

Interactive comment on “Multivariate Anomaly Detection for Earth Observations: A Comparison of Algorithms and Feature Extraction Techniques” by Milan Flach et al.

Anonymous Referee #1

Received and published: 9 December 2016

*** General comments**

The manuscript describes a systematic and comprehensive study of methods for extraction of anomalies and features from artificially-generated multivariate datasets. The presentation is clear, the manuscript is well written, and the study is sound as a comparison of methods for multivariate data analysis, though its value for earth observations is in my opinion not convincing.

Although I understand the rationale for using artificial data, particularly when comparing the performance of different methodological approaches, the artificial events that are considered in the study seem to be unrealistically exaggerated, particularly the amplitude change in the seasonal cycle (Fig. 2 c) and and the change in variance (Fig.

C1

2 d). For example climate related changes in the seasonal cycle or in variance are far more subtle (in terms of magnitude) and much more difficult to identify in real data than the ones exemplified in Fig. 2.

I'm uncomfortable with the term "Complication" used throughout the manuscript to refer to specific characteristics of the artificial data. For example a seasonal cycle can hardly be seen as a complication, it's a feature of the data, not necessarily something complex as it is implicit from denoting it a "complication".

I think that the comprehensiveness of the study is a strength and paradoxically maybe the greatest weakness of the work, because the results need to be necessarily presented in a highly summarized way, here as difference in AUC values (which itself are already a reduction of a ROC curve to a single number) to a univariate approach without "complications" (UNIV). I don't doubt the technical correctness of the results, but in my opinion it's difficult to assess their relevance, particularly in the context of real earth observation data. I find the conclusions of the study quite obvious and realistic (the importance of deseasoning or dimensionality reduction), whether they would require such a wide statistical study on a artificial data farm is not obvious to me.

*** Specific comments**

If I understood correctly the length of the generated time series is only of 300 time steps (appendix B), which may be in itself a major factor influencing the performance of some of the methods.

Although I'm keen on the transference of methodological approaches across different areas, and in this case the use of statistical process control (SPC) methods typically used in other contexts (e.g. industry), the restriction of feature extraction methods to the ones used in classical multivariate SPC seems to me an unnecessary restriction. Many feature extraction methods, e.g. wavelets, are routinely used with earth observations precisely because they perform very well in that kind of data.

C2

* Technical corrections

Page 9, line 31: cdot notation

Interactive comment on Earth Syst. Dynam. Discuss., doi:10.5194/esd-2016-51, 2016.