loading 22

3.2.3 Stress-drop-based criterion 22

3.3 Estimate of stress state on faults in the underlying Carboniferous 23

3.3.1 Fault dip 24

3.3.2 Vertical total stress 24

3.3.3 Minimum horizontal stress in the Carboniferous 25

3.3.4 Covariance between horizontal and vertical stresses 27

3.3.5 Pore pressure 27

3.4 Possibility of Carboniferous rupture for base case assumptions 28

3.5 Sensitivity of results and additional considerations 30

3.6 Summary: Probability that all events will be confined to the vicinity of the reservoir 32

4.0 Events confined to the vicinity of the reservoir 32

4.1 Definition of a confined event 32

4.2 Magnitude determination from earthquake scaling relations 33

4.3 Relating compaction to earthquake magnitude 36

4.4 Magnitude of earthquakes determined from a 3-D fault-based geomechanical model 37

4.5 Summary: Maximum magnitude of confined events 39
5.0 Triggered out-of-reservoir events 40
5.1 Fault-length-based correlations 40
5.2 Fault-area-based correlations 42
5.3 Representation in PSHA analyses 43
5.4 Summary: Maximum magnitude of unconfined (triggered) events 44
6.0 Conclusions 45
References 47
Appendix A: Determination of the stress state in the Carboniferous 49
A.1 Total vertical stress 49
A.2 Minimum horizontal stress in the Carboniferous 50
A.2.1 Initial minimum horizontal stress in the reservoir 50
A.2.2 Stress contrast between the reservoir and the Carboniferous 51
A.3 Pore pressure 54
Appendix B: Wave propagation simulation of the Huizinge event 55
Appendix C: Uncertainty in the triggered event maximum magnitude 58

0.0 Executive Summary

As of January 1, 2016, 271 events of magnitude 1.5 or larger have been recorded in the Groningen gas field, Netherlands, the largest onshore gas field in the world. The largest earthquake observed to date is the 2012, M_L 3.6, Huizinge event. The occurrence of this event and the hazard analysis that followed highlighted issues surrounding the magnitude of the largest possible earthquake in the Groningen field.

Upon examination of the geologic setting of the Groningen gas field and the observations of earthquakes to date, a fundamental question arises: Is it possible for an earthquake to rupture out of the reservoir interval? If the rupture is confined to the vicinity of the reservoir, then only a small portion of the pre-existing fault is capable of seismic slip and therefore the maximum magnitude is limited by that fault area. If the rupture is able to propagate out of the reservoir interval (and into the underlying Carboniferous formation), then a larger portion of the fault is able to coseismically slip and the maximum magnitude of an event could correspondingly be larger. This report attempts to provide support for establishing the probability that a rupture could propagate out of the reservoir interval and determining the range of possible maximum event magnitudes.

There is no record of historical seismicity in the area, which suggests that pre-production (and away from the perturbed reservoir zone), the faults were not critically stressed (not close to failure). However, the stress state in the reservoir has been altered due to depletion, bringing a subset of faults to a critical state and initiating events in the reservoir interval. Once started, an earthquake has the potential to propagate into an area that is not critically stressed (i.e. below the reservoir), and an analysis of the stress state can indicate how likely this is to occur. An analysis considering the field-wide variation and uncertainty of key parameters, finds that many ruptures cannot propagate out of the reservoir perturbed zone. Approximately 20% of scenarios could result in Carboniferous rupture, so the possibility of large magnitude events cannot be eliminated. The same conclusion is reached from analysis of the frequency-magnitude distribution of the recent catalog: While the catalog is more consistent with a low truncation magnitude (M_W 4-4.5), a high truncation magnitude cannot be ruled out.

If an earthquake is confined to the vicinity of the reservoir, then a M_W 4.5 upper bound should be considered. This upper bound is based on the statistics of the earthquake catalog and is consistent with multiple geomechanical and seismological considerations. An important assumption in the M_W 4.5 truncation magnitude is that ruptures in the vicinity of the reservoir should not extend much more than 350 m in the down-dip direction and would be limited to 2 km in the strike dimension. This assumption is based on observations of fault rupture aspect ratios observed for tectonic, dip-slip earthquakes.

To address the possible magnitude of a rupture out of the reservoir, relationships from tectonic earthquakes are considered. Earthquakes occur due to slip on faults of a variety of lengths; some faults are capable of hosting a M_W 5.5 while some are capable of larger events. In Groningen, there are a small number of faults that are capable of hosting a M_W 6.5 event, but most faults are of short enough lengths that an event larger than a M_W 5.5 is not possible. The relative proportion of small and large faults is used to assign the relative likelihoods of larger magnitude truncations. The high ratio of smaller

to larger faults in Groningen gives a low likelihood of large magnitude truncation values. The results are presented in a logic tree and this provides the probability that any given earthquake will be bounded by a certain maximum magnitude.

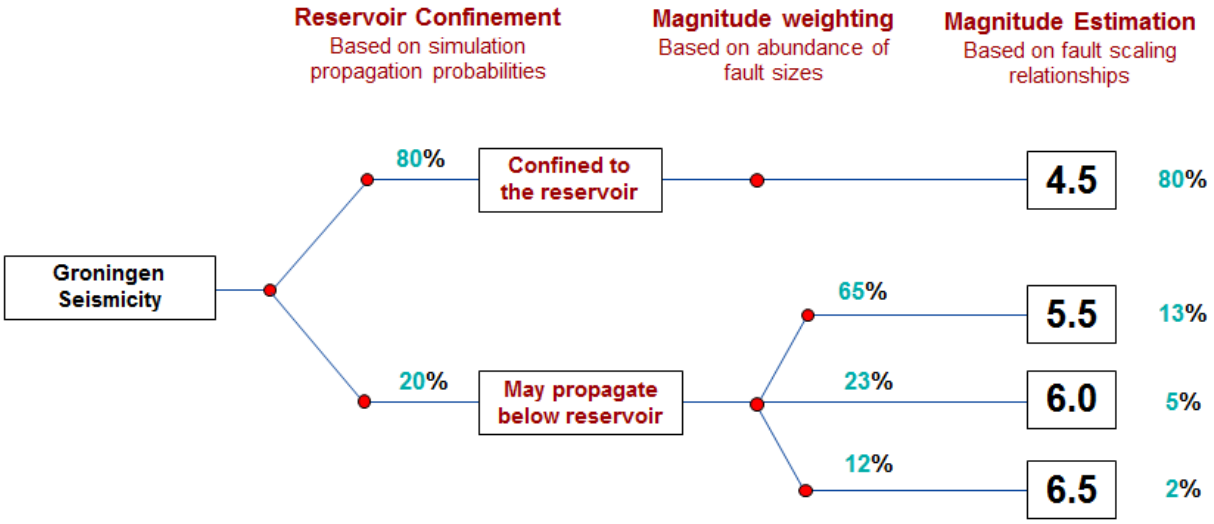

Figure E1. Logic tree representation of the probability that a given maximum magnitude should be applied to any given earthquake

The logic tree as shown assigns a 20% probability that earthquakes could rupture out of the vicinity of reservoir. This does not mean that 20% of the earthquakes will propagate out of the reservoir zone. Even if rupture out of the reservoir is possible, earthquake magnitudes still obey the Gutenberg-Richter power-law distribution, where many more events will be small magnitude events that nucleate in and are confined to the reservoir vicinity. For truncation magnitudes higher than M_W 4.5, it is expected that 1 in 1000 events would be larger than a M_W 4.5 (271 events have been observed to date).

1.0 Introduction

The first reports of seismic activity in the historically aseismic northeastern Netherlands occurred in the mid 1980's. Since the early 90's it was recognized that the earthquakes were induced by hydrocarbon extraction from the subsurface. There are many producing gas fields in the area that are associated with induced seismic events, but most of the events are in the vicinity of the Groningen field, the largest gas field in Europe, and the tenth largest in the world.

Gas from the Groningen field provides fuel for cooking and heating in many of the homes in the Netherlands. The Dutch infrastructure and appliances are designed for the low BTU Groningen gas and maintaining production from Groningen plays a role in maintaining Dutch energy security. The importance of this asset has led to a large effort to better characterize the seismic activity, understand its causes, and accurately assess the seismic hazard posed to the population in the area. One aspect of this effort is to have an accurate evaluation of the largest possible event size that could occur (the maximum magnitude).

Given the concern about the possibility of larger magnitude events, the existing data has been examined to determine the constraints on the maximum possible event size. This is a difficult task since it requires an extension of what is understood from an earthquake catalog to a description of events that have not yet occurred. At the request of the operator of the field, NAM, many institutions and organizations have recently tried to answer this question. This report documents the body of work performed by ExxonMobil (a shareholder of NAM) to address this topic and suggests a framework for considering the probability of large magnitude events.

1.1 Overview of Groningen seismicity

The Groningen gas field is located in a tectonically quiet area of little to no historic seismicity. By the 1980s, some 20 years after first gas production, felt earthquake events occurred in the area. By 1991, seismicity had increased and was understood to be a consequence of fault reactivation due to pressure depletion from gas production. The number of events by magnitude and year shows an increasing number of events per year and overall trend of increasing magnitude of events (figure 1). The largest magnitude event was a M_L 3.6 in August 2012 near Huizinge (Dost et al., 2013). The magnitude of the Huizinge event raised concerns about the potential hazard and risk for larger magnitude events.

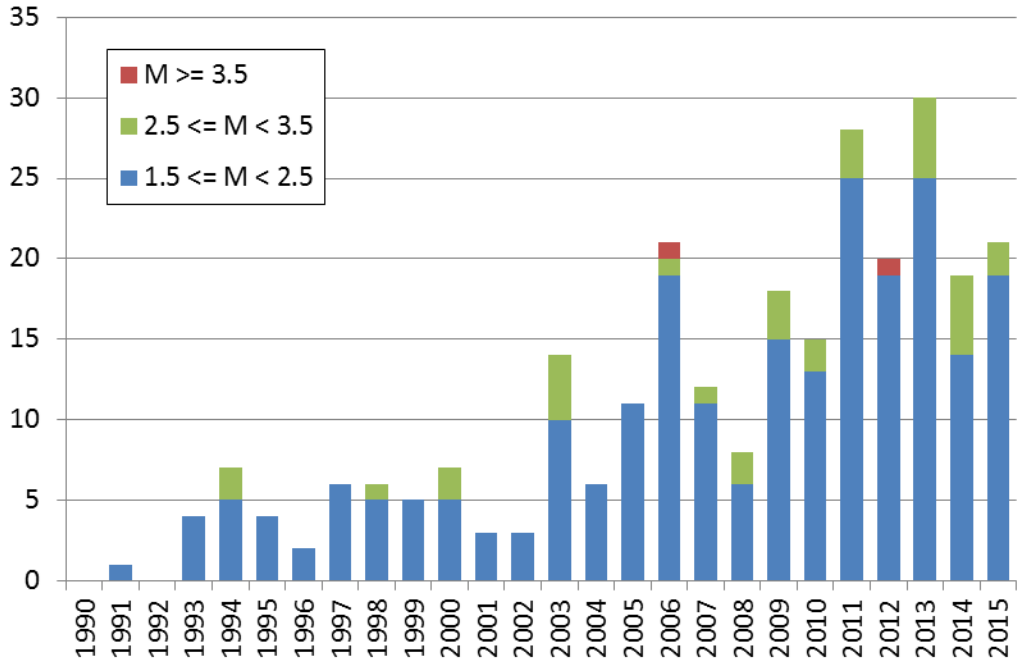

Figure 1. Number of events observed per year, categorized by magnitude (reported as local magnitude M_L).

As of January 1, 2016, 271 earthquakes of magnitude $M_L \geq 1.5$ have been observed. 12 of those events occurred before April of 1995, when the installation of new seismic stations established catalog completeness down to M_L 1.5, with local areas of completeness down to ~M_L 1.2 (KNMI 2016). The map view of earthquakes that have occurred is shown in figure 2 (field outline in thin light blue line). Most of the earthquakes are located in a large cluster towards the central-north end of the field, with two of the largest magnitude events observed to date (the M_L 3.6 (M_W 3.4) Huizinge event, and the M_L 3.5 Westeremden event) located in the northwest of this cluster. This area with larger magnitude and more frequent events is called the Loppersum region (named for the nearby town of Loppersum).

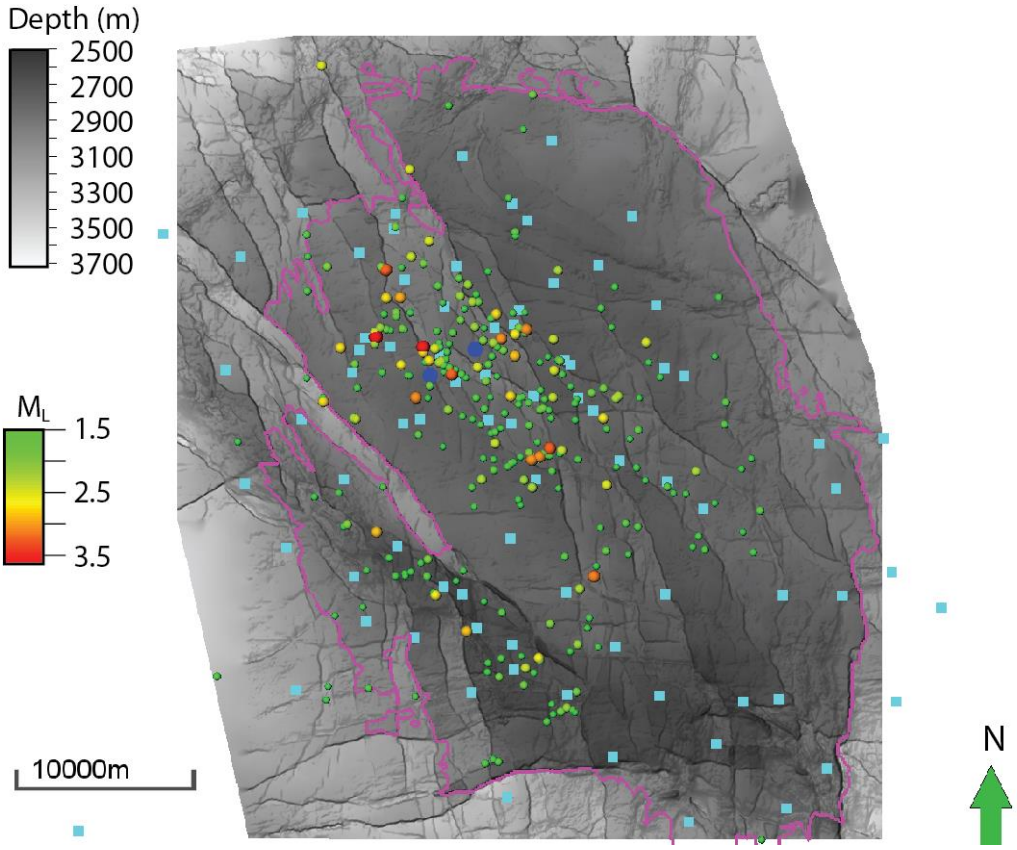

Figure 2. The current earthquake catalog for the Groningen field, through end of 2015. The field is outlined in pink and the depth map is the top of the Rotliegend reservoir. Current surface seismic stations are light blue squares and the downhole array locations are dark blue circles. Earthquake epicenters are circles and scaled/colored according to magnitude. Earthquake locations were obtained from the Royal Netherlands Meteorological Institute's website (KNMI, 2016). Lateral error in the location is estimated by KNMI to be ~1 km, while depth is assigned to be 3 km – within the reservoir.

In addition to a surface array of seismometers operated by the Royal Meteorological Institute of the Netherlands (KNMI), data also exists from NAM operated downhole arrays of geophones deployed in the reservoir interval. The downhole arrays are able to detect smaller magnitude events than the surface array is capable of, but only in the immediate vicinity of the boreholes, therefore the catalogs collected by these two arrays are different but overlap. Earthquake observations from both the surface array and downhole arrays provide constraints on a few key earthquake seismology parameters (figure 3):

- **Focal mechanisms** – Events are predominantly normal faulting with a small strike-slip component (figure 3a)

  There is a limited collection of derived focal mechanisms but the results show similar mechanisms. While there is no historical seismicity to compare to, normal faulting events are consistent with the expected deformation in a compacting reservoir and with the background stress field. Stress measurements in the area indicate that it is an extensional environment where the vertical stress is the maximum stress.

- **Earthquake hypocenters** – Earthquakes are nucleating in the reservoir

  KNMI sets the depth of all earthquakes to 3 km in the seismic inversion because the data is not sufficient to constrain depth from the surface array data. Data from the downhole arrays of geophones, however, locates all but 6 of 563 events in the reservoir interval (figure 3b). It is believed that the 6 events located out of the reservoir interval were mis-locations due to sparse data (geophone failures) with large uncertainties (the errors on these events are 600-2000 m in the vertical dimension).

- **Static stress drop** – Stress drops range from 0.1-10 MPa

  Bommer et al., (2015a) performed an analysis of the frequency content of 18 Groningen events and used the observed corner frequency to solve for the static stress drop (figure 3c). They found that the static stress drop generally ranges from 0.1 -10 MPa, with the best fit single value solution of 3 MPa. Stress drops from tectonic earthquakes are generally in the range of 0.1-100 MPa so these observations are similar to tectonic observations, if on the low side.

Both local magnitude, M_L, and moment magnitude, M_W, are referenced in this report. All event magnitudes are calculated as local magnitude and a handful of events have had a moment magnitude calculation. For events larger than M_L 2.5, $M_L \approx M_W + 0.2$ (Bommer et al., 2016). Since this is a discrete and not a continuous adjustment, a catalog distortion occurs if moment magnitudes are used. Because of this, local magnitudes are used for the G-R frequency magnitude analysis, and moment magnitudes are used elsewhere with the knowledge that they are not equivalent. All final statements about maximum magnitude are moment magnitudes.

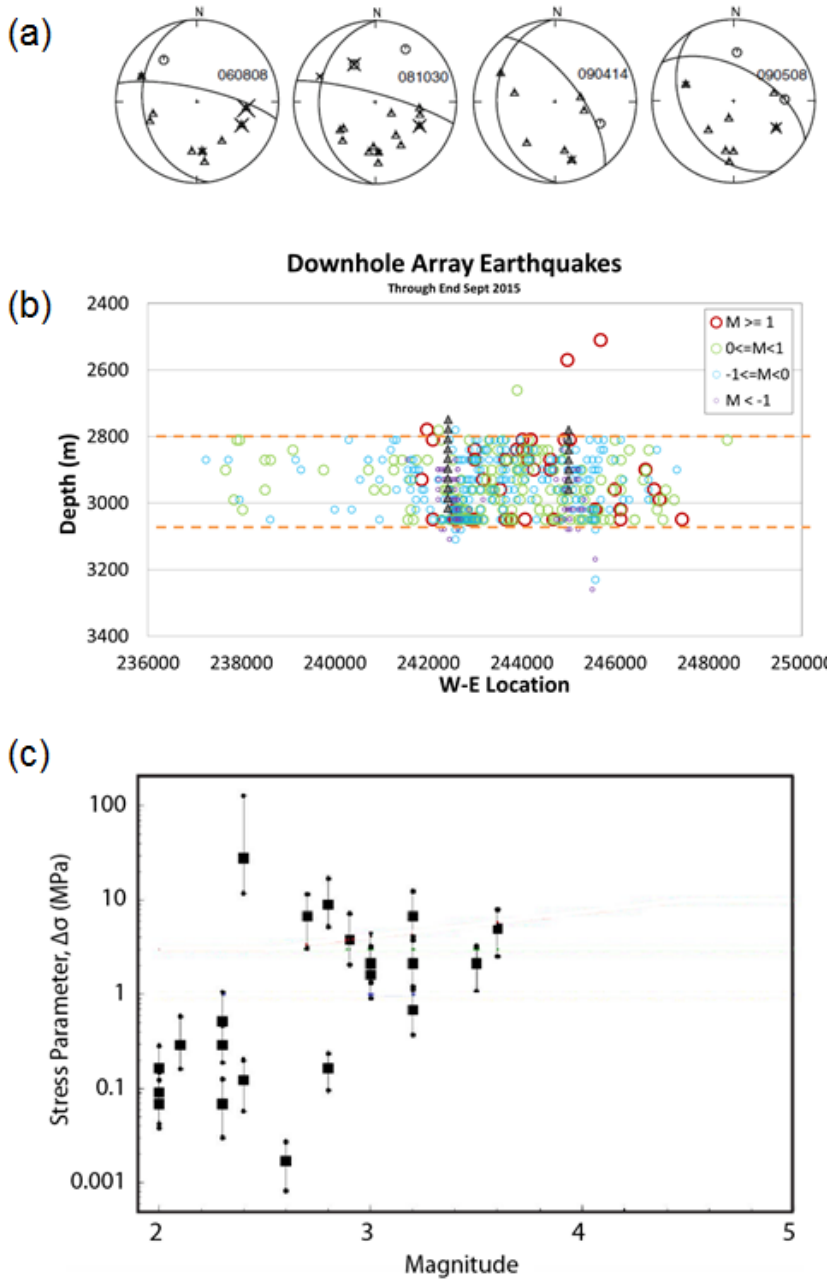

Figure 3. (a) Focal mechanisms indicate predominantly normal faulting with a small component of strike-slip motion (Kraajipoel & Dost, 2013). (b) Location of events from the downhole array. Earthquakes are nucleating in the reservoir interval. The few events located out of the reservoir have large location errors and could easily fall within the reservoir interval (data posted to NAM platform) (c) Stress drops calculated from observed earthquake spectra range from ~0.1-10 MPa (modified from Bommer et al, 2015a)

1.2 Geology of the Groningen field and the structure of the fault network

Earthquakes are related to the depletion of the Rotliegend sandstone at ~3 km depth. An overview of the Rotliegend geology and the fault network is necessary to understand the current characteristics and potential future behavior of the induced seismicity.

The Groningen gas field is depleting the sandstone of the Rotliegend formation which extends into Germany, Poland and the North Sea. The Groningen high is located on the margin of the Southern Permian Basin (basin center to the North) in an area of aeolian dunes (although not preserved), alluvial fans, and wadies (braided rivers) heading north to a large saline lake. The reservoir is predominantly sand, but there are also 4 interbedded shale intervals that can be seen field wide, corresponding to high-stands. The deposition of this sand facies was followed by the deposition of the Ten Boer claystone (~60 m thick) during a long duration high-stand and then the Zechstein salt (the seal for the hydrocarbon system) as the saline lake dried up. The Rotliegend thickens to the north of the field, toward the center of the Southern Permian basin. The thickness of the interval varies from 125 m in the southeast to 300 m in the northwest. 210 meters is a representative thickness of the sand in the Loppersum area of highest seismicity, and 270 m if the Ten Boer claystone is included. Below the Rotliegend is a thick (~3 km) interval of Carboniferous shales with interbedded sand bodies. The interface between Carboniferous and Rotliegend is the Saalian unconformity (Doornenbal, 2010).

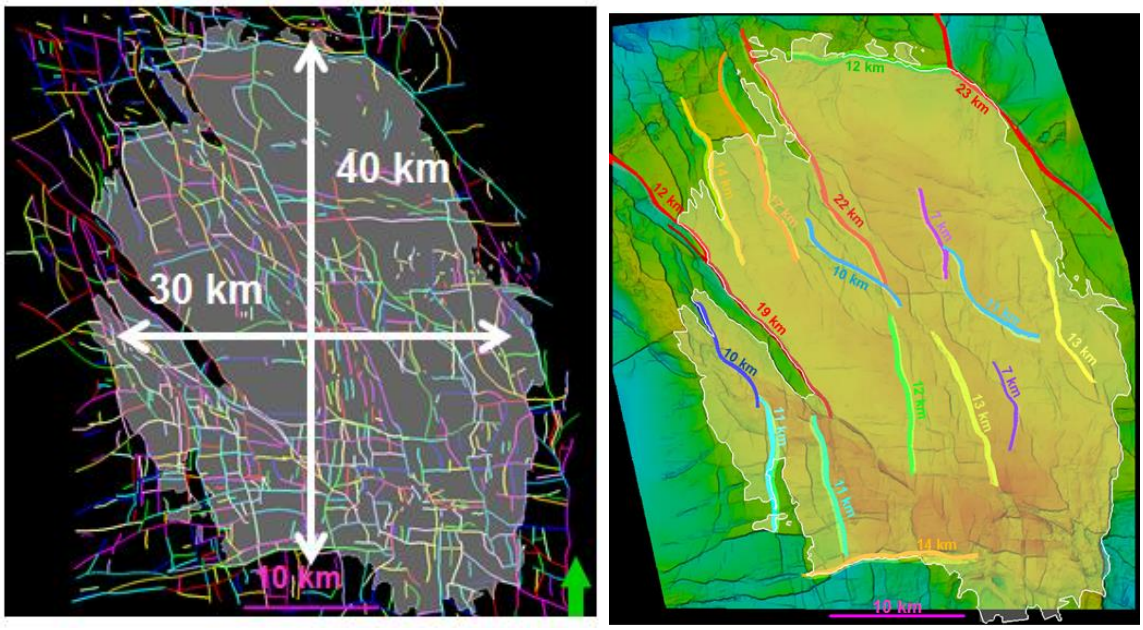

Figure 4. (a) Extent of the fault network mapped in the Groningen field (colored randomly by individual fault mapping) (b) Contour map of the depth of the top of the Rotliegend reservoir. Field outline shown in transparency with the longest faults in the field highlighted and labeled with the mapped length.

Many faults cutting through the Rotliegend reservoir are easily identifiable in seismic data due to their large throw (10s of meters). NAM has done extensive work interpreting the seismic data, identifying faults, and determining reservoir offset (i.e. the throw of the faults). The fault offset observed in the Rotliegend reservoir is a minimum bound on the amount of slip that has occurred over the life of the fault. The Carboniferous rocks below the reservoir were faulted and deformed during a period of wrench faulting prior to Rotliegendes deposition and this resulted in both strike-slip and dip-slip components of offset at depth. The Carboniferous was then exposed and eroded just prior to the Rotliegend deposition. Later, during a time of rifting, some of the faults in the Carboniferous fault

network were re-activated in predominantly normal motion. Figure 4a shows the network of faults that were re-activated to offset the Rotliegend.

Out of the more than 1000 mapped faults, only 43 are longer than 7 km in the lateral dimension. There are no continuous fault segments that span the entire length of the field. The longest fault segments are shown in figure 4b, with no individual segment reaching 25 km in length. The throw on the faults varies throughout the field and can be as high as 640 m on the reservoir bounding faults (figure 5). While the offset can be large, 80 m is a representative value of maximum fault throw for a given fault. This makes 250-300 m (~210 m sand thickness + ~80 m throw) a representative value for the reservoir sand thickness plus fault offset.

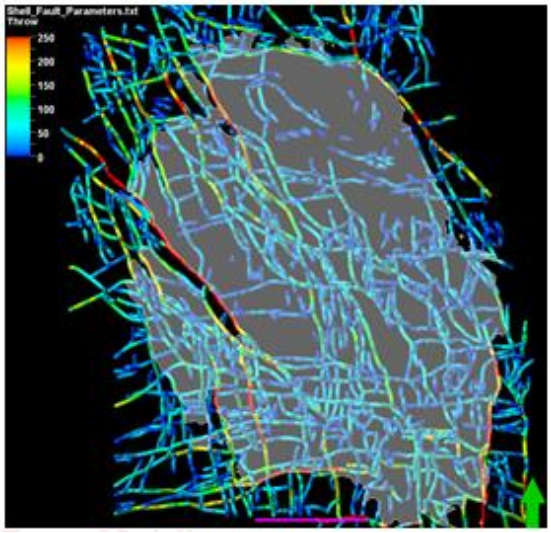

Figure 5. Fault throw (m) of all mapped faults used in the reservoir model. The faults with largest throw occur on the edges of the field (bounding faults).

The depth extent of the faults is also important for understanding the seismic hazard. A 3D seismic survey over Groningen field provides adequate imaging down to 6 km depth, but imaging below that depth is of low quality. Figure 6 shows an interpreted seismic section in the north of the field and crossing one of the largest faults (labeled M1 in figure 6).

Salt overlies and seals the hydrocarbon system and while it is non-uniform in thickness, it averages ~1 km over the field. Faults do not extend through the salt. Salt is buoyant compared to the overlying deposits and this buoyancy results in deformation of the overburden due to salt tectonics. Over geologic time the salt is mobile and can deform in response to movement of the faults or to differential loading. This can result in faulting in the shallow strata (e.g. fault from 0-2 km depth on far right side of figure 6) but this deformation is not the result of a through-going fault extending from the reservoir interval (~3 km depth) up into the shallow strata.

All faults are truncated by the salt at their up-dip end but the down-dip extent is more difficult to constrain. Many of the smaller faults have a depth extent of < 3 km, but the M1 fault has a large

amount of offset at 6 km depth (see offset of red horizon in figure 6), so likely the fault continues to a greater unknown depth.

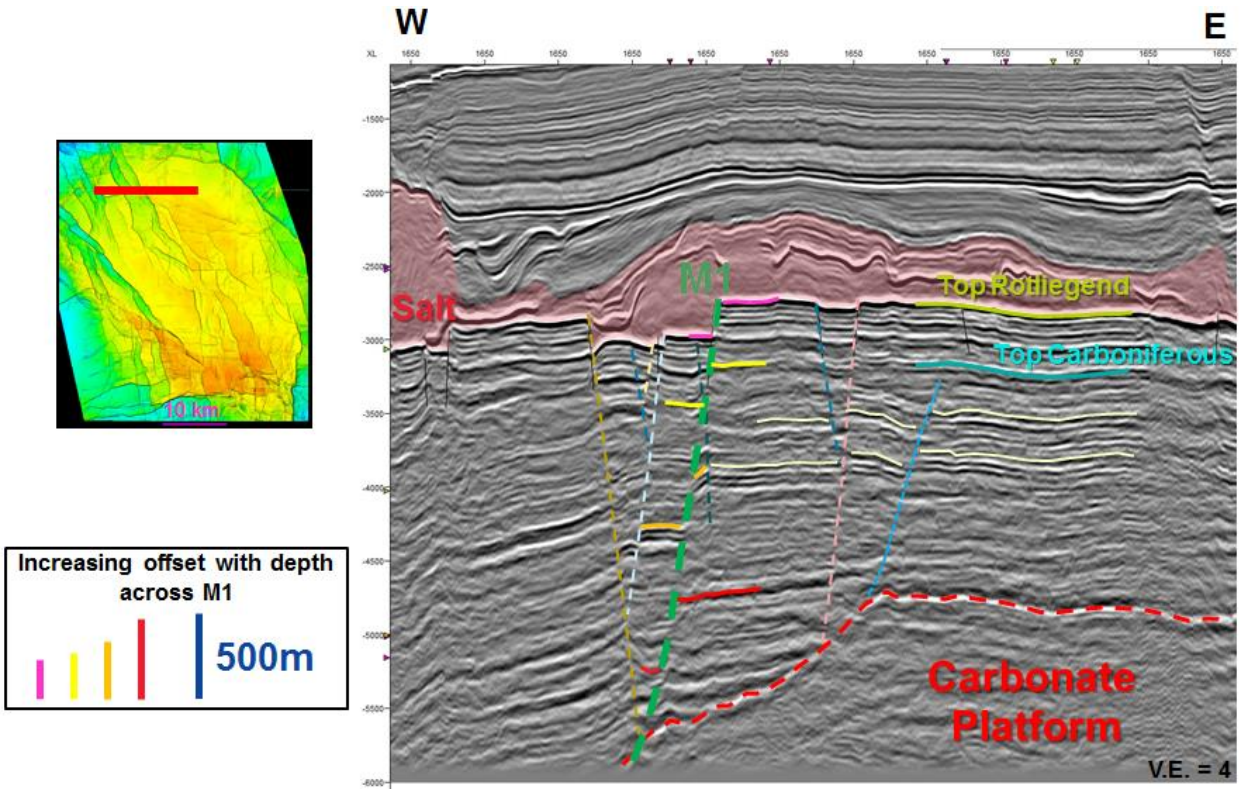

Figure 6. Interpreted seismic line across one of the larger structures in the Groningen field. The salt body is shaded in red. Some faults are 1-3 km tall but others (Green fault labeled M1) extend off of the seismic data to an unconstrained depth. Panel in lower left shows the amount of vertical offset of the pink, yellow, orange and red horizons marked in the seismic cross-section.

1.3 Definition of maximum magnitude

Given the finite number of earthquakes that will occur during the life of the Groningen field, it is very unlikely that the largest possible earthquake that could occur will occur, therefore the largest *expected* magnitude of an earthquake is smaller than the largest *possible* earthquake magnitude. Probabilistic seismic hazard assessment (PSHA) is an accepted framework to address the hazard posed by earthquakes, and the methodology is designed to handle very low probability events (i.e. large earthquakes) that can have substantial consequences for a given region. A standard component of the PSHA framework is the maximum possible earthquake magnitude (not the maximum event size that is likely to be observed), so here the discussion of maximum magnitude is limited to the maximum possible (limited by the fault rupture area), not the maximum probable (affected by the number of events assumed to occur).

The distribution of earthquake magnitudes follows a power law known as the Gutenberg-Richter (G-R) relationship. This relationship was first observed for naturally occurring, tectonic earthquakes, and has subsequently been shown to hold for induced earthquakes in a variety of settings (hydraulic fracturing, wastewater injection, reservoir depletion, and geothermal injection). Figure 7 shows this relationship as

the number of events larger than a specified magnitude that are observed. A lower bound threshold magnitude is specified so only the number of earthquakes above this magnitude is considered. In the case of Groningen the lower threshold is M_L 1.5, because the local seismic network is capable of detecting all M_L 1.5 events and higher, making the catalog "complete" down to this magnitude. In addition to this minimum magnitude, a maximum (truncation) magnitude is also specified based on local geologic conditions. There are three key components to the truncated Gutenberg-Richter relationship:

1. a-value – the number of events larger than the minimum threshold magnitude that are observed over some period of time. In tectonic settings this is generally thought to be constant in time but this can be time varying. Rather than a cumulative number, a frequency can also be specified as an a-value.
2. b-value – the slope of the line describes the relative number of large versus small earthquakes. Large b-values indicate a more steeply dipping probability distribution and therefore fewer large earthquakes. A low b-value means that more large earthquakes are expected.
3. M_{max} – the theoretical maximum possible event size that could be observed (many functional forms can be used to specify how this truncation is implemented; here the exponential form is used).

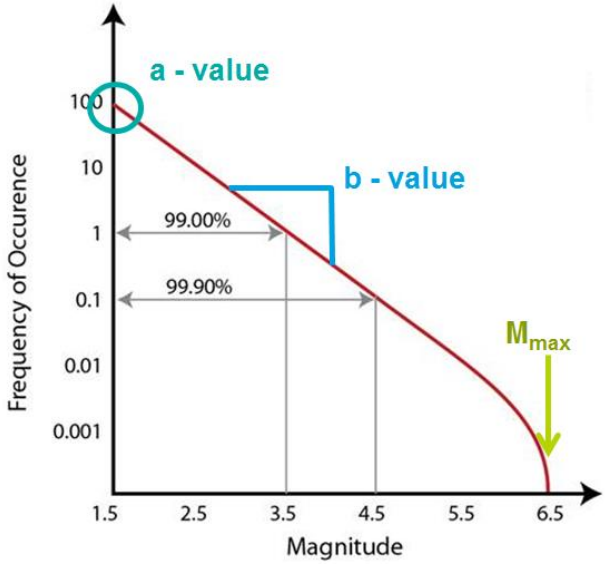

Figure 7. Example of a Gutenberg-Richter relationship and its three key components (a-value, b-value and truncation magnitude)

The straight portion of the G-R relation is described by the equation:

$$\log_{10} N = a - bM$$

where N is the number of events greater than or equal to a given magnitude M. It is standard practice to truncate the G-R relationship with a smooth transition at a large magnitude (as seen in figure 7). The probability density function that describes the exponentially truncated G-R equation is:

$$f_M(M) = \frac{b\ ln(10)10^{-b(M-M_{min})}}{1-10^{-b(M_{max}-M_{min})}}$$

where M_{min} is the minimum threshold magnitude and M_{max} is the upper bound truncation magnitude. For induced seismicity within the United States, the United States Geological Survey (USGS) truncates M_{max} to 6.0 – 6.5 (Petersen et al., 2016). For tectonic earthquakes, M_{max} can be much larger (Petersen et al., 2014 and Petersen et al., 2016).

With a minimum threshold magnitude of 1.5, and a standard b-value of 1.0, 99% of the magnitudes of all events are smaller than 3.5, 99.9% are smaller than 4.5, 99.99% are smaller than 5.5 and so on. Statistically, for 100 M 1.5 events in a given (complete) catalog, there will be 10 M 2.5 earthquakes, and 1 M 3.5. These probabilities are insensitive to the truncation magnitude specified unless the magnitude under consideration is within 1-2 magnitude units of the truncation specified. Generally, ~1000+ events are required to have a statistically complete catalog (Stork et al., 2014), and in comparison Groningen has 271 events detected above the minimum threshold magnitude.

1.4 Framework of this study

Upon examination of the geologic setting in the Groningen field, and the observations of earthquakes to date, a fundamental question arises: Is it possible for an earthquake to propagate out of the reservoir interval? (figure 8)

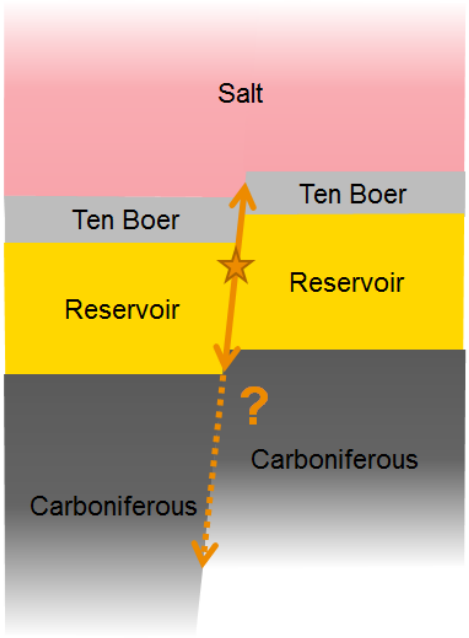

Figure 8. Schematic representation of an earthquake nucleating in the reservoir (orange star) and propagating across the reservoir zone (solid orange line). While fault ruptures cannot propagate through the salt, an important question is whether the rupture could propagate deeper, into the Carboniferous (dashed orange line).

If the rupture is confined to the vicinity of the reservoir, then only a small portion of the pre-existing fault is capable of seismic slip and the maximum magnitude is limited by that fault area. If the rupture is able to propagate out of the reservoir interval (and into the underlying Carboniferous formation), then a much larger portion of the fault could coseismically slip and the maximum magnitude of an event would correspondingly be larger. This report provides a criterion to determine the probability that a rupture could propagate out of the reservoir interval and determines the maximum event magnitudes for earthquakes confined and unconfined to the vicinity of the reservoir.

There is a large amount of uncertainty regarding the details of the subsurface, so a logic tree approach is implemented here to account for the epistemic uncertainty that surrounds the issue of maximum magnitude. There is a probability that the maximum possible magnitude of events is ~ M_W 4.5 and there is a probability that the maximum event size is larger than M_W 4.5. This study attempts to provide support for establishing those probabilities and determining the range of possible maximum event magnitudes.

This document is structured as follows: In section 2, an analysis of the existing seismic catalog is performed to determine if an indication of the truncation magnitude can be observed; however, it is found that the existing catalog is too small to be statistically conclusive. It is possible that earthquakes are confined to the vicinity of the reservoir, so in section 3 methods of confinement are considered and an analysis of the stress state outside the reservoir is performed. This analysis leads to the conclusion that there is an ~80% probability that conditions are not favorable for earthquake propagation and growth outside of the reservoir zone. In section 4, multiple lines of reasoning are presented to set a maximum magnitude of M 4.5 for earthquakes confined to the vicinity of the reservoir. In section 5, the potential magnitude of earthquakes that propagate out of the vicinity of the reservoir is considered. Given the sizes of the faults, most earthquakes will occur on faults that are not capable of the field-wide maximum magnitude. In section 6, a few concluding remarks are made as well as suggestions for further study.

2.0 Maximum magnitude constraints using the existing earthquake catalog

An important source of information about what is seismologically possible in the Groningen field is the catalog of earthquakes that have been observed. The largest event observed to date was a $M_L = 3.6$ ($M_w = 3.4$), therefore the maximum magnitude must be larger than M_L 3.6. The Groningen-based Gutenberg-Richter plot, figure 9, appears to follow a distribution with a truncation magnitude around M_L 4.0. This could be an artifact of the small size of the catalog and short period of observation. In tectonic settings it is common to initially observe fewer large magnitude events than predicted by the G-R relationship because it is difficult to observe all the possible sizes of earthquakes on a fault (or in a fault network) within the limited timespan of human observation. Often with more observation time, large events occur and this apparent deficit disappears.

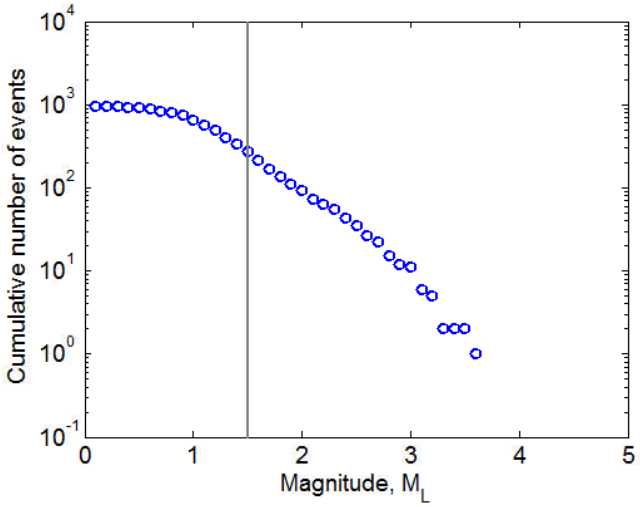

Figure 9. The Groningen earthquake catalog follows the Gutenberg-Richter frequency-magnitude relationship. Catalog contains 271 events of magnitude $M_L \geq 1.5$ (vertical gray line). The catalog appears to be following a G-R distribution with a truncation around ML 4.0, but this could be an artifact of the short duration of observation and the limited sample size.

The catalog that has been observed is just one of many possible earthquake catalogs that could be the result of sampling from a G-R distribution with a given b-value and Mmax. If 271 earthquake magnitudes ($M_L \geq 1.5$) are randomly sampled from a G-R distribution, and this is done twice, the catalogs will not be identical. Due to the random nature of the process, the largest magnitude event simulated in the two catalogs will not be identical. By specifying the b-value and Mmax, it is possible to calculate the probability that any magnitude, M, would be the maximum observed out of 271 events. One can think of this probability as the result of a Monte-Carlo simulation – if 50,000 earthquake catalogs of 271 earthquakes are generated from an assumed distribution, what percentage of the catalogs have their largest event equal to a given magnitude M? From this information it is possible to address several important questions:

- What is the most likely maximum observed magnitude? Is it consistent with our observation of a M_L 3.6 as the largest event?
- What is the probability that an event larger than the M 3.6 Huizinge should have been observed by now?
- Can the observation of a maximum of M_L 3.6 help inform a decision about the maximum magnitude?

To answer these questions a G-R relation must be assumed. The choice of b-value is very important, as it determines whether large earthquakes are more or less likely to occur. A low b-value, for example, would suggest that larger magnitude events should occur more often. For the purposes of answering the above questions, an analysis based on a constant, uniform b-value was performed.

2.1 Statistical analysis of the earthquake catalog

The b-value fit is made only to events that occurred after April 1995, the start of catalog completeness down to M_L 1.5, which reduces the number of events to 259. Based on the current Groningen catalog,

the Gutenberg-Richter distribution is best-fit with $b = 0.96 \pm 0.06$ (95% confidence bounds). Note that this b-value is based on the local magnitude, not moment magnitude, and the $M_L 1.5$ bin extends from 1.45-1.55 so the catalog is complete down to M_L 1.45 in the b-value maximum likelihood calculation.

For the best fit b-value of 0.96 and an assumed maximum magnitude of 6.5, the probability that any magnitude M will be the maximum observed for 271 events is shown in figure 10a. The b-value fit was made on the "complete" set of data (259 events) because an incomplete catalog that does not capture all small events will result in an erroneous calculation of the b-value. That b-value was then used as the accurate representation of the distribution that describes all observed events (271 events), including those that occurred prior to 1995. The maximum observed event to date is M_L 3.6 (red bar in figure 10) and a M_L 4.0 is the most likely observed maximum magnitude. While M_L 3.6 is not the most likely observed maximum magnitude, there is a 5% probability that a M_L 3.6 event is the largest event that should have been observed to date. This non-zero probability indicates that the observed catalog could be the result of a high (M_L 6.5) truncation.

The sum of all the probabilities for magnitudes larger than M_L 3.6 (3.65 due to the binning of reported magnitudes) is the probability that an event larger than the Huizinge event should have been observed by now (this is the same as the percentage of catalog realizations that have a largest event larger than Huizinge). If this value was extremely high or low (e.g. 99% or 1%), it would indicate that our observed catalog realization was not consistent with the underlying assumptions. However, for the scenario shown in figure 10a (88%), the probability is not high enough to rule out the assumption of a M_L 6.5 truncation.

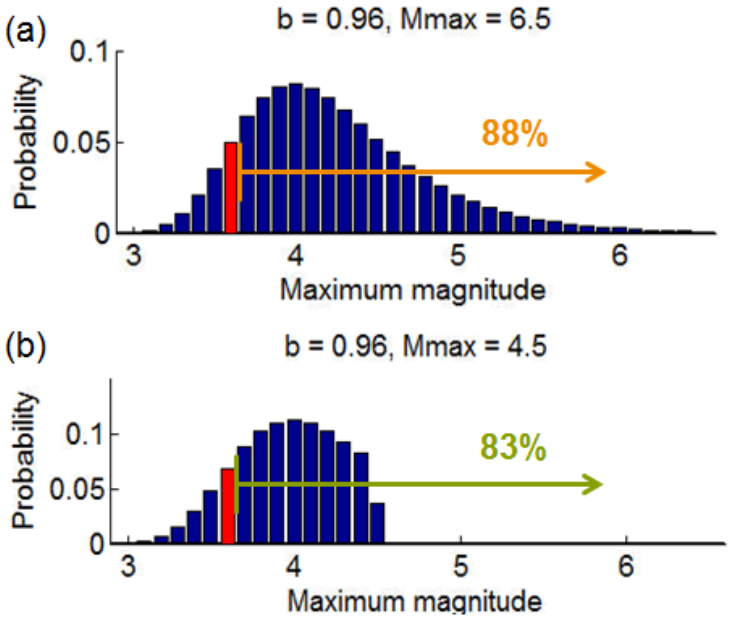

Figure 10. For the assumption of a b-value (0.96), a maximum magnitude ((a) 6.5 (b) 4.5) and a given number of events (271), there is a probability that any given magnitude will be the largest event observed in the catalog. The largest event to date (3.6) is marked by the red bar. The sum of all magnitudes larger than 3.6 is the probability that an event larger than 3.6 should have been observed by now (88% for Mmax = 6.5 and 83% for Mmax = 4.5).

If the assumed truncation is changed to a lower value, the relative probabilities (i.e. relative heights of the bars) do not change, but the probabilities above that specified truncation value are zero (see figure 10 for a comparison of two truncation magnitudes considered). The relative probabilities only change if the b-value assumption is modified. Higher b-values mean that large events are less likely, so the observation of a maximum magnitude to date of M_L 3.6 would be more consistent with the assumption of a higher b-value.

The probability that an event larger than M_L 3.6 should have occurred by now depends on both the b-value and the maximum magnitude assumed. Table 1 presents this probability for a range of assumptions. High probabilities indicate that a larger event likely should have been seen by now. Higher truncation magnitudes have higher probabilities and are therefore less consistent with the data. These probabilities are not high enough to rule out a high truncation magnitude, but they do indicate that the catalog observed is more likely to have come from a distribution with maximum magnitude truncation lower than 6.5.

Table 1. Probability of observing an event > M_L 3.6 for different b-value and truncation magnitude assumptions using a catalog of 271 events (bf – denotes the best fit value).

	Truncation Magnitude			
b value	4.0	4.5	5.5	6.5
0.85	84%	95%	97%	97%
0.96 (bf)	68%	83%	87%	88%
1.07	50%	65%	70%	70%

2.1.1 Uncertainty in the statistical analysis

For the probability calculation, the total number of events is an important number. The more events observed without an event larger than a M_L 3.6 occurring, the higher the probability that a low truncation magnitude is applicable. The probabilities presented in table 1 are lower bounds because the catalog is not complete prior to 1995, so there were likely more earthquakes that were not recorded. In addition to unrecorded Groningen events, there were also 78 events observed in adjacent Rotliegend gas fields (the producing reservoir interval in Groningen). Including these observed events that contain nothing larger than M_L 3.6 makes the catalog less consistent with the assumption of a large truncation magnitude.

There is some evidence to support the idea that the b-value varies in space. The data suggests that the Loppersum area (or the high compaction area in general) has a lower b-value, meaning most of the earthquakes occur in the low b-value region (as low as 0.8). If the majority of earthquakes occur in a low b-value area, it means that there is a very high probability that an event larger than M_L 3.6 should have been observed by now if the distribution was truncated at a high maximum magnitude. If low b-values

are appropriate for most of the earthquakes, then the top row of the table is more representative, and the observed catalog starts to become more and more inconsistent with a high magnitude truncation.

2.2 Extreme-value analysis

In addition to the G-R analysis, an extreme value analysis can be performed. For each year there is a maximum magnitude observed. The distribution of the yearly observed maximum magnitude can be fit assuming several forms of probability distributions. Once this fit to the data is made, the 99th percentile value can be used to determine the maximum that should be observed in any year (table 2). This analysis shows that the catalog is consistent with a low truncation magnitude in the M 4-4.5 range, rather than a high truncation (~M 6.5).

Table 2. An extreme value analysis evaluates truncation magnitudes at the 99th percentile for different assumptions of distribution shapes.

Distribution	Mag
Weibull	3.8
Largest Extreme Value	4.6
Gamma	4.1

2.3 Conclusions drawn from the observed catalog

With this analysis it is possible to answer the posed questions

- What is the most likely maximum observed magnitude?

  A $\boldsymbol{M_L\ 4.0}$ has the highest probability of being the observed maximum magnitude, regardless of truncation magnitude, but the maximum observed could range from $\boldsymbol{M_L\ 3.1}$ – Mmax. This makes a $\boldsymbol{M_L\ 3.6}$ a probable maximum observed magnitude to date, even if a high truncation magnitude (such as 6.5) is assumed.

- What is the probability that an event larger than the M 3.6 Huizinge should have been observed by now?

  Distributions with high truncation magnitudes have a higher probability that a larger event should have been seen by now. If the catalog is sampling from a high truncation magnitude distribution there is a high probability that a larger event should have been seen by now (p > 87% for a truncation of $\boldsymbol{M_L\ 5.5}$ +).

- Can the observation of a maximum of $M_L\ 3.6$ help inform a decision on maximum magnitude?

  The observed catalog is more consistent with a low truncation magnitude (M 4.5) than a high truncation magnitude (M 6.5) but the observed catalog cannot definitively constrain the maximum magnitude

3.0 Confinement of rupture to the vicinity of the reservoir

If earthquakes are confined to the vicinity of the reservoir then very large magnitude events are not possible. A limited fault area exists in the reservoir, so the maximum size of a confined rupture must be smaller than if the entire fault area outside of the reservoir is also able to slip. The earthquake hypocenter locations (figure 3b) indicate that earthquakes are only initiating in the reservoir; therefore earthquakes starting outside of the reservoir interval are not considered here and if such events are observed in the future, a re-visit of this topic would be warranted.

To evaluate the likelihood that earthquakes cannot access the full fault area, a method for confining the rupture to the vicinity of the reservoir must exist and be evaluated. An unfavorable stress state in the Carboniferous is the most likely candidate for confinement. Ruptures can propagate through areas of the fault that are not close to failure, but there must be a sufficient amount of shear stress to sustain the rupture growth. In the following sections, a stress-drop-based criterion for sufficiency is discussed and a summary of the analysis of the stress state in the Carboniferous is presented (more detail is presented in Appendix A).

There is heterogeneity in the state of stress throughout the field, along with varying fault dips, so one fault and its neighbor can have differences in stress loading. In addition to this heterogeneity, there is also uncertainty – uncertainty in a measured quantity (e.g. a stress value) or uncertainty due to incomplete information. Because of this uncertainty and variation, a Monte-Carlo approach is used to assess the likelihood of rupture confinement. Each scenario sampled from a variety of input distributions results either in rupture confinement to the vicinity of the reservoir or the possibility of rupture away from the reservoir. Throughout the following section histograms will be presented that are the basis of the Monte-Carlo inputs. These input distributions are sampled to generate scenarios that represent the stress and failure state of the fault. The percentage of scenarios that have a potential stress drop determines the likelihood that a given earthquake could propagate out of the reservoir into the underlying Carboniferous. Sensitivity to each of the input parameter distributions are also examined by considering multiple input distributions (a base case, a low side and a high side estimate are made). From this analysis it is found that ~80% of the time the stress drop sufficiency criteria is not satisfied indicating that ~80% of the time, stress conditions are unfavorable and ruptures must be confined to the vicinity of the reservoir.

3.1 Methods of Confinement

For an earthquake to be confined to an area there must be a mechanism to stop the rupture propagation. Generally, ruptures stop when they run out of energy or when they reach the end of a fault. They are thought to stop by one of three mechanisms: (1) geometric complexities (2) changes in fault properties, or (3) stress barriers. Regarding (1), the faults are relatively planar and continuous in the dip direction, so geometric complexities are unlikely to stop the propagation from the reservoir into the Carboniferous. Regarding (2), there is no indication that the properties of the Carboniferous will only allow deformation by aseismic creep and it is unlikely that ~100m of throw accumulated on faults strictly via aseismic processes. Assuming that earthquakes occurred on these faults in the geologic past makes it very unlikely that frictional properties can't sustain an earthquake, and therefore variations in

friction coefficient are unlikely to stop an earthquake. While fault geometry and fault properties are probably not going to stop rupture propagation, the state of stress (3) has a large potential to stop rupture propagation. A rupture cannot propagate into a region where there is no stress to be released (for example, an earthquake is unlikely to rupture the same portion of fault that hosted a recent event because the prior event released most of the built-up shear stress). An evaluation of the stress state on the faults in the Carboniferous can indicate if there is shear stress to be released.

For a rupture to propagate into the Carboniferous, there must be enough shear stress to sustain rupture propagation. The historically seismically quiet area of the northern Netherlands is consistent with faults subjected to a low background shear stress/strength ratio (i.e. low shear stress/normal stress). The shear stress/strength ratio is often referred to as the shear capacity utilization (SCU). Modeling work by Shell indicates that there is a local extreme decrease in the SCU at the base of the reservoir as the reservoir is depleted (figure 11). In addition to a low background level of SCU, this local decrease may provide an additional barrier to rupture propagation. However, it is difficult to quantify the magnitude and extent of the local barrier given the simplified model geometry and the dependence on fault offset, fault dip and depletion, and their variation throughout the field. Therefore, the focus here is on the background stress level and it is acknowledged that the local barrier may further decrease the probability of rupture into the Carboniferous beyond what is considered here.

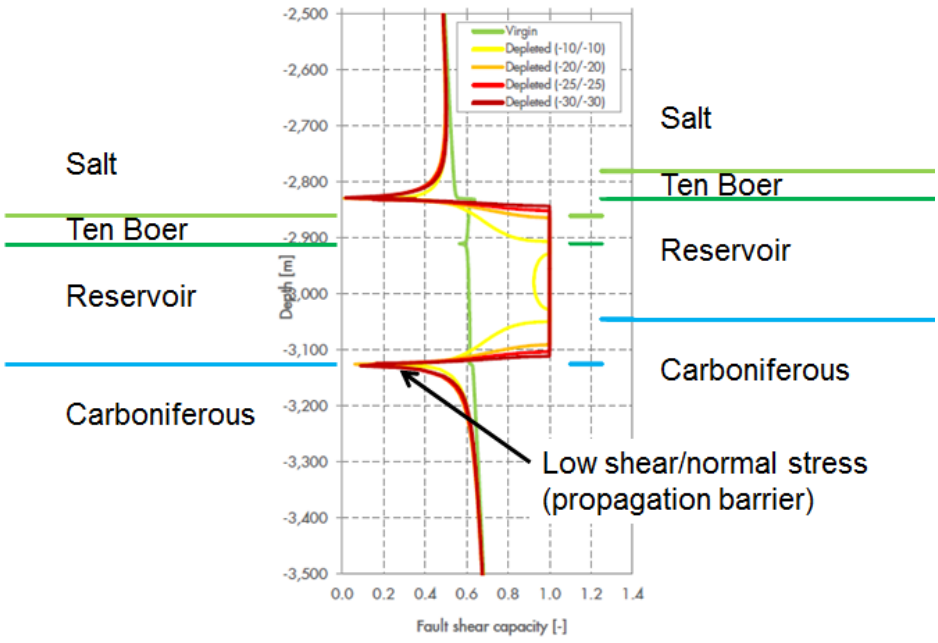

Figure 11. Relative stressing level on the down-dip dimension of the fault as the reservoir is depleted (green is virgin conditions and yellow to red are the depleted state). A fault shear capacity of 1.0 indicates that the shear stress level has reached the strength. A representative fault dip and fault offset are assumed and the offset results in a low fault shear capacity at the base of the reservoir. Legend indicates bars of depletion imposed in the reservoir on each side of the fault (figure adapted from van den Bogert, 2013)

3.2 Criteria for determining if stress will inhibit rupture propagation

3.2.1 Fault friction during sliding

Earthquakes are able to propagate in low shear stress environments (i.e. on faults that are not optimally oriented for failure and through areas that are not critically stressed). Earthquakes start because the stress state in a nucleation area reached the failure strength of the fault, generally characterized by a static coefficient of friction ($\mu_s \sim 0.6 - 0.8$). However, once the fault starts to slip, the coefficient of friction that characterizes the fault surface can change drastically. During an earthquake, the fault begins to slip at very high slip velocities (~ 1 m/s) and at these high velocities, a range of thermal weakening mechanisms come into effect, causing the coefficient of friction to decrease drastically. Figure 12 shows a compilation of laboratory data collected at seismic slip velocities for a range of geologic materials (clay gouge, marble, dolomite gouge, gypsum gouge, etc.). All studies exhibit a weakening of the friction coefficient to a dynamic value of $\mu_d \sim 0.1 - 0.3$ (highlighted by the red "SS" box for Steady State). Before this highly weakened state is reached, the material responds with the static coefficient, $\mu_s \sim 0.6 - 1.0$ (highlighted by the red "P" box for Peak strength), spanning a larger range than the generally considered traditional values of $\mu_s \sim 0.6 - 0.8$. This extreme weakening has also been validated by field studies that have drilled through a fault zone after an earthquake occurred and examined the heat anomaly that persisted. If the fault slid at a high coefficient of friction for meters of slip, in theory a large amount of heat should have been created. These field studies consistently detect only a small heat anomaly, however, so the seismic friction value is determined to be in the range of 0.08-0.12.

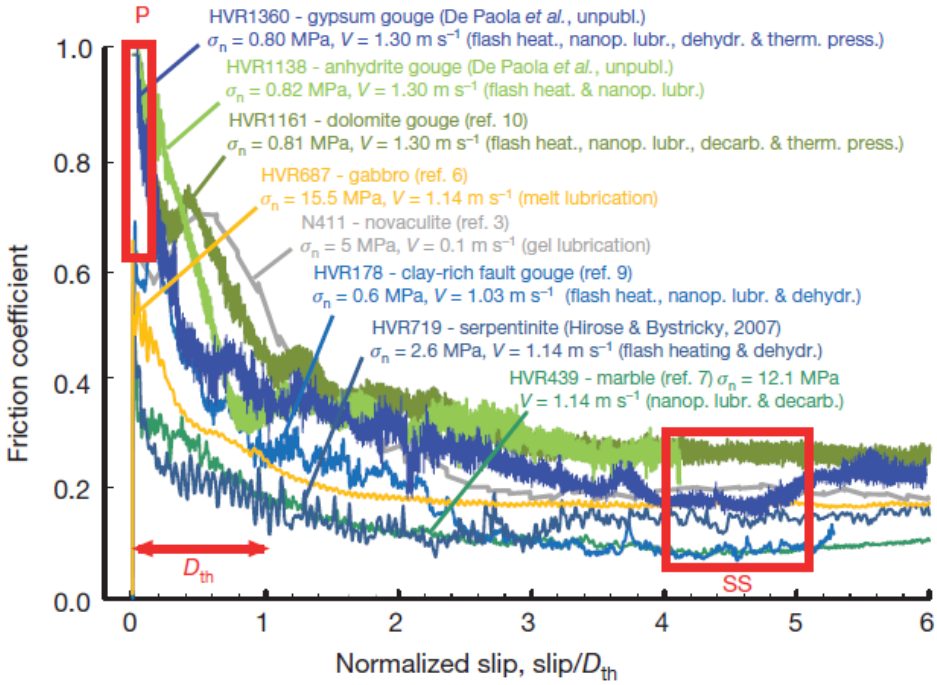

Figure 12. Laboratory data compilation presented in Di Toro et al., (2011). Results show the extreme frictional weakening that occurs at high slip velocities for a variety of materials. All materials reach a steady-state friction value in the range of 0.1-0.3 after reaching a peak value of 0.6-0.8.

Some amount of slip must accumulate before this weakened state is reached. In the laboratory setting this slip value is on the order of a meter, but the laboratory cannot duplicate in situ stress conditions. It is observed that the slip weakening distance decreases with increasing normal stress and since in situ stress states have a larger normal stress than those reproduced in the laboratory, the slip weakening distance is shorter than the lab measured value. The degree of weakening, the slip distance to weakening, and the modeling of the mechanisms that control it are areas of active study. This uncertainty means that by the time an earthquake reaches the size of a M 3 event, ~ 1 cm of slip has occurred and this may not be enough slip for full dynamic weakening to occur. At this time the dynamic friction values that are relevant to the Groningen earthquakes are not known.

Given the laboratory data presented in figure 12, a range of possible dynamic coefficients of friction ranging from 0.05 to 0.3 are considered. The relative likelihoods of these values are shown in figure 13, with a most likely value of ~0.15. Additional distributions for the possible dynamic friction value are shown in figure 13 (base case scenario denoted with yellow star). One case considers the possibility that the dynamic coefficient of friction cannot reach fully weakened values after only ~1 cm of slip, so high dynamic friction values up to 0.45 are included. In contrast, recent laboratory studies have found that lower and lower values for the dynamic coefficient of friction are possible, so an alternative representation that reduces the likelihood of the larger dynamic values is shown in figure 13.

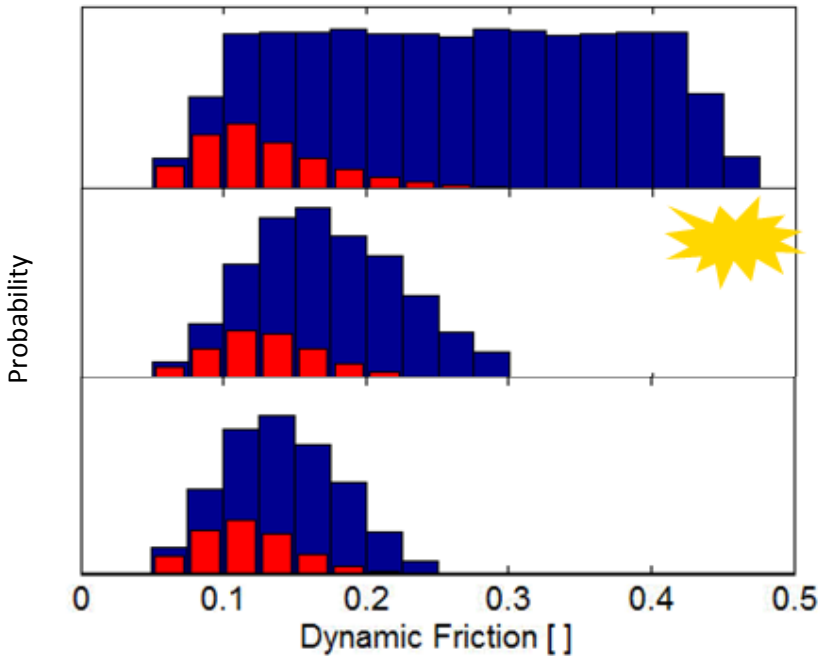

Figure 13. Three versions of assumed dynamic friction distributions. Distribution with the yellow star is the base case, base case distribution. The blue bars are the input assumed distribution for the Monte-Carlo analysis. For comparison, the red bars are the distributions of the cases that resulted in a stress drop during the Monte Carlo analysis.

3.2.2 Rupture propagation into areas of low fault loading

Before an earthquake, a fault is subject to an average level of background stress. There is heterogeneity in the stress field on the fault due to surface roughness and fault memory of previous slip events, but here these variations are neglected and treated with an average value. The background level of shear stress is labeled τ_0 in figure 14; ahead of the crack tip, the fault is at this background level. Far behind the crack tip, the fault has fully weakened and is at a residual shear stress level, $\tau_r = \mu_d \sigma$, determined by the dynamic coefficient of friction, $\mu_d \sim 0.1 - 0.3$, and the effective normal stress, σ. At the crack tip, the material is breaking, so peak shear stress values $\tau_p = \mu_s \sigma$ are reached, determined by the static coefficient of friction, $\mu_s \sim 0.6 - 0.8$. These peak shear stress values are only reached because of the presence of the crack, which creates a stress concentration. Stresses increase as the crack tip is approached. If the crack is long enough, the stresses are high enough that the strength of the material is exceeded (τ_p is reached) and the crack becomes unstable. This is similar to the processes that take place after a rock hits a windshield. The rock impact creates a crack, which at some point in the future can become unstable and grow. The rest of the windshield, away from the crack, is nowhere near failure, but the presence of the crack is able to bring the glass to failure and this failure results in crack propagation.

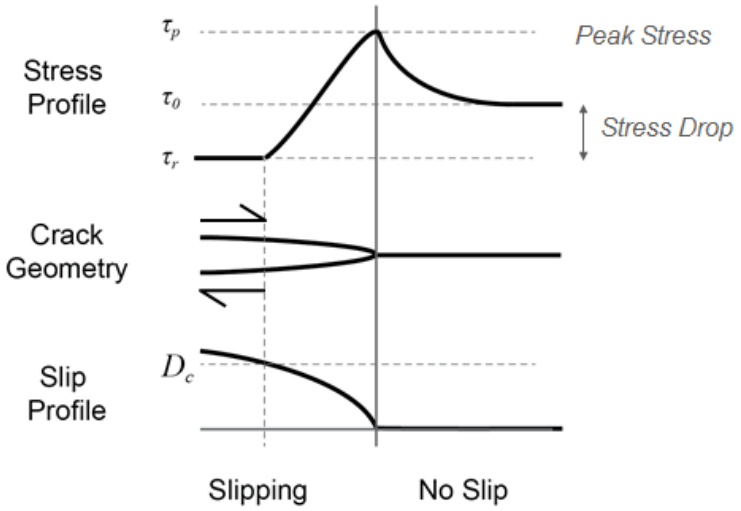

Figure 14. Slip and stress distribution around a propagating crack tip. The dynamic coefficient of friction specifies the residual stress level. At the crack tip the material is breaking at a stress level specified by the static coefficient of friction. Ahead of the crack tip the material is not at failure. The difference between the pre-crack stress and the post-crack stress is the stress drop.

3.2.3 Stress-drop-based criterion

The growth of a crack is governed by an energy balance between the energy consumed by the fracture process and the release of energy as the fault slips and the off-fault material is unloaded. Here, a simplified criterion built around the static stress drop is considered. The static stress drop is the difference between the pre-earthquake and the post-earthquake shear stress on the fault ($\tau_0 - \tau_r$). The peak stress τ_p need not be considered in determining the static stress drop. If there is no stress drop, the fault slip occurs quasi-statically and there is no possibility of seismic rupture. By examining the stress state in the Carboniferous, τ_0 and σ can be determined, and by assuming a value for the dynamic

coefficient of friction, τ_r can be determined. The scenarios that could result in a stress drop can then be constrained.

Some assumptions are made during this methodology that can affect the results. It is assumed that faults are cohesionless and that the stress state on a fault in the Carboniferous is the same as the bulk stress state in the Carboniferous (i.e. stress rotation in the vicinity of faults and heterogeneity in the stress state on the fault are not considered). Additional depletion-induced stress barriers that result from the geometry at the base of the reservoir are neglected and it is assumed that very small stress drops (< 0.1 MPa) could sustain rupture propagation.

The existence of a stress drop is a necessary but not sufficient condition for rupture propagation. Stress analyses can conclude that there is a potential stress drop, but earthquake propagation frequently arrests due to heterogeneity and effects not captured by this analysis. It is not possible to know why every individual earthquake rupture arrests, or why a fault patch slips in 10 small events rather than one larger event, but seismic slip cannot occur without a stress drop. For these reasons, the existence of a stress drop does not mean that a fault will rupture (not a sufficient condition) but it means that the fault could rupture (necessary condition).

3.3 Estimate of stress state on faults in the underlying Carboniferous

To determine the state of stress on a fault in the Carboniferous, the most important quantities are the minimum horizontal effective stress, the vertical effective stress, and the dip of the fault. To estimate the effective stress state, both total stress and pore pressure estimates are made. All stress values are presented as depth normalized values (bar/10m or .01 MPa/m). To evaluate the magnitudes of the stresses a depth is assumed and multiplied by the stress gradient. A summary of the stress analysis is presented here and Appendix A contains additional information used to determine state of stress.

To the best of our knowledge, events are normal faulting earthquakes so the primary concern is the shear stress on the fault that will promote dip-slip faulting below the reservoir. The minimum (Sh) and maximum (SH) total horizontal stresses are very similar in magnitude (SH/Sh ~1.07) (van Eijs, 2015). This similarity means that for a constant dip, the strike of the fault will have little effect on the shear and normal stress levels resolved on the fault. The contrast between the horizontal and vertical stresses is much larger, resulting in larger possible shear stress variation dependent upon the fault dip. This makes the dip the most important factor in determining the relative levels of shear and normal stress resolved on the fault.

In normally stressed regimes (vertical stress is the maximum) high vertical stresses are associated with high horizontal stresses, indicating that some of the input parameter distributions are correlated. To address this, a criterion for admissible Sh/Sv ratios is discussed and the application of this criterion prevents the consideration thrust stress regime scenarios.

There is uncertainty and variation throughout the field regarding the orientation of the minimum horizontal stress, so as a base-case all faults are assumed to be optimally oriented for dip-slip motion (i.e. the strike is perpendicular to the minimum horizontal stress direction). As a sensitivity, the conversion of a pure dip-slip event into an event with a strike-slip component at depth is also

considered. Accounting for this possibility allows the small horizontal stress anisotropy to modify the shear stress, which has an effect on the likelihood of having a stress drop. For this variation, the optimal strike orientation for slip is again considered.

3.3.1 Fault dip

NAM undertook an extensive fault mapping project using the 3D seismic volume, thus the dip range and variation throughout the field, at the reservoir depth, is constrained by the seismic data. Figure 15 shows the distribution of mapped fault dips for two subsets of faults with differing lower bounds: all faults longer than 2km and all faults longer than 7 km. Many faults shorter than 2 km are mapped as a purely vertical fault with a dip of 90 degrees, possibly due to a resolution issue on small faults with little observable fault offset, making them difficult to represent accurately. Since their apparent geometries are likely an artifact of interpretation, these small faults are not included. Bigger, longer faults tend to be more shallowly dipping and are also the faults that are capable of hosting the biggest earthquakes. Earthquakes nucleate on both big and small faults so ignoring all faults shorter than 7 km is not a proper representation. However, this subset of faults is considered here as a variation to examine the sensitivity to dip of the likelihood of rupture into the Carboniferous.

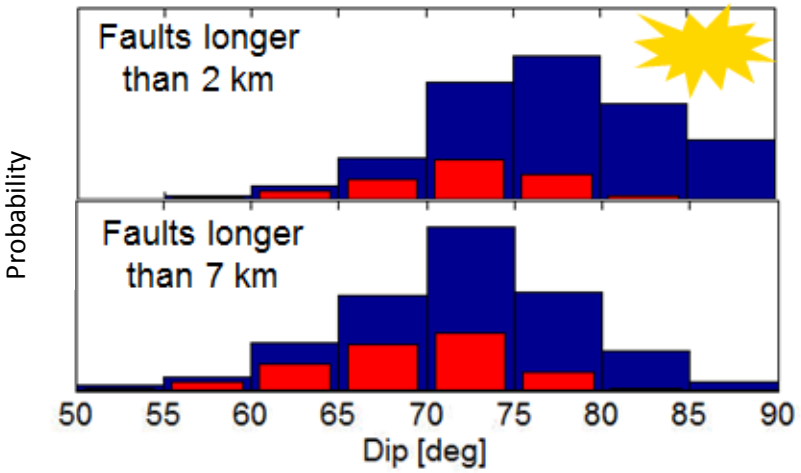

Figure 15. Two versions of assumed fault dip distributions. Distribution with the yellow star is the base case, base case distribution. The blue bars are the input assumed distribution for the Monte-Carlo analysis. For comparison, the red bars are the distributions of the cases that resulted in a stress drop during the Monte Carlo analysis.

3.3.2 Vertical total stress

The total vertical stress is the weight of all the material above a point, and can be determined by integrating the density log of wells in the field. In the shallow subsurface (~0-400 m depth) logs are not generally collected so approximations must be made. However, even with this lack of data, the vertical total stress is the most constrained of the three principal stresses. The vertical stress gradient varies from 2.15-2.33 bar/10 m at the reservoir depth (van Eijs, 2016). Most wells show a vertical stress gradient in the range of 2.15-2.25 bar/10m, but there are a few measurements on the high side in the 2.25-2.33 bar/10m range. The variation throughout the field is represented with the histogram shown in figure 16a (histogram with yellow star is the base case distribution). Since the vertical stress is the

maximum stress, high values will result in higher shear stresses on the Carboniferous faults at depth and will promote the possibility of out-of-reservoir rupture. The alternate distributions in figure 16 are other possible interpretations of the limited data to capture the variation and uncertainty throughout the field. The alternate distributions are considered in the sensitivity analysis (section 3.5).

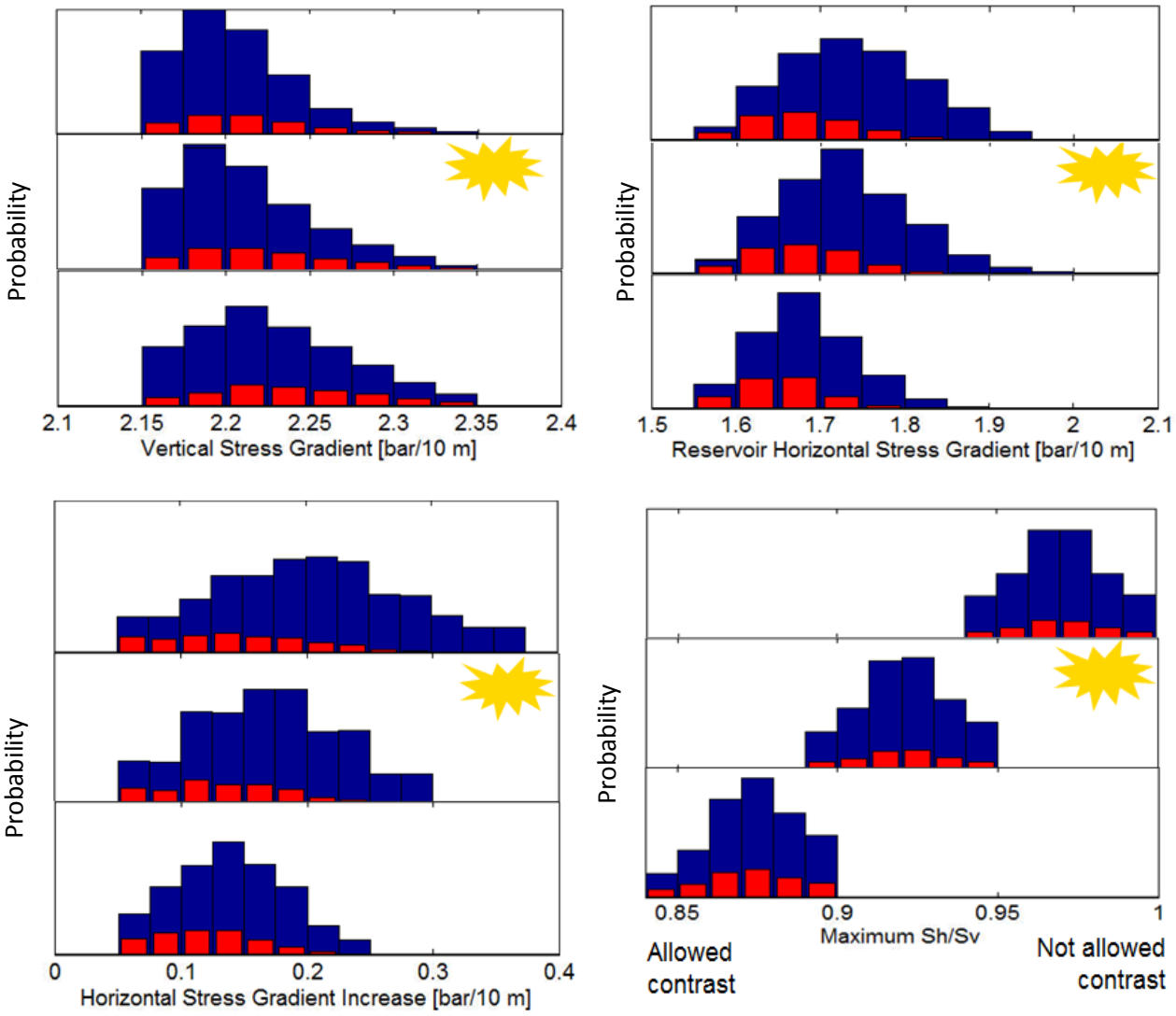

Figure 16. (a) Three versions of assumed vertical stress gradient distributions. (b) Three versions of assumed initial horizontal stress gradient distributions in the reservoir interval. (c) Three versions of assumed increase in the horizontal stress gradient distributions in the Carboniferous compared to the reservoir. (d) Three versions of assumed maximum contrast between horizontal and vertical stress distributions (this ensures that all scenarios result in normal faulting regimes, consistent with observations). Distribution with the yellow star is the base case, base case distribution. The blue bars are the input assumed distribution for the Monte-Carlo analysis. For comparison, the red bars are the distributions of the cases that resulted in a stress drop during the Monte Carlo analysis.

3.3.3 Minimum horizontal stress in the Carboniferous

There are no present day or pre-production stress measurements in the Carboniferous. We must estimate the current minimum horizontal total stress in the Carboniferous based on the limited number

of stress measurements that exist, which were all made in the Rotliegend reservoir under various stages of depletion. The present day stress state in the reservoir is highly altered from virgin conditions due to the 300 bars of depletion that has occurred. The Carboniferous is predominantly shale, and due to its low permeability there should be a minimal pressure decrease, implying a minimal change in the state of stress from virgin conditions due to production from the Carboniferous. Local perturbations to the stress state due to isolated depleted sand bodies in the Carboniferous are neglected here.

The present day minimum horizontal stress in the Carboniferous should be related to (but not equal to) the initial, pre-production, minimum horizontal stress in the reservoir. Due to differing lithologies and geologic histories, the pre-production state of stress in the Carboniferous shale likely differed from the pre-production state of stress in the reservoir by some increment and this increment must be estimated. To estimate the present day minimum horizontal stress in the Carboniferous, this increment is added to an estimate of the initial minimum horizontal stress in the reservoir. The initial state of stress in the reservoir is estimated by extrapolating the current day stress state back to initial conditions based on assumptions about the rock properties.

3.3.3.1 Initial minimum horizontal stress in the reservoir

Recent analyses of the stress measurements/inferences that have been made in the Groningen field have revealed that the observations of lost returns early in the field life are likely not an accurate measure of the stress state, therefore the only reliable measures of stress in the reservoir have been made after a substantial amount of depletion (van Eijs, 2015). Five measurements were made after 150-250 bars of depletion (initial reservoir pressure was 350 bars).

It is generally observed and understood that the total horizontal stress decreases as the pore pressure decreases. The relative magnitude of these effects is captured by the depletion constant, γ,

$$\gamma = \frac{dS_h}{dP_p} \approx \alpha \frac{1-2\nu}{1-\nu}$$

where ν is Poisson's ratio and α is Biot's coefficient. Measurements from the Groningen field are unable to constrain the value for this parameter, but it is likely that γ is bounded by an assumed range of $\gamma = 0.5 - 0.8$. Since the depletion is known, by assuming a value for γ the change in horizontal stress from initial conditions to the time of the measurement can be determined. This can then be used to determine the initial value for the minimum horizontal stress.

The assumed variation and uncertainty in the initial horizontal stress state in the reservoir is shown in figure 16 (distribution with yellow star is the base case distribution). The stress state is likely in the range of 1.6-1.8 bar/10 m, but given the uncertainty in γ, there is also a possibility for larger values. Note that larger values result in a more isotropic stress state and decrease the likelihood of rupture out of the reservoir. To examine the sensitivity of the result to this assumption, additional distributions are considered in that allow for higher or lower horizontal stress distribution assumptions (figure 16).

3.3.3.2 Stress contrast between the reservoir and the Carboniferous

Present day horizontal stresses should be higher in the Carboniferous than the reservoir due to several mechanisms:

- **Sand-shale contrast** – Horizontal stresses are generally higher in shales (Carboniferous) than in sands (Rotliegend reservoir) (contrasts of 0.07-0.27 bar/10 m measured in nearby Rotliegend reservoirs)
- **Elevated Carboniferous measurements** – Higher horizontal stresses were measured in a Carboniferous sand 40 km to the south (~ 0-0.15 bar/10m)
- **Stress-arching** – Reservoir contraction due to depletion can alter the horizontal stress above and below the reservoir (Segall, 1989) (second order effect, increase of ~0.03 bar/10 m or less)
- **Coal** – Coals in the Carboniferous can alter the stress state (this effect can be large but is estimated to be a second order effect here due to the low coal content, ~0-0.05 bar/10 m)

Of all the effects considered it is clear that the sand-shale contrast is the largest. The distribution shown in figure 16 (base case distribution marked with a yellow star) is used to estimate the stress increase in the Carboniferous. It is consistent with the sand-shale effect and includes ~30% of the increase in the 0.2-0.3 bar/10 m range that is relevant if multiple effects are additive. Additional distributions are considered and shown in figure 16 , allowing for more or less increase in the horizontal stress in the reservoir.

3.3.4 Covariance between horizontal and vertical stresses

A degree of covariance between stress parameters is expected. In a passive margin setting (vertical stress is the maximum), the earth is loaded with a vertical stress due to sedimentation and horizontal stresses develop in response to this loading and lateral constraint. For given material properties, higher vertical stresses imply that higher horizontal stresses should develop. All drilling experience and data gathered in the Netherlands is consistent with a normal stress regime (vertical stress is the maximum) and normal faulting events are observed in the Groningen field. Based on this, it is inconsistent to sample a stress state from the above probability distributions that would result in thrust faulting. If a low vertical stress is sampled but a high original horizontal stress and a high Carboniferous stress increase are sampled, the result is a higher horizontal stress than vertical. This is considered an inadmissible scenario not consistent with the expected covariance.

To prevent scenarios that result in thrust faulting, a maximum value for the ratio of horizontal to vertical stress is imposed. Thrust faulting results if the ratio is larger than 1.0, but horizontal to vertical stress ratios of near one are rarely observed in normal faulting regimes. As a base case, a maximum value for the stress ratio of 0.89-0.95 is imposed, with 0.92 as the median maximum (figure 16 shows the distribution of maximum ratios with the base case distribution marked with a yellow star). Additional values for the maximum contrast are also considered as sensitivities.

3.3.5 Pore pressure

Except in a few isolated places, the Carboniferous rocks are located in the water leg (below the gas-water contact). While pore pressure in the gas column is well constrained, there are few pressure

measurements in the water leg and they vary throughout the field from 1.15-1.17 bar/10 m. A value of 1.17 bar/10 m is used as the base case pore pressure gradient. Values ranging from 1.15-1.22 bar/10 m are also considered as variations. The high value of 1.22 bar/10 m would be representative of a scenario in which substantially more overpressure (elevated pore pressure) is able to develop within the Carboniferous shale.

3.4 Possibility of Carboniferous rupture for base case assumptions

Variations in the stress state throughout the field, as well as epistemic uncertainty, are taken into account by implementing a Monte Carlo sampling. Each scenario sampled from a variety of input distributions results either in the possibility of a stress drop, or else the shear stress on the fault is small enough that a stress drop is not possible. The percentage of scenarios that have a potential stress drop determines the likelihood that a given earthquake could propagate out of the reservoir into the underlying Carboniferous.

50,000 trials are sampled from each of the distributions and if a trial violates the maximum admissible Sh/Sv, that trial is dis-regarded and a new trial is sampled (50,000 scenarios were used to reach a stable solution). Figure 17 shows the results of a sub-sample of the 50,000 scenarios. Each blue dot is plotted at the shear and normal stress resolved on the fault for a fault dip and stress state scenario. A purely vertical fault will have no dip-slip shear stress resolved on it, so it will plot on the x-axis (i.e. shear = 0). A depth of 3500 m was chosen to evaluate the magnitude of the stresses. The light gray dashed lines represent different coefficients of friction. Cohesion is not considered here, so these friction lines also represent the ratio of shear to normal stress that exists on the fault, which is an indication of stress vs. strength. High realization values (exceeding higher frictional failure lines) indicate that the fault is closer to failure and low values that it is further away.

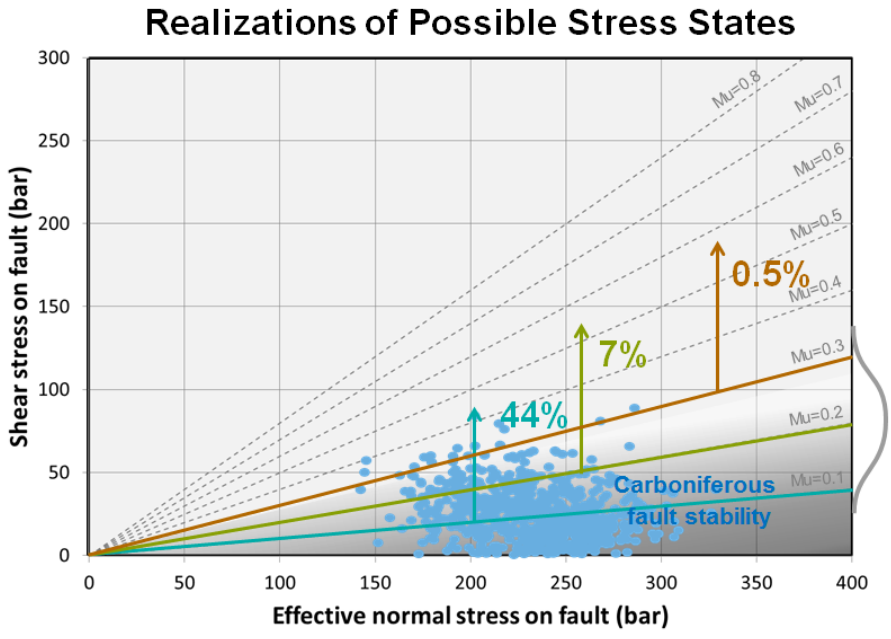

Figure 17. Blue dots are a subset of the realizations of the state of stress on faults in the Carboniferous as 3500 m depth. Most realizations indicate that the faults will be at a low level of shear stressing.

None of the blue dots are close to the traditional frictional failure line characterized by $\mu = 0.6$. This indicates that the faults in the Carboniferous experience a background stress state far from failure (i.e. it would be very difficult to nucleate an earthquake in the Carboniferous). This is consistent with the historical observation of a lack of seismicity and with the present day observation that events are nucleating in the reservoir interval. Figure 18 shows realizations of the initial stress state in the reservoir (a depth of 3000 m was used). These red dots are also not approaching the traditional $\mu = 0.6$ failure line. Therefore, in the initial state, neither the reservoir nor the Carboniferous were close to failure. This is consistent with the lack of historical seismicity.

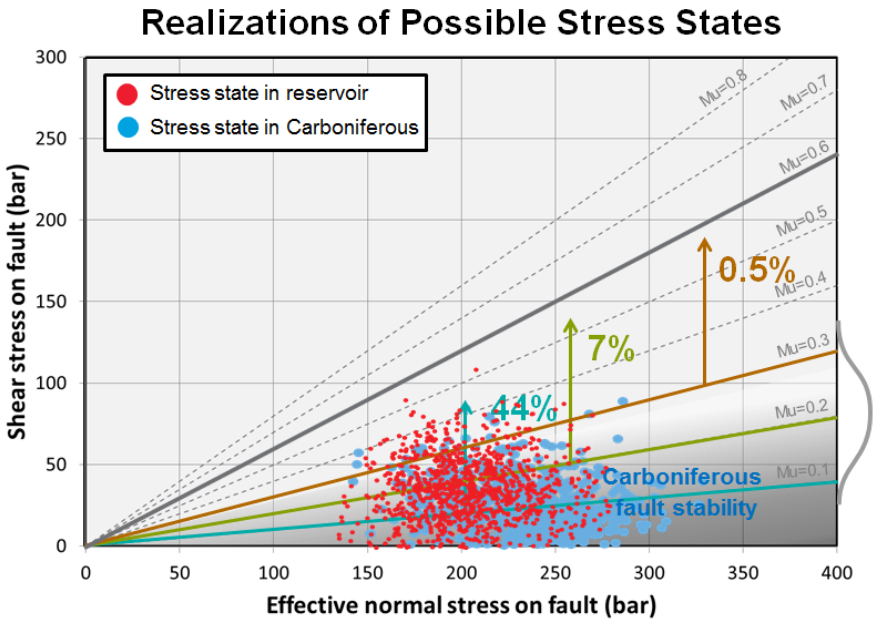

Figure 18. Blue dots are the realizations of the present day state of stress on faults in the Carboniferous (3500 m depth) and red dots represent the initial, pre-production state of stress in the reservoir (300 m depth) for all faults longer than 2 km. All realizations indicate that the faults are at a low level of shear stressing, far from the 0.6 level. This is consistent with the lack of historic seismicity in the area. Gray line on the right side indicates the uncertainty in the value of the dynamic friction coefficient

If the dynamic friction coefficient is 0.1, then after an earthquake occurs a given point would reside on the $\mu = 0.1$ line. 44% of the Carboniferous realizations reside at a shear to normal stress ratio larger than 0.1, meaning 44% of scenarios have shear stresses that could be released in the process of getting to the 0.1 line. The 56% of scenarios that reside below the 0.1 line have so little shear stress that there is not enough stress for an earthquake to release, meaning rupture propagation could not be sustained and the rupture would die out. 7% of the scenarios result in a shear to normal stress ratio larger than 0.2 and only 0.5% result in a shear to normal stress ratio larger than 0.3. Although informative, none of these percentages can be explicitly used because there is uncertainty in the applicable value for the dynamic coefficient of friction.

Figure 19 shows the comparison of the distribution of the Carboniferous stress realizations with the distribution of the base case dynamic friction distribution. If the base case distribution for friction coefficient is used, 18% of the points are found to lie above the dynamic friction value for that scenario.

This can be interpreted as saying that 18% of the Carboniferous scenarios could result in a stress drop > 0 MPa. The static stress drop during an earthquake is simply the vertical distance between the blue dot that represents the state of stress and the relevant dynamic friction coefficient line. Normal stress does not change after an earthquake so only the vertical dimension on the graph is important. 13% of the scenarios have a stress drop > 0.5 MPa, 1% have a stress drop > 3 MPa and 0% have a stress drop > 10 MPa. While this is a real quantity that represents an actual change in shear stress on a fault, it should be noted that this is not exactly the same as the stress drop inferred from seismic recordings. They should be the same in theory, but given the many assumptions in the inversion process, the seismic stress drop is often referred to as the stress parameter to avoid over-interpretation.

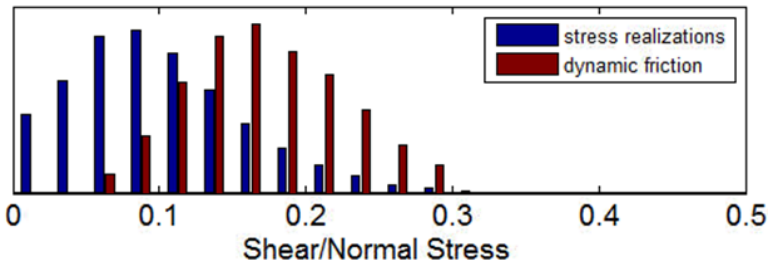

Figure 19. Distributions of stress realizations and the assumed dynamic friction coefficient. If the realization is above the dynamic friction level it is an admissible scenario that could allow for rupture propagation into the Carboniferous.

Because it is difficult to know what stress drop value should be used as a cutoff for allowing rupture to propagate, a conservative choice is made using stress drop > 0 MPa. This means that 18% of the time an earthquake exists in a world in which rupture propagation into the Carboniferous is possible, which is not the same thing as saying that 18% of the time a rupture propagates into the Carboniferous. Earthquakes occur over a range of magnitudes and it is difficult to assess why any individual rupture stops. Even if bulk representation of the conditions in the Carboniferous did not prevent rupture, any given earthquake is still subject to a heterogeneous environment and governed by the same rules that result in a Gutenberg-Richter distribution. In 18% of the scenarios the rupture could propagate out of the reservoir zone, but according to the G-R relationship (minimum magnitude of M 1.5), 1 in 100 events would be larger than a M 3.5 and 1 in 1000 events would be larger than a M 4.5 (and require out-of-reservoir propagation). 18% of the time the earthquake scenarios are sampling from a G-R distribution with a larger truncation magnitude (possible unconfined, full fault rupture) and 82% of the time earthquake scenarios are sampling from a G-R distribution with a smaller (confined-event) truncation magnitude.

3.5 Sensitivity of results and additional considerations

The results presented in the previous section are highly dependent on the assumed input distributions. Figure 20 is a tornado chart that illustrates the sensitivity of the 18% result to the assumed inputs. The dynamic friction coefficient and the fault dip are some of the more important parameters.

While restricting the dip to the longest faults (> 7 km) results in a large change in this probability (18% unconfined becomes 30% unconfined), earthquakes are not limited to initiating only on long faults,

making the ≥2 km long fault set a better representation of the base case. The dynamic friction coefficient distribution is uncertain and with further study some of the uncertainty around this parameter may be reduced. Other parameters, such as horizontal stress increase in the Carboniferous, have an effect on the base case number but do not change the result substantially. Therefore, 20% is used as the logic tree weight to express the probability that the system is described with a large truncation magnitude, and 80% is the probability that a smaller truncation magnitude describes the system.

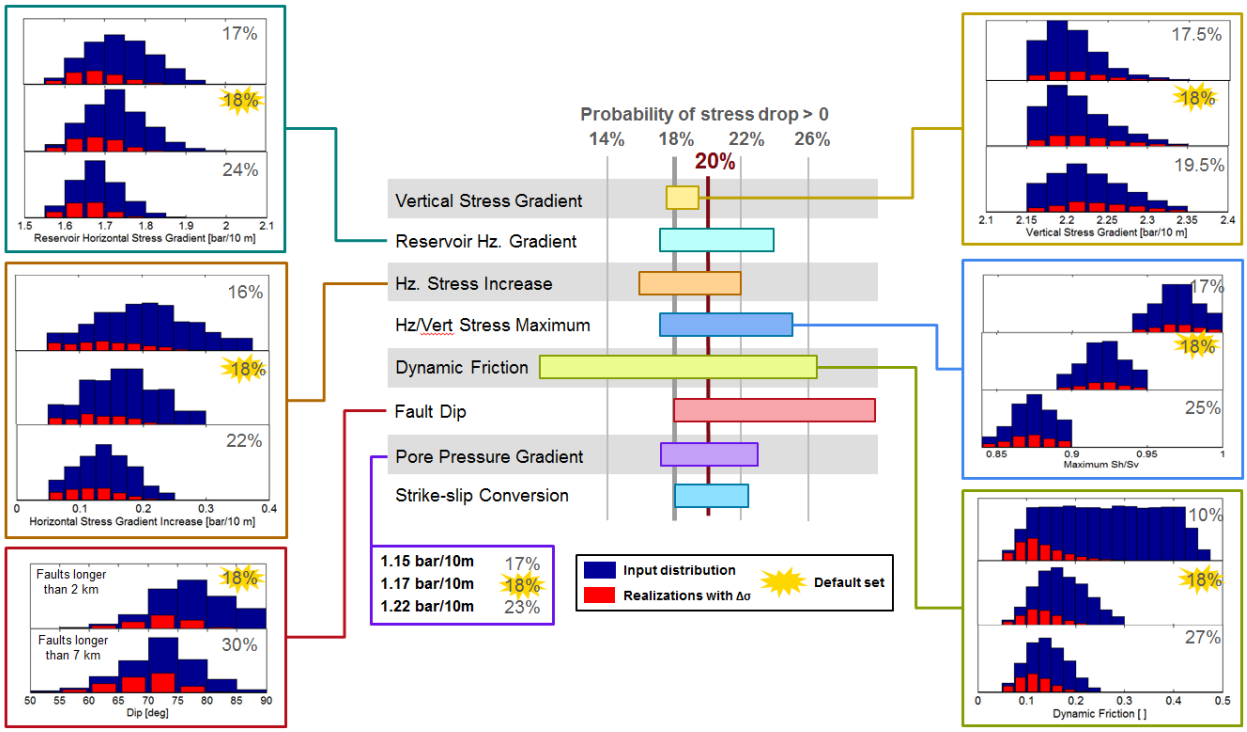

Figure 20. Tornado chart showing the variation of the results to the assumed input distributions. Based on the variation to the parameter inputs, a 20% probability of having a stress drop > 0 is assumed.

In addition to variations in the input distributions, a mode conversion of the rupture as it enters the Carboniferous was considered. In the preceding analysis, optimal fault orientation for dip-slip stressing was assumed (i.e. fault strike parallel to the maximum horizontal stress direction). The result of this assumption is that a vertical fault has no shear stress resolved on it, making Carboniferous rupture impossible. However, there is some anisotropy in the horizontal stress magnitudes so a vertical fault would have shear stress resolved on it, promoting lateral slip (i.e. strike-slip motion). To determine the magnitude of this effect, the best estimate of the horizontal stress anisotropy was assumed (SH max/ Sh min = 1.07). Additionally, all faults were assumed to have an optimal strike maximizing the combined strike-slip and dip-slip shear stress magnitude. The result is that all faults have at least some component of shear stress, meaning no points lie on the shear stress equals zero axis in figures 17 and 18. All points shift up in figures 17 and 18 but the increase is small. If strike-slip motion is allowed to occur, rather than 18%, 23% of the scenarios could possibly host an out-of-reservoir event.

3.6 Summary: Probability that all events will be confined to the vicinity of the reservoir

Earthquakes are nucleated in the reservoir due to depletion-induced stress changes. It is believed that these stress changes are only large enough within the reservoir to initiate an earthquake. Once nucleated, ruptures have the potential to propagate into lower loading environments (i.e. out of the reservoir area), but the shear stress state below the reservoir may be too low to sustain rupture. Given the uncertainty in the state of stress away from the depleted interval and the variation throughout the field, ~**20%** of Monte Carlo realizations have the potential to allow out of reservoir rupture propagation. This does not mean that 20% of ruptures *will* propagate out of the vicinity of the reservoir, but rather 20% of the time an earthquake *could* propagate out of the vicinity. Small earthquakes occur more frequently than large earthquakes (Gutenberg-Richter relationship) and ruptures often arrest for reasons unknown to researchers. Therefore, even if a rupture could propagate into the Carboniferous, most events will be small and terminate before that point is reached.

This result is highly dependent on the degree of dynamic weakening (the dynamic friction coefficient), a quantity that is poorly constrained. Future work could determine that the assumed range was too high or too low and alter the probability that a rupture could propagate deep into the Carboniferous. Additional measurements could be made to determine the current state of stress in the Carboniferous and this also has the potential to change the ~20% determination.

A 20% likelihood of events propagating out of the reservoir supports a low truncation magnitude Gutenberg-Richter and is consistent with the observations of catalog to date.

4.0 Events confined to the vicinity of the reservoir

It is possible that ruptures cannot propagate deep into the Carboniferous. If ruptures are stopped, what would the largest confined event look like and what would be the resulting maximum magnitude?

4.1 Definition of a confined event

In the case that rupture cannot be sustained in the Carboniferous, the phrase “confined” can be applied. This does not mean that slip on the fault is confined to that area where the reservoir sand is in contact with the fault, but rather that the rupture cannot be sustained out of the immediate vicinity of the reservoir. A confined event would start to propagate out of the reservoir, but would encounter conditions that were unfavorable for continued rupture propagation. In this case, the rupture would rapidly slow down as it lost energy but would still propagate some distance into the Carboniferous. There is no physical reason that a rupture would encounter the top of the Carboniferous (the Saalian unconformity) and instantaneously lose all inertia and stop. When the term “confined” is used, it means that the rupture is confined to the vicinity of the reservoir – it does not mean that the slip patch is strictly confined to the sand zone of the reservoir.

The Carboniferous is predominately shale but there are isolated bodies of sand with unknown lateral extent. There is an angular unconformity between the Carboniferous and the Rotliegend so areas exist where these Carboniferous sand bodies are in contact with the reservoir (figure 21). These

Carboniferous sands will be depleted and this will change the stress state in the sands and the shale around them. If a fault is located near one of these sand bodies then this depletion is another mechanism by which a rupture could propagate a short distance out of the reservoir.

To estimate the magnitude of an earthquake, an assumption about the down dip dimension of the slip patch must be made. The dip of the fault implies that a confined event could span a dimension larger than the reservoir thickness and the pre-existing throw on the fault could also make a larger area available for rupture (the Rotliegend averages ~210 m thickness with an offset of ~80 m). The gradual loss of inertia and the depleted Carboniferous sands are mechanisms that allow a rupture to propagate out of the zone of depleted Rotliegend sandstone but ultimately still lead to rupture termination. Because of these considerations a down-dip slip dimension of 350 m is assumed although smaller or larger values could apply in different locations throughout the field.

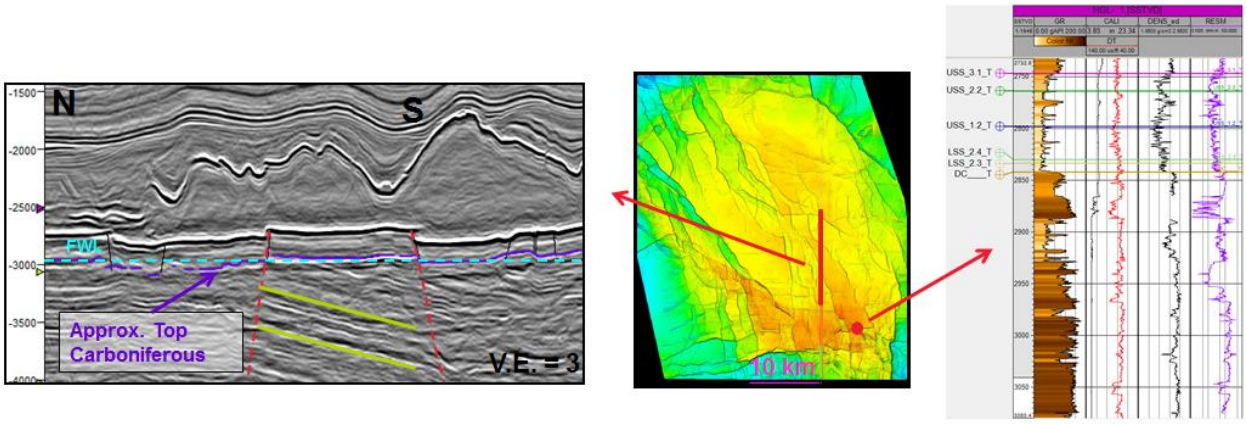

Figure 21. The structure in the Carboniferous puts some isolated sandy intervals in contact and pressure communication with the depleting reservoir. The log on the right shows one of these sandy intervals below the top of the Carboniferous (DC_T).

4.2 Magnitude determination from earthquake scaling relations

Earthquake scaling relations can be used to understand the size of a rupture area for earthquakes of different magnitudes. These scaling relations depend on the shape of rupture assumed (circle, square or rectangle) and most importantly the stress drop, $\Delta\sigma$. Larger stress drops result in a larger amount of slip, d,(i.e. $d \propto \Delta\sigma$) over the same rupture dimension and hence a larger magnitude event. Stiffer materials will have less slip for the same stress drop, so slip is inversely proportional to the shear modulus, G (i.e. $d \propto 1/G$). Earthquake moment magnitude is determined from the seismic moment, $M_0 = GAd$, where A is the fault area and by substitution it can be found that $M_0 \propto A\Delta\sigma$. The shear modulus cancels out, so for an assumed stress drop, the rupture area is unaffected by the assumed shear modulus, but the slip is affected. The form of the proportionality depends on the geometry of the slipping patch assumed. For a rectangular dip-slip fault with down-dip dimension, *w*, the relationship is defined by the following equation (Stein and Wysession, 2003):

$$\Delta\sigma = \left(\frac{8}{3\pi}\right) * \left(\frac{G\,d}{w}\right) = \left(\frac{8}{3\pi}\right) * \left(\frac{M_o}{A\ w}\right)$$

From corner frequency analysis performed on 18 Groningen events, it is seen that a range in stress drops could be appropriate, ranging from 0.1 to 10 MPa with 3 MPa as the best all-around fit to the data (Bommer et al., 2015a). As discussed earlier, this inversion from the data may not be exactly the same as the true in situ stress drop, but these values are consistent with the shear stress that exists at this depth in the subsurface. Therefore this 3 MPa stress drop is used to determine the rupture area associated with an earthquake of a given magnitude.

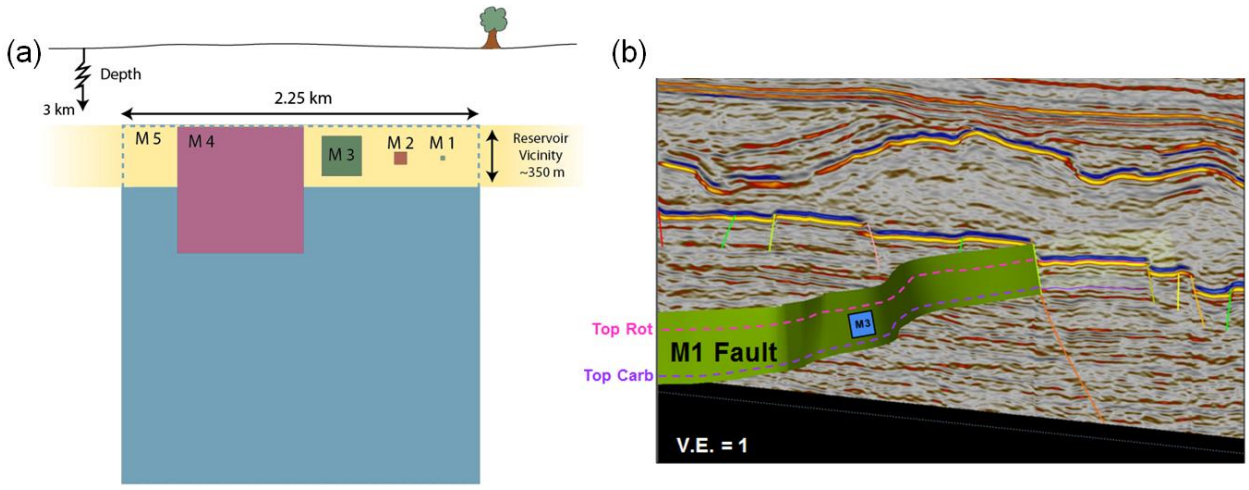

Figure 22. Relative sizes of the reservoir and square fault ruptures of different magnitudes (stress drop of 3 MPa assumed to determine fault dimensions). Fault image from the 3D seismic volume indicating the size of a M3 event on the mapped fault.

Figure 22 compares the square fault rupture areas for a 3 MPa stress drop to the 350 m dimension (larger than the reservoir thickness). Up to a magnitude M_W 3.5 event could be contained within the vicinity of the reservoir. Even allowing for the assumptions of the model shown in figure 22, a magnitude M_W 5 + event cannot reasonably be contained in this zone and must involve a significant portion of slip in the underlying Carboniferous. A M_W $3.5 - 4.5$ event is an intermediate size event that, depending on the stress drop and rupture dimensions assumed, could be confined to the vicinity of the reservoir or involve a significant portion of the Carboniferous.

Table 3. Possible magnitude of events for assumed stress drops and rupture lengths that could be confined to a 350m tall zone. Events smaller than M 3 are in green, events of the Huizinge size or larger ($M_W 3.4, M_L 3.6$) are in red and intermediate sized events are in orange.

	Earthquake Magnitude, M_w				Ribbon-like ruptures Aspect ratio > 6:1		
$\Delta\sigma$ / *length*	250 m	500 m	1 km	2 km	5 km	10 km	25 km
0.1 MPa	2.3	2.5	2.7	2.9	3.2	3.4	3.6
1.0 MPa	3.0	3.2	3.4	3.6	3.8	4.0	4.3
(bf) 3.0 MPa	3.3	3.5	3.7	3.9	4.2	4.4	4.6
10 MPa	3.6	3.8	4.0	4.2	4.5	4.7	5.0

Mw <= 3.0
3.0 < Mw < 3.4
Mw >= 3.4
Unlikely Scenarios

For a range of stress drops and rupture lengths, table 3 shows the resulting magnitude for a given rupture dimension, w, of 350 m. If a 400-450 m tall dimension was assumed the resulting slip and

rupture area would be larger, increasing the magnitudes reported in table 3 by 0.1-0.2, respectively. Events of magnitude greater than or equal to the largest event observed to date (M_W 3.4) are colored in red. The abundance of scenarios allowing this magnitude makes it clear that all events to date are consistent with the assumption of confined events. Additional examination of the Huizinge event (Appendix B) is also consistent with this statement because there is no data definitively showing that a rupture has left the reservoir zone.

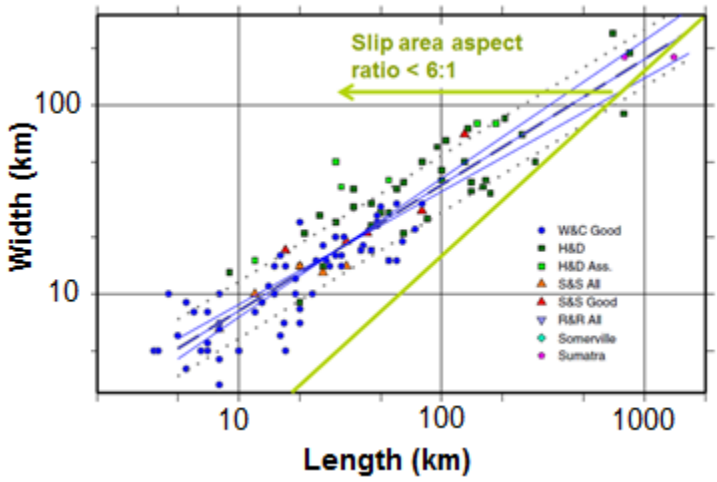

Figure 23. Observations of fault rupture length vs. down-dip width for dip-slip faulting events from Leonard (2010). An aspect ratio of 6:1 fully encompasses the observations at the length scale of the Groningen field.

The right three columns of table 3 would require an extremely thin and long rupture area but slip patch dimension aspect ratios of 1000:1 are not observed for dip-slip faults (Leonard, 2010) (figure 23). All well-constrained observations of dip-slip fault rupture area aspect ratios are bounded by 6:1, at down-dip width dimensions < 90 km (green line in figure 23). In fact, 6:1 is a conservative upper bound to the observations and this value is only used because there is a possibility that more "ribbon-like" ruptures could occur in the induced setting. In strike-slip settings, extreme ribbon-like ruptures occur because the down-dip width of the fault area is limited by the base of the seismogenic zone. When this limit is reached, the rupture continues to expand laterally and a large aspect ratio develops. Here, if the fault is confined to the vicinity of the reservoir, the Carboniferous could be acting in the same way as the base of the seismogenic zone, forcing the lateral expansion of the rupture until it loses energy. However, a laterally propagating strike-slip event is a mode II propagating crack (shear slip in the direction of propagation) while a laterally expanding dip-slip event is a mode III propagating crack (shear slip perpendicular to the direction of propagation). Mode III propagation is less efficient than mode II propagation which may inhibit extreme ribbons from developing.

This 6:1 upper bound on the fault aspect ratio assumption plays a significant role in the maximum magnitude determination. Based on this ratio, for a 350 m down-dip fault dimension, rupture lengths longer than 2 km are not expected (labeled "unlikely scenarios" in table 3). If rupture lengths longer than 2 km are ignored, all events are less than a M_W 4.5, making 4.5 a reasonable truncation magnitude given the uncertainty and range of parameter inputs.

4.3 Relating compaction to earthquake magnitude

The discussion in the previous section specifies a fault area, and then uses the stress drop as a parameter to determine the amount of slip. An alternative approach considers the compaction of the reservoir when determining the amount of slip that could occur. Once slip and area are known, the moment magnitude is easily calculated. Figure 24 illustrates a simple conceptual model, and with some approximations, an earthquake magnitude can be estimated. In the case of the vertical fault shown in figure 24, the left side of the reservoir is undergoing depletion and will compact by an amount, Δz, in response to this depletion. However, in this model the fault is locked so fault slip cannot occur to accommodate the change in vertical dimension associated with the compaction (the pre-earthquake blue dashed line). It is assumed that in the ensuing earthquake the full amount of slip needed to accommodate the compaction occurs (resulting in green post-earthquake dashed line).

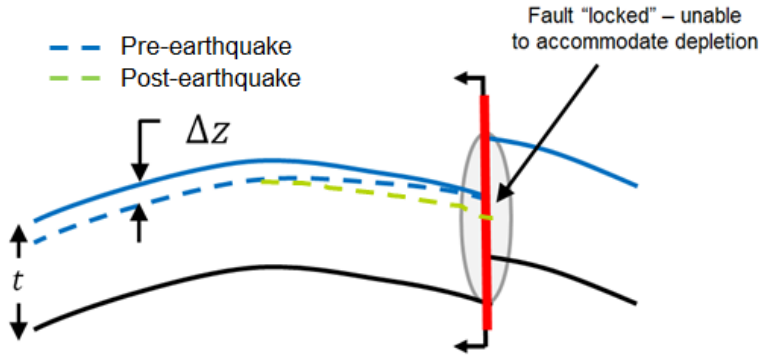

Figure 24. Cross-section schematic representation of compaction resulting in slip on a vertical fault.

The amount of reservoir compaction, Δz, is proportional to the pore pressure drop due to depletion, Δp. The two quantities are related by the Poisson's ratio, ν, the shear modulus, *G*, and the thickness of the interval considered, t.

$$\Delta z = t\frac{\Delta p}{2G}\left(\frac{1-2\nu}{1-\nu}\right)$$

350m is used as the height dimension, but this is an approximation to the geometry of the problem (the fault dip and the amount of pre-existing fault offset would alter the vertical dimension of the slipping patch). Additional assumptions are that all compaction is released as slip and that the slip is uniform over the slipping patch.

By substituting $d = \Delta z$ as the fault slip in the seismic moment equation, moment is once again found to be independent of the shear modulus.

$$M_o = G \cdot (tL) \cdot t\frac{\Delta p}{2G}\left(\frac{1-2\nu}{1-\nu}\right) = t^2 L\,\frac{\Delta p}{2}\left(\frac{1-2\nu}{1-\nu}\right)$$

where L is the fault length. Figure 25 shows the moment magnitude (calculated from the seismic moment) as a function of the length of the slipping patch considered and the pressure depletion ($\nu = 0.2$ assumed). As discussed in the previous section, fault lengths longer than 2 km are not

admissible based on the upper limit on fault aspect ratio. At 300 bars of depletion (more than the projected end of field life depletion) the maximum moment magnitude that could be expected is approximately a M_W 4.3, once again making 4.5 a reasonable upper bound truncation magnitude.

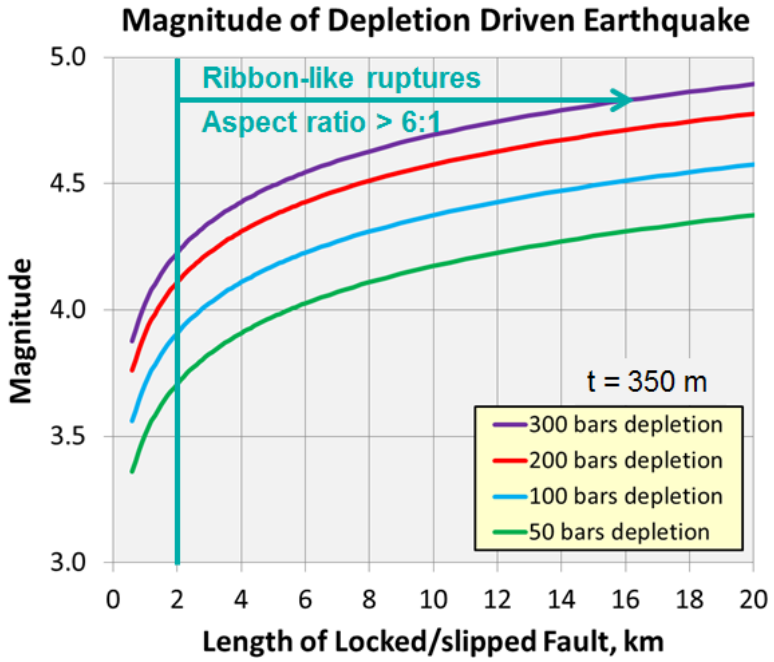

Figure 25. Earthquake magnitudes vary with fault length and reservoir depletion for a simple conceptual model. Earthquakes should be less than Mw 4.5 after depletion given the assumptions stated in the text, including a 2 km maximum length (all lengths greater than 2 km are considered inadmissible).

4.4 Magnitude of earthquakes determined from a 3-D fault-based geomechanical model

A more complete representation of the concepts presented in the previous section involves the modeling of slip evolution on explicitly modeled fault surfaces. ExxonMobil has used the commercial finite element program, ABAQUS, to develop a quasi-static 3D geomechanical model covering large portions of the field and including ~90% of the faults mapped in those areas (Lele, 2016). The 3D model imposes pore pressure changes in a global model that does not explicitly include faults (reservoir layers are draped across faults to approximate the pre-existing fault offsets) and extends far beyond the boundaries of the field. The prescribed pore pressures are taken from the reservoir simulation model that is history matched to the production and forecasts future pore pressure changes for different production scenarios. The deformations calculated in this global model are then applied as boundary conditions to three overlapping submodels that cover domains smaller than the field and explicitly include the faults. The faults are modeled as contact surfaces that are able to slide past one another. The complex geometry (e.g. surface roughness) cannot be fully captured in these submodels, but the spatially variable average strike and dip of the fault are well represented. Both the global model and the submodels have porosity (location) dependent elastic moduli. The porosity variation throughout the field is prescribed by the geologic model and the moduli dependence on the porosity is determined from lab data. Figure 26 is an example of the slip magnitude that accumulates on the modeled faults in one of

the submodels. This is a much more accurate representation of the field geometry than the vertical fault model of the previous section.

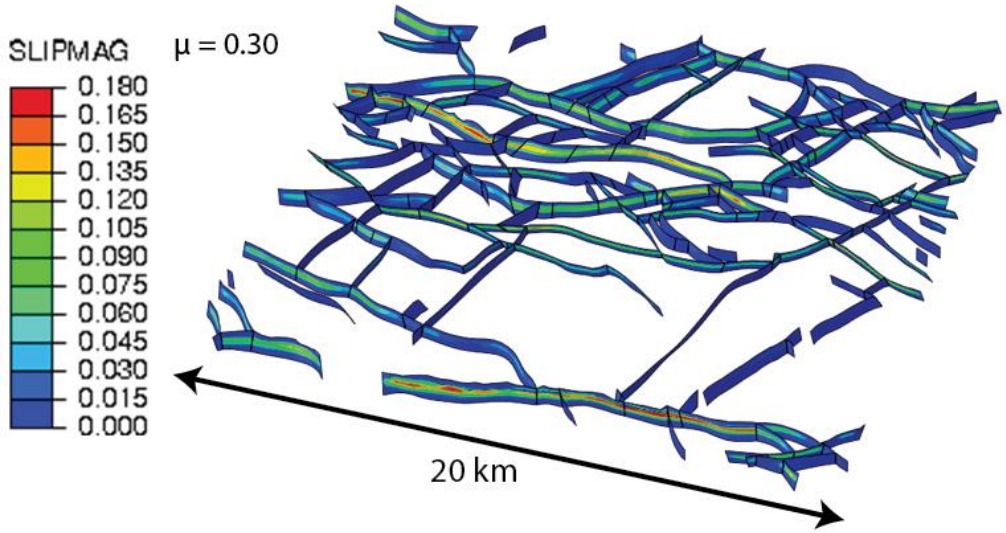

Figure 26. Oblique view of one of three geomechanical submodels. Contoured value is the magnitude of slip for the case of $\mu = 0.35$.

The submodels cover the portion of the field with the most seismicity as well as some portions of the field with relatively low seismicity. The faults included in the model can be seen as the black lines in figure 27. The colored box outlines show where results are taken from each of the three submodels and stitched together to form one dataset. The submodels span the reservoir interval and faults extend ~200-300m below into the Carboniferous (varies throughout the model domain). Since faults only exist in the vicinity of the reservoir, slip cannot propagate deep into the Carboniferous, making the results of this model fall within the definition of confined events.

This quasi static model outputs slip at each node on the fault surface and how that slip evolves with time. A fault area is associated with each node, so in conjunction with an assumption for shear modulus (20 GPa), the seismic moment release can be determined. It is not physically reasonable for all fault slip on all faults to be simultaneously released throughout the field, so all moment released is not summed to determine a maximum magnitude. The 2 km fault length dimension can be used to determine a spatial variation in the maximum moment. At each point in space, all seismic moment released on a fault within a 1 km radius of the observation point is summed and converted to a moment magnitude. The resulting map is a diffuse version of the fault map (figure 27b). The diffusion is due to the summing of moment at an off fault observation location. While the spatial distribution is informative, the peak value (less than M_W 4.5) provides the most insight.

The model chosen for this analysis is a conservative case for fault slip. All fault slip that occurs between the onset of production and the end of field life (2060) is released in one event. A high shear modulus is assumed (12 GPa may be a more representative value for the reservoir sands based on the P and S wave velocities inferred from the seismic data acquisition). Additionally, a low coefficient of friction is

used ($\mu_s = 0.15$) so that all slip that would occur dynamically occurs quasi-statically. This is an approximation because dynamic effects can result in different slip distributions than a quasi-static determination would find. This difference should not change the result by more than ~0.1-0.2 magnitude units and the other conservatism should offset this effect. Given the model chosen for this analysis, M_W 4.5 is a conservative upper bound.

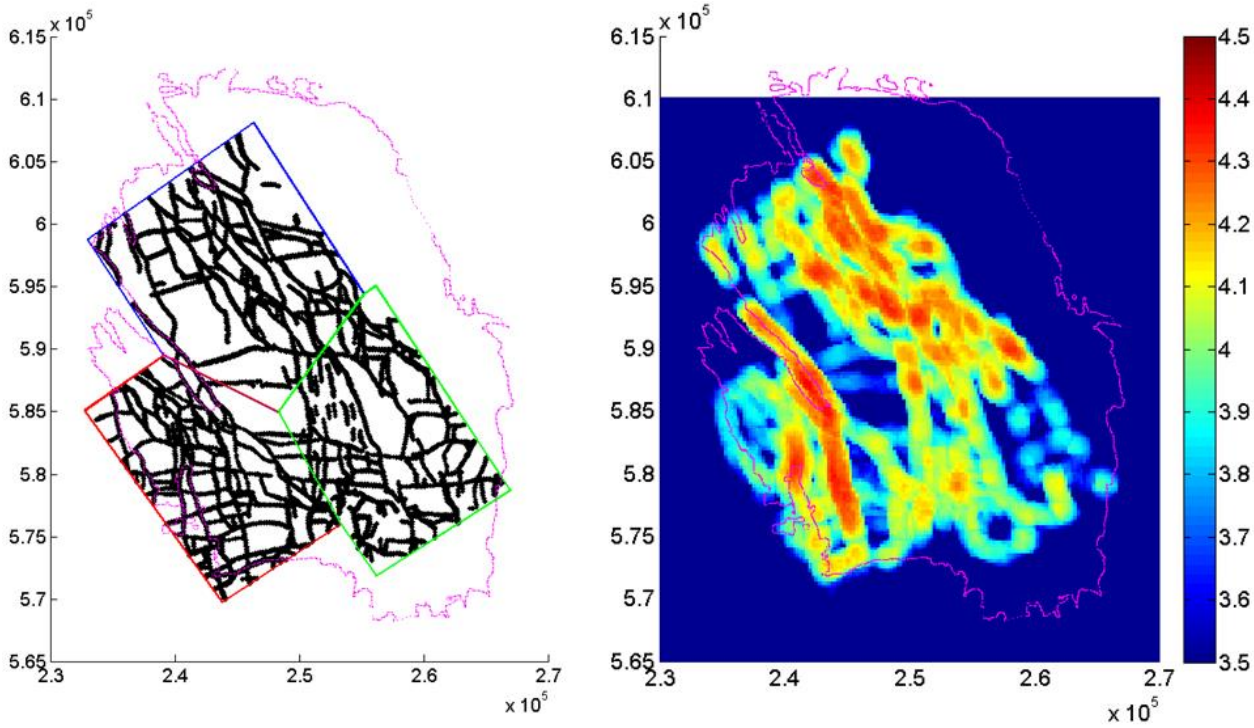

Figure 27. (a) Map of the faults included in the geomechanical submodels with axes in meters (the global model extends well beyond the domain shown here). Magenta line – field outline; black lines - location of faults included in the three sub-models; red, blue and green lines – lines used to stitch the three submodels together (b) Heat map of the magnitude of an earthquake that results from all modeled slip through 2060 in a 1 km radius being released in one earthquake event. Peak magnitudes are less than M_W 4.5.

4.5 Summary: Maximum magnitude of confined events

If a rupture cannot propagate deep into the Carboniferous because of unfavorable conditions (estimated to be likely ~80% of the time), it is termed a confined event. This does not mean that slip on the fault is confined to that area where the reservoir sand is in contact with the fault, but rather that the rupture cannot propagate out of the immediate vicinity of the reservoir. Fault scaling relations along with geomechanical model estimates, all indicate that a M_W 4.5 event is a conservative upper bound and a reasonable truncation magnitude (Mmax) to apply. A key assumption in this determination is that given the limited dip-dimension of a rupture (~350 m), a confined event should not be longer than 2 km in the strike dimension. To date, all ruptures in the Groningen field have been small and are unlikely to have left the reservoir vicinity, making them confined events.

5.0 Triggered out-of-reservoir events

If events are not confined to the vicinity of the reservoir then the whole fault area must be considered in determining the maximum magnitude of earthquakes. The earth is stressed so there is elastic strain energy in the system that could be tapped into and released during an earthquake. The resulting event would release both pre-existing tectonic strain energy as well as the strain energy that was added to the system via depletion.

5.1 Fault-length-based correlations

To determine the earthquake magnitudes that could occur on the existing fault structures, observations of tectonic earthquakes are used (e.g. Wells and Coppersmith, 1994; Leonard, 2010) since equivalent correlations for induced events do not exist. As an example, figure 28 shows how the magnitude of an earthquake varies with the sub-surface rupture length of the fault. Subsurface fault lengths can be determined from the faults mapped by NAM, but due to the resolution of the seismic data, it is not possible to map many of the faults to their down-dip terminus. Offsets less than 20 m are difficult to detect and therefore rarely mapped. Based on the assumption of a 100:1 ratio of fault length to maximum fault throw, ~2 km of fault length could be added to the mapped fault lengths (1 km at each fault tip). However, many mapped faults truncate at other faults so this fault length addition is only valid in some situations.

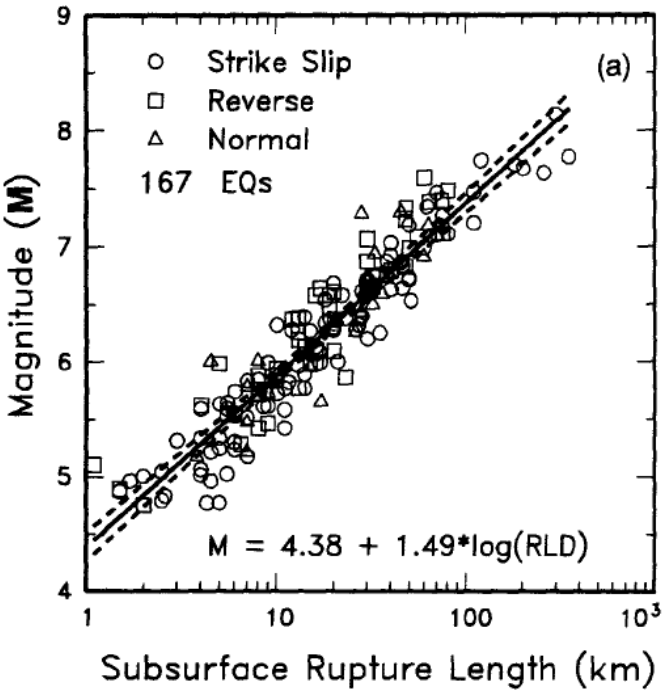

Figure 28. Figure from Wells and Coppersmith, 1994 illustrating the relationship between subsurface rupture length and earthquake magnitude. If the full length of a fault is used as the rupture length, this relationship can specify the size of event that would occur.

The location and lengths of some of the largest faults in the field are shown in figure 4b. 88% of mapped faults are shorter than 5 km, but the longest faults are on the order of 22-23 km. The best fit

relationship from Wells and Coppersmith for dip-slip faults indicates that a 23 km long fault is capable of producing a M_W 6.4 event (M_W 6.9 at the 1 sigma level). Based on the best-fit line, faults longer than 12 km are capable of producing a M_W 6.0 + event and faults shorter than 5.6 km can at most produce a M_W 5.5 event. Ruptures are able to propagate though geometric complexities and link multiple fault segments together, but here this is assumed to be an unlikely scenario so only single fault ruptures are considered.

Figure 29 shows the percentage of mapped fault length that exists on faults of a certain length range. 9% of fault length exists on structures longer than 12 km (capable of producing an event between Mw 6.0 and 6.5), 68% of fault length exists on structures shorter than 5.6 km (maximum Mw 5.5) and 23% of fault length exists on intermediate length structures capable of up to a M 6.0 event. Due to resolution and limitations there is a lack of very short faults mapped (faults < 2 km) so in reality there exists more fault length than is captured in the model. The implication of not capturing all the small faults is that the length percentages in all categories longer than 2 km are actually upper bounds.

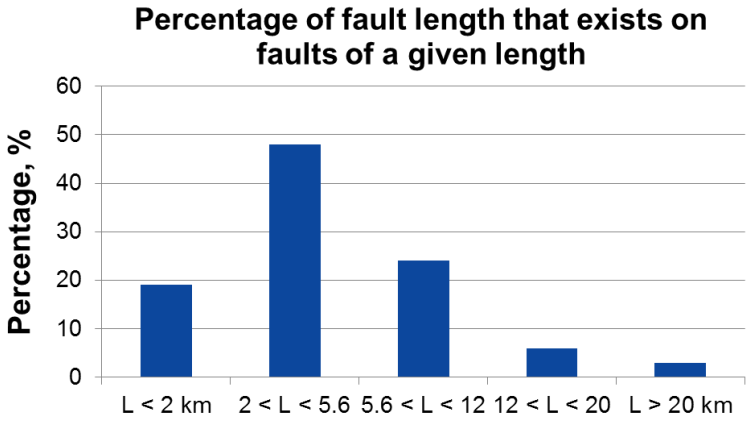

Figure 29. The mapped faults in the fault model can be categorized by fault length, L. Due to seismic data resolution limitations there are fewer short faults mapped. In reality, the L < 2 km bin should be the largest.

To determine these percentages some QA/QC of the fault interpretation is required because some short faults in the fault model are actually part of a larger fault. They can be identified as faults with a throw to fault length ratio that is unrealistically high (~0.01 is an average number). Faults with a ratio >0.06 were deemed to be part of a much longer fault and a pseudo-length was determined based on the throw and an assumed ratio of 0.03. The pseudo length determines the correct bin for the fault length, but to avoid double-counting fault length, only the original mapped fault length quantity is assigned to the bin in determining the percentages (e.g. a 18 km fault is mapped as two faults: segment A is 15 km in length and segment B is 3 km in length. Data QC results in a pseudo length of 14 km for segment B so both segments A and B belong in the 12 < L < 20 km bin which is accurate because the original fault is 18 km long. There is 29 km of total pseudo length and only 18 km of actual fault length, so the pseudo length is ignored in the summation of total fault length).

5.2 Fault-area-based correlations

As an alternative to the fault length - magnitude relationship, a fault-area/stress-drop-based relationship can be used to evaluate the magnitude. Similar to the discussion in the confined event section, an equation can be used to relate stress drop to the seismic moment of an earthquake. The down-dip fault dimension, w, is fundamental in determining the seismic moment.

$$\Delta\sigma = \left(\frac{8}{3\pi}\right) * \left(\frac{G\,d}{w}\right) = \left(\frac{8}{3\pi}\right) * \left(\frac{M_o}{A\ w}\right)$$

Based on the imaging of faults in the seismic data, to determine the possible size of earthquakes, a 3 km and an 8 km down-dip width are considered. The 8 km dimension is consistent with a 10 km seismogenic zone depth.

The sizes of earthquakes that could occur for various stress drops, fault lengths and down-dip fault dimensions are shown in table 4. Some earthquake ruptures are of an unlikely rupture dimension (>6:1 or < 1:1 aspect ratio) and are labeled "unlikely scenarios". Additionally, the 10 MPa stress drop is less likely to occur because most faults have less than 10 MPa of shear stress resolved on them at depth. This determination is made based on the estimates of the stress state in the Carboniferous (only 2.5% of realizations that have a stress drop have 10 MPa or more).

Table 4. Moment magnitude of possible earthquakes for 3 or 8 km down-dip dimension for a variety of stress drops and fault lengths. Some fault length:width ratios are unlikely and are shaded a darker color. Given the state of stress in this area at depth, a large 10 MPa stress drop is less likely than a smaller stress drop.

Earthquake Magnitude, M_w

3 km down-dip extent

$\Delta\sigma$ / *length*	2 km	5 km	10 km	25 km
0.1 MPa	4.2	4.4	4.6	4.9
1.0 MPa	4.8	5.1	5.3	5.6
(bf) 3.0 MPa	5.1	5.4	5.6	5.9
10 MPa	5.5	5.8	6.0	6.2

8 km down-dip extent

$\Delta\sigma$ / *length*	2 km	5 km	10 km	25 km
0.1 MPa	4.7	5.0	5.2	5.5
1.0 MPa	5.4	5.7	5.9	6.1
(bf) 3.0 MPa	5.7	6.0	6.2	6.4
10 MPa	6.1	6.3	6.5	6.8

Unlikely Scenarios

$M < 4.5$
$4.5 \le M < 5.0$
$5.0 \le M < 5.5$
$5.5 \le M < 6.0$
$6.0 \le M < 6.5$
$M \ge 6.5$

Using the range of possible magnitudes in table 4 and the observed fault lengths, a fault-area-based relationship can be established to determine the percentage of earthquake ruptures that could occur on faults capable of a range of maximum magnitudes. Unsurprisingly, the resulting percentages are similar to those found by the fault length correlation. 63% of earthquakes would occur on a fault theoretically capable of up to a magnitude 5.5 event, 23% would be capped by a 6.0 and 14% would be capped by a 6.5. If larger down-dip fault dimensions or a 10 MPa stress drop were considered, the Mmax would increase to consider a Mw 7.0, and a higher percentage of faults would be capable of larger magnitude (see Appendix C for a discussion of the variation in these percentages).

5.3 Representation in PSHA analyses

The maximum event size for a fault is an input to a PSHA analysis. Some faults in the Groningen field are large enough to host an earthquake up to a Mw 6.5 but many faults are smaller could host no more than a Mw 5.5. One way to represent this variability in a PSHA analysis is by using a fault-based approach where each fault is assigned an activity rate and a maximum magnitude. The PSHA implementation currently used for the Groningen field is not a fault-based approach because location errors and the large number of faults make it difficult to assign past earthquakes to individual faults and determine an activity rate. An area source analysis is currently implemented for the Groningen field with spatial variations in the activity-based on the underlying inputs of the model (i.e. variations in compaction). In this implementation, the same maximum magnitude choice is applied to the entire area.

If a maximum magnitude of 6.5 is used, any earthquake that occurs in the field is capable of becoming a M 6.5 event. However, from the fault map this can clearly be seen as an overestimation of the hazard since there are only a small number of faults that could produce an earthquake of that magnitude. Figure 30 schematically illustrates how the Groningen fault system could be represented in an area-based PSHA analysis. There are faults of multiple sizes with different maximum magnitudes and this can be represented as three overlying area sources with different maximum magnitudes. The relative activity rates applied to each area source should be based on the percentage of earthquakes that are likely to occur on faults of different sizes.

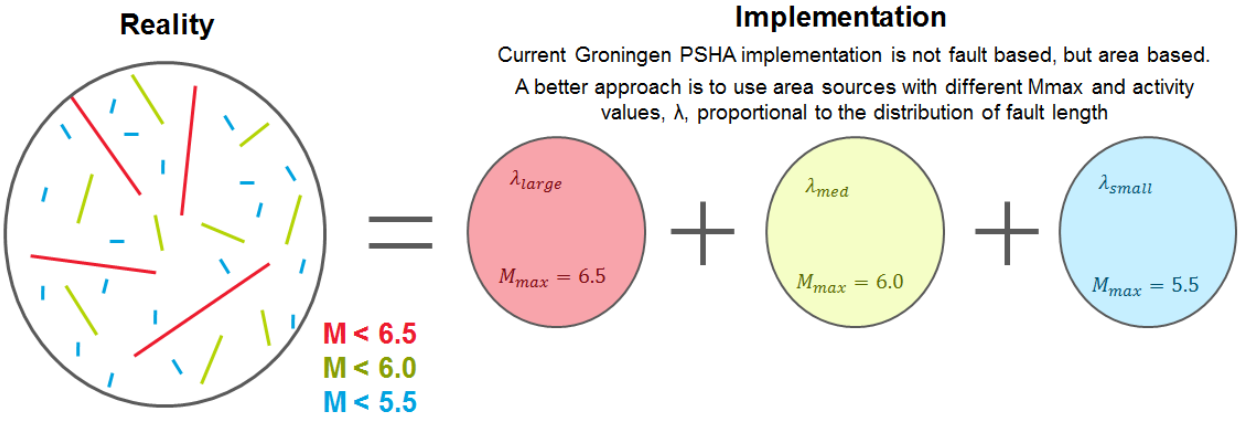

Figure 30. The Groningen system has faults of different lengths, capable of different magnitudes. Ideally this variability would be captured by area sources with different maximum magnitudes and different activity rates.

Large, long faults have a much larger fault area because they extend deeper into the crust, but this down-dip extent is not relevant for the earthquake nucleation. Earthquakes are believed to be nucleating in the reservoir, making the fault area available in the reservoir the relevant quantity for scaling. Since all faults extend the thickness of the reservoir, the quantity of interest can be reduced to the fault length available on small vs. medium vs. long faults. If any given patch of fault is equally likely to host an earthquake (regardless of whether it is part of a small or large fault) then the activity rate for the large magnitude area source should be weighted by the percentage of fault length that exists on long faults. Likewise, the activity rate for the small magnitude area source should be weighted by the percentage of fault length that exists on small faults. Based on the description in the previous sections, if the Wells and Coppersmith best fit relations are used, then 9% of the activity should occur on faults with a Mmax of 6.5, 23% with an Mmax of 6.0 and 68% with an Mmax of 5.5. The equality of activity on small versus large faults is an assumption and counter observations to this could change the results presented here.

The correct representation of the Groningen system with area sources is using three overlapping area sources with different activities, but the result is mathematically identical to a single area source representation with a distribution of maximum magnitudes applied (see figure 31). While this co-mingles the representation of epistemic and aleatoric uncertainty, the final hazard calculation is the same.

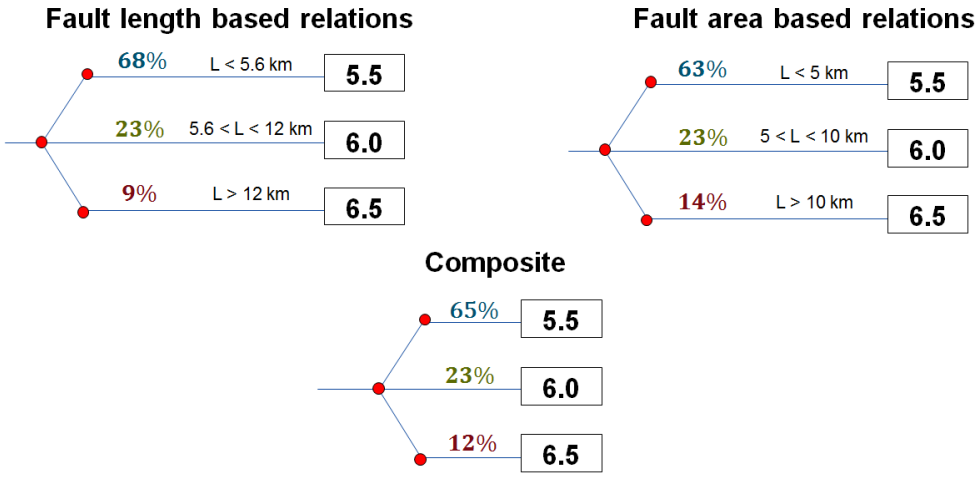

Figure 31. Few faults are large enough to host a large magnitude event, therefore the probability of applying a given truncation magnitude should depend on the relative abundance of faults capable of producing that magnitude event. Different scaling relationships will result in different relative likelihoods of the truncation magnitudes.

5.4 Summary: Maximum magnitude of unconfined (triggered) events

Tectonic earthquake scaling relations can be used to determine the size of events that are possible in the Groningen filed if entire fault areas were to rupture. Some faults in the Groningen field are capable of hosting an earthquake up to a Mw 6.5, but many faults are smaller and could host no more than a Mw 5.5. If any given patch of fault is equally likely to host an earthquake (regardless of whether it

belongs to a small or large fault), then the activity rate for a given Mmax should be weighted by the percentage of fault length that exists that is capable of that given Mmax.

6.0 Conclusions

A standard component of the PSHA analysis is the maximum *possible* earthquake magnitude. The analysis presented in this report is intended to document the probability that a specific maximum possible magnitude is the correct maximum magnitude. Given the relatively small number of earthquakes that may occur during the life of the Groningen field, it is very unlikely that the largest possible earthquake that could occur will occur, therefore the largest *expected* magnitude is smaller than the largest *possible* earthquake magnitude.

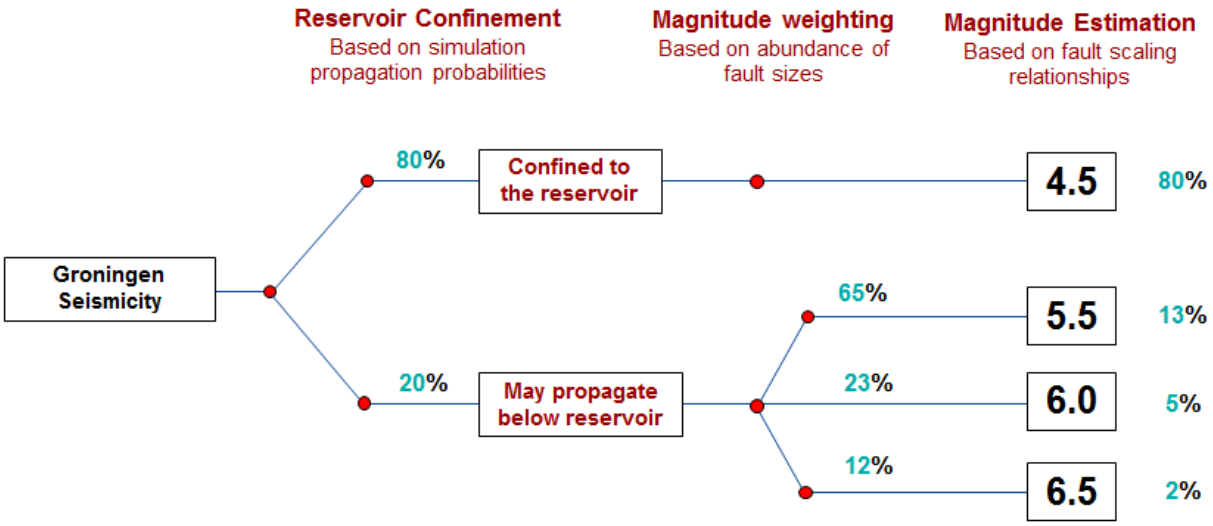

Figure 32. Logic tree representation of the probability that a given maximum magnitude should be applied to any given earthquake.

The largest possible event that could occur is the maximum magnitude, also referred to as the truncation magnitude for the Gutenberg-Richter relationship. The truncation magnitude is not affected by the production because it is determined by the fault size (unaffected by production) or the stress state below the reservoir (where production induced stress changes, like stress arching, are second order to the other effects considered). Based on the analyses presented here, figure 32 as the probability that a given truncation magnitude is the relevant truncation magnitude.

Earthquakes nucleate in the reservoir due to depletion-induced stress changes. Once nucleated, ruptures can propagate into lower loading environments (i.e. out of the reservoir area), but the stress state below the reservoir may be too low to sustain rupture. Given the uncertainty in the state of stress away from the depleted interval and the variation throughout the field, it is found that ~**20%** of realizations in a Monte Carlo simulation have the potential to allow a rupture to propagate away from the vicinity of the reservoir. This does not mean that 20% of ruptures will propagate out of the reservoir vicinity, it means that in 20% of the cases, an earthquake could propagate out of the vicinity. Heterogeneity not accounted for here will still result in many more small earthquakes than large ones.

To date, all observed events have been small and are unlikely to have propagated out of the reservoir vicinity, supporting a low truncation magnitude Gutenberg-Richter.

If a rupture cannot propagate deep into the Carboniferous because of unfavorable conditions, fault scaling relations and geomechanical model estimates all indicate that a M_W 4.5 event is a conservative upper bound and a sensible truncation magnitude, or Mmax, to apply. A key assumption in this determination is that given the limited dip-dimension of a rupture (~350 m), a confined event should not be longer 2 km in the strike dimension.

If events are not confined to the vicinity of the reservoir then the whole fault area must be considered in determining the maximum magnitude of earthquakes. In this case, tectonic earthquake scaling relations can be used to determine the size of events that are possible in the Groningen. Some faults in the Groningen field are capable of hosting an earthquake up to a Mw 6.5, but many faults are smaller and are not capable of hosting an earthquake larger than a Mw 5.5. If any given patch of fault is equally likely to host an earthquake (regardless of whether it belongs to a small or large fault) then the activity rate for a given Mmax should be weighted by the percentage of fault length that exists which is capable of that Mmax.

The distribution shown in figure 32 is highly dependent on the degree of dynamic weakening (the dynamic friction coefficient), which is poorly constrained. Future work could determine that the assumed range was inappropriate and alter the probability that a rupture could propagate deep into the Carboniferous. Additional measurements could be made to measure the current state of stress in the Carboniferous and this also has the potential to change the ~20% determination. Additional factors that could significantly affect this analysis include an observation that events are nucleating away from the vicinity of the reservoir (current assumption is that all events start within the reservoir).

References

Bommer, J.J., P.J. Stafford, B. Edwards, M. Ntinalexis, B. Dost, and D. Kraaijpoel, 2015a, Development of Version 1 GMPEs for Response Spectral Accelerations and for Strong-Motion Durations: Groningen Seismic Hazard and Risk Assessment Model.

Bommer, J.J, B. Dost, B. Edwards, P. P. Kruiver, P. Meijers, M. Ntinalexis, B. Polidoro, A. Rodriguez-Marek, and P.J. Stafford, 2015b, Development of Version 2 GMPEs for Response Spectral Accelerations and Significant Durations from Induced Earthquakes in the Groningen Field.

Di Toro, G., R. Han, T. Hirose, N. De Paola, S. Nielsen, K. Mizoguchi, F. Ferri , M. Cocco & T. Shimamoto, 2011, Fault lubrication during earthquakes, *Nature* (471), 494-498.

Doornenbal, H., and A. Stevenson, 2010, Petroleum Geologic Atlas of the Southern Permian Basin Area, EAGE Publications b.v.

Dost, B. and D. Kraaijpoel, 2013, The August 16, 2012 earthquake near Huizinge (Groningen): KNMI Report, Koninklijk Nederlands Meteorologisch Instituut.

Dost, B., Edwards, B., and J.J. Bommer, 2016, Local and Moment Magnitudes in the Groningen Field: White Paper

Gulden, L. and D. Burch, 2015, Probabalistic estimation of the Gutenberg-Richter *b* value and associated prediction of future seismic events in Groningen, URC Report R027833.

Hanks, T.C. and H. Kanamori, 1979, A moment magnitude scale, *J. Geophys. Res.* (84) 2348-2350

Kraaijpoel, D., and B. Dost, 2013, Implications of salt-related propagation and mode conversion effects on the analysis of induced seismicity, Journal of Seismology, 17, 95-107.

Lele, S., J. L. Garzon, S.-Y. Hsu, N. L. DeDontney, K. H. Searles, P. F. Sanz, 2016, Groningen 2015 Geomechanical Analysis, URC Report xxxxxx.

Leonard, M., 2012, Earthquake Fault Scaling: Self-Consistent Relating of Rupture Length, Width, Average Displacement, and Moment Release, *Bull. Seismol. Soc. Am.* (102), 1971-1988.

Petersen, M.D., Moschetti, M.P., Powers, P.M., Mueller, C.S., Haller, K.M., Frankel, A.D., Zeng, Yuehua, Rezaeian, Sanaz, Harmsen, S.C., Boyd, O.S., Field, Ned, Chen, Rui, Rukstales, K.S., Luco, Nico, Wheeler, R.L., Williams, R.A. and A. H. Olsen, 2014, Documentation for the 2014 update of the United States national seismic hazard maps: U.S. Geological Survey Open-File Report 2014–1091, 243 p., *http://dx.doi.org/10.333/ofr20141091*.

Petersen, Mark D., Charles S. Mueller, Morgan P. Moschetti, Susan M. Hoover, Andrea L. Llenos,

William L. Ellsworth, Andrew J. Michael, Justin L. Rubinstein, Arthur F. McGarr, and Kenneth S. Rukstales, 2016 One-Year Seismic Hazard Forecast for the Central and Eastern United States from Induced and Natural Earthquakes: U.S. Geological Survey Open-File Report 2016–1035.

Segall, P., 1989, Earthquakes triggered by fluid Extraction, *Geology* (17), 942-946

Stein, S. and M. Wysession, 2003, An introduction to seismology, earthquakes, and earth structure, *Blackwell Pub.*, 498 p.269-270.

van den Bogert, P., 2013, Simulation of fault slip by 2-dimensional Finite-Element modelling, Shell internal report, SR.13

van Eijs, R., 2015, Neotectonic stresses in the Groningen Field (Rotliegendes Formation), NAM Internal Report EP201510210531

Wells, D.L. and K.J. Coppersmith, 1994, New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seis. Soc. Am.* (84) 974-1002.

KNMI (Koninklijk Nederlands Meteorologisch Institut) website, 2016, http://www.knmi.nl/kennis-en-datacentrum/dataset/aardbevingscatalogus, accessed January 15, 2016

Appendix A: Determination of the stress state in the Carboniferous

A.1 Total vertical stress

The total vertical stress is the weight of all the material above a given point and can be determined by integrating the density log of wells in the field. In the shallow subsurface, logs are not generally collected so approximations must be made. Figure A1 presents the measurements as equivalent mud weight (EMW) with units of specific gravity. If this value is multiplied by 0.98 (0.1*gravitational constant, g) then the results are in the units of bar/10 m (a stress measurement normalized by the depth) (van Eijs, 2015). After applying this conversion, it is observed that the vertical stress gradient varies from 2.15-2.33 bar/10 m at the bottom of the wells.

The reservoir is at ~3000 m depth and considerable variation in vertical stress gradient can be seen throughout the field. Much of this variation is due to the variable thickness of salt in the field. Most wells penetrate the salt at ~2000 m depth, the point at which a deviation in the EMW is observed due to the low density material encountered.

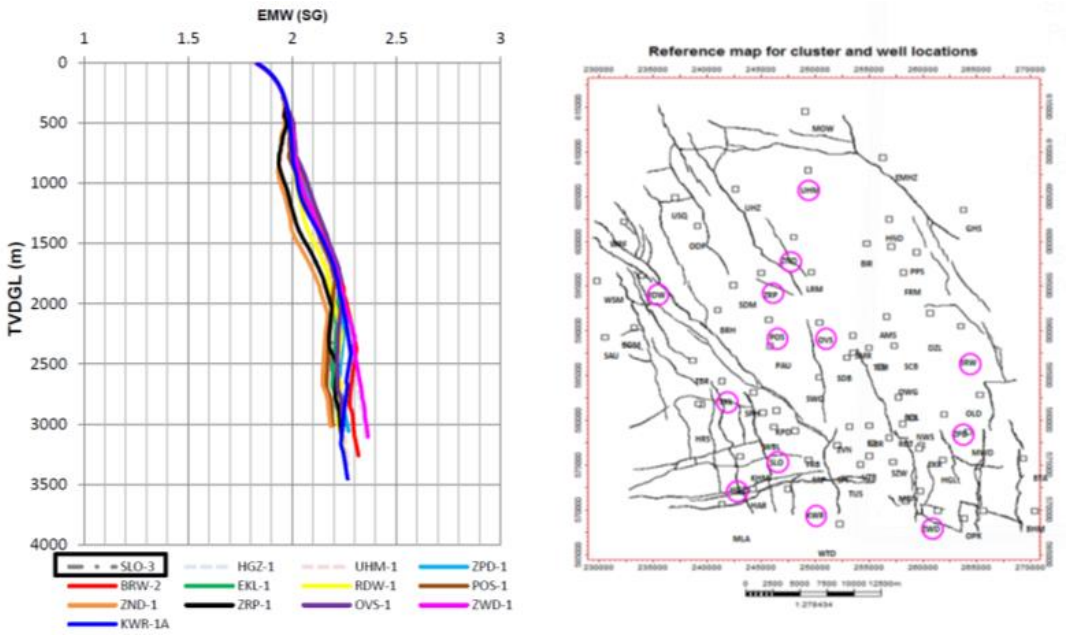

Figure A1. Integral of density logs from various wells throughout Groningen field (values given in units of specific gravity (SG)). Much of the variation is due to differences in salt thickness throughout the field (van Eijs, 2015)

The variation in the field is represented with the histogram shown in figure 16 (histogram with yellow star is the base case distribution). Most wells show a vertical stress gradient in the range of 2.15-2.25 bar/10m, but in figure A1 there are a few points on the higher side. Since the vertical stress is the maximum stress, high values will result in higher shear stresses on the Carboniferous faults at depth and will promote the possibility of out-of-reservoir rupture. The alternate distributions in figure 16 are other possible interpretations of the variation and uncertainty throughout the field that are considered in the sensitivity analysis.

A.2 Minimum horizontal stress in the Carboniferous

To estimate the present day minimum horizontal stress in the Carboniferous, the initial minimum horizontal stress in the reservoir is estimated and an estimate of an increment of stress change is added to that value.

A.2.1 Initial minimum horizontal stress in the reservoir

The only reliable measurements of stress in the reservoir have been made after a substantial amount of depletion. The stress data are presented in van Eijs (2015) and a summary is shown in figure A2 (dark blue diamonds). Five measurements were made after 150-250 bars of depletion (initial reservoir pressure was 350 bars). It is expected that the total horizontal stress should decrease as a result of reservoir production, but this trend is not observed in the data. This is likely due to uncertainty and variation throughout the field, so the data cannot be used to constrain the depletion constant, γ. The change in stress with depletion is due to the poroelastic effect – high pore pressure (P_p) acts to expand the rock, but because the rock cannot laterally expand a high horizontal stress results (S_h). Upon depletion, less horizontal stress is required to keep the rock from expanding laterally, so it is observed that the total horizontal stress decreases as the pore pressure decreases. The relative magnitude of these effects is captured by the depletion constant, γ,

$$\gamma = \frac{dS_h}{dP_p} = \alpha \frac{1-2\nu}{1-\nu}$$

where ν is Poisson's ratio and α is Biot's coefficient. The right hand side of the equation is the poroelastic representation of γ under uniaxial strain boundary conditions (i.e. no lateral deformation). Commonly observed values for the depletion constant in a competent sand reservoir are in the range of 0.5-0.7. ExxonMobil laboratory data to date has measured $\gamma \sim 0.6$ and Shell has measured $\gamma \sim 0.7-0.8$. Reasonable values for elastic parameters ($\nu = 0.2$ and $\alpha = 0.7$) result in $\gamma = 0.52$. 3D geomechanical modeling of the stress evolution in response to the decrease in reservoir pressure (without imposing uniaxial strain boundary conditions) shows that most of the reservoir undergoes a stress path characterized by a depletion constant of $\gamma = 0.6-0.7$. With these constraints, it is likely that γ is bounded by the assumed range of $\gamma = 0.5-0.8$.

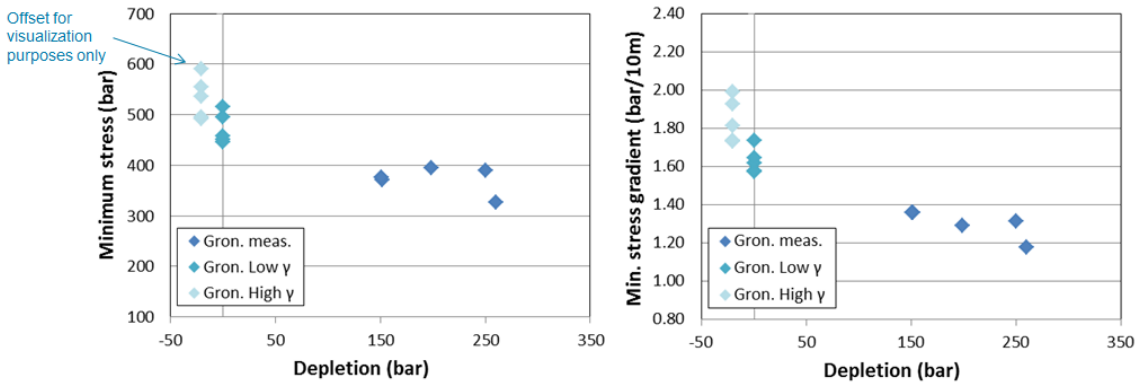

Figure A2. Stress measurements from the reservoir made in a depleted state (blue points). Measurements are extrapolated back to the initial conditions using various assumptions (teal points). Figure on the left shows stress measurements and figure on the right shows the depth normalized values.

Since the depletion is known, by assuming a value for γ the change in horizontal stress from initial conditions to the time of the measurement can be determined. This can then be used to determine the initial value for the minimum horizontal stress. The teal points in figure A2a show the values for the initial state of stress (zero depletion) for the assumption of $\gamma = 0.5$ (dark teal diamonds) and $\gamma = 0.8$ (light teal diamonds). The light teal diamonds belong on the x=0 axis, but they are shown as offset at depletion = -20 bar for visualization purposes. These measurements were taken at different depths so the depth normalized values (in bar/10 m) are also shown in figure A2b.

Based on the data in figure A2, the assumed variation and uncertainty in the initial stress state in the reservoir is shown in figure 16 (distribution with yellow star is the base case distribution). The stress state is likely in the range of 1.6-1.8 bar/10 m, but given the uncertainty in γ, there is also a possibility for larger values. Note that larger values result in a more isotropic stress state and decrease the likelihood of rupture out of the reservoir. To examine the sensitivity of the result to this assumption, additional distributions are considered in figure 16 that allow for higher or lower horizontal stress distribution assumptions.

A.2.2 Stress contrast between the reservoir and the Carboniferous

Present day horizontal stresses should be higher in the Carboniferous than the reservoir due to several mechanisms:

- **Sand-shale contrast** – Horizontal stresses are generally higher in shales (Carboniferous) than in sands (Rotliegend reservoir)
- **Elevated Carboniferous measurements** – Higher horizontal stresses were measured in a Carboniferous sand 40 km to the south
- **Stress-arching** – Reservoir contraction due to depletion can increase the horizontal stress above and below the reservoir
- **Coal** – Coals in the Carboniferous can alter the stress state

Sand-shale contrast

The horizontal stress in shales is generally higher than in adjacent sands, often thought to be due to a contrast in their elastic properties. Both the Poisson's ratio and the Young's modulus of shale are generally higher than that of sand. If a layered medium is loaded from above (i.e. burial) under uniaxial strain boundary conditions, the material with the higher Poisson's ratio will develop higher horizontal stresses (i.e. the shale). Alternatively, if horizontal loading is applied such that a horizontal displacement boundary condition is imposed, the higher horizontal stresses will develop in the stiffer material (i.e. the shale).

There are measurements of the horizontal stress contrast between sands and shales in un-depleted states in nearby fields. The Blija field (65 km to the west) also produces from a Rotliegend reservoir. Measurements made in two sand intervals and two shale intervals at 2750 m depth show a 20-50 bar increase in minimum stress in the shales. Normalizing this difference by depth results in an increase of 0.07-0.18 bar/10 m. Additionally, measurements in the Coevorden field, 50 km to the south, indicate a stress contrast of 50-75 bar at 2825 m depth (0.18-0.27 bar/10 m). Coevorden does not produce from

the Rotliegend reservoir, but rather from a large sand body in the Carboniferous. The shale measurements were made in interbedded shales, not the massive shale surrounding the sand body.

Carboniferous measurements

In addition to the above mentioned measurements in the Carboniferous, additional measurements in the Coevorden sands were made. The Coevorden data can be seen in comparison to the Groningen data in figure A3. The red dots are the measurements (some of which were made in an un-depleted state) and the orange dots are the extrapolation of the depleted measurements back to initial conditions under the assumption of $\gamma = 0.5 - 0.8$.

The Coevorden Carboniferous sand measurements suggest a horizontal stress state that is higher than the Groningen measurements. This could be an effect of higher horizontal stresses in the Carboniferous, geography (Coevordan is further south, closer to the Alps), local structural geology, or simply scatter in the data. The difference suggests an increase in stress in the Carboniferous that ranges from 0-0.15 bar/10m, and since these measurements were made in sand intervals, this could potentially be in addition to an effect of the sand-shale contrast.

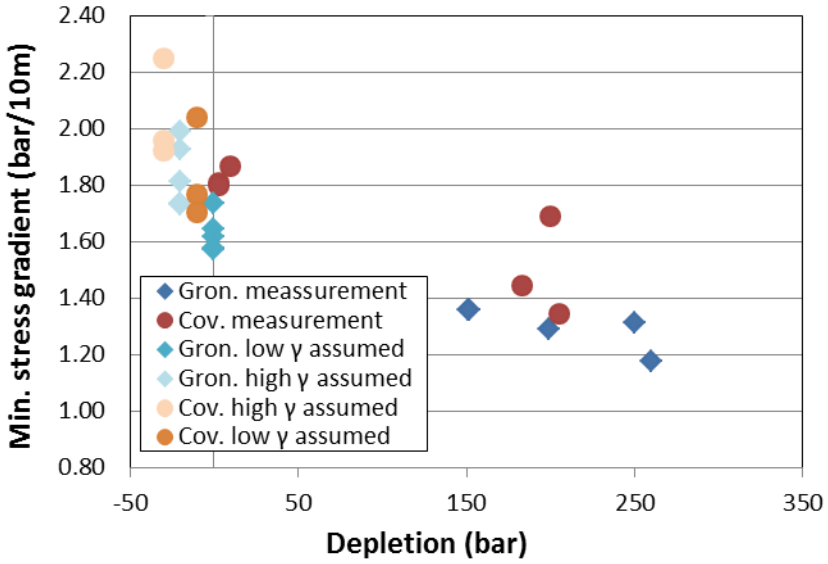

Figure A3. Stress measurements from sands in both the Groningen (blue) and Coevordan fields (red) Measurements are extrapolated back to the initial conditions (teal and orange points) using various assumptions. Measurements are generally higher in the Carboniferous Coevordan field. This may be due to a difference within the Carboniferous or a difference in geography.

Stress arching

The depletion of a reservoir embedded in an elastic medium will alter the state of stress in the surrounding rock. Above and below the reservoir, increases in horizontal stress, known as stress arching, are expected (Segall, 1992). The ExxonMobil global geomechanical model solves for these stress changes. Figure A4 shows the depth normalized (bar/10 m) change in the magnitude of the horizontal minimum stress between 1963 (start of production) and the present day, for the top of the

Carboniferous. Local areas of horizontal stress increase of ~0.07 bar/10 m occur in areas where there are large offsets in the reservoir due to pre-existing faults. However, most of the reservoir undergoes only a small change in horizontal stress, on the order of 0.03 bar/10 m or less. Based on these numbers, stress-arching does alter the stress state, but it is a second order effect.

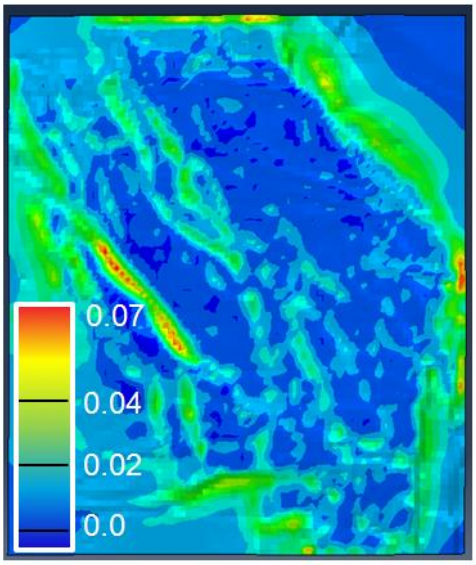

Figure A4. Modeled change in horizontal stress gradient (bar/10 m) from the beginning of reservoir production to present day. Results are shown as a stress change at the top of the Carboniferous.

Coals in the Carboniferous

The Rotliegend reservoir in the Groningen field is unconformably overlying a coal-bearing interval of the Carboniferous Westphalian sequence (WBCL). The limited well penetrations into the Carboniferous suggest that there are coal beds, but there is little lateral continuity and the concentration is likely only 3-5% coal.

In the Piceance basin, which is ~25% coal, a substantial increase in horizontal stress is observed upon entering the coal rich intervals. Coal is a weak unit that cannot support large deviatoric stresses. As a result, the horizontal stress is higher (closer to the vertical stress), making the stress state more isotropic. The depth normalized increase in Piceance is ~0.5 bar/10 m. However, since the coal concentration is much lower in Groningen, the local effect in the area of substantial coals will be much smaller. It is likely a second order effect on the order of 0-0.05 bar/10 m.

Distribution of horizontal stress increases in the Carboniferous

Of all the effects considered, it is clear that the sand-shale contrast is the largest. The distribution shown in figure 16 (base case distribution marked with a yellow star) is used to estimate the stress increase in the Carboniferous. It is consistent with the sand-shale effect and includes ~30% of the increase in the 0.2-0.3 bar/10 m range if multiple effects are additive. Additional distributions are considered (shown in figure 16) that allow for more or less increase in the horizontal stress in the reservoir.

A.3 Pore pressure

While pore pressure in the gas column is well constrained, there are few pressure measurements in the water leg (below the gas-water contact). In the northern part of the field, the Carboniferous is located in the water leg. However, in the southern part of the field, the top of the Carboniferous is in the gas leg (above the gas-water contact). The original pore pressure in the gas column was 350 bar, and due to the low density of gas, this pressure was mostly uniform from the top to the bottom of the gas column. The original gas water contact was at ~2970 m so the depth normalized pore pressure stress gradient is 1.18 bar/10 m. This value should represent the pore pressure gradient at the top of the water column as well. Measurements of the water leg pressure gradient vary throughout the field from 1.15-1.17 bar/10 m. The crest of the Groningen closure in the Loppersum area is at ~2700 m depth, so the depth normalized pressure gradient at the top of the structure is 1.30 bar/10 m. Given that some portions of the top Carboniferous are in the water leg and some are in the gas leg, a value of 1.17 bar/10 m is used as the base case pore pressure gradient. Values ranging from 1.15-1.22 bar/10 m are also considered as variations to examine the sensitivity to this parameter. The high value of 1.22 bar/10 m would be representative of a scenario in which substantially more overpressure (elevated pore pressure) is able to develop within the Carboniferous shale.

Appendix B: Wave propagation simulation of the Huizinge event

A complex waveform was observed in some of the M_L=3.6 Huizinge records near the epicenter (figure B1). Recordings of previous nearby events with comparable magnitudes had much simpler waveforms that did not show multiple, closely-spaced large amplitude arrivals (referred to as the doublet). This complex waveform could be the result of the path (wave refractions and multiples) or rupture source properties (e.g. non-uniform rupture propagation velocity, unilateral propagation, source area geometry, etc.). The experiments described below were designed to help understand the cause of the recorded waveform because it may contain information about whether or not the Huizinge rupture propagated into the Carboniferous, and therefore has implications for the maximum magnitude discussion.

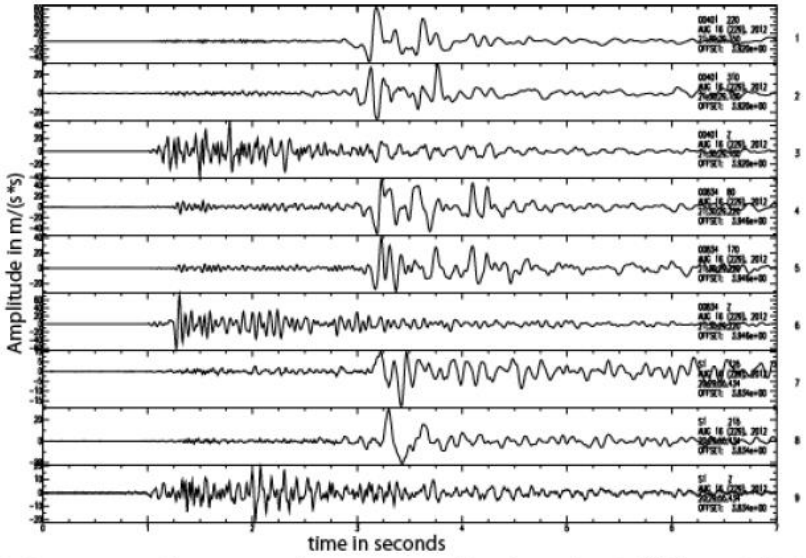

Figure 8. From top to bottom: accelerometer recording in station Middelstum-1 (1-3, radial, transverse and vertical)), Westeremden (4-6, rad, trans., vert.) and Stedum (7-9, rad, trans., vert.) for the August 16, 2012 Huizinge event.

Figure B1. Figure and caption from Dost & Kraaijpoel, 2013

3D elastic finite-difference wavefield modeling was applied to the 3D subsurface velocity model provided by NAM. Point-source simulations were used to determine if the doublet could be the result of an internal reflection off of a velocity contrast (a multiple). These simulations used the KNMI source mechanism that was calculated from a nearby event (a focal mechanism could not be determined from the Huizinge recordings) (Dost & Kraaijpoel, 2013). The intent was to identify arrival phases at the Westeremden station and compare the timing and amplitudes of these multiple arrivals to the doublet signature. The timing of the two pulses in the doublet is approximately equal to the time of an internal reflection within the reservoir, but amplitudes predicted for this arrival phase do not match observations (figure B2a).

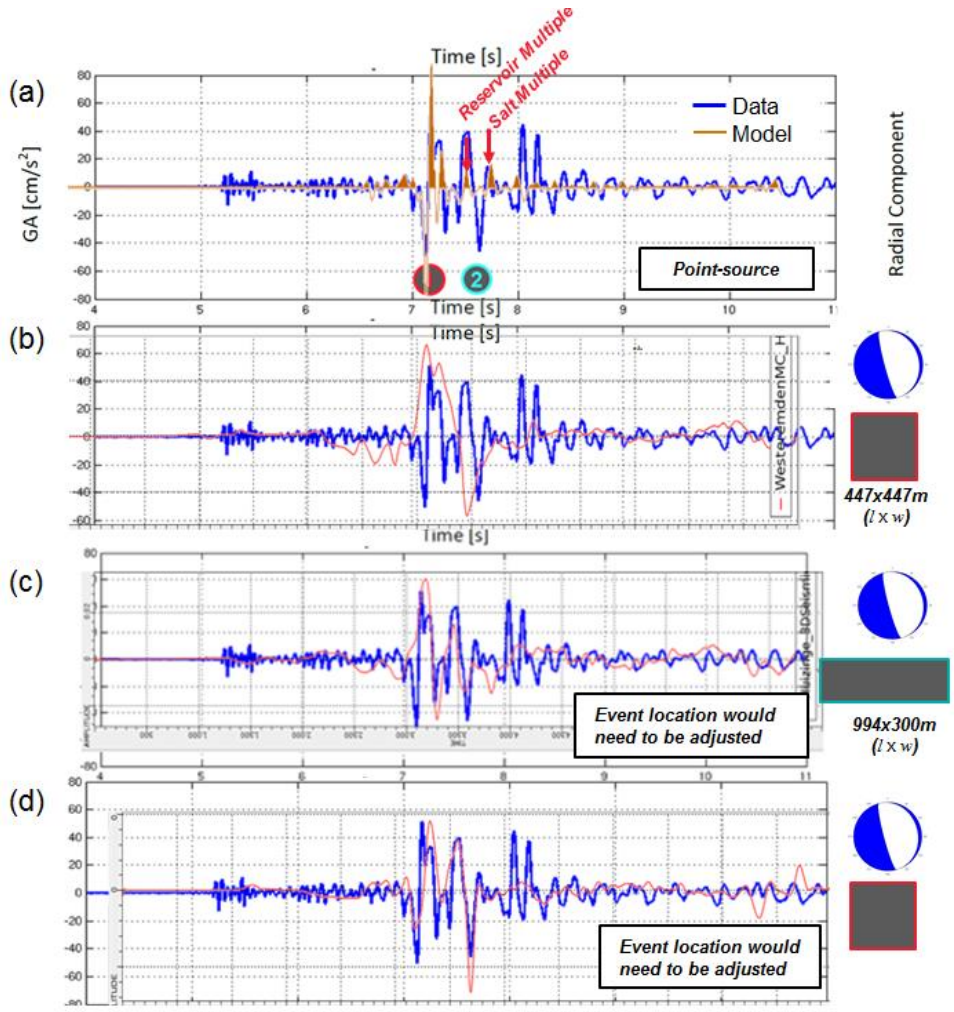

Figure B2. Comparison of simulations (red lines) and recording (blue line) of the radial component recording of the Huizinge event at the Westeremden seismic station. Comparison made for multiple reservoir geometries and a source mechanism based on the fault orientation in the 3D seismic data.

Two source mechanisms were also tested—one calculated by KNMI for a previous event, and one based on the strike and dip of the nearest fault to the calculated Huizinge epicenter, based on the interpreted 3D seismic data. Both sources were then used along with the fault scaling relation for rectangular fault rupture areas (discussed in sections 4 & 5) to design finite sources for wavefield modeling. Finite sources were simulated using an array of discrete point sources, spaced at 22 m to avoid source aliasing up to 20 Hz, at an assumed rupture velocity, V_r, of 80% of the shear wave speed (V_r=1,783 m/s). Two rupture patches with different aspect ratios were examined, both oriented along the same strike and dip as the source mechanism. Using an assumed stress drop of 3 MPa for a $M_W 3.6$ event results in a square rupture patch with 447 m on each side, and a "reservoir-confined" patch of 300 m width, and 994 m length (new data indicates that the Huizinge event was $M_W 3.4$ but at the time of this study the event was thought to be $M_W 3.6$). Initial tests to understand the effects of rupture velocity, hypocenter location uncertainty, and rupture directivity were also performed.

These tests demonstrate that there are several ways to generate a complex ("doublet") waveform without requiring that rupture propagated into the Carboniferous. While many simulated rupture specifications could generate a comparable waveform, none provided a perfect match (figure B2 b-d). The best match overall was obtained using a rectangular rupture with unilateral propagation toward the south, and the hypocenter shifted from the estimated location onto the nearest fault. Through these simulations intuitive effects were observed (e.g. slower rupture velocities produce doublets for smaller rupture areas) and more subtle effects were also observed (e.g. faster rupture velocities tend to increase the peak ground motion because the waveform has a higher-frequency content and signals from the rupture nucleation and termination interfere more constructively).

While the timing of the doublet arrivals is consistent with simulated internal reflection times, the path effects do not reproduce the observed amplitudes so the complex signature appears to be primarily a signature of the rupture source parameters. In the simulation results, the waveform signature shows substantial variation with spatial coordinates near the epicenter—so rapidly, that any single station within a few kilometers of the epicenter cannot be used alone to determine details about the rupture dimensions (figure B3).

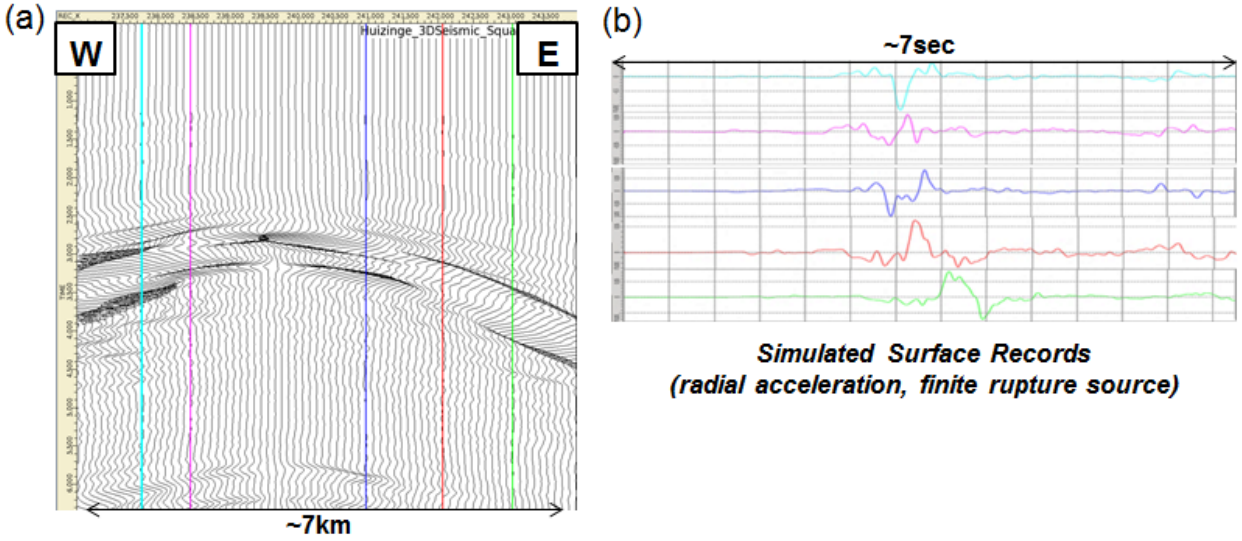

Figure B3. In the epicentral area there is significant variation in the waveform recording with small changes in observation location. (a) Seismic gather of many recordings along an east-west transect (b) A few highlighted recordings to demonstrate the large differences in waveforms.

Multi-component waveforms from several stations should be analyzed together, but there are too many degrees of freedom in the source specification to easily obtain a good fit to all of the data. The factors considered in this study alone include: path effects from the local overburden, source mechanism, source location, rupture aspect ratio, rupture velocity, and directivity. These parameters strongly influence the waveforms and investigation into the problem is inefficient due to a manual approach. Therefore, at this time it is not possible to use the Huizinge event recordings to determine if the rupture propagated into the Carboniferous.

Appendix C: Uncertainty in the triggered event maximum magnitude

There are multiple sets of empirical relationships that can be used to estimate the magnitude of an earthquake on a fault of a given length (e.g. Wells & Coppersmith, 1994; Leonard, 2010). For each, the estimated magnitude will change if the best fit or the 1 standard deviation line is considered. Additionally, if all faults are assumed to be longer than their mapped length (additional 2 km per fault) due to seismic resolution limitations and mapping constraints, then the estimated magnitude will also increase. Figure C1 illustrates the effect of these uncertainties on the percentage of earthquakes that would nucleate on a fault capable of a given maximum magnitude. In addition to the empirical relationships, an analysis based on the stress drop and assumed rupture area can estimate the earthquake magnitude. In section 5.2 it was assumed that a 10 MPa stress drop was unlikely in the Groningen setting. Figure C1 also illustrates the effect of including the 10 MPa stress drop in the magnitude estimation. Accounting for these uncertainties could decrease the probability that the maximum magnitude is 5.5, increase the probability that it is 6.5, and introduce the possibility that the maximum magnitude is approaching M 7.0.

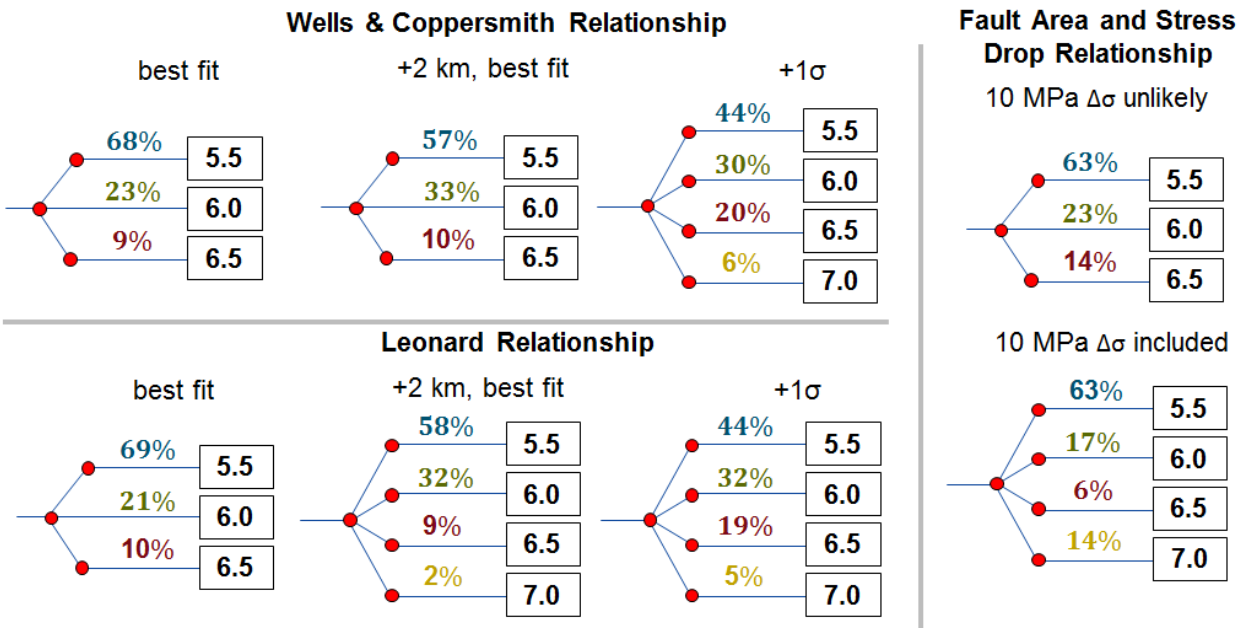

Figure C1. Effect of different empirical relationships, uncertainty, fault length and stress drop assumptions on the maximum magnitude probabilities.