*Article*

# Characterizing Water Composition with an Autonomous Robotic Team Employing Comprehensive In Situ Sensing, Hyperspectral Imaging, Machine Learning, and Conformal Prediction

John Waczak, Adam Aker, Lakitha O. H. Wijeratne, Shawhin Talebi, Ashen Fernando, Prabuddha M. H. Dewage, Mazhar Iqbal, Matthew Lary, David Schaefer and David J. Lary *

Hanson Center for Space Sciences, University of Texas at Dallas, Richardson, TX 75080, USA;
john.waczak@utdallas.edu (J.W.); adam.aker@utdallas.edu (A.A.); lhw150030@utdallas.edu (L.O.H.W.);
shawhintalebi@gmail.com (S.T.); ashen.fernando@utdallas.edu (A.F.); pxh180012@utdallas.edu (P.M.H.D.);
mazhar.iqbal@utdallas.edu (M.I.); MDL210001@utdallas.edu (M.L.); captdaveschaefer@gmail.com (D.S.)
* Correspondence: david.lary@utdallas.edu

**Abstract:** Inland waters pose a unique challenge for water quality monitoring by remote sensing techniques due to their complicated spectral features and small-scale variability. At the same time, collecting the reference data needed to calibrate remote sensing data products is both time consuming and expensive. In this study, we present the further development of a robotic team composed of an uncrewed surface vessel (USV) providing in situ reference measurements and an unmanned aerial vehicle (UAV) equipped with a hyperspectral imager. Together, this team is able to address the limitations of existing approaches by enabling the simultaneous collection of hyperspectral imagery with precisely collocated in situ data. We showcase the capabilities of this team using data collected in a northern Texas pond across three days in 2020. Machine learning models for 13 variables are trained using the dataset of paired in situ measurements and coincident reflectance spectra. These models successfully estimate physical variables including temperature, conductivity, pH, and turbidity as well as the concentrations of blue–green algae, colored dissolved organic matter (CDOM), chlorophyll-a, crude oil, optical brighteners, and the ions $Ca^{2+}$, $Cl^-$, and $Na^+$. We extend the training procedure to utilize conformal prediction to estimate 90% confidence intervals for the output of each trained model. Maps generated by applying the models to the collected images reveal small-scale spatial variability within the pond. This study highlights the value of combining real-time, in situ measurements together with hyperspectral imaging for the rapid characterization of water composition.

**Keywords:** water quality; robotic teams; hyperspectral imaging; machine learning; conformal prediction

## 1. Introduction

For decades, remote sensing imagery has been used for environmental monitoring, with applications ranging from resource mapping, land type classification, and urban growth assessment to wildfire monitoring, natural disaster tracking, and many more [1,2]. Among these applications, the retrieval of water quality variables from remote sensing imagery remains challenging due to the difficulty of obtaining in situ reference data coincident with available satellite imagery. Traditional approaches to obtain these data have relied on serendipitous satellite passes over fixed sensing sites or sensor-equipped vessels. As a consequence, curating comprehensive datasets can require decades of observations [3,4]. This poses a significant challenge for assessing natural and anthropogenic changes to water composition in real time, for example, during oil spills [5].

Where remote sensing imagery and in situ measurements have been combined, studies have demonstrated successful extraction of optically active water quality variables such

as colored dissolved organic matter (CDOM), chlorophyll-a, and suspended sediment concentrations by using combinations of spectral bands [6–8]. These approaches can be further augmented by machine learning methods, which consist of nonlinear and non-parametric models designed to learn representations of functions directly from data [9]. For example, Petersen et al. utilized a deep neural network to successfully estimate blue–green algae, chlorophyll-a, CDOM, dissolved oxygen, specific conductance, and turbidity from a fused dataset of Landsat-8 and Sentinel-2 imagery [10]. Other methods such as support vector machines and decision trees have also been successfully applied to retrieve water quality parameters from remote sensing imagery [11,12]. However, the low spatial and spectral resolution of available multiband remote sensing satellites makes it difficult to analyze inland waters with small spatial features and complicated spectral signatures.

Advances in multispectral and hyperspectral imaging technology have led to considerable reductions in size, making it possible to incorporate these cameras into the payloads of autonomous aerial vehicles (UAVs) [13]. Flying at low altitudes enables the collection of centimeter-scale imagery whilst limiting the need for complicated atmospheric corrections to account for scattering by atmospheric aerosols and gasses [14,15]. Already, UAVs equipped with multispectral and hyperspectral imagers are being used in a variety of domains to great effect, including for biomass estimation, forest management, precision agriculture, and, recently, water quality monitoring [14,16–19]. Despite the superior spectral and spatial resolution enabled by UAV platforms, these improvements alone do not address the limited spatial coverage of the in situ reference data used for the associated water composition retrieval. For instance, Lu et al. used a UAV-born hyperspectral imager to develop machine learning models for the inversion of chlorophyll-a and suspended solids using samples from 33 fixed locations [20]. Similarly, Zhang et al. utilized a UAV equipped with a hyperspectral imager to estimate water quality parameters by collecting imagery coincident with samples taken from 18 sampling sites [21]. In both of these examples, the collection of in situ reference data remains the key challenge for the application of UAV systems to water quality quantification.

To address this gap and enable comprehensive, real-time evaluation of water composition, we have developed a robotic team comprised of an autonomous uncrewed surface vessel (USV) equipped with a collection of reference grade instruments together with a UAV carrying a hyperspectral imager. By incorporating reference instruments on a maneuverable USV platform, we are able to rapidly collect large volumes of water quality data for a comprehensive suite of physical, biochemical, and chemical variables that are precisely collocated with spectra captured by the UAV. Critically, the USV enables the collection of reference data with significantly improved spatial resolution compared to other approaches. In our previous work, we introduced this paradigm and described how in situ measurements collected by the USV can be used to provide the ground-truth data for machine learning models that map the reflectance spectra captured by the hyperspectral imager to the desired water quality variables [22].

The main objective of this study is to expand on our previous work in three new ways: The first is to explore the breadth of water quality variables that can be inferred from collected hyperspectral imagery. With the goal of comprehensive measurement in mind, we demonstrate the ability to accurately predict physical variables such as temperature, conductivity, pH, and turbidity as well as biochemical constituents, including blue–green algae pigments, CDOM, and chlorophyll-a, in addition to concentrations of crude oil, optical brighteners, and a variety of ions. The second addition is to demonstrate that observations from separate collections can effectively be combined by carefully accounting for variability in the viewing and illumination geometries of the scene. Finally, we expand our machine learning approach to take advantage of the considerable volume of collected data in order to determine reliable confidence intervals for each predicted parameter using conformal prediction.

## 2. Materials and Methods

The robotic team presented in this study consists of two key sensing sentinels: an uncrewed surface vessel (USV) used to collect in situ reference measurements and an unmanned aerial vehicle (UAV) for performing rapid, wide-area surveys to gather remote sensing data products. Both platforms are coordinated using open-source QGroundControl version 4.0.11 software for flight control and mission planning and are equipped with high-accuracy GPS and INS such that all data points collected are uniquely geolocated and time-stamped [23]. Both the USV and UAV include long-range Ubiquiti 5 GHz LiteBeam airMAX WiFi to enable streaming of data products to a ground station with network-attached storage to provide redundancy.

### 2.1. USV: In Situ Measurements

The USV employed in the robot team is a Maritime Robotics Otter equipped with an in situ sensing payload consisting of a combination of Eureka Manta + 40 multiprobes. These sensors include fluorometers, ion-selective electrodes, and other physical sensors and are mounted on the underside of the boat as illustrated in Figure 1. Together, this sensor array enables the collection of comprehensive near-surface measurements including colored dissolved organic matter (CDOM), crude oil, blue–green algae (phycoerythrin and phycocyanin), chlorophyll-a, $Na^+$, $Ca^{2+}$, $Cl^-$, temperature, conductivity, and many others. The full list of measurements utilized in this study is outlined in Table 1 and is categorized into four distinct types: physical measurements, ion measurements, biochemical measurements, and chemical measurements. Additionally, the USV is equipped with an ultrasonic weather monitoring sensor for measuring air speed and direction as well as a BioSonics MX Aquatic Habitat Echosounder sonar, which are not explored in this study.
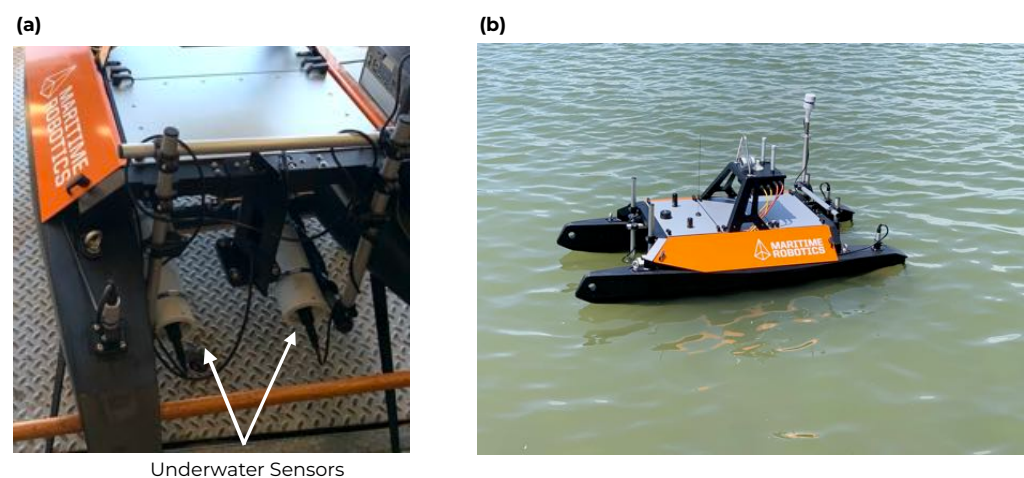


**(a)**  **(b)**

Underwater Sensors

**Figure 1.** Configuration of the USV: (**a**) Frontal view of the USV showing the Eureka Manta + 40 multiprobes mounted on the underside of the boat. (**b**) The USV deployed in the water.

As shown in Table 1, the physical measurements and ion sensors are largely based on different electrode configurations, while the chemical and biochemical measurements are all optically significant in UV and visible light, enabling their determination by fluorometry [24–26]. The pigments phycoerythrin and phycocyanin are used to determine the blue–green algae content, which together with chlorophyll-a enables us to assess the distribution of photosynthetic life in the pond [27,28]. In addition, we also measured the concentration of colored dissolved organic matter (CDOM), which impacts light penetration and serves as a primary source of bioavailable carbon. Crude oil (natural petroleum) and optical brightener concentrations are also measured with fluorometers and are relevant for identifying sources of industrial contamination, natural seepage, and sewage [29,30].

**Table 1.** In situ reference sensors utilized in this study.

| Sensor | Units | Resolution | Sensor Type | Target Category |
|---|---|---|---|---|
| Temperature | °C | 0.01 | Thermistor | Physical |
| Conductivity | µS/cm | 0.01 | Four-Electrode Graphite Sensor | Physical |
| pH | logarithmic (0–14) | 0.01 | Flowing-Junction Reference Electrode | Physical |
| Turbidity | FNU | 0.01 | Ion-Selective Electrode | Physical |
| $Ca^{2+}$ | mg/L | 0.1 | Ion-Selective Electrode | Ions |
| $Cl^-$ | mg/L | 0.1 | Ion-Selective Electrode | Ions |
| $Na^+$ | mg/L | 0.1 | Ion-Selective Electrode | Ions |
| Blue–Green Algae (phycoerythrin) | ppb | 0.01 | Fluorometer | Biochemical |
| Blue–Green Algae (phycocyanin) | ppb | 0.01 | Fluorometer | Biochemical |
| CDOM | ppb | 0.01 | Fluorometer | Biochemical |
| Chlorophyll-a | ppb | 0.01 | Fluorometer | Biochemical |
| Optical Brighteners | ppb | 0.01 | Fluorometer | Chemical |
| Crude Oil | ppb | 0.01 | Fluorometer | Chemical |

The inclusion of optically inactive variables such as conductivity, pH, and ion concentrations was motivated by a desire to be comprehensive. Multiple studies have classified inland water bodies according to their ionic compositions [31,32]. Other research indicates that the structure of dissolved organic matter is affected by changes in pH and cation concentration [33]. Therefore, changes in physical parameters and ionic content can be expected to be related to the observed distribution of optically active variables in the water. It is therefore reasonable to expect that these parameters may be estimated from hyperspectral images at a given site.

### 2.2. UAV: Hyperspectral Data Cubes

A Freefly Alta-X autonomous quadcopter was used as a UAV platform for the robotic team. The Alta-X is specifically designed to carry cameras and has a payload of up to 35 lbs. We equipped the UAV with a Resonon Pika XC2 visible+near-infrared (VNIR) hyperspectral imager. For each image pixel, this camera samples 462 wavelengths ranging from 391 to 1011 nm. Additionally, the UAV includes an upward facing Ocean Optics UV-Vis-NIR spectrometer with a cosine corrector to capture the incident downwelling irradiance spectrum. Data collection by the hyperspectral imager is controlled by an attached Intel NUC small-form-factor computer. A second processing NUC is also included for onboard georectification and generation of data products. The collected hyperspectral images (HSIs) are stored locally on a solid state drive that is simultaneously mounted by the processing computer. The configuration of the drone is shown in Figure 2.
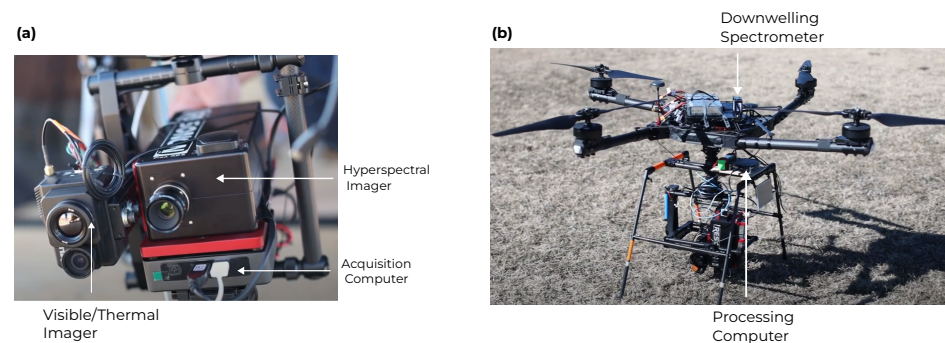


**Figure 2.** Configuration of the UAV: (**a**) The hyperspectral imager and acquisition computer. (**b**) The assembled UAV with secondary processing computer and (upward facing) downwelling irradiance spectrometer.

To effectively utilize the spectra collected by our UAV, we must account for the variability of the incident light that illuminates the water and transform the raw hyperspectral data cubes from their native imaging reference frame to a chosen coordinate system compatible with the data collected by the USV. This procedure is illustrated in Figure 3.
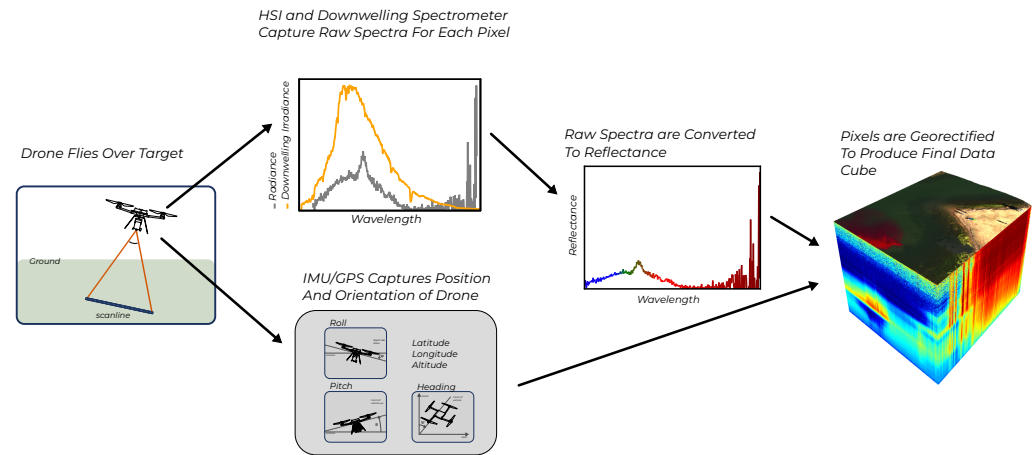


**Figure 3.** Hyperspectral image processing: Hyperspectral data cubes are collected one scan-line at a time (**left**). By utilizing downwelling irradiance spectra, we convert each pixel from spectral radiance to reflectance. By using orientation and position data from the on-board GPS and INS, we georeference each pixel to assign it a latitude and longitude on the ground. The final data product is the georectified hyperspectral reflectance data cube (**right**) visualized as a pseudo-color image with reflectance as a function of wavelength along the z-axis.

The hyperspectral imager utilized in our robot team is in a so-called pushbroom configuration: that is, each image captured by the drone is formed one scan line at a time as the UAV flies. Each scan line consists of 1600 pixels, for which incoming light is diffracted into 462 wavelength bins. In the collection software, a regular cutoff of 1000 lines is chosen so that each resulting image forms an array of size $462 \times 1600 \times 1000$ called a hyperspectral data cube. Initially, the captured spectra are in units of spectral radiance (measured in microflicks); however, this does not account for the variability of incident light illuminating the water. To this end, we convert the hyperspectral data cubes into units of reflectance by utilizing the skyward-facing downwelling irradiance spectrometer. When the camera is normal to the water surface, the reflectance is given by

$$R(\lambda) = \pi L(\lambda) / E_d(\lambda) \tag{1}$$

where $L$ is the spectral radiance, $E_d$ is the downwelling irradiance, and a factor of $\pi$ steradians results from assuming the water surface is Lambertian (diffuse) [34].

Having converted the hyperspectral data cube to units of reflectance, we must also georeference each pixel into a geographic coordinate system so that each image pixel can be assigned a latitude and longitude corresponding to the location on the ground from which it was sampled. During our three surveys, the UAV was flown at an altitude of approximately 50 m above the water. At this scale, the surface is essentially flat, so the hyperspectral data cube can be reliably georectified without the need for an on-board digital elevation map (DEM). We adopt the approach outlined in [35–37] whereby each scan line is georeferenced using the known field of view (30.8°) together with the position and orientation of the UAV as provided by the on-board GPS/INS. After a sequence of coordinate transformations, the pixel coordinates are obtained in the relevant UTM zone (in meters). The resulting image is then re-sampled to a final output resolution. For these collections, a resolution of 10 cm was utilized; however, this can be adjusted to optimize the processing time and final resolution for real-time applications. Finally, the UTM pixel coordinates obtained are transformed to latitude and longitude for easy comparison with

in situ data collected by the USV. The final result is a georectified hyperspectral reflectance data cube. In Figure 4, we visualize one such data cube, highlighting a selection of exemplar pixel spectra as well as the incident downward irradiance spectrum. A pseudo-color image is generated (plotted on the top of the data cube) to illustrate the scene.
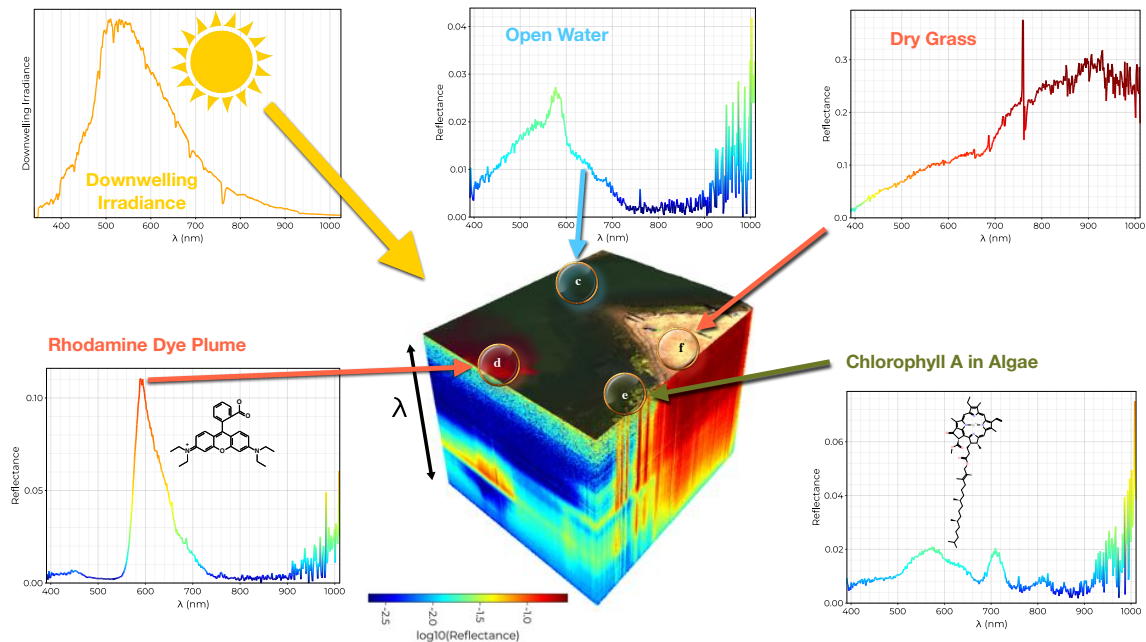


**Figure 4.** A georectified reflectance data cube is visualized (center) with the $\log_{10}$ reflectance along the z-axis and a pseudo-color image on the top. In the top left, we visualize the downwelling irradiance spectrum (the incident light). The surrounding plots showcase exemplar pixel reflectance spectra for open water, dry grass, algae, and a rhodamine dye plume used to test the system.

The above processing workflow was implemented using the Julia programming language: a just-in-time compiled language with native multi-threading support [38]. By running this pipeline on the onboard computer, we are able to process the collected hyperspectral data cubes in near real time. This feature is critical for time-sensitive applications wherein we need to quickly assess if an area is safe and cannot afford to wait to download and post-process collected imagery after a flight.

## 2.3. Data Collection

The robot team was deployed at a private pond in Montague, Texas, close to the Oklahoma border for three separate collections on 23 November 2020, 9 December 2020, and 10 December 2020. The pond spans an area $< 0.1$ km$^2$ and has a maximum depth of 3 m. As shown in Figure 5, the area includes multiple distinct regions with significant small-scale variability. For each acquisition, the UAV first completed a broad survey of the pond, capturing multiple hyperspectral data cubes. Subsequently, the USV sampled across the pond, collecting in situ reference measurements. Each of these reference measurements was then collocated with individual pixel spectra, whereby the USV track overlapped with the UAV's pixels. To account for any time lag in the values measured by the in situ instruments and to account for the USV's size in comparison to the data cube's spatial resolution, each in situ measurement is associated with a $3 \times 3$ grid of HSI pixels: that is, a 30 cm $\times$ 30 cm square. These combined data form the tabular dataset on which we train regression models; pixel spectra form input features, and each separate USV sensor forms a distinct target variable.

**Figure 5.** The pond in Montague, Texas, where the robot team was deployed. The pond includes multiple distinct regions separated by small islands and grasses.

Each collection was performed near solar noon to maximize the amount of incident sunlight illuminating the water. For the site in northern Texas, this corresponded to average solar zenith angles of 54.9°, 56.7°, and 56.75° for 23 November 2020, 9 December 2020, and 10 December 2020, respectively. Given the hyperspectral imager acquires data cubes at nadir, there was little concern for effects due to sunglint. However, to account for any potential variation in lighting conditions between data cubes, we augment the training set with additional features including the drone's viewing geometry (roll, pitch, and heading) and solar illumination geometry (solar azimuth, solar elevation, and solar zenith) as well as the total at-pixel intensity before reflectance conversion, the total downwelling intensity, and the drone's altitude. Further feature engineering is performed to add additional spectral indices that utilize combinations of specific wavelength bands such as the normalized difference vegetation index (NDVI), normalized difference water index (NDWI), simple ratio (SR), photochemical reflectance index (PRI), and more, as outlined in [39–42]. A comprehensive list of these added features is provided in Supplementary Table S1. The final dataset includes a total of 526 features (462 reflectance bands plus 64 additional features) with over 120,000 records.

### 2.4. Machine Learning Methods

For each of the 13 target variables listed in Table 1, the data were randomly partitioned into a 90:10 training/testing split. To model the data, we chose to use the random forest regressor (RFR) as implemented in the Machine Learning framework for Julia (MLJ) [43,44]. Random forests are an ensembling technique based on bagged predictions of individual decision tree regressors trained using the classification and regression trees (CART) algorithm. Each tree in an RFR is trained on a random subset of features and a random subset of training records [45,46]. Random forests are particularly attractive due to their fast training and inference times. Furthermore, studies continue to observe that tree-based models like random forest remain superior for tabular datasets [47,48].

As reflectance values in adjacent wavelength bins tend to correlate with each other and, therefore, may not necessarily contribute additional information content to the final model, it is desirable to evaluate the relative importance of each feature to the trained model predictions. This is useful both for identifying the most relevant features and for performing feature selection to reduce the final model size. By default, tree-based methods such as RFR allow for impurity-based ranking as described in [46,49]. However, these methods have been shown to be biased towards high cardinality and correlated features [50]. Therefore, we choose to use the model-agnostic permutation importance as described by [51]. To do this, we further partition the training dataset, resulting in an 80:10:10 split with 80% of the points used for model training, 10% of the points used for validation and determination of feature importance, and the final 10% held out as an independent testing set. The importance of the *j*th feature is then computed as

$$\text{Imp}_j = R^2(f(X_{\text{val}}), y_{\text{val}}) - R^2(f(X_{\text{val}}^{(j)}), y_{\text{val}}) \tag{2}$$

where $(X_{\text{val}}, y_{\text{val}})$ is the validation set, $f(\cdot)$ is the trained model, $R^2(\cdot, \cdot)$ is the coefficient of determination, and $X_{\text{val}}^{(j)}$ is the validation feature set with the $j$th column randomly permuted. The permutation importance is therefore understood to be the decrease in model performance when the $j$th feature is replaced by random values from the validation set.

To assess the uncertainty of the final model's predictions, we employed inductive conformal prediction as described in [52–55]. To do this, we computed a set of nonconformity scores of the trained model on the validation set using an uncertainty heuristic: in this case, the absolute error:

$$s_i = s(X_i, y_i) = |f(X_i) - y_i| = |\hat{y}_i - y_i| \tag{3}$$

where $f(X_i) = \hat{y}_i$ denotes the trained ML model applied to the $i$th calibration record. These n-many scores are sorted, and the interval half width, $d$, is calculated as the $\frac{[(1-\alpha)(n+1)]}{n}$ quantile of this set in order to achieve coverage of $1 - \alpha$ on the calibration set. Prediction intervals for new data are then formed as $f(X) \pm d$. For this study, we chose $\alpha = 0.1$ for coverage corresponding to a 90% confidence interval.

Using these tools, the training procedure for each model was as follows: First, each model was trained using six-fold cross-validation on the full 526-feature training set with default hyperparameter values. Feature importances for the trained model were then computed, and the top 25 features were identified. A second model was then trained using the same six-fold cross-validation scheme with only these 25 most important features together with hyperparameter optimization using a random search over the number of trees and sampling fraction. The number of trees was optimized, as it has a significant impact on both model performance and inference time. The sampling ratio determines the fraction of training records that each individual tree is exposed to during training. Tuning this parameter helps limit overfitting by increasing the diversity of trees in the ensemble. Additionally, we chose to fix the maximum tree depth to 20 to control the final model size such that each trained model can fit in-memory on the onboard UAV processing computer. The remaining hyperparameters were left to their default values in order to constrain the total optimization space. The performance of each model was evaluated by computing out-of-fold scores for the coefficient of determination as well as the

$$\text{RMSE} = \sqrt{\frac{\sum\limits_{i=1}^{N} (\hat{y}_i - y_i)^2}{N}}, \tag{4}$$

$$\text{MAE} = \frac{\sum\limits_{i=1}^{N} |\hat{y}_i - y_i|}{N}, \tag{5}$$

where RMSE is the root-mean-square error, and MAE is the mean absolute error between the true measurements $y_i$ and the predictions $\hat{y}_i$.

Having identified the model with the best out-of-fold performance, we proceeded to train the final hyperparameter-optimized model on the full training set with the associated uncertainty estimated using conformal prediction. Then, each final model was evaluated on the previously untouched testing set. We visualize model performance across the distribution of the testing data with a scatter diagram and a quantile–quantile plot, for which successful model predictions should lie close to a 1:1 line.

These trained models can then easily be deployed on the onboard processing computer so that during subsequent surveys, target concentrations can be inferred as imagery are collected and processed. The application of each model to the collected hyperspectral data cubes results in a map of the distribution of the water composition across the pond.

## 3. Results

The final dataset of combined observations from each of the three separate collections contains more than 120,000 individual records. Based on the size of the UAV payload and

the available battery capacity, the collection on 23 November 2020 was chosen to cover the broadest possible area, resulting in small horizontal gaps between flight tracks. The collections on 9 December 2020 and 10 December 2020 were designed to complement this collection by sampling a smaller spatial extent with uniform coverage.

The variability of incident lighting conditions for each collection is visualized in Figure 6, wherein the distribution of the total downwelling intensity measured by the downwelling irradiance spectrometer across all hyperspectral data cubes is visualized. Despite performing all UAV flights near solar noon, there were differences in the minimum solar zenith angles between collections due to the time of year. Additionally, there was some slight cloudiness during the 9 December and 10 December collections.
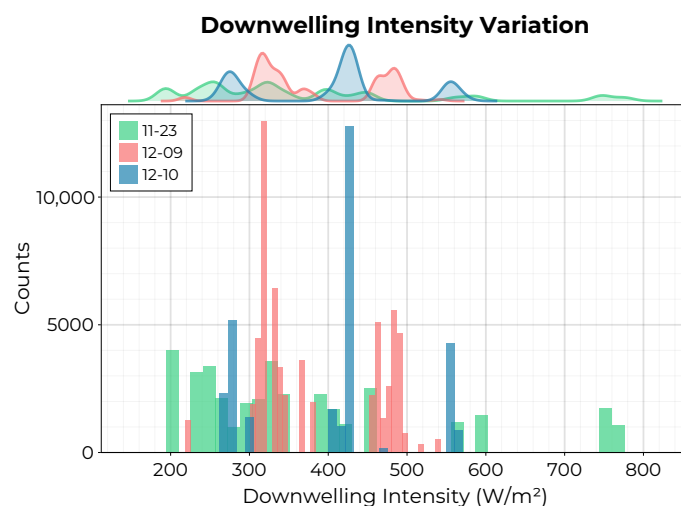


**Figure 6.** Distribution of total downwelling intensity during each of the three HSI collection flights. The multi-modal nature of these distributions reflects the impact of the relative orientation of the drone to the sun as well as potential occlusion due to the presence of clouds.

The results of the model training procedure are presented in Table 2. The performance of the model is identified by the $R^2$, RMSE, and MAE out-of-fold estimates (mean $\pm$ standard deviation) of the final hyperparameter-optimized model in the training set, with the target variables ranked in descending order by the $R^2$ value and separated by sensor type (physical variables, ions, biochemical variables, and chemical variables). The final hyperparameter values for each model are listed in Table A1 in Appendix A. The small variation in values across folds confirms that the reported performance is independent of how the training set was sampled. Furthermore, we report the interval width that yields a 90% confidence interval on the holdout validation set determined by the conformal prediction procedure. We then evaluate how the estimated uncertainty generalizes by computing the empirical coverage on the holdout testing set: that is, we compute the percentage of predictions in the test set that actually fall within the estimated 90% confidence interval.

From Table 2, we see that the empirical coverage achieved by the inferred confidence interval evaluated in the independent test set is within 1% of our desired coverage for each target modeled. This indicates that the uncertainties obtained by the conformal prediction procedure are reliable—at least within the bounds of the collected dataset. We also note that in all cases, the inferred model uncertainties are larger than the resolution of the in situ sensors. This lends further credence to the inferred uncertainty estimates, as we should not expect to be able to have lower uncertainty than the smallest resolvable difference in reference sensor measurements.

To further examine the differences in model performance between the target variables, we can consider the difference between the RMSE and MAE scores. The MAE is less sensitive to the impact of outliers than the RMSE, and as a consequence, any large difference between the two is indicative of impacts due to the distribution of target values. Indeed,

this is the case for turbidity, for which almost all measurements were below 10 FNU, with only a small fraction of observations from a small area near the shore being above this value. The rest of the models all show mean per-fold RMSE values with sizes comparable to the mean per-fold MAEs.

In the remainder of this section, we compare the models within each target category.

**Table 2.** Summary of fitting statistics for each target measurement. Models were evaluated using 6-fold cross-validation on the training set. The estimated uncertainty is evaluated so that a prediction $\hat{y} \pm \Delta y$ achieves 90% coverage on the calibration holdout set. The empirical coverage is the percentage of predictions in the testing set that fall within the inferred confidence interval.

| Target | Units | $R^2$ | RMSE | MAE | Estimated Uncertainty | Empirical Coverage (%) |
|---|---|---|---|---|---|---|
| Temperature | °C | $1.0 \pm 6.04 \times 10^{-6}$ | $0.0289 \pm 0.000466$ | $0.0162 \pm 0.00016$ | ±0.039 | 90.3 |
| Conductivity | µS/cm | $1.0 \pm 1.54 \times 10^{-5}$ | $0.574 \pm 0.0128$ | $0.322 \pm 0.00579$ | ±0.76 | 90.6 |
| pH | 0–14 | $0.994 \pm 0.000288$ | $0.0145 \pm 0.000304$ | $0.00739 \pm 9.49 \times 10^{-5}$ | ±0.017 | 89.5 |
| Turbidity | FNU | $0.897 \pm 0.00611$ | $3.13 \pm 0.084$ | $0.736 \pm 0.0156$ | ±1.1 | 89.8 |
| $Ca^{2+}$ | mg/L | $1.0 \pm 1.06 \times 10^{-5}$ | $0.285 \pm 0.00357$ | $0.137 \pm 0.00224$ | ±0.33 | 89.8 |
| $Cl^-$ | mg/L | $0.995 \pm 0.000196$ | $0.895 \pm 0.0202$ | $0.516 \pm 0.00759$ | ±1.2 | 90.1 |
| $Na^+$ | mg/L | $0.993 \pm 0.000229$ | $6.16 \pm 0.102$ | $2.83 \pm 0.0303$ | ±7.3 | 90.0 |
| Blue–Green Algae (Phycoerythrin) | ppb | $0.995 \pm 0.000601$ | $0.783 \pm 0.0489$ | $0.287 \pm 0.00959$ | ±0.73 | 89.3 |
| CDOM | ppb | $0.965 \pm 0.00352$ | $0.248 \pm 0.0142$ | $0.0921 \pm 0.0024$ | ±0.15 | 89.9 |
| Chlorophyll-a | ppb | $0.908 \pm 0.00664$ | $0.37 \pm 0.00934$ | $0.131 \pm 0.00228$ | ±0.27 | 89.2 |
| Blue–Green Algae (Phycocyanin) | ppb | $0.708 \pm 0.00689$ | $0.749 \pm 0.0129$ | $0.446 \pm 0.00405$ | ±0.93 | 89.8 |
| Crude Oil | ppb | $0.949 \pm 0.00267$ | $0.247 \pm 0.00597$ | $0.0935 \pm 0.00114$ | ±0.17 | 89.8 |
| Optical Brighteners | ppb | $0.943 \pm 0.00122$ | $0.0806 \pm 0.0014$ | $0.0481 \pm 0.000416$ | ±0.095 | 89.8 |

### 3.1. Physical Variables

Physical variables included temperature, conductivity, pH, and turbidity. In the combined dataset, the distributions for temperature and conductivity had two distinct, nonoverlapping regions corresponding to the measurements from 23 November and the measurements from 9 December and 10 December, respectively. The pH value of the pond was slightly alkaline, showing a multimodal spatial distribution with values ranging from 8.0 to 8.6. As mentioned above, the pond water was very clear for each observation period, with most turbidity values ranging between 1 and 3 FNU and very few above 10.

The results of the hyperparameter-optimized RFR fits are shown in Figure 7. Temperature and conductivity show the best performance in the independent training set, with $R^2$ values of 1.0 (to three decimal places). Similarly, the pH model achieves an excellent fit, with most predictions falling close to the 1:1 line. Quantile–quantile plots for these three models further confirm that the distributions of the true and predicted values match. The turbidity model also achieves a strong fit, with a $R^2$ value of 0.905. The scatter diagram and quantile–quantile plot for this target show that the model performance degrades with larger values, for which deviation in the predicted distribution is apparent past 25 FNU.
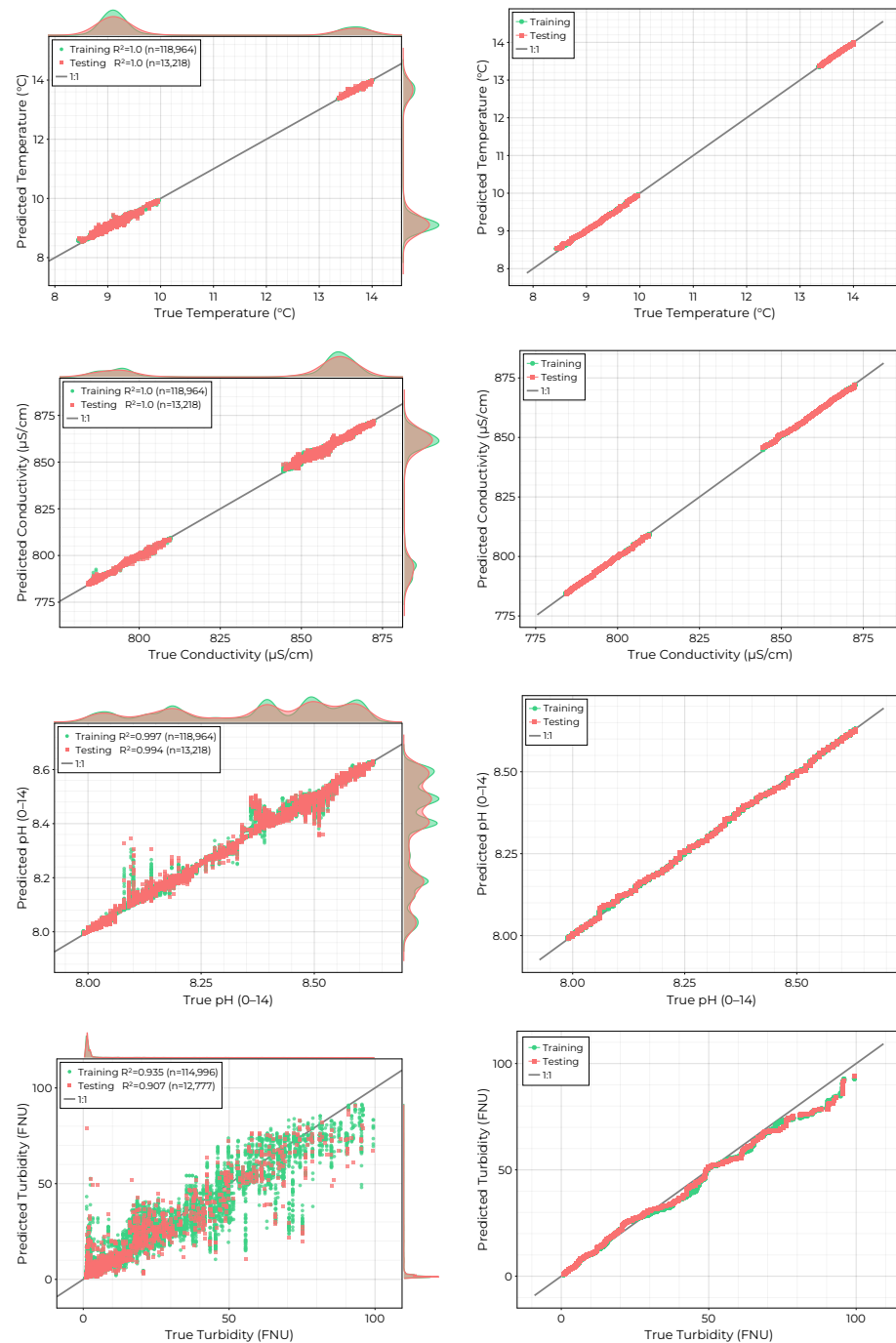
**Figure 7.** Scatter diagrams (**left**) and quantile–quantile plots (**right**) for the hyperparameter-optimized RFR models for the physical variables measured by the USV.

The permutation importance of the top 25 features for each of the models of the physical variables is shown in Figure 8. All four models show strong dependence on the solar illumination geometry (solar azimuth, elevation, and solar zenith) as well as the viewing geometry (pitch, altitude, heading, etc.). All four models also include the total downwelling intensity and the total pixel intensity as highly important features. The temperature, conductivity, and pH models all include red-to-infrared reflectance bins and a combination of spectral indices within their most important features. Finally, the turbidity model relies mainly on blue wavelengths from 462 to 496 nm and did not include any spectral indices amongst the 25 most important features.

By applying the trained models to the full hyperspectral data cubes, we can produce maps of the distributions of the target variables as in Figure 9. Here, we have chosen to show the map produced from the imagery collected during the 23 November collection period, as it showcases the largest spatial extent. The temperature map shows lower values near the shore, which is to be expected as the air temperature was below the water temperature. The temperature, conductivity, and pH maps all show a distinction between the main body of water and the alcove to the east, which receives little flow from the main body. The turbidity map confirms that the water is largely clear but has elevated levels near the shore.



**Figure 8.** Ranked permutation importance for each feature in the physical variable models. Permutation importance measured the decrease in the model's $R^2$ value after replacing each feature in the prediction set with a random permutation of its values.

**Figure 9.** Maps generated by applying each of the physical variable models to the hyperspectral data cubes collected on 23 November. Overlaid over the predictions are color-filled squares showing the associated in situ reference data for the same collection period. The size of the squares has been exaggerated for visualization. We note that there is good agreement between the model predictions and the reference data.

*3.2. Ions*

The measured ions include $Ca^{2+}$, $Cl^-$, and $Na^+$. All three measurements showed multimodal spatial distributions throughout the pond on each of the three collections. The scatter diagrams and quantile–quantile graphs for the resulting fits are shown in Figure 10. All three models achieved excellent fits, with $R^2$ values of 1.0, 0.996, and 0.993 on the independent testing set, respectively. Furthermore, there is no clear decrease in model performance for low or high concentrations; rather, for $Cl^-$ and $Na^+$, the models have the most difficulty in the middle of the target distributions.

The permutation importance rankings for the top 25 features of each of the ion models is shown in Figure 11. Here, we see that all three models depend on the solar illumination and viewing geometries as well as the total downwelling intensity and the total pixel intensity measured by the hyperspectral imager. All three models utilize a combination of spectral indices that combine green, red, and infrared reflectance bins. The $Ca^+$ and $Cl^-$ models depend on specific red wavelengths of 740 to 769 nm. $Cl^-$ and $Na^+$ also depend on green and yellow reflectance bins of 541 to 589 nm.

The maps produced by applying the fitted models to the hyperspectral data cubes for 23 November are shown in Figure 12. Both positive ions $Ca^{2+}$ and $Na^+$ show high concentrations in the northwest portion of the pond, with lower values being measured in the alcove on the eastern side. Positive ion concentrations also appear to decrease near the shore. The negative ion $Cl^-$ shows the opposite distribution, with larger values in the alcove to the east and the lowest values on the western side of the pond. The $Cl^-$ ion concentration also appears to increase near the shore.
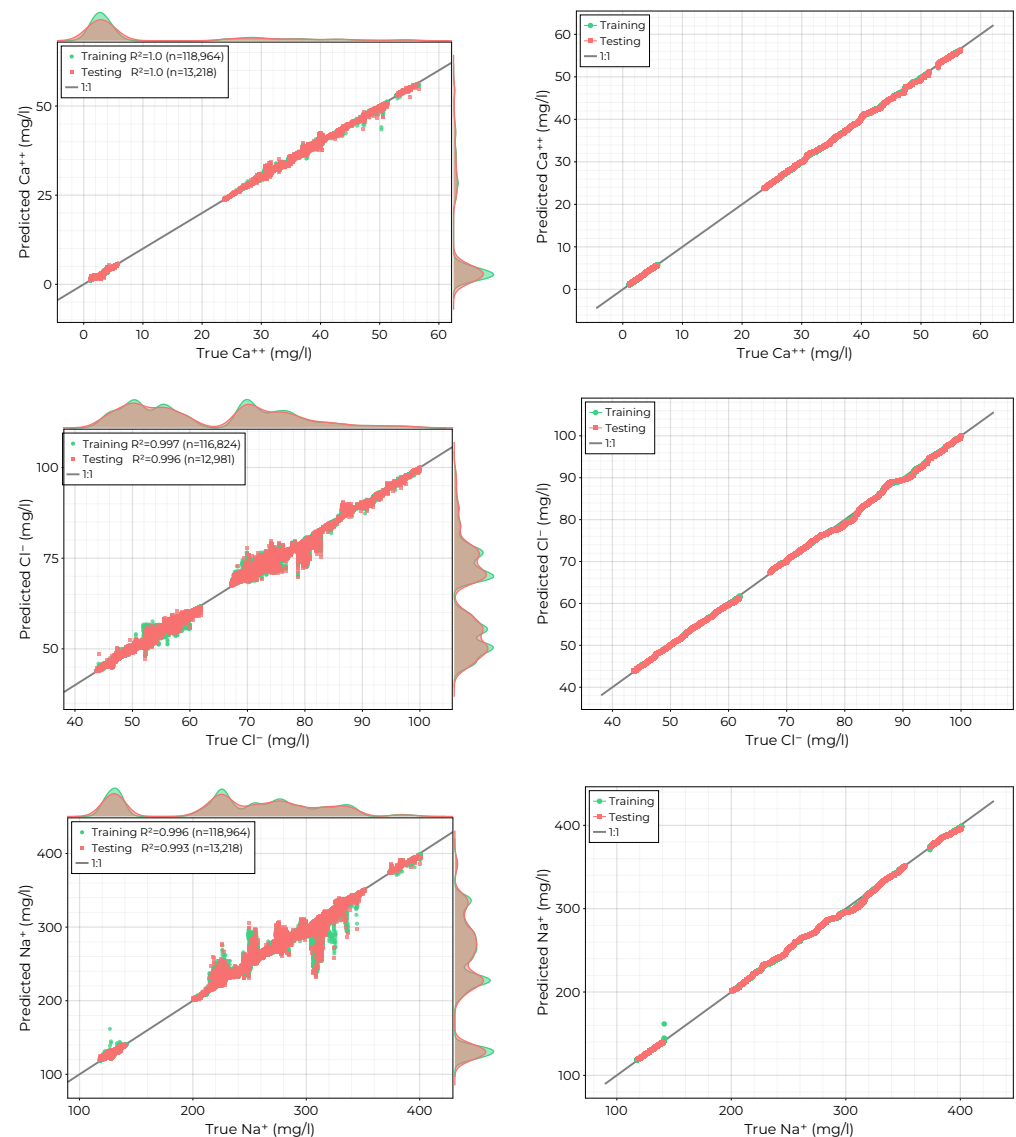
**Figure 10.** Scatter diagrams (**left**) and quantile–quantile plots (**right**) for the hyperparameter-optimized RFR models for the ion measurements made by the USV.
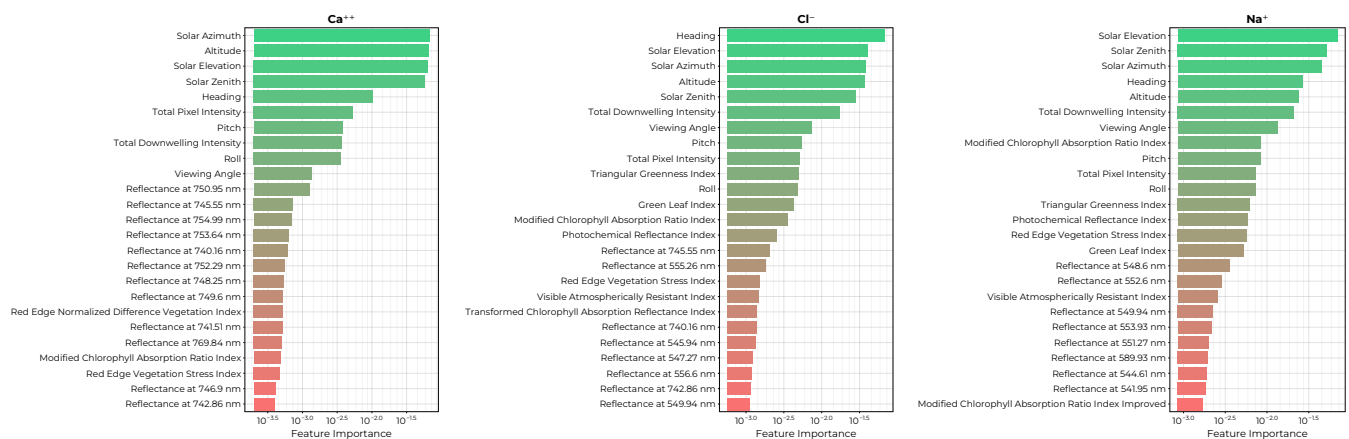


**Figure 11.** Ranked permutation importance for the top 25 features of the ion models. The permutation importance measures the decrease in the model's $R^2$ value when each feature is replaced by a random permutation of its values.
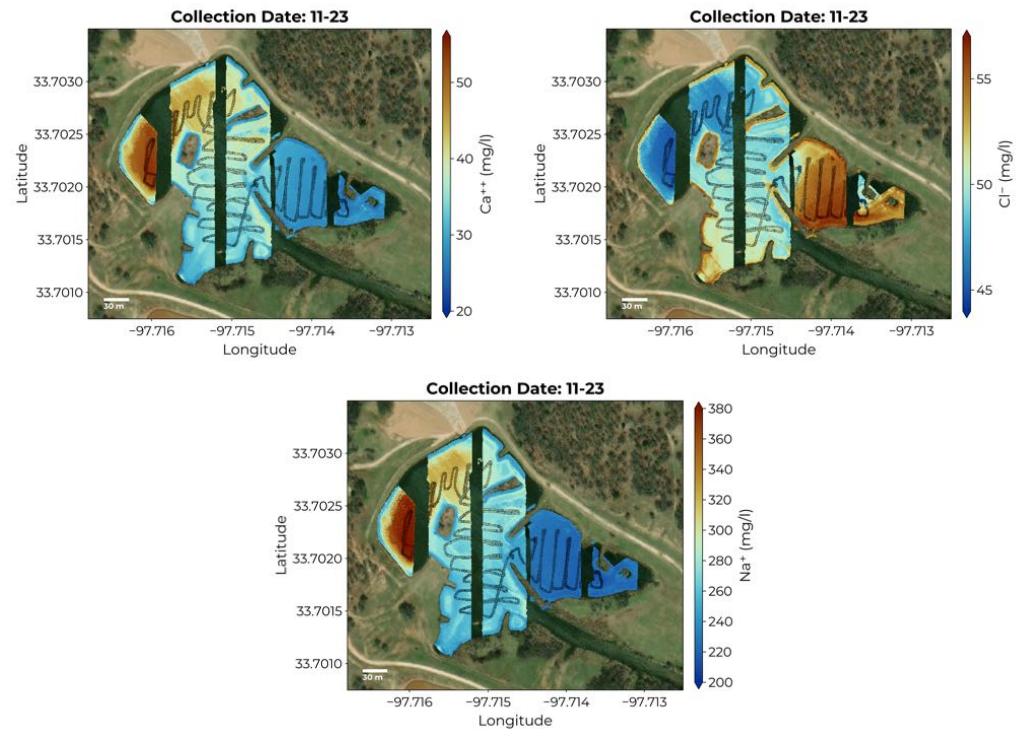
**Figure 12.** Maps generated by applying the trained ion models to the data cubes collected on 23 November. Overlaid on the maps are the in situ reference measurements for the same collection period. The size of the squares has been exaggerated for the visualization. We note that there is good agreement between the generated map and the reference data.

### 3.3. Biochemical Variables

The measured biochemical variables include the pigments phycoerythrin, phycocyanin, and chlorophyll-a, as well as CDOM. Phycoerythrin and phycocyanin are both present in blue–green algae, and chlorophyll-a is found in all photosynthetic organisms except bacteria. In the combined dataset, the three pigments showed multimodal distributions separated by the collection day and with little spatial variation within each individual collection. CDOM showed a variable spatial distribution throughout the pond between the main water body and the eastern alcove on 23 November.

The results of the RFR fits for the biochemical variables are shown in Figure 13. Phycoerythrin showed the best model performance, with an $R^2$ value of 0.995 in the training set. Both CDOM and chlorophyll-a achieved good performance, with $R^2$ values of 0.967 and 0.917 in the training set. Quantile–quantile plots indicate that the CDOM model degrades for values below 16 ppb, where data are sparse. The chlorophyll-a model shows the opposite trend, with poorer performance for concentrations above 5 ppb, for which there are very few records. The phycocyanin model had the lowest performance of the biochemical sensors, with an $R^2$ value of 0.727 and with model predictions rapidly decreasing in quality for concentrations greater than 3 ppb.

The permutation importance ranking of the top 25 features of each biochemical model is shown in Figure 14. Again, all four models include the solar illumination and viewing geometries amongst their most important features as well as the total downwelling intensity and total pixel intensity at the imager. Additionally, all four models include some vegetation indices amongst the top features, which utilize combinations of blue, green, yellow, red, and infrared reflectance bands. The phycoerythrin model shows a preference for green reflectance bins from 544 to 556 nm, while the phycocyanin model prefers blue and red reflectance bins. The CDOM model uses mainly red reflectance values, whereas the chlorophyll-a model includes red, green, and blue reflectance bins.

The maps generated for the 23 November collection by applying trained biochemical models are shown in Figure 15. The three pigments show low concentrations in the body of water but elevated levels near the shore. The CDOM distribution shows spatial variability, with higher values in the eastern alcove—similar to the separation seen in the maps for temperature, conductivity, $Ca^{2+}$, $Cl^-$, and $Na^+$.
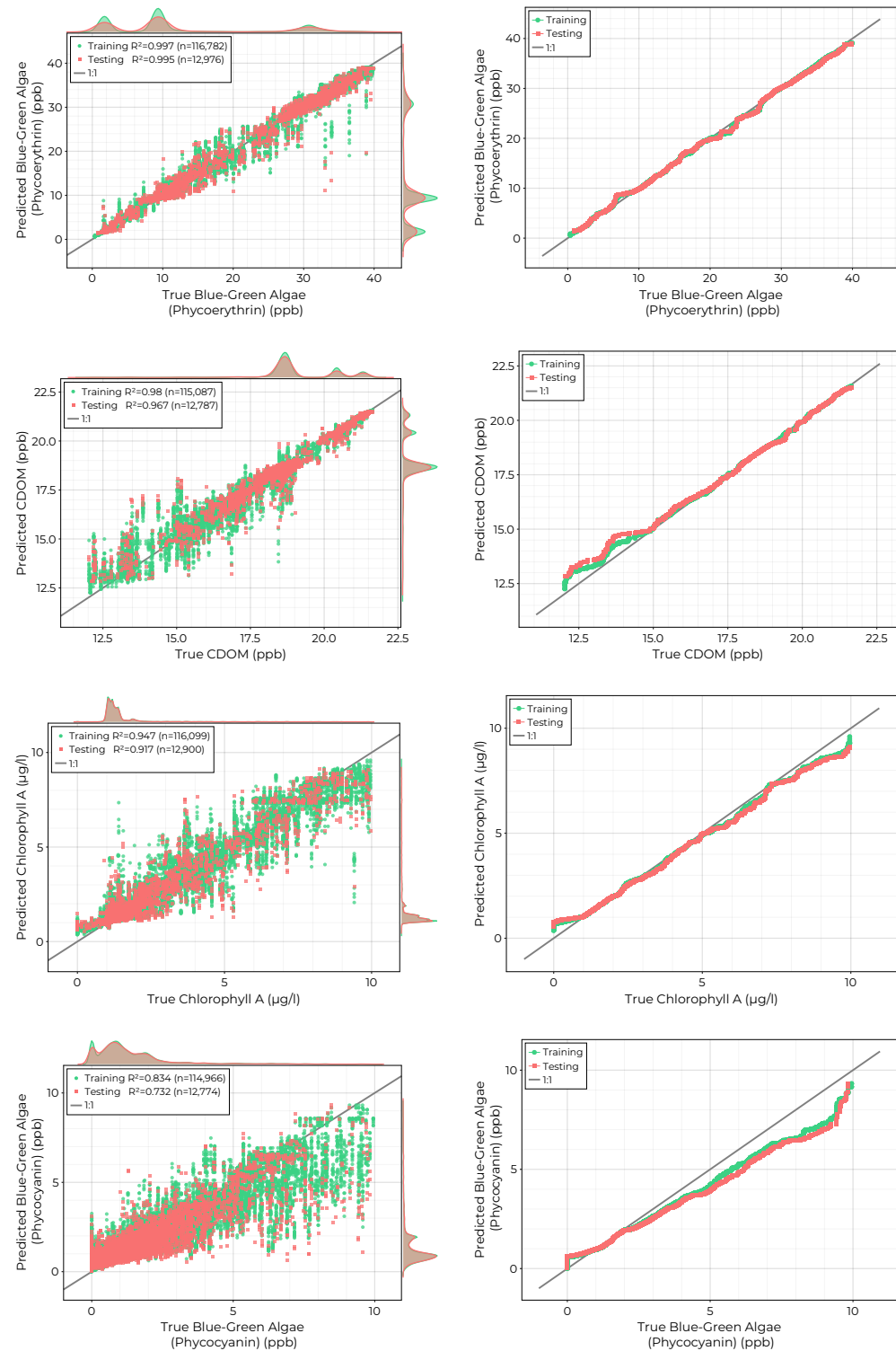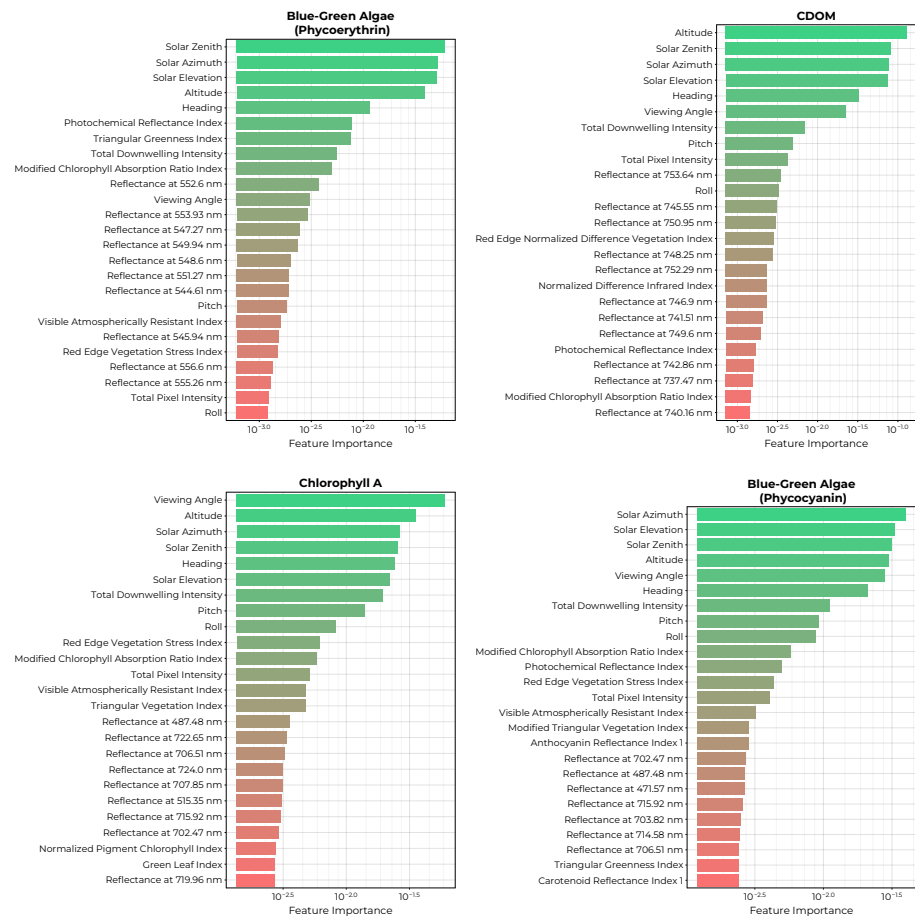


**Figure 13.** Scatter plots (**left**) and quantile–quantile plots (**right**) for the final hyperparameter-optimized models for the biochemical targets blue–green algae (phycoerythrin), CDOM, chlorophyll-a, and blue–green algae (phycocyanin).

**Figure 14.** Ranked permutation importance for each feature in the trained biochemical models. The permutation importance measures the decrease in the model's $R^2$ value after replacing each feature with a random permutation of its values.
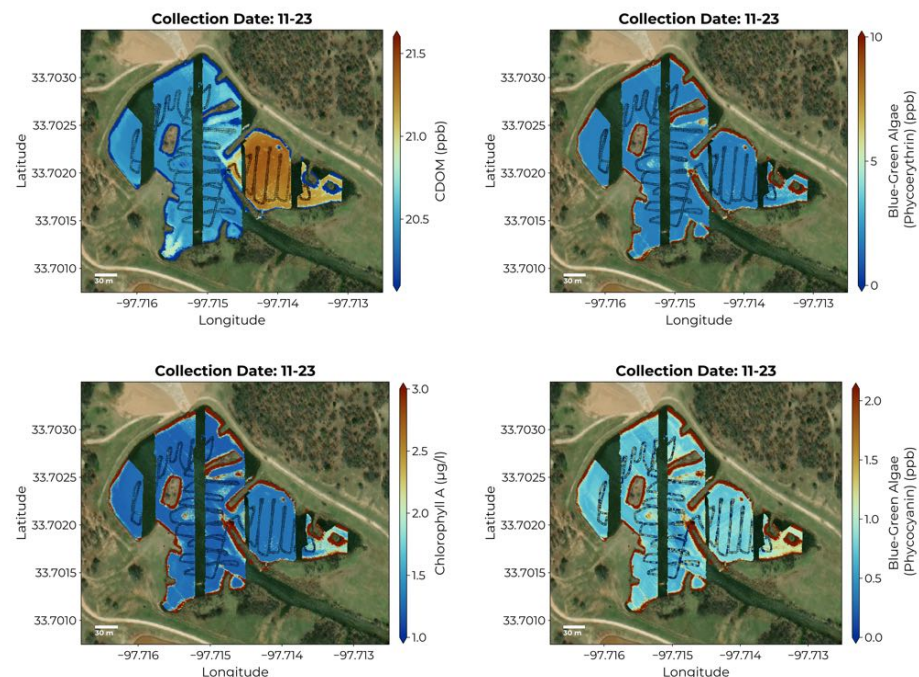


**Figure 15.** Maps generated by applying the trained biochemical models to the data cubes collected on 23 November. Overlaid are the in situ reference data for the same collection period. The size of the squares has been exaggerated for the visualization. We note there is good agreement between the predicted map and the reference data.

### 3.4. Chemical Variables

The final two models to consider are for the measured chemical concentrations of crude oil (CO) and optical brighteners (OB). The crude oil measurement includes natural unprocessed petroleum, whereas optical brighteners consist of whitening agents that are often added to products such as soaps, detergents, and cleaning agents. Both the crude oil and optical brightener measurements show multi-modal spatial distributions across each collection period. Scatter diagrams and quantile–quantile plots for the fitted models are shown in Figure 16. Both models achieve good performance, with $R^2$ values of 0.957 and 0.941 for CO and OB on the holdout test set. The performance of the CO model degrades for concentrations below 24 ppb, for which there are few records. Similarly, the OB model shows worse performance for concentrations below roughly 3.5 ppb.

The ranked permutation importances of the top 25 features for each model are shown in Figure 17. Both models rank the solar illumination and viewing geometries together with the total downwelling intensity and total pixel intensities amongst the top features. Both models include a combination of spectral indices using blue, green, yellow, red, and infrared reflectance bins. Additionally, the CO model includes green–yellow reflectances from 539 to 589 nm as well as red reflectances from 749 to 769 nm. The OB model includes yellow reflectance at 584.6 nm and red reflectance bins.

The maps generated by applying the CO and OB models to the 23 November data cubes are shown in Figure 18. Both models show a distinct spatial distribution, with elevated values in the eastern alcove of the pond—similar to the CDOM distribution in Figure 15.
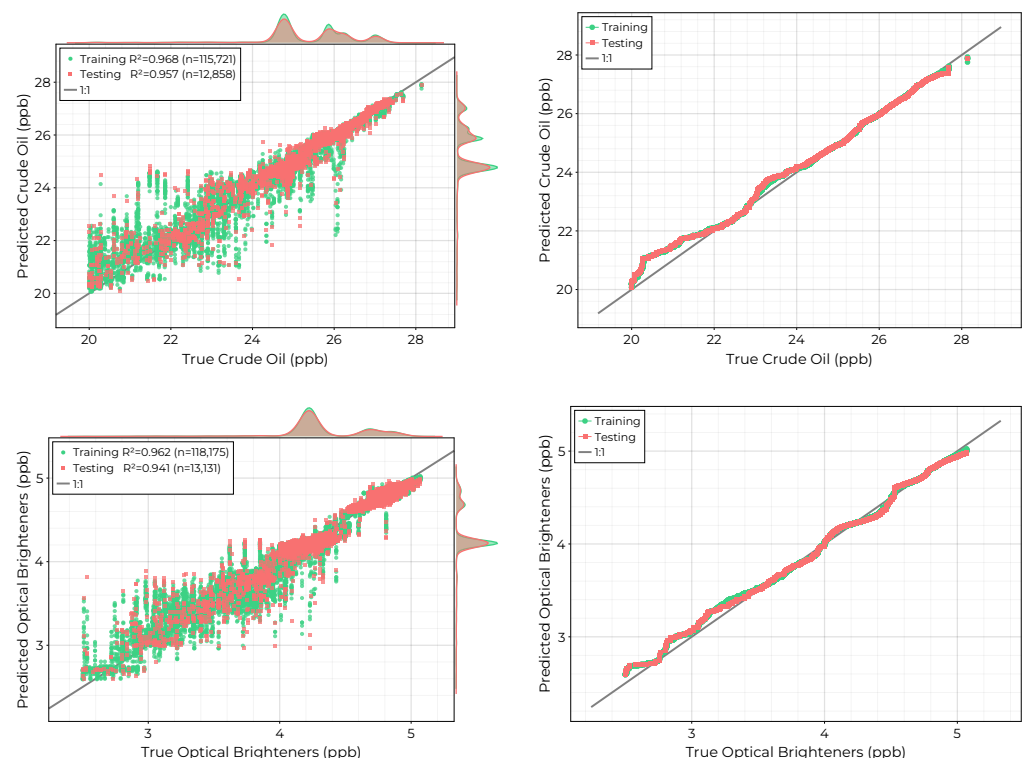


**Figure 16.** Scatter diagrams (**left**) and quantile–quantile plots (**right**) for the hyperparameter-optimized RFR models for the chemical variables measured by the USV.
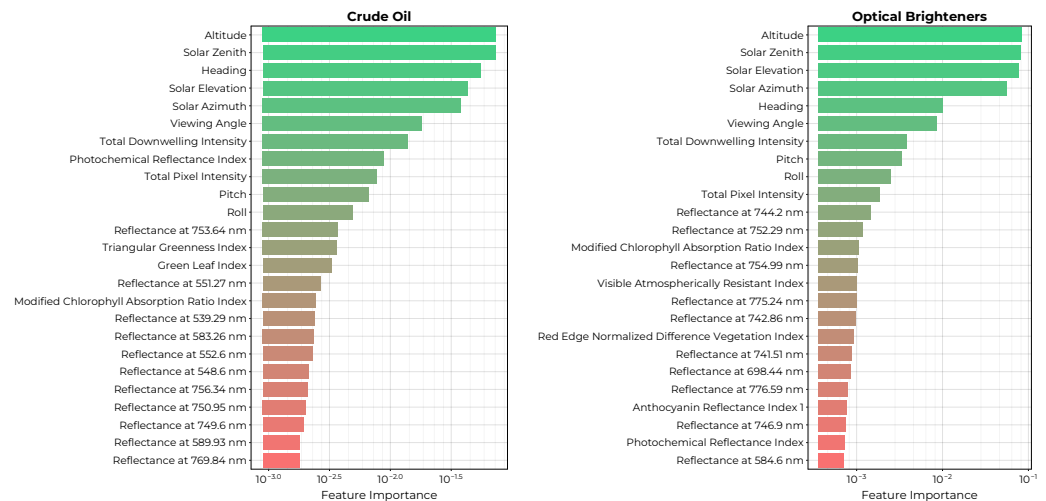
**Figure 17.** Ranked permutation importance for the top 25 features of the chemical models. The permutation importance measures the decrease in the model's $R^2$ value after replacing each feature in the prediction set with a random permutation of its values.
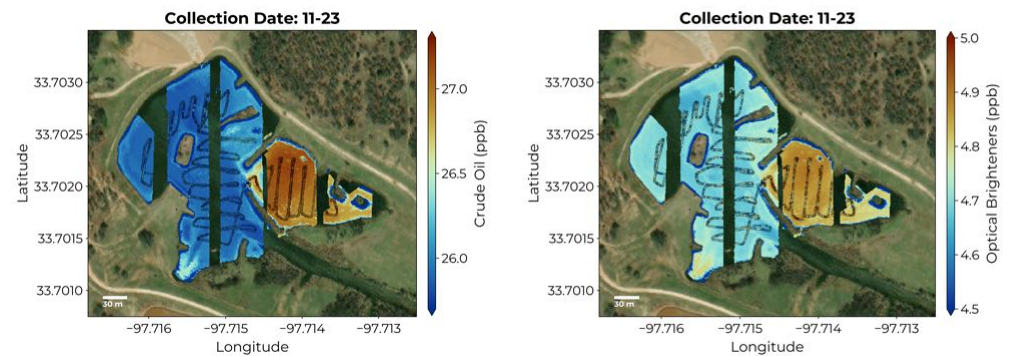


**Figure 18.** Maps generated by applying the trained chemical variable models to the hyperspectral data cubes collected on 23 November. Overlaid are color-filled squares showing the in situ reference data for the same collection period. The size of the squares is exaggerated for the visualization. We note that there is good agreement between the model predictions and reference data.

## 4. Discussion

In recent years, much effort has been spent on the curation of comprehensive datasets combining water quality records with decades of satellite imagery to enable the development of new methods for retrieving water quality parameters. For example, Aurin et al. curated over 30 years of oceanographic field campaign data with associated coincident satellite imagery [3]. Similarly, Ross et al. combined more than 600,000 records of dissolved organic carbon, chlorophyll-a and other water quality variables with historical Landsat reflectance data for the period 1984–2019 [4]. The sensing paradigm we have demonstrated here was able to rapidly collect comparable volumes of data within the span of just three observation periods. Therefore, despite the fact that individual UAV tracks cover far less spatial extent than remote sensing imagery, the ability to collect coordinated in situ measurements together with detailed hyperspectral images offers a significant improvement over these traditional approaches. With a coordinated robot team, one does not need to rely on infrequent satellite overpasses when planning data collection. Furthermore, the time offset between reference measurements and remote sensing is significantly reduced from days to minutes.

This study is not the first to employ UAVs equipped with multispectral or hyperspectral imagers for the purpose of assessing water composition and quality. Indeed there are many such examples focused on inferring optically active and inactive water quality parameters using band ratios and machine learning methods [20,21,56]. The key advance-

ment demonstrated by our robot team is the ability to combine UAV-borne hyperspectral imagers together with comprehensive, in situ sensing for a significant improvement in data volume. Purposefully coordinating USV sampling with the flight tracks of the UAV greatly accelerates data collection by removing the need to acquire individual samples for the calibration of water quality models. Additionally, the USV facilitates rapid validation of model predictions. When a trained model applied to collected hyperspectral imagery suggests elevated levels of a particular water quality parameter, the USV can quickly be provisioned to confirm these estimates with its reference instruments.

In [22], we introduced this paradigm. In this study, we have built on this approach in three new ways. First, we have demonstrated the ability to effectively combine observations from disparate collections by augmenting the machine learning models with sufficient features describing the illumination and viewing geometries. As Figure 6 indicates, we observed variation in the total downwelling intensity between the images collected on the same day and between each separate collection period. These within-collection variations are due to a combination of the stability of the UAV (on which the upward facing downwelling irradiance spectrometer is mounted) together with the occasional interference of clouds. Moreover, the assumption that the water's surface can be treated as Lambertian is clearly violated when the water is not perfectly still. Despite the potential impact of these limitations on the quality of the resulting reflectance data cubes, the smoothness of the maps generated by our models suggests that we have provided sufficient context by including the relevant solar illumination and viewing angles as additional features in the final dataset. This fact is reinforced by the position of these variables as the most important features for *each* of the estimated water quality variables. As long as we are primarily interested in these values and not the reflectances themselves, we are able to successfully account for these lighting effects when combining data from multiple collections.

The second contribution of this study is to explore the breadth of possible water quality and composition parameters that can be accurately mapped by hyperspectral imagery collected by the UAV. The results presented here confirm the ability of the robot team to predict optically active parameters including blue–green algae, chlorophyll-a, CDOM, crude oil, optical brighteners, turbidity, and temperature. Additionally, we are also able to infer the distributions of optically inactive variables including conductivity, pH, and ion concentrations. Other studies using multispectral and hyperspectral remote sensing imagery have also estimated optically inactive water quality parameters, with the ability to do so stemming from the relationship of these variables to optically active properties of the water [57–59]. We note that in our investigation, the models trained for many of these variables outperformed their optically active counterparts. As the abundance of these variables is likely tied to the specific composition and content of the pond, it is unlikely that models trained for these optically inactive variables will generalize to other bodies of water.

The third contribution of this work is the extension of our machine learning approach to enable uncertainty quantification through conformal prediction. For water quality risk assessment, the trustworthiness of model predictions is of equal or greater importance to the values themselves. However, robust uncertainty quantification has historically been challenging for many machine learning models, which behave like black boxes. Conformal prediction is an attractive approach to enable model-agnostic uncertainty estimation and has recently seen adoption to remote sensing classification tasks such as land-type classification and object identification [60,61]. In this setting, the goal is to produce predictive sets guaranteed to contain the correct class labels at a predetermined confidence level. Nevertheless, conformal prediction works equally well for regression tasks. By leveraging the large data volume collected by the robot team, we are able to simultaneously train predictive models and evaluate confidence intervals for their predictions. As the final column of Table 2 confirms, the empirical coverage on the holdout testing set provided by the inferred confidence intervals achieves the desired coverage to within 1%. We chose

to use a 90% confidence interval for this study, but this can easily be adapted to suite the needs of a specific application if greater confidence is required.

Despite the wealth of information provided by the increased resolution of hyperspectral images, their considerable size impedes their complete utilization in real-time applications. Often, much of the available spectral information is discarded in favor of indices like the NDVI, which can be quickly computed as images are captured [62]. Utilizing machine learning allows us to take advantage of the full spectrum captured by each pixel while simultaneously reducing the size of the final data product to single-band "images" of selected water quality variables. We note that training a reduced-feature model without further hyperparameter optimization takes roughly one minute per target variable of interest using the processing computer included on the UAV. This means that, in principle, training data can be collected by the USV, imagery can be acquired and processed by the UAV, coincident records can be selected, and the resulting dataset can be used to train machine learning models all while investigators are still in the field. Analyzing the maps produced by applying each trained model enables areas of interest to be readily pinpointed, as demonstrated by the identification of slightly elevated levels of crude oil, optical brighteners, and CDOM in the eastern alcove of the pond on 23 November.

Finally, we note that the high spectral resolution of the UAV imagery together with the ability to collect precisely co-located reference measurements provides fertile ground for the development of new spectral indices targeted towards water quality variables. In this paper, we have shown that permutation importance ranking for trained machine learning models enables a straightforward interpretation of the relative values of each reflectance bin to the final model predictions. In future work, we plan to utilize this information to identify combinations of spectral bands that can be applied to remote sensing imagery captured by satellites equipped with hyperspectral imagers. The recently launched Environmental Mapping and Analysis Program (EnMAP) is one such example and includes over 91 spectral bands in the VNIR that overlap with those of our hyperspectral imager [63].

## 5. Conclusions

In this study, we address two key limitations of current remote sensing approaches to characterize water quality: namely, the limited spatial, spectral, and temporal resolution provided by existing satellite platforms and the lack of comprehensive in situ measurements needed to validate remote sensing data products. By equipping an autonomous USV with a suite of reference sensors, we rapidly collect significantly more data than existing approaches that rely on the collection of individual samples for lab analysis or are constrained to continuous sensing at fixed sites. Utilizing an autonomous UAV equipped with a hyperspectral imager in tandem with the USV allows us to quickly generate aligned datasets that are used to train machine learning models mapping measured reflectance spectra to the desired water quality variables. By virtue of this increased data volume, we are able to simultaneously estimate the uncertainty of our models by using conformal prediction. Finally, the hyperspectral data cube processing workflow employed onboard the UAV makes it possible to deploy these trained models to swiftly generate maps of the target variables across bodies of water. The rapid turnaround time from data collection to model deployment is critical for real-time water quality evaluation and risk assessment.

**Author Contributions:** Conceptualization, D.J.L.; methodology, J.W. and D.J.L.; software, J.W.; field deployment and preparation, J.W., A.A., L.O.H.W., S.T., A.F., P.M.H.D., M.I., M.L., D.S. and D.J.L.; validation, J.W.; formal analysis, J.W.; investigation J.W.; resources, D.J.L.; data curation, J.W., A.A., L.O.H.W. and D.J.L.; writing—original draft preparation, J.W.; writing—review and editing, J.W. and D.J.L.; visualization, J.W.; supervision, D.J.L.; project administration, D.J.L.; funding acquisition, D.J.L. All authors have read and agreed to the published version of the manuscript.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| GPS | Global Positioning System |
| INS | Inertial Navigation System |
| UTM | Universal Transverse Mercator |
| UV | Ultraviolet |
| ML | Machine Learning |
| USV | Uncrewed Surface Vessel |
| UAV | Unmanned Aerial Vehicle |
| CDOM | Colored Dissolved Organic Matter |
| CO | Crude Oil |
| OB | Optical Brighteners |
| FNU | Formazin Nephelometric Unit |
| RFR | Random Forest Regressor |
| MLJ | Machine Learning framework for Julia |
| RMSE | Root Mean Square Error |
| MAE | Mean Absolute Error |
| RENDVI | Red-Edge Normalized Difference Vegetation Index |

## Appendix A

**Table A1.** Final hyperparameter values for each target model. The number of trees and the sampling ratio were optimized using a random search. The maximum tree depth was fixed to 20 to limit overfitting and to control the size of the final model. The number of sub-features was set to the square root of the total number of features, and the minimum samples per leaf and minimum samples per split were left to their default values.

| Target | Number of Trees | Sampling Ratio | Maximum Tree Depth | Number of Sub-Features | Minimum Samples per Leaf | Minimum Samples per Split |
|---|---|---|---|---|---|---|
| Temperature | 153 | 0.979 | 20 | 5 | 1 | 2 |
| Conductivity | 154 | 0.992 | 20 | 5 | 1 | 2 |
| pH | 103 | 0.972 | 20 | 5 | 1 | 2 |

**Table A1.** *Cont.*

| Target | Number of Trees | Sampling Ratio | Maximum Tree Depth | Number of Sub-Features | Minimum Samples per Leaf | Minimum Samples per Split |
|---|---|---|---|---|---|---|
| Turbidity | 158 | 0.998 | 20 | 5 | 1 | 2 |
| $Ca^{2+}$ | 172 | 0.984 | 20 | 5 | 1 | 2 |
| $Cl^-$ | 110 | 0.999 | 20 | 5 | 1 | 2 |
| $Na^+$ | 103 | 0.972 | 20 | 5 | 1 | 2 |
| Phycoerythrin | 158 | 0.998 | 20 | 5 | 1 | 2 |
| CDOM | 157 | 0.982 | 20 | 5 | 1 | 2 |
| Chlorophyll-a | 158 | 0.998 | 20 | 5 | 1 | 2 |
| Phycocyanin | 142 | 0.995 | 20 | 5 | 1 | 2 |
| Crude Oil | 154 | 0.992 | 20 | 5 | 1 | 2 |
| Optical Brighteners | 157 | 0.982 | 20 | 5 | 1 | 2 |

## References

1. Melesse, A.M.; Weng, Q.; Thenkabail, P.S.; Senay, G.B. Remote sensing sensors and applications in environmental resources mapping and modelling. *Sensors* **2007**, *7*, 3209–3241. [CrossRef]
2. Joyce, K.E.; Belliss, S.E.; Samsonov, S.V.; McNeill, S.J.; Glassey, P.J. A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. *Prog. Phys. Geogr.* **2009**, *33*, 183–207. [CrossRef]
3. Aurin, D.; Mannino, A.; Lary, D.J. Remote sensing of CDOM, CDOM spectral slope, and dissolved organic carbon in the global ocean. *Appl. Sci.* **2018**, *8*, 2687. [CrossRef]
4. Ross, M.R.; Topp, S.N.; Appling, A.P.; Yang, X.; Kuhn, C.; Butman, D.; Simard, M.; Pavelsky, T.M. AquaSat: A data set to enable remote sensing of water quality for inland waters. *Water Resour. Res.* **2019**, *55*, 10012–10025. [CrossRef]
5. Fingas, M.; Brown, C.E. A review of oil spill remote sensing. *Sensors* **2017**, *18*, 91. [CrossRef]
6. Koponen, S.; Attila, J.; Pulliainen, J.; Kallio, K.; Pyhälahti, T.; Lindfors, A.; Rasmus, K.; Hallikainen, M. A case study of airborne and satellite remote sensing of a spring bloom event in the Gulf of Finland. *Cont. Shelf Res.* **2007**, *27*, 228–244. [CrossRef]
7. Bonansea, M.; Rodriguez, M.C.; Pinotti, L.; Ferrero, S. Using multi-temporal Landsat imagery and linear mixed models for assessing water quality parameters in Río Tercero reservoir (Argentina). *Remote Sens. Environ.* **2015**, *158*, 28–41. [CrossRef]
8. Absalon, D.; Matysik, M.; Woźnica, A.; Janczewska, N. Detection of changes in the hydrobiological parameters of the Oder River during the ecological disaster in July 2022 based on multi-parameter probe tests and remote sensing methods. *Ecol. Indic.* **2023**, *148*, 110103. [CrossRef]
9. Lary, D.J. Artificial intelligence in geoscience and remote sensing. In *Geoscience and Remote Sensing New Achievements*; IntechOpen: London, UK, 2010.
10. Peterson, K.T.; Sagan, V.; Sloan, J.J. Deep learning-based water quality estimation and anomaly detection using Landsat-8/Sentinel-2 virtual constellation and cloud computing. *GIScience Remote Sens.* **2020**, *57*, 510–525. [CrossRef]
11. 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]
12. Sagan, V.; Peterson, K.T.; Maimaitijiang, M.; Sidike, P.; Sloan, J.; Greeling, B.A.; Maalouf, S.; Adams, C. Monitoring inland water quality using remote sensing: Potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Sci. Rev.* **2020**, *205*, 103187. [CrossRef]
13. Hruska, R.; Mitchell, J.; Anderson, M.; Glenn, N.F. Radiometric and geometric analysis of hyperspectral imagery acquired from an unmanned aerial vehicle. *Remote Sens.* **2012**, *4*, 2736–2752. [CrossRef]
14. Adão, T.; Hruška, J.; Pádua, L.; Bessa, J.; Peres, E.; Morais, R.; Sousa, J.J. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sens.* **2017**, *9*, 1110. [CrossRef]
15. Banerjee, B.P.; Raval, S.; Cullen, P. UAV-hyperspectral imaging of spectrally complex environments. *Int. J. Remote Sens.* **2020**, *41*, 4136–4159. [CrossRef]
16. Pádua, L.; Vanko, J.; Hruška, J.; Adão, T.; Sousa, J.J.; Peres, E.; Morais, R. UAS, sensors, and data processing in agroforestry: A review towards practical applications. *Int. J. Remote Sens.* **2017**, *38*, 2349–2391. [CrossRef]
17. Arroyo-Mora, J.P.; Kalacska, M.; Inamdar, D.; Soffer, R.; Lucanus, O.; Gorman, J.; Naprstek, T.; Schaaf, E.S.; Ifimov, G.; Elmer, K.; et al. Implementation of a UAV–hyperspectral pushbroom imager for ecological monitoring. *Drones* **2019**, *3*, 12. [CrossRef]
18. Kurihara, J.; Ishida, T.; Takahashi, Y. Unmanned Aerial Vehicle (UAV)-based hyperspectral imaging system for precision agriculture and forest management. In *Unmanned Aerial Vehicle: Applications in Agriculture and Environment*; Springer International Publishing: Cham, Switzerland, 2020; pp. 25–38.
19. Ehmann, K.; Kelleher, C.; Condon, L.E. Monitoring turbidity from above: Deploying small unoccupied aerial vehicles to image in-stream turbidity. *Hydrol. Process.* **2019**, *33*, 1013–1021. [CrossRef]
20. Lu, Q.; Si, W.; Wei, L.; Li, Z.; Xia, Z.; Ye, S.; Xia, Y. Retrieval of water quality from UAV-borne hyperspectral imagery: A comparative study of machine learning algorithms. *Remote Sens.* **2021**, *13*, 3928. [CrossRef]

21. Zhang, D.; Zeng, S.; He, W. Selection and Quantification of Best Water Quality Indicators Using UAV-Mounted Hyperspectral Data: A Case Focusing on a Local River Network in Suzhou City, China. *Sustainability* **2022**, *14*, 16226. [CrossRef]

22. Lary, D.J.; Schaefer, D.; Waczak, J.; Aker, A.; Barbosa, A.; Wijeratne, L.O.H.; Talebi, S.; Fernando, B.; Sadler, J.Z.; Lary, T.; et al. Autonomous Learning of New Environments with a Robotic Team Employing Hyper-Spectral Remote Sensing, Comprehensive In-Situ Sensing and Machine Learning. *Sensors* **2021**, *21*, 2240. [CrossRef]

23. Meier, L. QGroundControl. MAVLink Micro Air Vehicle Communication Protocol. 2010. Available online: http://qgroundcontrol.org/mavlink/start (accessed on 30 January 2019).

24. De Marco, R.; Clarke, G.; Pejcic, B. Ion-selective electrode potentiometry in environmental analysis. *Electroanal. Int. J. Devoted Fundam. Pract. Asp. Electroanal.* **2007**, *19*, 1987–2001. [CrossRef]

25. Trees, C.C.; Bidigare, R.R.; Karl, D.M.; Van Heukelem, L.; Dore, J. Fluorometric chlorophyll a: Sampling, laboratory methods, and data analysis protocols. In *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, NASA/TM-2002-210004/Rev3-Vol2*; Mueller, J.L., Fargion, G.S., Eds.; NASA Goddard Space Flight Center: Greenbelt, MD, USA, 2002; pp. 269–283.

26. Tillman, E.F. *Evaluation of the Eureka Manta2 Water-Quality Multiprobe Sonde*; Technical Report; US Geological Survey: Reston, VA, USA, 2017.

27. Brient, L.; Lengronne, M.; Bertrand, E.; Rolland, D.C.; Sipel, A.; Steinmann, D.; Baudin, I.; Legeas, M.; Rouzic, B.L.; Bormans, M. A phycocyanin probe as a tool for monitoring cyanobacteria in freshwater bodies. *J. Environ. Monit. JEM* **2008**, *10*, 248–255. [CrossRef] [PubMed]

28. Boyer, J.N.; Kelble, C.R.; Ortner, P.B.; Rudnick, D.T. Phytoplankton bloom status: Chlorophyll a biomass as an indicator of water quality condition in the southern estuaries of Florida, USA. *Ecol. Indic.* **2009**, *9*, S56–S67. [CrossRef]

29. Brown, C.E.; Fingas, M.F. Review of the development of laser fluorosensors for oil spill application. *Mar. Pollut. Bull.* **2003**, *47*, 477–484. [CrossRef]

30. Cao, Y.; Griffith, J.F.; Weisberg, S.B. Evaluation of optical brightener photodecay characteristics for detection of human fecal contamination. *Water Res.* **2009**, *43*, 2273–2279. [CrossRef] [PubMed]

31. Piper, A.M. A graphic procedure in the geochemical interpretation of water-analyses. *Eos Trans. Am. Geophys. Union* **1944**, *25*, 914–928.

32. Dordoni, M.; Zappalà, P.; Barth, J.A. A preliminary global hydrochemical comparison of lakes and reservoirs. *Front. Water* **2023**, *5*, 1084050. [CrossRef]

33. Pace, M.L.; Reche, I.; Cole, J.J.; Fernández-Barbero, A.; Mazuecos, I.P.; Prairie, Y.T. pH change induces shifts in the size and light absorption of dissolved organic matter. *Biogeochemistry* **2012**, *108*, 109–118. [CrossRef]

34. Ruddick, K.G.; Voss, K.; Banks, A.C.; Boss, E.; Castagna, A.; Frouin, R.; Hieronymi, M.; Jamet, C.; Johnson, B.C.; Kuusk, J.; et al. A review of protocols for fiducial reference measurements of downwelling irradiance for the validation of satellite remote sensing data over water. *Remote Sens.* **2019**, *11*, 1742. [CrossRef]

35. Müller, R.; Lehner, M.; Muller, R.; Reinartz, P.; Schroeder, M.; Vollmer, B. A Program for Direct Georeferencing of Airborne and Spaceborne Line Scanner Images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2002**, *34*, 148–153.

36. Bäumker, M.; Heimes, F. New calibration and computing method for direct georeferencing of image and scanner data using the position and angular data of an hybrid inertial navigation system. In Proceedings of the OEEPE Workshop, Integrated Sensor Orientation, Hannover, Germany, 17–18 September 2001; pp. 1–17.

37. Mostafa, M.M.R.; Schwarz, K.P. A Multi-Sensor System for Airborne Image Capture and Georeferencing. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 1417–1424.

38. Bezanson, J.; Karpinski, S.; Shah, V.B.; Edelman, A. Julia: A fast dynamic language for technical computing. *arXiv* **2012**, arXiv:1209.5145.

39. Vegetation Indices Background. 2023. Available online: https://www.nv5geospatialsoftware.com/docs/backgroundvegetationindices.html (accessed on 3 January 2024).

40. Thenkabail, P.S.; Lyon, J.G.; Huete, A. *Hyperspectral Indices and Image Classifications for Agriculture and Vegetation*; CRC Press: Boca Raton, FL, USA, 2018.

41. Kaufman, Y.J.; Tanre, D. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 261–270. [CrossRef]

42. Zheng, Q.; Huang, W.; Cui, X.; Shi, Y.; Liu, L. New Spectral Index for Detecting Wheat Yellow Rust Using Sentinel-2 Multispectral Imagery. *Sensors* **2018**, *18*, 868. [CrossRef] [PubMed]

43. Blaom, A.D.; Kiraly, F.; Lienart, T.; Simillides, Y.; Arenas, D.; Vollmer, S.J. MLJ: A Julia package for composable machine learning. *J. Open Source Softw.* **2020**, *5*, 2704. [CrossRef]

44. Sadeghi, B.; Chiarowongse, P.; Squire, K.; Jones, D.C.; Noack, A.; St-Jean, C.; Huijzer, R.; Schätzle, R.; Butterworth, I.; Peng, Y.; et al. DecisionTree.jl—A Julia implementation of the CART Decision Tree and Random Forest algorithms. *Zenodo* **2022**. [CrossRef]

45. Breiman, L. *Classification and Regression Trees*; CRC Press: Boca Raton, FL, USA, 2017.

46. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

47. Grinsztajn, L.; Oyallon, E.; Varoquaux, G. Why do tree-based models still outperform deep learning on typical tabular data? *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 507–520.

48. Shwartz-Ziv, R.; Armon, A. Tabular data: Deep learning is not all you need. *Inf. Fusion* **2022**, *81*, 84–90. [CrossRef]

49. Louppe, G.; Wehenkel, L.; Sutera, A.; Geurts, P. Understanding variable importances in forests of randomized trees. *Adv. Neural Inf. Process. Syst.* **2013**, *26*, 431–439.

50. Strobl, C.; Boulesteix, A.L.; Kneib, T.; Augustin, T.; Zeileis, A. Conditional variable importance for random forests. *BMC Bioinform.* **2008**, *9*, 307. [CrossRef]

51. Parr, T.; Turgutlu, K.; Csiszar, C.; Howard, J. Beware default random forest importances. *March* **2018**, *26*, 2018.

52. Shafer, G.; Vovk, V. A Tutorial on Conformal Prediction. *J. Mach. Learn. Res.* **2008**, *9*, 371–421.

53. Angelopoulos, A.N.; Bates, S. A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv* **2021**, arXiv:2107.07511.

54. Fontana, M.; Zeni, G.; Vantini, S. Conformal prediction: A unified review of theory and new challenges. *Bernoulli* **2023**, *29*, 1–23. [CrossRef]

55. Papadopoulos, H. Inductive conformal prediction: Theory and application to neural networks. In *Tools in Artificial Intelligence*; IntechOpen: London, UK, 2008.

56. Vogt, M.C.; Vogt, M.E. Near-remote sensing of water turbidity using small unmanned aircraft systems. *Environ. Pract.* **2016**, *18*, 18–31. [CrossRef]

57. Vakili, T.; Amanollahi, J. 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.* **2020**, *247*, 119134. [CrossRef]

58. Guo, H.; Huang, J.J.; Chen, B.; Guo, X.; Singh, V.P. A machine learning-based strategy for estimating non-optically active water quality parameters using Sentinel-2 imagery. *Int. J. Remote Sens.* **2021**, *42*, 1841–1866. [CrossRef]

59. 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]

60. Valle, D.; Izbicki, R.; Leite, R.V. Quantifying uncertainty in land-use land-cover classification using conformal statistics. *Remote Sens. Environ.* **2023**, *295*, 113682. [CrossRef]

61. Zhu, N.; Xi, Z.; Wu, C.; Zhong, F.; Qi, R.; Chen, H.; Xu, S.; Ji, W. Inductive Conformal Prediction Enhanced LSTM-SNN Network: Applications to Birds and UAVs Recognition. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 3502705. . [CrossRef]

62. Horstrand, P.; Guerra, R.; Rodríguez, A.; Díaz, M.; López, S.; López, J.F. A UAV platform based on a hyperspectral sensor for image capturing and on-board processing. *IEEE Access* **2019**, *7*, 66919–66938. [CrossRef]

63. Storch, T.; Honold, H.P.; Chabrillat, S.; Habermeyer, M.; Tucker, P.; Brell, M.; Ohndorf, A.; Wirth, K.; Betz, M.; Kuchler, M.; et al. The EnMAP imaging spectroscopy mission towards operations. *Remote Sens. Environ.* **2023**, *294*, 113632. [CrossRef]