THE UNIVERSITY *of* EDINBURGH

The University of Edinburgh
School of GeoSciences

Resolving bias due to short- and long-term catalogue incompleteness and improving accuracy by optimal sampling in the Epidemic-Type Aftershock Sequence (ETAS) model

Farnaz Kamranzad

Thesis submitted for the degree of Doctor of Philosophy

2024

Abstract

Earthquake catalogues are fundamental tools for understanding seismic processes and forecasting future events. However, these catalogues are often associated with different degrees of incompleteness, including data gaps immediately after the occurrence of large earthquakes due to waveform overlaps and seismogram saturation, as well as long-term gaps arising from technological and logistical limitations in the development of seismic networks. Such incompleteness in data can lead to substantial errors and biases in earthquake models and therefore distort our understanding of seismicity behaviour in a region. Furthermore, statistical modelling is highly sensitive to how we select data samples from these catalogues, and this can also significantly impact the estimations of seismicity model parameters.

This thesis presents significant advances in the field of seismicity modelling by enhancing the Epidemic-Type Aftershock Sequence (ETAS) model. The primary goal is to address critical issues of temporal incompleteness and systematic biases that affect the estimation of ETAS parameters, which are essential for accurate earthquake forecasting and seismic hazard assessments.

In this thesis, I first address the issue of short-term incompleteness commonly observed following large mainshocks. This period of incompleteness is critical because it involves the initial aftershock sequence, which can provide valuable insight into the seismicity rates and properties of a seismic sequence. I introduce a methodological enhancement to the inversion algorithm of the ETAS model, which enables it to effectively handle incomplete data during these crucial early stages. Theoretically, this adjustment involves defining a censorship function and integrating it into the ETAS conditional intensity and likelihood functions to create a modified inversion solution. For model implementation, I use a Bayesian framework with the inlabru package, which

leverages the Integrated Nested Laplace Approximation (INLA) method to provide posterior distributions of the model parameters, rather than conventional point estimates. The performance of the modified ETAS model is extensively tested through synthetic experiments designed to simulate realistic aftershock sequences with short-term incompleteness, demonstrating its ability to accurately capture ETAS parameters and actual aftershock rates even with significant data gaps.

Further, I explore optimising the selection of representative samples for the ETAS inversions, which is critical for reducing bias in parameter estimation. Various sampling strategies and their potential biases are examined, proposing a comprehensive approach to optimise survey design. This includes evaluating the sensitivity of the ETAS model to temporal binning strategies, conditioning the model on the run-in history before an earthquake sequence, the role of combination of different earthquake magnitudes, and the trade-offs between ETAS productivity parameters. I also consider the choice of incompleteness model parameters, the impact of secondary large aftershocks, and the spatial and temporal size of the modelling domain. By systematically analysing these factors, I establish guidelines that identify and minimise biases and enhance the reliability and robustness of modelling seismicity patterns when fitting the ETAS model to real earthquake data.

Finally, I expand the model's applicability to address long-term data incompleteness, which stems from sparse network coverage and technological limitations over extended periods. This comprehensive approach significantly improves the predictive accuracy of the ETAS model, as evidenced by its application to both simulated data and real earthquake sequences. By applying the same censorship approach used for addressing short-term incompleteness, I extend the model to handle incomplete long-term data, such as century-long records from the instrumental era, that also lack distinct aftershock sequences. As a result, with this new generic framework, the ETAS model adapts to varying degrees of data completeness, ensuring robust parameter estimation even when faced with extensive temporal gaps. The enhancements to the ETAS model introduced in this thesis significantly strengthen its theoretical foundations and practical utility in delivering more reliable and flexible operational earthquake forecasts and seismic hazard assessments.

Lay Summary

Earthquake catalogues are crucial for forecasting earthquakes and understanding their hazards. However, these catalogues often have gaps, especially after a large earthquake or due to long-term technological and logistical limitations. These gaps can lead to errors in earthquake models and skew our understanding of how earthquakes behave.

My Ph.D. thesis focuses on improving earthquake models by enhancing a specific model called the Epidemic-Type Aftershock Sequence (ETAS) model. The main goal is to tackle the issues caused by missing data and biases that affect the accuracy of this model, which is essential for predicting earthquakes and assessing seismic hazards.

Firstly, I address the problem of missing data right after a large earthquake. This period is critical because the initial aftershock sequence provides valuable information on earthquake rates. I propose a new method to improve the ETAS model's ability to handle incomplete data during these crucial early stages. This involves modifying the model's inversion algorithm to account for missing data and using a statistical technique to provide more accurate estimates of the model's parameters. I test this improved model with simulated aftershock sequences and show that it can accurately capture earthquake patterns even with significant data gaps in early aftershocks.

Next, I explore how to select the best data samples for the ETAS model to reduce bias in parameter estimation. I examine different sampling strategies and propose an approach to optimise survey design. This includes looking at how the model reacts to different ways of grouping data, considering the history of earthquakes before a sequence, combining data from different earthquake magnitudes, and balancing various factors affecting the model's parameters. By systematically analysing these factors, I provide guidelines to identify and minimise biases, making the model more reliable

when applied to real earthquake data.

Finally, I extend the model to handle long-term missing data, such as century-long records with sparse network coverage. Using the same approach for short-term missing data, I adapt the model to deal with long-term gaps, ensuring robust parameter estimation even with extensive temporal gaps. This comprehensive approach significantly improves the predictive accuracy of the ETAS model, as demonstrated by its application to both simulated and real earthquake sequences.

In summary, the enhancements introduced in my thesis strengthen the ETAS model, making it more reliable and flexible for operational earthquake forecasts and seismic hazard assessments. This work contributes to a better understanding of earthquakes and improves our ability to forecast future events, ultimately helping to mitigate the impact of earthquakes on society.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Mark Naylor and Ian Main. Thank you, Mark, for your profound knowledge and enthusiasm, which have been a constant source of motivation and have greatly enriched my research experience. Thank you, Ian, for your detailed feedback, guidance, and support that have helped shape this work into what it is today.

I extend my thanks to Finn Lindgren, not just for tackling maths problems with me, but for fostering a collaborative spirit that led to our joint publications.

I am very grateful to my viva examiners, Stefan Wiemer and Kathy Whaler, for the interesting discussions and for their valuable comments and constructive feedback.

Special thanks go to my advisor, Lara Kalnins, who provided invaluable support throughout this journey.

I would like to thank Kirsty Bayliss, whose friendship and encouragement have meant so much along the way. Thanks also go to Ester Manganiello, who came to Edinburgh as a short-term visiting student from Naples and became a dear friend, leaving behind nice memories and deep appreciation for her kindness and friendship.

I also want to thank my wonderful friends in our office, Anya, Ashrika, Calum, Conner, Eliot, Henry, Himanshu, Morag, Nick, Philipp, Prakash and Saraswati, who made this journey so much more enjoyable.

Last but not least, my heartfelt thanks go to my lovely family for always believing in me and for their encouragement and support of my academic dreams.

Thank you all for your support and for making this journey unforgettable.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Farnaz Kamranzad

Contents

Chapter 1

Introduction

1.1 An overview and problem statement

Seismicity models, which statistically describe the distribution of earthquakes in space, time, and magnitude, provide a fundamental basis for operational earthquake forecasting (OEF) and probabilistic seismic hazard assessments (PSHA) (Marzocchi and Jordan 2018), which are essential tools for assessing earthquake risk and improving public safety. OEF refers to the practice of providing real-time forecasts of earthquake probabilities within a specific time frame and region, thus aiding disaster preparation and response (Jordan and Jones 2010; Jordan et al. 2011; Jordan et al. 2014; Field et al. 2016; Schorlemmer et al. 2018). PSHA, on the other hand, involves quantifying the likelihood that various levels of earthquake-induced ground shaking occur at a site over a given period, which is crucial for designing earthquake-resistant structures and mitigating potential damage (Cornell 1968; McGuire and Arabasz 1990; Bommer 2002; Baker et al. 2021). A reliable seismicity model must accurately represent the spatio-temporal evolution, occurrence patterns, clustering, and characteristics of earthquakes within fault systems (Guo and Ogata 1997; Essing and Poli 2022). Despite significant advancements in this field, several gaps, challenges, and limitations persist. This section outlines these issues, setting the stage for the research presented in this thesis.

Earthquake catalogues form the backbone of seismicity modelling and provide valuable data on the occurrence and characteristics of seismic events (Chen et al. 1998; Weatherill et al. 2016). Such data usually contain two types of seismicity behaviour:

(1) a long-term background which consists of independent earthquakes that form the general seismicity rate in a region over extended periods, reflecting the tectonic loading processes and fault interactions; and (2) a short-term clustering part which consists of groups of earthquakes that occur dependently in the form of aftershock sequences, often triggered by a mainshock. Earthquake catalogues are necessary for researchers who want to understand the temporal and spatial patterns of earthquakes (Nievas et al. 2020). However, the quality and completeness of these data are often compromised by various factors. These limitations not only hinder the accuracy of seismicity models, but also affect the reliability of earthquake forecasts and seismic hazard models.

One major challenge in seismicity modelling is the temporal incompleteness in earthquake catalogues. This temporal incompleteness can be short-term or long-term. In the short term, incompleteness usually occurs after large mainshocks when overlapping seismic waves from numerous aftershocks saturate seismic networks, making it difficult to detect smaller events (Kagan 2004; Omi et al. 2016; Bachura et al. 2021). This leads to under-reporting of early aftershocks and biases the data (Hardebeck et al. 2024). In the long term, there is a step-like evolution in the development of seismic networks, leading to an increase in recorded smaller earthquakes. Improvements in seismic network coverage and technology mean that older data are less complete than recent records, creating inconsistencies that challenge long-term seismicity modelling (Arrowsmith et al. 2022). In addition to types of temporal incompleteness, the choice of seismicity datasets used to develop and fit our statistical models can also significantly impact their accuracy, depending on how representative the datasets are.

Widely used for modelling aftershock sequences, the Epidemic-Type Aftershock Sequence (ETAS) model (Ogata 1988; Ogata 1998) is one of the most effective temporal and spatio-temporal seismicity models. The ETAS model relies on earthquake catalogues to estimate seismicity behaviour parameters such as background seismicity rate, aftershock productivity, and the decay rate of aftershock activity over time and space. The existence of temporal incompleteness in earthquake catalogues can lead to significant errors in the ETAS parameter estimates, reducing the model's reliability and predictive skill (Naylor et al. 2023). Another challenge is optimising the selection of representative samples for ETAS model inversions. Different sampling strategies

can introduce biases that affect the accuracy of parameter estimates. For instance, the amount of historical seismicity data prior to earthquake sequences, or the spatio-temporal size of modelling domain, can influence the results of ETAS model inversions. In practice, it is easy to define the spatio-temporal domain without undertaking a critical assessment of how the domain affects the ETAS parameters and the epistemic uncertainty associated with this choice. So, developing an adaptive model capable of dealing with different types of incompleteness while also identifying optimal strategies to minimise biases in estimations is crucial to improving the reliability of ETAS models.

Current approaches to addressing temporal incompleteness and biases in seismicity modelling often fall short. Traditional methods typically involve either excluding incomplete data or attempting to fill in the gaps through interpolation or other estimation techniques. However, these methods can introduce additional biases or inaccuracies, as they rely on assumptions about missing data that may not hold true. Furthermore, many existing models do not adequately account for the uncertainties inherent in sparse or incomplete data. This can lead to overconfidence in model predictions, which can be particularly problematic in seismic hazard assessments, where the consequences of inaccurate forecasts can be severe. In this thesis, I develop a novel method to correct ETAS parameterisations by explicitly defining the censoring process and modifying the inversion algorithm through the application of a temporal censorship function. In this context, censoring refers to the omission or partial observation of data, particularly aftershock events, because smaller seismic events can go undetected due to limitations like overlapping seismic waves or saturated signals following large earthquakes. The censorship function that I define accounts for these missing or incomplete data points by incorporating a mathematical adjustment within the ETAS model, which helps compensate for gaps in the earthquake catalogue. By incorporating this process into the ETAS model, the method aims to address inherent biases and uncertainties more effectively than traditional approaches. Simultaneously, a contemporaneous study by Li et al. (2024) also tackles censorship bias by analysing how the apparent branching ratio changes with different cutoff magnitudes. The branching ratio in the ETAS model refers to the proportion of earthquakes that are aftershocks, triggered by previous events, as opposed to independent background seismicity. It is called 'apparent' because it is an observed ratio that depends on the detection thresh-

old of earthquakes. As cutoff magnitudes vary, the ratio can change, reflecting how many smaller undetected aftershocks might be missing from the data. This indicates the ongoing advances in the field and the relevance of addressing censorship bias. In addition to the theoretical modifications, I will apply my approach to understand its effects on ETAS parameters, implications for earthquake triggering, and improvements in model interpretations and bias reduction. This new censoring approach will be critically evaluated to improve the practical application of the ETAS model. These topics will be explored in more specific detail in research Chapters 4 to 6.

1.2 The claim

The overarching scope of this thesis is to enhance ETAS modelling by addressing critical issues of temporal incompleteness and systematic biases. By improving the accuracy and reliability of the ETAS model, this research aims to contribute significantly to better seismicity modelling, operational forecasting, and seismic hazard assessment. Here, I outline the key research questions, the overall aim, and the specific objectives that guide this research.

1.2.1 Central questions

This research is driven by several key questions:

1. How can the ETAS model be modified to better handle short-term incompleteness after large earthquakes?
2. What are the optimal sampling strategies for reducing biases when fitting the ETAS model?
3. Can we use longer-term earthquake data to fit the ETAS model? If so, how can long-term incompleteness in catalogues be addressed to improve the reliability of the ETAS model?
4. What are the benefits of integrating a Bayesian framework into the ETAS model?
5. How do the proposed methodological enhancements improve the predictive accuracy and reliability of the ETAS model in both synthetic and real-world applications?

1.2.2 Overall aim

The primary goal of this research is to develop and validate a methodological enhancement to the ETAS model that effectively handles the incompleteness in earthquake catalogues on different time scales and reduces biases in parameter estimation. This enhancement is designed to improve the model's predictive accuracy, resulting in more reliable forecasts of aftershock sequences and better-informed seismic hazard assessments. In addition to the methodological advancement, I apply the enhanced ETAS model to critically evaluate and challenge current practices within the field. By systematically analysing the results, I identify key areas where current approaches fall short and provide actionable suggestions for improvement. These contributions indicate the robustness and validity of my proposed method and also provide a comprehensive framework for refining seismicity modelling practices.

1.2.3 Specific objectives

This thesis focuses on the following objectives:

- **Addressing the short-term incompleteness in early aftershocks:** Enhancing the ETAS model to effectively handle short-term incompleteness following large earthquakes. This will involve adjusting the inversion algorithm using a censorship approach to account for the initial period of aftershock sequences, ensuring that early aftershocks are accurately addressed in the model.
- **Integrating a fast Bayesian method:** Implementing the modified ETAS model in a Bayesian framework using the Integrated Nested Laplace Approximation (INLA) method (details about this method are presented in Section 3.7). This will allow for more accurate estimation of model parameters by leveraging prior knowledge and providing posterior distributions with quantified uncertainties rather than point estimates.
- **Optimising the sampling considerations:** Investigating and optimising various sampling strategies for ETAS inversions to reduce biases in parameter estimation. This includes evaluating the sensitivity of the ETAS model to temporal binning strategies, conditioning the model on the run-in history before an earthquake se-

quence, the role of combination of different earthquake magnitudes, the trade-offs between ETAS productivity parameters, impact of the choice of incompleteness model parameters, the impact of secondary large aftershocks, and the spatial and temporal size of the modelling domain.

- **Extending the ETAS framework to incorporate long-term incompleteness:** Enhancing the ETAS model to handle long-term data gaps caused by limited network coverage and technological limitations. This improved model is intended for use when individual sequences of aftershocks are not available, making use of any accessible regional data. The updated long-term ETAS model aims to provide reliable forecasts across various time scales.
- **Validating the model's robustness through synthetic and real experiments:** Conducting an extensive testing of the modified ETAS model using synthetic datasets that simulate realistic aftershock sequences and long-term seismicity. In addition, I will apply the enhanced model to several real earthquake sequences to validate its performance and demonstrate its practical utility.

1.3 Thesis structure

This thesis is organised into seven chapters, each building on the previous to provide a comprehensive exploration of the advances made in enhancing the ETAS model.

In this regard, *Chapter 2* provides an introduction of background on seismicity modelling, earthquake forecasting, and seismic hazard analysis, with a focus on the specific challenges addressed in this thesis. It covers fundamental seismicity models, including the Gutenberg-Richter relation, declustering algorithms, the Omori law, and the ETAS model. This chapter also discusses the remaining challenges and gaps in current models, establishing the context for the research presented in this thesis.

In *Chapter 3*, the focus is on the statistical modelling techniques used in this research. It introduces Bayesian inference, particularly the INLA method and the inlabru package, which are central to the proposed methodological enhancements. This chapter explains the theoretical underpinnings of these methods.

Chapter 4 presents the first major contribution of this thesis: a methodological

enhancement to the ETAS model to account for short-term incompleteness in data following large mainshocks. It details the development process and the improved performance of the modified ETAS model through synthetic data experiments.

Chapter 5 explores optimising sampling strategies for the ETAS inversions. It examines sampling considerations, evaluates potential biases, and provides guidelines for selecting representative samples, crucial for accurate parameter estimation.

Chapter 6 extends the enhancements to address long-term incompleteness in data caused by sparse network coverage and technological limitations. The modified ETAS model is tested with both synthetic and real earthquake sequences, demonstrating its applicability and reliability over several temporal scales.

Finally, *Chapter 7* summarises the key findings and contributions of the research. It discusses the broader implications of the enhanced ETAS model for earthquake forecasting and seismic hazard assessment. In addition, this chapter outlines the remaining limitations and suggests directions for future research.

Chapter 2

Background

2.1 Introduction

There are two complementary perspectives in earthquake seismology directly linked to real-world challenges: 'earthquake forecasting' and 'earthquake hazard modelling'. The first, earthquake forecasting, falls primarily within the realm of statistical seismology, while the second, earthquake hazard modelling, is more closely related to engineering seismology, although it incorporates aspects of statistical seismology as well.

Although we cannot reliably predict earthquakes and their subsequent hazards in the deterministic sense at present, we can estimate the probability of earthquake occurrences and the expected hazards with some degree of uncertainty. This is achieved through Operational Earthquake Forecasting (OEF) (Jordan and Jones 2010; Jordan et al. 2011; Jordan et al. 2014; Field et al. 2016; Schorlemmer et al. 2018) and Probabilistic Seismic Hazard Assessment (PSHA) (Cornell 1968; McGuire and Arabasz 1990; Bommer 2002; Baker et al. 2021). In this regard, the OEF deals with the short-term forecast of earthquakes, relying on past seismicity to forecast future events. Its primary goal is to provide timely warnings and reduce the impact of imminent earthquakes. In contrast, PSHA looks at the long-term probabilities of different seismic hazards at specific locations. This is essential for designing buildings and infrastructure that can withstand potential earthquakes. By combining OEF and PSHA, we can harness the strengths of both short-term alerts and long-term planning.

At the heart of these approaches are fundamental spatio-temporal seismicity mod-

els. For example, the Gutenberg-Richter relation (Gutenberg and Richter 1944) helps us understand how often earthquakes of various magnitudes occur. De-clustering techniques are used to separate dependent earthquakes (clusters) from independent ones, while the Omori law (Omori 1895a) describes how the rate of aftershocks decreases over time. These models form the basis for the ETAS model (Ogata 1988; Ogata 1998), which is widely used in OEF and PSHA studies to model the distribution of earthquake sequences in space and time.

Despite many advances in seismicity modelling, significant challenges remain. Earthquake catalogues, which are the main input data for these models, are not fully representative of the underlying process. As a result, seismicity models may malfunction due to incomplete data and biases. This thesis aims to address these issues to reduce biases and thereby enhance the reliability of seismicity models.

This chapter sets the foundation for an exploration of earthquake forecasting and seismic hazard modelling. It highlights the need to integrate advanced methodologies to overcome existing limitations and improve the accuracy of forecasts. In the subsequent sections, I will explain hazard assessment, earthquake forecasting, and fundamental seismicity models, and then discuss the remaining challenges and the role of new methods in offering a complementary perspective.

2.2 Earthquake forecasting and seismic hazard analysis

2.2.1 Operational Earthquake Forecasting (OEF)

An earthquake forecasting study aims to determine the magnitude, location, and time of future earthquakes using past seismicity through a probabilistic scheme. In other words, a forecasting experiment tries to estimate the likelihood of future earthquakes within a specified time-space-magnitude window. The current best practice for making reliable and timely short-term forecasts involves OEF, which aims to provide a basis for decision making to reduce the seismic risk of earthquake sequences (Jordan et al. 2011; Field et al. 2017). OEF studies are commonly interested in timescales from days to years. In practice, the OEF provides actionable information to various stakeholders, including emergency response agencies, government authorities, and the general public,

about the potential for future seismic activity (Marzocchi and Zechar 2011; Jordan et al. 2014).

In line with OEF studies, and to provide a robust basis for running and validating fully prospective earthquake experiments, a global cyber infrastructure named the 'Collaboratory for the Study of Earthquake Predictability (CSEP)' was launched in 2009. The CSEP platform gathers testing centres around the world to actively collaborate in forecasting experiments (Zechar et al. 2010b; Jackson et al. 2016; Michael and Werner 2018; Schorlemmer et al. 2018). CSEP examines the accuracy of earthquake forecast models by offering robustness tests and aims to avoid the potential for retrospective bias when testing using past data.

Currently, there are several OEF studies in different geographic regions, for example, including forecasts for Italy (Marzocchi et al. 2014), Canterbury, New Zealand (Rhoades et al. 2016), Europe (Zechar et al. 2016), New Zealand (Christophersen et al. 2017), and California (Milner et al. 2020). The OEF studies and applications face ongoing challenges. For example, large earthquakes typically have long-term return periods, largely confining short-term forecasting using current seismicity-based models within a low-probability environment. This low probability can be a significant obstacle in trying to communicate risk effectively to the public and policy makers (Herovic 2016).

To improve the effectiveness of OEF, researchers suggest incorporating physical information and different types of data into forecasting models (e.g., in Mancini et al. 2019). Jordan et al. (2014) recommend integrating data such as earthquake rates, Coulomb stress changes, geodetic measurements, and electromagnetic transients to enhance the time-dependent probabilities of potentially damaging earthquakes. These multidisciplinary approaches can provide a more comprehensive understanding of the physical processes driving seismicity and improve the accuracy and reliability of forecasts (Dutta et al. 2013).

In addition, OEF models must address the inherent uncertainties in earthquake forecasting. These uncertainties stem from the complex nature of earthquake processes, the limitations of current models, and the quality and availability of seismic data. Addressing these uncertainties involves developing probabilistic models that can quantify

the range of possible outcomes and their associated probabilities. This probabilistic approach allows decision makers to understand the likelihood of various scenarios and plan accordingly (Field et al. 2017).

In practice, an operational earthquake forecast should involve continuous monitoring of seismic activity, updating probabilistic models in real-time, and disseminating forecasts through various channels. The effectiveness of OEF depends on the robustness of the models used, the precision of the data, and the efficiency of communication strategies to ensure that stakeholders receive relevant and timely information.

Moreover, any earthquake forecasting study should typically be validated and pass community-endorsed OEF tests. In this regard, CSEP conducts a variety of tests to evaluate the performance of forecasting models to ensure their reliability and accuracy in operational settings (Tsuruoka et al. 2012; Bayliss et al. 2022; Khawaja et al. 2023). Here are some key types of tests conducted by CSEP:

- **Number (N) Test**: The N-Test evaluates the accuracy of earthquake forecasts by comparing the total number of forecasted events with the number of observed events. The quantile score for the N-Test represents the probability of observing at least N_{obs} events, assuming a Poisson distribution. Zechar et al. (2010a) proposed an enhanced version that checks the probability of observing at least N_{obs} events with a score δ_1 and at most N_{obs} events with a score δ_2, providing a range for the expected number of events. It is recommended to report both scores to ensure a comprehensive evaluation of the forecast's performance.
- **Likelihood (L) Test**: This test assesses the performance of earthquake forecasts by evaluating the likelihood of observed events in individual spatial bins. Each bin's likelihood is calculated using a Poisson likelihood function, and the overall likelihood of the forecast is obtained by multiplying these individual likelihoods across all bins. The quantile score is determined by comparing the joint log-likelihoods from numerous simulations to the observed log-likelihood. This score represents the fraction of simulated joint log-likelihoods that are less than or equal to the observed log-likelihood, indicating how well the forecast matches the observed data (Harte 2015; Taroni et al. 2018).
- **Conditional Likelihood (CL) Test**: This is an adaptation of the L-Test, designed

to address the dependency of L-Test results on the total number of forecasted events. It normalises the number of events during the simulation phase to match the observed number of events, thereby reducing the impact of any significant discrepancy between the forecasted and observed event counts.

- **Magnitude (M) Test**: The M-Test evaluates the accuracy of earthquake forecasts by focusing on the observed magnitude distributions. This test isolates the magnitude component from the full likelihood by summing over the spatial bins, effectively separating the magnitude information from spatial variations. The simulation approach is used to generate expected distributions, and the final test statistic is the fraction of observed log-likelihoods that fall within the range of simulated log-likelihood values. This method ensures that the forecast's ability to predict the magnitude of earthquakes is rigorously tested.
- **Spatial (S) Test**: This test examines the accuracy of earthquake forecasts by concentrating on the observed spatial distributions. This test isolates the spatial component from the full likelihood by summing over all magnitude bins, thus focusing on the location information independently of the magnitude. The simulation approach is applied to produce expected spatial distributions, and the final test statistic is the fraction of observed log-likelihoods within the range of simulated log-likelihood values. This approach ensures a thorough evaluation of the forecast's spatial accuracy.
- **New CSEP tests:** In addition to the previous traditional CSEP tests, some new tests have also been established, including the negative-binomial N-test, the binary CL- and S-tests, and the binary T-test (Bayona et al. 2022). The Negative-binomial N-Test addresses the limitations of the Poisson distribution in capturing the variability of earthquake activity by using the negative binomial distribution, which better accounts for spatiotemporal clustering of earthquakes. The binary CL- and S-Tests use a binary likelihood function to reduce the sensitivity of CSEP tests to earthquake clustering by focusing on the probability of any activity occurring in a forecast bin rather than the likelihood of observing multiple events. The binary T-Test evaluates the likelihood of an observation being either zero or nonzero in a bin, providing a fairer measure to compare performance between models by concentrating on the presence or absence of activity rather than the exact number of

events.

By continuously validating and refining the models through these rigorous testing procedures, CSEP helps enhance the accuracy and utility of earthquake forecasts.

2.2.2 Probabilistic Seismic Hazard Assessment (PSHA)

In parallel to earthquake forecasting, there is seismic hazard assessment, which is a fundamental method in engineering seismology. The purpose of PSHA studies is not only to model the occurrence of future earthquakes, but also to assess their subsequent hazard through a probabilistic approach over longer timescales. By hazard, we mean the ground motion parameters (e.g., Peak Ground Acceleration (PGA), Spectral Acceleration (Sa), Peak Ground Velocity (PGV), etc.) produced by earthquakes. An earthquake hazard analysis translates seismic activity into understandable engineering parameters, and the results of this process are directly delivered to earthquake engineers to calculate the seismic loading on structures, or to decision makers to adopt appropriate strategies for construction developments and risk reduction plans (McGuire and Arabasz 1990; Baker et al. 2021; Papadopoulos et al. 2021; Bommer 2022; Douglas et al. 2024).

Introduced by Cornell (1968), PSHA is the most widely used method to analyse seismic hazards probabilistically. It integrates geological, seismological, and engineering data to quantify estimates of seismic hazard parameters. A traditional PSHA analysis usually consists of four main steps (see Fig. 2.1):

1. Identification of seismicity background and seismogenic sources (linear and area sources),
2. Source characterisation (geology, geometry, and seismicity) and establishment of recurrence relations and parameters (such as a-value which is the productivity parameter and it indicates the overall rate of earthquake occurrence, b-value which is a metric for the ratio between the number of large and small earthquakes, annual rates, etc).
3. Selection or development of proper empirical Ground Motion Models (GMMs), and
4. Combination of all probabilities and preparation of hazard maps and curves. The

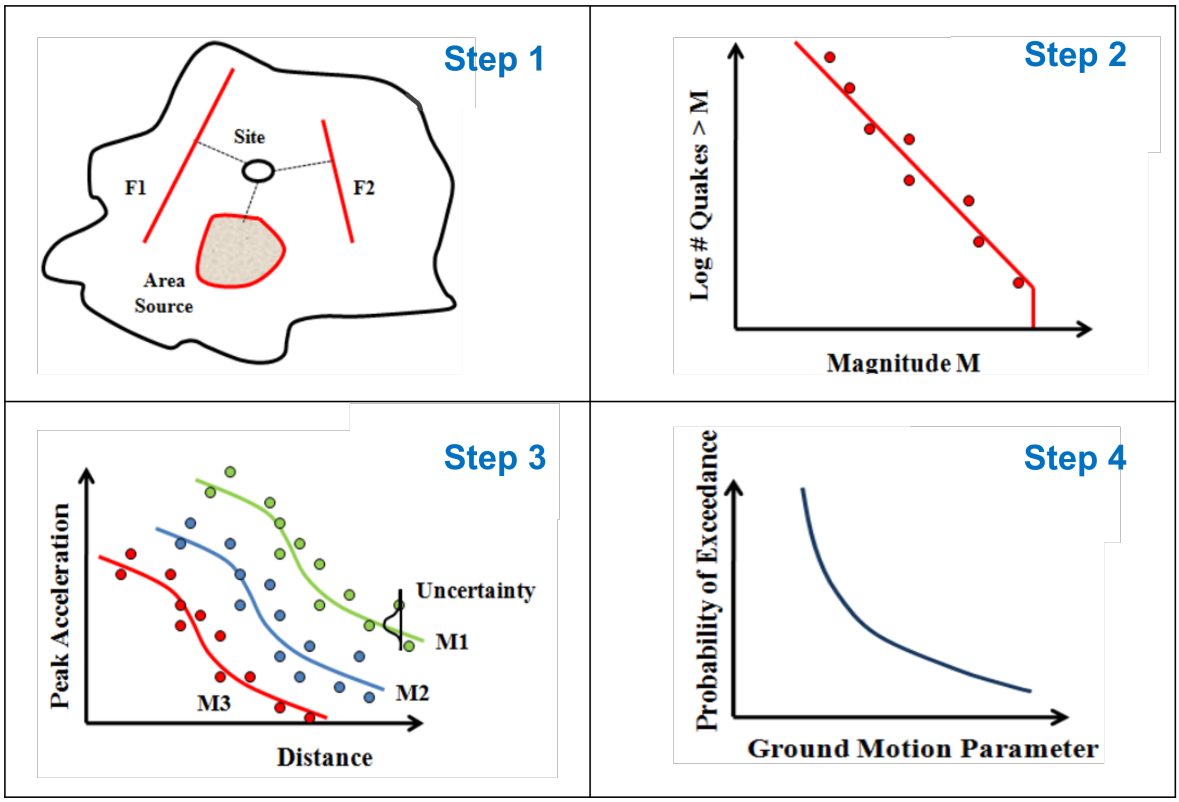

Figure 2.1: General steps of a PSHA study.

probability of exceeding a particular ground motion level at a site, considering all potential seismic sources and their associated uncertainties, is calculated using the total probability theorem:

$$P(Y > y) = \sum_{i=1}^{N} \int_{m_{min}}^{m_{max}} \int_{r_{min}}^{r_{max}} P(Y > y|m,r) f_{M_i}(m) f_{R_i}(r)\, dm\, dr \qquad (2.1)$$

where: $P(Y > y)$ is the probability of ground motion Y exceeding level y; $P(Y > y|m,r)$ is the conditional probability of exceeding y given an earthquake of magnitude m at distance r; $f_{M_i}(m)$ is the probability density function of magnitude m for source i; $f_{R_i}(r)$ is the probability density function of distance r from source i; and N is the number of seismic sources considered. This integration accounts for all possible magnitudes and distances, weighted by their probabilities, providing a comprehensive assessment of the seismic hazard at the site.

Currently, there are newer physics-based methods for hazard assessment which use numerical simulations for ground motions (e.g. Milner et al. 2021; Antonietti et al. 2021; Stupazzini et al. 2021; Field 2022). However, this procedure requires

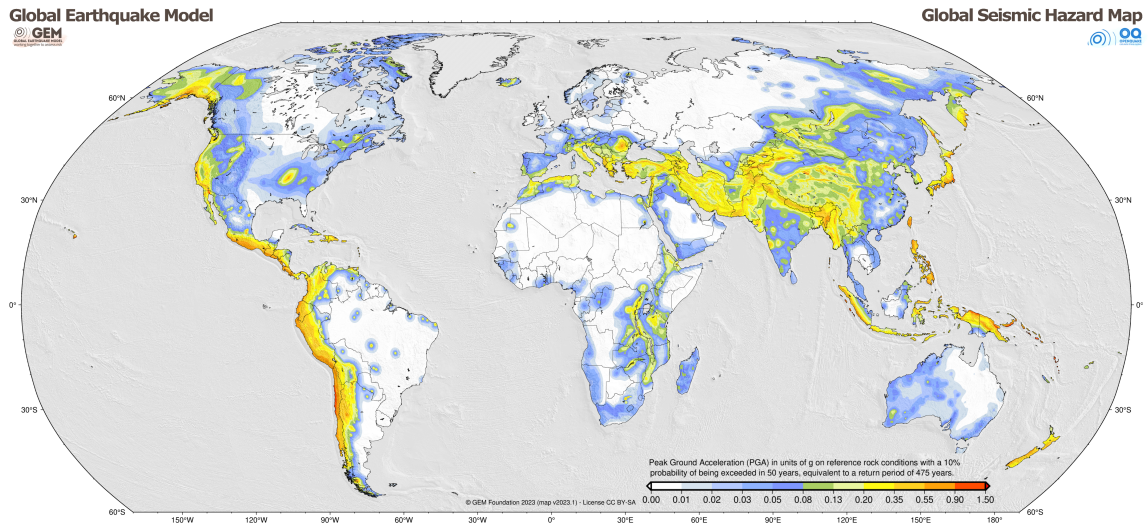

Figure 2.2: The global seismic hazard map produced by the Global Earthquake Model (GEM) Foundation (Pagani et al. 2020).

high-resolution datasets and remains a challenge for large-scale studies at present. Depending on the precision and aim of the assessment, PSHA can be implemented on various scales, including site-specific, microzonation, regional, national, and global scales (e.g. in Fig. 2.2).

Although PSHA is a common practice for hazard modelling, it remains challenging and has faced significant debates in recent years regarding the validity of such models. A notable example was raised by Stein et al. (2012), who demonstrated that earthquake hazard maps can fail considerably; for example, some recent great earthquakes, including the 2008 Wenchuan, 2010 Haiti, and 2011 Tohoku earthquakes, occurred in areas predicted by a priori earthquake hazard maps to be relatively safe. The authors suggest that these model/map failures may result from several underlying factors, such as incorrect physical models about the faulting process, characteristic earthquakes, seismic gaps, the seismic potential of sources, incorrect assumptions about fault activity rates, large uncertainties associated with return periods of large earthquakes, poorly recorded data, or perhaps only a 'black swan' event.

In a recent major review, Gerstenberger et al. (2020) explore the state-of-the-art and future challenges of PSHA on regional and national scales. They emphasise the importance of moving from traditional time-independent PSHA toward time-dependent hazard forecasting and more accurate, spatially precise hazard modelling balanced with

the quantification of uncertainty (both aleatory and epistemic uncertainties). They also mention that meeting current challenges requires the development of science-driven models using novel mathematical methods to quantify the different types of uncertainty in hazard models. In addition, recent advances in PSHA include the integration of time-dependent models, which consider temporal variability in seismicity, compared to traditional time-independent models that assume a Poisson process for earthquake occurrence. Time-dependent models provide more accurate hazard assessments, especially in regions with well-understood seismic cycles.

2.2.3 OEF and PSHA: a quick comparison

OEF and PSHA are two distinct, yet complementary approaches used in seismology to understand and mitigate earthquake risks. Although they share some common principles, they differ significantly in their objectives, methodologies, and applications. Here, I provide a comparison of these two approaches.

- **Objectives and time perspective**: The primary objective of PSHA is to estimate the likelihood of various levels of earthquake-induced ground motion at a specific site over a long period, typically spanning decades. This information is crucial to inform building codes, infrastructure design, and long-term risk mitigation strategies. The purpose of PSHA is to translate seismic hazards into actionable engineering parameters that can guide the construction and retrofitting of buildings and infrastructure to withstand potential earthquakes. In contrast, OEF aims to provide short-term forecasts of earthquake probabilities, focusing on timescales ranging from days to years. The main goal of OEF is to issue timely alerts and inform immediate risk reduction measures, such as public warnings, emergency response planning, and temporary evacuation or shutdown of critical facilities. OEF is more concerned with the imminent threat posed by earthquake sequences and aftershock activity.
- **Seismicity-based approaches**: Both PSHA and OEF rely heavily on observed seismicity. PSHA utilises long-term seismic records to identify stable, long-term patterns of earthquake occurrence, providing a probabilistic assessment of ground shaking over extended periods. This approach assumes that seismicity is relatively stationary over the long term, which is a reasonable approximation for many tectonic

settings. OEF, however, focuses on short-term variations in seismicity, particularly clustering patterns that may indicate increased likelihood of future events. This approach recognises that earthquake occurrence can be highly variable over short timescales, with periods of heightened activity following significant events.

- **Probabilistic approach and uncertainty**: Both PSHA and OEF employ probabilistic frameworks to quantify uncertainties inherent in earthquake forecasting and hazard assessment. In PSHA, uncertainties arise from the variability in earthquake occurrence, ground motion prediction, and incomplete knowledge of seismic sources. These uncertainties are addressed using statistical methods and probabilistic models that account for different scenarios and their associated probabilities. OEF also deals with significant uncertainties, particularly related to the short-term forecasting of earthquake occurrences. Probabilistic models used in OEF incorporate real-time seismic data and other geophysical information to provide estimates of earthquake probabilities, recognising the inherent variability and unpredictability of seismic activity over short timescales.
- **Physics-based approaches**: In addition to traditional probabilistic models, both PSHA and OEF increasingly incorporate physics-based approaches. These approaches utilise physical models of stress accumulation and release, fault interactions, and other geophysical processes to improve the accuracy and reliability of forecasts. For instance, OEF models may incorporate data on Coulomb stress changes, geodetic measurements, and electromagnetic transients to refine short-term forecasts. Similarly, PSHA can benefit from integrating physical constraints on fault behaviour and earthquake recurrence.
- **Applications and challenges**: PSHA is widely used in engineering seismology to inform building codes, design earthquake-resistant structures, and develop long-term risk mitigation strategies. Its long-term perspective and probabilistic nature make it a valuable tool for ensuring the safety and resilience of infrastructure. OEF, with its focus on short-term forecasts, is crucial for emergency preparedness and response. It provides timely information that can be used to implement immediate risk reduction measures, such as public warnings and evacuation plans. However, the inherent uncertainties and low probabilities associated with short-term forecasts present significant challenges in effectively communicating and utilising OEF

information.

- **Earthquake prediction vs. earthquake forecasting**: Earthquake prediction involves specifying in advance, within narrow limits and above chance, the time, magnitude, and location of a future earthquake. Such deterministic prediction is not considered a realistic scientific goal neither in PSHA nor in OEF at present (Main 1999; Jordan et al. 2014). In contrast, earthquake forecasting provides a statement of the likelihood of a population of future earthquakes within a space-time-magnitude window, expressed as a probability (Jordan et al. 2014). OEF provides forecasts of the likelihood of future earthquakes themselves, whereas PSHA propagates this likelihood into the resultant hazard of future earthquakes in terms of ground motions.

2.3 Fundamental seismicity models

Seismicity models, which elucidate how earthquakes occur and behave in space and time, are crucial to accurate earthquake forecasting and hazard assessments. In seismicity modelling, the distinction between temporal, spatial, and spatio-temporal approaches is important. Temporal models focus solely on the timing of seismic events, often neglecting where these events occur. This limitation can obscure patterns that are evident only when the spatial distribution of events is considered. Spatial models, in contrast, concentrate on the distribution of seismicity across regions but may ignore the timing of events. Spatio-temporal models aim to integrate these dimensions, providing a comprehensive view of seismicity patterns. However, these models are complex and are still under development.

In seismicity modelling, different parameters might be considered depending on the study's objectives. Common ones include the magnitude of completeness (M_c), which is the smallest earthquake magnitude reliably recorded in a region; and the $a-$ and $b-$values of the Gutenberg-Richter law. These parameters help determine annual rates and also return periods for earthquakes of different magnitudes.

Distinguishing between random seismic events and those that are causally linked, such as aftershocks or earthquake swarms, is also crucial. Background seismicity refers

to the random, unclustered earthquakes, while aftershocks and swarms are characterised by their spatio-temporal clustering. In PSHA, parameters such as maximum magnitude ($M_{\max}$) and maximum credible earthquake (MCE) are particularly important, as they provide upper limits on potential risk, especially if the earthquake distribution includes more large events than expected based on historical data (Main 1995). It should be noted that the maximum credible magnitude refers to the largest earthquake that is considered physically possible along a particular fault or in a specific seismic region, based on geological and seismological and geophysical evidence (Sokolov 2017). It represents an upper limit used for engineering and hazard assessment, taking into account fault characteristics and historical data.

Several well-established empirical statistical models are used in seismicity modelling to quantify these parameters and improve our understanding of earthquake processes. These include:

- **Gutenberg-Richter recurrence relation**: This describes the relationship between the magnitude and the total number of earthquakes over a given time period.
- **De-clustering algorithms**: These are used to separate background seismicity from foreshocks and aftershocks.
- **Omori-Utsu law**: This describes the decay rate of aftershocks following a mainshock, indicating how the frequency of aftershocks decreases over time.
- **ETAS model**: This model simulates the occurrence of aftershocks based on the concept that each earthquake can trigger additional earthquakes, creating a cascading effect.

Using these models, we can better forecast future seismic activity, assess potential hazards, and develop strategies to mitigate earthquake risks. In the following, I briefly explain some of these models in a bit more detail.

2.3.1 Gutenberg-Richter (G-R) relation

Introduced by Gutenberg and Richter (1944), the $G-R$ law is one of the most fundamental empirical models in seismicity. It describes the relationship between earthquake magnitudes and their frequency in a given region and time period. The $G-R$

recurrence relation states that earthquake magnitudes are distributed exponentially, according to the formula:

$$\log_{10} N = a - bM \tag{2.2}$$

where M is the magnitude of earthquakes, and N is the cumulative frequency of earthquakes with magnitude $\geq M$. The parameter a (known as the a-value) is the intercept of the regression model and is related to the seismotectonic features, indicating the rate of seismic events for $M \geq 0$. The parameter b (known as the b-value) is the negative slope of the model, which relates to the ratio between small and large events. The b-value typically ranges from 0.6 to 1.2 in different regions. A higher b-value (steep slope) indicates a higher number of small-to-moderate events relative to large ones, while a lower b-value (gentle slope) suggests that the energy of the region is not released through frequent small events, implying a higher likelihood of large events and thus greater hazards. The b-value is a critical parameter both in the PSHA and OEF models (Gulia and Wiemer 2019; Elst 2021).

Since the magnitude threshold for complete reporting (M_c) is an arbitrary reference point and it would rarely be equal to zero, we can reformulate the $G-R$ law to include (M_c) so that:

$$\log_{10} N = a - b(M - M_c) \tag{2.3}$$

One method for estimating M_c is the vertex of the curvature at the head of the $G-R$ plot using techniques such as Maximum Curvature (MAXC) (Wiemer and Wyss 2000). Another method of computing M_c is based on the stability of the b-value, which demonstrates more reliable estimates of M_c (Roberts et al. 2015). Fitting an appropriate model to magnitude-frequency data is challenging because the head and tail of the fitted curve are highly influenced by the deficiency of recorded data (low-magnitude events due to low-resolution coverage of seismic networks and large earthquakes due to return periods longer than the lifespan of seismic networks). According to Aki's method (Aki 1965), the b-value can be estimated by maximum likelihood estimation as:

$$\tilde{b} = \frac{\log_{10} e}{\overline{M} - (M_c - \Delta M/2)} \tag{2.4}$$

where $\tilde{b}$ is the estimate of the b-value, $\overline{M}$ is the mean magnitude, M_c is the magnitude of completeness, and ΔM is the magnitude bin size. The uncertainty of the b-value at

one standard deviation (67% confidence) can be estimated by:

$$\sigma_{\tilde{b}} = \frac{\tilde{b}}{\sqrt{N_c}} \tag{2.5}$$

where N_c is the number of events in the complete part of catalogue, or 1.96 times this value at 95% confidence (Roberts et al. 2015). For this equation to be valid, the underlying data must follow a $G-R$ distribution, which is not always clear in small-bandwidth datasets.

2.3.2 De-clustering algorithms

Seismicity is characterised by a complex spatio-temporal pattern of earthquakes (Utsu 2002). It is well understood that earthquakes do not usually occur as individual events, but tend to trigger other earthquakes, forming clusters in space and time to generate seismic sequences. Such sequences typically consist of two parts: (1) independent earthquakes or 'mainshocks' which are primarily caused by tectonic loadings or stress transients, and (2) dependent events, including 'foreshocks' and 'aftershocks', which are triggered by static or dynamic stress changes, seismically activated fluid flow, after slip, post-seismic stress relaxation, etc. Consequently, the time correlation and spatial clustering of an earthquake sequence can reveal different aspects of the dynamic behaviour of seismicity in a region (Telesca et al. 2016).

PSHA assumes a spatially inhomogeneous, temporally homogeneous Poisson process model. Therefore, it is sometimes necessary to separate the independent (mainshocks) and dependent (foreshocks and aftershocks) parts of seismic clusters in seismicity datasets. This process of event-type separation is called 'declustering'. Several traditional declustering approaches have been proposed over the years. These include simple space-time window methods (Gardner and Knopoff 1974), link-based methods (Reasenberg 1985), a game-theory-based clustering approach (Molchan and Dmitrieva 1992), stochastic modelling (Kagan and Jackson 1991; Zhuang et al. 2002; Zhuang et al. 2004; Zhuang et al. 2006), model-independent stochastic declustering (Marsan and Lengline 2008; Marsan and Lengliné 2010), single-link cluster analysis (Frohlich and Davis 1990; Davis and Frohlich 1991), model-based inter-event-time distributions (Hainzl et al. 2006), the coefficient of variation of inter-event times (Bottiglieri et al.

2009), the ratios method (Frohlich and Davis 1985), correlation metric-based methods (Baiesi and Paczuski 2004; Zaliapin et al. 2008), and the waveform similarity approach (Barani et al. 2007). More recent studies have used additional techniques such as the nearest-neighbour distance method, graph theory, and probabilistic identification of clusters (Zaliapin and Ben-Zion 2013; Bayliss et al. 2019).

Declustering is crucial in seismic hazard analysis because it aims to isolate independent seismic events, providing a clearer picture of the underlying seismicity, and improving the accuracy of hazard assessments. One of the main challenges in declustering is to effectively distinguish between mainshocks and aftershocks, especially in complex seismic sequences where the boundaries are not clear-cut. This is particularly important in regions with high seismic activity, where the interaction between earthquakes can significantly influence hazard assessments.

Different methods offer various strengths and weaknesses. For instance, space-time window methods are simple to implement but may not capture the complexity of seismic sequences accurately. Stochastic modelling approaches, such as the ETAS model, offer a more sophisticated representation by considering the probability of aftershock occurrence based on the magnitude and location of previous earthquakes. However, these models require extensive computational resources and detailed seismicity data. Finally, the choice of declustering method depends on the specific requirements of the study, including the complexity of the seismic region, the available data, and the computational resources. As seismic data and modelling techniques continue to advance, ongoing research is essential to refine these methods and enhance their applicability in various contexts of seismic hazard assessment.

2.3.3 Omori law and ETAS model

The temporal decay in the rate of aftershocks following a mainshock is a crucial aspect in seismicity modelling. This decay can be empirically described using the Omori law (Omori 1895b; Utsu 1961), a widely-accepted relationship in seismology as:

$$n(t) = \frac{K}{(c+t)^p} \tag{2.6}$$

where t denotes the time after the mainshock, and $n(t)$ represents the number of aftershocks per unit time (usually per day). This implies that the number of events decays as an inverse power law of the exponent p after a characteristic time c, with amplitude K proportional to the total number of aftershocks in the sequence. The parameter c is necessary to avoid a numerical singularity at zero time, but is constrained by the number of events that can be detected when wave trains of separate aftershocks overlap early in the sequence. The value of c is typically a fraction of an hour or a day (usually between 0.5 and 20 hours after the mainshock) and is inversely related to the aftershock threshold magnitude (M_z) rather than being an intrinsic time constant in the physical process (Utsu and Ogata 1995; Peng et al. 2007). Since small aftershocks are not recorded in the initial part of seismic sequences, the estimation of c and K might be strongly biased (Lippiello et al. 2019b). The parameter p represents the rate of aftershock decay and typically has a value between 0.7 and 1.5. A higher p-value indicates a steeper decay slope and thus a shorter time to return to the background seismicity rate. Studies indicate no systematic relation between the p-value and the energy released or the magnitude of the mainshock, but the p-value is more influenced by the tectonic environment, such as structural heterogeneity, stress and temperature (Utsu and Ogata 1995).

Although earthquakes are finite objects in space and time, earthquake catalogues list points in space (location) and time (origin time), marked by attributes such as magnitude. This allows seismicity to be modelled using spatio-temporal, marked point process models. In this context, the magnitude serves as the 'mark'. Point process models are a class of statistical models that define the behaviour of a set of randomly located points in a mathematical space ($\mathbb{R}^d$). Examples of point processes include Poisson, Cox, Hawkes (self-exciting), determinantal, and geometric processes (Schoenberg 2010). Introduced by Ogata (1988) and Ogata (1998), the Epidemic-Type Aftershock Sequence (ETAS) model is a point process belonging to the class of self-exciting Hawkes processes (Hawkes 1971) adapted for aftershock sequences. The ETAS model can be fitted to seismicity data from a catalogue using maximum likelihood estimation. It is based on the idea that each earthquake, whether spontaneous or triggered, can trigger its own aftershocks. Such point processes can be defined by a conditional intensity function, λ, which represents the infinitesimal expected occurrence rate at a specific

point in space and time and can be compared or optimised to those observed. The intensity function of the ETAS model is formulated as:

$$\lambda(s,t \mid H_t) = \mu(s) + \sum_{i:t_i<t} f(m_i) \cdot g(t-t_i) \cdot h(s-s_i, m_i) \tag{2.7}$$

where s and t denote the space and time domains, respectively, and $H_t = \{(s_i, t_i, m_i) \in X : t_i < t\}$ symbolises the history of the point process. The parameter $\mu(s)$ represents the background seismicity rate for independent events, assumed to be stationary, that is, constant and independent of time. The three functions within the summation kernel $\{f(m_i), g(t-t_i), h(s-s_i, m_i)\}$ represent the probability density functions of magnitude, time, and location for a direct aftershock triggered by the parent event i, respectively.

The ETAS model is a highly efficient method for characterising aftershock sequences in a probabilistic sense, and therefore, it serves as a powerful alternative to identifying the background rate using de-clustering algorithms. Furthermore, the ETAS model has been used not only for determining aftershock sequences, but also to other areas in a practical manner. Recent studies highlight various applications of the ETAS model in seismic hazard assessment (Reyes Canales and Baan 2020; Papadopoulos et al. 2021; Zhang et al. 2021), probabilistic forecasting of seismicity (Gordon et al. 2021), physics-based forecasting models (Mancini et al. 2020; Hardebeck 2021), and its application within a Bayesian framework using prior information (Ebrahimian and Jalayer 2017; Shcherbakov et al. 2019).

2.4 Challenges and gaps

2.4.1 Earthquake catalogues and data quality

All the seismicity models and therefore PSHA and OEF models mainly depend on earthquake catalogues that report information about occurrence time, location, magnitude, depth and other seismic parameters about earthquakes. The evolution of these catalogues reflects significant advancements in seismological methods and technologies. Before 1900, earthquake data were predominantly derived from historical accounts rather than scientific instruments, and these records often lacked the precision and completeness necessary for accurate analysis. From 1900 to 1964, the early instrumental era introduced seismographs that improved the accuracy and systematic recording

of earthquakes, although the initial sparsity of seismic networks and limited sensitivity of instruments often missed smaller earthquakes. Over time, as technology improved and networks expanded, the recording of these smaller events became more consistent. Since 1964, the modern era of seismology has seen significant improvements in seismic monitoring. The introduction of the World Wide Standardised Seismograph Network (WWSSN) and advances in digital recording have enabled real-time, global monitoring of seismic activity, characterised by detailed data that allow for comprehensive spatial and temporal analyses of seismicity.

A key concept in analysing seismicity catalogues is the 'magnitude of completeness' (M_c), which is the lowest magnitude at which all earthquakes in a region are reliably detected. The magnitude of completeness is not a static value, but varies between different regions and over time, influenced by changes in network sensitivity and the deployment of new seismic monitoring technologies (Wiemer and Wyss 2000). This variability can significantly affect seismic hazard assessments and requires regular updates of the M_c value as seismic networks evolve.

Several techniques have been developed to estimate M_c, such as the maximum curvature method, which identifies M_c at the peak of the frequency-magnitude distribution curve (Woessner and Wiemer 2005). This method is particularly useful in regions with well-established seismic networks and can provide a quick estimate of M_c for ongoing seismic monitoring. Furthermore, a multiscale approach has been introduced to map M_c across different regions, acknowledging that M_c can vary not only temporally but also spatially within a single catalogue. This approach accommodates variations in network coverage and geological characteristics that might affect earthquake detection capabilities (Mignan et al. 2011).

Regular evaluation and updating of M_c are essential to maintain the integrity of seismicity catalogues and to ensure that historical and contemporary earthquake data are accurately interpreted for seismic hazard analysis and earthquake forecasting. In addition, there are some particular patterns in the completeness of earthquake catalogues. Temporally, in the short term, especially after large earthquakes, earthquake catalogues often become less complete. This is due to the saturation of the seismograms by numerous aftershocks, making it difficult to detect and accurately locate

smaller aftershocks (this issue will be addressed in research Chapter 4). Over long periods, gradual improvements in seismic detection technology have led to step-like increases in the detection of smaller earthquakes (I will address this issue in the research Chapter 6).

The quality and completeness of the earthquake catalogues significantly influence the accuracy of the seismicity models used in PSHA and OEF. Incomplete or biased data can lead to incorrect assessments of seismic hazard and forecasts. Therefore, understanding the limitations and uncertainties inherent in these catalogues is crucial to developing robust models that provide reliable forecasts and assessments. As such, ongoing improvements in data collection and processing techniques, along with the development of sophisticated statistical methods to address data incompleteness, are vital for advancing our understanding of seismicity patterns and spatio-temporal evolution of earthquakes.

2.4.2 Remaining challenges for spatio-temporal seismicity, OEF and PSHA models

Despite significant advances, both earthquake forecasting and seismic hazard models face numerous challenges that can hinder their accuracy and reliability. Although forecasting studies have made considerable progress in the last decade, the predominant approach in earthquake hazard modelling remains the traditional PSHA method, which has seen little change for more than half a century.

Developing seismicity models that effectively combine spatial and temporal data is a critical but challenging goal in earthquake science. Ideally, these models would provide a comprehensive understanding of how earthquakes occur and interact over time and in different regions. However, these models are still a work in progress, with several significant challenges that limit their practical use in OEF and PSHA. Specifically, PSHA models often simplify to stationary spatial models that do not account for temporal variations and typically rely on incomplete data sets. This simplification can obscure the true variability and timing of seismic hazards. In contrast, OEF models mainly focus on temporal data to forecast short-term earthquake likelihood, which significantly reduces their ability to use spatial information effectively. This tempo-

ral emphasis diminishes the potential to leverage spatial patterns that could elucidate the sequences and clustering of seismic events, thus limiting the models' overall predictive power. The challenge remains to enhance these models by integrating robust spatio-temporal analytics that can dynamically reflect both the spatial distribution and temporal evolution of seismic activity for better accuracy in earthquake forecasting and hazard assessment.

Another challenge is the inherent uncertainty in earthquake forecasting. Current models often struggle to account for the complex and non-linear nature of seismic activity. Enhancing the precision of earthquake forecasts requires the integration of recent scientific advances and novel methodologies. These include the development of more sophisticated mathematical and computational models (specifically the next generation of ETAS models with a nonlinear kernel) that can better capture the dynamics of seismic processes. In addition, quantifying uncertainties in traditional PSHA models often relies on simplifications and assumptions that do not fully reflect the variability in seismic activity. Therefore, improved statistical methods and computational power now allow for more comprehensive uncertainty quantification, which is essential to produce more reliable hazard assessments.

In OEF models, we need to distinguish between triggered seismic events and background seismicity. Triggered events are typically aftershocks or earthquakes that occur as a direct result of a previous mainshock, whereas background seismicity includes earthquakes that are not directly linked to specific triggering events and may occur randomly over time. Accurately partitioning between these two types of events is essential for developing reliable seismicity models. Mistakes in this partitioning can lead to significant errors in the forecasting. For example, overestimating the influence of triggering can result in overstated earthquake probabilities, while underestimating it can lead to unpreparedness for potential aftershocks. The challenge lies in refining the forecasting models to better recognise and separate these phenomena, enhancing the accuracy of both short-term earthquake forecasts and long-term seismic hazard assessments. This requires a deep understanding of the underlying geophysical processes and continuous improvements in both data collection and analytical techniques.

Building more robust seismicity models is essential. This involves not only refining

existing models but also integrating new data sources and techniques. For example, the combination of statistical and physics-based approaches can provide a more holistic understanding of seismic hazards (Iacoletti et al. 2021; Zhu et al. 2023). Time-dependent modelling, as opposed to the traditional time-independent approach, offers a way to incorporate temporal variations in seismicity, leading to more dynamic and responsive hazard assessments.

Moreover, current practice in seismic source modelling often involves subjective decisions by experts, which can introduce biases. Moving towards automatic, objective approaches in constructing seismic source models can help eliminate these biases and improve the reproducibility and reliability of hazard assessments.

Another challenge lies in the effective communication and application of these models. Stakeholders, including policymakers, engineers, and the general public, need accessible and actionable information derived from these advanced models. Ensuring that the outputs of OEF and PSHA are effectively translated into practical measures for earthquake preparedness and risk mitigation is crucial.

In summary, while OEF and PSHA have laid the foundation for understanding and mitigating seismic hazards, addressing their limitations through advanced modelling techniques, better uncertainty quantification, integration of new approaches, and objective seismic source modelling is essential. These advances promise to improve the accuracy and reliability of earthquake hazard forecasting, ultimately contributing to improved safety and resilience against seismic events.

2.5 How can inlabru provide a complementary view?

A novel approach that can address some of the current gaps in spatio-temporal seismicity modelling is the application of the INLA method and the inlabru package. As will be explained in Chapter 3, INLA is a robust method for fitting Bayesian models, particularly complex hierarchical models, and for calculating uncertainties. Inlabru is an extension of INLA and is an open-source R package specifically designed to facilitate spatial modelling. Here, I will address how inlabru can provide a complementary perspective to spatio-temporal seismicity modelling and advance the current boundaries

in earthquake hazard assessment.

As a recently developed approach, inlabru has proven to be a valuable tool for spatial modelling of point processes, particularly in ecological datasets (Bachl et al. 2019). Borrowing from ecological applications, Bayliss et al. (2020) recently applied inlabru to seismology, using Log-Gaussian Cox Process (LGCP) models to describe the spatially varying intensity of earthquake locations in a 2D time-independent framework. This approach can be used in OEF studies to improve the spatial accuracy of earthquake forecasts.

Given that earthquake occurrences exhibit clustering behaviour in both space and time, they can be effectively modelled by the Hawkes process (Hawkes 1971), an extension of the Poisson process with a self-exciting property. By customising and extending the inlabru model, we can implement time-dependent analyses of seismicity, enabling us to model both 'background' seismicity and 'clustering' effects with their associated uncertainties (Naylor et al. 2023). This is crucial for accurate forecasting.

Current declustering algorithms often rely on assumptions about data clustering, leading to biased representations rather than models with estimated uncertainties. This issue is also present in background seismicity models, which often lack expressions of uncertainty. To address these deficits, we can develop declustering models by calculating the posterior distributions of parameters for background seismicity (e.g., abundance rate) and clustering (e.g., ETAS parameters and anisotropic spatial kernels) using INLA and inlabru. The Matérn covariance function, in particular, can enhance the accuracy of event occurrence modelling based on spatial distance.

Specific Applications of inlabru in OEF:

- *Improving Spatial Accuracy*: By utilising LGCP models within inlabru, we can achieve a more detailed and accurate spatial representation of earthquake intensity. This is particularly useful for identifying regions of higher seismic hazard and refining spatial forecasts.
- *Time-Dependent Modelling*: Extending inlabru to handle time-dependent seismicity allows for dynamic updating of forecasts as new data become available. This enhances the responsiveness of OEF models to recent seismic activity.

- *Uncertainty Quantification*: inlabru's Bayesian framework naturally incorporates uncertainty quantification. This provides probabilistic forecasts that can inform risk assessments and decision-making processes more effectively than deterministic models.
- *Integration of Spatial Covariates*: inlabru facilitates the inclusion of spatial covariates (e.g., fault lines, strain rates, soil types) into seismicity models. This integration can improve the understanding of how these factors influence seismic activity and contribute to more accurate hazard assessments.

Specific Applications of inlabru in PSHA:

- *Probabilistic Characterisation of Background Seismicity*: inlabru allows for the modelling of background seismicity rates and abundances in a probabilistic manner. This includes estimating uncertainties, which is critical for robust PSHA.
- *Modelling Critical Parameters*: Parameters such as the variability of the magnitude of completeness over time and space, and the b-value, can be incorporated into inlabru models. This provides a more nuanced understanding of seismic hazards.
- *Historical and Modern Data Integration*: inlabru can combine historical and modern seismicity data to create comprehensive models that account for long-term trends and recent activity. This results in more reliable hazard assessments.
- *Enhanced Declustering Models*: By using INLA and inlabru, we can develop declustering models that accurately represent both background seismicity and clustering effects, complete with uncertainty estimates. This reduces biases inherent in traditional declustering algorithms.
- *Spatial Pattern Analysis*: inlabru can model spatio-temporal patterns of seismicity, revealing how anisotropy orientations vary over time and space. This information is crucial for understanding the underlying processes driving seismicity and for improving hazard models.

Integration into PSHA and OEF frameworks:

By integrating inlabru capabilities with seismicity models from a time-dependent perspective, we can, in principle, enhance the PSHA and OEF frameworks. Figure 2.3 illustrates how inlabru could be integrated with OEF and PSHA studies.

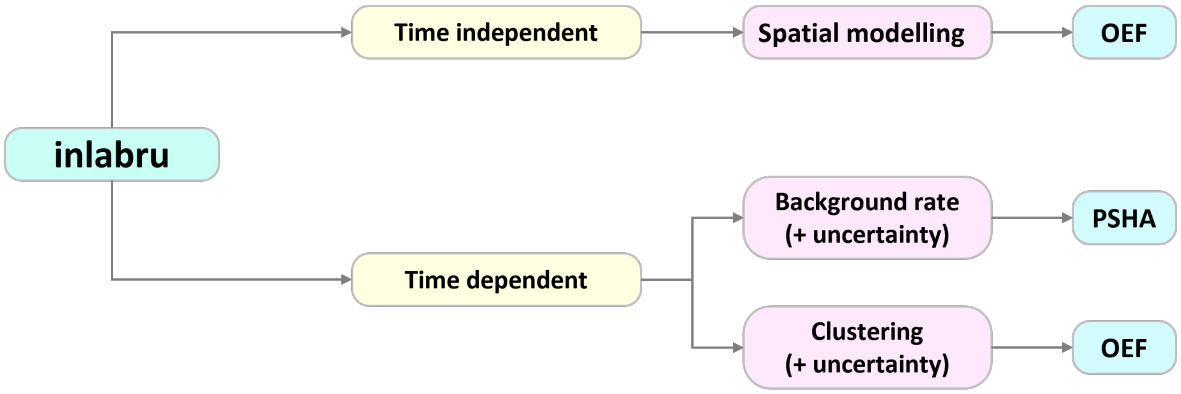

Figure 2.3: Integrating inlabru with OEF and PSHA studies.

2.6 Contribution to knowledge

This thesis makes several innovative contributions to the existing knowledge in the field of seismicity modelling:

- **Methodological enhancements:** The proposed enhancements to the ETAS model improve its theoretical foundations and practical utility, providing more accurate and reliable forecasts of aftershock sequences.
- **Bayesian inference:** The application of a Bayesian method using the inlabru package and the INLA method represents a novel approach to address data incompleteness and bias in seismicity modelling.
- **Comprehensive testing:** The validation of the modified ETAS model through extensive tests with synthetic and real data provides a new standard to evaluate seismicity models.
- **Guidelines for sampling strategies:** This research establishes guidelines for optimising sampling strategies in ETAS model inversions, contributing to more robust and unbiased parameter estimation.
- **Addressing long-term incompleteness:** The expansion of the model's applicability to long-term data incompleteness fills a critical gap in seismicity modelling, enhancing its relevance for seismic hazard assessments over extended periods.

Chapter 3

Methods

3.1 Statistical Modelling

Statistical modelling plays an important role in understanding and predicting complex phenomena in various scientific disciplines. At its core, statistical modelling involves the construction and analysis of mathematical models to represent data and underlying processes. These models are useful tools for extracting information, making predictions and informing decision makers. Statistical modelling usually involves several general steps:

1. **Observation**: Recording observations of the phenomena of interest (e.g. earthquakes) as a dataset.
2. **Dependency**: Assuming/diagnosing dependency structures between variables in the dataset.
3. **Data visualisation**: Creating data plots for visual inspection of distributions.
4. **Model fitting**: Determining an optimal statistical model, fitting this to the data, determining the model parameters, and understanding the relations between variables.
5. **Goodness-of-fit**: Assessing how well the model fits the data.
6. **Prediction**: Using the optimal model to predict outcomes outside the data set used to condition the model.
7. **Validation**: Assessing the accuracy of the predictions made by the fitted model.

8. **Verification**: Checking whether the model assumptions are still in place.
9. **Model visualisation**: Providing the final visualisations of the model (maps and diagrams).
10. **Interpretation**: Interpreting the outputs of the model and its strengths and weaknesses.
11. **Simulation**: Making simulations from a model.

Statistical modelling is crucial for several reasons. Firstly, it offers a structured way to quantify the relationships between different variables, helping researchers to test their hypotheses and understand the underlying processes better. Secondly, it enables us to predict future outcomes by analysing historical data, which is invaluable in many fields. In addition, these models help us to grasp the variability and uncertainty present in the data, which guide better decision-making in uncertain situations. Statistical models are especially important when dealing with complex and high-dimensional data, where simple analysis falls short. By applying advanced statistical techniques, we can discover patterns, spot anomalies, and identify the key factors driving the observed phenomena. This is particularly useful in areas like spatio-temporal modelling of earthquake patterns, where understanding the dynamics over space and time is essential.

Despite their benefits, statistical modelling presents several challenges. Selecting and validating the right model is crucial because an inappropriate model can lead to biased results and wrong conclusions. Overfitting is another common issue, where the model ends up capturing noise instead of the actual signal, especially with complex models. It is also important to ensure that models are interpretable and transparent, particularly in fields where decisions based on model results have significant consequences. The development of advanced computational methods and the availability of large datasets have led to more sophisticated statistical models. Techniques like Bayesian inference and Machine Learning (ML) provide powerful tools for handling complex data and making robust predictions. In a nutshell, statistical modelling is a cornerstone of modern data analysis, helping us understand and predict complex phenomena across various fields. As data becomes increasingly abundant and complex, the

need for advanced methodologies and computational tools to fully utilise their potential becomes even more critical.

3.2 Maximum Likelihood Estimation (MLE)

Maximum Likelihood Estimation (MLE) (Fisher 1922) is a commonly used method for estimating parameters in many statistical models (Dastile et al. 2020). In MLE, the goal is to find the best fit model and to determine the parameter values that make the observed data the most likely (Struben et al. 2015). An MLE analysis consists of three elements:

- **Statistical model**: A mathematical representation that describes the relationship between observed data and model parameters through a specific distribution or functional form. For example, the Omori relation (Eq. 2.6) has a power-law model.
- **Parameters**: The model parameters are the values we want to estimate. In my Omori example, these parameters are K, c, and p.
- **Likelihood**: This is a function that measures how likely it is to observe the given data under different parameter values.

To fit a model using the MLE method, we usually take the following general steps (Pan et al. 2002):

1. **Choosing a proper probability distribution**: we start by selecting a probability distribution that we believe best describes the data. For example, a Hawkes process model for aftershock sequences.
2. **Writing the likelihood function**: This function is based on the chosen distribution and the data. For a set of observations $X = \{x_1, x_2, \ldots, x_n\}$, the likelihood function $L(\theta)$ is the product of the probabilities of each data point in the model. If the data points are independent, it is given by:

$$L(\theta) = P(X|\theta) = \prod_{i=1}^{n} P(x_i|\theta) \tag{3.1}$$

   Here, θ represents the parameters of the distribution.

3. **Log-likelihood**: It is often easier to work with the natural logarithm of the likelihood function, called the log-likelihood, because it converts the product into a sum:

$$\log L(\theta) = \sum_{i=1}^{n} \log P(x_i|\theta) \tag{3.2}$$

4. **Maximising the log-likelihood**: Finding the parameter values that maximise the log-likelihood function. This involves taking the derivative of the log-likelihood with respect to the parameters and setting it to zero:

$$\frac{d}{d\theta} \log L(\theta) = 0 \tag{3.3}$$

By solving this set of equations, we can find the maximum-likelihood estimation of the model parameters. Consequently, the MLE method treats parameter estimation as an optimisation problem, seeking the set of parameters that best fits the joint probability of the data sample.

This is a powerful and popular method because of its flexibility, consistency, and efficiency. It is flexible because it can be used for a wide range of models. In terms of consistency, as the sample size increases, the MLE estimates converge to the true parameter values. Regarding efficiency, MLE makes full use of the data and often exhibits desirable statistical properties, such as asymptotic unbiasedness and minimal variance among unbiased estimators.

Although MLE is a general and effective approach that underlies many statistical algorithms, it has some limitations. One notable limitation is that it assumes that the data set is complete, which means that all relevant variables are present. This does not mean that the model has all the data, but that all important variables for the problem are included. However, this is not always true. Sometimes, some relevant variables are hidden and cannot be observed in the dataset, though they still affect other variables. These hidden variables are known as 'latent variables'. Many real-world problems have latent variables that are not observable in the available data. Typically, latent variables are indirectly inferred through their relationships with observable variables. In statistical models, this identification process often involves defining the structure of dependencies between the observed data and the latent variables, based on prior knowledge or assumptions about the system being modelled. For example, in mod-

els such as factor analysis or hidden Markov models, latent variables are identified by their influence on the variance of observed variables or the sequential structure (Bartholomew et al. 2011). When we have missing data or latent variables, computing the MLE becomes difficult. In such cases, an alternative method to MLE is needed. The Expectation-Maximisation algorithm is one such approach, which I will explain in the next section.

3.3 Expectation-Maximisation (EM) algorithm

As mentioned above, the MLE approach requires the selected dataset to be complete, which means that all relevant variables must be included. This becomes intractable if there are variables that interact with those in the dataset but are hidden or not observed, the so-called latent variables. To deal with this, the Expectation Maximisation (EM) technique (Dempster et al. 1977) is used to perform MLE when there are missing or incomplete data with latent variables, which are variables that are not directly observed but are 'inferred' from the observed data.

The EM algorithm is an iterative process that operates in two main steps: the Expectation (E) step and the Maximisation (M) step (Moon 1996; Gupta and Chen 2011). These steps are applied iteratively until the algorithm converges to a set of parameter values (Ng et al. 2012; Daskalakis et al. 2017). So the process includes:

1. **Initialisation:** Starting with initial guesses for the model parameters.

2. **E-Step (Expectation):** Estimating the expected value of the latent variables with respect to the observed data and the current estimate of the parameters. In this step, we, in fact, compute the expectation of the log-likelihood function of the complete data, given the observed data and the current parameter estimates. This step effectively fills in missing data or latent variables with their expected values. Mathematically, if $\theta^{(t)}$ are the current parameter estimates at iteration t, the E-step computes the expected value:

$$Q(\theta|\theta^{(t)}) = \mathbb{E}_{Z|X,\theta^{(t)}}[\log L(\theta; X, Z)] \tag{3.4}$$

where X are the observed data; Z represents the latent variables; θ are the pa-

rameters to be estimated; and $L(\theta; X, Z)$ is the likelihood function of the complete data.

3. **M-Step (Maximisation):** Optimising the model's parameters to best fit the available data and maximise the expected log-likelihood found in the E-Step. In this step, we find the parameter values θ that maximise the expected log-likelihood computed in the E-Step.

$$\theta^{(t+1)} = \arg\max_{\theta} Q(\theta|\theta^{(t)}) \tag{3.5}$$

4. **Convergence:** Repeating the E-step and M-step until the parameters converge, meaning that the changes in the parameter estimates are smaller than a predefined threshold.

A common application of the EM algorithm is in Gaussian Mixture Models (GMMs) (Bilmes et al. 1998; Sridharan 2014; Santos et al. 2017), which assume that the data are generated from a combination of Gaussian (Normal) probability distributions with unknown parameters. Each Gaussian represents a cluster in the data. For a single Gaussian distribution, the probability density function (PDF) is given by:

$$f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3.6}$$

where x is the data point; μ is the mean of the Gaussian, and σ is the standard deviation of the Gaussian. Using the EM algorithm, we follow the estimation steps as:

- **Initialisation**: We start by guessing the number of Gaussian distributions (clusters) K and their initial parameters: means μ_k, covariances Σ_k, and mixing coefficients π_k.

- **E-Step**: For each data point x_i, we calculate the probability that it belongs to each Gaussian component k using Bayes' theorem:

$$\gamma_{ik} = \frac{\pi_k \, \mathcal{N}(x_i|\mu_k, \Sigma_k)}{\sum_{j=1}^{K} \pi_j \, \mathcal{N}(x_i|\mu_j, \Sigma_j)} \tag{3.7}$$

where π_k is the mixing coefficient for cluster k, and $\mathcal{N}(x_i|\mu_k, \Sigma_k)$ is the probability density function of Gaussian k.

- **M-Step**: Updating the parameters of each Gaussian to maximise the expected log-likelihood with respect to the current estimates from the E-step:

$$\begin{aligned}
\pi_k^{(t+1)} &= \frac{1}{N}\sum_{i=1}^{N}\gamma_{ik} \\
\mu_k^{(t+1)} &= \frac{\sum_{i=1}^{N}\gamma_{ik}x_i}{\sum_{i=1}^{N}\gamma_{ik}} \\
\Sigma_k^{(t+1)} &= \frac{\sum_{i=1}^{N}\gamma_{ik}(x_i-\mu_k^{(t+1)})(x_i-\mu_k^{(t+1)})^T}{\sum_{i=1}^{N}\gamma_{ik}}
\end{aligned} \tag{3.7}$$

- **Convergence**: Repeating the E-step and M-step until the parameters stabilise.

The EM algorithm provides several key advantages, making it a popular tool in statistical modelling. Firstly, it is highly effective in handling datasets with missing or incomplete data, providing solutions even when significant portions of the data are unavailable. Secondly, it is very efficient in modelling scenarios involving latent variables. This capability simplifies complex problems by estimating these latent variables, thereby enhancing the overall accuracy of the model. Furthermore, the EM algorithm is versatile and can be applied to a wide range of statistical models beyond GMMs. Its iterative approach, characterised by alternating steps of expectation and maximisation until convergence, ensures robust parameter estimation, even in the presence of incomplete data or latent variables (O'Hagan et al. 2012). However, the algorithm's performance can depend on the initial parameter values, as it may converge to local rather than global optima. To mitigate this, the EM algorithm is often run with multiple initialisations to explore different solutions and ensure robustness. Initial guesses for parameter values can be generated randomly or based on domain knowledge. For example, in GMMs, K-means clustering results are often used for initialisation. Running the algorithm multiple times reduces the risk of convergence to suboptimal solutions (McLachlan and Krishnan 2007). In summary, while initial parameter choices can influence the outcome, the EM algorithm remains a potent and adaptable tool for maximum likelihood estimation, capable of addressing complex modelling challenges involving missing data or latent variables.

3.4 Fitting ETAS with EM: approach and limitations

The ETAS models are commonly fitted using the EM method. In the context of ETAS, the EM method addresses the incomplete data problem by treating the branching structure of aftershocks (i.e., which event triggers which) as latent variables. This approach was notably advanced by Veen and Schoenberg (2008), who demonstrated its robustness and accuracy for seismic data analysis. Consequently, the EM algorithm is well suited for ETAS models due to the latent nature of aftershock-triggering relationships. Several studies have used the EM method as a reliable point estimation technique for ETAS modelling using different formulations (e.g., in Wang et al. 2010; Chu et al. 2011; Lombardi 2015; Nandan et al. 2017; Stindl and Chen 2023; Mizrahi et al. 2023; Li et al. 2024). The following steps are usually undertaken for the EM algorithm in ETAS modelling:

1. **Initialisation:** Starting with initial guesses for the parameters (μ, K, α, c, p).

2. **E-Step:** For each earthquake, computing the probability that it is a background event or an aftershock. This is done using the current parameter estimates. For each earthquake, if it is considered an aftershock, estimate how much it contributes to the triggering of subsequent aftershocks.

3. **M-Step:** Using the probabilities calculated in the E-Step, updating the parameters to maximise the likelihood of observing the given earthquake sequence. Specifically, updating μ based on the expected number of background events and updating K, α, c and p based on the expected number and distribution of aftershocks.

4. **Iteration:** Repeating the E-Step and M-Step until the parameters converge.

More technical details on ETAS formulation and the use of the EM algorithm can be found in Veen and Schoenberg (2008).

The EM algorithm offers several advantages for ETAS modelling. It is robust, less sensitive to initial parameter estimates compared to direct maximisation of the incomplete data likelihood, and provides stable estimates even when the likelihood function is flat or multimodal. Flat or multimodal likelihood functions can occur in complex

models like ETAS (Kolev and Ross 2019), particularly when dealing with incomplete or noisy data, sparse datasets, or when the parameters are poorly constrained. Multimodality can also arise due to the existence of multiple plausible parameter sets that fit the data equally well. These challenges are more common when modelling seismicity in regions with low event counts or when aftershock sequences are difficult to distinguish. In addition, the algorithm converges relatively quickly, making it efficient for complex models such as the ETAS model.

Despite its advantages, the EM algorithm also has limitations. One significant drawback is its reliance on point estimates, which do not provide a full probabilistic description of uncertainty. Although the EM algorithm handles uncertainty through probabilistic expectations and iterative refinement during the estimation process, it ultimately provides single-point estimates for the parameters. The algorithm considers uncertainty by summarising the variability of the estimated parameters with summary statistics and measures such as the mean, standard deviation, quantiles, interquartile range, etc. These can be illustrated using common tools like violin plots or box plots, which show the distribution and spread of the parameter estimates, indicating their variability. However, these plots do not provide the full posterior distribution, which gives a complete probabilistic description of the parameter uncertainty. This limitation is critical, especially in ETAS modelling, where the ability to provide detailed uncertainty quantification of parameter estimates can significantly impact the reliability and robustness of seismic hazard assessments and forecasts.

Furthermore, the EM algorithm does not naturally incorporate prior information, which can be beneficial in modelling. The available information and expert knowledge can inform the model parameters, leading to more accurate and reliable estimates. The EM framework makes it challenging to integrate such prior information effectively.

Another limitation is the flexibility of the model specification. The EM algorithm can struggle with complex dependencies and hierarchical structures that are often present in seismic data. These complexities might be cumbersome or infeasible to model accurately using the EM approach.

Given these limitations, a Bayesian approach is preferable for ETAS modelling. Bayesian methods provide a robust framework for uncertainty quantification, offering

posterior distributions for all parameters. This probabilistic description is essential for a complete analysis. Additionally, Bayesian methods allow for the incorporation of prior information, enhancing the accuracy and reliability of the model. Bayesian approaches also offer greater flexibility in model specification, accommodating complex dependencies and hierarchical structures with ease. Spatial and temporal smoothing, which is crucial for accurately modelling the space-time dynamics of aftershocks, is more naturally handled within a Bayesian framework. In the following sections, I will introduce Bayesian inference in more detail.

3.5 Bayesian thinking and uncertainty

'Bayesian' analysis is a statistical modelling technique, named after the statistician Thomas Bayes (1702–1761), who was the first to use probability in statistics. The term 'Bayesian inference' does not clearly indicate what it involves, so a better term might be 'probabilistic inference' (Robert et al. 2009). In general, there are some differences between the classic 'frequentist' statistics and the 'Bayesian' approach (Table 3.1). Two main differences are:

1. In the traditional 'frequentist' approach, models are built using only data from the current experiments without any prior judgment. In contrast, the 'Bayesian' approach incorporates relevant prior knowledge (expressed as an inferred probability distribution). We then analyse the data, learn from them, update our knowledge, and adjust our model accordingly.

2. In the 'frequentist' analysis, model parameters are calculated as fixed values, while in the 'Bayesian' method, these parameters are expressed as probability distributions. This means that we view each parameter as a random variable and obtain a posterior distribution for it, from which we can provide summary statistics (mean, median, mode) and quantiles to directly obtain credible intervals and uncertainties. Therefore, in a Bayesian approach, we generally aim to obtain posterior distributions for our models, representing the distribution of the parameters given the data.

To better understand how the 'Bayesian' approach works, we first need to look

Table 3.1: Comparison between the "Frequentist" and the "Bayesian" inferences.

	Frequentist analysis	**Bayesian analysis**
Inputs	Data	Data + prior knowledge/belief
Outputs (model parameters)	Fixed values	Posterior probability distribution
Essential characteristic	Data driven	Explicit use of probability
Advantages	Being objective	Full uncertainty quantification (aleatory and epistemic)
Disadvantages	Partial uncertainty quantification (aleatory only)	Being semi-subjective
Major challenge	Handling the effect of sampling	Specifying appropriate priors, and computationally more demanding

at Bayes' theorem. This theorem is derived from the fundamental laws of conditional probabilities in probability theory, particularly the sum rule and the multiplication rule (Von Toussaint 2011). In its final formulation:

$$P(\theta \mid \text{data}) = \frac{P(\text{data} \mid \theta)P(\theta)}{P(\text{data})} \propto P(\text{data} \mid \theta)P(\theta) \tag{3.8}$$

in which θ, P, and L symbolise the model parameters, probability, and likelihood, respectively. So:

- $P(\theta)$ = prior = probability of different parameter values before seeing the data.
- $P(\theta \mid \text{data})$ = posterior = probability of different parameter values given the data.
- $P(\text{data} \mid \theta)$ = likelihood of the data given different parameter values.
- $P(\text{data})$ = This is only a scaling factor (normalising constant) and is often computationally prohibitive to calculate. It is sometimes called the 'evidence' or 'marginal', and indicates how probable the new evidence is under all possible hypotheses. It can be used in model selection, notably to define the Bayes factor between competing models (Jeffreys 1998). As in many contemporary applica-

tions, we simplify the calculation by treating it as an undetermined constant and using the proportionality sign $\propto$ instead.

The steps involved in Bayesian modelling are illustrated in Fig. 3.1. It is worth mentioning that the likelihood function of the data and the probability distributions of priors are usually selected from common statistical distributions (e.g., Gaussian, Poisson, gamma, Bernoulli, etc.). Ultimately, we sample from the posterior distribution and calculate its mean value and summary statistics. Currently, there are several methods for fitting Bayesian models (i.e., calculating posterior probability distributions), including the analytical method (conjugate priors) for the simplest models, machine learning approaches (variational Bayes), numerical simulation Markov chain Monte Carlo or MCMC) and an approximation solution (INLA). Among these, MCMC has been the most widely used approach, but it has significant limitations in terms of computational costs, as well as model convergence and performance for some classes of models. In the next sections (3.6 and 3.7), I will explain the logic behind the two MCMC and INLA methods and compare them in section 3.8.

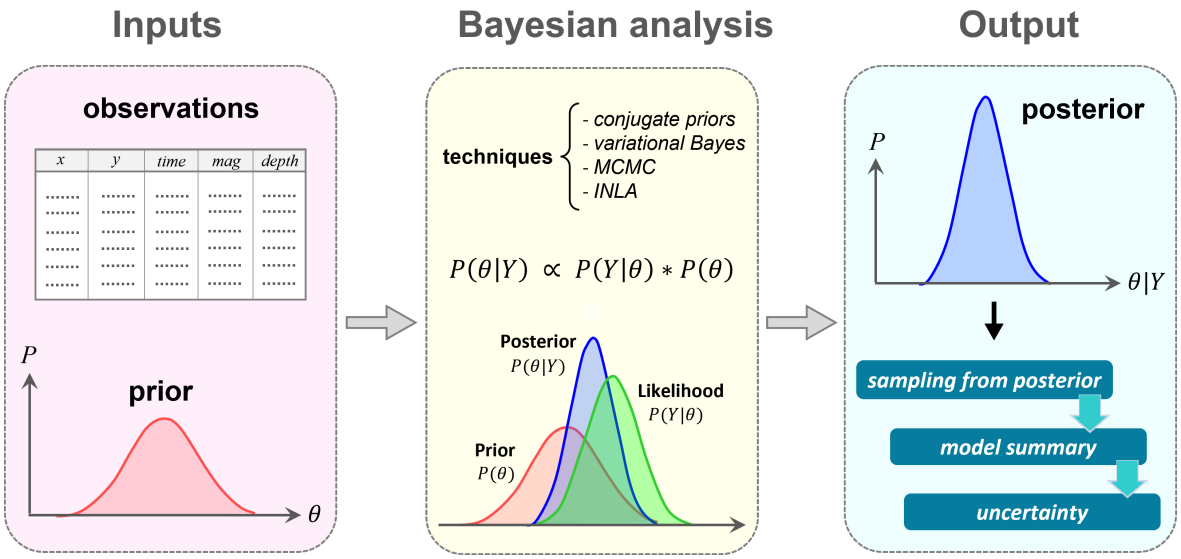

Figure 3.1: The overall framework of Bayesian approach in statistical modelling.

3.6 Markov Chain Monte Carlo (MCMC)

Markov Chain Monte Carlo (MCMC) is a powerful computational technique used in Bayesian statistics to estimate the posterior distribution of a model's parameters.

Bayesian methods involve updating the probability of a hypothesis as more evidence or information becomes available. The goal is to compute the posterior distribution, which combines the prior distribution (initial beliefs about the parameters) and the likelihood (the probability of the observed data given the parameters).

MCMC methods generate samples from the posterior distribution without requiring the distribution to be analytically tractable. The core idea is to create a Markov chain that has the desired posterior distribution as its equilibrium distribution. A Markov chain is a stochastic process in which the next state depends only on the current state, not on the sequence of events that preceded it.

Two widely used MCMC algorithms are the Metropolis-Hastings algorithm and the Gibbs sampler.

- **Metropolis-Hastings algorithm**: This algorithm generates a sequence of sample values from the posterior distribution. It works by proposing a new state and then deciding whether to accept or reject it on the basis of a probability ratio. The accepted values form a Markov chain that converges to the posteriors.
- **Gibbs sampling**: This is a special case of the Metropolis-Hastings algorithm, particularly useful when the posterior distribution is high-dimensional. In Gibbs sampling, each parameter is updated in turn by sampling from its conditional distribution given the current values of the other parameters.

Here is a simplified outline of the steps involved in using MCMC to estimate the posterior distribution of model parameters:

1. **Define the model**: Specify the prior distribution and the likelihood function. The prior reflects any existing beliefs about the parameters before observing the data, and the likelihood represents the probability of the observed data given the parameters.
2. **Initialise the chain**: Start with an initial set of parameter values. These can be chosen randomly or based on some heuristic.
3. **Propose a new state**: For the Metropolis-Hastings algorithm, generate a new set of parameters (the 'proposal') from a proposal distribution, which is often a normal distribution centred around the current state.

4. **Calculate the acceptance probability**: Compute the acceptance probability, which is the ratio of the posterior probabilities of the proposed and current states. This step ensures that the chain converges to the target posterior distribution.
5. **Accept or reject the proposal**: The decision to accept the new state is based on the Metropolis-Hastings acceptance criterion. Specifically, the new state is accepted if the computed probability ratio (the ratio of the posterior probability of the new state to the current state) is greater than or equal to 1. If the ratio is less than 1, the new state is accepted with a probability equal to this ratio. The Metropolis-Hastings criterion ensures that states with higher probabilities are accepted more frequently, but even lower-probability states can be accepted to explore the parameter space and avoid getting stuck in local maxima. If rejected, the current state is retained. This ensures that the chain spends more time in regions of higher posterior probability, thereby approximating the posterior distribution more accurately.
6. **Iterate**: Repeat the proposal and acceptance steps for a large number of iterations. The initial samples (burn-in period) are usually discarded to allow the chain to reach equilibrium.
7. **Collect samples**: The remaining samples after the burn-in period are used to approximate the posteriors. These samples can be used to estimate summary statistics, such as the mean and variance, or to make probabilistic predictions.

Fig. 3.2 shows a schematic illustration of the MCMC method to estimate posteriors of the model parameters. MCMC methods, through algorithms like Metropolis-Hastings and Gibbs sampling, provide a practical means to approximate posterior distributions in Bayesian inference. There are recent studies that use MCMC to provide Bayesian estimation of ETAS parameters (e.g. Ross 2021; Ross and Kolev 2022). MCMC methods often assume a latent structure for parent events, which helps to understand parent-daughter relationships in the ETAS model and also allows for uncertainty in this assignment. Although MCMC techniques can be accurate, they are often computationally expensive and less robust for some complex hierarchical models. To address these limitations, INLA was introduced as an alternative method, which I will explain in the next section.

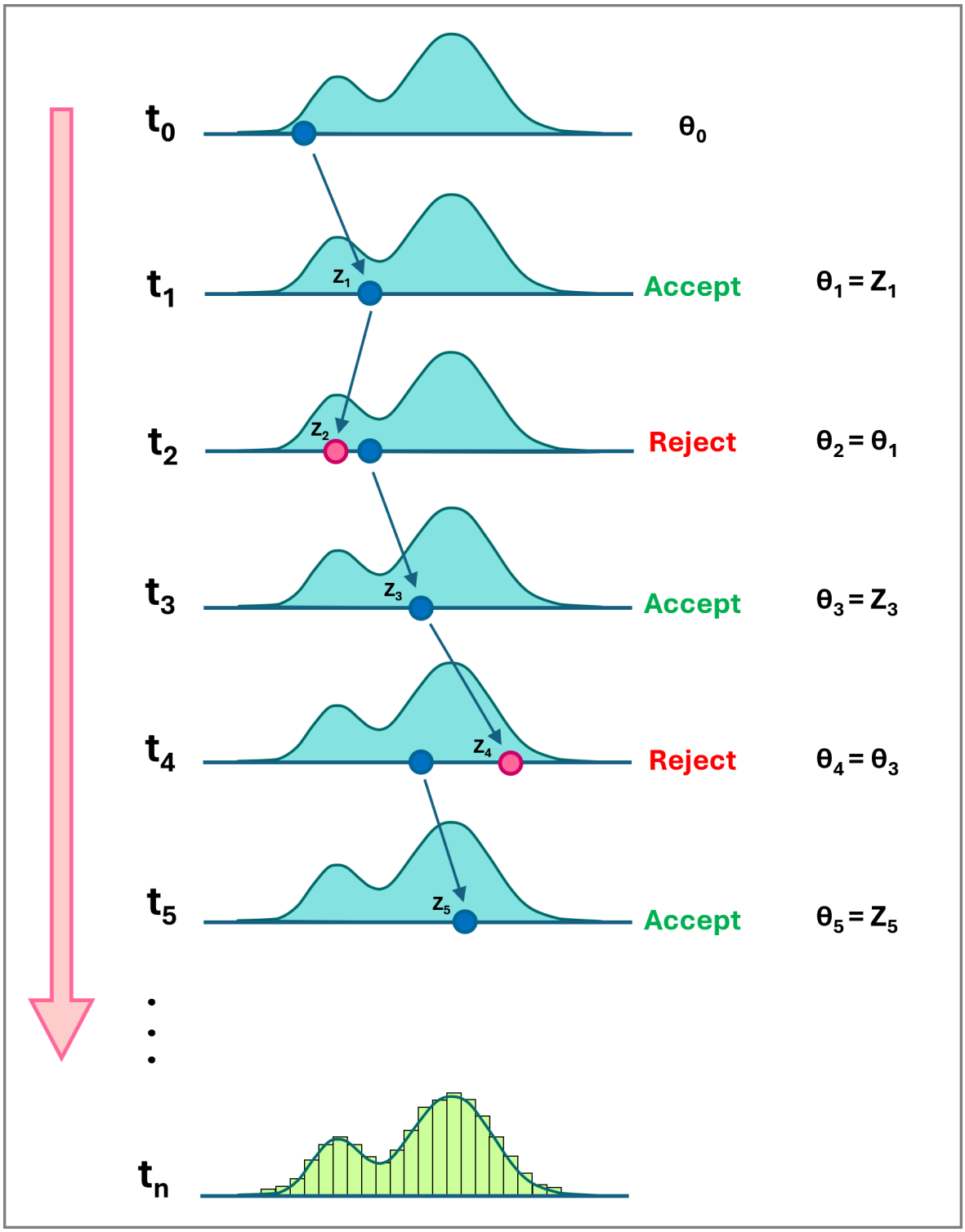

Figure 3.2: The MCMC approach for estimation of model posteriors.

3.7 Integrated Nested Laplace Approximation (INLA)

In the last decade, a novel approach called 'Integrated Nested Laplace Approximation' or INLA has been developed and proposed as a substitute for MCMC (Rue et al. 2009). INLA is a capable and fast alternative to MCMC for Bayesian analysis, especially for fitting latent Gaussian models (LGM). Many familiar models can be recast to resemble LGMs, such as generalised linear models (GLM), generalised linear mixed-effect models (GLMM), generalised additive models (GAM), generalised additive mixed-effect models (GAMM), spatial models, time series, and zero-inflated models (Morrison 2017).

INLA incorporates several mathematical concepts. In the following sections, I will introduce these concepts which may initially appear unrelated but are, in fact, complementary pieces of a larger puzzle. To avoid delving into the intricate details and complexities of INLA, which are beyond the scope of this thesis, I will keep the explanations concise and clear. More details can be found in the original paper by Rue et al. (2009) and other references, such as Blangiardo and Cameletti (2015), Krainski et al. (2018), and Zuur et al. (2017).

3.7.1 Latent Gaussian Models (LGM) and Gaussian Markov Random Fields (GMRF)

As mentioned earlier, INLA is particularly effective for the class of latent Gaussian models (LGMs), which include a wide range of statistical models. LGMs are especially useful when there are complex dependency structures between observations and covariates, as well as among the observations themselves. An LGM consists of three hierarchical components:

- **Likelihood:**

$$y \mid x, \theta_1 \sim \prod_i p(y_i \mid x_j, \theta_1) \tag{3.9}$$

- **Latent Field:**

$$x \mid \theta_2 \sim N(\mu(\theta_2), \Sigma(\theta_2)) = p(x \mid \theta_2) \tag{3.10}$$

- **Hyperpriors:**

$$\theta = [\theta_1, \theta_2]^T \sim p(\theta) \tag{3.11}$$

Hyperpriors are priors placed on the parameters θ_1 and θ_2, which control the uncertainty in both likelihood (through θ_1) and latent field (through θ_2). These hyperparameters reflect higher-level assumptions about the model and are crucial in Bayesian analysis because they allow the model to account for uncertainty in the underlying parameters.

In these equations, y represents the observed data described by likelihood, x denotes the latent field, which is the joint distribution of all model parameters in the linear predictor, and θ are the non-Gaussian hyperparameters (variance parameters) of the latent field. The distinction between θ_1 and θ_2 is that θ_1 represents non-latent parameters directly influencing the likelihood, while θ_2 governs the latent field, which indirectly influences the likelihood through its relationship with the unobserved latent variables. Since we are performing a Bayesian analysis, our goal is to find the posterior distribution of the model, which is the product of the prior for hyperparameters, the prior for model parameters given the hyperparameters, and the likelihood of observations given all parameters and hyperparameters.

Therefore, the joint posterior distribution is:

$$p(x,\theta \mid y) \propto p(\theta)\ p(x \mid \theta_2)\ \prod_i p(y_i \mid x_j,\theta_1) \tag{3.12}$$

It is important to note that we are interested in obtaining the marginal posterior distribution of each model parameter, namely $p(x_j \mid y)$ and $p(\theta_k \mid y)$, rather than the joint posterior distribution of all parameters together, $p(x,\theta \mid y)$. Later, I will explain how we use the numerical Laplace method to derive marginal posteriors from the joint posterior distribution.

In an LGM, observations (the vector y) are assumed to be independent of each other (e.g., y_i does not directly affect y_j), but are correlated through an unobserved latent field (the vector x), which has dependencies among its elements (x_i and x_j). The latent field x includes the joint distribution of all model parameters in the linear predictor (η_i). This predictor can be expressed in the general form of a generalised additive mixed model as follows:

$$\eta_i = \alpha + \sum_{j \in J} \beta_j z_{ij} + \sum_{k \in K} f_{k,j_k(i)} + \epsilon_i \tag{3.13}$$

In this equation, α is the model intercept, z is the vector of covariates with coefficients β in a linear regression, and f represents a set of functions on some covariates that may be non-linear. Therefore, the latent field x will be a set of the model parameters as:

$$x = (\eta, \alpha, \beta, f_1, f_2, \ldots) \tag{3.14}$$

The latent field x is also assumed to be a Gaussian Markov Random Field (GMRF). Thus, it takes the form:

$$x \mid \theta_2 \sim p(x \mid \theta_2) = N(\mu(\theta_2), Q^{-1}(\theta_2)) \tag{3.15}$$

Here, θ_1 is a vector of hyperparameters of the model. Σ and Q are the covariance and precision matrices, respectively, which are the inverses of each other ($\Sigma = Q^{-1}, Q = \Sigma^{-1}$) and describe the dependency structures of the process and, possibly, the observations. When two elements i and j are independent, $Q_{ij} = 0$. Later, I will discuss how the elements of a covariance matrix are specified by a Matérn covariance function for spatial models.

3.7.2 Laplace approximations

Introduced by Pierre-Simon Laplace in 1774, Laplace's method is an established numerical technique to approximate integrals of exponential functions. The method applies to integrals of the form:

$$I_n = \int e^{nf(x)}\, dx \tag{3.16}$$

where $f(x)$ is a twice-differentiable function, and n is a large number ($n \to \infty$). If we assume that x_0 is the point at which $f(x)$ attains its maximum, we can use Taylor's series expansion to approximate $f(x)$ by its first three terms. Since $f(x)$ has a global maximum at x_0 (a stationary point, not an endpoint), the first derivative $f'(x_0)$ in the second term of Taylor's expansion vanishes at x_0. Therefore, the function $f(x)$ can be simplified and approximated in quadratic order:

$$I_n \approx \int \exp\left(n\left(f(x_0) + \frac{(x - x_0)^2}{2} f''(x_0)\right)\right) dx \tag{3.17}$$

The integral above is in the form of a Gaussian integral and can be evaluated as:

$$I_n \approx e^{nf(x_0)} \sqrt{\frac{2\pi}{-nf''(x_0)}} = \tilde{I}_n \tag{3.18}$$

3.7.3 Gluing all together: INLA approach

As mentioned earlier, in Bayesian analysis, we are interested in computing the marginal posterior distribution of each model parameter rather than the joint posterior distribution of all parameters together. The Laplace approximation is particularly useful for this purpose, as it allows us to extract the marginals from the joint posteriors. Assume that $p(x_j \mid y)$ and $p(\theta_k \mid y)$ are the marginal posterior distributions of the latent field and the hyperparameters, respectively ($x_j \in x$ and $\theta_k \in \theta$). For examples of model parameters, see the vector x in Eq. 3.14. The vector θ is low-dimensional, typically consisting of 2 to 5 hyperparameters for LGM models, and not exceeding 20. The marginal posterior distributions are given by:

$$p(x_j \mid y) = \int p(x_j, \theta \mid y)\, d\theta = \int p(x_j \mid \theta, y)\ p(\theta \mid y)\, d\theta \tag{3.19}$$

$$p(\theta_k \mid y) = \int p(\theta \mid y)\, d\theta_{-k} \tag{3.20}$$

According to the Laplace approximation for integrals ($I_n \approx \tilde{I}_n$ as in Eqs. 3.17 and 3.18), the final form of the posteriors can be rewritten as:

$$p(x_j \mid y) = \int p(x_j \mid \theta, y)\ p(\theta \mid y)\, d\theta \approx \tilde{p}(x_j \mid y) = \sum_k \tilde{p}(x_j \mid \theta_k, y)\ \tilde{p}(\theta_k \mid y) \Delta_k \tag{3.21}$$

$$p(\theta_k \mid y) \approx \tilde{p}(\theta_k \mid y) = \int p(\theta \mid y)\, d\theta_{-k} \tag{3.22}$$

$$p(\theta \mid y) \approx \tilde{p}(\theta \mid y) = \left. \frac{p(\theta)\ p(x \mid \theta)\ p(y \mid x, \theta)}{\tilde{p}(x \mid \theta, y)} \right|_{x = x_0(\theta)} \tag{3.23}$$

$$p(x_j \mid \theta, y) \propto \frac{p(\theta)\ p(x \mid \theta)\ p(y \mid x)}{p(x_{-j} \mid x_j, \theta, y)} \tag{3.24}$$

As can be seen, these expressions involve nested structures, and we need to use the Laplace approximation to obtain the marginal posterior distributions for model parameters and hyperparameters.

3.7.4 Dependency structures: Mesh, Matérn covariance and SPDE approach

Some statistical datasets not only exhibit dependencies between their values but also may contain spatial, temporal or spatio-temporal dependency structures. This means that there are correlations between the time and / or location of the observations and covariates (Jiang 2020). For example, earthquakes often occur in sequences and clusters within fault zones, with their locations influenced by underlying factors such as tectonic stress fields, stress redistribution by large events, fluid injection, and the rheological structure of the lithosphere (Hager et al. 2021). Here, I briefly explain the spatial models that capture this systematic variability, which can then be extended to spatio-temporal models through similar procedures.

When considering a spatial model, it is essential to use a mesh to investigate the relationships between points, discover dependencies, and predict the Gaussian Markov random field (GMRF) of our model (Hristopulos 2020). In INLA, a mesh is constructed in the study area, commonly using Delaunay triangulation (Lee and Schachter 1980; Liu and Yin 2020). It is crucial to construct a mesh that can resolve the intensity of the GMRF model, estimated at the vertices of the mesh, at an appropriate resolution (Juan Verdoy 2021).

The latent field x in an LGM model is assumed to be a Gaussian Markov Random Field (GMRF), as defined in Section 2.4.3, with a covariance matrix Σ (see Eq. 3.15). Since this covariance matrix is usually large, calculating it can be time consuming. To simplify the calculation, a mathematical function is imposed on its structure. Specifically, a Matérn correlation function is used to define the covariance:

$$\Sigma = \sigma^2 \text{Corr}_{\text{Matérn}} \tag{3.25}$$

where σ^2 is the variance parameter and $\text{Cor}_{\text{Matérn}}$ is the Matérn correlation function

defined by:

$$\text{Cov}_{\text{Matérn}}(s_i, s_j) = \frac{\sigma^2 2^{1-\nu}}{\Gamma(\nu)}(\kappa\|s_i - s_j\|)^\nu K_\nu(\kappa\|s_i - s_j\|) \tag{3.26}$$

In this equation, s_i and s_j are the spatial positions of observations i and j with a Euclidean distance $\|s_i - s_j\|$. The symbol Γ denotes the gamma function and K_ν is a Bessel function of the second kind. The parameter ν is usually fixed and determines the smoothness of the process, while κ is a scaling parameter related to the range over which the correlation diminishes, analogous to the range parameter in a variogram in geostatistics. If we assume $\nu = 1$, then $\Gamma(\nu) = 1$, and the equation simplifies to depend only on κ and the distance between points:

$$\text{Cov}_{\text{Matérn}}(s_i, s_j) = \sigma^2\kappa\|s_i - s_j\|K_\nu(\kappa\|s_i - s_j\|) \tag{3.27}$$

To quantify the covariance of the underlying Gaussian random field, we only need the two parameters σ^2 and κ. However, because the covariance matrix Σ is typically large, computations can be time-consuming and impractical. This is where the Stochastic Partial Differential Equation (SPDE) approach becomes useful. Lindgren et al. (2011) proposed that random field models could be described by solutions to sets of stochastic partial differential equations (SPDEs), expressed as:

$$(\kappa^2 - \Delta)^{\alpha/2} S(x) = W(x) \tag{3.28}$$

In this equation, κ is the same parameter as in the Matérn correlation function, Δ is the Laplace operator, $S(x)$ is the Gaussian field, and $W(x)$ is the Gaussian spatial white noise process, indicating that there is no spatial correlation in this term. The parameters of the SPDEs are directly linked to the parameters of the Matérn correlation, so solving the SPDEs provides the covariance matrix of the GMRF. The SPDE method also works well for log-Gaussian Cox process models (Rue et al. 2009; Krainski et al. 2018).

It should be noted that, similar to using the Matérn covariance to determine dependency structures in spatial models, the first-order autoregressive process (known as the AR1 process) can be used to calculate the residual covariance matrix and understand dependency structures in temporal models (time series) (Yan et al. 2021).

3.8 Comparison between MCMC and INLA

While both MCMC and INLA have their own advantages and disadvantages, INLA has demonstrated several benefits over MCMC. Below is a detailed comparison between INLA and MCMC:

- **Computational approach:** MCMC is a simulation-based method that is intuitive to learn and apply. It involves generating samples from the posterior distribution through iterative sampling methods such as the Metropolis-Hastings algorithm or Gibbs sampling. In contrast, INLA employs a deterministic mathematical algorithm that solves nested posterior structures and integrals using the Laplace approximation, making the process more direct and often less computationally intensive.
- **Computational costs:** MCMC can be computationally expensive, especially for complex models with high-dimensional parameter spaces. The computation can take weeks to months, depending on the complexity of the model and the convergence rate of the algorithm. INLA, however, performs calculations much faster, often in seconds to minutes, by approximating the integrals required for Bayesian inference, thus significantly reducing computational time and resource requirements.
- **Accuracy:** MCMC is considered an exact method for Bayesian inference as it fully explores the posterior distribution given enough time and iterations. INLA is an approximate method that uses the Laplace approximation to estimate posteriors. Despite being an approximation, studies (e.g. in Gómez-Rubio (2020) and Khan et al. (2021)) have shown that INLA achieves an accuracy comparable to MCMC, with similar errors in the final models.
- **Generality:** INLA is highly efficient and automated, making it particularly suitable for large-scale problems and complex hierarchical models. MCMC, although very flexible, may not scale well with the complexity of the model and the number of parameters, especially when there are many correlated parameters, which can lead to slow convergence and the need for extensive tuning.
- **Robustness:** When constructing hierarchical models with complex dependency structures, such as Latent Gaussian Models (LGMs), MCMC may encounter convergence issues and poor performance due to the intricacies of the model and the dependency among parameters. However, INLA has been specifically designed to

support LGMs and facilitate the direct computation of their posterior marginals, offering a more robust solution for these types of models.

- **Model comparison:** INLA not only computes posterior marginals, but also provides tools for model comparison and selection using common information criteria such as the Deviance Information Criterion (DIC) (Spiegelhalter et al. 2002) and the Watanabe-Akaike Information Criterion (WAIC) (Watanabe and Opper 2010). These criteria allow for practical model comparison and ranking, compensating for the fact that the Bayes factor cannot be used because the evidence term is not computed.
- **Predictive measures:** INLA utilises its marginal posteriors to compute Bayesian predictive measures, which are essential for model validation and comparison. These measures help detect outliers or surprising observations and ensure the model's predictive performance. Unlike the common practice of using the predictive density for each observation y_i based on all other observations, INLA approximates this quantity without reanalysing the model, making the process more efficient and straightforward.

3.9 The inlabru package

3.9.1 Introduction to inlabru

The inlabru package (Bachl et al. 2019) is an extension of the INLA methodology, specifically designed to facilitate the implementation of complex spatial and spatio-temporal models. Developed with a user-friendly interface, inlabru aims to simplify the specification of Bayesian hierarchical models, making advanced statistical tools more accessible to researchers.

The core strength of inlabru lies in its ability to handle latent Gaussian models (LGMs) with greater flexibility compared to traditional INLA approaches. It integrates seamlessly with the R-INLA package, leveraging its computational efficiency while providing an intuitive syntax for model formulation. This combination allows users to define intricate dependency structures and incorporate various types of data and covariates without the need for extensive programming knowledge.

One of the key features of inlabru is its ability to model complex spatial dependencies using Gaussian Markov Random Fields (GMRFs). It supports the construction of spatial meshes, which are essential for approximating continuous spatial processes. The package also facilitates the inclusion of different types of boundary conditions and constraints, enhancing the accuracy of spatial predictions.

In addition, inlabru offers tools for handling large datasets and high-dimensional models, which are common in ecological and environmental studies. Its integration with the R ecosystem allows easy manipulation of data and results, providing a streamlined workflow for data analysis and model validation.

To illustrate its capabilities, consider a scenario in which researchers need to model the distribution of a species across a heterogeneous landscape. Using inlabru, they can easily incorporate environmental covariates, spatial random effects, and measurement error into their model. The syntax of inlabru enables the specification of these components in a concise manner, facilitating the interpretation and communication of the model results.

In recent years, several studies have used inlabru for modelling in various fields, including ecology and environmental sciences (Adde et al. 2020; Bell et al. 2021; Cunningham et al. 2021a; Cunningham et al. 2021b; Jullum et al. 2020; Williamson et al. 2022), biology (Martino et al. 2021), seismology (Bayliss et al. 2020), and social sciences (Becker et al. 2021; D'Angelo et al. 2020; Ferraccioli et al. 2021; Hu et al. 2021). These applications demonstrate the versatility and robustness of inlabru in handling complex spatial data. Fig. 3.3 shows a general workflow for 2D time-independent seismicity modelling with inlabru (Bayliss et al. 2022).

3.9.2 Model implementation using iterative linearised method

The iterative linearised method in inlabru is designed to handle Bayesian generalised additive models with non-linear predictors. This approach first linearises the model and, once linearised, systematically seeks the solution through iterative refinement. This method builds on the standard INLA, which is efficient for models with linear predictors, and extends it to manage non-linear predictors effectively.

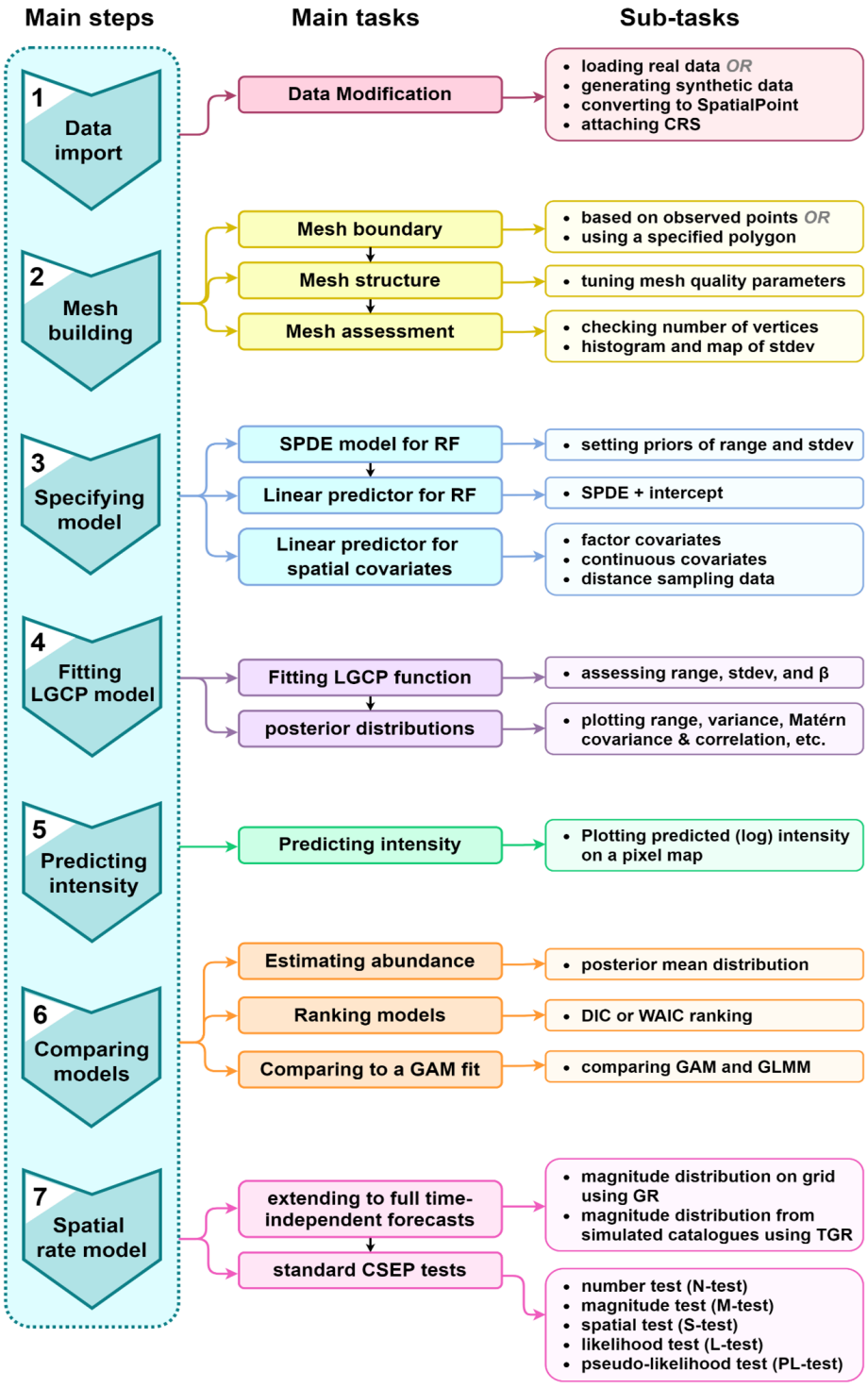

Figure 3.3: The workflow for 2D time-independent seismicity modelling with inlabru. This figure is the same as Figure 1 of Bayliss et al. (2022). Abbreviations: CRS (Coordinate Reference System), GR (Gutenberg-Richter), TGR (Truncated Gutenberg-Richter), SPDE (Stochastic Partial Differential Equation), RF (Random Field), LGCP (Log-Gaussian Cox Process), GAM (Generalised Additive Model), GLMM (Generalised Linear Mixed Model), DIC (Deviance Information Criterion), WAIC (Watanabe-Akaike Information Criterion), CSEP (Collaboratory for the Study of Earthquake Predictability).

3.9.2.1 Approximation for linear predictors

INLA is used to compute fast approximate posterior distributions for Bayesian models with latent Gaussian components. The hierarchical structure involves latent Gaussian variables u, covariance parameters θ, and observed response variables y. Typically, the linear predictor $\eta_i(u)$ links to the distribution's location parameter for observation y_i through a link function $g^{-1}(\cdot)$. In R-INLA, the observations are conditionally independent given η and θ.

3.9.2.2 Approximation for non-linear predictors

The inlabru method extends the INLA framework by introducing a linearisation step for nonlinear predictors. For a non-linear predictor $\tilde{\eta}(u)$, a first-order Taylor approximation around an initial point u_0 is used:

$$\bar{\eta}(u) = \tilde{\eta}(u_0) + B(u - u_0)$$

where B is the derivative matrix of the non-linear predictor evaluated at u_0.

The nonlinear observation model is approximated by this linearised model, allowing the use of standard INLA for the resulting linearised predictors. The posterior distribution is then factorised and approximated iteratively:

$$\tilde{p}(\theta, u \mid y) \approx \bar{p}(\theta, u \mid y)$$

Fixed point iteration

The fixed point iteration seeks an optimal linearisation point u^*. The process iterates as follows:

1. **Initial linearisation**: Start with an initial linearisation point u_0.
2. **Linearised posterior computation**: Compute the linearised INLA posterior at u_0.
3. **Posterior mode calculation**: Determine the posterior mode for θ and u given the linearised model.
4. **Line search**: Adjust the linearisation point using a line search method to minimise the difference between the non-linear and linearised predictors.

The iteration continues until convergence to a specified tolerance.

Line search

The line search step involves finding an optimal value α that minimises the norm of the difference between the non-linear and linearised predictors. This is achieved by considering the posterior variance of the predictors and using a quadratic approximation to the non-linear predictor as a function of α.

3.9.2.3 Posterior non-linearity checks

To ensure the accuracy of the approximation, the method evaluates the degree of nonlinearity in the vicinity of the optimal u^*. This can be done using sampling from the approximate posterior distribution and checking the variance-normalised squared deviation between the nonlinear and linearised predictors.

Accuracy

The accuracy of the approximation is assessed by comparing the true log-likelihood with the linearised log-likelihood. A Taylor expansion is used to derive the difference, focusing on the second-order terms. This leads to the evaluation of the Kullback-Leibler (K-L) divergence (Kullback and Leibler 1951) between the exact and approximate posterior distributions.

Initialisation and well-posedness

The linearisation point u_0 is crucial. Any value can be used unless the gradient at u_0 is zero, which might indicate a saddle point. In such cases, changing the initialisation point or the predictor parameterisation can help. This ensures that the predictor parameterisation does not lead to a multimodal posterior distribution or other ill-posed issues.

In summary, the iterative linearised INLA method in inlabru efficiently handles nonlinear predictors by iteratively refining a linear approximation, ensuring accuracy through fixed-point iterations and posterior nonlinearity checks. This approach extends the capabilities of INLA to a broader range of Bayesian models, maintaining computational efficiency and robust approximation accuracy. More details can be

found in https://inlabru-org.github.io/inlabru/articles/method.html. Fig 3.4 suggests how the inlabru workflow iteratively updates a set of trial parameters to approximate the ETAS model (Naylor et al. 2023).

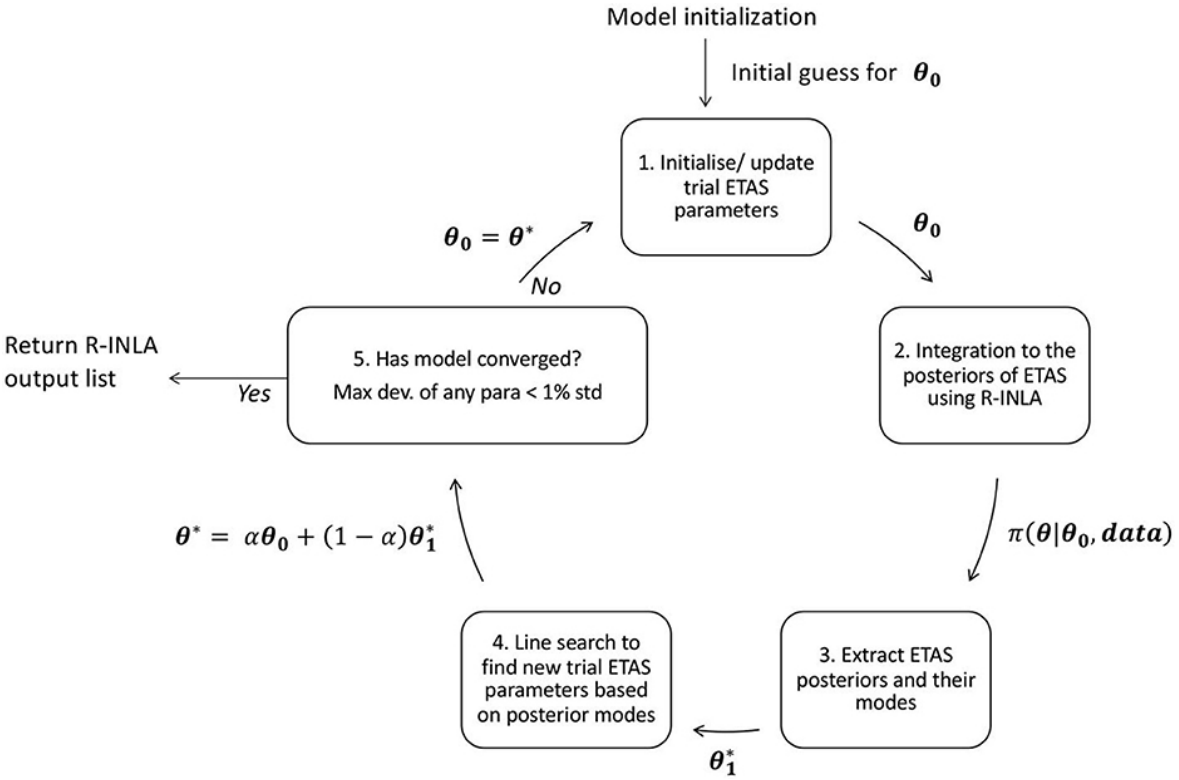

Figure 3.4: Schematic diagram of the inlabru workflow for iteratively updating a set of trial ETAS parameters. This figure is the same as Figure 2 of Naylor et al. (2023).

Chapter 4

Enhancing the ETAS model by incorporating short-term incompleteness in early aftershocks

4.1 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 and Richter 1944), which describes the distribution of earthquake magnitudes and their corresponding frequencies of occurrence; (2) the *modified Omori law* (Omori 1895a; 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 and 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.

During 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 parame-

ters through an optimisation algorithm. Important algorithms include gradient-based methods (e.g. Ogata 1998; Jalilian 2019), expectation maximisation (EM) (e.g. Veen and Schoenberg 2008; Mizrahi et al. 2023; Stindl and Chen 2023), nonlinear methods (e.g. Kanazawa and 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 the application of prior constraints on the model parameters and facilitates a more comprehensive exploration of the uncertainties associated with these parameters. Examples include Omi et al. 2015; Ebrahimian and Jalayer 2017; Shcherbakov et al. 2019; Ross 2021; Schneider and Guttorp 2021; Shcherbakov 2021; Laub et al. 2021; Ebrahimian et al. 2022; Molkenthin et al. 2022; Ross and Kolev 2022; Naylor et al. 2023; Nishikawa and Nishimura 2023.

The widespread use of the ETAS model comes with a significant challenge: assessing parameter estimations accurately when applied to real data is difficult. This difficulty arises because the methods listed above usually return a set of parameters 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; Arcangelis et al. 2018; Hardebeck et al. 2019; Lippiello et al. 2019a; Hainzl 2021; Mizrahi et al. 2021; Grimm et al. 2022; Iacoletti et al. 2022; 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 and 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 and Chiodi 2021; Chiodi et al. 2021).

In this chapter, I aim to enhance the accuracy of parameter estimations for the ETAS model when dealing with datasets characterised by short-term, time-varying incompleteness. I will use the (`INLA`) method (Rue et al. 2009), along with its extension, the `inlabru` package (Bachl et al. 2019), for computing posterior estimates of the parameters of the ETAS model 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 this study aim to address the 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. I begin by introducing the fundamental concepts and modifications that I have applied to the inversion algorithm of the ETAS model to address the issue of short-term incompleteness in the data, as detailed in Section 4.2. In Section 4.3, I assess and compare the performance of the original and my modified ETAS models using synthetic earthquake catalogues, demonstrating how my modifications enhance the accuracy of ETAS parameter estimation in the presence of incomplete data.

4.2 Methodology

4.2.1 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, I 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, I specifically focus on the temporal model with general form

$$\lambda_{\text{Hawkes}}(t|\mathcal{H}_t) = \mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} g(t) \tag{4.1}$$

where $\lambda_{\text{Hawkes}}(t|\mathcal{H}_t)$ represents the expected rate of events at time t, taking into account the history of the 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 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 called 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 commonly used in practice. Here, I consider one of the most common forms as

$$\lambda_{\text{Hawkes}}(t|\mathcal{H}_t) = \mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}, \tag{4.2}$$

where K, α, c, and p are the model parameters to be estimated along with μ (see Table 4.1).

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

$$\lambda_{\text{ETAS}}(t,|\mathcal{H}_t,m) = \left[\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}\right]\beta e^{\beta(m-M_0)}, \tag{4.3}$$

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, I focus primarily on the Hawkes part of the model as none of the G-R parameters are optimised in the inversion, but I will later use the properties of the magnitude model in addressing the censoring data in section 4.2.4.

Looking at Eq. (4.2), 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 the Utsu scaling law. This factor determines the increase in the 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 the triggering of 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, 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 an increase in intensity with each event and the temporal decay of it. In modelling aftershocks, this 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 the aftershocks, but in different ways. Conceptually, K is the base productivity parameter that quantifies the average number of direct aftershocks produced by an earthquake of reference magnitude M_0. To be exact, K is the change in intensity caused by a new event of 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 the model domain, that is, the expected rate increases at $t \simeq t_i$ for the 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 for 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. I will explore this issue in more detail in Section 5.4.

The other two parameters, c and p, control the decay of the Omori law. 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 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 delay period immediately after each triggering earthquake. It is a small, positive value that is used to avoid the singularity at $t = t_i$, ensuring finite rates at 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, lower values of c lead to a 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 milder 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. The

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 higher values of p indicating a faster decay in the rate of aftershocks, while lower values denote a slower decay. Physically, p is considered a region-based parameter and can vary depending on factors such as tectonic environment, temperature, magnitude, and depth of the mainshock, etc.

4.2.2 Approximation of parameters in the original ETAS model

In this section, I 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, I will further develop and adapt the solution for my modified version of the ETAS model, which specifically addresses the short-term incompleteness issue. This will be explored in detail in Sections 4.2.3 and 4.2.4.

In statistical modelling, the likelihood function plays a pivotal role in estimating the unknown model parameters. It quantifies how likely it is that 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

$$L(\boldsymbol{\theta}|\mathcal{H}) = \exp\left(-\int_{T_1}^{T_2} \lambda(t|\mathcal{H}_t)\, dt\right) \prod_{(t_i,m_i)\in\mathcal{H}} \lambda(t_i|\mathcal{H}_{t_i}), \tag{4.4}$$

or equivalently in logarithmic form as

$$\mathcal{L}(\boldsymbol{\theta}|\mathcal{H}) = \log L(\boldsymbol{\theta}|\mathcal{H}) = -\int_{T_1}^{T_2} \lambda(t|\mathcal{H}_t)\, dt + \sum_{(t_i,\, m_i)\, \in\, \mathcal{H}} \log \lambda(t_i\, |\mathcal{H}_{t_i})\ . \tag{4.5}$$

Here, $\lambda(t|\mathcal{H}_t)$ represents the intensity function as detailed in Eq. (4.2), and $\boldsymbol{\theta}$ denotes the vector of model parameters that I aim to estimate. For temporal ETAS modelling $\boldsymbol{\theta} = (\mu\ ,\ K,\ \alpha,\ c$ and $p)$. I use the logarithmic form of the likelihood function as it effectively transforms multiplications into additions, making complex calculations simpler and more numerically stable. By substituting Eq. (4.2) into Eq. (4.5) and then

solving the integral, the log-likelihood function is obtained as

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}|\mathcal{H}) = &-\mu\ (T_2 - T_1) \\ &- \sum_{(t_i,m_i)\in\mathcal{H}} K e^{\alpha(m_i - M_0)} \frac{c}{p-1} \left[\left(\frac{\max(T_1, t_i) - t_i}{c} + 1 \right)^{1-p} - \left(\frac{T_2 - t_i}{c} + 1 \right)^{1-p} \right] \\ &+ \sum_{(t_i,m_i)\in\mathcal{H}} \log \left(\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} K e^{\alpha(m_i - M_0)} \left(\frac{t - t_i}{c} + 1 \right)^{-p} \right), \end{aligned} \tag{4.6}$$

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 as 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 and 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, I adopt an exponential binning strategy for creation of a temporal mesh as proposed in Naylor et al. (2023):

$$\left\{ t_i\,, \quad t_i + \Delta\,, \quad t_i + \Delta(1+\delta)\,, \quad t_i + \Delta(1+\delta)^2\,, \quad \ldots\,, \quad t_i + \Delta(1+\delta)^{n_i}\,, \quad T_2 \right\}, \tag{4.7}$$

where $n_i \leq n_{\max}$, $\Delta > 0$ 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

$$\begin{aligned}\overline{\mathcal{L}}(\boldsymbol{\theta}|\mathcal{H}) = &-\exp\Big\{\log\mu+\log\left(T_2-T_1\right)\Big\}\\ &-\sum_{(t_i,m_i)\in\mathcal{H}}\sum_{j=0}^{B_i-1}\exp\Big\{\log K+\alpha(m_i-M_0)+\log\left(\frac{c}{p-1}\right)\\ &\quad+\log\left[\left(\frac{t_j^{b_i}-t_i}{c}+1\right)^{1-p}-\left(\frac{t_{j+1}^{b_i}-t_i}{c}+1\right)^{1-p}\right]\Big\}\\ &+\sum_{(t_i,m_i)\in\mathcal{H}}\log\left(\mu+\sum_{(t_i,m_i)\in\mathcal{H}_t}Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}\right).\end{aligned}\tag{4.8}$$

This formulation is then input into the `bru` function, which implements the `inlabru` workflow (Bachl et al. 2019) to estimate posterior density functions from the product of the prior distribution and likelihood function. The workflow takes the initial trial parameters for ETAS and iteratively updates these parameters based on the likelihood of observed earthquake data within a Bayesian context using `INLA`, where it makes 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, I will define a modified form of the conditional intensity function for the ETAS model that will accommodate short-term incompleteness. Then, I will modify the solution for the likelihood function, following a process similar to the steps explained above.

4.2.3 Model for transient short-term incompleteness in early aftershocks, $m_c(t)$

During an aftershock sequence, the overlap of numerous earthquake waveforms leads to censoring of smaller events, and hence to 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 the events. Helmstetter et al. (2006) proposed a model describing the evolution of the completeness magnitude of the form

$$m_c(t) = m_i - G - H\log_{10}(t-t_i)\tag{4.9}$$

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. (4.9) computed over all previous earthquakes (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, I 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

$$m_c(t) = M_m - G - H\log_{10}(t - T_m) \tag{4.10}$$

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. (4.10) and substituting $m_c(t) = M_0$, I can derive a formula to calculate the end of an incompleteness period following a mainshock, so that

$$T_e = T_m + 10^{(M_m - G - M_0)/H}. \tag{4.11}$$

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

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

4.2.4.1 Defining a time-dependent censorship function

My approach is to define a censorship factor between 0 and 1 that reduces the rates in the ETAS intensity function, and then modify the likelihood function accordingly, so that I can estimate the expected number of observed events, which can be directly compared to the catalogue. Building on section 4.2.3, I 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,

$$\begin{aligned} N(m \geq M_0) &= a10^{-bM_0}, \\ N(m \geq m_c(t)) &= a10^{-bm_c(t)}. \end{aligned} \tag{4.12}$$

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

$$\frac{N(m \geq m_c(t))}{N(m \geq M_0)} = 10^{-b(m_c(t)-M_0)}. \tag{4.13}$$

A similar approach was used by Stallone and Falcone (2021), who attempted to fill in the gaps and restore missing earthquakes, assuming that the Gutenberg-Richter law holds with the same exponent b in the censored part of the data. Unlike their method, my approach employs this ratio as a time-dependent censoring function, not aiming at recovering the missing events in the data, but rather to quantify the proportion of events observed given the censorship. Therefore, I use this to correct the estimate of the number of expected events above M_0, provided that we have a reasonable estimate of $m_c(t)$ at that 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 vary from underestimation to overestimation, and thereby could inadvertently introduce artefacts into the analysis. Substituting Eq. (4.10) into Eq. (4.13) for $m_c(t)$, I get

$$\pi(t) = \begin{cases} 1, & \text{if } t \leq T_m, \text{ or } t > T_e, \\ 10^{-b(M_m - G - H\log_{10}(t-T_m) - M_0)}, & \text{if } T_m < t \leq T_e, \end{cases} \tag{4.14}$$

where $\pi(t)$ is a piecewise function that determines the time-dependent censorship coefficient for my modified model. T_m is the time of the mainshock and T_e is the end of the incompleteness interval calculated in Eq. (4.11). For the period of incompleteness $(T_m, T_e]$, the ratio varies between $0 < \pi(t) < 1$ and represents the apparent (observed or recorded) rates as a fraction 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.

4.2.4.2 Modifying the intensity and likelihood functions

To incorporate the incompleteness model into the ETAS framework, I initially modify the ETAS conditional intensity function. This modification represents apparent rates

rather than actual rates, as

$$\lambda_{\text{apparent}}(t|\mathcal{H}_t, m_c(t)) = \left[\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}\right]\pi(t) \tag{4.15}$$

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

$$\lambda_{\text{modified}}(t|\mathcal{H}_t, m_c(t)) \simeq \mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}\pi(t). \tag{4.16}$$

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. (4.5), we have

$$\begin{aligned}\mathcal{L}_{\text{modified}}(\boldsymbol{\theta}|\mathcal{H}) \;=& -\int_{T_1}^{T_2}\mu\,dt \\ & -\int_{T_1}^{T_2}\sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}\pi(t)\,dt \\ & +\sum_{(t_i,m_i)\in\mathcal{H}}\log\left(\mu+\sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}\pi(t)\right)\end{aligned} \tag{4.17}$$

By solving the internal integral for the triggering part of Eq. (4.17) and subsequently incorporating the time binning strategy, along with linearisation (as previously

explained in Section 4.2.2), the modified log-likelihood is

$$
\begin{aligned}
\overline{\mathcal{L}}_{\text{modified}}(\boldsymbol{\theta}|\mathcal{H}) = &-\exp\left\{\log\mu + \log\left(T_2 - T_1\right)\right\} \\
&- \sum_{(t_i,m_i)\in\mathcal{H}} \sum_{j=0}^{B_i-1} \exp\left\{\log K + \alpha(m_i - M_0) + \log\left(\frac{c}{p-1}\right) + \right. \\
&\left. \log\left[\left(\frac{t_j^{b_i} - t_i}{c} + 1\right)^{1-p} - \left(\frac{t_{j+1}^{b_i} - t_i}{c} + 1\right)^{1-p}\right]\right\} \cdot I_1(t) \\
&- \sum_{(t_i,m_i)\in\mathcal{H}} \sum_{j=0}^{B_i-1} \exp\left\{\log K + \alpha(m_i - M_0) + \log\left(\frac{c}{p-1}\right) + \log\left(10^{-b(M_m - G - M_0)}\right) + \right. \\
&\log\left[\left[\left(\frac{t_j^{b_i} - t_i}{c} + 1\right)^{1-p} \quad (t_j^{b_i} - T_m)^{bH} \quad {}_2F_1\left(-bH, 1, 2-p, \frac{t_j^{b_i} - t_i + c}{t_j^{b_i} - T_m}\right)\right] - \right. \\
&\left.\left.\left[\left(\frac{t_{j+1}^{b_i} - t_i}{c} + 1\right)^{1-p} \quad (t_{j+1}^{b_i} - T_m)^{bH} \quad {}_2F_1\left(-bH, 1, 2-p, \frac{t_{j+1}^{b_i} - t_i + c}{t_{j+1}^{b_i} - T_m}\right)\right]\right]\right\} \cdot I_2(t) \\
&+ \sum_{(t_i,m_i)\in\mathcal{H}} \log\left(\mu + \sum_{(t_i,m_i)\in\mathcal{H}_t} K e^{\alpha(m_i - M_0)} \left(\frac{t - t_i}{c} + 1\right)^{-p} \pi(t)\right)
\end{aligned}
\tag{4.18}
$$

where ${}_2F_1$ denotes a Gaussian hypergeometric function (Ponnusamy and Vuorinen 2001; Aomoto 2011). This solution represents a joint demonstration of likelihood comprising the previous solution (Eq. 4.8) and the new one. Within the incompleteness interval, I adhere to the new solution with the applied censorship. Outside of this interval, where $\pi(t) = 1$, I switch to the original solution. To determine the appropriate solution, I use the indicator functions $I_1(t)$ and $I_2(t)$, ensuring that the correct solution is applied as needed. These indicators are defined as follows:

$$
I_1(t) = \begin{cases} 1, & \text{if} \quad T_1 \le t \le T_m \quad \text{or} \quad T_e < t \le T_2 \\ 0, & \text{otherwise,} \end{cases}
\tag{4.19}
$$

and

$$
I_2(t) = \begin{cases} 1, & \text{if} \quad T_m < t \le T_e \\ 0, & \text{otherwise.} \end{cases}
\tag{4.20}
$$

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

4.2.4.3 Considerations for time binning in the modified ETAS

As described above, I 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, I 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 the binning defined by Eq. (4.7), the temporal effect domain of each triggering event i is divided 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 is adequate in terms of accuracy and computational costs (Naylor et al. 2023). In the modified ETAS model, the computation of integral values over 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 by which the integration solution is amalgamated, depending on the positions of the bins, further contributes to these challenges.

Here, I 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. 4.1 illustrates general examples of triggering events that can occur 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, I 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. 4.1-a, b, c, and d), such that $[T_{1b}, T_{2b}] \leq T_m$ or $[T_{1b}, T_{2b}] > T_e$, I 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. 4.1-b and c), such that $T_m < [T_{1b}, T_{2b}] \leq T_e$, I 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. 4.1-a), so that $T_{1b} < T_m$ and $T_{2b} > T_e$, I 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. 4.1-b), so that $T_{1b} < T_m$ and $T_m < T_{2b} < T_e$, I split the bin into two sub-bins with length of $[T_{1b}, T_m]$ and $(T_m, T_{2b}]$, respectively. I 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. 4.1-b and c), so that $T_m < T_{1b} < T_e$ and $T_{2b} > T_e$, I 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.

To ensure that the hypergeometric function remains well defined and the computations do not diverge when a bin starts precisely at the mainshock time ($t_j^{b_i} = T_m$), as specified in Eq. (4.18), a small positive value, $\epsilon = 10^{-10}$, is added to the bin's starting time. This addition effectively shifts the starting point slightly, preventing the likelihood calculations from approaching infinity or becoming undefined.

Conceptually, in the modified model, the bins that fall within the incompleteness period show lower event rates compared to the original model, due to 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 a full capture of seismic activity without

censorship. However, within the incompleteness period, 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. At the end of the incompleteness period (T_e), the detection ratio reaches 1, indicating complete sampling of all events without missing data.

In Section 5.2, I will further investigate the model's sensitivity to time binning and demonstrate its practical implications.

4.2.5 Synthetic catalogue generation using the ETAS model

Since I will assess the performance of the ETAS model using both synthetic and real earthquake catalogues in the next sections, I first detail the process of generating synthetic data here. Synthetic catalogues are created using the ETAS model, with specific values selected as the true parameters. I use the same algorithm applied by Naylor et al. (2023), which follows these steps:

1. **Background events**: A Poisson process is used to simulate background seismicity. The expected number of background events is determined by the ETAS parameter μ and the total time interval $[T_1, T_2]$. Event times are uniformly sampled within this interval, and the magnitudes are drawn from a Gutenberg-Richter (GR) distribution.

2. **Aftershock generation**: Each background event can trigger aftershocks based on the ETAS triggering function. The number of aftershocks is determined by the productivity parameter K, magnitude, and time of the parent event. Aftershock times are sampled from Omori's law, which describes the temporal decay of aftershock rates, and magnitudes are again drawn from the GR distribution.

3. **Historic events**: If a set of known events is provided (e.g., historic or significant seismic events), these are used to precondition the sequence. The algorithm generates aftershocks for these events, ensuring they influence the synthetic catalogue. This allows for the inclusion of real-world events as drivers of synthetic aftershock sequences.

Table 4.1: Glossary and description of variables and parameters.

in text	description
t	a time point at which we evaluate the intensity
t_i	time of the i^{th} triggering event
$t_j^{b_i}$	start time of j^{th} bin of the i^{th} triggering event (left edge)
$t_{j+1}^{b_i}$	end time of j^{th} bin of the i^{th} triggering event (right edge)
T_1	start time of the modelling domain
T_2	end time of the modelling domain
$\max(T_1, t_i)$	time for either 'including' or 'conditioning' on history
T_m	time of the mainshock (also start time of the incompleteness interval)
T_e	end time of the incompleteness interval
c	time shift to avoid infinity at $t = t_i$
m_i	magnitude of the i^{th} triggering event
M_m	magnitude of the mainshock
M_0	triggering threshold: minimum magnitude capable of triggering events
m_c	detection threshold: magnitude above which all events are detected
$m_c(t)$	short-term time-varying magnitude of incompleteness
μ	background seismicity
K	base productivity parameter
α	magnitude scaling productivity parameter
p	decay speed of aftershock rates
b	b-value of the Gutenberg-Richter relation
G	baseline magnitude shift in incompleteness model
H	log-time scaling parameter in incompleteness model
λ	intensity or rate of aftershocks
$g(t)$	ETAS triggering function
$\mathcal{H}$	time history of aftershock evolution
$\mathcal{L}$	likelihood of model
θ	vector of all model parameters
Δ	base time increment in time binning
δ	growth factor in time binning
n_{max}	maximum number of bins for each event in temporal mesh

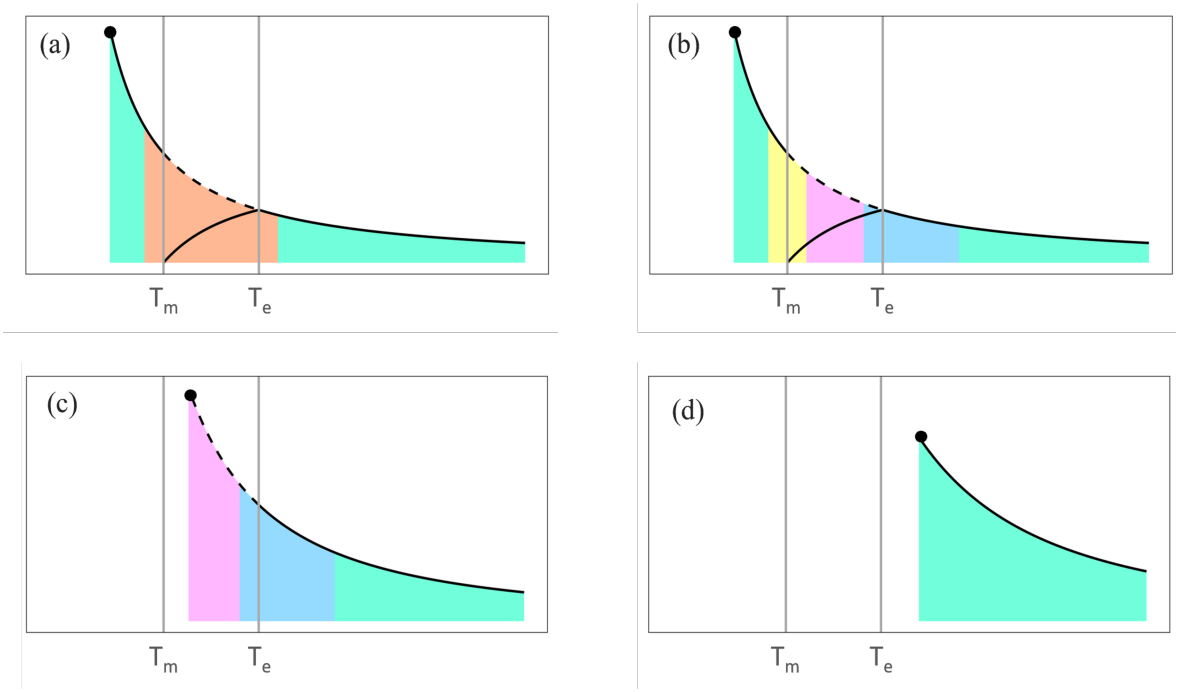

Figure 4.1: A schematic representation of time-binning considerations in the modified ETAS model. T_m and T_e represent the time of the mainshock and the end of the incompleteness period, respectively. Panel (a) and (b) illustrate triggering events occurring before the mainshock; Panel (c) demonstrates 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 4.2.4.3.

4. **Generational structure**: The model assigns each event a generational label:

- **Generation -1**: Events from the historic catalogue.
- **Generation 0**: Aftershocks of historic events.
- **Generation 1**: Background events.
- **Generation 2 and beyond**: Successive aftershocks triggered by first-order aftershocks or background events.

The process continues iteratively, generating higher-order aftershocks until no new

events are triggered within the time interval. The final catalogue is returned as a list or data frame containing the event times, magnitudes, and generation numbers.

4.3 Performance assessment of the modified ETAS model using synthetic data

Here, I present the results of my modified ETAS model, which extends the standard ETAS framework by incorporating a time-dependent censorship function to address the challenge of short-term, time-varying incompleteness in early aftershocks.

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. 4.10) estimates this evolving completeness threshold. When we look at real data in Section 5.7, my 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, I first demonstrate the efficiency of my modified ETAS model through synthetic experiments. In doing so, I generate sets of synthetic catalogues by creating complete catalogues and then remove all events below the completeness predicted by the Helmstetter model. Hereafter, I refer to the data before removal as the 'complete' catalogue and to the data after removal as the 'incomplete' catalogue. 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 (I will discuss this later in Section 5.3), 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 and Knopoff (1974).

These catalogues are generated with a background rate of $\mu = 0.1$ events per day, a rate consistent with moderate to highly seismic regions. The Gutenberg-Richter b-value parameter is also set to $b = 1$ for this study. In addition, I set a constant magnitude threshold at which the catalogue is complete except below $M_0 = 2.5$. The true ETAS parameters were established as $K = 0.15$, $\alpha = 2.29$, $c = 0.05$, and $p = 1.08$,

shown by the vertical dashed lines in Figure 4.7. For the incompleteness models, I adopted the 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 can vary based on the seismicity characteristics of a particular region and the ability of seismic networks to record and discriminate between events. In subsequent sections, I will explore different parameters to demonstrate my model's capability to adapt to different incompleteness behaviours. My modified model is versatile, accommodating a wide range of mainshock magnitudes and incompleteness parameters. As a representation, I provide four synthetic data samples with mainshock magnitudes of 6.0, 6.5, 7.0, and 7.5, as depicted in Fig. 4.2. 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.

Then, the original ETAS model was fitted to both the complete and incomplete datasets, while the modified ETAS model was applied exclusively to the incomplete dataset. For the starting values in the fitting process, I used the initial guesses for the ETAS parameters as follows: $\mu = 0.3$, $K = 0.1$, $\alpha = 1$, $c = 0.2$, and $p = 1.01$. These initial values were chosen to be in a reasonable range usually calculated for each parameter but not exactly the same as the true values to avoid overfitting based on prior knowledge.

The pair plot depicted in Fig. 4.3 provides a detailed visualisation of the univariate distributions and bivariate relationships among the parameters μ, K, α, c, and p. The marginal distributions of each parameter are displayed on the diagonal of the plot, illustrating their individual characteristics. Off the diagonal, scatter plots and contour plots reveal the interactions and correlations among these parameters. This plot helps to understand the dependencies and relationships within the model parameters. The contour and scatter plots between c and K suggest that higher values of c are associated with slower values of K, indicating a potential link between the productivity of the aftershock sequences and the time adjustment required to prevent singularities in the model. Furthermore, a clear correlation between c and p suggests that when the time shift c is smaller, the aftershock activity tends to decay more quickly. This relationship

is important for understanding how the time shift parameter influences the overall duration of aftershock sequences.

Figures 4.4 to 4.6 illustrate the convergence behaviour of the ETAS model parameters over multiple iterations, focusing on key parameters. The top-left panel tracks the mode estimates of each parameter across iterations, showing how they stabilise. The mode represents the most probable value of each parameter after each update. The top-right panel compares the current mode estimates with the linearisation-based estimates, helping assess the accuracy of the linear approximations. The bottom left panel shows the normalised changes in mode estimates relative to the standard deviation, highlighting how rapidly each parameter converges. Finally, the bottom-right panel compares the absolute changes in the mode estimates with the standard deviation, indicating whether the parameter estimates are stabilising within a reasonable range of uncertainty. These plots demonstrate the robustness of the optimisation process, as most parameters stabilise after a small number of iterations. It should be noted that in the convergence plots, the ETAS parameters are represented on the θ scale, which is the internal scale used by INLA to estimate the model and follows a Gaussian distribution with a default mean of 1 and a precision of 10. For presenting the posterior distributions, I converted the parameters from the INLA scale to the ETAS scale.

The estimated posteriors are illustrated in Fig. 4.7, with detailed information provided in Table 4.2. As the results indicate, the original ETAS model, when fitted to 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 fitted to incomplete data (red posteriors), a noticeable bias in parameter estimations occurs. This bias becomes more pronounced as the magnitude of the mainshock increases and the original model fails to accurately retrieve the true parameters. In contrast, the modified model, fitted to 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 from the original ETAS model fitted to complete data. This indicates that, despite being fed incomplete data, the modified ETAS model can achieve accurate posteriors akin to those of the original ETAS model fitted to complete data, provided we can parameterise the censoring pro-

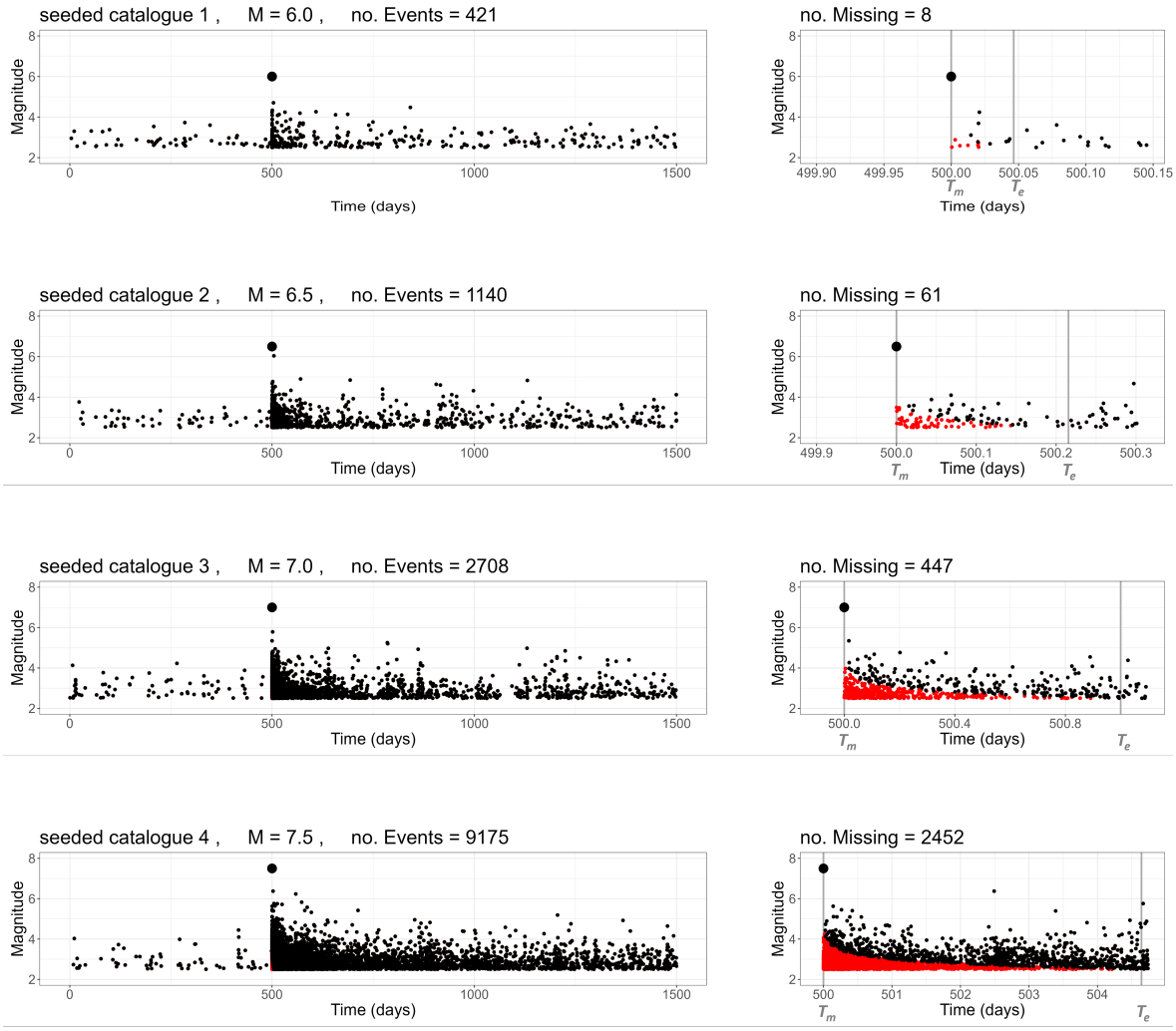

Figure 4.2: (Left): Four generated 1500-day synthetic catalogues with mainshock magnitudes of 6.0, 6.5, 7.0, and 7.5, seeded on day 500. (Right): Zoomed-in representations showing recorded events in black and missing events in red within the short-term incompleteness interval for each sequence. The grey vertical lines represent the time of the mainshock (T_m) and the end of the incompleteness period (T_e). 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.

cess. The modified model's posteriors not only align closely with the true values but also fall within the uncertainty range of the original model (blue posteriors), although they exhibit a shorter peak, indicating greater certainty in the estimates compared to the broader uncertainty seen in the original model when fitted to incomplete data.

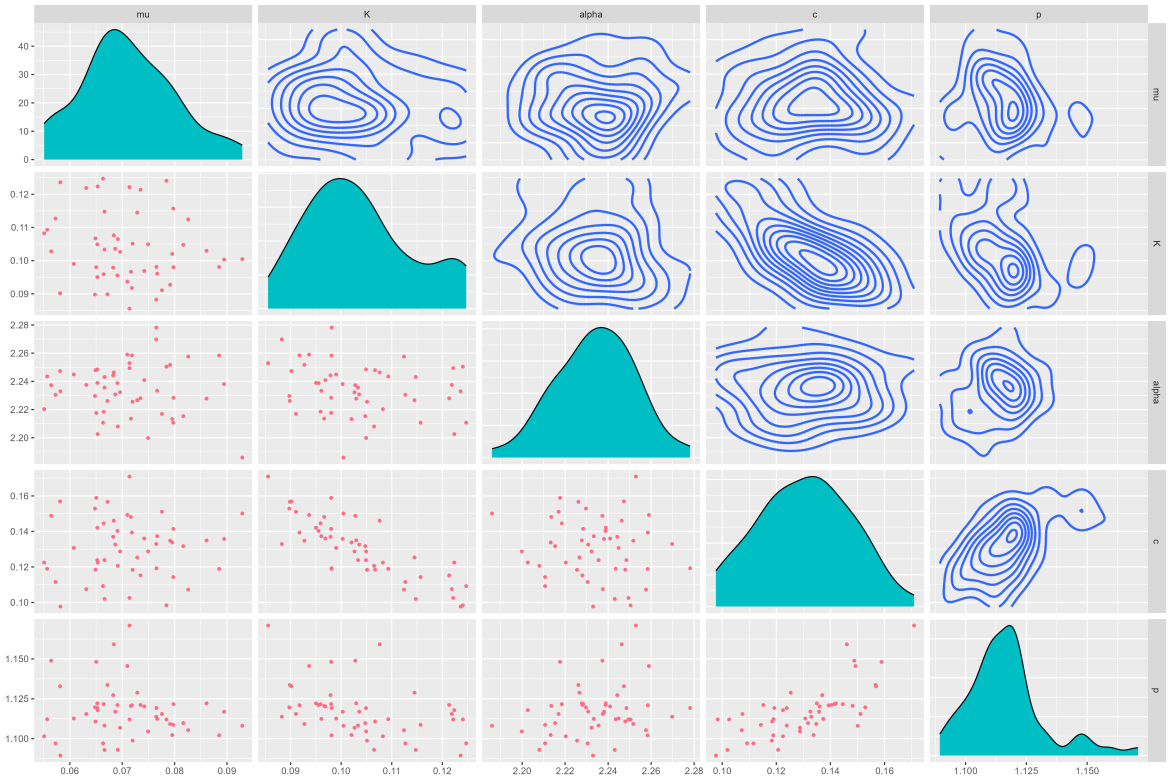

Figure 4.3: Pair plot of ETAS model parameters, showcasing the univariate distributions and bivariate relationships. This plot illustrates the distributional shapes of parameters μ, K, α, c, and p, along with their pairwise scatter plots and contour plots that reveal correlation patterns. Each parameter's marginal distribution is displayed along the diagonal, with scatter plots and contour lines capturing the interdependencies between parameters off the diagonal. This visualisation aids in understanding the interactions and potential correlations within the model parameters.

Although there are some small systematic differences between the true and posterior mode values, these differences are within an acceptable range. In Bayesian analysis, the mode is a representative point estimate, but the true parameters typically lie within the uncertainty range of the full posterior distributions of the modified model. As the magnitude increases, the original ETAS model exhibits a greater bias, and true values often fall outside of its uncertainty range. This means that sampling from the original model's posterior may not recover the true parameters, unlike the modified model, which consistently provides estimates closer to the true values even with incomplete data.

The triggering function (Eq. 4.2) is presented as another piece of evidence sup-

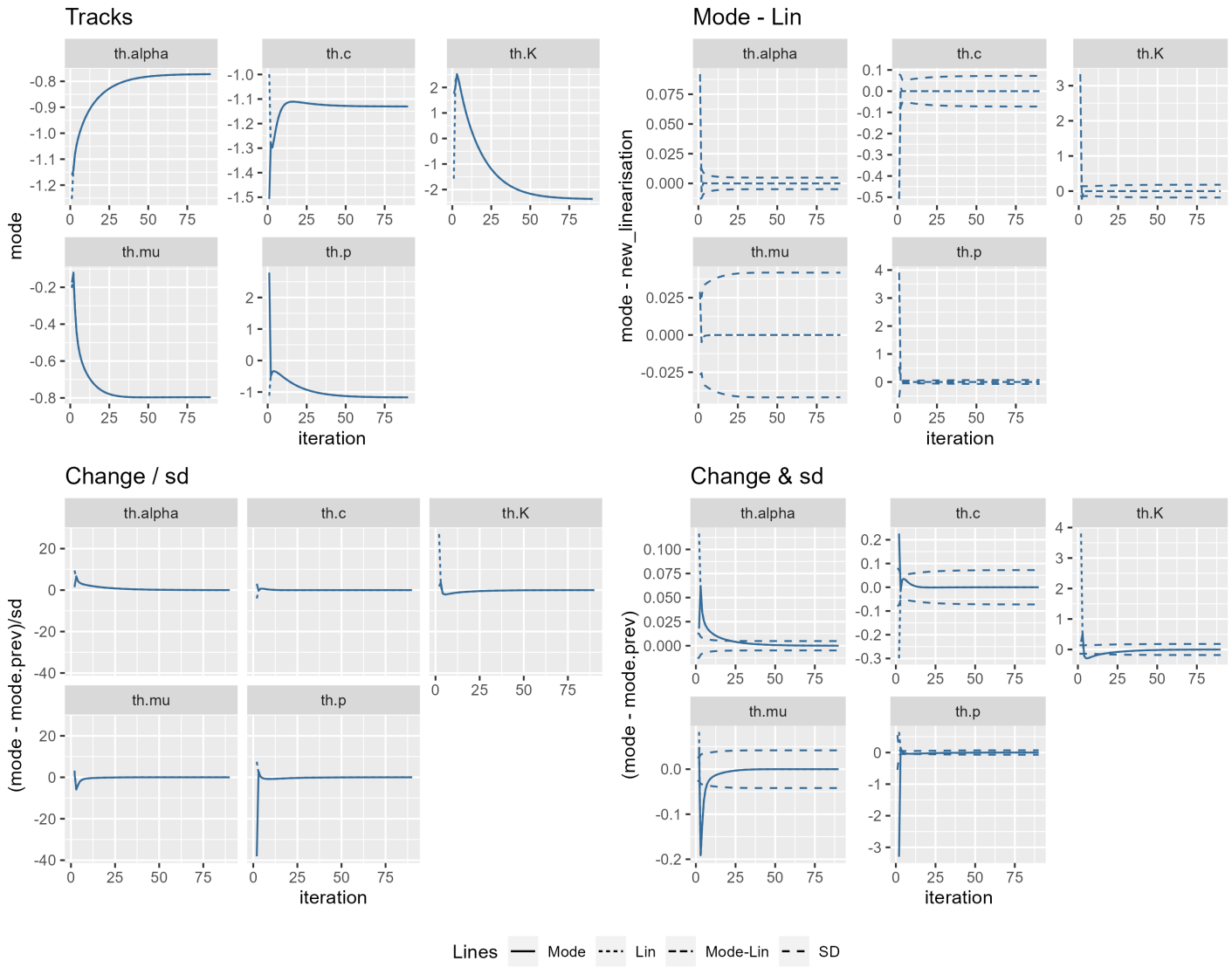

Figure 4.4: Convergence plot for the original ETAS model fitted to complete data. The panels show: (top-left) the mode estimates across iterations, (top-right) the difference between the mode and linearised estimates, (bottom-left) normalised changes in mode relative to the standard deviation, and (bottom-right) absolute changes in mode compared to the standard deviation.

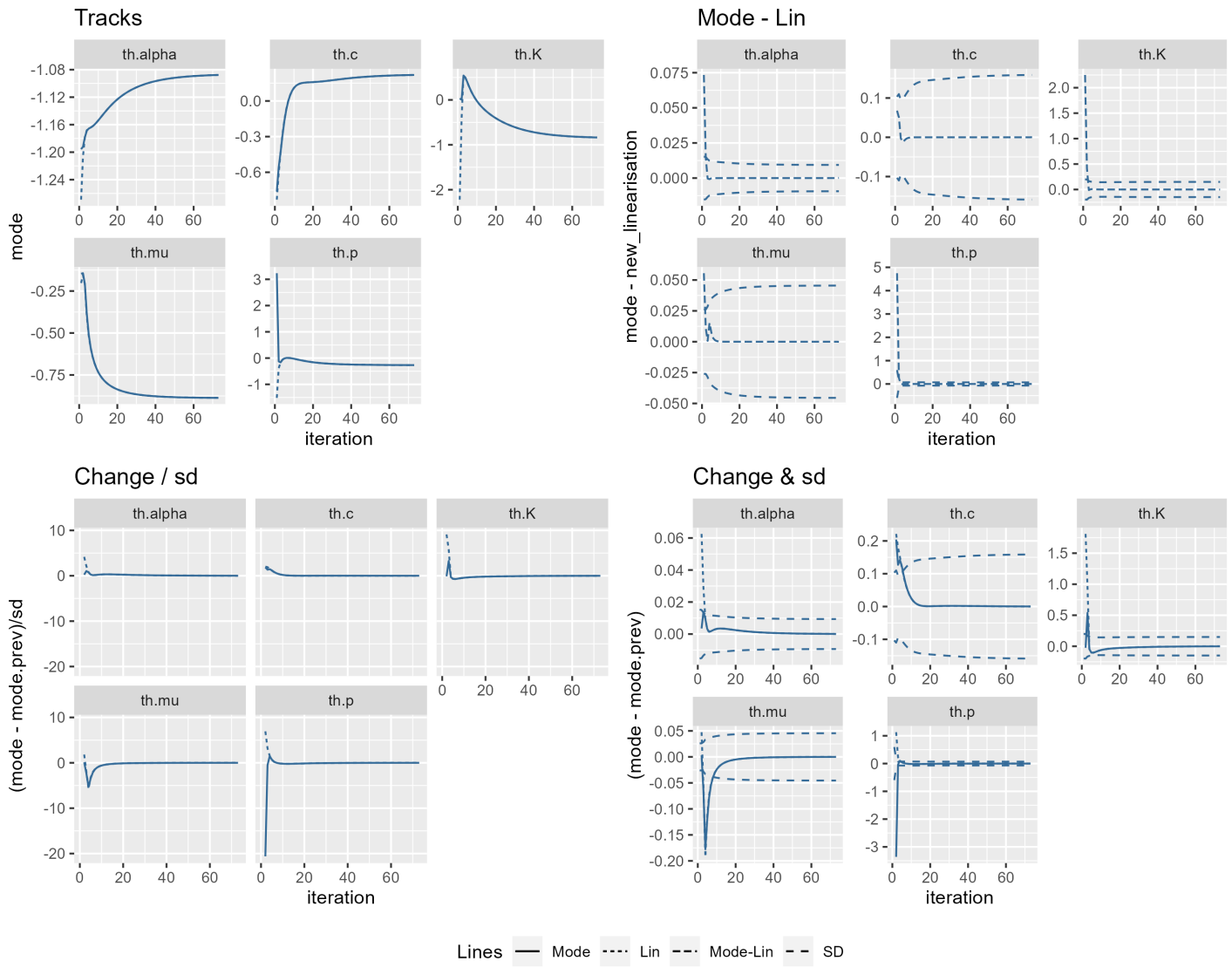

Figure 4.5: Convergence plot for the original ETAS model fitted to incomplete data. The panels show: (top-left) the mode estimates across iterations, (top-right) the difference between the mode and linearised estimates, (bottom-left) normalised changes in mode relative to the standard deviation, and (bottom-right) absolute changes in mode compared to the standard deviation.

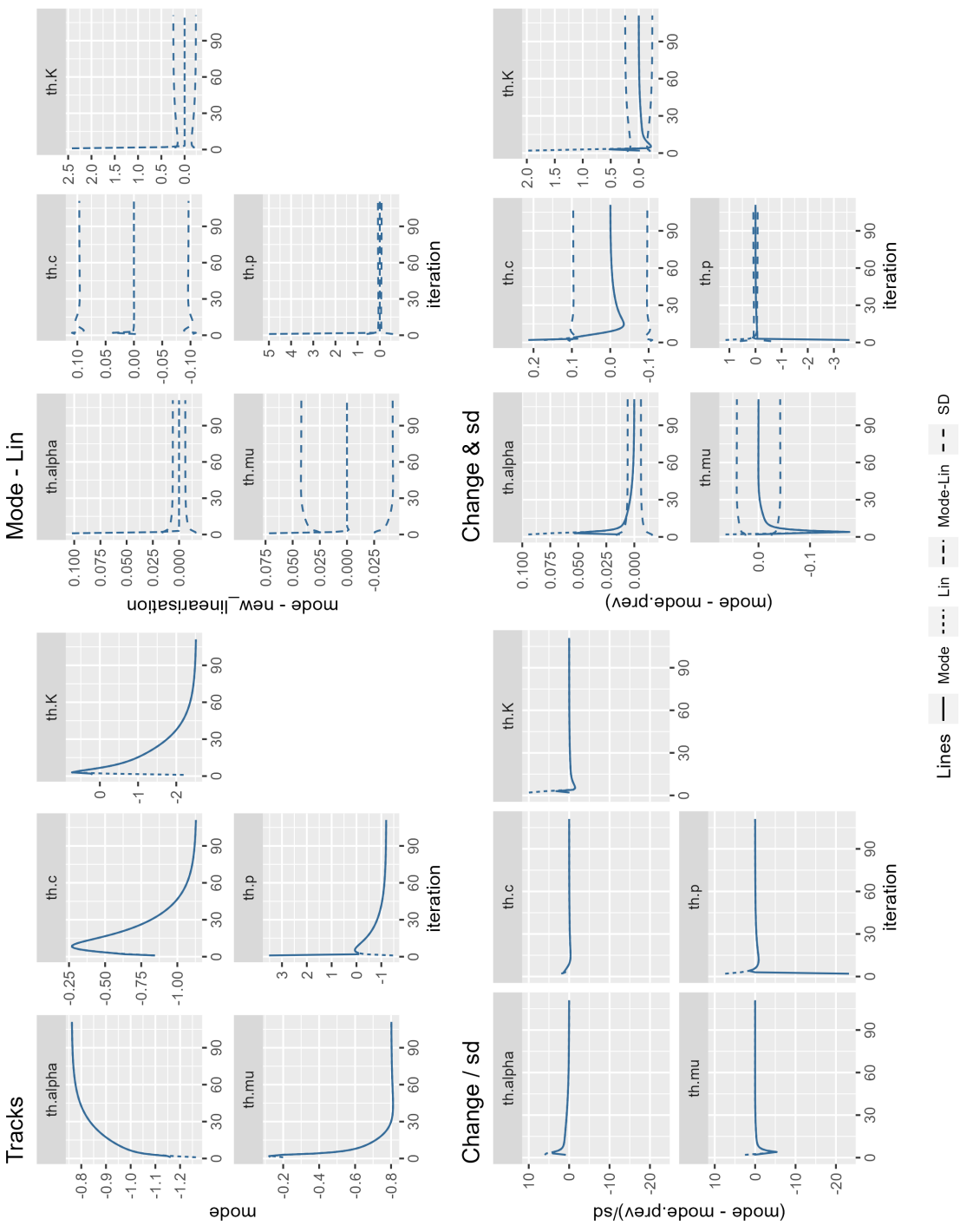

Figure 4.6: Convergence plot for the modified ETAS model fitted to incomplete data. The panels show: (top-left) the mode estimates across iterations, (top-right) the difference between the mode and linearised estimates, (bottom-left) normalised changes in mode relative to the standard deviation, and (bottom-right) absolute changes in mode compared to the standard deviation.

porting the improved performance of the modified model, as it demonstrates how the model captures the rate of triggered events across different magnitude thresholds. An illustrative example of this function, using the synthetic catalogue with a mainshock magnitude of $M = 7.0$, is presented in Fig. 4.8. The triggering function derived from the original ETAS model, which was fitted to incomplete data, shows a considerable underestimation of event rates (Fig. 4.8 - middle column). In contrast, the triggering function of the modified ETAS model fitted to incomplete data (Fig. 4.8 - right column) closely mirrors that of the original model when fitted to complete data (Fig. 4.8 - left column). Due to the missing portion of events in the incomplete data, the triggering function of the modified ETAS model, fitted to these data, displays a slightly wider uncertainty band than that of the original model fitted to complete data. However, it still remains well within the latter's uncertainty bounds. This demonstrates that my 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 4.9 displays the actual rates (in black) and the apparent (observed/recorded) rates (in purple) for my four selected synthetic catalogues. Then, I 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 ETAS and the modified ETAS models when fitted to incomplete data, highlighting the improvements made in the modified version.

4.4 Conclusion

In this chapter, I addressed the challenge of short-term, time-varying incompleteness in early aftershocks by enhancing the Epidemic Type Aftershock Sequence (ETAS) model. Traditional ETAS models often fail to account for the transient period follow-

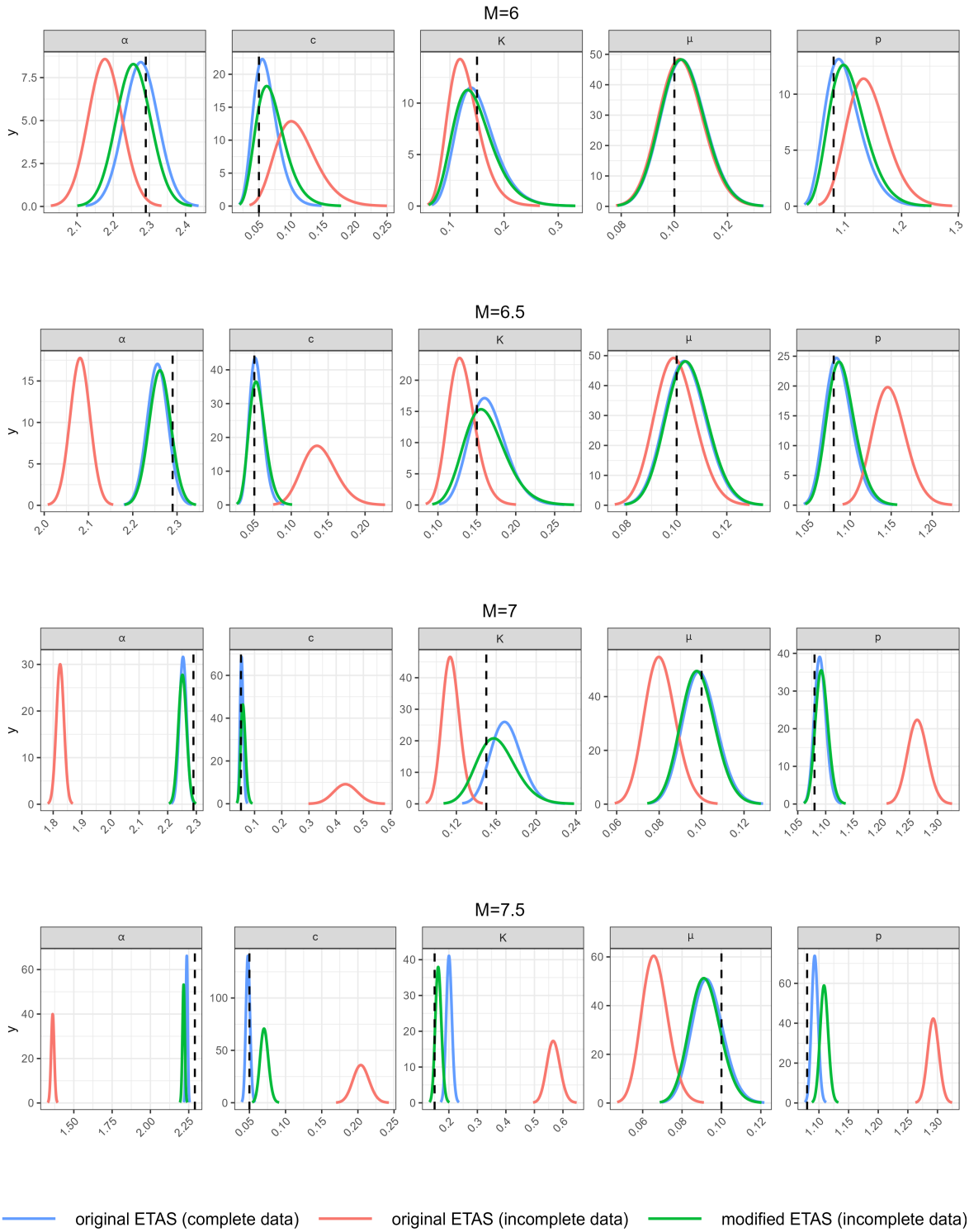

Figure 4.7: Posteriors presented for four simulated catalogues corresponding to mainshock magnitudes of 6.0, 6.5, 7.0, and 7.5. The blue and red posteriors are derived from the original ETAS model using complete and incomplete catalogues, respectively. The green posteriors are obtained by fitting the modified ETAS model to 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 4.2: Details on data incompleteness and parameter estimates from the original and the modified ETAS models fitted to 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

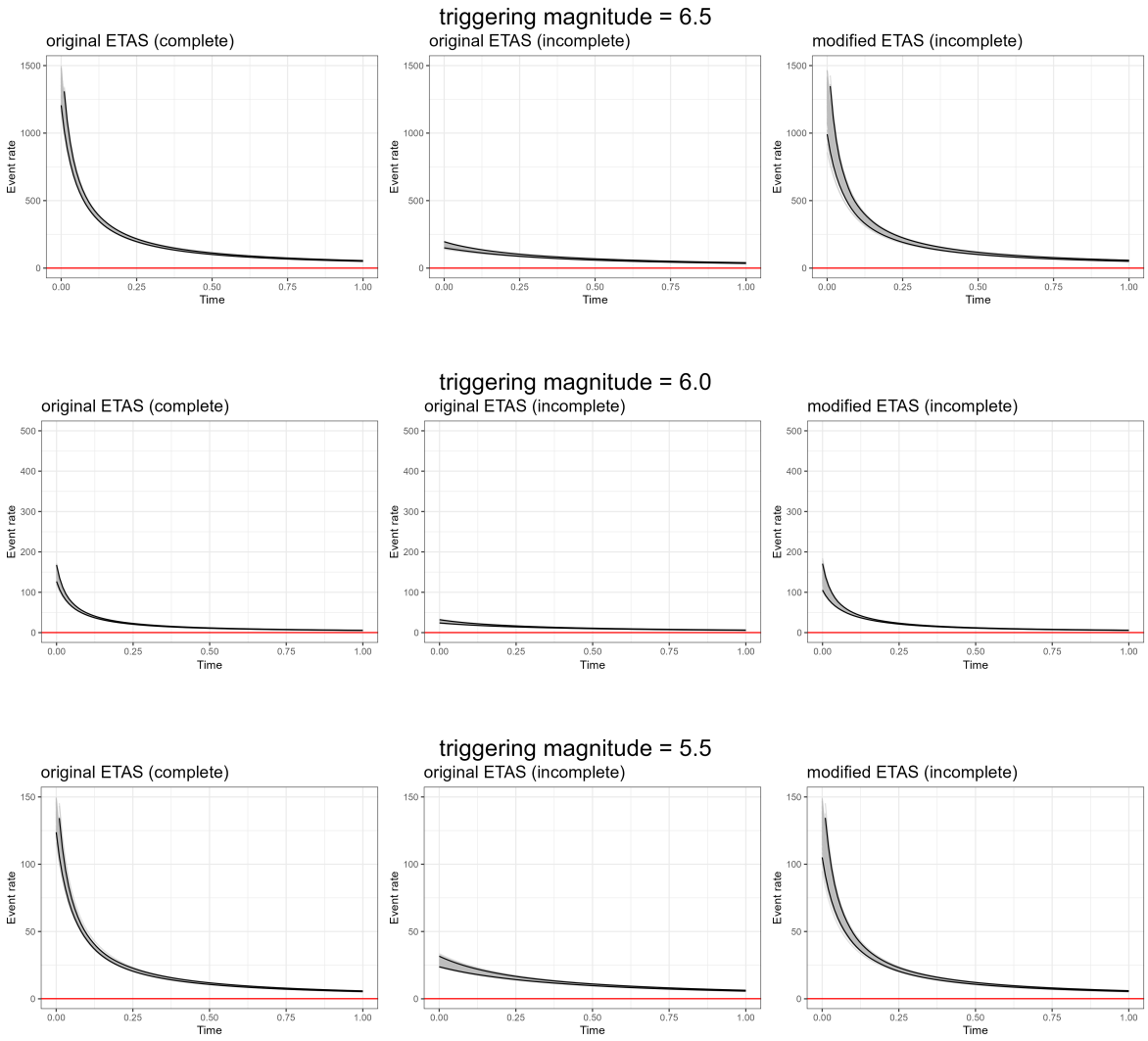

Figure 4.8: Triggering functions for the synthetic catalogue with mainshock magnitude of 7.0. The functions display rates and associated uncertainties for 1 day after the mainshock at different triggering magnitude levels (6.5, 6.0, and 5.5). These triggering functions are obtained from three scenarios: the original ETAS model fitted to complete data (left column), the original ETAS model fitted to incomplete data (middle column), and the modified ETAS model fitted to 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.

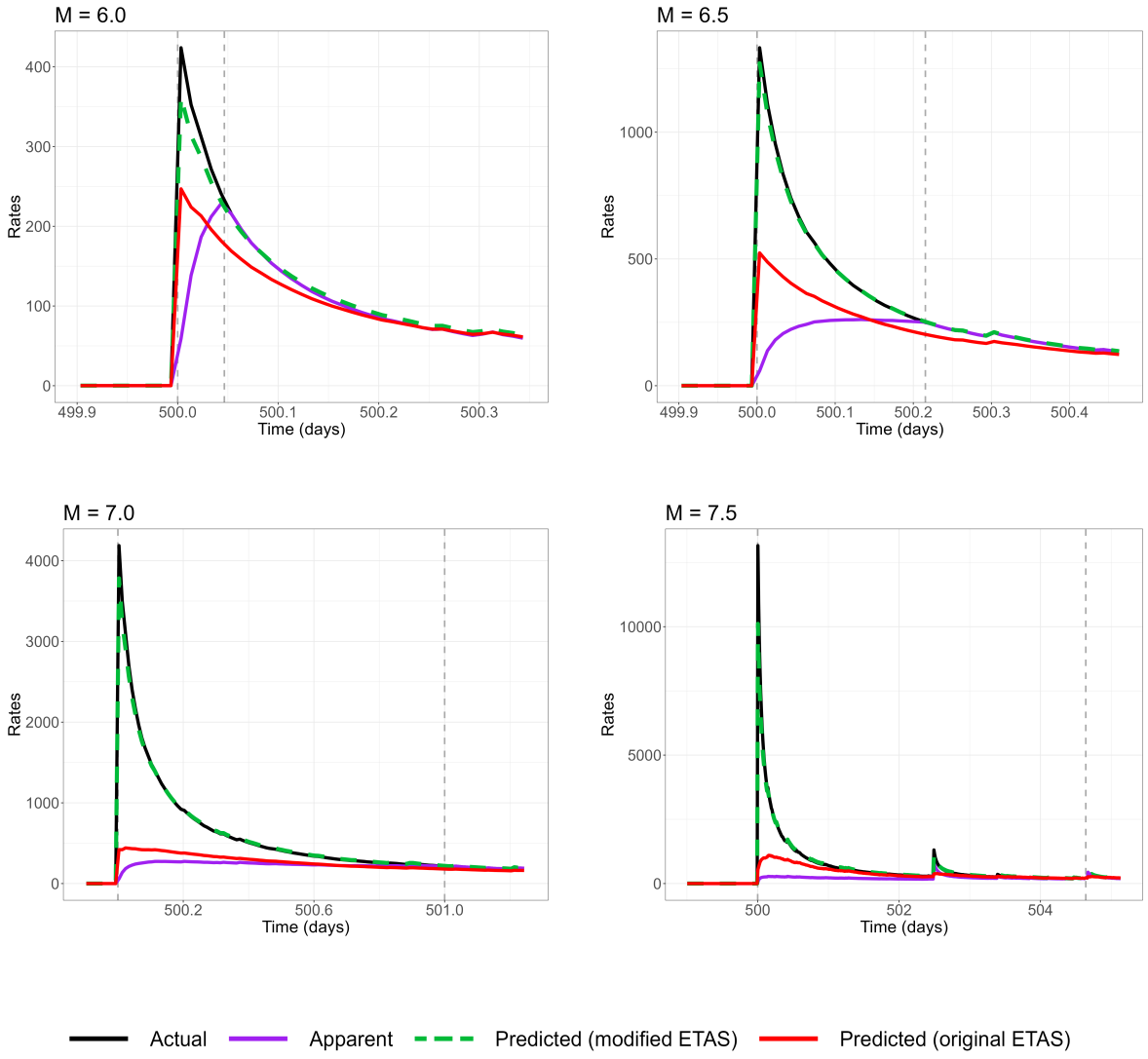

Figure 4.9: The actual (in black), apparent (in purple), and predicted intensities for four simulated catalogues with mainshock magnitudes of 6.0, 6.5, 7.0, and 7.5. The predicted intensities are calculated using posterior modes obtained from both the original and the modified ETAS models, each fitted to 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.

ing a significant mainshock during which smaller aftershocks are under-recorded. My approach introduced a time-dependent censorship function into the ETAS framework, significantly improving the model's accuracy in such scenarios.

The core of my modification involved defining a time-varying completeness magnitude, $m_c(t)$, which dynamically adjusts based on the rate of seismic activity. This model, introduced by Helmstetter et al. (2006), estimates the completeness level as a function of time after a mainshock, incorporating parameters G and H to describe the baseline change and the log-time scaling, respectively. In doing so, I effectively captured the short-term increase in the magnitude threshold to record smaller aftershocks.

To integrate the censorship function into the ETAS framework, I modified the conditional intensity function to reflect apparent rates rather than actual rates during the incompleteness period. This adjustment necessitated changes to the likelihood function, which I then solved using a combination of temporal binning and linearisation strategies. The binning strategy, crucial for handling rapid rate fluctuations post-mainshock, provides precise calculation of expected event numbers within each bin.

Through synthetic experiments, I demonstrated that the modified ETAS model significantly reduces bias in parameter estimations compared to the original model when fitted to incomplete data. The results showed that my model could accurately retrieve the true ETAS parameters (by which I generated synthetic data) even in the presence of incompleteness, closely aligning with the results obtained from the original ETAS fitted to complete data. This was evident in both the posterior distributions of ETAS parameters and the triggering functions, where my modified model consistently outperformed the original model in estimating rates during the incompleteness period.

Additionally, I showed that the modified model's predicted intensities are close to actual rates, especially for larger mainshock magnitudes associated with longer periods of incompleteness and a higher number of missing events. This indicates that the time-dependent censorship function effectively corrected for the underrecording of smaller aftershocks, providing more reliable forecasts of aftershock sequences.

The enhancements I introduced to the ETAS model address a significant gap in current seismicity modelling by accounting for short-term, time-varying incompleteness.

These improvements strengthen the model's accuracy and reliability and offer a framework adaptable to various seismicity characteristics and incompleteness behaviours, paving the way for more robust operational earthquake forecasting and seismic hazard assessment.

In the next chapter, I apply this model to optimise sampling strategies for ETAS inversions. This involves addressing key considerations for sampling, evaluating potential biases, and establishing guidelines for selecting representative samples. By refining these strategies, the goal is to achieve accurate parameter estimation, which is crucial for the practical application of the modified ETAS model developed in this chapter. This optimisation will be demonstrated through synthetic experiments and also real data applications, ensuring that the methodological advances lead to tangible improvements in seismic hazard assessments and forecast reliability.

Chapter 5

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

5.1 Introduction

Here, I 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. In addition, 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 ETAS parameters can be well constrained. In practice, individual case studies may not have enough data to allow this. However, by being aware of the deficits in specific case-study data, we can anticipate the limitations of our parameter estimations.

Here, I 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. I 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.

In this section, I start with a preliminary assessment of the sensitivity of my modified ETAS model to different time-binning choices in Section 5.2, offering guidance

in selecting an optimal temporal mesh to ensure reliable posteriors. In Sections 5.3 to 5.9, I introduce six key considerations to help practitioners select a representative sample that results in more accurate and unbiased ETAS inversions. These include the impacts of including and conditioning on the historic run-in period, combination of magnitudes and trade-off between K and α, choice of incompleteness model parameters, presence of a secondary large aftershock, as well as size of temporal and spatial modelling domains. In addition, in Section 5.7, I compare the efficacy of both the original ETAS and the modified ETAS models through their application to three real earthquake sequences.

5.2 Sensitivity of the ETAS model to time-binning

To assess the sensitivity of the original ETAS and modified ETAS models to time-binning choices, I fit both models to an incomplete synthetic dataset using different binning options. I first simulate a complete synthetic catalogue that spans 1500 days, including a mainshock with magnitude $M6.7$ on day 500. Then, I filter out events with magnitudes below the incompleteness model (Eq. 4.10) from the catalogue. The parameters of the incompleteness model are specified as $G = 3.8$ and $H = 1$. Then, I consider five temporal meshes with different binning designs and resolutions. Table 5.1 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 run-time 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. 5.1 illustrates how binning affects the posteriors of the 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 the posteriors in the modified model increases with the refinement of mesh resolution, highlighting the significant impact of the binning choice. In contrast, the original model remains unaffected by changes in binning, yet it consistently shows significant biases in its posteriors (which I will discuss in more detail in the next section). Here, my goal was merely to conduct

Table 5.1: 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 $H = 1$. I 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

a preliminary assessment of the model's sensitivity to time binning, demonstrating the importance 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 thesis.

5.3 Impact of 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, I 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, the same principle applies to the modified ETAS model, illustrating the degree of bias that can be reduced by including and conditioning on past seismicity. This shows the extent to which the parameters should change under natural variability in simulations. Here, I examine two scenarios: (1) extending the modelling domain to include run-in history prior to a mainshock, and

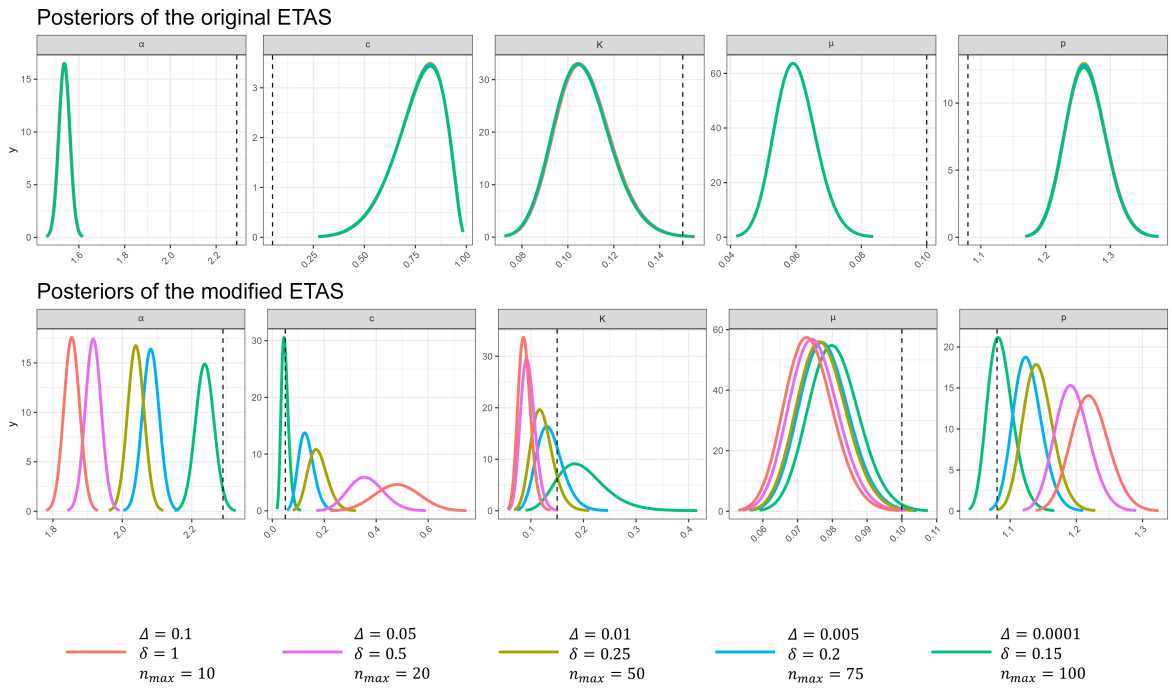

Figure 5.1: Posterior distributions obtained for the original ETAS model (top row) and the modified ETAS model (bottom row) using five temporal meshes with different binning parameters. Vertical dashed lines mark the true ETAS parameters used to generate a synthetic earthquake catalogue for this study. This catalogue covers a 1500-day period, featuring a mainshock of magnitude 6.7 on day 500. I 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 are binned, with the inaccuracies in estimating ETAS parameters significantly reduced as the mesh resolution becomes finer.

(2) conditioning the model on the historical data available before the start of the modelling domain.

Within the modelling domain $[T_1, T_2]$, there are certain events that may not be directly related to 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, I take into account events before T_1 and include their intensities in the prediction of the rate without evaluating themselves in the process.

To analyse both scenarios, I generate a catalogue that spans 1,500 days, with an M6.7 mainshock occurring on day 500.01 and M_0=2.5. Then, I create several subcatalogues by truncating the first 250, 400, 500, and 501 days (Fig. 5.2). Subsequently, I conduct two experiments. In the first experiment, I fit the modified ETAS model to the five subcatalogues 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, I repeat the same procedure as in the first one but additionally 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.

Fig. 5.3 and Table 5.2 display the results of the inversions using my 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. I 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 most aftershock rates - I am still able to retrieve accurate estimates of the ETAS parameters.

My interpretation of the bias trends with respect to the duration of the run-in period before the mainshock, for these synthetics, depends on how well the background rate, μ, is resolved. If μ is poorly estimated, I argue that the triggering model will need to compensate, and hence the parameters in the triggering model will also be biased; in these scenarios I have shown that the background rate tends to be high and the

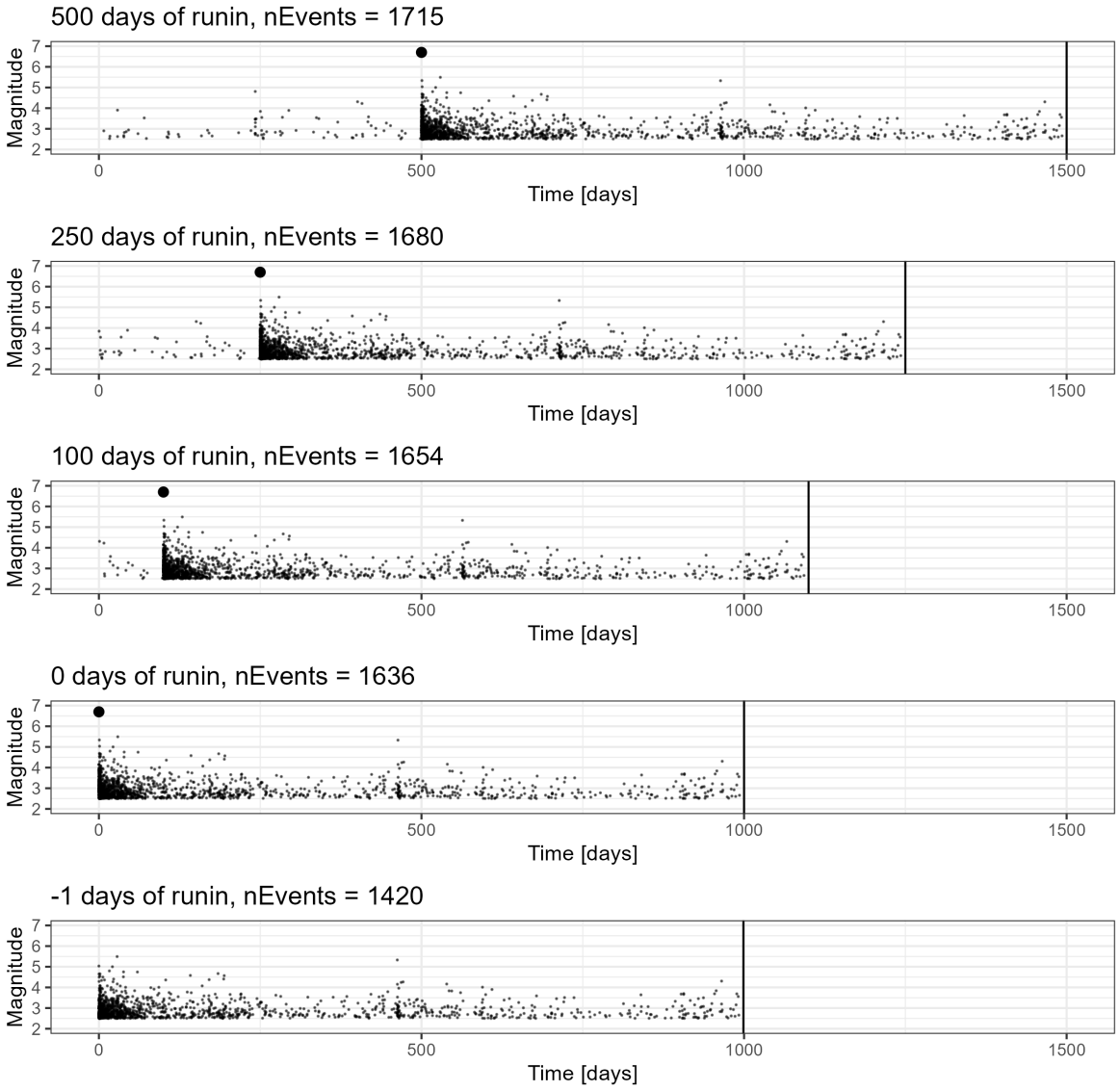

Figure 5.2: Setting different run-in periods for a 1500-day catalogue, by removing the first 0, 250, 400, 500, and 501 days of data.

triggering effects tend to be underestimated, which has implications for forecasting. In my 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 allows μ to have high estimates. As the duration 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 that is otherwise unaccounted for. This

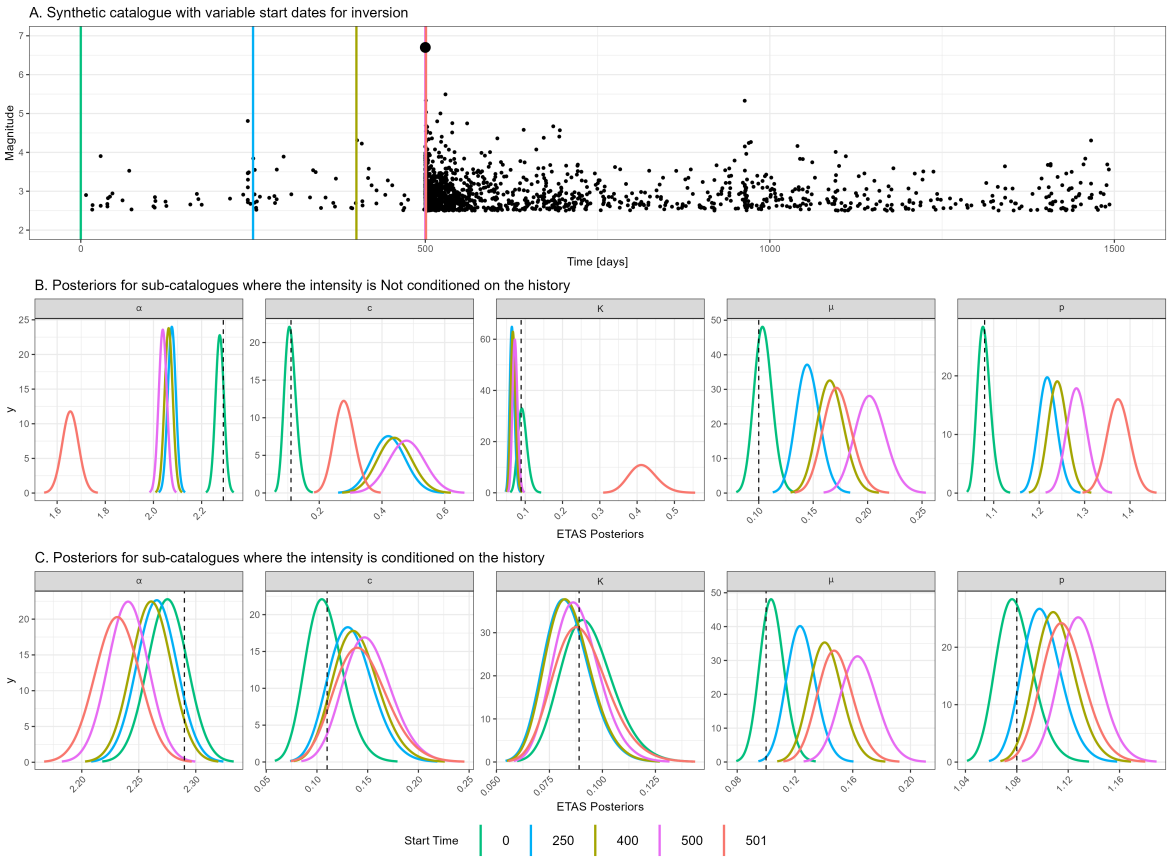

Figure 5.3: (A) A 1500-day catalogue with short-term incompleteness simulated to fit the modified ETAS model, considering "including" and "conditioning on" run-in history. The mainshock with M6.7 was seeded on day 500.01. I extract five sub-catalogues with different starting dates by excluding the first 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.

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, I recommend spending the time to constrain μ well either through an external constraint on the prior or through

Table 5.2: 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

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 trade-offs between the triggering parameters.

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

I investigated this effect using four synthetic earthquake catalogues, each spanning 5000 days (Fig. 5.4 - 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, I intentionally selected catalogues that did not contain other very large events in the sequences so I could isolate the impact of the magnitudes I prescribed.

Although the catalogue with three $M6$ events contains the largest number of events, I hypothesise that the catalogue containing three different $M6, M5, M4$ mainshock magnitudes will have the greatest power to resolve α and therefore do the best job in resolving the trade-off in the productivity parameters. Upon fitting the ETAS model to these data and analysing the posteriors (Fig. 5.4 - bottom), I can infer the following results:

1. 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.

2. Comparing the catalogue that combines different mainshock magnitudes of $M6, M5, M4$ to the catalogues of $3 \times M6$ and $3 \times M5$, I 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 underestimated or overestimated.

Where only a single sequence is studied, we should be aware 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 selected catalogue that contains sufficient number of mainshocks of different magnitude for the catalogue to be truly representative. I 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.

To further validate the findings with a larger sample, I expand the analysis by generating additional synthetic earthquake catalogues. Specifically, I generate ten new synthetic catalogues: five featuring three mainshocks of magnitude 6 ($3 \times M6$) and five with mixed mainshock magnitudes of $M6$, $M5$, and $M4$. This allows for a more comprehensive examination of the parameter estimation's stability and accuracy across different mainshock scenarios. Fig. 5.5, presents the magnitude time series of these ten

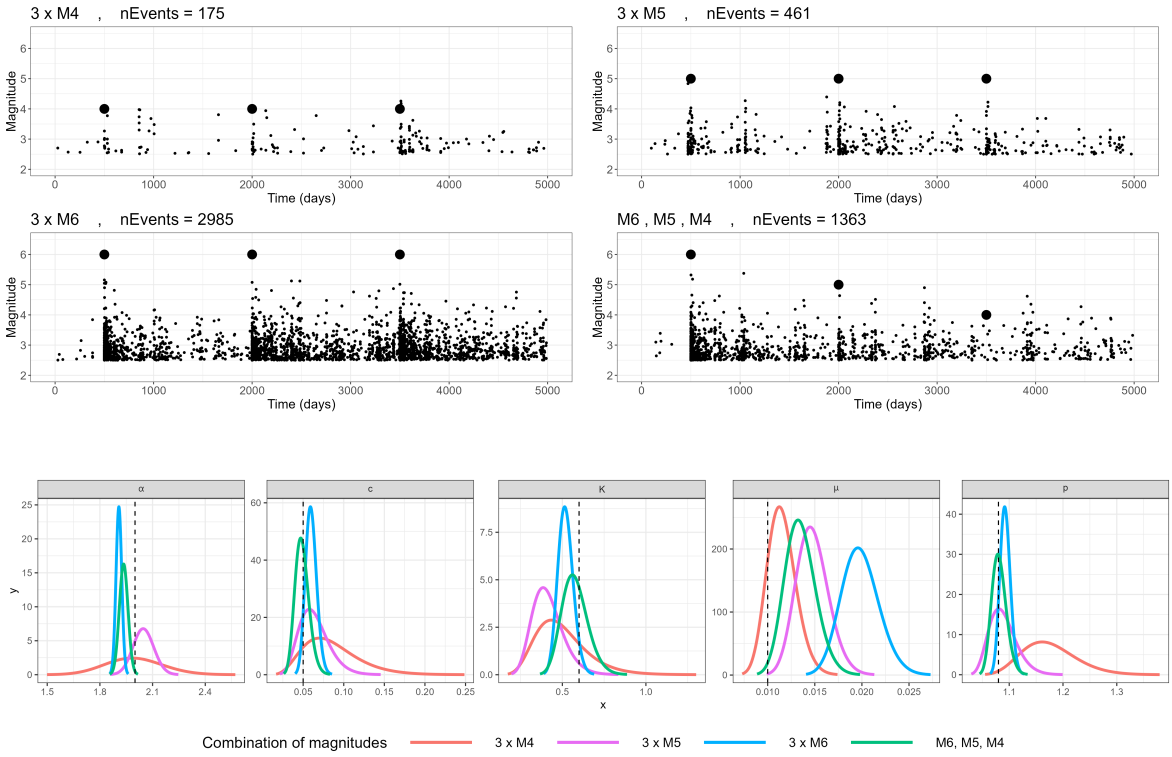

Figure 5.4: Four different simulated catalogues with different combination of magnitudes with $3 \times M4$, $3 \times M5$, $3 \times M6$, and combination of $M6, M5, M4$. Posteriors of each case are shown below the magnitude time series.

simulated catalogues. The left panel illustrates the five $3 \times M6$ catalogues, while the right panel depicts the five mixed-magnitude ($M6$, $M5$, $M4$) catalogues. The number of earthquakes in each catalogue is also mentioned above each plot. Fig. 5.6 shows the posterior distributions for the ETAS model parameters derived from these new catalogues. The posteriors for the $3 \times M6$ catalogues are shown in magenta and the posteriors for the mixed-magnitude catalogues are shown in turquoise.

In addition, the convergence plots for these examples are shown in Figs. 5.7 to 5.16. Interestingly, the different mainshock magnitude models ($M6, M5, M4$) converge much more cleanly and with far fewer iterations. The $3 \times M6$ plots, in all cases, are initially spiky and unstable until they find a stable convergence path, which they then follow. It seems that the initial conditions matter more for the $3 \times M6$ models than for the $M6, M5, M4$ models. This result again demonstrates that it is preferable to have a less biased sample with fewer events than more events in a biased sample, at

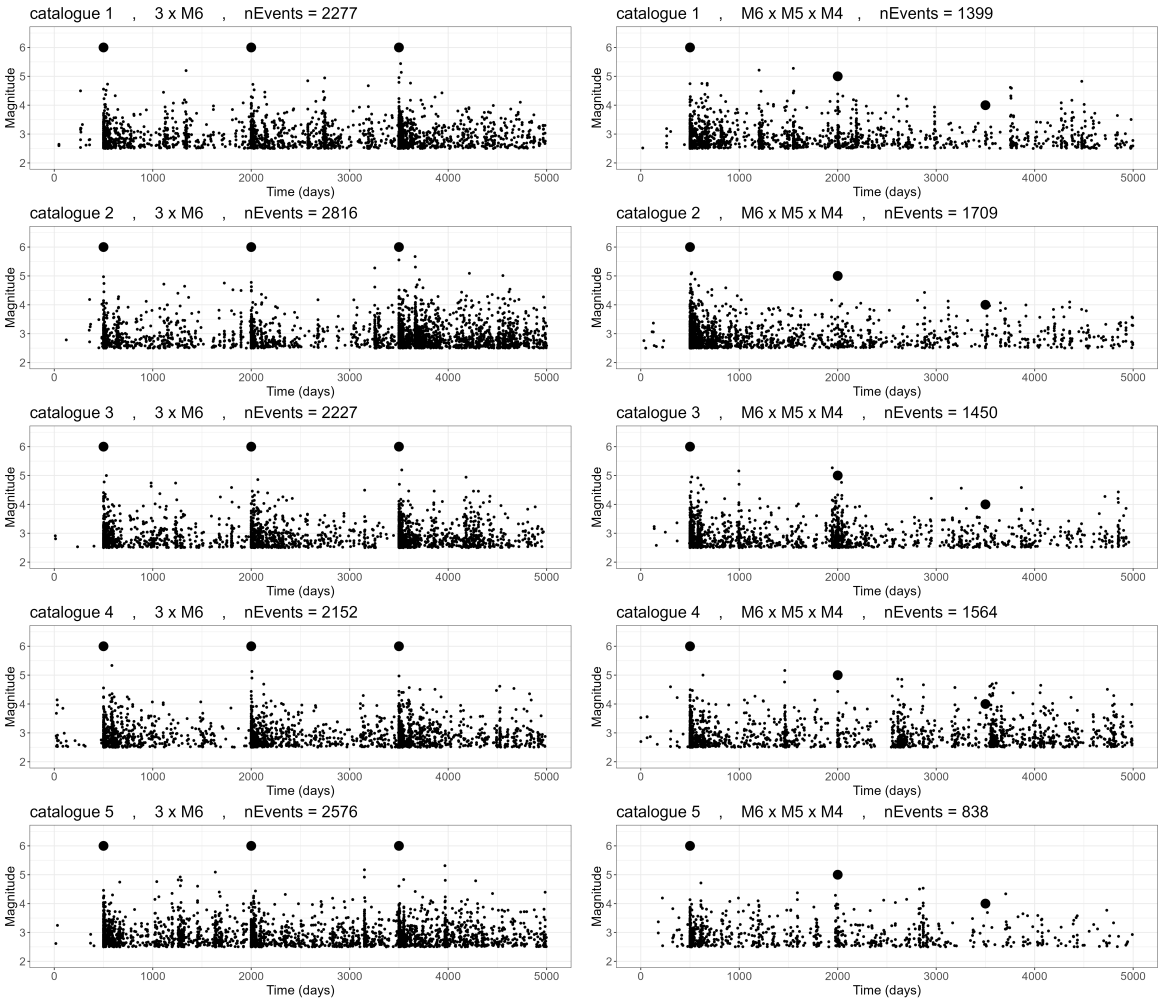

Figure 5.5: Magnitude plot of ten simulated catalogues with different combination of magnitudes with $3 \times M6$ (left panel) and combination of $M6, M5, M4$ (right panel).

least for the comparison of $3 \times M6$ versus the combination of various magnitudes. The former is less likely, and hence the latter more representative. This extended analysis corroborates the findings, reinforcing the hypothesis that diversity in the magnitudes of the mainshock significantly improves the resolution of the $K - \alpha$ trade-off.

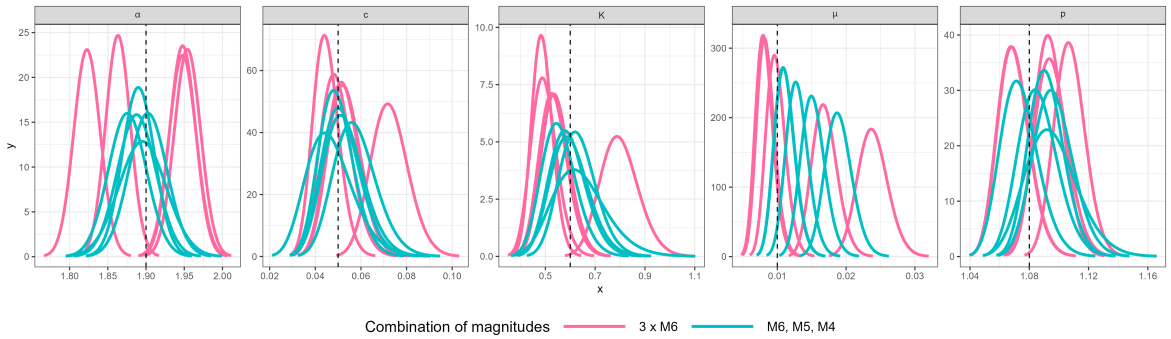

Figure 5.6: Posteriors of ten simulated catalogues with different combination of magnitudes with $3 \times M6$ (in magenta) and combination of $M6, M5, M4$ (in turquoise).

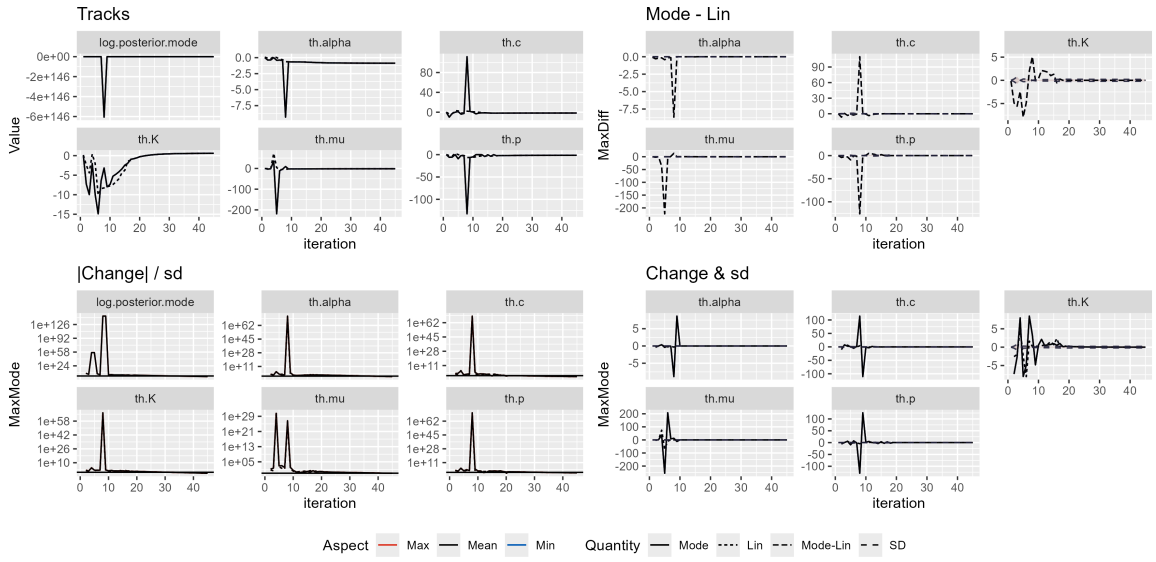

Figure 5.7: Convergence plots for the model from simulated catalogue 1 with $3 \times M6$.

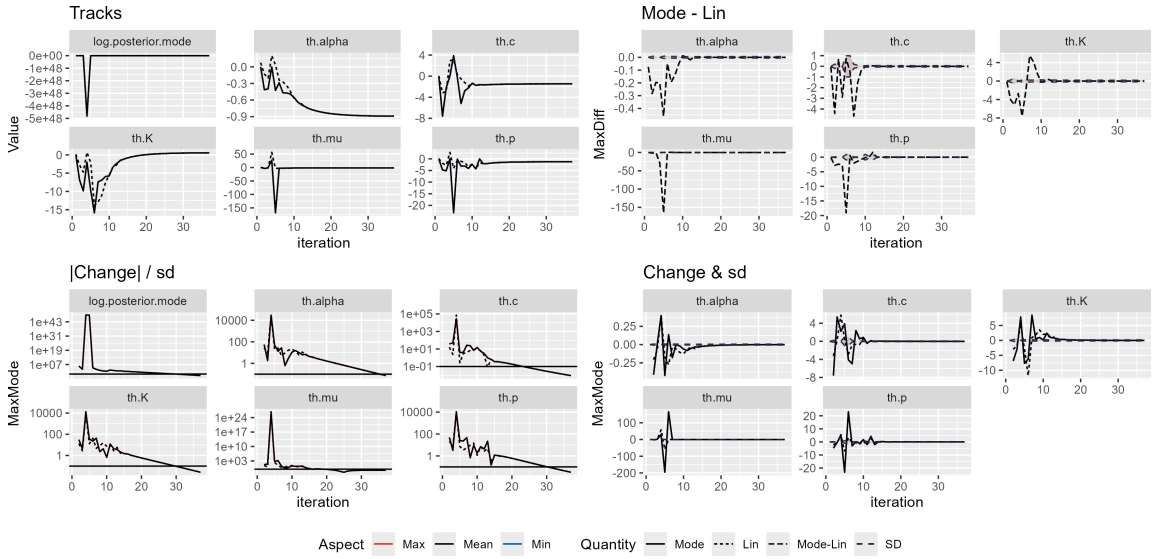

Figure 5.8: Convergence plots for the model from simulated catalogue 2 with $3 \times M6$.

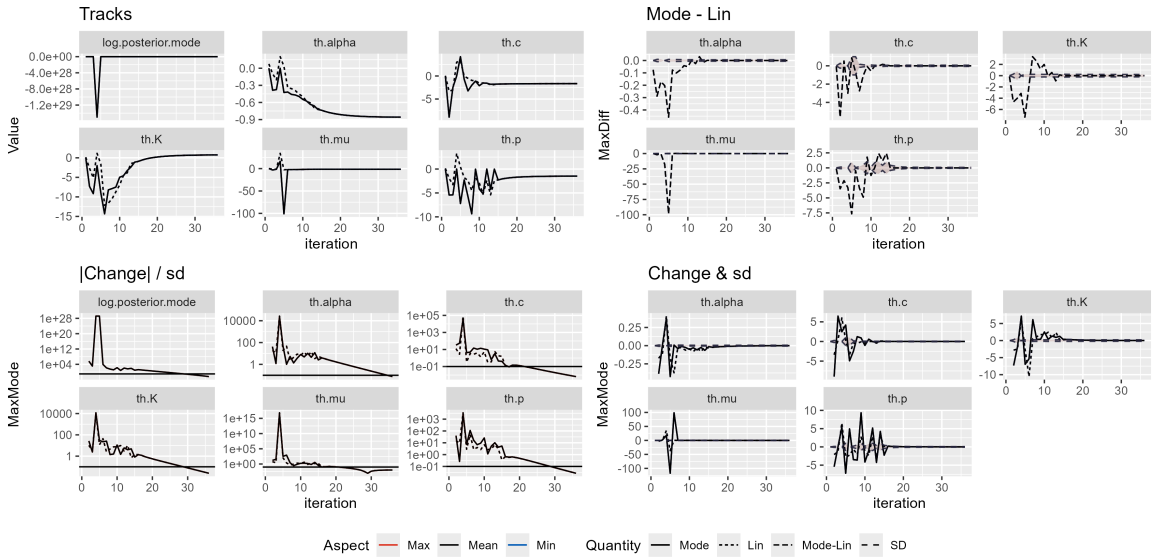

Figure 5.9: Convergence plots for the model from simulated catalogue 3 with $3 \times M6$.

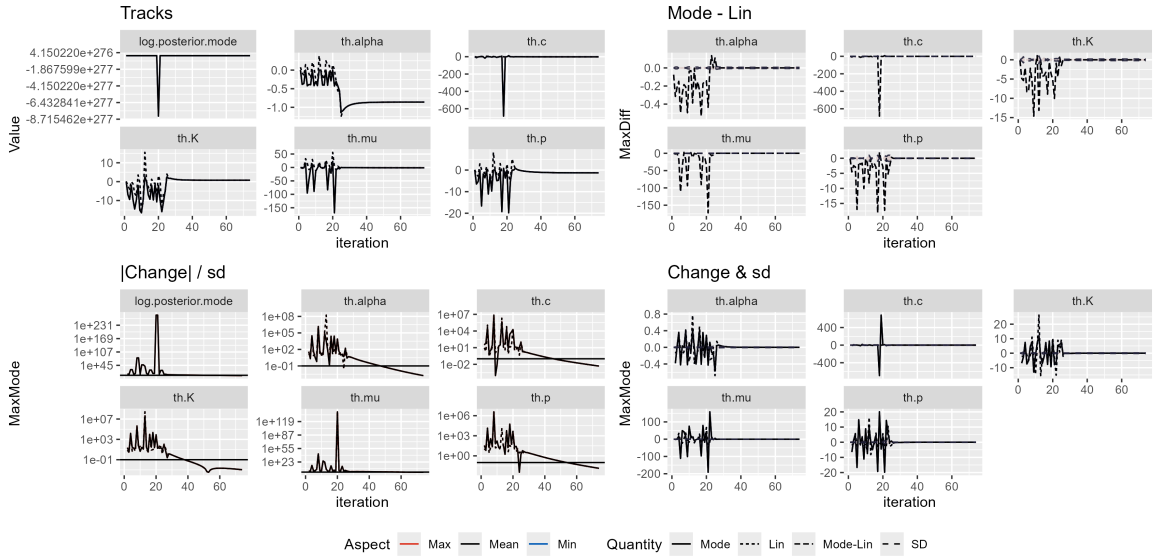

Figure 5.10: Convergence plots for the model from simulated catalogue 4 with $3 \times M6$.

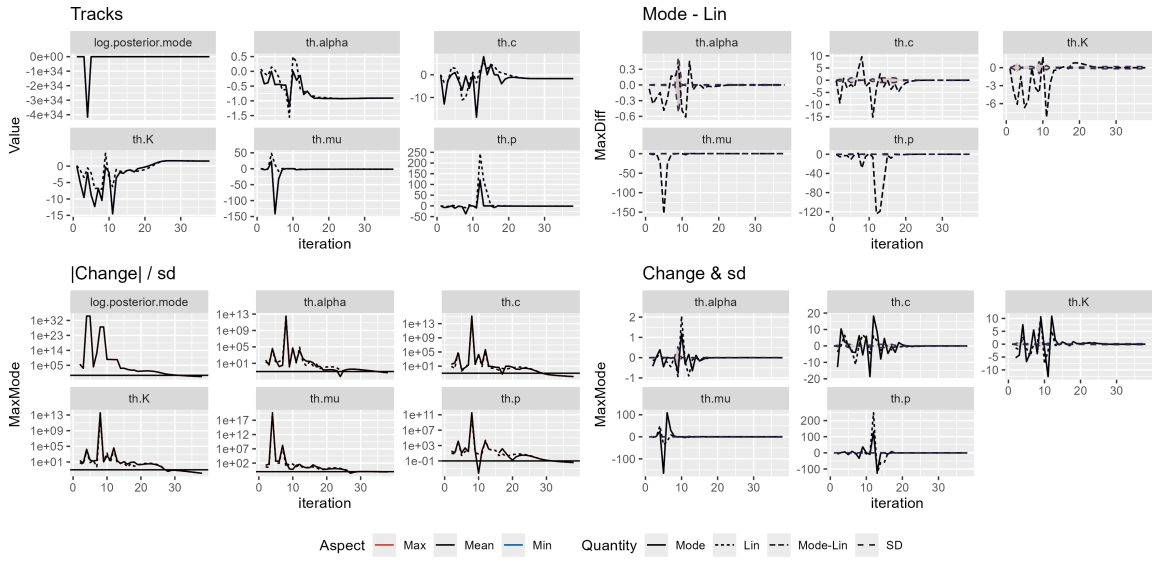

Figure 5.11: Convergence plots for the model from simulated catalogue 5 with $3 \times M6$.

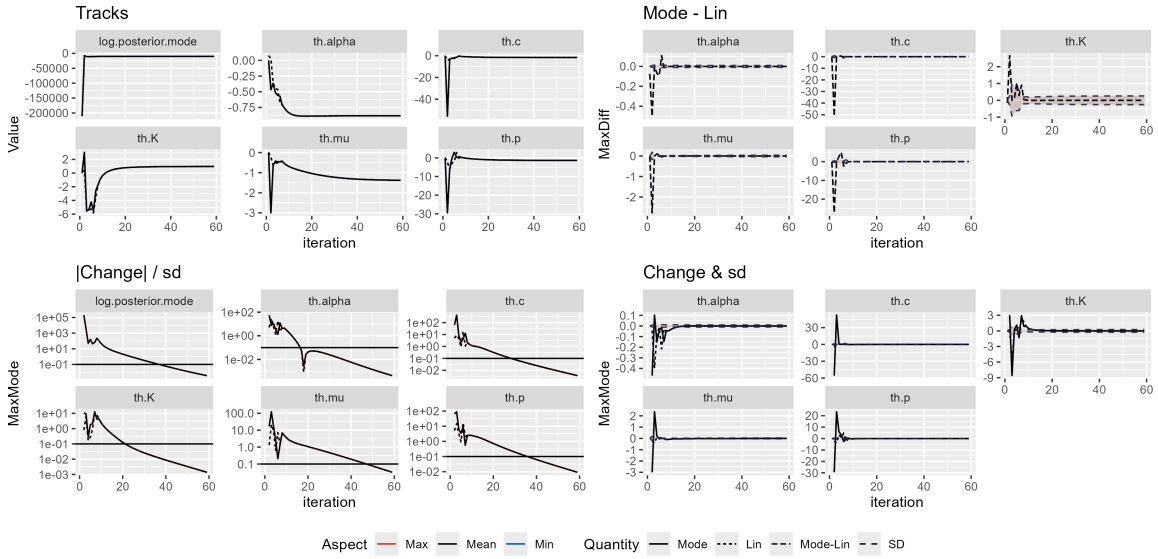

Figure 5.12: Convergence plots for the model from simulated catalogue 1 with $M6, M5, M4$.

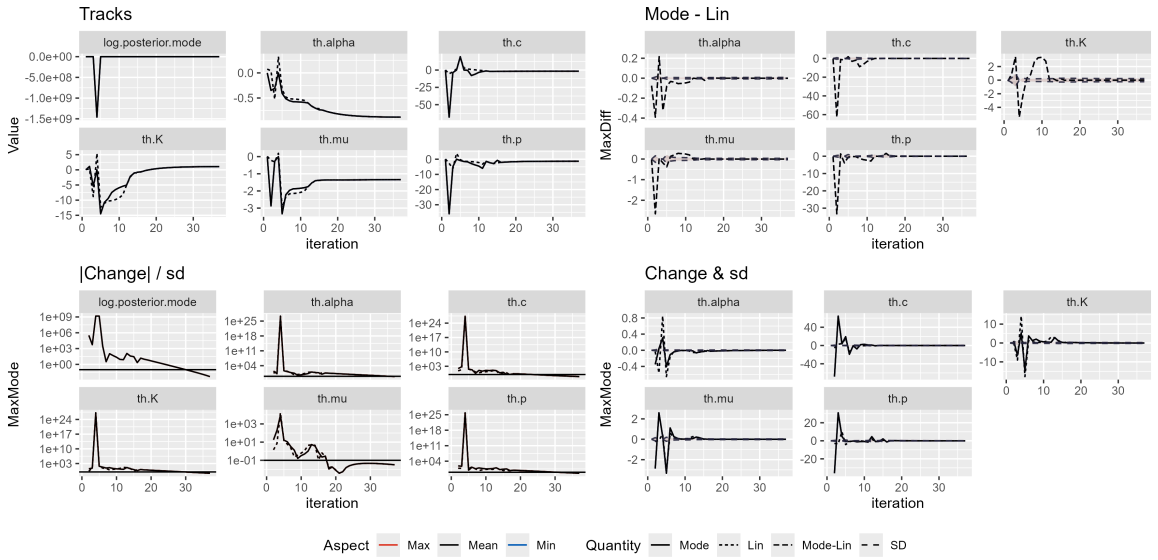

Figure 5.13: Convergence plots for the model from simulated catalogue 2 with $M6, M5, M4$.

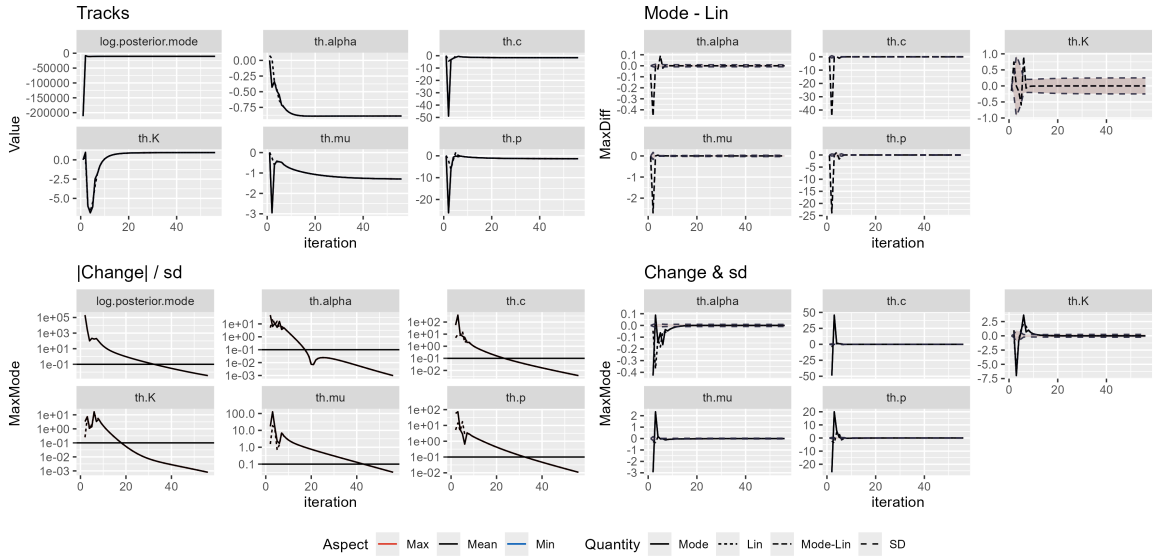

Figure 5.14: Convergence plots for the model from simulated catalogue 3 with $M6, M5, M4$.

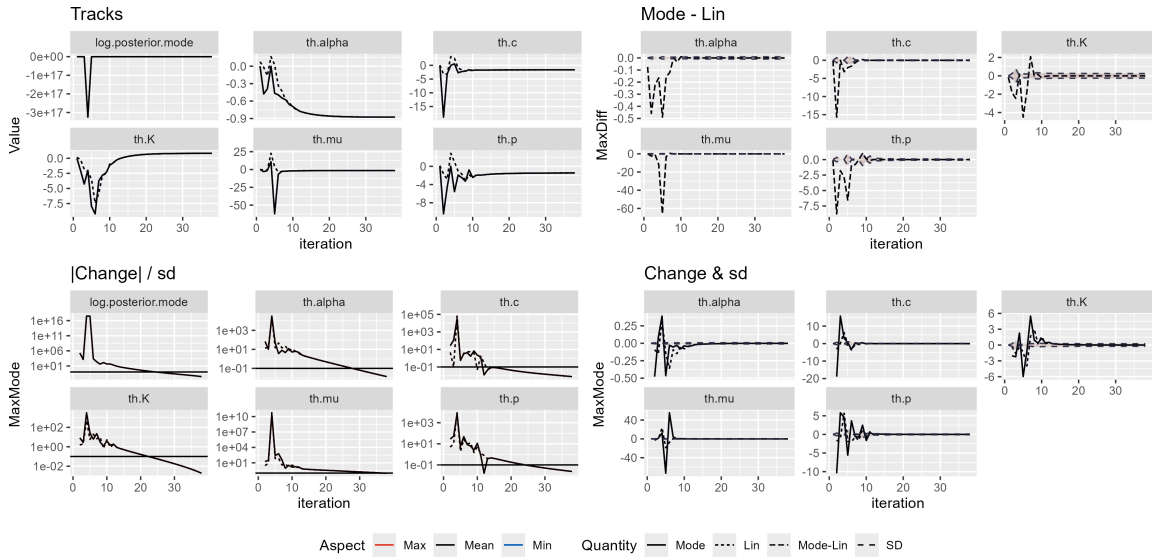

Figure 5.15: Convergence plots for the model from simulated catalogue 4 with $M6, M5, M4$.

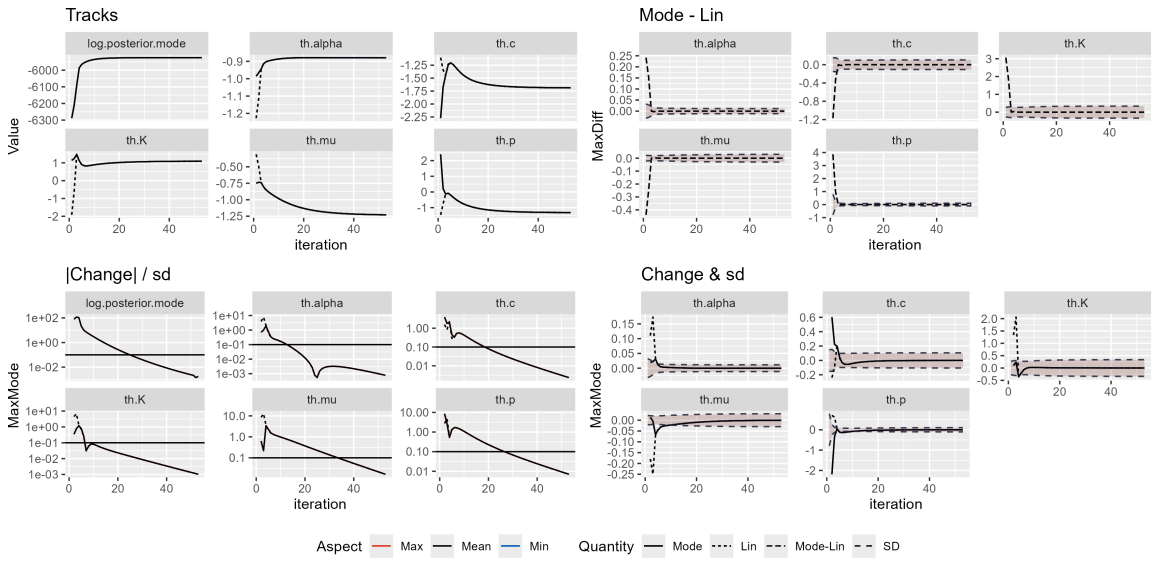

Figure 5.16: Convergence plots for the model from simulated catalogue 5 with $M6, M5, M4$.

5.5 Impact of choice of incompleteness model parameters

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

I investigate how varying the parameters G and H of the incompleteness model influences the posteriors of the modified ETAS model. To accomplish this, I generate a catalogue that spans 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. Then, I consider five distinct scenarios in which I 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 ($G = 3.8$, $H = 2$). I deliberately select these parameter combinations to cover a variety of choices in the incompleteness model. This results in different decay patterns and different endpoints for the incompleteness period, as illustrated in Fig. 5.17 (top). Then, I fit my modified ETAS model to the incomplete catalogue considering the different incompleteness parameter sets. I also run the original ETAS model with the true incompleteness parameters.

The posteriors for all scenarios are shown in Fig. 5.17 (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 as a result of 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 the estimates. Examining the purple ('inferred 2') and salmon ('inferred 3') scenarios, which significantly differ in incompleteness curves, shows bias in the ETAS posteriors. Interestingly, all five scenarios with the modified ETAS models, including

the extreme scenarios (purple and salmon), still outperform the original ETAS model (dashed dark grey posteriors), calibrated with true incompleteness parameters.

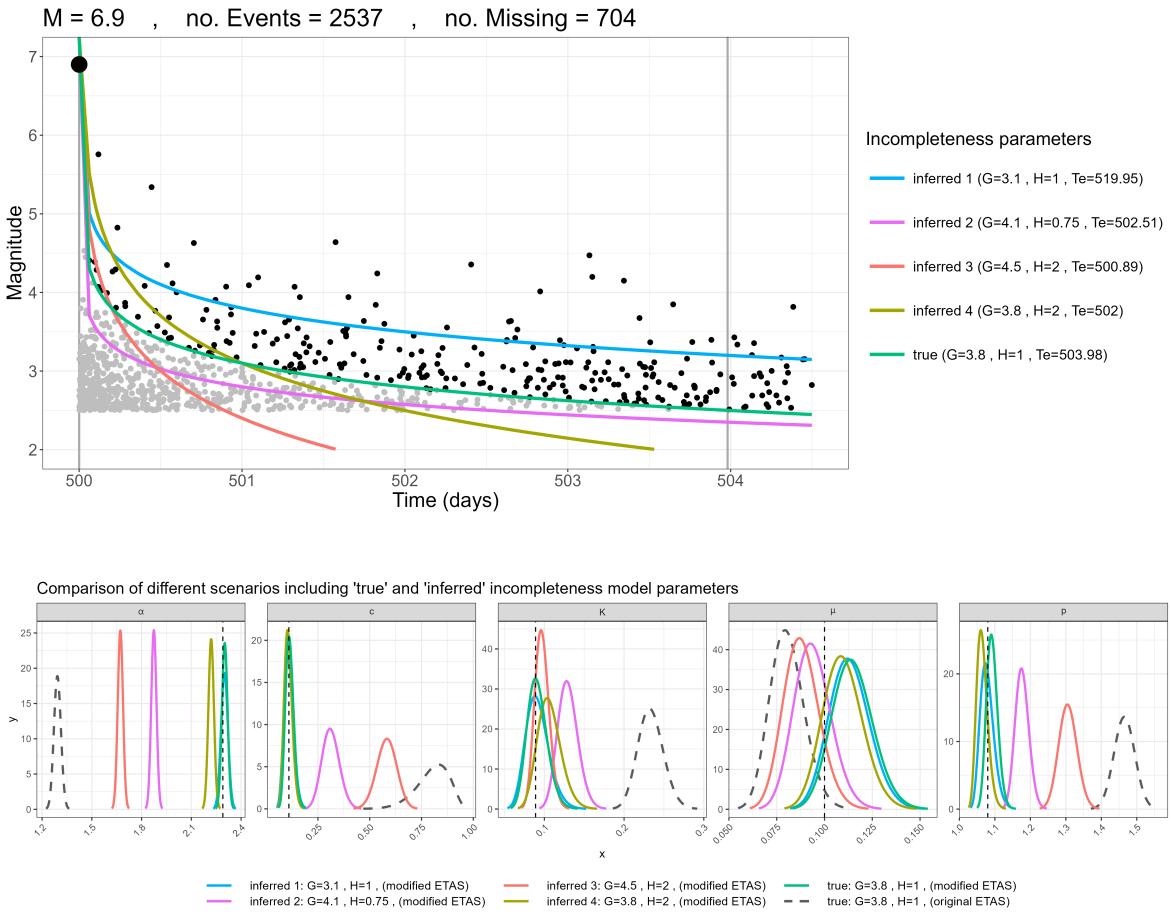

Figure 5.17: Impact of the choice of incompleteness model parameters on the ETAS estimates. (top): five incompleteness models with different decay curves and endpoints. The green curve represents the true incompleteness model, utilised for data truncation and preparation of the incomplete catalogue. 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.

5.6 Impact of presence of a secondary large aftershock

A special case in which there are mainshocks of different magnitudes is where a second large event occurs within an aftershock sequence. Real earthquake data reveal instances of sequences where either a significant aftershock with magnitude close to the mainshock occurs or twin/multiple large events take place. Recent examples include the 2012 M_w6.4 and M_w6.2 Ahar-Varzeghan doublet earthquakes in Iran; the 2017 M_w6.0 triplets in Hojedk, Iran; the 2019 M_w6.4 and M_w7.1 Ridgecrest sequence in the US; the 2019 quadruplet of M6.4, 6.6, 6.5, and 6.8 in Cotabato, Philippines; the 2021 M_w6.2 and M_w6.3 Hormozgan doublet earthquakes in Iran; the 2022 M_w6.0 and M_w6.0 Hormozgan doublet earthquakes in Iran; the 2023 M_w7.8 and M_w7.5 Kahramanmaras earthquakes in Turkey–Syria, and the recent quadruplet of M6.2-6.3 in 2023 in Herat, Afghanistan.

I examine this scenario using a 1500-day long synthetic catalogue with a mainshock magnitude of M6.7 occurring on day 500, followed by a large aftershock with magnitude of M5.9 on day 570 well within the triggered aftershock sequence. The sequence comprises a total of 1921 events. The data were then truncated using the incompleteness parameters $G = 4.5$ and $H = 0.75$, resulting in the removal of 164 events, leaving 1757 events in the dataset. Fig. 5.18 shows the time series and histogram of the simulated sequence. It is important to note that the incompleteness was only applied to the mainshock event (M6.7), and I deliberately chose not to apply it to the second large event (M5.9), based on the assumption that most of the rates are more influenced by the mainshock.

Then, I fitted both the original and modified models to the incomplete data, and the results are displayed in Fig. 5.18. Upon comparing the results, it becomes evident that the posteriors of the modified model generally offer improved estimates of the ETAS parameters compared to the original model, resulting in less bias. This indicates that even in the specific case of secondary large aftershocks or doublet/multiple events in real experiments (as exemplified above), accurate estimates of the ETAS parameters can still be obtained. However, I suggest applying the incompleteness model to the second large event to investigate its impact on bias reduction, though this falls beyond the scope of this thesis.

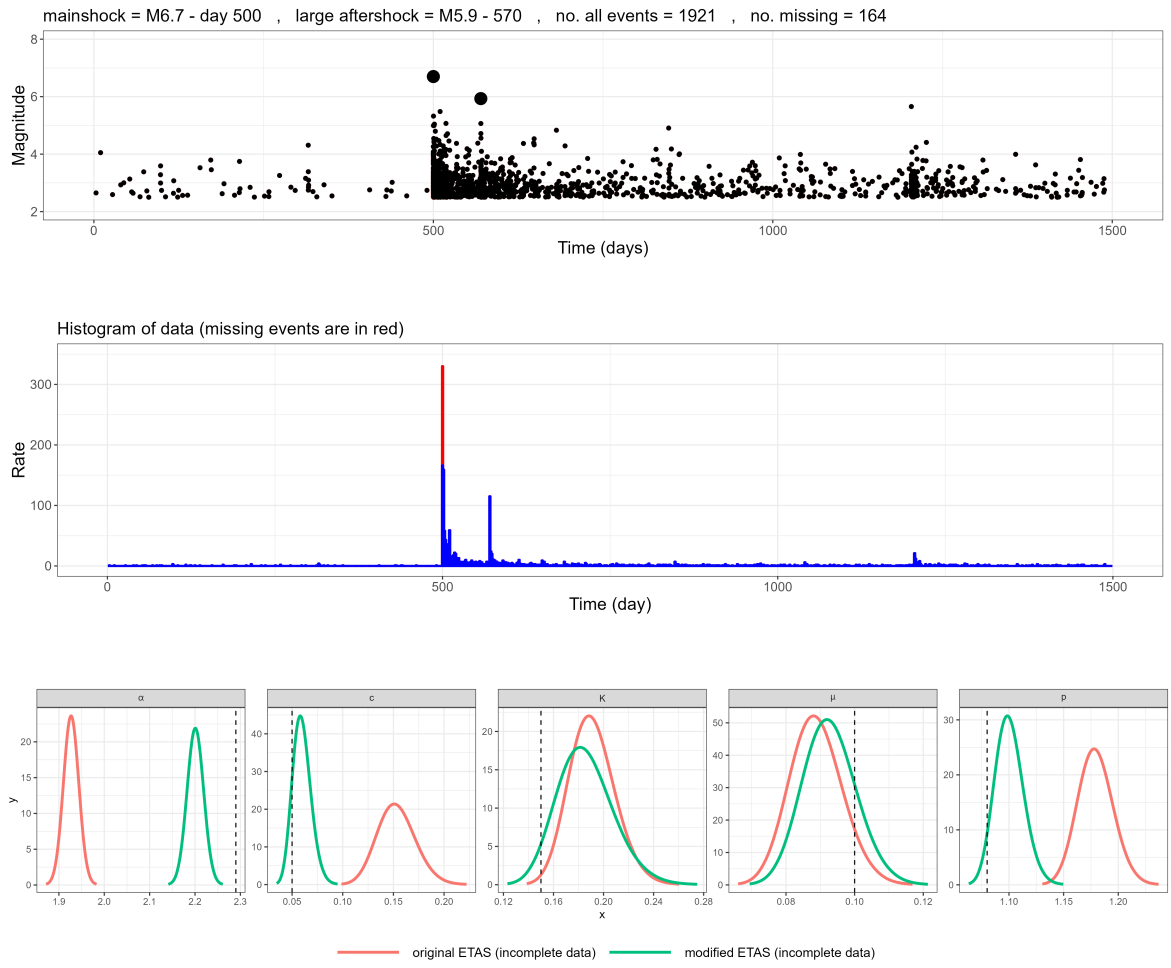

Figure 5.18: (top): magnitude-time plot of an earthquake sequence with a mainshock ($M6.7$) and a secondary large aftershock ($M5.9$) following the mainshock. (middle): Histogram of the events. The red bar shows the portion of data that was eliminated as the unobserved data in the incompleteness period. (bottom): posteriors of ETAS parameters using the modified (in green) and original ETAS (in red) versions, respectively.

5.7 Application to real earthquake case studies

Here, I 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 United States. To minimise bias and select a representative sample, I chose data 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 I analysed in Section 4.3. In addition, I also conduct my 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 that I analysed in Section 5.3. The spatial domain for each case was also determined based on published shake maps and seismicity maps.

The time series of magnitudes and the magnitude-event number plots for each earthquake are shown in Fig. 5.19. 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 of different magnitudes. Despite considering three years of data after the M.7.3 mainshock, the sequence is still active at the end, so I 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 the $K-\alpha$ tradeoff (as discussed in Section 5.4).

To find the incompleteness model parameters for each sequence, I plot zoomed-in

magnitude time plots that clearly reveal the incompleteness in the early aftershock period. I 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. 5.20. In addition, the b-value of each sequence is estimated using the stability of b-values plotted against the completeness threshold, as shown in Fig. 5.21. The figure demonstrates how the b-value fluctuates over different magnitude thresholds. To determine the final b-value, I focus on the more stable and horizontal portion of the diagram, which typically occurs between the middle magnitude ranges (2.6-4.4). At both the lower and upper ends of the diagram (corresponding to smaller and larger magnitudes), fluctuations are more pronounced due to fewer available data, making these regions less reliable for estimating the b-value. Therefore, the chosen b-value represents the average calculated in the middle stable section of the diagram. The average b-values calculated for different earthquake sequences are as follows: 0.85 for the 2016 Amatrice, Italy event, 0.77 for the 2017 Kermanshah, Iran event, and 0.82 for the 2019 Ridgecrest, California event.

Subsequently, I apply both the original and the modified ETAS models to the earthquake catalogues. In Fig. 5.22, I present the posterior outcomes of both the modified (in green) and the original (in red) ETAS models. These models are fitted to 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. Furthermore, I extracted the triggering functions for both models, as illustrated in Fig. 5.23, which indicates higher rates and narrower plots (less uncertainty) for the modified model compared to the original version.

Since I do not have access to the true parameters for the real datasets, I can assess the performance of both original and modified models based on two key aspects: (1) visual inspection: This involves examining whether models fitted to real data produce patterns of underestimation and overestimation similar to those observed in synthetic

experiments. Here, I 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, I use the Widely Applicable Information Criterion (WAIC), also known as the 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 5.3 displays the WAIC values for the three real sequences using both the original ETAS and modified ETAS models. The lower WAIC values of the modified model indicate better performance compared to those of the original model. Therefore, I can conclude that the modified ETAS model provides 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, although this falls outside the scope of the current study.

In all cases, correcting for short-term incompleteness changes the estimates of magnitude-dependent productivity α. The background rate is consistent in all case studies, presumably because of the adequate sampling of the pre-mainshock period. Ridgecrest data have the greatest short-term incompleteness, and this propagates to significant reductions in c and p; this is consistent with the changes seen in synthetics in Fig. 4.7, so I believe the corrected estimates are more reliable.

Table 5.3: WAIC score obtained from the original and the modified ETAS models fitted to 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	13980	12695	10593
original	13998	12698	10660

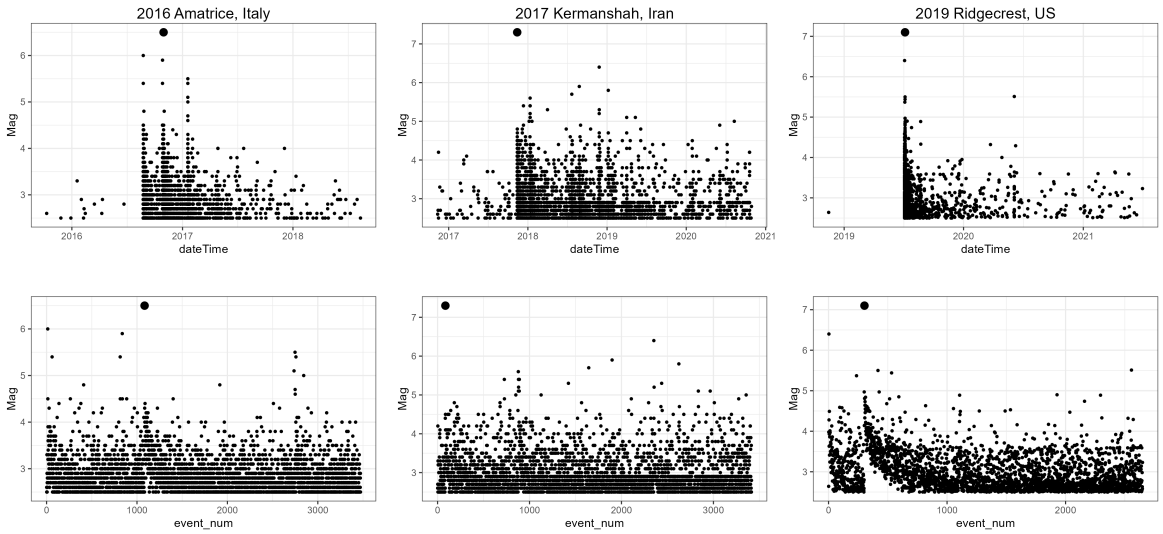

Figure 5.19: The three selected real earthquake sequences, including the 2016 Amatrice, Italy; the 2017 Kermanshah, Iran; and the 2019 Ridgecrest, US earthquakes. (top): magnitude - time plots and (bottom): magnitude - event number plots.

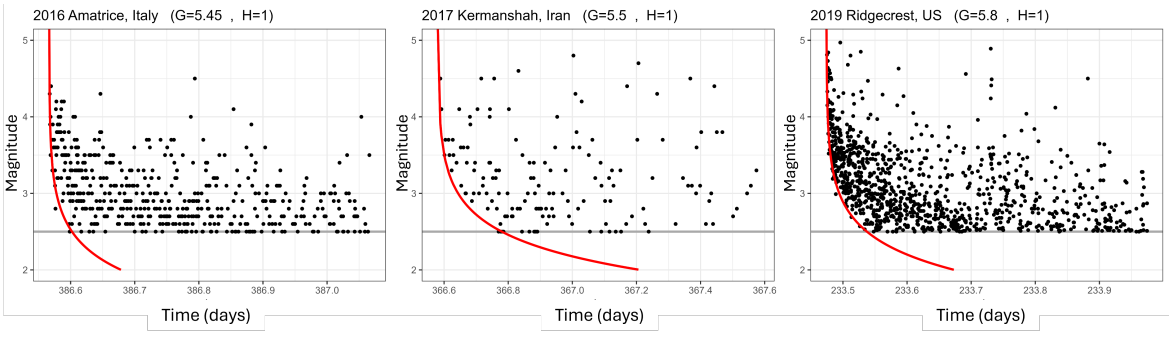

Figure 5.20: Extracting parameters G and H with the best fit of the incompleteness model to the three selected real earthquake sequences, including the 2016 Amatrice, Italy; the 2017 Kermanshah, Iran; and the 2019 Ridgecrest, US earthquakes. The red curves show the best exponential fit in order to derive the incompleteness parameters.

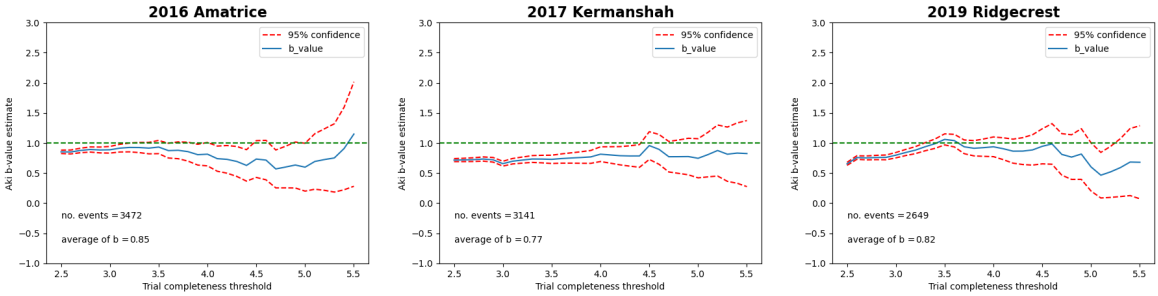

Figure 5.21: b-value estimate as a function of the trial completeness threshold for the 2016 Amatrice (Italy), 2017 Kermanshah (Iran), and 2019 Ridgecrest (US) sequences. The x-axis represents the trial completeness threshold, which is the minimum magnitude above which the earthquake catalogue is considered complete. The y-axis shows the Aki b-value estimate, indicating the relative frequency of smaller to larger earthquakes. The solid blue line shows the b-value estimate, while the dashed red lines represent the 95% confidence intervals. The dashed green line marks $b = 1$, providing a reference to assess how much the b-value deviates from this benchmark.

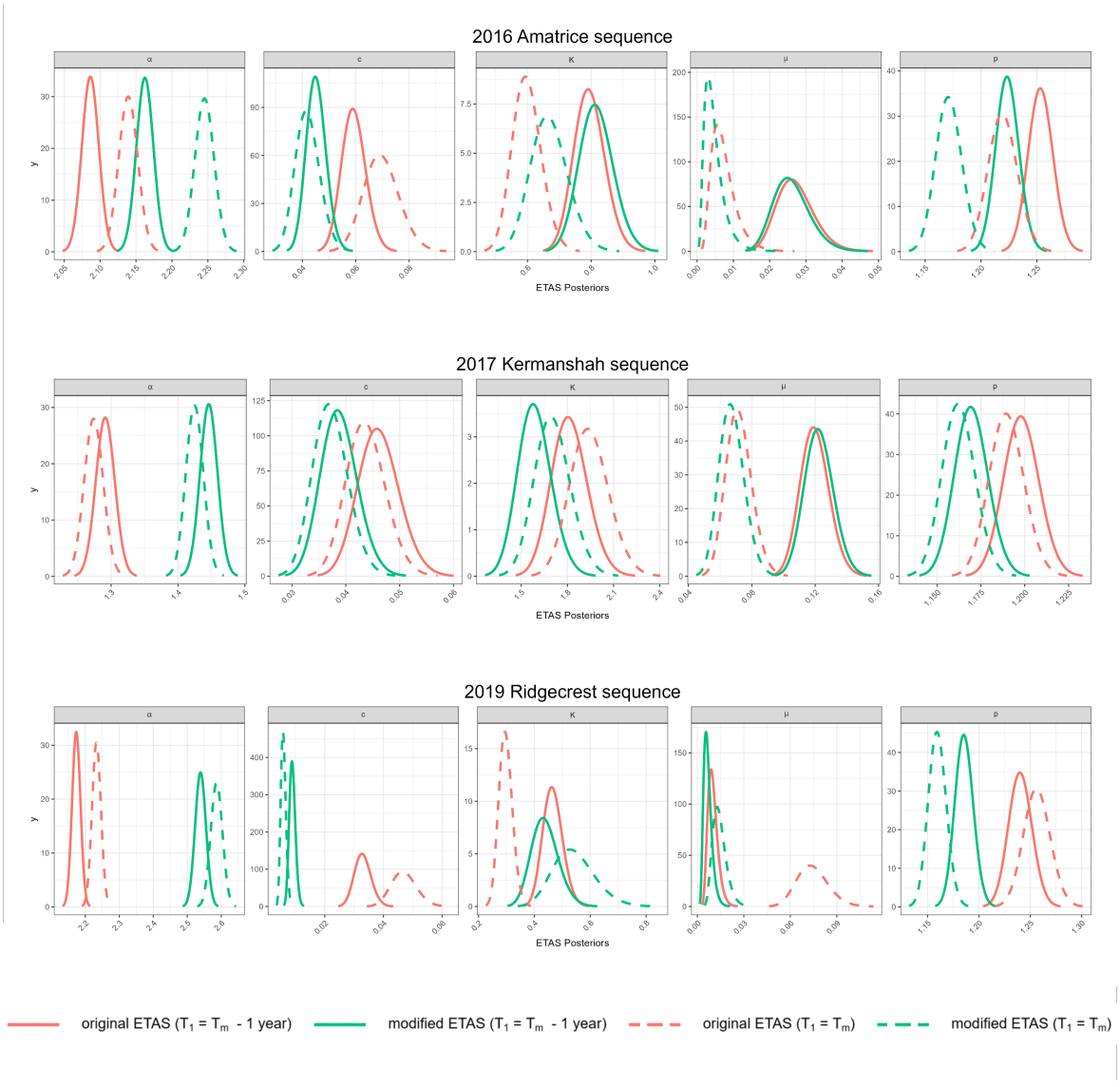

Figure 5.22: Posteriors of the three selected real earthquake sequences, including the 2016 Amatrice, Italy; the 2017 Kermanshah, Iran; and the 2019 Ridgecrest, US earthquakes. These posteriors are obtained using the modified (green) and the original (red) ETAS models, respectively. The models are fitted to 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).

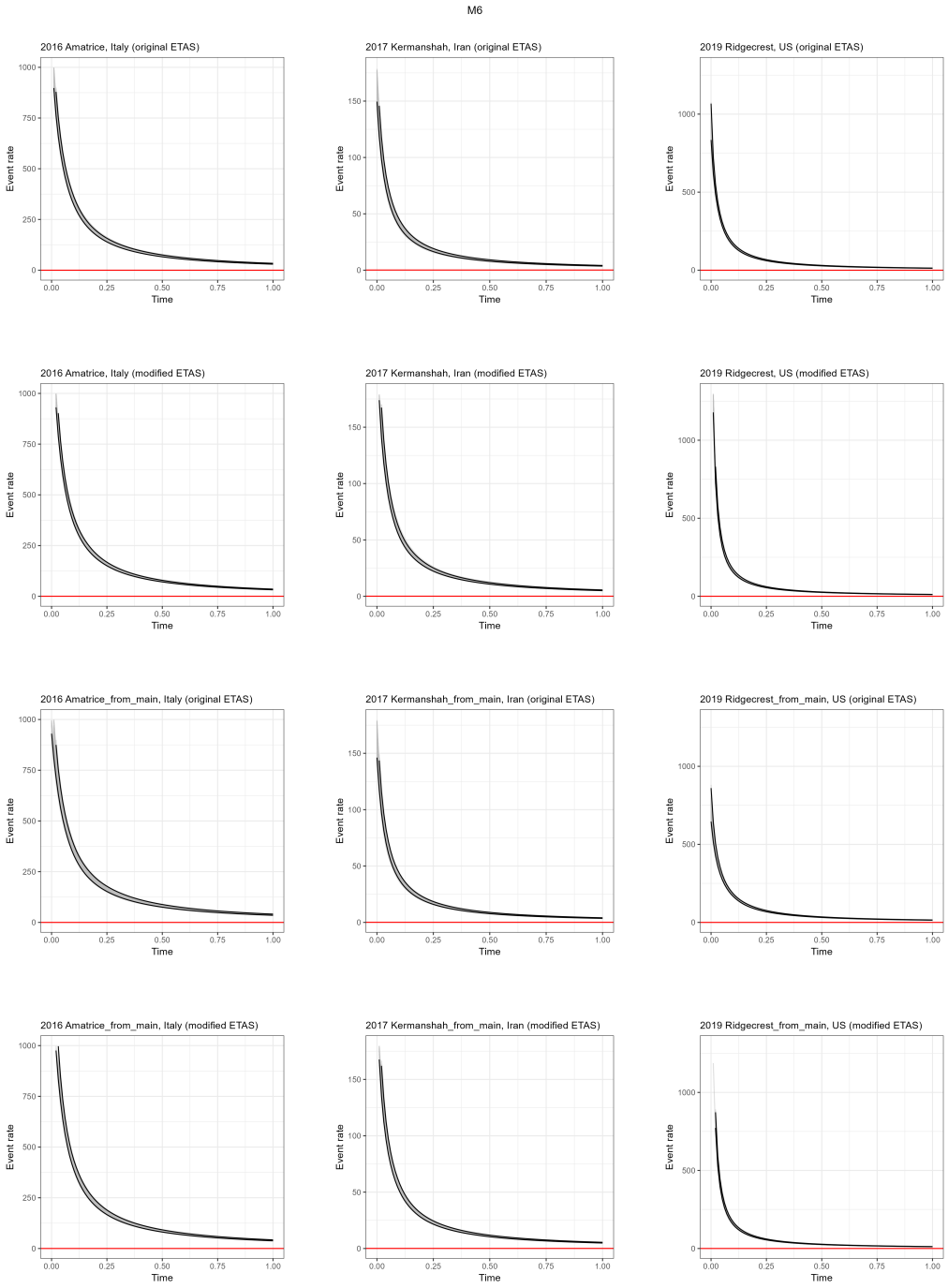

Figure 5.23: Triggering functions of the three selected real earthquake sequences, including the 2016 Amatrice, Italy (column 1), the 2017 Kermanshah, Iran (column 2), and the 2019 Ridgecrest, US (column 3) earthquakes. The figure is divided into four rows: the first row presents results from the original ETAS model, while the second row shows the modified ETAS model, both derived with 1-year data prior to the mainshock event included in the analysis. The third and fourth rows repeat the same comparison, but the modelling domain starts from the mainshock (indicated by 'from main'). In each plot, the solid black line represents the event rate over time and the grey shaded area shows the 95% confidence interval.

5.8 Impact of length of temporal modelling domain

Another consideration that we should take into account to accurately retrieve the ETAS model parameters is understanding the impact of the length of the temporal modelling domain on the ETAS inversions. In this regard, I select and analyse three major earthquake sequences that occurred in Central Italy, including the 1997 Colfiorito (Mw = 6.0), 2009 L'Aquila (Mw = 6.3) and 2016 Amatrice (Mw = 6.2) events (Fig. 5.24). I use data from the HOmogenised instRUmental Seismic Catalogue (HORUS) of Italy (Lolli et al. 2020). For each sequence, data were taken from one year before the mainshock up to the end of the Gardner-Knopoff (Gardner and Knopoff 1974) temporal windows, based on each mainshock's magnitude. This corresponds to 510, 678, and 622 days for the 1997 Colfiorito, 2009 L'Aquila , and 2016 Amatrice mainshocks, respectively. The magnitude time series for these sequences are also shown in Fig. 5.25.

I then consider seven temporal windows post-mainshock to observe how the ETAS parameters evolve: 1-day, 1-week, 2-weeks, 1-month, 3-months, 6-months, and 1-year (Fig. 5.26). Such temporal windows are commonly used by seismologists to assess aftershock sequences (Reasenberg and Jones 1989; Wiemer and Katsumata 1999). My aim here is to determine whether the ETAS parameters exhibit fluctuating or stable behaviours as the length of the temporal window varies. When considering different periods after the mainshock, the one-year period before the mainshock is still included in the time domain to take into account the run-in history as discussed in Section 5.3. I also consider the spatial window radius of the Gardner and Knopoff (1974) method to select the data for each sequence.

The results are presented in Figs. 5.27 and 5.28. Fig. 5.27 shows the posterior distributions, while Fig. 5.28 illustrates the temporal variations of ETAS parameters for each scenario. In the latter, the solid lines represent the mode of the posterior distributions for each parameter, and the shaded areas indicate the uncertainty bands using quantiles. In all scenarios, the background seismicity rate (μ) remained relatively stable in all scenarios, indicating that it is less sensitive to the length of the aftershock sequence used and is better resolved by the run-in period before the mainshocks. However, the parameters of the triggering function (α, K, c, and p) show significant changes in the shorter windows but tend to stabilise as the window length

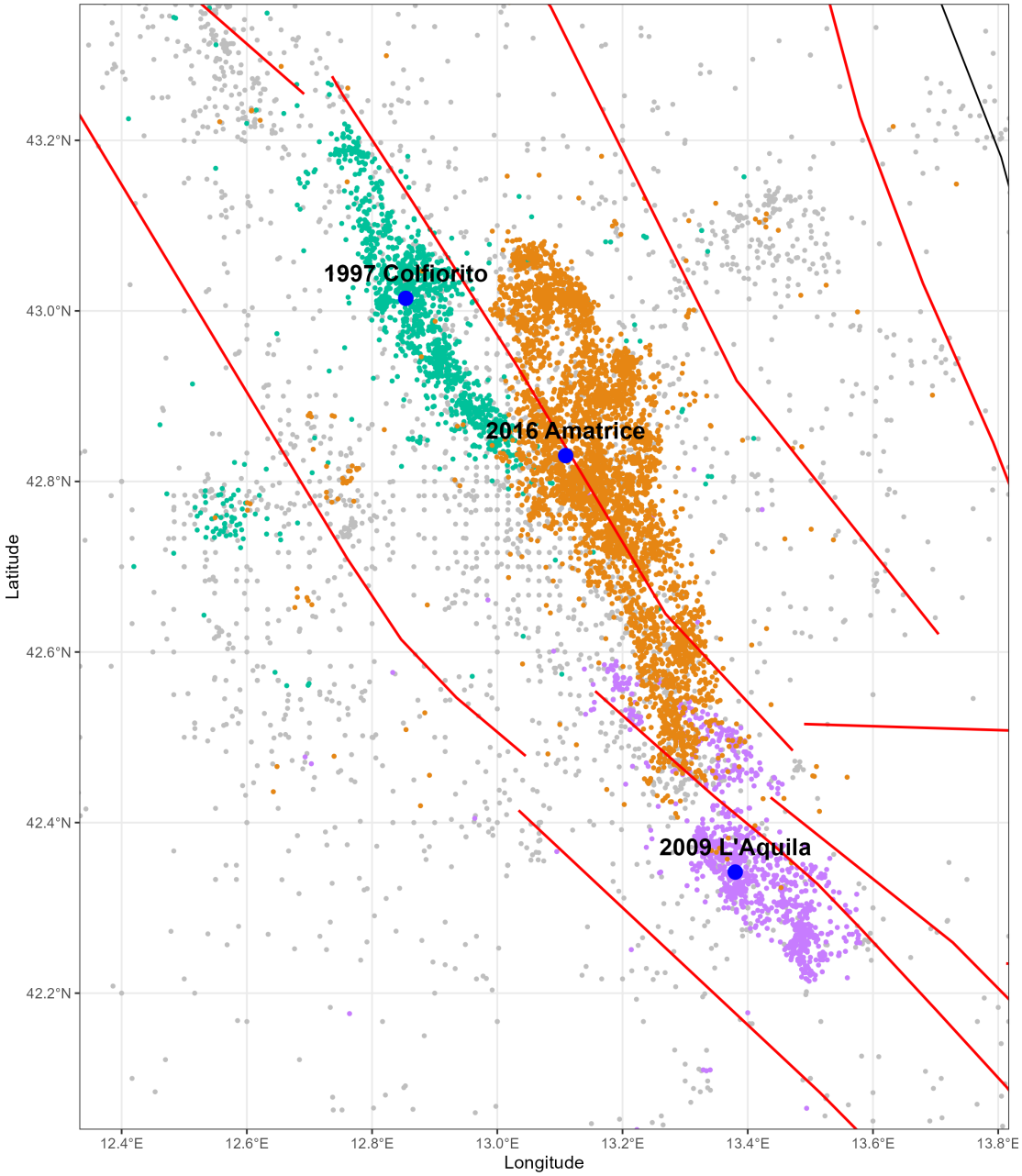

Figure 5.24: The distribution of the 1997 Colfiorito (Mw = 6.0), 2009 L'Aquila (Mw = 6.3), and 2016 Amatrice (Mw = 6.2) earthquake sequences in Central Italy. The coloured dots represent seismic events from different sequences: green for the 1997 Colfiorito sequence, purple for the 2009 L'Aquila sequence, and orange for the 2016 Amatrice sequence. Grey dots represent background seismicity unrelated to these main sequences. The red lines indicate mapped fault lines in the region. The blue circles mark the epicentres of the respective mainshocks for each sequence.

Figure 5.25: The magnitude time series for the 1997 Colfiorito (Mw = 6.0), 2009 L'Aquila (Mw = 6.3) and 2016 Amatrice (Mw = 6.2) earthquake sequences in Central Italy. The data for each sequence was taken from one year before the mainshock up to the end of the Gardner-Knopoff (Gardner and Knopoff (1974)) temporal windows, based on each mainshock's magnitude.

increases. In particular, after approximately six months, the parameters became more constant, suggesting that longer records of aftershock sequences are necessary for stable parameter estimation. This is particularly important for both 'retrospective' and 'prospective' earthquake forecasting. In prospective analysis, where the full sequence is not available, longer-term records might be used, which I will address later in Chapter 6. In summary, the findings indicate the importance of careful consideration in selecting the temporal modelling domain for ETAS analysis, emphasising the balance between data availability and the temporal stability of parameter estimates.

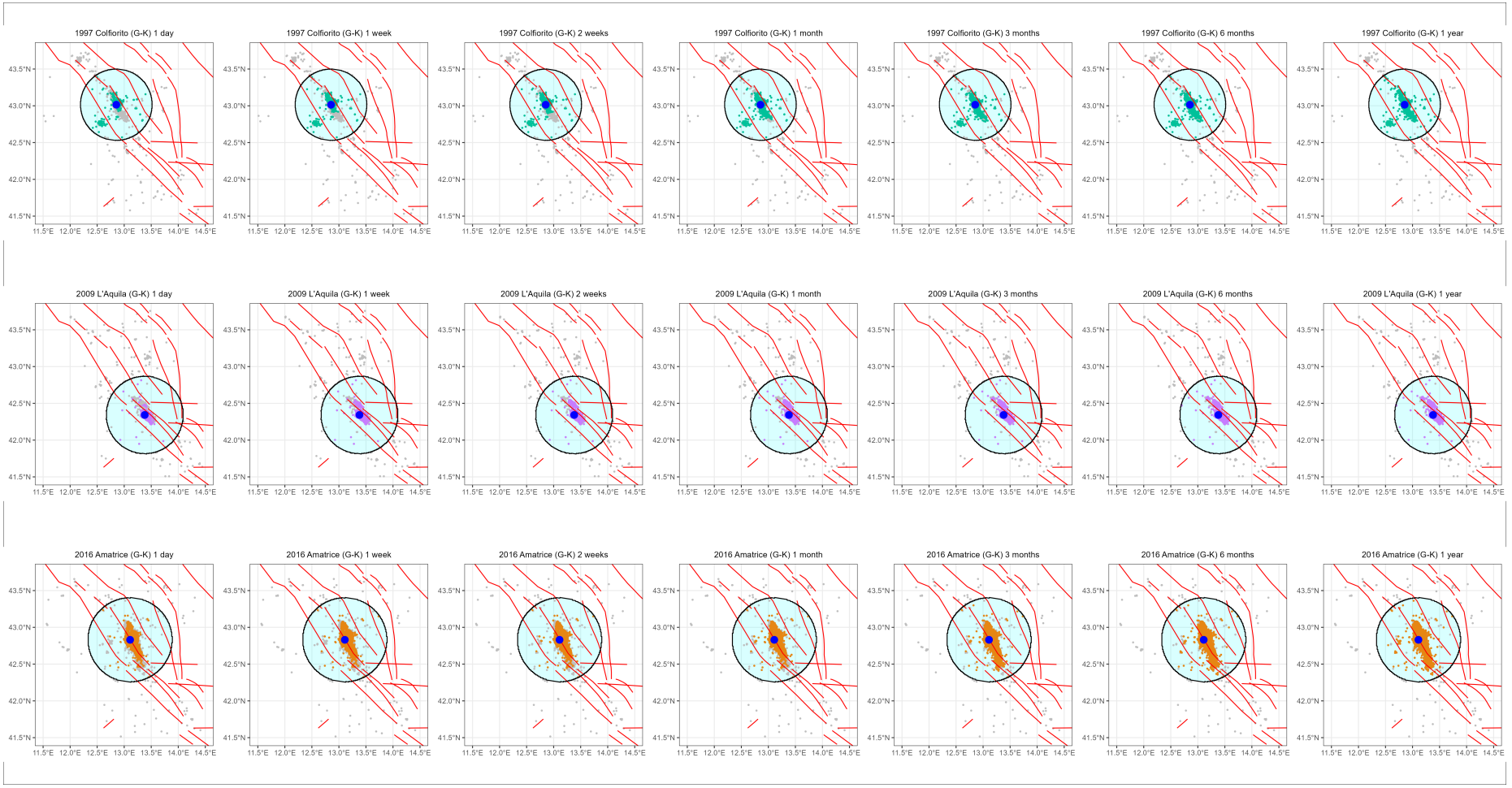

Figure 5.26: Temporal 1-day, 1-week, 2-week, 1-month, 3-month, 6-month, and 1-year windows for three central Italy sequences with different selected temporal windows after the mainshocks.

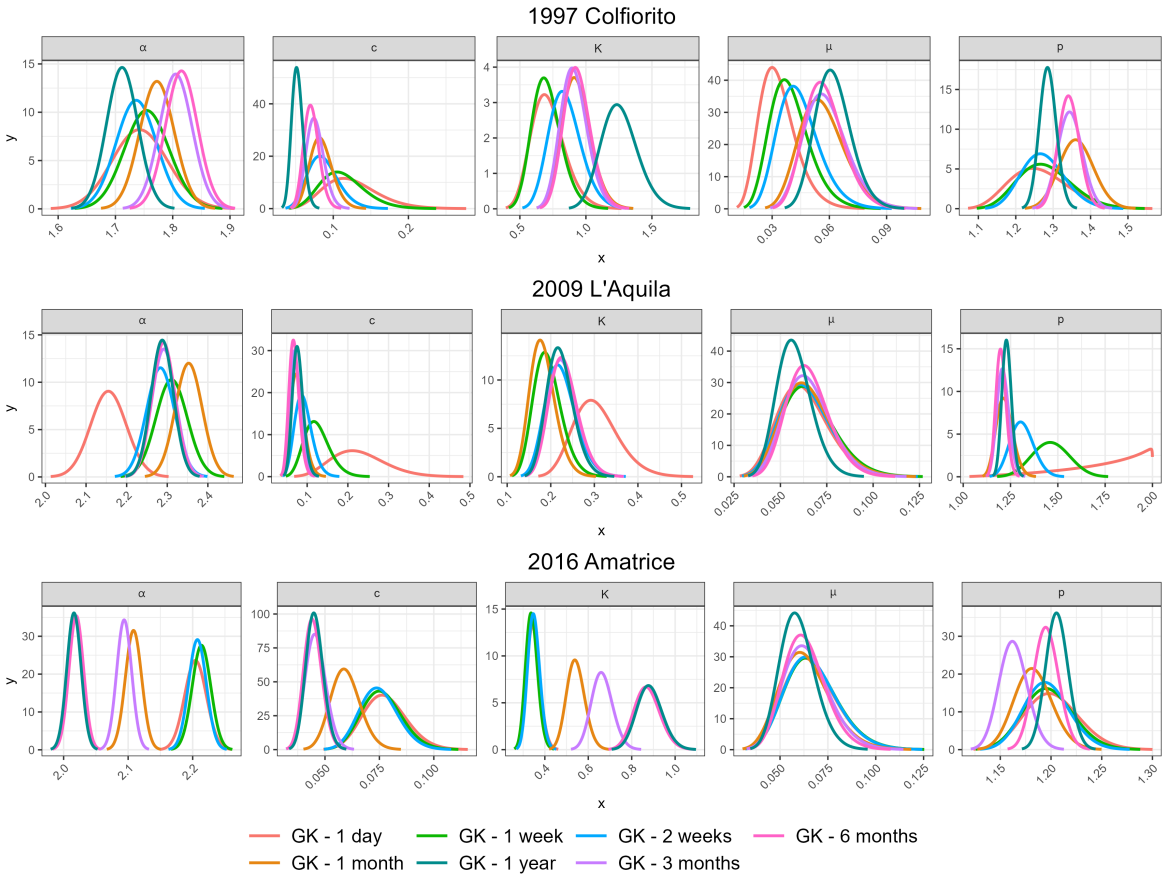

Figure 5.27: Posteriors for three central Italy sequences with different selected temporal windows after the mainshocks.

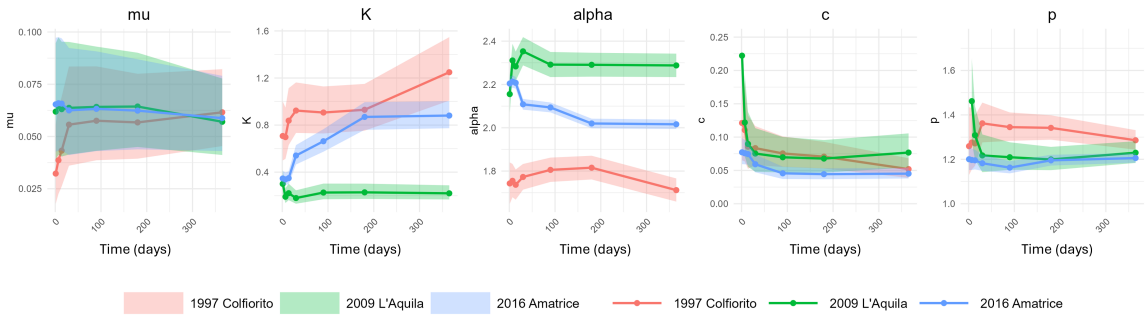

Figure 5.28: Temporal variation of ETAS parameters for three central Italy sequences, analysed using different temporal windows following the mainshocks. The solid lines represent the mode of the posteriors for each parameter, while the shaded areas indicate the uncertainty bands using quantiles.

5.9 Impact of extent of spatial modelling domain

In addition to previous considerations for selecting representative samples and reducing bias in the temporal ETAS model, there are also two other key factors significantly influencing the model's performance in the context of spatial modelling: 'extent of the spatial domain' and 'spatial anisotropy in earthquake distributions'. Here, I briefly address the former, as the choice of spatial area matters even in temporal modelling, but analysing the impact of spatial anisotropy falls beyond the scope of this thesis.

The spatial size of the domain selected to fit the ETAS model can have a marked impact on the estimated parameters. Fitting the model to data from different spatial scales, such as a single fault, a fault system, a seismogenic source, or a tectonic macrozone, can yield different values of the ETAS parameter. This variation arises because seismicity characteristics, including the background seismicity rate and aftershock productivity, can differ substantially between these scales. A smaller domain might capture specific local seismic behaviours but miss broader regional trends, leading to potentially biased parameter estimates. Conversely, a larger domain can encompass diverse seismicity patterns, offering a more comprehensive understanding but introducing complexity in model calibration and higher computational demands. For example, fitting the ETAS model to data from a single fault might provide precise local aftershock parameters, yet it may fail to generalise to other faults or regions. On the other hand, a tectonic macrozone that includes multiple fault systems can capture a wide range of seismic activities, enhancing the robustness of parameter estimates, but potentially diluting specific local behaviours.

In hazard and forecast studies, the spatial extent and shape of study areas are often arbitrarily selected, with researchers using different approaches depending on the study objectives and available data. These approaches can include circular polygons, such as those used in declustering methods, rectangular regions based on the coverage of the seismic network, or irregular polygons shaped by the spatial distribution of earthquake sources. Typically, spatial extents range from as small as 10 km to as large as 200 km. However, there is no standardised or systematic method for determining the optimal spatial extent or shape, which can lead to significant variation in results depending on the chosen domain.

To examine the impact of spatial extent on ETAS inversions, I focus again on the three earthquake sequences in central Italy: 1997 Colfiorito, 2009 L'Aquila, and 2016 Amatrice (Fig. 5.24). Data are collected from one year before the mainshock until the end of the Gardner-Knopoff temporal windows. I select several spatial windows to reflect common practices in seismic analysis. First, I use a circular window based on the Gardner-Knopoff method, widely employed in declustering to remove aftershocks and isolate mainshocks. Additionally, I define box-shaped windows with sides of 10 km, 25 km, 40 km, and 70 km to test the model's sensitivity to the inclusion or exclusion of events farther from the epicenter. I also consider rectangular windows aligned with the elongation of each sequence, accounting for the directional spread of seismicity along fault lines or geological structures. Finally, I define a non-regular polygon based on the spatial density and elongation of the seismic events, providing a more flexible and natural representation of earthquake distributions. These different spatial configurations allow me to explore how varying assumptions about the region's geometry affect the performance of the ETAS model. Table Table 5.4 compares the number of earthquakes within each spatial window.

Table 5.4: The number of earthquakes in each sequence considering the associated spatial windows. For each case, data were temporally collected from one year before the mainshock up to the end of the Gardner-Knopoff temporal windows.

Spatial window	1997 Colfiorito	2009 L'Aquila	2016 Amatrice
Gardner-Knopoff	1273	986	4256
Box (10 km)	716	594	1716
Box (25 km)	1170	947	3551
Box (40 km)	1260	976	4180
Box (70 km)	1373	1049	4283
Rectangle	1168	963	4165
Polygon	1133	960	4131

Then, I fit the ETAS model to each scenario. The modelling was started using initial values that are in reasonable ranges, as explained in Section 4.2.1. Specifically,

the initial values for the parameters were set as $\mu = 0.3$, $K = 0.1$, $\alpha = 1.0$, $c = 0.2$, and $p = 1.01$. The convergence plots are shown in Figs. 5.30 to 5.50. These plots show how parameter estimates evolve over the iterations of the algorithm. Specifically, the plots track changes in the parameter values (mode) against the iteration number. In the top-left panel of the figure ('Tracks'), each plot shows the progression of a parameter (μ, K, α, c, and p) over time, with the mode of the parameter stabilising as the algorithm converges. The 'Mode - Lin' panel further illustrates the difference between the mode and the new linearisation, with convergence indicated by the decreasing change between iterations. The bottom panels, 'Change / SD' and 'Change & SD' show the relative change in the mode of each parameter normalised by the standard deviation (SD) over iterations. As the iteration progresses, the changes in parameter estimates become smaller and stabilise, indicating that the model has converged to stable values. In these graphs, the ETAS parameters are represented on the θ scale, which is the internal parametric scale used by INLA to estimate the model. This scale follows a Gaussian distribution with a default mean of 1 and a precision of 10. The black horizontal lines in the convergence plots of change/sd represent the acceptance threshold, indicating that all values must drop below this threshold. The iterations continue until the standard deviation (sd) is less than 10% and the linear search is inactive, hence it should definitely go below 0.1. The loop terminates when the outstanding curve of change/sd drops below the threshold. Interestingly, the parameters c and p resolve quickly in fewer iterations compared to other parameters. It appears that the resolution of K and α primarily controls the time for convergence. Typically, the final iteration involves α dropping below the 0.1 threshold, but it is evident that K and α are co-varying. For presenting posterior distributions, the parameters were converted from the INLA scale to the ETAS scale. The posterior distributions are shown in Fig. 5.51.

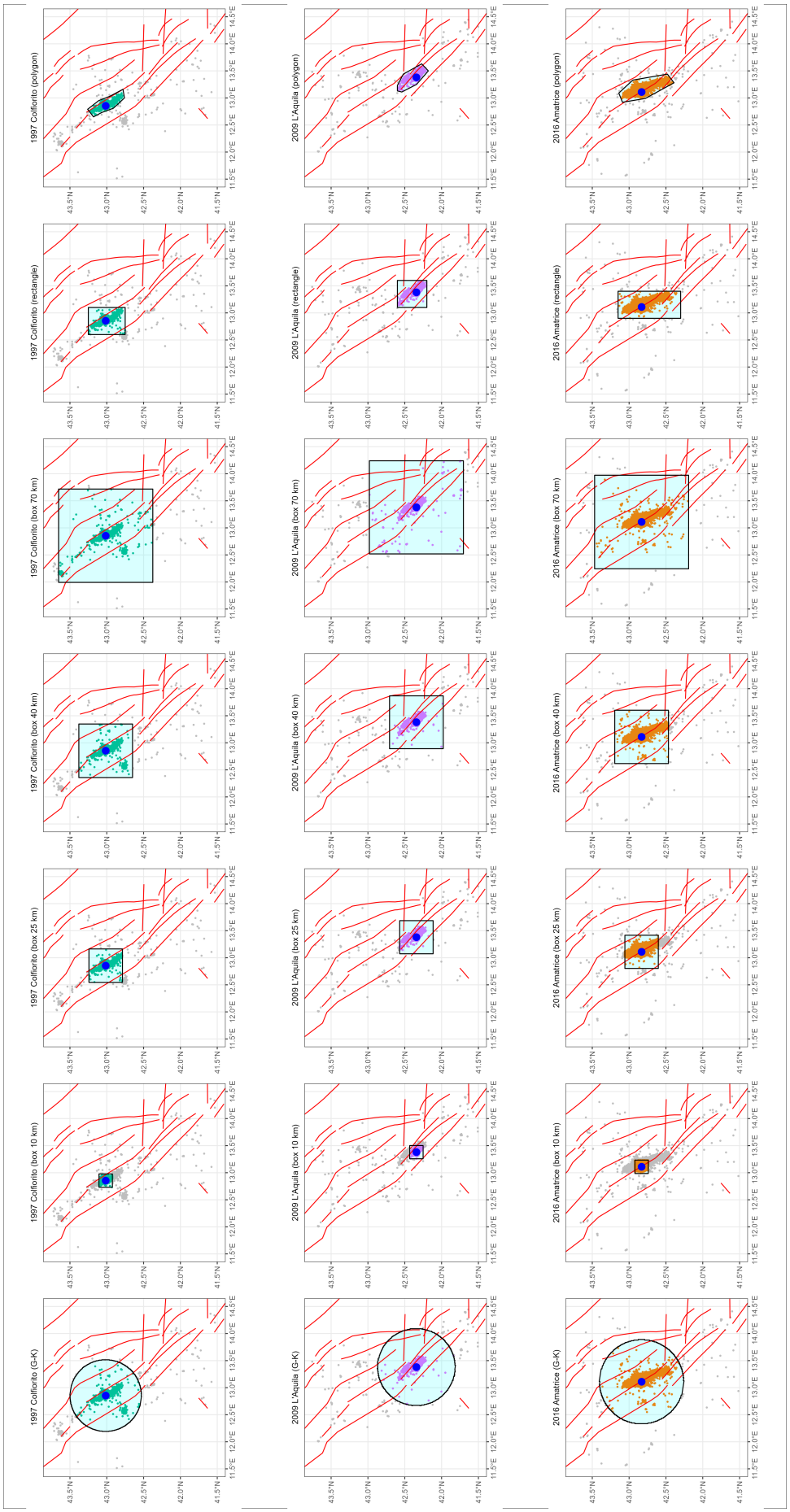

Figure 5.29: Spatial windows with different area and geometry considered for the three sequences in central Italy.

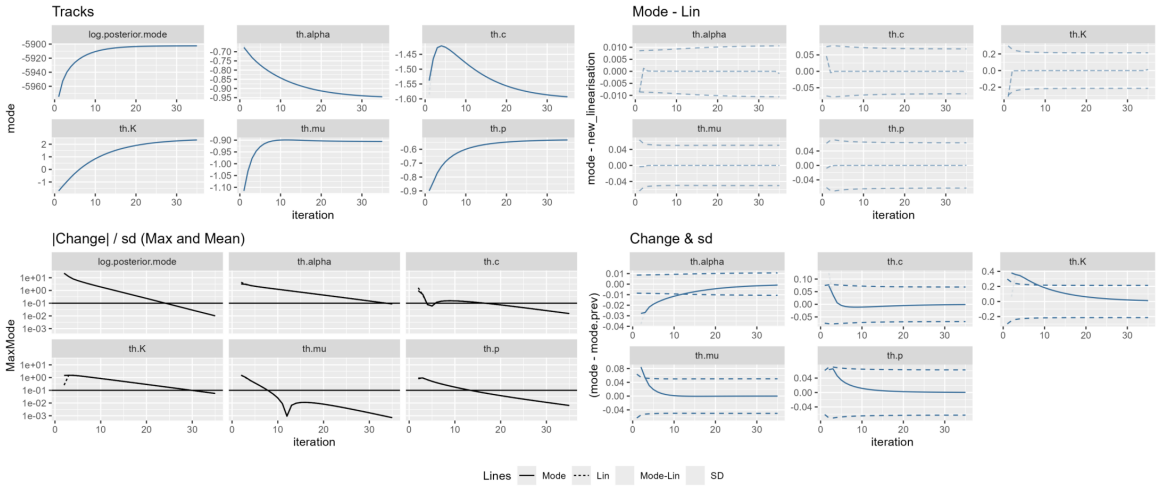

Figure 5.30: Convergence plots for the 1997 Colfiorito earthquake with a spatial G-K window. The grey boxes at the top show the parameter estimates tracked over iterations. See text for further details.

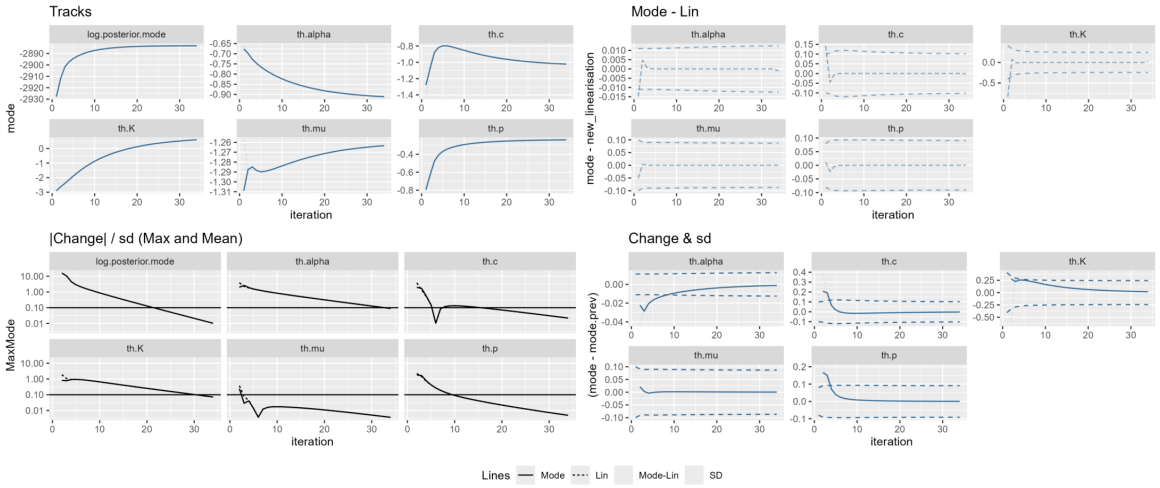

Figure 5.31: As for Fig. 5.30 but with a spatial 10-km box window.

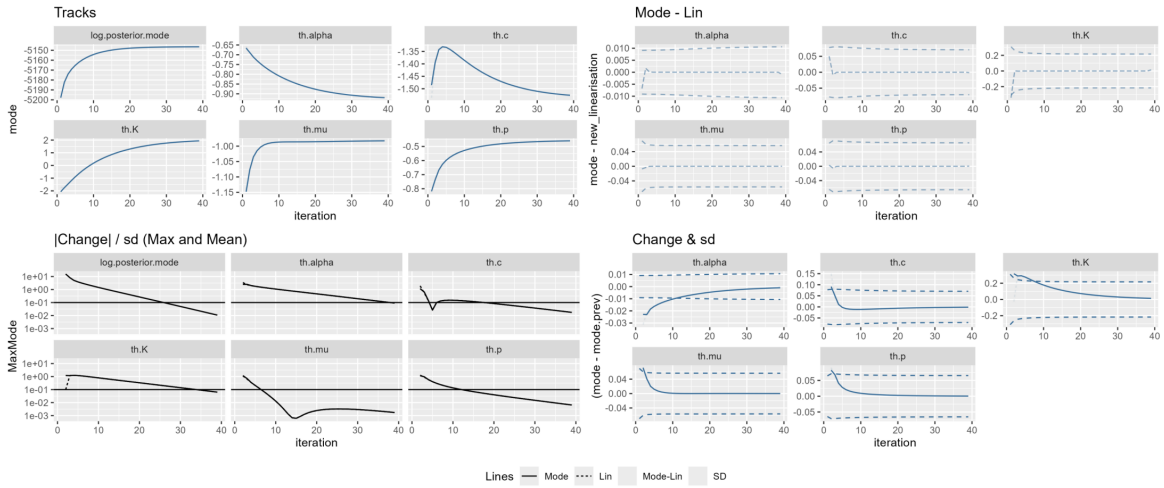

Figure 5.32: As for Fig. 5.30 but with a spatial 25-km box window.

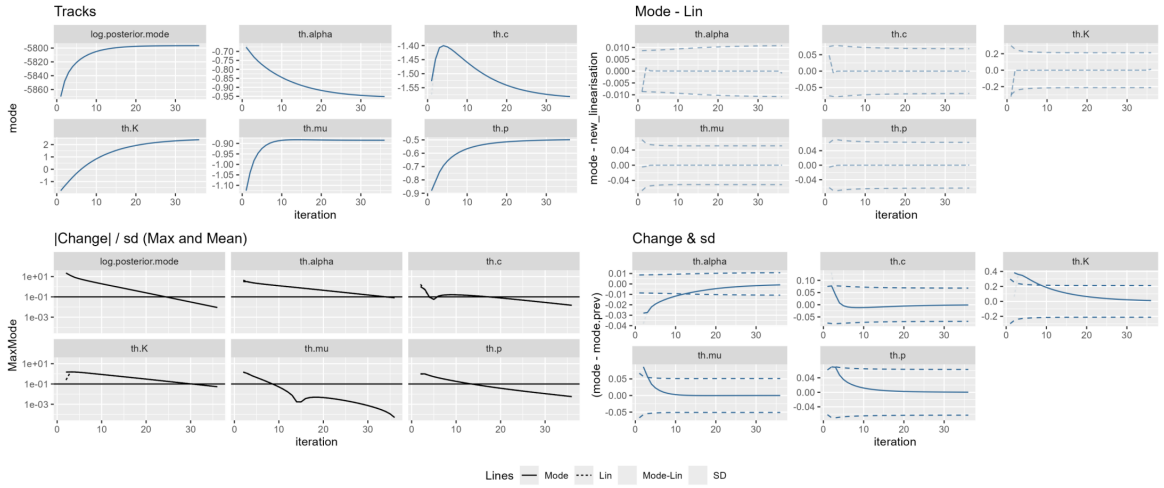

Figure 5.33: As for Fig. 5.30 but with a spatial 40-km box window.

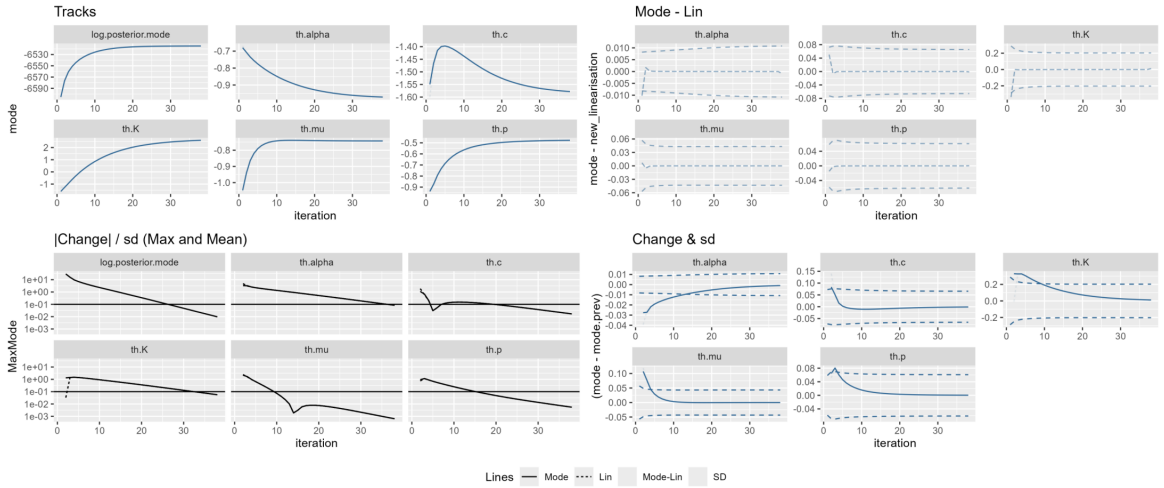

Figure 5.34: As for Fig. 5.30 but with a spatial 70-km box window.

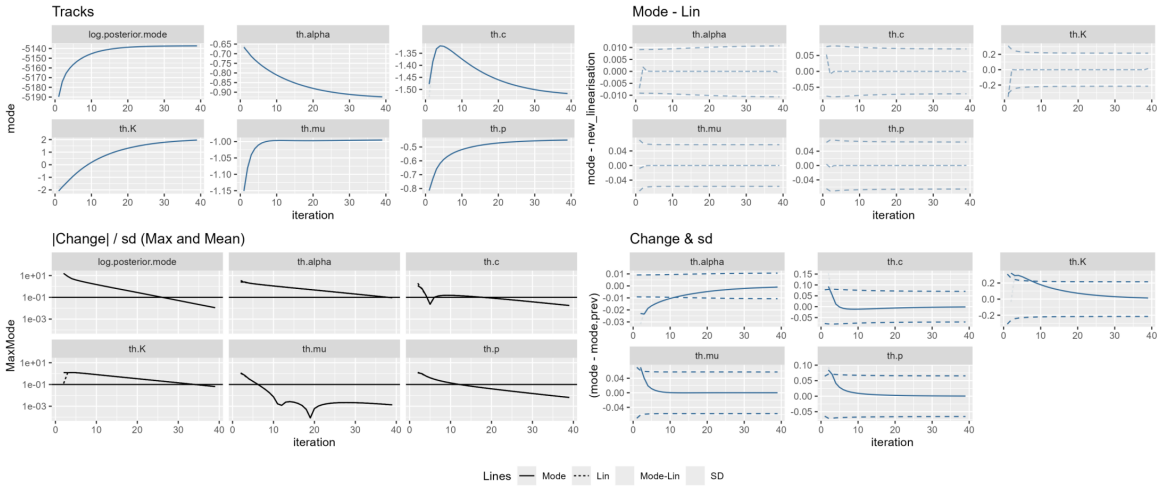

Figure 5.35: As for Fig. 5.30 but with a spatial rectangular window.

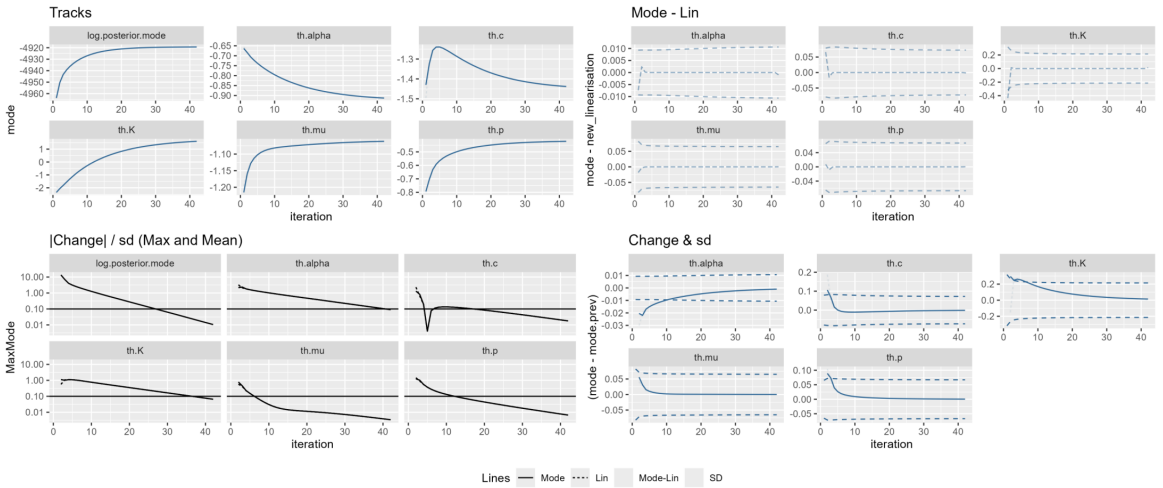

Figure 5.36: As for Fig. 5.30 but with a spatial polygon window.

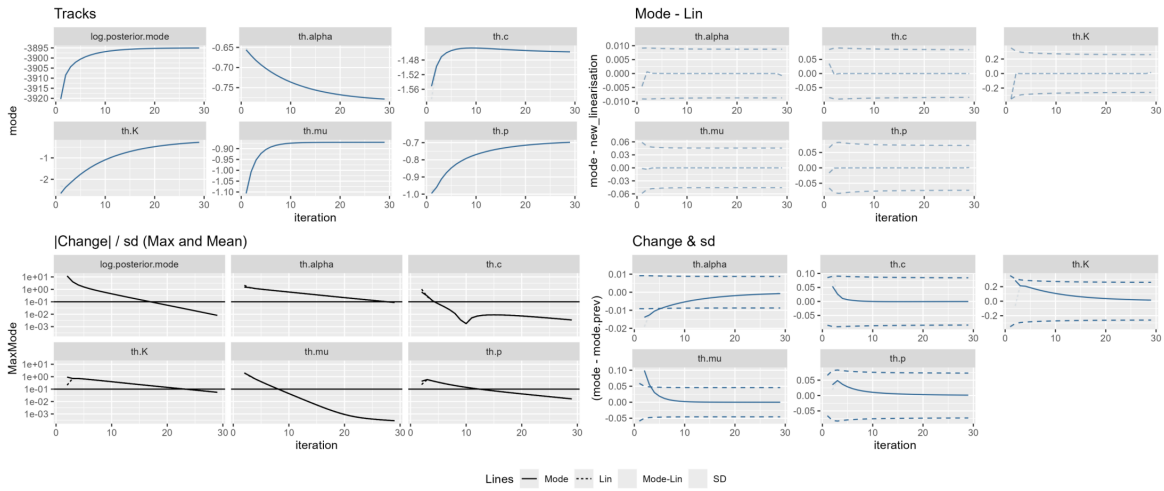

Figure 5.37: Convergence plots for the 2009 LAquila earthquake with a spatial G-K window. The grey boxes at the top show the parameter estimates tracked over iterations. See text for further details.

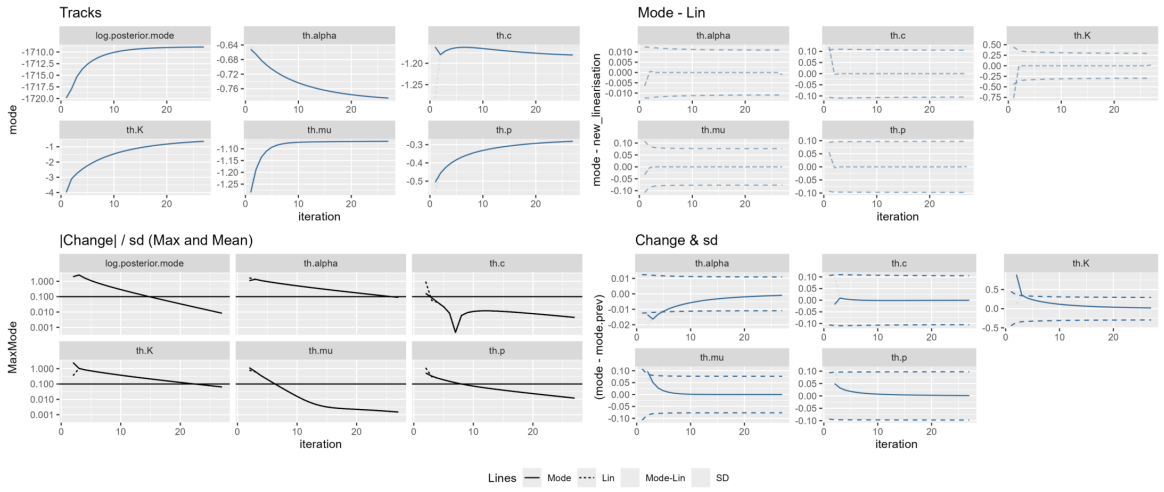

Figure 5.38: As for Fig. 5.37 but with a spatial 10-km box window.

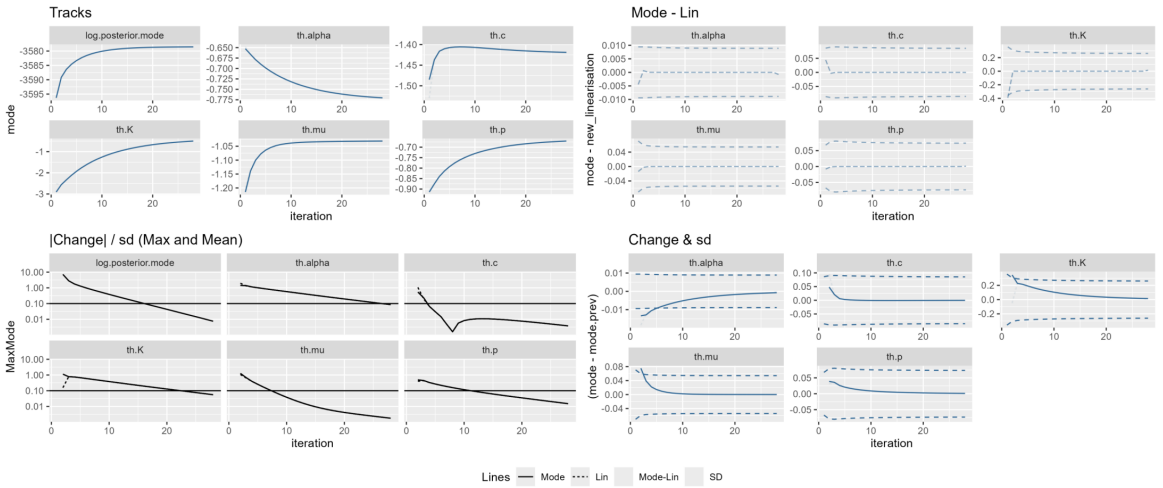

Figure 5.39: As for Fig. 5.37 but with a spatial 25-km box window.

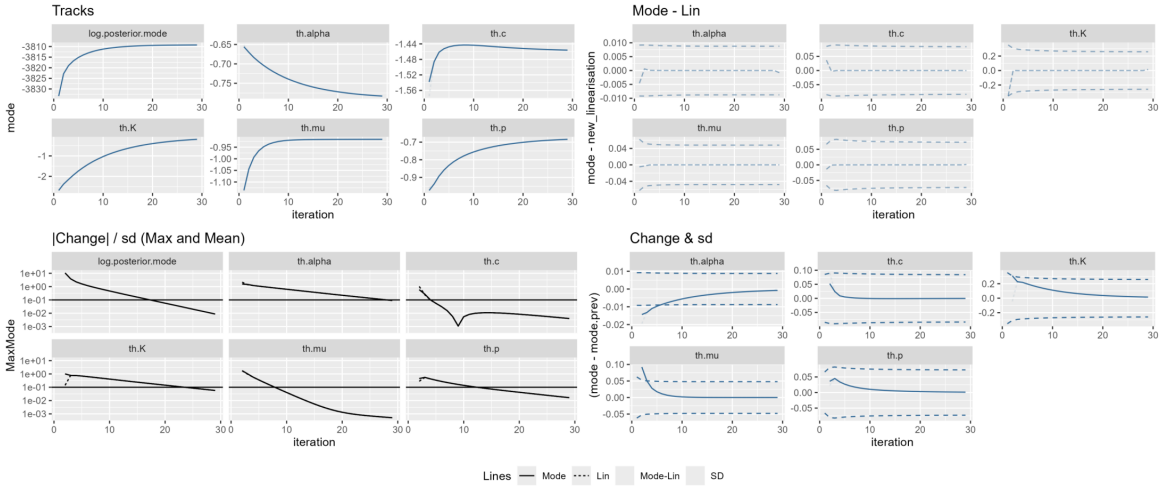

Figure 5.40: As for Fig. 5.37 but with a spatial 40-km box window.

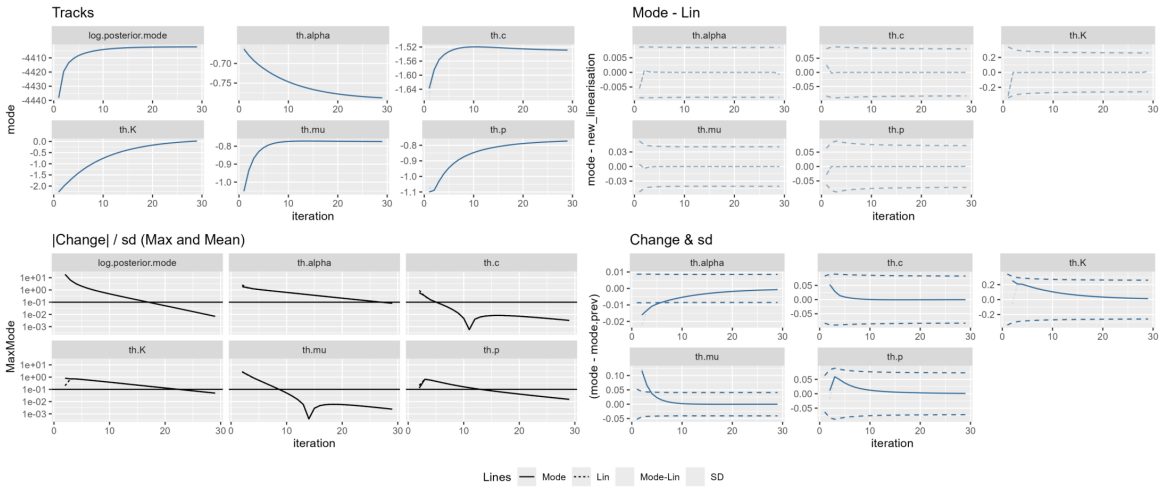

Figure 5.41: As for Fig. 5.37 but with a spatial 70-km box window.

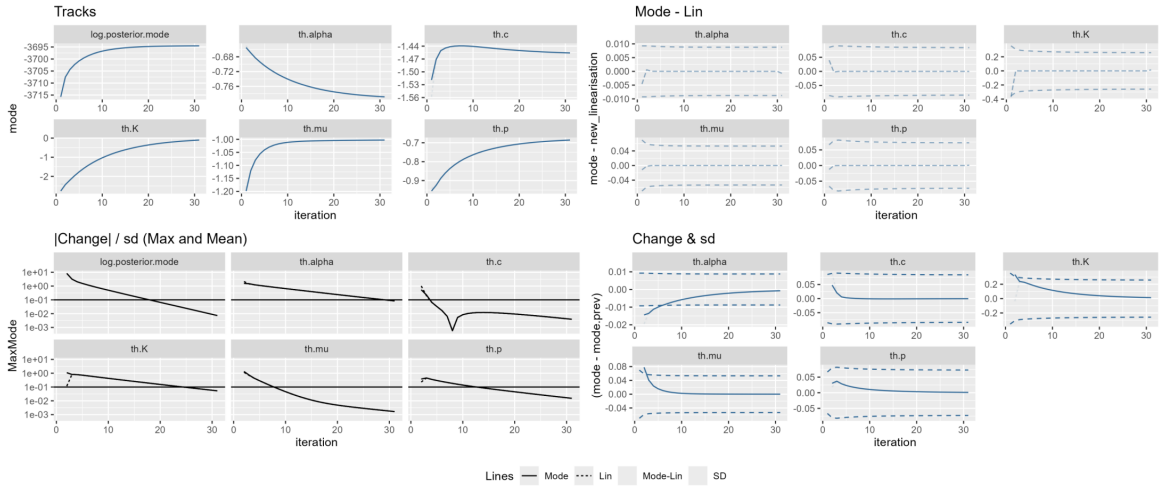

Figure 5.42: As for Fig. 5.37 but with a spatial rectangular window.

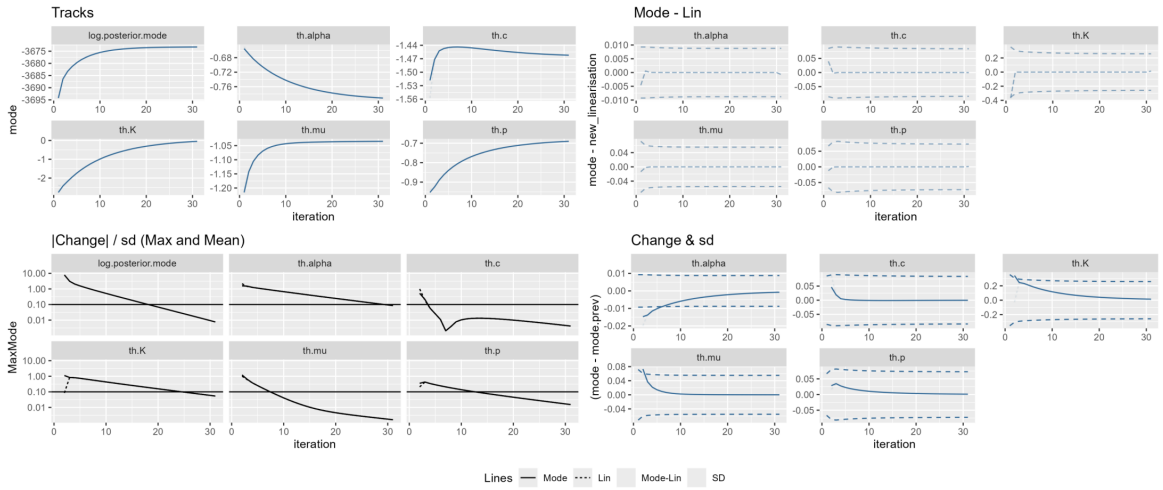

Figure 5.43: As for Fig. 5.37 but with a spatial polygon window.

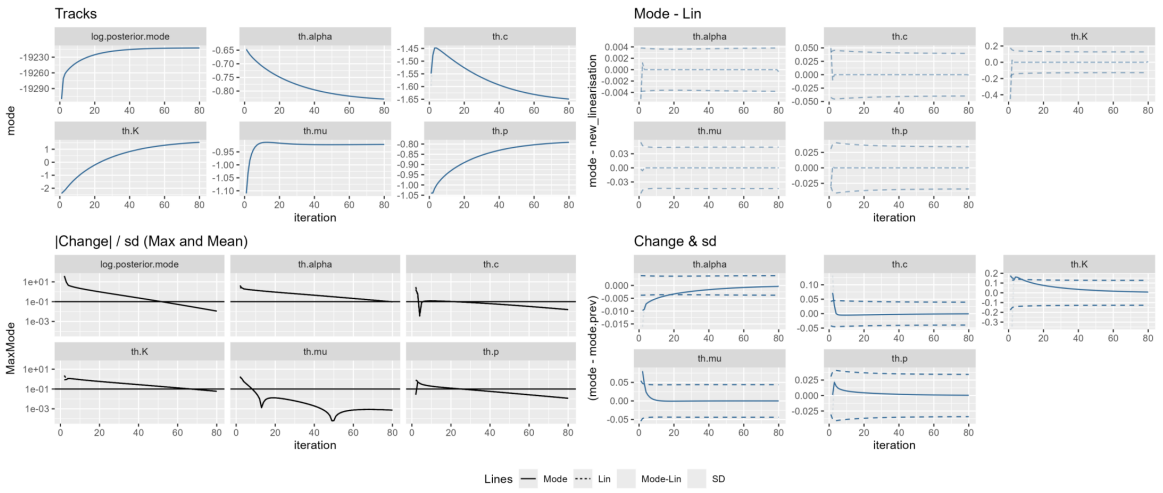

Figure 5.44: Convergence plots for the 2016 Amatrice earthquake with a spatial G-K window. The grey boxes at the top show the parameter estimates tracked over iterations. See text for further details.

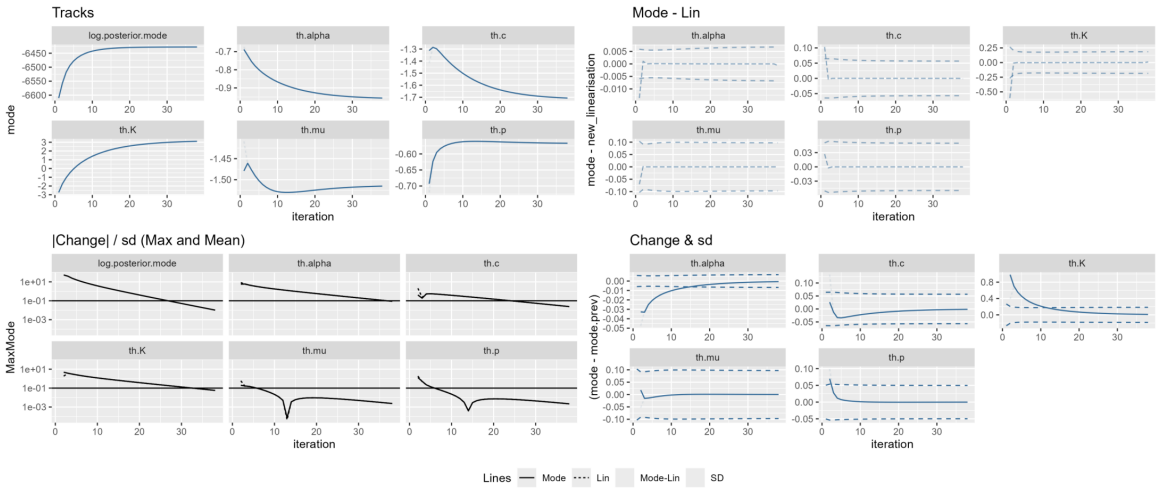

Figure 5.45: As for Fig. 5.44 but with a spatial 10-km box window.

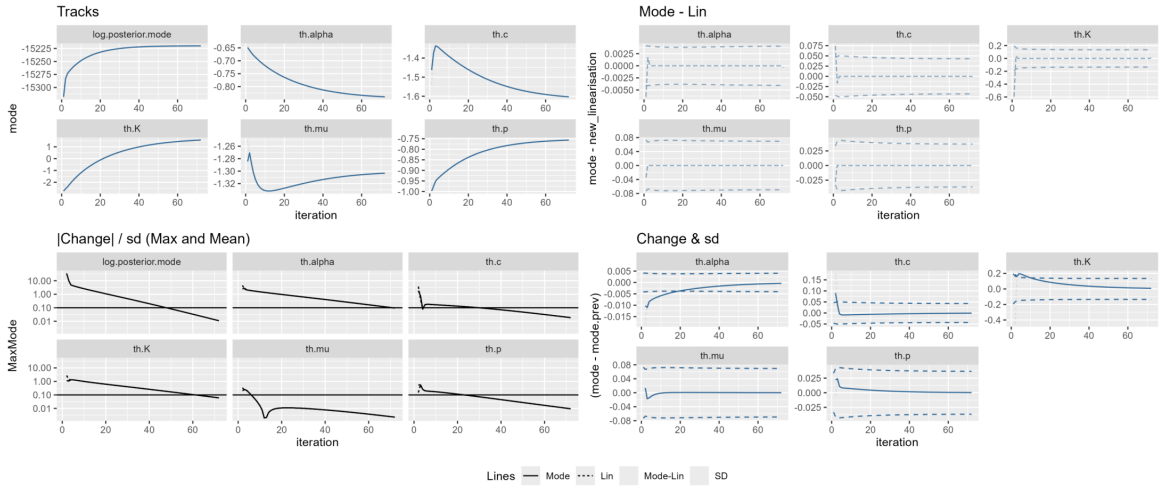

Figure 5.46: As for Fig. 5.44 but with a spatial 25-km box window.

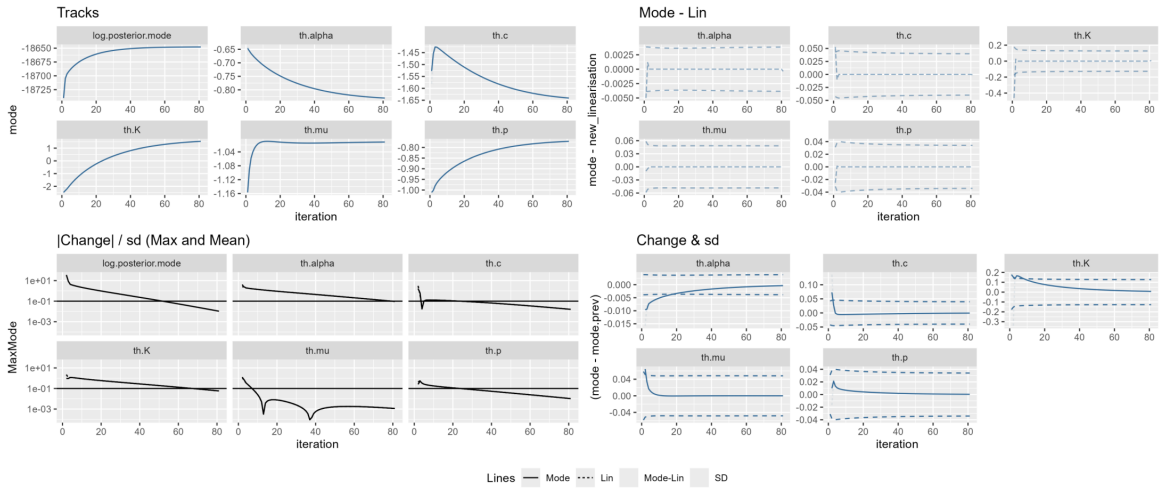

Figure 5.47: As for Fig. 5.44 but with a spatial 40-km box window.

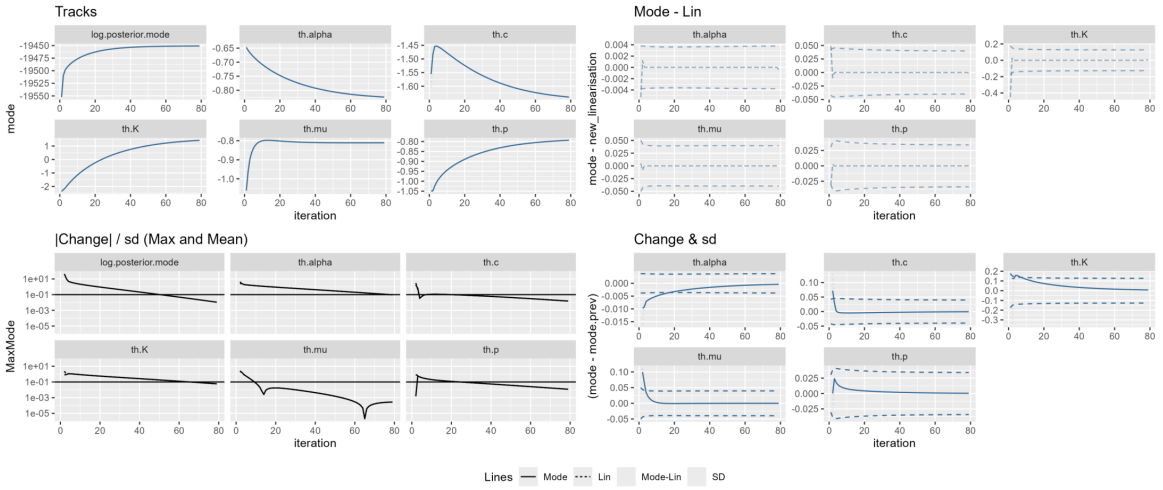

Figure 5.48: As for Fig. 5.44 but with a spatial 70-km box window.

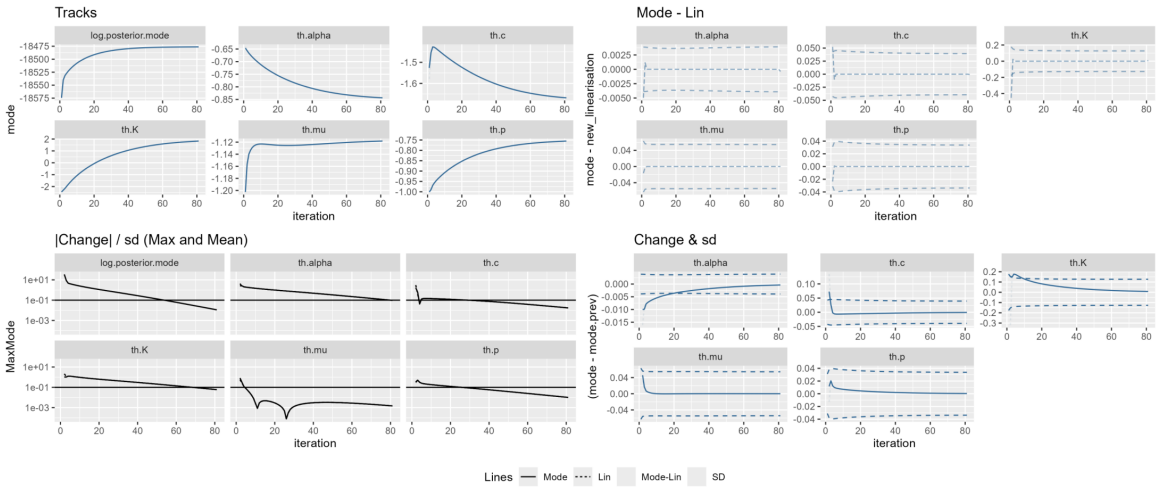

Figure 5.49: As for Fig. 5.44 but with a spatial rectangular window.

As the posteriors indicate (Fig. 5.51), the smallest box, with a 10-km length side, introduces a noticeable bias in all ETAS parameter estimates. While one might argue that we do not know the true parameters and this could be the correct estimate, it is important to consider the underlying assumptions of the ETAS model. The model is based on a mathematical framework that assumes a triggering-triggered relationship between events, meaning that missing data disrupt this interaction, which is fundamental to the model's accuracy. A 10-km box may work well for sequences with smaller mainshock magnitudes, where the spatial extent of the events is naturally limited. However, as shown in the real data in this section, for larger sequences, a 10-km box excludes a significant portion of events that fall both within the temporal window and

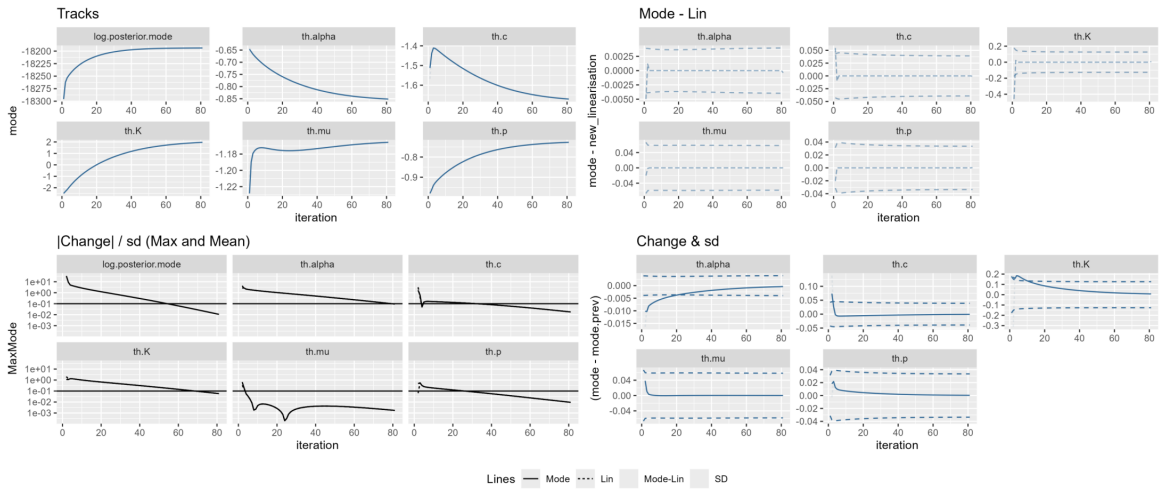

Figure 5.50: As for Fig. 5.44 but with a spatial polygon window.

spatially cluster around the main sequence. The missing data lead to an incomplete representation of the triggering process, as it disregards potential interactions between the excluded events and those within the box. By selecting such a small portion of the data, we compromise both the mathematical basis of the ETAS model and the physical and observational characteristics of the earthquake sequence. This not only biases the ETAS estimates, but also overlooks a critical part of the event distribution, which is essential for accurately modelling the triggering mechanisms that are central to the ETAS framework. When different scenarios are compared, the background parameter (μ) shows the most variation among the parameters. The results also indicate that as long as we consider a sufficiently large modelling domain that encompasses the majority of aftershocks and captures the elongation of distributed points (e.g. by a rectangle or a polygon), the posteriors remain similar regardless of the shape of the domain. In this context, a 'sufficiently large' domain refers to one that visually fits the clustered data, ensuring that the majority of events fall within the selected area and that the time window is consistent with those defined by declustering algorithms.

I also analysed the entire Central Italy region to investigate the relationship between the regional ETAS model parameters and those of individual earthquake sequences. The analysis encompassed data within the geographical coordinates of 11.5°E to 14.5°E and 41.5°N to 43.75°N, covering the period from one year before the 1997

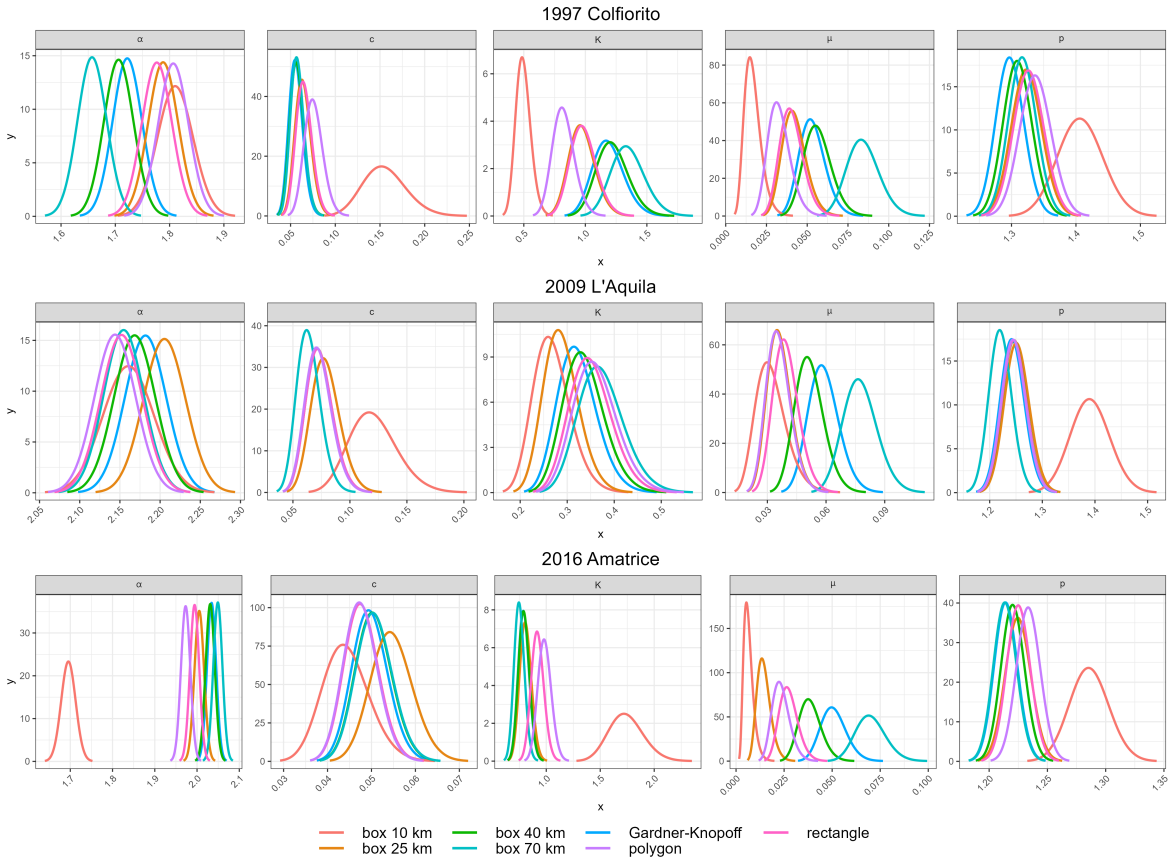

Figure 5.51: Posteriors for three central Italy sequences with different selected spatial domains. The shape and size of the domains are shown in the legend.

Colfiorito earthquake to the end of the 2016 Amatrice sequence. This regional dataset includes 11,937 earthquakes in total. The posterior distributions shown in Fig. 5.52 indicate that the ETAS model fitted to the entire area (represented in grey) differs significantly from the posterior distributions of individual sequences. Consequently, regional ETAS parameters cannot be generalised to specific individual sequences.

Given these considerations, it is crucial for modelers and practitioners to carefully select the spatial extent of the selected domain based on the study's specific objectives and the seismic region's characteristics. Testing multiple domain sizes can provide information on their impact on the model parameters. In addition, it is advisable to use spatial kernels that account for the anisotropic nature of earthquake distributions. This can involve integrating directional information into the model to better capture the true patterns of aftershock clusters. For any specific case study, a thorough validation and sensitivity analysis should be performed by applying the ETAS model to various

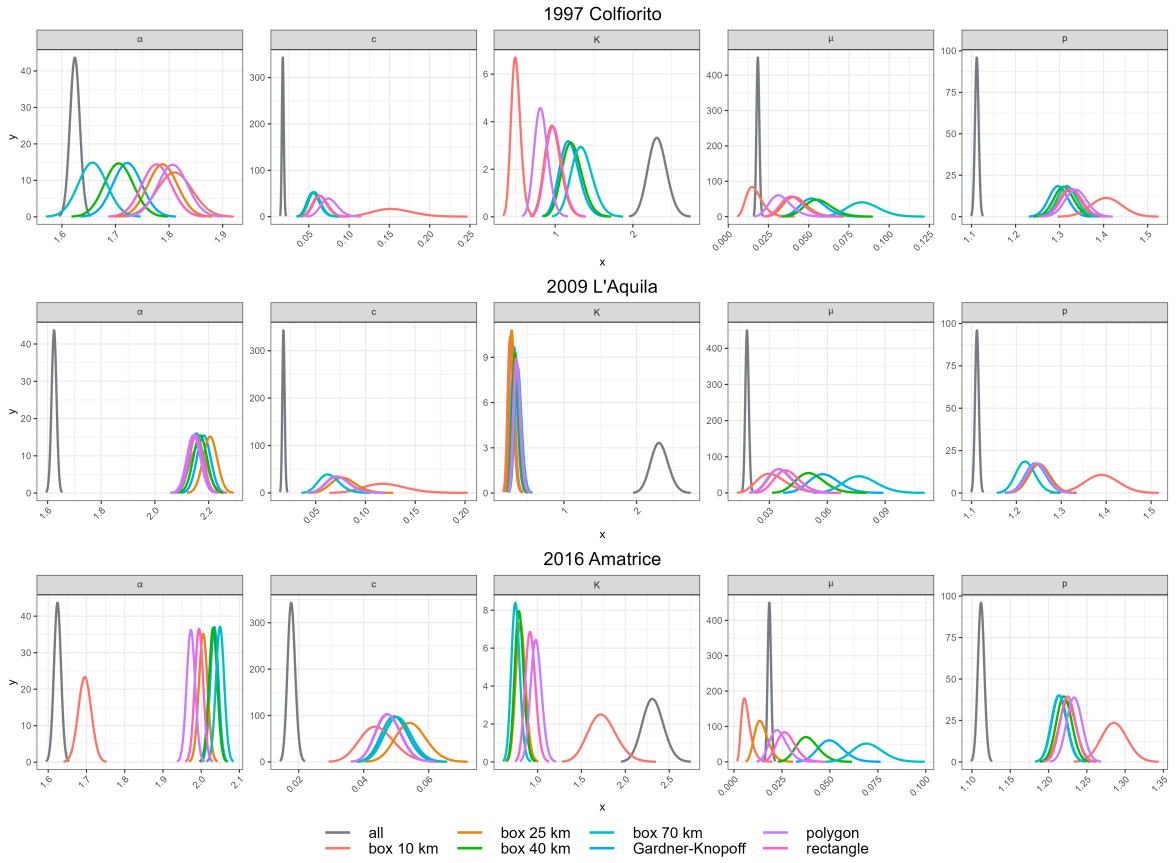

Figure 5.52: Posteriors for three central Italy sequences with different selected spatial domains along with the added regional model (in grey).

spatial domains and comparing the results. This practice helps identify and mitigate potential biases, enhancing the robustness of the ETAS inversions.

5.10 Conclusion

In this chapter, I focused on the critical task of selecting representative samples to minimise biases in the ETAS inversions. The study began by investigating the model's sensitivity to temporal binning, revealing that an appropriate temporal mesh is essential for accurate parameter estimations. By including and conditioning on the historical run-in period, the model's ability to capture the complexities of seismic sequences improved significantly. A comprehensive analysis of the combination of magnitudes within the sample data highlighted that incorporating a diverse range of magnitudes, including both small and large events, enhances the precision of key productivity parameters. This approach ensures that the model is well calibrated in different seismic scenarios.

Furthermore, the impact of the choice of incompleteness model parameters was evaluated, demonstrating that careful selection of these parameters can substantially reduce biases. The presence of secondary large aftershocks was also considered, indicating that by using the modified ETAS model, accurate estimations of the model parameters can still be achieved, even in cases involving doublet or multiple large events within an earthquake sequence. I also provided recommendations about considering the impact of temporal and spatial extent of modelling domains and also spatial anisotropy in ETAS inversions. The application of these optimised sampling strategies to real earthquake case studies validated their effectiveness, resulting in more robust and reliable modelling process.

Chapter 6

An ETAS framework for addressing long-term catalogue incompleteness

6.1 Introduction

The practice of documenting earthquakes has a long history, especially in countries with extensive written records, where the first records of seismic activities might date back thousands of years. With the development of seismology as a scientific discipline in the late 19^{th} and early 20^{th} centuries, and the establishment of the first seismographs around 1900, modern seismology was founded. This advancement enabled systematic recording of earthquakes and provided access to earthquake catalogues. The earthquakes documented in written references before 1900 are referred to as *historical earthquakes*, while those recorded by seismic instruments are classified as *instrumental earthquakes*.

One of the most significant advances in earthquake recording was the establishment of the World-Wide Standardised Seismograph Network (WWSSN) in the 1960s. This network revolutionised seismic monitoring by standardising data collection methods and expanding the geographic coverage of seismographic stations. Since the establishment of seismic networks, there have been gradual improvements in the recording of events, both in terms of increasing the number of stations and enhancing the precision and accuracy of seismograms. In light of this evolution, the instrumental era is divided into the early instrumental era (1900-1963), characterised by primary network installations, and the modern instrumental era (1964-present), marked by significant improvements. With the advent of digital technology, real-time monitoring became possible in 1978, allowing more accurate recording of seismic events.

Seismic catalogues provide the foundation for hazard assessments and earthquake forecasting. In operational earthquake forecasting and ETAS modelling, we mostly take into account recent earthquake data over short-term periods, such as years and a few decades, to focus on data with the highest resolution of completeness. However, relying solely on short-term data overlooks the importance of long-term data, particularly because large mainshocks, aftershock sequences, and earthquakes with various combinations of magnitudes typically occur over extended periods. In this regard, developing long-term models is crucial to understand the potential recurrence intervals of large earthquakes and longer-term earthquake patterns and regional seismicity behaviour, which are critical for both PSHA and OEF.

Given that we are interested in the largest events that have occurred, we should want to be able to include this information within our hazards workflow, and it would be good for it to be included within the analysis of earthquake catalogues. In a temporal model, some events will affect the expected intensity; and if building towards a spatio-temporal model, it would highlight regions of potentially high hazard within an integrated workflow.

With regard to longer catalogues, the development of long-term models requires a nuanced understanding of the interplay between seismicity forecasting and PSHA. Although each field has its distinct focus, both fundamentally need to understand the same concepts of triggering mechanisms, even if they use these concepts in different ways. Seismicity forecasting, often operating on shorter timescales, relies on detailed temporal models to capture the dynamic nature of earthquake sequences and immediate aftershock probabilities. In contrast, PSHA requires models that integrate both short-term seismicity data and long-term historical records to provide a comprehensive view of seismic hazards. By revisiting and clearly delineating the boundaries between these two approaches, we can juxtapose their methodologies, revealing how seismicity forecasting informs the immediate response to seismic events, while PSHA underpins structural design and long-term risk mitigation strategies. This integration of approaches allows us to develop sophisticated models that leverage the strengths of both fields, thus enhancing our ability to predict and mitigate the impacts of earthquakes over varying timescales. This holistic approach is pivotal in the advancement of both operational forecasting and strategic hazard assessments, ultimately contributing to a more flexible understanding of seismicity.

On the other hand, one of the major challenges in developing models for operating over longer periods is the need for models that can address specific forms of long-term incompleteness in earthquake catalogues. Due to various factors, such as technological limitations, changes in network coverage, and historical recording practices, earthquake catalogues often suffer from gaps and inconsistencies. In the early days of modern seismology, many smaller earthquakes were not recorded due to the limited sensitivity of the instruments and the sparse distribution of seismographic stations. In the long term, gradual improvements in seismic detection technology have led to step-like increases in

the detection of smaller earthquakes. An example of such development and the step-like increase in data recording, and consequently, the step-like decrease in the magnitude of completeness, is shown for the Iranian seismic network in Fig. 6.1. In this figure, the gradual temporal decrease in the magnitude of completeness is observed with the increase in seismic stations over 123 years of instrumental recording from 1900 to 2023. Interestingly, there is a clear decrease in temporal peak in 1962, coinciding with the occurrence of the Buin Zahra earthquake (Ambraseys and Melville 2005; IRSC 2024). It is likely that a local seismic network was installed after this earthquake. Installing temporary seismic networks a few hours after large earthquakes is a common practice to record aftershocks and monitor seismicity over several weeks, months, or even a few years (Margheriti et al. 2011). This is also evident in the example of the 2016 Amatrice earthquake in Italy, shown in Fig. 6.2, where after the mainshock, four Italian institutions installed a dense temporary network of 50 seismic stations in the epicentral area (Cara et al. 2019). In this figure, a more complicated record of events is depicted. Initially, a short-term exponential incompleteness is observed in the early aftershocks during the first hours following the mainshock (the red curve). As expected, once this rate-dependent short-term incompleteness decays, the data level is expected to return to the previous long-term completeness level of the network (the green dashed line). But less than 48 hours after the mainshock, the temporary local seismic network was installed. Consequently, with the increase in the number of stations in the epicentral area, more aftershocks with smaller magnitudes were detected, resulting in a lower long-term completeness level (the purple line). This complex case, which includes both short-term and long-term changes in the magnitude of completeness, should be carefully considered when developing ETAS and forecasting models.

In this chapter, I aim to develop an extended version of the ETAS model that can operate on long-term earthquake catalogues and address step-like incompleteness in the data. One might argue that the ETAS model is traditionally designed to only model aftershock sequences, which typically last only months to a few years, making its extension to longer periods seem unreasonable. However, in real-world scenarios and in many regions, even in high-seismic areas, we usually lack a single clear individual sequence of aftershocks for fitting the ETAS model. Therefore, we need to settle for whatever data we have and explore the possibility of using longer-term datasets to

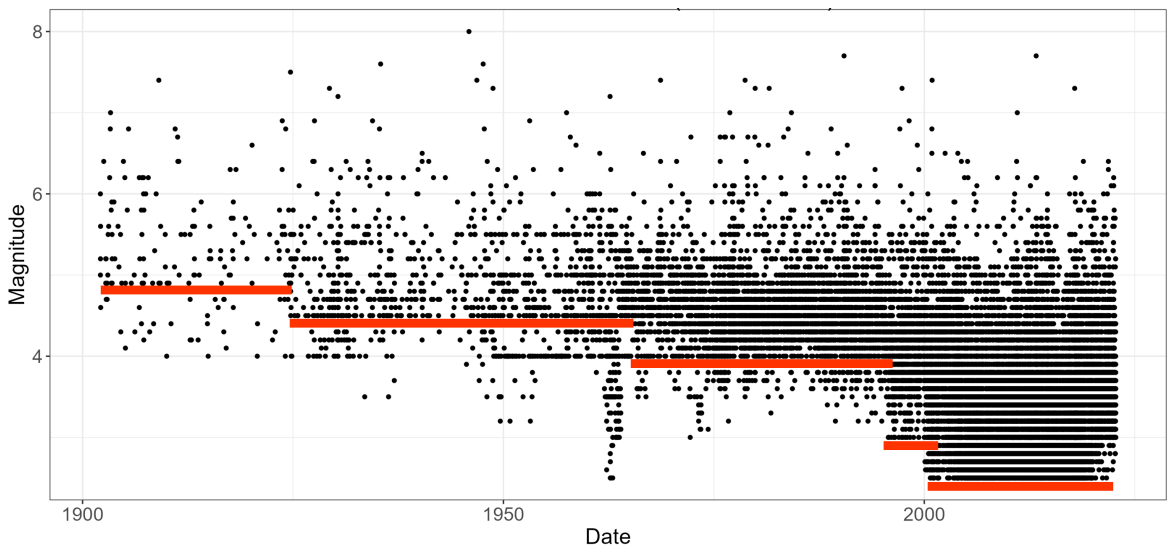

Figure 6.1: Development of earthquake recording and the gradual increase in the detection of smaller events from 1900 to 2023 for the Iranian seismic network. The red lines show the step-like decrease in the magnitude of completeness as the number of seismic stations increased over 123 years of instrumental recording. (Source of catalogue: Iranian Seismological Centre. Data publicly available at http://irsc.ut.ac.ir/, last accessed August 2023).

retrieve ETAS parameters. Moreover, some studies suggest that the ETAS model can be applied to long-term seismic hazard assessments, hypothesising that the ETAS parameters should remain relatively constant over different time scales within a specific area. In this section, I aim to develop a long-term version of the ETAS model and investigate whether we can observe the same ETAS behaviour over different time scales, improving our understanding of short-term and long-term seismicity behaviour and earthquake hazards.

Before testing this hypothesis, I first need to modify the inversion algorithm of the ETAS model to handle long-term data incompleteness, similar to the approach I used to accommodate short-term incompleteness in Chapter 4. Therefore, the new version of the ETAS model developed in this chapter will expand the temporal scope of the modelling domain, making it adaptable to datasets that are incomplete over longer periods. By developing both short-term and long-term modified ETAS models, we can obtain model parameters that ultimately depict a relatively comprehensive picture of seismicity behaviour in an area.

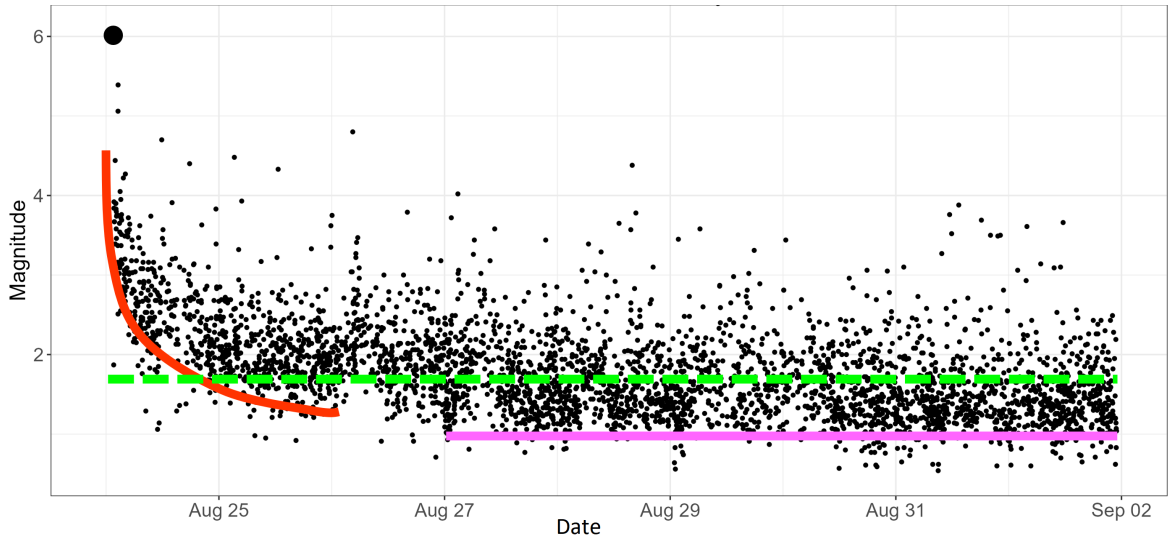

Figure 6.2: Short-term and long-term incompleteness levels during the first week after the 24 August 2016 Amatrice earthquake in Italy. The red curve represents the short-term exponential incompleteness observed in the early aftershocks immediately following the mainshock. The green dashed line marks the original long-term completeness level of the network before the mainshock. The purple line indicates the lowered long-term completeness level, which resulted from the installation of a dense temporary local seismic network within 48 hours of the mainshock. This figure highlights the importance of considering both short-term and long-term changes in completeness when developing ETAS and forecasting models. (Source of catalogue: INGV. Data publicly available at https://terremoti.ingv.it/en, last accessed August 2023).

6.2 Methodology

6.2.1 Model for long-term incompleteness in earthquake records, $M_c(t)$

As explained in the previous section, there is an intermittent decrease in the level of magnitude of completeness (M_c) over long-term periods. Since the establishment of seismic networks, and as we have moved forward in time, our ability to detect more events with lower magnitudes has improved. These improvements commonly result in a step-like decrease in M_c, above which we can sample all earthquakes of magnitude $\geq M_c$. So, the step function for the long-term time-varying M_c is defined as:

$$M_c(t) = \begin{cases} M_{c1} & \text{if} \quad t_1 \leq t < t_2 \\ M_{c2} & \text{if} \quad t_2 \leq t < t_3 \\ M_{c3} & \text{if} \quad t_3 \leq t < t_4 \\ ... & ... \quad ... \\ M_{cn} & \text{if} \quad t_n \leq t < t_{n+1} \end{cases} \tag{6.1}$$

where t is the given time and $M_c(t)$ is the corresponding magnitude of completeness at that time. A simple example of long-term changes in M_c can be:

$$M_c(t) = \begin{cases} M_{c1} & \text{if} \quad t < 500B.C. \quad \text{(paleoseismological data)} \\ M_{c2} & \text{if} \quad 500BC. \leq t < 1900 \quad \text{(historical records)} \\ M_{c3} & \text{if} \quad 1900 \leq t < 1964 \quad \text{(early instrumental records)} \\ M_{c4} & \text{if} \quad 1964 \leq t < 2005 \quad \text{(modern instrumental records)} \\ M_{c5} & \text{if} \quad 2005 \leq t < 2023 \quad \text{(latest instrumental records)} \end{cases} \tag{6.2}$$

With decreasing values of $M_c(t)$ over time, we have $M_{c1} > M_{c2} > M_{c3} > ... > M_{cn}$. It should be noted that this step-like decrease in M_c is generic but might have some exceptions. For example, in real catalogues for some regions, there might be downward steps at times when, after major earthquakes, temporary local seismic stations are deployed in the epicentral region. These values will return to their previous levels once the temporary stations are removed. Therefore, when dealing with real data, we need to consider such exceptions on a case-by-case basis. In the next section, I will explain how I will use this stepwise model to modify ETAS for longer-term modelling schemes.

6.2.2 Long-term censorship function

Similar to what I defined as a time-dependent censorship function in section 4.2.4.1 and Eq. 4.13, I use the same definition here for the long-term version. Specifically, I define the ratio between the number of events above $M_c(t)$ and the number of events above M_0 (as illustrated in Fig. 6.3). This ratio provides us with the proportion of events above M_0 that have been observed.

$$\frac{N(m \geq M_c(t))}{N(m \geq M_0)} = 10^{-b(M_c(t)-M_0)}. \tag{6.3}$$

where $M_c(t)$ is a step model as defined in the previous section. Substituting Eq. 6.1 into Eq. 6.3 for $Mc(t)$, we get:

$$\pi(t) = \begin{cases} 10^{-b(M_{c1}-M_0)} & \text{if} \quad t_1 \leq t < t_2 \\ 10^{-b(M_{c2}-M_0)} & \text{if} \quad t_2 \leq t < t_3 \\ 10^{-b(M_{c3}-M_0)} & \text{if} \quad t_3 \leq t < t_4 \\ \dots & \dots \quad \dots \\ 10^{-b(M_{cn}-M_0)} & \text{if} \quad t_n \leq t \quad t_{n+1} \end{cases} \tag{6.4}$$

For each incompleteness step, the ratio represents the apparent (observed or recorded) rates as a fraction of the actual rates (rates occurring in reality, encompassing both observed and unobserved events with smaller magnitudes). Over time, the apparent rate has an increasing step-like trend. In the next section, I will use this censorship factor to modify the ETAS for longer-term modelling.

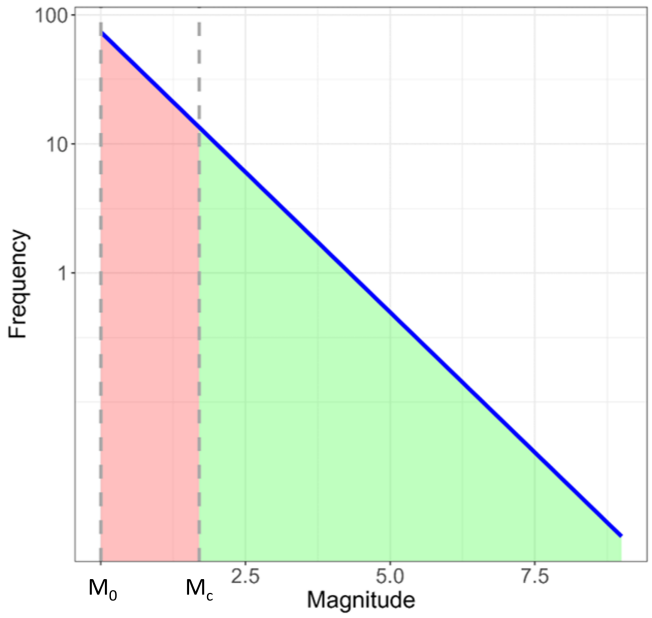

Figure 6.3: Schematic concept of censorship using the Gutenberg-Richter diagram. The two dashed grey lines show the magnitude levels for M_0 (left) and $M_c(t)$ (right). The censorship is defined as the ratio between the number of events with magnitudes larger than $M_c(t)$ (the green area) to the the number of events with magnitudes above M_0 (the red + green area).

6.2.3 Modifying the inversion algorithm for long-term ETAS

To accommodate the new censorship function for long-term incompleteness in the ETAS inversion algorithm, I first modify the ETAS conditional intensity function, as explained in Eq. 4.15. This modification represents apparent rates rather than actual rates. Unlike short-term incompleteness, where the correction factor had a slight effect on the background rate and was therefore disregarded (assuming μ was constant), for the long-term version, I apply the censorship to the background term since its apparent changes over long periods are not negligible. Consequently, the modified conditional intensity becomes:

$$\lambda_{\text{modified}}(t|\mathcal{H}_t, M_c(t)) = \mu 10^{-b(M_c(t)-M_0)} + \sum_{(t_i,m_i)\in\mathcal{H}_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p} 10^{-b(M_c(t)-M_0)}. \tag{6.5}$$

Similarly to the process for modifying short-term incompleteness in section 4.2.4.2, this change in the intensity function for long-term incompleteness also leads to changes in the approximation of the likelihood function. By substituting for λ from Eq. (6.5) in Eq. (4.5), I obtain the new modified likelihood function for the long-term version of the step ETAS model as:

$$\begin{aligned}\mathcal{L}(\theta|\mathcal{H}) = &-\int_{T_1}^{T_2} \mu 10^{-b(M_c(t)-M_0)}\,dt \\ &-\int_{T_1}^{T_2} \sum_{(t_i,m_i)\in H_t} Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p} 10^{-b(M_c(t)-M_0)}\,dt \\ &+\sum_{(t_i,m_i)\in H_t} \log\left(\mu + \sum_{(t_j,m_j)\in H_t} Ke^{\alpha(m_j-M_0)}\left(\frac{t-t_j}{c}+1\right)^{-p}\right) 10^{-b(M_c(t)-M_0)}\end{aligned} \tag{6.6}$$

The censorship factor has been applied to all likelihood terms, including the background rate (first line), the triggering part (second line), and the sum of log λ (third line). Then, by solving the internal integral for the triggering part of Eq. 6.6 and subsequently incorporating the time binning strategy, along with linearisation (as previously was implemented in Section 4.2.4.2), the modified log-likelihood for the long-term ver-

sion is:

$$\begin{aligned}\overline{\mathcal{L}}_{\text{modified}}(\boldsymbol{\theta}|\mathcal{H}) = &-\exp\left\{\log\left(\mu\sum_{k=1}^{n}(T_{2_{sk}}-T_{1_{sk}})\,10^{-b(M_{ck}-M_0)}\right)\right\}\\ &-\sum_{(t_i,m_i)\in H_t}\sum_{j=0}^{B_i-1}\exp\left\{\log K+\alpha(m_i-M_0)+\log\left(\frac{c}{1-p}\right)\right.\\ &\left.+\log\left(10^{-b(M_{cj}-M_0)}\right)+\log\left[\left(\frac{t_{j+1}^{b_i}-t_i}{c}+1\right)^{1-p}-\left(\frac{t_j^{b_i}-t_i}{c}+1\right)^{1-p}\right]\right\}\\ &+\sum_{(t_i,m_i)\in H_t}\log\left(\mu+\sum_{(t_i,m_i)\in H_t}Ke^{\alpha(m_i-M_0)}\left(\frac{t-t_i}{c}+1\right)^{-p}\right)10^{-b(M_c(t)-M_0)}\end{aligned} \tag{6.7}$$

6.2.4 Considerations for time binning in the implementation

Similar to the implementation for short-term incompleteness, where the positions of bins were compared to the completeness/incompleteness periods, I use a similar approach here. However, unlike the short-term version, which had only one incompleteness period, this long-term model has several incompleteness steps. Therefore, the scenarios for bins of length $[T_{1b}, T_{2b}]$, as explained in Section 4.2.4.3, are more complex. In general, I have defined a maximum of 10 steps (as an upper bound) for long-term modelling (which can be reduced to fewer steps or even extended if necessary). Considering the temporal mesh, defined with a chosen resolution strategy and the length of bins, each bin might completely fall within one step or span multiple steps (requiring splitting into sub-bins). The considerations are similar to those explained in Section 4.2.4.3, but there are about 55 possible scenarios for a 10-step incompleteness model. To avoid detail of each one, I have illustrated them with a schematic plot in Fig. 6.4.

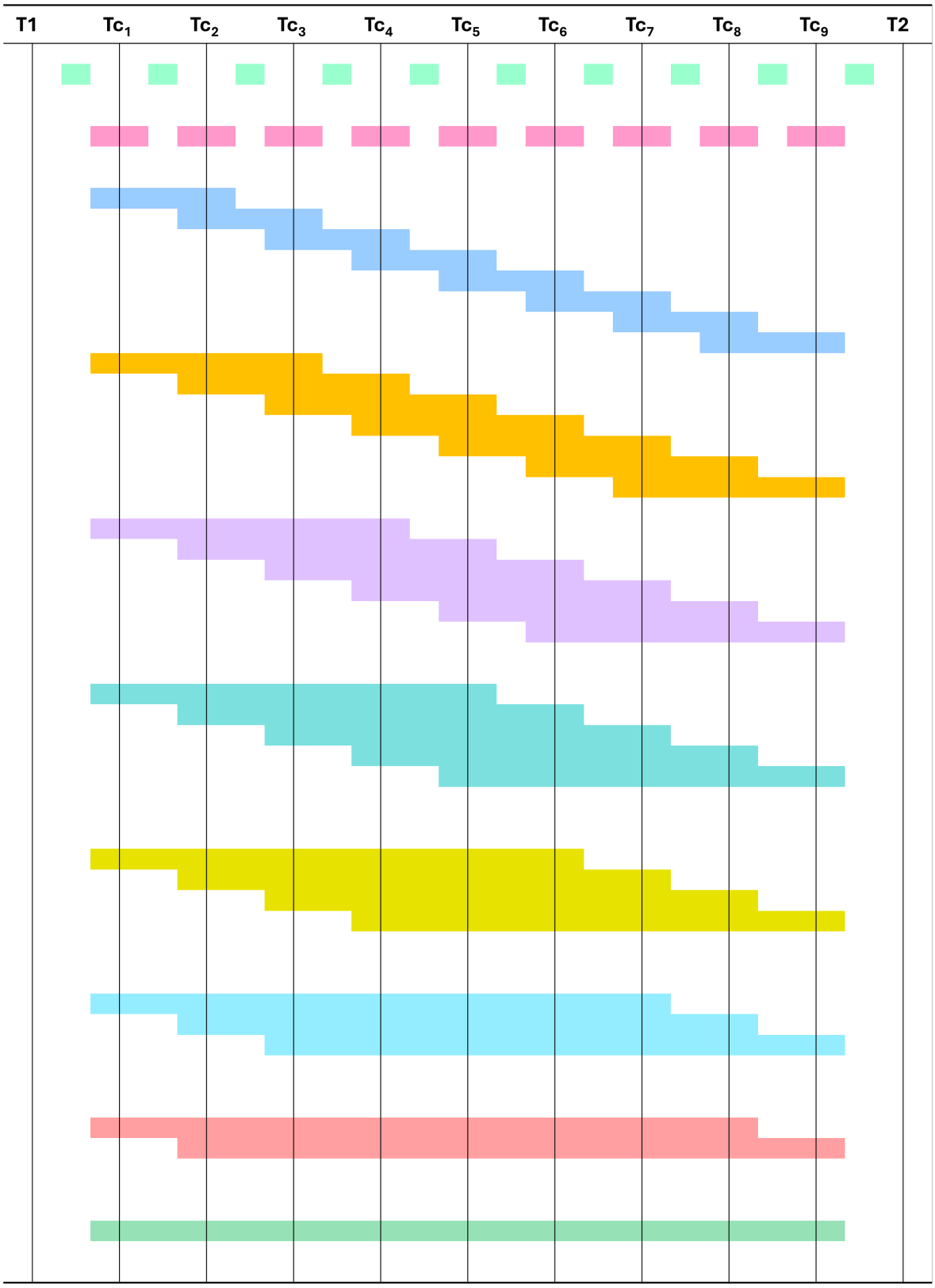

Figure 6.4: Schematic possible scenarios for time binning for a model with 10-step long-term incompleteness. T_1 and T_2 indicate the start and end of the modelling domain, and Tc_1 to Tc_9 represent the times of step changes. Coloured lines illustrate the length of possible bins located within or passing through different steps of a 10-step incompleteness model.

6.3 Assessing model performance by synthetic experiments

6.3.1 Comparing the 'original' and the 'modified' ETAS models

Here, I compare the performance of the original ETAS model with the modified ETAS model using long-term synthetic data. The comparison focuses on how well each model handles long-term incompleteness in earthquake catalogues and the impact of this on parameter estimation.

First, I simulate a sample synthetic catalogue spanning 124 years, equivalent to the instrumental era from 1900 to 2024. In generating these data, I deliberately kept the seismicity of recent decades almost quiet, without any clear aftershock sequences. The magnitude time series of this synthetic catalogue is shown in Fig. 6.5. Since the presence of individual aftershock sequences helps reduce bias, the objective is to test the model in the absence of such specific patterns to evaluate its performance under conditions where clear triggering behaviour of aftershocks is unavailable. For the simulated catalogue, I seed one large event ($M7.1$) in the early years around 1905. This catalogue is generated with a background rate of $\mu = 0.05$, a constant magnitude threshold of $M_0 = 2.5$, and $b = 0.85$. Furthermore, the true ETAS parameters used to generate the catalogue are set as $K = 0.4$, $\alpha = 1.7$, $c = 0.05$, and $p = 1.08$.

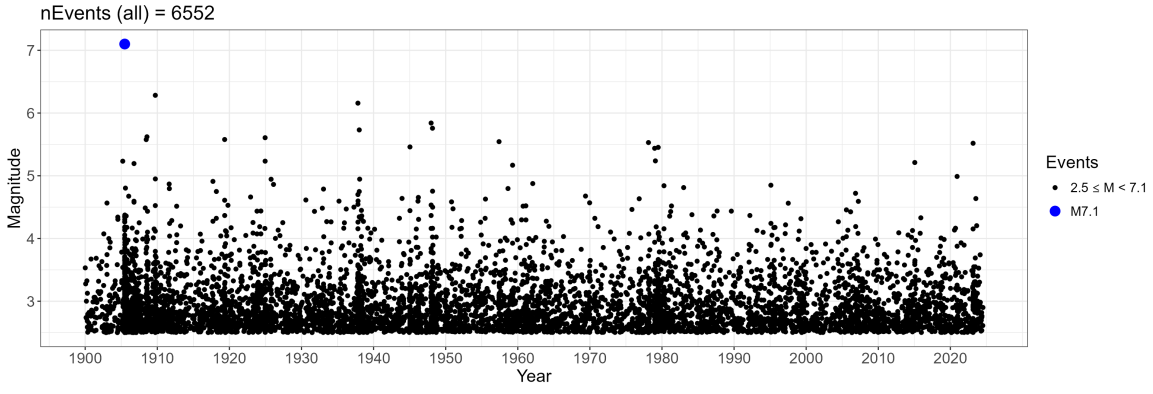

Figure 6.5: Magnitude time series for a synthetic catalogue spanning 124 years (1900-2024), containing 6552 events, with a large event (M7.1) seeded around the year 1905.

Then, I truncate this complete catalogue to create subcatalogues with long-term incompleteness. To achieve this, I define different long-term incompleteness models with various step lengths and values, including 3-step, 4-step, and 5-step models. The M_c values over the specified year ranges for each step model are presented in Table 6.1. For each scenario, I remove all events below the steps, assuming these to be unobserved (unrecorded) earthquakes. The magnitude time-series plots for each subcatalogue are shown in Fig. 6.6. In these plots, the removed and remaining earthquakes are shown by grey and black points, respectively. Above each plot, the number of all occurred earthquakes and the number of observed (recorded) earthquakes during 1900-2024 are also mentioned. The number of observed earthquakes is a small portion of all earthquakes that occurred, comprising about 5-8% of all data for different scenarios, meaning that 92-95% of events remain unrecorded. These scenarios may well reflect real-world conditions, where seismic networks have recorded only a small portion of events over the past century.

Table 6.1: Mc values over specified year ranges for different scenarios including 3-step, 4-step, and 5-step long-term incompleteness models.

Scenario	**Year Range**	**Mc**
Scenario 1: 3-step incompleteness model	1900-1940	5.0
	1940-1980	4.1
	1980-2024	3.3
Scenario 2: 4-step incompleteness model	1900-1930	5.0
	1930-1960	4.4
	1960-1990	3.8
	1990-2024	3.2
Scenario 3: 5-step incompleteness model	1900-1925	5.0
	1925-1950	4.5
	1950-1975	4.0
	1975-2000	3.5
	2000-2024	3.0

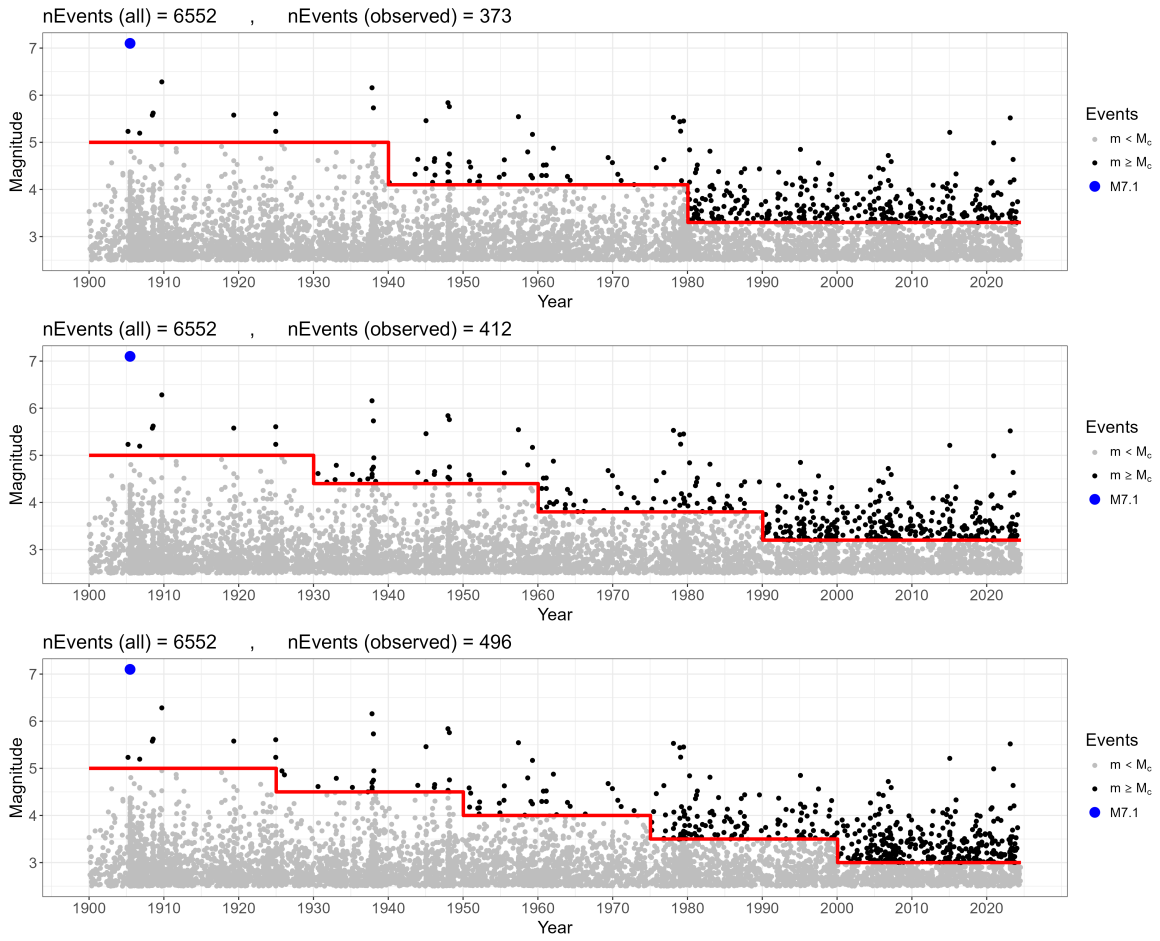

Figure 6.6: Magnitude time series plots for three different scenarios of 3-step, 4-step, and 5-step long-term incompleteness model.

Subsequently, I fit both the original and the modified ETAS models to the incomplete catalogues (with only observed earthquakes). During the model fitting and iterations, the original model struggled to converge for all three scenarios (Fig. 6.7), while the modified model converged well (Fig. 6.8).

The results of both models are presented as estimated posteriors in Fig. 6.9. For each scenario, the true ETAS parameters used to generate the synthetic catalogue are marked by vertical dashed lines. The results show that the original ETAS model, when applied to incomplete long-term data, produces parameter estimates with significant bias (red posteriors). Specifically, the original model has difficulty recovering α, μ, and p, yielding posteriors that are pushed towards the edges of the prior distributions. In contrast, the modified model, even when fitted to incomplete data, substantially reduces this bias, with its estimates (green posteriors) closely aligning with the true

parameters (vertical dashed lines). However, in Fig. 6.9, the biased red posteriors obscure the green posteriors of μ and p, making them difficult to see. To address this, I have provided a separate plot showing only the posteriors for the modified model in Fig. 6.10. As shown in this figure, in all cases, except for μ, the posterior distributions of the triggering parameters encompass the true values, keeping them within the uncertainty bounds of these distributions. For μ, the bias is minimal and still provides a close approximation to the true parameter. This indicates that despite being fed such incomplete long-term data, my inversion algorithm has performed well, enabling the modified ETAS model to accurately estimate the ETAS parameters.

Figure 6.7: Convergence plots from fitting the *'original'* ETAS model under three different scenarios of incompleteness models: 3-step (top), 4-step (middle), and 5-step (bottom). This is an example of the model struggling to converge.

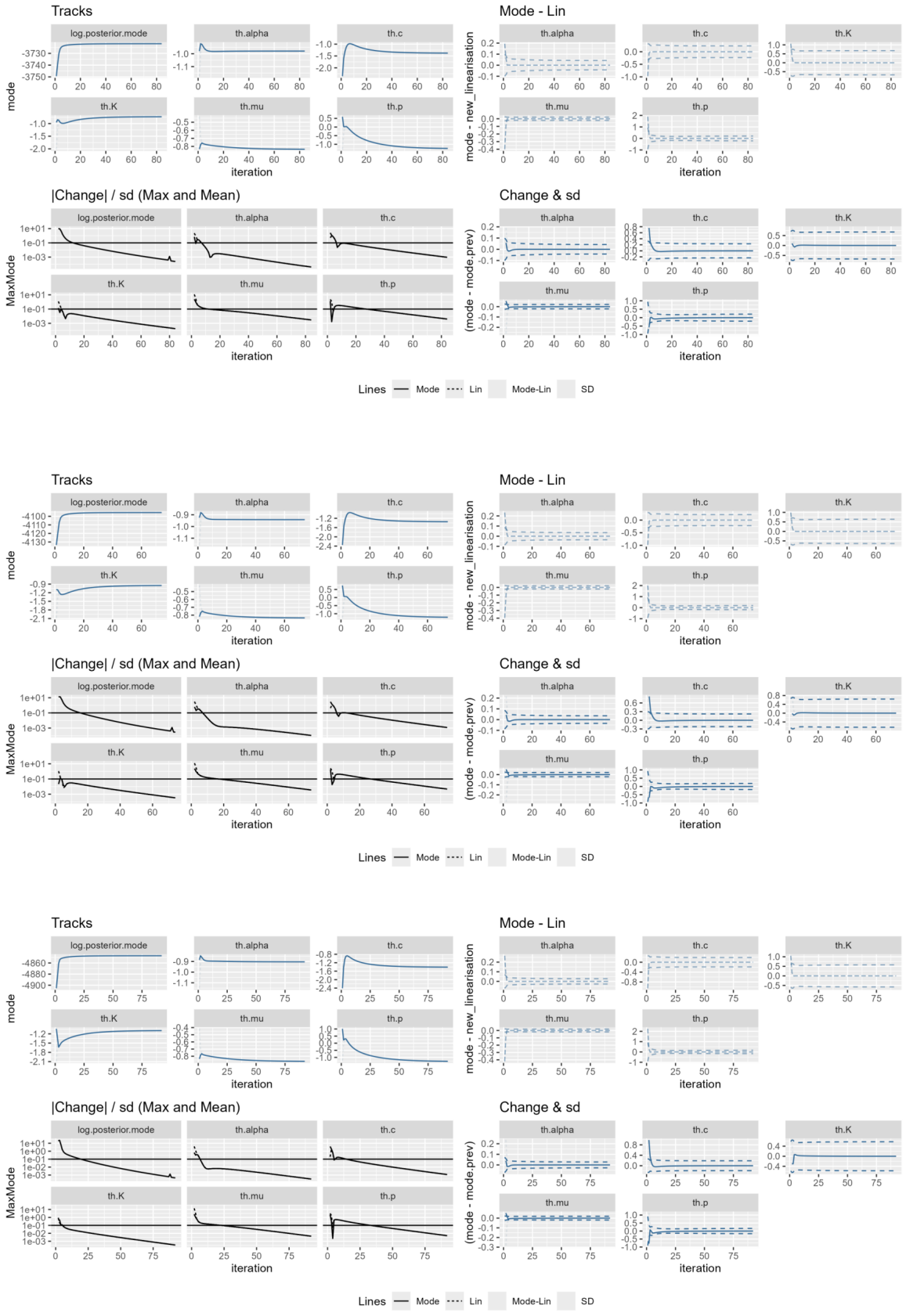

Figure 6.8: Convergence plots from fitting the *'modified'* ETAS model under three different scenarios of incompleteness models: 3-step (top), 4-step (middle), and 5-step (bottom). This is an example of the model converging well.

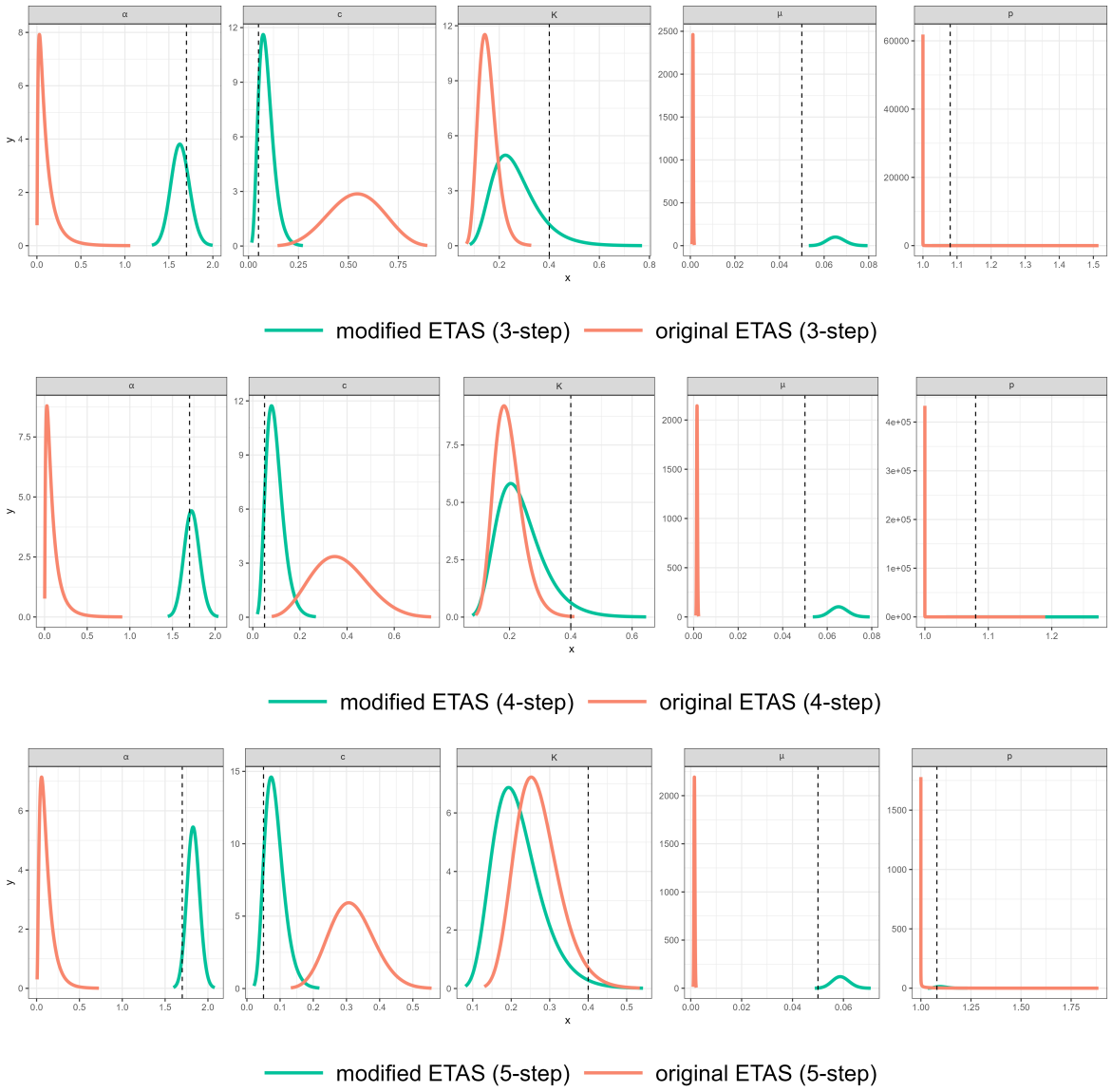

Figure 6.9: Posteriors from fitting the original (red) and the modified (green) ETAS models under three different scenarios of incompleteness models: 3-step (top), 4-step (middle), and 5-step (bottom) as shown in Fig. 6.6.

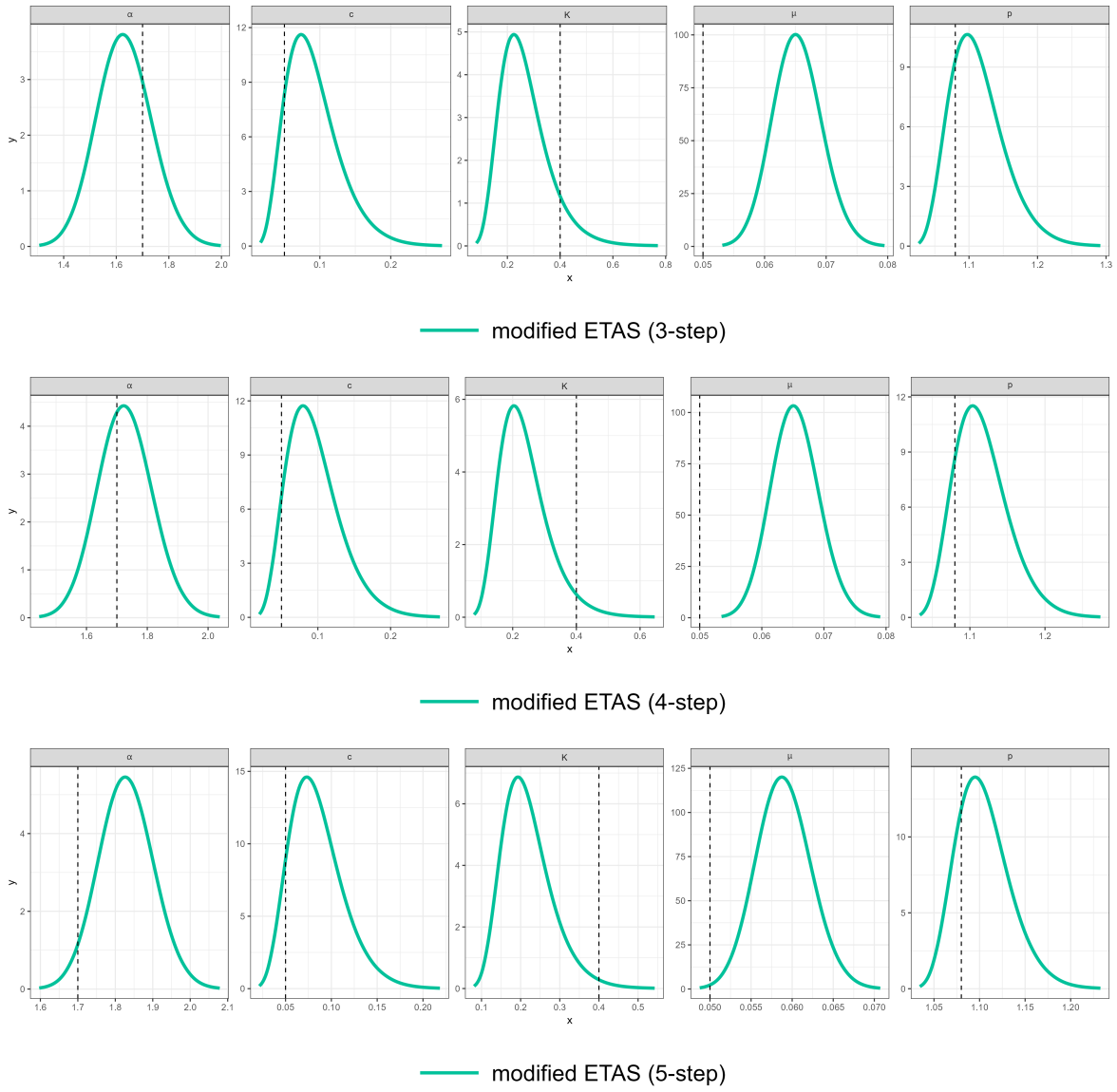

Figure 6.10: Posteriors from fitting the modified ETAS model under three different scenarios of incompleteness models: 3-step (top), 4-step (middle), and 5-step (bottom) as shown in Fig. 6.6.

6.3.2 Testing the robustness of the long-term modified ETAS model with diverse catalogue scenarios

To further test and demonstrate the efficiency and applicability of the long-term version of my modified ETAS model, I fit it to various synthetic catalogues with different seismicity patterns and validate the results by comparing the parameter estimates to the true values used to generate the synthetic catalogues. First, I simulate 10 different synthetic datasets for an instrumental period of 124 years, each with a large event of magnitude 7.1, seeded in different years (1905, 1922, 1929, 1945, 1952, 1970, 1979, 1996, 2004, 2021). These catalogues represent a range of possible scenarios during the instrumental era. The true ETAS parameters used to simulate the catalogues are fixed for all scenarios and include $\mu = 0.05$, $K = 0.4$, $\alpha = 1.7$, $c = 0.05$, and $p = 1.08$. As in the previous section, I define a 5-step incompleteness model for $M_c(t)$ and use it to truncate the catalogues.

The magnitude time series of the synthetic catalogues are shown in Figs. 6.11 and 6.12. In each case, the red line indicates the long-term incompleteness model that has five steps. The grey points below the step models represent earthquakes that were not detected because their magnitude was lower than the corresponding M_c at the time of their occurrence. Earthquakes above the step models represent the observed (recorded) events, shown as black points, with one large event ($M7.1$) highlighted in blue. The total number of events (observed + unobserved) and the number of observed events are also provided above each time series for each catalogue.

Then, I fit the modified ETAS model to evaluate the accuracy of the parameter estimates. The posteriors for the ten scenarios are shown in Fig. 6.13. The results for most synthetic catalogues demonstrate a high level of consistency and accuracy, with the estimated parameters closely aligning with the true values used to generate the catalogues, as indicated by the vertical dashed lines. When large events occur in more recent years, allowing part of their aftershock sequence to be captured above the step model, the triggering parameters are resolved more accurately. For large events in earlier years, where much of their aftershock activity is censored, the model shows slight bias but still provides acceptable parameter estimates. This highlights the importance of large events, even without their dependents, in improving parameter estimation. In

all cases, the model consistently captures large events well, demonstrating its ability to handle significant seismicity accurately. Additionally, when background events are better captured, the posteriors of μ are more precise, reflecting the model's robustness in retrieving background seismicity rates for long-term periods.

The results suggest that the long-term modified ETAS model is highly effective in accurately estimating parameters for a wide range of catalogue scenarios. The model's ability to produce reliable parameter estimates despite variations in data completeness indicates its potential applicability to real-world seismic data, where such incompleteness is common. This robust performance validates the modifications made to the original ETAS model, confirming that it can be a valuable tool specifically for long-term seismicity analysis, hazard assessment, and long-term forecasting.

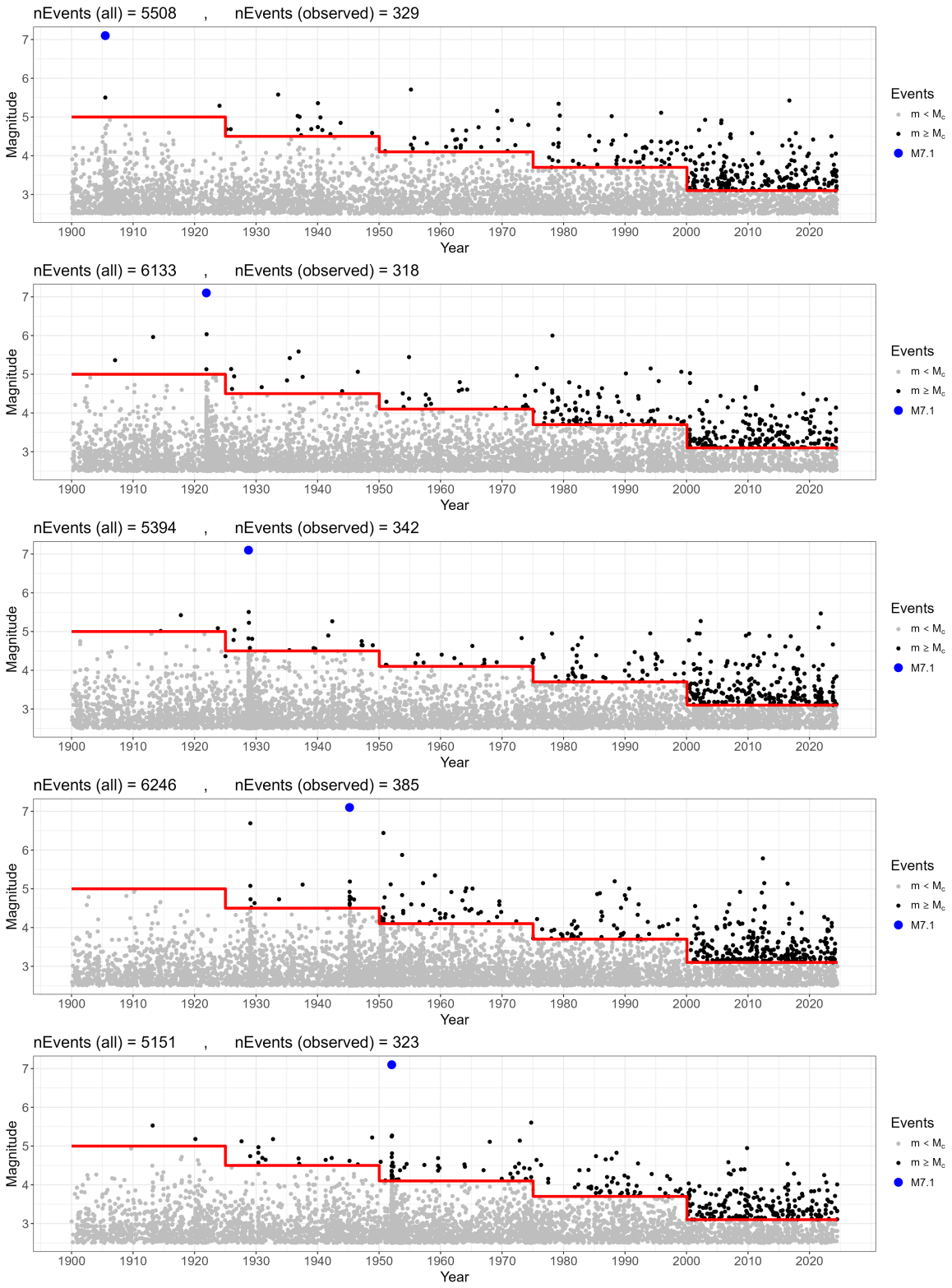

Figure 6.11: Magnitude time series of 10 different synthetic catalogues over a 124-year instrumental period. Each catalogue includes one large event with a magnitude of 7.1, seeded in different years: 1905, 1922, 1929, 1945, 1952, 1970, 1979, 1996, 2004, and 2021.

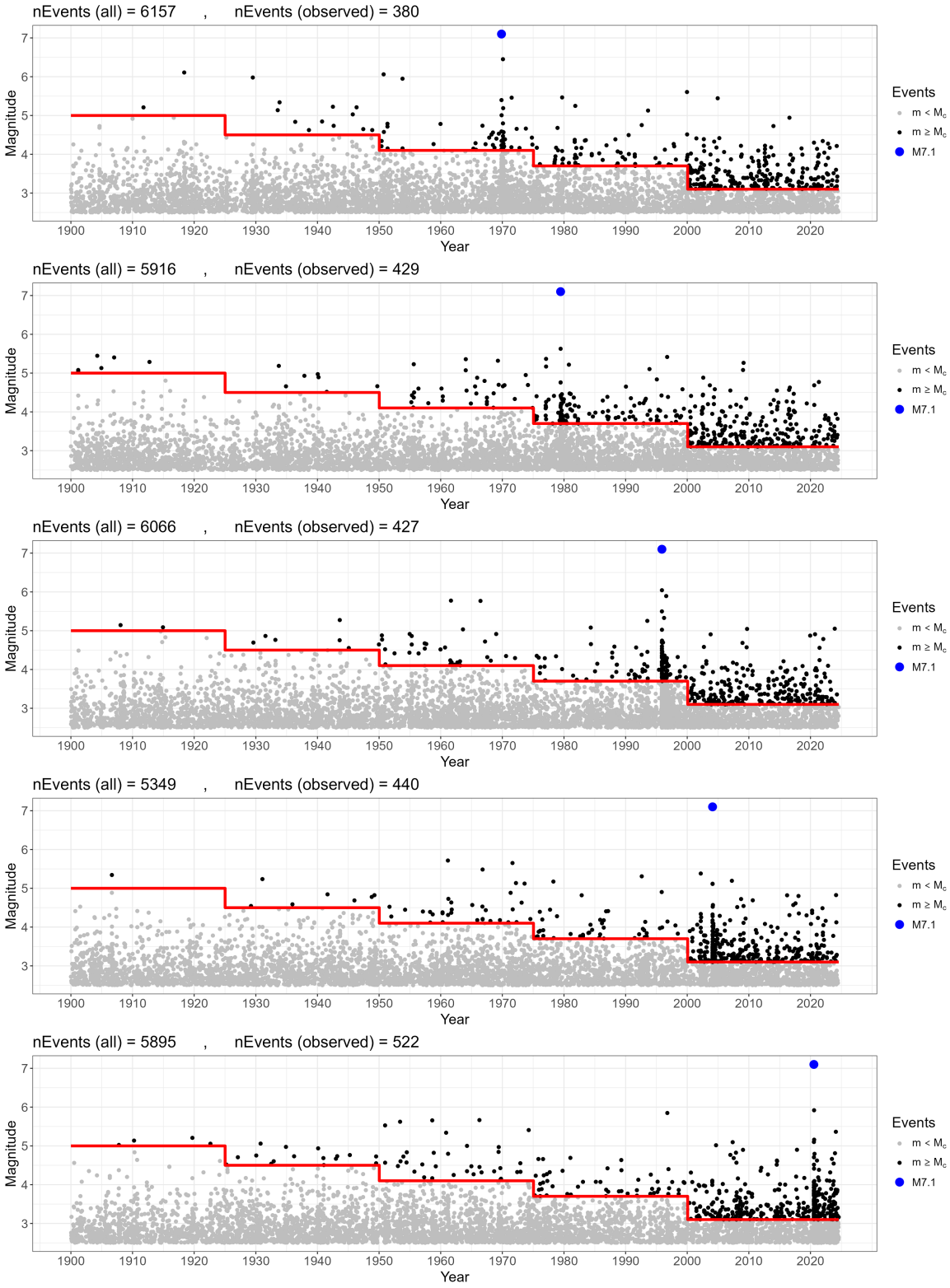

Figure 6.12: Continuation of Fig. 6.11.

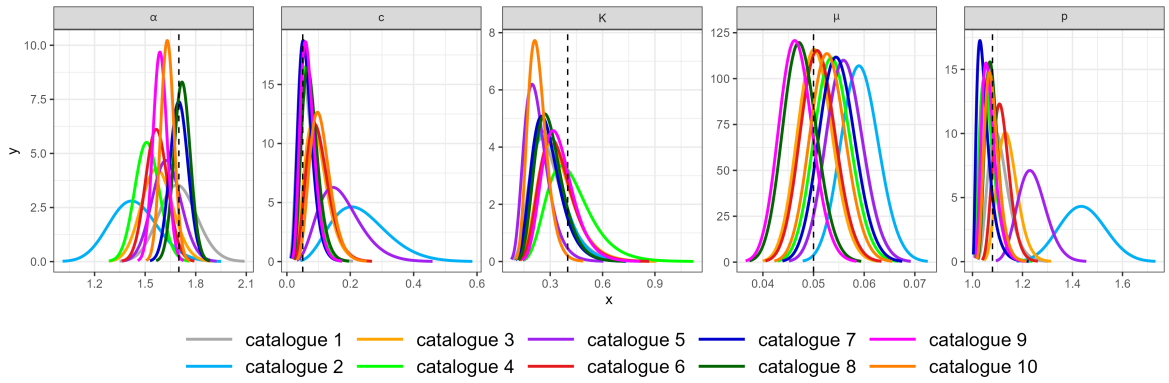

Figure 6.13: Posteriors of the modified ETAS model using 10 different synthetic catalogues over a 124-year instrumental period. Each catalogue includes one large event with a magnitude of 7.1, seeded in different years: catalogue 1 (1905), catalogue 2 (1922), catalogue 3 (1929), catalogue 4 (1945), catalogue 5 (1952), catalogue 6 (1970), catalogue 7 (1979), catalogue 8 (1996), catalogue 9 (2004), and catalogue 10 (2021).

6.3.3 Contribution level of step intervals

Although I demonstrated in the previous section that the long-term version of the modified ETAS model performs well on different sequences with a large mainshock at various time positions, one might hypothesise that the final steps with lower values of M_c have a greater influence on the ETAS inversions, since the probability of capturing aftershock behaviour in the last steps is higher than in the earlier steps, which include only a few events. To test this hypothesis, I first generate 10 random synthetic catalogues. Then, I define and apply a fixed 5-step incompleteness model to all the catalogues to extract incomplete sub-catalogues. The five steps are defined for 1900-1925, 1925-1950, 1950-1975, 1975-2000, and 2000-2024, with M_c values of 5.5, 5, 4.6, 4, and 3, respectively. All catalogues include one or two large events seeded within the first step, and seismicity is kept relatively quiet in the subsequent steps.

Next, I divide the entire modelling domain and select data with different levels of completeness to create six scenarios for each catalogue: 1900-2024 (all 5 steps), 1925-2024 (the last 4 steps), 1950-2024 (the last 3 steps), 1975-2024 (the last 2 steps), 2000-2024 (the last step), and 1900-2000 (the first 4 steps). This approach allows me to assess which level of incompleteness has the greatest impact on the modelling and

whether the final step alone can accurately retrieve the ETAS parameters, or if data from other levels are also essential. Figures 6.14 to 6.23 show the magnitude plots for the six scenarios, with black points representing the data considered in each scenario. In addition, the total number of events in the catalogues and the number of events observed in each selected period are indicated above each plot.

Then, I fit the modified ETAS model to the 6 sliced datasets for each of the 10 synthetic catalogues. The results of the posteriors are shown below the magnitude plots in Figs. 6.14 to 6.23. The results reveal some interesting patterns that are consistent in all cases. Firstly, the posteriors of the ETAS model fitted only to the data from the last step (2000-2024), which has the lowest M_c, show considerable bias. Interestingly, this period has the highest number of remaining data after excluding the other steps in all the examples. For example, after applying the step model to the entire dataset from 1900 to 2024 (8061 events) in Fig. 6.14, 401 events remain, of which 347 belong to 2000-2024, accounting for about 87% of the data. This pattern is consistent for all the 10 examples. However, the posteriors for this period (in orange) show the highest bias, particularly when retrieving the background (μ) and temporal decay (p) parameters. In all cases, the posteriors for μ show underestimations compared to the true value, and the posteriors for p are not correctly retrieved, skewed towards the left end of the specified prior distribution. This indicates the importance and role of having large events in the data for accurately retrieving the ETAS parameters.

The next model that exhibits some degree of bias is the one fitted to data from 1900-2000, which encompasses the highest levels of M_c and the lowest amount of data. For example, of the 401 events in 1900-2024 in Fig. 6.14, only 54 events are located within 1900-2000. As expected, the posteriors for this period are wider (due to fewer data) with higher uncertainty than for the other periods. This also indicates the importance and role of having smaller events in the data for accurately retrieving the ETAS parameters.

All other scenarios (1900-2024, 1925-2024, 1950-2024, and 1975-2024), which include large events and a fair number of smaller events, show proper estimates. This suggests that a balanced inclusion of both large and small events is crucial for accurate parameter estimation in the long-term ETAS modelling.

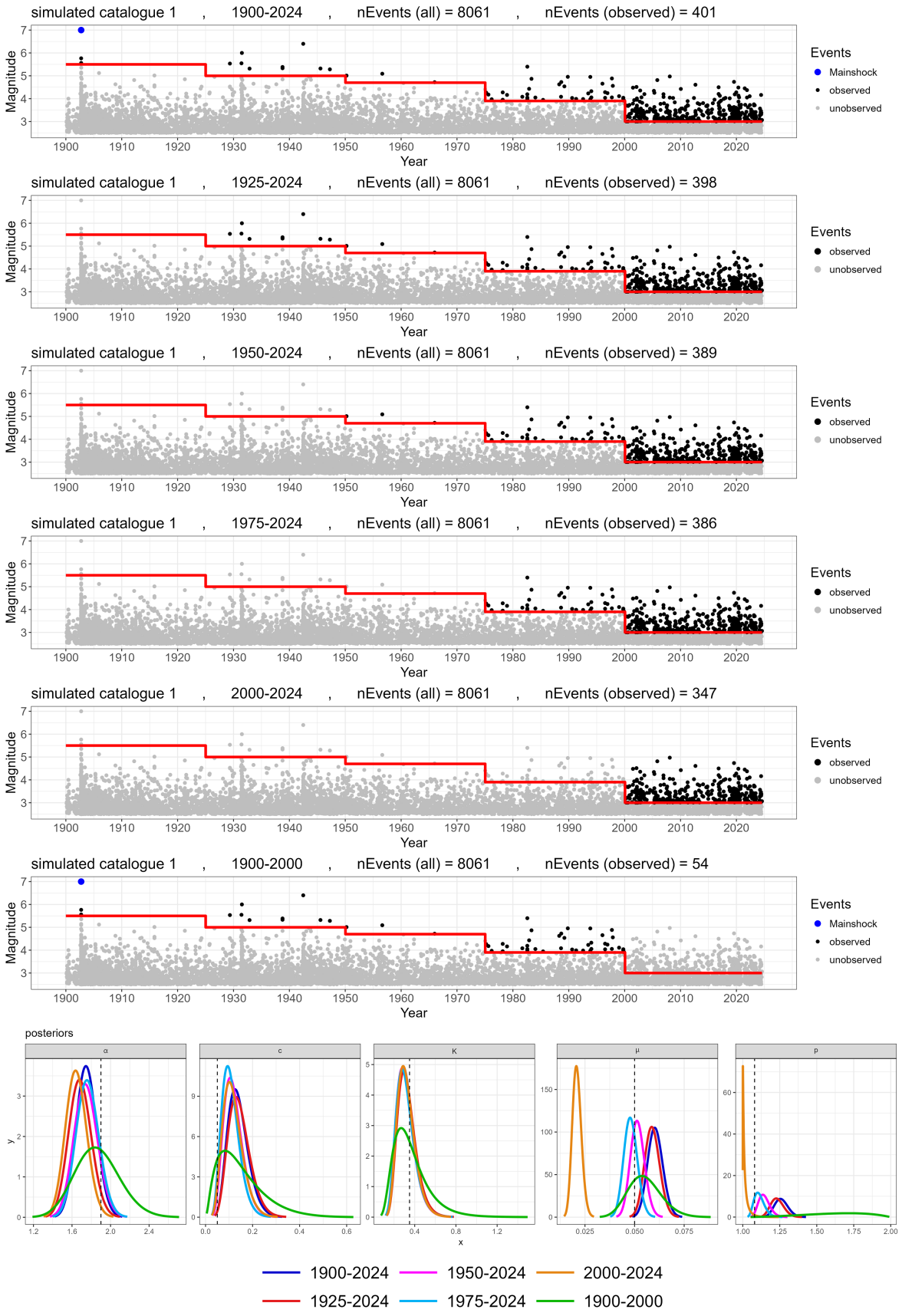

Figure 6.14: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 1: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

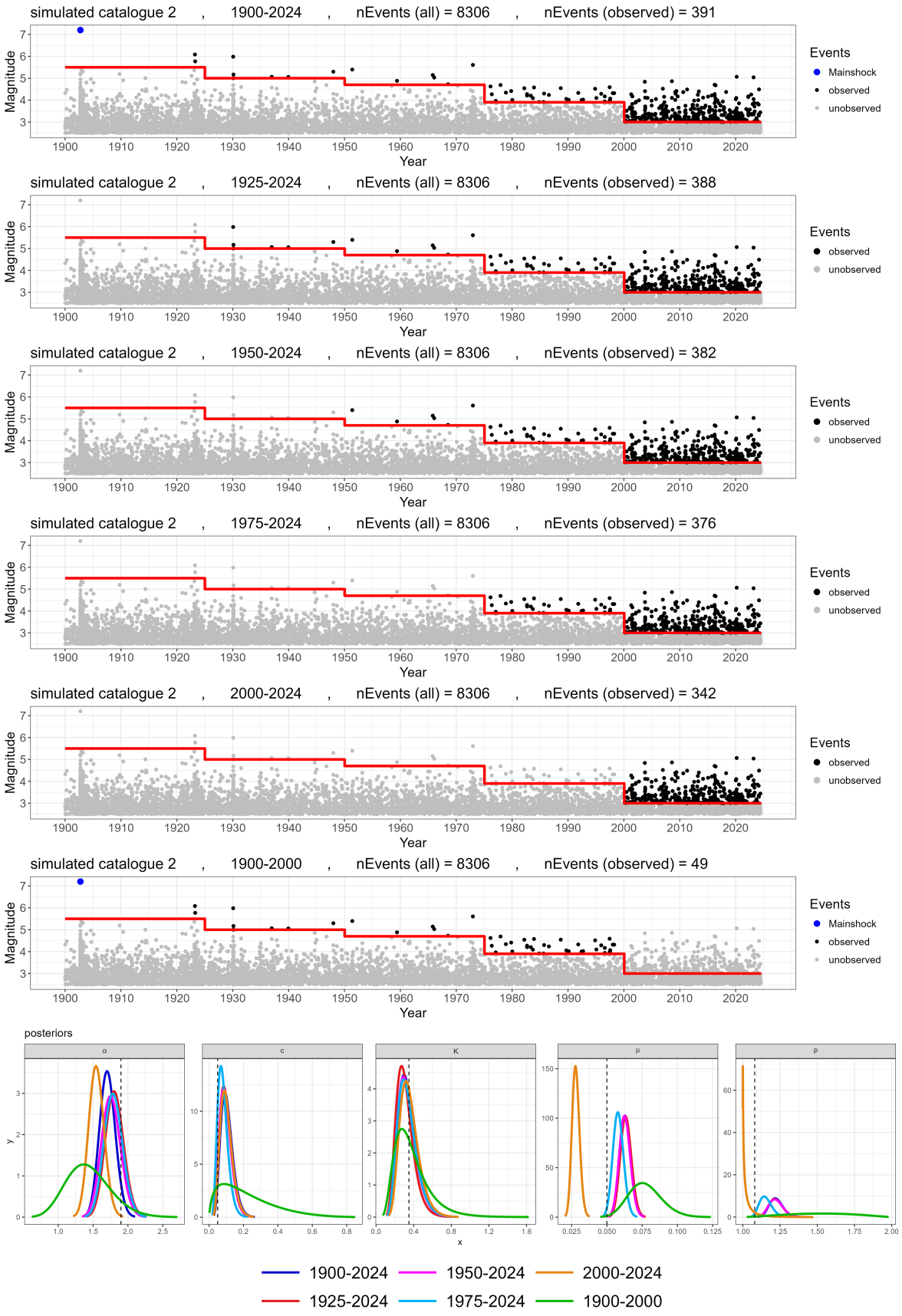

Figure 6.15: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 2: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

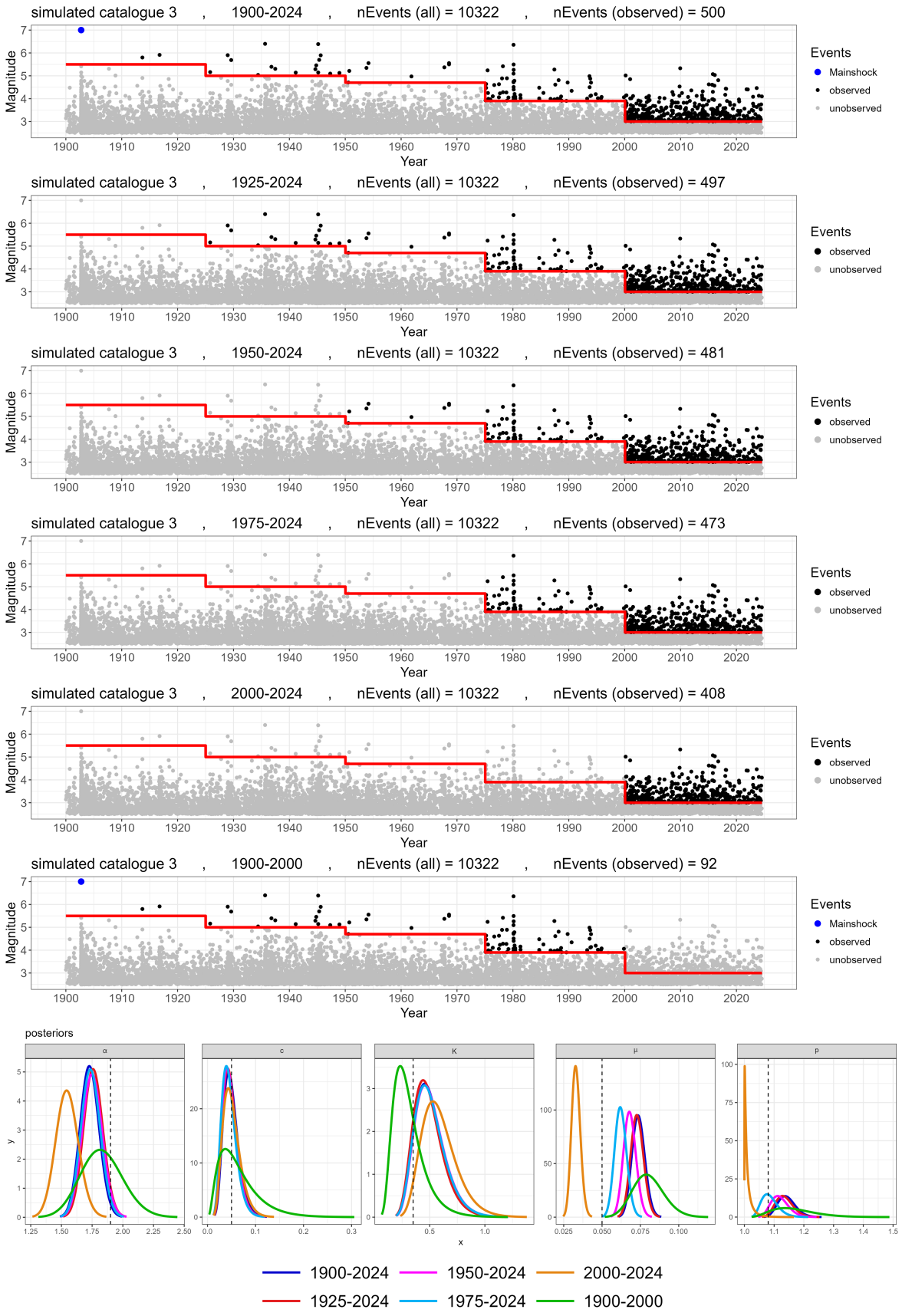

Figure 6.16: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 3: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

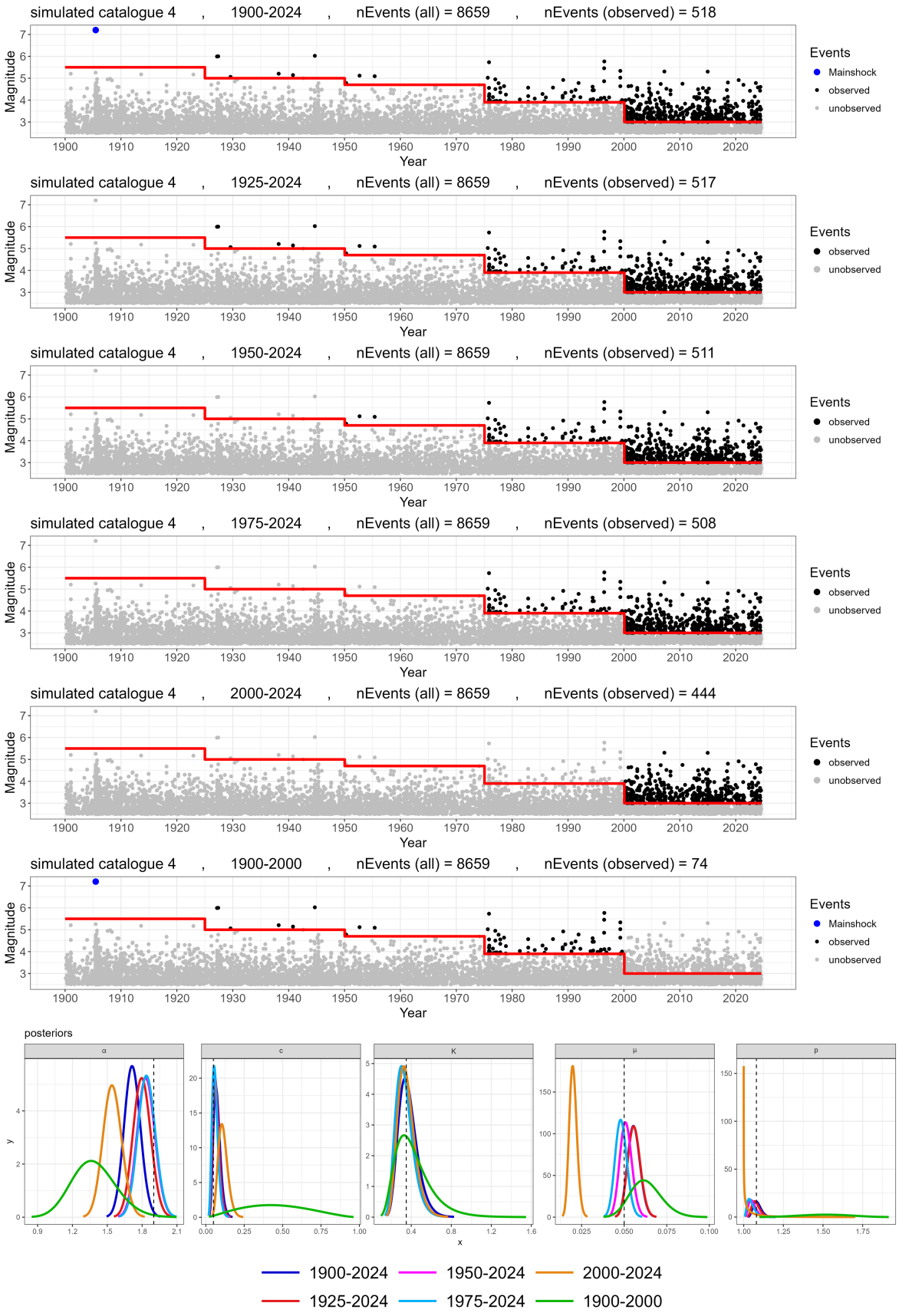

Figure 6.17: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 4: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

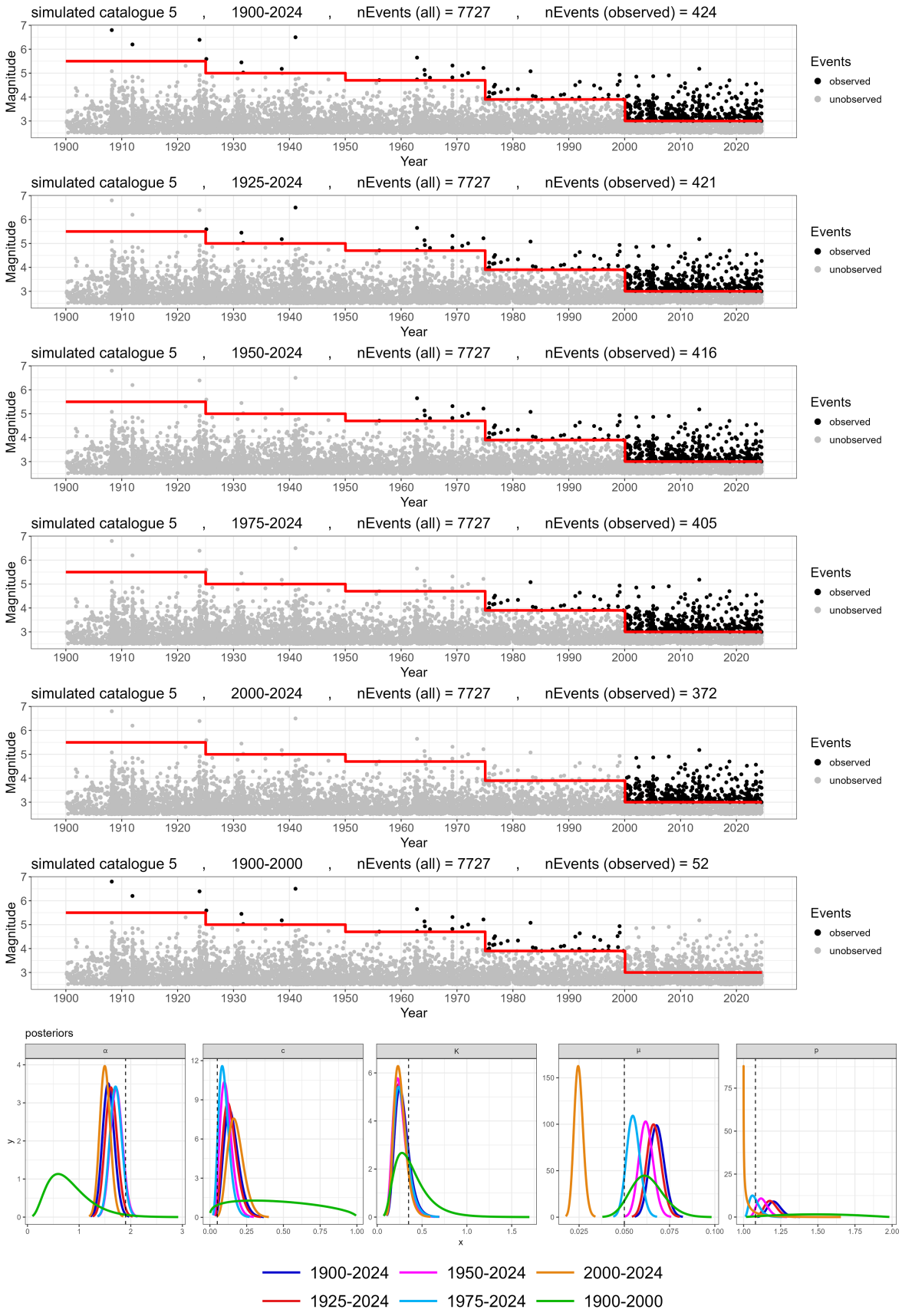

Figure 6.18: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 5: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

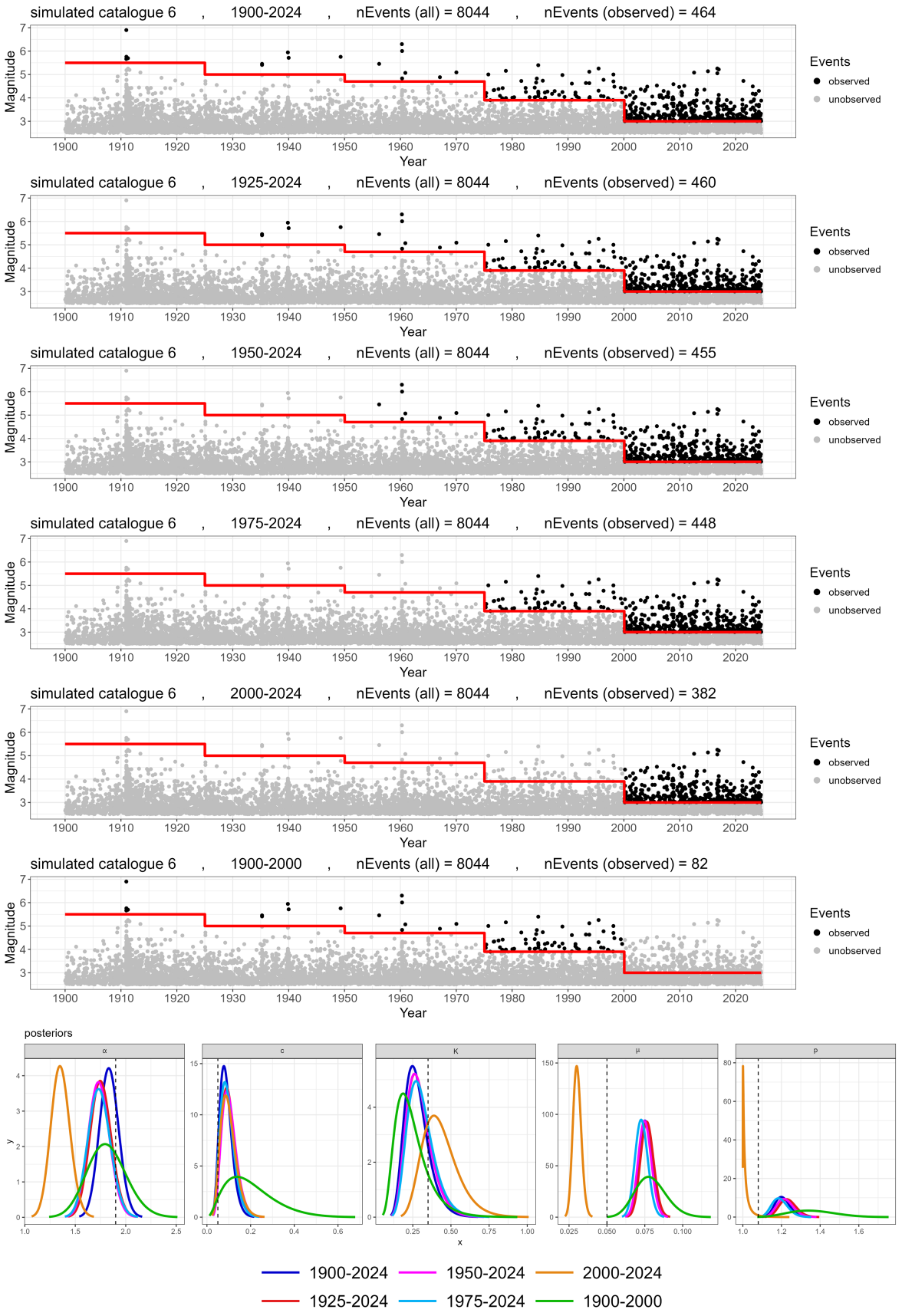

Figure 6.19: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 6: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

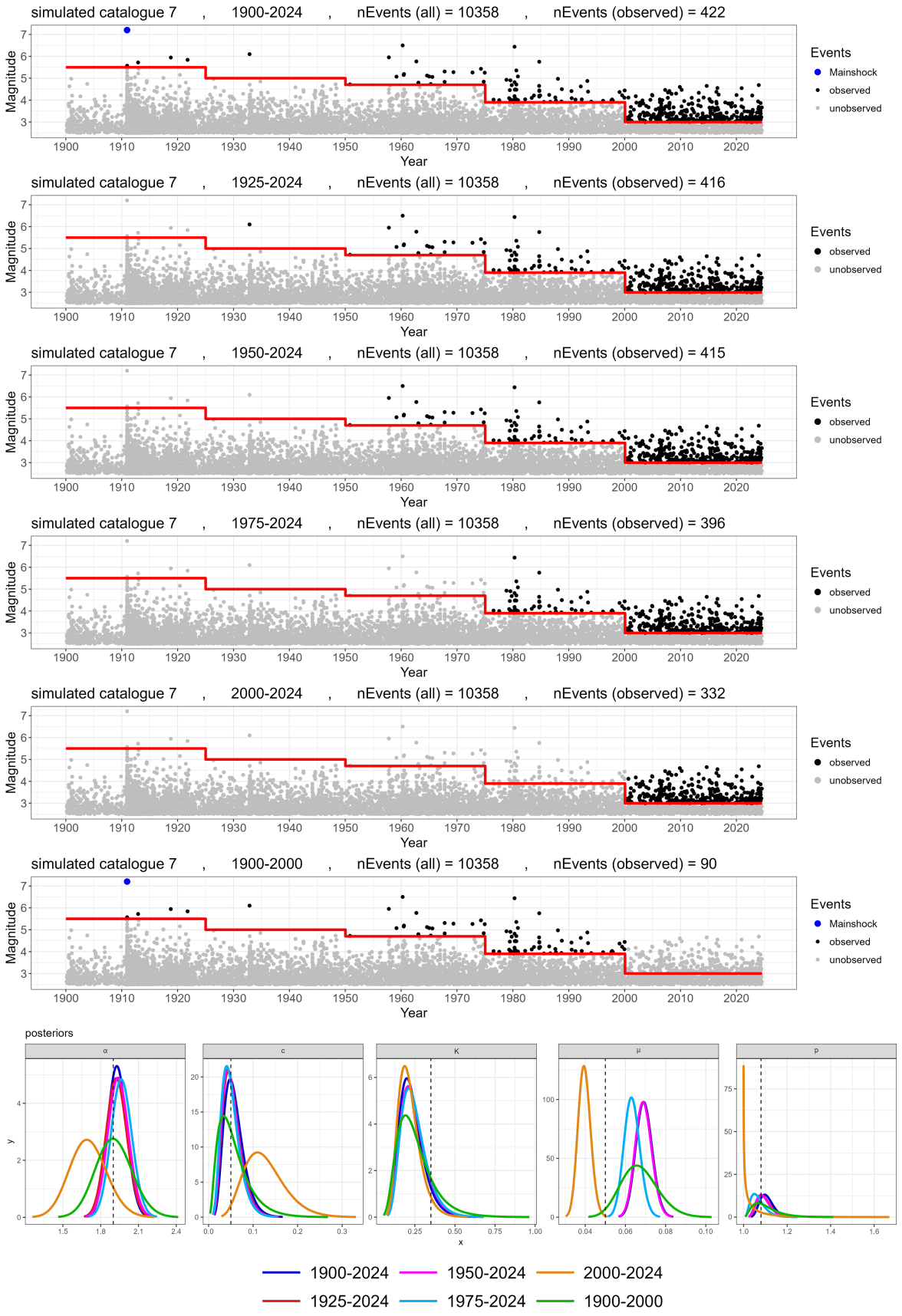

Figure 6.20: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 7: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

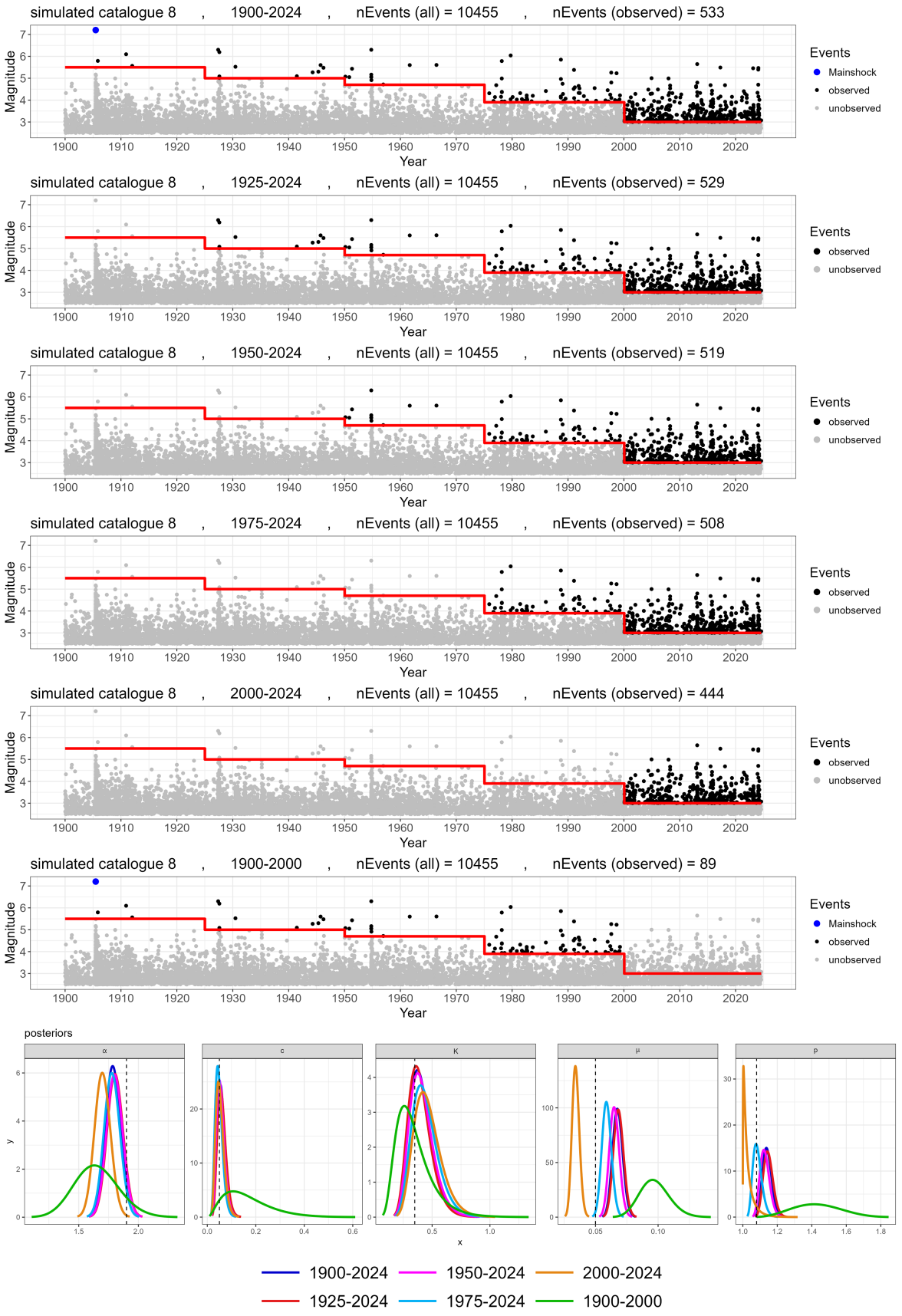

Figure 6.21: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 8: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

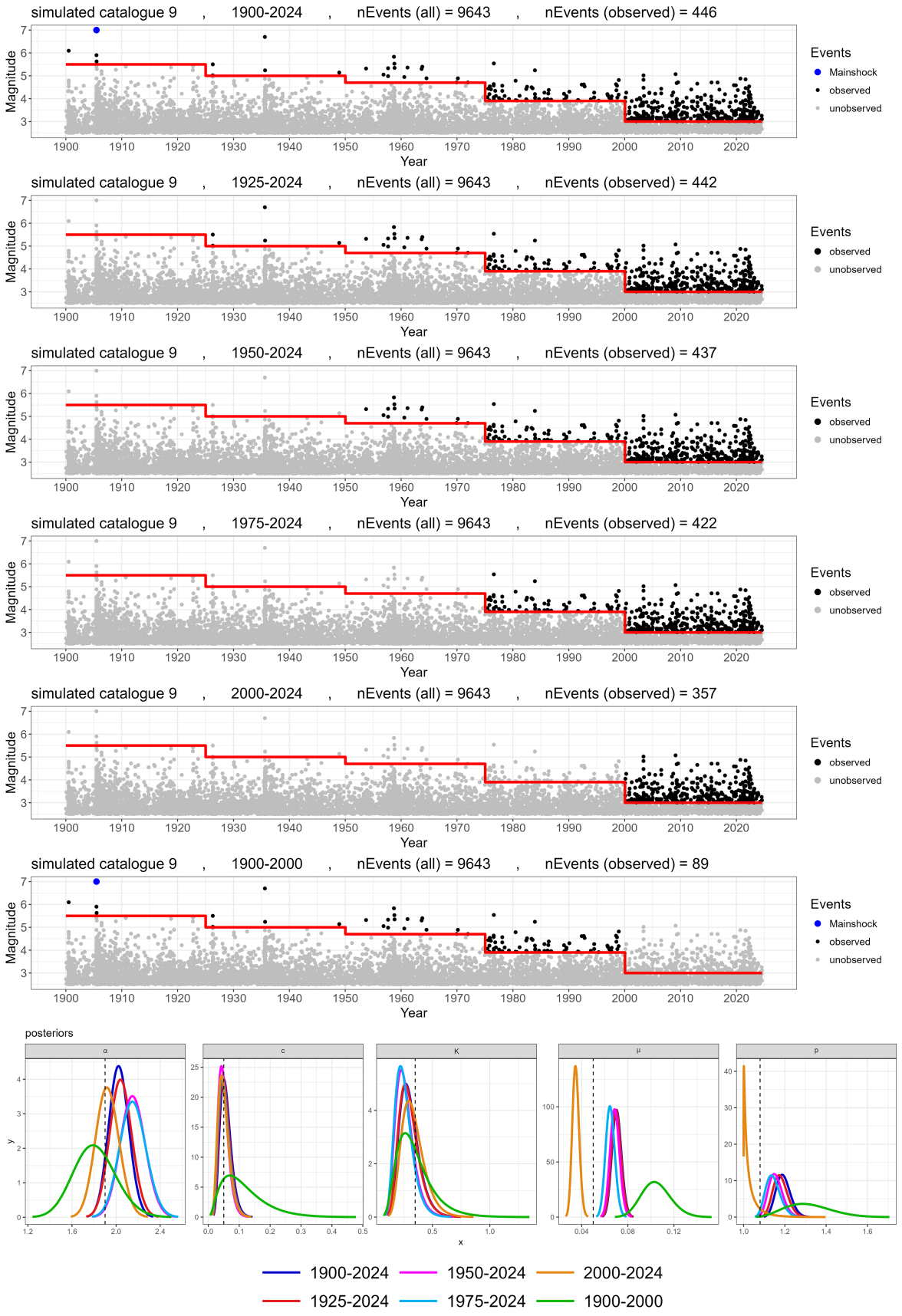

Figure 6.22: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 9: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

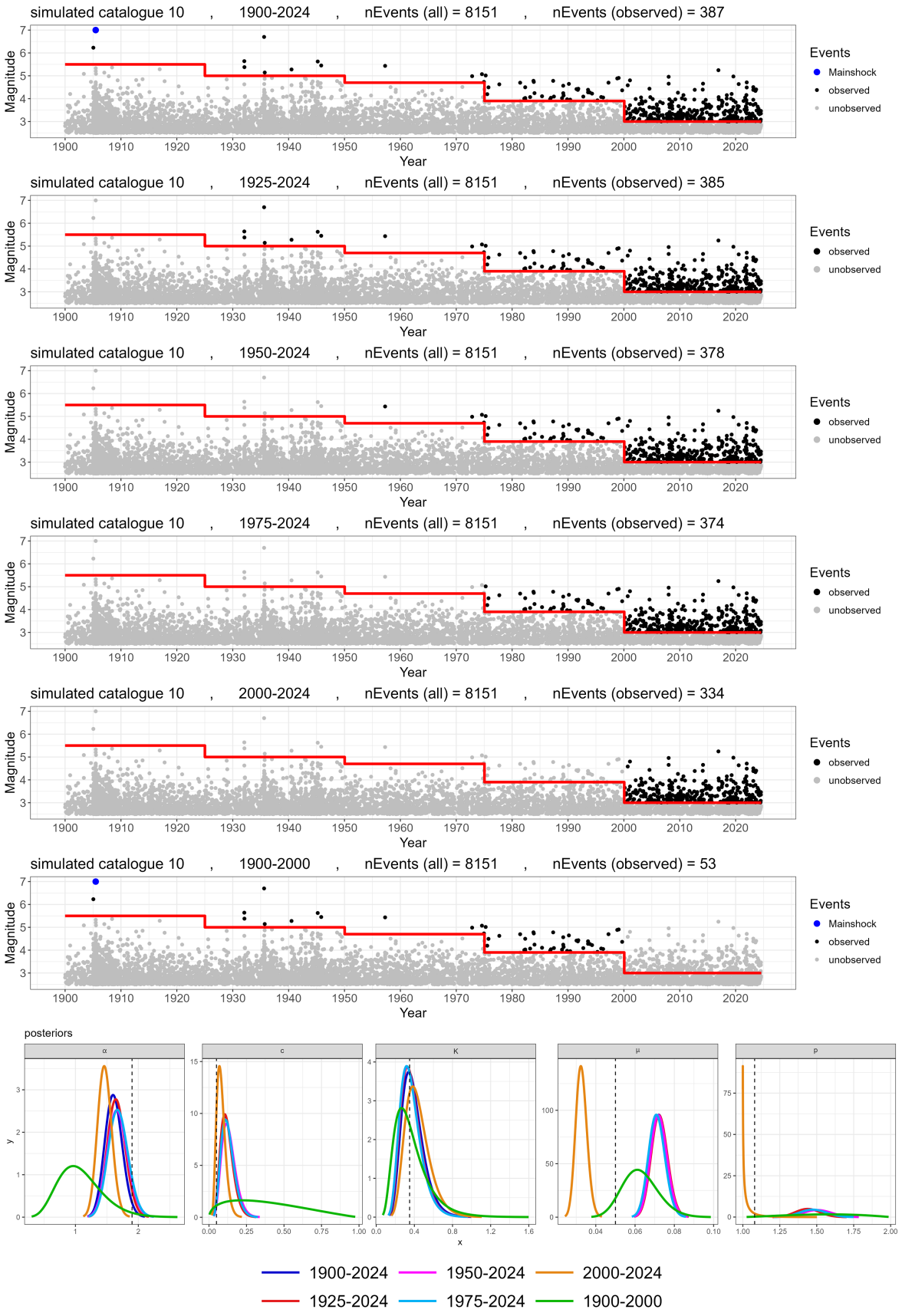

Figure 6.23: Magnitude plots and posterior distributions for six scenarios of data selected from the synthetic catalogue 10: 1900-2024 (all 5 steps), 1925-2024 (last 4 steps), 1950-2024 (last 3 steps), 1975-2024 (last 2 steps), 2000-2024 (last step), and 1900-2000 (first 4 steps).

6.4 Real tests: short- and long-term modified ETAS inversions for central Italy

With the two developed versions of the modified ETAS model, which are capable of addressing both short-term and long-term incompleteness in earthquake catalogues, the objective is to test real datasets to examine the similarities and differences between short-term and long-term modelling. This will allow me to determine the consistency or variability of seismicity behaviour in different time scales.

To achieve this, I analyse the seismicity of central Italy as a pilot region. Generally, Italy does not have records of large events ($M \geq 7$), and there are only about ten earthquakes with magnitudes ≥ 6. However, there is a good record of instrumental seismicity mainly concentrated in the central area of the country. In this area, there have been three significant sequences in recent decades, including the 26 September 1997 Colfiorito (M_w6), 6 April 2009 L'Aquila ($M_w6.3$), and 24 August 2016 Amatrice sequences ($M_w6.2$). The instrumental seismicity of Italy along with the location of these three recent major sequences is illustrated in Fig. 6.24.

For these known sequences, I aim to: (1) test the similarity or differences between the ETAS parameters of the sequences in terms of spatial consistency, and (2) determine whether the ETAS parameters calculated from the major aftershock sequences are similar to or different from those derived using long-term seismicity data prior to these sequences. Given the incompleteness in both short- and long-term periods, I can now test this hypothesis with my newly modified ETAS models, which have made this analysis possible.

I begin by selecting the appropriate spatial and temporal data for each sequence. I first apply the window-based declustering algorithm of Gardner and Knopoff (1974), which defines the temporal duration and spatial extent of an aftershock sequence based on the mainshock magnitude. For the temporal window, I deliberately consider a 1-year run-in period before each mainshock (for the reasons I discussed in Sections 5.3 and 5.7). Based on the mainshock magnitudes, I calculated temporal windows of 497, 673, and 610 days for the aftershocks following the 1997 Colfiorito, 2009 L'Aquila, and 2016 Amatrice mainshocks, respectively. The spatial extent for each case is a circular

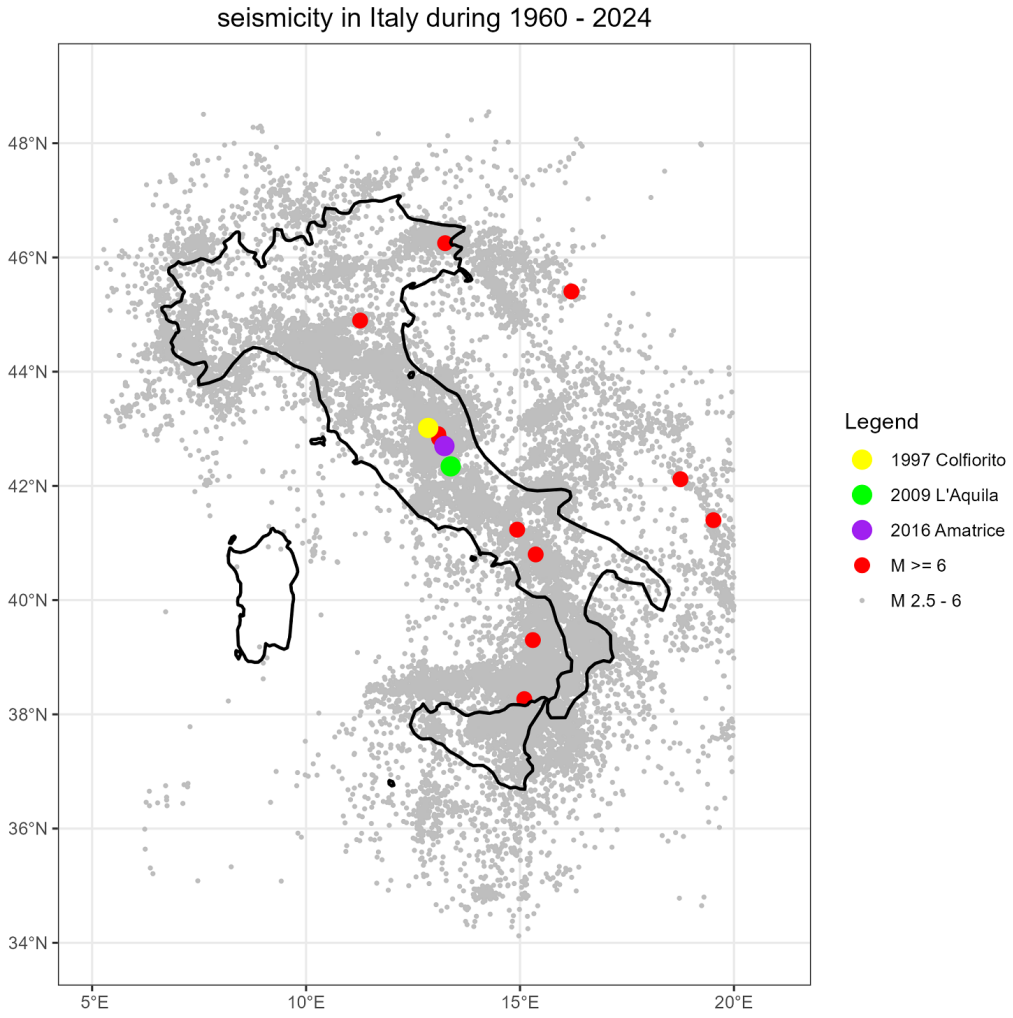

Figure 6.24: Instrumental seismicity in Italy from 1960 to 2024. Earthquakes with magnitudes between 2.5 and 6 are shown in grey, while those with magnitudes greater than 6 are depicted in red. The three selected sequences in Central Italy, including the 1997 Colfiorito, 2009 L'Aquila, and 2016 Amatrice events, are highlighted in yellow, green, and purple, respectively.

area with a radius of 53.6, 58, and 56.5 km, respectively. These details are illustrated in Fig. 6.25, where the sequences (each with a 1-year run-in period and corresponding spatio-temporal windows) are shown in different colours: 1997 Colfiorito in green, 2009 L'Aquila in purple, and 2016 Amatrice in orange. The black points represent the instrumental seismicity from 1900 up to the 1-year run-in period.

Earthquakes typically tend to cluster anisotropically along fault lines. By a visual inspection in Fig. 6.25, we can see a NW-SE trend in tectonic structures and aftershock sequences in the Central Italy region. Therefore, the choice of circular spatial windows

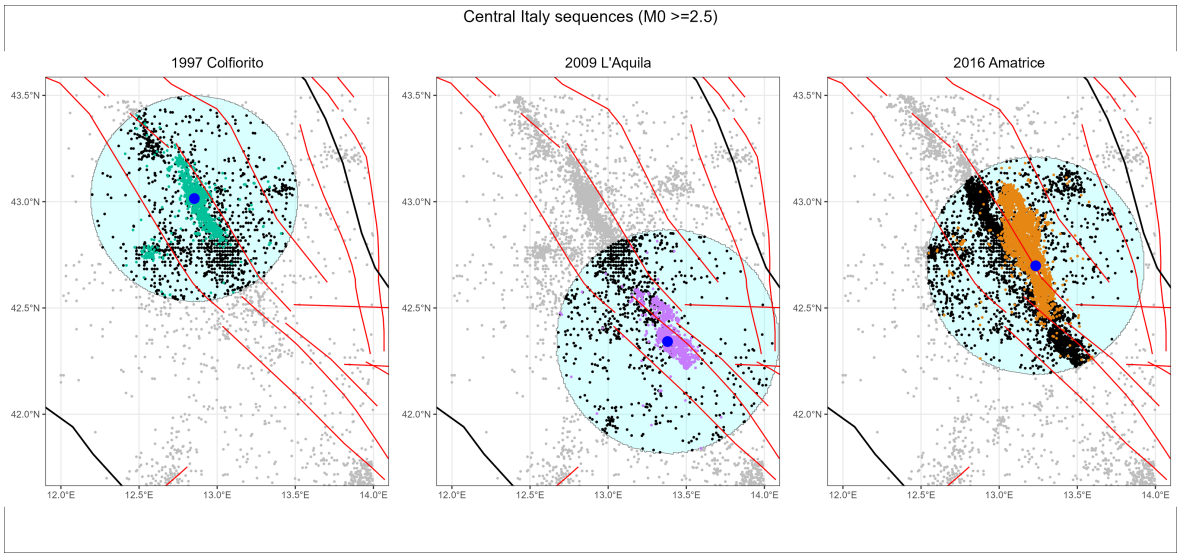

Figure 6.25: The three selected sequences in central Italy, including the 1997 Colfiorito, 2009 L'Aquila, and 2016 Amatrice events, with Gardner-Knopoff spatio-temporal windows.

is not ideal and might introduce bias in the modelling for two reasons: (1) During each sequence, events that are spatially distant from the main cluster but fall within the declustering algorithm's spatial window may be included, despite not being inherently related to the sequences. (2) Considering a long-term dataset (from 1900 to the start of the 1-year run-in before the mainshocks), and given that the three sequences occurred in close proximity, for the last mainshock (2016 Amatrice), we can clearly see numerous events from the 1997 Colfiorito and 2009 L'Aquila sequences included within the Amatrice spatial window. Thus, the influence of these past sequences should be excluded from the triggering process of subsequent sequences. Therefore, I delineate polygons (Fig. 6.26) that better fit the spatial cluster of each sequence, removing the contributions of unwanted distributions from previous clusters.

After selecting appropriate spatial and temporal windows for each sequence and their long-term past seismicity, I need to determine the short-term and long-term incompleteness models in the data. To this end, I detect the short-term incompleteness in the initial aftershocks of each sequence using the exponential model of Helmstetter et al. (2006) as described in Eq. 4.10 as well as the long-term incompleteness in instrumental past seismicity using a step model (as described in Eq. 6.1). For short-term incompleteness, we can fit the proposed exponential model to the real datasets and select the best-fit values for the parameters G and H. However, a straightforward ap-

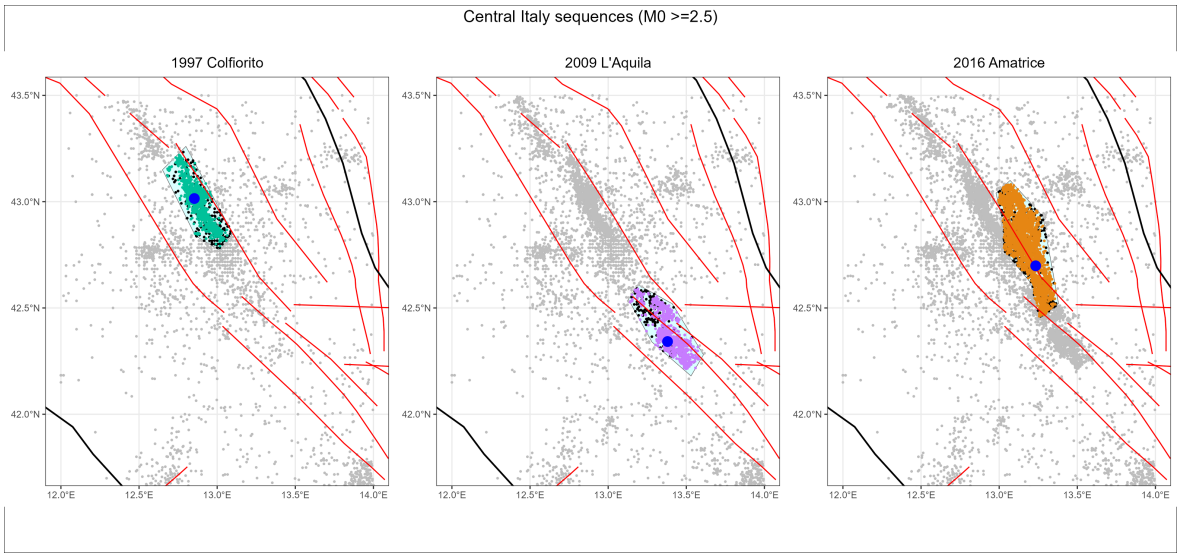

Figure 6.26: The three selected sequences in Central Italy, including the 1997 Colfiorito, 2009 L'Aquila, and 2016 Amatrice events with adjusted spatial windows.

proach is to perform exploratory data analysis and through visualisations, since visual inspections in real cases help to better determine the boundaries of completeness and incompleteness (similar to what I explained for real sequences in Section 5.7). The selected models are shown in Fig. 6.27 for the three sequences.

For the long-term model, I need to define appropriate step models. In Fig. 6.28, I have plotted the three sequences in different colours (1997 Colfiorito in green, 2009 L'Aquila in purple, and 2016 Amatrice in orange), and the past seismicity for each case from 1960 up to their 1-year run-in periods is shown with black points. Schorlemmer et al. (2010) suggests that the Italian National Seismic Network is complete at $M = 2.9$ during 2005-2008 for most of the country. The plot of the data in central Italy in Fig. 6.28 reveals three possible improvements in the network and the long-term M_c values. Consequently, I divide the instrumental seismicity into three steps with $M_c = 4$ for 1960-1979, $M_c = 2.9$ for 1979-2009, and $M_c = 2.5$ for 2009-2019. This is an approximation, and one might suggest 1-2 more steps and slightly higher or lower values for M_c. However, as I demonstrated in Section 6.3.1 for synthetic data, as long as there is no significant difference in the defined step model, the estimates from the modified ETAS model remain robust.

When examining the long-term data in Fig. 6.28, we observe a significant sequence around 1980 in both the 1997 Colfiorito and 2016 Amatrice plots. According to the

catalogue, a large event with a magnitude of 5.8 occurred on 19 September 1979, and some of its aftershocks are spatially located within the polygons of both the 1997 Colfiorito and 2016 Amatrice events (Fig. 6.29). This is important because part of the 1979 sequence falls within the spatial windows of the two later events, indicating that the triggering properties of this sequence are partially captured. Incorporating such incomplete spatial data with partial triggering behaviour can significantly bias the posteriors of the 1997 Colfiorito and 2016 Amatrice sequences. To avoid this, I remove data from the 1979 mainshock and the following 422 days (based on the Gardner-Knopoff temporal windows) that are located within the 1997 Colfiorito and 2016 Amatrice polygons. The removed data are shown in blue in Fig. 6.30, and the magnitude time series after the removal of the 1979 data are presented in Fig. 6.31.

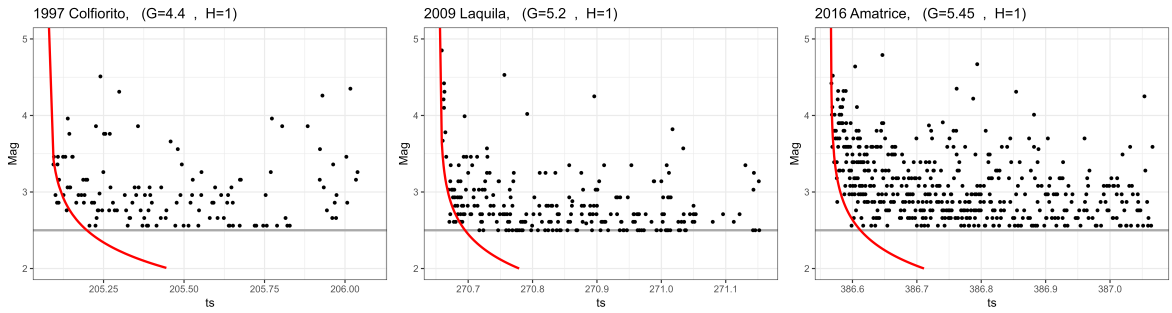

Figure 6.27: The selected parameters of the short-term incompleteness model for the 1997 Colfiorito, 2009 L'Aquila, and 2016 Amatrice sequences. The red curves represent the best-fitted Helmstetter model to the data over the short-term incompleteness interval. The grey lines indicate the constant magnitude of completeness, representing the threshold above which all events are assumed to be detected.

Then, I fit the short- and long-term modified ETAS models to the three sequences and the past seismicity data for each case. The results are presented as posterior distributions in Figs. 6.32, 6.33, and 6.34 for the three cases. For the 1997 Colfiorito and 2916 Amatrice sequences, I fit the long-term model once with (in black posteriors) and once without (in blue posteriors) the 1979 data. As observed, the existence of the 1997 data can slightly shift the posteriors in both cases. The modes of the posteriors are also provided in Table 6.2.

This study is the first to analyse both the short-term and long-term behaviour of seismicity using the censorship factor, accounting for the existing incompleteness over

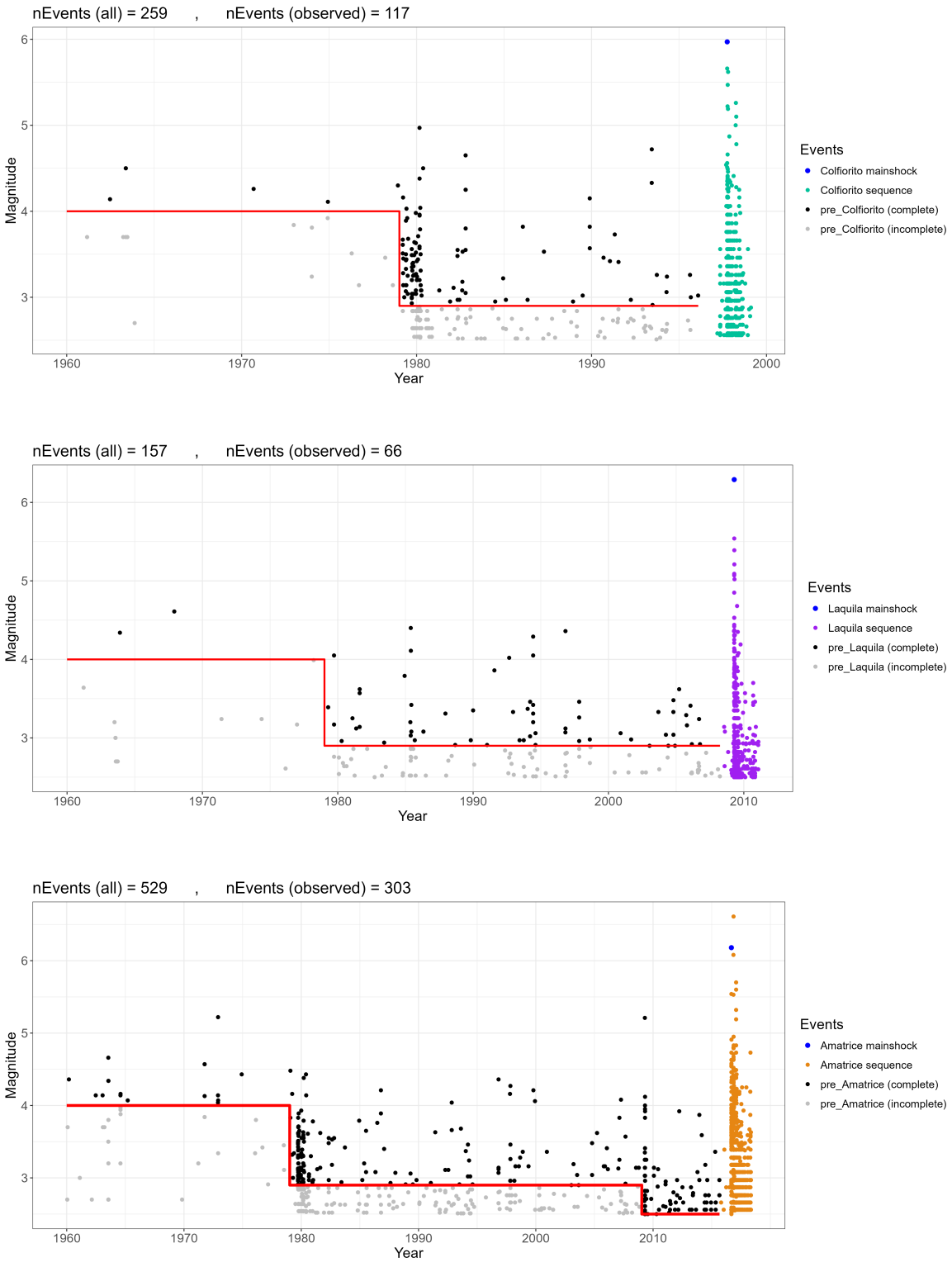

Figure 6.28: Seismic sequences of the 1997 Colfiorito (green, top panel), 2009 L'Aquila (purple, middle panel), and 2016 Amatrice (orange, bottom panel) and their past instrumental seismicity (black points). The selected step models are also illustrated with red lines.

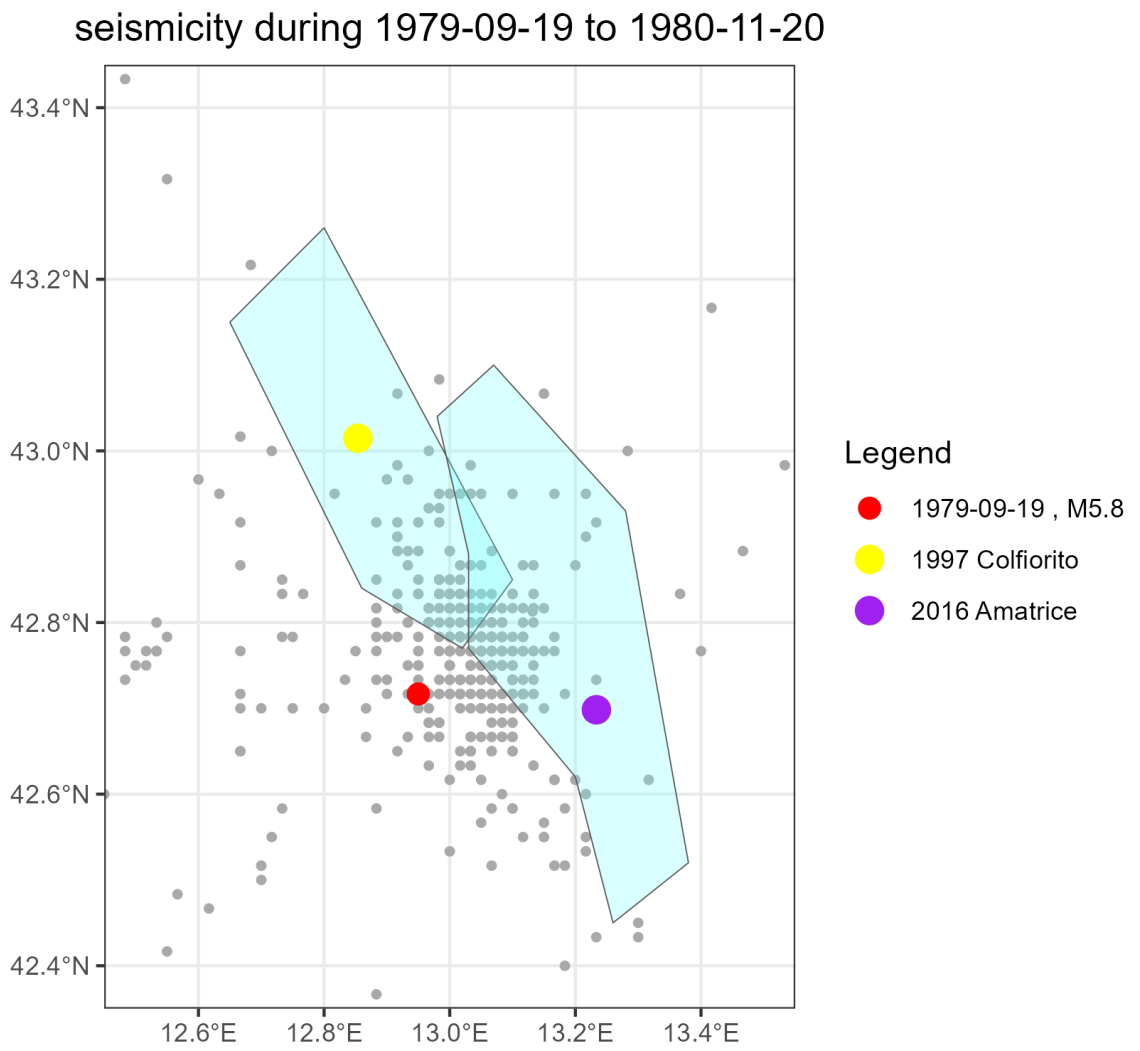

Figure 6.29: Spatial distribution of seismicity from 19/09/1979 to 20/11/1980, analysed to investigate the aftershock sequence of the M5.8 event in 1979. Portions of the aftershock sequence fall within the 1997 Colfiorito and 2016 Amatrice polygons. These need to be removed to avoid biased contributions of partial triggering behaviour from the 1979 sequence in the analyses of the 1997 Colfiorito and 2016 Amatrice events.

different time scales. In discussing the results, there are generally two hypotheses about seismicity behaviour over short- and long-term scales and its regional consistency or variations. One hypothesis posits that seismicity should exhibit stable spatio-temporal behaviour over different times, as tectonic cycles are assumed to follow regular patterns despite their stochastic nature. Therefore, we would expect to see almost similar ETAS estimates not only for the three closely located aftershock sequences in the region but also over extended periods. The other hypothesis suggests that the background and

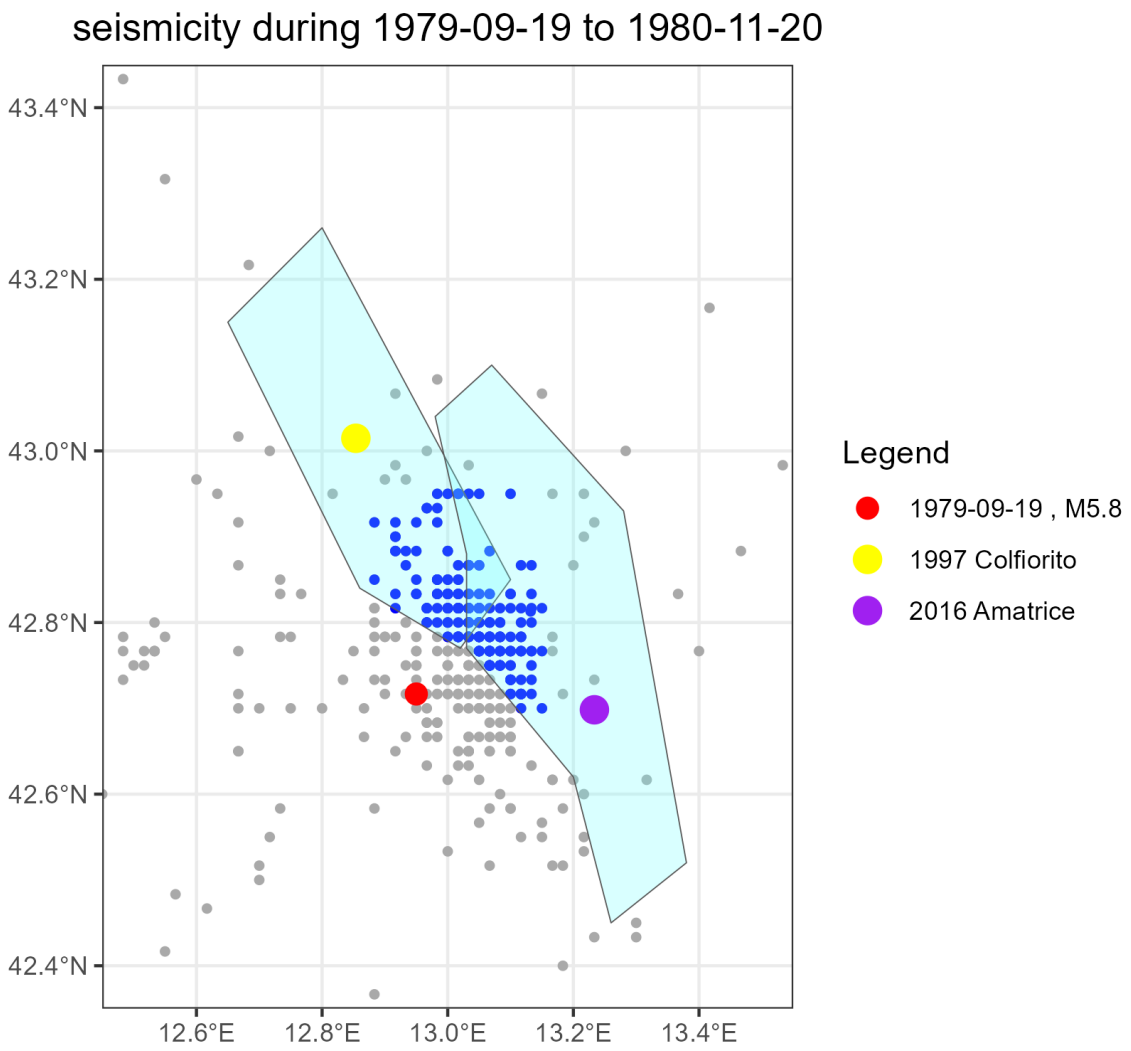

Figure 6.30: Spatial distribution of seismicity similar to those in Fig. 6.29. Data likely related to the 1997 sequence, shown in blue, are intended to be excluded from the 1997 Colfiorito and 2016 Amatrice polygons.

triggering patterns of earthquakes differ during short- and long-term intervals, so we might observe significant variations in ETAS estimates between individual sequences and over longer timescales.

My analysis of the three individual sequences of the 1997 Colfiorito, 2009 L'Aquila, and 2016 Amatrice earthquakes, along with their long-term past seismicity in Central Italy, confirms the latter hypothesis. The differences in the ETAS parameters across the short-term and long-term datasets indicate significant variability in seismicity behaviour over different timescales. Specifically, the parameter μ, which represents the background seismicity rate, shows higher values for individual sequences compared to

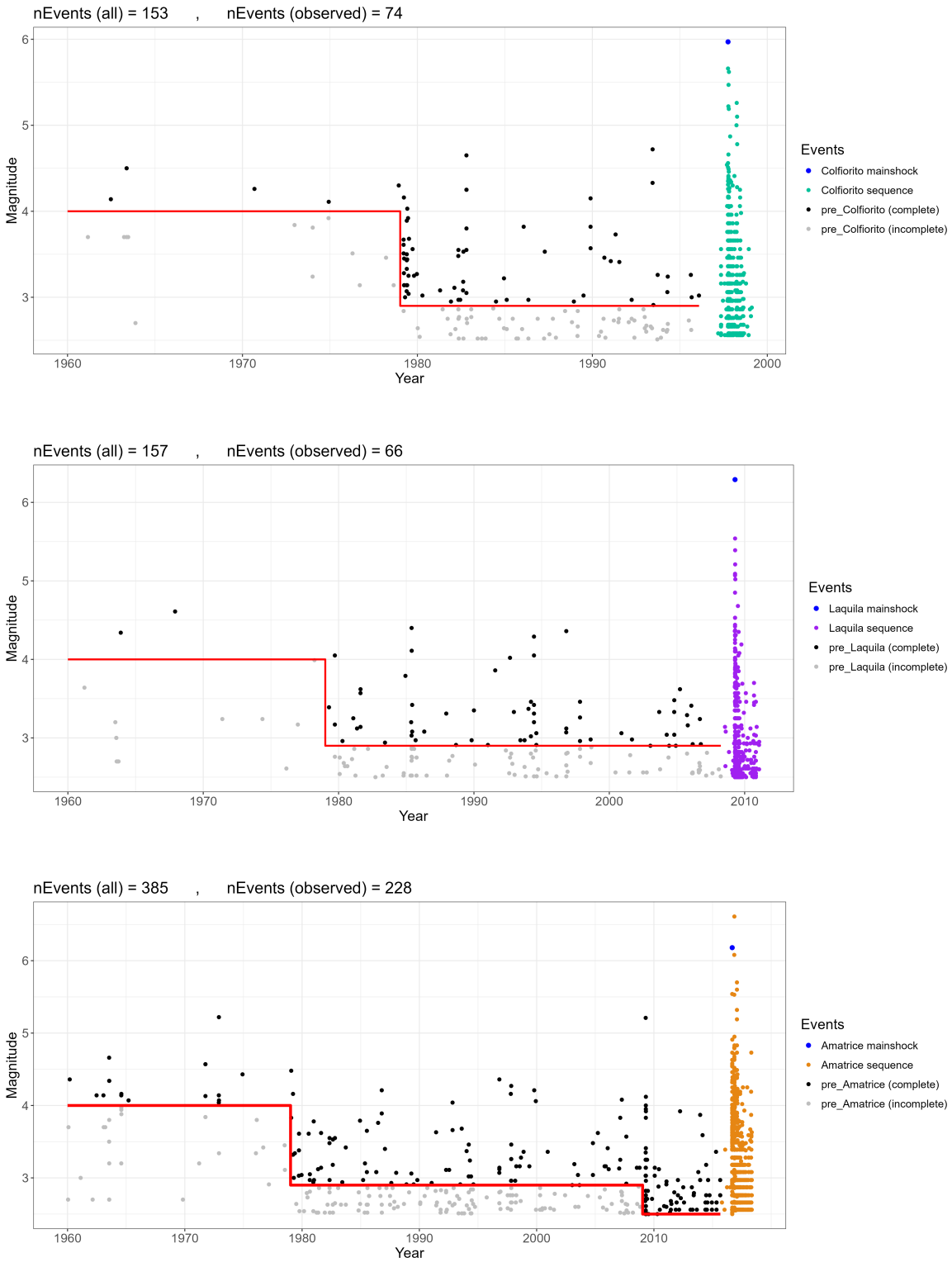

Figure 6.31: Seismic sequences similar to those in Fig. 6.28. In this figure, the 1979 sequence has been removed from both the 1997 Colfiorito and the 2016 Amatrice data.

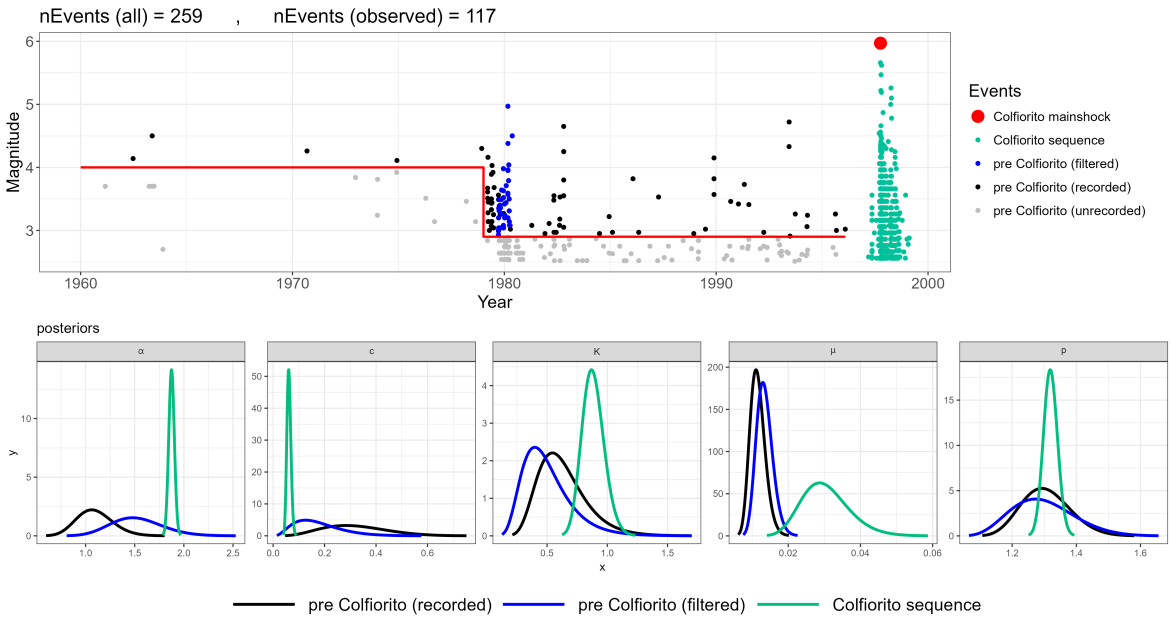

Figure 6.32: Posterior distributions for the 1997 Colfiorito sequence (in green) and the long-term seismicity model fitted twice: once including the 1979 data (in black) and once excluding the 1979 data (in blue). The inclusion of the 1979 data slightly shifts the posteriors.

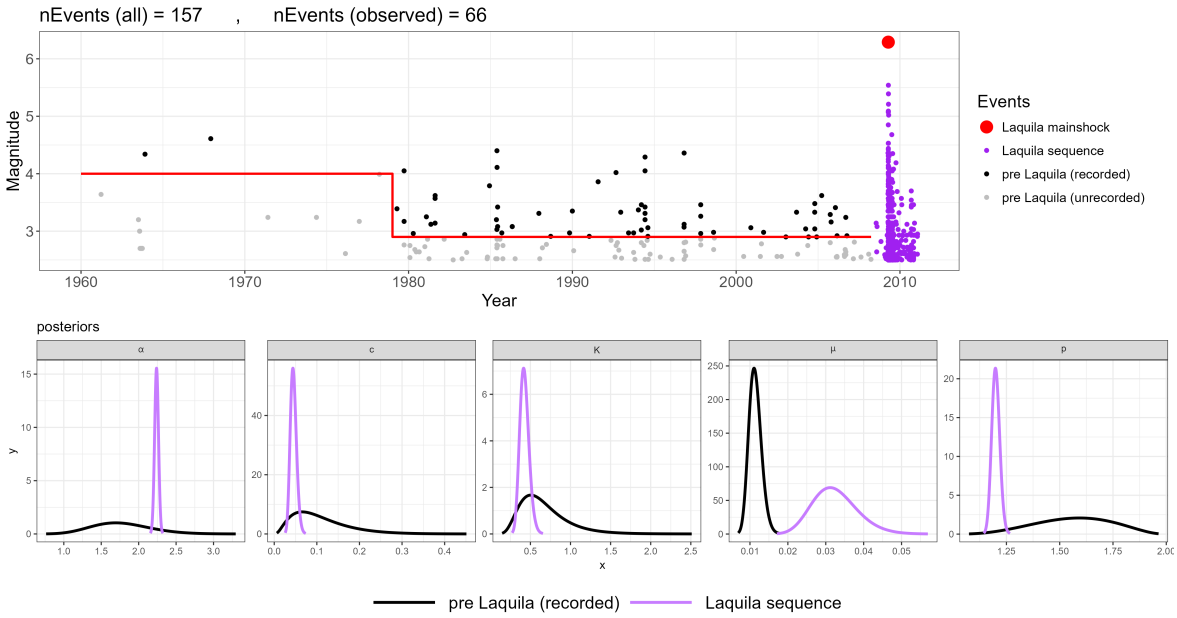

Figure 6.33: Posterior distributions for the 2009 L'Aquila sequence (in purple) and the long-term seismicity prior to the sequence (in black).

long-term data. This suggests an elevated level of activity during aftershock sequences, which is not mirrored in the long-term seismicity records. The triggering coefficient K shows notable differences between the datasets without any specific pattern. The

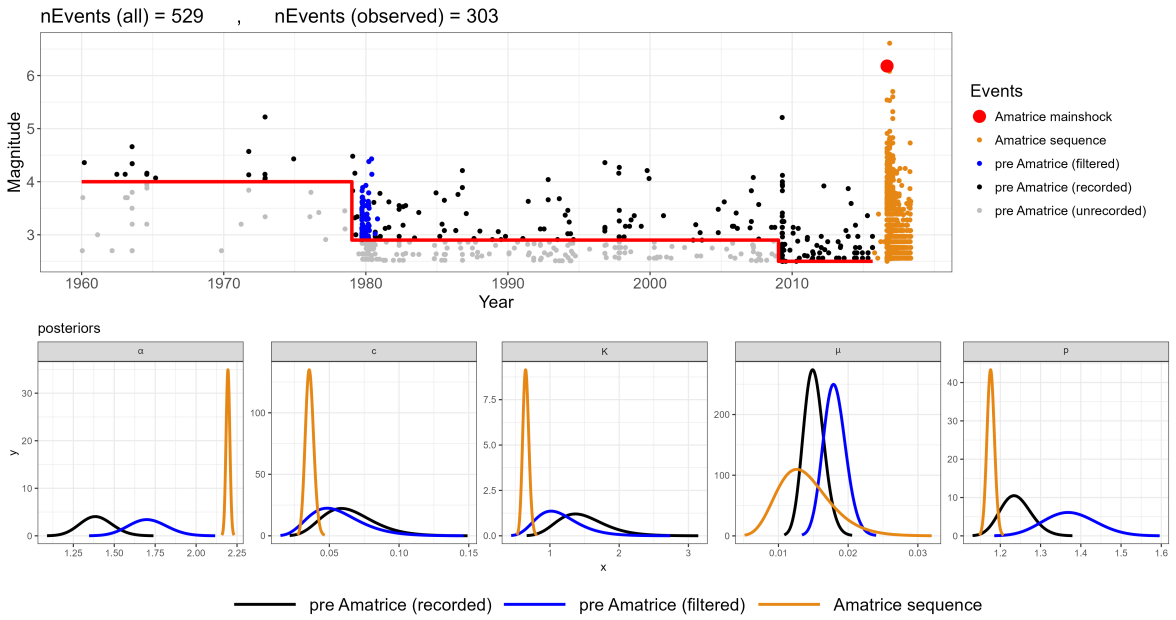

Figure 6.34: Posterior distributions for the 2016 Amatrice sequence (in orange) and the long-term seismicity model fitted twice: once including the 1979 data (in black) and once excluding the 1979 data (in blue). The inclusion of the 1979 data slightly shifts the posteriors.

productivity law exponent α shows variations, with consistently higher values of +0.5 to +0.8 for the phases of aftershock sequences compared to the long-term data. These variations highlight the different dynamics of aftershock generation immediately following a mainshock, compared to the seismicity patterns observed during extended periods of seismic activity. Furthermore, the decay parameters of the Omori law c and p display variations that further support the hypothesis. Short-term sequences have lower values of c, indicating a more immediate onset of aftershocks, while p values, although generally consistent, still exhibit slight differences.

These findings have profound implications for earthquake forecasting and seismic hazard assessment. The observed variabilities in the ETAS parameters between short-term and long-term datasets indicate that models based solely on long-term data may not accurately capture the dynamics of aftershock sequences, and models based solely on short-term data of aftershock sequences may not accurately reflect the long-term seismicity behaviour. Consequently, incorporating the modified ETAS model, which accounts for both short- and long-term incompleteness, and tailored modelling approaches that consider these differences are crucial for improving our understanding of seismic hazards and improvement of forecasting models.

Table 6.2: Mode of ETAS posteriors for the short-term (individual sequences) and long-term (past seismicity) scenarios.

Scenario	μ	**K**	α	**c**	**p**
past seismicity (pre 1997 Colfiorito sequence)	0.011	0.60	1.08	0.31	1.30
1997 Colfiorito sequence	0.029	0.88	1.87	0.06	1.32
past seismicity (pre 2009 L'Aquila)	0.011	0.61	1.75	0.09	1.57
2009 L'Aquila	0.032	0.42	2.24	0.04	1.20
past seismicity (pre 2016 Amatrice)	0.015	1.45	1.39	0.06	1.24
2016 Amatrice	0.013	0.65	2.19	0.04	1.18

Similar to the approach in Section 4.3 and Fig. 4.9, which tested the ability of the estimated ETAS parameters to predict intensities for short-term incompleteness periods, I apply the same method to assess how well the ETAS estimates for long-term past seismicity can predict earthquake rates over 124 years. This is accomplished by using the mode of posteriors and real seismicity in the modified conditional intensity for the long-term version in Eq. 6.5. As a result, the magnitude time series and the rate time series are shown in Figs. 6.35, 6.36, and 6.37 for the long-term past seismicity preceding the 1997 Colfiorito, 2009 L'Aquila, and 2016 Amatrice sequences, respectively. In these figures, the black lines show the observed (recorded) seismicity and the red lines show the seismicity predicted by the modified ETAS model.

As is evident, there is a strong correlation between the peaks and densities in both the 'observed' and 'predicted' seismicity rates. The long-term version of the ETAS model effectively captures the occurrence of significant seismic events, demonstrating its predictive power in regions where a large number of events are observed. Although some discrepancies, in the form of underestimations or overestimations of peak heights, may occur, these differences are expected in any statistical model and do not detract from the overall performance. The model remains robust in capturing the general patterns and rates of seismicity, providing a reliable representation of the underlying seismic processes.

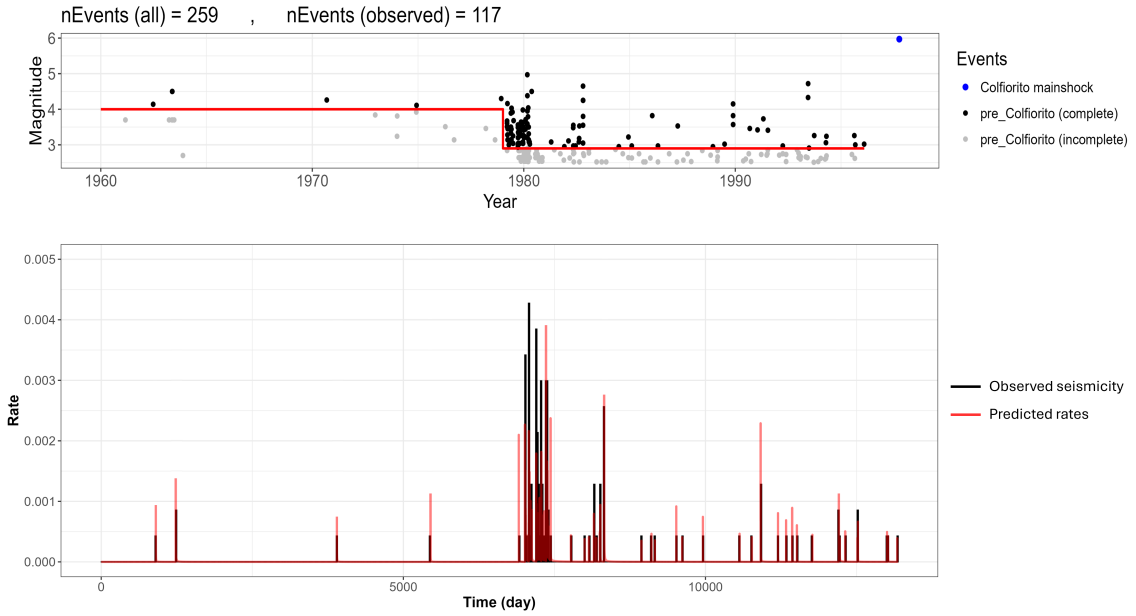

Figure 6.35: Magnitude plot (top) and rate plot (bottom) for the long-term seismicity preceding the 1997 Colfiorito earthquake. The bottom plot shows the histogram of observed seismicity (in black) and the rates predicted by the modified ETAS model (in red).

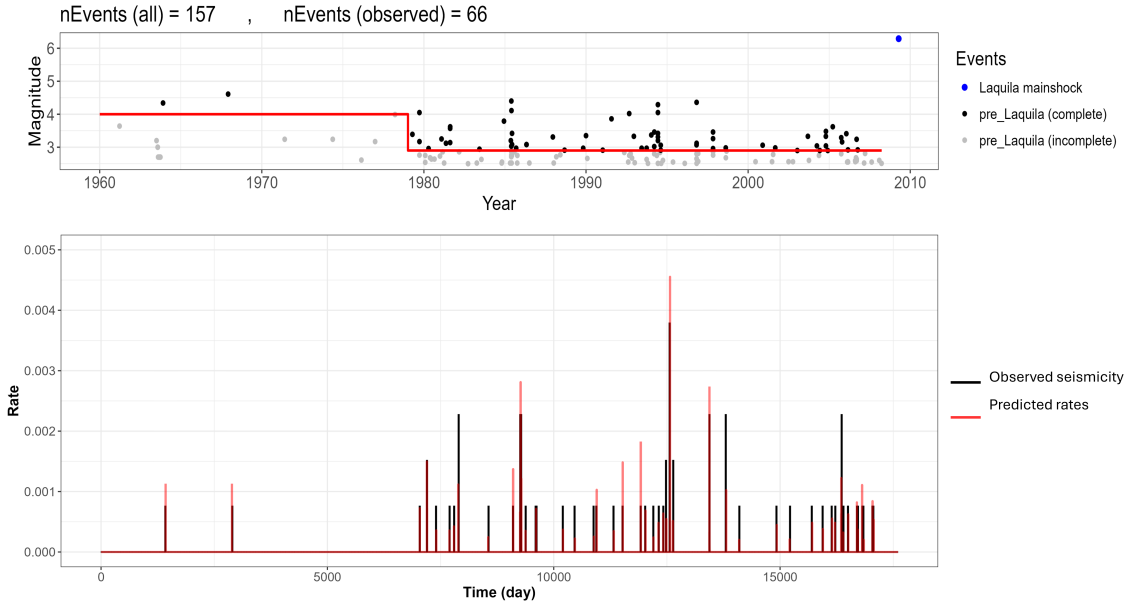

Figure 6.36: Magnitude plot (top) and rate plot (bottom) for the long-term seismicity preceding the 2009 L'Aquila earthquake. The bottom plot shows the histogram of observed seismicity (in black) and the rates predicted by the modified ETAS model (in red).

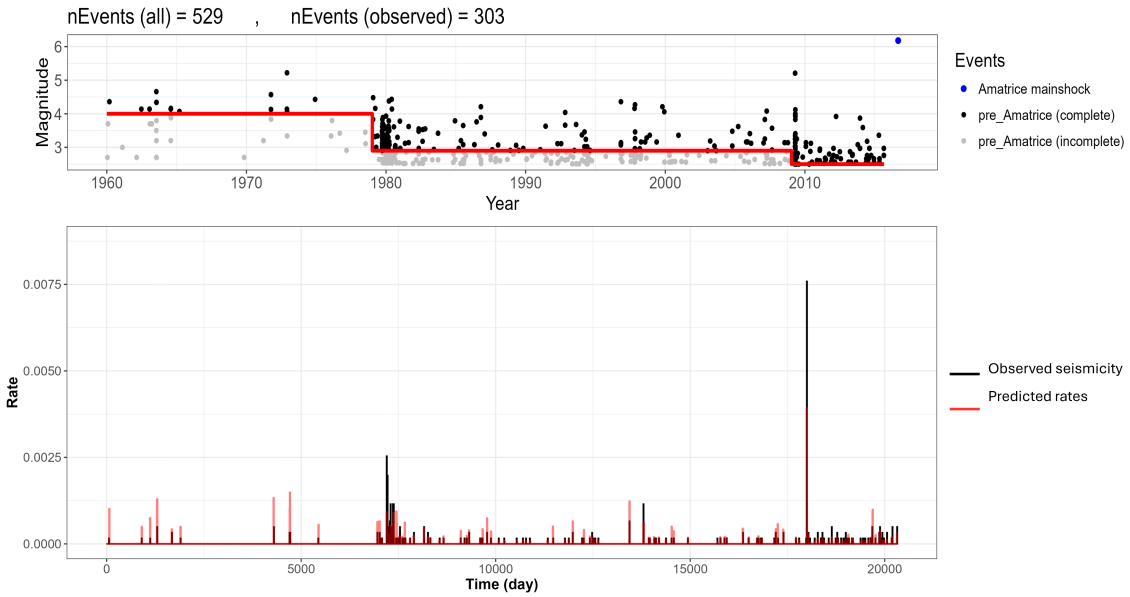

Figure 6.37: Magnitude plot (top) and rate plot (bottom) for the long-term seismicity preceding the 2016 Amatrice earthquake. The bottom plot shows the histogram of observed seismicity (in black) and the rates predicted by the modified ETAS model (in red).

6.5 Conclusion

In this chapter, I developed an ETAS model to address long-term catalogue incompleteness in earthquake records. This work is essential because it allows for the inclusion of long-term seismicity data in hazard assessments and earthquake forecasting, providing a more comprehensive understanding of seismic activity.

I began by examining the historical context of seismic recording, from written evidence to instrumental data, highlighting the persistent challenges of gaps and inconsistencies in long-term earthquake catalogues. Recognising these challenges, I introduced a step-like model for long-term incompleteness, which accounts for the gradual decrease in the magnitude of completeness (M_c) over time due to advancements in seismic detection technology and network expansions.

To accommodate long-term data incompleteness, I developed a long-term censorship function that adjusts the conditional intensity and likelihood functions of the ETAS model. This ensures that the modified ETAS model accurately reflects the apparent rates of seismic events, providing a more realistic depiction of earthquake

activity over extended periods.

I validated the robustness of the long-term version of the modified ETAS model using synthetic earthquake catalogues with defined step-model incompleteness scenarios. Synthetic experiments involved creating earthquake catalogues with various levels of incompleteness to mimic real-world conditions. The model was first tested by comparing the original and modified ETAS models. This comparison demonstrated that the modified model significantly reduced bias and produced parameter estimates that closely aligned with the true values used to generate synthetic data, highlighting its robustness in handling incomplete long-term data. Next, I tested the robustness of the modified ETAS model with various catalogue scenarios. This involved generating synthetic datasets with different seismicity patterns and levels of completeness. The modified model consistently provided accurate parameter estimates across these varied scenarios, showing its adaptability and potential for application in diverse seismic settings. In addition, I investigated the contribution level of step intervals by generating synthetic catalogues with a fixed five-step incompleteness model. I examined how different completeness levels influenced the ETAS parameter estimates. The results revealed that including large and small events across various completeness levels is crucial for reasonable parameter estimation. The modified model effectively handled these variations, further validating its comprehensive applicability in real-world conditions where data completeness can vary significantly over time.

Real-world applications of the model involved analysing data from central Italy, specifically the 1997 Colfiorito, 2009 L'Aquila, and 2016 Amatrice sequences. I used both the short-term and long-term incompleteness models to assess the seismicity in this region. The real-world data presented a more complex picture, with variations in the completeness of records over different periods. The analysis revealed notable differences in ETAS parameters between short- and long-term datasets. For example, the background seismicity rate (μ) was generally higher during individual aftershock sequences compared to long-term periods, indicating increased activity immediately following major events. The overall productivity coefficient (K) and the magnitude-based productivity exponent (α) also varied, reflecting the different dynamics of the aftershock triggering process and the long-term past seismicity. Furthermore, the Omori

law parameters (c and p) also displayed variations, further indicating the different behaviours observed on the short-term and long-term scales.

These findings highlight the variability in seismicity behaviour over different time scales and emphasise the need for tailored modelling approaches that account for both short- and long-term data. The modified ETAS models developed in this thesis, which address short- and long-term incompleteness, provide a more nuanced and accurate tool for understanding seismic hazards and improving earthquake forecasting.

Chapter 7

Discussion and Conclusions

7.1 Remaining limitations and possible improvements

Although my new modified ETAS model, which handles different forms of short- to long-term catalogue 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, I sum over all past events. However, I do not explicitly correct for the contributions to the intensity from events that lie below the time-varying completeness thresholds. In the short-term modelling, these contributions are likely small in total because they inherently come from less productive, smaller magnitude aftershocks. This could be accounted for analytically using the assumptions presented in this thesis, but since I recover the posteriors well with the current approach, I did not implement it here. In long-term modelling as well, we face a greater risk as larger earthquakes might be missing, and their triggering contributions to subsequent events are overlooked. Nevertheless, the model with the currently applied censorship function works well. The next phase of my study will involve developing the ETAS model to include the contributions of missing events in the ETAS estimations.

In the modified ETAS for the short-term incompleteness, I 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.

I was only able to consider a limited number of scenarios within this thesis. I highly recommend others conduct 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 specifically for short-term incompleteness. While the model introduced by Helmstetter et al. (2006) remains widely adopted, my experience indicates that its incorporation into the computational process of likelihood can be challenging, especially when resulting in the hypergeometric function. Looking for 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 the 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 the dynamics of earthquakes, taking into account the incompleteness issue for both the timing and location of earthquake events, potentially improving predictive accuracy and offering a better understanding of the underlying processes.

7.2 Discussion

In this thesis, I explored the impacts of different sources of bias and used this information to explore the data requirements for a selected catalogue to be sufficiently representative of the governing processes so that it can recover the key parameters unbiasedly.

The need for exploring these issues arises from the limitations of real datasets and the fact that an ETAS inversion generally returns a set of parameters but provides little information that helps us decide whether the selected data were sufficient to produce

unbiased estimates of these parameters in the first place. Consequently, it is easy to perform an inversion and work unquestionably with the parameters that were returned. For example, in practice, it is easy to define the spatio-temporal domain without undertaking a critical assessment of how the domain affects the ETAS parameters and the epistemic uncertainty associated with this choice. This thesis revisited that concept, emphasising that the true challenge lies in ensuring that the catalogue derived from the definition of a domain meets the necessary criteria for it to be representative. The intellectual starting point for this discussion in my thesis is representativeness, while for practitioners and operators, the starting point is often defining a region and then analysing it. So, this work aimed to bridge the gap between these two different starting points by developing methods that ensure the modelling domain is critically assessed for its representativeness, thereby minimising biases and improving the reliability of the ETAS inversions.

I 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 actively considering 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 my 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. To evaluate whether the data from quieter periods are 'sufficient' for reliably estimating μ, specific measures and indicators can be employed. To do so, resampling techniques such as bootstrap or jackknife are useful for testing the stability of μ estimates against variations in the dataset. Moreover, assessing the stability of information criteria like the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) helps gauge the impact of additional data on model fit. If these values plateau as more data are incorporated, it suggests

that further data would not substantially refine μ estimates. Consistent results across these assessments indicate that the dataset is sufficiently robust for accurate μ estimation.

- Do we have distinct mainshocks of different magnitudes in the selected 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? I have here 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.

- Are the short- and long-term incompleteness accurately modelled? The modified ETAS 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.

It is reassuring that the short-term and long-term incompleteness models asymptotically improve as the censored events are accounted for, so that a model that approximately corrects for incompleteness will be almost as good as one that perfectly finds the boundary for the incompleteness. This makes the method robust to uncertainties in the modelling of incompleteness, which is a desirable attribute.

Given that an ETAS inversion will generally return a set of parameters irrespective of their 'correctness', it is important that we can independently assess whether selected catalogues contain sufficient information to fit a model well. Here, I provided a blueprint to achieving this, and I believe that a Bayesian approach, where we can compare posteriors derived from different assumptions about the data, provides a natural, flexible, and intuitive tool for achieving this.

In addition, it is difficult to get an intuitive understanding of when differences in triggering parameters matter individually. Some other studies (e.g., by Li et al. 2024) have partially gone down this route in considering branching ratios where the average number of daughters is considered, but taking this approach does not take into account the timescales upon which those daughters are produced. Therefore, I strongly recommend routinely plotting out the triggering functions derived from sampling the posteriors in order to decide when these differences are significant. This provides a much better guide than simply plotting the posteriors of K, α, c, and p, as it visualises the influence of each parameter on the overall rate of aftershocks over time, thus offering a clearer interpretation of their practical implications.

I hope that this study gives an intuitive indication of where bias may arise in ETAS inversions, how these biases would propagate to systematic errors in operational earthquake forecasts, and how to correct for this bias in some cases. I believe that the analyses I have shown offer a way forward 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.

The modifications that I made to the ETAS model in this thesis demonstrated substantial improvements in capturing the intricacies of aftershock sequences and long-term seismic trends using a novel way of accounting for censorship of data. These advancements mark a significant step forward in seismicity modelling. Although my work is robust and promising, there are further avenues that could complement and enhance the applicability of the model.

Firstly, a comprehensive evaluation of the model's forecasting ability using the CSEP tests would be a valuable addition. These tests, elaborated on in Section 2.2.1, provide a rigorous framework for assessing the predictability of seismicity models. Implementing these tests would offer an empirical basis for further validation of the predictive power of my modified ETAS model. Specifically, running simulations with the enhanced model and subjecting the forecasts to N-Test, L-Test, and other relevant CSEP evaluations would provide additional information about its operational reliability.

Secondly, a comparative analysis between the modified ETAS model and other existing models (as mentioned in Section 4.1), such as those by Hainzl (2016), Mizrahi et al. (2021), Elst et al. (2022), and Li et al. (2024), could provide further benchmarks. Each of these models has a different ETAS formulation and functional form, as well as unique approaches to dealing with data incompleteness and seismicity patterns. Efforts to homogenise these models for a side-by-side performance evaluation would yield valuable insights. Standardising input datasets and testing conditions to ensure a fair comparison could highlight the strengths and potential integration points for the modified ETAS model.

Moreover, the integration of spatio-temporal models represents a crucial area for further improvement. Many tests, including those required by the CSEP, require the use of spatial bins, underscoring the importance of incorporating spatial dimensions into temporal forecasting models. Spatio-temporal models can significantly enhance the accuracy and reliability of seismic predictions by accounting for both spatial and temporal variations in seismic activity. In addition, the use of covariates, such as geological and geophysical parameters, can further refine these models. Previous studies suggest the practical advantages of incorporating geological and geophysical covariates. Research by Bayliss et al. (2020) and Bayliss et al. (2022) and Mancini et al. (2019) and Mancini et al. (2020) on spatial seismicity modelling, together with work on spatio-temporal ETAS models by Adelfio and Chiodi (2021) and Chiodi et al. (2021), demonstrate that adding covariates enhances model accuracy by integrating physical insights into statistical frameworks. Moreover, the INLA and inlabru frameworks employed in this study offer straightforward methods to implement and evaluate the effects of covariates, similar to their use in GLMs for various phenomena. This makes the integration of covariates both feasible and practical. Therefore, by incorporating covariates, we can better understand the underlying factors that influence seismicity patterns and improve the predictive abilities of the model.

7.3 Conclusion

The primary goal of this thesis was to address critical issues of temporal incompleteness and systematic biases that affect the estimation of ETAS model parameters, which are

essential for accurate earthquake forecasting and assessment of seismic hazards. This research has led to the development of significant methodological enhancements to the ETAS model, specifically targeting short-term and long-term catalogue incompleteness.

Initially, I addressed the issue of short-term incompleteness commonly observed following large mainshocks. This period of incompleteness is critical because it involves the initial aftershock sequence, which can provide valuable insight into the seismicity rates and properties of a seismic sequence. Traditional ETAS models often fail to account for this transient period, leading to under-reporting of smaller aftershocks due to waveform overlaps and seismogram saturation. To overcome this, I introduced a methodological enhancement that involves the definition of a time-dependent censorship function. This function was integrated into the ETAS conditional intensity and likelihood functions, creating a modified inversion solution capable of handling incomplete data during the crucial early stages. By implementing this enhancement within a Bayesian framework using the inlabru package and the INLA method, I was able to provide more accurate posterior distributions of the model parameters, rather than conventional point estimates. Extensive testing through synthetic experiments demonstrated that the modified ETAS model significantly reduced bias in parameter estimations compared to the original model when fitted to incomplete data. The results showed that the modified model could accurately capture ETAS parameters and actual aftershock rates even with significant data gaps, aligning closely with the results obtained from the original ETAS model fitted to complete data.

Beyond short-term incompleteness, I explored optimising the selection of representative samples for the ETAS inversions, which is critical for reducing bias in parameter estimation. Various sampling strategies and their potential biases were examined, leading to the proposal of a comprehensive approach to optimise survey design. This included evaluating the sensitivity of the ETAS model to temporal binning strategies, conditioning the model on the run-in history before an earthquake sequence, the role of combining different earthquake magnitudes, and the trade-offs between ETAS productivity parameters. Furthermore, considerations such as the choice of incompleteness model parameters, the impact of secondary large aftershocks, and the spatial and temporal size of the modelling domain were systematically analysed. Through this

analysis, I established guidelines that identify and minimise biases, enhancing the reliability and robustness of modelling seismicity patterns when fitting the ETAS model to real earthquake data.

Furthermore, the applicability of the ETAS model was expanded to address the incompleteness of long-term data, which arises from lack of coverage of the network and technological limitations over extended periods. The same censorship approach used for short-term incompleteness was adapted to handle incomplete long-term data, such as century-long records from the instrumental era that lack distinct aftershock sequences. This comprehensive approach significantly improved the ETAS model's predictive accuracy, as evidenced by its application to both simulated data and real earthquake sequences. The modified ETAS model was tested with synthetic earthquake catalogues with defined step-model incompleteness scenarios, showing that it significantly reduced bias and produced parameter estimates closely aligned with the true values used to generate the synthetic data. This demonstrated the robustness of the model in handling incomplete long-term data, ensuring robust parameter estimation even when faced with extensive temporal gaps.

The robustness of the modified ETAS in handling both short- and long-term incompleteness was further demonstrated through a real experiment in Central Italy. The earthquake records of this region, which span several decades, present significant long-term data gaps because of the evolution of seismic recording. By applying my modified ETAS model, I was able to address these gaps and produce reliable parameter estimates. The results showed that the modified model well captured the seismicity patterns in Central Italy, validating its effectiveness in regions with extensive temporal incompleteness.

Finally, this thesis has made significant advances in the field of seismicity modelling by developing and validating modified ETAS models that effectively handle incomplete seismic data and reduce biases in parameter estimation. These enhancements improve the model's predictive accuracy, providing more reliable forecasts of aftershock sequences and better-informed seismic hazard assessments. The contributions of this research improve our understanding of earthquake behaviour and improve the accuracy of earthquake forecasts, ultimately contributing to more effective seismic hazard

mitigation and risk reduction efforts. Future research can build on these findings by exploring additional modifications and extensions to the ETAS model, incorporating spatial variability, refining incompleteness models, and expanding real-world applications to further validate their effectiveness and generalisability.

Appendix

All data, source codes, and notebooks developed for this thesis are available on GitHub at: https://github.com/Farnaz-Kamranzad?tab=repositories.

Bibliography

Adde, A., Darveau, M., Barker, N., and Cumming, S. (2020). "Predicting spatiotemporal abundance of breeding waterfowl across Canada: A Bayesian hierarchical modelling approach". *Diversity and Distributions*, 2020, 26, pp. 1248–1263

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

Aki, K. (1965). "Maximum likelihood estimate of b in the formula log N= a-bM and its confidence limits". *Bull. Earthquake Res. Inst., Tokyo Univ.*, 1965, 43, pp. 237–239

Ambraseys, N. N. and Melville, C. P. (2005). *A history of Persian earthquakes*. Cambridge university press.

Antonietti, P. F., Mazzieri, I., Melas, L., Paolucci, R., Quarteroni, A., Smerzini, C., Stupazzini, M., et al. (2021). "Three-dimensional physics-based earthquake ground motion simulations for seismic risk assessment in densely populated urban areas". *Mathematics in Engineering*, 2021, 3, pp. 1–31

Aomoto, K. (2011). *Theory of hypergeometric functions*. Springer.

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

Arrowsmith, S. J., Trugman, D. T., MacCarthy, J, Bergen, K. J., Lumley, D., and Magnani, M. B. (2022). "Big data seismology". *Reviews of Geophysics*, 2022, 60, e2021RG000769

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

Bachura, M, Fischer, T, Doubravová, J, and Horálek, J (2021). "From earthquake swarm to a main shock–aftershocks: The 2018 activity in west Bohemia/Vogtland". *Geophysical Journal International*, 2021, 224, pp. 1835–1848

Baiesi, M. and Paczuski, M. (2004). "Scale-free networks of earthquakes and aftershocks". *Physical review E*, 2004, 69, p. 066106

Baker, J., Bradley, B., and Stafford, P. (2021). *Seismic hazard and risk analysis*. Cambridge University Press.

Barani, S, Ferretti, G, Massa, M, and Spallarossa, D (2007). "The waveform similarity approach to identify dependent events in instrumental seismic catalogues". *Geophysical Journal International*, 2007, 168, pp. 100–108

Bartholomew, D. J., Knott, M., and Moustaki, I. (2011). *Latent variable models and factor analysis: A unified approach*. John Wiley & Sons.

Bayliss, K., Naylor, M., Illian, J., and Main, I. G. (2020). "Data-driven optimization of seismicity models using diverse data sets: Generation, evaluation, and ranking using Inlabru". *Journal of Geophysical Research: Solid Earth*, 2020, 125, e2020JB020226

Bayliss, K., Naylor, M., Kamranzad, F., and Main, I. (2022). "Pseudo-prospective testing of 5-year earthquake forecasts for California using inlabru". *Natural Hazards and Earth System Sciences*, 2022, 22, pp. 3231–3246

Bayliss, K., Naylor, M., and Main, I. G. (2019). "Probabilistic identification of earthquake clusters using rescaled nearest neighbour distance networks". *Geophysical Journal International*, 2019, 217, pp. 487–503

Bayona, J. A., Savran, W. H., Rhoades, D. A., and Werner, M. (2022). "Prospective evaluation of multiplicative hybrid earthquake forecasting models in California". *Geophysical Journal International*, 2022, 229, pp. 1736–1753

Becker, D. G., Woolford, D. G., and Dean, C. B. (2021). "Algorithmically deconstructing shot locations as a method for shot quality in hockey". *Journal of Quantitative Analysis in Sports*, 2021, 17, pp. 107–115

Bell, O., Jones, M. E., Cunningham, C. X., Ruiz-Aravena, M., Hamilton, D. G., Comte, S., Hamede, R. K., Bearhop, S., and McDonald, R. A. (2021). "Isotopic niche variation in Tasmanian devils Sarcophilus harrisii with progression of devil facial tumor disease". *Ecology and Evolution*, 2021, 11, pp. 8038–8053

Bilmes, J. A. et al. (1998). "A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models". *International computer science institute*, 1998, 4, p. 126

Blangiardo, M. and Cameletti, M. (2015). *Spatial and spatio-temporal Bayesian models with R-INLA*. John Wiley & Sons.

Bommer, J. J. (2002). "Deterministic vs. probabilistic seismic hazard assessment: an exaggerated and obstructive dichotomy". *Journal of earthquake engineering*, 2002, 6, pp. 43–73

— (2022). "Earthquake hazard and risk analysis for natural and induced seismicity: towards objective assessments in the face of uncertainty". *Bulletin of earthquake engineering*, 2022, 20, pp. 2825–3069

Bottiglieri, M, Lippiello, E., Godano, C., and Arcangelis, L. de (2009). "Identification and spatiotemporal organization of aftershocks". *Journal of Geophysical Research: Solid Earth*, 2009, 114

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

Cara, F., Cultrera, G., Riccio, G., Amoroso, S., Bordoni, P., Bucci, A., D'Alema, E., D'Amico, M., Cantore, L., Carannante, S., et al. (2019). "Temporary dense seismic network during the 2016 Central Italy seismic emergency for microzonation studies". *Scientific data*, 2019, 6, p. 182

Chen, Liu, J., Chen, L, Chen, Q, and Chan, L. (1998). "Global seismic hazard assessment based on area source model and seismicity data". *Natural Hazards*, 1998, 17, pp. 251–267

Cheysson, F. and Lang, G. (2022). "Spectral estimation of Hawkes processes from count data". *Annals of Statistics*, 2022, 50, pp. 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*, 2021, 11, 13 pages

Christophersen, A, Rhoades, D., Gerstenberger, M., Bannister, S, Becker, J, Potter, S., and McBride, S (2017). "Progress and challenges in operational earthquake forecasting in New Zealand". *2017 New Zealand society for earthquake engineering annual technical conference*, 9 pages.

Chu, A., Schoenberg, F. P., Bird, P., Jackson, D. D., and Kagan, Y. Y. (2011). "Comparison of ETAS parameter estimates across different global tectonic zones". *Bulletin of the Seismological Society of America*, 2011, 101, pp. 2323–2339

Cornell, C. A. (1968). "Engineering seismic risk analysis". *Bulletin of the seismological society of America*, 1968, 58, pp. 1583–1606

Cunningham, C. X., Comte, S., McCallum, H., Hamilton, D. G., Hamede, R., Storfer, A., Hollings, T., Ruiz-Aravena, M., Kerlin, D. H., Brook, B. W., et al. (2021a). "Quantifying 25 years of disease-caused declines in Tasmanian devil populations: host density drives spatial pathogen spread". *Ecology Letters*, 2021, 24, pp. 958–969

Cunningham, C. A., Thomas, C. D., Morecroft, M. D., Crick, H. Q., and Beale, C. M. (2021b). "The effectiveness of the protected area network of Great Britain". *Biological Conservation*, 2021, 257, p. 109146

D'Angelo, N., Abbruzzo, A., and Adelfio, G. (2020). "Spatial bayesian hierarchical modelling with integrated nested laplace approximation". *arXiv preprint arXiv:2009.03712*, 2020

Daskalakis, C., Tzamos, C., and Zampetakis, M. (2017). "Ten steps of EM suffice for mixtures of two Gaussians". *Conference on Learning Theory*. PMLR, pp. 704–710.

Dastile, X., Celik, T., and Potsane, M. (2020). "Statistical and machine learning models in credit scoring: A systematic literature survey". *Applied Soft Computing*, 2020, 91, p. 106263

Davis, S. and Frohlich, C. (1991). "Single-link cluster analysis, synthetic earthquake catalogues, and aftershock identification". *Geophysical Journal International*, 1991, 104, pp. 289–306

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). "Maximum likelihood from incomplete data via the EM algorithm". *Journal of the royal statistical society: series B (methodological)*, 1977, 39, pp. 1–22

Douglas, J., Crowley, H., Silva, V., Marzocchi, W., Danciu, L., and Pinho, R. (2024). "Methods for evaluating the significance and importance of differences amongst probabilistic seismic hazard results for engineering and risk analyses: a review and insights". *Bulletin of Earthquake Engineering*, 2024, pp. 1–28

Dutta, P., Mishra, O., and Naskar, M. (2013). "A review of operational earthquake forecasting methodologies using linguistic fuzzy rule-based models from imprecise data with weighted regression approach". *Journal of Sustainability Science and Managemen*, 2013, 8, pp. 220–235

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

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

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

Elst, van-der, 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*, 2022, 92, pp. 620–640

Essing, D. and Poli, P. (2022). "Spatiotemporal evolution of the seismicity in the alto tiberina fault system revealed by a high-resolution template matching catalog". *Journal of Geophysical Research: Solid Earth*, 2022, 127, e2022JB024845

Ferraccioli, F., Arnone, E., Finos, L., Ramsay, J. O., and Sangalli, L. M. (2021). "Nonparametric density estimation over complicated domains". *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 2021, 83, pp. 346–368

Field, E. H. (2022). *Some systemic risks to progress on seismic hazard assessment.*

Field, E. H., Jordan, T., Jones, L. M., Michael, A. J., Blanpied, M. L., and Participants, W. (2016). "The potential uses of operational earthquake forecasting". *Seismological research letters*, 2016, 87, pp. 313–322

Field, E. H., Jordan, T., Page, M. T., Milner, K. R., Shaw, B. E., Dawson, T. E., Biasi, G. P., Parsons, T., Hardebeck, J. L., Michael, A. J., et al. (2017). "A synoptic view of the third Uniform California Earthquake Rupture Forecast (UCERF3)". *Seismological Research Letters*, 2017, 88, pp. 1259–1267

Fisher, R. A. (1922). "On the mathematical foundations of theoretical statistics". *Philosophical transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, 1922, 222, pp. 309–368

Frohlich, C. and Davis, S. (1985). "Identification of aftershocks of deep earthquakes by a new ratios method". *Geophysical Research Letters*, 1985, 12, pp. 713–716

— (1990). "Single-link cluster analysis as a method to evaluate spatial and temporal properties of earthquake catalogues". *Geophysical Journal International*, 1990, 100, pp. 19–32

Gardner, J. and Knopoff, L. (1974). "Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?" *Bulletin of the seismological society of America*, 1974, 64, pp. 1363–1367

Gerstenberger, M. C., Marzocchi, W., Allen, T., Pagani, M., Adams, J., Danciu, L., Field, E. H., Fujiwara, H., Luco, N., Ma, K.-F., et al. (2020). "Probabilistic seismic hazard analysis at regional and national scales: State of the art and future challenges". *Reviews of Geophysics*, 2020, 58, e2019RG000653

Gómez-Rubio, V. (2020). *Bayesian inference with INLA*. Chapman and Hall/CRC.

Gordon, J. S., Fox, E. W., and Schoenberg, F. P. (2021). "A nonparametric hawkes model for forecasting California seismicity". *Bulletin of the Seismological Society of America*, 2021, 111, pp. 2216–2234

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*, 2022, 36, pp. 2133–2152

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*, 2021, 112, pp. 474–493

Gulia, L. and Wiemer, S. (2019). "Real-time discrimination of earthquake foreshocks and aftershocks". *Nature*, 2019, 574, pp. 193–199

Guo, Z. and Ogata, Y. (1997). "Statistical relations between the parameters of aftershocks in time, space, and magnitude". *Journal of Geophysical Research: Solid Earth*, 1997, 102, pp. 2857–2873

Gupta, M. and Chen, Y. (2011). "Theory and use of the EM algorithm". *Foundations and Trends® in Signal Processing*, 2011, 4, pp. 223–296

Gutenberg, B. and Richter, C. (1944). "Frequency of earthquakes in California". *Bulletin of the Seismological society of America*, 1944, 34, pp. 185–188

Hager, B. H., Dieterich, J., Frohlich, C., Juanes, R., Mantica, S., Shaw, J. H., Bottazzi, F., Caresani, F., Castineira, D., Cominelli, A., et al. (2021). "A process-based approach to understanding and managing triggered seismicity". *Nature*, 2021, 595, pp. 684–689

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

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

Hainzl, S., Scherbaum, F., and Beauval, C. (2006). "Estimating background activity based on interevent-time distribution". *Bulletin of the Seismological Society of America*, 2006, 96, pp. 313–320

Hardebeck, Llenos, A. L., Michael, A. J., Page, M. T., Schneider, M., and Elst, N. J. van der (2024). "Aftershock Forecasting". *Annual Review of Earth and Planetary Sciences*, 2024, 52

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

Hardebeck, J. L. (2021). "Spatial clustering of aftershocks impacts the performance of physics-based earthquake forecasting models". *Journal of Geophysical Research: Solid Earth*, 2021, 126, e2020JB020824

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

— (2016). "Model parameter estimation bias induced by earthquake magnitude cut-off". *Geophysical Journal International*, 2016, 204, pp. 1266–1287

Harte, D. (2015). "Log-likelihood of earthquake models: evaluation of models and forecasts". *Geophysical Journal International*, 2015, 201, pp. 711–723

Hawkes, A. G. (1971). "Point spectra of some mutually exciting point processes". *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1971, 33, pp. 438–443

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*, 2006, 96, pp. 90–106

Herovic, E. (2016). "The challenges of communicating low probability and high consequence risk: recommendations for earthquake pre-crisis and emergency-risk communication", 2016

Hristopulos, D. T. (2020). *Random fields for spatial data modeling*. Springer.

Hu, G., Yang, H.-C., and Xue, Y. (2021). "Bayesian group learning for shot selection of professional basketball players". *Stat*, 2021, 10, e324

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*, 2022, 93, pp. 1601–1618

Iacoletti, S., Cremen, G., and Galasso, C. (2021). "Advancements in multi-rupture time-dependent seismic hazard modeling, including fault interaction". *Earth-Science Reviews*, 2021, 220, p. 103650

IRSC (2024). "Iranian seismological centre (IRSC)". *Institute of Geophysics, the University of Tehran*, 2024. URL: http://irsc.ut.ac.ir/

Jackson, D, Werner, C, Liukis, M, Schorlemmer, D, Maechling, P, Rhoades, D, Zechar, J, Marzocchi, W, and Jordan, T (2016). "Recent Achievements of the Collaboratory for the Study of Earthquake Predictability". *SSA 2016 Annual Meeting*, 2016, pp. 220–235

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

Jeffreys, H. (1998). *The theory of probability*. OuP Oxford.

Jiang, Z. (2020). "A survey on spatial and spatiotemporal prediction methods". *arXiv preprint arXiv:2012.13384*, 2020

Jordan, Chen, Y.-T., Gasparini, P., Madariaga, R., Main, I., Marzocchi, W., Papadopoulos, G., Sobolev, G., Yamaoka, K., and Zschau, J. (2011). "Operational earthquake forecasting: State of knowledge and guidelines for utilization". *Annals of Geophysics*, 2011, 54, pp. 315–391

Jordan and Jones, L. (2010). "Operational earthquake forecasting: Some thoughts on why and how". *Seismological Research Letters*, 2010, 81, pp. 571–574

Jordan, Marzocchi, W, Michael, A., and Gerstenberger, M. (2014). *Operational earthquake forecasting can enhance earthquake preparedness*.

Juan Verdoy, P. (2021). "Enhancing the SPDE modeling of spatial point processes with INLA, applied to wildfires. Choosing the best mesh for each database". *Communications in Statistics-Simulation and Computation*, 2021, 50, pp. 2990–3030

Jullum, M., Thorarinsdottir, T., and Bachl, F. E. (2020). "Estimating seal pup production in the Greenland Sea by using Bayesian hierarchical modelling". *Journal of the Royal Statistical Society Series C: Applied Statistics*, 2020, 69, pp. 327–352

Kagan, Y. (2004). "Short-term properties of earthquake catalogs and models of earthquake source". *Bulletin of the Seismological Society of America*, 2004, 94, pp. 1207–1228

Kagan, Y. and Jackson (1991). "Long-term earthquake clustering". *Geophysical Journal International*, 1991, 104, pp. 117–133

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

Khan, K., Luo, H., and Xi, W. (2021). "Computing with R-INLA: Accuracy and reproducibility with implications for the analysis of COVID-19 data". *arXiv preprint arXiv:2111.01285*, 2021

Khawaja, A. M., Hainzl, S., Schorlemmer, D., Iturrieta, P., Bayona, J. A., Savran, W. H., Werner, M., and Marzocchi, W. (2023). "Statistical power of spatial earthquake forecast tests". *Geophysical Journal International*, 2023, 233, pp. 2053–2066

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

Kolev, A. A. and Ross, G. J. (2019). "Inference for ETAS models with non-Poissonian mainshock arrival times". *Statistics and Computing*, 2019, 29, pp. 915–931

Krainski, E., Gómez-Rubio, V., Bakka, H., Lenzi, A., Castro-Camilo, D., Simpson, D., Lindgren, F., and Rue, H. (2018). *Advanced spatial modeling with stochastic partial differential equations using R and INLA*. Chapman and Hall/CRC.

Kullback, S. and Leibler, R. A. (1951). "On information and sufficiency". *The annals of mathematical statistics*, 1951, 22, pp. 79–86

Laub, P., Lee, Y., and Taimre, T. (2021). *The Elements of Hawkes Processes*. Springer, p. 133.

Lee, D.-T. and Schachter, B. J. (1980). "Two algorithms for constructing a Delaunay triangulation". *International Journal of Computer & Information Sciences*, 1980, 9, pp. 219–242

Li, J., Sornette, D., Wu, Z., Zhuang, J., and Jiang, C. (2024). "Revisiting Seismicity Criticality: A New Framework for Bias Correction of Statistical Seismology Model Calibrations". *arXiv preprint arXiv:2404.16374*, 2024

Lindgren, F., Rue, H., and Lindström, J. (2011). "An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach". *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 2011, 73, pp. 423–498

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

Lippiello, E., Petrillo, G., Godano, C., Tramelli, A., Papadimitriou, E., and Karakostas, V (2019b). "Forecasting of the first hour aftershocks by means of the perceived magnitude". *Nature communications*, 2019, 10, p. 2953

Liu, Y. and Yin, G. (2020). "The Delaunay triangulation learner and its ensembles". *Computational Statistics & Data Analysis*, 2020, 152, p. 107030

Lolli, B., Randazzo, D., Vannucci, G., and Gasperini, P. (2020). "The homogenized instrumental seismic catalog (HORUS) of Italy from 1960 to present". *Seismological Society of America*, 2020, 91, pp. 3208–3222

Lombardi, A. M. (2015). "Estimation of the parameters of ETAS models by Simulated Annealing". *Scientific reports*, 2015, 5, p. 8417

Main, I. (1995). "Earthquakes as critical phenomena: implications for probabilistic seismic hazard analysis". *Bulletin of the Seismological Society of America*, 1995, 85, pp. 1299–1308

— (1999). "Is the reliable prediction of individual earthquakes a realistic scientific goal?" *Nature*, 1999, 397

Mancini, S, Segou, M, Werner, M., and Cattania, C (2019). "Improving physics-based aftershock forecasts during the 2016–2017 Central Italy Earthquake Cascade". *Journal of Geophysical Research: Solid Earth*, 2019, 124, pp. 8626–8643

Mancini, S., Segou, M., Werner, M. J., and Parsons, T. (2020). "The predictive skills of elastic Coulomb rate-and-state aftershock forecasts during the 2019 Ridgecrest, California, earthquake sequence". *Bulletin of the Seismological Society of America*, 2020, 110, pp. 1736–1751

Margheriti, L., Chiaraluce, L., Voisin, C., Cultrera, G., Govoni, A., Moretti, M., Bordoni, P., Luzi, L., Azzara, R., Valoroso, L., et al. (2011). "Rapid response seismic networks in Europe: lessons learnt from the L'Aquila earthquake emergency". *Annals of Geophysics*, 2011, 54, pp–392

Marsan, D. and Lengline, O. (2008). "Extending earthquakes' reach through cascading". *Science*, 2008, 319, pp. 1076–1079

Marsan, D. and Lengliné, O. (2010). "A new estimation of the decay of aftershock density with distance to the mainshock". *Journal of Geophysical Research: Solid Earth*, 2010, 115

Martino, S., Pace, D. S., Moro, S., Casoli, E., Ventura, D., Frachea, A., Silvestri, M., Arcangeli, A., Giacomini, G., Ardizzone, G., et al. (2021). "Integration of presence-only data from several sources: a case study on dolphins' spatial distribution". *Ecography*, 2021, 44, pp. 1533–1543

Marzocchi, W. and Jordan (2018). "Experimental concepts for testing probabilistic earthquake forecasting and seismic hazard models". *Geophysical Journal International*, 2018, 215, pp. 780–798

Marzocchi, W., Lombardi, A. M., and Casarotti, E. (2014). "The establishment of an operational earthquake forecasting system in Italy". *Seismological Research Letters*, 2014, 85, pp. 961–969

Marzocchi, W. and Zechar, J. D. (2011). "Earthquake forecasting and earthquake prediction: different approaches for obtaining the best model". *Seismological Research Letters*, 2011, 82, pp. 442–448

McGuire, R. and Arabasz, W. (1990). "An introduction to probabilistic seismic hazard analysis". *Geotechnical an Environmental Geophysics: Volume I: Review and Tutorial*. Society of Exploration Geophysicists, pp. 333–354.

McLachlan, G. J. and Krishnan, T. (2007). *The EM algorithm and extensions*. John Wiley & Sons.

Michael, A. J. and Werner, M. J. (2018). "Preface to the focus section on the Collaboratory for the Study of Earthquake Predictability (CSEP): New results and future directions". *Seismological Research Letters*, 2018, 89, pp. 1226–1228

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

Mignan, A, Werner, M., Wiemer, S, Chen, C.-C., and Wu, Y.-M. (2011). "Bayesian estimation of the spatially varying completeness magnitude of earthquake catalogs". *Bulletin of the Seismological Society of America*, 2011, 101, pp. 1371–1385

Milner, K. R., Field, E. H., Savran, W. H., Page, M. T., and Jordan, T. H. (2020). "Operational earthquake forecasting during the 2019 Ridgecrest, California, earthquake sequence with the UCERF3-ETAS model". *Seismological Research Letters*, 2020, 91, pp. 1567–1578

Milner, K. R., Shaw, B. E., Goulet, C. A., Richards-Dinger, K. B., Callaghan, S., Jordan, T. H., Dieterich, J. H., and Field, E. H. (2021). "Toward physics-based nonergodic PSHA: A prototype fully deterministic seismic hazard model for southern California". *Bulletin of the Seismological Society of America*, 2021, 111, pp. 898–915

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

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

Molchan, G. and Dmitrieva, O. (1992). "Aftershock identification: methods and new approaches". *Geophysical Journal International*, 1992, 109, pp. 501–516

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*, 2022, 32

Moon, T. K. (1996). "The expectation-maximization algorithm". *IEEE Signal processing magazine*, 1996, 13, pp. 47–60

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

Morrison, K (2017). "A gentle inla tutorial". *Precision Analytics*, 2017. URL: https://www.precision-analytics.ca/articles/a-gentle-inla-tutorial/

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

Nandan, S., Ouillon, G., Wiemer, S., and Sornette, D. (2017). "Objective estimation of spatially variable parameters of epidemic type aftershock sequence model: Application to California". *Journal of Geophysical Research: Solid Earth*, 2017, 122, pp. 5118–5143

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*, 2023, 9, pp. 1–19

Ng, S. K., Krishnan, T., and McLachlan, G. J. (2012). "The EM algorithm". *Handbook of computational statistics: concepts and methods*, 2012, pp. 139–172

Nievas, C. I., Bommer, J. J., Crowley, H., Elk, J. van, Ntinalexis, M., and Sangirardi, M. (2020). "A database of damaging small-to-medium magnitude earthquakes". *Journal of Seismology*, 2020, 24, pp. 263–292

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*, 2023, 128, pp. 1–28

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

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

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

— (1998). "Space-time point-process models for earthquake occurrences". *Annals of the Institute of Statistical Mathematics*, 1998, 50, pp. 379–402

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

— (2015). "Intermediate-term forecasting of aftershocks from an early aftershock sequence: Bayesian and ensemble forecasting approaches". *Journal of Geophysical Research: Solid Earth*, 2015, 120, pp. 2561–2578

Omi, T., Ogata, Y., Shiomi, K., Enescu, B., Sawazaki, K., and Aihara, K. (2016). "Automatic aftershock forecasting: A test using real-time seismicity data in Japan". *Bulletin of the Seismological Society of America*, 2016, 106, pp. 2450–2458

Omori, F. (1895a). "On after-shocks of earthquakes". *The journal of the College of Science, Imperial University of Tokyo*, 1895, 7, pp. 111–200

Omori, F. (1895b). "On the after-shocks of earthquakes". PhD thesis. The University of Tokyo.

O'Hagan, A., Murphy, T. B., and Gormley, I. C. (2012). "Computational aspects of fitting mixture models via the expectation–maximization algorithm". *Computational Statistics & Data Analysis*, 2012, 56, pp. 3843–3864

Pagani, M., Garcia-Pelaez, J., Gee, R., Johnson, K., Poggi, V., Silva, V., Simionato, M., Styron, R., Viganò, D., Danciu, L., et al. (2020). "The 2018 version of the global earthquake model: hazard component". *Earthquake Spectra*, 2020, 36, pp. 226–251

Page, M., Elst, N. van der, 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*, 2016, 106, pp. 2290–2301

Pan, J.-X., Fang, K.-T., Pan, J.-X., and Fang, K.-T. (2002). "Maximum likelihood estimation". *Growth curve models and statistical diagnostics*, 2002, pp. 77–158

Papadopoulos, A. N., Bazzurro, P., and Marzocchi, W. (2021). "Exploring probabilistic seismic risk assessment accounting for seismicity clustering and damage accumulation: Part I. Hazard analysis". *Earthquake Spectra*, 2021, 37, pp. 803–826

Peng, Z., Vidale, J. E., Ishii, M., and Helmstetter, A. (2007). "Seismicity rate immediately before and after main shock rupture from high-frequency waveforms in Japan". *Journal of Geophysical Research: Solid Earth*, 2007, 112

Ponnusamy, S and Vuorinen, M (2001). "Univalence and convexity properties for Gaussian hypergeometric functions". *The Rocky Mountain Journal of Mathematics*, 2001, pp. 327–353

Reasenberg, P. (1985). "Second-order moment of central California seismicity, 1969–1982". *Journal of Geophysical Research: Solid Earth*, 1985, 90, pp. 5479–5495

Reasenberg, P. A. and Jones, L. M. (1989). "Earthquake hazard after a mainshock in California". *Science*, 1989, 243, pp. 1173–1176

Reyes Canales, M. and Baan, M. van der (2020). "Are aftershock sequences pertinent to long-term seismic hazard assessments? Insights from the temporal ETAS model". *Journal of Geophysical Research: Solid Earth*, 2020, 125, e2019JB019095

Rhoades, D., Liukis, M, Christophersen, A, and Gerstenberger, M. (2016). "Retrospective tests of hybrid operational earthquake forecasting models for Canterbury". *Geophysical Journal International*, 2016, 204, pp. 440–456

Robert, C. P., Chopin, N., and Rousseau, J. (2009). "Harold Jeffreys's theory of probability revisited", 2009

Roberts, N. S., Bell, A. F., and Main, I. G. (2015). "Are volcanic seismic b-values high, and if so when?" *Journal of Volcanology and Geothermal Research*, 2015, 308, pp. 127–141

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

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

Rue, H., Martino, S., 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)*, 2009, 71, pp. 319–392

Santos, A., Figueiredo, E., Silva, M., Santos, R., Sales, C., and Costa, J. C. (2017). "Genetic-based EM algorithm to improve the robustness of Gaussian mixture mod-

els for damage detection in bridges". *Structural Control and Health Monitoring*, 2017, 24, e1886

Schneider, M. I. and Guttorp, P. (2021). *Bayesian ETAS: towards improved operational aftershock forecasting*. Tech. rep. USGS.

Schoenberg, F. P. (2010). "Introduction to point processes". *Wiley Encyclopedia of Operations Research and Management Science*, 2010

Schorlemmer, D, Mele, F, and Marzocchi, W (2010). "A completeness analysis of the National Seismic Network of Italy". *Journal of Geophysical Research: Solid Earth*, 2010, 115

Schorlemmer, D., Werner, M. J., Marzocchi, W., Jordan, T. H., Ogata, Y., Jackson, D. D., Mak, S., Rhoades, D. A., Gerstenberger, M. C., Hirata, N., et al. (2018). "The collaboratory for the study of earthquake predictability: Achievements and priorities". *Seismological Research Letters*, 2018, 89, pp. 1305–1313

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*, 2017, 122, pp. 449–469

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

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

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

Shebalin, P., Narteau, C., and Baranov, S. (2020). "Earthquake productivity law". *Geophysical Journal International*, 2020, 222, pp. 1264–1269

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

Sokolov, V. (2017). "Seismic hazard analysis based on maximum credible earthquakes". *Bulletin of Earthquake Engineering*, 2017, 15, pp. 1831–1852

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). "Bayesian measures of model complexity and fit". *Journal of the royal statistical society: Series b (statistical methodology)*, 2002, 64, pp. 583–639

Sridharan, R. (2014). "Gaussian mixture models and the EM algorithm". *Avilable in: http://people. csail. mit. edu/rameshvs/content/gmm-em. pdf*, 2014

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

Stein, S., Geller, R. J., and Liu, M. (2012). "Why earthquake hazard maps often fail and what to do about it". *Tectonophysics*, 2012, 562, pp. 1–25

Stindl, T. and Chen, F. (2023). "EM algorithm for the estimation of the RETAS model". *Journal of Computational and Graphical Statistics*, 2023

Stockman, S., Lawson, D., and Werner, M. (2023). "Forecasting the 2016–2017 central Apennines earthquake sequence with a neural point process". *Earth's Future*, 2023

Struben, J., Sterman, J., and Keith, D. (2015). "Parameter estimation through maximum likelihood and bootstrapping methods". *Analytical methods for dynamic modelers*, 2015, 1, pp. 3–38

Stupazzini, M., Infantino, M., Allmann, A., and Paolucci, R. (2021). "Physics-based probabilistic seismic hazard and loss assessment in large urban areas: A simplified application to Istanbul". *Earthquake Engineering & Structural Dynamics*, 2021, 50, pp. 99–115

Taroni, M., Marzocchi, W., Schorlemmer, D., Werner, M. J., Wiemer, S., Zechar, J. D., Heiniger, L., and Euchner, F. (2018). "Prospective CSEP evaluation of 1-day, 3-month, and 5-yr earthquake forecasts for Italy". *Seismological Research Letters*, 2018, 89, pp. 1251–1261

Telesca, L., Lovallo, M., Golay, J., and Kanevski, M. (2016). "Comparing seismicity declustering techniques by means of the joint use of Allan Factor and Morisita index". *Stochastic environmental research and risk assessment*, 2016, 30, pp. 77–90

Tsuruoka, H., Hirata, N., Schorlemmer, D., Euchner, F., Nanjo, K. Z., and Jordan, T. H. (2012). "CSEP Testing Center and the first results of the earthquake forecast testing experiment in Japan". *Earth, planets and space*, 2012, 64, pp. 661–671

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

— (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*, 1972, 3, pp. 379–441

Utsu, T. (1961). "A statistical study on the occurrence of aftershocks". *Geophys. Mag.*, 1961, 30, pp. 521–605

— (2002). "Statistical features of seismicity". *International geophysics series*, 2002, 81, pp. 719–732

Utsu, T. and Ogata, Y. (1995). "The centenary of the Omori formula for a decay law of aftershock activity". *Journal of Physics of the Earth*, 1995, 43, pp. 1–33

Veen, A. and Schoenberg (2008). "Estimation of space–time branching process models in seismology using an em–type algorithm". *Journal of the American Statistical Association*, 2008, 103, pp. 614–624

Von Toussaint, U. (2011). "Bayesian inference in physics". *Reviews of Modern Physics*, 2011, 83, p. 943

Wang, Q., Schoenberg, F. P., and Jackson, D. D. (2010). "Standard errors of parameter estimates in the ETAS model". *Bulletin of the Seismological Society of America*, 2010, 100, pp. 1989–2001

Watanabe, S. and Opper, M. (2010). "Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory." *Journal of machine learning research*, 2010, 11

Weatherill, G., Pagani, M, and Garcia, J (2016). "Exploring earthquake databases for the creation of magnitude-homogeneous catalogues: tools for application on a regional and global scale". *Geophysical Journal International*, 2016, 206, pp. 1652–1676

Wiemer, S. and Katsumata, K. (1999). "Spatial variability of seismicity parameters in aftershock zones". *Journal of Geophysical Research: Solid Earth*, 1999, 104, pp. 13135–13151

Wiemer, S. and Wyss, M. (2000). "Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan". *Bulletin of the Seismological Society of America*, 2000, 90, pp. 859–869

Williamson, L. D., Scott, B. E., Laxton, M. R., Bachl, F. E., Illian, J. B., Brookes, K. L., and Thompson, P. M. (2022). "Spatiotemporal variation in harbor porpoise distribution and foraging across a landscape of fear". *Marine Mammal Science*, 2022, 38, pp. 42–57

Woessner, J. and Wiemer, S. (2005). "Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty". *Bulletin of the Seismological Society of America*, 2005, 95, pp. 684–698

Yan, Y., Huang, H.-C., and Genton, M. G. (2021). "Vector autoregressive models with spatially structured coefficients for time series on a spatial grid". *Journal of Agricultural, Biological and Environmental Statistics*, 2021, 26, pp. 387–408

Zaliapin, I. and Ben-Zion, Y. (2013). "Earthquake clusters in southern California II: Classification and relation to physical properties of the crust". *Journal of Geophysical Research: Solid Earth*, 2013, 118, pp. 2865–2877

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., and Wong, H. (2008). "Clustering analysis of seismicity and aftershock identification". *Physical review letters*, 2008, 101, p. 018501

Zechar, J. D., Gerstenberger, M. C., and Rhoades, D. A. (2010a). "Likelihood-based tests for evaluating space–rate–magnitude earthquake forecasts". *Bulletin of the Seismological Society of America*, 2010, 100, pp. 1184–1195

Zechar, J. D., Marzocchi, W., and Wiemer, S. (2016). "Operational earthquake forecasting in Europe: Progress, despite challenges". *Bulletin of Earthquake Engineering*, 2016, 14, pp. 2459–2469

Zechar, J. D., Schorlemmer, D., Liukis, M., Yu, J., Euchner, F., Maechling, P. J., and Jordan, T. H. (2010b). "The Collaboratory for the Study of Earthquake Predictability perspective on computational earthquake science". *Concurrency and Computation: Practice and Experience*, 2010, 22, pp. 1836–1847

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

Zhang, L., Goda, K., Werner, M. J., and Tesfamariam, S. (2021). "Spatiotemporal seismic hazard and risk assessment of M9. 0 megathrust earthquake sequences of wood-frame houses in Victoria, British Columbia, Canada". *Earthquake Engineering & Structural Dynamics*, 2021, 50, pp. 6–25

Zhu, C., Cotton, F., Kawase, H., and Nakano, K. (2023). "How well can we predict earthquake site response so far? Machine learning vs physics-based modeling". *Earthquake Spectra*, 2023, 39, pp. 478–504

Zhuang, J., Ogata, Y., and Vere-Jones, D. (2002). "Stochastic declustering of space-time earthquake occurrences". *Journal of the American Statistical Association*, 2002, 97, pp. 369–380

— (2004). "Analyzing earthquake clustering features by using stochastic reconstruction". *Journal of Geophysical Research: Solid Earth*, 2004, 109

— (2006). "Diagnostic analysis of space-time branching processes for earthquakes". *Case Studies in Spatial Point Process Modeling*, 2006, pp. 275–292

Zuur, A., Ieno, E., and Saveliev, A. (2017). *Beginner's Guide to Spatial, Temporal and Spatial-temporal Ecological Data Analysis with R-INLA*. Highland Statistics, 362 pages. ISBN: 9780957174191.