

A comprehensive analysis of retrieving optically inactive indicators from multi-level remote sensing product(s) in Irish waters using data science techniques

Abdul Majed Sajib ^{a,b,c}, Md. Galal Uddin ^{a,b,c*}, Agnieszka I. Olbert ^{a,b,c}

^a Civil Engineering, School of Engineering, College of Science and Engineering, University of Galway, Ireland

^b Ryan Institute, University of Galway, Ireland

^c Eco-HydroInformatics Research Group (EHIRG), School of Engineering, College of Science and Engineering, University of Galway, Ireland

*Corresponding author: Md Galal Uddin, Department of Civil Engineering, College of Science and Engineering, University of Galway, Ireland. Email:

mdgalal.uddin@universityofgalway.ie

List of Abbreviations

Abbreviation	Description	Abbreviation	Description
% sat	Percentage of saturation	MERIS	Medium resolution imaging spectrometer
µmol/L	Micromoles per liter	mg/L	Milligrams per liter
AdaB	Adaboost	ML	Machine learning
AI	Artificial intelligence	MLP	Multilayer perceptron
ANN	Artificial neural network	MODIS	Moderate resolution imaging spectroradiometer
BGR	Blending gradient boosting - random forest	MoD-Z	Modified Z-score
BPNN	Back propagation neural network	MS	Multi-sensors
BRR	Bayesian ridge regression	MSE	Mean square error
CatB	Catboost	NASA	National aeronautics and space administration
CDF	Cumulative distribution function	NEMO-PISCES/Modelled data	Nucleus for European modelling of the ocean-pelagic interaction scheme for carbon and ecosystem studies
CDOM	Coloured dissolved organic matter	nm	nanometer
CDSE	Copernicus data space scosystem	OA/OI	Optically active

CHL	Chlorophyll-a	OI	Optically inactive
CI	Confidence interval	OMP	Orthogonal matching pursuit
CMS	Copernicus marine service	PABE	Percentage of absolute bias error
CZCS	Coastal zone colour scanner	PBIAS	Percentage of bias
DNN	Deep neural network	PCR	Principal component regression
DOX	Dissolved oxygen	PDF	Probability distribution function
DT	Decision trees	PLS	Partial least square regression
EBK	Empirical Bayesian kriging	PREI	Percent of relative error Index
EPA	Environmental protection agency	PRF	Probabilistic random forest
ESA	European Space Agency	R ²	Coefficient of determination
ETR	Extra tree regression	RANSAC	Random sample consensus
EU	European Union	RF	Random forest
EUMETSAT	European organisation for the exploitation of meteorological satellites	RFECV	Recursive feature elimination
ExDT	Extra decision trees	Rhow	Water-leaving reflectance
GAM	Generalize adaptive model	RMSE	Root mean square error
GB	Gradient boosting	R _{rs}	Remote sensing reflectance
GBGS	Gradient boosted grid search	RS	Remote sensing
GPR	Gaussian process regression	S2-MSI	Sentinel-2 multi spectral instrument
GRU	Gated recurrent unit networks	S3-OLCI	Sentinel-3 Ocean and Land Colour Instrument
HM	Huber model	SFS-RF	Sequential feature selector–random forest
IsoF	Isolation Forest	SGD	Stochastic gradient descent
KNN	K-nearest neighbours	SHAP	SHapley Additive exPlanations
KNNR	Kernal ridge regression	SPM	Suspended particulate matter
L	Level	Sta-Gen	Stacked generalization
LARS	Least-angle regression	SVM/SVR	Support vector machine
LGBM	Light gradient boosting machine	TN	Total nitrogen
LoF	Local outlier factor	TP	Total phosphorus
LR	Linear regression	TrC	Transitional and coastal

LSTM	Long short-term memory	VIIRS	Visible infrared imaging radiometer suite
L_w	Water-leaving radiance	WFD	Water framework directive
MAE	Mean absolute error	WQ	Water quality
MDNN	Multi-modal deep neural network	XGB	Extreme gradient boosting

14 **Abstract**

15 The present research was carried out to retrieve dissolved oxygen (DOX) using the Copernicus
16 Marine Services products from the Irish transitional and coastal waters. To achieve the research
17 goal, the study developed and validated 2101 machine learning (ML)/artificial intelligence (AI)
18 (supervised learning, stacking-ensembles, equations, and voting-based ensembles) and
19 statistical models using multi-level (Level-3 and Level-4) Sentinel-3 OLCI (S3-OLCI) and
20 Multi-sensor (MS) remote sensing (RS) datasets in conjunction with in-situ and modelled DOX
21 datasets. While supervised models (e.g., K-nearest neighbours, Gradient boosting, and Extra
22 decision trees) excelled in the training phase (EPA: $MSE \leq 0.03$ with $CI \pm 0.02$; Modelled:
23 $MSE \approx 0$ with $CI \pm 0$) but showed limited generalizability on independent validation datasets
24 (2022-2023), indicating poor model accuracy and sensitivity (EPA-2022: $R^2 = -0.03 - 0.16$;
25 EPA-2023: $R^2 = -0.09 - 0.1$; Modelled-2022: $R^2 = 0.37 - 0.53$; Modelled-2023: $R^2 = -1.39 - -$
26 0.26). In terms of product, S3-OLCI outperformed MS data with low uncertainty, whereas
27 spatio-temporal analysis showed the highest DOX in inshore/semi-enclosed bays and the
28 lowest offshore. Overall, the results underscore that model performance is determined by
29 methodological characteristics rather than model quantity. Despite the validation challenges,
30 the results highlight key difficulties in retrieving optically inactive water quality (WQ)
31 indicators like DOX using RS and ML/AI approaches. The findings of the research could be
32 effective for supporting the mapping of baseline oxygen conditions, the application of ML/AI
33 techniques to retrieve WQ indicators from RS products and their further technological

34 advancement, such as managing the anthropogenic water cycle (i.e., human-altered
35 hydrological and nutrient dynamics).

36 **Keywords:** Optically Inactive Water Quality, Remote Sensing, Copernicus Marine Service,
37 Dissolved Oxygen, Data-driven model, Machine Learning

38 **1. Introduction**

39 Natural resources, such as water, are at risk of environmental deterioration due to growing
40 population, globalization, industrialization, and land-cover change (Gani et al., 2023). These
41 factors not only affect surface water quality (WQ) but also pose challenges for effective
42 monitoring and management. Traditional in-situ measurements are the most reliable approach
43 for routine WQ monitoring. However, this approach is quite expensive and laborious (e.g.,
44 requires a highly equipped laboratory and skilled manpower) and limits its scalability for large
45 or remote waterbodies (Diganta et al., 2024; Sajib et al., 2025a).

46 To overcome this issue, recently, scientists and organizations (e.g., the European Space Agency
47 - ESA and the National Aeronautics and Space Administration - NASA) have highlighted that
48 remote sensing (RS) technology could be a cost-effective and time-saving approach, which can
49 be utilized for measuring/retrieving WQ from various waterbodies (Diganta et al., 2024; Sajib
50 et al., 2025a). RS is a process of obtaining data/information about an object/area from a distance
51 by measuring the reflected and emitted radiation (Frouin et al., 2019), specifically in the visible
52 (≈ 400 to 700 nm) and near-infrared (≈ 700 to ≥ 2000 nm) wavelengths for ocean colour
53 applications (Mobley et al., 2022). The ocean colour research began with the launch of the
54 Coastal Zone Colour Scanner (CZCS) in 1978 to measure CHL (chlorophyll-a), CDOM
55 (coloured dissolved organic matter), SPM (suspended particulate matter), etc., from space
56 (Gordon, 2021; IOCCG, 2000; Martin Traykovski, 2003). Beyond WQ measurement, RS
57 technology is also utilized to study ocean carbon mapping (Liu et al., 2025), ecosystem change

58 mapping (Makri et al., 2025), fish zone mapping (Nadeem et al., 2025), coral reef mapping
59 (Moustafa et al., 2025), seagrass bed mapping (Lizcano-Sandoval et al., 2025), shallow-water
60 bathymetry mapping (Karimi & Torabi, 2025), harmful algal bloom mapping (Zarbipour et al.,
61 2026), and pollution source identification (Fang et al., 2025; Moon et al., 2025; Shen et al.,
62 2025).

63 Typically, WQ indicators measured by RS can be categorized into two types: (i) optically active
64 (OA) (e.g., CHL, CDOM, SPM); and (ii) optically inactive (OI) WQ indicators (e.g., DOX -
65 dissolved oxygen, TP - total phosphorus, TN - total nitrogen, Sajib et al., 2024a; Vakili &
66 Amanollahi, 2020). A detailed description of these classifications can be found in Diganta et
67 al. (2024) and Sajib et al. (2025a). Furthermore, in terms of optical complexity, waterbodies
68 are classified as case-1 waters (e.g., open ocean) and case-2 waters (e.g., coastal regions, rivers,
69 and lakes, Morel and Prieur, 1977; Gordon and Morel, 1983). To monitor these waterbodies,
70 various multispectral (e.g., MODIS-Aqua/Terra, VIIRS, Sentinel-3 (S3) OLCI, Sentinel-2 (S2)
71 MSI, Landsat-8 OLI) and hyperspectral satellites (e.g., Zhuhai-1, Geofan-5) with different
72 spatial resolutions have been launched into space in the last few decades (Chen et al., 2022;
73 Mohseni et al., 2022). Additionally, the Medium Resolution Imaging Spectrometer-MERIS
74 cloud also play a pivotal role in aquatic observation, particularly for assessing CHL and
75 turbidity (Lyu et al., 2015). Among these, the ESA encourages the use of the S2-MSI and the
76 S3-OLCI products for ocean colour observations in the European Union (EU) waterbodies
77 (Rixen et al., 2024). The S3-OLCI, which maintains the continuity of MERIS, provides high-
78 quality ocean colour data and is available at levels L1 (here 'L' refers to level, e.g., uncorrected
79 data), L2 (e.g., atmospherically corrected), L3 (e.g., gridded and atmospherically corrected),
80 and L4 (e.g., a combination of satellite data and model), with 300 m resolution, accessible via
81 the Copernicus Data Space Ecosystem (CDSE) (<https://dataspace.copernicus.eu/>) and the
82 Copernicus Marine Service (CMS) server (<https://dataspace.copernicus.eu/>). Similarly, the

83 CMS provides L3 S2-MSI and multi-sensor (MS) L3/L4 products (Sajib et al., 2025b). A
84 detailed description of these products can be found in Garnesson et al. (2024) and Colella et al.
85 (2023). Although these products are freely available to users through the CMS platform, very
86 few studies have tested the effectiveness of these products for retrieving both OA- and OI-WQ
87 indicators (Brando et al., 2024; Chatziantoniou et al., 2022; Laura et al., 2025; Marchese et al.,
88 2024; Uddin et al., 2025a).

89 Among OI-WQ indicators, DOX refers to the amount of oxygen dissolved in water, which
90 plays a critical role in the growth and survival of aquatic life (Cai et al., 2023). The level of
91 DOX in natural waters (typically ranging from 4 to 9.5 mg/L depending on various factors) is
92 influenced by physical, chemical, and biological activities (Boyd, 2015). Since the mid-20th
93 century, human-induced eutrophication, organic pollution, and ocean acidification have led to
94 increased rates of oxygen depletion, resulting in hypoxic zones in coastal waters worldwide
95 (Dai et al., 2023). These deoxygenation trends increase biodiversity loss, disrupt food chains,
96 and decrease ecological services in populated coastal areas (Limburg et al., 2020). Therefore,
97 maintaining standard DOX levels is essential for the survival of aquatic life, as low DOX can
98 develop hypoxia (Li et al., 2023a). Typically, hypoxia occurs when the water flows from
99 various sources injected into the waterbodies such as lakes, rivers, etc., and ends up in the
100 coastal oceans through the flow process (Dong et al., 2024; Pannard et al., 2024). For example,
101 regions like the Gulf of Mexico, the Chesapeake Bay, and the Saanich Inlet frequently
102 experience hypoxia events due to the influx of nutrient-rich water from various sources (Kudela
103 et al., 2024; Rabalais et al., 2010). In European coastal zones, the combined pressures of
104 agricultural runoff, rapid urbanization, and climate change are increasingly recognized as
105 threats to oxygen regimes (European Environment Agency, 2021). Therefore, in the EU, the
106 Water Framework Directive (WFD: 2000/60/EC) mandates monitoring of indicators, such as
107 DOX, to achieve “Good” status in all forms of waterbodies, including transitional and coastal

108 (TrC) waters (EU, 2019). As a result, countries like Ireland regularly assess DOX in TrC waters
109 via the Environmental Protection Agency (EPA, Ireland) to protect marine ecosystems (EPA,
110 2001; 2021a; 2023a). Additionally, due to increased human-induced pressures, there is an
111 urgent need for accurate and scalable monitoring techniques to identify early indicators of dead
112 zones and to facilitate mitigation efforts in human-affected coastal systems. However, the
113 available in-situ monitoring data for TrC waters from sources, such as EPA (Ireland) are often
114 insufficient for developing a RS-based model, specifically due to their sparse distribution and
115 high atmospheric disturbances (e.g., clouds, precipitation) during the monitoring period
116 (Delaney et al., 2023; Sajib et al., 2025b). In contrast, the CMS provides a full range of the
117 DOX data across Ireland via the NEMO-PISCES (Nucleus for European Modelling of the
118 Ocean-Pelagic Interaction Scheme for Carbon and Ecosystem Studies) numerical modelled
119 data, also known as the “Atlantic Iberia Biscay Ireland - Irish Ocean Biogeochemistry product”
120 (Gutknecht et al., 2019; Samlas et al., 2025). Although the NEMO-PISCES provides more
121 frequent coverage, but many researchers recently have criticized this data in terms of data
122 accuracy and reliability (Ford & Quiring, 2019; Mihailov et al., 2025; Zhao et al., 2024).

123 Recent advancements integrate machine learning/artificial intelligence (ML/AI) with RS data
124 to retrieve DOX concentrations (Li et al., 2023b; Luo et al., 2025; Shao et al., 2024). Among
125 the various satellites utilized to retrieve DOX, the Landsat-8 OLI and the S2-MSI are the most
126 widely utilized for retrieving the DOX concentrations from lake and river waters (Arias-
127 Rodriguez et al., 2023; Meng et al., 2022; Shi et al., 2024), followed by the MODIS-Aqua/Terra
128 (Guo et al., 2021; Liu et al., 2022), the VIIRS (Kim et al., 2020), and other hyperspectral
129 satellites (Wang et al., 2021; Yang et al., 2022). Furthermore, the existing DOX model often
130 uses band ratios from red, green, blue, and near-infrared spectra, combined with ML/AI
131 algorithms, such as support vector machine (SVM/SVR), extreme gradient boosting (XGB),
132 extra tree regression (ETR), multi-modal deep neural network (MDNN), Catboost (CatB),

133 artificial neural network (ANN), gaussian process regression (GPR), deep neural network
134 (DNN), and back propagation neural network (BPNN), etc., to retrieve DOX concentrations
135 from lakes, reservoirs, rivers, and coastal waterbodies (Dong et al., 2023, 2024; Gao et al.,
136 2024; Li et al., 2020; Luo et al., 2025; Peterson et al., 2020; Salas et al., 2022; Shi et al., 2023;
137 Tian et al., 2024; Toming et al., 2024). Additionally, several research utilized feature selection
138 algorithms to optimize the number of RS bands or band ratios during the training period (Guo
139 et al., 2021, 2023; Krishnaraj & Honnasiddaiah, 2022; Zhou et al., 2024). Yet, only about 38%
140 ($N = 16$) of studies showed strong performance (coefficient of determination - $R^2 \geq 0.8$) during
141 the training/testing period, and no universal band ratio or equation exists for DOX retrieval
142 (see summary of DOX-based retrieval studies in Table S1 - supplementary materials-1).

143 Despite these efforts, existing DOX retrieval approaches are insufficient in several ways.
144 Firstly, the RS studies have primarily focused on OA indicators, and the least importance has
145 been shown towards OI indicators like DOX (Diganta et al., 2024; Mohseni et al., 2022; Sajib
146 et al., 2025a). Secondly, most of the previous research relied on satellite-derived L1 and L2
147 products for retrieving DOX concentrations (Batur & Maktav, 2019; Karakaya & Evrendilek,
148 2011; Krishnaraj & Honnasiddaiah, 2022; Peterson et al., 2020), overlooking the potential of
149 L3/L4 products for better retrieval accuracy. Thirdly, to the best of the author's knowledge, the
150 developed model's validation is limited in the literature in terms of the model's generalization
151 and transferability aspects. Most studies have relied on co-located in-situ data from the same
152 study region, with limited use of completely independent datasets (unseen attributes) or testing
153 across diverse geographical regions to rigorously evaluate model transferability, sensitivity,
154 and uncertainty (Sajib et al., 2025a; 2025b; Uddin et al., 2025a). As a result, the reproducibility
155 of these models under different environmental conditions has become a concern in terms of
156 broader application and generalizability. Therefore, the research extensively explored and
157 evaluated a range of ML/AI, including statistical approaches with independent validation

158 datasets for evaluating how the model's generalization capabilities are influenced by their
159 temporal changes of input attributes. Lastly, the S3-OLCI product's potential for retrieving
160 DOX concentrations remains underexplored globally, despite its design for ocean colour
161 applications.

162 The present research advances existing knowledge by addressing these gaps through a
163 comprehensive evaluation of multi-level RS products (L3 and L4 products from S3-OLCI and
164 MS) integrating with advanced ML/AI and statistical models for DOX retrieval in Irish TrC
165 waterbodies. This study improves model reliability by testing independent datasets with the
166 best-performing model to understand its generalizability in terms of sensitivity and uncertainty,
167 which is essential for EU-WFD compliance and long-term marine environmental management.

168 The research achieved its goal through the following objectives:

- 169 • To compare the data accuracy between the EPA, Ireland, and the modelled DOX
170 concentrations, and explore their relationship with multi-level RS products band.
- 171 • To develop and evaluate the performance of various advanced ML/AI and statistical
172 models in retrieving DOX concentrations for Irish TrC waters.
- 173 • To compare the performance of various levels (L3 vs L4; and S3 OLCI vs MS data) of
174 RS data in retrieving DOX concentrations.
- 175 • To validate the best-performing model with independent datasets for retrieving DOX
176 concentrations more accurately in terms of improving the model(s) sensitivity and
177 uncertainty.

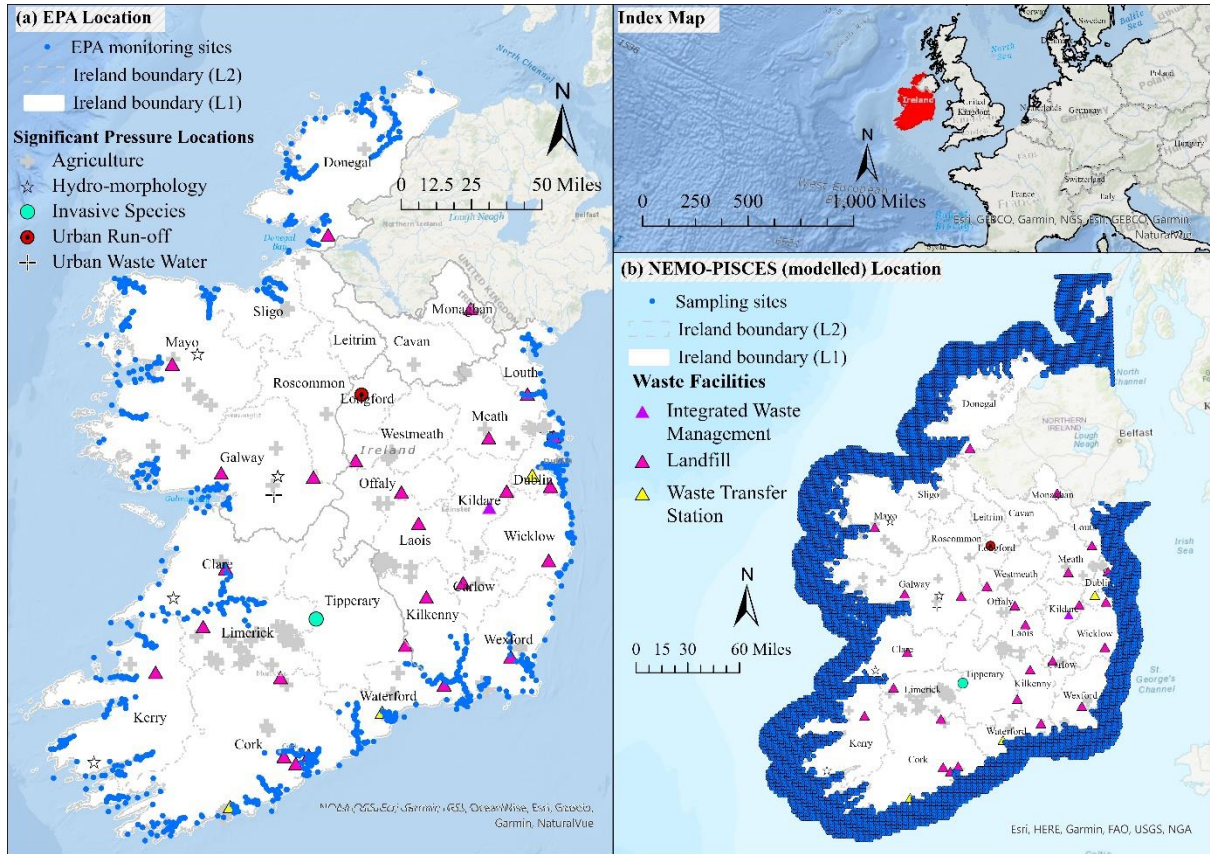
178 **2. Materials and methods**

179 **2.1 Study area**

180 The research utilized the Irish TrC waterbodies for the development of model(s) and validation
181 purposes (Fig. 1). The Ireland coastline (area \approx 4660.28 miles) is characterized by the strong
182 semi-diurnal tidal forces on the west coast, seasonal jet-like circulation on the east coast

183 (Brown et al., 2003; Creane et al., 2021; Fernand et al., 2006), and regional gyre-like
184 circulations due to the density gradients and stratification (Brown et al., 2003). Collectively,
185 these features lead to prolonged residence times in certain areas and impact the marine
186 environment through pollutant accumulation and oxygen depletion (Olbert et al., 2011).
187 Furthermore, seasonal stratification affects nutrient distribution and phytoplankton blooms,
188 resulting in oxygen production/consumption cycle (Sharples et al., 2020). A detailed
189 description of these features and their impact on the Irish TrC waters can be found in the
190 supplementary material-1 as a continuation of subsection 2.1.

191 Recent WQ reports found that approximately half of the Irish surface waters are in
192 unsatisfactory condition (EPA, 2023a; Trodd et al., 2022). The EPA has highlighted that the
193 increasing nutrient levels in TrC waters are damaging marine habitats (EPA, 2021b, 2023a,
194 2023b), and significant pressures, such as agriculture and hydro-morphological changes are the
195 primary causes of WQ degradation in Ireland (EPA, 2021b). Among them, nutrient gradients
196 from agricultural runoff increase the eutrophication processes and oxygen-depleting algal
197 blooms, resulting in the death of fish and benthic animals and plants (Sajib et al., 2025b).
198 Additionally, climate-driven changes (e.g., warmer temperatures and altered precipitation) may
199 further increase the risk of deoxygenation in transitional estuaries where salinity and
200 temperature gradients intersect (Mahaffey et al., 2023). These interactions between natural and
201 anthropogenic factors present the Irish TrC waters as an important system for RS-based
202 modelling of DOX.



203

204

205

Fig 1. Map of Irish EPA monitoring sites and the NEMO-PISCES data sampling sites utilized in this study to retrieve DOX concentrations.

206

2.2 Remote sensing data

207

The research utilized S3-OLCI and MS products from the CMS repository, which cover data

208

between 2016-2023 (Table S2 – supplementary materials-1). The justification for selecting

209

these products and their detailed description can be found in the supplementary material-1 as a

210

continuation of subsection 2.2. Typically, the CMS provides RS products in reflectance (R_{rs})

211

with various bands. The R_{rs} is defined as the ratio between water-leaving radiance (L_w) and

212

downwelling spectral irradiance (E_d) (Eq. 1) (Diganta et al., 2024; Sajib et al., 2025b). In this

213

study, the reflectance bands were converted into the water-leaving reflectance (R_{how}) using

214

the formula (Eq. 2) suggested by Mobley (2022). Furthermore, the methodological architecture

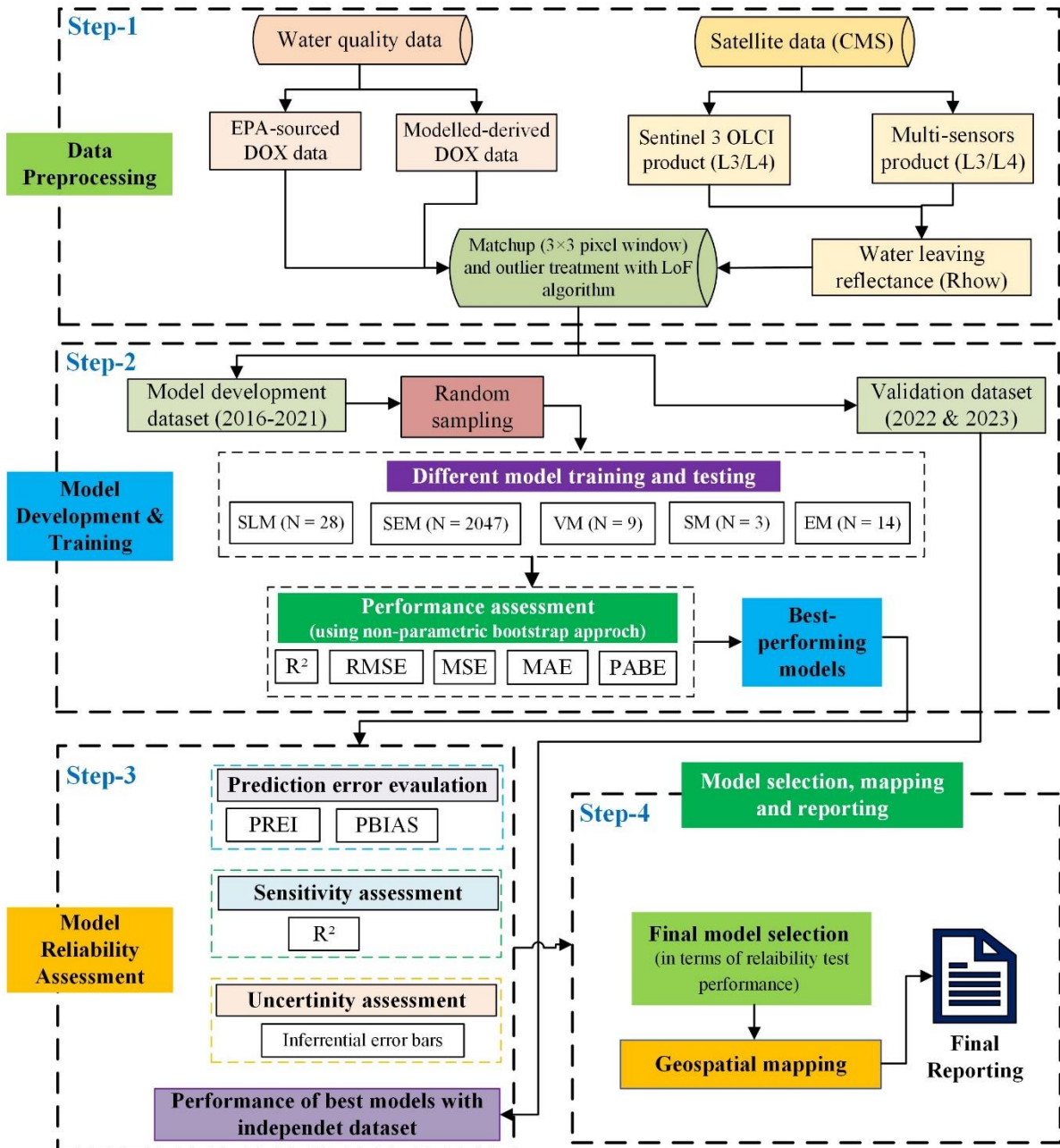
215

of this study can be found in Fig. 2.

216

$$R_{rs} = \frac{L_w}{E_s} \quad (1)$$

217 $R_{how} = R_{rs} \times \pi$ (2)



218 Fig. 2. Methodological framework of the research (here, SLM, SEM, EM, VM, and SM refer to supervised
 219 learning model, stacking-ensembles model, equations model, voting-based ensembles model, and statistical
 220 model, respectively).
 221

222 **2.3 Water quality data**

223 The research utilized in-situ DOX concentration data from the EPA, Ireland, and modelled
 224 DOX concentration data from the CMS repository. Similar to the RS data, it covers the period
 225 between 2016 and 2023. The details of the DOX concentration sources are as follows:

226 **(i) EPA data:** In this research, a total of 618 monitoring sites from the EPA were utilized for
227 matchup with RS data. Detailed description of the WQ program of the EPA can be found in
228 the supplementary materials-1 as a continuation of subsection 2.3. The EPA measured DOX
229 concentrations in both milligrams per liter (mg/L) and percentage of saturation (% sat), which
230 is openly accessible for all users at the following link: <https://www.catchments.ie/>. For
231 homogeneity, the current study converts % sat data into mg/L and considers 1m depth of data.

232 **(ii) NEMO-PISCES/ Modelled data:** The modelled product includes daily and monthly DOX
233 data, and other indicators (e.g., CHL, nitrates - NO₃, phosphates - PO₄, ammonium - NH₄).
234 Detailed description of the modelled data can be found in the supplementary materials-1 as a
235 continuation of subsection 2.3. In this research, a total of 6650 sampling sites were utilized for
236 matchups with RS data. The modelled DOX concentration is provided in micromoles per liter
237 (μmol/L). To ensure consistency with the EPA data, these concentrations were converted to
238 mg/L. A detailed description of modelled data can be found in McGovern et al. (2022) and
239 Gutknecht et al. (2019).

240 **2.4 Data understanding and preparation**

241 **2.4.1 Matchup procedure**

242 The research used a 3×3-pixel window to calculate the median value of each pixel for both S3-
243 OLCI and MS product(s) in the ESA SNAP Version 11.0.0 on the Windows 11 platform. The
244 selection of a 3×3-pixel window is based on the optical complexity and non-homogenous
245 condition of the Irish TrC waters, where a larger window (e.g., 5×5) may introduce high
246 variability in RS data (Sajib et al., 2025b). Moreover, the selection of median aggregation is
247 based on the recommendation of previous research conducted by different researchers and
248 organizations (IOCCG, 2019; Maciel et al., 2023). Furthermore, the median aggregation is
249 widely utilized in the field of ocean colour RS due to its robustness against outliers, accurate
250 representation of central tendency, and its suitability for skewed data (Coffer et al., 2025;

251 IOCCG, 2019). Additionally, in order to ensure data quality, a post-matchup filtering threshold
252 process was implemented, where only pixels with more than 50% valid data (e.g., 5 out of 9
253 pixels) were retained. The filtering procedures are consistent with the EUMETSAT (2021) and
254 Bailey & Werdell (2006) guidelines, which recommended maintaining a minimum matchup
255 threshold of 50%+1 valid pixels to ensure a balance between statistical robustness and data
256 usability. Further description of matchup procedure can be found in the supplementary
257 material-1 as a continuation of subsection 2.4.1.

258 **2.4.2 Outlier treatment**

259 The current research utilized four of the most popular outlier treatment procedures, including
260 Z-score (applied with a threshold value of 3), modified Z-score (MoD-Z, applied with a
261 threshold value of 3.5), local outlier factor (LoF, applied with a contamination factor value of
262 0.05 to 0.5), and isolation forest (IsoF, applied with a contamination factor value of 0.05 to
263 0.5) for data processing and cleaning, following the methodology of Uddin et al. (2024a). A
264 detailed description of these outlier's techniques and the rationale for selecting these outliers
265 in this research can be found in supplementary materials-1 as a continuation of section 2.4.2.

266 **2.4.3 Data splitting**

267 Prior to model development, temporal aggregation was utilized on both DOX and RS products
268 to reduce short-term variability and improve robustness. Daily/monthly DOX concentrations
269 and Rhow band values were aggregated into yearly average values. This approach ensures a
270 consistency between RS products and DOX observations by maintaining interannual
271 variability, which minimizes noise and seasonal impacts in the dataset (Alarcon Falconi et al.,
272 2020; Forkel et al., 2013). Furthermore, this kind of approach is well established in the field of
273 large-scale environmental assessment and RS-based studies (Sajib et al., 2025a; Uddin et al.,
274 2025a).

275 Although there is no clear guidance for best split/ratio practice in the data science field and
276 ML/AI applications, specifically for environmental modelling. Therefore, researchers use
277 common split ratios for training and testing, including 50:50, 70:30, 67:33, and 80:20,
278 depending on the sample sizes and model complexity (Joseph & Vakayil, 2022; Sarker, 2021).
279 Among these, the 80:20 split is justified by the renowned Pareto principle, which emphasizes
280 the critical aspect that produces the most significant results (Joseph, 2022), whereas the 70:30
281 splits are widely utilized for highly complex models (Sivakumar et al., 2024). These types of
282 splits/ratios are well established and frequently utilized in the field of environmental modelling
283 and ML/AI applications in order to reduce overfitting and more accurate parameter estimation
284 with the moderate sample size used in this research (Joseph & Vakayil, 2022; Luo et al., 2025;
285 Nguyen et al., 2021). In accordance with recommendations from previous studies (Krishnaraj
286 & Honnasiddaiah, 2022; Joseph, 2022; Luo et al., 2025; Peterson et al., 2020; Quang et al.,
287 2023; Zhou et al., 2024) and prior research conducted by the authors (Sajib et al., 2025b; Uddin
288 et al., 2022a; Uddin et al., 2023), the research used 80:20 and 70:30 random splits for
289 supervised, stacking-ensembles, voting-based ensemble, and statistical models (see section
290 2.6) to ensure a comprehensive model training and performance evaluation with the aggregated
291 multi-year dataset (2016–2021). To further evaluate model generalizability and temporal
292 transferability, the study used two independent datasets (2022 and 2023) for validation
293 purposes. On the other hand, since the equation models solely rely on explicit analytical
294 equations rather than data-driven techniques, and random train/test splitting may reduce
295 parameter stability, therefore, the full dataset was utilized for model training and tested with
296 the 2022 dataset. However, it's worth noting that the predictive capability of equation-based
297 models was evaluated using an independent dataset (2023) to provide an unbiased and thorough
298 evaluation of their generalizability and transferability.

299 **2.5 Feature engineering**

300 The research utilized five embedding-based models, including linear regression (LR), decision
301 trees (DT), extreme gradient boosting (XGB), k-nearest neighbor (KNN), random forest (RF);
302 and seven wrapper-based feature engineering models, such as recursive feature elimination –
303 Linear regression (RFE-LR), recursive feature elimination–random forest (RFE-RF), recursive
304 feature elimination cross validated – linear regression (RFECV-LR), recursive feature
305 elimination cross validated – random forest (RFECV-RF), sequential feature selector – random
306 forest (SFS-RF), extreme gradient boosting – grid search/random search (XGB-GS/XGB-RS),
307 and SHapley Additive exPlanations–K-nearest neighbors (SHAP-KNN), following the
308 methodology of Uddin et al. (2023) for feature optimization. A detailed description of these
309 approaches, implementation methodology, and rationale for selection can be found in the
310 supplementary materials-1 as a continuation of section 2.5.

311 **2.6 Model development process**

312 In this research, a total of 2101 advanced ML/AI and statistical models were developed to
313 retrieve DOX concentrations from RS products (Table 1). These models comprise five main
314 groups, including supervised learning, stacking-ensembles, equations, voting-based
315 ensembles, and statistical model. Each group of models builds on distinct architecture with
316 unique characteristics. A brief summary of these models can be found in Table 1. Additionally,
317 developed model(s) are selected based on prior research conducted by authors (Sajib et al.,
318 2024b; Uddin et al., 2022a; 2022b; 2023) and recommended by results of various RS-based
319 DOX retrieval studies (Batur & Maktav, 2019; Dong et al., 2023; Gao et al., 2024; Guo et al.,
320 2021; Li et al., 2020; Luo et al., 2025; Salas et al., 2022; Sharaf El Din & Zhang, 2017; Shi et
321 al., 2023; Tian et al., 2024; Toming et al., 2024; Yang et al., 2023). All models were trained on
322 average annual data from 2016 to 2021 for S3-OLCI and MS products. Moreover, in the
323 validation phase, two independent datasets (2022 and 2023) were utilized with the best-

324 performing model. A detailed description of the various model development processes can be
 325 found in the supplementary materials-1 as a continuation of section 2.6.

326 Table 1: Summary of models and dataset utilized during model training and validation phase in this research.

Model group	Definition	Base models	Feature engineering (Y = Yes /N = No)	No of develop models	Dataset used in training and testing phase	Dataset used in validation Phase	Outlier treatment
(i) Supervised learning model	refer to standalone ML/AI algorithms	AdaB, BPNN, BRR, CatB, DNN, DT, ExDT, GAM, GB, GPR, GRU, HM, KNN, KNNR, LARS, LGBM, LSTM, MLP, OMP, RANSAC, RF, SGD, SVM/SVR, XGB	N	28	32	64	Four outliers (Zscore, MoD-Z, LoF, IsoF) treated datasets
(ii) Stacking-ensemble model	refer to aggregation-based approaches (using “StackingRegressor” and “itertools.combinations” function in Python platform) that combined with a meta-learner (here, Linear Regression model) to produce the final output	AdaB, CatB, DT, GB, GPR, KNN, MLP, LR, RF, SVM, XGB	N	2047	8	16	LoF treated datasets (selected based on supervised model performance result)
(iii) Equation model	refer to non-linear pre-defined mathematical equation-based regression models	Arctan, Cosine, Equation, Exponential, Hyperbolic, Inverse power, Polynomial, Power, Quadratic, Cubic, Reciprocal,	Y	14	96	192	LoF treated datasets (selected based on supervised model performance result) with 12 feature selection methods

		Sigmoid, Sine, Tangent					
(iv) Voting-based ensemble model	refer to aggregation-based approaches (using VotingRegressor function in Python platform) that utilized multiple base regressor/learner models to develop an ensemble meta-estimator model	CatB, GPR, KNN, MLP, RF, SVM, XGB	N	9	8	16	LoF treated datasets (selected based on supervised model performance result)
(v) Statistical model	refer to statistical regression models	OLS, PCR, PLS	Y	3	8	16	LoF treated datasets (selected based on supervised model performance result) with 12 feature selection methods

327 [here = Adaboost (AdaB), Back propagation neural network (BPNN), Bayesian ridge regression (BRR), Catboost (CatB),
328 Decision trees (DT), Deep neural network (DNN), Extra decision trees (ExDT) , Extreme gradient boosting (XGB) , Gated
329 recurrent unit networks (GRU), Gaussian process regression (GPR), Generalize adaptive model (GAM), Gradient boosting
330 (GB), Huber model (HM), K-nearest neighbour (KNN), Kernel ridge regression (KNNR), Least-angle regression (LARS),
331 Light gradient boosting machine (LGBM), Long short-term memory (LSTM), Multilayer perceptron (MLP), Ordinary least
332 squares (OLS), Orthogonal matching pursuit (OMP), Probabilistic random forest (PRF), Principal components regression
333 (PCR), Partial least squares regression (PLS), Random forest (RF), Random sample consensus (RANSCA), Stochastic
334 Gradient Descent (SGD), Support vector machine/Support vector regression (SVM/SVR)]

335 2.7 Model performance assessment and optimal model selection

336 The research utilized four widely utilized statistical metrics, including root mean square error
337 (RMSE - Eq. 3), mean square error (MSE - Eq. 4), mean absolute error (MAE - Eq. 5), and
338 percentage of absolute bias error (PABE - Eq. 6), to measure the performance of various
339 models. A detailed description of these metrics and the rationale for their selection can be found
340 in supplemental materials-1 as a continuation of subsection 2.6.2. Additionally, the
341 performance metrics uncertainty was assessed using non-parametric bootstrap resampling with
342 2000 iterations and a 95% confidence interval (CI) percentile method (Rodrigues et al., 2019;
343 Thai et al., 2014). The equations of these metrics are as follows:

344
$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (DOX_{actual} - \widehat{DOX}_{predicted})^2}$$
 (3)

345
$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (DOX_{actual} - \widehat{DOX}_{predicted})^2$$
 (4)

346
$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |DOX_{actual} - \widehat{DOX}_{predicted}|$$
 (5)

347
$$\text{PABE} = \frac{1}{m} \sum_{i=1}^m \frac{|DOX_{actual} - \widehat{DOX}_{predicted}|}{DOX_{actual}} \times 100$$
 (6)

348 where DOX_{actual} is the actual value and $\widehat{DOX}_{predicted}$ is its corresponding predicted value.

349 m is the total number of observations.

350 Furthermore, a comprehensive rank-sum approach was utilized, following the methodology of
 351 Uddin et al. (2023), to identify the optimal model. This method involves assigning individual
 352 ranking values for each metric, where higher and lower values indicate excellent and poor
 353 performance, respectively (Sajib et al., 2024b; Uddin et al., 2022a). A detailed description of
 354 this approach can be found in Uddin et al. (2023). A major advantage of the rank sum approach
 355 is that it balances the consideration of multiple metrics (Diganta et al., 2025). In this study,
 356 selecting the best model was a crucial step because the best-performing model was
 357 subsequently utilized for model validation with independent datasets. The performance metrics
 358 (RMSE, MSE, MAE, and PABE) and sensitivity score (R^2) were utilized to evaluate the
 359 optimal model.

360 **2.8 Model retrieving error evaluation**

361 The research utilized the Percent of Relative Error Index (PREI) and Percentage of Bias
 362 (PBIAS) metrics to estimate model retrieval errors. Among them, PREI was utilized to measure
 363 the relative error at each monitoring/sampling site, whereas PBIAS measures the systematic
 364 bias between actual and retrieved DOX concentrations (Sajib et al., 2024b). These metrics
 365 express results in percentages. A PREI and PBIAS value close to 0 indicates low error, while
 366 a larger value indicates significant bias in retrieving data (Uddin et al., 2022a). Furthermore, a
 367 positive ('+') and negative ('-') PREI value indicates overestimation and underestimation of

368 the result, respectively (Uddin et al., 2022a). These metrics have been used in numerous studies
 369 to evaluate model prediction error (Aslan et al., 2022; Ozdemir et al., 2023). A detailed
 370 description of these metrics can be found in Sajib et al. (2024b). The PREI (Eq. 7) and PBIAS
 371 (Eq. 8) formulas are defined as follows:

$$372 \quad PREI = \left(\frac{DOX_{actual} - \widehat{DOX}_{predicted}}{DOX_{actual}} \right) \times 100 \quad (7)$$

373 where DOX_{actual} actual score for i^{th} sample and $\widehat{DOX}_{predicted}$ is the mean predicted score.

$$374 \quad PBIAS = \left(\frac{\sum_{i=1}^m (DOX_{actual} - \widehat{DOX}_{predicted})}{DOX_{actual}} \right) \times 100 \quad (8)$$

375 where DOX_{actual} actual score for i^{th} sample and $\widehat{DOX}_{predicted}$ is the mean predicted score.

376 **2.9 Model sensitivity and uncertainty analysis**

377 Typically, model sensitivity is utilized to assess how model input parameters affect the target
 378 variable (Šiljić Tomić et al., 2018). The study utilized the R^2 (Eq. 9) for model sensitivity
 379 assessment, a common approach for assessing model sensitivity (Diganta et al., 2025; Fan et
 380 al., 2018; Uddin et al., 2025a). R^2 value close to 1 indicates a high level of agreement, whereas
 381 a value close to 0 or $-\infty$ indicates a poor level of agreement between input and output variables
 382 (Chicco et al., 2021). The formula of R^2 is defined as follows:

$$383 \quad R^2 = 1 - \frac{\sum_{i=1}^m (DOX_{actual} - \widehat{DOX}_{predicted})^2}{\sum_{i=1}^m (DOX_{actual} - \overline{DOX}_{predicted})^2} \quad (9)$$

384 where m is the total number of observations, DOX_{actual} is the actual value, and the
 385 $\widehat{DOX}_{predicted}$ is the predicted value and $\overline{DOX}_{predicted}$ mean of the actual value.

386 On the other hand, the model uncertainty was quantified using inferential error bars with a 95%
 387 CI method, following the methodology of Cumming et al. (2007). In this method, the length of
 388 bars visually represents the amount of uncertainty, and the CI presents the possible range for
 389 the true mean. Similar methods have been utilized in numerous studies to measure model

390 uncertainty (Hofman et al., 2020; Uddin et al., 2024a; Van Allen et al., 2008). Furthermore,
391 following the methodology of Shaw (2017), a probability distribution function (PDF) and
392 cumulative distribution function (CDF) were utilized to visualize the uncertainty between
393 model development and validation datasets for both RS products.

394 **2.10 Geospatial mapping and statistical analysis**

395 In this research, empirical bayesian kriging (EBK) was utilized to illustrate the distribution of
396 DOX for modelled data, whereas the unclassed symbology (also known as choropleth
397 mapping) approach was used for the EPA dataset. Moreover, several statistical techniques were
398 utilized to understand the distribution (normal or non-normal), spatial autocorrelation, temporal
399 trends, and relationship within the utilized datasets using the Shapiro-Wilk statistics, the Mann-
400 Whitney U test, the Global Moran's I statistic, and the Mann-Kendall statistic. A detailed
401 description of these techniques can be found in the supplementary materials-1 as a continuation
402 of section 2.10.

403 **3. Result**

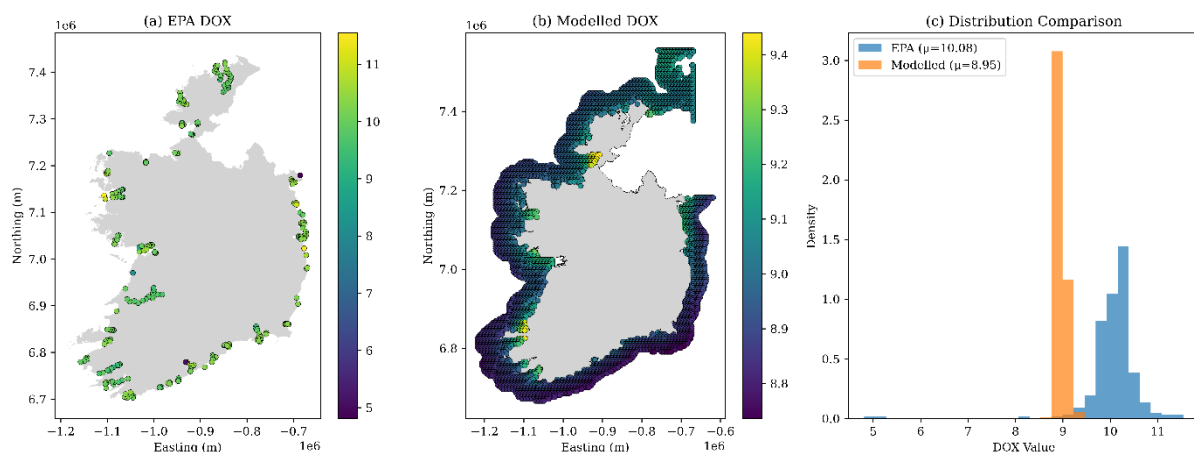
404 **3.1 Data distribution and outlier treatment**

405 **3.1.1 DOX concentration distribution**

406 In order to evaluate the differences between EPA and modelled DOX, the study measures the
407 descriptive statistics, spatial autocorrelation, and significance differences of the datasets used,
408 covering the periods of 2016-2021, 2022, and 2023. Fig. S1-S2 (supplementary materials-1)
409 show the descriptive statistics, whereas the spatial distribution of the S3-OLCI and MS
410 matchup DOX datasets are presented in Fig. 3, Fig. 4, and Fig. S3-S6 (supplementary materials-
411 1).

412 The research found a statistically significant positive spatial autocorrelation based on the
413 Global Moran's I statistics for S3-OLCI matchup EPA-sourced DOX datasets, including 2016-
414 2021 ($N = 289$, $I = 0.037$, $p\text{-value} < 0.05$), 2022 ($N = 214$, $I = 0.464$, $p\text{-value} < 0.05$), and 2023

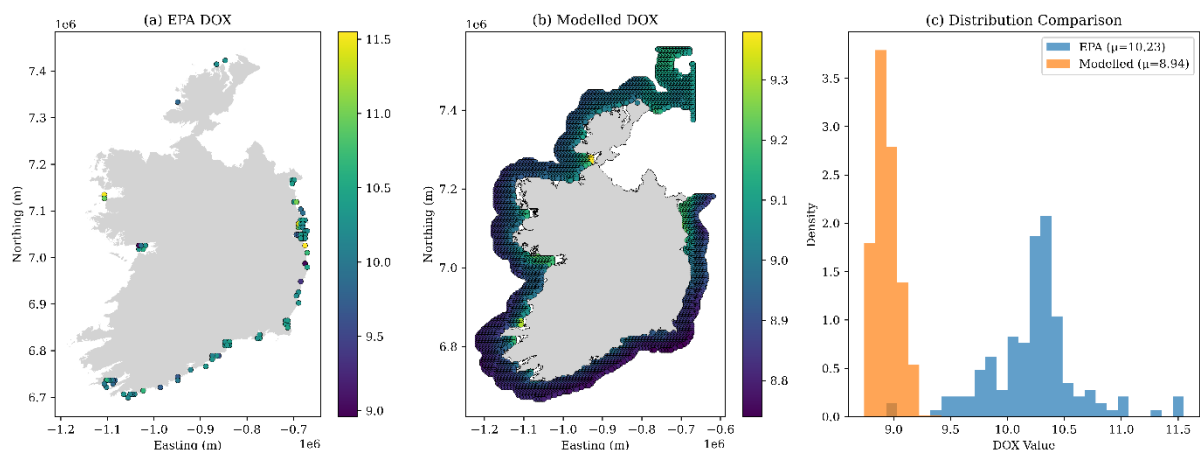
415 ($N = 179$, $I = 0.425$, $p\text{-value} < 0.05$) (Fig. 3 and Fig. S3-S4 – supplementary materials-1),
 416 indicating a spatial clustering rather than random spatial variation. Similarly, statistically
 417 significant positive spatial autocorrelation was also observed in modelled-derived DOX dataset
 418 (2016-2021: $N = 6466$, $I = 0.695$, $p\text{-value} < 0.05$; 2022: $N = 6495$, $I = 0.526$, $p\text{-value} < 0.05$;
 419 2023: $N = 6495$, $I = 0.5$, $p\text{-value} < 0.05$). Moreover, a statistically significant difference was
 420 found between the S3-OLCI matchup EPA and modelled data, which is justified by the Mann-
 421 Whitney U test result with a $p\text{-value} < 0.05$ (2016-2021: $U = 1844040.5$, $p\text{-value} < 0.05$; 2022:
 422 $U = 1374644$, $p\text{-value} < 0.05$; 2023: $U = 1155392$, $p\text{-value} < 0.05$), highlighting systematic
 423 differences in data despite similar spatial clustering characteristics.
 424 Furthermore, the Mann-Kendall test indicated there were no temporal trends during the study
 425 period in the EPA-sourced DOX (average Tau value = -0.01658 and average $p > 0.05$) and
 426 modelled-derived DOX (average Tau value = -0.54877 and average $p > 0.05$) datasets
 427 (supplementary materials-2). The absence of a significant temporal trend in the datasets could
 428 be due to temporal aggregation (2016-2021) applied during the model development, which
 429 limits the sensitivity of trend detection (Morbidelli et al., 2018). In the S3-OLCI matchup, the
 430 EPA data (L3-OLCI) for 2016-2021 range from 10.21 mg/L to 11.55 mg/L with a mean value
 431 of 10.08 ± 0.56 mg/L, whereas the modelled data range from 8.74 mg/L to 9.44 mg/L with a
 432 mean value of 8.95 ± 0.11 mg/L. It is noteworthy that similar variations were also observed for
 433 the 2022 and 2023 datasets (Fig. S1).



434

435 Fig. 3. Spatial distribution of S3-OLCI matchup DOX (mg/L) measurements for 2016-2021 dataset.
 436 On the other hand, for the MS products, both EPA-sourced (2016-2021: $N = 149$, $I = 0.098$, p -
 437 value < 0.05 ; 2022: $N = 117$, $I = 0.2$, p -value < 0.05 ; and 2023: $N = 95$, $I = 0.033$, p -value $<$
 438 0.05) and modelled-derived (2016-2021: $N = 6003$, $I = 0.589$, p -value < 0.05 ; 2022: $N = 5890$,
 439 $I = 0.49$, p -value < 0.05 ; and 2023: $N = 5890$, $I = 0.612$, p -value < 0.05) DOX showed
 440 statistically significant positive spatial autocorrelation across all study periods, indicating
 441 spatial clustering (Fig. 4 and Fig. S5-S6 – supplementary materials-1). In addition, the Mann-
 442 Whitney U test result indicated statistically significant differences between MS matchup EPA
 443 and modelled-derived DOX measurements (2016-2021: $U = 895604$, p -value < 0.05 ; 2022: U
 444 $= 680100$, p -value < 0.05 ; 2023: $U = 549959.5$, p -value < 0.05).

445 In the MS matchup, the EPA dataset 2016-2021 ranges from 8.96 mg/L to 11.55 mg/L with a
 446 mean value of 10.23 ± 0.56 mg/L, whereas the modelled data range from 8.74 mg/L to 9.44
 447 mg/L with a mean value of 8.95 ± 0.11 mg/L (Fig. S2 – supplementary materials-1), indicating
 448 high inconsistency between the EPA and the modelled data, which is further validated by the
 449 Mann-Whitney U test result. In EPA data for both products, a high standard deviation (SD) is
 450 observed across the products (L3-OLCI = 0.52 mg/L; L3-MS = 0.88 mg/L; L4-OLCI = 0.52
 451 mg/L; L3-MS = 0.44 mg/L), indicating high variation among the monitoring/sampling sites,
 452 whereas a low SD is observed for modelled data (SD = 0.13 for all datasets), indicating low
 453 variation in the data (Fig. S2 – supplementary materials-1).

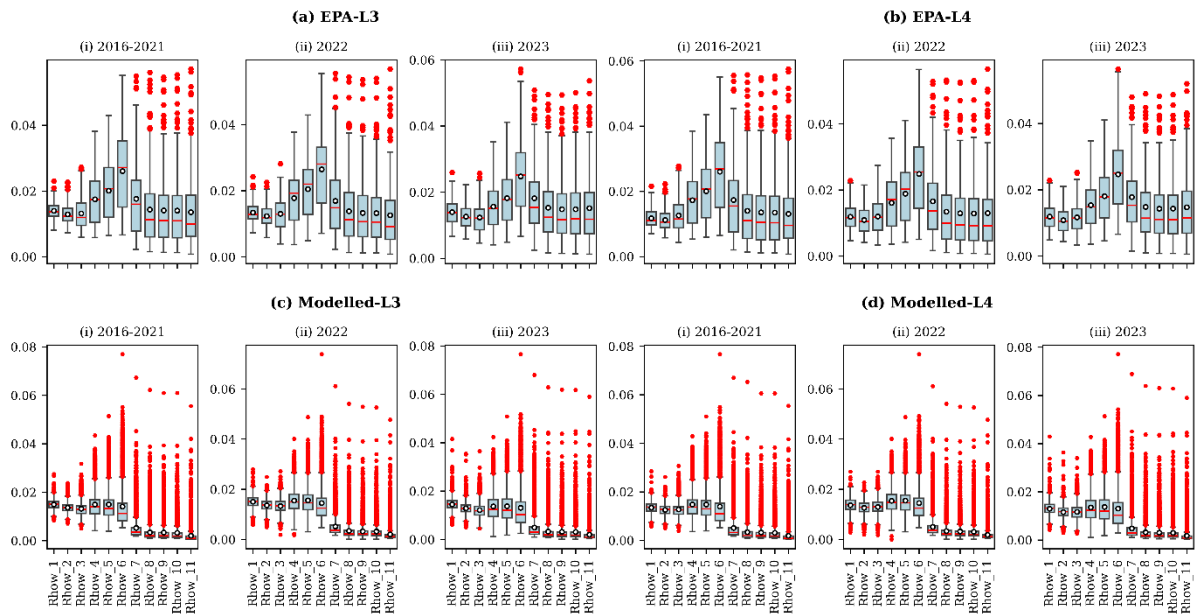


454

455 Fig. 4. Spatial distribution of MS matchup DOX (mg/L) measurements for 2016-2021 dataset.

456 **3.1.2 Characteristics of RS data**

457 Fig. 5 and Fig. 6 show the statistical distribution of RS bands (here referred to as Rhow) across
458 four different datasets - (i) L3-OLCI, (ii) L4-OLCI, (iii) L3-MS, and (iv) L4-MS, respectively.
459 It can be seen from Fig. 5 that, in the EPA matchup with S3-OLCI products, all Rhow bands
460 are distributed between values 0 and 0.06, whereas modelled matchup datasets are distributed
461 between values 0 and 0.08. Across the datasets, Rhow_5 and Rhow_6 show high variability
462 compared to other bands (Fig. 5). Typically, these bands are crucial for case-2 waters, as they
463 are highly sensitive to pigments, such as CHL (Diganta et al., 2024; Sajib et al., 2025b). In
464 contrast, the red bands (refer to Rhow_7 to Rhow_10) and the Rhow_11 band showed high
465 absorption compared to other bands. Similar characteristics for OLCI bands have been reported
466 in previous validation and algorithm assessment studies (Diganta et al., 2024; Gleratti et al.,
467 2024; Mikelsons et al., 2022; Sajib et al., 2025b; Sagan et al., 2020; Uddin et al., 2025a;
468 Vanhellefont & Ruddick, 2021). Among the different levels of data, a consistent trend was
469 observed for L4 datasets (Fig. 5b), whereas a wider distribution was found for L3 datasets,
470 indicating higher variability. Moreover, Rhow bands showed statistically significant positive
471 spatial autocorrelation across all the study periods (e.g., average Moran's I for the 2016-2021
472 dataset: EPA-L3 = 0.678, EPA-L4 = 0.687, Modelled-L3 = 0.645, Modelled-L4 = 0.644),
473 indicating spatial clustering characteristics in Rhow bands. Similar characteristics for RS data
474 have been reported in previous studies (Griffith & Chun, 2016; Karasiak et al., 2022; Warner
475 & Shank, 1997). Detailed descriptions of measured Moran's I with p-values for each Rhow
476 band (S3-OLCI) can be found in Table S3 (supplementary materials-1).



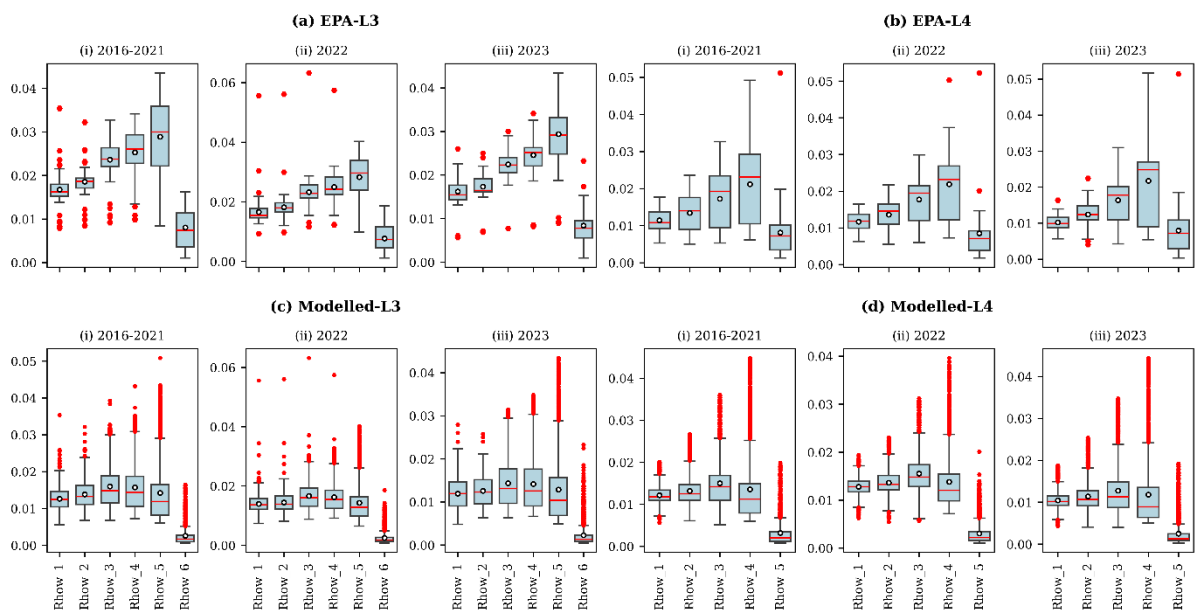
477

478 Fig. 5. Statistical summary of various bands (Rhow) of S3-OLCI products. (For a better interpretation of the colour
 479 references in this figure, readers are encouraged to read the web version of this article)

480 From Fig. 6, it can be seen that the L4 data bands had fewer outliers than the L3 data in EPA
 481 matchups. The presence of a few outliers in the L4 data is attributed to spatial and temporal
 482 averaging, interpolation, and smoothing techniques (Brando et al., 2024; Colella et al., 2023;
 483 Garnesson et al., 2024; Sajib et al., 2025b). In L3 data, the maximum and minimum reflectance
 484 values were found for the Rhow_4 (green band) and Rhow_5 (red band) bands. In contrast, in
 485 L4 data, the maximum and minimum reflectance values were found for the Rhow_5 (green
 486 band) and Rhow_6 (red band) bands, respectively (Fig. 6a, b). The findings are consistent with
 487 earlier studies where green bands showed high scattering and red bands showed low absorption
 488 in various waterbodies (Diganta et al., 2024; Gleratti et al., 2024; Mikelsons et al., 2022;
 489 Vanhellemont & Ruddick, 2021). The MS datasets show a consistent trend for the L4 datasets
 490 (Fig. 6b), whereas a wider distribution is found in the L3 datasets, indicating higher variability.
 491 However, similar to EPA datasets, a fluctuation in the distribution of Rhow bands was found
 492 in the modelled datasets for 2022 and 2023 compared to the 2016-2021 dataset. Such variability
 493 could be the combination of various factors, including sensor-specific issues, environmental

494 factors, averaging multi-year data, data harmonization issues, etc. (Chen et al., 2022; Janga et
 495 al., 2023; Stein et al., 2018; Zhu et al., 2022b).

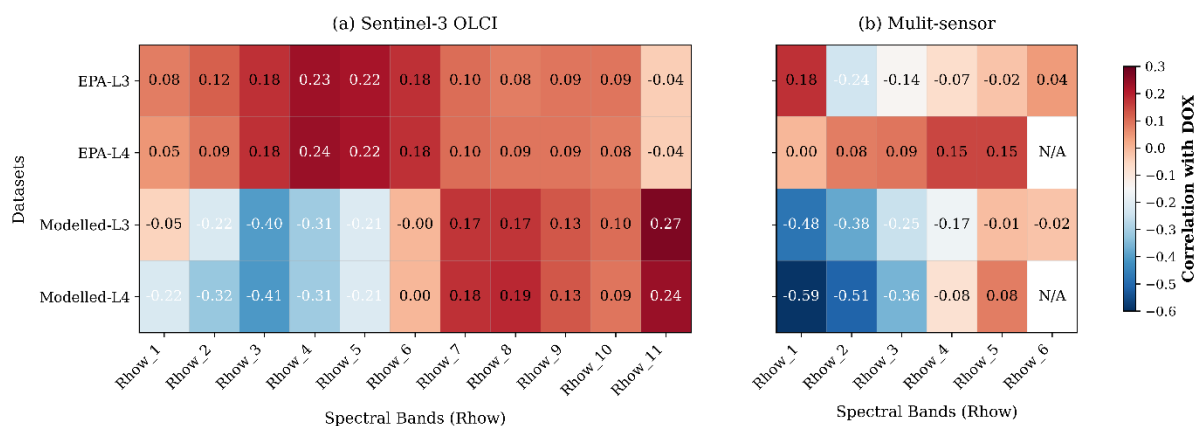
496 Based on the statistical overview of the datasets, the modelled L4 data for both products (OLCI
 497 and MS) found relatively less variability than the L3 data, indicating more consistent
 498 reflectance values during the study period. These characteristics of the L4 datasets are
 499 consistent with earlier studies, where L4 datasets are ascribed to be relatively stable due to their
 500 composite nature (Brando et al., 2024; Colella et al., 2023; Garnesson et al., 2024; Uddin et al.,
 501 2025a). Furthermore, a statistically significant positive spatial autocorrelation was found in the
 502 MS dataset for all Rhow bands (e.g., average Moran's I for the 2022 dataset: EPA-L3 = 0.386,
 503 EPA-L4 = 0.736, Modelled-L3 = 0.674, and Modelled-L4 = 0.674), indicating a non-random
 504 distribution. The findings are aligned with previous studies (Griffith & Chun, 2016; Karasiak
 505 et al., 2022; Warner & Shank, 1997). Detailed descriptions of measured Moran's I with p-
 506 values for each Rhow band (MS products) can be found in Table S4 (supplementary materials-
 507 1).



508
 509 Fig. 6. Statistical summary of various bands (Rhow) of MS products. (For a better interpretation of the colour
 510 references in this figure, readers are encouraged to read the web version of this article)

511 3.1.3 Relationship between Rhow bands and DOX concentrations

512 For the purposes of understanding the interrelationship between Rhow bands and DOX
 513 concentrations, the research utilized the Spearman rank correlation analysis that could be
 514 effective to mitigate the problem of multicollinearity issues. Fig. 7 shows the relationship
 515 between various Rhow bands and DOX concentrations for both OLCI and MS products. It can
 516 be seen from Fig. 7(a) that the Rhow_3 to Rhow_6 bands showed a slight positive correlation
 517 with EPA-sourced DOX, whereas Rhow_3 shows a moderate negative correlation with
 518 modelled-derived DOX in L3 ($r = -0.4$) and L4-OLCI ($r = -0.41$), which is not very significant.
 519 The negative correlation indicates that the Rhow_3 band could affect the DOX retrieval process
 520 using RS data. On the other hand, no notable association was found between MS products (L3
 521 and L4) with EPA-sourced DOX (Fig. 7b). However, a strong negative correlation (> 0.5) was
 522 found between Rhow bands (Rhow_1-Rhow_2) with MS products (L4) matchup modelled-
 523 derived DOX. Moreover, a moderate negative correction was observed for Rhow_1 ($r = -0.48$)
 524 with MS products (L3) matchup modelled-derived DOX. The above result indicates that,
 525 considering these relationships, a low multicollinearity problem could be expected in
 526 developing the DOX retrieval model.



527
 528 Fig. 7. Correlation between DOX concentrations and various bands (Rhow) for utilized RS products.

529 **3.1.4 Matchup datasets and outlier treatment result**

530 The research used the EUMETSAT (2021) recommended method for the matchup process and
531 four outlier treatment methods to execute extreme values. The matchup results for both the
532 EPA and the modelled data sites after removing the outlier can be found in Table S5
533 (supplementary materials-1). The same number of matchups (both EPA and modelled sites)
534 were found for both L3- and L4-OLCI data. However, the L3 and L4 datasets yielded different
535 matchup results for MS products. It could be due to the availability of the number of bands in
536 L3 (number of bands = 6) and L4-MS (number of bands = 5) products (Garnesson et al., 2019;
537 Saulquin et al., 2019). Across the study period for both EPA and modelled datasets (L3 and
538 L4), the Z-score, LoF, and IsoF methods retained around 90-95% of data, whereas the MoD-Z
539 method removed more than 15% of data (Table S5 - supplementary materials-1).

540 In this research, the best outlier treatment method was selected based on the balanced execution
541 of extreme values by outlier methods and its impact on the predictive performance of
542 supervised models. A total of 28 supervised models were evaluated for each outlier-treated
543 dataset to determine the suitable method that yields the most stable and accurate model
544 performance. Moreover, the LoF method is effective for heterogeneous and non-normally
545 distributed datasets (Alghushairy et al., 2021; Uddin et al., 2024a), which are the characteristics
546 of the data utilized in this study (see section 2.10). Furthermore, the selection of only
547 supervised models is to reduce computational time and manage the large volume of outcomes.
548 The performance results of various outlier treatment datasets with 28 supervised models can
549 be found in supplementary materials-3 and supplementary materials-4. Moreover, a summary
550 of the best-performing dataset based on various outlier methods with supervised models can be
551 found in Table S6 (supplementary materials-1). Finally, based on the best outlier method
552 outcomes (datasets), they were subsequently applied to train, test, and validate the stacking-
553 ensembles, equations, voting-based ensembles, and statistical models.

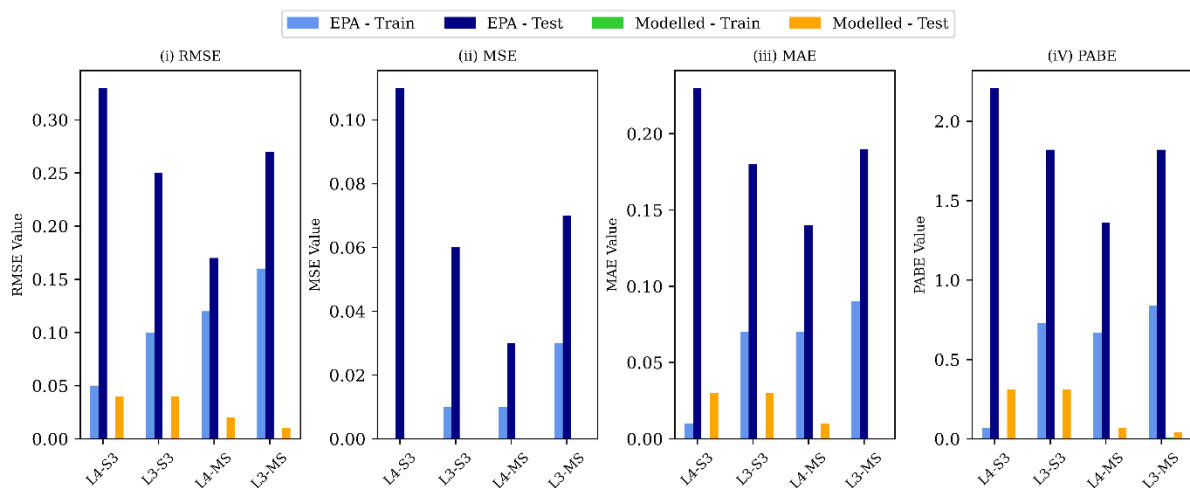
554 3.2 Performance of various model(s)

555 A group of models (supervised, stacking-ensembles, voting-based ensembles, equations, and
556 statistical model) were utilized in this research to retrieve DOX concentrations from various
557 RS products (see sub-section 2.6). Subsequently, performance metrics, such as RMSE, MSE,
558 MAE, and PABE with 2000 iterations and a 95% CI, were utilized to evaluate the performance
559 of these models. Overall, supervised learning models showed better performance and higher
560 accuracy compared to other models (Table S7-S9, supplementary materials-1). Therefore, this
561 section only highlights the results of supervised learning models, whereas other groups of
562 model results (Fig. S7-S9; Table S10-S13, supplementary materials-1) and performance of all
563 utilized model results with different split ratios (Table S14, supplementary materials-1) can be
564 found in the supplementary materials-1 as a continuation of section 3.2.

565 In this research, supervised models showed the highest accuracy compared to the other four
566 groups of models (stacking-ensembles, voting-based ensembles, equation, and statistical). For
567 instance, the KNN (k-nearest neighbour) model (configured with $n_neighbors = 11$, distance-
568 based weighting, and Euclidean L2 distance) performed best among the 28 supervised models
569 using the L4-S3 data (Table S7 - supplementary materials-1) in the EPA datasets. During the
570 training period, the model achieves an almost perfect fit (e.g., $MSE \approx 0$ with $CI: \pm 0.005$).
571 However, a slight overfitting issue was found in the testing dataset (e.g., $MSE = 0.11$ with CI
572 ± 0.07), indicating the model could memorize the training patterns but failed to generalize
573 unseen data (Fig. 8; Table S8-S9, supplementary materials-1). Moreover, a high fluctuation in
574 PABE from 0.07% ($CI \pm 0.06\%$) to 2.21% ($CI \pm 0.57\%$) is notable, indicating that the KNN
575 model may fail in predicting with new data. In contrast, the BGR (Blending gradient boosting-
576 random forest) model outperformed with the L3-S3-EPA data based on the cumulative ranking
577 score (Tables S8-S9, supplementary materials-1). However, similar to L4-S3, there is a notable
578 difference found between training (e.g., $RMSE = 0.1$ with $CI \pm 0.01$) and testing (e.g., $RMSE$

579 = 0.25 with CI \pm 0.06) performance, indicating that the model may struggle to maintain
 580 accuracy with unseen data. On the other hand, the GB (Gradient boosting) model showed
 581 excellent performance for the L4-MS datasets, whereas the KNN model (configured with
 582 $n_neighbors = 11$, distance-based weighting, and Manhattan L1 distance) showed excellent
 583 performance for the L3-MS datasets. Although an overfitting was found for both models during
 584 the testing phase (Table S8-S9, supplementary materials-1).

585 Further analysis with the modelled DOX concentration matchups, the ExDT (Extra decision
 586 tress, configured with $max_depth = 20-30$, $min_samples_leaf = 1$, $min_samples_split = 2$,
 587 $n_estimators = 200$) model showed an excellent performance across all products with very low
 588 uncertainty, which is justified by the narrow gap between confidence interval values (Table
 589 S7-S9, supplementary materials-1). However, the S3-OLCI products show slightly higher test
 590 errors and bias compared to MS products, indicating an overfitting issue during the testing
 591 period (Fig. 8). The results of ExDT demonstrated that ensemble models could be effective to
 592 handle the complex and nonlinear relationships of attributes compared to the standalone
 593 decision tree model(s) (Abbas et al., 2024; Sakib et al., 2025; Zhang et al., 2021).



594

595 Fig. 8. Performance metrics results of supervised learning models during training and testing period for S3-OLCI
 596 and MS products. (here L = Level, S3 – Sentinel 3 OLCI and MS = Multi sensor)

597 3.3 Model retrieving error assessment

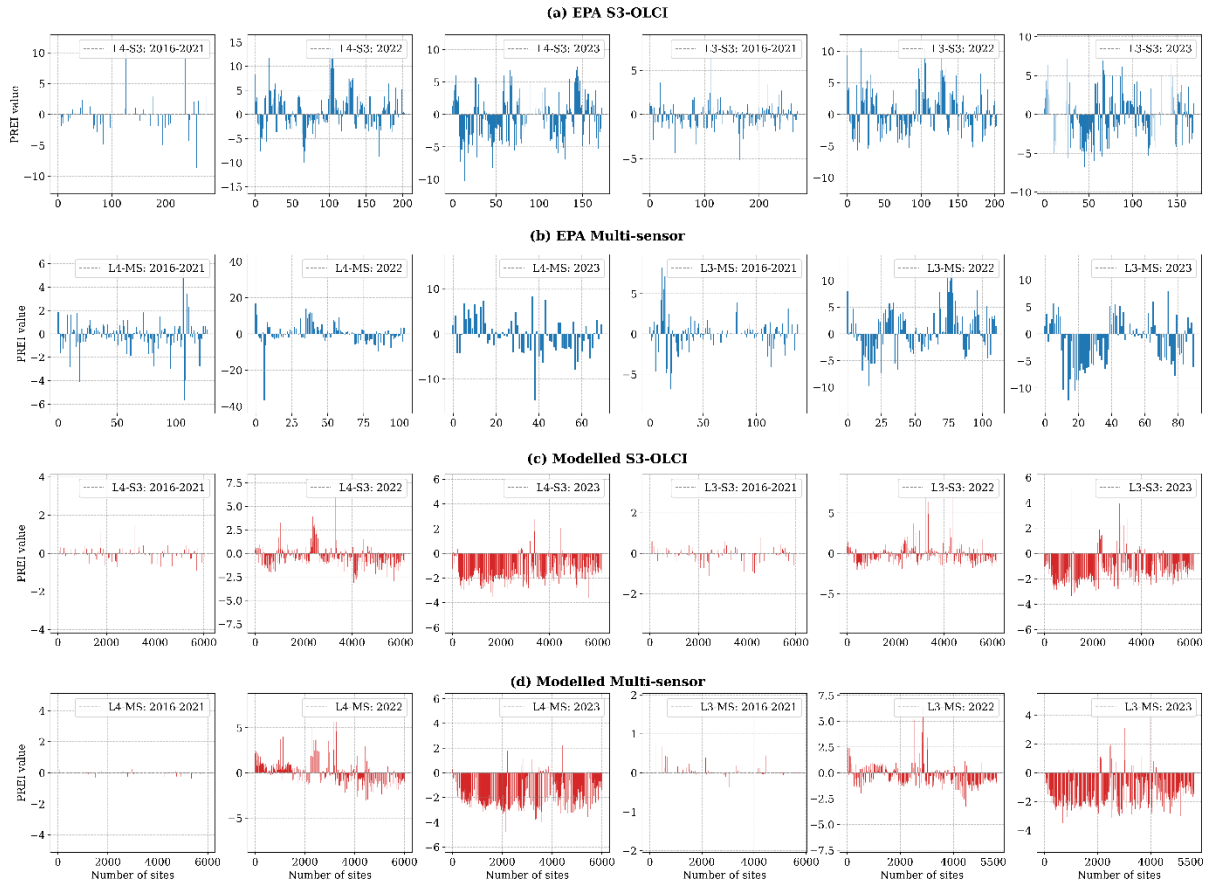
598 In order to assess the difference between actual and retrieval DOX concentrations, the study
599 utilized PREI at each monitoring/sampling site. Fig. 9 shows the monitoring/sampling site-
600 wise PREI score, while Fig. 10 represents the PBIAS score for each dataset based on the best-
601 performing supervised learning models. This section only highlights the PREI and PBIAS
602 results of supervised learning models as they showed the highest accuracy across different
603 datasets, whereas other groups of models retrieving error results (Fig. S10-S15, supplementary
604 materials 1) can be found in the supplemental materials-1 as a continuation of section 3.3.

605 It can be seen from Fig. 9 that the mean PREI values for both products in the EPA dataset range
606 from negative to positive or close to zero. In EPA datasets, the best-performing models showed
607 underestimation for 2016-2021 datasets (mean = -0.11 for L3-S3; mean = -0.16 for L4-MS)
608 and 2023 (mean = -0.48 for L4-S3; mean = -0.35 for L3-S3; mean = -0.44 for L4-MS; mean =
609 -1.51 for L3-MS), whereas overestimation was found for 2022 (mean = 0.67 for L4-S3; mean
610 = 0.35 for L3-S3; mean = 0.11 for L4-MS; mean = 0.59 for L3-MS). It is worth noting that L4-
611 S3 and L3-MS showed no overestimation or underestimation problem for the 2016-2021
612 dataset. Furthermore, a higher SD of PREI was found with increasing temporal resolution
613 (2016-2021, 2022, 2023) for the EPA datasets across products (Table S15-supplementary
614 materials 1).

615 Further analysis with modelled DOX concentration matchups showed that, compared to the
616 EPA datasets, modelled data showed lower error for the 2016-2021 dataset (e.g., -0.01 ± 0.26
617 for L4-S3; 0 ± 0.26 for L3-S3) across products, whereas an underestimation was found for the
618 2022 dataset (e.g., -0.12 ± 1.23 for L4-MS; -0.24 ± 1.1 for L3-MS) and 2023 dataset (e.g., -
619 1.42 ± 1.09 for L4-S3; -1.30 ± 1.1 for L3-S3) across products (Table S15-supplementary
620 materials 1). Moreover, similar to EPA datasets, a higher SD with increased temporal
621 distribution was found for the modelled dataset across products. A high SD indicates greater

622 inconsistency, whereas a low SD indicates consistent performance of the models (Barde &
623 Barde, 2012; Bella et al., 2005). It is worth highlighting that for both datasets across products,
624 a high variability was found for 2023, indicating poor performance with the independent
625 dataset.

626 Further error results based on PBIAS showed that S3-OLCI products represent a higher
627 negative bias (underestimation) in the modelled dataset over time, whereas there is less
628 negative bias for the EPA datasets (Fig. 10). In contrast, MS datasets show a more stable bias
629 value (with a smaller magnitude) over time. For example, in 2016–2021, both modelled (0.01%
630 for L4-MS; 0.00% for L3-MS) and EPA (-0.14% for L4-MS; 0.03% for L3-MS) biases were
631 very close to zero, indicating high accuracy. However, from 2022 to 2023, the resulting biases
632 for the EPA (0.76% for L4-MS: 2022; -1.33% for L3-MS: 2023) and modelled (-0.38% for L4-
633 MS: 2022; -1.39% for L4-MS: 2023) datasets become more negative, indicating a notable
634 difference between actual and retrieved DOX concentrations over time.



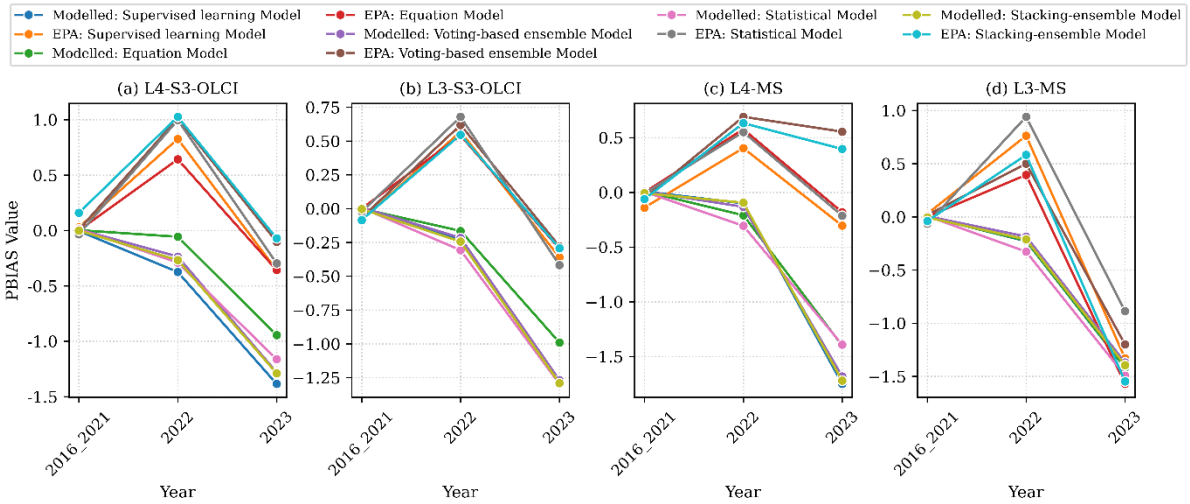
635

636

637

638

Fig. 9. Actual vs retrieved DOX concentrations difference at each monitoring/sampling site across the study domain for the best-performing supervised learning models. (For a better interpretation of the colour references in this figure, readers are encouraged to read the web version of this article)



639

640

641

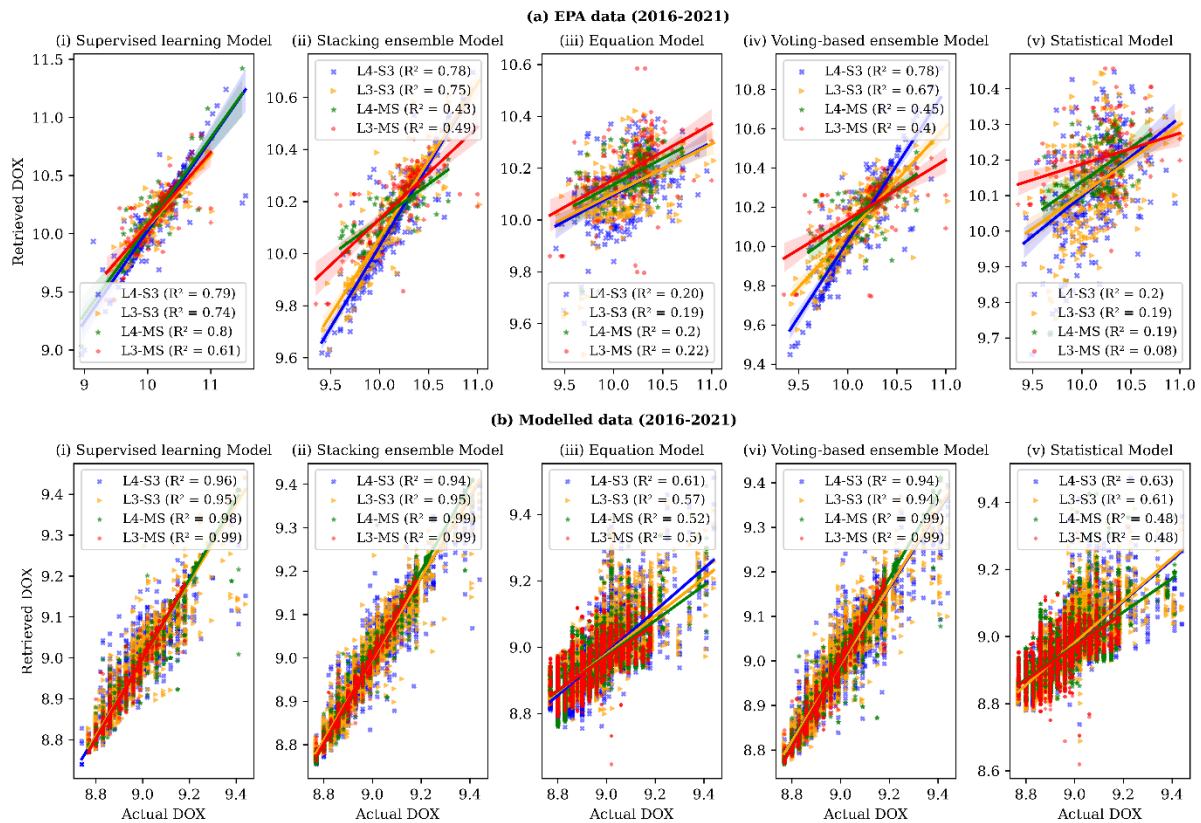
Fig. 10. Percentage of bias (PBIAS) between actual and retrieved DOX concentrations for various best-performing models.

642 3.4 Sensitivity of DOX retrieval model(s)

643 To assess the sensitivity of the best-performing models, R^2 statistics were used in this study.

644 Fig. 11 shows the scatter plot that compares the sensitivity of different best-performing models

645 in retrieving DOX. It can be seen from Fig. 11 that the supervised, stacking-ensembles, and
646 voting-based ensembles models show a consistent sensitivity across all datasets. In contrast,
647 the equations and statistical models show poor sensitivity, indicating these models may not be
648 suitable for retrieving DOX concentrations using RS data. In terms of DOX matchup data,
649 modelled data showed higher sensitivity ($R^2 > 0.90$) than the EPA data ($R^2 < 0.48$), except for
650 equations and statistical models. Moreover, the S3-OLCI products outperform the MS
651 products, specifically for the L4 datasets in most cases. This result indicates that S3-OLCI
652 products provide more reliable data for DOX measurement than the MS products. However,
653 an exception occurred in the voting-based ensemble models, where the MS products ($R^2 >$
654 0.95) outperformed the S3-OLCI ($R^2 > 0.9$) products. The notable gap in performance between
655 the EPA and the modelled data indicates that the models are not well fitted for effectively
656 retrieving the ground-measured DOX concentrations from the RS products in TrC waters
657 (Uddin et al., 2025a), as well as highlights potential limitations in applying these models to
658 real-world data without optimization.



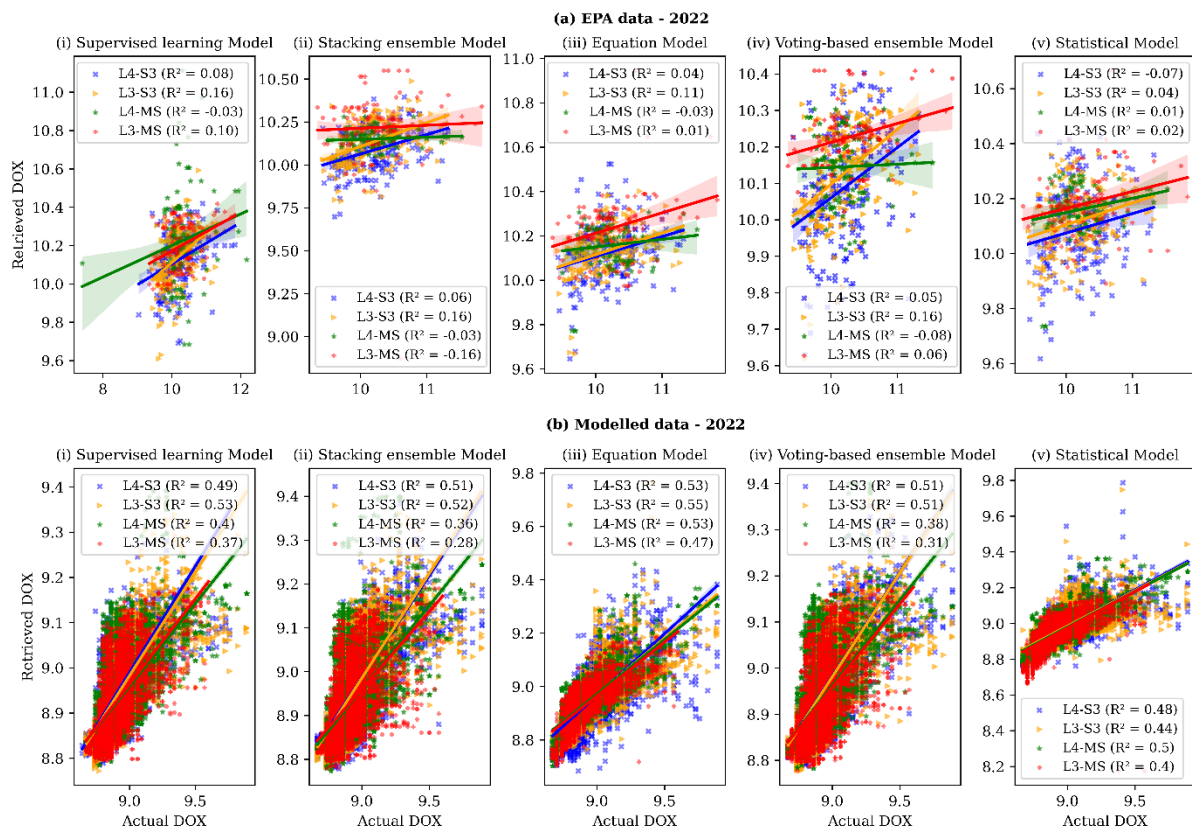
659

660 Fig. 11. Relationship between actual and retrieved DOX concentrations for various models for the period of 2016-
 661 2021 during the training phase. (For a better interpretation of the colour references in this figure, readers are
 662 encouraged to read the web version of this article)

663 3.5 Reliability of DOX retrieval model(s)

664 To address the reliability of best-performing DOX concentration retrieval models, two
 665 independent (2022 and 2023) datasets were utilized in the validation process. Fig. 12 and Fig.
 666 S15 (supplementary materials-1) provide validation performance for the best-performing
 667 model with independent datasets. It is apparent from these figures that none of the models
 668 showed stable performance during validation, indicating low sensitivity and high uncertainty
 669 among the best-performing models. In terms of 2022 DOX data, the modelled dataset showed
 670 better performance than the EPA datasets; however, the R^2 value was ≤ 0.6 for all models
 671 (supervised, stacking ensemble, equation, voting-based ensemble, and statistical), indicating
 672 marginal performance of the models (Chicco et al., 2021; Gupta et al., 2024). On the other
 673 hand, in terms of products, the S3-OLCI products showed more stable performance ($R^2 > 0.48$)
 674 than the MS products on modelled data in 2022. Turning now to the 2023 datasets, notable

675 poor results were found compared to the 2022 dataset across products, indicating the limitation
 676 of the best-performing models in generalizing unseen data.

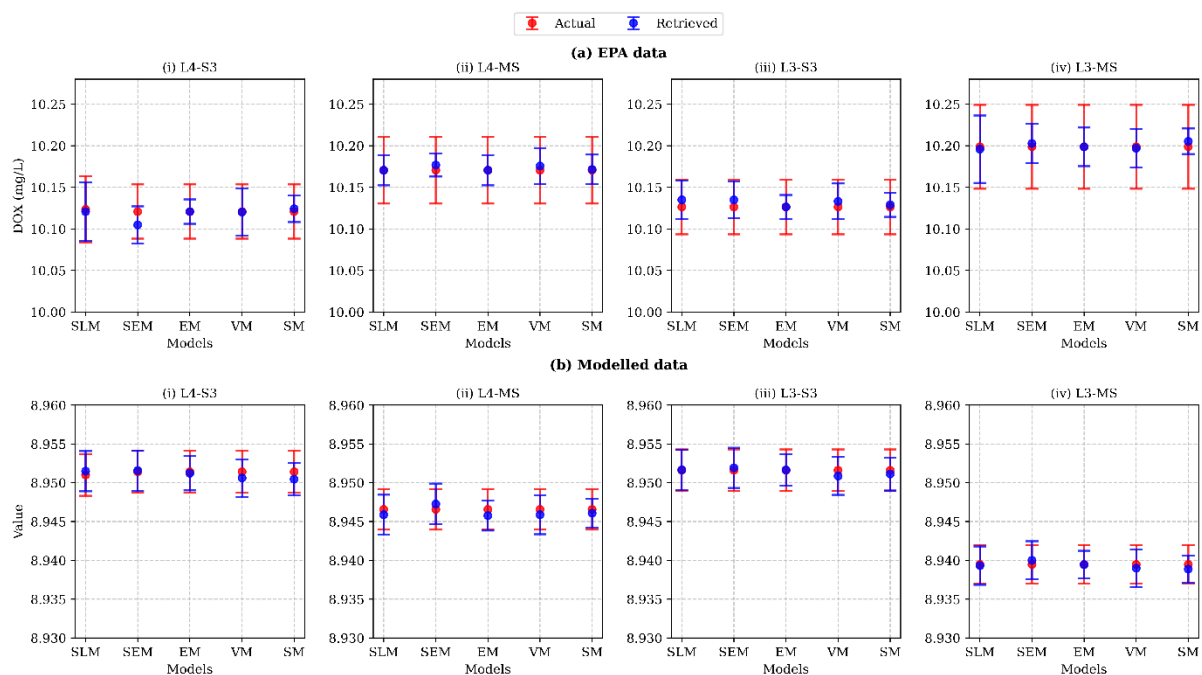


677
 678 Fig. 12. Relationship between actual and retrieved DOX concentrations for various models for the period of 2022
 679 during validation. (For a better interpretation of the colour references in this figure, readers are encouraged to read
 680 the web version of this article)

681 3.6 Uncertainty of DOX retrieval model(s)

682 In order to assess the uncertainty of best-performing DOX models, the present research used
 683 an inferential error bar approach with a 95% CI. Fig. 13 provides a summary of the uncertainties
 684 of the best-performing models across different products and datasets. In the EPA datasets, the
 685 retrieved DOX values are consistently lower than the actual DOX across all models. Moreover,
 686 the spread of CI varies, indicating different levels of uncertainty between the models. In terms
 687 of the variability of the datasets, the L4-S3 dataset shows a closer CI, whereas the L3-MS
 688 dataset shows a wide CI, indicating low and high uncertainty among the applied models
 689 (Cumming et al., 2007; IOCCG, 2019; Mikelsons et al., 2022). On the other hand, supervised
 690 learning models showed closer CI, suggesting relatively less uncertainty than other models

691 (stacking-ensembles, voting-based ensemble, equation, and statistical). However, product
 692 performance varied across different datasets. After supervised learning algorithms, stacking-
 693 ensembles and voting-based ensembles models showed less uncertainty based on actual and
 694 retrieved DOX in the EPA dataset. Notably, a significant uncertainty was found in the equation
 695 and statistical models, indicating an unsuitable approach for retrieving DOX concentrations
 696 from TrC waters. Further analysis with the modelled data shows closer CI, indicating lower
 697 uncertainty and better performance of the models (Cumming et al., 2007). For modelled DOX,
 698 a low and high uncertainty was found in supervised and equation models. Moreover, the S3-
 699 OLCI showed closer CI than the MS products for the modelled datasets, indicating less
 700 variability in the S3-OLCI product (IOCCG, 2019; Mikelsons et al., 2022; Sajib et al., 2025b).



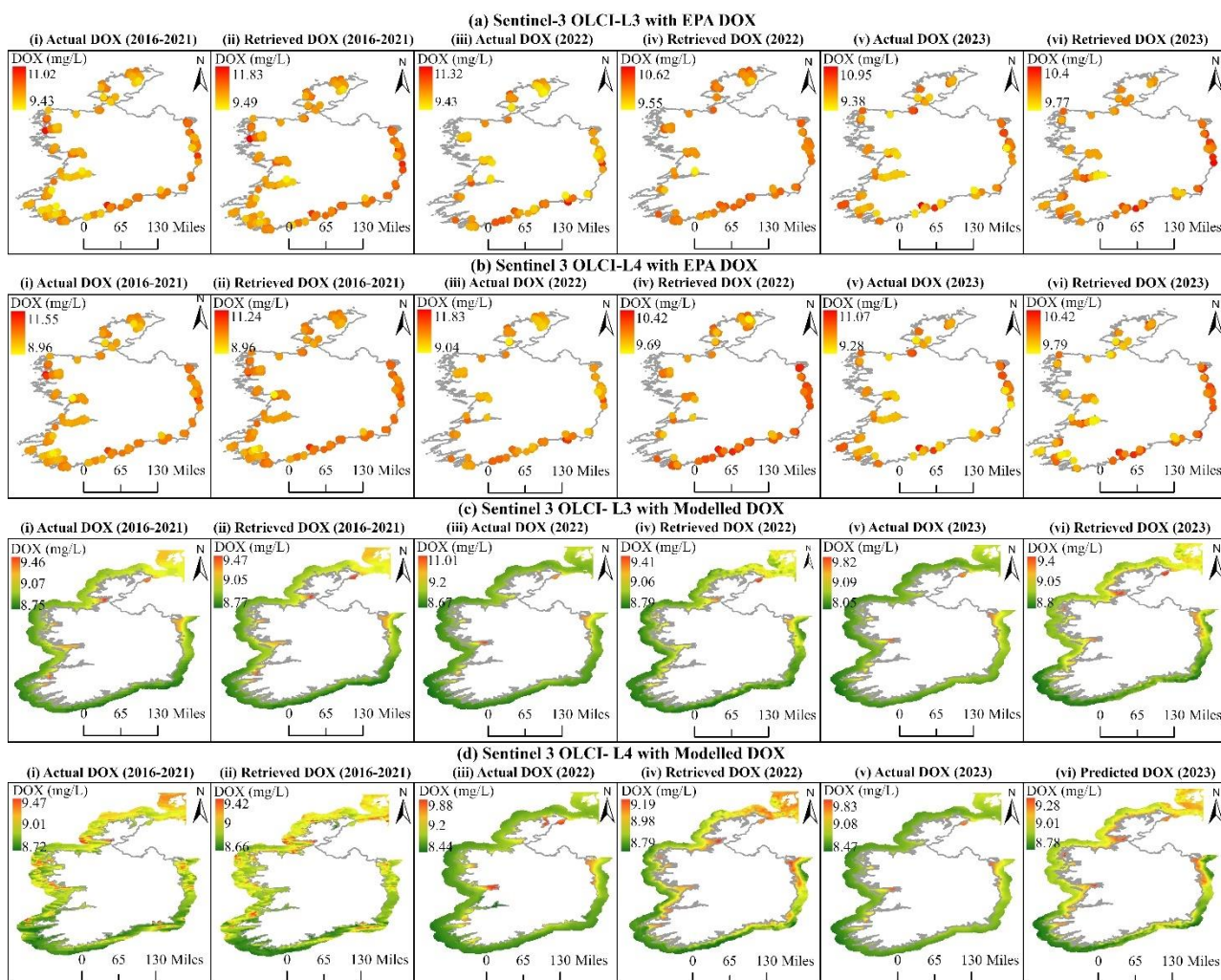
701
 702 Fig. 13. Uncertainty of various best performing models with 95 % CI at < 0.005 , where SLM, SEM, EM, VM,
 703 and SM on the x-axis refer to supervised learning model, stacking-ensembles model, equation model, voting-
 704 based ensembles model, and statistical model, respectively. (For a better interpretation of the colour references in
 705 this figure, readers are encouraged to read the web version of this article)

706 3.7 Spatio-temporal distribution of DOX concentrations in Irish TrC waters

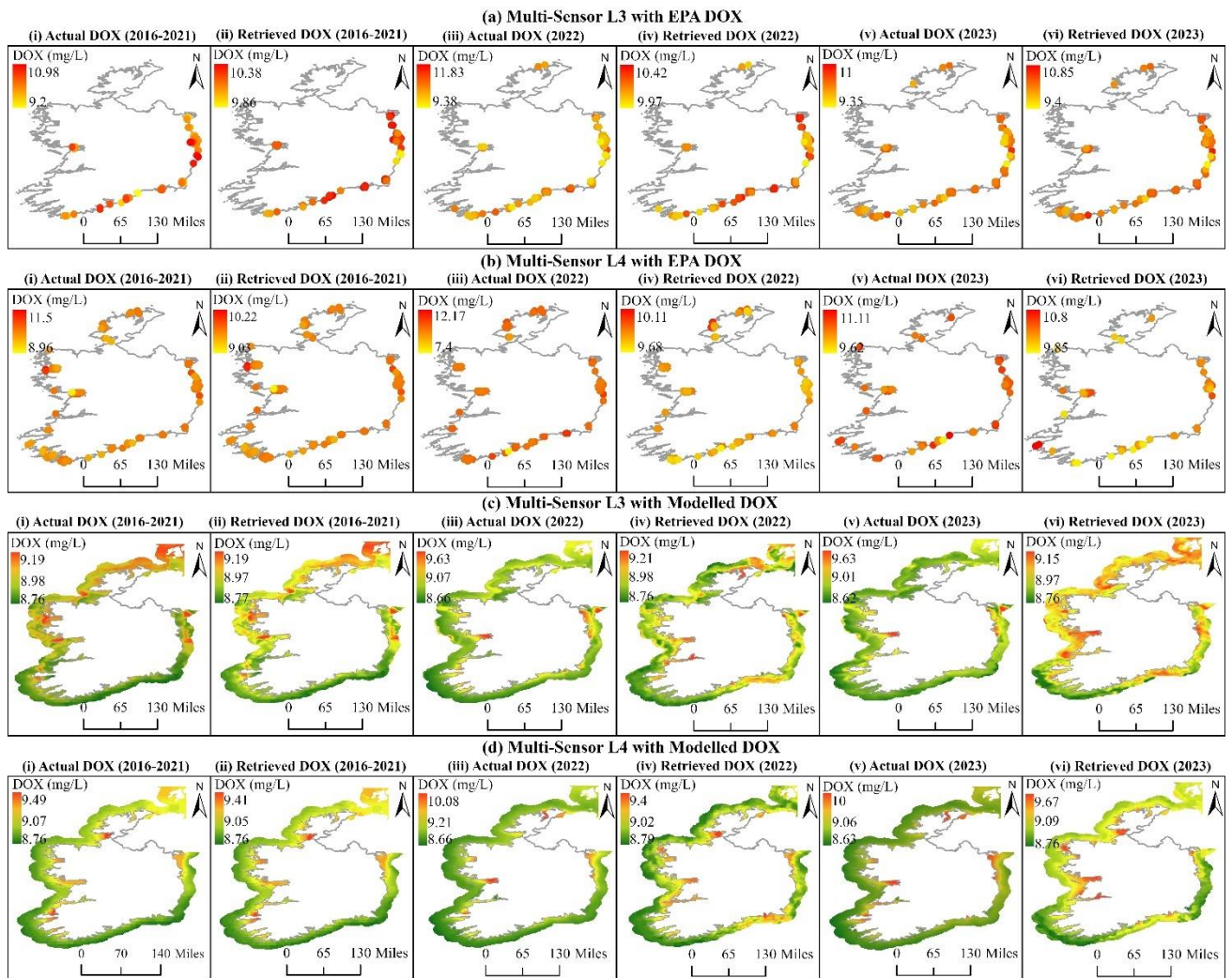
707 The supervised learning models showed higher accuracy than other models (stacking-
 708 ensembles, voting-based ensembles, equations, and statistical) in training, testing, and
 709 validation. Therefore, the present research visualized the results of supervised learning models

710 to demonstrate the spatial distribution of DOX concentrations over years in Irish waters. Fig.
711 14 and Fig. 15 show the spatial distribution of DOX level across different datasets and years.
712 It is apparent from these figures that DOX levels fluctuate with increasing temporal resolution
713 (2016-2021 to 2022), followed by a slight decrease in most datasets in 2023. This variability
714 may reflect the inter-annual changes in the tidal changes, stratification intensity, nutrient inputs,
715 and seasonal mixing dynamics (Brown et al., 2003; Fernand et al., 2006; Olbert et al., 2011).
716 Commonly, these factors highly influence the oxygen solubility and biological assimilation in
717 temperate coastal ecosystems like Ireland (Mahaffey et al., 2023). It should be mentioned that
718 consistently higher DOX concentrations were found in semi-enclosed bays such as River
719 Swilly, Donegal Bay, Clew Bay, Galway Bay, Dingle Bay, Dublin Bay, and Dundalk Bay over
720 the study periods, indicating that these waterbodies attributed to various physical factors,
721 including dynamic aeration (strong tidal mixing and wind-driven ventilation), high flushing
722 rates, short residence times, and wild Atlantic influences (Dabrowski et al., 2010; Olbert et al.,
723 2011; Sajib et al., 2024a). Accumulation of all these characteristics increased oxygen
724 replenishment and reduced thermal stratification in these zones (O'Boyle and Nolan, 2010).
725 Conversely, lower levels of DOX concentrations were found in the Oysterhaven, Spiddal, and
726 Ardmore, referring to the lower influences of these attributes on these waterbodies (Brierley &
727 Kingsford, 2009; Carey, 2023; Lønborg et al., 2021; Luo et al., 2024; Naylor et al., 2021).
728 Compared to the EPA datasets, the modelled datasets provide higher spatial resolution and
729 continuous monitoring results of DOX concentrations in offshore and inshore waterbodies.
730 However, the reliability of modelled DOX data remained questionable (Ford & Quiring, 2019;
731 Kosaka et al., 2022). In terms of retrieving DOX concentrations with the S3-OLCI and the MS
732 products, the retrieved DOX values are very close to the actual DOX concentrations for 2016-
733 2021, indicating the best-performing models performed well in DOX measurement. However,

734 with increasing years, the spatial patterns show inconsistencies, indicating poor performance
 735 of the model with validation datasets (2022 and 2023).



736
 737 Fig. 14. Spatio-temporal distribution of retrieved DOX concentrations (for S3-OLCI products and supervised
 738 learning models) across the Irish TrC waters for the years 2016–2021, 2022, and 2023. (For a better interpretation
 739 of the colour references in this figure, readers are encouraged to read the web version of this article.)



740

741 Fig. 15. Spatio-temporal distribution maps of retrieved DOX concentrations (for MS products and supervised
 742 learning models) across the Irish TrC waters for the years 2016–2021, 2022, and 2023. (For a better interpretation
 743 of the colour references in this figure, readers are encouraged to read the web version of this article.)

744 4. Discussion

745 4.1 Data and product selection

746 To date, several research have been conducted for retrieving DOX concentrations from
 747 different waterbodies (Table 2; Table S1-supplementary materials 1), but most studies have
 748 been criticised due to the reliable measurements of DOX concentrations (Cai et al., 2025; Sajib
 749 et al., 2025a). Moreover, a few studies have explored an in-depth relationship regarding how
 750 the Rhow bands affect retrieval accuracy (Dong et al., 2024; Quang et al., 2023). Furthermore,
 751 while most research on DOX concentration retrieval utilizes in-situ measurements with RS
 752 data, there are a notable absence of studies integrating modelled (NEMO-PISCES) DOX
 753 concentration data for developing RS-based retrieval models (Sajib et al., 2025a).

754 As mentioned earlier, the initial goal of the research was to assess the differences between the
755 EPA-sourced and modelled-derived DOX data and their respective relationships with various
756 Rhow bands across different RS products. The result revealed a significant difference between
757 the EPA and modelled DOX data, based on the Mann-Whitney U test ($p < 0.05$), aligning with
758 findings from previous literature (Ford and Quiring, 2019; Dekker et al., 2001). However, one
759 of the key challenges for the comparison between EPA and modelled DOX data is the limitation
760 of the literature, because there are no available studies that utilized modelled data (DOX) for
761 the development of the RS-based model(s).

762 Typically, modelled data is designed to address gaps where in-situ measurements are not
763 available (Ford & Quiring, 2019) and offer a cost-effective alternative. Despite these
764 advantages, recently the modelled data have received much more criticism over the years due
765 to data uncertainty, reliability, and its biasness (Ciavatta et al., 2025; Dekker et al., 2001; Ford
766 & Quiring, 2019; Kosaka et al., 2022; Mihailov et al., 2025; Skyllas, 2018; Skákala et al., 2024;
767 Séférian et al., 2013; Séférian et al., 2020; Vervatis et al., 2025; Zhao et al., 2024). Specifically,
768 $\approx 50\%$ underestimation was reported for the NEMO-PISCES model-derived WQ indicators,
769 such as nitrate (NO_3), phosphate (PO_4), and ammonium (NH_4) in the North Atlantic Ocean
770 (Skyllas, 2018). Similarly, Ciavatta et al. (2025) identified a substantial negative bias (-0.4
771 mg/m^3) in the simulated surface CHL concentration during blooming season. Moreover, the
772 NEMO-PISCES-derived DOX showed more complex characteristics, with both over- and
773 underestimation biases that vary with depth (Séférian et al., 2020). These criticisms raise
774 concerns about the uses of modelled data in this study and highlight the need for rigorous
775 validation with ground-measured data. On the other hand, regarding the association between
776 in-situ DOX and Rhow bands, there were no significant correlation found across various
777 datasets of EPA and modelled over the study periods (Fig. 7), consistent with previous studies
778 (Dong et al., 2024; Gao et al., 2024); although a moderate positive correlation (in the red-NIR

779 bands: Rhow_7–Rhow_11) was found between the modelled DOX and Rhow bands, indicating
 780 that relatively longer wavelengths could be associated with the modelled data attributes. In
 781 contrast relatively lower wavelengths (in the blue–green bands: Rhow_2–Rhow_6) showed
 782 moderate positive association with the EPA ground-measured data (Fig. 7).

783 In addition, to the best of the author’s knowledge, the research represented the first initiative
 784 to use the S3-OLCI data from the CMS programme for retrieving DOX, which is another
 785 strength of this research, as it is a novel contribution compared to the previous (Table S1-
 786 supplementary materials-1). For example, Chatziantoniou et al. (2022) demonstrated the
 787 potential of the MS products (L4) for DOX retrieval, but their model showed poor accuracy
 788 ($R^2 = 0.32$), and the band information utilized for model development was not clearly defined.
 789 The current research findings highlight that attention is required for developing a
 790 standardized/comprehensive approach(es) that could recommend more specific band
 791 engineering in terms of spectral resolution, temporal coverage, and environmental conditions
 792 for different space-borne sensors in retrieving DOX concentrations using state-of-the-art
 793 ML/AI technologies.

794 Table 2. A summary of previous studies based on DOX concentrations retrieving from TrC waters using RS data
 795 across globe.

Suggested algorithm	R ²	Data type	Water type	Domain	Reference	Independent data for validation (Y = Yes /N = No)
XGB	-	Landsat 9 OLI	Estuary	USA	Basirian et al., 2026	N
XGB-2D CNN	-	Sentinel 2 MSI	Coastal	China	Ali et al., 2025	N
Univariate LR	0.52	Zhuhai-1	Coastal	China	Pan et al., 2025	Y
TrAdaBoost.R2	0.81	Sentinel 2 MSI	Estuary	China	Li et al., 2025	N
MLR	0.59	Landsat 8 and 9 OLI	Coastal	China	Dong et al., 2024	N
RF	0.7	Sentinel 2 MSI	Coastal	Vietnam	Quang et al., 2023	N
XGB	-	Sentinel-2 MSI	Bay	China	Zhu et al., 2022a	N
SVR	0.32	CMS	Sea	Greece	Chatziantoniou et al., 2022	N
SMLR	0.8	MODIS Aqua, VIIRS	Sea	Korea	Kim et al., 2020	N
MLR	0.4	Landsat 7 ETM+	Bay	Turkey	Karakaya and Evrendilek, 2011	N
KNN	0.79	Sentinel 3 OLCI	Transitional	Ireland	The current study	Y
GB	0.8	Multi-Sensors	and Coastal			

796 **4.2 Model performance**

797 Another goal of this research was to develop, validate, and evaluate the performance of various
798 advanced ML/AI and statistical approach model(s) in retrieving DOX more accurately. This
799 research developed more than 2000 ML/AI and statistical models, where the results clearly
800 showed that model effectiveness is determined by methodological attributes rather than the
801 quantity and complexity of models. In this study, supervised learning models such as GB,
802 KNN, and ExDT outperformed compared to other models (stacking-ensembles, voting-based
803 ensembles, equations, and statistical-based approaches) because of their ability to handle high-
804 dimensional and non-linear relationships between data attributes of Rhow bands and DOX
805 concentrations, despite DOX being an OI-WQ indicator (Ahmadi et al., 2025; Fu et al., 2019;
806 Gafoor et al., 2022; Nezhlin et al., 2025). It can be seen from the result that, relatively, the tree-
807 based algorithms (ExDT, GB) can identify the complex relationships such as higher-order and
808 conditional dependencies between features (e.g., spectral bands) (Geurts et al., 2006;
809 Grinsztajn et al., 2022; Schiltz et al., 2018), while algorithms like KNN performed well by
810 using neighbour similarities within given dataset(s) (Halder et al., 2024). The findings are
811 consistent with previous research, which indicates the effectiveness of similar supervised
812 models for controlling the training scenarios (Chatziantoniou et al., 2022; Dong et al., 2023;
813 Gao et al., 2024; Guo et al., 2021; Krishnaraj & Honnasiddaiah, 2022; Li et al., 2020; Luo et
814 al., 2025; Quang et al., 2023; Salas et al., 2022; Shi et al., 2024; Toming et al., 2024).
815 Conversely, statistical and equation models are limited by strict functional assumptions and
816 fail to capture high-dimensional variability in RS data (Kumar et al., 2025; Johnstone et al.,
817 2009). Moreover, stacking-ensembles and voting-based ensembles models were developed to
818 improve the performance by combining multiple base regressors but failed to demonstrate good
819 performance during the validation phase (Fig. 12, and Fig. S15; Table S10, Table S12-
820 supplementary materials-1). These characteristics could be linked to the fact that the DOX is

821 OI-WQ, which means that weak and indirect spectral correlation makes ensemble aggregation
822 less effective (Raheli et al., 2024).

823 Recent advancements, such as Li et al. (2025), Ali et al. (2025), and Basirian et al. (2026), have
824 highlighted the increasing role of ML/AI, deep learning, and transfer learning using various
825 multi-sensor products for retrieving the DOX, particularly for TrC waters (Table 2). However,
826 to date, there are lack of studies that have been carried out to validate the model(s) performance
827 for completely unseen (independent) high-dimensional data attributes in order to evaluate their
828 proposed/developed model(s) in terms of the model's generalizability and transferability,
829 indicating that the existing ML/AI-based DOX retrieval approaches should be explored in
830 depth to generalize for the high-dimensional data like RS products for translating the unseen
831 dataset. In terms of the strength of this research, the proposed framework was tested and
832 validated for the unseen datasets, although the model(s) performance is not satisfactory to
833 retrieve DOX more accurately. The inconsistent performance during the validation phase raises
834 concerns about the robustness and generalizability of these models (Aldoseri et al., 2023; Kang,
835 2020; Rossi et al., 2024). The current research identified several factors that could contribute
836 to this discrepancy. These factors are overfitting/underfitting problems during the training and
837 testing period (Kang, 2020; Rossi et al., 2024) and distribution differences/data shifting
838 between the dependent and independent variables in both the training and validation datasets
839 (Cai et al., 2021; Maggio et al., 2022; Ötleş et al., 2021). In short, the results are not only
840 highlighting the need for further evaluation of factors affecting model transferability or
841 generalizability but also addressing the importance of developing more domain(s) adaptive
842 model(s) that could be capable of upholding the accuracy within different datasets.

843 **4.3 Uncertainty in features (Rhow) between model development and validation phases**

844 Several research have reported on the poor performance of models in retrieving the DOX
845 concentrations (see Table S1-supplementary materials 1). Although those studies have not

846 addressed the sources of the issues, why the model(s) obtained poor accuracy, nor have they
847 provided in detail the limitations or explanations of the models for both the training and testing
848 datasets. Therefore, it clearly noted that there were a lack of sufficient justification for why the
849 developed models failed to retrieve DOX accurately in existing studies. However, the third
850 goal of the study was to evaluate the accuracy of the model's performance with independent
851 validation datasets and to assess their sensitivity and uncertainty. The result of the research
852 reveals a notable difference between model development (2016-2021) and validation Rhow
853 bands (2022 and 2023) across all datasets, indicating high variability between the training and
854 validation datasets (Fig. S16-S23, supplementary materials-1). In particular, the PDF and CDF
855 plots show a notable change in the distribution over the years for most of the Rhow bands
856 (Rhow_1 to Rhow_11) across datasets (Fig. S16-S23, supplementary materials-1). Moreover,
857 the target variable DOX concentration (both modelled and EPA) also showed considerable
858 interannual variability within these datasets. These distributional issues reflect the poor
859 accuracy of the best-performing models with independent datasets, indicating the data shifting
860 could be an impact on model performance (Malinin et al., 2021; Ötleş et al., 2021;
861 Taghiyarrenani et al., 2023). To the best of the author's knowledge, the research represents one
862 of the first attempts to evaluate the distributional change between the training (2016-2021) and
863 independent datasets (2022-2023) within a ML/AI framework, specifically for RS-based DOX
864 retrieval from TrC waters. The notable decline in performance corresponds with the challenges
865 and limitations of the dataset shift in model generalization documented across various ML/AI
866 fields (Cai et al., 2021; Malinin et al., 2021; Ötleş et al., 2021; Dockès, 2021; Subbaswamy et
867 al., 2020; Taghiyarrenani et al., 2023; Turhan, 2012). These findings demonstrate that
868 distribution shifts in data not only reduced model accuracy but also increased error rates, biased
869 predictions, and created instability and overfitting/underfitting problems in this study.

870 **4.4 Implications for anthropogenic water cycle**

871 In this research, the Irish TrC waters exhibit natural DOX dynamics, which are primarily
872 influenced by seasonal stratification, tidal mixing, and primary production alongside
873 overlaying anthropogenic effects. This is evident by the high DOX levels (EPA matchups ~10–
874 11 mg/L) found in this study. During the study period, higher DOX levels were consistently
875 observed in semi-enclosed bays, whereas lower DOX levels were observed in less dynamic
876 areas (see Fig. 14 and Fig. 15). A recent EPA study found that excess nitrogen (mainly from
877 agricultural and wastewater runoff) causes notable variations in DOX levels in these regions
878 (Feeley et al., 2025; Trodd et al., 2021), leading to increased eutrophication. These pressures
879 correspond with larger European trends, where runoff from farming is a major threat to
880 eutrophication and deoxygenation, especially in TrC regions (European Environment Agency,
881 2021).

882 The research introduces a scalable and robust methodology for mapping baseline oxygen
883 conditions by utilizing RS and ML/AI techniques, which could provide critical synoptic
884 insights into the expansive TrC environment. The methodology also provides a valuable pre-
885 anthropogenic or natural benchmark that could accurately capture temporal and spatial DOX
886 variability at a low cost and minimal direct human intervention. Both the OSPAR Convention
887 commitments (Convention for the Protection of the Marine Environment of the North-East
888 Atlantic) and the Irish EPA assessment group could utilize this benchmark to detect future
889 changes in baseline DOX conditions caused by excessive nutrient inputs.

890 Moreover, the transferable nature of this framework could be crucial for addressing the
891 complexities of the anthropogenic water cycle, particularly by utilizing this framework to
892 different regions and contexts. Additionally, the independent validation data approach can be
893 further extended to scenario-based simulations. For example, it could be utilized to predict how

894 DOX levels would improve if agricultural runoff were reduced to meet the EU Nitrates
895 Directive standards (Musacchio et al., 2020). These types of transferable and predictive
896 capabilities could assist in the adaptive management plans required under the WFD and Marine
897 Strategy Framework Directive. In addition, this robust approach can be tested/adapted into
898 various global scenarios to combat ocean deoxygenation.

899 **4.5 Challenges, limitations and future research directions**

900 The research highlights the robust nature of integrating cutting-edge ML/AI with satellite data
901 for retrieving (predicting) OI-WQ indicators. However, as indicated in the author's prior
902 research (Diganta et al., 2024, 2025; Sajib et al., 2025a; Uddin et al., 2025a), a number of
903 persistent challenges hinder reliable and effective implementation of retrieving OI-WQ
904 indicators using these techniques. A primary issue is the indirect relationship between spectral
905 reflectance and OI-WQ indicators, which are further affected by optical complexity in TrC
906 waters (e.g., atmospheric disturbances and high variability in sediments) (Morel and Prieur,
907 1977; Gordon and Morel, 1983; Vanhellefont & Ruddick, 2021; Sajib et al., 2025a, 2025b).
908 Apart from this, high variability (e.g., due to algal bloom, eutrophication) and inherent noise
909 (e.g., aerosols, clouds, precipitation, fog, hue) in RS data increase uncertainty in the retrieval
910 models (Brando et al., 2024; Coffey et al., 2025; IOCCG, 2019; Maciel et al., 2023). All these
911 characteristics collectively may limit the physical interpretability and generalizability of
912 ML/AI models, specifically when validate model with independent datasets. These limitations
913 are highlighted in this study during the validation phase, indicating sensitivity to data
914 distribution shifts. Additionally, the study identified that standalone single R_{rs} bands are not
915 sufficient for developing OI-WQ indicators, as these bands are highly correlated with each
916 other, indicating multicollinearity issues (Fig. S24-S25, supplementary materials-1).

917 Another important consideration is the data quality issue, such as small datasets (e.g., $N = 289$
918 EPA matchups) and sparse feature characteristics of EPA-sourced DOX data (due to logistical

919 constraints). These limitations notably increase the risk of overfitting/underfitting, especially
920 when attempting to capture relationships between DOX and high-dimensional RS reflectance.
921 Furthermore, without integrating L2 RS data and seasonal variability analysis, the developed
922 model remains limited in its capacity to capture the short-term dynamics of DOX in the study
923 area.

924 In order to improve the model(s) performance, the research addressed a few strategies that
925 could be integrated into future research. First, exploring various spectral combinations, band
926 ratios, advanced feature engineering (e.g., Fast Fourier transformation, Discrete Wavelet
927 transformation, various ML/AI approaches like SHAP, feature engineering-based algorithms),
928 data augmentation (e.g., Gaussian noise to the predictor variable), and integration of L2 RS
929 data with atmospheric corrections to capture short- and long-term dynamics. Expand to multi-
930 source fusion (combination of S3-OLCI, MS, and hyperspectral data) and larger datasets from
931 various geographical domains to boost the generalization of the model. Second, hybrid
932 modelling approaches with physics-informed constraints (e.g., Physics-Informed Neural
933 Networks) or transfer learning to reduce overfitting and improve transferability across various
934 domains and time periods. Finally, model evaluation in various aquatic environments (e.g.,
935 lakes, rivers, seas, and wetlands) will be essential for assessing the model's performance,
936 uncertainty, and reliability. However, to the best of the author's knowledge, this is the first
937 effort to comprehensive analysis and assess 2101 ML/AI models, including independent
938 dataset(s) used for further validation of the model(s) performance in retrieving the DOX from
939 RS data. The results and findings of the research could be effective in selecting the ML/AI
940 model(s) and their architectural attributes with input customization for further improvement of
941 the framework.

942 **5. Conclusion**

943 The research was carried out to address the reliability of the existing relevant studies for
944 retrieving DOX concentrations using ML/AI techniques from the RS products. The main goal
945 of the research was to evaluate the reliability of various advanced ML/AI (supervised learning,
946 stacking-ensembles, equations, and voting-based ensemble models) and statistical models
947 (PCR, PLS, and OLS models) for retrieving DOX concentrations from in-situ (EPA, Ireland)
948 and modelled (NEMO-PISCES) using S3-OLCI and MS RS products. For achieving the
949 research goals, the study developed 2101 ML/AI with the integration of the advanced statistical
950 tools and techniques, whereas the model performance was evaluated using independent
951 dataset(s) by comparing multiple error metrics. From the results and findings, the summary of
952 the research are as follows:

- 953 • The research found a significant difference between the EPA and modelled DOX
954 concentrations over the study period, based on the Mann-Whitney U test ($p < 0.05$).
955 Also, the research identified the spatial clustering pattern in both matchups data
956 using Moran's I statistics.
- 957 • Among the five groups of models (>2000, including stacking-ensemble, voting-
958 based ensemble, statistical, and equation), the supervised learning models (KNN,
959 BGR, GB, and ExDT) showed marginal performance, which is consistent over the
960 years for all datasets; whereas the best-performing models (ranked 1) showed poor
961 generalizability on independent datasets.
- 962 • The best-performing models showed high sensitivity and low uncertainty during the
963 training period but failed to uphold consistent performance with the independent
964 datasets, linked to distributional shifts between datasets.
- 965 • In terms of the RS products, the S3-OLCI was identified as the most reliable RS
966 product compared to MS data for retrieving DOX using L3 and L4 data among the

967 tested products with less uncertainty and stable performance across datasets.
968 Relatively, using L4 data, the model(s) shows better performance compared to L3.
969 • In retrieving DOX, relatively higher variability in DOX concentrations was found
970 in bay areas such as River Swilly, Donegal Bay, Clew Bay, Galway Bay, Dingle
971 Bay, Dublin Bay, and Dundalk Bay, whereas lower DOX concentrations were
972 found in Oysterhaven, Spiddal, and Ardmore.

973 It should be highlighted that the novel contributions of this research include: (i) this was the
974 first initiative using the CMS S3-OLCI products for retrieving the DOX integrating the ML/AI
975 with advanced statistical approaches, and demonstrating their edge in accuracy for OI-WQ
976 indicator measurement; (ii) the research developed and validated 2101 models for DOX
977 retrieval, addressing the inherent limitations of advanced data-driven approaches in estimating
978 OI indicators using only RS spectral bands standalone; and (iii) an in-depth investigation of the
979 reasons behind the failure of the best-performing models at the validation stage via distribution
980 shifts, highlighting challenges in RS-based OI indicator estimation. Therefore, the results of
981 this research could be utilized for various operational monitoring, such as the EPA can utilize
982 S3-OLCI data for large-scale DOX assessment under the EU-WFD framework while
983 incorporating routine validation to reduce biases in the results. Moreover, the results of this
984 research have several important implications for future practice, especially for
985 researchers/organizations/scientist could utilize the present research outcomes for further
986 advancement of retrieving various OI-WQ indicators from the RS products using the advanced
987 data-driven approaches. Although the research has been carried out using the fundamental
988 ML/AI algorithms using stacking-ensembles, voting-based ensembles, statistical, and equation
989 approaches, future studies should incorporate deep learning and more sophisticated algorithms
990 like transformer-based time–space models, physics-informed and hybrid AI, etc., to improve
991 the retrieval accuracy of the model(s). However, the research reveals that the supervised

992 learning model(s) could be effective for retrieving the DOX within the coastal zone using the
993 S3-OLCI (L4) across highly cloud-coverage areas like Irish waterbodies.

994 **Acknowledgement**

995 The authors gratefully acknowledge the editor's and anonymous reviewers for their valuable
996 feedback. This research was funded by the Hardiman Research Scholarship of the University
997 of Galway, which funded the first author as part of his PhD program. The authors sincerely
998 acknowledge the Eco-HydroInformatics Research Group (EHIRG), School of Engineering,
999 College of Science and Engineering, University of Galway, Ireland for providing laboratory
1000 facilities to complete this research. The research was also supported by the Environmental
1001 Protection Agency, Ireland for the AquaCop project [Grant Ref No: 2022-NE-1128].

1002 **Supplementary materials**

1003 Supplementary information associated with this article can be found in supplementary
1004 materials 1-6.

1005 **References**

- 1006 Abbas, F., Zhang, F., Hussain, M. A., Abbas, H., Alrefaei, A. F., Albeshr, M. F., Iqbal, J., Ghani, J., & Ismail
1007 shah. (2024). Landslide susceptibility assessment along the Karakoram highway, Gilgit Baltistan,
1008 Pakistan: A comparative study between ensemble and neighbor-based machine learning algorithms.
1009 Science of Remote Sensing, 9(March 2023), 100132. <https://doi.org/10.1016/j.srs.2024.100132>
- 1010 Ahmadi, B., Gholamalifard, M., Ghasempouri, S. M., & Kutser, T. (2025). Comparative assessment of machine
1011 learning algorithms for retrieving colored dissolved organic matter (CDOM) from Sentinel-2/MSI
1012 images in the coastal waters of the Persian Gulf. Ecological Informatics, 89(January), 103171.
1013 <https://doi.org/10.1016/j.ecoinf.2025.103171>
- 1014 Alarcon Falconi, T. M., Estrella, B., Sempéregui, F., & Naumova, E. N. (2020). Effects of data aggregation on
1015 time series analysis of seasonal infections. International journal of environmental research and public
1016 health, 17(16), 5887. <https://doi.org/10.3390/ijerph17165887>
- 1017 Aldoseri, A., Al-Khalifa, K. N., & Hamouda, A. M. (2023). Re-Thinking Data Strategy and Integration for
1018 Artificial Intelligence: Concepts, Opportunities, and Challenges. Applied Sciences (Switzerland), 13(12).
1019 <https://doi.org/10.3390/app13127082>
- 1020 Alghushairy, O., Alsini, R., Soule, T., & Ma, X. (2021). A review of local outlier factor algorithms for outlier
1021 detection in big data streams. Big Data and Cognitive Computing, 5(1), 1–24.
1022 <https://doi.org/10.3390/bdcc5010001>
- 1023 Ali, A., Zhou, G., Jing, G., Lopez, F. P. A., Sun, K., Jiang, C., ... & Tan, Y. (2025). XGB-2DCNN based
1024 Dissolved Oxygen Inversion for Coastal Water. IEEE Journal of Selected Topics in Applied Earth
1025 Observations and Remote Sensing. <https://doi.org/10.1109/JSTARS.2025.3598875>
- 1026 Arias-Rodriguez, L. F., Tüzün, U. F., Duan, Z., Huang, J., Tuo, Y., & Disse, M. (2023). Global Water Quality
1027 of Inland Waters with Harmonized Landsat-8 and Sentinel-2 Using Cloud-Computed Machine Learning.
1028 Remote Sensing, 15(5). <https://doi.org/10.3390/rs15051390>
- 1029 Aslan, S., Zennaro, F., Furlan, E., & Critto, A. (2022). Recurrent neural networks for water quality assessment
1030 in complex coastal lagoon environments: A case study on the Venice Lagoon. Environmental Modelling
1031 and Software, 154(April), 105403. <https://doi.org/10.1016/j.envsoft.2022.105403>

- 1032 Bailey, S. W., & Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color
1033 satellite data products. *Remote sensing of environment*, 102(1-2), 12-23.
1034 <https://doi.org/10.1016/j.rse.2006.01.015>
- 1035 Barde, M. P., & Barde, P. J. (2012). What to use to express the variability of data: Standard deviation or
1036 standard error of mean?. *Perspectives in clinical research*, 3(3), 113-116. <https://doi.org/10.4103/2229-3485.100662>
1037
- 1038 Basirian, S., Najafzadeh, M., & Demir, I. (2026). Water quality monitoring for coastal hypoxia: Integration of
1039 satellite imagery and machine learning models. *Marine pollution bulletin*, 222, 118735.
1040 <https://doi.org/10.1016/j.marpolbul.2025.118735>
- 1041 Batur, E., & Maktav, D. (2019). Assessment of Surface Water Quality by Using Satellite Images Fusion Based
1042 on PCA Method in the Lake Gala, Turkey. *IEEE Transactions on Geoscience and Remote Sensing*,
1043 57(5), 2983–2989. <https://doi.org/10.1109/TGRS.2018.2879024>
- 1044 Bella, S., Fidler, F., Williams, J., & Cumming, G. (2005). Researchers misunderstand confidence intervals and
1045 standard error bars. *Psychological Methods*, 10(4), 389–396. <https://doi.org/10.1037/1082-989X.10.4.389>
- 1046 Boyd, C. E. (2015). Water quality: An introduction. In *Water Quality: An Introduction* (2nd ed.). Springer
1047 Cham. <https://doi.org/10.1007/978-3-319-17446-4>
- 1048 Brando, V. E., Santoleri, R., Colella, S., Volpe, G., Di Cicco, A., Sammartino, M., González Vilas, L., Lapucci,
1049 C., Böhm, E., Zoffoli, M. L., Cesarini, C., Forneris, V., La Padula, F., Mangin, A., Jutard, Q., Bretagnon,
1050 M., Bryère, P., Demaria, J., Calton, B., ... Lebreton, C. (2024). Overview of Operational Global and
1051 Regional Ocean Colour Essential Ocean Variables Within the Copernicus Marine Service. *Remote
1052 Sensing*, 16(23), 1–21. <https://doi.org/10.3390/rs16234588>
- 1053 Brierley, A. S., & Kingsford, M. J. (2009). Impacts of Climate Change on Marine Organisms and Ecosystems.
1054 *Current Biology*, 19(14), R602–R614. <https://doi.org/10.1016/j.cub.2009.05.046>
- 1055 Brown, J., Carrillo, L., Fernand, L., Horsburgh, K. J., Hill, A. E., Young, E. F., & Medler, K. J. (2003).
1056 Observations of the physical structure and seasonal jet-like circulation of the Celtic Sea and St. George's
1057 Channel of the Irish Sea. *Continental Shelf Research*, 23(6), 533-561. [https://doi.org/10.1016/S0278-4343\(03\)00008-6](https://doi.org/10.1016/S0278-4343(03)00008-6)
1058
- 1059 Cai, C., Liu, L., Wang, Z., Pang, W., Bai, C., & Zhang, H. (2025). Retrieval of non-optical active water quality
1060 parameters in complex Lake environments using a novel zoning-based ensemble modeling strategy.
1061 *Ecological Indicators*, 176(April), 113723. <https://doi.org/10.1016/j.ecolind.2025.113723>
- 1062 Cai, F., Ozdagli, A. I., & Koutsoukos, X. (2021). Detection of dataset shifts in learning-enabled cyber-physical
1063 systems using variational autoencoder for regression. *Proceedings - 2021 4th IEEE International
1064 Conference on Industrial Cyber-Physical Systems, ICPS 2021*, 104–111.
1065 <https://doi.org/10.1109/ICPS49255.2021.9468230>
- 1066 Cai, X., Li, Y., Lei, S., Zeng, S., Zhao, Z., Lyu, H., Dong, X., Li, J., Wang, H., Xu, J., Zhu, Y., Wu, L., &
1067 Cheng, X. (2023). A hybrid remote sensing approach for estimating chemical oxygen demand
1068 concentration in optically complex waters: A case study in inland lake waters in eastern China. *Science
1069 of the Total Environment*, 856(September 2022), 158869.
1070 <https://doi.org/10.1016/j.scitotenv.2022.158869>
- 1071 Carey, C. C. (2023). Causes and consequences of changing oxygen availability in lakes: Kilham Plenary Lecture
1072 Article. *Inland Waters*, 13(3), 316–326. <https://doi.org/10.1080/20442041.2023.2239110>
- 1073 Chatziantoniou, A., Charalampis Spondylidis, S., Stavrakidis-Zachou, O., Papandroulakis, N., & Topouzelis, K.
1074 (2022). Dissolved oxygen estimation in aquaculture sites using remote sensing and machine learning.
1075 *Remote Sensing Applications: Society and Environment*, 28(October), 100865.
1076 <https://doi.org/10.1016/j.rsase.2022.100865>
- 1077 Chen, J., Chen, S., Fu, R., Li, D., Jiang, H., Wang, C., Peng, Y., Jia, K., & Hicks, B. J. (2022). Remote Sensing
1078 Big Data for Water Environment Monitoring: Current Status, Challenges, and Future Prospects. *Earth's
1079 Future*, 10(2), 1–33. <https://doi.org/10.1029/2021EF002289>

- 1080 Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more
1081 informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ*
1082 *Computer Science*, 7, 1–24. <https://doi.org/10.7717/PEERJ-CS.623>
- 1083 Ciavatta, S., Lazzari, P., Álvarez, E., Bertino, L., Bolding, K., Bruggeman, J., Capet, A., Cossarini, G.,
1084 Daryabor, F., Nerger, L. and Popov, M., (2025). Control of simulated ocean ecosystem indicators by
1085 biogeochemical observations. *Progress in Oceanography*, 231, 103384
- 1086 Coffey, M. M., Schaeffer, B. A., Salls, W. B., Minucci, J. M., & Cronin-Golomb, O. (2025). Recommendations
1087 for temporal aggregation of water quality data from multi-platform satellite constellations. *International*
1088 *Journal of Remote Sensing*, 1-23. <https://doi.org/10.1080/01431161.2025.2575515>
- 1089 Colella, S., Böhm, E., Cesarini, C., Garnesson, P., Netting, J., & Calton, B. (2023). Product User Manual For
1090 Ocean Colour Products, Ref: CMEMS-OC-PUM-4.0. MERCATOR OCEAN INTERNATIONAL, 1–27.
- 1091 Creane, S., O'Shea, M., Coughlan, M., & Murphy, J. (2021). The Irish Sea bed load parting zone: Is it a mid-sea
1092 hydrodynamic phenomenon or a geological theoretical concept?. *Estuarine, Coastal and Shelf Science*,
1093 263, 107651. <https://doi.org/10.1016/j.ecss.2021.107651>
- 1094 Cumming, G., Fidler, F., & Vaux, D. L. (2007). Error bars in experimental biology. *Journal of Cell Biology*,
1095 177(1), 7–11. <https://doi.org/10.1083/jcb.200611141>
- 1096 Dabrowski, T., Hartnett, M., & Olbert, A. I. (2010). Influence of seasonal circulation on flushing of the Irish
1097 Sea. *Marine pollution bulletin*, 60(5), 748-758. <http://dx.doi.org/10.1016/j.marpolbul.2009.11.018>
- 1098 Dai, M., Zhao Y., Chai, F., Chen, M., Chen, N., Chen Y., Cheng, D., Gan, J., Guan, D., Hong, Y., Huang, J.,
1099 Lee, Y., Leung, K. M. Y., Lim, P. E., Lin, S., Lin, X., Liu, X., Liu, Z., Luo, Y-W., Meng, F., Sangmanee,
1100 C., Shen, Y, Uthaiapan, K., Wan Talaat, W. I. A., Wan, X. S., Wang, C., Wang D., Wang G., Wang S.,
1101 Wang Y., Wang, Y., Wang, Z., Wang, Z., Xu, Y., Yang, J-YT., Yang, Y., Yasuhara, M., Yu, D., Yu, J.,
1102 Yu, L., Zhang, Z., and Zhang, Z. (2023). Persistent eutrophication and hypoxia in the coastal ocean.
1103 *Cambridge Prisms: Coastal Futures*, 1, e19, 1–28. <https://doi.org/10.1017/cft.2023.7>
- 1104 Dekker, A. G., Vos, R. J., & Peters, S. W. M. (2001). Comparison of remote sensing data, model results and in
1105 situ data for total suspended matter (TSM) in the southern Frisian lakes. *Science of the Total*
1106 *Environment*, 268(1–3), 197–214. [https://doi.org/10.1016/S0048-9697\(00\)00679-3](https://doi.org/10.1016/S0048-9697(00)00679-3)
- 1107 Delaney, C., Mccarthy, V., French, K., Karki, S., Veerkamp, V., Ahmed, M., Ghaffar, A., Hanafin, J.,
1108 Mckinstry, A., Jennings, E., & Golden, A. (2023). Remote Sensing of Irish Surface Waters (Issue 443).
1109 https://www.epa.ie/publications/research/water/Final-CORRECT-Research-Report_443.pdf
- 1110 Diganta, M. T. M., Uddin, M. G., Rahman, A., & Olbert, A. I. (2024). A comprehensive review of various
1111 environmental factors' roles in remote sensing techniques for assessing surface water quality. *Science of*
1112 *the Total Environment*, 957(July). <https://doi.org/10.1016/j.scitotenv.2024.177180>
- 1113 Diganta, M.T.M., Uddin, M.G., Olbert, A.I., 2025. Optimization of Chlorophyll-a retrieval algorithm(s) for
1114 assessing tropic state in marine waters. (Submitted to publication).
- 1115 Dong, L., Gong, C., Huai, H., Wu, E., Lu, Z., Hu, Y., Li, L., & Yang, Z. (2023). Retrieval of Water Quality
1116 Parameters in Dianshan Lake Based on Sentinel-2 MSI Imagery and Machine Learning: Algorithm
1117 Evaluation and Spatiotemporal Change Research. *Remote Sensing*, 15(20).
1118 <https://doi.org/10.3390/rs15205001>
- 1119 Dong, L., Wang, D., Song, L., Gong, F., Chen, S., Huang, J., & He, X. (2024). Monitoring Dissolved Oxygen
1120 Concentrations in the Coastal Waters of Zhejiang Using Landsat-8/9 Imagery. *Remote Sensing*, 16(11),
1121 1–23. <https://doi.org/10.3390/rs16111951>
- 1122 Dockès, J., Varoquaux, G., & Poline, J. B. (2021). Preventing dataset shift from breaking machine-learning
1123 biomarkers. *GigaScience*, 10(9), giab055. <https://doi.org/10.1093/gigascience/giab055>
- 1124 EPA. (2001). PARAMETERS OF WATER QUALITY Interpretation and Standards. 2001, 5(3), 248–253.
- 1125 EPA. (2021a). Ireland's National Water Framework Directive Monitoring Programme 2019–2021.
1126 https://www.epa.ie/pubs/reports/water/waterqua/EPA_WFD_MonitoringProgramme_2019_2021.pdf

- 1127 EPA. (2021b). Water quality monitoring report on nitrogen and phosphorus concentrations in Irish waters 2020.
 1128 [https://www.epa.ie/publications/monitoring--assessment/freshwater--](https://www.epa.ie/publications/monitoring--assessment/freshwater--marine/EPA_NITrogenandPhosporous_Concentrations_2020.pdf)
 1129 [marine/EPA_NITrogenandPhosporous_Concentrations_2020.pdf](https://www.epa.ie/publications/monitoring--assessment/freshwater--marine/EPA_NITrogenandPhosporous_Concentrations_2020.pdf)
- 1130 EPA. (2023a). Ireland's National Water Quality Monitoring Programme 2022–2027. In Environmental
 1131 Protection Agency Ireland. [http://www.ctic.org/media/Detailed NWQMC Agenda.pdf](http://www.ctic.org/media/Detailed_NWQMC_Agenda.pdf)
- 1132 EPA. (2023b). Water Quality in 2022. In EPA Research Report. [https://www.epa.ie/publications/monitoring--](https://www.epa.ie/publications/monitoring--assessment/freshwater--marine/water-quality-in-2020.php)
 1133 [assessment/freshwater--marine/water-quality-in-2020.php](https://www.epa.ie/publications/monitoring--assessment/freshwater--marine/water-quality-in-2020.php)
- 1134 EU. (2019). EUROPEAN UNION ENVIRONMENTAL OBJECTIVES (SURFACE WATERS)
 1135 (AMENDMENT) REGULATIONS 2019. European Union, 3(1), 18–23.
- 1136 EUMETSAT. (2021). Recommendations for Sentinel-3 OLCI Ocean Colour product validations in comparison
 1137 with in situ measurements – Matchup Protocols. EUM/SEN3/D. Matchup Protocols, EUM/SEN3/D(v7),
 1138 1–10. <http://www.eumetsat.int>
- 1139 European Environment Agency (2021). Open ocean — ocean chemistry: dissolved oxygen and ocean acidity.
 1140 Accessed on 14 March 2025. From [https://www.eea.europa.eu/publications/europes-changing-climate-](https://www.eea.europa.eu/publications/europes-changing-climate-hazards-1/open-ocean/open-ocean-ocean-chemistry#:~:text=The%20chemical%20properties%20of%20seawater,makes%20the%20seawater%20more%20acidic)
 1141 [hazards-1/open-ocean/open-ocean-ocean-](https://www.eea.europa.eu/publications/europes-changing-climate-hazards-1/open-ocean/open-ocean-ocean-chemistry#:~:text=The%20chemical%20properties%20of%20seawater,makes%20the%20seawater%20more%20acidic)
 1142 [chemistry#:~:text=The%20chemical%20properties%20of%20seawater,makes%20the%20seawater%20m](https://www.eea.europa.eu/publications/europes-changing-climate-hazards-1/open-ocean/open-ocean-ocean-chemistry#:~:text=The%20chemical%20properties%20of%20seawater,makes%20the%20seawater%20more%20acidic)
 1143 [ore%20acidic](https://www.eea.europa.eu/publications/europes-changing-climate-hazards-1/open-ocean/open-ocean-ocean-chemistry#:~:text=The%20chemical%20properties%20of%20seawater,makes%20the%20seawater%20more%20acidic).
- 1144 Fan, J., Li, M., Guo, F., Yan, Z., Zheng, X., Zhang, Y., Xu, Z., & Wu, F. (2018). Prioritization of river restoration
 1145 by coupling soil and water assessment tool (SWAT) and support vector machine (SVM) models in the
 1146 Taizi river basin, northern China. *International Journal of Environmental Research and Public Health*,
 1147 15(10). <https://doi.org/10.3390/ijerph15102090>
- 1148 Fang, C., Song, K., Yan, Z., & Liu, G. (2025). Monitoring phycocyanin in global inland waters by remote
 1149 sensing: Progress and future developments. *Water Research*, 123176.
 1150 <https://doi.org/10.1016/j.watres.2025.123176>
- 1151 Feeley, H., McConigley, C. and Deakin, J. (2025). Water Quality in Ireland 2019-2024. ENVIRONMENTAL
 1152 PROTECTION AGENCY. PO Box 3000, Johnstown Castle, Co. Wexford, Ireland
- 1153 Fernand, L., Nolan, G. D., Raine, R., Chambers, C. E., Dye, S. R., White, M., & Brown, J. (2006). The Irish
 1154 coastal current: A seasonal jet-like circulation. *Continental Shelf Research*, 26(15), 1775-1793.
 1155 <https://doi.org/10.1016/j.csr.2006.05.010>.
- 1156 Ford, T. W., & Quiring, S. M. (2019). Comparison of Contemporary In Situ, Model, and Satellite Remote
 1157 Sensing Soil Moisture With a Focus on Drought Monitoring. *Water Resources Research*, 55(2), 1565–
 1158 1582. <https://doi.org/10.1029/2018WR024039>
- 1159 Forkel, M., Carvalhais, N., Verbesselt, J., Mahecha, M. D., Neigh, C. S., & Reichstein, M. (2013). Trend change
 1160 detection in NDVI time series: Effects of inter-annual variability and methodology. *Remote sensing*,
 1161 5(5), 2113-2144. <https://doi.org/10.3390/rs0502113>
- 1162 Frouin, R. J., Franz, B. A., Ibrahim, A., Knobelspiesse, K., Ahmad, Z., Cairns, B., Chowdhary, J., Dierssen, H.
 1163 M., Tan, J., Dubovik, O., Huang, X., Davis, A. B., Kalashnikova, O., Thompson, D. R., Remer, L. A.,
 1164 Boss, E., Coddington, O., Deschamps, P. Y., Gao, B. C., ... Zhai, P. W. (2019). Atmospheric Correction
 1165 of Satellite Ocean-Color Imagery During the PACE Era. *Frontiers in Earth Science*, 7(July), 1–43.
 1166 <https://doi.org/10.3389/feart.2019.00145>
- 1167 Fu, Y., He, H. S., Hawbaker, T. J., Henne, P. D., Zhu, Z., & Larsen, D. R. (2019). Evaluating k-nearest neighbor
 1168 (kNN) imputation models for species-level aboveground forest biomass mapping in Northeast China.
 1169 *Remote Sensing*, 11(17). <https://doi.org/10.3390/rs11172005>
- 1170 Gafoor, F. A., Al-Shehhi, M. R., Cho, C. S., & Ghedira, H. (2022). Gradient Boosting and Linear Regression for
 1171 Estimating Coastal Bathymetry Based on Sentinel-2 Images. *Remote Sensing*, 14(19).
 1172 <https://doi.org/10.3390/rs14195037>
- 1173 Gani, M. A., Sajib, A. M., Siddik, M. A., & Moniruzzaman, M. (2023). Assessing the impact of land use and
 1174 land cover on river water quality using water quality index and remote sensing techniques.
 1175 *Environmental Monitoring and Assessment*, 195(4). <https://doi.org/10.1007/s10661-023-10989-1>

- 1176 Gao, L., Shangguan, Y., Sun, Z., Shen, Q., & Shi, Z. (2024). Estimation of Non-Optically Active Water Quality
1177 Parameters in Zhejiang Province Based on Machine Learning. *Remote Sensing*, 16(3), 1–19.
1178 <https://doi.org/10.3390/rs16030514>
- 1179 Garnesson, P., Mangin, A., Bretagnon, M., & Jutard, Q. (2024). Satellite Observation Copernicus-GlobColour
1180 Products, Ref: CMEMS-OC-QUID-009-101to104-111-113- 116-118. MERCATOR OCEAN
1181 INTERNATIONAL, 1–93. [https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-OC-
1182 QUID-009-030-032-033-037-081-082-083-085-086-098.pdf](https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-OC-QUID-009-030-032-033-037-081-082-083-085-086-098.pdf)
- 1183 Garnesson, P., Mangin, A., D’Andon, O. F., Demaria, J., & Bretagnon, M. (2019). The CMEMS GlobColour
1184 chlorophyll a product based on satellite observation: Multi-sensor merging and flagging strategies. *Ocean
1185 Science*, 15(3), 819–830. <https://doi.org/10.5194/os-15-819-2019>
- 1186 Geurts, P., Ernst, D. & Wehenkel, L. (2006). Extremely randomized trees. *Mach Learn* 63, 3–42.
1187 <https://doi.org/10.1007/s10994-006-6226-1>
- 1188 Gleratti, G., Martinez-Vicente, V., Atwood, E. C., Simis, S. G. H., & Jackson, T. (2024). Validation of full
1189 resolution remote sensing reflectance from Sentinel-3 OLCI across optical gradients in moderately turbid
1190 transitional waters. *Frontiers in Remote Sensing*, 5(July), 1–18.
1191 <https://doi.org/10.3389/frsen.2024.1359709>
- 1192 Gordon, H. R. (2021). Evolution of ocean color atmospheric correction: 1970–2005. *Remote Sensing*, 13(24),
1193 1970–2005. <https://doi.org/10.3390/rs13245051>
- 1194 Gordon, H.R. and Morel, A.Y. (1983) *Remote Assessment of Ocean Color for Interpretation of Satellite Visible
1195 Imagery: A Review*. Springer-Verlag, New York.
1196 <https://doi.org/10.1029/LN004>
- 1197 Griffith, D. A., & Chun, Y. (2016). Spatial autocorrelation and uncertainty associated with remotely-sensed
1198 data. *Remote Sensing*, 8(7), 535. <https://doi.org/10.3390/rs8070535>
- 1199 Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning
1200 on typical tabular data?. *Advances in neural information processing systems*, 35, 507-520.
- 1201 Guo, H., Huang, J. J., Zhu, X., Wang, B., Tian, S., Xu, W., & Mai, Y. (2021). A generalized machine learning
1202 approach for dissolved oxygen estimation at multiple spatiotemporal scales using remote sensing.
1203 *Environmental Pollution*, 288(March), 117734. <https://doi.org/10.1016/j.envpol.2021.117734>
- 1204 Guo, H., Zhu, X., Jeanne Huang, J., Zhang, Z., Tian, S., & Chen, Y. (2023). An enhanced deep learning
1205 approach to assessing inland lake water quality and its response to climate and anthropogenic factors.
1206 *Journal of Hydrology*, 620(PA), 129466. <https://doi.org/10.1016/j.jhydrol.2023.129466>
- 1207 Gupta, A., Stead, T. S., & Ganti, L. (2024). Determining a meaningful R-squared value in clinical medicine.
1208 *Academic Medicine & Surgery*. <https://doi.org/10.62186/001c.125154>
- 1209 Gutknecht, E., Reffray, G., Mignot, A., Dabrowski, T., & Sotillo, M. G. (2019). Modelling the marine
1210 ecosystem of Iberia-Biscay-Ireland (IBI) European waters for CMEMS operational applications. *Ocean
1211 Science*, 15(6), 1489–1516. <https://doi.org/10.5194/os-15-1489-2019>
- 1212 Halder, R. K., Uddin, M. N., Uddin, M. A., Aryal, S., & Khraisat, A. (2024). Enhancing K-nearest neighbor
1213 algorithm: a comprehensive review and performance analysis of modifications. *Journal of Big Data*,
1214 11(1), 113. <https://doi.org/10.1186/s40537-024-00973-y>
- 1215 Hofman, J. M., Goldstein, D. G., & Hullman, J. (2020). How Visualizing Inferential Uncertainty Can Mislead
1216 Readers about Treatment Effects in Scientific Results. *Conference on Human Factors in Computing
1217 Systems - Proceedings*. <https://doi.org/10.1145/3313831.3376454>
- 1218 IOCCG. (2000). *Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters*.
1219 Sathyendranath, S. (ed.), In *Reports of the International Ocean-Colour Coordinating Group*, No. 3,
1220 IOCCG, Dartmouth, Canada.
- 1221 IOCCG. (2019). *Uncertainties in Ocean Colour Remote Sensing*. Mélin F. (ed.), IOCCG Report Series, No. 18,
1222 International Ocean Colour Coordinating Group, Dartmouth, Canada. <http://dx.doi.org/10.25607/OBP696>

- 1223 Janga, B., Asamani, G. P., Sun, Z., & Cristea, N. (2023). A Review of Practical AI for Remote Sensing in Earth
1224 Sciences. *Remote Sensing*, 15(16). <https://doi.org/10.3390/rs15164112>
- 1225 Johnstone, I. M., & Titterton, D. M. (2009). Statistical challenges of high-dimensional data. *Philosophical
1226 transactions of the Royal Society A: Mathematical, physical and engineering sciences*, 367(1906), 4237-
1227 4253. <https://doi.org/10.1098/rsta.2009.0159>
- 1228 Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining*, 15(4), 531–538.
1229 <https://doi.org/10.1002/sam.11583>
- 1230 Joseph, V. R., & Vakayil, A. (2022). SPlit: An Optimal Method for Data Splitting. *Technometrics*, 64(2), 166–
1231 176. <https://doi.org/10.1080/00401706.2021.1921037>
- 1232 Kang, S. (2020). Model validation failure in class imbalance problems. *Expert Systems with Applications*, 146.
1233 <https://doi.org/10.1016/j.eswa.2020.113190>
- 1234 Karakaya, N., & Evrendilek, F. (2011). Monitoring and validating spatio-temporal dynamics of biogeochemical
1235 properties in Mersin Bay (Turkey) using Landsat ETM+. *Environmental Monitoring and Assessment*,
1236 181(1–4), 457–464. <https://doi.org/10.1007/s10661-010-1841-5>
- 1237 Karasiak, N., Dejoux, J. F., Monteil, C., & Sheeren, D. (2022). Spatial dependence between training and test
1238 sets: another pitfall of classification accuracy assessment in remote sensing. *Machine Learning*, 111(7),
1239 2715-2740. <https://doi.org/10.1007/s10994-021-05972-1>
- 1240 Karimi, N., & Torabi, O. (2025). Remote sensing-based bathymetry mapping in shallow lakes: comparative
1241 analysis of Sentinel-2 and Landsat-8 imagery integrated with machine learning techniques. *Advances in
1242 Space Research*. <https://doi.org/10.1016/j.asr.2025.10.028>
- 1243 Kim, Y. H., Son, S., Kim, H. C., Kim, B., Park, Y. G., Nam, J., & Ryu, J. (2020). Application of satellite remote
1244 sensing in monitoring dissolved oxygen variabilities: A case study for coastal waters in Korea.
1245 *Environment International*, 134(October 2019), 105301. <https://doi.org/10.1016/j.envint.2019.105301>
- 1246 Kosaka, N., Iizuka, T., Umemiya, Y., Itami, G., Murata, A., & Mitarai, S. (2022). A Comparison of Satellite and
1247 Model with In-Situ Ocean Observation Utilizing an Autonomous Wave Glider. *International Geoscience
1248 and Remote Sensing Symposium (IGARSS)*, 2022-July, 6864–6867.
1249 <https://doi.org/10.1109/IGARSS46834.2022.9883349>
- 1250 Krishnaraj, A., & Honnasiddaiah, R. (2022). Remote sensing and machine learning based framework for the
1251 assessment of spatio-temporal water quality in the Middle Ganga Basin. *Environmental Science and
1252 Pollution Research*, 29(43), 64939–64958. <https://doi.org/10.1007/s11356-022-20386-9>
- 1253 Kudela, R. M., Senn, D. B., Richardson, E. T., Bouma-Gregson, K., Bergamaschi, B. A., & Sim, L. (2024).
1254 Evaluation and Refinement of Chlorophyll-a Algorithms for High-Biomass Blooms in San Francisco Bay
1255 (USA). *Remote Sensing*, 16(6). <https://doi.org/10.3390/rs16061103>
- 1256 Kumar, S., Singh, S. K., & Nelson, L. (2025). Computational intelligence in decision support: Scope and
1257 techniques. In *Uncertainty in Computational Intelligence-Based Decision Making* (pp. 219-238).
1258 Academic Press. <https://doi.org/10.1016/B978-0-443-21475-2.00019-9>
- 1259 Laura, Z. M., Vittorio, B., Gianluca, V., Luis, G. V., Rowe, D. B. F., Robert, F., Jaime, P., Simon, O., Jing, T.,
1260 Simone, C., & Christian, M. (2025). CIAO: A Machine-Learning Algorithm for Mapping Arctic Ocean
1261 Chlorophyll-a from Space. *Science of Remote Sensing*, 112852.
1262 <https://doi.org/10.1016/j.srs.2025.100212>
- 1263 Li, L., Gu, M., Gong, C., Hu, Y., Wang, X., Yang, Z., & He, Z. (2023b). An advanced remote sensing retrieval
1264 method for urban non-optically active water quality parameters: An example from Shanghai. *Science of
1265 the Total Environment*, 880(November 2022), 163389. <https://doi.org/10.1016/j.scitotenv.2023.163389>
- 1266 Li, L., Sun, B., Xue, C., Zheng, Z., & Ma, Y. (2025). An Enhanced Transfer Learning Remote Sensing
1267 Inversion of Coastal Water Quality: A Case Study of Dissolved Oxygen. *IEEE Journal of Selected
1268 Topics in Applied Earth Observations and Remote Sensing*.
1269 <https://doi.org/10.1109/JSTARS.2025.3599644>

- 1270 Li, W., Fang, H., Qin, G., Tan, X., Huang, Z., Zeng, F., Du, H., & Li, S. (2020). Concentration estimation of
 1271 dissolved oxygen in Pearl River Basin using input variable selection and machine learning techniques.
 1272 *Science of the Total Environment*, 731, 139099. <https://doi.org/10.1016/j.scitotenv.2020.139099>
- 1273 Li, Y., Robinson, S. V. J., Nguyen, L. H., & Liu, J. (2023a). Satellite prediction of coastal hypoxia in the
 1274 northern Gulf of Mexico. *Remote Sensing of Environment*, 284(May 2022), 113346.
 1275 <https://doi.org/10.1016/j.rse.2022.113346>
- 1276 Limburg, K. E., Breitburg, D., Swaney, D. P., & Jacinto, G. (2020). Ocean deoxygenation: a primer. *One Earth*,
 1277 2(1), 24-29. <https://doi.org/10.1016/j.oneear.2020.01.001>
- 1278 Liu, H., Liu, H., Peng, H., Zheng, T., Zhao, Y., Wang, S., ... & Luo, J. (2025). Climate change poses risks to
 1279 water retention and carbon sequestration capacity in the source area of the Yangtze River. *Water*
 1280 *Research*, 124461. <https://doi.org/10.1016/j.watres.2025.124461>
- 1281 Liu, M., Wang, L., & Qiu, F. (2022). Using MODIS data to track the long-term variations of dissolved oxygen
 1282 in Lake Taihu. *Frontiers in Environmental Science*, 10(December), 1–15.
 1283 <https://doi.org/10.3389/fenvs.2022.1096843>
- 1284 Lyu, H., Li, X., Wang, Y., Jin, Q., Cao, K., Wang, Q., & Li, Y. (2015). Evaluation of chlorophyll-a retrieval
 1285 algorithms based on MERIS bands for optically varying eutrophic inland lakes. *Science of the Total*
 1286 *Environment*, 530, 373-382. <https://doi.org/10.1016/j.scitotenv.2015.05.115>
- 1287 Lizcano-Sandoval, L., Beck, M. W., Scolaro, S., Sherwood, E. T., & Muller-Karger, F. (2025). Cloud-based
 1288 satellite remote sensing for enhancing seagrass monitoring and ecosystem management. *ISPRS Journal*
 1289 *of Photogrammetry and Remote Sensing*, 227, 508-518. <https://doi.org/10.1016/j.isprsjprs.2025.06.034>
- 1290 Lønborg, C., Müller, M., Butler, E. C. V., Jiang, S., Ooi, S. K., Trinh, D. H., Wong, P. Y., Ali, S. M., Cui, C.,
 1291 Siong, W. B., Yando, E. S., Friess, D. A., Rosentreter, J. A., Eyre, B. D., & Martin, P. (2021). Nutrient
 1292 cycling in tropical and temperate coastal waters: Is latitude making a difference? *Estuarine, Coastal and*
 1293 *Shelf Science*, 262(April). <https://doi.org/10.1016/j.ecss.2021.107571>
- 1294 Luo, A., Chen, H., Gao, X., Carvalho, L., Zhang, H., & Yang, J. (2024). The impact of rainfall events on
 1295 dissolved oxygen concentrations in a subtropical urban reservoir. *Environmental Research*,
 1296 244(December 2023), 117856. <https://doi.org/10.1016/j.envres.2023.117856>
- 1297 Luo, X., Li, N., Zhang, Y., Zhang, Y., Shi, K., Qin, B., Zhu, G., Jeppesen, E., Brookes, J. D., & Sun, X. (2025).
 1298 Real-Time Monitoring of Dissolved Oxygen Using a Novel Ground-Based Hyperspectral Proximal
 1299 Sensing System. *ACS ES and T Water*. <https://doi.org/10.1021/acsestwater.4c00896>
- 1300 Maciel, D. A., Pahlevan, N., Barbosa, C. C. F., Martins, V. S., Smith, B., O'Shea, R. E., Balasubramanian, S.
 1301 V., Saranathan, A. M., & Novo, E. M. L. M. (2023). Towards global long-term water transparency
 1302 products from the Landsat archive. *Remote Sensing of Environment*, 299(October), 113889.
 1303 <https://doi.org/10.1016/j.rse.2023.113889>
- 1304 Maggio, S., Bouvier, V., & Dreyfus-Schmidt, L. (2022). Performance Prediction Under Dataset Shift.
 1305 *Proceedings - International Conference on Pattern Recognition, 2022-Augus*, 2466–2474.
 1306 <https://doi.org/10.1109/ICPR56361.2022.9956676>
- 1307 Mahaffey, C., Hull, T., Hunter, W., Greenwood, N., Palmer, M., Sharples, J., ... & Williams, C. (2023). Climate
 1308 change impacts on dissolved oxygen concentration in marine and coastal waters around the uk and
 1309 ireland. *MCCIP Sci. Rev*, 31(10.14465). <https://doi.org/10.14465/2023.reu07.oxy>
- 1310 Makri, D., Kalogirou, E., Iqbal, A., Hadjimitsis, D., & Mettas, C. (2025, September). Habitat types mapping
 1311 using remote sensing techniques in Cyprus. In *Eleventh International Conference on Remote Sensing and*
 1312 *Geoinformation of the Environment (RSCy2025)* (Vol. 13816, pp. 34-40). SPIE.
 1313 <https://doi.org/10.1117/12.3072248>
- 1314 Malinin, A., Band, N., Ganshin, Alexander, Chesnokov, G., Gal, Y., Gales, M. J. F., Noskov, A., Ploskonosov,
 1315 A., Prokhorenkova, L., Provilkov, I., Raina, V., Raina, V., Roginskiy, Denis, Shmatova, M., Tigas, P., &
 1316 Yangel, B. (2021). Shifts: A Dataset of Real Distributional Shift Across Multiple Large-Scale Tasks.
 1317 *NeurIPS*, 1–44. <http://arxiv.org/abs/2107.07455>

- 1318 Marchese, C., Colella, S., Brando, V. E., Zoffoli, M. L., & Volpe, G. (2024). Towards accurate L4 ocean colour
1319 products: Interpolating remote sensing reflectance via DINEOF. *International Journal of Applied Earth*
1320 *Observation and Geoinformation*, 135(August 2024), 104270. <https://doi.org/10.1016/j.jag.2024.104270>
- 1321 Martin Traykovski, L. V. (2003). Feature-based classification of optical water types in the Northwest Atlantic
1322 based on satellite ocean color data. *Journal of Geophysical Research*, 108(C5), 1–18.
1323 <https://doi.org/10.1029/2001jc001172>
- 1324 McGovern, J. V., Dabrowski, T., Pereiro, D., Gutknecht, E., Lorente, P., Reffray, G., & Sotillo, M. G. (2022).
1325 Quality Information Document for Atlantic IBI - Iberia Biscay Ireland - Biogeochemical Multi-year
1326 Product: IBI-MULTIYEAR-BGC-005-003, Ref: CMEMS-IBI-QUID-005-003. *MERCATOR OCEAN*
1327 *INTERNATIONAL*, 62. [https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-IBI-QUID-](https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-IBI-QUID-005-003.pdf)
1328 [005-003.pdf](https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-IBI-QUID-005-003.pdf)
- 1329 Meng, H., Zhang, J., & Zheng, Z. (2022). Retrieving Inland Reservoir Water Quality Parameters Using Landsat
1330 8-9 OLI and Sentinel-2 MSI Sensors with Empirical Multivariate Regression. *International Journal of*
1331 *Environmental Research and Public Health*, 19(13). <https://doi.org/10.3390/ijerph19137725>
- 1332 Mihailov, M. E., Chiroasca, A. V., & Chiroasca, G. (2025). Fusion of In-Situ and Modelled Marine Data for
1333 Enhanced Coastal Dynamics Prediction Along the Western Black Sea Coast. *Journal of Marine Science*
1334 *and Engineering*, 13(2), 1–20. <https://doi.org/10.3390/jmse13020199>
- 1335 Mikelsons, K., Wang, M., Kwiatkowska, E., Jiang, L., Dessailly, D., & Gossn, J. I. (2022). Statistical evaluation
1336 of sentinel-3 OLCI Ocean Color data retrievals. *IEEE Transactions on Geoscience and Remote Sensing*,
1337 60, 1-19. <https://doi.org/10.1109/TGRS.2022.3226158>
- 1338 Mobley, C. D. (2022). *The Oceanic Optics Book*. International Ocean Colour Coordinating Group (IOCCG)
1339 Dartmouth, NS, Canada, 924. <https://doi.org/10.25607/OBP-1710>
- 1340 Mohseni, F., Saba, F., Mirmazloumi, S. M., Amani, M., Mokhtarzade, M., Jamali, S., & Mahdavi, S. (2022).
1341 Ocean water quality monitoring using remote sensing techniques: A review. *Marine Environmental*
1342 *Research*, 180(August), 105701. <https://doi.org/10.1016/j.marenvres.2022.105701>
- 1343 Moon, J., Jung, S., Suh, S., & Pyo, J. (2025). Development of deep learning quantization framework for remote
1344 sensing edge device to estimate inland water quality in South Korea. *Water Research*, 123760.
1345 <https://doi.org/10.1016/j.watres.2025.123760>
- 1346 Morbidelli, R., Saltalippi, C., Flammini, A., Corradini, C., Wilkinson, S. M., & Fowler, H. J. (2018). Influence
1347 of temporal data aggregation on trend estimation for intense rainfall. *Advances in Water Resources*, 122,
1348 304-316. <https://doi.org/10.1016/j.advwatres.2018.10.027>
- 1349 Morel, A., & Prieur, L. (1977). Analysis of variations in ocean color. *Limnology and Oceanography*, 22(4),
1350 709–722. <https://doi.org/10.4319/lo.1977.22.4.0709>
- 1351 Moustafa, M., El-Gharabawy, S., Hamouda, A., El Maghraby, M., & Abdel-Fattah, T. (2025). Monitoring coral
1352 reefs by remote sensing techniques and acoustic data as a ground truth reference in Hurghada, Red Sea,
1353 Egypt. *Egyptian Journal of Aquatic Research*. <https://doi.org/10.1016/j.ejar.2025.05.008>
- 1354 Musacchio, A., Re, V., Mas-Pla, J., & Sacchi, E. (2020). EU Nitrates Directive, from theory to practice:
1355 Environmental effectiveness and influence of regional governance on its performance. *Ambio*, 49(2),
1356 504-516.
- 1357 Nadeem, A., Rana, A. D., Batool, S. A., Azhar, A., Ishfaq, U. B. E., & Hameed, A. (2025). Satellite based
1358 assessment of potential fishing zones (PFZs) within the exclusive economic zone (EEZ) of Pakistan.
1359 *Continental Shelf Research*, 286, 105410. <https://doi.org/10.1016/j.csr.2025.105410>
- 1360 Naylor, R. L., Hardy, R. W., Buschmann, A. H., Bush, S. R., Cao, L., Klinger, D. H., Little, D. C., Lubchenco,
1361 J., Shumway, S. E., & Troell, M. (2021). A 20-year retrospective review of global aquaculture. *Nature*,
1362 591(7851), 551–563. <https://doi.org/10.1038/s41586-021-03308-6>
- 1363 Nezlin, N. P., Son, S., Salem, S. I., & Ondrusek, M. E. (2025). Chlorophyll-a in the Chesapeake Bay Estimated
1364 by Extra-Trees Machine Learning Modeling. *Remote Sensing*, 17(13), 1–23.
1365 <https://doi.org/10.3390/rs17132151>

- 1366 Nguyen, Q.H., Ly, H.B., Ho, L.S., Al-Ansari, N., Le, H.V., Tran, V.Q., Prakash, I. and Pham, B.T., (2021).
 1367 Influence of data splitting on performance of machine learning models in prediction of shear strength of
 1368 soil. *Mathematical Problems in Engineering*, 2021(1), 4832864. <https://doi.org/10.1155/2021/4832864>
- 1369 O'Boyle, S., & Nolan, G. (2010). The influence of water column stratification on dissolved oxygen levels in
 1370 coastal and shelf waters around Ireland. In *Biology and Environment: Proceedings of the Royal Irish*
 1371 *Academy* (Vol. 110, No. 3, pp. 195-209). Royal Irish Academy.
- 1372 Olbert, A. I., Hartnett, M., Dabrowski, T., & Mikolajewicz, U. (2011). Long-term inter-annual variability of a
 1373 cyclonic gyre in the western Irish Sea. *Continental Shelf Research*, 31(13), 1343-1356.
 1374 <https://doi.org/10.1016/j.csr.2011.05.010>
- 1375 Ötles, E., Oh, J., Li, B., Bochinski, M., Joo, H., Ortwine, J., Shenoy, E., Washer, L., Young, V. B., Rao, K., &
 1376 Wiens, J. (2021). Mind the Performance Gap: Examining Dataset Shift During Prospective Validation.
 1377 *Proceedings of Machine Learning Research*, 149, 506–534.
- 1378 Ozdemir, S., Yaqub, M., & Yildirim, S. O. (2023). A systematic literature review on lake water level prediction
 1379 models. In *Environmental Modelling and Software* (Vol. 163, Issue December 2022, p. 105684). Elsevier
 1380 Ltd. <https://doi.org/10.1016/j.envsoft.2023.105684>
- 1381 Pan, C., Luo, Z., Wei, Z., Wang, L., Wang, M., Peng, Y., & Xia, L. (2025). Remote Sensing Inversion
 1382 Technology for the Evaluation of Coastal Water Eutrophication with the Pressure-State-Response
 1383 Framework. *Journal of Cleaner Production*, 145771. <https://doi.org/10.1016/j.jclepro.2025.145771>
- 1384 Pannard, A., Souchu, P., Chauvin, C., Delabuis, M., Gascuel-Oudou, C., Jeppesen, E., Le Moal, M.,
 1385 Ménesguen, A., Pinay, G., Rabalais, N. N., Souchon, Y., & Gross, E. M. (2024). Why are there so many
 1386 definitions of eutrophication? *Ecological Monographs*, 94(3), 1–25. <https://doi.org/10.1002/ecm.1616>
- 1387 Peterson, K. T., Sagan, V., & Sloan, J. J. (2020). Deep learning-based water quality estimation and anomaly
 1388 detection using Landsat 8 Sentinel 2 virtual constellation and cloud computing (pp. 510–525).
 1389 <https://doi.org/10.1080/15481603.2020.1738061>
- 1390 Quang, N. H., Dinh, N. T., Dien, N. T., & Son, L. T. (2023). Calibration of Sentinel-2 Surface Reflectance for
 1391 Water Quality Modelling in Binh Dinh's Coastal Zone of Vietnam. *Sustainability*, 15(2), 1410.
 1392 <https://doi.org/10.3390/su15021410>
- 1393 Rabalais, N. N., Díaz, R. J., Levin, L. A., Turner, R. E., Gilbert, D., & Zhang, J. (2010). Dynamics and
 1394 distribution of natural and human-caused hypoxia. *Biogeosciences*, 7(2), 585–619.
 1395 <https://doi.org/10.5194/bg-7-585-2010>
- 1396 Raheli, B., Talabbeydokhti, N., & Nourani, V. (2024). Uncertainty assessment of optically active and inactive
 1397 water quality parameters predictions using satellite data, deep and ensemble learnings. *Journal of*
 1398 *Hydrology*, 644, 132091. <https://doi.org/10.1016/j.jhydrol.2024.132091>
- 1399 Rixen, M., Dowell, M., Immler, F., Dusart, J., & Kristopaitis, E. (2024). Earth Observation Strategic Research
 1400 and Innovation Agenda (E. Commission, Ed.; Issue KJ-NA-31-903-EN-N (online)). Publications Office
 1401 of the European Union, Luxembourg. <https://doi.org/10.2760/18985>
- 1402 Rodrigues, M., Santana, P. & Rocha, A. (2019). Bootstrap approach to validate the performance of models for
 1403 predicting mortality risk temperature in Portuguese Metropolitan Areas. *Environ Health* 18, 25.
 1404 <https://doi.org/10.1186/s12940-019-0462-x>
- 1405 Rossi, L., Fiorentino, M. C., Mancini, A., Paolanti, M., Rosati, R., & Zingaretti, P. (2024). Generalizability and
 1406 robustness evaluation of attribute-based zero-shot learning. *Neural Networks*, 175(September 2023),
 1407 106278. <https://doi.org/10.1016/j.neunet.2024.106278>
- 1408 Sagan, V., Peterson, K. T., Maimaitjiang, M., Sidike, P., Sloan, J., Greeling, B. A., Maalouf, S., & Adams, C.
 1409 (2020). Monitoring inland water quality using remote sensing: potential and limitations of spectral
 1410 indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Science Reviews*,
 1411 205(August 2019), 103187. <https://doi.org/10.1016/j.earscirev.2020.103187>
- 1412 Sajib, A. M., Diganta, M. T. M., Jordan, C., Cusack, C., Croot, P. L., Uddin, M. G., Olbert, A. I. (2025b).
 1413 Validation of Sentinel-3 OLCI Water Product(s) in the Irish Shelf. (Submitted to publication).

- 1414 Sajib, A. M., Diganta, M. T. M., Moniruzzaman, M., Rahman, A., Dabrowski, T., Uddin, M. G., & Olbert, A. I.
1415 (2024b). Assessing water quality of an ecologically critical urban canal incorporating machine learning
1416 approaches. *Ecological Informatics*, 80(August 2023), 102514.
1417 <https://doi.org/10.1016/j.ecoinf.2024.102514>
- 1418 Sajib, A. M., Uddin, M. G., Jordan, C., Dabrowski, T., Cusack, C., & Olbert, A. I. (2024a). Assessing Remote
1419 Sensing Data Quality and Alignment with In-Situ Measurements in Ireland: A Cloud Coverage Analysis
1420 Using Multi-Dataset Integration Abdul. *Civil Engineering Research in Ireland 2024 (CERI2024)*
1421 Conference, 2024(August), 5–11.
- 1422 Sajib, A. M., Uddin, M. G., Rahman, A., Olbert, A. I. (2025a). Remote sensing applications for monitoring
1423 optically inactive water quality: A Comprehensive Review. *Earth Science Reviews*.
1424 <https://doi.org/10.1016/j.earscirev.2025.105259>
- 1425 Sakib, M., Mustajab, S., & Alam, M. (2025). Ensemble deep learning techniques for time series analysis: a
1426 comprehensive review, applications, open issues, challenges, and future directions. In *Cluster Computing*
1427 (Vol. 28, Issue 1). Springer US. <https://doi.org/10.1007/s10586-024-04684-0>
- 1428 Salas, E. A. L., Kumaran, S. S., Partee, E. B., Willis, L. P., & Mitchell, K. (2022). Potential of mapping
1429 dissolved oxygen in the Little Miami River using Sentinel-2 images and machine learning algorithms.
1430 *Remote Sensing Applications: Society and Environment*, 26, 100759.
1431 <https://doi.org/10.1016/j.rsase.2022.100759>
- 1432 Samlas, O., Luik, S. T., Korabel, V., She, J., & Lips, U. (2025). Applicability of Copernicus marine service
1433 products for the eutrophication status assessment of the Baltic Sea. *Marine Pollution Bulletin*, 216(April),
1434 117975. <https://doi.org/10.1016/j.marpolbul.2025.117975>
- 1435 Sarker, I. H. (2021). Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-
1436 Making and Applications Perspective. *SN Computer Science*, 2(5), 1–22. <https://doi.org/10.1007/s42979-021-00765-8>
- 1438 Saulquin, B., Gohin, F., & Fanton d'Andon, O. (2019). Interpolated fields of satellite-derived multi-algorithm
1439 chlorophyll-a estimates at global and European scales in the frame of the European Copernicus-Marine
1440 Environment Monitoring Service. *Journal of Operational Oceanography*, 12(1), 47–57.
1441 <https://doi.org/10.1080/1755876X.2018.1552358>
- 1442 Schiltz, F., Masci, C., Agasisti, T., & Horn, D. (2018). Using regression tree ensembles to model interaction
1443 effects: a graphical approach. *Applied Economics*, 50(58), 6341–6354.
1444 <https://doi.org/10.1080/00036846.2018.1489520>
- 1445 S  f  rian, R., Berthet, S., Yool, A., Palmi  ri, J., Bopp, L., Tagliabue, A., ... & Yamamoto, A. (2020). Tracking
1446 improvement in simulated marine biogeochemistry between CMIP5 and CMIP6. *Current Climate*
1447 *Change Reports*, 6(3), 95-119. <https://doi.org/10.1007/s40641-020-00160-0>
- 1448 S  f  rian, R., Bopp, L., Gehlen, M., C. Orr, J., Eth  , C., Cadule, P., Aumont, O., M  lia, D. S. Voltaire, A., and
1449 Madec, G. (2013). Skill assessment of three earth system models with common marine biogeochemistry.
1450 *Climate Dynamics* 40(9), 2549-2573. <https://doi.org/10.1007/s00382-012-1362-8>
- 1451 Shao, J., Huang, S., Chen, Y., Qi, J., Wang, Y., Wu, S., Liu, R., & Du, Z. (2024). Satellite-Based Global Sea
1452 Surface Oxygen Mapping and Interpretation with Spatiotemporal Machine Learning. *Environmental*
1453 *Science and Technology*, 58(1), 498–509. <https://doi.org/10.1021/acs.est.3c08833>
- 1454 Sharaf El Din, E., & Zhang, Y. (2017). Estimation of both optical and nonoptical surface water quality
1455 parameters using Landsat 8 OLI imagery and statistical techniques. *Journal of Applied Remote Sensing*,
1456 11(04), 1. <https://doi.org/10.1117/1.jrs.11.046008>
- 1457 Sharples, J., Holt, J., & Wakelin, S. (2020). Impacts of climate change on shelf sea stratification, relevant to the
1458 coastal and marine environment around the UK. *MCCIP Science Review 2020*, 103-115.
1459 <https://doi.org/10.14465/2020.arc05.str>
- 1460 Shaw, B. D. (2017). *Uncertainty Analysis of Experimental Data with R*. (1st ed.). Chapman and Hall/CRC.
1461 <https://doi.org/10.1201/9781315366715>

- 1462 Shen, A., Ma, Y., Li, Y., Hong, P., Niu, Z., Zhang, Y., ... & Zhang, D. (2025). RFAGB model: A new machine
1463 learning model for microplastic inversion based on remotely sensed data in Bohai Sea. *Water Research*,
1464 124490. <https://doi.org/10.1016/j.watres.2025.124490>
- 1465 Shi, K., Lang, Q., Wang, P., Yang, W., Chen, G., Yin, H., Zhang, Q., Li, W., & Wang, H. (2023). Dissolved
1466 oxygen concentration inversion based on Himawari-8 data and deep learning: a case study of lake Taihu.
1467 *Frontiers in Environmental Science*, 11(17), 1–15. <https://doi.org/10.3389/fenvs.2023.1230778>
- 1468 Shi, L., Gao, C., Wang, T., Liu, L., Wu, Y., & You, X. (2024). Information extraction of seasonal dissolved
1469 oxygen in urban water bodies based on machine learning using sentinel-2 imagery: An open access
1470 application in Baiyangdian Lake. *Ecological Informatics*, 82(August), 102782.
1471 <https://doi.org/10.1016/j.ecoinf.2024.102782>
- 1472 Šiljić Tomić, A., Antanasijević, D., Ristić, M., Perić-Grujić, A., & Pocajt, V. (2018). A linear and non-linear
1473 polynomial neural network modeling of dissolved oxygen content in surface water: Inter- and
1474 extrapolation performance with inputs' significance analysis. *Science of the Total Environment*, 610–
1475 611, 1038–1046. <https://doi.org/10.1016/j.scitotenv.2017.08.192>
- 1476 Sivakumar, M., Parthasarathy, S., & Padmapriya, T. (2024). Trade-off between training and testing ratio in
1477 machine learning for medical image processing. *PeerJ Computer Science*, 10, e2245.
1478 <https://doi.org/10.7717/peerj-cs.2245>
- 1479 Skákala, J., Ford, D., Fowler, A., Lea, D., Martin, M. J., & Ciavatta, S. (2024). How uncertain and observable
1480 are marine ecosystem indicators in shelf seas?. *Progress in Oceanography*, 224, 103249.
1481 <https://doi.org/10.1016/j.pocean.2024.103249>
- 1482 Skyllas, N. (2018). Evaluation of NEMO–PISCES v2 biogeochemical model with field data from a north-south
1483 gradient in the Northeast Atlantic Ocean. Utrecht University. (Master's thesis).
1484 <https://studenttheses.uu.nl/handle/20.500.12932/30621>
- 1485 Stein, A., Ge, Y., & Fabris-Rotelli, I. (2018). Introduction to the special issue “uncertainty in remote sensing
1486 image analysis.” *Remote Sensing*, 10(12), 4–6. <https://doi.org/10.3390/rs10121975>
- 1487 Subbaswamy, A., Schulam, P., & Saria, S. (2020). Preventing failures due to dataset shift: Learning predictive
1488 models that transport. *AISTATS 2019 - 22nd International Conference on Artificial Intelligence and
1489 Statistics*, 89. <https://doi.org/10.48550/arXiv.1812.04597>
- 1490 Taghiyarrenani, Z., Nowaczyk, S., Pashami, S., & Bouguelia, M. R. (2023). Multi-domain adaptation for
1491 regression under conditional distribution shift. *Expert Systems with Applications*, 224(August 2022).
1492 <https://doi.org/10.1016/j.eswa.2023.119907>
- 1493 Thai, H. T., Mentré, F., Holford, N. H., Veyrat-Follet, C., & Comets, E. (2014). Evaluation of bootstrap
1494 methods for estimating uncertainty of parameters in nonlinear mixed-effects models: a simulation study
1495 in population pharmacokinetics. *Journal of pharmacokinetics and pharmacodynamics*, 41(1), 15-33.
1496 <https://doi.org/10.1007/s10928-013-9343-z>
- 1497 Tian, D., Zhao, X., Gao, L., Liang, Z., Yang, Z., Zhang, P., Wu, Q., Ren, K., Li, R., Yang, C., Li, S., Wang, M.,
1498 He, Z., Zhang, Z., & Chen, J. (2024). Estimation of water quality variables based on machine learning
1499 model and cluster analysis-based empirical model using multi-source remote sensing data in inland
1500 reservoirs, South China. *Environmental Pollution*, 342(November 2023), 123104.
1501 <https://doi.org/10.1016/j.envpol.2023.123104>
- 1502 Toming, K., Liu, H., Soomets, T., Uemaa, E., Nöges, T., & Kutser, T. (2024). Estimation of the
1503 Biogeochemical and Physical Properties of Lakes Based on Remote Sensing and Artificial Intelligence
1504 Applications. *Remote Sensing*, 16(3), 1–28. <https://doi.org/10.3390/rs16030464>
- 1505 Trodd, W., O'Boyle, & Gurrie, M. (2022). *Water Quality in Ireland: A Summary*. 1–112.
1506 [https://www.epa.ie/publications/monitoring--assessment/freshwater--
1507 marine/EPA_WaterQualityReport2016_2021.pdf](https://www.epa.ie/publications/monitoring--assessment/freshwater--marine/EPA_WaterQualityReport2016_2021.pdf)
- 1508 Turhan, B. (2012). On the dataset shift problem in software engineering prediction models. *Empirical Software
1509 Engineering*, 17(1–2), 62–74. <https://doi.org/10.1007/s10664-011-9182-8>

- 1510 Uddin, G., Rahman, A., Rosa, F., & Olbert, A. I. (2024a). Data-driven evolution of water quality models: An in-
 1511 depth investigation of innovative outlier detection approaches-A case study of Irish Water Quality Index
 1512 (IEWQI) model. *Water Research*, 255(March), 121499. <https://doi.org/10.1016/j.watres.2024.121499>
- 1513 Uddin, M. G., Diganta, M. T. M., Sajib, A. M., Rahman, A., Olbert A. I., (2025a). A methodological framework
 1514 for the development of scalable data-driven model for retrieving chlorophyll-a in transitional and coastal
 1515 waters using remote sensing technology. (Submitted to publication).
- 1516 Uddin, M. G., Nash, S., Mahammad Diganta, M. T., Rahman, A., & Olbert, A. I. (2022a). Robust machine
 1517 learning algorithms for predicting coastal water quality index. *Journal of Environmental Management*,
 1518 321(June), 115923. <https://doi.org/10.1016/j.jenvman.2022.115923>
- 1519 Uddin, M. G., Nash, S., Rahman, A., & Olbert, A. I. (2022b). A comprehensive method for improvement of
 1520 water quality index (WQI) models for coastal water quality assessment. *Water Research*, 219(May),
 1521 118532. <https://doi.org/10.1016/j.watres.2022.118532>
- 1522 Uddin, M. G., Nash, S., Rahman, A., & Olbert, A. I. (2023). Assessing optimization techniques for improving
 1523 water quality model. *Journal of Cleaner Production*, 385, 135671.
 1524 <https://doi.org/10.1016/j.jclepro.2022.135671>
- 1525 Vakili, T., & Amanollahi, J. (2020). Determination of optically inactive water quality variables using Landsat 8
 1526 data: A case study in Geshlagh reservoir affected by agricultural land use. *Journal of Cleaner Production*,
 1527 247, 119134. <https://doi.org/10.1016/j.jclepro.2019.119134>
- 1528 Van Allen, T., Singh, A., Greiner, R., & Hooper, P. (2008). Quantifying the uncertainty of a belief net response:
 1529 Bayesian error-bars for belief net inference. *Artificial Intelligence*, 172(4–5), 483–513.
 1530 <https://doi.org/10.1016/j.artint.2007.09.004>
- 1531 Vanhellemont, Q., & Ruddick, K. (2021). Atmospheric correction of Sentinel-3/OLCI data for mapping of
 1532 suspended particulate matter and chlorophyll-a concentration in Belgian turbid coastal waters. *Remote
 1533 Sensing of Environment*, 256(January), 112284. <https://doi.org/10.1016/j.rse.2021.112284>
- 1534 Vervatis, V. D., de Mey-Frémaux, P., Karagiorgos, J., Lemieux-Dudon, B., Ayoub, N. K., & Sofianos, S.
 1535 (2025). Regional ocean model uncertainties using stochastic parameterizations and a global atmospheric
 1536 ensemble. *Ocean Modelling*, 194, 102501.
- 1537 Wang, X., Gong, C., Ji, T., Hu, Y., & Li, L. (2021). Inland water quality parameters retrieval based on the VIP-
 1538 SPCA by hyperspectral remote sensing. *Journal of Applied Remote Sensing*, 15(04), 1–17.
 1539 <https://doi.org/10.1117/1.jrs.15.042609>
- 1540 Warner, T. A., & Shank, M. C. (1997). Spatial autocorrelation analysis of hyperspectral imagery for feature
 1541 selection. *Remote Sensing of Environment*, 60(1), 58-70. [https://doi.org/10.1016/S0034-4257\(96\)00138-1](https://doi.org/10.1016/S0034-4257(96)00138-1)
- 1543 Yang, W., Fu, B., Li, S., Lao, Z., Deng, T., He, W., He, H., & Chen, Z. (2023). Monitoring multi-water quality
 1544 of internationally important karst wetland through deep learning, multi-sensor and multi-platform remote
 1545 sensing images: A case study of Guilin, China. *Ecological Indicators*, 154(July), 110755.
 1546 <https://doi.org/10.1016/j.ecolind.2023.110755>
- 1547 Yang, Z., Gong, C., Ji, T., Hu, Y., & Li, L. (2022). Water Quality Retrieval from ZY1-02D Hyperspectral
 1548 Imagery in Urban Water Bodies and Comparison with Sentinel-2. *Remote Sensing*, 14(19).
 1549 <https://doi.org/10.3390/rs14195029>
- 1550 Zarbipour, P., Nikoo, M. R., Akbari, H., Nazari, R., & Karimi, M. (2026). Bridging Causality and Deep
 1551 Learning for Harmful Algal Bloom Prediction. *Water Research*, 125492.
 1552 <https://doi.org/10.1016/j.watres.2026.125492>
- 1553 Zhang, Y., Li, Z., Bai, K., Wei, Y., Xie, Y., Zhang, Y., Ou, Y., Cohen, J., Zhang, Y., Peng, Z., Zhang, X., Chen,
 1554 C., Hong, J., Xu, H., Guang, J., Lv, Y., Li, K., & Li, D. (2021). Satellite remote sensing of atmospheric
 1555 particulate matter mass concentration: Advances, challenges, and perspectives. *Fundamental Research*,
 1556 1(3), 240–258. <https://doi.org/10.1016/j.fmre.2021.04.007>
- 1557 Zhao, X., Fioletov, V., Griffin, D., Mclinden, C., Staebler, R., Strawbridge, K., Davies, J., Abboud, I., Lee, S.
 1558 C., Cede, A., & Swap, R. (2024). The differences between remote sensing and in situ air pollutants
 1559 measurements over the Canadian Oil Sands. *Atmospheric Measurement Techniques*, 2, 1–40.

- 1560 Zhou, X., Liu, C., Carrion, D., Akbar, A., & Wang, H. (2024). Spectro-environmental factors integrated
1561 ensemble learning for urban river network water quality remote sensing. *Water Research*,
1562 267(September), 122544. <https://doi.org/10.1016/j.watres.2024.122544>
- 1563 Zhu, X., Guo, H., Huang, J. J., Tian, S., Xu, W., & Mai, Y. (2022a). An ensemble machine learning model for
1564 water quality estimation in coastal area based on remote sensing imagery. *Journal of Environmental*
1565 *Management*, 323(February), 116187. <https://doi.org/10.1016/j.jenvman.2022.116187>
- 1566 Zhu, Z., Qiu, S., & Ye, S. (2022b). Remote sensing of land change: A multifaceted perspective. *Remote Sensing*
1567 *of Environment*, 282(September). <https://doi.org/10.1016/j.rse.2022.113266>