



Modelling the Dynamic Relationship Between Mining Induced Seismic Activity and Production Rates, Depth and Size: A Mine-Wide Hierarchical Model

J. MARTINSSON^{1,2} and W. TÖRNMAN^{2,3}

Abstract—The dynamic properties of mining induced seismic activity with respect to production rate, depth and size are studied in seven orebodies in the same underground iron ore mine. The objective is to understand the relationship between the measured seismic activity and the: seismic decay time, planned production rate, production size and mining depth. This relationship is the first step to individually customise the production rate for each orebody in the mine, make short-term predictions of future seismicity given planned productions, and to find out in what way the available predictors affect the seismicity. The seismic response with respect to the dependent variables is parametrised and the estimated decay times for each orebody, which are of particular interest here, are compared. An autoregressive model is proposed to capture the dynamic relationship between the induced seismic activity, the current production rate and the past seismic activity. Bayesian estimation of the parameters is considered and parameter constraints are incorporated in the prior distributions. The models for all orebodies are tied together and modelled hierarchically to capture the underlying joint structure of the problem, where the mine-wide parameters are learnt together with the individual orebody parameters from the observed data. Comparisons between the parameters from the hierarchical model and independent models are given. Group-level regressions reveal dependencies on size and mining depth. Model validation with posterior predictive checking using several discrepancy measures could not detect any model deficiencies or flaws. Posterior predictive intervals are evaluated and inference of model parameters are presented.

Keywords: Induced seismicity, Bayesian hierarchical model, Probabilistic forecasting, Time-series analysis, Statistical seismology.

List of Symbols

y_{ij}	Number of detected events per week within the i th week and in the j th orebody
p_{ij}	Planned accumulative rock mass in megatonne (Mt) retrieved by the production blasts during the i th week and in the j th orebody. This estimate is based on planned production and is not measured
μ_j	Expected number of detected events in the j th orebody
v_j	Dispersion parameter (or scale parameter) describing the width of the distribution modelling the activity in the j th orebody
θ	First level model parameters, describing the individual orebody
θ_{1j}	Autoregressive parameter describing the seismic decay in the j th orebody
θ_{2j}	Effect of the production in the j th orebody
θ_{3j}	Seismic exposure in the j th orebody
$\theta_{4j} = v_j$	Dispersion parameter (or scale parameter) describing the width of the distribution
τ_j	Half-life
ϕ	Second level model parameters, i.e. the hyperparameters, describing the behaviour of the group of orebodies
ϕ_1, ϕ_2	Hyperparameters for prior distribution for the autoregressive parameter θ_{1j}
ϕ_3, ϕ_4, ϕ_5	Hyperparameters for prior distribution for the effect of the production θ_{2j} , where ϕ_4 is the effect of the production size s_j

¹ Department of Engineering Sciences and Mathematics, Mathematical Sciences, Luleå University of Technology (LTU), 97187 Luleå, Sweden. E-mail: jesper.martinsson@ltu.se

² Luossavaara-Kiirunavaara AB (LKAB), R&D, 98186 Kiruna, Sweden.

³ Department of Civil, Environmental and Natural Resources Engineering, Luleå University of Technology (LTU), 97187 Luleå, Sweden.

ϕ_6, ϕ_7, ϕ_8	Hyperparameters for prior distribution for the seismic exposure θ_{3j} , where ϕ_7 is the effect of the production depth d_j
ϕ_9, ϕ_{10}	Hyperparameters for prior distribution for the dispersion θ_{4j}
s_j	Planned production size in the j th orebody in megatonne (Mt/week), estimated based on the average weekly production rate in the orebody. This estimate is based on planned production and is not measured
d_j	Production depth in the j th orebody

1. Introduction

To individually customise the mining operations with respect to the seismic impact in a certain orebody, it is important to understand how mining operations and mining conditions affect the measured seismic activity in the volume. Even though this effect is different for each orebody in the mine considered here (LKAB's iron-ore mine in Malmberget, northern Sweden) the orebodies are still related. They share not only the geological features present in the region (Bergman et al. 2001; Debras 2010), but they are also subjected to the same mining method (sub-level caving) and fairly similar mining conditions (Wettainen and Martinsson 2014). This type of higher level relationship suggests the use of *hierarchical modelling* (see e.g. Kruschke 2014; Gelman and Hill 2007; Gelman et al. 2004) with the ability to treat the model parameters for each orebody in the mine as one observation from a mine-wide distribution that describes the group of orebodies. The mine-wide model has in turn *group-level predictors* and *hyperparameters* (Gelman and Hill 2007) that are estimated from the data by treating the estimation problem jointly. The group-level predictors are orebody specific data (e.g. mining depth, average production rate, rock mechanic properties, spatial orientation and shape) that might affect the lower level parameters describing the seismic activity in each orebody. As opposed to non-hierarchical Bayesian models, data can enter a hierarchical model at

different levels of the hierarchy and is not limited to the likelihood function. The model parameters and corresponding predictors for each orebody are embedded in the likelihood function in the same way as with non-hierarchical models, while hyperparameters and corresponding group-level predictors are embedded in the prior distributions.

The hierarchical structure allows the data to inform us of similarities across the orebodies and to use this information in the estimation process for the individual models and group-level parameters. If the orebodies behave similar given the group-level predictors, this similarity information is used to shrink the posterior distributions of the parameters towards a mine-wide average (Kruschke 2014; Gelman and Hill 2007). The prior distributions of the model parameters are then more informative and their shape are defined by the group-level predictors and hyperparameters which are jointly estimated from the data. This means that with access to group-level predictors the prior distributions are sharper and more informative which helps in reducing the prediction uncertainty when predicting new outcomes. This process is analogous to non-hierarchical Bayesian or maximum likelihood regression where access to individual-level predictors often contracts the likelihood around a regression line.

For example, in the time interval we are studying here, mining depth and the size of the production are constant (or nearly constant) and can not be distinguished from the constant term in the model if the orebodies are modelled independently. However, constant orebody-specific properties can be analysed with group-level (or mine-wide) regression models by adapting a hierarchical model structure.

This is interesting, not just from an analysis point of view, but it will also yield more certain predictions when data is sparse. We will show that in the most extreme case, predictions can be made without any measurements of seismicity, based only on prior distributions and group-level predictors. This is convenient in areas not yet covered by the seismic system and the Bayesian reliance on prior distributions when data is sparse gives an intuitive solution. The best we can do when we have no data, is to base our prediction on the similarity information of the orebodies in the mine.

Depth dependencies can also be explored as in Vallejos and McKinnon (2011) by pooling the orebodies and consider a single model for the entire mine but at different mining levels. This will, however, eliminate the possibility to explore dissimilarities among the orebodies and our aim to customise the production for each one of them would be lost. Also, if the orebodies are not equal, pooling will inflate the overall uncertainty to handle the differences between the orebodies and produce less certain individual predictions.

For weekly or monthly predictions, the data shows a dynamic behaviour and, e.g., a high seismic activity this week increases the probability of high seismicity the following week. Dynamic behaviours that remembers past mining induced seismicity have been studied in Polish mines (Węglarczyk and S. Lasocki 2009) by evaluating the auto-correlation function. In this paper we are more interested in a model that includes this property in its structure to obtain better predictions. The straightforward way to capture this is with an autoregressive model (see e.g. Smyth and Mori 2011), where the past week's activity is included in the model as a predictor. Using the previous week as a predictor, the memory of the past activity will decrease exponentially with time and the autoregressive parameter can also be represented as the seismic half-life (see Appendix for more details). The impulse response of the autoregressive process will not follow Omori's Law (Omori 1894) but instead decay exponentially similar to many physical relaxation processes. This was also one of the early criticisms of the Omori's law (see e.g. Burridge and Knopoff 1967; Richter 1958). For longer time periods, those of interest in this study, it has been shown that the exponential decay predominates over power-law decay (Otsuka 1985, 1987; Souriau et al. 1982) for aftershocks. Also, the decay times in this study are in the order of weeks as opposed to hours or days for major events in mines (Vallejos and Mckinnon 2010; Olivier et al. 2015). Furthermore, as no single earthquake in our sequences is a main shock but the result of weakly production in the mine, the process may have a closer relation to that of earthquake swarms as opposed to aftershocks (see e.g. Špičák et al. 1999). The decay of earthquake swarms has been shown to follow both

exponential decays (Mogi 1987, 1991; Tsukuda 1991, 1993) and power-law decays (Mogi 1991; Tsukuda 1993) and may depend on the magnitude of strength at the source (Ogasawara 2002) or a change in fracture condition (Hirata 1987).

The aim of this study is to find a suitable model that describes the observed seismicity in each orebody in the mine, given a set of the most commonly available predictors in a mine: weekly production, past seismicity, mining depth, production size. From this model we then want to:

- Make short-term predictions of future seismicity with realistic prediction uncertainties, given planned productions. From predictions it is possible to customise the production rates for each orebody under the constraints of limited induced seismicity.
- Estimate the effects of the different predictors on the observed seismicity.
- Find possible similarities between the orebodies.
- Find important group-level predictors to obtain more informative prior distributions for the individual parameters.

Additional predictors can easily be included in the model, but we limit our analysis to data that are commonly available in most mines; making the proposed model applicable in similar settings. A deeper understanding of the underlying causes of the effects and their correlation to rock mechanical properties is not the objective here as it would require much more comprehensive data that are absent in most settings. The magnitude information is not included in the analysis, but this information can easily be combined with the predicted activity from the model proposed here to give a more comprehensive probabilistic earthquake hazard assessment in the mine (see e.g. Martinsson and Jonsson 2018). Planned weekly production (M_t), estimated using production layouts and expected densities, is used here as a predictor when modelling mining induced seismicity and it is roughly proportional to the excavated volume per week that in turn is related to the accumulated moment McGarr (1976). This is a natural choice as information of planned production is commonly available in mines and relationships between induced seismicity and related predictors

such as the total injected volume or the injection rate has been established before (see e.g. Shapiro et al. 2007, 2010; Bachmann et al. 2011; Mena et al. 2013; Leptokaropoulos et al. 2017; Garcia-Aristizabal 2018).

2. Data and Methods

The mine under investigation is LKAB's underground iron ore mine in Malmberget, located in northern Sweden. Long-term predictions of mining induced seismicity in this mine has been studied earlier (Wettainen and Martinsson 2014), aiming to deliver long-term prognosis of future ground vibration levels as a function of the production years ahead. The work here is instead devoted to short-term predictions (1 week ahead), exploiting the short-term dynamic behaviour of the observed induced seismicity. The Malmberget mine consists of around twenty orebodies of varying sizes, shapes and depths. The seven major orebodies are covered by the seismic measurement system and considered in the analysis here. The other orebodies are smaller, partly mined out, and not expected to cause significant seismic events (Wettainen and Martinsson 2014).

2.1. Data

The seismic and mine specific data used in this analysis was collected between 2010.01.01 and 2013.10.14 under which the seismic system remained constant. The measured seismic events are all manually processed and accepted as seismic events in the catalogue, with average location errors of ≈ 33 meters (Martinsson 2012). The data set used here is exported from the analysis software mXrap (formerly MS-RAP), developed by the Australian Centre for Geomechanics. In this software the events are automatically placed, based on their location, in clusters that are manually associated with one of the seven major magnetite orebodies covered by the measurement system. These orebodies are indexed by $j = 1, \dots, 7$ and correspond to Alliansen, Dennewitz, Fabian, Kapten, Parta, Printzsköld and Viri (Vitåfors/ridderstolpe), respectively, shown in Figs. 1 and 2. The spatial clustering enables us also to remove

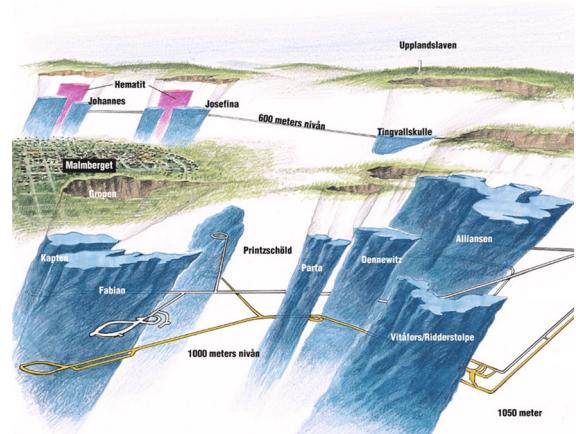


Figure 1
Illustrating the ore bodies at the Malmberget mine where the major orebodies shown in blue are all magnetite and plunging 40–50° towards SSW (LKAB photo archive)

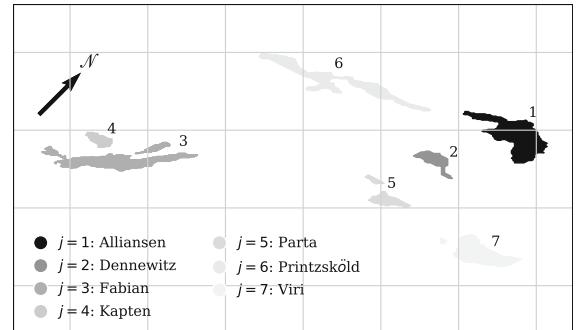


Figure 2
Cross sections of the seven major magnetite orebodies shown in Fig. 1 based on planned production at the individual depths shown in Fig. 6. The orebodies are coloured and indexed by $j = 1, \dots, 7$ and correspond to Alliansen, Dennewitz, Fabian, Kapten, Parta, Printzsköld and Viri (Vitåfors/ridderstolpe), respectively. The grid size is 500 × 500 m and the coordinate system is rotated 45° (relative to the original mine coordinates) to match the view in Fig. 1

events caused by stationary mining noises (such as ore-passes and crushers) that occasionally slips through manual rejection. This is important as it may otherwise affect the analysis with respect to the production.

2.2. Inferential Methods

The main reasons for applying Bayesian hierarchical analysis are: (1) that the Bayesian hierarchical structure matches our situation with one mine

containing several orebodies (Kruschke 2014; Gelman and Hill 2007); (2) to account for individual- and group-level variation in estimating group-level regression coefficients (Gelman and Hill 2007); (3) to obtain reasonable estimates for orebodies with small sample sizes; (4) to provide richer inference and avoid the common problems associated with p-values (e.g. dependencies on the sampling and testing intentions (Kruschke 2013); (5) to make direct probability statements about the effect given the observed evidence (Kruschke 2014).

The inference is based on Markov Chain Monte Carlo (MCMC) methods (Gelman et al. 2004; Kruschke 2014; Chen et al. 2000) providing realistic estimates of credible intervals of parameter values (Martinsson 2012) compared to, e.g., approximations based on derivatives. Slice sampling (Neal 2003) is used and inference is calculated from 100,000 samples from the posterior distribution with converged chains. However, alternative open-source statistical packages such as Stan (Carpenter et al. 2017) or JAGS (Plummer 2003) are available for both sampling and modelling, (see e.g. Kruschke (2014), for a nice introduction.)

Credible intervals are here based on equal-tail intervals using empirical percentiles and point estimates are obtained using the sample median rather than the sample mean. The main reason is to preserve point estimates (location, and interval values) under monotonic transformations. For example, this means that point estimates and intervals obtained for the autoregressive parameter can easily be recalculated to the seismic half-life (or the other way around).

3. Model Selection

In ideal cases the occurrence of the seismic events in time can be described by a Poissonian process with an area-specific mean activity rate (see e.g. Kijko and Graham 1999). In practice, the distribution of count data are often *overdispersed* (Gelman and Hill 2007) with variance greater than the mean value. The Poisson distribution, commonly used to describe counts, requires the variance to be equal to the mean value and this restriction often results in underestimated uncertainties when applied to real data. This is

also the case for the count data in this study (i.e. measured number of events per week). Instead, we apply the negative binomial distribution (Campbell 1982; Rao and Kaila 1986) with the ability to describe overdispersed count data. This distribution allows the variance to be greater than the mean value by adding a separate dispersion parameter (Gelman and Hill 2007) to adjust the width of the distribution. The dispersion parameter is estimated jointly with the other parameters based on the observed data, resulting in realistic uncertainty estimates.

Except for the autoregressive term, we use an exponential *link function* (Kruschke 2014; Gelman and Hill 2007) for the other predictors for two reasons. Firstly, the average seismic activity must be positive so it makes sense to fit a regression on the logarithmic scale following Kruschke (2014) and Gelman and Hill (2007). Secondly, the exponential link function will result in a multiplicative effect of the predictors on the average seismic activity, and the constant term (i.e. the intercept in linear regression) will act as an *exposure term* in relation to the other predictors (Gelman and Hill 2007). This means that the sensitivity of the measurement system (i.e. the probability of detecting an event of a certain magnitude) in the specific volume will act as an exposure term and will not affect the other estimands of interest and limit its role as a potential confounding variable, as opposed to using e.g. a linear link function. Also, discarding events below the magnitude of completeness threshold in our catalogue to compensate against differences in sensitivity would in our case lead to a significant loss in information and the usefulness of our data. Including all events in the analysis provides us with significantly more information for predicting the activity, which is the focus in this paper, but it has also been shown to give richer inference and significantly improved *b*-value estimates in terms of bias and uncertainty (Martinsson and Jonsson 2018).

The autoregressive term is of special interest here since it describes how long the current mining operation will be “remembered” by the orebody in the future. For example, the autoregressive term can be expressed as the seismic *half-life* and is here defined as the amount of time required for the activity to fall half ways to its steady state value if we would

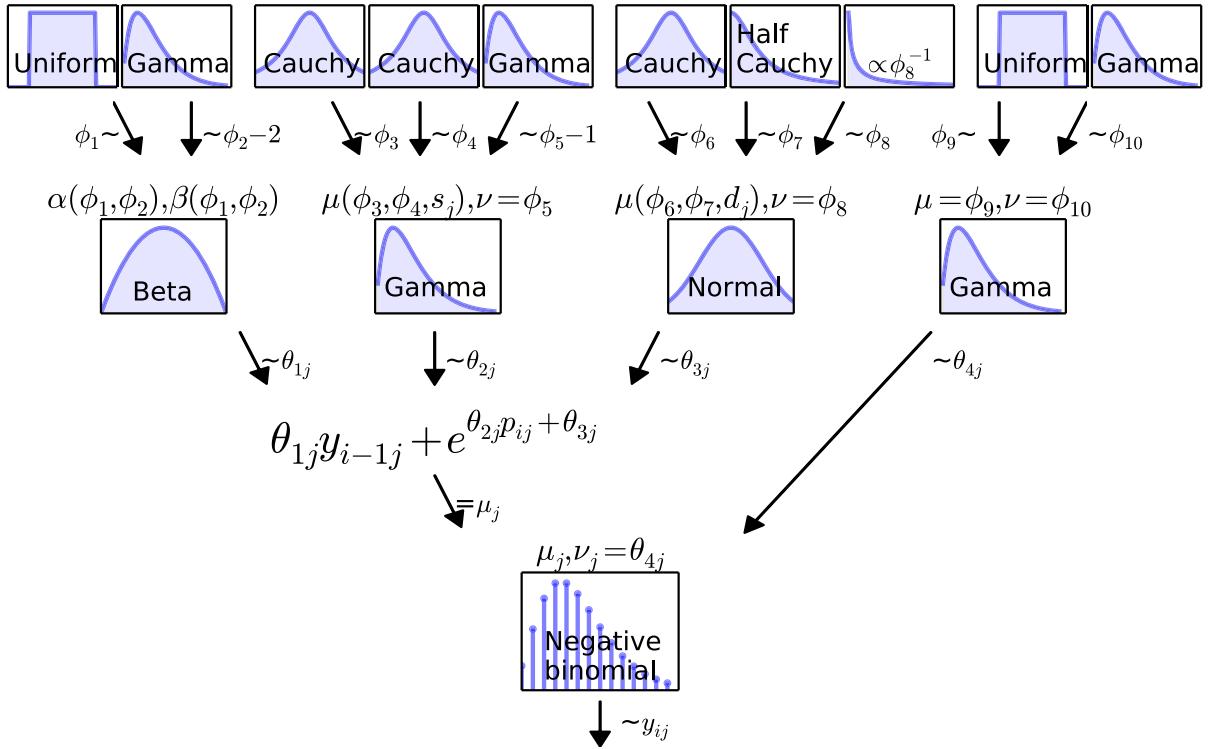


Figure 3

Overview of the proposed three-level hierarchical Bayesian model. The first level (bottom part) shows the individual model for each orebody. The second level (middle part) shows the prior distributions for the parameters at the first layer and describes the mine-wide distribution. The third level (top part) shows the hyperprior distributions for the parameters in the prior distributions at the second level

decrease the production by a certain amount at present time. Production in orebodies with long seismic half-life must be less frequent to avoid accumulated activity.

3.1. The Hierarchical Model Structure

A three-level hierarchical Bayesian model is needed for the data in this study and an overview of the model is given in Fig. 3. The graphical representation illustrates the complete hierarchy (following Kruschke 2014), and summarises the model in simple graphs and figures without extensive notations of the distributions and the dependencies between parameters. Figure 3 can directly be used in combination with open-source statistical packages, such as STAN (Carpenter et al. 2017) or JAGS (Plummer 2003), to implement the model we propose, with limited programming efforts (see e.g.

Kruschke 2014), for a introductory explanation on how to do it). The choice of model structure and distributions follow standard references in Bayesian hierarchical modelling (see e.g. Kruschke 2014; Gelman and Hill 2007; Gelman et al. 2004) and we also adopt standard notations to more easily match the proposed model structure with standard models found in these resources.

The first level in the hierarchy is the individual model for each orebody with model parameters denoted by θ . This model is simple, preserves physical constraints, and describes the activity in the orebody using four basic parameters: one constant, one parameter that describes the effect of past seismicity, one parameter that describes the effect of the planned production in the orebody, and one that governs the width (or dispersion or scale) of the distribution of activity. The second-level model describes the mine-wide characteristics and contains

the prior distributions for the parameters in level one. The prior parameters are often called hyperparameters (Gelman et al. 2004; Kruschke 2014) and are here denoted ϕ . The mine-wide model describes the behaviour of the group of orebodies in the studied mine and, in contrast to non-hierarchical models who's second-level parameters (hyperparameters) for the prior distributions are fixed, we use the data to estimate their values given group-level predictors. This means that the observed data will reveal similarities within the group of orebodies. The second-level mine-wide information are then shared by the individual models in the first level to obtain better estimates. Group-level predictors and similarities between the orebodies are of interest, not only for inference on a mine-wide scale, but essential for robust estimation in areas where the data is poor or observations are sparse (Gelman and Hill 2007). Below we present the proposed Bayesian hierarchical model in detail. The model includes the group-level relationships that we found: (1) agree with the individual observations (Sect. 4.1) and group-level behaviour (Sect. 4.2); (2) have intuitive group-level relationships (Sect. 4.2) and (3) generate good (Sect. 4.1 and Sect. 4.3) and valid predictions (Sect. 4.4).

In Sect. 5 we mention other options and additional parameters that did not result in modelling improvements for our data set.

3.1.1 Level 1: The Individual Model

The individual model at the first level describes the measured seismic activity y_{ij} (i.e. number of detected events per week) within the i th week and in the j th orebody. Given the model parameters θ_j and covariates, y_{ij} is assumed independently negative binomial distributed

$$p(y_{ij}|\theta_j, y_{i-1,j}, p_{ij}) = \frac{\Gamma(v_j + y_{ij})}{y_{ij}! \Gamma(v_j)} \left(\frac{v_j}{v_j + \mu_{ij}} \right)^{v_j} \times \left(\frac{\mu_{ij}}{v_j + \mu_{ij}} \right)^{y_{ij}}, \quad (1)$$

$$\mu_{ij} = \theta_{1j} y_{i-1,j} + \exp(\theta_2 p_{ij} + \theta_3), \quad (2)$$

$$v_j = \theta_4, \quad (3)$$

during $i = 1, 2, \dots, n_j$ observed weeks, in $j = 1, 2, \dots, J$ orebodies. The covariates are the measured seismic activity from the prior week $y_{i-1,j}$ and the planned accumulative rock mass in megatonne (Mt) retrieved by the production blasts p_{ij} during the current week. The values of p_{ij} is not measured but based on planned production. In Eq. (1) we adopt the standard reparametrisation of the negative binomial distribution in terms of the mean value μ_{ij} and dispersion parameter v_j , that is suitable for regression (Hilbe 2011). The model parameter θ_{1j} describes the seismic decay, θ_2 the multiplicative effects of the production, θ_3 the seismic exposure term and $v_j = \theta_4$ the dispersion parameter.

The seismic decay can also be characterised by the *half-life*

$$\tau_j = -\log(2)/\log(\theta_{1j}), \quad (4)$$

and is defined here as the amount of time required for the activity to fall (or rise) halfway to its steady state value, if we would decrease (or increase) the production by a certain amount at the present time. If θ_{1j} is close to one, the half-life is very long and the j th orebody will accumulate past seismic activity. See Appendix for more details.

The expected value in (2) is primarily applicable to orebodies subjected to active mining. In this state the activity will approach a steady-state level (or background seismicity) rather than zero activity in periods with no production, see Appendix. We discuss possible model extensions in Sect. 5 to additionally include transitions from active mining to post mining.

3.1.2 Level 2: The Mine Wide Model

The second-level model describes the group of orebodies and is the prior distributions for the individual model parameters θ in (1) given a set of hyperparameters ϕ .

The autoregressive parameter θ_{1j} , given the hyperparameters ϕ_1 and ϕ_2 , is assumed independent identically beta distributed

$$p(\theta_{1j}|\phi_1, \phi_2) = \text{Beta}(\alpha, \beta), \quad (5)$$

$$\alpha = \phi_1(\phi_2 - 2) + 1, \quad (6)$$

$$\beta = (1 - \phi_1)(\phi_2 - 2) + 1, \quad (7)$$

with the restricted support $\theta_{1j} \in [0; 1[$ to guarantee stability using autoregression (see e.g. Porat 1997). Here we use the more intuitive parametrisation with the mode $\phi_1 = (\alpha - 1)/(\alpha + \beta - 2)$ and the samples size $\phi_2 = \alpha + \beta$ (see e.g. Kruschke 2014), where Beta(α, β) is the standard parametrisation of the beta distribution.

The effect of the production is described by θ_{2j} and is non-negative and gamma distributed

$$p(\theta_{2j}|\phi_3, \phi_4, \phi_5, s_j) = \frac{1}{\Gamma(v)\theta_{2j}} \left(\frac{v\theta_{2j}}{\mu} \right)^v \exp \left(-\frac{v\theta_{2j}}{\mu} \right), \quad (8)$$

$$\mu = \exp(\phi_3 + \phi_4 s_j), \quad (9)$$

$$v = \phi_5, \quad (10)$$

where the mean effect shows an exponential behaviour with respect to the size of the production s_j . The size s_j is here based on the *average* weekly production rate in the orebody, following Gelman and Hill (2007), but similar results are obtained using the production area. The exponential relationship (i.e. a regression in log-scale) makes sense as the mean effect μ must be positive (Gelman and Hill 2007; Gelman et al. 2004).

The constant term θ_{3j} in (2), or the *exposure* (Gelman and Hill 2007), is normally distributed

$$p(\theta_{3j}|\phi_6, \phi_7, \phi_8, d_j) = N(\phi_6 + \phi_7 d_j, \phi_8), \quad (11)$$

where the mean depends linearly on the production depth.

As size s_j and depth d_j are nearly constants during the studied time frame, it makes sense to use their mean values as group-level predictors in the model (see e.g. Kruschke 2014; Gelman and Hill 2007).

The exponential link-function in (2) means that the dependency on depth in θ_{3j} will have a positive exponential effect on the average seismicity (2) and follows the results obtained by Vallejos and McKinnon (2011).

The last parameter is the non-negative dispersion term θ_{4j} in (3) and is gamma distributed

$$p(\theta_{4j}|\phi_9, \phi_{10}) = \frac{1}{\Gamma(v)\theta_{4j}} \left(\frac{v\theta_{4j}}{\mu} \right)^v \exp \left(-\frac{v\theta_{4j}}{\mu} \right), \quad (12)$$

$$\mu = \phi_9, \quad (13)$$

$$v = \phi_{10}. \quad (14)$$

3.1.3 Level 3: The Distributions for the Mine-Wide Parameters

The data consist of only seven individuals in the group and some regularisation is needed in the structural level to obtain stable inference (Gelman and Hill 2007). We follow the hyperprior design for ϕ_1, ϕ_2 in (5) proposed by Kruschke (2014), and we put a uniform hyperprior distribution for the mode ϕ_1 and a gamma distribution for the dispersion $\phi_2 - 2$ with mean 1 and standard deviation 5. Because the value of $\phi_2 - 2$ must be non-negative, the prior distribution on $\phi_2 - 2$ must not allow negative values (Kruschke 2014).

For the regression hyperparameters $\phi_3, \phi_4, \phi_6, \phi_7$ we consider weakly-informative Cauchy hyperprior distributions proposed in Gelman et al. (2008), where the effect of depth ϕ_7 is restricted to be positive (Vallejos and McKinnon 2011) using a half-Cauchy prior distribution (Kruschke 2014) and illustrated in Fig. 3.

The dispersion parameter $\phi_5 + 1$ is gamma distributed and restricted to be non-negative to force zero probability density at zero. The motivation for this is that we know that the production must have a positive effect on the observed seismicity. With this restriction the resulting hyperprior distribution matches both the independent non-hierarchical and the hierarchical estimation result. For the standard deviation ϕ_8 in the normal regression model in (11) we consider the standard non-informative prior $p(\phi_8) \propto \phi_8^{-1}$ (see e.g. Gelman et al. 2004; Kruschke 2014; Gelman and Hill 2007). The last dispersion parameter ϕ_{10} is gamma distributed (following Kruschke 2014).

3.1.4 The Joint Posterior Distribution

The joint posterior distribution is given by collecting the therm in the three levels under the assumption of

independence and conditional independence (see e.g. Kruschke 2014; Gelman et al. 2004; Gelman and Hill 2007). Starting with the hyperprior distributions in level 3 at the top layer in Fig. 3, and moving down towards the prior distributions in level 2, and finally ending with the likelihood in level 1, the joint posterior can be expressed as

$$p(\boldsymbol{\beta}|\mathbf{d}) \propto p(\boldsymbol{\phi}) \prod_{j=1}^J p(\boldsymbol{\theta}_j|s_j, d_j) \prod_{i=1}^{n_j} p(y_{ij}|\boldsymbol{\theta}_j, y_{i-1,j}, p_{ij}), \quad (15)$$

where \mathbf{d} represents both the measured data and covariates, and $\boldsymbol{\beta}$ is a 38 dimensional parameter vector containing all parameters in the joint model. The hyper-prior (level 3) and prior distributions (level 2) are given by

$$p(\boldsymbol{\phi}) = \prod_{k=1}^{10} p(\phi_k), \quad (16)$$

$$\begin{aligned} p(\boldsymbol{\theta}_j|s_j, d_j) &= p(\theta_{1j}|\phi_1, \phi_2)p(\theta_{2j}|\phi_3, \phi_4, \phi_5, s_j) \\ &\times p(\theta_{3j}|\phi_6, \phi_7, \phi_8, d_j)p(\theta_{4j}|\phi_9, \phi_{10}), \end{aligned} \quad (17)$$

respectively, and summarises the terms in levels 2 and 3. As opposed to standard non-hierarchical Bayesian models we have here an additional term with hyperprior distributions $p(\boldsymbol{\phi})$.

4. Results

In this section we present inferential results obtained from the model. The results mainly address our aim previously declared in Sect. 1. The model's predictability is evaluated in Sect. 4.1 and general results are initially discussed based on visual comparisons. A more detailed analysis of the specific effects of the different predictors are later presented in Sect. 4.2 along with estimates of the seismic half-life. Finally, the model's ability to predict new areas are demonstrated in Sect. 4.3.

4.1. Predicting the Activity

Figures 4 and 5 show the proposed model's weekly predictions of the induced seismicity in the seven studied orebodies. The predictions shown in

red summarises the posterior predictive distributions (see e.g. Gelman and Hill 2007; Gelman et al. 2004; Kruschke 2014) and are compared against the actual observations shown as black dots. The green dashed curve represents the weekly production in the orebody and is used as a predictor in the model. The red line shows the median prediction and the light red and dark red regions are the 50% and 95% prediction intervals, respectively. The number of observations inside the prediction intervals matches the specified percentages fairly well. A more detailed evaluation of the model's predictability is given later in Sect. 4.1 using cross-validation (Söderström and Stoica 1989) and Sect. 4.4 when validating the model. A few interesting observations can be made from a visual inspection:

- The effect of the production on the seismicity varies between the orebodies. The production in the second and the fifth orebody seems to correlate more to the observed seismicity, compared to the first, the third and the sixth orebody. There is only a short production interval in the fourth orebody.
- There are indications of seismic decays at intervals succeeding production stops. These are visible in the second orebody between weeks 60–85 and 140–165, and in the fifth orebody between weeks 125–150. Also, successive seismic buildups are visible as the production starts, e.g., in the fifth orebody after week 150 and in the seventh orebody after week 140. The seismic time constants associated with changes in weekly production are difficult to quantify visually, but indicate weeks rather than hours reported in Vallejos and Mckinnon (2010) or days reported in Olivier et al. (2015) when associated with single events or single blasts. We return to a more detailed analysis in Sect. 4.2 and Fig. 7.
- There are bursts of activity in the third and fourth orebody in Fig. 4 not directly driven by the current production. We discuss causes likely related to the applied mining method in Sect. 5.

4.2. Parameter Estimation

Estimation results for the individual parameters θ_{kj} in (2–3) and their corresponding prior

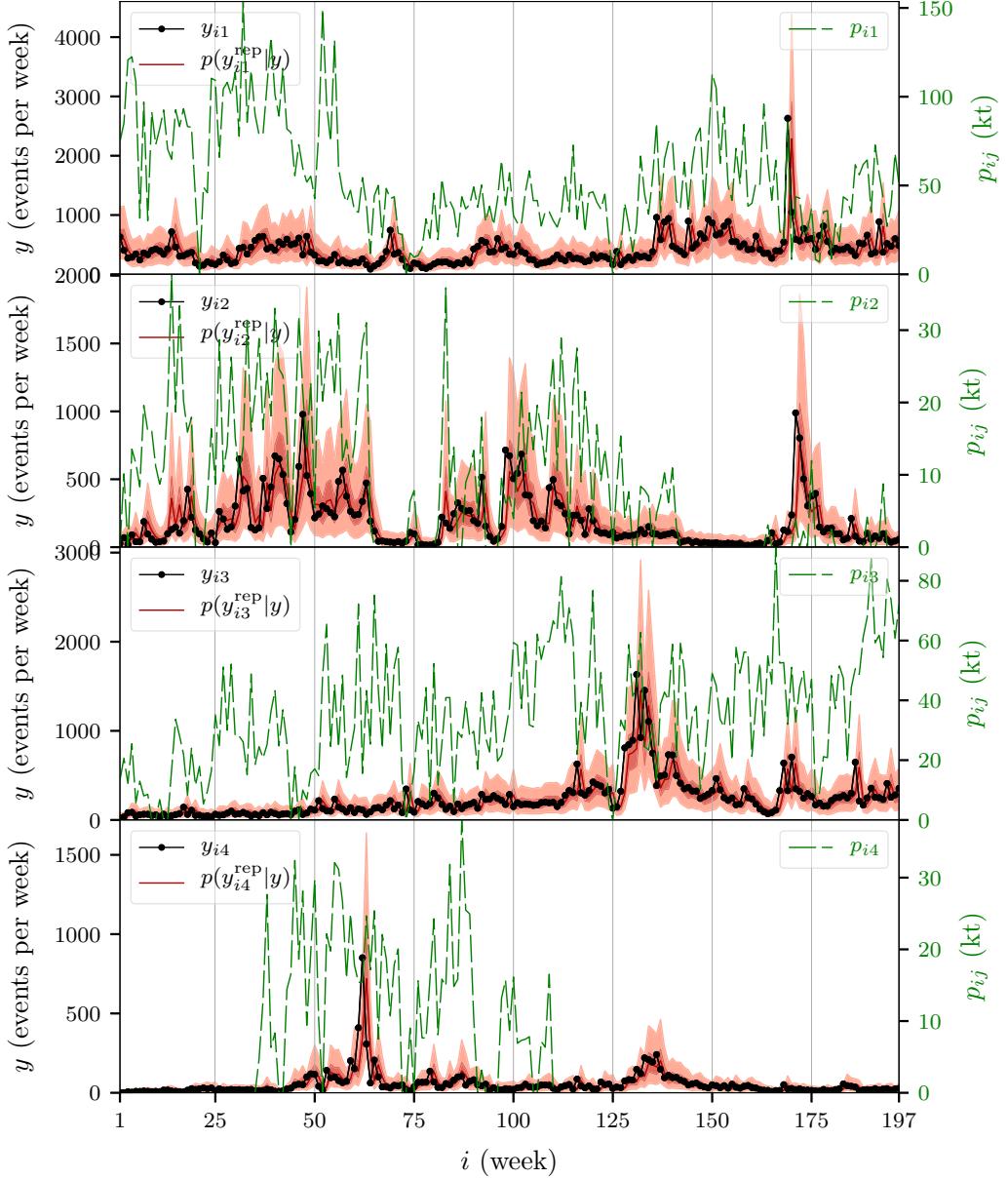


Figure 4

Weekly predictions of the induced seismicity in the first four ($j = 1, \dots, 4$) of our seven ($J = 7$) studied orebodies. The predictions summarises the posterior predictive distribution and are compared against the actual observations shown as black dots. The green dashed curve represents the planned weekly accumulative rock mass retrieved by the production blasts in the orebody, and is used as a predictor in the model. Note that the production is for visualisation proposes shown here in kilotonne (kt) rather than megatonne (Mt). The red curve shows the median prediction and the light red are dark red regions are the 50% and 95% prediction intervals, respectively

distributions (5–12) are shown in Fig. 6. The red squares $p(\theta_{kj}|y_j)$ are estimates of the parameters when treating the orebodies independently and the blue circles $p(\theta_{kj}|y)$ are estimates obtained using the

hierarchical model structure. The numbers $j = 1, \dots, 7$ inside the markers represent the seven orebodies in our data set. The corresponding bold and thin horizontal whiskers represent the 50% and

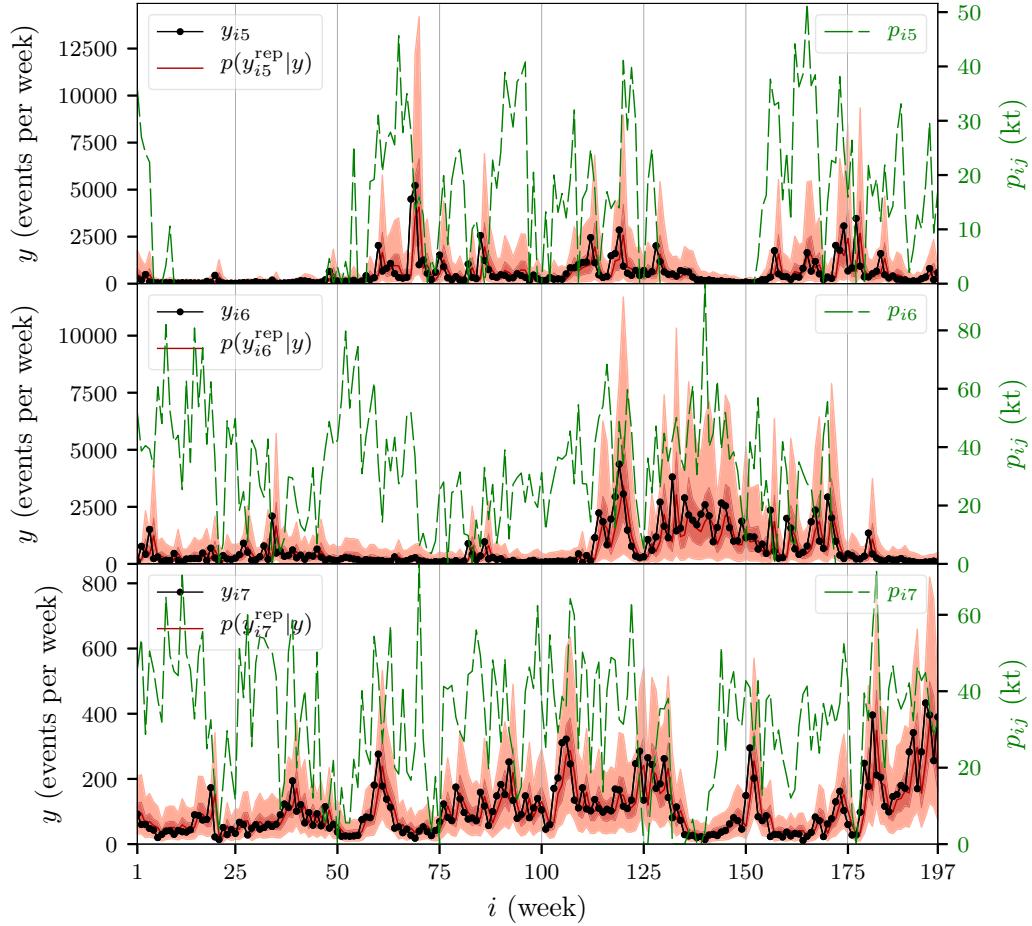


Figure 5

Weekly predictions of the induced seismicity in the last three ($j = 5, 6, 7$) of our seven ($J = 7$) studied orebodies. The predictions summarises the posterior predictive distribution and are compared against the actual observations shown as black dots. The green dashed curve represents the weekly accumulative rock mass retrieved by the production blasts in the orebody, and is used as a predictor in the model. Note that the production is for visualisation proposes shown here in kilotonne (kt) rather than megatonne (Mt). The red curve shows the median prediction and the light red and dark red regions are the 50% and 95% prediction intervals, respectively

95% credible intervals, respectively. The blue curves are the jointly estimated prior distributions for the parameters in the hierarchical model, and describes the group of orebodies (i.e. the mine-wide distribution). The green curves represent group-level regression curves in (9) and (11) given size and depth, respectively.

The first parameter θ_{lj} represent the seismic decay term (or the autoregressive parameter). The similarity of the seven orebodies means that the prior distribution in (5) contracts around its median value of 0.890 with the 95% interval (0.665, 0.985). Some shrinkage (with respect to both location and uncertainty)

towards the mine-wide distribution is observed when comparing the hierarchical estimated (blue circles) against the independent estimates (red squares). This rather small effect of the prior distribution is expected as the likelihood is much more informative compared to the similarity of the orebodies captured by the prior distribution. The whiskers spanning the credible intervals are considerably smaller than the estimated prior distribution. Figure 7 shows the same estimation results but viewed from the seismic half-life using the relationship in (4). The median half-life for the mine-wide distribution is 5.960 weeks and the 95% interval is (1.697, 45.556). We could not see

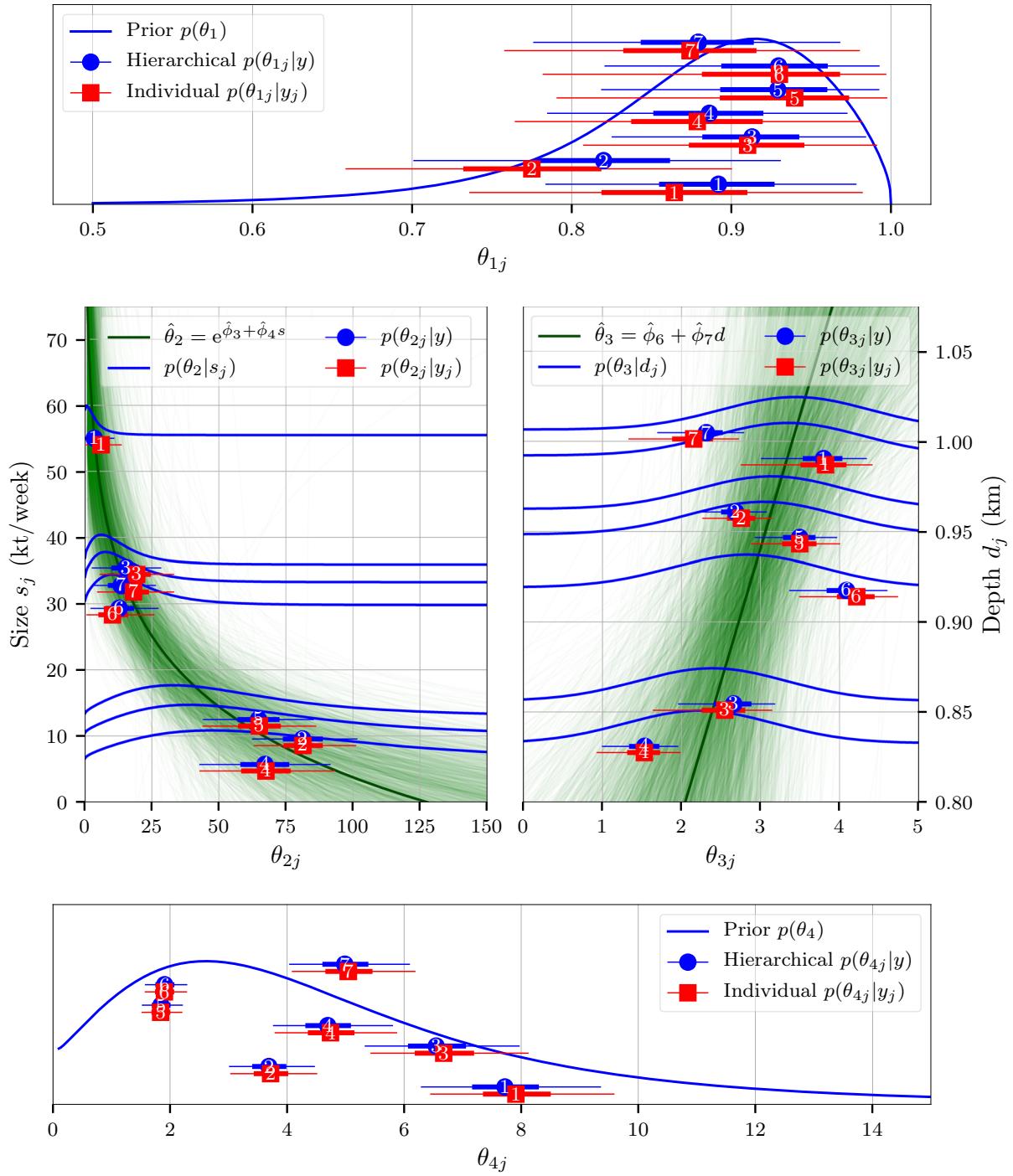


Figure 6

Estimation results for the individual parameters in (2) and (3) and their prior distributions defined in (5–12). The red squares are the independent estimates and the blue circles are estimates obtained using the hierarchical model structure for the seven orebodies. The corresponding bold and the thin horizontal whiskers represent the 50% and 95% credible intervals, respectively. The blue curves are the estimated prior distributions for the parameters the green curves represent group-level regression curves. The size s_j is for visualisation proposes shown here in kilotonne (kt/week) rather than megatonne (Mt/week)

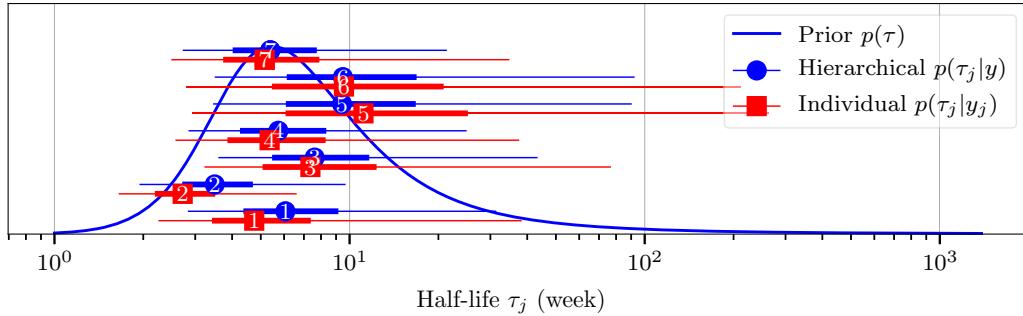


Figure 7
Estimation results for the seismic half-life τ_j using the relationship in (4)

any group-level dependence of the decay times with respect to depth (nor size) for the studied data, as was found in Vallejos and McKinnon (2011).

The second parameter θ_{2j} is the effect of the weekly accumulative rock mass retrieved by the production blasts in (2) on the activity. The rock mass is not based on actual measurements but on planned production. The exponential effect in (9), motivated in Sect. 3.1.2, seems reasonable and is shown as the green curves in Fig. 6. The bold green curve represent the median regression curve and the thin curves are credible regression curves generated using a subset of 1000 posterior samples (Sect. 2.2). Note also that the prior distributions in blue are more informed as the location and the dispersion follows the group-level relationship with respect to size s_j . Neglecting the dependency on size results in a much wider prior distribution describing all sizes.

The third parameter θ_{3j} in Fig. 6 is the exposure term in (2), Sect. 3.1.2, and depends linearly on depth d_j . The exponential link-function (2) means that depth will have a positive exponential effect on the average seismicity (Vallejos and McKinnon 2011). As with the previous parameter, the prior distributions in blue are more informed as the location of the normal distribution in blue follows the group-level relationship. The bold green curve represents the median regression line and, e.g., increasing the depth d_j by 100 m results in a factor 1.944 more events on average. This is in agreement with Vallejos and McKinnon (2011) reporting a factor 1.896 more events for the same increase in Creighton Deep (or more precisely $0.0001 \exp(0.0064Z)$ where Z is

depth). As indicated by the scattering of credible regression lines with different slopes in Fig. 6, the uncertainty of this effect is high and so is the corresponding 95% credible interval (1.046, 7.331) for this increase in depth. Access to only seven orebodies with depths relatively close to each other is one reason for this uncertainty.

The fourth parameter θ_{4j} in Fig. 6 is the dispersion term (3) and shows overdispersion with respect to a Poisson distribution. Similar to the autoregressive term, no group-level dependencies are observed.

4.3. Predicting the Activity in New Areas Using Cross-Validation

Except for inferring group-level dependencies, the Bayesian hierarchical model allows us to predict seismicity when data is sparse or even missing. In Fig. 8, cross-validation (Söderström and Stoica 1989) is applied where one orebody is removed from the data set and we use the similarity information of the other orebodies to predict its activity from start. We demonstrate this for the two most extreme orebodies, in both size and depth, as these two will be the most influential points in the group-level regressions in Fig. 6.

In Fig. 8, the red box represents this prediction and is compared against the blue circles that is the prediction when the orebody is included as in Figs. 4 and 5. The prediction of the first week is only based on the prior distributions and the similarity of the six other orebodies, using cross-validation. As the week progresses the uncertainties contracts as more data is

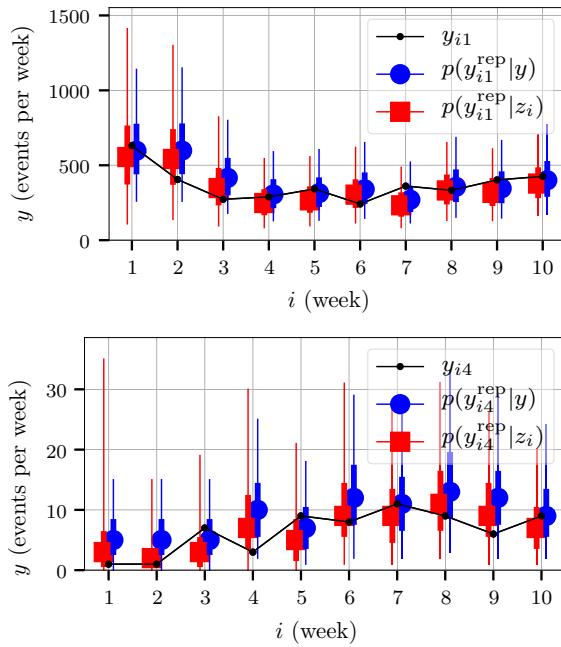


Figure 8

Predictions of weekly activities based on the similarity information of the other six orebodies. The top and bottom panel shows predictions of weekly activities for the largest and smallest orebody, respectively, using cross-validation. The red squares represent predictions when the data for the particular orebody is removed from the data set (and this new data set is denoted z_i). The blue circles are the prediction when the orebody is included in the data set and is identical to the prediction shown in Figs. 4 and 5. The black dots are the observed seismicity

available for this new orebody and the results approaches the predictions in Figs. 4 and 5, for which both types of predictions are identical when $i = 197$. The rapid convergence towards the results shown in Figs. 4 and 5 is expected as each individual orebody has only four parameters and after only a few weeks we have enough data to estimate the individual four parameters without the otherwise necessary support from the prior distributions provided by the similarity information of the other orebodies.

The small systematic differences seen between the two types of predictions reduces as more weeks progress but needs more data from this new orebody to vanish. The reason for this difference is that these are the two most extreme orebodies in the group and therefore the most influential points in the group-level regressions in Fig. 6, and we need more data to influence the group-level relationships, compared to

the data that is available for the other orebodies. In a hierarchical Bayesian model, the influence of the added data from the new individual is weighted both with respect to sample size n_j and with respect to the number of individuals in the group J and their corresponding sample sizes (Kruschke 2014; Gelman et al. 2004; Gelman and Hill 2007).

4.4. Model Validation

Model validation is a central part in justifying the proposed model (Kruschke 2014; Gelman et al. 2004; Pintelton and Schoukens 2001). In order to trust inference obtained using the model, the predicted data generated using the same covariates should look similar to the actual observations in Figs. 4 and 5 (Gelman et al. 2004). In the model validation step we try to detect any model deficiencies that may lead to false inferences about parameters or predictions of interest.

We start in Table 1 by evaluating the number of actual observations that are inside the 50% and 95% prediction intervals shown in Figs. 4 and 5 as light red and dark red regions. We expect about half of the 50% intervals and about 5% of the 95% intervals to exclude the true measured value (Gelman et al. 2004) shown as black dots in Figs. 4 and 5. We conclude that the fractions in Table 1 of observations inside these prediction intervals are realistic. None out of the fourteen intervals, or 0.0 (0.0, 21.8)%, contains

Table 1
Evaluation of the 1-week-ahead prediction intervals in Figs. 4 and 5

Orebody	In 95% region	In 50% region
$j = 1$	0.970 (0.935, 0.986)	0.513 (0.443, 0.582)
$j = 2$	0.964 (0.929, 0.982)	0.487 (0.418, 0.557)
$j = 3$	0.964 (0.929, 0.982)	0.487 (0.418, 0.557)
$j = 4$	0.964 (0.929, 0.982)	0.543 (0.473, 0.611)
$j = 5$	0.970 (0.935, 0.986)	0.569 (0.499, 0.636)
$j = 6$	0.980 (0.949, 0.992)	0.447 (0.379, 0.517)
$j = 7$	0.970 (0.935, 0.986)	0.467 (0.399, 0.537)

The table shows the fraction of observations inside the 95% and inside the 50% prediction intervals. The numbers in the parentheses are 95% credible intervals of this fraction. All fractions are within the specified uncertainty

Table 2

Evaluation of the prediction intervals based on replicated data without access to past measurements

Orebody	In 95% region	In 50% region
$j = 1$	0.964 (0.929, 0.982)	0.533 (0.463, 0.601)
$j = 2$	0.970 (0.935, 0.986)	0.503 (0.433, 0.572)
$j = 3$	*0.893 (0.842, 0.929)	0.477 (0.408, 0.547)
$j = 4$	0.949 (0.909, 0.972)	0.569 (0.499, 0.636)
$j = 5$	0.964 (0.929, 0.982)	0.548 (0.478, 0.616)
$j = 6$	0.970 (0.935, 0.986)	*0.416 (0.350, 0.486)
$j = 7$	0.964 (0.929, 0.982)	0.503 (0.433, 0.572)

The table shows the fraction of observations inside the 95% and inside the 50% prediction intervals. The numbers in the parentheses are 95% credible intervals of this fraction. Two out of fourteen (marked with *) gives a fraction with a corresponding 95% credible interval that do not contain the specified uncertainty

more or less observations than the interval was specified to do. This matches the predefined risk of 5%, considering multiple comparisons (see e.g. Gelman and Hill 2007).

Good short-term forecasts generally rely on recent measurements. Figures 4 and 5 showed weekly predictions given measurements of the seismicity the week before $y_{i-1,j}$ in (2), along with the other covariates explained in Sect. 3.

Longer range forecasts (e.g. 2 weeks, 3 weeks, months, years) from the currently known state are less certain. The most extreme case is predictions without any recent measurements at all. Table 2 shows an evaluation of the prediction intervals for this extreme case, using the same parameter estimates and covariates as in Figs. 4 and 5 but without access to any observations. The observed value of the seismicity the week before $y_{i-1,j}$ in (2) is replaced with a replicated (or simulated) value the week before. To trust the model for long-term predictions the corresponding prediction intervals should be realistic. One out of fourteen intervals in Table 2 contain more observations than specified, and we draw the same conclusion as for Table 1 above.

To further test the validity of the model we use posterior predictive p-value (Gelman et al. 2004) for various discrepancy measures that are important for our inferential goals. Although criticised (see e.g. Kruschke 2013, 2014) it still gives a hint of possible

discrepancies between the model and the observations if we take multiple comparisons into account (Gelman 2013).

Based on 100,000 predicted data sets, given the same covariates as in the observations, the posterior predictive p-vales of different measures are shown in Table 3. Five out of 70 measures, or 7.1 (3.2, 15.7)%, are more extreme and matches a predefined risk of 5% (Gelman 2013).

We also evaluate 100,000 replicated data sets without access to the observed value of the seismicity the week before $y_{i-1,j}$ in (2). The posterior predictive p-vales of different measures are shown in Table 4. One out of 70 measures, or 1.4 (0.3, 7.6)%, is more extreme and matches a predefined risk of 5%.

Based on the evaluation of the prediction intervals and the posterior predictive p-values of predicted and replicated data, there is little evidence to indicate significant discrepancies and the actual observations are typical of the replicated observations generated using the proposed model (Gelman et al. 2004).

5. Discussion

Good models are essential to understand the relationship between induced seismicity and the production rate in the mine. We need them to: make predictions of induced seismicity, estimate the effects of different predictors, customise production plans that limits induced seismicity. In this study we present a Bayesian hierarchical model that captures the dynamic relationships and group-level dependencies. We found no evidence that would indicate model deficiencies based on posterior predictive checks on standard discrepancy measures, and the actual observations are typical of the replicated observations generated using the model. The prediction intervals are realistic and the fraction of observations within them matches their predefined probability.

We found that the seismic half-life ranges from weeks up to 2 months for the studied orebodies and a visual examination of the observations agrees with these findings. As opposed to sequences after a large event or a single blast, that shows decay times of hours reported in Vallejos and Mckinnon (2010) or several days reported in Olivier et al. (2015), we

Table 3

Posterior predictive p-values (Gelman et al. 2004) of different discrepancy measures of predicted data

Orebody	P _{2.5}	P ₂₅	P ₅₀	P ₇₅	P _{97.5}
j = 1	0.084	0.390	0.320	0.724	0.947
j = 2	0.346	*0.980	0.836	0.704	0.901
j = 3	0.231	0.858	0.276	0.582	0.560
j = 4	0.199	0.338	0.733	0.957	0.708
j = 5	0.044	0.171	0.124	0.811	0.797
j = 6	*0.003	0.667	0.967	0.073	0.910
j = 7	0.099	0.862	0.392	0.424	0.765

Orebody	Min	Max	Mean	Std	Iqr
j = 1	0.156	0.350	0.793	0.771	0.771
j = 2	0.065	0.924	0.810	0.906	0.545
j = 3	0.419	0.616	0.629	0.661	0.433
j = 4	0.837	0.373	0.728	0.490	*0.978
j = 5	*0.011	0.831	0.907	0.911	0.866
j = 6	*0.000	0.972	0.757	0.940	0.056
j = 7	0.221	0.934	0.648	0.813	0.296

Here P_x is the x th percentile, std the sample standard deviation and $iqr = P_{75} - P_{25}$ is the interquartile range. Five out of 70 measures (marked with *) are more extreme

Table 4

Posterior predictive p-values (Gelman et al. 2004) of different discrepancy measures of replicated data

Orebody	P _{2.5}	P ₂₅	P ₅₀	P ₇₅	P _{97.5}
j = 1	0.214	0.466	0.540	0.709	0.881
j = 2	0.539	0.904	0.751	0.680	0.834
j = 3	0.575	0.756	0.509	0.575	0.513
j = 4	0.367	0.484	0.684	0.846	0.711
j = 5	0.210	0.513	0.504	0.774	0.853
j = 6	0.039	0.608	0.796	0.413	0.688
j = 7	0.368	0.779	0.574	0.557	0.703

Orebody	Min	Max	Mean	Std	Iqr
j = 1	0.214	0.390	0.727	0.777	0.786
j = 2	0.187	0.884	0.767	0.818	0.598
j = 3	0.623	0.409	0.577	0.494	0.501
j = 4	0.846	0.240	0.704	0.492	0.900
j = 5	0.059	0.850	0.852	0.873	0.792
j = 6	*0.001	0.826	0.629	0.719	0.384
j = 7	0.366	0.811	0.657	0.697	0.476

The data was replicated without access to the observed value of the seismicity the week before. Here P_x is the x th percentile, std the sample standard deviation and $iqr = P_{75} - P_{25}$ is the interquartile range. One out of 70 measures (marked with *) are more extreme

model the accumulated retrieval of rock mass during an entire week that is associated with several production blasts. The larger decay times found in this data set may partly be explained by a much larger redistribution of stresses compared to a single event or a single blast. For example, the weekly excavated volume in the largest orebody ($j = 1$) is 1600 times larger compared to the blasted volume (20 m^3) in the mining tunnel investigated in Olivier et al. (2015).

The different caving processes associated with the applied mining method (sub-level caving, Wettainen and Martinsson 2014) may be another explanation for the larger decay times and visible bursts of seismicity not directly correlated with the production. In sub-level caving, the hangingwall gradually caves into the excavated volume as the production continuous (see e.g. Svartsjaern 2018; Svartsjaern et al. 2016). This rather complex caving processes are known to progress over weeks (or up to years for crown pillars), and depend on size, location, geometry and geology of the hangingwall (Malmgren, private communication, 2017). These caving mechanism often affects large volumes and may be present in several seismic volumes at the same time, and may explain the correlated burst, seen e.g. in Fig. 4 for the two neighbouring orebodies and volumes $j = 3$ and $j = 4$. As the caving moves upwards, it could also affect neighbouring seismic volumes at slightly different times as may be the case for $j = 3$ and $j = 4$. These mechanisms are not directly related to the current production but a result of decades of mining, hence there is no direct correlation to the current production term in Figs. 4 and 5. Bursts of seismicity that is not directly caused by the recent production, may also be a result of internal interrelations among seismic events (Kijko and Funk 1996) that can modify future activity rate (Orlecka-Sikora 2010). Additional predictors retrieved from the seismic system, such as coseismic stress changes (Orlecka-Sikora 2010), may improve the model's predictability.

The model we propose models the seismic behaviour of the orebodies under active mining. In this state the activity will approach a steady-state level when there is no production rather than zero activity, see Appendix for more details. This can also be seen in our data in Fig. 4 and for $j = 4$. At the end of this data there is no production for a period of more than

20 weeks and the activity level is not decaying to zero. When mining stops, the caving mechanism in the mine-out areas will continue for many years to come and for some orebodies the caving will eventually reach the surface as subsidence (Malmgren, private communication, 2017) and is visible in Fig. 1 for the orebodies: Kapten, Parta, Dennewitz and Alliansen. We could extend the model and include a diminishing term to reduce the steady-state level, i.e. $\exp(\theta_3)$, over time with an additional parameter to govern the rate of it (i.e. a time constant). However, in this study we only have observations during active mining and we have no information of the post-mining seismic behaviour in our material. Hence, we have no information to estimate such a parameter (i.e. time constant). Modelling and studying a transition from an active to passive mining state is not our objective. Modelling it would require observations from such a transition together with longer observation time for sub-level caving.

Another finding was that the effect of the production on the induced seismicity depends exponentially on the size of the production. The weekly production has a larger direct effect in smaller orebodies with lower average production, and this effect decreases exponentially with production size. As noted from the visual comparison between the measured seismicity and the production in Sect. 4.1, the production in the second and the fifth orebody seems to correlate more to the observed seismicity, compared to the first, third and the sixth orebody. This agrees well with the estimation results shown in Fig. 6 for the production parameter θ_{2j} and its group-level dependency on size s_j .

The size s_j is estimated based on the average planned production rate in the orebody, and we may question the correctness of this information in the future, as planned production p_{ij} may change and affect the average production rate s_j . For long term predictions this may be an issue if the actual production deviates heavily from planned production. However, for short-term predictions (e.g. predicting the activity next week) this is not an issue as an estimate can easily be obtained by looking at historical data (either planned or measured if available). For long-term predictions we may incorporate uncertainties associated with both p_{ij} and s_j that may

be based on historical evaluations of deviations between planned and actual production. A straightforward Monte-Carlo implementation is to base our predictions on not just one specific production sequence, as we did in Figs. 4 and 5, but on many credible production realisations taking these uncertainties into account (following the discussion in Mishra et al. 2017).

The seismic exposure term reveals a dependency on depth. For example, an increase in depth by 100 m doubles the seismic activity. Although rather uncertain, the effect agrees well with the findings in Creighton Deep (Vallejos and McKinnon 2011). The uncertainty in our estimate is likely to be caused by only seven orebodies at relatively similar depths.

In the analysis we have mainly followed Kruschke (2014), Gelman et al. (2004, 2008) and considered either non-informative, vague or weakly informed prior distributions for the hyperparameters. However, one strength with Bayesian analysis is to use informative priors and reuse findings from other research in the current analysis. This is especially fruitful when the accessible data is not informative enough. In our case, we could have included the depth dependencies found in Vallejos and McKinnon (2011) and assigned a more informative hyperprior distribution for ϕ_7 centred around the value estimated in Vallejos and McKinnon (2011) and with a scale reflecting reasonable uncertainties. This would not only reduce the uncertainties in our estimate of ϕ_7 to more realistic intervals, but also contribute to more certain estimates of other parameters (Kruschke 2014) and more certain predictions when data is sparse (Sect. 4.3).

More autoregressive terms (past week, past 2 weeks, etc.) have been included in the evaluation process but without any significant modelling improvements for the studied material. Different combinations of the proposed group-level predictors and additions of other predictors (e.g. the angle towards the main stress field, the dip of the orebody, the mayor horizontal length, the horizontal length perpendicular to the main stress field, the size of the hanging wall, and other size and volume measures) have been evaluated in this study but without significant modelling improvements. Also, an alternative production term has been evaluated that is a weighted

combination of the production rate in all orebodies, where the weights are the inverse squared distance from the production point to the centre of the volume, but without improvements. Knowledge of rock mechanic properties for the specified volumes are currently limited to only a few locations in a small subset of the investigated volumes. The limited information and absence of prior knowledge of such effects means that we can not test or include it as group-level predictors.

A possible extensions to the model would be to incorporate additional autoregressive terms from the neighbouring orebodies. For instance, the first and the second orebody are spatially close and share a common burst around week $i = 171$ in Fig. 4 and such an extension may improve inference. However, covariation in other time intervals are absent. The third and fourth orebody are also spatially close and share a common burst around week $i = 130$, but also seem to lack covariation elsewhere. The lack of a more consistent covariation in our material speaks against such modelling extension. For other mines that show clear covariation and have spatially closer volumes, orebodies or blocks, this type of model extension may improve inference.

No major events can be found in the vicinity of the volumes for the time periods specified above, that may have explained the common bursts. Also, the decay times for these bursts are in the order of weeks as opposed to hours or several days for major events (Vallejos and Mckinnon 2010; Olivier et al. 2015). However, the time periods are during or just after the snow melting period that is known to increase seismicity (Ersholm, private communication, 2017).

The reported inference (e.g. the predictability, the specific effects of the covariates, group-level dependencies, etc.) is limited to the specific study setting and the mine under investigation. However, the proposed Bayesian hierarchical model structure and the procedures used here are applicable in evaluating, analysing and understanding the effects of similar settings. Size and depth are clearly important group-level predictors but additional predictors that are missing in this study, might improve the model's predictability and reduce uncertainty.

The time steps investigated here are weeks, but could be reduced to days or even hours depending on

the detailedness of the data. For hourly predictions, more autoregressive terms might be needed to describe seismic decays from both single events or blasts (Vallejos and Mckinnon 2010) and decays from daily or weekly productions (Fig. 7). For production plan optimisations, which is the objective here, time steps in weeks are appropriate.

We also consider the same seismically active volumes used in Wettainen and Martinsson (2014) in this study. These volumes can be redefined or divided into sub-volumes to increase the detailedness in specific areas of interest. Dividing the larger volumes and corresponding production areas into smaller sizes might also increase the effect of the production term on the induced seismicity in a smaller nearby sub-volume, Fig. 6. This, however, means that the resulting sub-volumes are spatially closer and dependencies between neighbouring volumes should be investigated and included in the model.

The magnitude information of the events and the corresponding b -values for the volumes are not investigated in this study. It is an ongoing research (see e.g. Martinsson and Jonsson 2018) that combined with the results presented here can be used for more comprehensive probabilistic earthquake hazard assessment in the mine.

6. Conclusion

We present a Bayesian hierarchical model that captures the dynamic relationships between seismic activity and production rate, depth and size of the orebodies. The induced seismic behaviour of seven orebodies in the same iron-ore mine are modelled jointly using hierarchical levels, where the mine-wide parameters are learnt together with the individual orebody parameters from the observed seismic data. Model validation with posterior predictive checking using several discrepancy measures could not detect any model deficiencies or flaws. The weekly prediction with corresponding intervals are realistic and the fraction of observations within the intervals matches their predefined probability. The model can be used to evaluate the short and long-term impact of a production plan on the induced seismicity in a specific orebody. Inference from the model reveals that the

seismic half-life ranges from weeks up to 2 months for the studied orebodies. The effect of the weekly production on the induced seismicity depends exponentially on the average production size in the orebody. The seismic exposure term reveals a dependency on depth and, for example, an increase in depth by 100 m doubles the seismic activity. Except for inferring group-level dependencies on depth and size, we show that the Bayesian hierarchical model can predict seismicity when data is sparse or even missing.

Acknowledgements

Open access funding provided by Lulea University of Technology. The authors acknowledge the financial support from Luossavaara-Kiirunavaara AB (LKAB). The authors also thank Fredrik Ersholm at LKAB for valuable discussions during this work and Lars Malmgren at LKAB for useful suggestions and comments on this manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Steady State Level, Exponential Decay and Seismic Half-Life

The expected steady state level of the activity in (2) under constant production $p_{ij} = p_0$ is given by

$$\mu_{ij} = \exp(\theta_2 p_0 + \theta_3 j) / (1 - \theta_1 j). \quad (18)$$

Assume that for week $i = 0$ the expected activity is at this steady state level, i.e. $\mu_{0j} = \exp(\theta_2 p_0 + \theta_3 j) / (1 - \theta_1 j)$, and that the production is suddenly increased one unit at week $i = 1$ as $p_{ij} = p_0 + \delta_{i1}$, where

$$\delta_{i1} = \begin{cases} 0 & \text{if } i \neq 1, \\ 1 & \text{if } i = 1, \end{cases} \quad (19)$$

is the Kronecker delta function. The impulse response of such scenario is given by

$$\mu_{ij} = \mu_{0j} + \theta_{1j}^i C, \quad (20)$$

where $C = \mu_{0j} \left(1 + \theta_{1j}^{-1} (\exp(\theta_2 j) (1 - \theta_{1j}) - 1) \right)$ is a constant term independent of the week index i . The impulse response resulting from the increased production at week $i = 1$ decreases exponentially with respect to i towards the steady state level μ_{0j} . The half-life of the second term in (20) is determined by $\theta_{1j}^\tau = 1/2$ and leads to the seismic half-life in (4) under the stability constraints provided by the beta prior distribution in (5).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Bachmann, C. E., Wiemer, S., Woessner, J., & Hainzl, S. (2011). Statistical analysis of the induced basel 2006 earthquake sequence: Introducing a probability-based monitoring approach for enhanced geothermal systems. *Geophysical Journal International*, 186(2), 793–807.
- Bergman, S., Kibler, L., & Martinsson, O. (2001). Description of regional geological and geophysical map of northern Norrbotten county (east of the Caledonian Orogen). *Geological survey of Sweden Ba*, 56(110)
- Burridge, R., & Knopoff, L. (1967). Model and theoretical seismicity. *Bulletin of the Seismological Society of America*, 57(3), 341–371.
- Campbell, K. W. (1982). Bayesian analysis of extreme earthquake occurrences. Part I. Probabilistic hazard model. *Bulletin of the Seismological Society of America*, 72(1), 1689–1705.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., & Betancourt, M., et al. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1)
- Chen, M. H., Shao, Q. M., & Ibrahim, J. G. (2000). *Monte Carlo Methods in Bayesian Computation* (Springer Series in Statistics). Berlin: Springer.
- Debras, C. (2010). Petrology, geochemistry and structure of the host rock for the Printzsköld ore body in the Malmberget deposit. Master's thesis, Luleå University of Technology.
- Garcia-Aristizabal, A. (2018). Modelling fluid-induced seismicity rates associated with fluid injections: Examples related to fracture stimulations in geothermal areas. *Geophysical Journal International*, 215(1), 471–493.
- Gelman, A. (2013). Two simple examples for understanding posterior p-values whose distributions are far from uniform. *Electronic Journal of Statistics*, 7, 2595–2602.

- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian Data Analysis* (2nd ed.). London: Chapman and Hall/CRC.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models. Analytical methods for social research*. New York: Cambridge University Press.
- Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, 2(4), 1360–1383.
- Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge: Cambridge University Press.
- Hirata, T. (1987). Omori's power law aftershock sequences of microfracturing in rock fracture experiment. *Journal of Geophysical Research: Solid Earth*, 92(B7), 6215–6221.
- Kijko, A., & Funk, C. W. (1996). Space-time interaction amongst clusters of mining induced seismicity. *Pure and Applied Geophysics*, 147(2), 277–288.
- Kijko, A., & Graham, G. (1999). "Parametric-historic" procedure for probabilistic seismic hazard analysis part ii: Assessment of seismic hazard at specified site. *Pure and Applied Geophysics*, 154, 1–22.
- Kruschke, J. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, 142(2), 573–603.
- Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial introduction with R, JAGS and Stan* (2nd ed.). Amsterdam: Elsevier Science.
- Leptokaropoulos, K., Staszek, M., Lasocki, S., Martínez-Garzón, P., & Kwiatek, G. (2017). Evolution of seismicity in relation to fluid injection in the North-Western part of The Geysers geothermal field. *Geophysical Journal International*, 212(2), 1157–1166.
- Martinsson, J. (2012). Robust Bayesian hypocentre and uncertainty region estimation: The effect of heavy-tailed distributions and prior information in cases with poor, inconsistent and insufficient arrival times. *Geophysical Journal International*, 192, 1156–1178.
- Martinsson, J., & Jonsson, A. (2018). A new model for the distribution of observable earthquake magnitudes and applications to b-value estimation. *IEEE Geoscience and Remote Sensing Letters*.
- McGarr, A. (1976). Seismic moments and volume changes. *Journal of Geophysical Research*, 81(8), 1487–1494.
- Mena, B., Wiemer, S., & Bachmann, C. (2013). Building robust models to forecast the induced seismicity related to geothermal reservoir enhancement. *Bulletin of the Seismological Society of America*, 103(1), 383–393.
- Mishra, M., Martinsson, J., Rantatalo, M., & Goebel, K. (2017). Bayesian hierarchical model-based prognostics for lithium-ion batteries. *Journal of Reliability Engineering and System Safety*.
- Mogi, K. (1987). A generation mechanism of matsushiro earthquake swarm and its relation to the Niigata earthquake (in Japanese). *Abstract Seismological Society of Japan*, 2, 97.
- Mogi, K. (1991). Recent irreversible crustal activity in east Izu peninsula (in Japanese). *Chikyu Monthly*, 13, 55–62.
- Neal, R. M. (2003). Slice sampling. *Annals of Statistics*, 31(3), 705–767.
- Ogasawara, H. (2002). A simple analogue experiment to account for power-law and exponential decays of earthquake sequences. *Pure and Applied Geophysics*, 159(1), 309–343.
- Olivier, G., Brenguier, F., Campillo, M., Roux, P., Shapiro, N. M., & Lynch, R. (2015). Investigation of coseismic and postseismic processes using in situ measurements of seismic velocity variations in an underground mine. *Geophysical Research Letters*, 42(21), 9261–9269.
- Omori, F. (1894). On the aftershocks of earthquakes. *Journal of the College of Science, Imperial University of Tokyo*, 7, 111–200.
- Orlecka-Sikora, B. (2010). The role of static stress transfer in mining induced seismic events occurrence, a case study of the rudna mine in the legnica-glogow copper district in poland. *Geophysical Journal International*, 182(2), 1087.
- Otsuka, M. (1985). Studies on aftershock sequences, part 1: Physical interpretation of omori formula (in Japanese). *Scientific Reports Shimabara Earthquake Volcano Observatory*, 21, 11–20.
- Otsuka, M. (1987). A simulation of earthquake occurrence, part 8: On omori's law to express aftershock seismicity (in Japanese). *Zisin, Ser*, 2(40), 65–75.
- Pintelon, R., & Schoukens, J. (2001). *System identification: A frequency domain approach*. New Jersey: IEEE Press.
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing*.
- Porat, B. (1997). *A course in digital signal processing*. New York: Wiley.
- Rao, N. R., & Kaila, K. (1986). Application of the negative binomial to earthquake occurrences in the Alpide-Himalayan belt. *Geophysical Journal International*, 85, 283–290.
- Richter, C. F. (1958). *Elementary seismology*. San Francisco: Freeman.
- Shapiro, S. A., Dinske, C., & Kummerow, J. (2007). Probability of a given-magnitude earthquake induced by a fluid injection. *Geophysical Research Letters*, 34(22).
- Shapiro, S. A., Dinske, C., Langenbruch, C., & Wenzel, F. (2010). Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulations. *The Leading Edge*, 29(3), 304–309.
- Smyth, C., & Mori, J. (2011). Statistical models for temporal variations of seismicity parameters to forecast seismicity rates in Japan. *Earth, Planets and Space*, 63(3), 7.
- Söderström, T., & Stoica, P. (1989). *System identification*. Englewood Cliffs: Prentice-Hall.
- Souriau, M., Souriau, A., & Gagnepain, J. (1982). Modeling and detecting interactions between earth tides and earthquakes with application to an aftershock sequence in the Pyrenees. *Bulletin of the Seismological Society of America*, 72(1), 165–180.
- Špičák, A., Horálek, J., Boušková, A., Tomek, Č., & Vaněk, J. (1999). Magma intrusions and earthquake swarm occurrence in the western part of the Bohemian massif. *Studia Geophysica et Geodaetica*, 43(1), 87–106.
- Svartsjærn, M. (2018). A prognosis methodology for underground infrastructure damage in sublevel cave Mining. *Rock Mechanics and Rock Engineering*.
- Svartsjærn, M., Saiang, D., Nordlund, E., & Eitzenberger, A. (2016). Conceptual numerical modeling of large-scale footwall behavior at the Kiirunavaara Mine, and implications for deformation monitoring. *Rock Mechanics and Rock Engineering*, 49(3), 943–960.
- Tsukuda, T. (1991). Earthquake swarm. *The Earth Monthly*, 13, 673–678.
- Tsukuda, T. (1993). West iriomote island earthquake swarm viewed as time series (in Japanese). *The Earth Monthly*, 15(4), 218–224.

- Vallejos, J. A., & Mckinnon, S. D. (2010). Omori's law applied to mining-induced seismicity and re-entry protocol development. *Pure and Applied Geophysics*, 167, 91–106.
- Vallejos, J. A., & McKinnon, S. D. (2011). Correlations between mining and seismicity for re-entry protocol development. *International Journal of Rock Mechanics and Mining Sciences*, 48(4), 616–625.
- Węglarczyk, S., & Lasocki, S. (2009). Studies of short and long memory in mining-induced seismic processes. *Acta Geophysica*, 57(3), 696–715.
- Wettainen, T., & Martinsson, J. (2014). Estimation of future ground vibration levels in Malmberget town due to mining-induced seismic activity. *Journal of the Southern African Institute of Mining and Metallurgy*, 114, 835–843.

(Received May 1, 2019, revised September 24, 2019, accepted November 22, 2019, Published online December 4, 2019)