

Probabilistic earthquake locations of induced seismicity in the Groningen region, the Netherlands

Jonathan D. Smith^{1,2}, Robert S. White,¹ Jean-Philippe Avouac² and Stephen Bourne³

¹*Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, UK. E-mail: jdsmith@caltech.edu*

²*Department of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA*

³*Shell Global Solutions International, 1031 HW Amsterdam, The Netherlands*

Accepted 2020 April 4. Received 2020 March 24; in original form 2019 January 14

SUMMARY

The Groningen gas reservoir, situated in the northeast of the Netherlands, is western Europe's largest producing gas field and has been in production since 1963. The gas production has induced both subsidence and seismicity. Seismicity is detected and located using the Koninklijk Nederlands Meteorologisch Instituut shallow-borehole array for the period 2015–2017, incorporating the back projection techniques of QuakeMigrate and the nonlinear location procedure to constrain earthquake locations and depths. The uncertainties on the estimated depths are estimated taking into account velocity model, changes in station array geometry and uncertainties in the measurement of arrival times of the *P* and *S* waves. We show that the depth distribution of seismicity is consistent with nucleation within the reservoir (28 per cent) or in the overburden (60 per cent) within ~500 m from the top of the reservoir. Earthquakes with hypocentres in the overburden likely originate from overlying Zechstein anhydrite caprock. Based on their depth distribution, it seems like the earthquakes are primarily driven by the elastic strain in the reservoir and overburden, induced by the reservoir compaction. We estimate the probability of earthquakes nucleating beneath the reservoir in the underlying Carboniferous limestone and basement, to be no more than 12 per cent.

Key words: Earthquake hazards; Earthquake source observations; Induced seismicity.

1 INTRODUCTION

The Groningen gas field, situated in the northeast of the Netherlands (Fig. 1), has been in production since 1963, with 70 per cent of the estimated 2800 billion cubic metres (bcm) of gas now extracted (Bourne *et al.* 2014). Prior to gas extraction the region was considered aseismic with no recorded historical events. However, since the 1990s small magnitude earthquakes have been detected, with these shallow events causing non-structural building damage and some public concern (Dost *et al.* 2012). The last decade has shown an increase in both the mean frequency and the magnitudes of earthquakes, leading to the Netherlands Minister of Economic Affairs instructing a reduction in gas extraction since 2014, with the aim of reducing the frequency of future earthquakes (van Thienen-Visser *et al.* 2015). On 2014 January 17, gas production of the Groningen field was limited to 42 bcm yr⁻¹, in response to the increasing frequency and magnitude of seismicity across the region. The regulator of the reservoir, Nederlandse Aardolie Maatschappij, has released a series of publications investigating the induced seismicity, surface subsidence and geomechanical response from gas production. The publication from Nederlandse Aardolie Maatschappij (2016) summarize the advances

made since the initial Nederlandse Aardolie Maatschappij (2013) publication.

Seismicity is attributed to changes in induced strain caused by the bulk reservoir volume decrease (Bourne *et al.* 2015; Nederlandse Aardolie Maatschappij 2016; Dempsey & Suckale 2017), with this induced strain expected to be dependent on the possible spatial and temporally varying reservoir properties. Accurate determinations of the hypocentral locations of the generated earthquakes are required to understand the geomechanical behaviour of the reservoir in response to gas extraction. The latest seismicity catalogues of Spetzler & Dost (2017) and Willacy *et al.* (2019) are relatively consistent, with epicentral locations typically within 1 km and hypocentral locations within agreement of 300 m. The epicentral location difference is larger than the depth uncertainty due to differences in the imposed regional 1-D or 3-D velocity models and is discussed further in Section 3. Given the typical 300 m thickness of the reservoir, these hypocentral locations are not sufficiently precise to determine whether the earthquakes occur within the reservoir, or just external to the reservoir in the overburden or underburden basement. As a result the mechanism by which earthquakes are induced or triggered remains uncertain. The purpose of this study is to improve the accuracy of the hypocentral

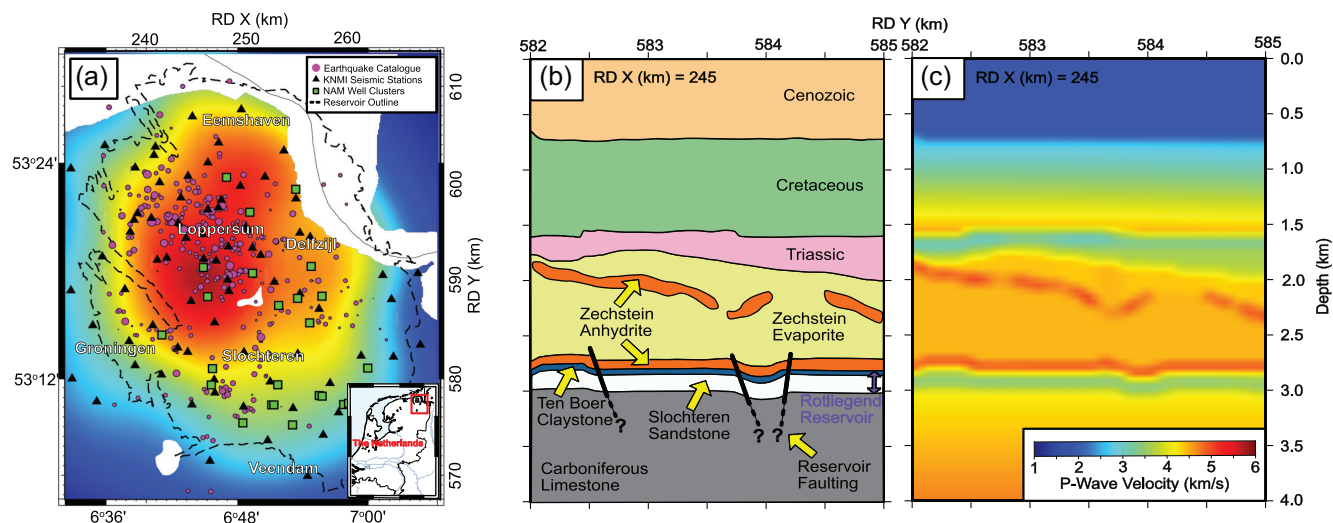


Figure 1. Surface subsidence and seismicity across the Groningen region. (a) Cumulative modelled surface subsidence from 1964 to 2017 (Smith *et al.* 2019). Seismicity between 1964 and 2017, combining KNMI locations from 1964 to 2015 and our catalogue for 2015–2017, shown by pink circles with size scaled by magnitude (maximum magnitude $M_L = 3.1$). Well cluster locations shown by green squares. KNMI network of accelerometer and seismometer locations shown by black triangles. Gas extent within the reservoir shown by the black dashed outline. (b) Schematic cross-section from this region displaying the different lithologies. Black lines represent mapped fault structures from seismic reflection (Nederlandse Aardolie Maatschappij 2013). (c) P -wave velocity along the same cross-section from NAM 3-D velocity model (Nederlandse Aardolie Maatschappij 2017).

locations and evaluate the associated uncertainties and any possible methodological bias.

Hereafter we present an alternative method to determine the earthquake hypocentres compared to those used in previous studies. The results from this procedure are used to investigate the depth distribution of the seismicity and compare it to the reservoir depth range. We then draw implications regarding the mechanism by which earthquakes are induced or triggered, determining whether there is possibility for earthquakes originating external to the depleting reservoir. These findings have important consequences for the maximum possible earthquake magnitude.

2 GEOLOGICAL SETTING

The Groningen field is situated in the Groningen structural high, formed during the late Kimmerian tectonic phase. The subsurface is well known from various geophysical investigations (seismic reflection and seismic refraction), borehole core samples and borehole logging (Fig. 1b). The reservoir is located within the Upper Rotliegend Group, at depths between 2.6 and 3.2 km, with a thickness varying between 100 and 300 m (Nederlandse Aardolie Maatschappij 2013). The Upper Rotliegend comprises the Slochteren sandstone and Ten Boer claystone formation. The gas migrated into these formations from its source in the highly faulted underlying Pennsylvanian Carboniferous limestone which includes coal layers (Fig. 1b). The overlying thick and impermeable evaporite and anhydrite layers of the Permian Zechstein group provide the seal for the reservoir. This caprock formation comprises a laterally discontinuous ~ 50 m thick basal anhydrite and 0.2–1 km thick evaporite deposit with an internal disconnected anhydrite layer of 50 m thickness (De Jager & Visser 2017). Seismic reflection, seismic refraction, sonic log and well core samples have been combined to produce a 3-D-elastic model of the Groningen region derived in Nederlandse Aardolie Maatschappij (2013) and expanded in Nederlandse Aardolie Maatschappij (2017). Sonic log data gave a P -wave velocity structure at well locations, with the relative S -wave and

density determined using the core-sample data. These 1-D velocity profiles were combined with horizon changes determined from seismic reflection and refraction profiles to generate a 3-D velocity model at 50 m spacing. The expected uncertainty in the derived velocity models has been estimated to be ± 5 per cent of the total expected traveltimes (Nederlandse Aardolie Maatschappij 2017). A cartoon schematic representation of the separate lithologies, and a cross-section of the P -wave velocity models are shown in Figs 1(b) and (c).

3 RESULTS FROM PRIOR MICROSEISMIC MONITORING

The Royal Netherlands Meteorological Institute (KNMI) has provided field-wide coverage of the Groningen region for the detection and location of earthquakes since 1995. The array initially consisted of 8 borehole stations, further extended to 14 in 2010, to monitor the induced seismicity of events with local magnitudes $M_{LN} > 1.5$ (Dost *et al.* 2017). However, the large station spacing of this array meant it could produce only poorly constrained epicentral locations with an estimated ± 500 m uncertainty (Dost *et al.* 2012). In late 2014, this network was expanded with the deployment of an additional 59 passive borehole geophone stations, each equipped with 4 three-channel geophones and a surface accelerometer, spaced at 50 m depth intervals between 50 and 200 m depth. The present station geometry is shown in Fig. 1(a), with the variation in seismic array geometry and KNMI epicentral locations displayed in Supporting Information Video S1.

Initial location methods by the KNMI used the P - and S -wave arrival time difference following the HYPOCENTER Method of Lienert *et al.* (2005). This inversion procedure uses a 1-D regional velocity model determined for the whole of the North Netherlands regional, comprising sharp velocity changes that are not necessarily present in the Groningen reservoir (Dost *et al.* 2012). Due to the high uncertainty in the depth estimation only epicentral latitude and longitude were determined, with a default depth of 3 km imposed. The

later method of Spetzler & Dost (2017) expands on this by adopting the equal-differential time (EDT) method of Lomax (2005), incorporating traveltimes determined by full-waveform modelling with a local 1-D velocity model defined at the initial epicentral location, embedded in a coarse 3-D location grid. However, this method neglects lateral 3-D variations in velocity and possible boundary attenuation effects, resulting in high uncertainties in horizontal position and depth. Willacy *et al.* (2019) improves on this processing procedure by deriving locations using 3-D full-waveform inversion, solving for errors in the estimation of location. They find that optimal earthquake locations fall within the Rotliegend reservoir, but their procedure lacks a detection framework and uses earthquakes only from the KNMI catalogue (Dost *et al.* 2012). We expand on this analysis using a probabilistic detection approach to assess the probability of the earthquakes being confined within or locating outside the depleting reservoir interval.

4 METHODS

4.1 Microseismic monitoring

Prior earthquake detection techniques across this region have relied on a short-term-average to long-term-average approach (STA/LTA) for the detection of coherent signals on individual (Dost *et al.* 2012). In a built up region such as Groningen, a station by station STA/LTA can lead to many false triggers from anthropogenic noise. To expand on the prior methods, we adopt a coherence based location procedure as outlined in Drew *et al.* (2013). To that effect, we developed an open source software called QuakeMigrate (which is publicly available at <https://github.com/QuakeMigrate/QuakeMigrate>).

We determine onset functions for each seismic station by using an STA (0.2 s) to LTA (1.0 s) function, representing a form of noise normalization of the original seismic signal. The onset function for the *P* wave is determined from the vertical component of each seismic station and the *S*-wave onset function determined from the combined horizontal components. The onset function is meant to represent the probability density function (pdf) of the timing of an arrival. Drew *et al.* (2013) demonstrated that the onset function can be approximated by a Gaussian distribution. The estimated arrival time picks are determined by the maximum of the onset function, with an associated uncertainty determined as the standard deviation about this peak value. To find the coherence of signals between seismic stations a coalescence function is computed as the sum of the logarithm of the onset functions for the seismic stations. The coalescence function f_c is given by

$$f_c(t, \vec{s}) = \prod f_{r_i}(t + t_{g_i}(\vec{s})) = \exp\left(\sum_{i=1}^n L_{R_i}(t + t_{g_i}(\vec{s}))\right) \quad (1)$$

where t and \vec{s} are the time and space variables, f_{r_i} the onset function for station i , t_{g_i} the traveltime from a spatial location to station i and L_{R_i} is the natural logarithm of f_{r_i} . The coalescence function is a 4-D function dependent on spatial and temporal terms. As such it can be represented by instantaneous spatial maps, $f_c(t, \vec{s}(t))$ at each time step, and the distribution with time of maximum value $\hat{f}_c(t)$.

An event is triggered when the maximum coalescence value through time exceeds a median-averaged deviation from the background. The formulation of the detection threshold is given by

$$T_h = \text{median}(\hat{f}_c) + A * \text{MAD}(\hat{f}_c), \quad (2)$$

where T_h is the detection threshold for the time interval, $\text{median}(\hat{f}_c)$ is the median of the coalescence value during the interval, $\text{MAD}(\hat{f}_c)$

is the median-averaged deviation of the coalescence value during this interval and A is a user defined scaling factor. The detection threshold is determined across periods of 30 min, with the scaling factor A set at a value of 11 to minimize the number of false detections. As the exact origin time of the event is unknown, a marginal time window of 0.5 s is taken about the maximum coalescence value equal to the expected traveltime uncertainties in the model. The event location and uncertainty can be represented by the integration of the instantaneous spatial map over the marginal time window,

$$f_e(\vec{s}) = \int_{t_D - t_a/2}^{t_D + t_a/2} f_c(\tau, \vec{s}) d\tau, \quad (3)$$

where f_e is the event spatial map, t_D is the maximum coalescence time, t_a is the marginal window and f_c is the spatial-temporal coalescence value. The optimal hypocentral location is then given as the location of the maximum event spatial map, with location uncertainty given by the fitting of a 3-D Gaussian function to the event spatial map.

An example of the output from the coherence procedure for the 2015 October 30 18:49:02 $M_{LN} = 2.3$ earthquake is shown in Fig. 2. Fig. 2(a) represents the normalized coalescence for the earthquake location with the higher coalescence values representing a more likely earthquake source location. Fig. 2(b) represents the automated *P*- and *S*-wave arrival picks determined by fitting the maximum onset function for each station around the expected arrival time for each station from the maximum coalescence location.

To improve on the initial earthquake locations detected using the coherence technique, we refine the seismic phase arrivals by manually picking the first arrival of the *P*- and *S*-wave phases. We adopt the nonlinear location method (Lomax 2005), using the first arrival traveltime formulation of Podvin & Lecomte (1991). This approach produces a misfit function representing points drawn from the posterior pdf for the earthquake hypocentre. This location procedure implements the EDT method (Lomax 2005), improving the hypocentral location through the incorporation of an error rejection method of arrival time outliers. Traveltime uncertainties are accounted for and are estimated to 5 per cent error of the values determined by Nederlandse Aardolie Maatschappij (2017). The complete catalogue of events during September 2015 to January 2017 for the initial coherence method and improved NonLinLoc method are shown in Fig. 3, with the NonLinLoc earthquake catalogues given in Supporting Information Table S1. The station arrival time estimates and uncertainties for all events are given in Supporting Information Table S2. Uncertainty in the multiphase arrivals of the *P* and *S* waves is assessed by determining on the first arrival of each of the phases, consistent with the traveltime formulation. The robustness of this procedure with the formulation of the traveltime model is further demonstrated in the following section.

4.2 Reliability tests of eikonal procedures

In order to test the reliability of the eikonal method with the complex velocity structure in the Groningen region, we applied a simple 2-D waveform modelling calculation using SPECSEM2D (Tromp *et al.* 2008), taking a transect in the 3-D velocity model from Nederlandse Aardolie Maatschappij (2017). Synthetic seismograms were then determined for a hypothetical network of shallow-borehole stations, spaced every 1 km along a transect going through the epicentral location (Fig. 4). The eikonal method was applied for each

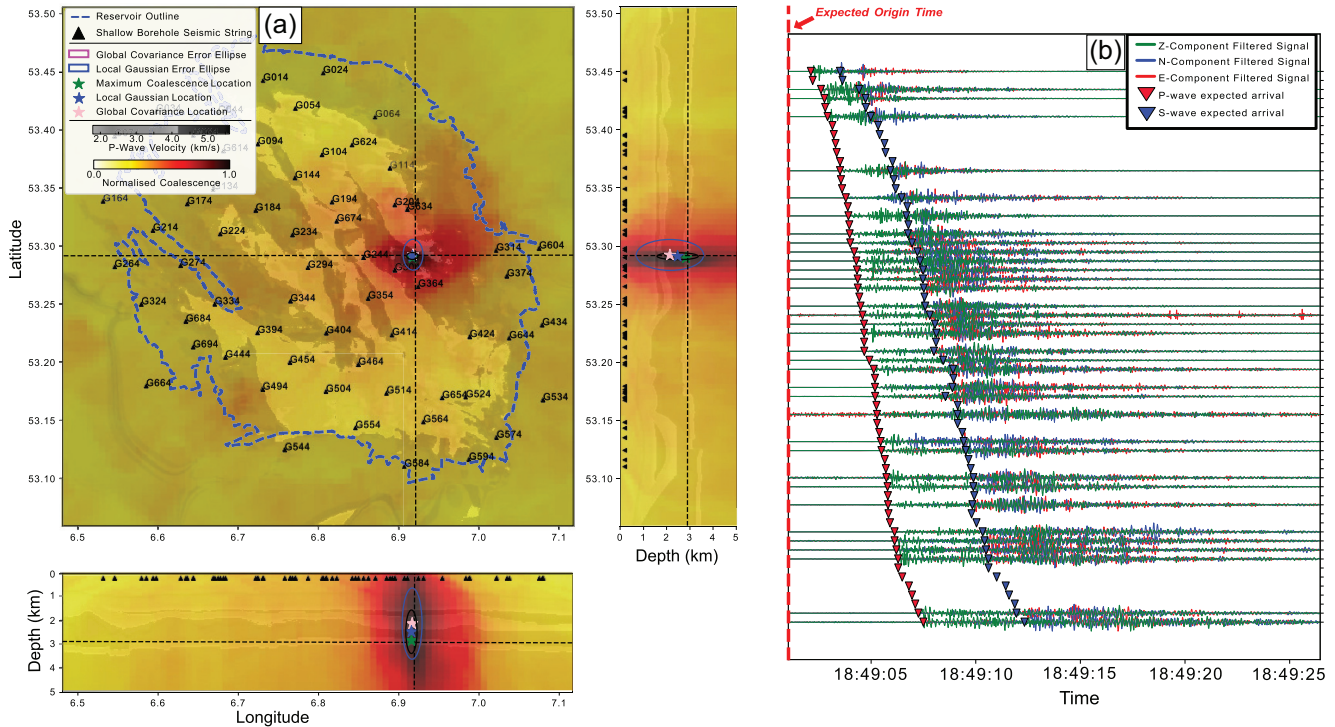


Figure 2. Coherence-based earthquake location for the 2015 October 30 18:49 $M_{LN} = 2.3$ earthquake. (a) Normalized maximum coalescence of the energy of the earthquake, with the maximum value representing the most likely earthquake location. (b) Automated P - and S -wave arrival times for each of the seismic stations from the optimal hypocentral location.

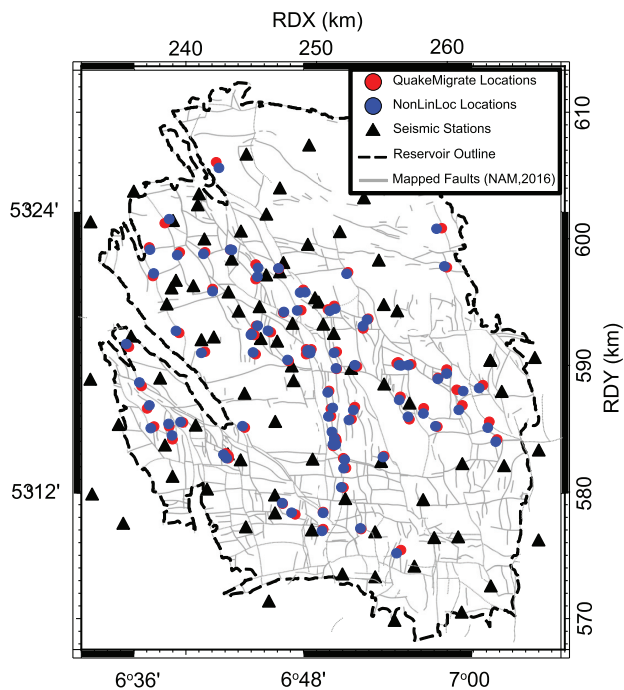


Figure 3. Comparison of the QuakeMigrate and NonLinLoc earthquake locations. Red dots show the initial QuakeMigrate locations. Blue dots show the refined NonLinLoc locations. Black triangles represent the seismic stations. Black dashed line the reservoir outline. Grey lines represented the mapped faults.

of these synthetic stations to determine model P - and S -wave traveltimes. A 5 per cent traveltime uncertainty was incorporated for the event location (shown by dashed lines in Fig. 4), in accordance

with the uncertainty given in Nederlandse Aardolie Maatschappij (2017). There is good agreement between the first arrival picks and the synthetic full-waveform arrival times for the P -wave arrivals, with some disparity in S -wave arrivals at distal stations with epicentral distances > 12 km. We therefore use an eikonal based method, but expect the attributed location error to be affected by S -wave picking at distal stations. Once an initial location is determined, we neglect S -wave arrivals for stations with an epicentral distance > 12 km, and iterate the procedure to refine the earthquake location accordingly.

5 SYNTHETIC EARTHQUAKE TESTING

Synthetic tests are used to evaluate the impact of the changing geometry of the seismic stations, uncertainties on the velocity structure, and arrival times picking on the determination of hypocenters. Using the traveltimes calculated based on the assumed velocity structure, we impose earthquake locations and investigate the impact from various sources of errors. The locations are resolved using the NonLinLoc EDT-octree method (Lomax 2005). The probability distribution is represented with a cloud of scattered points which samples possible hypocentral locations. The uncertainty on the location is smaller where the scatter points make a more compact distribution. We define the error as the difference between the imposed and recovered earthquake locations, with a value determined by fitting a 3-D Gaussian ellipse to encompass 95 per cent of the scattered points (corresponding to two standard deviations). The following paragraphs describe how earthquake locations are affected by changes in the number and distribution of stations, velocity model errors, and P - and PS -location methods.

A regional seismic network was first operational in 1995, later expanded with the addition of three more borehole strings. In 2014,

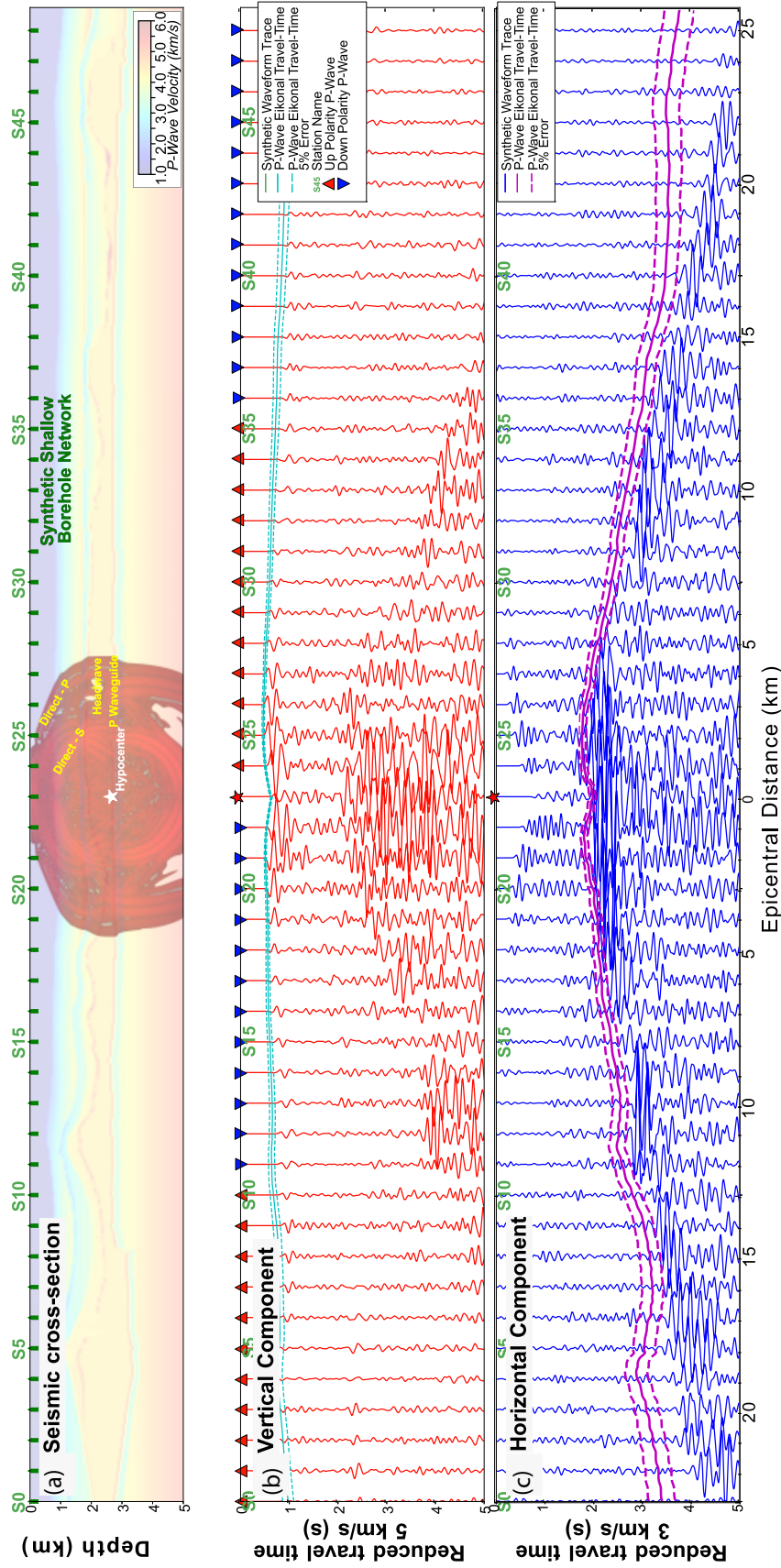


Figure 4. Full-waveform modelling for synthetic testing of the eikonal traveltime method in the complex velocity structure of the Groningen reservoir. (a) Imposed hypocentre for the 2016 November 1 00:57:47 earthquake. Shallow borehole geometry given at 1 km spacing away from the epicentral location. A waveform snapshot is superimposed on a 2-D P -wave velocity slice taken from the Nederlandse Aardolie Maatschappij (2013) 3-D velocity model, projected to a depth of 10 km. (b) Example slice from wavefield simulation. Reduction velocity 5 km s^{-1} for vertical component of synthetic seismogram record section. Predicted waveforms are calculated for the deepest station for each borehole. Overlain line shows expected eikonal P -wave traveltime for each of the stations and 5 per cent bounds for these traveltimes. (c) Example slice from wavefield simulation. Reduction velocity 3 km s^{-1} for horizontal component of synthetic seismograms' record section. Predicted waveforms are calculated for the deepest station for each borehole. Overlain line shows expected S - S - S - S - S eikonal traveltimes for each of the stations and 5 per cent bounds for these traveltimes.

a network of 17 surface accelerometer network was installed across the region for evaluation of damage and hazard assessment. Borehole strings were then added at 70 sites between 2014 and 2016, with an average spacing of 4–5 km. The changing deployment history, and KNMI earthquake locations are shown in Supporting Information Video S1, with more information outlined in Dost *et al.* (2017). In order to assess the effect of changing deployment patterns on epicentral locations, we conduct synthetic tests considering three main periods, representing the densification across the region; 1995–2014, 2014–2015 and 2015–2018. In order to compare our locations with the KNMI event locations, we impose the same 1-D regional velocity model as in Dost *et al.* (2012). We introduce traveltimes uncertainties on both *P*- and *S*-wave picks with the same size as the sampling interval of each instrument. This allows us to assess the minimum uncertainty of epicentre locations. Synthetic earthquake locations were imposed at 5 km spacing in epicentral locations, with a depth of 2.7 km fixed to the average depth of the reservoir. The pdf of the returned hypocentral locations for 1996, 2014 and 2017 array geometries are shown in Fig. 5. The estimated locations of earthquakes with epicentres within the gas field are not biased significantly but there is a significant change in the distribution of epicentral location uncertainties: the initial array geometry of 1995 results in ± 800 m uncertainties; the 2014 geometry in ± 300 m uncertainties; and the 2017 geometry in ± 150 m uncertainties. This synthetic test neglects velocity model uncertainty and traveltimes picking uncertainty, so represents a lower bound on the epicentral uncertainties.

Seismic reflection, seismic refraction, sonic log and well core samples have been combined to produce a 3-D-elastic model of the Groningen region (Nederlandse Aardolie Maatschappij 2017). This 3-D-elastic model has been used to generate a *P*-wave velocity model, with the *S*-wave velocity and density model determined from core-samples and the *P*-wave velocities. This data improves on the prior regional velocity model outlined in Dost *et al.* (2012). The relative uncertainty on traveltimes is estimated at 5 per cent due to the limited resolution of the shallow subsurface velocity structure and the lack of an independent *S*-wave velocity model (Nederlandse Aardolie Maatschappij 2017). In order to assess the impact on earthquake locations, we conducted several synthetic tests by varying randomly the theoretical traveltimes calculated based on our *a priori* velocity model by 5 per cent and 1 per cent, with the phase arrival uncertainty assigned equal to the instrument sample interval (0.015 s). The synthetic earthquakes have epicentral locations spaced every 5 km and a fixed depth of 2.7 km. The pdf of the estimated hypocentral locations for 5 per cent and 1 per cent traveltimes errors for the 2017 array geometry are shown in Fig. 5. The decrease in traveltimes errors from 5 per cent to 1 per cent reduces the uncertainty on estimated depth from ± 300 to ± 100 m at the 2σ confidence interval. Epicentral location uncertainties are significantly reduced from ± 150 to ± 50 m. The uncertainties on hypocentral locations obtained with a 5 per cent random noise on traveltimes is more representative of the uncertainties on the velocity model currently available for Groningen. It could be reduced further in the future by improving the velocity model.

Prior analyses of earthquake locations have been conducted using only *P*-wave traveltimes. In order to assess how the addition of *S*-wave traveltimes affects the earthquake depth resolution, we conduct a synthetic test for earthquakes with hypocentral depth at the reservoir level and determine the vertical depth errors between just *P*- and combined *P*&*S*-wave picks. Synthetic earthquakes have epicentral locations spaced by 5 km and a fixed depth of 2.7 km. The returned hypocentral locations and event scatter for *P* and *PS*

picked phases for the 2017 array geometry with a fixed 5 per cent velocity model error is displayed in Fig. 5. The *S*-wave arrivals significantly reduce the vertical uncertainty in the earthquake location. The uncertainty on hypocentral depth reduces from ± 800 m using only *P*-waves to ± 300 m using both *P*- and *S*-wave arrivals. These synthetic tests demonstrate that picking *S*-wave arrival times is required to constrain earthquake depth with an uncertainty comparable or smaller than the reservoir thickness.

6 RESULTS

We use the coherence method described above to generate an earthquake catalogue. To minimize errors on earthquake locations we selected the period from September 2015 to January 2017. Once the initial catalogue is generated, first arrivals are manually picked for *P*- and *S*-wave arrivals across the entire shallow-borehole array. These manual picks are then used to determine the earthquake locations using the NonLinLoc procedure. This stage incorporates the 5 per cent uncertainty on traveltimes due to uncertainties from the velocity model and manually picked arrival times. The corresponding location uncertainties estimated based on our synthetic tests are displayed in Fig. 5. The synthetic tests demonstrate that the uncertainty is typically ± 300 m on the hypocentral depth and ± 150 m on the horizontal location. This procedure is computationally efficient requiring a maximum of 2 min on a single CPU system. A catalogue of 82 earthquakes was generated for Sept-2015 to Jan-2017, with expected magnitudes between $M_{LN} = 0.0$ and 3.1 (magnitudes determined from Dost *et al.* 2016). The hypocentral locations and location uncertainties are listed in Supporting Information Table S1. Uncertainty in the earthquake locations are defined by the 95 per cent confidence intervals on the hypocentral location.

Our procedure yields well constrained hypocentral locations with vertical uncertainties < 500 m and epicentral uncertainties < 600 m; with the highest probability hypocentral locations situated within the Slochteren, Ten Boer or Lower Zechstein anhydrite levels. The best constrained location in the catalogue (2015/10/30 18:49:02 $M_{LN} = 2.3$; Fig. 6) has a 2-sigma vertical uncertainty of 270 m and a horizontal uncertainty of 210 m with the highest probability hypocentral location within the Zechstein Basal anhydrite, and an epicentral location close to the intersection of the mapped fault planes given by Nederlandse Aardolie Maatschappij (2013). The proximity of the seismicity to the mapped faults suggests that some of them reactivated pre-existing faults. For events with lower magnitudes $M_{LN} < 1.5$, the location uncertainties are larger with epicentral uncertainties ≤ 1 km and hypocentral uncertainties ≤ 1.3 km. The hypocentral locations for these events show some scatter, but those with smaller uncertainties display a most probable hypocentral depth centred near the top of the reservoir.

To assess how the earthquake depths compare to the reservoir depth the posterior pdfs of hypocentral depth can be combined to determine the depth distribution of the earthquakes relative to the top of the reservoir. The depth was determined from seismic reflection, seismic refraction, core sample and sonic log data, with a lateral resolution of 50 m (Nederlandse Aardolie Maatschappij 2017). For each sample drawn from the probability distribution the depth difference between reservoir top surface and sample are determined at the epicentral location. We conducted this procedure for all samples across all earthquakes, and combined all the pdfs to determine a global pdf for the different magnitude threshold of $M_{LN} > 0.0$ and $M_{LN} > 1.5$ (Fig. 7c). Both distributions

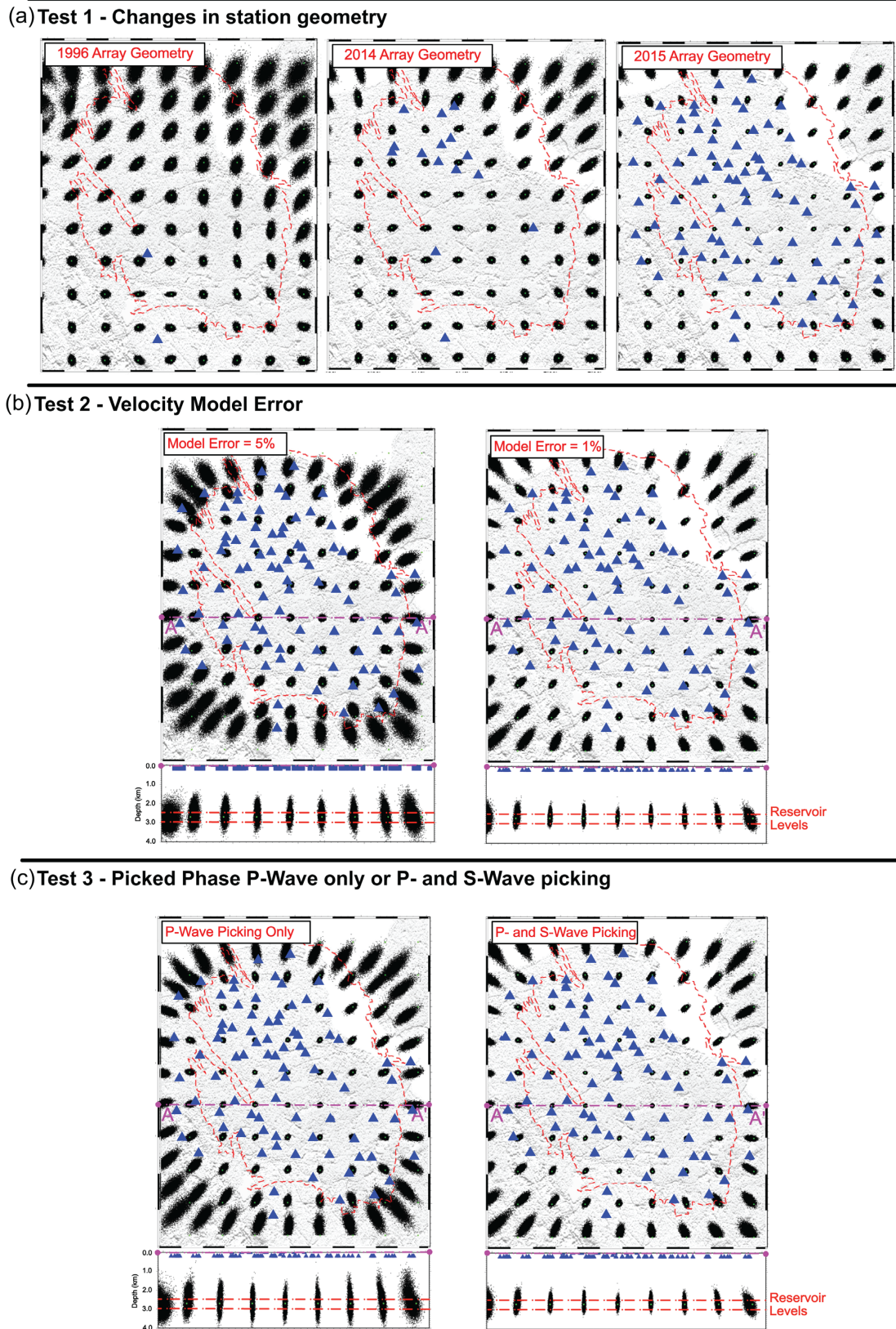


Figure 5. Synthetic earthquake testing for imposed locations at 5 km spacing and fixed depth of 2.7 km. Red dashed line represents reservoir outline. Blue triangles are station locations, green stars the imposed earthquake locations and black points the recovered location error. A to A' represents a cross-section selected across the middle of the reservoir. (a) The recovered location uncertainty with changing array geometry in 1996, 2014 and 2015. (b) The recovered location uncertainty with changing percentage traveltime error of 5 per cent and 1 per cent. (c) The recovered location uncertainty with picked phases of *P*-wave only, or *P*-wave plus *S*-wave.

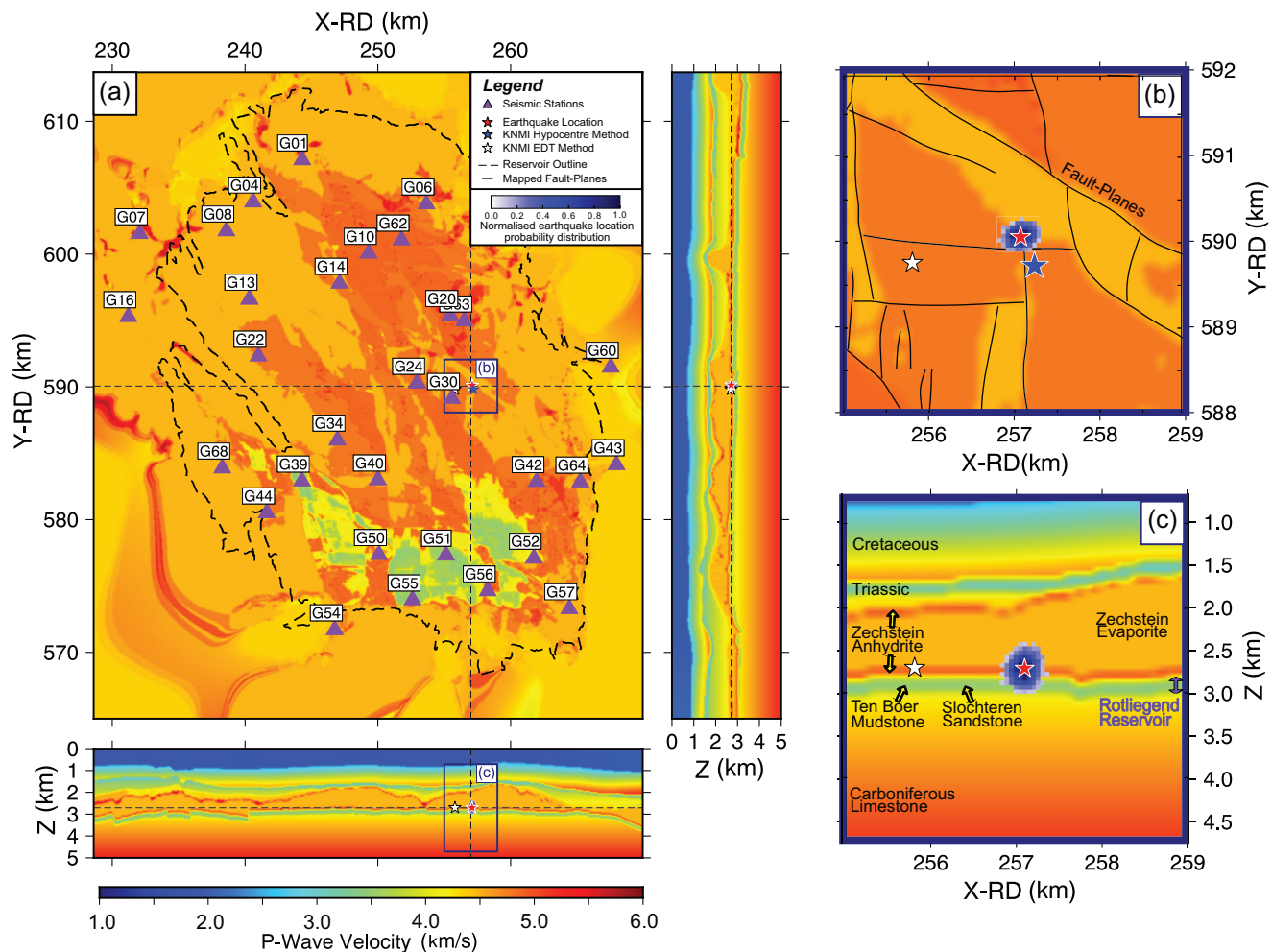


Figure 6. Earthquake hypocentral location and uncertainty for the 2015 October 30 18:49:02 $M_{LN} = 2.3$ earthquake, in comparison with KNMI earthquake locations. (a) Epicentral location of earthquake within the Groningen region. Seismic stations used to determine location are labelled. The sections on the side show NS and EW slices of the P -wave velocity model through the hypocentral location. Black dashed line represents the gas extent within the reservoir. (b) Zoom-in of epicentral location. Black lines show mapped faults at the reservoir level (Nederlandse Aardolie Maatschappij 2013). Red star represents the epicentre of the earthquake from this study, white star the epicentre for the event from the catalogue of Spetzler & Dost (2017) and blue star the epicentre for the event from the KNMI catalogue (Dost *et al.* 2012). Blue shading represents the posterior probability density function (pdf) for hypocentral location from this study. (c) Zoom-in depth slice. Red star shows the location from this study; white star the location from the catalogue of Spetzler & Dost (2017). Blue shading represents the pdf for the hypocentral location from this study.

peak at a depth about 100m above the top of the reservoir with the 2-sigma uncertainty of ± 500 m. For earthquakes with $M_{LN} > 1.5$, the probability of earthquakes originating within the reservoir is ~ 28 per cent. The probability that they originate above the reservoir is ~ 60 per cent and only ~ 12 per cent from beneath the reservoir.

We compared our locations with $M_{LN} > 1.5$ earthquakes with those reported by Dost *et al.* (2012), Spetzler & Dost (2017) and Willacy *et al.* (2019). We constructed the pdf of the difference of epicentral locations (Fig. 8a) and of hypocentral depths (Fig. 8b). Fig. 8(a) shows that our locations agree best with those of Willacy *et al.* (2019). The epicentral locations differ by less than 0.36km at the 95 per cent confidence level. All catalogues report epicentral locations within 1.3km at the 95 per cent confidence level, demonstrating consistency between all earthquake epicentral locations. Fig. 8(b) show that hypocentral depths, plotted again relative to the top of the reservoir, are also consistent within uncertainties. This comparison does not include the catalogue of Dost *et al.* (2012), as this catalogue imposes earthquake locations at a hypocentral

depth of 3km. The pdf distributions consistently show hypocentral depth within or toward the top of the reservoir, with our determination suggesting systematically shallower hypocentral depth by about 200m.

Our catalogue and the catalogue of Willacy *et al.* (2019) both yield a low probability of hypocentral locations in the Carboniferous Limestone underburden. All catalogues seem consistent with earthquake originating within the reservoir or from the anhydrite caprock.

7 DISCUSSION AND CONCLUSIONS

This study describes a probabilistic back-projection workflow for the detection and location of seismicity with an application to the Groningen region. Synthetic earthquake testing has allowed us to investigate changing station geometry and velocity model uncertainties on hypocentral locations. Current velocity model uncertainties still suggest the use of a 5 per cent velocity model error, due to the

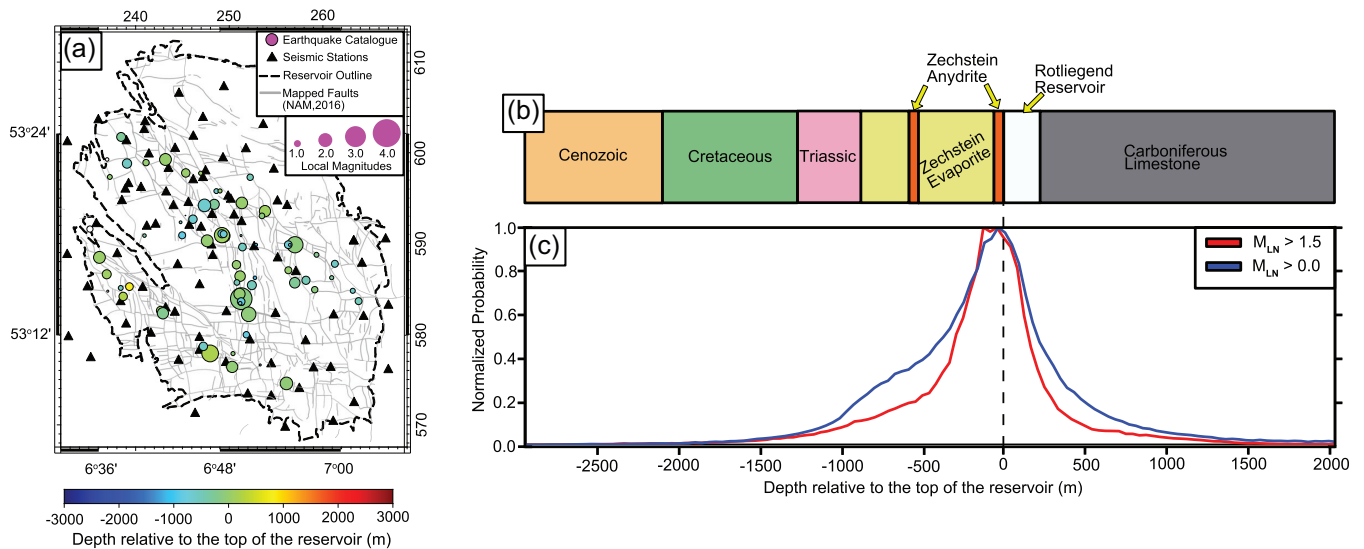


Figure 7. Earthquake location and depth distribution for events between 2015 and 2017. (a) Epicentral locations of earthquakes shown by circles. Mapped faults from Nederlandse Aardolie Maatschappij (2016) shown in grey. Reservoir outline shown as black dashed line. Seismic station locations shown by black triangles. (b) Schematic representation of the stratigraphy relative to the top of the reservoir. Thickness of each unit is the mean value across the region (Nederlandse Aardolie Maatschappij 2013). (c) Combined probability density function (pdf) for earthquake depth relative to the top of the reservoir. Red line represents earthquakes with $M_{LN} \geq 1.5$, and blue line earthquakes with $M_{LN} \geq 0.0$.

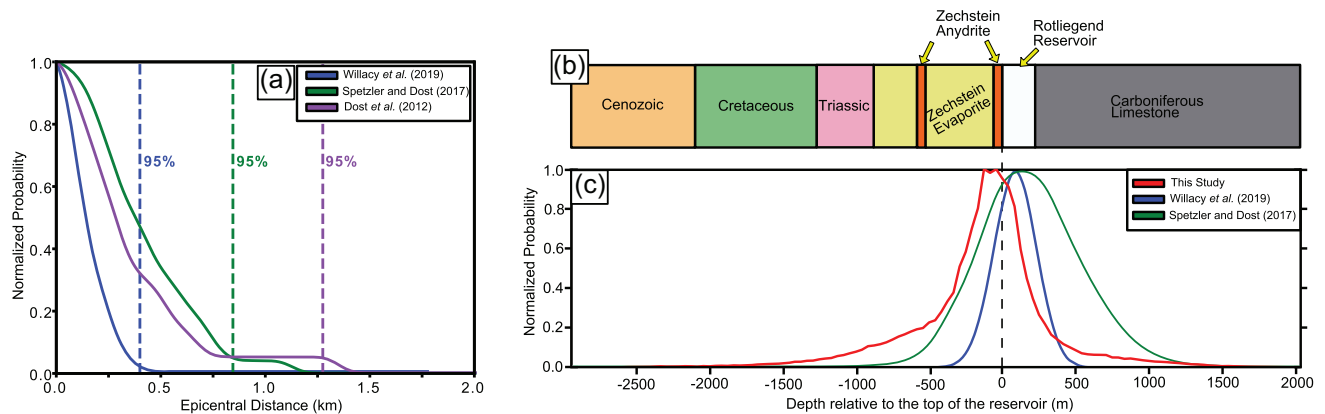


Figure 8. Comparison of epicentral locations and hypocentral depths obtained from this and previous studies of Dost *et al.* (2012), Spetzler & Dost (2017) and Willacy *et al.* (2019) for $M_{LN} \geq 1.5$ between 2015 and 2017. (a) Cumulative distribution of relative earthquake epicentral difference. The 95 per cent percentile for each difference is marked. (b) Probability distribution function of the depth distribution of earthquakes in each catalogue, relative to the top of the reservoir.

poorly defined shallow subsurface velocity structure and no independent S -wave velocity structure. We demonstrated using synthetic testing that if the relative uncertainty on the velocity model can be reduced to 1 per cent hypocentral locations uncertainty could be reduced to $\pm 100\text{m}$, allowing a better understanding of the nucleation lithologies.

Better constrained hypocentral depth is of particular importance to assess the lithologies from which earthquakes are originating, and also to forecast seismicity and assess the possible magnitudes of future earthquakes. The prior studies of Bourne *et al.* (2018), Dempsey & Suckale (2017) and Bourne *et al.* (2015) were assuming earthquakes located within the Rotliegend reservoir. Our refined hypocentral depths suggest that a large fraction (~ 60 per cent) of the earthquakes could nucleate in the overburden, most probably from the Zechstein Basal Anhydrite. By contrast it is unlikely, with a probability less than 12 per cent, that they originate from the underburden.

These results are consistent with the earthquakes occurring in response to the strain induced within the overburden and the reservoir itself, by the reservoir compaction. The compaction of the reservoir implies shear strain of the overburden and reservoir where the reservoir is offset by pre-existing faults (Bourne *et al.* 2018). It is possible that the earthquakes would then be confined to the domain that has been strained as a result of the cumulated compaction since the onset of gas production. In that case, following the logic Shapiro *et al.* (2017), the strain due to the reservoir compaction being bounded, the possible magnitude of earthquakes would be bounded as well. Shapiro *et al.* (2017) suggested a maximum possible magnitude of 4.2, based on the assumption of earthquakes confined to a 400 m thick layer. Our study would call for accounting for the strain of the overburden as well as within the reservoir. A more conservative estimate of the seismogenic thickness would be of the order of $\sim 1\text{km}$, the depth range within which 95 per cent of the hypocenters are located (Fig. 6). With the same other assumptions

as Shapiro *et al.* (2017), such an hypothesis would yield a larger maximum magnitude of 5.0. This reasoning neglects the possibility that future earthquakes would release stresses of tectonic origin in the underburden. It has indeed been proposed that the basal Carboniferous limestone and underlying basement could be critically stressed and that a triggered earthquake could eventually rupture the entire seismogenic zone reaching a magnitude possibly as high 7.5 (Nederlandse Aardolie Maatschappij 2016). The seismological data that we have analysed show no indication of earthquakes nucleating below the reservoir and we therefore consider such a scenario unlikely. We however recommend that this possibility be taken into account in seismic hazard studies.

ACKNOWLEDGEMENTS

J.S. was supported by the Engineering and Physical Sciences Research Council (EPSRC; Grant RG81432; Geomechanical Properties of the Groningen Reservoir) with CASE support from Shell Global Solutions International B.V. The data for this paper can be found in the Supporting Information. Waveform data were provided by Koninklijk Nederlands Meteorologisch Instituut (<http://www.knmi.nl/>). We gratefully acknowledge data and support from Nederlandse Aardolie Maatschappij (Jan Van Elk, Gini Ketelaar and Dirk Doornhof) and Shell Global Solutions (Stijn Bierman, Steve Oates, Rick Wentinck, Xander Campman, Alexander Droujinine and Chris Willacy).

REFERENCES

- Bourne, S.J., Oates, S.J., Van Elk, J. & Doornhof, D., 2014. A seismological model for earthquakes induced by fluid extraction from a subsurface reservoir, *J. geophys. Res.*, **119**, 8991–9015.
- Bourne, S.J., Oates, S.J., Bommer, J.J., Dost, B., van Elk, J. & Doornhof, D., 2015. A Monte Carlo method for probabilistic hazard assessment of induced seismicity due to conventional natural gas production, *Bull. seism. Soc. Am.*, **105**, 1721–1738.
- Bourne, S.J., Oates, S.J. & van Elk, J., 2018. The exponential rise of induced seismicity with increasing stress levels in the Groningen gas field and its implications for controlling seismic risk, *Geophys. J. Int.*, **213**, 1693–1700.
- De Jagar, J. & Visser, C., 2017. Geology of the Groningen field - an overview, *Neth. J. Geosci.*, **96**, s3–s15.
- Dempsey, D. & Suckale, J., 2017. Physics-based forecasting of induced seismicity at Groningen gas field, the Netherlands, *Geophys. Res. Lett.*, **44**, 7773–7782.
- Dost, B., Goutbeek, F., Eck, T.V. & Kraaijpoel, D., 2012. Monitoring induced seismicity in the north of the Netherlands, Scientific status report, <http://www.sciamachy-validation.org/bibliotheek/knmipub WR/WR2012-03.pdf>.
- Dost, B., Edwards, B. & Bommer, J.J., 2016. *Local and Moment Magnitudes in the Groningen Field*, NAM, <https://nam-feitenencijfers.data-app.nl/download/rapport/62253e3b-4843-4e65-aa8b-6d5186cc00df?open=true>.
- Dost, B., Ruigrok, E. & Spetzler, J., 2017. Development of seismicity and probabilistic hazard assessment for the Groningen gas field, *Neth. J. Geosci.*, **96**, s235–s245.
- Drew, J., White, R.S., Tilmann, F. & Tarasewicz, J., 2013. Coalescence microseismic mapping, *J. geophys. Int.*, **195**, 1773–1785.
- Lienert, B.R., Berg, E. & Frazer, N., 1986. HYPOCENTER: An earthquake location method using centered, scaled and adaptively damped least squares, *Bull. seism. Soc. Am.*, **76**, 771–783.
- Lomax, A., 2005. A reanalysis of the hypocentral location and related observations for the great 1906 California earthquake, *Bull. seism. Soc. Am.*, **95**, 861–877.
- Nederlandse Aardolie Maatschappij, 2013. *Technical Addendum to the Winningsplan Groningen 2013*, NAM, <https://www.rijksoverheid.nl/binarie/s/rijksoverheid/documenten/rapporten/2014/01/17/bijlage-1-analyse-voer-verzakkingen-geinduceerde-aardbevingen-en-seismische-risico-s/2-2-a-technical-addendum-to-the-winningsplan-groningen-2013.pdf>.
- Nederlandse Aardolie Maatschappij, 2016. *Technical Addendum to the Winningsplan Groningen 2016*, NAM, https://www.rvo.nl/sites/default/files/2016/06/4_Technical%20addendum.pdf.
- Nederlandse Aardolie Maatschappij, 2017. *Groningen Velocity Model 2017 - Groningen Full Elastic Velocity Model September 2017*, NAM, <https://nam-feitenencijfers.data-app.nl/download/rapport/9a5751d9-2ff5-4b6a-9c25-e37e76976bc1?open=true>.
- Podvin, P. & Lecomte, I., 1991. Finite difference computation of travel times in very contrasted velocity models: a massively parallel approach and its associated tools, *Geophys. J. Int.*, **105**, 271–284.
- Shapiro, S.A., Dinske, C. & Krueger, O., 2017. Maximum magnitudes of earthquakes induced by fluid injections and productions: Controlling parameters, estimations, and case study examples, in *SEG Technical Program Expanded Abstracts*, doi:10.1029/2018JB016801.
- Smith, J.D., Avouac, J.P., White, R.S., Copley, A., Gualandi, A. & Bourne, S.J., 2019. Reconciling the long-term relationship between reservoir pore pressure depletion and compaction in the Groningen region, *J. geophys. Res.*, **124**, 6165–6178.
- Spetzler, J. & Dost, B., 2017. Hypocentre estimation of induced earthquakes in Groningen, *J. geophys. Int.*, **209**, 453–465.
- Tromp, J., Komatitsch, D. & Liu, Q., 2008. Spectral-element and adjoint methods in seismology, *Commun. Comput. Phys.*, **3**, 1–32.
- van Thienen-Visser, K., Pruiksma, J.P. & Breunese, J.N., 2015. Compaction and subsidence of the Groningen gas field in the Netherlands, *Proc. Int. Assoc. Hydrol. Sci.*, **372**, 367–373.
- Willacy, C., van Dedem, E., Minisini, S., Li, J., Blokland, J.W., Das, I. & Droujinine, A., 2019. Full-waveform event location and moment tensor inversion for induced seismicity, *Geophysics*, **84**, 92–99.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

SupplementaryTable1.pdf
SupplementaryTable2.docx
video.zip

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.