Globally validated non-unique inversion framework to estimate optically active water quality indicators using *in situ* and space-borne hyperspectral datasets

Appendix A: Metadata of the GLORIA benchmark data set.

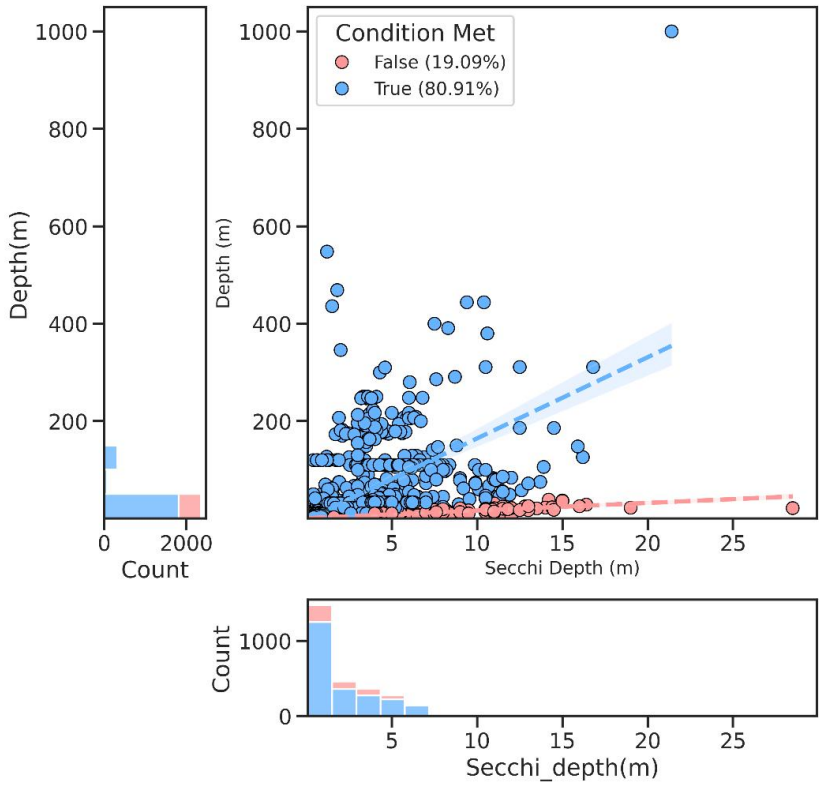

Fig. A1 The plot reflects the percentage of data within the GLORIA data set, which meets the condition for the bottom reflectance to be negligible. Here, the condition refers to the water column depth being three times or more than the Secchi disk depth. The proportion of data which doesn't meet the criterion has been removed from the analysis.

设置格式[军歌]: 非突出显示

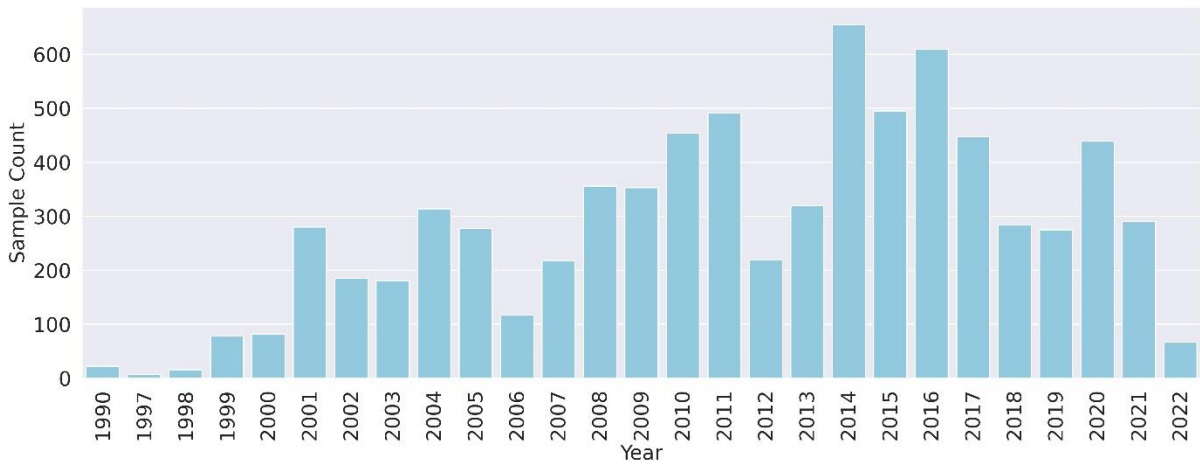

Fig. A2 Temporal distribution of the in situ samples within the GLORIA data set.

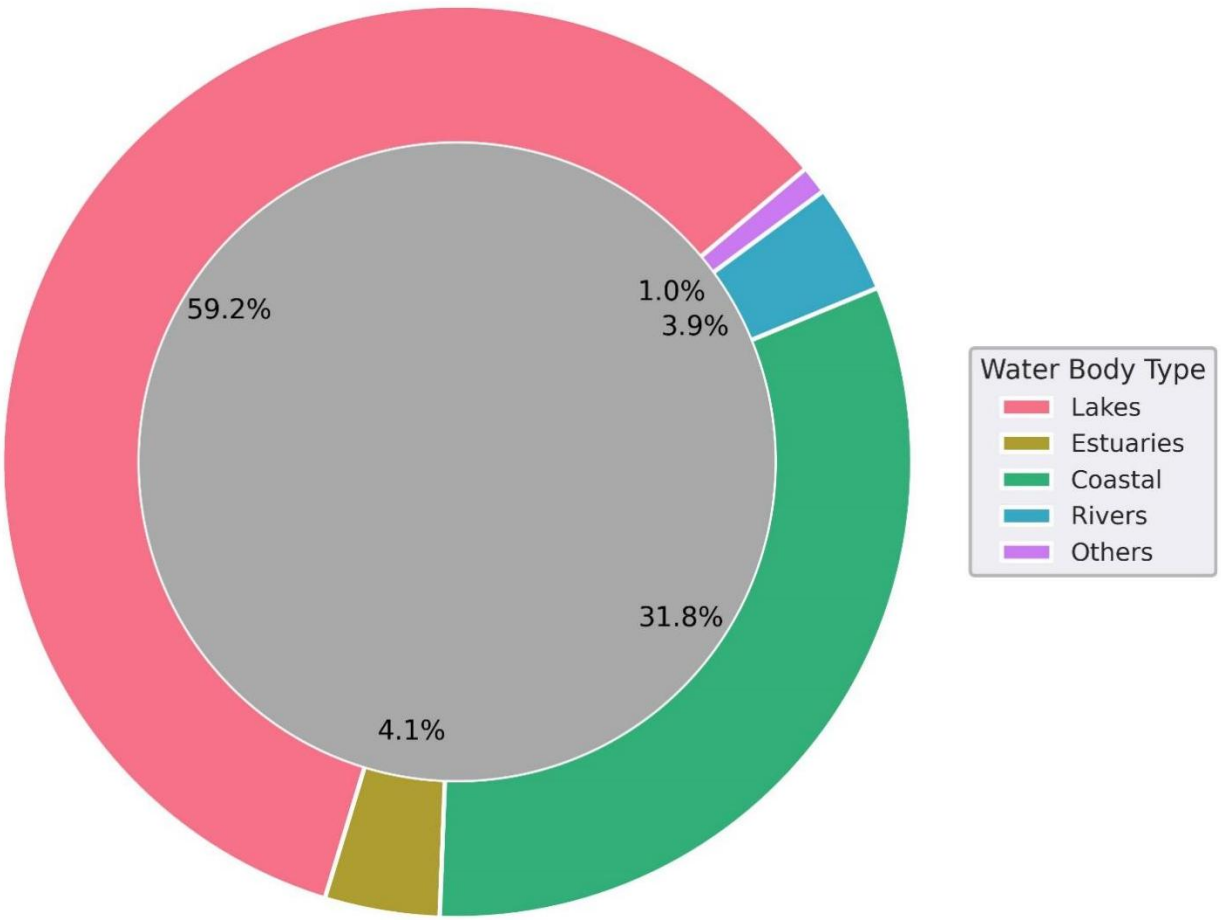

Fig. A3 Distribution of the in situ samples in terms of percentage corresponding to the various water body types in the GLORIA data set.

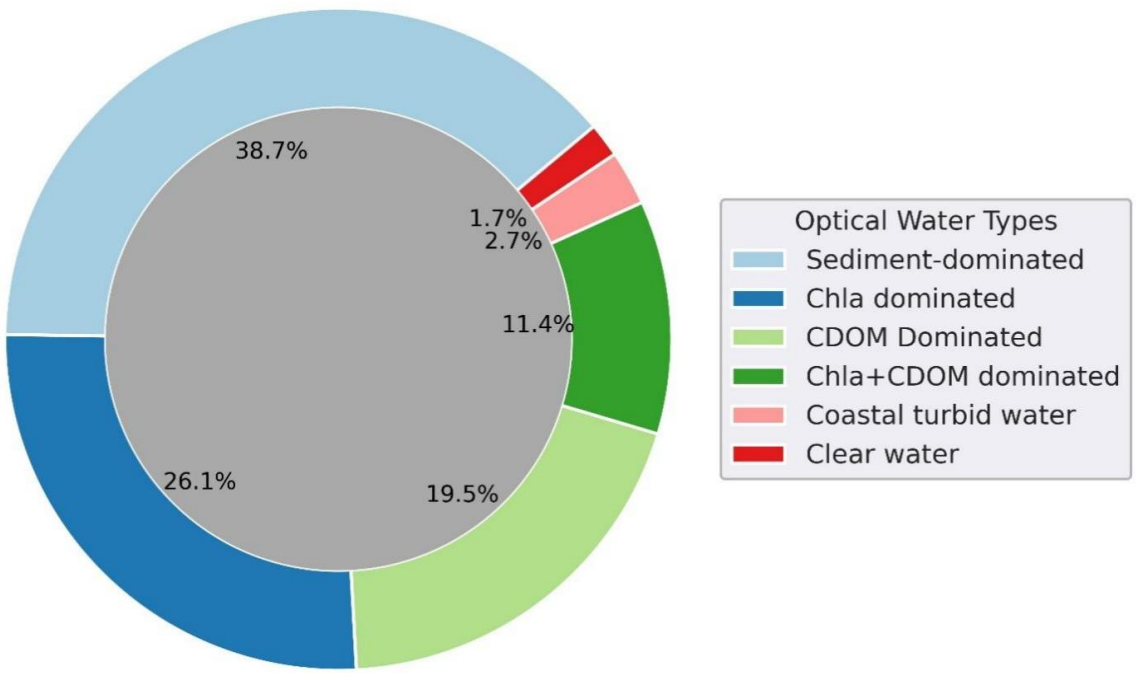

Fig. A4 Distribution of the in situ samples in terms of percentage corresponding to the six optical water types in the benchmark data set.

设置格式[军歌]：非突出显示

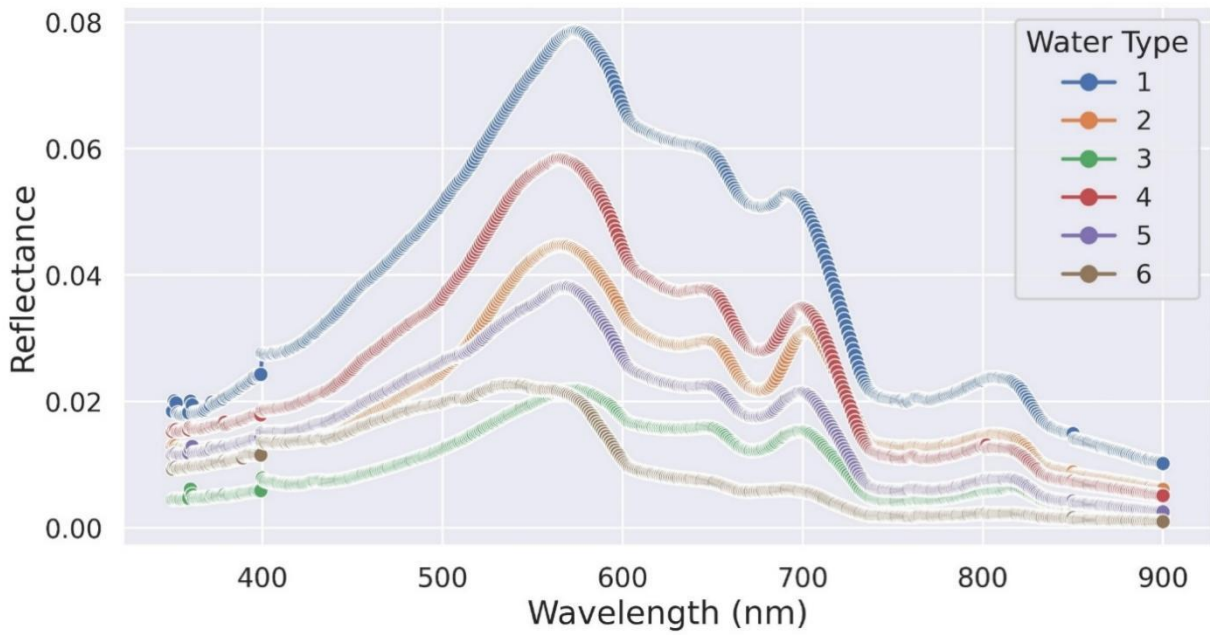

Fig. A5 Mean R_{rs} spectra corresponding to the six optical water types (OWT) labeled within the GLORIA data set.

Appendix B ACOLITE Settings

Table B1 Settings implemented within the ACOLITE module for the atmospheric correction of the space borne hyperspectral data cube PRISMA.

Parameters	Values
l2w_parameters	None
l2w_mask	TRUE
l2w_mask_wave	1600
l2w_mask_threshold	0.0215
l2w_mask_water_parameters	TRUE
l2w_mask_negative_rhow	TRUE
l2w_mask_negative_wave_range	400,900
l2w_mask_cirrus	TRUE
l2w_mask_cirrus_threshold	0.005
l2w_mask_cirrus_wave	1373
l2w_mask_high_toa	TRUE
l2w_mask_high_toa_threshold	0.3
l2w_mask_high_toa_wave_range	400,2500
l2w_mask_smooth	TRUE
l2w_mask_smooth_sigma	3
rgb_rhos	TRUE

Appendix C: Simulated bottom reflectance

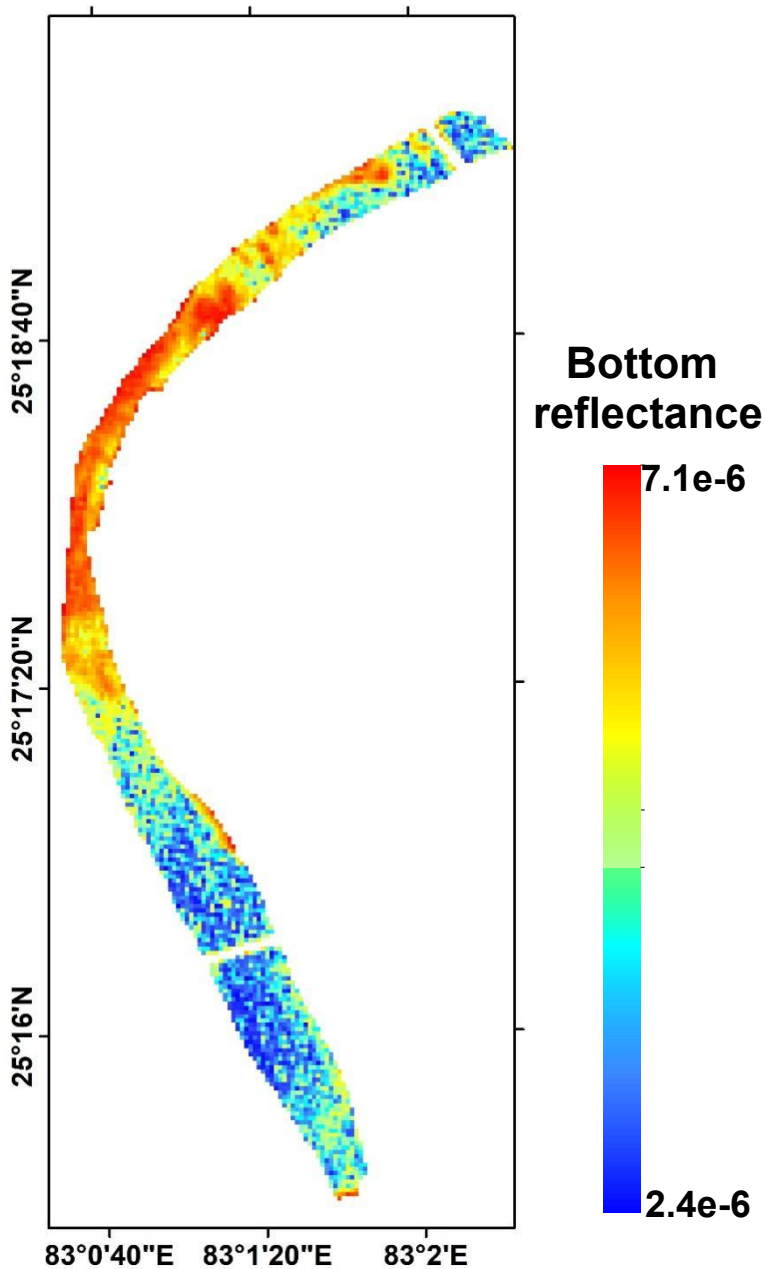

Fig. C1 Pixel-wise representation of simulated bottom reflectance using WASI.

设置格式[军歌]：非突出显示

Appendix D Hyperparameter tuning.

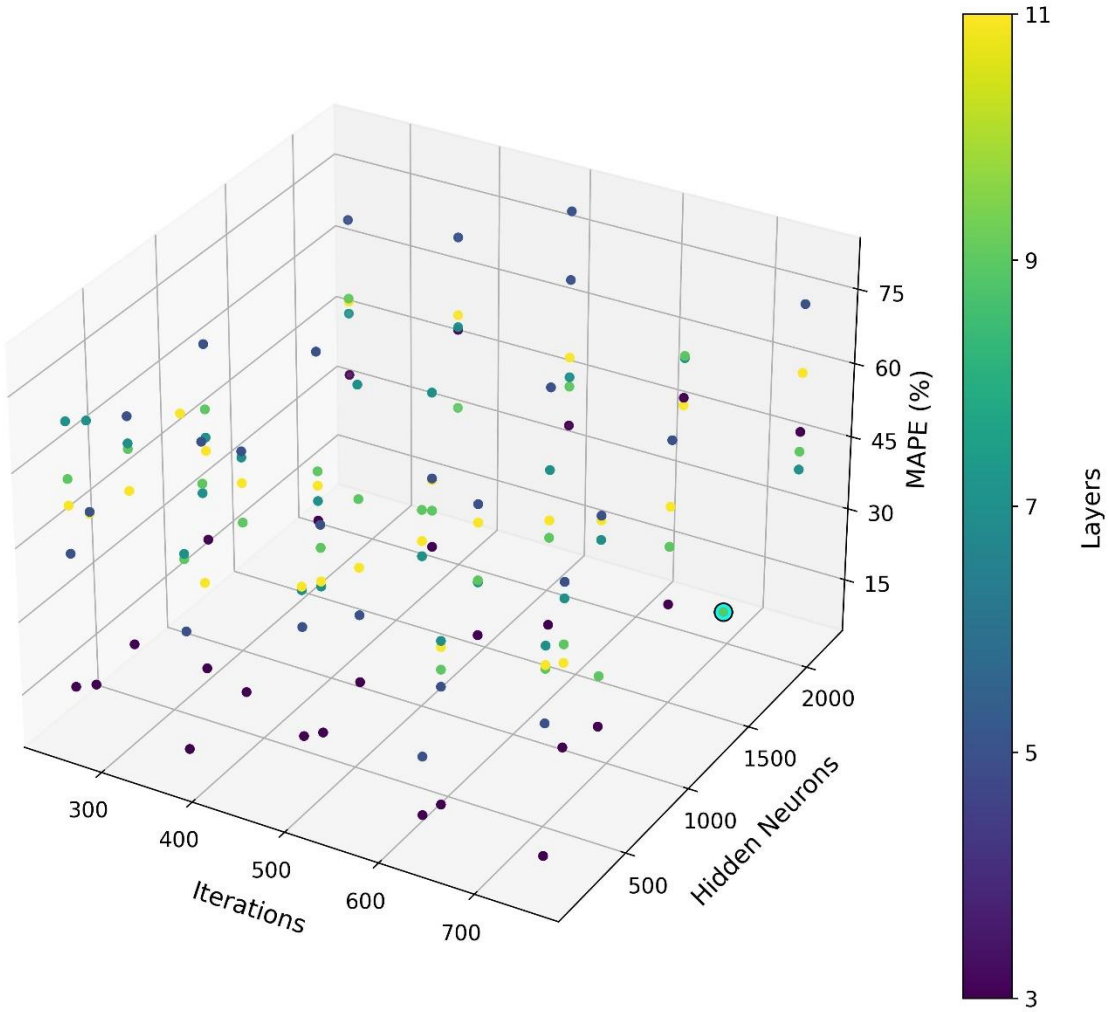

Fig. D1 Illustration of the hyperparameter optimization during the training phase of the proposed GMDN architecture for the concurrent retrieval of the WQ indicators. Variations in the iterations, neuron counts, and hidden layers have been explored, with the most optimal configuration highlighted in cyan based on the lowest MAPE value.