Supplement of Earth Syst. Dynam., 8, 677–696, 2017
https://doi.org/10.5194/esd-8-677-2017-supplement
© Author(s) 2017. This work is distributed under
the Creative Commons Attribution 3.0 License.

Supplement of

Multivariate anomaly detection for Earth observations: a comparison of algorithms and feature extraction techniques

Milan Flach et al.

Correspondence to: Milan Flach (milan.flach@bgc-jena.mpg.de)

The copyright of individual parts of the supplement might differ from the CC BY 3.0 License.
Supplementary Material 1  Parameterization of Recurrences and Kernel Density Estimation

**Figure S1.** We test different choices of $\sigma$ (Recurrences (REC)) or $\varepsilon$ (for Kernel Density Estimation (KDE)) in a small simulation (500 repetitions) trying to detect a BaseShift. $\sigma$ (or $\varepsilon$, respectively) is varied between the 0.05 and 0.95 quantile of the distribution of values of the distance matrix. The Area Under the receiver operator characteristics Curve (AUC) is computed for each parameterization. Results exhibit constant AUC values for KDE within the testing range of $\sigma$. In contrast REC is more sensitive to the choice of $\varepsilon$, although it might yield slightly higher AUC values in case of optimal chosen $\varepsilon$. 
Figure S2. Effect of different data properties on the 3 best detection algorithms (KDE, REC, KNN-Gamma) presented as AUC difference to the UNIV control for the event types (a-d). Details for each algorithm reveal that KNN-Gamma is often less affected by 'difficult' data properties like CorrelatedNoise or MoreIndepComponents, i.e. KNN-Gamma is more adaptive than the other two algorithms.