



## Review Article

# Remote sensing applications for monitoring optically inactive water quality indicators: A comprehensive review

Abdul Majed Sajib<sup>a,b,c,d</sup>, Md Galal Uddin<sup>a,b,c,d,\*</sup>, Azizur Rahman<sup>e,f</sup>, Reza Ahmadian<sup>g</sup>, Agnieszka I. Olbert<sup>a,b,d</sup>

<sup>a</sup> Civil Engineering, School of Engineering, College of Science and Engineering, University of Galway, Ireland

<sup>b</sup> Ryan Institute, University of Galway, Ireland

<sup>c</sup> MaREI Research Centre, University of Galway, Ireland

<sup>d</sup> Eco-Hydroinformatics Research Group (EHIRG), School of Engineering, College of Science and Engineering, University of Galway, Ireland

<sup>e</sup> School of Computing, Mathematics and Engineering, Charles Sturt University, Wagga Wagga, Australia

<sup>f</sup> The Gulbali Institute of Agriculture, Water and Environment, Charles Sturt University, Wagga Wagga, Australia

<sup>g</sup> School of Engineering, Cardiff University, The Parade, Cardiff CF24 3AQ, UK

## ARTICLE INFO

## Keywords:

Remote sensing

Water quality

Optically inactive indicators

Retrieval algorithms

Machine learning/artificial intelligence

## ABSTRACT

Monitoring water quality (WQ) is crucial to ensure the safety and health of our water resources. Despite their importance, contemporary WQ monitoring programs are struggling with challenges such as high costs, limited spatio-temporal resolution, and data reliability issues. A promising solution to these challenges is the integration of remote sensing (RS) techniques with machine learning (ML) and artificial intelligence (AI) algorithms, which can significantly improve the efficiency and accuracy of WQ monitoring. Based on the literature, most of the studies have focused on optically active (OA)-WQ indicators like chlorophyll-a and colored dissolved organic matter, etc., while a few studies have been carried out focusing on optically inactive (OI)-WQ indicators. But WQ

**Abbreviations:** AI, Artificial intelligence; ALK, Alkalinity; ANN, Artificial neural network; AWEInsh, Automated water extraction index with no shadow features; BOD<sub>5</sub>, Biological oxygen demand; BPNN, Backpropagation neural network; BPNN-RF, Backpropagation neural network-random forest; BP-VIP-SPCA, Back-propagation-variable importance projection-segmented principal component analysis; BR, Binomial regression; CDOM, Colored dissolved organic matter; CHL, Chlorophyll-a; COD, Chemical oxygen demand; Cond, Conductivity; ConvLSTM, Convolutional long-short-term-memory; COST, Cosine of solar zenith angle; CO<sub>2</sub>, Carbon dioxide; DDA, Direct derivation algorithm; DI, Difference index; DIC, Dissolved inorganic carbon; DIN, Dissolved inorganic nitrogen; DNN, Deep neural network; DNS-DNNs, Dynamic network surgery-deep neural networks; DOS, Dark spectrum fitting; DOX, Dissolved oxygen; DRF, Data regression analysis and fitting; EC, Electrical conductivity; ELU, Exponential linear unit; EM, Empirical model; ETR, Extra tree regression; EU, European Union; FC, Fecal Coliform; GA-PLS, Genetic algorithm and partial least square; GA-XGB, Genetic algorithm-extreme gradient boosting; GP, Genetic programming; GPR, Gaussian process regression; HF-DFM, Hybrid feedback deep factorization machine; Hybrid BPNN, Hybrid Bayesian probabilistic neural network; LassoR, Lasso regression; Leaky-ReLu, Leaky-rectified linear unit; LOOCV-GB, Leave-one-out validation- gradient boosting; LOOCV-XGB, Leave-one-out validation-extreme gradient boosting; LR, Linear regression; MDL, Multimodal deep learning; MDNN, Multimodal Deep Neural Network; ML, Machine learning; MLD, Mixed layer depth; MLP, Multilayer perceptron; MLR, Multiple linear regression; MNDWI, Modified normalized difference water index; MRP, Molybdate reactive phosphorus; MSMCF, Multi-spectral scale morphological combined feature; NDI, Normalized difference index; NDVI, Normalized vegetation difference index; NDWI, Normalized difference water index; NIR, Near-infrared; NN, Neural network; OA, Optically active; OA-WQ, Optically active water quality; OI, Optically inactive; OI-WQ, Optically inactive water quality; OLS, Ordinary least squares; PCA, Principal Component Analysis; PCA-RSR, PCA-based response surface regression; pCO<sub>2</sub>, Partial pressure of CO<sub>2</sub> in surface seawater; PCR, Power curvilinear regression; pDNN, Progressively decreasing deep neural network; PLSR, Partial least squares regression; QLR, Quadratic linear regression; QPR, Quadratic polynomial regression; QSVM, Quadratic support vector machine; R<sup>2</sup>, Coefficient of determination; ReLU, Rectified linear unit; RF, Random forest; RI, Ratio index; RMSE, Root mean square error; RQ, Research question; RS, Remote sensing; SAL, Salinity; SEM, Semi-empirical model; SI, Subtraction index; SLR, Stepwise linear regression; SMLR, Stepwise multiple linear regression; SMPE-GCN, Graph Convolution Network with Superposition of Multi-point Effect; SPIM, Suspended particulate inorganic matter; SPM, Suspended particulate matter; SPOM, Suspended particulate organic matter; SS, Suspended solid; SSD, Secchi disk depth; SST, Sea surface temperature; ST-DBN, Spatiotemporal-integrated deep belief network; SVM, Support vector machine; SVR, Support vector regression; SWIR, Short-wave infrared; TA, Total alkalinity; TC, Total Coliform; TEMP, Temperature; TN, Total nitrogen; TOC, Total organic carbon; TON, Total organic nitrogen; TP, Total phosphorus; TRAN, Transparency; TrC, Transitional and Coastal; TSM, Total suspended matter; TSS, Total suspended solid; TUR, Turbidity; V-LSTM, Vanilla long short-term memory; VNIR, Visible near infrared; WFD, Water framework directive; WOA, World oceanic atlas; WQ, Water quality; WQI, Water quality index; XGB, Extreme gradient boosting.

\* Corresponding author at: Civil Engineering, College of Science and Engineering, University of Galway, Ireland,

E-mail address: [mdgalal.uddin@universityofgalway.ie](mailto:mdgalal.uddin@universityofgalway.ie) (M.G. Uddin).

<https://doi.org/10.1016/j.earscirev.2025.105259>

Received 6 April 2024; Received in revised form 19 August 2025; Accepted 24 August 2025

Available online 3 September 2025

0012-8252/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

monitoring requires a number of OA- and OI-WQ indicators; for instance, the European Union Water Framework Directive (WFD) recommend 11 fundamental WQ indicators, which include both OA- and OI-WQ. Therefore, it is essential to consider both types of indicators in a regular WQ monitoring program to develop an effective water resources management plan. However, several recent studies have shown that the development of RS-based OI-WQ indicator retrieval model(s) introduces considerable uncertainty in the final retrieval results due to various factors. Additionally, these studies highlight that most of the retrieval models may not be suitable for global application. To highlight these challenges, the goal of the research is to conduct a comprehensive analysis of various RS data and existing techniques in order to more accurately retrieve OI-WQ indicators such as pH, dissolved oxygen (DOX), biological oxygen demand (BOD<sub>5</sub>), total phosphorus (TP), total nitrogen (TN), and dissolved inorganic nitrogen (DIN) in different waterbodies. To achieve the research objectives, this study conducted a critical review analysis of 105 research publications, including journal papers and conference papers, from 2005 to 2023. The study not only identified different types of satellite data, such as Landsat, Sentinel, and Aqua/Terra (MODIS), which are widely used, but also identified the advantages and disadvantages of different models, including empirical, semi-empirical, and ML/AI-based methods that are widely used in developing RS-driven retrieval model(s) for various OI-WQ indicators. Additionally, the study identified a range of opportunities (e.g., proposing a structural framework, reliable global model, etc.) and limitations (e.g., lack of in-situ data, structural framework, optimal RS wavelength for different OI-WQ indicators, etc.) in existing retrieval models. Moreover, the analysis suggests that advanced ML/AI approaches can be effective in retrieving OI-WQ indicators compared to other techniques in terms of retrieval data accuracy and reliability. The study also highlights current limitations of RS data and retrieval methods, such as spatial and temporal constraints, the need for improved calibration, and the demand for broader and more diverse training and testing datasets. Finally, the findings emphasize the significant potential of ML/AI algorithms in improving RS-based techniques for WQ monitoring, which may be more useful for water resource management and sustainable development strategies in the future.

## Contents

|                                                                                    |    |
|------------------------------------------------------------------------------------|----|
| 1. Introduction                                                                    | 2  |
| 2. Method of review                                                                | 3  |
| 2.1. Research questions                                                            | 4  |
| 2.2. Selection criteria for reviewed OI-WQ indicators                              | 5  |
| 2.3. Literature search                                                             | 5  |
| 2.4. Inclusion and exclusion criteria                                              | 5  |
| 2.5. Conducting the review                                                         | 5  |
| 2.6. Obtained information                                                          | 5  |
| 3. RS data and atmospheric corrections                                             | 6  |
| 4. RS based retrieval methods for OI-WQ indicators                                 | 7  |
| 5. OI-WQ Indicators                                                                | 8  |
| 5.1. pH                                                                            | 8  |
| 5.1.1. RS data and bands (wavelength) used for pH retrieval models                 | 8  |
| 5.1.2. Performance of various RS data in retrieving pH concentration               | 8  |
| 5.1.3. Comparison of various RS models for retrieving pH                           | 8  |
| 5.1.4. pH retrieval algorithm(s) optimization processes                            | 11 |
| 5.1.5. Limitations                                                                 | 12 |
| 5.2. Dissolved Oxygen (DOX)                                                        | 12 |
| 5.2.1. RS data and bands (wavelength) used for DOX retrieval models                | 12 |
| 5.2.2. Performance of various RS data in retrieving DOX concentration              | 13 |
| 5.2.3. Comparison of various RS models for retrieving DOX                          | 13 |
| 5.2.4. DOX algorithm optimization techniques                                       | 13 |
| 5.2.5. Limitations                                                                 | 15 |
| 5.3. Biochemical oxygen demand (BOD <sub>5</sub> )                                 | 15 |
| 5.3.1. RS data and bands (wavelength) used for BOD <sub>5</sub> retrieval models   | 15 |
| 5.3.2. Performance of various RS data in retrieving BOD <sub>5</sub> concentration | 15 |
| 5.3.3. Comparison of various RS models for retrieving BOD <sub>5</sub>             | 15 |
| 5.3.4. BOD <sub>5</sub> algorithm optimization techniques                          | 17 |
| 5.3.5. Limitations                                                                 | 17 |
| 5.4. Total phosphorus (TP)                                                         | 17 |
| 5.4.1. RS data and bands (wavelength) used for TP retrieval models                 | 17 |
| 5.4.2. Performance of various RS data in retrieving TP concentration               | 17 |
| 5.4.3. Comparison of various RS models for retrieving TP                           | 17 |
| 5.4.4. TP algorithm optimization techniques                                        | 18 |
| 5.4.5. Limitations                                                                 | 18 |
| 5.5. Total Nitrogen (TN)                                                           | 18 |
| 5.5.1. RS data and bands (wavelength) used for TN retrieval models                 | 18 |
| 5.5.2. Performance of various RS data in retrieving TN concentration               | 20 |
| 5.5.3. Comparison of various RS models for retrieving TN                           | 21 |
| 5.5.4. TN algorithm optimization techniques                                        | 21 |
| 5.5.5. Limitations                                                                 | 21 |
| 5.6. Dissolved inorganic nitrogen (DIN)                                            | 21 |
| 5.6.1. RS data and bands (wavelength) used for DIN retrieval models                | 21 |

|        |                                                                                               |    |
|--------|-----------------------------------------------------------------------------------------------|----|
| 5.6.2. | Performance of various RS data in retrieving DIN concentration                                | 21 |
| 5.6.3. | Comparison of various RS models for retrieving DIN                                            | 23 |
| 5.6.4. | DIN algorithm optimization techniques                                                         | 24 |
| 5.6.5. | Limitations                                                                                   | 24 |
| 5.7.   | Relationship between RS bands and in-situ WQ indicators                                       | 24 |
| 6.     | Discussion                                                                                    | 25 |
| 6.1.   | Improve the performance of retrieval algorithms in terms of model sensitivity and uncertainty | 26 |
| 6.2.   | Develop OI-WQ retrieval indicators model for transitional and coastal (TrC) waters            | 26 |
| 6.3.   | Generalized OI-WQ indicator model(s) using ML/AI technique(s)                                 | 27 |
| 6.4.   | Performance evaluation of the RS data in retrieving OI-WQ indicators                          | 27 |
| 7.     | Conclusions                                                                                   | 27 |
|        | Declaration of competing interest                                                             | 27 |
|        | Acknowledgement                                                                               | 27 |
|        | Supplementary data                                                                            | 28 |
|        | Data availability                                                                             | 28 |
|        | References                                                                                    | 28 |

## 1. Introduction

Water quality (WQ) is essential for ecosystems, human health, and economic development. Typically, WQ is defined by its suitability for a specific use, which is determined by physical, chemical, and biological properties of water (EPA, 2007). However, disruptions in any WQ indicators can lead to various issues, for example, hypoxia caused by low dissolved oxygen (DOX) (Post et al., 2018); organic pollution indicated by high biological oxygen demand (BOD<sub>5</sub>) (Nafsin and Li, 2022); acidic aquatic systems due to high pH (Ascani et al., 2022; Boyd, 2015); and eutrophication caused by high nutrients (e.g., total phosphorus (TP), total nitrogen (TN), and dissolved inorganic nitrogen (DIN), etc.) (Siriwardana et al., 2024; van Wijk et al., 2024). These WQ indicators affect the suitability of water for ecosystems and human use (Gani et al., 2023). Over the past few decades, WQ degradation has been driven by both natural factors, such as climate change, and anthropogenic factors, including population growth, urbanization, and industrialization (Gani et al., 2023; Sajib et al., 2024; Uddin et al., 2021, 2023a, 2023b). Recently, UNESCO (2021) highlighted WQ as a critical issue for the 21st century because it has profound impacts on human health, food production, ecosystems, and economic growth. For example, in 2017, unsafe water caused over 1.1 million deaths, which was 2.2 % of global deaths (Ritchie and Roser, 2021). As a result, worldwide efforts are required for regular, timely, and sustainable WQ monitoring.

Several developed countries have implemented management and action strategies to maintain their “Good” WQ standards. Within this, the European Union (EU) Water Framework Directive (WFD; 2000/60/EC) has been established to standardize the assessment of all waterbodies across Europe (EU, 2019). Specifically, the WFD aim is to ensure good water for wildlife and humans by assessing the ecological and chemical status in all types of waterbodies (WFD, 2000). To achieve this, the WQ indicators under the WFD have been classified into several groups (EU, 2019). These groups include hydro-morphological elements (e.g., water flow dynamics, connection to aquifers); hydrodynamics (e.g., depth, velocity, residence time); physical components (e.g., TEMP, Colour, Taste, Odor, TUR, Solids, TRAN, EC); chemical components (e.g., pH, DOX, BOD<sub>5</sub>, COD, TOC, TSS); biological components (e.g., CHL, FC, TC, phytoplankton); and nutrient elements (e.g., TON, DIN, TN, MRP, TP) (Pattnaik et al., 2021; Diganta et al., 2024).

Furthermore, according to the WFD standards, the ecological status of surface waterbodies assesses the health of the overall ecosystem as determined by biological quality elements (e.g., fish, aquatic flora, macroinvertebrates, and phytoplankton) (EU, 2019). To maintain the quality of these biological components, it is essential to establish standard levels for the physico-chemical indicators (e.g., TEMP, DOX, BOD<sub>5</sub>, SAL, pH, NH<sub>4</sub><sup>+</sup>, TON, DIN, MRP, TP, and TRAN). Moreover, regular monitoring is crucial to ensure compliance with these standards (EPA,

2023). Notably, any disturbances in these physico-chemical indicators directly affect the biological quality properties (Samarinas et al., 2023; Uddin et al., 2021, 2023a). However, it should be noted that some of these WQ indicators are inherently interrelated (Mohseni et al., 2022) and their reciprocal relationship can significantly affect the overall WQ of an aquatic ecosystem. For example, high nutrient levels can cause eutrophication, which reduces water transparency. As a result, BOD<sub>5</sub> levels increase while DOX levels decrease, potentially leading to hypoxia (Ahmad et al., 2024; Gao et al., 2019). Furthermore, physical factors like TEMP and Cond can also affect chemical processes and the solubility of compounds in a waterbody (Mamun and An, 2021).

Generally, two types of techniques are widely used for assessing and monitoring WQ indicators: (1) in-situ techniques - WQ assessment and monitoring involves real-time data collection using sensors and probes placed directly in waterbodies, supplemented by laboratory analyses (Diganta et al., 2024; Uddin et al., 2021, 2023a, 2023b). These methods offer valuable insights into dynamic processes, which facilitate informed decision-making and effective management of water resources (Pollard et al., 2017); and (2) remote sensing (RS) techniques - provide valuable, non-invasive tools for assessing and monitoring WQ at various spatial and temporal scales (Sagan et al., 2020; Tao et al., 2025; Wilson et al., 2025). These methods contribute to a better understanding and management of aquatic ecosystems and resources (Ogashawara et al., 2017; Sajib et al., 2024; Uddin et al., 2023g). While both in-situ and RS techniques have unique benefits, but they also have distinct constraints. For instance, in-situ approaches provide accurate, localized measurements of WQ indicators, but they can be time-consuming, labour-intensive, and have limited spatial coverage (Agarwal et al., 2018; Cao et al., 2023; Mohseni et al., 2022). In contrast, RS techniques offer broad spatial and temporal coverage with reduced fieldwork requirements, enabling continuous monitoring of large waterbodies, such as inland, coastal, and transitional (He et al., 2008; Matthews, 2011; Soomets et al., 2022; Sun et al., 2022). However, effects such as atmospheric disturbance, sun-glitter, adjacency effects, etc., significantly affect retrieval model development using RS technology in inland, coastal and open waters (Ansper-Toomsalu et al., 2024; Concha et al., 2021; Frouin et al., 2019; Gleratti et al., 2024; González Vilas et al., 2023; Kutser, 2012; Pahlevan et al., 2021a, 2021b; Vanhellemont and Ruddick, 2018; Zibordi et al., 2022). Additionally, hydrodynamic factors, particularly surface waves and tides further complicate coastal and transitional and inland river-based WQ retrieval model development (Pan and Lou, 2023; Wei et al., 2016). Moreover, recent studies have highlighted a range of limitations of RS techniques, including image acquisition error (s), impacts of various atmospheric factors, cloud coverage, image processing error(s), challenges in capturing fine-scale variations, etc. (Adjovu et al., 2023; Gholizadeh et al., 2016; Sagan et al., 2020; Wang and Yang, 2019; Yang et al., 2022b). Despite these challenges, recent

research has suggested that combining in-situ and RS methods may improve understanding of RS attribute interactions for effective WQ management (Arabi et al., 2020; Diganta et al., 2024; Sajib et al., 2024; Sheffield et al., 2018). As a result, despite their limitations, RS techniques have gained much more attention from scientific communities, including the various national and international environmental organizations, in the field of WQ model development (Gholizadeh et al.,

2016; Sagan et al., 2020; Wang and Yang, 2019; Yang et al., 2022b).

Typically, WQ assessment using RS involves optically active (OA – e.g., CHL, TSM, CDOM, etc.) and optically inactive (OI – e.g., DO, BOD<sub>5</sub>, DIN, TP, TN, etc.) WQ indicators, each presenting unique advantages and challenges (Adjovu et al., 2023; Gholizadeh et al., 2016; He et al., 2008). For instance, OA indicators, such as CHL, TSM, and CDOM, can be retrieved from the RS data due to the inherent capabilities of this

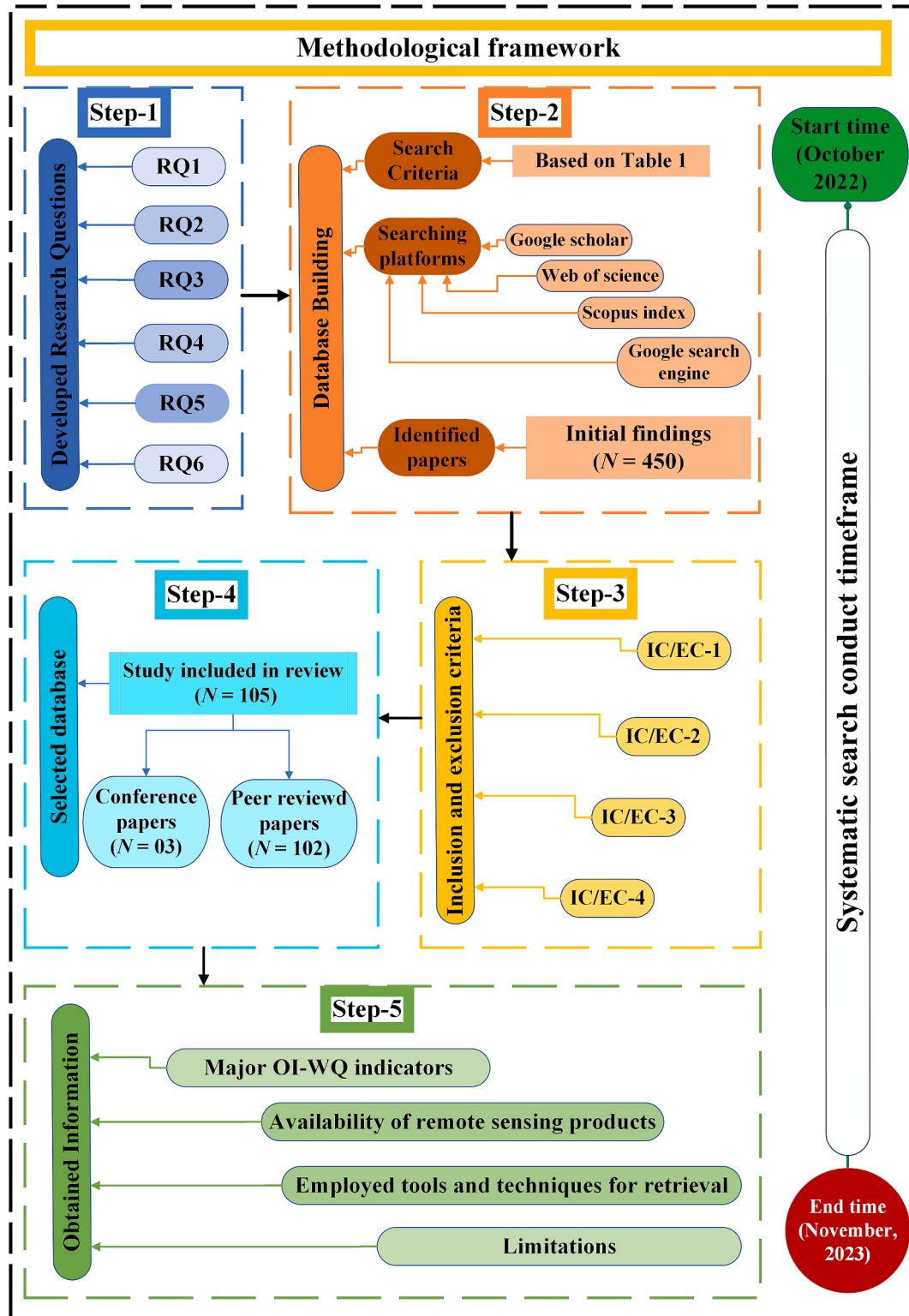


Fig. 1. Methodological framework of the research.

technology (Diganta et al., 2024; Frouin et al., 2019; Gleratti et al., 2024; González Vilas et al., 2023; Kutser, 2012; Pahlevan et al., 2021a; Zhang et al., 2021a). In particular, the spectral properties of water around 440, 675 and 700 nm are highly sensitive to CHL concentration (Gilerson et al., 2010; Diganta et al., 2024; Pahlevan et al., 2021b), whereas the NIR regions are highly sensitive to the concentration of TSM (Du et al., 2020; Wen et al., 2022; Odermatt et al., 2010). On the other hand, CDOM is characterized by the absorption coefficient of solar radiation at 440 nm (Dall'Olmo et al., 2017; Morel and Gentili, 2009; Hooker et al., 2020; Houskeeper et al., 2021). However, recent studies have identified several challenges in retrieving OA-WQ retrieval algorithms, including atmospheric conditions (Ansper-Toomsalu et al., 2024; Concha et al., 2021; Frouin et al., 2019; Gao et al., 2006; Gleratti et al., 2024; González Vilas et al., 2023; Kutser, 2012; Pahlevan et al., 2021a; Vanhellemont and Ruddick, 2018; Zibordi et al., 2022), surface reflection (IOCCG, 2010; Wei et al., 2016; Scheirer et al., 2018), adjacency effect (Ansper-Toomsalu et al., 2024; Tao and Hill, 2019; Tessin et al., 2024), and sensor characteristics (Gholizadeh et al., 2016; Chawla et al., 2020), etc. To address these issues, particularly the atmospheric problems, several atmospheric correction (AC) algorithms have been developed, such as ACOLITE (Vanhellemont and Ruddick, 2018), POLYMER (Steinmetz et al., 2011), C2RCC (Brockmann et al., 2016), 6 s (Kotchenova et al., 2006), Sen2Cor (Louis et al., 2016), etc. (Li et al., 2023b; Li et al., 2023c). Additionally, numerous studies have also emphasized the need for extensive in-situ data to accurately retrieve OA-WQ indicators from optically complex waters (Chen et al., 2022; Diganta et al., 2024; Ogashawara et al., 2017; Sajib et al., 2024). A detailed description of the progress and challenges in retrieving OA-WQ indicators from satellite data can be found in Cao et al. (2023), Chawla et al. (2020), Diganta et al. (2024), Dey and Vijay (2021), Dörnhöfer and Oppelt (2016), Yang et al. (2022a), and Yang et al. (2022b).

On the other hand, OI-WQ indicators also face various challenges, including complex mathematical models for indirect estimation, which require large amounts of ancillary data to validate the models (Gholizadeh et al., 2016; Niu et al., 2021; Sagan et al., 2020; Vakili and Amanollahi, 2020). Additionally, OI-WQ indicator models heavily depend on the input data's accuracy, calibration, and validation with in-situ measurements (Chen et al., 2022; Gholizadeh et al., 2016; Sajib et al., 2024). Moreover, numerous retrieval models, such as empirical models (EM) (Portela et al., 2024), semi-empirical models (SEM) (Chen et al., 2022; Goyens and Ruddick, 2023; Mobley, 2001; Mobley et al., 2004), and machine learning (ML)/artificial intelligence (AI)-based models (Chen et al., 2022; Han et al., 2023) have been developed for OA-WQ indicators. However, OI-WQ-based models (e.g., DOX model by Liu et al., 2022b; Salas et al., 2022; Sharaf El Din et al., 2017) have faced much criticism due to the considerable uncertainty in the final retrieval results (Sajib et al., 2024). A detailed description of uncertainty associated with retrieval algorithms can be found in Chen et al. (2022), Yang et al. (2022a), and Yang et al. (2022b).

To the best of the authors' knowledge, while there have been a considerable number of studies focusing on monitoring OA-WQ indicators (Chen et al., 2022; Gholizadeh et al., 2016; Diganta et al., 2024; Sagan et al., 2020; Wang and Yang, 2019; Yang et al., 2022b), but research on the OI-WQ remains limited. Therefore, the aim of the research was to identify the suitable RS data and retrieval methods for retrieving the OI-WQ indicators from different waterbodies by analysing the literature. To achieve the goal of the research, the following objectives were considered:

- To review and evaluate the existing RS data and retrieval methods for retrieving OI-WQ indicators.
- To identify the key spectral bands and indices that are widely used and most effective in retrieving OI-WQ indicators from RS data.
- To assess the accuracy and reliability of different RS data and retrieval methods for OI-WQ indicators by comparing the literature.

- To compare the performance of different retrieval methods, including EM, SEM, and ML/AI-based methods, by conducting a critical analysis of literature.
- To determine the limitations and potential avenues for further improvement of current RS data and retrieval methods for monitoring OI-WQ indicators.
- To provide recommendations for the selection of appropriate RS data, band(s), and combination of band(s) for the development of effective retrieval model(s) in order to retrieve various OI-WQ indicators in terms of certain environmental conditions.

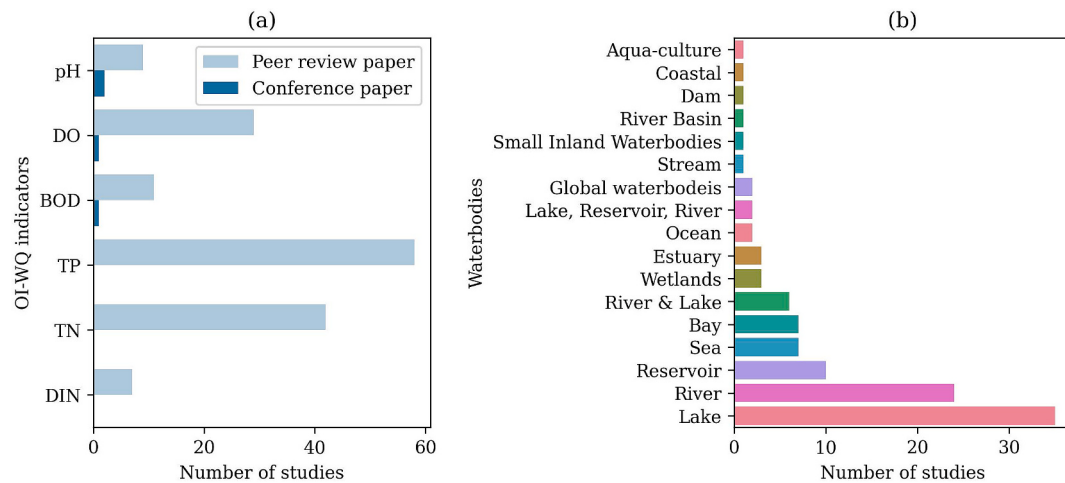
The paper consists of seven sections. Following the introduction, Section 2 explains the techniques utilized in the review process. Section 3 not only describes the parameters and satellite data used in OI-WQ estimation but also addresses different AC methods. Section 4 addresses RS-based EM, SEM, and ML/AI retrieval models and limitations of those retrieval models. Section 5 provides a brief description of the RS data, parameters, algorithms, and optimization process utilized to retrieve OI-WQ indicators. This section also addresses the existing limitations in the retrieval of OI-WQ indicators. On the other hand, section 6 discussed the limitations and suggested future scope for OI-WQ research. Finally, a conclusion is presented in Section 7.

## 2. Method of review

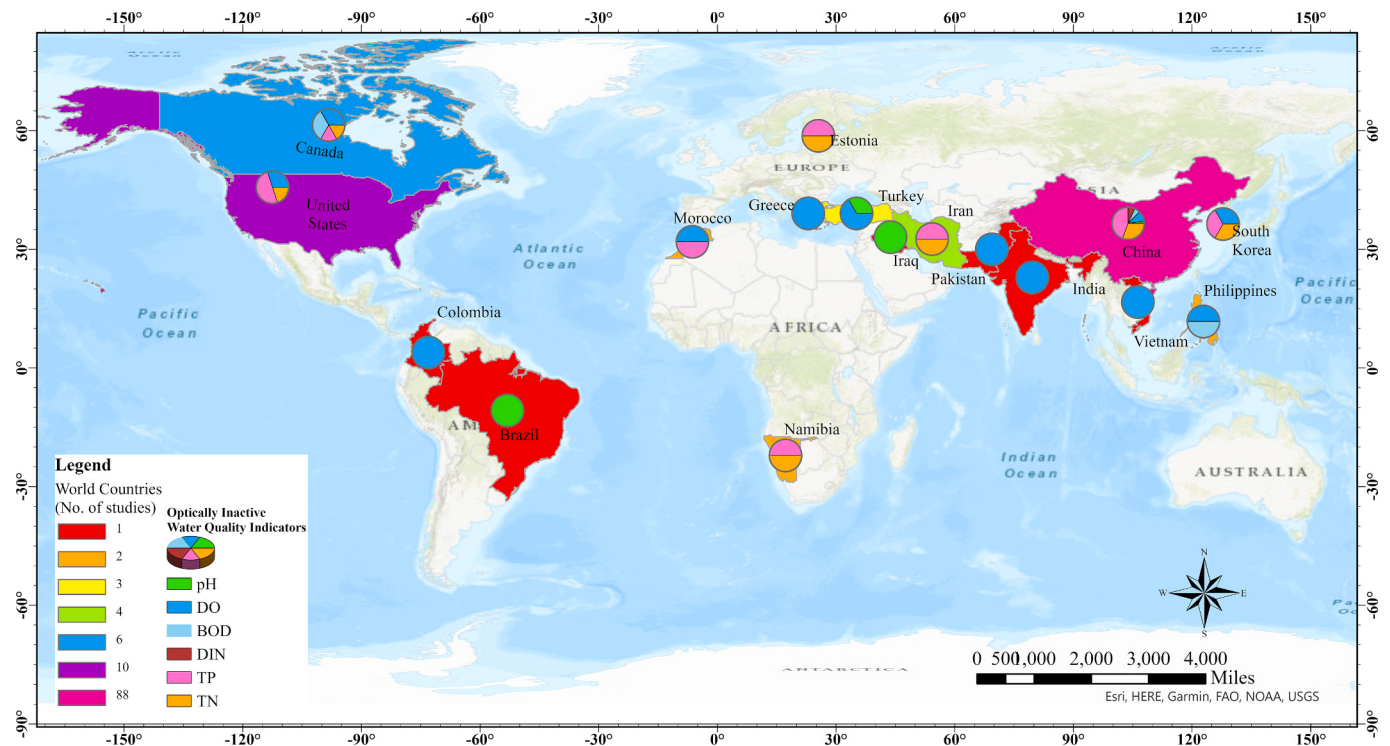
There are several established guidelines for conducting literature reviews (Snyder, 2019) and researchers commonly utilized these methodologies based on their review goals. The current study was conducted using a systematic literature review process (Fig. 1). A systematic literature review differs from traditional narrative reviews in that it uses a well-defined search strategy and inclusion/exclusion process to identify, select, and critically evaluate existing literature (Rousso et al., 2020). The advantage of utilizing is that it reduces uncertainty at every step of the review process by using an organized and clear approach (Mengist et al., 2020). Details of this methodology can be found in Ozdemir et al. (2023) and Rousso et al. (2020). However, systematic search in this study was challenging because most published studies of OI-WQ did not mention "non-OA/OI-WQ" in their publications. On the other hand, several studies have focused on both OA- and OI-WQ indicators (Arias-Rodriguez et al., 2023; Tian et al., 2023; Zhang et al., 2020). Therefore, the present study reviewed both categories of RS-based WQ indicator monitoring studies to identify the focused WQ indicators.

**Table 1**  
Keywords used for articles and conference proceedings.

| Keyword category                        | Search query                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-----------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RS-based OI-WQ indicators               | <ul style="list-style-type: none"> <li>• Water AND quality AND using AND remote AND sensing</li> <li>• Optically AND active AND water AND quality AND monitoring</li> <li>• Optically AND inactive AND water AND quality AND monitoring</li> <li>• Non-optically AND active AND water AND quality AND monitoring</li> <li>• Water AND quality AND retrieval using AND Machine AND learning AND RS</li> <li>• Water AND quality AND retrieval AND using AND deep AND learning AND RS</li> </ul> |
| ML/AI based OI-WQ indicators prediction | <ul style="list-style-type: none"> <li>• pH AND prediction AND in AND water</li> <li>• DO AND prediction AND in AND water</li> <li>• BOD<sub>5</sub> AND prediction AND in AND water</li> <li>• TP AND prediction AND in AND water</li> <li>• DIN AND prediction AND in AND water</li> <li>• TN AND prediction AND in AND water</li> </ul>                                                                                                                                                     |



**Fig. 2.** (a) Number of published papers on OI-WQ indicators in peer-reviewed papers and conference proceedings; (b) Number of studies conducted on different waterbodies.



**Fig. 3.** Published research on OI-WQ indicators in different waterbodies around the world.

## 2.1. Research questions

The study critically reviews the current global trends and research progress in the field of OI-WQ indicators, guided by the following research questions (RQs). This RQ determines the scope of the research.

RQ1: What are the basic principles of retrieving OI-WQ from different waterbodies?

RQ2: What are the tools and techniques commonly used for retrieving OI-WQ from various waterbodies?

RQ3: What types of RS data are widely utilized to retrieve OI-WQ indicators from different types of waterbodies, such as lakes, rivers, estuaries, bays, and seas?

RQ4: What parameters (e.g., OA-WQ indicators, spectral bands, indices, etc.) are typically utilized in the retrieval process?

RQ5: What types of algorithms and optimization methods, including

EM, SEM, and ML/AI techniques, are utilized in the retrieval process of OI-WQ indicators?

RQ6: What is the accuracy and reliability of different RS data and retrieval methods for OI-WQ indicators?

RQ7: What are the limitations and potential for improvement to current RS data and retrieval methods for OI-WQ indicators?

## 2.2. Selection criteria for reviewed OI-WQ indicators

The management of WQ has received significant attention due to growing environmental concerns and inadequate freshwater resources (Sajib et al., 2023). As a result, this aspect has directed various stakeholders, including researchers, organizations, and nations, to develop cost-effective methods for managing water resources (Mustafa et al., 2021; Uddin et al., 2022a, 2023c, 2023d, 2023e). Although several

policies and regulations have been established to monitor the WQ in various waterbodies (Diganta et al., 2024; Uddin et al., 2021), these programs often rely on conventional monitoring systems, which are time-consuming and require considerable manpower (Uddin et al., 2021). Furthermore, these measures often fail to provide a concise scenario of the current state of the waterbodies (Uddin et al., 2022a). To address this issue, different stakeholders utilize various approaches; the water quality index (WQI) model is one of them, which can provide a concise view of the waterbody by incorporating a range of WQ indicators; and a well-known methodology in the field of water resource management (Uddin et al., 2022a, 2022b, 2023f, 2024). For instance, the Irish Water Quality Index (IEWQI) model serves as a notable example in order to assess and monitor the TrC waters (Uddin et al., 2023e). Moreover, the WFD is another fundamental approach to maintain the “Good” WQ condition within the EU extent. As the aim of the research, the study focuses on the WFD regulations for selecting the OI-WQ variables that are guided by the framework to consider the regular monitoring programme in EU member states. The framework which outlined 11 WQ indicators, including five OI-WQ indicators like DOX, BOD<sub>5</sub>, pH, DIN, TP, and TN for monitoring the WQ in the EU waterbodies (WFD, 2000). A detailed description of these WQ indicators and their standard values can be found in Table S2. Therefore, the study considered these OI-WQ indicators for a comprehensive analysis of their existing tools and techniques.

### 2.3. Literature search

In order to achieve the primary research objective, the current study utilized a systematic search on the Google Search Engine, Google Scholar, Web of Science, and Scopus Index with the keywords listed in Table 1. The search method is implemented according to methodology provided in the studies by Ozdemir et al. (2023) and Yang et al. (2022b). Moreover, the study considered articles published from 2006 to late 2023. This selected timeframe allows for the examination of techniques utilized in OI-WQ retrieval, ranging from traditional EM to recent ML/AI and deep learning (DL) techniques and widely utilized RS data. By following this rigorous search and selection process, this study provides a comprehensive overview of the current state-of-the-art techniques and advances in RS-based OI-WQ retrieval and prediction.

### 2.4. Inclusion and exclusion criteria

The following inclusions and exclusions (IC/EC) were utilized in the review process to identify the relevant studies.

- IC/EC 1: The study must be related to the OI-WQ (e.g., pH, DOX, BOD<sub>5</sub>, TP, TN, and DIN) assessment and monitoring.
- IC/EC 2: The study must include various types of approaches (e.g., EM, SEM, ML/AI, and DL) that use RS data to develop algorithms for pH, DOX, BOD<sub>5</sub>, TP, TN, and DIN.
- IC/EC 3: The study must include different types of sensors (e.g., ground, UAV and space-borne) used for monitoring pH, DOX, BOD<sub>5</sub>, TP, TN, and DIN in different waterbodies.
- IC/EC 4: Preprints or early versions of articles are excluded from the review process.

### 2.5. Conducting the review

The systematic search was conducted between October 2022 and November 2023. In the first phase, a total of 450 papers were found based on search criteria. After removing duplicates and applying inclusion and exclusion criteria, a total of 105 papers are selected, including 102 peer-reviewed papers and 3 full-length conference proceedings. The details of individual numbers of identified OI-WQ indicator studies are shown in Fig. 2(a), and tested waterbodies are illustrated in Fig. 2(b), while Fig. 3 shows the country-wise published

studies.

### 2.6. Obtained information

In order to achieve the research goal, the review encompassed several key attributes relevant to OI-WQ indicator retrieval. These attributes are summarized as follows:

#### (i) Major OI-WQ indicators:

The review focused on identifying and analysing the primary indicators used for OI-WQ assessment. This involved examining the commonly studied parameters and characteristics that are indicative of WQ, such as DOX, pH, and BOD<sub>5</sub>.

#### (ii) Availability of RS data:

The review investigated the existing RS datasets that are commonly utilized for OI-WQ retrieval. This involved assessing the availability and accessibility of satellite imagery, aerial photographs, and other RS-derived data sources that enable the monitoring of WQ indicators.

#### (iii) Tools and techniques:

The review encompassed an exploration of the various tools and techniques employed for OI-WQ retrieval. This included:

- Statistical approaches: The review examined statistical methods, such as regression analysis and multivariate analysis, that have been utilized to establish relationships between RS data and OI-WQ indicators.
- ML/AI techniques: The review investigated the application of ML/AI techniques, such as SVM, NN, and RF algorithms for OI-WQ retrieval.
- Others: In addition to statistical and ML/AI approaches, the review considered other methods and techniques that have been employed, such as fuzzy logic, expert systems, and hybrid approaches combining multiple methodologies.
- Comparison of various tools and techniques: The review included a comparative analysis of the different tools and techniques used for OI-WQ retrieval. This involved assessing their strengths, limitations, and performance in different scenarios.

#### (iv) Limitations:

The review acknowledged and discussed the limitations associated with OI-WQ retrieval. This encompassed considerations such as uncertainties in RS data, limitations of specific tools and techniques, and challenges in accurately interpreting and validating the retrieved WQ information.

By comprehensively addressing these attributes, the review provides valuable insights into the list of major OI-WQ indicators, the availability of RS data, and the range of tools and techniques employed for OI-WQ retrieval. It also highlights the limitations that need to be considered when interpreting the results of such retrieval models.

## 3. RS data and atmospheric corrections

In contemporary times, the implementation of a free and open data policy has significantly enhanced access to vast amounts of RS data (Gholizadeh et al., 2016; Sajib et al., 2024). Additionally, combined with advanced cloud computing services, this has greatly accelerated the analysis of extensive time series data (Arias-Rodriguez et al., 2023; Doxani et al., 2018). Specifically, spaceborne, airborne, and modern ground-based sensors are utilized in the RS process to estimate the radiation frequency across various wavelengths (e.g., visible, infrared, and microwave) reflected from the water surface to assess WQ (Diganta et al., 2024; Wagle et al., 2020). For instance, the optimal spectral range for total suspended matter (TSM) concentration lies between 580 nm - 680 nm and 700 nm - 900 nm (Mohseni et al., 2022). Similarly, in the microwave domain, the ocean surface TEMP and SAL can be estimated by microwave radiometers and synthetic aperture radars (Guo et al., 2022a). The most commonly used RS data for WQ retrieval are the Landsat series (e.g., Landsat-5 TM, Landsat-7 ETM+, Landsat-8 OLI, and

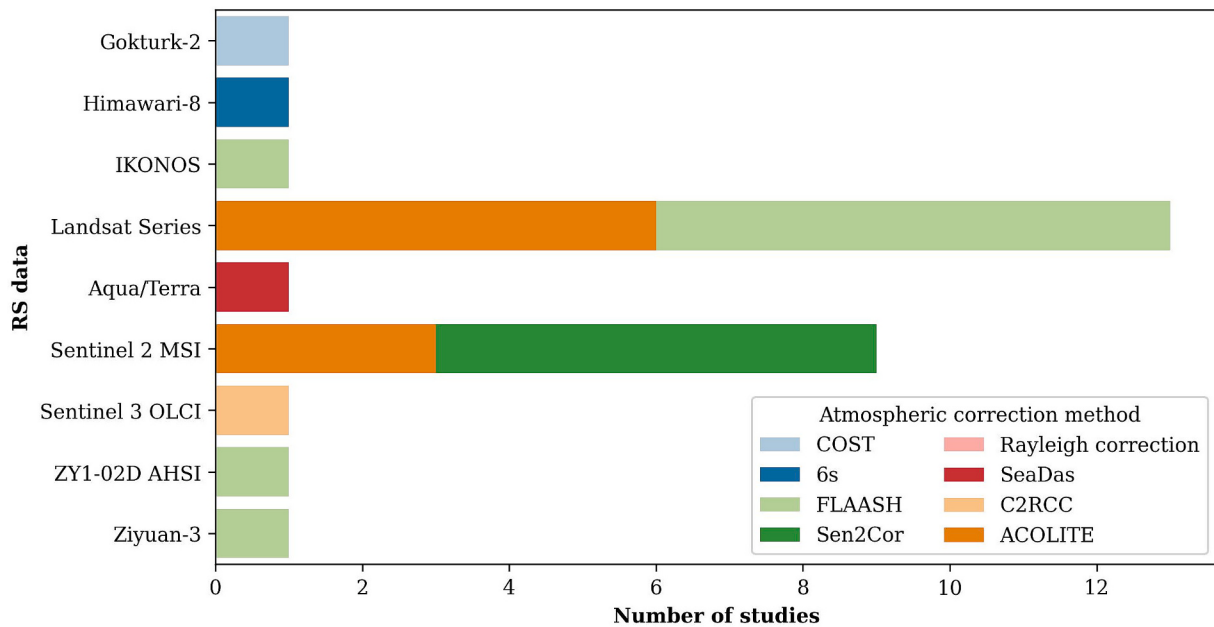


Fig. 4. Atmospheric correction methods utilized in the OI-WQ indicator retrieval models.

Landsat-9 OLI), the Copernicus programme imagery (e.g., Sentinel-2 MSI, Sentinel-3 OLCI), the MODIS (Aqua/Terra), and the VIIRS (Chen et al., 2022; Llodrà-Llabrés et al., 2023; Matthews, 2011; Portela et al., 2024; Radeloff et al., 2024; Sagan et al., 2020; Samarinas et al., 2023; Wagle et al., 2020; Wang and Yang, 2019; Yang et al., 2022a; Yang et al., 2022b). Moreover, unmanned aerial vehicles (UAV) and ground-borne (proximal and portable) high-resolution hyperspectral data have gained popularity in recent years (Gholizadeh et al., 2016; Diganta et al., 2024; Wasehun et al., 2024). However, since hyperspectral data collection and processing is quite complex, multispectral satellite data are preferable for large-scale waterbody monitoring (Li et al., 2022c; Raghul and Porchelvan, 2024). A detailed description of frequently used satellites for WQ retrieval can be found in Table S3 and Table S4.

Over the past four decades, notable progress has been achieved in RS-based WQ retrieval through advances in sensors, image correction processes, and retrieval algorithms (Ansper-Toomsalu et al., 2024; Concha et al., 2021; Guo et al., 2022a). However, numerous challenges complicate the use of RS data for WQ monitoring (Gleratti et al., 2024; González Vilas et al., 2023; Kutser, 2012; Pahlevan et al., 2021a; Vanhellemont and Ruddick, 2018; Zibordi et al., 2022). Particularly, water reflectance observations conducted from space are subject to various interfering processes that arise as a result of the two conditions within the atmosphere-surface system; condition one encompasses various phenomena such as particles (e.g., soot), gaseous absorption, molecular scattering, aerosol scattering and absorption, and water reflection in a clear sky; on the other hand, condition two includes scattering, specifically scattering caused by cloud droplets (Frouin et al., 2019). To accurately retrieve WQ, it is essential to eliminate the impact of the atmosphere and surface from RS data (Diganta et al., 2024; Kim et al., 2023; Kutser, 2012; Sajib et al., 2024). This process is commonly known as atmospheric correction (AC) in scientific literature. The implementation of AC is deemed essential for RS retrieval applications, although it may not always be necessary for RS classification applications (Zhu and Xia, 2023). The conventional method for AC involves the assumption of a “black pixel” (Song et al., 2023; Wang et al., 2023a). Typically, there are two ways to correct this issue: firstly, the estimation of aerosol/surface reflectance is performed over totally absorbing water in the red and NIR regions, and secondly, applying extrapolation techniques to determine the aerosol/surface reflectance at shorter wavelengths (Frouin et al., 2019). Alternatively, another method for AC

involves simultaneously determining the aerosols and water constituent's properties by comparing the top-of-atmosphere (TOA) reflectance with the output of a radiative transfer model (Shi and Nakajima, 2018). Consequently, several methods for AC have been developed and extensively utilized in various RS applications (Ansper-Toomsalu et al., 2024; Concha et al., 2021; Frouin et al., 2019; Gleratti et al., 2024; González Vilas et al., 2023; Kutser, 2012; Pahlevan et al., 2021a; Vanhellemont and Ruddick, 2018; Zibordi et al., 2022).

The AC methods commonly utilized include, Sen2Cor, C2RCC, POLYMER, iCOR, ACOLITE, FLAASH, 6s, DOS, COST, and L2gen (Ansper-Toomsalu et al., 2024; Concha et al., 2021; Frouin et al., 2019; Gleratti et al., 2024; González Vilas et al., 2023; Kutser, 2012; Pahlevan et al., 2021a; Vanhellemont and Ruddick, 2018; Zibordi et al., 2022). The specific characteristics and properties of these AC methods can be found in Li et al. (2023b) and Diganta et al. (2024). Fig. 4 illustrates the various AC methods utilized in different research studies for the OI-WQ retrieval across diverse waterbodies. Notably, the FLAASH AC method is commonly utilized for the Landsat series (Cruz-Montes et al., 2023; Fu et al., 2022; Kapalanga et al., 2021; Krishnaraj and Honnasiddaiah, 2022; Mohandas and Brema, 2023), whereas the Sen2Cor method is considered the most suitable for AC in the context of Sentinel 2 A/2B MSI missions (Guo et al., 2021a; Mohandas and Brema, 2023; Salas et al., 2022). However, specific algorithms, such as C2RCC, ACOLITE, POLYMER, and L2gen designed for water applications (Ansper-Toomsalu et al., 2024; Diganta et al., 2024; Vanhellemont and Ruddick, 2018), while Sen2Cor and 6s are specifically tailored for land applications (Li et al., 2023b; Li et al., 2023c). Furthermore, numerous studies have employed these algorithms to identify the most appropriate AC method for quantifying WQ in various regions (Ansper-Toomsalu et al., 2024; Concha et al., 2021; Frouin et al., 2019; Gleratti et al., 2024; González Vilas et al., 2023; Kutser, 2012; Li et al., 2023b; Li et al., 2023c; Pahlevan et al., 2021a; Vanhellemont and Ruddick, 2018; Zibordi et al., 2022). Nevertheless, numerous researchers have also highlighted that all these AC methods have certain limitations (Arená et al., 2024; Diganta et al., 2024; Soppa et al., 2021; Warren et al., 2019), and the outcome of AC varies in terms of waterbodies.

#### 4. RS based retrieval methods for OI-WQ indicators

The current study focused on the OI-WQ indicators and techniques

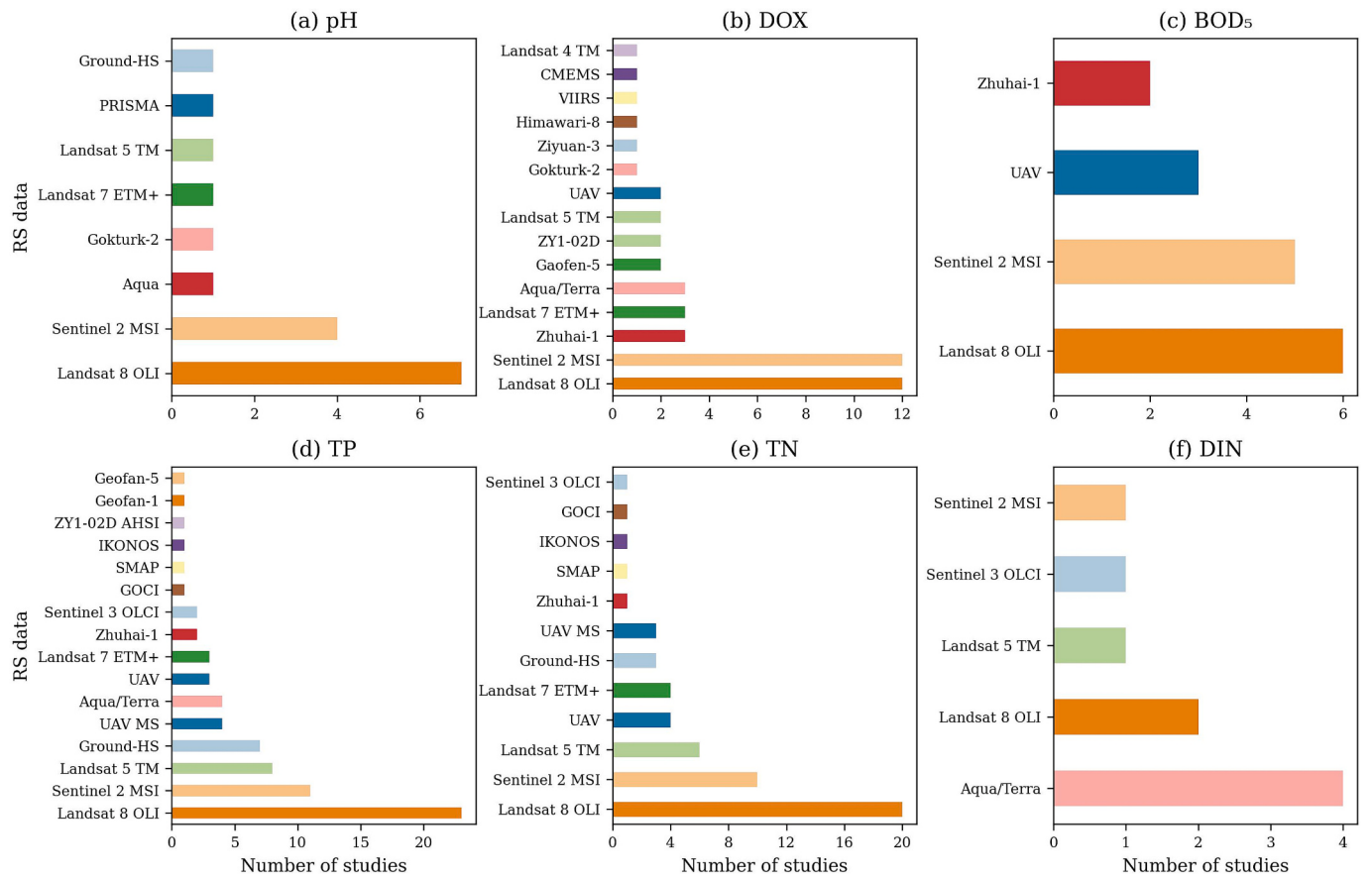


Fig. 5. Frequently used RS data in OI-WQ indicator retrieval models. (here, HS = Hyperspectral; UAV = Unmanned Aerial Vehicle; MS = Multi Spectral).

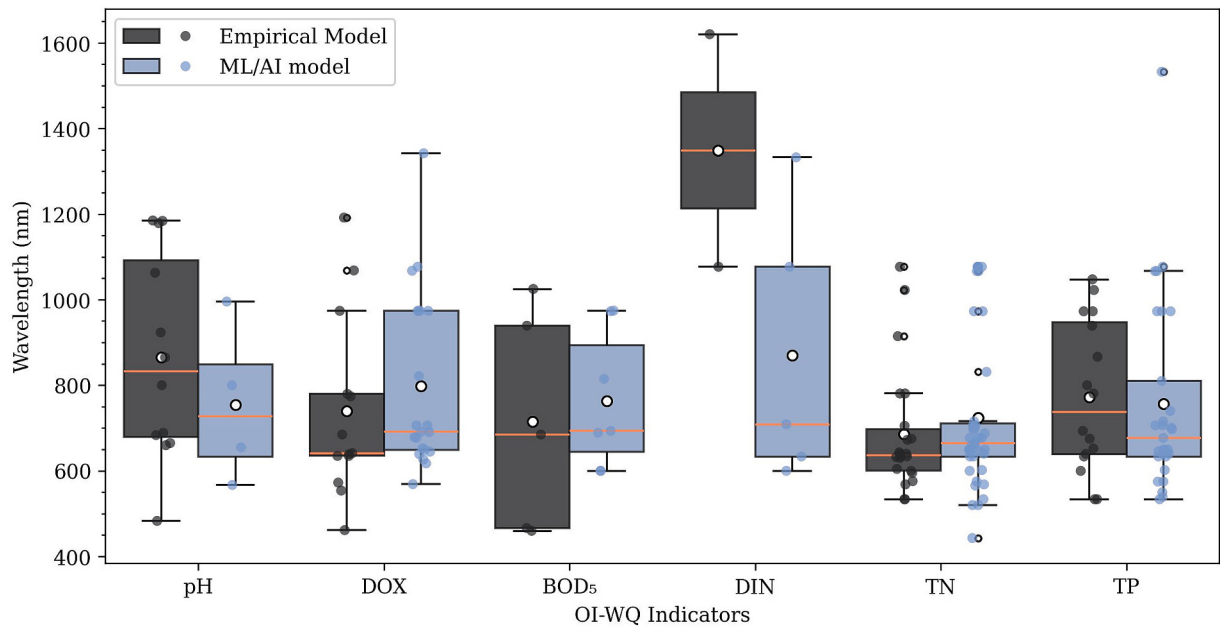
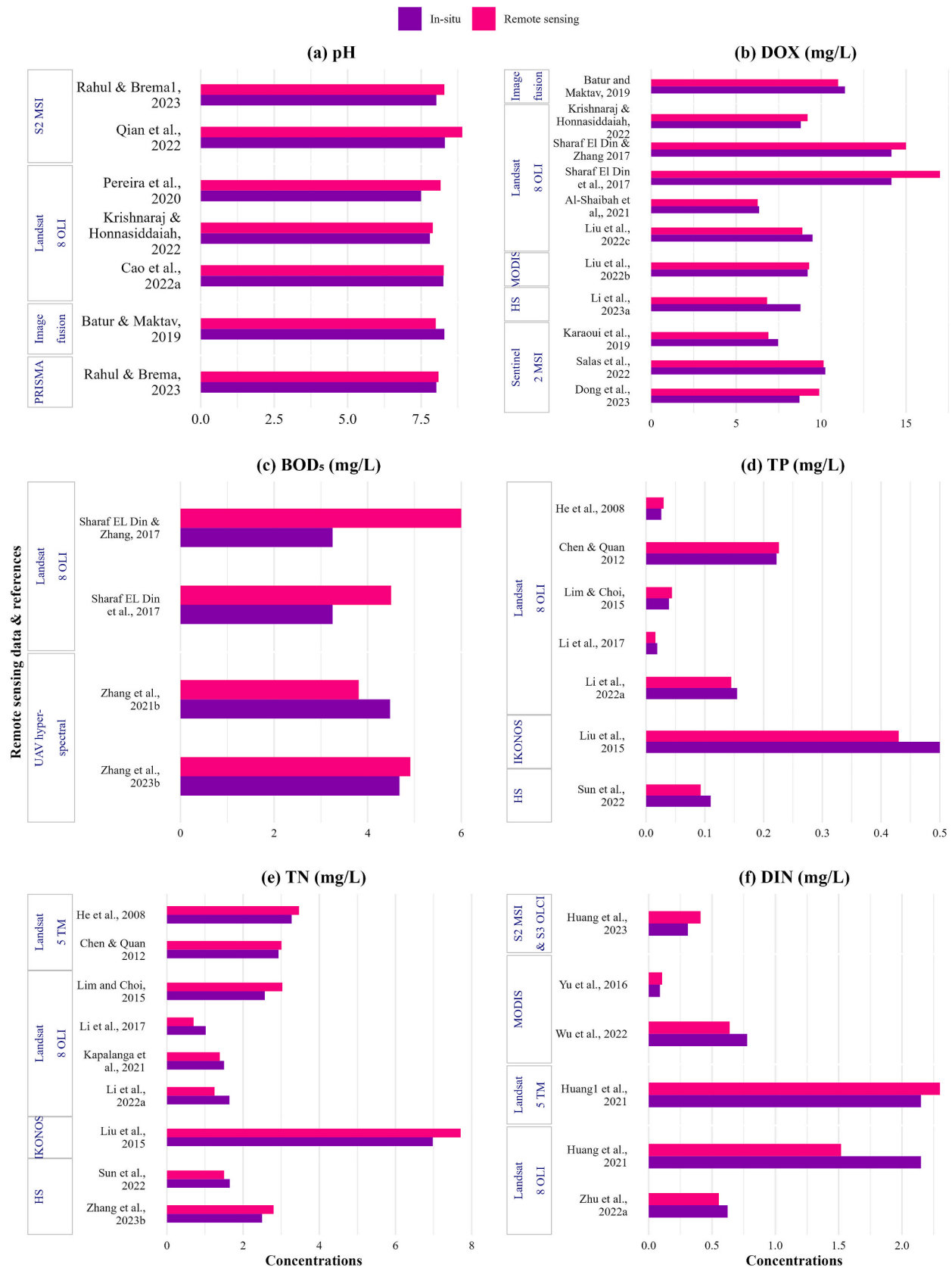


Fig. 6. Boxplot of the RS data wavelength used in different OI-WQ indicator retrieval processes.

available for retrieving these indicators using RS data. Recent advancements in RS technology have notably enhanced the ability to estimate OI-WQ indicators using various methods, such as EM, SEM and ML/AI-based models (Chen et al., 2022; Sagan et al., 2020; Wagle et al., 2020; Wang and Yang, 2019; Yang et al., 2022a; Yang et al., 2022b; Zhu and Xia, 2023). These approaches are widely utilized to

assess spatial and temporal variations in WQ indicators (Chen et al., 2022; Sagan et al., 2020; Wagle et al., 2020; Wang and Yang, 2019; Yang et al., 2022a; Yang et al., 2022b; Zhu and Xia, 2023). Furthermore, researchers have recently made notable progress in improving algorithms for estimating OI-WQ using RS technology (Diganta et al., 2024; Sajib



**Fig. 7.** A comparison between in-situ and RS-retrieved OI-WQ concentrations in different studies. (here, S2 MSI = Sentinel 2 MSI, S3 OLCI = Sentinel 3 OLCI).

et al., 2024); however, despite these advancements, most of the developed models still struggle to achieve high accuracy when applied across different regions (Chen et al., 2022; Diganta et al., 2024; Sagan et al., 2020).

Typically, the EM is a correlation-based regression analysis between in-situ WQ indicators and the corresponding RS band value (Shang et al., 2021). For instance, common EM methods include linear regression (such as stepwise linear, multiple linear, quadratic polynomial, power curvilinear, etc.), single-band approach, band-combination method (such as band ratio and band difference), principal component analysis, and so on (Wang and Yang, 2019). However, several researchers have questioned the applicability of EM due to its high uncertainty and reliance on extensive in-situ data (Aladejare and Idris, 2020; Kumar et al., 2024; Panchanathan et al., 2023). Nevertheless, one of the key advantages of EM techniques is their user-friendliness (Ding et al., 2020; Shang et al., 2021; Sun et al., 2022; Wu et al., 2010; Zhang et al., 2022a).

Moreover, the application of hyperspectral RS technology has accelerated the development of SEM (Wang and Yang, 2019), which is based on optical physics (inherent optical properties (IOPs) and apparent optical properties (AOPs)) with traditional statistical methods (Cao et al., 2023; Chen et al., 2022; Shin et al., 2020; Sagan et al., 2020). On the one hand, SEM is primarily applicable to local contexts (for most optically complex Case 2 waters) due to their high sensitivity to changes in water composition and concentration (Lednicka and Kubacka, 2022). On the other hand, compared to other models, these models offer high performance with minimal computational costs (Han et al., 2022). Additionally, to date, the most widely used RS estimation model is based on cutting-edge ML/AI technology (Cao et al., 2022a; Cao et al., 2021; Karimi et al., 2023; Zhang et al., 2021b). By combining big data technology and high-performance computing, ML/AI techniques not only enhance the ability to develop a relationship between WQ and RS reflectance (Xiong et al., 2022) but also significantly reduce time and cost while generating accurate results for large datasets (Wagle et al., 2020). A detailed description of ML/AI and RS-based estimation methods for WQ indicators can be found in Chen et al. (2022), Yang et al. (2022b) and Wagle et al. (2020).

## 5. OI-WQ Indicators

### 5.1. pH

The pH of water plays a critical role in the aquatic environment, as it indicates whether the water is acidic or alkaline, ranging from zero (very acidic) to fourteen (very alkaline) (EPA, 2001). A detailed description of its significance as a WQ indicator and its traditional measurement method can be found in the supplementary materials as a continuation of Section 5.1. Beside traditional methods, ML/AI is widely utilized for water pH prediction in different waterbodies (DeSimone et al., 2020; Flecha et al., 2022; Fu et al., 2021; Hu et al., 2019; Son et al., 2021; Stackelberg et al., 2021). However, a key limitation of these approaches is their reliance on in-situ measurements, which fail to capture the spatio-temporal distribution of pH in different waterbodies. To address this challenge, researchers have recently integrated RS data with ML/AI and statistical techniques to predict and estimate water pH more effectively (Abbas et al., 2021; Batur and Maktav, 2019; Jiang et al., 2022; Nakano and Watanabe, 2005; Pereira et al., 2020; Sabia et al., 2015).

#### 5.1.1. RS data and bands (wavelength) used for pH retrieval models

To date, several RS data have been utilized for developing pH retrieval models. Notably, the Landsat 8 OLI stands out as the most widely utilized satellite (Abbas et al., 2021; Cruz-Montes et al., 2023). Additionally, medium-resolution satellite (10 m - 250 m) data such as the Landsat series (Landsat 5 TM, Landsat 7 ETM+), the Sentinel 2A MSI, and the MODIS-Aqua/Terra are commonly utilized for assessing pH in different waterbodies (Fig. 5a). Furthermore, high-resolution (30 cm - 5 m) satellite images from the Gokturk-2 and hyperspectral data are also used in pH retrieval models (Batur and Maktav, 2019). Moreover, single or multi-parameter-based satellite images, which are the combination of in-situ and RS data such as the SMOS-L3, the L4-OSTIA, the ESA-OceanFlux-GHG project climatology (Sabia et al., 2015), WOA (Nakano and Watanabe, 2005), etc., are used in pH retrieval models. Additionally, satellite-derived indicators, such as CHL, DOX, TEMP, ALK, pCO<sub>2</sub>, and MLD, have been used to develop RS-based pH algorithms (Batur and Maktav, 2019; Jiang et al., 2022; Nakano and Watanabe, 2005; Sabia et al., 2015). In parallel, different indices, such as

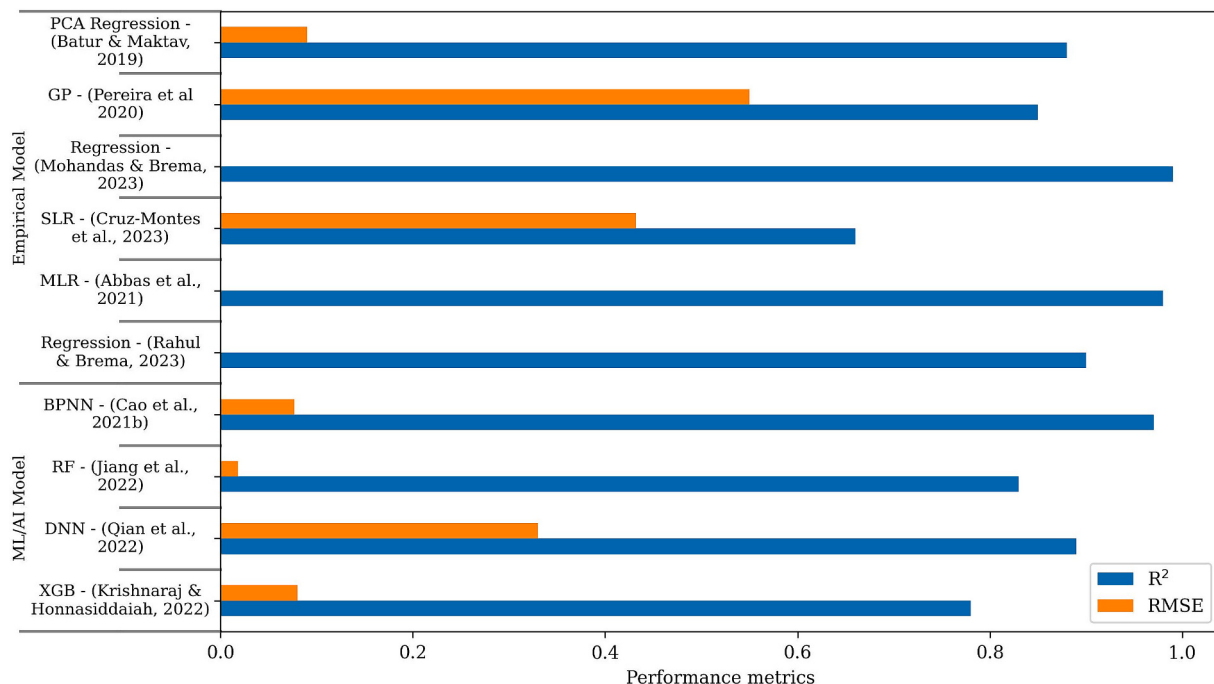


Fig. 8. Performance evaluation results of different pH retrieval models.

**Table 2**  
Application of regression-based pH retrieval models.

| RS data type   | Waterbody type | Domain   | Equation                                                                                                                                                                                                                  | R <sup>2</sup> | Reference                 |
|----------------|----------------|----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|---------------------------|
| Landsat 8 OLI  | Wetlands       | Columbia | $pH = 10.454 + 18.352(B_4) + 3.1284 \left( \frac{B_4 \cdot B_5}{B_4 + B_5} \right) + 15.115 \left( \frac{B_6}{B_2} \right) + 10.2 \left( \frac{B_4 \cdot B_2}{B_3} \right) + 26.32 \left( \frac{B_6}{B_2} \right)^2$      | 0.66           | Cruz-Montes et al. (2023) |
| Sentinel 2 A   | Lake           | India    | $pH = 5.0159 \left( \frac{B_2}{B_4} \right) + 5.6782 \left( \frac{B_1}{B_2} \right)$                                                                                                                                      | 0.98           | Mohandas and Brema (2023) |
| Landsat 8 OLI  |                |          |                                                                                                                                                                                                                           | 0.99           |                           |
| Landsat 5 TM   |                |          | $pH(GP) = 5.502868058 + 0.000268732 \cdot NDWI + 5.669416074e^{-8} \cdot MNDWI + 4.014937768e^{-6} \cdot (MNDWI)^2 + 0.485758883 \cdot \sin(5.035647409 + 5.502868058) - 0.375986531 \cdot \sin(2.502132245 \cdot MNDWI)$ | 0.81           |                           |
| Landsat 7 ETM+ | Lake           | Brazil   | $pH(SMLR) = 1.3633 - 0.00110 \cdot AWEInsh + 0.00818 \cdot B_2 - 0.00392 \cdot B_4 + 0.00120 \cdot NDWI$                                                                                                                  | 0.73           | Pereira et al. (2020)     |
| Landsat 8 OLI  |                |          | $pH(Autum) = (-642.7787152 + (74118.07298 \cdot (B_5)) + (-3454146.477 \cdot (B_5)^2) + (84300360.47 \cdot (B_5)^3) + (-1136293090 \cdot (B_5)^4) + (8022005949.148 \cdot (B_5)^5) + (-23184321457.0779 \cdot (B_5)^6))$  | 0.97           |                           |
|                |                |          | $pH(Summer) = (478.2633 + (-540486.840 \cdot (B_4)) + (12516620.16 \cdot (B_4)^2) + (-121653098.3 \cdot (B_4)^3) + (-1136293090 \cdot (B_4)^4) + (557708124.0187 \cdot (B_4)^5) + (-993832250.974 \cdot (B_4)^6))$        | 0.91           | Abbas et al. (2021)       |
|                |                |          | $pH(Spring) = (22.865 + (-15703.293 \cdot (B_5)^2) + (335329.655 \cdot (B_5)^3) + (-2931384.848 \cdot (B_5)^4) + (11821314.64 \cdot (B_5)^5) + (-18168435.54 \cdot (B_5)^6))$                                             | 0.99           |                           |
|                |                |          | $pH(Winter) = ((388.0475 + (-22957.45678 \cdot (B_4)) + (563449.015 \cdot (B_4)^2) + (-7184093.563 \cdot (B_4)^3) + (50158611.13 \cdot (B_4)^4) + (-181769305.2 \cdot (B_4)^5) + (267101357.1 \cdot (B_4)^6))$            | 0.98           |                           |

NDVI, AWEInsh, NDWI, and MNDWI, have been utilized to develop pH retrieval algorithms (Pereira et al., 2020). According to the existing literature, most studies found a positive correlation between pH values and green, red, blue, NIR, and VNIR spectral bands (spectral reflectance in the range from 480 nm to 2210 nm) for medium-resolution satellites (Fig. 6). In contrast, the spectral reflectance ranges between 400 nm and 800 nm in hyperspectral images utilized for pH retrieval from lake and river waterbodies (Table S5).

### 5.1.2. Performance of various RS data in retrieving pH concentration

Fig. 7a shows the differences between in-situ pH measurements and RS retrieval pH values from different studies. It can be seen from the Fig. 7a that the Landsat 8 OLI and the Sentinel 2 MSI demonstrated high accuracy in RS-based pH retrieval. For instance, the difference between the measured and predicted values of the Landsat 8 OLI-based pH retrieval model was 0.11 (Cao et al., 2022a; Krishnaraj and Honnasiddaiah, 2022; Pereira et al., 2020). Furthermore, using a recent cutting-edge technology known as the image fusion process, researchers have achieved high accuracy in pH measurements. For example, Batur and Maktav (2019) utilized the Gokturk-2, the Landsat 8 OLI, and the Sentinel 2 MSI to retrieve pH from Turkish lakes, where the difference between measured and predicted pH was  $\approx 0.29$ . However, it should be noted that image fusion is a complex process that requires multi-sensor data. Additionally, Qian et al. (2022) obtained high accuracy from the Sentinel 2 MSI-based pH retrieval model.

### 5.1.3. Comparison of various RS models for retrieving pH

Although, in the field of RS-based WQ indicators, it is widely acknowledged that water pH values cannot be directly measured from RS data (Jiang et al., 2022; Pereira et al., 2020). However, numerous studies have explored the efficacy of RS and ML/AI-based models for pH estimation (Abbas et al., 2021; Batur and Maktav, 2019; Cao et al., 2022a; Cao et al., 2022b; Cruz-Montes et al., 2023; Jiang et al., 2022; Krishnaraj and Honnasiddaiah, 2022; Mohandas and Brema, 2023; Pereira et al., 2020; Qian et al., 2022). Additionally, many studies have shown that water pH value can be estimated based on the statistical relationship between pH and other WQ indicators, such as CHL, TEMP (Nakano and Watanabe, 2005), pCO<sub>2</sub>, DIC, TA (Sabia et al., 2015) and the spectral reflectance of the RS data (Abbas et al., 2021; Cruz-Montes et al., 2023; Pereira et al., 2020; Qian et al., 2022). According to existing literature, ML/AI and EM models are widely utilized for pH retrieval using RS technology. Notably, based on the evaluation metrics (R<sup>2</sup> and RMSE score), EM have consistently performed better than ML/AI models (Fig. 8). It can be seen from Fig. 8 that there was a significant difference between the EM and ML/AI retrieval models performance.

In the field of pH estimation using RS technology, Mohandas and Brema (2023) developed two different pH retrieval equations using the wavelength ranges of 443 nm – 665 nm from the Landsat 8 OLI and the Sentinel 2 MSI, respectively. These models demonstrated high performance, with R<sup>2</sup> values ranging from 0.98 to 0.99. Similarly, Abbas et al. (2021) utilized the MLR model to retrieve river pH in Iraq across different seasons. This study utilized the B4 - B5 (655 nm – 865 nm) bands of Landsat 8 OLI as input variables and achieved outstanding results during both training (R<sup>2</sup> > 0.90) and validation (R<sup>2</sup> > 0.85) periods. Furthermore, Table 2 summarizes various regression models for retrieving pH in different waterbodies using different RS data. Additionally, the study by Pereira et al. (2020) has utilized the cloud-based approach for retrieving pH in Brazilian lakes using time series data from the Landsat satellites. In this study, the authors employed the GP and SMLR models for retrieving water pH by utilizing the Google Earth Engine platform in various lakes. Their findings showed that, compared to the SMLR (R<sup>2</sup> = 0.73), the GP model could be effective (R<sup>2</sup> = 0.81) for predicting pH in lakes (Table 2).

To the best of the authors' knowledge, there are no tools or techniques for directly retrieving the pH from satellite data. However, recent studies have notably employed ML/AI technologies to predict or retrieve

pH concentrations directly from RS data. For instance, Jiang et al. (2022) proposed a framework to retrieve pH concentration from the RS data using ML/AI methods. In this study, the authors utilized the RF model to predict monthly sea surface pH for the years 2004–2019 using the MODIS-Aqua data. Consequently, the study revealed that the RF model performed better in both the training and testing phases with a  $R^2$  value of 0.80 (Fig. 8 and Table S5). Moreover, evidence from different studies suggest that NN-based models performed better with complex nonlinear variables than EM techniques. For example, Cao et al. (2022a) developed an ANN-based retrieval model based on the correlation of the Landsat 8 OLI data and in-situ pH measurements, which showed relatively low errors compared to a regression model (relative error ANN model = 1.25 % and relative error regression model = 1.58 %). Similarly, the BPNN model showed higher effectiveness than the RF model with  $R^2$  value of 0.90 during training and testing periods (Cao et al., 2022b). Additionally, to investigate the spatio-temporal distribution of pH in the Ganga River basin in India, Krishnaraj and Honnasiddaiah (2022) developed two ML-based models (XGB and MLP), where they utilized B1–B4 bands (440 nm - 655 nm) of the Landsat 8 OLI, along with their ratios, as input variables and validated the model using the in-situ measurements. Consequently, both models show high ( $R^2 = 0.72$ ) accuracy in the training and testing phase (Fig. 8 and Table S5). Furthermore, recent studies have explored alternative approaches, such as image fusion combined with a PCA regression model to retrieve pH in lakes (Batur and Maktav, 2019). The study showed that, compared to other ML/AI models (e.g., MLR, ANN and SVM), the PCA regression model offered a reliable alternative for retrieving pH (Table S5).

#### 5.1.4. pH retrieval algorithm(s) optimization processes

Publications focusing on the development of ML/AI models for retrieving water pH using RS technology often incorporate various hyperparameter optimization techniques to enhance the model performance in order to achieve the optimal results. Table S6 summarizes the optimal hyperparameter utilized across various pH retrieval models in the literature. To date, several studies have implemented the hyperparameters tuning approaches to improve the performance of pH retrieval models (Cao et al., 2022a, 2022b; Jiang et al., 2022; Krishnaraj and Honnasiddaiah, 2022). For instance, Jiang et al. (2022) have utilized 1200 decision trees during training to enhance the RF model's accuracy. Their study reported that the using 1200 trees, a maximum tree depth of 20, and 30 iterations could be potentially improved the model performance; whereas Krishnaraj and Honnasiddaiah (2022) adopted a grid search approach to optimize the XGB model for pH retrieval. Additionally, the study by Krishnaraj and Honnasiddaiah (2022) recommended a learning rate of 0.26, a max depth of trees 4, L1 regularization term on weights of 0.22, and L2 regularization term on weights of 0.00059588 to enhance retrieval accuracy (Table S6). Similarly, Cao et al. (2022a; 2022b) have reported that neuron sizes of seven and five were effective in improving the performance of the ANN and BPNN models, respectively, for pH retrieval.

#### 5.1.5. Limitations

- Most of the ML/AI-based pH retrieval models are focused on inland waters and are not suitable for long-term prediction.
- The developed pH retrieval models are not generalized and require further parameterizations and algorithm updates based on the application domain.
- In large waterbodies, EM techniques failed to achieve good accuracy in pH retrieval.
- The data gaps in the RS data reduced the pH retrieval model's accuracy.
- The uncertainty and reliability of the developed retrieval models are not specified.

- Another major constraint of the developed retrieval model(s) is that their sensitivity has not been specified or validated in terms of geo-spatial resolution.

### 5.2. Dissolved Oxygen (DOX)

DOX is a fundamental chemical property of aquatic ecosystems. Typically, the sources, consumption, and water solubility of oxygen determine the DOX concentration (Guo et al., 2021b). A brief description of DOX associations with different water catastrophes (e.g., heavy rainfall, flooding, typhoons, water temperature) and traditional measurement approaches can be found in the supplementary materials as a continuation of Section 5.2. In addition to traditional approaches, several investigations of DOX have utilized complex computational models such as QUAL2E (Palmieri and De Carvalho, 2006), QUAL2K (Ye et al., 2013) and QUASAR (Cox, 2003). However, these models have failed to provide a generalized approach for all users due to their complex nature (Zhang et al., 2020). Similarly, statistical models often struggle to accurately predict DOX concentrations (Moghadam et al., 2021). As a result, ML/AI and RS techniques have been proposed as an alternative method for predicting and quantifying DOX in waterbodies (Chatziantoniou et al., 2022; Ziyad Sami et al., 2022).

#### 5.2.1. RS data and bands (wavelength) used for DOX retrieval models

According to the literature, many studies have utilized the Landsat and Sentinel data to develop DOX retrieval model(s) (Ahmed et al., 2022; Cruz-Montes et al., 2023; Karaoui et al., 2019; Padilla-Mendoza et al., 2023; Peterson et al., 2020; Tanjung et al., 2023; Tian et al., 2023). Among these, the Landsat 8 OLI and the Sentinel 2 MSI sensors are most widely used for retrieving DOX concentrations (Fig. 5b). Additionally, several recent researchers have also explored the use of the MODIS-Aqua/Terra data to validate their potential of retrieving DOX concentrations in different waterbodies (Guo et al., 2022b; Kim et al., 2020; Liu et al., 2022b). Moreover, there has been a growing trend in utilizing hyperspectral RS data to assess DOX concentrations (Li et al., 2023a; Wang et al., 2020; Wang et al., 2021; Yang et al., 2022c). In addition to these frequently used satellite products, such as the Gokturk-2 (Batur and Maktav, 2019), the CMEMS (Copernicus Marine Environment Monitoring Service) ocean products (Chatziantoniou et al., 2022), and the VIIRS (Kim et al., 2020) have also been employed to develop DOX retrieval algorithms for estimating DOX concentrations in aquatic environments (Fig. 5b). Regarding input variables during model development, studies have frequently utilized blue, green, red, and VNIR of the Sentinel 2 MSI data for rivers (Salas et al., 2022; Yang et al., 2022c), while green, red, and VNIR bands are commonly used for lakes, respectively (Batur and Maktav, 2019; Dong et al., 2023; Peterson et al., 2020). Moreover, the B1–B8a (Coastal-VNIR) bands of the Sentinel 2 MSI are often employed to estimate DOX in transitional and coastal (TrC) waters (Batur and Maktav, 2019; Zhu et al., 2022a) (Fig. 6). Recent studies have also identified a positive association between DOX and blue, green, red and NIR bands of the Landsat 8 OLI for lakes and rivers (Table S7) (Al-Shaibah et al., 2021; Cao et al., 2022a; Mohandas and Brema, 2023). Additionally, the study by Karakaya and Evrendilek (2011) showed that the B1–B7 (Coastal-SWIR) bands of the Landsat 8 OLI can be effective for retrieving DOX concentrations in TrC waters. On the other hand, in terms of hyperspectral data, most DOX retrieval models developed using the spectral reflectance range between 450 nm – 810 nm and reported that this particular spectral range could be effective for retrieving DOX in aquatic environments, especially for lake and river waters (Table S7).

#### 5.2.2. Performance of various RS data in retrieving DOX concentration

For the purposes of understanding the model performance, the study has carried out a comparative analysis of various retrieval models using different RS data to estimate DOX in rivers, lakes and TrC waters. Fig. 6 presents a summary of the wavelengths of RS data; these are widely

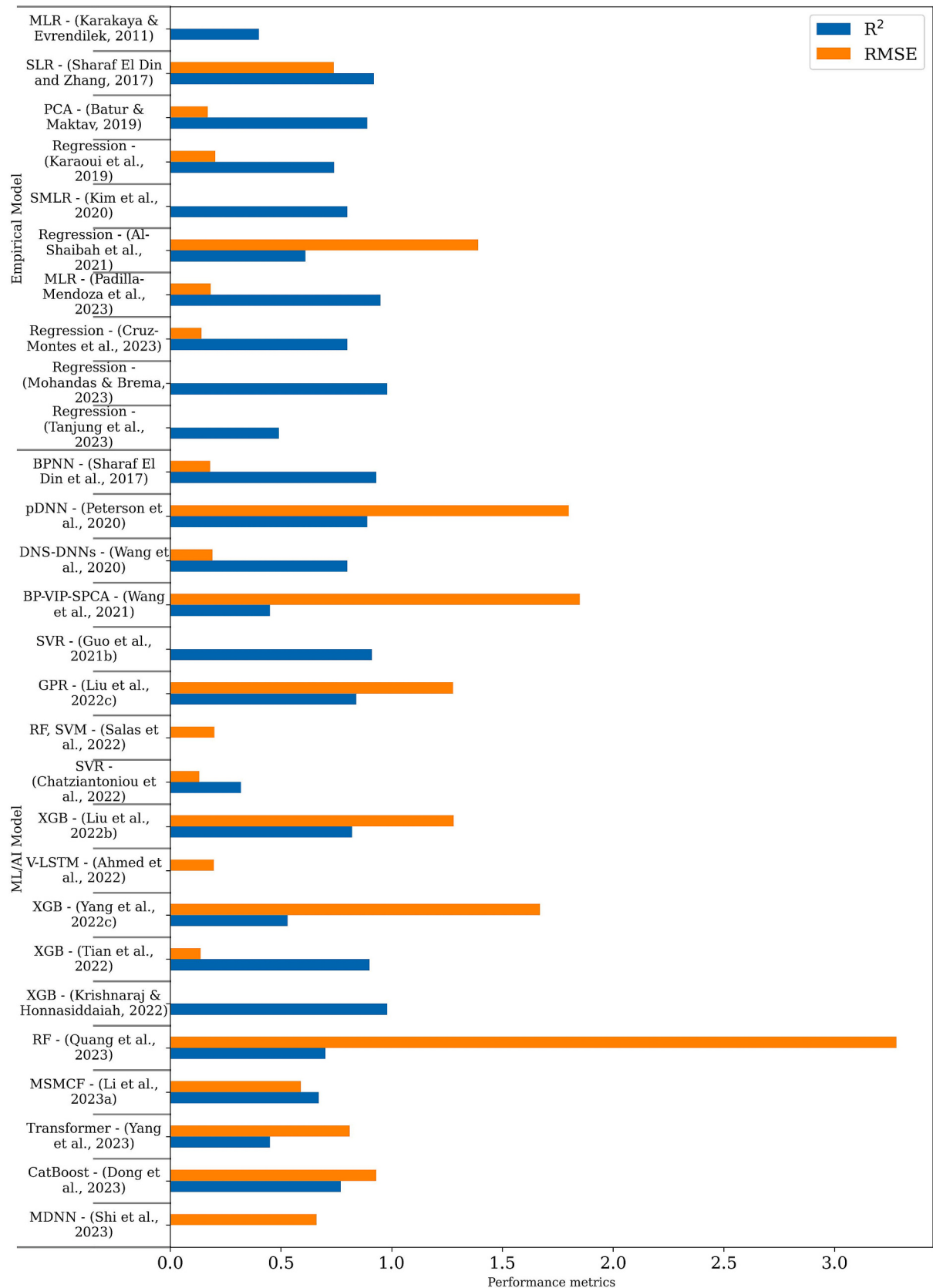


Fig. 9. Performance evaluation results of different DOX retrieval models.

**Table 3**  
Application of regression-based DOX retrieved models.

| RS Data Type                                      | Waterbody Type | Domain   | Equation                                                                                                                                                                                                        | R <sup>2</sup> | Reference                      |
|---------------------------------------------------|----------------|----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|--------------------------------|
| Sentinel 2 MSI                                    | Wetlands       | India    | $DOX = 4.7068 * \left( \frac{B_2}{B_4} \right)$                                                                                                                                                                 | 0.81           | Mohandas and Brema (2023)      |
| Landsat 8 OLI                                     |                |          | $DOX = 5.3111 * \left( \frac{B_1}{B_2} \right)$                                                                                                                                                                 | 0.99           |                                |
| Landsat 8 OLI                                     | Lake           | Colombia | $DOX = \left[ 3.940628.099(B_4)^2 * 0.32123 \left( \frac{B_4}{B_2} \right) * 1.08(10)^9 * (B_{11}) + 32.083 \left( \frac{B_4}{B_2} \right)^3 \right]^2$                                                         | 0.80           | Cruz-Montes et al. (2023)      |
| Sentinel 2 MSI                                    | Wetlands       | Colombia | $DOX = -39.2556 + 0.8061/B_4 + 4288.3263 * (B_4 * B_5) + 19.4829 * (B_4/B_5)$                                                                                                                                   | 0.95           | Padilla-Mendoza et al. (2023)  |
| Landsat 5 TM,<br>Landsat 7 ETM+,<br>Landsat 8 OLI | Lake           | China    | $DOX_{Landsat-8OLI} = (B_5 - B_4) * \left( \left( \frac{B_5}{B_3} \right) * 80 \right) + 8.3$<br>$DOX_{Landsat 5TM \text{ or } 7ETM+} = (B_4 - B_3) * \left( \left( \frac{B_4}{B_2} \right) * 80 \right) + 8.3$ | 0.61           | Al-Shaibah et al. (2021)       |
| MODIS Aqua and<br>VIIRS                           | Sea            | Korea    | $DOX = -0.131 * SST - 0.132 - SST_{m-1} - 0.066 * Chla_{m-1} + 12.343$                                                                                                                                          | 0.80           | Kim et al. (2020)              |
| Sentinel 2 MSI                                    | Reservoir      | Morocco  | $DOX = -0.0167 * B_8 + 0.0067 * B_9 + 0.0162 * B_{10} + 0.0162 * B_{11} + 9.577$                                                                                                                                | 0.74           | Karaoui et al. (2019)          |
| Landsat 8 OLI                                     | River          | Canada   | $DOX = -4.635 \left( \frac{B_6}{B_3} \right) - 51.305 * \left( \frac{B_1}{B_2} \right) + 75.409$                                                                                                                | 0.91           | Sharaf El Din and Zhang (2017) |

utilized for developing DOX retrieval models. Since DOX is an OI-WQ indicator, its concentration is estimated indirectly through various OA-WQ indicators. Notably, the most frequently utilized indicators include TEMP, CHL, TSS and SDD (Chatziantoniou et al., 2022; Guo et al., 2022b; Kim et al., 2020; Liu et al., 2022b). Furthermore, recent studies have explored different bands' arithmetic formulas, such as NDI, SI, DI index, etc., to develop models (Salas et al., 2022; Peterson et al., 2020; Zhu et al., 2022a). However, most studies have primarily relied on correlations between different RS bands and DOX or other WQ indicators (Al-Shaibah et al., 2021; Cao et al., 2022a; Liu et al., 2022c; Quang et al., 2023; Salas et al., 2022). Fig. 7b shows a comparison result between in-situ and RS-retrieved DOX concentrations (unit = mg/L). It can be seen from the Fig. 7b that the use of the MODIS-Aqua, the Sentinel 2 MSI, and the Landsat 8 OLI demonstrated excellent performance in retrieving DOX concentrations (Fig. 7b). For example, Al-Shaibah et al. (2021) reported a minimal variation of  $\pm 0.08$  mg/L between measured and predicted DOX using the Landsat 8 OLI. Additionally, the study by Mohandas and Brema (2023) showed that compared to the Sentinel 2 MSI, the Landsat 8 OLI provided better accuracy in order to retrieve DOX. Similarly, a comparative study has been conducted between three different hyperspectral satellites (e.g., GF5-01, ZY1-02D, and Zhuhai-1), where the Zhuhai-1 performed better than other hyperspectral satellites (Li et al., 2023a). Moreover, several studies have highlighted the effectiveness of fusion satellite data and UAV hyperspectral data in retrieving DOX from different waterbodies (Batur and Maktav, 2019; Li et al., 2023a; Peterson et al., 2020). Nevertheless, compared to commonly used datasets, the image fusion method resulted in a relatively high variation of  $\pm 4.0$  mg/L between measured and retrieved DOX concentrations (Batur and Maktav, 2019). Additionally, several studies have reported that the overestimation is a common problem in retrieving DOX from RS data-driven approaches (Dong et al., 2023; Krishnaraj and Honnasiddaiah, 2022; Sharaf El Din et al., 2017).

### 5.2.3. Comparison of various RS models for retrieving DOX

Recently a series of studies have widely utilized the state-of-the-art ML/AI techniques for retrieving DOX concentration using RS data across different waterbodies (Shi et al., 2023; Dong et al., 2023; Quang et al., 2023; Tian et al., 2023). This rapid advance of ML/AI techniques has facilitated accurate measurements of DOX in various waterbodies, incorporating complex hydro-environmental issues (Maroufpoor et al., 2022). However, it is widely acknowledged that retrieving DOX

concentrations using RS data remains challenging, mainly because DOX is an OI-WQ indicator (Chatziantoniou et al., 2022; Salas et al., 2022). Despite these challenges, numerous studies have demonstrated that DOX concentrations can be successfully retrieved and predicted by integrating satellite-derived OI-WQ indicators (e.g., CHL, TEMP, and TSS) with ML/AI algorithms (Liu et al., 2022c; Quang et al., 2023; Sharaf El Din et al., 2017).

Generally, EM techniques are widely utilized to retrieve DOX concentrations from satellite data (Sharaf El Din and Zhang, 2017; Karakaya and Evrendilek, 2011; Kim et al., 2020). Moreover, regression-based algorithms, such as MLR, SMLR, LR, etc., have consistently demonstrated strong performance in retrieving DOX concentrations from various waterbodies (Table S7). For instance, Kim et al. (2020) developed the first generalized RS-based DOX retrieval model using the SMLR model, which effectively correlated the measured DOX with satellite-derived OA-WQ indicators (SST and CHL) to capture long-term DOX concentration trends. Similarly, Sharaf El Din and Zhang (2017) developed a DOX retrieval algorithm using the SMLR model and achieved high accuracy with R<sup>2</sup> value of 0.92 (Fig. 9). The derived equation from the regression-based study can be found in Table 3. Consequently, numerous studies have employed a similar approach to develop RS-based DOX retrieval model (Al-Shaibah et al., 2021; Batur and Maktav, 2019; Cruz-Montes et al., 2023; Escoto et al., 2021; Karaoui et al., 2019; Mohandas and Brema, 2023; Padilla-Mendoza et al., 2023). However, it is worth noting that some studies, such as Al-Shaibah et al. (2021), Escoto et al. (2021), and Karakaya and Evrendilek (2011), reported unsatisfactory results compared to the research conducted by Batur and Maktav (2019), Cruz-Montes et al. (2023), Kim et al. (2020), Mohandas and Brema (2023), and Padilla-Mendoza et al. (2023) (Fig. 9 and Table S7).

Furthermore, numerous studies have demonstrated the effectiveness of ML/AI-based tools and techniques in retrieving DOX concentrations (Table S7). For example, Chatziantoniou et al. (2022) utilized satellite-derived CHL and SST data coupled with the SVR model to estimate DOX concentrations in Greece. Although the proposed model yields an unsatisfactory R<sup>2</sup> value of 0.32 (Fig. 9). In contrast, Guo et al. (2021a) utilized the Landsat series and the MODIS-Aqua/Terra data, also coupled with the SVR model for long-term DOX retrieval from lakes, achieving a robust model performance with R<sup>2</sup> value exceeding 0.90. In both studies by Chatziantoniou et al. (2022) and Guo et al. (2021a), the SVR model showed excellent performance in retrieving DOX using RS data. However, the SVR model failed to demonstrate high performance

in the study by Liu et al. (2022b). On the other hand, the study by Quang et al. (2023) found that the RF model performed most accurately ( $R^2 = 0.70$ ) in the DOX concentration retrieval process using the Sentinel 2 MSI data. Additionally, another ML model, the XGB, has demonstrated robust performance for DOX estimation across different waterbodies worldwide using the Landsat 8 OLI, the Sentinel 2 MSI, and hyperspectral data (Krishnaraj and Honnasiddaiah, 2022; Tian et al., 2023; Yang et al., 2022c; Zhu et al., 2022a) (Table S7).

Moreover, the concentration of DOX can be estimated based on the empirical relationship between satellite band reflectance, synthetic band combinations, and in-situ DOX measurements (Table S7). For example, Salas et al. (2022) used the Sentinel 2 MSI and ML algorithms (RF and SVM models) to map DOX in the Little Miami River, USA. Likewise, Sharaf El Din et al. (2017) utilized the BPNN algorithm and the Landsat 8 OLI to retrieve river DOX in Canadian rivers, achieving high accuracy (Fig. 9). Conversely, Liu et al. (2022c) found that both the BPNN and SVM models underperformed compared to the GPR model. Additionally, DL/hybrid models have also demonstrated strong performance in estimating DOX levels across various waterbodies. Specifically, models such as V-LSTM (Ahmed et al., 2022), BP-VIP-SPCA (Wang et al., 2021), DNS-DNNs (Wang et al., 2020), pDNN (Peterson et al., 2020) and MSMCF (Li et al., 2023a) have significantly outperformed different standalone models in DOX estimation using RS-based approaches (Table S7).

#### 5.2.4. DOX algorithm optimization techniques

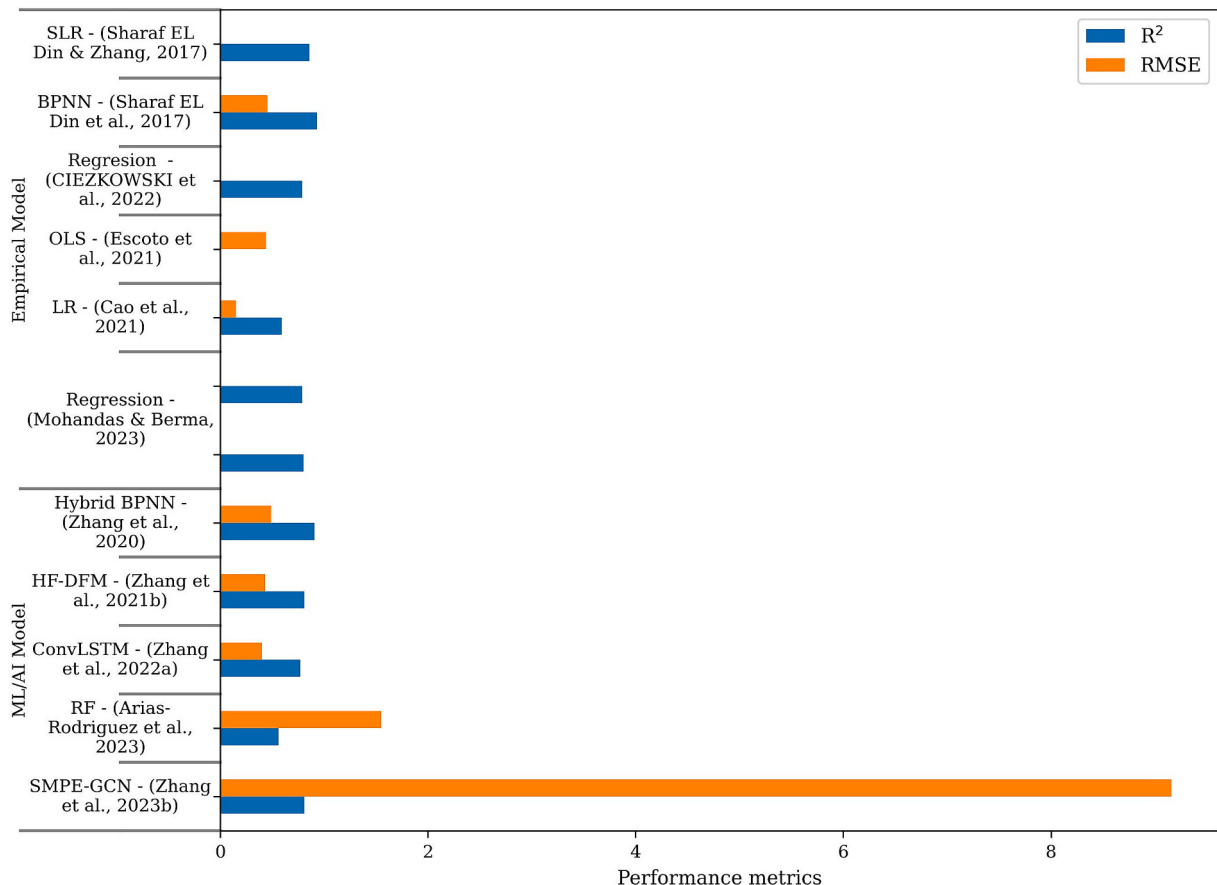
In order to achieve high performance and reliable predictions, several studies have modified the default hyperparameters in the RS-based ML/AI models for DOX retrieval (Cao et al., 2022a; Chatziantoniou et al., 2022; Guo et al., 2021b; Liu et al., 2022a; Liu et al., 2022c; Sharaf El Din et al., 2017). For example, grid search techniques have

**Table 4**

Application of regression-based BOD<sub>5</sub> retrieval models.

| RS data type   | Waterbody type | Domain | Equation                                                                  | R <sup>2</sup> | Reference                      |
|----------------|----------------|--------|---------------------------------------------------------------------------|----------------|--------------------------------|
| Sentinel 2 MSI | Lake           | India  | $BOD_5 = 2.8759 * \left( \frac{B_1}{B_2} \right)$                         | 0.79           | Mohandas and Brema (2023)      |
| Zhuhai-1       | Lake           | China  | $BOD_5 = 1.0504 \left( \frac{B_{12}}{B_{13}} \right) - 0.0028$            | 0.59           | Cao et al. (2021)              |
| Landsat 8 OLI  | River          | Canada | $BOD_5 = 0.568 * \left( \frac{B_1}{B_6} \right) + 189.29 * (B_6) - 4.821$ | 0.86           | Sharaf El Din and Zhang (2017) |

been employed by Guo et al. (2021b), Krishnaraj and Honnasiddaiah (2022), and Liu et al. (2022b) to identify the optimal hyperparameter values. Additionally, the utilization of a radial basis kernel function enhances the performance of the RS-based SVR model (Chatziantoniou et al., 2022; Guo et al., 2021b). Table S8 presents the hyperparameters discussed in different studies. It can be seen from Table S8 that Sharaf El Din et al. (2017) developed a BPNN-based DOX model by employing 20 hidden layers, a learning rate of 0.01, and a sigmoid activation function to minimize model error. Similarly, the XGB model utilized a regression lambda value of 0.22, a learning rate of 0.23, and a maximum depth of 4 in order to achieve favourable performance (Table S8). Furthermore, Ahmed et al. (2022) and Tian et al. (2023) employed hyperparameter tuning techniques to optimize the performance of the V-LSTM and XGB



**Fig. 10.** Performance evaluation results of different BOD<sub>5</sub> retrieval models.

models, respectively.

### 5.2.5. Limitations

- Existing DOX retrieval models are developed based on a specific region and cannot be generalized to another domain.
- Most of the DOX concentration retrieval models have not focused on the spatio-temporal variability.
- Cloud cover had a substantial negative impact on the accuracy of DOX retrieval models.
- Researchers did not recommend any optimal RS wavelengths for retrieving DOX from waterbodies.
- Over- and underestimation of DOX concentrations was a common phenomenon in DOX retrieval models.

### 5.3. Biochemical oxygen demand (BOD<sub>5</sub>)

BOD<sub>5</sub> refers to the amount of DOX required by microorganisms to break down organic matter in water, usually measured at a TEMP of 20 °C over a specific period (Li and Zhang, 2020). This method is commonly employed to assess the quantity of biodegradable organic pollution present in waterbodies (Abyaneh, 2014). A detailed description of the significance of BOD<sub>5</sub> as a WQ indicator, its impact on aquatic systems, and various measurement techniques can be found in the supplementary materials as a continuation of Section 5.3. In addition to traditional and biosensor methods, ML/AI-based prediction models are commonly used for BOD<sub>5</sub> estimation (Baki et al., 2019). According to Alsulaili and Refaie (2021), measuring BOD<sub>5</sub> with a prediction model saves time and enables online management. Therefore, a series of recent studies on BOD<sub>5</sub> (Arias-Rodriguez et al., 2023; Cao et al., 2021; Escoto et al., 2021; Sharaf El Din et al., 2017; Sharaf El Din and Zhang, 2017; Zhang et al., 2022a; Zhang et al., 2021b) explored the effectiveness of RS technology for predicting BOD<sub>5</sub> in various waterbodies, even though BOD<sub>5</sub> is an OI-WQ indicator.

#### 5.3.1. RS data and bands (wavelength) used for BOD<sub>5</sub> retrieval models

BOD<sub>5</sub> is one of the most commonly analysed WQ indicators in the planning and management of wastewater treatment plants (Baki et al., 2019). Typically, RS-based BOD<sub>5</sub> retrieval models utilized satellite-derived band combination and band reflectance to retrieve BOD<sub>5</sub> from different waterbodies (Table S9). Among the RS data, the Sentinel 2 MSI and the Landsat 8 OLI are most widely utilized to retrieve BOD<sub>5</sub> concentrations (Fig. 5c). Fig. 6 shows a statistical summary of frequently utilized RS wavelengths in BOD<sub>5</sub> estimation. In literature, different studies used different bands to retrieve BOD<sub>5</sub> for lakes and rivers; however, there are no defined conclusion (Table S9). For example, both Mohandas and Brema (2023) and Arias-Rodriguez et al. (2023) have utilized the Sentinel 2 MSI and the Landsat 8 OLI data to retrieve BOD<sub>5</sub> from lakes. In these studies, Mohandas and Brema (2023) retrieved BOD<sub>5</sub> using the spectral reflectance range of 440 nm – 490 nm, whereas Arias-Rodriguez et al. (2023) retrieved BOD<sub>5</sub> utilizing 560 nm – 865 nm (Table S9). Similarly, numerous studies reported a positive correlation of BOD<sub>5</sub> with the 440 nm – 1610 nm spectral reflectance range of the Landsat 8 OLI data in rivers (Sharaf El Din et al., 2017; Sharaf El Din and Zhang, 2017; Zhang et al., 2022a). On the other hand, a number of studies proposed that the UAV hyperspectral image reflectance range from 400 nm to 930 nm can be suitable for retrieving BOD<sub>5</sub> from rivers (Zhang et al., 2020; Zhang et al., 2021b; Zhang et al., 2023b).

#### 5.3.2. Performance of various RS data in retrieving BOD<sub>5</sub> concentration

Several researchers, including Arias-Rodriguez et al. (2023), Escoto et al. (2021), Sharaf El Din and Zhang (2017), and Zhang et al. (2022c), developed a RS-based BOD<sub>5</sub> retrieval model using medium-resolution satellites, such as the Landsat 8 OLI and the Sentinel 2 MSI (Fig. 7c). Additionally, hyperspectral RS images were employed to develop the BOD<sub>5</sub> retrieval models (Cao et al., 2021; Zhang et al., 2021b). However,

hyperspectral RS data performed better than medium-resolution satellites in terms of retrieval accuracy (Fig. 7c). For example, the difference between measured and predicted BOD<sub>5</sub> in hyperspectral and the Landsat 8 OLI-based retrieval models is  $\pm 0.67$  mg/L (Zhang et al., 2021b) and  $\pm 1.25$  mg/L (Sharaf El Din et al., 2017), respectively (Fig. 7c).

#### 5.3.3. Comparison of various RS models for retrieving BOD<sub>5</sub>

Data-driven ML/AI approaches are most widely used for BOD<sub>5</sub> concentration prediction and simulation, especially in wastewater treatment applications (Baki et al., 2019; Qambar and Al Khalidy, 2023). Although BOD<sub>5</sub> is an OI-WQ indicator, a number of RS-based approaches combined with ML/AI methods were used to retrieve BOD<sub>5</sub> concentrations from various waterbodies (Table S9). For instance, Sharaf El Din et al. (2017) quantified BOD<sub>5</sub> concentrations using the BPNN algorithm and the Landsat 8 OLI data. In this study, compared to the SVM model, the BPNN model achieved higher accuracy ( $R^2 = 0.93$ ). Similarly, the Landsat 8 OLI band ratios and the SMLR model were employed to determine BOD<sub>5</sub> concentrations in rivers, where the coastal aerosol and SWIR-2 bands of the Landsat 8 OLI were particularly effective in retrieving BOD<sub>5</sub> concentrations (Sharaf El Din and Zhang, 2017). The performance of different BOD<sub>5</sub> retrieval models is shown in Fig. 10, and the derived regression equations presented in Table 4. Unlike Sharaf El Din and Zhang (2017), the empirical regression model has not performed strongly in the study of Escoto et al. (2021). The study by Escoto et al. (2021) has developed a BOD<sub>5</sub> retrieval model coupling the OLS model and the Sentinel 2 MSI in the Pasig River, Philippines. While most of the available literature on BOD<sub>5</sub> concentration deals with single satellites, Arias-Rodriguez et al. (2023) developed a cloud-based RF model for retrieving BOD<sub>5</sub> concentrations at a global scale by combining data from the Landsat 8 OLI and the Sentinel 2 MSI. Although this model didn't perform well, it is still a first step towards developing the global-scale BOD<sub>5</sub> retrieval model for lakes using the image fusion process.

Regarding the DL approach for estimating BOD<sub>5</sub> concentrations, Zhang et al. (2022a) developed a novel DL model driven by spectral characteristics to determine monthly six-year BOD<sub>5</sub> changes at Dongping Lake in China. Specifically, this study utilized 26 Landsat 8 OLI band arithmetic formulas as model inputs, and the ConvLSTM model performed notably well, with a  $R^2$  of 0.77 (Fig. 10). Similarly, Zhang et al. (2021b) developed the HF-DFM model to quantify BOD<sub>5</sub> concentrations utilizing UAV hyperspectral data. Notably, performance evaluation metrics demonstrated that this model outperformed other models, such as OLS and BPNN, achieving a  $R^2$  value of 0.81 (Fig. 10). Furthermore, Zhang et al. (2023b) proposed the novel SMPE-GCN model that incorporates feature engineering to predict BOD<sub>5</sub> concentrations in rivers using UAV hyperspectral data. Meanwhile, Fu et al. (2022) demonstrated that the LOOCV-XGB model is an optimal method for estimating BOD<sub>5</sub> in lakes when combined with multispectral and hyperspectral data. Nevertheless, the application of this model is more complex than other models.

Moreover, in hyperspectral-based BOD<sub>5</sub> retrieval model development studies, Cao et al. (2021) and Zhang et al. (2020, 2021b) investigated the relationship between BOD<sub>5</sub> and the spectral reflectance of hyperspectral data, where they found a strong positive correlation in the wavelength range of 400 nm to 800 nm. Specifically, Cao et al. (2021) estimated BOD<sub>5</sub> concentrations in Chinese lakes using the B12/B13 (640 nm/500 nm) composite band of Zuhai-1 hyperspectral data. The derived regression equation can be found in Table 4. Similarly, Zhang et al. (2020) introduced a hybrid BPNN to estimate BOD<sub>5</sub> concentration using UAV hyperspectral images. The proposed method showed high accuracy ( $R^2 = 0.91$ ) despite the small size of the training data (Fig. 10). Moreover, the model demonstrated superior performance compared to the hybrid ANN-BPNN model, SEM, and empirical regression methods (Table S9).

#### 5.3.4. BOD<sub>5</sub> algorithm optimization techniques

Data from several studies showed that, compared to traditional ML/

AI models, optimized models performed better in developing BOD<sub>5</sub> retrieval models (Arias-Rodriguez et al., 2023; Sharaf El Din et al., 2017; Zhang et al., 2020; Zhang et al., 2021b; Zhang et al., 2023b). Table S10 shows the hyperparameters discussed in various BOD<sub>5</sub> retrieval studies. For example, ReLU, Leaky-ReLU and sigmoid activation functions were employed to maximize the performance from the RS-based BPNN model (Sharaf El Din et al., 2017) and the hybrid BPNN model (Zhang et al., 2020) for BOD<sub>5</sub> concentration retrieval. However, in NN-based models, selecting the appropriate number of hidden layers remains a challenge, as no constant number of hidden layers is suggested in any studies. For instance, Zhang et al. (2023b) set a hidden layer of 3 and 15 input layers to develop the SMPE-GCN model, whereas Sharaf El Din et al. (2017) used 4 hidden layers and 7 input layers to develop the BPNN algorithm. Additionally, Arias-Rodriguez et al. (2023) used a grid search technique to optimize the model hyperparameters, as this method assesses model accuracy at each grid. Moreover, in order to minimize model errors, Fu et al. (2022) and Arias-Rodriguez et al. (2023) utilized 5-fold and 10-fold cross-validation approaches, respectively (Table S10).

### 5.3.5. Limitations

- Most of BOD<sub>5</sub> retrieval models have been developed for inland waters.
- Standalone ML/AI-based retrieval models failed to achieve high accuracy of BOD<sub>5</sub> concentrations.
- No researcher has suggested any unified RS wavelength for estimating BOD<sub>5</sub> concentrations.
- Models trained using small datasets often overestimate BOD<sub>5</sub> concentrations in various retrieval models.

## 5.4. Total phosphorus (TP)

TP is the total amount of orthophosphate, phosphorus monoester, phosphorus diester, phosphonate, polyphosphate monoester, polyphosphate, and pyrophosphate (all P-content) present in a water sample (Ma et al., 2017). A detailed description of the importance of TP as a WQ indicator, its traditional measurement methods, and its role in eutrophication can be found in the supplementary materials as a continuation of Section 5.4. Generally, traditional TP measurements are time-consuming, labour-intensive and cannot illustrate the spatio-temporal scenario of TP in different waterbodies (Quinlan et al., 2021; van Wijk et al., 2024). With the advancement of RS technology, numerous studies have demonstrated that TP can be estimated using RS approaches (Du et al., 2018; Vakili and Amanollahi, 2020; Soomets et al., 2022; Yang and Jin, 2023; Zhao et al., 2023). Nevertheless, the techniques are quite complex since TP does not exhibit spectral features (Xiong et al., 2022). As a result, OA-WQ indicators such as SPM, SPOM, TSM, CHL, etc., are often used to determine the TP in waterbodies (Xiong et al., 2019).

### 5.4.1. RS data and bands (wavelength) used for TP retrieval models

A considerable amount of literature has been published on TP concentration estimation using different sensors (Table S11). Notably, the Landsat series, the Sentinel series and hyperspectral RS data are frequently used to measure TP concentrations (Fig. 5d). Additionally, the MODIS-Aqua/Terra, the SMAP, the GOCI, the IKONOS, and the Gaofen-1 were also utilized in the TP retrieval models. Fig. 5d shows the frequently used RS data to retrieve TP concentration, whereas Fig. 6 presents a summary of the satellite band spectral reflectance utilized as input parameters in ML/AI and RS-based models. As previously noted, TP is an OI-WQ indicator. However, several studies have shown that it can be retrieved using strongly correlated OA-WQ indicators (Table S11). For instance, studies by Du et al. (2018), Gholizadeh and Melesse (2017), Lu et al. (2020), Vakili and Amanollahi (2020), Wang et al. (2018), and Xiong et al. (2019, 2022) have demonstrated strong correlations between TP and WQ indicators, such as SPM, SPIM, SPOM, TSM, SST, SS, TSS, and CHL. As a result, band RI, NDI, DI, and individual band

reflectance, which are employed to calculate SPM, SPIM, SPOM, TSM, SSL, SS, TSS, and CHL, can be utilized in the TP retrieval models (Table S11).

Moreover, numerous studies have investigated the relationship between TP and different spectral reflectance's using both statistical and ML models (Table S11); however, there is no unified conclusion. For example, Karimi et al. (2023) utilized blue and green bands of the Landsat 8 OLI, whereas Vakili and Amanollahi (2020) utilized red, green and blue bands of the same satellite to retrieve TP concentrations from rivers. Additionally, some studies found a correlation between TP in rivers and the spectral range of 400 nm – 800 nm hyperspectral RS data (Hou et al., 2023; Zhang et al., 2021b; Zhang et al., 2023b). Similarly, the blue, green, and red bands of the Landsat series have shown the highest potential for estimating TP in lakes and reservoirs (Table S11). While most TP retrieval models concentrate on inland waterbodies, only a few have tried to estimate TP in coastal waters. For instance, Li et al. (2022a) identified a strong correlation between TP and spectral reflectance ranges from 440 nm to 655 nm. Similarly, Wang et al. (2018) found that 443 nm, 490 nm, 555 nm, 660 nm, and 680 nm can be used to estimate TP concentrations in Chinese coastal waterbodies.

### 5.4.2. Performance of various RS data in retrieving TP concentration

In terms of the accuracy of the satellite-derived TP concentrations, only a few researchers have published comparison of actual and predicted TP values (Fig. 7d). Fig. 7d shows the difference between in-situ and RS-derived TP concentrations in different studies. To date, numerous studies have reported notable accuracy in TP concentration retrieval models using the Landsat 8 OLI (Table S11). For example, the difference between the measured and predicted values of the Landsat 8 OLI-based TP retrieval model was approximately  $\pm 0.004$  mg/L (Chen and Quan, 2012; Li et al., 2022c). Additionally, IKONOS and hyperspectral data have shown strong performance in predicting TP concentrations in lakes and rivers (Liu et al., 2015; Sun et al., 2022). For instance, difference between measured and predicted values of the IKONOS and hyperspectral data-derived TP concentration was  $\pm 0.03$  mg/L and  $\pm 0.017$  mg/L, respectively (Fig. 7d). Furthermore, many researchers have demonstrated the efficacy of the Sentinel series products in the TP retrieval models (Li et al., 2022b; Fu et al., 2022; Padilla-Mendoza et al., 2023).

### 5.4.3. Comparison of various RS models for retrieving TP

Thus far, numerous researchers have developed TP retrieval algorithms using RS techniques (Table 11), whereas most studies relied on statistical techniques to identify spectral bands that can be used to determine TP concentrations (Xiong et al., 2019). However, EM and ML/AI methods have shifted the trend towards more accurate TP concentration estimation in waterbodies than other statistical models (Table S11). For instance, Du et al. (2018), Gholizadeh and Melesse (2017), Li et al. (2017), Lu et al. (2020), Vakili and Amanollahi (2020), Wang et al. (2018), and Xiong et al. (2019, 2022) employed an EM-based direct derivation method that uses the statistical linear or non-linear connections between the band reflectance and the in-situ P concentration to estimate the TP levels. In these studies, regression models were utilized to develop a direct retrieval method of TP concentrations from hyperspectral images, the Landsat 8 OLI and the MODIS-Aqua/Terra satellites, respectively. Consequently, these studies reported  $R^2$  values ranging from 0.58 to 0.79 (Fig. 11).

On the other hand, some studies utilized different satellites to determine the most effective RS data for TP retrieval (Table S11). For example, Liang et al. (2022) used two distinct satellites, such as the Landsat 8 OLI and the Sentinel 2 MSI, coupled with the PLSR model to develop a direct TP retrieval model, where the Sentinel 2 MSI ( $R^2 = 0.77$ ) data performed better than the Landsat 8 OLI ( $R^2 = 0.63$ ). Similarly, Li et al. (2022a) utilized the BPNN model coupling the Landsat 5 TM and the Landsat 8 OLI images to find the best algorithms for TP retrieval, in which the Landsat 5 TM showed high accuracy with  $R^2 =$

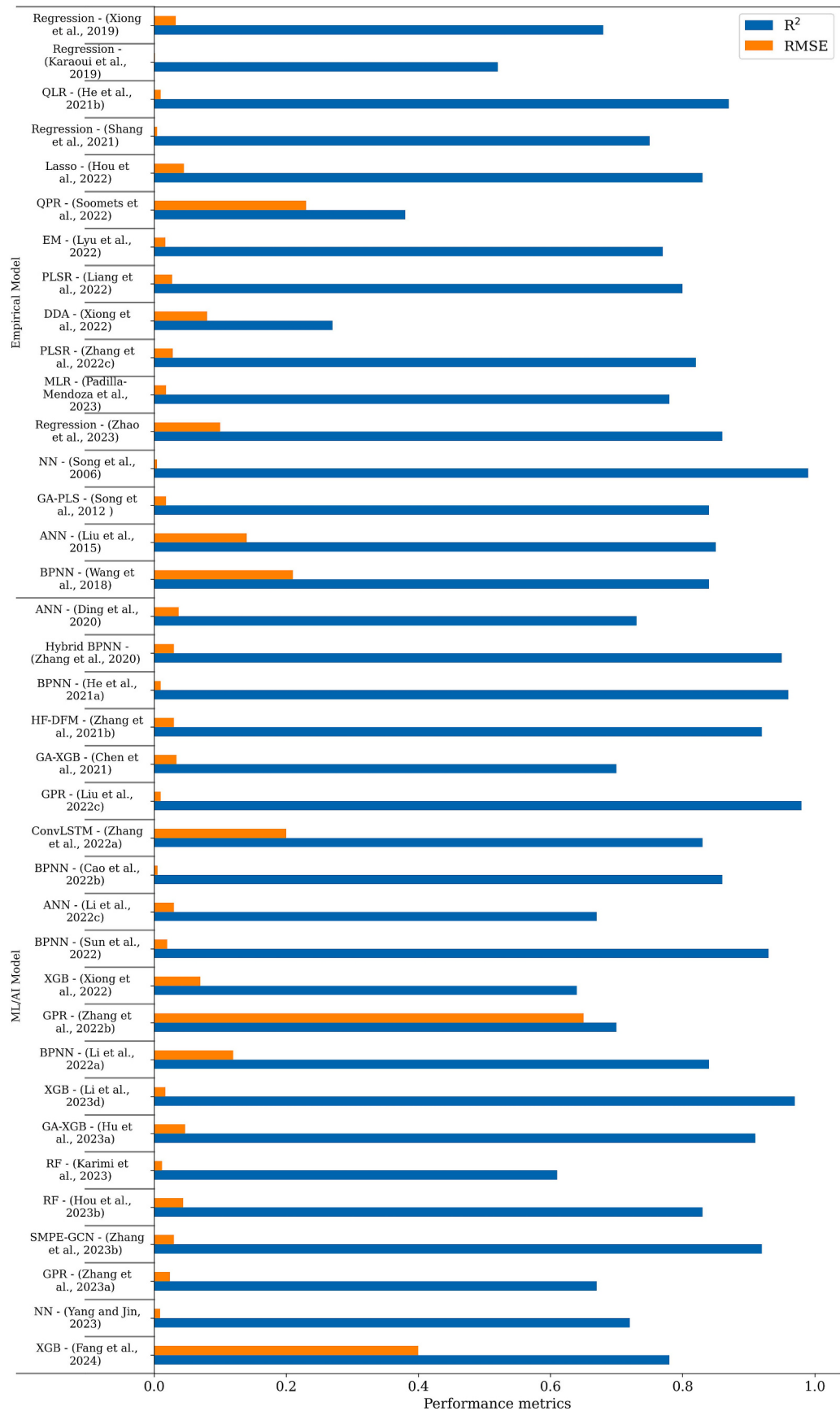


Fig. 11. Performance evaluation results of different TP retrieval models.

**Table 5**

Application of regression-based TP retrieval models.

| RS Data Type       | Waterbody Type  | Domain   | Equation                                                                                                                                                                                                                                                  | R <sup>2</sup> | Reference                     |
|--------------------|-----------------|----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|-------------------------------|
| Sentinel 2 MSI     | Wetlands        | Colombia | $TP = 1.6130 - 0.030 \times (1/B_2) - 12.1973 \times B_6 - 0.2562 \times (B_4/B_8)$                                                                                                                                                                       | 0.78           | Padilla-Mendoza et al. (2023) |
| Landsat 8 OLI      |                 |          | $\ln(TP) = 0.8 \times B_2 + 15 \times B_3 - 4.2$                                                                                                                                                                                                          | 0.5            |                               |
| Sentinel 2 MSI     | Lake            | Iran     | $\ln(TP) = 2.5 \times \left( \frac{B_3 - (B_2 + B_4 + B_{11})}{B_3 + (B_2 + B_4 + B_{11})} \right)^2 + 3.1 \times \left( \frac{B_3 - (B_2 + B_4 + B_{11})}{B_3 + (B_2 + B_4 + B_{11})} \right) - 2.4$                                                     | 0.61           | Karimi et al. (2023)          |
| Sentinel 2 MSI     | River           | China    | $TP = 0.542 \times \left( \frac{B_4 + B_5}{B_3} \right)^{6.693}$                                                                                                                                                                                          | –              | Wang et al. (2022)            |
| Landsat 8 OLI      |                 |          | $TP = 0.0782 + 0.00004 \times (B_1 - B_4) - 0.00007 \times (B_1 - B_5) - 0.0001 \times (B_2 - B_3) - 0.00001 \times (B_2 - B_4) - 0.00009 \times (B_2 - B_5) + 0.00021 \times (B_5 - B_6)$                                                                | 0.63           |                               |
| Sentinel 2 MSI     | Lake            | China    | $TP = 0.0527 + 0.0231 \times \left( \frac{B_1 + B_8}{B_1 - B_8} \right) + 0.0039 \times \left( \frac{B_2 + B_8}{B_2 - B_8} \right) - 0.0536 \times \left( \frac{B_3 + B_8}{B_3 - B_8} \right) + 0.0149 \times \left( \frac{B_4 + B_8}{B_4 - B_8} \right)$ | 0.77           | Liang et al. (2022)           |
|                    |                 |          | $TP_{Dry} (Landsat-5 TM) = 0.001(TSS) - 0.202(B_2/B_3) - 2.56(B_4) + 0.468$                                                                                                                                                                               | 0.92           |                               |
| Landsat 5 TM       | Lake            | USA      | $TP_{Dry} (Landsat-8 OLI) = 0.001(TSS) - 0.202(B_3/B_4) - 2.56(B_5) + 0.468$                                                                                                                                                                              |                | Hajigholizadeh et al. (2021)  |
| Landsat 8 OLI      |                 |          | $TP_{wet} (Landsat-5 TM) = 0.002(TSS) + 0.154(B_1/B_3) + 1.66(B_4/B_2) - 1.23(B_4/B_3) - 0.232$                                                                                                                                                           | 0.89           |                               |
|                    |                 |          | $TP_{wet} (Landsat-8 OLI) = 0.002(TSS) + 0.154(B_2/B_4) + 1.66(B_5/B_3) - 1.23(B_5/B_4) - 0.232$                                                                                                                                                          |                |                               |
| Landsat 8          | Dam             | Namibia  | $TP = -0.00309 - 8.78139(B_2) - 4.99958(B_3) + 15.37113(B_4) + -0.3916(B_5)$                                                                                                                                                                              | 0.90           | Kapalanga et al. (2021)       |
| Hyperspectral data | Lake            | China    | $TP = 0.136 \left[ \frac{R_{rs}(443) + R_{rs}(710)}{R_{rs}(575)} \right]^2 - 0.284 \left[ \frac{R_{rs}(443) + R_{rs}(710)}{R_{rs}(575)} \right] + 0.234$                                                                                                  | 0.77           | Lyu et al. (2022)             |
|                    |                 |          | $TP = -5.3248 \times \frac{\ln(B_4)}{B_4} + 0.0885 (Lake - 1)$                                                                                                                                                                                            | 0.75           |                               |
| Landsat 8 OLI      | Lake            | China    | $TP = 0.0037 \times \left( \frac{B_4 - B_3}{B_5} \right)^2 + 0.0146 \times \left( \frac{B_4 - B_3}{B_5} \right) + 0.0772 (Lake - 2)$                                                                                                                      | 0.58           | Shang et al. (2021)           |
|                    |                 |          | $TP = 0.0292 \times \left( \frac{B_5}{B_1} \right)^2 + 0.0979 \times \left( \frac{B_5}{B_1} \right) + 0.0332 (Lake - 3)$                                                                                                                                  | 0.71           |                               |
| Landsat 8 OLI      | River           | China    | $TP = 0.0577 \left( \frac{B_4 + B_6 + B_7}{3} \right)^2 + 0.707 \left( \frac{B_4 + B_6 + B_7}{3} \right) + 0.0735$                                                                                                                                        | 0.87           | He et al. (2021b)             |
| Gaofen-1           | River           | China    | $TP = -0.531B_1 - 0.9224B_2 - 5.41B_3 + 12.638B_4 + 0.4083$                                                                                                                                                                                               | 0.76           |                               |
|                    |                 |          | $TP = 0.0126COD_{Mn} + 0.00124CHL + 0.0047SS + 0.02296$                                                                                                                                                                                                   | 0.91           | Lu et al. (2020)              |
| MODIS-Aqua         | Lake            | China    | $TP = 0.2553 \times \left( \frac{B_2 - B_5}{B_2 + B_5} \right) - 0.0084$                                                                                                                                                                                  | 0.75           | Xiong et al. (2019)           |
|                    |                 |          | $TP = 2.203(R_{rs}798 + R_{rs}803)^2 + 1.903(R_{rs}798 + R_{rs}803) (Type - 1)$                                                                                                                                                                           | 0.80           |                               |
| Hyperspectral      | River and Lakes | China    | $\ln(TP) = 0.99 \ln(R_{rs}730) + 2.199 (Type - 2)$                                                                                                                                                                                                        | 0.66           | Du et al. (2018)              |
|                    |                 |          | $TP = 2.514(R_{rs}827)^{0.679} (Full data)$                                                                                                                                                                                                               | 0.45           |                               |
| Landsat 5 TM       | Reservoir       | China    | $TP = -0.023 + 0.0055 \times e^{(0.67098 \times ((B_1 + B_3 + B_4)/B_2))}$                                                                                                                                                                                | –              | Li et al. (2017)              |
| Landsat 8 OLI      |                 |          | $TP_{wet season-Landsat 5 TM} = 0.163 + 0.002 \times CHL - 0.068 \times \left( \frac{B_1}{B_2} \right) - 0.086 \times \left( \frac{B_2}{B_1} \right) + 0.002 \times \left( \frac{B_2}{B_3} \right) - 0.001 \left( \frac{B_2}{B_4} \right)$                | 0.69           |                               |
|                    |                 |          | $TP_{wet season-Landsat 8 OLI} = 0.163 + 0.002 \times CHL - 0.068 \times \left( \frac{B_2}{B_3} \right) - 0.086 \times \left( \frac{B_3}{B_2} \right) + 0.002 \times \left( \frac{B_3}{B_4} \right) - 0.001 \left( \frac{B_3}{B_5} \right)$               |                |                               |
| Landsat 5 TM       | Bay             | USA      | $TP_{Dry season-Landsat 5 TM} = -0.101 + 0.002 \times CHL + 0.022 \times \left( \frac{B_1}{B_3} \right) + 0.105 \times \left( \frac{B_3}{B_1} \right) + 0.051 \times \left( \frac{B_4}{B_1} \right)$                                                      | 0.74           | Gholizadeh and Melesse (2017) |
|                    |                 |          | $TP_{Dry season-Landsat 8 OLI} = -0.101 + 0.002 \times CHL + 0.022 \times \left( \frac{B_2}{B_4} \right) + 0.105 \times \left( \frac{B_4}{B_2} \right) + 0.051 \times \left( \frac{B_5}{B_2} \right)$                                                     |                |                               |
| Landsat 8 OLI      | River           | Korea    | $TP = 0.043 + 0.152 \times B_3 - 0.168 \times B_5$                                                                                                                                                                                                        | 0.57           | Lim and Choi (2015)           |
| Landsat 5 TM       | Lake            | China    | $TP = 0.042B_1 + 0.0162B_2 - 0.0001B_3 - 0.175$                                                                                                                                                                                                           | 0.65           | Song et al. (2006)            |

0.84 compared to the Landsat 8 OLI data. Moreover, the PLSR and MLR models also showed high accuracy in the TP retrieval process in the studies by Hajigholizadeh et al. (2021) and Padilla-Mendoza et al. (2023). The equations of various direct and indirect retrieval algorithms are presented in Table 5, while the performance of different models is shown in Fig. 11. Although, in various studies, direct retrieval models demonstrated high accuracy, the estimation mechanism is still quite complex and challenging to interpret (Xiong et al., 2019). Therefore,

researchers, such as Guo et al. (2022b), Lu et al. (2020), and Zhang et al. (2022b) used EM and various RS data to develop indirect algorithms for TP retrieval and achieved satisfactory results (Table S11).

With advances in algorithm development, computing power, and data availability, ML/AI models have been widely used for estimating TP concentrations using satellite data (Li et al., 2022a,b). Specifically, popular ML/AI algorithms, such as the GPR (Zhang et al., 2023a), XGB (Xiong et al., 2019; Yang et al., 2022c), BPNN (Guo et al., 2021a; Li

et al., 2022a; Sun et al., 2022; Wang et al., 2018; Xiao et al., 2022), ETR (Li et al., 2022b; Qiao et al., 2021), RF (Karimi et al., 2023; Li et al., 2022b), and ANN (Ding et al., 2020; Vakili and Amanollahi, 2020; Zhang et al., 2022c) have been extensively used in TP estimation along with various RS data such as the Landsat 5/7/8, the MODIS-Aqua/Terra, the Sentinel 2/3, and hyperspectral imagery (Table S11). Among these studies, around 75 % focus on retrieving TP from lakes, while 25 % focus on rivers, oceans and seas. Additionally, some researchers used hybrid/DL models, such as the SMPE-GCN (Zhang et al., 2023b), DRF (Du et al., 2018), MDL (Guo et al., 2022b), GA-XGB (Chen et al., 2021; Hu et al., 2023), LOOCV-GB (Fu et al., 2022), BP-VIP-SPCA (Wang et al., 2021), HF-DFM (Zhang et al., 2021b) and ConvLSTM (Zhang et al., 2022a), to retrieve TP concentrations from various waterbodies and achieved robust model performance (Fig. 11). Furthermore, an interesting study by Xiong et al. (2022) compared the conventional models (direct and indirect derivation methods) and ML/AI models to understand the efficacy of the optimal method for TP prediction in eutrophic lakes. The result demonstrated that the ML framework performed better than the conventional methods with  $R^2$  value of 0.64 (Table S11).

#### 5.4.4. TP algorithm optimization techniques

Similar to other OI-WQ indicators, studies that concentrate on TP model development using RS and ML/AI technologies often adopt optimization strategies to enhance model performance (Table S12). For instance, Cao et al. (2022b), Ding et al. (2020), Vakili and Amanollahi (2020), and Wang et al. (2018) employed different hidden layers, activation functions and neuron numbers to achieve high accuracy from the NN-based TP retrieval models (Table S12). In these studies, the Levenberg-Marquardt algorithm was used in the ANN model to optimize the model performance using the Landsat 8 OLI (Vakili and Amanollahi, 2020) and the IKONOS data (Liu et al., 2015). In addition to optimizing standalone models, researchers also apply various optimization techniques to hybrid models in order to enhance their accuracy (Table S12). For instance, activation functions, such as ReLU and ELU, were employed in the HF-DFM (Zhang et al., 2021b) and the MDL (Guo et al., 2022b) to achieve high accuracy. Moreover, 10-fold and 5-fold cross-validation approaches were applied by Guo et al. (2021a) and Xiong et al. (2022) to optimize model performance. Similarly, Qiao et al. (2021) set the tree number to 125 and the maximum depth size to 25 to enhance the accuracy of the ETR model.

#### 5.4.5. Limitations

- Development of TP retrieval models is limited due to lack of in-situ datasets and high cloud cover in various waterbodies.
- Lack of a generalized TP retrieval model for different waterbodies.
- Region-based TP retrieval models failed to achieve high accuracy in the testing domain.
- Only a few studies highlighted the sensitivity and uncertainty of the developed TP retrieval models.

### 5.5. Total Nitrogen (TN)

TN is the composition of nitrate-nitrogen ( $\text{NO}_3\text{-N}$ ), nitrite-nitrogen ( $\text{NO}_2\text{-N}$ ), ammonium-nitrogen ( $\text{NH}_4\text{-N}$ ), and organic nitrogen found in a water sample (LAWA, 2021). Recently, numerous studies (e.g., He et al., 2008; Li et al., 2017; Liu et al., 2015; Shang et al., 2021; Wang et al., 2022) have shown the possibilities of combining RS with ML/AI models in retrieving TN from various waterbodies as an alternative to traditional methods. A detailed description of the importance of TN as a WQ indicator, its impact on the aquatic environment, and traditional measurement methods can be found in the supplementary materials as a continuation of Section 5.5.

#### 5.5.1. RS data and bands (wavelength) used for TN retrieval models

The RS data utilized in TN retrieval are illustrated in Fig. 5e. It is

apparent from Fig. 5e that the Landsat 8 OLI satellite is most frequently utilized for TN level estimation. Moreover, several studies used other Landsat series data, such as the Landsat-5 TM and the Landsat-7 ETM+, to obtain long-term TN concentrations from different waterbodies (Guo et al., 2022b, 2022c; Zhang et al., 2022a). Similar to the Landsat data, the Copernicus RS data (e.g., Sentinel 2 MSI and Sentinel 3 OLCI) are also the most utilized RS data for TN retrieval (Table S13). Recently, hyperspectral and UAV RS data gained popularity in the TN retrieval process due to their spectral advantages (Cao et al., 2022b; Hou et al., 2023; Sun et al., 2022; Xiao et al., 2022; Zhang et al., 2023b).

Furthermore, RS and ML/AI are data-driven models that rely on various types of input parameters to predict and quantify TN levels in waterbodies. Particularly, band combinations (one-band, two-band and three-band statistics), band reflectance, and spectral indices are used as input parameters in the TN retrieval models (Cao et al., 2022b; Chen and Quan, 2012; Kapalanga et al., 2021; Sun et al., 2022; Wang et al., 2022). Additionally, WQ indicators, such as TUR, SAL, and TP, are also utilized by different researchers to develop RS-based TN models (Gholizadeh and Melesse, 2017; Wang et al., 2018). Fig. 6 shows an overview of the RS wavelength utilized to measure TN concentrations in EM and ML/AI models. Notably, several studies have identified a positive correlation between TN concentration and spectral reflectance across various wavelength ranges (Table S13). For instance, several studies demonstrated that wavelengths from 440 nm to 2200 nm (Landsat 8 OLI) are effective for lakes (Chen and Quan, 2012; Shang et al., 2021). Similarly, wavelength ranges from 480 nm to 865 nm (Landsat 8 OLI) are suitable for reservoir, dam and river waters (Kapalanga et al., 2021; Li et al., 2017; Lim and Choi, 2015; Liu et al., 2022c; Vakili and Amanollahi, 2020). Additionally, He et al. (2008) showed that wavelengths from 660 nm to 2215 nm (Landsat 5 TM) can be used for reservoirs, while Guo et al. (2021a) showed that 490 nm to 842 nm of the Sentinel 2 MSI are effective for rivers. Furthermore, numerous studies demonstrated that high-resolution RS data (e.g., hyperspectral and IKONOS) in the wavelength range of 500 nm to 900 nm can be utilized for TN retrieval in lakes and rivers (Liu et al., 2015; Sun et al., 2022).

#### 5.5.2. Performance of various RS data in retrieving TN concentration

Recent evidence from various studies indicates the feasibility of RS-based TN retrieval (Table S13); however, there is no specific framework to achieve high accuracy from the retrieval models. Fig. 7e shows the difference between in-situ and RS-derived TN concentrations across various studies. Previously, several researchers have reported high accuracy using the Landsat 8 OLI-based models (Lim and Choi, 2015; Li et al., 2017; Kapalanga et al., 2021). For example, Liu et al. (2022c) found a difference of  $\pm 0.28$  mg/L between the measured and retrieval TN values. In contrast, hyperspectral RS data accuracy varies by location. Particularly, Zhang et al. (2023b) observed that their UAV hyperspectral-based retrieval model overestimated TN concentrations, whereas Sun et al. (2022) noted their hyperspectral-based retrieval model underestimated TN concentrations. Additionally, high-resolution-based RS data demonstrated high accuracy in the TN retrieval model in urban narrow rivers (Table S13). The difference between measured and predicted data from the IKONOS-based TN retrieval model was  $\pm 0.02$  mg/L (Liu et al., 2015).

#### 5.5.3. Comparison of various RS models for retrieving TN

TN is an OI-WQ indicator, but considerable evidence indicates that RS-based models can be utilized to quantify TN levels because it is significantly correlated with OA-WQ indicators, such as CHL, SPM, CDOM, etc. (Liu et al., 2022c; Soomets et al., 2022; Vakili and Amanollahi, 2020). Furthermore, several published studies (Cao et al., 2022b; Chen and Quan, 2012; Guo et al., 2021a; Sun et al., 2022) have explored the relationship between RS reflectance and TN levels. Most of the RS-based TN retrieval empirical algorithms (direct and indirect) rely on regression-based approaches (Table S13). These include models such as simple LR (Chen and Quan, 2012), PCR (Wang et al., 2022), lassoR

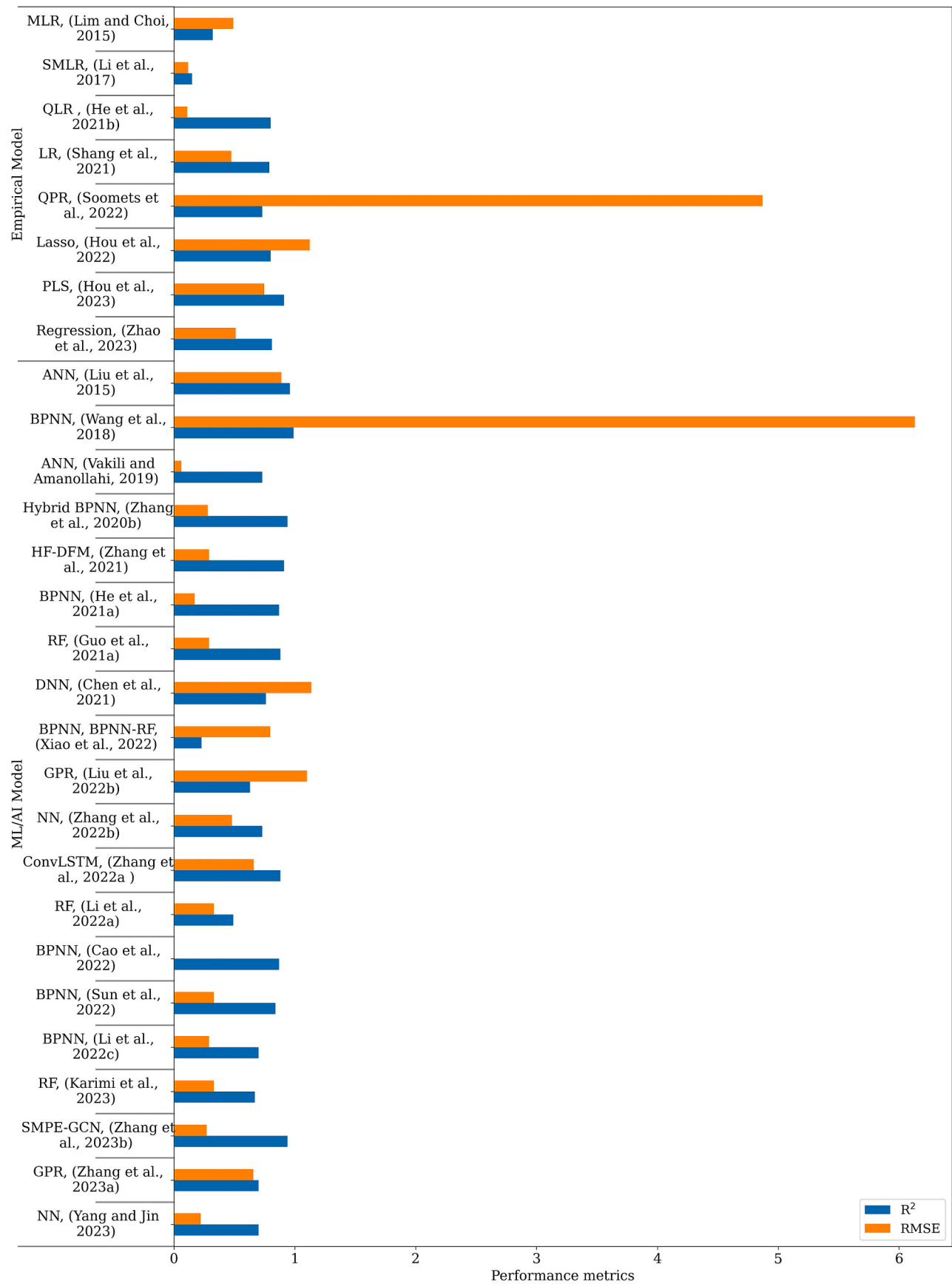


Fig. 12. Performance evaluation results of different TN retrieval models.

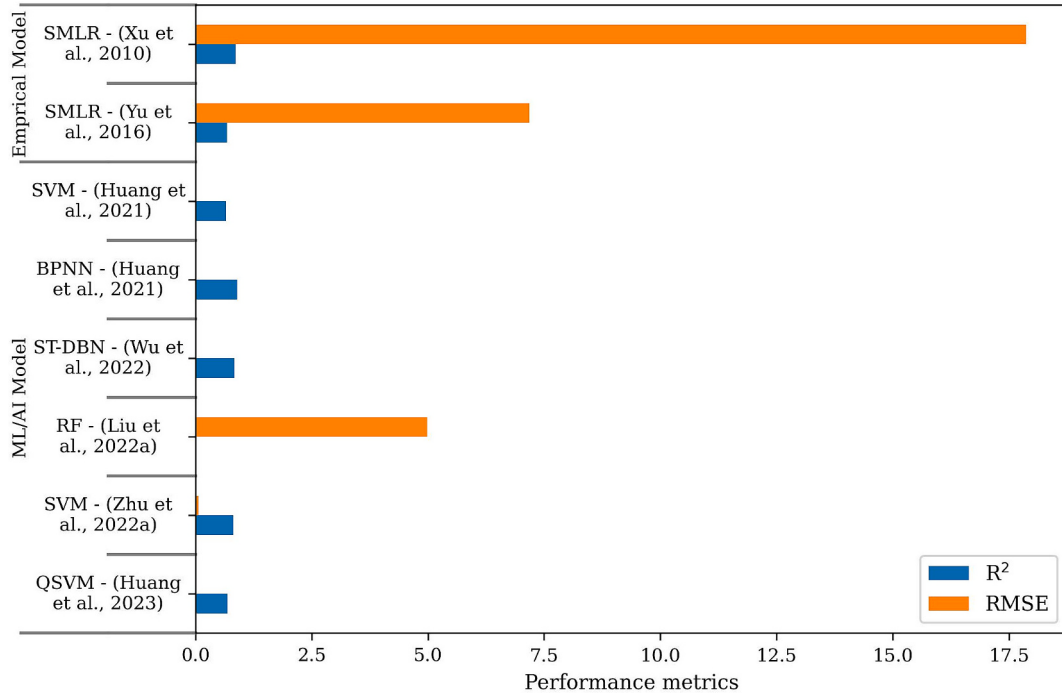
**Table 6**

Application of regression-based TN retrieval models.

| RS Data Type                  | Waterbody Type | Domain  | Equation                                                                                                                                                                                                                                | R <sup>2</sup> | Reference                     |
|-------------------------------|----------------|---------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|-------------------------------|
| Landsat 8 OLI                 | Lake           | Iran    | $\ln(TN) = -6.3 * \left( \frac{B_5 - B_7}{B_5 + B_7} \right)^2 + 0.9 \left( \frac{B_5 - B_7}{B_5 + B_7} \right) + 0.12$                                                                                                                 | 0.67           | Karimi et al. (2023)          |
| Sentinel 2 MSI                |                |         | $\ln(TN) = 0.4 * \left( \frac{B_3 - (B_5 + B_6 + B_7)}{B_3 + (B_5 + B_6 + B_7)} \right)^2 + \left( \frac{B_3 - (B_5 + B_6 + B_7)}{B_3 + (B_5 + B_6 + B_7)} \right) - 0.12$                                                              | 0.84           |                               |
| Sentinel 2 MSI                | River          | China   | $TN = 0.066 * \left( \frac{B_6}{B_8} \right)^{11.165}$                                                                                                                                                                                  | –              | Wang et al. (2022)            |
| Landsat 8 OLI                 | Dam            | Namibia | $TN = 1.047532 - 54.928(B_2) - 46.2947(B_3) + 120.8943(B_4) - 19.223(B_5)$                                                                                                                                                              | 0.79           | Kapalanga et al. (2021)       |
| Landsat 8 OLI                 | Lake           | China   | $TN = 0.233 * (B_4 - B_1)^2 + 1.2714 * (B_4 - B_1) + 1.3499 \text{ (Lake - 1)}$                                                                                                                                                         | 0.79           | Shang et al. (2021)           |
|                               |                |         | $TN = 0.5914 * (B_5 + B_2) + 1.1997 \text{ (Lake - 2)}$                                                                                                                                                                                 | 0.65           |                               |
|                               |                |         | $TN = -3.219 * (B_3 - B_7) + 5.712 \text{ (Lake - 3)}$                                                                                                                                                                                  | 0.75           |                               |
| Landsat 8 OLI                 | Reservoir      | China   | $TN = 0.11608 + 2.716 * \left( \frac{B_4}{B_2 + B_5} \right)^{1.26477}$                                                                                                                                                                 | 0.58           | Li et al. (2017)              |
| IKONOS                        | River and lake | China   | $TN = -276.02 * (B_2) + 54.56$                                                                                                                                                                                                          | 0.88           | Liu et al. (2015)             |
| Landsat 5 TM<br>Landsat 8 OLI | Bay            | USA     | $TN_{dry\ season - Landsat\ 8\ OLI} = -4.51 - 5.18 * (TP) - 0.01 * TUR + 0.68 * \left( \frac{B_2}{B_3} \right) + 0.64 * \left( \frac{B_3}{B_4} \right) + 1.67 * \left( \frac{B_3}{B_2} \right) + 3.02 * \left( \frac{B_4}{B_2} \right)$ | 0.69           | Gholizadeh and Melesse (2017) |
|                               |                |         | $TP_{dry\ season - Landsat\ 5\ TM} = -4.51 - 5.18 * (TP) - 0.01 * TUR + 0.68 * \left( \frac{B_1}{B_2} \right) + 0.64 * \left( \frac{B_1}{B_3} \right) + 1.67 * \left( \frac{B_2}{B_1} \right) + 3.02 * \left( \frac{B_3}{B_1} \right)$  |                |                               |
|                               |                |         | $TP_{wet\ season - Landsat\ 8\ OLI} = 1.35 + 25.6 * (TP) + 0.05 * CHL - 12.93 * B_4 + 25.93 * B_5 - 2.64 * \left( \frac{B_5}{B_4} \right)$                                                                                              | 0.74           |                               |
|                               |                |         | $TP_{wet\ season - Landsat\ 5\ TM} = 1.35 + 25.6 * (TP) + 0.05 * CHL - 12.93 * B_3 + 25.93 * B_4 - 2.64 * \left( \frac{B_4}{B_3} \right)$                                                                                               |                |                               |
| Landsat 8 OLI                 | River          | Korea   | $TN = 2.89 - 20.054 * (B_3) + 15.137 * (B_4) + 8.257 * (B_5)$                                                                                                                                                                           | –              | Lim and Choi (2015)           |
| Landsat 5 TM                  | Lake           | China   | $TN = -275.26(B_1) - 6.85(B_2) + 224.43(B_3) + 7.86(B_4) + 3.48$                                                                                                                                                                        | 0.24           | Chen and Quan (2012)          |
| Landsat 5 TM                  | Reservoir      | China   | $\ln(TN) = 4.3907 - \frac{177.9298}{B_6} - 0.1362 \left( \frac{B_6}{B_7} \right) - 1.211 \left( \frac{B_3}{B_6} \right)$                                                                                                                | 0.56           | He et al. (2008)              |

(Hou et al., 2022), SMLR (Guo et al., 2022c; Li et al., 2017), BR (Yuan et al., 2022), MLR (He et al., 2008; Lim and Choi, 2015), and QLR (He et al., 2021b; Soomets et al., 2022). The performance of these models is shown in Fig. 12, while the derived equations are summarized in

Table 6. It is apparent from Table 6 that the Landsat 8 OLI-based regression model showed high accuracy ( $R^2 = 0.58$  to  $0.80$ ) in Chinese waterbodies compared to the Landsat 5 TM-based regression model ( $R^2 = 0.24$  to  $0.56$ ). Moreover, a study by Liu et al. (2015) used the

**Fig. 13.** Performance evaluation results of different DIN inversion models.

Sentinel 2 MSI and the IKONOS satellite-based regression model, where the IKONOS showed high accuracy with  $R^2 = 0.88$  (Liu et al., 2015). Similarly, the Sentinel 3 OLCI-based TN retrieval regression model exhibited a high accuracy with  $R^2$  value of 0.73, but the RMSE (4.87) value is quite high compared to other regression models (Table S13).

It should be acknowledged that in terms of accuracy, the ML/AI-based models performed better than the regression models (Fig. 12). For instance, NN-based models, such as BPNN (Cao et al., 2022b; He et al., 2021a; Xiao et al., 2022), NN (Zhang et al., 2022a) and ANN (Vakili and Amanollahi, 2020), performed notably better with a  $R^2$  value ranging from 0.96 to 0.73. Although Wang et al. (2018) reported a BPNN model with a high RMSE (6.13) compared to other BPNN models (Cao et al., 2022b; Sun et al., 2022), but it showed high accuracy with a  $R^2$  value of 0.99. In comparison, when evaluating various RS data, the RF model exhibited better performance ( $R^2 = 0.62$  to  $0.88$ ) with the Sentinel 2 MSI (Guo et al., 2021a; Li et al., 2022b) than the Landsat 8 OLI ( $R^2 = 0.49$ ) for TN estimation. Moreover, the GPR model showed higher accuracy ( $R^2 = 0.84$ ) with hyperspectral RS data than the Landsat 8 OLI ( $R^2 = 0.63$ ) in TN estimation (Liu et al., 2022c). In addition to standalone models, numerous studies have assessed the efficacy of various hybrid/DL models in TN retrieval (Table S13). For instance, Guo et al. (2022b) utilized the Landsat series data coupled with the MDL model to estimate long-term TN in lakes and achieved excellent results compared to other DL, ML/AI and EM algorithms (Table S13). Additionally, hybrid models, such as the SMPE-GCN (Zhang et al., 2023b), GA-XGB (Chen et al., 2021), BPNN-RF (Fu et al., 2022), and LOOCV-GB (Fu et al., 2022), showed high accuracy than standalone models in the RS-based TN retrieval processes (Table S13).

#### 5.5.4. TN algorithm optimization techniques

Several studies have considered ML/AI algorithm parameter tuning to enhance accuracy and prediction results (Table S14). Table S14 provides the hyperparameters discussed in different studies. For instance, to optimize the performance of the BPNN model, Cao et al. (2022b) and He et al. (2021a) set the hidden layer to 5 and 10 neuron nodes, respectively, while setting the input layer to 1221 and 7 nodes, respectively. Similarly, a high processing power-based Levenberg-Marquardt training algorithm was employed in the ANN model to achieve high accuracy in TN estimation models utilizing the Landsat 8 OLI and the IKONOS imagery (Liu et al., 2015; Vakili and Amanollahi, 2020). Furthermore, in the GPR model, the maximum likelihood kernel function is used as a hyperparameter value because this function performed better with medium-sized datasets (Liu et al., 2022c). Additionally, Guo et al. (2022b, 2021a) and Li et al. (2022a) showed the effectiveness of hyperparameter tuning in model performance (Table S14).

#### 5.5.5. Limitations

- The efficacy of TN retrieval models is constrained by the range of TN concentrations used in the model training, leading to high uncertainties when these models are applied to different waterbodies with varying TN concentrations.

- Most researchers have not focused on the physical and chemical processes of waterbodies, which produced high uncertainties in TN retrieval models.
- Existing studies show that publicly accessible RS data such as the Landsat, the Sentinel, etc. are not suitable for retrieving TN concentrations from narrow waterbodies.
- Accuracy of the TN retrieval model is constrained by the limited sample size.
- No optimal RS wavelengths have been suggested by any studies in order to retrieve TN from different waterbodies.
- Lack of comparative studies of different RS data for retrieving TN from different waterbodies.

### 5.6. Dissolved inorganic nitrogen (DIN)

DIN is comprised of nitrite ( $\text{NO}_2$ ), nitrate ( $\text{NO}_3$ ) and ammonia ( $\text{NH}_3$ ) (Louis et al., 2015). However, obtaining continuous long-term DIN concentration data is costly (Wu et al., 2022). Furthermore, researchers have noted that, due to their complex nature, nutrient-based numerical models are also unsuitable for DIN prediction (Liu et al., 2022a). In comparison, RS-based technology is becoming an alternative solution for monitoring DIN concentration across larger waterbodies (Huang et al., 2021). A detailed description of DIN, its impact on aquatic ecosystems, and its measurement approaches can be found in the supplementary materials as a continuation of Section 5.6.

#### 5.6.1. RS data and bands (wavelength) used for DIN retrieval models

There are relatively few articles that are involved with DIN concentration estimation using ML/AI and RS imagery (Table S15). Generally, WQ indicators such as TEMP, SAL, and CHL are often used to estimate DIN concentrations in combination with RS data (Huang et al., 2021; Liu et al., 2022a; Wu et al., 2022; Xu et al., 2010; Yu et al., 2016). Moreover, most of the current literature on RS-based DIN models investigates long-term trends in DIN concentrations in various waterbodies (Table S15). Therefore, RS data such as the Aqua/Terra (MODIS) and the Landsat series, which have continuity in constant satellite observation and historical data records, are used for DIN estimation (Huang et al., 2021; Liu et al., 2022a; Wu et al., 2022; Xu et al., 2010; Yu et al., 2016). Fig. 5f shows different RS data utilized in various studies, whereas Fig. 6 presents the wavelength utilized to assess DIN concentrations in various waterbodies. For instance, many recent studies have identified a positive correlation between DIN concentration and the MODIS-Aqua/Terra B1-B7 (655 nm – 2130 nm) bands to retrieve time series data of DIN from saline waterbodies (Wu et al., 2022; Xu et al., 2010; Yu et al., 2016). In contrast, coastal, red, green, blue, and NIR bands of the Landsat series satellites are frequently utilized to retrieve DIN concentrations from coastal waterbodies (Huang et al., 2021; Zhu et al., 2022b).

#### 5.6.2. Performance of various RS data in retrieving DIN concentration

A comparison between in-situ and RS-retrieved DIN concentrations from different RS data is presented in Fig. 7f. Among the RS data, the MODIS-Aqua satellite showed high accuracy in retrieving DIN

**Table 7**  
Application of regression-based DIN retrieval models.

| RS data type       | Waterbody type | Domain | Equation                                                                                                                               | $R^2$ | Reference        |
|--------------------|----------------|--------|----------------------------------------------------------------------------------------------------------------------------------------|-------|------------------|
| MODIS-Aqua Level 2 | Sea            | China  | $DIN_{Entire\ Bohai\ Sea} = -69.562 + 0.008 * F_{20}(8, 10) + 0.044 * F_{20}(1, 20)$                                                   | 0.68  | Yu et al., 2016  |
|                    |                |        | $DIN_{Bohai-lai\ zhou\ Bay} = -4.514 + 0.011 * F_{20}(8, 10) - 0.159 * F_{20}(3, 4) + 0.008 * F_{20}(14, 21) - 0.001 * F_{20}(12, 15)$ | 0.82  |                  |
|                    |                |        | $DIN_{B\ Liaodong\ Bay} = 30.960 + 0.000009199 * F_{20}(11, 12)$                                                                       | 0.99  |                  |
|                    |                |        | $DIN_{Inner\ sea} = 13.532 + 0.160 * F_{20}(1, 5) - 0.007 * F_{20}(8, 19)$                                                             | 0.79  |                  |
| MODIS (Aqua/Terra) | Sea            | China  | $DIN_{all} = 7.86 \left( \frac{R_3 + R_4}{R_3 - R_4} \right) + 9.83$                                                                   | 0.87  | Xu et al. (2010) |

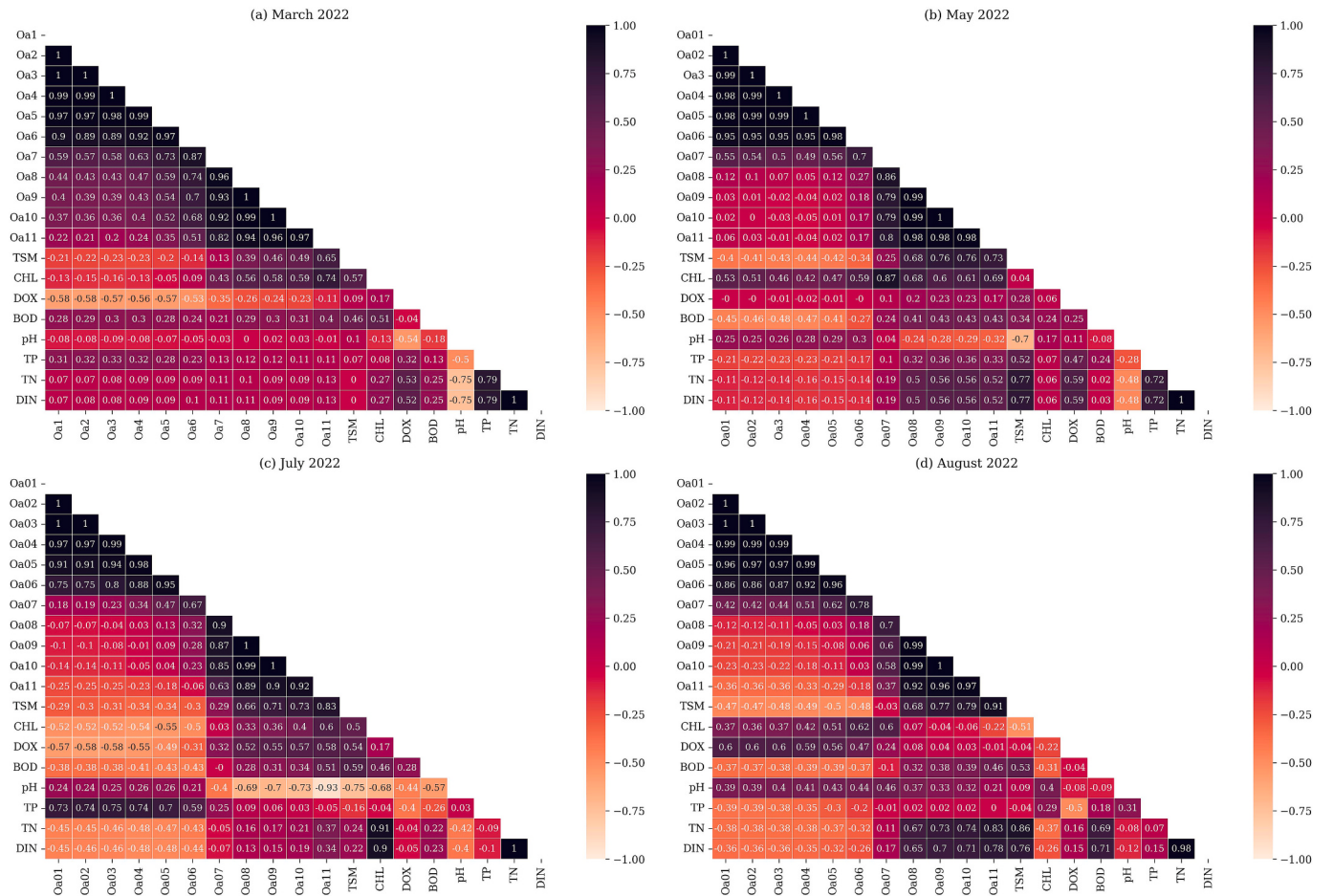


Fig. 14. Relationship between Sentinel 3 OLCI bands, in-situ and RS-derived WQ indicators.

concentrations compared to the Landsat 8 OLI. For example, the difference between in-situ measured and retrieved DIN concentrations was  $\pm 0.63$  mg/L for the Landsat 8 OLI, and whereas it was only  $\pm 0.13$  mg/L for the MODIS-Aqua/Terra-based DIN retrieval model (Huang et al., 2021; Wu et al., 2022).

### 5.6.3. Comparison of various RS models for retrieving DIN

Nutrients, such as DIN, are OI-WQ indicators. In order to estimate DIN concentrations using satellite data, many studies have emphasized the development of empirical models that correlate DIN value with OI-WQ indicators such as CHL, CDOM, and TSS (Yu et al., 2016). Thus far, there has been notable progress made in quantifying DIN values in waterbodies using RS technology. For instance, combinations of visible and near-infrared bands from the MODIS-Aqua/Terra data have been analysed with the SMLR model to determine DIN concentrations in China (Xu et al., 2010; Yu et al., 2016). The performance of the different DIN retrieval models is illustrated in Fig. 13, whereas the DIN retrieval model equations are presented in Table 7. The  $R^2$  value of these studies range between 0.68 and 0.99 (Fig. 13).

Table S15 represents the WQ indicators and RS parameters utilized to develop the DIN retrieval models. It can be seen from Table S15 that ML/AI models provide a sophisticated solution to develop the DIN retrieval models using RS technology. For example, Zhu et al. (2022b) developed a DIN retrieval model for Yueqing Bay, China, by employing the SVM model. In this study, time series of in-situ (e.g., TEMP and DIN) and the Landsat 8 OLI satellite data (e.g., B2, B5, and B6 bands) were used as input data. Consequently, this model showed high accuracy with a  $R^2$  value exceeding 0.80 for both the training and validation phases. Furthermore, this study examined the spatio-temporal fluctuations and

key influential factors of DIN concentrations in the bay over the past seven years (from 2013 to 2020). Similarly, Liu et al. (2022a) established a ML approach to reconstruct monthly sea-surface DIN and other nutrients in the Chinese sea waters from 2003 to 2019. In this study, a large dataset of in-situ and environmental data (including SST, SAL, CHL, and Kd490) was utilized to train ML/AI models (SVR, RF, and ANN models) in order to select the optimal ML/AI model and accurately predict DIN concentrations (Table S15). The study identified that changes in SST and SAL influenced the surface DIN concentration differently in nearshore and offshore waters. Prior to these studies, Huang et al. (2021) and Wu et al. (2022) utilized the ML/AI and RS reflectance data to retrieve DIN concentrations for estuary and bay waters. Particularly, Huang et al. (2021) employed two ML/AI models (BPNN and SVM) to develop DIN retrieval models and track 32-years (from 1988 to 2020) of changes in DIN concentration at Shenzhen Bay, China. In this study, the BPNN ( $R^2 = 0.90$ ) based model performed better than the SVM model ( $R^2 = 0.66$ ). Moreover, DL models also performed better in RS-based DIN retrieval. For example, Wu et al. (2022) proposed a ST-DBN model that reduces estimation errors above 40 % compared to other ML/AI models, such as MLR, MPNN, GRNN, DBN, ST-MLR, ST-MPNN, ST-GRNN, and ST-DBN.

### 5.6.4. DIN algorithm optimization techniques

Numerous published studies, including Huang et al. (2021), Liu et al. (2022a), and Wu et al., 2022) have explored optimal hyperparameter values to retrieve better accuracy from the DIN retrieval models (Table S16). For example, Huang et al. (2021) employed the RF algorithm with 500 trees to develop a Landsat series RS-based DIN retrieval model. Moreover, the performance of ML models can be estimated using a cross-

validation approach (Sajib et al., 2023). Specifically, Wu et al. (2022) utilized a 10-fold cross-validation approach to assess the model's performance. Similarly, Liu et al. (2022a) set 4-10 hidden layers and Levenberg-Marquardt activation functions to achieve high performance from the BPNN-based DIN retrieval model (Table S16).

### 5.6.5. Limitations

- Lack of performance evaluation of different satellites based on DIN retrieval results.
- The accuracy of DIN retrieval models was challenging due to limited validation datasets.
- Lack of uncertainty determination of DIN retrieval models.
- Lack of recommended RS wavelengths for determining DIN concentrations from different waterbodies.
- Lack of attention to the spatio-temporal variability of DIN concentrations when developing DIN retrieval models.

### 5.7. Relationship between RS bands and in-situ WQ indicators

In order to develop RS-driven retrieval model(s), it is essential to explore the interactions and relationships among WQ indicators with different RS bands and their combinations (Diganta et al., 2024; Sajib et al., 2024). To the best of the author's knowledge, most RS-driven retrieval model(s) have been developed based on the correlation between specific WQ indicators and various RS bands. For instance, numerous studies have employed these approaches to assess OI-WQ indicators, such as DOX (Kim et al., 2020; Sharaf El Din and Zhang, 2017), pH (Mohandas and Brema, 2023; Rahul and Brema, 2023), and TN (Chen and Quan, 2012; Isenstein and Park, 2014). However, several studies indicate that these models often exhibit poor prediction when employed in different regions (Cao et al., 2023; Hajigholizadeh et al., 2021). Nevertheless, a number of studies successfully utilized similar processes along with different band combinations to develop various OI-WQ retrieval models and obtained high accuracy (Arias-Rodriguez et al., 2023; Du et al., 2022; Jiang et al., 2022; Lyu et al., 2022; Padilla-Mendoza et al., 2023; Qiao et al., 2021; Soomets et al., 2022).

In this study, we explored the relationship between different multi-spectral satellite surface reflectance's and OI-WQ indicators using Pearson correlation. As a case study, we selected Cork Harbour, Ireland, for its dynamics and anthropogenic pressures. Details of the utilized in-situ and satellite data are available in Table S17 and Table S18. The correlation assessment revealed that there is no notable relationship between in-situ pH and surface reflectance of the Landsat 8/9 OLI and the Sentinel 3 OLCI, except for the Sentinel 2 MSI bands B1 to B8A ( $r \geq 0.5$ ) (Fig. 14; Fig. S1-S2). For the Sentinel 3 OLCI, in-situ DOX concentration showed a strong negative correlation with Band 1 to Band 6 in March and July 2022 (Fig. 14a, c) and a strong positive correlation ( $r \geq 0.5$ ) was shown in August 2022 (Fig. 14d). However, no notable relationship was found in May 2022 (Fig. 14b). Similarly, Bands B3 to B8A of the Sentinel 2 MSI showed a strong positive relationship with DOX in July 2022 ( $r \geq 0.5$ ), whereas a strong negative relationship was found for Band B1 to B5 in March 2022 (Fig. S1). Additionally, DOX concentration showed a strong positive relationship with the Landsat 8/9 OLI Bands B1 to B5 in February and March 2022 (Fig. S2).

Furthermore, similar to pH, in-situ BOD<sub>5</sub> showed no notable relationship with the Sentinel 3 OLCI and the Sentinel 2 MSI bands (Fig. 14; Fig. S1); however, an exception was observed for B1 to B3 bands of the Landsat 8/9 OLI (Fig. S2). Several studies utilized the correlation process with RS-based surface reflectance with CHL and TSM to estimate nutrients like TP, TN, and DIN (Gao et al., 2015; Hajigholizadeh et al., 2021; Lu et al., 2020; Peterson et al., 2020). Moreover, a very strong relationship ( $r \geq 0.7$ ) with in-situ TN and DIN shows for TSM obtained from the Sentinel 3 OLCI (May and August 2020) bands Oa08 to Oa11 (Fig. 14). Similarly, a strong relationship ( $r \geq 0.5$ ) was found for the Sentinel 2 MSI and the Landsat 8/9 OLI bands B1 to B3 (See Fig. S1; Fig.

S2). On the other hand, TP did not show any notable relationship with any of the bands. In summary, the relationship between the OI-WQ indicator and surface reflectance of RS data varies over time and space.

## 6. Discussion

Since the development of the RS technique, its application has increased tremendously in various fields, including water resource management and aquatic environmental monitoring. As a result, numerous successive applications of the RS technique exist in the field of OA-WQ indicators, such as TSM, CHL, CDOM, etc. (Ansper-Toomsalu et al., 2024; Concha et al., 2021; Diganta et al., 2024; Frouin et al., 2019; Gleratti et al., 2024; González Vilas et al., 2023; Kutser, 2012; Pahlevan et al., 2021a; Vanhellemont and Ruddick, 2018; Zibordi et al., 2022). In contrast, there is a lack of successive retrieval models for OI-WQ indicators. To address this knowledge gap, the study comprehensively reviews the current global trends and research progress in using RS technologies for retrieving OI-WQ indicators, such as pH, DOX, BOD<sub>5</sub>, TP, TN, and DIN across various waterbodies. The current research analysed 105 research papers based on OI-WQ indicators to identify the various RS data utilized in different waterbodies, including rivers, lakes, TrC waters, etc. Additionally, the study summarizes various retrieval models, their limitations and explores the potential of available tools and techniques to improve the existing general monitoring programs.

Typically, in order to maintain "Good" WQ status in all forms of waterbodies within the EU regions, it is required to regularly monitor eleven WQ indicators (WFD, 2000). Among them, pH is one of the most important chemical properties of water. According to NOAA (2020), ocean acidification is significantly affecting fish and seaweed habitats. Therefore, several studies have focused on developing pH retrieval algorithms (Batur and Maktav, 2019; Pereira et al., 2020; Qian et al., 2022). The shortcomings of these retrieval models are that they are not suitable for long-term prediction and require more parameterizations and algorithm tuning based on region. Additionally, the uncertainty of these retrieval models remains unquantified. In the pH retrieval process, compared to ML/AI-based models, EM techniques showed high accuracy in small-sized waterbodies (Krishnaraj and Honnasiddaiah, 2022; Mohandas and Brema, 2023; Pereira et al., 2020). However, the accuracy level decreases as the size of the waterbodies increases. Furthermore, regarding satellite wavelengths for retrieving pH from different waterbodies, no standardized wavelength has been suggested by researchers. On the other hand, in terms of RS data performance, high-resolution satellite data (e.g., Gokturk-2) and hyperspectral data (ground and ship-borne) showed the utmost performance in pH retrieval (Batur and Maktav, 2019; Cao et al., 2022a; Rahul and Brema, 2023). Nevertheless, it should be noted that in most cases, processing and acquiring ground and shipborne hyperspectral data is quite complex and costly.

In the case of DOX-based studies, the current review identified a lack of consideration of the spatio-temporal variability of DOX in retrieval models. As a result, the uncertainty of these retrieval model(s) can vary seasonally. To address this issue and mitigate model(s) uncertainties, especially seasonal variability of retrieval results, a few studies have recommended considering the wavelength range of 400 nm–800 nm for hyperspectral RS data (Li et al., 2023d; Yang et al., 2023). However, no specific wavelength for multi-spectral images was suggested or proposed in any studies. Regarding the retrieval models, the regression-based models yielded more favourable outcomes than ML/AI algorithms (Batur and Maktav, 2019; Karaoui et al., 2019; Mohandas and Brema, 2023; Sharaf El Din and Zhang, 2017). Additionally, most DOX retrieval models were developed based on regional aspects. As a result, it is quite complex and challenging to utilize these models in a global context. According to the basic principles of the RS technology, numerous physical and anthropogenic factors affect the accuracy of the satellite sensors. Therefore, it is assumed that the RS technology may not be suitable across the globe due to the imbalance of the physical attributes

of the atmosphere (Arias-Rodriguez et al., 2023). Despite these challenges, several recent studies have emphasized that it is essential to develop a DOX retrieval model focused on global configurations that can be utilized across different geographical contexts (Arias-Rodriguez et al., 2023; Sajib et al., 2024). Furthermore, it is also important to acknowledge that most of the developed DOX algorithms have not been validated across different geographical regions. Therefore, when applying these algorithms in other waterbodies, it is necessary to recalibrate the model using different datasets specific to the local region.

In the BOD<sub>5</sub> retrieval process, DL models showed high accuracy compared to other standalone ML/AI models. However, the primary drawbacks of DL models include a slow training rate, their complex nature, overfitting and underfitting issues (Arias-Rodriguez et al., 2023; Zhang et al., 2021b; Zhang et al., 2022a). In the BOD<sub>5</sub> retrieval process, ground-based and UAV hyperspectral data performed better than multi-spectral satellite data. Nevertheless, this type of hyperspectral data requires preprocessing with domain expertise, and representative samples are essential for effective model training in order to capture spectral features accurately (Zhang et al., 2023b). Furthermore, numerous studies have hypothesized that it is possible to estimate BOD<sub>5</sub> based on its relationship with OA-WQ indicators (Fu et al., 2022; Zhang et al., 2022a; Zhang et al., 2023b). However, no studies have yet assessed BOD<sub>5</sub> using OA-WQ indicators. Additionally, a notable limitation of RS-based BOD<sub>5</sub> studies is that most focus on smaller waterbodies while large waterbodies, such as estuaries, coastal waters, or seas, remain unexplored.

In the case of TP, a number of models have been developed to retrieve TP from different waterbodies (Table S11). However, no researchers have suggested any ideal wavelength range to retrieve TP concentrations. Additionally, several studies hypothesize that there is a strong correlation between TP and OA-WQ indicators (Fang et al., 2024; Yuan et al., 2022; Zhang et al., 2023b), but there is a lack of studies utilizing OA-WQ indicators to develop RS-based TP retrieval algorithms. A notable drawback for developing TP algorithms for the sea is a limitation of in-situ monitoring datasets (Soomets et al., 2022). Moreover, most of the ML/AI-based TP models showed high accuracy in retrieving TP across different waterbodies (Fig. 11). However, these models often fail to provide an optimal result when applied in other waterbodies. Moreover, while DL models demonstrated high accuracy in TP retrieval models (Fig. 11); they are highly susceptible to overfitting problems due to their complex nature.

In RS-based TN retrieval studies, ML/AI and DL algorithms demonstrated higher accuracy than EM techniques (Fig. 12). However, the limited size of training and validation datasets limits the use of ML/AI and DL models in RS-based TN retrieval. Additionally, the limited sample size introduces significant uncertainty into the model results. Although most developed models exhibited high accuracy in the local domain(s) but requires additional band selection and hyperparameter tuning when applied to the global domain(s). Furthermore, the output TN inversion models were constrained by the TN concentration range, and there were notable uncertainties observed when applying the developed model(s) to a global scale. Additionally, frequently used satellite data, such as the Landsat series, the Sentinel series, etc., have failed to achieve high accuracy for narrow waterbodies (Liu et al., 2015). Finally, the complexity of the physical and chemical processes of waterbodies further contributes to notable uncertainty in estimating TN concentration using RS-based retrieval model(s).

For the purposes of DIN retrieval, several researchers utilize various band combinations from different RS data (Table S15). However, these models are constrained within the seawater. Additionally, a significant drawback of the DIN retrieval models is that most of the developed models are based on the Chinese marine waters. As a result, these model(s) are not suitable for global applications. Furthermore, in the DIN retrieval process, retrieval models were developed using continuous satellite observation and historical data records (Huang et al., 2023; Huang et al., 2021; Xu et al., 2010). Nevertheless, the comparative

performance of various satellites in terms of DIN retrieval results has not been explored in any study.

The above discussion highlights a continuum of uncertainty factors in developing RS-based retrieval model(s) for various OI-WQ indicators. To enhance the reliability of these retrieval models, future studies should consider the holistic approaches that incorporate EM with advanced ML/AI techniques. Furthermore, for the purposes of the global configuration of these retrieval models, the developed algorithms should be tested and validated using completely independent datasets that represent the global attributes of WQ. Therefore, to develop global model(s), the researchers should not only focus on the tools, techniques and RS data but also consider the quality of RS and WQ data in terms of the geospatial-temporal resolution of domains. To address the challenges outlined above, future studies could incorporate the following considerations.

#### 6.1. Improve the performance of retrieval algorithms in terms of model sensitivity and uncertainty

Several factors may contribute to high uncertainty in RS-based retrieval models. These factors include the physico-chemical processes of waterbodies, the insignificant performance of retrieval models, atmospheric-derived errors in RS data, non-continuous datasets, and the low temporal resolution of RS data. As a result, various studies have observed the occurrence of both over- and underestimation of OI-WQ indicator concentrations (Zhang et al., 2022b; Sun et al., 2022). To address these challenges, continuous atmospherically corrected RS datasets, such as Level 3 and Level 4 data of the Sentinel 3 OLCI and the Multi-Sensors (comprised of SeaWiFS, MERIS, MODIS-Aqua/Terra, VIIRS-SNPP, JPPS1, Sentinel 3 A OLCI & Sentinel 3B OLCI) from the Copernicus Marine Service (<https://marine.copernicus.eu/>) can be utilized in developing models. Among these, Level 3 is a daily atmospherically corrected grid-based RS data, whereas Level 4 is monthly interpolated atmospherically corrected data. A detailed description of these data can be found in Garnesson et al. (2024) and Colella et al. (2024). The availability of continuous RS data will not only contribute to developing a robust model but also to assessing its reliability with different datasets from other regions. Additionally, researchers can assess the model sensitivity and uncertainty by using independent datasets during model validation. This type of framework may help the researchers to develop a novel retrieval algorithm for each OI-WQ indicator on a global scale with high sensitivity.

#### 6.2. Develop OI-WQ retrieval indicators model for transitional and coastal (TrC) waters

Estuaries or TrC waters are complex ecological systems that act as the last filter for the transfer of dissolved and particulate matter between freshwater and the open coast (Uddin et al., 2023a). The dynamics of WQ indicators in TrC waters exhibit distinct characteristics compared to other waterbodies (Uddin et al., 2022a). Additionally, sediment and nutrient loadings, such as DIN, TP, and TN, are high in this area (Li et al., 2022a; Soomets et al., 2022). Consequently, these nutrients are the primary contributors to coastal eutrophication, which leads to reduced water transparency and hypoxic conditions (Wang et al., 2023b; Yuan et al., 2022). In order to conserve the marine environment effectively, it is essential to maintain “Good” WQ status in the TrC zone. Since regular in-situ WQ monitoring in the TrC zone is challenging due to its complex nature, integrating RS and ML/AI techniques could be an alternative solution for regular monitoring and sustainable WQ management in this region. However, current research findings reveal that only 18 % of studies are focused on developing OI-WQ indicator models for the TrC waters (Fig. 2). Therefore, more emphasis is required on TrC waterbodies by utilizing RS techniques to ensure good WQ and conserve marine species in these waters.

### 6.3. Generalized OI-WQ indicator model(s) using ML/AI technique(s)

Based on the literature review conducted in this study, less than 2 % of the literature focused on developing RS-based global OI-WQ indicator retrieval models (Fig. 2). In order to implement a sustainable WQ monitor system globally, generalized retrieval algorithms are required for each WQ indicator. Additionally, not every country has the capacity to monitor WQ regularly, as it requires significant time, financial resources, well-equipped laboratories, and expert individuals. In this case, RS and cutting-edge ML/AI techniques offer a sustainable solution, enabling long-term, rapid, and online monitoring facilities for WQ assessment. Although developing a global algorithm is very challenging in the field of RS-based WQ quantification due to the limited in-situ data available for validation. Therefore, it is noteworthy that researchers from different countries may collaborate in this crucial area to effectively optimize this challenge.

### 6.4. Performance evaluation of the RS data in retrieving OI-WQ indicators

Overall, a series of customized bands or band combinations of various satellites, including the MODIS-Aqua/Terra, the Envisat, the Sentinel 3 OLCI, the GOCI, etc., could be effective for developing the OA-WQ indicators (Yang et al., 2022a). However, developing retrieval model(s) for OI-WQ indicators requires further investigation. To the best of the author's knowledge, no prior research has investigated the suitability of the various RS data, along with their associated tools and techniques, for developing OI-WQ indicator model(s) across different waterbodies. As a result, this study emphasizes that in order to improve model(s) performance, it is necessary to validate the developed model(s) in terms of spatio-temporal resolution of the domain for different publicly available RS data. Additionally, incorporating image fusion systems into the performance evaluation process can be advantageous because it can reduce data volume and improve data quality for both human and machine perception.

## 7. Conclusions

The current study has explored the significant advances in the development of OI-WQ indicator retrieval models, particularly pH, DOX, BOD<sub>5</sub>, TP, TN, and DIN, using RS technology, while also identifying limitations of the models. These OI-WQ indicators are often utilized for broad-spectrum WQ monitoring and were selected for this study in accordance with EU-WFD regulations. In this review, the study focused on identifying the suitable RS data in terms of area of interest (AOI), key spectral bands, and various tools and techniques that can be employed in the development of OI-WQ indicator retrieval models. The main findings of this study are summarized as follows:

- Among the RS data utilized for OI-WQ indicator assessment, the Landsat and the Copernicus missions stand out as particularly suitable RS data for developing retrieval models. Their open distribution policy, multispectral resolution, and broad spatial and temporal coverage make them suitable for all researchers. Additionally, hyperspectral RS data (ground- and shipborne) are gaining popularity in the OI-WQ indicator retrieval field due to their spectral resolution and narrowband data acquisition system.
- The research has identified that, generally, spectral reflectance indices (e.g., RI, NDI, and SI) derived from the red, green, blue, and NIR (400-700 nm) bands are prevalent in the assessment of OI-WQ indicators across various satellites. Nevertheless, reaching a consensus or unified conclusion in this regard continues to pose a challenge.
- Researchers widely utilize EM, SEM, ML/AI, and DL-based models to develop RS-based OI-WQ indicator retrieval models. Among them, ML/AI-based models received notable attention due to their cost-

effectiveness and higher accuracy in estimating pH, DOX, BOD<sub>5</sub>, TP, TN, and DIN. Although EM and SEM methods have shown less accuracy and high uncertainty, several studies have used these models for OI-WQ retrieval. However, most studies overlooked the retrieval model(s) uncertainty and sensitivity, which directly affects their reliability.

- The estimation of WQ using RS data is susceptible to various interfering factors related to the surface and atmosphere, such as gaseous absorption, molecular scattering, aerosol scattering, absorption, and cloud droplets. These surface-atmosphere effects can be mitigated through different AC methods, such as C2RCC, ACOLITE, POLYMER, and L2gen. However, the effectiveness of these AC methods varies in terms of space and time. On the other hand, high cloud cover is another major challenge in the field of optical RS.

This study lays a foundation for future research on OI-WQ indicators and serves as a valuable resource for future studies that seek to deepen their understanding of the relationship between OI-WQ indicators and various wavelengths in different waterbodies in order to improve OI-WQ retrieval algorithms. Notably, this study systematically presents all relevant RS data, appropriate wavelengths, and algorithms for retrieving OI-WQ indicators in diverse waterbodies, including lakes, rivers, reservoirs, estuaries, seas, and more. Furthermore, the study provides a comprehensive analysis that not only emphasizes effectiveness but also acknowledges the limitations associated with each approach. Consequently, these valuable insights will play a critical role in guiding future researchers, organizations, and institutions towards selecting the most suitable tools, techniques, and parameters, such as RS wavelengths and WQ indicators, while simultaneously considering the inherent constraints of OI-WQ determination in different aquatic environments. Additionally, it is noteworthy that this comprehensive assessment will significantly advance typical monitoring programs that utilize RS data. Furthermore, this research will contribute to achieving Sustainable Development Goal (SDG-6), which further emphasizes the importance of preserving "Good" WQ status in all types of waterbodies and improving sustainable WQ monitoring systems for a sustainable future.

## Declaration of competing interest

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

## Acknowledgement

The authors gratefully acknowledge the editor's and anonymous reviewers' contributions to the improvement of this paper. This research was funded by the Hardiman Research Scholarship of the University of Galway, which funded the first author as part of his PhD program. The authors wish to thank Dr. Tomasz Dabrowski, an ocean modeller at the Marine Institute of Ireland, for valuable guidance and support. The authors would like to acknowledge support from MaREI, the SFI Research Centre for Energy, Climate, and Marine research. The authors also sincerely acknowledge the Eco-HydroInformatics Research Group (EHIRG), School of Engineering, College of Science and Engineering, University of Galway, Ireland for providing laboratory facilities to complete this research. The research was also supported by the Environmental Protection Agency, Ireland for the AquaCop project [Grant Ref No: 2022-NE-1128].

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.earscirev.2025.105259>.

## Data availability

Data will be made available on request.

## References

- Abbas, M.R., Bin Rasib, A.W., Ahmad, B. Bin, Abbas, T.R., 2021. Statistical Remote Sensing for Prediction of Inland Water Quality Parameters for Shatt Al-Arab River in Iraq. *IOP Conf. Ser. Earth Environ. Sci.* 722 (1). <https://doi.org/10.1088/1755-1315/722/1/012014>.
- Abyanah, H.Z., 2014. Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters. *J. Environ. Health Sci. Eng.* 12 (1), 6–13. <https://doi.org/10.1186/2052-336X-12-40>.
- Adjovu, G.E., Stephen, H., James, D., Ahmad, S., 2023. Overview of the application of remote sensing in effective monitoring of water quality parameters. *Remote Sens* 15 (7). <https://doi.org/10.3390/rs15071938>.
- Agarwal, A., Tavenneau, A., Olbert, A.I., 2018. Remote Sensing of Surface Waters in Ireland. CERI 2018. University College Dublin. <http://hdl.handle.net/10379/16430>.
- Ahmad, N.B., Jaafaru, M.S., Isa, Z., Abdulhamid, Y., Kakudi, R.A., Ugya, A.Y., Meguellati, K., 2024. High pollution loads engineer oxygen dynamics, ecological niches, and pathogenicity shifts in freshwater environments. *J. Hazard. Mater. Adv.* 14, 100425. <https://doi.org/10.1016/j.hazadv.2024.100425>.
- Ahmed, M., Mumtaz, R., Anwar, Z., Shaikat, A., Arif, O., Shafait, F., 2022. A multi-step approach for optically active and inactive water quality parameter estimation using deep learning and remote sensing. *Water (Switzerland)* 14 (13). <https://doi.org/10.3390/w14132112>.
- Aladejare, A.E., Idris, M.A., 2020. Performance analysis of empirical models for predicting rock mass deformation modulus using regression and Bayesian methods. *J. Rock Mech. Geotech. Eng.* 12 (6), 1263–1271. <https://doi.org/10.1016/j.jrmge.2020.03.007>.
- Al-Shaibah, B., Liu, X., Zhang, J., Tong, Z., Zhang, M., El-Zeiny, A., Faichia, C., Hussain, M., Tayyab, M., 2021. Modeling water quality parameters using landsat multispectral images: a case study of erlong lake, Northeast China. *Remote Sens* 13 (9). <https://doi.org/10.3390/rs13091603>.
- Alsulaili, A., Refaie, A., 2021. Artificial neural network modeling approach for the prediction of five-day biological oxygen demand and wastewater treatment plant performance. *Water Supply* 21 (5), 1861–1877. <https://doi.org/10.2166/ws.2020.199>.
- Anspers-Toomsalu, A., Uusöue, M., Kangro, K., Hieronymi, M., Alikas, K., 2024. Suitability of different in-water algorithms for eutrophic and absorbing waters applied to Sentinel-2 MSI and Sentinel-3 OLCI data. *Front. Remote Sens.* 5 (July), 1–24. <https://doi.org/10.3389/frsen.2024.1423332>.
- Arabi, B., Salama, M.S., Pitarch, J., Verhoef, W., 2020. Integration of in-situ and multi-sensor satellite observations for long-term water quality monitoring in coastal areas. *Remote Sens. Environ.* 239, 111632. <https://doi.org/10.1016/j.rse.2020.111632>.
- Arena, M., Pratolongo, P., Loisel, H., Tran, M.D., Jorge, D.S.F., Delgado, A.L., 2024. Optical water characterization and atmospheric correction assessment of estuarine and coastal waters around the AERONET-OC Bahia Blanca. *Front. Remote Sens.* 5, 1–17. <https://doi.org/10.3389/frsen.2024.1305787>.
- Arias-Rodriguez, L.F., Tüzün, U.F., Duan, Z., Huang, J., Tuo, Y., Disse, M., 2023. Global water quality of inland waters with harmonized Landsat-8 and Sentinel-2 using cloud-computed machine learning. *Remote Sens* 15 (5). <https://doi.org/10.3390/rs15051390>.
- Ascani, M., Pabsch, D., Klinksiek, M., Gajardo-Parra, N., Sadowski, G., Held, C., 2022. Prediction of pH in multiphase multicomponent systems with ePC-SAFT advanced. *Chem. Commun.* 58 (60), 8436–8439. <https://doi.org/10.1039/d2cc02943j>.
- Baki, O.T., Aras, E., Akdemir, U.O., Yilmaz, B., 2019. Biochemical oxygen demand prediction in wastewater treatment plant by using different regression analysis models. *Desalin. Water Treat.* 157, 79–89. <https://doi.org/10.5004/dwt.2019.24158>.
- Batur, E., Maktav, D., 2019. Assessment of surface water quality by using satellite images fusion based on PCA method in the Lake Gala, Turkey. *IEEE Trans. Geosci. Remote Sens.* 57 (5), 2983–2989. <https://doi.org/10.1109/TGRS.2018.2879024>.
- Boyd, C.E., 2015. Water quality: An introduction. In: *Water Quality: An Introduction*, 2nd ed. Springer Cham. <https://doi.org/10.1007/978-3-319-17446-4>.
- Brockmann, C., Doerffer, R., Peters, M., Stelzer, K., Embacher, S., Ruescas, A., 2016. Evolution of the C2RCC neural network for sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters. *Living Planet Sympos.* 3 (2), 13–22.
- Cao, L., Zhang, D., Guo, Q., Zhan, J., 2021. Inversion of water quality parameter bod5 based on hyperspectral remotely sensed data in Qinghai Lake. In: *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 5036–5039. <https://doi.org/10.1109/IGARSS47720.2021.9553783>.
- Cao, J., Wen, X., Luo, D., Tan, Y., 2022a. Study on water quality inversion model of Dianchi lake based on Landsat 8 data. *J. Spectrosc.* 2022. <https://doi.org/10.1155/2022/3341713>.
- Cao, Q., Yu, G., Sun, S., Dou, Y., Li, H., Qiao, Z., 2022b. Monitoring water quality of the haihe river based on ground-based hyperspectral remote sensing. *Water (Switzerland)* 14 (1), 1–13. <https://doi.org/10.3390/w14010022>.
- Cao, Q., Yu, G., Qiao, Z., 2023. Application and recent progress of inland water monitoring using remote sensing techniques. *Environ. Monit. Assess.* 195 (1). <https://doi.org/10.1007/s10661-022-10690-9>.
- Chatziantoniou, A., Charalampis Spondylidis, S., Stavrakidis-Zachou, O., Papandroulakis, N., Topouzelis, K., 2022. Dissolved oxygen estimation in aquaculture sites using remote sensing and machine learning. *Remote Sens. Appl. Soc. Environ.* 28, 100865. <https://doi.org/10.1016/j.rsase.2022.100865>.
- Chawla, I., Karthikeyan, L., Mishra, A.K., 2020. A review of remote sensing applications for water security: quantity, quality, and extremes. *J. Hydrol.* 585, 124826. <https://doi.org/10.1016/j.jhydrol.2020.124826>.
- Chen, J., Quan, W., 2012. Using Landsat/TM imagery to estimate nitrogen and phosphorus concentration in Taihu Lake, China. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 5 (1), 273–280. <https://doi.org/10.1109/JSTARS.2011.2174339>.
- Chen, B., Mu, X., Chen, P., Wang, B., Choi, J., Park, H., Xu, S., Wu, Y., Yang, H., 2021. Machine learning-based inversion of water quality parameters in typical reach of the urban river by UAV multispectral data. *Ecol. Indic.* 133. <https://doi.org/10.1016/j.ecolind.2021.108434>.
- Chen, J., Chen, S., Fu, R., Li, D., Jiang, H., Wang, C., Peng, Y., Jia, K., Hicks, B.J., 2022. Remote sensing big data for water environment monitoring: current status, challenges, and future prospects. *Earth's Future* 10 (2), 1–33. <https://doi.org/10.1029/2021EF002289>.
- Colella, S., Böhm, E., Cesarini, C., Jutard, Q., Brando, V.E., 2024. Copernicus Marine Product User Manual for Ocean Colour Products. Mercator Ocean, CMEMS-OC-P, pp. 1–27.
- Concha, J.A., Bracaglia, M., Brando, V.E., 2021. Assessing the influence of different validation protocols on Ocean Colour match-up analyses. *Remote Sens. Environ.* 259, 112415. <https://doi.org/10.1016/j.rse.2021.112415>.
- Cox, B.A., 2003. A review of currently available in-stream water-quality models and their applicability for simulating dissolved oxygen in lowland rivers. *Sci. Total Environ.* 314–316 (03), 335–377. [https://doi.org/10.1016/S0048-9697\(03\)00063-9](https://doi.org/10.1016/S0048-9697(03)00063-9).
- Cruz-Montes, E.E., Durango-Banquett, M.M., Torres-Bejarano, F.M., Campo-Daza, G.A., Padilla-Mendoza, C., 2023. Remote sensing application using Landsat 8 images for water quality assessments. *J. Phys. Conf. Ser.* 2475 (1). <https://doi.org/10.1088/1742-6596/2475/1/012007>.
- Dall'Olmo, G., Brewin, R.J.W., Nencioli, F., Organelli, E., Lefering, I., McKee, D., Röttgers, R., Mitchell, C., Boss, E., Bricaud, A., Tilstone, G., 2017. Determination of the absorption coefficient of chromophoric dissolved organic matter from underway spectrophotometry. *Opt. Express* 25 (24), A1079. <https://doi.org/10.1364/oe.25.0a1079>.
- DeSimone, L.A., Pope, J.P., Ransom, K.M., 2020. Machine-learning models to map pH and redox conditions in groundwater in a layered aquifer system, Northern Atlantic Coastal Plain, eastern USA. *J. Hydrol. Region. Stud.* 30, 100697. <https://doi.org/10.1016/j.ejrh.2020.100697>.
- Dey, J., Vijay, R., 2021. A critical and intensive review on assessment of water quality parameters through geospatial techniques. *Environ. Sci. Pollut. Res.* 28 (31), 41612–41626. <https://doi.org/10.1007/s11356-021-14726-4>.
- Diganta, M.T.M., Uddin, M.G., Rahman, A., Olbert, A.I., 2024. A comprehensive review of various environmental factors' roles in remote sensing techniques for assessing surface water quality. *Sci. Total Environ.* 957 (July). <https://doi.org/10.1016/j.scitotenv.2024.177180>.
- Ding, C., Pu, F., Li, C., Xu, X., Zou, T., Li, X., 2020. Combining artificial neural networks with causal inference for total phosphorus concentration estimation and sensitive spectral bands exploration using MODIS. *Water (Switzerland)* 12 (9). <https://doi.org/10.3390/w12092372>.
- Dong, L., Gong, C., Huai, H., Wu, E., Lu, Z., Hu, Y., Li, L., Yang, Z., 2023. Retrieval of water quality parameters in dianshan lake based on Sentinel-2 MSI imagery and machine learning. *Algor. Eval. Spatiotemp. Change Res. Remote Sens.* 15 (20). <https://doi.org/10.3390/rs15205001>.
- Dörnhöfer, K., Oppelt, N., 2016. Remote sensing for lake research and monitoring - recent advances. *Ecol. Indic.* 64, 105–122. <https://doi.org/10.1016/j.ecolind.2015.12.009>.
- Doxani, G., Vermote, E., Roger, J.C., Gascon, F., Adriaenssens, S., Frantz, D., Hagolle, O., Hollstein, A., Kirches, G., Li, F., Louis, J., Mangin, A., Pahlevan, N., Pflug, B., Vanhelmont, Q., 2018. Atmospheric correction inter-comparison exercise. *Remote Sens* 10 (2), 1–18. <https://doi.org/10.3390/rs10020352>.
- Du, C., Wang, Q., Li, Y., Lyu, H., Zhu, L., Zheng, Z., Wen, S., Liu, G., Guo, Y., 2018. Estimation of total phosphorus concentration using a water classification method in inland water. *Int. J. Appl. Earth Obs. Geoinf.* 71, 29–42. <https://doi.org/10.1016/j.jag.2018.05.007>.
- Du, Y., Song, K., Liu, G., Wen, Z., Fang, C., Shang, Y., Zhao, F., Wang, Q., Du, J., Zhang, B., 2020. Quantifying total suspended matter (TSM) in waters using Landsat images during 1984–2018 across the Songnen Plain, Northeast China. *J. Environ. Manag.* 262, 110334. <https://doi.org/10.1016/j.jenvman.2020.110334>.
- Du, C., Shi, K., Liu, N., Li, Y., Lyu, H., Yan, C., Pan, J., 2022. Remote Estimation of the Particulate Phosphorus Concentrations in Inland Water Bodies: A Case Study in Hongze Lake. *Remote Sens* 14 (16). <https://doi.org/10.3390/rs14163863>.
- EPA, 2001. Parameters of Water Quality: Interpretation and Standards. Environmental Protection Agency, Ireland. <https://thewaternetwork.com/document-XAs/parameters-of-water-quality-interpretation-and-standards-2001-epa-835aGBDasoKppIzyGUWHg>, accessed on 07 October 2022.
- EPA. (2007). Proposed Quality Standards for Surface Water Classification - FINAL. July, 92. [https://www.google.ie/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0CCUQFjAB&url=http://www.wfdireland.ie/docs/34\\_Public\\_Participation/Proposed\\_Quality\\_Standards\\_for\\_Surface\\_Water\\_Classification](https://www.google.ie/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0CCUQFjAB&url=http://www.wfdireland.ie/docs/34_Public_Participation/Proposed_Quality_Standards_for_Surface_Water_Classification).
- EPA, 2023. The EPA's Monitoring and Assessment Role. <https://www.epa.ie/our-services/monitoring-assessment/> accessed on 21 December 2023.
- Escoto, J.E., Blanco, A.C., Argamosa, R.J., Medina, J.M., 2021. Pasig river water quality estimation using an empirical ordinary least squares regression model of sentinel-2 satellite images. In: *International Archives of the Photogrammetry, Remote Sensing*

- and Spatial Information Sciences - ISPRS Archives, 46(4/W6-2021), pp. 161–168. <https://doi.org/10.5194/isprs-archives-XLVI-4-W6-2021-161-2021>.
- EU, 2019. European Union environmental objectives (surface waters) (amendment) regulations 2019. Eur. Union 3 (1), 18–23.
- Fang, C., Song, C., Wang, X., Wang, Q., Tao, H., Wang, X., Ma, Y., Song, K., 2024. A novel total phosphorus concentration retrieval method based on two-line classification in lakes and reservoirs across China. *Sci. Total Environ.* 906, 167522. <https://doi.org/10.1016/j.scitotenv.2023.167522>.
- Flecha, S., Giménez-Romero, À., Tintoré, J., Pérez, F.F., Alou-Font, E., Matías, M.A., Hendriks, I.E., 2022. pH trends and seasonal cycle in the coastal Balearic Sea reconstructed through machine learning. *Sci. Rep.* 12 (1), 1–11. <https://doi.org/10.1038/s41598-022-17253-5>.
- Frouin, R.J., Franz, B.A., Ibrahim, A., Knobelspiesse, K., Ahmad, Z., Cairns, B., Chowdhary, J., Dierssen, H.M., Tan, J., Dubovik, O., Huang, X., Davis, A.B., Kalashnikova, O., Thompson, D.R., Remer, L.A., Boss, E., Coddington, O., Deschamps, P.Y., Gao, B.C., Zhai, P.W., 2019. Atmospheric correction of satellite ocean-color imagery during the PACE Era. *Front. Earth Sci.* 7, 1–43. <https://doi.org/10.3389/feart.2019.00145>.
- Fu, Y., Hu, Z., Zhao, Y., Huang, M., 2021. A long-term water quality prediction method based on the temporal convolutional network in smart mariculture. *Water* (Switzerland) 13 (2907). <https://doi.org/10.3390/w13202907>.
- Fu, B., Lao, Z., Liang, Y., Sun, J., He, X., Deng, T., He, W., Fan, D., Gao, E., Hou, Q., 2022. Evaluating optically and non-optically active water quality and its response relationship to hydro-meteorology using multi-source data in Poyang Lake, China. *Ecol. Indic.* 145. <https://doi.org/10.1016/j.ecolind.2022.109675>.
- Gani, M.A., Sajib, A.M., Siddik, M.A., Moniruzzaman, Md, 2023. Assessing the impact of land use and land cover on river water quality using water quality index and remote sensing techniques. *Environ. Monit. Assess.* 195 (4). <https://doi.org/10.1007/s10661-023-10989-1>.
- Gao, B.C., Davis, C.O., Goetz, A.F.H., 2006. A review of atmospheric correction techniques for hyperspectral remote sensing of land surfaces and ocean color. *Int. Geosci. Remote Sens. Sympos.* 1979–1981. <https://doi.org/10.1109/IGARSS.2006.512>.
- Gao, Y., Gao, J., Yin, H., Liu, C., Xia, T., Wang, J., Huang, Q., 2015. Remote sensing estimation of the total phosphorus concentration in a large lake using band combinations and regional multivariate statistical modeling techniques. *J. Environ. Manag.* 151, 33–43. <https://doi.org/10.1016/j.jenvman.2014.11.036>.
- Gao, K., Beardall, J., Häder, D.P., Hall-Spencer, J.M., Gao, G., Hutchins, D.A., 2019. Effects of ocean acidification on marine photosynthetic organisms under the concurrent influences of warming, UV radiation, and deoxygenation. *Front. Mar. Sci.* 6, 1–18. <https://doi.org/10.3389/fmars.2019.00322>.
- Garsson, P., Mangin, A., Bretagnon, M., Jutard, Q., 2024. Satellite observation copernicus-globcolour products, Ref: CMEMS-OC-QUID-009-101to104-111-113-116-118. Mercator Ocean Int. 1–93. <https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-OC-QUID-009-030-032-033-037-081-082-083-085-086-098.pdf>.
- Gholizadeh, M.H., Melesse, A.M., 2017. Study on spatiotemporal variability of water quality parameters in florida bay using remote sensing. *J. Remote Sens. GIS* 06 (03). <https://doi.org/10.4172/2469-4134.1000207>.
- Gholizadeh, M.H., Melesse, A.M., Reddi, L., 2016. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors* (Switzerland) 16 (8). <https://doi.org/10.3390/s16081298>.
- Gilerson, A.A., Gitelson, A.A., Zhou, J., Gurlin, D., Moses, W., Ioannou, I., Ahmed, S.A., 2010. Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Opt. Express* 18 (23), 24109. <https://doi.org/10.1364/oe.18.024109>.
- Gleratti, G., Martinez-Vicente, V., Atwood, E.C., Simis, S.G.H., Jackson, T., 2024. Validation of full resolution remote sensing reflectance from Sentinel-3 OLCI across optical gradients in moderately turbid transitional waters. *Front. Remote Sens.* 5, 1–18. <https://doi.org/10.3389/frsen.2024.1359709>.
- González Vilas, L., Brando, V.E., Di Cicco, A., Colella, S., D'Alimonte, D., Kajiyama, T., Attila, J., Schroeder, T., 2023. Assessment of ocean color atmospheric correction methods and development of a regional ocean color operational dataset for the Baltic Sea based on Sentinel-3 OLCI. *Front. Mar. Sci.* 10, 1–26. <https://doi.org/10.3389/fmars.2023.1256990>.
- Goyens, C., Ruddick, K., 2023. Improving the standard protocol for above-water reflectance measurements: 1. Estimating effective wind speed from angular variation of sunglint. *Appl. Opt.* 62 (10), 2442. <https://doi.org/10.1364/ao.481787>.
- Guo, H., Huang, J.J., Chen, B., Guo, X., Singh, V.P., 2021a. A machine learning-based strategy for estimating non-optically active water quality parameters using Sentinel-2 imagery. *Int. J. Remote Sens.* 42 (5), 1841–1866. <https://doi.org/10.1080/01431161.2020.1846222>.
- Guo, H., Huang, J.J., Zhu, X., Wang, B., Tian, S., Xu, W., Mai, Y., 2021b. A generalized machine learning approach for dissolved oxygen estimation at multiple spatiotemporal scales using remote sensing. *Environ. Pollut.* 288, 117734. <https://doi.org/10.1016/j.envpol.2021.117734>.
- Guo, C., Ai, W., Liu, M., Feng, M., Qiao, J., Hu, S., 2022a. Sea surface temperature retrieval based on simulated space-borne one-dimensional multifrequency synthetic aperture microwave radiometry. *Front. Environ. Sci.* 10, 1–12. <https://doi.org/10.3389/fenvs.2022.1054076>.
- Guo, H., Tian, S., Jeanne Huang, J., Zhu, X., Wang, B., Zhang, Z., 2022b. Performance of deep learning in mapping water quality of Lake Simcoe with long-term Landsat archive. *ISPRS J. Photogramm. Remote Sens.* 183, 451–469. <https://doi.org/10.1016/j.isprsjprs.2021.11.023>.
- Guo, Y., Deng, R., Li, J., Hua, Z., Wang, J., Zhang, R., Liang, Y., Tang, Y., 2022c. Remote sensing retrieval of total nitrogen in the Pearl River Delta based on landsat8. *Water* (Switzerland) 14 (22). <https://doi.org/10.3390/w14223710>.
- Hajjizadeh, M., Moncada, A., Kent, S., Melesse, A.M., 2021. Land-lake linkage and remote sensing application in water quality monitoring in Lake Okeechobee, Florida, USA. *Land* 10 (2), 1–17. <https://doi.org/10.3390/land10020147>.
- Han, Z., Gu, X., Zuo, X., Bi, K., Shi, S., 2022. Semi-empirical models for the bidirectional water-leaving radiance: an analysis of a Turbid Inland Lake. *Front. Environ. Sci.* 9, 1–17. <https://doi.org/10.3389/fenvs.2021.818557>.
- Han, W., Zhang, X., Wang, Y., Wang, L., Huang, X., Li, J., Wang, S., Chen, W., Li, X., Feng, R., Fan, R., Zhang, X., Wang, Y., 2023. A survey of machine learning and deep learning in remote sensing of geological environment: challenges, advances, and opportunities. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 202. Elsevier B.V., pp. 87–113. <https://doi.org/10.1016/j.isprsjprs.2023.05.032>.
- He, W., Chen, S., Liu, X., Chen, J., 2008. Water quality monitoring in a slightly-polluted inland water body through remote sensing - case study of the Guanting Reservoir in Beijing, China. *Front. Environ. Sci. Eng. China* 2 (2), 163–171. <https://doi.org/10.1007/s11783-008-0027-7>.
- He, Y., Gong, Z., Zheng, Y., Zhang, Y., 2021a. Inland reservoir water quality inversion and eutrophication evaluation using bp neural network and remote sensing imagery: a case study of dashahe reservoir. *Water* (Switzerland) 13 (20). <https://doi.org/10.3390/w13202844>.
- He, Y., Jin, S., Shang, W., 2021b. Water quality variability and related factors along the yangtze river using landsat-8. *Remote Sens* 13 (12), 1–20. <https://doi.org/10.3390/rs13122241>.
- Hooker, S.B., Matsuoka, A., Kudela, R.M., Yamashita, Y., Suzuki, K., Houskeeper, H.F., 2020. A global end-member approach to derive aCDOM(440) from near-surface optical measurements. *Biogeosciences* 17 (2), 475–497. <https://doi.org/10.5194/bg-17-475-2020>.
- Hou, Y., Zhang, A., Lv, R., Zhao, S., Ma, J., Zhang, H., Li, Z., 2022. A study on water quality parameters estimation for urban rivers based on ground hyperspectral remote sensing technology. *Environ. Sci. Pollut. Res.* 29 (42), 63640–63654. <https://doi.org/10.1007/s11356-022-20293-z>.
- Hou, Y., Zhang, A., Lv, R., Zhang, Y., Ma, J., Li, T., 2023. Machine learning algorithm inversion experiment and pollution analysis of water quality parameters in urban small and medium-sized rivers based on UAV multispectral data. *Environ. Sci. Pollut. Res.* 30 (32), 78913–78932. <https://doi.org/10.1007/s11356-023-27963-6>.
- Houskeeper, H.F., Hooker, S.B., Kudela, R.M., 2021. Spectral range within global aCDOM (440) algorithms for oceanic, coastal, and inland waters with application to airborne measurements. *Remote Sens. Environ.* 253, 112155. <https://doi.org/10.1016/j.rse.2020.112155>.
- Hu, Z., Zhang, Y., Zhao, Y., Xie, M., Zhong, J., Tu, Z., Liu, J., 2019. A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture. *Sensors* (Switzerland) 19 (6). <https://doi.org/10.3390/s19061420>.
- Hu, W., Liu, J., Wang, H., Miao, D., Shao, D., Gu, W., 2023. Retrieval of TP concentration from UAV multispectral images using IOA-ML models in small inland waterbodies. *Remote Sens* 15 (5). <https://doi.org/10.3390/rs15051250>.
- Huang, J., Wang, D., Gong, F., Bai, Y., He, X., 2021. Changes in nutrient concentrations in Shenzhen bay detected using landsat imagery between 1988 and 2020. *Remote Sens* 13 (17), 1–20. <https://doi.org/10.3390/rs13173469>.
- Huang, J., Wang, D., Pan, S., Li, H., Gong, F., Hu, H., He, X., Bai, Y., Zheng, Z., 2023. A new high-resolution remote sensing monitoring method for nutrients in coastal waters. *IEEE Trans. Geosci. Remote Sens.* 1. <https://doi.org/10.1109/TGRS.2023.3294436>.
- IOCCG, 2010. IOCCG Report Number 10: Atmospheric correction for remotely-sensed ocean-colour products. In: *IOCCG Web Page*. <http://www.ioccg.org/reports/report10.pdf>.
- Isenstein, E.M., Park, M.H., 2014. Assessment of nutrient distributions in Lake Champlain using satellite remote sensing. *J. Environ. Sci.* 26 (9), 1831–1836. <https://doi.org/10.1016/j.jes.2014.06.019>.
- Jiang, Z., Song, Z., Bai, Y., He, X., Yu, S., Zhang, S., Gong, F., 2022. Remote sensing of global sea surface pH based on massive underway data and machine learning. *Remote Sens* 14 (10), 1–24. <https://doi.org/10.3390/rs14102366>.
- Kapalanga, T.S., Hoko, Z., Gumindoga, W., Chikwiramakomo, L., 2021. Remote-sensing-based algorithms for water quality monitoring in olushandja dam, north-Central Namibia. *Water Supply* 21 (5), 1878–1894. <https://doi.org/10.2166/ws.2020.290>.
- Karakaya, N., Evrendilek, F., 2011. Monitoring and validating spatio-temporal dynamics of biogeochemical properties in Mersin Bay (Turkey) using Landsat ETM+. *Environ. Monit. Assess.* 181 (1–4), 457–464. <https://doi.org/10.1007/s10661-010-1841-5>.
- Karaoui, I., Arioua, A., Boudhar, A., Hssaisoune, M., El Mouatassime, S., Ouhamchich, K. A., Elhamedouni, D., 2019. Evaluating the potential of Sentinel-2 satellite images for water quality characterization of artificial reservoirs: the Bin El Ouidane Reservoir case study (Morocco). *Meteorol. Hydrol. Water Manage.* 7 (1), 1–10.
- Karimi, B., Hashemi, S.H., Aghighi, H., 2023. Development of the best retrieval models of non-optically active parameters for an artificial shallow lake by random forest algorithm. *Remote Sens. Appl. Soc. Environ.* 29, 100926. <https://doi.org/10.1016/j.rsase.2023.100926>.
- Kim, Y.H., Son, S., Kim, H.C., Kim, B., Park, Y.G., Nam, J., Ryu, J., 2020. Application of satellite remote sensing in monitoring dissolved oxygen variabilities: a case study for coastal waters in Korea. *Environ. Int.* 134, 105301. <https://doi.org/10.1016/j.envint.2019.105301>.
- Kim, Y.J., Kim, W., Im, J., Choi, J., Lee, S., 2023. Atmospheric-correction-free red tide quantification algorithm for GOCI based on machine learning combined with a radiative transfer simulation. *ISPRS J. Photogramm. Remote Sens.* 199, 197–213. <https://doi.org/10.1016/j.isprsjprs.2023.04.007>.

- Kotchenova, S.Y., Vermote, E.F., Matarrese, R., Klemm, F.J., 2006. Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: path radiance. *Appl. Opt.* 45 (26), 6762–6774. <https://doi.org/10.1364/AO.45.006762>.
- Krishnaraj, A., Honnasiddaiah, R., 2022. Remote sensing and machine learning based framework for the assessment of spatio-temporal water quality in the Middle Ganga Basin. *Environ. Sci. Pollut. Res.* 29 (43), 64939–64958. <https://doi.org/10.1007/s11356-022-20386-9>.
- Kumar, S., Imen, S., Sridharan, V.K., Gupta, A., McDonald, W., Ramirez-Avila, J.J., Abdul-Aziz, O.I., Talchabhadhel, R., Gao, H., Quinn, N.W.T., Weiss, W.J., Poulou, T., Palmate, S.S., Lee, C.M., Baskaran, L., 2024. Perceived barriers and advances in integrating earth observations with water resources modeling. *Remote Sens. Appl. Soc. Environ.* 33, 101119. <https://doi.org/10.1016/j.rsase.2023.101119>.
- Kutser, T., 2012. The possibility of using the Landsat image archive for monitoring long time trends in coloured dissolved organic matter concentration in lake waters. *Remote Sens. Environ.* 123, 334–338. <https://doi.org/10.1016/j.rse.2012.04.004>.
- LAWA, 2021. Factsheet: Nitrogen. Land Air Water AOTEAROA, New Zealand. Retrieved on 12 November 2022 from. <https://www.lawa.org.nz/learn/factsheets/nitrogen/>.
- Lednicka, B., Kubacka, M., 2022. Semi-empirical model of remote-sensing reflectance for chosen areas of the Southern Baltic. *Sensors* 22 (3), 1–23. <https://doi.org/10.3390/s22031105>.
- Li, W., Zhang, J., 2020. Prediction of BOD concentration in wastewater treatment process using a modular neural network in combination with the weather condition. *Appl. Sci.* 10 (21), 1–14. <https://doi.org/10.3390/app10217477>.
- Li, Y., Zhang, Y., Shi, K., Zhu, G., Zhou, Y., Zhang, Y., Guo, Y., 2017. Monitoring spatiotemporal variations in nutrients in a large drinking water reservoir and their relationships with hydrological and meteorological conditions based on Landsat 8 imagery. *Sci. Total Environ.* 599–600, 1705–1717. <https://doi.org/10.1016/j.scitotenv.2017.05.075>.
- Li, H., Zhang, G., Zhu, Y., Kaufmann, H., Xu, G., 2022a. Inversion and driving force analysis of nutrient concentrations in the ecosystem of the Shenzhen-Hong Kong Bay Area. *Remote Sens.* 14 (15). <https://doi.org/10.3390/rs14153694>.
- Li, J., Wang, J., Wu, Y., Cui, Y., Yan, S., 2022b. Remote sensing monitoring of total nitrogen and total phosphorus concentrations in the water around Chaoahu Lake based on geographical division. *Front. Environ. Sci.* 10, 1–16. <https://doi.org/10.3389/fenvs.2022.1014155>.
- Li, N., Ning, Z., Chen, M., Wu, D., Hao, C., Zhang, D., Bai, R., Liu, H., Chen, X., Li, W., Zhang, W., Chen, Y., Li, Q., Zhang, L., 2022c. Satellite and machine learning monitoring of optically inactive water quality variability in a tropical river. *Remote Sens.* 14 (21), 5466. <https://doi.org/10.3390/rs14215466>.
- Li, L., Gu, M., Gong, C., Hu, Y., Wang, X., Yang, Z., He, Z., 2023a. An advanced remote sensing retrieval method for urban non-optically active water quality parameters: an example from Shanghai. *Sci. Total Environ.* 880, 163389. <https://doi.org/10.1016/j.scitotenv.2023.163389>.
- Li, S., Song, K., Li, Y., Liu, G., Wen, Z., Shang, Y., Lyu, L., Fang, C., 2023b. Performances of atmospheric correction processors for sentinel-2 msi imagery over typical lakes across China. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 16, 2065–2078. <https://doi.org/10.1109/JSTARS.2023.3238713>.
- Li, W., Huang, Y., Shen, Q., Yao, Y., Xu, W., Shi, J., Zhou, Y., Li, J., Zhang, Y., Gao, H., 2023c. Assessment of seven atmospheric correction processors for the sentinel-2 multi-spectral imager over lakes in Qinghai Province. *Remote Sens.* 15 (22), 5370. <https://doi.org/10.3390/rs15225370>.
- Li, N., Zhang, Y., Shi, K., Zhang, Y., Sun, X., Wang, W., Qian, H., Yang, H., Niu, Y., 2023d. Real-time and continuous tracking of total phosphorus using a ground-based hyperspectral proximal sensing system. *Remote Sens.* 15 (2), 1–18. <https://doi.org/10.3390/rs15020507>.
- Liang, Y., Yin, F., Xie, D., Liu, L., Zhang, Y., Ashraf, T., 2022. Inversion and monitoring of the TP Concentration in Taihu Lake using the Landsat-8 and Sentinel-2 Images. *Remote Sens.* 13 (6284), 1–19.
- Lim, J., Choi, M., 2015. Assessment of water quality based on Landsat 8 operational land imager associated with human activities in Korea. *Environ. Monit. Assess.* 187 (6), 1–17. <https://doi.org/10.1007/s10661-015-4616-1>.
- Liu, J., Zhang, Y., Yuan, D., Song, X., 2015. Empirical estimation of total nitrogen and total phosphorus concentration of urban water bodies in China using high resolution IKONOS multispectral imagery. *Water (Switzerland)* 7 (11), 6551–6573. <https://doi.org/10.3390/w7116551>.
- Liu, H., Lin, L., Wang, Y., Du, L., Wang, S., Zhou, P., Yu, Y., Gong, X., Lu, X., 2022a. Reconstruction of monthly surface nutrient concentrations in the yellow and Bohai Seas from 2003–2019 using machine learning. *Remote Sens.* 14 (19). <https://doi.org/10.3390/rs14195021>.
- Liu, M., Wang, L., Qiu, F., 2022b. Using MODIS data to track the long-term variations of dissolved oxygen in Lake Taihu. *Front. Environ. Sci.* 10, 1–15. <https://doi.org/10.3389/fenvs.2022.1096843>.
- Liu, X., Al-Shaibah, B., Zhao, C., Tong, Z., Bian, H., Zhang, F., Zhang, J., Pei, X., 2022c. Estimation of the key water quality parameters in the surface water, Middle of Northeast China, based on Gaussian process regression. *Remote Sens.* 14 (24). <https://doi.org/10.3390/rs14246323>.
- Llodrà-Llabrés, J., Martínez-López, J., Postma, T., Pérez-Martínez, C., Alcaraz-Segura, D., 2023. Retrieving water chlorophyll-a concentration in inland waters from Sentinel-2 imagery: review of operability, performance and ways forward. In: *Int. J. Appl. Earth Obs. Geoinf.* 125. <https://doi.org/10.1016/j.jag.2023.103605>.
- Louis, J., Bressac, M., Pedrotti, M.L., Guieu, C., 2015. Dissolved inorganic nitrogen and phosphorus dynamics in seawater following an artificial Saharan dust deposition event. *Front. Mar. Sci.* 2, 1–11. <https://doi.org/10.3389/fmars.2015.00027>.
- Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., Gascon, F., 2016. Sentinel-2 SEN2COR: L2A processor for users. Eur. Space Agency 9–13 (Special Publication) ESA SP, SP-740(May). [https://elib.dlr.de/107381/1/LPS2016\\_sm10\\_3louis.pdf](https://elib.dlr.de/107381/1/LPS2016_sm10_3louis.pdf).
- Lu, S., Deng, R., Liang, Y., Xiong, L., Ai, X., Qin, Y., 2020. Remote sensing retrieval of total phosphorus in the pearl river channels based on the GF-1 remote sensing data. *Remote Sens.* 12 (9). <https://doi.org/10.3390/rs12091420>.
- Lyu, L., Song, K., Wen, Z., Liu, G., Shang, Y., Li, S., Tao, H., Wang, X., Hou, J., 2022. Estimation of the lake trophic state index (TSI) using hyperspectral remote sensing in Northeast China. *Opt. Express* 30 (7), 10329. <https://doi.org/10.1364/oe.453404>.
- Ma, J., Yuan, Y., Zhou, T., Yuan, D., 2017. Determination of total phosphorus in natural waters with a simple neutral digestion method using sodium persulfate. *Limnol. Oceanogr. Methods* 15 (4), 372–380. <https://doi.org/10.1002/lom3.10165>.
- Mamun, M., An, K.G., 2021. Application of multivariate statistical techniques and water quality index for the assessment of water quality and apportionment of pollution sources in the yeongsan river, South Korea. *Int. J. Environ. Res. Public Health* 18 (16), 1–23. <https://doi.org/10.3390/ijerph18168268>.
- Maroufpoor, S., Sammen, S.S., Alansari, N., Abba, S.I., Malik, A., Shahid, S., Mokhtar, A., Maroufpoor, E., 2022. A novel hybridized neuro-fuzzy model with an optimal input combination for dissolved oxygen estimation. *Front. Environ. Sci.* 10 (August), 1–17. <https://doi.org/10.3389/fenvs.2022.929707>.
- Matthews, M.W., 2011. A current review of empirical procedures of remote sensing in Inland and near-coastal transitional waters. *Int. J. Remote Sens.* 32 (21), 6855–6899. <https://doi.org/10.1080/01431161.2010.512947>.
- Mengist, W., Soromessa, T., Legese, G., 2020. Method for conducting systematic literature review and meta-analysis for environmental science research. *MethodsX* 7, 100777. <https://doi.org/10.1016/j.mex.2019.100777>.
- Mobley, C.D., 2001. Radiative transfer in the ocean. *Encyclop. Ocean Sci.* 4, 379–388. <https://doi.org/10.1016/B978-0-12-409548-9.04311-6>.
- Mobley, C.D., Stramski, D., Paul Bissett, W., Boss, E., 2004. Optical modeling of ocean waters: is the case 1 - case 2 classification still useful? *Oceanography* 17 (SPLISS. 2), 60–67. <https://doi.org/10.5670/oceanog.2004.48>.
- Moghadam, S.V., Sharafati, A., Feizi, H., Marjaie, S.M.S., Asadollah, S.B.H.S., Motta, D., 2021. An efficient strategy for predicting river dissolved oxygen concentration: application of deep recurrent neural network model. *Environ. Monit. Assess.* 193 (12). <https://doi.org/10.1007/s10661-021-09586-x>.
- Mohandas, K.A., Brema, J., 2023. Comparative analysis of regression models for remote sensing-based water quality assessment. *J. Sci. Ind. Res.* 82 (4), 466–474. <https://doi.org/10.56042/jsir.v82i04.61646>.
- Mohseni, F., Saba, F., Mirmazloumi, S.M., Amani, M., Mokhtarzade, M., Jamali, S., Mahdavi, S., 2022. Ocean water quality monitoring using remote sensing techniques: a review. *Mar. Environ. Res.* 180, 105701. <https://doi.org/10.1016/j.marenvres.2022.105701>.
- Morel, A., Gentili, B., 2009. A simple band ratio technique to quantify the colored dissolved and detrital organic material from ocean color remotely sensed data. *Remote Sens. Environ.* 113 (5), 998–1011. <https://doi.org/10.1016/j.rse.2009.01.008>.
- Mustafa, H.M., Mustapha, A., Hayder, G., Salisu, A., 2021. Applications of IoT and artificial intelligence in water quality monitoring and prediction: a review. In: *Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021*, pp. 968–975. <https://doi.org/10.1109/ICICT50816.2021.9358675>.
- Nafsin, N., Li, J., 2022. Prediction of 5-day biochemical oxygen demand in the Buriganga River of Bangladesh using novel hybrid machine learning algorithms. *Water Environ. Res.* 94 (5), 1–17. <https://doi.org/10.1002/wer.10718>.
- Nakano, Y., Watanabe, Y.W., 2005. Reconstruction of pH in the surface seawater over the North Pacific basin for all seasons using temperature and chlorophyll-a. *J. Oceanogr.* 61 (4), 673–680. <https://doi.org/10.1007/s10872-005-0075-6>.
- Niu, C., Tan, K., Jia, X., Wang, X., 2021. Deep learning based regression for optically inactive inland water quality parameter estimation using airborne hyperspectral imagery. *Environ. Pollut.* 286, 117534. <https://doi.org/10.1016/j.envpol.2021.117534>.
- NOAA, 2020. Ocean Acidification. Retrieved on 21 December 2023 from. <https://www.noaa.gov/education/resource-collections/ocean-coasts/ocean-acidification>.
- Odermatt, D., Giardino, C., Heege, T., 2010. Chlorophyll retrieval with MERIS Case-2-Regional in perialpine lakes. *Remote Sens. Environ.* 114 (3), 607–617. <https://doi.org/10.1016/j.rse.2009.10.016>.
- Ogashawara, I., Mishra, D.R., Gitelson, A.A., 2017. Remote sensing of inland waters: background and current state-of-the-art. In: *Bio-optical Modeling and Remote Sensing of Inland Waters*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-804644-9.00001-X>.
- Ozdemir, S., Yaqub, M., Yildirim, S.O., 2023. A systematic literature review on lake water level prediction models. In: *Environmental Modelling and Software*, Vol. 163. Elsevier Ltd, p. 105684. <https://doi.org/10.1016/j.envsoft.2023.105684>.
- Padilla-Mendoza, C., Torres-Bejarano, F., Campo-Daza, G., González-Márquez, L.C., 2023. Potential of sentinel images to evaluate physicochemical parameters concentrations in water bodies—application in a Wetlands System in Northern Colombia. *Water (Switzerland)* 15 (4). <https://doi.org/10.3390/w15040789>.
- Pahlevan, N., Mangin, A., Balasubramanian, S.V., Smith, B., Alikas, K., Arai, K., Barbosa, C., Bélanger, S., Binding, C., Bresciani, M., Giardino, C., Gurlin, D., Fan, Y., Harmel, T., Hunter, P., Ishikawa, J., Kratzer, S., Lehmann, M.K., Ligi, M., Warren, M., 2021a. ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. *Remote Sens. Environ.* 258. <https://doi.org/10.1016/j.rse.2021.112366>.
- Pahlevan, N., Smith, B., Binding, C., Gurlin, D., Li, L., Bresciani, M., Giardino, C., 2021b. Hyperspectral retrievals of phytoplankton absorption and chlorophyll-a in inland and nearshore coastal waters. *Remote Sens. Environ.* 253, 112200. <https://doi.org/10.1016/j.rse.2020.112200>.

- Palmieri, V., De Carvalho, R.J., 2006. Qual2e model for the Corumbataí River. *Ecol. Model.* 198 (1–2), 269–275. <https://doi.org/10.1016/j.ecolmodel.2006.04.018>.
- Pan, Y., Lou, S., 2023. Editorial: Hydrodynamics and water environment characteristics in coastal areas under the influences of climate change and human activities. *Front. Mar. Sci.* 10, 1–3. <https://doi.org/10.3389/fmars.2023.1199807>.
- Panchanathan, A., Hossein Ahari, A., Md Touhidul Mustafa, S., Torabi Haghighi, A., 2023. An Overview of Approaches for Reducing Uncertainties in Hydrological Forecasting: Progress, and Challenges, Vol. 258, pp. 1–51. <https://doi.org/10.1016/j.earscirev.2024.104956>.
- Pattanaik, B.S., Pattanayak, A.S., Udgata, S.K., Panda, A.K., 2021. Machine learning based soft sensor model for BOD estimation using intelligence at edge. *Complex Intell. Syst.* 7 (2), 961–976. <https://doi.org/10.1007/s40747-020-00259-9>.
- Pereira, O.J.R., Merino, E.R., Montes, C.R., Barbiero, L., Rezende-Filho, A.T., Lucas, Y., Melfi, A.J., 2020. Estimating water pH using cloud-based landsat images for a new classification of the Nhecolândia Lakes (Brazilian Pantanal). *Remote Sens.* 12 (7). <https://doi.org/10.3390/rs12071090>.
- Peterson, K.T., Sagan, V., Sloan, J.J., 2020. Deep learning-based water quality estimation and anomaly detection using Landsat-8/Sentinel-2 virtual constellation and cloud computing. *GISci. Remote Sens.* 57 (4), 510–525. <https://doi.org/10.1080/15481603.2020.1738061>.
- Pollard, C.J., Stockwell, M.P., Bower, D.S., Clulow, J., Mahony, M.J., 2017. Combining ex situ and in situ methods to improve water quality testing for the conservation of aquatic species. *Aquat. Conserv. Mar. Freshwat. Ecosyst.* 27 (2), 559–568. <https://doi.org/10.1002/aqc.2700>.
- Portela, C.F., Martins, V.S., Novo, E.M.L.M., Paulino, R.S., Barbosa, C.C.F., 2024. Recent advances in geostationary satellites for inland and coastal aquatic systems: scientific research and applications ABSTRACT. *Int. J. Remote Sens.* 45 (5), 1574–1607. <https://doi.org/10.1080/01431161.2024.2314007>.
- Post, C.J., Cope, M.P., Gerard, P.D., Mastro, N.M., Vine, J.R., Stiglitz, R.Y., Hallstrom, J. O., Newman, J.C., Mikhailova, E.A., 2018. Monitoring spatial and temporal variation of dissolved oxygen and water temperature in the Savannah River using a sensor network. *Environ. Monit. Assess.* 190 (5). <https://doi.org/10.1007/s10661-018-6646-y>.
- Qambar, A.S., Al Khalidi, M.M.M., 2023. Development of local and global wastewater biochemical oxygen demand real-time prediction models using supervised machine learning algorithms. *Eng. Appl. Artif. Intell.* 118, 1–42. <https://doi.org/10.1016/j.engappai.2022.105709>.
- Qian, J., Liu, H., Qian, L., Bauer, J., Xue, X., Yu, G., He, Q., Zhou, Q., Bi, Y., Norra, S., 2022. Water quality monitoring and assessment based on cruise monitoring, remote sensing, and deep learning: a case study of Qingcaosha Reservoir. *Front. Environ. Sci.* 10, 1–13. <https://doi.org/10.3389/fenvs.2022.979133>.
- Qiao, Z., Sun, S., Jiang, Q., Xiao, L., Wang, Y., Yan, H., 2021. Retrieval of total phosphorus concentration in the surface water of miyun reservoir based on remote sensing data and machine learning algorithms. *Remote Sens.* 13 (22). <https://doi.org/10.3390/rs13224662>.
- Quang, N.H., Dinh, N.T., Dien, N.T., Son, L.T., 2023. Calibration of Sentinel-2 surface reflectance for water quality modelling in binh Dinh's Coastal Zone of Vietnam. *Sustainability* 15 (2), 1410. <https://doi.org/10.3390/su15021410>.
- Quinlan, R., Filazzola, A., Mahdian, O., Shuvo, A., Blagrove, K., Ewins, C., Moslenko, L., Gray, D.K., O'Reilly, C.M., Sharma, S., 2021. Relationships of total phosphorus and chlorophyll in lakes worldwide. *Limnol. Oceanogr.* 66 (2), 392–404. <https://doi.org/10.1002/lno.11611>.
- Radeloff, V.C., Roy, D.P., Wulder, M.A., Anderson, M.C., Cook, B.D., Crawford, C.J., Friedl, M.A., Gao, F., Gorelick, N., Hansen, M.C., Healey, S.P., Hostert, P., Hulley, G., Huntington, J.L., Johnson, D.M., Neigh, C., Lyapustin, A., Lymburner, L., Pahlevan, N., Zhu, Z., 2024. Need and vision for global medium-resolution Landsat and Sentinel-2 data products. *Remote Sens. Environ.* 300, 113918. <https://doi.org/10.1016/j.rse.2023.113918>.
- Raghul, M., Porchelvan, P., 2024. A critical review of remote sensing methods for inland water quality monitoring: progress, limitations, and future perspectives. *Water Air Soil Pollut.* 235 (2), 1–21. <https://doi.org/10.1007/s11270-024-06957-1>.
- Rahul, T.S., Brema, J., 2023. Assessment of water quality parameters in Muthupet estuary using hyperspectral PRISMA satellite and multispectral images. *Environ. Monit. Assess.* 195 (7). <https://doi.org/10.1007/s10661-023-11497-y>.
- Ritchie, H., Roser, M., 2021. Clean Water. <https://ourworldindata.org/water-access> accessed on 20 February 2023.
- Roussio, B.Z., Bertone, E., Stewart, R., Hamilton, D.P., 2020. A systematic literature review of forecasting and predictive models for cyanobacteria blooms in freshwater lakes. *Water Res.* 182, 115959. <https://doi.org/10.1016/j.watres.2020.115959>.
- Sabia, R., Fernandez-Prieto, D., Shuter, J., Donlon, C., Land, P., Reul, N., 2015. Remote sensing of surface ocean PH exploiting sea surface salinity satellite observations. In: *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 106–109. <https://doi.org/10.1109/IGARSS.2015.7325709>.
- Sagan, V., Peterson, K.T., Maimaitijiang, M., Sidike, P., Sloan, J., Greeling, B.A., Maalouf, S., Adams, C., 2020. Monitoring inland water quality using remote sensing: potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth Sci. Rev.* 205, 103187. <https://doi.org/10.1016/j.earscirev.2020.103187>.
- Sajib, A.M., Diganta, M.T.M., Rahman, A., Dabrowski, T., Olbert, A.I., Uddin, M.G., 2023. Developing a novel tool for assessing the groundwater incorporating water quality index and machine learning approach. *Groundw. Sustain. Dev.* 23, 101049. <https://doi.org/10.1016/j.gsd.2023.101049>.
- Sajib, A.M., Uddin, M.G., Jordan, C., Dabrowski, T., Cusack, C., Olbert, A.I., 2024. Assessing remote sensing data quality and alignment with in-situ measurements in ireland: a cloud coverage analysis using multi-dataset integration. In: *Civil Engineering Research in Ireland 2024 (CERI2024) Conference, 2024*, pp. 5–11.
- Salas, E.A.L., Kumaran, S.S., Partee, E.B., Willis, L.P., Mitchell, K., 2022. Potential of mapping dissolved oxygen in the Little Miami River using Sentinel-2 images and machine learning algorithms. *Remote Sens. Appl. Soc. Environ.* 26, 100759. <https://doi.org/10.1016/j.rsase.2022.100759>.
- Samarinas, N., Spiliotopoulos, M., Tziolas, N., Loukas, A., 2023. Synergistic use of earth observation driven techniques to support the implementation of water framework directive in Europe: a review. *Remote Sens.* 15 (8). <https://doi.org/10.3390/rs15081983>.
- Scheirer, R., Dybbroe, A., Raspaud, M., 2018. A general approach to enhance short wave satellite imagery by removing background atmospheric effects. *Remote Sens.* 10 (4), 560. <https://doi.org/10.3390/rs10040560>.
- Shang, W., Jin, S., He, Y., Zhang, Y., Li, J., 2021. Spatial-temporal variations of total nitrogen and phosphorus in poyang, dongting and taihu lakes from landsat-8 data. *Water (Switzerland)* 13 (12). <https://doi.org/10.3390/w13121704>.
- Sharaf El Din, E., Zhang, Y., 2017. Estimation of both optical and nonoptical surface water quality parameters using Landsat 8 OLI imagery and statistical techniques. *J. Appl. Remote Sens.* 11 (04), 1. <https://doi.org/10.1117/1.jrs.11.046008>.
- Sharaf El Din, E., Zhang, Y., Suliman, A., 2017. Mapping concentrations of surface water quality parameters using a novel remote sensing and artificial intelligence framework. *Int. J. Remote Sens.* 38 (4), 1023–1042. <https://doi.org/10.1080/01431161.2016.1275056>.
- Sheffield, J., Wood, E.F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., Verbist, K., 2018. Satellite remote sensing for water resources management: potential for supporting sustainable development in data-poor regions. *Water Resour. Res.* 54 (12), 9724–9758. <https://doi.org/10.1029/2017WR022437>.
- Shi, C., Nakajima, T., 2018. Simultaneous determination of aerosol optical thickness and water-leaving radiance from multispectral measurements in coastal waters. *Atmos. Chem. Phys.* 18 (6), 3865–3884. <https://doi.org/10.5194/acp-2017-901>.
- Shi, K., Lang, Q., Wang, P., Yang, W., Chen, G., Yin, H., Zhang, Q., Li, W., Wang, H., 2023. Dissolved oxygen concentration inversion based on Himawari-8 data and deep learning: a case study of lake Taihu. *Front. Environ. Sci.* 11 (17), 1–15. <https://doi.org/10.3389/fenvs.2023.1230778>.
- Shin, M., Kang, Y., Park, S., Im, J., Yoo, C., Quackenbush, L.J., 2020. Estimating ground-level particulate matter concentrations using satellite-based data: a review. *GISci. Remote Sens.* 57 (2), 174–189. <https://doi.org/10.1080/15481603.2019.1703288>.
- Siriwardana, H., Samarasekera, R.S.M., Anthony, D., Vithanage, D., 2024. Measurements and analysis of nitrogen and phosphorus in oceans: practice, frontiers, and insights. *Heliyon* 10 (7), e28182. <https://doi.org/10.1016/j.heliyon.2024.e28182>.
- Snyder, H., 2019. Literature review as a research methodology: an overview and guidelines. *J. Bus. Res.* 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>.
- Son, M., Yoon, N., Jeong, K., Abass, A., Logan, B.E., Cho, K.H., 2021. Deep learning for pH prediction in water desalination using membrane capacitive deionization. *Desalination* 516, 115233. <https://doi.org/10.1016/j.desal.2021.115233>.
- Song, K., Zhang, B., Wang, Z., Li, F., Duan, H., Guo, Y., 2006. Water TOC and TP concentration estimation using Landsat TM data with empirical algorithms in Chagan Lake, China. In: *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3421–3424. <https://doi.org/10.1109/IGARSS.2006.882>.
- Song, Z., He, X., Bai, Y., Dong, X., Wang, D., Li, T., Zhu, Q., Gong, F., 2023. Atmospheric correction of absorbing aerosols for satellite ocean color remote sensing over coastal waters. *Remote Sens. Environ.* 290, 113552. <https://doi.org/10.1016/j.rse.2023.113552>.
- Soomets, T., Toming, K., Jefimova, J., Jaanus, A., Põllumäe, A., Kutser, T., 2022. Deriving nutrient concentrations from Sentinel-3 OLCI data in North-Eastern Baltic Sea. *Remote Sens.* 14 (6). <https://doi.org/10.3390/rs14061487>.
- Soppa, M.A., Silva, B., Steinmetz, F., Keith, D., Scheffler, D., Bohn, N., Bracher, A., 2021. Assessment of polymer atmospheric correction algorithm for hyperspectral remote sensing imagery over coastal waters. *Sensors* 21 (12), 1–16. <https://doi.org/10.3390/s21124125>.
- Stackelberg, P.E., Belitz, K., Brown, C.J., Erickson, M.L., Elliott, S.M., Kauffman, L.J., Ransom, K.M., Reddy, J.E., 2021. Machine learning predictions of pH in the glacial aquifer system, Northern USA. *Groundwater* 59 (3), 352–368. <https://doi.org/10.1111/gwat.13063>.
- Steinmetz, F., Deschamps, P.-Y., Ramon, D., 2011. Atmospheric correction in presence of sun glint: application to MERIS. *Opt. Express* 19 (10), 9783. <https://doi.org/10.1364/oe.19.009783>.
- Sun, X., Zhang, Y., Shi, K., Zhang, Y., Li, N., Wang, W., Huang, X., Qin, B., 2022. Monitoring water quality using proximal remote sensing technology. *Sci. Total Environ.* 803. <https://doi.org/10.1016/j.scitotenv.2021.149805>.
- Tanjung, R.H.R., Indrayani, E., Hamuna, B., Agamawan, L.P.I., Alianto, A., 2023. Spatial assessment and mapping of water quality in lake sentani (indonesia) using in-situ data and satellite imagery. *Ecol. Eng. Environ. Technol.* 24 (9), 71–83. <https://doi.org/10.12912/27197050/172916>.
- Tao, J., Hill, P.S., 2019. Correlation of Remotely Sensed Surface Reflectance with Forcing Variables in six Different Estuaries. *J. Geophys. Res. Oceans* 124 (12), 9439–9461. <https://doi.org/10.1029/2019JC015336>.
- Tao, H., Song, K., Wen, Z., Liu, G., Shang, Y., Fang, C., Wang, Q., 2025. Remote sensing of total suspended matter of inland waters: past, current status, and future directions. *Eco. Inform.* 86 (February), 103062. <https://doi.org/10.1016/j.ecoinf.2025.103062>.
- Tessin, E., Hamre, B., Kristoffersen, A.S., 2024. Testing the limits of atmospheric correction over turbid norwegian fjords. *Remote Sens.* 16 (21). <https://doi.org/10.3390/rs16214082>.
- Tian, S., Guo, H., Xu, W., Zhu, X., Wang, B., Zeng, Q., Mai, Y., Huang, J.J., 2023. Remote sensing retrieval of inland water quality parameters using Sentinel-2 and multiple machine learning algorithms. *Environ. Sci. Pollut. Res.* 30 (7), 18617–18630. <https://doi.org/10.1007/s11356-022-23431-9>.

- Uddin, M.G., Nash, S., Olbert, A.I., 2021. A review of water quality index models and their use for assessing surface water quality. *Ecol. Indic.* 122, 107218. <https://doi.org/10.1016/j.ecolind.2020.107218>.
- Uddin, M.G., Nash, S., Mahammad Diganta, M.T., Rahman, A., Olbert, A.I., 2022a. Robust machine learning algorithms for predicting coastal water quality index. *J. Environ. Manag.* 321 (June), 115923. <https://doi.org/10.1016/j.jenvman.2022.115923>.
- Uddin, M.G., Nash, S., Rahman, A., Olbert, A.I., 2022b. A comprehensive method for improvement of water quality index (WQI) models for coastal water quality assessment. *Water Res.* 219 (May), 118532. <https://doi.org/10.1016/j.watres.2022.118532>.
- Uddin, M.G., Diganta, M.T.M., Sajib, A.M., Rahman, A., Nash, S., Dabrowski, T., Ahmadian, R., Hartnett, M., Olbert, A.I., 2023a. Assessing the impact of COVID-19 lockdown on surface water quality in Ireland using advanced Irish water quality index (IEWQI) model. *Environ. Pollut.* 336 (May), 122456. <https://doi.org/10.1016/j.envpol.2023.122456>.
- Uddin, M.G., Jackson, A., Nash, S., Rahman, A., Olbert, A.I., 2023b. Comparison between the WFD approaches and newly developed water quality model for monitoring transitional and coastal water quality in Northern Ireland. *Sci. Total Environ.* 901 (July), 165960. <https://doi.org/10.1016/j.scitotenv.2023.165960>.
- Uddin, M.G., Nash, S., Diganta, M.T.M., Rahman, A., Olbert, A.I., 2023c. A comparison of geocomputational models for validating geospatial distribution of water quality index. *Comput. Stat. Methodol. Model. Artif. Intell.* <https://doi.org/10.1201/9781003253051-16>.
- Uddin, M.G., Nash, S., Rahman, A., Olbert, A.I., 2023d. A novel approach for estimating and predicting uncertainty in water quality index model using machine learning approaches. *Water Res.* 229 (2022), 119422. <https://doi.org/10.1016/j.watres.2022.119422>.
- Uddin, M.G., Nash, S., Rahman, A., Olbert, A.I., 2023e. A sophisticated model for rating water quality. *Sci. Total Environ.* 868, 161614. <https://doi.org/10.1016/j.scitotenv.2023.161614>.
- Uddin, M.G., Nash, S., Rahman, A., Olbert, A.I., 2023f. Assessing optimization techniques for improving water quality model. *J. Clean. Prod.* 385, 135671. <https://doi.org/10.1016/j.jclepro.2022.135671>.
- Uddin, M.G., Rahman, A., Nash, S., Diganta, M.T.M., Sajib, A.M., Moniruzzaman, M., Olbert, A.I., 2023g. Marine waters assessment using improved water quality model incorporating machine learning approaches. *J. Environ. Manag.* 344, 118368. <https://doi.org/10.1016/j.jenvman.2023.118368>.
- Uddin, M.G., Nash, S., Rahman, A., Dabrowski, T., Olbert, A.I., 2024. Data-driven modelling for assessing trophic status in marine ecosystems using machine learning approaches. *Environ. Res.* 242, 117755. <https://doi.org/10.1016/j.envres.2023.117755>.
- UNESCO (2021). International Initiative on Water Quality (IIWQ). <https://en.unesco.org/waterquality-iiwq/wq-challenge#:~:text=One%20in%20nine%20people%20wor,ldwide%20uses%20drinking%20water%20from%20unimproved%20and%20unsafe%20sources&text=2.4%20billion%20people%20live%20without%20any%20form%20of%20sanitation&text=Lack%20of%20sanitation%20is%20one,untreated%20directly%20into%20water%20bodies> (accessed on 20 February 2023).
- Vakili, T., Amanollahi, J., 2020. Determination of optically inactive water quality variables using Landsat 8 data: a case study in Geshlagh reservoir affected by agricultural land use. *J. Clean. Prod.* 247, 119134. <https://doi.org/10.1016/j.jclepro.2019.119134>.
- van Wijk, D., Janse, J.H., Wang, M., Kroeze, C., Mooij, W.M., Janssen, A.B.G., 2024. How nutrient retention and TN:TP ratios depend on ecosystem state in thousands of Chinese lakes. *Sci. Total Environ.* 918, 170690. <https://doi.org/10.1016/j.scitotenv.2024.170690>.
- Vanhelmont, Q., Ruddick, K., 2018. Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications. *Remote Sens. Environ.* 216, 586–597. <https://doi.org/10.1016/j.rse.2018.07.015>.
- Wagle, N., Acharya, T.D., Lee, D.H., 2020. Comprehensive review on application of machine learning algorithms for water quality parameter estimation using remote sensing data. *Sensors Mater.* 32 (11), 3879–3892. <https://doi.org/10.18494/SAM.2020.2953>.
- Wang, X., Yang, W., 2019. Water quality monitoring and evaluation using remote-sensing techniques in China: a systematic review. *Ecosyst. Health Sustain.* 5 (1), 47–56. <https://doi.org/10.1080/20964129.2019.1571443>.
- Wang, D., Cui, Q., Gong, F., Wang, L., He, X., Bai, Y., 2018. Satellite retrieval of surface water nutrients in the coastal regions of the East China Sea. *Remote Sens.* 10 (12). <https://doi.org/10.3390/rs10121896>.
- Wang, L., Yue, X., Wang, H., Ling, K., Liu, Y., Wang, J., Hong, J., Pen, W., Song, H., 2020. Dynamic inversion of inland aquaculture water quality based on UAVs-WSN spectral analysis. *Remote Sens.* 12 (3), 1–19. <https://doi.org/10.3390/rs12030402>.
- Wang, X., Gong, C., Ji, T., Hu, Y., Li, L., 2021. Inland water quality parameters retrieval based on the VIP-SPCA by hyperspectral remote sensing. *J. Appl. Remote. Sens.* 15 (04), 1–17. <https://doi.org/10.1117/1.jrs.15.042609>.
- Wang, S., Shen, M., Liu, W., Ma, Y., Shi, H., Zhang, J., Liu, D., 2022. Developing remote sensing methods for monitoring water quality of alpine rivers on the Tibetan Plateau. *GISci. Remote Sens.* 59 (1), 1384–1405. <https://doi.org/10.1080/15481603.2022.2116078>.
- Wang, D., Xiang, X., Ma, R., Guo, Y., Zhu, W., Wu, Z., 2023a. A novel atmospheric correction for turbid water remote sensing. *Remote Sens.* 15 (8). <https://doi.org/10.3390/rs15082091>.
- Wang, W., Zhai, X., Yang, P., Xia, J., Hu, S., Zhou, L., Fu, C., 2023b. Evaluating the driving forces of spectral inversion methods used for assessing water quality parameters in Poyang Lake, China. *Ecohydrology* 16 (7). <https://doi.org/10.1002/eeco.2577>.
- Warren, M.A., Simis, S.G.H., Martinez-Vicente, V., Poser, K., Bresciani, M., Alikas, K., Spyarakos, E., Giardino, C., Ansper, A., 2019. Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters. *Remote Sens. Environ.* 225, 267–289. <https://doi.org/10.1016/j.rse.2019.03.018>.
- Wasehun, E.T., Hashemi Beni, L., Di Vittorio, C.A., 2024. UAV and satellite remote sensing for inland water quality assessments: a literature review. *Environ. Monit. Assess.* 196 (3). <https://doi.org/10.1007/s10661-024-12342-6>.
- Wei, J., Lee, Z., Shang, S., 2016. A system to measure the data quality of spectral remote-sensing reflectance of aquatic environments. *J. Geophys. Res. Oceans* 121, 3372–3380. <https://doi.org/10.1002/2016JC012126>. Received.
- Wen, Z., Wang, Q., Liu, G., Jacinthe, P.A., Wang, X., Lyu, L., Tao, H., Ma, Y., Duan, H., Shang, Y., Zhang, B., Du, Y., Du, J., Li, S., Cheng, S., Song, K., 2022. Remote sensing of total suspended matter concentration in lakes across China using Landsat images and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* 187, 61–78. <https://doi.org/10.1016/j.isprsjprs.2022.02.018>.
- WFD, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 Establishing a Framework for Community Action in the Field of Water Policy. Retrieved on 20 March 2023 from. <http://data.europa.eu/eli/dir/2000/60/2014-11-20>.
- Wilson, H., Raasakka, N., Spyarakos, E., Millar, D., Neely, M.B., Salyani, A., Pawar, S., Chernov, I., de Ague, S.K.L., Aguilar Vega, X., Akinsemolu, A., Baltodano Martinez, A., Cillero Castro, C., Del Valle, M., Fadlseed, M., Ferral, A., Hassen, J. M., Jiang, D., Mubambi, T.K., Tyler, A., 2025. Unlocking the global benefits of Earth Observation to address the SDG 6 in situ water quality monitoring gap. *Front. Remote Sens.* 6 (March), 1–8. <https://doi.org/10.3389/frsen.2025.1549286>.
- Wu, C., Wu, J., Qi, J., Zhang, L., Huang, H., Lou, L., Chen, Y., 2010. Empirical estimation of total phosphorus concentration in the mainstream of the Qiantang River in China using Landsat TM data. *Int. J. Remote Sens.* 31 (10), 2309–2324. <https://doi.org/10.1080/01431160902973873>.
- Wu, S., Qi, J., Yan, Z., Lyu, F., Lin, T., Wang, Y., Du, Z., 2022. Spatiotemporal assessments of nutrients and water quality in coastal areas using remote sensing and a spatiotemporal deep learning model. *Int. J. Appl. Earth Obs. Geoinf.* 112, 102897. <https://doi.org/10.1016/j.jag.2022.102897>.
- Xiao, Y., Guo, Y., Yin, G., Zhang, X., Shi, Y., Hao, F., Fu, Y., 2022. UAV multispectral image-based urban river water quality monitoring using stacked ensemble machine learning algorithms—a case study of the Zhanghe River, China. *Remote Sens.* 14 (14). <https://doi.org/10.3390/rs14143272>.
- Xiong, J., Lin, C., Ma, R., Cao, Z., 2019. Remote sensing estimation of Lake total phosphorus concentration based on MODIS: a case study of Lake Hongze. *Remote Sens.* 11 (17), 1–19. <https://doi.org/10.3390/rs11172068>.
- Xiong, J., Lin, C., Cao, Z., Hu, M., Xue, K., Chen, X., Ma, R., 2022. Development of remote sensing algorithm for total phosphorus concentration in eutrophic lakes: conventional or machine learning? *Water Res.* 215, 118213. <https://doi.org/10.1016/j.watres.2022.118213>.
- Xu, Y., Zhang, Y., Zhang, D., 2010. Retrieval of dissolved inorganic nitrogen from multi-temporal MODIS data in Haizhou bay. *Mar. Geod.* 33 (1), 1–15. <https://doi.org/10.1080/01490410903530257>.
- Yang, Y., Jin, S., 2023. Long-Time water quality variations in the Yangtze River from landsat-8 and sentinel-2 images based on neural networks. *Water (Switzerland)* 15 (21), 1–22. <https://doi.org/10.3390/w15213802>.
- Yang, H., Kong, J., Hu, H., Du, Y., Gao, M., Chen, F., 2022a. A review of remote sensing for water quality retrieval: progress and challenges. *Remote Sens.* 14 (8). <https://doi.org/10.3390/rs14081770>.
- Yang, L., Driscoll, J., Sarigai, S., Wu, Q., Lippitt, C.D., Morgan, M., 2022b. Towards synoptic water monitoring systems: a review of AI methods for automating water body detection and water quality monitoring using remote sensing. *Sensors* 22 (6). <https://doi.org/10.3390/s22062416>.
- Yang, Z., Gong, C., Ji, T., Hu, Y., Li, L., 2022c. Water quality retrieval from ZY1-02D hyperspectral imagery in urban water bodies and comparison with sentinel-2. *Remote Sens.* 14 (19). <https://doi.org/10.3390/rs14195029>.
- Yang, W., Fu, B., Li, S., Lao, Z., Deng, T., He, W., He, H., Chen, Z., 2023. Monitoring multi-water quality of internationally important karst wetland through deep learning, multi-sensor and multi-platform remote sensing images: a case study of Guilin, China. *Ecol. Indic.* 154 (July), 110755. <https://doi.org/10.1016/j.ecolind.2023.110755>.
- Ye, H., Guo, S., Li, F., Li, G., 2013. Water quality evaluation in Tidal River reaches of Liaohe River Estuary, China using a revised QUAL2K model. *Chin. Geogr. Sci.* 23 (3), 301–311. <https://doi.org/10.1007/s11769-013-0586-9>.
- Yu, X., Yi, H., Liu, X., Wang, Y., Liu, X., Zhang, H., 2016. Remote-sensing estimation of dissolved inorganic nitrogen concentration in the Bohai Sea using band combinations derived from MODIS data. *Int. J. Remote Sens.* 37 (2), 327–340. <https://doi.org/10.1080/01431161.2015.1125555>.
- Yuan, X., Wang, S., Fan, F., Dong, Y., Li, Y., Lin, W., Zhou, C., 2022. Spatiotemporal dynamics and anthropologically dominated drivers of chlorophyll-a, TN and TP concentrations in the Pearl River Estuary based on retrieval algorithm and random forest regression. *Environ. Res.* 215. <https://doi.org/10.1016/j.envres.2022.114380>.
- Zhang, Y., Wu, L., Ren, H., Deng, L., Zhang, P., 2020. Retrieval of water quality parameters from hyperspectral images using hybrid Bayesian probabilistic neural network. *Remote Sens.* 12 (10). <https://doi.org/10.3390/rs12101567>.
- Zhang, H., Yao, B., Wang, S., Wang, G., 2021a. Remote sensing estimation of the concentration and sources of coloured dissolved organic matter based on MODIS: A case study of Erhai lake. *Ecol. Indic.* 131, 108180. <https://doi.org/10.1016/j.ecolind.2021.108180>.
- Zhang, Y., Wu, L., Deng, L., Ouyang, B., 2021b. Retrieval of water quality parameters from hyperspectral images using a hybrid feedback deep factorization machine model. *Water Res.* 204, 117618. <https://doi.org/10.1016/j.watres.2021.117618>.

- Zhang, H., Xue, B., Wang, G., Zhang, X., Zhang, Q., 2022a. Deep learning-based water quality retrieval in an impounded lake using landsat 8 imagery: an application in Dongping Lake. *Remote Sens* 14 (18), 4505. <https://doi.org/10.3390/rs14184505>.
- Zhang, L., Zhang, L., Cen, Y., Wang, S., Zhang, Y., Sultan, M., Tong, Q., Zhang, L., Zhang, L., Wang, S., Zhang, Y., Sultan, M., Zhang, L., Huang, Y., 2022b. Prediction of total phosphorus concentration in macrophytic lakes using chlorophyll-sensitive bands: a case study of Lake Baiyangdian. *Remote Sens* 14 (13). <https://doi.org/10.3390/rs14133077>.
- Zhang, Y., Jin, S., Wang, N., Zhao, J., Guo, H., Pellikka, P., 2022c. Total phosphorus and nitrogen dynamics and influencing factors in dongting lake using landsat data. *Remote Sens* 14 (22), 5648. <https://doi.org/10.3390/rs14225648>.
- Zhang, Y., He, X., Lian, G., Bai, Y., Yang, Y., Gong, F., Wang, D., Zhang, Z., Li, T., Jin, X., 2023a. Monitoring and spatial traceability of river water quality using Sentinel-2 satellite images. *Sci. Total Environ.* 894, 164862. <https://doi.org/10.1016/j.scitotenv.2023.164862>.
- Zhang, Y., Kong, X., Deng, L., Liu, Y., 2023b. Monitor water quality through retrieving water quality parameters from hyperspectral images using graph convolution network with superposition of multi-point effect: a case study in Maozhou River. *J. Environ. Manag.* 342 (May), 118283. <https://doi.org/10.1016/j.jenvman.2023.118283>.
- Zhao, J., Jin, S., Zhang, Y., 2023. Dynamic water quality changes in the main stream of the yangtze river from multi-source remote sensing data. *Remote Sens* 15 (10). <https://doi.org/10.3390/rs15102526>.
- Zhu, W., Xia, W., 2023. Effects of atmospheric correction on remote sensing statistical inference in an aquatic environment. *Remote Sens* 15 (7), 1–9. <https://doi.org/10.3390/rs15071907>.
- Zhu, X., Guo, H., Huang, J.J., Tian, S., Xu, W., Mai, Y., 2022a. An ensemble machine learning model for water quality estimation in coastal area based on remote sensing imagery. *J. Environ. Manag.* 323 (February), 116187. <https://doi.org/10.1016/j.jenvman.2022.116187>.
- Zhu, B., Bai, Y., Zhang, Z., He, X., Wang, Z., Zhang, S., Dai, Q., 2022b. Satellite remote sensing of water quality variation in a semi-enclosed Bay (Yueqing Bay) under strong anthropogenic impact. *Remote Sens* 14 (3). <https://doi.org/10.3390/rs14030550>.
- Zibordi, G., Kwiatkowska, E., Mélin, F., Talone, M., Cazzaniga, I., Dessailly, D., Gossn, J. I., 2022. Assessment of OLCI-A and OLCI-B radiometric data products across European seas. *Remote Sens. Environ.* 272 (August 2021). <https://doi.org/10.1016/j.rse.2022.112911>.
- Ziyad Sami, B.F., Latif, S.D., Ahmed, A.N., Chow, M.F., Murti, M.A., Suhendi, A., Ziyad Sami, B.H., Wong, J.K., Birima, A.H., El-Shafie, A., 2022. Machine learning algorithm as a sustainable tool for dissolved oxygen prediction: a case study of Feitsui Reservoir, Taiwan. *Sci. Rep.* 12 (1), 1–12. <https://doi.org/10.1038/s41598-022-06969-z>.