

This manuscript investigates the potential mitigation of induced seismicity in the Groningen gas field through nitrogen injection into the depleted reservoir following the cessation of gas production. The study employs the existing TNO modelling framework, integrating reservoir flow simulations, a seismic source model (SSM), ground motion modelling, and seismic hazard and risk assessment to evaluate the potential impact of several injection scenarios on seismicity and associated societal risk.

The topic is relevant to NHESS and timely. Although gas production in the Groningen field ceased in October 2023, seismicity is expected to persist due to delayed reservoir compaction and ongoing stress redistribution. Exploring potential mitigation strategies for post-production seismicity is therefore scientifically and societally important. The integrated modelling chain used in this study provides a structured framework to link reservoir pressure evolution to seismicity forecasts and seismic risk metrics.

The manuscript is generally well written and logically structured. The scenario-based analysis provides useful insight into how reservoir repressurization through nitrogen injection may influence the temporal evolution of seismicity and associated seismic risk. The results indicate that nitrogen injection may reduce the predicted seismicity rate and corresponding risk metrics compared to a reference case without injection.

Several modelling assumptions and simplifications are already acknowledged and discussed by the authors in the discussion section. Since these limitations are transparently presented in the manuscript, it is not necessary to reiterate them in this review. My main comment concerns the validation of the modelling framework against historical observations, which would help strengthen confidence in the predictive capability of the model.

The seismic source model (SSM) described in Section 2.2 is a central component of the modelling framework, as it converts reservoir pressure changes into forecasts of seismic activity. The manuscript states that the SSM has been calibrated against historical seismicity, and the comparison between observed and simulated number of events per year appears to show a good agreement (e.g., the green shaded part in Figures 6–13). However, it is less clear how well the calibrated SSM reproduces other important characteristics of the historical seismicity. In particular, the manuscript does not appear to show comparisons of the spatial distribution of events between observed and simulated seismicity. Since fault geometry and spatial stress redistribution play an important role in induced seismicity, demonstrating that the model reproduces the spatial pattern of seismicity would strengthen confidence in the calibration.

In addition, it would be useful to clarify how the model represents larger events, particularly the Huizinge earthquake (M3.6 in 2012), which represents the largest recorded event in the Groningen field. It is not clear from the manuscript whether events of this magnitude can be reproduced by the calibrated SSM, or whether the calibration mainly focuses on reproducing the overall seismicity rate. Demonstrating whether events with magnitude larger than 3 can be reproduced within the calibrated SSM would be helpful for assessing the model's capability to represent the historical seismicity.

Furthermore, since the seismicity in Groningen is closely related to reservoir compaction and resulting surface subsidence, it would also be useful to show how well the reservoir model reproduces the historical subsidence evolution observed in the Groningen field. Subsidence measurements provide an important constraint on reservoir compaction behaviour and therefore indirectly on the stress changes driving seismicity. A brief comparison between

observed and simulated subsidence evolution would therefore further strengthen the validation of the modelling framework.

Overall, a figure or additional comparison showing (i) spatial seismicity patterns, (ii) magnitude distribution including larger events such as the Huizinge earthquake, and (iii) subsidence history matching would provide a more comprehensive validation of the modelling framework against historical observations.

RESPONSE

We thank the anonymous referee for their time to read and review our manuscript and for the thorough and constructive suggestions for improvement and clarification.

The central comments focus on establishing the predictive capabilities of the modelling framework, for which calibration and validation against historical observations is required. In particular, the referee suggests:

- Showing the historical predictions (hindcast) of the calibrated source model against observations in the spatial and magnitude domain.
- Showing a comparison between modelled reservoir compaction and resulting subsidence against historically observed subsidence.

Both are valid and valuable suggestions. Regarding the first suggestion, we will directly follow the suggestion to add additional figures regarding the modelled spatial seismicity pattern and magnitude distribution against historical observations, and add additional references to more detailed manuscripts describing the match. The model used is widely considered ‘state-of-the-art’, both in terms of spatial performance and magnitude prediction. We agree that showing this more thoroughly will indeed provide more support for the validity of the modelling approach.

Regarding the second point, in the revised manuscript, we will add an expanded description explaining the origin and calibration of this compaction grid, including its performance. While we acknowledge the importance of subsidence as a validation constraint on the compaction model, we prefer to keep the discussion concise. Given that the primary objective of this paper is to evaluate seismic risk mitigation strategies rather than subsidence modeling itself, we believe that providing the calibration context and referencing the comprehensive foundational studies is the most effective way to establish the model’s validity without detracting from the manuscript's core focus.

The study explores whether injecting gas back into the Groningen gas field in the Netherlands could reduce earthquake activity. Using an adapted modeling framework, TNO Model Chain for Groningen, that accounts for pressure increases and cooling effects from injected gas, the study tested several hypothetical injection scenarios and compared them to a no-injection case. The results suggest that gas injection can significantly reduce seismicity, though benefits may diminish if injection stops before pressure equilibrium is fully restored across the reservoir. Overall, this study is well-designed and comprehensive, and the findings are meaningful for earthquake risk management for gas production areas. However, I still have several comments and suggestions as follows.

- 1) Figure 1: It is suggested to discuss how the model and data uncertainties/errors propagate through the workflow.
- 2) In the Methodology section, how was the grid size of the models determined, and how would the different spatial resolution of the model grid affect the simulated results?
- 3) Equations 2.a and 2.b: What is the difference between the subscripts “H” and “h”?
- 4) Equation (5) may not be correct because the Young’s modulus was not included.
- 5) Figure 2 and other maps: It is suggested to add a north arrow and a scale bar to the maps, and it is also better to add units to the color bars of the simulated results.
- 6) Line 322: Could you justify if the calibrated parameters based on historical data are still applicable to the future scenarios?
- 7) Line 497: It is suggested to clearly define the “personal risk” or “seismic risk” in the Methodology section. Is the probability of property loss also considered? Remove the repeated explanations in Lines 521-522.
- 8) Line 25: A typo, change “The are two main reasons for this” to “There are...”
- 9) Lines 130 and 492: Correct the punctuation marks such as “,”.
- 10) Line 149 and Table 1: Please keep the notation of the variables (e.g., P_p) or scenarios (e.g., LO-10ZB or LO_10ZB) consistent throughout the manuscript.
- 11) Lines 203, 209, and 257: “Table 2” should be “Table A2”.
- 12) Lines 238-239: Correct the format of the units, cubic meters.
- 13) Section 4.1: Use either “MPa” or “bar” throughout the discussion.
- 14) Line 333: “Fig. 7 to 13” or “Fig. 6 to 13”?

RESPONSE

We thank the anonymous referee for their time to read and review our manuscript and for the thorough and constructive suggestions for improvement and clarification. We will address all comments below, following the original numbering used in the review.

1. We agree that this point requires additional attention in the manuscript. We will add a section, as uncertainty/error is central to probabilistic seismic hazard and risk analysis. We carefully deal with both aleatory variability and epistemic uncertainty in the entire modelling chain, and this will be more clearly reflected in the revised manuscript.
2. We will add a couple of lines related to the discretization. While some models we've applied do not easily allow changing the resolution, where possible we've chosen all levels of discretization such that further refinement does not meaningfully affect the results. This will be made more explicit.
3. These are the maximum and minimum horizontal stress and should have been mentioned. We will fix this omission.
4. The Young's Modulus is in fact included since it's a factor in H . However, we agree that this should be clarified.
5. We will add these.
6. This aligns partially with the comments of the other referee. We will add additional justification for this.
7. We will clarify the definition of personal risk/seismic risk more carefully. It is the risk of dying due to building collapse resulting from an induced earthquake. This is the main metric used in the Netherlands for regulating seismic risk, and should be clearly defined in this manuscript.
8. We will fix this.
9. We will fix this
10. We will double-check for consistent notation throughout
11. We will fix this
12. We will fix this
13. We will double-check for consistent units throughout
14. We will fix this