remote sensing

Review

Progress in Remote Sensing of Heavy Metals in Water

Xiaoling Xu [1,2], Jiayi Pan [1,3,*], Hua Zhang [1,2] and Hui Lin [1,2]

[1] School of Geography and Environment, Jiangxi Normal University, Nanchang 330022, China; xuxiaoling@jxnu.edu.cn (X.X.); zhangh2013@jxnu.edu.cn (H.Z.); huilin@cuhk.edu.hk (H.L.)
[2] Key Laboratory of Poyang Lake Wetland and Watershed Research, Ministry of Education, Nanchang 330022, China
[3] Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, Hong Kong, China
* Correspondence: panj@cuhk.edu.hk

Abstract: This review article details the advancements in detecting heavy metals in aquatic environments using remote sensing methodologies. Heavy metals are significant pollutants in aquatic environment, and their detection and monitoring are crucial for predicting water quality. Traditional in situ water sampling methods are time-consuming and costly, highlighting the advantages of remote sensing techniques. Analysis of the reflectance and absorption characteristics of heavy metals has identified the red and near-infrared bands as the sensitive wavelengths for heavy metal detection in aquatic environments. Several studies have demonstrated a correlation between total suspended matter and heavy metals, which forms the basis for retrieving heavy metal content from TSM data. Recent developments in hyperspectral remote sensing and machine (deep) learning technologies may pave the way for developing more effective heavy metal detection algorithms.

Keywords: heavy metal; remote sensing algorithm; hyperspectral remote sensing; machine learning method

check for updates

Citation: Xu, X.; Pan, J.; Zhang, H.; Lin, H. Progress in Remote Sensing of Heavy Metals in Water. *Remote Sens.* **2024**, *16*, 3888. https://doi.org/10.3390/rs16203888

Academic Editors: Teodosio Lacava and Gabriel Senay

Received: 10 September 2024
Revised: 23 September 2024
Accepted: 17 October 2024
Published: 19 October 2024

1. Overview of Heavy Metal Pollution

Heavy metals are recognized for their significant biotoxicity and resistance to degradation and are prevalent across various environmental media in nature. These pollutants are characterized by their mobility, persistence, and propensity for bioaccumulation, posing substantial threats to ecological systems [1]. Due to their non-biodegradable nature, toxicity, endurance, and accumulation in the food chain, heavy metals can infiltrate the human body through dietary intake, ultimately compromising human health [2,3]. In aquatic environments, heavy metals manifest in three primary forms: particulate, where they adhere to suspended particles; dissolved, where they are complexed by dissolved organic matter; and biological, where they are assimilated by phytoplankton [4–6]. Consequently, the distribution of heavy metal concentrations in water is governed by three components: suspended sediments, phytoplankton, and dissolved organic substances. Human activities, particularly through sediment deposition in rivers, render heavy metals a critical ecological and public health concern globally. Common heavy metals include cadmium (Cd), arsenic (As), lead (Pb), mercury (Hg), copper (Cu), zinc (Zn), and nickel (Ni). The environmental concern largely centers on highly biotoxic metals such as Hg, Cd, Pb, Cr, and As, as well as other toxic metals like Zn, Cu, cobalt (Co), Ni, and tin (Sn) [7].

The origins of heavy metals in aquatic systems are multifaceted, deriving from both the natural weathering of rocks and various dispersed pollution sources, including dust, rainfall, and water quality transformations [8]. Predominantly, mining and industrial wastewater significantly elevate metal concentrations in riverine environments. Sediments in river–lake systems are chiefly influenced by anthropogenic activities, encompassing industrial emissions, chemical fertilizer applications, and aquaculture. Domestic wastewater also contains heavy metals, and the combustion of liquid fuels, vehicle wear, and the weathering of track or road materials release particulates laden with heavy metals. Thus, industrial and agricultural waste, runoff, and atmospheric deposition serve as primary contributors to surface water contamination [9–12]. Heavy metals in water can manifest in various forms, sometimes occurring as ions. These metals can react with S_2 to form metal sulfides, a critical depositional form of toxic heavy metals. In high-oxygen environments, these metal sulfides may dissolve, re-releasing heavy metal ions and causing pollution [13–16].

Effective monitoring of heavy metals in water is essential for protecting aquatic environments from contamination by these pollutants. Traditionally, this involves collecting water samples and subsequently analyzing heavy metal concentrations in laboratories. However, this method is laborious, costly, and inefficient, making it impractical for extensive water body surveillance. Satellite remote sensing technology has emerged as a formidable alternative, demonstrating its efficacy in detecting key water quality indicators such as chlorophyll a (Chl-a), suspended particulate matter (SPM), colored dissolved organic matter (CDOM), turbidity, etc. The adoption of remote sensing techniques offers a promising substitute for traditional methods, enabling more comprehensive monitoring of heavy metal concentrations in aquatic environments. Therefore, this paper presents a review of the latest advancements in the remote sensing of heavy metals in water bodies, and some polluting light metals are also discussed, highlighting remote sensing's potential as a transformative tool for environmental protection.

2. Spectral Characteristics of Heavy Metal

Theoretically, remote sensing reflectance of a certain substance in water can be linked to the absorption and backscattering by [17].

$$R(\lambda) = G\frac{b_{b(\lambda)}}{a(\lambda) + b_b(\lambda)}, \tag{1}$$

where $a(\lambda)$ and $b_b(\lambda)$ represent absorption and backscattering, respectively; G is a constant that depends on the incident light field. Equation (1) elucidates that the optimal bands for detecting specific water constituents (including heavy metals) are those characterized by significant backscattering, facilitating the remote sensing sensor's ability to discern the backscattering signal emanating from the water content. Bands with high absorption properties are also valuable for detection purposes, as absorption can attenuate remote sensing reflectance, thereby manifesting the influence of the specific water constituents on the reflectance spectra. In practical applications, the presence of complex optically sensitive materials in water necessitates careful selection of the optimal bands, taking into account the spectral properties of other aqueous substances to circumvent signal interference among these components.

Understanding reflection and absorption spectra of the heavy metals is a key step in detecting or quantifying heavy metal concentrations in water bodies. It is a base for constructing theoretical and empirical algorithms in retrieving heavy metal concentrations using remote sensing spectral reflectance data. As early as the 1940s, detailed observations on the spectral characteristics of heavy metal began to appear [18,19]. For copper ammonia solution, it was found that there is an absorption peak between 600 and 650 nm [20]. The absorption spectra of copper sulfate ($CuSO_4$) solutions were measured in the wavelength range of 400–1100 nm, and the results revealed that the maximum of molal absorption coefficient is at 780 nm of the red band [21].

Recently, Liang et al. [22] improved a spectral measurement system including an Analytical Spectral Devices (ASD) (Malvern PANalytical, Longmont, CO, USA) spectrometer, with the adjustable path length of passing light that can provide accurate results at the visible and near-infrared wavelength ranges. The absorption coefficient spectrum (400–900 nm) of copper ions in water was measured by this spectrometer system. The absorption coefficient in the blue and green solutions is very low but rapidly increases from the red to near-infrared region; maximum absorption is observed at 810 nm, and the absorption coefficient is larger in red than in blue and green [22]. Deng et al. [23] further investigated the absorption coefficient spectrum in 400–900 nm of hydrated and complex ferric ion in water with the same spectrum measurement system. The results indicate that all the absorption coefficients of three types of the types of ferric ions in blue are larger than green and red and slightly change from red to near-infrared. The value of the absorption coefficient of three types of ferric ion from large to small (potassium ferricyanide, ferric sulfate, and ferric chloride) ranges from 400 nm to 455 nm. The absorption coefficient of ferric sulfate is larger than potassium ferricyanide and ferric chloride and ranges from 460 nm to 900 nm. The best wavelengths of absorption spectra used to estimate concentrations are 417, 469, and 624 nm for ferric chloride, potassium ferricyanide, and ferric sulfate, respectively [23].

The spectroscopic examination of cadmium compounds uncovers distinctive attributes within the absorption (or extinction) coefficient spectrum of CdS, marked by two depressions at 550 nm and 830 nm and a prominent peak at 675 nm [24]. The extinction coefficient spectrum of CdO demonstrates a reduction from the purple to the near-infrared spectrum. For both substances, the extinction coefficients in the blue spectrum exceed those in the green and red spectra. In terms of reflectance, CdS demonstrates elevated levels in the yellow and red zones, in stark contrast to the lower values noted in the purple and blue regions. Meanwhile, CdO's reflectance exhibits a consistent upward trend from 525 to 900 nm. These observations emphasize that the highest reflectance values are found in the yellow and red spectra, indicating these bands' heightened sensitivity to fluctuations in cadmium levels within water bodies [24]. By using multi-concentration pollution standard solution samples of Pb and Cd in a single water body, the reflectance spectra were measured; optimal spectral bands for Pb were identified at 554 nm and 748 nm, and for Cd at 1102 nm and 1137 nm [25].

In addition to laboratory measurements, field observations were undertaken to delineate the spectral properties of aquatic environments contaminated with heavy metals in real-world scenarios. Utilizing an Analytical Spectral Devices (ASD) (Malvern PANalytical, Longmont, CO, USA) spectrometer, Liang et al. [26] documented the water-leaving reflectance spectrum from mine drainage at Dabaoshan Mountain. They observed a stable reflective peak within the 600 to 700 nm range (red light spectrum). Generally, the findings demonstrated that the drainage water exhibited maximal absorption in the purple band and minimal absorption in the red band. Beginning at 400 nm, the absorption coefficient decreased precipitously, moderated in the blue and green bands, then diminished rapidly again approaching the yellow band until it reached a nadir at 676 nm. This specific spectral pattern was attributed to the presence of ferric sulfate and its hydrolysate.

These studies suggest that most heavy metal ions in water exhibit sensitive bands in absorption within the red spectral bands. This recurring pattern indicates that the red bands could be the most effective choice for detecting and quantifying heavy metal concentrations in aquatic environments. Enhancing our understanding of these spectral characteristics can significantly improve the accuracy of remote sensing methods used in monitoring and assessing pollution levels, ultimately aiding in the effective management and remediation of contaminated water bodies. Table 1 lists sensitive bands of heavy metal solutions measured in the laboratory.

Table 1. Sensitive bands of heavy metal solutions measured in laboratory.

Heavy Metal Solutions (Concentration)	Absorption	Reflectance	References
Cu -NH_3 solution (1–5 mol/m^3)	600–650 nm		Bjerrum and Agarwal [20]
$CuSO_4$	780 nm		Jancso [21]
Cu ions	810 nm		Liang et al. [22]
$FeCl_3$ (0.138~0.390 g/L)	417 nm		Deng et al. [23]
C6FeK3N6 (potassium ferricyanide) (0.023 g/L)	469 nm		Deng et al. [23]
$Fe_2(SO_4)_3$ (0.123~0.161 g/L)	624 nm		Deng et al. [23]
CdS (0.048~0.494 g/L)	675 nm	Yellow, red	Liang et al. [24]
CdO (0.106~0.129 g/L)	Blue band	Red	Liang et al. [24]
Standard solution of Pb (0.01~2.0 mg/L)		654, 535 nm	Luo et al. [25]
Standard solution of Cd (0.01~2.0 mg/L)		1067, 1079 nm	Luo et al. [25]

3. Empirical Remote Sensing Algorithms of Heavy Metal in Waters

Although certain heavy metal solutions exhibit distinct spectral characteristics amenable to remote sensing detection, the majority of heavy metals do not manifest detectable spectral signals. Consequently, direct spectral analysis often fails to accurately quantify heavy metal concentrations. Fortunately, there exist optically sensitive materials in water that demonstrate a strong correlation with heavy metal content. For instance, in a tidally influenced wetland, methylmercury (MeHg) exhibits a significant correlation with dissolved organic matter (DOM). Thus, by quantifying DOM concentrations through spectral reflectance, it becomes feasible to detect MeHg using remote sensing techniques [27]. Additionally, in the same region, factors such as turbidity and fluorescent dissolved organic matter (FDOM) have been successfully employed to estimate mercury concentration in unfiltered water samples (UTHg) with a 94% accuracy across different seasons [28].

By analyzing sampled data in the Pearl River estuary (PRE), Chen et al. [29] found that the spectral band data at 711 nm have a correlation of 0.51 with both Cu and Zn concentrations, the band ratio B711/B406 has a high correlation (0.75) with Cu, and band ratio B711/B416 has the highest correlation (0.80) with Zn. These high correlations revealed that empirical algorithms could be developed to estimate the Zn and Cu concentrations with the spectral data. Thus, Zn, Pb, and Cu concentrations in the PRE were estimated using empirical models with single band and band ratio as input features, and the retrieval correlation coefficient was higher than 0.7 [4]. Based on these algorithms and Landsat data, the spatial patterns of the Zn, Pb, and Cu concentration were derived; this study shows that the red band of 660 nm is a key band to estimate the concentrations of Zn, Pb, and Cu [4]. A substantial majority (over 80%) of heavy metals are sequestered and complexed by suspended sediments, with the degree of metal adsorption being contingent upon the load and size of these sediments. Therefore, information regarding the load and size of suspended sediments could serve as critical indicators for the assessment of heavy metal concentrations via remote sensing methods and thus the concentration of the heavy metal, Zn, can be written as follows [30]:

$$C_{Zn} = f(C_{SS}, \gamma) \tag{2}$$

where C_{Zn} represents the Zn concentration in the particulate phase, C_{ss} is the suspended sediment (SS) concentration, and γ is the particulate scattering spectral slope, reflecting size information of suspended sediments, which could be estimated by the remote sensing reflectance. Liu et al. [30] revealed there are strong linear relationships of C_{Zn} in logarithm with SS concentration and γ, and therefore an empirical model could be developed based on this linear relationship and in situ data with the correlation coefficient of ~0.8. In this empirical algorithm, the red–near-infrared (NIR) bands of 620~779 nm were selected as the major band and the 520 nm as the reference band to generate band ratios [30].

With the development of remote sensing technology, satellite hyperspectral sensors have been launched with operational observations of the earth surface, providing more choices of wavelength bands for water quality parameter retrieval [31–33]. Corresponding to these hyperspectral data, more spectral bands could be selected for heavy metal concentration retrieval in field spectrometer measurements. Rostom et al. [34] used a FieldSpec-3 ASD(Malvern PANalytical, Longmont, CO, USA) Spectroradiometer to observe the hyperspectral reflectance in Mariut Lake of Egypt, and based on the spectral data and the sampled heavy metal concentrations, they established the empirical models for retrieval of the concentrations of Mn, Co, Cr, Ni, Cu, Cd, Pb, Fe, and Zn. Except Co, Cd, and Zn, which had correlation coefficients less than 0.6, the correlation coefficients of other heavy metals are higher than 0.7, and more than 10 wavelength banks were employed to estimate the concentration of these heavy metals in a range of 366–1013 nm [34]. The key spectral wavelengths associated with copper ion concentrations in solutions predominantly fall within the 900 to 1100 nm range [35–37]. Utilizing laboratory spectral reflectance measurements for copper ion solutions, Huang et al. [35] constructed an algorithm designed to predict copper ion concentrations in the NIR bands at 901.11 and 1079.08 nm using logarithmically transformed reflectance data, which achieved an adjusted R^2 of 0.827. This result implies that applying a logarithmic transformation to the spectral data enhances the model's predictive performance by compressing the data while maintaining its fundamental spectral attributes and reducing low-frequency noise [35,38].

The transport and distribution of heavy metals are associated with suspended particles due to their high binding affinity [39–42]. Thus, it is a practical way to estimate the heavy metal concentrations in water bodies by measuring Total Suspended Solids (TSS) [39,43,44]. Based on this method, Swain and Sahoo [39] developed an empirical algorithm to calculate the turbidity from the red (L_3) (631–692 nm) and NIR (L_4) (772–898 nm) Landsat bands and the relationship between the TSS and turbidity. Then, the empirical algorithms of heavy metals of Fe, Zn, Cu, Cr, Pb, and Cd from TSS were built up, and finally the algorithms of heavy metal from input features of L_3 and L_4 were established with correlation coefficients between 0.76 and 0.93 in the validation [39].

By measuring in situ spectral reflectance data and heavy metal concentrations in sampled waters, Rajesh et al. [45] found linear relationships between the spectral data and the concentrations of heavy metals, such as As, Cd, Cr, Cu, Fe, Pb, Ni, Co, Mg, Be, B, Li, and Mo, with correlation coefficients higher than 0.7 using recursive linear regression analysis. The highest correlation coefficient, 0.94, was observed for Cd and Cr, indicating that spectral measurements can effectively assess heavy metal pollution [45]. In a real case of polluted wastewater containing heavy metals, it was revealed that as heavy metal pollution increased, the spectral reflectance pattern changed significantly from orange-red to brownish-red. This demonstrates that spectral characteristics can be used to discern the extent of heavy metal pollution [46].

4. Machine Learning Algorithms

4.1. Machine Learning Versus Traditional Algorithms

In recent years, machine learning (including deep learning) technology has become increasingly prominent for water quality retrieval in various aquatic environments, including rivers, lakes, estuaries, and coastal waters [47–49]. Machine learning methodologies markedly elevate the precision of remote sensing algorithms in the retrieval of aquatic parameters by surmounting a range of challenges that conventional approaches often encounter. A predominant issue lies in the intricate and non-linear associations between water quality indicators and the spectral data captured by remote sensing platforms. Traditional models frequently falter in accurately representing these complexities, whereas machine learning algorithms, particularly sophisticated non-linear models such as neural networks, demonstrate superior proficiency in encapsulating these nuanced relationships. This capability leads to enhanced predictive accuracy for critical parameters, including chlorophyll concentration and turbidity [50–52]. Conventional remote sensing algorithms typically hinge on predefined assumptions, which may not exhibit robust generalizability across diverse environmental contexts. In contrast, machine learning models, driven by empirical data, eschew these limitations by adapting to a multitude of conditions, thereby yielding more precise and reliable predictions. The strength of these models is further amplified by their adeptness at integrating heterogeneous datasets from multiple sources, such as various satellites and sensors, thereby refining overall predictive accuracy through ensemble learning techniques [53–55].

Machine learning excels in the domain of automatic feature extraction, facilitating the identification of salient patterns within the data that might elude traditional manual analysis. This enhances the accuracy of water parameter retrieval, particularly in complex or data-rich scenarios. The intrinsic flexibility of machine learning models allows for them to be retrained or fine-tuned with new datasets, ensuring their adaptability to evolving conditions or region-specific variations—a critical attribute for sustained monitoring efforts [56–59]. In addition to their adaptability, machine learning models exhibit resilience against noise and data uncertainties, with techniques such as regularization and probabilistic modeling fortifying their robustness. This resilience ensures that predictions remain reliable even in the presence of data imperfections. Finally, the computational efficiency of machine learning models enables rapid inference post-training, rendering them highly suitable for real-time or near-real-time applications, such as the monitoring of water quality during natural disasters or other dynamic environmental events [60–62]. In summation, machine learning significantly enhances the accuracy, adaptability, and efficiency of remote sensing algorithms in the retrieval of aquatic parameters by addressing inherent challenges with advanced, data-driven technique.

4.2. Machine Learning Methods

The machine learning technology application has not only advanced the retrieval of non-optically sensitive parameters but also offered fresh perspectives. For example, Zhou et al. [63] employed a backpropagation (BP) neural network machine learning method to retrieve some metal ion concentrations in West Taijinar Lake, Qaidam Basin, China from satellite spectral band data. The BP neural network is a form of supervised learning algorithm characterized by its multilayer feedforward structure, which includes an input layer, several hidden layers, and an output layer. The learning process in a BP neural network involves modifying its weights to minimize the discrepancy between the predicted and actual outputs observed during training. This process of adjusting weights based on errors is conducted in a reverse manner, flowing from the output layer back to the input layer, which is why it is termed backpropagation [64]. At first, 20 optimal bands were identified from 115 bands of the HJ-1A hyperspectral dataset with the Adaptive Band Selection (ABS) method. Employing these selected bands in the ABS analysis, they developed the BP neural network model capable of estimating K^+, Mg^{2+}, Na^+, Cl^-, and ${SO_4}^{2-}$ ion concentrations with an accuracy exceeding 85% [63].

Sentinel-2A data were utilized to estimate lithium concentrations within the Arizaro Salt Lake in Salta Province, Argentina [65]. The remote sensing models for lithium concentration in the lake were established using selected Sentinel-1A bands (R4, R6, R8, and R8a), with their corresponding wavelengths listed in Table 2. Three distinct remote sensing models were developed: the Principal Component Analysis (PCA) regression model, BP neural network model, and Random Forrest (RF) model. The RF model is an ensemble learning method primarily used for classification and regression tasks. It works by constructing multiple decision trees during training and merging their results to improve accuracy and prevent overfitting. Each tree in the forest is built on a different random subset of the data, and when making predictions, the model takes the majority vote (for classification) or the average (for regression) of the trees' predictions. This method enhances the model's robustness and generalization by reducing the variance often seen in individual decision trees [66].

Table 2. Sentinel-2A bands and wavelengths for Li retrieval models.

Bands	R4	R6	R8	R8a
Wavelength (nm)	665	740	842	865

The PCA regression model yielded a determination coefficient (R^2) of 0.725, the BP neural network model achieved an R^2 of 0.731, while the RF model demonstrated the highest R^2 of 0.771. These results indicate that the RF model exhibited relatively lower relative errors, ranging from a minimum of 5.2% to a maximum of 21.3% for the detection range of 6.8 to 96.0 mg/L [65].

Utilizing Landsat-8 remote sensing imagery (Band1–7), the estimation of lithium concentration in the North Lake of Zabuye Salt Lake in Tibet was conducted through remote sensing observations, employing the Light Gradient Boosting Machine (LightGBM) model [67]. The LightGBM is a high-performance, gradient boosting framework developed by Microsoft. It is known for its efficiency and speed, particularly in handling large datasets.

LightGBM employs a unique leaf-wise tree growth strategy, as opposed to the traditional level-wise growth, which can lead to faster training times and improved accuracy. It utilizes histogram-based algorithms for optimizing splits, thereby reducing memory usage and increasing computational efficiency. LightGBM supports categorical features directly, eliminating the need for manual feature encoding. Additionally, it offers capabilities for parallel and GPU learning, making it highly scalable and suitable for a variety of machine learning tasks including classification, regression, and ranking problems [68].

Based on many advantages, the LightGBM integrates multiple decision trees to synthesize decision results, overcomes the limitations of single learners, and addresses the issues of the limited bands in Landsat remote sensing data [69–71]. The model can well estimate the Li concentrations in the range of 424.7–568.1 mg/L, with the accuracy of an average relative error (AVE) of 5.3%, and a coefficient of determination (R^2) of 0.87 [67]. Using Landsat-8 remote sensing imagery (Band1–7), the wavelength range of Landsat bands is shown in Table 3.

Table 3. Wavelength range of Landsat bands.

Landsat Band	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Wavelength (nm)	433–453	450–515	525–600	630–680	845–885	560–1660	2100–2300
Resolution (m)	30	30	30	30	30	30	30

A recent study conducted in Dalian Lake, Shanghai, China, reported the retrieval of Cu and Fe ions utilizing in situ hyperspectral data and a machine learning model [72]. Initially, optimal features for estimating concentrations of total organic carbon (TOC), Chl-a, and total phosphorus (TP) were identified using the random forest (RF) approach, considering the original spectrum (OR), first-order derivative reflectance (FDR), and second-order derivative reflectance (SDR). Subsequently, the Genetic Algorithm–Partial Least Squares Regression (GA-PLSR) approach was used to develop an effective algorithm. The Genetic Algorithm–Partial Least Squares Regression (GA-PLSR) approach combines genetic algorithms (GAs) with partial least squares regression (PLSR) to optimize feature selection and predictive modeling in high-dimensional data. This method uses the evolutionary capabilities of GAs to select the most relevant variables, which are then used in PLSR to model the relationship between predictors and response variables. The hybrid approach is particularly effective in handling datasets with multicollinearity or more predictors than observations, enhancing both model simplicity and predictive accuracy [73]. The GA-PLSR was applied for the inversion of Cu and Fe ions, utilizing in-situ hyperspectral bands at 497, 665, 686, 831, and 935 nm for Cu retrieval, and 700, 746, 801, 948, and 993 nm for Fe retrieval. The results indicated that the GA-PLSR model surpassed both the PLSR and RF models, achieving an R^2 of 0.75 in a range of 0.002–0.025 mg/L for Cu inversion. Similarly, for Fe inversion, the GA-PLSR model demonstrated superior performance with an R^2 of 0.73 in a range of 0.014–0.262 mg/L [72].

Based on these studies, we can find that compared with empirical algorithms, the machine learning algorithms allow for more input features for model establishment, which can better make use of hyperspectral data with more than 100 bands. Some empirical and machine learning algorithms are summarized in Table 4.

Table 4. Algorithms for retrieval of polluted metals from remote sensing reflectance.

Sources	Metals	Significant Band and Band Ratio	Algorithm	Detection Range	Sample No.	R^2 (Validation)	MRE	RMSE
Chen et al. [4].	Cu, Pb, Zn	B660/B560, B660/B485, B560/B485	$\log Cu = 2.23454\log\left(R_{660}/R_{560}\right) + 1.4445$ $\log Pb = 0.5310\log\left[\left(R_{660} + R_{560}\right)/R_{485}\right] + 0.0777$ $\log Zn = 1.784\log\left(R_{660}/R_{560}\right) + 1.7742$		Cu: 80 Pb: 56 Zn: 83	Cu: 0.75 Pb: 0.83 Zn: 0.75		
Liu, et al. [30].	Zn	B681/B560, B754/B560, B620/B560, B779/B865	$\log Zn = \begin{cases} 0.4 \times \left(\frac{2.3776R_{620}}{R_{560}} - 0.7467\right) + \frac{0.6R_{779}}{R_{865}} & \frac{R_{620}}{R_{560}} < 0.9 \\ 0.01 \times \left[0.8971 \times \frac{(R_{681}+R_{754})}{R_{560}} + 0.5125\right] + 0.9\frac{R_{779}}{R_{865}} & \frac{R_{620}}{R_{560}} \geq 0.9 \end{cases}$	5.67–86.62 μg/L	65	0.88	24.7%	5.51 μg/L
Rostom et al. [34].	Mn Co Cr Ni Cu Cd Pb Fe Zn	B531, B904 B989, B997, and B985 B989, B990, and B945, B989 and B1013 B704 B1001, B952, and B887 B989, and B990 B366 B1001, B977, & B996	$Mn = 0.021 + 1.157R_{531} - 0.966R_{904}$ $Co = 0.134 - 61.536R_{989} + 115.889R_{997} - 69.808R_{985}$ $Cr = 0.086 - 57.461R_{989} + 62.272R_{997} - 3.492R_{945}$ $Ni = 0.577 - 47.625R_{989} + 14.589R_{1013}$ $Cu = -0.037 + 0.303R_{704}$ $Cd = -0.013 + 10.096R_{1001} - 5.888R_{952} + 0.219R_{887}$ $Pb = 1.354 - 243.647R_{989} + 207.788R_{990}$ $Fe = 0.018 + 4.376R_{366}$ $Zn = -0.587 + 125.347R_{1001} + 392.603R_{977} - 299.516R_{996}$	Mn: 0.097–0.245 mg/L Co: 0.069–0.199 mg/L Cr: 0.006–0.120 mg/L Ni: 0.395–0.516 mg/L Cu: 0.006–0.094 mg/L Cd: 0.003–0.068 mg/L Pb: 0.980–1.527 mg/L Fe: 0.034–1.261 mg/L Zn: 0.021–1.160 mg/L	22	Mn: 0.70 Co: 0.58 Cr: 0.86 Ni: 0.82 Cu: 0.97 Cd: 0.27 Pb: 0.88 Fe: 0.87 Zn: 0.27		
Huang et al. [35].	Cu ion	B901.11, B1079.08	$Cu_{ion} = 7829.05 \times [\log(\frac{1}{R_{901.11}}) - 8085.861 \times [\log(\frac{1}{R_{1079.08}})] + 208.703$	100–1000 mg/L,	98	0.74 (adjusted)		
Swain and Sahoo [39].	Fe, Zn, Cu, Cr, Pb, Cd	L_3: 631–692 nm, L_4: 772–898 nm, L_3/L_4	$\log T_u = -30{L_3}^{1.02} + 29.99{L_4}^{0.62} - 4.67\left(\frac{L_3}{L_4}\right)^{-0.72} + 2.24$ $TSS = \begin{cases} 2.73{T_u}^{0.98} + 4.63, & T_u \leq 50 \\ 4.37{T_u}^{0.83}, & 50 < T_u < 200 \\ 0.22{T_u}^{1.12} + 135.84, & T_u \geq 200 \end{cases}$ $Fe = 20.27 \times (TSS)^{1.008}$ $Zn = 3.98 \times (TSS)^{0.43} + 0.16$ $Cu = 1.084 \times \exp[0.013(TSS)]Cr = 0.0005 \times (TSS)^{2.20} + 3.69$ $Pb = 2.87 \times \exp[0.01(TSS)]$ $Cd = 0.306 \times \exp[0.017(TSS)]$	Fe: 95.01–7538.03 μg/L Zn: 7.85–49.86 μg/L Cu: 1.15–109.21 μg/L Cr: 3.70–207.38 μg/L Pb: 3.01–207.38 μg/L Cd: 0.33–127.45 μg/L	250	From L_n to T_u: 0.90 From T_u to TSS: 0.86 Form TSS to metal: 0.76–0.91 for all, among which, Cd: 0.91 Fe: 0.76		From L_n to T_u: 27.28 NTU From T_u to TSS: 28.27 mg/L Form TSS to Metal: 3.76–38.72 μg/L for all metals
Zhou et al. [63].	K^+, Mg^{2+}, Na^+	HJ–1A bands: 105, 104, 106, 103, 107, 46, 45, 109, 102, 47, 48, 108, 49, 101, 44, 50, 100, 70, 51, 99	BP Neural Network	K: 2.9–3.5 g/L Mg: 6.4–7.4 g/L Na: 110–116 g/L	21	>0.85	<10%	
Wang et al. [65].	Li	Sentinel-2A Bands 4 (665 nm), 6 (740 nm), 8 (842 nm), 8a (865 nm)	BP Neural Network Random Forrest (RF)	6.8–96 mg/L	30	BP: 0.731 RF: 0.771		(mg/L) 15.719 12.822
Liu et al. [67].	Li	Landsat Bands 1–7	LightGBM	424.7 -568.1 mg/L	27	0.876	5.3%	10.287 mg/L
Lin et al. [72].	Cu, Fe	Cu: 497, 665, 686, 831, and 935 nm Fe: 700, 746, 801, 948, and 993 nm	Genetic Algorithm–Partial Least Squares Regression (GA-PLSR) PLSR RF	Cu: 0.002–0.025 mg/L Fe: 0.014–0.262 mg/L	35	Cu: RF 0.53 PLSR 0.57 GA-PLSR 0.75 Fe: RF 0.47 PLSR 0.57 GA-PLSR 0.73	Cu: RF 42.4% PLSR 48.6% GA-PLSR 38.2% Fe: RF 56.3% PLSR 46.6% GA-PLSR 46.4%	Cu: RF 0.004 mg/L PLSR 0.004 mg/L GA-PLSR 0.004 mg/L Fe: RF 0.045 mg/L PLSR 0.037 mg/L GA-PLSR 0.036 mg/L

5. Discussion

Heavy metal concentrations serve as a crucial indicator of water quality. While significant advancements have been made in optical detection of water quality parameters such as chlorophyll-a (Chl-a), colored dissolved organic matter (CDOM), and total suspended matter (TSM), quantifying heavy metals through optical remote sensing remains challenging due to the limited research in this area. Most heavy metals are not optically active on their own; however, some of them may associate with optically active substances like TSM, Chl-a, and CDOM. These associations create a correlation between the concentrations of these water quality parameters and the presence of heavy metals. Therefore, by accurately determining the concentrations of TSM, Chl-a, or CDOM, it is possible to indirectly estimate heavy metal concentrations in water. Existing studies primarily examine the correlation between TSM and heavy metals, with most algorithms for remote sensing of heavy metals employing the red and near-infrared (NIR) bands. These wavelengths are ideal for retrieving TSM concentrations. For instance, Chen et al. [29] utilized remote sensing reflectance at 660 nm and 560 nm for the detection of copper (Cu), lead (Pb), and zinc (Zn) concentrations, while [30] used NIR bands at 754, 779, and 865 nm to analyze TSM particle size distribution. More recent studies by Huang et al. [35], Rostom et al. [34], and Swain and Sahoo [39] have employed longer NIR wavelengths to detect a wider range of heavy metals in aquatic environments, suggesting that increased spectral resolution in the NIR could enhance the detection of heavy metals in water.

Some heavy metal solutions exhibit distinct spectral characteristics, such as Cu ions in water, which typically give a green color. To accurately determine the concentrations of these metals, it is essential to identify their spectral patterns in aquatic environments, especially for common pollutant metals. Laboratory measurements can be conducted to control concentrations and establish empirical relationships between concentration levels and spectral data. However, many heavy metals do not exhibit clear spectral properties or correlations with total suspended matter (TSM) or other optically active water components, making it challenging to estimate their concentrations. New mechanisms are needed to explain the correlation between spectral data and heavy metal concentrations, which would enable indirect methods of retrieval. Beyond spectral data, multi-source information, such as hydrodynamic parameters, may also be useful for detecting non-optically active water components [74], particularly if a connection between heavy metal concentrations and hydrodynamic properties exists. It is important to recognize that detecting heavy metals in water through remote sensing is a relatively new area of study. Current research is still in its early stages, with many limitations, but there is considerable potential for further development.

Hyperspectral remote sensing offers an extensive array of bands across both the NIR and visible light spectra, significantly enhancing the fidelity of remote sensing studies. These studies typically leverage in situ spectral measurements to generate data that accurately reflects satellite hyperspectral sensor readings. Hyperspectral satellite missions, initiated in the 1980s, have progressively advanced our observational capabilities [75,76]. For instance, NASA's HyperSpectral Imager (HSI) in the LEWIS mission featured 128 bands spanning the 0.4–1.0 µm range in visible light and NIR, and 256 bands in the SWIR range of 0.9–2.5 µm, with spectral resolutions of 5.0 nm and 6.5 nm, respectively [77]. Although the HSI mission concluded in 2008, it yielded a substantial dataset for heavy metal detection in aquatic surfaces. Subsequently, in 2000, NASA launched the Hyperion sensor aboard the Earth Observing-1 (EO-1) satellite, which included 220 spectral bands across 0.4–2.5 µm with a spectral resolution of 10 nm and spatial resolution of 30 m [78]. More recently, in 2022, Germany deployed the Environmental Mapping and Analysis Program (EnMAP) hyperspectral satellite, which mirrors Hyperion's spectral range but with enhanced spectral resolutions of 6.5 nm in the VNIR and 10 nm in the SWIR, facilitating ongoing stable earth observations crucial for monitoring aqueous heavy metal concentrations [79]. Another significant hyperspectral mission is NASA's Plankton, Aerosol, Cloud, ocean Ecosystem (PACE), launched in February 2024, which includes the Ocean Color Instrument (OCI) that spans from 315 nm in the ultraviolet to 895 nm in the NIR, with a spectral resolution of 2.5 nm [80].

China has also been active in hyperspectral imaging, launching missions like Zhuhai-1 and GaoFen-5, which feature high-resolution hyperspectral sensors. The Orbita hyperspectral satellites (OHS) of the Zhuhai-1 mission provide about 256 bands in the 0.4–1.0 µm range with a spectral resolution of 2.5 nm, while the Advanced Hyperspectral Imager (AHSI) aboard GaoFen-5 offers 330 bands across 0.39 to 2.51 µm with spectral resolutions of 5 nm in the VNIR and 10 nm in the SWIR ranges [81,82].

As technological advancements continue, we anticipate significant enhancements in hyperspectral missions, characterized by increasingly sophisticated sensors that offer finer spectral and spatial resolutions. With the proliferation of bands in reflectance data, specialized approaches are necessary to process or select pertinent band data for developing algorithms targeting heavy metal detection. In this context, machine learning and deep learning technologies are pivotal in integrating as much reflectance data as possible observed in hyperspectral remote sensing. Prominent machine and deep learning methodologies include random forest (RF) [83], Gradient Boosting Decision Tree (XGBoost) [47], Support Vector Machine (SVM) [70], Deep Neural Network (DNN) [84], and Mixture Density Network (MDN) [85]. These techniques are systematically classified into ensemble tree methods, kernel-based approaches, and neural network strategies. Diverging from traditional empirical methods, machine learning foregoes a static model framework; rather, it continuously refines models through a process of "learning" from errors, thereby deciphering the complex interactions between independent variables (spectral reflectance) and dependent variables (heavy metal concentrations) [86]. This iterative learning process is poised to enhance the accuracy of heavy metal retrieval, skillfully navigating the varied optical properties of aquatic environments. To bolster the application of machine and deep learning in water heavy metal detection, an expanded dataset is essential, necessitating increased water sample collection for robust model development. With the influx of satellite hyperspectral data coupled with advanced machine learning techniques, the detection of heavy metals at acceptable accuracies is increasingly feasible.

6. Conclusions

Heavy metal pollution represents a growing threat to both ecological systems and human health due to the biotoxicity and persistence of these pollutants across various environmental media. In aquatic environments, heavy metals exhibit complex behaviors, shifting between particulate, dissolved, and biological forms as they interact with suspended sediments, phytoplankton, and dissolved organic matter. Recent advancements in water color remote sensing have facilitated the detection of heavy metal concentrations in aquatic environments. Notably, the reflectance and absorption characteristics of metals like Cu, Fe, and Cd are primarily within the red and near-infrared (NIR) spectra, ranging from 600 to 810 nm, suggesting that these spectral ranges can be effectively utilized for detecting these heavy metals.

Empirical algorithms have been developed to estimate concentrations of metals such as Cu, Pb, and Zn by combining in situ spectral data from red bands with heavy metal measurements from water samples. Additionally, research has highlighted the relationship between heavy metals and total suspended matter in water, which has led to the development of algorithms that enhance the retrieval of total suspended matter and improve heavy metal detection capabilities.

Hyperspectral remote sensing, which offers a broader range of spectral bands, has further enhanced the precision of heavy metal detection in aquatic environments by capturing the complex and subtle spectral characteristics of these pollutants. The integration of artificial intelligence (AI) with hyperspectral remote sensing holds promise for significantly improving heavy metal retrieval practices. Studies have demonstrated that machine learning, a data-driven approach, can enhance retrieval accuracy. However, the effectiveness of AI in this context is currently limited by the availability of extensive datasets on heavy metal concentrations and their associated spectra. As more data become available, AI technologies are expected to play a crucial role in accurately estimating heavy metal concentrations in aquatic environments, paving the way for more robust monitoring and management strategies.

Author Contributions: Conceptualization, X.X. and J.P.; methodology, X.X. and J.P.; investigation, X.X. and J.P.; resources, X.X. and J.P.; data curation, X.X., J.P., H.Z. and H.L.; writing—original draft preparation, X.X., J.P., H.Z. and H.L.; writing—review and editing, J.P.; visualization, X.X., J.P. and H.Z.; supervision, J.P.; project administration, J.P. and H.L.; funding acquisition, J.P. and H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by National Key R&D Program of China, grant number 2021YFB3900400, and the National Natural Science Foundation of China, grant number: 42330108.

Data Availability Statement: Data are contained within this review article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gong, J.; Wang, Y.; Li, Q.; Liu, X.; Cao, Y.; He, L. Distribution characteristics, pollution assessment and source analysis of heavy metals in a drinking water source area. *Environ. Chem.* **2022**, *41*, 2276–2288.
2. Wang, H.; Gou, W.; Wu, Y.; Li, W. Progress in remediation technologies of heavy metals contaminated soil: Principles and technologies. *Chin. J. Ecol.* **2021**, *40*, 2277–2288. [CrossRef]
3. Dai, B.; Lv, J.; Zhan, J.; Zhang, Z.; Liu, Y.; Zhou, R. Assessment of sources, spatial distribution and ecological risk of heavy metals in soils in a typical industry-based city of Shandong Province, Eastern China. *Environ. Sci.* **2015**, *36*, 507–515. [CrossRef]
4. Chen, C.; Liu, F.; Tang, S. Estimation of Heavy Metal Concentration in the Pearl River Estuarine Waters from Remote Sensing Data. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 2575–2578. [CrossRef]
5. Zheng, J.; Wang, Z.; Lin, Z.; Li, Z.; Zhu, Z.; Chen, J. A study of estuarine chemistry in the Zhujiang River, I. Trace metal species in water phase. *Oceanol. Limnol. Sin.* **1982**, *13*, 19–25.
6. Lin, Z.; Zheng, J.; Wang, Z.; Chen, J. A study of estuarine chemistry in the Zhujiang River II. Chemical forms of heavy metals in the suspended particulate. *Oceanol. Limnol. Sin.* **1982**, *13*, 523–530.
7. Tong, Y.; Yue, T.; Gao, J.; Wang, K.; Wang, C.; Zuo, P.; Zhang, X.; Tong, L.; Liang, Q. Partitioning and Emission Characteristics of Hg, Cr, Pb, and As Among Air Pollution Control Devices in Chinese Coal-Fired Industrial Boilers. *Energy Fuels* **2020**, *34*, 7067–7075. [CrossRef]
8. Greenfield, R.; van Vuren, J.; Wepener, V. Heavy Metal Concentrations in the Water of the Nyl River System, South Africa. *Afr. J. Aquat. Sci.* **2012**, *37*, 219–224. [CrossRef]
9. Taghinia Hejabi, A.; Basavarajappa, H.T.; Karbassi, A.R.; Monavari, S.M. Heavy Metal Pollution in Water and Sediments in the Kabini River, Karnataka, India. *Environ. Monit. Assess.* **2011**, *182*, 1–13. [CrossRef]
10. Suzuki, K.; Yabuki, T.; Ono, Y. Roadside Rhododendron Pulchrum Leaves as Bioindicators of Heavy Metal Pollution in Traffic Areas of Okayama, Japan. *Environ. Monit. Assess.* **2009**, *149*, 133–141. [CrossRef]
11. Zhang, H.; Wang, Z.; Zhang, Y.; Ding, M.; Li, L. Identification of Traffic-Related Metals and the Effects of Different Environments on Their Enrichment in Roadside Soils along the Qinghai–Tibet Highway. *Sci. Total Environ.* **2015**, *521–522*, 160–172. [CrossRef]
12. Bhuiyan, M.; Suruvi, N.; Dampare, S.; Islam, M.; Quraishi, S.; Ganyaglo, S.; Suzuki, S. Investigation of the Possible Sources of Heavy Metal Contamination in Lagoon and Canal Water in the Tannery Industrial Area in Dhaka, Bangladesh. *Environ. Monit. Assess.* **2011**, *175*, 633–649. [CrossRef] [PubMed]
13. Rickard, D. Kinetics of FeS Precipitation: Part 1. Competing Reaction Mechanisms. *Geochim. Cosmochim. Acta* **1995**, *59*, 4367–4379. [CrossRef]
14. Caetano, M.; Madureira, M.-J.; Vale, C. Metal Remobilisation during Resuspension of Anoxic Contaminated Sediment: Short-Term Laboratory Study. *Water Air Soil Pollut.* **2003**, *143*, 23–40. [CrossRef]
15. Zhuang, Y.; Allen, H.E.; Fu, G. Effect of Aeration of Sediment on Cadmium Binding. *Environ. Toxicol. Chem.* **1994**, *13*, 717–724. [CrossRef]
16. Hong, Y.S.; Kinney, K.A.; Reible, D.D. Effects of Cyclic Changes in pH and Salinity on Metals Release from Sediments. *Environ. Toxicol. Chem.* **2011**, *30*, 1775–1784. [CrossRef]
17. Stewart, R.H. *Methods of Satellite Oceanography*; University of California Press: Berkeley, CA, USA; London, UK, 1985; p. vii + 360.
18. Vosburgh, W.C.; Cooper, G.R. Complex Ions. I. The Identification of Complex Ions in Solution by Spectrophotometric Measurements. *J. Am. Chem. Soc.* **1941**, *63*, 437–442. [CrossRef]
19. Bjerrum, J.; Ballhausen, C.J.; Jørgensen, C.K.; Sörensen, N.A. Studies on Absorption Spectra. I. Results of Calculations on the Spectra and Configuration of Copper(II) Ions. *Acta Chem. Scand.* **1954**, *8*, 1275–1289. [CrossRef]

20. Bjerrum, J.; Agarwala, B.V.; Ljungström, E.; Tørneng, E.; Woldbæk, T.; Strand, T.G.; Sukhoverkhov, V.F. Metal Ammine Formation in Solution. XIX. On the Formation of Tetraamminedi-Mu-Hydroxodicopper(II) and Hydroxotetraammine Complexes in Ammoniacal Copper(II) Solutions. *Acta Chem. Scand.* **1980**, *34a*, 475–481. [CrossRef]
21. Jancsó, G. Effect of D and 18O Isotope Substitution on the Absorption Spectra of Aqueous Copper Sulfate Solutions. *Radiat. Phys. Chem.* **2005**, *74*, 168–171. [CrossRef]
22. Liang, Y.; Deng, R.; Gao, Y.; Qin, Y.; Liu, X. Measuring absorption coefficient spectrum (400–900 nm) of copper ions in water. *J. Remote Sens.* **2016**, *20*, 27–34. [CrossRef]
23. Deng, R.; Liang, Y.; Gao, Y.; Qin, Y.; Liu, X. Measuring absorption coefficient spectrum (400–900 nm) of hydrated and complex ferric ion in water. *J. Remote Sens.* **2016**, *20*, 35–44. [CrossRef]
24. Liang, Y.; Deng, R.; Liu, Y.; Lin, L.; Qin, Y.; He, Y. Measuring the spectrum of extinction coefficient and reflectance for cadmium compounds from 400 to 900 nm. *Spectrosc. Spectr. Anal.* **2016**, *36*, 4006–4012.
25. Luo, Z.; Tang, Y.; Zou, B.; Feng, H.; Han, T.; Zhou, S. Inversion of indoor hyperspectral characteristics of heavy metal Pb and Cd pollution in water. *Geomat. Spat. Inf. Technol.* **2024**, *47*, 104–108.
26. Liang, Y.; Deng, R.; Huang, J.; Xiong, L.; Qin, Y.; Liu, Z. The special characteristic analysis of typical heavy metal polluted water-a cases study of mine drainage in Dabaoshan Mountain, Guangdong province, China. *Spectrosc. Spectr. Anal.* **2019**, *39*, 3237–3244.
27. Bergamaschi, B.A.; Fleck, J.A.; Downing, B.D.; Boss, E.; Pellerin, B.; Ganju, N.K.; Schoellhamer, D.H.; Byington, A.A.; Heim, W.A.; Stephenson, M.; et al. Methyl Mercury Dynamics in a Tidal Wetland Quantified Using in Situ Optical Measurements. *Limnol. Oceanogr.* **2011**, *56*, 1355–1371. [CrossRef]
28. Bergamaschi, B.A.; Fleck, J.A.; Downing, B.D.; Boss, E.; Pellerin, B.A.; Ganju, N.K.; Schoellhamer, D.H.; Byington, A.A.; Heim, W.A.; Stephenson, M.; et al. Mercury Dynamics in a San Francisco Estuary Tidal Wetland: Assessing Dynamics Using In Situ Measurements. *Estuaries Coasts* **2012**, *35*, 1036–1048. [CrossRef]
29. Chen, C.; Liu, F.; He, Q.; Shi, H. The Possibility on Estimation of Concentration of Heavy Metals in Coastal Waters from Remote Sensing Data. In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI, USA, 25–30 July 2010; pp. 4216–4219. [CrossRef]
30. Liu, F.; Tang, S.; Chen, C. Estimation of Particulate Zinc Using MERIS Data of the Pearl River Estuary. *Remote Sens. Lett.* **2013**, *4*, 813–821. [CrossRef]
31. Ritchie, J.C.; Zimba, P.V.; Everitt, J.H. Remote Sensing Techniques to Assess Water Quality. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 695–704. [CrossRef]
32. Niu, C.; Tan, K.; Jia, X.; Wang, X. Deep Learning Based Regression for Optically Inactive Inland Water Quality Parameter Estimation Using Airborne Hyperspectral Imagery. *Environ. Pollut.* **2021**, *286*, 117534. [CrossRef]
33. Pyo, J.; Kwon, Y.S.; Min, J.-H.; Nam, G.; Song, Y.-S.; Ahn, J.M.; Park, S.; Lee, J.; Cho, K.H.; Park, Y. Effect of Hyperspectral Image-Based Initial Conditions on Improving Short-Term Algal Simulation of Hydrodynamic and Water Quality Models. *J. Environ. Manag.* **2021**, *294*, 112988. [CrossRef]
34. Rostom, N.G.; Shalaby, A.A.; Issa, Y.M.; Afifi, A.A. Evaluation of Mariut Lake Water Quality Using Hyperspectral Remote Sensing and Laboratory Works. *Egypt. J. Remote Sens. Space Sci.* **2017**, *20*, S39–S48. [CrossRef]
35. Huang, C.; Chen, X.-Y.; Lee, M. An Improved Hyperspectral Sensing Approach for the Rapid Determination of Copper Ion Concentrations in Water Environment Using Short-Wavelength Infrared Spectroscopy. *Environ. Pollut.* **2023**, *333*, 121984. [CrossRef] [PubMed]
36. Hung, S.-C.; Lu, C.-C.; Wu, Y.-T. An Investigation on Design and Characterization of a Highly Selective LED Optical Sensor for Copper Ions in Aqueous Solutions. *Sensors* **2021**, *21*, 1099. [CrossRef] [PubMed]
37. Yin, F.; Wu, M.; Liu, L.; Zhu, Y.; Feng, J.; Yin, D.; Yin, C.; Yin, C. Predicting the Abundance of Copper in Soil Using Reflectance Spectroscopy and GF5 Hyperspectral Imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102420. [CrossRef]
38. Jiang, G.; Zhou, S.; Cui, S.; Chen, T.; Wang, J.; Chen, X.; Liao, S.; Zhou, K. Exploring the Potential of HySpex Hyperspectral Imagery for Extraction of Copper Content. *Sensors* **2020**, *20*, 6325. [CrossRef]
39. Swain, R.; Sahoo, B. A Simplified Modelling Framework for Real-Time Assessment of Conservative Pollutants in Ungauged Rivers during Cloudy Periods. *J. Environ. Manag.* **2021**, *293*, 112821. [CrossRef]
40. Yuan, Y.; Hall, K.; Oldham, C. A Preliminary Model for Predicting Heavy Metal Contaminant Loading from an Urban Catchment. *Sci. Total Environ.* **2001**, *266*, 299–307. [CrossRef]
41. Liebens, J. Heavy Metal Contamination of Sediments in Stormwater Management Systems: The Effect of Land Use, Particle Size, and Age. *Environ. Geol.* **2001**, *41*, 341–351. [CrossRef]
42. Ujevic, I. Trace Metal Accumulation in Different Grain Size Fractions of the Sediments from a Semi-Enclosed Bay Heavily Contaminated by Urban and Industrial Wastewaters. *Water Res.* **2000**, *34*, 3055–3061. [CrossRef]
43. Herngren, L.; Goonetilleke, A.; Ayoko, G.A. Understanding Heavy Metal and Suspended Solids Relationships in Urban Stormwater Using Simulated Rainfall. *J. Environ. Manag.* **2005**, *76*, 149–158. [CrossRef]
44. Hallberg, M.; Renman, G.; Lundbom, T. Seasonal Variations of Ten Metals in Highway Runoff and Their Partition between Dissolved and Particulate Matter. *Water Air Soil Pollut.* **2007**, *181*, 183–191. [CrossRef]
45. Rajesh, A.; Jiji, G.W.; Raj, J.D. Estimating the Pollution Level Based on Heavy Metal Concentration in Water Bodies of Tiruppur District. *J. Indian Soc. Remote Sens.* **2020**, *48*, 47–57. [CrossRef]

46. Cui, W.; Ma, X. Remote sensing monitoring method for heavy metal polluted wastewater based on spectral analysis. *Adm. Tech. Environ. Monit.* **2022**, *34*, 53–56. [CrossRef]
47. Cao, Z.; Ma, R.; Duan, H.; Pahlevan, N.; Melack, J.; Shen, M.; Xue, K. A Machine Learning Approach to Estimate Chlorophyll-*a* from Landsat-8 Measurements in Inland Lakes. *Remote Sens. Environ.* **2020**, *248*, 111974. [CrossRef]
48. Ye, H.; Tang, S.; Yang, C. Deep Learning for Chlorophyll-a Concentration Retrieval: A Case Study for the Pearl River Estuary. *Remote Sens.* **2021**, *13*, 3717. [CrossRef]
49. Balasubramanian, S.V.; Pahlevan, N.; Smith, B.; Binding, C.; Schalles, J.; Loisel, H.; Gurlin, D.; Greb, S.; Alikas, K.; Randla, M.; et al. Robust Algorithm for Estimating Total Suspended Solids (TSS) in Inland and Nearshore Coastal Waters. *Remote Sens. Environ.* **2020**, *246*, 111768. [CrossRef]
50. Bangira, T.; Matongera, T.N.; Mabhaudhi, T.; Mutanga, O. Remote Sensing-Based Water Quality Monitoring in African Reservoirs, Potential and Limitations of Sensors and Algorithms: A Systematic Review. *Phys. Chem. Earth Parts A/B/C* **2024**, *134*, 103536. [CrossRef]
51. Batina, A.; Krtalić, A. A Review of Remote Sensing Applications for Determining Lake Water Quality. *Preprints* **2023**, 2023090489. [CrossRef]
52. Jakovljevic, G.; Álvarez-Taboada, F.; Govedarica, M. Long-Term Monitoring of Inland Water Quality Parameters Using Landsat Time-Series and Back-Propagated ANN: Assessment and Usability in a Real-Case Scenario. *Remote Sens.* **2024**, *16*, 68. [CrossRef]
53. Shen, H.; Jiang, M.; Li, J.; Zhou, C.; Yuan, Q.; Zhang, L. Coupling Model- and Data-Driven Methods for Remote Sensing Image Restoration and Fusion: Improving Physical Interpretability. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 231–249. [CrossRef]
54. Malakar, N.K.; Lary, D.J.; Moore, A.; Gencaga, D.; Roscoe, B.; Albayrak, A.; Wei, J. Estimation and Bias Correction of Aerosol Abundance Using Data-Driven Machine Learning and Remote Sensing. In Proceedings of the 2012 Conference on Intelligent Data Understanding, Boulder, CO, USA, 24–26 October 2012; pp. 24–30. [CrossRef]
55. Geng, L.; Che, T.; Ma, M.; Tan, J.; Wang, H. Corn Biomass Estimation by Integrating Remote Sensing and Long-Term Observation Data Based on Machine Learning Techniques. *Remote Sens.* **2021**, *13*, 2352. [CrossRef]
56. Nagaraj, R.; Kumar, L.S. Multi Scale Feature Extraction Network with Machine Learning Algorithms for Water Body Extraction from Remote Sensing Images. *Int. J. Remote Sens.* **2022**, *43*, 6349–6387. [CrossRef]
57. Li, A.; Fan, M.; Qin, G.; Xu, Y.; Wang, H. Comparative Analysis of Machine Learning Algorithms in Automatic Identification and Extraction of Water Boundaries. *Appl. Sci.* **2021**, *11*, 10062. [CrossRef]
58. Huang, X.; Xie, C.; Fang, X.; Zhang, L. Combining Pixel- and Object-Based Machine Learning for Identification of Water-Body Types From Urban High-Resolution Remote-Sensing Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2097–2110. [CrossRef]
59. Peterson, K.T.; Sagan, V.; Sidike, P.; Hasenmueller, E.A.; Sloan, J.J.; Knouft, J.H. Machine Learning-Based Ensemble Prediction of Water-Quality Variables Using Feature-Level and Decision-Level Fusion with Proximal Remote Sensing. *Photogramm. Eng. Remote Sens.* **2019**, *85*, 269–280. [CrossRef]
60. Jian, L.; Gao, F.; Ren, P.; Song, Y.; Luo, S. A Noise-Resilient Online Learning Algorithm for Scene Classification. *Remote Sens.* **2018**, *10*, 1836. [CrossRef]
61. Yuan, Q.; Shen, H.; Li, T.; Li, Z.; Li, S.; Jiang, Y.; Xu, H.; Tan, W.; Yang, Q.; Wang, J.; et al. Deep Learning in Environmental Remote Sensing: Achievements and Challenges. *Remote Sens. Environ.* **2020**, *241*, 111716. [CrossRef]
62. Li, W.; Fan, X.; Huang, F.; Chen, W.; Hong, H.; Huang, J.; Guo, Z. Uncertainties Analysis of Collapse Susceptibility Prediction Based on Remote Sensing and GIS: Influences of Different Data-Based Models and Connections between Collapses and Environmental Factors. *Remote Sens.* **2020**, *12*, 4134. [CrossRef]
63. Zhou, Y.; Zhang, R.; Ma, H.; Zhang, J.; Zhang, X. Retrieving salt lake mineral ions salinity from hyper-spectral data based on BP neural network. *Remote Sens. Land Resour.* **2016**, *28*, 34–40. [CrossRef]
64. Li, J.; Cheng, J.; Shi, J.; Huang, F. Brief Introduction of Back Propagation (BP) Neural Network Algorithm and Its Improvement. In *Advances in Computer Science and Information Engineering*; Jin, D., Lin, S., Eds.; Advances in Intelligent and Soft Computing; Springer: Berlin/Heidelberg, Germany, 2012; Volume 169, pp. 553–558. ISBN 978-3-642-30222-0.
65. Wang, D. Models for Predicting Lithium Content in Salt Lake Based on Remote Sensing: A Case Study of Argentina's Arizaro Salt Lake. Master's Thesis, Jilin University, Changchun, China, 2019.
66. Belgiu, M.; Drăguţ, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
67. Liu, T.; Dai, J.; Zhao, Y.; Tian, S.; Ye, C. Remote sensing inversion of lithium concentration in salt lake using LightGBM: A case study of northern Zabuye Salt Lake in Tibet. *Acta Geol. Sin.* **2021**, *95*, 2249–2256. [CrossRef]
68. Fan, J.; Ma, X.; Wu, L.; Zhang, F.; Yu, X.; Zeng, W. Light Gradient Boosting Machine: An Efficient Soft Computing Model for Estimating Daily Reference Evapotranspiration with Local and External Meteorological Data. *Agric. Water Manag.* **2019**, *225*, 105758. [CrossRef]
69. Zhang, Y.; Wang, Y.; Gao, M.; Ma, Q.; Zhao, J.; Zhang, R.; Wang, Q.; Huang, L. A Predictive Data Feature Exploration-Based Air Quality Prediction Approach. *IEEE Access* **2019**, *7*, 30732–30743. [CrossRef]
70. Li, L.; Lin, Y.; Yu, D.; Liu, Z.; Gao, Y.; Qiao, J. A Multi-Organ Fusion and LightGBM Based Radiomics Algorithm for High-Risk Esophageal Varices Prediction in Cirrhotic Patients. *IEEE Access* **2021**, *9*, 15041–15052. [CrossRef]

71. Song, J.; Liu, G.; Jiang, J.; Zhang, P.; Liang, Y. Prediction of Protein–ATP Binding Residues Based on Ensemble of Deep Convolutional Neural Networks and LightGBM Algorithm. *Int. J. Mol. Sci.* **2021**, *22*, 939. [CrossRef]
72. Lin, Y.; Gao, J.; Tu, Y.; Zhang, Y.; Gao, J. Estimating Low Concentration Heavy Metals in Water through Hyperspectral Analysis and Genetic Algorithm-Partial Least Squares Regression. *Sci. Total Environ.* **2024**, *916*, 170225. [CrossRef]
73. Kaneko, H. Genetic Algorithm-Based Partial Least-Squares with Only the First Component for Model Interpretation. *ACS Omega* **2022**, *7*, 8968–8979. [CrossRef]
74. Xia, X.; Pan, J.; Pei, J. A New Approach to Estimate Total Nitrogen Concentration in a Seasonal Lake Based on Multi-Source Data Methodology. *Ecol. Inform.* **2024**, *83*, 102807. [CrossRef]
75. Goetz, A.F.H. Three Decades of Hyperspectral Remote Sensing of the Earth: A Personal View. *Remote Sens. Environ.* **2009**, *113*, S5–S16. [CrossRef]
76. Dierssen, H.; Ackleson, S.; Joyce, K.; Hestir, E.; Castagna, A.; Lavender, S.; McManus, M. Living up to the Hype of Hyperspectral Aquatic Remote Sensing: Science, Resources and Outlook. *Front. Environ. Sci.* **2021**, *9*, 649528. [CrossRef]
77. Qian, S.-E. Hyperspectral Satellites, Evolution, and Development History. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7032–7056. [CrossRef]
78. Ungar, S.G.; Pearlman, J.S.; Mendenhall, J.A.; Reuter, D. Overview of the Earth Observing One (Eo-1) Mission. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1149–1159. [CrossRef]
79. Mission-EnMAP. Available online: https://www.enmap.org/mission/ (accessed on 16 October 2024).
80. NASA PACE-Home. Available online: https://pace.gsfc.nasa.gov/ (accessed on 16 October 2024).
81. Jiang, Y.; Wang, J.; Zhang, L.; Zhang, G.; Li, X.; Wu, J. Geometric Processing and Accuracy Verification of Zhuhai-1 Hyperspectral Satellites. *Remote Sens.* **2019**, *11*, 996. [CrossRef]
82. Liu, Y.-N.; Sun, D.-X.; Hu, X.-N.; Ye, X.; Li, Y.-D.; Liu, S.-F.; Cao, K.-Q.; Chai, M.-Y.; Zhou, W.-Y.-N.; Zhang, J.; et al. The Advanced Hyperspectral Imager: Aboard China's GaoFen-5 Satellite. *IEEE Geosci. Remote Sens. Mag.* **2019**, *7*, 23–32. [CrossRef]
83. DeLuca, N.M.; Zaitchik, B.F.; Curriero, F.C. Can Multispectral Information Improve Remotely Sensed Estimates of Total Suspended Solids? A Statistical Study in Chesapeake Bay. *Remote Sens.* **2018**, *10*, 1393. [CrossRef]
84. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat. Deep Learning and Process Understanding for Data-Driven Earth System Science. *Nature* **2019**, *566*, 195–204. [CrossRef]
85. Pahlevan, N.; Smith, B.; Schalles, J.; Binding, C.; Cao, Z.; Ma, R.; Alikas, K.; Kangro, K.; Gurlin, D.; Hà, N.; et al. Seamless Retrievals of Chlorophyll-*a* from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in Inland and Coastal Waters: A Machine-Learning Approach. *Remote Sens. Environ.* **2020**, *240*, 111604. [CrossRef]
86. Lary, D.J.; Alavi, A.H.; Gandomi, A.H.; Walker, A.L. Machine Learning in Geosciences and Remote Sensing. *Geosci. Front.* **2016**, *7*, 3–10. [CrossRef]