

SeismoStats: A Python Package for Statistical Seismology

Aron Mirwald^{*1}, Nicolas Schmid¹, Leila Mizrahi¹, Marta Han¹, Alicia
Rohnacher¹, Vanille A. Ritz¹, and Stefan Wiemer¹

¹Swiss Seismological Service, ETH Zurich, Zurich, Switzerland

^{*}Email: aron.mirwald@sed.ethz.ch

November 7, 2025

Abstract

We introduce SeismoStats, a Python package that enables essential statistical seismology analyses, with a focus on well-established methods. The package provides user-friendly tools to download and manipulate earthquake catalogs, but also plotting functionalities to visualize them, as well as means to perform analyses such as estimating the a- and b-value of the Gutenberg-Richter law, or estimating the magnitude of completeness of any earthquake catalog. This is the first well-tested, well-documented, and openly accessible Python package with all these features. It is intended to serve as the nucleus of a long-term community effort, continually expanding in functionality through shared contributions. We invite seismologists and developers to contribute ideas and code to support and shape its future development.

1 Introduction

Earthquake Catalogs are a fundamental resource for seismological research and serve as the basis for a wide range of analyses. Among these analyses, some are performed routinely by most seismologists, especially those who statistically characterize seismicity. However, these analysis methods are often implemented independently, with researchers developing their own code for commonly used methods. This not only leads to duplicated work across the community but also increases the chance for inconsistencies and implementation errors.

Here, we introduce the software package SeismoStats, which aims to alleviate this duplication of effort and reduce the risk of inconsistent or incorrect implementations. The package is designed to provide functionalities such as downloading a catalog, visualizing the seismicity in space and time, as well as its distribution of magnitudes, and estimating statistically relevant parameters of the catalog. Figure 1 shows a map of seismicity in Switzerland that was created by applying one of the standard plotting functions of SeismoStats to the Swiss catalog that was downloaded using SeismoStats. Note that all figures of seismicity in this article are generated using the Swiss catalog of the year 2024 (Swiss Seismological Service, 1983).

The main focus of SeismoStats relies on one of the most important relations in statistical seismology, the Gutenberg-Richter (GR) law (Gutenberg and Richter, 1944). Estimation and analysis of the three parameters of the GR law, magnitude of completeness (m_c), a-value, and b-value, have become an ubiquitous practice in statistical seismology. Many seismologists continue to use the ZMAP software (Wiemer, 2001). However, this software is only available for the proprietary software MATLAB. In addition, it is no longer maintained, and since its last update, there have been scientific advances relevant for the estimation of the a- and b-value (Van Der Elst, 2021; Van Der Elst and Page, 2023; Lippiello and Petrillo, 2024) that are not implemented in ZMAP.

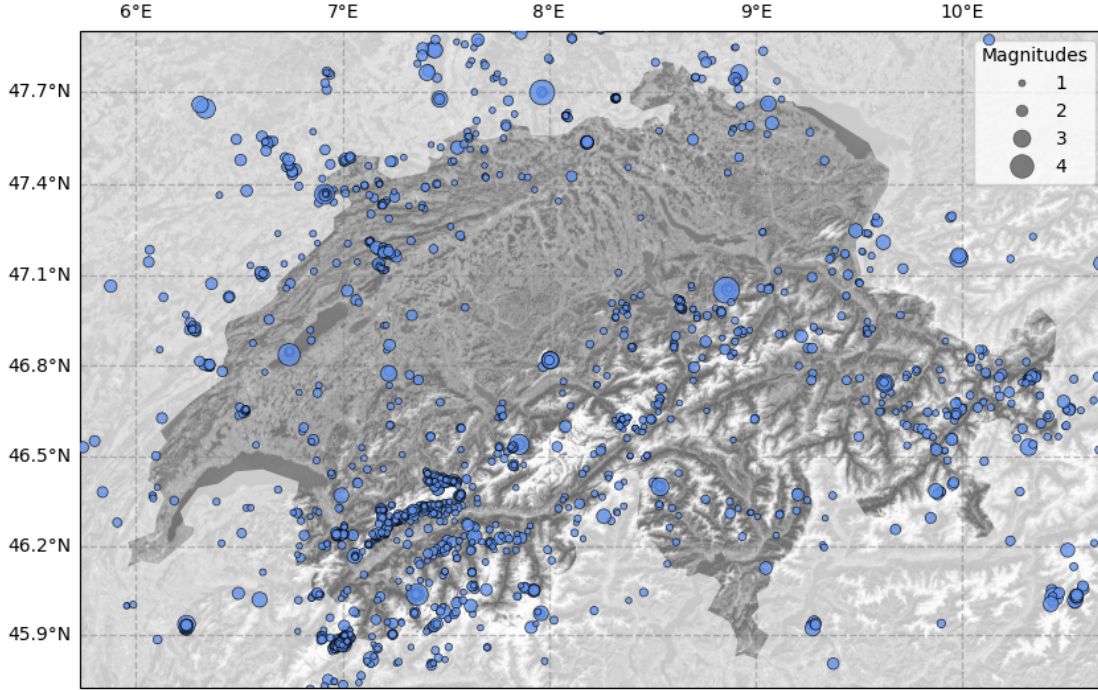


Figure 1: An example of an earthquake catalog: The Swiss earthquake catalog for the year 2024 (Swiss Seismological Service, 1983). Figure generated using the SeismoStats function `plot_in_space`.

Alt-text: Swiss map with dots whose size corresponds to the earthquake magnitude.

SeismoStats fills this gap by providing robust, well-tested, and well-documented open source tools for estimating GR parameters using the widely adopted Python programming language. The package incorporates several recent methodological improvements for estimating a - and b -values, while also offering additional functionality for common tasks in catalog-based analysis, such as downloading and visualizing earthquake catalogs, and evaluating temporal variation in the b -value. Although GR parameter estimation is currently a central feature, the package is not limited to this application. Its modular and extensible architecture supports the integration of further statistical methods, and as an open-source project under active development, SeismoStats welcomes contributions from the scientific community to ensure it continues to grow with the evolving needs of seismologists.

SeismoStats complements other Python libraries for seismology, such as ObsPy (Beyreuther et al., 2010) for waveform processing, SeisComP (GFZ Data Services, 2008) for real-time

analysis, pyCSEP (Savran et al., 2022; Graham et al., 2024) for forecast testing, SeisBench (Woollam et al., 2022) for ML applications, and OpenQuake (Pagani et al., 2014) for hazard modeling, by addressing a widely used but under-served area: statistical analysis of earthquake catalogs. Together, these tools provide a powerful and flexible foundation for modern seismological research.

This article is structured into two parts: In the first part, we present the fundamental structure of the package and its general capabilities. In this part, we also explain how the community can contribute to the package in the future. In the second part, we discuss the estimation methods of GR parameters (m_c , b-value and a-value) that are implemented in SeismoStats in detail.

2 Philosophy and Implementation of SeismoStats

In Figure 2, the overarching structure of the package is shown. SeismoStats is designed to facilitate the statistical analysis of earthquake catalogs. To achieve this, we follow a two-fold approach. First, SeismoStats simplifies access to and handling of earthquake catalogs by providing the Catalog class (see panel (b) of Figure 2) that facilitates catalog analysis, even for users with limited experience in the field of statistical seismology. It integrates seamlessly with functions to download earthquake catalogs and return them as Catalog objects, as well as with tools to convert between the standard QuakeML format (Danijel Schorlemmer et al., 2011) and the Catalog format. Both the Catalog class and the download client are located at the very top level of the package. Second, SeismoStats provides flexible, standalone functions that can operate on common Python data structures such as lists, NumPy arrays (Harris et al., 2020), and pandas series (The pandas development team, 2020). This design ensures flexibility and allows users to incorporate the functionality of the package into their existing workflows with minimal adaptation. These functions are located within three sub-packages: `analysis`, `utils` and `plots`.

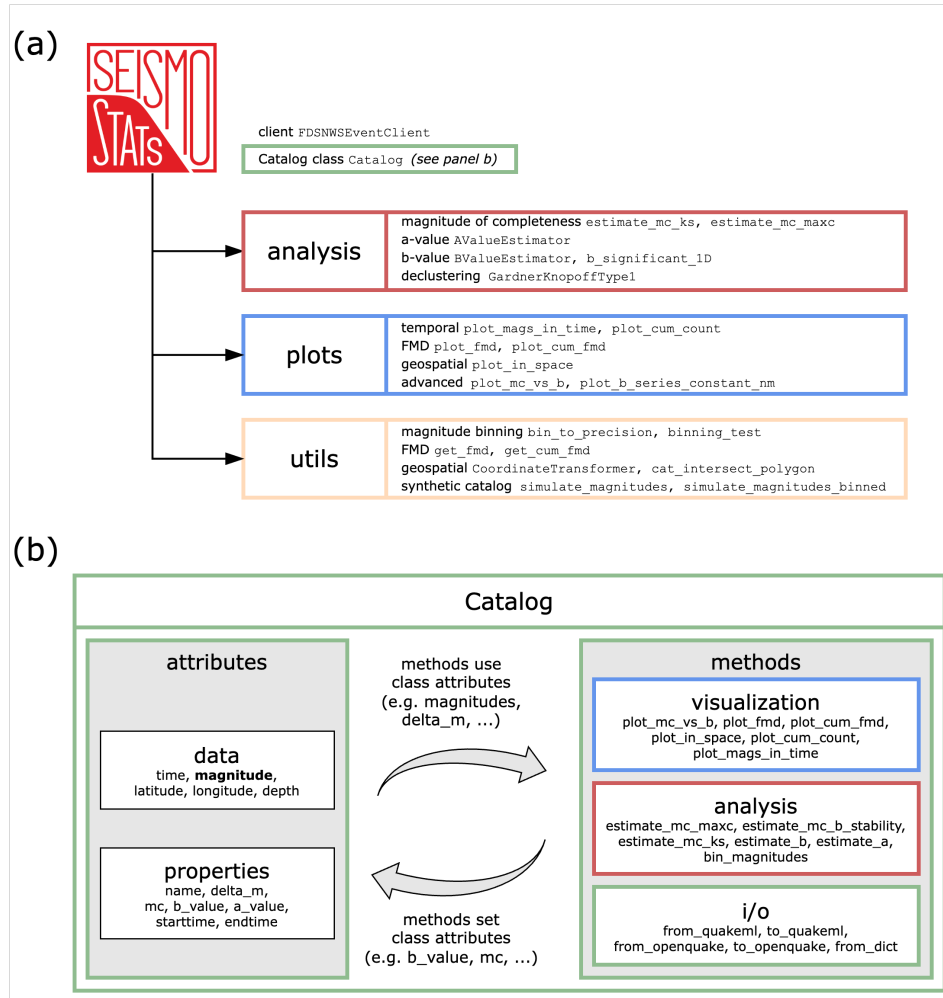


Figure 2: Overview of the SeismoStats package. a) Organisation of the subpackages with selected functions. b) Detailed architecture of the Catalog class in which methods can both use and update the catalog's properties and data.

Alt-text: Schematic diagram of the structure of SeismoStats.

2.1 The Catalog class

At the core of the package is the Catalog class, a subclass of the pandas DataFrame (The pandas development team, 2020). This design choice enables users to benefit from the powerful data manipulation capabilities of pandas while adding specialized methods and properties for earthquake analysis.

A Catalog object must contain at least a ‘magnitude’ column, and additional columns, like ‘latitude’, ‘longitude’, ‘depth’, and ‘time’, can be added as required by the user. The initialization and manipulation of the Catalog are identical to those of a pandas DataFrame. However, the Catalog class provides additional methods and properties, each of which requires certain columns to be present in order to be available. This ensures seamless integration with existing pandas workflows while extending functionality for catalog-specific tasks.

Many of the functions within the three sub-packages are available as methods of the Catalog class, with the added benefit that they can use and update the catalogs properties and data. Examples of such methods include the estimation of the magnitude of completeness m_c (which stores m_c as a property) and the b-value (which uses m_c as an input automatically), as well as visualization of the seismicity.

The usefulness of the Catalog class is further enhanced by the package’s methods to facilitate the import and analysis of earthquake data. For instance, QuakeML files can be converted into a Catalog, making it easy to ingest data from standard formats. In addition, SeismoStats supports downloading earthquake catalogs from FDSN Web Services using the `FDSNWSEventClient`. These features simplify the process of acquiring and working with earthquake data from local or international sources.

2.2 Standalone functions

The standalone functions in `SeismoStats` are organized into three sub-packages. `SeismoStats.analysis` provides tools for the statistical analysis of earthquake catalogs. This includes functions for estimating and evaluating Gutenberg–Richter parameters, such as the magnitude of completeness, a-value, and b-value, as well as a method for declustering seismic events.

Estimation methods for the a-value and b-value are implemented as a class to ensure consistency across the different methods and to enhance usability. For example, all a- and b-value methods require the input of magnitudes, magnitude of completeness and bin-size of the discretized magnitudes, and have the property n for the number of magnitudes used for the analysis. In contrast, 111 functions to estimate completeness are implemented as standalone functions, since they share fewer common features. Tools to evaluate the quality and robustness of estimates include the Shi and Bolt (Shi and Bolt, 1982) standard deviation of the b-value, the Lilliefors test (Lilliefors, 1969; Herrmann and Marzocchi, 2021) to assess whether the Gutenberg–Richter law is a valid assumption, and the b-significance method (Mirwald et al., 2024) to test whether the b-value varies significantly. Finally, the declustering method proposed by Gardner and Knopoff (1974) is implemented.

`SeismoStats.plots` offers functions for visualizing seismic data and analysis results. Built on modern plotting libraries such as `matplotlib`, `cartopy`, and `geopandas` (Hunter, 2007; Met Office, 2010; The pandas development team, 2020), this module includes functions for mapping seismicity in space and time, plotting frequency-magnitude distributions (FMDs), and visualizing parameter estimates. All figures in this article are created using the plotting functionality of `SeismoStats`.

`SeismoStats.utils` contains utility functions that support statistical workflows. These include tools for testing the discretization of magnitudes, generating synthetic magnitudes,

and preparing data for analysis, for example, by binning the magnitudes. Further capabilities are the transformation of coordinates to local ones, and filtering catalogs based on a polygon.

2.3 Development Practices

SeismoStats is developed following best practices in software development, ensuring high quality and maintainability. Continuous integration and continuous development (CI/CD) workflows are implemented to maintain stability, supported by extensive unit tests with high coverage that catch errors early and improve reliability. Semantic versioning is used to clearly manage changes and updates, and PEP 8 style guides are followed to maintain code style consistency across the package. Documentation is generated automatically, ensuring that it is always up-to-date, while a detailed user guide complements this by explaining the most relevant functionalities in depth with numerous examples.

To build on a solid foundation, SeismoStats leverages established Python libraries such as Cartopy (Met Office, 2010), Geopandas (The pandas development team, 2020), Matplotlib (Hunter, 2007), Numpy (Harris et al., 2020), Pandas (The pandas development team, 2020), Scipy (Virtanen, Pauli et al., 2020), Shapely (Sean Gillies et al., 2022), and Statsmodels (Seabold and Perktold, 2010).

As an open-source project hosted on GitHub, SeismoStats is designed to be a long-lasting community effort, and we invite fellow seismologists to actively contribute ideas, feedback and code. There are no strict limitations: any method may be implemented, provided that it is useful for more than one person and that it is not excessively complex. We believe that this collaborative approach best benefits the earthquake research community by continuously improving and expanding the package’s capabilities.

3 Implemented methods to estimate GR parameters

The GR law states that earthquake magnitudes are exponentially distributed (Gutenberg and Richter, 1944; Ishimoto and Iida, 1939).

$$N(m) = 10^{a-b(m-m_c)}, \quad (1)$$

where $N(m)$ is the number of earthquakes with a magnitude equal to or larger than magnitude m above a magnitude threshold m_c . The so-called a- and b-value are parameters that quantify the total number of earthquakes and the relative frequency of small vs. large earthquakes, respectively. The threshold m_c in this context is the magnitude of completeness above which we assume to be able to detect all earthquakes.

In this section, we introduce each implemented method to estimate the magnitude of completeness, b-value, and a-value.

3.1 Methods to evaluate completeness

Earthquakes below a certain size often remain undetected by seismic networks (Wiemer and Wyss, 2000). Whether an earthquake can be detected by a network depends on various factors. These include the type and distribution of instruments present in a network, their signal-to-noise ratio (Schorlemmer and Woessner, 2008; Wiemer and Wyss, 2000), as well as the algorithms used for event detection (Herrmann and Marzocchi, 2021; Schaff and Waldhauser, 2010).

The determination of the magnitude of completeness (m_c) in seismic catalog analysis is a crucial step in assessing the reliability and sensitivity of recorded earthquake events. Without it, the evaluation of the a- and b-value is futile (Eq. 1), as the estimates will be affected by assuming a wrong completeness level (i.e. the underestimation of m_c leads to underestimation

of the b-value due to the lower-than-expected count of low-magnitude events above m_c). Over the years, many methods have been proposed to determine m_c (Woessner and Wiemer, 2005; Schorlemmer and Woessner, 2008; Ogata and Katsura, 1993; Mignan and Woessner, 2012). The SeismoStats package includes three commonly used approaches. These are described in the following, although the list is not intended to be exhaustive. An example visualization of the three different m_c estimation methods applied to the Swiss catalog is shown in Figure 5. Further methods can be added by the community as they are developed.

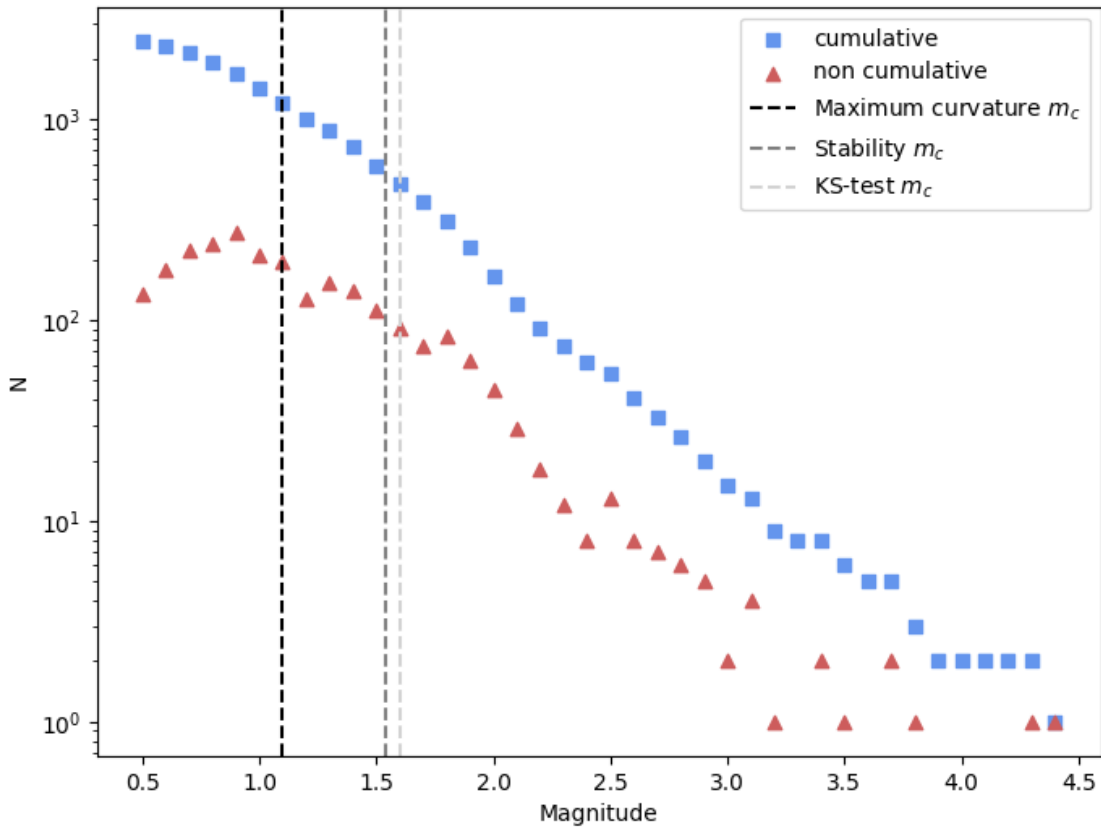


Figure 3: Visualization of the application of the three completeness estimation methods to the Swiss catalog in the year 2024. The blue squares represent the cumulative FMD, the red triangles the non-cumulative one. The vertical dashed lines indicate the completeness magnitude that was estimated using the different available methods. Figure generated using SeismoStats functions `plot_fmd` and `plot_cum_fmd`.

Alt-text: The Frequency of earthquakes in relation to their magnitude is shown as scatter plot, together with dotted lines that the result of three different completeness methods.

In SeismoStats, all completeness methods are implemented as functions with a tuple as

output, where the first entry is the estimated magnitude of completeness and the second entry provides additional outputs in a dictionary. This ensures that completeness functions can be used interchangeably despite their variable outputs. Although each function has different input parameters, the minimally required input for all is an array of magnitudes and the bin size of the discretized magnitudes, Δm (hereafter referred to as bin size).

The user should be aware that the methods implemented assume the magnitude of completeness to be time-independent. Although an unchanging magnitude of completeness can be a useful model in many cases, it is well known that the level of completeness can change over time (Helmstetter, 2006; Kagan, 2004). This can occur, for example, due to changes in the seismic network (Schorlemmer and Woessner, 2008), the overlap of waveforms (Mignan and Woessner, 2012), varying noise levels between days and nights (Rydelek and Sacks, 1989), or changes in the event detection methods.

Maximum Curvature

The Maximum Curvature method (MAXC) defines the magnitude of completeness of a catalog as the magnitude at which the non-cumulative frequency magnitude distribution (FMD) is maximal (Wiemer and Wyss, 2000), plus a correction term to avoid underestimation,

$$m_c = \underset{m_i}{\operatorname{argmax}} (N(m_i)) + \delta, \quad (2)$$

where $N(m_i)$ is the number of earthquakes within a magnitude bin of width Δm^* (hereafter called bin width) centered at m_i , and δ is the correction term, which is typically set to $\delta = 0.2$ (Woessner and Wiemer, 2005). Note that the bin width Δm^* considered to calculate m_c can differ from the bin size of the discretized magnitudes Δm . That is, a catalog can have magnitudes discretized to bins of $\Delta m = 0.01$, but one might be interested in estimating m_c with a precision of $\Delta m^* = 0.1$. Within SeismoStats, the bin size of the discretized

magnitudes Δm is called `delta_m`, while bin width Δm^* , which depends on the user's needs and is not a property of the data, is called `fmd_bin`.

Implementation in SeismoStats: The function `estimate_mc_maxc` requires the array of magnitudes and the bin width Δm^* (`fmd_bin`) as input. The correction factor δ is set to 0.2 by default, as in the original article (Woessner and Wiemer, 2005), but can be specified otherwise by the user.

m_c by b-value stability.

Introduced by Cao and Gao (2002) and refined by Woessner and Wiemer (2005), this method is based on the experience that the b-value for incomplete catalogs is underestimated. When excluding successively the smallest events by increasing the lower magnitude cutoff threshold, the b-value therefore increases. Under the assumption that the magnitudes are exponentially distributed, the b-value is expected to stabilize when the completeness is reached. The completeness threshold is thus defined as the magnitude at which the b-value is relatively stable in relation to its theoretical standard deviation (Shi and Bolt, 1982). In practice, this is achieved as follows:

$$m_c = \min \left\{ m_i \mid \text{abs} \left(\frac{1}{K} \sum_{k=1}^K b(m_i + k \cdot \Delta m^*) - b(m_i) \right) < \sigma_{b(m_i)} \right\}, \quad (3)$$

where $b(m)$ is the b-value estimate of all magnitudes above m , σ_b is its uncertainty (see Eq. 6), Δm^* is the magnitude bin size considered for m_c precision and K is the number of m_c bins that is considered for stability. $L = K \cdot \Delta m^*$ is then the length of the magnitude range considered for stability. Note that in SeismoStats, the input required for b-value estimation is L , not K .

Implementation in SeismoStats: The function `estimate_mc_b_stability` requires the array of magnitudes and the bin size of the discretized magnitudes Δm as input. The magnitudes

at which the stability should be tested, as well as the stability range can be additionally provided. By default, the function tests at all magnitudes between the minimum and maximum magnitude of the sample, and the stability range L is 0.5, as in the original definition of the method (Woessner and Wiemer, 2005).

KS-distance.

The determination of the magnitude of completeness through the Kolmogorov-Smirnov (KS) distance method is a statistical approach in which observed and expected cumulative distribution functions (CDFs) are compared (Clauset et al., 2009; Mizrahi et al., 2021). The underlying assumption is that the magnitudes are, in fact, exponentially distributed. The completeness threshold is then defined as the magnitude at which the deviation of the observed CDF compared to the expected one is within a certain threshold,

$$m_c = \min \{m_i \mid p(D_{KS}^i) \geq p_{th}\}, \quad (4)$$

where D_{KS}^i is the KS-distance of the observed CDF from the theoretical expected distribution when using m_i as a lower cutoff, p_{th} is the threshold that can be chosen freely, and $p(D_{KS})$ is the probability of a KS-distance equal or larger than the observed one under the assumption that the magnitudes follow the theoretical expected distribution.

It is important to note that KS tests are generally sensitive to the sample size. This means that with an increasingly large sample of magnitudes, increasingly minor deviations from the exponential distribution can be detected. As many real earthquake catalogs exhibit some kind of small artifacts in the magnitude distribution, this might make the KS method impractical for very large datasets.

Implementation in SeismoStats: The function `estimate_mc_ks` requires the array of magnitudes and the bin size of the discretized magnitudes Δm as input. The KS test is imple-

mented such that $p(D_{KS})$ is estimated through the simulation of n (set to 10,000 by default) samples of discretized magnitudes. These are drawn from an exponential distribution with the b-value that was estimated from the observed sample, unless the b-value is given as input. The threshold is set to $p_{th} = 0.1$ by default, which is a commonly used value. The smaller the p-value threshold is, the stricter the KS test.

3.2 b-value Estimation Methods

The b-value of the GR law (Eq. 1) is a measure that quantifies the relative frequency of large vs. small earthquakes. The long-term global b-value is around one (El-Isa and Eaton, 2014), indicating that the frequency of earthquakes increases tenfold for each unit decrease of magnitude. A lower b-value relates to a larger share of large magnitudes, while a higher b-value relates to a smaller share of large magnitudes.

In the case of continuous magnitudes (i.e., no discretization), the b-value can be estimated by the maximum likelihood estimation (MLE) (Aki, 1965):

$$\hat{b} = \frac{\log e}{\frac{1}{n} \sum_{i=1}^n m_i - m_c}, \quad (5)$$

where m_c is the magnitude of completeness and m_i are the individual magnitudes within the catalog. All the b-value estimation methods described in this article are based on the above approach. However, they differ from it in that they consider the discretization of magnitudes, and two of the methods use magnitude differences instead of the magnitudes themselves (see Sections 3.2 and 3.2). In Figure 4, we show an example of an estimation of the b-value. In Figure 5, we show an example of the b-value variation with time, estimated with two different methods.

When magnitudes are considered to be random variables, the estimate of the b-value (Eq. 5) is also a random variable, with a mean and variance that depend on the number n_m

of earthquake magnitudes used for the estimate. For a sufficiently large n_m , the b-value estimate can be considered to follow a normal distribution (Shi and Bolt, 1982), with a standard deviation of

$$\sigma(\hat{b}) = \frac{\ln(10)b^2}{\sqrt{n_m - 1}} \sqrt{\text{var}(m)}. \quad (6)$$

where b is the true b-value, \hat{b} is its estimate, and $\text{var}(m)$ is the variance of the magnitudes which can be estimated from the data.

All b-value methods are implemented as classes in SeismoStats. This ensures homogeneity between the implementation of different methods and has the advantage that, apart from the b-value itself (`estimator.b_value` or `estimator.value`), other parameters and products of the estimation can be easily accessed. These include the standard deviation of the estimate (`estimator.std`), the number of earthquakes used for the estimation (`estimator.n`), and the number of magnitudes above completeness used for the estimate (`estimator.magnitudes`). For the standard deviation, all functions use the estimate of Eq. 6, with the exception of the b-more-positive method, where a bootstrap method was implemented due to the incompatibility of the former. Note that although Eq. 6 defines the uncertainty of the estimate assuming continuous magnitudes (Eq. 5), in SeismoStats it is used even if the magnitudes are discretized. This is due to the lack of an alternative that does account for the discretization and the fact that it is commonly used. Finally, the Lilliefors test which tests if a given sample is exponentially distributed (Herrmann and Marzocchi, 2021; Lilliefors, 1969) is implemented as a method which returns a p-value.

Classical maximum likelihood estimate

In most catalogs of natural seismicity, the magnitudes are discretized. If not accounted for, this leads to a bias in the b-value estimate. Tinti and Mulargia (1987) and Marzocchi and Sandri (2003) derived the exact MLE for the b-value in the case of discretized magnitudes:

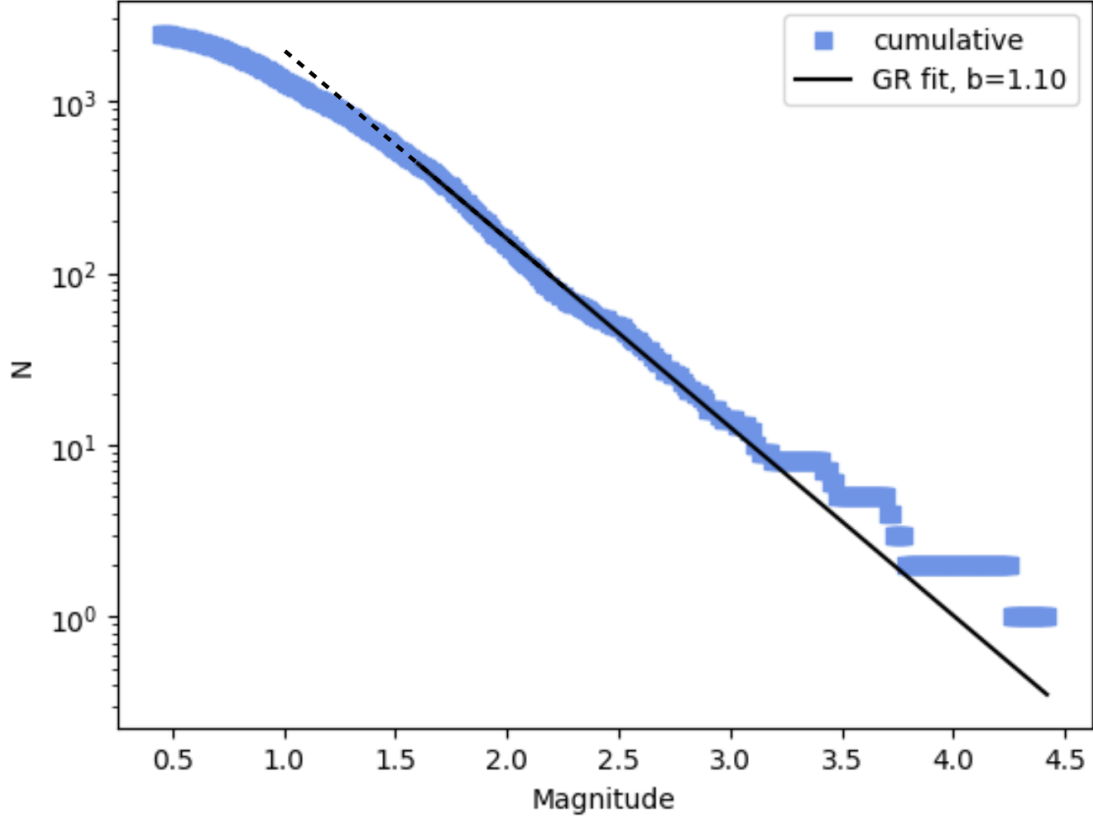


Figure 4: Example of an estimated b-value fit. The estimator `ClassicBValueEstimator` was applied, with a bin size of $\Delta m = 0.01$ and assuming an incompleteness of $m_c = 1.6$. The GR-law with the estimated b-value is shown in black, the corresponding cumulative FMD is shown in blue. The dotted line is the extension of the GR to the reference magnitude $M=1$. Figure generated using SeismoStats function `plot_cum_fmd`.

Alt-text: Plot of the frequency vs the magnitude of earthquakes, together with a line that fits the slope.

$$\hat{b} = \frac{1}{\ln(10)\Delta m} \ln p, \text{ where} \quad (7)$$

$$p = 1 + \frac{\Delta m}{\frac{1}{n} \sum_{i=1}^n m_i - m_c}.$$

Note that for the case of continuous magnitudes ($\Delta m = 0$), Eq. 7 is reduced to Eq. 5.

Implementation in SeismoStats: For the estimation of the b-value, the class `ClassicBValueEstimator` requires the array of magnitudes, the magnitude of completeness

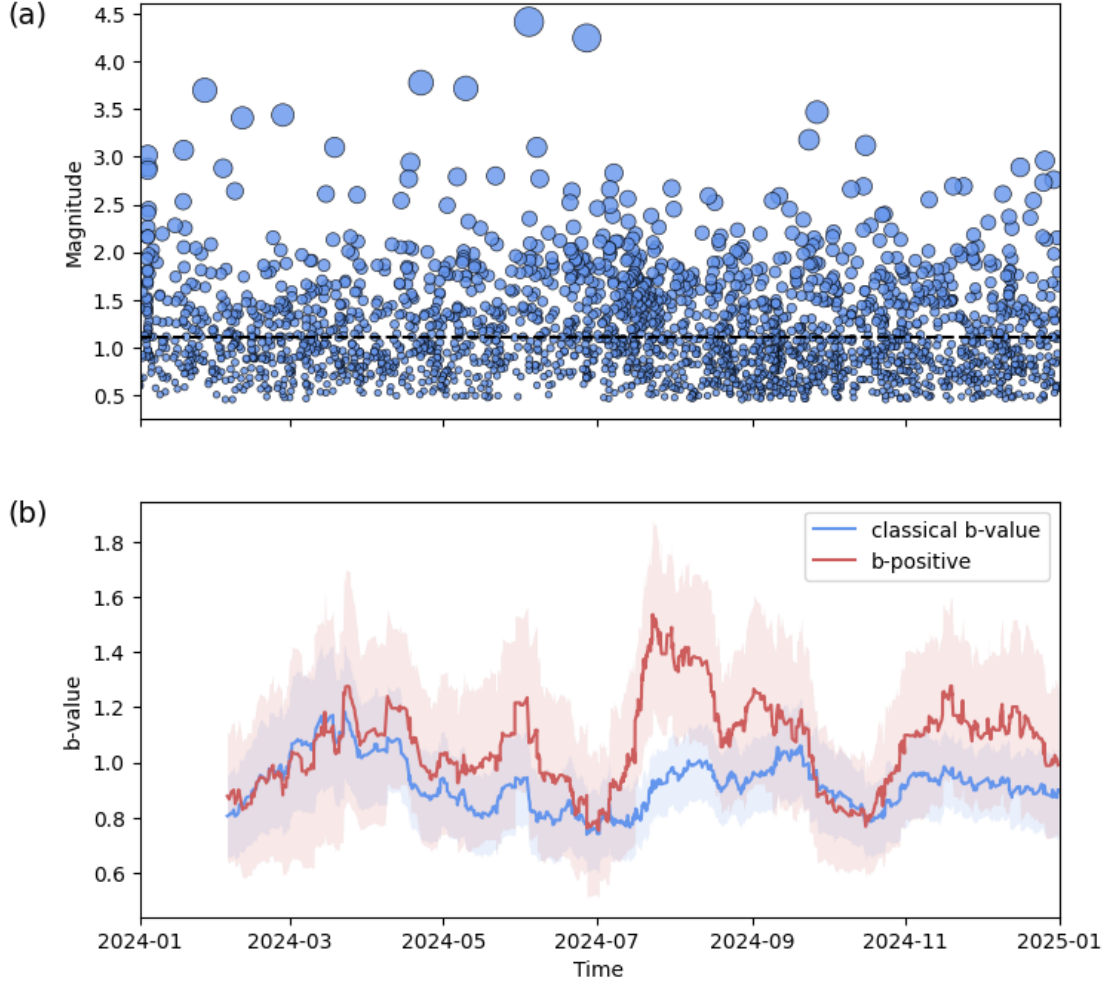


Figure 5: Illustration of temporal variation of the b-value both for the classical b-value estimate and b-positive using the Swiss earthquake catalog for 2024. On the top panel, the magnitudes over time are shown, figure generated using SeismoStats function `plot_mag_in_time`. On the bottom panel, the classical b-value and b-positive over time, each b-value is estimated using 100 magnitudes above m_c . The b-value plotted at each time is estimated with magnitudes up to that time. This figure was generated using SeismoStats function `plot_b_series_constant_nm`.

Alt-text: (a): scatter plot of magnitudes vs. time, with a dotted line indicating the completeness of 1.1. (b): b-value as a function of time, estimated with the classical and the b-positive method.

and the bin size Δm as input. Weights can be given optionally.

Utsu’s approximate estimate

As an alternative to the exact result, SeismoStats also provides an approximate introduced by Utsu (1965):

$$\hat{b} = \frac{\log e}{\frac{1}{n} \sum_{i=1}^n m_i - m_c + \Delta m/2}. \quad (8)$$

This estimate, although not exact, is very commonly used and accurate for small bins.

Implementation in SeismoStats: For the estimation of the b-value, the class `UtsuBValueEstimator` requires the array of magnitudes, the magnitude of completeness, and the bin size Δm as input.

b-positive

After large earthquakes, smaller earthquakes are harder to detect, leading to a temporary drop in the fraction of small earthquakes detected. This effect is known as short-term aftershock incompleteness (STAI). STAI effectively increases the magnitude of completeness after a large event has occurred. In an effort to reduce the impact of STAI on the b-value estimate, Van Der Elst (2021) introduced a b-value estimation method that utilizes the positive differences of subsequent magnitudes. The method, called b-positive, effectively assumes that the completeness magnitude at each point in time is at least as large as the last earthquake magnitude recorded plus some margin (δm_c).

The b-positive estimate is performed as follows. First, the differences between consecutive magnitudes ($m_i - m_{i-1}$) are calculated. Then, only the differences that are positive and larger than a threshold, δm_c , are selected (hence the name 'b-positive'). With the obtained magnitude differences, the b-value is estimated using the classical MLE (see section 3.2). This is based on the fact that the difference between two exponentially distributed variables

follows a Laplace distribution.

Although the b-positive method effectively eliminates the effect of short-term aftershock incompleteness, it is still affected by normal incompleteness resulting from the detection capabilities of the seismic network (Lippiello and Petrillo, 2024). This effect can be reduced by considering only the complete catalog (above m_c) before taking the differences, or by increasing the threshold δm_c .

Implementation in SeismoStats: For the estimation of the b-value, the class `BPositiveBValueEstimator` requires the array of magnitudes, the magnitude of completeness and the bin size Δm as input. The user can also specify the threshold δm_c can be given as input, with the default value being one magnitude bin. If the user supplies a time array, it will be used to order the events in time, otherwise it is assumed that they are already ordered. Weights can be given optionally.

b-more-positive

Based on the b-positive method, Lippiello and Petrillo (2024) developed the b-more-positive method. This method also uses magnitude differences, but the differences are constructed in an alternative way. For each earthquake of magnitude m_i , the next earthquake with a magnitude greater or equal than $m_i + \delta m_c$ is used to calculate the differences, effectively resulting in a higher number of magnitude differences compared to b-positive. However, this does not necessarily translate into a smaller uncertainty of the estimate. In our synthetic tests, we found that the actual standard deviation of the estimate is larger than the expected standard deviation (Eq. 6). Therefore, SeismoStats uses a bootstrap method to estimate the standard deviation for the b-more-positive estimate.

Implementation in SeismoStats: For the estimation of the b-value, the class `BMorePositiveBValueEstimator` requires the array of magnitudes, the magnitude of completeness and the bin size Δm as input. The user can also specify the threshold δm_c can

be given as input, with the default value being one magnitude bin. If the user supplies a time array, it will be used to order the events in time, otherwise it is assumed that they are already ordered. Weights can be given optionally.

3.3 a-value Estimation Methods

The a-value of the GR law (Eq. 1) quantifies the number of earthquakes that occurred in a given volume and time. Together with the b-value, it is typically used to estimate the probability of an earthquake larger than m (Tormann et al., 2014; Ritz et al., 2022)

Similarly to the b-value estimation, the a-value estimation methods are implemented as classes in SeismoStats. One difference compared to the b-value estimator classes is that they lack the standard deviation property. The reason for this is that most published literature does not provide any error estimation for the a-value. The aim of SeismoStats is to provide implementations of commonly used, already existing methods. Defining an a-value uncertainty estimation method is thus beyond the scope of this work, though one might consider implementing the standard deviation of the a-value under the assumption that seismicity follows a Poisson process. Otherwise, the classes have the properties `estimator.a_value` or `estimator.value`, `estimator.n` and `estimator.magnitudes`, similarly to the b-value estimator classes.

Classical estimate

With Eq. 1, we can estimate the a-value as the base-10 logarithm of the number of earthquakes above completeness,

$$a = \log N(m_c). \quad (9)$$

There are, however, two commonly used modifications of the definition above, which are

relevant when comparing different a-values to each other.

First, the completeness magnitudes, relative to which different a-values are calculated, may not be constant. Therefore, in practice, the a-value is often estimated with respect to a certain reference magnitude, such that $10^{a_{m_{\text{ref}}}} = N(m_{\text{ref}})$. Here, $N(m_{\text{ref}})$ is not the actual number of earthquakes above m_{ref} , but the extrapolated number if the GR law were perfectly valid above and below m_c , as shown in Fig. 4. The new a-value can now be estimated using the a-value defined in Eq. 9: $a_{m_{\text{ref}}} = a - b(m_{\text{ref}} - m_c)$.

Second, the time intervals for which different a-values are compared often differ, making the comparison more difficult. In many cases, researchers scale the a-value so that $N(m)$ refers to the number of earthquakes above m within a year, which is effectively a rate rather than an absolute number. This possibility is implemented in SeismoStats in the form of a scaling factor that can be provided as an input to the a-value estimator. This scaling factor quantifies how many time units fit into the observation interval. For example, if the catalog spans 10 years and the a-value is to be scaled to one year, the scaling factor is 10. Note that the scaling factor can analogously be applied to facilitate a comparison between catalogs spanning different spatial extents: To compare the number of earthquakes in two different volumes, one may be interested in the number of earthquakes per cubic km. For a volume of 100 cubic km, the scaling factor would thus be 100.

Implementation in SeismoStats: For the estimation of the a-value, the class `ClassicAValueEstimator` requires the array of magnitudes, the magnitude of completeness and the bin size Δm as input. The user can also specify the threshold δm_c can be given as input, with the default value being one magnitude bin. Weights can be given optionally.

a-positive

Similarly to the b-value estimation, we have implemented "positive" a-value estimation methods. These are built on the implicit assumption that the momentary completeness magnitude

is given by the magnitude of the last detected earthquake (plus a margin, called δm_c). We based our definitions on Van Der Elst and Page (2023), with the difference that we homogenized the naming convention with the b-value methods, so that "positive" means that the differences of subsequent events are used for the rate estimation, and "more positive" means that for each earthquake, the difference to the next larger one in the catalog is used for the rate estimation. The latter method is described by Van Der Elst and Page (2023).

For both methods, the main idea is to estimate the share of time covered by the inter-event times used to estimate the a-value. For the a-positive method, this is fairly straightforward: for each positive difference $m_{i+1} - m_i > \delta m_c$, the corresponding time difference is $\Delta t_i = t_{i+1} - t_i$. Then, the a-value can be estimated as:

$$a^+ = \log n^+ - \log \frac{\sum_{i=1}^{n^+} \Delta t_i}{T}, \quad (10)$$

where n^+ is the number of positive differences in the catalog, and T is the length of the entire observation interval. The result of this computation can be directly compared with the classical a-value.

Implementation in SeismoStats: For the estimation of the a-value, the class `APositiveBValueEstimator` requires the array of magnitudes, the array of times, the magnitude of completeness and the bin size Δm as input. The user can also specify the threshold δm_c , with the default value being one magnitude bin. Weights can be given optionally.

a-more-positive

For the a-more-positive method (as described in Van Der Elst and Page (2023)), time differences Δt_i to the next larger event are calculated for each event. The time differences are conditioned on the magnitude being larger than the previous one, so in order to estimate

the rate above m_c , they have to be scaled according to the GR law,

$$\tau_i = \Delta t_i 10^{-b(m_i + \delta m_c)}, \quad (11)$$

where m_i is the magnitude of the first earthquake of the pair of earthquakes with the time difference Δt_i . Therefore, this method depends on the b-value as an input.

Next, we must take into account the earthquakes that did not have a larger counterpart within the data in order to avoid an estimation bias. This is done by including the scaled open intervals in the estimate.

$$T_j = (T - t_j) 10^{-b(m_j + \delta m_c)}, \quad (12)$$

where t_j are times of events of magnitude m_j that are not followed by a larger event (by the margin δm_c) within the observation interval.

Finally, the a-more-positive a-value estimate is given as

$$a^{++} = \log n^{++} - \log \frac{\sum_{i=1}^{n^{++}} \Delta \tau_i + \sum_{j=1}^m T_j}{T}, \quad (13)$$

where n^{++} is the number of closed intervals, m is the number of open intervals, and $n^{++} + m$ is the total number of events above the completeness magnitude. Note that this equation is the same as put forward by Van Der Elst and Page (2023), with the exception that the total length of the time period T is included in our formulation. This detail has the effect that the classical a-value estimator and the a-more-positive estimator have the same expected value and can therefore be directly compared.

Implementation in SeismoStats: For the estimation of the a-value, the class `AMorePositiveBValueEstimator` requires the array of magnitudes, the array of times, the

magnitude of completeness, the bin size Δm and the b-value as input. The user can also specify the threshold δm_c as input, with the default value being one magnitude bin. Weights can be given optionally.

3.4 Pitfalls and Limitations

Although estimating the parameters of the Gutenberg-Richter (GR) law is straightforward, and the methods implemented in the SeismoStats package have been thoroughly tested, several limitations and potential sources of bias must be considered when applying these methods (Herrmann and Marzocchi, 2021; Marzocchi et al., 2020; Gulia and Gasperini, 2021).

For example, the m_c estimation methods currently implemented in SeismoStats do not account for spatial or temporal variations in m_c , despite evidence that such variability is common in seismic catalogs (Schorlemmer and Woessner, 2008). Catalogs may also contain human-made events, such as quarry blasts (Gulia and Gasperini, 2021). This can introduce artifacts both in the a- and b-value estimation. Furthermore, many catalogs contain different magnitude types, and conversion between magnitude types is not straightforward. Failing to address this heterogeneity can significantly affect the estimated magnitude distribution. (e.g., Deichmann, 2017).

Although the SeismoStats package provides robust and well-tested methods, it does not automatically correct for data quality issues. Users of SeismoStats are therefore encouraged to carefully and critically examine the data before applying statistical methods.

4 Conclusion

This article has outlined the key functionalities and features of SeismoStats, an open-source Python package for performing statistical analysis on earthquake catalogs. By simplifying the

process of statistical analysis and improving accessibility to valuable insights, SeismoStats aims to advance seismological research and contribute to collective efforts to mitigate and prepare for seismic hazards.

Currently, the primary capability of SeismoStats is the estimation of parameters of the Gutenberg–Richter law. However, this represents only the beginning of a long-term community effort, with plans to expand the package’s functionalities over time. Planned future developments include tools for estimating parameters of the Omori-Utsu law (Utsu et al., 1995), calculating the fractal dimension of seismicity (Kagan, 2007), and functionalities to rapidly compare different earthquake catalogs. Furthermore, our objective is to improve interoperability and compatibility with other established seismological packages such as pyCSEP (Savran et al., 2022; Graham et al., 2024) to improve collaborative development across the seismological community.

We strongly encourage the community to actively contribute to the ongoing enhancement of SeismoStats. Researchers and developers are invited to share suggestions, propose new features, or contribute code through our GitHub repository. Those interested in discussing ideas or potential collaborations are also welcome to contact the authors directly. Together, we hope to advance SeismoStats and support the continued progress of seismological research.

5 Data and Resources

All figures provided in this article are based on the Swiss catalog between 1 January 2024 and 1 January 2025 (Swiss Seismological Service, 1983).

The complete documentation can be found here: <https://seismostats.readthedocs.io/>. The corresponding code repository can be found here: <https://github.com/swiss-seismological-service/SeismoStats>

6 Declaration of Competing Interests

The authors declare no competing interests.

7 Acknowledgments

References

- Keiiti Aki. Maximum Likelihood Estimate of b in the Formula $\log N = a - bM$ and its Confidence Limits. *Bulletin of the Earthquake Research Institute*, 43(2):237–239, September 1965. doi: 10.15083/0000033631. URL <https://repository.dl.itc.u-tokyo.ac.jp/records/33631>.
- M. Beyreuther, R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3):530–533, May 2010. ISSN 0895-0695. doi: 10.1785/gssrl.81.3.530. URL <https://pubs.geoscienceworld.org/srl/article/81/3/530-533/143693>.
- A. M. Cao and S. S. Gao. Temporal variation of seismic b -values beneath northeastern Japan island arc. *Geophysical Research Letters*, 29(9):art. no.–1334, 2002. URL <GotoISI>://000178888000029. Type: Journal article.
- Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. Power-Law Distributions in Empirical Data. *SIAM Review*, 51(4):661–703, November 2009. ISSN 0036-1445, 1095-7200. doi: 10.1137/070710111. URL <http://epubs.siam.org/doi/10.1137/070710111>.
- Danijel Schorlemmer, Fabian Euchner, Philipp Kästli, and Joachim Saul. QuakeML: status of the XML-based seismological data exchange format. *Annals of Geophysics*, 54(1), April 2011. ISSN 2037416X. doi: 10.4401/ag-4874. URL <http://www.annalsofgeophysics.eu/index.php/annals/article/view/4874>.

- Nicholas Deichmann. Theoretical Basis for the Observed Break in M_1 / M_w Scaling between Small and Large Earthquakes. *Bulletin of the Seismological Society of America*, 107(2):505–520, April 2017. ISSN 0037-1106, 1943-3573. doi: 10.1785/0120160318. URL <https://pubs.geoscienceworld.org/bssa/article/107/2/505-520/354180>.
- Z.H. El-Isa and David W. Eaton. Spatiotemporal variations in the b-value of earthquake magnitude–frequency distributions: Classification and causes. *Tectonophysics*, 615-616: 1–11, March 2014. ISSN 00401951. doi: 10.1016/j.tecto.2013.12.001. URL <https://linkinghub.elsevier.com/retrieve/pii/S0040195113007063>.
- J. K. Gardner and Leon Knopoff. Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the seismological society of America*, 64(5):1363–1367, 1974. Publisher: Citeseer.
- GFZ Data Services. The SeisComP seismological software package, 2008. doi:10.5880/GFZ.2.4.2020.003.
- Kenny M. Graham, José A. Bayona, Asim M. Khawaja, Pablo Iturrieta, Francesco Serafini, Emanuele Biondini, David A. Rhoades, William H. Savran, Philip J. Maechling, Matthew C. Gerstenberger, Fabio Silva, and Maximilian J. Werner. New Features in the pyCSEP Toolkit for Earthquake Forecast Development and Evaluation. *Seismological Research Letters*, 95(6):3449–3463, November 2024. ISSN 0895-0695, 1938-2057. doi: 10.1785/0220240197. URL <https://pubs.geoscienceworld.org/srl/article/95/6/3449/648802/New-Features-in-the-pyCSEP-Toolkit-for-Earthquake>.
- Laura Gulia and Paolo Gasperini. Contamination of Frequency–Magnitude Slope (b-Value) by Quarry Blasts: An Example for Italy. *Seismological Research Letters*, 92(6):3538–3551, November 2021. ISSN 0895-0695, 1938-2057. doi: 10.1785/0220210080. URL <https://pubs.geoscienceworld.org/srl/article/92/6/3538/605744/Contamination-of-Frequency-Magnitude-Slope-b-Value>.

- B. Gutenberg and C. F. Richter. Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4):185–188, October 1944. ISSN 0037-1106. URL <https://resolver.caltech.edu/CaltechAUTHORS:20140731-150249818>. Number: 4 Publisher: Seismological Society of America.
- Charles R. Harris, K. Jarrod Millman, Stéfan J. Van Der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, and Nathaniel J. Smith. Array programming with NumPy. *Nature*, 585(7825):357–362, 2020. URL <https://www.nature.com/articles/s41586-020-2649-2>. Publisher: Nature Publishing Group UK London.
- A. Helmstetter. Comparison of Short-Term and Time-Independent Earthquake Forecast Models for Southern California. *Bulletin of the Seismological Society of America*, 96(1):90–106, February 2006. ISSN 0037-1106. doi: 10.1785/0120050067. URL <https://pubs.geoscienceworld.org/bssa/article/96/1/90-106/146804>.
- Marcus Herrmann and Warner Marzocchi. Inconsistencies and Lurking Pitfalls in the Magnitude–Frequency Distribution of High-Resolution Earthquake Catalogs. *Seismological Research Letters*, 92(2A):909–922, March 2021. ISSN 0895-0695, 1938-2057. doi: 10.1785/0220200337. URL <https://pubs.geoscienceworld.org/ssa/srl/article/92/2A/909/593280/Inconsistencies-and-Lurking-Pitfalls-in-the>.
- John D. Hunter. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3):90–95, 2007. ISSN 1521-9615. doi: 10.1109/MCSE.2007.55. URL <http://ieeexplore.ieee.org/document/4160265/>.
- M. Ishimoto and K. Iida. Observations sur les séismes enregistrés par le micro-séismographe construit dernièrement. *Bulletin of the Earthquake Research Institute*, 17(2):443–478, June 1939. URL <https://repository.dl.itc.u-tokyo.ac.jp/records/34504>. (In Japanese with french abstract).

- Y. Y. Kagan. Short-Term Properties of Earthquake Catalogs and Models of Earthquake Source. *Bulletin of the Seismological Society of America*, 94(4):1207–1228, August 2004. ISSN 0037-1106. doi: 10.1785/012003098. URL <https://pubs.geoscienceworld.org/bssa/article/94/4/1207-1228/120955>.
- Yan Y. Kagan. Earthquake spatial distribution: the correlation dimension. *Geophysical Journal International*, 168(3):1175–1194, March 2007. ISSN 0956540X, 1365246X. doi: 10.1111/j.1365-246X.2006.03251.x. URL <https://academic.oup.com/gji/article-lookup/doi/10.1111/j.1365-246X.2006.03251.x>.
- Hubert W. Lilliefors. On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown. *Journal of the American Statistical Association*, 64(325):387–389, March 1969. ISSN 0162-1459, 1537-274X. doi: 10.1080/01621459.1969.10500983. URL <http://www.tandfonline.com/doi/abs/10.1080/01621459.1969.10500983>.
- E. Lippiello and G. Petrillo. b-More-Incomplete and b-More-Positive: Insights on a Robust Estimator of Magnitude Distribution. *Journal of Geophysical Research: Solid Earth*, 129(2):e2023JB027849, 2024. ISSN 2169-9356. doi: 10.1029/2023JB027849. URL <https://onlinelibrary.wiley.com/doi/abs/10.1029/2023JB027849>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2023JB027849>.
- W. Marzocchi and L. Sandri. A review and new insights on the estimation of the b-value and its uncertainty. *Annals Of Geophysics*, 46(6):1271–1282, 2003. URL <GotoISI>://000221367700007. Type: Journal article.
- W Marzocchi, I Spassiani, A Stallone, and M Taroni. How to be fooled searching for significant variations of the b-value. *Geophysical Journal International*, 220(3):1845–1856, March 2020. ISSN 0956-540X, 1365-246X. doi: 10.1093/gji/ggz541. URL <https://academic.oup.com/gji/article/220/3/1845/5647355>.

- Met Office. Cartopy: a cartographic python library with a Matplotlib interface, 2010. URL <https://scitools.org.uk/cartopy>.
- Arnaud Mignan and Jochen Woessner. Estimating the magnitude of completeness for earthquake catalogs. *Community online resource for statistical seismicity analysis*, pages 1–45, 2012. URL <https://episodesplatform.eu/eprints/1587/>. Publisher: Community Online Resource for Statistical Seismicity Analysis.
- Aron Mirwald, Leila Mizrahi, and Stefan Wiemer. How to b -Significant When Analyzing b -Value Variations. *Seismological Research Letters*, September 2024. ISSN 0895-0695, 1938-2057. doi: 10.1785/0220240190. URL <https://pubs.geoscienceworld.org/srl/article/doi/10.1785/0220240190/648709/How-to-b-Significant-When-Analyzing-b-Value>.
- Leila Mizrahi, Shyam Nandan, and Stefan Wiemer. The Effect of Declustering on the Size Distribution of Mainshocks. *Seismological Research Letters*, 92(4):2333–2342, July 2021. ISSN 0895-0695, 1938-2057. doi: 10.1785/0220200231. URL <https://pubs.geoscienceworld.org/srl/article/92/4/2333/594794/The-Effect-of-Declustering-on-the-Size>.
- Yosihiko Ogata and Koichi Katsura. Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues. *Geophysical Journal International*, 113(3):727–738, June 1993. ISSN 0956540X, 1365246X. doi: 10.1111/j.1365-246X.1993.tb04663.x. URL <https://academic.oup.com/gji/article-lookup/doi/10.1111/j.1365-246X.1993.tb04663.x>.
- M. Pagani, D. Monelli, G. Weatherill, L. Danciu, H. Crowley, V. Silva, P. Henshaw, L. Butler, M. Nastasi, L. Panzeri, M. Simionato, and D. Vigano. OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model. *Seismological Research*

Letters, 85(3):692–702, May 2014. ISSN 0895-0695, 1938-2057. doi: 10.1785/0220130087. URL <https://pubs.geoscienceworld.org/srl/article/85/3/692-702/315386>.

Vanille A. Ritz, Antonio P. Rinaldi, and Stefan Wiemer. Transient evolution of the relative size distribution of earthquakes as a risk indicator for induced seismicity. *Communications Earth & Environment*, 3(1):249, October 2022. ISSN 2662-4435. doi: 10.1038/s43247-022-00581-9. URL <https://www.nature.com/articles/s43247-022-00581-9>.

Paul A. Rydelek and I. Selwyn Sacks. Testing the completeness of earthquake catalogues and the hypothesis of self-similarity. *Nature*, 337(6204):251–253, January 1989. ISSN 0028-0836, 1476-4687. doi: 10.1038/337251a0. URL <https://www.nature.com/articles/337251a0>.

William H. Savran, José A. Bayona, Pablo Iturrieta, Khawaja M. Asim, Han Bao, Kirsty Bayliss, Marcus Herrmann, Danijel Schorlemmer, Philip J. Maechling, and Maximilian J. Werner. pyCSEP: A Python Toolkit for Earthquake Forecast Developers. *Seismological Research Letters*, 93(5):2858–2870, September 2022. ISSN 0895-0695, 1938-2057. doi: 10.1785/0220220033. URL <https://pubs.geoscienceworld.org/srl/article/93/5/2858/615824/pyCSEP-A-Python-Toolkit-for-Earthquake-Forecast>.

D. P. Schaff and F. Waldhauser. One Magnitude Unit Reduction in Detection Threshold by Cross Correlation Applied to Parkfield (California) and China Seismicity. *Bulletin of the Seismological Society of America*, 100(6):3224–3238, December 2010. ISSN 0037-1106. doi: 10.1785/0120100042. URL <https://pubs.geoscienceworld.org/bssa/article/100/6/3224-3238/325368>.

D. Schorlemmer and J. Woessner. Probability of Detecting an Earthquake. *Bulletin of the Seismological Society of America*, 98(5):2103–2117, October 2008. ISSN 0037-1106. doi: 10.1785/0120070105. URL <https://pubs.geoscienceworld.org/bssa/article/98/5/2103-2117/341891>.

- Skipper Seabold and Josef Perktold. Statsmodels: Econometric and Statistical Modeling with Python. pages 92–96, Austin, Texas, 2010. doi: 10.25080/Majora-92bf1922-011. URL <https://doi.curvenote.com/10.25080/Majora-92bf1922-011>.
- Sean Gillies, Casper van der Wel, Joris Van den Bosche, Mike W. Taves, Joshua Arnott, Brendan C. Ward, and Others. Shapely: Manipulation and analysis of geometric objects in the Cartesian plane, December 2022.
- Y. Shi and B. A. Bolt. The standard error of the magnitude-frequency b value. *Bulletin of the Seismological Society of America*, 72:1677–1687, 1982. Type: Journal article.
- Swiss Seismological Service. National Seismic Networks of Switzerland, 1983. Seismic Network, doi: 10.12686/SED/NETWORKS/CH.
- The pandas development team. pandas-dev/pandas: Pandas, 2020.
- S. Tinti and F. Mulargia. Confidence-intervals of B -values for grouped magnitudes. *Bulletin of the Seismological Society of America*, 77(6):2125–2134, 1987. ISSN 0037-1106. URL <GotoISI>://WOS:A1987K990500014.
- T. Tormann, S. Wiemer, and A. Mignan. Systematic survey of high-resolution b value imaging along Californian faults: Inference on asperities. *Journal of Geophysical Research: Solid Earth*, 119(3):2029–2054, March 2014. ISSN 2169-9313, 2169-9356. doi: 10.1002/2013JB010867. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/2013JB010867>.
- Tokuji Utsu. A Method for Determining the Value of b in a Formula $\log n = bM$ showing the Magnitude-Frequency Relation for Earthquakes. *Geophysical bulletin of the Hokkaido University*, 13:99 – 103, February 1965. doi: 10.14943/gbhu.13.99. URL <https://doi.org/10.14943/gbhu.13.99>.

- Tokuji Utsu, Yosihiko Ogata, Ritsuko S, and Matsu'ura. The Centenary of the Omori Formula for a Decay Law of Aftershock Activity. *Journal of Physics of the Earth*, 43(1):1–33, 1995. ISSN 1884-2305, 0022-3743. doi: 10.4294/jpe1952.43.1. URL http://www.jstage.jst.go.jp/article/jpe1952/43/1/43_1_1/_article.
- Nicholas J. Van Der Elst. *B-Positive* : A Robust Estimator of Aftershock Magnitude Distribution in Transiently Incomplete Catalogs. *Journal of Geophysical Research: Solid Earth*, 126(2):e2020JB021027, February 2021. ISSN 2169-9313, 2169-9356. doi: 10.1029/2020JB021027. URL <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020JB021027>.
- Nicholas J. Van Der Elst and Morgan T. Page. *a -Positive*: A Robust Estimator of the Earthquake Rate in Incomplete or Saturated Catalogs. *Journal of Geophysical Research: Solid Earth*, 128(10):e2023JB027089, October 2023. ISSN 2169-9313, 2169-9356. doi: 10.1029/2023JB027089. URL <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2023JB027089>.
- Virtanen, Pauli, Gommers, Ralf, Oliphant, Travis E., Haberland, Matt, Reddy, Tyler, Cournapeau, David, Burovski, Evgeni, Peterson, Pearu, Weckesser, Warren, Bright, Jonathan, van der Walt, Stéfan J., Brett, Matthew, Wilson, Joshua, Millman, K. Jarrod, Mayorov, Nikolay, Nelson, Andrew R. J., Jones, Eric, Kern, Robert, Larson, Eric, Carey, C J, Polat, İlhan, Feng, Yu, Moore, Eric W., VanderPlas, Jake, Laxalde, Denis, Perktold, Josef, Cimrman, Robert, Henriksen, Ian, Quintero, E. A., Harris, Charles R., Archibald, Anne M., Ribeiro, Antônio H., Pedregosa, Fabian, {van Mulbregt}, Paul, and {SciPy 1.0 Contributors}. *{{scipy}} 1.0: fundamental algorithms for scientific computing in python*. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2. URL <https://rdcu.be/b08Wh>.
- S. Wiemer. A Software Package to Analyze Seismicity: ZMAP. *Seismological Research*

Letters, 72(3):373–382, May 2001. ISSN 0895-0695. doi: 10.1785/gssrl.72.3.373. URL <https://pubs.geoscienceworld.org/srl/article/72/3/373-382/142637>.

S. Wiemer and M. Wyss. Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. *Bulletin of the Seismological Society of America*, 90(4):859–869, 2000. ISSN 0037-1106. doi: Doi10.1785/0119990114. URL <GotoISI>://WOS:000089060700005. Type: Journal article.

J. Woessner and S. Wiemer. 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, 2005. ISSN 0037-1106. doi: Doi10.1785/0120040007. URL <GotoISI>://WOS:000229028400022. Type: Journal article.

Jack Woollam, Jannes Münchmeyer, Frederik Tilmann, Andreas Rietbrock, Dietrich Lange, Thomas Bornstein, Tobias Diehl, Carlo Giunchi, Florian Haslinger, and Dario Jozinović. SeisBench—A Toolbox for Machine Learning in Seismology. 93(3), 2022.

List of Figures

- 1 An example of an earthquake catalog: The Swiss earthquake catalog for the year 2024 (Swiss Seismological Service, 1983). Figure generated using the SeismoStats function `plot_in_space`. Alt-text: Swiss map with dots whose size corresponds to the earthquake magnitude. 3
- 2 Overview of the SeismoStats package. a) Organisation of the subpackages with selected functions. b) Detailed architecture of the Catalog class in which methods can both use and update the catalog’s properties and data. Alt-text: Schematic diagram of the structure of SeismoStats. 5

3	Visualization of the application of the three completeness estimation methods to the Swiss catalog in the year 2024. The blue squares represent the cumulative FMD, the red triangles the non-cumulative one. The vertical dashed lines indicate the completeness magnitude that was estimated using the different available methods. Figure generated using SeismoStats functions <code>plot_fmd</code> and <code>plot_cum_fmd</code> . Alt-text: The Frequency of earthquakes in relation to their magnitude is shown as scatter plot, together with dotted lines that the result of three different completeness methods.	10
4	Example of an estimated b-value fit. The estimator <code>ClassicBValueEstimator</code> was applied, with a bin size of $\Delta m = 0.01$ and assuming an incompleteness of $m_c = 1.6$. The GR-law with the estimated b-value is shown in black, the corresponding cumulative FMD is shown in blue. The dotted line is the extension of the GR to the reference magnitude $M=1$. Figure generated using SeismoStats function <code>plot_cum_fmd</code> . Alt-text: Plot of the frequency vs the magnitude of earthquakes, together with a line that fits the slope.	16
5	Illustration of temporal variation of the b-value both for the classical b-value estimate and b-positive using the Swiss earthquake catalog for 2024. On the top panel, the magnitudes over time are shown, figure generated using SeismoStats function <code>plot_mag_in_time</code> . On the bottom panel, the classical b-value and b-positive over time, each b-value is estimated using 100 magnitudes above m_c . The b-value plotted at each time is estimated with magnitudes up to that time. This figure was generated using SeismoStats function <code>plot_b_series_constant_nm</code> . Alt-text: (a): scatter plot of magnitudes vs. time, with a dotted line indicating the completeness of 1.1. (b): b-value as a function of time, estimated with the classical and the b-positive method. . .	17

List of Tables