Coastal process understanding through automated identification of recurring surface dynamics in permanent laser scanning data of a sandy beach

Author Comment - EGUSPHERE-2025-4964

Daan Hulskemper, José A. Á. Antolínez, Roderik Lindenbergh, Katharina Anders

This comment is a response to both anonymous reviewers. First, we want to thank both for taking the time to read the manuscript and provide us with many constructive comments and suggestions. We feel that incorporating these suggestions benefits the manuscript greatly. Additionally, thanks to the associate editor, Giulia Sofia, for ensuring the review.

In the following, we reply to all major comments and summarize how we address these and the in-text comments in our revision.

After this, we will prepare the revised submission, which we hope will be accepted and published in *Earth Surface Dynamics*.

Reviewer #1:

1. ***However, while the data-driven methodology is a strong asset, the manuscript would benefit from a more rigorous geomorphological justification for the specific thresholds and technical choices made during the analysis.***

This is a good suggestion, which also comes back in comments of reviewer #2. We provide a justification for the choice per threshold where needed, as suggested in the in-line comments of both reviewers. Additionally, the choice for algorithms will be further clarified in the introduction and methods, and reflected on in the discussion, as suggested by reviewer #2. Justification will be elaborated in response to later comments.

2. ***One area for improvement concerns the integration of environmental variables. While the authors claim to combine these variables with the 4D-OBCs, the current analysis remains largely qualitative, relying on manual comparisons and interpretations for only eight selected clusters. To avoid giving readers the false impression of a fully automated or integrated analysis, the abstract and introduction should be rephrased to accurately reflect this level of manual intervention.***

This is a valid point; the extraction is automated; the interpretations are not. We will revise the abstract and introduction to reflect this as follows:

L.14–15: " This approach demonstrates the potential of integrating PLS and unsupervised learning to characterize complex surface dynamics, through a fully automated extraction and classification workflow. While the interpretation of clusters and their relation to environmental variables in this study is performed through expert-based analysis, the methods provide a framework for targeted, data-driven investigation and prediction of morphodynamic processes in high-resolution 4D remote sensing datasets. "

And:

L.116: "In future work, the structured dataset of grouped surface dynamics can be linked to environmental drivers through automated data-driven methods, such as threshold analysis, time series correlation, or physics-informed machine learning, enabling systematic and scalable investigation of causal relationships of surface dynamics."

3. ***Furthermore, the authors should explore or at least highlight a bit earlier whether/how a more data-driven comparison, such as calculating direct correlations between clusters and environmental drivers. At the moment, some discussion is performed solely towards the end of the manuscript.***

Through the edits mentioned above we already reflect on these points in the abstract and introduction.

4. ***The spatial relevance of the environmental data requires further clarification. The monitoring stations used appear quite distant from the PLS site, raising questions about whether these data streams remain spatially correlated with the local beach forms.***

This is a good observation. The use of monitoring stations relatively far from the PLS site is in our opinion justified with respect to the scale and meaning of analysis in our application. The meteo and hydrodynamic parameters have appropriate temporal correlation with the location of the PLS, as we do find clear relations between 4D-OBC activity and the environmental variables. The meteorological and hydrodynamic variables recorded at Hoek van Holland, IJmuiden, and Scheveningen are, however, not identical to those at Noordwijk on a minute-by-minute basis, but they show similar trends and magnitudes of variation at the (sub)daily timescales relevant to our analysis.

We will mention this in the data section:

*L154: The monitoring stations are located at distances up to ~35 km from the PLS site (Fig. 2). Meteorological and hydrodynamic conditions along this stretch of the Dutch North Sea coast are, however, highly spatially coherent at the timescales most relevant to this study. Wind speed over the southern North Sea has a decorrelation length on the order of hundreds of kilometres for daily variability (*Sušelj *et al., 2010), with Hoek van Holland and IJmuiden showing similar wind statistics (Coelingh et al., 1998). Wave conditions along the Dutch coast are also spatially homogeneous, with observed deviations in mean significant wave height on the order of only 0.2 m (Wijnberg, 2002)*

For future data-driven correlation and prediction studies, locally measured or spatially interpolated/downscaled environmental variables would however be interesting and preferable. We will add this as recommendation to the Discussion:

L674: A limitation of the environmental analysis is the use of monitoring stations located up to 35 km from the PLS site. While the large-scale spatial coherence of wind and wave conditions along this coast justifies this for our scale of analysis, local effects and subdaily temporal variations will not be fully captured. For future data-driven correlation and prediction studies, locally measured or spatially interpolated/downscaled environmental variables would be preferable.

5. ***The authors should also address whether the model needs to account for sediment availability, specifically considering inputs from along-shore transport or foreshore regions, which are critical drivers of geomorphic change.***

This is an important point. The present study is fully observation-based. We will add a recommendation to include additional measurements in the Discussion:

L674: Sediment availability from along-shore transport or the foreshore is not directly measured as potential environmental drivers in this framework but may be implicitly reflected in the extracted 4D-OBCs and history of environmental variables. In a future PLS setup, explicit monitoring of these boundary conditions (through e.g., regular monitoring with bathymetric sensors, wider-area UAV surveying, or acoustic sediment transport sensors), or exploration of the prediction thereof through the history of environmental variables in combination with for example LSTMs, could aid further explanation of found variations in surface dynamics.

6. ***Finally, regarding data management, the authors suggest partitioning large datasets based on existing gaps. However, for continuous datasets where gaps are few or non-existent, a clearer strategy is needed; the authors might***

consider discussing whether a sliding window approach would be more appropriate than arbitrary partitioning to avoid artificial hard breaks in the process analysis.

We agree that this is an important consideration of partitioning strategies. We will mention this in the manuscript:

L222: " In cases where data gaps are few or non-existent, the gap-based subsetting strategy applied here would result in very long subsets or a single subset, which may exceed available memory. For such datasets, alternative strategies should be considered, for example, a sliding window breakpoint detection approach or an event-based partitioning triggered by periods of low morphodynamic activity, detected with, for example, trend analysis (Kuschnerus et al., 2024). These would provide a more data-driven approach; however, they come at the cost of the parallelisation efficiency that gap-based partitioning provides."

Reviewer #2:

1. ***However, I am missing some information and rationales for technical aspects that were not presented (see comments below). Occasionally, a more critical discussion on how meaningful the clustering is, and how well it might generalize to other settings or data, could be included.***

These are good points which will be processed in the revision. Details are elaborated below point-by-point.

2. ***One beneficial addition would be the rationale for using Self-Organizing Maps (SOMs) and hierarchical clustering. While SOMs seem a natural choice given the spatial nature of the data, other options for dimensionality reduction exist. Even more alternatives for hierarchical clustering are conceivable (e.g., DBSCAN, DBCV, k-means); therefore, arguments for why the specific methods (and their specific settings/parameters) were chosen would strengthen the study.***

We agree that the reasoning behind this choice is an important aspect of the study that has not been explicitly explained. We will add these justifications to the methods:

L294: "SOMs are chosen because they provide a largely topology-preserving projection of high-dimensional data onto a 2D lattice, making inter-sample similarity visually inspectable without requiring a prior definition of the number of clusters. This is particularly *suited to our dataset where surface dynamics form a continuum of types depending on research focus rather than discrete natural classes. Hierarchical clustering is subsequently applied to SOM nodes because it allows cluster boundaries to be explored at multiple levels, which makes them especially useful in the scale-dependent coastal setting, where different research questions call for different levels of morphodynamic detail to be explored."*

Additional clarification of SOM size parameter selection and transferability:

L324: "The SOM lattice size is chosen to provide sufficient resolution to distinguish between sub-types of dynamics within the intertidal, berm, and backshore zones, while keeping the number of nodes visually interpretable. A larger lattice could be beneficial for longer time series with a large diversity in surface dynamics."

Additional comment on transferability of hierarchical clustering settings:

L349: "In other environmental settings, the value of a distance threshold at which clustering is useful can be different, depending on dataset size and the diversity of dynamics. The silhouette-based threshold selection provides a data-driven means to adapt this threshold accordingly."

3. ***Another aspect that could be strengthened is the confidence in the robustness and meaningfulness of the clustering in this study. A key element would be to add information on how interpretable the other clusters are, beyond the 8 analyzed clusters. This is also related to the previous point, because if some of the clusters are not really interpretable or well defined, improvements with different algorithms or parameters (or even feature engineering) might be possible. Therefore, the process of selecting and tuning these becomes more central. A way to allow for visual evaluation of the clustering could be i) adding a dimensionality reduced plot of the 4D-OBCs as a PCA and/or t-SNE; and ii) adding a dendrogram for the clusters. Such representations have been widely used in other fields (e.g., biology or biomedical research) for similar clustering approaches of high-dimensional data.***

We agree that extending this discussion in the paper provides clearer direction for improvement. A clustering dendrogram and projection of SOM nodes and 4D-OBC feature vectors in PCA spaces will be made and added as supplement. Additional interpretation of these in terms of the validity, robustness and interpretation of the clusters will be discussed in section 5.2, after L568.

4. ***A more critical discussion of the role of algorithm choice and parameterization, as well as the inherent limits of clustering without external validation, could be added in section 5.4. Likewise, the limitations from the need for manual validation (and the resulting avenues for future studies) could be expanded (currently in lines 672-674).***

In line with our responses to previous comments, we add further methodological reasoning in the methods. In the discussion, we will extend the evaluation and comparison to other methods (currently in Section 5.2, 562-568), to aid direction of improvement of the framework.

L565: "Future clustering applications may benefit from using density-based alternatives such as Hierarchical Density-Based Spatial Clustering of Applications with Noise (Campello et al., 2015), which can identify clusters of varying density and flag low-density nodes as noise rather than forcing them into a cluster as occurs in our workflow. Such methods would however require alternative validation metrics suited to density-based clustering, such as the Density-Based Clustering Validation score (DBCV, Moulavi et al., 2014), in place of the silhouette score used here. Other simpler methods like k-means clustering offer another alternative but requires the number of clusters to be predefined and assumes approximately spherical, equal-variance clusters, assumptions unlikely to hold for the complex varying dynamics found here."

Limits (but also advantages) of clustering without labels is indeed also an important point to elaborate on. We will add an extended discussion on this in place of the last paragraph of Section 5.4:

L672: "The reliance on unsupervised clustering avoids the need to predefine surface dynamic classes, which on the one hand is an important advantage in a natural coastal system where class boundaries are continuous and scale-dependent. This, however, means we require manual validation, and there is no quantified external validation against labelled ground truth. This makes it difficult to assess whether a given parameterisation is optimal or to compare methods extensively. Future work developing benchmark datasets of labelled 4D-OBCs, derived for example using video imagery or expert annotation of known events, could enable quantitative validation, data-driven feature selection, and greater confidence in the transferability of the workflow to new sites."

In-line comments

The specific in-line comments by reviewer #1 and #2 will all be processed in the revised manuscript. The comments on specification/phrasing will be handled in the revised manuscript directly and not be replied to here. Also, minor comments related to major comments addressed above will not be elaborated here but handled in the revised manuscript. The larger specific questions/comments are further addressed here:

Reviewer #1

1. **L. 117-118: It needs to be clearly demonstrated how the proposed method is transferable to other monitoring platforms and different spatial scales.**

Agreed, the transferability lies in that it only requires a time series of topographical point clouds, which can be acquired in any way. Edit:

L117:" The dataset used in this study is acquired in a PLS setup, but the method is generalizable to any time series of topographic point clouds, regardless of acquisition platform, such as multitemporal UAV, airborne or satellite topographic measurements, as long as sufficient temporal frequency and spatial resolution are available to detect the target surface dynamics."

2. **Fig. 2: Please add a scale bar to the map so that the reader can assess the actual distances between the meteorological stations and the PLS site.**

We will add this, also in Fig. 3.

1. ***Reviewer #1:*** **L. 177-178: Could you clarify the reasoning behind the assumption that smaller scale changes are not considered geomorphological changes?**

This is a valid observation; our phrasing here was misleading. We fully agree that smaller-scale changes are also geomorphological changes. Our statement in the manuscript is rather meant to indicate that we do not consider them in our analysis at the scale that is below the observation capability of our PLS monitoring. We will revise the phrasing in the manuscript to:

L177-178: "This spatial resolution is deemed appropriate for the surface dynamics we aim to capture (>10 m^2). Morphological changes below this resolution fall within the positional uncertainty of the PLS and cannot be reliably distinguished from sensor noise. We therefore do not consider them in the present analysis."

2. ***Reviewer #1: L. 180: Regarding the choice of a one-week averaging window, please justify this specific timeframe and address whether there is a risk of smoothing out significant short-term signals.***

The one-week averaging window follows the approach of Anders et al. (2020, 2021), where it was found to effectively suppress sub-daily noise while preserving multi-day surface dynamics of geomorphological relevance. We will edit the manuscript as follows:

L179-183: "The obtained time series are smoothed to further ensure outliers like objects and humans on the beach, and measurement errors are not affecting the 4D-OBC extraction. This smoothing is done using a temporal median averaging window of 168 h (1 week), following Anders et al. (2020, 2021), where it was found to effectively suppress sub-daily noise while preserving multi-day surface dynamics of geomorphological relevance. By design, changes occurring on timescales shorter than the window are not captured in the smoothed time series, this is a consequence of the chosen observation scale, consistent with the minimum detectable event duration of 24 h. Where higher-frequency reliable data is available, a shorter filtering window could be used. This filtering also interpolates temporal gaps in the data, e.g., when data is missing due to rainfall, tidal water level, or technical problems."

3. ***L. 187: Please provide specific numerical values to define what is considered too large in this context.***

We realize that "too large" is a misleading phrasing here, as there is no direct threshold used in the breakpoint detection. The discrepancy threshold depends on a penalty parameter that determines how strict the change point detection is. We will rephrase it as follows:

L187: "If the discrepancy between the median of the first half of the window and second half of the window exceeds a penalty-driven discrepancy, a breakpoint is detected (Truong et al., 2020). "

4. ***Reviewer #1: L. 188: How does the model account for geomorphic processes that do not return to the original elevation, such as permanent erosion or dune formation? It might be beneficial to mention this briefly here, even if the authors expanded upon this in the discussion, to resolve potential reader confusion early on.***

Processes that result in net permanent elevation change (e.g. dune growth) are indeed not captured as 4D-OBCs by design, since the algorithm is designed to detect temporary deviations that return to a baseline. We will add a short note:

L193: "It should be noted that processes resulting in net elevation change such as dune growth or longer-term beach erosion trends are thus not captured as 4D-OBCs by design. The algorithm detects temporary deviations that return to a baseline elevation. The study of such longer-term or permanent geomorphic trends requires complementary analysis methods, such as bitemporal DEM differencing or elevation time series trend assessment (Kuschnerus et al., 2024)."

5. **L. 195: Please explicitly define the parameter used for the homogeneity criteria.**
   Edit:

L.195: "This is done by first sorting all the seeds based on their neighbourhood homogeneity, i.e., the mean similarity of the eight neighbour locations based on the Dynamic Time Warping distance (DTW, Anders et al., 2021)"

6. ***Reviewer #1: L. 200: Why was an area of 10 m^2 selected, and how does this specific size relate to the scale of the expected geomorphic processes?***

The 10 m^2 threshold is a practical lower boundary rather than a geomorphically derived one. In theory, valid 4D-OBCs could be defined at arbitrarily small scales, down to the limit imposed by point cloud density and accuracy. However, reducing this threshold leads to a rapid increase in the number of detected 4D-OBCs (Vos et al., 2023), the majority of which represent small-scale fluctuations at the margin of sensor precision. A 10 m^2 minimum was selected to focus the analysis on surface changes of a spatial scale relevant to meso-scale beach morphodynamics, consistent with the resolution and accuracy of the dataset, while keeping the number of objects computationally and analytically usable.

This will be mentioned shortly in-text:

L200: If the region is smaller … it is not used, which is a practical choice in terms of the spatial data resolution at hand and the size of target change processes. Reducing this threshold leads to a rapid increase in detected 4D-OBCs, most of which represent small-scale fluctuations at the margin of sensor precision (Vos et al., 2023).

7. ***Reviewer #1: L. 236-238: Would it be possible to consider filtering the 4D-OBCs individually, such as filtering the margins of an object rather than only using internal variability for noise identification and subsequent removal?***

This is an interesting idea, that could be explored in future work. We will add in the discussion:

L639: "Which could involve filtering individual 4D-OBCs at their spatial margins after or during region growing. For example, testing whether added corepoints show a coherent direction of elevation change relative to the seed rather than relying on the DTW distance and aggregate internal variability for filtering of the whole object as currently done. This could improve the coherence of the extracted objects and reduce the proportion of ambiguous 4D-OBCs"

8. ***Reviewer #1: L. 259-260: Might a centroid be more suitable for the analysis.***

A centroid could also be suitable and its incorporation could be tested in future work. However, as the seed is conceptually the most representative time series of the 4D-OBC (the highest ranked temporal segment), this should indicate the location of the most defining part / major change of the surface dynamics, which we found gave the suitable subsequent grouping. A weighted centroid could also be considered in future work.

9. ***Reviewer #1: L. 271-275: I am afraid that this paragraph is unclear to me; would it be possible to scale or normalize the volumeTS to improve interpretability?***

To clarify we will rephrase it as follows:

L271-285: "The volume time series (volumeTS) of individual 4D-OBCs differ in length because their durations differ. To enable comparison in the clustering step, all volumeTS are resampled to a common length N (the average duration across all 4D-OBCs) using linear interpolation. As a result, the volumeTS contributes N dimensions to the feature vector, whereas the remaining scalar features (duration, magnitude, etc.) each contribute only one dimension. To prevent the volumeTS from dominating the distance computation only through its higher dimensionality, the scalar features are multiplied by N, so that all features contribute equally during clustering. All features are then scaled from 0 to 1 using min-max normalisation to ensure equal weighting across features of different units: EQUATION. Where [..] scaling."

10. **L. 308: Please specify how the weight vectors are initialized, e.g., are they generated randomly?**

The weight are initialized randomly. This will be mentioned in-text.

11. **L. 302: What specific metric is used for the maximum dissimilarity sampling? Is it, e.g., based on Euclidean distance.**

The dissimilarity is based on the euclidean distance. We will mention this in-text.

12. **L. 320: When you mention radius are you referring to sigma?**

Yes, we refer to sigma. We will rewrite radii to sigma in-text.

13. ***Reviewer #1: L. 335-336: Could you confirm if you are using average linkage for the clustering? If so, did you also consider the Ward linkage approach, which might provide more balanced clusters by accounting for cluster variance?***

We use average linkage for clustering. We did consider Ward linkage, however, given the feature engineering we use (with variations in weight of features) we have to rely on Manhattan distance, which ward linkage does not support as it minimizes within-cluster sum of squared Euclidean distances. Edit:

L335: "The intercluster distance of these is quantified by computing the Manhattan distances between the mean feature vectors (i.e., using average linkage)."

14. ***Reviewer #1: L. 351-357: I am afraid that this entire paragraph is difficult for me to follow.***

To clarify, we will restructure this paragraph as follows:

L351-357: "The conditions and temporal patterns of the surface dynamics in the identified clusters are analysed using the active count time series of 4D-OBCs per cluster. The active count at a given epoch is defined as the number of 4D-OBCs that have started before that epoch but have not yet concluded. By computing this separately per cluster, we can compare the temporal activity of different surface dynamic types. The active count serves as a proxy for process energy: a higher count implies that more sand transport events of that type are simultaneously ongoing, which can be related to the prevailing environmental conditions. It does not, however, quantify the absolute volume of sand transported, which would require integration of the volumeTS across all active 4D-OBCs."

15. ***Reviewer #1: L. 362: Why was only a single cluster selected for manual investigation, and what were the specific criteria used to choose this one?***

The selection of a single cluster for detailed manual inspection was a deliberate scope choice for the purpose of demonstrating the workflow in this study. In practice, we visually inspected the spatial characteristics and time series of several surface dynamics clusters to assess whether the groupings are physically coherent; these broader checks lead to the interpretation of all eight clusters discussed in the results. Cluster D4 was selected for the detailed example because berm deposition is one of the most morphodynamically significant and recognisable processes at the study site, allowing readers to assess whether the grouped 4D-OBCs are indeed similar events. We will clarify this reasoning in the manuscript.

L364: "Finally, [..] season. In practice, the spatial characteristics and time series of multiple clusters were inspected by the authors to assess physical coherence; these broader checks inform the interpretation of all eight clusters discussed in the results. A single cluster was selected for detailed illustration that represents berm deposition, because berm deposition is one of the most morphodynamically apparent and visually recognisable processes at the study site, providing a clear basis to evaluate whether the grouped 4D-OBCs represent similar type of events."

16. ***L. 364: Please clarify what is meant by "relative frequency by season".***

The relative frequency by season refers to the fraction of 4D-OBCs assigned to a given SOM node that were initiated during a particular season, normalised by the total number of 4D-OBCs initiated in that node across all seasons. We will add a definition to the figure caption of Figures 12 and 13.

Reviewer #2

1. **L104: What were the main considerations that led to using the algorithms that are presented here?'**

We add additional explanation for the choice of algorithms in the methods, see our reply to the major comments.

2. ***L. 113: Was there a process to test whether cluster were previously unidentified types of dynamics, or whether some might be artifacts?***

We did not test this specifically as part of our analysis. We will additionally mention that these can be unidentified groups of surface dynamics, but also data artefacts that could be filtered out in an extra step.

L113: "We demonstrate the effectiveness of the method in classifying and organizing the PLS dataset into known and previously unidentified types of short-term surface dynamics, or data artefacts."

3. **L130: It might be good to give a summarizing statement on how large the systematic errors were before this correction and how large residual systematic errrors were (if any persisted).**

The characterisation of scanner errors and their correction is described in detail in Vos et al. (2022) and Voordendag et al. (2023), to which we refer the reader. Repeating this quantification here would in our view go beyond the scope of the present manuscript, which focuses on the extraction and classification of surface dynamics based on the data quality at hand.

4. ***L. 134: Is there a justification/reference for this specific threshold and the value of -60m? How noisy were the remaining point clouds that passed this filter? Was some noise filtering applied?***

We chose this threshold as the scanner was located at 55.757 m above NAP (approx. MSL in The Netherlands, Kuschnerus et al., 2024). Thus, even under low-tide there should not be real measurements below –60 m (relative to the scanner) by a margin (-4.25 m MSL). Therefore, if there is a large part of the point cloud that falls below this threshold, it cannot be deemed accurate. A selection of remaining point clouds have been inspected around periods where this noise occurred. It was observed that this noise increased over a couple of scans and then improved at once. With the filtering applied, the noisy clouds were removed around dates that were inspected. No extra noise filtering was done at this stage, but this was largely accounted for by temporal smoothing.

5. ***just semantics: the bars are the features; their build-up and erosion/destruction and migration would be the dynamics. It might be good to specify these dynamics individually.***

Agreed, we will change the wording throughout the manuscript accordingly, e.g., to intertidal bar build-up and erosion/migration.

6. ***L158: "conditions" refers here (and in line164) to enviromental condition during the dynamic change in the topography; and not to conditions before the dynamic change?***

Correct, in this study we only consider the conditions during the time between the initiation and finalization of 4D-OBCs. We have added recommendations to also consider the history of environmental variables in future work.

7. ***L177-178: Would be good to specify what the target size of changes was. Larger than 1x1 m (= 1 corepoint), or were changes expected to be larger than multiple***

corepoints (similar to the threshold mentioned in line 200)? Sampling distance refers to distance in 2D (x,y), or in 3D?

The surface dynamics under consideration were chosen to be larger than 10 m^2, which is a choice mainly based on practicality, which will be shortly elaborated on in-text (see the reply to in-line comment 6 by reviewer #1). However, these dynamics should show coherence over smaller areas, which is why the space-time array was constructed on a smaller grid. Sampling distance indeed is in 2D, which will be mentioned in-text.

"L176: We choose the corepoints as a regular grid with a 2D (planimetric) sampling distance of 1 m covering the beach and intertidal area."

8. ***L186-187: Is there a specific reason for choosing these 12h windows?***

The 12 h windows ensure that breakpoints are detected at high sensitivity with respect to the time scale of analysis (daily or larger). Breakpoints are then used to construct temporal seeds, and in case of over-detection, they are integrated within the timespan of a temporal seed. Shorter windows are not deemed useful, because the subdaily scale is already mostly averaged through temporal smoothing. This choice follows Anders et al. (2020). We add a short explanation:

L187: "The 24 h sliding window width ensures that breakpoints are detected at high sensitivity, however detecting changes at shorter windows is not deemed useful, because the subdaily scale is already mostly averaged through temporal smoothing."

9. ***L207: HPC node setups can vary widely. Could you provide a ball-park number on computational demand? E.g., how many GPUs/CPUs of which type and with how much RAM were used.***

We will add a concise statement of the computational resources used for the analysis, and rough estimation of minimum requirements necessary to the manuscript:

L222: "Computations are performed on an HPC node equipped with 48 cores (2× Intel Xeon E5-6248R, 3.0 GHz) and 768 GB RAM. The complete processing of all 9 subsets and their seed batches into 4D-OBCs with derived features completed in approximately 18 h of wall-clock time. The memory-intensive nature of loading a full subset point cloud time series into RAM is the primary driver for HPC use. A single subset, in our case, requires at maximum 23 GB of RAM. Parallel subset computation could thus in principle be performed on a workstation with sufficient memory, while the seed-batch step completion speed is CPU-bound and serial. Individual seed batches (of approximately 5,000 seeds) take up to *several hours. Using these smaller batches is found to be more efficient due to higher scheduler priority on shared HPC systems."*

And:

L536: " [..], we achieved extraction of 4D-OBCs and derivation of features from a set of 21,194 point clouds in only 18 h on an HPC (48 cores, 2× Intel Xeon E5-6248R, 3.0 GHz, and 768 GB RAM), with subsequent clustering on a relatively minimal workstation within an hour (32GB RAM, 24 threads)."

10. ***L220: How long did these jobs take? In relation with the used setup, this would give people a rough estimation of the computational power that is required for the workflow.***

See our previous reply. These batch jobs (of each 5000 seeds) took a maximum of several hours, however, as shorter jobs get higher priorities, the use of smaller batches worked more efficiently.

11. ***L243: Were there specific criteria to determine what constitues "unnatural outlying" or is this described in lines 241-244?***

Indeed, this is described in 241-244. We extend this to:

L243: "The value of 1 m is chosen as the inspected intertidal bar deposits in the dataset did not exceed this value. These intertidal bar deposits are deemed to be the largest natural dynamics occurring in the intertidal area."

12. ***L255: What was the reasoning for choosing these specific features. I think there is a brief statement in the discussion but it would be good to state this already here.***

We add a short statement here on the selection criteria:

L256: "These eight features were selected based on prior knowledge of the processes that we aim to distinguish and are chosen such that they describe the varying dimensions in which these differ."

13. ***L295: SOMs are sensitive to the initialization and training order. Were different hyperparamters/training configurations as well the stability of results (i.e., several runs with the same setup) tested?***

We have tested a random input order, but we did not perform any stability testing. However, the MDA input order used here gave a better representation of rare surface dynamics types.

14. ***L330: A sentence on the rationale behind choosing hierarchical clustering would be helpful. Were other clustering algorithms considered?***

This is addressed in our response to the major comment by Reviewer #2 on algorithm choice, and the proposed manuscript addition at L.294 and in Section 5.2. The rationale for hierarchical clustering and the consideration of alternatives (k-means, HDBSCAN) are discussed there.

15. ***L335: Were other distances tested as well? Which linkage criterion was used?***

This is addressed in our response to minor comment L.335–336 by Reviewer #1. Average linkage with Manhattan distance is used, Ward linkage was considered but is incompatible with Manhattan distance. The choice of Manhattan distance over Euclidean is motivated by the feature weighting applied to the volumeTS, as described in Section 3.2.3.

16. ***Fig.8 Caption: Do you mean edge colors do not imply similar clusters in general or similar edge colors between a and b do not imply similar clusters?***

We mean the latter indeed. We rephrase to:

Corresponding edge colours between a) and b) do not imply similar clusters, only similar colours within a) or b) imply similar clusters.

17. ***L412: Why 8 specifically? Becasue of the characteristics listed below? Were other (not analyzed) clusters less interpretable?***

The eight clusters are selected as a representative subset for illustration, not because the remaining clusters are less interpretable. In practice, the spatial characteristics and average time series of all 31 clusters were visually inspected. The eight clusters discussed were chosen to illustrate the main cross-shore zones and process types present in the dataset. We have addressed this comment further as a response to the comment of R*eviewer #1 at L. 362*.

18. ***Fig.9 Caption: Were there clusters with near uniform distributions for these features or were there clusters with identical feature distirbutions?***

The density plots in Figures 9 and 10 show a representative selection of clusters chosen for their clear feature separation, as noted in the manuscript. A full characterisation of the feature distributions of all 31 clusters is beyond the scope of this study. The dendrogram that we will add to the revised manuscript will provide an additional overview of the relative distances between all clusters, giving the reader an indication of which clusters are well-separated and which are more similar to each other.

19. ***L526: Is this statement also valid for the groups that were not analyzed in-depth?***

This is a fair question. The statement that the activity of the clusters is physically interpretable is demonstrated for the eight investigated clusters only. We will rephrase:

L526:" The variations in activity of the 8 selected groups are physically interpretable, and relate to changes in environmental conditions."

20. ***L536: Is this bracketing the entire workflow (incl. pre-processing) or referring to the 4D-OBC segmentation? Also, as mentioned above, what were the broad hardware specs?***

The 18 h refers to the full 4D-OBC extraction and feature derivation workflow, including the parallel subset processing and serial seed batch computation, but not the subsequent clustering which is completed within an hour on a standard workstation. The hardware specifications are now provided in the manuscript following our response to Reviewer #2's comment at L.207.

21. ***L553: This statement might be a bit exaggerated/too general. The workflow worked well for the presented classes in this dataset. Some of these classes might be ambigous, and in other settings, the dynamics might cluster differently.***

This is a valid point. We will rephrase to:
L553: "These findings highlight how unsupervised clustering methods applied to the 4D-OBC feature space can, for this environmental setting, extract morphodynamically meaningful patterns without the need for predefined class labels or thresholds. This is demonstrated for the 8 clusters investigated in this study. "

22. ***L632: Also, potentially, high-frequency (< 12h) dynamics with no net change would be missed. Although, I am not sure if any such dynamics exist that would be relevant for this study setting.***

Agreed, as per reply to earlier comments, we now already mention this in the methods section.

23. ***L685: That is one cluster, others might be different. This could be communicated here as well.***

We agree that the statement at L.686 could be read as a broader claim. We will change the sentence to:

L685: "4D-OBCs in a single cluster are manually interpreted and indeed yield 4D-OBCs of similar appearance at different moments in time throughout the three years of hourly data, while broader inspection of feature distributions of a further 8 clusters supports this coherence of surface dynamics type. Consequently, we demonstrate the ability to identify if a surface dynamic with a certain set of characteristics appears at different moments in time."

References

- Anders, Katharina, Lukas Winiwarter, Roderik Lindenbergh, Jack G. Williams, Sander E. Vos, and Bernhard Höfle. "4D Objects-by-Change: Spatiotemporal Segmentation of Geomorphic Surface Change from LiDAR Time Series." *ISPRS Journal of Photogrammetry and Remote Sensing* 159 (January 2020): 352–63. https://doi.org/10.1016/j.isprsjprs.2019.11.025.
- Anders, Katharina, Lukas Winiwarter, Hubert Mara, Roderik Lindenbergh, Sander E. Vos, and Bernhard Höfle. "Fully Automatic Spatiotemporal Segmentation of 3D LiDAR Time Series for the Extraction of Natural Surface Changes." *ISPRS Journal of Photogrammetry and Remote Sensing* 173 (March 2021): 297–308. https://doi.org/10.1016/j.isprsjprs.2021.01.015.
- Coelingh, J. P., A. J. M. Van Wijk, and A. A. M. Holtslag. "Analysis of Wind Speed Observations on the North Sea Coast." *Journal of Wind Engineering and Industrial Aerodynamics* 73, no. 2 (1998): 125–44. https://doi.org/10.1016/S0167-6105(97)00285-7.
- Kuschnerus, Mieke, Sierd De Vries, José A. Á. Antolínez, Sander Vos, and Roderik Lindenbergh. "Identifying Topographic Changes at the Beach Using Multiple Years of Permanent Laser Scanning." *Coastal Engineering* 193 (October 2024): 104594. https://doi.org/10.1016/j.coastaleng.2024.104594.
- Moulavi, Davoud, Pablo A. Jaskowiak, Ricardo J. G. B. Campello, Arthur Zimek, and Jörg Sander. "Density-Based Clustering Validation." *Proceedings of the 2014 SIAM International Conference on Data Mining*, April 28, 2014, 839–47. https://doi.org/10.1137/1.9781611973440.96.
- Sušelj, Kay, Abha Sood, and Detlev Heinemann. "North Sea Near-Surface Wind Climate and Its Relation to the Large-Scale Circulation Patterns." *Theoretical and Applied Climatology* 99, nos. 3–4 (2010): 403–19. https://doi.org/10.1007/s00704-009-0149-2.

- Truong, Charles, Laurent Oudre, and Nicolas Vayatis. "Selective Review of Offline Change Point Detection Methods." *Signal Processing* 167 (February 2020): 107299. https://doi.org/10.1016/j.sigpro.2019.107299.
- Voordendag, Annelies, Brigitta Goger, Christoph Klug, Rainer Prinz, Martin Rutzinger, Tobias Sauter, and Georg Kaser. "Uncertainty Assessment of a Permanent Long-Range Terrestrial Laser Scanning System for the Quantification of Snow Dynamics on Hintereisferner (Austria)." *Frontiers in Earth Science* 11 (March 2023): 1085416. https://doi.org/10.3389/feart.2023.1085416.
- Vos, Sander, Katharina Anders, Alain De Wulf, Sierd De Vries, and Roderik Lindenbergh. "SPATIO-TEMPORAL VARIATION OF AEOLIAN SHOREWARD SAND TRANSPORT MEASURED USING NEAR-CONTINUOUS LASER SCANNING." *Coastal Engineering Proceedings*, no. 37 (September 2023): 24. https://doi.org/10.9753/icce.v37.papers.24.
- Wijnberg, Kathelijne M. "Environmental Controls on Decadal Morphologic Behaviour of the Holland Coast." *Marine Geology* 189, nos. 3–4 (2002): 227–47. https://doi.org/10.1016/S0025-3227(02)00480-2.