

# JGR Solid Earth

## RESEARCH ARTICLE

10.1029/2024JB029973

### Key Points:

- We propose to use all the data available from earthquake catalogs and invert conjointly for  $b_{value}$  and the earthquake detectability
- We retrieve the full posterior density function of  $b_{value}$  with a proper evaluation of its uncertainties than classical approaches
- b-Bayesian performs a transdimensional inversion to recover the temporal changes of the  $b_{value}$  and detectability

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

M. Laporte,  
[marine.laporte@univ-lyon1.fr](mailto:marine.laporte@univ-lyon1.fr)

### Citation:

Laporte, M., Durand, S., Bodin, T., Gardonio, B., & Marsan, D. (2025). b-bayesian: The full Probabilistic estimate of b-value Temporal variations for non-truncated catalogs. *Journal of Geophysical Research: Solid Earth*, 130, e2024JB029973. <https://doi.org/10.1029/2024JB029973>

Received 20 JUL 2024

Accepted 3 MAR 2025

### Author Contributions:

**Conceptualization:** S. Durand, T. Bodin, B. Gardonio

**Formal analysis:** M. Laporte, S. Durand, T. Bodin, B. Gardonio, D. Marsan

**Funding acquisition:** S. Durand, B. Gardonio

**Investigation:** M. Laporte, S. Durand, T. Bodin, B. Gardonio

**Methodology:** M. Laporte, S. Durand, T. Bodin, B. Gardonio, D. Marsan

**Software:** M. Laporte

**Supervision:** S. Durand, T. Bodin, B. Gardonio, D. Marsan




**Validation:** M. Laporte, S. Durand, T. Bodin, B. Gardonio, D. Marsan

**Visualization:** M. Laporte

© 2025. The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

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

M. Laporte<sup>1</sup> , S. Durand<sup>1</sup>, T. Bodin<sup>1,2</sup> , B. Gardonio<sup>1,3</sup>, and D. Marsan<sup>3</sup> 

<sup>1</sup>LGL-TPE, Université Claude Bernard, Lyon, France, <sup>2</sup>Instituto de Ciencias del Mar (ICM) - CSIC, Barcelona, Spain, <sup>3</sup>Université de Savoie Mont Blanc, Chambéry, France

**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 modeled 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 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 around 0.8, albeit with larger uncertainties. This  $b_{value}$  lower than 1 is expected in such a region with large interseismic strain accumulation. This confirms that 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.

**Plain Language Summary** Above a magnitude of completeness ( $M_c$ ), all earthquakes are considered to have been detected, and the frequency distribution of earthquakes per unit of magnitudes follows an exponential law, known as the Gutenberg-Richter law, whose exponent is the so-called  $b_{value}$ . The study of the spatio-temporal variations of  $b_{value}$  has attracted much attention in recent years. In particular, it is thought to vary with increasing stress state and has recently been proposed to discriminate foreshock sequences. Therefore, a robust estimation of its variations and uncertainties is needed. In this paper, we present a new method, called b-Bayesian, to properly estimate the temporal variations of  $b_{value}$  within a probabilistic framework. The b-Bayesian method has the following advantages: (a) it uses all available earthquake data from an earthquake catalog without the need to truncate them above the completeness magnitude, (b) it jointly considers the temporal variations of detectability and  $b_{value}$ , and (c) it operates within a Bayesian framework. As a first application, we chose an earthquake catalog from far-western Nepal with significant variations in detectability. We compared the full probability distributions obtained with our new approach with maximum likelihood estimates from classical approaches.

## 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 catalogs, the  $b_{value}$  is typically close to 1, but has been shown to vary in both space and time when focusing on earthquake catalogs for specific seismogenic regions or time periods (e.g., Ogata & Katsura, 2006; Wiemer & Wyss, 1997). With the rapid growth of seismological instruments and recording capabilities, there is a need for advanced statistical

**Writing – original draft:** M. Laporte  
**Writing – review & editing:** M. Laporte,  
S. Durand, T. Bodin, B. Gardonio,  
D. Marsan

methods to analyze earthquake catalogs. Here, we analyze 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., Beauval & Scotti, 2004; Keller et al., 2014; Taroni & Akinci, 2020). For accurate hazard assessment,  $b_{value}$  biases due to the incompleteness of earthquake catalogs need to be addressed (e.g., Beauval & Scotti, 2004; Dutoy, 2020; Plourde, 2023; Weichert, 1980) 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., Carter & Berg, 1981; Herrmann et al., 2022; Mogi, 1962; Scholz, 1968). Based on observations from laboratory earthquake simulations, which are commonly used as analogs to study natural earthquake behavior, it has been proposed that  $b_{value}$  is inversely related to the normal and shear stress applied to the fault (Scholz, 1968). At the scale of the seismic cycle, reproduced in stick-slip experiments with controlled stress and friction properties,  $b_{value}$  has been observed to decrease linearly with stress build-up and to increase abruptly with the release of stress drops during earthquake rupture (Avlonitis & Papadopoulos, 2014; Bolton et al., 2020; Goebel et al., 2017; Rivière et al., 2018). Extending this observation to real earthquake systems is not straightforward because real earthquake catalogs contain additional uncertainties, and the estimation of the actual state of the stress field is another inverse problem.

However,  $b_{value}$  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 earthquake depths or within different stress regimes (Morales-Yáñez et al., 2022; Mori & Abercrombie, 1997; Petrucci et al., 2019; Scholz, 2015; Schorlemmer et al., 2005). Low  $b_{value}$  (<0.8), associated with a larger number of large magnitude 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; Kwiątek 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., Geffers et al., 2022; Godano et al., 2024; Lombardi, 2021; Spassiani et al., 2023; Yin & Jiang, 2023).

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

$$\beta = \frac{1}{\bar{m} - M_c} \quad (2)$$

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

Unnoticed changes in completeness over time are the main source of bias when studying  $b_{value}$  temporal variations (e.g., Godano et al., 2023; Helmstetter et al., 2006; Lombardi, 2021; Mignan & Woessner, 2012; Plourde, 2023; Woessner & Wiemer, 2005). Two main sources of incompleteness are usually identified (e.g., Lippiello & Petrillo, 2024): (a) the background incompleteness coming from momentary changes in the detectability of the seismic network, and (b) 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., Hainzl & Fischer, 2002; Helmstetter et al., 2006). 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.



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

with  $\overline{m'}^+$  the mean of positive consecutive 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 obtained with the original b-positive approach is computed from less than half of the available data. Recently, the novel b-more positive approach has been proposed to overcome some of the shortcomings of the b-positive approach by including not only consecutive pairs of earthquake magnitudes (Lippiello & Petrillo, 2024). In addition,  $1\sigma$  uncertainties of the b-positive approach can be derived through a theoretical formulation (Tinti & Gasperini, 2024).

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 catalog, 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 et al., 2012; Bodin & Sambridge, 2009). 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 data set. 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 in far western Nepal. In this region, where the last great earthquake has been dated to 1,505, the strain from the India-Asia collision has been building up for more than 500 years. According to the recent studies relating  $b_{value}$  and stress, a low and potentially decreasing  $b_{value}$  could be expected in such a region undergoing a long-term strain accumulation. 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 2 years of microseismicity.

## 2. Method: A Bayesian Framework

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

$$\mathbf{d} = [m_1, m_2, \dots, m_N] \quad (4)$$

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

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

which is true only for  $m_i > M_c$ . Then, assuming that the magnitudes of the seismic events are independent, we can write the probability of observing the entire earthquake data set  $\mathbf{d}$  with  $m_i \geq M_c$ ,  $p(\mathbf{d}|\beta)$ , as:

$$p(\mathbf{d}|\beta) = \prod_{i=1}^N p(m_i|\beta) = \beta^N e^{-\beta N(\bar{m}_i - M_c)} \quad (6)$$

where  $\bar{m}_i$  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  $\beta$  that maximizes Equation 6 is the maximum likelihood solution given by Aki's formula in Equation 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., Mignan & Woessner, 2012; Ogata & Katsura, 1993; Ringdal, 1975; Woessner & Wiemer, 2005) or to correct the data set for its temporal variations (e.g., Cao & Gao, 2002; Chan et al., 2012; Helmstetter et al., 2006).

However, defining such a completeness magnitude always implies to ignore a significant portion of a data set that may contain valuable information about the statistics of seismicity. Here instead, we propose to analyze the entire data set by modeling the entire frequency-magnitude distribution of earthquakes including magnitudes below the completeness  $M_c$ . To do so, the Gutenberg-Richter law is modulated by a detection law  $q(m)$  such that now:

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

where  $M_{\min}$  is the minimum earthquake magnitude in the catalog, a cut-off threshold below which no earthquakes have been reported. This minimum magnitude allows to restrict the distribution to the range of observed magnitudes and to include negative magnitudes. There,  $q(m)$  defines the probability density of detecting an event and,  $K$  a constant to insure that the probability distribution integrates to one:

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

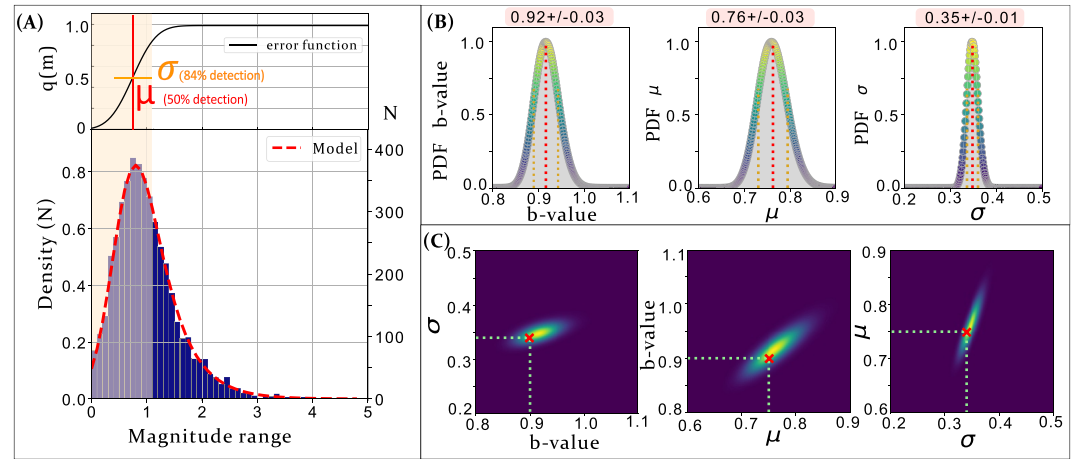
In this way, the probability of observing an event is the product of the probability of occurrence (given by the Gutenberg Richter law) and the probability of detection (given by the detection law  $q(m)$  that varies from 0, no detection, to 1, 100% detection). The error function has been proposed in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

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

where  $\mu$  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:

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



**Figure 1.** (a) [Top] Detection function  $q(m)$  associated with the synthetic data set [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) Marginal 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 red crosses.

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

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

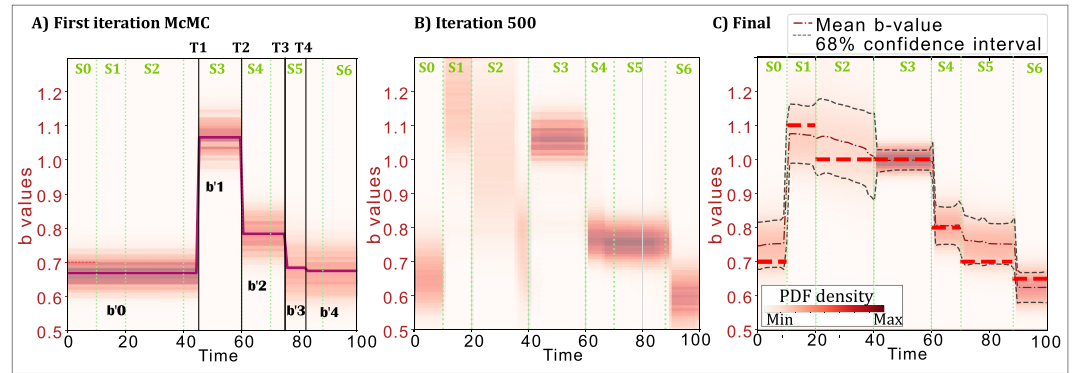
Now assuming that magnitudes are independent, the probability of observing a full data set  $\mathbf{d}$  is:

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

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  $p(\mathbf{d}|\omega)$  (Equation 12).

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

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  $\mu$  and  $\sigma$  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  $\mu$  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 large mainshock/aftershock sequence, this range should also be increased. The prior distribution on  $\sigma$  can be set between 0.01 and 0.5 and does not need to be adjusted depending on the context.



**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  $d_j (j = 1 \dots 5)$ . The bold line is the mean posterior density function of  $\beta_j$  over time. Here, the synthetic earthquake data set 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 data set which are represented in bold dashed horizontal lines.

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

As an example, we construct a synthetic data set of 4,460 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  $\mu = 0.75$  and  $\sigma = 0.34$  (see Figure 1a). We then estimate our set of parameters  $\omega = [\beta, \mu, \sigma]$  from the catalog, 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 (PDF) can be projected onto each parameter (Figure 1b) to derive 1D and 2D marginal distributions (Figure 1c). For example, the marginal distribution for  $\beta$  is simply obtained by integrating the posterior over  $\mu$ , and  $\sigma$ :

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

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  $1\sigma$  confidence interval.

Compared to optimization approaches where only the best fitting (i.e., maximum likelihood)  $\omega$  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 data set 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.

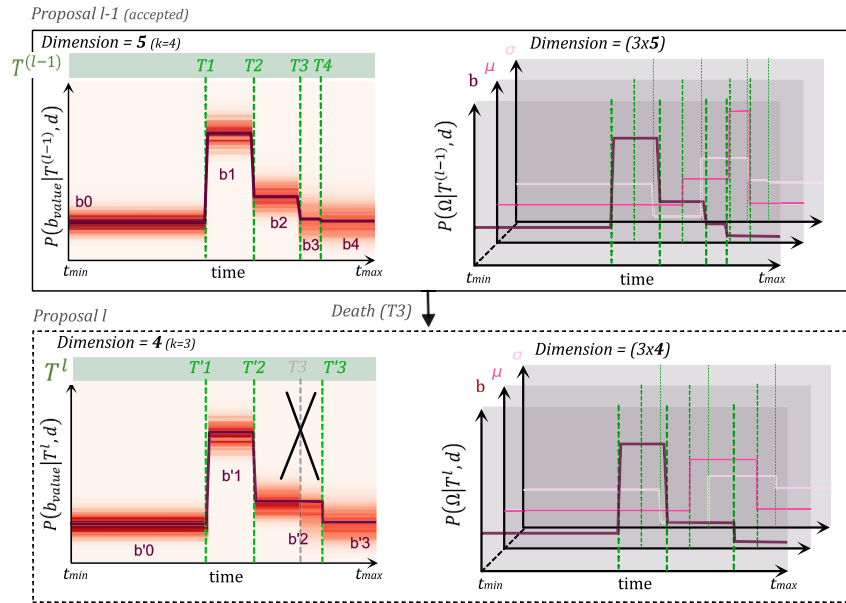
## 2.2. Temporal Variations of b-Value

### 2.2.1. A Transdimensional Parametrization

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

$$\mathbf{T} = [T_1, T_2, \dots, T_k] \quad (15)$$

where,  $k$  is the number of temporal discontinuities and  $T_j (j = 1 \dots k)$  the times at which the frequency-magnitude distribution changes. The unknown models vectors of the time varying frequency-magnitude distribution will be denoted:



**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  $p(\beta|\mathbf{T}_{l-1}, \mathbf{d})$ ,  $P(\mu|\mathbf{T}_{l-1}, \mathbf{d})$  and  $P(\sigma|\mathbf{T}_{l-1}, \mathbf{d})$ . (bottom) New proposal model  $\mathbf{T}_l$  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 ( $k = 3$ ).

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

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  $p(\mathbf{T}|\mathbf{d})$ :

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

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

### 2.2.2. The Conditional Posterior $p(\mathbf{\Omega}|\mathbf{d}, \mathbf{T})$

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:

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

Since all magnitudes are independent, the likelihood is the product of likelihoods for every sub-dataset  $\mathbf{d}_j$  given by the temporal model  $\mathbf{T}$ :



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

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:

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

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 12.

The prior distribution for  $\mathbf{\Omega}$  given a fixed temporal model  $\mathbf{T}$ ,  $p(\mathbf{\Omega}|\mathbf{T})$  from Equation 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 (Equation 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  $\mathbf{T} = [T_1, T_2, \dots, T_k]$  given the full data set of observed magnitudes. It can be obtained by integrating the full posterior  $p(\mathbf{\Omega}, \mathbf{T}|\mathbf{d})$  over the parameters  $\mathbf{\Omega} = [\omega_1, \omega_2, \dots, \omega_{k+1}]$ :

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

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

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

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 22, we get:

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

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 (Equation 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.

The integral of Equation 23 can be decomposed as a product of integrals between each temporal subsets. Thus, the full posterior can be expressed as:

$$p(\mathbf{T}|\mathbf{d}) \propto p(\mathbf{T}) \times \int_{\omega_1} \dots \int_{\omega_{k+1}} p(\mathbf{d}, \omega_1|\mathbf{T}) p(\omega_1|\mathbf{T}) \dots p(\mathbf{d}, \omega_{k+1}|\mathbf{T}) p(\omega_{k+1}|\mathbf{T}) d\omega_1 \dots d\omega_{k+1} \quad (24)$$

Considering that the sub-datasets  $\mathbf{d}_j$  of a time model  $\mathbf{T}$  are independent, and since data set  $\mathbf{d}_j$  only depends on parameters  $\omega_j$ , then Equation 25 can be simplified as:

$$p(\mathbf{T}|\mathbf{d}) \propto p(\mathbf{T}) \times \prod_{j=1}^{k+1} \left( \int_{\omega_j} p(\mathbf{d}_j|\omega_j, \mathbf{T}) p(\omega_j|\mathbf{T}) d\omega_j \right) \quad (25)$$

where,  $p(\mathbf{d}_j|\omega_j, \mathbf{T})$  is the likelihood of observing the subset  $\mathbf{d}_j$  between  $[T_{j-1}, T_j]$  according to the local model  $\omega_j$  and can be estimated using Equation 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  $p(x)$ :

$$\int p(x)f(x)dx \approx \frac{1}{N} \times \sum_{i=1}^N f(x_i) \quad (26)$$

Applied to Equation 26, we have:

$$p(\mathbf{T}|\mathbf{d}) \propto p(\mathbf{T}) \times \left( \frac{1}{N_w} \right)^{(k+1)} \prod_{j=1}^{k+1} \left( \sum_{i=1}^{N_w} p(d_j|\omega_{j(i)}, \mathbf{T}) \right) \quad (27)$$

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

From Equation 27 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  $N_w$  realizations 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  $p(d_j|\omega_{j(i)}, \mathbf{T})$  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 27 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 transdimensional function and cannot be explored using a standard Metropolis algorithm (Hastings, 1970; Metropolis et al., 1953). One of the most popular technique for exploring a transdimensional posterior is the rj-McMC method (e.g., Green, 1995; Sambridge et al., 2006, 2013) and more specifically the birth-death Markov-chain Monte Carlo (McMC) algorithm (e.g., Geyer & Møller, 1994). The rj-McMC algorithm, used in many geophysical inverse problems (e.g., Bodin et al., 2012; Gallagher et al., 2009; Gallagher et al., 2011), 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 toward a stable value and when the rate of accepted models falls in the range of 20%–40%. Details about the algorithm are given in Supporting Information S1. We also refer the reader to (Bodin et al., 2012) for further details.

### 2.2.5. Appraising the Full Posterior Distribution $p(\Omega, \mathbf{T}|\mathbf{d})$

As a reminder, the solution to our inverse problem is the full posterior solution  $p(\Omega, \mathbf{T}|\mathbf{d})$  that describes the temporal changes of  $\beta$ ,  $\mu$ , and  $\sigma$ . As shown in Equation 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  $p(\Omega|\mathbf{T}, \mathbf{d})$  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 transdimensional 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  $p(\mathbf{T}|\mathbf{d})$  (Equation 27). 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  $p(\Omega|\mathbf{d}, \mathbf{T}^l)$ : the probability distribution of  $\beta$ ,  $\mu$ , and  $\sigma$  for the given model  $\mathbf{T}^l$  (Equation 20).

Thus, at each iteration, the code proposes a new model depicting constant periods separated by discrete changes in parameters, which is either accepted or rejected based on the Metropolis-Hastings criteria. If the proposed model is accepted, it is added to the ensemble of previously accepted models. As the process continues over many iterations, the cumulative effect of all accepted stair-step changes is built up.

At the completion, the full distribution for  $\beta$ ,  $\mu$ , and  $\sigma$  can therefore be obtained by summing the all the distributions  $p(\Omega|\mathbf{d}, \mathbf{T}^l)$  for all the sampled models  $\mathbf{T}^l \in T^{(c)}$  (Figures 2b and 2c). Although each individual model shows discrete steps, the time of changes is variable, and the averaging process over many iterations results in a smoothing effect. The ensemble solution of accepted models creates a composite distribution that more accurately reflects the information contained in the underlying data. This sampling process results in a final posterior density where the average varies smoothly. In this way, the mean and the standard deviation for our three parameters can be obtained as a smooth function of time (see Figure 2c).

## 2.3. Synthetic Test

### 2.3.1. Generated Data

We simulate a synthetic earthquake catalog of 5,683 independent events following frequency-magnitude distributions with some temporal variations in  $b_{\text{value}}$  and detectability. In this section, for demonstration purposes, we only consider a data set with abrupt and discontinuous changes for the three parameters of the frequency-magnitude distribution which is more adapted to our “step-changes” parametrization. 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 catalog 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  $b_{\text{value}}$  specified in Table 1.
- Earthquake occurrence is generated according to a basic epidemic-type aftershock sequence (ETAS) (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 modeled by the Omori power law (Omori, 1894). The ETAS parametrization does not vary temporally. In particular, to characterize the temporal occurrence of aftershocks, we keep a constant  $p_{\text{etas}}$  value of 1.1 and an  $\alpha_{\text{etas}}$  value of 1.5. For now, we do not generate the detectability variations coming from the short-term incompleteness (e.g., Helmstetter et al., 2006; Ogata & Katsura, 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  $\mu$  and  $\sigma$  (see Equation 9). Each pair of  $\mu$  and  $\sigma$  for each of the seven sub-datasets are specified in Table 1.

The data set is made to test the capabilities of our algorithm and approach. To that aim, there are some periods where  $b_{\text{value}}$  remains constant while the detectability varies (S2 and S3) (Figure 4). We also include small

**Table 1**

Set of Chosen Values for  $b_{value}$ ,  $\mu$  and  $\sigma$  for Each of the 7 Data Subsets.  $N_{events}$  Corresponds to the Number of Earthquakes Detected of Each Sub-Set (After the Detection Thinning)

Subset	S0	S1	S2	S3	S4	S5	S6
$N_{events}$	514	647	261	2,538	727	672	324
$b_{value}$	0.70	1.10	1.00	1.00	0.80	0.70	0.65
$\mu$	0.80	0.80	1.50	0.50	0.75	0.90	0.90
$\sigma$	0.35	0.30	0.20	0.15	0.30	0.40	0.15

detectability or  $b_{value}$  changes (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.

### 2.3.2. Results and Comparison

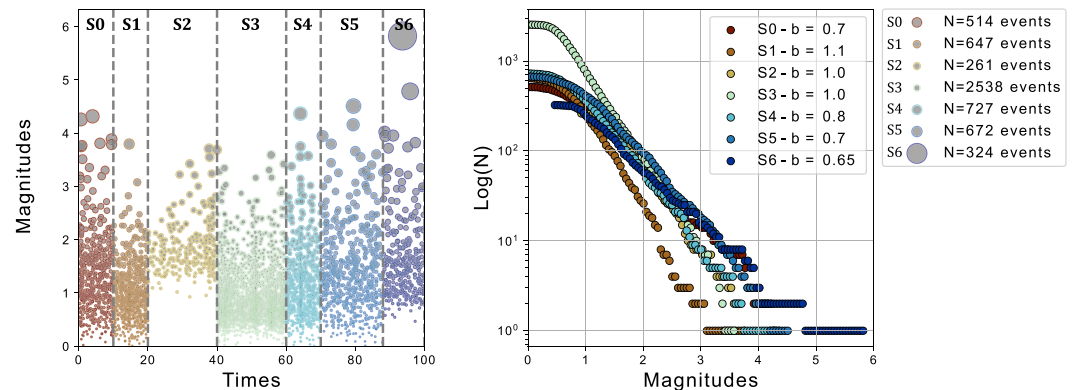
Let us start by applying classical approaches on this synthetic data set. Ignoring temporal variations of  $b_{value}$  and considering a constant completeness magnitude of 1.8, the frequentist approach (Aki, 1965) gives the maximum likelihood estimate over the complete catalog of  $b_{value} = 0.81 \pm 0.03$ . This illustrates how the uncertainties can be under-

estimated 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  $(i = 0, \dots, 6)$ , the true  $b_{value}$  can be recovered by the maximum likelihood estimate within its uncertainties. However, for applications to real earthquake catalogs, 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.

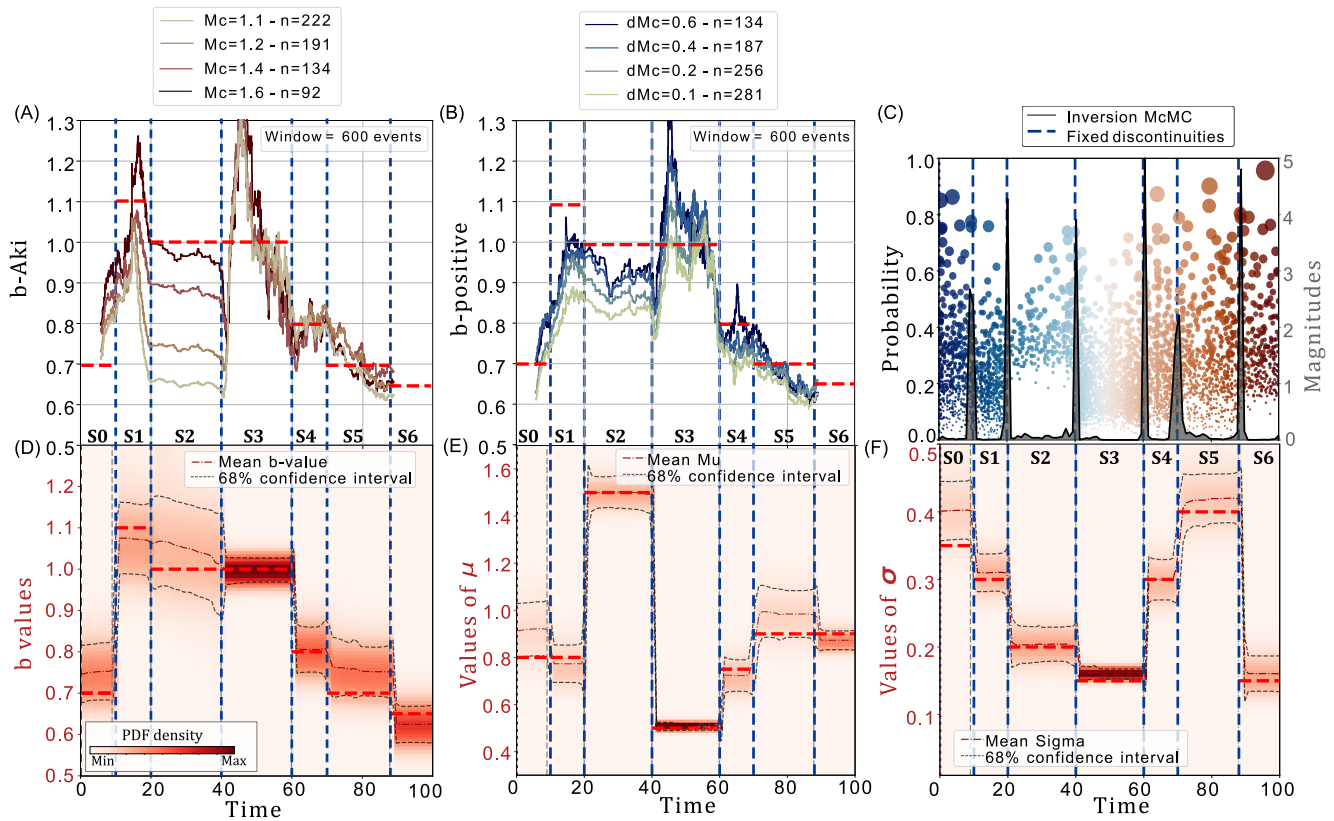
Using a moving-window of 600 events and a constant completeness magnitude, Figure 5a shows that depending on the temporal sub-dataset the maximum likelihood estimate tends to under-estimate or over-estimate the  $b_{value}$  depending on the choice of  $M_c$  and the true detectability. In particular for (S2) with a low detectability, the choice of the completeness magnitude has a major influence on the results. This large variability of solutions can be partly corrected by the use of the b-positive approach (Van der Elst, 2021) (Figure 5b). 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 data sets, 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 data set.

We applied our method to the synthetic data set and conducted 50 parallel rj-McMCMC explorations with different initializations for  $\mathbf{T}$  in order to efficiently explore the model space. We present the results derived from the final



**Figure 4.** (a) Temporal distribution of magnitudes over time for the 7 data subsets. Each subset is represented with a different color. Gray 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  $b_{value}$ .



**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  $\mu$  over time (f) Marginal density distribution of  $\sigma$  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.

stack of posterior densities coming from the 50 runs, each encompassing 5,000 proposed models of temporal discontinuities (Figures 5c–5f). In total, 250,000 temporal models were proposed out of which approximately 41,800 were accepted upon completion. The six temporal discontinuities are well retrieved and clearly identified (Figure 5c). Analysis of the number of discontinuities at each step indicates that the algorithm converged toward the correct value in less than 2,000 iterations, even though all initialisations began with random values of  $k$  ranging from 4 to 12. A burn-in period of 1,000 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 (Figures 5d–5f). 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 (Figures 5d–5f). 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  $\pm 0.1$ . As this period involves the fewest events, the relative probability is the lowest. This confidence intervals narrows to  $\pm 0.02$  for the period S3 containing 2,538 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.

Here we have shown the efficiency of the b-Bayesian code to effectively resolve stair-step variations from a synthetic catalog. In the case of a continuous evolution of  $b_{value}$  or detectability as in the STAI model (Helmstetter et al., 2006), the b-Bayesian approach will model the continuous evolution by creating and moving discontinuities in the continuous period to represent all the stair-steps that fit the evolution. In this way, the final posterior density function will be a composite of all these accepted stair-step models, resulting in a smooth temporal evolution of the marginal density. An example of such continuous evolution retrieved by the b-Bayesian code can be found in Supporting Information S1 and in the next section on a real earthquake data set. For a more accurate description of a continuous evolution, one perspective could be to change the current parameterization by modeling linear trends for  $\beta$ ,  $\mu$  or  $\sigma$ , between each step. Note also that a more physical parameterization could be used based on the STAI model. Such a parameterization would increase the computational time without any evidence that the results would be significantly more accurate.

### 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

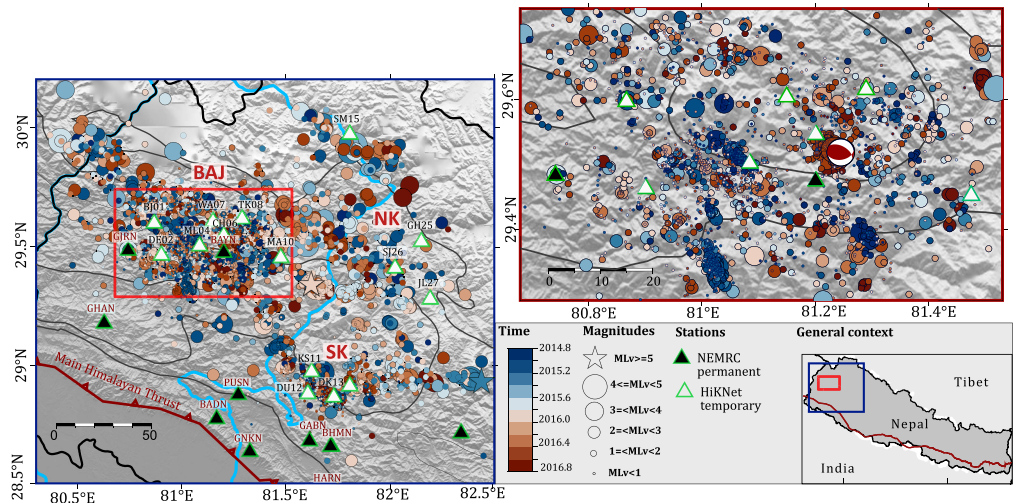
Our first application of the transdimensional approach to investigate  $b_{value}$  variations focus on the temporal evolution of seismicity in a very seismically active region of Nepal. We use the earthquake data collected during two years by the Himalayan Karnali Network (HiKNet), the first dense seismological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

Between December 2014 and September 2016, the temporary experiment recorded almost 4,500 earthquakes in this region. The seismicity is structured into three seismic belts: one large belt in the westernmost part (BAJ, Bajhang region) and two separate belts in the east (SK and NK, South and North Karnali sectors) (Figure 6). Each belt contains several seismic clusters of different size and spatio-temporal behavior. 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.

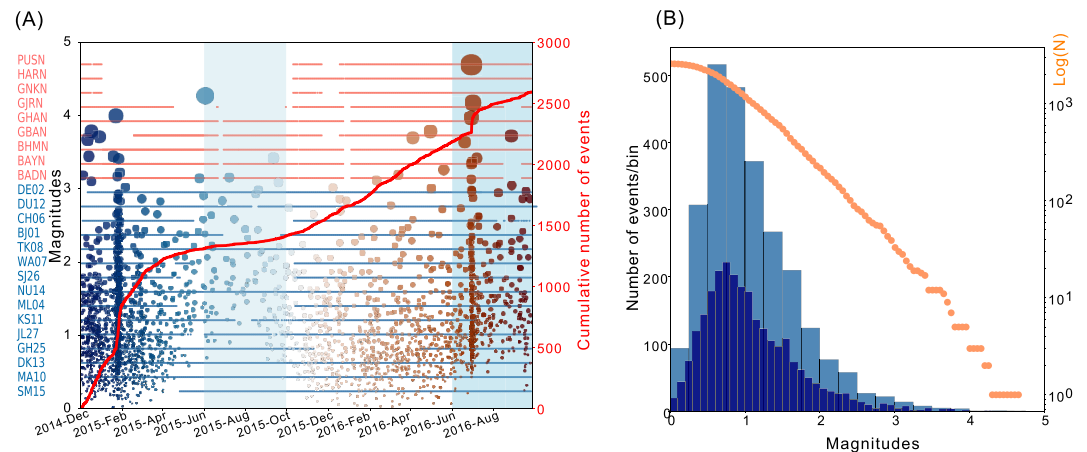
In Nepal, the monsoon season typically spans from early June to September, peaking in July and August. During this period, some small seasonal strain and stress variations have been evidenced as well as a correlation between such stress rate and the seismicity rate (Bettinelli et al., 2008; Bollinger et al., 2007). However, during the same



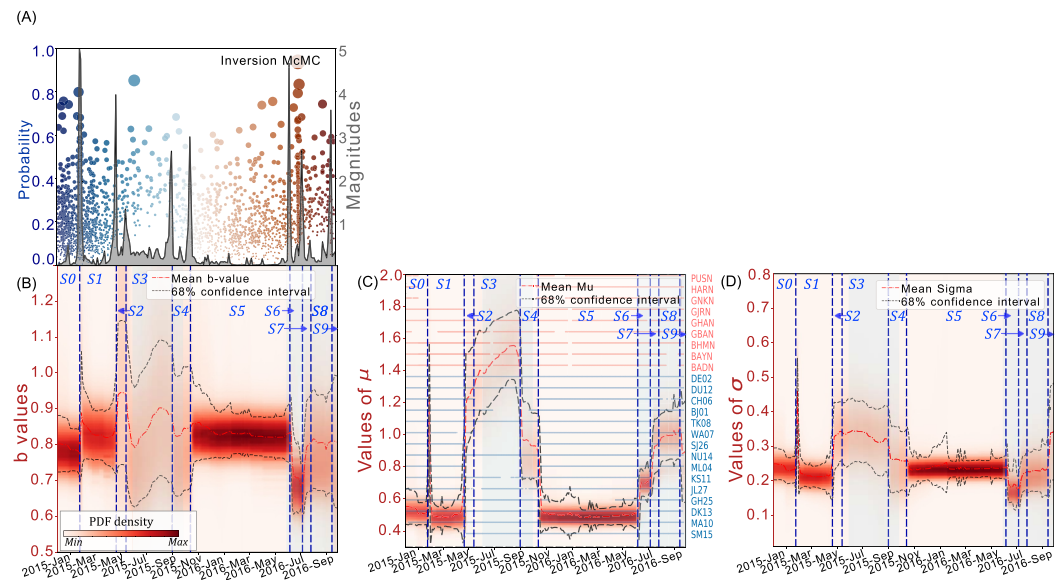
**Figure 6.** Far-western Nepal seismicity recorded during 2 years by the temporary seismological experiment of the Himalayan-Karnali network (HiKNet) (white triangles) and the permanent seismological network of the National Earthquake Monitoring Research Center (black triangles). We use the earthquake 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 2,593 events: the Bajhang region (BAJ).

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 7a, along with the availability of the 24 instruments throughout the experiment.

Seasonal variations in the  $b_{value}$  could represent such modulation of seismicity due to seasonal stress perturbations. Given the challenges posed by temporal variations in detectability due to station losses and monsoon periods, classical approaches cannot adequately address this issue. We chose this data set to study the potential temporal variations in  $b_{value}$  using the transdimensional approach, which accounts for uncertainties arising from detectability. To ensure spatial homogeneity of  $b_{value}$ , we focused on a geographical region of 94 km by 57 km, corresponding to the seismicity of the western belt, which is the most instrumented. 2,593 earthquakes were recorded within this specific region. The cumulative number of earthquakes reveals distinct periods of seismic activity (see Figure 7a). Using the frequentist approach of Aki for a completeness magnitude of 1.5 on the full



**Figure 7.** (a) Temporal distribution of magnitudes during 2 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.



**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  $\mu$  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  $\sigma$  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).

time period, we get a  $b_{value}$  of  $0.82 \pm 0.08$  (Figure 7b) 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 toward 9 temporal discontinuities (defined as being over a probability of 15%) after 5,000 iterations (Figure 8a). 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 data set. We interpret and discuss each of them with respect to the stacked marginal densities of probability for  $b_{value}$ ,  $\mu$  and  $\sigma$  (Figure 8).

- S0 (15/10/2014–27/01/2015) corresponds to the first three months of the HiKNet temporary seismological experiment. During that time the  $b_{value}$  and the two detectability parameters  $\mu$  and  $\sigma$  are constant and equal to  $0.79 \pm 0.06$ ,  $0.55 \pm 0.07$  and  $0.24 \pm 0.04$ , respectively.
- S1 (27/01/2015–19/04/2015) The January 2015 discontinuity follows the occurrence of a seismic crisis after a  $M_{L_v}$  4.0 earthquake on 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 occurring in 9 days (Hoste-Colomer et al., 2018; Laporte et al., 2021). For this seismic crisis, characterized as a large seismic swarm by Hoste-Colomer et al. (2018), the  $b_{value}$  and the two detectability parameters increase suddenly and then return within the uncertainties of their previous value:  $0.82 \pm 0.07$ ,  $0.49 \pm 0.07$  and  $0.21 \pm 0.05$ .
- S2 (19/04/2015–11/05/2015) April 2015 corresponds to the loss of three stations (ML04, SJ26 and GJRN) (Figure 8c), two of which were located in the considered region (Figure 6). The confidence interval for each

parameter becomes wider, corresponding to fewer events detected. Only  $\mu$  and  $\sigma$  show a clear discontinuous change. Specifically, in April 2015,  $\mu$  increases from  $0.49 \pm 0.07$  to  $1.24 \pm 0.27$  and  $\sigma$  from  $0.21 \pm 0.05$  to  $0.33 \pm 0.09$ . During this month the parameter  $\mu$  is not constant and seems to increase. The mean of the marginal density distribution of  $b_{\text{value}}$  increases from 0.82 to 0.94 with twice the uncertainties of  $\pm 0.2$ . Given the large uncertainties, such an increase in  $b_{\text{value}}$  is not statistically significant...

- S3 (11/05/2015–01/09/2015) May 2015 corresponds to the loss of two stations in the center of the region (WA07 and CH06) (Figure 8). In the region, only 5 out of 9 stations are available during this period. Between June and September the monsoon occurs and the detectability decreases progressively, as shown by the progressive increase of the  $\mu$  parameter from  $1.30 \pm 0.25$  to  $1.55 \pm 0.23$ . The  $\sigma$  parameter is not affected as much as  $\mu$  and remains approximately constant. The mean  $b_{\text{value}}$  stays within the confidence interval of the previous periods,  $0.84 \pm 0.17$ . Similar to period S2, we cannot statistically resolve the  $b_{\text{value}}$  variations during the monsoon period due to the large uncertainties resulting from poor detectability.
- S4 (01/09/2015–12/10/2015) In September 2015 the monsoon stopped but 6 stations are still unavailable. At that time, we notice a brutal improvement in the detectability:  $\mu$  decreases from  $1.55 \pm 0.23$  to  $0.94 \pm 0.21$  and  $\sigma$  from  $0.33 \pm 0.09$  to  $0.26 \pm 0.09$ . However, the detectability has not returned to its pre-monsoon value and the confidence interval of the  $b_{\text{value}}$  remains as large  $0.84 \pm 0.17$ .
- S5 (12/10/2015–04/06/2016) In October 2015, a field mission recovers the 6 missing stations. Both the  $b_{\text{value}}$  and the detectability parameters return to the pre-monsoon periods S1 and S2,  $0.50 \pm (0.04\text{min}/0.13\text{max})$  and  $0.26 \pm (0.02\text{min}/0.09\text{max})$ , respectively. This long period shows a stable frequency-magnitude distribution of seismicity for 8 months until early June 2016. The  $b_{\text{value}}$  remains constant around  $0.82 \pm 0.13$  during this period, with a confidence interval that narrows to  $\pm 0.05$  at the end of December 2015.
- S6 (04/06/2016–01/07/2016) This period corresponds to the first days of the expected monsoon period in 2016,  $\mu$  increases significantly ( $0.70 \pm 0.1$ ) and  $\sigma$  decreases ( $0.18 \pm 0.05$ ). The density probability of  $b_{\text{value}}$  decreases from  $0.82 \pm 0.13$  to  $0.69 \pm 0.08$ , which is the most significant decrease of the 2 years. Meanwhile, this period includes the onset of the second largest seismic crisis ever recorded in this region. On June 29, a  $M_{\text{L}} 4.8$  earthquake occurred and was followed by several aftershocks, including two  $M_{\text{L}} > 4$  earthquakes. The last largest aftershock of  $M_{\text{L}} 4.1$  occurred on July 2 (00h23).
- S7 (01/07/2016–24/07/2016) This period begins during the 29th of June crisis that lasted 12 days. At the same time, a station (KS11) 50 km away from the area of interest interrupts its recording, making it difficult to trace the possible origin of the discontinuity. The  $\mu$  parameter increases sharply ( $0.94 \pm 0.15$ ) and  $\sigma$  slightly increases ( $0.23 \pm 0.05$ ). The  $b_{\text{value}}$  seems to come back to its probably constant value of  $0.81 \pm 0.12$  (S1,S2, S3,S4,S5).
- S8 (24/07/2016–13/09/2016) Another station from the area of interest becomes unavailable (CH06) and the confidence intervals become wider.  $b_{\text{value}}$  and  $\sigma$  are constant during that period and approximately equal to the mean value they had during the 2 years around 0.8 and 0.25 for  $b_{\text{value}}$  and  $\sigma$  respectively. The mean detectability  $\mu$  seems to increase slightly to reach  $1.00 \pm 0.15$ . This increase is supported by the simultaneous widening of the confidence intervals, which could be due to fewer events being detected during the monsoon. Detectability is not as bad as during the 2015 monsoon, as only 2 stations stopped recording.
- S9 (13/09/2016–25/09/2016) The last two weeks of the HiKNet temporary experiment. The  $\mu$  parameter decreases with the end of the monsoon period but  $\sigma$  increases probably because less events are picked. The  $b_{\text{value}}$  remains constant.

Looking back at these results, our transdimensional inversion  $b_{\text{value}}$  and detectability (Figure 8) shows that the potential temporal variations of the  $b_{\text{value}}$  that may arise from the stress-loading accompanying monsoon is not fully resolved due to the high detectability variations during the summers of 2015 and 2016 in far western Nepal. However, during the 2 years of the experiment, and especially outside of the monsoon season, the transdimensional approach identifies a constant  $b_{\text{value}}$  around  $0.8 \pm 0.1$ , except for a short period preceding the monsoon but which also includes a seismic crisis (S6). The origin of this very low  $b_{\text{value}}$  (0.6) cannot be determined based on those results. However, such a low  $b_{\text{value}}$  (0.8) is coherent with previous studies for the same data set (Hoste-Colomer et al., 2018; Laporte et al., 2021) and more globally is coherent with the overall seismotectonic context in a thrust regime with large interseismic accumulation (Schorlemmer et al., 2005). 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  $\mu$  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  $1\sigma$  uncertainty of the  $b_{value}$  is reduced to  $\pm 0.05$ . 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 exist. A much longer data set over several years, with fewer stations shut down during the monsoons, could help provide a more detailed answer on such seasonal modulations. 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., Cao & Gao, 2002; Gulia & Wiemer, 2019; Nanjo et al., 2012; Nuannin et al., 2005). 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 9a). 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 data set only. The uncertainty in  $b_{value}$  is often estimated using theoretical formulas: from Aki (1965), Utsu (1966) for Aki's approach or Tinti and Gasperini (2024) for the b-positive. These uncertainties are computed within the sliding window. Importantly, this uncertainty is consistently estimated independently of the uncertainty associated with the completeness magnitude nor their correlation.

Comparing the outcomes derived from traditional and b-positive 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 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 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  $b_{value}$ . In particular, the apparent decrease of the classical approach and b-positive method in April 2016 seems to be due to the sliding moving-window that artificially creates a decrease when events of July 2016 are beginning to be included.

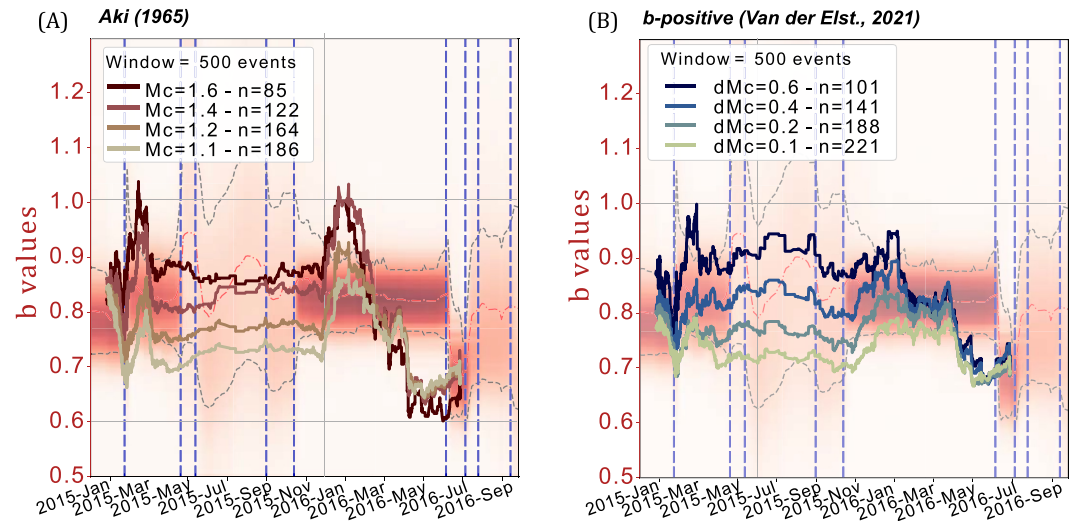
Despite using a moving-window of 500 events, this number is significantly reduced by truncation at  $M_c$  (up to only 17% of magnitudes retained for Aki's approach with  $M_c = 1.6$ ) or by using only positive magnitude differences for the b-positive method (>50%) (Figure 9), while the b-Bayesian approach 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 traditional and the b-positive approaches and the novel b-Bayesian approach presented in this paper. Some of the short-comings of the b-positive approach, notably the quantity of data used, have been recently addressed within the novel b-more-positive approach which is an adaptation of the b-positive (Lippiello & Petrillo, 2024). 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.

### 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., Gulia & Wiemer, 2019; Schorlemmer et al., 2005). This is also supported by experimental studies on the micro-failure





**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.

(e.g., Scholz, 1968; Scholz, 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,  $b_{value}$  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., De Barros et al., 2019; Hainzl & Fischer, 2002; Shelly et al., 2016).

In the present case, for studying Nepalese seismicity, the application of the method on a larger interseismic time scale to see the potential  $b_{value}$  decrease with increasing stress accumulation or on a post-seismic sequences such as the Gorkha earthquake data set could have wider implications for seismic hazard and seismotectonic interpretations. 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. Another perspective of this work would be to change the parameterization of b-Bayesian to invert for a linear trend for  $b_{value}$  instead of constant discrete steps which could be more representative of the continuous temporal evolution of  $b_{value}$ . An even more physical parameterization could be considered by modeling the short-term  $b_{value}$  changes according to a theoretical STAI formula for detectability (Helmstetter et al., 2006).

**Table 2**

Table of Comparison Between the Two Classical Approaches From Aki (1965), Van der Elst (2021) and b-Bayesian

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

\* Arbitrary \*\* for One Parametrization of  $M_c/dM_c$  and Choice of Moving-Window (seconds) \*\*\* No Arbitrary Parametrization (days).

The b-Bayesian method addresses temporal variations in  $b_{value}$  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 proper selection of a spatial subset of earthquakes with a homogeneous  $b_{value}$ , making it challenging to determine the adequacy of the data set. 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.

## Data Availability Statement

The first version of the b-Bayesian software described in this manuscript will be preserved at (Laporte, 2024). The seismicity data set from the far west of Nepal used for the real application in the study is available at (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 2,030 of the French government operated by the National Research Agency (ANR). The research leading to these results has received funding from the Spanish Ministry of Science and Innovation ATR2023-143517/AEI/10.13039/501100011033. We would like to thanks the Editor and Associate Editor and the two anonymous reviewers that contributed to the improvement of this article.

## References

- Ader, T., Avouac, J.-P., Liu-Zeng, J., Lyon-Caen, H., Bollinger, L., Galetzka, J., et al. (2012). Convergence rate across the Nepal Himalaya and interseismic coupling on the main Himalayan thrust: Implications for seismic hazard. *Journal of Geophysical Research*, 117(B4), B04403. <https://doi.org/10.1029/2011JB009071>
- 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(10), 2569–2580. <https://doi.org/10.1007/s00024-014-0799-6>
- 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), 1892. <https://doi.org/10.1029/2003gl017576>
- Beauval, C., & Scotti, O. (2004). Quantifying sensitivities of  $p$ -shallow 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. <https://doi.org/10.1785/012003246>
- Bettinelli, P., Avouac, J.-P., Flouzat, M., Bollinger, L., Ramillien, G., Rajaure, S., & Sapkota, S. (2008). Seasonal variations of seismicity and geodetic strain in the Himalaya induced by surface hydrology. *Earth and Planetary Science Letters*, 266(3–4), 332–344. <https://doi.org/10.1016/j.epsl.2007.11.021>
- Bodin, T., & Sambridge, M. (2009). Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, 178(3), 1411–1436. <https://doi.org/10.1111/j.1365-246x.2009.04226.x>
- Bodin, T., Sambridge, M., Tkalcic, 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), B02301. <https://doi.org/10.1029/2011JB008560>
- Bollinger, L., Perrier, F., Avouac, J.-P., Sapkota, S., Gautam, U., & Tiwari, D. (2007). Seasonal modulation of seismicity in the Himalaya of Nepal. *Geophysical Research Letters*, 34(8), L08304. <https://doi.org/10.1029/2006gl029192>
- Bolton, D. C., Shreedharan, S., Riviere, 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. <https://doi.org/10.1029/2019jb018975>
- Cao, A., & Gao, S. S. (2002). Temporal variation of seismic  $b$ -values beneath northeastern Japan island arc. *Geophysical Research Letters*, 29(9), 48–148–3. <https://doi.org/10.1029/2001gl013775>
- 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. [https://doi.org/10.1016/0040-1951\(81\)90100-1](https://doi.org/10.1016/0040-1951(81)90100-1)
- 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. <https://doi.org/10.1016/j.tecto.2012.02.004>
- Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58(5), 1583–1606. <https://doi.org/10.1785/bssa0580051583>
- Dal Zilio, L., van Dinther, Y., Gerya, T., & Avouac, J. P. (2019). Bimodal seismicity in the Himalaya controlled by fault friction and geometry. *Nature communications*, 10(1), 48.
- Daniel, G., Marsan, D., & Bouchon, M. (2008). Earthquake triggering in southern Iceland following the June 2000  $M_s$  6.6 doublet. *Journal of Geophysical Research*, 113(B5), B05310. <https://doi.org/10.1029/2007jb005107>
- De Barros, L., Baques, M., Godano, M., Helmstetter, A., Deschamps, A., Larroque, C., & Courboux, F. (2019). Fluid-induced swarms and coseismic stress transfer: A dual process highlighted in the aftershock sequence of the 7 April 2014 earthquake ( $M_l$  4.8, Ubaye, France). *Journal of Geophysical Research: Solid Earth*, 124(4), 3918–3932. <https://doi.org/10.1029/2018jb017226>
- 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. <https://doi.org/10.1007/s10518-020-00790-7>
- 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. <https://doi.org/10.1007/s00024-020-02551-8>
- 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. <https://doi.org/10.1016/j.jvolgeores.2009.08.008>
- 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. <https://doi.org/10.1016/j.epsl.2011.09.015>
- 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. <https://doi.org/10.1016/j.marpetgeo.2009.01.003>
- Geffers, G.-M., Main, I. G., & Naylor, M. (2022). Biases in estimating  $b$ -values from small earthquake catalogues: How high are high  $b$ -values? *Geophysical Journal International*, 229(3), 1840–1855. <https://doi.org/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). Evaluating the incompleteness magnitude using an unbiased estimate of the  $b$ -value. *Geophysical Journal International*, 236(2), 994–1001. <https://doi.org/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). <https://doi.org/10.26443/seismica.v3i1.1084>
- 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. <https://doi.org/10.1130/g39147.1>
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4), 711–732. <https://doi.org/10.2307/2337340>
- 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. <https://doi.org/10.1785/0220190356>
- Gulia, L., & Wiemer, S. (2019). Real-time discrimination of earthquake foreshocks and aftershocks. *Nature*, 574(7777), 193–199. <https://doi.org/10.1038/s41586-019-1606-4>
- 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. <https://doi.org/10.1785/0220190307>
- 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*, 107(B12), ESE5-1–ESE5-9. <https://doi.org/10.1029/2002jb001865>
- 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. <https://doi.org/10.1785/0120050067>
- 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. <https://doi.org/10.1038/s41467-022-32755-6>
- Hoste-Colomer, R., Bollinger, L., Lyon-Caen, H., Adhikari, L., Baillard, C., Benoit, A., et al. (2018). Lateral variations of the midcrustal seismicity in western Nepal: Seismotectonic implications. *Earth and Planetary Science Letters*, 504, 115–125. <https://doi.org/10.1016/j.epsl.2018.09.041>
- Hossler, T., Bollinger, L., Sapkota, S. N., Lavé, J., Gupta, R. M., & Kandel, T. P. (2016). Surface ruptures of large Himalayan earthquakes in Western Nepal: Evidence along a reactivated strand of the Main Boundary Thrust. *Earth and Planetary Science Letters*, 434, 187–196.
- Ito, R., & Kaneko, Y. (2023). 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. <https://doi.org/10.1029/2023JB027413>
- Iwata, T. (2013). Estimation of completeness magnitude considering daily variation in earthquake detection capability. *Geophysical Journal International*, 194(3), 1909–1919. <https://doi.org/10.1093/gji/ggt208>
- Keller, M., Pasanisi, A., Marcihac, M., Yalavas, 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. <https://doi.org/10.1002/qre.1735>
- Kwiatek, G., Martínez-Garzón, P., Becker, D., Dresen, G., Cotton, F., Beroza, G. C., et al. (2023). Months-long seismicity transients preceding the 2023 Mw 7.8 Kahramanmaraş earthquake, Türkiye. *Nature Communications*, 14(1), 7534.
- Laporte, M. (2024).  $b$ -Bayesian-2024: The fortran code for probabilistic estimate of  $b$ -value temporal variations for non-truncated catalog [software]. Retrieved from. <https://doi.org/10.5281/zenodo.14803788>
- Laporte, M., Bollinger, L., Lyon-Caen, H., Hoste-Colomer, R., Duverger, C., Letort, J., et al. (2021). Seismicity in far western Nepal reveals flats and ramps along the Main Himalayan Thrust [Dataset]. *Geophysical Journal International*, 226(3), 1747–1763. <https://doi.org/10.1093/gji/ggab159>
- Li, Y., & Chen, X. (2021). Variations in apparent stress and  $b$  value preceding the 2010 Mw8.8 Bio-Bio, Chile earthquake. *Pure and Applied Geophysics*, 178(12), 1–17. <https://doi.org/10.1007/s00024-020-02637-3>
- 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. <https://doi.org/10.1029/2023jb027849>
- Lombardi, A. M. (2021). 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. <https://doi.org/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. <https://doi.org/10.1063/j.1.1699114>
- 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. *Bulletin of the Earthquake Research Institute, University of 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. <https://doi.org/10.1038/s41598-022-25338-4>
- Mori, J., & Abercrombie, R. E. (1997). Depth dependence of earthquake frequency-magnitude distributions in California: Implications for rupture initiation. *Journal of Geophysical Research*, 102(B7), 15081–15090. <https://doi.org/10.1029/97jb01356>
- Nanjo, K., Hirata, N., Obara, K., & Kasahara, K. (2012). Decade-scale decrease in  $b$ -value prior to the M9 2011 Tohoku and 2004 Sumatra quakes. *Geophysical Research Letters*, 39(20), L20304. <https://doi.org/10.1029/2012gl052997>
- 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), L11307. <https://doi.org/10.1029/2005gl022679>
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401), 9–27. <https://doi.org/10.2307/2288914>
- 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. <https://doi.org/10.1111/j.1365-246x.1993.tb04663.x>
- Ogata, Y., & Katsura, K. (2006). Immediate and updated forecasting of aftershock hazard. *Geophysical Research Letters*, 33(10), L10305. <https://doi.org/10.1029/2006gl025888>
- Omori, F. (1894). On aftershocks. *Seismological journal of Japan*, 19, 71–80.
- Öncel, A. O., & Wyss, M. (2000). The major asperities of the 1999 Mw= 7.4 Izmit earthquake defined by the microseismicity of the two decades before it. *Geophysical Journal International*, 143(3), 501–506. <https://doi.org/10.1046/j.1365-246x.2000.00211.x>

- Petrucelli, 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. <https://doi.org/10.1016/j.epsl.2019.115791>
- Plourde, A. (2023). Assessing earthquake rates and b-value given spatiotemporal variation in catalog completeness: Application to Atlantic Canada. *Seismica*, 2(2). <https://doi.org/10.26443/seismica.v2i2.384>
- Riesner, M., Bollinger, L., Hubbard, J., Guérin, C., Lefèvre, M., Vallage, A., et al. (2021). Localized extension in megathrust hanging wall following great earthquakes in western Nepal. *Scientific Reports*, 11(1), 21521.
- Ringdal, F. (1975). On the estimation of seismic detection thresholds. *Bulletin of the Seismological Society of America*, 65(6), 1631–1642. <https://doi.org/10.1785/bssa0650061631>
- 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. <https://doi.org/10.1016/j.epsl.2017.11.036>
- Sambridge, M., Bodin, T., Gallagher, K., & Tkalcic, H. (2013). Transdimensional inference in the geosciences. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 371(1984), 20110547. <https://doi.org/10.1098/rsta.2011.0547>
- Sambridge, M., Gallagher, K., Jackson, A., & Rickwood, P. (2006). Transdimensional inverse problems, model comparison and the evidence. *Geophysical Journal International*, 167(2), 528–542. <https://doi.org/10.1111/j.1365-246x.2006.03155.x>
- Scholz, C. H. (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. <https://doi.org/10.1785/bssa0580010399>
- Scholz, C. H. (2015). On the stress dependence of the earthquake b-value. *Geophysical Research Letters*, 42(5), 1399–1402. <https://doi.org/10.1002/2014gl062863>
- Schorlemmer, D., Wiemer, S., & Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes. *Nature*, 437(7058), 539–542. <https://doi.org/10.1038/nature04094>
- 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. <https://doi.org/10.1002/2015jb012719>
- Shi, H., Meng, L., Zhang, X., Chang, Y., Yang, Z., Xie, W., et al. (2018). Decrease in b-value prior to the Wenchuan earthquake (ms 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. <https://doi.org/10.1785/bssa0720051677>
- Spasiani, I., Taroni, M., Murru, M., & Falcone, G. (2023). Real time gutenbergrichter 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. <https://doi.org/10.1093/gji/ggad134>
- Taroni, M., & Akinci, A. (2020). Good practices in PSHA: Declustering, b-value estimation, foreshocks and aftershocks inclusion; a case study in Italy. *Geophysical Journal International*, 224(2), 1174–1187. <https://doi.org/10.1093/gji/ggaa462>
- Tinti, S., & Gasperini, P. (2024). The estimation of b-value of the frequency–magnitude distribution and of its 1 $\sigma$  intervals from binned magnitude data. *Geophysical Journal International*, 238(1), 433–458. <https://doi.org/10.1093/gji/ggae159>
- 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. <https://doi.org/10.4294/jpe1952.14.37>
- Utsu, T. (1972). Aftershocks and earthquake statistics: Analyses of the distribution of earthquakes in magnitude, time and space with special consideration to clustering characteristics of earthquake occurrence. *Journal of the Faculty of Science, Hokkaido University. 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. <https://doi.org/10.1029/2020jb021027>
- 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. <https://doi.org/10.1785/bssa0700041337>
- Wetzler, N., Lay, T., & Brodsky, E. E. (2023). Global characteristics of observable foreshocks for large earthquakes. *Seismological Research Letters*, 94(5), 2313–2325. <https://doi.org/10.1785/0220220397>
- Wiemer, S., & Wyss, M. (1997). Mapping the frequency-magnitude distribution in asperities: An improved technique to calculate recurrence times? *Journal of Geophysical Research*, 102(B7), 15115–15128. <https://doi.org/10.1029/97jb00726>
- 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. <https://doi.org/10.1785/0120040007>
- Yin, F., & Jiang, C. (2023). Enhanced b-value time series calculation method using data-driven approach. *Geophysical Journal International*, 236(1), 78–87. <https://doi.org/10.1093/gji/ggad419>