

Satellite Remote Sensing of Inland Water Quality: Sensor Evaluation, Atmospheric Correction, and Empirical Modeling

Mohsen Ansari

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the degree of
Doctor of Philosophy in Geography

Department of Geography, Environment and Geomatics
Faculty of Arts
University of Ottawa

Supervisor:
Anders Knudby

© Mohsen Ansari, Ottawa, Canada, 2026

Abstract

Inland waters represent a critical yet understudied environment in satellite Remote Sensing (RS), occupying a unique position between terrestrial and open oceans systems that makes them particularly challenging to monitor from space. Despite their ecological and socioeconomic importance, only 60% of known inland waters meet good ambient water quality standards, underscoring the need for cost-effective, large-scale monitoring solutions. Satellite RS offers such a solution, but its effectiveness depends on three interconnected components: suitable sensors, accurate atmospheric correction, and robust retrieval models. This thesis advances each of these components through four studies.

Chapter 1 provides a comprehensive review of satellite RS for inland Water Quality Parameter (WQP) retrieval, synthesizing developments in sensor technologies, atmospheric correction algorithms, and Machine Learning (ML) models. The review identifies three key research gaps that motivate the subsequent chapters: the potential of hyperspectral data for inland WQP retrieval, the impact of Adjacency Effect (AE) correction on WQP retrieval, and the influence of spatial autocorrelation on ML model evaluation.

Chapter 2 addresses hyperspectral sensor capability by evaluating Wyvern Dragonette-001 against Sentinel-2 MSI for retrieving chlorophyll-*a*, colored dissolved organic matter, dissolved organic carbon, suspended solids, and turbidity in the St. Lawrence River, Canada. Wyvern demonstrated competitive performance for turbidity retrieval but underperformed relative to Sentinel-2 for chlorophyll-*a*, colored dissolved organic matter, presumably due to the absence of blue spectral bands. These results highlight that hyperspectral capability alone does not guarantee superior WQP retrieval performance.

Chapter 3 addresses atmospheric correction by investigating the impact of AE correction on turbidity retrieval from Landsat 8/9 and Sentinel-2 imagery in the St. Lawrence River. Applying T-Mart AE correction prior to atmospheric correction with ACOLITE's Dark Spectrum Fitting technique improved retrieval accuracy for both sensors, with Root Mean Square Error reductions of 48.8% and 28.4% for Landsat 8/9 and Sentinel-2 respectively, compared to atmospheric correction alone. The most pronounced

improvements were observed within 300 meters of the shoreline, where AE from surrounding land are strongest, demonstrating that AE correction is a meaningful and necessary preprocessing step in riverine environments.

Chapter 4 addresses WQP retrieval modeling through a comprehensive comparison of 24 statistical and ML models for salinity mapping in the Karun River, Iran, using Landsat 8 OLI imagery. Decision tree-based ensemble models, particularly Gradient Boosted Decision Trees, achieved the highest accuracy while also exhibiting non-significant residual spatial autocorrelation. A key finding is that strong statistical performance metrics do not necessarily indicate spatially uncorrelated residuals, and vice versa, demonstrating that residual spatial autocorrelation analysis is an essential complement to conventional performance evaluation when satellite imagery is used to map WQP.

Together, these contributions demonstrate that advances in inland WQP retrieval depend on progress across three distinct but interconnected components: sensor design, atmospheric correction algorithms, and retrieval methodology. Moreover, the findings provide actionable guidance for researchers and practitioners developing satellite-based monitoring systems for inland waters.

Acknowledgements

This work would not have been possible without the generous financial support of the Shallow Water Earth Observation Lab, the Canadian Space Agency, and the University of Ottawa Admission and Doctoral International Scholarships. I am also grateful to Mitacs and the Ottawa Hospital Research Institute, which allowed me to pursue a project outside my thesis. I sincerely thank the International Ocean Colour Coordinating Group (IOCCG) for facilitating and funding my attendance at the 6th International Ocean Colour Science Meeting in 2025, and INRS for organizing and funding my participation in the IEEE GRSS Summer School on Advanced SAR Polarimetry in 2025. I am also thankful for travel grants provided by the Faculty of Arts, GSAED, IOCCG, and the Canadian Remote Sensing Society, which enabled me to present my work at various conferences.

I am immensely grateful to my supervisor, Dr. Anders Knudby, for creating an environment where I could truly learn, not simply complete a project. Your unwavering support, scientific vision, and dedication to your students' growth have shaped my academic journey in ways I will carry forward long after this thesis. To Dr. Saeid Homayouni, my Master's supervisor, thank you for your continued confidence in me and for always being a source of guidance and encouragement. I also want to thank Dr. Michael Sawada, from whom I gained foundational skills in coding and GIS, and whose attention to detail has been instrumental in my development as a researcher. Thank you to Dr. Meisam Amani for your critical feedback, your leadership of the smartHarbour project, and the motivation your engagement brought to my work. Special thanks to Dr. Bahram Salehi for your constructive comments during the thesis evaluation, which genuinely improved this work.

Lastly, I would like to thank my parents and my sister for your unconditional love and support. None of my achievements, past or present, would have been possible without you. I am also deeply grateful to God for the knowledge, opportunities, and blessings I have been given, and I hope to always use what I have learned in a way that is worthy of them.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iv
List of Figures.....	ix
List of Tables.....	xiii
Introduction.....	1
Chapter 1 : Retrieving Inland Water Quality Parameters via Satellite Remote Sensing: Sensor Evaluation, Atmospheric Correction, and Machine Learning Approaches.....	12
1.1 Introduction.....	13
1.2 Methodology for Article Selection and Filtering.....	17
1.3 Earth Observation Satellite Sensors for Modeling WQPs in Inland Waters.....	17
1.3.1 Ranking Satellite Sensors for Retrieving Chl-a, CDOM, and NAP.....	22
1.4 Atmospheric Correction for Inland Waters.....	25
1.4.1 Additive and Multiplicative Atmospheric Effects	26
1.4.2 Sun Glint and Air–Water Interface Correction	27
1.4.3 Water Vapor Absorption, Rayleigh Scattering, and Gas Absorption	28
1.4.4 Aerosol Contributions	28
1.4.5 Adjacency Effect.....	31
1.4.6 Comparison of Atmospheric Correction Algorithms for OLI and MSI Sensors	33
1.5 Machine Learning Models for Retrieving Water Quality Parameters	34
1.5.1 Improving Accuracy and Generalization in Machine Learning.....	37
1.5.2 Challenges and Solutions for Addressing Spatial and Temporal Autocorrelation in Machine Learning Models	38
1.5.3 Enhancing Dimensionality in Inland Water Remote Sensing.....	40
1.6 Conclusions.....	41
1.7 Appendix A.....	46
Chapter 2 : Evaluating Wyvern Dragonette-001 and Sentinel-2 for Inland Water Quality Retrieval in the St. Lawrence River.....	48
2.1 Introduction.....	49
2.2 Study Areas and Data.....	50
2.2.1 Study Areas.....	50
2.2.2 In-situ and Satellite Data	51
2.3 Data Processing.....	53

Table of contents

2.3.1 Atmospheric Correction.....	54
2.3.2 Data Splitting	55
2.3.3 Regression Model	55
2.3.4 Performance metrics	56
2.4 Results.....	56
2.4.1 Inter-sensor comparison of ρ_w and spectral signatures	56
2.4.2 WQP Retrieval and Model Performance Assessment.....	60
2.4.3 Spatial distribution of retrieved WQPs	61
2.5 Discussion.....	63
2.5.1 ρ_w	63
2.5.2 CDOM and Chl- <i>a</i>	64
2.5.3 DOC and SS.....	66
2.5.4 Turbidity	66
2.6 Conclusion.....	67
2.7 Appendix A.....	68
2.8 Appendix B.....	69
Chapter 3 : Assessing the Impact of T-Mart Adjacency Effect Correction on Turbidity Retrieval from Landsat 8/9 and Sentinel-2 Imagery (Case Study: St. Lawrence River, Canada).....	70
3.1 Introduction.....	71
3.2 Study area and data	74
3.2.1 Study area.....	74
3.2.2 In-situ and satellite data	75
3.3 Methodology.....	76
3.3.1 Atmospheric correction.....	77
3.3.2 Dataset creation.....	77
3.3.3 Regression Model	80
3.3.4 Performance metrics	81
3.4 Results.....	82
3.4.1 Matchups.....	82
3.4.2 Turbidity retrieval	82
3.4.3 Variability of correction results with distance to shore	83
3.4.4 Visual comparison	84
3.5 Discussion.....	87
3.5.1 Performance Comparison Among pre-processing options.....	87

Table of contents

3.5.2 Variation of Scenario Performance with Shoreline Distance.....	87
3.5.3 Visual Assessment of Turbidity Maps	88
3.6 Conclusion.....	89
Chapter 4 : River Salinity Mapping through Machine Learning and Statistical Modeling using Landsat 8	
OLI Imagery.....	91
4.1 Introduction.....	92
4.2 Study area and data	98
4.2.1 Study area.....	98
4.2.2 Landsat-8 OLI images.....	100
4.2.3 In situ salinity data	101
4.3 Methodology.....	103
4.3.1 Feature Engineering	103
4.3.2 Regression Methods	104
4.3.2.1 Statistical Models.....	105
4.3.2.2 Kernel-Based Models.....	106
4.3.2.3 Decision Tree-Based Models	106
4.3.2.4 Neural Network-Based Models.....	106
4.3.3 Model Evaluation.....	107
4.3.3.1 Evaluation Metrics	108
4.3.3.2 Residual Spatial Autocorrelation	109
4.4 Results.....	110
4.4.1 Feature Selection.....	110
4.4.2 Performance Assessment	114
4.4.3 Spatial Analysis	115
4.4.4 Salinity Maps	117
4.5 Discussion.....	119
4.5.1 Key Indices and Band Combinations for Salinity Mapping	119
4.5.2 Performance Evaluation across Regression Models: Accuracy and Error Analysis.....	120
4.5.3 Residual Spatial Autocorrelation Analysis	121
4.5.4 Salinity Map Validation.....	123
4.6 Conclusion	124
4.7 Appendix C	125
4.7.1 Computation of Salinity.....	125
4.7.2 Feature Selection.....	126

Table of contents

4.7.2.1 Random Forest Feature Importance Score.....	126
4.7.2.2 Sobol' Sensitivity Analysis.....	126
4.7.3 Regression Models' Hyperparameters.....	127
4.7.3.1 Statistical Models.....	127
4.7.3.2 Kernel-Based Models.....	128
4.7.3.3 Decision Tree-Based Models.....	129
4.7.3.4 Neural Network-Based Models.....	131
4.7.4 Genetic Algorithm.....	132
4.7.4.1 Overview of GA.....	132
4.7.4.2 Genetic Algorithm for Hyperparameter Tuning	133
4.7.5 Moran's I Reference Distribution.....	137
Conclusion	139
References.....	147

List of Figures

Figure 1.1. Yearly publications from 1984 to 2024 were indexed by Scopus, the largest peer-reviewed article citation database. Data were downloaded on 15 March 2025. The vertical dashed lines indicate the launch years of Landsat 8 (2013) and Sentinel-2 (2015). The full Queries (1–4) used for data collection are detailed in Appendix A.	14
Figure 1.2. Overview of selection and filtering process.	17
Figure 1.3. Representative R_{rs} spectrum (dark green solid line) for inland waters in the 400 to 750 nm range (Gurlin, 2012; Mishra et al., 2017) and the spectral range of the sensors listed in Table 1.2. The spectral range corresponding to the visible region (400–700 nm) is shown using true colors, while wavelengths above 700 nm are displayed in dark red and those below 400 nm in dark blue. The SNR values are annotated within the corresponding spectral ranges. For bands where exact SNR values were not available, approximate values are indicated using the symbol "~" or expressed as being greater than (>) or less than (<) a certain threshold. For hyperspectral sensors, a general representative SNR value is reported. Vertical lines mark the key spectral bands relevant for modeling Chl- <i>a</i> , CDOM, and NAP. Labels 1–4 indicate critical points of the R_{rs} spectrum characteristic of inland waters.	20
Figure 1.4. Methodology for ranking satellite sensors based on their suitability for monitoring concentration of Chl- <i>a</i> , NAP, and CDOM.	23
Figure 1.5. Concept diagram showing the pathways through which sunlight reaches a remote sensing detector. The image was generated with the help of DeepAI (DeepAI, 2025).	26
Figure 1.6. Workflow of supervised machine learning models. The image was generated with the help of DeepAI (DeepAI, 2025).	36
Figure 2.1. Study area. (a) Geographic overview of the study region within the St. Lawrence River, southern Quebec, Canada. (b) Spatial extent of the Contrecoeur and Lake Saint Pierre sites.	51
Figure 2.2. Spatial distribution of in-situ measurements for (a) CDOM, (b) Chl- <i>a</i> , (c) DOC, (d) SS, and (e) turbidity. Symbol color and size are scaled according to the magnitude of the measured WQP values. The black solid line in Contrecoeur site (c-e) separates the northern and southern sections associated with the 06 September 2024 and 02 August 2024 field campaigns, respectively.	53
Figure 2.3. ρ_w product derived from Wyvern imagery over the Contrecoeur site on 06 September 2024. Panels (a-w) correspond to Wyvern spectral bands. The central wavelength (λ , nm) of each band is indicated at the top of each panel.	57

List of Figures

Figure 2.4. ρ_w product derived from S2 imagery over the Contrecoeur site on 06 September 2024. Panels (a-k) correspond to S2 spectral bands. The central wavelength (λ , nm) of each band is indicated at the top of each panel.	57
Figure 2.5. Pixel-wise comparison of S2 and Wyvern ρ_w for seven spectrally corresponding band pairs (panels a-g). Each panel shows hexbin density plots of matched observations aggregated from all three acquisition dates. The dashed line represents the 1:1 relationship, and RMSE values are reported for each band pair.	58
Figure 2.6. Sensor-specific spectral responses of WQPs derived from Wyvern and S2 imagery. For each parameter, observations were divided into two concentration ranges based on the midpoint between the minimum and maximum measured values, and the mean spectrum of each range was plotted. The left column corresponds to Wyvern and the right column to S2. CDOM, Chl-a, DOC, SS, and turbidity are shown in the first through fifth rows, respectively.	59
Figure 2.7. Scatterplots of observed versus estimated WQP values derived from the RF model for Wyvern (top row; panels (a-e), blue dots) and S2 (bottom row; panels (f-j), red dots), with CDOM, Chl-a, DOC, SS, and turbidity shown in the first to fifth columns, respectively. The dashed diagonal line represents the 1:1 line.	61
Figure 2.8. Spatial distribution of CDOM and Chl- <i>a</i> derived from Wyvern and S2 imagery over the Lake Saint Pierre site on 12 August 2025. The left column corresponds to Wyvern and the right column to S2. CDOM and Chl-a are shown in the first and second rows, respectively.	62
Figure 2.9. Spatial distribution of DOC, SS, and turbidity derived from Wyvern and S2 imagery over the Contrecoeur site on 02 August 2024. The left column corresponds to Wyvern and the right column to S2. DOC, SS, and turbidity are shown in the first, second, and third rows, respectively.	63
Figure 3.1. Study area: the St. Lawrence River, Montréal, Québec, Canada. (a) Sampling locations where turbidity measurements were collected during four field campaigns; each campaign is shown in a different color. (b) Location of the case study area. (c) Map showing turbidity values at each measurement location. (d) Example water samples collected at Sites 1 and 2 (in panel a) on August 07, 2024.	75
Figure 3.2. Spatial distribution of in-situ turbidity measurements corresponding to L8/9 acquisitions (a) and their experimental semivariogram (b). Yellow points indicate the locations of in-situ turbidity measurements, while the overlaid grid represents the spatial blocking scheme used for data splitting. Cells outlined in red denote the test dataset, those outlined in blue represent the validation dataset, and the remaining cells were assigned for model training.	78
Figure 3.3. Spatial distribution of in-situ turbidity measurements corresponding to S2 acquisitions (a) and their experimental semivariogram (b). Yellow points indicate the locations of in-situ turbidity measurements, while the overlaid grid represents the spatial blocking scheme used for data splitting. Cells	

List of Figures

outlined in red denote the test dataset, those outlined in blue represent the validation dataset, and the remaining cells were assigned for model training.....	79
Figure 3.4. Ridge-density plots of in-situ turbidity measurements in the training, validation, and test datasets for (a) L8/9 and (b) S2 matchups. Kernel density estimates are shown for values up to 45 NTU, with densities encoded using a continuous color scale. Samples exceeding 45 NTU are represented as aggregated bars. Summary statistics and sample counts for each dataset are provided within each subplot.	80
Figure 3.5. RMSE values on the test dataset for (a) L8/9 and (b) S2, obtained from the LightGBM model trained using per-pixel and kernel-based (mean and median with 3×3, 5×5, and 7×7 window sizes) reflectance inputs across the three pre-processing options. The star symbol marks the kernel that achieved the lowest RMSE for each pre-processing option.....	82
Figure 3.6. Scatterplots comparing measured and estimated turbidity values obtained from the LightGBM model for L8/9 (top row; panels (a)–(c)) and S2 (bottom row; panels (d)–(f)). Each column corresponds to one pre-processing option: No (first column; panels (a) and (d)), AC (second column; panels (b) and (e)), and AE (third column; panels (c) and (f)). Text shows RMSE, β , ϵ , and number of samples for the test dataset.	83
Figure 3.7. Variability of model performance metrics with distance to the shore, combining test data from L8/9 and S2 sensors. (a) Changes in RMSE with distance to the shore, where the numbers above the bars indicate the number of samples within each distance interval. (b) ϵ (left y-axis) alongside the minimum, maximum, and mean turbidity values (right y-axis) represented by the green shaded polygon. (c) β (left y-axis) with the interquartile range and median turbidity (right y-axis) illustrated by the blue shaded polygon.	84
Figure 3.8. Spatial distribution of turbidity estimates on August 2, 2024, derived from L8/9 (left column; panels (a), (c), and (e)) and S2 (right column; panels (b), (d), and (f)). Each row represents a different pre-processing option: AE (top row; panels (a) and (b)), AC (middle row; panels (c) and (d)), and No (bottom row; panels (e) and (f)). Black labels indicate estimated turbidity values, while white labels on the upper-right map (panel (b)) show in-situ turbidity measurements.	85
Figure 3.9. Spatial distribution of turbidity estimates on September 11, 2024, derived from L8/9 (left column; panels (a), (c), and (e)) and S2 (right column; panels (b), (d), and (f)). Each row represents a different pre-processing option: AE (top row; panels (a) and (b)), AC (middle row; panels (c) and (d)), and No (bottom row; panels (e) and (f)). Black labels indicate estimated turbidity values, while white labels on the upper-right map (panel (b)) show in-situ turbidity measurements.	86

List of Figures

Figure 4.1. Study area (Karun River, Khuzestan, Iran). Overview of 8 salinity sampling stations (red) along the Karun River. Google Earth Pro and Landsat 8 OLI images of each station is shown on the right side of the map. The yellow line on each image indicates the river's width at the respective station.....	99
Figure 4.2. Correlation of features with salinity measurements and within each other.	112
Figure 4.3. R^2 values for each model using different feature combinations. The #Feature refers to the feature number in Table 4.5, indicating that the feature with this number, along with those with lower numbers, yields the highest R^2 , as shown in the top-left corner of each panel.	113
Figure 4.4. Scatter plot illustrates the relationship between measured and estimated salinity using test data (N=30). Each point represents a model prediction, with the name and class of each model displayed at the bottom right corner. The top right and left corners display the statistical metrics: R^2 , RMSE, RMSLE, β , and ϵ . Asterisks denote Pearson correlations: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. The bottom left corner shows the number of selected features in red. A dashed line represents the 1:1 diagonal line for reference, while a solid-colored line depicts the fitted regression line. Contour lines, derived from kernel density estimates, show data distributions.	114
Figure 4.5. Box plot depicting the distribution of test data (N = 30). Each box displays the median and interquartile range, while the whiskers represent values within 1.5 times the interquartile range from the lower and upper quartiles. White circles indicate outliers beyond this range.....	115
Figure 4.6. Moran scatterplot for each of the 24 regression models, showing standardized residuals on the x-axis and their spatially lagged values on the y-axis. The slope of the OLS fit line (solid black line) equals the global Moran's I for each model. The global Moran's I value, the raw permutation p-value (p_r), and the Benjamini-Hochberg FDR-corrected p-value (p_{prf}) are reported in each panel. An asterisk (*) indicates significance at $\alpha = 0.05$ after FDR correction. Model names are presented at the bottom right.	116
Figure 4.7. Karun River salinity maps generated by the GBDT model using the entire dataset (N=102) on (a) 13 June 2013, (b) 01 February 2015, (c) 05 September 2018, and (d) 11 May 2022. Measured salinity at stations with available data is shown in red, and estimated salinity is shown in blue.	118

List of Tables

Table 1.1. Optically active and inactive WQPs.	18
Table 1.2. Overview of satellite sensors for modeling WQPs in inland waters. Green dot = highly suitable; blue dot = suitable; yellow dot = potential; red dot = not suitable; Free? = available for research. X denotes sensors without SWIR band, while ✓ indicates those with SWIR band.	23
Table 1.3. Overview of atmospheric correction algorithms for inland water remote sensing, highlighting their main advantages and disadvantages along with example algorithms.	34
Table 2.1. Summary of field campaign dates, corresponding satellite overpasses, acquisition times (Eastern Daylight Time), and associated scene identifiers for the study areas.	52
Table 2.2. Semivariogram ranges, the number of training and testing samples, and the training:testing ratio for each WQP.	55
Table 2.3. Performance metrics of the RF models for each WQP and sensor.	60
Table 3.1. Satellite overpasses with coincident in-situ turbidity measurements in 2024.	76
Table 3.2. Tuned hyperparameters of the LightGBM model for turbidity estimation under the three pre-processing options for both L8/9 and S2 datasets.	83
Table 4.1. Summary of studies that have used moderate spatial-resolution sensors to map salinity, in chronological order. NIR: near-infrared band, SWIR: shortwave infrared band, MSS: Landsat Multispectral Scanner, TM: Landsat Thematic Mapper, OLI: Landsat Operational Line Imager, ETM+: Landsat Enhanced Thematic Mapper Plus, and MSI: Sentinel-2 MSI.	95
Table 4.2. Salinity indices used in this study. ρ represents the surface reflectance.	100
Table 4.3. Date of satellite image acquisition and <i>in situ</i> data collection. Matchups for each date are available at the stations listed in the Station(s) column. For other stations, either field data were not collected, or the data were masked due to cloud cover.	102
Table 4.4. Minimum, maximum, and average salinity values (psu) and number of observations (N) at the eight salinity measurement stations.	103
Table 4.5. Results of feature engineering.	111
Table 4.6. Number of observations (N) and RMSE for each station between measured and estimated salinity using GBDT trained on all dataset (N = 102).	117
Table 4.7. Average salinity (psu) within a 1 km buffer around 5 cities in the Karun River basin, derived from salinity maps generated by the GBDT model on 4 different dates.	119

Introduction

The design of satellite Remote Sensing (RS) sensors is largely driven by the objectives of their missions. As a result, most satellite sensors have been developed and optimized either for open oceans or terrestrial applications, which together cover most of the Earth's surface (IOCCG, 2018). Consequently, research and technological developments in RS have mainly focused on these two environments. Inland waters, however, have received less attention despite their importance. Inland waters represent a unique environment between oceans and land. Although they are water bodies, they are strongly influenced by nearby land surfaces. This makes them more challenging to study from space than either open oceans or land alone, and methods and sensors developed for ocean or terrestrial studies are often not fully suitable for inland water applications.

At the same time, monitoring inland waters is becoming increasingly important. Inland waters play a key role in biogeochemical cycles, ecology, and human history (Likens, 2009). Lakes, rivers, and reservoirs support biodiversity, regulate climate, supply freshwater for drinking and agriculture, and sustain the livelihoods of billions of people worldwide (Likens, 2009; Mishra et al., 2017). However, these ecosystems are under pressure from human activities including urbanization, agricultural intensification, industrial discharge, and infrastructure development, as well as from the accelerating effects of climate change such as altered precipitation patterns, increased drought frequency, and rising water temperatures (Christodoulou et al., 2020). Additionally, only 60% of known inland waters meet good ambient water quality standards (United Nations, 2023). Inland waters act as indicators of environmental change, including eutrophication, the excessive enrichment of water bodies with nutrients, typically nitrogen and phosphorus, that stimulates algal growth and depletes dissolved oxygen (Klemas, 2011), as well as increasing turbidity, salinization, and the accumulation of Colored Dissolved Organic Matter (CDOM) (Mishra et al., 2017). Therefore, low-cost enforcement and effective monitoring of inland Water Quality Parameters (WQPs) is needed to support sustainable water management (National Research Council, 2002).

Introduction

Satellite RS provides a cost-effective way to monitor inland water quality over large areas. Compared to traditional field-based approaches, which are spatially limited, resource-intensive, and often restricted to accessible locations and infrequent sampling intervals, satellite observations can provide synoptic, spatially continuous, and temporally consistent estimates of WQPs across entire water bodies and over extended time periods. This capability is particularly valuable for large or remote water bodies where frequent in-situ sampling is logistically and economically prohibitive, and for detecting and characterizing spatially heterogeneous processes such as algal blooms, sediment plumes, and salinity gradients that cannot be adequately resolved with point-based sampling networks. The archive of satellite data now spans several decades for some sensors, enabling the detection of long-term trends in water quality that would be impossible to reconstruct from sparse field records alone. Furthermore, the increasing availability of open-access satellite data, including the Landsat and Sentinel series, and commercial satellite imagery have made satellite-based water quality monitoring increasingly accessible to researchers and water resource managers alike (IOCCG, 2026, 2018; Matthews, 2011).

WQPs can be broadly categorized into optically active and non-optically active parameters (Mishra et al., 2017), and this distinction has fundamental implications for how they can be retrieved from satellite imagery. Optically active parameters are those that directly influence the absorption and/or scattering of light within the water column and therefore leave a measurable imprint on the water-leaving reflectance spectrum observed by satellite sensors. The three main optically active constituents are Chlorophyll-*a* (Chl-*a*), CDOM, and Non-Algal Particles (NAP). Chl-*a* absorbs strongly in the blue (~438 nm) and red (~667 nm) spectral regions; it serves as the primary proxy for phytoplankton biomass and primary productivity, and elevated Chl-*a* concentrations are a defining indicator of eutrophication (IOCCG, 2018; Mishra et al., 2017). CDOM absorbs predominantly in the blue and ultraviolet spectral regions, with absorption decreasing exponentially toward longer wavelengths (Mishra et al., 2017); it plays an important role in light attenuation, carbon cycling, and the photochemical degradation of pollutants, and its concentration is often

closely linked to terrestrial carbon inputs from watershed runoff. NAP encompasses mineral particles, detrital organic matter, and other suspended materials that scatter light broadly across the visible and near-infrared spectrum (Mishra et al., 2017), directly controlling water turbidity and Secchi disk depth. Each of these constituents interacts with light, which in principle allows satellite sensors with appropriate spectral coverage and resolution to detect, separate, and quantify them. In practice, the spectral overlap between the absorption and scattering signatures of different constituents makes this inversion problem ill-posed, particularly in optically complex waters where multiple constituents co-vary (Pahlevan et al., 2022).

Non-optically active parameters, by contrast, do not directly affect the optical properties of water at concentrations typically encountered in natural environments, and therefore cannot be detected through their own spectral signature (Gholizadeh et al., 2016; Mishra et al., 2017). Parameters such as salinity, total dissolved solids, dissolved organic carbon, electrical conductivity, and nutrient concentrations fall into this category. Their monitoring with RS relies entirely on indirect approaches. One such approach is to exploit empirical relationships between the target parameter and one or more optically active constituents, for example, the inverse relationship between salinity and CDOM in many estuarine and riverine systems (Blough et al., 1993; Bowers et al., 2004; D'Sa et al., 2000; Siddorn et al., 2001), where higher salinity corresponds to lower CDOM from terrestrial sources. Another is to establish purely empirical statistical relationships between the satellite-observed spectral signal and the target parameter across a training dataset, without requiring a mechanistic link. The latter approach, which is widely adopted in ML-based retrieval, is powerful but inherently site-specific: the empirical relationships learned from one location or time period may not transfer to another where the covariance structure between spectral reflectance and the target parameter differs (Gray et al., 2024; Mishra et al., 2017).

Despite the promise of satellite RS for inland water quality monitoring, its application in inland water environments remains constrained by several challenges that do not arise, or arise less acutely, in ocean color or terrestrial RS. First, spatial resolution is a fundamental limitation. Many

inland water bodies are small or narrow: rivers frequently have widths of less than 10 meters (IOCCG, 2018), and 63.48% of global lakes cover areas less than 100 km² (Cao et al., 2019). Resolving individual water pixels without contamination from adjacent land requires imagery at sufficient spatial resolution, which is available from only a limited number of satellite platforms. At the same time, sensors with high spatial resolution often sacrifice Signal-to-Noise Ratio (SNR), spectral resolution, or temporal revisit frequency, creating trade-offs that must be carefully considered for each application (Mishra et al., 2017). Second, inland waters are optically complex, typically containing high and variable concentrations of Chl-*a*, CDOM, NAP, and other constituents that vary independently in response to different environmental drivers (Mishra et al., 2017; Pahlevan et al., 2022). This optical complexity means that the spectral signal observed by the sensor is the combined result of multiple overlapping absorption and scattering contributions, making it difficult to separate the effects of individual constituents without sufficient spectral resolution and the application of appropriate retrieval algorithms. Third, the proximity of inland waters to surrounding land creates a strong Adjacency Effect (AE), whereby light reflected from nearby vegetation, soil, built surfaces, or clouds is scattered by the atmosphere into the sensor's field of view, increasing the apparent reflectance of water pixels. The magnitude of the AE depends on the optical properties of the surrounding land, the aerosol loading and composition, and the geometry of illumination and observation (Mishra et al., 2017; Wu et al., 2024), and can be detectable at distances of up to 36 kilometers (Bulgarelli and Zibordi, 2018). Fourth, the atmospheric correction methods developed for ocean color RS, which rely on assumptions such as negligible water-leaving radiance in the near-infrared or short-wave infrared spectral regions (IOCCG, 2010; Mishra et al., 2017), are frequently violated in turbid inland waters, where particle scattering can produce reflectance at these wavelengths. Applying standard ocean color atmospheric correction algorithms to turbid inland waters can therefore result in overcorrection of the visible bands, producing negative or unrealistically low water-leaving reflectance values and ultimately degrading WQP retrieval accuracy.

Introduction

The landscape of satellite sensors available for inland water quality monitoring has evolved considerably over the past three decades. Early studies relied predominantly on coarse-resolution ocean color sensors such as Moderate Resolution Imaging Spectroradiometer (MODIS), which were designed for open ocean applications and offered limited spatial resolution, typically several hundred meters to a few kilometers, that restricted their applicability to large lakes and reservoirs. The launch of the Landsat series, beginning with Landsat 1 in 1972 and continuing with Landsat 8 and 9 carrying the Operational Land Imager (OLI), provided moderate spatial resolution (30 meters) with multispectral coverage and a 16-day revisit frequency, making it suitable for monitoring medium to large inland water bodies. The launch of Sentinel-2 MSI in 2015, with 10-20 meter resolution, a 5-day revisit time, and improved spectral coverage including red-edge bands, represented a significant advance for inland water quality monitoring and has since become one of the most widely used sensors in this field. More recently, the development of hyperspectral satellite platforms has opened new possibilities for inland water quality retrieval by providing continuous spectral coverage across the visible and near-infrared, enabling finer discrimination of optically active constituents than is possible with multispectral sensors (IOCCG, 2026). Commercial hyperspectral platforms with high spatial resolution, such as PRISMA, Environmental Mapping and Analysis Program (EnMAP), Tanager, Pixxel, and the recently launched Wyvern Dragonette series, represent the latest generation of these developments, combining the spectral richness of hyperspectral imaging with the spatial detail needed to resolve narrow rivers and complex water body geometries.

Addressing the aforementioned atmospheric correction challenges has been a major focus of research in inland water RS over the past decade (IOCCG, 2010; Pahlevan et al., 2021). A variety of algorithms have been developed or adapted for inland water applications, each making different assumptions about the optical properties of the atmosphere and the water surface. Image-based methods such as the Dark Object Subtraction (Chavez, 1988) approach and its variants provide simple corrections for additive atmospheric effects but do not account for multiplicative effects or

the AE. Physics-based radiative transfer algorithms such as ATCOR (Richter and Schlapfer, 2023) apply more rigorous corrections but require accurate characterization of the atmospheric state, which is often unavailable over inland water environments. Algorithms specifically designed for aquatic applications, such as ACOLITE (Vanhellemont and Ruddick, 2018), address some of the limitations of standard terrestrial correction algorithms by using image-derived dark targets or neural network inversion, but each retains its own assumptions. Neural network-based approaches, trained on large datasets of simulated or measured spectra, offer the advantage of simultaneously retrieving the atmospheric correction and the water-leaving reflectance without explicit assumptions about aerosol type, but their performance depends heavily on the representativeness of their training data.

The retrieval of WQP concentrations from atmospherically corrected satellite imagery constitutes another major component of the satellite-based water quality monitoring workflow. Retrieval methods can be broadly categorized into three classes (Mishra et al., 2017). Physics-based methods, including semi-analytical and quasi-analytical algorithms, use radiative transfer models that relate the water-leaving reflectance to the inherent optical properties of the water column, and from these properties infer the concentrations of optically active constituents. These methods have the advantage of being grounded in physical principles and are therefore more readily transferable across different water bodies and conditions, but they require accurate parameterization of the bio-optical relationships between constituent concentrations and their inherent optical properties, which can vary substantially across different water types. Empirical methods establish statistical relationships between satellite-observed spectral indices or band ratios and WQP concentrations measured in the field, typically using regression models applied to spectral features. These methods are highly site-specific and may not generalize beyond the conditions under which they were calibrated. ML models, which represent a class of empirical methods, use flexible data-driven algorithms, including random forests, gradient boosting, support vector regression, and neural networks, to learn complex non-linear relationships between spectral features and WQP

concentrations from large training datasets. ML methods have demonstrated good performance in a wide range of inland water quality retrieval applications and can in principle be applied to both optically active and non-optically active parameters (Gholizadeh et al., 2016), but they inherit the generalization limitations of all empirical approaches and introduce additional methodological challenges related to model selection, hyperparameter optimization, overfitting, and the evaluation of model performance on spatially structured data.

A particularly important methodological challenge in ML-based WQP retrieval, which has received increasing attention in recent years (Mahboobi et al., 2023; Mainali et al., 2019; Peng et al., 2024; Zheng et al., 2024), is the treatment of spatial autocorrelation in model training and evaluation. Spatial autocorrelation, the tendency for observations at nearby locations to be more similar to each other than to observations at distant locations, as formalized by Tobler's First Law of Geography (Tobler, 1970), is a pervasive feature of satellite-derived WQP datasets, since WQPs vary continuously in space and adjacent pixels or nearby sampling locations share environmental conditions (Kelly, 2019a; Liu et al., 2022). When ML models are trained and evaluated using randomly split datasets, spatially autocorrelated observations frequently appear on both sides of the train-test boundary, allowing the model to exploit spatial proximity rather than genuine spectral-WQP relationships to make accurate predictions on the test set (Gray et al., 2024; Kelly, 2019a). This information leakage can produce performance metrics that are substantially more optimistic than the model's true generalization ability. Autocorrelated residuals indicate the presence of unmodeled spatial processes and can lead to systematic prediction biases in specific areas of the study domain. Despite its importance, the evaluation of residual spatial autocorrelation has remained largely absent from published studies on ML-based inland water quality retrieval, and its implications for the reliability and interpretation of reported performance metrics are not yet widely appreciated.

Progress in satellite RS of inland waters has accelerated in recent years, driven by the launch of new sensors with improved spectral and spatial capabilities (IOCCG, 2018), the development of

atmospheric correction algorithms tailored to inland waters (IOCCG, 2010; Pahlevan et al., 2021), and the growing adoption of ML methods for empirical WQP retrieval (Deng et al., 2024; Yan et al., 2024). Notable recent developments include the emergence of commercial hyperspectral satellite platforms with spatial resolutions of a few meters, which combine fine spectral discrimination with sufficient spatial detail to monitor narrow rivers and small lakes (IOCCG, 2026); the development of physics-based AE correction methods using three-dimensional Monte Carlo simulations that can model the complex spatial pattern of land-induced contamination more accurately than earlier approaches (Wu et al., 2024); and the application of spatially aware model evaluation strategies that account for the autocorrelated structure of satellite-derived datasets (Liu et al., 2022; Mainali et al., 2019). However, despite these advances, key knowledge gaps remain. The performance of newly launched commercial hyperspectral sensors for inland water quality retrieval has not been systematically evaluated against established benchmarks. The impact of AE correction on WQP retrieval accuracy has not been quantitatively assessed across multiple sensors and WQPs. And the relationship between conventional statistical performance metrics and residual spatial autocorrelation in ML-based WQP retrieval models remains poorly understood. Addressing these gaps is essential for realizing the full potential of satellite RS for inland water quality monitoring.

Successful retrieval of inland WQPs depends on three main components (Mishra et al., 2017):

1. Suitable satellite sensors with the spatial and spectral characteristics required to study inland WQPs.
2. Accurate atmospheric correction methods capable of separating the weak water reflectance from atmospheric effects and contamination from nearby land areas.
3. Robust retrieval and modeling approaches that can reliably estimate WQPs.

These three components are closely connected. Even the most advanced models cannot accurately estimate WQPs from space-based radiometric measurements if the sensor lacks the necessary spectral information or if atmospheric correction introduces significant errors. Likewise,

improvements in sensor design or atmospheric correction can only be fully realized when appropriate retrieval methods are used. Therefore, advances in inland water quality monitoring require progress in all three areas rather than in a single component alone.

This thesis is built around these three components. By reviewing previous studies and identifying current research gaps, my thesis research aims to improve the use of satellite RS for inland water quality monitoring through contributions in sensor evaluation, atmospheric correction, and retrieval modeling. The thesis follows an article-based format, where each chapter is written as an independent scientific paper. The chapters focus on the three key components described above, and in combination serve to improve our ability to monitor inland WQPs with satellite RS.

Chapter 1: Comprehensive review of satellite RS studies for inland WQPs (Ansari et al., 2025a)

- Title: Retrieving Inland Water Quality Parameters via Satellite Remote Sensing: Sensor Evaluation, Atmospheric Correction, and Machine Learning Approaches
- Co-authors: Anders Knudby, Meisam Amani, and Michael Sawada
- Published in Remote Sensing; DOI: 10.3390/rs17101734
- This chapter provides a comprehensive review of satellite RS studies related to inland WQP retrieval. The review examines developments in sensor technologies, atmospheric correction methods, and WQP retrieval approaches, corresponding to the three key components identified in this thesis. Through this analysis, several research gaps were identified that motivate the subsequent chapters.

Chapter 2: Wyvern Dragonette-001 vs Sentinel-2 (in preparation)

- Title: Evaluating Wyvern Dragonette-001 and Sentinel-2 for Inland Water Quality Retrieval in the St. Lawrence River.
- Co-authors: Anders Knudby
- This chapter addresses the first component, sensor capabilities. The recently launched Wyvern Dragonette-001 sensor combines hyperspectral observations with high spatial

resolution and has strong potential for inland water quality applications; however, its suitability for retrieving inland WQPs has not yet been evaluated. This chapter evaluates the Wyvern Dragonette-001 sensor for inland water applications by comparing its performance with Sentinel-2 MSI, a widely used benchmark sensor for inland water quality studies.

Chapter 3: Assessing the Impact of T-Mart AE Correction (Ansari et al., 2025d)

- Title: Assessing the Impact of T-Mart Adjacency Effect Correction on Turbidity Retrieval from Landsat 8/9 and Sentinel-2 Imagery (Case Study: St. Lawrence River, Canada)
- Co-authors: Anders Knudby, Yulun Wu
- Published in Remote Sensing; DOI: 10.3390/rs18010127
- This chapter addresses the second component, atmospheric correction. T-Mart, a recently developed AE correction approach, has shown promise for reducing land contamination in aquatic observations, but its impact on inland WQPs retrieval has not been thoroughly investigated. This chapter investigates the influence of AE correction on turbidity retrieval. By evaluating the T-Mart approach across Sentinel-2 MSI and Landsat 8 OLI and Landsat 9 OLI-2, the study examines whether reducing land-induced contamination can improve the accuracy of turbidity retrieval.

Chapter 4: Statistical Performance vs Residual Spatial Autocorrelation (Ansari et al., 2025b). This chapter has been updated relative to what was published.

- Title: River salinity mapping through machine learning and statistical modeling using Landsat 8 OLI imagery
- Co-authors: Anders Knudby, Saeid Homayouni
- Published in Advances in Space Research; DOI: 10.1016/j.asr.2025.03.037
- This chapter addresses the third component, retrieval and modeling approaches. Machine Learning (ML) models are widely used for inland WQPs retrieval, yet relatively little

attention has been given to how spatial autocorrelation influences model evaluation and the interpretation of performance metrics. This chapter develops ML-based models for river salinity estimation and examines how residual spatial autocorrelation affects model evaluation. The results provide insights into the reliability of commonly used performance metrics and highlight considerations for future inland WQPs retrieval studies.

Together, these chapters address the three fundamental requirements for successful inland water quality monitoring from space: suitable sensors, accurate atmospheric correction, and reliable retrieval models. By contributing to each of these areas, this thesis aims to improve the accuracy, reliability, and practical application of satellite-based inland water quality observation.

Chapter 1: Retrieving Inland Water Quality Parameters via Satellite Remote Sensing: Sensor Evaluation, Atmospheric Correction, and Machine Learning Approaches

Mohsen Ansari¹, Anders Knudby¹, Meisam Amani² and Michael Sawada¹

¹Department of Geography, Environment and Geomatics, University of Ottawa, Ottawa, Canada

²Canada Centre for Mapping and Earth Observation, Natural Resources Canada, Ottawa, Canada

Abstract

Satellite remote sensing provides a cost-effective and large-scale alternative to traditional methods for retrieving water quality parameters for inland waters. Effective water quality parameter retrieval via optical satellite remote sensing requires three key components: (1) a sensor whose measurements are sensitive to variations in water quality; (2) accurate atmospheric correction to eliminate the effect of absorption and scattering in the atmosphere and retrieve the water-leaving radiance/reflectance; and (3) a bio-optical model used to estimate water quality from the optical signal. This study provides a literature review and an evaluation of these three components. First, a review of decommissioned, active, and upcoming satellite sensors is presented, highlighting their advantages and limitations, and a ranking method is introduced to assess their suitability for retrieving chlorophyll-a, colored dissolved organic matter, and non-algal particles in inland waters. This ranking can aid in selecting appropriate sensors for future studies. Second, the strengths and weaknesses of atmospheric correction algorithms used over inland waters are examined. The results show that no atmospheric correction algorithm performed consistently across all conditions. However, understanding their strengths and weaknesses allows users to select the most suitable algorithm for a specific use case. Third, the challenges, limitations, and recent advances of machine learning use in bio-optical models for inland water quality parameter retrieval are discussed. Machine learning models have limitations, including low generalizability, low dimensionality, spatial/temporal autocorrelation, and information leakage. These issues highlight the importance of locally trained models, rigorous cross-validation methods, and integrating auxiliary data to enhance dimensionality. Finally, recommendations for promising research directions are provided.

Keywords: satellite remote sensing; inland water quality parameters; water quality retrieval; machine learning models; atmospheric correction; bio-optical methods

DOI: 10.3390/rs17101734

1.1 Introduction

Limnology, the study of inland waters, focuses on the 1% of the Earth's surface covered by lakes, rivers, wetlands, and other inland water bodies (Likens, 2009). Inland waters are crucial for biogeochemical cycles, ecology, and human history (Bastviken et al., 2011; Mishra et al., 2017). However, only 60% of known inland waters meet good ambient water quality standards, based on five Water Quality Parameters (WQPs): oxygen, salinity, nitrogen, phosphorus, and acidity (United Nations, 2023). This finding underscores the need for sustainable management of WQPs. Additionally, the risk of water crises, including the decline in freshwater quality, is a global and regional concern (Franco, Emilio Granados et al., 2020). For example, water crises are widely evident in countries like the United States (Mueller and Gasteyer, 2021), Iran (Barati et al., 2023), and even Brazil (Getirana et al., 2021), which has the world's largest supply of freshwater.

Inland waters also act as indicators of environmental change, including climate change (Christodoulou et al., 2020), land use changes (Yao et al., 2023), and pollution (Parmar et al., 2016), all of which can lead to harmful impacts such as eutrophication (Klemas, 2011). One of the key challenges in environmental resource management is low-cost enforcement and monitoring (National Research Council, 2002). Remote Sensing (RS) offers cost-effective, large-scale, and frequent observations compared to traditional WQP monitoring methods. Given the importance of mapping and monitoring WQPs in inland waters and the advantages of RS methods, as well as the diverse methodologies available, research on retrieving inland WQPs using RS has gained increasing attention (Figure 1.1).

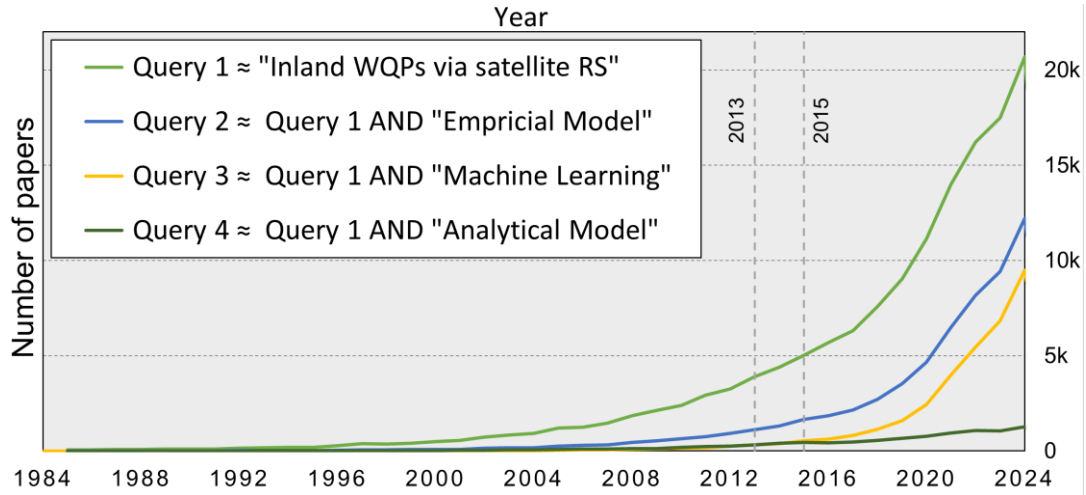


Figure 1.1. Yearly publications from 1984 to 2024 were indexed by Scopus, the largest peer-reviewed article citation database. Data were downloaded on 15 March 2025. The vertical dashed lines indicate the launch years of Landsat 8 (2013) and Sentinel-2 (2015). The full Queries (1–4) used for data collection are detailed in Appendix A.

WQPs can be categorized into three groups: (1) physical parameters, such as temperature, turbidity, Secchi disk depth, salinity, and concentrations of Total Suspended Matter (TSM) or Total Suspended Solids (TSSs), (2) chemical parameters, such as pH and concentrations of Total Dissolved Solid (TDS), nitrogen, Colored Dissolved Organic Matter (CDOM), and phosphorus, and (3) biological parameters, such as concentrations of Chlorophyll-*a* (Chl-*a*) and total bacteria. RS methods are used to monitor these WQPs by detecting the optical properties of WQPs. These optical properties are divided into two categories: (1) Apparent Optical Properties (AOPs), which depend on the light field, and (2) Inherent Optical Properties (IOPs), which are independent of the light field.

Accurately monitoring inland WQPs using satellite sensors necessitates a sensor with the following five characteristics:

1. A spatial resolution sufficient for the size of the targeted inland water bodies, which vary from large lakes to narrow rivers;
2. A high radiometric sensitivity and Signal-to-Noise Ratio (SNR) to capture weak water-leaving radiance/reflectance;

3. Spectral bands that are sensitive to variations in WQPs;
4. Frequent temporal coverage to capture the dynamic nature of inland waters;
5. Design features such as a tilting mechanism to minimize sun glint effects (the specular reflection of direct sunlight from the water surface (Kay et al., 2009)).

However, due to operational limitations, no single sensor currently meets all these requirements. Reviewing past, present, and upcoming sensors will help determine which sensors meet these requirements and to what extent. Such an evaluation also highlights the potential for combining complementary sensors, enabling the integration of multi-sensor data. This multi-sensor integration approach can enhance inland water monitoring by providing higher temporal (Sagan et al., 2020) or spatial resolution (Declaro and Kanae, 2024) observations compared to those obtained from a single sensor.

The radiance captured at the Top of Atmosphere (TOA) by sensors is affected by atmospheric attenuation and surface reflection at the air–water interface, making robust Atmospheric Correction (AC) essential to mitigate these effects and to enable accurate retrieval of remote sensing reflectance (R_{rs}) and in turn WQPs from satellite data. While AC methods for open ocean waters are well established (IOCCG, 2010), additional challenges exist for AC over inland waters. These challenges include the influence of atmospheric aerosols from urban pollution, high turbidity, and the Adjacency Effect (AE) (IOCCG, 2018; Mishra et al., 2017), leading to uncertainties in WQP modeling. Several AC algorithms have thus been developed for retrieving inland WQPs, each with their own strengths and limitations.

Bio-optical models (Smith and Baker, 1978) describe the relationship between optical properties and water constituents and can be classified as physics-based models (e.g., analytical models) and empirical models (including machine learning (ML) models). Physics-based models work by inverting radiative transfer models based on mechanistic principles. They offer physical interpretations and require several parameters like attenuation coefficients of optically active WQPs (e.g., Chl-*a*, NAP, and CDOM). Physics-based approaches only model optically active WQPs. This

is because non-optically active WQPs (e.g., concentrations of TDS, nitrogen, and phosphorus), do not directly affect the attenuation of light in water (Gholizadeh et al., 2016; Tesfaye, 2024). As a result, non-optically active WQPs cannot be modeled with radiative transfer models.

Empirical models, which are typically based on correlations between in situ measurements of WQPs and radiometric data from satellite sensors, can estimate both optically active and non-optically active WQPs (Gholizadeh et al., 2016). For non-optically active WQPs, modeling can be performed explicitly by indirectly estimating them from satellite-derived optically active WQPs (Chen et al., 2025; Vakili and Amanollahi, 2020). Alternatively, it can be carried out implicitly, assuming that non-optically active WQPs do not attenuate light in the visible spectrum but still influence the overall optical characteristics of water through correlations with optically active WQPs (Gholizadeh et al., 2016; Tesfaye, 2024). Empirical models are widely used in inland WQP studies (Matthews, 2011) and have surpassed physics-based models in publication volume in recent years (Figure 1.1). Among empirical models, ML models have become the most popular approach for retrieving WQPs in inland waters (Figure 1.1). However, the performance of ML models depends on the quality, quantity, and dimensionality of the input data. They may be prone to overfitting if the dataset is insufficient or if overfitting prevention techniques (Deng et al., 2024; Elmotawakkil et al., 2025)—such as cross-validation and regularization—are not properly applied. Additionally, ML models do not inherently incorporate physical principles; instead, they infer patterns from data that may implicitly reflect such principles (Koldasbayeva et al., 2024). Consequently, they are vulnerable to spatial and temporal autocorrelation issues and often exhibit limited generalizability (Gray et al., 2024; Liu et al., 2022).

This paper provides a comprehensive and up-to-date overview of past, present, and future satellite sensors capable of retrieving inland WQPs, evaluating their suitability for retrieving optically active WQPs (Section 3). It also explores AC algorithms used in the context of monitoring inland WQPs (Section 4). Furthermore, it discusses the strengths and challenges of ML models for retrieving inland WQPs (e.g., generalizability, spatial and temporal autocorrelation between the data used for

training and/or testing, and data dimensionality) and recent methods to address these challenges (Section 5). The review is limited in scope; it is not a meta-analysis, and it does not address the integration of in situ field data with WQP monitoring, nor does it provide a practical evaluation of AC algorithms or bio-optical models through direct implementation and comparison.

1.2 Methodology for Article Selection and Filtering

Articles were selected for inclusion in this review using a multi-stage process. First, articles were retrieved from the following databases: Scopus™, Web of Science™, Multidisciplinary Digital Publishing Institute (MDPI), and Google Scholar™. Next, filtering was performed using Query 1 and Query 2 (detailed in Appendix A), along with Query 1 AND “atmospheric correction” to refine results. Priority was given to studies published since 2013. Next, articles were skimmed, and those relevant to Sections 3–5 were selected. The chosen papers were then read in detail, and only those deemed suitable (number of articles = 164) were included in the review. The overall selection process is summarized in Figure 1.2.

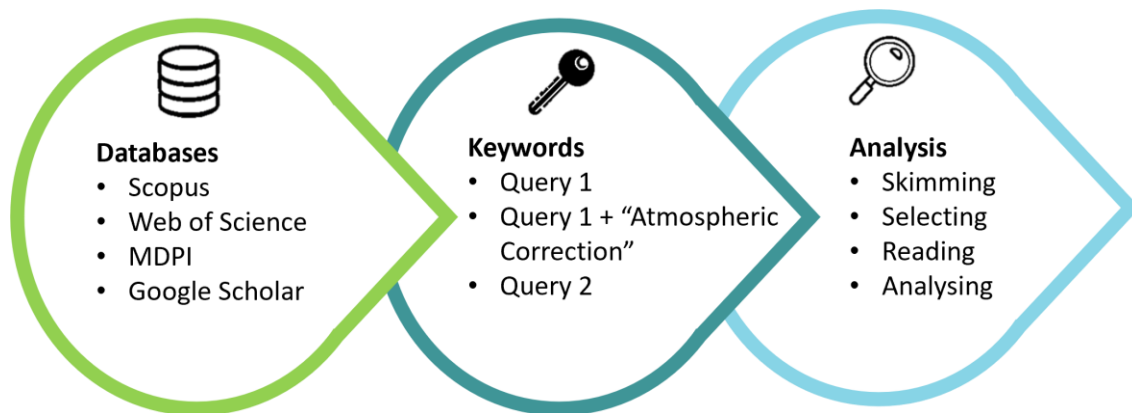


Figure 1.2. Overview of selection and filtering process.

1.3 Earth Observation Satellite Sensors for Modeling WQPs in Inland Waters

The initial step in modeling WQPs in inland waters is to choose a sensor that meets the requirements for modeling the target WQP. Since the design of a satellite sensor is determined by the objectives of its mission, most sensors are optimized for oceanic or terrestrial applications rather than inland

WQP monitoring. This section outlines the key sensor specifications required for inland WQP monitoring and concludes with a summary of past, current, and upcoming satellites, detailing their characteristics and suitability for estimating optically active WQPs. The names of the satellites and sensors and their corresponding acronyms are listed in the Abbreviations Section.

First, to obtain a valid pixel over a water body, it is generally recommended to use a square configuration of approximately 3×3 or 4×4 pixels (IOCCG, 2018). This criterion makes sensor selection dependent on the size of the study area. For example, 36.52% of global lakes cover areas larger than 100 km^2 (Cao et al., 2019), meaning that even sensors with coarse spatial resolution, such as the Ocean Color Imager (OCI) on the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) satellite (1200 m), could be suitable. However, for narrow inland water bodies like most rivers, which are often under 10 m wide (IOCCG, 2018) (requiring an approximately 3 m spatial resolution), sensors with mid-to-high spatial resolution are essential.

Second, a suitable sensor must include spectral bands sensitive to variations in the WQP of interest. Ideally, the sensor should have a spectral resolution of 5 nm from 380 to 737 nm (Dekker et al., 2018). Table 1.1 lists the key optically active WQPs, some of the main non-optically active WQPs that can be indirectly estimated from optically active WQPs, and the associated references.

Table 1.1. Optically active and inactive WQPs.

Optically Active WQPs	Correlated Non-Optically Active WQPs	References
Chl- <i>a</i>	Phosphorus, nitrogen, dissolved oxygen, and chemical oxygen	(Fu et al., 2024; Sagan et al., 2020; Y. Zhou et al., 2024)
CDOM	TDS, salinity, and electrical conductivity	(Ansari et al., 2025c; Blough et al., 1993; D'Sa et al., 2002; Siddorn et al., 2001)
TSM and TSS	TDS, salinity, and electrical conductivity	(Adjovu et al., 2023b; Khadim et al., 2017; Lavery et al., 1993)

Figure 1.3 illustrates a typical R_{rs} spectrum for inland waters within the 400–750 nm range (Gurlin, 2012), identifying four critical points where optically active WQP attenuation is pronounced. Due to the absorption of Chl-*a* at around 438 and 676 nm, two prominent local minima (i.e., points 1 and 3 in Figure 1.3) are observed in R_{rs} around these wavelengths. Thus, spectral bands around

418, 438, and 458 nm (referred to as series 1), as well as 647, 667, and 697 nm (referred to as series 2), are useful for monitoring the concentration of Chl-*a* and any parameters that covary with it (IOCCG, 2018). Unlike Chl-*a*, CDOM does not scatter light measurably and shows a smooth decrease in absorption from blue to red wavelengths. As a result, CDOM does not have a distinct spectral band for RS estimation (Mishra et al., 2017). To monitor the concentration of CDOM, spectral bands outside the pigment absorption regions, particularly in the blue-to-green spectrum, such as 380, 412, 425, and 440 nm, are recommended (IOCCG, 2018). For monitoring Non-Algal Particle (NAP) concentration, it is recommended to use spectral bands that fall outside the algal pigment absorption and fluorescence regions (IOCCG, 2018). In Figure 1.3, points 2 and 4 correspond to wavelengths where algal pigment absorption is minimal and where only pure water exhibits strong absorption, respectively (Gurlin, 2012; IOCCG, 2018; Mishra et al., 2017). Accordingly, wavelengths near point 2, such as 475, 510, 532, 542, 555, and 640 nm (IOCCG, 2018; Wolanin et al., 2016) (referred to as Series 1), and near point 4, such as 710, 748, and 775 nm (Gilerson et al., 2008; IOCCG, 2018; Ruddick et al., 2006) (referred to as Series 2), are effective for NAP concentration monitoring (IOCCG, 2018; Mishra et al., 2017). This is because they minimize residual pigment absorption (Babin et al., 2003). Given that NAPs are a key component of TSM, which also includes algal biomass (IOCCG, 2018), selecting appropriate wavelengths is essential for accurate TSM concentration monitoring. In this regard, while longer wavelengths (>700 nm) are more effective for high TSS concentrations, shorter wavelengths, particularly in the green region, yield better results for lower TSS levels (Ruddick et al., 2006).

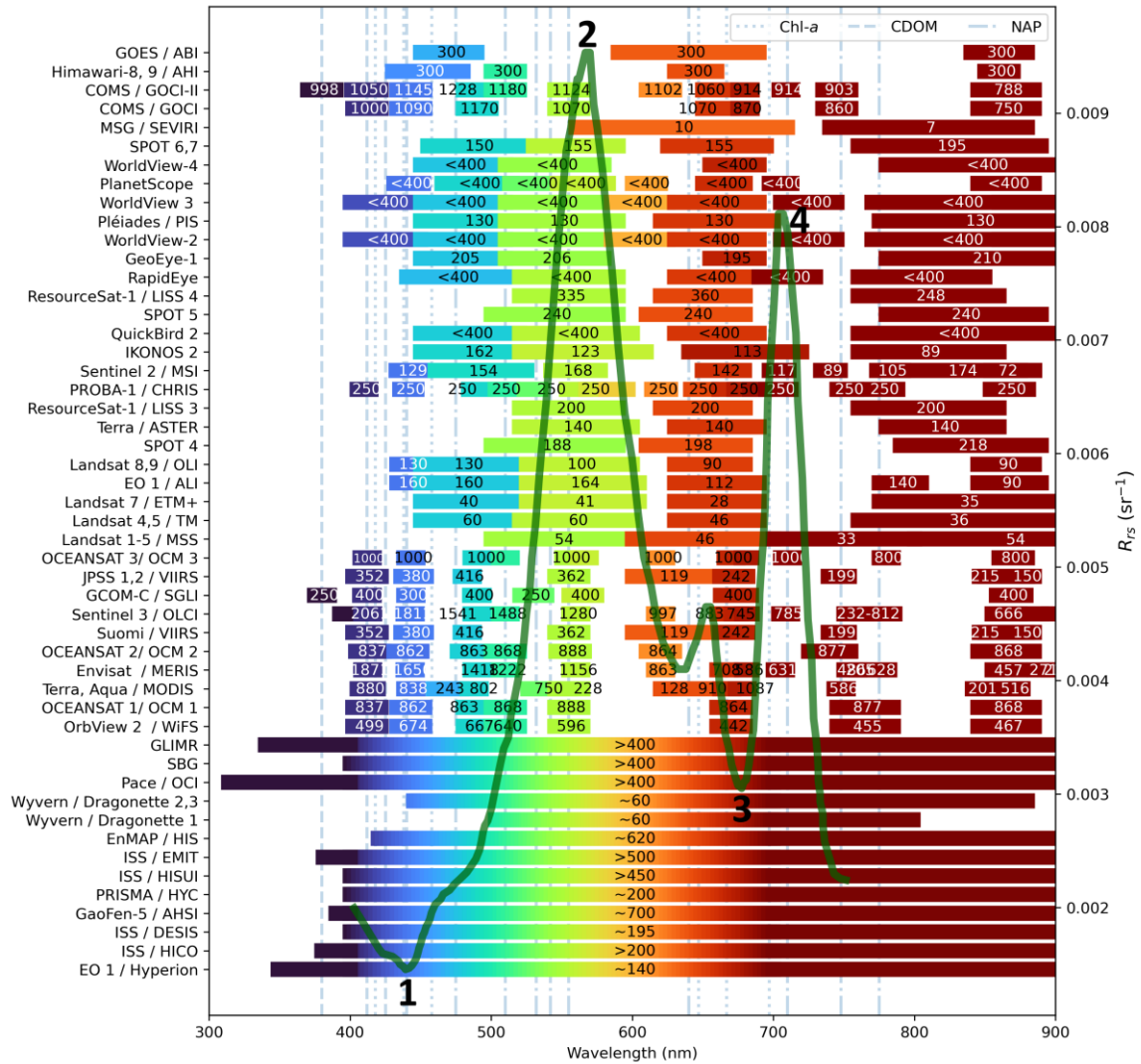


Figure 1.3. Representative R_{rs} spectrum (dark green solid line) for inland waters in the 400 to 750 nm range (Gurlin, 2012; Mishra et al., 2017) and the spectral range of the sensors listed in Table 1.2. The spectral range corresponding to the visible region (400–700 nm) is shown using true colors, while wavelengths above 700 nm are displayed in dark red and those below 400 nm in dark blue. The SNR values are annotated within the corresponding spectral ranges. For bands where exact SNR values were not available, approximate values are indicated using the symbol "~" or expressed as being greater than (>) or less than (<) a certain threshold. For hyperspectral sensors, a general representative SNR value is reported. Vertical lines mark the key spectral bands relevant for modeling Chl-a, CDOM, and NAP. Labels 1–4 indicate critical points of the R_{rs} spectrum characteristic of inland waters.

Third, given that water absorbs most light and contributes only about 10% of the radiance detected by sensors (Gordon, 1978), a sensor with high radiometric sensitivity and SNR is required. This

ensures the detection of subtle variations in the weak water-leaving radiance. The sensor SNR value is typically determined using a standard target with a 5% spectrally uniform albedo (Moses et al., 2012a). However, water reflectivity often falls below 5%, particularly in the NIR region, and the optical complexity of inland waters leads to a non-uniform reflectance pattern. Consequently, the actual SNR observed in inland waters tends to be lower than the prescribed SNR (Moses et al., 2012a; Pahlevan et al., 2014). On the other hand, spectral bands with narrow bandwidths, which are sensitive to the absorption peaks of water constituents, are essential for retrieving WQPs in inland waters. However, both these narrower bandwidths, and the high spatial resolution required for small water bodies, reduce the SNR. To reduce noise and improve the SNR, spatial resolution can be degraded through spatial aggregation (Schröder et al., 2024) where possible, but in studies of small inland waters, where water pixel availability is limited (Jorge et al., 2017; Vanhellemont and Ruddick, 2015), spatial resolution degradation may not be a feasible solution.

For effective retrieval of WQPs, an SNR above 1,000 is recommended for visible bands and at least 600 for NIR bands (IOCCG, 2012a). Meanwhile, in practice, more uncertainty (related to the SNR) in retrieving R_{rs} arises from the AC process (using the NIR bands; see Section 4 for more details) rather than sensor noise (Qi et al., 2017). Additionally, the uncertainty caused by the AC process using NIR bands decreases only slightly once the NIR SNR exceeds ~600 (Cetinic et al., 2018). Consequently, a study determined the minimum SNR for visible bands, aiming to achieve an uncertainty level comparable to that associated with the AC process (Qi et al., 2017). A minimum SNR of ~400 for visible bands was recommended when the NIR SNR was above 600 (Qi et al., 2017). This is because increasing the visible band SNR beyond this threshold does not significantly improve the overall uncertainty reduction. However, this criterion does not account for AC algorithms that utilize short-wave infrared (SWIR) bands (see Section 4 for more details).

Fourth, sufficient temporal resolution is necessary to consistently monitor dynamic inland waters and capture phenomena that vary on time scales from hourly (e.g., algal cycles) to monthly (e.g., seasonal biomass changes) (IOCCG, 2018). Geostationary satellites are positioned ~36,000 km

from the Earth and offer long integration times by maintaining a fixed view of the location. Their high temporal resolution (e.g., 5 min) makes them ideal for monitoring hourly or daily phenomena. These satellites can address challenges such as cloud cover over inland waters and AE correction by capturing images at varying sun angles throughout the day (IOCCG, 2012b).

Fifth, similar to ocean color sensors, appropriate sensors for inland waters typically incorporate features such as a depolarizer to minimize polarization sensitivity from atmospheric Rayleigh radiance, have elements that include tilting mechanisms to avoid sun glint, and employ methods for tracking sensor (e.g., internal lamps, solar diffusers, and lunar views) (Werdell and McClain, 2019).

1.3.1 Ranking Satellite Sensors for Retrieving Chl-a, CDOM, and NAP

Table 1.2, adapted and updated from (Dekker and Hestir, 2012; IOCCG, 2018; Matthews, 2011), presents a list of satellite sensors capable of retrieving WQPs, along with their specifications. Each sensor is assigned a suitability rating from “not suitable” to “highly suitable” for retrieving three optically active inland WQPs based on its spectral bands and SNR (see Figure 1.3). The ranking methodology is illustrated in Figure 1.4. As explained earlier, monitoring Chl-*a* and NAP concentration typically involves two series of spectral bands, while CDOM concentration typically involves one series. First, the spectral coverage of the sensor is checked to ensure it includes the useful bands. Next, the suitability of the spectral bands is assessed by verifying whether their SNR meets the minimum requirement: 400 in the visible spectrum when the NIR SNR is above 600 (Qi et al., 2017). It is important to emphasize that the ranking assumes the selected sensor’s spatial and temporal resolutions are adequate for the intended use case. To categorize satellite sensors, a prioritization hierarchy is applied to address overlapping classifications. Sensors classified as ocean color are prioritized under the ocean color category, followed by hyperspectral sensors, and then sensors with mid-to-high spatial resolution. The remaining sensors are all geostationary.

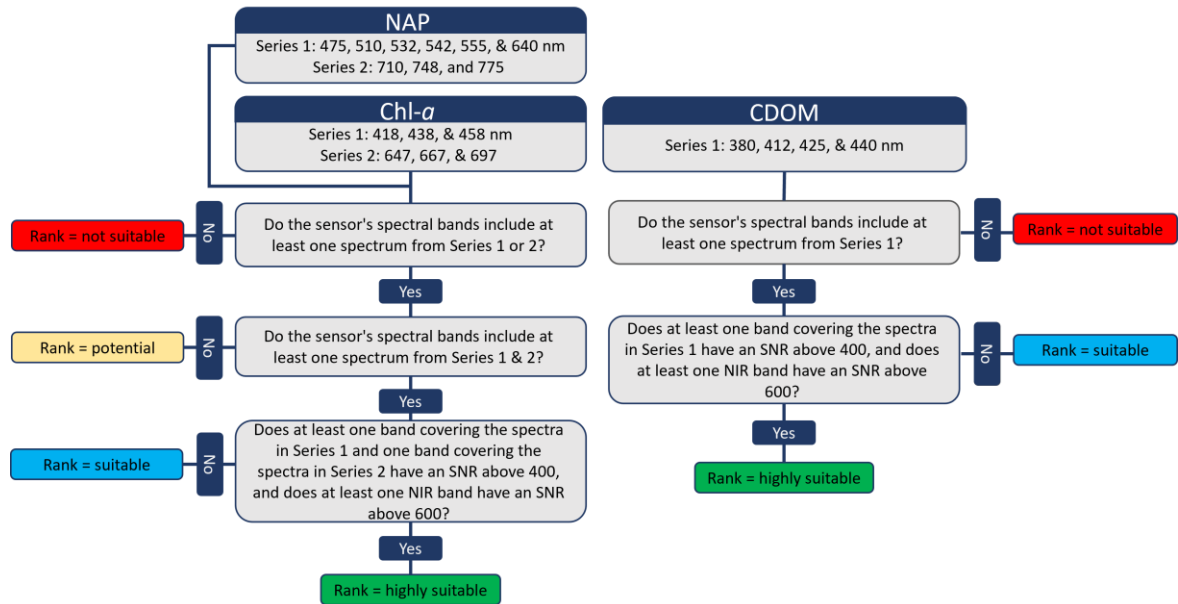


Figure 1.4. Methodology for ranking satellite sensors based on their suitability for monitoring concentration of Chl-a, NAP, and CDOM.

Table 1.2. Overview of satellite sensors for modeling WQPs in inland waters. Green dot = highly suitable; blue dot = suitable; yellow dot = potential; red dot = not suitable; Free? = available for research. X denotes sensors without SWIR band, while ✓ indicates those with SWIR band.

Type	Satellite/Sensor	Spatial Res. (m)	Spectral Res. between ~400–900 (nm) (Number of Bands)	SWIR Band	Temporal Res. (days)	Data Record Period (year)	Data Cost	WQPs		
								CDOM	Chl-a	NAP
Ocean-Color	OrbView-2/SeaWiFS	1100	402–885 (8)	X	1–2	1997–2010	Free	●	●	●
	OCEANSAT 1/OCM 1	360	402–885 (8)	X	2	1999–2010	Free	●	●	●
	Terra, Aqua/MODIS	250, 500, 1k	405–877 (13)	✓	1–2	1999–now	Free	●	●	●
	Envisat/MERIS	300	407–905 (15)	X	2–3	2002–2012	Free	●	●	●
	OCEANSAT 2/OCM 2	360	404–885 (8)	X	2	2009–2022	Free	●	●	●
	Suomi/VIIRS	375, 750	402–885 (9)	✓	1	2011–2018	Free	●	●	●
	Sentinel 3/OLCI	300	392–905 (19)	✓	2–3	2016–now	Free	●	●	●
	GCOM-C/SGLI	250	374–878 (8)	✓	2–4	2017–2022	Free	●	●	●
	JPSS 1,2/VIIRS	375, 750	402–885 (9)	✓	1	2011–2018	Free	●	●	●
	OCEANSAT 3/OCM 3	360, 1080	407–880 (12)	✓	2	2022–now	Free	●	●	●
	Pace/OCI	1200	314–895 (280)	✓	1–2	2024–now	Free	●	●	●
	SBG	30	400–2500 (63)	✓	16	2028	Free	●	●	●
	GLIMR	300	340–1040 (250)	✓	4 h	2026	Free	●	●	●

Chapter 1: Comprehensive review of satellite RS studies for inland WQPs

Hyperspectral	EO-1/Hyperion	60	349–896 (60)	✓	16	2000–2017	Free	●	●	●
	ISS/HICO	90	380–960 (100)	✓	~3	2009–2014	Free	●	●	●
	ISS/DESI	30	400–1000 (235)	✓	3–5	2018–2023	Free?	●	●	●
	GaoFen-5/AHSI	30	390–900 (100)	✓	5	2018–now	Free?	●	●	●
	PRISMA/HYC	30	400–1010 (66)	✓	29	2019–now	Free?	●	●	●
	ISS/HISUI	30	400–970 (60)	✓	2–60	2019–2023	Free?	●	●	●
	ISS/EMIT	60	381–1001 (84)	✓	~1	2022–now	Free	●	●	●
	EnMAP/HIS	30	420–900 (90)	✓	4, 27	2022–now	Free?	●	●	●
	Wyvern/Dragonette 1	5.3	503–799 (23)	✗	~2	2023–now	Charge	●	●	●
	Wyvern/Dragonette 2,3	5.3	445–880 (32)	✗	~2	2023–now	Charge	●	●	●
Mid Spatial Resolution	Landsat 1–5/MSS	60	500–1100 (4)	✗	16	1972–2013	Free	●	●	●
	Landsat 4,5/TM	30	450–900 (4)	✓	16	1982–2013	Free	●	●	●
	SPOT 4	20	500–890 (3)	✓	26	1998–2013	Charge	●	●	●
	Terra/ASTER	15	520–860 (3)	✓	16	1999–now	Free	●	●	●
	Landsat 7/ETM+	30	450–900 (4)	✓	16	1999–2022	Free	●	●	●
	EO 1/ALI	30	433–890 (6)	✓	16	2000–2017	Free	●	●	●
	PROBA-1/CHRIS	18	405–880 (17)	✓	7	2001–2021	Free	●	●	●
	ResourceSat-1/LISS 3	23.5	520–860 (3)	✓	24	2003–2013	Free?	●	●	●
	Landsat 8,9/OLI	30	433–885 (5)	✓	16	2013–now	Free	●	●	●
	Sentinel-2/MSI	10, 20, 60	442–875 (9)	✓	5	2015–now	Free	●	●	●
High Spatial Resolution	IKONOS 2	4	450–860(4)	✗	3	1999–2015	Charge	●	●	●
	QuickBird 2	2.4	450–900 (4)	✗	3	2001–2015	Charge	●	●	●
	SPOT 5	10	500–890 (3)	✓	26	2002–2015	Charge	●	●	●
	ResourceSat-1/LISS 4	5.8	520–860 (3)	✗	24	2003–2013	Free?	●	●	●
	RapidEye	6.5	440–850 (5)	✗	1	2008–2020	Charge	●	●	●
	GeoEye-1	1.64	450–920 (4)	✗	4	2008–now	Charge	●	●	●
	WorldView 2	1.8	400–1040 (8)	✗	1	2009–now	Charge	●	●	●
	Pléiades/HiRI	2	450–915 (4)	✗	1	2011–now	Charge	●	●	●
	WorldView 3	1.24	400–1040 (8)	✓	1	2014–now	Charge	●	●	●
	PlanetScope	3	431–885 (8)	✗	1	2014–2023	Charge	●	●	●

Geostationary	WorldView 4	1.24	450–920 (4)	×	1	2016– now	Charge	●	●	●
	SPOT 6,7	6	455–890 (4)	×	26	2021– now	Charge	●	●	●
	MSG/SEVIRI	1000	560–880 (2)	×	15 min	2002– now	Free	●	●	●
	COMS/GOCI	500	402–885 (8)	×	1 h	2010– 2021	Free	●	●	●
	COMS/GOCI-II	250	370–885 (12)	×	1 h	2020– now	Free	●	●	●
	Himawari-8, 9/AHI	1000	430–870 (4)	✓	10 min	2014– now	Free	●	●	●
	GOES/ABI	1000	450–880 (3)	✓	10 min	2016– now	Free	●	●	●

1.4 Atmospheric Correction for Inland Waters

Figure 1.5 illustrates the pathways through which sunlight can reach a satellite sensor, highlighting the complexity of AC. AC models the absorption of gasses and water vapor, Rayleigh scattering, aerosol absorption and scattering, and surface reflection at the air–water interface distorting the radiation reflected from the Earth’s surface detected at the TOA (IOCCG, 2010). This process is challenging for water bodies, as water absorbs most light, contributing only about 10% of the radiance detected by sensors, while 90% originates from the atmosphere (Gordon, 1978). AC for inland waters is particularly challenging due to three factors (Mishra et al., 2017):

Light from surrounding land areas or floating objects can reflect into the sensor’s view, causing non-negligible water reflectance in the NIR region. This interference disrupts AC algorithms that use NIR to derive the aerosol type and optical thickness, potentially resulting in overcorrection of R_{rs} in visible wavelengths (Hu et al., 2000; Mishra et al., 2017; Ruddick et al., 2000).

The atmosphere over inland waters is often heterogeneous due to atmospheric advection and pollution from terrestrial sources.

Inland waters typically exhibit high turbidity, leading to non-negligible reflectance in the NIR and even SWIR bands.

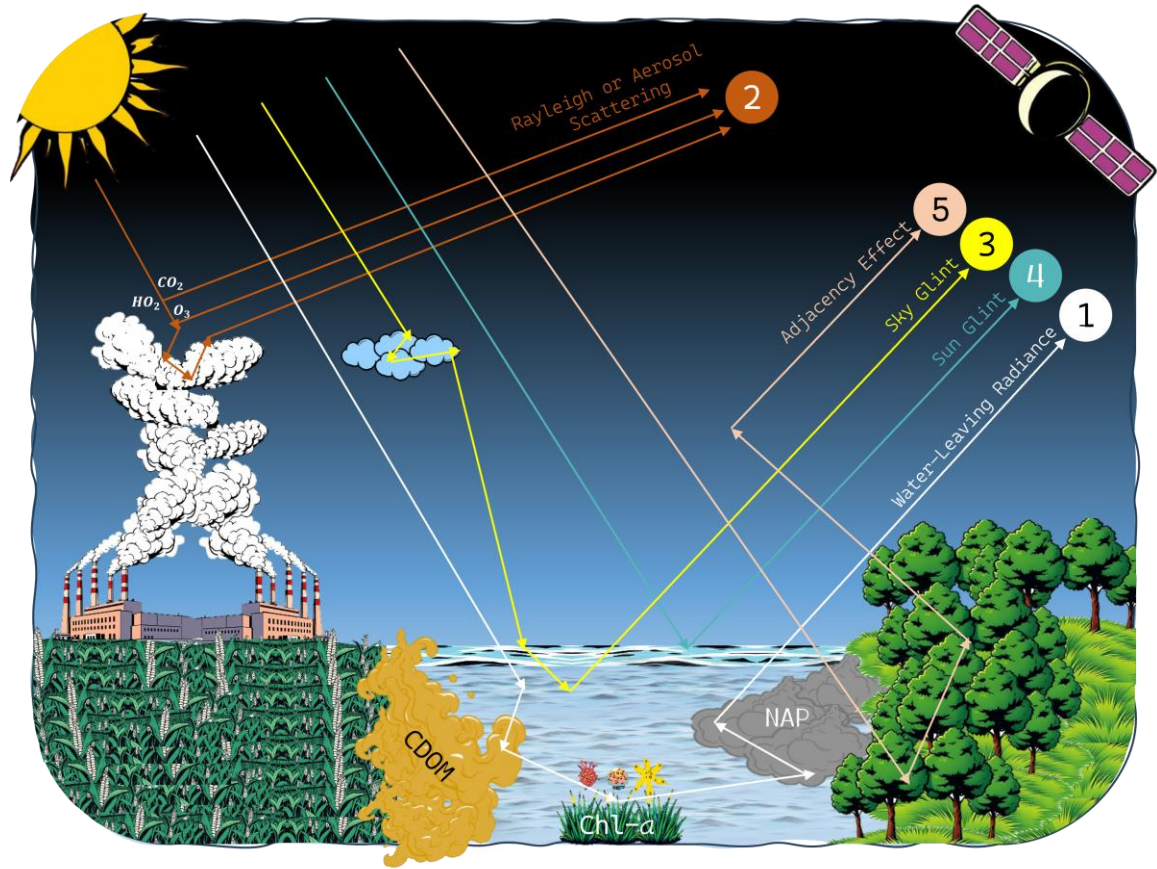


Figure 1.5. Concept diagram showing the pathways through which sunlight reaches a remote sensing detector. The image was generated with the help of DeepAI (DeepAI, 2025).

1.4.1 Additive and Multiplicative Atmospheric Effects

Atmospheric effects vary by wavelength and include additive components, such as haze, and multiplicative components, such as atmospheric transmittance (Chavez, 1996; El Alem et al., 2021). To address additive components, simple image-based methods like Dark Object Subtraction (DOS) (Chavez, 1988) are commonly used. DOS identifies dark objects (e.g., deep clear water or shadows) manually or through a histogram method. In the histogram approach, a sharp increase in the number of pixels at a nonzero radiance value is used to detect the dark object. The dark object radiance values are then subtracted from each pixel in the image, reducing the influence of atmospheric effects without requiring auxiliary data. Image-based methods are thus simple to

implement, but their performance is often limited in inland waters (Gong et al., 2008) potentially due to atmospheric heterogeneity. To handle multiplicative effects such as downward atmospheric transmittance, the cosine of the sun zenith angle (COST) (Chavez, 1996) method can be used. A more advanced image-based approach is the QUick AC (QUAC) algorithm (Bernstein et al., 2005a, 2005b), which relies on assumptions such as the presence of at least 10 endmembers and sufficiently dark pixels for baseline correction. However, these assumptions may not always hold in practice. While image-based methods have been used in various studies (Mancino et al., 2009; Moses et al., 2012b; Zhou et al., 2006), AC algorithms based on radiative transfer models are generally preferred when meteorological data are available (Allan et al., 2011). Examples of such models include the Second Simulation of a Satellite Signal in the Solar Spectrum (6S) (Vermote et al., 1997) and libRadtran (Emde et al., 2016).

1.4.2 Sun Glint and Air–Water Interface Correction

AC algorithms like Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH), ATCOR (Richter and Schlapfer, 2023), Land Surface Reflectance Code (LaSRC) (U.S. Geological Survey, 2019a), and Sentinel-2 Correction (Sen2Cor) (Main-Knorn et al., 2017) were designed to correct both additive and multiplicative atmospheric effects over land and water. However, they perform poorly in the removal of surface reflection at the air–water interface (e.g., sky glint and sun glint, see path #3 and #4 in Figure 1.5, respectively) (Cao et al., 2021), making them ineffective for inland waters.

To mitigate sun glint effects, statistical models of the sea surface, such as the Cox and Munk model (Cox and Munk, 1954), calculate the probability of glint based on wind speed and direction. This probability is then used to predict glint for a given wind vector, sun, and sensor position. Such methods rely on predicting glint using a probability distribution of sea surface slopes, which is effective at the 100–1000 m scale of ocean color sensors (Kay et al., 2009). However, for higher resolution images (1–10 m), suitable for some inland waters, the pixel size may be comparable to

the sea surface slopes, making statistical models of the sea surface less accurate (Kay et al., 2009). For high-spatial-resolution images, some glint correction methods assume negligible water-leaving radiance in the NIR (Goodman et al., 2008; Hedley et al., 2005; Lyzenga et al., 2006) or SWIR (Harmel et al., 2018a) bands or relate glint to the depth of the oxygen absorption feature (Kutser et al., 2009). Alternatively, in situations where high spatial resolution is not required, spatial aggregation could be used to facilitate statistical glint correction models to be applied to the data.

1.4.3 Water Vapor Absorption, Rayleigh Scattering, and Gas Absorption

Beyond sun glint, other atmospheric effects, such as water vapor absorption, Rayleigh scattering, and gas absorption could impact the accuracy of AC (see path #2 in Figure 1.5). Water vapor absorption can be modeled using specific absorption bands, depending on the available sensor bands (Emberton et al., 2016; Griffin and Burke, 2003). Rayleigh scattering and gas absorption, which are generally stable (Mobley et al., 2016), can be addressed with radiative transfer models like Successive Order of Scattering (SOS) (Steinmetz et al., 2011) or 6S (Vanhellemont and Ruddick, 2018).

1.4.4 Aerosol Contributions

Accurately quantifying spatial and temporal variability of aerosol contributions to TOA radiance (see path #2 in Figure 1.5) is crucial for reliable AC (Emberton et al., 2016). One common method for estimating aerosol contributions is the black-pixel assumption in the NIR region because the water-leaving radiance (see path #1 in Figure 1.5) is negligible in the NIR. The NIR black-pixel method estimates aerosol type and optical thickness by analyzing the reflectance ratio in two NIR bands at the TOA, which is then extended to the visible spectrum. Over open oceans, where aerosols are largely non-absorptive and water absorption in NIR is high, this method can reach an accuracy of up to 95% (McClain et al., 1998). However, its effectiveness is reduced for inland waters, where turbidity makes water-leaving radiance in the NIR region non-negligible (D. Wang et al., 2024),

leading to potential overestimation of aerosol effects. Thus, three alternatives are commonly employed:

Using the SWIR black-pixel assumption for wavelengths like 1240, 1640, and 2130 nm to retrieve aerosol contributions (Shi and Wang, 2009). The AC for the Operational Land Imager lite (ACOLITE) exponential extrapolation mode (Vanhellemont and Ruddick, 2015) employs this approach. However, SWIR bands often have a low SNR, especially in sensors designed for land observation like the Operational Land Imager (OLI); this can be improved by a spatially averaged filter (Wang and Shi, 2012) or by using a cross-calibration method from less turbid waters (Chen et al., 2013). Additionally, not all sensors have a SWIR band. For example, as shown in Table 1.2, 20 out of 50 sensors lack a SWIR band, which limits the use of this approach.

Modeling marine contributions to NIR bands, a method critical for sensors like the Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Medium Resolution Imaging Spectrometer (MERIS), and Geostationary Ocean Color Imager (GOCI), which lack SWIR bands. Marine contributions to NIR can be modeled by identifying the aerosol type over clear waters and transferring it to turbid waters using the nearest neighbor approach (Hu et al., 2000). Another method involves using a bio-optical model to estimate the backscattering of particles in the NIR band from the backscattering of particles in the green band (e.g., 670 nm) and subsequently calculating the water-leaving radiance in the NIR, after which the NIR black-pixel assumption AC algorithm is applied (Stumpf et al., 2003). Alternatively, aerosol scattering in NIR bands can be calculated by assuming spatial homogeneity in the NIR band ratios for aerosol and water-leaving reflectance (Ruddick et al., 2000).

Combining or switching between NIR and SWIR bands is a method where turbid pixels are processed using the SWIR-based AC algorithms, while non-turbid pixels are handled with the NIR-based AC algorithms. This method has been applied in the SeaDAS (Pahlevan et al., 2017) and the Level 2 generator (L2gen) (Shi and Wang, 2009). However, the success of L2gen depends on accurately determining the aerosol type.

Besides NIR/SWIR black-pixel assumption methods, other methods have been developed for AC over inland waters. For instance, the ACOLITE's Dark Spectrum Fitting (DSF) mode (Vanhellemont, 2019; Vanhellemont and Ruddick, 2018), an image-based method, estimates the aerosol type and optical thickness by selecting black targets with the lowest observed TOA reflectance across all available bands. This approach differs from traditional methods that rely on predefined dark bands, such as NIR and SWIR. Thus, it minimizes the risk of overcorrected reflectance after the AC process.

Land-based methods, such as Image correction for atmospheric effects (iCOR) (De Keukelaere et al., 2018) and the AC algorithm for inLand and Nearshore Coastal waters (ACLANC) (Zhao et al., 2022), extend aerosol information from dark land pixels (like dark vegetation) or adjacent areas to water areas. These methods, however, perform poorly due to the limited availability or absence of dark land pixels and the variable aerosol properties across different terrains due to terrestrial sources.

Neural Network (NN)-based AC algorithms simultaneously retrieve R_{rs} and optically active WQPs by analyzing the visible and NIR TOA radiance. The Ocean Color Simultaneous Marine and Aerosol Retrieval Tool (OC-SMART) (Fan et al., 2017), Case 2 Regional Coast Color (C2RCC) (Brockmann et al., 2016), Case2eXtreme (C2X) (Hieronymi et al., 2016), and C2X Complex Net (C2XC) are examples of NN-based AC algorithms. A recent study has shown that the use of NN-based AC algorithms for Multispectral Imagers (MSIs) is increasing (Llodrà-Llabrés et al., 2023), likely due to their assumption-free nature, ease of implementation, or applicability across a range of water conditions, from very clear to ultra-oligotrophic (Radin et al., 2020). NN-based AC algorithms perform according to the dataset they are trained on, and they therefore tend to produce better results when the IOP values in the areas where they are applied match those in the training dataset. For example, a study comparing C2RCC, C2X, and C2XC to retrieve Chl-*a* using MSIs in 32 small reservoirs across Spain found that C2RCC outperformed the other models (Cuartero et al., 2023), likely due to differences in training datasets. Because NN-based AC algorithms'

performance depends on the training data, their usefulness can be limited in heterogeneous inland waters. The NN-based AC algorithms mentioned above depend on several auxiliary data —such as salinity, temperature, ozone concentration, atmospheric pressure, and elevation—limiting their use as these auxiliary data are not always readily available. Spectral-based algorithms employ either spectral optimization, such as POLYMER (Steinmetz et al., 2011), or spectral fitting, such as Glint Removal for Sentinel-2 (GRS) (Harmel et al., 2018a), to simultaneously correct for aerosol and sun glint contributions. These algorithms require substantial computational resources. Alternatively, the Empirical Line Method (ELM) (Karpouzli and Malthus, 2003) uses in situ reflectance spectra and applies linear regression for each band to adjust the surface reflectance. While effective, ELM requires concurrent in situ spectra, which are not always available.

1.4.5 Adjacency Effect

The AE, which occurs when light from surrounding land areas reflects into the sensor view (see #5 in Figure 1.5), complicates AC, especially over inland waters (Mishra et al., 2017). This effect, influenced by aerosol optical thickness, scene geometry, and topography, can be observed on water pixels up to 36 km from the coast (Bulgarelli and Zibordi, 2018). To resolve this issue, the Improving Contrast between Ocean and Land (ICOL) processor, designed for the MERIS sensor, applies AE correction to the TOA signal before the AC process. However, it remains unadopted for mid-to-high spatial resolution sensors like OLI or MSI, widely used in inland water studies. The SIMilarity Environment Correction (SIMEC) method (Sterckx et al., 2011), used in Sen2Cor and iCOR, assumes that water's spectral shape in the NIR range is stable. However, it relies on assumptions related to red-edge and NIR bands, making it sensitive to factors such as shallow water, turbidity, and algal blooms. The SIMEC method may perform poorly due to its insufficient consideration of the aerosol type and concentration, the complexity of three-dimensional (3D) geometry, and the variability in spectral reflectance across different land cover types. Simple AE

correction is supported by 6S, but it is limited to isotropically reflecting objects, making it less suitable for inland water studies.

Correcting for the AE requires defining the point spread function of the atmosphere, which explains the origin of scattered radiation from the Earth's surface and how radiation from different surface areas contributes to the measurements made by the sensor (Wu et al., 2024). Some methods have been developed to address the point spread function, such as (Sei, 2015), which calculates azimuthally symmetric point spread functions for Lambertian surfaces. However, this method performs poorly with off-nadir observations and water specular reflectance. Additionally, the Adaptive Window by Proportion (AWP-Inland Water) model uses the point spread function to compute the contribution of the AE for the MSI (Paulino et al., 2022). The AWP-Inland Water model estimates the horizontal range of the AE required to compute reflectance due to the AE on a pixel-by-pixel basis, considering the proportion of land and water pixels within a 5 km window. A recent study introduced the Remote Sensing Adjacency Correction (RAdCor) method (Castagna and Vanhellemont, 2025a), a physics-based approach for correcting adjacency effects. RAdCor constructs the point spread function by simulating the top-of-atmosphere radiance generated by a single unit-reflectance pixel surrounded by non-reflective neighbors. Using the 6SV radiative transfer model, RAdCor estimates how radiance from surrounding land pixels is scattered into the sensor's view. By repeating this simulation across different distances and directions, RAdCor derives a scene-specific 2D point spread function, enabling correction over heterogeneous inland water environments. Recent studies have employed 3D Monte Carlo methods (Bulgarelli et al., 2014; Wu et al., 2024, 2023) to account for the AE by incorporating diverse atmospheric profiles and topographies, simulating light transfer in a 3D ocean–land–atmosphere system. Additionally, the Genetic Algorithm for AC (GAAC) method (Pan and Bélanger, 2025) retrieves the AE by solving a set of physics-based equations with seven unknown variables, including the aerosol optical depth and type, using a genetic algorithm. The number of equations matches the number of sensor spectral bands, excluding SWIR bands.

1.4.6 Comparison of Atmospheric Correction Algorithms for OLI and MSI Sensors

Recent studies have compared AC algorithms across different sensors to evaluate their effectiveness in retrieving R_{rs} and WQPs. The OLI onboard Landsat 8 and 9 and the MSI onboard Sentinel-2 have been frequently used in inland water studies. As shown in Figure 1.1, the number of publications related to these sensors has notably increased since their launch. The following highlights recent studies comparing AC algorithms in this context. For the OLI sensor, two image-based AC algorithms (DOC and COST), three physics-based algorithms (ACOLITE's exponential extrapolation mode, ATCOR, and FLAASH), and TOA reflectance were compared to estimate suspended particulate matter (El Alem et al., 2021). The image-based methods outperformed TOA reflectance and physics-based methods, with the latter underperforming due to poor atmospheric inputs and overcorrection. In another study with OLI data, ACOLITE's DSF mode, C2RCC, iCOR, L2gen, and POLYMER were evaluated for retrieving R_{rs} by comparing them with a global set of in situ R_{rs} measurements (Yan et al., 2023), finding that L2gen was the most accurate method.

For the MSI sensor, AC algorithms used for monitoring Chl-a concentration have been summarized (Llodrà-Llabrés et al., 2023), highlighting the frequent application of Sen2Cor, ACOLITE's DSF mode, C2RCC, and POLYMER. However, TOA reflectance has remained frequently used in most ecosystems except marine environments (Llodrà-Llabrés et al., 2023). POLYMER and C2RCC have been found to provide the best statistical outcomes when compared with ACOLITE's DSF mode, iCOR, L2gen, and Sen2Cor for retrieving R_{rs} (Pereira-Sandoval et al., 2019; Warren et al., 2019). Additionally, ACOLITE's DSF mode, C2RCC, FLAASH, iCOR, and L2gen have been evaluated by comparing observed and estimated R_{rs} (D. Wang et al., 2024); iCOR performed best in the Blue and NIR bands, while ACOLITE's DSF mode performed best in the green and red bands.

In one study, algorithms such as L2gen, ACOLITE's DSF mode, Sen2Cor, LaSRC, iCOR, C2RCC, C2X, and POLYMER were compared for R_{rs} retrieval from for both OLI and MSI (Pan et al., 2022). The results showed that ACOLITE's DSF mode preserved the spectral shape the best, while

NN-based AC algorithms had the lowest errors and bias. Another similar study (Pahlevan et al., 2021) evaluated ACOLITE's DSF mode, OC-SMART, SeaDAS, C2X, POLYMER, and iCOR, identifying iCOR, OC-SMART, and ACOLITE's DSF mode as top performers.

In conclusion, various AC algorithms have been developed, each with unique strengths and limitations, as summarized in Table 1.3. No single approach consistently performs the best across all inland waters and sensors. Consequently, in some cases, researchers use raw TOA data without AC for modeling WQPs via empirical models, such as cyanobacteria (Begliomini et al., 2023; Ge et al., 2024a) and Chl-*a* (Llodrà-Llabrés et al., 2023). Furthermore, while various AE algorithms have been introduced recently, their performance in retrieving WQPs remains largely unassessed and warrants further investigation.

Table 1.3. Overview of atmospheric correction algorithms for inland water remote sensing, highlighting their main advantages and disadvantages along with example algorithms.

AC Algorithms	Example	Main Advantage	Main Disadvantage
Image-based methods	DOS, COST, QUAC	Simple to implement	Provides only partial correction
Radiative transfer models	FLAASH, ATCOR, LaSRC, Sen2Cor	Corrects both additive and multiplicative atmospheric effects	Has limitations in surface reflection removal (e.g., sky glint, sun glint)
SWIR black-pixel assumption	ACOLITE's exponential extrapolation mode	Addresses nonnegligible water-leaving radiance in NIR	Limited to sensors with SWIR bands
Modeling marine contributions to NIR	(Hu et al., 2000; Ruddick et al., 2000; Stumpf et al., 2003)	Useful for sensors lacking SWIR bands	Based on assumptions that may not always hold
Combining/switching between NIR and SWIR	SeaDAS, L2gen	Addresses nonnegligible water-leaving radiance in NIR	Requires precise aerosol type determination
Land-based methods	iCOR, ACLANC	Addresses nonnegligible water-leaving radiance in NIR	Limited by availability of dark land pixels and variable aerosol properties
NN-based AC algorithms	OC-SMART, C2RCC, C2X, C2XC	Simultaneously retrieves R_{rs} and optically active WQPs; assumption-free	Requires multiple auxiliary data; highly data-dependent
Spectral-based algorithms	POLYMER, GRS	Corrects aerosol and sun glint simultaneously	Computationally intensive

1.5 Machine Learning Models for Retrieving Water Quality Parameters

ML models are algorithms that make predictions of a variable (the response) based on data related to that variable (the predictors); these models are trained to recognize relationships between the predictors and the response through optimizing model parameters to minimize a cost function, e.g.,

the difference between the predictions and a set of validation data. One of the primary types of ML models is the supervised model, where labeled data are used to train the model. In supervised machine learning, as illustrated in Figure 1.6, the process begins with satellite and in situ data. These are then combined through a matchup process to align the corresponding observations. Next, optional data preprocessing steps may be performed to clean the data, remove outliers, normalize and transform variables, and handle missing values. This is followed by feature engineering, where meaningful features, such as value from spectral bands or indices, are derived from the raw data. The dataset is then split into training and test sets, with the training data used for model training to learn patterns. The trained model is then assessed through model evaluation using the test data to measure its performance. Finally, the model is applied for prediction. Supervised models can be divided into four major categories:

1. Statistical models, which include linear models like ordinary least squares regression and non-linear models like generalized additive models;
2. Kernel-based models, such as support vector regression, which operate by mapping input variables into higher-dimensional feature spaces using a kernel function;
3. Tree-based models, such as Decision Trees (DTs), which are structured hierarchically, with each node representing a decision based on a specific feature;
4. NN models, such as multilayer perceptrons, which process raw data through multiple layers, each transforming the data into more abstract representations than the previous layer.

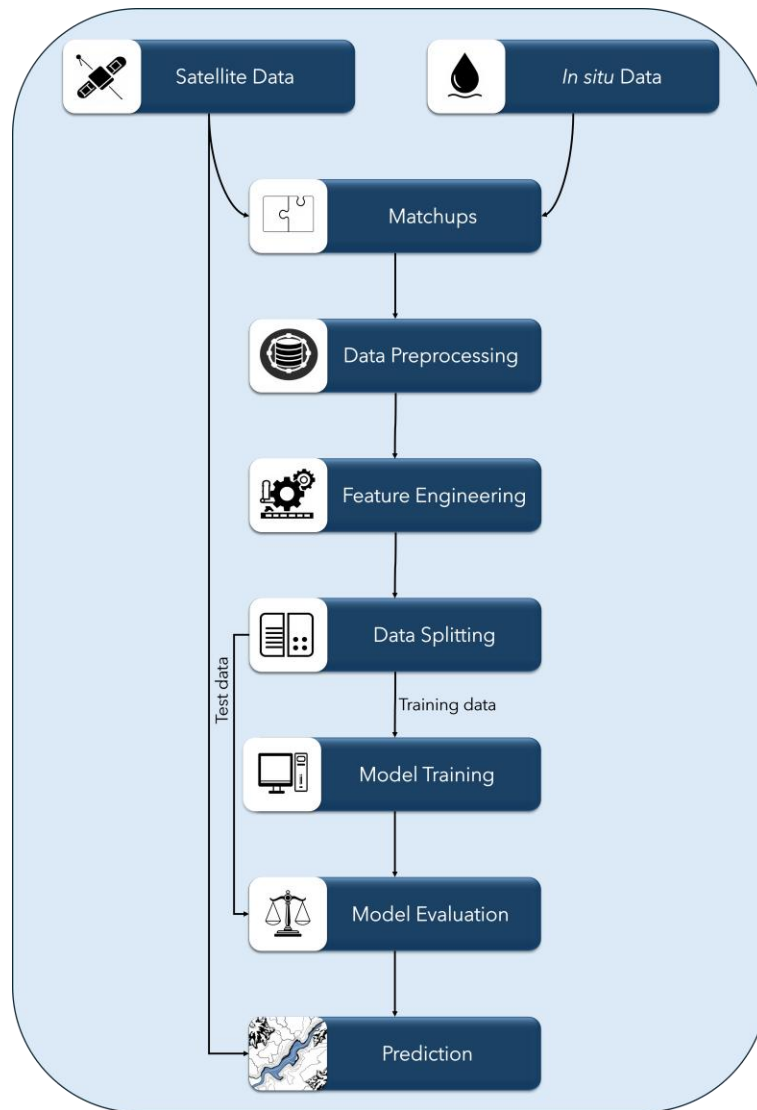


Figure 1.6. Workflow of supervised machine learning models. The image was generated with the help of DeepAI (DeepAI, 2025).

Numerous studies have compared ML model performance for retrieving inland WQPs (Ansari et al., 2025c; Ansari and Akhoondzadeh, 2019; M. Ansari and Akhoondzadeh, 2020; Ansari and Akhoondzadeh, 2021; Feng et al., 2024; Fooladi et al., 2024; Fu et al., 2024; Gao et al., 2024; Ge et al., 2024b; Guo et al., 2024; Jiang et al., 2024; R.M. Khan et al., 2024; Leggesse et al., 2024; Mamun et al., 2024; Nikoo et al., 2024; Tian et al., 2024; Q. Wang et al., 2024; Yan et al., 2025; Zhao et al., 2024; Yadong Zhou et al., 2024; Zhu et al., 2024). Recent review studies have examined ML models used for estimating organic matter and harmful algal blooms (Nguyen et al., 2024),

TDS and TSS (Adjovu et al., 2023b), and Chl-*a* (Llodrà-Llabrés et al., 2023). These reviews generally conclude that no single ML model consistently outperforms others across all WQPs and study areas (Yan et al., 2024). However, a key takeaway is that empirical methods, particularly ML models, generally outperform other bio-optical models in estimating WQPs (Llodrà-Llabrés et al., 2023; Sagan et al., 2020; Yan et al., 2024). It should be noted, however, that the performance of empirical methods is often limited to the specific regions or conditions on which they were trained, as discussed in Section 5.1. In contrast, other bio-optical models may offer better performance for extrapolation. This review focuses on the challenges associated with ML models and recent efforts to overcome them.

1.5.1 Improving Accuracy and Generalization in Machine Learning

One approach for improving ML model performance in modeling WQPs is ensemble modeling, where the outputs of several ML models are combined. For example, multiple DTs can form an ensemble model like random forest (Xia et al., 2024) or extreme gradient boosting (Li et al., 2024; Sillen et al., 2024). The underlying assumption is that an ensemble model can yield more reliable predictions in regression tasks and enhance the model's generalization by capturing diverse relationships that might be missed when relying on a single regression method. It is also possible to combine the outputs of different ML methods using simple mathematical operators, such as averaging (Park et al., 2024), or statistical methods like least absolute shrinkage and selection operator (Peng et al., 2024) and BayesianRidge (Zhang et al., 2022). Another method for combining ML outputs is Bayesian maximum entropy-based fusion (Fooladi et al., 2024; Nikoo et al., 2024), which integrates the predictions of different ML models using weighted averaging.

The main advantage of ML is its lack of assumptions (e.g., about atmospheric composition). However, due to the absence of a physical basis, ML models often perform poorly when applied beyond their calibrated range, resulting in unreliable predictions during extrapolation in ML models. To address this, a global model was developed using a Mixture Density Network (MDN)

trained on optically active WQPs from various locations worldwide (Pahlevan et al., 2022, 2020). However, the MDN has certain shortcoming: (1) it is sensitive to uncertainties in the AC process, which can affect its performance under different atmospheric or water conditions (Chegoonian et al., 2023; Pahlevan et al., 2020; Smith et al., 2021), and (2) it may not perform well when the target WQPs fall outside its calibrated range. For example, the globally trained MDN overestimated C-Phycocyanin concentration in the Billings Reservoir (Begliomini et al., 2023), likely due to a lower C-Phycocyanin to Chl-*a* ratio compared to the global training dataset used for MDN calibration (O'Shea et al., 2021). Similarly, another study (Rabia Munsaf Khan et al., 2024) trained a random forest model using both local and global datasets (i.e., GLORIA dataset (Lehmann et al., 2023)) to map Secchi disk depth in the Finger Lakes, USA. The results showed that the locally trained model outperformed the global one, possibly because the GLORIA dataset is skewed toward eutrophic water bodies, and as a result, the global model may be biased when applied to oligotrophic waters like the Finger Lakes. Therefore, it appears that locally trained ML models generally outperform globally trained ones when evaluated within the same region that the locally trained ML models were trained on. However, this performance advantage may not hold when applied to new or unseen locations.

1.5.2 Challenges and Solutions for Addressing Spatial and Temporal Autocorrelation in Machine Learning Models

Spatial and temporal autocorrelation occurs when nearby data points influence each other or are impacted by the same spatially/temporally variable factors. This is generally present in spatial data, such as satellite imagery. Autocorrelation indicates that unmodeled spatial/temporal processes might be influencing the variable of interest (Mainali et al., 2019) and violates the independence assumption, leading to biased estimates and unreliable conclusions. Moreover, based on Tobler's First Law of Geography (Tobler, 1970), assessing spatial autocorrelation in model residuals is crucial for spatial regression models (Kelly, 2019a; Ansari et al., 2025c) because identification of

spatial patterns in residuals helps determine whether the model captures all spatial dependencies in the data.

Importantly, for estimating spatial/temporal parameters like WQPs, data-driven methods need to be evaluated to understand the extent to which their outputs are influenced by spatial autocorrelation (Liu et al., 2022). This is crucial because these methods rely solely on data, without a physical foundation built in to the model structure (Gray et al., 2024). Moran's I is the most commonly used statistical tool to evaluate spatial autocorrelation (Mahboobi et al., 2023; Mainali et al., 2019; Peng et al., 2024; Zheng et al., 2024). To address spatial autocorrelation, several methods have been proposed. The spatial lag and spatial error models (Anselin, 2001) account for autocorrelation by incorporating spatially lagged variables or spatially correlated error terms. Additionally, the geographically weighted regression model (Brunsdon et al., 1998) enables localized modeling by incorporating spatially varying coefficients based on geographical coordinates. Another strategy is to include spatial and temporal information as input features. Although this may reduce the model's capacity for extrapolation, it enhances the ability to capture spatial and temporal patterns, leading to improved model performance for the training area. This was demonstrated in multiple studies (Khadim et al., 2017; Nguyen et al., 2018; Zhao et al., 2024; Zhu et al., 2024) where longitude, latitude, month, day of the week, and day of the year were used as spatial and temporal parameters.

A related issue is information leakage between training and test sets. This occurs when spatial autocorrelation exists between the training set samples and the surrounding test set samples (Feng et al., 2023). For example, if test and training data are split randomly, adjacent spatially autocorrelated data points will often end up in both sets, leading to overestimation of model performance caused by incomplete independence between the datasets (Gray et al., 2024). Other methods for splitting training and test datasets, such as leave-one-out (Pahlevan et al., 2022), K-fold cross-validation (Wang et al., 2023; Zhang et al., 2022; Zhao et al., 2023), or Monte Carlo simulations (Begliomini et al., 2023), have been employed in previous studies. However, like

random splitting, they do not ensure the complete prevention of information leakage. Methods like spatial or temporal cross-validation (Zhao et al., 2024), checkerboard evaluation (Rolf et al., 2021), and block or buffered cross-validation (Knudby and Richardson, 2023) have been employed in previous studies to reduce or prevent information leakage. These methods train models on data from certain regions or periods and evaluate their performance on data from different regions or periods (Rolf, 2023).

1.5.3 Enhancing Dimensionality in Inland Water Remote Sensing

Since ML models are data-driven, their accuracy largely depends on the quality, quantity, and dimensionality of training data. Generally, more training data lead to better model performance. Moreover, the quality, quantity, and dimensionality of the training dataset exert a higher influence on model performance than the data-driven modeling approach itself (Gray et al., 2024). The quality of data can be improved during the “data preprocessing” step (see Figure 1.6). This step involves cleaning the data, removing outliers, normalizing and transforming variables, and handling missing values to enhance overall data quality. To increase the quantity of training data, one possible approach is the use of generative artificial intelligence to create synthetic data. For instance, a recent study (Thopanaiah et al., 2024) employed a Generative Adversarial Network (GAN) for data augmentation, generating synthetic training data to enhance water quality management via RS.

Unlike terrestrial RS, which utilizes high-dimensional data such as spectral bands, spatial patterns, 3D modeling, and polarization, RS of inland waters is constrained to fewer dimensions, primarily spectral bands. This limitation reduces the ability to develop the multiple layers of representation needed to train complex ML models effectively. To overcome this limitation and enhance data dimensionality, recent studies have incorporated auxiliary data and spectral indices in inland and coastal water studies. For instance, some studies have utilized water salinity, water currents, and bathymetry data (Zhu et al., 2024) or water temperature and wind speed (Sillen et al., 2024) to map

eutrophication indicators. For example, a study on Lake Simcoe in Toronto, Canada (Guo et al., 2023), employed a wide range of auxiliary data, including air temperature, sea level, precipitation, wind speed and direction, water temperature, urban and farmland areas, nighttime light (as an indicator of economic development and human activity), and the number of mobile devices connected to cellular towers (as a proxy for population distribution) to map Chl-*a*, total phosphorus, total nitrogen, Secchi disk depth, dissolved organic carbon, and dissolved oxygen. In this context, the degree of polarization provided by the OCI could be considered a new dimension for enhancing future studies, adding another layer of information that could improve the accuracy of inland WQP estimation.

Another approach is to extract more information from spectral bands through feature engineering. For example, spectral indices combine different spectral bands through mathematical operations, either linearly or nonlinearly, and have been used in recent studies to improve the performance of ML models (Adilakshmi and Venkatesan, 2024; Jiang et al., 2024; Shahvaran et al., 2024; Xia et al., 2024). Furthermore, derivative spectroscopy has emerged as a novel method for extracting spectral features in inland water studies using hyperspectral data. Derivative spectroscopy leverages integer-order (Hou et al., 2022) or fractional-order derivatives (Fu et al., 2024; Xia et al., 2024), enabling the representation of higher-order derivatives.

1.6 Conclusions

This study reviewed decommissioned, active, and future sensors that have the potential to estimate WQPs in inland waters. We also discussed proposed AC algorithms and the challenges of estimating WQPs using ML approaches, along with proposed solutions. Below, we summarize key findings and offer recommendations for future research on WQP modeling in inland waters:

1. Not all sensors are suitable for retrieving Chl-*a*, CDOM, and NAP concentrations. Before initiating modeling, it is important to address whether the selected sensor can effectively model the target WQP (Table 1.2). Leveraging a multi-sensor integration strategy -

especially combining sensors with complementary strengths - can help overcome individual limitations.

2. High spatial resolution sensors may lack the necessary spectral resolution and SNR for inland WQP estimation, but they are the only option for small inland waters.
3. The Surface Biology and Geology (SBG) sensor, set to launch in 2028, is highly suitable for modeling Chl-*a*, NAP, and CDOM concentrations (Table 1.2). It has spatial and temporal resolutions similar to the Landsat OLI but offers higher spectral resolution, an improved SNR, and a tilting mechanism to minimize sun glint, making it highly suitable for inland WQP monitoring. However, its performance remains to be validated once operational data become available.
4. No AC algorithm has consistently outperformed others across all atmospheric and water conditions. Therefore, it is important to evaluate the suitability of a given algorithm for the specific sensor, water type, and atmospheric context before implementation. Moreover, although new AE correction methods have been proposed in recent research, their performance has not been thoroughly evaluated.
5. Recent studies demonstrate that ensemble methods achieve higher accuracy and robustness compared to single machine learning models.
6. Locally trained ML models generally outperform globally trained ones when evaluated within the same region the local model was trained on. This is because ML models tend to perform poorly in regions where they have not been calibrated, due to differences in atmospheric conditions or variations in WQP concentrations between the training data and the target area.
7. Addressing the impact of spatial and temporal autocorrelation in WQP modeling is important, particularly when using data-driven models, as it can result in biased estimates and unreliable conclusions. Additionally, identifying spatial patterns in residuals is key to assessing whether the model has captured all spatial dependencies in the data.

8. Ensuring the prevention of information leakage during the separation of training and test data is necessary for reliable performance evaluation. Such leakage can lead to inflated performance estimates and undermine the validity of the results. Methods like spatial/temporal cross-validation, checkerboard evaluation, and buffered cross-validation help mitigate the risk of information leakage.
9. Increasing data dimensionality through the integration of auxiliary variables—such as meteorological parameters—can improve the performance of ML models. This approach compensates for the limited spectral and spatial information typically available in inland water RS and allows models to capture more complex patterns. However, this may lead to overfitting if cross-validation and regularization techniques are not properly applied.

Improvements in WQP monitoring over the coming years are likely to result from gradual improvements in sensor characteristics, data quality, AC, and other pre-processing factors, as well as improvements in ML models and an understanding of how to apply such models effectively to the available data. Improvements may come from these components in isolation; for example, improved ML models, e.g., from deep learning, may produce better WQP monitoring on their own, independently of other factors. Or improvement may come from combinations of factors, such as ML models being increasingly able to extract information from the additional dimensionality of hyperspectral data or the varied information provided through multi-sensor integration. In this study, we have identified some of the most promising avenues for future improvement in WQP monitoring using remote sensing technology, and real-world testing will demonstrate which of these will provide the greatest value over the coming years.

Author Contributions

M.A. (Mohsen Ansari): conceptualization, data curation, formal analysis, investigation, methodology, resources, visualization, writing—original draft. A.K.: conceptualization, methodology, project administration, supervision, writing—review and editing. M.A. (Meisam

Amani); writing—review and editing. M.S.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Data Availability Statement

No new data were created or analyzed in this study.

Conflicts of Interest

The authors declare no conflict of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, we used OpenAI's ChatGPT (GPT-4-turbo) for English grammar corrections. For Figures 1.5 and 1.6, image generation was assisted by DeepAI. After using these tools, we thoroughly reviewed and edited the content as needed and take full responsibility for the content of the publication.

Abbreviations

The following abbreviations are used in this manuscript:

SeaWiFS	Sea-viewing Wide Field-of-view Sensor
OCM	Ocean Color Monitor
JPSS	Joint Polar Satellite System
VIIRS	Visible Infrared Imaging Radiometer Suite
OLCI	Ocean and Land Color Instrument
SGLI	Second Generation Global Imager
OCI	Ocean Color Imager
GLIMR	Geostationary Littoral Imaging and Monitoring Radiometer
HICO	Hyperspectral Imager for the Coastal Ocean
AHI	Advanced Himawari Imager

Chapter 1: Comprehensive review of satellite RS studies for inland WQPs

AHSI	Advanced Hyperspectral Imager
HYC	Hyperspectral Camera
COMS	Communication, Ocean, and Meteorological Satellite
EMIT	Earth Surface Mineral Dust Source Investigation
MSS	Multispectral Scanner
TM	Thematic Mapper
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
ETM+	Enhanced Thematic Mapper Plus
ALI	Advanced Land Imager
CHRIS	Compact High Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
SPOT	Satellite Pour l'Observation de la Terre
MSI	Multispectral Imager
HiRI	High Resolution Imager
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
GOCI	Geostationary Ocean Color Imager
OLI	Operational Land Imager
ABI	Advanced Baseline Imager
MERIS	Medium Resolution Imaging Spectrometer
PACE	Plankton, Aerosol, Cloud, ocean Ecosystem
SBG	Surface Biology and Geology
EO	Earth Observing
ISS	International Space Station
EnMAP	Environmental Mapping and Analysis Program
LISS	Linear Imaging and Self-Scanning Sensor

MSG	Meteosat Second Generation
HISUI	Hyperspectral Imager Suite
GOES	Geostationary Operational Environmental Satellite
PROBA	Project for On-Board Autonomy
HIS	Hyperspectral Imager
DESI	DLR Earth Sensing Imaging Spectrometer

1.7 Appendix A

We defined the search terms to minimize temporal bias and ensure broad coverage of relevant studies. Nonetheless, we acknowledge that shifts in terminology, such as the transition from “machine learning” to “AI”, may influence search trends.

The following query formulations were used to retrieve publications for trend analysis in Figure 1.1:

- Query 1: (“inland water*” OR “river*” OR “lake*” OR “reservoir*” OR “wetland*” OR “freshwater” OR “estuary*” OR “aquatic system*”) AND (“remote sensing” OR “satellite imagery” OR “Earth observation” OR “hyperspectral imaging” OR “multispectral imaging”) AND (“water quality” OR “turbidity” OR “chlorophyll*” OR “total suspended solid*” OR “total dissolved solid*” OR “salinity” OR “colored dissolved organic matter” OR “dissolved organic carbon” OR “particulate organic carbon” OR “electrical conductivity” OR “Secchi disk depth” OR “eutrophication” OR “harmful algal bloom*” OR “phytoplankton” OR “cyanobacteria” OR “non-algal particle*” OR “water transparency” OR “biogeochemical cycle*” OR “optically active constituent*” OR “optically inactive constituent*” OR “non-optically active constituent*”);
- Query 2: (“empirical model*” OR “data-driven model*” OR “statistical model*” OR “supervised learning” OR “unsupervised learning” OR “regression” OR “machine

learning” OR “deep learning” OR “statistical analysis” OR “linear regression” OR “Bayesian” OR “ensemble learning “);

- Query 3: (“machine learning” OR “deep learning” OR “Bayesian” OR “ensemble learning “);
- Query 4: (“physical model*” OR “mechanistic model*” OR “analytical model*” OR “semi-analytical model*” OR “quasi-analytical model*” OR “radiative transfer model*” OR “radiative transfer code*” OR “radiative transfer equation*”).

Chapter 2: Evaluating Wyvern Dragonette-001 and Sentinel-2 for Inland Water Quality Retrieval in the St. Lawrence River

Mohsen Ansari¹, and Anders Knudby¹

¹Department of Geography, Environment and Geomatics, University of Ottawa, Ottawa, Canada

Abstract

Satellite remote sensing offers a cost-effective and spatially extensive approach for monitoring inland Water Quality Parameters (WQPs). However, its application remains constrained by spectral and spatial limitations, particularly in optically complex and narrow inland waters. In this regard, Wyvern Dragonette-001, a recently launched hyperspectral sensor with 23 spectral bands and 5.3 m spatial resolution, has potential to retrieve inland WQPs. This study evaluates the quality of imagery acquired with Dragonette-001, as well as its performance for retrieving Chlorophyll-*a* (Chl-*a*), Colored Dissolved Organic Matter (CDOM), Dissolved Organic Carbon (DOC), Suspended Solids (SS), and turbidity in the St. Lawrence River, Québec, Canada, using Sentinel-2 (S2) Multispectral Imager (MSI) as a reference benchmark due to its well-established performance for inland water quality retrieval. Atmospheric correction was performed using the Remote sensing Adjacency Correction (RAdCor) algorithm, and a Random Forest (RF) regression framework was used to model empirical relationships between water-leaving spectral reflectance (ρ_w) and in-situ WQP measurements. Wyvern imagery showed systematic artifacts, including elevated ρ_w at $\lambda \geq 700$ nm and striping noise, particularly at $\lambda \geq 712$ nm. Inter-sensor ρ_w comparison across seven corresponding spectral bands revealed RMSE values exceeding 0.01 for all band pairs, with the best agreement at 550 nm. For WQP retrieval, turbidity showed the strongest performance for both sensors, with Wyvern achieving lower relative RMSE (0.84% vs. 1.02% for S2), lower ϵ (1.60% vs. 2.10%), and near-zero bias. For DOC, Wyvern showed some indication of stronger performance, while SS results were comparable between the two sensors; however, the limited sample size ($N = 21$) for these two WQPs limit robust interpretation of the results. CDOM and Chl-*a* retrievals were poor overall, with S2 outperforming Wyvern: for CDOM, normalized RMSE was 7.4% (S2) vs. 13.5% (Wyvern), and for Chl-*a* 1.9% vs. 7.3%, respectively, accompanied by lower ϵ and β values for S2. Newer Dragonette missions (Dragonette-002 and Dragonette-003), which include blue spectral bands, may improve the retrieval of WQPs, particularly CDOM and Chl-*a*.

Keywords: Inland water quality retrieval, hyperspectral, high resolution, multi-sensor comparison, optically complex waters

2.1 Introduction

Satellite remote sensing provides a cost-effective and spatially extensive approach for monitoring inland WQPs. However, spatial resolution is one of the primary limitations in the application of satellite remote sensing to inland waters (Olmanson et al., 2015). Although ~36.52% of global lakes cover areas larger than 100 km² (Cao et al., 2019) and can therefore be monitored using moderate- or coarse-resolution sensors, rivers and narrow inland water bodies often require higher spatial resolution because their widths are commonly less than 10 m (IOCCG, 2018). Consequently, sensors with medium-to-high spatial resolution have been widely used for inland WQP retrieval (Ansari et al., 2025a).

In addition to spatial resolution, spectral resolution influences the capability of a sensor to retrieve inland WQPs. WQPs can generally be categorized into optically active and non-optically active WQPs (Mishra et al., 2017). Optically active WQPs, such as Chl-*a*, CDOM, and turbidity, directly influence light absorption and/or scattering within the water column. In contrast, non-optically active parameters, such as DOC and salinity, do not directly affect water optical properties and are typically retrieved indirectly through covariance with optically active constituents (Ansari et al., 2025c; Valerio et al., 2018). Accurate retrieval of optically active WQPs requires spectral bands sensitive to subtle variations in water optical properties. Multispectral sensors (e.g., S2) acquire data within a limited number of relatively broad spectral bands across the visible, Near-Infrared (NIR), and shortwave infrared regions. However, spectral resolutions on the order of ~5 nm from 380 to 737 nm can improve the detection of absorption and scattering features associated with optically active constituents (Dekker et al., 2018). Consequently, recent developments in satellite remote sensing have shifted from multispectral toward hyperspectral imaging systems (IOCCG, 2026).

Recent advances in hyperspectral satellite technology have created new opportunities for inland water monitoring. The Wyvern Dragonette-001 satellite (hereafter referred to as Wyvern), launched in April 2023, provides hyperspectral imagery with a spatial resolution of 5.3 m and 23 spectral

bands spanning the visible to NIR spectral range (503-800 nm) (Wyvern, 2026). This combination of high spatial and spectral resolution has potential for mapping and monitoring inland WQPs. However, similar to several current hyperspectral high-spatial-resolution satellite sensors (e.g., PRISMA) (IOCCG, 2026), Wyvern exhibits relatively low Signal-to-Noise Ratio (SNR) performance ($\text{SNR} < 400$), with estimated average nadir SNR values of ~ 60 , well below the recommended requirements for aquatic remote sensing (~ 400 in the visible, assuming NIR $\text{SNR} > \sim 600$ (Qi et al., 2017)).

However, the applicability of Wyvern data for inland WQP retrieval remains unexplored. A recent review (Ansari et al., 2025a) suggested that the spectral characteristics of Wyvern may be suitable for estimating non-algal particles and potentially Chl-*a*, whereas limitations may exist for CDOM retrieval because of the absence of blue spectral bands. Nevertheless, the capability of Wyvern imagery for inland WQP retrieval has not yet been assessed using coincident in-situ observations. In this study we therefore evaluated the capability of Wyvern imagery for retrieving Chl-*a*, CDOM, turbidity, SS, and DOC in the St. Lawrence River, Québec, Canada.

2.2 Study Areas and Data

2.2.1 Study Areas

The study was conducted at two sites along the St. Lawrence River (Figure 2.1): a $\sim 50 \text{ km}^2$ area near Contrecoeur (hereafter referred to as the Contrecoeur site) and a $\sim 41 \text{ km}^2$ area within Lake Saint Pierre (hereafter referred to as the Lake Saint Pierre site). The St. Lawrence River is frequently characterized by elevated turbidity levels associated with shoreline erosion (Vachon, 2023), tributary runoff (EcoLogic LLC., 2019), and anthropogenic activities (Merriman, 1997), including agriculture and urbanization.

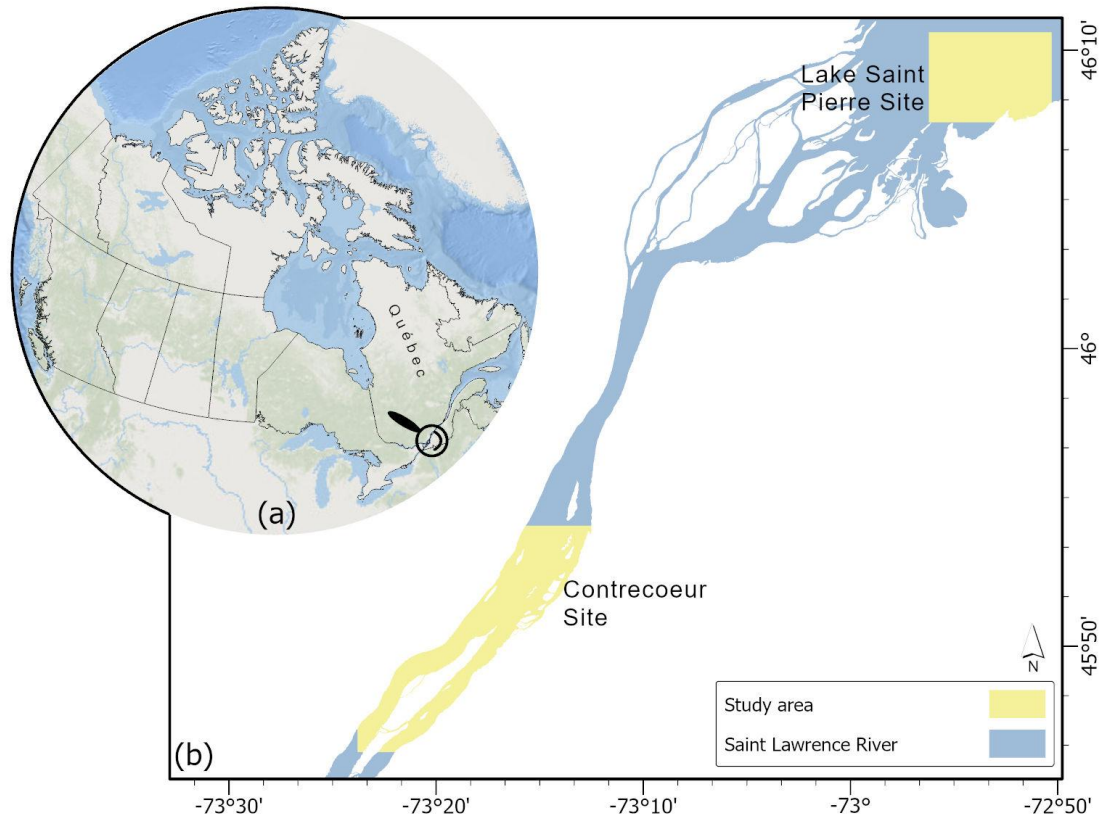


Figure 2.1. Study area. (a) Geographic overview of the study region within the St. Lawrence River, southern Quebec, Canada. (b) Spatial extent of the Contrecoeur and Lake Saint Pierre sites.

2.2.2 In-situ and Satellite Data

Two field campaigns were conducted on 02 August 2024 and 06 September 2024 at the Contrecoeur site, and another field campaign was conducted on 12 August 2025 at the Lake Saint Pierre site. All field campaigns coincided with overpasses of Wyvern and one of the S2 satellites (S2A, S2B, or S2C). In-situ measurements were collected between approximately 09:00 and 14:00 local time (Eastern Daylight Time). The acquisition dates, satellite overpass times, scene identifiers, and corresponding local times are summarized in Table 2.1. The temporal differences between Wyvern and S2 overpasses were less than 2 hours across all field campaigns. In all cases, in-situ observations were acquired within a few hours before or after the satellite overpasses.

Table 2.1. Summary of field campaign dates, corresponding satellite overpasses, acquisition times (Eastern Daylight Time), and associated scene identifiers for the study areas.

Date	Overpass	Local Time	Scene IDs
02 Aug 2024	Wyvern	10:34:58	wyvern dragonette-001 20240802T143458 6eac673b
	S2B	11:48:09	S2B MSIL1C 20240802T154809 N0511 R054 T18TXR 20240802T192935
06 Sep 2024	Wyvern	10:25:22	wyvern dragonette-001 20240906T142522 32257400
	S2A	11:48:11	S2A MSIL1C 20240906T154811 N0511 R054 T18TXR 20240906T210647
12 Aug 2025	Wyvern	09:53:46	wyvern dragonette-001 20250812T135346 70b1f00a
	S2C	11:49:51	S2C MSIL1C 20250812T154951 N0511 R054 T18TXS 20250812T192837

At the Lake Saint Pierre site, Chl-*a* ($\mu\text{g/L}$) and CDOM (ppb) measurements were collected along nine transects (Figure 2.2a and 2.2b). At the Contrecoeur site, turbidity (NTU), SS (mg/L), and DOC (mg/L) were measured (Figure 2.2c-e). Turbidity measurements were collected along 13 transects, and water samples were collected at or near the start and end points of selected transects. Duplicate measurements were also acquired at one randomly selected location during each of the two field campaigns. In total, 21 water samples were collected from 19 distinct locations within the Contrecoeur site, stored in a cooler with ice packs, and transported to a laboratory within a few hours for SS and DOC analyses. Measurements acquired on 06 September 2024 and 02 August 2024 were located north and south of the gray dashed line in Figure 2.2c-e, respectively.

In-situ measurements of turbidity, Chl-*a*, and CDOM were acquired using a YSI multiparameter sonde operated by WSP Canada Inc. (Montreal, QC, Canada). Unlike laboratory-derived CDOM absorption coefficients, the sonde reports CDOM as fluorescence intensity calibrated to relative concentration units (ppb) (Solinst, 2026). Laboratory analyses for SS and DOC were conducted by Bureau Veritas-Environmental Services Laboratory (Montreal, QC, Canada). All field observations were georeferenced using a Trimble R1 integrated GNSS system.

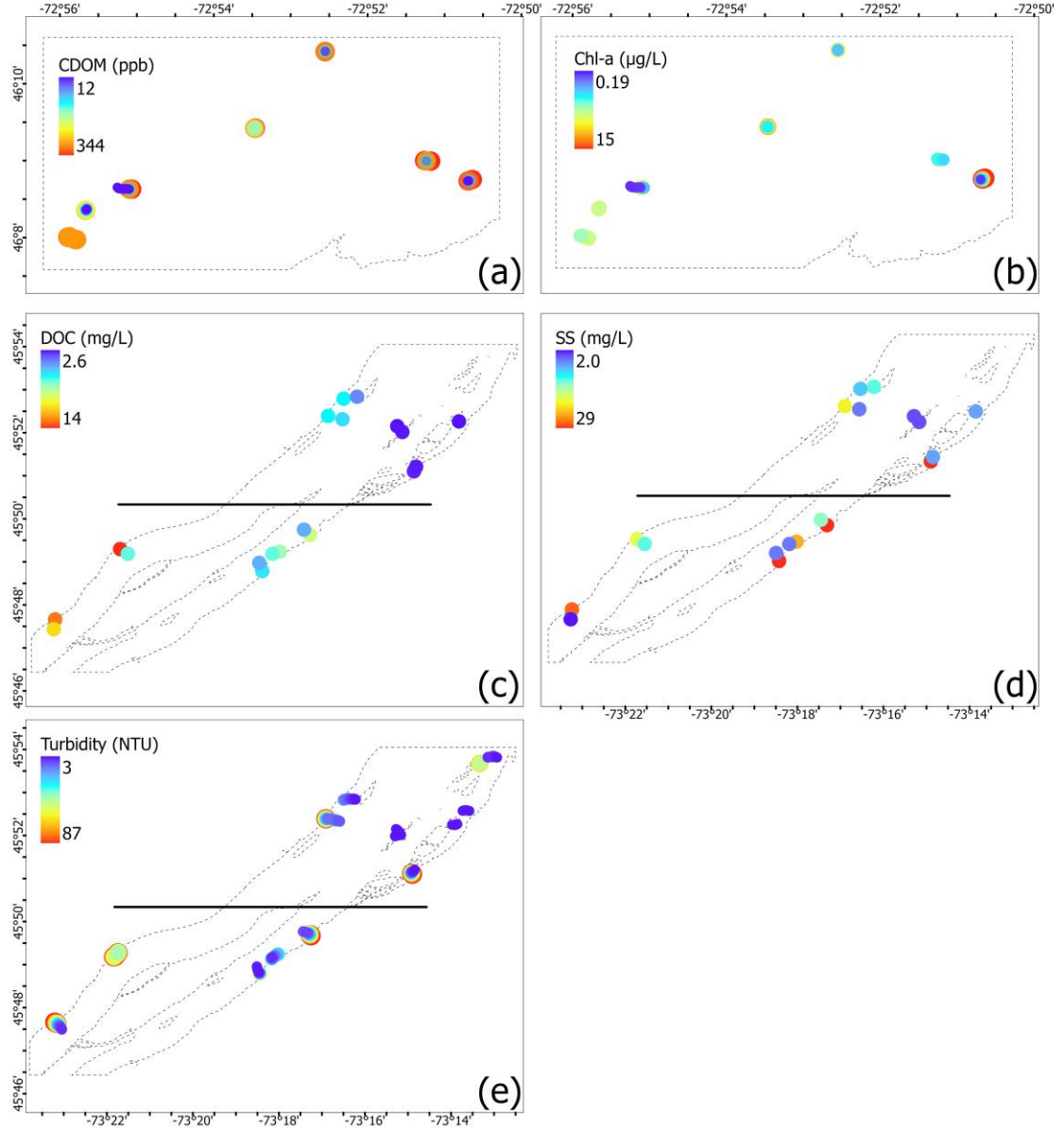


Figure 2.2. Spatial distribution of in-situ measurements for (a) CDOM, (b) Chl-*a*, (c) DOC, (d) SS, and (e) turbidity. Symbol color and size are scaled according to the magnitude of the measured WQP values. The black solid line in Contrecoeur site (c-e) separates the northern and southern sections associated with the 06 September 2024 and 02 August 2024 field campaigns, respectively.

2.3 Data Processing

This section describes the methodology adopted for WQP estimation using Wyvern and S2 imagery. First, atmospheric correction procedures applied to derive ρ_w are presented (Section 2.3.1). Atmospheric correction for both sensors was performed using the RAdCor algorithm (Castagna and Vanhellemont, 2025a) because it is one of the few atmospheric correction algorithms

available for both Wyvern and S2 imagery that accounts for adjacency effects, and a previous study conducted using S2 imagery in the same environment showed that reducing adjacency effects improved turbidity retrieval performance (Ansari et al., 2025d). In addition, for S2 inland water applications, a previous study demonstrated that RAdCor outperforms several commonly used atmospheric correction algorithms (Castagna and Vanhellemont, 2025a). Next, the spatial blocking strategy used for splitting the matchups into training and test datasets is described (Section 2.3.2). The development of the RF regression models, including hyperparameter optimization, is then explained (Section 2.3.3). RF regression was used for WQP retrieval because of its straightforward implementation and ability to estimate non-optically active WQPs (e.g., DOC) as an empirical model. Finally, the statistical metrics used for model evaluation are introduced (Section 2.3.4). Given the well-established performance and widespread use of S2 for inland WQP retrieval (Ansari et al., 2025a; Dekker and Hestir, 2012; IOCCG, 2018; Llodrà-Llabrés et al., 2023; Matthews, 2011), WQP retrievals derived from Wyvern imagery were compared against those obtained from S2 imagery to assess relative performance.

2.3.1 Atmospheric Correction

The RAdCor algorithm (Castagna and Vanhellemont, 2025a) implemented in Operational Land Imager lite (ACOLITE) (version 20260421.0) (Vanhellemont and Ruddick, 2018) was applied to derive ρ_w from S2 and Wyvern imagery. Default parameter settings were used, with two modifications. First, sun-glint correction (Harmel et al., 2018a) was enabled. Second, the negative- ρ_w filtering option was deactivated in ACOLITE to preserve negative reflectance values. These values may arise from sensor noise or overcorrection in the atmospheric correction process and were retained, as they may still retain physically meaningful spectral information. For inter-sensor pixel-wise ρ_w comparison, Wyvern and S2 10 m bands were resampled to a common 20 m grid for quantitative analyses. The comparison was restricted to seven spectrally corresponding band pairs based on the closest central wavelengths. In addition, for each WQP, the matchups were partitioned

into two concentration ranges using the midpoint between the minimum and maximum observed values ($\text{midpoint} \approx (\text{maximum} + \text{minimum})/2$) as a threshold. The average spectral signature of each concentration range was subsequently examined to assess the response of the sensors to changes in WQP concentrations.

2.3.2 Data Splitting

To minimize information leakage and reduce the effect of spatial autocorrelation between training and test datasets, a spatial blocking approach (Knudby and Richardson, 2023) was applied. A regular grid was first generated over the study area, and grid cells were randomly assigned to the test dataset until at least 20% of the total samples were included. The remaining cells were assigned to the training dataset (80%). The grid size was defined based on the semivariogram range of the in-situ measurements, ensuring reduced spatial dependence between neighboring blocks. Figure A1 illustrates the spatial distribution of in-situ measurements together with the allocation of samples into training and test datasets. Table 2.2 summarizes the semivariogram range, the number of training and testing samples, and the training:testing ratio for each WQP. The final training:testing ratios slightly deviated from the intended 80:20 split because the number of samples per grid cell varied. This is an inevitable consequence of the spatial blocking approach.

Table 2.2. Semivariogram ranges, the number of training and testing samples, and the training:testing ratio for each WQP.

WQPs	CDOM	Chl- <i>a</i>	DOC	SS	Turbidity
Semivariogram Range (m)	246	691	1,512	462	364
N of training Samples	521	463	14	16	9,874
N of testing Samples	175	233	7	5	2,520
Training:Testing Ratio (%)	75:25	67:33	67:33	76:24	80:20

2.3.3 Regression Model

The RF (Breiman, 2001) model was used for WQP estimation. RF is a tree-based ensemble machine learning method that combines multiple decision trees to improve predictive performance and reduce model variance. Model development and training were implemented in Python 3.10.17 using the *scikit-learn* library (Buitinck et al., 2013) version 1.7.1. To improve model performance

and reduce overfitting, hyperparameters were optimized using the *Optuna* library (Akiba et al., 2019) version 4.3.0 with 5-fold cross-validation. *Optuna* applies a Bayesian optimization strategy to efficiently explore the hyperparameter space. The following search ranges were defined: *n_estimators* (3-4,000), *max_depth* (2-400), *min_samples_split* (2-70), *min_samples_leaf* (1-70), and *max_features* (sqrt or log2). The optimization process was conducted over 200 trials. Model performance was evaluated using RMSE and the hyperparameter set corresponding to the lowest cross-validation RMSE was selected for the final model. The final models were then used to generate turbidity, DOC, and SS maps for 02 August 2024 at the Contrecoeur site, as well as Chl-*a* and CDOM maps for 12 August 2025 at the Lake Saint Pierre site.

2.3.4 Performance metrics

Three statistical metrics were employed as evaluation metrics: RMSE, ε (Morley et al., 2018), and β (Morley et al., 2018) (Equation 2.1-3):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_i - O_i)^2} \quad (2.1)$$

$$\beta = 100 \times \text{sign} \left[\text{median} \left(\log \frac{E_i}{O_i} \right) \right] \times \left(e^{|\text{median}(\log \frac{E_i}{O_i})|} - 1 \right) \quad (2.2)$$

$$\varepsilon = 100 \times \left(e^{\text{median} \left[\left| \log \left(\frac{E_i}{\bar{O}} \right) \right| \right]} - 1 \right) \quad (2.3)$$

N is the number of observations, and O , E , $\bar{O}$, and $\bar{E}$ are the observations, estimated values, the average of the observations, and the average of the estimated values, respectively.

2.4 Results

2.4.1 Inter-sensor comparison of ρ_w and spectral signatures

Figures 2.3 and 2.4 present the ρ_w product retrieved using RAdCor for Wyvern and S2 imagery acquired on 06 September 2024 over the Contrecoeur site. In general, the Wyvern bands exhibit higher ρ_w values relative to the corresponding S2 bands. In addition, diagonal striping noise patterns are evident in the Wyvern ρ_w product, particularly in wavelengths ≥ 700 nm (Figure 2.3p-w).

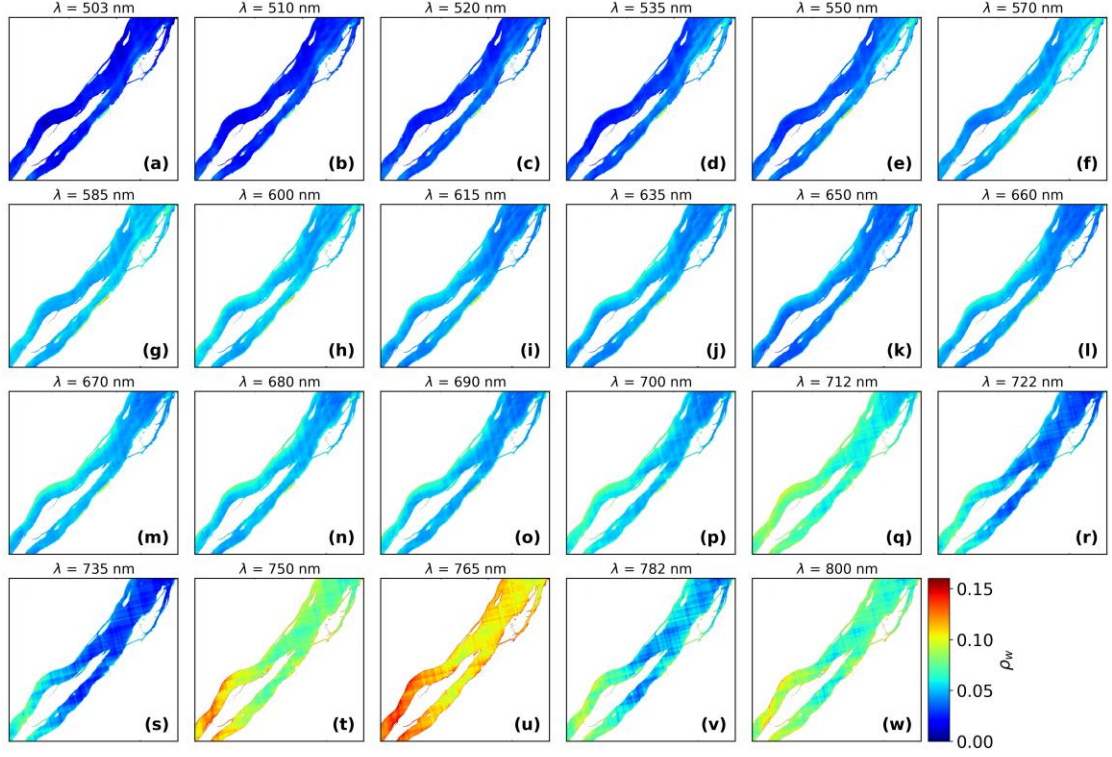


Figure 2.3. ρ_w product derived from Wyvern imagery over the Contrecoeur site on 06 September 2024. Panels (a-w) correspond to Wyvern spectral bands. The central wavelength (λ , nm) of each band is indicated at the top of each panel.

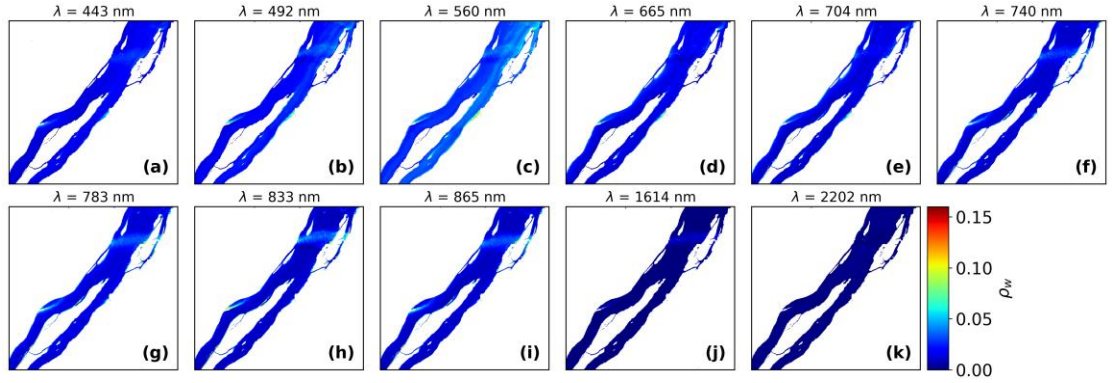


Figure 2.4. ρ_w product derived from S2 imagery over the Contrecoeur site on 06 September 2024. Panels (a-k) correspond to S2 spectral bands. The central wavelength (λ , nm) of each band is indicated at the top of each panel.

Figure 2.5 shows the pixel-wise comparison between S2 and Wyvern ρ_w for the seven spectrally corresponding band pairs. Overall discrepancies are evident from the RMSE values and data spread.

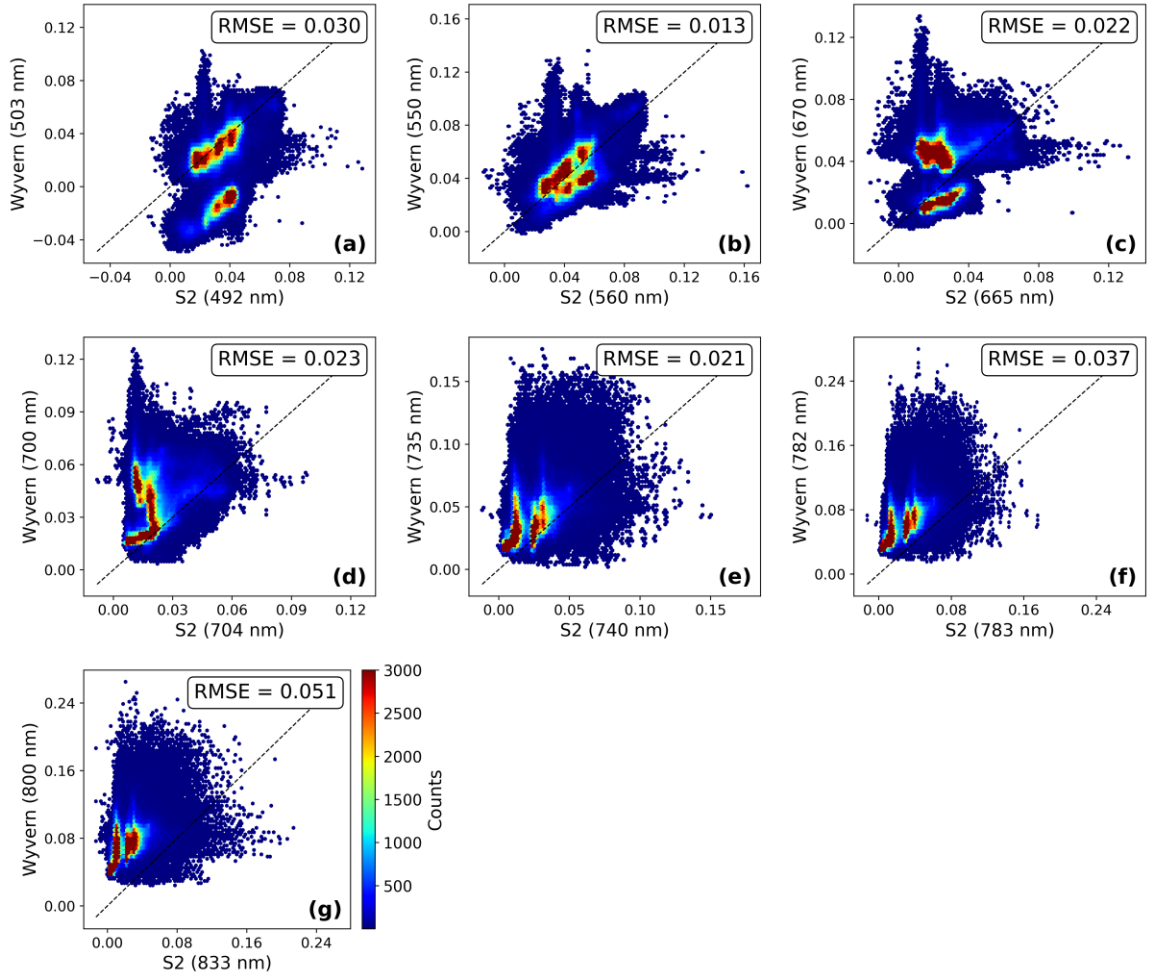


Figure 2.5. Pixel-wise comparison of S2 and Wyvern ρ_w for seven spectrally corresponding band pairs (panels a-g). Each panel shows hexbin density plots of matched observations aggregated from all three acquisition dates. The dashed line represents the 1:1 relationship, and RMSE values are reported for each band pair.

Figure 2.6 presents the mean spectral responses of the two concentration ranges defined for each WQP (see Section 2.3.1). Overall, no clear differences in the spectral signatures of CDOM and Chl-*a* are observed between the two sensors. For the remaining WQPs, a general increase in ρ_w with increasing concentration is evident across both sensors

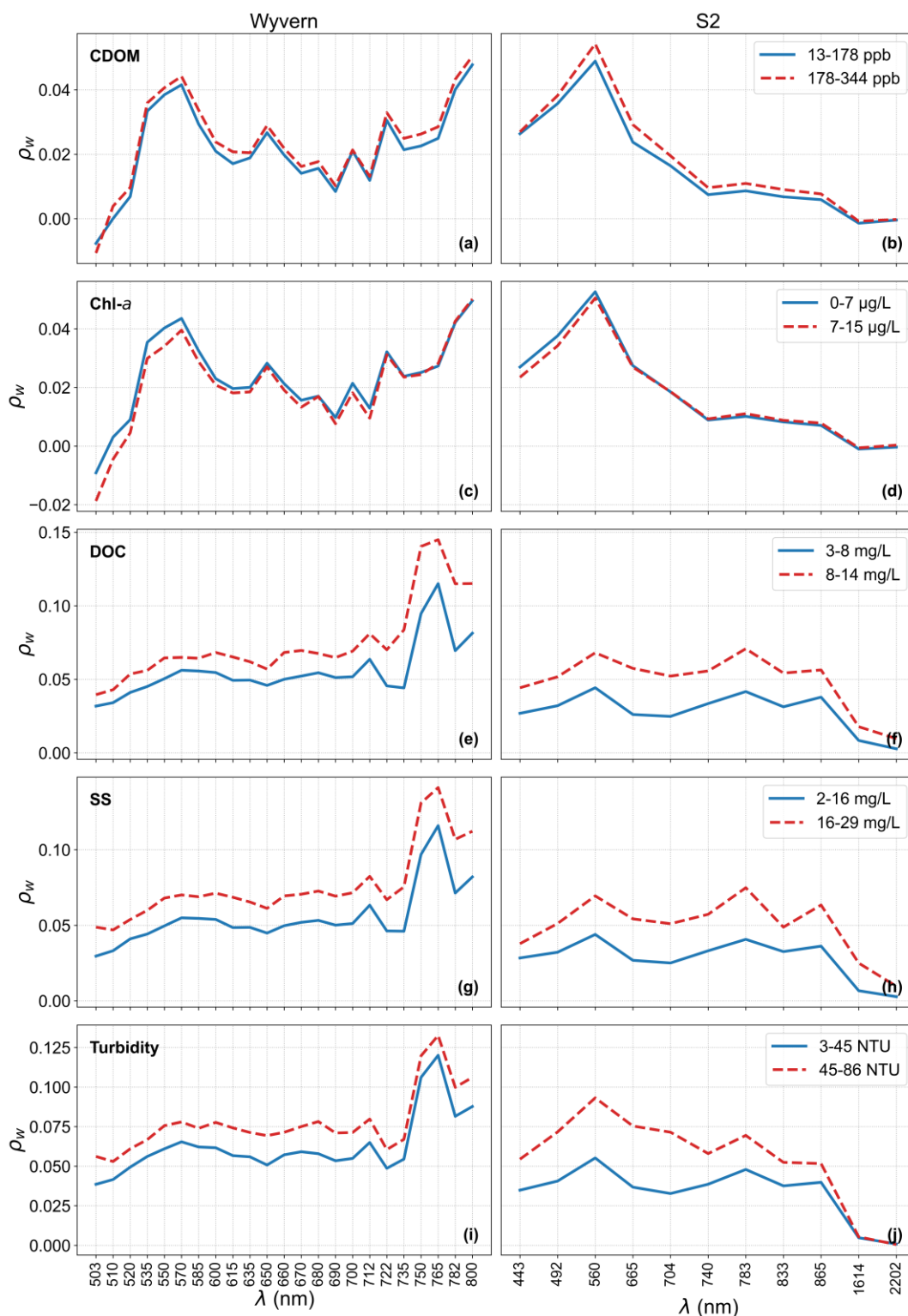


Figure 2.6. Sensor-specific spectral responses of WQPs derived from Wyvern and S2 imagery. For each parameter, observations were divided into two concentration ranges based on the midpoint between the minimum and maximum measured values, and the mean spectrum of each range was plotted. The left column corresponds to Wyvern and the right column to S2. CDOM, Chl-a, DOC, SS, and turbidity are shown in the first through fifth rows, respectively.

2.4.2 WQP Retrieval and Model Performance Assessment

The optimized RF hyperparameters for each WQP and sensor configuration are summarized in Table B1. Model performance metrics across sensors are reported in Table 2.3. S2 achieved lower RMSE values than Wyvern for CDOM and Chl-*a* retrieval, whereas Wyvern showed lower RMSE values than S2 for DOC. For SS and turbidity, similar performance is obtained for both Wyvern and S2.

Table 2.3. Performance metrics of the RF models for each WQP and sensor.

Metric	Sensor	CDOM	Chl-<i>a</i>	DOC	SS	Turbidity
RMSE	Wyvern	44.60	1.07	0.88	5.31	0.72
	S2	24.41	0.28	3.10	5.28	0.88
β (%)	Wyvern	-1.42	125.35	-0.60	-32.81	0.07
	S2	-0.03	1.92	38.90	-19.90	0.01
ϵ (%)	Wyvern	1.64	125.35	8.86	32.81	1.60
	S2	0.44	3.98	70.01	29.77	2.10

Figure 2.7 presents scatter plots of observed versus predicted values for each WQP and sensor. For CDOM, closer alignment with the 1:1 line is observed at higher concentration ranges for both sensors (Figures 2.7a and 2.7f) compared with lower concentrations. For Chl-*a*, Wyvern shows weaker agreement at low concentrations (0-1 $\mu\text{g/L}$), whereas S2 shows closer alignment in this range. For DOC and SS, interpretation is limited by the small number of observations. For turbidity, values from both sensors are closely distributed around the 1:1 line (Figures 2.7e and 2.7j).

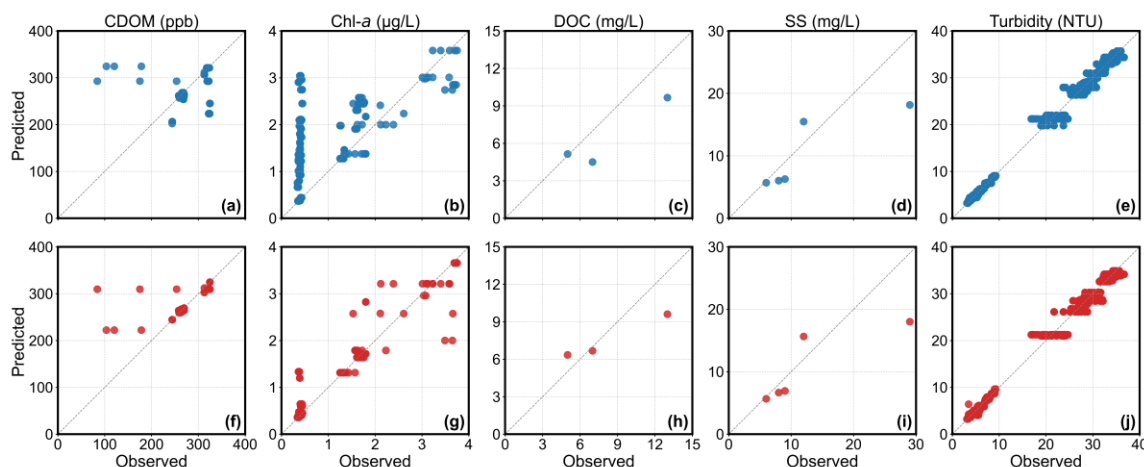


Figure 2.7. Scatterplots of observed versus estimated WQP values derived from the RF model for Wyvern (top row; panels (a-e), blue dots) and S2 (bottom row; panels (f-j), red dots), with CDOM, Chl-*a*, DOC, SS, and turbidity shown in the first to fifth columns, respectively. The dashed diagonal line represents the 1:1 line.

2.4.3 Spatial distribution of retrieved WQPs

Figure 2.8 presents the spatial distributions of CDOM and Chl-*a* over the Lake Saint Pierre site on 12 August 2025 for both Wyvern and S2 imagery. For CDOM (Figures 2.8a and 2.8b), partially consistent spatial patterns are observed between the two sensors, although differences in magnitude are evident. For example, a high-CDOM plume in the northwestern portion of the Lake Saint Pierre site appears in both maps, but with different magnitudes (~280 ppb for Wyvern versus ~250 ppb for S2). For Chl-*a* (Figures 2.8c and 2.8d), spatial patterns are more divergent. In particular, Wyvern estimates are largely concentrated within ~2-3 µg/L across most pixels.

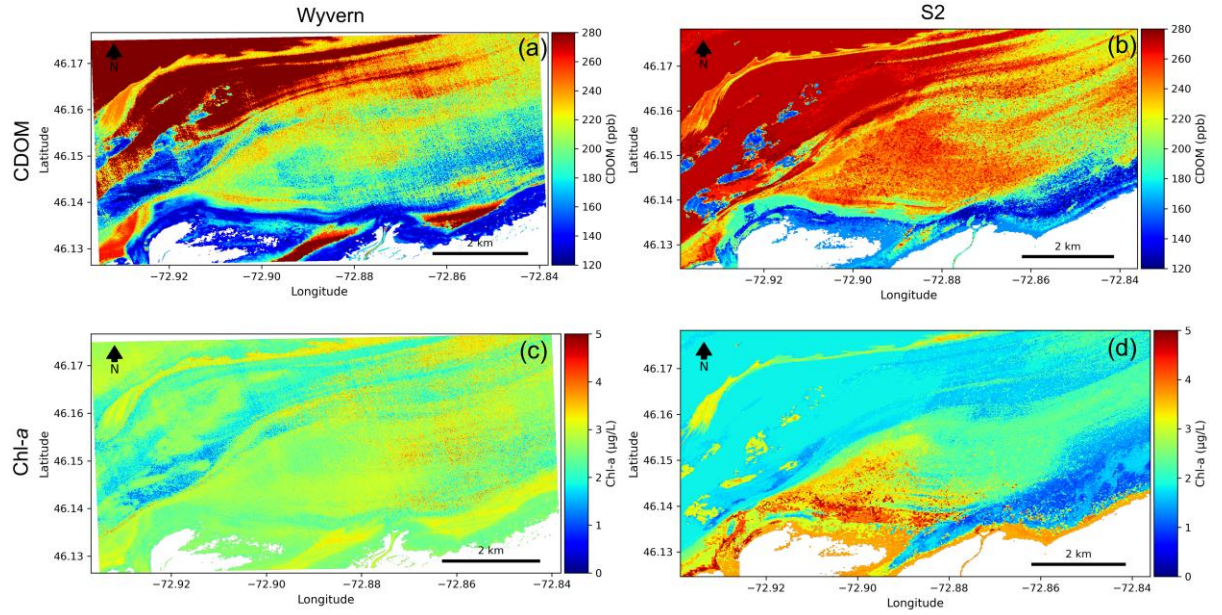


Figure 2.8. Spatial distribution of CDOM and Chl-a derived from Wyvern and S2 imagery over the Lake Saint Pierre site on 12 August 2025. The left column corresponds to Wyvern and the right column to S2. CDOM and Chl-a are shown in the first and second rows, respectively.

Figure 2.9 presents the spatial distributions of DOC, SS, and turbidity over the Contrecoeur site for 02 August 2024. Higher concentrations of DOC, SS, and turbidity are generally observed along the northern bank of the river compared with the southern bank in both dates. This spatial pattern is consistent with field observations (Figure 2.2 c-e).

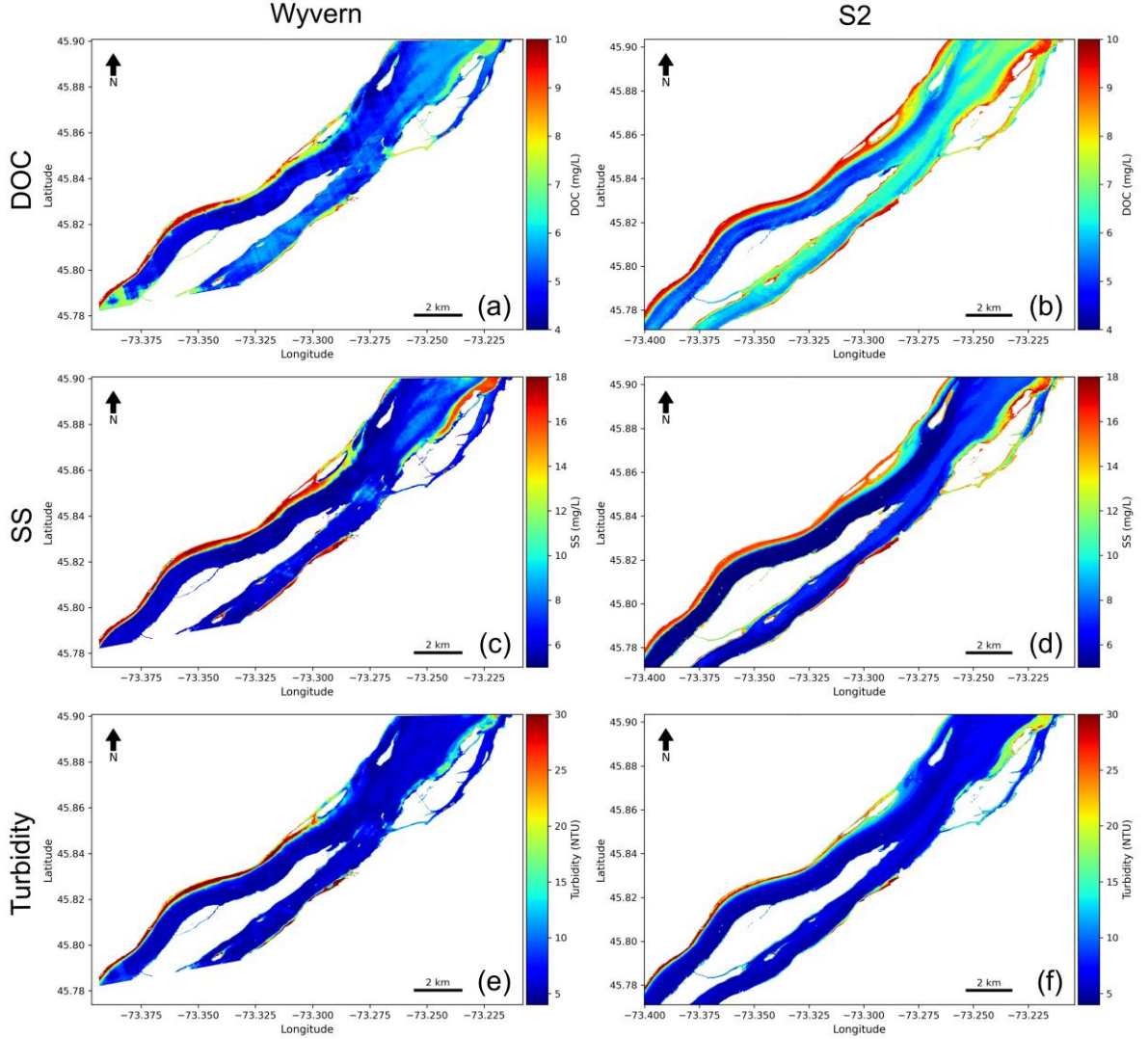


Figure 2.9. Spatial distribution of DOC, SS, and turbidity derived from Wyvern and S2 imagery over the Contrecoeur site on 02 August 2024. The left column corresponds to Wyvern and the right column to S2. DOC, SS, and turbidity are shown in the first, second, and third rows, respectively.

2.5 Discussion

2.5.1 ρ_w

Diagonal striping noise patterns are evident in the Wyvern-derived ρ_w product, particularly for $\lambda \geq 700$ nm (Figures 2.3p-w). These artifacts are subsequently propagated to the Wyvern-derived WQP maps and are especially visible in the DOC maps (Figures 2.9a). In the NIR wavelength range, ρ_w values are generally expected to be negligible due to strong absorption by water, except in high-turbidity conditions (IOCCG, 2010; Mishra et al., 2017). This is consistent with the S2-derived ρ_w

products for the NIR bands (Figures 2.4e-g), while elevated ρ_w values are observed in the Wyvern imagery at $\lambda \geq 712$ nm (Figure 2.3q-w), compared both to shorter-wavelength bands and to the corresponding S2 bands. Several factors can produce elevated NIR reflectance values in satellite imagery, including high turbidity, strong adjacency effects, errors in aerosol contribution estimation during atmospheric correction, or sensor-specific effects such as calibration bias or stray light. Specifically the elevated ρ_w values observed near 765 nm may also be associated with atmospheric correction uncertainties related to the O₂ absorption feature around 762 nm (Greenblatt et al., 1990; IOCCG, 2026). However, the comparison to concurrent S2 observations that do not show similarly elevated NIR ρ_w values strongly suggests that the cause of elevated ρ_w values in the Wyvern imagery is sensor-related and not caused by environmental or processing-related issues. Moreover, the spectral behavior between the two sensors is not closely aligned across the seven corresponding spectral bands, as indicated by RMSE values exceeding 0.01 for all band pairs (Figure 2.5). Specifically, assuming the S2-derived ρ_w as a baseline reference, Wyvern generally overestimates ρ_w at 670, 700, 735, 783, and 800 nm (Figure 2.5c-g), underestimates it at 503 nm (Figure 2.5a), while the best agreement between the two sensors is observed at 550 nm (Figure 2.5b).

2.5.2 CDOM and Chl-*a*

As shown in Figures 2.6a-d, the spectral signatures of CDOM and Chl-*a* under low and high concentration conditions are very similar, suggesting that neither sensor is particularly sensitive to the range of concentrations observed in this study. This is further supported by the performance metrics (Table 2.3), which indicate generally weak retrieval performance for both sensors, although S2 consistently outperformed Wyvern. For CDOM, S2 yields a lower RMSE (24.41 ppb), corresponding to 7.4% of the total CDOM range, compared to Wyvern (44.60 ppb; 13.5%). In addition, the ϵ metric reveals larger overall errors for Wyvern, ~3.7 times higher than those of S2. The scatterplots (Figures 2.7a and 2.7f) further support the superior performance of S2. For S2,

agreement with the 1:1 line improves at higher concentrations (>300 ppb), is moderate at intermediate levels (200-300 ppb), and weakens at lower concentrations (<200 ppb). In contrast, Wyvern shows acceptable agreement only within a narrow range (~ 280 -310 ppb). A similar pattern is observed for Chl-*a*, where S2 has superior performance relative to Wyvern. The RMSE for S2 is $0.28 \mu\text{g/L}$ (1.9% of the observed range), whereas Wyvern exhibits a higher RMSE of $1.07 \mu\text{g/L}$ (7.4% of the range). The β metric indicates a slight overestimation for S2 ($\beta = 1.92\%$), while Wyvern shows a pronounced positive bias ($\beta = 125.35\%$), ~ 65 times larger than that of S2. Consistently, ϵ values for Wyvern are ~ 31 times greater than those of S2. Scatterplots (Figures 2.7b and 2.7g) confirm these trends: S2 shows reasonable agreement with the 1:1 line at low ($\sim 0.3 \mu\text{g/L}$) and intermediate ($\sim 1.8 \mu\text{g/L}$) concentrations, while dispersion increases elsewhere. In contrast, Wyvern overestimates low concentrations ($<1 \mu\text{g/L}$), producing a near-vertical pattern around predicted values of $\sim 0.3 \mu\text{g/L}$, with only partial improvement at higher concentrations.

Both sensors capture similar spatial gradients in CDOM across most of the study area (Figures 2.8a and 2.8b), however, noticeable differences in magnitude are evident. In particular, S2-derived estimates show closer agreement with in-situ observations (Figure 2.2a). Wyvern maps also exhibit a slightly noisier spatial pattern, which may be attributed to lower retrieval stability, higher noise levels in Wyvern imagery compared to S2, and the higher spatial resolution of Wyvern compared to S2. For Chl-*a* (Figures 2.8c and 2.8d), Wyvern-derived estimates are largely concentrated within the ~ 2 - $3 \mu\text{g/L}$ range across most pixels. Given the superior performance of S2 (Table 2.3) and agreement with in-situ observations (Figure 2.2b), the S2-derived map can be considered more reliable. Under this assumption, a consistent spatial feature between both products is the elongated Chl-*a* plume ($\sim 3 \mu\text{g/L}$; yellow in the maps) located in the northwestern portion of Lake Saint Pierre, extending toward the north-central region. Beyond this feature, the spatial distributions differ substantially between the two sensors. Overall, neither sensor demonstrates strong reliability for CDOM and Chl-*a* retrieval, though S2 consistently outperforms Wyvern. Previous studies (Ansari et al., 2025a; Mishra et al., 2017) have shown that CDOM retrieval is primarily driven by blue

wavelengths (e.g., 380, 412, 425, and 440 nm), while Chl-*a* retrieval relies on both blue (~438 nm) and red (~667 nm) wavelengths. Therefore, newer Dragonette sensor versions (e.g., Dragonette-002 and Dragonette-003), which include blue spectral bands, may perform better for CDOM and Chl-*a* retrieval.

2.5.3 DOC and SS

Considering the limited number of DOC and SS matchups available in this study, the results should be interpreted with caution and do not allow for a robust assessment of sensor performance. For DOC, Wyvern yielded lower RMSE, ϵ , and $|\beta|$ values than S2 (Table 2.3), and the corresponding scatterplot showed closer agreement with the 1:1 line (Figures 2.7c and 2.7h). In addition, both sensors captured elevated DOC concentrations along the northern bank of the river (Figures 2.9a and 2.9b), consistent with the in-situ observations (Figure 2.2c). For SS, the performance of the two sensors was generally similar, with S2 showing slightly better performance. RMSE values were comparable between the sensors, while relatively high ϵ and $|\beta|$ values were observed for both (Table 2.3). The β metric indicated a tendency toward underestimation, which is also evident in the scatterplots (Figures 2.7d and 2.7i). Spatially (Figures 2.9c and 2.9d), both sensors produced similar SS distributions, including elevated concentrations along the northern and southern banks of the river, in agreement with field observations (Figure 2.2d). Overall, Wyvern showed some indication of improved DOC retrieval, whereas both sensors exhibited comparable performance for SS. However, due to the limited availability of in situ DOC and SS measurements, these findings should be considered preliminary. Further validation using additional field observations is required to better assess the relative performance of the two sensors for DOC and SS mapping.

2.5.4 Turbidity

As shown in Figures 2.6i and 2.6j, increasing turbidity concentrations are associated with higher ρ_w values for both sensors. This pattern is consistent with previous studies (IOCCG, 2010; Mishra et al., 2017; Ruddick et al., 2006), which have shown that increasing turbidity enhances particle

scattering and consequently increases ρ_w , particularly in the red and NIR spectral regions. From a performance perspective (Table 2.3), both sensors demonstrated good turbidity retrieval capabilities. Wyvern achieved slightly lower RMSE and ϵ values than S2, whereas S2 exhibited a marginally lower $|\beta|$ value. These results are consistent with the scatterplots (Figures 7e and 7j), where predictions from both sensors generally cluster around the 1:1 line. In addition, the turbidity maps generated for 02 August 2024 (Figures 2.9e and 2.9f) show similar spatial patterns and are broadly consistent with the in-situ observations (Figure 2.2e). Overall, both sensors performed well for turbidity retrieval, with Wyvern showing slightly better performance than S2.

2.6 Conclusion

This study evaluated the performance of Wyvern Dragonette-001 and S2 MSI imagery for the retrieval of CDOM, Chl-*a*, DOC, SS, and turbidity at two sites in the St. Lawrence River. Atmospheric correction was performed using the RAdCor algorithm, and an RF model was applied to establish relationships between ρ_w and in-situ measurements. The analysis revealed elevated ρ_w values in Wyvern imagery at $\lambda \geq 700$ nm and diagonal striping noise, particularly at $\lambda \geq 712$ nm. Comparison with concurrent S2 observations, which did not exhibit similar behavior, suggests that these artifacts are sensor-related. In addition, notable discrepancies were observed between S2 and Wyvern ρ_w across seven spectrally corresponding bands. Among the investigated WQPs, turbidity showed the most robust retrieval performance for both sensors, with Wyvern performing slightly better than S2. For DOC, Wyvern showed some indication of improved performance, while both sensors produced comparable results for SS. However, the limited number of in-situ DOC and SS measurements prevents robust conclusions regarding their relative performance. In contrast, neither sensor achieved strong retrieval performance for CDOM or Chl-*a*, although S2 consistently outperformed Wyvern for both parameters. Overall, the results highlight both the potential and current limitations of Wyvern Dragonette-001 for inland water quality monitoring. Newer

Dragonette missions (Dragonette-002 and Dragonette-003), which include blue spectral bands, may improve the retrieval of optically complex WQPs, particularly CDOM and Chl-*a*.

2.7 Appendix A

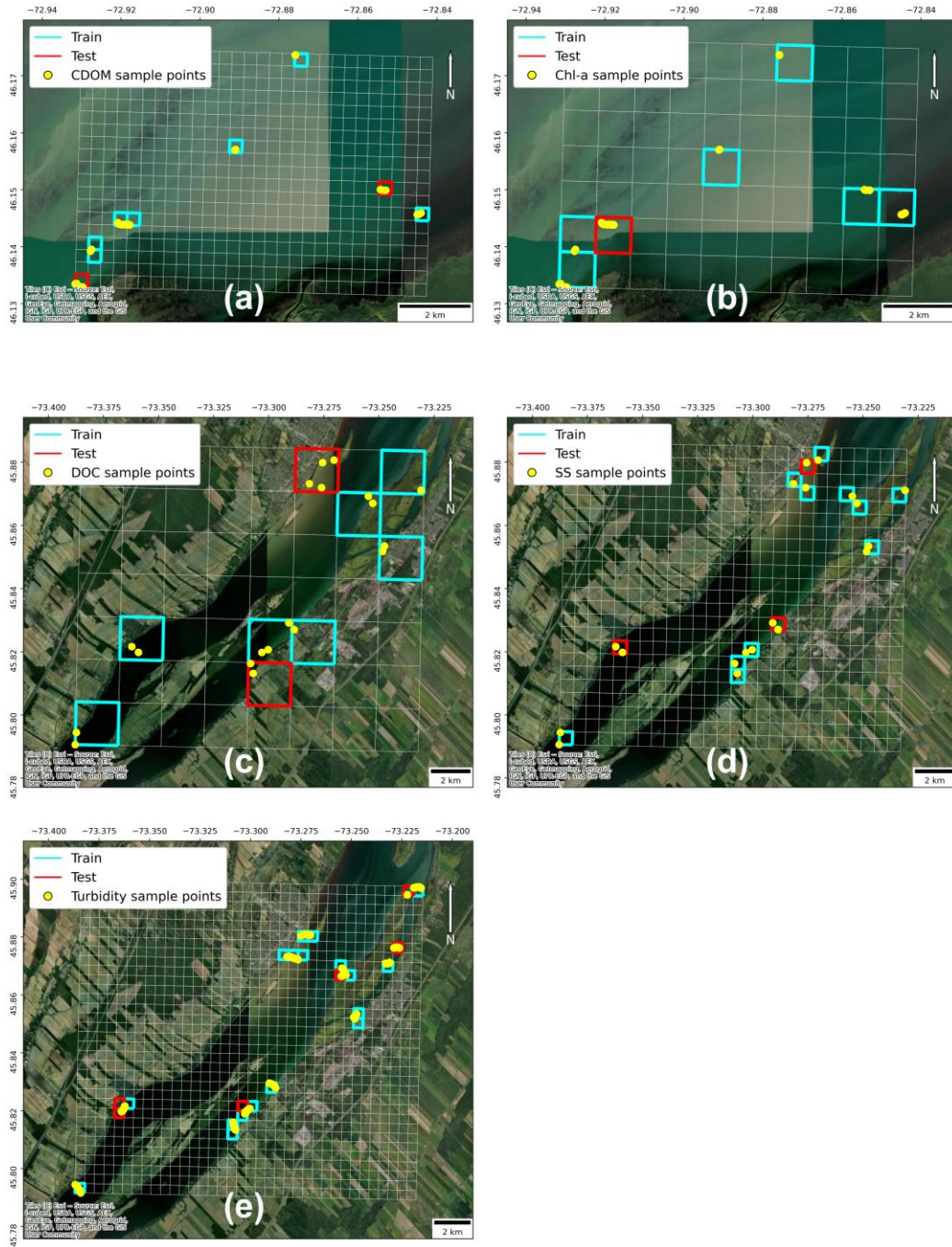


Figure A1. Spatial distribution of in-situ measurements and the corresponding data splitting scheme for (a) CDOM, (b) Chl-*a*, (c) DOC, (d) SS, and (e) turbidity. Yellow points indicate field sampling

locations, while grid cells with cyan and red borders represent training and test samples, respectively.

2.8 Appendix B

Table B1. Optimized RF hyperparameters for each WQP and sensor.

Metric	Sensor	CDOM	Chl-<i>a</i>	DOC	SS	Turbidity
n_estimators	Wyvern	2,616	1,935	1,221	3,023	2,536
	S2	1,832	3,655	2,527	3,343	816
max_depth	Wyvern	272	40	38	22	181
	S2	141	353	14	327	280
min_samples_split	Wyvern	2	5	4	9	3
	S2	8	2	5	8	9
min_samples_leaf	Wyvern	1	1	2	1	1
	S2	1	1	3	1	1
max_features	Wyvern	log2	log2	log2	log2	sqrt
	S2	log2	log2	sqrt	sqrt	log2

Chapter 3: Assessing the Impact of T-Mart Adjacency Effect Correction on Turbidity Retrieval from Landsat 8/9 and Sentinel-2 Imagery (Case Study: St. Lawrence River, Canada)

Mohsen Ansari¹, Yulun Wu¹, and Anders Knudby¹

¹Department of Geography, Environment and Geomatics, University of Ottawa, Ottawa, Canada

Abstract

In inland waters, Atmospheric Correction (AC), including Adjacency Effect (AE) correction, is a major challenge for water quality retrieval using optical satellite data. This study evaluated three image pre-processing options for turbidity retrieval in the St. Lawrence River using Sentinel-2 (S2) and Landsat 8/9 (L8/9) imagery with the Light Gradient Boosting Machine (LightGBM) model: (1) No pre-processing, i.e. use of Top-of-Atmosphere (TOA) reflectance, (2) AC pre-processing, obtaining water-leaving reflectance (R_w) from AC for the Operational Land Imager lite (ACOLITE)'s Dark Spectrum Fitting (DSF) technique, and (3) AE pre-processing, correcting for the AE using T-Mart before obtaining R_w from DSF. Results demonstrated that AE pre-processing outperformed the other two options. For L8/9, AE pre-processing reduced the Root Mean Square Error (RMSE) and improved the median symmetric accuracy (ϵ) by 48.8% and 19.0%, respectively, compared with AC pre-processing, and by 48.5% and 50.7%, respectively, compared with No pre-processing. For S2, AE pre-processing performed better than AC pre-processing and also outperformed No pre-processing, reducing RMSE by 28.4% and ϵ by 50.8%. However, No pre-processing yielded the lowest absolute symmetric signed percentage bias ($|\beta|$) among all pre-processing options. Analysis indicated that AE pre-processing yielded superior performance within 0-300 m from shore than other options, where the AE influence is strongest. Turbidity maps generated using AE pre-processing were smoother and less noisy compared to the other pre-processing options, particularly in cloud-adjacent regions. Overall, our findings suggest that incorporating AE correction through T-Mart improves the performance of the LightGBM model for turbidity retrieval from both L8/9 and S2 imagery in the St. Lawrence River, compared to the alternative pre-processing options.

Keywords: Turbidity; Sentinel-2; Landsat; Atmospheric correction; Adjacency effect correction; Machine learning; Aquatic remote sensing.

DOI: 10.3390/rs18010127

3.1 Introduction

Monitoring inland water quality is essential to water resource management (Ansari et al., 2025a), and turbidity is an important water quality indicator as it affects light penetration, photosynthetic activity, primary productivity, and aquatic life (Cloern, 1987), while also reflecting the concentration of suspended materials (Liu et al., 2019). Traditional field-based measurements can provide accurate turbidity data but are spatially limited and time-consuming to acquire, especially in large or dynamic inland water bodies. Satellite remote sensing offers a complementary approach, enabling spatially continuous and temporally consistent turbidity estimation across extensive areas. Satellite sensors such as the Landsat 8/9 (L8/9) Operational Land Imager (OLI) and the Sentinel-2 (S2) MultiSpectral Imager (MSI) have been used to map inland water quality parameters, including turbidity, due to their moderate spatial, temporal, and spectral resolution (Ansari et al., 2025a). However, the radiance measured at the Top-of-Atmosphere (TOA) is influenced by atmospheric attenuation and surface reflection at the air-water interface (Mishra et al., 2017), therefore, a robust Atmospheric Correction (AC) method is needed to remove the atmospheric and water surface signals and extract R_w , the fraction of sunlight that penetrates the water column and is reflected upward out of the water surface and serves as the basis for estimating water quality parameters from satellite imagery. While such AC methods for open-ocean waters are well established (IOCCG, 2010), AC for inland water is complicated for several reasons (Mishra et al., 2017):

- Atmospheric heterogeneity, caused by urban aerosols and pollution, invalidates the atmospheric conditions assumed in radiative transfer modeling;
- Inland waters typically exhibit moderate-to-high turbidity, which causes non-negligible reflectance in the near-infrared (NIR) region. This invalidates assumptions used in AC methods that rely on NIR bands to derive the aerosol type (i.e., the composition and size distribution of atmospheric particles) and optical thickness (a measure of the extent to which aerosols and molecules attenuate light as it passes through the atmosphere),

potentially leading to the overcorrection of R_w in the visible spectrum (Hu et al., 2000; Ruddick et al., 2000); and

- The Adjacency Effect (AE), the scattering of light from nearby bright surfaces (e.g., land, clouds) into the sensor's field of view, thereby contaminating the reflectance of water pixels, is strong for most inland waters due to their close proximity to land (Tanré et al., 1987).

To address the first two challenges, several AC methods have been developed for inland waters, each with its own advantages and limitations (Ansari et al., 2025a; Pahlevan et al., 2021). Among them, AC for the Operational Land Imager lite (ACOLITE)'s DSF technique (Vanhellemont, 2019; Vanhellemont and Ruddick, 2018) is widely used for L8/9 and S2 imagery (Ansari et al., 2025a). This AC method first identifies the darkest pixels across all spectral bands within the image, assuming these pixels have minimal water-leaving radiance. It then fits an aerosol model by adjusting the aerosol type and optical thickness so that the simulated TOA reflectance matches the observed values over these dark targets. Once the optimal aerosol parameters are determined, they are applied across the image to correct for atmospheric effects and derive R_w . By relying on image-derived dark spectra instead of predefined dark bands (e.g., NIR), DSF minimizes the risk of overcorrection in turbid or complex inland waters. However, no single AC method performs consistently across all inland waters and sensors (Ansari et al., 2025a). Consequently, in some cases, researchers have used raw TOA reflectance directly in empirical models (e.g., machine learning models) to estimate water quality parameters such as turbidity (Ouma, 2020; Q. Wang et al., 2024), cyanobacteria (Begliomini et al., 2023; Ge et al., 2024a) and chlorophyll-*a* (Hieronymi et al., 2016).

Correcting for the AE requires modeling the atmospheric point spread function, which quantifies how light scattered from different surface areas contributes to the signal received by the sensor (Wu et al., 2024). Several methods have been proposed to account for this effect. For instance, the Remote sensing Adjacency Correction (RAdCor) method (Castagna and Vanhellemont, 2025b)

constructs the point spread function through simulations of TOA radiance from reflective and non-reflective neighboring pixels. The Genetic Algorithm for AC (GAAC) (Pan and Bélanger, 2025) retrieves AE parameters, such as aerosol optical depth and type, by solving a set of radiative transfer equations using a genetic optimization algorithm. The T-Mart approach employs 3D Monte Carlo simulations (Wu et al., 2024) to explicitly model AE under various atmospheric and topographic conditions within a coupled ocean–land–atmosphere system.

The present study investigates whether applying AC using DSF and AE correction using T-Mart can improve turbidity retrieval from S2 and L8/9 imagery for a portion of the St. Lawrence River, using a machine learning model. To this end, turbidity was estimated with three pre-processing options:

1. using raw TOA reflectance (No pre-processing option);
2. using R_w derived from DSF (AC pre-processing option); and
3. using R_w derived from DSF after pre-processing with T-Mart (AE pre-processing option).

Although several alternative AE correction approaches exist, such as RAdCor and GAAC, each is developed under specific assumptions and tailored to particular environmental conditions. The present study concentrates on assessing the performance of T-Mart within a portion of the St. Lawrence River, rather than providing a comprehensive comparison of all available AE correction methods. A larger-scale intercomparison involving T-Mart, RAdCor, and GAAC is currently underway, supported by a community-contributed dataset from the aquatic remote sensing network (<https://odnature.naturalsciences.be/radcor>; personal communication with A. Castagna, Q. Vanhellemont, and S. Bélanger).

It should be noted that, although river depth, nearshore shallow-water conditions, and bottom sediment reflectance can influence observed turbidity (Kwon et al., 2023; Wu et al., 2025), particularly in optically shallow and nearshore environments, these factors were not explicitly incorporated into the modeling framework of the present study. The analysis focuses on evaluating the impact of atmospheric and adjacency effect corrections on turbidity retrieval, assuming that in

situ turbidity measurements implicitly capture the combined effects of water column properties within the sampled conditions. A more explicit treatment of bathymetry and bottom reflectance would require additional datasets and radiative transfer modeling, which is beyond the scope of this study.

Machine learning models have demonstrated capabilities in capturing nonlinear relationships between spectral data and water quality parameters (Adjovu et al., 2023b, 2023a; Ansari et al., 2025c; Ansari and Akhoondzadeh, 2021; Mohsen Ansari and Akhoondzadeh, 2020; R.M. Khan et al., 2024; Mamun et al., 2024; Richardson et al., 2025; Q. Wang et al., 2024; Yadong Zhou et al., 2024), including turbidity . This study employed the LightGBM model to estimate turbidity from L8/9 and S2 imagery under the three pre-processing options described above.

3.2 Study area and data

3.2.1 Study area

The study area is a ~100 km² portion of the St. Lawrence River (Figure 3.1a), north of Montréal (Figure 3.1b), Québec, Canada, where there has been a proposal for the expansion of the Port of Montreal's Contrecoeur Terminal. The St. Lawrence River is often turbid due to shoreline erosion (Vachon, 2023), runoff from tributaries (EcoLogic LLC., 2019), and human activity (Merriman, 1997) such as agriculture, shoreline development, and urbanization.

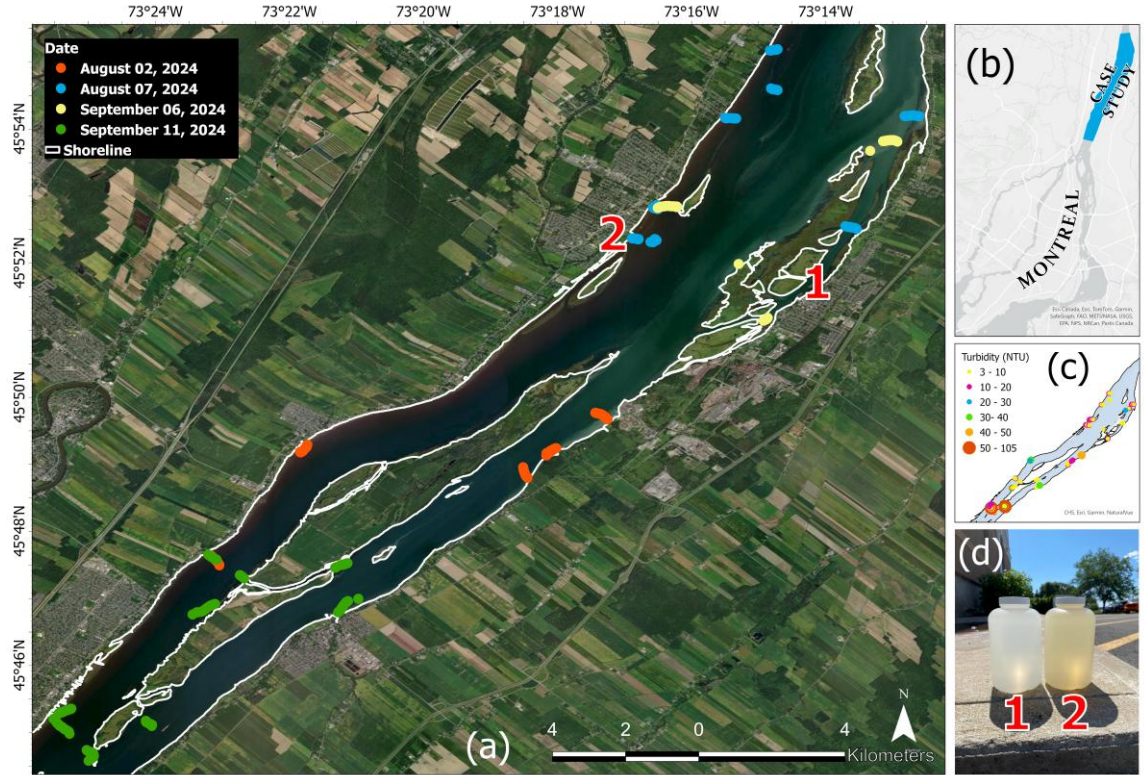


Figure 3.1. Study area: the St. Lawrence River, Montréal, Québec, Canada. (a) Sampling locations where turbidity measurements were collected during four field campaigns; each campaign is shown in a different color. (b) Location of the case study area. (c) Map showing turbidity values at each measurement location. (d) Example water samples collected at Sites 1 and 2 (in panel a) on August 07, 2024.

3.2.2 In-situ and satellite data

Four field campaigns were conducted during August and September 2024, resulting in 16,074 in-situ turbidity measurements collected along 30 transects (Figure 3.1a). These campaigns coincided with six satellite overpasses from S2A, S2B, L8, and L9 (Table 3.1 and Figure 3.1). Turbidity was measured using a YSI multiparameter sonde operated by WSP Canada Inc., and each observation was georeferenced with a Trimble R1 integrated GNSS system. The recorded turbidity values ranged from 1.89 NTU to 104.7 NTU, reflecting the spatial variability across the study area.

Table 3.1. Satellite overpasses with coincident in-situ turbidity measurements in 2024.

Date	Overpass	Scene IDs
August 02, 2024	L8, S2B	L8: LC08_L1TP_014028_20240802_20240808_02_T1 S2B: S2B_MSIL1C_20240802T154809_N0511_R054_T18TXR_20240802T192935
August 07, 2024	S2A	S2A_MSIL1C_20240807T154941_N0511_R054_T18TXR_20240807T211007
September 06, 2024	S2A	S2A_MSIL1C_20240906T154811_N0511_R054_T18TXR_20240906T210647
September 11, 2024	L9, S2B	L9: LC09_L1TP_014028_20240911_20240911_02_T1 S2B: S2B_MSIL1C_20240911T154809_N0511_R054_T18TXR_20240911T192755

Sampling locations were selected to capture areas where variations in water color and turbidity were visually apparent. In addition, for safety reasons, data collection avoided the main navigation channel used by large commercial vessels. As a result, most measurements were acquired along the river margins rather than the central channel, where turbidity showed relatively limited variability. Notably, during the field campaigns, high turbidity was observed near the northern shore at Site 2 (Figure 3.1a), while the southern shore at Site 1 (Figure 3.1a) exhibited much lower turbidity, as seen from the color contrast in the water samples in Figure 3.1d.

3.3 Methodology

Three pre-processing options were considered in this study (section 3.3.1). For each pre-processing option, matchups between satellite and in-situ turbidity measurements were prepared and split into training, validation, and test datasets (section 3.3.2), and a LightGBM model was trained using the training dataset (section 3.3.3). Model performance metrics (section 3.3.4) were subsequently calculated after applying the model to the test dataset, and additional evaluations were conducted to examine how results vary with distance to the shore, using the combined test datasets from both L8/9 and S2. Lastly, turbidity maps for each pre-processing option were produced for the two acquisition dates, August 2, 2024, and September 11, 2024, which were selected because images were available from both satellite sensor types on these dates.

3.3.1 Atmospheric correction

TOA reflectance was computed directly from the raw satellite imagery without any AC. ACOLITE's DSF technique (Vanhellemont, 2019; Vanhellemont and Ruddick, 2018) version 20231023.0 was then employed to retrieve R_w from S2 and L8/9 images, both with and without prior AE correction using T-Mart version 2.4.5 (Wu et al., 2024). T-Mart uses Global Modeling and Assimilation Office (GMAO) MERRA2 ancillary data (Global Modeling and Assimilation Office (GMAO), 2015) to characterize the optical properties of the atmosphere, including aerosol optical thickness at 550 nm. It then generates atmospheric point-spread functions and applies a convolution of the satellite imagery to correct for AE one band at a time.

Default settings were used for both ACOLITE and T-Mart, with the following modifications: Sun-glint correction (Harmel et al., 2018b) and GMAO MERRA-2 atmospheric ancillary datasets (OBPG., 2023) were enabled in ACOLITE, and the negative- R_w filter in ACOLITE was deactivated to retain negative-reflectance retrievals affected by sensor noise and overcorrections in AC as the retrieved spectra, although with negative values, can contain useful information.

3.3.2 Dataset creation

For model development, all spectral bands of L8/9 OLI, except the panchromatic band, and all spectral bands of Sentinel-2 MSI were used in this study. In addition to extracting per-pixel reflectance, mean and median kernels were computed using Window Sizes (WSs) of 3×3 , 5×5 , and 7×7 pixels. To reduce the impact of information leakage and spatial autocorrelation between training, validation, and test dataset, a spatial blocking approach (Knudby and Richardson, 2023) was adopted for data splitting. Specifically, a grid was superimposed over the study area, and random cells were progressively assigned to the test dataset until it contained at least 20% of the total samples. Using the same random selection process, another 20% of the samples were allocated to the validation dataset, while the remaining 60% were used for training. The cell size for spatial blocking was defined based on the semivariogram range of the in-situ turbidity measurements,

ensuring statistical independence between neighboring blocks. Semivariogram analyses were performed separately for the in-situ turbidity measurements collected simultaneously with the S2 and L8/9 overpasses (Figures 3.2a and 3.3a), revealing similar range values of ~400 m for both sensors. Consequently, a cell size of 400 m was adopted. Figures 3.2b and 3.3b illustrate the spatial distribution of the in-situ turbidity measurements corresponding to L8/9 and S2 acquisitions.

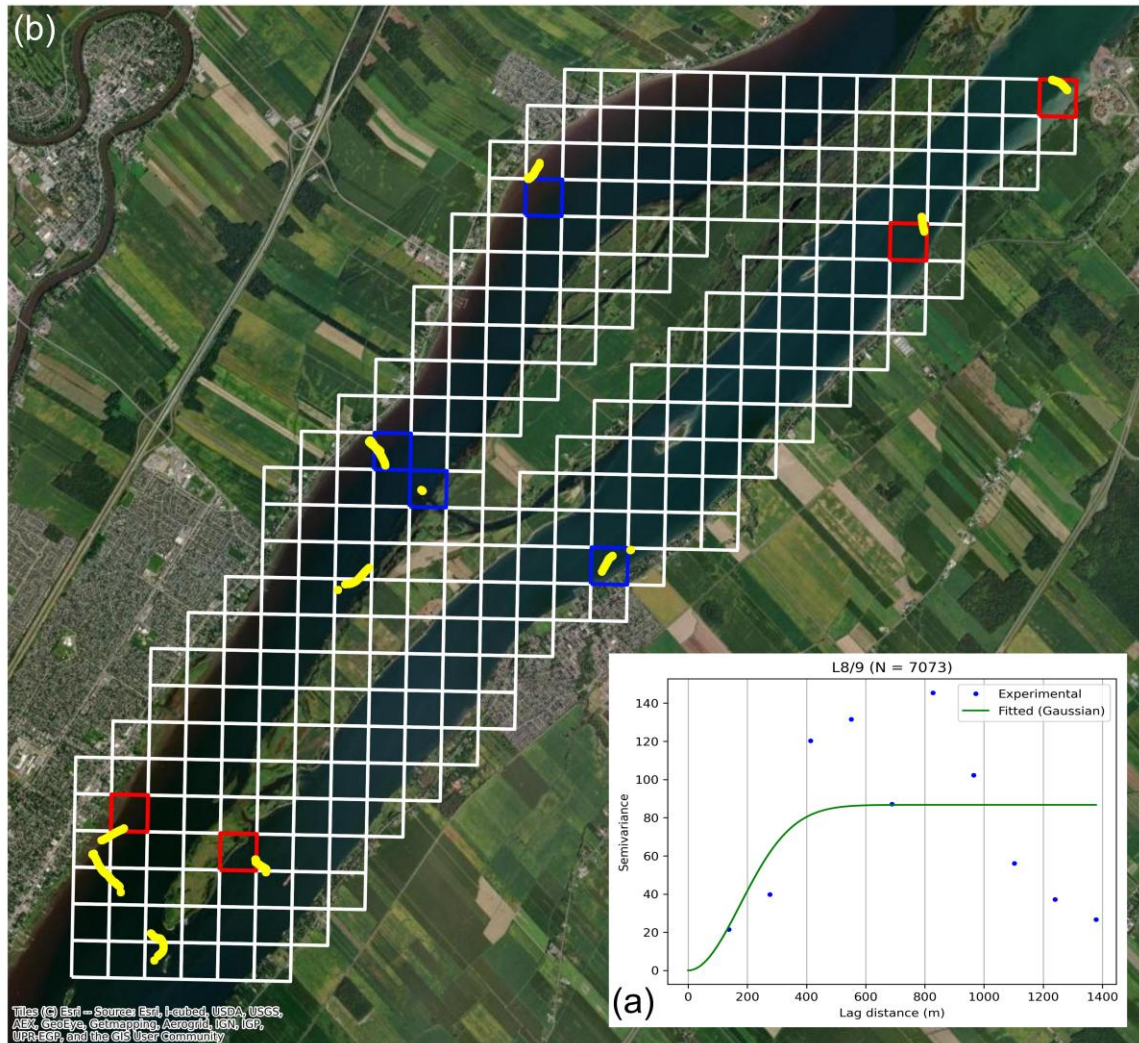


Figure 3.2. Spatial distribution of in-situ turbidity measurements corresponding to L8/9 acquisitions (a) and their experimental semivariogram (b). Yellow points indicate the locations of in-situ turbidity measurements, while the overlaid grid represents the spatial blocking scheme used for data splitting. Cells outlined in red denote the test dataset, those outlined in blue represent the validation dataset, and the remaining cells were assigned for model training

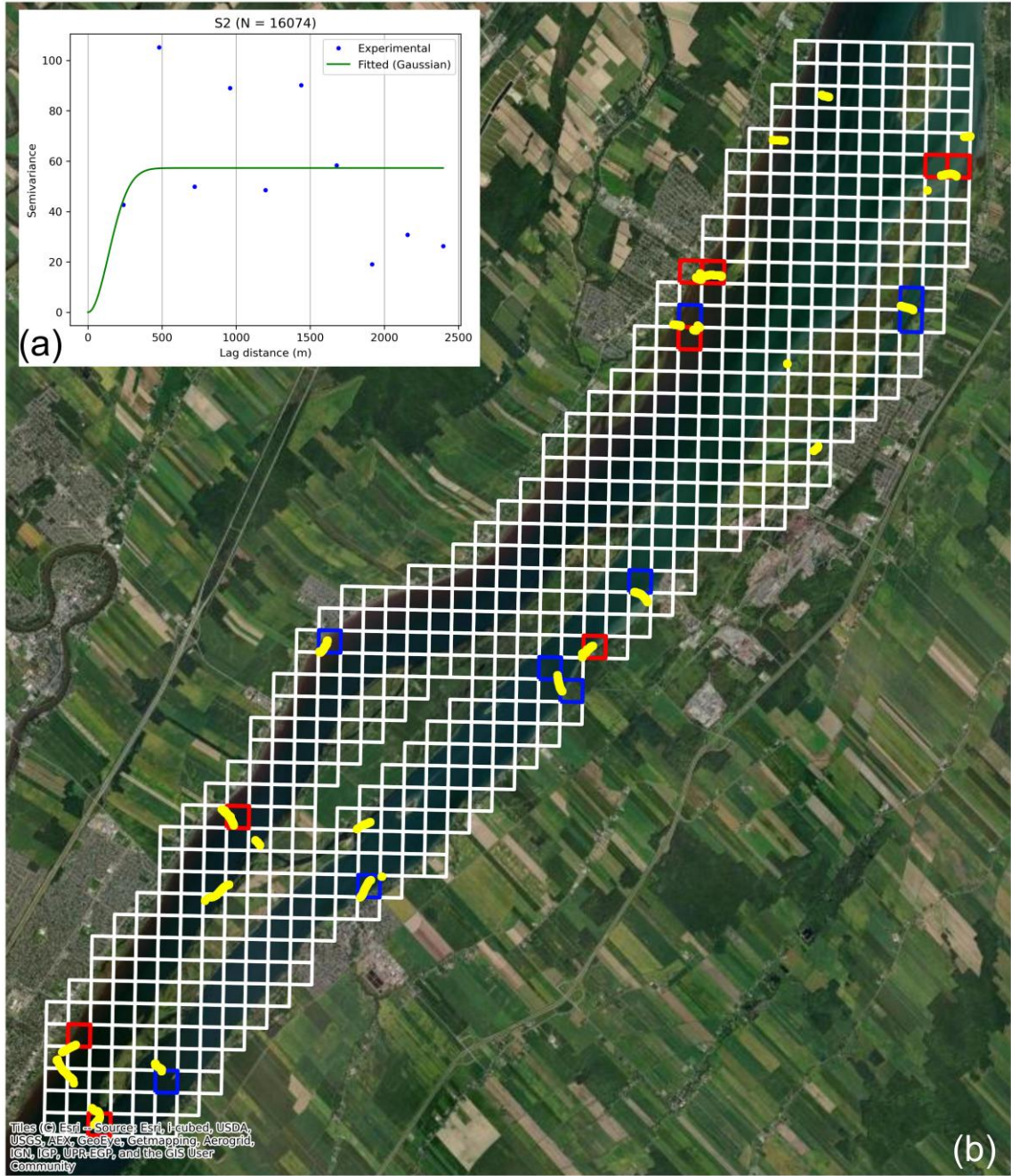


Figure 3.3. Spatial distribution of in-situ turbidity measurements corresponding to S2 acquisitions (a) and their experimental semivariogram (b). Yellow points indicate the locations of in-situ turbidity measurements, while the overlaid grid represents the spatial blocking scheme used for data splitting. Cells outlined in red denote the test dataset, those outlined in blue represent the validation dataset, and the remaining cells were assigned for model training.

The number of turbidity measurements with turbidity values exceeding 45 NTU was relatively limited compared to the number of turbidity measurements within the 0–40 NTU range (Figure

3.4); consequently, model training and performance evaluation are more reliable for turbidity values between 0 and 40 NTU, where data density is substantially higher than for turbidity values exceeding 45 NTU.

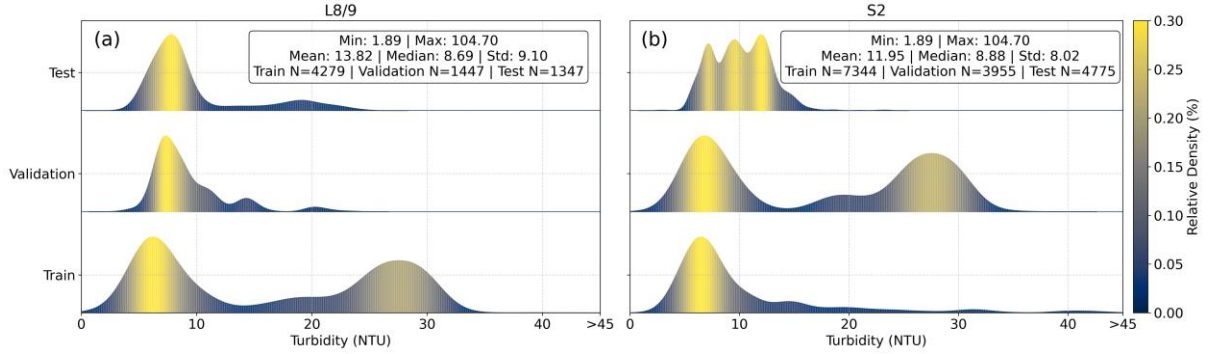


Figure 3.4. Ridge-density plots of in-situ turbidity measurements in the training, validation, and test datasets for (a) L8/9 and (b) S2 matchups. Kernel density estimates are shown for values up to 45 NTU, with densities encoded using a continuous color scale. Samples exceeding 45 NTU are represented as aggregated bars. Summary statistics and sample counts for each dataset are provided within each subplot.

3.3.3 Regression Model

The LightGBM (Ke et al., 2017) model was used for turbidity estimation. LightGBM is a tree-based ensemble learning model that uses gradient boosting to iteratively minimize prediction errors. The model optimizes a loss function, which measures the difference between the estimated and measured turbidity values.

Model development and training procedures were implemented in Python, using the *LightGBM* library (Ke et al., 2017). To enhance model performance and prevent overfitting, key hyperparameters were optimized through the *Optuna* library (Akiba et al., 2019), which applies a Bayesian optimization approach to explore the parameter space. The model was trained using the training dataset, while hyperparameter tuning was performed on the validation dataset to prevent overfitting. The tuning process aimed to minimize the RMSE using the validation dataset. After the model's hyperparameters were finalized, the test dataset was used for final performance evaluation. The hyperparameters adjusted during tuning included the number of boosting iterations

($n_estimators$, varied from 1 to 200), the maximum tree depth (max_depth , between -1 and 200), the learning rate ($learning_rate$, between 0.0001 and 1), the fraction of features used per iteration ($feature_fraction$, between 0.0001 and 1.0), and the maximum number of leaves per tree (num_leaves , between 2 and 6). All other hyperparameters were retained at their default values.

The optimization process was executed 200 times in *Optuna*. After running 200 trials in *Optuna*, the set of hyperparameters corresponding to the trial with the lowest RMSE on the validation dataset was selected for the final model.

Finally, regression models were trained and evaluated using each kernel (section 3.3.2) for the three pre-processing options. For each pre-processing option, the kernel yielding the lowest RMSE for test dataset was selected for comparison. After completing the evaluation phase, the entire dataset was used to train the final model, which was then applied to generate turbidity maps.

3.3.4 Performance metrics

Three statistical metrics were employed as evaluation metrics: RMSE, ε (Morley et al., 2018), and β (Morley et al., 2018) (Equation 3.1-3):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_i - O_i)^2} \quad (3.1)$$

$$\beta = 100 \times \text{sign} \left[\text{median} \left(\log \frac{E_i}{O_i} \right) \right] \times \left(e^{|\text{median}(\log \frac{E_i}{O_i})|} - 1 \right) \quad (3.2)$$

$$\varepsilon = 100 \times \left(e^{\text{median}[\log(\frac{E_i}{O_i})]} - 1 \right) \quad (3.3)$$

N is the number of observations, and O , E , $\bar{O}$, and $\bar{E}$ are the observations, estimated values, the average of the observations, and the average of the estimated values, respectively.

3.4 Results

3.4.1 Matchups

Figure 3.5 shows the RMSE values on the test dataset for L8/9 and S2, obtained from the LightGBM model trained using mean and median kernels with different WSs across the three pre-processing options. For L8/9, the median kernel with WS = 7 produced the lowest RMSE under the AE pre-processing option, WS = 5 under the AC pre-processing, and WS = 3 under the No pre-processing. For S2, the per-pixel reflectance (WS = 1) resulted in the lowest RMSE for both the No and AC pre-processing options, whereas the median kernel with WS = 7 achieved the best performance for the AE option.

As shown in Figure 3.5, the choice of kernel influenced RMSE. For example, under the AE pre-processing using the L8/9 dataset, replacing the median kernel (WS = 7) with a mean kernel (WS = 5) increased RMSE by $\sim 156\%$.

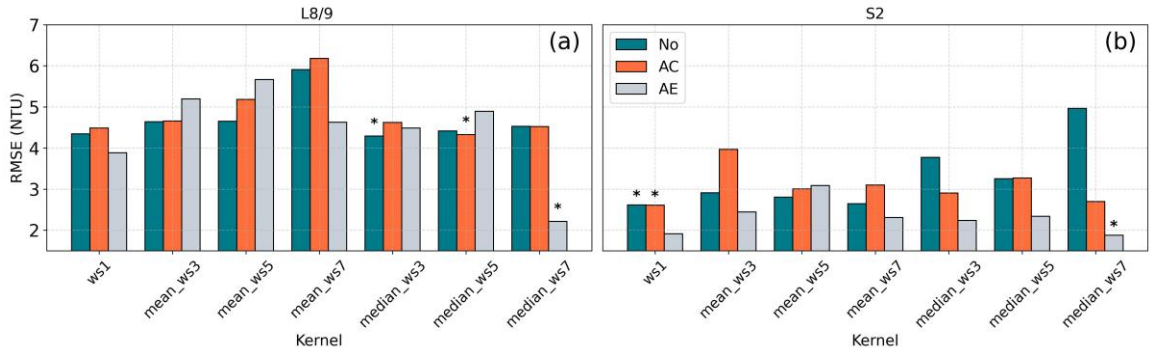


Figure 3.5. RMSE values on the test dataset for (a) L8/9 and (b) S2, obtained from the LightGBM model trained using per-pixel and kernel-based (mean and median with 3×3 , 5×5 , and 7×7 window sizes) reflectance inputs across the three pre-processing options. The star symbol marks the kernel that achieved the lowest RMSE for each pre-processing option.

3.4.2 Turbidity retrieval

The tuned hyperparameters used for the LightGBM models are listed in Table 3.2. Figure 3.6 presents the scatterplots between the measured and estimated turbidity values using the test dataset. Each panel displays the RMSE, β , and ϵ metrics, calculated for the test dataset.

Table 3.2. Tuned hyperparameters of the LightGBM model for turbidity estimation under the three pre-processing options for both L8/9 and S2 datasets.

Sensor	Scenario	$n_estimators$	max_depth	$learning_rate$	$feature_fraction$	num_leaves
L8/9	No	710	21	0.41	0.15	115
	AC	962	14	0.004	0.12	102
	AE	594	1	0.19	0.47	27
S2	No	223	0	0.22	0.22	20
	AC	259	49	0.10	0.75	5
	AE	362	37	0.05	0.22	2

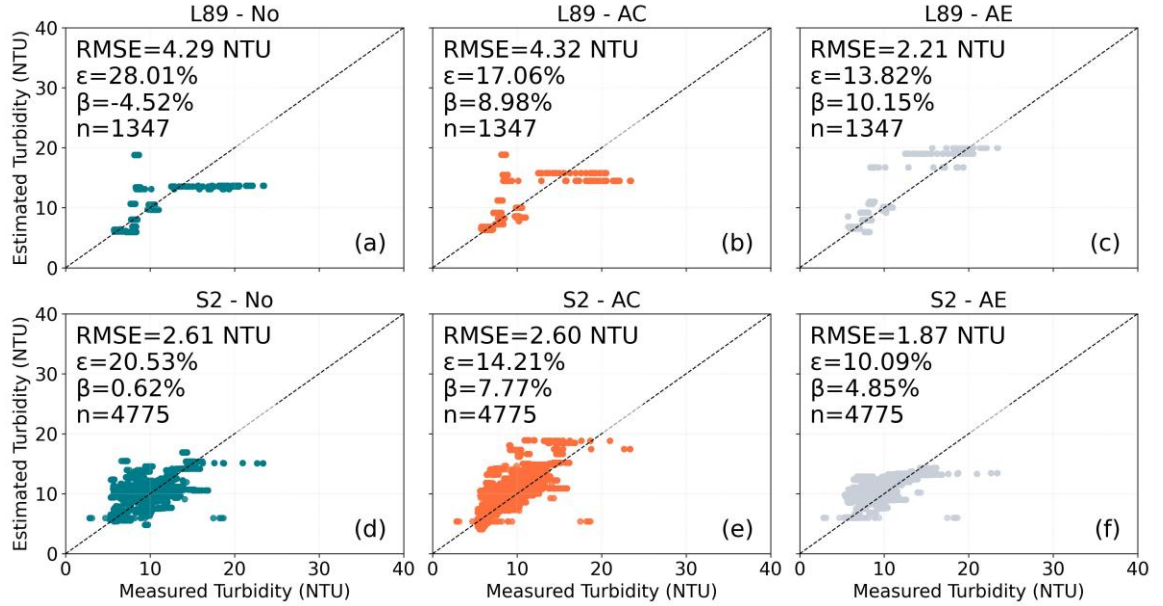


Figure 3.6. Scatterplots comparing measured and estimated turbidity values obtained from the LightGBM model for L8/9 (top row; panels (a)–(c)) and S2 (bottom row; panels (d)–(f)). Each column corresponds to one pre-processing option: No (first column; panels (a) and (d)), AC (second column; panels (b) and (e)), and AE (third column; panels (c) and (f)). Text shows RMSE, β , ϵ , and number of samples for the test dataset.

3.4.3 Variability of correction results with distance to shore

Figure 3.7 shows the variation of model performance with distance from the shore. As shown in Figure 3.7, the AE pre-processing exhibits superior performance compared to the other pre-processing options across all three metrics (RMSE, ϵ , and β) within 0–300 m from the shore.

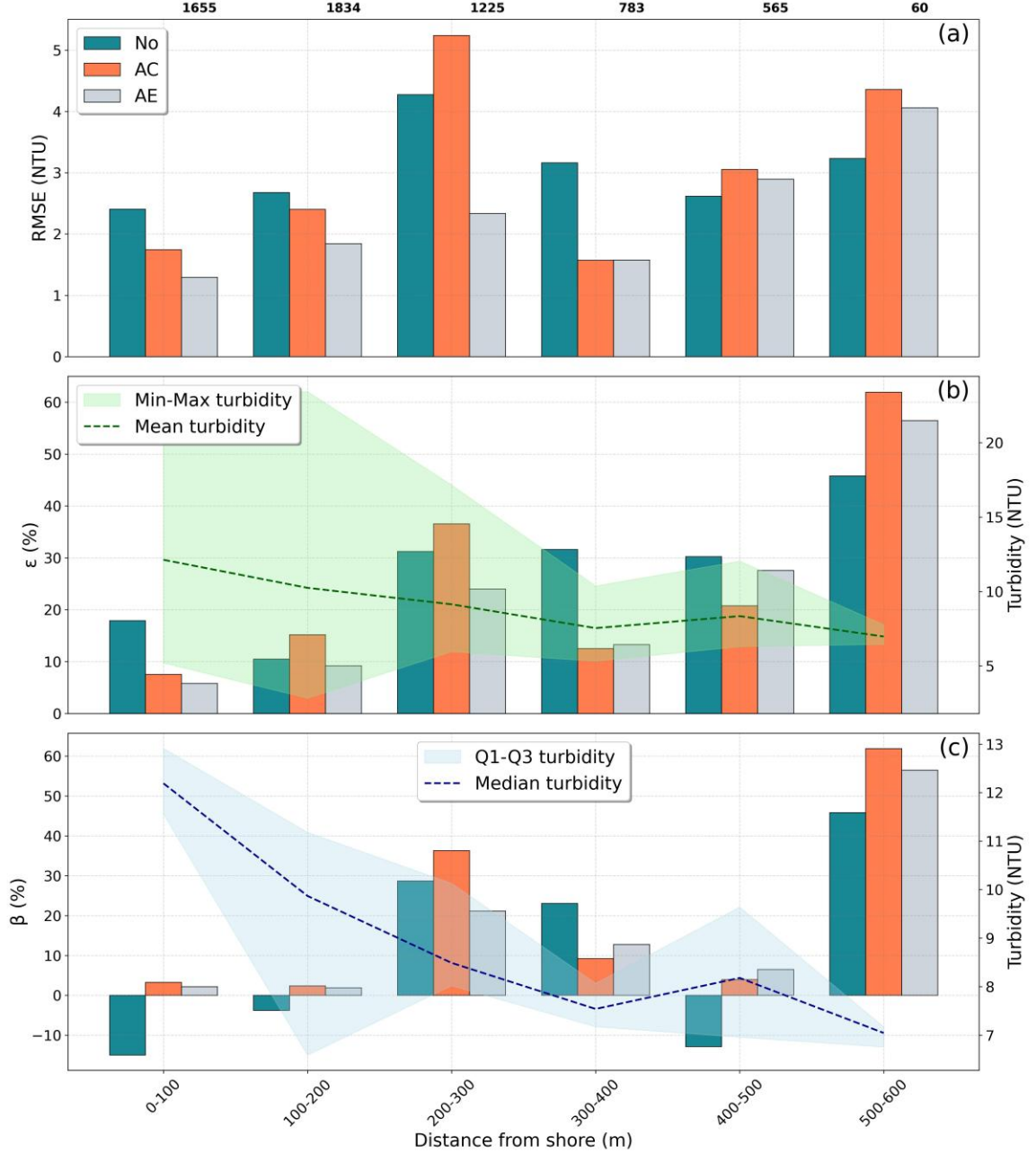


Figure 3.7. Variability of model performance metrics with distance to the shore, combining test data from L8/9 and S2 sensors. (a) Changes in RMSE with distance to the shore, where the numbers above the bars indicate the number of samples within each distance interval. (b) ϵ (left y-axis) alongside the minimum, maximum, and mean turbidity values (right y-axis) represented by the green shaded polygon. (c) β (left y-axis) with the interquartile range and median turbidity (right y-axis) illustrated by the blue shaded polygon.

3.4.4 Visual comparison

Figures 3.8 and 3.9 present the turbidity maps generated using the three pre-processing options for August 2, 2024, and September 11, 2024, respectively. All maps show agreement with the field

observations, capturing the overall spatial pattern of low turbidity along the southern shore (Site 1 in Figure 3.1a) and high turbidity along the northern shore (Site 2 in Figure 3.1a).

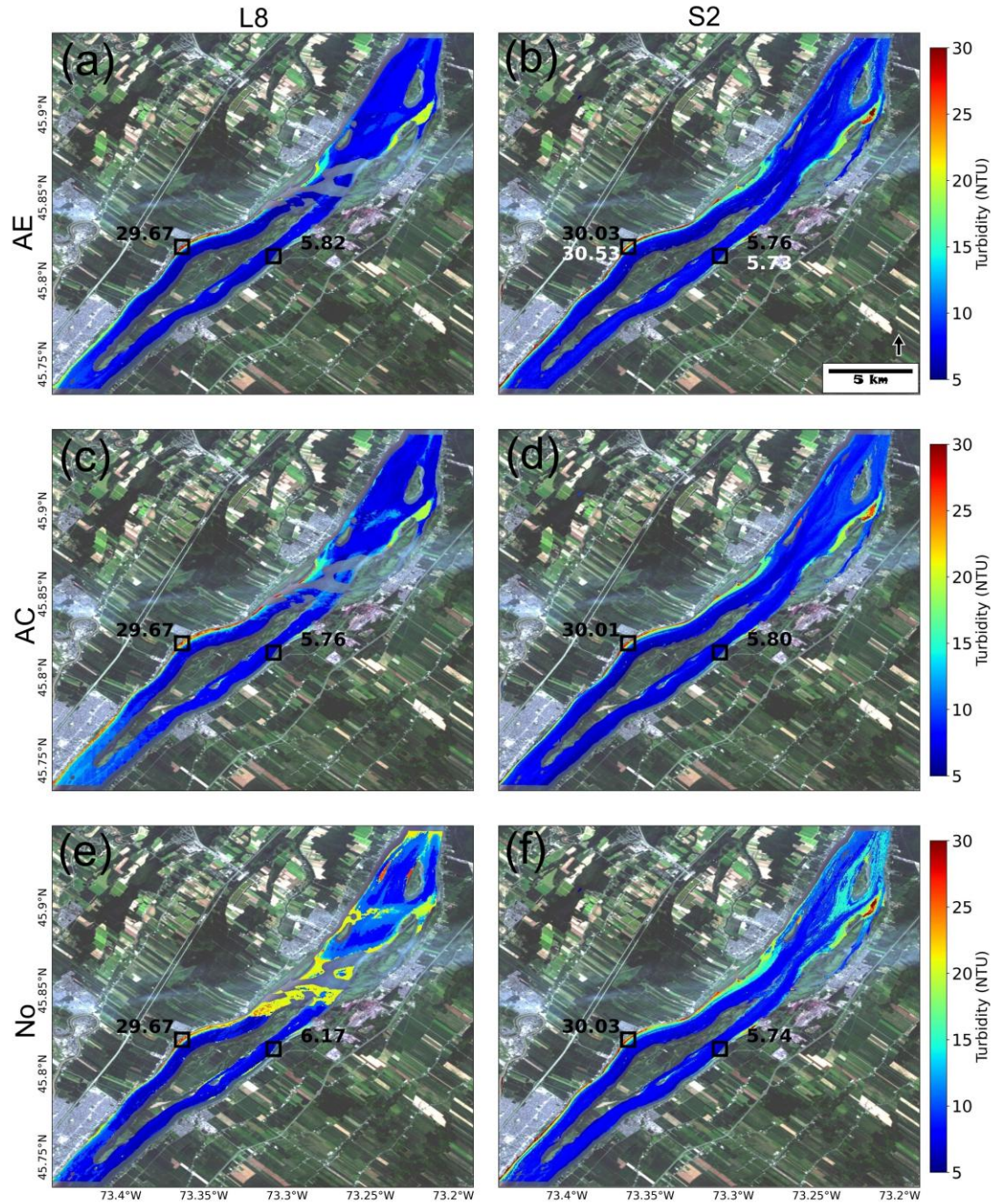


Figure 3.8. Spatial distribution of turbidity estimates on August 2, 2024, derived from L8/9 (left column; panels (a), (c), and (e)) and S2 (right column; panels (b), (d), and (f)). Each row represents a different pre-processing option: AE (top row; panels (a) and (b)), AC (middle row; panels (c) and (d)), and No (bottom row; panels (e) and (f)). Black labels indicate estimated turbidity values, while white labels on the upper-right map (panel (b)) show in-situ turbidity measurements.

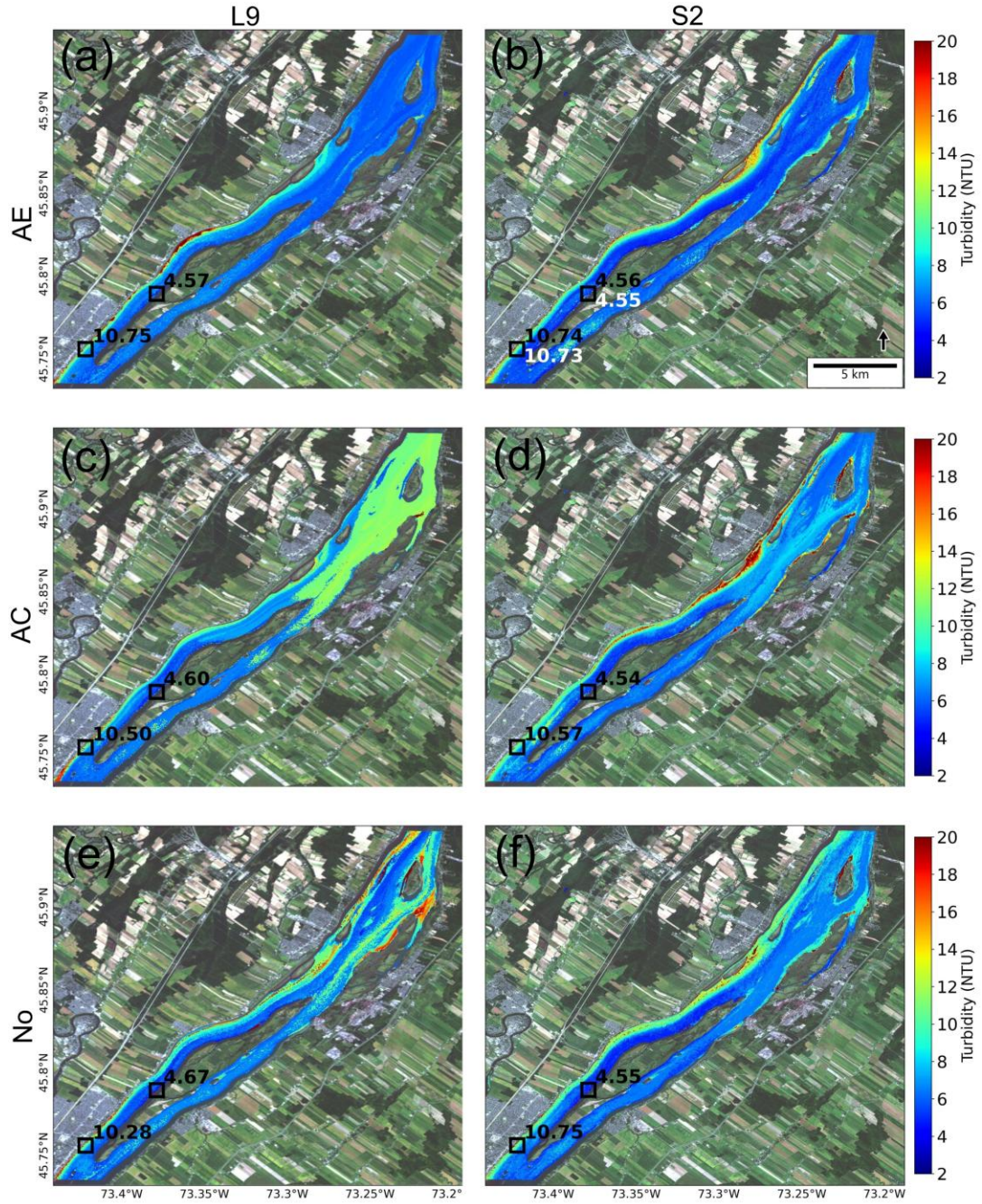


Figure 3.9. Spatial distribution of turbidity estimates on September 11, 2024, derived from L8/9 (left column; panels (a), (c), and (e)) and S2 (right column; panels (b), (d), and (f)). Each row represents a different pre-processing option: AE (top row; panels (a) and (b)), AC (middle row; panels (c) and (d)), and No (bottom row; panels (e) and (f)). Black labels indicate estimated turbidity values, while white labels on the upper-right map (panel (b)) show in-situ turbidity measurements.

3.5 Discussion

3.5.1 Performance Comparison Among pre-processing options

As shown in Figures 3.6a and 3.6b, for L8/9, applying DSF only (AC pre-processing) did not improve RMSE or $|\beta|$ relative to TOA (No pre-processing): RMSE increased from 4.29 to 4.32 NTU, and $|\beta|$ increased from 4.52% to 8.98%, corresponding to a 98.7% deterioration. However, the ε metric improved under the AC pre-processing, decreasing from 28.01% to 17.06%, which corresponds to a 39.1% reduction. Conversely, as illustrated in Figures 3.6b and 3.6c, introducing the AE correction with T-Mart improved performance, reducing RMSE from 4.32 to 2.21 NTU (an improvement of 48.8%) and ε from 17.06% to 13.82% (an improvement of 19.0%) compared to the AC pre-processing. Relative to the No pre-processing (Figures 3.6a and 3.6c), the AE pre-processing also enhanced performance: RMSE decreased by 48.5% and ε by 50.7%. Therefore, in terms of RMSE and ε , the AE pre-processing outperformed both the AC and No pre-processing options. However, for $|\beta|$, the No pre-processing yielded the lowest value among all ($\beta = -4.52\%$). For S2, the AE pre-processing achieved the best performance in terms of RMSE and ε , followed by the AC pre-processing, while the No pre-processing exhibited the weakest performance. Specifically, RMSE and ε decreased from 2.61 NTU and 20.53% under the No pre-processing to 1.87 NTU and 10.09% under the AE pre-processing (Figures 3.6d and 3.6e), corresponding to improvements of approximately 28.4% and 50.8%, respectively. However, the No pre-processing yielded the lowest $|\beta|$ compared to the other pre-processing options ($\beta = 0.62\%$).

3.5.2 Variation of Scenario Performance with Shoreline Distance

As shown in Figure 3.7, within 0-300 m from the shoreline, where AE and turbidity levels are generally higher than those observed at 300-600 m based on mean and median values (Figure 3.7b and 3.7c), the AE pre-processing outperformed the other two pre-processing options. The superior performance of the AE option was most evident 200-300 m from shore, where RMSE decreased by >45% compared to the others.

From 0-300 m, RMSE increased with distance from the shoreline under all pre-processing options. However, the rate of increase was substantially lower for the AE pre-processing compared to the others. For instance, between 100-200 m and 200-300 m, RMSE increased by ~126% under the AC pre-processing but only by ~21% under the AE pre-processing. This indicates that model performance was more stable and consistent within 0-300 m when AE pre-processing was applied compared to when it was not.

Between 300–500 m, differences among the three pre-processing options become less clear, with performance varying across metrics rather than showing a single dominant approach. At distances greater than 500 m from the shoreline, the No pre-processing performed best across all metrics. However, the 500-600 m bin contained only 60 samples, which limits the strength of this conclusion. Overall, these results indicate that AE pre-processing is most important near the shoreline (0–300 m), performance becomes mixed in the middle range (300-500 m), and No pre-processing performs best farther offshore (500-600 m).

3.5.3 Visual Assessment of Turbidity Maps

In Figure 3.8, a cloud is visible in the L8 image, which introduces noise in the turbidity maps generated by the AC and No pre-processing options (Figure 3.8c and 3.8e). The turbidity map derived from the AE pre-processing (Figure 3.8a) was less affected by this cloud, likely due to T-Mart treating clouds as reflecting surfaces and thus correcting for the AE from them in the processing. Additionally, in the lower-left part of the L8 turbidity maps, some artifacts were observed in the AC and No pre-processing options (Figure 3.8c and 3.8e), while the AE pre-processing remained cleaner and smoother in that region (Figure 3.8a).

The differences between the pre-processing options were more evident in the L9 maps shown in Figure 3.9. In the upper-right part of these maps, the AC and No pre-processing options (Figure 3.9c and 3.9e) exhibited noise, whereas the AE pre-processing (Figure 3.9a) produced a smoother spatial distribution. In the AC pre-processing (Figure 3.9c), the model appeared to fail in accurately

predicting turbidity in the northeast section of the map. For S2, in Figure 3.8, the turbidity maps generated by the AE and AC pre-processing options (Figure 3.8b and 3.8d) were similar and both performed better than the No pre-processing (Figure 3.8f), which showed some noise in the upper-right area. In contrast, Figure 3.9 clearly demonstrated that the AE pre-processing (Figure 3.9b) yielded a smoother and less noisy turbidity map compared to the other two pre-processing options.

3.6 Conclusion

This study compared three pre-processing options for turbidity estimation in the St. Lawrence River using S2 and L8/9 imagery with a LightGBM model: (1) the use of TOA reflectance data, (2) AC with DSF, and (3) AE correction using T-Mart followed by atmospheric correction with ACOLITE. The results demonstrated that the AE pre-processing improved RMSE and ϵ compared to the other two options, while the No pre-processing yielded the lowest $|\beta|$. Moreover, in terms of RMSE, the No and AC pre-processing options produced very similar values, whereas the AE pre-processing achieved improvements for both the S2 and L8/9 datasets. Additionally, the AE pre-processing led to smoother and more realistic turbidity maps, particularly in regions influenced by cloud contamination. Performance patterns across shoreline distances further revealed that AE correction outperformed the alternatives within 0-300 m from the shore, where the AE is strong.

Author Contributions

M.A. (Mohsen Ansari): conceptualization, data curation, formal analysis, investigation, methodology, resources, validation, visualization, software, writing original draft. Y.W. (Yulun Wu): data curation, software, writing – review and editing. A.K.: conceptualization, funding acquisition, project administration, supervision, writing – review and editing. All authors have read and agreed to the published version of the manuscript.

Funding

This study received financial support from the Canadian Space Agency under the smartHarbour initiative and the Mitacs Accelerate program (IT40023).

Acknowledgements

The authors would like to acknowledge the Canadian Space Agency for the financial support under the smartHarbour initiative, implemented in collaboration with Public Services & Procurement Canada and the Montréal Port Authority to foster the development and adoption of best practices for environmental monitoring as well as the protection of terrestrial and aquatic ecosystems. Additional support was provided by WSP, the Mitacs Accelerate program, and members of the University of Ottawa's Shallow Water Earth Observation Laboratory. Assistance from ChatGPT-4o was used for portions of code development (e.g. code structure and debugging) and improving readability (e.g. text editing and word suggestions). All outputs were reviewed and edited by the authors as needed.

Data Availability Statement

The field data presented in this article are not publicly available; they were collected jointly by the University of Ottawa and WSP Canada Inc. for the smartHarbour project, conducted under the oversight of the Montreal Port Authority (MPA). Data access requests should be directed to WSP Canada Inc. All satellite data used in this article are publicly available.

Conflicts of Interest

The authors declare no conflict of interest.

Chapter 4: River Salinity Mapping through Machine Learning and Statistical Modeling using Landsat 8 OLI Imagery

Mohsen Ansari¹, Anders Knudby¹, and Saeid Homayouni²

¹Department of Geography, Environment and Geomatics, University of Ottawa, Ottawa, Canada

²Institut National de la Recherche Scientifique, Centre Eau Terre Environnement, Quebec, Canada

Abstract

This study uses Landsat 8 OLI imagery and 102 in situ salinity data points to investigate salinity mapping in the Karun River, southwestern Iran. A total of 24 features, including salinity indices and Landsat 8 OLI spectral bands, were assessed using the Random Forest Feature Importance Score (RFFIS), Sobol' sensitivity analysis, and correlation with salinity to identify the most sensitive features for salinity estimation. These included the Red and Green bands, Salinity index 2–6, Normalized Suspended Material Index (NSMI), and Enhanced Green Ratio Index (EGRI). A total of 24 regression models, including statistical, kernel-based, Neural Network (NN)-based, and Decision Tree (DT)-based models, were evaluated using statistical error metrics and global Moran's I measure of residual spatial autocorrelation. The DT-based models, specifically Gradient Boosted DT (GBDT), outperformed other models, demonstrating low errors, bias, and non-significant residual spatial autocorrelation. Kernel-based models performed better than conventional linear models, while NN models tended to underfit. Residual spatial autocorrelation analysis indicated that models incorporating spatial information reduced residual autocorrelation. Landsat 8 OLI imagery effectively mapped salinity dynamics, revealing increased salinity from Gotvand to Ahvaz city due to agricultural activities and the Gachsaran formation within the reservoir.

Keywords: Water Salinity; Spatial Autocorrelation; Multispectral Imagery; Regression Analysis; Rivers and Estuaries; Machine Learning.

DOI: 10.1016/j.asr.2025.03.037

4.1 Introduction

The salinity of rivers and estuaries is an important water quality parameter (WQP) that influences freshwater supply and human well-being (Cañedo-Argüelles et al., 2016), the suitability of water for irrigation and agriculture (Tessema et al., 2022), the animal species survival (Jualaong et al., 2019), and the growth of aquatic vegetation (Sebilian Wittingham et al., 2019). Salinity can vary across space and time due to fluctuations in precipitation and runoff in rivers (Sumner and Belaine, 2005), water circulation (Bowden and Sharaf El Din, 1966), tidal and wind-driven mixing of water masses (Uncles and Stephens, 2011), as well as changes in the volume of freshwater discharge in estuaries (Jiang et al., 2024). Monitoring changes in salinity, therefore, has many practical applications for water management (Dasgupta et al., 2013; Fichez et al., 2017; James et al., 2003). Salinity is the concentration of dissolved salts in a volume of water. The most common salts in water include the cations sodium (Na^+), potassium (K^+), magnesium (Mg^{2+}), and calcium (Ca^{2+}), and the anions chloride (Cl^-), sulfate (SO_4^{2-}), and bicarbonate (HCO_3^-) (El-Swaify, 2000). These ions increase both the salinity and the Electrical Conductivity (EC) of the water, creating a close relationship between salinity and EC that allows salinity to be measured indirectly via EC measurements (Dahaan et al., 2016; El-Swaify, 2000; U.S. Geological Survey, 2019b; Wagner et al., 2006a), an approach that is typically used instead of direct measurement of salinity (Khorram, 1985; Rusydi, 2018; Wagner et al., 2006a). Measurement of EC does not require lab access and can be automated to acquire data at regular intervals, allowing for high-frequency monitoring. However, such tracking is spatially limited to the location at which probes are deployed; effort has therefore also gone into the development of methods to monitor salinity using satellite imagery that can provide less frequent but spatially complete coverage of an area of interest.

The absorption and scattering of light in water, and therefore water colour, is directly influenced by salinity (Röttgers et al., 2014; Zhang et al., 2009; Zhang and Hu, 2021). However, for the salinity variations that are relevant to natural waters, this influence is relatively small compared to that of

more optically active water constituents such as chlorophyll-a found in algal particles, coloured dissolved organic matter (CDOM), and total suspended sediment (TSS) (Mishra et al., 2017). While knowledge of salinity can improve the retrieval of other water quality variables from optical satellite imagery (Werdell et al., 2013) by fine-tuning retrieval algorithms, direct mapping of salinity with inversion algorithms applied to optical satellite data is not currently possible, and direct mapping of salinity from space is limited to microwave instruments that operate at spatial resolutions too coarse to have value for rivers and estuaries (Kerr et al., 2010; Lagerloef et al., 2008). As a result, practical approaches to salinity mapping with optical data have relied on indirect associations between salinity and other water quality parameters that are i) empirically related to salinity and ii) easier to observe with satellite data. For example, many studies report a linear, inverse correlation between salinity and the CDOM absorption coefficient in coastal waters (Blough et al., 1993; Bowers et al., 2004; D'Sa et al., 2000; Hu et al., 2004; Siddorn et al., 2001). Several studies have also observed a relationship between salinity and TSS (Adjovu et al., 2023b, 2023c; Wang and Xu, 2008), noting that they are negatively correlated (Fang et al., 2010; Wang and Xu, 2008). The successful remote sensing of CDOM and TSS and their observed relationships with salinity suggests that salinity can be mapped indirectly using satellite data, either explicitly using TSS and CDOM as proxies or implicitly by using observations at wavelengths where water color is influenced by TSS and CDOM concentrations (Khadim et al., 2017; Nguyen et al., 2018). From a practical perspective, mapping salinity via an explicit relationship with CDOM and/or TSS is demanding because it requires establishing an empirical relationship between salinity and CDOM and/or TSS, which in turn requires suitable data on those variables. The easier approach is to establish an empirical regression relationship directly between salinity and water colour, with the understanding that CDOM and/or TSS are an implicit link between the two.

Three important considerations involve using empirical regression models to monitor salinity in rivers and estuaries: i) *Selecting appropriate satellite data*. Due to the dynamic nature of rivers, salinity varies spatially and temporally. The spatial resolution of the satellite imagery must be

sufficient to allow differentiation of salinity between different parts of the area of interest, the sensor should possess spectral bands sensitive to variations in CDOM and/or TSS, and the temporal resolution must be sufficient for specific monitoring requirements. ii) *Feature engineering*. Selecting spectral bands and indices sensitive to variations in CDOM and/or TSS is important to optimize model performance, but including an excessive number of variables can lead to model overfitting (Nguyen et al., 2018). iii) *Model selection and parameter optimization*. A finely tuned and flexible empirical model can help produce accurate salinity estimates because the spectral behavior of CDOM and/or TSS, and the relationships between CDOM/TSS and salinity, can be complex. Especially when *in situ* data are scarce, employing a regression method that prevents underfitting and avoids overfitting is important.

Previous salinity mapping studies have used sensors such as SeaWiFS (D'Sa et al., 2002), Aquarius (Kao et al., 2018), Visible Infrared Imaging Radiometer Suite (VIIRS) (Vandermeulen et al., 2014), and Moderate Resolution Imaging Spectroradiometer (MODIS) - Aqua (DeLuca et al., 2020). However, due to their low spatial resolution, these sensors are unsuitable for most inland waters, especially rivers (Priyadarshini et al., 2023). For salinity mapping of inland waters, previous studies have used the Landsat series of sensors as well as Sentinel-2 MSI and the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (Table 4.1). These sensors have bands suitable for detection of variations in CDOM absorption (Li et al., 2023) and TSS concentration (Adjovu et al., 2023b), and thus indirectly salinity, and they also allow for the generation of spectral indices correlated with salinity (Mejía Ávila et al., 2022) including dedicated salinity indices (Adjovu et al., 2023b, 2022; Ferdous and Rahman, 2020; Maliki et al., 2020), NSMI (Montalvo, 2010), and the normalized difference suspended sediment index (NDSSI) (Hossain et al., 2010). While visible and NIR bands are used in many studies due to their association with absorption peaks of CDOM and/or TSS, the bands and spectral indices used in combination to study salinity differ between studies.

Table 4.1. Summary of studies that have used moderate spatial-resolution sensors to map salinity, in chronological order. NIR: near-infrared band, SWIR: shortwave infrared band, MSS: Landsat Multispectral Scanner, TM: Landsat Thematic Mapper, OLI: Landsat Operational Line Imager, ETM+: Landsat Enhanced Thematic Mapper Plus, and MSI: Sentinel-2 MSI.

Sensor(s)	Band(s) and Feature(s)	Empirical Model(s)	Reference
MSS	Red and NIR 1	LR	(Khorram, 1985)
TM	NIR and SWIR 2	LR	(Lavery et al., 1993)
MSS and TM	Blue, NIR, and SWIR 2	LR	(Vuille and Baumgartner, 1993)
TM	Blue, Green, Red, NIR, and SWIR 1	LR and ridge regression	(Wang and Xu, 2008)
TM	Blue, Red, and NIR	LR	(Zhang et al., 2012)
TM	First three principal component analysis components	GWR	(Xie et al., 2013)
TM	Blue, Green, Red, NIR, SWIR 1, and SWIR 2	SWOM	(Khadim et al., 2017)
OLI	Coastal Aerosol, Blue, Green, and Red	LR	(Zhao et al., 2017)
OLI	Red	Non-linear Regression	(Mushtaq and Nee Lala, 2017)
TM	Blue, Red, and NIR	LR and GWR	(Nazeer and Bilal, 2018)
OLI	Blue, Green, Red, and SWIR 2	MLR, DT, and RF	(Nguyen et al., 2018)
ASTER	Red, SWIR 2, and SWIR 4	LR	(Abdelmalik, 2018)
OLI	Coastal Aerosol, Blue, and Green	SVR	(Ansari and Akhoondzadeh, 2019)
MSI	All Sentinel-2 bands and image properties	ANN	(Medina-Lopez and Ureña-Fuentes, 2019)
OLI	Blue, Green, and Red	LR	(Ferdous et al., 2019)
OLI	Salinity index 2	LR	(Maliki et al., 2020)
OLI	Coastal Aerosol, Blue, and Green	LR, SVR, and MLP	(Mohsen Ansari and Akhoondzadeh, 2020)
TM and OLI	Blue, Green, and Red	MLR	(Ferdous and Rahman, 2020)
MSI	All Sentinel-2 bands and image properties	ANN	(Medina-Lopez, 2020)
OLI	Coastal Aerosol, Blue, and Green	MLP and SVR	(Ansari and Akhoondzadeh, 2021)
OLI and MSI	Blue, Green, and Red	ANN, MLR, and ANFIS	(Bayati and Danesh-Yazdi, 2021)
MSI	Red edge 2, Red edge 3, NIR 1, NIR 2, SWIR 2, and SWIR 1	LR, RF and AdaBoost	(Borovskaya et al., 2022)
OLI	Salinity index 5	LR	(Adjovu et al., 2022)
TM and OLI	Blue, NIR, and SWIR 2	LR	(Chatterjee and Bhadra, 2023)
ETM+ and OLI	Blue, Green, Red, NIR, SWIR 1, and SWIR 2	LightGBM, RF, and ANN	(Dai et al., 2023)
MSI	All Sentinel-2 bands except from SWIR 3 and SWIR 4	DT, RF, XGBoost, AdaBoost	(Jiang et al., 2024)

Table 4.1 lists statistical and Machine Learning (ML) models used to map salinity. These include Linear Regression (LR), Geographically Weighted Regression (GWR), Spatially Weighted Optimization Model (SWOM), Multiple Linear Regression (MLR), DT, Random Forest (RF), Support Vector Regression (SVR), Artificial Neural Network (ANN), Multilayer Perceptron (MLP), Adaptive Neuro-Fuzzy Inference System (ANFIS), Adaptive Boosting (AdaBoost), Light Gradient Boosting Machine (LightGBM), and Extreme Gradient Boosting (XGBoost). LR and MLR are commonly used; however, a key limitation of these models is their inability to generalize across large spatial and temporal scales due to variability in atmospheric and water composition that affects the spectral signatures of WQPs, including CDOM/TSS (Topp et al., 2020). Specifically, riverine WQPs, including CDOM/TSS, vary within and between river catchments (Chang, 2008; Lintern et al., 2018), and a range of relationships exist between spectral signatures, CDOM/TSS, and salinity across different areas. This requires the development of site-specific empirical models and highlights salinity as a non-stationary variable (Xie et al., 2013) due to interactions between land and water, and most importantly saltwater intrusion into estuaries (Jiang et al., 2024). GWR (Xie et al., 2013) SWOM (Khadim et al., 2017) address this challenge by considering spatial dependencies in modeling salinity. While studies have identified geographical location as a crucial variable in some models (Khadim et al., 2017; Nguyen et al., 2018; Urquhart et al., 2012; Xie et al., 2013), such models often lack transferability to different locations.

ML models such as DT (Nguyen et al., 2018), SVR (Mohsen Ansari and Akhoondzadeh, 2020), and ensemble models like RF (Nguyen et al., 2018), AdaBoost (Borovskaya et al., 2022), LightGBM (Dai et al., 2023), and NN-based models like ANN (Bayati and Danesh-Yazdi, 2021), MLP (Mohsen Ansari and Akhoondzadeh, 2020), and ANFIS (Bayati and Danesh-Yazdi, 2021) typically outperform the traditional linear and non-linear statistical approaches for modeling salinity (Table 4.1).

Drawing from Tobler's First Law of Geography (Tobler, 1970), the evaluation of spatial autocorrelation in regression residuals is essential in spatial regression modeling (Feng et al., 2021;

Kelly, 2019b; Lu et al., 2023; Mabula et al., 2023; Mahboobi et al., 2023; Mainali et al., 2019; Moran, 1950; Peralta et al., 2016; Sotomayor et al., 2023). Firstly, spatial autocorrelation in residuals violates the independence assumption fundamental to regression analysis, leading to biased parameter estimates and incorrect inference. This undermines the reliability of regression results (Higazi et al., 2013). Secondly, detecting spatial patterns in residuals provides insights into the validity of the regression model by revealing whether it adequately captures all spatial dependencies in the data (Kelly, 2019b). The presence of spatial autocorrelation in residuals suggests potential unmodeled spatial processes influencing the response variable, pointing to areas for model refinement (Higazi et al., 2013; Kelly, 2019b; Khorram, 1985; Metulini et al., 2018; Xie et al., 2013). However, despite its significance, the investigation of spatial autocorrelation in regression residuals from statistical and ML models remains underexplored, including in salinity modeling studies.

This study collected *in situ* water salinity data from 2013 to 2018 for the Karun River, Iran's longest and highest-discharge river. Agricultural effluents and industrial sewage, often linked to sugarcane development projects (Nahvi et al., 2018), combine with salt dissolution from the Gachsaran Formation within the reservoir behind the Upper Gotvand Dam (Jalali et al., 2019) to contribute to the salinity of the river, generally resulting in increased salinity downstream. Using these data, we modeled the river's salinity with Landsat 8 OLI imagery. Landsat 8 was selected due to its free availability, spatial resolution suitable for our case study, unlike Aquarius, VIIRS, and MODIS, and its full operational status during our field data collection, unlike Sentinel-2, which became operational only after 2015. For this purpose, we leveraged 14 established spectral indices and 24 predictive models from previous studies. Due to the limited number of matchups (102 samples), developing complex novel models, such as those based on NNs, was impractical. Instead, our efforts were directed toward improving input features to achieve better results with this suite of existing models. We comprehensively evaluated the statistical performance metrics associated with each model. Moreover, we investigated the presence of global spatial autocorrelation in the

residuals of each model. Specifically, we addressed the following questions within the context of Landsat-based salinity modeling in the Karun River: i) Which predictive model performs best regarding statistical metrics and spatial autocorrelation of residuals? ii) What is the relationship between these two aspects of model performance? iii) How do the influences of the Gachsaran Formation versus other factors (e.g., sugar cane projects) contribute to downstream salinity increases?

4.2 Study area and data

4.2.1 Study area

The Karun River basin is located in southwestern Iran. It covers an area of ~67,000 km² (Shahraki et al., 2023) (Figure 4.1). It runs through 16 cities (e.g., Gotvand, Shushtar, Ahvaz, Khorramshahr, and Abadan are shown in green in Figure 4.1) and numerous villages and serves various industrial and agricultural purposes, irrigating over 380,000 hectares of the surrounding plains (Shahraki et al., 2023). The Karun River converges with the Arvand (or Shatt al-Arab) River at Khorramshahr before flowing into the Persian Gulf (Figure 4.1). Saline discharges originating from arid land, irrigation practices, fish hatcheries, and urban sources (Adib and Javdan, 2015; Nouraki et al., 2021) contribute to elevated salinity levels in the Karun River. Pollution in the Karun River primarily stems from agricultural and industrial waste, often linked to sugar cane projects (Nahvi et al., 2018). Moreover, the construction of the Upper Gotvand Dam (located near Gotvand station in Figure 4.1) has created a reservoir that includes the Gachsaran Formation, known for its deposits of gypsum and halite (rock salt) (Bahadori et al., 2011). Since its completion in 2011, the dam has facilitated prolonged contact between the Gachsaran Formation and reservoir water, increasing salinity in the reservoir (Jalali et al., 2019) and hence in the Karun River.

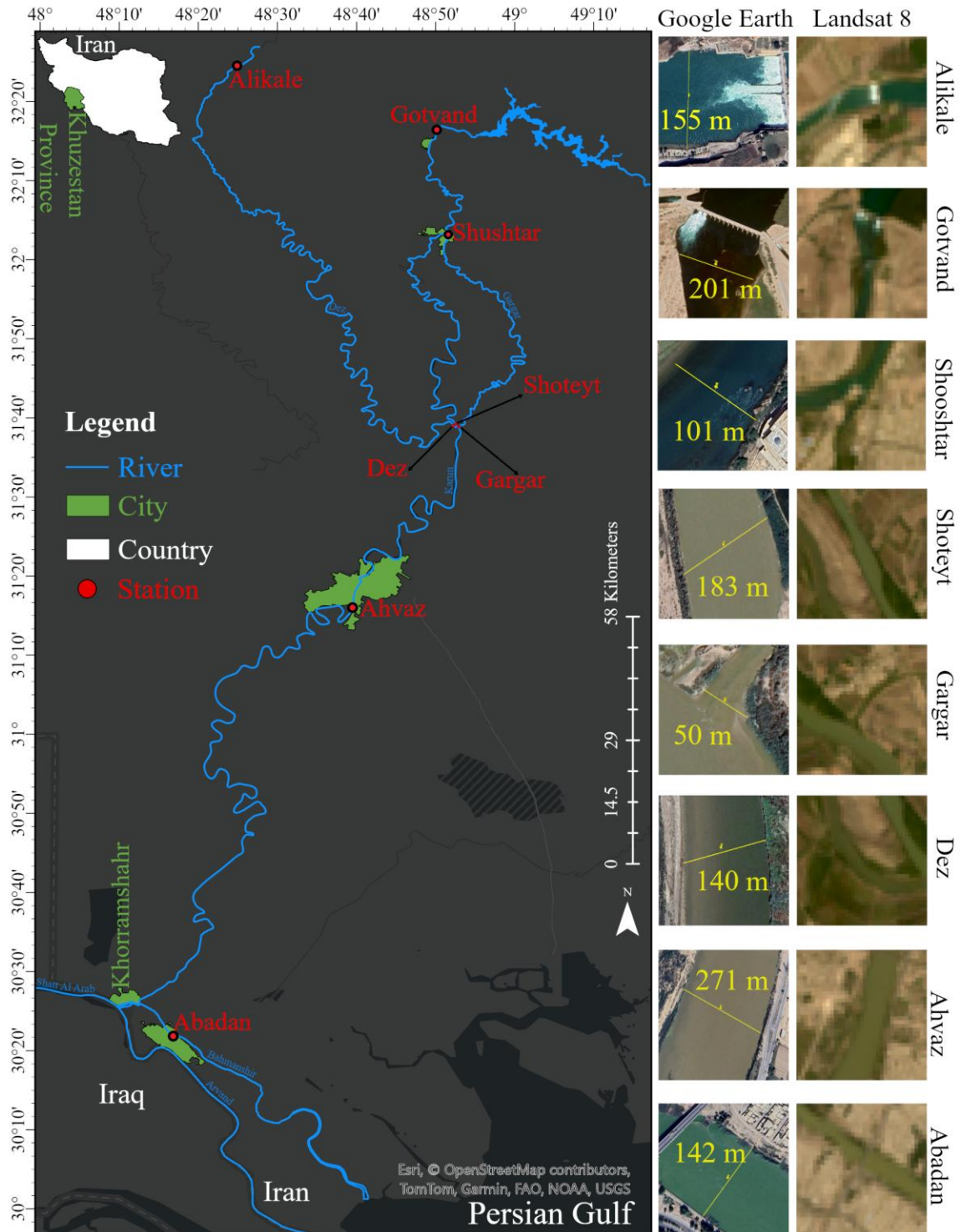


Figure 4.1. Study area (Karun River, Khuzestan, Iran). Overview of 8 salinity sampling stations (red) along the Karun River. Google Earth Pro and Landsat 8 OLI images of each station is shown on the right side of the map. The yellow line on each image indicates the river's width at the respective station.

4.2.2 Landsat-8 OLI images

We used 37 Landsat 8 OLI L1 images that are cloud-free at the 8 stations shown in Figure 4.1, for the path/row combinations 165/38, 165/39, and 166/38, covering the period from June 2013 to May 2022. With Landsat 8 OLI's 30-meter spatial resolution, and considering the width of the Karun River at the 8 sampling stations (> 30 m) (refer to Figure 4.1), the Karun River at these stations can be observed in at least one pixel on a Landsat 8 OLI image. All Landsat 8 OLI images were atmospherically corrected using the Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) module (Cooley et al., 2002), and land, defined as pixels with positive Normalized Difference Vegetation Index (NDVI) values, was masked out. Fourteen indices, previously used for modeling soil/water salinity or TSS, were calculated for each pixel. These indices were selected because recent studies have demonstrated that they correlate with soil/water salinity or TSS (Adjovu et al., 2023b, 2022; Aksoy et al., 2022). Although soil salinity differs from water salinity, indices related to soil salinity have been used in previous water salinity studies and have contributed to improving model performance (Adjovu et al., 2022; Ferdous and Rahman, 2020). While other salinity and TSS indices exist, this study focused on those listed by Adjovu et al.(2023b). The selected indices include Salinity indices 1-7, the Normalized Difference Salinity Index (NDSI), NSMI, NDSSI, the Green-Blue Band Ratio (BR), EGRI, the Water-Sediment Ratio Index (WSRI), and NDVI (Table 4.2).

Table 4.2. Salinity indices used in this study. ρ represents the surface reflectance.

indices	Formula	Reference
Salinity index 1	$\sqrt{\rho_{Blue} \times \rho_{Red}}$	(Khan et al., 2001)
Salinity index 2	$\sqrt{\rho_{Blue} \times \rho_{Green}}$	(Maliki et al., 2020)
Salinity index 3	$\sqrt{\rho_{Green}^2 + \rho_{Red}^2}$	(Douaoui et al., 2006)
Salinity index 4	$\sqrt{\rho_{Green}^2 + \rho_{Red}^2 + \rho_{NIR}^2}$	(Douaoui et al., 2006)
Salinity index 5	$\rho_{Red} \times \rho_{Green}$	(Maliki et al., 2020)
Salinity index 6	$\sqrt{\rho_{Green}^2 \times \rho_{NIR}^2}$	(Bouaziz et al., 2011)
Salinity index 7	ρ_{Blue}/ρ_{Green}	(Ferdous and Rahman, 2020)
NDSI	$(\rho_{Red} - \rho_{NIR})/(\rho_{Red} + \rho_{NIR})$	(Khan et al., 2005)

NSMI	$(\rho_{Red} + \rho_{Green} - \rho_{Blue})/(\rho_{Red} + \rho_{Green} + \rho_{Blue})$	(Montalvo, 2010)
NDSSI	$(\rho_{Blue} - \rho_{NIR})/(\rho_{Blue} + \rho_{NIR})$	(Hossain et al., 2010)
BR	ρ_{Green}/ρ_{Blue}	(Montalvo, 2010)
EGRI	$\rho_{Green}/(\rho_{Blue} + \rho_{Coastal\ Aerosol})$	(Malahlela, 2019)
NDVI	$(\rho_{NIR} - \rho_{Red})/(\rho_{NIR} + \rho_{Red})$	(Rouse et al., 1974)
WSRI	$1 - [(\rho_{SWIR1} - \rho_{Blue})/\rho_{Red}]$	(Malahlela, 2019)

4.2.3 In situ salinity data

In situ measurements conducted by the Iran Water and Power Resources Development Company (2023) were performed using a WTW - Portable conductivity meter ProfiLine Cond 3310, which is regularly calibrated to ensure accuracy. This device measures EC, which is temperature-dependent (Dahaan et al., 2016; Wagner et al., 2006a), and adjusts readings to 25 °C to report *specific conductance* in $\mu\text{S}/\text{cm}$ (U.S. Geological Survey, 2019b). Salinity was calculated in practical salinity units (psu) from specific conductance using equations from Schemel (2001) (see Equation C1 and Equation C2 in Appendix C); we did not correct for pressure because all measurements were conducted near the water surface (Wagner et al., 2006a).

In situ salinity data acquired within $\pm$ two days of a Landsat 8 OLI image (Table 4.3) were extracted from each station for 102 matchups. It is worth mentioning that although Sentinel-2 has better spatial and spectral resolution compared to Landsat 8 because it was non-operational before 2015 or due to a mismatch with the timing of the field data collection, its matchups would have been limited to only 21, which is insufficient for training machine learning models. As shown in Table 4.3, matchups were prepared for specific stations on each date, as cloud cover or the absence of field data on the target date limited availability. In 2013 and 2014, 31 and 30 matchups were prepared across various dates and stations, respectively, while in some years, such as 2018 and 2015, the number of matchups decreased to 1 and 9, respectively. Given the extensive length of the Karun River and the five-year timeframe covered by the data, field data collection was economically constrained.

Table 4.3. Date of satellite image acquisition and *in situ* data collection. Matchups for each date are available at the stations listed in the Station(s) column. For other stations, either field data were not collected, or the data were masked due to cloud cover.

Date of satellite image acquisition	Date of <i>in situ</i> data collection	Station(s)
19 June 2013	19 June 2013	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz
26 June 2013	25 June 2013	Alikale
12 July 2013	14 July 2013	Alikale
13 August 2013	13 August 2013	Alikale
22 August 2013	24 August 2013	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
23 September 2013	25 September 2013	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
25 October 2013	26 October 2013	Gotvand
28 December 2013	29 December 2013	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
02 March 2014	01 March 2014	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
03 April 2014	05 April 2014	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
26 April 2014	28 April 2014	Alikale
10 September 2014	08 September 2014	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
03 October 2014	04 October 2014	Alikale
15 December 2014	14 December 2014	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
01 February 2015	01 February 2015	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
18 July 2015	16 July 2015	Alikale
04 September 2015	03 September 2015	Alikale
24 April 2016	26 April 2016	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz
29 July 2016	28 July 2016	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
01 October 2016	29 September 2016	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz
21 January 2017	23 January 2017	Abadan
13 May 2017	14 May 2017	Gotvand, Shushtar, Shoteyt, Gargar, Dez, Ahvaz, Abadan
07 December 2017	05 December 2017	Gotvand, Gargar, Ahvaz, Abadan
10 July 2018	12 July 2018	Alikale

Table 4.4 provides summary statistics for the salinity observed at each of the eight stations. To evaluate the performance of the salinity models, the matchups were randomly split, allocating ~70% (N = 72) for model training and ~30% (N = 30) for model testing (Shahin et al., 2004). This single partition was held constant across all 24 regression models, ensuring that performance metrics and residual spatial autocorrelation statistics are directly comparable across models on an identical test set.

Table 4.4. Minimum, maximum, and average salinity values (psu) and number of observations (N) at the eight salinity measurement stations.

Station	Min.	Ave.	Max.	N
Alikale	0.19	0.22	0.28	10
Gotvand	0.38	0.57	0.76	15
Shushtar	0.42	0.66	0.85	11
Shoteyt	0.65	0.85	1.00	13
Gargar	1.13	1.56	1.93	14
Dez	0.36	1.11	1.64	13
Ahvaz	0.89	1.11	1.51	14
Abadan	1.28	1.50	1.90	12

4.3 Methodology

4.3.1 Feature Engineering

This study extracted and analyzed 21 features, including the Coastal Aerosol, visible, NIR, and SWIR spectral bands of Landsat 8 OLI and the 14 indices listed in Table 4.2. Three feature selection methods were employed to address potential overfitting and reduced interpretability from this large number of features (Bayati and Danesh-Yazdi, 2021; Jiang et al., 2024): RFFIS, Sobol' sensitivity analysis, and correlation matrix.

To calculate the RFFIS, we first trained an RF model (Breiman, 2001) using the training dataset, considering all 21 features. The RF's hyperparameters were fine-tuned using a Genetic Algorithm (GA) from the *PyGAD* Python library (Gad, 2023); implementation details are available in Table C2 in the supplementary material. The RF model constructs multiple regression trees from random training data samples and averages their predictions. Model accuracy is assessed using data not used in building the trees. Feature importance is determined by measuring the increase in prediction error when the values of a feature are permuted, keeping others unchanged (Abdel-Rahman et al., 2013; Menze et al., 2009). RFFIS scores, which range between 0 and 1 and sum to 1 across all features, are then used to identify the most important features for the regression process. Our implementation of RF was carried out in Python using the *Scikit-learn* library.

Sobol' is a global sensitivity analysis method that decomposes variance to quantify each feature's impact on the model output relative to others (Sobol', 1993). The problem formulation is $Y = f(X) = f(X_1, \dots, X_{21})$, where Y is the model output (i.e., salinity) and $X = (X_1, \dots, X_{21})$ is the

feature set (Verrelst and Rivera, 2017). An Ordinary Least Squares (OLS) model with all 21 features and all matchups ($N = 102$) was used to establish the problem formulation using the *statsmodels* library in Python. In the Sobol' method, the total unconditional variance of Y , $V(Y)$, is decomposed into $V_i, V_{ij}, V_{ijk}, \dots, V_{1\dots 21}$, representing contributions from individual factors X_i , interactions X_i and X_j , and higher-order interactions (Homma and Saltelli, 1996; Nossent et al., 2011). Variance is calculated using Monte Carlo integrals and Sobol' quasi-random sampling (Homma and Saltelli, 1996; Nossent et al., 2011), for which 10,000 samples of independent random variables, uniformly distributed within the range of each feature, are generated. The total-order sensitivity indices, $S_{T_i} = 1 - \frac{V_{\sim i}}{V}$ where $V_{\sim i}$ is obtained from variation of all parameters, except X_i , ranging from 0 to 1 and sum to 1 across all features, measure the main effect and interactions of X_i with other parameters, allowing for the ranking of feature importance (Nossent et al., 2011). Our implementation of Sobol' sensitivity analysis was conducted in Python using the *SALib* library. The correlation matrix was generated by calculating the coefficient of determination (R^2) values between all 21 features and salinity. The 21 features were then ranked based on their RFFIS, S_T values, and R^2 values with salinity, from 1 to 21, and the average of these three ranks was computed to determine the final rank. Subsequently, regression models were then trained using the feature with the best average rank, with its hyperparameters optimized by GA, and its R^2 was calculated using the test dataset. This process was repeated, adding one top-ranked feature until all features were included. Finally, each model's features that resulted in the highest R^2 were selected.

4.3.2 Regression Methods

We tested 24 different regression methods for salinity modeling, including statistical, kernel-based, DT-based, and NN-based models. To ensure reproducibility, a fixed random seed of 42 was applied to all stochastic components of model training and hyperparameter optimization across all 24 models. The hyperparameters of all methods, as detailed in Section S3 of the Supplementary Material, were tuned using a GA to ensure optimal performance. All regression models were

implemented using Python. We utilized DT, GBDT, Gaussian Process Regression (GPR), Kernel Ridge Regression (KRR), MLP, Partial Least Squares Regression (PLSR), RF, Ridge Linear Regressor (RLR), and SVR from the *Scikit-learn* library (Buitinck et al., 2013), One-Dimensional Convolutional Neural Network (1D-CNN), ANN, Gated Recurrent Units (GRU), Long Short-Term Memory (LSTM), and Simple Recurrent Neural Network (RNN) from the *Keras* library (Chollet, 2015), OLS and Generalized Linear Model (GLM) from the *statsmodels* library (Seabold and Perktold, 2010), Spatial Lag Model (SLM) and Spatial Error Model (SEM) from the *sperg* library, GWR from the *MGWR* library (Oshan et al., 2019), Generalized Additive Model (GAM) from the *pyGAM* library (Servén et al., 2018), Cubist from the *Cubist* library (Kuhn et al., 2023), LightGBM from *LightGBM* library (Ke et al., 2017), XGBoost from *XGBoost* library (Chen and Guestrin, 2016), and ANFIS from *Scikit-Fuzzy* (Warner et al., 2019).

4.3.2.1 Statistical Models

OLS assumes that the errors are independent and identically distributed, which may not be the case, especially for spatial data (Xie et al., 2013). To address this limitation, spatial regression methods such as the SLM and SEM (Anselin, 2001) account for spatial autocorrelation by incorporating spatially lagged variables or allowing for spatially correlated error terms (Mainali et al., 2019), respectively. A weight matrix is required to implement the SLM and SEM, reflecting each observation's influence on its neighbors. We calculate the weight matrix using the data's spatial coordinates, with closer observations receiving higher weights. Furthermore, the GWR method (Brunsdon et al., 1998) offers localized modeling capabilities by incorporating spatially varying coefficients based on spatial coordinates. A common challenge in linear modeling is multicollinearity, i.e., highly correlated predictor variables. Models like PLSR (Geladi and Kowalski, 1986) and RLR (Hoerl and Kennard, 1970) address multicollinearity by either performing dimensionality reduction or incorporating a regularization term into the OLS objective function. GAM (Hastie, 1992) and GLM (Gotway and Stroup, 1997) extend the traditional linear

modeling framework by accommodating non-linear relationships between predictors and response variables.

4.3.2.2 Kernel-Based Models

A key component of kernel-based models is the *kernel function*, which computes the similarity between pairs of data points in the original input space. We test the Radial Basis Function (RBF), linear, sigmoid, and second-degree polynomial kernels for all kernel-based methods, selecting the one that produces the best results.

GPR (Williams and Rasmussen, 1995) adopts a non-parametric Bayesian approach to model the relationship between input variables and outputs as a distribution over functions. KRR (Vovk, 2013) is a regularized linear regression version, employing kernel functions to map input variables into high-dimensional feature spaces. SVR (Smola and Schölkopf, 2004) builds on support vector machines and seeks a hyperplane that maximizes the margin around the predicted values.

4.3.2.3 Decision Tree-Based Models

DT-based models recursively split the data into subsets based on features, resulting in a tree structure. Ensemble learning models like RF (Breiman, 2001), XGBoost (Chen and Guestrin, 2016), LightGBM (Ke et al., 2017), and GBDT (Friedman, 2001) try to increase predictive accuracy by using the collective outputs of multiple DTs. Cubist (Quinlan, 1992) builds a series of regression trees to partition the data into smaller subsets similar to the other tree-based models. Within each subset, it fits a linear model, allowing for the capture of local trends.

4.3.2.4 Neural Network-Based Models

Regarding NN-based models, the Adam optimizer was used universally, and the ReLU activation function was applied to the ANN, 1D-CNN, and MLP models, while the linear activation function was used for the GRU, LSTM, and SimpleRNN models, as determined through trial and error. The MLP (Atkinson and Tatnall, 1997) forms the foundation of NNs, consisting of Fully Connected

(FC) layers where each neuron in a layer is connected to every neuron in the subsequent layer. The ANN (Paoletti et al., 2019) builds upon the MLP structure but typically incorporates more layers, allowing them to learn increasingly complex data representations. The ANFIS (Jang, 1993) is a fuzzy logic-based ML model that combines fuzzy logic with an NN's structure. The 1D-CNN (Albawi et al., 2017) consists of one-dimensional convolutional layers, enabling them to learn hierarchical features automatically. The RNN (Elman, 1990), LSTM (Hochreiter and Schmidhuber, 1997), and GRU (Cho et al., 2014) models use feedback loops to incorporate previous data as they process new input.

4.3.3 Model Evaluation

Model performance was assessed using i) evaluation metrics and box plots (range, quartiles, and median) of model estimates and observations and ii) examination of residual spatial autocorrelation using Moran's I. The model with the best overall performance in evaluations (i) and (ii) was identified, retrained using all matchups ($N = 102$), and applied to Landsat 8 OLI images on 19 June 2013; 01 February 2015; 05 September 2018; and 11 May 2022. These dates were selected because there was no cloud cover in the Landsat 8 OLI images of the study area, *in situ* data were available on 19 June 2013 and 01 February 2015, at salinity sampling stations (Figure 4.1) for validation, to allow comparison with salinity maps from previous studies (Ansari and Akhoondzadeh, 2021; Mohsen Ansari and Akhoondzadeh, 2020), and to evaluate the model across different years and seasons under varying weather conditions. A salinity map was generated in ESRI ArcGIS Pro 3.1.0 for each date. The average salinity within a 1 km buffer around five cities, shown in green in Figure 4.1, was calculated. The cities were selected because Gotvand and Abadan are located far upstream and downstream, respectively, and the areas around Shushtar, Ahvaz, and Khorramshahr have sugarcane cultivation that may influence salinity in the Karun River (Moradi-Majd et al., 2022).

4.3.3.1 Evaluation Metrics

6 statistical metrics were employed as evaluation metrics: linear regression slope, R^2 , Root Mean Square Error (RMSE), Root Mean Squared Logarithmic Error (RMSLE), Median Symmetric Accuracy (ϵ) (Morley et al., 2018), and Symmetric Signed Percentage Bias (β) (Morley et al., 2018) (Equation 4.1-5):

$$R^2 = \left[\frac{\sum_{i=1}^N (E_i - \bar{E})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^N (E_i - \bar{E})^2} \sqrt{\sum_{i=1}^N (O_i - \bar{O})^2}} \right]^2 \quad (4.1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_i - O_i)^2} \quad (4.2)$$

$$RMSLE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\log E_i - \log O_i)^2} \quad (4.3)$$

$$\beta = 100 \times \text{sign} \left[\text{median} \left(\log \frac{E_i}{O_i} \right) \right] \times \left(e^{|\text{median}(\log \frac{E_i}{O_i})|} - 1 \right) \quad (4.4)$$

$$\epsilon = 100 \times \left(e^{\text{median} \left[\left| \log \left(\frac{E_i}{O_i} \right) \right| \right]} - 1 \right) \quad (4.5)$$

N is the number of observations, and O , E , $\bar{O}$, and $\bar{E}$ are the observations, estimated values, the average, and the average of the estimated values, respectively.

R^2 , a metric ranging from 0 to 1, quantifies the proportion of variance explained by the model, with 1 indicating a perfect fit and values near 0 suggesting no correlation, as frequently used in previous studies in Table 4.1. RMSE measures the average magnitude of errors between predicted and actual values, with lower values indicating better performance. RMSLD, unlike RMSE, considers the logarithmic difference between predicted and actual values, which can be beneficial when dealing with non-normally distributed errors or outliers as it scales down the impact of outliers (Morley et al., 2018). Additionally, ϵ represents a symmetric percentage error, penalizing over- and under-

estimation equally. Lower values indicate superior performance, with perfect accuracy assigned a value of 0%. Similarly, β , akin to ϵ , represents a percentage bias that maintains symmetry between over- and under-estimation, with values closer to zero indicating better performance. Previous studies (Pahlevan et al., 2022; Smith et al., 2021), have demonstrated that median measures, linear regression slopes, and RMSLD collectively provide a comprehensive framework for evaluating models.

4.3.3.2 Residual Spatial Autocorrelation

Spatial autocorrelation (Tobler, 1970) occurs when neighboring data points influence each other or are influenced by the same spatially variable process, potentially leading to bias in predictions and a violation of the assumption of independence in standard statistical procedures (Mainali et al., 2019). Autocorrelated model residuals, stemming from this spatial dependence, can lead to overly optimistic conclusions about model performance (Kelly, 2019b). Ensuring the validity of regression results, therefore, entails assessing whether residuals are independent and identically distributed (Anselin et al., 2004; Lu et al., 2023; Mabula et al., 2023; Mahboobi et al., 2023; Sotomayor et al., 2023). Moran's I (Moran, 1950) is commonly employed to analyze residual spatial autocorrelation (Mainali et al., 2019). Interpreting Moran's I involves calculating the associated p-value and Z-Score under the null hypothesis of no spatial autocorrelation. A non-significant p-value indicates that the residuals are spatially independent, meaning the assumption of independence underlying regression analysis is upheld.

Conversely, a significant p-value indicates significant spatial patterns in the residuals. Positive Z-Scores indicate positive spatial autocorrelation of high and/or low values, while negative Z-Scores indicate negative spatial autocorrelation. Moran's I is calculated as follows:

$$\text{Moran's I} = \frac{N \sum_{i=1}^N \sum_{j=1}^N w_{ij} (r_i - \bar{r})(r_j - \bar{r})}{\sum_{j=1}^N (r_j - \bar{r})^2 (\sum_{i=1}^N \sum_{j=1}^N w_{ij})} \quad (4.6)$$

where r represents the residual values at locations i and j , $\bar{r}$ denotes the mean of r , and w_{ij} is a matrix of spatial weights.

To calculate Moran's I, we created a weight matrix using the inverse of Haversine distance, giving higher values to closer neighbors. The weight matrix was row-standardized, and self-weights on the diagonal were set to zero to ensure that each observation is not treated as its own neighbor. Moran's I was computed on the test-set residuals ($N = 30$) for each of the 24 models. Significance was assessed using permutation pseudo p-values derived from 1,000,000 random permutations of residuals across the fixed sampling locations, under the null hypothesis of no spatial autocorrelation (i.e., random assignment of residual values to locations). To account for multiple testing across 24 models, Benjamini-Hochberg false discovery rate (FDR) correction was applied to the full set of permutation p-values. FDR-corrected p-values are reported as the primary criterion for significance. The reference distribution for each model, used to evaluate statistical significance, is provided in Figure C1 in the Appendix C. All spatial analyses were implemented using Python's *PySAL* library (Rey and Anselin, 2010).

4.4 Results

4.4.1 Feature Selection

Figure 4.2 illustrates the correlation matrix for 21 features and salinity measurements. Table 4.5 provides the values and ranks of RFFIS, S_{Ti} , and the correlation with salinity for Landsat 8 OLI spectral bands and indices from Table 4.2, sorted by mean rank. Details of the OLS model used in Sobol' analysis and the tuned hyperparameters obtained through GA for the RF model are presented in Table C1. According to Table 4.5, the Red band is identified as the most sensitive feature to salinity.

Table 4.5. Results of feature engineering.

No.	Feature	S_T	S_T Rank	RFFIS	RFFIS Rank	Correlation	Correlation Rank	Mean Rank
1	Red	2.77E-01	1	9.20E-02	3	1.79E-01	4	2.7
2	Salinity index 3	2.66E-01	2	6.33E-02	6	1.82E-01	2	3.3
3	Salinity index 4	1.86E-01	3	7.13E-02	4	1.43E-01	5	4.0
4	NSMI	1.85E-03	10	1.38E-01	1	2.36E-01	1	4.0
5	Salinity index 6	8.82E-02	5	4.59E-02	8	1.17E-01	7	6.7
6	Salinity index 2	1.93E-03	9	5.17E-02	7	8.85E-02	9	8.3
7	Green	3.79E-03	8	3.17E-02	14	1.80E-01	3	8.3
8	Salinity index 5	2.64E-04	15	1.20E-01	2	8.24E-02	10	9.0
9	EGRI	7.13E-04	12	6.75E-02	5	3.15E-02	14	10.3
10	Salinity index 7	1.31E-01	4	9.98E-03	21	1.37E-01	6	10.3
11	Salinity index 1	1.53E-02	7	2.23E-02	17	1.07E-01	8	10.7
12	Blue	2.44E-02	6	1.30E-02	19	3.38E-02	13	12.7
13	SWIR 1	4.26E-04	14	3.83E-02	11	5.89E-03	18	14.3
14	SWIR 2	1.86E-04	16	4.31E-02	9	2.67E-03	20	15.0
15	NDVI	1.20E-05	19	3.09E-02	15	3.71E-02	11	15.0
16	BR	6.60E-04	13	3.23E-02	12	1.28E-03	21	15.3
17	NDSSI	1.35E-04	17	4.01E-02	10	3.50E-03	19	15.3
18	NDSI	1.20E-05	19	2.52E-02	16	3.71E-02	11	15.3
19	WSRI	1.03E-04	18	3.17E-02	13	2.40E-02	16	15.7
20	Coastal Aerosol	7.36E-04	11	1.24E-02	20	1.66E-02	17	16.0
21	NIR	1.00E-06	21	1.91E-02	18	2.94E-02	15	18.0

Chapter 4: Statistical Performance vs Residual Spatial Autocorrelation

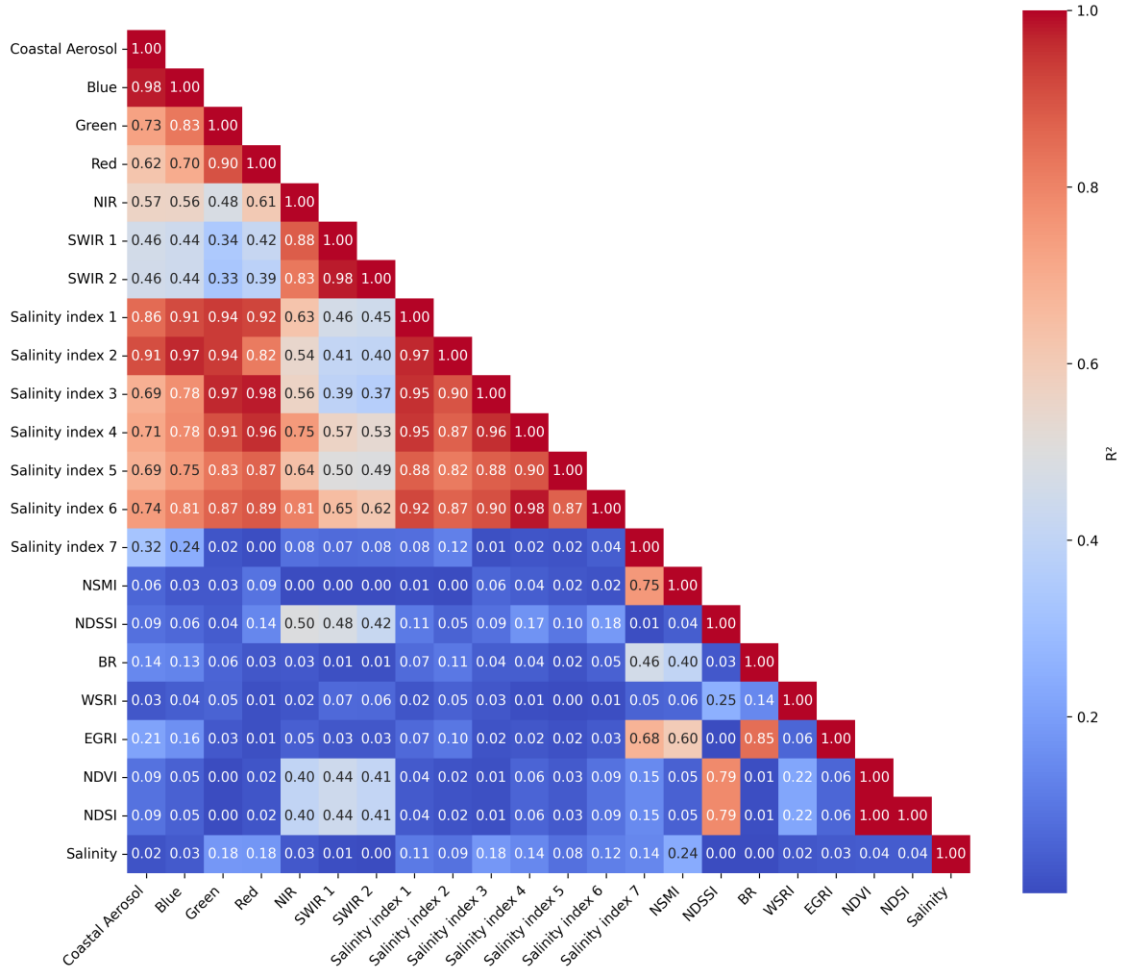


Figure 4.2. Correlation of features with salinity measurements and within each other.

Consequently, all 24 regression models were trained using the Red band, with hyperparameters optimized via GA and R^2 values calculated on the test dataset. This process was repeated iteratively, adding the next top-ranked feature until all features were included. The R^2 values at each step for each model are shown in Figure 4.3. Ultimately, the bands yielding the highest R^2 for each model were selected. Table C3 in the Appendix C lists the fine-tuned hyperparameters for each model using the best feature combination.

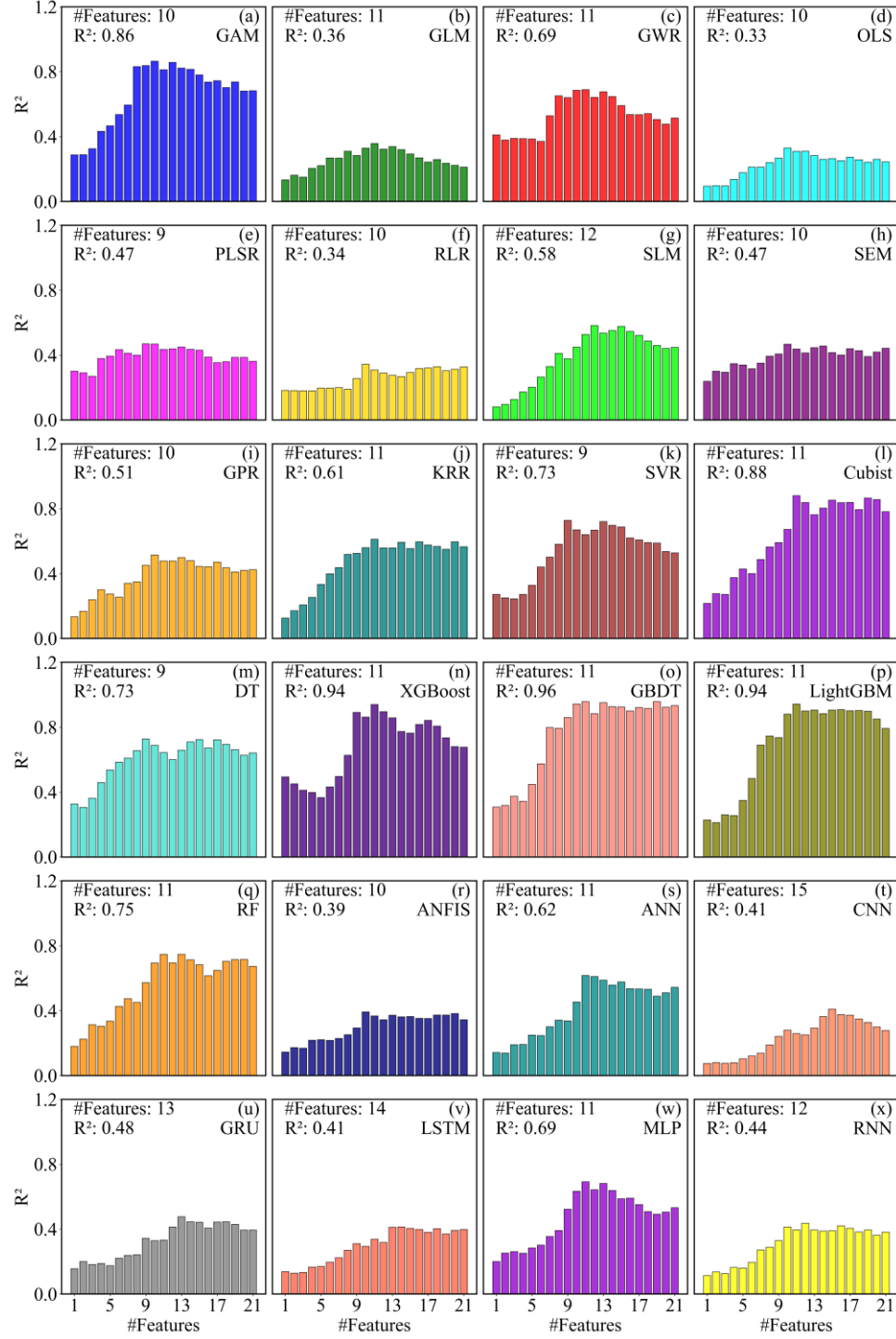


Figure 4.3. R^2 values for each model using different feature combinations. The #Feature refers to the feature number in Table 4.5, indicating that the feature with this number, along with those with lower numbers, yields the highest R^2 , as shown in the top-left corner of each panel.

4.4.2 Performance Assessment

The RBF kernel achieved better R^2 than other kernels across all models. Hence, we use the RBF kernel for all kernel-based models (see Table C3 in the Appendix C). Figure 4.4 shows that the Pearson correlation coefficients for all methods were statistically significant. Cubist, XGBoost, and GBDT had the lowest β values ($\sim 0\%$), with GBDT also having the lowest RMSE (0.10 psu) and ϵ (4.62%) and the highest R^2 (0.96). GWR had the slope closest to 1 (1.03), while XGBoost, GBDT, and LightGBM had the lowest RMSLE (0.11).

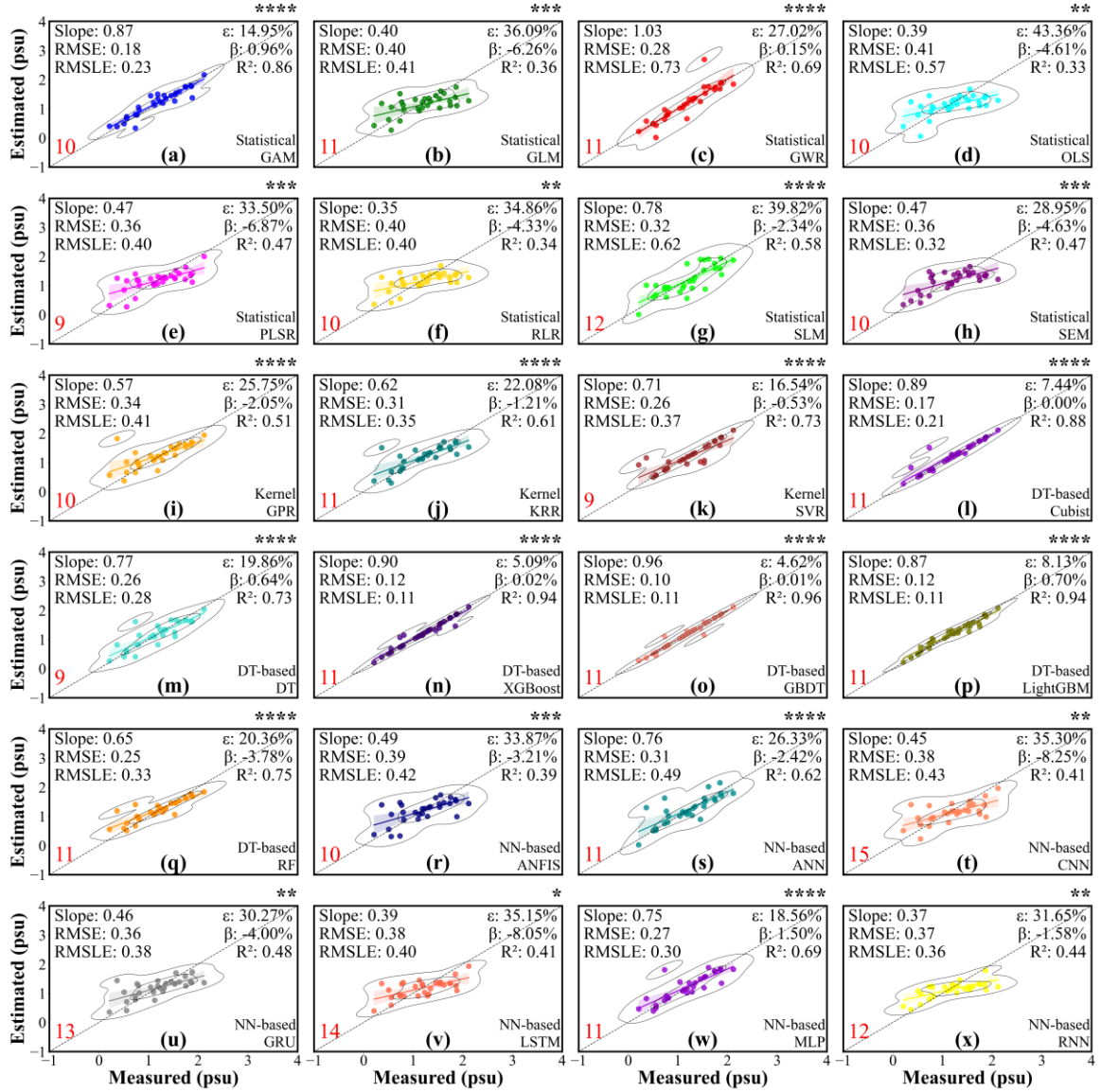


Figure 4.4. Scatter plot illustrates the relationship between measured and estimated salinity using test data ($N=30$). Each point represents a model prediction, with the name and class of each model

displayed at the bottom right corner. The top right and left corners display the statistical metrics: R^2 , RMSE, RMSLE, β , and ϵ . Asterisks denote Pearson correlations: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. The bottom left corner shows the number of selected features in red. A dashed line represents the 1:1 diagonal line for reference, while a solid-colored line depicts the fitted regression line. Contour lines, derived from kernel density estimates, show data distributions.

Figure 4.5's box plot reveals that the GBDT, XGBoost, and Cubist models' predictions are closest to the observations based on their median, quartiles, and range.

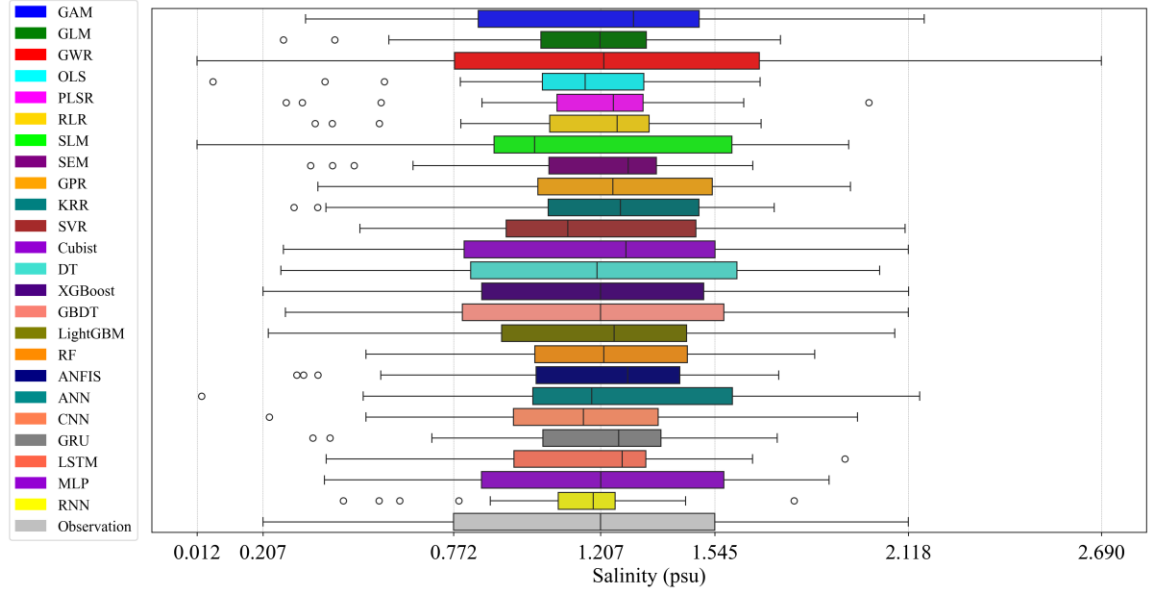


Figure 4.5. Box plot depicting the distribution of test data ($N = 30$). Each box displays the median and interquartile range, while the whiskers represent values within 1.5 times the interquartile range from the lower and upper quartiles. White circles indicate outliers beyond this range.

4.4.3 Spatial Analysis

Figure 4.6 presents the Moran scatterplot for each of the 24 models, showing standardized residuals on the x-axis and their spatially lagged values on the y-axis. The slope of the OLS fit line in each panel equals the global Moran's I for that model. Based on raw permutation p-values, 14 out of 24 models showed significant residual spatial autocorrelation at $\alpha = 0.05$. After applying Benjamini-Hochberg FDR correction to account for multiple testing, 11 models retained significance: GAM, GLM, OLS, PLSR, RLR, SEM, RF, CNN, GRU, LSTM, and RNN.

The remaining 13 models showed non-significant residual spatial autocorrelation after FDR correction, indicating spatially independent residuals: GWR ($I = -0.083$, $p_{pr\beta} = 0.403$), SLM ($I = -0.045$, $p_{pr\beta} = 0.486$), GPR ($I = -0.013$, $p_{pr\beta} = 0.424$), SVR ($I = -0.019$, $p_{pr\beta} = 0.423$), DT ($I = -0.080$, $p_{pr\beta} = 0.423$), Cubist ($I = 0.061$, $p_{pr\beta} = 0.232$), XGBoost ($I = 0.169$, $p_{pr\beta} = 0.091$), GBDT ($I = 0.169$, $p_{pr\beta} = 0.091$), LightGBM ($I = 0.169$, $p_{pr\beta} = 0.091$), RF ($I = 0.169$, $p_{pr\beta} = 0.091$), ANFIS ($I = 0.169$, $p_{pr\beta} = 0.091$), ANN ($I = 0.169$, $p_{pr\beta} = 0.091$), and CNN ($I = 0.169$, $p_{pr\beta} = 0.091$).

$= 0.011$, $p_{pr\beta} = 0.403$), LightGBM ($I = 0.209$, $p_{pr\beta} = 0.091$), ANFIS ($I = 0.206$, $p_{pr\beta} = 0.091$), ANN ($I = 0.049$, $p_{pr\beta} = 0.311$), MLP ($I = 0.117$, $p_{pr\beta} = 0.171$), and KRR ($I = 0.206$, $p_{pr\beta} = 0.088$).

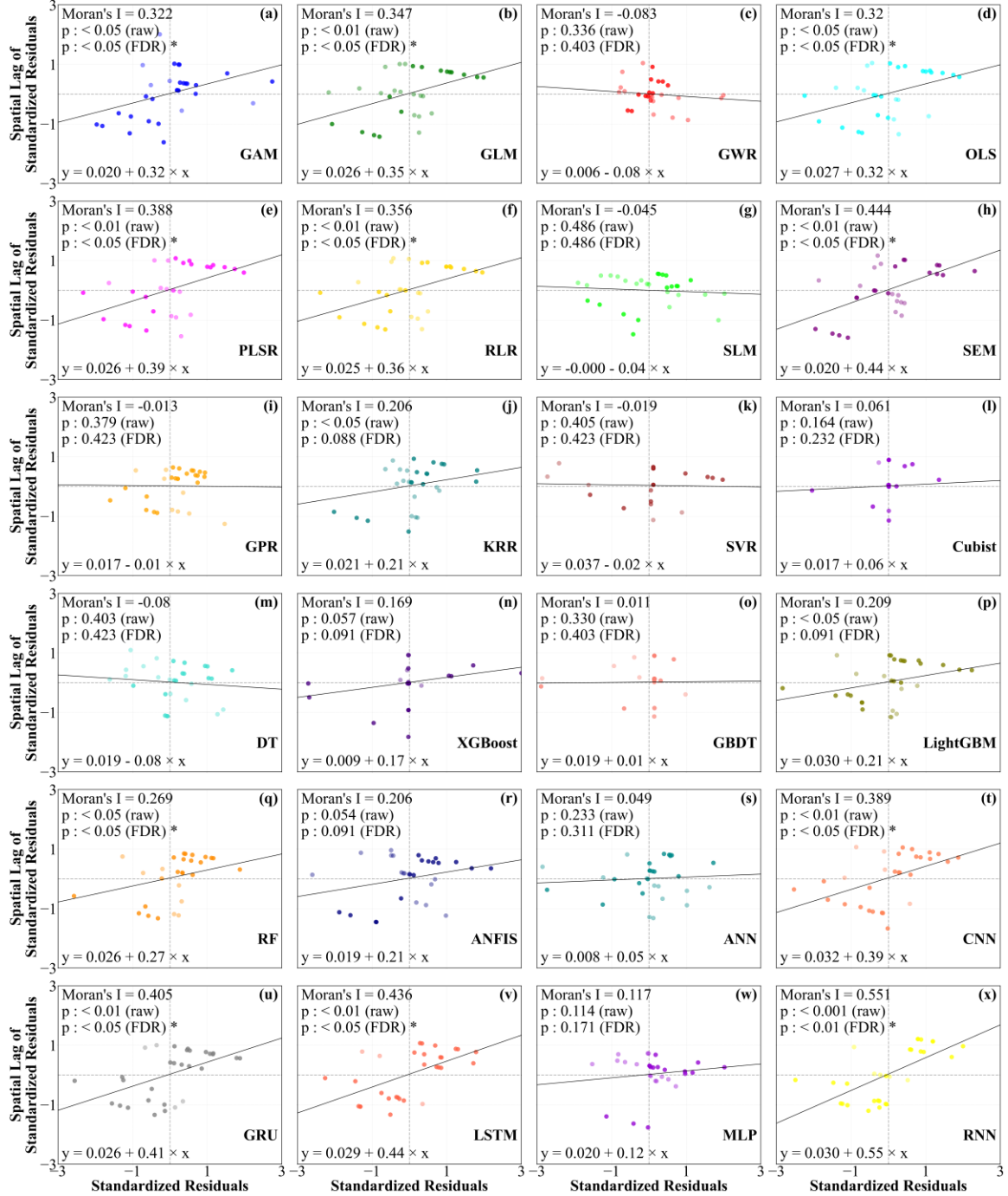


Figure 4.6. Moran scatterplot for each of the 24 regression models, showing standardized residuals on the x-axis and their spatially lagged values on the y-axis. The slope of the OLS fit line (solid black line) equals the global Moran's I for each model. The global Moran's I value, the raw permutation p-value (p_r), and the Benjamini-Hochberg FDR-corrected p-value ($p_{pr\beta}$) are reported in each panel. An asterisk (*) indicates significance at $\alpha = 0.05$ after FDR correction. Model names are presented at the bottom right.

4.4.4 Salinity Maps

Figure 4.7 shows salinity maps of the Karun River generated by GBDT. The GBDT was selected because, as discussed in Section 4.4.2, Cubist, XGBoost, and GBDT generally perform better than other models in terms of β , ϵ , slope, R^2 , error, and box plot metrics (predictions' median, quartiles, and range). XGBoost and GBDT specifically outperform Cubist in ϵ , slope, R^2 , error, and predictions' median. Although both XGBoost and GBDT showed non-significant residual spatial autocorrelation after FDR correction, GBDT exhibited a lower Moran's I (0.011 vs 0.169) and a higher FDR-corrected p-value (0.403 vs 0.091), indicating weaker spatial structure in its residuals. We therefore chose GBDT to generate the salinity map. This model, trained on the entire dataset (N=102) and compared to the data from each station, achieved an RMSE of 0.11, ranging from 0.06 at Ahvaz to 0.19 at Abadan (Table 4.6). Figure 4.7(a, b) show the estimated and measured salinity at stations with available data. Table 4.7 presents the average salinity at 1 km buffers around five cities in the Karun River basin [Figure 4.7(c)]. Table 4.7 shows salinity increases from Gotvand to Ahvaz city on all four dates.

Table 4.6. Number of observations (N) and RMSE for each station between measured and estimated salinity using GBDT trained on all dataset (N = 102).

Station	RMSE (psu)	N
Alikale	0.10	8
Gotvand	0.08	15
Shushtar	0.10	13
Shoteyt	0.13	13
Gargar	0.11	14
Dez	0.11	13
Ahvaz	0.06	14
Abadan	0.19	12

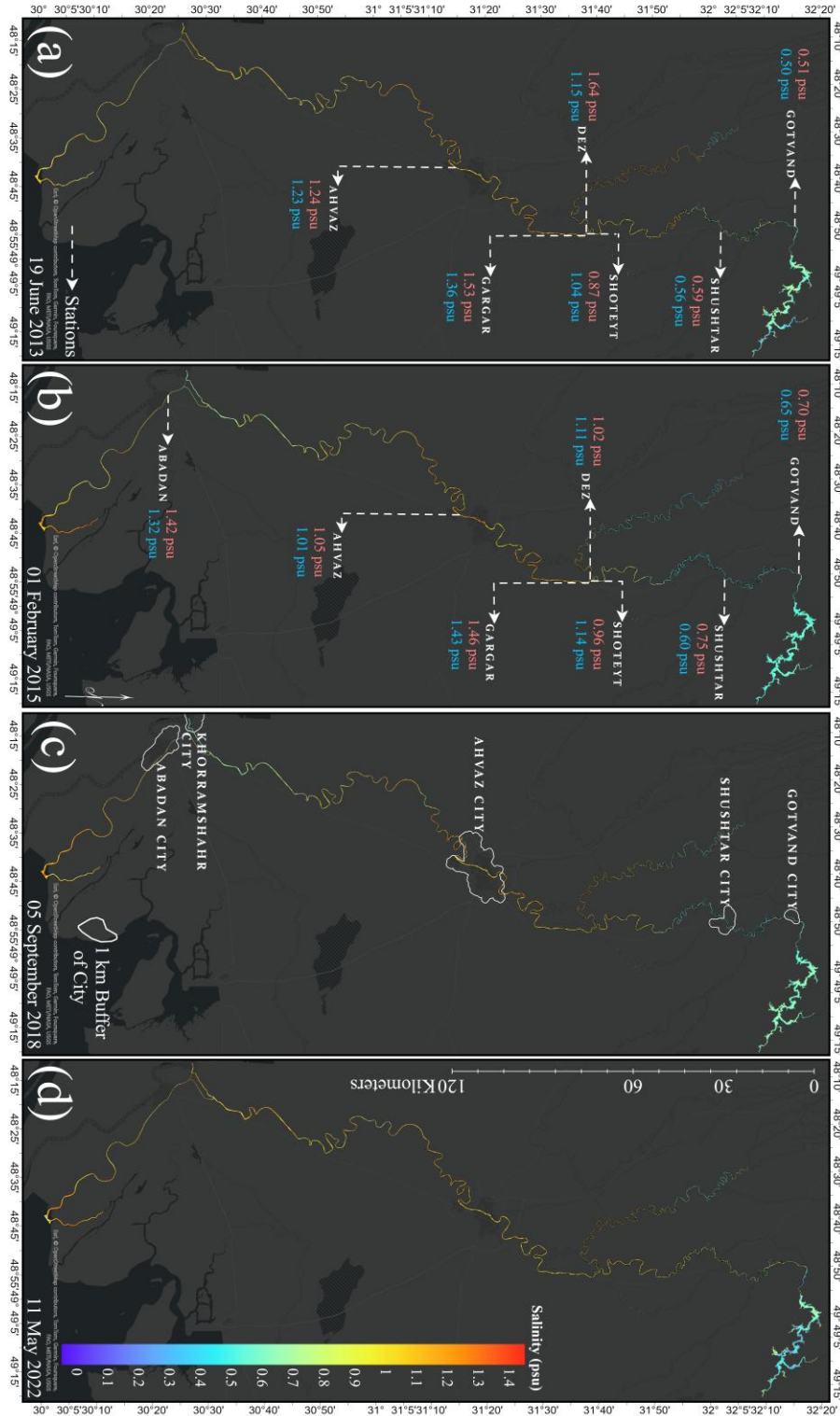


Figure 4.7. Karun River salinity maps generated by the GBDT model using the entire dataset (N=102) on (a) 13 June 2013, (b) 01 February 2015, (c) 05 September 2018, and (d) 11 May 2022. Measured salinity at stations with available data is shown in red, and estimated salinity is shown in blue.

Table 4.7. Average salinity (psu) within a 1 km buffer around 5 cities in the Karun River basin, derived from salinity maps generated by the GBDT model on 4 different dates.

Date	City				
	Gotvand	Shushtar	Ahvaz	Khorramshahr	Abadan
19 June 2013	0.58	0.79	1.07	1.05	0.72
01 February 2015	0.47	0.51	1.08	0.69	0.97
05 September 2018	0.56	0.54	1.06	0.61	0.96
11 May 2022	0.79	0.91	1.06	1.13	0.93

4.5 Discussion

4.5.1 Key Indices and Band Combinations for Salinity Mapping

As shown in Table 4.5, three different feature selection approaches ranked the features differently: Sobol' by variance analysis, RFFIS by permutation, and the correlation matrix by correlation with salinity. For example, NSMI is ranked first in both RFFIS and correlation but 10th in terms of S_T . Similarly, Salinity Index 5 is ranked 15th and 10th for S_T and correlation with salinity, respectively, but 2nd in RFFIS. Therefore, sorting the features by their mean rank allows for identifying those most sensitive to salinity, considering the different criteria used by the three feature selection methods. As shown in Fig. 3, all models ultimately used features ranked 1 to 9, indicating that these are the most sensitive features. Among these nine features, NSMI and EGRI have relatively low correlation with the others, as illustrated in Figure 4.2.

Most of the indices in Table 4.2 rely on visible spectral bands, which respond to variations in CDOM absorption and TSS (Mishra et al., 2017). As CDOM moves through an estuary, its photochemical reactivity and optical properties may change based on salinity due to changes in CDOM conformation, iron CDOM photochemistry, or both (Minor et al., 2006). This might explain why the Red band was the most sensitive feature in our study (Table 4.5) and previous studies (Khorram, 1985; Mushtaq and Nee Lala, 2017). The NSMI index, also selected, estimates TSS (Hossain et al., 2010; Malahlela, 2019), which is inversely correlated with salinity. Our study also used salinity indices 2-6, which highly correlate with each other (Figure 4.2), consistent with previous studies (Adjovu et al., 2022; Chatterjee and Bhadra, 2023; Liu et al., 2023), indicating that the NIR band is valuable for salinity mapping, especially in turbid waters, where longer

wavelengths are more sensitive and saturate at higher TSS concentrations than shorter wavelengths (Mishra et al., 2017).

Previous studies of the Karun River using the same dataset (Ansari and Akhoondzadeh, 2021; Mohsen Ansari and Akhoondzadeh, 2020) used the Coastal Aerosol, Blue, and Green bands, but this study selected the EGRI index, which combines these bands. Therefore, it is recommended that future research consider both spectral bands and indices followed by feature selection to determine the best combination, as band combinations (i.e., indices) might have higher feature importance than individual bands.

4.5.2 Performance Evaluation across Regression Models: Accuracy and Error Analysis

Regarding the statistical models, GAM showed good performance with higher R^2 and nearly half the RMSE and RMSLD compared to other statistical models. However, the linear regression fit for predicted versus observed salinity relationships showed a closer-to-unity slope for GWR (1.03), indicating superior fit across the entire range [Figure 4.4 (c)] compared to other models. Scatterplots in Figure 4.4 show a trend where, except for GAM, SLM, and GWR, other statistical models tended to overestimate salinity for values below one psu and underestimate it for values above 1.5 psu, consistent with their respective β values ranging from -6.87% (PLSR) to -4.33% (RLR).

While the performance of all kernel-based models was similar, SVR slightly outperformed the others across all six metrics. All three kernel-based models outperformed 5 out of 8 statistical models implemented in the study (except for GAM, GWR, and SLM). Notably, despite KRR being a linear model leveraging kernel methods, it improved R^2 by ~50% on average over statistical models (except for GAM and GWR), demonstrating the effectiveness of kernel methods in practice. The NN-based models exhibited weaker performance than other regression model types, characterized by large negative β values (on average -3.72%), indicating underfitting likely due to the scarce training dataset ($N = 72$). Reported slopes (on average 0.42) indicate poor performance

of these models, particularly evident at the tails of data distributions (salinity < 0.5 psu and salinity > 1.5 psu). Due to the limited training dataset, which was insufficient for optimal performance of NN-based models, certain outcomes deviated from theoretical expectations. For instance, while LSTM is designed to address the gradient anomaly issue in RNN effectively and is theoretically expected to outperform RNN, it performed slightly worse than RNN. Other models (e.g., DT-based models) were also anticipated to yield better results, as the available data was not ideal for effective NN training. It is worth noting that NN-based models like ANN performed better in a previous study (Dai et al., 2023), where a larger dataset of 2,000 matchups was available; their performance relative to other model types might also be improved for the Karun River with a larger dataset. Among the NN-based models implemented, MLP outperformed the others across all six metrics. Overall, DT-based models outperformed the other model classes. XGBoost and GBDT performed similarly, with negligible β , low ε ($< 5.1\%$), slope values close to unity (0.9), R^2 values > 0.93 , and low error values (RMSE < 0.13 psu, and RMSLD = 0.11). Scatterplots of XGBoost and GBDT in Figure 4.4(n, o) indicate a good fit across the entire value range. LightGBM also had good accuracy with low error, though its median and range of predictions differed more from the observations than XGBoost and GBDT (Figure 4.5). Interestingly, although XGBoost and GBDT improved R^2 and nearly halved RMSE compared to Cubist, Cubist had the lowest possible bias ($\beta = 0.00\%$), with its boxplot range and quartiles closely matching the observation boxplot (Figure 4.5). Overall, XGBoost, GBDT, and Cubist outperformed other regression models in terms of statistical metrics.

4.5.3 Residual Spatial Autocorrelation Analysis

Figure 4.6 presents the Moran scatterplot for each model. Among the statistical models, GWR and SLM showed non-significant residual spatial autocorrelation after FDR correction (GWR: $I = -0.083$, $p_{pr\beta} = 0.403$; SLM: $I = -0.045$, $p_{pr\beta} = 0.486$), which was expected given that both models explicitly incorporate spatial location in their structure. In contrast, GAM showed good accuracy but significant residual autocorrelation after FDR correction. Among kernel-based models, KRR

showed marginally non-significant autocorrelation ($I = 0.206$, $p_{pr\beta} = 0.088$), while GPR and SVR produced non-significant results ($p_{pr\beta} > 0.40$). Regarding neural network models, most showed significant autocorrelation after FDR correction suggesting that these models are spatially misspecified despite their predictive accuracy.

The performance of LightGBM is notable, as shown in Figure 4.4 (p), with good results in terms of accuracy ($R^2 = 0.94$ and $\varepsilon = 8.13\%$), bias ($\beta = 0.7\%$), and error (RMSE and RMSLD < 0.13). However, LightGBM showed residual spatial autocorrelation ($I = 0.209$, $p_r < 0.05$), though this did not survive FDR correction ($p_{pr\beta} = 0.091$) and should be interpreted cautiously. Conversely, DT showed a low Moran's I ($I = -0.080$, $p_{pr\beta} = 0.424$) despite not having the best accuracy. XGBoost and GBDT performed similarly in accuracy and both showed non-significant autocorrelation after FDR correction.

In summary, different models have different performances. This chapter demonstrates that several specific ML models are spatially misspecified despite achieving good aggregate accuracy, a finding that is consequential because published inland water quality retrieval studies routinely report only accuracy metrics and omit residual autocorrelation checks, leaving spatial misspecification undetected. Residual spatial autocorrelation is a diagnostic of model misspecification, specifically an omitted spatial effect, and not a second measure of predictive accuracy. A model can therefore be accurate and spatially misspecified at the same time, or inaccurate and spatially well-specified. Checking residual autocorrelation alongside conventional accuracy metrics is therefore recommended as standard practice in inland water quality retrieval studies. If significant autocorrelation is found, identifying and incorporating the omitted spatially structured variable may further improve both spatial specification and predictive accuracy.

4.5.4 Salinity Map Validation

The salinity maps in Figure 4.7 (b, c) are consistent with previous observations, where Ansari and Akhoondzadeh (2021; 2020) demonstrated that the Karun River became increasingly saline from upstream to downstream on 01 February 2015, and 05 September 2018.

Considering that there is no sampling station in the reservoir and only one sampling station after the confluence with the Arvand River (i.e., Abadan), the lack of local validation renders the regression model output more uncertain in these areas. As shown in Table 4.6, the RMSE for Abadan (0.19 psu) is the highest of any station, illustrating the increased uncertainty in this area. However, from Gotvand to Ahvaz, given the six sampling stations and the good performance of GBDT in terms of statistical metrics and residual autocorrelation analysis, the salinity maps can be considered more reliable ($RMSE < 0.13$ psu).

Based on Figure 4.7, and especially Figure 4.7 (d), a general increase in salinity is observed from the upstream end of the reservoir to the dam, suggesting that water becomes saltier due to prolonged contact with the Gachsaran formation, which is consistent with a previous study (Jalali et al., 2019). However, the Gachsaran formation is unlikely to be the sole cause of salinity in the Karun River because of the observed increase in salinity from Gotvand to Ahvaz city. According to Table 4.7, salinity increased on average by more than 34% from Gotvand to Ahvaz city across four different dates. The salinity increase from the reservoir to Gotvand across these dates is notably smaller than that observed from Gotvand to Ahvaz city, indicating that additional factors influence salinity levels in the river at Ahvaz. In this regard, the widespread cultivation of sugarcane in Shushtar and Ahvaz city (Moradi-Majd et al., 2022), which produces agricultural waste, may also contribute to the increased salinity observed from Gotvand to Ahvaz. Specifically, sugarcane production in the Karun basin exacerbates salinity due to several interconnected factors (Ghadiri, 2016): the land selected for sugarcane cultivation is characterized by saline and shallow groundwater, which, when used for irrigation, introduces salt into the soil and eventually into the river. To manage these saline lands, a drainage system is employed, involving the leaching of salt from the soil by applying large

volumes of water, which then drains into a collection system. This drainage water, containing both irrigation water and naturally saline groundwater, is subsequently pumped back into the Karun River, further increasing its salinity.

Regarding Khorramshahr city, as shown in Figure 4.7 and Table 4.7, the salinity on four different dates is inconsistent compared to Ahvaz city. The average salinity in Khorramshahr was relatively similar to Ahvaz on 19 June 2013, was lower by over 56% on 01 February 2015, and 05 September 2018, and was higher on 11 May 2022. This variability could be due to the model's lack of calibration at downstream points or the influence of water inflow from the Arvand River (see Figure 4.1) (Maliki et al., 2020) near Khorramshahr, affecting downstream WQP in the Karun River.

4.6 Conclusion

ML models, especially DT-based models like GBDT, outperformed traditional statistical models for mapping salinity in the Karun River using Landsat 8 OLI imagery. The Red band, Green band, NSMI, EGRI, and Salinity indices 2-6 were the most important spectral features for salinity prediction.

This chapter demonstrates that several specific ML models are spatially misspecified despite achieving good aggregate accuracy, as evidenced by significant residual spatial autocorrelation after FDR correction. The Red band, NSMI, EGRI, and Salinity indices 2–6 were the most important spectral features for salinity prediction. Residual spatial autocorrelation analysis, using permutation pseudo p-values with Benjamini-Hochberg FDR correction across 24 models, is recommended as a standard complement to conventional performance evaluation in inland water quality retrieval studies. Models like GWR and SLM, which incorporate spatial information, effectively addressed residual autocorrelation. Landsat 8 OLI imagery successfully mapped salinity dynamics in the Karun River.

Landsat 8 OLI imagery successfully mapped salinity dynamics in the Karun River. Salinity increased from Gotvand to Ahvaz city, likely due to agricultural and industrial activities such as sugarcane plantations. Future research should incorporate additional datasets to provide a more detailed examination of the influence of various natural (e.g., geology) and anthropogenic (e.g., land use) factors affecting salinity along the Karun River.

Acknowledgments

The authors acknowledge the Iran Water and Power Resources Development Company for providing the *in situ* water salinity data and the US Geological Survey (USGS) for providing Landsat-8 OLI images.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used OpenAI's ChatGPT for English grammar corrections. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

4.7 Appendix C

4.7.1 Computation of Salinity

Salinity in psu can be determined from specific conductance in $\mu\text{S}/\text{cm}$ using Equation C(1, 2) (Schemel, 2001; Wagner et al., 2006b):

$$\begin{aligned} \text{Salinity (psu)} = & 0.0120 - \left(0.2174 \times R^{1/2} \right) + (25.3283 \times R) + \left(13.7714 \times R^{3/2} \right) \\ & - (6.4788 \times R^2) + \left(2.5842 \times R^{5/2} \right) \end{aligned} \quad \text{C1}$$

$$R = \frac{\text{specific conductance } (\mu\text{S}/\text{cm})}{53087 \mu\text{S}/\text{cm}} \quad \text{C2}$$

The variable R represents the ratio of specific conductance at 25 °C to that of standard seawater with a salinity of 35 psu at 25 °C, i.e. 53,087 $\mu\text{S}/\text{cm}$.

4.7.2 Feature Selection

4.7.2.1 Random Forest Feature Importance Score

In order to find the best combination of bands/features, we employed the Random Forest Feature Importance (RFFI) score. To determine the optimal hyperparameter values of the Random Forest (RF) model, a Genetic Algorithm (GA) was utilized, with the fitness function being R^2 . The optimization results are presented in Table C1.

4.7.2.2 Sobol' Sensitivity Analysis

To implement the Sobol' sensitivity analysis for finding the best bands/features combination, a linear model is required. Therefore, the Ordinary Least Squares (OLS) model was employed. The dependent parameter is the salinity, and the independent parameters consist of spectral bands of Landsat 8 OLI and 14 salinity indices calculated using all available matchups ($N=102$). The output of OLS is presented in Table C1.

Table C1. The output of OLS and tuned hyperparameters obtained through GA for RF model. R^2 served as a fitness function for GA. ε is median symmetric accuracy.

Model	Hyperparameter(s) or Model	R^2	ε
RF	$n_estimators_{RF} = 40$, $max_depth_{RF} = 97$, $min_samples_split_{RF} = 8$, $min_samples_leaf_{RF} = 1$, and $max_features_{RF} = 6$	0.51	39.94
OLS	Salinity = $10.0048 + 17.4752 \times \text{Coastal Aerosol}$ $- 95.1472 \times \text{Blue} + 30.3542 \times \text{Green}$ $+ 219.0711 \times \text{Red} - 0.4421 \times \text{NIR} - 8.9355$ $\times \text{SWIR 1}$ $+ 6.7681 \times \text{SWIR 2} - 62.7162$ $\times \text{Salinity index 1}$ $+ 195.6363 \times \text{Salinity index 2} - 166.6795$ $\times \text{Salinity index 3} - 108.6336 \times \text{Salinity index 4}$ $- 20.1965 \times \text{Salinity index 5}$ $+ 92.8954 \times \text{Salinity index 6} - 8.1160$ $\times \text{Salinity index 7} - 10.1899 \times \text{NSMI} - 1.0215$ $\times \text{NDSSI} - 0.3527 \times \text{BR} + 0.1549 \times \text{WSRI}$ $+ 1.2351 \times \text{EGRI} - 0.3855 \times \text{NDVI}$ $+ 0.3855 \times \text{NDSI}$	0.60	24.69

4.7.3 Regression Models' Hyperparameters

4.7.3.1 Statistical Models

OLS regression minimizes the sum of squared differences between observed and predicted values in linear regression. OLS assumes that the errors are independent and identically distributed, which may not be true, especially for spatial data (Xie et al., 2013). To address this limitation, spatial regression methods such as the Spatial Lag Model (SLM) and Spatial Error Model (SEM) (Anselin, 2001) account for spatial autocorrelation by incorporating spatially lagged variables or allowing for spatially correlated error terms (Mainali et al., 2019), respectively. A weight matrix is required to implement the SLM and SEM, reflecting each observation's influence on its neighbors. This matrix is calculated using the data's spatial coordinates, with closer observations receiving higher weights. Furthermore, the Geographically weighted regression (GWR) method (Brunsdon et al., 1998) offers localized modeling capabilities by incorporating spatially varying coefficients based on spatial coordinates. GWR utilizes weighted linear regression to develop a distinct model for each geographical location. The choice of weighting is determined by employing either a Gaussian or biweight kernel function (Xie et al., 2013). The bandwidth hyperparameter which allows for adjustments to accommodate varying sample densities, need to be tuned.

A common challenge in linear modeling is multicollinearity; highly correlated predictor variables. Models like Partial Least Squares Regression (PLSR) (Geladi and Kowalski, 1986) and Ridge Linear Regressor (RLR) (Hoerl and Kennard, 1970) provide solutions to this issue. PLSR performs dimensionality reduction to capture maximal variance in both predictors and response variables, thereby mitigating the effects of multicollinearity. The “number of components” parameter ($n_{\text{components}_{\text{PLSR}}}$), which determines the number of latent variables used in the dimensionality reduction process, needs to be tuned. On the other hand, RLR introduces a regularization term to the OLS objective function, penalizing large coefficients and reducing the impact of multicollinearity on parameter estimates. For the implementation of RLR, different solver algorithms can be utilized, like least squares, Cholesky, Singular Value Decomposition, and

Average Gradient descent. These solvers determine the computational approach used to solve the optimization problem in RLR. Additionally, the regularization strength, controlled by the α_{RLR} parameter, needs to be tuned. The α_{RLR} parameter penalizes the size of the coefficients, with higher α_{RLR} values leading to greater regularization.

Lastly, Generalized Additive Model (GAM) (Hastie, 1992) and Generalized Linear Model (GLM) (Gotway and Stroup, 1997) extend the traditional linear modeling framework by accommodating non-linear relationships between predictors and response variables. GAM allows for flexible modeling of complex data patterns by incorporating various types of terms including spline terms_{GAM}, factor terms_{GAM}, tensor products_{GAM}, and intercept_{GAM}. Spline terms are used to model smooth non-linear relationships, while factor terms are used to incorporate categorical variables into the model. Tensor products allow for interactions between variables. Additionally, the intercept term represents the baseline level of the response variable. GLM handles non-normal error distributions and models non-linear relationships through appropriate link functions. In GLM, the choice of link function determines how the mean of the response variable depends on the linear predictor. Commonly used link functions include the identity, log, and power transform functions, among others. The choice of link function should be based on the distributional assumptions of the response variable and the nature of the relationship between predictors and response. Additionally, GLM allows for the specification of the error distribution, including Gaussian, binomial, Poisson, and gamma distributions, among others.

4.7.3.2 Kernel-Based Models

Kernel-based models operate by implicitly transforming input variables into high-dimensional feature spaces, enabling them to model complex relationships in data (Pal, 2009). A key component of kernel-based models is the kernel function, which computes similarity between pairs of data points in the original input space. In kernel-based models, selecting the appropriate kernel function and tuning its hyperparameters is critical. For all kernel-based methods, we test the Radial Basis

Function (RBF), linear, sigmoid, and second-degree polynomial kernels, selecting the one that produces the best results. The Radial Basis Function (RBF) kernel is commonly used among kernel options (Pal, 2009). RBF and sigmoid kernels use the γ_{RBF} and $\gamma_{Sigmoid}$ parameter, respectively, which must be tuned, to regulate the influence of individual data points on the decision boundary or regression function in kernel-based models.

Gaussian Process Regression (GPR) (Williams and Rasmussen, 1995) adopts a non-parametric Bayesian approach to model the relationship between input variables and outputs as a distribution over functions. By assuming a multivariate Gaussian distribution for any finite number of points from the function, GPR captures non-linear relationships and prediction uncertainties. To utilize GPR effectively, tuning $n_restarts_optimizer$ and α_{GPR} hyperparameters which govern the model's flexibility and regularization is crucial.

Kernel Ridge Regression (KRR) (Vovk, 2013) serves as a regularized version of linear regression, employing kernel functions to map input variables into high-dimensional feature spaces. By incorporating a regularization term that penalizes large coefficients, KRR mitigates overfitting and enhances generalization performance. For KRR the α_{KRR} hyperparameter requires tuning, which regulates the balance between fitting the training data and preventing overfitting.

Support Vector Regression (SVR) (Smola and Schölkopf, 2004) builds on support vector machines seeks a hyperplane that maximizes the margin around the predicted values. Hyperparameters to tune in SVR are C , controlling the regularization strength, and ϵ_{SVR} , determining the width of the epsilon-tube around the regression function.

4.7.3.3 Decision Tree-Based Models

Tree-based models, based on Decision Tree (DT) (Suthaharan, 2016) as their fundamental units, characterized by hierarchical structures where each node represents a decision based on a feature, offer an intuitive way to analyze and interpret data patterns. They recursively split the data into subsets based on features, resulting in a tree structure. When implementing DT (and similarly for

other DT-based models), it is crucial to tune the $\text{max_depth}_{\text{DT}}$ hyperparameter, which determines the maximum depth of the tree, controlling the complexity of the model and its ability to capture intricate patterns. Other DT hyperparameters [common to both RF and Light Gradient-Boosting Machine (LightGBM)] that require tuning are $\text{min_samples_split}_{\text{DT}}$, which specifies the minimum number of samples required to split an internal node and influences the tree's tendency to split further, as well as $\text{min_samples_leaf}_{\text{DT}}$, which sets the minimum number of samples required to be at a leaf node and affects the granularity of the tree and limits overfitting. The last DT hyperparameters shared with RF is $\text{max_features}_{\text{DT}}$, which determines the maximum number of features to consider when looking for the best split at each node.

Ensemble learning models like RF (Breiman, 2001), Extreme Gradient Boosting (XGBoost) (Chen and Guestrin, 2016), LightGBM (Ke et al., 2017), and Gradient Boosted Decision Trees (GBDT) (Friedman, 2001) try to elevate predictive accuracy by harnessing the collective wisdom of multiple DTs. They construct an ensemble of DTs, each trained on a random subset of the training data and features. By averaging the predictions from these trees, they try to reduce overfitting and enhance model robustness. When implementing RF and other ensemble learning DT-based models, the number of trees in the forest ($\text{n_estimators}_{\text{RF}}$) hyperparameter, which influences model complexity and generalization ability, should also be fine-tuned. Additional hyperparameters to tune in XGBoost are: $\text{subsample}_{\text{XGBoost}}$, which specifies the fraction of samples used in each boosting round, $\text{colsample_bytree}_{\text{XGBoost}}$, which specifies the fraction of features to be randomly sampled for each tree, $\text{learning_rate}_{\text{XGBoost}}$ (shared with GBDT and LightGBM), which determines the step size shrinkage during the optimization process and influence the contribution of each tree, $\text{min_child_weight}_{\text{XGBoost}}$, which sets the minimum sum of instance weight required in a child and affect the tree's partitioning, γ_{XGBoost} , which is a regularization parameter that controls the complexity of the tree and penalize splits that do not significantly reduce loss, and α_{XGBoost} , which introduces L1 regularization on leaf weights and promote sparsity in the model.

LightGBM requires tuning the following parameters: $\text{num_leaves}_{\text{LightGBM}}$, which sets the maximum number of leaves in one tree, $\text{feature_fraction}_{\text{LightGBM}}$, which specifies the fraction of features to be used in each boosting round, and $\text{bagging_freq}_{\text{LightGBM}}$, which controls bagging frequency.

Among the tree-based models implemented in this study, Cubist (Quinlan, 1992) stands out as a rule-based approach. Cubist combines regression trees with boosting techniques to enhance predictive accuracy. To implement Cubist effectively, three hyperparameters, n_rules , $n_committees$, and $neighbors$ should be tuned. The n_rules parameter determines the maximum number of rules generated by the model, influencing the complexity of the decision-making process, $n_committees$ specifies the number of boosting committees used by the model, affecting the overall boosting process, and the $neighbors$ parameter controls the number of nearest neighbors considered during the prediction process, influencing the model's sensitivity to local patterns in the data.

4.7.3.4 Neural Network-Based Models

Neural Network (NN)-based models (Atkinson and Tatnall, 1997; Paoletti et al., 2019) come in various forms but share the common goal of capturing data patterns. To work well, they need special settings, like the activation function that decides how neurons behave, like ReLU. Another important setting is the learning rate, which controls how much the model learns from data each time it updates. Also, there is solver or optimizer, like Adam, which helps the model adjust itself to get better results. Multilayer Perceptron (MLP) (Atkinson and Tatnall, 1997) form the foundation of NNs, consisting of fully connected (FC) layers where each neuron in a layer is connected to every neuron in the subsequent layer. Tuning the number of FC layers and neurons in each layer, along with the learning rate, is crucial for optimizing MLP performance. Artificial Neural Network (ANN) (Paoletti et al., 2019) build upon the structure of MLP but typically incorporate more layers, allowing them to learn increasingly complex representations of the data.

Adaptive Neuro Fuzzy Inference System (ANFIS) (Jang, 1993) is a fuzzy, logic-based, ML model that inherently is a Takagi-Sugeno-Kang fuzzy inference system, but is built and trained using the structure of a NN. It works as an optimization framework by tuning all the Takagi-Sugeno-Kang parameters (i.e., parameters of the Gaussian function). Membership in ANFIS refers to the characterization of input-output relationships using fuzzy sets, and it's important to determine the number of memberships to capture the underlying complexity of the data effectively.

One-Dimensional Convolutional Neural Network (1D-CNN) (Albawi et al., 2017) consist of one dimensional convolutional layers enabling them to automatically learn hierarchical features. Tuning the number of one-dimensional convolutional layers, kernel size, FC layers, neurons, and learning rate is essential for optimizing 1D-CNN performance. The kernel size is particularly related to the dimension of the input data, influencing the extraction of spatial features.

Moving to recurrent NN, Simple Recurrent Neural Network (RNN) (Elman, 1990), Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997), and Gated Recurrent Units (GRU) (Cho et al., 2014) models use feedback loops to use previous data as they process new input. This two-way flow of information helps the model learn to predict outcomes at each layer, with each node storing ongoing calculations. Tuning the number of layers, units, and learning rate ensures optimal performance in capturing long-term dependencies. LSTM and GRU address the vanishing gradient problem, with LSTM utilizing memory cells and gating mechanisms, while GRU exhibits simpler architectures.

4.7.4 Genetic Algorithm

4.7.4.1 Overview of GA

The Genetic Algorithm (GA) (Holland, 1992) , which is inspired by Darwin's theory is one of the optimization techniques widely used to find the optimal parameters to solve optimization problems. Regarding Darwin's theory, there are different organisms in nature called *populations*. Each population has members who can reproduce new members as *offspring*. As good offspring is

produced by better members, in the long run only the good members are selected to reproduce the new generation, and others die.

The GA is hierarchically composed of three smaller components: (i) *population*, (ii) *chromosome*, and (iii) *gene* (Goldberg and Holland, 1988). A population that is randomly initialized is a collection of individuals. An *individual* or a chromosome is a solution vector (i.e., one of the solutions in solution space). Finally, all the chromosomes consist of discrete genes that can control one or more chromosome features.

To generate new solutions from existing solutions, GA uses two operators: *crossover* and *mutation*. On the one hand, in a crossover operator, two chromosomes are selected from the existing chromosomes based on their goodness (Holland, 1992). The selected chromosomes, called *parents*, combine to reproduce offspring, which inherit good genes from parents. Therefore, good genes appear more frequently in the population by applying the crossover operator iteratively. On the other hand, in a mutation operator, random changes are made to the part of genes of some of the population's chromosomes. The *mutation rate* depends on the chromosome's length and determines the probability of changing the gene's characteristics (Piszc and Soule, 2006).

GA selects the fit chromosomes (i.e., solutions) based on their fitness as new parents or the final solution (selection). The selection procedure is performed using the *fitness function*, which can determine the probability of a chromosome's survival for the next generation. To summarize, GA searches in the solution space using crossover and mutation operators to maximize the fitness function and find the optimal parameters. Table C2 describes the procedure of GA.

4.7.4.2 Genetic Algorithm for Hyperparameter Tuning

We employed a GA to optimize the hyperparameters of our regression models using the *PyGAD* library (Gad, 2023) in Python. The following parameters were set for the GA:

Number of Generations (`num_generations = 300`): The algorithm iterates through 300 generations to evolve the population towards optimal solutions.

Number of Parents Mating (`num_parents_mating = 20`): In each generation, 20 solutions are selected as parents to produce offspring for the next generation.

Solutions per Population (`sol_per_pop = 50`): Each generation consists of 50 potential solutions (chromosomes).

Mutation Percentage of Genes (`mutation_percent_genes = 100`): Every gene in the population has a 100% chance of undergoing mutation. This high mutation rate ensures diversity in the population, preventing premature convergence to suboptimal solutions.

Parent Selection Type (`parent_selection_type = "sss"`): We used steady-state selection (sss) for parent selection. This method provides a balanced approach to selecting parents based on their fitness, ensuring that even less fit individuals have a chance to contribute to the next generation.

Number of Parents Kept (`keep_parents = 1`): The fittest parent from the current generation is retained in the next generation to maintain good solutions.

Crossover Type (`crossover_type = "single_point"`): Single-point crossover (default) was employed, where a random crossover point is selected, and the genes from parent solutions are exchanged at this point to produce new offspring.

Mutation Type (`mutation_type = "random"`): Random mutation was applied to introduce variability in the genes, which helps in exploring the solution space more thoroughly.

Fitness Function (`fitness = R2`): The fitness of each solution was evaluated using the R^2 providing a measure of the model's accuracy.

This approach enabled us to explore the hyperparameter space and identify the optimal settings for our regression models. The results of GA for each model are listed in Table C3.

Table C2. Genetic Algorithm.

Genetic Algorithm	
begin	Initialization: set $Gen = 1$. Introduce the first population (Pop_1) by randomly generating m chromosomes ($sol_per_pop = m$). Calculate the fitness function of n chromosomes in Pop_1 .
<i>Step 1</i>	Crossover: generate a new population (new_pop_{Gen}) as follows: 2.1. From Pop_{Gen} , choose two chromosomes ($chrom'$ and $chrom''$), which have the highest fitness function value. 2.2. Generate new chromosomes (offspring) using crossover operator and add them to new_pop_{Gen} .
<i>Step 2</i>	Mutation: based on the mutation rate ($mutation_percent_genes$), mute each chromosome in new_pop_{Gen} .
<i>Step 3</i>	Fitness calculation: based on the fitness function, calculate the fitness function value for each chromosome in new_pop_{Gen} .
<i>Step 4</i>	Selection: from new_pop_{Gen} select m chromosomes, which have the highest fitness function value and copy them to Pop_{Gen+1} .
<i>Step 5</i>	Evaluation: From Pop_{Gen+1} selected the best chromosomes which has the highest fitness value. <i>if</i> , the best chromosomes ($chrom_{best}$) can be considered the proper solution for optimization problem: stop search. return $chrom_{best}$ <i>else</i> , set $Gen = Gen + 1$ go to <i>Step 2</i> .
end	

Table C3. Tuned hyperparameters obtained through Genetic Algorithm for linear, kernel-based, and tree-based models, as well as architectures for neural network-based models. R^2 served as a fitness function for Genetic Algorithm. The index of kernel-based methods represents the kernel function, which includes Radial Basis Function (RBF), linear, second-degree polynomial (Poly2), and sigmoid.

Class	Model	Hyperparameter(s) or Architecture	R^2
Statistical Models	GAM	spline terms $_{GAM} = 9$, factor terms $_{GAM} = 0$, tensor products $_{GAM} = 9$, and intercept $_{GAM} = 1 \times 10^{-3}$	0.86
	GLM	link function $_{GLM} = \log$ and lerror distribution $_{GLM} = Gaussian$	0.36
	GWR	bandwidth $_{GWR} = 28$	0.69
	OLS	Salinity = $4.15 - 2.30 \times \text{Salinity index 2} - 22.22 \times \text{Salinity index 3} - 65.69 \times \text{Salinity index 4} - 25.94 \times \text{Salinity index 5} + 53.59 \times \text{Salinity index 6} - 3.68 \times \text{Salinity index 7} - 3.12 \times \text{NSMI} + 63.17 \times \text{Red} + 20.10 \times \text{Green} - 0.20 \times \text{EGRI}$	0.33
	PLSR	n_components $_{PLSR} = 6$	0.47
	RLR	solver algorithm $_{RLR} = \text{least squares}$ and $\alpha_{RLR} = 4.9 \times 10^{-4}$	0.34
	SLM	weight matrix = inverse of Haversine distance	0.58
	SEM	weight matrix = inverse of Haversine distance	0.47
Kernel-based Models	GPR _{RBF}	$\gamma_{RBF} = 6.3 \times 10^{-2}$, n_restarts_optimizer = 10, and $\alpha_{GPR} = 0.1$	0.51
	GPR _{Linear}	n_restarts_optimizer = 10, and $\alpha_{GPR} = 2.3 \times 10^{-1}$	0.34
	GPR _{Poly2}	n_restarts_optimizer = 10, and $\alpha_{GPR} = 2.8 \times 10^{-2}$	0.41
	GPR _{Sigmoid}	$\gamma_{Sigmoid} = 6.8 \times 10^{-7}$, n_restarts_optimizer = 10, and $\alpha_{GPR} = 2.7 \times 10^{-4}$	0.12
	KRR _{RBF}	$\gamma_{RBF} = 8.25$ and $\alpha_{KRR} = 1.03 \times 10^{-3}$	0.61
	KRR _{Linear}	$\alpha_{KRR} = 2.25 \times 10^{-1}$	0.29
	KRR _{Poly2}	$\alpha_{KRR} = 6.1 \times 10^{-2}$	0.33
	KRR _{Sigmoid}	$\gamma_{Sigmoid} = 675.51$ and $\alpha_{KRR} = 3.3 \times 10^{-1}$	0.14
	SVR _{RBF}	$\gamma_{RBF} = 417.32$, $C = 8345.61$, and $\epsilon_{SVR} = 14.21$	0.73
	SVR _{Linear}	$C = 6839.35$, and $\epsilon_{SVR} = 63.31$	0.66
	SVR _{Poly2}	$C = 2314.05$, and $\epsilon_{SVR} = 935.77$, degree = 2	0.25
	SVR _{Sigmoid}	$\gamma_{Sigmoid} = 0.35$, $C = 980.75$, and $\epsilon_{SVR} = 920.28$	0.11
	Cubist	n_rules = 2, n_committees = 5, and neighbors = 1	0.88
Decision Tree-based Models	DT	max_depth $_{DT} = 67$, min_samples_split $_{DT} = 4$, min_samples_leaf $_{DT} = 4$, and max_features $_{DT} = 7$	0.73
	XGBoost	colsample_bytree $_{XGBoost} = 0.07$, learning_rate $_{XGBoost} = 0.85$, min_child_weight $_{XGBoost} = 5$, subsample $_{XGBoost} = 0.77$, $\gamma_{XGBoost} = 10$, max_depth $_{XGBoost} = 90$, $\alpha_{XGBoost} = 0$, and n_estimators $_{XGBoost} = 614$	0.94
	GBDT	learning_rate $_{GBDT} = 1.12$, n_estimators $_{GBDT} = 83$, max_depth $_{GBDT} = 2$, min_samples_split $_{GBDT} = 2$, and min_samples_leaf $_{GBDT} = 1$	0.96
	LightGBM	num_leaves $_{LightGBM} = 4$, learning_rate $_{LightGBM} = 0.85$, feature_fraction $_{LightGBM} = 35$, bagging_freq $_{LightGBM} = 57$, max_depth $_{LightGBM} = 2$, and n_estimators $_{LightGBM} = 0.24$	0.94
	RF	n_estimators $_{RF} = 64$, max_depth $_{RF} = 92$, min_samples_split $_{RF} = 2$, min_samples_leaf $_{RF} = 1$, and max_features $_{RF} = 1$	0.75
	ANFIS	number of memberships = 8	0.39
Neural Network-based Models	ANN	ANN _{Architecture} : [FC(128), FC(64), FC(16), FC(16), FC(9), FC(1)], learning_rate $_{ANN} = 1.6 \times 10^{-3}$, optimizer $_{ANN} = adam$, activation function $_{ANN} = ReLU$	0.62
	1D-CNN	1D_CNN _{Architecture} : [Conv1D(32 $\times$ 8), Conv1D(64 $\times$ 8), FC(128), FC(6432), FC(1)], learning_rate $_{1D-CNN} = 1 \times 10^{-3}$, optimizer $_{1D-CNN} = adam$, activation function $_{1D-CNN} = ReLU$	0.41
	GRU	GRU _{Architecture} : [GRU _{layer} (64), GRU _{layer} (32), FC(256), FC(128), FC(1)],	0.48

		learning_rate _{GRU} = 1×10^{-4} , optimizer _{GRU} = <i>adam</i> , activation function _{GRU} = Linear	
	LSTM	LSTM _{Architecture} : [LSTM _{layer} (64), LSTM _{layer} (128), FC(128), FC(32), FC(1)], learning_rate _{LSTM} = 1×10^{-3} , optimizer _{LSTM} = <i>adam</i> , activation function _{LSTM} = Linear	0.41
	MLP	MLP _{Architecture} : [FC(96), FC(20), FC(41), FC(1)], learning_rate _{MLP} = 1×10^{-3} , optimizer _{MLP} = <i>adam</i> , activation function _{MLP} = <i>ReLU</i>	0.69
	RNN	SimpleRNN _{Architecture} : [SimpleRNN _{layer} (128), SimpleRNN _{layer} (32), FC(64), FC(1)], learning_rate _{SimpleRNN} = 1×10^{-3} , optimizer _{SimpleRNN} = <i>adam</i> , activation function _{SimpleRNN} = Linear	0.44

4.7.5 Moran's I Reference Distribution

To calculate Moran's I, we created a weight matrix ($N \times N$, where $N = 30$ test observations) using the inverse of Haversine distance between sampling locations. The weight matrix was row-standardized, and self-weights on the diagonal were set to zero so that no observation is treated as its own neighbor. Moran's I values were obtained through 1,000,000 random permutations of residuals across the fixed sampling locations under the null hypothesis of no spatial autocorrelation. To account for multiple testing across 24 models, Benjamini-Hochberg FDR correction was applied.

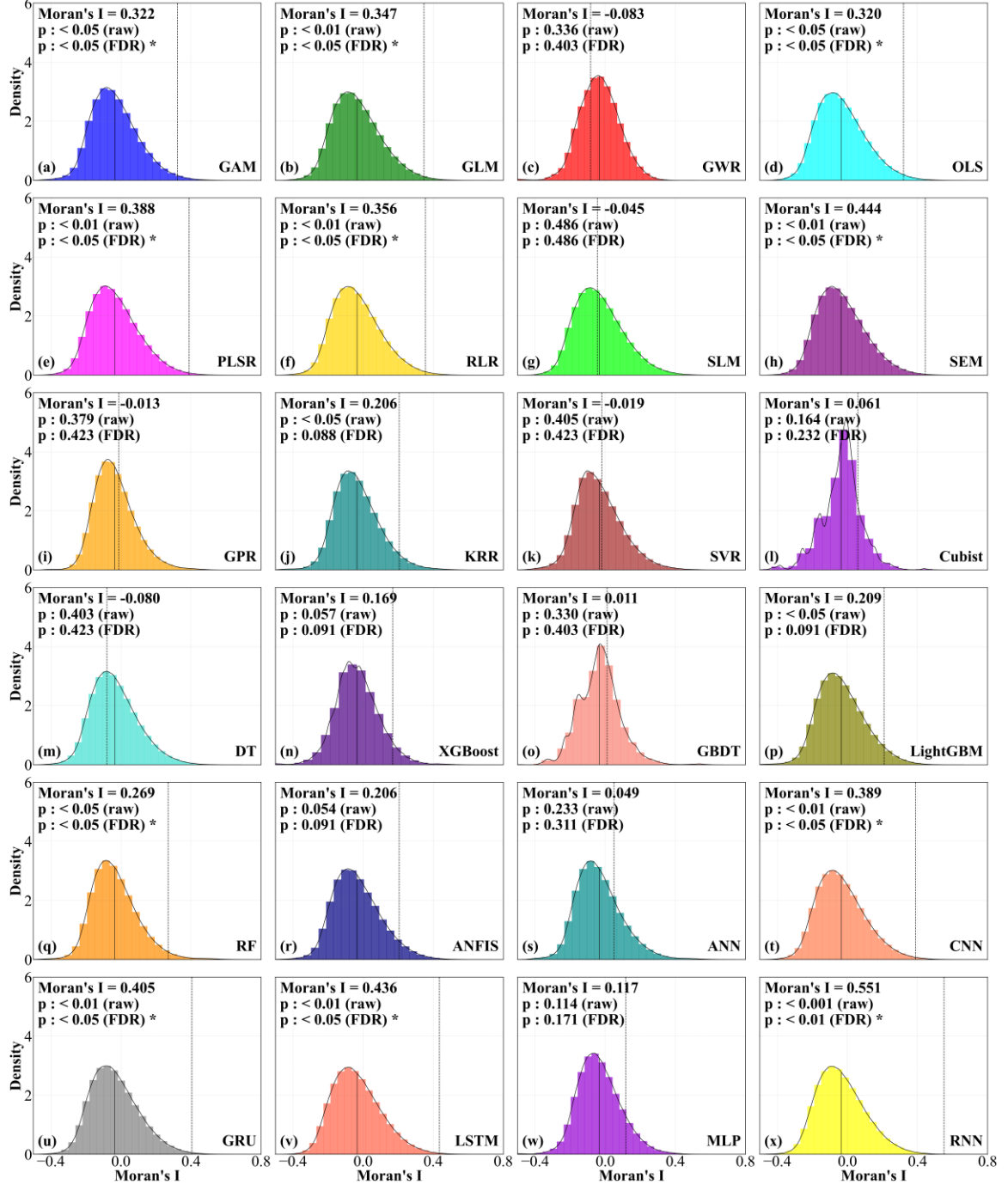


Fig. C1. Histograms illustrate the distribution of Moran's I values obtained through 1,000,000 permutations of residuals for each of the 24 models. The dashed vertical line shows the observed Moran's I for each model. The solid vertical line shows the expected value of Moran's I under the null hypothesis of no spatial autocorrelation, equal to $-1/(n-1) \approx -0.034$ for $n = 30$. Raw permutation p-values (p_r) and Benjamini-Hochberg FDR-corrected p-values (p_{prf}) are reported in each panel.

Conclusion

Effective satellite-based monitoring of inland WQPs depends on three main components: the spectral and spatial capabilities of the satellite sensor, the accuracy of the atmospheric correction algorithms, and the reliability of the WQPs retrieval approach. This thesis was structured around these three components, with each chapter contributing empirical evidence or methodological insight relevant to one of them. Taken together, the four chapters advance the state of knowledge in satellite RS of inland water quality and provide a coherent framework for evaluating and improving future monitoring systems.

Chapter 1 provided the conceptual foundation for this thesis through a systematic review of satellite RS studies for inland WQP retrieval, spanning sensor technologies, atmospheric correction algorithms, and ML models, and identified three research gaps that directly motivated the subsequent chapters. With respect to sensor technologies, the review highlighted that the suitability of a given sensor depends on the target WQP, since optically active constituents such as Chl-a, CDOM, and NAP interact with light at distinct wavelength regions, Chl-a retrieval typically relies on blue (~438 nm) and red (~667 nm) wavelengths, CDOM absorption is primarily driven by blue wavelengths, and NAP are associated with red and NIR scattering, and sensors that lack coverage of these specific regions or that acquire data at insufficient spectral resolution are necessarily limited in their retrieval capability. At the same time, high spatial resolution, which is critical for monitoring narrow rivers, often comes at the cost of reduced SNR, limiting sensitivity to subtle variations in water-leaving reflectance. These trade-offs between spatial resolution, spectral resolution, and SNR reflect fundamental physical constraints in sensor design rather than incidental technological limitations: a sensor that excels in one dimension will inevitably sacrifice performance in another; multi-sensor integration and the upcoming SBG mission, with its hyperspectral design, improved SNR, and sun-glint minimization capability, were identified as promising paths forward. The review also noted that Wyvern Dragonette-001, which combines

Conclusion

hyperspectral observations with high spatial resolution, had strong theoretical potential for inland water quality applications, but that its actual suitability for retrieving WQPs had not yet been evaluated, a gap directly addressed in Chapter 2.

With respect to atmospheric correction, the review found that no algorithm consistently outperforms others across all sensors, water conditions, and atmospheric environments. Several factors compound the AC challenge for inland waters: atmospheric heterogeneity caused by urban aerosols invalidates uniform atmospheric assumptions; moderate-to-high turbidity produces non-negligible NIR reflectance that violates the dark-pixel assumptions used by widely applied algorithms for open oceans; and the AE, arising from light scattered from nearby land surfaces into the sensor's field of view, can contaminate water pixel reflectance. A particularly important conclusion was that AE correction methods, though increasingly available, had not been evaluated for their effect on actual WQP retrieval accuracy, a gap directly addressed in Chapter 3. Finally, the review identified two frequently overlooked methodological issues in ML-based WQP retrieval: information leakage between training and test sets, arising when spatially autocorrelated observations appear on both sides of a randomly drawn split, which can cause performance metrics to be overly optimistic; and spatial autocorrelation in model residuals, which indicates that unmodeled spatial processes are influencing the target variable and can produce biased predictions in areas where residuals are positively clustered. These issues motivated the use of spatial blocking for data splitting in Chapters 2 and 3, and the explicit residual autocorrelation analysis in Chapter 4

Chapter 2 conducted the first evaluation of the Wyvern Dragonette-001 hyperspectral sensor for inland water quality retrieval, using Sentinel-2 MSI as a benchmark in the St. Lawrence River across three field campaigns and two study sites. Atmospheric correction was performed with RAdCor, and RF regression models were trained on in-situ measurements of Chl-a, CDOM, DOC, SS, and turbidity, with spatial blocking applied to prevent information leakage between training and test sets. The inter-sensor spectral comparison revealed weak agreement across all seven

Conclusion

corresponding band pairs, with RMSE exceeding 0.01 for all pairs; Wyvern overestimated reflectance at 670, 700, 735, 783, and 800 nm while underestimating it at 503 nm relative to Sentinel-2, with the best agreement at 550 nm. Two artifacts were identified in Wyvern imagery: elevated reflectance at wavelengths above 700 nm, where strong water absorption should produce near-zero values, and diagonal striping noise at longer wavelengths. Because concurrent Sentinel-2 imagery did not show similar patterns, both artifacts were attributed to the Dragonette sensor itself rather than to environmental or atmospheric conditions.

For turbidity, both sensors achieved strong results and Wyvern performed slightly better, with RMSE of 0.72 NTU (0.84% of the observed range) versus 0.88 NTU (1.02% of the observed range) for Sentinel-2. For DOC, Wyvern showed some indication of improved retrieval, and both sensors produced comparable results for SS; however, only 21 in-situ samples were available for these two parameters, which constrains the strength of these conclusions, and further validation with larger datasets is needed before the relative performance of the two sensors can be considered reliable. For CDOM and Chl-a, both sensors showed weak overall retrieval performance, but Sentinel-2 consistently outperformed Wyvern. For CDOM, Sentinel-2 achieved an RMSE of 24.41 ppb (7.4% of the observed range) compared to 44.60 ppb (13.5% of the observed range) for Wyvern. For Chl-a, Sentinel-2 achieved an RMSE of 0.28 $\mu\text{g/L}$ (1.9% of the observed range) compared to 1.07 $\mu\text{g/L}$ (7.3% of the observed range) for Wyvern. This performance gap is plausibly related to Wyvern's lack of blue spectral bands, which are known to be important for CDOM and Chl-a retrieval, but this explanation could not be directly confirmed in this study and should be treated as a hypothesis rather than an established cause, since the sensor artifacts described above may also have contributed. More recent Dragonette missions incorporating blue bands may help clarify and potentially resolve the observed limitations for CDOM and Chl-a retrieval.

Chapter 3 examined whether correcting for the AE using T-Mart, a physics-based Monte Carlo simulation approach, improves turbidity retrieval from Landsat 8/9 and Sentinel-2 imagery in the St. Lawrence River. Three pre-processing options were compared using a LightGBM model trained

Conclusion

on a large in-situ turbidity dataset with spatial blocking applied to prevent information leakage: raw TOA reflectance, atmospheric correction alone (ACOLITE's DSF), and AE correction with T-Mart followed by DSF. AE pre-processing consistently outperformed both alternatives. For Landsat 8/9, it reduced RMSE by 48.8% and ϵ by 19.0% compared to DSF alone, and by 48.5% and 50.7% respectively compared to the TOA baseline; for Sentinel-2, it reduced RMSE by 28.4% and ϵ by 50.8% relative to TOA. AE pre-processing showed its strongest advantage within 0-300 m of the shoreline, where AE is strongest, with RMSE more than 45% lower than the alternatives in the 200-300 m of the shoreline, while beyond 500 m, where the AE is lower, the uncorrected TOA approach performed best. Turbidity maps generated with AE pre-processing were also visibly smoother than the alternatives, with reduced noise near clouds, since T-Mart explicitly accounts for cloud-related scattering. These results suggest that AE correction can improve turbidity retrieval accuracy in nearshore areas of the St. Lawrence River, where the AE influence is strongest, though the generalizability of this finding to other river systems and WQPs remains to be evaluated.

Chapter 4 evaluated 24 statistical and ML models for salinity mapping in the Karun River, Iran, using Landsat 8 OLI imagery. Salinity in the Karun River is driven by both geological factors, including evaporite dissolution in the Gachsaran Formation, and agricultural drainage from sugarcane cultivation downstream. Feature selection identified the Red band, Salinity Indices 2-6, NSMI, and EGRI as the most informative predictors. Models were evaluated using both conventional accuracy metrics and global Moran's I permutation pseudo p-values with Benjamini-Hochberg FDR correction applied across all 24 models to account for multiple testing. Decision tree-based ensemble models substantially outperformed other model classes, with GBDT and XGBoost achieving the highest accuracy ($R^2 \geq 0.93$, $RMSE \leq 0.13$ psu, near-zero bias), while neural network-based models underfit the data due to the limited training set ($N = 72$). The central methodological contribution of Chapter 4 was demonstrating that several specific ML models, including most neural network-based approaches and several statistical models, are spatially misspecified despite achieving good aggregate accuracy, as evidenced by significant residual

Conclusion

spatial autocorrelation after FDR correction. This is consequential because published inland water quality retrieval studies routinely report only accuracy metrics and omit residual autocorrelation checks. LightGBM, for example, achieved good accuracy ($R^2 = 0.94$, $\varepsilon = 8.13\%$) but showed residual spatial autocorrelation ($I = 0.209$, $p_r < 0.05$) that did not survive FDR correction ($p_{pr\beta} = 0.091$) and should be interpreted cautiously. GWR and SLM, which explicitly incorporate spatial location, produced non-significant residual autocorrelation after FDR correction. GBDT was selected as the best overall model because it combined high accuracy with the lowest Moran's I among top performers ($I = 0.011$, $p_{pr\beta} = 0.403$), indicating the weakest spatial structure in its residuals. The salinity maps generated by GBDT across four Landsat 8 OLI acquisition dates confirmed a consistent increase in salinity from the Gotvand station downstream to Ahvaz city, with an average increase exceeding 34% across dates, reflecting the combined influence of salt dissolution from the Gachsaran Formation and saline agricultural drainage from sugarcane cultivation. These results reinforce the recommendation that residual spatial autocorrelation analysis using permutation pseudo p-values with FDR correction should be adopted as a standard complement to conventional accuracy evaluation in inland water quality retrieval studies.

A source of uncertainty that is relevant to all chapters of this thesis, and that warrants explicit recognition, is the temporal mismatch between in-situ field measurements and satellite observations. Satellite sensors capture the optical state of the water surface at a single instantaneous overpass, whereas in-situ measurements are collected along transects or at sequential stations during field campaigns that may span several hours around that overpass. In the field campaigns conducted for Chapters 2 and 3 in the St. Lawrence River, the temporal offset between the start of in-situ data collection and the satellite overpass was at most approximately three hours. During this interval, river flow continually advects water masses downstream and wind-driven mixing redistributes suspended sediment and other optically active constituents laterally and vertically, so the water observed by the instrument and the water captured by the satellite are not necessarily the same parcel.

Conclusion

The magnitude of water quality change over a two-to-three-hour window depends on the parameter of interest, the prevailing hydrodynamic conditions, and proximity to dynamic sources such as tributary inputs, shoreline erosion zones, or navigation-induced resuspension. For turbidity, the primary parameter evaluated in Chapter 3, short-term variability is driven largely by wind speed, vessel traffic, and episodic sediment resuspension events. A study has documented uncertainty of in-situ measurement of Chl-*a* and suspended particulate matter concentrations within the $\pm$ hour of satellite overpass is comparable to the uncertainties in satellite-derived water quality products (Salama et al., 2022). Current match-up protocols for satellite validation of water quality products typically recommend temporal windows of ± 1 to ± 4 hours around the satellite overpass as a reasonable compromise between data availability and temporal consistency (IOCCG, 2019). The temporal offsets in the field campaigns of this thesis fall within these recommended windows, suggesting that the associated uncertainty, while non-negligible, is consistent with accepted community practice. Nevertheless, temporal mismatch remains an unquantified component of the overall matchup uncertainty in this thesis, alongside spatial mismatch between the point-scale in-situ measurement and the area-integrated satellite pixel. Future studies in dynamic riverine environments would benefit from deploying continuously recording sensors at fixed stations to characterize sub-hourly variability at each sampling location, enabling explicit uncertainty propagation in matchup analyses rather than relying on community-accepted temporal windows. Another dimension that this thesis has not fully addressed is the pathway from research-grade satellite retrieval to operational use by private sector entities and water resource management organizations. The research presented here was conducted within two applied project contexts, the smartHarbour initiative in the St. Lawrence River, supported by the Canadian Space Agency and the Montreal Port Authority, and environmental monitoring of the Karun River in Iran, both of which involve real decision-making contexts in which water quality information has direct economic and regulatory consequences. Bridging the gap between the retrieval methods developed

Conclusion

in this thesis and their adoption in operational monitoring systems requires attention to several practical considerations that go beyond retrieval accuracy alone.

First, operational systems must be robust across a wider range of atmospheric, hydrological, and seasonal conditions than those represented in the training and validation datasets of this thesis, whose field campaigns were conducted during a limited number of seasons and environmental states. Second, the computational infrastructure required to apply these methods routinely, including access to near-real-time satellite imagery, automated atmospheric correction pipelines, and model inference systems, represents a non-trivial investment for most operational users, particularly for smaller organizations or those without dedicated remote sensing expertise. Third, the outputs of satellite-based water quality retrieval carry uncertainty that must be communicated clearly to decisionmakers who may not have a remote sensing background; uncertainty maps and confidence intervals are as important as the point estimates themselves, and their absence can lead to misplaced confidence in retrieved values. Fourth, private sector entities operating under regulatory frameworks, such as port authorities monitoring dredging-induced turbidity plumes, or agricultural operators managing saline irrigation water quality, require retrieval systems validated against the specific water quality thresholds and spatial scales relevant to their regulatory obligations, which may differ substantially from the research objectives addressed in this thesis. Addressing these operationalization challenges, encompassing system design, uncertainty communication, and regulatory alignment, represents a natural next step for translating the scientific contributions of this thesis into sustained, decision-relevant environmental monitoring capacity.

In summary, this thesis makes four distinct contributions to the field of satellite-based inland water quality monitoring. First, it provides a comprehensive review of sensor technologies, atmospheric correction algorithms, and ML retrieval approaches, synthesizing the state of the art and identifying key research gaps. Second, it presents the first empirical evaluation of the Wyvern Dragonette-001 hyperspectral sensor for inland water quality retrieval, characterizing its spectral data quality and

Conclusion

comparing its WQP retrieval performance against Sentinel-2 MSI across multiple parameters and field campaigns. Third, it provides a quantitative assessment of the impact of T-Mart AE correction on turbidity retrieval from Landsat 8/9 and Sentinel-2 imagery, demonstrating that AE correction can improve retrieval accuracy, particularly near the shoreline where AE contamination is strongest. Fourth, it demonstrates that several specific ML models are spatially misspecified despite good aggregate accuracy, and recommends residual spatial autocorrelation analysis as a standard complement to conventional accuracy evaluation. Together, these contributions provide actionable guidance for researchers and practitioners working to improve the accuracy, reliability, and practical application of satellite-based inland water quality monitoring.

References

- Abdelmalik, K.W., 2018. Role of statistical remote sensing for Inland water quality parameters prediction. *Egypt. J. Remote Sens. Space Sci.* 21, 193–200.
<https://doi.org/10.1016/j.ejrs.2016.12.002>
- Abdel-Rahman, E.M., Ahmed, F.B., Ismail, R., 2013. Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. *Int. J. Remote Sens.* 34, 712–728.
<https://doi.org/10.1080/01431161.2012.713142>
- Adib, A., Javdan, F., 2015. Interactive approach for determination of salinity concentration in tidal rivers (Case study: The Karun River in Iran). *Ain Shams Eng. J.* 6, 785–793.
<https://doi.org/10.1016/j.asej.2015.02.005>
- Adilakshmi, A., Venkatesan, V., 2024. Effective monitoring of Noyyal River surface water quality using remote sensing and machine learning and GIS techniques. *Desalination Water Treat.* 320, 100630. <https://doi.org/10.1016/j.dwt.2024.100630>
- Adjovu, G.E., Stephen, H., Ahmad, S., 2023a. A Machine Learning Approach for the Estimation of Total Dissolved Solids Concentration in Lake Mead Using Electrical Conductivity and Temperature. *Water* 15, 2439. <https://doi.org/10.3390/w15132439>
- Adjovu, G.E., Stephen, H., Ahmad, S., 2022. Monitoring of Total Dissolved Solids Using Remote Sensing Band Reflectance and Salinity Indices: A Case Study of the Imperial County Section, AZ-CA, of the Colorado River, in: *World Environmental and Water Resources Congress 2022*. American Society of Civil Engineers, pp. 1132–1145.
<https://doi.org/10.1061/9780784484258.106>
- Adjovu, G.E., Stephen, H., James, D., Ahmad, S., 2023b. Measurement of Total Dissolved Solids and Total Suspended Solids in Water Systems: A Review of the Issues, Conventional, and Remote Sensing Techniques. *Remote Sens.* 15, 3534.
<https://doi.org/10.3390/rs15143534>
- Adjovu, G.E., Stephen, H., James, D., Ahmad, S., 2023c. Overview of the Application of Remote Sensing in Effective Monitoring of Water Quality Parameters. *Remote Sens.* 15, 1938.
<https://doi.org/10.3390/rs15071938>
- Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: A Next-generation Hyperparameter Optimization Framework, in: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19*. Association for Computing Machinery, New York, NY, USA, pp. 2623–2631.
<https://doi.org/10.1145/3292500.3330701>
- Aksoy, S., Yildirim, A., Gorji, T., Hamzehpour, N., Tanik, A., Sertel, E., 2022. Assessing the performance of machine learning algorithms for soil salinity mapping in Google Earth Engine platform using Sentinel-2A and Landsat-8 OLI data. *Adv. Space Res.* 69, 1072–1086. <https://doi.org/10.1016/j.asr.2021.10.024>
- Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network, in: *2017 International Conference on Engineering and Technology (ICET)*. Presented at the 2017 International Conference on Engineering and Technology (ICET), pp. 1–6. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
- Allan, M.G., Hamilton, D.P., Hicks, B.J., Brabyn, L., 2011. Landsat remote sensing of chlorophyll a concentrations in central North Island lakes of New Zealand. *Int. J. Remote Sens.* 32, 2037–2055. <https://doi.org/10.1080/01431161003645840>
- Ansari, M., Akhoondzadeh, M., 2021. Generation of Karun River Water Salinity Map from Landsat-8 Satellite Images using Support Vector Regression, Multilayer Perceptron and

References

- Genetic Algorithm. *J. Environ. Stud.* 46, 581–600.
<https://doi.org/10.22059/jes.2021.317955.1008130>
- Ansari, M., Akhoondzadeh, M., 2020. Mapping water salinity using Landsat-8 OLI satellite images (Case study: Karun basin located in Iran). *Adv. Space Res.* 65, 1490–1502.
<https://doi.org/10.1016/j.asr.2019.12.007>
- Ansari, Mohsen, Akhoondzadeh, M., 2020. Mapping water salinity using Landsat-8 OLI satellite images (Case study: Karun basin located in Iran). *Adv. Space Res.* 65, 1490–1502.
<https://doi.org/10.1016/j.asr.2019.12.007>
- Ansari, M., Akhoondzadeh, M., 2019. Water Salinity Mapping of Karun Basin Located In Iran Using the SVR Method, in: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. pp. 97–101. <https://doi.org/10.5194/isprs-archives-XLII-4-W18-97-2019>
- Ansari, M., Knudby, A., Amani, M., Sawada, M., 2025a. Retrieving Inland Water Quality Parameters via Satellite Remote Sensing: Sensor Evaluation, Atmospheric Correction, and Machine Learning Approaches. *Remote Sens.* 17, 1734.
<https://doi.org/10.3390/rs17101734>
- Ansari, M., Knudby, A., Homayouni, S., 2025b. River salinity mapping through machine learning and statistical modeling using Landsat 8 OLI imagery. *Adv. Space Res.* 75, 6981–7002.
<https://doi.org/10.1016/j.asr.2025.03.037>
- Ansari, M., Knudby, A., Homayouni, S., 2025c. River Salinity Mapping through Machine Learning and Statistical Modeling using Landsat 8 OLI Imagery. *Adv. Space Res.*
<https://doi.org/10.1016/j.asr.2025.03.037>
- Ansari, M., Wu, Y., Knudby, A., Ansari, M., Wu, Y., Knudby, A., 2025d. Assessing the Impact of T-Mart Adjacency Effect Correction on Turbidity Retrieval from Landsat 8/9 and Sentinel-2 Imagery (Case Study: St. Lawrence River, Canada). *Remote Sens.* 18.
<https://doi.org/10.3390/rs18010127>
- Anselin, L., 2001. Rao's score test in spatial econometrics. *J. Stat. Plan. Inference*, Rao's Score Test 97, 113–139. [https://doi.org/10.1016/S0378-3758\(00\)00349-9](https://doi.org/10.1016/S0378-3758(00)00349-9)
- Anselin, L., Bongiovanni, R., Lowenberg-DeBoer, J., 2004. A Spatial Econometric Approach to the Economics of Site-Specific Nitrogen Management in Corn Production. *Am. J. Agric. Econ.* 86, 675–687. <https://doi.org/10.1111/j.0002-9092.2004.00610.x>
- Atkinson, P.M., Tatnall, A.R.L., 1997. Introduction Neural networks in remote sensing. *Int. J. Remote Sens.* 18, 699–709. <https://doi.org/10.1080/014311697218700>
- Babin, M., Stramski, D., Ferrari, G.M., Claustre, H., Bricaud, A., Obolensky, G., Hoepffner, N., 2003. Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. *J. Geophys. Res. Oceans* 108. <https://doi.org/10.1029/2001JC000882>
- Bahadori, A., Carranza, E.J.M., Soleimani, B., 2011. Geochemical analysis of evaporite sedimentation in the Gachsaran Formation, Zeloi oil field, southwest Iran. *J. Geochem. Explor., Geochemistry for Geological-Environmental Studies* 111, 97–112.
<https://doi.org/10.1016/j.gexplo.2011.02.007>
- Barati, A.A., Pour, M.D., Sardooei, M.A., 2023. Water crisis in Iran: A system dynamics approach on water, energy, food, land and climate (WEFLC) nexus. *Sci. Total Environ.* 882, 163549. <https://doi.org/10.1016/j.scitotenv.2023.163549>
- Bastviken, D., Tranvik, L.J., Downing, J.A., Crill, P.M., Enrich-Prast, A., 2011. Freshwater Methane Emissions Offset the Continental Carbon Sink. *Science* 331, 50–50.
<https://doi.org/10.1126/science.1196808>
- Bayati, M., Danesh-Yazdi, M., 2021. Mapping the spatiotemporal variability of salinity in the hypersaline Lake Urmia using Sentinel-2 and Landsat-8 imagery. *J. Hydrol.* 595, 126032.
<https://doi.org/10.1016/j.jhydrol.2021.126032>

References

- Begliomini, F.N., Barbosa, C.C.F., Martins, V.S., Novo, E.M.L.M., Paulino, R.S., Maciel, D.A., Lima, T.M.A., O'Shea, R.E., Pahlevan, N., Lamparelli, M.C., 2023. Machine learning for cyanobacteria mapping on tropical urban reservoirs using PRISMA hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* 204, 378–396. <https://doi.org/10.1016/j.isprsjprs.2023.09.019>
- Bernstein, L.S., Adler-Golden, S.M., Sundberg, R.L., Levine, R.Y., Perkins, T.C., Berk, A., Ratkowski, A.J., Felde, G., Hoke, M.L., 2005a. Validation of the QUick atmospheric correction (QUAC) algorithm for VNIR-SWIR multi- and hyperspectral imagery. Presented at the Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI, SPIE, pp. 668–678. <https://doi.org/10.1117/12.603359>
- Bernstein, L.S., Adler-Golden, S.M., Sundberg, R.L., Levine, R.Y., Perkins, T.C., Berk, A., Ratkowski, A.J., Felde, G., Hoke, M.L., 2005b. A new method for atmospheric correction and aerosol optical property retrieval for VIS-SWIR multi- and hyperspectral imaging sensors: QUAC (QUick atmospheric correction). Presented at the Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05., pp. 3549–3552. <https://doi.org/10.1109/IGARSS.2005.1526613>
- Blough, N.V., Zafiriou, O.C., Bonilla, J., 1993. Optical absorption spectra of waters from the Orinoco River outflow: Terrestrial input of colored organic matter to the Caribbean. *J. Geophys. Res. Oceans* 98, 2271–2278. <https://doi.org/10.1029/92JC02763>
- Borovskaya, R., Krivoguz, D., Chernyi, S., Kozhurin, E., Khorosheltseva, V., Zinchenko, E., 2022. Surface Water Salinity Evaluation and Identification for Using Remote Sensing Data and Machine Learning Approach. *J. Mar. Sci. Eng.* 10, 257. <https://doi.org/10.3390/jmse10020257>
- Bouaziz, M., Matschullat, J., Gloaguen, R., 2011. Improved remote sensing detection of soil salinity from a semi-arid climate in Northeast Brazil. *Comptes Rendus Géoscience* 343, 795–803. <https://doi.org/10.1016/j.crte.2011.09.003>
- Bowden, K.F., Sharaf El Din, S.H., 1966. Circulation, Salinity and River Discharge in the Mersey Estuary. *Geophys. J. Int.* 10, 383–399. <https://doi.org/10.1111/j.1365-246X.1966.tb03066.x>
- Bowers, D.G., Evans, D., Thomas, D.N., Ellis, K., Williams, P.J. le B., 2004. Interpreting the colour of an estuary. *Estuar. Coast. Shelf Sci.* 59, 13–20. <https://doi.org/10.1016/j.ecss.2003.06.001>
- Breiman, L., 2001. Random Forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brockmann, C., Doerffer, R., Peters, M., Kerstin, S., Embacher, S., Ruescas, A., 2016. Evolution of the C2RCC Neural Network for Sentinel 2 and 3 for the Retrieval of Ocean Colour Products in Normal and Extreme Optically Complex Waters. *Living Planet Symp.* 740, 54.
- Brunsdon, C., Fotheringham, S., Charlton, M., 1998. Geographically Weighted Regression. *J. R. Stat. Soc. Ser. Stat.* 47, 431–443. <https://doi.org/10.1111/1467-9884.00145>
- Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., Vanderplas, J., Joly, A., Holt, B., Varoquaux, G., 2013. API design for machine learning software: experiences from the scikit-learn project. <https://doi.org/10.48550/arXiv.1309.0238>
- Bulgarelli, B., Kiselev, V., Zibordi, G., 2014. Simulation and analysis of adjacency effects in coastal waters: a case study. *Appl. Opt.* 53, 1523–1545. <https://doi.org/10.1364/AO.53.001523>
- Bulgarelli, B., Zibordi, G., 2018. On the detectability of adjacency effects in ocean color remote sensing of mid-latitude coastal environments by SeaWiFS, MODIS-A, MERIS, OLCI, OLI and MSI. *Remote Sens. Environ.* 209, 423–438. <https://doi.org/10.1016/j.rse.2017.12.021>

References

- Cañedo-Argüelles, M., Hawkins, C.P., Kefford, B.J., Schäfer, R.B., Dyack, B.J., Brucet, S., Buchwalter, D., Dunlop, J., Frör, O., Lazorchak, J., Coring, E., Fernandez, H.R., Goodfellow, W., Achem, A.L.G., Hatfield-Dodds, S., Karimov, B.K., Mensah, P., Olson, J.R., Piscart, C., Prat, N., Ponsá, S., Schulz, C.-J., Timpano, A.J., 2016. Saving freshwater from salts. *Science* 351, 914–916. <https://doi.org/10.1126/science.aad3488>
- Cao, Z., Ma, R., Duan, H., Xue, K., 2019. Effects of broad bandwidth on the remote sensing of inland waters: Implications for high spatial resolution satellite data applications. *ISPRS J. Photogramm. Remote Sens.* 153, 110–122. <https://doi.org/10.1016/j.isprsjprs.2019.05.001>
- Cao, Z., Ma, R., Liu, J., Ding, J., 2021. Improved Radiometric and Spatial Capabilities of the Coastal Zone Imager Onboard Chinese HY-1C Satellite for Inland Lakes. *IEEE Geosci. Remote Sens. Lett.* 18, 193–197. <https://doi.org/10.1109/LGRS.2020.2971629>
- Castagna, A., Vanhellemont, Q., 2025a. A generalized physics-based correction for adjacency effects. *Appl. Opt.* 64, 2719–2743. <https://doi.org/10.1364/AO.546766>
- Castagna, A., Vanhellemont, Q., 2025b. A generalised physics-based correction for adjacency effects. *Appl. Opt.* <https://doi.org/10.1364/AO.546766>
- Cetinic, I., McClain, C.R., Werdell, P.J., 2018. Pre-Aerosol, Clouds, and Ocean Ecosystem (PACE) Mission Science Definition Team Report: PACE Technical Report Series - Volume 2 (No. GSFC-E-DAA-TN65844).
- Chang, H., 2008. Spatial analysis of water quality trends in the Han River basin, South Korea. *Water Res.* 42, 3285–3304. <https://doi.org/10.1016/j.watres.2008.04.006>
- Chatterjee, S., Bhadra, T., 2023. Assessing the Water Quality of Indian Sundarban Estuaries using Remote Sensing Techniques, in: *IOP Conference Series: Earth and Environmental Science*. IOP Publishing, p. 012006. <https://doi.org/10.1088/1755-1315/1164/1/012006>
- Chavez, P.S., 1996. Image-Based Atmospheric Corrections - Revisited and Improved. *Photogramm. Eng. Remote Sens.* 62, 1025–1035.
- Chavez, P.S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sens. Environ.* 24, 459–479. [https://doi.org/10.1016/0034-4257\(88\)90019-3](https://doi.org/10.1016/0034-4257(88)90019-3)
- Chegoonian, A.M., Pahlevan, N., Zolfaghari, K., Leavitt, P.R., Davies, J.-M., Baulch, H.M., Duguay, C.R., 2023. Comparative Analysis of Empirical and Machine Learning Models for Chla Extraction Using Sentinel-2 and Landsat OLI Data: Opportunities, Limitations, and Challenges. *Can. J. Remote Sens.* 49. <https://doi.org/10.1080/07038992.2023.2215333>
- Chen, J., Cui, T., Lin, C., 2013. An Improved SWIR Atmospheric Correction Model: A Cross-Calibration-Based Model. *IEEE Trans. Geosci. Remote Sens.* 52, 3959–3967. <https://doi.org/10.1109/TGRS.2013.2278340>
- Chen, L., Liu, L., Liu, S., Shi, Z., Shi, C., 2025. The Application of Remote Sensing Technology in Inland Water Quality Monitoring and Water Environment Science: Recent Progress and Perspectives. *Remote Sens.* 17, 667. <https://doi.org/10.3390/rs17040667>
- Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*. Association for Computing Machinery, New York, NY, USA, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. <https://doi.org/10.48550/arXiv.1406.1078>
- Chollet, F., 2015. keras.
- Christodoulou, A., Christidis, P., Bisselink, B., 2020. Forecasting the impacts of climate change on inland waterways. *Transp. Res. Part Transp. Environ.* 82, 102159. <https://doi.org/10.1016/j.trd.2019.10.012>

References

- Cloern, J.E., 1987. Turbidity as a control on phytoplankton biomass and productivity in estuaries. *Cont. Shelf Res., Dynamics of Turbid Coastal Environments* 7, 1367–1381.
[https://doi.org/10.1016/0278-4343\(87\)90042-2](https://doi.org/10.1016/0278-4343(87)90042-2)
- Cooley, T., Anderson, G.P., Felde, G.W., Hoke, M.L., Ratkowski, A.J., Chetwynd, J.H., Gardner, J.A., Adler-Golden, S.M., Matthew, M.W., Berk, A., Bernstein, L.S., Acharya, P.K., Miller, D., Lewis, P., 2002. FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation, in: *IEEE International Geoscience and Remote Sensing Symposium*. Presented at the IEEE International Geoscience and Remote Sensing Symposium, pp. 1414–1418 vol.3.
<https://doi.org/10.1109/IGARSS.2002.1026134>
- Cox, C., Munk, W., 1954. Statistics of the sea surface derived from sun glitter. *J. Mar. Res.* 13.
- Cuartero, A., Cáceres-Merino, J., Torrecilla-Pinero, J.A., 2023. An application of C2-Net atmospheric corrections for chlorophyll-a estimation in small reservoirs. *Remote Sens. Appl. Soc. Environ.* 32, 101021. <https://doi.org/10.1016/j.rsase.2023.101021>
- Dahaan, S.A.M.A., Al-Ansari, N., Knutsson, S., 2016. Influence of Groundwater Hypothetical Salts on Electrical Conductivity Total Dissolved Solids. *Engineering* 8, 823–830.
<https://doi.org/10.4236/eng.2016.811074>
- Dai, J., Liu, T., Zhao, Y., Tian, S., Ye, C., Nie, Z., 2023. Remote sensing inversion of the Zabuye Salt Lake in Tibet, China using LightGBM algorithm. *Front. Earth Sci.* 10.
<https://doi.org/10.3389/feart.2022.1022280>
- Dasgupta, S., Akhter Kamal, F., Huque Khan, Z., Choudhury, S., Nishat, A., 2013. River Salinity and Climate Change: Evidence from Coastal Bangladesh, in: *World Scientific Reference on Asia and the World Economy*. World Scientific, pp. 205–242.
https://doi.org/10.1142/9789814578622_0031
- De Keukelaere, L., Sterckx, S., Adriaensen, S., Knaeps, E., Reusen, I., Giardino, C., Bresciani, M., Hunter, P., Neil, C., Van der Zande, D., Vaiciute, D., 2018. Atmospheric correction of Landsat-8/OLI and Sentinel-2/MSI data using iCOR algorithm: validation for coastal and inland waters. *Eur. J. Remote Sens.* 51, 525–542.
<https://doi.org/10.1080/22797254.2018.1457937>
- Declaro, A., Kanae, S., 2024. Enhancing Surface Water Monitoring through Multi-Satellite Data-Fusion of Landsat-8/9, Sentinel-2, and Sentinel-1 SAR. *Remote Sens.* 16, 3329.
<https://doi.org/10.3390/rs16173329>
- DeepAI, 2025. DeepAI [WWW Document]. URL <https://deepai.org/> (accessed 4.5.25).
- Dekker, A.G., Hestir, E.L., 2012. Evaluating the feasibility of systematic inland water quality monitoring with satellite remote sensing. *CSIRO Water Healthy Ctry. Natl. Res. Flagship*.
- Dekker, A.G., Pinnel, N., Gege, P., Briottet, X., Court, A., Peters, S., Turpie, K.R., Sterckx, S., Costa, M., Giardino, C., Brando, V.E., Braga, F., Bergeron, M., Heege, T., Pflug, B., 2018. Feasibility Study for an Aquatic Ecosystem Earth Observing System Version 1.2., in: *Feasibility Study for an Aquatic Ecosystem Earth Observing System*. CEOS.
- DeLuca, N.M., Zaitchik, B.F., Guikema, S.D., Jacobs, J.M., Davis, B.J.K., Curriero, F.C., 2020. Evaluation of remotely sensed prediction and forecast models for *Vibrio parahaemolyticus* in the Chesapeake Bay. *Remote Sens. Environ.* 250, 112016.
<https://doi.org/10.1016/j.rse.2020.112016>
- Deng, Y., Zhang, Y., Pan, D., Yang, S.X., Gharabaghi, B., 2024. Review of Recent Advances in Remote Sensing and Machine Learning Methods for Lake Water Quality Management. *Remote Sens.* 16, 4196. <https://doi.org/10.3390/rs16224196>
- Douaoui, A.E.K., Nicolas, H., Walter, C., 2006. Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. *Geoderma* 134, 217–230.
<https://doi.org/10.1016/j.geoderma.2005.10.009>

References

- D'Sa, E.J., Hu, C., Muller-Karger, F.E., Carder, K.L., 2002. Estimation of colored dissolved organic matter and salinity fields in case 2 waters using SeaWiFS: Examples from Florida Bay and Florida Shelf. *J. Earth Syst. Sci.* 111, 197–207. <https://doi.org/10.1007/BF02701966>
- D'Sa, E.J., Zaitzeff, J.B., Steward, R.G., 2000. Monitoring water quality in Florida Bay with remotely sensed salinity and in situ bio-optical observations. *Int. J. Remote Sens.* 21, 811–816. <https://doi.org/10.1080/014311600210605>
- EcoLogic LLC., 2019. St. Lawrence River Watershed Characterization Report. Franklin County Soil & Water Conservation District, Louisburg, NC, USA.
- El Alem, A., Lhissou, R., Chokmani, K., Oubennaceur, K., 2021. Remote Retrieval of Suspended Particulate Matter in Inland Waters: Image-Based or Physical Atmospheric Correction Models? *Water* 13, 2149. <https://doi.org/10.3390/w13162149>
- Elman, J.L., 1990. Finding Structure in Time. *Cogn. Sci.* 14, 179–211. https://doi.org/10.1207/s15516709cog1402_1
- Elmotawakkil, A., Enneya, N., Bhagat, S.K., Ouda, M.M., Kumar, V., 2025. Advanced machine learning models for robust prediction of water quality index and classification. *J. Hydroinformatics* 27, 299–319. <https://doi.org/10.2166/hydro.2025.290>
- El-Swaify, S.A., 2000. Soil and Water Salinity, in: *Plant Nutrient Management in Hawaii's Soils: Approaches for Tropical and Subtropical Agriculture*, Chapter 17. University of Hawaii, pp. 151–158.
- Emberton, S., Chittka, L., Cavallaro, A., Wang, M., 2016. Sensor Capability and Atmospheric Correction in Ocean Colour Remote Sensing. *Remote Sens.* 8, 1. <https://doi.org/10.3390/rs8010001>
- Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., Kylling, J., Richter, B., Pause, C., Dowling, T., Bugliaro, L., 2016. The libRadtran software package for radiative transfer calculations (version 2.0.1). *Geosci. Model Dev.* 9, 1647–1672. <https://doi.org/10.5194/gmd-9-1647-2016>
- Fan, Y., Li, W., Gatebe, C.K., Jamet, C., Zibordi, G., Schroeder, T., Stamnes, K., 2017. Atmospheric correction over coastal waters using multilayer neural networks. *Remote Sens. Environ.* 199, 218–240. <https://doi.org/10.1016/j.rse.2017.07.016>
- Fang, L., Chen, S., Wang, H., Qian, J., Zhang, L., 2010. Detecting marine intrusion into rivers using EO-1 ALI satellite imagery: Modaomen Waterway, Pearl River Estuary, China. *Int. J. Remote Sens.* 31, 4125–4146. <https://doi.org/10.1080/01431160903229218>
- Feng, H., Wang, Y., Li, Z., Zhang, N., Zhang, Y., Gao, Y., 2023. Information Leakage in Deep Learning-Based Hyperspectral Image Classification: A Survey. *Remote Sens.* 15, 3793. <https://doi.org/10.3390/rs15153793>
- Feng, L., Wang, Y., Zhang, Z., Du, Q., 2021. Geographically and temporally weighted neural network for winter wheat yield prediction. *Remote Sens. Environ.* 262, 112514. <https://doi.org/10.1016/j.rse.2021.112514>
- Feng, P., Song, K., Wen, Z., Tao, H., Yu, X., Shang, Y., 2024. Remote Sensing Estimation of CDOM for Songhua River of China: Distributions and Implications. *Remote Sens.* 16, 4608. <https://doi.org/10.3390/rs16234608>
- Ferdous, J., Rahman, M.T.U., 2020. Developing an empirical model from Landsat data series for monitoring water salinity in coastal Bangladesh. *J. Environ. Manage.* 255, 109861. <https://doi.org/10.1016/j.jenvman.2019.109861>
- Ferdous, J., Rahman, M.T.U., Ghosh, S.K., 2019. Detection of Total Dissolved Solids from Landsat 8 OLI Image in Coastal Bangladesh, in: *The International Conference on Climate Change*, 1. pp. 35–44. <https://doi.org/10.17501/2513258X.2019.3103>
- Fichez, R., Archundia, D., Grenz, C., Douillet, P., Gutiérrez Mendieta, F., Origel Moreno, M., Denis, L., Contreras Ruiz Esparza, A., Zavala-Hidalgo, J., 2017. Global climate change and local watershed management as potential drivers of salinity variation in a tropical

References

- coastal lagoon (Laguna de Terminos, Mexico). *Aquat. Sci.* 79, 219–230.
<https://doi.org/10.1007/s00027-016-0492-1>
- Fooladi, M., Nikoo, M.R., Mirghafari, R., Madramootoo, C.A., Al-Rawas, G., Nazari, R., 2024. Robust clustering-based hybrid technique enabling reliable reservoir water quality prediction with uncertainty quantification and spatial analysis. *J. Environ. Manage.* 362. <https://doi.org/10.1016/j.jenvman.2024.121259>
- Franco, Emilio Granados, Melinda Kuritzky, Richard Lukacs, Saadia Zahidi, 2020. The global risks report 2020. World Economic Forum.
- Friedman, J.H., 2001. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* 29, 1189–1232.
- Fu, B., Li, S., Lao, Z., Wei, Y., Song, K., Deng, T., Wang, Y., 2024. A novel hierarchical approach to insight to spectral characteristics in surface water of karst wetlands and estimate its non-optically active parameters using field hyperspectral data. *Water Res.* 257, 121673. <https://doi.org/10.1016/j.watres.2024.121673>
- Gad, A.F., 2023. PyGAD: an intuitive genetic algorithm Python library. *Multimed. Tools Appl.* <https://doi.org/10.1007/s11042-023-17167-y>
- Gao, L., Shangguan, Y., Sun, Z., Shen, Q., Shi, Z., 2024. Estimation of Non-Optically Active Water Quality Parameters in Zhejiang Province Based on Machine Learning. *Remote Sens.* 16. <https://doi.org/10.3390/rs16030514>
- Ge, Y., Shen, F., Sklenička, P., Vymazal, J., Baxa, M., Chen, Z., 2024a. Machine learning for cyanobacteria inversion via remote sensing and AlgaeTorch in the Třeboň fishponds, Czech Republic. *Sci. Total Environ.* 947. <https://doi.org/10.1016/j.scitotenv.2024.174504>
- Ge, Y., Shen, F., Sklenička, P., Vymazal, J., Baxa, M., Chen, Z., 2024b. Machine learning for cyanobacteria inversion via remote sensing and AlgaeTorch in the Třeboň fishponds, Czech Republic. *Sci. Total Environ.* 947. <https://doi.org/10.1016/j.scitotenv.2024.174504>
- Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. *Anal. Chim. Acta* 185, 1–17. [https://doi.org/10.1016/0003-2670\(86\)80028-9](https://doi.org/10.1016/0003-2670(86)80028-9)
- Getirana, A., Libonati, R., Cataldi, M., 2021. Brazil is in water crisis — it needs a drought plan. *Nature* 600, 218–220. <https://doi.org/10.1038/d41586-021-03625-w>
- Ghadiri, H., 2016. Salinization of Karun River in Iran by shallow groundwater and seawater encroachment. *Adv. Hydro-Sci. -Eng.* 4, 1–9.
- Gholizadeh, M.H., Melesse, A.M., Reddi, L., 2016. A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. *Sensors* 16, 1298. <https://doi.org/10.3390/s16081298>
- Gilerson, A., Zhou, J., Hlaing, S., Ioannou, I., Gross, B., Moshary, F., Ahmed, S., 2008. Fluorescence component in the reflectance spectra from coastal waters. II. Performance of retrieval algorithms. *Opt. Express* 16, 2446–2460. <https://doi.org/10.1364/OE.16.002446>
- Global Modeling and Assimilation Office (GMAO), 2015. MERRA-2 tavg1_2d_aer_Nx: 2d,1-Hourly,Time-averaged,Single-Level,Assimilation,Aerosol Diagnostics V5.12.4. Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD, USA. <https://doi.org/10.5067/KLICLTZ8EM9D>
- Goldberg, D.E., Holland, J.H., 1988. Genetic Algorithms and Machine Learning. *Mach. Learn.* 2–3, 95–99. <https://doi.org/10.1023/A:1022602019183>
- Gong, S., Huang, J., Li, Y., Wang, H., 2008. Comparison of atmospheric correction algorithms for TM image in inland waters. *Int. J. Remote Sens.* 29, 2199–2210. <https://doi.org/10.1080/01431160701422262>
- Goodman, J.A., Lee, Z., Ustin, S.L., 2008. Influence of atmospheric and sea-surface corrections on retrieval of bottom depth and reflectance using a semi-analytical model: a case study in Kaneohe Bay, Hawaii. *Appl. Opt.* 47, F1–F11. <https://doi.org/10.1364/AO.47.0000F1>

References

- Gordon, H.R., 1978. Removal of atmospheric effects from satellite imagery of the oceans. *Appl. Opt.* 17, 1631–1636. <https://doi.org/10.1364/AO.17.001631>
- Gotway, C.A., Stroup, W.W., 1997. A Generalized Linear Model Approach to Spatial Data Analysis and Prediction. *J. Agric. Biol. Environ. Stat.* 2, 157–178. <https://doi.org/10.2307/1400401>
- Gray, P.C., Boss, E., Prochaska, J.X., Kerner, H., Demeaux, C.B., Lehahn, Y., 2024. The Promise and Pitfalls of Machine Learning in Ocean Remote Sensing.
- Greenblatt, G.D., Orlando, J.J., Burkholder, J.B., Ravishankara, A.R., 1990. Absorption measurements of oxygen between 330 and 1140 nm. *J. Geophys. Res. Atmos.* 95, 18577–18582. <https://doi.org/10.1029/JD095iD11p18577>
- Griffin, M.K., Burke, H.K., 2003. Compensation of Hyperspectral Data for Atmospheric Effects. *Linc. Lab. J.* 14, 29–54.
- Guo, H., Huang, J.J., Zhu, X., Tian, S., Wang, B., 2024. Spatiotemporal variation reconstruction of total phosphorus in the Great Lakes since 2002 using remote sensing and deep neural network. *Water Res.* 255. <https://doi.org/10.1016/j.watres.2024.121493>
- Guo, H., Zhu, X., Jeanne Huang, J., Zhang, Z., Tian, S., Chen, Y., 2023. An enhanced deep learning approach to assessing inland lake water quality and its response to climate and anthropogenic factors. *J. Hydrol.* 620. <https://doi.org/10.1016/j.jhydrol.2023.129466>
- Gurlin, D., 2012. Near Infrared-red Models for the Remote Estimation of Chlorophyll- α Concentration in Optically Complex Turbid Productive Waters: From in Situ Measurements to Aerial Imagery. *Sch. Nat. Resour. Diss. Theses Stud. Res.*
- Harmel, T., Chami, M., Tormos, T., Reynaud, N., Danis, P.-A., 2018a. Sun glint correction of the Multi-Spectral Instrument (MSI)-SENTINEL-2 imagery over inland and sea waters from SWIR bands. *Remote Sens. Environ.* 204, 308–321. <https://doi.org/10.1016/j.rse.2017.10.022>
- Harmel, T., Chami, M., Tormos, T., Reynaud, N., Danis, P.-A., 2018b. Sun glint correction of the Multi-Spectral Instrument (MSI)-SENTINEL-2 imagery over inland and sea waters from SWIR bands. *Remote Sens. Environ.* 204, 308–321. <https://doi.org/10.1016/j.rse.2017.10.022>
- Hastie, T.J., 1992. Generalized Additive Models, in: *Statistical Models in S*. Routledge.
- Hedley, J.D., Harborne, A.R., Mumby, P.J., 2005. Technical note: Simple and robust removal of sun glint for mapping shallow-water benthos. *Int. J. Remote Sens.* 26, 2107–2112. <https://doi.org/10.1080/01431160500034086>
- Hieronymi, M., Krasemann, H., Mueller, D., Brockmann, C., Ruescas, A., Stelzer, K., Nechad, B., Ruddick, K., Simis, S., Tilstone, G., Steinmetz, F., Regner, P., 2016. Ocean Colour Remote Sensing of Extreme Case-2 Waters. *Proc. ESA Living Planet Symp.* 2016 SP-740, Sco 3-1.
- Higazi, S.F., Abdel-Hady, D.H., Al-Oulfi, S.A., 2013. Application of Spatial Regression Models to Income Poverty Ratios in Middle Delta Contiguous Counties in Egypt. *Pak. J. Stat. Oper. Res.* 9, 93. <https://doi.org/10.18187/pjsor.v9i1.272>
- Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. *Neural Comput.* 9, 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hoerl, A.E., Kennard, R.W., 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* 12, 55–67. <https://doi.org/10.1080/00401706.1970.10488634>
- Holland, J.H., 1992. Genetic Algorithms. *Sci. Am.* 267, 66–73.
- Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. *Reliab. Eng. Syst. Saf.* 52, 1–17. [https://doi.org/10.1016/0951-8320\(96\)00002-6](https://doi.org/10.1016/0951-8320(96)00002-6)
- Hossain, A.K.M.A., Chao, X., Jia, Y., 2010. Development of Remote Sensing Based Index for Estimating/Mapping Suspended Sediment Concentration in River and Lake Environments, in: *Proceedings of 8th International Symposium on ECOHYDRAULICS (ISE 2010)*. pp. 578–585.

References

- Hou, Y., Zhang, A., Lv, R., Zhao, S., Ma, J., Zhang, H., Li, Z., 2022. A study on water quality parameters estimation for urban rivers based on ground hyperspectral remote sensing technology. *Environ. Sci. Pollut. Res.* 29, 63640–63654. <https://doi.org/10.1007/s11356-022-20293-z>
- Hu, C., Carder, K.L., Muller-Karger, F.E., 2000. Atmospheric Correction of SeaWiFS Imagery over Turbid Coastal Waters: A Practical Method. *Remote Sens. Environ.* 74, 195–206. [https://doi.org/10.1016/S0034-4257\(00\)00080-8](https://doi.org/10.1016/S0034-4257(00)00080-8)
- Hu, C., Chen, Z., Clayton, T.D., Swarzenski, P., Brock, J.C., Muller-Karger, F.E., 2004. Assessment of estuarine water-quality indicators using MODIS medium-resolution bands: Initial results from Tampa Bay, FL. *Remote Sens. Environ.* 93, 423–441. <https://doi.org/10.1016/j.rse.2004.08.007>
- IOCCG, 2026. A Scientific Roadmap of Aquatic Hyperspectral Remote Sensing: Overview of Status, Challenges and Future Perspectives, Version 1.0 (No. 22). International Ocean Colour Coordinating Group (IOCCG), Dartmouth, NS, Canada. <https://doi.org/https://doi.org/10.25607/OBP-2062>
- IOCCG, 2019. Uncertainties in Ocean Colour Remote Sensing (No. 18), IOCCG Report Series. International Ocean Colour Coordinating Group, Dartmouth, Canada. <https://doi.org/http://dx.doi.org/10.25607/OBP696>
- IOCCG, 2018. Earth Observations in Support of Global Water Quality Monitoring (No. 17), IOCCG Report Series. International Ocean Colour Coordinating Group, Dartmouth, Canada.
- IOCCG, 2012a. Mission Requirements for Future Ocean-Colour Sensors (No. 13), IOCCG Report Series. International Ocean Colour Coordinating Group, Dartmouth, Canada.
- IOCCG, 2012b. Ocean-Colour Observations from a Geostationary Orbit (No. 12), IOCCG Report Series. International Ocean Colour Coordinating Group, Dartmouth, Canada.
- IOCCG, 2010. Atmospheric Correction for Remotely-Sensed Ocean-Colour Products (No. 10), IOCCG Report Series. International Ocean Colour Coordinating Group, Dartmouth, Canada.
- Iran Water and Power Resources Development Company [WWW Document], 2023. URL <https://iwpc.ir/?l=EN> (accessed 11.23.23).
- Jalali, L., Zarei, M., Gutiérrez, F., 2019. Salinization of reservoirs in regions with exposed evaporites. The unique case of Upper Gotvand Dam, Iran. *Water Res.* 157, 587–599. <https://doi.org/10.1016/j.watres.2019.04.015>
- James, K.R., Cant, B., Ryan, T., 2003. Responses of freshwater biota to rising salinity levels and implications for saline water management: a review. *Aust. J. Bot.* 51, 703–713. <https://doi.org/10.1071/bt02110>
- Jang, J.-S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* 23, 665–685. <https://doi.org/10.1109/21.256541>
- Jiang, D., Dong, C., Ma, Z., Wang, X., Lin, K., Yang, F., Chen, X., 2024. Monitoring saltwater intrusion to estuaries based on UAV and satellite imagery with machine learning models. *Remote Sens. Environ.* 308, 114198. <https://doi.org/10.1016/j.rse.2024.114198>
- Jorge, D.S.F., Barbosa, C.C.F., De Carvalho, L.A.S., Affonso, A.G., Lobo, F.D.L., Novo, E.M.L.D.M., 2017. SNR (Signal-To-Noise Ratio) Impact on Water Constituent Retrieval from Simulated Images of Optically Complex Amazon Lakes. *Remote Sens.* 9, 644. <https://doi.org/10.3390/rs9070644>
- Jualaong, S., Songnui, A., Thongprajukaew, K., Ninwat, S., Khwanmaung, S., Hahor, W., Khunsaeng, P., Kanghae, H., 2019. Optimal Salinity for Head-Starting Northern River Terrapins (*Batagur baska* Gray, 1831). *Animals* 9, 855. <https://doi.org/10.3390/ani9110855>

References

- Kao, H.-Y., Lagerloef, G.S.E., Lee, T., Melnichenko, O., Meissner, T., Hacker, P., 2018. Assessment of Aquarius Sea Surface Salinity. *Remote Sens.* 10, 1341. <https://doi.org/10.3390/rs10091341>
- Karpouzli, E., Malthus, T., 2003. The empirical line method for the atmospheric correction of IKONOS imagery. *Int. J. Remote Sens.* 24, 1143–1150. <https://doi.org/10.1080/0143116021000026779>
- Kay, S., Hedley, J.D., Lavender, S., 2009. Sun Glint Correction of High and Low Spatial Resolution Images of Aquatic Scenes: a Review of Methods for Visible and Near-Infrared Wavelengths. *Remote Sens.* 1, 697–730. <https://doi.org/10.3390/rs1040697>
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree, in: *Advances in Neural Information Processing Systems*. Curran Associates, Inc.
- Kelly, M., 2019a. The Standard Errors of Persistence.
- Kelly, M., 2019b. The Standard Errors of Persistence. [WWW Document]. SSRN. URL <https://papers.ssrn.com/abstract=3398303> (accessed 5.27.24).
- Kerr, Y.H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J., Reul, N., Gruhier, C., Juglea, S.E., Drinkwater, M.R., Hahne, A., Martín-Neira, M., Mecklenburg, S., 2010. The SMOS Mission: New Tool for Monitoring Key Elements of the Global Water Cycle. *Proc. IEEE* 98, 666–687. <https://doi.org/10.1109/JPROC.2010.2043032>
- Khadim, F.K., Su, H., Xu, L., 2017. A spatially weighted optimization model (SWOM) for salinity mapping in Florida Bay using Landsat images and in situ observations. *Phys. Chem. Earth Parts ABC, Physics and Economics of Ecosystem Services Flows* 101, 86–101. <https://doi.org/10.1016/j.pce.2017.06.001>
- Khan, N.M., Rastoskuev, V.V., Sato, Y., Shiozawa, S., 2005. Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agric. Water Manag., Special Issue on Land and Water Use: Environmental Management Tools and Practices* 77, 96–109. <https://doi.org/10.1016/j.agwat.2004.09.038>
- Khan, N.M., Rastoskuev, V.V., Shalina, E.V., Sato, Y., 2001. Mapping Salt-affected Soils Using Remote Sensing Indicators - A Simple Approach With the Use of GIS IDRISI. Presented at the 22nd Asian Conference on Remote Sensing, Singapore.
- Khan, R.M., Salehi, B., Niroumand-Jadidi, M., Mahdianpari, M., 2024. Mapping Water Clarity in Small Oligotrophic Lakes Using Sentinel-2 Imagery and Machine Learning Methods: A Case Study of Canandaigua Lake in Finger Lakes, New York. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 17, 4674–4688. <https://doi.org/10.1109/JSTARS.2024.3359648>
- Khan, Rabia Munsaf, Salehi, B., Niroumand-Jadidi, M., Mahdianpari, M., 2024. Global vs Local Random Forest Model for Water Quality Monitoring: Assessment in Finger Lakes Using Sentinel-2 Imagery and Gloria Dataset. Presented at the IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, pp. 4389–4392. <https://doi.org/10.1109/IGARSS53475.2024.10641536>
- Khorram, S., 1985. Development of water quality models applicable throughout the entire San Francisco Bay and Delta. *Photogramm. Eng. Remote Sens.* 51, 53–62.
- Klemas, V., 2011. Remote Sensing of Algal Blooms: An Overview with Case Studies. *J. Coast. Res.* 28, 34–43. <https://doi.org/10.2112/JCOASTRES-D-11-00051.1>
- Knudby, A., Richardson, G., 2023. Incorporation of neighborhood information improves performance of SDB models. *Remote Sens. Appl. Soc. Environ.* 32, 101033. <https://doi.org/10.1016/j.rsase.2023.101033>
- Koldasbayeva, D., Tregubova, P., Gasanov, M., Zaytsev, A., Petrovskaya, A., Burnaev, E., 2024. Challenges in data-driven geospatial modeling for environmental research and practice. *Nat. Commun.* 15, 10700. <https://doi.org/10.1038/s41467-024-55240-8>

References

- Kuhn, M., Weston, S., Keefer, C., Coulter, N., Quinlan, R., 2023. Cubist: Rule- And Instance-Based Regression Modeling.
- Kutser, T., Vahtmäe, E., Praks, J., 2009. A sun glint correction method for hyperspectral imagery containing areas with non-negligible water leaving NIR signal. *Remote Sens. Environ.* 113, 2267–2274. <https://doi.org/10.1016/j.rse.2009.06.016>
- Kwon, S., Noh, H., Seo, I.W., Park, Y.S., 2023. Effects of spectral variability due to sediment and bottom characteristics on remote sensing for suspended sediment in shallow rivers. *Sci. Total Environ.* 878, 163125. <https://doi.org/10.1016/j.scitotenv.2023.163125>
- Lagerloef, G., Colomb, F.R., Le Vine, D., Wentz, F., Yueh, S., Ruf, C., Lilly, J., Gunn, J., Chao, Y., Decharon, A., Feldman, G., Swift, C., 2008. The Aquarius/Sac-D Mission: Designed to Meet the Salinity Remote-Sensing Challenge. *Oceanography* 21, 68–81.
- Lavery, P., Pattiaratchi, C., Wyllie, A., Hick, P., 1993. Water quality monitoring in estuarine waters using the Landsat thematic mapper. *Remote Sens. Environ.* 46, 268–280. [https://doi.org/10.1016/0034-4257\(93\)90047-2](https://doi.org/10.1016/0034-4257(93)90047-2)
- Leggesse, E., Zimale, F., Sultan, D., Enku, T., Tilahun, S.A., 2024. Advancing non-optical water quality monitoring in Lake Tana, Ethiopia: insights from machine learning and remote sensing techniques. *Front. Water* 6, 1432280. <https://doi.org/10.3389/frwa.2024.1432280>
- Lehmann, M.K., Gurlin, D., Pahlevan, N., Alikas, K., Conroy, T., Anstee, J., Balasubramanian, S.V., Barbosa, C.C.F., Binding, C., Bracher, A., Bresciani, M., Burtner, A., Cao, Z., Dekker, A.G., Di Vittorio, C., Drayson, N., Errera, R.M., Fernandez, V., Ficek, D., Fichot, C.G., Gege, P., Giardino, C., Gitelson, A.A., Greb, S.R., Henderson, H., Higa, H., Rahaghi, A.I., Jamet, C., Jiang, D., Jordan, T., Kangro, K., Kravitz, J.A., Kristoffersen, A.S., Kudela, R., Li, L., Ligi, M., Loisel, H., Lohrenz, S., Ma, R., Maciel, D.A., Malthus, T.J., Matsushita, B., Matthews, M., Minaudo, C., Mishra, D.R., Mishra, S., Moore, T., Moses, W.J., Nguyễn, H., Novo, E.M.L.M., Novoa, S., Odermatt, D., O'Donnell, D.M., Olmanson, L.G., Ondrusek, M., Oppelt, N., Ouillon, S., Pereira Filho, W., Plattner, S., Verdú, A.R., Salem, S.I., Schalles, J.F., Simis, S.G.H., Siswanto, E., Smith, B., Somlai-Schweiger, I., Soppa, M.A., Spyarakos, E., Tessin, E., van der Woerd, H.J., Vander Woude, A., Vandermeulen, R.A., Vantrepotte, V., Wernand, M.R., Werther, M., Young, K., Yue, L., 2023. GLORIA - A globally representative hyperspectral in situ dataset for optical sensing of water quality. *Sci. Data* 10, 100. <https://doi.org/10.1038/s41597-023-01973-y>
- Li, Z., Yang, X., Zhou, T., Cai, S., Zhang, W., Mao, K., Ou, H., Ran, L., Yang, Q., Wang, Y., 2024. Monitoring Chlorophyll-a Concentration Variation in Fish Ponds from 2013 to 2022 in the Guangdong-Hong Kong-Macao Greater Bay Area, China. *Remote Sens.* 16. <https://doi.org/10.3390/rs16112033>
- Li, Z., Zhang, F., Shi, J., Chan, N.W., Tan, M.L., Kung, H., Liu, C., Cheng, C., Cai, Y., Wang, W., Li, X., 2023. Remote sensing for chromophoric dissolved organic matter (CDOM) monitoring research 2003–2022: A bibliometric analysis based on the web of science core database. *Mar. Pollut. Bull.* 196, 115653. <https://doi.org/10.1016/j.marpolbul.2023.115653>
- Likens, G.E., 2009. Inland Waters, in: *Encyclopedia of Inland Waters*. Academic Press, Oxford, pp. 1–5.
- Lintern, A., Webb, J. a., Ryu, D., Liu, S., Bende-Michl, U., Waters, D., Leahy, P., Wilson, P., Western, A.W., 2018. Key factors influencing differences in stream water quality across space. *Wiley Interdiscip. Rev. Water* 5, e1260. <https://doi.org/10.1002/wat2.1260>
- Liu, T., Dai, J., Zhao, Y., Tian, S., Nie, Z., Ye, C., 2023. Using remote sensing technology to monitor salt lake changes caused by climate change and melting glaciers: insights from Zabuye Salt Lake in Xizang. *J. Oceanol. Limnol.* 41, 1258–1276. <https://doi.org/10.1007/s00343-022-2138-6>

References

- Liu, W., Wang, S., Yang, R., Ma, Y., Shen, M., You, Y., Hai, K., Baqa, M.F., 2019. Remote Sensing Retrieval of Turbidity in Alpine Rivers based on high Spatial Resolution Satellites. *Remote Sens.* 11, 3010. <https://doi.org/10.3390/rs11243010>
- Liu, X., Kounadi, O., Zurita-Milla, R., 2022. Incorporating Spatial Autocorrelation in Machine Learning Models Using Spatial Lag and Eigenvector Spatial Filtering Features. *ISPRS Int. J. Geo-Inf.* 11, 242. <https://doi.org/10.3390/ijgi11040242>
- Llodrà-Llabrés, J., Martínez-López, J., Postma, T., Pérez-Martínez, C., Alcaraz-Segura, D., 2023. Retrieving water chlorophyll-*a* concentration in inland waters from Sentinel-2 imagery: Review of operability, performance and ways forward. *Int. J. Appl. Earth Obs. Geoinformation* 125, 103605. <https://doi.org/10.1016/j.jag.2023.103605>
- Lu, Y., Chen, Q., Yu, M., Wu, Z., Huang, C., Fu, J., Yu, Z., Yao, J., 2023. Exploring spatial and environmental heterogeneity affecting energy consumption in commercial buildings using machine learning. *Sustain. Cities Soc.* 95, 104586. <https://doi.org/10.1016/j.scs.2023.104586>
- Lyzenga, D.R., Malinas, N.P., Tanis, F.J., 2006. Multispectral bathymetry using a simple physically based algorithm. *IEEE Trans. Geosci. Remote Sens.* 44, 2251–2259. <https://doi.org/10.1109/TGRS.2006.872909>
- Mabula, M.J., Kisanga, D., Pamba, S., 2023. Application of machine learning algorithms and Sentinel-2 satellite for improved bathymetry retrieval in Lake Victoria, Tanzania. *Egypt. J. Remote Sens. Space Sci.* 26, 619–627. <https://doi.org/10.1016/j.ejrs.2023.07.003>
- Mahboobi, H., Shakiba, A., Mirbagheri, B., 2023. Improving groundwater nitrate concentration prediction using local ensemble of machine learning models. *J. Environ. Manage.* 345, 118782. <https://doi.org/10.1016/j.jenvman.2023.118782>
- Mainali, J., Chang, H., Chun, Y., 2019. A review of spatial statistical approaches to modeling water quality. *Prog. Phys. Geogr. Earth Environ.* 43, 801–826. <https://doi.org/10.1177/0309133319852003>
- Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., Gascon, F., 2017. Sen2Cor for Sentinel-2. Presented at the Image and Signal Processing for Remote Sensing XXIII, SPIE, pp. 37–48. <https://doi.org/10.1117/12.2278218>
- Malahlela, O.E., 2019. Spatio-temporal assessment of inland surface water quality using remote sensing data in the wake of changing climate, in: *IOP Conference Series: Earth and Environmental Science*, 6. IOP Publishing, p. 062012. <https://doi.org/10.1088/1755-1315/227/6/062012>
- Maliki, A.A.A., Chabuk, A., Sultan, M.A., Hashim, B.M., Hussain, H.M., Al-Ansari, N., 2020. Estimation of Total Dissolved Solids in Water Bodies by Spectral Indices Case Study: Shatt al-Arab River. *Water. Air. Soil Pollut.* 231, 482. <https://doi.org/10.1007/s11270-020-04844-z>
- Mamun, M., Hasan, M., An, K.-G., 2024. Advancing reservoirs water quality parameters estimation using Sentinel-2 and Landsat-8 satellite data with machine learning approaches. *Ecol. Inform.* 81. <https://doi.org/10.1016/j.ecoinf.2024.102608>
- Mancino, G., Nolè, A., Urbano, V., Amato, M., Ferrara, A., 2009. Assessing water quality by remote sensing in small lakes: the case study of Monticchio lakes in southern Italy. *IForest - Biogeosciences For.* 2, 154. <https://doi.org/10.3832/ifor0507-002>
- Matthews, M.W., 2011. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *Int. J. Remote Sens.* 32, 6855–6899. <https://doi.org/10.1080/01431161.2010.512947>
- McClain, C.R., Cleave, M.L., Feldman, G.C., Gregg, W.W., Hooker, S.B., Kuring, N., 1998. Science quality SeaWiFS data for global biosphere research: Ocean color is back, providing global estimates of oceanic chlorophyll-*a*, other bio-optical quantities - Data critical for understanding temporal variability of marine ecosystems. *Sea Technol.* 39, 10–16.

References

- Medina-Lopez, E., 2020. Machine Learning and the End of Atmospheric Corrections: A Comparison between High-Resolution Sea Surface Salinity in Coastal Areas from Top and Bottom of Atmosphere Sentinel-2 Imagery. *Remote Sens.* 12, 2924. <https://doi.org/10.3390/rs12182924>
- Medina-Lopez, E., Ureña-Fuentes, L., 2019. High-Resolution Sea Surface Temperature and Salinity in Coastal Areas Worldwide from Raw Satellite Data. *Remote Sens.* 11, 2191. <https://doi.org/10.3390/rs11192191>
- Mejía Ávila, D., Torres-Bejarano, F., Martínez Lara, Z., 2022. Spectral indices for estimating total dissolved solids in freshwater wetlands using semi-empirical models. A case study of Guartínaja and Momil wetlands. *Int. J. Remote Sens.* 43, 2156–2184. <https://doi.org/10.1080/01431161.2022.2057205>
- Menze, B.H., Kelm, B.M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., Hamprecht, F.A., 2009. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics* 10, 213. <https://doi.org/10.1186/1471-2105-10-213>
- Merriman, J., 1997. Water quality in the St. Lawrence River at Wolfe Island (No. EHD 97-01/I). Environment Canada.
- Metulini, R., Patuelli, R., Griffith, D.A., 2018. A Spatial-Filtering Zero-Inflated Approach to the Estimation of the Gravity Model of Trade. *Econometrics* 6, 9. <https://doi.org/https://doi.org/10.3390/econometrics6010009>
- Minor, E.C., Pothén, J., Dalzell, B.J., Abdulla, H., Mopper, K., 2006. Effects of salinity changes on the photodegradation and ultraviolet—visible absorbance of terrestrial dissolved organic matter. *Limnol. Oceanogr.* 51, 2181–2186. <https://doi.org/10.4319/lo.2006.51.5.2181>
- Mishra, D.R., Ogashawara, I., Gitelson, A.A., 2017. Bio-optical Modeling and Remote Sensing of Inland Waters. Elsevier.
- Mobley, C.D., Werdell, J., Franz, B., Ahmad, Z., Bailey, S., 2016. Atmospheric Correction for Satellite Ocean Color Radiometry (No. GSFC-E-DAA-TN35509).
- Montalvo, L.G., 2010. Spectral analysis of suspended material in coastal waters: A comparison between band math equations. *Dep. Geol. Univ. P. R. Mayaguez PR USA* 1–6.
- Moradi-Majd, N., Fallah-Ghalhari, G., Chatrenor, M., 2022. Estimation of greenhouse gas emission flux from agricultural lands of Khuzestan province in Iran. *Environ. Monit. Assess.* 194, 811. <https://doi.org/10.1007/s10661-022-10497-8>
- Moran, P.A.P., 1950. Notes on Continuous Stochastic Phenomena. *Biometrika* 37, 17–23. <https://doi.org/10.2307/2332142>
- Morley, S.K., Brito, T.V., Welling, D.T., 2018. Measures of Model Performance Based On the Log Accuracy Ratio. *Space Weather* 16, 69–88. <https://doi.org/10.1002/2017SW001669>
- Moses, W.J., Bowles, J.H., Lucke, R.L., Corson, M.R., 2012a. Impact of signal-to-noise ratio in a hyperspectral sensor on the accuracy of biophysical parameter estimation in case II waters. *Opt. Express* 20, 4309–4330. <https://doi.org/10.1364/OE.20.004309>
- Moses, W.J., Gitelson, A.A., Perk, R.L., Gurlin, D., Rundquist, D.C., Leavitt, B.C., Barrow, T.M., Brakhage, P., 2012b. Estimation of chlorophyll-*a* concentration in turbid productive waters using airborne hyperspectral data. *Water Res.* 46, 993–1004. <https://doi.org/10.1016/j.watres.2011.11.068>
- Mueller, J.T., Gasteyer, S., 2021. The widespread and unjust drinking water and clean water crisis in the United States. *Nat. Commun.* 12, 3544. <https://doi.org/10.1038/s41467-021-23898-z>
- Mushtaq, F., Nee Lala, M.G., 2017. Remote estimation of water quality parameters of Himalayan lake (Kashmir) using Landsat 8 OLI imagery. *Geocarto Int.* 32, 274–285. <https://doi.org/10.1080/10106049.2016.1140818>

References

- Nahvi, A., Daghighi, A., Nazif, S., 2018. The environmental impact assessment of drainage systems: a case study of the Karun river sugarcane development project. *Arch. Agron. Soil Sci.* 64, 185–195. <https://doi.org/10.1080/03650340.2017.1340641>
- National Research Council, 2002. *The Drama of the Commons*. National Academies Press, Washington, D.C., USA. <https://doi.org/10.17226/10287>
- Nazeer, M., Bilal, M., 2018. Evaluation of Ordinary Least Square (OLS) and Geographically Weighted Regression (GWR) for Water Quality Monitoring: A Case Study for the Estimation of Salinity. *J. Ocean Univ. China* 17, 305–310. <https://doi.org/10.1007/s11802-018-3380-6>
- Nguyen, P.T.B., Koedsin, W., McNeil, D., Van, T.P.D., 2018. Remote sensing techniques to predict salinity intrusion: application for a data-poor area of the coastal Mekong Delta, Vietnam. *Int. J. Remote Sens.* 39, 6676–6691. <https://doi.org/10.1080/01431161.2018.1466071>
- Nguyen, X.C., Bui, V.K.H., Cho, K.H., Hur, J., 2024. Practical application of machine learning for organic matter and harmful algal blooms in freshwater systems: A review. *Crit. Rev. Environ. Sci. Technol.* 54, 953–975. <https://doi.org/10.1080/10643389.2023.2285691>
- Nikoo, M.R., Zamani, M.G., Zadeh, M.M., Al-Rawas, G., Al-Wardy, M., Gandomi, A.H., 2024. Mapping reservoir water quality from Sentinel-2 satellite data based on a new approach of weighted averaging: Application of Bayesian maximum entropy. *Sci. Rep.* 14. <https://doi.org/10.1038/s41598-024-66699-2>
- Nossent, J., Elsen, P., Bauwens, W., 2011. Sobol’ sensitivity analysis of a complex environmental model. *Environ. Model. Softw.* 26, 1515–1525. <https://doi.org/10.1016/j.envsoft.2011.08.010>
- Nouraki, A., Alavi, M., Golabi, M., Albaji, M., 2021. Prediction of water quality parameters using machine learning models: a case study of the Karun River, Iran. *Environ. Sci. Pollut. Res.* 28, 57060–57072. <https://doi.org/10.1007/s11356-021-14560-8>
- OBPG., 2023. Ancillary Data. [WWW Document]. NASA Ocean Biol. Process. Group. URL <https://oceancolor.gsfc.nasa.gov/resources/docs/ancillary/>
- Olmanson, L.G., Brezonik, P.L., Bauer, M.E., 2015. Remote Sensing for Regional Lake Water Quality Assessment: Capabilities and Limitations of Current and Upcoming Satellite Systems, in: Younos, T., Parece, T.E. (Eds.), *Advances in Watershed Science and Assessment*. Springer International Publishing, Cham, pp. 111–140. https://doi.org/10.1007/978-3-319-14212-8_5
- Oshan, T.M., Li, Z., Kang, W., Wolf, L.J., Fotheringham, A.S., 2019. mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *ISPRS Int. J. Geo-Inf.* 8, 269. <https://doi.org/10.3390/ijgi8060269>
- O’Shea, R.E., Pahlevan, N., Smith, B., Bresciani, M., Egerton, T., Giardino, C., Li, L., Moore, T., Ruiz-Verdu, A., Ruberg, S., Simis, S.G.H., Stumpf, R., Vaičiūtė, D., 2021. Advancing cyanobacteria biomass estimation from hyperspectral observations: Demonstrations with HICO and PRISMA imagery. *Remote Sens. Environ.* 266. <https://doi.org/10.1016/j.rse.2021.112693>
- Ouma, Y.O., 2020. Modelling reservoir turbidity from medium resolution Sentinel-2A/MSI and Landsat-8/OLI satellite imagery, in: *Remote Sensing for Agriculture, Ecosystems, and Hydrology XXII*. Presented at the Remote Sensing for Agriculture, Ecosystems, and Hydrology XXII, SPIE, pp. 111–117. <https://doi.org/10.1117/12.2579319>
- Pahlevan, N., Lee, Z., Wei, J., Schaaf, C.B., Schott, J.R., Berk, A., 2014. On-orbit radiometric characterization of OLI (Landsat-8) for applications in aquatic remote sensing. *Remote Sens. Environ.* 154, 272–284. <https://doi.org/10.1016/j.rse.2014.08.001>
- Pahlevan, N., Mangin, A., Balasubramanian, S.V., Smith, B., Alikas, K., Arai, K., Barbosa, C., Bélanger, S., Binding, C., Bresciani, M., Giardino, C., Gurlin, D., Fan, Y., Harmel, T.,

References

- Hunter, P., Ishikaza, J., Kratzer, S., Lehmann, M.K., Ligi, M., Ma, R., Martin-Lauzer, F.-R., Olmanson, L., Oppelt, N., Pan, Y., Peters, S., Reynaud, N., Sander de Carvalho, L.A., Simis, S., Spyrakos, E., Steinmetz, F., Stelzer, K., Sterckx, S., Tormos, T., Tyler, A., Vanhellemont, Q., Warren, M., 2021. ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. *Remote Sens. Environ.* 258, 112366. <https://doi.org/10.1016/j.rse.2021.112366>
- Pahlevan, N., Schott, J.R., Franz, B.A., Zibordi, G., Markham, B., Bailey, S., Schaaf, C.B., Ondrusek, M., Greb, S., Strait, C.M., 2017. Landsat 8 remote sensing reflectance (Rrs) products: Evaluations, intercomparisons, and enhancements. *Remote Sens. Environ.* 190, 289–301. <https://doi.org/10.1016/j.rse.2016.12.030>
- Pahlevan, N., Smith, B., Alikas, K., Anstee, J., Barbosa, C., Binding, C., Bresciani, M., Cremella, B., Giardino, C., Gurlin, D., Fernandez, V., Jamet, C., Kangro, K., Lehmann, M.K., Loisel, H., Matsushita, B., Hà, N., Olmanson, L., Potvin, G., Simis, S.G.H., VanderWoude, A., Vantrepotte, V., Ruiz-Verdù, A., 2022. Simultaneous retrieval of selected optical water quality indicators from Landsat-8, Sentinel-2, and Sentinel-3. *Remote Sens. Environ.* 270, 112860. <https://doi.org/10.1016/j.rse.2021.112860>
- Pahlevan, N., Smith, B., Schalles, J., Binding, C., Cao, Z., Ma, R., Alikas, K., Kangro, K., Gurlin, D., Hà, N., Matsushita, B., Moses, W., Greb, S., Lehmann, M.K., Ondrusek, M., Oppelt, N., Stumpf, R., 2020. Seamless retrievals of chlorophyll-*a* from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach. *Remote Sens. Environ.* 240, 111604. <https://doi.org/10.1016/j.rse.2019.111604>
- Pal, M., 2009. Kernel Methods in Remote Sensing: A Review. *ISH J. Hydraul. Eng.* 15, 194–215. <https://doi.org/10.1080/09715010.2009.10514975>
- Pan, Y., Bélanger, S., 2025. Genetic Algorithm for Atmospheric Correction (GAAC) of water bodies impacted by adjacency effects. *Remote Sens. Environ.* 317, 114508. <https://doi.org/10.1016/j.rse.2024.114508>
- Pan, Y., Bélanger, S., Huot, Y., 2022. Evaluation of Atmospheric Correction Algorithms over Lakes for High-Resolution Multispectral Imagery: Implications of Adjacency Effect. *Remote Sens.* 14, 2979. <https://doi.org/10.3390/rs14132979>
- Paoletti, M.E., Haut, J.M., Plaza, J., Plaza, A., 2019. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS J. Photogramm. Remote Sens.* 158, 279–317. <https://doi.org/10.1016/j.isprsjprs.2019.09.006>
- Park, J., Khanal, S., Zhao, K., Byun, K., 2024. Remote Sensing of Chlorophyll-*a* and Water Quality over Inland Lakes: How to Alleviate Geo-Location Error and Temporal Discrepancy in Model Training. *Remote Sens.* 16, 2761. <https://doi.org/10.3390/rs16152761>
- Parmar, T.K., Rawtani, D., Agrawal, Y.K., 2016. Bioindicators: the natural indicator of environmental pollution. *Front. Life Sci.* 9, 110–118. <https://doi.org/10.1080/21553769.2016.1162753>
- Paulino, R.S., Martins, V.S., Novo, E.M.L.M., Barbosa, C.C.F., de Carvalho, L.A.S., Begliomini, F.N., 2022. Assessment of Adjacency Correction over Inland Waters Using Sentinel-2 MSI Images. *Remote Sens.* 14, 1829. <https://doi.org/10.3390/rs14081829>
- Peng, C., Xie, Z., Jin, X., 2024. Using Ensemble Learning for Remote Sensing Inversion of Water Quality Parameters in Poyang Lake. *Sustain. Switz.* 16. <https://doi.org/10.3390/su16083355>
- Peralta, N.R., Assefa, Y., Du, J., Barden, C.J., Ciampitti, I.A., 2016. Mid-Season High-Resolution Satellite Imagery for Forecasting Site-Specific Corn Yield. *Remote Sens.* 8, 848. <https://doi.org/10.3390/rs8100848>
- Pereira-Sandoval, M., Ruescas, A., Urrego, P., Ruiz-Verdú, A., Delegido, J., Tenjo, C., Soria-Perpinyà, X., Vicente, E., Soria, J., Moreno, J., 2019. Evaluation of Atmospheric

References

- Correction Algorithms over Spanish Inland Waters for Sentinel-2 Multi Spectral Imagery Data. *Remote Sens.* 11, 1469. <https://doi.org/10.3390/rs11121469>
- Piszcz, A., Soule, T., 2006. Genetic Programming: Analysis of Optimal Mutation Rates in a Problem with Varying Difficulty. Presented at the FLAIRS conference, pp. 451–456.
- Priyadarshini, R., Sudhakara, B., Sowmya Kamath, S., Bhattacharjee, S., Pruthviraj, U., Gangadharan, K.V., 2023. Water Salinity Assessment Using Remotely Sensed Images—A Comprehensive Survey, in: *Computer Vision and Machine Intelligence, Lecture Notes in Networks and Systems*. Springer Nature, Singapore, pp. 577–589. https://doi.org/10.1007/978-981-19-7867-8_46
- Qi, L., Lee, Z., Hu, C., Wang, M., 2017. Requirement of minimal signal-to-noise ratios of ocean color sensors and uncertainties of ocean color products. *J. Geophys. Res. Oceans* 122, 2595–2611. <https://doi.org/10.1002/2016JC012558>
- Quinlan, J.R., 1992. Learning with continuous classes, in: 5th Australian Joint Conference on Artificial Intelligence. Presented at the 5th Australian joint conference on artificial intelligence, pp. 343–348.
- Radin, C., Sòria-Perpinyà, X., Delegido, J., 2020. Estudio multitemporal de calidad del agua del embalse de Sitjar (Castelló, España) utilizando imágenes Sentinel-2. *Rev. Teledetec.* 117–130. <https://doi.org/10.4995/raet.2020.13864>
- Rey, S.J., Anselin, L., 2010. PySAL: A Python Library of Spatial Analytical Methods, in: Fischer, M.M., Getis, A. (Eds.), *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*. Springer, Berlin, Heidelberg, pp. 175–193. https://doi.org/10.1007/978-3-642-03647-7_11
- Richardson, G., Knudby, A., Wu, Y., Ansari, M., 2025. A case study comparing approaches to mask satellite-derived bathymetry. *Discov. Geosci.* 3, 103. <https://doi.org/10.1007/s44288-025-00219-1>
- Richter, R., Schlapfer, D., 2023. Atmospheric and Topographic Correction (ATCOR theoretical background document).
- Rolf, E., 2023. Evaluation Challenges for Geospatial ML. <https://doi.org/10.48550/arXiv.2303.18087>
- Rolf, E., Proctor, J., Carleton, T., Bolliger, I., Shankar, V., Ishihara, M., Recht, B., Hsiang, S., 2021. A generalizable and accessible approach to machine learning with global satellite imagery. *Nat. Commun.* 12, 4392. <https://doi.org/10.1038/s41467-021-24638-z>
- Röttgers, R., McKee, D., Utschig, C., 2014. Temperature and salinity correction coefficients for light absorption by water in the visible to infrared spectral region. *Opt. Express* 22, 25093–25108. <https://doi.org/10.1364/OE.22.025093>
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ.* 351, 309–317.
- Ruddick, K.G., De Cauwer, V., Park, Y.-J., Moore, G., 2006. Seaborne measurements of near infrared water-leaving reflectance: The similarity spectrum for turbid waters. *Limnol. Oceanogr.* 51, 1167–1179. <https://doi.org/10.4319/lo.2006.51.2.1167>
- Ruddick, K.G., Ovidio, F., Rijkeboer, M., 2000. Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Appl. Opt.* 39, 897–912. <https://doi.org/10.1364/AO.39.000897>
- Rusydi, A.F., 2018. Correlation between conductivity and total dissolved solid in various type of water: A review, in: *IOP Conference Series: Earth and Environmental Science*. IOP publishing, p. 012019. <https://doi.org/10.1088/1755-1315/118/1/012019>
- Sagan, V., Peterson, K.T., Maimaitijiang, M., Sidike, P., Sloan, J., Greeling, B.A., Maalouf, S., Adams, C., 2020. Monitoring inland water quality using remote sensing: potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Sci. Rev.* 205, 103187. <https://doi.org/10.1016/j.earscirev.2020.103187>

References

- Salama, M.S., Spaias, L., Poser, K., Peters, S., Laanen, M., 2022. Validation of Sentinel-2 (MSI) and Sentinel-3 (OLCI) Water Quality Products in Turbid Estuaries Using Fixed Monitoring Stations. *Front. Remote Sens.* 2. <https://doi.org/10.3389/frsen.2021.808287>
- Schemel, L.E., 2001. Simplified conversions between specific conductance and salinity units for use with data from monitoring stations. *Interag. Ecol. Program Newsl.* 14, 17–18.
- Schröder, T., Schmidt, S.I., Kutzner, R.D., Bernert, H., Stelzer, K., Friese, K., Rinke, K., 2024. Exploring Spatial Aggregations and Temporal Windows for Water Quality Match-Up Analysis Using Sentinel-2 MSI and Sentinel-3 OLCI Data. *Remote Sens.* 16, 2798. <https://doi.org/10.3390/rs16152798>
- Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and Statistical Modeling with Python, in: *The 9th Python in Science Conference*. Austin, Texas, pp. 92–96.
- Sebilian Wittingham, S., Moderan, J., Boyer, K.E., 2019. Temperature and salinity effects on submerged aquatic vegetation traits and susceptibility to grazing. *Aquat. Bot.* 158, 103119. <https://doi.org/10.1016/j.aquabot.2019.05.004>
- Sei, A., 2015. Efficient correction of adjacency effects for high-resolution imagery: integral equations, analytic continuation, and Padé approximants. *Appl. Opt.* 54, 3748–3758. <https://doi.org/10.1364/AO.54.003748>
- Servén, D., Brummitt, C., Abedi, H., Hlink, 2018. pyGAM: Generalized Additive Models in Python. *Zenodo*. <https://doi.org/10.5281/zenodo.1208723>
- Shahin, M.A., Maier, H.R., Jaksa, M.B., 2004. Data Division for Developing Neural Networks Applied to Geotechnical Engineering. *J. Comput. Civ. Eng.* 18, 105–114. [https://doi.org/10.1061/\(ASCE\)0887-3801\(2004\)18:2\(105\)](https://doi.org/10.1061/(ASCE)0887-3801(2004)18:2(105))
- Shahraki, M.Z., Keivany, Y., Dorche, E.E., Blocksom, K., Bruder, A., Flotemersch, J., Bănăduc, D., 2023. Distribution and Expansion of Alien Fish Species in the Karun River Basin, Iran. *Fishes* 8, 538. <https://doi.org/10.3390/fishes8110538>
- Shahvaran, A.R., Kheyrollah Pour, H., Van Cappellen, P., 2024. Comparative Evaluation of Semi-Empirical Approaches to Retrieve Satellite-Derived Chlorophyll-a Concentrations from Nearshore and Offshore Waters of a Large Lake (Lake Ontario). *Remote Sens.* 16. <https://doi.org/10.3390/rs16091595>
- Shi, W., Wang, M., 2009. An assessment of the black ocean pixel assumption for MODIS SWIR bands. *Remote Sens. Environ.* 113, 1587–1597. <https://doi.org/10.1016/j.rse.2009.03.011>
- Siddorn, J.R., Bowers, D.G., Hogue, A.M., 2001. Detecting the Zambezi River Plume using Observed Optical Properties. *Mar. Pollut. Bull.* 42, 942–950. [https://doi.org/10.1016/S0025-326X\(01\)00053-4](https://doi.org/10.1016/S0025-326X(01)00053-4)
- Sillen, S.J., Ross, M.R.V., Collins, S.M., 2024. Long-Term Trends in Productivity Across Intermountain West Lakes Provide No Evidence of Widespread Eutrophication. *Water Resour. Res.* 60. <https://doi.org/10.1029/2023WR034997>
- Smith, B., Pahlevan, N., Schalles, J., Ruberg, S., Errera, R., Ma, R., Giardino, C., Bresciani, M., Barbosa, C., Moore, T., Fernandez, V., Alikas, K., Kangro, K., 2021. A Chlorophyll-a Algorithm for Landsat-8 Based on Mixture Density Networks. *Front. Remote Sens.* 1.
- Smith, R.C., Baker, K.S., 1978. The bio-optical state of ocean waters and remote sensing. *Limnol. Oceanogr.* 23, 247–259. <https://doi.org/10.4319/lo.1978.23.2.0247>
- Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. *Stat. Comput.* 14, 199–222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
- Sobol', I.M., 1993. Sensitivity Estimates for Nonlinear Mathematical Models. *Math Model Comput Exp* 1, 407–414.
- Solinst, 2026. (CDOM) Colored Dissolved Organic Matter Sensors. Solinst Groundw. Surf. Water Monit. Instrum. Water Level Meters Interface Meters Dataloggers. URL <https://www.solinst.com/instruments/solinst-eureka/water-quality-sensors/cdom-colored-dissolved-organic-matter-sensors/> (accessed 6.7.26).

References

- Sotomayor, L.N., Cracknell, M.J., Musk, R., 2023. Supervised machine learning for predicting and interpreting dynamic drivers of plantation forest productivity in northern Tasmania, Australia. *Comput. Electron. Agric.* 209, 107804. <https://doi.org/10.1016/j.compag.2023.107804>
- Steinmetz, F., Deschamps, P.-Y., Ramon, D., 2011. Atmospheric correction in presence of sun glint: application to MERIS. *Opt. Express* 19, 9783–9800. <https://doi.org/10.1364/OE.19.009783>
- Sterckx, S., Knaeps, E., Ruddick, K., 2011. Detection and correction of adjacency effects in hyperspectral airborne data of coastal and inland waters: the use of the near infrared similarity spectrum. *Int. J. Remote Sens.* 32, 6479–6505. <https://doi.org/10.1080/01431161.2010.512930>
- Stumpf, R.P., Arnone, R.A., Gould Jr., R.W., Martinolich, P.M., Ransibrahmanakul, V., 2003. A partially coupled ocean-atmosphere model for retrieval of water-leaving radiance from SeaWiFS in coastal waters. NASA Tech. Memo. - SeaWiFS Postlaunch Tech. Rep. Ser. 51–59.
- Sumner, D.M., Belaineh, G., 2005. Evaporation, precipitation, and associated salinity changes at a humid, subtropical estuary. *Estuaries* 28, 844–855. <https://doi.org/10.1007/BF02696014>
- Suthaharan, S., 2016. Decision Tree Learning, in: *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning*. Springer US, Boston, MA, pp. 237–269. https://doi.org/10.1007/978-1-4899-7641-3_10
- Tanré, D., Deschamps, P.Y., Duhaut, P., Herman, M., 1987. Adjacency effect produced by the atmospheric scattering in thematic mapper data. *J. Geophys. Res. Atmos.* 92, 12000–12006. <https://doi.org/10.1029/JD092iD10p12000>
- Tesfaye, A., 2024. Remote Sensing-Based Water Quality Parameters Retrieval Methods: A Review. *East Afr. J. Environ. Nat. Resour.* 7, 80–97. <https://doi.org/10.37284/eajenr.7.1.1753>
- Tessema, N., Yadeta, D., Kebede, A., Ayele, G.T., 2022. Soil and Irrigation Water Salinity, and Its Consequences for Agriculture in Ethiopia: A Systematic Review. *Agriculture* 13, 109. <https://doi.org/10.3390/agriculture13010109>
- Thopanaiah, C.K., Gireesh Babu C N, Gurani, V., Rajani, T., Pavani, A., Muthukumaran, D., Deivasigamani, S., 2024. Advanced Remote Sensing and Generative Models for Comprehensive Water Quality Management in a Changing Climate. *Remote Sens. Earth Syst. Sci.* 7, 596–611. <https://doi.org/10.1007/s41976-024-00149-5>
- Tian, D., Zhao, X., Gao, L., Liang, Z., Yang, Z., Zhang, P., Wu, Q., Ren, K., Li, R., Yang, C., Li, S., Wang, M., He, Z., Zhang, Z., Chen, J., 2024. Estimation of water quality variables based on machine learning model and cluster analysis-based empirical model using multi-source remote sensing data in inland reservoirs, South China. *Environ. Pollut.* 342. <https://doi.org/10.1016/j.envpol.2023.123104>
- Tobler, W.R., 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ. Geogr.* 46, 234–240. <https://doi.org/10.2307/143141>
- Topp, S.N., Pavelsky, T.M., Jensen, D., Simard, M., Ross, M.R.V., 2020. Research Trends in the Use of Remote Sensing for Inland Water Quality Science: Moving Towards Multidisciplinary Applications. *Water* 12, 169. <https://doi.org/10.3390/w12010169>
- Uncles, R.J., Stephens, J.A., 2011. The Effects of Wind, Runoff and Tides on Salinity in a Strongly Tidal Sub-estuary. *Estuaries Coasts* 34, 758–774. <https://doi.org/10.1007/s12237-010-9365-3>
- United Nations, 2023. The United Nations World Water Development Report 2023: Partnerships and Cooperation for Water. UNESCO, Paris.
- Urquhart, E.A., Zaitchik, B.F., Hoffman, M.J., Guikema, S.D., Geiger, E.F., 2012. Remotely sensed estimates of surface salinity in the Chesapeake Bay: A statistical approach. *Remote Sens. Environ.* 123, 522–531. <https://doi.org/10.1016/j.rse.2012.04.008>

References

- U.S. Geological Survey, 2019a. Landsat 8 Surface Reflectance Code (LASRC) Product Guide.
- U.S. Geological Survey, 2019b. Specific conductance: U.S. Geological Survey Techniques and Methods, Techniques and Methods. Reston, VA, book 9, chap. A6.3, 15 p.
<https://doi.org/10.3133/tm9A6.3>
- Vachon, D., 2023. Water quality in the fluvial section Physicochemical and bacteriological parameters, 5th edition. ed. Canada Water Agency.
- Vakili, T., Amanollahi, J., 2020. Determination of optically inactive water quality variables using Landsat 8 data: A case study in Geshlagh reservoir affected by agricultural land use. *J. Clean. Prod.* 247, 119134. <https://doi.org/10.1016/j.jclepro.2019.119134>
- Valerio, A. de M., Kampel, M., Vantrepotte, V., Ward, N.D., Sawakuchi, H.O., Less, D.F.D.S., Neu, V., Cunha, A., Richey, J., 2018. Using CDOM optical properties for estimating DOC concentrations and pCO₂ in the Lower Amazon River. *Opt. Express* 26, A657–A677. <https://doi.org/10.1364/OE.26.00A657>
- Vandermeulen, R.A., Arnone, R., Ladner, S., Martinolich, P., 2014. Estimating sea surface salinity in coastal waters of the Gulf of Mexico using visible channels on SNPP-VIIRS, in: *Ocean Sensing and Monitoring VI. Presented at the Ocean Sensing and Monitoring VI, SPIE*, pp. 56–62. <https://doi.org/10.1117/12.2053417>
- Vanhellemont, Q., 2019. Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives. *Remote Sens. Environ.* 225, 175–192. <https://doi.org/10.1016/j.rse.2019.03.010>
- Vanhellemont, Q., Ruddick, K., 2018. Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications. *Remote Sens. Environ.* 216, 586–597. <https://doi.org/10.1016/j.rse.2018.07.015>
- Vanhellemont, Q., Ruddick, K., 2015. Advantages of high quality SWIR bands for ocean colour processing: Examples from Landsat-8. *Remote Sens. Environ.* 161, 89–106. <https://doi.org/10.1016/j.rse.2015.02.007>
- Vermote, E.F., Tanre, D., Deuze, J.L., Herman, M., Morcette, J.-J., 1997. Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: an overview. *IEEE Trans. Geosci. Remote Sens.* 35, 675–686. <https://doi.org/10.1109/36.581987>
- Verrelst, J., Rivera, J.P., 2017. Chapter 16 - A Global Sensitivity Analysis Toolbox to Quantify Drivers of Vegetation Radiative Transfer Models, in: *Petropoulos, G.P., Srivastava, P.K. (Eds.), Sensitivity Analysis in Earth Observation Modelling. Elsevier*, pp. 319–339. <https://doi.org/10.1016/B978-0-12-803011-0.00016-1>
- Vovk, V., 2013. Kernel Ridge Regression, in: *Schölkopf, B., Luo, Z., Vovk, V. (Eds.), Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik. Springer, Berlin, Heidelberg*, pp. 105–116. https://doi.org/10.1007/978-3-642-41136-6_11
- Vuille, M., Baumgartner, M.F., 1993. Hydrologic investigations in the north Chilean Altiplano using Landsat - MSS and - TM data. *Geocarto Int.* 8, 35–45. <https://doi.org/10.1080/10106049309354418>
- Wagner, R.J., Jr, R.W.B., Oblinger, C.J., Smith, B.A., 2006a. Guidelines and standard procedures for continuous water-quality monitors: Station operation, record computation, and data reporting. *Tech. Methods*. <https://doi.org/10.3133/tm1D3>
- Wagner, R.J., Jr, R.W.B., Oblinger, C.J., Smith, B.A., 2006b. Guidelines and standard procedures for continuous water-quality monitors: Station operation, record computation, and data reporting. *Tech. Methods*. <https://doi.org/10.3133/tm1D3>
- Wang, D., Tang, B.-H., Li, Z.-L., 2024. Evaluation of five atmospheric correction algorithms for multispectral remote sensing data over plateau lake. *Ecol. Inform.* 82, 102666. <https://doi.org/10.1016/j.ecoinf.2024.102666>
- Wang, F., Xu, Y.J., 2008. Development and application of a remote sensing-based salinity prediction model for a large estuarine lake in the US Gulf of Mexico coast. *J. Hydrol.* 360, 184–194. <https://doi.org/10.1016/j.jhydrol.2008.07.036>

References

- Wang, M., Shi, W., 2012. Sensor Noise Effects of the SWIR Bands on MODIS-Derived Ocean Color Products. *IEEE Trans. Geosci. Remote Sens.* 50, 3280–3292.
<https://doi.org/10.1109/TGRS.2012.2183376>
- Wang, Q., Liu, G., Song, K., Wen, Z., Shang, Y., Li, S., Fang, C., Tao, H., 2024. Comparison of Machine Learning Algorithms for Estimating Global Lake Clarity With Landsat TOA Data. *IEEE Trans. Geosci. Remote Sens.* 62, 1–14.
<https://doi.org/10.1109/TGRS.2024.3400221>
- Wang, X., Jiang, Y., Jiang, M., Cao, Z., Li, X., Ma, R., Xu, L., Xiong, J., 2023. Estimation of Total Phosphorus Concentration in Lakes in the Yangtze-Huaihe Region Based on Sentinel-3/OLCI Images. *Remote Sens.* 15, 4487. <https://doi.org/10.3390/rs15184487>
- Warner, J., Sexauer, J., scikit-fuzzy, twmeggs, alexsavio, Unnikrishnan, A., Castelh o, G., Pontes, F.A., Uelwer, T., pd2f, laurazh, Batista, F., alexbuy, Broeck, W.V. den, Song, W., Badger, T.G., P rez, R.A.M., Power, J.F., Mishra, H., Trullols, G.O., H rteborn, A., 99991, 2019. JDWarner/scikit-fuzzy: Scikit-Fuzzy version 0.4.2.
<https://doi.org/10.5281/zenodo.3541386>
- Warren, M.A., Simis, S.G.H., Martinez-Vicente, V., Poser, K., Bresciani, M., Alikas, K., Spyarakos, E., Giardino, C., Ansper, A., 2019. Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters. *Remote Sens. Environ.* 225, 267–289. <https://doi.org/10.1016/j.rse.2019.03.018>
- Werdell, P.J., Franz, B.A., Lefler, J.T., Robinson, W.D., Boss, E., 2013. Retrieving marine inherent optical properties from satellites using temperature and salinity-dependent backscattering by seawater. *Opt. Express* 21, 32611–32622.
<https://doi.org/10.1364/OE.21.032611>
- Werdell, P.J., McClain, C.R., 2019. Satellite Remote Sensing: Ocean Color, in: *Encyclopedia of Ocean Sciences*, GSFC-E-DAA-TN65587. Elsevier, pp. 443–455.
<https://doi.org/10.1016/B978-0-12-409548-9.10817-6>
- Williams, C., Rasmussen, C., 1995. Gaussian Processes for Regression, in: *Advances in Neural Information Processing Systems*. MIT Press.
- Wolanin, A., Soppa, M.A., Bracher, A., 2016. Investigation of Spectral Band Requirements for Improving Retrievals of Phytoplankton Functional Types. *Remote Sens.* 8, 871.
<https://doi.org/10.3390/rs8100871>
- Wu, Y., Knudby, A., Lapen, D., 2023. Topography-adjusted Monte Carlo simulation of the adjacency effect in remote sensing of coastal and inland waters. *J. Quant. Spectrosc. Radiat. Transf.* 303, 108589. <https://doi.org/10.1016/j.jqsrt.2023.108589>
- Wu, Y., Knudby, A., Pahlevan, N., Lapen, D., Zeng, C., 2024. Sensor-generic adjacency-effect correction for remote sensing of coastal and inland waters. *Remote Sens. Environ.* 315, 114433. <https://doi.org/10.1016/j.rse.2024.114433>
- Wu, Z., Pang, J., Li, J., Wang, Y., Ruan, J., Zhang, X., Yang, L., Pang, Y., Gao, Y., 2025. A review of remote sensing-based water quality monitoring in turbid coastal waters. *Intell. Mar. Technol. Syst.* 3, 24. <https://doi.org/10.1007/s44295-025-00075-2>
- Wyvern, 2026. Product Specifications | Wyvern Data Product Guide [WWW Document]. URL <https://guide.wyvern.space/data-product-guide/technical-specifications/product-specifications/> (accessed 5.26.26).
- Xia, K., Wu, T., Li, X., Wang, S., Tang, H., Zu, Y., Yang, Y., 2024. A novel method for assessing water quality status using MODIS images: A case study of large lakes and reservoirs in China. *J. Hydrol.* 638. <https://doi.org/10.1016/j.jhydrol.2024.131545>
- Xie, Z., Zhang, C., Berry, L., 2013. Geographically weighted modelling of surface salinity in Florida Bay using Landsat TM data. *Remote Sens. Lett.* 4, 75–83.
<https://doi.org/10.1080/2150704X.2012.693218>

References

- Yan, N., Sun, Z., Huang, W., Jun, Z., Sun, S., 2023. Assessing Landsat-8 atmospheric correction schemes in low to moderate turbidity waters from a global perspective. *Int. J. Digit. Earth* 16, 66–92. <https://doi.org/10.1080/17538947.2022.2161651>
- Yan, X., Zhang, T., Du, W., Meng, Q., Xu, X., Zhao, X., 2024. A Comprehensive Review of Machine Learning for Water Quality Prediction over the Past Five Years. *J. Mar. Sci. Eng.* 12, 159. <https://doi.org/10.3390/jmse12010159>
- Yan, Z., Fang, C., Song, K., Wang, X., Wen, Z., Shang, Y., Tao, H., Lyu, Y., 2025. Spatiotemporal variation in biomass abundance of different algal species in Lake Hulun using machine learning and Sentinel-3 images. *Sci. Rep.* 15, 2739. <https://doi.org/10.1038/s41598-025-87338-4>
- Yao, S., Chen, C., He, M., Cui, Z., Mo, K., Pang, R., Chen, Q., 2023. Land use as an important indicator for water quality prediction in a region under rapid urbanization. *Ecol. Indic.* 146, 109768. <https://doi.org/10.1016/j.ecolind.2022.109768>
- Zhang, C., Xie, Z., Roberts, C., Berry, L., Chen, G., 2012. Salinity Assessment in Northeast Florida Bay Using Landsat TM Data. *Southeast. Geogr.* 52, 267–281.
- Zhang, J., Fu, P., Meng, F., Yang, X., Xu, J., Cui, Y., 2022. Estimation algorithm for chlorophyll-a concentrations in water from hyperspectral images based on feature derivation and ensemble learning. *Ecol. Inform.* 71. <https://doi.org/10.1016/j.ecoinf.2022.101783>
- Zhang, X., Hu, L., 2021. Light Scattering by Pure Water and Seawater: Recent Development. *J. Remote Sens.* 2021. <https://doi.org/10.34133/2021/9753625>
- Zhang, X., Hu, L., He, M.-X., 2009. Scattering by pure seawater: Effect of salinity. *Opt. Express* 17, 5698–5710. <https://doi.org/10.1364/OE.17.005698>
- Zhao, D., Feng, L., Sun, K., 2022. Development of a Practical Atmospheric Correction Algorithm for Inland and Nearshore Coastal Waters. *IEEE Trans. Geosci. Remote Sens.* 60, 1–15. <https://doi.org/10.1109/TGRS.2021.3092839>
- Zhao, J., Temimi, M., Ghedira, H., 2017. Remotely sensed sea surface salinity in the hyper-saline Arabian Gulf: Application to Landsat 8 OLI data. *Estuar. Coast. Shelf Sci.* 187, 168–177. <https://doi.org/10.1016/j.ecss.2017.01.008>
- Zhao, R., Yang, Q., Wen, Z., Fang, C., Li, S., Shang, Y., Liu, G., Tao, H., Lyu, L., Song, K., 2023. Satellite Estimation of pCO₂ and Quantification of CO₂ Fluxes in China's Chagan Lake in the Context of Climate Change. *Remote Sens.* 15. <https://doi.org/10.3390/rs15245680>
- Zhao, Y., He, X., Pan, S., Bai, Y., Wang, D., Li, T., Gong, F., Zhang, X., 2024. Satellite retrievals of water quality for diverse inland waters from Sentinel-2 images: An example from Zhejiang Province, China. *Int. J. Appl. Earth Obs. Geoinformation* 132, 104048. <https://doi.org/10.1016/j.jag.2024.104048>
- Zheng, Y., Wei, C., Fu, H., Li, H., Fu, M., Zheng, Y., He, Q., Yu, D., 2024. Spatial-temporal evolution analysis of pollutants in Daitou River watershed based on Sentinel-2 satellite images. *Ecol. Indic.* 166. <https://doi.org/10.1016/j.ecolind.2024.112436>
- Zhou, W., Wang, S., Zhou, Y., Troy, A., 2006. Mapping the concentrations of total suspended matter in Lake Taihu, China, using Landsat-5 TM data. *Int. J. Remote Sens.* <https://doi.org/10.1080/01431160500353825>
- Zhou, Yadong, Li, W., Cao, X., He, B., Feng, Q., Yang, F., Liu, H., Kutser, T., Xu, M., Xiao, F., Geng, X., Yu, kai, Du, Y., 2024. Spatial-temporal distribution of labeled set bias remote sensing estimation: An implication for supervised machine learning in water quality monitoring. *Int. J. Appl. Earth Obs. Geoinformation* 131, 103959. <https://doi.org/10.1016/j.jag.2024.103959>
- Zhou, Y., Li, W., Cao, X., He, B., Feng, Q., Yang, F., Liu, H., Kutser, T., Xu, M., Xiao, F., Geng, X., Yu, K., Du, Y., 2024. Spatial-temporal distribution of labeled set bias remote sensing estimation: An implication for supervised machine learning in water quality monitoring. *Int. J. Appl. Earth Obs. Geoinformation* 131. <https://doi.org/10.1016/j.jag.2024.103959>

References

- Zhu, L., Cui, T., Runa, A., Pan, X., Zhao, W., Xiang, J., Cao, M., 2024. Robust remote sensing retrieval of key eutrophication indicators in coastal waters based on explainable machine learning. *ISPRS J. Photogramm. Remote Sens.* 211, 262–280.
<https://doi.org/10.1016/j.isprsjprs.2024.04.007>