Induced Seismicity Indicated by Cross Correlation*

Ivo Oprsal[1] and Leo Eisner[2]

Search and Discovery Article #41318 (2014)
Posted April 14, 2014

*Adapted from extended abstract prepared in conjunction with oral presentation at GEO-2014, 11th Middle East Geosciences Conference and Exhibition, 10-12 March 2014, GEO-2014©2014

[1]Seismik s.r.o., Prague, Czech Republic
[2]Institute of Rock Structure and Mechanics/Seismik s.r.o, Prague, Czech Republic

Abstract

The qualitative correlation between earthquake rates (i.e., the number of observed earthquakes larger than the magnitude of completeness) and the injected volume, has been an established tool for investigating possible induced or triggered seismicity (Evans 1966, Healy et al 1968, Hsieh and Bredehoeft, 1981), however, the method using direct values of the earthquake rates and the injected volumes for normalized cross-correlation is a recent development (Horton 2012). The use of cross-correlation indicates e.g., the temporal dependence of seismic activity resulting from man-made actions without implying physical dependence of the phenomena. At the same time, the maximum cross-correlation value of positive functions can inherently have a relatively high maximum regardless of its physical meaning.

Determining the relationship between injection rates and seismicity in an area where injection is occurring through normalized cross-correlation (NCC), could have the potential to be an objective tool for investigating the causal relationship between seismicity and injection. However, we show that a direct cross-correlation between the injection rates and seismicity would result in high cross-correlation values even for random functions. The injection volumes as well as the seismicity (event count), are both positive functions. We show theoretically and by numerical example that the mathematical definition of NCC provided by a set of positive random functions, exhibits high cross-correlation values with a limit equal to one for large mean and low standard deviation time histories.

Instead of positive-value time histories, we use their “Useful Functions” (the original functions with their weighted running mean subtracted). Then the normalized cross-correlation (NCC) of such functions may be close to zero or oscillating between positive and negative values of +0.5 and -0.5 in cases where seismicity is probably not induced by injection. NCC dominated by a positive maximum with a zero (or small) time lag, indicates that seismicity is induced by the injection.

We present two of several test cases with well-known induced or triggered seismicity: Blackpool, UK, and Rocky Mountain Arsenal, Denver, Colorado, USA. They show a statistically significant positive cross-correlation peak with no significant negative peaks. Furthermore, in all test cases we found a nearly zero time lag between induced seismicity and injection. An example of very probable natural seismicity near the Guy-

Greenbrier fault shows that NCC for corresponding useful functions gives values with statistically insignificant peaks ranging between +0.5 and -0.5.

Method and Theory

Let us investigate the normalized cross-correlation of time series A and B with non-zero means.

$$A = N_A + \mu_A \ , B = N_B + \mu_B$$

Where (e.g.): is a function of zero mean, μ_A is mean of A and the respective standard deviation reads:

$$E(N_A^2) = \sigma_A^2 \ , \ E(N_A) = 0$$

E(X) is the expected value of X. There are plenty of definitions of cross-correlation. The normalized cross-correlation (as in MATLAB Signal Processing Toolbox, or ‘reflective correlation’, Wikipedia) reads:

$$Corr_n(A,B) = \frac{E(AB)}{\sqrt{E(A^2)E(B^2)}} = \frac{E((N_A + \mu_A)(N_B + \mu_B))}{\sqrt{E((N_A + \mu_A)^2)E((N_B + \mu_B)^2)}} == \frac{E(N_A N_B) + \mu_A \mu_B}{\sqrt{(\sigma_A^2 + \mu_A^2)(\sigma_B^2 + \mu_B^2)}} \qquad (1)$$

For statistically independent random functions NA, NB:

$$\frac{\mu_A \mu_B}{\sqrt{(\sigma_A^2 + \mu_A^2)(\sigma_B^2 + \mu_B^2)}} \qquad (2)$$

Therefore for two functions with non-zero means, the smaller the deviation around that mean the higher the cross-correlation (assuming both means have the same signs). For histories with relatively small σ_A and σ_B and independent functions N_A, N_B:

$$\lim_{\frac{\sigma_A}{\mu_A}, \frac{\sigma_B}{\mu_B} \to 0,} \left(\frac{\mu_A \mu_B}{\sqrt{(\sigma_A^2 + \mu_A^2)(\sigma_B^2 + \mu_B^2)}} \right) = 1 \qquad (3)$$

For histories with $\mu_A = \mu_B = 0$:

$$Corr_{n_{zeromeans}}(A,B) = \frac{E(N_A N_B)}{\sigma_A^2 \sigma_B^2} \qquad (4)$$

Only in this case the normalized cross-correlation ranges between -1 and +1 and may indicate physical relationship if close to 1. Note that 'xcorr' is different from MATLAB Statistics Toolbox's 'corr', which computes correlation as specified in Equation (4).

Formulae (1) - (4) give maximum (or plateau) values of the cross-correlation and are valid as regards to the average of the realizations or infinite time-series limit, being an estimate for the current realization of independent functions (see Figure 1A, bottom plot). The theoretical value of NCC (Equation 2, 'max_xcross(theor)' value in all figures) is computed for every time series assuming independent functions A and B (E(AB)=0).

Synthetic Dataset

The first example shows the normalized cross-correlation (NCC) of two positive (daily value) random functions with boxcar distribution in <0,1> (μ=1/2), Figure 1A. The random realization of these two variables (the upper two rows of panels) is the same, but the statistical features of the second row are computed to include the zero-padded interval as in MATLAB. The panels in the third row of Figures 1A and 1B show that the plateau amplitudes are consistent (in average) with the theoretically obtained values of the normalized cross-correlation. Had the functions been of equal length, then the envelope of the NCC would have been triangular. In this study, the envelopes of the NCC are trapezoids with their maximum amplitudes decreased due to normalization. Note the change in mean value and standard deviation of the second panel of Figure 1B time history from the non-zero part expected μ~1/2 and $\sqrt{1/12}$ =0.289 due to zero-padding and normalization. The time series of Figure 1B is 'weekly' values created by summation of the Figure 1A random 'daily' realizations. If normalized then the mean value of their non-zero parts would remain unchanged while their standard deviation would decrease to $\sqrt{\frac{1}{7}\frac{1}{12}}$ =0.109 following the Central Limit Theorem (Chow and Teicher, 1978, Fischer 2010). The plateau rises from an average (or theoretical) 0.56 (the lower panel of Figure 1A) to approximately 0.7 (1B). Such a plateau-like maximum is very unstable in terms of the lag value for individual realizations. The Pearson's (Rodgers and Nicewander, 1988) test of this random realization in Figure 1C shows a very low NCC (0.21) which would be significantly lower for a longer time series or an average of multiple realizations. The lag of 19 weeks is obviously a very formal, unstable value for random functions.

Case Studies

We study the relationship between injection and seismicity rates for the Rocky Mountain Arsenal (USA) and Blackpool (UK, 2011) cases of induced seismicity (Figures 2A and B, respectively) by a method called the 'extended Pearson's test'. The useful functions are obtained by subtracting the low-frequency information from the seismicity and injection, where we subtract the weighted (w=0.85) running window average from the data. The NCC functions for these two induced-seismicity cases are slightly oscillating functions with positive zero-lagged peaks with values of ~0.7 and very small negative cross-correlation values. The third case presented here is the activity near the Guy-

Greenbrier fault (Arkansas) in 2010-2011 (Figure 3). The Pearson's cross-correlation shown in Figure 3B exhibits two important features: 1) the maximum of normalized cross-correlation is lower than it is for the original positive functions; and 2) the negative part of Pearson's NCC (NCCP) for the data set (anti-correlation) is of the same amplitude as the positive part. This would imply that injection of fluid reduces earthquake activity, which is physically improbable. A simpler explanation is that cross-correlation peaks are just random values and the two phenomena are unrelated.

Conclusions

The normalized cross-correlation (NCC) of positive functions, such as injection rates and seismicity, has a limit equal to 1 even for purely random (independent) functions. The long-term average values (e.g., weekly) of injection rates and seismicity have a smaller deviation than the daily values' higher maxima (Central Limit Theorem). This is the reason for very high NCC values (Equation 3) even for a random time series of daily values.

Cross-correlation of the high-frequency part of a time series (useful function) is an objective tool to investigate the temporal relationship of the time series. The NCC lag between injection and seismicity rates for proven cases of induced seismicity shows a nearly zero time lag between injection and seismicity. It peaks at statistically significant values above 0.5 and has very small values of negative cross-correlation. Such cross-correlation may be regarded as an objective criterion of induced seismicity. The NCC function for induced-seismicity is expected to be a slightly oscillating function with a positive zero-lagged peak.

Acknowledgements

We thank Jessica Griffin (Chesapeake Energy Corp.) and anonymous colleagues for their useful discussion and support, and Clive A Cowan of CCR Editorial for technical reading and editing.

References Cited

Chow, Y.S., and H. Teicher, 1978, Probability Theory: Independence, Interchangeability, Martingales: New York and Heidelberg, Springer-Verlag, 455 p.

Fischer, H., 2010, A History of the Central Limit Theorem: From Classical to Modern Probability Theory, Sources and Studies in the History of Mathematics and Physical Sciences: New York, Springer, 1st ed., 418 p.

Healy, J.T., W.W. Rubey, D.T. Griggs, and C.B. Raleigh, 1968, The Denver earthquakes: Science, v. 161, p 1301-1,310.

Horton, S., 2012, Disposal of Hydrofracking Waste Fluid by Injection into Subsurface Aquifers Triggers Earthquake Swarm in Central Arkansas with Potential for Damaging Earthquake: Seismological Research Letters, v. 83/2, p. 250-260.

Hsieh, P.A., and J.D. Bredehoeft, 1981, A Reservoir Analysis of the Denver Earthquakes: A Case of Induced Seismicity: J. Geophys. Res., v. 86(B2), p. 903–920, doi:10.1029/JB086iB02p00903.

Rodgers, J.L., and W.A. Nicewander, 1988, Thirteen ways to look at the correlation coefficient: The American Statistician, v. 42/1, p. 59–66.

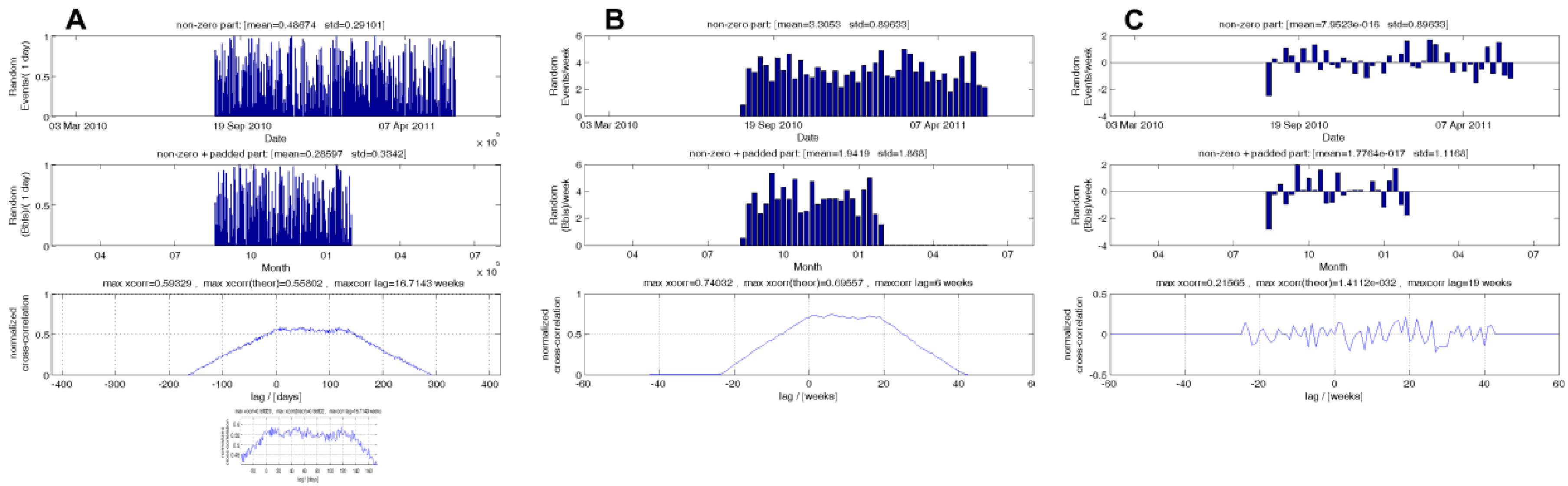

Figure 1. Normalized cross-correlation (bottom panels) of two random functions (top two rows of panels) substituting frequency of events and injection in daily rates (A) and in weekly rates (B). Panels C depict the Pearson's cross-correlation of weekly data as presented in panels B. The mean values of the non-zero-padded parts are subtracted before computing the normalized correlation. For a longer series or more realizations (in terms of average), the correlation limits to zero for independent variables. The theoretically obtained value of the normalized cross-correlation (of presumably independent functions) is denoted as max_xcorr(theor), Equation (2).

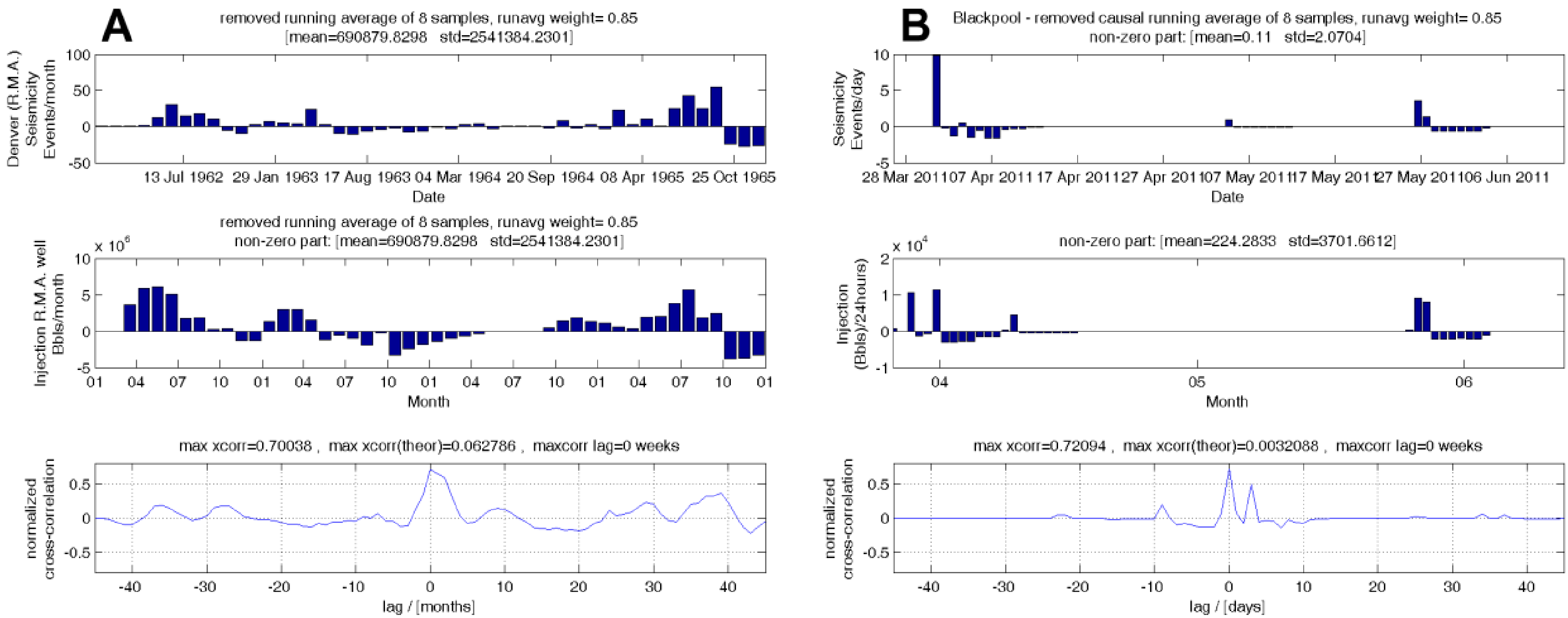

Figure 2. NCC for useful functions (extended Pearson's test). A: For number of earthquakes at the Rocky Mountain Arsenal waste injection (upper panel) versus injected volume (lower panel) (Data for useful functions taken from Healy et al, 1968), B: For number of events in the Blackpool (UK) site (upper panel) versus injected volume (lower panel).

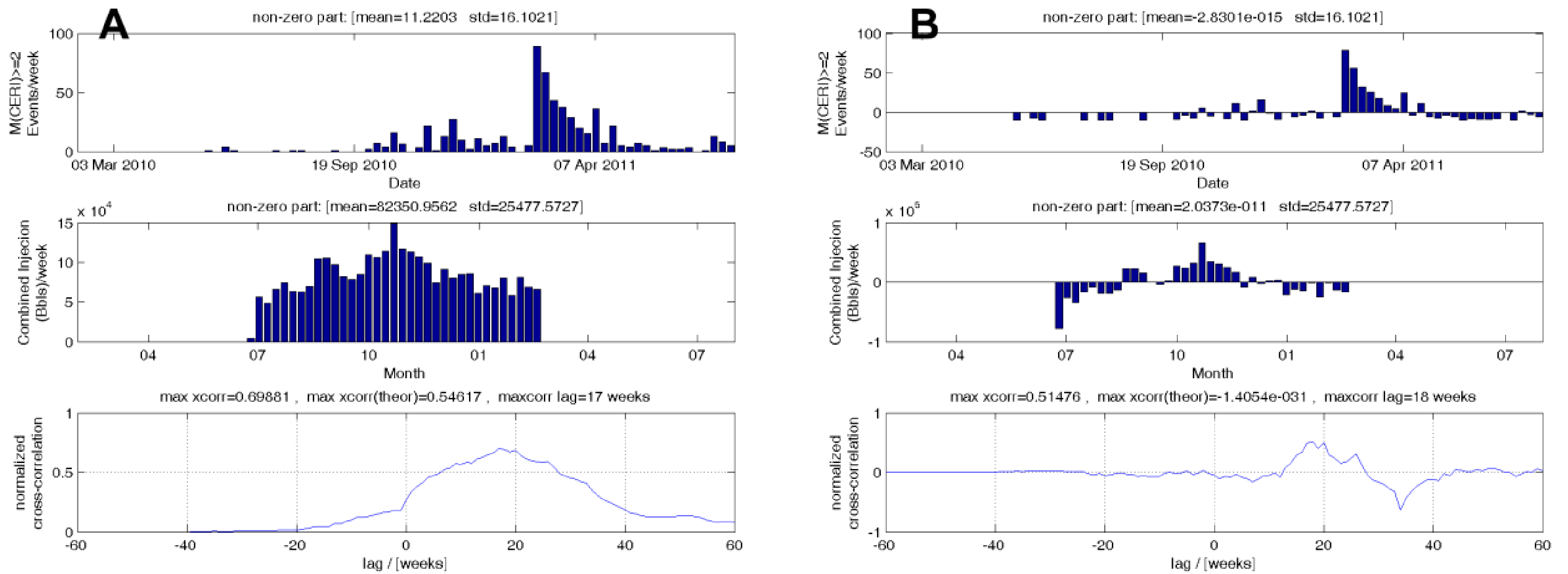

Figure 3 A: Normalized cross-correlation of earthquake frequency (USGS catalogue, M≥2) and combined injection volume at Edgmon and SRE wells per week. B: The Pearson's cross-correlation for the same data (combined injection rates taken from Horton, 2012, seismicity from Ceri catalogue: www.ceri.memphis.edu/seismic/catalogs/cat_nm.html).