Available online at www.sciencedirect.com
SCIENCE DIRECT®
TECTONOPHYSICS
Tectonophysics 413 (2006) 33–38
www.elsevier.com/locate/tecto

Discussion

Comment on 'Testing earthquake prediction methods: "The West Pacific short-term forecast of earthquakes with magnitude MwHRV ≥ 5.8"' by V.G. Kossobokov

Y.Y. Kagan *, D.D. Jackson

Department of Earth and Space Sciences, University of California, Los Angeles, California 90095-1567, USA

Received 21 March 2005; received in revised form 5 May 2005; accepted 4 October 2005
Available online 28 November 2005

Abstract

In his paper Kossobokov investigates the efficiency of our short-term forecast for two western Pacific regions. Although we agree with the basic results of his evaluation that the forecast statistics is much better than a random guess, we have reservations about his definition of earthquake prediction, some of his tests, and his interpretation of the test results. We distinguish between deterministic earthquake predictions and statistical forecasts. We argue that some techniques used by Kossobokov may not be appropriate for testing our forecasts and discuss other testing methods, based on the likelihood function. We demonstrate that Kossobokov's null hypothesis may be biased, and this bias can influence some of his conclusions. We show that contrary to Kossobokov's statement, our algorithm predicts mainshocks when they are preceded by foreshocks.
© 2005 Elsevier B.V. All rights reserved.

1. Introduction

In earlier papers (Jackson and Kagan, 1999; Kagan and Jackson, 2000), we introduced two earthquake forecast methods based on past seismicity, with an open invitation for others to test these forecasts on future earthquakes. Since then we have been updating these forecasts daily and posting results on the Web. We are grateful to V. G. Kossobokov for taking up our invitation. We have no fundamental disagreement with his Paper (Kossobokov, 2006—this issue, henceforth referred to as K2005), but we must point out a few considerations affecting its interpretation.

In our papers we rigorously distinguished between "forecasts" and "predictions." We defined statistical forecasts in terms of the rate-density, or probability density per unit time and area of earthquakes as a function of magnitude. Predictions which are usually specified as deterministic or quasi-deterministic statements, are much more specific, implying considerably higher probability and more certainty than a forecast. Our models explicitly produce forecasts, not predictions.

Earthquake occurrence is usually represented as a stochastic point process (Kagan and Jackson, 2000; Daley and Vere-Jones, 2003), and event occurrence rate is one of the most important parameters in the theory of point processes. In west Pacific regions we estimate the earthquake rate for cells $0.5\times0.5^\circ$. The daily rate values for $M_w \geq 5.8$ earthquakes usually are less than 0.1 eq/day, commonly they are of the order 10^{-2}–10^{-4} eq/day and lower. For such a small rate

* Corresponding author.
E-mail addresses: ykagan@ucla.edu (Y.Y. Kagan), djackson@ucla.edu (D.D. Jackson).

0040-1951/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.tecto.2005.10.007

value, the rate is almost equivalent to the probability of an earthquake occurrence, thus when it cannot lead to confusion we sometimes use a more common expression "probability" instead of rate.

We produced "long-term" and "short-term" forecasts. Long-term forecasts employ an assumption that each earthquake provides information about a rate-density field independent of time. Short-term forecasts are based on the assumption that the rate-density results from individual earthquakes, with an Omori type time-dependence (i.e., decreasing with reciprocal time). K2005 deals with short-term forecasts only.

We designed our model to optimize the likelihood score: a scalar measure of the probability density at the locations of future quakes. K2005 employs the error diagram: a different testing methodology, originally developed to evaluate deterministic predictions. He further introduces a scalar quantity to measure the goodness of fit in a particular result. There is nothing wrong with using these yardsticks, but they are somewhat arbitrary. The error diagram has a certain advantage when risk mitigation actions require turning the probabilistic language of scientists into concrete actions, but for measuring the efficiency of forecasts other yardsticks may be more appropriate. We discuss below some reasons why error diagrams are not a good measure of the essential features of our forecasts.

Kossobokov has been testing our results for a few years now, and his conclusions have evolved from the beginning. Previously Kossobokov (2003a) analyzed our forecast record for one half-year and in his next publication (2003b) apparently for a similar period. The previous results, shown in posters at meetings, were similar to Fig. 1 of K2005 (except that these early results had fewer data and thus higher random fluctuations). Kossobokov (2003b) wrote that

> ". . . the effectiveness of prediction came out to be hardly better than random guessing. The conclusion holds even when evident aftershocks are included into statistic of successes."

Kossobokov (2003a) expresses a similar opinion.

Kossobokov (2004) evaluated our forecast for the period 2002/04/10–2003/04/09. His Fig. 1a,b is similar both in display methods and results to Figs. 1 and 2 of K2005, respectively. However, the 2004 abstract says that

> ". . . the conclusion is drawn that the Jackson-Kagan method's [effectiveness] of locating the space–time of large earthquakes is rather doubtful . . ."

(we quote from the English abstract of the paper with a slight modification to make the text more compatible to the Russian original).

K2005 comments that in Fig. 1a by Kossobokov (2004) the difference between the observational curves and the random prediction is spread over a smaller range than in the K2005 plot. These figures represent the "error diagrams" (Molchan and Kagan, 1992; Molchan, 1997, 2003) with two variables that characterize prediction efficiency: the fraction of space–time alarm area μ, and the fraction of failures to predict, ν. However, in any short-term forecast test the most important part of the error diagram is for small values of μ. This is because according to Omori's law, high values of earthquake rates are limited to relatively minute space–time areas, these areas correspond to small μ's. In other words, highly localized forecasts are more useful, since they allow to use very short-term earthquake mitigation measures. The behavior of the curves in Kossobokov (2004) and K2005 is almost identical in the range $0.1 \geq \mu \geq 0$ (see more discussion in Section 4). However, Fig. 1 elicits the following comment in K2005:

> "It is rather evident that the achieved statistics are much better than random guessing through wide ranges of μ . . ."

Although we have no problem with his statement, it is inconsistent with the statements quoted earlier from Kossobokov (2003a,b, 2004). Since the test results in all of the preceding publications were almost identical for small values of space–time ($0.1 \geq \mu \geq 0$), it is not likely that this evaluation difference was caused by insufficient data.

We would like to discuss several issues concerning the K2005 test. Some of these issues transcend K2005 comments and may be of interest to all who attempt earthquake forecasting or prediction. In particular, we discuss the following items here:

- General definition and aims of earthquake forecasting or prediction;
- Short description of our algorithm and its objectives;
- K2005 short-term forecast test;
- K2005 comparison of our forecast with the *M*8 algorithm;
- Prediction of mainshocks.

2. Earthquake prediction definition

K2005 comments that our forecast contradicts the definition of earthquake prediction proposed by Allen et al. (1976). Though practical and necessary at that time, the 1976 definition is clearly inappropriate for *statistical* or *probabilistic* earthquake forecasts. In our opinion, only statistical forecasts should be considered in modern earthquake prediction research. The prediction definition by Allen et al. (1976) is no longer applicable for modern earthquake forecast efforts, because it did not consider probabilities.

In K2005, Kossobokov suggests that our use of the term "short-term" is misleading because some regions have high forecasted results persisting for years. Our usage was always clear, consistent and remains so. By "short-term," we mean a forecast updated frequently enough to respond to temporary clustering. For practical reasons we update only once per day.

Our results show that earthquake rates decrease approximately as the reciprocal of time since a previous event, as in Omori's law. Thus, it would be scientifically and strategically advantageous to update immediately after each earthquake reported.

In our forecast some regions of high seismicity do persist for a long time because frequent events keep stimulating future ones. In those regions the estimate rates may vary strongly with time, even though they always exceed those of neighboring regions.

3. Probabilistic forecasting of earthquakes

The forecast algorithm running since 1999/1/1 in a test mode for the west Pacific regions (Jackson and Kagan, 1999; Kagan and Jackson, 2000), combines two procedures: the short-term statistical method (Kagan and Knopoff, 1987) and the long-term program (Kagan and Jackson, 1994). When we started this experiment, we hoped to test the algorithm thoroughly in both short- and long-term applications, and then develop a probabilistic procedure to combine both methods in a comprehensive model. We tested the long-term forecast (Jackson and Kagan, 1999; Kagan and Jackson, 2000); see also additional results for years 2000–2002 at http://scec.ess.ucla.edu/~ykagan/tests_index.html. Because of funding limitations, we were unable to develop the forecast method fully, create a permanent archive of forecasts, or thoroughly test the method's effectiveness.

We have been using the Harvard CMT catalog (Ekström et al., 2005, and references therein) for our forecast experiments. We update the catalog every day, using email messages from Harvard. This email information has a preliminary form with higher uncertainties and completeness threshold. The final catalog is released in monthly installments with a delay of a few months. We retrieve the finished dataset from Harvard once every few months. Therefore, the forecast tables and figures which appear on our Web site around midnight Los Angeles time are partially based on preliminary data. For the long-term forecast, the lower accuracy of the recent (a few months) part of the catalog should not significantly influence the result. For the short-term forecast, the use of the temporary catalog may present a problem for testing and evaluation.

The short-term earthquake rate values change very fast following any earthquake. They follow a power-law dependence on time (Jackson and Kagan, 1999; Kagan and Jackson, 2000). This rapid change necessitates frequent updating of forecasts. Ideally, a new forecast should be calculated after each event as discussed by Helmstetter et al. (in press), but practical considerations prompted us to renew the forecast only once daily. Moreover, a delay of at least a few hours occurs between an earthquake and its processing by the Harvard team. This delay also degrades the performance of our short-term algorithm.

Helmstetter et al. (in press) applied a similar technique for short- and long-term forecasts of seismicity in southern California. They use a local earthquake catalog with the magnitude threshold 2.0. The forecast is also updated daily and available on the Web http://moho.ess.ucla.edu/~helmstet/forecast.html.

4. K2005 short-term forecast test

As mentioned above, we agree in principle with the K2005 characterization of short-term forecast testing: compared to a random guess it has a clear advantage in evaluating the probability of earthquake occurrence quickly following other events.

However, the K2005 method to calculate the null (Poisson) hypothesis threshold for testing may have a bias, which can favor rejecting the null hypothesis. K2005 used the global 1963–1984 earthquake catalog with magnitude cut-off $M_c = 4.0$ to calculate the Poisson rate of earthquake occurrence, which was then compared to short-term earthquake rate. The same method was employed in testing the $M8$ method (Kossobokov et al., 1997, 1999).

This test procedure in each Kossobokov's paper has two problems: 1) global catalogs are incomplete for earthquakes $5.0 \geq M \geq 4.0$, and 2) as a rule, the magnitude threshold decreases with time, as technology improved over the 1963–1984 period. For example, Willemann (1999) shows that the body-wave magnitude (m_b) threshold changes in time and space for the ISC (International Seismological Centre) catalog. He estimates that M_c-values fluctuate between 4.0 and

5.0 (see also Fig. 3 by Kossobokov et al., 1997). Marzocchi et al. (2003) suggest that this magnitude incompleteness may lead to the unwarranted higher performance estimate of the *M*8 method.

Obviously, catalog incompleteness weakens the null hypothesis, and comparison to a weak null hypothesis makes the test hypothesis look stronger than it is (Stark, 1997). This spurious benefit will be greater for a long-term method like *M*8 for two reasons. First, the null hypothesis must rely on an older and less complete catalog. Second, the rate variations in a long-term forecast like *M*8 are relatively small, so deficiencies in the null hypothesis are proportionally larger in that case.

To apply the error diagram in an earthquake mitigation problem, the curves could be made concave by randomizing optimal response strategies (Molchan and Kagan, 1992, their Fig. 8). Thus, if one connects by a straight line the points $\mu=1$, $\nu=0$ (the pessimist strategy) with the observational curve value at $\mu=0.1$, the results in K2005 and Kossobokov (2004, Fig. 1a) are almost identical.

Molchan and Kagan (1992, their Fig. 8) created a similar error diagram for San Andreas fault earthquakes, using the retrospective short-term forecast algorithm proposed by Kagan and Knopoff (1987). Because we can calculate the probability immediately following each earthquake in the retrospective application, the prediction efficiency measured by how the observational curves deviate from a random guess is much higher than that in our prospective forecast (see Fig. 1 in K2005). For the west Pacific CMT catalog, many aftershocks (20–30%) follow a large event within one day. Because we calculate the forecast tables once per day and messages from Harvard are delayed under the best conditions by a few hours, the higher values of short-term probabilities do not enter our results.

5. Short-term forecast versus *M*8 comparison

K2005 compares the sum of errors ($S=\mu+\nu$) in two predictions: Jackson and Kagan (1999) and the *M*8 algorithm, quoting the values 84% and 50.6%, respectively. Lower *S*-values correspond to a higher forecasting efficiency. However, the sum of errors is neither the only nor the best measure of forecast efficiency. For example, if we had two sets of error diagram results $\mu_1=0.1$, $\nu_1=0.7$ (similar to Fig. 1 in K2005) and $\mu_2=0.7$, $\nu_2=0.1$, the sum would be the same in both cases: $S=0.8$. However, in terms of possible earthquake warning strategies, the first forecast would require an alert with a duration of 10% of space–time, predicting 30% of events. The second forecast would declare 70% space–time dangerous, with 90% of events in alert zones. One way to evaluate the effectiveness of a forecast is to calculate its 'predictive ratio' (Kagan and Knopoff, 1977, p. 101) or 'probability gain' and 'information gain' (Daley and Vere-Jones, 2004; Helmstetter et al., in press). The predictive ratio is the ratio of short-term forecasted earthquake occurrence rate to the Poisson rate. If the rates are much smaller than 1.0, as is the case for small space–time regions, they are almost equal to earthquake probabilities, hence the other name for this ratio is *probability gain*. For the first example, the gain is 3.0; for the second forecast it is $0.9/0.7\approx1.3$. Measures to reduce earthquake damage could likely be applied more effectively if the prediction gain or predictive skill of a method is higher (Molchan and Kagan, 1992; Molchan, 1997, 2003).

As we mentioned in Section 4, there is a possibility that calculating error diagrams for the *M*8 program by Kossobokov et al. (1999) may bias the results towards lower μ-values and hence lower *S* estimates. An independent rigorous test of the *M*8 program is needed to verify its performance. But the *S*-values would become lower if the time delay between earthquake occurrence and report were significantly reduced and the short-term forecast issued immediately after each earthquake. K2005 claims that $S\approx50\%$ is the *M*8 performance estimate. Molchan and Kagan's (1992, their Fig. 8) diagram suggests that $S\approx50\%$ is a reachable goal for our short-term algorithm.

6. Mainshock forecast

K2005 asserts that the success of our forecasts is due solely to their ability to forecast aftershocks, and that when aftershocks are removed from the tests the forecasts are not successful. These statements miss the point in a few respects.

Forecasting aftershocks has a value in its own right. While seismologists all agree that aftershocks can be expected after most earthquakes, a quantitative probabilistic description has substantial value for those who must respond to earthquakes.

Depending on how aftershocks are defined, Kossobokov's statement about mainshock prediction, based on Fig. 2 of K2005, could be a tautology. There is no standard definition of aftershocks, but several have been used for specific purposes. Molchan and Dmitrieva (1992) argue that aftershock identification depends on research goals. K2005 uses an aftershock definition in which any event within a time and space window of a previous larger event is labeled an "aftershock." The size of the window is magnitude dependent. Kagan and

Jackson (1991, p. 119) used a more sophisticated scheme in which each earthquake is assigned a probability of being independent. The probability is evaluated by a likelihood analysis of the earthquake catalog (Kagan and Knopoff, 1987; Kagan and Jackson, 2000). Zhuang et al. (2004) employed a similar procedure. All of these methods have a common goal: to distinguish those events that could be expected, based on previous events, from independent ones. Thus, the very events that we forecast with high likelihood could be labeled "aftershocks," and of course the remainder, by definition, is not well forecasted by our model.

The declustering algorithm which K2005 applies (based on Keilis-Borok et al., 1988) creates holes in a catalog after each strong earthquake, because all weaker events inside an aftershock window are deleted. However, this procedure does not take into account that an event belonging to background seismicity in a window may occur by chance. Even in a Poisson process some events may be close in space or time from mere randomness.

Because K2005 uses a different aftershock recognition scheme than would be obtained by our forecast model, the remaining (after declustering) events are not completely independent. In fact, the remaining events are anti-correlated with our forecast model, as shown in Fig. 2 of K2005. We believe this fact results from a quirk in Kossobokov's aftershock definition. Fig. 2 of K2005 suggests that when aftershocks are removed, our algorithm performance is even worse than a random guess. It is likely, the remaining 'mainshocks' form a residual process which, as we discussed in the previous paragraph, may exhibit even less clustering than a Poisson process. This may explain the unusual features of the diagram.

By many definitions, including the one used in K2005, aftershocks are smaller than the mainshock that preceded them. This could lead to an assumption that following an apparent main shock, no larger event will occur. In fact, it is not uncommon for a large earthquake to be followed by an even larger one. Two tables of earthquake pairs in NW and SW Pacific polygons have been calculated and posted at the URL address: http://moho.ess.ucla.edu/~kagan/tables-NW-SW-Pacific.txt. For 5 pairs out of 30 in the NW and 27 pairs out of 89 in the SW, the second event is larger than the first. Such events, while not frequent compared with all the other events in a catalog, are important targets in our forecast. Kossobokov's test does not measure this quality of the forecast.

However, some large earthquakes which would be recognized by a majority of seismologists as mainshocks are unquestionably preceded by foreshocks. Reasenberg (1999a,b) determines that between 8% and 17% of $M \geq 6.0$ earthquakes in the Harvard catalog have foreshocks. These clear mainshocks should be forecasted by our method if the time delay is greater than one day. As we explained in Section 3, our forecasts could be even more successful with more frequent updates, because possible triggering events that occur after the last update are not considered. Our test of a short-term forecast algorithm in California (Kagan and Knopoff, 1987, their Fig. 3) indicates that a significant fraction of mainshocks is preceded by foreshocks and thus in principle can be predicted by the method.

In our 2002/04/10–2004/09/13 record, for example, the Irian Jaya earthquake sequence provides an example of mainshock prediction (all data below are from Harvard emails). The sequence started with a M5.9 earthquake on 23:09:35 UTC 2004/02/03 (epicenter coordinates: 3.59S; 140.69E), followed by M6.9 event on 21:05:14 2004/02/05 (3.63S; 135.45E), and followed by M7.3 shock 02:42:43 2004/02/06 (4.02S; 134.71E). The first earthquake raised the earthquake rate value by 8% near the epicenter of the second event. For the third earthquake, the earthquake rate was increased by a factor of 44.

7. Conclusions

As always we fully support objective and prospective efforts for testing forecasting and prediction experiments. K2005 is a constructive contribution to these efforts. However, in the testing method employed there is not well suited to our forecasts. Our forecasts are most valuable in identifying local regions where the probability of damaging earthquakes, especially those larger than a previous trigger, may be modest but temporarily elevated. K2005 tests more generically, and the S criterion is not very sensitive to the probability gain in those local regions.

Acknowledgements

The authors appreciate support from the National Science Foundation through grant EAR 04-09890 and from the Southern California Earthquake Center (SCEC). SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. We are grateful to reviewers P. Stark, A. Michael, and an anonymous reviewer, as well as to the Guest Associate Editor M. Anghel for very helpful comments and suggestions. Kathleen Jackson edited the final version. Publication 934, SCEC.

References

Allen, C.R. (Chairman), Edwards, W., Hall, W.J., Knopoff, L., Raleigh, C.B., Savit, C.H., Toksoz, M.N., Turner, R.H., 1976. Predicting earthquakes: a scientific and technical evaluation with implications for society. Panel on Earthquake Prediction of the Committee on Seismology, Assembly of Mathematical and Physical Sciences, National Research Council, U.S. National Academy of Sciences, Washington, D.C.

Daley, D.J., Vere-Jones, D., 2003. An Introduction to the Theory of Point Processes, 2-nd ed., vol. 1. Springer-Verlag, New York, p. 469.

Daley, D.J., Vere-Jones, D., 2004. Scoring probability forecasts for point processes: the entropy score and information gain. J. Appl. Probab. 41A, 297–312 (Sp. Iss.).

Ekström, G., Dziewonski, A.M., Maternovskaya, N.N., Nettles, M., 2005. Global seismicity of 2003: centroid-moment-tensor solutions for 1087 earthquakes. Phys. Earth Planet. Inter. 148 (2–4), 327–351.

Helmstetter, A., Kagan, Y.Y., Jackson, D.D., in press. Comparison of short-term and long-term earthquake forecast models for Southern California, Bull. Seismol. Soc. Amer. http://scec.ess.ucla.edu/~ykagan/agnes_index.html.

Jackson, D.D., Kagan, Y.Y., 1999. Testable earthquake forecasts for 1999. Seismol. Res. Lett. 70 (4), 393–403.

Kagan, Y.Y., Jackson, D.D., 1991. Long-term earthquake clustering. Geophys. J. Int. 104, 117–133.

Kagan, Y.Y., Jackson, D.D., 1994. Long-term probabilistic forecasting of earthquakes. J. Geophys. Res. 99, 13,685–13,700.

Kagan, Y.Y., Jackson, D.D., 2000. Probabilistic forecasting of earthquakes. Geophys. J. Int. 143, 438–453.

Kagan, Y., Knopoff, L., 1977. Earthquake risk prediction as a stochastic process. Phys. Earth Planet. Inter. 14 (2), 97–108.

Kagan, Y.Y., Knopoff, L., 1987. Statistical short-term earthquake prediction. Science 236, 1563–1567.

Keilis-Borok, V.I., Knopoff, L., Rotwain, I.M., Allen, C.R., 1988. Intermediate-term prediction of occurrence times of strong earthquakes. Nature 335, 690–694.

Kossobokov, V.G., 2003a. West Pacific short-term earthquake forecast: an independent assessment. 2003 European Geophys. Soc. Meeting, Nice, France, Abstracts. http://www.cosis.net/abstracts/EAE03/06539/EAE03-J-06539.pdf.

Kossobokov, V.G., 2003b. An independent evaluation of the west Pacific short-term earthquake forecast. IUGG 2003 Meeting Abstracts, Sapporo, Japan. http://www.jamstec.go.jp/jamstec-e/iugg/htm/abstract/abst/jws01_p/019775-3.html.

Kossobokov, V.G., 2004. "The West Pacific short-term forecast of earthquakes with magnitude MwHRV ≥5.8" by D. Jackson-Y. Kagan: an independent evaluation of efficiency. Comput. Seismol. 35, 201–211 (in Russian with English abstract).

Kossobokov, V.G., 2006—this issue. Testing earthquake prediction methods: "The West Pacific short-term forecast of earthquakes with magnitude MwHRV ≥5.8". Tectonophysics (K2005) 413, 25–31.

Kossobokov, V.G., Healy, J.H., Dewey, J.W., 1997. Testing an earthquake prediction algorithm. Pure Appl. Geophys. 149, 219–232.

Kossobokov, V.G., Romashkova, L.L., Keilis-Borok, V.I., Healy, J.H., 1999. Testing earthquake prediction algorithms: statistically significant advance prediction of the largest earthquakes in the Circum-Pacific, 1992–1997. Phys. Earth Planet. Inter. 111, 187–196.

Marzocchi, W., Sandri, L., Boschi, E., 2003. On the validation of earthquake-forecasting models: the case of pattern recognition algorithms. Bull. Seismol. Soc. Am. 93 (5), 1994–2004.

Molchan, G.M., 1997. Earthquake prediction as a decision-making problem. Pure Appl. Geophys. 149, 233–247.

Molchan, G.M., 2003. Earthquake prediction strategies: a theoretical analysis. In: Keilis-Borok, V.I., Soloviev, A.A. (Eds.), Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer, Heidelberg, pp. 208–237.

Molchan, G.M., Dmitrieva, O.E., 1992. Aftershock identification: methods and new approaches. Geophys. J. Int. 109, 501–516.

Molchan, G.M., Kagan, Y.Y., 1992. Earthquake prediction and its optimization. J. Geophys. Res. 97, 4823–4838.

Reasenberg, P.A., 1999a. Foreshock occurrence before large earthquakes. J. Geophys. Res. 104, 4755–4768.

Reasenberg, P.A., 1999b. Foreshock occurrence rates before large earthquakes worldwide. Pure Appl. Geophys. 155 (2–4), 355–379.

Stark, P.B., 1997. Earthquake prediction: the null hypothesis. Geophys. J. Int. 131, 495–499.

Willemann, R.J., 1999. Regional thresholds of the ISC bulletin. Seismol. Res. Lett. 70 (3), 313–321 (Errata, p.529).

Zhuang, J.C., Ogata, Y., Vere–Jones, D., 2004. Analyzing earthquake clustering features by using stochastic reconstruction. J. Geophys. Res. 109 (B5), B05301.