

Description of Additional Supplementary Files

File Name: Supplementary Data 1

Description: Information of 953 metagenomic samples collected from Bio-GO-SHIP.

File Name: Supplementary Data 2

Description: Information of biogeochemical marker genes.

File Name: Supplementary Data 3

Description: Environmental factors linked to each sample.

File Name: Supplementary Data 4

Description: Dataset for machine learning used to quantitatively predict each microbial index.

File Name: Supplementary Data 5

Description: Machine learning performance in algorithm selection and hyperparameter tuning for regression.

File Name: Supplementary Data 6

Description: Hyperparameter tuning for Hierarchical clustering.

File Name: Supplementary Data 7

Description: Results of Nonparametric test and Cohen's d between each pair of ecological status.

File Name: Supplementary Data 8

Description: Dataset for machine learning classifiers used to predict ecological status

File Name: Supplementary Data 9

Description: Machine learning performance in algorithm selection and hyperparameter tuning for ecological status prediction.