

Smoothing seismicity techniques applied to seismic sources characterization and probabilistic seismic hazard analysis in Brazil

Marlon Pirchiner^{1,2}

Corresponding author: Marlon Pirchiner, Department of Geophysics, IAG, USP, São Paulo, Brazil. (marlon.pirchiner@iag.usp.br)

¹Department of Geophysics, Institute of Astronomy, Geophysics and Atmospherical Sciences (IAG), University of São Paulo (USP), São Paulo, Brazil.

²Applied Math School (EMAp), Getúlio Vargas Foundation (FGV), Rio de Janeiro, Brazil.

Intraplate seismicity is a hard problem to solve in tectonics, mainly during the lack of observed seismicity in many areas where seismic hazard assessment are also difficult and the methods applied are often close correlated to the geometries defined by the experts with valuable criteria.

Three smoothing seismicity techniques are reviewed in order to investigate their results when applied over the Brazilian seismicity and suggesting a seismic hazard assesment.

All this methods are based on a kernel density estimation. The main difference between them is the bandwidth selection process.

Frankel in their method uses a *fixed* bandwidth in space. Woo's method suggest a bandwidth as function of the *magnitude*. Helmstetter and Werner proposes a *local-adaptive* in space and time kernel bandwidths, chosen for each event. This last one was proposed initialy to take care of the background seismicity rateon long-term earthquake forecast context and the time-independent assumption about the annual background seismic rate could be used (not indiscriminately REFERENCE) as a grid of points seismic sources.

As conclusions the results shown that all methods gives defines better some known seismic regions and could be used in the Brazilian seismicity case, all of them, improoving the spatial resolution of the GSHAP model, the last one made on the region.

The Last two methods was implemented and are available as a free software on the Hazard Modeling Toolkit (HMTK) from GEM Foundation.

All the hazard calculations was performed on OpenQuake open source software.

1. Introduction

Brazil have no recent seismic hazard study at national scale despite some important and localized recent studies focused mainly on specific buildings.

The Brazilian seismic hazard building code was made based on GSHAP model which is the well accepted state of art model for the whole region.

A significant number of earthquakes has been recorded in the last years, mainly after the national seismic network (RSBR) operation, and an actualized hazard map whould be computed.

Although the standard Cornell McGuire zoning methodolgy for the classical hazard seismic assesment, the goal is review a few smoothing seismicity methods based on kernel density estimation. This estimation allow a seismic point sources grid definition. The seismic rate (a -value) associated to each source is definided by the smoothed seismicity methods, but the more parameters like b -values, minimum and maximum magnitudes, rupture and depth distributions must to be assigned in other ways.

In the kernel density estimation theory two things there always be handled. The kernel shape and bandwidth. Almost kernels used are gaussian or power-laws in one of their variants. The mainly distinction between methods decribed is the proposed bandwidth selection process.

Frankel smoothing method applies a gaussian Nadaraya-Watson smoothing method with a *fixed* smoothing-distance bandwidth to estimate seismicity rate. Gordon Woo proposes an *magnitude-dependent* kernel bandwidth based on the nearest neighbour mean distance in magnitude bins. Helmstetter proposes a *local-adaptive* bandwidth, based on kernel estimations on both space and time dimentions, estimation of the background seismicity rate for long-term forecasts. Using one-year forecast and assume

time-independence, the same assumption of PSHA, of seismicity rate is possible to use these values on seismic sources characterization.

To perform the hazard computation, the OpenQuake suite was choosed. The modeller environment used was the HMTK also available by GEM scientifical board. The Frankel method was already implemented on the toolkit and the other two was implemented in the context of this paper and they are free available on a public repository.

The earthquake catalog data comes from the Brazilian seismological research authorities.

The Brazilian seismicity data is presented including some overview plots. Next the method's theory is presented follow by the specific decisions and optimizations performed on each method modelling. And last, the results, conclusions and further considerations are discussed.

1.1. Intraplate Seismicity

1.2. Previous Brazilian Studies

1.3. Previous Smoothing Seismicity (Frankel, Woo, Helmstetter...)

1.4. Specifical Purpose

2. Brazilian seismicity data

The main source of brazilian intraplate seismicity data is current the Brazilian Seismic Bulletin maintained mainly by the Seismological Certer, Institute of Astronomy, Ge-

physics and Atmospheric Sciences, University of São Paulo [BSB, 2014]. This catalog is a compilation from different sources.

Until 1981, the catalog corresponds to *Berrocal et al.* [1984] which is a huge compilation effort from historic and instrumental earthquake data with participation of Universidade de São Paulo (USP), Universidade de Brasilia (UnB), Universidade Federal do Rio Grande do Norte (UFRN), Observatório Nacional (ON), Universidade Estadual de São Paulo (UNESP) and Instituto de Pesquisas Tecnológicas do Estado de São Paulo (IPT). From 1982 to 1995 the information comes from Revista Brasileira de Geofísica. Since 1995, the Bulletin is maintained mainly by University of São Paulo under the same cooperation network. Earlier versions of this catalog was used to compose the well-known *CERESIS* [1985, 1995] catalogs. The number of earthquakes on the catalog

Today the catalog is distributed in two ways. One could be called *raw* and contains all compiled information even events outside country which effects was felt in Brazil, and other could be called *clean*, since that events with high error and low quality was removed.

This catalog only cover crustal events. Andean deep subduction earthquakes with depth higher than 50km at west was discarded. This deep earthquakes contributions to the ground shaking hope to be consider in the future based on some new comprehensive south-american catalog.

Almost magnitude values was computed as *mb* type or equivalent *mR* [*Assumpção, 1983*]. Part of them was computed from macroseismic felt area data [*Berrocal et al., 1984*]. Since almost GMPE are based on *M_W* moment scale magnitude, on this present study, a new version of the catalog with *M_W* values as delivered. *Scordilis* [2006] was discarded by its magnitude range definition, over than 6, whilst almost magnitude values on *BSB* [2014] catalog have magnitude values lower than that. On this context the

conversion was made following the guidelines on table 1 from an *ad-hoc* communication distributed with the catalog.

3. Smoothing methods

Under this work, three smoothing methods was considered. On this section brief explanation of each one will be done.

3.1. Frankel

Arthur *Frankel* [1995] smoothing seismicity proposal consists originally in use a called correlation distance d_F as the kernel *fixed* bandwidth and next to apply the Nadaraya-Watson [Nadaraya, 1964; Watson, 1964] estimator to smooth a 2D seismicity histogram using a gaussian kernel:

$$\tilde{n}_j = \frac{\sum_i n_i e^{-\left(\frac{d_{ij}}{d_F}\right)^2}}{\sum_i e^{-\left(\frac{d_{ij}}{d_F}\right)^2}}, \quad (1)$$

where \tilde{n}_j is the smoothed seismicity (number of earthquakes with magnitude m above the minimum magnitude M_d in the catalog) on cell j . n_i is the earthquake counting in each other cell i and d_{ij} is distance between grid cells i and j .

3.2. Woo

Gordon *Woo* [1996] suggested to evaluate the contribution of each earthquake i , far \mathbf{r}_i from its kernel, on the cell centered on \mathbf{r} in dependence of their magnitude m :

$$R(\mathbf{r}, m) = \sum_{i=1}^N \frac{K(\mathbf{r} - \mathbf{r}_i, m)}{T(\mathbf{r}_i)}, \quad (2)$$

where N is the number of earthquakes i on the catalog and $T(\mathbf{r}_i)$ is the timeframe which all earthquakes with magnitude above m is completely observed on \mathbf{r}_i .

Any kernel could be applied on that definition. In practice Woo used a *Kagan and Knopoff* [1980] for a infinity spatial domain:

$$K_{KK}(\mathbf{r}, m | a_W) = \frac{a_W - 1}{\pi h(m)^2} \left[1 + \frac{\mathbf{r}^2}{h(m)^2} \right]^{-a_W}, \quad (3)$$

where a_W is, accordingly *Vere-Jones* [1992], fractal dimension factor, generally about 1.5 and 2.

To compute the magnitude-dependent bandwidth function $h(m)$, Gordon Woo suggested the follow relation

$$h(m|a_0, a_1) = a_0 e^{a_1 m}, \quad (4)$$

where a_0 and a_1 are computed by the regression of the mean nearest distance h from each magnitude bin $m \pm dm$.

3.3. Helmstetter and Werner

Even work on a forecast perspective, the background seismicity rate for a long-term time-independent forecast could be used to characterize a diffused seismicity under the common assumption that this background seismic rate is invariant over time.

Helmstetter and Werner [2012] proposes a seismicity model space and time dependent using kernels for space and time independently:

$$R(\mathbf{r}, t) = \sum_{i=1}^N \frac{1}{h_i d_i^2} K_t \left(\frac{t - t_i}{h_i} \right) K_r \left(\frac{\|\mathbf{r} - \mathbf{r}_i\|}{d_i} \right), \quad (5)$$

where $R(\mathbf{r}, t)$ is the taxa de sismicidade na localização \mathbf{r} e no instante t , K_t is the função de núcleo na dimensão do tempo, onde t_i é a time location of earthquake i e h_i é a temporal bandwidth to earthquake i , K_r is the função de núcleo na dimensão do espaço, onde \mathbf{r}_i é a spatial location of earthquake i e d_i é a spatial bandwidth to earthquake i .

As they are interested in forecast, just past time $t_i < t$ need to be considered. Also the observation completeness is taken in account by a set of weights w in this follow way:

$$R(\mathbf{r}, t) = R_{min} + \sum_{t_i < t} \frac{2 w(\mathbf{r}_i, t_i)}{h_i d_i^2} K_t \left(\frac{t - t_i}{h_i} \right) K_r \left(\frac{\|\mathbf{r} - \mathbf{r}_i\|}{d_i} \right), \quad (6)$$

where R_{min} is the minimum seismic rate, positive, allowing earthquakes to occur where their never occurred yet.

The weights $w(\mathbf{r}_i, t_i)$, computed for each earthquake i , are the Gutenberg-Richter a -value projection. These weights increase the seismicity contribution from earthquakes occurred where and when the completeness magnitude M_c was greater than minimum catalog magnitude M_d . These weights could easily consider space and time completeness and b -value fluctuations. The expression used to compute it:

$$w(\mathbf{r}, t) = 10^{b(\mathbf{r}, t)[M_c(\mathbf{r}, t) - M_d]}, \quad (7)$$

where w is the weight at location \mathbf{r} on instant t , $b(\mathbf{r}, t)$ is the space and time dependent b -value, $M_c(\mathbf{r}, t)$ is the completeness magnitude on \mathbf{r} and t , M_d is the minimum magnitude value on the catalogue.

3.3.1. Local bandwidth computation

The method implemented to compute the space and time bandwidths for each earthquake was the Coupled-Nearest-Neighbour [Helmstetter and Werner, 2012]:

$$h_i, d_i = \arg \min_{\substack{h_i \geq h_k \\ d_i \leq d_k}} [s(h_i, d_i | k_{cnn}, a_{cnn}) := h_i + a_{cnn}d_i], \quad (8)$$

where k_{cnn} is the k^{th} nearest neighbour, a_{cnn} is a space-time coupling factor, d_k is the $\max \{d_j\}, j = 1, \dots, k_{cnn}$ e h_k is the $\max \{h_j\}, j = 1, \dots, k_{cnn}$.

The bandwidths are defined locally. Could be small in high earthquake density regions and higher where earthquakes are rarely or regions with information lack.

3.3.2. Stationary seismic rate

After compute the model parameters and completely define it, the time-independent or stationary seismic rate \bar{R} on each location \mathbf{r}_0 have to be computed. The way to do it proposed by Helmstetter and Werner [2012] was to get the median on \mathbf{r}_0 over all

considered time window:

$$\bar{R}(\mathbf{r}_0) = \text{Median}[R(\mathbf{r}_0, t)]. \quad (9)$$

The median should avoid the seismicity rate fluctuations derived by fore and aftershocks and consequently the decluster procedures.

3.3.3. Maximum likelihood optimization

The model is completely defined by k_{cnn} , a_{cnn} and R_{min} . To optimize these parameters the catalog is divided in two parts: one for learning about the parameters and other to test them.

If the earthquake occurrence could be modelled by a Poisson process with rate N_p , then the probability to observe exactly n events on the consider time frame is

$$p(N_p, n) = \frac{N_p^n e^{-N_p}}{n!}. \quad (10)$$

Then, over all cells, the (log) likelihood, to be maximized, between model (from the learning catalog) prediction and earthquakes observed on testing catalog is written as

$$L = \sum_{i_x=1}^{N_x} \sum_{i_y=1}^{N_y} \log p[N_p(i_x, i_y), n(i_x, i_y)] \quad (11)$$

where $N_p(i_x, i_y)$ is the predicted seismic rate on cell (i_x, i_y) and $n(i_x, i_y)$ is the number of target earthquakes observed on cell (i_x, i_y) .

The model parameters R_{min} , a_{cnn} and k_{cnn} should be optimized by the maximization of log-likelihood L .

4. Results

Referencing equation. Mauris vel lorem magna, tristique auctor ipsum. Aliquam pharetra eleifend massa. Donec porttitor sagittis luctus. Aliquam pretium luctus leo quis congue. Morbi vel felis mi. Referencing Table 3. Referencing Figure 1.

4.1. Simulations

4.1.1. Simulation 1

Vivamus magna enim, aliquet id cursus a, pharetra ut purus. Phasellus suscipit nisi iaculis mi vulputate id interdum velit dictum. Nam ullamcorper elit in lectus ultrices vitae volutpat massa gravida. Etiam sagittis commodo neque eget placerat. Sed et nisi faucibus metus interdum adipiscing id nec lacus. Donec ipsum diam, malesuada at euismod consectetur, placerat quis diam. Phasellus cursus semper viverra. Proin magna tortor, blandit in ultricies id, facilisis at nibh. Proin eu neque est. Etiam euismod auctor ante. Mauris mauris sem, tincidunt a placerat rutrum, porta id est. Aenean non velit porta eros condimentum facilisis at in nibh. Etiam cursus purus ut orci rhoncus sit amet semper eros porttitor. Etiam ac leo at ipsum tincidunt consequat ac non sapien. Aenean sed leo diam, venenatis pharetra odio.

4.1.2. Simulation 2

Suspendisse viverra eleifend nulla at facilisis. Nullam eget tellus orci. Cras sit amet lorem velit. Maecenas rhoncus pellentesque orci eget vulputate. Phasellus massa nisi, mattis nec elementum accumsan, blandit non neque. In ac enim elit, sit amet luctus ante. Cras feugiat commodo lectus, vitae convallis dui sagittis id. In in tellus lacus, sed lobortis eros. Phasellus sit amet eleifend velit. Duis ornare dapibus porttitor. Maecenas eros velit, dignissim at egestas in, tincidunt lacinia erat. Proin elementum mi vel lectus suscipit fringilla. Mauris justo est, ullamcorper in rutrum interdum, accumsan eget mi. Maecenas ut massa aliquet purus eleifend vehicula in a nisi. Fusce molestie cursus lacinia.

4.2. Real Data

Aliquam interdum pellentesque scelerisque. Sed tincidunt suscipit purus, id aliquet nulla vehicula quis. Duis sed nisl lorem. Vivamus erat ante, dignissim et aliquam vel, adipiscing vitae magna. Cras id dapibus metus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Proin ut lectus ut nisi congue ullamcorper. Ut ac turpis ligula. Sed faucibus bibendum nunc eget gravida.

5. Discussion

Nam fermentum sapien at enim varius consectetur. Quisque lobortis imperdiet mauris, et accumsan libero vulputate vitae. Integer lacinia purus vel metus tempus suscipit. Curabitur ac sapien quis mauris euismod commodo. Sed pharetra sem elit. Fusce ultrices, mauris eu fermentum tempor, tellus sem ornare lectus, in convallis nunc urna id dolor. Donec convallis ligula vitae sem viverra fermentum. Mauris in ullamcorper erat. Donec ultrices tempus nibh quis vestibulum.

Praesent volutpat, nibh in dignissim commodo, tellus justo consequat erat, vel consequat mi arcu vel lectus. Aliquam a tellus nec felis sagittis consequat. Quisque convallis imperdiet neque a tempor. Nulla non erat urna. Mauris vel lorem magna, tristique auctor ipsum. Aliquam pharetra eleifend massa. Donec porttitor sagittis luctus. Aliquam pretium luctus leo quis congue. Morbi vel felis mi. Suspendisse viverra tortor pretium orci lacinia eleifend. Phasellus aliquam, nunc eu cursus feugiat, erat odio porttitor libero, quis accumsan orci ipsum ut lorem. Vestibulum pharetra malesuada egestas. Sed non orci sit amet erat suscipit fringilla in et diam. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nunc ut rhoncus nulla. Aenean porta rhoncus suscipit.

Vivamus magna enim, aliquet id cursus a, pharetra ut purus. Phasellus suscipit nisi iaculis mi vulputate id interdum velit dictum. Nam ullamcorper elit in lectus ultrices vitae volutpat massa gravida. Etiam sagittis commodo neque eget placerat. Sed et nisi faucibus metus interdum adipiscing id nec lacus. Donec ipsum diam, malesuada at euismod consectetur, placerat quis diam. Phasellus cursus semper viverra. Proin magna tortor, blandit in ultricies id, facilisis at nibh. Proin eu neque est. Etiam euismod auctor ante. Mauris mauris sem, tincidunt a placerat rutrum, porta id est. Aenean non velit porta eros condimentum facilisis at in nibh. Etiam cursus purus ut orci rhoncus sit amet semper eros porttitor. Etiam ac leo at ipsum tincidunt consequat ac non sapien. Aenean sed leo diam, venenatis pharetra odio.

Suspendisse viverra eleifend nulla at facilisis. Nullam eget tellus orci. Cras sit amet lorem velit. Maecenas rhoncus pellentesque orci eget vulputate. Phasellus massa nisi, mattis nec elementum accumsan, blandit non neque. In ac enim elit, sit amet luctus ante. Cras feugiat commodo lectus, vitae convallis dui sagittis id. In in tellus lacus, sed lobortis eros. Phasellus sit amet eleifend velit. Duis ornare dapibus porttitor. Maecenas eros velit, dignissim at egestas in, tincidunt lacinia erat. Proin elementum mi vel lectus suscipit fringilla. Mauris justo est, ullamcorper in rutrum interdum, accumsan eget mi. Maecenas ut massa aliquet purus eleifend vehicula in a nisi. Fusce molestie cursus lacinia.

Appendix A: Appendix Title

Vivamus magna enim, aliquet id cursus a, pharetra ut purus. Phasellus suscipit nisi iaculis mi vulputate id interdum velit dictum. Nam ullamcorper elit in lectus ultrices vitae volutpat massa gravida. Etiam sagittis commodo neque eget placerat. Sed et nisi faucibus metus interdum adipiscing id nec lacus. Donec ipsum diam, malesuada at

euismod consectetur, placerat quis diam. Phasellus cursus semper viverra. Proin magna tortor, blandit in ultricies id, facilisis at nibh. Proin eu neque est. Etiam euismod auctor ante. Mauris mauris sem, tincidunt a placerat rutrum, porta id est. Aenean non velit porta eros condimentum facilisis at in nibh. Etiam cursus purus ut orci rhoncus sit amet semper eros porttitor. Etiam ac leo at ipsum tincidunt consequat ac non sapien. Aenean sed leo diam, venenatis pharetra odio.

Acknowledgments. The author thanks specially to Marcelo Assumpção for the support and critical discussion as well all IAG-USP seismological team. It is also grateful to Stéphane Drouet and Vincent Guigues for all helpful discussions.

References

- Assumpção, M. (1983), A regional magnitude scale for brazil, *Bulletin of the Seismological Society of America*, 73(1), 237–246.
- Berrocal, J., U. de São Paulo. Instituto Astronômico e Geofísico, and B. C. N. de Energia Nuclear (1984), *Sismicidade do Brasil*, Instituto Astronômico e Geofísico, Universidade de São Paulo.
- BSB (2014), Boletim sísmico brasileiro (BSB) version 2014.11. institute of astronomy, geophysics and atmospherical sciences, university of são paulo, iag-usp.
- CERESIS (1995), Catalogue for south america and the caribbean prepared in the framework of gshap, *Tech. rep.*, Centro Regional de Sismología para América del Sur.
- CERESIS, S. T. A., Bonny L. Askew (1985), Catalog of earthquakes for south america, *Tech. rep.*, CERESIS (Centro Regional de Sismología para América del Sur). Earthquake mitigation program in the Andean Region (SISRA project).
- Frankel, A. (1995), Mapping seismic hazard in the central and eastern united states, *Seismological Research Letters*, 66(4), 8–21.

Helmstetter, A., and M. J. Werner (2012), Adaptive spatiotemporal smoothing of seismicity for long-term earthquake forecasts in California, *Bulletin of the Seismological Society of America*, 102(6), 2518–2529.

Kagan, Y. Y., and L. Knopoff (1980), Spatial distribution of earthquakes: the two point correlation function, *Geophys. J. R. Astr. Soc.*, 62.

Nadaraya, E. A. (1964), On estimating regression, *Theory of Probability & Its Applications*, 9(1), 141–142, doi:10.1137/1109020.

Scordilis, E. (2006), Empirical global relations converting m s and m b to moment magnitude, *Journal of Seismology*, 10(2), 225–236, doi:10.1007/s10950-006-9012-4.

Vere-Jones, D. (1992), Statistical methods for the description and display of earthquake catalogs, *Statistics in the Environmental and Earth Sciences*.

Watson, G. S. (1964), Smooth regression analysis, *Sankhya: The Indian Journal of Statistics, Series A (1961-2002)*, 26(4), pp. 359–372.

Woo, G. (1996), Kernel estimation methods for seismic hazard area source modeling, *Bulletin of the Seismological Society of America*, 86(2), 353–362.

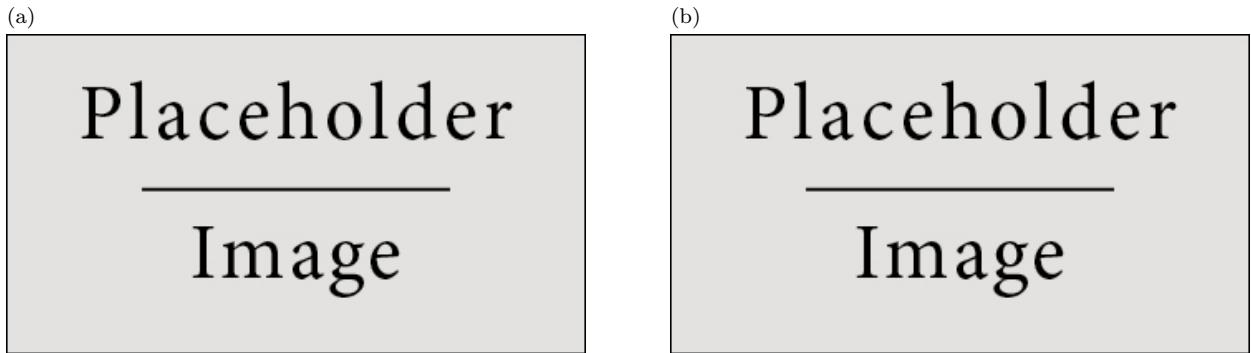


Figure 1. This tabled figure shows the first (a) image and the 2nd one (b). Even with the same bla ba.

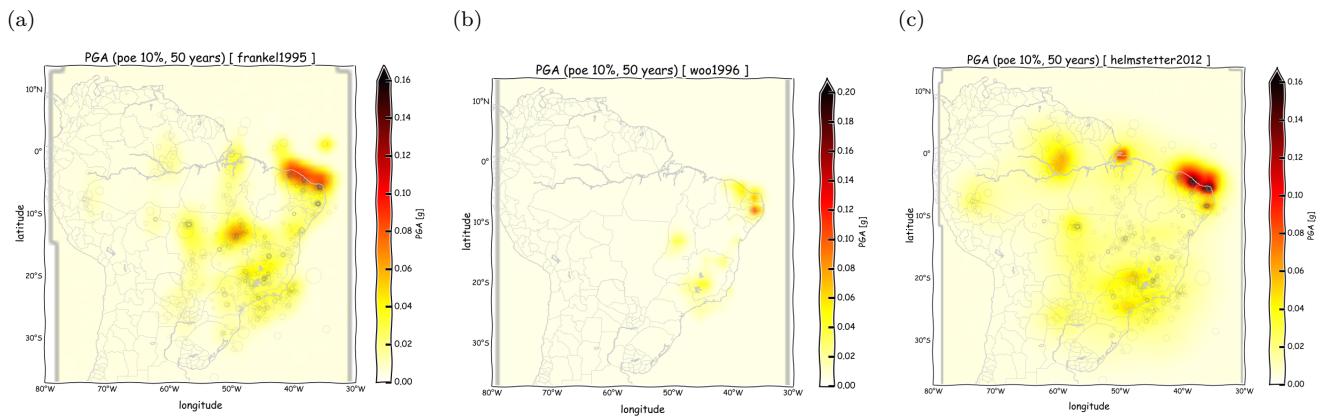


Figure 2. This tabled figure shows the first (a) image and the 2nd one (b). Even with the same bla ba.

Table 1. BSB 2014.11 M_W proxies *ad-hoc* guidelines.

magnitude source	rule	uncertainty
mb or mR	$M_W(m) = 1.21m - 0.76$	0.3
Area felt A_f	$M_W(A_f) = 0.81 + 0.639 \log(A_f) + 0.00084\sqrt{A_f}$	0.4
Maximum intensity I_0	$mb(I_0) = 1.21 + 0.45I_0$ then $M_W(m)$	0.6
mb and A_f	$M_W(m, A_f) = 0.7M_W(m) + 0.3M_W(A_f)$	0.33

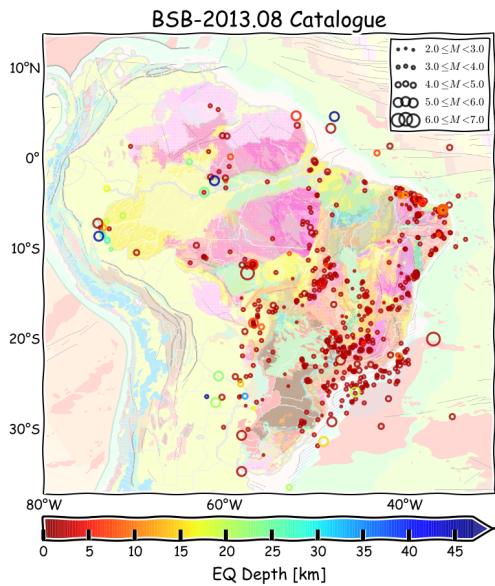


Figure 3. Brazilian earthquake catalog.

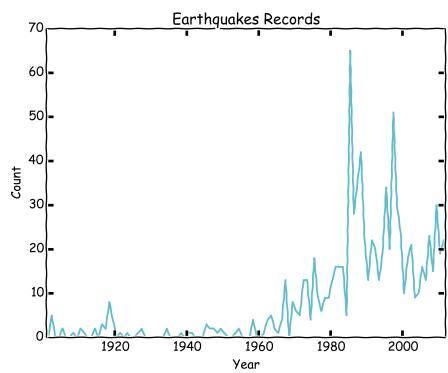


Figure 4. Earthquake records by year

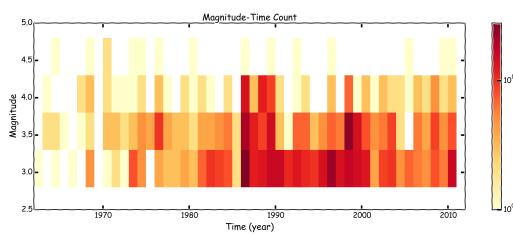


Figure 5. Magnitude count by year.

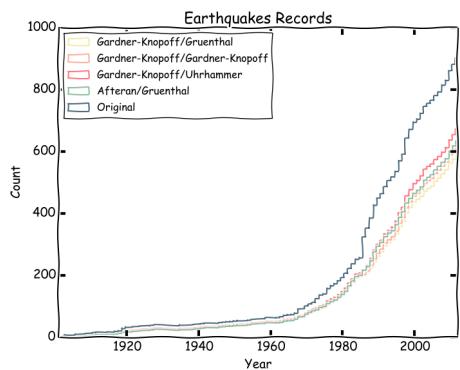


Figure 6. Catalog declustering.

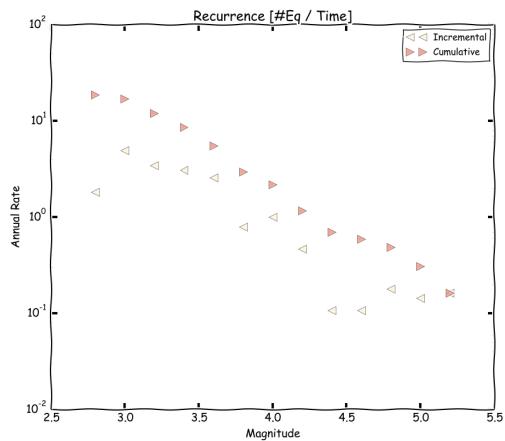


Figure 7. Annual earthquake recurrence.

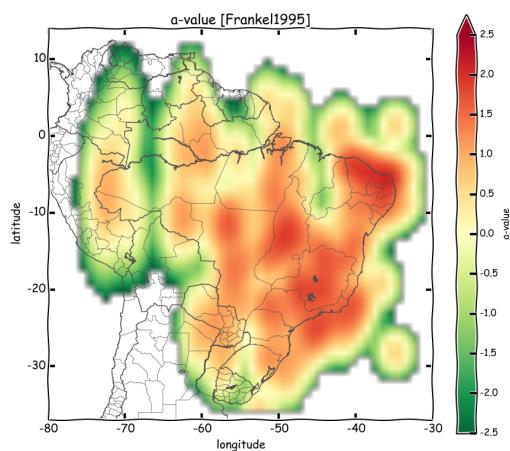


Figure 8. Smoothed seismicity: Frankel's method.

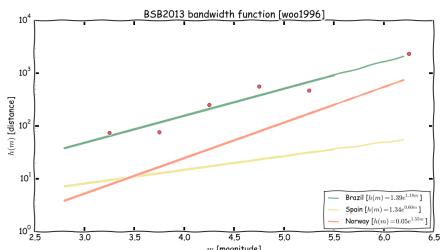


Figure 9. Magnitude dependence bandwidth.

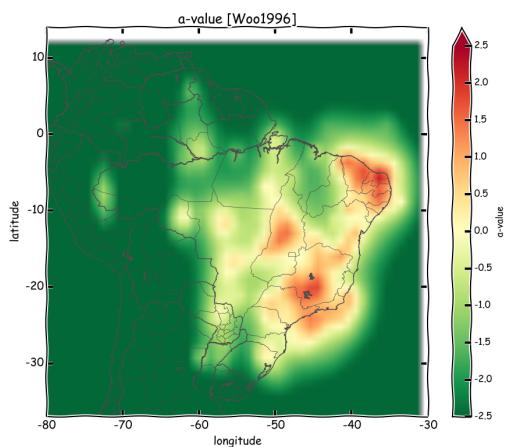


Figure 10. Smoothed seismicity: Woo's method.

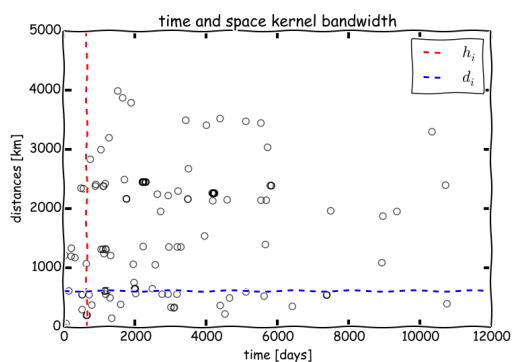


Figure 11. Local bandwidth example.

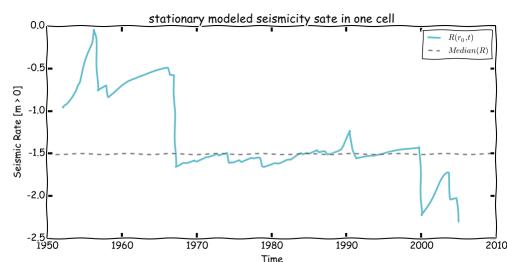


Figure 12. Stationary seismic rate

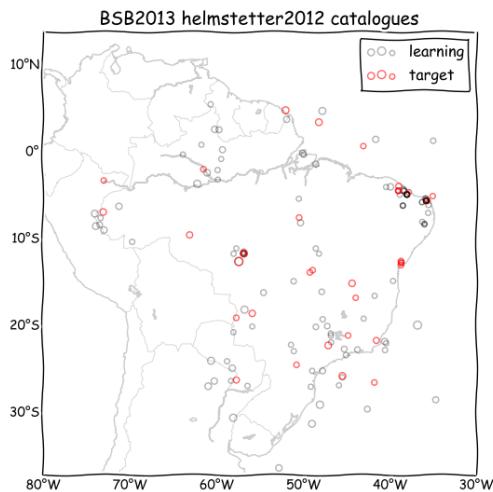


Figure 13. Learning and target catalogs.

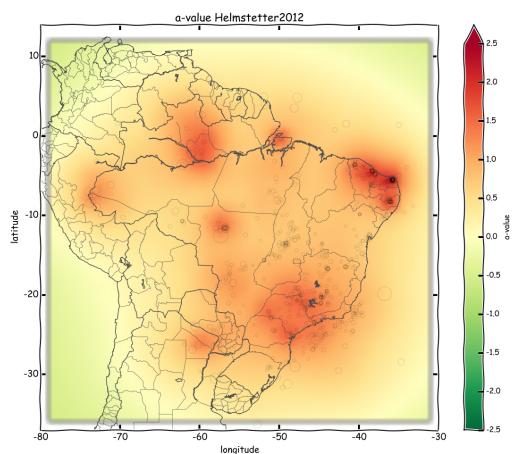


Figure 14. Smoothed seismicity: Helmstetter's method.

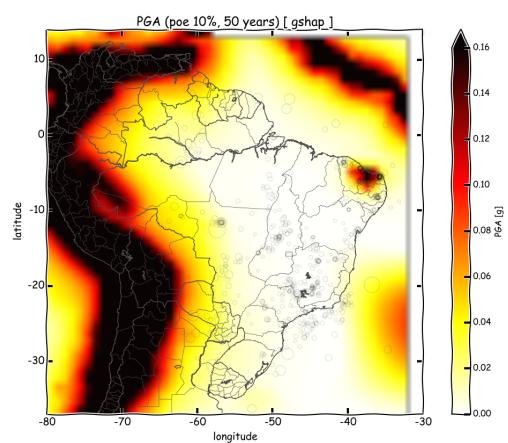


Figure 15. GSHAP results: PGA (poe 10%/50y)

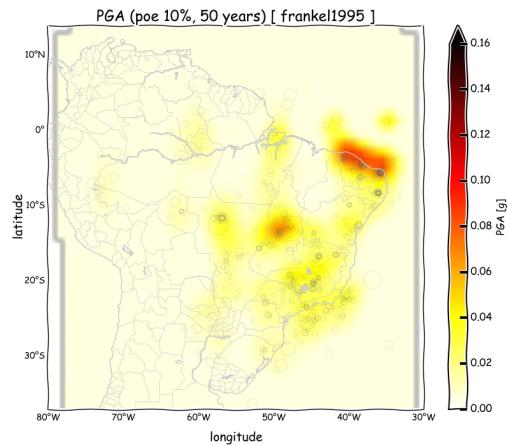


Figure 16. Frankel's results: PGA (poe 10%/50y)

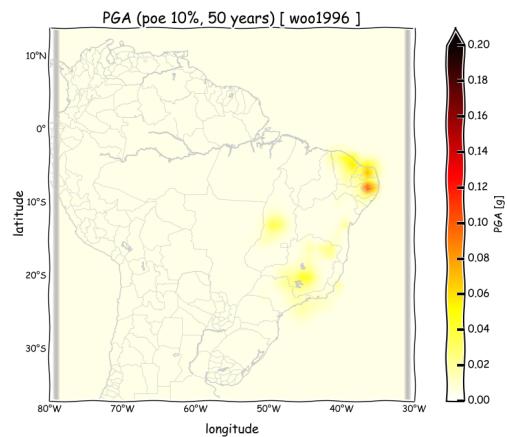


Figure 17. Woo's results: PGA (poe 10%/50y)

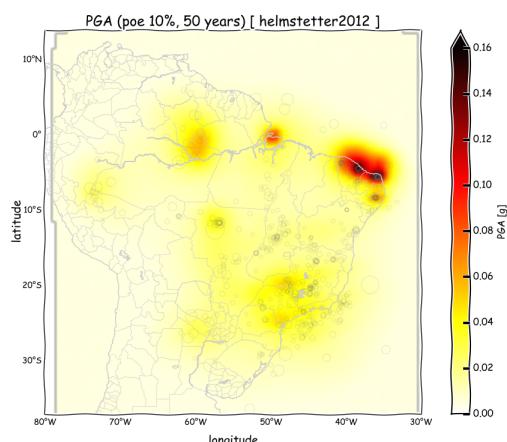


Figure 18. Helmstetter's results: PGA (poe 10%/50y)