Review of 3 reports by NAM

Dirk Kraaijpoel, Jesper Spetzler and Bernard Dost, KNMI, May 26, 2015

SR.15.BinnedParetoMLE

The report describes a method for the MLE estimation of b-values given a seismic catalogue with rounded (binned) magnitude values, as well as its application on the Groningen dataset in various subsets.

- The method has been derived and described rigorously, and seems to be the proper way of doing MLE for binned magnitudes. It uses a discrete multinomial distribution (Eq. 2.8), rather than the continuous exponential distribution as used in common practice. The latter requires approximate corrections for the binning to prevent bias (e.g., Marzocchi and Sandri, Annals of Geophysics, 2003).
- The method is expressed in terms of seismic moments and the Pareto distribution, rather than magnitudes and the exponential (Gutenberg-Richter) distribution. This choice is not essential, since after binning the resulting multinomial distribution is exactly the same. The detour through seismic moments seems to be an unnecessary distraction, but this may be a matter of perspective or taste. We note that the magnitudes communicated by KNMI are local rather than moment magnitudes, so definition Eq. (1.1) is formally not valid, although we usually assume the equivalence of local and moment magnitude.
- The report perhaps shows a bit too much respect for the "accepted value b=1" (page 7). It is not clear why the MLE estimate of b=0.966 is replaced by the exact number b=1 for further analysis on the errors. There is no reason why b=1 would be better or preferred over b=0.966.
- To analyze the uncertainty ("error bar") of the MLE estimate the report describes a Monte Carlo experiment that provides an empirical distribution for the MLE estimate of b given a fixed model value of b=1. This procedure provides confidence bounds on the range of MLE estimates to be expected for a single model b-value. However, in fact we are more interested in the confidence bounds on the range of model b-values that may explain the single "observed" MLE estimate of b. Therefore, model value b should be varied to obtain its likelihood profile. This may possibly give the same result but this is not clear from the report.
- Apart from sets of disjoint catalogues, also sliding windows may provide more insight into temporal variations.
- We have tested a number of reported b-values with our own MLE procedures and obtained equivalent results.
- The observation of the low b-value in the Loppersum area is interesting. Perhaps other areas may be of interest as well (Hoogezand?).
- Hypothesis testing may be sharpened. It is concluded that a b-value of 1 can be rejected for the Loppersum area. But is there another constant value of b (not the entire field MLE estimate; lower) that may explain both the Loppersum area and the rest? In other words, can we reject the null hypothesis that the Loppersum area has the same b-value as the rest of the field (or as the Ten Boer area)?
- Suggestion for further work: study variations with other "time-like" variables such as local reservoir pressure or local compaction, etc.

Statistical methodology to test for evidence of seasonal variation in rates of earthquakes in the Groningen field

The report studies possible seasonal variations in earthquake rates.

- The motivation for the statistical models and algorithms in Section 4 is a bit too brief. Also the use of index i for both “julian months” and weeks is confusing. The results of the report are difficult to understand on the basis of this Section.
- Figure 5.1 shows smoothed monthly deviations in rate. In the figures there is no smooth connection between the end and the beginning of the year. The transition from December to January should be just as smooth as from, e.g., May to June. The solution is to use periodic smoothing and perhaps circular statistics. The Figures are hard to interpret now because some of the most prominent features may be artefacts.
- Seasonal and diurnal variations in small-magnitude events may indeed be well explained by environmental circumstances, especially background noise. It would be interesting to look at correlation between background noise and event rate (note to self).

A2.1 Activity rate Loppersum

The report looks for evidence for a change in seismicity rate after the production measures of Jan 2014, by looking at the inter-event times of sets of consecutive, declustered events.

- The use of lambda for inter-event time is confusing, since usually lambda is used for its inverse: the event rate
- The use of a Markov Chain Monte Carlo Method for a 1D parameter characterization seems a bit cumbersome.
- The statement that an “aftershock” is not a “true event” is a bit bold. It is a true event but it may not be independent from other events in time and place. If aftershocks are not removed the assumption of a Poisson process is invalid.
- The visual inspection of distributions to assess if hypotheses can be rejected is a bit subjective. This can be quantified with significance levels or Bayesian equivalents. However, the combination of kernel density figures and confidence levels do provide useful information. Especially the difference between results for all events compared to M>1.5 is striking (compare figure 20 and 28). Clearly the interevent time goes up in the last period, contrary to the situation for the M>1.5 events. This is an important result.
- It may be interesting to also look at the numbers of events recorded in the deep boreholes.
- It may be interesting to look at inhomogeneous rate models.
- The remark in the summary that “we don’t know the exact location of the geophones” is unclear. We do know the exact position of the geophones very well. Maybe they meant “events” instead.
- Compared to the heat maps published in the TNO report, taking into account only one year of seismicity prior to the change in production, this study compares a number of years prior to the change in production and gives a quantification of the change in activity rate.