Supplementary Materials: Estimation of water quality parameters in oligotrophic coastal waters using UAV-obtained hyperspectral data

Morena Galešić Divić [1]*, Marija Kvesić Ivanković [2,3] Vladimir Divić [1], Mak Kišević [2], Marko Panić [4], Predrag Lugonja [4], Vladimir Crnojević [4] and Roko Andričević [1,2]

1. Sampling locations and field campaign conditions

Field campaigns consist of 32 measurement points considered for the analysis due to their overlap with the UAV overflight and depicted with a red polyline in Figure S1. Details of field conditions can be found in Table S1 where river flow corresponds to weekly averaged data, while wind speed refers to hourly averaged data.
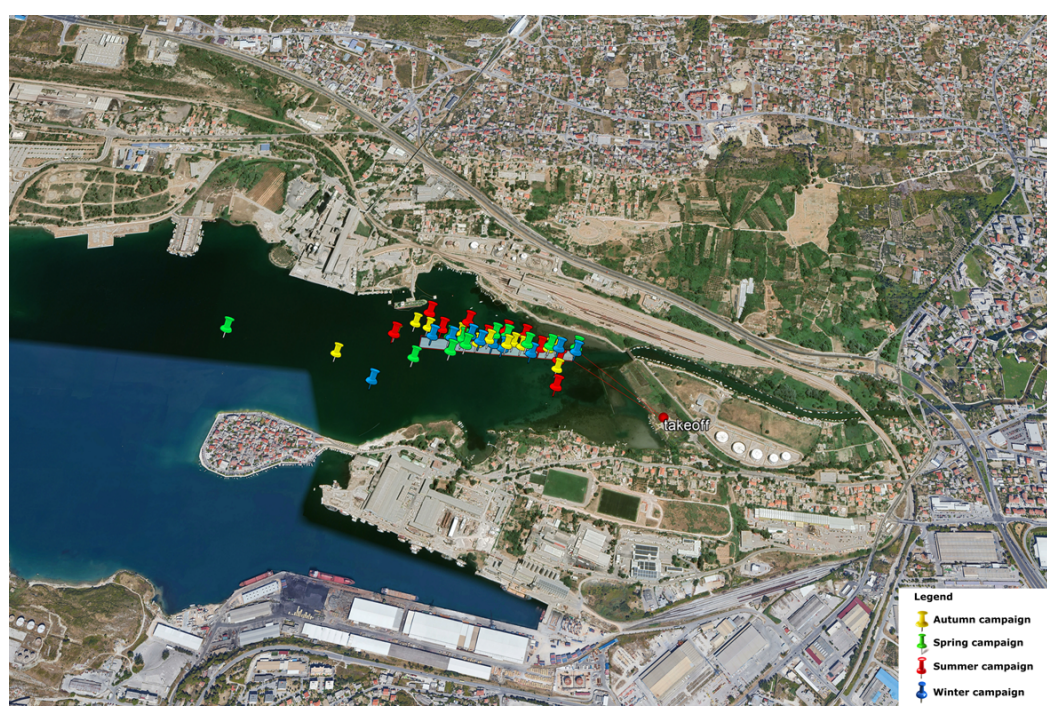

Figure S1. Sampling locations with the red polyline indicating flight path.

Table S1. Environmental conditions for field campaigns.

Field campaign date	River flow [m3/s]	Wind speed [m/s]	Wind direction	Air temperature [C°]	Humidity [%]	Sentinel-2 overpass date
25.06.2021.	4.1	3.1	southwest	30	45	26.06.2021.
02.11.2021.	4.3	5.6	northwest	19	60	29.10.2021.
03.03.2022.	7.6	1.7	east-southeast	11	30	03.03.2022.
27.04.2022.	6.1	3.1	southwest	20	56	27.04.2022.

2. Sensitivity analysis of hyperspectral data

Different preprocessing options for noise and scatter removal were tested, as mentioned in the Data processing sections. The same accuracy assessment was provided for reflectance data without any smoothing (RAW dataset), SG dataset with Savitzky-Golay filter (window 15 and 3rd-order polynomial) and the dataset containing both filtered and

normalized reflectance (SG SNV dataset). Results are given for tested existing algorithms for *Chl a* in Table S2 , TUR in Table S3 and CDOM in Table S4, respectively.

Table S2. Performance of hyperspectral reflectance datasets with different levels of preprocessing for *Chl a* evaluation using 32 ground truth points.

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	HS bands
RAW dataset						
$\widetilde{\theta}^1_{Chl\,a}$	$y = 0.03x + 0.40$	0.4375	0.2202	25.5151	10.5104	700,655
$\widetilde{\theta}^2_{Chl\,a}$	$y = 1.19x - 0.27$	0.1379	0.2749	39.8458	19.2964	432,562
$\widetilde{\theta}^3_{Chl\,a}$	$y = 0.65x + 0.10$	0.0468	0.2891	43.7987	21.1794	467,562
$\widetilde{\theta}^4_{Chl\,a}$	$y = 0.09x + 1.15$	0.5035	0.2086	24.8139	8.0996	680,660
$\widetilde{\theta}^5_{Chl\,a}$	$y = -0.004x + 0.72$	0.0002	0.2961	45.267	22.1892	498,518
$\widetilde{\theta}^6_{Chl\,a}$	$y = 2.06x + 1.13$	0.4658	0.2164	25.4393	10.0728	702,649
$\widetilde{\theta}^7_{Chl\,a}$	$y = 0.05x + 0.63$	0.3821	0.2328	30.4758	12.1707	677,664
$\widetilde{\theta}^8_{Chl\,a}$	$y = 1.25x - 0.08$	0.4915	0.2111	23.9894	9.4523	708,649
$\widetilde{\theta}^9_{Chl\,a}$	$y = 1.54x + 1.14$	0.4529	0.21	27.4971	10.9558	655,711,854

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	S2 bands
SG SNV dataset						
$\widetilde{\theta}^1_{Chl\,a}$	$y = 0.05x + 0.27$	0.3782	0.2335	30.8434	12.1893	700,655
$\widetilde{\theta}^2_{Chl\,a}$	$y = 0.88x + 0.0005$	0.0781	0.2843	42.5374	20.8805	432,562
$\widetilde{\theta}^3_{Chl\,a}$	$y = 0.72x + 0.03$	0.0687	0.2858	43.2213	20.8812	467,562
$\widetilde{\theta}^4_{Chl\,a}$	$y = 0.09x + 1.10$	0.3305	0.2423	32.4618	13.1433	680,660
$\widetilde{\theta}^5_{Chl\,a}$	$y = -0.0x + 0.86$	0.0673	0.286	43.9063	20.7531	498,518
$\widetilde{\theta}^6_{Chl\,a}$	$y = 2.77x + 1.29$	0.3575	0.2374	31.4312	12.6041	702,649
$\widetilde{\theta}^7_{Chl\,a}$	$y = 0.04x + 0.83$	0.251	0.2563	36.0757	15.3286	677,664
$\widetilde{\theta}^8_{Chl\,a}$	$y = 1.69x - 0.29$	0.3043	0.247	33.6407	13.9456	708,649
$\widetilde{\theta}^9_{Chl\,a}$	$y = 2.02x + 1.24$	0.4183	0.2258	30.7807	11.0438	655,711,854

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	S2 bands
SG dataset						
$\widetilde{\theta}^1_{Chl\,a}$	$y = 0.03x + 0.38$	0.4612	0.2174	25.1206	10.3064	700,655
$\widetilde{\theta}^2_{Chl\,a}$	$y = 1.20x - 0.28$	0.1437	0.274	39.7642	19.2927	432,562
$\widetilde{\theta}^3_{Chl\,a}$	$y = 0.76x - 0.02$	0.0722	0.2852	42.7849	20.6709	467,562
$\widetilde{\theta}^4_{Chl\,a}$	$y = 0.10x + 1.10$	0.5773	0.1925	21.2113	6.6163	680,660
$\widetilde{\theta}^5_{Chl\,a}$	$y = -0.06x + 0.85$	0.0481	0.2889	44.2392	21.1015	498,518
$\widetilde{\theta}^6_{Chl\,a}$	$y = 2.07x + 1.11$	0.4799	0.2135	24.9076	9.7163	702,649
$\widetilde{\theta}^7_{Chl\,a}$	$y = 0.06x + 0.83$	0.5941	0.1886	21.5512	5.99	677,664
$\widetilde{\theta}^8_{Chl\,a}$	$y = 1.31x - 0.11$	0.4803	0.2135	24.3119	9.7059	708,649
$\widetilde{\theta}^9_{Chl\,a}$	$y = 1.72x + 1.17$	0.4633	0.2169	26.1958	10.651	655,711,854

Table S3. Performance of hyperspectral reflectance datasets with different levels of preprocessing for TUR evaluation using 32 ground truth points.

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	HS bands
RAW dataset						
$\widetilde{\theta}^1_{TUR}$	$y = 0.05x + 2.05$	0.0172	0.8152	26.7126	11.2505	660
$\widetilde{\theta}^2_{TUR}$	$y = 0.002x + 2.18$	0.0024	0.8213	27.1614	11.5237	868
$\widetilde{\theta}^3_{TUR}$	$y = 0.11x + 1.94$	0.0188	0.8145	27.2474	11.3504	863,662
$\widetilde{\theta}^4_{TUR}$	$y = -2.39 \cdot 10^{-4}x + 2.37$	0.006	0.8198	27.3952	11.6292	784,772

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	S2 bands
SG SNV dataset						
$\widetilde{\theta}^1_{TUR}$	$y = 0.04x + 4.35$	0.0152	0.816	26.2873	11.3474	660
$\widetilde{\theta}^2_{TUR}$	$y = -4.23 \cdot 10^{-6}x + 2.23$	0.0008	0.822	27.3881	11.5819	868
$\widetilde{\theta}^3_{TUR}$	$y = 0.19x + 1.56$	0.0844	0.7868	24.6546	10.1796	863,662
$\widetilde{\theta}^4_{TUR}$	$y = -3.47 \cdot 10^{-4}x + 2.16$	0.0694	0.7932	26.2756	10.4028	784,772

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	S2 bands
SG dataset						
$\widetilde{\theta}^1_{TUR}$	$y = 0.52x + 2.05$	0.0175	0.8151	26.6984	11.2439	660
$\widetilde{\theta}^2_{TUR}$	$y = 0.003x + 2.17$	0.004	0.8206	27.1284	11.4889	868
$\widetilde{\theta}^3_{TUR}$	$y = 0.09x + 1.99$	0.0133	0.8168	27.3662	11.4465	863,662
$\widetilde{\theta}^4_{TUR}$	$y = -0.001x + 2.24$	0.093	0.7831	26.522	10.1447	784,772

Table S4. Performance of hyperspectral reflectance datasets with different levels of preprocessing for CDOM evaluation using 32 ground truth points.

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	HS bands
RAW dataset						
$\widetilde{\theta}^1_{CDOM}$	$y = -0.43x + 0.66$	0.0133	0.324	88.5857	61.5671	677,516
$\widetilde{\theta}^2_{CDOM}$	$y = 7.23x - 0.15$	0.0318	0.3209	89.0914	63.6139	525,562
$\widetilde{\theta}^3_{CDOM}$	$y = -0.99x + 0.65$	0.0135	0.324	88.5453	61.3957	461,677
$\widetilde{\theta}^4_{CDOM}$	$y = -0.17x + 0.64$	0.0254	0.322	88.5961	61.2461	576,662
$\widetilde{\theta}^5_{CDOM}$	$y = 6.50x - 0.48$	0.0384	0.3199	83.0367	57.0468	443,461

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	S2 bands
SG SNV dataset						
$\widetilde{\theta}^1_{CDOM}$	$y = -0.62x + 0.68$	0.0027	0.3257	89.4496	62.2277	677,516
$\widetilde{\theta}^2_{CDOM}$	$y = 15.17x - 0.95$	0.1243	0.3052	84.6498	58.8787	525,562
$\widetilde{\theta}^3_{CDOM}$	$y = -0.92x + 0.63$	0.0014	0.326	89.8084	62.5381	461,677
$\widetilde{\theta}^4_{CDOM}$	$y = -0.74x + 0.70$	0.0235	0.3223	86.9412	60.3223	576,662
$\widetilde{\theta}^5_{CDOM}$	$y = 10.65x - 1.15$	0.0941	0.3104	72.1338	48.3684	443,461

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	S2 bands
SG dataset						
$\widetilde{\theta}^1_{CDOM}$	$y = -0.49x + 0.67$	0.0164	0.3235	88.1263	61.1435	677,516
$\widetilde{\theta}^2_{CDOM}$	$y = 15.25x - 0.96$	0.1244	0.3052	84.1851	58.6932	525,562
$\widetilde{\theta}^3_{CDOM}$	$y = -1.15x + 0.66$	0.018	0.3232	88.0562	60.9208	461,677
$\widetilde{\theta}^4_{CDOM}$	$y = -0.17x + 0.64$	0.0271	0.3217	88.3406	60.9985	576,662
$\widetilde{\theta}^5_{CDOM}$	$y = 12.07x - 1.35$	0.1015	0.3092	72.6353	48.61	443,461

3. Comparison of UAV and satellite obtained data

The SNV-processed reflectance for both satellite and UAV data are presented in Figure S2. Point marks in satellite data indicate the central wavelength of corresponding bands.
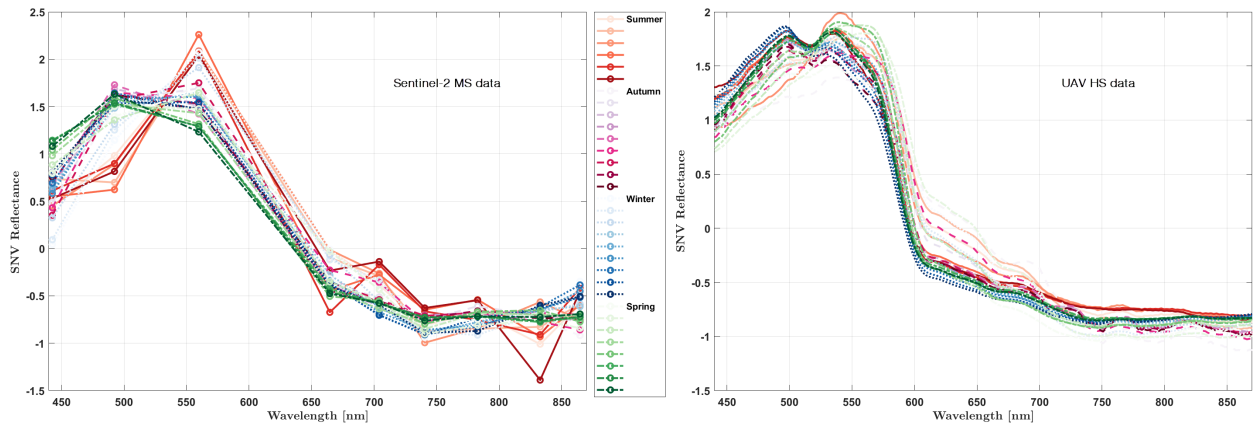

Figure S2. Comparison of normalized reflectance data obtained from Sentinel MSI (left panel) and UAV- mounted hyperspectral camera (right panel).

4. Statistical evaluation for tested algorithms

The performance of existing algorithms has been assessed by calculating the coefficient of determination (R^2), the root mean square error (RMSE), the absolute percentage difference (APD), and the relative percentage difference (RPD) for each of the algorithm (where applicable) and corresponding water quality parameter. Results are presented in the following tables.

Table S5. Performance summary of selected 9 algorithms for the *Chl a* derived from UAV HS and UAV MS datasets in comparison to 32 *in-situ* measurements.

Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	HS bands
UAV HS						
$\widetilde{\theta}^1_{Chl\,a}$	$y = 0.03x + 0.38$	0.4612	0.2174	25.1206	10.3064	700,655
$\widetilde{\theta}^2_{Chl\,a}$	$y = 1.20x - 0.28$	0.1437	0.274	39.7642	19.2927	432,562
$\widetilde{\theta}^3_{Chl\,a}$	$y = 0.76x - 0.02$	0.0722	0.2852	42.7849	20.6709	467,562
$\widetilde{\theta}^4_{Chl\,a}$	$y = 0.10x + 1.10$	0.5773	0.1925	21.2113	6.6163	680,660
$\widetilde{\theta}^5_{Chl\,a}$	$y = -0.06x + 0.85$	0.0481	0.2889	44.2392	21.1015	498,518
$\widetilde{\theta}^6_{Chl\,a}$	$y = 2.07x + 1.11$	0.4799	0.2135	24.9076	9.7163	702,649
$\widetilde{\theta}^7_{Chl\,a}$	$y = 0.06x + 0.83$	0.5941	0.1886	21.5512	5.99	677,664
$\widetilde{\theta}^8_{Chl\,a}$	$y = 1.31x - 0.11$	0.4803	0.2135	24.3119	9.7059	708,649
$\widetilde{\theta}^9_{Chl\,a}$	$y = 1.72x + 1.17$	0.4633	0.2169	26.1958	10.651	655;711;854
UAV MS						
Algorithm	Regression	R^2	RMSE	APD (%)	RPD (%)	S2 bands
$\overline{\theta}^1_{Chl\,a}$	$y = 0.03x + 0.38$	0.4504	0.2195	25.6372	10.5391	B5,B4
$\overline{\theta}^2_{Chl\,a}$	$y = 1.04x - 0.16$	0.0969	0.2814	41.5835	20.3101	B1,B3
$\overline{\theta}^3_{Chl\,a}$	$y = 0.73x - 0.03$	0.0515	0.2884	43.748	21.0595	B2,B3
$\overline{\theta}^4_{Chl\,a}$	/	/	/	/	/	B4,B4
$\overline{\theta}^5_{Chl\,a}$	/	/	/	/	/	B2,B2
$\overline{\theta}^6_{Chl\,a}$	$y = 2.60x + 1.09$	0.4597	0.2176	25.4758	10.1877	B5,B4
$\overline{\theta}^7_{Chl\,a}$	/	/	/	/	/	B4,B4
$\overline{\theta}^8_{Chl\,a}$	$y = 1.70x - 0.57$	0.4544	0.2187	25.3355	10.3945	B5,B4
$\overline{\theta}^9_{Chl\,a}$	$y = 2.40x + 1.12$	0.4153	0.2264	28.2316	11.998	B4,B5,B8a

Table S6. Summary of 5 tested algorithms for the TUR evaluation using 32 ground truth points.

	UAV HS					
Algorithm	*Regression*	R^2	*RMSE*	*APD (%)*	*RPD (%)*	*HS bands*
$\widetilde{\theta}^1_{TUR}$	$y = 0.52x + 2.05$	0.0175	0.8151	26.6984	11.2439	660
$\widetilde{\theta}^2_{TUR}$	$y = 0.003x + 2.17$	0.004	0.8206	27.1284	11.4889	868
$\widetilde{\theta}^3_{TUR}$	$y = 0.09x + 1.99$	0.0133	0.8168	27.3662	11.4465	863,662
$\widetilde{\theta}^4_{TUR}$	$y = -0.001x + 2.24$	0.093	0.7831	26.522	10.1447	784,772
Algorithm	*Regression*	R^2	*RMSE*	*APD (%)*	*RPD (%)*	*S2 bands*
$\overline{\theta}^1_{TUR}$	$y = 0.05x + 2.06$	0.0158	0.8157	26.7495	11.2725	B4
$\overline{\theta}^2_{TUR}$	$y = 0.003x + 2.17$	0.004	0.8206	27.1273	11.4903	B8a
$\overline{\theta}^3_{TUR}$	$y = 0.08x + 2.01$	0.0107	0.8179	27.3192	11.48	B8a,B4
$\overline{\theta}^4_{TUR}$	$y = 0.0002x + 2.11$	0.0378	0.8066	27.4884	11.073	B8,B7

Table S7. Summary of 5 tested algorithms for the CDOM evaluation using 32 ground truth points.

	UAV HS					
Algorithm	*Regression*	R^2	*RMSE*	*APD (%)*	*RPD (%)*	*HS bands*
$\widetilde{\theta}^1_{CDOM}$	$y = -0.49x + 0.67$	0.0164	0.3235	88.1263	61.1435	677,516
$\widetilde{\theta}^2_{CDOM}$	$y = 15.25x - 0.96$	0.1244	0.3052	84.1851	58.6932	525,562
$\widetilde{\theta}^3_{CDOM}$	$y = -1.15x + 0.66$	0.018	0.3232	88.0562	60.9208	461,677
$\widetilde{\theta}^4_{CDOM}$	$y = -0.17x + 0.64$	0.0271	0.3217	88.3406	60.9985	576,662
$\widetilde{\theta}^5_{CDOM}$	$y = 12.07x - 1.35$	0.1015	0.3092	72.61	48.6099	443,461
	UAV MS					
Algorithm	*Regression*	R^2	*RMSE*	*APD (%)*	*RPD (%)*	*S2 bands*
$\overline{\theta}^1_{CDOM}$	$y = -0.29x + 0.64$	0.0062	0.3252	89.2442	62.0692	B4,B2
$\overline{\theta}^2_{CDOM}$	$y = 6.74x - 0.14$	0.0585	0.3165	87.7103	62.2816	B2,B3
$\overline{\theta}^3_{CDOM}$	$y = -0.77x + 0.63$	0.0075	0.3249	89.1821	61.9696	B2,B4
$\overline{\theta}^4_{CDOM}$	$y = -0.14x + 0.62$	0.015	0.3237	88.9626	61.8989	B3,B4
$\overline{\theta}^5_{CDOM}$	$y = 2.87x + 0.02$	0.0298	0.3213	83.8034	58.4464	B1,B2

5. Visualized performance of different functions implemented in local empirical algorithms

We have tested four simple functions to obtain the water quality parameters based on available combinations of hyperspectral bands. The simple relation between bands was used: the $\widetilde{R}_{Mband}/\widetilde{R}_{Nband}$, where both bands belong to the relevant spectral domain.
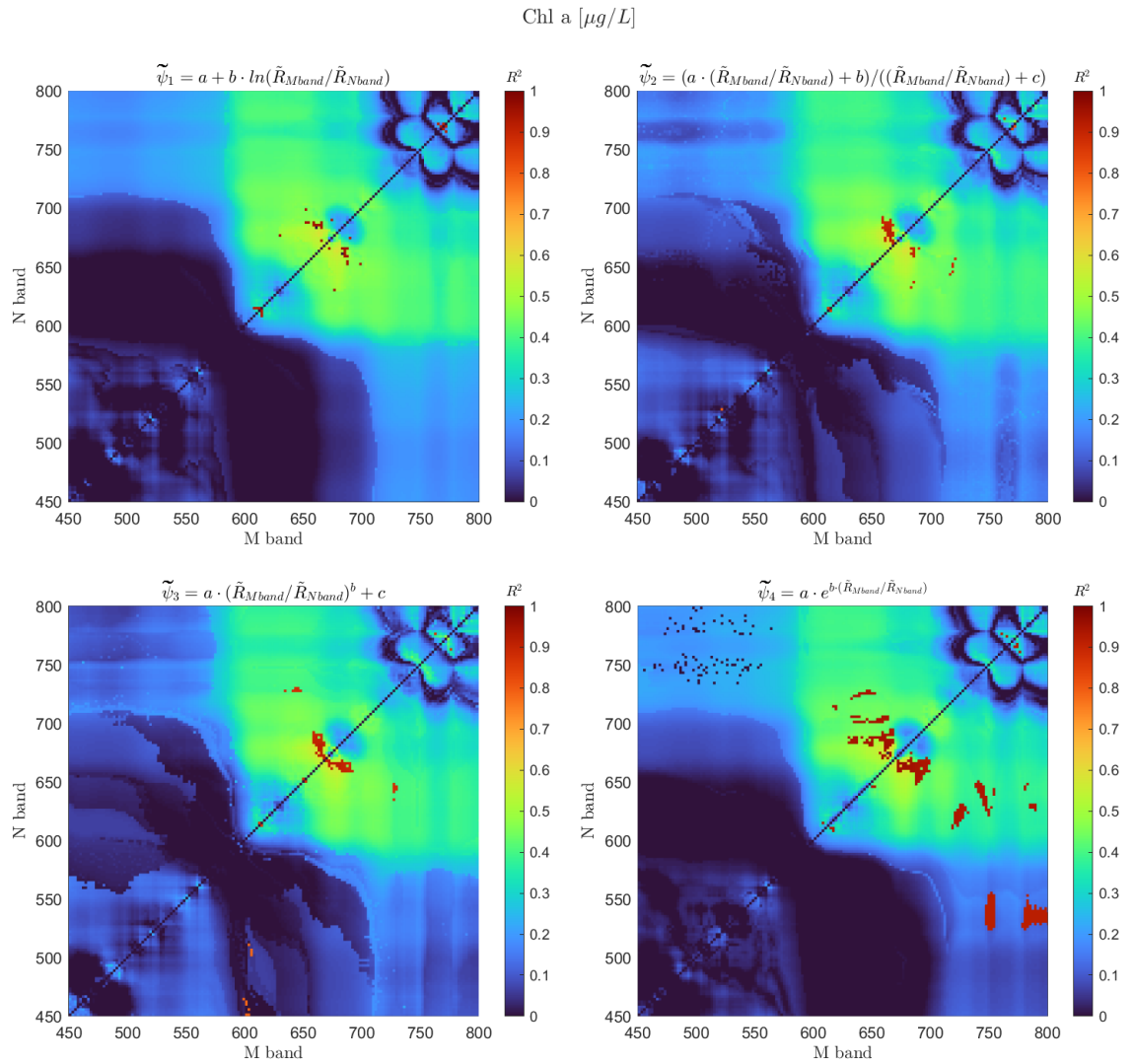

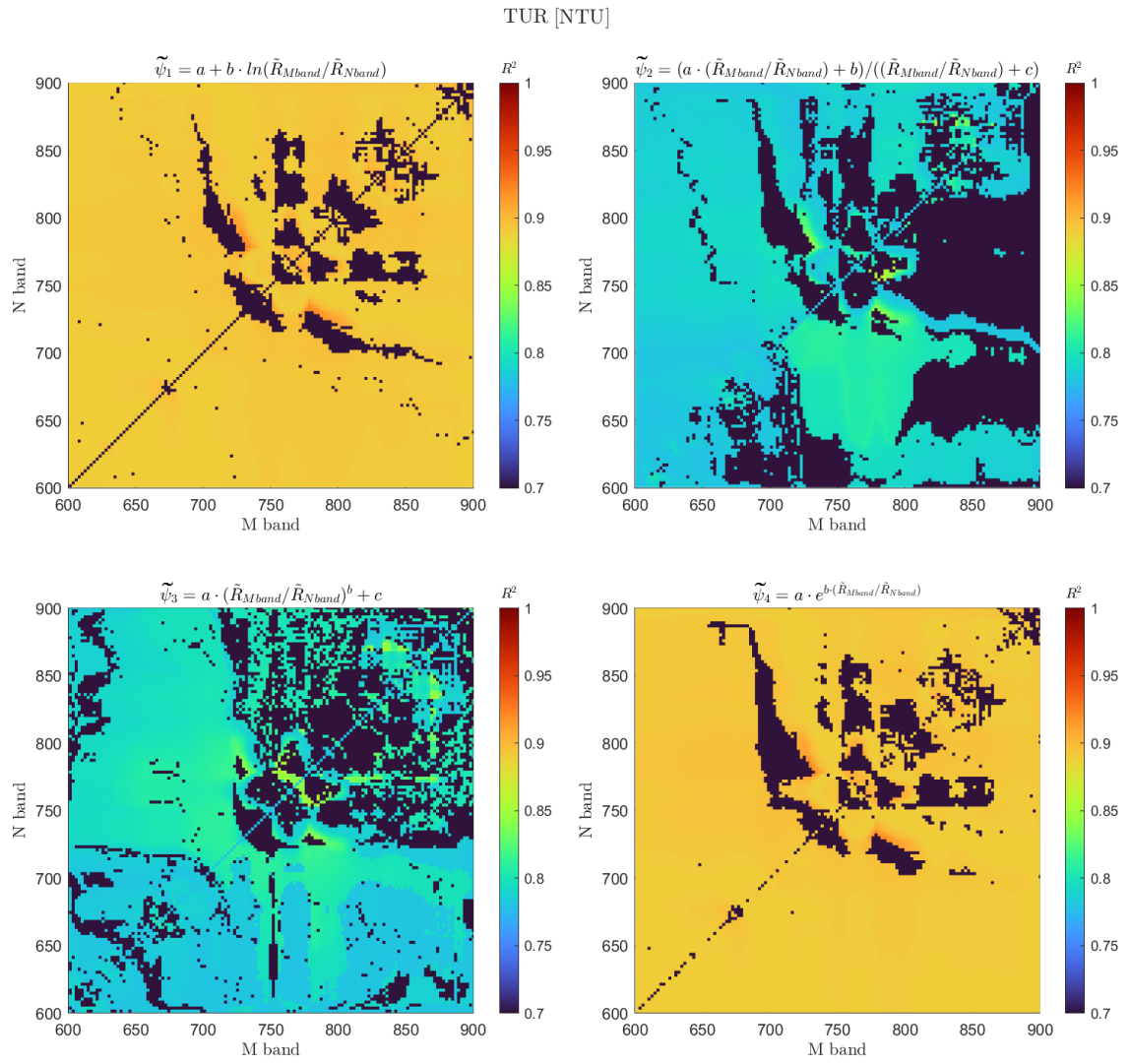**Figure S3.** R^2 Performance of different functions for *Chl a*.

TUR [NTU]

Figure S4. R^2 Performance of different functions for TUR.

CDOM [ppb]

Figure S5. R^2 Performance of different functions for CDOM.

6. Robust Nonlinear Fitting for Regression Analysis

Adequate regression-based models were obtained using robust nonlinear fitting integrated into the MATLAB computing environment. In this section, we present the impact of utilizing weights (*w*) obtained within the MATLAB Curve Fitting Toolbox with the Least Absolute Residuals (LAR) method. This method serves as a computational alternative to manual outlier removal, chosen to mitigate potential bias and optimize computational resources during the iterative search for the optimal band ratio and regression function.

The following figures illustrate the specific weights and the selected model with automatically detected outliers for three parameters: chlorophyll-*a* (*Chl a*) (Figure S6), turbidity (Figure S7), and colored dissolved organic matter (Figure S8).
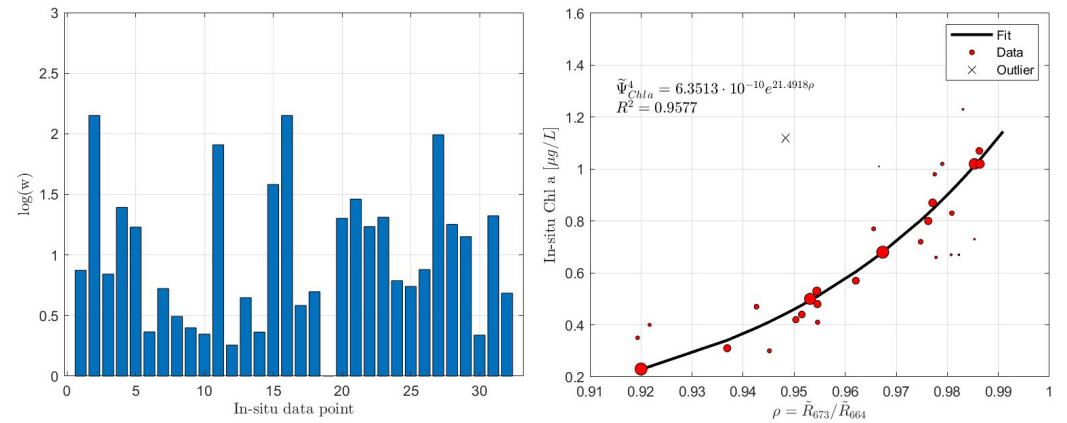

Figure S6. Left Panel: Specific weights (for 32 data points) derived from LAR are used to assess the R^2 of the proposed *Chl a* model. Each bar represents the weight of a data point used in the model. Right Panel: Model visualization with data points indicated by circles. The size of each circle corresponds to the specific weight shown in the left panel. One outlier is denoted by cross marker.
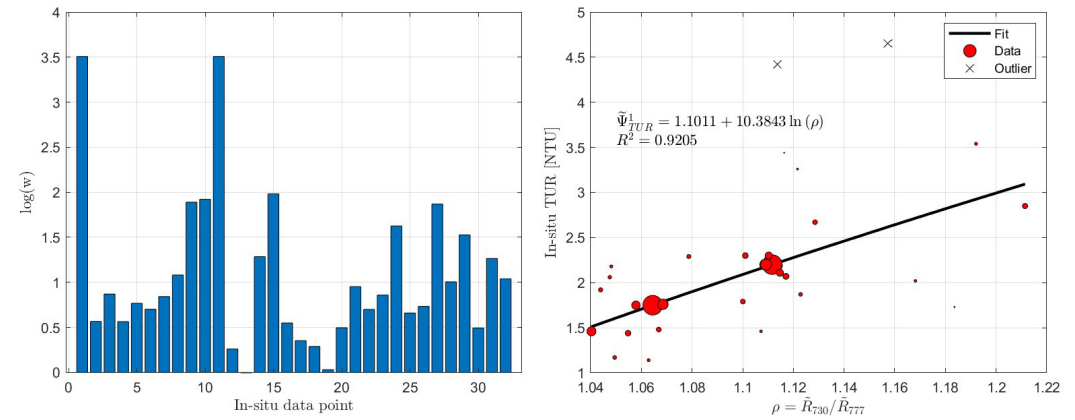

Figure S7. Left Panel: Specific weights (for 32 data points) derived from LAR are used to assess the R^2 of the proposed TUR model. Each bar represents the weight of a data point used in the model. Right Panel: Model visualization with data points indicated by circles. The size of each circle corresponds to the specific weight shown in the left panel. Two outliers are marked with cross markers.
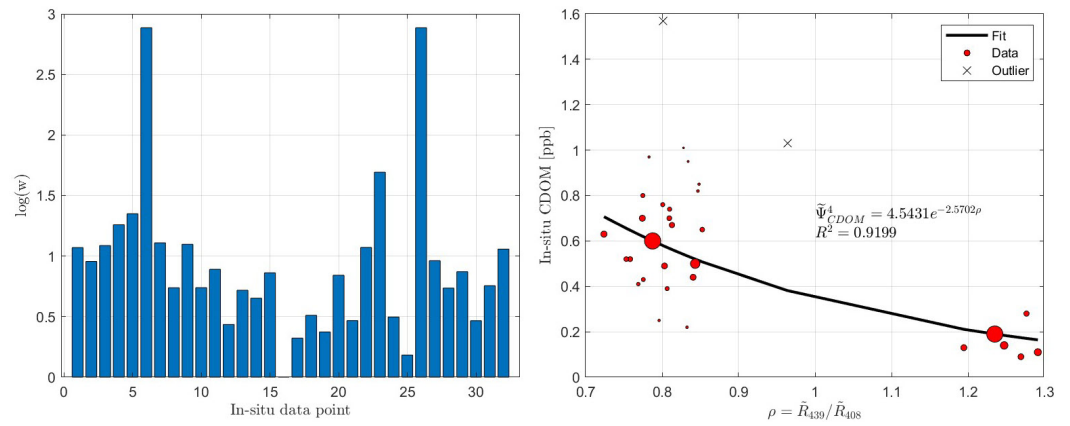
Figure S8. Left Panel: Specific weights (for 32 points) derived from LAR are used to assess the R^2 of the proposed CDOM model. Each bar represents the weight of a data point used in the model. Right Panel: Model visualization with data points indicated by circles. Circle sizes correspond to the specific weights shown in the left panel. Two outliers are indicated by cross markers.

7. Statistical evaluation for newly proposed regression-based algorithms

Similarly to existing algorithms, the performance of newly proposed regression-based algorithms has been assessed for each algorithm and corresponding water quality parameter. However, among all the possible and conducted combinations, we present only the results of best-performing band ratios for corresponding regression models.

Table S8. Statistical summary of regression models with the best performance for *Chl a* retrieval.

Regression model	R^2	RMSE	APD (%)	RPD (%)	M band	N band
			Chl a			
$\widetilde{\Psi}^1_{Chl\,a}$	0.9528	0.2038	21.9418	-3.7816	664	671
$\widetilde{\Psi}^2_{Chl\,a}$	0.9190	0.1821	17.5841	3.3349	664	673
$\widetilde{\Psi}^3_{Chl\,a}$	0.9175	0.1832	17.6008	3.7347	664	673
$\widetilde{\Psi}^4_{Chl\,a}$	0.9577	0.1853	17.5460	0.3254	673	664

Table S9. Statistical summary of regression models with the best performance for TUR retrieval.

Regression model	R^2	RMSE	APD (%)	RPD (%)	M band	N band
			TUR			
$\widetilde{\Psi}^1_{TUR}$	0.9205	0.7079	19.8099	1.4558	730	777
$\widetilde{\Psi}^2_{TUR}$	0.8616	0.6613	21.8016	8.9411	784	755
$\widetilde{\Psi}^3_{TUR}$	0.8610	0.6559	21.1616	6.2539	784	753
$\widetilde{\Psi}^4_{TUR}$	0.9194	0.7399	18.5130	-3.5596	784	728

Table S10. Statistical summary of regression models with the best performance for CDOM retrieval.

Regression model	R^2	RMSE	APD (%)	RPD (%)	M band	N band
			CDOM			
$\widetilde{\Psi}^1_{CDOM}$	0.9045	0.3165	49.4806	20.1483	456	439
$\widetilde{\Psi}^2_{CDOM}$	0.7938	0.3160	81.2970	56.1188	516	531
$\widetilde{\Psi}^3_{CDOM}$	0.7889	0.3197	86.7708	61.2583	456	540
$\widetilde{\Psi}^4_{CDOM}$	0.9199	0.2904	39.6592	7.9484	439	408