Earth Surf. Dynam. Discuss.,
https://doi.org/10.5194/esurf-2020-34-RC2, 2020

Interactive comment on "Coastal Change Patterns from Time Series Clustering of Permanent Laser Scan Data" *by* Mieke Kuschnerus et al.

Anonymous Referee #2

Received and published: 21 October 2020

This paper presents informative results of a study to automatically differentiate geomorphic surface change dynamics on a coastal back beach. The analysis is comprehensive and robust and the findings are well presented. The manuscript requires some minor revisions before it can be accepted, but I would also urge the authors to better highlight (if possible) what exactly the advantages are of applying a clustering technique. I actually found the results in the end rather underwhelming, because if you remove the two clusters associated with the bulldozer works the results really just yield four different trends of geomorphic change: stable surfaces, steadily eroding and steadily accreting surfaces, and fluctuating surfaces (noise). These four trends are hardly surprising and can just as easily be detected from simple erosion/deposition mapping. The limitation of the application here may be related to the fact that the anal-

Printer-friendly version

Discussion paper

ysis was directed to 6 clusters. The manuscript does not give objective or quantitative justification for this decision, other than that the results with 6 clusters seemed good or reasonable to the authors (L304), but this then effectively preempts the possibility of finding something new or interesting. Maybe 10 clusters could have revealed some more interesting trends, for example.

Below follows a list of issues and queries that can be resolved in a minor revision.

L39: define what you mean with "epochs" L39: why "high-dimensional"? This is just a 4D dataset. At this point in the paper it is not clear to the reader yet that you are going to define your data in a multi-dimensional space with the dimensions defined by the snapshots (epochs?). L40: citation here is 15 yrs old, can you refer to more recent literature on the challenges of data mining? L85: why the second representation in terms of range data? The fact that it's the native data format doesn't really give us an actual justification. Your second criterion is about geomorphic deformation, which presumably relates specifically to height changes, hence the cartesian grid seems most suitable to that. Your results later on essentially show that the range format is simply not useful, so you could achieve a great simplification and a more focused message here if you strip out all this stuff about the range format and just report the results related to the cartesian data... Fig.1 says data from 2019, but text says data was from Dec-2016 to May-2017? L122 and earlier: not suitable to use x_0 and y_0 for identifying a test location as it has nothing to do with zero. Suggest subscript 't' or even 'test'. Table 1 and associated text: we really need more info on these test areas: are they single points? areas? If latter, what size. Then, the test statistics is not informative. Stdev is not sufficient, you should be able to calculate the standard error and the associated 95% confidence intervals around the mean height. Also, why the difference in N vs S? This requires discussion. In the manuscript you present these test results here, but in the results and discussion there should be further reflection on the potential impact of the error on the cluster classifications. L145: you remove the mean in the cartesian format, but not in the range format, why do this in the cartesian grid?

The same logic you use there should somehow apply to the range time-series? More crucially however, in the later results it seems as if the mean was in fact NOT removed, for example in figure 9b, the centroids all have distinct absolute elevations, surely this can only be possible if the mean was not removed? Otherwise the centroids should all be fluctuating around zero? Eq 2 and L145: notation is not suitable; delta usually refers to a real discreet difference; suggest using prime ‘ as the fluctuating component. Section 3.2: from later on in the results I get the impression that euclideanw as used for k-means and aggregation, while correlation distance was only used for DBSCAN, si this correct? If so, this needs to be stated here. L179-180: then why don’t you standardize your data? L245: but you also evaluate Euclidean distance in DBSCAN? (or not, see above?) so what are the clusters in that case? Figure 7: something is seriously going wrong at the white polygons left of centre. The height changes don’t match at all. In A the original elevation in this area is around 5.6 m, in B those two white polygons appear to be at 6.4, so this should yield a height change of 0.9 (red) in C. Or are these polygons areas with No Data? If so the colour scales need to be completely different so as to avoid white being part of the scale (so that it can then indicate no-data). Fig7c: show contour lines of height changes beyond significant (as basis for additional clusters?) Stable points not shown? But not clear then how much of area has been allocated properly? L304: why six? Not enough justification. What is ‘good’? Why not 8 or 10? Isn’t there a statistic to tell you when to stop clustering? Figure 9b: needs horizontal gridlines; labels should be added to lines, rather than a legend (because the sequence in the legend doesn’t match the sequence from top to bottom). Vertical axis labels don’t make sense: I don’t understand how you can have real values asl here for the centroids when the original time-series was mean-subtracted? L341: this is the only place where you really say that there are no results like 4.2 for the range data; this needs to be made more explicit in Line ∼300

L372-374: please elaborate a bit on this here, please summarize or give us a taste of the cause for this ‘curse’. 5.2: it only gets clear to me here that you use Euclidean for k-means and aggregation, and corr for dbscan! Conclusion: not really clear what all this work benefits; if you remove the clusters associated with the bulldozer work you basically end up with 4 obvious trends: stable, erosion, accretion, noise. This is a bit underwhelming. . .

Interactive comment on Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2020-34, 2020.