Geophysical Research Abstracts, Vol. 8, 02644, 2006
SRef-ID: 1607-7962/gra/EGU06-A-02644

A semi-parametric approach to the identification of extreme events prior to automatic mapping.

D. D'Alimonte and **D. Cornford**

NCRG, SEAS, Aston University (d.cornford@aston.ac.uk)

Automatic mapping, particularly for emergency situations is a complex task. The Spatial Interpolation Comparison 2004 (SIC2004) exercise challenged participants to automatically map radioactivity measurements from a series of observations, including a simulated nuclear release. In the simple interpolation cases all algorithms did a reasonable job, but in the simulated release, which contained a very small number of extreme values most algorithms strongly over-smoothed the release, which could have serious consequences in any operational early-warning system. An answer, which many participants suggested, is to first identify the novel or unusual observations. In this talk we outline a new approach to the identification of 'novel' data given a known population distribution. The approach is based on using a flexible Gaussian mixture model to represent the known 'background', or population, process. The 'background' process is assumed to be known from existing measurements. In the context of radioactivity mapping this represents the point probability distribution of background radioactivity levels. In this initial work we do not consider the spatial aspect, however it is a natural extension to incorporate this. Once the 'background' probability distribution is known, we assume that another unknown process (which may represent several physical effects) could be operating. Our aim is identify observations from the 'novel' process. We accomplish this using a Kolmogorov-Smirnov test to identify the proportion of observations that are derived from the 'novel' process. We are then able to separate the 'novel' and 'background' observations in a Bayes optimal manner. We show how on synthetic data the scheme outperforms standard approaches based on novelty detection methods, or outlier detection approaches. We explore how the scheme behaves as a function of the Kullback Leibler distance (relative entropy) between the 'background' and 'novel' probability distributions. We apply the algorithm to data used in the SIC2004 competition, and show its ability to detect the nuclear release.