Geochemistry, Geophysics, Geosystems

Supporting Information for

Mantle flow and anisotropy in subduction zones: modeling and clustering of olivine textures

Yijun Wang[1], Ágnes Király[1], Clinton P. Conrad[1], Valerie Maupin[1]

[1]Center for Planetary Habitability (PHAB), University of Oslo, Oslo, Norway

Contents of this file

Text S1. Features for clustering analysis and feature selection

Wang et al. (2024) showed that the effect of anisotropic viscosity, represented by the AV/IV ratio, is related to olivine textures, accumulated strains, and principal stresses. In this study, we aim to categorize tracers based on the time evolution of the parameters mentioned above. Our initial feature pool includes scores that describe olivine textures (pointiness, girdle-ness, and randomness), orientation of olivine a-axis (represented with the sine and cosine of the horizontal and vertical orientations), scores that describe olivine texture strength (the M-index and J-index), deformation history (accumulated strain and principal stresses), tracer path (x, y, and z coordinates), which together sum to 16 features in total. These features are not independent of each other. We use PCA (principal component analysis) to study the contribution of each feature for clustering (feature importance Fig. S1), selecting important features to keep in the following time-series clustering analysis. The most important features are always the x, y, and z coordinates and the P and G scores. This suggests that the tracer flow paths and CPO evolution contribute the most to clustering. AV/IV is usually one of the least important features, which shows its dependence on the other features. This aligns well with our objective to study CPO and CPO-induced AV evolution in different regions of subduction systems, represented by the

clusters. The results shown in the main text come from the set of features selected with a threshold of 0.33 (above the red line in Fig. S1) for time-series cluster analysis.

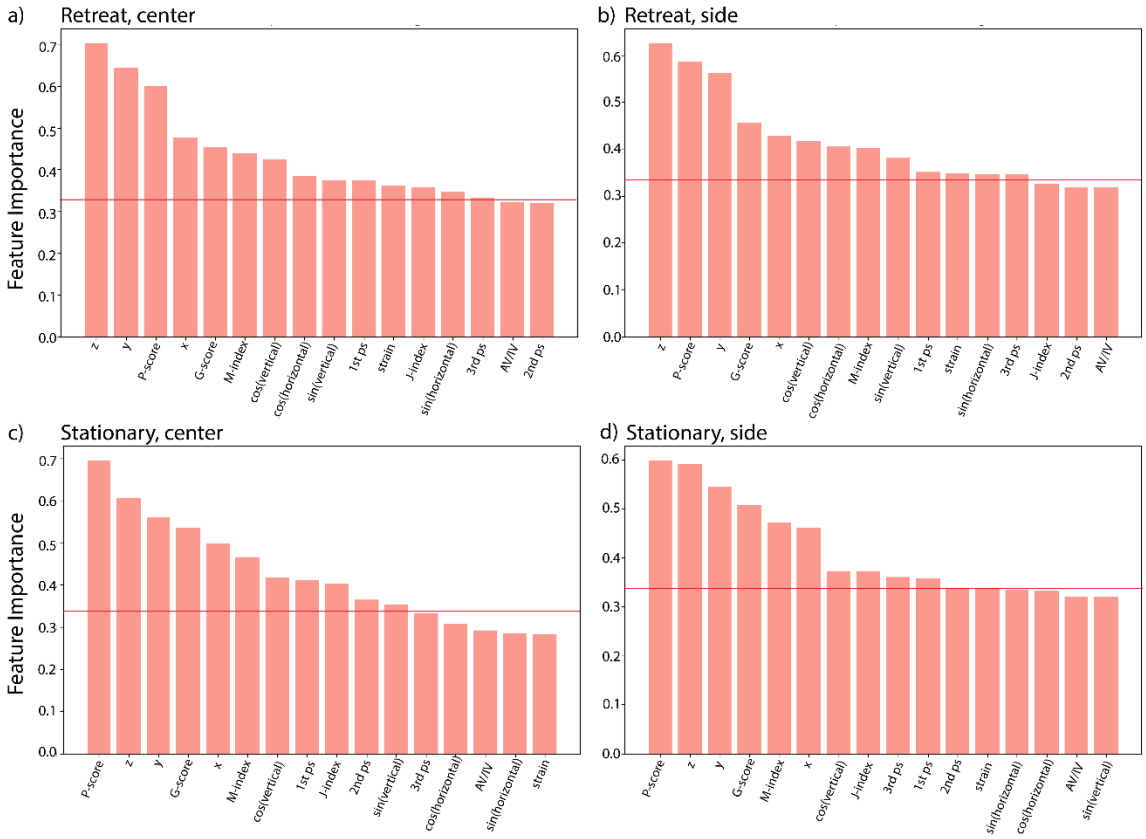

Figure S1. Feature importance sorted from the most important (greatest contribution to PCA) on the left to the least important on the right, for the retreating-trench model (a and b) and the stationary-trench model (c and d). Along the x-axis, the features include: P-score, G-score, M-index, J-index, sine, and cosine of the horizontal and vertical orientations (e.g. sin(horizontal)), three principal stresses (e.g. 1st ps), accumulated strain (strain), AV/IV, x-, y-, and z-coordinates.

Text S2. Selection of the number of clusters

Choosing the optimal number of clusters k for the analysis is important to obtain reasonable clustering results. As described in the main text, this choice should be based on what we expect about the natural separation of the dataset, as well as criteria that can be used to assess the clustering results with different k. For tracers placed in the subduction model, we expect the deformation and texture information to be separated into at least three large groups: the mantle wedge above the slab, the slab, and the sub-

slab region beneath the slab. As a result, we test k values with a minimum of 3 clusters. The Davies-Boudin index and Silhouette score for different k values are shown in Figure S2 for a clustering analysis of tracers in the subduction model with a retreating trench and a stationary trench in the center of the model and near the slab edge. The optimal k (denoted by the red line) is chosen by achieving a balance between maximizing the silhouette score and minimizing the Davies-Bouldin index values.

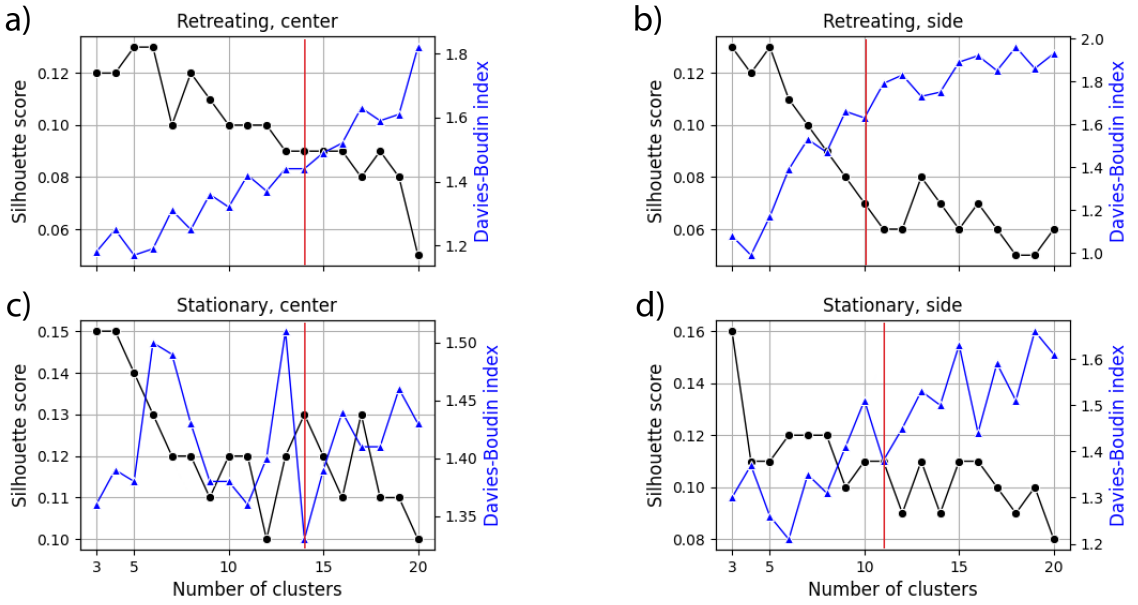

Figure S2. The Silhouette score (black lines, left axes) and Davis-Boudin index (blue lines, right axes) plotted against the number of clusters, k, for clustering analyses of tracers in the subduction model with a retreating trench (a and b) and a stationary trench (c and d), in the center of the model (a and c) and near the slab edge (b and d). The selected k value is marked with the red line.

Text S3. Cluster mean and its spread

We use the cluster mean as a representation of the general evolution of each cluster, while the spread varies for different features and clusters. Therefore, we discuss mainly the general trend (increasing, decreasing, constant) in the main text. Among all the features, the orientation of the olivine a-axis is least characterized by the mean. As a result, we describe the orientation as trench-parallel (TP), trench-normal (TN), and scattered in all directions (random, R), as represented by the majority of the orientations. The cluster mean for accumulated strain, P-score, G-score, and AV/IV, plotted on top of individual data lines from the cluster, are included in the Zenodo file Results folder, as well as the individual

evolution of olivine a-axis orientations and a histogram for the last time step for each cluster.