b-Bayesian : The Full Probabilistic Estimate of b-value Temporal Variations for Non-Truncated Catalogs

Marine Laporte[1], Stephanie Durand[2], Thomas Bodin[3], Blandine Gardonio[4], and David Marsan[5]

[1]Université Claude Bernard Lyon 1
[2]Laboratoire de Géologie de Lyon
[3]CNRS, ENS Lyon, Université de Lyon-1
[4]LGL TPE
[5]STerre, Univ. Grenoble Alpes, CNRS, France

July 29, 2024

Abstract

The frequency/magnitude distribution of earthquakes can be approximated by an exponential law whose exponent (the so-called b-value) is routinely used for probabilistic seismic hazard assessment. The b-value is commonly measured using Aki's maximum likelihood estimation, although biases can arise from the choice of completeness magnitude (i.e. the magnitude below which the exponential law is no longer valid). In this work, we introduce the b-Bayesian method, where the full frequency-magnitude distribution of earthquakes is modelled by the product of an exponential law and a detection law. The detection law is characterized by two parameters, which we jointly estimate with the b-value within a Bayesian framework. All available data are used to recover the joint probability distribution. The b-Bayesian approach recovers temporal variations of the b-value and the detectability using a transdimensional Markov chain Monte Carlo (McMC) algorithm to explore numerous configurations of their time variations. An application to a seismic catalog of far-western Nepal shows that detectability decreases significantly during the monsoon period, while the b-value remains stable, albeit with larger uncertainties. This confirms that variations in the b-value can be estimated independently of variations in detectability (i.e. completeness). Our results are compared with those obtained using the maximum likelihood estimation, and using the b-positive approach, showing that our method avoids dependence on arbitrary choices such as window length or completeness thresholds.

b-Bayesian : The Full Probabilistic Estimate of b-value Temporal Variations for Non-Truncated Catalogs

M. Laporte[1], S. Durand[1], T. Bodin[1], B. Gardonio[1], and D. Marsan[2],

[1]LGL-TPE, Université Claude Bernard, Lyon, France
[2]Université de Savoie Mont Blanc, Isterre, Chambéry, France

Key Points:

- The b_{value} from the Gutenberg-Richter law is usually inferred by truncating earthquake catalogs above a completeness magnitude. We propose to use all the data available and invert conjointly for b_{value} and two parameters describing the detectability.
- Using a Bayesian framework we retrieve the full posterior density function of b_{value} with a more realistic evaluation of its uncertainties than classical approaches.
- b-Bayesian performs a transdimensional inversion to recover the temporal changes of the b_{value} and detectability.

Corresponding author: Marine Laporte, `marine.laporte@univ-lyon1.fr`

Abstract

The frequency/magnitude distribution of earthquakes can be approximated by an exponential law whose exponent (the so-called b_{value}) is routinely used for probabilistic seismic hazard assessment. The b-value is commonly measured using Aki's maximum likelihood estimation, although biases can arise from the choice of completeness magnitude (i.e. the magnitude below which the exponential law is no longer valid). In this work, we introduce the b-Bayesian method, where the full frequency-magnitude distribution of earthquakes is modelled by the product of an exponential law and a detection law. The detection law is characterized by two parameters, which we jointly estimate with the b_{value} within a Bayesian framework. All available data are used to recover the joint probability distribution. The b-Bayesian approach recovers temporal variations of the b_{value} and the detectability using a transdimensional Markov chain Monte Carlo (McMC) algorithm to explore numerous configurations of their time variations. An application to a seismic catalog of far-western Nepal shows that detectability decreases significantly during the monsoon period, while the b-value remains stable, albeit with larger uncertainties. This confirms that variations in the b_{value} can be estimated independently of variations in detectability (i.e. completeness). Our results are compared with those obtained using the maximum likelihood estimation, and using the b-positive approach, showing that our method avoids dependence on arbitrary choices such as window length or completeness thresholds.

1 Introduction

Classically, the probability density function of an earthquake of magnitude m above a magnitude of completeness M_c follows the Gutenberg-Richter law (Aki, 1965):

$$p(m) = \beta e^{-\beta(m-M_c)} \tag{1}$$

where $\beta = b_{value} \times \ln(10)$. The b_{value} is the seismic parameter that describes the relative number of large magnitude earthquakes versus smaller magnitude earthquakes. For global earthquake catalogues, the b_{value} is typically close to 1, but has been shown to vary in both space and time when focusing on earthquake catalogues for specific seismogenic regions or time periods (e.g. Wiemer & Wyss, 1997; Ogata & Katsura, 2006). With the rapid growth of seismological instruments and recording capabilities, there is a need for advanced statistical methods to analyze earthquake catalogs. Here, we analyse the distribution of earthquake magnitudes, focusing on the possibility to observe and interpret temporal variations in this distribution.

The Gutenberg-Richter law is widely used as an earthquake recurrence model for Probabilistic Seismic Hazard Assessment (PSHA) studies (e.g. Cornell, 1968; Drouet et al., 2020). Consequently, the accurate estimation of b_{value} and its uncertainties play a crucial role in the accuracy and robustness of seismic hazard estimates (e.g. Keller et al., 2014; Beauval & Scotti, 2004; Taroni & Akinci, 2020). For accurate hazard assessment, b_{value} biases due to the incompleteness of earthquake catalogues need to be addressed (e.g. Weichert, 1980; Plourde, 2023; Dutfoy, 2020; Beauval & Scotti, 2004) as well as possible temporal or spatial variations of the b_{value} (e.g. Beauval & Scotti, 2003; Yin & Jiang, 2023).

The physical interpretation of these spatio-temporal variations in the frequency-magnitude distribution of earthquakes has been a subject of ongoing debate for years (e.g. Mogi, 1962; Scholz, 1968; Carter & Berg, 1981; Herrmann et al., 2022). Based on observations from laboratory earthquake simulations, which are commonly used as analogues for studying natural earthquake behaviour, it has been proposed that the b_{value} is inversely related to the normal and shear stress applied to the fault (Scholz, 1968). At the scale of the seismic cycle, which is reproduced in stick-slip experiments with controlled stress and friction properties, the b_{value} has been observed to decrease linearly with stress build-up and to increase abruptly with the stress-drop release during earthquake rupture (Avlonitis & Papadopoulos, 2014; Goebel et al., 2017; Rivière et al., 2018; Bolton et al., 2020). Extending this observation to real earthquake systems is not straightforward because real earthquake catalogues contain additional uncertainties and the estimation of the actual state of the stress field is another inverse problem.

However, the b_{value} is also widely used to characterize real earthquake catalogs. It is commonly estimated to characterize earthquake clusters and discriminate between seismic swarms (e.g. De Barros et al., 2019). Some variations in b_{value} have been observed for different earthquakes depths or within different stress regimes (Mori & Abercrombie, 1997; Schorlemmer et al., 2005; Scholz, 2015; Petruccelli et al., 2019; Morales-Yáñez et al., 2022). Low b_{value} (< 0.8), associated to a larger number of larger magnitudes earthquakes compared to the normal regime, have been observed for several earthquake sequences occurring before a large earthquake rupture (e.g. Nanjo et al., 2012; Chan et al., 2012; H. Shi et al., 2018; Li & Chen, 2021; Van der Elst, 2021; Kwiatek et al., 2023; Wetzler et al., 2023). This observation has a major impact for the identification of precursory phases before large mainshocks. Recently, b_{value} monitoring has been proposed to serve as a stress-meter for discrimination of foreshock sequences (e.g. Gulia & Wiemer, 2019; Ito & Kaneko, 2023). This topic remains under debate due to large uncertainties that could arise either from earthquake catalogs or from b_{value} estimation approaches (e.g. Lombardi, 2021; Spassiani et al., 2023; Yin & Jiang, 2023; Geffers et al., 2022; Godano et al., 2024).

The most classical approach for estimating b_{value} from a catalog of earthquake magnitudes is the maximum likelihood estimation of Aki and its generalization (Aki, 1965; Utsu, 1966), which depends on the arbitrary choice of the completeness magnitude M_c :

$$\beta = \frac{1}{\overline{m} - M_c} \tag{2}$$

with $\overline{m}$ the mean of magnitudes greater than M_c. Using this formula, only events with magnitudes larger than M_c are used to estimate β.

Unnoticed changes in completeness over time are the main source of bias when studying b_{value} temporal variations (e.g. Woessner & Wiemer, 2005; Helmstetter et al., 2006; Mignan & Woessner, 2012; Lombardi, 2021; Plourde, 2023; Godano et al., 2023) Two main sources of incompleteness are usually identified (e.g. Lippiello & Petrillo, 2024) : (1) the background incompleteness coming from momentary changes in the detectability of the seismic network, and (2) the short-term aftershock incompleteness (STAI) which describes the short but large changes in completeness that occur during mainshock-aftershock sequences where large earthquakes mask smaller ones (e.g. Helmstetter et al., 2006; Hainzl & Fischer, 2002). The b-positive approach (Van der Elst, 2021) is a variant of Aki's maximum likelihood approach, using differences in the magnitudes of successive earthquakes to propose a moving-window estimate of temporal changes in the b_{value} during mainshock-aftershock sequences without being biased by STAI.

$$\beta = \frac{1}{\overline{m'}^{+} - dM_c} \tag{3}$$

with $\overline{m'}^{+}$ the mean of positive magnitudes differences greater than dM_c, which is a chosen value that should be greater than twice the minimum magnitude difference. Based on the fact that two consecutive events in a mainshock-aftershock sequence share the same completeness, this approach is now frequently used for a more accurate estimation of temporal variations of b_{value}.

Even though the b-positive approach provides a major advantage in comparison to Aki's classical approach, it still suffers from its dependence on the choice of dM_c and to the size of the moving-window (e.g. Lippiello & Petrillo, 2024). Furthermore, the b_{value} estimate is computed on less than half of the available data and uncertainties are usually assessed using a bootstrap approach (Van der Elst, 2021).

In this paper, we introduce the b-Bayesian approach to explore the temporal variation of b_{value}, while addressing the problems of classical frequentist approaches. We propose to invert for b_{value} using the entire catalogue, taking into account a detectability function. By adopting this approach, our results are independent of the arbitrary choice of a completeness magnitude. Instead of traditional methods that compare frequency-magnitude distributions over random data subsets or that recover pseudo-continuous temporal variations using moving time windows, we address temporal variations in b_{value} and detectability by considering the number and positions of temporal discontinuities where b_{value} or detectability changes. The inversion of temporal discontinuities is achieved using a transdimensional framework.

Transdimensional inversion is commonly used in seismic tomography to allow the data to determine the level of spatial complexity in the recovered tomographic model (e.g. Bodin & Sambridge, 2009; Bodin et al., 2012). It has recently been adapted to estimate variations in the b_{value} from truncated catalogs along one dimension, such as time or depth (Morales-Yáñez et al., 2022). Here, we use transdimensional inversion to recover one-dimensional partitions of the entire earthquake dataset. A Bayesian framework provides a global formulation of the inverse problem and allows for the probabilistic estimation of temporal changes of b_{value}, and detectability. The complexity of the model does not depend on any arbitrary parameter, but is determined by the complexity of the data.

This paper is organized as follows. First, we describe the novel b-Bayesian approach for a time-invariant case, including the assessment of detectability using all available magnitude data. We describe how we extend this approach using a transdimensional framework in order to invert for temporal variations based on the complexity of the data. We present the results obtained using a synthetic catalog generated to mimic real-world scenarios. A first application of the b-Bayesian method is presented using a real earthquake catalog. We compare the results of b-Bayesian with the two frequentist approaches : the maximum likelihood estimate and the b-positive to describe the temporal variation of b_{value} and the temporal variations of detectability for an earthquake catalog of far-western Nepal spanning two years of microseismicity.

2 Method : A Bayesian framework

In this study, a dataset is an earthquake catalogue which corresponds to a set of N observations of (non-discrete) magnitudes m_i $(i = 1 \dots N)$ that we note :

$$\mathbf{d} = [m_1, m_2, \dots m_N] \tag{4}$$

From hereafter, we refer to conditional probabilities using $p(a|b)$. We know from the Gutenberg-Richter law (eq.1) that the probability density of observing one earthquake i of magnitude $m_i \geq M_c$ for a given β is :

$$p(m_i|\beta) = \beta e^{-\beta(m_i - M_c)} \tag{5}$$

and is zero if $m_i < M_c$. Then, assuming that the magnitudes of the seismic events are independent, we can write the probability of observing the entire earthquake dataset $\mathbf{d}$ with $m_i \geq M_c$, $p(\mathbf{d}|\beta)$, as :

$$p(\mathbf{d}|\beta) = \prod_{i=1}^{N} p(m_i|\beta) = \beta^N e^{-\beta N(\overline{m_i} - M_c)} \tag{6}$$

where $\overline{m_i}$, is the mean magnitude of events with $m_i \geq M_c$ and M_c, the magnitude of completeness. Note here that the value of β that maximizes (eq.6) is the maximum likelihood solution given by Aki's formula in (eq.2).

2.1 The temporally invariant case

In practise, seismic catalogs are truncated at the completeness magnitude M_c in order to avoid biases due to the detection capacity (e.g. Aki, 1965; Utsu, 1966). As a result, the classical approach for b-value estimation strongly depends on the choice made for M_c. Various methods have been developed for assessing the completeness magnitude (e.g. Ringdal, 1975; Ogata & Katsura, 1993; Woessner & Wiemer, 2005; Mignan & Woessner, 2012) or to correct the dataset for its temporal variations (e.g. Helmstetter et al., 2006; Cao & Gao, 2002; Chan et al., 2012). However, defining such a completeness magnitude always implies to ignore a significant portion of a dataset that may contain valuable information about the statistics of seismicity. Here instead, we propose to analyse the entire dataset by modelling the entire frequency-magnitude distribution of earthquakes. To do so, the Gutenberg-Richter law is modulated by a detection law q(m) such that now :

$$p(m_i|\beta) = \frac{1}{K} q(m_i)\beta e^{-\beta m_i} \tag{7}$$

where $q(m)$ defines the probability density of detecting an event, and K a constant to insure that the probability distribution integrates to one :

$$\int_{M_{min}}^{\infty} p(m_i|\beta) dm = 1 \tag{8}$$

with M_{min} the smallest earthquake magnitude in the catalog.

In this way, the probability of observing an event is the product of the probability of occurrence (given by the Gutenberg Richter law) and the probability of detection (given by the detection law $q(m)$ that varies from 0, no detection, to 1, 100% detection).The error function has been proposed in the literature to represent the probability of detection of an event in the presence of log-normal seismic noise (e.g. Ringdal, 1975; Ogata & Katsura, 1993; Daniel et al., 2008). The error function (see Figure 1.A) depends on two parameters μ and σ such as :

$$q(m) = \frac{1}{2} + \frac{1}{2} erf\left(\frac{m-\mu}{\sqrt{2}\sigma}\right) \tag{9}$$

where μ represents 50% of probability of detection for an earthquake of magnitude $m = \mu$, and becomes 84% for $m = \mu + \sigma$. The magnitude of completeness is the equivalent of the 84% threshold. This function fits adequately the frequency-magnitude distribution for a variety of cases (Ogata & Katsura, 1993; Woessner & Wiemer, 2005).

From equations (7) and (8), we can write :

$$K = \int_{M_{min}}^{\infty} q(m)\beta e^{-\beta(m-M_{min})} dm \tag{10}$$

Fortunately, this integral for the error function q(m) in (eq.9) has a closed form solution between M_{min} and infinity :

$$K = q(M_{min}) + \left(1 - q(M_{min} + \beta\sigma^2)\right) \exp\left(\frac{\beta^2\sigma^2}{2} - \beta(\mu - M_{min})\right) \tag{11}$$

Now assuming that magnitudes are independent, the probability of observing a full dataset **d** is :

$$p(\mathbf{d}|\omega) = \prod_{i=1}^{N} p(m_i|\omega) = \left(\frac{\beta}{K}\right)^N \prod_{i=1}^{N} q(m_i) \times e^{-\beta N(\overline{m_i} - M_{min})} \tag{12}$$

where $\omega = [\beta, \mu, \sigma]$ is our set of three unknown model parameters. Our goal here is to estimate these parameters from a set of realizations **d**. This is an inverse problem that can be formulated in a Bayesian framework, where the posterior solution $p(\omega|\mathbf{d})$ is the product between the model priors and the likelihood function $p(\mathbf{d}|\omega)$ (eq.12).

$$p(\omega|\mathbf{d}) = p(\beta, \mu, \sigma) p(\mathbf{d}|\omega) \tag{13}$$

Here we set independent uniform prior distributions for the three parameters, partly because they are not related to the same physics: b_{value} is related to seismicity and μ and σ to network detectability. Although we can expect correlations between these parameters from the data (i.e. *a posteriori*), our level of knowledge is independent for each parameter. This independence greatly facilitates Bayesian inference. For each parameter, we use a simple uniform prior distribution, independently defined between a fixed range of realistic values. The choice of the bounds is guided by the literature and should be chosen depending on the seismotectonic context and the mean detectability of the network. We advise to impose a relatively wide range of values for the b_{value} inference, allowing both high (> 1) and low (< 1) b_{value} for an earthquake catalog. In the context of geothermal or volcanic activity, this range may be extended to allow larger values (up to 2.5) for the b_{value}. The choice of bounds for the μ parameter should be guided by the level of detectability of the seismic network and the "expected" variations in completeness. For a local network (seismicity included within 50 km), which essentially records microseismicity, we can set this range of values between 0.5 and 2. In the presence of at least one mainshock/aftershock sequence, this range should also be increased. The prior distribution on σ can be set between 0.01 and 0.5 and does not need to be adjusted depending on the context.

The posterior distribution (eq.13) can be numerically approximated using a classical Monte-Carlo approach.

As an example, we construct a synthetic dataset of 4460 independent magnitudes, randomly drawn from a Gutenberg-Richter law characterized by $b_{value} = \frac{\beta}{log(10)} = 0.9$

and modulated by an error detection function characterized by μ=0.75 and σ=0.34 (see Figure 1.A). We then estimate our set of parameters $\omega = [\beta, \mu, \sigma]$ from the catalogue, by approximating the posterior distribution $p(\omega|\mathbf{d})$ with a standard Monte Carlo scheme by randomly sampling the model priors. The resulting 3D posterior density function can be projected onto each parameter (Figure 1.B) to derive 1D and 2D marginal distributions (Figure 1.C). For example, the marginal distribution for β is simply obtained by integrating the posterior over μ, and σ :

$$p(\beta|\mathbf{d}) = \int\int p(\beta, \mu, \sigma|\mathbf{d})\, d\mu\, d\sigma \tag{14}$$

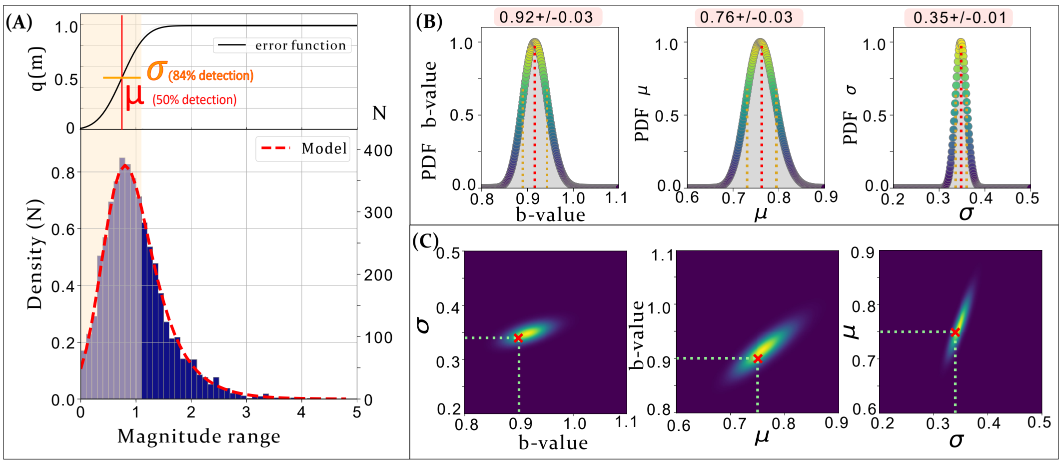

Figure 1. (A) [Top] Detection function q(m) associated with the synthetic dataset [Bottom] Frequency-magnitude distribution of a synthetic catalog with values of $b_{value} = 0.9$, $\mu = 0.75$ and $\sigma = 0.34$. The red dotted line corresponds to the model $\omega = [0.92, 0.76, 0.35]$ that best fits the observations. The yellow area corresponds to the distribution of magnitudes inferior to $(\mu + \sigma)$ that are usually removed by classical approaches (about 60% of available data). (B) Marginals distributions of the posterior function: $p(b_{value}|\mathbf{d})$, $p(\mu|\mathbf{d})$ and $p(\sigma|\mathbf{d})$ from left to right, respectively. Here, posterior functions are normalized by their maximum. (C) 2D marginal distributions of the posterior function. True values are represented by the red cross.

Note that the 2D marginals are useful to show the correlations between pairs of parameters. Uncertainty estimates of the three parameters can be obtained with the 1σ confidence interval.

Compared to optimization approaches where only the best fitting (i.e. maximum likelihood) ω is obtained, our method provides the 3D posterior density distribution, $p(\omega|\mathbf{d})$, from which parameter correlations and uncertainties can be estimated. Moreover, the dataset is no longer truncated above a completeness magnitude, instead, the full frequency-magnitude distribution is now used to jointly invert for b_{value} and detectability.

2.2 Temporal variations of b-value

2.2.1 A transdimensional parametrization

Going one step further, we now consider that ω can vary with time and our goal is to recover the location of temporal changes. Our three parameters in ω are considered constant in periods separated by abrupt changes (see Figure 3). To that aim, temporal variations are modeled with a set of discontinuities $\mathbf{T}$:

$$\mathbf{T} = [T_1, T_2, \dots T_k] \tag{15}$$

where, k is the number of temporal discontinuities and T_j $(j = 0 \dots k)$ the times at which the frequency-magnitude distribution changes. The unknown models vectors of the time varying frequency-magnitude distribution will be denoted :

$$\mathbf{\Omega} = [\omega_1, \omega_2, \dots \omega_{k+1}] \tag{16}$$

where $\omega_j = [\beta_j, \mu_j, \sigma_j]$ is the local model predicting the sub-dataset d_j between two discontinuities $[T_{j-1}, T_j]$. Note that $T_0 = min(t_{obs})$ and $T_{k+1} = max(t_{obs})$. The full posterior solution $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ describes the joint probability for local models $[\omega_1, \omega_2, \dots \omega_{k+1}]$ predicting events between each temporal discontinuities $[T_1, T_2, \dots T_k]$ of the temporal model, $\mathbf{T}$. Since the dimension of the model varies with the number of discontinuities, k, which is unknown, the inverse problem is so-called transdimensional. The posterior $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ does not have an analytical solution but can be sampled with a Monte Carlo algorithm. In this work, we propose to isolate the part of the posterior solution that is transdimensional (and to sample it with an appropriate algorithm), and to separate it from a part where the dimension is fixed.

That is, the full posterior solution $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ can be developed as a product of a conditional term $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T})$ and a marginal term $p(\mathbf{T}|\mathbf{d})$:

$$p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d}) = p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T}) \times P(\mathbf{T}|\mathbf{d}) \tag{17}$$

The following sections describe each of these terms in detail and how they can be approximated.

2.2.2 The conditional posterior $p(\boldsymbol{\Omega}|\boldsymbol{d}, \boldsymbol{T})$

The conditional term $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T})$ describes the probability distribution for parameters $\mathbf{\Omega}$ for a given time partition $\mathbf{T}$ (Figure. 2). It can be itself decomposed with the Bayes theorem into the product of a likelihood distribution and a prior distribution :

$$p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T}) = p(\mathbf{d}|\mathbf{\Omega}, \mathbf{T}) \times p(\mathbf{\Omega}|\mathbf{T}) \tag{18}$$

Since all magnitudes are independent, the likelihood is the product of likelihoods for ev-

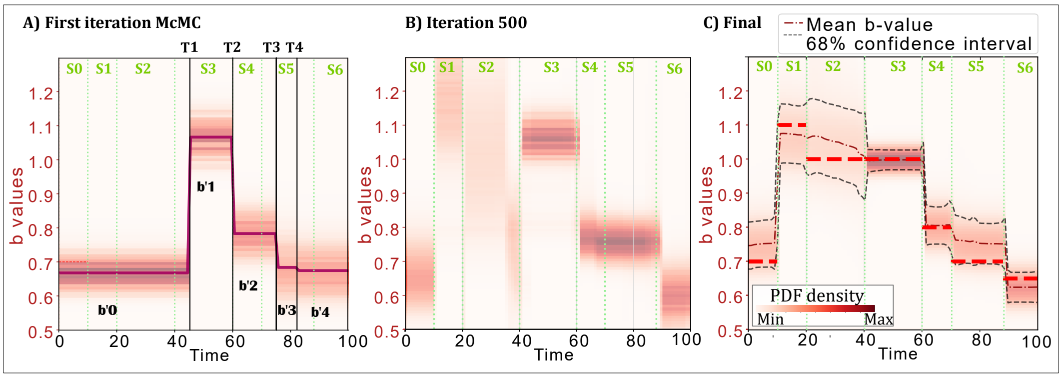

Figure 2. (A) Initial iteration of the Markov chain : plot of the conditional probability $p(\mathbf{d}|b_{value}, \mathbf{T})$, the probability of b_{value} for a fixed time model $\mathbf{T}$ of dimension $k = 4$. The 4 black vertical lines are the discontinuities of the proposed time model $\mathbf{T}$. The 3D posterior density function is computed for each data subset $T_j (j = 0 \ldots 4)$. The bold line is the mean posterior density function of β_j over time. Here, the synthetic earthquake dataset has been constructed to represent 6 discontinuities. The six green vertical dashed lines are the theoretical discontinuities. (B) Iteration 500 of the McMC: preliminary result of the b_{value} time variations, sum of the marginal density functions of the accepted models. The six discontinuities are almost all retrieved. (C) Final iteration of the Markov Chain : sum of marginal density functions of the totality of accepted models. The final temporal evolution of b_{value} fits the true b_{value} of the synthetic dataset which are represented in bold dashed horizontal lines.

ery sub-dataset $\mathbf{d}_j$ given by the temporal model $\mathbf{T}$:

$$p(\mathbf{d}|\mathbf{\Omega}, \mathbf{T}) = \prod_{j=1}^{k+1} p(\mathbf{d}_j|\mathbf{\Omega}, \mathbf{T}) \tag{19}$$

And since the magnitudes of events $\mathbf{d}_j$ occurring between two discontinuities $[T_{j-1}, T_j]$ only depend on the local parameters $\omega_j = [\beta_j, \mu_j, \sigma_j]$, we can write the likelihood :

$$p(\mathbf{d}|\mathbf{\Omega}, \mathbf{T}) = \prod_{j=1}^{k+1} p(\mathbf{d}_j|\omega_j, \mathbf{T}). \tag{20}$$

where $p(\mathbf{d}_j|\omega_j, \mathbf{T})$ is the likelihood of the data within a time period j which is simply given by equation (eq.12).

The prior distribution for $\mathbf{\Omega}$ given a fixed temporal model $\mathbf{T}$, $p(\mathbf{\Omega}|\mathbf{T})$ from equation (eq.18), is chosen to be the same within each partition $p(\omega_j|\mathbf{T})$, and simply corresponds to the uniform prior distribution used in the temporally invariant case (eq. 13). Thus, the conditional posterior $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T})$ is easy to sample as different periods j can be independently sampled with the same algorithm described in the previous section and used to produce results in Figure 1. Therefore, for any partition of the time $\mathbf{T}$, we know how to probabilistically estimate the parameters $\mathbf{\Omega}$.

The question now is to estimate the number and the position of discontinuities $\mathbf{T}$. This is given by the marginal posterior $p(\mathbf{T}|\mathbf{d})$.

2.2.3 The marginal posterior $p(\mathbf{T}|\mathbf{d})$

$p(\mathbf{T}|\mathbf{d})$ describes the probability of the time partition $\mathbf{T} = [T_0, T_2, \dots T_k]$ given the full dataset of observed magnitudes. It can be obtained by integrating the full posterior $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ over the parameters $\mathbf{\Omega} = [\omega_1, \omega_2, ..., \omega_{k+1}]$:

$$p(\mathbf{T}|\mathbf{d}) = \int_{\mathbf{\Omega}} p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d}) d\Omega \tag{21}$$

According to Bayes' rule, the posterior density function $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ is proportional to the product of the likelihood $p(d|\mathbf{\Omega}, T)$ times the joint prior $p(\mathbf{\Omega}, \mathbf{T})$.

$$p(\mathbf{T}|\mathbf{d}) \propto \int_{\mathbf{\Omega}} p(\mathbf{d}|\mathbf{\Omega}, \mathbf{T}) p(\mathbf{\Omega}, \mathbf{T}) d\Omega \tag{22}$$

The joint prior $p(\mathbf{\Omega}, \mathbf{T})$ can be decomposed according to the property of joint density distributions $p(\mathbf{\Omega}, \mathbf{T}) = p(\mathbf{T}) \times p(\mathbf{\Omega}|\mathbf{T})$. Applied to equation (eq.22), we get :

$$p(\mathbf{T}|\mathbf{d}) \propto p(\mathbf{T}) \times \int_{\Omega} p(\mathbf{d}, \mathbf{\Omega}|\mathbf{T}) p(\mathbf{\Omega}|\mathbf{T}) d\Omega \tag{23}$$

with $p(\mathbf{T})$, the prior distribution for the time partitions and $p(\mathbf{\Omega}|\mathbf{T})$, the prior for the models $\mathbf{\Omega}$ given a fixed temporal model which is also present in the conditional posterior (eq.18). The prior $p(\mathbf{T})$ is a joint distribution, where the prior for each discontinuity T_j is given by a uniform distribution bounded between $min(t_{obs})$ and $max(t_{obs})$. The prior $p(\mathbf{\Omega}|\mathbf{T})$ corresponds to the model priors described in the temporally-invariant case as described for the section above.

Considering that the sub-datasets $textbfd_j$ of a time model $\mathbf{T}$ are independent, and since dataset $\mathbf{d}_j$ only depends on parameters ω_j the full posterior can be expressed as :

$$p(\mathbf{T}|\mathbf{d}) \propto p(\mathbf{T}) \times \prod_{j=1}^{k+1} \left(\int_{\omega_j} p(\mathbf{d}_j|\omega_j, \mathbf{T}) p(\omega_j|\mathbf{T}) d\omega_j \right) \tag{24}$$

where, $p(\mathbf{d}_j|\omega_j, \mathbf{T})$ is the likelihood of observing the subset $\mathbf{d}_j$ between $[T_j, T_{j+1}]$ according to the local model ω_j and can be estimated using equation (eq.12) obtained in the time-invariant case.

These integrals can be estimated using importance sampling. That is, for a large number of realizations x_i, $i = (1, \ldots, N)$, randomly drawn from a distribution p(x) :

$$\int p(x) f(x) dx \approx \frac{1}{N} \times \sum_{i=1}^{N} f(x_i) \tag{25}$$

Applied to (eq.25), we have:

$$p(\mathbf{T}|\mathbf{d}) \propto p(\mathbf{T}) \times \left(\frac{1}{N_\omega} \right)^{(k+1)} \prod_{j=0}^{k} \left(\sum_{i=1}^{N_\omega} p(d_j|\omega_{j_{(i)}}, \mathbf{T}) \right) \tag{26}$$

where for each period j, $\omega_{j_{(i)}} = [\beta_j, \mu_j, \sigma_j]_{(i)}$ for $i = (1, \ldots, N_\omega)$ are a set of realizations randomly drawn from the uniform prior distributions $p(\omega_j|\mathbf{T})$.

From equation (eq.26) we see that the marginal posterior is proportional to the product of the prior at temporal discontinuities $p(\mathbf{T})$ and the product of the mean likelihoods $p(d_j|\omega_{j_{(i)}}, \mathbf{T})$, computed over N_w realisations of the model priors, between each temporal discontinuity. In this way, adding an extra discontinuity will be valuable only if it sufficiently increases the local likelihood $p(d_j|w_{j_{(i)}}, \mathbf{T})$ to counterbalance this first effect.

Therefore, this methodology based on a Bayesian framework inherently follows the principle of parsimony, finding a balance between finding a simple model with a low number of temporal discontinuities, k, and maximizing the overall likelihood $p(\mathbf{d}|\Omega, \mathbf{T})$.

2.2.4 The reversible-jump Markov-chain Monte-Carlo algorithm (rj-McMC)

The marginal posterior $p(\mathbf{T}|\mathbf{d})$ can be numerically approximated with equation (eq.26) but only for a given partition $\mathbf{T}$. One way to estimate the full distribution $p(\mathbf{T}|\mathbf{d})$ is through a Monte Carlo exploration over the space of temporal discontinuities $\mathbf{T}$. The solution is then a large ensemble of partition vectors $\mathbf{T}^l(l = 1 \dots N_l)$, with N_l the number of realizations $\mathbf{T}^l$, whose distribution approximates the target solution $p(\mathbf{T}|\mathbf{d})$.

As the dimension of $\mathbf{T}$ varies with the number of discontinuities k, $p(\mathbf{T}|\mathbf{d})$ is a trans-dimensional function and cannot be explored using a standard Metropolis algorithm (Metropolis et al., 1953; Hastings, 1970). One of the most popular technique for exploring a trans-dimensional posterior is the rj-McMC method (e.g. Green, 1995; Sambridge et al., 2006, 2013) and more specifically the birth-death McMC algorithm (e.g. Geyer & Møller, 1994). The rj-McMC algorithm, used in many geophysical inverse problems (e.g. Gallagher et al., 2009, 2011; Bodin et al., 2012), allows both the model parameters and the model dimension (i.e. the number of parameters) to be inferred. The rj-McMC follows the general principles of the McMC approach by generating samples from the target distribution. A Markov chain follows a random walk, where at each step, a proposed model $\mathbf{T}^{(p)}$ is generated by randomly modifying a current model $\mathbf{T}^{(c)}$ (Figure 3). This proposed model is then either accepted (and replaces the current model) or rejected. In this way, each step of the rj-McMC is a part of a chain converging to the target distribution.

The convergence is considered sufficient by monitoring the evolution of the number of discontinuities towards a stable value and when the rate of accepted models falls in the range of 20% to 40%. Details about the algorithm are given in Appendix. We also refer the reader to (Bodin et al., 2012) for further details.

2.2.5 Appraising the full posterior distribution $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$

As a reminder, the solution to our inverse problem is the full posterior solution $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ that describes the temporal changes of β, μ, and σ. As shown in equation (eq.17), this posterior can be written as the product of the marginal distribution $p(\mathbf{T}|\mathbf{d})$ describing the probability of temporal changes and a conditional distribution $p(\mathbf{\Omega}|\mathbf{T}, \mathbf{d})$ for the parameters of the frequency-magnitude distribution, given a set of temporal changes.

By decomposing in such a way the posterior distribution into a conditional and a marginal distribution, the Metropolis-Hastings rule is simplified by only simulating a trans-dimensional temporal point process for vector $\mathbf{T}$ (e.g. Geyer & Møller, 1994; Green, 1995). With the rj-McMC algorithm, we have a numerical way to approximate the marginal probability distribution about the number and position of temporal changes $p(\mathbf{T}|\mathbf{d})$ (eq.26).

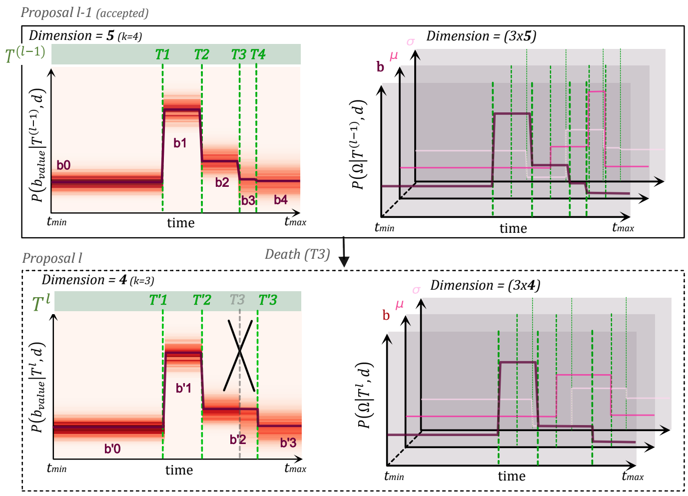

Figure 3. (Top)Left, $P(\beta|\mathbf{T}_{l-1}, \mathbf{d})$ over the 5 temporal segments ($k = 4$) of the proposed temporal model $\mathbf{T}_{l-1}$ for proposition $l-1$ of the rj-McMC. On the right, the mean likelihood over the 5 temporal segments for the three marginals posterior $p(\beta|\mathbf{T}_{l-1}, \mathbf{d})$, $P(\mu|T_{l-1}, \mathbf{d})$ and $P(\sigma|\mathbf{T}_{l-1}, \mathbf{d})$. (Bottom) New proposal model T_l in case of a death proposition of the rj-McMC. The two figures are the same as above : $P(\beta|\mathbf{T}_l, \mathbf{d})$ and the marginals computed for the proposed temporal model with a lower dimension (k=3).

In addition, for each sampled temporal model $\mathbf{T}^l$ proposed at iteration l of the rj-McMC, we are able to easily compute the conditional probability $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T}^l)$: the probability distribution of β, μ, and σ for the given model $\mathbf{T}^l$ (eq. 20).

At the completion, the full distribution for β, μ, and σ can therefore be obtained by summing the all the distributions $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T}^l)$ for all the sampled models $\mathbf{T}^l \in T^{(c)}$ (Figure 2.B,C). In practice, at each time-bin over an arbitrarily fine grid, the full probability distribution of β is the sum of all the marginal densities at this time over the ensemble solution for $\mathbf{T}$(Figure 2.B,C). In this way, the mean and the standard deviation for our three parameters can be obtained as a smooth function of time (see Figure 2.C).

2.3 Synthetic test

2.3.1 Generated Data

We simulate a synthetic earthquake catalog of 5683 independent events following frequency-magnitude distributions as realistic as possible with some temporal variations in b_{value} and detectability. In this section, we only consider a dataset with abrupt and discontinuous changes in the three parameters of the frequency-magnitude distribution. More precisely, the earthquake catalog is generated as follows :

- A discontinuity corresponds to the time when at least one of the three parameters of the frequency-magnitude distribution changes. We generate a catalogue with six temporal discontinuities that we aim to recover. Therefore, the catalog is the combination of seven temporal subsets.
- Within each temporal sub-dataset, earthquake magnitudes are drawn from a Gutenberg-Richter law characterized by a b_{value} specified in the Table 1.
- Earthquake occurrence is generated according to a basic epidemic-type aftershock sequence (Ogata, 1988) with a constant background rate. Each generated earthquake can be followed by aftershocks according to the aftershock productivity law (Utsu, 1972). Aftershock occurrence time is modelled by the Omori power law (Omori, 1894). The ETAS parametrization does not vary temporally. In particular, to characterise the temporal occurrence of aftershocks, we keep a constant p_{etas} value of 1.1 and an α_{etas} value of 1.5. For now, we do not generate the detectability variations coming from the short-term incompleteness (e.g. Ogata & Katsura, 2006; Helmstetter et al., 2006) following large earthquakes.
- We thin this ETAS earthquake catalog using the error detection law. Each event has a probability of being detected and preserved, or undetected and removed, according to its magnitude and some chosen μ and σ (see eq.9). Each pair of μ and σ for each of the seven sub-datasets are specified in the Table 1.

The dataset is made to test the capabilities of our algorithm and approach. To that aim, there are some periods where b_{value} remains constant while the detectability varies (S2 and S3) (Figure 4). We also include small detectability or b_{value} changes (ex:S5 and S6). For the period with the highest detectability (S3), earthquakes with magnitudes as low as 0.01 are detected. Considering a definition of $M_c = \mu + \sigma$, the completeness magnitude varies at most between subsets from 1.7 to 0.65.

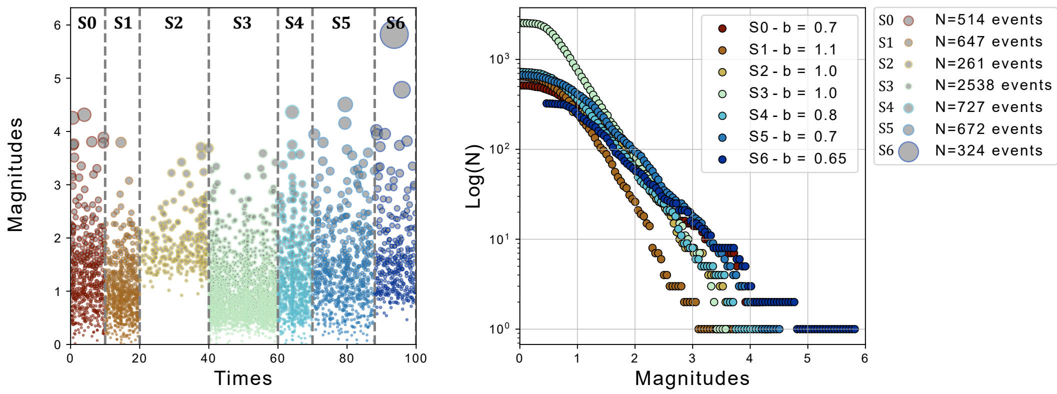

Figure 4. (A) Temporal distribution of magnitudes over time for the 7 data subsets. Each subset is represented with a different color. Grey vertical lines are the positions of the fixed temporal discontinuities. (B) Frequency-magnitude distribution of each synthetic data subset in logarithmic scale, respective colors from (A) are conserved. The slopes are related to the respective b_{value}.

Subset	S0	S1	S2	S3	S4	S5	S6
N_{events}	514	647	261	2538	727	672	324
b_{value}	0.70	1.10	1.00	1.00	0.80	0.70	0.65
μ	0.80	0.80	1.50	0.50	0.75	0.90	0.90
σ	0.35	0.30	0.20	0.15	0.30	0.40	0.15

Table 1. Set of chosen values for b_{value}, μ and σ for each of the 7 data subsets. N_{events} corresponds to the number of earthquakes detected of each sub-set (after the detection thinning).

2.3.2 Results and comparison

Let us start by applying classical approaches on this synthetic dataset. Ignoring temporal variations of b_{value} and considering a constant completeness magnitude of 1.8, the frequentist approach (Aki, 1965) gives the maximum likelihood estimate over the complete catalogue of $b_{value} = 0.81 \pm 0.03$. This illustrates how the uncertainties can be underestimated when using the classical approach without considering temporal variations. However, if we assume that we know the position of temporal discontinuities and that the completeness magnitude is correctly estimated using $M_{ci}^{true} = \mu_i + \sigma_i$ for $(i = 0, \ldots, 6)$, the true b_{value} can be recovered by the maximum likelihood estimate within its uncertainties. However, for applications to real earthquake catalogues, the temporal discontinuities are mostly unknown or at least ambiguous. Classical approaches therefore deal with temporal variations by estimating the b-value over a moving window.

Using a moving-window of 600 events and a constant completeness magnitude, Figure (5.A) shows that depending on the temporal sub-dataset the maximum likelihood estimate tends to under-estimate or over-estimate the b_{value} depending on the choice of M_c and the true detectability. In particular for (S2) with a low detectability, the choice of the completeness magnitude has a major influence on the results. This large variabil-

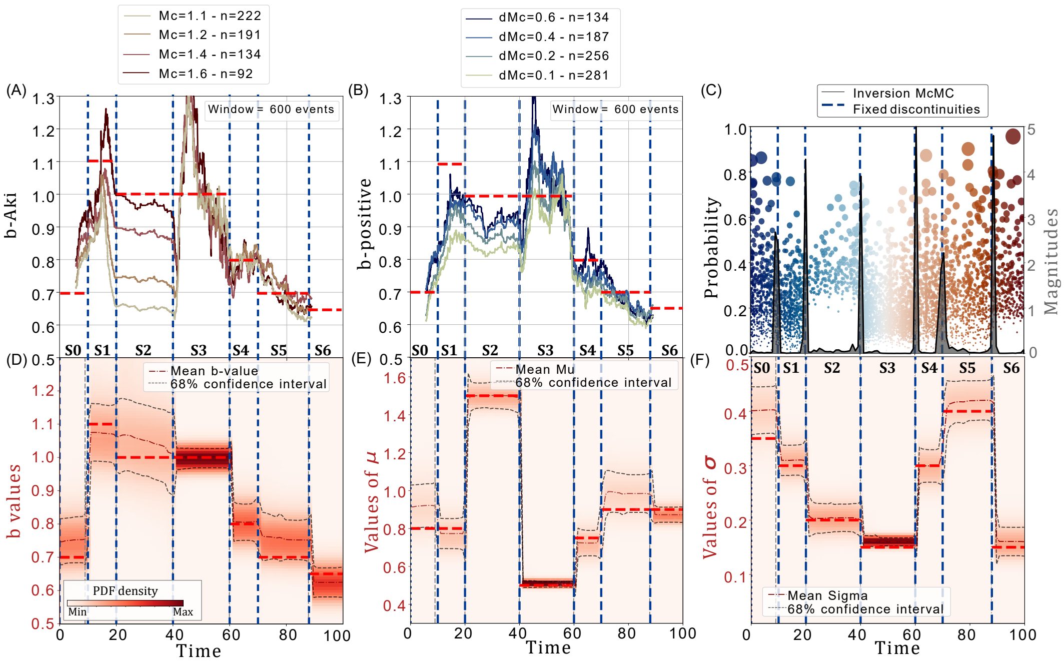

Figure 5. (A)Temporal variations of b_{value} using the classical maximum likelihood estimate (Aki, 1965; Utsu, 1966). The continuous curves are the moving-window estimate for a window size of 600 events with different choices of completeness magnitudes depending on the color-scale. (B) Temporal variations of b_{value} using the b-positive approach (Van der Elst, 2021). The continuous curves are the moving-window estimate for a window size of 600 events with different choices of difference threshold dM_c depending on the color-scale. (C) Probability of temporal discontinuities at the completion of the rj-McMC algorithm and the temporal distribution of magnitudes in the background (D) Marginal density distribution of b_{value} over time (E) Marginal density distribution of μ over time (F) Marginal density distribution of σ over time. For each subplot, the red horizontal lines represent the true parameter values fixed for each data subset, while the blue vertical lines denote the true values of temporal discontinuities as per Table 1.

ity of solutions can be partly corrected by the use of the b-positive approach (Van der Elst, 2021) (5.B). The b-positive approach (Van der Elst, 2021) uses moving-windows to infer temporal variations of b_{value} without being biased by the continuous decay of completeness for mainshock-aftershocks sequences. However, we show, that for discontinuous changes of completeness, the b-positive depends on the choice of the difference threshold dM_c. Thus, dM_c also needs to be adapted over time to correct for large variations in 'background' incompleteness, such as those between (S2) and (S3), to mitigate the risk of over-interpreting some temporal variations. This observation has also been highlighted by several recent studies (e.g. Lippiello & Petrillo, 2024).

These methods can be very efficient and fast but they use truncated datasets, which can result in b_{value} estimations based on very few events for periods with low detectability. With our new transdimensional approach, b_{value} and completeness are jointly inferred together with a probabilistic estimate of temporal changes from the entire dataset.

We applied our method to the synthetic dataset and conducted 50 parallel rj-McMC explorations with different initialisations for $\mathbf{T}$ in order to efficiently explore the model space. We present the results derived from the final stack of posterior densities coming from the 50 runs, each encompassing 5000 proposed models of temporal discontinuities (Figure 5.C,D,E,F). In total, 250000 temporal models were proposed out of which approximately 41800 were accepted upon completion. The six temporal discontinuities are well retrieved and clearly identified (Figure 5.C). Analysis of the number of discontinuities at each step indicates that the algorithm converged towards the correct value in less than 2000 iterations, even though all initialisations began with random values of k ranging from 4 to 12. A burn-in period of 1000 iterations has been set to disregard initial iterations. The maximum number of discontinuities allowed was set to 40 and did not affect the random-walk. For the 50 runs, the mean acceptance rate is around 30% for the move cases and 10% for the birth and death cases. These values are consistent with acceptance values obtained by other applications of rj-McMC algorithm (Gallagher et al., 2009).

Our approach allows to display the temporal evolution of the b_{value} of the Gutenberg-Richter law along with the two parameters describing the detectability (Figure 5.D,E,F). At each time-bin over a 100 grid, the full probability distribution of b_{value} is the sum of all the marginal densities that have been accepted. For the three parameters, the probability distribution comprises the true value even for periods with low detectability (Figure 5.D.E.F). Specifically, for period S2 containing 261 events, the b_{value} estimated by the transdimensional approach between 1.05 and 1.0 is close to the true value 1.0 with a 68% confidence interval of ± 0.1. As this period involves the fewest events, the relative probability is the lowest. This confidence intervals narrows to ± 0.02 for the period S3 containing 2538 events. Despite a large increase of detectability between these two periods, the b_{value} remains stable. We demonstrate that the transdimensional framework retrieves the true values of the three parameters governing the frequency-magnitude distribution of earthquakes over time for a synthetic case. This approach gives larger uncertainties compared to those proposed by classical methods, primarily because it accounts, in addition to the number of earthquakes used, for existing correlations between parameters that classical approaches fail to resolve.

3 Application to a real earthquake catalog

A significant challenge is thus to select a region where the b_{value} and the completeness are homogeneous, and the Gutenberg-Richter law is valid for the chosen sub-dataset of earthquakes. The transdimensional approach presented here addresses this issue in time, yet the selection of a catalog with spatially homogeneous b_{value} remains crucial to better understand the physical meaning of b_{value} temporal variations.

Another challenge addressed by the transdimensional approach is the temporal variations in completeness magnitude, influenced by mainshock-aftershock sequences and seasonal variations due to anthropogenic or meteorological factors(e.g. Iwata, 2013). Here, we choose an earthquake catalog with expected variations in completeness to evaluate the approach's efficiency and compare results with other methods.

3.1 Data : Far-Western Nepal seismicity

Our first application of the transdimensional approach to investigate b_{value} variations focus on the temporal evolution of seismicity in a very seismically active region of Nepal. We use the earthquake data collected during two years by the Himalayan Karnali Network (HiKNet), the first dense seismological network of 15 temporary stations deployed in far-western Nepal. This earthquake catalog is derived from two studies (Hoste

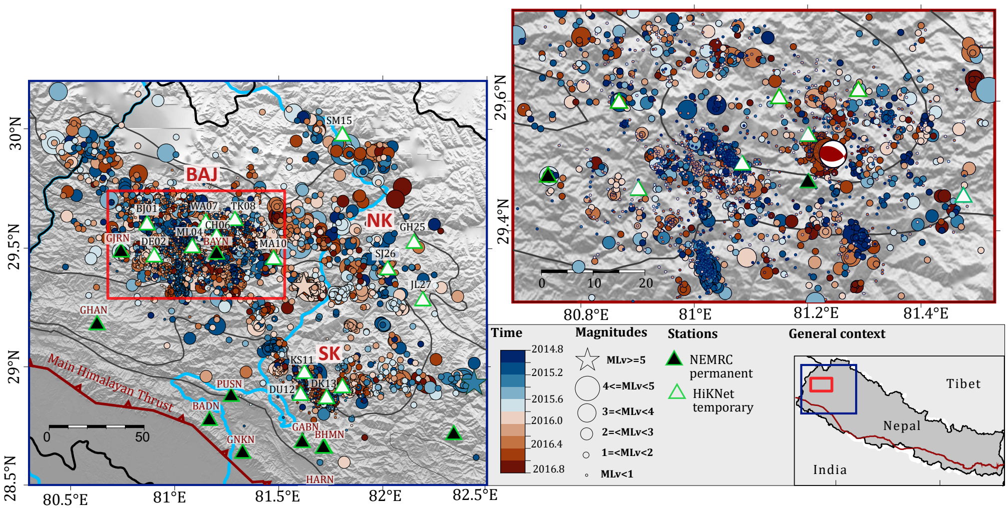

Figure 6. Far-western Nepal seismicity recorded during two years by the temporary seismological experiment of the Himalayan-Karnali network (HiKNet) (white triangles) and the permanent seismological network of the National Earthquake Monitoring Research Center (NEMRC) (black triangles). We use the earthquake catalog and focal mechanism from Laporte et al. (2021). NK: North Karnali sector, SK: South Karnali sector. In this study, we focus our analysis on the geographical subset represented by the inner red rectangle comprising 2593 events : the Bajhang region (BAJ).

et al., 2018; Laporte et al., 2021) which focuses on the spatio-temporal analysis for seismotectonic interpretation.

In Nepal, the main feature of seismicity is a belt of intense microseismicity which is located at depth on the locked portion of the Main Himalayan Thrust (MHT), (e.g. Ader et al., 2012). The MHT is the main active thrust fault which accommodates most of the shortening between the Indian plate and the Tibetan plateau. The seismicity is interpreted as resulting from stress build-up on the locked portion of the MHT. It exhibits a multimodal behavior, generating intermediate earthquakes ($M > 5$) that partially rupture the MHT, as well as large ($M > 7$) and great earthquakes ($M > 8$) that may rupture several lateral segments of the MHT, sometimes up to the surface (Dal Zilio et al., 2019).

In the area of interest in this paper, the most recent great earthquake occurred in 1505 A.D. according to historical records supported by paleoseismological evidence (Hossler et al., 2016; Riesner et al., 2021).

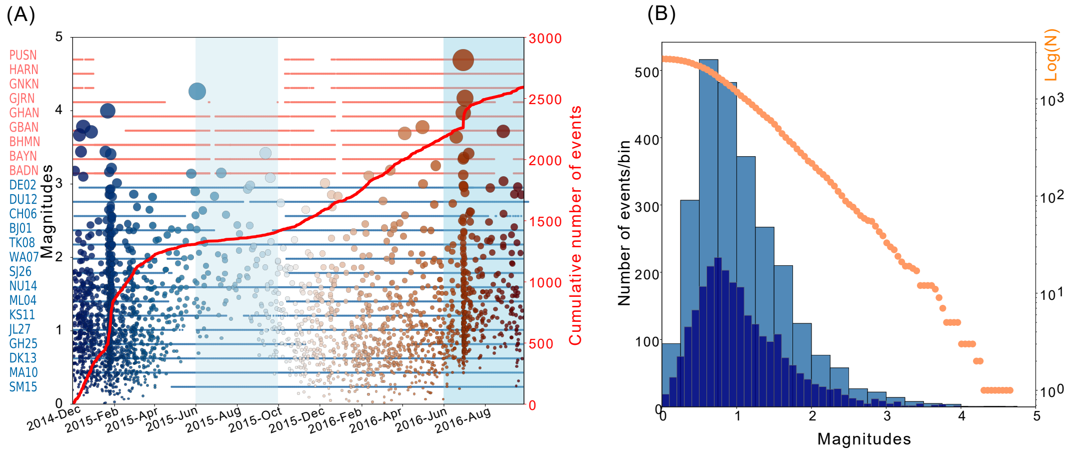

Figure 7. (A) Temporal distribution of magnitudes during two years of the HiKNet experiment in the Bajhang region. The color scale is the same as Figure 6. The red curve is the temporal evolution of the cumulative number of events. Horizontal lines correspond to station availability. Blue shades are the usual monsoon periods in Nepal. (B) Frequency-magnitude distribution in normal (blue histograms) and logarithmic (orange curve) scale. The two histograms are histograms for bins of 0.2 and 0.1 in magnitude, respectively.

Between December 2014 and September 2016, the temporary experiment recorded almost 4500 earthquakes in this region. The seismicity is structured into three seismic belts: one large belt in the westernmost part (BAJ) and two separate belts in the east (SK and NK) (Figure 6). Each belt contains several seismic clusters of different size and spatio-temporal behaviour. Most of them are located at mid-crustal depths (15−20km) close to mid-crustal structures such as the toe of mid-crustal ramps, which accumulate most of the interseismic strain on the fault. The geometry of the MHT fault is considered to be the primary factor influencing microseismicity. The focal mechanisms of the largest earthquakes are consistent with thrust faulting.

Facing challenges due to temporal variations of the detectability caused by station losses and monsoon periods, we chose this dataset to study temporal variations in b_{value} using the transdimensional approach. In Nepal, the monsoon season typically spans from early June to September, peaking in July and August. During this period, the high-frequency seismic noise increases, leading to fewer detected earthquakes. Additionally, in 2015, 9 out of 24 instruments were successively disconnected by storms. These typical monsoon periods are highlighted in blue in Figure 7.A, along with the availability of the 24 instruments throughout the experiment

To ensure spatial homogeneity of b_{value}, we focused on a geographical region of 94km by 57km, corresponding to the seismicity of the western belt, which is the most instrumented. 2593 earthquakes were recorded within this specific region. The cumulative number of earthquakes reveals distinct periods of seismic activity (see Figure 7.A). Using the frequentist approach of Aki for a completeness magnitude of 1.5 on the full time period, we get a b_{value} of 0.82±0.08 (Figure 7.B) consistent with the b_{value}obtained in far-western Nepal and more generally in Central and Eastern Nepal (e.g. Laporte et al., 2021). This low b_{value} is also consistent with the thrust-type faulting style (e.g. Schorlemmer et al., 2005).

3.2 Results

For this real application, we configure the rj-McMC algorithm to conduct 15,000 iterations for each individual run. However, as for the synthetic case, we initiate 50 parallel runs, totaling 750,000 models tested starting from distinct random seeds. For each of them, we burn 4,000 iterations and thin the chain by keeping only 1 out of 5 accepted models.

The algorithm converges towards 9 temporal discontinuities (defined as being over a probability of 15%) after 5000 iterations (Figure 8.A). The total acceptance rate is 23% and is above 20% for the three case scenarios births/deaths/moves.

Moreover, comparing the position of these discontinuities with the magnitude distribution of the seismicity, we can see that these nine discontinuities are coincident with some specific time periods of the dataset. We interpret and discuss each of them with respect to the stacked marginal densities of probability for b_{value}, μ and σ (Figure 8).

- S1 corresponds to the time of the M_{Lv} 4.0 earthquake of 22 January 2015, located at the base of the mid-crustal ramp of the MHT and followed by an increased seismic rate of about 300 events occurrence in 9 days (Hoste-Colomer et al., 2018; Laporte et al., 2021). For this seismic crisis, characterised as a large seismic swarm

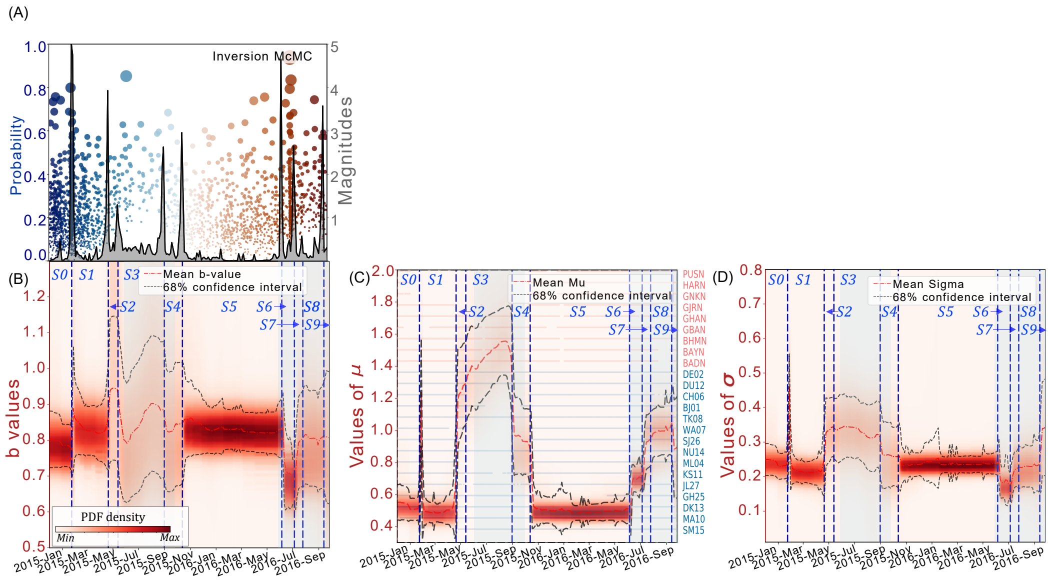

Figure 8. (A)Probability of temporal discontinuities at the completion of the rj-McMC algorithm and the temporal distribution of magnitudes in the background. The color-scale is the same as Figure 7. Here, a discontinuity is defined by the 15% probability threshold. (B) Marginal density distribution of b_{value} over time. (C) Marginal density distribution of μ over time. Horizontal lines are the stations availability in red for the permanent network and blue for the temporary network. (D) Marginal density distribution of σ over time For B), C) and D), the thin red dashed line is the mean probability for the respective marginal density distribution. The thin black lines are the $\pm 1\sigma$ uncertainty. The blue vertical dashed lines are the position to the solution of temporal discontinuities according to A).

by Hoste-Colomer et al. (2018), the b_{value} and the two detectability parameters increase suddenly and then return to their previous value.

- S2 corresponds to the loss of three stations (ML04, SJ26 and GJRN)(Figure 8.C), two of which were sited in the considered region (Figure 6). The confidence interval for every parameter becomes wider which is in accordance with fewer events detected. Only μ and σ present a clear discontinuous change. Specifically, μ increases from 0.5 ± 0.1 to 1.3 ± 0.3, and σ from 0.25 ± 0.05 to 0.32 ± 0.1. The b_{value} remains within the uncertainties of S1 while its uncertainty becomes larger.

- S3 corresponds to the loss of two stations in the center of the considered region (WA07 and CH06) (Figure 8). In the region, only 5 out of 9 stations are available during the monsoon period. Between S3 and S4, the monsoon periods starts and the μ parameter increases along with degradation in detectability, the σ parameter is not affected as much as μ. The mean b_{value} stays within the confidence interval of previous periods with a larger uncertainty. There is no statistical evidence of b_{value} variations during the monsoon period.
- S4 corresponds to a brutal improvement in the detectability. μ decreases from 1.6 ± 0.2 to 1.0 ± 0.2. This time likely corresponds to the early end of the monsoon period at the beginning of September 2015.
- S5 Both detectability parameters come back to the pre-monsoon values with the return of all 6 lost stations in October 2015. The b_{value} remains constant around 0.85 ± 0.1 during that time with a narrower confidence interval.
- S6 The density probability of b_{value} has a significant decrease at the beginning of June 2016 from 0.85 ± 0.1 to 0.7 ± 0.1. This time also corresponds to the expected beginning of the monsoon period, μ increases significantly and σ decreases.
- S7 corresponds to the onset of the second largest seismic crisis recorded in that region. On the 29th of June a $M_{Lv}4.8$ earthquake occurred and was followed by several aftershocks including two $M_{Lv} > 4$ earthquakes. It seems that after this crisis that lasted 12 days, the b_{value} comes back to its value of 0.85 ± 0.15. However this variation is not statistically significant because the confidence interval also increases due to the decrease in detectability. In fact, the μ parameter describing the mean of the detectability function keeps increasing in steps. One station (KS11) at 50 km from the area of interest is also interrupted.
- S8 Another station from the area of interest becomes unavailable (CH06) and the confidence intervals become wider. b_{value} and σ are constant during that period and approximately equal to the mean value they had during the two years 0.8 and 0.25 for b_{value} and σ respectively. The mean detectability μ seems to increase slightly but this increase is within the confidence interval. However this increase can be attested by the simultaneous widening of confidence intervals that might be due to fewer events detected during the monsoon.
- S9 in September 2015 corresponds to the end of the experiment, uncertainties are becoming wider with the disconnection of the firsts temporary stations.

Looking back at these results, the analysis of temporal variations of b_{value} (Figure 8) shows that there is no statistical evidence of temporal variations during the two years of the temporary experiment at the exception of a short period preceding one seismic crisis (S6). The b-Bayesian approach recovers as well the large variations of background detectability which can be explained by the loss of stations during summer 2015 and by the higher seismic noise during the monsoon periods. In particular, the μ parameter from the detectability function is the most sensitive to detectability changes and can be used as a proxy for deciphering detectability variations. This additional information, which is not given by traditional approaches, can be very useful for the characterization of the efficiency of a seismic network over time. Moreover, these large variations of detectability are taken into account in the Bayesian estimation of the b_{value} and act in the spread of its uncertainties along time. When all stations of the temporary network were available, the 1σ uncertainty of the b_{value} is reduced to ± 0.05. The widening of the posterior density function during the monsoon periods shows that during these periods the information contained in this earthquake catalog is not good enough to decipher some seasonal variations of the b_{value} even though it might exists. Outside the scope of this study, as the temporary network recorded data between 2014 and 2016, this particular sector of far-western Nepal has experienced 5 moderate ($M_L > 5$) damaging earthquakes since 2022, while the temporary experiment recorded none in two years.

4 Discussion and Perspectives

4.1 Comparison between classical approaches

Most of the time, temporal variations in b_{value} have been studied by applying the maximum likelihood approach (Aki, 1965; Utsu, 1966; Y. Shi & Bolt, 1982) over sliding time windows (e.g. Nuannin et al., 2005; Cao & Gao, 2002; Nanjo et al., 2012; Gulia & Wiemer, 2019). These techniques are dependent on the accuracy of the estimation of the time variations of the completeness magnitude (e.g. Helmstetter et al., 2006; Woessner & Wiemer, 2005) (Figure 9.A). More recently, the b-positive approach introduced by Van der Elst (2021) has been shown to be insensitive to the short-term incompleteness coming from mainshock-aftershocks sequences. Both approaches are very efficient and do not require any prior information on the b_{value} but they use a small part of the dataset only. The uncertainty in b_{value} is often estimated using formulas from Aki (1965) or Utsu (1966) within the sliding window, or by employing the bootstrap method(e.g.

Woessner & Wiemer, 2005). Importantly, this uncertainty is consistently estimated independently of the uncertainty associated with the completeness magnitude nor their correlation.

Comparing the outcomes derived from traditional approaches with those obtained using our b-Bayesian method on the earthquake catalog of far-western Nepal reveals intrinsic differences among these methodologies (Figure 9). Notably, both classical approaches exhibit sensitivity to the selection of the magnitude threshold (M_c or dM_c), whereas the b-Bayesian method effectively captures the uncertainty arising from fluctuations in detectability. Our findings indicate that most b_{value} variations shown by the classical moving-window techniques fall within the uncertainties accounted for by the b-Bayesian approach. These small-scale variations can sometimes lead to over-interpretations of the temporal variations in b_{value}.

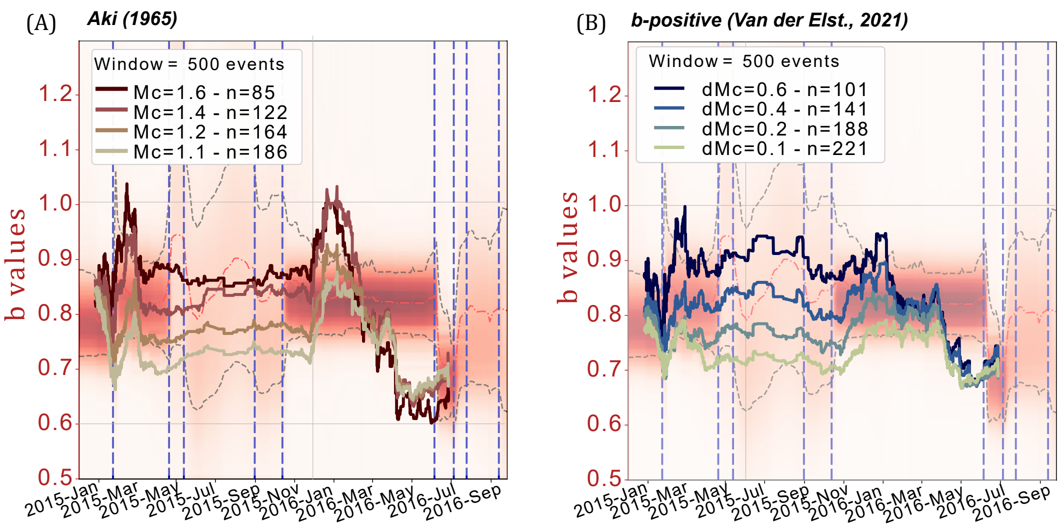

Figure 9. (A) Comparison of the temporal variations of b_{value} obtained using the b-Bayesian approach for far-western Nepal in red-shaded and the frequentist approach from Aki applied on a moving-window of 500 events as the 4 bold lines, depending on 4 values of completeness magnitude M_c. (B) Same as (A) but compared to the b-positive approach with 4 different values for dM_c. For both sub-figures, the legend gives the mean number of magnitudes kept for the b_{value} estimates in the windows, depending on the magnitude cut-off.

Despite using a moving-window of 500 events, this number is significantly reduced by truncation at M_c (up to only 17% of magnitudes retained for Aki's approach with $M_c =$ 1.6) or by using only positive magnitude differences for the b-positive method ($> 50\%$)

(Figure 9), while the b-Bayesian approach does not require any truncation and uses 100% of all the available data. Moreover, b-Bayesian proves to be the only tool for deciphering jointly the variations in detectability and can be used as a preliminary step before applying classical approaches to ensure that detectability is adequately considered.

In Table 2, we propose a comparison between the two classical approaches presented in this paper and the novel b-Bayesian approach. In conclusion, b-Bayesian proposes to use all the data available to invert for the temporal variations of three parameters related to the frequency-magnitude distribution. It uses Bayesian inference to capture the full density distributions and does not require any parametrization. This can be done at the cost of the computational time which is currently being reduced.

	Aki (1965)	b-Positive (van der Elst., 2021)	b-Bayesian (this study)
Inversion param.	b_{value}	b_{value}	b_{value}, μ, σ
Approach type	MLE	MLE	Bayesian
Uncertainties estimates	MLE	Bootstrap	Full PDF
Truncation	Mc $^{(*)}$	dMc	None
Used Data	$< 40\%$	$< 50\%$	100%
Temporal	Moving-window	Moving-window	Probabilistic
Comp. Time	Immediate $^{(**)}$	Immediate $^{(**)}$	Long $^{(***)}$

Table 2. Table of comparison between the two classical approaches from Aki (1965), van der Elst (2021) and b-Bayesian. $^{(*)}$ arbitrary $^{(**)}$ for one parametrization of Mc/dMc and choice of moving-window $^{(***)}$ no arbitrary parametrization

4.2 Perspectives

The frequency-magnitude distribution of earthquakes varies temporally and spatially. At the regional scale, b_{value} is thought to reflect the faulting style and the evolution of the state of stress on the faults (e.g. Schorlemmer et al., 2005; Gulia & Wiemer, 2019). This is also supported by experimental studies on the micro-failure (e.g. Scholz, 1968, 2015). Consequently, numerous studies have focused on monitoring the b_{value} at regional scales to discriminate foreshock and mainshocks sequences (e.g. Gulia et al., 2020;

Van der Elst, 2021). At the local scale, the b_{value} has also been proven valuable for describing the spatio-temporal behavior of seismic clusters (e.g. Farrell et al., 2009; Gui et al., 2020; Herrmann et al., 2022) or characterizing swarm-like sequences in relation to fluid-pressure (e.g. Hainzl & Fischer, 2002; Shelly et al., 2016; De Barros et al., 2019). We are now looking forward to applying b-Bayesian in these different contexts to discuss our results in comparison with previous studies and infer valuable information for characterizing seismic sequences.

The b-Bayesian method addresses temporal variations in the b_{value} using Bayesian inference. However, these variations are generally considered to be secondary compared to spatial variations (Wiemer & Wyss, 1997; Öncel & Wyss, 2000). Currently, both the b-Bayesian and classical approaches require a spatial subset of earthquakes with a homogeneous b_{value}, making it challenging to determine the adequacy of the dataset. Similar to the inversion of seismic velocities in tomography (e.g. Bodin & Sambridge, 2009), adapting the transdimensional approach to account for 2D spatial partitions could enable the capture of both temporal and spatial variations in the frequency-magnitude distribution. This is a future development of the method.

5 Appendix

5.1 Method : the Markov-chain Monte-Carlo

The Markov-chain is initialised by a randomized choice of a temporal model $T^{(c)} = [T_1^{(c)}, T_2^{(c)}, \dots T_k^{(c)}]$ with a number of discontinuities $k^{(c)}$ drawn between two values k_{min} and k_{max}. In practice, temporal discontinuities are considered as floating values in the range between $min(t_{obs})$ and $max(t_{obs})$. After this initialisation, at each iteration of the reversible-jump algorithm a new temporal model is proposed by making a random choice among three possibilities :

- (i) a death of a random temporal discontinuity $T_j^{(p)}$. Then, the proposed dimension is $k^{(p)} = k^{(c)} - 1$.
- (ii) a birth of a random temporal discontinuity $T_j^{(p)}$. Then, the proposed dimension is $k^{(p)} = k^{(c)} + 1$.
- (iii) a move of a random temporal discontinuity $T_j^{(p)}$ around its previous location. Then the proposed dimension remains $k^{(p)} = k^{(c)}$.

Each proposal (birth/death/move) has the same uniform probability of being drawn. For the moves, we randomly draw a time offset δ_t from a normal distribution around the previous time of a random time discontinuity of the current model. The standard deviation σ_{δ_t} of this normal distribution controls the efficiency of the exploration. A large standard deviation will produce large jumps with many rejected models and poor precision while a small standard deviation will produce very similar models and accept most of them. We follow the approach of (Gallagher et al., 2009) which tunes the value of σ_{δ_t} in order to get close a 20% acceptance rate. Every five hundred iterations, we monitor the acceptance rate and increase or decrease the σ_{δ_t} linearly with the deviation of the acceptance rate from the 20%.

The acceptance criterion is computed according to the Metropolis-Hastings rule in order to guide the chain towards the target distribution. Its general form for trans-dimensional functions is written as follows :

$$\alpha = min(1, \textit{prior ratio} \times \textit{likelihood ratio} \times \textit{proposal ratio} \times |J|) \tag{27}$$

such as :

$$\alpha = min(1, \frac{p(T^{(p)})}{p(T^{(c)})} \times \frac{p(d|T^{(p)})}{p(d|T^{(c)})} \times \frac{q(T^{(c)}|T^{(p)})}{q(T^{(p)}|T^{(c)})} \times |J|) \tag{28}$$

where $|J|$ is the Jacobian of the transformation from the current model to the proposal and can be shown as equal to 1 (e.g. Gallagher et al., 2009; Bodin & Sambridge, 2009). For moves of discontinuities, the dimension remains fixed, proposals are symmetric, and the prior ratio is one. The acceptance criterium is simply given by the usual Metropolis criterion :

$$\alpha = min(1, \frac{p(d|T^{(p)})}{p(d|T^{(q)})}) \tag{29}$$

For birth proposals, we randomly draw the location of a new discontinuity from the prior distribution. In this way, the prior and proposal ratios cancel out for birth and death steps, and thus the acceptance criterion (**??**) also conveniently becomes the usual Metropolis criterion.

6 Open Research

Codes of the b-Bayesian will be made available on github upon publication. The earthquake catalog of far-western Nepal used in Section 3 is available in Supplementary Material of Laporte et al. (2021).

Acknowledgments

The authors are grateful to the LabeX Lyon Institute of Origins (ANR-10-LABX-0066) Lyon for its financial support within the Plan France 2030 of the French government operated by the National Research Agency (ANR).

References

Ader, T., Avouac, J.-P., Liu-Zeng, J., Lyon-Caen, H., Bollinger, L., Galetzka, J., ... others (2012). Convergence rate across the nepal himalaya and interseismic coupling on the main himalayan thrust: Implications for seismic hazard. *Journal of Geophysical Research: Solid Earth*, *117*(B4).

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.*, *43*, 237–239.

Avlonitis, M., & Papadopoulos, G. (2014). Foreshocks and b value: bridging macroscopic observations to source mechanical considerations. *Pure and Applied Geophysics*, *171*, 2569–2580.

Beauval, C., & Scotti, O. (2003). Mapping b-values in france using two different magnitude ranges: Possible non power-law behavior. *Geophysical research letters*, *30*(17).

Beauval, C., & Scotti, O. (2004). Quantifying sensitivities of psha for france to earthquake catalog uncertainties, truncation of ground-motion variability, and magnitude limits. *Bulletin of the Seismological Society of America*, *94*(5), 1579–1594.

Bodin, T., & Sambridge, M. (2009). Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, *178*(3), 1411–1436.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., & Rawlinson, N. (2012). Transdimensional inversion of receiver functions and surface wave dispersion. *Journal of geophysical research: solid earth*, *117*(B2).

Bolton, D. C., Shreedharan, S., Rivière, J., & Marone, C. (2020). Acoustic energy release during the laboratory seismic cycle: Insights on laboratory earthquake precursors and prediction. *Journal of Geophysical Research: Solid Earth*, *125*(8), e2019JB018975.

Cao, A., & Gao, S. S. (2002). Temporal variation of seismic b-values beneath northeastern japan island arc. *Geophysical research letters*, *29*(9), 48–1.

Carter, J. A., & Berg, E. (1981). Relative stress variations as determined by b-values from earthquakes in circum-pacific subduction zones. *Tectonophysics*, *76*(3-4), 257–271.

Chan, C.-H., Wu, Y.-M., Tseng, T.-L., Lin, T.-L., & Chen, C.-C. (2012). Spatial and temporal evolution of b-values before large earthquakes in taiwan. *Tectonophysics*, *532*, 215–222.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the seismological society of America*, *58*(5), 1583–1606.

Daniel, G., Marsan, D., & Bouchon, M. (2008). Earthquake triggering in southern iceland following the june 2000 ms 6.6 doublet. *Journal of Geophysical Research: Solid Earth*, *113*(B5).

De Barros, L., Baques, M., Godano, M., Helmstetter, A., Deschamps, A., Larroque, C., & Courboulex, F. (2019). Fluid-induced swarms and coseismic stress transfer: A dual process highlighted in the aftershock sequence of the 7 april 2014 earthquake (ml 4.8, ubaye, france). *Journal of Geophysical Research: Solid Earth*, *124*(4), 3918–3932.

Drouet, S., Ameri, G., Le Dortz, K., Secanell, R., & Senfaute, G. (2020). A probabilistic seismic hazard map for the metropolitan france. *Bulletin of Earthquake Engineering*, *18*(5), 1865–1898.

Dutfoy, A. (2020). Estimation of the gutenberg–richter earthquake recurrence parameters for unequal observation periods and imprecise magnitudes. *Pure and Applied Geophysics*, *177*(10), 4597–4606.

Farrell, J., Husen, S., & Smith, R. B. (2009). Earthquake swarm and b-value characterization of the yellowstone volcano-tectonic system. *Journal of Volcanology and Geothermal Research*, *188*(1-3), 260–276.

Gallagher, K., Bodin, T., Sambridge, M., Weiss, D., Kylander, M., & Large, D. (2011). Inference of abrupt changes in noisy geochemical records using transdimensional changepoint models. *Earth and Planetary Science Letters*, *311*(1-2),

182–194.

Gallagher, K., Charvin, K., Nielsen, S., Sambridge, M., & Stephenson, J. (2009). Markov chain monte carlo (mcmc) sampling methods to determine optimal models, model resolution and model choice for earth science problems. *Marine and Petroleum Geology*, *26*(4), 525–535.

Geffers, G.-M., Main, I. G., & Naylor, M. (2022, February). Biases in estimating *b* -values from small earthquake catalogues: how high are high *b* -values? *Geophysical Journal International*, *229*(3), 1840–1855. Retrieved 2024-04-04, from https://academic.oup.com/gji/article/229/3/1840/6515950 doi: 10.1093/gji/ggac028

Geyer, C. J., & Møller, J. (1994). Simulation procedures and likelihood inference for spatial point processes. *Scandinavian journal of statistics*, 359–373.

Godano, C., Petrillo, G., & Lippiello, E. (2023, December). Evaluating the incompleteness magnitude using an unbiased estimate of the *b* value. *Geophysical Journal International*, *236*(2), 994–1001. Retrieved 2024-04-04, from https://academic.oup.com/gji/article/236/2/994/7464075 doi: 10.1093/gji/ggad466

Godano, C., Tramelli, A., Petrillo, G., & Convertito, V. (2024). Testing the predictive power of b value for italian seismicity. *Seismica*, *3*(1).

Goebel, T. H., Kwiatek, G., Becker, T. W., Brodsky, E. E., & Dresen, G. (2017). What allows seismic events to grow big?: Insights from b-value and fault roughness analysis in laboratory stick-slip experiments. *Geology*, *45*(9), 815–818.

Green, P. J. (1995). Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, *82*(4), 711–732.

Gui, Z., Bai, Y., Wang, Z., Dong, D., Wu, S., & Li, T. (2020). Spatiotemporal seismotectonic implications for the izu–bonin–mariana subduction zone from b-values. *Seismological Research Letters*, *91*(3), 1679–1693.

Gulia, L., & Wiemer, S. (2019). Real-time discrimination of earthquake foreshocks and aftershocks. *Nature*, *574*(7777), 193–199.

Gulia, L., Wiemer, S., & Vannucci, G. (2020). Pseudoprospective evaluation of the foreshock traffic-light system in ridgecrest and implications for aftershock hazard assessment. *Seismological Research Letters*, *91*(5), 2828–2842.

Hainzl, S., & Fischer, T. (2002). Indications for a successively triggered rupture growth underlying the 2000 earthquake swarm in vogtland/nw bohemia. *Journal of Geophysical Research: Solid Earth*, *107*(B12), ESE–5.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications.

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

Herrmann, M., Piegari, E., & Marzocchi, W. (2022). Revealing the spatiotemporal complexity of the magnitude distribution and b-value during an earthquake sequence. *Nature Communications*, *13*(1), 5087.

Hoste-Colomer, R., Bollinger, L., Lyon-Caen, H., Adhikari, L., Baillard, C., Benoit, A., … others (2018). Lateral variations of the midcrustal seismicity in western nepal: Seismotectonic implications. *Earth and Planetary Science Letters*, *504*, 115–125.

Ito, R., & Kaneko, Y. (2023, December). Physical Mechanism for a Temporal Decrease of the Gutenberg-Richter b -Value Prior to a Large Earthquake. *Journal of Geophysical Research: Solid Earth*, *128*(12), e2023JB027413. Retrieved 2024-04-04, from `https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2023JB027413` doi: 10.1029/2023JB027413

Iwata, T. (2013). Estimation of completeness magnitude considering daily variation in earthquake detection capability. *Geophysical Journal International*, *194*(3), 1909–1919.

Keller, M., Pasanisi, A., Marcilhac, M., Yalamas, T., Secanell, R., & Senfaute, G. (2014). A bayesian methodology applied to the estimation of earthquake recurrence parameters for seismic hazard assessment. *Quality and Reliability Engineering International*, *30*(7), 921–933.

Kwiatek, G., Martínez-Garzón, P., Becker, D., Dresen, G., Cotton, F., Beroza, G. C., … Bohnhoff, M. (2023). Months-long seismicity transients preceding the 2023 mw 7.8 kahramanmaraş earthquake, türkiye. *Nature Communications*, *14*(1), 7534.

Laporte, M., Bollinger, L., Lyon-Caen, H., Hoste-Colomer, R., Duverger, C., Letort, J., … others (2021). Seismicity in far western nepal reveals flats and ramps along the main himalayan thrust. *Geophysical Journal International*, *226*(3), 1747–1763.

Li, Y., & Chen, X. (2021). Variations in apparent stress and b value preceding the 2010 m w 8.8 bio-bío, chile earthquake. *Pure and Applied Geophysics*, 1–17.

Lippiello, E., & Petrillo, G. (2024). b-more-incomplete and b-more-positive: Insights on a robust estimator of magnitude distribution. *Journal of Geophysical Research: Solid Earth*, *129*(2), e2023JB027849.

Lombardi, A. M. (2021, March). A Normalized Distance Test for Co-Determining the Completeness Magnitude and *b* -Value of Earthquake Catalogs. *Journal of Geophysical Research: Solid Earth*, *126*(3), e2020JB021242. Retrieved 2024-04-04, from `https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020JB021242` doi: 10.1029/2020JB021242

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, *21*(6), 1087–1092.

Mignan, A., & Woessner, J. (2012). Theme iv—understanding seismicity catalogs and their problems. *Community online resource for statistical seismicity analysis*.

Mogi, K. (1962). Magnitude-frequency relation for elastic shocks accompanying fractures of various materials and some related problems in earthquakes. *Bull. Earthq. Res. Inst., Univ. Tokyo*, *40*, 831–853.

Morales-Yáñez, C., Bustamante, L., Benavente, R., Sippl, C., & Moreno, M. (2022). B-value variations in the central chile seismic gap assessed by a bayesian transdimensional approach. *Scientific Reports*, *12*(1), 21710.

Mori, J., & Abercrombie, R. E. (1997). Depth dependence of earthquake frequency-magnitude distributions in california: Implications for rupture initiation. *Journal of Geophysical Research: Solid Earth*, *102*(B7), 15081–15090.

Nanjo, K., Hirata, N., Obara, K., & Kasahara, K. (2012). Decade-scale decrease inb value prior to the m9-class 2011 tohoku and 2004 sumatra quakes. *Geophysical Research Letters*, *39*(20).

Nuannin, P., Kulhanek, O., & Persson, L. (2005). Spatial and temporal b value anomalies preceding the devastating off coast of nw sumatra earthquake of december 26, 2004. *Geophysical research letters*, *32*(11).

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

Ogata, Y., & Katsura, K. (1993). Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues. *Geophysical Journal International*, *113*(3), 727–738.

Ogata, Y., & Katsura, K. (2006). Immediate and updated forecasting of aftershock hazard. *Geophysical research letters*, *33*(10).

Omori, F. (1894). On after-shocks. *Seismological journal of Japan*, *19*, 71–80.

Öncel, A. O., & Wyss, M. (2000). The major asperities of the 1999 m w= 7.4 izmit earthquake defined by the microseismicity of the two decades before it. *Geophysical Journal International*, *143*(3), 501–506.

Petruccelli, A., Schorlemmer, D., Tormann, T., Rinaldi, A. P., Wiemer, S., Gasperini, P., & Vannucci, G. (2019). The influence of faulting style on the size-distribution of global earthquakes. *Earth and Planetary Science Letters*, *527*, 115791.

Plourde, A. (2023, September). Assessing earthquake rates and b-value given spatiotemporal variation in catalog completeness: Application to Atlantic Canada. *Seismica*, *2*(2). Retrieved 2024-04-04, from `https://seismica.library.mcgill.ca/article/view/384` doi: 10.26443/seismica.v2i2.384

Ringdal, F. (1975). On the estimation of seismic detection thresholds. *Bulletin of the Seismological Society of America*, *65*(6), 1631–1642.

Rivière, J., Lv, Z., Johnson, P., & Marone, C. (2018). Evolution of b-value during the seismic cycle: Insights from laboratory experiments on simulated faults. *Earth and Planetary Science Letters*, *482*, 407–413.

Sambridge, M., Bodin, T., Gallagher, K., & Tkalčić, H. (2013). Transdimensional inference in the geosciences. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*(1984), 20110547.

Sambridge, M., Gallagher, K., Jackson, A., & Rickwood, P. (2006). Trans-dimensional inverse problems, model comparison and the evidence. *Geophysical Journal International*, *167*(2), 528–542.

Scholz. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bulletin of the seismological society of America*, *58*(1),

399–415.

Scholz. (2015). On the stress dependence of the earthquake b value. *Geophysical Research Letters*, *42*(5), 1399–1402.

Schorlemmer, D., Wiemer, S., & Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes. *Nature*, *437*(7058), 539–542.

Shelly, D. R., Ellsworth, W. L., & Hill, D. P. (2016). Fluid-faulting evolution in high definition: Connecting fault structure and frequency-magnitude variations during the 2014 long valley caldera, california, earthquake swarm. *Journal of Geophysical Research: Solid Earth*, *121*(3), 1776–1795.

Shi, H., MENG, L., ZHANG, X., CHANG, Y., YANG, Z., XIE, W., . . . HAN, P. (2018). Decrease in b value prior to the wenchuan earthquake (m s 8.0). *Chinese Journal of Geophysics*, *61*(5), 1874–1882.

Shi, Y., & Bolt, B. A. (1982). The standard error of the magnitude-frequency b value. *Bulletin of the Seismological Society of America*, *72*(5), 1677–1687.

Spassiani, I., Taroni, M., Murru, M., & Falcone, G. (2023, March). Real time Gutenberg–Richter *b* -value estimation for an ongoing seismic sequence: an application to the 2022 marche offshore earthquake sequence (*M* L 5.7 central Italy). *Geophysical Journal International*, *234*(2), 1326–1331. Retrieved 2024-04-04, from `https://academic.oup.com/gji/article/234/2/1326/7086114` doi: 10.1093/gji/ggad134

Taroni, M., & Akinci, A. (2020, November). Good practices in PSHA: declustering, b-value estimation, foreshocks and aftershocks inclusion; a case study in Italy. *Geophysical Journal International*, *224*(2), 1174–1187. Retrieved 2024-04-04, from `https://academic.oup.com/gji/article/224/2/1174/5911581` doi: 10.1093/gji/ggaa462

Utsu, T. (1966). A statistical significance test of the difference in b-value between two earthquake groups. *Journal of Physics of the Earth*, *14*(2), 37–40.

Utsu, T. (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 (1). *Journal of the Faculty of Science, Hokkaido University. Series 7, Geophysics*, *3*(5), 379–441.

Van der Elst, N. J. (2021). B-positive: A robust estimator of aftershock magnitude distribution in transiently incomplete catalogs. *Journal of Geophysical*

Research: Solid Earth, *126*(2), e2020JB021027.

Weichert, D. H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes. *Bulletin of the Seismological Society of America*, *70*(4), 1337–1346.

Wetzler, N., Lay, T., & Brodsky, E. E. (2023). Global characteristics of observable foreshocks for large earthquakes. *Seismological Research Letters*, *94*(5), 2313–2325.

Wiemer, S., & Wyss, M. (1997). Mapping the frequency-magnitude distribution in asperities: An improved technique to calculate recurrence times? *Journal of Geophysical Research: Solid Earth*, *102*(B7), 15115–15128.

Woessner, J., & Wiemer, S. (2005). Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bulletin of the Seismological Society of America*, *95*(2), 684–698.

Yin, F., & Jiang, C. (2023, October). Enhanced b -value time series calculation method using data-driven approach. *Geophysical Journal International*, ggad419. Retrieved 2024-04-04, from `https://academic.oup.com/gji/advance-article/doi/10.1093/gji/ggad419/7328941` doi: 10.1093/gji/ggad419

b-Bayesian : The Full Probabilistic Estimate of b-value Temporal Variations for Non-Truncated Catalogs

M. Laporte[1], S. Durand[1], T. Bodin[1], B. Gardonio[1], and D. Marsan[2],

[1]LGL-TPE, Université Claude Bernard, Lyon, France
[2]Université de Savoie Mont Blanc, Isterre, Chambéry, France

Key Points:

- The b_{value} from the Gutenberg-Richter law is usually inferred by truncating earthquake catalogs above a completeness magnitude. We propose to use all the data available and invert conjointly for b_{value} and two parameters describing the detectability.
- Using a Bayesian framework we retrieve the full posterior density function of b_{value} with a more realistic evaluation of its uncertainties than classical approaches.
- b-Bayesian performs a transdimensional inversion to recover the temporal changes of the b_{value} and detectability.

Corresponding author: Marine Laporte, `marine.laporte@univ-lyon1.fr`

Abstract

The frequency/magnitude distribution of earthquakes can be approximated by an exponential law whose exponent (the so-called b_{value}) is routinely used for probabilistic seismic hazard assessment. The b-value is commonly measured using Aki's maximum likelihood estimation, although biases can arise from the choice of completeness magnitude (i.e. the magnitude below which the exponential law is no longer valid). In this work, we introduce the b-Bayesian method, where the full frequency-magnitude distribution of earthquakes is modelled by the product of an exponential law and a detection law. The detection law is characterized by two parameters, which we jointly estimate with the b_{value} within a Bayesian framework. All available data are used to recover the joint probability distribution. The b-Bayesian approach recovers temporal variations of the b_{value} and the detectability using a transdimensional Markov chain Monte Carlo (McMC) algorithm to explore numerous configurations of their time variations. An application to a seismic catalog of far-western Nepal shows that detectability decreases significantly during the monsoon period, while the b-value remains stable, albeit with larger uncertainties. This confirms that variations in the b_{value} can be estimated independently of variations in detectability (i.e. completeness). Our results are compared with those obtained using the maximum likelihood estimation, and using the b-positive approach, showing that our method avoids dependence on arbitrary choices such as window length or completeness thresholds.

1 Introduction

Classically, the probability density function of an earthquake of magnitude m above a magnitude of completeness M_c follows the Gutenberg-Richter law (Aki, 1965):

$$p(m) = \beta e^{-\beta(m-M_c)} \quad (1)$$

where $\beta = b_{value} \times \ln(10)$. The b_{value} is the seismic parameter that describes the relative number of large magnitude earthquakes versus smaller magnitude earthquakes. For global earthquake catalogues, the b_{value} is typically close to 1, but has been shown to vary in both space and time when focusing on earthquake catalogues for specific seismogenic regions or time periods (e.g. Wiemer & Wyss, 1997; Ogata & Katsura, 2006). With the rapid growth of seismological instruments and recording capabilities, there is a need for advanced statistical methods to analyze earthquake catalogs. Here, we analyse the distribution of earthquake magnitudes, focusing on the possibility to observe and interpret temporal variations in this distribution.

The Gutenberg-Richter law is widely used as an earthquake recurrence model for Probabilistic Seismic Hazard Assessment (PSHA) studies (e.g. Cornell, 1968; Drouet et al., 2020). Consequently, the accurate estimation of b_{value} and its uncertainties play a crucial role in the accuracy and robustness of seismic hazard estimates (e.g. Keller et al., 2014; Beauval & Scotti, 2004; Taroni & Akinci, 2020). For accurate hazard assessment, b_{value} biases due to the incompleteness of earthquake catalogues need to be addressed (e.g. Weichert, 1980; Plourde, 2023; Dutfoy, 2020; Beauval & Scotti, 2004) as well as possible temporal or spatial variations of the b_{value} (e.g. Beauval & Scotti, 2003; Yin & Jiang, 2023).

The physical interpretation of these spatio-temporal variations in the frequency-magnitude distribution of earthquakes has been a subject of ongoing debate for years (e.g. Mogi, 1962; Scholz, 1968; Carter & Berg, 1981; Herrmann et al., 2022). Based on observations from laboratory earthquake simulations, which are commonly used as analogues for studying natural earthquake behaviour, it has been proposed that the b_{value} is inversely related to the normal and shear stress applied to the fault (Scholz, 1968). At the scale of the seismic cycle, which is reproduced in stick-slip experiments with controlled stress and friction properties, the b_{value} has been observed to decrease linearly with stress build-up and to increase abruptly with the stress-drop release during earthquake rupture (Avlonitis & Papadopoulos, 2014; Goebel et al., 2017; Rivière et al., 2018; Bolton et al., 2020). Extending this observation to real earthquake systems is not straightforward because real earthquake catalogues contain additional uncertainties and the estimation of the actual state of the stress field is another inverse problem.

However, the b_{value} is also widely used to characterize real earthquake catalogs. It is commonly estimated to characterize earthquake clusters and discriminate between seismic swarms (e.g. De Barros et al., 2019). Some variations in b_{value} have been observed for different earthquakes depths or within different stress regimes (Mori & Abercrombie, 1997; Schorlemmer et al., 2005; Scholz, 2015; Petruccelli et al., 2019; Morales-Yáñez et al., 2022). Low b_{value} (< 0.8), associated to a larger number of larger magnitudes earthquakes compared to the normal regime, have been observed for several earthquake sequences occurring before a large earthquake rupture (e.g. Nanjo et al., 2012; Chan et al., 2012; H. Shi et al., 2018; Li & Chen, 2021; Van der Elst, 2021; Kwiatek et al., 2023; Wetzler et al., 2023). This observation has a major impact for the identification of precursory phases before large mainshocks. Recently, b_{value} monitoring has been proposed to serve as a stress-meter for discrimination of foreshock sequences (e.g. Gulia & Wiemer, 2019; Ito & Kaneko, 2023). This topic remains under debate due to large uncertainties that could arise either from earthquake catalogs or from b_{value} estimation approaches (e.g. Lombardi, 2021; Spassiani et al., 2023; Yin & Jiang, 2023; Geffers et al., 2022; Godano et al., 2024).

The most classical approach for estimating b_{value} from a catalog of earthquake magnitudes is the maximum likelihood estimation of Aki and its generalization (Aki, 1965; Utsu, 1966), which depends on the arbitrary choice of the completeness magnitude M_c :

$$\beta = \frac{1}{\overline{m} - M_c} \tag{2}$$

with $\overline{m}$ the mean of magnitudes greater than M_c. Using this formula, only events with magnitudes larger than M_c are used to estimate β.

Unnoticed changes in completeness over time are the main source of bias when studying b_{value} temporal variations (e.g. Woessner & Wiemer, 2005; Helmstetter et al., 2006; Mignan & Woessner, 2012; Lombardi, 2021; Plourde, 2023; Godano et al., 2023) Two main sources of incompleteness are usually identified (e.g. Lippiello & Petrillo, 2024) : (1) the background incompleteness coming from momentary changes in the detectability of the seismic network, and (2) the short-term aftershock incompleteness (STAI) which describes the short but large changes in completeness that occur during mainshock-aftershock sequences where large earthquakes mask smaller ones (e.g. Helmstetter et al., 2006; Hainzl & Fischer, 2002). The b-positive approach (Van der Elst, 2021) is a variant of Aki's maximum likelihood approach, using differences in the magnitudes of successive earthquakes to propose a moving-window estimate of temporal changes in the b_{value} during mainshock-aftershock sequences without being biased by STAI.

$$\beta = \frac{1}{\overline{m'}^{+} - dM_c} \tag{3}$$

with $\overline{m'}^{+}$ the mean of positive magnitudes differences greater than dM_c, which is a chosen value that should be greater than twice the minimum magnitude difference. Based on the fact that two consecutive events in a mainshock-aftershock sequence share the same completeness, this approach is now frequently used for a more accurate estimation of temporal variations of b_{value}.

Even though the b-positive approach provides a major advantage in comparison to Aki's classical approach, it still suffers from its dependence on the choice of dM_c and to the size of the moving-window (e.g. Lippiello & Petrillo, 2024). Furthermore, the b_{value} estimate is computed on less than half of the available data and uncertainties are usually assessed using a bootstrap approach (Van der Elst, 2021).

In this paper, we introduce the b-Bayesian approach to explore the temporal variation of b_{value}, while addressing the problems of classical frequentist approaches. We propose to invert for b_{value} using the entire catalogue, taking into account a detectability function. By adopting this approach, our results are independent of the arbitrary choice of a completeness magnitude. Instead of traditional methods that compare frequency-magnitude distributions over random data subsets or that recover pseudo-continuous temporal variations using moving time windows, we address temporal variations in b_{value} and detectability by considering the number and positions of temporal discontinuities where b_{value} or detectability changes. The inversion of temporal discontinuities is achieved using a transdimensional framework.

Transdimensional inversion is commonly used in seismic tomography to allow the data to determine the level of spatial complexity in the recovered tomographic model (e.g. Bodin & Sambridge, 2009; Bodin et al., 2012). It has recently been adapted to estimate variations in the b_{value} from truncated catalogs along one dimension, such as time or depth (Morales-Yáñez et al., 2022). Here, we use transdimensional inversion to recover one-dimensional partitions of the entire earthquake dataset. A Bayesian framework provides a global formulation of the inverse problem and allows for the probabilistic estimation of temporal changes of b_{value}, and detectability. The complexity of the model does not depend on any arbitrary parameter, but is determined by the complexity of the data.

This paper is organized as follows. First, we describe the novel b-Bayesian approach for a time-invariant case, including the assessment of detectability using all available magnitude data. We describe how we extend this approach using a transdimensional framework in order to invert for temporal variations based on the complexity of the data. We present the results obtained using a synthetic catalog generated to mimic real-world scenarios. A first application of the b-Bayesian method is presented using a real earthquake catalog. We compare the results of b-Bayesian with the two frequentist approaches : the maximum likelihood estimate and the b-positive to describe the temporal variation of b_{value} and the temporal variations of detectability for an earthquake catalog of far-western Nepal spanning two years of microseismicity.

2 Method : A Bayesian framework

In this study, a dataset is an earthquake catalogue which corresponds to a set of N observations of (non-discrete) magnitudes m_i $(i = 1 \dots N)$ that we note :

$$\mathbf{d} = [m_1, m_2, \dots m_N] \tag{4}$$

From hereafter, we refer to conditional probabilities using $p(a|b)$. We know from the Gutenberg-Richter law (eq.1) that the probability density of observing one earthquake i of magnitude $m_i \geq M_c$ for a given β is :

$$p(m_i|\beta) = \beta e^{-\beta(m_i - M_c)} \tag{5}$$

and is zero if $m_i < M_c$. Then, assuming that the magnitudes of the seismic events are independent, we can write the probability of observing the entire earthquake dataset $\mathbf{d}$ with $m_i \geq M_c$, $p(\mathbf{d}|\beta)$, as :

$$p(\mathbf{d}|\beta) = \prod_{i=1}^{N} p(m_i|\beta) = \beta^N e^{-\beta N(\overline{m_i} - M_c)} \tag{6}$$

where $\overline{m_i}$, is the mean magnitude of events with $m_i \geq M_c$ and M_c, the magnitude of completeness. Note here that the value of β that maximizes (eq.6) is the maximum likelihood solution given by Aki's formula in (eq.2).

2.1 The temporally invariant case

In practise, seismic catalogs are truncated at the completeness magnitude M_c in order to avoid biases due to the detection capacity (e.g. Aki, 1965; Utsu, 1966). As a result, the classical approach for b-value estimation strongly depends on the choice made for M_c. Various methods have been developed for assessing the completeness magnitude (e.g. Ringdal, 1975; Ogata & Katsura, 1993; Woessner & Wiemer, 2005; Mignan & Woessner, 2012) or to correct the dataset for its temporal variations (e.g. Helmstetter et al., 2006; Cao & Gao, 2002; Chan et al., 2012). However, defining such a completeness magnitude always implies to ignore a significant portion of a dataset that may contain valuable information about the statistics of seismicity. Here instead, we propose to analyse the entire dataset by modelling the entire frequency-magnitude distribution of earthquakes. To do so, the Gutenberg-Richter law is modulated by a detection law q(m) such that now :

$$p(m_i|\beta) = \frac{1}{K} q(m_i) \beta e^{-\beta m_i} \tag{7}$$

where $q(m)$ defines the probability density of detecting an event, and K a constant to insure that the probability distribution integrates to one :

$$\int_{M_{min}}^{\infty} p(m_i|\beta) dm = 1 \tag{8}$$

with M_{min} the smallest earthquake magnitude in the catalog.

In this way, the probability of observing an event is the product of the probability of occurrence (given by the Gutenberg Richter law) and the probability of detection (given by the detection law $q(m)$ that varies from 0, no detection, to 1, 100% detection).The error function has been proposed in the literature to represent the probability of detection of an event in the presence of log-normal seismic noise (e.g. Ringdal, 1975; Ogata & Katsura, 1993; Daniel et al., 2008). The error function (see Figure 1.A) depends on two parameters μ and σ such as :

$$q(m) = \frac{1}{2} + \frac{1}{2} erf\left(\frac{m-\mu}{\sqrt{2}\sigma}\right) \tag{9}$$

where μ represents 50% of probability of detection for an earthquake of magnitude $m = \mu$, and becomes 84% for $m = \mu + \sigma$. The magnitude of completeness is the equivalent of the 84% threshold. This function fits adequately the frequency-magnitude distribution for a variety of cases (Ogata & Katsura, 1993; Woessner & Wiemer, 2005).

From equations (7) and (8), we can write :

$$K = \int_{M_{min}}^{\infty} q(m) \beta e^{-\beta(m - M_{min})} dm \tag{10}$$

Fortunately, this integral for the error function q(m) in (eq.9) has a closed form solution between M_{min} and infinity :

$$K = q(M_{min}) + \left(1 - q(M_{min} + \beta\sigma^2)\right) \exp\left(\frac{\beta^2\sigma^2}{2} - \beta(\mu - M_{min})\right) \tag{11}$$

Now assuming that magnitudes are independent, the probability of observing a full dataset $\mathbf{d}$ is :

$$p(\mathbf{d}|\omega) = \prod_{i=1}^{N} p(m_i|\omega) = \left(\frac{\beta}{K}\right)^N \prod_{i=1}^{N} q(m_i) \times e^{-\beta N(\overline{m_i} - M_{min})} \tag{12}$$

where $\omega = [\beta, \mu, \sigma]$ is our set of three unknown model parameters. Our goal here is to estimate these parameters from a set of realizations $\mathbf{d}$. This is an inverse problem that can be formulated in a Bayesian framework, where the posterior solution $p(\omega|\mathbf{d})$ is the product between the model priors and the likelihood function $p(\mathbf{d}|\omega)$ (eq.12).

$$p(\omega|\mathbf{d}) = p(\beta, \mu, \sigma)p(\mathbf{d}|\omega) \tag{13}$$

Here we set independent uniform prior distributions for the three parameters, partly because they are not related to the same physics: b_{value} is related to seismicity and μ and σ to network detectability. Although we can expect correlations between these parameters from the data (i.e. *a posteriori*), our level of knowledge is independent for each parameter. This independence greatly facilitates Bayesian inference. For each parameter, we use a simple uniform prior distribution, independently defined between a fixed range of realistic values. The choice of the bounds is guided by the literature and should be chosen depending on the seismotectonic context and the mean detectability of the network. We advise to impose a relatively wide range of values for the b_{value} inference, allowing both high (> 1) and low (< 1) b_{value} for an earthquake catalog. In the context of geothermal or volcanic activity, this range may be extended to allow larger values (up to 2.5) for the b_{value}. The choice of bounds for the μ parameter should be guided by the level of detectability of the seismic network and the "expected" variations in completeness. For a local network (seismicity included within 50 km), which essentially records microseismicity, we can set this range of values between 0.5 and 2. In the presence of at least one mainshock/aftershock sequence, this range should also be increased. The prior distribution on σ can be set between 0.01 and 0.5 and does not need to be adjusted depending on the context.

The posterior distribution (eq.13) can be numerically approximated using a classical Monte-Carlo approach.

As an example, we construct a synthetic dataset of 4460 independent magnitudes, randomly drawn from a Gutenberg-Richter law characterized by $b_{value} = \frac{\beta}{log(10)} = 0.9$

and modulated by an error detection function characterized by μ=0.75 and σ=0.34 (see Figure 1.A). We then estimate our set of parameters $\omega = [\beta, \mu, \sigma]$ from the catalogue, by approximating the posterior distribution $p(\omega|\mathbf{d})$ with a standard Monte Carlo scheme by randomly sampling the model priors. The resulting 3D posterior density function can be projected onto each parameter (Figure 1.B) to derive 1D and 2D marginal distributions (Figure 1.C). For example, the marginal distribution for β is simply obtained by integrating the posterior over μ, and σ :

$$p(\beta|\mathbf{d}) = \int\int p(\beta, \mu, \sigma|\mathbf{d})\, d\mu\, d\sigma \tag{14}$$

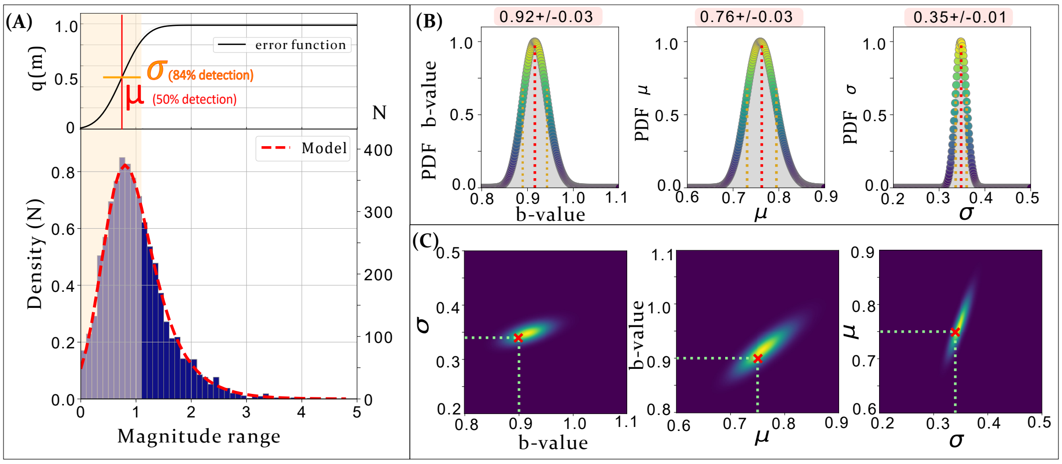

Figure 1. (A) [Top] Detection function q(m) associated with the synthetic dataset [Bottom] Frequency-magnitude distribution of a synthetic catalog with values of $b_{value} = 0.9$, $\mu = 0.75$ and $\sigma = 0.34$. The red dotted line corresponds to the model $\omega = [0.92, 0.76, 0.35]$ that best fits the observations. The yellow area corresponds to the distribution of magnitudes inferior to $(\mu + \sigma)$ that are usually removed by classical approaches (about 60% of available data). (B) Marginals distributions of the posterior function: $p(b_{value}|\mathbf{d})$, $p(\mu|\mathbf{d})$ and $p(\sigma|\mathbf{d})$ from left to right, respectively. Here, posterior functions are normalized by their maximum. (C) 2D marginal distributions of the posterior function. True values are represented by the red cross.

Note that the 2D marginals are useful to show the correlations between pairs of parameters. Uncertainty estimates of the three parameters can be obtained with the 1σ confidence interval.

Compared to optimization approaches where only the best fitting (i.e. maximum likelihood) ω is obtained, our method provides the 3D posterior density distribution, $p(\omega|\mathbf{d})$, from which parameter correlations and uncertainties can be estimated. Moreover, the dataset is no longer truncated above a completeness magnitude, instead, the full frequency-magnitude distribution is now used to jointly invert for b_{value} and detectability.

2.2 Temporal variations of b-value

2.2.1 A transdimensional parametrization

Going one step further, we now consider that ω can vary with time and our goal is to recover the location of temporal changes. Our three parameters in ω are considered constant in periods separated by abrupt changes (see Figure 3). To that aim, temporal variations are modeled with a set of discontinuities $\mathbf{T}$:

$$\mathbf{T} = [T_1, T_2, \dots T_k] \tag{15}$$

where, k is the number of temporal discontinuities and T_j $(j = 0 \dots k)$ the times at which the frequency-magnitude distribution changes. The unknown models vectors of the time varying frequency-magnitude distribution will be denoted :

$$\mathbf{\Omega} = [\omega_1, \omega_2, \dots \omega_{k+1}] \tag{16}$$

where $\omega_j = [\beta_j, \mu_j, \sigma_j]$ is the local model predicting the sub-dataset d_j between two discontinuities $[T_{j-1}, T_j]$. Note that $T_0 = min(t_{obs})$ and $T_{k+1} = max(t_{obs})$. The full posterior solution $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ describes the joint probability for local models $[\omega_1, \omega_2, \dots \omega_{k+1}]$ predicting events between each temporal discontinuities $[T_1, T_2, \dots T_k]$ of the temporal model, $\mathbf{T}$. Since the dimension of the model varies with the number of discontinuities, k, which is unknown, the inverse problem is so-called transdimensional. The posterior $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ does not have an analytical solution but can be sampled with a Monte Carlo algorithm. In this work, we propose to isolate the part of the posterior solution that is transdimensional (and to sample it with an appropriate algorithm), and to separate it from a part where the dimension is fixed.

That is, the full posterior solution $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ can be developed as a product of a conditional term $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T})$ and a marginal term $p(\mathbf{T}|\mathbf{d})$:

$$p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d}) = p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T}) \times P(\mathbf{T}|\mathbf{d}) \tag{17}$$

The following sections describe each of these terms in detail and how they can be approximated.

2.2.2 *The conditional posterior* $p(\boldsymbol{\Omega}|\boldsymbol{d}, \boldsymbol{T})$

The conditional term $p(\boldsymbol{\Omega}|\mathbf{d}, \mathbf{T})$ describes the probability distribution for parameters $\boldsymbol{\Omega}$ for a given time partition $\mathbf{T}$ (Figure. 2). It can be itself decomposed with the Bayes theorem into the product of a likelihood distribution and a prior distribution :

$$p(\boldsymbol{\Omega}|\mathbf{d}, \mathbf{T}) = p(\mathbf{d}|\boldsymbol{\Omega}, \mathbf{T}) \times p(\boldsymbol{\Omega}|\mathbf{T}) \tag{18}$$

Since all magnitudes are independent, the likelihood is the product of likelihoods for ev-

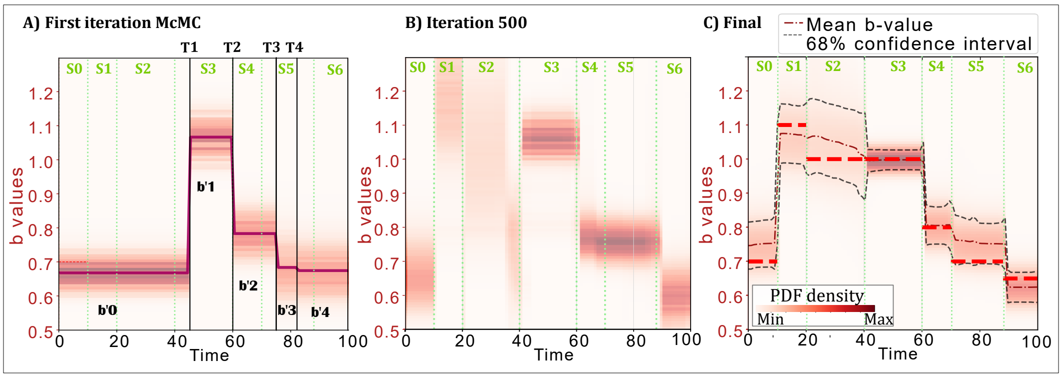

Figure 2. (A) Initial iteration of the Markov chain : plot of the conditional probability $p(\mathbf{d}|b_{value}, \mathbf{T})$, the probability of b_{value} for a fixed time model $\mathbf{T}$ of dimension $k = 4$. The 4 black vertical lines are the discontinuities of the proposed time model $\mathbf{T}$. The 3D posterior density function is computed for each data subset $T_j (j = 0 \ldots 4)$. The bold line is the mean posterior density function of β_j over time. Here, the synthetic earthquake dataset has been constructed to represent 6 discontinuities. The six green vertical dashed lines are the theoretical discontinuities. (B) Iteration 500 of the McMC: preliminary result of the b_{value} time variations, sum of the marginal density functions of the accepted models. The six discontinuities are almost all retrieved. (C) Final iteration of the Markov Chain : sum of marginal density functions of the totality of accepted models. The final temporal evolution of b_{value} fits the true b_{value} of the synthetic dataset which are represented in bold dashed horizontal lines.

ery sub-dataset $\mathbf{d}_j$ given by the temporal model $\mathbf{T}$:

$$p(\mathbf{d}|\boldsymbol{\Omega}, \mathbf{T}) = \prod_{j=1}^{k+1} p(\mathbf{d}_j|\boldsymbol{\Omega}, \mathbf{T}) \tag{19}$$

And since the magnitudes of events $\mathbf{d}_j$ occurring between two discontinuities $[T_{j-1}, T_j]$ only depend on the local parameters $\omega_j = [\beta_j, \mu_j, \sigma_j]$, we can write the likelihood :

$$p(\mathbf{d}|\mathbf{\Omega}, \mathbf{T}) = \prod_{j=1}^{k+1} p(\mathbf{d}_j|\omega_j, \mathbf{T}). \tag{20}$$

where $p(\mathbf{d}_j|\omega_j, \mathbf{T})$ is the likelihood of the data within a time period j which is simply given by equation (eq.12).

The prior distribution for $\mathbf{\Omega}$ given a fixed temporal model $\mathbf{T}$, $p(\mathbf{\Omega}|\mathbf{T})$ from equation (eq.18), is chosen to be the same within each partition $p(\omega_j|\mathbf{T})$, and simply corresponds to the uniform prior distribution used in the temporally invariant case (eq. 13). Thus, the conditional posterior $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T})$ is easy to sample as different periods j can be independently sampled with the same algorithm described in the previous section and used to produce results in Figure 1. Therefore, for any partition of the time $\mathbf{T}$, we know how to probabilistically estimate the parameters $\mathbf{\Omega}$.

The question now is to estimate the number and the position of discontinuities $\mathbf{T}$. This is given by the marginal posterior $p(\mathbf{T}|\mathbf{d})$.

2.2.3 The marginal posterior $p(\boldsymbol{T}|\boldsymbol{d})$

$p(\mathbf{T}|\mathbf{d})$ describes the probability of the time partition $\mathbf{T} = [T_0, T_2, \ldots T_k]$ given the full dataset of observed magnitudes. It can be obtained by integrating the full posterior $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ over the parameters $\mathbf{\Omega} = [\omega_1, \omega_2, ..., \omega_{k+1}]$:

$$p(\mathbf{T}|\mathbf{d}) = \int_{\mathbf{\Omega}} p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d}) d\Omega \tag{21}$$

According to Bayes' rule, the posterior density function $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ is proportional to the product of the likelihood $p(d|\mathbf{\Omega}, T)$ times the joint prior $p(\mathbf{\Omega}, \mathbf{T})$.

$$p(\mathbf{T}|\mathbf{d}) \propto \int_{\mathbf{\Omega}} p(\mathbf{d}|\mathbf{\Omega}, \mathbf{T}) p(\mathbf{\Omega}, \mathbf{T}) d\Omega \tag{22}$$

The joint prior $p(\mathbf{\Omega}, \mathbf{T})$ can be decomposed according to the property of joint density distributions $p(\mathbf{\Omega}, \mathbf{T}) = p(\mathbf{T}) \times p(\mathbf{\Omega}|\mathbf{T})$. Applied to equation (eq.22), we get :

$$p(\mathbf{T}|\mathbf{d}) \propto p(\mathbf{T}) \times \int_{\Omega} p(\mathbf{d}, \mathbf{\Omega}|\mathbf{T}) p(\mathbf{\Omega}|\mathbf{T}) d\Omega \tag{23}$$

with $p(\mathbf{T})$, the prior distribution for the time partitions and $p(\mathbf{\Omega}|\mathbf{T})$, the prior for the models $\mathbf{\Omega}$ given a fixed temporal model which is also present in the conditional posterior (eq.18). The prior $p(\mathbf{T})$ is a joint distribution, where the prior for each discontinuity T_j is given by a uniform distribution bounded between $min(t_{obs})$ and $max(t_{obs})$. The prior $p(\mathbf{\Omega}|\mathbf{T})$ corresponds to the model priors described in the temporally-invariant case as described for the section above.

Considering that the sub-datasets $textbfd_j$ of a time model $\mathbf{T}$ are independent, and since dataset $\mathbf{d}_j$ only depends on parameters ω_j the full posterior can be expressed as :

$$p(\mathbf{T}|\mathbf{d}) \propto p(\mathbf{T}) \times \prod_{j=1}^{k+1} \left(\int_{\omega_j} p(\mathbf{d}_j|\omega_j, \mathbf{T}) p(\omega_j|\mathbf{T}) d\omega_j \right) \tag{24}$$

where, $p(\mathbf{d}_j|\omega_j, \mathbf{T})$ is the likelihood of observing the subset $\mathbf{d}_j$ between $[T_j, T_{j+1}]$ according to the local model ω_j and can be estimated using equation (eq.12) obtained in the time-invariant case.

These integrals can be estimated using importance sampling. That is, for a large number of realizations x_i, $i = (1, \dots, N)$, randomly drawn from a distribution p(x) :

$$\int p(x) f(x) dx \approx \frac{1}{N} \times \sum_{i=1}^{N} f(x_i) \tag{25}$$

Applied to (eq.25), we have:

$$p(\mathbf{T}|\mathbf{d}) \propto p(\mathbf{T}) \times \left(\frac{1}{N_\omega} \right)^{(k+1)} \prod_{j=0}^{k} \left(\sum_{i=1}^{N_\omega} p(d_j|\omega_{j_{(i)}}, \mathbf{T}) \right) \tag{26}$$

where for each period j, $\omega_{j_{(i)}} = [\beta_j, \mu_j, \sigma_j]_{(i)}$ for $i = (1, \dots, N_\omega)$ are a set of realizations randomly drawn from the uniform prior distributions $p(\omega_j|\mathbf{T})$.

From equation (eq.26) we see that the marginal posterior is proportional to the product of the prior at temporal discontinuities $p(\mathbf{T})$ and the product of the mean likelihoods $p(d_j|\omega_{j_{(i)}}, \mathbf{T})$, computed over N_w realisations of the model priors, between each temporal discontinuity. In this way, adding an extra discontinuity will be valuable only if it sufficiently increases the local likelihood $p(d_j|w_{j_{(i)}}, \mathbf{T})$ to counterbalance this first effect.

Therefore, this methodology based on a Bayesian framework inherently follows the principle of parsimony, finding a balance between finding a simple model with a low number of temporal discontinuities, k, and maximizing the overall likelihood $p(\mathbf{d}|\Omega, \mathbf{T})$.

2.2.4 The reversible-jump Markov-chain Monte-Carlo algorithm (rj-McMC)

The marginal posterior $p(\mathbf{T}|\mathbf{d})$ can be numerically approximated with equation (eq.26) but only for a given partition $\mathbf{T}$. One way to estimate the full distribution $p(\mathbf{T}|\mathbf{d})$ is through a Monte Carlo exploration over the space of temporal discontinuities $\mathbf{T}$. The solution is then a large ensemble of partition vectors $\mathbf{T}^l(l = 1 \ldots N_l)$, with N_l the number of realizations $\mathbf{T}^l$, whose distribution approximates the target solution $p(\mathbf{T}|\mathbf{d})$.

As the dimension of $\mathbf{T}$ varies with the number of discontinuities k, $p(\mathbf{T}|\mathbf{d})$ is a trans-dimensional function and cannot be explored using a standard Metropolis algorithm (Metropolis et al., 1953; Hastings, 1970). One of the most popular technique for exploring a trans-dimensional posterior is the rj-McMC method (e.g. Green, 1995; Sambridge et al., 2006, 2013) and more specifically the birth-death McMC algorithm (e.g. Geyer & Møller, 1994). The rj-McMC algorithm, used in many geophysical inverse problems (e.g. Gallagher et al., 2009, 2011; Bodin et al., 2012), allows both the model parameters and the model dimension (i.e. the number of parameters) to be inferred. The rj-McMC follows the general principles of the McMC approach by generating samples from the target distribution. A Markov chain follows a random walk, where at each step, a proposed model $\mathbf{T}^{(p)}$ is generated by randomly modifying a current model $\mathbf{T}^{(c)}$ (Figure 3). This proposed model is then either accepted (and replaces the current model) or rejected. In this way, each step of the rj-McMC is a part of a chain converging to the target distribution.

The convergence is considered sufficient by monitoring the evolution of the number of discontinuities towards a stable value and when the rate of accepted models falls in the range of 20% to 40%. Details about the algorithm are given in Appendix. We also refer the reader to (Bodin et al., 2012) for further details.

2.2.5 *Appraising the full posterior distribution* $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$

As a reminder, the solution to our inverse problem is the full posterior solution $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$ that describes the temporal changes of β, μ, and σ. As shown in equation (eq.17), this posterior can be written as the product of the marginal distribution $p(\mathbf{T}|\mathbf{d})$ describing the probability of temporal changes and a conditional distribution $p(\mathbf{\Omega}|\mathbf{T}, \mathbf{d})$ for the parameters of the frequency-magnitude distribution, given a set of temporal changes.

By decomposing in such a way the posterior distribution into a conditional and a marginal distribution, the Metropolis-Hastings rule is simplified by only simulating a trans-dimensional temporal point process for vector $\mathbf{T}$ (e.g. Geyer & Møller, 1994; Green, 1995). With the rj-McMC algorithm, we have a numerical way to approximate the marginal probability distribution about the number and position of temporal changes $p(\mathbf{T}|\mathbf{d})$ (eq.26).

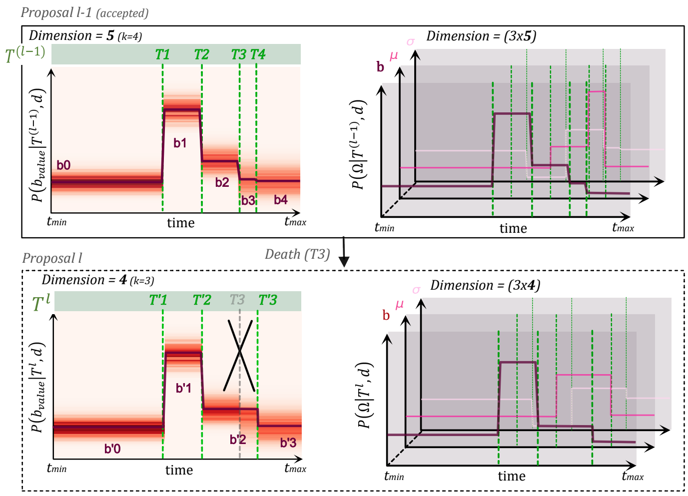

Figure 3. (Top)Left, $P(\beta|\mathbf{T}_{l-1}, \mathbf{d})$ over the 5 temporal segments ($k = 4$) of the proposed temporal model $\mathbf{T}_{l-1}$ for proposition $l-1$ of the rj-McMC. On the right, the mean likelihood over the 5 temporal segments for the three marginals posterior $p(\beta|\mathbf{T}_{l-1}, \mathbf{d})$, $P(\mu|T_{l-1}, \mathbf{d})$ and $P(\sigma|\mathbf{T}_{l-1}, \mathbf{d})$. (Bottom) New proposal model T_l in case of a death proposition of the rj-McMC. The two figures are the same as above : $P(\beta|\mathbf{T}_l, \mathbf{d})$ and the marginals computed for the proposed temporal model with a lower dimension (k=3).

In addition, for each sampled temporal model $\mathbf{T}^l$ proposed at iteration l of the rj-McMC, we are able to easily compute the conditional probability $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T}^l)$: the probability distribution of β, μ, and σ for the given model $\mathbf{T}^l$ (eq. 20).

At the completion, the full distribution for β, μ, and σ can therefore be obtained by summing the all the distributions $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T}^l)$ for all the sampled models $\mathbf{T}^l \in T^{(c)}$ (Figure 2.B,C). In practice, at each time-bin over an arbitrarily fine grid, the full probability distribution of β is the sum of all the marginal densities at this time over the ensemble solution for $\mathbf{T}$(Figure 2.B,C). In this way, the mean and the standard deviation for our three parameters can be obtained as a smooth function of time (see Figure 2.C).

2.3 Synthetic test

2.3.1 Generated Data

We simulate a synthetic earthquake catalog of 5683 independent events following frequency-magnitude distributions as realistic as possible with some temporal variations in b_{value} and detectability. In this section, we only consider a dataset with abrupt and discontinuous changes in the three parameters of the frequency-magnitude distribution. More precisely, the earthquake catalog is generated as follows :

- A discontinuity corresponds to the time when at least one of the three parameters of the frequency-magnitude distribution changes. We generate a catalogue with six temporal discontinuities that we aim to recover. Therefore, the catalog is the combination of seven temporal subsets.
- Within each temporal sub-dataset, earthquake magnitudes are drawn from a Gutenberg-Richter law characterized by a b_{value} specified in the Table 1.
- Earthquake occurrence is generated according to a basic epidemic-type aftershock sequence (Ogata, 1988) with a constant background rate. Each generated earthquake can be followed by aftershocks according to the aftershock productivity law (Utsu, 1972). Aftershock occurrence time is modelled by the Omori power law (Omori, 1894). The ETAS parametrization does not vary temporally. In particular, to characterise the temporal occurrence of aftershocks, we keep a constant p_{etas} value of 1.1 and an α_{etas} value of 1.5. For now, we do not generate the detectability variations coming from the short-term incompleteness (e.g. Ogata & Katsura, 2006; Helmstetter et al., 2006) following large earthquakes.
- We thin this ETAS earthquake catalog using the error detection law. Each event has a probability of being detected and preserved, or undetected and removed, according to its magnitude and some chosen μ and σ (see eq.9). Each pair of μ and σ for each of the seven sub-datasets are specified in the Table 1.

The dataset is made to test the capabilities of our algorithm and approach. To that aim, there are some periods where b_{value} remains constant while the detectability varies (S2 and S3) (Figure 4). We also include small detectability or b_{value} changes (ex:S5 and S6). For the period with the highest detectability (S3), earthquakes with magnitudes as low as 0.01 are detected. Considering a definition of $M_c = \mu + \sigma$, the completeness magnitude varies at most between subsets from 1.7 to 0.65.

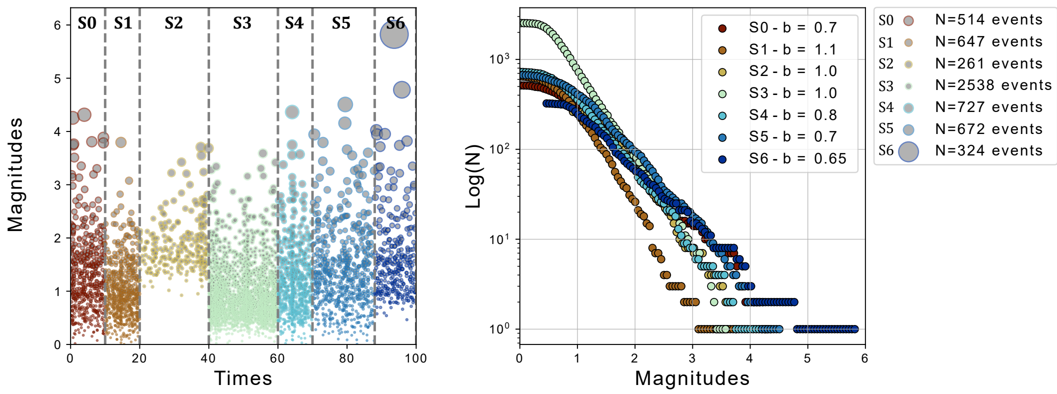

Figure 4. (A) Temporal distribution of magnitudes over time for the 7 data subsets. Each subset is represented with a different color. Grey vertical lines are the positions of the fixed temporal discontinuities. (B) Frequency-magnitude distribution of each synthetic data subset in logarithmic scale, respective colors from (A) are conserved. The slopes are related to the respective b_{value}.

Subset	S0	S1	S2	S3	S4	S5	S6
N_{events}	514	647	261	2538	727	672	324
b_{value}	0.70	1.10	1.00	1.00	0.80	0.70	0.65
μ	0.80	0.80	1.50	0.50	0.75	0.90	0.90
σ	0.35	0.30	0.20	0.15	0.30	0.40	0.15

Table 1. Set of chosen values for b_{value}, μ and σ for each of the 7 data subsets. N_{events} corresponds to the number of earthquakes detected of each sub-set (after the detection thinning).

2.3.2 Results and comparison

Let us start by applying classical approaches on this synthetic dataset. Ignoring temporal variations of b_{value} and considering a constant completeness magnitude of 1.8, the frequentist approach (Aki, 1965) gives the maximum likelihood estimate over the complete catalogue of $b_{value} = 0.81 \pm 0.03$. This illustrates how the uncertainties can be underestimated when using the classical approach without considering temporal variations. However, if we assume that we know the position of temporal discontinuities and that the completeness magnitude is correctly estimated using $M_{ci}^{true} = \mu_i + \sigma_i$ for $(i = 0, \ldots, 6)$, the true b_{value} can be recovered by the maximum likelihood estimate within its uncertainties. However, for applications to real earthquake catalogues, the temporal discontinuities are mostly unknown or at least ambiguous. Classical approaches therefore deal with temporal variations by estimating the b-value over a moving window.

Using a moving-window of 600 events and a constant completeness magnitude, Figure (5.A) shows that depending on the temporal sub-dataset the maximum likelihood estimate tends to under-estimate or over-estimate the b_{value} depending on the choice of M_c and the true detectability. In particular for (S2) with a low detectability, the choice of the completeness magnitude has a major influence on the results. This large variabil-

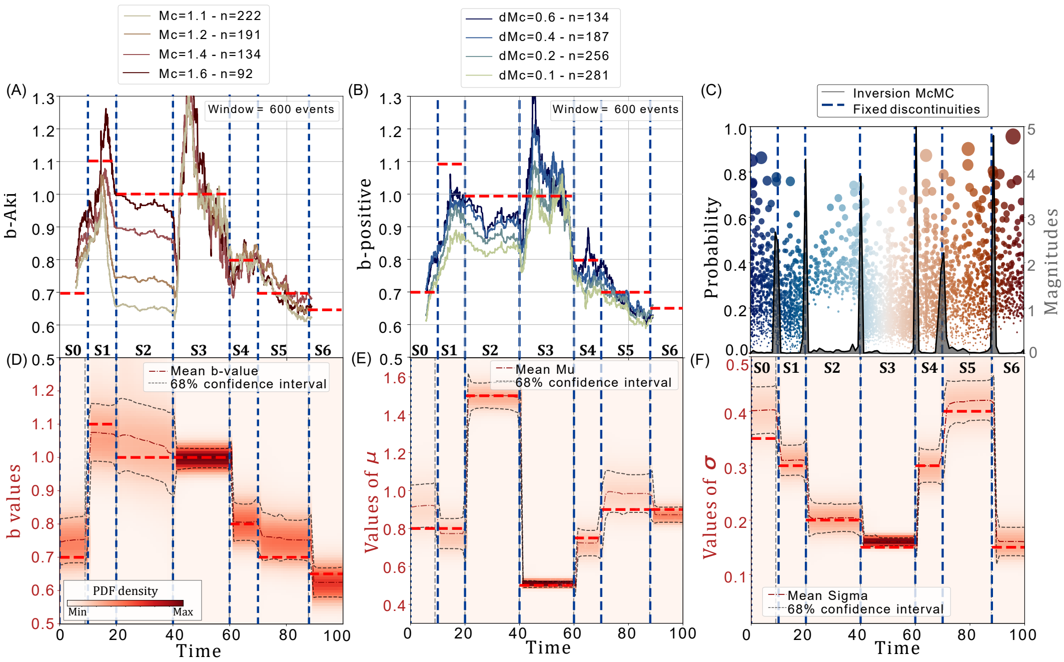

Figure 5. (A)Temporal variations of b_{value} using the classical maximum likelihood estimate (Aki, 1965; Utsu, 1966). The continuous curves are the moving-window estimate for a window size of 600 events with different choices of completeness magnitudes depending on the color-scale. (B) Temporal variations of b_{value} using the b-positive approach (Van der Elst, 2021). The continuous curves are the moving-window estimate for a window size of 600 events with different choices of difference threshold dM_c depending on the color-scale. (C) Probability of temporal discontinuities at the completion of the rj-McMC algorithm and the temporal distribution of magnitudes in the background (D) Marginal density distribution of b_{value} over time (E) Marginal density distribution of μ over time (F) Marginal density distribution of σ over time. For each subplot, the red horizontal lines represent the true parameter values fixed for each data subset, while the blue vertical lines denote the true values of temporal discontinuities as per Table 1.

ity of solutions can be partly corrected by the use of the b-positive approach (Van der Elst, 2021) (5.B). The b-positive approach (Van der Elst, 2021) uses moving-windows to infer temporal variations of b_{value} without being biased by the continuous decay of completeness for mainshock-aftershocks sequences. However, we show, that for discontinuous changes of completeness, the b-positive depends on the choice of the difference threshold dM_c. Thus, dM_c also needs to be adapted over time to correct for large variations in 'background' incompleteness, such as those between (S2) and (S3), to mitigate the risk of over-interpreting some temporal variations. This observation has also been highlighted by several recent studies (e.g. Lippiello & Petrillo, 2024).

These methods can be very efficient and fast but they use truncated datasets, which can result in b_{value} estimations based on very few events for periods with low detectability. With our new transdimensional approach, b_{value} and completeness are jointly inferred together with a probabilistic estimate of temporal changes from the entire dataset.

We applied our method to the synthetic dataset and conducted 50 parallel rj-McMC explorations with different initialisations for $\mathbf{T}$ in order to efficiently explore the model space. We present the results derived from the final stack of posterior densities coming from the 50 runs, each encompassing 5000 proposed models of temporal discontinuities (Figure 5.C,D,E,F). In total, 250000 temporal models were proposed out of which approximately 41800 were accepted upon completion. The six temporal discontinuities are well retrieved and clearly identified (Figure 5.C). Analysis of the number of discontinuities at each step indicates that the algorithm converged towards the correct value in less than 2000 iterations, even though all initialisations began with random values of k ranging from 4 to 12. A burn-in period of 1000 iterations has been set to disregard initial iterations. The maximum number of discontinuities allowed was set to 40 and did not affect the random-walk. For the 50 runs, the mean acceptance rate is around 30% for the move cases and 10% for the birth and death cases. These values are consistent with acceptance values obtained by other applications of rj-McMC algorithm (Gallagher et al., 2009).

Our approach allows to display the temporal evolution of the b_{value} of the Gutenberg-Richter law along with the two parameters describing the detectability (Figure 5.D,E,F). At each time-bin over a 100 grid, the full probability distribution of b_{value} is the sum of all the marginal densities that have been accepted. For the three parameters, the probability distribution comprises the true value even for periods with low detectability (Figure 5.D.E.F). Specifically, for period S2 containing 261 events, the b_{value} estimated by the transdimensional approach between 1.05 and 1.0 is close to the true value 1.0 with a 68% confidence interval of ± 0.1. As this period involves the fewest events, the relative probability is the lowest. This confidence intervals narrows to ± 0.02 for the period S3 containing 2538 events. Despite a large increase of detectability between these two periods, the b_{value} remains stable. We demonstrate that the transdimensional framework retrieves the true values of the three parameters governing the frequency-magnitude distribution of earthquakes over time for a synthetic case. This approach gives larger uncertainties compared to those proposed by classical methods, primarily because it accounts, in addition to the number of earthquakes used, for existing correlations between parameters that classical approaches fail to resolve.

3 Application to a real earthquake catalog

A significant challenge is thus to select a region where the b_{value} and the completeness are homogeneous, and the Gutenberg-Richter law is valid for the chosen sub-dataset of earthquakes. The transdimensional approach presented here addresses this issue in time, yet the selection of a catalog with spatially homogeneous b_{value} remains crucial to better understand the physical meaning of b_{value} temporal variations.

Another challenge addressed by the transdimensional approach is the temporal variations in completeness magnitude, influenced by mainshock-aftershock sequences and seasonal variations due to anthropogenic or meteorological factors(e.g. Iwata, 2013). Here, we choose an earthquake catalog with expected variations in completeness to evaluate the approach's efficiency and compare results with other methods.

3.1 Data : Far-Western Nepal seismicity

Our first application of the transdimensional approach to investigate b_{value} variations focus on the temporal evolution of seismicity in a very seismically active region of Nepal. We use the earthquake data collected during two years by the Himalayan Karnali Network (HiKNet), the first dense seismological network of 15 temporary stations deployed in far-western Nepal. This earthquake catalog is derived from two studies (Hoste

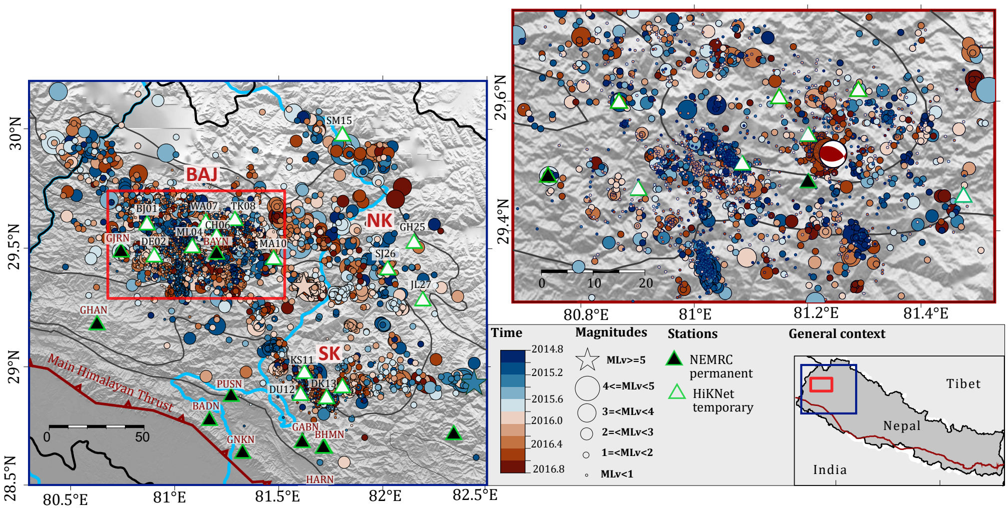

Figure 6. Far-western Nepal seismicity recorded during two years by the temporary seismological experiment of the Himalayan-Karnali network (HiKNet) (white triangles) and the permanent seismological network of the National Earthquake Monitoring Research Center (NEMRC) (black triangles). We use the earthquake catalog and focal mechanism from Laporte et al. (2021). NK: North Karnali sector, SK: South Karnali sector. In this study, we focus our analysis on the geographical subset represented by the inner red rectangle comprising 2593 events : the Bajhang region (BAJ).

et al., 2018; Laporte et al., 2021) which focuses on the spatio-temporal analysis for seismotectonic interpretation.

In Nepal, the main feature of seismicity is a belt of intense microseismicity which is located at depth on the locked portion of the Main Himalayan Thrust (MHT), (e.g. Ader et al., 2012). The MHT is the main active thrust fault which accommodates most of the shortening between the Indian plate and the Tibetan plateau. The seismicity is interpreted as resulting from stress build-up on the locked portion of the MHT. It exhibits a multimodal behavior, generating intermediate earthquakes ($M > 5$) that partially rupture the MHT, as well as large ($M > 7$) and great earthquakes ($M > 8$) that may rupture several lateral segments of the MHT, sometimes up to the surface (Dal Zilio et al., 2019).

In the area of interest in this paper, the most recent great earthquake occurred in 1505 A.D. according to historical records supported by paleoseismological evidence (Hossler et al., 2016; Riesner et al., 2021).

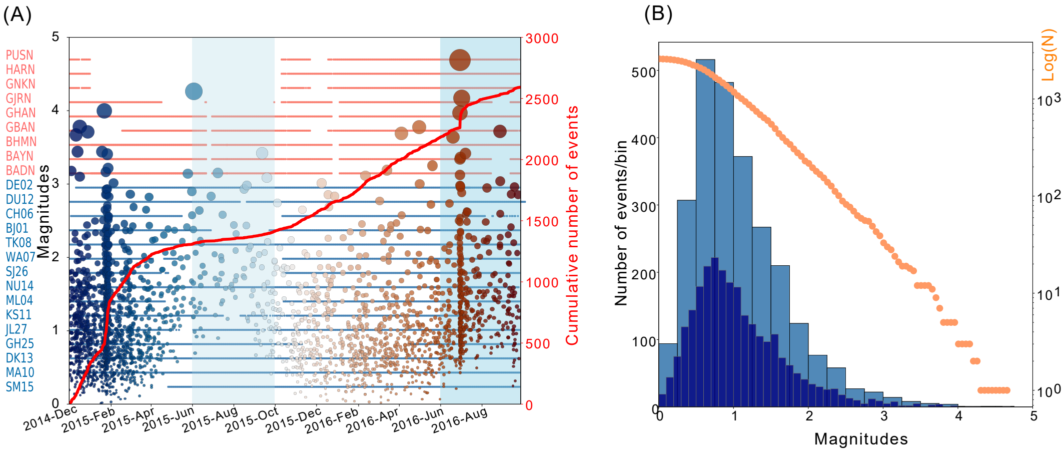

Figure 7. (A) Temporal distribution of magnitudes during two years of the HiKNet experiment in the Bajhang region. The color scale is the same as Figure 6. The red curve is the temporal evolution of the cumulative number of events. Horizontal lines correspond to station availability. Blue shades are the usual monsoon periods in Nepal. (B) Frequency-magnitude distribution in normal (blue histograms) and logarithmic (orange curve) scale. The two histograms are histograms for bins of 0.2 and 0.1 in magnitude, respectively.

Between December 2014 and September 2016, the temporary experiment recorded almost 4500 earthquakes in this region. The seismicity is structured into three seismic belts: one large belt in the westernmost part (BAJ) and two separate belts in the east (SK and NK) (Figure 6). Each belt contains several seismic clusters of different size and spatio-temporal behaviour. Most of them are located at mid-crustal depths (15−20km) close to mid-crustal structures such as the toe of mid-crustal ramps, which accumulate most of the interseismic strain on the fault. The geometry of the MHT fault is considered to be the primary factor influencing microseismicity. The focal mechanisms of the largest earthquakes are consistent with thrust faulting.

Facing challenges due to temporal variations of the detectability caused by station losses and monsoon periods, we chose this dataset to study temporal variations in b_{value} using the transdimensional approach. In Nepal, the monsoon season typically spans from early June to September, peaking in July and August. During this period, the high-frequency seismic noise increases, leading to fewer detected earthquakes. Additionally, in 2015, 9 out of 24 instruments were successively disconnected by storms. These typical monsoon periods are highlighted in blue in Figure 7.A, along with the availability of the 24 instruments throughout the experiment

To ensure spatial homogeneity of b_{value}, we focused on a geographical region of 94km by 57km, corresponding to the seismicity of the western belt, which is the most instrumented. 2593 earthquakes were recorded within this specific region. The cumulative number of earthquakes reveals distinct periods of seismic activity (see Figure 7.A). Using the frequentist approach of Aki for a completeness magnitude of 1.5 on the full time period, we get a b_{value} of 0.82±0.08 (Figure 7.B) consistent with the b_{value}obtained in far-western Nepal and more generally in Central and Eastern Nepal (e.g. Laporte et al., 2021). This low b_{value} is also consistent with the thrust-type faulting style (e.g. Schorlemmer et al., 2005).

3.2 Results

For this real application, we configure the rj-McMC algorithm to conduct 15,000 iterations for each individual run. However, as for the synthetic case, we initiate 50 parallel runs, totaling 750,000 models tested starting from distinct random seeds. For each of them, we burn 4,000 iterations and thin the chain by keeping only 1 out of 5 accepted models.

The algorithm converges towards 9 temporal discontinuities (defined as being over a probability of 15%) after 5000 iterations (Figure 8.A). The total acceptance rate is 23% and is above 20% for the three case scenarios births/deaths/moves.

Moreover, comparing the position of these discontinuities with the magnitude distribution of the seismicity, we can see that these nine discontinuities are coincident with some specific time periods of the dataset. We interpret and discuss each of them with respect to the stacked marginal densities of probability for b_{value}, μ and σ (Figure 8).

- S1 corresponds to the time of the M_{Lv} 4.0 earthquake of 22 January 2015, located at the base of the mid-crustal ramp of the MHT and followed by an increased seismic rate of about 300 events occurrence in 9 days (Hoste-Colomer et al., 2018; Laporte et al., 2021). For this seismic crisis, characterised as a large seismic swarm

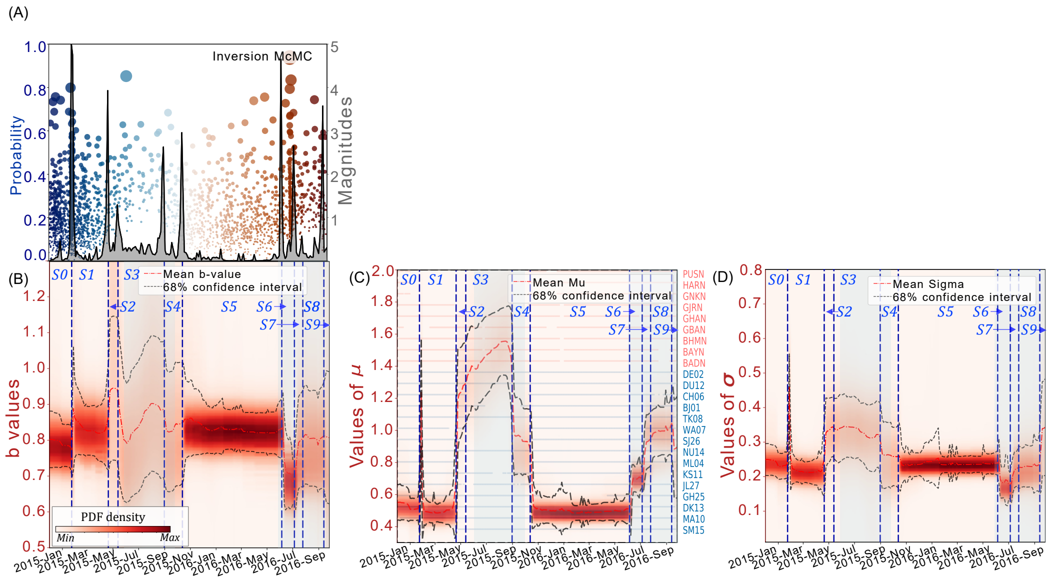

Figure 8. (A)Probability of temporal discontinuities at the completion of the rj-McMC algorithm and the temporal distribution of magnitudes in the background. The color-scale is the same as Figure 7. Here, a discontinuity is defined by the 15% probability threshold. (B) Marginal density distribution of b_{value} over time. (C) Marginal density distribution of μ over time. Horizontal lines are the stations availability in red for the permanent network and blue for the temporary network. (D) Marginal density distribution of σ over time For B), C) and D), the thin red dashed line is the mean probability for the respective marginal density distribution. The thin black lines are the $\pm 1\sigma$ uncertainty. The blue vertical dashed lines are the position to the solution of temporal discontinuities according to A).

by Hoste-Colomer et al. (2018), the b_{value} and the two detectability parameters increase suddenly and then return to their previous value.

- S2 corresponds to the loss of three stations (ML04, SJ26 and GJRN)(Figure 8.C), two of which were sited in the considered region (Figure 6). The confidence interval for every parameter becomes wider which is in accordance with fewer events detected. Only μ and σ present a clear discontinuous change. Specifically, μ increases from 0.5 ± 0.1 to 1.3 ± 0.3, and σ from 0.25 ± 0.05 to 0.32 ± 0.1. The b_{value} remains within the uncertainties of S1 while its uncertainty becomes larger.

- S3 corresponds to the loss of two stations in the center of the considered region (WA07 and CH06) (Figure 8). In the region, only 5 out of 9 stations are available during the monsoon period. Between S3 and S4, the monsoon periods starts and the μ parameter increases along with degradation in detectability, the σ parameter is not affected as much as μ. The mean b_{value} stays within the confidence interval of previous periods with a larger uncertainty. There is no statistical evidence of b_{value} variations during the monsoon period.
- S4 corresponds to a brutal improvement in the detectability. μ decreases from 1.6 ± 0.2 to 1.0 ± 0.2. This time likely corresponds to the early end of the monsoon period at the beginning of September 2015.
- S5 Both detectability parameters come back to the pre-monsoon values with the return of all 6 lost stations in October 2015. The b_{value} remains constant around 0.85 ± 0.1 during that time with a narrower confidence interval.
- S6 The density probability of b_{value} has a significant decrease at the beginning of June 2016 from 0.85 ± 0.1 to 0.7 ± 0.1. This time also corresponds to the expected beginning of the monsoon period, μ increases significantly and σ decreases.
- S7 corresponds to the onset of the second largest seismic crisis recorded in that region. On the 29th of June a $M_{Lv}4.8$ earthquake occurred and was followed by several aftershocks including two $M_{Lv} > 4$ earthquakes. It seems that after this crisis that lasted 12 days, the b_{value} comes back to its value of 0.85 ± 0.15. However this variation is not statistically significant because the confidence interval also increases due to the decrease in detectability. In fact, the μ parameter describing the mean of the detectability function keeps increasing in steps. One station (KS11) at 50 km from the area of interest is also interrupted.
- S8 Another station from the area of interest becomes unavailable (CH06) and the confidence intervals become wider. b_{value} and σ are constant during that period and approximately equal to the mean value they had during the two years 0.8 and 0.25 for b_{value} and σ respectively. The mean detectability μ seems to increase slightly but this increase is within the confidence interval. However this increase can be attested by the simultaneous widening of confidence intervals that might be due to fewer events detected during the monsoon.
- S9 in September 2015 corresponds to the end of the experiment, uncertainties are becoming wider with the disconnection of the firsts temporary stations.

Looking back at these results, the analysis of temporal variations of b_{value} (Figure 8) shows that there is no statistical evidence of temporal variations during the two years of the temporary experiment at the exception of a short period preceding one seismic crisis (S6). The b-Bayesian approach recovers as well the large variations of background detectability which can be explained by the loss of stations during summer 2015 and by the higher seismic noise during the monsoon periods. In particular, the μ parameter from the detectability function is the most sensitive to detectability changes and can be used as a proxy for deciphering detectability variations. This additional information, which is not given by traditional approaches, can be very useful for the characterization of the efficiency of a seismic network over time. Moreover, these large variations of detectability are taken into account in the Bayesian estimation of the b_{value} and act in the spread of its uncertainties along time. When all stations of the temporary network were available, the 1σ uncertainty of the b_{value} is reduced to ± 0.05. The widening of the posterior density function during the monsoon periods shows that during these periods the information contained in this earthquake catalog is not good enough to decipher some seasonal variations of the b_{value} even though it might exists. Outside the scope of this study, as the temporary network recorded data between 2014 and 2016, this particular sector of far-western Nepal has experienced 5 moderate ($M_L > 5$) damaging earthquakes since 2022, while the temporary experiment recorded none in two years.

4 Discussion and Perspectives

4.1 Comparison between classical approaches

Most of the time, temporal variations in b_{value} have been studied by applying the maximum likelihood approach (Aki, 1965; Utsu, 1966; Y. Shi & Bolt, 1982) over sliding time windows (e.g. Nuannin et al., 2005; Cao & Gao, 2002; Nanjo et al., 2012; Gulia & Wiemer, 2019). These techniques are dependent on the accuracy of the estimation of the time variations of the completeness magnitude (e.g. Helmstetter et al., 2006; Woessner & Wiemer, 2005) (Figure 9.A). More recently, the b-positive approach introduced by Van der Elst (2021) has been shown to be insensitive to the short-term incompleteness coming from mainshock-aftershocks sequences. Both approaches are very efficient and do not require any prior information on the b_{value} but they use a small part of the dataset only. The uncertainty in b_{value} is often estimated using formulas from Aki (1965) or Utsu (1966) within the sliding window, or by employing the bootstrap method(e.g.

Woessner & Wiemer, 2005). Importantly, this uncertainty is consistently estimated independently of the uncertainty associated with the completeness magnitude nor their correlation.

Comparing the outcomes derived from traditional approaches with those obtained using our b-Bayesian method on the earthquake catalog of far-western Nepal reveals intrinsic differences among these methodologies (Figure 9). Notably, both classical approaches exhibit sensitivity to the selection of the magnitude threshold (M_c or dM_c), whereas the b-Bayesian method effectively captures the uncertainty arising from fluctuations in detectability. Our findings indicate that most b_{value} variations shown by the classical moving-window techniques fall within the uncertainties accounted for by the b-Bayesian approach. These small-scale variations can sometimes lead to over-interpretations of the temporal variations in b_{value}.

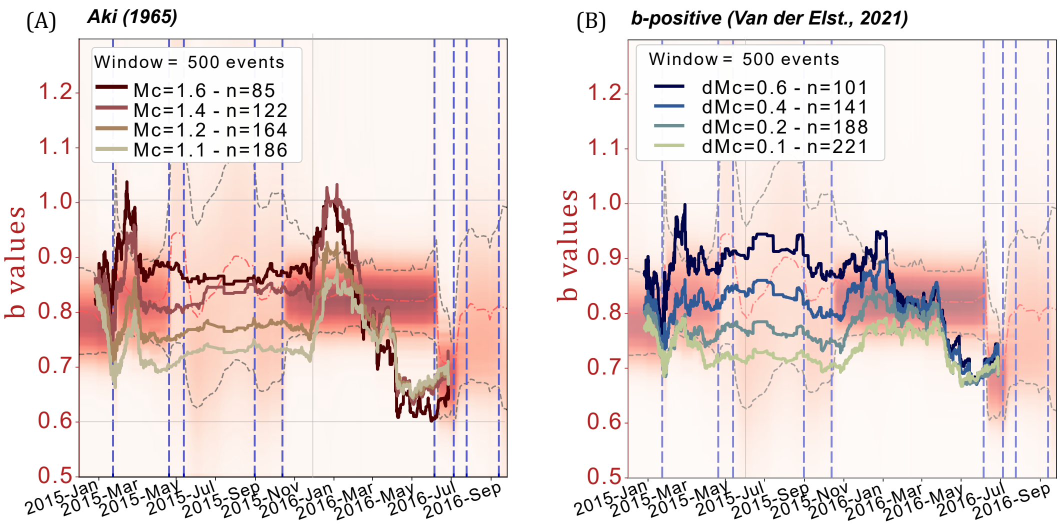

Figure 9. (A) Comparison of the temporal variations of b_{value} obtained using the b-Bayesian approach for far-western Nepal in red-shaded and the frequentist approach from Aki applied on a moving-window of 500 events as the 4 bold lines, depending on 4 values of completeness magnitude M_c. (B) Same as (A) but compared to the b-positive approach with 4 different values for dM_c. For both sub-figures, the legend gives the mean number of magnitudes kept for the b_{value} estimates in the windows, depending on the magnitude cut-off.

Despite using a moving-window of 500 events, this number is significantly reduced by truncation at M_c (up to only 17% of magnitudes retained for Aki's approach with $M_c =$ 1.6) or by using only positive magnitude differences for the b-positive method ($> 50\%$)

(Figure 9), while the b-Bayesian approach does not require any truncation and uses 100% of all the available data. Moreover, b-Bayesian proves to be the only tool for deciphering jointly the variations in detectability and can be used as a preliminary step before applying classical approaches to ensure that detectability is adequately considered.

In Table 2, we propose a comparison between the two classical approaches presented in this paper and the novel b-Bayesian approach. In conclusion, b-Bayesian proposes to use all the data available to invert for the temporal variations of three parameters related to the frequency-magnitude distribution. It uses Bayesian inference to capture the full density distributions and does not require any parametrization. This can be done at the cost of the computational time which is currently being reduced.

	Aki (1965)	b-Positive (van der Elst., 2021)	b-Bayesian (this study)
Inversion param.	b_{value}	b_{value}	b_{value}, μ, σ
Approach type	MLE	MLE	Bayesian
Uncertainties estimates	MLE	Bootstrap	Full PDF
Truncation	Mc $^{(*)}$	dMc	None
Used Data	$< 40\%$	$< 50\%$	100%
Temporal	Moving-window	Moving-window	Probabilistic
Comp. Time	Immediate $^{(**)}$	Immediate $^{(**)}$	Long $^{(***)}$

Table 2. Table of comparison between the two classical approaches from Aki (1965), van der Elst (2021) and b-Bayesian. $^{(*)}$ arbitrary $^{(**)}$ for one parametrization of Mc/dMc and choice of moving-window $^{(***)}$ no arbitrary parametrization

4.2 Perspectives

The frequency-magnitude distribution of earthquakes varies temporally and spatially. At the regional scale, b_{value} is thought to reflect the faulting style and the evolution of the state of stress on the faults (e.g. Schorlemmer et al., 2005; Gulia & Wiemer, 2019). This is also supported by experimental studies on the micro-failure (e.g. Scholz, 1968, 2015). Consequently, numerous studies have focused on monitoring the b_{value} at regional scales to discriminate foreshock and mainshocks sequences (e.g. Gulia et al., 2020;

Van der Elst, 2021). At the local scale, the b_{value} has also been proven valuable for describing the spatio-temporal behavior of seismic clusters (e.g. Farrell et al., 2009; Gui et al., 2020; Herrmann et al., 2022) or characterizing swarm-like sequences in relation to fluid-pressure (e.g. Hainzl & Fischer, 2002; Shelly et al., 2016; De Barros et al., 2019). We are now looking forward to applying b-Bayesian in these different contexts to discuss our results in comparison with previous studies and infer valuable information for characterizing seismic sequences.

The b-Bayesian method addresses temporal variations in the b_{value} using Bayesian inference. However, these variations are generally considered to be secondary compared to spatial variations (Wiemer & Wyss, 1997; Öncel & Wyss, 2000). Currently, both the b-Bayesian and classical approaches require a spatial subset of earthquakes with a homogeneous b_{value}, making it challenging to determine the adequacy of the dataset. Similar to the inversion of seismic velocities in tomography (e.g. Bodin & Sambridge, 2009), adapting the transdimensional approach to account for 2D spatial partitions could enable the capture of both temporal and spatial variations in the frequency-magnitude distribution. This is a future development of the method.

5 Appendix

5.1 Method : the Markov-chain Monte-Carlo

The Markov-chain is initialised by a randomized choice of a temporal model $T^{(c)} = [T_1^{(c)}, T_2^{(c)}, \dots T_k^{(c)}]$ with a number of discontinuities $k^{(c)}$ drawn between two values k_{min} and k_{max}. In practice, temporal discontinuities are considered as floating values in the range between $min(t_{obs})$ and $max(t_{obs})$. After this initialisation, at each iteration of the reversible-jump algorithm a new temporal model is proposed by making a random choice among three possibilities :

- (i) a death of a random temporal discontinuity $T_j^{(p)}$. Then, the proposed dimension is $k^{(p)} = k^{(c)} - 1$.
- (ii) a birth of a random temporal discontinuity $T_j^{(p)}$. Then, the proposed dimension is $k^{(p)} = k^{(c)} + 1$.
- (iii) a move of a random temporal discontinuity $T_j^{(p)}$ around its previous location. Then the proposed dimension remains $k^{(p)} = k^{(c)}$.

Each proposal (birth/death/move) has the same uniform probability of being drawn. For the moves, we randomly draw a time offset δ_t from a normal distribution around the previous time of a random time discontinuity of the current model. The standard deviation σ_{δ_t} of this normal distribution controls the efficiency of the exploration. A large standard deviation will produce large jumps with many rejected models and poor precision while a small standard deviation will produce very similar models and accept most of them. We follow the approach of (Gallagher et al., 2009) which tunes the value of σ_{δ_t} in order to get close a 20% acceptance rate. Every five hundred iterations, we monitor the acceptance rate and increase or decrease the σ_{δ_t} linearly with the deviation of the acceptance rate from the 20%.

The acceptance criterion is computed according to the Metropolis-Hastings rule in order to guide the chain towards the target distribution. Its general form for trans-dimensional functions is written as follows :

$$\alpha = min(1, prior\ ratio \times likelihood\ ratio \times proposal\ ratio \times |J|) \tag{27}$$

such as :

$$\alpha = min(1, \frac{p(T^{(p)})}{p(T^{(c)})} \times \frac{p(d|T^{(p)})}{p(d|T^{(c)})} \times \frac{q(T^{(c)}|T^{(p)})}{q(T^{(p)}|T^{(c)})} \times |J|) \tag{28}$$

where $|J|$ is the Jacobian of the transformation from the current model to the proposal and can be shown as equal to 1 (e.g. Gallagher et al., 2009; Bodin & Sambridge, 2009). For moves of discontinuities, the dimension remains fixed, proposals are symmetric, and the prior ratio is one. The acceptance criterium is simply given by the usual Metropolis criterion :

$$\alpha = min(1, \frac{p(d|T^{(p)})}{p(d|T^{(q)})}) \tag{29}$$

For birth proposals, we randomly draw the location of a new discontinuity from the prior distribution. In this way, the prior and proposal ratios cancel out for birth and death steps, and thus the acceptance criterion (**??**) also conveniently becomes the usual Metropolis criterion.

6 Open Research

Codes of the b-Bayesian will be made available on github upon publication. The earthquake catalog of far-western Nepal used in Section 3 is available in Supplementary Material of Laporte et al. (2021).

Acknowledgments

The authors are grateful to the LabeX Lyon Institute of Origins (ANR-10-LABX-0066) Lyon for its financial support within the Plan France 2030 of the French government operated by the National Research Agency (ANR).

References

Ader, T., Avouac, J.-P., Liu-Zeng, J., Lyon-Caen, H., Bollinger, L., Galetzka, J., ... others (2012). Convergence rate across the nepal himalaya and interseismic coupling on the main himalayan thrust: Implications for seismic hazard. *Journal of Geophysical Research: Solid Earth*, *117*(B4).

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.*, *43*, 237–239.

Avlonitis, M., & Papadopoulos, G. (2014). Foreshocks and b value: bridging macroscopic observations to source mechanical considerations. *Pure and Applied Geophysics*, *171*, 2569–2580.

Beauval, C., & Scotti, O. (2003). Mapping b-values in france using two different magnitude ranges: Possible non power-law behavior. *Geophysical research letters*, *30*(17).

Beauval, C., & Scotti, O. (2004). Quantifying sensitivities of psha for france to earthquake catalog uncertainties, truncation of ground-motion variability, and magnitude limits. *Bulletin of the Seismological Society of America*, *94*(5), 1579–1594.

Bodin, T., & Sambridge, M. (2009). Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, *178*(3), 1411–1436.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., & Rawlinson, N. (2012). Transdimensional inversion of receiver functions and surface wave dispersion. *Journal of geophysical research: solid earth*, *117*(B2).

Bolton, D. C., Shreedharan, S., Rivière, J., & Marone, C. (2020). Acoustic energy release during the laboratory seismic cycle: Insights on laboratory earthquake precursors and prediction. *Journal of Geophysical Research: Solid Earth*, *125*(8), e2019JB018975.

Cao, A., & Gao, S. S. (2002). Temporal variation of seismic b-values beneath northeastern japan island arc. *Geophysical research letters*, *29*(9), 48–1.

Carter, J. A., & Berg, E. (1981). Relative stress variations as determined by b-values from earthquakes in circum-pacific subduction zones. *Tectonophysics*, *76*(3-4), 257–271.

Chan, C.-H., Wu, Y.-M., Tseng, T.-L., Lin, T.-L., & Chen, C.-C. (2012). Spatial and temporal evolution of b-values before large earthquakes in taiwan. *Tectonophysics*, *532*, 215–222.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the seismological society of America*, *58*(5), 1583–1606.

Daniel, G., Marsan, D., & Bouchon, M. (2008). Earthquake triggering in southern iceland following the june 2000 ms 6.6 doublet. *Journal of Geophysical Research: Solid Earth*, *113*(B5).

De Barros, L., Baques, M., Godano, M., Helmstetter, A., Deschamps, A., Larroque, C., & Courboulex, F. (2019). Fluid-induced swarms and coseismic stress transfer: A dual process highlighted in the aftershock sequence of the 7 april 2014 earthquake (ml 4.8, ubaye, france). *Journal of Geophysical Research: Solid Earth*, *124*(4), 3918–3932.

Drouet, S., Ameri, G., Le Dortz, K., Secanell, R., & Senfaute, G. (2020). A probabilistic seismic hazard map for the metropolitan france. *Bulletin of Earthquake Engineering*, *18*(5), 1865–1898.

Dutfoy, A. (2020). Estimation of the gutenberg–richter earthquake recurrence parameters for unequal observation periods and imprecise magnitudes. *Pure and Applied Geophysics*, *177*(10), 4597–4606.

Farrell, J., Husen, S., & Smith, R. B. (2009). Earthquake swarm and b-value characterization of the yellowstone volcano-tectonic system. *Journal of Volcanology and Geothermal Research*, *188*(1-3), 260–276.

Gallagher, K., Bodin, T., Sambridge, M., Weiss, D., Kylander, M., & Large, D. (2011). Inference of abrupt changes in noisy geochemical records using transdimensional changepoint models. *Earth and Planetary Science Letters*, *311*(1-2),

182–194.

Gallagher, K., Charvin, K., Nielsen, S., Sambridge, M., & Stephenson, J. (2009). Markov chain monte carlo (mcmc) sampling methods to determine optimal models, model resolution and model choice for earth science problems. *Marine and Petroleum Geology*, *26*(4), 525–535.

Geffers, G.-M., Main, I. G., & Naylor, M. (2022, February). Biases in estimating *b* -values from small earthquake catalogues: how high are high *b* -values? *Geophysical Journal International*, *229*(3), 1840–1855. Retrieved 2024-04-04, from `https://academic.oup.com/gji/article/229/3/1840/6515950` doi: 10.1093/gji/ggac028

Geyer, C. J., & Møller, J. (1994). Simulation procedures and likelihood inference for spatial point processes. *Scandinavian journal of statistics*, 359–373.

Godano, C., Petrillo, G., & Lippiello, E. (2023, December). Evaluating the incompleteness magnitude using an unbiased estimate of the *b* value. *Geophysical Journal International*, *236*(2), 994–1001. Retrieved 2024-04-04, from `https://academic.oup.com/gji/article/236/2/994/7464075` doi: 10.1093/gji/ggad466

Godano, C., Tramelli, A., Petrillo, G., & Convertito, V. (2024). Testing the predictive power of b value for italian seismicity. *Seismica*, *3*(1).

Goebel, T. H., Kwiatek, G., Becker, T. W., Brodsky, E. E., & Dresen, G. (2017). What allows seismic events to grow big?: Insights from b-value and fault roughness analysis in laboratory stick-slip experiments. *Geology*, *45*(9), 815–818.

Green, P. J. (1995). Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, *82*(4), 711–732.

Gui, Z., Bai, Y., Wang, Z., Dong, D., Wu, S., & Li, T. (2020). Spatiotemporal seismotectonic implications for the izu–bonin–mariana subduction zone from b-values. *Seismological Research Letters*, *91*(3), 1679–1693.

Gulia, L., & Wiemer, S. (2019). Real-time discrimination of earthquake foreshocks and aftershocks. *Nature*, *574*(7777), 193–199.

Gulia, L., Wiemer, S., & Vannucci, G. (2020). Pseudoprospective evaluation of the foreshock traffic-light system in ridgecrest and implications for aftershock hazard assessment. *Seismological Research Letters*, *91*(5), 2828–2842.

Hainzl, S., & Fischer, T. (2002). Indications for a successively triggered rupture growth underlying the 2000 earthquake swarm in vogtland/nw bohemia. *Journal of Geophysical Research: Solid Earth*, *107*(B12), ESE–5.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications.

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

Herrmann, M., Piegari, E., & Marzocchi, W. (2022). Revealing the spatiotemporal complexity of the magnitude distribution and b-value during an earthquake sequence. *Nature Communications*, *13*(1), 5087.

Hoste-Colomer, R., Bollinger, L., Lyon-Caen, H., Adhikari, L., Baillard, C., Benoit, A., ... others (2018). Lateral variations of the midcrustal seismicity in western nepal: Seismotectonic implications. *Earth and Planetary Science Letters*, *504*, 115–125.

Ito, R., & Kaneko, Y. (2023, December). Physical Mechanism for a Temporal Decrease of the Gutenberg-Richter *b* -Value Prior to a Large Earthquake. *Journal of Geophysical Research: Solid Earth*, *128*(12), e2023JB027413. Retrieved 2024-04-04, from `https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2023JB027413` doi: 10.1029/2023JB027413

Iwata, T. (2013). Estimation of completeness magnitude considering daily variation in earthquake detection capability. *Geophysical Journal International*, *194*(3), 1909–1919.

Keller, M., Pasanisi, A., Marcilhac, M., Yalamas, T., Secanell, R., & Senfaute, G. (2014). A bayesian methodology applied to the estimation of earthquake recurrence parameters for seismic hazard assessment. *Quality and Reliability Engineering International*, *30*(7), 921–933.

Kwiatek, G., Martínez-Garzón, P., Becker, D., Dresen, G., Cotton, F., Beroza, G. C., ... Bohnhoff, M. (2023). Months-long seismicity transients preceding the 2023 mw 7.8 kahramanmaraş earthquake, türkiye. *Nature Communications*, *14*(1), 7534.

Laporte, M., Bollinger, L., Lyon-Caen, H., Hoste-Colomer, R., Duverger, C., Letort, J., ... others (2021). Seismicity in far western nepal reveals flats and ramps along the main himalayan thrust. *Geophysical Journal International*, *226*(3), 1747–1763.

Li, Y., & Chen, X. (2021). Variations in apparent stress and b value preceding the 2010 m w 8.8 bio-bío, chile earthquake. *Pure and Applied Geophysics*, 1–17.

Lippiello, E., & Petrillo, G. (2024). b-more-incomplete and b-more-positive: Insights on a robust estimator of magnitude distribution. *Journal of Geophysical Research: Solid Earth*, *129*(2), e2023JB027849.

Lombardi, A. M. (2021, March). A Normalized Distance Test for Co-Determining the Completeness Magnitude and *b* -Value of Earthquake Catalogs. *Journal of Geophysical Research: Solid Earth*, *126*(3), e2020JB021242. Retrieved 2024-04-04, from `https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020JB021242` doi: 10.1029/2020JB021242

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, *21*(6), 1087–1092.

Mignan, A., & Woessner, J. (2012). Theme iv—understanding seismicity catalogs and their problems. *Community online resource for statistical seismicity analysis*.

Mogi, K. (1962). Magnitude-frequency relation for elastic shocks accompanying fractures of various materials and some related problems in earthquakes. *Bull. Earthq. Res. Inst., Univ. Tokyo*, *40*, 831–853.

Morales-Yáñez, C., Bustamante, L., Benavente, R., Sippl, C., & Moreno, M. (2022). B-value variations in the central chile seismic gap assessed by a bayesian transdimensional approach. *Scientific Reports*, *12*(1), 21710.

Mori, J., & Abercrombie, R. E. (1997). Depth dependence of earthquake frequency-magnitude distributions in california: Implications for rupture initiation. *Journal of Geophysical Research: Solid Earth*, *102*(B7), 15081–15090.

Nanjo, K., Hirata, N., Obara, K., & Kasahara, K. (2012). Decade-scale decrease inb value prior to the m9-class 2011 tohoku and 2004 sumatra quakes. *Geophysical Research Letters*, *39*(20).

Nuannin, P., Kulhanek, O., & Persson, L. (2005). Spatial and temporal b value anomalies preceding the devastating off coast of nw sumatra earthquake of december 26, 2004. *Geophysical research letters*, *32*(11).

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

Ogata, Y., & Katsura, K. (1993). Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues. *Geophysical Journal International*, *113*(3), 727–738.

Ogata, Y., & Katsura, K. (2006). Immediate and updated forecasting of aftershock hazard. *Geophysical research letters*, *33*(10).

Omori, F. (1894). On after-shocks. *Seismological journal of Japan*, *19*, 71–80.

Öncel, A. O., & Wyss, M. (2000). The major asperities of the 1999 m w= 7.4 izmit earthquake defined by the microseismicity of the two decades before it. *Geophysical Journal International*, *143*(3), 501–506.

Petruccelli, A., Schorlemmer, D., Tormann, T., Rinaldi, A. P., Wiemer, S., Gasperini, P., & Vannucci, G. (2019). The influence of faulting style on the size-distribution of global earthquakes. *Earth and Planetary Science Letters*, *527*, 115791.

Plourde, A. (2023, September). Assessing earthquake rates and b-value given spatiotemporal variation in catalog completeness: Application to Atlantic Canada. *Seismica*, *2*(2). Retrieved 2024-04-04, from `https://seismica.library.mcgill.ca/article/view/384` doi: 10.26443/seismica.v2i2.384

Ringdal, F. (1975). On the estimation of seismic detection thresholds. *Bulletin of the Seismological Society of America*, *65*(6), 1631–1642.

Rivière, J., Lv, Z., Johnson, P., & Marone, C. (2018). Evolution of b-value during the seismic cycle: Insights from laboratory experiments on simulated faults. *Earth and Planetary Science Letters*, *482*, 407–413.

Sambridge, M., Bodin, T., Gallagher, K., & Tkalčić, H. (2013). Transdimensional inference in the geosciences. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*(1984), 20110547.

Sambridge, M., Gallagher, K., Jackson, A., & Rickwood, P. (2006). Trans-dimensional inverse problems, model comparison and the evidence. *Geophysical Journal International*, *167*(2), 528–542.

Scholz. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bulletin of the seismological society of America*, *58*(1),

399–415.

Scholz. (2015). On the stress dependence of the earthquake b value. *Geophysical Research Letters*, *42*(5), 1399–1402.

Schorlemmer, D., Wiemer, S., & Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes. *Nature*, *437*(7058), 539–542.

Shelly, D. R., Ellsworth, W. L., & Hill, D. P. (2016). Fluid-faulting evolution in high definition: Connecting fault structure and frequency-magnitude variations during the 2014 long valley caldera, california, earthquake swarm. *Journal of Geophysical Research: Solid Earth*, *121*(3), 1776–1795.

Shi, H., MENG, L., ZHANG, X., CHANG, Y., YANG, Z., XIE, W., … HAN, P. (2018). Decrease in b value prior to the wenchuan earthquake (m s 8.0). *Chinese Journal of Geophysics*, *61*(5), 1874–1882.

Shi, Y., & Bolt, B. A. (1982). The standard error of the magnitude-frequency b value. *Bulletin of the Seismological Society of America*, *72*(5), 1677–1687.

Spassiani, I., Taroni, M., Murru, M., & Falcone, G. (2023, March). Real time Gutenberg–Richter *b* -value estimation for an ongoing seismic sequence: an application to the 2022 marche offshore earthquake sequence (*M* L 5.7 central Italy). *Geophysical Journal International*, *234*(2), 1326–1331. Retrieved 2024-04-04, from https://academic.oup.com/gji/article/234/2/1326/7086114 doi: 10.1093/gji/ggad134

Taroni, M., & Akinci, A. (2020, November). Good practices in PSHA: declustering, b-value estimation, foreshocks and aftershocks inclusion; a case study in Italy. *Geophysical Journal International*, *224*(2), 1174–1187. Retrieved 2024-04-04, from https://academic.oup.com/gji/article/224/2/1174/5911581 doi: 10.1093/gji/ggaa462

Utsu, T. (1966). A statistical significance test of the difference in b-value between two earthquake groups. *Journal of Physics of the Earth*, *14*(2), 37–40.

Utsu, T. (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 (1). *Journal of the Faculty of Science, Hokkaido University. Series 7, Geophysics*, *3*(5), 379–441.

Van der Elst, N. J. (2021). B-positive: A robust estimator of aftershock magnitude distribution in transiently incomplete catalogs. *Journal of Geophysical*

Research: Solid Earth, *126*(2), e2020JB021027.

Weichert, D. H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes. *Bulletin of the Seismological Society of America*, *70*(4), 1337–1346.

Wetzler, N., Lay, T., & Brodsky, E. E. (2023). Global characteristics of observable foreshocks for large earthquakes. *Seismological Research Letters*, *94*(5), 2313–2325.

Wiemer, S., & Wyss, M. (1997). Mapping the frequency-magnitude distribution in asperities: An improved technique to calculate recurrence times? *Journal of Geophysical Research: Solid Earth*, *102*(B7), 15115–15128.

Woessner, J., & Wiemer, S. (2005). Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bulletin of the Seismological Society of America*, *95*(2), 684–698.

Yin, F., & Jiang, C. (2023, October). Enhanced b -value time series calculation method using data-driven approach. *Geophysical Journal International*, ggad419. Retrieved 2024-04-04, from `https://academic.oup.com/gji/advance-article/doi/10.1093/gji/ggad419/7328941` doi: 10.1093/gji/ggad419