Proceedings FORGE Symposium, Luxembourg 5 to 7 February 2013

'Second Order' Exploratory Data Analysis Of The Large Scale Gas Injection Test (Lasgit) Dataset, Focused Around Known Gas Migration Events

Dan Bennett[1], Rob Cuss[2], Phil Vardon[1,3], Jon Harrington[2], Majid Sedighi[1], Richard Shaw[2], Hywel Thomas[1]

[1]Geoenvironmental Research Centre, Wales
[2]British Geological Survey, England
[3]Delft University of Technology, Netherlands

Summary

Within large-scale experimental datasets a wealth of small scale information can typically be found. An example of such an experiment is the Large Scale Gas Injection Test (Lasgit). A toolkit has been developed to facilitate the investigation of the small scale or 'second order' detail contained within Lasgit's dataset. Results obtained through application include: quantifications of trends and seasonal effects; and second order gas transmission spurring from a primary macro scale transmission event.

1. Introduction

The Large Scale Gas Injection Test (Lasgit) is a field scale experiment run by the British Geological Survey (BGS) and is located approximately 420m underground at SKB's Äspö Hard Rock Laboratory (HRL) in Sweden. It has been designed to study the impact on safety of gas build up within a KBS-3V concept high level radioactive waste repository. Lasgit has been in almost continuous operation for approximately eight years and is still underway. Detailed information about the Lasgit experimental setup and procedures can be found in Harrington et al. (2008) [7] and Cuss et al. (2010) [5].

Lasgit is highly instrumented and frequently sampled. This has lead to a substantial dataset containing in excess of 14.7 million datum points. The data is anticipated to include a wealth of information, including many smaller scale or 'second order' features of interest that are not immediately apparent when considering the dataset as a whole. Given the magnitude of the dataset and the detail required in investigation, computational analysis is essential; however it is hampered by the non-uniform nature of the Lasgit dataset. To facilitate the second order analysis of such long-term, large-scale datasets a toolkit has been developed capable of accommodating non-uniform input [2].

2. Methodology

While specifically developed to be applied to the Lasgit dataset, the toolkit was designed to have generic applicability to long-term, large-scale time series based datasets, possibly containing non-uniformities. The resultant toolkit was applied to the Lasgit dataset both globally and locally in order to first identify and then investigate second order events.

The choice of capability incorporated into the toolkit and the development of methodologies used to achieve the chosen capability were guided by a requirement for minimisation of pre-requisite knowledge of the nature of information contained within a dataset before application. As such, the capability selection process favoured non-parametric techniques when information form is considered, and parameters related to the scales of interest to investigation (i.e. the desired second order scale) when they could not be avoided.

The toolkit’s capability includes:

- Event detection sensitive to second order scale behaviour
- Detection of disparate datum points (spike detection) for QC purposes
- Frequency content analysis
- Trend detection / signal component derivation
- A range of smoothing and averaging capabilities

Modification (generalisation) of a number of established signal processing procedures (e.g. [1,3,4]) was required to accommodate non-uniform input in order to provide the capability listed above. Notably, windowing functions, typically used in smoothing functions and moving averages, were adapted from spanning a fixed number of datum points to spanning a fixed length in time and operating on the variable number of datum points hence encompassed. This allows the toolkit to be applied on a scale fixed relative to the processes of interest in the system studied, regardless of data sampling rate. Additionally a Discrete Fourier Transform (DFT) procedure was generalised to accommodate an arbitrary inspection range in the frequency domain along with the removal of reliance on a fixed input sample rate.

Event and spike detection made use of simple statistical indicators (such as local standard deviation) coupled with the time windowing approach described above. This approach removes the need for knowledge of the form of events. This approach however does only highlight fluctuations in the data, essentially highlighting ‘candidates’ for events of interest warranting further investigation.

Singular spectrum analysis (SSA) [6] was used for trend detection / signal component analysis due to its non-parametric estimation of form within a time series. Macro scale trend detection was primarily used to determine a long-term behaviour of the system that could subsequently be subtracted from the dataset allowing smaller scale features to appear more prominently, thus potentially exposing information otherwise obscured.

Application of the toolkit to the dataset was multi-stage, with a global application for reference and to assess the second order behaviour in the context of the macro scale behaviour, a trend removal stage, and subsequent reapplication of the second order assessment with macro obscurity removed. A number of second order event candidates were then chosen for a ‘drill down’, in which a more classical ‘by hand’, focused investigation was undertaken assisted by higher resolution output from the toolkit. The event candidates chosen for investigation were prioritised by relative prominence and proximity to known macro scale gas movement events.

3. Results

Exemplar results are presented below depicting both the general output generated by the application of the toolkit and specific discoveries made as a direct result of the information exposure achieved.

Figure 1 provides an example of typical toolkit outputs when applied to a pore-water pressure sensor installed down hole in Lasgit. Identified second order events (depicted by grey impulses in Subfigure a) are identified by the prominence of local noise in the system with time. Frequency content (Subfigure b) clearly depicts an annual (seasonal) cycle inherent in the magnitude of the sensor record of approximately ±10kPa. Subfigures c and d depicts the SSA derived trend and the residual signal after subtraction respectively.

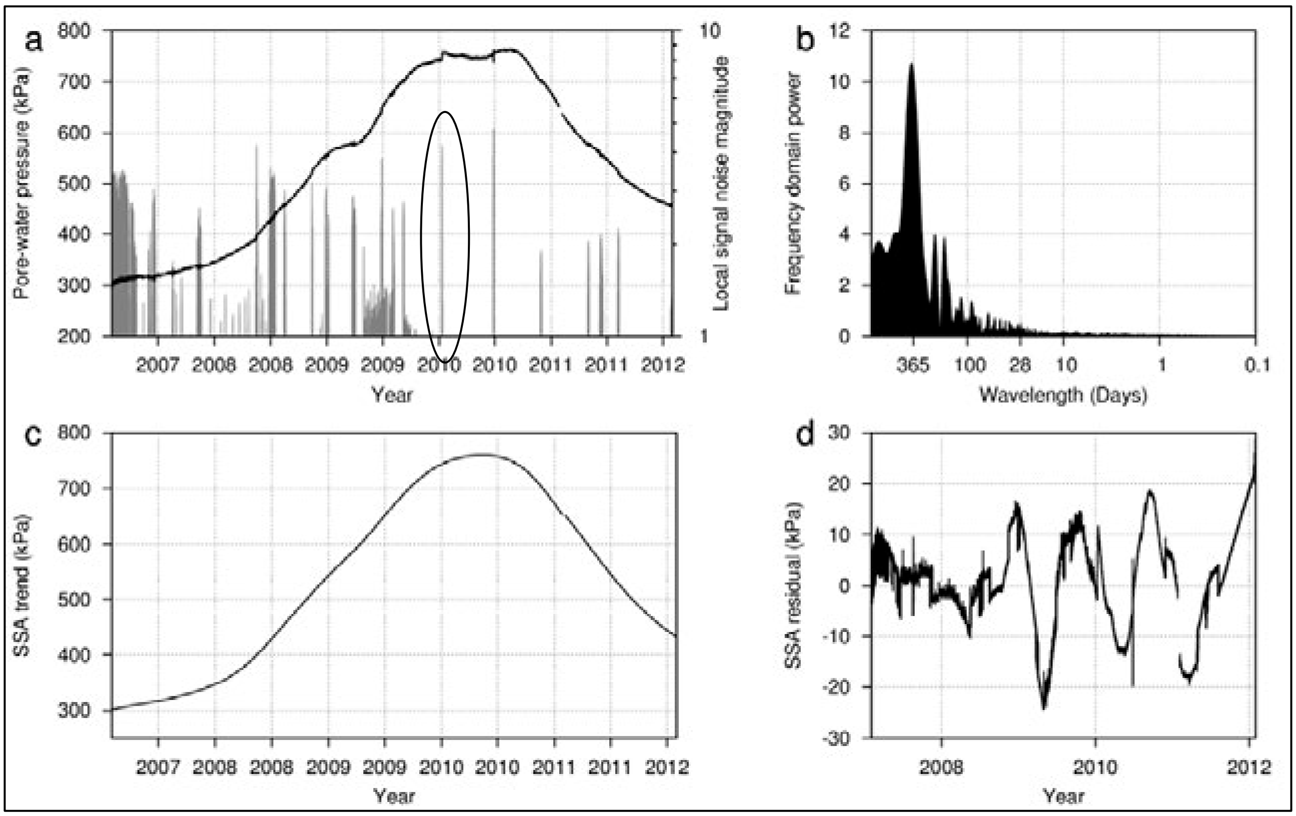

Figure 1. Example toolkit output when applied to a Lasgit pore-water pressure sensor. (a) The original time series with detected second order events depicted by impulses. (b) Frequency content of the de-trended time series showing annual (seasonal) variation. (c) and (d) the SSA derived trend and residual of the original signal respectively.

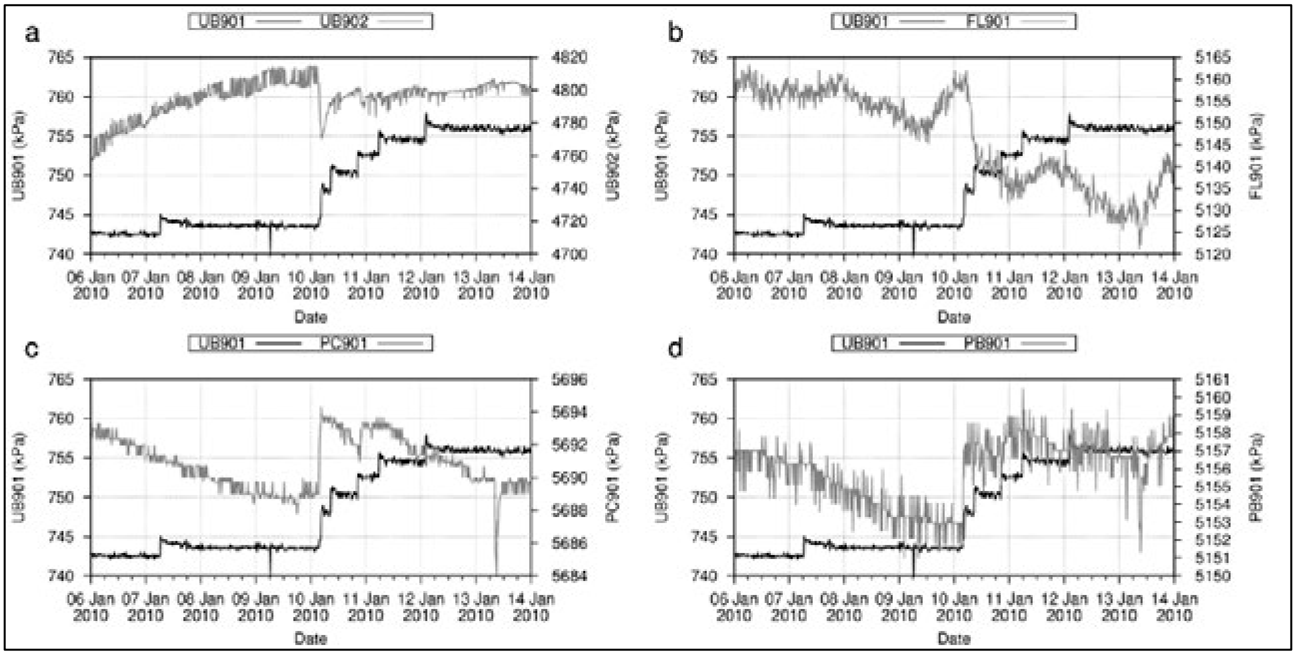

Figure 2. Sensors with second order event indicators in close proximity to the primary pore-water sensor being investigated. (a) a neighbouring pore-water pressure sensor that has experienced macro scale gas transmission. (b) The active gas injection filter. (c) and (d) nearby stress sensors.

Detailed scrutiny of the ten day period surrounding the impulse highlighted in Figure 1a is presented as an exemplar. Comparison of impulse results in the period identified four other sensors with second order events: the active gas injection filter; a neighbouring pore-water pressure sensor; and two nearby stress sensors. The records of these sensors are juxtaposed with the sensor record of the originally investigated pore-water sensor in Figure 2.

4. Discussion

The episodic pressure increase in the exemplar pore-water pressure sensor coupled with a decrease in pressure at the active gas injection filter and neighbouring pore-water pressure sensor that was previously the recipient of a macro scale gas transmission indicate a second order spurring flow has occurred from the original transmission path. The coinciding increases in stress locally indicate that dilatant flow was the mechanism for the spur. Other stress sensors in the region show no response to this event, suggesting it is a highly localised phenomenon.

5. Conclusions

The successful development of a toolkit capable of analysing a long-tern, large-scale, non-uniform dataset has been instrumental in facilitating the investigation of the Lasgit dataset. Small scale or 'second order' events have been identified as warranting detailed investigation with a high level of relevancy. Investigation of these events has lead to discovery of phenomenon that may otherwise be overlooked. An exemplar of such a discovery is spurring gas flow has been indentified within the Lasgit system.

6. Acknowledgements, References

The research leading to these results has received funding from the European Atomic Energy Community's Seventh Framework Programme (FP7/2007-2011) under Grant Agreement No. 230357, the FORGE project.

[1] Bagchi, S. and Mitra, S.K. (1999) The nonuniform discrete Fourier transform and its applications in signal processing. Kluwer Academic Publishers. ISBN: 0-7923-8281-1.

[2] Bennett DP, Cuss RJ, Vardon PJ, Harrington JF, Philp RN and Thomas HR (In Press) Data analysis toolkit for long-term, large-scale experiments. Accepted for publication in Special Edition of Mineralogical Magazine - Geological Disposal of Radioactive Waste: Underpinning Science and Technology for Radioactive Waste.

[3] Box, G.E.P. and Jenkins, G.M. (1976) Time series analysis: forecasting and control. Revised Edition. Holden-Day.

[4] Chatfield, C. (1989) The analysis of time series, an introduction. Fourth Edition. Chapman & Hall.

[5] Cuss, R.J., Harrington, J.F. and Noy, D.J. (2010) Large scale gas injection test (Lasgit) performed at the Äspö Hard Rock Laboratory: Summary report 2008. Technical Report TR-10-38, Svensk Kärnbränslehantering AB.

[6] Golyandina, N., Nekrutkin, V. and Zhigljavsky, A. (2001) Analysis of Time Series Structure: SSA and Related Techniques. Chapman & Hall/CRC.

[7] Harrington, J.F., Birchall, D.J., Noy, D.J., and Cuss, R.J. (2008) Large scale gas injection test (Lasgit) performed at the Äspö Hard Rock Laboratory: Summary report 2007. Technical Report CR/07/211, British Geological Survey.