An operational earthquake forecasting system (OEFS) to control induced seismicity in the Groningen natural gas field

Jansen, Jan Dirk

DOI
10.4233/uuid:403d77b5-0e2f-4978-a9cd-8c5789871b33

Publication date
2018

Document Version
Final published version

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.
An operational earthquake forecasting system (OEFS) to control induced seismicity in the Groningen natural gas field

Jan Dirk Jansen, Delft University of Technology, Department of Geoscience and Engineering
j.d.jansen@tudelft.nl
White paper, version 2, 8 May 2018
https://doi.org/10.4233/uuid:403d77b5-0e2f-4978-a9cd-8c5789871b33

Abstract
This ‘white paper’ proposes to develop a data-informed operational system to forecast induced seismicity, and the associated seismic hazard in the Groningen natural gas field. The goals of developing and maintaining such an Operational Earthquake Forecasting System (OEFS) are 1) to provide a unified environment to test and align research efforts from, e.g., the DeepNL and KEM national research programs; 2) to improve the quality of seismic hazard forecasts by combining a broad range of measured data with ‘evergreen’ models based on large-scale numerical simulation and systematic data assimilation; and 3) to provide a testbed for the development of operational procedures to minimize seismic hazard and risk such as (adaptive) traffic light systems, or optimized spatial and temporal production and/or re-injection rates. Like weather forecasting systems, OEFS depends on data assimilation, e.g. the systematic combination of uncertain measured data with uncertain models such that the combined result has a better forecasting capability than the data or models on their own. The aim is to use physics-based models in OEFS whenever possible, in combination with probabilistic models whenever necessary. The wide variety in temporal and spatial scales that govern the physical processes behind induced seismicity imply that OEFS will have to be based on a combination of multiple computational models, if necessary using simplified physics, which will be computationally intensive. Moreover, to systematically capture uncertainties in geology and physical parameters, OEFS will need to make use of ensembles of realizations which further increases the computational requirements. Developing and maintaining OEFS will therefore be a major exercise which is probably best done by one or more large technological institutes (KNMI, TNO, Deltares) with input from academic institutions.

Introduction
We consider induced seismicity related to natural gas production in the Groningen field; for an overview see, e.g., Van Thienen-Visser and Breunese (2015) and the recently published “Special Issue on Induced Seismicity in the Groningen Gas Field” of the Netherlands Journal of Geosciences (Van Geuns and Van Thienen-Visser, 2018).

In line with the convention in seismology, we will use the notion of ‘forecasting’ to indicate the probability of occurrence of earthquakes in a given magnitude range within a given time window in a given area. This is as opposed to ‘prediction’ which implies a much more precise statement about the occurrence of a seismic event of a specific magnitude at a specific moment and location (Jordan et al. 2011).

We also make a distinction between ‘triggered’ earthquakes and ‘induced’ earthquakes. Both are related to anthropogenic activities, e.g. natural gas production or injection, but for the former, the source of energy stems from natural plate tectonics whereas for the latter the energy originates from the production or injection process itself, through compaction or expansion of reservoir rock (Dahm et al., 2015). However, the difference between triggered and induced earthquakes is not always clear-cut.

Moreover, in line with the accepted practice in seismology we will use the term ‘seismic hazard’ to refer to a forecast of ground motions, e.g. peak ground accelerations or velocities. ‘Seismic risk’ then refers to a forecast of fatalities or damage, and depends on combining the seismic hazard with forecasts of the dynamic response of buildings and other surface infrastructure.
Due to the complex physics of stick-slip friction processes in seismicity and the inherent uncertainties in subsurface material properties and in-situ stress conditions prior to human intervention, the exact prediction of (the triggering of) induced seismic events is, at present, generally impossible. However, forecasting the probability of seismic events in the Groningen field is possible within statistical bounds. Currently, the most successful methods to do so appear to be those based on statistical models (also referred to as probabilistic models) which correlate gas production over time with the occurrence of seismic events using catalogues of recorded earthquakes. In particular the operator of the field has proposed different seismic models in other to quantify the seismic hazard for various production scenarios (Bourne et al. 2014, 2015; Bourne and Oates 2015; Nederlandse Aardolie Maatschappij 2016; Van Elk and Doornhof, 2017). In addition there have been attempts to forecast seismicity in the Groningen field with aid of physics-based models (also referred to as geomechanics-based models) or combinations of probabilistic and physics-based models; see Sanz et al. (2015), Dempsey and Suckale (2017), Bourne and Oates, 2017, and van Wees et al. 2018. Note that physics-based subsurface models unavoidably also contain many uncertain parameters and therefore always need to be calibrated with the aid of measured data.

The use of large-scale physics-based models of atmospheric fluid dynamics is commonplace in the meteorological community for what is loosely known as ‘weather prediction’, but what is in fact ‘weather forecasting’, i.e. determining the probability of occurrence of atmospheric state variables (pressure, temperatures, humidity and wind velocity) in given magnitude ranges within a given time window in a given area; see, e.g., ECMWF (2018). A key elements in numerical weather forecasting is the use of ‘uncertain models’. These are either models with uncertain parameters described with statistical distributions, or ensembles of model ‘realizations’ generated with randomly selected parameter values from these distributions. Another key element is the use of near-real time ‘data assimilation’ of large amounts of measured data to keep the underlying models ‘evergreen’. The underlying idea is that the systematic combination of uncertain measured data with uncertain models provides a better forecasting capability than the data or models can provide on their own. In particular the use of ensembles of models to capture uncertainty and the development of associated data assimilation methods has seen major developments in meteorology, oceanography and hydrocarbon reservoir simulation over the past decades (Gneiting and Raftery, 2005, Evensen, 2009).

Conceptually, OEFS would be similar to the large-scale atmospheric models used in weather forecasting. The latter can be used to, e.g., forecast thunderstorms, i.e. to compute their probability of occurrence in a certain period and region without being able to exactly pinpoint the time and location of lightning bolts. Similarly, OEFS could be used to compute the probability of occurrence of seismicity in a certain period and region, although the exact prediction of the time and location of seismic events will remain out of reach. The aim is to use physics-based models in OEFS whenever feasible. However, because it has until now not been possible to provide reliable forecast using physics-based models alone (and it is unlikely that this will be possible in the near future), a combination with probabilistic models is expected to be used.

OEFS can serve to formulate ‘measurement and control protocols’ or ‘traffic light systems’ with threshold values for observables that have a sound statistical basis. In particular OEFS could form the basis for the development of an ‘adaptive traffic light system (ATLS)’ similar to those proposed over the past years for the reactive control of injection-induced seismicity in geothermal applications, especially in the group of Wiemer et al. (ETH Zürich); see, e.g., Kiraly-Proag et al. 2016, 2018 and Mignan et al. 2017. In these ATLS, frequent measurements are used to continuously adapt threshold values for operational parameters to keep the seismic hazard below a predetermined value over a predetermined time window. In the most recent paper (Kiraly-Proag et al. 2018), use is made of ensembles of different probabilistic models (with more or less physics-based elements) to capture model uncertainties. In OEFS we propose to use, in addition, ensembles to capture parameter uncertainties (like in meteorological applications).

Moreover, unlike weather forecasting models, OEFS can be used for more than forecasting: gas production operations are human-controlled, and the system can therefore also be used to
determine the best operational strategy to minimize seismic hazards. In particular, we propose to move beyond reactive control by optimizing production rates, or combined production and re-injection rates for pressure maintenance, or even well locations, in a pro-active manner. Such an operational use of OEFS would then be similar to ‘closed-loop reservoir management’ and ‘closed-loop field development’ approaches as developed in the petroleum and resources extraction industries over the past decade (Jansen et al. 2008, 2009; Sarma et al. 2008; Chen and Oliver, 2010; Chen et al. 2012; Shirangi and Durlofsky, 2015; Benndorf and Jansen, 2017).

**Operational Forecasting and Earthquake Simulation**

In an extensive review report commissioned by the Italian Government, Jordan et al. (2011) give an overview of probabilistic long-term (on a time scale of years) and short-term (days) forecasting methods. Long-term forecast are typically probabilistic, time-independent and based on long histories (catalogues) of recorded earthquakes. Sometimes time dependency is introduced through physical arguments based on elastic rebound theory which accounts for the effect that after a major earthquake it will take considerable time before the fault becomes critically stressed again. Short-term forecasts are always time-dependent, mostly probabilistic, and are usually based on the observation that earthquakes are often clustered in time and space which may allow for a forecast of the probability of aftershocks and sometimes even of the probability of a main event.

A particularly interesting form of short-term forecasting makes use of earthquake simulators which model the dynamics of interacting faults with the aid of (simplified) physical simulations of wave propagation. A well-documented example is the Virtual California earthquake simulator and its successor Virtual Quake; see, e.g., Rundle et al. (2006); Yikilmaz et al. (2010); Sachs et al. (2102) and Yoder et al. (2015). These and other simulators aim at increased insight in earthquake occurrence by combining physical models with measured data; for an overview see the ‘Focused issue on earthquake simulators’ of Seismological Research Letters (Tullis, 2012 and Tullis et al., 2012).

Earthquake simulators have the potential to improve forecasts of seismicity by combining physics-based models and data but typically suffer from a lack of understanding of the physical processes, a lack of knowledge of many of the relevant physical and geological properties, and the necessity to strongly approximate the full elastodynamic behavior of the subsurface because of computational limits. The same potential and shortcomings will hold for OEFS as proposed in the present paper. An important difference is, however, that the key source of seismic energy in the Groningen field is human-controlled gas production, whereas in case of natural seismicity the source is plate tectonics which is completely outside human control. Nevertheless, a combination of physics-based and probabilistic models will be necessary, but with the aim to expand the inclusion of physical mechanisms as far as possible.

**OEFS - Specifics**

**Goal**

The goals of developing and maintaining OEFS are:

- to provide a unified environment to test and align research efforts from, e.g., the DeepNL program administered by Netherlands Organisation for Scientific Research (NWO, 2017) and the Kennisprogramma Effecten Mijnbouw (KEM) (knowledge program effects of mining) administered by the Netherlands Regulatory Authority (SodM, 2017);
- to improve the quality of seismic hazard forecasts by combining a broad range of measured data with ‘evergreen’ physics-based models (if necessary in combination with probabilistic ones) based on large-scale numerical simulation and systematic data assimilation;
- to provide a testbed for the development of operational procedures to minimize seismic hazard and risk such as (adaptive) traffic light systems, or optimized spatial and temporal production and/or re-injection rates.

It should be noted that the proposed scope of OEFS is currently limited to seismic hazard forecasts in line with the emphasis of the DEEP NL program on the subsurface aspects of induced seismicity. An
extension to seismic risk would require a major additional effort to quantify the response of buildings and other surface-based infrastructure along the lines of the Seismic Hazard and Risk assessment procedure as currently routinely performed by NAM (Van Elk and Doornhof, 2017).

**General set-up**

Figure 1 displays a schematic of OEFS as an operational forecast system with data assimilation to continuously improve system models. The white box at the top of the figure represent reality (i.e., the Groningen natural gas reservoir, overburden, underburden and neighboring rock as far as relevant to induced seismicity.) The white boxes at the bottom represent a variety of system models. These could be ensembles of flow, geomechanical and seismic (wave propagation) models at different spatial and temporal scales and at varying levels of detail and physical simplification, possibly in combination with probabilistic models. The input to the real system (green box top left) consists of operational variables in the form of gas production rates (and possibly gas re-injection rates). Measured values of these variables form the input to the various models. Both the real system and the models will produce outputs: for the real system this will be acoustic signals observed by surface- and well-based seismometers and geophones, subsidence as measured by remote sensing and in situ techniques (INSAR, gravimetry, etc.), and (well head and/or bottom hole) pressures, temperatures, and water rates (green box top right); for the models this will be the same geophysical and production variables but then as predicted rather than as measured (green box bottom right).

Inevitably there will be discrepancies between the measured and the predicted outputs (green box middle right) which serve as a basis for (near-) continuous model improvement with aid of systematic data assimilation techniques (Lewis et al. 2006; Evensen, 2009). The frequency of the updates will depend on the dynamics of the processes involved (notably pressure transients in the reservoir) and the seismicity rate. The updated models could be used on-line or off-line to create seismic hazard forecasts, test operational strategies, provide a basis for developing ATLS, and optimize operational strategies as schematically indicated by the dotted arrow at the bottom of the figure.

**Physics-based, ensemble based, multi-scale and multi-model**

The current seismic hazard and risk assessment procedure of the operator uses a probabilistic approach (Van Elk and Doornhof, 2017) based on the classic paper of Cornell (1968). OEFS is meant to complement this approach, at least for the hazard assessment, with the aid of physics-based models wherever feasible in combination with reduced-physics and/or probabilistic models whenever necessary. The physics of production-induced seismicity is governed by processes at widely varying scales in space and time in a domain with strongly heterogeneous flow and mechanical properties. Reservoir depletion and the resulting compaction and subsidence take place over several decades and cover a large rock volume around the entire reservoir and its overburden. Pore pressure transients and the resulting stress changes are governed by poroelastic (‘slow’ Biot) waves and multi-phase (gas/condensate/water) effects on time scales ranging from hours to years and are spatially...
confined to the reservoir. Seismic events are governed by poroelasto-dynamic ('fast' Biot) waves in the order of seconds to minutes which travel widely outside the reservoir including the overburden where heterogeneities and specific soil properties may strongly influence the dynamics and thus the seismic hazard.

A single computational model capturing all these scales would require an astonishing number of grid cells (elements) and time steps which is at present not a realistic option. Although the rapid development of multi-scale models with adaptive grid refinement and adaptive time stepping schemes offers future perspectives for the development of such uniform computational models, it is foreseen that OEFS will have to be based on different families of models: those for the (quasi-static) depletion process, those for the transient pore-pressure propagation in response to operational control, and those for the simulation of seismic events and the corresponding fault interactions. The latter could be full rupture models or simplified-physics models as, e.g., employed in earthquake simulators used for natural seismicity (Tullis et al. 2012). In addition to serving as forward simulators, the various models will also be of relevance for inverse modelling, e.g. to pin point source locations using measured seismic data or to infer reservoir properties using pressure transient analysis.

An essential element of OEFS will be the capturing of uncertainty in geological structures (facies, stratigraphy, faulting, fracturing), constitutive behavior (fault friction, compaction) and other reservoir flow/geomechanical variables (permeability, porosity, stiffness, viscosity). This will require ensembles of realizations (simulation models) based on sampling of statistical distributions or being itself the result of simulation models with stochastic components (e.g. process-based simulations of geology). The simulation requirements scale linearly with the number of ensemble members (typically in the order of hundreds to thousands), but allow for trivial parallelization. Notwithstanding the persistent increase in computing power, especially through massive parallelization, the computational requirements for OEFS will be very large.

**Parties involved**
Developing and maintaining OEFS will be a major exercise which is probably best done by one or more major technological institutes (KNMI, TNO, Deltares) with input from academic institutions. Operation of OEFS will require near-continuous access to measured geophysical data (seismic, strain, subsidence) as well as measured production data (rates, pressures, composition) which will have to be made available by the operator of the field (NAM). Of course a system similar to OEFS could be developed and maintained by the operator. Such a parallel development, preferably using different physical models and assumptions, would allow for a comparison of results and could serve as a basis for informed discussions between operator and regulator (SodM) about control of production (and re-injection) operations to minimize seismic hazard and risk in the Groningen field. Moreover, systems similar to OEFS as proposed for Groningen could be developed for control of gas production or storage operations in other natural gas fields.

**References**


ECMWF, 2018: *European Centre for Medium Range Weather Forecasting*. https://www.ecmwf.int/


Acknowledgments

The first version of this white paper was published on 5 March 2018. This second version has benefitted from comments and suggestions made by Giovanni Bertotti (TU Delft), Pavel Ditmar (TU Delft), Olwijn Leeuwenburgh (TNO and TU Delft), Annemarie Muntendam-Bos (SodM), Simone van Schijndel (Deltares), Karin van Thienen-Visser (TNO), Jan-Diederik van Wees (TNO and Utrecht University), and Karl-Heinz Wolf (TU Delft). I like to thank all these people for their contributions. I also like to make clear that they have not seen the revised version of the document prior to its publication and therefore do not necessarily fully agree with its contents.