AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2017GL073878

Key Points:
- We model prior induced earthquakes in the Groningen reservoir and forecast future seismicity under different production scenarios
- Parameters estimated by model calibration to earthquake data indicate that the reservoir was not critically stressed prior to gas production
- A model forecast for February 2017 to 2024 suggests that the largest induced event has a 5% likelihood of exceeding M 4

Supporting Information:
- Supporting Information S1
- Text S2
- Data Set S1
- Data Set S2
- Data Set S3

Correspondence to:
D. Dempsey,
d.dempsey@auckland.ac.nz

Citation:
Dempsey, D., and J. Suckale (2017), Physics-based forecasting of induced seismicity at Groningen gas field, the Netherlands, *Geophys. Res. Lett.*, *44*, doi:10.1002/2017GL073878.

Received 25 APR 2017
Accepted 20 JUL 2017
Accepted article online 25 JUL 2017

Physics-based forecasting of induced seismicity at Groningen gas field, the Netherlands

David Dempsey[1] and Jenny Suckale[2]

[1]Department of Engineering Science, University of Auckland, Auckland, New Zealand, [2]Department of Geophysics, Stanford University, Stanford, California, USA

Abstract Earthquakes induced by natural gas extraction from the Groningen reservoir, the Netherlands, put local communities at risk. Responsible operation of a reservoir whose gas reserves are of strategic importance to the country requires understanding of the link between extraction and earthquakes. We synthesize observations and a model for Groningen seismicity to produce forecasts for felt seismicity ($M > 2.5$) in the period February 2017 to 2024. Our model accounts for poroelastic earthquake triggering and rupture on the 325 largest reservoir faults, using an ensemble approach to model unknown heterogeneity and replicate earthquake statistics. We calculate probability distributions for key model parameters using a Bayesian method that incorporates the earthquake observations with a nonhomogeneous Poisson process. Our analysis indicates that the Groningen reservoir was not critically stressed prior to the start of production. Epistemic uncertainty and aleatoric uncertainty are incorporated into forecasts for three different future extraction scenarios. The largest expected earthquake was similar for all scenarios, with a 5% likelihood of exceeding M 4.0.

Plain Language Summary Earthquakes have been triggered by natural gas extraction at the Groningen reservoir in the Netherlands since 1991 with the largest, a M 3.6, occurring in 2012. These earthquakes pose a risk to local communities and raise questions about future management of a resource that is strategically important to the country. In this study, we describe a computer model for the earthquakes at Groningen and the projections this model makes for future seismicity. We adapt new techniques to train our model on data and to quantify where uncertainty arises. Then, embracing the inherent unpredictability of earthquakes, we calculate the future earthquake rate for three different scenarios of gas extraction. Our model suggests that the largest earthquake in the period beginning February 2017 to the end of 2024 will be in the range M 3.0 to 4.0, with a 5% likelihood to exceed 4.0. However, it is not possible to predict when or where any particular event will occur.

1. Introduction

Earthquakes induced by human activities have been known for decades [e.g., *Healy et al.*, 1968] but have become more frequent in recent years [e.g., *Ellsworth*, 2013; *Zang et al.*, 2014]. Events large enough to be felt are at a minimum a nuisance to the local population and, at larger magnitude, a hazard to public safety. Model-based forecasting can assist in managing the risk posed by induced seismicity through either responsible operation or project abandonment.

In seismology, a prediction carries a different connotation to a forecast. We adopt the domain definition of earthquake prediction as a high-confidence statement regarding the location, timing, and size of a future event [*Marzocchi and Zechar*, 2011]. This is different than a forecast, which is a probabilistic statement about future earthquake outcomes. This study concerns itself with the latter: forecasting induced seismicity.

Forecasting is routinely undertaken for natural earthquakes. For example, the return time of characteristic events on a particular fault can be estimated on the basis of paleoearthquake observations [e.g., *Ogata*, 1999]. Or, statistical clustering models incorporating empirical seismicity relations (Omori and Gutenberg-Richter laws) and calibrated using post main shock seismicity can be extrapolated to forecast diminishing aftershock rates [e.g., *Gerstenberger et al.*, 2005]. In both cases, the forecast is (i) conditioned on data and (ii) inherits a probabilistic character from the (apparently) random qualities of earthquake triggering.

We take a similar approach to forecast induced seismicity: first, we calibrate on prior earthquake data and reservoir operations, which enables us to learn about the model parameters, in this case mainly the

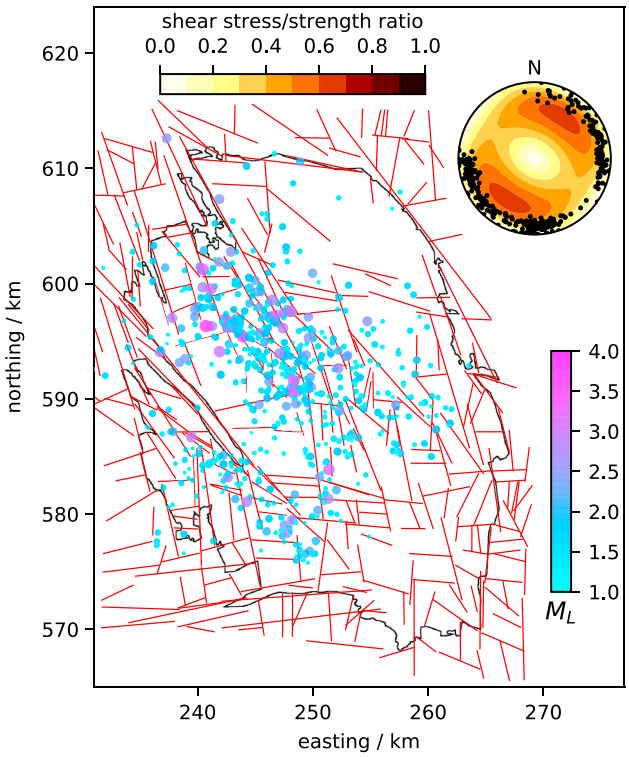

Figure 1. Groningen reservoir (black) with modeled fault traces (red) and earthquakes (size and color indicate magnitude). Inset shows best fit initial stress state with black points indicating fault orientations.

tectonic stress. Then, we extrapolate for future operating scenarios, incorporating epistemic (model) and aleatoric (random process) uncertainty. We apply this approach to forecast the induced seismicity at Groningen natural gas field, the Netherlands, for three potential extraction scenarios provided to us by the operator of the field. For this case, the data are reasonably complete and findings from the study may assist decision-makers and the local community in understanding the nature of the hazard and the steps that can be taken to mitigate it.

2. Induced Seismicity at Groningen

Groningen reservoir, located in the northeast of the Netherlands, is western Europe's largest gas field and particularly significant to the Dutch economy. Production of gas at Groningen began in 1963, and the first earthquake induced by gas extraction was recorded in 1991 [*van Eck et al.*, 2006]. The main gas-bearing interval is ~3 km deep and overlain by the Zechstein salt. Numerous faults have been mapped in the reservoir (Figure 1), which have previously been activated in extension [*Bourne et al.*, 2014].

In August 2012, an M 3.6 event occurred near Huizinge, which created public concern and catalyzed a renewed effort to understand the induced seismicity. Forecasting ongoing seismicity—its rate and the likelihood of large events—helps inform decision-making regarding future extraction rates at the field.

Earthquake observations are an essential input for model-based forecasting, allowing us to train models on data and thus learn the values of otherwise ill-constrained parameters. At Groningen, these data are sourced from a catalog maintained by the Royal Netherlands Meteorological Institute (KNMI), the seismology agency in the Netherlands, which has operated a dedicated borehole network at Groningen since 1995 [*Dost et al.*, 2012]. Formally, we define the catalog to comprise $\hat{N}$ events, each with a time, $\hat{t}$, location, $\hat{\boldsymbol{x}}$, and local magnitude, $\hat{M}$. Between 1991 and February 2017, ~600 events were recorded larger than the $M > 1$ detection threshold reported by *Dost et al.* [2012]. Seismicity fluctuated about 10 events per year until about 2003 when it started to increase (69 events were recorded in 2013). In recent years, government-mandated reductions of the extraction rate have been associated with reduced seismicity.

Prior studies have used models to understand the Groningen seismicity. *Bourne et al.* [2014] presented a model for reservoir compaction correlated with field seismicity and used this to forecast, for the 2013 gas production plan, a 10% likelihood (P_{10}) that the largest event in the period 2013–2023 would exceed M_w 4.6. In 2016, a workshop was convened at which models of future Groningen seismicity were presented in the context of a wider hazard assessment [*Nederlandse Aardolie Maatschappij (NAM)*, 2016a; *Bommer and van Elk*, 2017]. One of these studies [*Zöller and Holschneider*, 2016] used a nonhomogeneous Poisson point process with rate parameter, $\lambda(t)$, proportional to the gas extraction rate: forecasts for scenarios with mandated reductions from the 2013 plan yielded P_{10} estimates for the largest event in the period 2016–2024 between M_L 3.9 and 4.3. Other models have addressed specific geomechanical [*Sanz et al.*, 2015] and ground motion aspects [*Bommer et al.*, 2016] of the seismicity. Our model extends previous work to incorporate the physical processes

linking reservoir operation to earthquake triggering and rupture, which improves our understanding of the reservoir stress state and future seismicity.

2.1. Seismogenic Processes in the Groningen Reservoir

Since 1963, more than 2000 billion cubic meters (bcm) of gas have been extracted at Groningen, with a further 800 bcm estimated to remain [*Bourne et al.*, 2014]. The resulting pressure drop, Δp, averages about 22 MPa in the reservoir, which, from a Coulomb failure perspective, should stabilize faults by increasing the effective normal stress on them. Thus, observed seismicity must be attributed to a different mechanism, which we assume is poroelastic contraction and associated horizontal stress reduction, $\Delta\sigma_h$, induced in the reservoir [*Segall and Fitzgerald*, 1998]. As a result, the resolved shear stress on faults is increased and, if this exceeds the concurrent strength increase, the fault is destabilized.

We use a modification of the *Segall and Fitzgerald* [1998] model, which introduces the stress path coefficient, A [*Zoback*, 2010]

$$A = \alpha \frac{1-2\nu}{1-\nu}, \tag{1}$$

where ν is Poisson's ratio and α is the Biot coefficient. Although the stress path coefficient is formulated in terms of the measurable material parameters ν and α, in our study, its value is determined by model calibration to data. Thus, values for ν and α are not required as model inputs nor are their values independently determined by calibration.

Note that whereas *Segall and Fitzgerald* [1998] assume homogeneous pressure decline throughout an elliptical reservoir, using a stress path allows us to express the horizontal stress changes in terms of just the local pressure change, i.e., $\Delta\sigma_h(\mathbf{x}, t) = A\Delta p(\mathbf{x}, t)$. Hence, it is better suited to modeling the Groningen reservoir where pressure decline is heterogeneous. We note that this approximate model neglects nonlocal stress effects (such as can induce earthquakes outside a reservoir [*Segall et al.*, 1994]) that could only be captured by a relatively slow 3-D numerical model. However, as most of the earthquakes at Groningen occur within the lateral extent of the reservoir (Figure 1), this first-order description of reservoir poromechanics will suffice. We also assume that changes in the vertical stress are negligible, which is reasonable for reservoirs whose lateral extent greatly exceeds their thickness. At Groningen, the reservoir aspect ratio is of the order 10^2 [*Bourne et al.*, 2014], well above the limiting value given by *Segall and Fitzgerald* [1998] to ignore vertical stress changes (~5).

2.2. Tectonic Stress

Although stress directions and magnitudes at Groningen are not well constrained, in accordance with previous studies [e.g., *Orlic and Wassing*, 2013] and published focal mechanisms [*Kraaijpoel and Dost*, 2013], we assume that the tectonic setting is extensional. Given the uncertainty, it is useful to parameterize a general extensional stress state:

1. Vertical stress, σ_1, is given by integration of the overburden (~74 MPa at 3km depth, assuming an average overburden density 2500kgm^{-3}).
2. The minimum horizontal stress, σ_3, is bounded from above by isotropy (σ_1) and from below by the strength of a critically stressed crust [*Jaeger et al.*, 2009] in frictional equilibrium with reservoir pressure

$$\sigma_{\text{crit}} = (\sigma_1 - p_0)\left[f_s + \sqrt{1+f_s^2}\right]^{-2} + p_0, \tag{2}$$

   where f_s is the static friction coefficient and p_0 is ambient reservoir fluid pressure prior to production (~34 MPa in 1958). This stress component is expressed on the normalized interval [0,1] by the parameter, $h = (\sigma_3 - \sigma_{\text{crit}})/(\sigma_1 - \sigma_{\text{crit}})$; i.e., $h = 0$ indicates a critical stress state, and $h = 1$ indicates isotropy.
   Note that if it is assumed that the maximum shear stress on an optimally oriented fault in a critically stressed crust is within one earthquake stress drop from failure, then we can define a critically stressed interval for h. For an average Groningen stress drop of ~1 MPa, and Byerlee friction ($f_s \sim 0.6$), this interval is $h_{\text{crit}} = [0, 0.06]$.
3. The maximum horizontal stress, σ_2, lies between σ_1 and σ_3. Its magnitude is expressed on the interval [0,1] by the parameter $H = (\sigma_2 - \sigma_3)/(\sigma_1 - \sigma_3)$. As H tends to 0, faults not well oriented with regard to σ_3 become more likely to fail.

Finally, the angle through which σ_3 is rotated clockwise from north is denoted ϕ. This parameter determines which faults in the Groningen reservoir are the first to be activated.

3. Induced Seismicity Models

We develop a hierarchy of models to describe seismicity at Groningen. The first model uses stress and strength changes on faults to approximate the seismicity rate for the entire field. The second model builds on this to output synthetic earthquake catalogs.

We restrict our analysis to the 325 largest reservoir faults (length >2 km) at Groningen, which is a trade-off that limits computational expense while still representing the structures capable of generating the largest earthquakes. Our modeling approach approximates a fault as a 1-D line segment embedded in a 2-D flattened reservoir (as opposed to a flat or rough 2-D surface in a 3-D rock volume). We chose to orient the fault segments in the plane of the reservoir so as to capture both the dominant extent of large ruptures confined to the reservoir and the spatial heterogeneity in pressure drawdown (Figure 1). An initial shear, τ, and normal stress, σ_n, are resolved on each fault based on its orientation within the regional stress field. Initial fault strength is then computed from the effective normal stress, $\tau_s = f_s(\sigma_n - p)$. The most likely tectonic stress state at Groningen, determined during model calibration, is shown in Figure 1.

Spatiotemporally varying pressure, strength, and stress changes are resolved on each of the modeled faults. Spatiotemporal pressure changes throughout the Groningen gas field and surrounding area are derived from a two-component (water and gas) reservoir simulation maintained by the field operators (see supporting information Data Sets S1–S3). The model incorporates historic extraction and injection from all wells in the field and is calibrated using pressure and subsidence data [*NAM*, 2016b]. Three hypothetical scenarios of future extraction were simulated—21, 27, and 33 bcm/yr—that, respectively, correspond to larger modeled pressure drawdown within the field. Shear and normal stress changes are computed from the pressure field using the stress path coefficient.

Static stress transfer, which can be important for aftershock triggering, is not captured by our model. However, as there is not much evidence for strong clustering in the Groningen catalog, this simplification should have only a minimal impact on our study. Dynamic wave propagation effects are also neglected.

3.1. Total Field Seismicity Rate

This model considers induced seismicity in the context of fault criticality. Modeled faults are discretized into patches of equal length, δx, and associated an initial stress and strength, τ and τ_s. We define patches as either stable, where $\tau < \tau_s$, or at the limit of frictional failure (critical), in which case $\tau = \tau_s$. We do not permit $\tau > \tau_s$.

Imposing an arbitrary stress, $\delta\tau$, or strength change, $\delta\tau_s$, on a patch has one of two effects. If the patch is not critical, it becomes either closer to or farther from failure, depending on the sign of $\delta\tau$ or $\delta\tau_s$. If it is critical, an incremental change that would otherwise cause $\tau > \tau_s$ instead generates a proportional amount of seismicity, $\delta N = k[\delta\tau - \delta\tau_s]\delta x$, where k is an empirical constant with unknown value. This formulation recovers the steady state limit of the *Dieterich* [1994] seismicity rate model, in which stressing and seismicity rates are proportional. Furthermore, this proportionality essentially captures the ability of fault patches to slip multiple times.

Thus, a fault of length, L, subject to spatiotemporal changes in stress and strength, $\Delta\tau(x,t)$ and $\Delta\tau_s(x,t)$, generates N earthquakes during the period $[0, T]$

$$N = k\int_0^L\int_0^T \left[\Delta\tau(x,t) - \Delta\tau_s(x,t)\right]\delta\left(\tau(x,t) - \tau_s(x,t)\right)\mathrm{d}x\mathrm{d}t, \tag{3}$$

where $\delta(x)$ is the Dirac delta function. The spatiotemporal stress and strength evolution are given

$$\tau(x,t) = \tau(x,0) + \int_0^t \Delta\tau(x,t')u\left(\frac{\tau_s(x,t')}{\tau(x,t')} - 1\right)\mathrm{d}t', \tag{4}$$

$$\tau_s(x,t) = \tau_s(x,0) + \int_0^t \Delta\tau_s(x,t')\mathrm{d}t' \tag{5}$$

where $u(x)$ is the Heaviside step function. The seismicity rate for the ith fault is $\lambda_i(t) = \mathrm{d}N/\mathrm{d}t$

$$\lambda_i(t) = k\int_0^L \left[\dot{\tau} - \dot{\tau}_s\right]\delta(\tau - \tau_s)\mathrm{d}x, \quad \dot{\tau} = \frac{\partial\Delta\tau}{\partial t}u\left(\frac{\tau_s}{\tau} - 1\right), \quad \dot{\tau}_s = \frac{\partial\Delta\tau_s}{\partial t}. \tag{6}$$

The proportionality constant, k (equation (3)), is a parameter assumed to incorporate unknown relationships between stressing and earthquake triggering, such as seismogenic depth of the fault, elastic properties, and degree of aseismic creep. We estimate its value by requiring that the seismicity rate integrated up to the present day matches the total number of observed earthquakes. We assume that k is the same for all faults and constant over time, which is equivalent to assuming that the underlying parameters are homogeneous across the field and constant in time.

We approximate the total seismicity rate at Groningen as the sum of rates on all faults, $\lambda = \Sigma \lambda_i$. This model for the seismicity rate depends on the parameters defining the tectonic stress state as well as A because these control both the initial stress and the stress changes on the fault. Aggregating the behavior of all faults ignores both spatial variability in $\lambda(t)$ and the parameters defining it. In the future, it would be worthwhile relaxing this restriction to explore the possibility of parameter heterogeneity.

3.2. Synthetic Earthquakes

We use a physics-based model to capture magnitude frequency properties of the Groningen seismicity. In contrast to the deterministic seismicity rate, $\lambda(t)$, returned by the model in section 3.1, this second model constructs catalogs of synthetic earthquakes—magnitudes and times—that are inherently random. It does this by using an ensemble approach to represent uncertainty about the precise state of stress, which we assume to be spatially heterogeneous on a particular fault. For each fault, we construct an ensemble of stress realizations, each one a random draw from a statistical model that accounts for spatiotemporally evolving stress and strength on the fault.

On each realization, we apply a slip-weakening instability criterion [*Uenishi and Rice*, 2003] to determine where and when earthquakes are induced as a result of the resolved pressure and stress changes. Then, we use 1-D crack dynamic equations to simulate rupture propagation and arrest, thus yielding the size of the rupture (which can be less than the fault length) and moment magnitude of the event [*Dempsey and Suckale*, 2016]. We set the rupture strength drop to about ~10%, which results in modeled stress drops of about 0.5–1.0 MPa, consistent with observations from the field [*Bommer et al.*, 2016]. Shear modulus is set to 37.5 GPa. Although our model outputs a 1-D rupture length, L, we require the moment of an equivalent circular source to approximate the event magnitude. Assuming that the source radius is $L/2$, then for stress drop, $\Delta\tau$, the moment magnitude is given [*Kanamori and Anderson*, 1975; *Hanks and Kanamori*, 1979]

$$M_w = \frac{2}{3}\log_{10}\left(\frac{16}{7}\Delta\tau\frac{L^3}{8}\right) - 6. \tag{7}$$

While this magnitude definition is not wholly satisfactory given the 1-D nature of our model, it nevertheless generates event magnitudes that exhibit a power law scaling very similar to the Groningen events (Figure 2f). For further discussion of source scaling and stress drop independence in our model, see *Dempsey et al.* [2016]. Faults are discretized to $\delta x \sim 10$ m, which is sufficient to resolve earthquakes down to M 1.

Although this second model outputs earthquake catalogs, we emphasize that it does not attempt to capture specific historic earthquakes at Groningen or those that could occur in the future. Rather, we are replicating the collective behavior of all events, which in turn reflects the underlying physical drivers, such as elevated seismicity during periods of rapid pressure or stress change. In addition, the ensemble approach ensures that the synthetic earthquakes replicate some empirical features of real seismicity, such as an exponential interevent time distribution and the Gutenberg-Richter magnitude frequency distribution (MFD) [*Dempsey et al.*, 2016] (see also supporting information). The MFD shape is sensitive to the parameterization of stress heterogeneity on faults, and calibration of this distribution to replicate Gutenberg-Richter power law scaling is described in *Dempsey et al.* [2016].

Performing computations on 1-D faults allows us to simulate tens of thousands of events in minutes on just a single core, which permits a robust exploration of model uncertainty. However, this also limits our ability to model complex ruptures, such as those that propagate into the basement below the reservoir or which jump from one fault segment to another. As a result, the largest earthquake we can model is M_{max} 6.4, corresponding to complete rupture of the longest structure. In a later section, we show that earthquakes forecast by our model are much smaller than this limit and, as such, it is not obvious that increasing M_{max} would affect our results. Nevertheless, complex ruptures at Groningen are important to study, and other codes are suited to the task [e.g., *Richards-Dinger and Dieterich*, 2012]. However, all models face the same difficulties characterizing

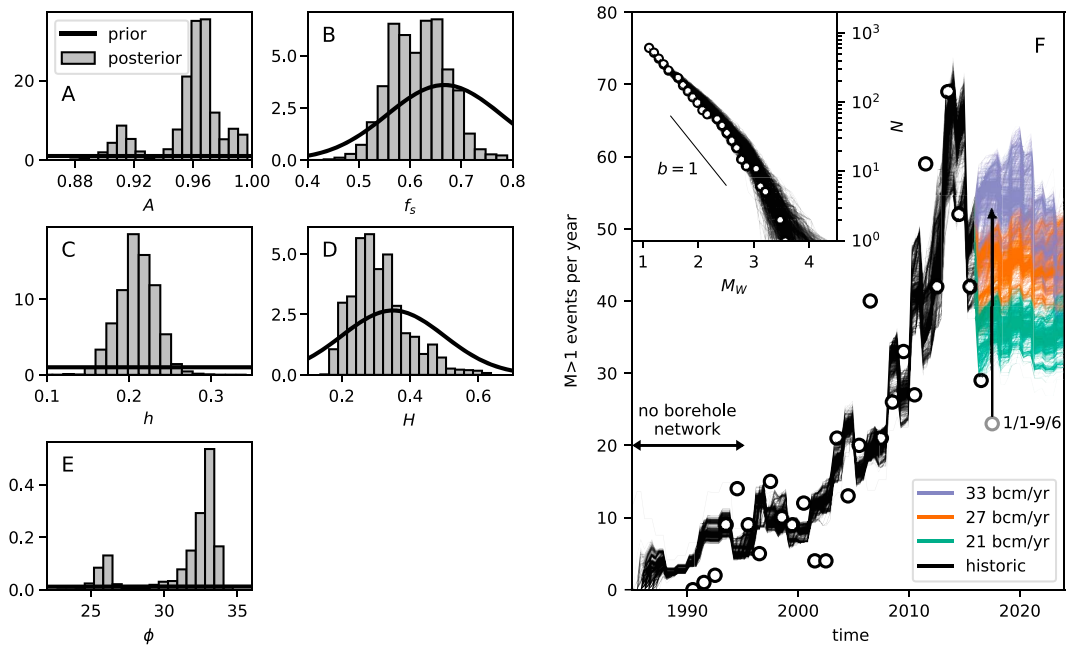

Figure 2. Calibration of the Groningen seismicity model. Parameter priors and posteriors for (a) stress path coefficient, A; (b) friction coefficient, f_s; (c) minimum horizontal stress parameter, h; (d) maximum horizontal stress parameter, H; and (e) angle between σ_3 and north, ϕ. Priors are A: $\mathcal{U}(0.6, 1)$; f_s: $\mathcal{N}(0.67, 0.11)$; h: $\mathcal{U}(0, 0.5)$; H: $\mathcal{N}(0.35, 0.15)$; and ϕ: $\mathcal{N}(30, 30)$, where $\mathcal{U}(a, b)$ is a uniform distribution between a and b and $\mathcal{N}(\mu, \mathrm{SD})$ is a normal distribution with mean, μ, and standard deviation, SD. Posteriors in Figures 2a–2e are computed by MCMC using 30 walkers (reasonable for five parameters), a burn-in of 100 jumps, and 2600 jumps per walker, stopping the exploration when distributions had obtained a reasonably stable shape. (f) Sample of 1000 seismicity rate models drawn from $P(\boldsymbol{\theta})$ and extrapolated for each of the three future scenarios. The historic annual seismicity rate at Groningen is shown by the open circles. The grey marker indicates the events to date in 2017 (from 1/1 to 9/6), and the arrow shows the extrapolated event count for 2017. Inset: MFD, measured (open circles) and modeled (black), shows good agreement.

system heterogeneity and capturing uncertainty. Incorporating more physics and longer runtimes may preclude the type of uncertainty analysis conducted here.

4. Calibration

Through calibration, we seek parameter values for the seismicity rate model: $\lambda(\boldsymbol{\theta}, t)$ where $\boldsymbol{\theta} = [A, f_s, h, H, \phi]$ is the parameter vector. However, the presence of uncertainty prevents unique identification of $\boldsymbol{\theta}$ and, instead, the best we can do is identify a probability distribution of the likely parameter values, called the posterior, $P(\boldsymbol{\theta})$.

To calibrate $\lambda(\boldsymbol{\theta}, t)$, we must decide which observations to compare against the model. An approximate seismicity rate can be calculated from the observations, $\hat{\lambda}(t) \approx \hat{N}(t, \mathrm{d}t)/\mathrm{d}t$, where $\hat{N}(t, \mathrm{d}t)$ is the number of events detected between $[t, t + \mathrm{d}t]$. However, there are two problems defining the observed seismicity rate this way:

1. The interval $\mathrm{d}t$ is arbitrary, i.e., as $\mathrm{d}t \to 0$, then $\hat{\lambda}$ becomes bimodal, either equal to 0 (no events in $\mathrm{d}t$) or $1/\mathrm{d}t$. As $\mathrm{d}t \to \infty$, time variability is lost, $\hat{\lambda}(t) = \hat{\lambda}_0 = \hat{N}/T_{\mathrm{obs}}$, with T_{obs} the observation interval.
2. Inherent randomness in earthquake triggering can be misleading. For example, a Poisson triggering process with constant rate parameter $\lambda = 10.5$ events per year could yield event counts of 13, 11, and 9 in consecutive years. It might be tempting to infer temporal changes in the earthquake rate that do not exist.

We avoid these problems by using a log likelihood expression that embeds $\lambda(t)$ as the rate parameter of a nonhomogeneous Poisson process. The expression quantifies the relative likelihood that $\lambda(t)$ could have generated the earthquake event time observations $\hat{\mathbf{t}} = [t_1, \cdots t_n]$ [*Lindqvist and Taraldsen*, 2013]

$$\mathrm{LLK}(\boldsymbol{\theta}) = \sum_{j=1}^{n} \log \lambda(\boldsymbol{\theta}; t_j) - \int_0^{t_n} \lambda(\boldsymbol{\theta}; t')\mathrm{d}t', \tag{8}$$

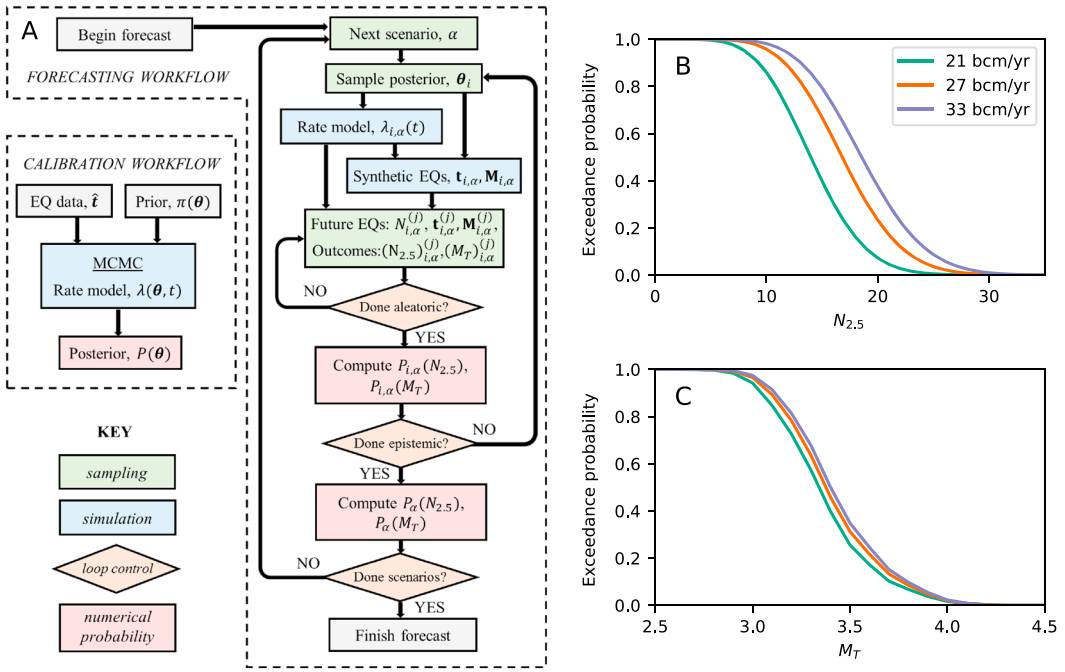

Figure 3. Forecasts from the Groningen seismicity model. (a) Model calibration and forecast workflows. See text for symbol definitions. Forecasts for (b) $N_{2.5}$, the number of felt events ($M > 2.5$), and (c) M_T, the largest expected event in the interval February 2017 to end 2024 for three future scenarios.

where θ are the parameters on which λ depends. As LLK is a probabilistic expression, it explicitly allows for aleatoric uncertainty in the earthquake observations (point 2 above) and thus permits that multiple rate functions, $\lambda(t)$, could be correct.

For the Groningen model, calibration yields a posterior distribution, $P(\theta)$, that constructs the best fitting rate models for the ~600 earthquakes recorded between 1991 and February 2017 as well as the absence of events prior to this period. We determine $P(\theta)$ using Markov chain Monte Carlo (MCMC) [*Foreman-Mackey et al.*, 2013], a standard method for exploring model parameter space. This is a Bayesian method, and we therefore need to define a prior distribution, $\pi(\theta)$, that reflects our preanalysis understanding of the parameters. Here we have used uniform priors for A and h (no knowledge) and normal distributions for H, f_s, and ϕ (Figures 2a–2e).

One limitation of our analysis is that it does not address uncertainty inherent in the reservoir model and the pressure changes it predicts. These are essentially accepted as "truth" when, in fact, although the reservoir model is calibrated against field data to ensure its consistency with observation, there will always be some error in its prediction of the pressure. The principle outcome for our study is that parameter distributions estimated during calibration of the seismicity model will be different to those estimated if we had access to the true pressure evolution. A secondary limitation is that we are unable to link prior or forecasted seismicity to the reservoir operation at the level of the individual well. Accordingly, our discussion focuses on bulk extraction rates from Groningen.

5. Forecasting

In this phase, we project the parameterized model beyond the calibration period to the end of 2024. Epistemic uncertainty quantified by $P(\theta)$ is propagated through sampling and construction of a family of models, each characterized by different parameters and thus a different likelihood of being correct, LLK(θ). The forecast is the set of predictions from all models in the family. Instead of presenting a single, most likely outcome, we present a numerically approximated probability distribution.

We forecast two quantities for the period February 2017 to 2024: (1) the number of felt events ($M > 2.5$), denoted as $N_{2.5}$, and (2) the largest event during this period, M_T. To be clear, M_T is the largest expected event and is different from the largest possible event, M_{max} [*Zöller and Holschneider*, 2016]. In our model, M_{max} is limited by the size of modeled faults to ~6.4 and an event this large did not occur in over 10^6 simulated earth-

quakes. Nevertheless, as discussed in section 3.2, there are limitations in our model that preclude modeling even larger events, such as those that propagate into basement or that jump between multiple fault segments.

To capture epistemic and aleatoric uncertainty within a forecast, we use the following workflow (see also Figure 3a):

1. *Choose hypothetical future scenario*. In this case, the reservoir pressure changes that drive seismicity correspond to an extraction rate, α, of 21, 27, or 33 bcm/yr.
2. *Predict seismicity rate*. Sample a parameter set, $\boldsymbol{\theta}_i$, from $P(\boldsymbol{\theta})$, and from this construct the model for field seismicity rate, $\lambda_{i,\alpha}$, in the forecast period (Figure 3).
3. *Model earthquakes*. Use $\boldsymbol{\theta}_i$ and $\lambda_{i,\alpha}$ to model a large pool of "potential" earthquakes for the forecast period, $\sim 10^5$ events with times, $\mathbf{t}_{i,\alpha}$, and magnitudes, $\mathbf{M}_{i,\alpha}$.
4. *Sample the earthquake pool*. Sample the nonhomogeneous Poisson process with rate, $\lambda_{i,\alpha}$, generating a count of $N_{i,\alpha}^{(j)}$ earthquakes in the forecast period. Sample $N_{i,\alpha}^{(j)}$ events from the pool of potential future earthquakes, yielding one possible realization of future seismicity at Groningen, $\left[\mathbf{t}_{i,\alpha}^{(j)}, \mathbf{M}_{i,\alpha}^{(j)}\right]$. From this catalog, compute the forecast outcomes, $N_{2.5}$ and M_T.
5. *Incorporate aleatoric uncertainty*. Repeat step 4 to approximate probability distributions for seismicity outcomes, $P_{i,\alpha}(N_{2.5})$ and $P_{i,\alpha}(M_T)$, specific to parameter, $\boldsymbol{\theta}_i$, and scenario α. We found that 300 samples produced stable distributions.
6. *Incorporate epistemic uncertainty*. Repeat steps 2–5 to approximate probability distributions for seismicity outcomes, $P_{\alpha}(N_{2.5})$ and $P_{\alpha}(M_T)$, specific to scenario α (Figures 3b and 3c). For this step, about 300 samples produced stable distributions.

6. Discussion

Model calibration enables us to learn about the likely properties of the Groningen reservoir. This is analogous to inverse modeling, wherein the model acts as a lens through which to interpret the (earthquake) observations. For instance, it appears from Figure 2d that σ_2 is likely nearer to σ_3 than it is to σ_1, although the broad nature of the posterior indicates some uncertainty. The posterior for ϕ is bimodal (Figure 2e) and indicates an extension direction approximately 30° from north.

The posterior distribution for h (Figure 2c) suggests that prior to extraction, the reservoir was not critically stressed ($h > 0.06$). If it had been, then earthquakes should have been induced earlier than 1991. From our modeling, the most likely initial stress state is $\sigma_3 = 51.3$ MPa and $\sigma_2 = 57.1$ MPa, which corresponds to $f_s = 0.67$, $\sigma_1 = 74$ MPa, and $p_0 = 34$ MPa.

We note that our model provides no indication about whether the basement below the reservoir is critically stressed. The stress environment at Groningen is complicated by the presence of a plastically deforming salt body above the reservoir [*Sanz et al.*, 2015], which could be causing relaxation of the stresses below. Whether the basement is critically stressed or not has important implications for the maximum size of earthquakes induced in the reservoir.

During calibration, a stress path coefficient is preferred (Figure 2a) that is higher than the typically reported range (0.5–0.8 [*Chan and Zoback*, 2002]). From a model perspective, this occurs because the steeply dipping reservoir faults (mean dip ~78°) require quite large stress changes to reactivate in extension. It seems unlikely that A has such a large value at Groningen, as this would require an unreasonably low Poisson's ratio. More likely, this indicates a deficiency of the model physics, for instance, our ignoring nonlocal stresses or parameter heterogeneity or possibly inaccuracies in the reservoir model. Thus, to attain a good match with the observations, A assumes an unrealistically large value: an instance of parameter surrogacy [*Doherty and Christensen*, 2011]. While we do not expect that this will strongly impact our model forecasts—after all, the missing physics are being represented, after a fashion—this nevertheless points to further lines of investigation in future work.

The estimated friction coefficient is consistent with the range of Byerlee friction [*Byerlee*, 1978].

In Figure 2f, we plot 1000 seismicity rate models drawn from the posterior and compare these to the annual seismicity rate (setting $dt = 1$ year is admittedly arbitrary). There is some variability amongst the models, but they are generally consistent with each other. The model predicts the onset of seismicity at Groningen slightly earlier than the first recorded event in 1991. The trend suggests that there has been a modest rate

increase since onset that then accelerated in ~2005 and started to decrease from ~2014. However, seismicity to date (9 June) in 2017 is on track to exceed 2016. Whether this indicates a stabilization of the seismicity rate (as suggested by our model) or the beginning of a sustained increase will require more observations to determine.

Our model does not replicate short-term departures from long-term trends. For instance, during 2006 and 2011, recorded seismicity exceeded the modeled rate ($N_{2006} = 40$, $\lambda_{2006} \approx 20$ year^{-1} and $N_{2011} = 59$, $\lambda_{2011} \approx 40$ year^{-1}) to a degree not readily explained by Poisson uncertainty (the results, respectively, have a <0.01% and 0.2% chance of occurrence under a correct model). One possibility is that these spikes are due to a nonlocal poroelastic effect caused by rapid changes in the extraction rate, similar to the mechanism suggested by *Segall and Lu* [2015] for injection.

Increased extraction rates induce more earthquakes, but not much larger ones. Of the three scenarios, extraction at 33 bcm causes the largest stress perturbation to the Groningen reservoir and hence our forecast indicates that this should lead to the highest seismicity rate (Figure 2f) and the greatest number of felt events (Figure 3b). However, the three scenarios are nearly indistinguishable in their forecast of M_T, which has a 5% likelihood of exceeding 4.0. It seems counterintuitive that comparing the 21 and 33 bcm/yr scenarios (Figure 2f), an ~50% increase in the seismicity rate over almost 8 years barely impacts the largest likely event, M_T. This insensitivity occurs partly because shifts in the distribution $P_\alpha(M_T)$ are obscured by sources of uncertainty. The nature of power law MFDs also contributes: assuming a fixed b value, an ~50% increase of catalog size from 21 to 33 bcm barely affects the ordinate (which is given on a log scale) and abscissa intercepts (the largest event; see also supporting information).

Zöller and Holschneider [2016] constructed a forecast of Groningen seismicity for the same time period using a nonhomogeneous Poisson process with $\lambda(t)$ proportional to the gas production rate. They report a P_{10} exceedance for M_T for the three scenarios 21, 27, and 33 bcm/yr of 4.14, 4.26, and 4.31, respectively (using their results for an untruncated MFD). The equivalent results from our forecast for the three scenarios are 3.82, 3.87, and 3.89, which are about a half a magnitude unit smaller. The discrepancy could be attributed either to different triggering mechanisms of the two models or the different calibration and forecast periods: respectively, 1991–2015 and 2016–2024 for *Zöller and Holschneider* [2016], compared with 1991–2017.2 and 2017.2–2024 for our model.

7. Conclusion

Induced earthquakes at the Groningen natural gas reservoir may put local communities at risk and raise challenging questions about how to balance public safety against economic concerns. Physics-based models provide a framework for understanding the relationship between induced earthquakes, the tectonic and material conditions particular to a site, and the manner in which a reservoir is operated. We have developed such a model for production-induced seismicity at the Groningen natural gas field.

We used Bayesian calibration to identify distributions of stress and material parameters from the earthquake data. Calibration is probabilistic and explicitly recognizes the inherent limitations of incomplete earthquake data. The parameter posteriors we obtained indicate that the Groningen reservoir was not critically stressed prior to production.

We then introduced a workflow to forecast future seismicity at Groningen, incorporating epistemic and aleatoric uncertainty. Our forecast was constructed for the period February 2017 to the end of 2024 for three hypothetical extraction scenarios. We show that the largest event induced at Groningen during this period has a 5% likelihood of exceeding M 4.0. Considering the relatively shallow hypocentral depths of the events (~3 km) and the regional building stock which reflects low levels of historic seismicity, events of this magnitude will pose a greater hazard to local communities than similar-sized events elsewhere in the world.

Acknowledgments
This study was one of several contributed to the workshop on "Maximum Magnitude Estimates for Probabilistic Seismic Hazard and Risk Modeling in Groningen Gas Field" from 8 to 10 March 2016, in Amsterdam. We thank the organizers for inviting us. The earthquake catalog was obtained from KNMI at www.knmi.nl. The earthquake catalog, fault segments, and annual reservoir pressure snapshots used in this study are available as supporting information. Funding for an earlier version of the earthquake simulation code was provided by the Stanford Center for Induced and Triggered Seismicity, and a copy of the code can be obtained from the authors. We thank Oliver Maclaren for particularly enlightening conversations on sources of and strategies to quantify uncertainty, as well as Max Werner and an anonymous reviewer whose comments greatly improved this contribution.

References

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 Mattias Holschneider, *Bull. Seismol. Soc. Am.*, *107*, 1564–1567, doi:10.1785/0120170040.

Bommer, J. J., B. Dost, B. Edwards, P. J. Stafford, J. van Elk, D. Doornhof, and M. Ntinalexis (2016), Developing an application-specific ground-motion model for induced seismicity, *Bull. Seismol. Soc. Am.*, *106*, 158–173.

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, *J. Geophys. Res. Solid Earth*, *119*, 8991–9015, doi:10.1002/2014JB011663.

Byerlee, J. D. (1978), Friction of rocks, *Pure Appl. Geophys.*, *116*, 393–402.
Chan, A. W., and M. D. Zoback (2002), *Deformation Analysis in Reservoir Space (DARS): A Simple Formalism for Prediction of Reservoir Deformation With Depletion*, Soc. of Pet. Eng., SPE-78174-MS, Irving, Tex.
Dempsey, D., and J. Suckale (2016), Collective properties of injection-induced earthquake sequences: 1. Model description and directivity bias, *J. Geophys. Res. Solid Earth*, *121*, 3609–3637, doi:10.1002/2015JB012550.
Dempsey, D., J. Suckale, and Y. Huang (2016), Collective properties of injection-induced earthquake sequences: 2. Spatiotemporal evolution and magnitude frequency distributions, *J. Geophys. Res. Solid Earth*, *121*, 3638–3665, doi:10.1002/2015JB012551.
Dieterich, J. (1994), A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, *99*, 2601–2618.
Doherty, J., and S. Christensen (2011), Use of paired simple and complex models to reduced predictive bias and quantify uncertainty, *Water Resour. Res.*, *47*, W12534, doi:10.1029/2011WR010763.
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.*, Royal Netherlands Meteorol, Inst., KNMI.
Ellsworth (2013), Injection-induced earthquakes, *Science*, *341*, 1225942, doi:10.1126/science.1225942.
Foreman-Mackey, D., D. W. Hogg, D. Lang, and J. Goodman (2013), EMCEE: The MCMC hammer, *Publ. Astron. Soc. Pac.*, *125*(925), 306.
Gerstenberger, M. C., S. Wiemer, L. M. Jones, and P. A. Reasenberg (2005), Real-time forecasts of tomorrow's earthquakes in California, *Lett. Nat.*, *435*, 328–331.
Hanks, T. C., and H. Kanamori (1979), A moment magnitude scale, *J. Geophys. Res.*, *84*, 2348–2350.
Healy, J. H., W. W. Rubey, D. T. Griggs, and C. B. Raleigh (1968), The Denver earthquakes, *Science*, *161*(3848), 1301–1310.
Jaeger, J. C., N. G. Cook, and R. W. Zimmerman (2009), *Fundamentals of Rock Mechanics*, John Wiley, New York.
Kanamori, H., and D. L. Anderson (1975), Theoretical basis of some empirical relations in seismology, *Bull. Seismol. Soc. Am.*, *65*, 1073–1095.
Kraaijpoel, D., and B. Dost (2013), Implications of salt-related propagation and mode conversion effects on the analysis of induced seismicity, *J. Seismol.*, *17*, 95–107.
Lindqvist, B. H., and G. Taraldsen (2013), Exact statistical inference for some parametric nonhomogeneous Poisson processes, *J. Iranian Stat. Soc.*, *12*(1), 113–126.
Marzocchi, W., and J. D. Zechar (2011), Earthquake forecasting and earthquake prediction: Different approaches for obtaining the best model, *Seismol. Res. Lett.*, *82*, 442–448.
Nederlandse Aardolie Maatschappij (NAM) (2016a), *Groningen Seismic Hazard and Risk Assessment: Report on M_{max} Workshop*, 481 pp. [Available at http://www.nam.nl/feiten-en-cijfers/onderzoeksrapporten.html, last accessed March 2017.]
Nederlandse Aardolie Maatschappij (NAM) (2016b), *Technical Addendum to the Winningsplan Groningen 2016: Production, Subsidence, Induced Earthquakes and Seismic Hazard and Risk Assessment in the Groningen Field*, 57 pp. [Available at http://www.nam.nl/algemeen/mediatheek-en-downloads/winningsplan-2016.html, last accessed March 2017.]
Ogata, Y. (1999), Estimating the hazard of rupture using uncertain occurrence times of paleoearthquakes, *J. Geophys. Res.*, *104*(B8), 17,995–18,014.
Orlic, B., and B. B. T. Wassing (2013), A study of stress change and fault slip in producing gas reservoirs overlain by elastic and viscoelastic caprocks, *Rock Mech. Rock Eng.*, *46*, 421–435.
Richards-Dinger, K., and J. H. Dieterich (2012), RSQSim earthquake simulator, *Seismol. Res. Lett.*, *83*, 983–990.
Sanz, P. F., S. P. Lele, K. H. Searles, S. Y. Hsu, J. L. Garzon, J. A. Burdette, W. E. Kline, B. A. Dale, and P. D. Hector, (2015), Geomechanical analysis to evaluate production-induced fault reactivation at Groningen gas field, SPE Ann. Tech. Conf. Exhib. SPE-174942-MS, Houston, Tex.
Segall, P., and S. D. Fitzgerald (1998), A note on induced stress changes in hydrocarbon and geothermal reservoirs, *Tectonophys.*, *289*, 117–128.
Segall, P., and S. Lu (2015), Injection-induced seismicity: Poroelastic and earthquake nucleation effects, *J. Geophys. Res. Solid Earth*, *120*, 5082–5103, doi:10.1002/2015JB012060.
Segall, P., J.-R. Grasso, and A. Mossop (1994), Poroelastic stressing and induced seismicity near the Lacq gas field, southwestern France, *J. Geophys. Res.*, *99*, 15,23–15,438.
Uenishi, K., and J. R. Rice (2003), Universal nucleation length for slip-weakening rupture instability under nonuniform fault loading, *J. Geophys. Res.*, *108*(B1), 2042, doi:10.1029/2001JB001681.
van Eck, T., F. Goutbeek, H. Haak, and B. Dost (2006), Seismic hazard due to small-magnitude, shallow-source, induced earthquakes in the Netherlands, *Eng. Geol.*, *87*, 105–121.
Zang, A., V. Oye, P. Jousset, N. Deichmann, R. Gritto, A. McGarr, E. Majer, and D. Bruhn (2014), Analysis of induced seismicity in geothermal reservoirs—An overview, *Geothermics*, *52*, 6–21.
Zoback, M. D. (2010), *Reservoir Geomechanics*, Cambridge Univ. Press.
Zöller, G., and M. Holschneider (2016), The maximum possible and the maximum expected earthquake magnitude for production-induced earthquakes at the gas field in Groningen, the Netherlands, *Bull. Seismol. Soc. Am.*, *106*(6), 2917–2921.