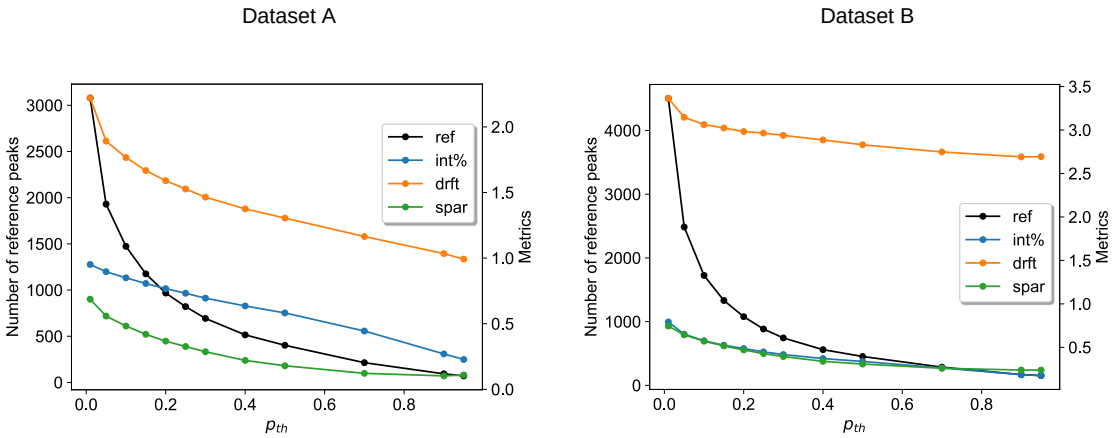
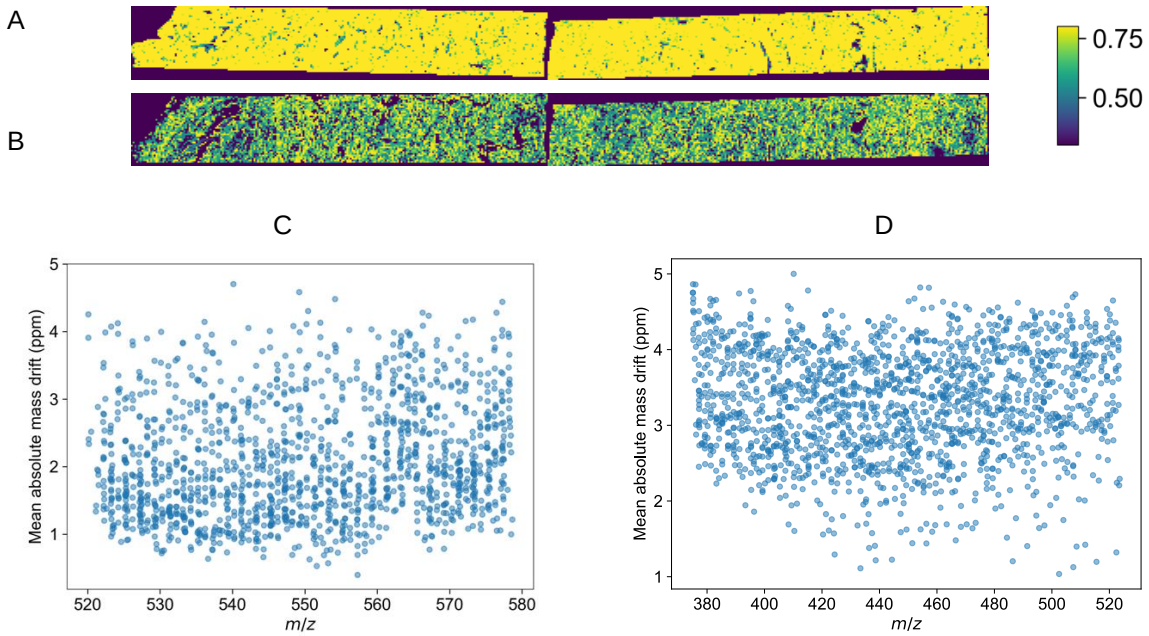
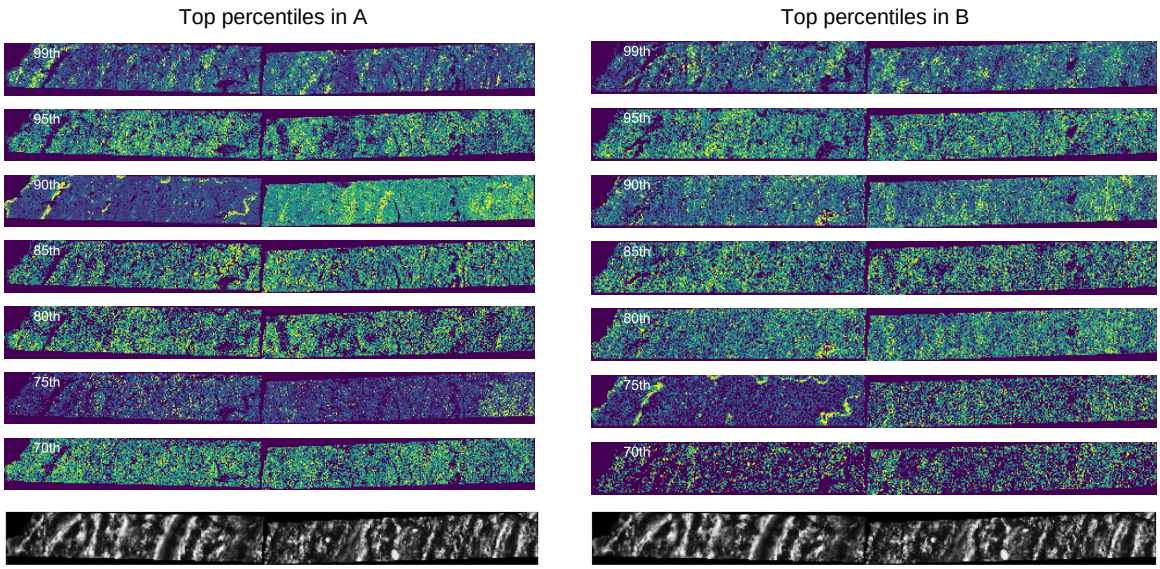


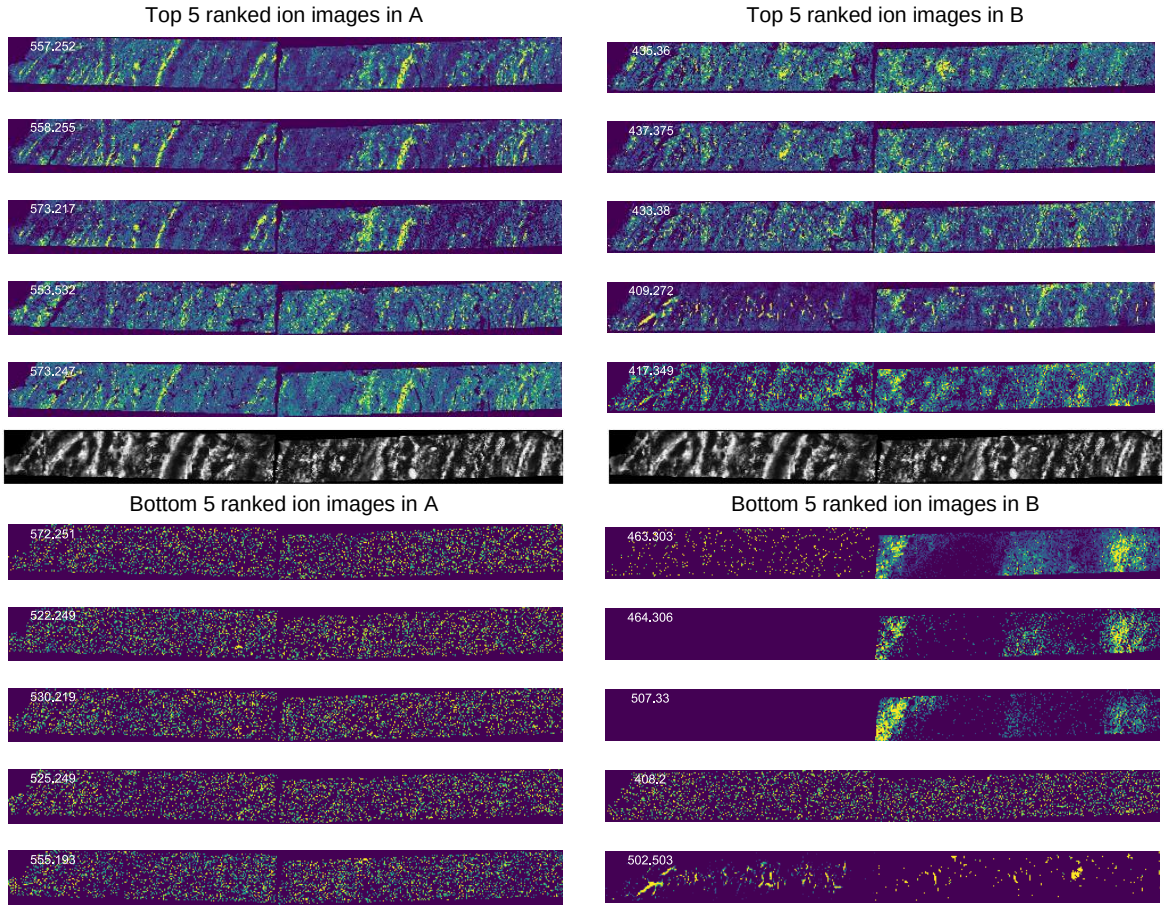
Supplementary Figure S1. Performance of bin-wise KDE with varying p_{th} values. *ref* on the left vertical axis denotes the number of reference peaks detected, *int%* on the right vertical axis denotes the percentage of peak intensity recovered from total peak intensity of the spectra, *drft* on the right vertical axis denotes the mean mass drifts in ppm of measured peaks around reference peaks, and *spar* on the right vertical axis denotes the sparseness of resulting data matrix after aligning measured peaks to detected reference peaks.



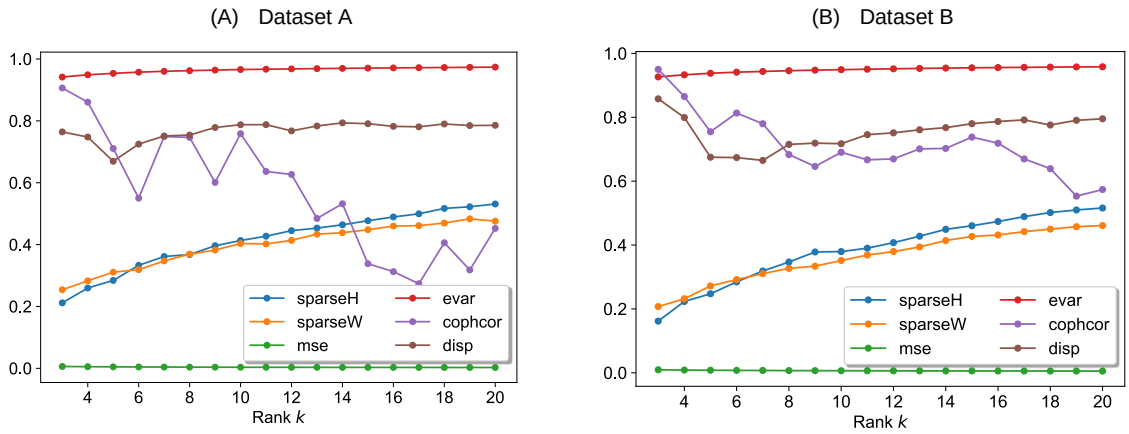
Supplementary Figure S2. Performance of bin-wise KDE with p_{th} values of 0.1 in datasets A and B. (A) and (B) show the percentage of recovered ion intensity per spot in dataset A and B, respectively; (C) and (D) show the mean absolute mass drifts of each reference peak in dataset A and B.



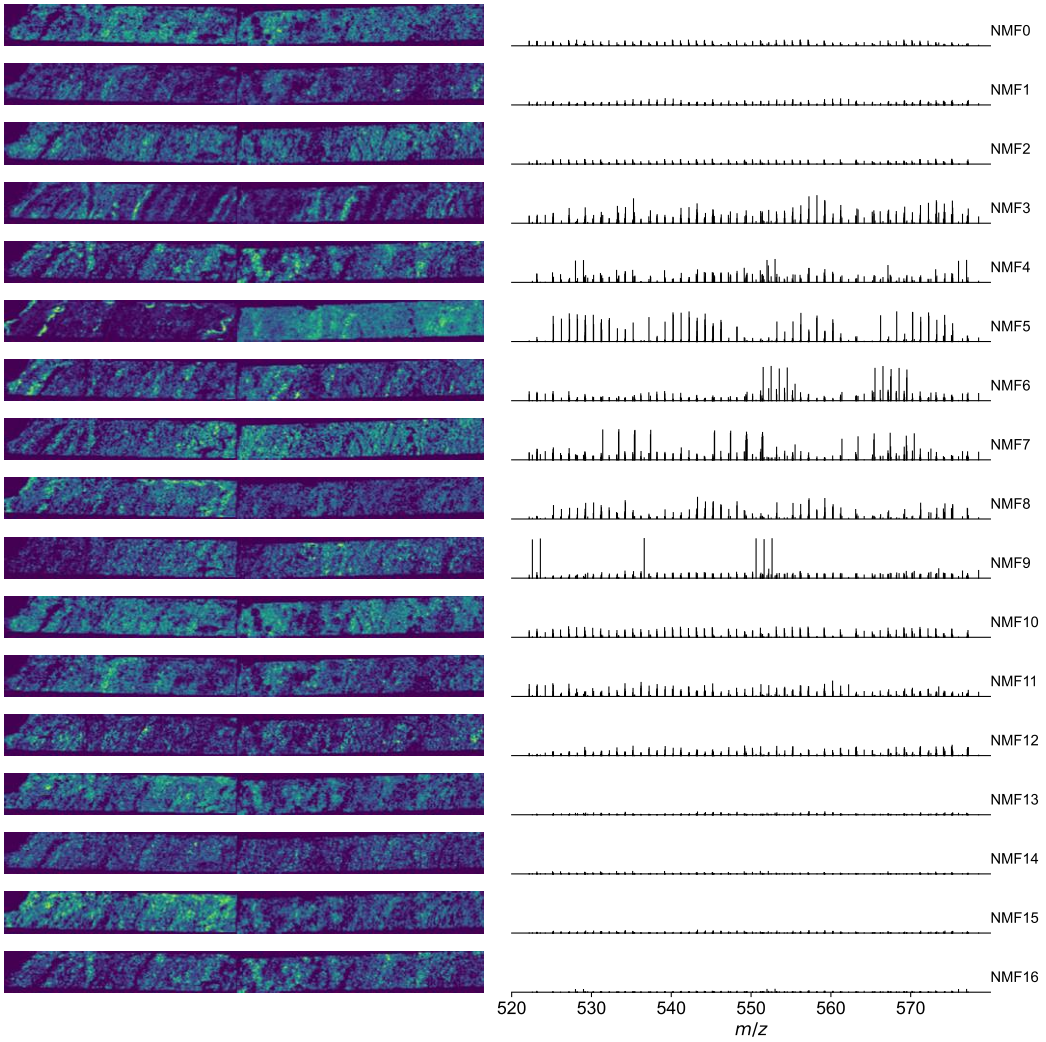
Supplementary Figure S3. The ion intensity maps of peaks at top percentiles after ranking by their similarities from the X-ray photograph (the pictures at the bottom) of the sediment.



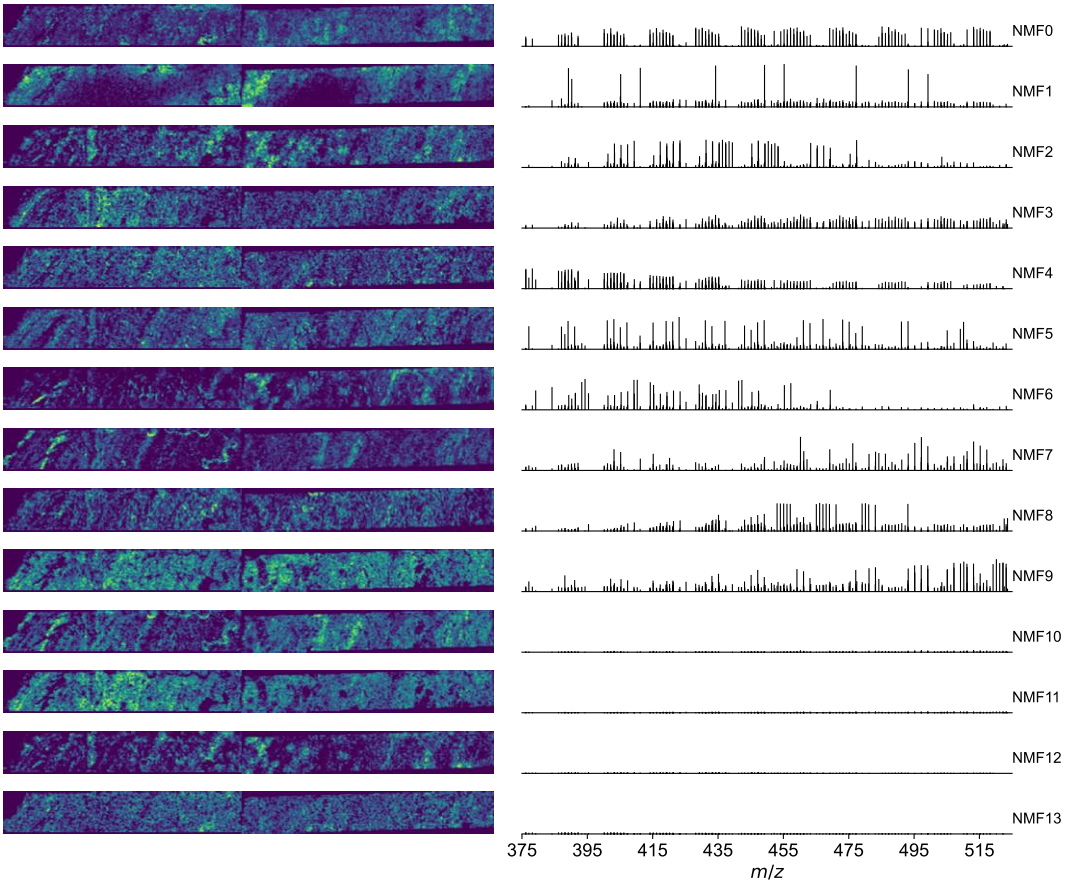
Supplementary Figure S4. The ion intensity maps of the five top ranked and the five bottom ranked peaks after ranking by their similarities from the X-ray photograph (the pictures in the middle) of the sediment.



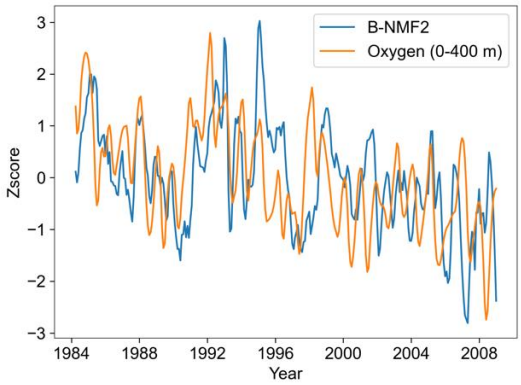
Supplementary Figure S5. NMF performance for dataset A (A) and B (B) at varying ranks between 3 and 20. Given $V \sim WH$, sparseW denotes the sparsity of resulting W matrix, which describes the unique spatial signatures; sparseH refers to the sparsity of resulting H matrix, which describes the most representative molecules associated with the spatial signatures. *mse* denotes the mean squared error between the original matrix and the reconstructed matrix. *evar* shows the proportion of the overall variance of original matrix that can be explained by the reconstructed matrix. *cophcor* and *disp* denote respectively the cophenetic correlation coefficients (Brunet et al., 2004) and the dispersion coefficient (Kim and Park, 2007) of the consensus matrix resulting from 30 repeated NMF runs.

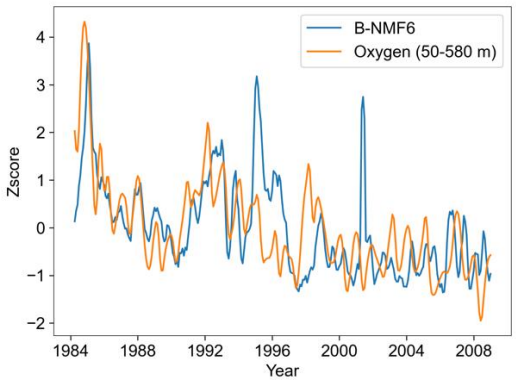


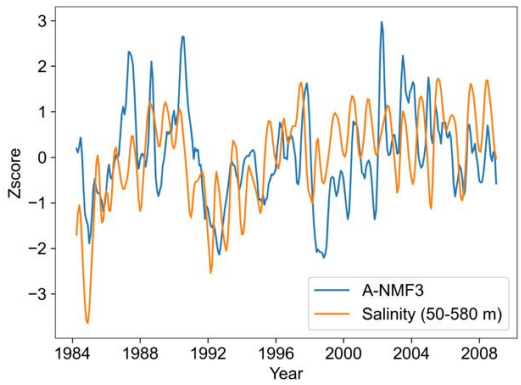
Supplementary Figure S6. All spatial signatures discovered in datasets A. The left pictures show the ion images of the spatial signatures, i.e., the columns of W, the weighted average of all ion images in the cluster. The ion images are in color scale “viridis”, and lighter colors denote higher abundances of the signatures. The right pictures show the pseudo mass spectra constructed using the contributions of molecules, i.e., in which peak intensities denote the entries in the rows of H. Peaks heights in different pseudo mass spectra are comparable: Higher peak heights indicate that they are accumulated over more NMF runs, i.e., having more distinctive spatial patterns.

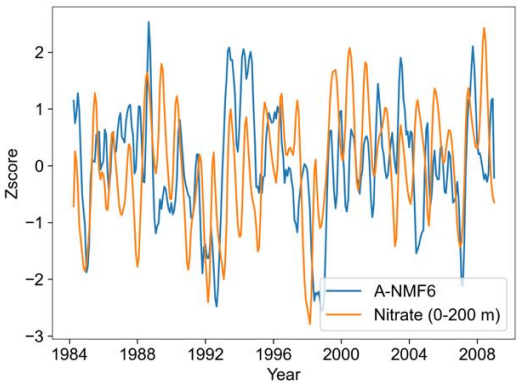


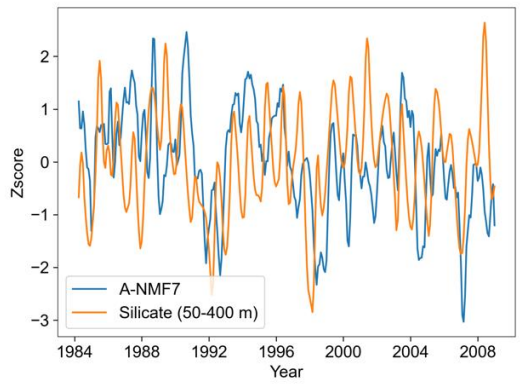
Supplementary Figure S7. All spatial signatures discovered in datasets B. The left pictures show the ion images of the spatial signatures, i.e., the columns of W, the weighted average of all ion images in the cluster. The ion images are in color scale “viridis”, and lighter colors denote higher abundances of the signatures. The right pictures show the pseudo mass spectrum constructed using the contributions of molecules, i.e., in which peak intensities denote the entries in the rows of H. Peaks heights in different pseudo mass spectra are comparable: Higher peak heights indicate that they are accumulated over more NMF runs, i.e., having more distinctive spatial patterns.

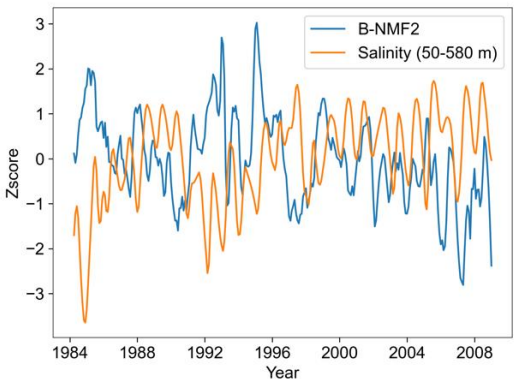

B-NMF2
Oxygen (0-400 m)
Zscore
Year

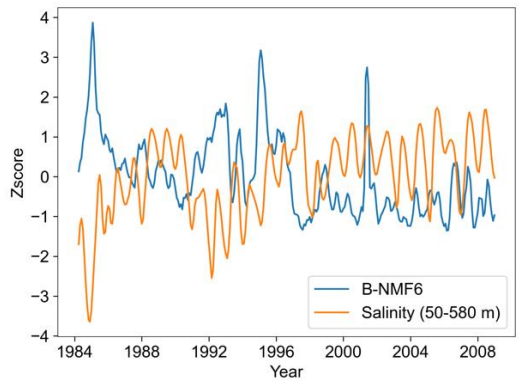

B-NMF6
Oxygen (50-580 m)
Zscore
Year

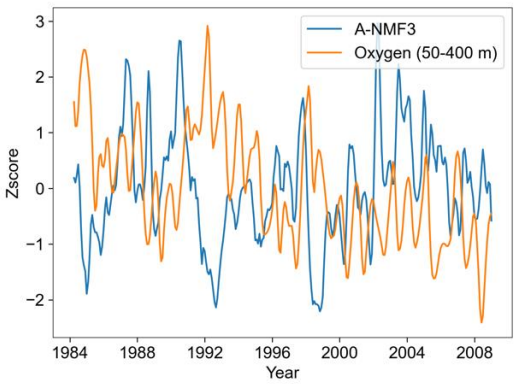

A-NMF3
Salinity (50-580 m)
Zscore
Year

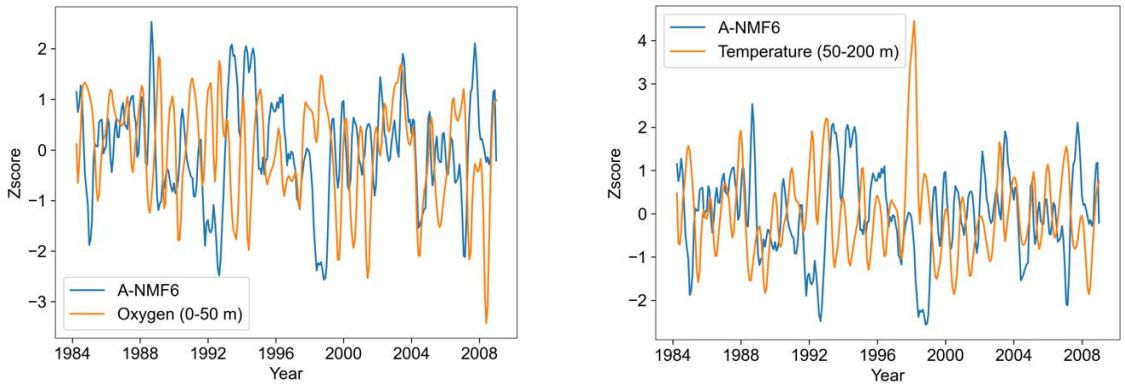

A-NMF6
Nitrate (0-200 m)
Zscore
Year


A-NMF7
Silicate (50-400 m)
Zscore
Year

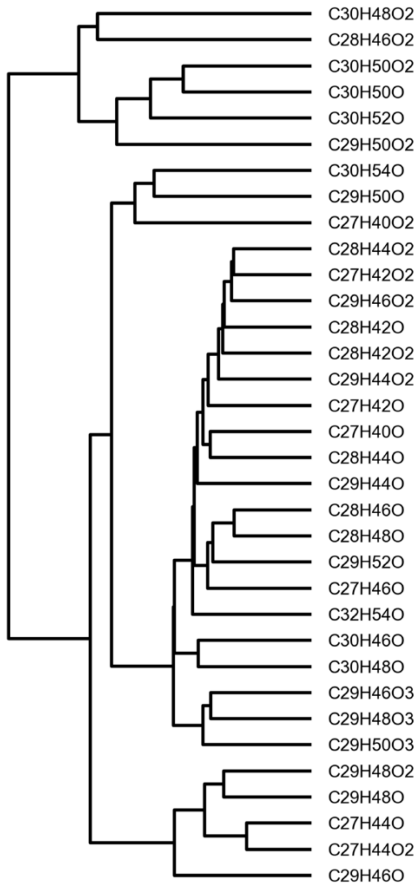

B-NMF2
Salinity (50-580 m)
Zscore
Year


B-NMF6
Salinity (50-580 m)
Zscore
Year


A-NMF3
Oxygen (50-400 m)
Zscore
Year



Supplementary Figure S8. Comparison between selected quarterly averaged signatures with environmental parameters.



Supplementary Figure S9. Hierarchical clustering of B-NMF2 illustrating its internal structure. The possible chemical formulas listed all have mass deviations smaller than 2 ppm.