



The inference of abrupt changes and noise levels in multivariate datasets

Kerry Gallagher (1), Malcolm Sambridge (2), Thomas Bodin (3), and Anya Reading (4)

(1) Université de Rennes 1, UMR 6118 Géosciences Rennes, Rennes Cedex, France (kerry.gallagher@univ-rennes1.fr), (2) Research School of Earth Sciences, Australian National University, Canberra, ACT 0200, Australia, (3) Seismological Laboratory, UC Berkeley, CA 94720-4767, USA, (4) School of Earth Sciences and CODES Centre of Excellence, University of Tasmania, Private Bag 79, Hobart, TAS, 7001, Australia

We present a method to quantify abrupt changes (or changepoints) in data series, as a function of depth or time. These changes are often the result of environmental variations and can be manifested differently in multiple data sets, but all data can have the same changepoint locations. The method uses transdimensional Markov chain Monte Carlo to infer pdfs on the number and locations of changepoints, the function values between changepoints and the level of noise associated with each dataset. This latter point is important when we have estimates only of measurement uncertainty, and it is not practical to make repeat measurements to assess other contributions to the data variability. We describe the main features of the approach and demonstrate its validity using synthetic datasets, with known changepoint structure (number and locations of changepoints) and distribution of noise for each dataset. We show that when using multiple data, we expect better resolution of the changepoint structure than when we use each dataset individually. This is conditional on the assumption of common changepoints between different datasets.

We then apply the method to real multivariate data, two geochemical data sets from peat cores, taken from NE Australia and eastern Tibet and a set of geophysical well logs from Australia. Under the assumption that changes occur at the same time for all data sets, we recover solutions consistent with those previously inferred qualitatively from independent data and interpretations. However, our approach provides a quantitative estimate of the relative probability of the inferred changepoints and the magnitude of noise on each data set, allowing an objective assessment of the significance of each change.