Farnaz Kamranzad[1], Mark Naylor[1], Finn Lindgren[2], Kirsty Bayliss[3], and Ian Main[1]

[1]School of GeoSciences, University of Edinburgh, UK
[2]School of Mathematics, University of Edinburgh, UK
[3]Global Earthquake Model (GEM) Foundation, Italy

April 12, 2024

Abstract

The development of reliable operational earthquake forecasts is dependent upon managing uncertainty and bias in the parameter estimations obtained from models like the Epidemic-Type Aftershock Sequence (ETAS) model. Given the intrinsic complexity of the ETAS model, this paper is motivated by the questions: "What constitutes a representative sample for fitting the ETAS model?" and "What biases should we be aware of during survey design?". In this regard, our primary focus is on enhancing the ETAS model's performance when dealing with short-term temporally transient incompleteness, a common phenomenon observed within early aftershock sequences due to waveform overlaps following significant earthquakes. We introduce a methodological modification to the inversion algorithm of the ETAS model, enabling the model to effectively operate on incomplete data and produce accurate estimates of the ETAS parameters. We build on a Bayesian approach known as *inlabru*, which is based on the Integrated Nested Laplace Approximation (*INLA*) method. This approach provides posterior distributions of model parameters instead of point estimates, thereby incorporating uncertainties. Through a series of synthetic experiments, we compare the performance of our modified version of the ETAS model with the original (standard) version when applied to incomplete datasets. We demonstrate that the modified ETAS model effectively retrieves posterior distributions across a wide range of mainshock magnitudes and can adapt to various forms of data incompleteness, whereas the original model exhibits bias. In order to put the scale of bias into context, we compare and contrast further biases arising from different scenarios using simulated datasets. We consider: (1) sensitivity analysis of the modified ETAS model to a time binning strategy; (2) the impact of including and conditioning on the historic run-in period; (3) the impact of combination of magnitudes and trade-off between the two productivity parameters K and α; and (4) the sensitivity to incompleteness parameter choices. Finally, we explore the utility of our modified approach on three real earthquake sequences including the 2016 Amatrice earthquake in Italy, the 2017 Kermanshah earthquake in Iran, and the 2019 Ridgecrest earthquake in the US. The outcomes suggest a significant reduction in biases, underlining a marked improvement in parameter estimation accuracy for the modified ETAS model, substantiating its potential as a robust tool in seismicity analysis.

submitted to *Geophys. J. Int.*

Enhancing the ETAS model: incorporating rate-dependent incompleteness, constructing a representative dataset, and reducing bias in inversions

Farnaz Kamranzad[1], Mark Naylor[1], Finn Lindgren[2], Kirsty Bayliss[3], and Ian Main[1]

[1] *School of GeoSciences, University of Edinburgh, EH9 3FE, Edinburgh, UK.*
[2] *School of Mathematics, University of Edinburgh, EH9 3FD, Edinburgh, UK.*
[3] *Global Earthquake Model (GEM) Foundation, via Ferrata 1, 27100 Pavia, Italy.*

Received 2024 April 07; in original form 2024 April 07

SUMMARY

The development of reliable operational earthquake forecasts is dependent upon managing uncertainty and bias in the parameter estimations obtained from models like the Epidemic-Type Aftershock Sequence (ETAS) model. Given the intrinsic complexity of the ETAS model, this paper is motivated by the questions: "What constitutes a representative sample for fitting the ETAS model?" and "What biases should we be aware of during survey design?". In this regard, our primary focus is on enhancing the ETAS model's performance when dealing with short-term temporally transient incompleteness, a common phenomenon observed within early aftershock sequences due to waveform overlaps following significant earthquakes. We introduce a methodological modification to the inversion algorithm of the ETAS model, enabling the model to effectively operate on incomplete data and produce accurate estimates of the ETAS parameters. We build on a Bayesian approach known as *inlabru*, which is based on the Integrated Nested Laplace Approximation (*INLA*) method. This approach provides posterior distributions of model parameters instead of point estimates, thereby incorporating uncertainties. Through a series of synthetic experiments, we compare the performance of our modified version of the ETAS model with the original (standard) version when applied to incomplete datasets. We demonstrate that the modified ETAS model effectively retrieves posterior distributions across a wide range of mainshock magnitudes and can adapt to various forms of data incompleteness, whereas the original model exhibits bias. In order to put the scale of bias into context, we compare and contrast further biases arising from different scenarios using simulated datasets. We consider: (1) sensitivity analysis of the modified ETAS model to a time binning strategy; (2) the impact of including and conditioning on the historic run-in period; (3) the impact of combination of magnitudes and trade-off between the two productivity parameters K and α; and (4) the sensitivity to incompleteness parameter choices. Finally, we explore the utility of our modified approach on three real earthquake sequences including the 2016 Amatrice earthquake in Italy, the 2017 Kermanshah earthquake in Iran, and the 2019 Ridgecrest earthquake in the US. The outcomes suggest a significant reduction in biases, underlining a marked improvement in parameter estimation accuracy for the modified ETAS model, substantiating its potential as a robust tool in seismicity analysis.

Key words: Statistical seismology; Theoretical seismology; Earthquake interaction, forecasting, and prediction; Statistical methods; Bayesian inference.

1 INTRODUCTION

Seismicity modelling plays a crucial role in understanding the behaviour of earthquake sequences. This process involves fitting appropriate statistical models to effectively describe and forecast the spatial, temporal, spatio-temporal and magnitude patterns of earthquakes. These models build on well-recognised empirical relations, most commonly: (1) the *Gutenberg-Richter law* (Gutenberg & Richter 1944), which describes the distribution of earthquake magnitudes and their corresponding frequencies of occurrence; (2) the *modified Omori law* (Omori 1895; Utsu 1957; Shcherbakov et al. 2004), which explains the decay rate of aftershocks over time following a mainshock; (3) *Utsu's scaling productivity law* (Utsu 1972; Mignan 2018; Shebalin et al. 2020), which estimates aftershock productivity based on mainshock magnitude; (4) *Båth's law* (Båth 1965), which determines the magnitude difference between a mainshock and its largest aftershock; and (5) the *ETAS model* (Ogata 1988; Ogata & Zhuang 2006; Ogata 2011), which amalgamates elements from the aforementioned models, and expands the modelling framework by capturing the effect of complex inter-event interactions.

Over its 35-year evolution, the ETAS model has established itself as a core tool for retrospective seismicity analysis and prospective operational earthquake forecasting. Central to the ETAS model is the concept that earthquake populations can be modelled as a marked point process and that any earthquake has the potential to trigger subsequent aftershocks, initiating a branching pattern of seismic activity — this class of statistical model is referred to as a self-exciting point process, or a marked Hawkes process. The ETAS model is a specific example. It characterises aftershock sequences through two components: the background seismicity rate, representing the average baseline rate of independent earthquakes within a specified spatial and temporal domain, and the triggered seismicity, which encompasses the additional seismic activity triggered by preceding earthquakes. Thus, the ETAS model offers a dynamic representation of earthquake occurrences, facilitating the analysis and forecast of aftershock sequences, and enhancing the understanding of the clustered nature of seismic events.

Many flavours of the ETAS model exist. The majority employ a maximum-likelihood estimation (MLE) method to produce point estimates of the model parameters through an optimisation algorithm. Important algorithms include gradient-based methods e.g. (Ogata 1998; Jalilian 2019), expectation-maximisation (EM) e.g. (Veen & Schoenberg 2008; Mizrahi et al. 2023; Stindl & Chen 2023), non-linear methods e.g. (Kanazawa & Sornette 2023), machine-learning likelihood-free inference e.g. (Stockman et al. 2023), etc. Recent research studies have adopted Bayesian inference, focusing on providing posterior probability distributions instead of point estimates for ETAS parameters. This shift allows for applying prior constraints on the model parameters and facilitates a more comprehensive exploration of uncertainties associated with these parameters. Examples include (Omi et al. 2015; Ebrahimian & Jalayer 2017; Shcherbakov et al. 2019; Ross 2021; Schneider & Guttorp 2021; Shcherbakov 2021; Laub et al. 2021; Ebrahimian et al. 2022; Molkenthin et al. 2022; Ross & Kolev 2022; Naylor et al. 2023; Nishikawa & Nishimura 2023).

The widespread use of the ETAS model comes with a significant challenge: accurately assessing parameter estimations when applied to real data is difficult. This difficulty arises because the methods listed above usually return a parameter set without flagging potential issues of bias. To address this, synthetic experiments offer a solution by allowing us to understand how issues, such as incomplete datasets, influence the accuracy and precision of parameter estimations, as well as the efficacy of the ETAS model. Once important sources of bias leading to epistemic uncertainty (i.e. that which cannot be quantified by the random error or aleatory uncertainty) are identified, we can identify routes to accommodate or correct such biases. Then, when we return to real datasets, where the true underlying model remains unknown, we are restricted to making comparative estimates of accuracy or bias against synthetic data, where the underlying parameters are known. Consequently, by ensuring that the corrections applied are consistent with those made in the synthetic experiments, we can bolster our confidence in these corrective measures.

A number of studies have investigated some limitations, considerations, and advancements related to the ETAS model including the effect of short-term time-varying incompleteness (Moradpour et al. 2014; Omi et al. 2014; Hainzl 2016; Page et al.2016; de Arcangelis et al. 2018; Hardebeck et al. 2019; Lippiello et al. 2019; Hainzl 2021; Mizrahi et al. 2021; Grimm et al. 2022; Iacoletti et al. 2022; van der Elst et al. 2022; Naylor et al. 2023), model under-fitting for major mainshock-aftershock sequences and over-fitting for regions with normal seismicity (Harte 2013), impact of triggering boundary magnitude (Harte 2016), the impact of sample size and temporal finiteness of catalogues on background rate and branching ratio estimators (Seif et al. 2017), time-varying background rates (Muir & Ross 2023), the impact of run-in history before mainshock (Naylor et al. 2023), incorporating anisotropic spatial kernels (Ogata 2011; Moradpour et al. 2014; Zhang et al. 2018; Grimm et al. 2021; Grimm et al. 2022), restricting infinite spatial extent (Grimm et al. 2021; Grimm et al. 2022), and including extra covariates (Adelfio & Chiodi 2021; Chiodi et al. 2021).

In this study, we aim to enhance the accuracy of parameter estimations for the ETAS model, when dealing with datasets characterised by short-term, time-varying incompleteness. We also try to identify some potential sources of bias and introduce ways to select appropriate representative samples to minimise the bias during the training of the ETAS model. We use the Integrated Nested Laplace Approximation (`INLA`) method (Rue et al. 2009), along with its extension, the `inlabru` package (Bachl et al. 2019), for computing posterior estimates of the ETAS model parameters within a Bayesian framework. Compared to the traditionally used Markov Chain Monte Carlo (`MCMC`) method, `INLA` and `inlabru` provide substantial computational benefits, notably in efficiency and speed. However, it is important to note that the methodological improvements and investigations proposed in our study aim to address broader issues inherent in the ETAS model. These improvements are applicable regardless of the specific estimation technique used, whether it involves point estimate methods or Bayesian implementations of the ETAS model. This paper is structured as follows: We begin by introducing the fundamental concepts and modifications we have applied to the inversion algorithm of the ETAS model to tackle the issue of short-term incompleteness in data, as detailed in Section 2. In Section 3, we assess and compare the performance of both the original and our modified ETAS models using synthetic earthquake catalogues, demonstrating how our modifications enhance the accuracy of ETAS parameter estimation in the presence of incomplete data. We then apply both models to three real aftershock sequences to examine the consistency of the results. Additionally, we introduce some considerations to minimise further biases. In Section 4 we elaborate on remaining limitations and possible improvements. We then discuss the process of selecting representative samples for ETAS estimations in Section 5. Finally, in Section 6, we conclude by summarising the main findings of our research.

2 METHODS

2.1 Concept and formulation of the ETAS model

The ETAS model is a spatio-temporal statistical model used to describe and forecast the occurrence rate of aftershocks. Aftershocks are modelled as a *self-exciting point process*, often referred to as *Hawkes process* in statistics. Hawkes processes are non-Markovian, meaning that the memory of the previously occurred events changes the probability of the upcoming events. Conceptually, this means that in a sequence of aftershocks, every earthquake can trigger other future earthquakes, which in turn generate more earthquakes and so on, creating a "cascade" or "epidemic" of events. Consequently, unlike more basic models that assume aftershocks are directly triggered only by the mainshock, the ETAS model takes into account the secondary, tertiary, etc., aftershocks as well, and assumes that aftershocks can act as "parents" to further "generations" of aftershocks (also known as "offspring", "descendants", or "daughters") in a branching process, leading to an inter-connected sequence of earthquakes. Here, we refer to them as "triggering" and "triggered" events, respectively.

A Hawkes process is mathematically represented by its conditional intensity function, which provides the rate of events at any given point in time and space. In this study, we specifically focus on the temporal model with general form

$$\lambda_{\text{Hawkes}}(t|\mathcal{H}_t) = \mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} g(t) \tag{1}$$

where $\lambda_{\text{Hawkes}}(t|\mathcal{H}_t)$ represents the expected rate of events at time t, taking into account the history of process up to that point, denoted by $\mathcal{H}_t$. The history includes the set of past events as $\mathcal{H}_t = \{(t_i, m_i) : t_i < t,\ m_i \geq M_0,\ i = 1, \ldots, n\}$. m_i and t_i correspond to the magnitude and time of the i^{th} earthquake in the history, respectively. M_0 represents the explicit constant reference magnitude, ensuring that the model parameters remain constant. μ is the background rate, and it can be regarded as the "base level" of earthquakes in a region, representing the rate of spontaneous earthquake occurrences that are independent of each other, i.e. are not triggered by other events. $\sum$ is the sum over all triggering earthquakes that happened before time t; The function $g(t)$ inside the summation is referred to as the "triggering function" and determines the triggering contribution from all previous events to the occurrence of future events; $g(t)$ can take various functional forms, with exponential and power-law functions being commonly used in practice. Here, we consider one of the most-commonly used form as

$$\lambda_{\text{Hawkes}}(t|\mathcal{H}_t) = \mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}, \tag{2}$$

where K, α, c, and p are the model parameters to be estimated along with μ. (see Table A1).

The ETAS model is a specific type of marked Hawkes process with a conditional intensity function that can be expressed as

$$\lambda_{\text{ETAS}}(t, |\mathcal{H}_t, m) = \left[\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}\right]\beta e^{\beta(m-M_0)}, \tag{3}$$

where $\beta e^{\beta(m-M_0)}$ is the probability density form of the Gutenberg-Richter (G-R) law added to the Hawkes model. In this study, we focus primarily on the Hawkes part of the model as none of the G-R parameters are optimised in the inversion, but we will later use the properties of the magnitude model in addressing the censoring data in section 2.4.

Looking at Eq. (2), the first factor of the triggering function, $Ke^{\alpha(m_i-M_0)}$ is often referred to as the "exponential magnitude-based productivity", and is equivalent to Utsu scaling law. This factor determines the increase in seismicity rate after the i^{th} earthquake based on its magnitude m_i. This implies that a larger earthquake will have a greater influence on triggering subsequent events. The second factor, $\left(\frac{t-t_i}{c}+1\right)^{-p}$, also known as the "temporal triggering kernel", represents the decay of this influence over time. It follows a power-law function equivalent to the Omori law, and captures the dependence on time since the triggering event and makes the rate decay over time. The interplay between these two ingredients of the triggering function ensures a balance between a rise of intensity with each event and the temporal decay of it. In modelling aftershocks, such a balance is handled by a quantity called "branching ratio" which controls the average number of aftershocks directly triggered by any given earthquake.

Of the model parameters, both K and α jointly contribute to the productivity of aftershocks but in different ways. Conceptually, K is the base productivity parameter which quantifies the average number of direct aftershocks produced by an earthquake of a reference magnitude M_0.

To be exact, K is the change in intensity by a new event with magnitude $m_i = M_0$. K usually ranges from 0.01 to 10 or more, depending on the magnitude range implied by the choice of model domain, i.e. expected rate increases at $t \simeq t_i$ for parent of M_0. This baseline productivity is then adjusted by $e^{\alpha(m_i - M_0)}$ which is a factor that increases this productivity for larger earthquakes. α is the magnitude scaling productivity parameter, dictating how much more productive an earthquake becomes for each unit increase in its magnitude. This allows a magnitude dependent increase in the intensity. There is always a trade-off between K and α when contributing to the productivity of an earthquake sequence. We will explore this issue in more detail in section 3.3.2.

The other two parameters, c and p, control the Omori-law decay. In the Omori law, c is a characteristic time that represents a short temporal delay after the mainshock during which the rate of aftershocks does not exhibit a decay trend. c can range from a few minutes to several days, depending on the magnitude of the mainshock and the capabilities of the seismic network involved. However, in the context of the ETAS modelling, c has a slightly different meaning and applies to all events, not just the mainshock. Here, c represents a short-term offset in time or a lag period immediately after each triggering earthquake. It is a small, positive value that is used to avoid singularity at $t = t_i$, ensuring finite rates for all times. Typical values for c are very small, often in the range of 0.001 to 0.1 days. Note that the ETAS model is highly sensitive to the choice of parameter c, so that a small change in c can significantly affect the predicted earthquake rates. Specifically, smaller values of c lead to sharper initial increase in aftershock rates immediately after a parent event, accompanied by a rapid temporal decay. In contrast, larger values of c result in a gentler initial increase in aftershocks, followed by a slower decrease over time. Because c directly influences the temporal evolution of the aftershock sequence, precise estimation of this parameter is crucial for accurate modelling and forecasting sequences. Some studies use temporarily varying c to model incompleteness but this mixes a physical and a network design constraint so it is not an ideal implementation. Parameter p is simply the Omori law's exponent and measures how quickly the fading of aftershocks happens. Empirical studies of various aftershock sequences suggest that p typically ranges between 0.8 and 1.5, with larger values of p indicating a faster decay in the rate of aftershocks, while smaller values denote a slower decay. Physically, p is considered a region-based parameter and may vary based on factors such as tectonic environment, temperature, magnitude, and depth of the mainshock, etc.

2.2 Approximation of parameters in the original ETAS model

In this section, we explain the approximation of the model parameters for the original (standard or traditional) ETAS model. This serves as a preliminary step, examining a version of the model before incorporating adjustments for the transient short-term incompleteness observed in early aftershocks. Building on this foundation, we will further develop and adapt the solution for our modified version of the ETAS model, which specifically addresses the short-term incompleteness issue. This will be thoroughly explored in Sections 2.3 and 2.4.

In statistical modelling, the likelihood function plays a pivotal role in estimating the unknown model parameters. It quantifies how likely a given set of model parameters would produce the observed data. For the Hawkes process model, the likelihood function in the interval $t \in [T_1, T_2]$ is defined as

$$L(\boldsymbol{\theta}|\mathcal{H}) = \exp\left(-\int_{T_1}^{T_2} \lambda(t|\mathcal{H}_t)\, dt\right) \prod_{(t_i,m_i)\in\mathcal{H}} \lambda(t_i|\mathcal{H}_{t_i}), \tag{4}$$

or equivalently in logarithmic form as

$$\mathcal{L}(\boldsymbol{\theta}|\mathcal{H}) = \log L(\boldsymbol{\theta}|\mathcal{H}) = -\int_{T_1}^{T_2} \lambda(t|\mathcal{H}_t)\, dt + \sum_{(t_i,\, m_i)\,\in\,\mathcal{H}} \log \lambda(t_i\,|\mathcal{H}_{t_i}). \tag{5}$$

Here, $\lambda(t|\mathcal{H}_t)$ represents the intensity function as detailed in Eq. (2), and $\boldsymbol{\theta}$ denotes the vector of model parameters that we aim to estimate. For the temporal ETAS modelling $\boldsymbol{\theta}$ = (μ , K, α, c, and p). We use the logarithmic form of likelihood function as it effectively transforms multiplications into additions, making complex calculations simpler and more numerically stable. By substituting Eq. (2) into Eq. (5) and then solving the integral, the log-likelihood function is obtained as

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}|\mathcal{H}) =& -\mu\,(T_2 - T_1)\\ &- \sum_{(t_i,m_i)\in\mathcal{H}} K e^{\alpha(m_i-M_0)} \frac{c}{p-1}\left[\left(\frac{\max(T_1,t_i)-t_i}{c}+1\right)^{1-p} - \left(\frac{T_2-t_i}{c}+1\right)^{1-p}\right]\\ &+ \sum_{(t_i,m_i)\in\mathcal{H}} \log\left(\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} K e^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}\right),\end{aligned} \tag{6}$$

where the 1st term represents the expected background rate, the 2nd term is the expected number of triggered earthquakes by each triggering event, and the 3rd term indicates the sum of log-intensities. However, the approximation of the 2nd term, which considers the role of each triggering event, is not adequately precise as it is. This is primarily due to the fact that a Hawkes process is naturally impulsive, and it is a summation of exponential functions that spike after each event. Also, for each event, the triggering function varies most rapidly for the times close to t_i and becomes nearly constant moving away from it. So, to properly handle such rate fluctuations, a time binning strategy is usually applied e.g. in (Kirchner 2017; Cheysson & Lang 2022; Shlomovich et al. 2022). This involves partitioning the impact interval of each event, $[t_i, T_2]$, into several discrete bins and then counting the rate in each bin, within the model domain. To balance rapidly decreasing rates whilst maintaining reasonable bin occupancy, we adopt an exponential binning strategy for creation of a temporal mesh as proposed in (Naylor et al. 2023):

$$\Big\{ t_i \,, \quad t_i + \Delta \,, \quad t_i + \Delta(1+\delta) \,, \quad t_i + \Delta(1+\delta)^2 \,, \quad \ldots \,, \quad t_i + \Delta(1+\delta)^{n_i} \,, \quad T_2 \Big\}, \tag{7}$$

where $n_i \le n_{\max}$, $\Delta > 0$ and $\delta > 0$. $n_{\max}$ controls the maximum number of bins and the two constants Δ and δ regulate the length of the first bin, and the length ratio between consecutive bins, respectively. By incorporating the binning strategy and linearisation into calculations, the likelihood function undergoes reformulation, resulting in

$$\begin{aligned}
\overline{\mathcal{L}}(\boldsymbol{\theta}|\mathcal{H}) = &- \exp\Big\{ \log \mu + \log{(T_2 - T_1)} \Big\} \\
&- \sum_{(t_i,m_i)\in\mathcal{H}} \sum_{j=0}^{B_i-1} \exp \Bigg\{ \log K + \alpha(m_i - M_0) + \log\left(\frac{c}{p-1}\right) \\
&+ \log\left[\left(\frac{t_j^{b_i} - t_i}{c} + 1\right)^{1-p} - \left(\frac{t_{j+1}^{b_i} - t_i}{c} + 1\right)^{1-p} \right] \Bigg\} \\
&+ \sum_{(t_i,m_i)\in\mathcal{H}} \log\left(\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} K e^{\alpha(m_i - M_0)} \left(\frac{t-t_i}{c} + 1\right)^{-p} \right) .
\end{aligned} \tag{8}$$

This formulation is then input into the `bru` function, which implements the `inlabru` workflow (Bachl et al. 2019) to estimate Bayesian posterior density functions from the product of the prior distributions and the likelihood function. The workflow takes initial trial parameters for ETAS, and iteratively updating these parameters based on the likelihood of observed earthquake data within a Bayesian context using `INLA`, where it does the calculations around the mode of the posteriors (Rue et al. 2009). Once the model parameters stop significantly changing between iterations, it returns the estimated ETAS parameters and their approximate posterior distributions.

In the following section, we will define a modified form of the conditional intensity function for the ETAS model which will accommodate short-term incompleteness. We will then modify the solution for the likelihood function, following a process similar to the steps explained above.

2.3 Transient short-term incompleteness in early aftershocks

2.3.1 Model for time-varying magnitude of completeness, $m_c(t)$

During an aftershock sequence, the overlap of numerous earthquake waveforms leads to censoring of smaller events and hence an upward temporary shift in the magnitude of completeness. This implies that the level of completeness, which is otherwise a constant (M_0), now varies with the activity rate and magnitude of events. Helmstetter et al. (2006) proposed a model describing the evolution of the completeness magnitude in the form

$$m_c(t) = m_i - G - H \log_{10}(t - t_i) \tag{9}$$

following some previous earthquake i. Here, $m_c(t)$ is an estimate of the level of completeness magnitude at time t, and is the maximum value of Eq. (9) computed over all previous earthquakes (van der Elst 2021). G and H are the model parameters ($G, H > 0$). In this study, to simplify the complexity of combining the ETAS and the incompleteness models, we focus solely on the incompleteness caused by a significant mainshock, disregarding the influence of other events. This is reasonable as many sequences include a single significant event. Thus, the incompleteness model is re-written as

$$m_c(t) = M_m - G - H \log_{10}(t - T_m) \tag{10}$$

where t denotes the time after the mainshock ($t > T_m$), and M_m and T_m correspond to the magnitude and occurrence time of the mainshock, respectively. By rearranging Eq. (10) and substituting $m_c(t) = M_0$, we can derive a formula to calculate the end of an incompleteness period following a mainshock, so that

$$T_e = T_m + 10^{(M_m - G - M_0)/H}. \tag{11}$$

where T_e denotes the specific point in time when the time-varying $m_c(t)$ returns to its constant baseline value M_0.

2.4 Modified ETAS: incorporating short-term incompleteness in the model

2.4.1 Defining a time-dependent censorship function

Our approach is to define a censorship factor that is added to the ETAS intensity function, and then modify the likelihood function accordingly, so that we can estimate the expected number of observed events which can be directly compared to the catalogue. Building on section 2.3, we consider a catalogue which is generally complete down to a constant threshold of M_0 but is temporarily complete at a higher threshold of $m_c(t)$. This scenario is common for a short period following large earthquakes. Assuming constant b-value and activity rate, the Gutenberg-Richter law provides an estimate of the expected number of events above those thresholds,

$$\begin{aligned} N(m \geq M_0) &= a10^{-bM_0}, \\ N(m \geq m_c(t)) &= a10^{-bm_c(t)}. \end{aligned} \tag{12}$$

The ratio of these allows us to estimate the proportion of events above M_0 that have been observed,

$$\frac{N(m \geq m_c(t))}{N(m \geq M_0)} = 10^{-b(m_c(t) - M_0)}. \tag{13}$$

A similar approach was used by Stallone and Falcone (2021), who attempted to fill the gaps and restore the missing earthquakes, assuming that the Gutenberg-Richter law holds with the same exponent *b* in the censored part of the data. Unlike their method, our approach employs this ratio as a time-dependent censoring function, not aiming at recovering the missing events in the data, but rather to quantify what proportion of events are observed given the censorship. Therefore, we use this to correct the estimate of the number of expected events above M_0 provided we have a reasonable estimate of $m_c(t)$ at that point in time. This approach avoids the potential inaccuracies that may arise from trying to explicitly reconstruct the missing data, which necessitates assuming a specific pattern of data omission— which might be varying from underestimation to overestimation—and thereby could inadvertently introduce artifacts into the analysis. Substituting Eq. (10) into Eq. (13) for $m_c(t)$, we get

$$\pi(t) = \begin{cases} 1, & \text{if } t \leq T_m, \text{ or } t > T_e, \\ 10^{-b(M_m - G - H\log_{10}(t-T_m) - M_0)}, & \text{if } T_m < t \leq T_e, \end{cases} \tag{14}$$

where $\pi(t)$ is a piece-wise function determining the time-dependent censorship coefficient for our modified model. T_m is the time of mainshock and T_e is the end of the incompleteness interval as calculated in Eq. (11). For the incompleteness period $(T_m, T_e]$, the ratio varies between $0 < \pi(t) < 1$ and it represents the apparent (observed or recorded) rates as a percentage of the actual rates (rates occurring in reality, encompassing both observed and unobserved events with smaller magnitudes). In other words, over the incompleteness period, the apparent rate has an increasing trend until it fully reaches the actual rates.

2.4.2 *Modifying the intensity and likelihood functions*

To incorporate the incompleteness model into the ETAS framework, we initially modify the ETAS conditional intensity function. This modification represents apparent rates rather than actual rates, as

$$\lambda_{\text{apparent}}(t|\mathcal{H}_t, m_c(t)) = \left[\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i - M_0)} \left(\frac{t-t_i}{c} + 1\right)^{-p}\right] \pi(t) \tag{15}$$

We can assume that the correction factor has a slight effect on the background rate within the short-term incompleteness interval; thus, we can disregard adjustments to μ and treat it as constant. Hence,

$$\lambda_{\text{modified}}(t|\mathcal{H}_t, m_c(t)) \simeq \mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i - M_0)} \left(\frac{t-t_i}{c} + 1\right)^{-p} \pi(t). \tag{16}$$

This change in the intensity function leads to changes in approximation of the likelihood function as well. Substituting $\lambda_{\text{modified}}$ for λ in Eq. (5), we have

$$\begin{aligned}\mathcal{L}_{\text{modified}}\left(\boldsymbol{\theta}|\mathcal{H}\right) &= -\int_{T_1}^{T_2} \mu\, dt \\ &- \int_{T_1}^{T_2} \sum_{(t_i,m_i)\in\mathcal{H}_t} K e^{\alpha(m_i-M_0)} \left(\frac{t-t_i}{c}+1\right)^{-p} \pi(t)\, dt \\ &+ \sum_{(t_i,m_i)\in\mathcal{H}} \log\left(\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} K e^{\alpha(m_i-M_0)} \left(\frac{t-t_i}{c}+1\right)^{-p} \pi(t)\right)\end{aligned} \tag{17}$$

By solving the internal integral for the triggering part of Eq. (17) and subsequently incorporating the time binning strategy, along with linearisation (as previously explained in Section 2.2), the modified log-likelihood is

$$\begin{aligned}\overline{\mathcal{L}}_{\text{modified}}(\boldsymbol{\theta}|\mathcal{H}) = &- \exp\left\{\log\mu + \log\left(T_2 - T_1\right)\right\} \\ &- \sum_{(t_i,m_i)\in\mathcal{H}} \sum_{j=0}^{B_i-1} \exp\left\{\log K + \alpha(m_i - M_0) + \log\left(\frac{c}{p-1}\right) + \right. \\ &\left. \log\left[\left(\frac{t_j^{b_i}-t_i}{c}+1\right)^{1-p} - \left(\frac{t_{j+1}^{b_i}-t_i}{c}+1\right)^{1-p}\right]\right\} \cdot I_1(t) \\ &- \sum_{(t_i,m_i)\in\mathcal{H}} \sum_{j=0}^{B_i-1} \exp\left\{\log K + \alpha(m_i - M_0) + \log\left(\frac{c}{p-1}\right) + \log\left(10^{-b(M_m-G-M_0)}\right) + \right. \\ &\log\left[\left[\left(\frac{t_j^{b_i}-t_i}{c}+1\right)^{1-p} \left(t_j^{b_i}-T_m\right)^{bH} \; {}_2F_1\left(-bH,\, 1,\, 2-p,\, \frac{t_j^{b_i}-t_i+c}{t_j^{b_i}-T_m}\right)\right] - \right. \\ &\left.\left.\left[\left(\frac{t_{j+1}^{b_i}-t_i}{c}+1\right)^{1-p} \left(t_{j+1}^{b_i}-T_m\right)^{bH} \; {}_2F_1\left(-bH,\, 1,\, 2-p,\, \frac{t_{j+1}^{b_i}-t_i+c}{t_{j+1}^{b_i}-T_m}\right)\right]\right]\right\} \cdot I_2(t) \\ &+ \sum_{(t_i,m_i)\in\mathcal{H}} \log\left(\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} K e^{\alpha(m_i-M_0)} \left(\frac{t-t_i}{c}+1\right)^{-p} \pi(t)\right)\end{aligned} \tag{18}$$

where ${}_2F_1$ denotes a Gaussian hypergeometric function. This solution represents a joint demonstration of likelihood comprising the previous solution (Eq. 8) and the new one. Within the incompleteness interval, we adhere to the new solution with the applied censorship. Outside this interval, where $\pi(t) = 1$, we switch to the original solution. To determine the appropriate solution, we use indicator functions, $I_1(t)$ and $I_2(t)$, ensuring the correct solution is applied as needed.

These indicators are defined as follows:

$$I_1(t) = \begin{cases} 1, & \text{if} \quad T_1 \leq t \leq T_m \quad \text{or} \quad T_e < t \leq T_2 \\ 0, & \text{otherwise,} \end{cases} \tag{19}$$

and

$$I_2(t) = \begin{cases} 1, & \text{if} \quad T_m < t \leq T_e \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

In the following section, we elaborate on the practical implementation of the transition between the two solutions considering the time binning.

2.4.3 Considerations for time binning in the modified ETAS

As described above, we have adopted a temporal binning strategy designed specifically to enhance the accuracy of calculating the integral of triggered events, especially when the intensity changes rapidly following each triggering event. Here, we investigate the sensitivity of the ETAS model parameters to the choice of time binning, and offer insights into making an informed selection for an optimal binning strategy when fitting the ETAS model to datasets.

Based on the exponential form of binning defined by Eq. (7), the temporal effect domain of each triggering event i is split into several bins, such that bins closer to t_i are narrower (higher resolution), and progressively become wider as the distance from t_i increases. This approach is taken because the triggering function shows the greatest variations at times t close to t_i, and tends to stabilise, or remain nearly constant, at times further away from t_i. In the original ETAS model, having approximately 10 bins for each observed point proves adequate in terms of accuracy and computational costs (Naylor et al. 2023). In the modified ETAS model, the computation of integral values across bins introduces additional complexities. These complexities are primarily due to increased variations in the new modified triggering function and are exacerbated by overlaps between binning intervals and critical temporal markers, T_m and T_e. Moreover, the way how we amalgamate the integration solution, depending on the bins' positions, further contributes to these challenges.

Here, we divide the entire modelling domain $[T_1\ ,\ T_2]$ into two phases: the complete phase, which includes $[T_1,\ T_m]$ and $(T_e, T_2]$, and the incompleteness period which includes $(T_m, T_e]$. A triggering event with the effect domain $[t_i\ ,\ T_2]$ can occur in any of these phases. Fig. 1 illustrates general examples of triggering events that may happen either before the mainshock ($t_i < T_m$), within the incompleteness interval ($T_m < t_i \leq T_e$), or thereafter ($t_i > T_e$). Given that the effect domain of each triggering event is divided into several bins, a single bin may fall either entirely within the incompleteness period, or completely outside of it, or it may cross the boundaries of the incompleteness period and encompass parts of areas within and outside of the period. Based on this, we identify five distinct scenarios for bins of length $[T_{1b}, T_{2b}]$:

(1): If the whole bin is inside the complete phase (green bins in Fig. 1-a, b, c, and d), such that $[T_{1b}, T_{2b}] \leq T_m$ or $[T_{1b}, T_{2b}] > T_e$, we use the integral solution of the original ETAS in order to count the expected number of events within that bin.

(2): If the whole bin is inside the incompleteness interval (purple bins in Fig. 1-b and c), such that $T_m < [T_{1b}, T_{2b}] \leq T_e$, we apply the integral solution of the modified ETAS.

(3): If the bin is long enough to encompass the whole incompleteness period and parts of the complete intervals (orange bin in Fig. 1-a), so that $T_{1b} < T_m$ and $T_{2b} > T_e$, we split the bin into three sub-bins with length of $[T_{1b}, T_m]$, $(T_m, T_e]$, and $(T_e, T_{2b}]$, and then consider the integral solution of the original, modified, and original ETAS models, respectively.

(4): If the bin crosses the left border of the incompleteness period (yellow bin in Fig. 1-b), so that $T_{1b} < T_m$ and $T_m < T_{2b} < T_e$, we split the bin into two sub-bins with length of $[T_{1b}, T_m]$ and $(T_m, T_{2b}]$, respectively. We then consider the integral solution of the original and the modified ETAS, respectively.

(5): If the bin crosses the right border of the incompleteness period (blue bins in Fig. 1-b and c), so that $T_m < T_{1b} < T_e$ and $T_{2b} > T_e$, we split the bin into two sub-bins with length of $[T_{1b}, T_e]$ and $(T_e, T_{2b}]$, and then consider the integral solution of the modified and original ETAS, respectively.

It is noteworthy that in order to maintain the model finite where a bin starts exactly from the

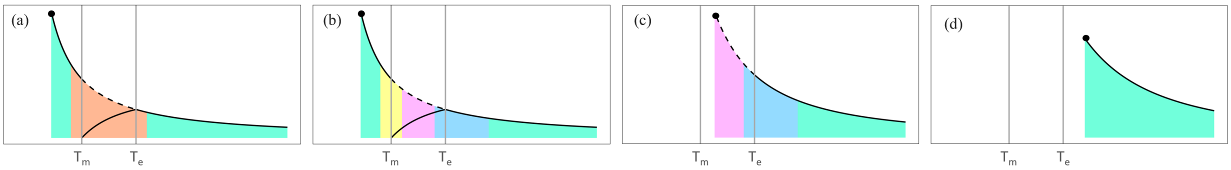

Figure 1. A schematic representation of time-binning considerations in the modified ETAS model. T_m and T_e represent the time of the mainshock and the end of the incompleteness period, respectively. Panel (a) and (b) illustrate triggering events occurring before the mainshock; Panel (c) demonstrates a triggering event occurring within the incompleteness period; and Panel (d) shows a triggering event occurring after the incompleteness period. The bins are colour-coded based on the binning strategy: the green bins entirely fall within the complete interval, the purple bins entirely fall within the incompleteness period, the orange bin starts before the mainshock and ends after the incompleteness period, the yellow bin starts before the mainshock and ends within the incompleteness period, and the blue bins start within the incompleteness period and end after this period. We then divide each bin into sub-bins to apply appropriate integral solutions, as detailed in Section 2.4.3.

mainshock time (where $t_j^{b_i} = T_m$ in the hypergeometric function in Eq. (18)), we add an epsilon value, which in our implementation is set to 10^{-10}.

Conceptually, in the modified model, those bins that fall within the incompleteness period show lower event rates compared to the original model, due to the applied censorship. To quantify this difference, we can calculate the expected number of events inside each bin as a fraction of the count predicted by the original model. For bins in the complete phase, this ratio equals 1, indicating full capture of seismic activity without censorship. Within the incompleteness period, however, the ratio varies between 0 and 1. Near the mainshock time (T_m), the detected events are significantly fewer than the actual rates, resulting in greater censorship and consequently, lower detection ratios. As time progresses, our detection of individual events improves, narrowing the gap between the observed and the actual rates, and thus, the ratio increases. By the end of the incompleteness period (T_e), the detection ratio reaches 1, signifying complete sampling of all events with no missing data.

In Section 3.1, we will further investigate the model's sensitivity to time binning and demonstrate its practical implications.

3 RESULTS

In this section, we present the results from our modified ETAS model which extends the standard ETAS framework by incorporating a time-dependent censorship function in order to address the challenge of short-term, time-varying incompleteness in early aftershocks. We start by a preliminary assessment of the sensitivity of our model to different time-binning choices in Section 3.1, offering guidance for selecting an optimal temporal mesh to secure reliable posteriors. In Section 3.2, we highlight the enhanced parameter estimation accuracy achieved by our model through the analysis of several synthetic datasets, contrasting these findings against those derived from the standard ETAS model. In Section 3.3, we introduce three key considerations to help practitioners in selecting a representative sample that results in more accurate and unbiased ETAS inversions. Lastly, in Section 3.4, we compare the efficacy of both the original and the modified ETAS models through their application to three real earthquake sequences.

3.1 Sensitivity analysis to time binning strategy

To assess the sensitivity of the original and the modified ETAS models to time-binning choices, we fit both models to an incomplete synthetic dataset using different binning options. We first simulate a complete synthetic catalogue spanning 1500 days, including a mainshock with magnitude $M6.7$ on day 500. Then, we filter out events with magnitudes below the incompleteness model (Eq. 10) from the catalogue. We specify the parameters of the incompleteness model as $G = 3.8$, and $H = 1$. We then consider five temporal meshes with different binning designs and resolutions. Table 1 summarises the results regarding the effects of time binning on the run-time, and the number of iterations required for model convergence. As expected, the runtime increased for both the original and the modified models with higher binning resolutions. However, the number of iterations only increased for the modified model. In addition, Fig. 2 exemplifies how binning affects the posteriors of model parameters (μ , K, α, c, and p) in both the original and the modified ETAS models. In this figure, the vertical dashed lines represent the true ETAS parameters that were used to generate the synthetic catalogue for this analysis. Therefore, any deviations from these lines indicate biases in parameter estimations. Clearly, the accuracy of posteriors in the mod-

Table 1. Runtime (in minutes) and number of iterations required for model convergence using different binning options. The analysis was conducted on an incomplete synthetic catalogue spanning 1500 days, featuring a mainshock of magnitude 6.7 on day 500, with incompleteness parameters set at $G = 3.8$ and $H = 1$. We ran the models on a Windows-10 laptop with 16-GB RAM, 4 cores, and 8 logical processors.

parameters of time binning			run time (minutes)		no. iterations for convergence	
Δ	δ	n_{max}	original ETAS	modified ETAS	original ETAS	modified ETAS
0.1	1.0	10	1.7	1.8	37	38
0.05	0.50	20	2	1.8	37	38
0.01	0.25	50	2.5	3.1	37	41
0.005	0.20	75	3.0	4.3	37	42
0.0001	0.15	100	5.3	8.8	36	63

ified model increases with the refinement of mesh resolution, highlighting the significant impact of binning choice. In contrast, the original model remains unaffected by changes in binning, yet it consistently shows significant biases in its posteriors (we will discuss this in more detail in the next section). Here, our goal was merely to conduct a preliminary assessment of the model's sensitivity to time binning, demonstrating the significance of selecting an appropriate mesh resolution to achieve the best performance when fitting the ETAS models. Identifying an optimal binning strategy is beyond the scope of this paper.

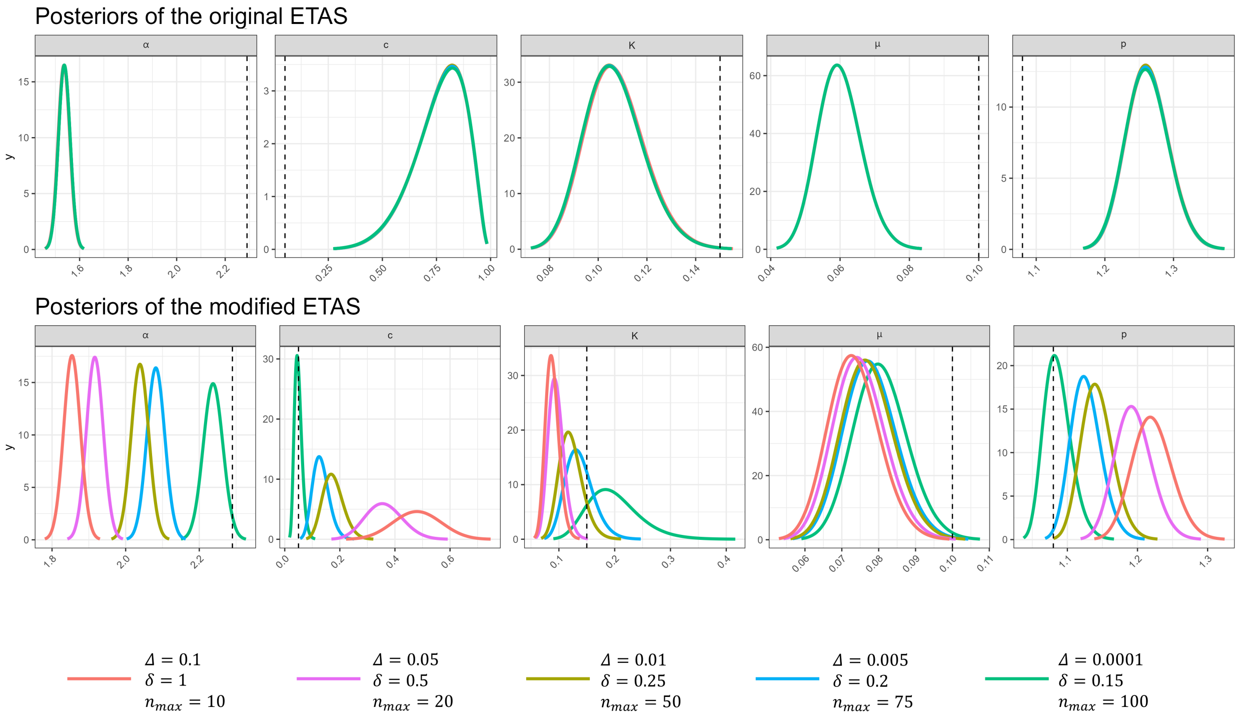

Figure 2. Posterior distributions obtained for the original ETAS model (top row) and the modified ETAS model (bottom row) using five temporal meshes with different binning parameters. Vertical dashed lines mark the true ETAS parameters used to generate a synthetic earthquake catalogue for this study. This catalogue covers a 1500-day period, featuring a mainshock of magnitude 6.7 on day 500. We then removed the incomplete data portion using incompleteness parameters set at $G = 3.8$ and $H = 1$, and fitted both the original and the modified models to the new incomplete catalogue. As the posterior distributions illustrate, the original ETAS model's performance is not influenced by the choice of binning strategy, but shows consistent biases in the estimation of parameters across different resolutions. In contrast, the modified ETAS model is highly sensitive to how data is binned, with the inaccuracies in estimating ETAS parameters significantly reduced as the mesh resolution becomes finer.

3.2 Performance assessment of the modified ETAS model using synthetic data

In real catalogues that exhibit short-term incompleteness after large events, the catalogues are incomplete in the sense that there is partial observation of events below a time-evolving threshold. The Helmstetter model (Eq. 10) estimates this evolving completeness threshold. When we look at real data in Section 3.4, our strategy will be to remove all events below this threshold and correct for this censoring using the apparent intensity function which tells us the proportion of events that should remain above that threshold. Here, we first demonstrate the efficiency of our modified ETAS model through synthetic experiments. In doing so, we generate sets of synthetic catalogues by creating complete catalogues and then remove all events below the completeness predicted by the Helmstetter model. Hereafter, we refer to the data before removal as the 'complete' and the data after removal as the 'incomplete' catalogues. Each synthetic catalogue spans 1500 days, with a mainshock seeded on day 500. The 500-day pre-mainshock period is designed to ensure sufficient background before the emergence of the aftershock cluster (we will discuss this later in Section 3.3.1), and the 1000-day sequence ensures that the sequence has ended and returned to the background, aligning with the temporal windows for $M \leq 8$ introduced by Gardner & Knopoff (1974).

These catalogues are generated with a background rate of $\mu = 0.1$ events per day, a rate consistent with moderately to highly seismic regions. The Gutenberg-Richter b-value parameter is also set to $b = 1$ for this study. Also, we set a constant magnitude threshold at which the catalogue is complete except below $M_0 = 2.5$. The true ETAS parameters were set as $K = 0.15$, $\alpha = 2.29$, $c = 0.05$, and $p = 1.08$, shown by the vertical dashed lines on Figure 4. For the incompleteness models, we adopted parameters proposed by Helmstetter et al. (2006), with $G = 4.5$ and $H = 0.75$. It is worth mentioning that these parameters are not universal, and they can vary based on the seismicity characteristics of a particular region and the capability of seismic networks to record and discriminate between events. In subsequent sections, we will explore different parameters to demonstrate our model's capability to adapt to different incompleteness behaviours. Our modified model is versatile, accommodating a broad range of mainshock magnitudes and incompleteness parameters. As a representation, we provide four synthetic data samples with mainshock magnitudes of 6.0, 6.5, 7.0, and 7.5, as depicted in Fig. 3. In this figure, the left panel displays each sequence's complete data, while the right panel illustrates a close-up view around the incompleteness period. Unobserved (missing) data points are shown with red circles and their count is provided at the top for each case.

We then fitted the original ETAS model to both the complete and incomplete datasets, while the modified ETAS model was applied exclusively to the incomplete dataset (pair plots of ETAS parameters and convergence plots are provided in appendix in Fig. A1 to Fig. A4). The estimated posteriors are illustrated in Fig. 4, with detailed information provided in Table 2. As the results indicate, the original ETAS model, when trained on complete data (blue posteriors), adeptly retrieves the true parameters. However, as explained before, real earthquake sequences often exhibit incompleteness, and when the original ETAS is trained on incomplete data (red posteriors), there is a noticeable bias in parameter estimations. In contrast, the modified model, trained on incomplete data (green posteriors), demonstrates a significant reduction in this bias, so that its estimations closely align with the true parameters and the blue posteriors. This indicates that, despite being fed the incomplete data, the modified ETAS model can achieve accurate posteriors akin to those of the original ETAS model trained on complete data, providing we can parameterise the censoring process. Also, the modified model's posteriors fall within the uncertainty range of the original model but exhibit a shorter peak.

Another evidence supporting our findings can be observed in the triggering function (Eq. 2), which indicates rates for various magnitude thresholds. An illustrative example of this function, using the synthetic catalogue with mainshock magnitude of $M = 7.0$, is presented in Fig. 5. The triggering function derived from the original ETAS model, which was trained on incomplete data, shows a considerable underestimation of event rates (Fig. 5 - middle column). In contrast, the triggering function of the modified ETAS model trained on incomplete data (Fig. 5 - right column) closely mirrors that of the original model when trained on complete data (Fig. 5 - left column). Due to the missing portion of events in the incomplete data, the triggering function of the modified ETAS model, trained on this data, displays a slightly wider uncertainty band than that of the original model trained on complete data. However, it still remains well within the latter's uncertainty bounds. This demonstrates that our modifications have enhanced the model's ability to accurately capture the triggering patterns of aftershocks.

Further evidence of the models' performance, along with a consistency check, is provided by predicted intensities within the short-term incompleteness periods. Figure 6 displays the actual rates (in black) and the apparent (observed/recorded) rates (in purple) for our four selected synthetic catalogues. We then predict the modelled intensities using the posterior modes of both the original (in red) and the modified (in dashed green) ETAS models. The modified ETAS model notably outperforms the original ETAS model in reproducing actual rates. This superiority becomes even more pronounced for larger mainshock magnitudes, which are associated with longer periods of incompleteness and a higher number of missing events. These findings underscore significant differences in the results between the original and the modified ETAS models when trained on incomplete data, thereby highlighting the improvements made in the modified version.

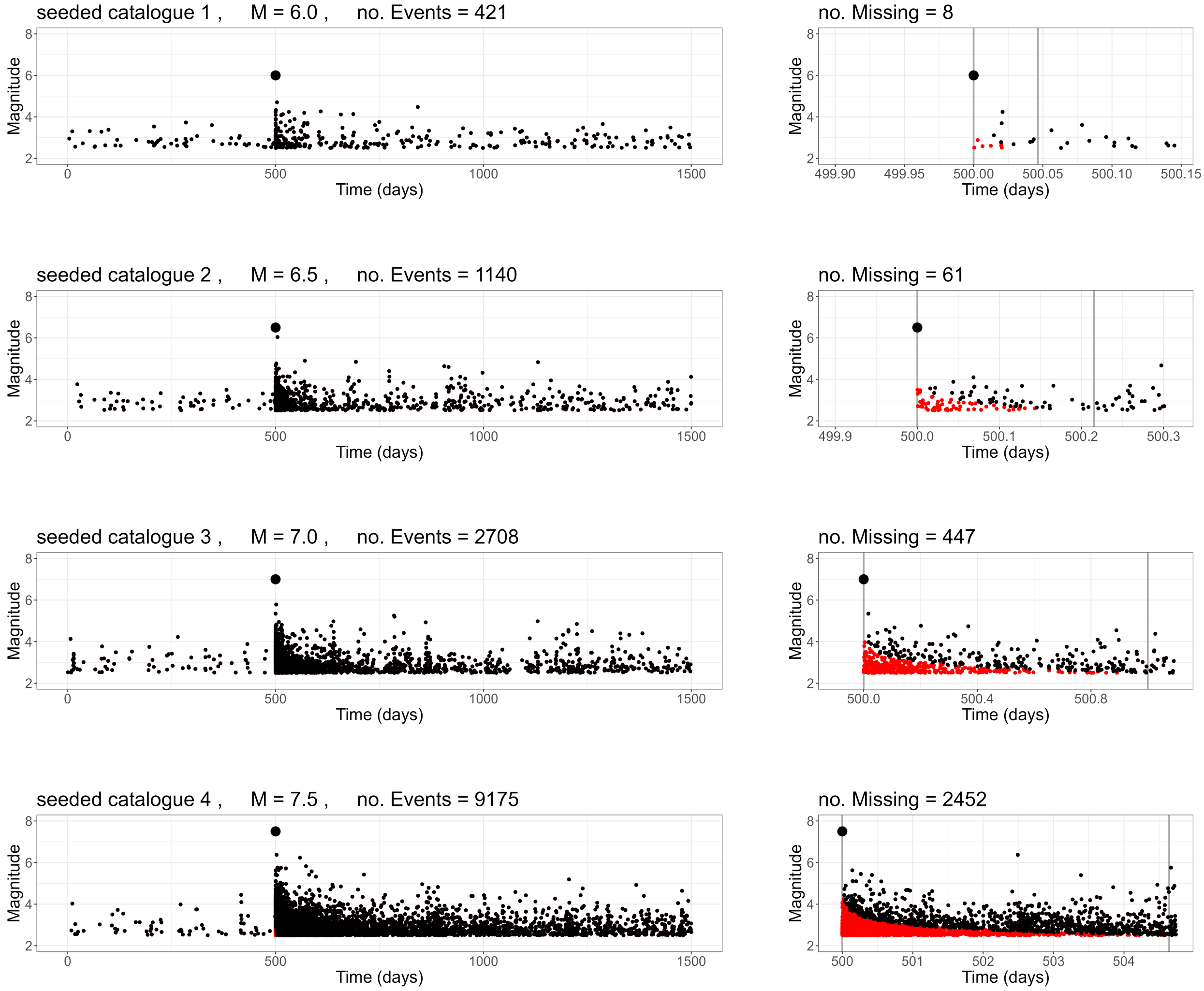

Figure 3. (Left): Four generated 1500-day synthetic catalogues with mainshock magnitudes of 6.0, 6.5, 7.0, and 7.5, seeded on day 500. (Right): Zoomed-in representations showing recorded events in black and missing events in red within the short-term incompleteness interval for each sequence. The incompleteness model parameters are set as $G = 4.5$, $H = 0.75$, and $b = 1$. The number of events is also indicated above each figures. Obviously, under the same parameterisation, the duration of incompleteness and the portion of missing events increase with the magnitude of mainshock.

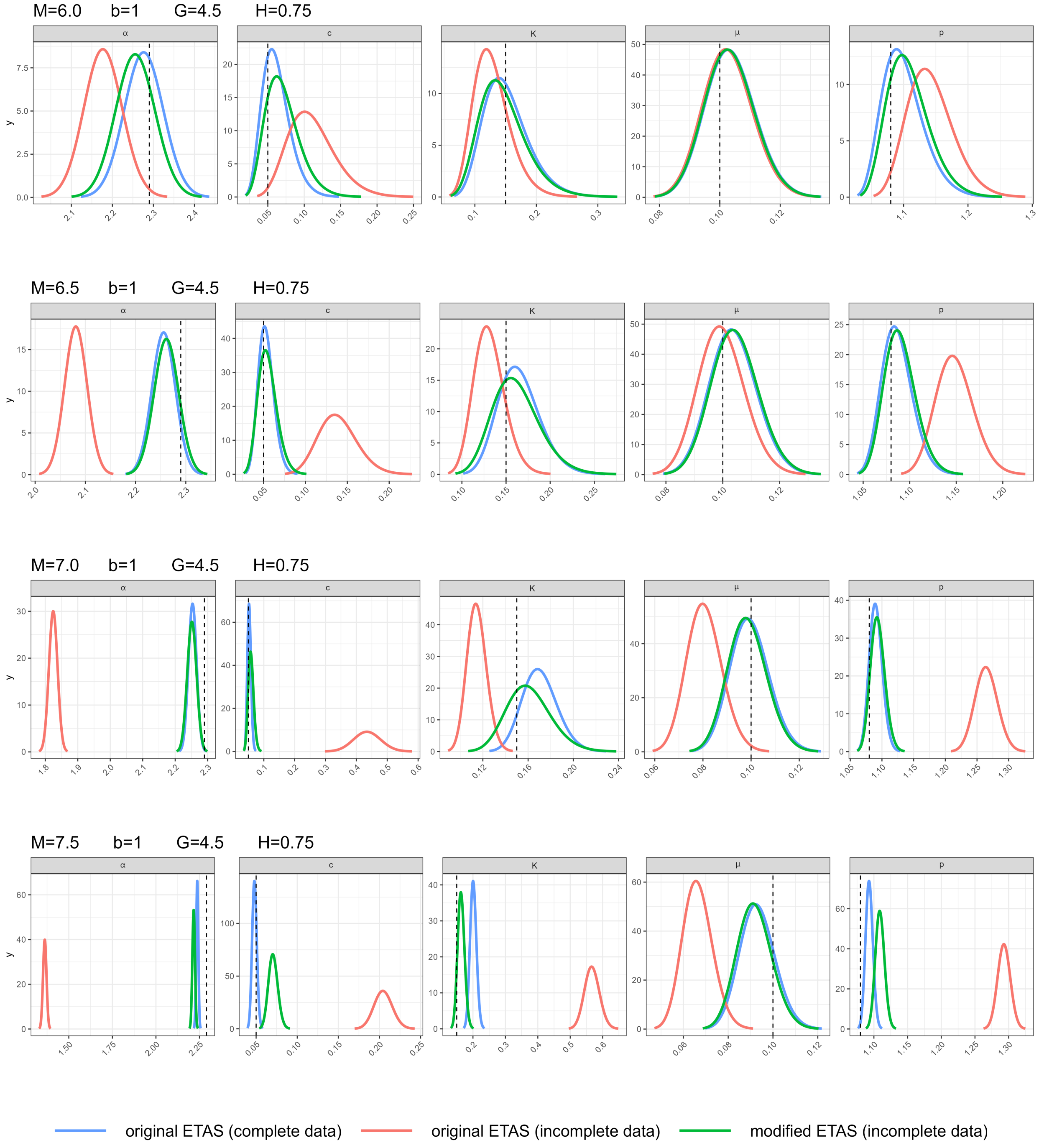

Figure 4. Posteriors presented for four simulated catalogues corresponding to mainshock magnitudes of 6.0, 6.5, 7.0, and 7.5. The blue and red posteriors are derived from the original ETAS model using complete and incomplete catalogues, respectively. The green posteriors are obtained by training the modified ETAS model on the incomplete data. It is evident that when working with data exhibiting short-term incompleteness, the original ETAS model struggles to accurately estimate parameters. In contrast, the modified model performs well, closely approximating the blue posterior, as if working with complete data.

Table 2. Details on data incompleteness and parameter estimates from the original and the modified ETAS models trained on incomplete synthetic catalogues with different mainshock magnitudes. True values are included for reference. Comparison of values indicates that the modified ETAS model clearly outperforms the ETAS original model.

Mainshock magnitude		6.0	6.5	7.0	7.5
number of all events		421	1140	2708	9175
number of recorded events		413	1079	2261	6723
number of missing events		8	61	447	2452
incompleteness period (days)		0.05	0.22	1	4.64
number of events in the incompleteness period		16	106	673	3435
% missing events in the incompleteness period		50%	57.5%	66.4%	71.4%
μ	true value	0.1	0.1	0.1	0.1
	posterior mode (modified ETAS)	0.103	0.104	0.098	0.092
	posterior mode (original ETAS)	0.103	0.099	0.080	0.066
K	true value	0.15	0.15	0.15	0.15
	posterior mode (modified ETAS)	0.14	0.16	0.16	0.16
	posterior mode (original ETAS)	0.12	0.13	0.11	0.57
α	true value	2.29	2.29	2.29	2.29
	posterior mode (modified ETAS)	2.26	2.26	2.25	2.22
	posterior mode (original ETAS)	2.18	2.08	1.82	1.36
c	true value	0.05	0.05	0.05	0.05
	posterior mode (modified ETAS)	0.07	0.05	0.06	0.07
	posterior mode (original ETAS)	0.11	0.14	0.43	0.20
p	true value	1.08	1.08	1.08	1.08
	posterior mode (modified ETAS)	1.10	1.09	1.09	1.11
	posterior mode (original ETAS)	1.14	1.15	1.26	1.29

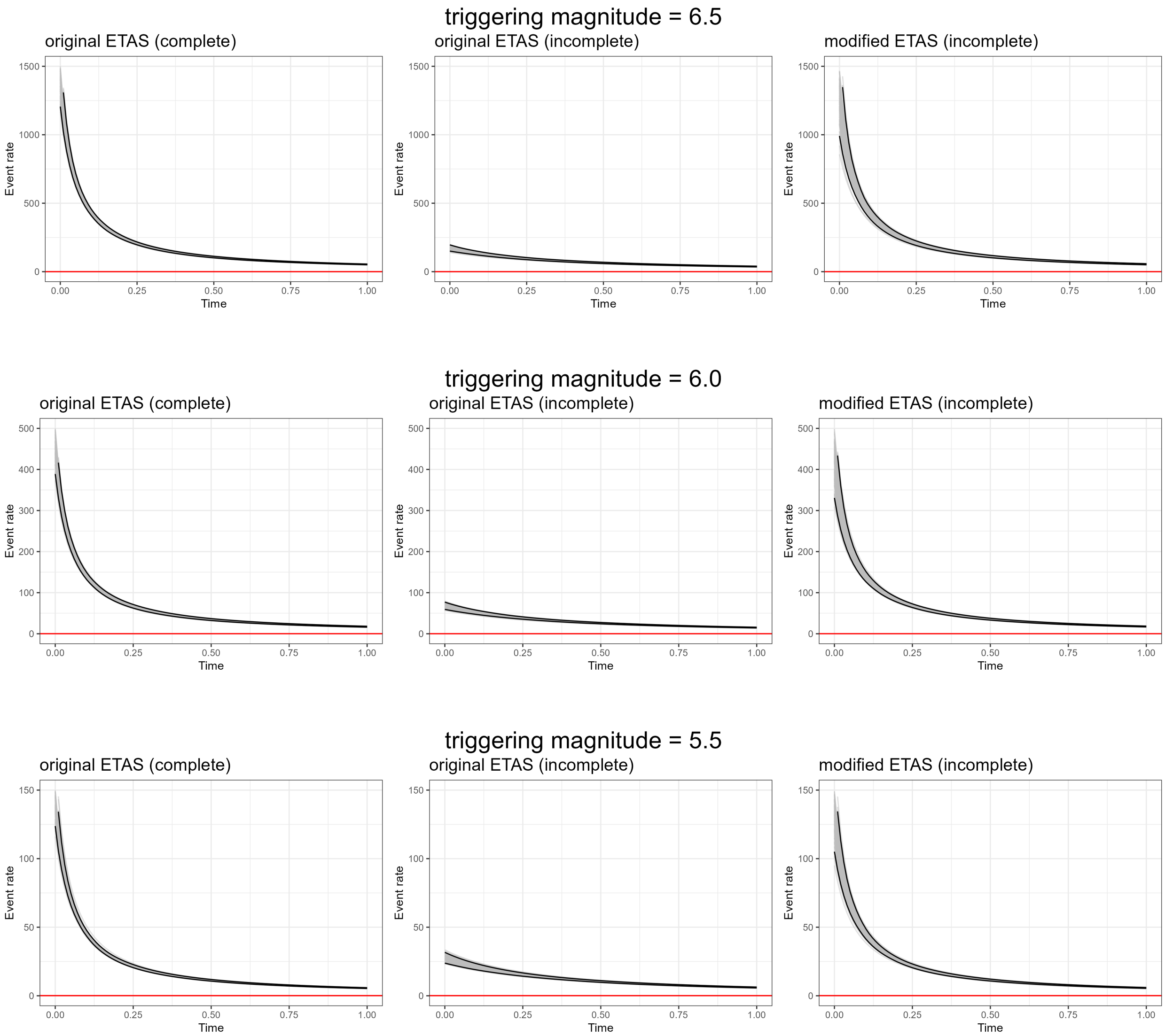

Figure 5. Triggering functions for the synthetic catalogue with mainshock magnitude of 7.0. The functions display rates and associated uncertainties for 1 day after the mainshock at different triggering magnitude levels (6.5, 6.0, and 5.5). These triggering functions are obtained from three scenarios: the original ETAS model trained on complete data (left column), the original ETAS model trained on incomplete data (middle column), and the modified ETAS model trained on incomplete data (right column). The original model exhibits significant underestimation when dealing with incomplete data. In contrast, the modified model accurately estimates rates, resulting in slightly wider plots than the original model with complete data, yet laying within the uncertainty range of the latter.

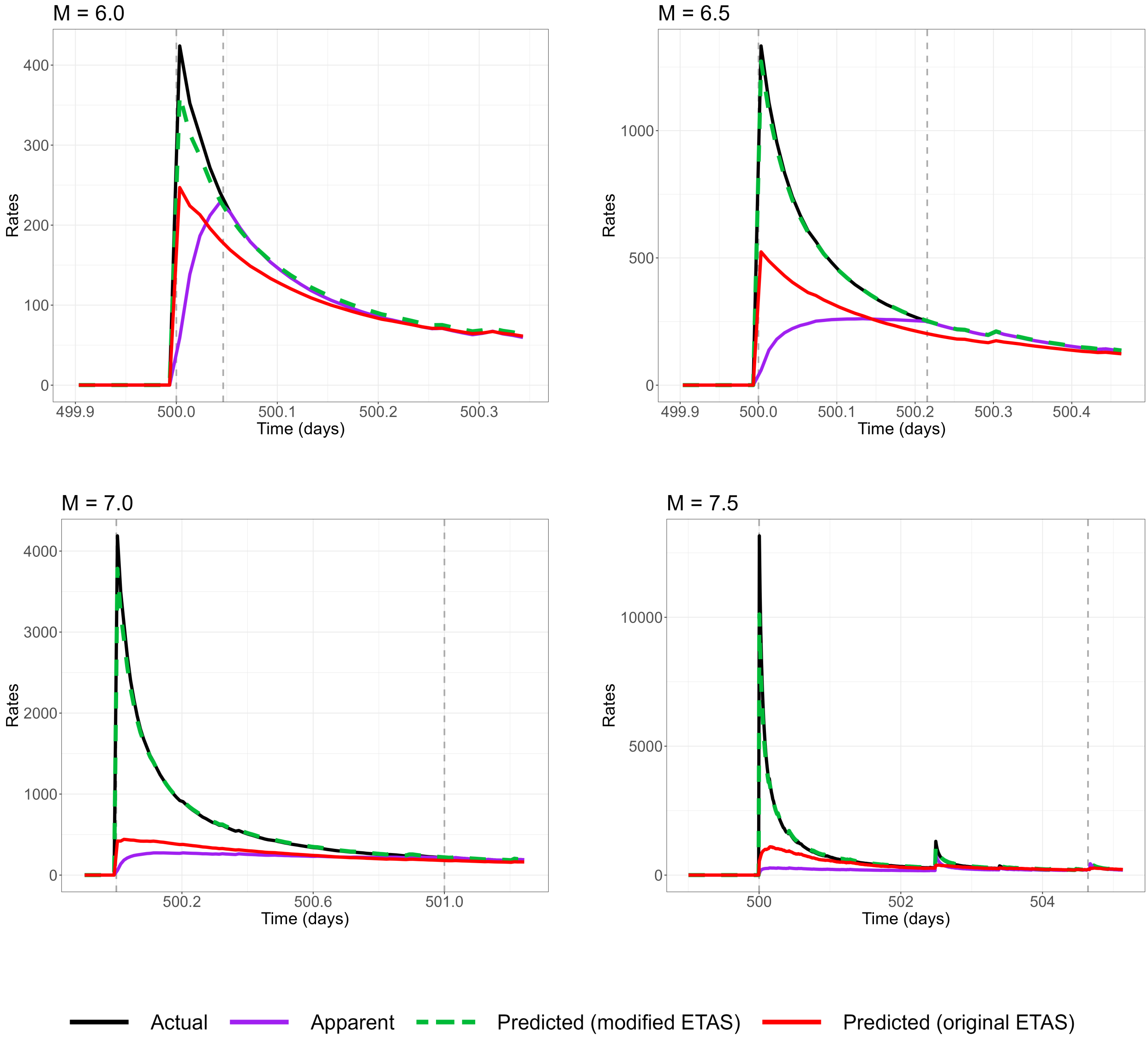

Figure 6. The actual (in black), apparent (in purple), and predicted intensities for four simulated catalogues with mainshock magnitudes of 6.0, 6.5, 7.0, and 7.5. The predicted intensities are calculated using posterior modes obtained from both the original and the modified ETAS models, each trained on incomplete data, and applied to the conditional intensity formula. Two grey vertical dashed lines mark the beginning (mainshock time) and end of the incompleteness period. Clearly, the modified ETAS model (in dashed green) yields predicted intensities that are closer to the actual ones, whereas the original ETAS model (in red) significantly underestimates the intensities.

3.3 Considerations for selecting representative samples and reducing bias in the ETAS inversions

Here we consider the concept of having a representative sample so that the data being analysed contains sufficient information to understand and parameterise the generative processes. This is an intuitive problem when we want to, for example, understand the distribution of heights in the adult population where what is important is that we have a random sample of the population from which we can estimate means and standard deviations, etc.

However, defining a representative sample for a Hawkes process is non-trivial and has important implications for survey design and data analysis. This is evidence in how we choose the spatial-temporal domain to be analysed. Even a purely temporal ETAS model has a spatial component in the sense that we chose to draw a box within which we extract a catalogue to be modelled; and in drawing this box, we are biased towards areas in which there are interesting active sequences. Further, it is common to start the analysis close to the start of the mainshock. The acts of defining domains containing interesting sequences, and excluding regions with lower productivity inherently biases model parameters.

A truly representative sample would contain sufficient diversity that all of the ETAS parameters can be well constrained. In practice, individual case studies may not have sufficient data to permit this. However, by being aware of the deficits in specific case study data, we can anticipate the limitations of our parameter estimations.

Here, we compare the biases that arise from (i) only analysing the active sequences and how this can be mitigated by including historic events to condition a more recent temporal domain, (ii) what properties catalogues require in order to resolve tradeoffs in the productivity parameters, and (iii) exploring the sensitivity of the modified ETAS model to the accuracy of the incompleteness (Helmstetter) model parameters. We hope that these analyses will build intuition regarding the reliability of ETAS inversion on real data. This has important consequences for those attempting to forecast evolving aftershock sequences.

3.3.1 Including and conditioning on the historic run-in period

When dealing with real datasets, it is common practice to calibrate the ETAS model using individual earthquake sequences that start from a mainshock and extend to an inferred endpoint, cropping out the remaining data. Here, we explain how incorporating and also conditioning the model on the history preceding the mainshock impacts the quality of the ETAS inversions. This idea was previously demonstrated for the original ETAS model by Naylor et al. (2023). Here, we apply the same principle to the modified ETAS model, illustrating the degree of bias that can be reduced by including and conditioning on the past seismicity. This shows the extent to which parameters should change under natural variability in simulations. Here, we examine two scenarios: (1) extending the modelling domain to include run-in history prior to a mainshock, and (2) conditioning the model on the history prior to modelling domain.

Within the modelling domain $[T_1, T_2]$, there are some certain events that may not be directly linked to either the triggering events in the sequence or the background activity. Instead, they are triggered by and linked to events that occurred before T_1. This implies that the intensities of these preceding events are still strongly effective, still producing earthquakes, and thereby affecting the overall rate. Therefore, by conditioning the model on the history prior to the modelling domain, we take into account events before T_1 and include their intensities into the rate prediction without evaluating them in the process.

To analyse both scenarios, we generate a catalogue spanning 1,500 days, with an M6.7 mainshock occurring on day 500.01, and M_0=2.5. We then create several sub-catalogues by truncating the first 250, 400, 500, and 501 days (Fig. 7). Subsequently, we conduct two experiments. In the first experiment, we fit the modified ETAS model to the five sub-catalogues within their time intervals $[T_1, T_2]$ with different starting dates T_1. This approach extends the modelling domain each time and incorporates some historical data before the mainshock event. For the second experiment, we repeat the same procedure as in the first one, but additionally, we condition the model on the history before the modelling domain $[0, T_1)$ for each case. This accounts for capturing the influence of previous events that occurred prior to the modelling domain without including them in the modelling process.

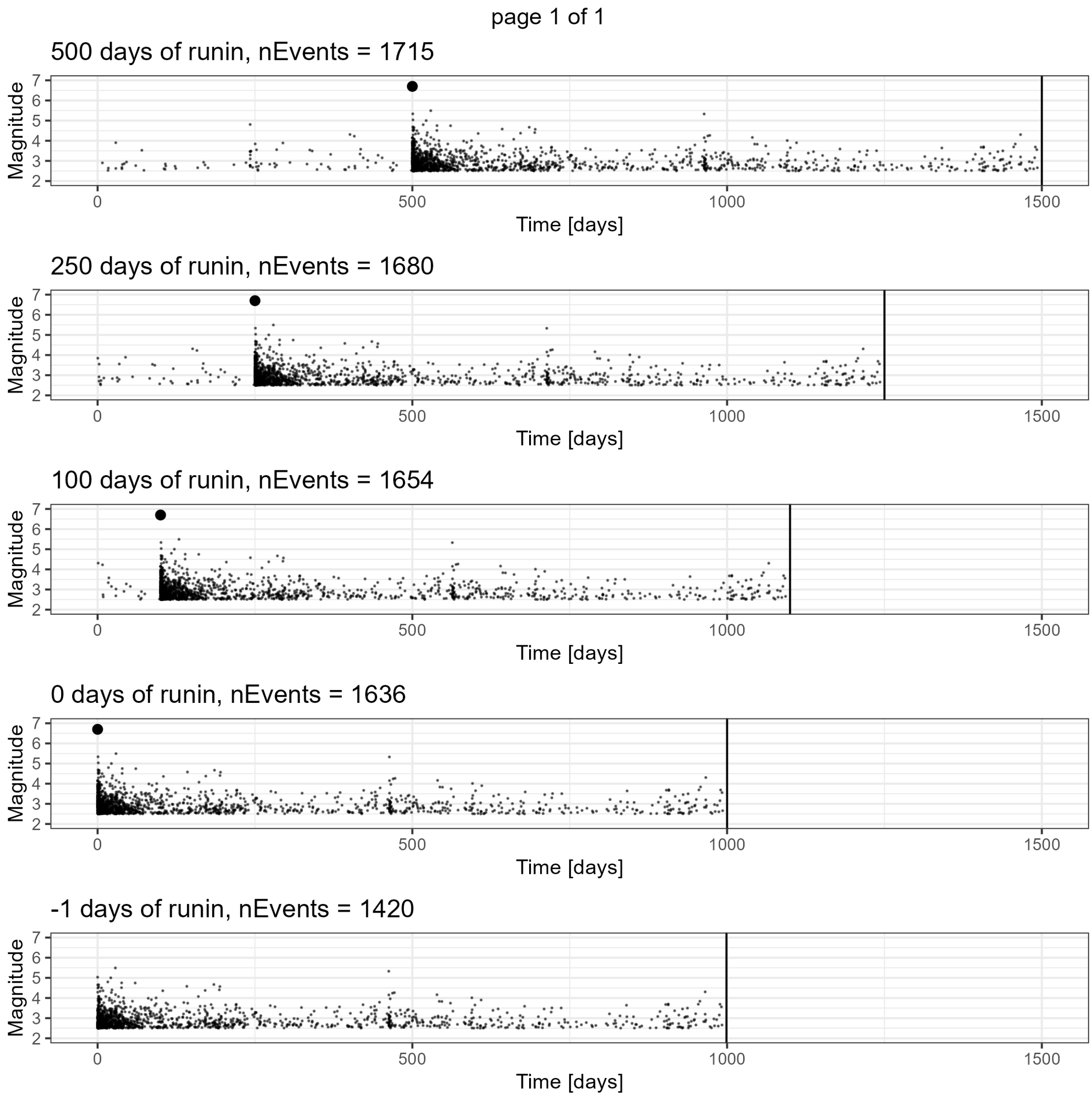

Figure 7. Setting different run-in periods for a 1500-day catalogue, by removing the first 0, 250, 400, 500, and 501 days of data.

Fig. 8 and Table 3 display the results of the inversions using our modified ETAS model. The results indicate that including a run-in history before the mainshock event, and conditioning on the past seismicity before our modelling domain, significantly impacts estimations of the model parameters. With an adequate run-in period, we can reduce bias in the estimates of ETAS parameters, bringing the posteriors closer to their true values. This consideration becomes more crucial in the presence of missing data in the incompleteness interval, where productivity is more affected.

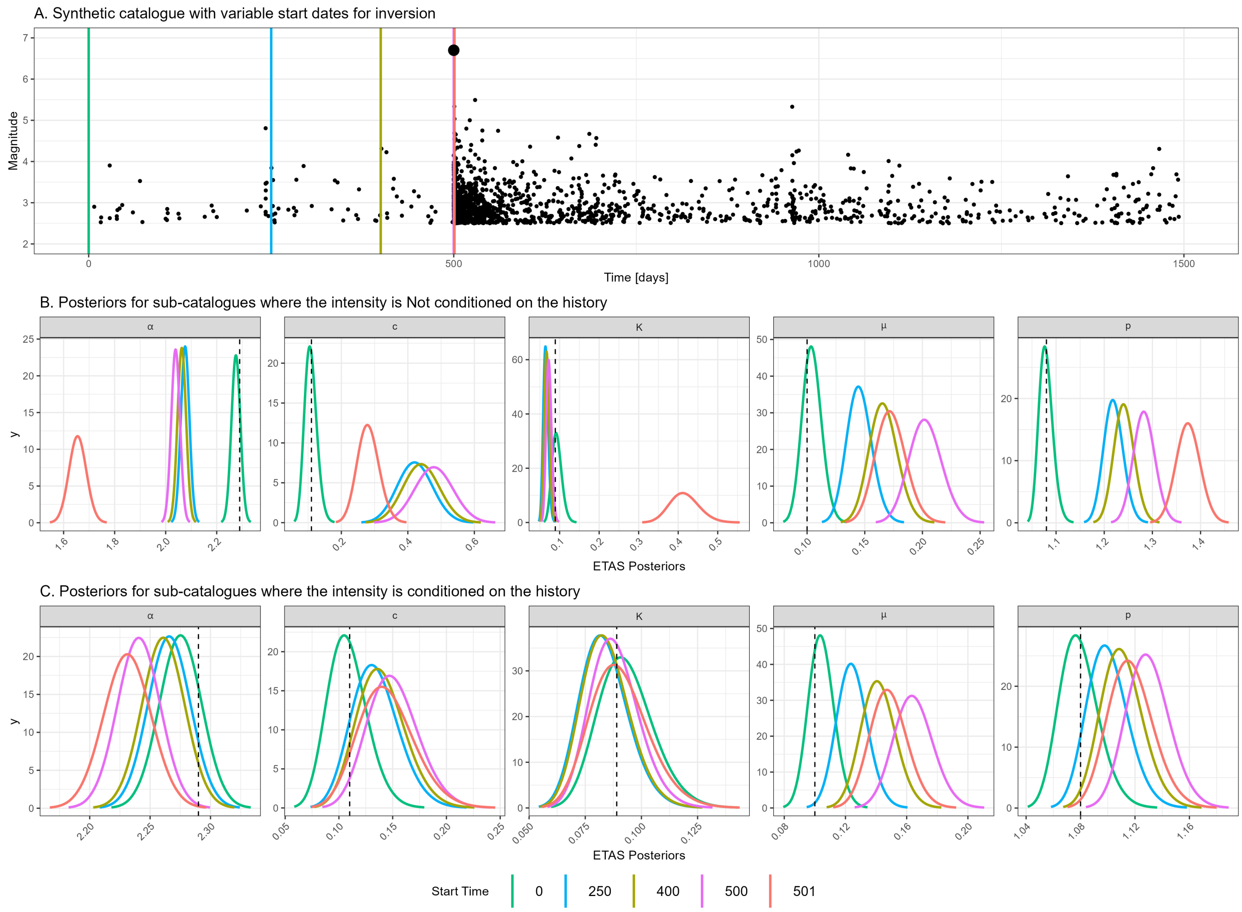

Figure 8. (A) A 1500-day catalogue with short-term incompleteness simulated to train the modified ETAS model, considering "including" and "conditioning on" run-in history. The mainshock with M6.7 was seeded on day 500.01. We extract five sub-catalogues with different starting dates by excluding the first 0 days, 250 days, 400 days, 500 days, and 501 days. (B) Posteriors of the ETAS parameters for each sub-catalogue without conditioning the model on run-in history. (C) Posteriors of the ETAS parameters for each sub-catalogue with the model conditioned on run-in history prior to T_1.

We can conclude that conditioning on the past history significantly enhances model performance, revealing that even for the shortest modelling domain with start date on day 501 that lacks the presence of the large event within the model domain - which is supposed to be the dominant event responsible for majority of aftershock rates - we are still able to retrieve accurate estimates of the ETAS parameters.

Our interpretation of the trends in the bias with respect to the duration of runin period before the mainshock, for these synthetics, hangs on how well the background rate, μ, is resolved. If μ is poorly estimated, we argue that the triggering model will need to compensate, and hence the

Table 3. Estimation of the ETAS parameters for the five synthetic sub-catalogues with a starting point of T_1 for two scenarios: not conditioning and conditioning on history prior to T_1. The mainshock was imposed on day 500.01. The true values for each parameter are also shown below them in the first row of the table.

Scenario	T_1	μ	K	α	c	p
		0.1	0.089	2.29	0.11	1.08
not-conditioned on the run-in history	0	0.10	0.092	2.28	0.11	1.08
	250	0.15	0.065	2.08	0.42	1.22
	400	0.17	0.067	2.06	0.44	1.24
	500	0.20	0.073	2.04	0.48	1.28
	501	0.17	0.414	1.66	0.28	1.37
conditioned on $[0, T_1]$	0	0.10	0.092	2.28	0.11	1.08
	250	0.12	0.083	2.27	0.13	1.10
	400	0.14	0.084	2.26	0.14	1.11
	500	0.16	0.088	2.24	0.15	1.13
	501	0.15	0.090	2.23	0.14	1.12

parameters in the triggering model will also be biased; in these scenarios we have shown that the background rate tends to be high and the triggering effects tend to be underestimated which has implications for forecasting. In our example, at the end of the model domain, T_2, the rate of events has not yet decayed to the background rate, hence the model is entirely dependent on the information in the period prior to the mainshock to calibrate the background rate; this is what permits μ to have high estimates. As the length of the period prior to the mainshock increases, the accuracy of the estimate of μ improves as the intensity tends towards the background intensity. In this synthetic example, μ is biased to high values even when it contains 250 days of data before the mainshock - presumably because there was a larger event just prior to the 250 day mark which is otherwise unaccounted for. This highlights an important operational consideration. In an evolving aftershock sequence, the intensity of events on any given day after the mainshock will tend to decay, and if we do not have a sufficiently long period of data prior to the mainshock to calibrate μ well, this decay in intensity as the sequence evolves will gradually draw the estimate of μ down as the sequence evolves - this means that the triggering parameters will also evolve to compensate for the bias. This would make it appear that the parameters are time dependent. This is problematic -

particularly since we know that in this example the parameters were actually fixed. Consequently, we recommend spending the time to constrain μ well either through an external constraint on the prior or through careful selection of the model domain. Conditioning on a history, has the potential to increase the stationarity of the analysis and account for triggers unobserved in the target catalogue, but cannot correct the background rate inaccuracies significantly. In total, the ETAS model contains 5 parameters. One background rate and four within the triggering function. If we can satisfactorily partition the background and triggered events, we then have the opportunity to resolve the tradeoffs between the triggering parameters.

3.3.2 Impact of combination of magnitudes and trade-off between K and α

There exists a clear trade-off between the two productivity parameters in the triggering component of the ETAS model. K describes the productivity at M_0, and α describes a magnitude dependent productivity for parent events with magnitudes greater than M_0.

Here, we explore the requirements for a catalogue to have sufficient information to resolve the tradeoff between α and K. Our hypothesis is that resolution of the magnitude dependence (i.e. α) in triggering requires a sufficient number of mainshocks of different magnitude in the training catalogue. This will inform what a sufficiently representative catalogue would look like if we expect to resolve all the parameters unbiasedly.

We investigate this effect using four synthetic earthquake catalogues, each spanning 5000 days (Fig. 9 - top). These catalogues each feature three mainshocks seeded on days 500, 2000, and 3500. The first catalogue has three mainshocks, each with a magnitude of 4; the second one has three mainshocks, each of magnitude 5; the third one has three mainshocks of magnitude 6; and the fourth one is a mix with magnitudes of 6, 5, and 4. In generating the synthetics, we intentionally selected catalogues which did not contain other very large events in the sequences so we could isolate the impact of the magnitudes we prescribed.

Whilst the catalogue with three $M6$ events contains the greatest number of events, we hypothesise that the catalogue containing three different $M6, M5, M4$ mainshock magnitudes will have the greatest power at resolving α and hence do the best job at resolving the tradeoff in the productivity parameters. Upon fitting the ETAS model to these data and analysing the posteriors (Fig. 9 - bottom), we can infer the following results:

(i) Although the catalogue with $3 \times M4$ better retrieves accurate background rates, it is less precise when estimating ETAS triggering parameters. This leads to broader posteriors, which exhibit biased estimates and higher uncertainty compared to other catalogues. Consequently, sequences with lower mainshock magnitudes and longer quiet periods provide better conditioning for the parameter μ, but not for the triggering parameters. As the mainshock magnitude increases, the posteriors for triggering parameters become tighter and exhibit less bias.

(ii) Comparing the catalogue that combines different mainshock magnitudes of $M6, M5, M4$ to the catalogues of $3 \times M6$ and $3 \times M5$, we find that the catalogue which varies the mainshock magnitudes provides more accurate estimates than those with identical mainshock magnitudes. Thus, resolving the $\alpha - K$ trade-off when there are not sequences of different sizes is more challenging. However, diversity in mainshock magnitudes allows for effective conditioning of α and K, even with fewer data points.

This tradeoff has an important consequence for operational earthquake forecasting. Since α controls the magnitude dependent productivity, a biased estimate means that the scaling of the number of triggered events from a future larger event could be significantly under- or over-estimated.

Where only a single sequence is studied, we should be conscious of this bias. To mitigate it, we should try to use more representative samples by increasing the size of the spatial-temporal domain. At the same time, in many regions, it would be impractical to draw a geographical and temporal boundary around a training catalogue that contains sufficient number of mainshocks of different magnitude for the catalogue to be truly representative. We recommend that practitioners need to actively recognise this tradeoff as part of their workflow, identify mitigating strategies where possible, and acknowledge the residual uncertainty if it is not possible to analyse a more representative catalogue.

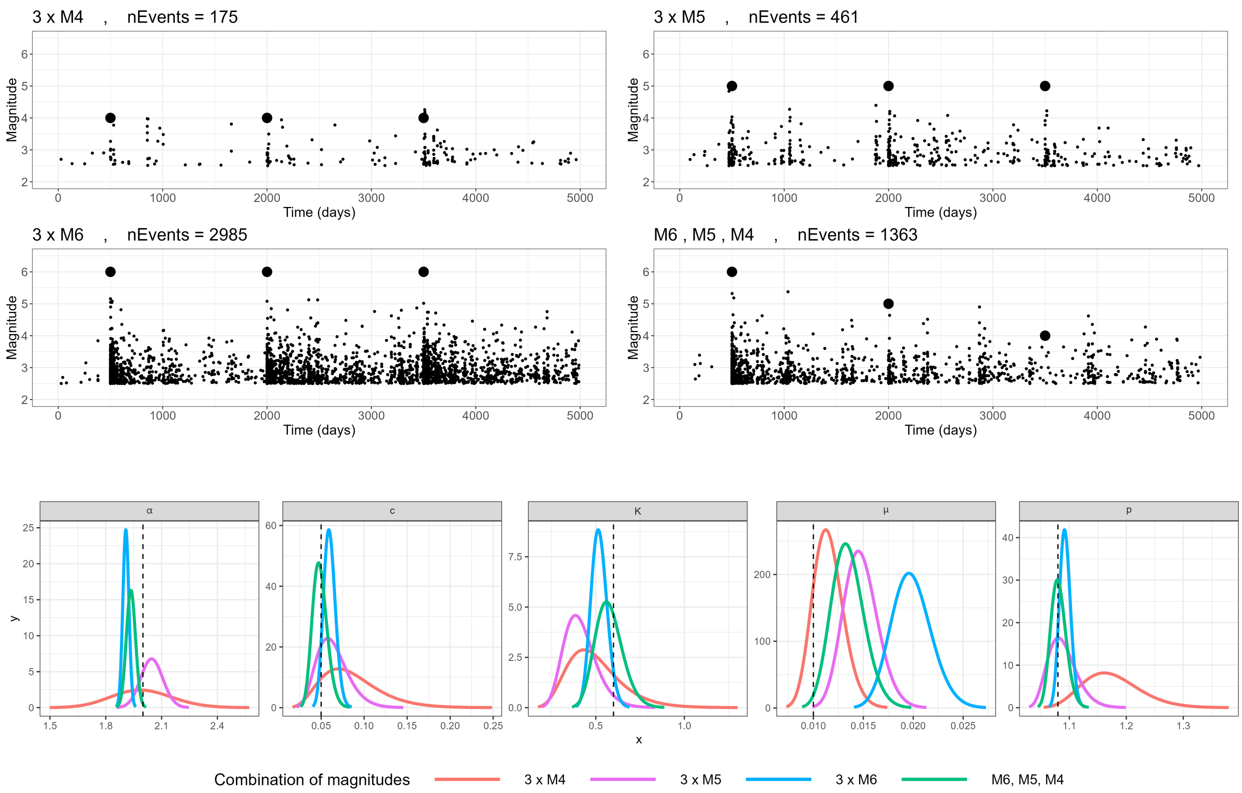

Figure 9. Four different simulated catalogues with different combination of magnitudes with $3 \times M4$, $3 \times M5$, $3 \times M6$, and combination of $M6$, $M5$, $M4$. Posteriors of each case are shown below the magnitude time series.

3.3.3 Impact of choice of incompleteness model parameters

In this section, we consider the consequences of mis-specifying the incompleteness model. Estimating $m_c(t)$ to be too high will reduce the total amount of reliable data from the inversion. Underestimating will mean that there remains some incomplete data, but perhaps we should expect asymptotic improvement as the completeness threshold is approached from either direction.

We investigate how varying the parameters G and H of the incompleteness model influences the posteriors of the modified ETAS model. To accomplish this, we generate a catalogue spanning a 1500-day period with a mainshock magnitude of 6.9 and apply truncation to create an incomplete catalogue, using $G = 3.8$ and $H = 1$ as the true incompleteness parameters. We then consider five distinct scenarios in which we fit the ETAS model with various choices for the incompleteness parameters. These scenarios include the use of the true parameters ($G = 3.8$, $H = 1$), inferred parameters 1 ($G = 3.1$, $H = 1$), inferred parameters 2 ($G = 4.1$, $H = 0.75$), inferred parameters 3 ($G = 4.5$, $H = 2$), and inferred parameters 4 ($G = 3.8$, $H = 2$). We deliberately select these parameter combinations to cover a variety of choices of the incompleteness model. This results in different decay patterns and different endpoints for the incompleteness period, as illustrated in Fig. 10 (top). Then, we fit our modified ETAS model to the incomplete catalogue considering the different incompleteness parameter sets. We also run the original ETAS model with the true incompleteness parameters.

Posteriors for all scenarios are shown in Fig. 10 (bottom). In this figure, the green posteriors, derived from the true incompleteness model, accurately capture the true ETAS values. The blue posteriors, corresponding to the 'inferred 1' scenario, wherein more events are eliminated compared to the green model, closely mirror the ETAS estimates of the 'true' model but exhibit somewhat shorter peaks. This indicates a similar yet slightly increased uncertainty, expected due to discarding a larger dataset above the actual incompleteness threshold, leading to precise yet more uncertain estimates. The amber posteriors from the 'inferred 4' scenario show slight deviations in estimates. Examining the purple ('inferred 2') and salmon ('inferred 3') scenarios, which significantly differ in incompleteness curves, show bias in the ETAS posteriors. Interestingly, all five scenarios with the modified ETAS models, including the extreme scenarios (purple and salmon),

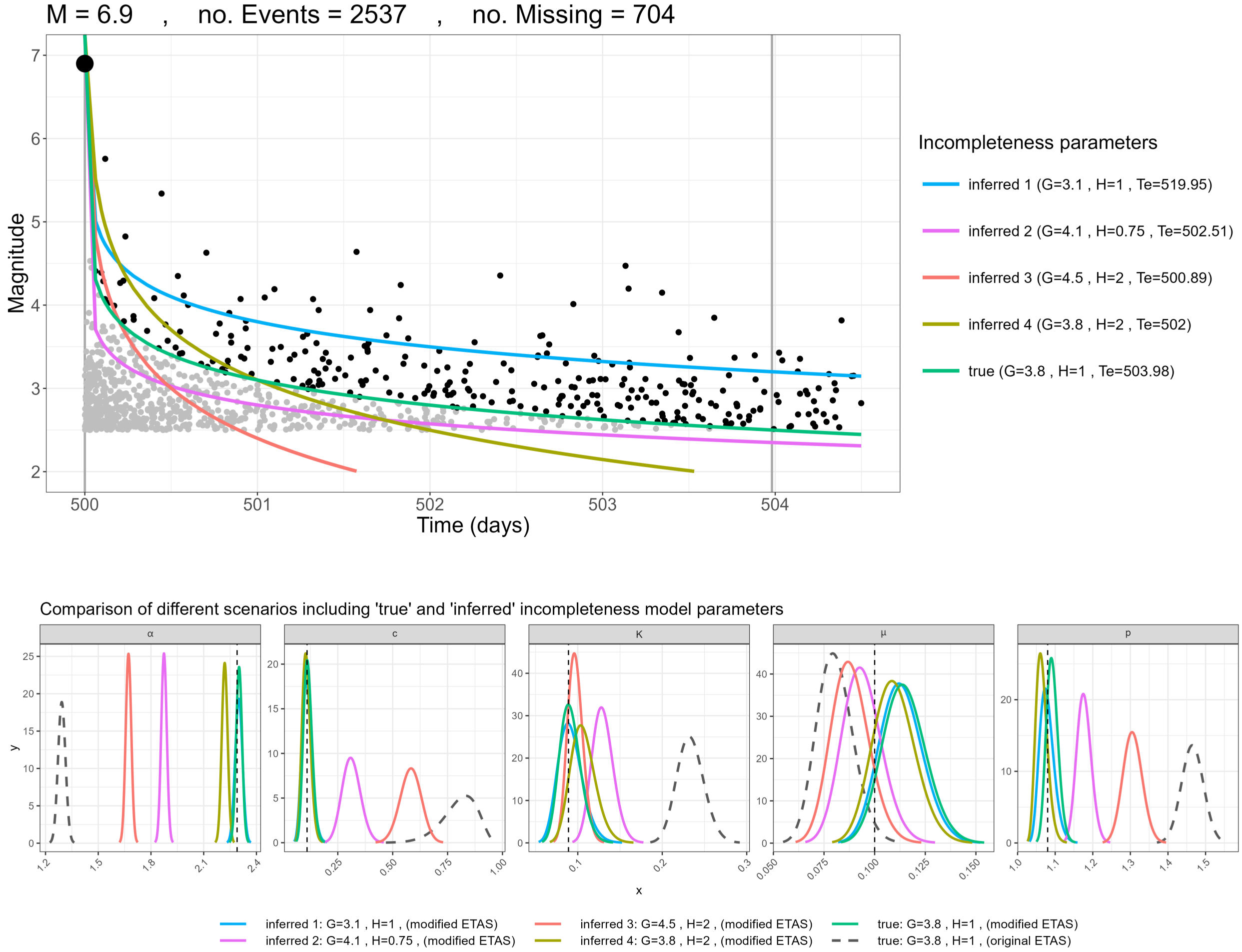

Figure 10. Impact of the choice of incompleteness model parameters on the ETAS estimates. (top): five incompleteness models with different decay curves and endpoints. The green curve represents the true incompleteness model, utilised for data truncation and preparation of the incomplete catalogue. The other four incompleteness models ('inferred 1' to 'inferred 4') are intentionally selected with visually different curves to assess the impact of the choice of incompleteness parameters. (bottom): Posteriors derived from the modified ETAS model using the five incompleteness scenarios, and posteriors obtained from the original ETAS model (in dashed dark grey) using the true incompleteness parameters.

still outperform the original ETAS model (dashed dark grey posteriors), calibrated with true incompleteness parameters.

3.4 Application to real earthquake case studies

We now apply both the original and the modified ETAS models to three real earthquake sequences: the 2016 M6.5 Amatrice earthquake in Italy, the 2017 M7.3 Kermanshah earthquake in Iran, and the 2019 M7.1 Ridgecrest earthquake in the US. To minimise bias and select a representative sample, we chose data with one year before and two years after the mainshock for the 2016 Amatrice and 2019 Ridgecrest earthquake sequences, and one year before and three years after the mainshock for the 2017 Kermanshah earthquake sequence. This choice provides a fair amount of background data and allows sufficient time for the sequences to return to background rates and decay from the triggering effects and it permits a comparison with the synthetic models we analysed in section 3.2. In addition, we also conduct our analysis without the one-year period prior to the mainshocks, where the modelling domain directly starts from the mainshcok event. This also allows a comparison with the synthetic models we analysed in section 3.3.1. The spatial domain for each case was also determined based on published shake maps and seismicity maps.

The time series of magnitudes and magnitude-event number plots for each earthquake are shown in Fig. 11. There are some positive and negative aspects of the nature of the sequences in terms of being able to compensate for the short-term censoring effects. For the 2016 Amatrice sequence, the pre-mainshock period has few events but looks reasonable for resolving background rate, and the sequence has nearly decayed back to this by the end. The sequence encompasses a good mix of magnitudes and only the strong period of incompleteness needs correcting. For the 2017 Kermanshah sequence, the pre-mainshock phase looks quite active, and this seismicity level is typical of the longer term within the Zagros mountains in Iran. There are also quite a few events at different magnitudes. Despite considering three years of data after the M.7.3 mainshock, the sequence is still active at the end so we expect that the pre-mainshock period dominates the estimate of μ. For the 2019 Ridgecrest sequence, there are not many events in the pre-mainshock phase, and there is also only one larger event in the aftershock sequence which might not be sufficient to resolve $K - \alpha$ tradeoff (as discussed in section 3.3.2).

To find the incompleteness model parameters for each sequence, we plot zoomed-in magnitude-time plots that clearly reveal the incompleteness in the early aftershock period. We fit the most appropriate model to the observed events, and obtain values of G and H for the 2016 Amatrice earthquake as $G = 5.45$ and $H = 1$, for the 2017 Kermanshah earthquake as $G = 5.5$ and $H = 1$, and for the 2019 Ridgecrest earthquake as $G = 5.8$ and $H = 1$, as shown in Fig. 12. Also, the b-value of each sequence is estimated using stability of b versus completeness threshold shown in Fig. 13. Subsequently, we apply both the original and the modified ETAS models to the earthquake catalogues. In Fig. 14, we present the posterior outcomes of both the modified (in green) and the original (in red) ETAS models. These models are trained on two distinct datasets: The first dataset comprises one-year of data prior to the mainshock, defining the modelling domain as $T_1 = T_m - 1$, with the corresponding results represented by solid lines. The second dataset involves modelling without pre-mainshock data, initiating directly from the mainshock event. In this case, the modelling domain is set as $T_1 = T_m$, and the results are illustrated using dashed lines. Additionally, we extract the triggering functions for both models, as illustrated in Fig. 15 which indicates higher rates and narrower plots (less uncertainty) for the modified model in comparison to the original version.

Since we do not have access to the true parameters for the real datasets, we can assess the performance of both original and modified models based on two key aspects: (1) visual inspection: This involves examining whether models trained on real data produce patterns of underestimation and overestimation similar to those observed in synthetic experiments. Here, we see that the results for the real sequences are consistent with changes seen in synthetics data. (2): goodness-of-fit metrics: utilising quantitative goodness-of-fit metrics can provide an additional measure of how well each model fits the real data. Here, we use the Widely Applicable Information Criterion (WAIC), also known as Watanabe–Akaike Information Criterion, (WAIC) score which is a measure used in statistical modelling, particularly in the context of Bayesian analysis, to assess the goodness of fit of a model to the data. It is used for model comparison, where lower WAIC values generally indicate a better-fitting model.

Table 4 displays the WAIC values for the three real sequences using both the original and the modified ETAS models. The modified model's lower WAIC values indicate better performance compared to the original model. Therefore, we can conclude that the modified ETAS model pro-

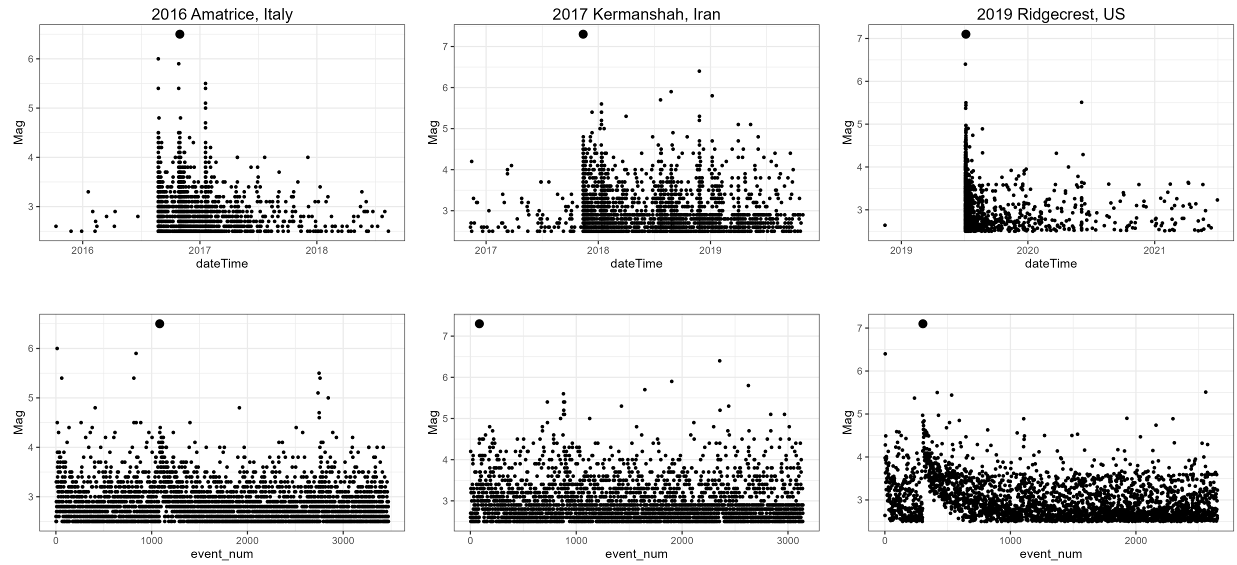

Figure 11. The three selected real earthquake sequences, including the 2016 Amatrice, Italy; the 2017 Kermanshah, Iran; and the 2019 Ridgecrest, US earthquakes. (top): magnitude - time plots and (bottom): magnitude - event number plots.

vides more reliable estimates of the ETAS parameters and offers a better representation of the underlying processes. Alternatively, the forecasting ability of each model could be examined using CSEP tests, though this falls beyond the scope of the current study.

In all cases correcting for short term incompleteness changes the estimates of the magnitude dependent productivity α. The background rate is consistent in all case studies, presumably because of the adequate sampling of the pre-mainshock period. The Ridgecrest data has the greatest short-term incompleteness and this propagates through to significant reductions in c and p; this is consistent with the changes seen in the synthetics in Fig. 4, so we believe the corrected estimates to be more reliable.

Table 4. WAIC score obtained from the original and the modified ETAS models trained on the three selected real earthquake sequences: the 2016 Amatrice, Italy; the 2017 Kermanshah, Iran; and the 2019 Ridgecrest, US.

ETAS version	WAIC		
	2016 Amatrice	2017 Kermanshah	2019 Ridgecrest
modified	13979.98	12694.9	10592.66
original	13997.91	12697.5	10659.72

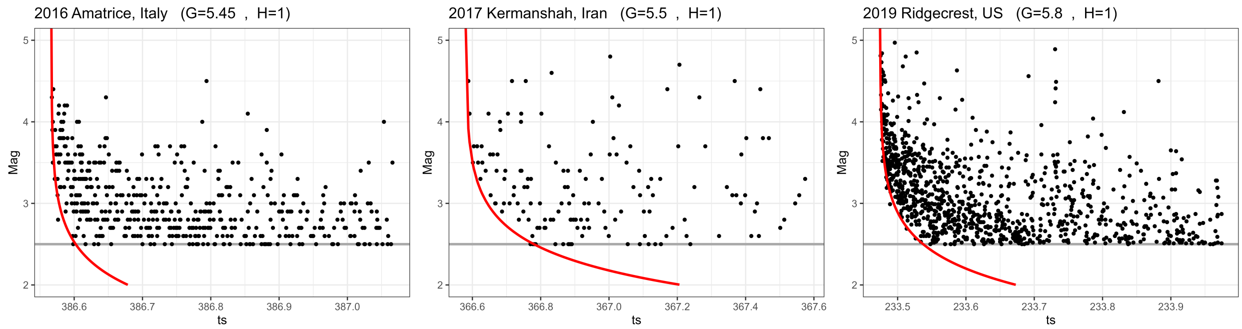

Figure 12. Extracting parameters G and H with the best fit of the incompleteness model to the three selected real earthquake sequences, including the 2016 Amatrice, Italy; the 2017 Kermanshah, Iran; and the 2019 Ridgecrest, US earthquakes.

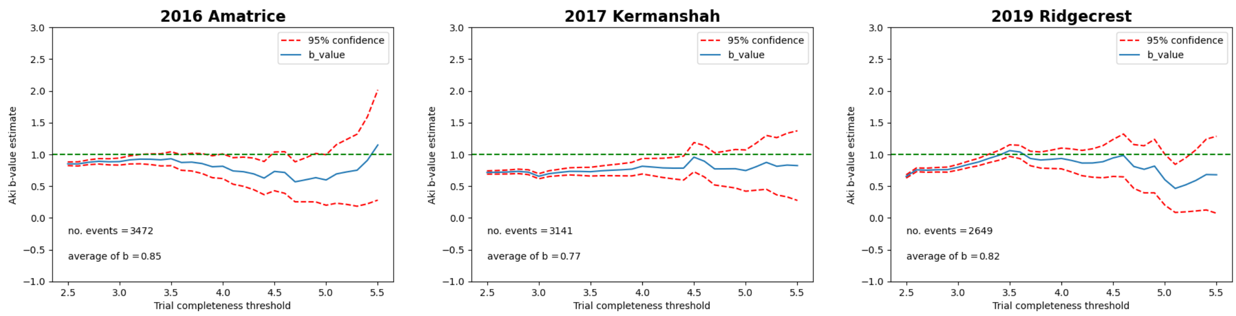

Figure 13. Estimation of b-value of the Gutenberg-Richter law for the three selected real earthquake sequences, including the 2016 Amatrice, Italy; the 2017 Kermanshah, Iran; and the 2019 Ridgecrest, US earthquakes.

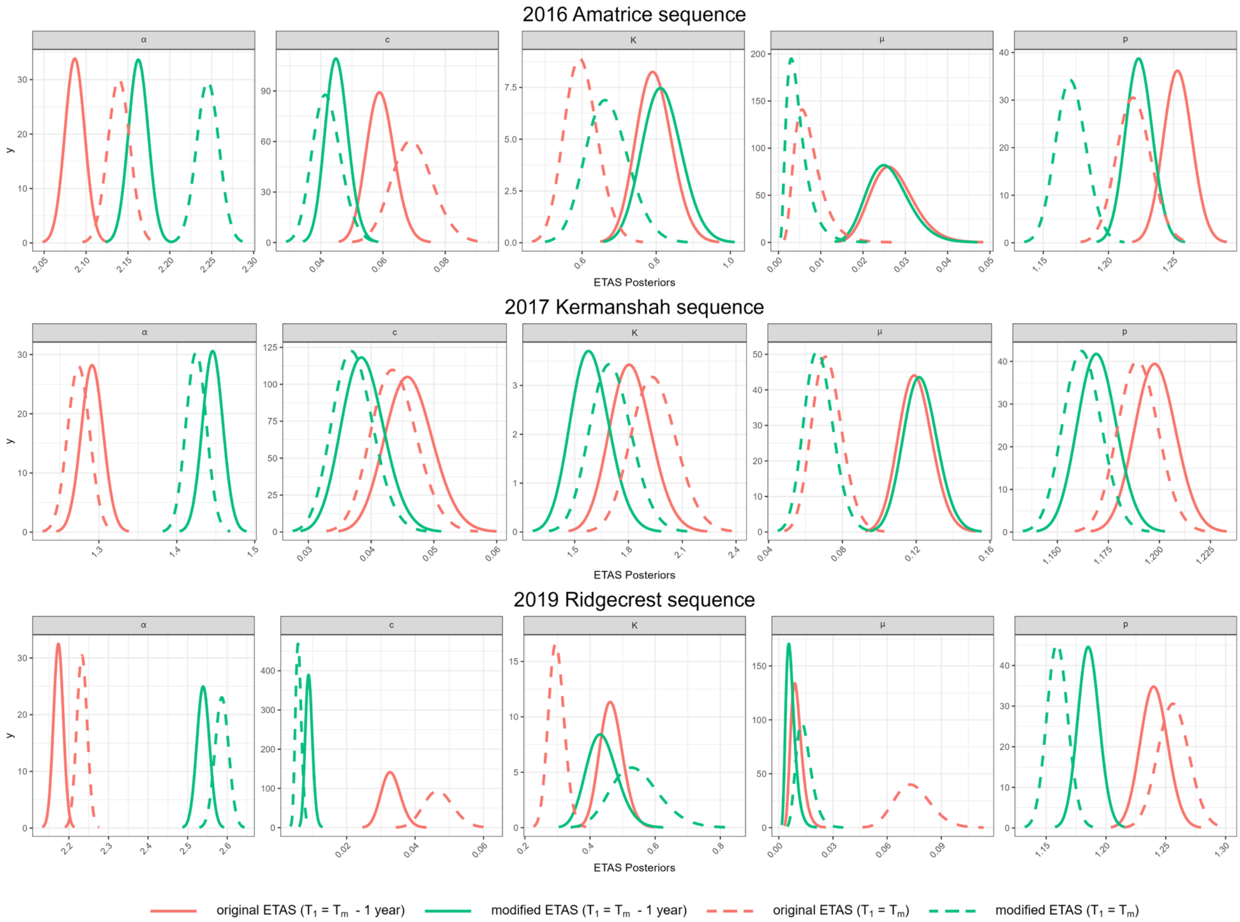

Figure 14. Posteriors of the three selected real earthquake sequences, including the 2016 Amatrice, Italy; the 2017 Kermanshah, Iran; and the 2019 Ridgecrest, US earthquakes. These posteriors are obtained using the modified (green) and the original (red) ETAS models, respectively. The models are trained on two different datasets: one with one year of data prior to the mainshock (solid lines), and the other without a pre-mainshock period, starting directly from the mainshock event (dashed lines).

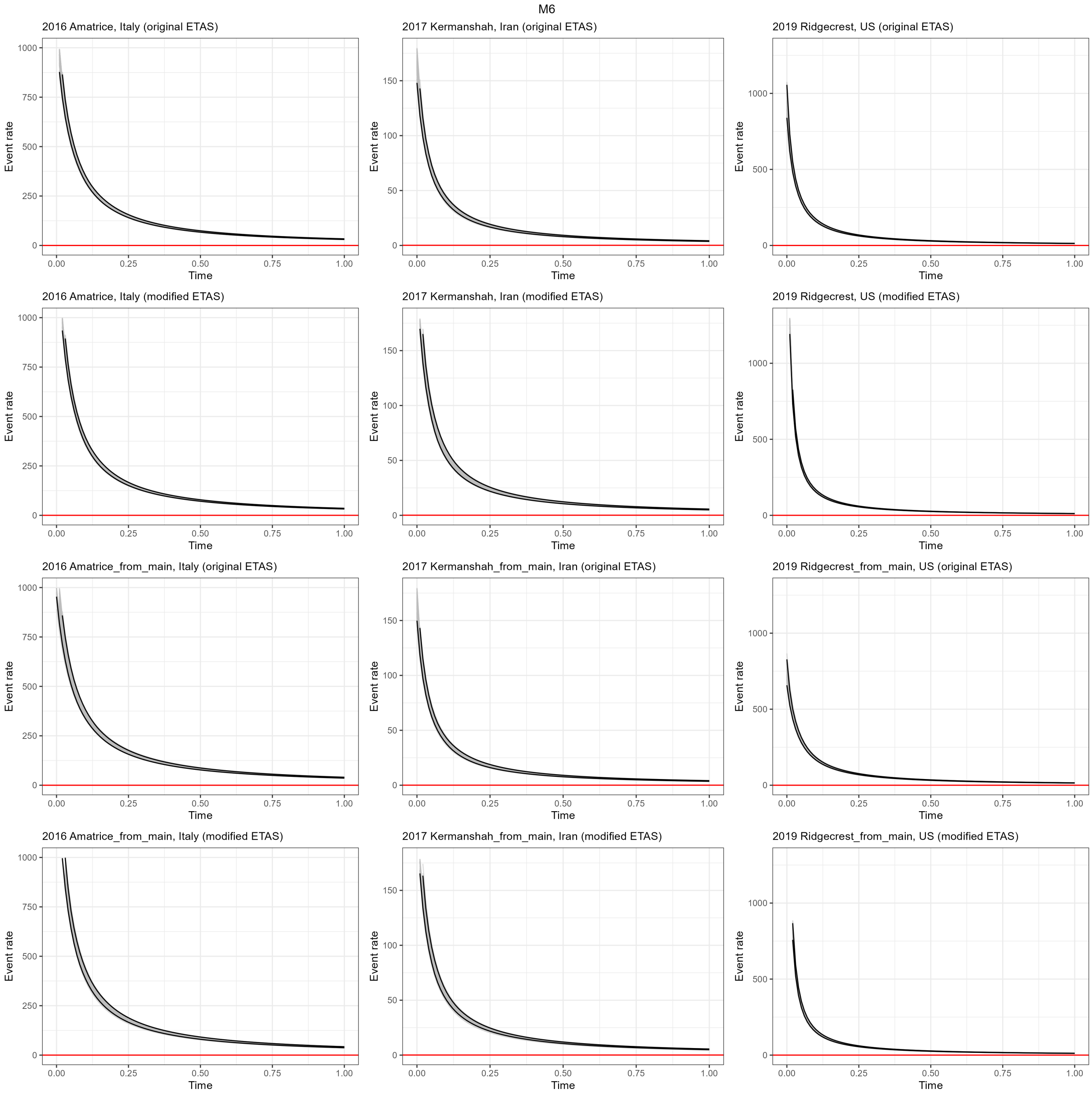

Figure 15. Triggering functions of the three selected real earthquake sequences, including the 2016 Amatrice, Italy; the 2017 Kermanshah, Iran; and the 2019 Ridgecrest, US earthquakes obtained from the original and the modified ETAS models.

4 REMAINING LIMITATIONS AND POSSIBLE IMPROVEMENTS

While our modified model to correct for temporal incompleteness has demonstrated its efficacy in enhancing the accuracy of ETAS estimates, ensuring generality, and mitigating bias in synthetic realisations where underlying parameters are known, there remain certain limitations that warrant attention in future developments.

In calculating the intensity at any point, we sum over all past events. However, we do not explicitly correct for the contributions to the intensity from events that lie below the time varying completeness threshold. It is likely that these contributions are in total small because they are inherently from the less productive, smaller magnitude aftershocks. This could be accounted for analytically using the assumptions presented in this paper, but since we recover the posteriors well with the current approach, we did not implement it here.

We also make the simplifying assumption that the background rate during the incompleteness period was not affected by censoring. Statistically, this is reasonable because the background rate is extremely small compared to the rate of triggered events during the period of temporal incompleteness. However, it is not strictly physically correct.

We were only able to consider a limited number of scenarios within this publication. We highly recommend others undertake a similar study calibrated to their setting in order to understand potential sources of bias and their implications when performing ETAS inversions and seismicity forecasts.

A further area of improvement could involve the exploration of alternative models for incompleteness. While the model introduced by Helmstetter et al. (2006) remains widely adopted, our experience indicates that its incorporation into the computational process of likelihood can be challenging, especially when resulting in the hypergeometric function. Seeking models with different functional forms, such as exponential or power-law, could offer a more intuitive or computationally efficient representation. Such enhancements may elevate both the performance and flexibility of the ETAS model.

In terms of technical advances, there exists potential to expand the consideration of incompleteness from purely temporal ETAS models to encompass spatio-temporal models as well. This would provide a more comprehensive view of earthquake dynamics, taking into account the incompleteness issue for both the timing and location of earthquake events, potentially enhancing predictive accuracy and offering a richer understanding of the underlying processes.

5 DISCUSSION

Here we have explored the impacts of different sources of bias and used this information to explore the data requirements for a training catalogue to be sufficiently representative of the governing processes that it can recover the key parameters unbiasedly.

The need for exploring these questions arises from the limitations of real datasets and the fact that an ETAS inversion will generally return a set of parameters but little information that helps us decide whether the training data was sufficient to produce unbiased estimates of these parameters in the first place. Consequently, it is easy to perform an inversion and unquestioningly work with the parameters that were returned.

We believe that we can do better than this, and the starting point needs to be understanding potential sources of bias within synthetic datasets such as those presented here and to actively consider sources of potential bias.

In practice, our datasets will always be limited by the seismic history of a region and the practicalities of defining a space-time-magnitude domain within it. However, we can question whether we have sufficient data to constrain key components of the ETAS model. For example,

- Do we have sufficient data from quieter periods to constrain μ? If not, we should anticipate that it may be biased. In our experience, when modelling productive sequences not including sufficient background will produce systematically high estimates of μ and consequently underestimations of the productivity within the triggered sequences.
- Do we have distinct mainshocks of different magnitudes in the training data? If not, even if μ is well calibrated, we should anticipate that the forecasts may do a bad job when scaling to future mainshocks of very different magnitude.
- Is there short-term incompleteness following large events within the catalogue? We have presented an innovative solution for dealing with this as a censoring problem. If there is such

incompleteness, and we do not correct for it, the synthetics shown here suggest we will both underestimate the background rate and underestimate the number of triggered events. Again, this would affect the performance of a prospective forecast.

- Is the short term incompleteness accurately modelled? The modified model performs well provided the time varying incompleteness threshold is reasonably estimated. As the threshold is reached from the incomplete side, it provides asymptotic improvement. If the threshold is estimated at a higher level than necessary, we still see good recovery of the true triggering parameters in the synthetics.

We hope that this study gives an intuitive indication of where bias may arise in ETAS inversions and how these biases would propagate through to systematic errors in operational earthquake forecasts. We believe that the analyses we have shown offer a way forwards for critiquing the performance of ETAS inversions and can help practitioners anticipate how they can better define model domains for extracting catalogues that are sufficiently representative for producing forecasts that lie within uncertainty of real evolving sequences. Our study offers a roadmap for future research in earthquake sequence modelling, promising improved seismic hazard analysis accuracy and a better understanding of earthquake behaviour.

6 CONCLUSION

In this study, we demonstrated the importance of accounting for short-term incompleteness in aftershock sequences, which can lead to biased ETAS estimates and inaccurate forecasts if not properly addressed. To mitigate this issue, we proposed a modified ETAS model that incorporates a correction for the short-term incompleteness, thereby improving the accuracy of the ETAS parameter estimates and enhancing forecast performance. To achieve this, we defined a censorship function and applied it to the inversion algorithm of the ETAS model.

Through a series of synthetic experiments, we have shown that the modified ETAS model yields more reliable parameter estimates compared to the original ETAS model. In addition, we investigated the impact of time binning strategy on model performance, the impact of conditioning model on the run-in history, the trade-off between productivity parameters K and α, as well as the impact of choice of incompleteness model parameters. These analyses demonstrated the robustness and reliability of the modified ETAS model across various synthetic earthquake scenarios, thereby contributing to a better understanding of its efficacy in real-world applications.

Subsequently, we applied both original and modified ETAS models to earthquake sequences from Amatrice, Italy (2016); Kermanshah, Iran (2017); and Ridgecrest, US (2019), covering a period of one year before and two/three years after each mainshock. We observed distinct regional seismicity patterns and triggering mechanisms, with the Kermanshah sequence remaining active well beyond the others, highlighting the enduring seismic influence in the Zagros mountains in Iran. Conversely, the Ridgecrest sequence presented challenges in resolving the ETAS model parameters due to its limited pre-mainshock events and aftershock data. We found that the results from the posteriors and triggering functions for real scenarios are consistent with the patterns and changes observed in synthetic data. Indicated by lower WAIC scores, the modified ETAS model shows better performance than the original model, suggesting its enhanced ability to better capture real seismic processes.

Finally, we identified remaining limitations and proposed potential avenues for future research. These include exploring alternative models for incompleteness, expanding the consideration of incompleteness to spatio-temporal models, and further investigating sources of bias in ETAS parameter estimates. By refining our understanding of earthquake dynamics and the factors influencing parameter estimation, we can advance the state-of-the-art in earthquake forecasting and contribute to better-informed decision making for earthquake risk reduction.

ACKNOWLEDGMENTS

Farnaz Kamranzad has jointly been funded by the UKRI GCRF Tomorrow's Cities Hub (grant no. NE/S009000/1) and the School of GeoSciences internal funding at the University of Edinburgh.

DATA AVAILABILITY

The datasets for the three real earthquake sequences used in this study are publicly available and were directly downloaded from their respective websites: The 2016 Amatrice earthquake dataset was sourced from Istituto Nazionale di Geofisica e Vulcanologia (INGV) Terremoti at http://terremoti.ingv.it/en. The 2017 Kermanshah earthquake dataset was obtained from the Iranian Seismological Centre (IRSC) at http://irsc.ut.ac.ir/bulletin.php. And the 2019 Ridgecrest earthquake dataset was retrieved from the U.S. Geological Survey (USGS) Earthquake Catalogue at https://earthquake.usgs.gov/earthquakes/search/, all last accessed in August 2023.

REFERENCES

Adelfio, G, and Chiodi, M., 2021. Including covariates in a space-time point process with application to seismicity, *Statistical Methods & Applications*, **30**, 947–971.

Bachl, F.E., Lindgren, F., and Borchers, D.L., and Illian, J.B., 2019. inlabru: an R package for Bayesian spatial modelling from ecological survey data, *Methods in Ecology and Evolution*, **10**, 760–766.

Båth, M., 1965. Lateral in homogeneities of the upper mantle, *Tectonophysics*, **2**, 483–514.

Cheysson, F., and Lang, G., 2022. Spectral estimation of Hawkes processes from count data, *Annals of Statistics*, **50**(3), 1722–1746.

Chiodi, M., Nicolis, O., Adelfio, G., D'Angelo, N., and Gonzàlez, A., 2021. ETAS Space–Time Modeling of Chile Triggered Seismicity Using Covariates: Some Preliminary Results, *Applied Sciences*, **11**(19), 13 pages.

de Arcangelis, L., Godano, C., and Lippiello, E., 2018. The Overlap of Aftershock Coda Waves and Short-Term Postseismic Forecasting, *Journal of Geophysical Research: Solid Earth*, **23**, 5661–5674.

Ebrahimian, H., and Jalayer, F., 2017. Robust seismicity forecasting based on Bayesian parameter estimation for epidemiological spatio-temporal aftershock clustering models, *Scientific Reports*, **7**(9803), 1–15.

Ebrahimian, H., Jalayer, F., Maleki Asayesh, B., Hainzl, S., and Zafarani, H., 2022. Improvements to seismicity forecasting based on a Bayesian spatio-temporal ETAS model, *Scientific Reports*, **12**(20970), 1–27.

Gardner, J. K., & Knopoff, L., 1974. Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?, *Bulletin of the Seismological Society of America*, **64**, 1363–1367.

Grimm, C., Kaeser, M., Hainzl, S., Pagani, M., and Kuechenhoff, H., 2021. Improving Earthquake Doublet Frequency Predictions by Modified Spatial Trigger Kernels in the Epidemic-Type Aftershock Sequence (ETAS) Model, *Bulletin of the Seismological Society of America*, **112**(1), 474–493.

Grimm, C., Hainzl, S., Kaeser, M., and Kuechenhoff, H., 2022. Solving three major biases of the ETAS model to improve forecasts of the 2019 Ridgecrest sequence, *Stochastic Environmental Research and Risk Assessment*, **36**, 2133–2152.

Gutenberg, B., and Richter, C.F., 1944. Frequency of earthquakes in California, *Bulletin of the Seismological Society of America*, **4**, 185–188.

Hainzl, S., 2016. Apparent triggering function of aftershocks resulting from rate-dependent incompleteness of earthquake catalogs, *Journal of Geophysical Research: Solid Earth*, **121**, 6499–6509.

Hainzl, S., 2021. ETAS-Approach Accounting for Short-Term Incompleteness of Earthquake Catalogs, *Bulletin of the Seismological Society of America*, **112**(1), 494–507.

Hardebeck, J., Llenos, A., Michael, A.J., Page, M.T., and van der Elst, N., 2019. Updated California Aftershock Parameters, *Seismological Research Letters*, **90**(1), 262–270.

Harte, D., 2013. Bias in fitting the ETAS model: a case study based on New Zealand seismicity, *Geophysical Journal International*, **192**, 390–412.

Harte, D., 2016. Model parameter estimation bias induced by earthquake magnitude cut-off, *Geophysical Journal International*, **204**, 1266–1287.

Helmstetter, A., Kagan, Y., and Jackson, D., 2006. Comparison of short-term and time-independent earthquake forecast models for southern California, *Bulletin of the Seismological Society of America*, **96**(1), 90–106.

Iacoletti, S., Cremen, G., and Galasso, C., 2022. Validation of the epidemic-type aftershock sequence (ETAS) models for simulation-based seismic hazard assessments, *Seismological Research Letters*, **93**(3), 1601–1618.

Jalilian, A., 2019, ETAS: An R Package for Fitting the Space-Time ETAS Model to Earthquake Data, *Journal of Statistical Software, Code Snippets*, **88**(1), 1–39.

Kanazawa, K., and Sornette, D., 2023, Asymptotic solutions to nonlinear Hawkes processes: A systematic classification of the steady-state solutions, *Physical Review Research*, **5**, 013067-1–46.

Kirchner, M., 2017, ETAS: An estimation procedure for the Hawkes process, *Quantitative Finance*, **17**(4), 571–595.

Laub, P.J., Lee, Y., and Taimre, T., 2021, The Elements of Hawkes Processes, *Springer*, 133 pages.

Lippiello, E., Cirillo, A., Godano, C., Papadimitriou, E., and Karakostas, V., 2019, Post Seismic Catalog Incompleteness and Aftershock Forecasting, *Geosciences*, **9**, 1–12.

Mignan, A., 2018, Utsu aftershock productivity law explained from geometric operations on the permanent static stress field of mainshocks, *Nonlinear Processes in Geophysics*, **25** (1), 241–250.

Mizrahi, L., Nandan, S., and Wiemer, S., 2021, Embracing Data Incompleteness for Better Earthquake Forecasting, *Journal of Geophysical Research: Solid Earth*, **126**(12), 1–26.

Mizrahi, L., Nandan, S., Savran, W., Wiemer, S., and Ben-Zion, Y., 2023, Question-Driven Ensembles of Flexible ETAS Models, *Seismological Research Letters*, **94**(2A), 829–843.

Molkenthin, C., Donner, C., Reich, S., Zöller, G., Hainzl, S., Holschneider, M., and Opper, M., 2022, GP-ETAS: semiparametric Bayesian inference for the spatio-temporal epidemic type aftershock sequence model, *Statistics and Computing*, **32**:29

Moradpour, J., Hainzl, S., and Davidsen, J., 2014, Nontrivial decay of aftershock density with distance in Southern California, *Journal of Geophysical Research: Solid Earth*, **119**, 5518–5535.

Muir, J.,and Ross, Z., 2023, A deep Gaussian process model for seismicity background rates, *Geophysical Journal International*, **234**(1), 427–438.

Naylor, M., Serafini, F., Lindgren, F. and Main, I., 2023, Bayesian modelling of the temporal evolution of seismicity using the ETAS.inlabru R-package, *Frontiers in Applied Mathematics and Statistics*, **9**, 1–19.

Nishikawa, T., and Nishimura, T., 2023, Development of an Epidemic-Type Aftershock-Sequence Model Explicitly Incorporating the Seismicity-Triggering Effects of Slow Slip Events, *Journal of Geophysical Research: Solid Earth*, **128**(5), 1–28.

Ogata, Y., 1988, Statistical models for earthquake occurrences and residual analysis for point processes, *Journal of the American Statistical Association*, **83**(401), 9–27.

Ogata, Y., 1998, Space-time point-process models for earthquake occurrences, *The Annals of the Institute of Statistical Mathematics*, **50**(2), 379–402.

Ogata, Y., and Zhuang, J., 2006, Space–time ETAS models and an improved extension, *Tectonophysics*, **413**, 13–23.

Ogata, Y., 2011, Significant improvements of the space-time ETAS model for forecasting of accurate baseline seismicitys, *Earth, Planets and Space*, **63**, 217–229.

Omi, T., Ogata, Y., Hirata, Y., and Aihara, K., 2014, Estimating the ETAS model from an early aftershock sequence, *Geophysical Research Letters*, **41**, 850–857.

Omi, T., Ogata, Y., Hirata, Y., and Aihara, K., 2015, Intermediate-term forecasting of aftershocks from an early aftershock sequence: Bayesian and ensemble forecasting approaches, *Journal of Geophysical Research: Solid Earth*, **120**(4), 2561–2578.

Omori, F., 1985, On after-shocks of earthquakes, *The journal of the College of Science, Imperial University of Tokyo*, **7**, 111–200.

Page, M., van der Elst, N., Hardebeck, J., Felzer, K., and Michael, A., 2016, Three Ingredients for Improved Global Aftershock Forecasts:Tectonic Region, Time-Dependent Catalog Incompleteness, and Intersequence Variability, *Bulletin of the Seismological Society of America*, **106**(5), 2290–2301.

Ross, G.J., 2021, Bayesian Estimation of the ETAS Model for Earthquake Occurrences, *Bulletin of the Seismological Society of America*, **111**(3), 1473–1480.

Ross, G.J. and Kolev, A., 2022, Semiparametric Bayesian forecasting of spatiotemporal earthquake occurrences, *The Annals of Applied Statistics*, **16**(4), 2083–2100.

Rue, H., Martino, Sara., and Chopin, N., 2009, Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *Journal of the royal statistical society: Series b (statistical methodology)*, **71**(2), 319–392.

Schneider, M I., and Guttorp, P., 2021, Bayesian ETAS: towards improved operational aftershock forecasting, *USGS report*.

Shcherbakov, R., Turcotte, D.Dl., and Rundle, J.B., 2004, A generalized Omori's law for earthquake aftershock decay, *Geophysical Research Letters* **31**(11), 1–5.

Shcherbakov, R., Zhuang, J., Zöller G., and Ogata, Y., 2019, Forecasting the magnitude of the largest expected earthquake, *nature communications* **10**(1), 1–11.

Shcherbakov, R., 2021, Statistics and Forecasting of Aftershocks During the 2019 Ridgecrest, California, Earthquake Sequence, *Journal of Geophysical Research: Solid Earth* **126**(2), 1–25.

Shebalin, P.N., Narteau, C., Baranov, S.V., 2020, Earthquake productivity law, *Geophysical Journal International* **222**, 1264–1269.

Seif, S., Mignan, A., Zechar, J., Werner, M., and Wiemer, S., 2017, Estimating ETAS: the effects of truncation, missing data, and model assumptions, *Journal of Geophysical Research: Solid Earth*, **122**, 449–469.

Shlomovich, L., Cohen, E., Adams, N., and Patel, L., 2022, Parameter Estimation of Binned Hawkes Processes, *Journal of Computational and Graphical Statistics*, **31**(4), 990–1000.

Stallone, A., and Falcone1, G., 2021, Missing Earthquake Data Reconstruction in the SpaceTime-Magnitude Domain, *Earth and Space Science*, **8**(8), 1–13.

Stindl, T., and Chen, F., 2023, EM algorithm for the estimation of the RETAS model, *Journal of Computational and Graphical Statistics*, DOI: 10.1080/10618600.2023.2253293.

Stockman, S., Lawson, D.J. and Werner, M.J., 2023, Forecasting the 2016–2017 central Apennines earthquake sequence with a neural point process, *Earth's Future*, DOI: 10.1029/2023EF003777.

Utsu, 1957. Magnitude of earthquakes and occurrence of their aftershocks, *Zisin2*. **10**(1), 35–45.

Utsu, 1972. Aftershocks and earthquake statistics (3): Analyses of the distribution of earthquakes in magnitude, time and space with special consideration to clustering characteristics of earthquake occurrence, *Journal of the Faculty of Science, Hokkaido University*. **3**(5), 379–441.

van der Elst, N.J, 2021. B-Positive: A Robust Estimator of Aftershock Magnitude Distribution in Transiently Incomplete Catalogs, *Journal of Geophysical Research: Solid Earth*. **126**, 1–19.

van der Elst, N.J, Hardebeck, J., Michael, A., McBride, S., and Vanacore, E., 2022. Prospective and Retrospective Evaluation of the U.S. Geological Survey Public Aftershock Forecast for the 2019–2021 Southwest Puerto Rico Earthquake and Aftershocks, *Seismological Research Letters*. **92**(2A), 620–640.

Veen, A., and Schoenberg, F.P, 2008. Estimation of Space–Time Branching Process Models in Seismology Using an EM–Type Algorithm, *Journal of the American Statistical Association*. **103**(482), 614–624.

Wolfram—Alpha, 2023. *Wolfram Alpha LLC*. https://www.wolframalpha.com/input?i=

Integrate%5B%28%28%28%28t-a%29%2Fc%29%2B1%29%5E%28-p%29%29+*+%28%28t-T%29%5E%28b*H%29%29+%2C+t%5D, (last accessed 05 June 2023).

Zhang, L., Werner, M.J., and Goda, K., 2018. Spatiotemporal Seismic Hazard and Risk Assessment of Aftershocks of M 9 Megathrust Earthquakes, *Bulletin of the Seismological Society of America*. **108**(6), 3313–3335.

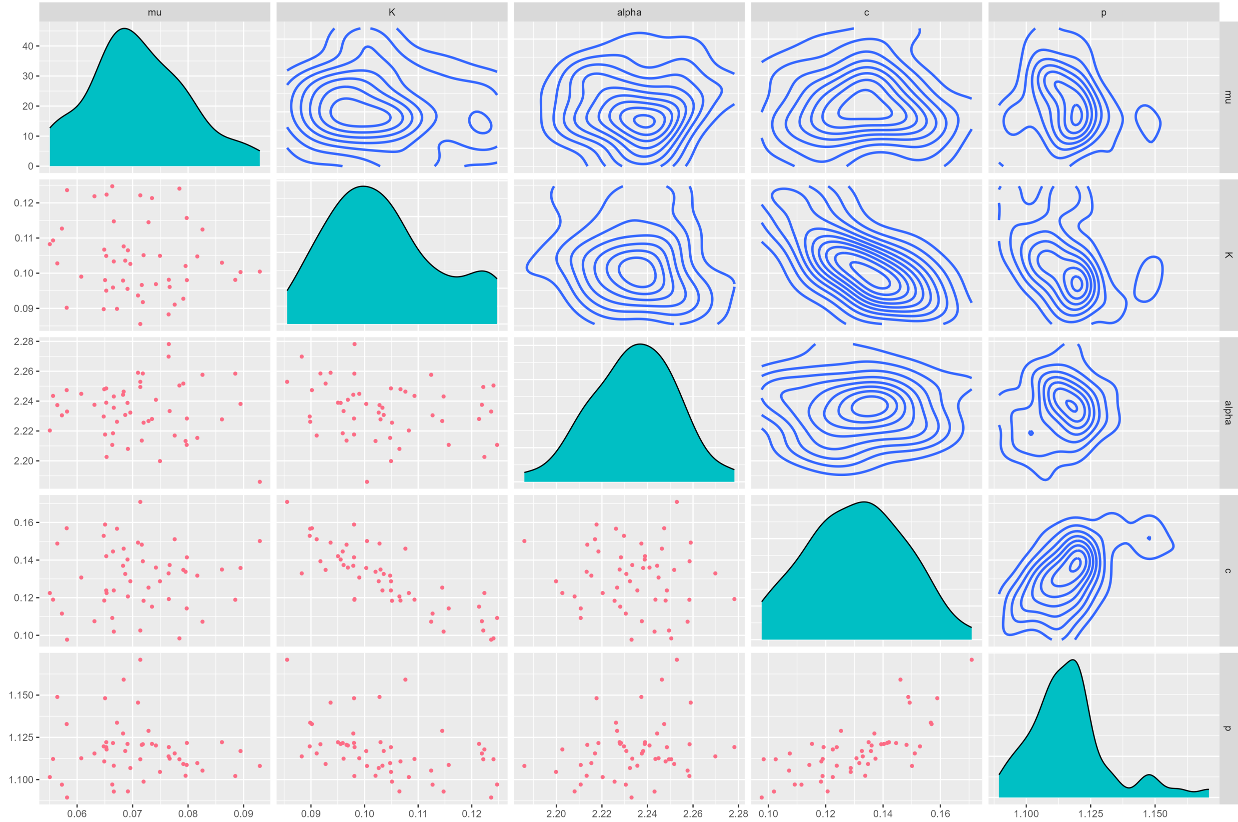

Figure A1. Pair plot of ETAS parameters

933 # APPENDIX A: EXTRA PLOTS

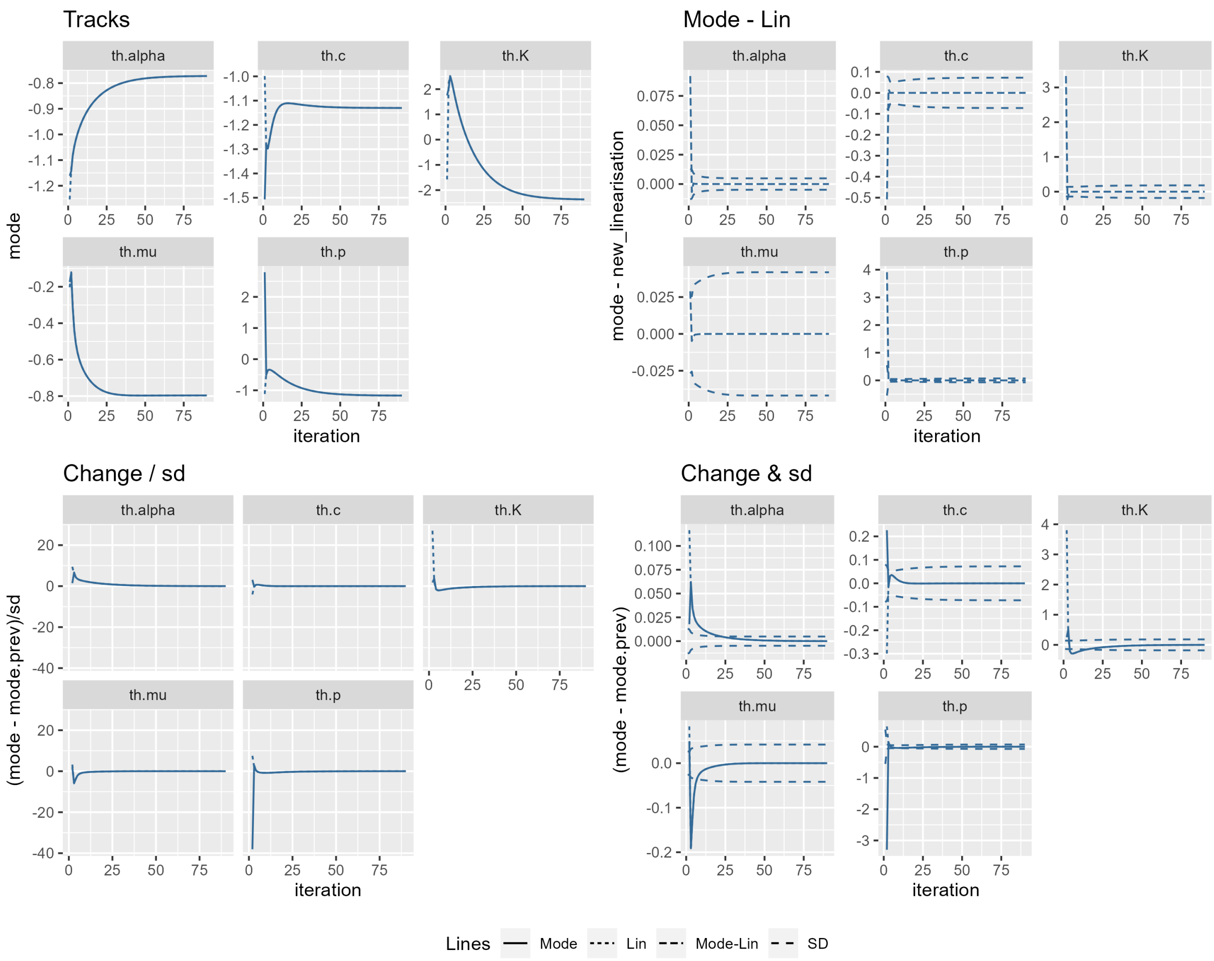

Figure A2. Convergence plot for the original ETAS model trained on complete data

Table A1. Glossary and description of variables and parameters.

in text	**in codes**	**description**
t	t	a time point at which we evaluate the intensity
t_i	th	time of the i^{th} triggering event
$t_j^{b_i}$	T1b	start time of j^{th} bin of the i^{th} triggering event (left edge)
$t_{j+1}^{b_i}$	T2b	end time of j^{th} bin of the i^{th} triggering event (right edge)
T_1	T1	start time of the modelling domain
T_2	T2	end time of the modelling domain
$\max(T_1, t_i)$	T.l	time for either 'including' or 'conditioning' on history
T_m	Tm	time of the mainshock (also start time of the incompleteness interval)
T_e	Te	end time of the incompleteness interval
c	c	time shift to avoid infinity at $t = t_i$
m_i	mh	magnitude of the i^{th} triggering event
M_m	Mm	magnitude of the mainshock
M_0	M0	fixed magnitude of incompleteness
$m_c(t)$	mct	short-term time-varying magnitude of incompleteness
μ	mu	background seismicity
K	K	base productivity parameter
α	alpha	magnitude scaling productivity parameter
p	p	decay speed of aftershock rates
b	b	b-value of the Gutenberg-Richter relation
G	G	baseline magnitude shift in incompleteness model
H	H	log-time scaling parameter in incompleteness model
λ	cond.lambda	intensity or rate of aftershocks
$g(t)$	gt	ETAS triggering function
$\mathcal{H}$	Ht	time history of aftershock evolution
$\mathcal{L}$	predictor.fun	likelihood of model
θ	theta	vector of all model parameters
Δ	delta.t	base time increment in time binning
δ	coef.t	growth factor in time binning
n_{max}	Nmax	maximum number of bins for each event in temporal mesh

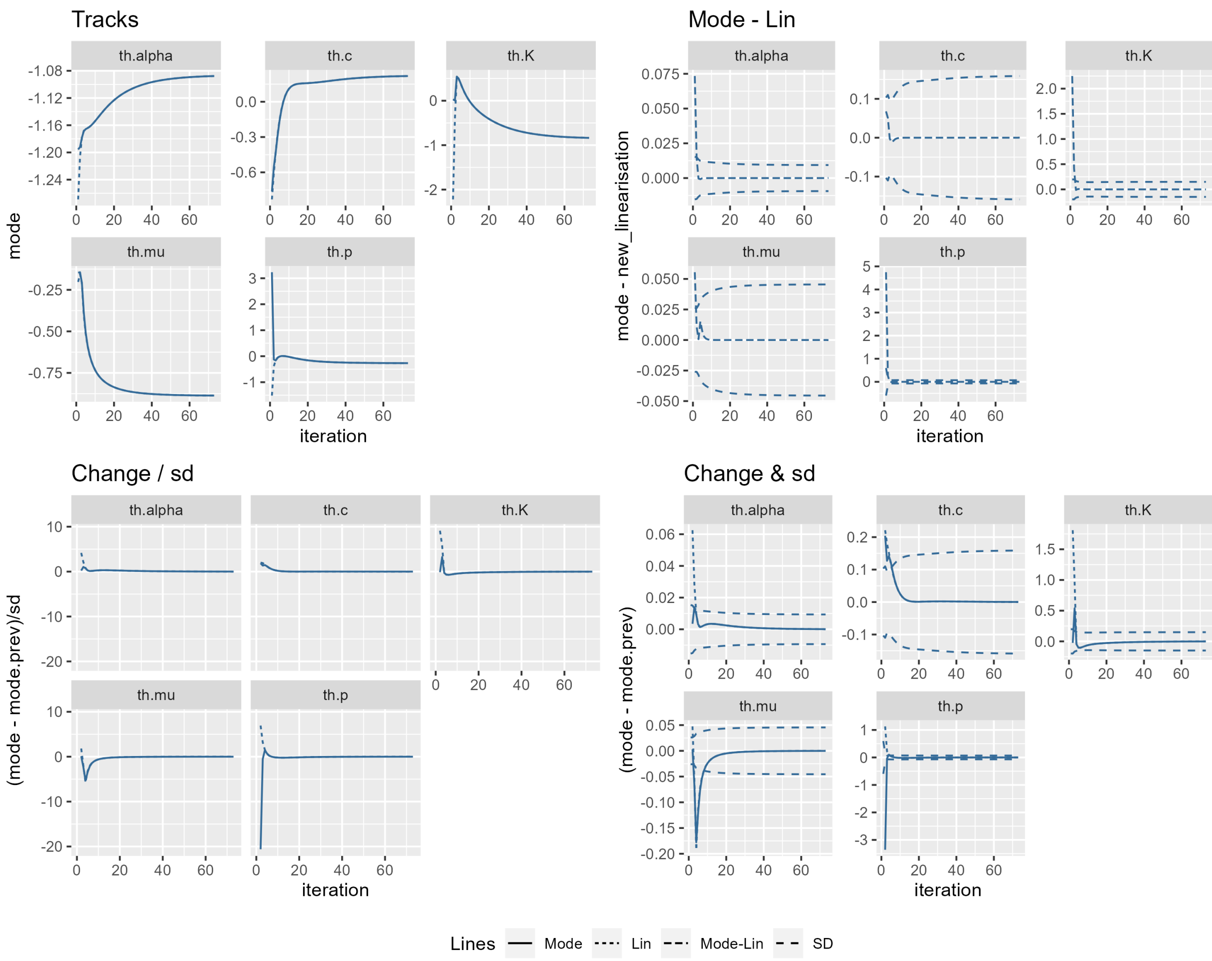

Figure A3. Convergence plot for the original ETAS model trained on incomplete data

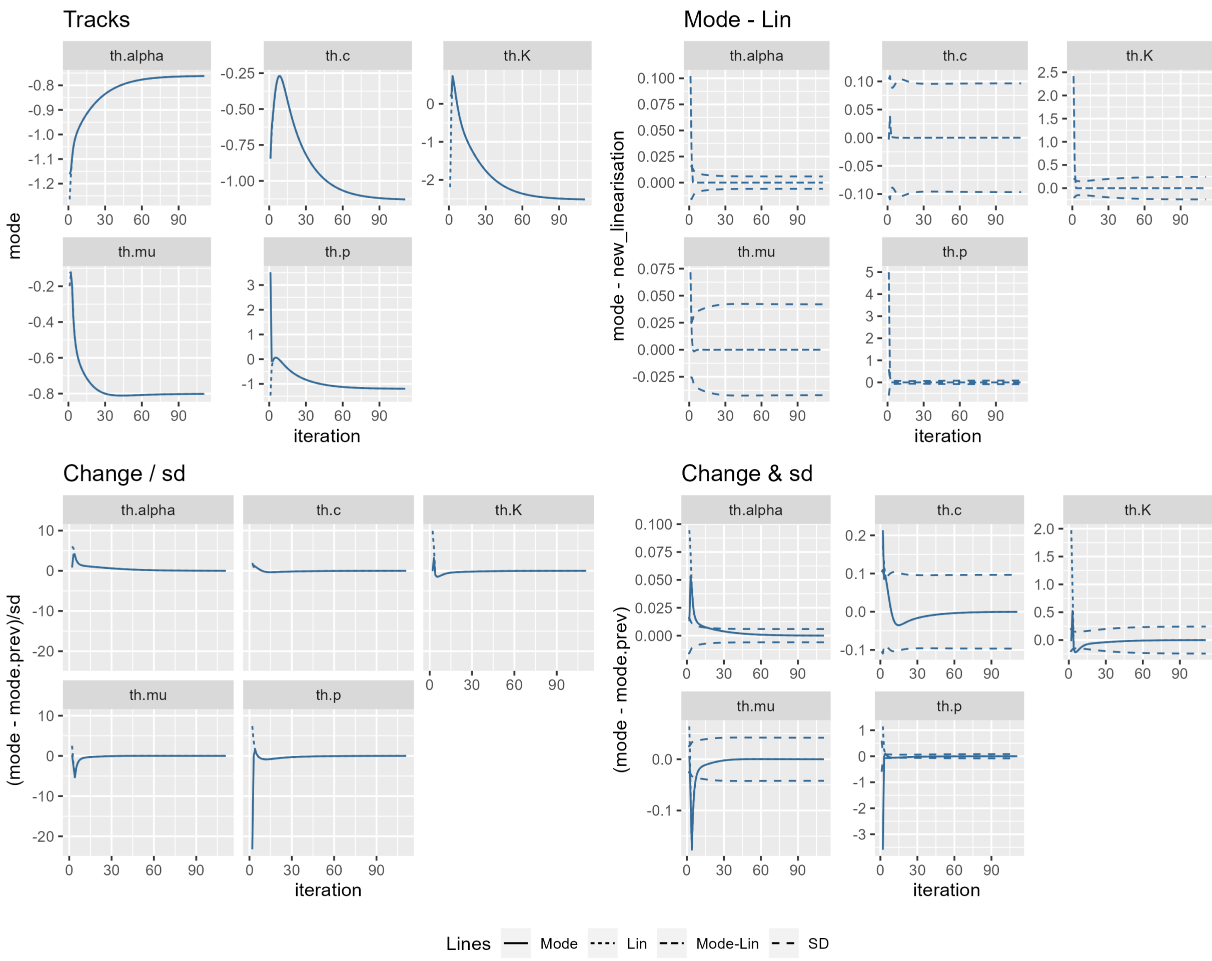

Figure A4. Convergence plot for the modified ETAS model trained on incomplete data