WP1: Review of the existing NAM seismogenic source model

Review of seismogenic source models for the Groningen gas reservoir

April 2020

GFZ

Helmholtz-Zentrum
POTSDAM

NORSAR

Contents

1 Executive summary

In the present report, we perform a comprehensive review of NAM's latest seismogenic source model for the Groningen gas reservoir on request of the Dutch State Supervision of Mines (SSM). The initial version of the model (Arup, 2013; Bourne et al., 2014), assumed that compaction is the main mechanism behind induced seismicity in the Groningen field, and used an empirical relation linking compaction strain and total seismic moment. A revision of the model (Bourne and Oates, 2014) changed the focus from modelling total seismic-moment budget to modelling seismic activity rate. The main benefit was a reduced uncertainty in the resulting seismicity forecasts. The model went through additional changes that eventually resulted in the most recent version reviewed in this report (Bourne and Oates, 2017b; Bourne et al., 2018). This version includes both physics-based and statistical elements connecting reservoir deformation due to gas extraction with Coulomb failure stresses of the fault network and probability of earthquake occurrence, including b-values. The probability density distribution of the criticality for failure of pre-existing faults is estimated in an empirical-statistical manner to be able to explain the long-term delay of the onset of seismicity after the start of production. Except for the initial version of the model, aftershock occurrence is also taken into account through the incorporation of the Epidemic Type Aftershock Sequence (ETAS) model (Ogata, 2011).

We rigorously reviewed all relevant elements of the different model versions, though without implementing the actual computations.

Our main conclusions are:

- The influence of the static and dynamic reservoir models on the final results of the seismogenic model is practically not traceable for readers of the publications. Also, the output of the dynamic modelling depends strongly on expert opinion.
- Reservoir compaction is the main input parameter to the models and may be biased by reservoir depth, the uncertainty in the estimate of the reservoir porosity and the uncertainty in the relationship between porosity and uniaxial compressibility of the bulk reservoir. These uncertainties allow for alternative compaction models to be compatible with the available data.
- The fit parameters for the b-value are not well constrained by the small data set that is available. In addition, the use of the maximum magnitude value estimated by the Expert Panel (Coppersmith et al., 2016) implicitly allows for earthquake ruptures that extend outside the reservoir.
- The data fit of the seismogenic models does not explicitly consider earthquake location errors, which may affect the optimised strain smoothing kernel with unknown consequences.
- In the initial model version. the estimation of the strain partitioning function has the strongest impact on the uncertainty and in consequence on the induced strains that will be seismogenic.
- For the activity rate models, the use of the ETAS model for aftershock triggering is reasonable,

but involves six additional fitting parameters, which are difficult to constrain by the few observed aftershocks in the Groningen field.

- For the most recent model version reviewed, the main assumption is that the failure threshold lies within the tail of the stress distribution. Once the reservoir transitions into a steady state, induced seismic activity rates become proportional to induced strains and the models will overestimate the seismicity rate.
- Also for the most recent model version reviewed, faults with the largest offsets are considered aseismic due to their contact with the Zechstein salt formation, which involves the risk of having excluded major faults that may become activated in the future.
- Nevertheless, we consider the approach in the most recent model reviewed to be generally reasonable and sophisticated. The epistemic uncertainties related to the compaction model can be addressed by different branches in the logic tree and aleatoric uncertainties can be quantified by a large number of Monte Carlo simulations.

2 NAM seismogenic source model - an overview

The Groningen gas field is located in the north-east of the Netherlands at a depth of approximately 3000 m. It consists of Slochteren sandstone sealed on the top by the Zechstein salt formation and is primarily limited in its lateral extents by fault closures. The net reservoir thickness increases from about 0 m to 280 m in a south-east to north-west direction, with seismic profiles showing the existence of faults with variability in fault-density and orientations (Bourne et al., 2014).

The presence of gas in Groningen was discovered in 1959 with the start of production taking place in 1963. Since the beginning, interest was placed in the monitoring of subsidence in the field. Subsidence had been linked to oil and gas production in the past, and had also produced significant negative outcomes for surface and near-surface infrastructure, for example, in several oil and gas fields in the US including the well-known Wilmington oil field (Mayuga and Allen, 1969; Yerkes and Castle, 1969).

The first seismic event related to gas extraction in Groningen was detected in 1991 (Van Eck et al., 2006). From this point on, seismic activity continued being observed, which prompted the need for the development of seismic hazard analyses. At first, these analyses focused on the estimation of the maximum-magnitude event that could possibly happen (e.g., see Van Eck et al., 2006). A later analysis incorporated also the probabilistic modelling of peak ground accelerations and velocities (PGA and PGV, respectively) for different earthquake sizes and distances from the source; i.e, using the formal framework of Probabilistic Seismic Hazard Analysis or PSHA (Cornell, 1968). (A detailed description of the evolution of the seismic hazard analysis at KNMI can be found in Goertz-Allmann et al., 2018.) This early analysis, however, did not elaborate on the causal connection between seismicity and production, which limited the possibility to forecast changes in seismic hazard with respect to changes in production scenarios (Van Eck et al., 2006).

The first seismic event related to gas extraction in Groningen was detected in 1991.

Forecast models for natural seismicity rely on statistical stationarity. However, induced seismicity requires a fundamentally different approach, since the operational activities causing or triggering seismicity are not stationary in time.

A seismogenic model is aimed at forecasting the distribution, rate and magnitudes of seismicity. Leading forecast models for natural seismicity rely on statistical stationarity, such that forecasts are based on historic frequencies (e.g., European Committee for Standardization, 2005). However, induced seismicity requires a fundamentally different approach, since the operational activities causing or triggering seismicity are not stationary in time and may even be adapted to reduce the seismic hazard (Bourne et al.,

2018). The NAM seismogenic model for Groningen introduced by Arup (2013) and Bourne et al. (2014) (Figure 2.1) is part of a PSHA where the state of the reservoir is connected through empirical and physical mechanisms to the induced seismicity; therefore, allowing to adapt hazard forecasts to changes in production strategies. The prediction of ensembles of earthquake catalogues in this seismogenic model uses a Monte Carlo simulation, which is perhaps the most general and easily adaptable method in this case (Bourne et al., 2014; Bourne et al., 2015). In this V0 version of the model (see Figure 2.1), an earthquake catalogue is created by exhausting a seismic-moment budget. A later revision (Bourne and Oates, 2015a, ; model V1 in Figure 2.1) found, however, that by modelling the seismicity rate rather than total seismic-moment budget, the uncertainties in the forecasts of seismicity were reduced. Thus, consecutive upgrades of the model continue along this line of development.

The missing element from existing poroelastic theories in order to match the statistical space-time-magnitude history seem to be heterogeneities that localise stress build-up.

Seismic models based only on the statistics of measured seismicity induced in the past may pose a problem when it comes to forecasting the seismic response in a load range not previously covered in the training data set. Therefore, physics-based models are of interest, which link the poroelastic, structural and seismogenic properties of the reservoir and have dynamic fields such as pressure and strain as input. Existing poroelastic theories succeed in describing the occurrence of induced seismicity (Segall, 1989; Segall et al., 1994; Zoback and Zinke, 2002; Segall and Lu, 2015; Shirzaei et al., 2016), but lack sufficient detail to quantitatively match the statistical space-time-magnitude history of observed seismicity or to assess seismic hazard and risk. The main missing element seems to be heterogeneities that localise and amplify stress build-up on pre-existing faults (Bourne and Oates, 2015b, ; model V2 and subsequent versions in Figure 2.1). Further, Bourne et al. (2018) point out that allowing for small-scale heterogeneities in fault strength and treating these as an uniformly distributed stochastic quantity predicts a constant rate of induced seismicity relative to incremental stresses. Bourne et al. (2018) note that some authors found this to be in good agreement with observed rates of induced seismicity for injected fluid volumes up to 10^7 m^3 (Shapiro et al., 2007; Langenbruch and Zoback, 2016). On the other hand, for larger induced deformations, where the bulk reservoir volume change exceeds 2 x 10^8 m^3 as in the case of the Groningen gas field, an exponential-like rise of induced seismicity rate is observed.

Fig. 2.1: *Details of the evolution of the seismogenic model developed by NAM for the Groningen gas field. Model labels were provided by NAM (pers. comm., 2020)*

Year of introduction	Label (NAM)	Relevant details	Aftershocks	Magnitudes	Mmax	NAM report	Journal publication
2013	V0	Strain partitioning model with seismic moment budget	None	Constant b-value	Determined by strain partitioning	NAM (2013)	Bourne et al. (2014) Bourne et al. (2015)
2014	V1	Activity rate with exponential compaction trend and PPP	ETAS	Constant b-value	Constant based on reservoir strain limit (Mmax = 6.5)	NAM (2015a)	Bourne and Oates (2017a)
2015	V2	Activity rate, exponential compaction trend, elastic thin-sheet model*, strain-thickness trend, PPP	ETAS	Hyperbolic tangent stress-dependent b-value	Constant based on reservoir strain limit (Mmax = 6.5)	NAM (2015b)	Bourne and Oates (2017a)
2017	V5	Activity rate with extreme threshold failures and elastic thin-sheet model*, Coulomb stress as PPP	ETAS	Inverse power-law stress-dependent b-value	From 2016 Mmax workshop probability distribution		Bourne and Oates (2017b) Bourne et al. (2018)
2019	V6**	Activity rate with extreme threshold failures and elastic thin-sheet model*, Coulomb stress as PPP	ETAS	Exponential stress-dependent zeta model and hyperbolic tangent stress-dependent beta model	From 2016 Mmax workshop probability distribution	NAM (pers. comm., 2020)	

PPP = Poisson Point Process
ETAS = Epidemic Type Aftershock Sequence
*Elastic thin-sheet model includes the influence of topographic gradients
** Not part of this review
Bourne and Oates (2017a) reviews both strain partitioning and activity rate type models

The new theory for fault reactivation induced by poroelastic deformations provides an operational, stochastic- and physics-based model to account for the early spatial-temporal evolution of induced seismicity rates and magnitudes as a function of reservoir pore pressure, strain and lateral changes in reservoir geometry.

The new theory for fault reactivation induced by poroelastic deformations presented by Bourne and Oates (2017b) and updated in 2018 (Bourne et al., 2018) to account for aftershocks (model V5 in Figure 2.1), provides an operational, stochastic-, and physics-based model to account for the early spatial-temporal evolution of induced seismicity rates and magnitudes as a function of reservoir pore pressure, strain, and lateral geometrical changes within a heterogeneous, thin-sheet reservoir geometry. In this model, elastic and geometric reservoir heterogeneities govern the incremental Coulomb stresses induced by pore pressure changes, and under these incremental Coulomb stress loads the weakest parts of pre-existing geological fault surfaces may experience frictional failure resulting in an induced earthquake. While structural heterogeneities act to localise shear stress development and seismicity within the regions of greatest structural gradients (which are often associated with faults that partially offset the reservoir), elastic heterogeneities act to localise induced shear stresses and seismicity within the regions of greatest reservoir compressibility (which often are associated with higher reservoir porosity; Bourne and Oates, 2017b; Bourne et al., 2018). Fault friction heterogeneities localise seismicity within the weakest fault segments and govern the exponential-like increase in the rate of induced seismicity and an increase in expected magnitudes as the fraction of fault segments that are close to failure increases (Bourne and Oates, 2017b; Bourne et al., 2018).

The current seismogenic model is checked approximately every six months against new earthquake data, but so far, the availability of new data never invalidated the model NAM (pers. comm., 2019).

In model version V5, resolvable elastic and geometric heterogeneities are represented by a deterministic, smoothed, poroelastic, stress-strain tensor field, which is derived using geodetic measurements of surface displacements, geophysical imaging of reservoir geometry and in-well monitoring of reservoir pore pressures (Bourne and Oates, 2017b; Bourne et al., 2018). Since the parameters of frictional fault strength and initial stress heterogeneities are not directly observable, they are represented in the model by a single invariant probability distribution of initial fault stress and a transient stochastic function for stress triggering due to previous earthquakes (Bourne et al., 2018). The resulting stochastic model for earthquake occurrences including aftershocks may be estimated using the observed earthquake catalogue (Bourne et al., 2018). The forecasts of the current seismogenic model are checked approximately every six months against new earthquake data (until replaced by a new model), but so far, the availability of new data never invalidated the model (NAM, pers. comm., 2019).

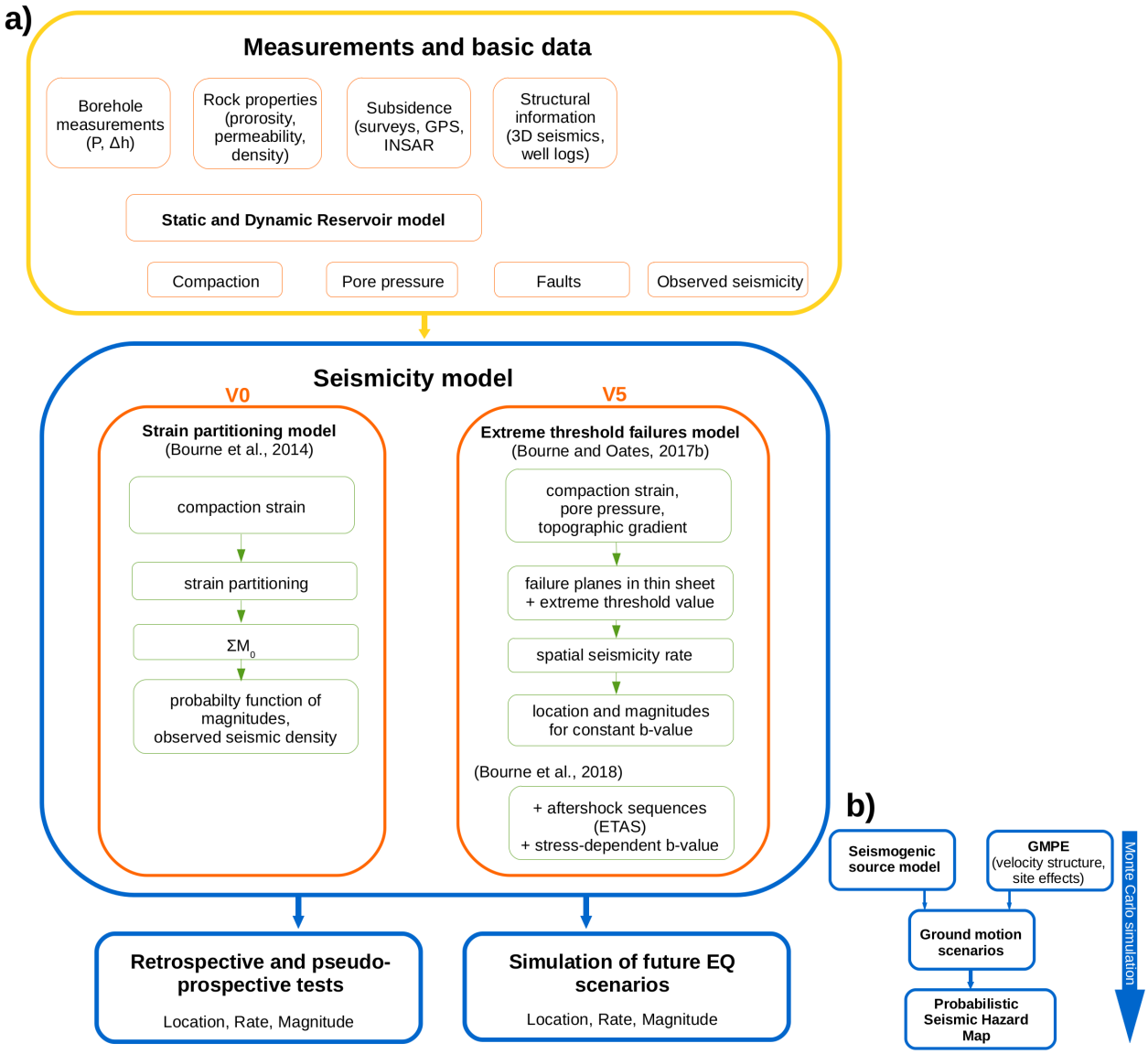

Fig. 2.2: *a) Comparison between the first version (V0) of the NAM seismogenic source model and the latest version reviewed in this report (V5) together with input and output data. b) The implementation of the seismogenic source model in the PSHA with the Monte Carlo simulation workflow.*

Thus, the seismogenic model has evolved from having an empirical basis (Bourne et al., 2014) to a more physical and complex formulation (Bourne and Oates, 2017b; Bourne et al., 2018). Note, nevertheless, that the variations to generate the results are based on almost the same input data (Figure 2.2). Independently of the formulation at hand, it is clear that if the input data is not reliable, the output from the model will not be representative of the process that it intends to reproduce. For this reason, Section 3 is dedicated to the analysis of the input data, paying particular attention to the seismic observations and the reservoir compaction model. The static and dynamic models have already been independently reviewed (SGS Horizon B.V., 2016), therefore, this report mostly sum-

marises relevant conclusions from those reviews. Section 4 reviews the different NAM seismogenic source models as they evolved from versions V0 to V5. It is interesting to note that, although uncertainties from the input data are certainly carried forward into the forecast of earthquake catalogues, Monte Carlo simulations offer the possibility to incorporate these uncertainties in the final estimated hazard (NAM, pers. comm., 2019).

The general structure of the PSHA proposed by NAM is displayed in Figure 2.2b. The focus of this report is the review of the current seismogenic model proposed by the authors including its evolution over time (see Figure 2.1, Section 4), but in addition, we also offer some general comments on the use of Monte Carlo simulations for PSHA (Section 5).

3 Input parameters

3.1 Static model

A static reservoir model is a representation of the subsurface that encompasses geological aspects that are stable over much longer time periods than the duration of the phenomena being investigated with the model. These geological aspects include the geometry of structures such as reservoir thickness, layers and faults, and their lithological properties.

The structural characteristics of the reservoir are normally interpreted from surface seismic and/or any other available geophysical data (e.g., gravity and magnetic surveys), and tied to information observed inside wells. The lithological information is obtained from well logs and cores. This information is then correlated to the attributes of the seismic data at the well locations, such that these correlations can be used as a reference to populate the model away from the wells guided by the seismic data and/or using statistical properties (i.e., geostatistics).

Static reservoir models are the input to dynamic models used to forecast production, depletion and, as in the case of the Groningen field, other processes such as surface subsidence and induced seismicity.

3.1.1 Measurements available for the model

Around the days of its discovery, only **2D seismic** data and information from a handful of wells was available to construct a model of the Groningen field. More well and seismic data was subsequently incorporated, which resulted in an improved model in 1969. However, it was not until after the introduction of **3D seismic** technology in Groningen that a more detailed model of the field became available by 2003 (Visser and Solano Viota, 2017).

In addition to seismic data, information from **well logs** and **cores** continuously became available with the drilling of new wells in the field. By 2012, more than 400 wells had been drilled through the reservoir layers contributing with information that has been incorporated into the static model (Visser and Solano Viota, 2017).

3.1.2 Model evolution

The 2003 static model consisted of a structural model that included tops of formations and location of faults. This model was updated in 2012, and then again in 2015. Both updates were followed by independent reviews in 2013 and 2016, respectively (SGS Horizon B.V., 2016).

The static model of Groningen was independently reviewed in 2013 and 2016, with the recommendations from the reviewers being incorporated afterwards.

The 2012 model included more than 1100 interpreted faults, but this number was reduced to around 700 to improve efficiency in the simulations where the model was being used. The reservoir was divided into 175 layers (Visser and Solano Viota, 2017). The 2015 model is an extended version towards the west of the field with no major changes in the structural components. Some faults were removed and some others were added in the extended area (Figure 3.1). Still, note that the independent review assessed that the density of faults in the extended area was too low from a geological perspective (SGS Horizon B.V., 2016).

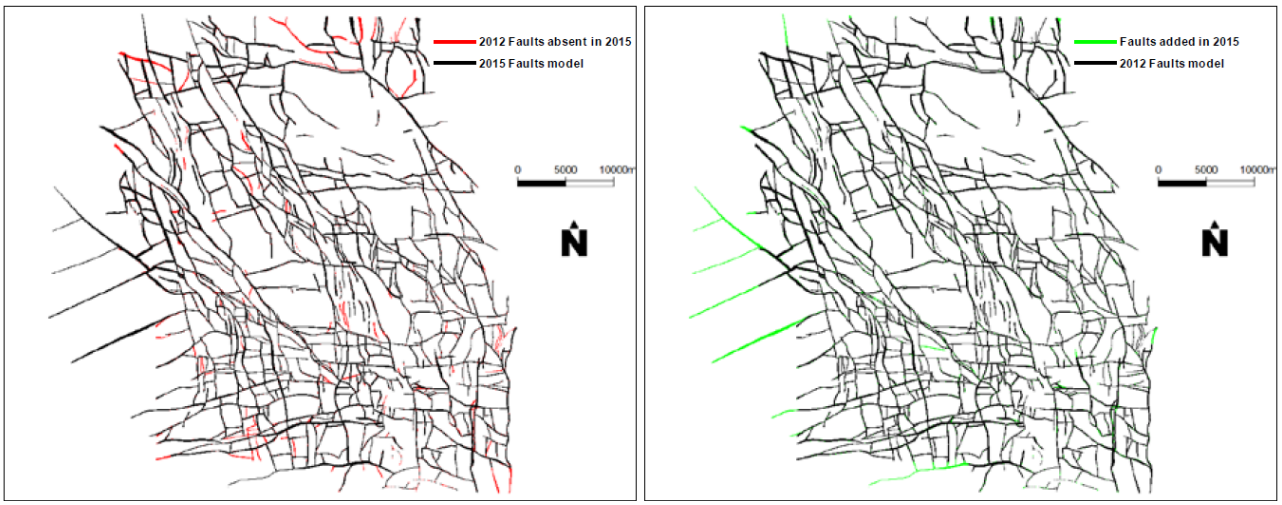

Fig. 3.1: *Map view of faults included in the 2012 and 2015 models of the Groningen field. Image from SGS Horizon B.V. (2016)*

The work from 2003 also included the results of a seismic inversion, which was incorporated into the construction of the 2012 and 2015 models. From the seismic inversion, an acoustic impedance cube was produced, and this was the input to generate porosity maps for the different units in which the reservoir was divided. An additional inversion process was then followed to improve resolution. As an output, maps of porosity, thickness, net-to-gross (i.e., part of the formation with reservoir quality) and an acoustic impedance cube were generated (SGS Horizon B.V., 2016).

Cores were used to estimate the net-to-gross values that were input to calibrate the second inversion. In this respect, the independent review observed that the cores of Slochteren sandstone were not evenly distributed over the field with several of them extracted from wells near the flanks of the reservoir. The central region of the reservoir, which contains most of the gas initially in place, was not representatively captured within the available core data. The information from the cores was correlated with gamma ray logs to determine Vclay (clay volume) cutoff thresholds necessary to specify net-to-gross values outside the location of the cores (SGS Horizon B.V., 2016).

SGS Horizon B.V. (2016) points out a high level of heterogeneity on properties derived from well logs that does not seem to be consistent with geology. It is observed that the source of the problem is related to measurements of density, and differences between service companies in terms of how

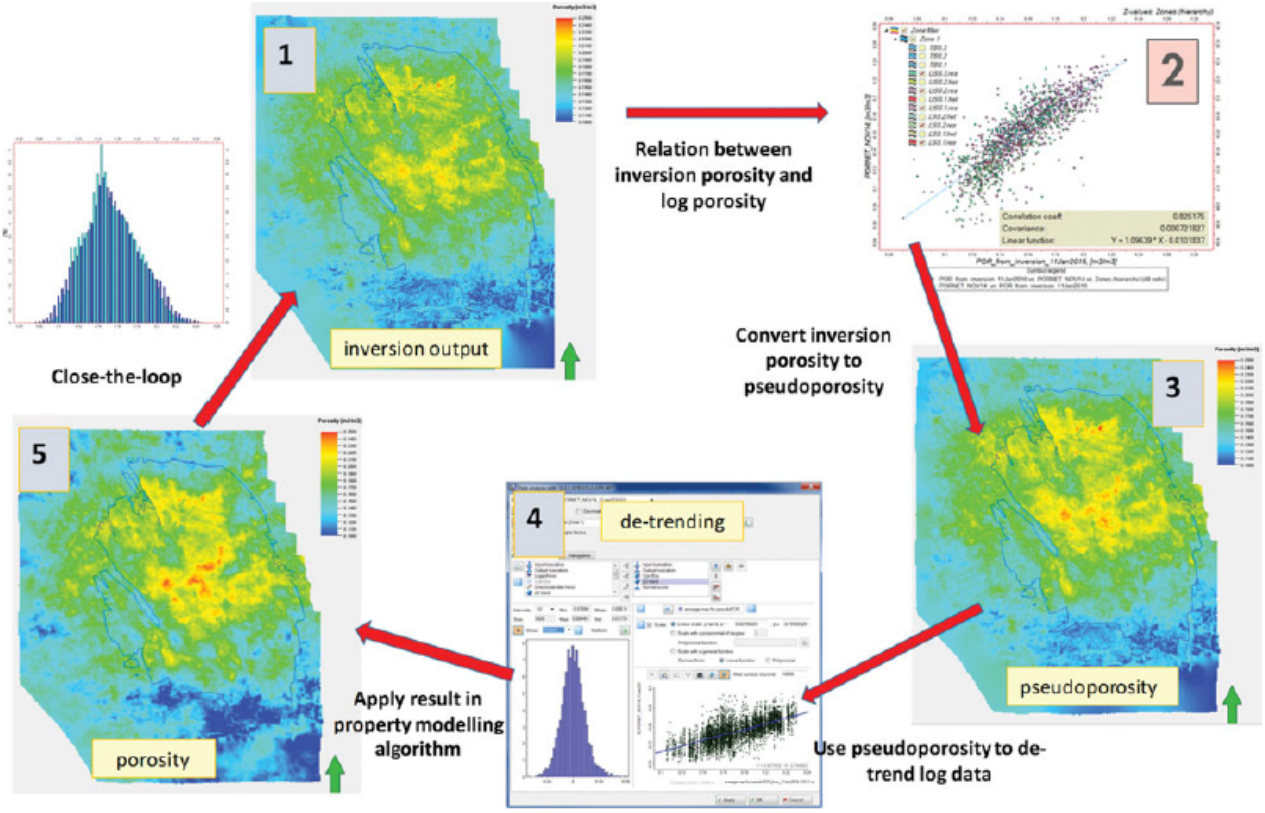

Fig. 3.2: *Workflow followed for the generation of a porosity cube in the 2015 model. Image from Visser and Solano Viota (2017)*

porosity is interpreted in the service company's respective log measurements.

Despite the identified inconsistencies, the independent review from SGS Horizon B.V. (2016) observes that the 2015 static model correctly captures *the main geological, geophysical and petrophysical features of the Groningen Rotliegend reservoir*, offering recommendations on further improvements for the derivation of porosities. Similarly, Visser and Solano Viota (2017) refer to ongoing efforts into incorporating additional details to the model but without expecting major modifications except if new modelling techniques or data become available.

3.1.3 Discussion

The porosity information in the static model influences the calculation of the pressure changes and the compaction values that are input to estimate surface subsidence, which in turn is used to model future seismicity within the Groningen field. In consequence, uncertainties in porosity maps within the static model may have a significant impact in seismicity forecasts. During a meeting with NAM representatives, it was mentioned that the recommendations from the 2016 review had been addressed in the current model (NAM, pers. comm., 2019).

Porosity and permeability change with compaction. It is not obvious from the consulted information

whether or how these changes are addressed within the models. Repeated well logs would be useful to verify and calibrate predictions of porosity and permeability changes from in-situ measurements. No information about repeated well log measurements was found during this review. The absence of this type of measurements was confirmed during a meeting with NAM representatives (NAM, pers. comm., 2019).

Finally, no information was found on the extent of data collection within the formations that confine the reservoir. This information is important not only to model how the behaviour of the confining formations can affect the reservoir response, but also, to determine whether seismicity could migrate outside the reservoir.

3.2 Dynamic model

The dynamic reservoir model is employed to assess the pressure response in the gas bearing formations to the extraction of gas and water, since pressure is an important driver for compaction and subsequently, subsidence. Because compaction-induced strain is considered to be of high importance for the origin of induced earthquakes, the model plays an important role in the optimisation of the gas withdrawal from the reservoir to reduce seismicity (Burkitov et al., 2016). Whereas parameter-fitting of the dynamic reservoir model to pressure data and water influx data is conventional, matching it to subsidence data is a novelty (van Oeveren et al., 2017).

The dynamic reservoir model is employed to assess the pressure response in the gas bearing formations to the extraction of gas and water, since pressure is an important driver for compaction and subsequently, subsidence.

A two-step method is used: first, field-average data is matched and then, the match is improved at regional level subsequently. The output from the dynamic reservoir model is employed as input to forecast compaction, subsidence and production. The modelling of subsidence is required, since direct measurements of compaction are sparse and only available from 5 wells; thus, the compaction dataset is insufficient to constrain the full reservoir model. For history-matching, the final product is a single best-matched model. For forecasting, 1000 model realisations are computed based on the uncertainty ranges. Afterwards, three models are selected (representing the P10/P50/P90 realisations with respect to ultimate recovery), which are then used as base for forecasts of gas recovery, pressure distribution and subsidence (van Oeveren et al., 2017).

3.2.1 Measurements available for the model

Input data to the dynamic reservoir model comprise reservoir pressure data, aquifer ingress data and subsidence data.

For history-matching of the dynamic reservoir model, data measured at 347 wells throughout the Groningen field are employed (van Oeveren et al., 2017). Wells - and thus temporal information to constrain the model - are scant north of the field and outside of the field limits (van Oeveren et al., 2017). Input data to the dynamic reservoir model comprise reservoir pressure data, aquifer ingress data and subsidence data. Field-wide subsidence has been measured by both satellite and levelling surveys starting from 1972.

The static geological subsurface pillar grid model by Visser and Solano Viota (2017) is upscaled from more than 50 million voxels to about 550 000 grid blocks with an approximate extend of 400 x 400 x 6 m each. A number of properties is either taken from the static geological model (net-to-gross, porosity, horizontal permeability, vertical permeability), measured directly in wells (temperature), assessed

in laboratory experiments (pressure-volume-temperature behaviour of the gas phase), or modelled (capillary pressure, relative permeability). Reservoir compaction in each grid cell is calculated based on reservoir pressure as well as matrix compressibility and subsequently converted to surface subsidence in order to constrain the dynamic model (Burkitov et al., 2016). This reservoir compaction, though, differs from the one used in NAM's official subsidence predictions performed by NAM's Geomechanics Department (van der Waal and van Eijs, 2016; van Oeveren et al., 2017). The total subsidence at the surface is obtained by superposition of the subsidence bowls associated with each individual grid cell (van Oeveren et al., 2017).

Due to the large number of over 1600 interpreted faults within the Groningen field, only a limited selection (630 faults; SGS Horizon B.V., 2016) is modelled as sealing in the dynamic reservoir model (van Oeveren et al., 2017). For history match, over 60 years of historic production data are available (including subsidence data, static pressure gradients, repeat formation tests, aquifer ingress from pulsed neutron log surveys) and over 96 model parameters constrained within realistic bounds are used in the history-matching workflow. For each observable, the most sensitive parameters are found to be faults separating the Groningen field from adjacent aquifers (subsidence), the permeability multiplier in the Rotliegendes (static pressure gradient) and the residual gas uncertainty (aquifer ingress; van Oeveren et al., 2017).

For history match, over 60 years of historic production data are available and over 96 model parameters constrained within realistic bounds are used in the history-matching workflow.

3.2.2 Model evolution

Both the initial dynamic reservoir model (GFR2012) and its successor (GFR2015) were reviewed by SGS Horizon in 2013 and 2016, respectively.

The initial dynamic reservoir model (GFR2012) was built in 2011 and 2012 and was used to support the Winningsplan 2013 (Burkitov et al., 2016; van Oeveren et al., 2017). For the model, an in-house 3D numerical, fully implicit, finite volume reservoir modelling simulator by Shell was used (MoReS and Reduce++; Burkitov et al., 2016; van Oeveren et al., 2017). Since the model was well suited for field development, but not for forecasting production-induced seismicity, and indeed had not been constrained to subsidence data, an update was performed as part of the 2015 Groningen Field Review (GFR2015) as input to the Groningen Winningsplan update 2016. In order to match subsidence, a 2D subsidence proxy model was built (Burkitov et al., 2016). Further, the model was extended to capture subsidence effects at the edge of the Groningen field, which required the inclusion of 9 neighbouring fields and their pressure depletion histories (van Oeveren et al., 2017). Both models were reviewed by SGS Horizon in

2013 and 2016, respectively. In addition to including subsidence data, several history match parameters have been updated in the GFR2015 version: fault seal multipliers, gross bulk volume multipliers, permeability multipliers, analytical aquifers, relative permeability parameters (residual gas saturation intercept and slope, end points for gas and water, Corey exponents for gas and water), salinity tuning factor, lowest porosity value for saturation function bins, free water levels, skin factors, aquifer water viscosity, gas and water density and subsidence parameters (Poisson ratio, measurement uncertainty, rock compressibility multiplier; SGS Horizon B.V., 2016). The number of layers in the Lower Slochteren sandstone was reduced to speed up the computations while preserving the relatively homogeneous reservoir characteristics, whereas the number of upscaled layers has been increased in the Upper Slochteren sandstone (Burkitov et al., 2016). More details on the changes from the GFR2012 to the GFR2015 model can be found in Burkitov et al. (2016).

Potential future studies to further improve the predictive modelling are listed by van Oeveren et al. (2017), however, we do not know which of these suggestions have been implemented so far. Among these are to include static reservoir pressure measurements based on shut-in tubing head pressure data, including historic and recent gravity surveys to better constrain reservoir depletion and aquifer ingress, measuring the occurrence of residual gas below the gas-water contact, inverting subsidence data and computed reservoir pressure to update the compressibility grid instead of inferring it from a single porosity transform, and to use the full subsidence data set that is available, instead of only the earliest and latest survey.

3.2.3 Discussion

An assessment of the dynamic reservoir model is not part of this review and the dynamic reservoir model is only described with regard to its importance in the modelling chain culminating in the seismogenic source model. Since the authors of this study are no specialists in the field of reservoir modelling, we mainly follow the argumentation of SGS Horizon B.V. (2016).

The main points criticised by SGS Horizon B.V. (2016) are the difference in gas volumes between the static and the dynamic model (which are within the uncertainty range of 5%, though), the pressure match in the peripheral areas of the field, and the missing sensitivity analyses of the fault seal modelling, the aquifer parameters and the permeability models. In addition, they suggest a more detailed documentation of parameters, since e.g. the water salinity is set to different values in the correlation for water density calculation and the water gas ratio computation. Further, due to upscaling, the reservoir property distribution is more homogeneous in the dynamic than in the static model. Also, due to the dynamic model being coarser, wells of a cluster may be penetrating the same or neighbouring grid cells and the accuracy of the pressure match might be limited (SGS Horizon B.V., 2016).

The dependency of 96 model parameters in the history-matching workflow on the agreement of the multidisciplinary subsurface team (van Oeveren et al., 2017) strikes us as unusual. Geologists, ge-

omechanics and petro-physicists are involved to constrain parameter ranges in realistic bounds and to filter out inconsistent or unrealistic combination of parameters. A space-filling approach is used subsequently to generate a set of 1000 models. For each model realisation, every model parameter is sampled within its allocated range. For each data type (pressure, aquifer ingress, subsidence), a field-wide root-mean-squared error is calculated. Models with low field averaged mismatch to subsidence, pressure and water influx are selected by a graphical method and not an iterative method using a numerical solver due to practicality (van Oeveren et al., 2017). To improve the match at well level, the local root-mean-squared error is computed and compared to all variable model parameters and in case that an apparent correlation between a parameter and a local mismatch is found, the parameter value is changed manually. In addition, in case of mutually exclusive combinations of parameter matches, parameters may be adapted by the subsurface team (van Oeveren et al., 2017).

Thus, the output of the dynamic modelling depends strongly on expert opinion. In addition, SGS Horizon B.V. (2016) points out that within the inherent range of uncertainty and the large set of model parameters, a single history match model is by definition a non-unique solution.

The output of the dynamic modelling depends strongly on expert opinion. SGS Horizon B.V. (2016) points out that within the inherent range of uncertainty and the large set of model parameters, a single history match model is by definition a non-unique solution.

Nevertheless, the dynamic model was deemed by SGS Horizon B.V. (2016) appropriate for preparing production forecasts of future reservoir pressure distribution over time within the Groningen field. Especially, the pressure match on the clusters is very good and does not demonstrate a significant trend of pressure mismatch against time, which is an indicator that the overall field depletion mechanism is honoured by the dynamic reservoir model. SGS Horizon B.V. (2016), though, did not assess the geomechanical aspects of the reservoir rock and potentially induced subsidence and seismicity.

3.3 Reservoir compaction model

The weight of the column of material located above the reservoir (i.e., the overburden) is supported by its rock frame and the fluids filling its pore space. As fluids are extracted from the reservoir, the pore pressure drops evenly or unevenly, and additional weight is placed on the rock frame. Eventually, the compressive strength of the reservoir can be exceeded, resulting in partially irreversible deformation (i.e., compaction) that reduces the pore space and permeability (Settari, 2002). Pore pressure gradients, e.g. caused by uneven weight changes, constitute a danger for localised stress concentrations that may result in seismicity.

Assuming a 1D elastic behaviour, with compaction taking place uniaxially in the vertical direction, the entire thickness change in the reservoir is integrally transferred to the surface as subsidence (Figure 3.3a). In reality, as a result of arching in the overburden and boundary effects on the sides of the reservoir, compaction is smaller at the edges, resulting in subsidence bowls (Figure 3.3b).

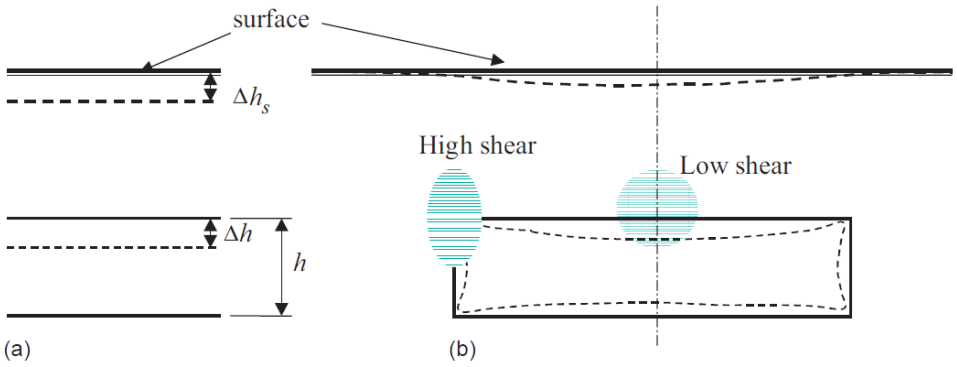

Fig. 3.3: *Compaction and related surface subsidence: (a) 1D elastic model and (b) more realistic behaviour. Image from Settari (2002).*

The state-of-the-art models that explain the connection between compaction and surface subsidence are numerical and extend the domain of modelling to the surroundings of the reservoir. This permits to take into account stress interactions between the reservoir and other geological features, such as faults, and the confining formations (Settari, 2002, and references therein). For the implementation of a model of this type, the data gathering would need to be extended as to permit the characterisation of the formations between the reservoir and the surface, and those limiting the sides of the reservoir. Nevertheless, for first order analyses, semi-analytical models have also been proven to work well for subsidence modelling, they require limited to no information about the overburding, and are more practical to implement (SodM, pers. comm., 2020).

3.3.1 Measurements available for the model

Surface subsidence above the Groningen reservoir has been monitored practically since the start of its production history (Bourne et al., 2014). The earliest measurements of surface subsidence date

back to 1964, and were taken with optical levelling instruments, shallow observation wells and in-situ monitoring wells where radioactive bullets were shot (de Waal, 1986).

Optical levelling measurements were initially available only in the central and southern area of the reservoir. However, from the time of the first repeat survey taken in 1972, the entire field was covered with them. Additional repeat surveys were conveyed in 1975, 1985, 1987 and then every 1 to 5 years until 2018. The last survey, conducted in 2018, had not yet been incorporated in the modelling at the time of preparation of this report (SodM, pers. comm., 2020). The number and spatial coverage of benchmark points varies significantly between surveys (Figure 3.4).

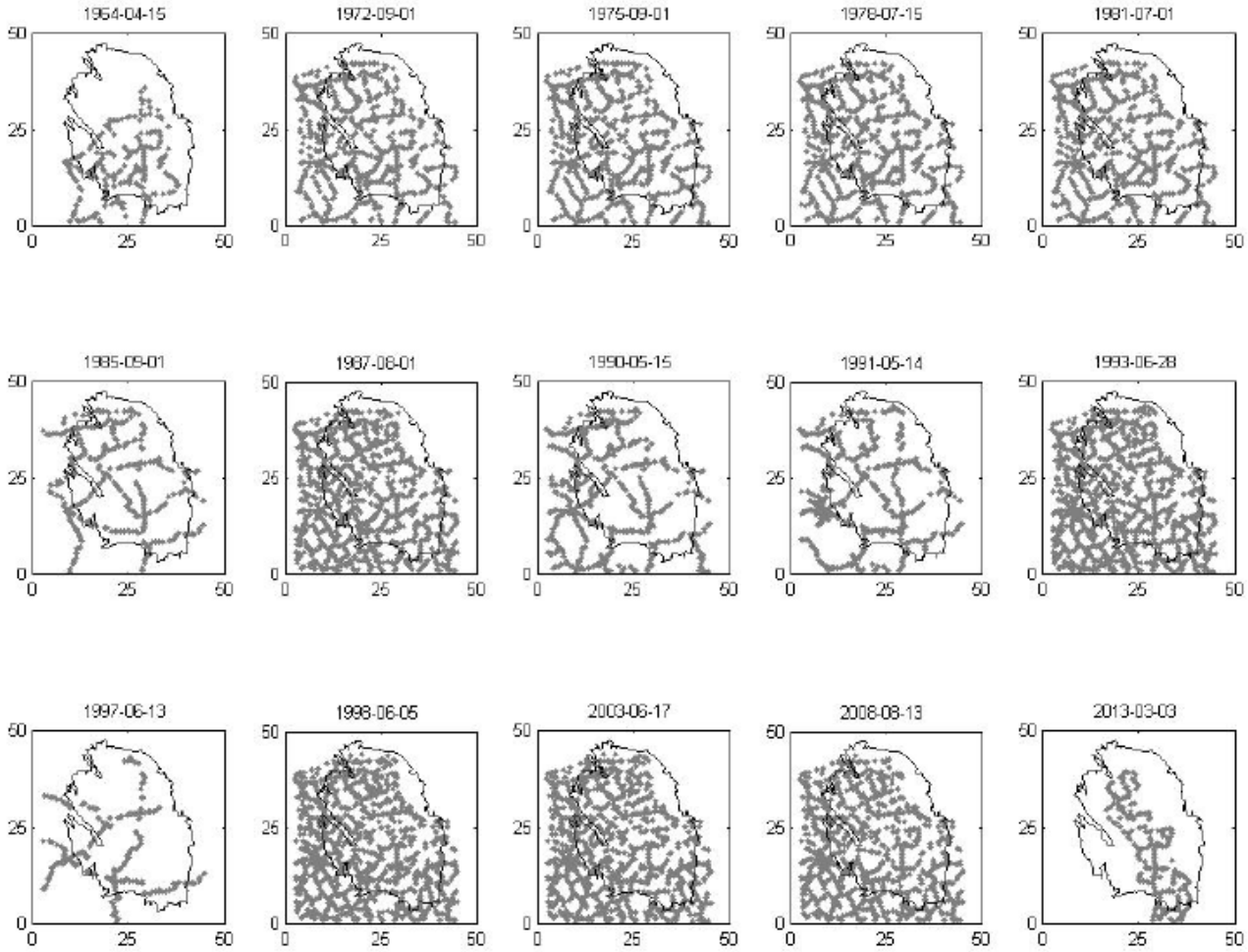

Fig. 3.4: *Coverage of benchmark points of different optical levelling surveys in Groningen. Figure from Bierman et al. (2015)*

InSAR (Interferometric Synthetic Aperture Radar) monitoring of the reservoir started in 1995. The InSAR data was processed to produce the equivalent of an optical levelling survey. Both optical levelling and InSAR measurements were compared against each other in 1997, 1998 and 2008 for corroboration. InSAR measurements are not available for the years 2000 to 2002 (Bourne et al., 2014).

GPS (global positioning system) monitoring started in 2013 with one station. Then, in 2014, 10 more stations were incorporated.

Cores taken from the reservoir have been used in laboratory experiments to estimate compressibility values. The main limitations in these measurements are related to sample contamination, up-scaling of results, and extrapolation away from the location where the cores were taken. However, it is worth noting that a **fibre optic cable** deployed in 2015 recorded in-situ compressibility measurements similar to those obtained in the laboratory (Van Wees et al., 2017).

Radioactive bullets were shot in 11 wells. One well was lost due to mechanical issues, and another one discontinued for logging in 1985 due to sticking problems. During the 1990's, logging in four more wells stopped (Kole, 2015). It was realised that the resolution of the measurements and data quality was not comparable to those from geodetic surveys, and generally insufficient to detect the desired changes within the competent reservoir rock (Mobach and Gussinklo, 1994).

3.3.2 Model evolution

An early investigation effort in Groningen postulated an almost linear relationship between compaction and pore pressure drop, furthermore predicting a maximum subsidence of 100 cm by the end of production (Geertsma and Van Opstal, 1973). However, this (**linear**) **model** did not capture appropriately the compaction behaviour of the reservoir, partly due to the overpressure present (SodM, pers. comm., 2020); hence, by 1974 it was clear that its predictions were an overestimation (de Waal, 1986).

Back then, it was thought that the inaccurate model predictions were related to contamination of rock samples used in the laboratory to obtain measurements of compressibility in the reservoir. Therefore, the compressibilities were manually tuned until the model matched the observations. After this, a new forecast of maximum subsidence of 30 cm was produced (Shoonbeek, 1976). The **tuned linear model** produced forecasts that matched the observations up until 1981. From this point, the accuracy of the predictions degraded, and by 1984 the observations were on the order of 15% underpredicted (de Waal, 1986).

Thus, de Waal (1986) concluded that sample disturbance, together with other factors, were not likely the reason for the inaccurate predictions. Instead, the author proposed that a model in which compaction depends on the type of loading rate (rate type compaction model or **RTCM**) was better suitable for Groningen. This model offered a better match of the observations of subsidence, resulting in an updated forecast of maximum subsidence of between 60 and 70 cm.

Mainly four models have been used to relate pore pressure to compaction in Groningen: linear, bilinear, isotach rate type (RTiCM) and time decay. The last two are non-linear.

Subsequent observations started to depict an increase in the rate of subsidence above Groningen after the initial years of production. This resulted in a temporal behaviour that could be better matched with a **bilinear model** (Figure 3.5). The bilinear model was used by NAM up until 2011 to provide forecasts of subsidence, when measurements above the Ameland field made its inadequacy evident (Van Eijs and Van der Wal, 2017).

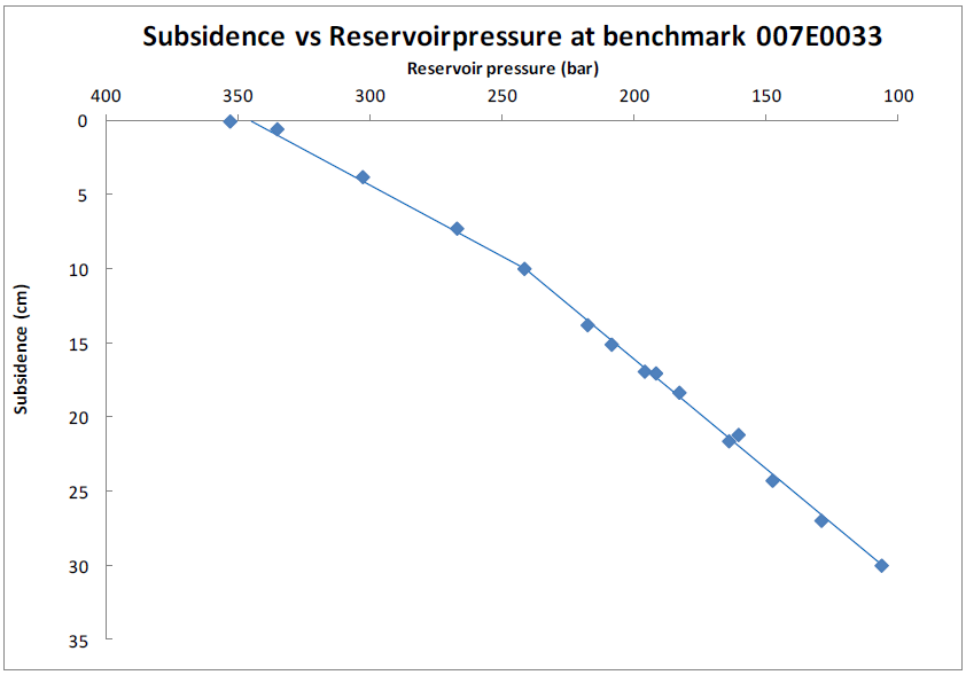

Fig. 3.5: *Example of the bilinear relationship between subsidence and reservoir pressure at one benchmark point in Groningen. Dots are observations. Image from Arup (2013)*

NAM then developed a **time-decay** model that fitted better the observations in the Ameland field (e.g., Mossop, 2012). Later, Pruiksma et al. (2015) also extended the **RTCM** to account for transitions between loading rates. The updated model was referred to as the Rate Type isotach Compaction Model (**RTiCM**). Van Thienen-Visser et al. (2015) observed that the temporal subsidence behaviour measured in Groningen could be fitted by both the time-decay and RTiCM models. The non-uniqueness in the explanatory model was an important finding because each model produced different forecasts of compaction at the end of production (being larger for RTiCM). It is relevant to note the assumptions made in this analysis: a 1D model translating compaction to subsidence, constant depth of the basement layer, and a homogeneous, non-compacting area surrounding the reservoir. Despite acknowledged weaknesses in these assumptions, the authors concluded that the misfits in their results could be better explained by porosity variations in the east of the reservoir, and aquifer activity in the southwest.

A systematic comparison of the predictions produced by the linear, bilinear, RTiCM, and time-decay models in the period between 1972 and 2008 concluded that the models performed similarly in predicting subsidence. All models overpredict subsidence in the northern area of the field and under-

predict it in the southern part (Bierman et al., 2015).

3.3.3 NAM's compaction model

NAM's compaction model is based on the solution of an inverse problem (Muntendam-Bos et al., 2008). It represents a shift from the logic followed by all the previous models in two ways:

1. It uses observations of subsidence to estimate compaction in the reservoir. Conversion from surface subsidence to compaction assumes a linear relationship (Bierman et al., 2015, equation 4.4), and is aided by regularisation constraints.
2. Estimated compaction values are fitted to observations of reservoir pressure by means of coefficients called *rate of compaction per unit pore pressure decline*. The fitting process assumes a linear relationship (Bierman et al., 2015, equation 4.11).

Regarding point 1, all the previous models work in the opposite direction. They first estimate compaction from reservoir pressure. Then, the estimated compaction values are converted into subsidence. In all cases, it is assumed that compaction and surface subsidence have a linear relationship, although this relationship can be different between models.

Regarding point 2, the connection between compaction and reservoir pressure is what sets each of the previous models apart. In some of the models, this relationship is non-linear and/or based on properties determined from laboratory experiments.

The predictions from the previous models are assessed by comparing them to measurements (i.e., surface subsidence), even after going through additional transformations (i.e., semi-analytical model). The compaction values estimated in NAM's approach are not strictly compared to other observations. Instead, they are fitted to the observations of reservoir pressure by solving an additional inverse problem.

NAM developed an additional compaction model based on the solution of two linear inverse problems. One of them also assumes a linear relationship between pore pressure and compaction.

The performance of NAM's compaction model is reported to be comparable to the previous models, although with smaller prediction errors. It is possible, however, that the smaller errors are related to the fact that the inverse problem is set up to fit both sets of observations of subsidence and reservoir pressure (the previous models fit only one set and then compare predictions to the other set). In such case, the prediction power of the model for future years depends on whether the inverted rate-of-compaction-per-unit-pore-pressure-decline coefficients are static properties of the reservoir, and the validity of the relationship on which the matrix of influence coefficients is based (Bierman et al., 2015, equation 4.5).

3.3.4 Discussion

Surface subsidence is an observation used to validate the models of reservoir compaction that are the input to the different versions of the seismological model. While different compaction models reproduce equally well past observations of subsidence, their forecasts are not unique. For example, the compaction forecasts from NAM's compaction model and one of the better accepted physical models (RTiCM) show clear differences in their 5-year predictions (NAM, 2016). This has an immediate impact on seismicity forecasts. Previous analyses concluded that this sensitivity was not of concern because of the Monte Carlo analysis of multiple models that took into account uncertainties in the estimations of compaction (NAM, pers. comm., 2019). On the other hand, this is not part of the Monte Carlo analysis anymore; thus, there is uncertainty on whether this is still the case (SodM, pers. comm., 2020).

Given the connection between subsidence and reservoir compaction, it is important to verify not only the models used to link reservoir pressure to compaction, but also the connection between compaction and surface subsidence. The following points have been identified from the literature review:

- Compaction field estimates are assessed based on their fitting to surface subsidence. Since the translation from compaction to subsidence uses a (semi)-analytical model with its own uncertainties and assumptions, it would be of advantage having access to in-situ measurements for direct comparison. For example, Kole (2015) and Van Thienen-Visser and Fokker (2017) make reference to real-time compaction monitoring with fibre optic. This could help to single out a better compaction model (linear, bilinear, RTiCM, time-decay), and then focus efforts on improving the (semi)-analytical model to match subsidence observations.

  Assessment of compaction estimates would be better done against in-situ references, such that uncertainties in the (semi-)analytical modelling to surface subsidence are not ascribed to the compaction models.

- Given the depth of the reservoir, it is possible that small-scale compaction details become spatially distributed over larger areas as they are propagated to surface subsidence. This constitutes a resolution limit which further supports the case for local measurements of compaction.
- Pijnenburg et al. (2018) reported a variation in the compaction response to pressure changes for Slochteren sandstone that depends on initial porosity, which is spatially variable at field scale. This means that the parameters that define the different compaction models in Groningen also vary in space and that assumptions such as single field-wide compressibility should be avoided (e.g., Bourne et al., 2014). This is taken into account in the more recent model with the use of spatially variable coefficients. These are meant to represent the unknown distribution of porosity and its uncertain relation with compaction (SodM, pers. comm., 2020).

- Pijnenburg et al. (2018) pointed out that poroelasticity is not adequate to represent the behaviour at reservoir level (e.g., observed horizontal stresses are larger than predicted by poroelasticity). The (semi-)analytical model assumes poroelasticity for the propagation of effects from reservoir to surface. It would be necessary to evaluate the impact of inaccurate reservoir modelling in the final subsidence predictions. On the other hand, the inelastic, time-varying behaviour of the reservoir rock is taken into account by some of the reservoir compaction models (e.g., RTiCM and time-decay).
- Hol et al. (2018) conducted laboratory experiments with samples from Groningen. The authors observed that skewness of the particle size distribution and mineral content are variables equally significant as porosity to estimate compressibility in the reservoir. Therefore, the population of compressibility values within the Groningen model could be refined in light of these findings.
- Hol et al. (2018) also recorded seismic activity during their experiments and confirmed its association with inelastic strain mechanisms. Therefore, it would be useful to separate elastic from inelastic compaction and use the inelastic compaction estimates for calculations related to predictions of future seismicity.
- Van Thienen-Visser and Fokker (2017) observed that different influence functions driving the (semi-)analytical model that translates compaction into surface subsidence offer similar outputs. The influence functions tested by Van Thienen-Visser and Fokker (2017) have in common that the overburden is represented by homogeneous layers. Settari (2002) and Van Thienen-Visser and Fokker (2017) observe that finite element modelling is better suited to approximate the behaviour of complex geology (e.g., faults, plastic behaviour of salt layers, and interactions with confining formations). Visser and Solano Viota (2017) and SGS Horizon B.V. (2016) agree in that the current static model is representative of the main structural aspects of the medium. Therefore, the good average fitting of subsidence observations accompanied by regions of systematic misfit obtained with the current (semi-)analytical model suggests that additional complexity may be required in the modelling approach to improve the fitting of all subsidence observations.
- As previously suggested in Van Thienen-Visser et al. (2015) as well as Fokker and Van Thienen-Visser (2016), it is relevant to evaluate the effect of faults in the (semi-)analytical model, not only in the form of porosity variations or discontinuities but also in its impact on local stress reorientation, which can affect the assumption of vertical uniaxial compaction.

3.4 Seismicity

The Monte Carlo simulation of synthetic catalogues employed for seismic hazard assessment requires three inputs: the probability distribution of the total seismic moment, a relative earthquake density map and estimates of the b-value (Bourne et al., 2014).

These can be extracted from observed seismicity by analysing the following parameters:

1. seismicity rate;
2. event locations (or epicentral density, respectively);
3. event magnitudes.

A good estimate of the seismicity rate depends on data homogeneity and completeness as well as efficiency of event detection, which in turn determines the magnitude of completeness. Whereas Bourne et al. (2014) employ a catalogue constricted to the period from 1 Apr 1995 to 30 Oct 2012, the catalogue is updated in later publications to the monitoring period of 1 April 1995 to 1 January 2017, thus containing 271 events with $M_L \geq 1.5$ (Bourne and Oates, 2017b). During this period, the station network underwent two major updates. In addition, the data recording mode changed from triggered to continuous in 2010 and the analysis software was replaced in 2015. Such updates, although necessary to ensure the functionality of a seismic network, may disturb the homogeneity of a catalogue.

NORSAR prepared a detailed review of the publicly available KNMI induced earthquake catalogue for the Groningen field, which is employed in the present model, for the Staatstoezicht op de Mijnen of the Netherlands in 2018 (Dando et al., 2018). This report contains sections reviewing the metadata and development of the station network, analysing existing data (completeness, homogeneity, magnitude of completeness), assessing data processing (including event detection, location and location uncertainty, magnitude and stress drop computation) and finally reviewing selected advanced data analyses (event relocation, cross-correlation of waveforms, source mechanisms). Therefore, we only provide a short overview of aspects relevant to the current report here. If not indicated otherwise, the information is taken from Dando et al. (2018); please refer to the report for further details.

3.4.1 Development of the station network

The very first seismometers in the Netherlands were installed in the seismically active southern part at the beginning of the 20th century (Dost and Haak, 2007). The Groningen field was discovered in 1959 and production started in 1963. During more than 20 years, no seismic activity was recorded or felt in the area. However, in December 1986, the Assen earthquake (M_L=2.8) occurred, followed one year later by another event (M_L=2.5) located only a few kilometres south. Although these events could not be related with certainty to the gas field exploitation, it was decided to instrument the area to monitor its seismicity. This was conveyed by the installation of a small-scale surface seismic

network around Assen in 1988, composed of 6 short-period and vertical-component sensors. This network was operational until 1994 (Dost, 2016). In-between, the occurrence of a third significant earthquake in 1989 (M_L=2.7) increased the need for even more monitoring capacities in the Groningen area (and more generally in Northern Netherlands). Since then, the seismic instrumentation has been greatly improved, keeping in mind two main goals: (1) a better characterisation of the seismicity (detection, location, source mechanism) and (2) seismic hazard mitigation.

The borehole network underwent 3 major upgrades in 1995, 2010 and 2015 and contains 84 stations nowadays.

The Northern part of the Netherlands can be considered to be rather noisy (Dost and Haak, 2007; Dost et al., 2017) and thus, the installation of instruments in boreholes is quite beneficial. The borehole network underwent three major upgrades in 1995, 2010 and 2015 and contains 84 stations nowadays. A first borehole equipped with 5 sensors with a 75 m spacing from 0 to 300 m depth was tested in 1991. In the end of 1995, 7 additional borehole seismometers were installed and formed a regional network covering 40x80 km with an average station spacing of 20 km. During the years 2009-2010, 3 new boreholes were instrumented in order to enlarge the network coverage towards the north. Within the individual boreholes, the sensor spacing was reduced to 30 m.

In August 2012, the largest event (M_L=3.6) that occurred in the Groningen region until today was recorded. The earthquake caused damage to buildings, and a major and massive effort was undertaken to extend the monitoring network and reduce the station spacing to only 4-5 km. Figure 3.6 illustrates the station distribution in the larger Groningen area. From 2014 to 2015, 70 new boreholes were instrumented, constituting the ***G-network*** (red dots in Figure 3.6).

The sensor configuration of the individual boreholes was the same as in 1995 (i.e. 4 sensors separated by 50 m up to 200 m depth), but accelerometers have been additionally installed at the surface. Three further boreholes (***N-network***, blue dots in Figure 3.6) were installed in 2016 slightly south of the Groningen field for the monitoring of a gas storage facility (Dost et al., 2017), and these are also useful for monitoring the Groningen field.

Before 2013, accelerometers have been deployed temporarily in various configurations. Today, 95 accelerometers are installed as part of the B- and G-networks.

Starting from 1997, surface accelerometers were installed in the Groningen region, primarily at places where felt events have been reported. The network configuration was modified several times until 2013 and the instruments were only installed temporarily. A total of 23 accelerometers was installed in 2010. These instruments are not operational any longer, but some of the previous sensor locations were re-used during the years 2013-2014 when the ***B-network*** was put in place (yellow stars in Figure 3.6). 9 accelerometers

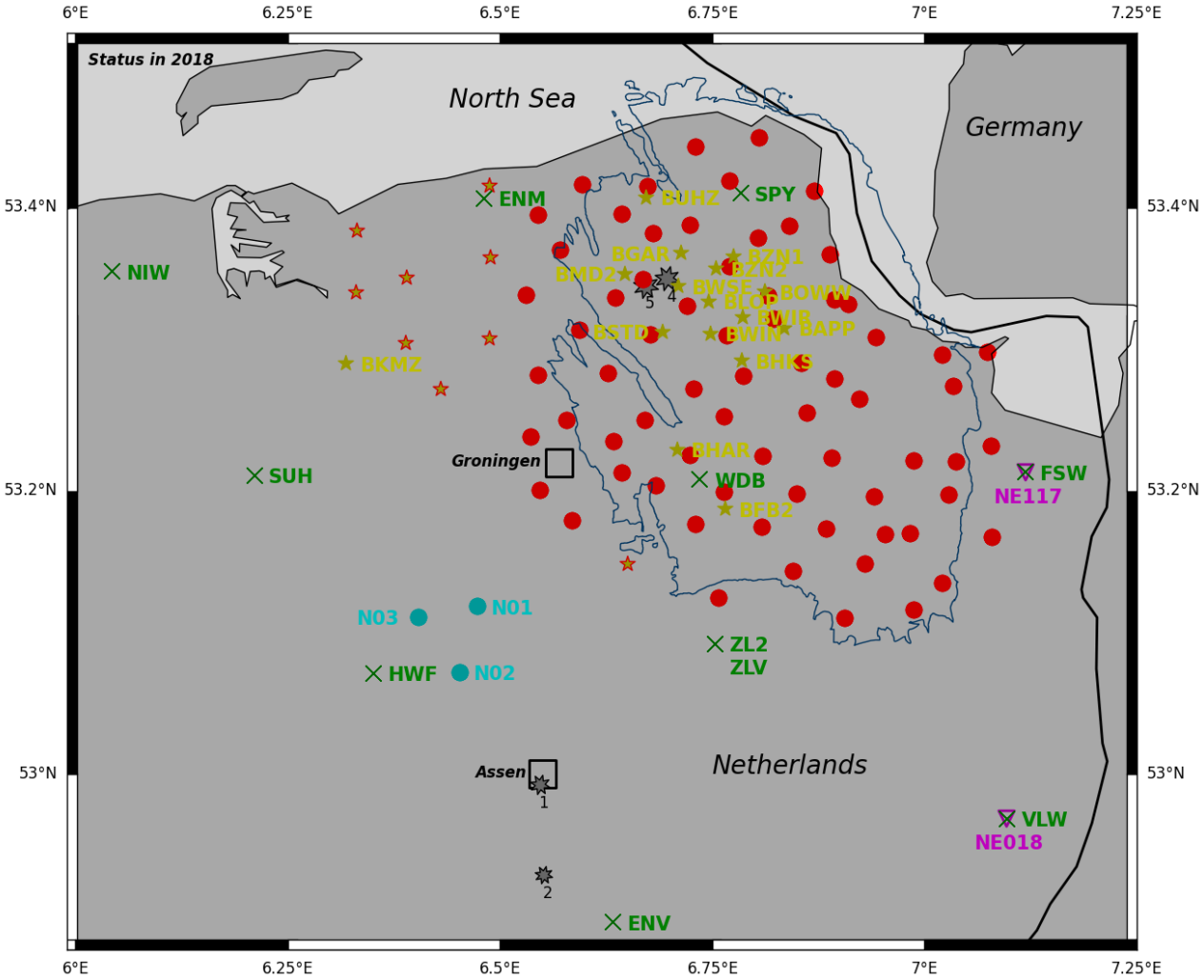

Fig. 3.6: *Seismic instruments operated by KNMI in the Groningen area in 2018. The dark blue line delimits the Groningen field. The dark grey symbols show the location of the major earthquakes (see Fig. /reffig:timeline for details). Symbol code: reversed triangle=broadband; star=accelerometer; cross=borehole string; circle=borehole string and accelerometer at the surface. Figure taken from Dando et al. (2018).*

were also installed as part of the ***G-network*** (yellow stars encompassed in red in Figure 3.6) in December 2017. Today, a total of 25 independent (i.e. not co-located with a borehole) accelerometers are operational in the Groningen region. Combined with the accelerometers placed at the surface of the boreholes of the ***G-network***, there is a total of 95 instruments.

Figure 3.7 shows an approximated time-line of the seismic instrumentation in the Netherlands as well as some of the decisive events which motivated the installation of new sensors over the years.

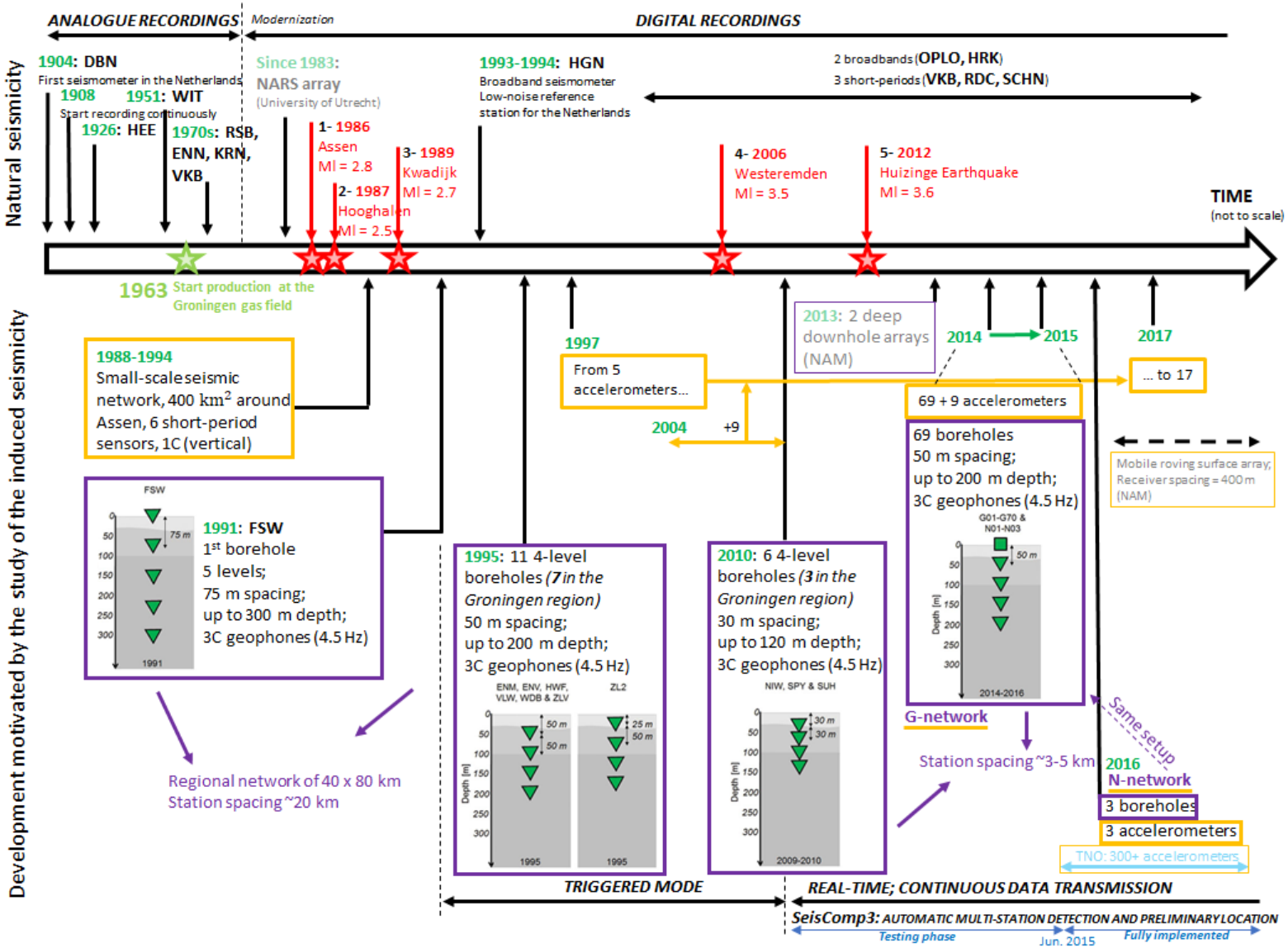

Fig. 3.7: *Time-line of the development of the seismic instrumentation across the Netherlands. A particular focus on the development related to induced seismicity is given in the lower part. Some important dates are also indicated. Figure taken from Dando et al. (2018).*

3.4.2 Data completeness and homogeneity

Typically, the majority of data gaps were fixed promptly within two days, which is very good, given the complexity and size of the network.

Data completeness refers to the availability of data for analysis in the period of interest. KNMI provides easily accessible metadata for qualitative evaluation of the completeness since 2015. The G- and B-networks show good completeness with only few stations having longer periods of data gaps. For multi-level stations, mostly all sensor levels of the given station show gaps; only on a few occasions there were failures of single sensors. This indicates problems with data transmission. Typically, the majority of data gaps were fixed promptly within two days, which is very good, given the complexity and size of the network.

In statistics, homogeneity relates to the assumption that the statistical properties of any part of a dataset are the same as any other part. In the Groningen area a variety of sensors are used, reflecting the technological advances and evolution of the requirements. Thus, sensor types are heterogeneous, but nevertheless, the design of the G-network provides a very homogeneous coverage of the Groningen area. The additional stations of the B-network and the old borehole network provide a denser coverage in some areas. Since KNMI's latest processing requires a minimum of 6 triggered stations, the network provides a homogeneous coverage that is exceptional in comparison with other regional networks. Despite the change of the attenuation function being used for magnitude calculation in 2004 (Dost et al., 2004), the catalogue after 1996 can be considered consistent. Also, event detection is affected by the development of a seismic network, for example, there is a noticeable change in the cumulative number of P-picks in July 2015, due to an increase in availability of data recorded by a significant number of new stations within the G-network. However, this increase is not related to an increase in the number of events, which suggests that the homogeneity of the catalogue has been preserved.

3.4.3 Magnitude computation and magnitude of completeness

For all induced events in the Groningen region, the determination of the local magnitude M_L at KNMI is part of the standard processing and the basis for the Groningen seismic hazard and risk models. Mainly for larger events, also moment magnitudes M_W have been determined from a spectral analysis. Scaling relations between M_L and M_W have been developed for the whole range of Groningen seismicity, which is in agreement with recent theoretical and empirical publications on scaling relations elsewhere.

The magnitude of completeness M_C is defined as the threshold value above which all earthquakes of equal or larger magnitudes are detected, within a given spatial area and during a specific time period (Rydelek and Sacks, 1989). Estimating the magnitude of completeness M_C of an earthquake

catalogue is important for any seismicity analysis and hazard studies. Statistical properties of the seismicity cloud, such as Gutenberg's b-value can only be assessed from earthquakes above the M_C.

Until 2014, KNMI adopted M_C=1.5 for the North of the Netherlands. An overall decrease of M_C from 1.2 to 0.5 was observed between 2003 and 2016.

In the Northern Netherlands, the local population starts to feel earthquakes from about magnitude 1.8 and above. As such, the objective for KNMI was to establish a network that is able to detect all earthquakes that may be felt by people in the Groningen area, and include slightly smaller events as well. Until 2014, KNMI adopted M_C=1.5 for the North of the Netherlands. An overall decrease of M_C from 1.2 to 0.5 was observed between 2003 and 2016 (Paleja and Bierman, 2016; Dost et al., 2017). Note, however, that NORSAR's own analysis led to a significantly higher M_C of 1.5 than reported by NAM (0.9) in the time period 2009-2012 and that NORSAR recommends a moving window analysis for M_C analysis instead of the selection of fixed time windows. Since the authors of the seismogenic model employ only events with $M_L \geq 1.5$, this finding is irrelevant in the current context, though.

3.4.4 Event detection

KNMI have used standard detection methods and made them evolve together with the upgrades in the station network.

KNMI have used standard detection methods and made them evolve together with the upgrades in the station network. Until 2010, the stations operated in triggered mode and event detection was performed on the boreholes only. Continuous data are sent to the KNMI data centre since 2010. SeisComp3 started to be tested at that time and has been fully implemented in June 2015. This transition did not result in a change in the number of events, which suggests that the new implementation has successfully preserved a certain homogeneity in the catalogue. For the Groningen field, the event detection threshold is M_L = 1.0 (Dost et al., 2012).

3.4.5 Event location and relocation

The event location methodology used by KNMI has consistently used an iteratively linearised inversion throughout the catalogue period.

Pick consistency for the bulletin data was judged to be good overall, with only two significant outliers corresponding to valid travel-time residuals. The use of S-waves for locating earthquakes changes with the introduction of the G-network in 2014. From this time, the use of S-waves that are notoriously difficult to pick within the Groningen field is phased out, with the event location relying solely on the P-wave arrival-times.

Throughout the publicly available catalogue, the event depth is fixed to 3 km. The event location methodology used by KNMI has consistently used an iteratively linearised inversion throughout the catalogue period.

Recent work has involved trying to improve both the lateral locations of the Groningen seismicity and resolving the depth of the events. The EDT method, which KNMI have investigated, shows good potential for use with the expanded network, and takes advantage of the 3D velocity model generated by NAM. However, the methodology relies on many station pairs making it incompatible with events occurring before the network expansion. It is noteworthy that although Spetzler and Dost (2017) find the majority of events to be located between 2600 and 3200 m and thus at reservoir depth, they also demonstrated that the full depth range is between 2200 and 3500 m, thus extending outside of the reservoir. Establishing the correct depth of the Groningen induced seismicity has important consequences in terms of understanding the seismic hazard. If earthquakes are occurring in the basement, there is the potential for higher magnitude events (DeDontney et al., 2016; Vlek, 2018), while events above the reservoir have implications for caprock integrity. Hypocentres of relocated events are not part of the publicly available induced seismicity catalogue.

No location uncertainties are provided in the publicly available catalogue or in the bulletin data. An indication of location uncertainty can be gained from rms residual values, providing a goodness of fit for each event. We understand that location uncertainties have been computed, but they are stored internally at KNMI.

3.4.6 Quality of earthquake catalogue

Over the years, huge progress has been made; certainly in instrumentation of the field, which is by now probably the best monitored gas field globally, but not least in data accessibility via KNMI's online data portal, the standardisation of their processing methodology and their continuous effort in research for further improvement (Dando et al., 2018). Overall, taking into account that the KNMI induced seismicity catalogue covers the complete time period of seismicity recorded within the Groningen field, in which both the instrumentation of the field as well as methodologies for data processing changed considerably, NORSAR considers the catalogue as extensive and of high quality. Nevertheless, NORSAR recommends that event depth, having important consequences in terms of understanding the seismic hazard, should be estimated whenever possible. In addition, location uncertainties should be determined, with all relevant parametric data

With more than 100 seismic stations operated by KNMI and an inter-station distance smaller than 5 km, the Groningen area is one of the most thoroughly instrumented sites in the world for the monitoring of induced seismicity. The dense station coverage and the catalogue's high quality provide a sound base to extract parameters required for the seismogenic source model.

stored in the publicly available bulletin (Dando et al., 2018). Notwithstanding, the dense station coverage of the field and the high quality of the catalogue provide a sound base to extract parameters required for the seismogenic source model.

3.4.7 Incorporation of observed seismicity characteristics into seismogenic source model

According to Bourne and Oates (2017b), the largest concentration of earthquake epicentres occurs close to the region of greatest compaction rather than greatest pressure depletion. In fact, 90% of $M_L \geq 1.5$ events happen at a time and place, when reservoir compaction was at least 0.18 m and the location of first such events in 1991 is within the region of greatest compaction (Bourne et al., 2014). However, no single threshold in compaction seems to exist for the occurrence of events, it rather transpires to be a continuous process where the likelihood of an event occurring increases according to the local reservoir compaction (Bourne et al., 2014). The number of events per unit volume change increases substantially with reservoir compaction in a log-linear fashion: doubling the reservoir compaction from 0.15 m to 0.3 m increased the number of $M \geq 1.5$ events per unit volume change by a factor of 10 (Bourne et al., 2014). This exponential-like increase of seismicity relative to the volume of fluids extracted is interpreted by Bourne and Oates (2017b) as progressive failure of the weakest fault patches with heterogeneously and randomly distributed frictional strengths among all reactivated faults.

Bourne et al. (2014) assume the moment magnitude to be similar to the local magnitude over the observed magnitude range, a fact confirmed by Dost et al. (2018) at least for events with magnitudes $M_L \geq 2$ (for lower magnitudes, a quadratic relationship is developed). Both the number of detected events with magnitudes larger than M_C as well as their magnitude estimates influence the computation of b-values. The initial work was based on the catalogue of the period 1 April 1995 to 31 December 2014 (Bourne and Oates, 2017a). Estimates of the b-value appear to depend on reservoir compaction, since there is a statistically significant decrease in b-value with increasing compaction, such that the central area of the reservoir experiences both the highest compaction, the greatest event density and the greatest proportion of larger magnitude events (Bourne et al., 2014). More recent calculations tested spatially varying b-values in more detail; although the catalogue is too sparse for accepted b-value mapping methods (Wiemer, 2001; Wiemer and Wyss, 2002), significantly different b-values were found for regions of 5 km radius centred on Loppersum and Ten Boer (Bourne and Oates, 2017a). On the other hand, no clear indications of systematic dependence of b-value on time or event rate could be recognised (Harris and Bourne, 2015; Bourne and Oates, 2017a).

In addition, the frequency magnitude distribution does not show statistically significant evidence for an upper bound (Bourne et al., 2014). The authors of the seismogenic model first employed a single estimate of M_{max} = 6.4 obtained from expected total bulk reservoir volume change (Bourne et al., 2014) and subsequently, a range of maximum magnitudes by drawing on broader range of site-specific and global analogue data and expert engineering judgement (Coppersmith et al., 2016;

Bourne and Oates, 2017a).

Over the monitoring period 1 Apr 1995 to 30 Oct 2012, less than 0.1% of the induced strain has been accommodated by earthquakes; either aseismic deformation such as fault creep and ductile flow is dominant or induced strain is elastic and remains available to be released by future earthquakes (Bourne et al., 2014).

Bourne et al. (2018) assume the standard epicentral errors to be 500 m, which is similar to the mapped fault spacing, since customized location errors can not be applied in the modelling as it is set up now (NAM, pers. comm., 2019).

4 Seismogenic source model

4.1 Initial model (V0): strain partitioning

The beginnings of the current NAM's seismogenic source model can be traced back to the work of Bourne et al. (2014). The model presented in that work (version V0) is based on the idea of strain partitioning. This approach was used to model total seismic-moment budget.

The average strain partitioning ($\bar{\alpha}$) is a coefficient that represents the fraction of shear strain induced in a deforming medium that is released through seismic sources. It is defined in Bourne et al. (2014, eq. 7) as

$$\bar{\alpha} = \frac{\bar{\epsilon_s}}{\bar{\epsilon}} \,, \tag{4.1}$$

where $\bar{\epsilon_s}$ is the average shear strain related to seismic sources (Kostrov, 1974), and $\bar{\epsilon}$ the average, total induced shear strain. The fraction of shear strain not released through earthquakes is accommodated by aseismic processes such as creep and permanent deformation, or further accumulates until failure takes place. The average is removed from the strain partitioning (i.e., $\bar{\alpha} \rightarrow \alpha$) if it is considered a local/point property of the medium. Based on Kostrov (1974), the total seismic moment available at time t to activate seismic events is expressed through the strain partitioning coefficient as (Bourne et al., 2014, eq. 10):

The initial version V0 of the current NAM's seismogenic model was based on a strain partitioning formulation. This was used to model the total seismic moment budget.

$$M_{0,T}(t) = \frac{4\mu}{3} \int_S \alpha |\Delta h(t)| \mathrm{d}S \,, \tag{4.2}$$

where μ is the rigidity, Δh is a change in thickness of the reservoir (i.e., compaction) and the integral is applied over the surface area of the reservoir S.

The modelling of an earthquake catalogue is then performed as follows:

1. Randomly select a strain partitioning coefficient α from a distribution.
2. Using estimations of compaction within a given time interval and the selected α, calculate a total seismic moment budget.
3. Repeat steps 1 and 2 in order to generate a probability distribution of the total seismic moment budget for the given time interval. This permits to translate the uncertainties in α to uncertainties in total seismic moment.
4. Draw a total seismic moment budget from its probability distribution.
5. Propose a random event location (spatial) and keep it if it agrees with a probability drawn from a map of event density created from the observed seismicity; otherwise, repeat the procedure until a location is selected.

6. Assign a magnitude to the event from a magnitude distribution also fitted to the observed seismicity. Remove the corresponding seismic moment from the total budget.
7. Repeat steps 5 and 6 until the total seismic moment budget is exhausted. Once this happens, a seismicity catalogue will have been created.
8. Generate multiple seismicity catalogues in order to capture the statistical properties of the simulation results.

4.1.1 Mathematical model for strain partitioning

The model proposed by Bourne et al. (2014) for the strain partitioning is constrained by mathematical limits and observational evidence. The mathematical limits are that the variable must fall within the range $[0, 1]$, where the lower limit corresponds to no seismicity, and the upper limit to all strain being released as seismicity. The observational evidence is that induced seismicity in Groningen has been increasing following an exponential trend. The mathematical model chosen by the authors was therefore

$$\alpha = \frac{e^{f+g\Delta h}}{1 + e^{f+g\Delta h}}\,, \tag{4.3}$$

where f and g are hyperparameters, which are fitted to observations of average thickness change of the reservoir layer $< \Delta h >$ and approximations of an average α (i.e., $< \alpha >$). The averages are considered within sections of the reservoir and within the time frame of the available observations of seismicity. In particular, the averaging in time presents the limitation of not capturing the temporal evolution of α, limiting thereby its forecasting power by possibly underestimating its value (e.g., if α is increasing over time as suggested by Bourne et al., 2014, Figure 15).

4.1.2 Discussion

Pijnenburg et al. (2018) and Hol et al. (2018) reported results from laboratory experiments on depletion within Slochteren sandstone where the partitioning between the induced elastic and inelastic strain is a function of the depletion rate and porosity of the rock sample (more specifically, initial porosity in the case of Pijnenburg et al., 2018). In these cases, the authors referred to axial strain rather than shear strain. Thus, the numbers observed by Pijnenburg et al. (2018) and Hol et al. (2018) cannot be translated directly into α values, but nevertheless their observations are valuable to point out bounds to α and physical properties that influence its magnitude.

Pijnenburg et al. (2018) argues that the elastic strain constitutes the budget from which the strain related to seismic events is withdrawn. Based on this argument, the authors concluded that at least 30 to 55% (and possibly more for porosities > 21.6%) of the total axial strain from pressure depletion is not available for seismicity. Furthermore, for strain rates relevant to field scale, the authors consider that these percentages could even be larger.

Hol et al. (2018) on the other hand recorded the acoustic emission (AE) activity generated during depletion experiments employing Groningen samples. From their data, the authors interpreted that AE activity *is related to the deformation mechanism responsible for the (observed) inelastic strain*. These apparently contradictory conclusions can be explained in terms of past and future seismicity. Some of the inelastic strain can be related to seismicity that has already taken place resulting in permanent deformation (past seismicity). On the other hand, as deformation continues, some of the elastic strain will eventually be released as earthquakes (future seismicity), thus becoming part of the inelastic fraction of the total strain. If the laboratory experiments are considered simulations of full depletion cycles of the field, then the total inelastic strain observed during those simulations represents the total budget for seismicity that took place during depletion. In the case of Hol et al. (2018), the authors estimated that up to 75% of the total strain they observed was inelastic. Unfortunately, there is no information reported regarding the amount of this inelastic strain released through AE activity. This could have suggested the scale of values for α. From observations of subsidence and seismicity at field scale, Bourne et al. (2014) estimated a field wide α at scales of 10^{-4}. This implies that even considering that 75% of the total strain is inelastic strain, the portion of total inelastic strain released as co-seismic shear strain in the form of induced earthquakes is very small.

Bourne et al. (2014, eq. 31) proposed a model for α that depends on the change in thickness of the reservoir (i.e., compaction). This general formulation comprehends both permanent (inelastic) and reversible (elastic) deformation. Hol et al. (2018) observed that elastic strain is poroelastic, and for their Groningen samples, porosity is well correlated with elastic compressibility. However, Hol et al. (2018) made the point that elastic strain is not related to seismic activity. On the other hand, the inelastic compressibility, which is related to the observed seismic activity, depends not only on porosity but also on the skewness of the particle size distribution and mineral content. If these results hold at field-scale, the model for α would need to be updated to take into account the variability of these rock properties together with the current models of compaction that consider only the variability of porosity.

Improvements to the model for strain partitioning include dependency on initial porosities (i.e., at the start of production), inelastic compressibility and spatial variability of specific rock properties.

Additional improvements to the model for α are to consider compaction values that reflect dependency on the porosity of the field at the start of production (Pijnenburg et al., 2018), and to consider only inelastic strain.

4.2 Seismicity rate models

4.2.1 A Poisson Point Process connected to compaction

Forecasting models related to seismicity rate have tighter confidence intervals compared to models related to total seismic moment budget.

The strain partitioning is an empirical parameter designed to fit the exponential temporal increase of seismicity observed in Groningen. In an effort to develop an alternative model that offered a better representation of the relationship between seismicity and reservoir compaction, Bourne and Oates (2014) and Bourne and Oates (2015a) proposed a formulation based on the modelling of the rate of seismicity in the context of a Poisson Point Process or PPP (model version V1). This shift in focus had also the benefit of a reduction in the uncertainty of forecasts, which was important for the assessment of hazards related to induced seismicity in Groningen.

In the standard PPP, the probability that n events occur within a time interval $[0, t_0]$ is

$$\Pr(N(0,t_0)=n)=\frac{\Lambda(t_0)^n}{n!}e^{-\Lambda(t_0)}\,, \tag{4.4}$$

where $\Lambda(t_0)=\int_0^{t_0}\lambda(u)\mathrm{d}u$. Bourne and Oates (2014) used the PPP formulation in 4.4 to represent a stochastic process in which seismic events were randomly allocated within a spatial region and period of time based on an intensity function $\lambda(u)$. In order to connect this stochastic process with reservoir compaction, the authors defined the intensity function as

$$\lambda=\alpha\dot{c}\,, \tag{4.5}$$

The seismicity observed in Groningen between 1995 and 2013 for events with $M \geq 1.5$ was better explained by a PPP with seismicity rate connected to compaction compared to a PPP that assumed a field-wide constant seismicity rate.

where α in this case is the number of events per unit reservoir volume change (i.e., $\frac{N}{\Delta V}$) and $\dot{c}$ is the rate of reservoir compaction (i.e., $\frac{\Delta V}{\Delta t}$). In other words, the intensity function specifies the seismicity rate (i.e., $\frac{N}{\Delta t}$) within a section of the reservoir. Through a relative likelihood estimate between assuming a constant intensity function (i.e., field-wide constant seismicity rate) and the intensity function connected to compaction (Equation 4.5), Bourne and Oates (2014) showed that the seismicity observed in Groningen between 1995 and 2013 for events with $M \geq 1.5$ was better explained by the PPP with seismicity rate connected to compaction. Furthermore, the authors observed that the increase in the number of seismic events with the amount of compaction was not linear in this set of observations; hence, they introduced a formulation for α based on an exponential relationship with compaction. Admittedly,

the limited availability of compaction data meant that other functions could also fit the relationship between number of events and unit volume decrease. Thus, the exponential relationship was a personal choice.

The intensity function modified to incorporate the exponential relationship between number of events and compaction is

$$\lambda = \beta_0 \dot{c}(1 + \beta_1 c)e^{\beta_1 c}\,, \tag{4.6}$$

where β_0 describes the background seismic activity and β_1 describes the intensity of the exponential increase of seismicity with compaction. As before, the limited observational data constituted a challenge to constrain unique estimations of β_0 and β_1. The variables were correlated within the range of values that could be constrained with an earthquake catalogue covering the years from 1995 to 2015, again with $M \geq 1.5$.

For the simulation of an earthquake catalogue, the following steps were followed:

1. Draw a total number of seismic events taking place within a time interval (t_s, t_0) from the Poisson distribution given by the PPP with seismicity rate connected to compaction (equation 4.4 with intensity function 4.6).

2. Extract event locations as samples extracted from an expected event density function given by $N(\mathbf{x}) = \beta_0(c_0 e^{\beta_1 c_0} - c_s e^{\beta_1 c_s})$, where c_0 and c_s are the compaction values at position $\mathbf{x}$ at times t_0 and t_s, respectively.

3. Assign an origin time to a given event as a sample drawn from the cumulative probability distribution
$$F(t) = \frac{c_t e^{\beta_1 c_t} - c_s e^{\beta_1 c_s}}{c_0 e^{\beta_1 c_0} - c_s e^{\beta_1 c_s}}\,.$$

4. Finally, draw seismic magnitudes from a truncated exponential frequency-magnitude distribution given by
$$F(M|M \geq M_{min}) = \frac{e^{-\beta(M - M_{min})} - e^{-\beta(M_{max} - M_{min})}}{1 - e^{-\beta(M_{max} - M_{min})}}\,,$$
where $M_{min} = 1.5$ was taken from the earthquake catalogue used to calibrate the model and $M_{max} = 6.5$ was estimated assuming that the total compaction strain at the end of production was released as a single earthquake. The parameter $\beta = \frac{2}{3} b \log 10$ with $b = 1$ (i.e., b-value constant in space and time) estimated from the observed earthquake catalogue as well.

4.2.2 Incorporation of aftershocks

From the comparison between earthquake observations and catalogues simulated with the PPP workflow described in the previous section for the years 1995 to 2014, it was noticed that the occurrence of consecutive events was more clustered in space and time in the observations with respect to the simulation results. This clustering was interpreted as underlying a connection between

consecutive events consistent with the behaviour of aftershocks (Bourne and Oates, 2014). This was not acknowledged in the formulated PPP, since it inherently assumed independent and randomly distributed events. The incorporation of the aftershock behaviour within the PPP model was then accomplished with the introduction of the Epidemic Type Aftershock Sequence (ETAS) model (Ogata, 2011).

The ETAS model includes two parameters for the temporal Omori-decay (c^E and p), two parameters for the aftershock productivity as a function of the mainshock magnitude (K and a), and two parameters for the spatial aftershock decay (d and q). The intensity function incorporating ETAS is represented as

Compared to the PPP simulation results, consecutive events were more clustered in space and time in the observations, which was interpreted as consistent with the behaviour of aftershocks.

$$\lambda = \lambda_p + \sum_{j=1}^{i-1} f(t_i - t_j, \mathbf{x}_i - \mathbf{x}_j | M_j)\,, \tag{4.7}$$

where λ_p is the intensity function for the independent events (i.e., equation 4.6) and M_j is the magnitude of the jth event with location $\mathbf{x}_j$ and origin time t_j. The same notation convention applies for the ith event. The function f is the aftershock triggering function given by

$$f = Kg(t)h(r)e^{a(M-M_{min})}\,, \tag{4.8}$$

with

$$g(t) = \frac{(p-1)}{c^E}\left(\frac{t}{c^E}+1\right)^{-p}$$

and

$$h(r) = \frac{(q-1)}{\pi d}\left(\frac{r^2}{d}+1\right)^{-q}\,,$$

where $r^2 = ||\mathbf{x}||_2^2$. The full set of parameters for the updated PPP model $\{\beta_0, \beta_1, K, p, c^E, q, d, a\}$ was estimated numerically from the seismicity and compaction observations in Groningen. The amount of observational data was not enough to constrain unique values for these parameters, which resulted in multiple covariances between them. Furthermore, c^E was manually specified with a value of 3 days, which was suggested by a graphical analysis of the clustering in the observed seismicity.

Although the model parameters of the PPP incorporating ETAS were not uniquely constrained by the available observations, the clustering behaviour identified in the observed seismicity was reproduced better.

Choosing a set of valid parameters to model the historical seismicity in Groningen, Bourne and Oates (2014) reported an improved reproduction of the clustering behaviour identified in the observed seismicity. Discrepancies with respect to the observations remained in terms of the spatial density of events and distribution of epicentral locations. The first type of discrepancy

mirrored the distribution of residuals observed in the estimation of subsidence from the compaction models. The second type of discrepancy (3 ± 2 km) was partly unresolved considering the uncertainty in the estimated epicentral locations in Groningen (0.5-1 km, Bourne et al., 2018). Therefore, these discrepancies could not be completely reconciled due to the limit of accuracy of the observations used to calibrate the model.

Subsequent versions of the seismogenic model continued incorporating ETAS. In Bourne et al. (2018), for example, the authors reported that during retrospective experiments the activity rate model (version V5 in this report) incorporating aftershocks produced 5-year forecasts in agreement with the observations. The observed seismic activity was found to fall within the 95% confidence intervals of the model (Figure 4.1).

Using statistical analysis, Bourne et al. (2018) concluded that the incorporation of aftershocks represented an improvement. The authors also argued that incorporating aftershocks into the model was consistent with the observations of spatiotemporal clustering of events, which could not be explained assuming only earthquakes triggered by compaction and pore pressure changes. However, a more recent clustering analysis concluded that the proportion of clustered events, and accordingly aftershocks, is much less than what was obtained by Bourne et al. (2018), and argues that the cause of this discrepancy stems from the fact that their model misses second order processes such as rate dependence of elastic vs inelastic strain and earthquake nucleation. These processes are erroneously corrected by the ETAS model parameters (Muntendam-Bos, 2020).

More recent clustering analyses (Muntendam-Bos, 2020), however, observe that the proportion of clustered events, and accordingly aftershocks, is much less than what was estimated by Bourne et al. (2018).

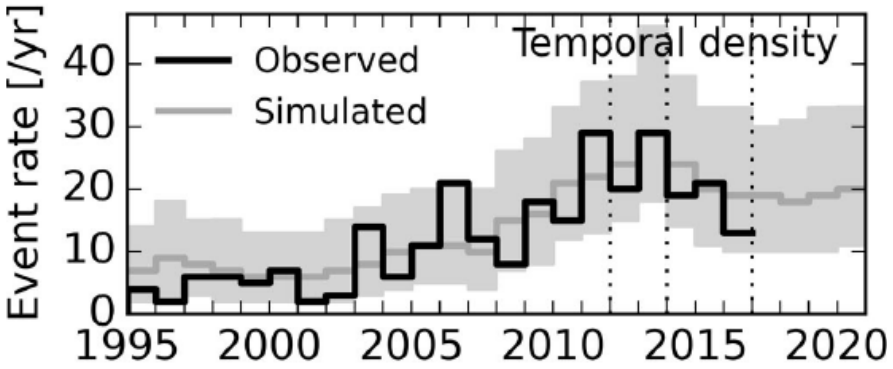

Fig. 4.1: *Example of modelling results using an activity rate model incorporating aftershocks (Figure 5a from Bourne et al., 2018). Years before 2012 were used for model training. Results after 2012 are forecasts. The shaded area is the 95% confidence interval.*

4.2.3 Elastic thin-sheet model

The seismogenic model version V1 considered only the vertical component of strain as it related to compaction within the intensity function (equations 4.6 and 4.7). In the following update of the model (i.e., version V2), Bourne and Oates (2015b) extended the formulation to consider the full strain tensor.

Using results obtained by Kostrov (1974) together with a reservoir model where lateral gradients in reservoir thickness and compaction were negligible relative to lateral gradients in the elevation of the base of the reservoir, the authors deduced an expression to estimate the total seismic moment per unit area with respect to the deviatoric strains within the reservoir, this is

$$\frac{1}{A}\sum_{k=1}^{N(t)} M_0^k m_{ij}^k = -\mu c(t)\left(\frac{1}{3}+\sqrt{1+|\nabla(z_0)|^2}\right)\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\lambda_2'}{\lambda_1'} & 0 \\ 0 & 0 & \frac{\lambda_3'}{\lambda_1'} \end{pmatrix}, \tag{4.9}$$

where A is the horizontal area of the reservoir volume under analysis, and M_0^k and m_{ij}^k are the seismic moment and moment tensor of the kth fault activated within that volume during the time interval t, respectively. On the right hand side of equation 4.9, λ_i' are the eigenvectors of the vertically-averaged deviatoric stress tensor rotated into its principal axes, μ is the shear modulus, and $|\nabla(z_0)|^2 = \left(\frac{\partial z_0}{\partial x}\right)^2 + \left(\frac{\partial z_0}{\partial y}\right)^2$, where z_0 describes the elevation of the base of the reservoir (Figure 4.2).

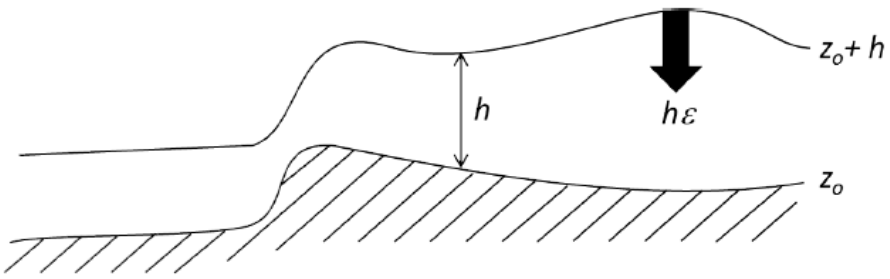

Fig. 4.2: *General model of uniaxial compaction within a thin-sheet reservoir geometry (Figure 3 from Bourne and Oates, 2015b).*

Assuming that lateral gradients in the elevation of the base of the reservoir are also negligible, equation 4.9 simplifies to

$$\sum_{k=1}^{N(t)} M_0^k m_{ij}^k = \frac{4}{3}\mu\Delta V(t)\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}, \tag{4.10}$$

where $\Delta V = A\bar{c}$ with $\bar{c}$ being the average reservoir compaction uniformly distributed within the volume V. This equation represents the type of compaction formulation used in model version V1. Comparing to equation 4.9, compaction in the case without neglecting lateral gradients in the eleva-

tion of the base of the reservoir becomes

$$c_t = c(t)\left(\frac{1}{3} + \sqrt{1 + |\nabla(z_0)|^2}\right) . \tag{4.11}$$

Furthermore, the authors incorporated a weighting scheme as

$$c_t = c(t)\left(\frac{1}{3}w_c + \sqrt{w_c + w_t|\nabla(z_0)|^2}\right) , \tag{4.12}$$

with $w_c = 1 - w_t$. For $w_t = 0$, equation 4.12 reduces to a c_t that represents an average reservoir compaction where strains are uniformly distributed over the reservoir (i.e., equation 4.10 or model version V1). For $w_t = 1$, equation 4.12 becomes $c_t = c(t)|\nabla(z_0)|$, which in the limit of lateral differentials being much smaller than reservoir thickness represents the effect of the offset induced between the two sides of a fault as a result of compaction. This limiting case represents a model where strains are localised on faults. Hence, the weighting scheme allowed to investigate which type of model (or combination thereof) better explained the observations in Groningen. Equation 4.12 was therefore used instead of c in the intensity function of equation 4.6 and consequently in equation 4.7 when incorporating aftershocks.

The formulation of the elastic thin-sheet model permitted to study the relative importance of effects related to uniformly distributed strains and to strains concentrated on faults to explain the seismicity observed in Groningen.

For the application of the model, fault throws and locations were extracted from seismic interpretation. Given the resolution of the seismic survey, fault throws of less than 25 m could not be resolved, although the power-law distribution of geological fault populations suggests that they exist in larger proportions than the mapped faults. Bourne and Oates (2015b) describe the calculation of the horizontal gradients of the top boundary of the reservoir (i.e., topographic gradients) via finite-difference modelling. As their derivations refer to the base of the reservoir (see Figure 4.2), this assumes that the gradients at the top and bottom of the reservoir are the same.

Comparing the horizontal location of mapped faults to the epicentres of observed seismicity, it was not straightforward to interpret which events could be associated to any one of the mapped faults. Therefore, the association of the observed seismicity to the mapped faults was an assumption and the probability density distribution constrained by the observed epicentres was considered a smoothed version of the higher resolution distribution connected to the faults.

The spatiotemporal analysis of the b-value of subsets of the observed seismicity produced inconclusive results on whether this parameter could be variant. Based on the possibility that the b-value could be related to strain, the authors proposed a model where the b-value varied with respect to c_t following a hyperbolic-tangent function. The type of function was selected to represent a monotonic transition between two bounding b-values set as 0.5 and 1.5. The tests reported by the authors

showed that models with constant b-value or b-values that increase or decrease with strain were all valid. As with other model parameters, the main limitation to constrain a particular b-value model was the lack of sufficient observational data.

The formulation of the elastic thin-sheet model improved the match of the spatial distribution of observed seismicity compared to the previous model version.

Compared to the activity rate model employing a uniform compaction distribution (V1), the updated model produced a better match of the event density maps of observed seismicity. Notice, however, that this improvement was not matched by improvements in the fitting of surface subsidence from the compaction model(s), which had been pointed out to show significant residuals where the seismogenic model version V1 also displayed mismatches in event density. Therefore, it is not clear to what extent the improvements in model version V2 were related to an increase in its complexity (i.e., more degrees of freedom), which can also be expected to improve fitting to the same amount of observational data. One of the suggestions from the authors in this regard was to work on improvements in the compaction model and the generation of more detailed subsidence maps using all available measurements.

4.2.4 Extreme threshold failures

The formulation of extreme threshold failures considers the activation of fault planes incorporating a Coulomb failure criterion and the principles of poroelasticity.

In the latest model version reviewed in this report (V5), Bourne and Oates (2017a) investigated an alternative formulation that explicitly took into account failure planes within the reservoir and their activation as a result of depletion and compaction following the principles of poroelasticity (Figure 4.3a). The location and geometry of the failure planes was again taken from the seismic interpretation. Activation was assumed in normal mode, with the initial stress conditions of the failure planes and their friction coefficients considered to vary with corresponding probability distributions. Assuming a Coulomb failure criterion, the planes with higher initial stress and lower friction coefficient failed first. Since these planes were located in the tail of the joint probability distribution, extreme value theory could be applied, which lead to exponential characteristics that reproduced the exponential temporal increase of seismicity in Groningen. Furthermore, by allowing lateral heterogeneities in the geometrical distribution of the failure planes and elastic properties of the reservoir represented by the topographic gradient (Figure 4.3b and c), the spatial foot print of the observed seismicity was reproduced better as well. The fault offset emerged as the crucial factor to switch between pore pressure dependent or compaction strain dependent stresses. The PPP intensity function incorporating extreme threshold failures considering a Coulomb failure criterion and taking into account reservoir heterogeneities within an elastic thin-

sheet reservoir took the form

$$\lambda = \rho h \frac{\partial P_f}{\partial t} \,, \tag{4.13}$$

where ρ is the volume density of fault segments and h the reservoir thickness. Variable P_f is the probability of failure deduced by the authors, which is affected by parameters related to the proximity to failure of the initial stress state and the sensitivity of the system to the stabilizing effect of pressure depletion as well as the destabilizing effect of shear strains induced by reservoir compaction.

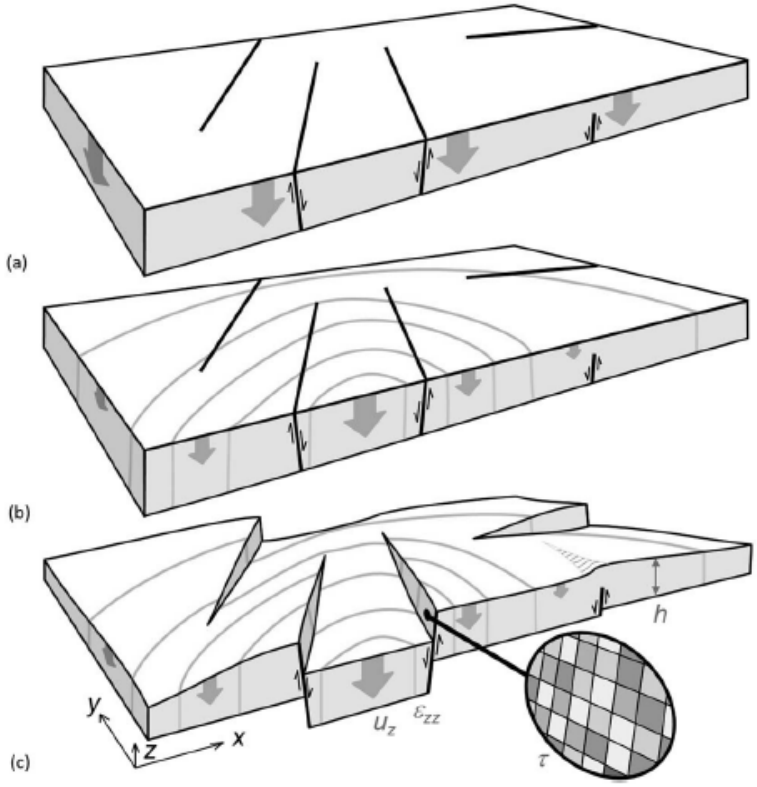

Fig. 4.3: *Different levels of complexity incorporated into the extreme threshold failures model. (a) failure planes within a uniform depleting thin sheet; (b) heterogeneous depletion and mechanical properties; (c) structural heterogeneity. Figure 1 from Bourne and Oates (2017b).*

Besides the application of extreme value theory, the model implied a smoothing of the strain data of the compaction model and the observed fault offsets. Faults with larger offsets were ignored because they were supposed to be aseismic due to their contact with the Zechstein salt formation. An initial limitation, also pointed out by the authors, was that this formulation was static, in the sense that every failure plane existed only as an isolated entity. In other words, stress interactions between failure planes and self-excitation were not represented. To alleviate this limitation, the incorporation of ETAS into this model, as had been the practice with the previous versions, is reported in Bourne et al. (2018).

In their assessment of different versions of the seismicity rate model (see Bourne et al., 2018, for details on the different scenarios investigated), the authors observed that the extreme threshold

failures model with exponential strain trend and b-values with an inverse power-law strain trend were best for the reproduction of the observations of seismicity in Groningen.

The other important aspect was the incorporation of geometrical, elastic and stress heterogeneities to better match the spatial distribution of hypocentres. As with previous activity rate models, the confidence intervals obtained with the extreme threshold failures model were tighter compared to those resulting from the earlier strain partitioning model (V0). The authors point out that an important reason for this better constraint was the switch from modelling the total seismic moment budget in the strain partitioning formulation to modelling the seismicity rate in the extreme threshold failures formulation (Bourne and Oates, 2017a, Figure 4.4 in this report).

Bourne et al. (2018) assess that geometrical, elastic and stress heterogeneities together with an exponential strain trend and b-values that follow an inverse power-law strain trend are important to improve the match of the observations of seismicity in Groningen.

4.2.5 Discussion

Aside from the first version of the seismogenic model (i.e., V0) proposed by Bourne and Oates and coauthors, the subsequent versions of the model are all based on a stochastic Poisson Point Process (PPP). Hence, starting from model version V1, the following versions constitute gradual increases in model complexity mostly related to the incorporation of more sophisticated physics-based formulations. These formulations are mainly incorporated through the intensity function that feeds the PPP.

Different valid scenarios have been investigated, mostly based on empirical evidence and its connection to plausible physical processes that could take place within the Groningen field. Unfortunately, the amount of observed data has been historically insufficient to constrain the parameters of the proposed models. Without properly constrained model parameters, the forecasts from the models can only be expected to reproduce the data used for calibration and possibly for short-term predictions. The results reported by the authors suggest that 5 years could be a reasonable limit to generate predictions.

Important elements necessary to validate the proposed models and that, upon validation, could improve the forecasting power are:

- Reducing uncertainty in epicentral earthquake locations will allow a better association between seismicity and mapped (active) faults, as well as with the subsidence maps. Upon improving event locations, it could also be possible to identify active faults not mapped in the seismic interpretation.
- Improving the compaction model, subsidence maps and modelling of the connection between

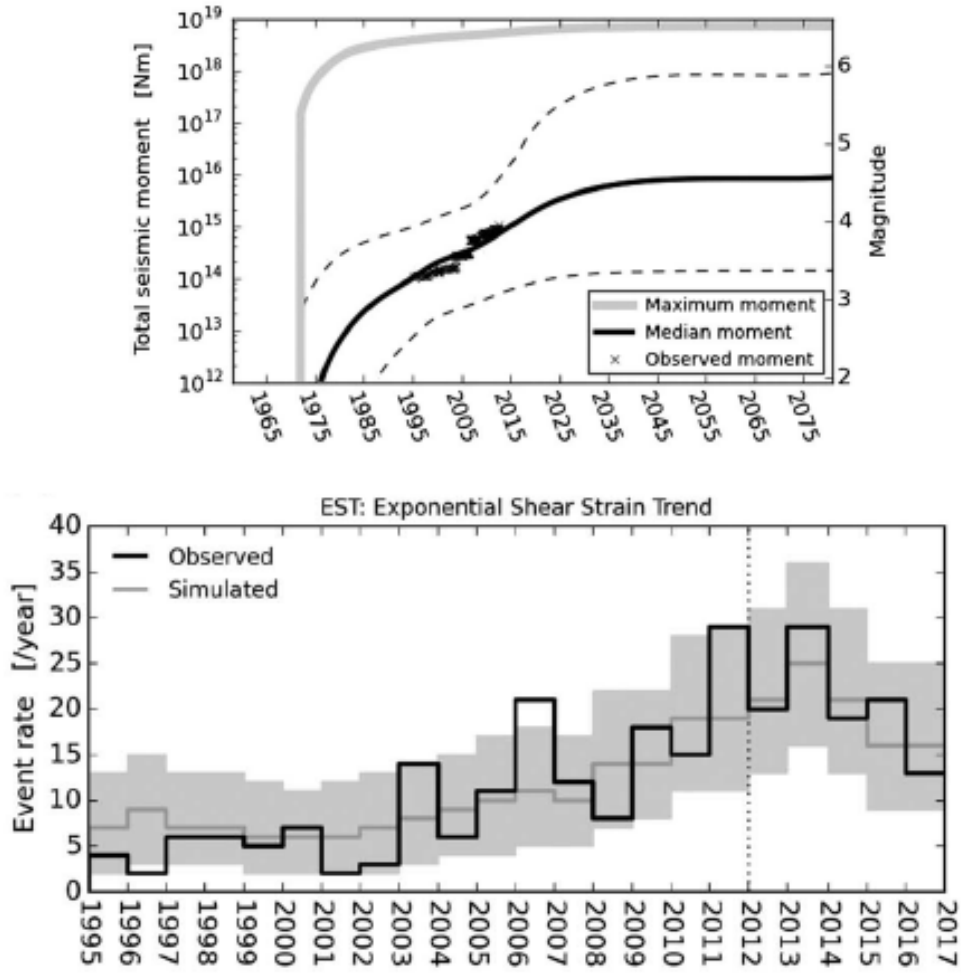

Fig. 4.4: *Fit and forecasting results. Top: strain partitioning modelling of total seismic-moment budget (Figure 19a from Bourne et al., 2014). Dashed lines are 95% confident intervals. Bottom: extreme threshold failure modelling of event rate (Figure 14d from Bourne and Oates, 2017b). Years before 2012 were used for model training. Results after 2012 are forecasts. Shaded area is the 95% confidence interval.*

both is critical to remove remaining spatial uncertainties on the connection between event density, compaction and faults.

- Reducing the minimum magnitude of the events included in the earthquake catalogue used to calibrate the model would increase the amount of observations, which can be expected to improve the constraint of model parameters (e.g., ETAS parameters, spatiotemporal variation of b-values, etc.).

5 Monte Carlo simulation of earthquake catalogues for seismic hazard assessment

The Monte Carlo simulation of earthquake catalogues applied for the probabilistic seismic hazard assessment (PSHA) of Groningen following the modelling of seismogenic sources is not part of the review, but we include a short description of the concept as we regard it to be important to understand potential consequences of its usage.

The Monte Carlo simulation of earthquake catalogues is perhaps the most general and easily adaptable method in order to characterise the number, magnitude and location of future earthquakes necessary to assess seismic hazard or risk (Bourne et al., 2014). The approach is particularly well suited to time-dependent induced seismicity in combination with ETAS-type activity, which cannot be treated as stationary in time, as is a standard assumption in PSHA (Bourne et al., 2015).

The Monte Carlo simulation approach is particularly well suited to incorporate time-dependency of induced seismicity in seismic hazard analysis.

The approach requires two essential inputs: a model for the location and magnitude of possible future earthquakes (for this, the described seismogenic source model is employed) as well as a ground motion model tailored to incorporate site effects (Bourne et al., 2015). The description of the applied ground motion model is outside of the scope of this report. However, a short description is given in Goertz-Allmann et al. (2018).

The ground motions are aggregated over an ensemble of simulated catalogues to result in a probabilistic representation of ground motion hazard. This can be done either by an integration over the distribution of earthquake locations and frequencies as well as the variability of predicted ground motions or by sampling via a Monte Carlo simulation (Ebel and Kafka, 1999; Musson, 1999; Musson, 2000; Assatourians and Atkinson, 2013; Pagani et al., 2014). The comparison of advantages and disadvantages of each approach is beyond the focus of this review.

The advantage of the Monte Carlo seismic hazard simulation is that it avoids the problem of overestimating the aggregated damage or losses.

Bourne et al. (2015) argue for the Monte Carlo seismic hazard simulation as it avoids the problem that the aggregated damage or losses are overestimated, when conventional PSHA is applied to multiple locations simultaneously (Crowley and Bommer, 2006). This behaviour is caused by the the nature of the ground motion variability, which can be decomposed into between-event and within-event components. The between-event variability describes the earthquake-to-earthquake variability originating from the

influence of source parameters that are not included in the ground motion prediction equations (GMPEs), e.g. stress drop and slip distribution on the fault. The within-event variability reflects the record-to-record variability caused by differences in travel paths and site response characteristics. The within-event variability is generally much larger than the between-event variability (Bourne et al., 2015). Since the conventional PSHA does not differentiate between both parts, parallel calculations for multiple sites treat all variability as between-event variability. Monte Carlo simulations, on the other hand, allow to sample a value for the between-event variability for each earthquake and values for the within-event variability for each location.

For the model version V0, a standard Monte Carlo procedure revealed the set of parameter pairs of the strain-partitioning function assumed to describe the fraction of the average total strain that is seismogenic (Bourne et al., 2015). For each acceptable parameter pair, a total seismic moment was calculated for a given reservoir compaction model and repetition for every set of acceptable model parameters yielded a probability distribution of the total seismic moment. Likewise, a relative probability map of event epicentres and a normalized seismic moment density map were computed (Bourne et al., 2014). For each event in the synthetic catalogue, a single random independent value for the event-specific epsilon (corresponding to the number of standard deviations away from the logarithmic value of, in this case, the between-event variability) was chosen. For each observation point, a single random independent value for the within-event epsilon was selected. For each combination of event and observation point, the ground motion was calculated respecting the chosen amount of variability (Bourne et al., 2015). To obtain the seismic hazard, the full procedure was reiterated, until the exceedance rates were sufficiently free from stochastic error for all surface observation points and ground motion thresholds of interest (Bourne et al., 2015).

From model version V1 on, modelling of the total seismic moment was replaced by the modelling of the seismicity rate based on a Poisson Point Process (PPP) as described in chapter 4. Therefore, the base data for sampling the earthquake catalogues must have changed accordingly. However, unfortunately, there is no mentioning of the ramifications of the model updates on the probabilistic seismic hazard analysis, since later publications and reports focus only on the seismicity model itself and its forecasts related to production scenarios.

Uncertainties related to seismic hazard computations may be subdivided into two classes: aleatory and epistemic uncertainty. Aleatory variability is the true statistical variability representing the stochastic nature of a process and is taken care of by the random sampling of the distributions during the Monte Carlo simulations (Bourne et al., 2015).

Uncertainties related to seismic hazard computations consist of aleatory and epistemic uncertainty. Aleatory variability is the true statistical variability representing the stochastic nature of a process. Epistemic uncertainties aer uncertainties related to choices of parameters or models that are only known to a certain accuracy or extent.

Epistemic uncertainty represents uncertainties related to choices of parameters or models that are only known to a certain accuracy or extent, and that for the Groningen field comprise for example - depending on the model version - the strain partitioning function, the reservoir compaction, the variation of the b-value with compaction, the maximum possible magnitude, the choice of compaction model, the assumed ground motion model, and future production levels. In the newest model, thirteen model parameters and their uncertainties are included: four parameters describe the extreme threshold model, three parameters characterise earthquake magnitudes and six parameters define the ETAS model (Bourne et al., 2018).

Disaggregation demonstrates that the largest contribution to seismic hazard results from earthquakes of M_W 4 - 5 at short hypocentral distances. Main uncertainties arise from uncertainty about the future fraction of induced strains that will be seismogenic and the unknown scaling of ground motions and their variability to larger magnitudes.

The disaggregation demonstrates that the largest contribution to the seismic hazard in the Groningen field are earthquakes of M_W 4 - 5 at short hypocentral distances of 3 km as well as ground motions with a size of two standard deviations above the median. The main uncertainties arise from the uncertainty about the future fraction of induced strains that will be seismogenic and the unknown scaling of ground motions and their variability to larger magnitudes (Bourne et al., 2015).

5.1 Assessment

Although not as commonly used as a conventional PSHA, the Monte Carlo method to compute seismic hazard by sampling distributions of locations and magnitudes of possible future earthquakes and ground motions at observation sites, seems to be well fitted in the present case of the Groningen field due to the following arguments by the authors:

- easily adaptable method in order to characterise the number, magnitude and location of future earthquakes;
- approach is particularly well suited to time-dependent induced seismicity, which cannot be treated as stationary in time, which is a standard assumption in PSHA;
- avoiding the problem that when the conventional PSHA is applied at multiple locations simultaneously, the aggregated damage or losses are overestimated.

6 Discussion of strengths and weaknesses of the approach

The influence of static and dynamic reservoir model on the final results of the seismogenic model is practically not traceable for readers of the publications.

The seismogenic model consists of empirical assumptions and physical calculations based on production data and statistical model components. All model components and choices seem reasonably chosen, but the total number and values of the involved model parameters are hard to assess. This is because parameters for the seismological model as well as parameters and assumptions in the static and dynamic reservoir model can have an influence on the results. Hence, the final result is practically only reproducible by the authors themselves and not traceable or quantitatively assessable for readers of the publications.

Reservoir compaction is the main input parameter to the model and may be biased by reservoir depth, the uncertainty in the estimate of the reservoir porosity and the uncertainty in the relationship between porosity and uniaxial compressibility of the bulk reservoir. These uncertainties allow for alternative compaction models to be compatible with the available data.

Reservoir compaction is the main input to the model, which may be biased by several factors. Firstly, the ability to laterally resolve reservoir compaction based on surface subsidence is limited by the reservoir depth, because reservoir compaction at any given location yields a dedicated bowl of surface subsidence with a radius related to reservoir depth (van Oeveren et al., 2017). Secondly, the uncertainty in the estimate of the reservoir porosity influences the distribution of uniaxial compressibility used in the reservoir compaction model, especially away from well controls. Furthermore, the uncertainty in the relationship between reservoir porosity and uniaxial compressibility of the bulk reservoir, which is due to limited sampling and differences in the length scale between laboratory measurements and field, is represented as a single field-wide scalar parameter constrained by minimising the misfit between computed surface subsidence and subsidence measurements. The subsurface mechanical properties surrounding the reservoir are extrapolated from measurements of petrophysical logs and core materials. This leads to uncertainties that influence the relationship between reservoir compaction and surface subsidence. The subsidence is predicted best in the Western and Central part of the Groningen field; the estimates are too high in the Northeast and too low in the South. Also the residuals show a clear areal trend that is most likely related to physical elements not included in the model (van Oeveren et al., 2017). In addition, the plastic behaviour of the salt layers above the reservoir is not being modelled by the dynamic compaction model. This could be responsible for parts of the errors in reproducing the subsidence observations and might also explain some delay in subsidence at the start of production (NAM, pers. comm., 2019). All these uncertainties and simplifications allow alternative compaction models that are as well compatible with available reservoir and geodetic data,

but yield different predictions for future reservoir compaction (de Waal, 1986; Mossop, 2012; Arup, 2013; Bourne et al., 2015).

The data fit of the seismogenic model does not explicitly consider earthquake location errors, which may affect the optimised strain smoothing kernel with unknown consequences.

The second important input data set, which is used to constrain the seismogenic forecast model, is the earthquake catalogue. Although no individual location uncertainties are provided in the publicly available catalogue or in the bulletin data (Dando et al., 2018), the location errors can be considerable depending on time and location with standard deviations between 500 and 1000 metre (Spetzler and Dost, 2017). However, the data fit of the NAM seismogenic model does not explicitly consider those earthquake location errors.

The most important uncertainty of the strain partitioning model is in the estimation of the strain partitioning function leading to large uncertainties about the future percentage of induced strains that will be seismogenic.

Apart from the uncertainties of the input data, some critical aspects concern the two NAM model types: the strain partitioning model (V0) and the extreme failure threshold model (V5), which is the latest version of the activity type models. The most important uncertainty of the strain partitioning model (V0; Bourne et al., 2014) lies in the estimation of the strain partitioning function. The fit of a generalised logistic function is subject to significant uncertainties due to the limited number of observed earthquakes; attempts to quantify this variability lead to estimates for a probability distribution of total seismic moment with a heavy tail of large values, partly due to the Pareto sum distribution characterising the random variability in observations of the total seismic moment. The Pareto sum distribution is approximated by a log normal distribution, where the standard deviation is equal to the approximation for Pareto sums obtained by Zaliapin et al. (2005). As indicated by the authors, this approximation leads to a scale bias in the residuals systematically underestimating the regions with the largest seismic moments due to the log normal distribution under-representing the heavy upper-tail of the Pareto sum distribution. The authors claim that this bias can be corrected by minimising the least squares difference between the observed and computed total seismic moment for the entire reservoir, if weighted in the same manner. However, it remains unclear how well this correction works. Furthermore, epistemic uncertainties about true factors governing the evolution of strain partitioning are involved; e.g. the role of compaction might be better described by alternative functional forms such as an inverse power law or strain partitioning may depend on other factors such as fault or reservoir properties. These unknowns lead to large uncertainties about the future fraction of induced strains that will be seismogenic and the unknown scaling of ground motions and their variability to larger magnitudes.

The extreme threshold failures model (V5; Bourne and Oates, 2017b) contains some assumptions as well that might be questionable. Perhaps the most important is the assumption that the average initial stress level is homogeneous in the whole reservoir, but that the initial stresses are heterogeneous in each location due to small-scale variability. Another major assumption is that the failure threshold lies in the tail of this stress distribution in each place. The latter allows the application of extreme value statistics, which leads to the generalised Pareto model. This model depends in general on three parameters, but the NAM model finally uses only its simplest form, an exponential function with two free parameters. The applicability of the extreme value statistics for the seismogenic model requires that the failure stress stays in the tail of the stress distribution from the beginning until the end of the production. While this might be a reasonable assumption for the beginning of the production, we doubt that this still holds after induced stresses of more than 20 MPa. If not, the reactivation of pre-existing faults with stresses in the body of distribution will lead to a deviation from the exponential trend as reservoir transitions to steady state, where induced seismic activity rates are proportional to induced strains. Because the absolute stresses are not known, the time of the transition is also unknown, which limits the use of the model for future seismicity. Once the stress state leaves the distribution tail, the model will overestimate the seismicity rate and thus will represent a conservative prediction.

For the extreme threshold failures model, the main assumption is that the failure threshold lies within the tail of the stress distribution. Once the reservoir transitions into a steady state, induced seismic activity rates become proportional to induced strains and the model will overestimate the seismicity rate.

Besides the application of the extreme value theory, the model implies a smoothing of the strain data of the compaction model and the observed fault offsets. Here, faults with largest offsets are ignored because they are considered aseismic due to contact with the Zechstein salt formation. This involves the risk of excluding major faults that may be activated in the future. Furthermore, the width of the smoothing kernel is optimised by the fit to the observed earthquake data. Because the epistemic earthquake location uncertainties are not considered in the fitting procedure as mentioned above, the optimised strain smoothing kernel is likely affected by these with unclear consequences.

Faults with the largest offsets are considered aseismic due to their contact with the Zechstein salt formation, which involves the risk of having excluded major faults that may become activated in the future.

Aftershock triggering is considered only in a stochastic way by the use of the ETAS model rather than as a consequence of the mainshock induced stress transfers.

The use of the ETAS model for aftershock triggering is reasonable, but involves six additional fitting parameters, which are difficult to constrain by the few observed aftershocks in the Groningen field.

Thus, the model is not consistent in this regard. However, a deterministic modelling of the coseismic stress transfer is unattainable due to the involved uncertainties and therefore, the use of the well-known stochastic ETAS model is reasonable. Nonetheless, the ETAS model involves six additional fitting parameters, which are almost impossible to constrain by the few observed aftershocks in the Groningen field. Large epistemic uncertainties are the result.

Also the fit parameters for the b-value are not well constrained by the small data set. In addition, use of the maximum magnitude value implicitly allows for earthquake ruptures to extend outside the reservoir.

The extreme threshold failures model (V5; Bourne et al., 2018) leads to significantly reduced uncertainties in the seismicity rates, but the seismic moment is not anymore controlled by the induced strains and stresses. Now the seismic moment release is determined by the Gutenberg-Richter b and M_{max} values. Here, the b-value and its dependence on stress is obtained by data fitting. However, the assumed functional form is not well established and the three fit parameters for b are not well constrained by the rather small observational data set. In addition, the seismic moment release depends on the maximum magnitude value, which is taken in the range between 3.75 and 7.25 according to independent estimations (Bommer and Van Elk, 2017). Therefore, this approach implicitly abandons the previous approach of Bourne et al. (2014) and partly allows for earthquake ruptures to extend outside the reservoir. Nevertheless, rupture extensions are not explicitly modelled and the question of whether or not basement faults might be activated is not addressed. This also holds for potential activity triggered in neighbouring regions related to depletion extending outside of the reservoir.

We consider the model approach to be generally reasonable and sophisticated. The epistemic uncertainties related to the compaction model can be addressed by different branches in the logic tree and aleatoric uncertainties can be quantified by a large number of Monte Carlo simulations.

Despite the involved assumptions, simplifications, and uncertainties, we consider the model approach to be generally reasonable and sophisticated. All models with some predictive power for variable future production scenarios need to be based on production data, particularly on compaction or pore pressure estimations, which always involves uncertainties. In seismic hazard studies, the epistemic uncertainties related to the compaction model can be simply addressed by different branches in the logic tree as discussed by Bourne et al. (2015). Furthermore, the assumption of an inhomogeneous Poisson model for the production-induced mainshocks and the ETAS model for aftershock activity allows the quan-

tification of the involved aleatoric uncertainties based on a large number of Monte Carlo simulations. Hazard estimations can be simply calculated by additionally sampling from ground-motion prediction equations within these calculations. The Monte Carlo approach to compute seismic hazard is particularly well suited to time-dependent induced seismicity, which cannot be treated as stationary in time, which is a standard assumption in PSHA. In addition, it avoids the problems that when the conventional PSHA is applied at multiple locations simultaneously, the aggregated damage or losses are overestimated.

The epistemic uncertainties of all involved parameters of the seismogenic model can be in principle considered in the logic tree approach for PSHA, although it is not fully done so far because of computational costs (NAM, pers. comm., 2019). Furthermore, the model has passed several internal retrospective tests (NAM, pers. comm., 2019) as well as a published pseudo-prospective test (V5; Bourne and Oates, 2017b). The latter was based on the model calibrated on the seismicity between 1995 and 2012 and tested for the spatiotemporal event density in the five subsequent years. The current model's forecasts are checked approximately every six months against new earthquake data, but so far, the availability of new data never invalidated the model NAM (pers. comm., 2019). These checks are out-of-sample likelihood tests to assess model performance; the model is re-evaluated against the updated out-of-sample event catalogue NAM (pers. comm., 2019). The evaluated model is not retrained as the in-sample event catalogue used for training remains unchanged NAM (pers. comm., 2019).

Thus, the extreme value assumption seems to be still holding (NAM, pers. comm., 2019). By regularly re-evaluating the forecast performance and restricting forecast period (typically to 5 years), the quality of the forecasts should be assured in the future.

7 Outlook

The following points were discussed during our meeting with NAM at Schiphol airport (4th June 2019) as potential future work on the seismogenic model to be conducted by NAM:

- The model adopts the recommendations from the expert panel workshop in a logic tree (Coppersmith et al., 2016). But one of the main questions remains open so far: is there a stress-dependent taper for the maximum expected magnitude?
- Well testing for radioactive bullets to measure in-situ deformation was reduced, because due to the decreased depletion, the signal becomes harder to interpret. Fiber optic cable measurements are an alternative; measurements are instantly available and the signal is measured constantly with higher quality. So far, however, fiber optic cable is only installed in one single well, as it has to be installed behind the casing (not possible for previously cased wells). Such measurements are a promising, but still new technology.
- The main faults within the Groningen field tend to strike in two general directions, and there are some indications that stress may be anisotropic in parts of the field. The model framework can be extended to incorporate focal mechanisms, but so far, only $\tilde{1}$00 mechanisms have been computed (Willacy et al., 2018).
- In the last years, the earthquake detection ability has improved significantly and the magnitude of completeness has lowered. The improved data set can be used to test the model with more statistical power.
- So far, no seasonality can be observed in the earthquake catalogue, but it is not clear for the geodetic measurements. In order to answer the question if an additional time-depending effect caused by the brine-saturated layer beneath the reservoir can be observed, the secular response versus the delay due to a seasonal depletion response must be separated (which is currently under research).
- Comprising the effects of an injection of nitrogen or CO_2 will be beyond the limits of the current model; especially, different fault systems might be activated.

Appendices

A Assumptions

The following is a list of some of the assumptions made by the authors for the development of the different versions of the seismogenic model. The list is not exhaustive. The model versions indicated are the first model version for which the assumption is stated; this does not exclude that the assumptions hold for later model versions as well.

General assumptions (in the order of publication of papers):

A.1 The large number of natural faults within and around the Groningen reservoir with density and strike varying throughout the field formed during an extensional phase during the Late Carboniferous and Early Permian. Faults have been reactivated during several phases, lately due to reservoir compaction within the field (V0; Bourne et al., 2014).

A.2 Following Houtgast (1995), the contributions of tectonic deformation and postglacial rebound are excluded in the model due to the very low natural seismicity (V0; Bourne et al., 2014).

A.3 The amount of strain induced by fluid extraction is the mechanism for governing the amount of slip induced on critically stressed faults (V0; Bourne et al., 2014).

A.4 The increase in both the number of events and total seismic moment per unit reservoir volume change is the suggested mechanism responsible for the observed escalation in activity rate and seismic moment release with production. The computed seismic moment release rate depends on two competing effects. First, the seismic moment released per unit gas production increases with reservoir compaction. Second, the rate of gas production eventually decreases due to reservoir pressure depletion (V0; Bourne et al., 2014).

A.5 In order to evaluate the total strain integral associated with the subsurface reservoir, the reservoir is assumed to have the geometry of a thin sheet, meaning that its lateral extents are significantly larger than its vertical extent. Thus, due to symmetry, deformations induced by an internal body force (such as a change in reservoir pore fluid pressure) will be approximately uniaxial (V0; Bourne et al., 2014).

A.6 Seismic strain increases more quickly than bulk strain (V0; Bourne et al., 2014) as can be suspected from laboratory measurements (Ojala et al., 2003; Heap et al., 2009) and volcanic eruptions (Bell and Kilburn, 2012).

A.7 Predictions about future changes in seismicity are based on predictions from the current gas production plans. Thus, forecasts of future seismicity should be limited to relatively short time periods (3 - 5 years) and the seismological model should be subjected to frequent revalidations and if necessary, adaptations (V0; Bourne et al., 2014).

A.8 The exponential-like trend between earthquake nucleation rate as function of spatially and temporally variable compaction is based on the observed earthquake activity rate trend relative to reservoir compaction inferred from geodetic monitoring of surface displacement (V1; Bourne and Oates, 2017a).

A.9 Aftershock sequences have been incorporated in the rate function using the Epidemic Type Aftershock Sequence (ETAS, Ogata, 1998; Ogata, 2011) model (V1; Bourne and Oates, 2017a).

A.10 Magnitudes of the modelled earthquake occurrences are drawn from a Cornell-Vanmarcke probability distribution (Cornell and Vanmarcke, 1969), which is a truncated exponential distribution characterised by a slope controlled by the b-value and an upper bound controlled by the maximum expected magnitude (V1; Bourne and Oates, 2017a).

A.11 Explanation of the observed trend of increasing seismicity rates with reservoir compaction by progressive failure of a heterogeneous fault population with an increasing fraction of fault systems becoming seismically active under a steadily increasing mechanical load (V5; Bourne and Oates, 2017b).

A.12 Description of the induced fault reactivations by means of an extreme threshold failure model (V5; Bourne and Oates, 2017b).

A.13 Use of a simple discrete fault strain model, where the seismic moment scales with the product of reservoir compaction and pre-existing reservoir offset across the fault (V5; Bourne and Oates, 2017b).

A.14 Assumption of a thin sheet continuum strain model, where seismicity only relates to the principal deviatoric strain tensor (i.e. earthquakes play not part in accommodating the isotropic strain) and where lateral gradients in the reservoir thickness and reservoir compaction are small relative to the lateral gradients in elevation of the base of the reservoir (V5; Bourne and Oates, 2017b).

A.15 Employment of a Poisson point process model for the induced seismicity (V5; Bourne and Oates, 2017b).

A.16 Fault model has been built from geometries of all pre-existing geological faults with pre-existing preproduction vertical offset of at least 25 m as imaged and mapped at the top reservoir surface with 3D reflection seismics. However, due to the presence of the Zechstein salt formation directly above the reservoir, pre-existing fault offsets above some threshold may not be seismogenic due to juxtaposition of the reservoir against a ductile material and are thus excluded, if the fault offset to reservoir thickness ratio exceeds a threshold (V5; Bourne and Oates, 2017b).

A.17 The achievable resolution of the seismological model is limited by the available monitoring data. This resolution is controlled and optimized by smoothing the total shear strain field with the smoothing kernel being assumed as normalized multivariant Gaussian (V5; Bourne and Oates,

2017b). (Both the standard deviation of this strain smoothing kernel as well as the fault offset to reservoir thickness ratio are included as two additional unknown model parameters).

Assumptions of the dynamic reservoir model:

B.1 All subsidence is assumed to be caused by compaction in the Groningen field or adjacent fields only (van Oeveren et al., 2017).

B.2 Upscaled properties from static reservoir model include net-to-gross, porosity and horizontal as well as vertical permeability (van Oeveren et al., 2017).

B.3 The 45 initialisation regions are based on observed free water levels and pressure lags (van Oeveren et al., 2017).

B.4 The temperature at datum depths varies laterally (van Oeveren et al., 2017).

B.5 Two further properties were not upscaled, since their distribution throughout the field is relatively homogeneous (Burkitov et al., 2016): capillary pressure is modelled using Brooks-Corey based on saturation logs and the relative permeability is modelled based on 19 core plugs (van Oeveren et al., 2017).

B.6 The Alvestad inflow equations are employed to model the wells (van Oeveren et al., 2017).

B.7 Matrix compressibility is modelled as a function of porosity based on trends observed in core experiments (van Oeveren et al., 2017).

B.8 Compaction is converted to surface subsidence using the Geertsma and van Opstal equation (Geertsma and van Opstal, 1972; van Oeveren et al., 2017) including a significant simplification of the overburden (Burkitov et al., 2016; van Oeveren et al., 2017).

B.9 A Poisson's ratio of 0.25 is assumed for subsidence computation (van Oeveren et al., 2017).

B.10 96 model parameters are used in the history-matching workflow; of those, parameters that are uncertain and are likely to impact the mismatch to the data are decided on by a multidisciplinary subsurface team (van Oeveren et al., 2017).

B.11 Due to the large number of variable parameters, the number of simulations has to be limited and a modified Van der Corput (van der Corput, 1935) space-filling approach is employed to select a set of 1000 models (van Oeveren et al., 2017).

B.12 Further assumptions on parameter choices and computation of parameters can be found in SGS Horizon B.V. (2016) and Burkitov et al. (2016). Since this review is concerned with the seismogenic source model and not the dynamic reservoir model, it is outside the scope to list them all.

Assumptions of the reservoir compaction model:

C.1 The pressure depletion throughout the reservoir is inferred from the dynamic reservoir model constrained to match the history of metered gas production and measured reservoir pressure depletion (V0; Bourne et al., 2014).

C.2 The initial simple model considered for reservoir compaction in response to pressure changes is linear poroelasticity representing compaction as a product of reservoir pressure depletion, net reservoir thickness and uniaxial compressibility of the bulk reservoir (V0; Bourne et al., 2014).

C.3 The net reservoir thickness is taken from the static reservoir model constrained by reflection seismic data and well control (V0; Bourne et al., 2014).

C.4 The distribution of uniaxial compressibility depends on the reservoir porosity taken from the static reservoir model constrained by petrophysical well logs (V0; Bourne et al., 2014).

C.5 The total reservoir pore volume is constrained to match the volume of gas initially in place obtained from the analysis of pressure depletion versus gas production data (V0; Bourne et al., 2014).

C.6 The relationship between reservoir porosity and uniaxial compressibility of the bulk reservoir is based on laboratory measurements of plug and core samples recovered from the reservoir (V0; Bourne et al., 2014).

Mathematical assumptions:

D.1 The Pareto sum distribution is approximated with a log normal distribution, where the standard deviation is equal to the approximation of Pareto sums obtained by Zaliapin et al. (2005), described in Bourne et al. (V0; 2014).

B Literature

References

Arup (2013). *Technical addendum to the winningsplan Groningen 2013: Subsidence, induced earthquakes and seismic hazard analysis in the Groningen field*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Assatourians, K. and G.M. Atkinson (2013). "EqHaz: An open-source probabilistic seismic-hazard code based on the Monte Carlo simulation approach". In: *Seismological Research Letters* 84.3, pp. 516–524.

Bell, A.F. and C.R.J. Kilburn (2012). "Precursors to dyke-fed eruptions at basaltic volcanoes: insights from patterns of volcano-tectonic seismicity at Kilauea volcano, Hawaii". In: *Bulletin of volcanology* 74.2, pp. 325–339.

Bierman, S., F. Kraaijeveld, and S. Bourne (2015). *Regularised direct inversion to compaction in the Groningen reservoir using measurements from optical leveling campaigns*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Bommer, J.J. and J. Van Elk (2017). "Comment on "The maximum possible and the maximum expected earthquake magnitude for production-induced earthquakes at the gas field in Groningen, The Netherlands" by Gert Zöller and Matthias Holschneider". In: *Bulletin of the Seismological Society of America* 107.3, pp. 1564–1567.

Bourne, S. and S. Oates (2014). *An activity rate model of induced seismicity within the Groningen field*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Bourne, S. and S. Oates (2015a). *An activity rate model of induced seismicity within the Groningen field (part 1)*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Bourne, S. and S. Oates (2015b). *An activity rate model of induced seismicity within the Groningen field (part 2)*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Bourne, S.J. and S.J. Oates (2017a). "Development of statistical geomechanical models for forecasting seismicity induced by gas production from the Groningen field". In: *Netherlands Journal of Geosciences* 96.5, s175–s182.

Bourne, S.J. and S.J. Oates (2017b). "Extreme threshold failures within a heterogeneous elastic thin sheet and the spatial-temporal development of induced seismicity within the Groningen gas field". In: *Journal of Geophysical Research: Solid Earth* 122, pp. 10299–10320.

Bourne, S.J., S.J. Oates, J.J. Bommer, B. Dost, J. Van Elk, and D. Doornhof (2015). "A Monte Carlo Method for Probabilistic Hazard Assessment of Induced Seismicity due to Conventional Natural Gas Production". In: *Bulletin of the Seismological Society of America* 105.3, pp. 1721–1738.

Bourne, S.J., S.J. Oates, and J. Van Elk (2018). "The exponential rise of induced seismicity with increasing stress levels in the Groningen gas field and its implications for controlling seismic risk". In: *Geophysical Journal International*.

Bourne, S.J., S.J. Oates, J. Van Elk, and D. Doornhof (2014). "A seismological model for earthquakes induced by fluid extraction from a subsurface reservoir". In: *Journal of Geophysical Research: Solid Earth* 119.12, pp. 8991–9015.

Burkitov, U., H. van Oeveren, and P. Valvatne (2016). *Groningen field review 2015: subsurface dynamic modelling report*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Coppersmith, K., H. Bungum, A. McGarr, I. Wong, J. Ake, T. Dahm, I. Main, and B. Youngs (2016). *Report from the expert panel on maximum magnitude estimates for probabilistic seismic hazard and risk modelling in Groningen gas field*. Mmax Expert Workshop, 8 - 10 March 2016, World Trade Centre, Schiphol Airport, the Netherlands.

Cornell, C. A. (1968). "Engineering seismic risk analysis". In: *Bulletin of the seismological society of America* 58.5, pp. 1583–1606.

Cornell, C.A. and E.H. Vanmarcke (1969). "The major influences on seismic risk". In: *Proceedings of the fourth world conference on earthquake engineering*. Vol. 1, pp. 69–83.

Crowley, H. and J.J. Bommer (2006). "Modelling seismic hazard in earthquake loss models with spatially distributed exposure". In: *Bulletin of Earthquake Engineering* 4.3, pp. 249–273.

Dando, B, B Goertz-Allmann, D Kühn, N Langet, V Oye, and A Wüstefeld (2018). *Review of the public KNMI induced earthquake catalogue from the Groningen gas field (report project phase 1, WP1: catalogue review*. Tech. rep. NORSAR.

de Waal, J. A. (1986). "On the rate type compaction behaviour of sandstone reservoir rock". PhD thesis. Technische Hogeschool Delft, The Netherlands.

DeDontney, N. et al. (2016). *Maximum Magnitude of Induced Earthquakes in the Groningen Gas Field*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Dost, B. (2016). *Evolution of the Groningen earthquake monitoring network and event catalogue*. Mmax Expert Workshop, 8 - 10 March 2016, World Trade Centre, Schiphol Airport, the Netherlands.

Dost, B., B. Edwards, and J.J. Bommer (2018). "The Relationship between M and M_L: a Review and Application to Induced Seismicity in the Groningen Gas Field, The Netherlands". In: *Seismological Research Letters* 89.3, pp. 1062–1074.

Dost, B., F. Goutbeek, T. Van Eck, and D. Kraaijpoel (2012). *Monitoring induced seismicity in the North of the Netherlands: status report 2010*. Tech. rep. KNMI.

Dost, B. and H.W. Haak (2007). "Natural and induced seismicity". In: *Geology of the Netherlands*. Ed. by T.E. Wong. Royal Netherlands Academy of Arts and Sciences, Amsterdam, The Netherlands, pp. 223–239.

Dost, B., E. Ruigrok, and J. Spetzler (2017). "Development of seismicity and probabilistic hazard assessment for the Groningen gas field". In: *Netherlands Journal of Geosciences* 96.5, s235–s245.

Dost, B., T. Van Eck, and H. Haak (2004). "Scaling of peak ground acceleration and peak ground velocity recorded in the Netherlands". In: *Bollettino di Geofisica Teorica ed Appliocata* 45.3, pp. 153–168.

Ebel, J.E. and A.L. Kafka (1999). "A Monte Carlo approach to seismic hazard analysis". In: *Bulletin of the Seismological Society of America* 89.4, pp. 854–866.

European Committee for Standardization (2005). *Eurocode 8: Design of structures for earthquake resistance-part 1: general rules, seismic actions and rules for buildings*. Tech. rep. Brussels.

Fokker, P. and K. Van Thienen-Visser (2016). "Inversion of double-difference measurements from optical leveling for the Groningen gas field". In: *International Journal of Applied Earth Observation and Geoinformation* 49, pp. 1–9.

Geertsma, J. and G. Van Opstal (1973). "A numerical technique for predicting subsidence above compacting reservoirs, based on the nucleus of strain concept". In: *Verhandelingen Kon. Ned. Geol. Mijnbouwk. Gen.* 28, pp. 63–78.

Geertsma, J. and G. van Opstal (1972). *A numerical technique for predicting subsidence above compacting reservoirs, based on the nucleus of strain concept*. Tech. rep. Koninklijke/Shell Exploratie en Productie Laboratorium (Rijswijk).

Goertz-Allmann, B, D Kühn, N Langet, C Lindholm, A Meslem, V Oye, and A Wüstefeld (2018). *Review of the public KNMI induced earthquake catalogue from the Groningen gas field (report project phase 1, WP2: qualitative sensitivity analysis*. Tech. rep. NORSAR.

Harris, C.K. and S.J. Bourne (2015). *Maximum likelihood estimates of b-value for induced seismicity in the Groningen field*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Heap, M.J., P. Baud, P.G. Meredith, A.F. Bell, and I.G. Main (2009). "Time-dependent brittle creep in Darley Dale sandstone". In: *Journal of Geophysical Research: Solid Earth* 114.B7.

Hol, S., A. van der Linden, S. Bierman, F. Marcelis, and A. Makurat (2018). "Rock physical controls on production-induced compaction in the Groningen field". In: *Scientific Reports* 8, p. 7156.

Houtgast, G. (1995). "Aardbevingen in Nederland". In: *Koninklijk Nederlands Meteorologisch Instituut, De Bilt, The Netherlands* 179, p. 166.

Kole, P. (2015). *In-situ compaction measurements using gamma ray markers*. Tech. rep. NAM.

Kostrov, V. (1974). "Seismic moment and energy of earthquakes, and seismic flow of rocks". In: *Izv. Acad. Sci. USSR Phys. Solid Earth* 1, pp. 23–44.

Langenbruch, C. and M.D. Zoback (2016). "How will induced seismicity in Oklahoma respond to decreased saltwater injection rates?" In: *Science advances* 2.11, e1601542.

Mayuga, M. and D. Allen (1969). "Subsidence in the Wilmington oil field, Long Beach, California". In: *1st International Symposium on Land Subsidence*.

Mobach, E. and H. Gussinklo (1994). "In-situ reservoir compaction monitoring in the Groningen field". In: *Eurock SPE/ISRM Rock Mechanics in Petroleum Engineering Conference*.

Mossop, A. (2012). "An explanation for anomalous time dependent subsidence". In: *46th US Rock Mechanics/Geomechanics Symposium, Chicago, USA*.

Muntendam-Bos, A. (2020). "Clustering characteristics of gas-extraction induced seismicity in the Groningen gas field". In: *Geophysical Journal International* under review.

Muntendam-Bos, A., I. Kroon, and P. Fokker (2008). "Time-dependent inversion of surface subsidence due to dynamic reservoir compaction". In: *Mathematical Geosciences* 40, pp. 159–177.

Musson, R.M.W. (1999). "Determination of design earthquakes in seismic hazard analysis through Monte Carlo simulation". In: *Journal of Earthquake Engineering* 3.4, pp. 463–474.

Musson, R.M.W. (2000). "The use of Monte Carlo simulations for seismic hazard assessment in the UK". In: *Annals of Geophysics* 43.1.

NAM (pers. comm., 2020). Personal Communication.

NAM (pers. comm., 2019). *June 4 meeting at Schiphol airport, Amsterdam, The Netherlands*. Personal Communication.

NAM (2016). *Technical Addendum to the Winningsplan Groningen 2016: Production, subsidence, induced earthquakes and seismic hazard and risk assessment in the Groningen field, Part II Subsidence*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Ogata, Y. (1998). "Space-time point-process models for earthquake occurrences". In: *Annals of the Institute of Statistical Mathematics* 50.2, pp. 379–402.

Ogata, Y. (2011). "Significant improvements of the space-time ETAS model for forecasting of accurate baseline seismicity". In: *Earth Planets Space* 63, pp. 217–229.

Ojala, I.O., B.T. Ngwenya, I.G. Main, and S.C. Elphick (2003). "Correlation of microseismic and chemical properties of brittle deformation in Locharbriggs sandstone". In: *Journal of Geophysical Research: Solid Earth* 108.B5.

Pagani, M., D. Monelli, G. Weatherill, L. Danciu, H. Crowley, V. Silva, P. Henshaw, L. Butler, M. Nastasi, L. Panzeri, et al. (2014). "OpenQuake engine: An open hazard (and risk) software for the global earthquake model". In: *Seismological Research Letters* 85.3, pp. 692–702.

Paleja, R. and S. Bierman (2016). *Measuring changes in earthquake occurrence rates in Groningen.* Tech. rep.

Pijnenburg, R., B. Verberne, S. Hangx, and C. Spiers (2018). "Deformation behavior of sandstones from the seismogenic Groningen gas field: Role of inelastic versus elastic mechanisms". In: *Journal of Geophysical Rsearch: Solid Earth* 123, pp. 5532–5558.

Pruiksma, J., J. Breunese, K. Van Thienen-Visser, and J. de Waal (2015). "Isotach formulation of the Rate Type Compaction Model for sandstone". In: *International Journal of Rock Mechanics & Mining Sciences* 78, pp. 127–132.

Rydelek, P.A. and I.S. Sacks (1989). "Testing the completeness of earthquake catalogues and the hypothesis of self-similarity". In: *Nature* 337, pp. 251–253.

Segall, P. (1989). "Earthquakes triggered by fluid extraction". In: *Geology* 17.10, pp. 942–946.

Segall, P., J.-R. Grasso, and A. Mossop (1994). "Poroelastic stressing and induced seismicity near the Lacq gas field, southwestern France". In: *Journal of Geophysical Research: Solid Earth* 99.B8, pp. 15423–15438.

Segall, P. and S. Lu (2015). "Injection-induced seismicity: poroelastic and earthquake nucleation effects". In: *Journal of Geophysical Research: Solid Earth* 120.7, pp. 5082–5103.

Settari, A. (2002). "Reservoir compaction". In: *Journal of Petroleum Technology* 54.08, pp. 62–69.

SGS Horizon B.V. (2016). *Independent review of Groningen subsurface modelling update for winningsplan 2016 with opinion letter*. Tech. rep. SGS Horizon B.V.

Shapiro, S.A., C. Dinske, and J. Kummerow (2007). "Probability of a given-magnitude earthquake induced by a fluid injection". In: *Geophysical research letters* 34.22.

Shirzaei, M., W.L. Ellsworth, K.F. Tiampo, P.J. González, and M. Manga (2016). "Surface uplift and time-dependent seismic hazard due to fluid injection in eastern Texas". In: *Science* 353.6306, pp. 1416–1419.

Shoonbeek, J. (1976). "Land subsidence as a result of natural gas extraction in the province of Groningen". In: *SPE European Spring Meeting, Amsterdam, Netherlands*.

SodM (pers. comm., 2020). Personal Communication.

Spetzler, J. and B. Dost (2017). "Hypocentre estimation of induced earthquakes in Groningen". In: *Geophysical Journal International* 209.1, pp. 453–465.

van der Corput, J.G. (1935). "Verteilungsfunktionen". In: *Proceedings of the Koninklijke Nederlandse Akademi van Wetenschappen*. Vol. 38, pp. 813–821.

van der Waal, O. and R. van Eijs (2016). *Subsidence inversion on Groningen using leveling data only*. Tech. rep. Nederlandse Aardolie Maatschappij BV.

Van Eck, T., F. Goutbeek, H. Haak, and B. Dost (2006). "Seismic hazard due to small-magnitude, shallow-source, induced earthquakes in The Netherlands". In: *Engineering Geology* 87.1-2, pp. 105–121.

Van Eijs, R. and O. Van der Wal (2017). "Field-wide reservoir compressibility estimation through inversion of subsidence data above the Groningen gas field". In: *Netherlands Journal of Geosciences* 96, s117–s129.

van Oeveren, H., P. Valvatne, L. Geurtsen, and J. Van Elk (2017). "History match of the Groningen field dynamic reservoir model to subsidence data and conventional subsurface data". In: *Netherlands Journal of Geosciences* 96.5, s47–s54.

Van Thienen-Visser, K. and P.A. Fokker (2017). "The future of subsidence modelling: compaction and subsidence due to gas depletion of the Groningen gas field in the Netherlands". In: *Netherlands Journal of Geosciences* 96.5, s105–s116.

Van Thienen-Visser, K., J. Pruiksma, and J. Breunese (2015). "Compaction and subsidence of the Groningen gas field in the Netherlands". In: *Proceedings of the International Association of Hydrological Sciences*.

Van Wees, J.-D., P.A. Fokker, K. Van Thienen-Visser, B.B.T. Wassing, S. Osinga, B. Orlic, S.A. Ghouri, L. Buijze, and M. Pluymaekers (2017). "Geomechanical models for induced seismicity in the Netherlands: inferences from simplified analytical, finite element and rupture model approaches". In: *Netherlands Journal of Geosciences* 96.5, s183–s202.

Visser, C.A. and J.L. Solano Viota (2017). "Introduction to the Groningen static reservoir model". In: *Netherlands Journal of Geosciences* 96.5, s39–s46.

Vlek, C. (2018). "Induced Earthquakes from Long-Term Gas Extraction in Groningen, the Netherlands: Statistical Analysis and Prognosis for Acceptable-Risk Regulation". In: *Risk Analysis*.

Wiemer, S. (2001). “A software package to analyze seismicity: ZMAP”. In: *Seismological Research Letters* 72.3, pp. 373–382.

Wiemer, S. and M. Wyss (2002). “Mapping spatial variability of the frequency-magnitude distribution of earthquakes”. In: *Advances in geophysics*. Vol. 45. Elsevier, pp. 259–V.

Willacy, C., E. Van Dedem, S. Minisini, J. Li, J.W. Blokland, I. Das, and A. Droujinine (2018). “Application of full-waveform event location and moment-tensor inversion for Groningen induced seismicity”. In: *The Leading Edge* 37.2, pp. 92–99.

Yerkes, R. and R. Castle (1969). “Surface deformation associated with oil and gas operations in the United States”. In: *1st International Symposium on Land Subsidence*.

Zaliapin, I.V., Y.Y. Kagan, and F.P. Schoenberg (2005). “Approximating the distribution of Pareto sums”. In: *Pure and Applied geophysics* 162.6-7, pp. 1187–1228.

Zoback, M.D. and J.C. Zinke (2002). “Production-induced normal faulting in the Valhall and Ekofisk oil fields”. In: *The mechanism of induced seismicity*. Springer, pp. 403–420.

Report Number: 19-007	Confidential: X Unlimited:	External: X Internal:	NORSAR Project No.: 10178
Title:	Review of seismogenic source models for the Groningen gas reservoir (WP1: Review of the existing NAM seismogenic source model)		
Client:	Staatstoezicht op de Mijnen, Netherlands		
Project manager:	D. Kühn		
Authors/prepared by (alphabetical order):	T. Dahm (GFZ), S. Hainzl (GFZ), D. Kühn (NORSAR), V. Oye (NORSAR), G. Richter (GFZ), I. Vera Rodriguez (NORSAR)		
Submitted to:	Staatstoezicht op de Mijnen, Netherlands		
Contract reference:			
Archive reference:			
Approved by: Department head: CEO:	Name: V. Oye A. Strømmen Lycke	Signature: Volker Oye	Date: 24.04.2020

NORSAR

info@norsar.no
www.norsar.no

GFZ
Helmholtz-Zentrum
POTSDAM