Earth Syst. Dynam. Discuss.,
doi:10.5194/esd-2016-51-AC1, 2017

Earth System Dynamics Discussions

Open Access

ESDD

Interactive comment

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

Milan Flach et al.

mflach@bgc-jena.mpg.de

Received and published: 3 March 2017

In the following we are responding to RC1 by an anonymous referee.

* General comments

Reviewer: 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.

Response: We would like to thank the reviewer for the positive feedback. Regarding the

Printer-friendly version

Discussion paper

reviewer's concern about the values for EOs, we consider our study as relevant in this context, because current anomaly or extreme detection in Earth observations is mostly done with peak-over-threshold techniques (p. 2, l. 22). These do not consider the multivariate and potentially non-linear correlation structure between multiple variables which we have in EOs. It is therefore an important motivation for our paper to provide a sound basis for alternative and more general approaches. This paper systematically analyses and proposes several algorithms and workflows which consider the structure among multiple variables and furthermore might also reveal novelties about so called compound extremes beyond known patterns, i.e. anomalies where none of the single variables is extreme itself, but their combination is anomalous and leading to an extreme impact. The consideration of compound events does play an increasing role within the community, but is typically confined to known compound events (e.g. heat and droughts) and not very generic. The comparison is performed on artificial data, which were explicitly built to mimic current EOs as ground truth is missing for 'real' EOs (p.3, l.25). Furthermore, available time periods as well as sample sizes are rather small for detecting anomalies in 'real' EOs, which empowers the use of an ensemble of artificial data for method comparison. The application of the proposed workflows to EOs will follow soon.

Reviewer: 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 that the ones exemplified in Fig. 2.

Response: Please consider that Figure 2 is only an illustration. As described in the manuscript (p. 6, l. 9), we analyse each type of anomaly across 20 different magnitudes - from very minor perturbations to entirely exaggerated values. The generic formula B1 shows that we are effectively exploring the full range of perturbations between very subtle changes (Appendix B: k = 0.2) to exceptionally high changes (k = 4.0) as if it was a model parameter sensitivity analysis. In the revised manuscript, we will add an additional sentence for clarification, explain it in the figure caption and show more realistic magnitudes in Figure 2.

Reviewer: 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".

Response: We agree with the Reviewer, that the term 'complication' is far away from being optimal. However, 'data features' might be misunderstood with 'feature extraction techniques', which we wanted to prevent. Therefore, we suggest to rephrase 'complications' into 'data properties' in the revised version.

Reviewer: 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.

Response: We thank the reviewer for this comment. It has two components: (1) the presentation of the results, and (2) the overall relevance. (1) Indeed, we decided to highly summarize the results in the main part of the study. However, more detailed results, for instance the effect of different magnitudes on specific event types and complications can still be inferred from the Appendix Fig. A1. (2) The importance of dimensionality reduction as one way to enhance the performance of anomaly detection algorithms (p.18, l.12) has not been shown before to the best of our knowledge, in particular not for EOs. We are convinced that a highly multivariate system like the Earth with seasonality and potentially non-linear dependencies among the variables requires specific workflows like the ones we propose, i.e. the results of our study are relevant in this context. Currently we are working on applying the algorithms on ‘real’ EOs with very promising results, i.e., results that capture the major known events globally. Our overarching objective is developing workflows to open a path to a series of scientific studies exploring extreme compound events in depth.

* Specific comments

Reviewer: 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.

Response: Indeed, the length of the time series is a factor influencing the performance of the multivariate anomaly detection algorithms. One crucial point of our study is, that Earth observations are typically short. We seek to understand the performance characteristics of various algorithms and feature extraction methods on short time series. Furthermore, please note that Ding et al. (2014) studied this effect in detail, changing the data set size between 50-30000. The only algorithm on which the size of the data set had a remarkable effect was the Support Vector Data Description (SVDD). SVDD performance increased with the size of the data set. However, even the best performance of SVDD was worse than the other algorithms. Therefore, we conclude that the size of the data set is not influencing the results of the top algorithms (KDE, KNN, REC, T2). We include this aspect in a second version of the paper

Reviewer: 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.

Response: We are aware that the list of feature extraction algorithms as well as the list of anomaly detection algorithms can hardly be complete. We did not restrict the feature extraction methods only to the ones used in classical statistical process control. We also included non-standard ones from process monitoring in industry (e.g., Independent Component Analysis) and of course from univariate extreme extreme event detection (e.g., subtracting the mean seasonal cycle) (p.6, l. 28). We agree with the reviewer that wavelets perform very well on EOs, e.g., for extracting information about dominant frequencies in the data. However, event detection is another task. We are not aware that wavelets improve the detection rate of multivariate anomalous events, but we will consider this as an interesting aspect for future research.

* Technical corrections

Reviewer: Page 9, line 31: cdot notation

Response: We changed it.

References: Ding, X., Li, Y., Belatreche, A., & Maguire, L. P. (2014). An experimental evaluation of novelty detection methods. Neurocomputing, 135(C), 313–327. http://doi.org/10.1016/j.neucom.2013.12.002

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