



Article

Evaluation and Refinement of Chlorophyll-a Algorithms for High-Biomass Blooms in San Francisco Bay (USA)

Raphael M. Kudela ^{1,*}, David B. Senn ², Emily T. Richardson ³, Keith Bouma-Gregson ³, Brian A. Bergamaschi ³ and Lawrence Sim ²

¹ Ocean Sciences Department, University of California Santa Cruz, Santa Cruz, CA 95064, USA

² San Francisco Estuary Institute, 4911 Central Avenue, Richmond, CA 94804, USA; davids@sfei.org (D.B.S.); lawrences@sfei.org (L.S.)

³ USGS California Water Science Center, 6000 J Street, Sacramento, CA 95819, USA; erichardson@usgs.gov (E.T.R.); kbouma-gregson@usgs.gov (K.B.-G.); bbergama@usgs.gov (B.A.B.)

* Correspondence: kudela@ucsc.edu; Tel.: +1-831-459-3290

Abstract: A massive bloom of the raphidophyte *Heterosigma akashiwo* occurred in summer 2022 in San Francisco Bay, causing widespread ecological impacts including events of low dissolved oxygen and mass fish kills. The rapidly evolving bloom required equally rapid management response, leading to the use of near-real-time image analysis of chlorophyll from the Ocean and Land Colour Instrument (OLCI) aboard Sentinel-3. Standard algorithms failed to adequately capture the bloom, signifying a need to refine a two-band algorithm developed for coastal and inland waters that relates the red-edge part of the remote sensing reflectance spectrum to chlorophyll. While the bloom was the initial motivation for optimizing this algorithm, an extensive dataset of in-water validation measurements from both bloom and non-bloom periods was used to evaluate performance over a range of concentrations and community composition. The modified red-edge algorithm with a simplified atmospheric correction scheme outperformed existing standard products across diverse conditions, and given the modest computational requirements, was found suitable for operational use and near-real-time product generation. The final version of the algorithm successfully minimizes error for non-bloom periods when chlorophyll *a* is typically $<30 \text{ mg m}^{-3}$, while also capturing bloom periods of $>100 \text{ mg m}^{-3}$ chlorophyll *a*.

Keywords: remote sensing; chlorophyll; San Francisco Bay; OLCI; *Heterosigma*; harmful algal boom



Citation: Kudela, R.M.; Senn, D.B.; Richardson, E.T.; Bouma-Gregson, K.; Bergamaschi, B.A.; Sim, L. Evaluation and Refinement of Chlorophyll-a Algorithms for High-Biomass Blooms in San Francisco Bay (USA). *Remote Sens.* **2024**, *16*, 1103. <https://doi.org/10.3390/rs16061103>

Academic Editor: SeungHyun Son

Received: 14 February 2024

Revised: 13 March 2024

Accepted: 18 March 2024

Published: 21 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

San Francisco Bay (Bay) is the largest estuary system on the west coast of North America, draining about 40% of California's land area [1], while the surrounding population of ~7 million people has the potential to put considerable stress on the ecosystem through nutrient pollution [2]. As a result, monitoring water quality is a vital aspect of Bay-wide management. One emerging concern is the proliferation of potentially harmful algal blooms (HABs). HABs have great potential to impact estuaries such as the Bay, because estuarine systems are both highly populated by humans and highly productive. Estuaries provide numerous valuable ecosystem functions and are sites of intensive aquaculture, subsistence, and commercial fisheries, all of which are threatened by HABs [3].

As described in [2], nutrient over-enrichment has led to ecosystem impairments in the majority of the world's estuaries [4–7]. This impairment often includes the presence or expansion of HAB organisms responding directly to nutrient enrichment [8]. The Bay has largely been resistant to proliferation of HABs, in part due to vigorous tidal mixing and flushing and light limitation that precludes biomass accumulation [2,9]. Despite the historical resistance to large HAB events in the Bay, multiple HAB threats exist. In particular, the dinoflagellates *Heterocapsa* (fish-killing; [10]), and *Akashiwo sanguinea* (harmful to

birds; [3,11]) have on occasion produced expansive blooms in the Bay [12]. The raphidophyte *Heterosigma akashiwo* (fish-killing; exact toxic mechanism is presently unknown; [13]) is also present in the Bay but has not bloomed significantly since summer 2002 [14,15]. *Heterosigma akashiwo* re-emerged in the Bay in summer 2022, and again in summer 2023, suggesting that recent conditions may be favoring more frequent blooms.

1.1. Remote Sensing of Coastal Waters

Several water quality indicators including turbidity, dissolved organic matter [16], and chlorophyll *a* (*chl_a*) directly affect the optical properties of water, and can therefore be estimated using satellites. However, satellite remote sensing has thus far generally been under-utilized by resource managers in the Bay region. This is due in part to the optical complexity of estuarine waters and the need for sensors with high spectral and spatial resolution to adequately disentangle the remote sensing signal and achieve quantifiable results (e.g., [16,17]). The most commonly used algorithms for retrieval of *chl_a*, including standard ocean color (OC) products from the National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA), generally rely on the ratio of blue to green reflectance spectrum and were developed for the open ocean [18]. Coastal waters exhibit unique challenges for these algorithms because of elevated levels of colored dissolved organic matter (CDOM) that absorbs strongly in blue reflectance bands but does not necessarily covary with *chl_a*. Additionally, elevated concentrations of suspended solids in coastal waters tend to elevate reflectance in all wavelengths, but particularly longer wavelengths, like green and red bands.

To address these issues for coastal and inland waters, *chl_a* algorithms of varying complexity (e.g., [17,19–22]) have been developed, and several have been proposed specifically for the identification of harmful algal blooms (e.g., [23–26]). A recent analysis [22] directly compared multiple models grouped loosely into blue-green band ratios, such as the standard OC products from NASA and ESA and Red–Near-Infrared approaches from various authors. That analysis demonstrated that blue-green models perform poorly in turbid (e.g., coastal) waters, while the family of “red-edge” algorithms exhibits similar performance after tuning with in situ matchups.

1.2. Red-Edge Algorithms

One particularly effective method for retrieving *chl_a* in coastal waters relies on the use of the red-edge feature, a peak in reflectance occurring around 700 nm, which is insensitive to CDOM and slightly sensitive to suspended solids [20,21]. The red part of the visible spectrum is also less sensitive to issues with atmospheric correction, making these algorithms easier to implement operationally while reducing potential sources of error. A significant disadvantage is that water strongly absorbs in the red and near-infrared part of the spectrum, so in the absence of significant *chl_a* biomass, there can be very little signal, resulting in poor performance below about 1 mg m^{-3} *chl_a* [20].

With this background in mind, we target one specific algorithm, referred to as Red Edge 10 (RE10), proposed by [20] and modified for Chesapeake Bay [21] to adapt for use in the Bay. This algorithm was chosen because it can be adjusted for the optical properties of specific blooms, it is computationally inexpensive, and it is amenable to operational use [21]. For comparison, we provide results from the standard OC-type algorithms [18], and also evaluate three atmospheric correction methods commonly used for operational products: the Case 2 Regional CoastColour (C2RCC) atmospheric correction, as implemented in the Sentinel Applications Platform (SNAP) toolbox, the standard neural network approach used by ESA, and a simple black pixel assumption [27–29]. We utilize ocean color data from the Ocean and Land Color Instrument (OLCI) aboard the Sentinel-3 series (S3A and S3B) of satellites. OLCI has reasonable spatial and temporal resolution (nominally 300 m, ~daily imagery from S3A and S3B), good signal-to-noise ratios, and appropriate bands for red-edge algorithms. Our primary goals are to demonstrate the efficacy of a modified

RE10 algorithm tuned for the Bay and to highlight potential for operational use to support management decisions.

2. Materials and Methods

2.1. Study Area

The Bay can be simplified into a series of connected study areas: North Bay, comprising San Pablo and Suisun Bays to the north, Central Bay, which provides direct connection to the Pacific Ocean through the Golden Gate, and South Bay (Figure 1). The Bay is connected to the Sacramento–San Joaquin Delta (Delta) to the northeast, and the Pacific Ocean to the west. North Bay is characterized by a strong salinity gradient, and river water is the primary regulator of bay-wide salinity [1]. North Bay is river-dominated, with salinities ranging from 0 to 15, while South Bay is a marine lagoon, with salinities ranging from 5 to 35. North Bay also tends to have the most riverine-derived suspended sediment.

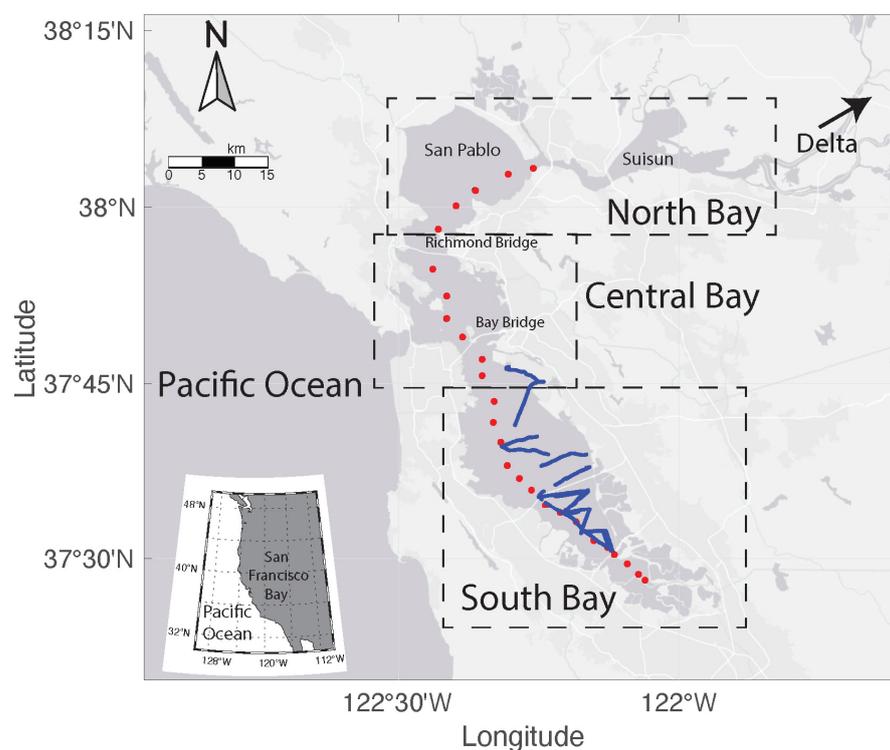


Figure 1. Map of the study site, showing San Francisco Bay (Bay) and the three major basins as defined in the text. Major U.S. Geological Survey (USGS) water quality monitoring stations are shown as filled circles, while USGS underway mapping data are shown as blue segments. The direction of the upstream location Sacramento–San Joaquin Delta is indicated by the arrow labeled “Delta”.

The North, Central, and South Bay regions can be approximately defined by the location of bridges crossing the Bay; North Bay is the region of the Bay north of the Richmond Bridge, Central Bay is the region between the Richmond and Bay Bridges, and South Bay is the region south of the Bay Bridge. The Bay overall is a shallow wetland system, with most of the Bay being less than 3 m deep, except for a dredged shipping channel which runs through the middle of the Bay from north to south, reaching depths greater than 10 m. Based on field observations, microscopic identification, and remote sensing imagery, the onset of the 2022 *H. akashiwo* bloom occurred in Central Bay and spread into South and North Bay.

2.2. Data Overview

To evaluate whether remote sensing could provide reasonable estimates of *chl*_a during extreme events in an operational context, we focus on the *H. akashiwo* bloom that occurred

from about August to September 2022. Satellite data were obtained from the European Space Agency as full-resolution L1A (no atmospheric correction) or L2 (atmospherically corrected, *chla*-derived) data products and processed in SNAP version 8.0 (see Supplementary Table S1). The ground-based data used for validation and algorithm tuning included data from discrete samples collected at the stations indicated in Figure 1, and high-resolution underway mapping data from the U.S. Geological Survey (USGS) collected for June, August, and September 2021, and July, August, and September 2022 [30,31].

2.3. Model Tuning Data

High-resolution mapping data include continuous (1 hz) underway chlorophyll fluorescence (fChl; YSI EXO2) collected from near the surface (~1 m depth) through a pressure-compensated manifold [26]. The in situ fChl measurements were median-filtered (20 s) and compared to results from discrete laboratory measurements of >0.7 micron *chla* (collected from ~1 m depth with a submerged centrifugal pump; ref. [31]) (Supplementary Figure S1; Supplementary Table S2). High-resolution mapping data were further median-binned to 300 m spatial resolution, and the corresponding point location to the satellite overpass was used for spatial matchups with a ± 15 min window applied. Pixels overlapping land in the satellite imagery were discarded, as well as pixels with Rayleigh-corrected top-of-atmosphere (TOA) reflectance (ρ , dimensionless) >0.5, indicative of pixels too bright to be water (e.g., physical structures). A total of 426 potential matchups from 2021–2022 were used in this analysis (see Supplementary Table S3). For any individual algorithm, some matchup points were removed due to various failures, typically related to atmospheric correction (see Supplementary Figure S2).

2.4. Model Validation Data

Water quality measurements used for qualitative model validation included fChl collected at regular point stations at 1 m depth for the USGS water quality cruises, corrected with discrete *chla* [32]. Quality control data from the USGS are available for public use [32].

2.5. Chla Algorithms

Three atmospheric correction schemes and four *chla* algorithms were evaluated. For the atmospheric correction, the standard ESA processing was used for Level 2 water products (*OL_2_WFR*), which uses a neural network approach as part of the “Alternative Atmospheric Correction”. Level 1 files (*OL_1_EFR*) for the same satellite overpasses were separately processed using the Case 2 Regional CoastColour (C2RCC) atmospheric correction as implemented in the Sentinel SNAP toolbox. SNAP was also used to provide Rayleigh-corrected TOA reflectances (ρ , dimensionless) for bands at 665, 708.75, and 885 nm from the C2RCC-processed imagery. The $\rho(885)$ values were subtracted from $\rho(665)$ and $\rho(708.75)$ prior to use in the red-edge algorithms described below. *Chla* was extracted from ESA *OL_2_WFR* standard products using the OC4Me maximum band ratio semi-analytical algorithm [33], and from ESA *OL_1_EFR* files using the C2RCC neural network-derived *chla* [28]. Variations of the red-edge algorithms were used to estimate *chla* with the reflectance data from the C2RCC processor.

For the red-edge algorithms, following [20], we started with the following equation:

$$[chla] = [35.75 \times R^2 - 19.30]^{1.124} \quad (1)$$

where R^2 is the ratio of dimensionless water reflectance (λ_2/λ_1), using the red bands $\lambda_1 = 665$ nm and $\lambda_2 = 708.75$ nm, and the reflectances are based on the TOA, dark-pixel corrected data. We refer to this original formulation as RE10.

Wynne et al. [21] adjusted RE10 by treating the offset correction (19.3 in RE10) as a tunable parameter and proposed 14.30 as optimal for Chesapeake Bay:

$$[chla] = [35.75 \times R^2 - 14.30]^{1.124} \quad (2)$$

We refer to this version of the algorithm as RE22. For this analysis, we treated the offset as a tunable parameter, and also considered the exponential term (1.124 in Equations (1) and (2)) as tunable based on the theoretical description provided by [20]. In the original formulation, the power function was derived from the relationship between the phytoplankton-specific absorption coefficient, a^*_{ph} ($m^2(mg\ chl)^{-1}$), and the measured chl_a concentration ($mg\ m^{-3}$). Those authors noted that the RE10 algorithm is sensitive to a^*_{ph} , and used representative spectra from various algal functional groups to derive the exponential term.

We directly measured a^*_{ph} using standard methods [34] for *H. akashiwo* cultures that were isolated from the bloom in 2022 (Figure 2). Lower a^*_{ph} values result in a proportionally higher exponential term. We therefore allowed the exponential to vary up to a value of 1.375. For the final algorithm, based on the optimization of multiple statistical criteria, we chose a switching version of the red-edge algorithm that applies different tuning parameters to $chl_a < 30\ mg\ m^{-3}$ and $> 30\ mg\ m^{-3}$, with a smoothing function to avoid a discontinuous transition when $28 < chl_a < 32\ mg\ m^{-3}$:

$$\begin{aligned} & \text{if } [RE10\ chl_a] < 28, [chl_a] = [35.75 \times R^2 - 20.15]^{1.124} \\ & \text{else if } [RE10\ chl_a] > 32, [chl_a] = [35.75 \times R^2 - 20.15]^{1.375} \\ & \text{else } [chl_a] = ([35.75 \times R^2 - 20.15]^{1.124} + [35.75 \times R^2 - 20.15]^{1.375})/2 \end{aligned} \quad (3)$$

We refer to this algorithm as red-edge-San Francisco Bay (RE-SFB).

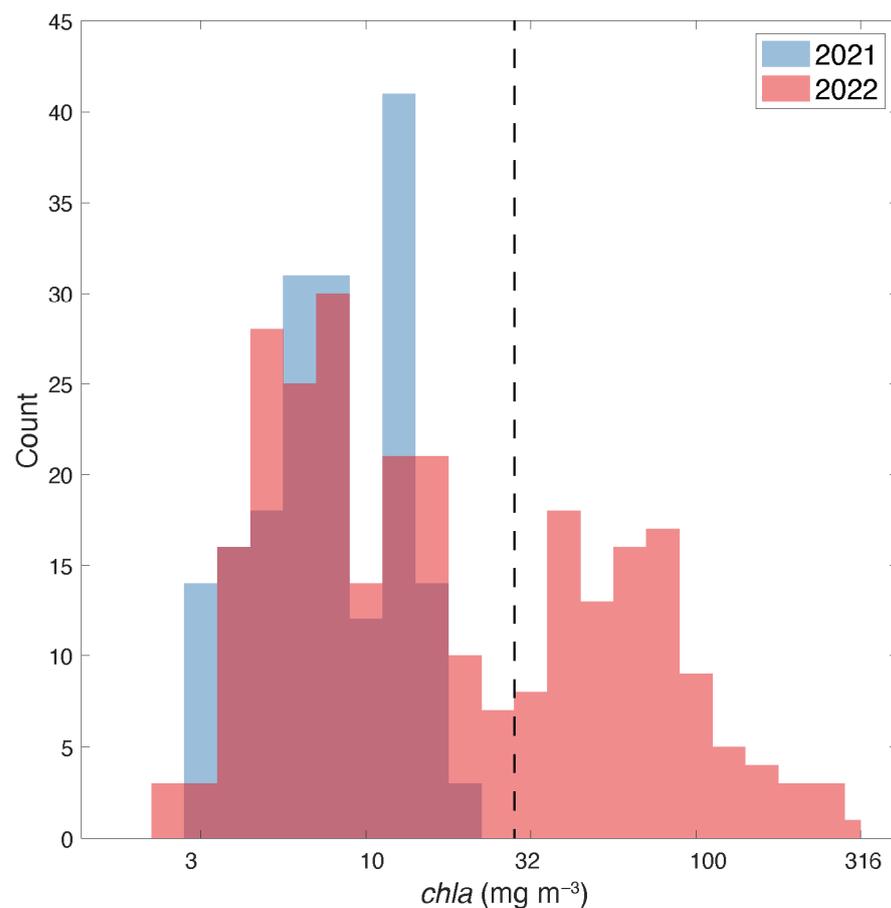


Figure 2. Distribution of field chl_a data from 2021 and 2022 (plotted on a \log_{10} x-axis). The cutoff of $30\ mg\ m^{-3}$ is shown as a vertical dashed line.

2.6. Algorithm Performance

The accuracy of each tested algorithm was evaluated by calculating a series of commonly used statistical methods. These include the coefficient of determination (R^2) for a linear fit, which also provides the slope and intercept for the relationship, plus five other performance parameters. The other methods are the root mean square error (RMSE), the mean absolute error (MAE), the mean bias (MBIAS), the median absolute error (MedAE), and the median bias (MedBIAS). These methods were selected following recommendations outlined by [35] to evaluate satellite and in situ matchups. As noted by [35], MAE is preferable to RMSE, as MAE is less sensitive to the distribution of error magnitudes and sample size; we report RMSE primarily for comparison with previously published methods. The MedAE and MedBIAS capture the typical error, while the MAE and MBIAS indicate whether a method has outlier errors.

The RMSE follows standard statistical formulation, while the MAE and MBIAS were derived following Seegers et al. 2018 [35]:

$$\text{RMSE} = \left[\frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2 \right]^{\frac{1}{2}} \quad (4)$$

$$\text{MAE} = 10^{\left[\frac{1}{N} \sum_{i=1}^N |(\log_{10}(X_i) - (\log_{10}(Y_i)))| \right]} \quad (5)$$

$$\text{MBIAS} = 10^{\left[\frac{1}{N} \sum_{i=1}^N ((\log_{10}(X_i) - (\log_{10}(Y_i))) \right]} \quad (6)$$

where X and Y are the fitted and independent variables, respectively (i.e., the derived *chl*_a and the field validation *chl*_a values). MedBIAS and MedAE were calculated using the same equations as for MBIAS and MAE, with the substitution of median values for mean values. Since the RE-SFB algorithm includes a 2-step fit, statistics were calculated for the ranges 0–30 mg m⁻³ and the full *chl*_a datasets.

3. Results

3.1. Field *chl*_a and Spectral Absorption Data

The *H. akashiwo* bloom resulted in considerably higher *chl*_a levels than typically observed in SFB, which led to the development of the RE-SFB algorithm. Figure 2 provides log-transformed histogram distributions of data from 2021, when there were no blooms, and 2022, during the anomalous bloom event.

The data from 2022 show a clear bimodal distribution with “typical” *chl*_a values < 30 mg m⁻³, similar to 2021, and strongly elevated *chl*_a concentrations associated with the bloom, exceeding 280 mg m⁻³ in surface waters. A value of 30 mg m⁻³ was therefore chosen as the breakpoint in the RE-SFB algorithm, allowing tuning for the unusually high bloom concentrations separate from the more typical *chl*_a concentrations recorded in the region.

The breakpoint (Equation (3)) allows tuning of the algorithm for high biomass that was dominated by a single species during these events. Gilerson et al. [20] optimized RE10 by adjusting the exponential term by ~6.5% from empirically fit data and provided confidence intervals that result in ~50% variability at high (>10 mg m⁻³) *chl*_a concentrations. Those authors noted that the exponential term is sensitive to the chlorophyll-specific absorption (*a**_{ph}) of the dominant biomass. Figure 3 shows the *a**_{ph} spectra from the 2022 *H. akashiwo* bloom versus the spectra used in the original RE10 algorithm (refer to [20] Figure 9 for more details). *H. akashiwo* cultures exhibited much lower *a**_{ph} (675) absorption, which would result in a proportionally higher exponential term when fitted to those data. We therefore used the published exponential term from RE10 below the transition, and the optimized exponential term for *chl*_a > 30 mg m⁻³.

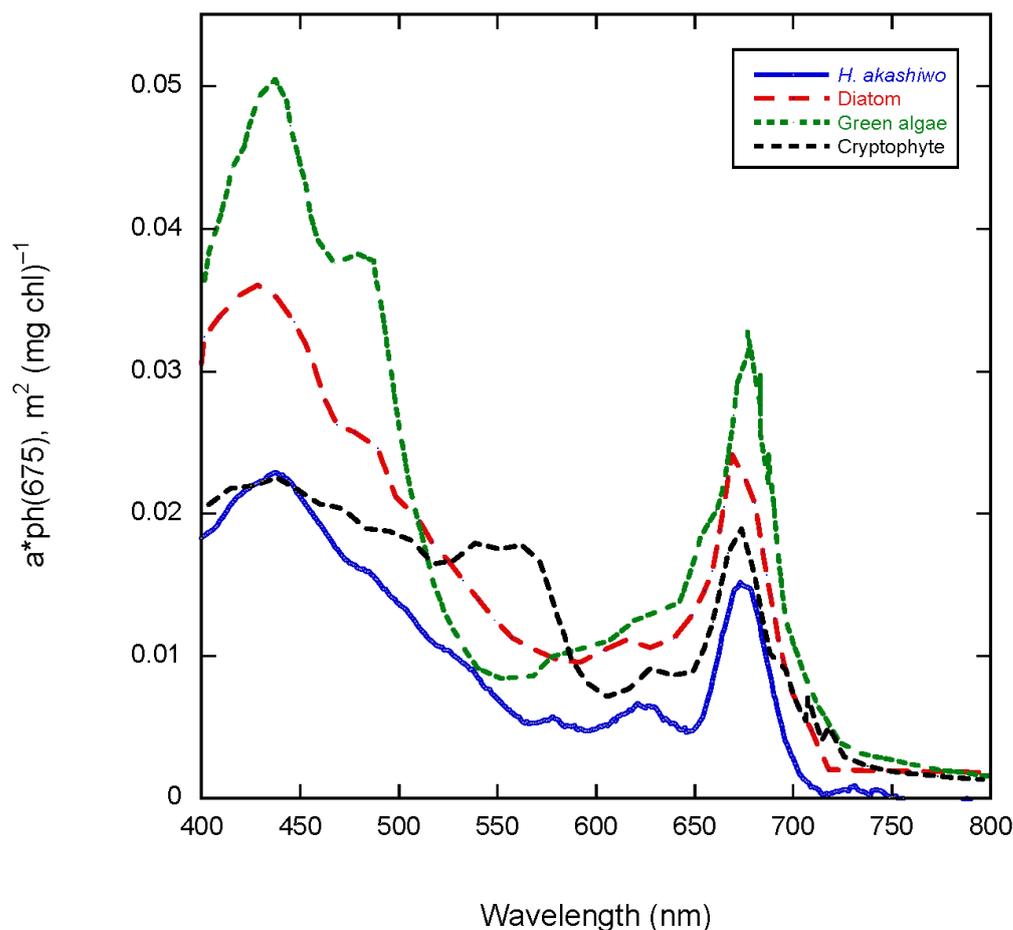


Figure 3. Specific absorption coefficient spectra of cryptophytes (black), diatoms (red), and green algae (green) compared to the raphidophyte *H. akashiwo* (blue), the species that bloomed in the Bay in 2022. The absorption for *H. akashiwo* beyond 750 nm was zero or negative, resulting in a discontinuous line.

3.2. Matchup between Field *chl a* and OC4Me, RE10, RE22, RE-SFB

The matchups between field data (2021–2022) for the Bay and previously published algorithms are shown in Figure 4 (observed versus modeled data) and Figure 5 (quantile–quantile plots of the same data). Three of the algorithms, OC4Me, C2RCC, and RE10, exhibit saturation at high *chl a* concentrations. RE22 is better but tends to underestimate values higher than $\sim 30 \text{ mg m}^{-3} \text{ chl a}$. OC4Me shows the least coherence (more random distribution about the 1:1 line) compared to the other algorithms.

The quantile–quantile plots (Figure 5) present the data with observed and modeled data sorted from lowest to highest values. These plots provide a graphical representation of the fidelity of the satellite algorithms across the range of observed *chl a*. Values falling below the 1:1 line represent underestimates, while values above the 1:1 line represent overestimates. This graphical representation highlights the saturation of OC4Me, C2RCC, and RE10, i.e., above a certain concentration, the data flatline. RE22 and C2RCC show the best performance with good coherence (close to the 1:1 line) but with overestimates at low concentrations and underestimates at high concentrations for RE22. C2RCC is robust at low concentrations (up to $\sim 10 \text{ mg m}^{-3} \text{ chl a}$) but underestimates high concentrations, while OC4Me consistently underestimates *chl a* across the range of observed values, and RE10 is reasonably accurate (but offset high) up to about $30 \text{ mg m}^{-3} \text{ chl a}$.

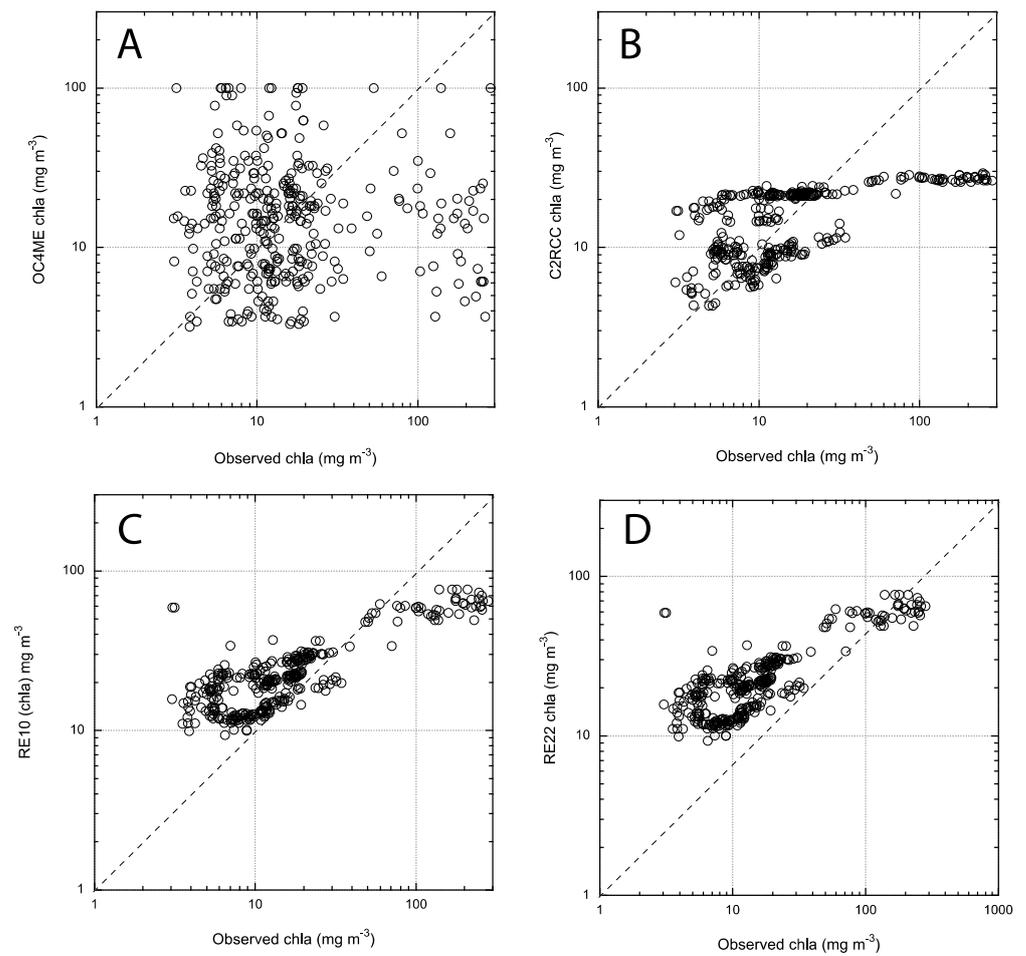


Figure 4. Matchups between field and modeled *chl a* for the standard OC4Me algorithm (A), the C2RCC processor (B), RE10 (C), and RE22 (D). The dashed line indicates the 1:1 fit.

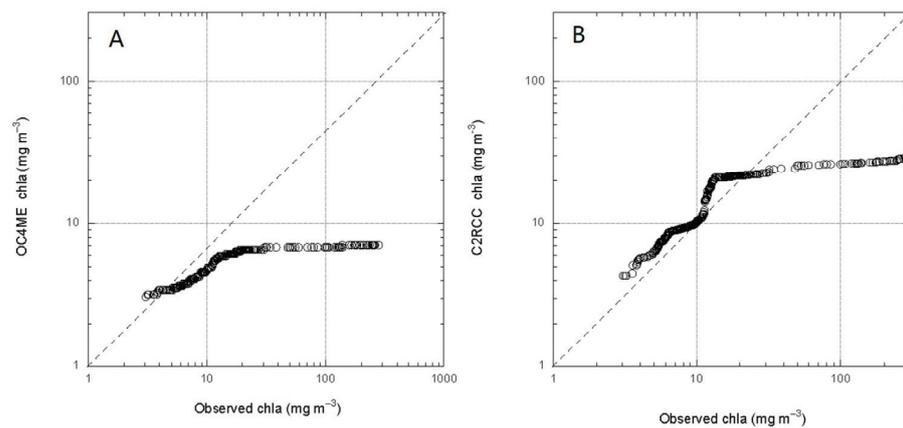


Figure 5. Cont.

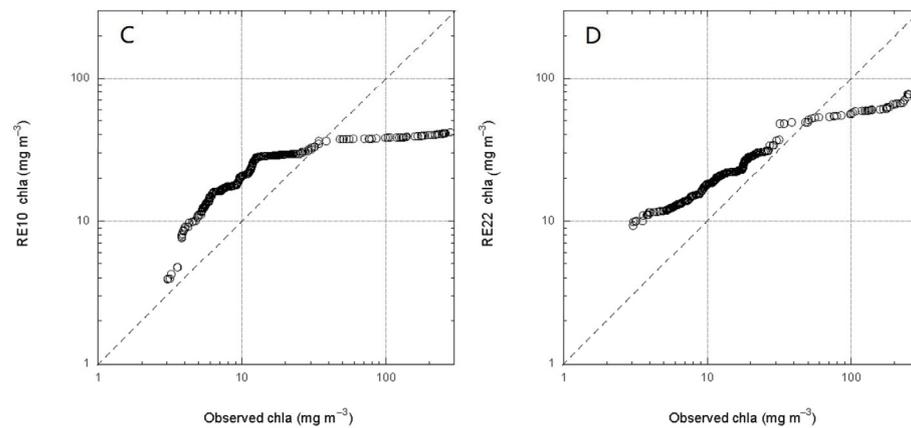


Figure 5. Quantile–quantile plots between field and modeled *chl*_a for the standard OC4Me algorithm (A), the C2RCC processor (B), RE10 (C), and RE22 (D). The dashed line indicates the 1:1 fit.

Figure 6 provides the same plots for the optimized RE-SFB algorithm. Compared to the other algorithms, the fit is qualitatively better, with no saturation at high *chl*_a and with the closest correspondence to the 1:1 line for the quantile–quantile plots. There is slight underestimation at the lowest *chl*_a values and slight to moderate overestimation for values >30 mg m⁻³.

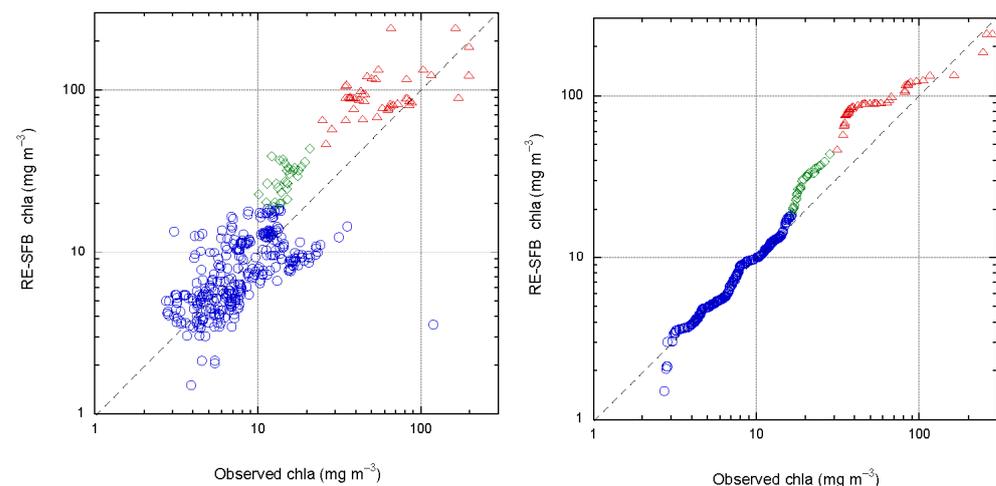


Figure 6. Scatter and quantile–quantile plots between field and modeled *chl*_a for the optimized RE-SFB algorithm; symbols represent <28 (blue circles), 28–32 (green diamonds), and >32 (red triangles) values from RE10, used to partition the data, as described in Equation (3). The dashed line indicates the 1:1 fit.

3.3. Error Metrics for the Algorithms

Summary statistics (described in Section 2.5) are provided in Tables 1–3 for each algorithm, with Table 1 including the full dataset, Table 2 using data up to 30 mg m⁻³, and Table 3 using data greater than 30 mg m⁻³. For the full dataset, RE-SFB consistently outperforms the other algorithms, with low MBIAS and MedBIAS (values close to 1.0). The median bias describes the central tendency of the data, while the mean bias describes whether matchups with large errors are unevenly distributed. The MAE and MedAE provide metrics of overall algorithm performance (MAE for the full range, MedAE for typical *chl*_a concentrations). The MAE and MedAE can be interpreted as approximately the percent error associated with a particular algorithm; for example, the value of MAE = 1.469 for RE-SFB (Table 1) implies that a typical matchup has ~46.9% error when all matchups are considered.

Table 1. Summary statistics for the five algorithms plotted in Figures 4–6 using the full dataset.

Algorithm	n *	R ²	Slope	RMSE	MAE	MBIAS	MedAE	MedBIAS
OC4Me	415	0.615	0.654	0.370	2.235	1.707	2.127	1.840
C2RCC	351	−0.003	−0.012	0.544	3.314	0.870	3.421	1.266
RE10	426	0.319	1.786	0.425	2.316	1.843	2.301	2.084
RE22	426	0.380	1.418	0.381	2.136	1.488	2.074	1.799
RE-SFB	425	0.891	0.865	0.203	1.469	0.964	1.398	0.991

* Variable n is a result of (typically) atmospheric correction failure.

Table 2. Summary statistics for the five algorithms plotted in Figures 4–6 using the data < 30 mg m^{−3}.

Algorithm	n *	R ²	Slope	RMSE	MAE	MBIAS	MedAE	MedBIAS
OC4Me	354	0.019	0.171	0.370	1.934	1.450	1.760	1.760
C2RCC	257	0.001	0.082	0.544	2.815	1.026	2.534	2.534
RE10	359	0.350	0.385	0.425	2.310	1.998	2.408	2.408
RE22	359	0.477	0.390	0.381	2.111	1.636	2.106	2.106
RE-SFB	359	0.489	0.841	0.197	1.460	0.963	1.395	0.927

* Variable n is a result of (typically) atmospheric correction failure.

Table 3. Summary statistics for the five algorithms plotted in Figures 4–6 using the data > 30 mg m^{−3}.

Algorithm	n *	R ²	Slope	RMSE	MAE	MBIAS	MedAE	MedBIAS
OC4Me	61	0.001	0.026	0.984	7.324	0.155	6.470	0.155
C2RCC	94	0.568	0.077	0.666	3.972	0.297	4.702	0.213
RE10	67	0.625	0.067	0.549	3.138	0.432	3.323	0.308
RE22	67	0.771	0.335	0.390	2.172	0.569	2.350	0.458
RE-SFB	66	0.801	0.588	0.233	1.548	1.178	1.412	1.012

* Variable n is a result of (typically) atmospheric correction failure.

An algorithm used operationally should show good performance during typical (non-bloom) conditions as well as during bloom conditions. Table 2 shows that RE-SFB outperforms the other algorithms for this dataset over the typical range (Figure 2) of <30 mg m^{−3} *chl*_a. MBIAS and MedBIAS are close to unity, while the overall fit (MAE) improves slightly compared to the full dataset (Table 1).

Table 3 again shows the best performance at high *chl*_a using RE-SFB, with considerably improved RMSE and bias compared to the other algorithms. Overall, the error metrics provide quantitative evaluation of what can be observed qualitatively in Figures 4–6. The unoptimized algorithms suffer from saturation at high *chl*_a, and both over- and underestimation across the full range of values. RE-SFB minimizes overall error and performs well at both moderate and elevated *chl*_a. Further improvements could be made to optimize any specific subset of the error metrics, but we elected to generally optimize RE-SFB so that it is not over-fitted to a specific bloom event.

3.4. Spatial Distribution of the Bloom

Figure 7 provides an example of the spatial variability of the bloom, as observed using RE22 and RE-SFB. The differences in spatial patterns are consistent with algorithm performance, i.e., RE-SFB reduces *chl*_a relative to RE10 and RE22 below 30 mg m^{−3}, while better representing extreme bloom concentrations for values > 30 mg m^{−3}.

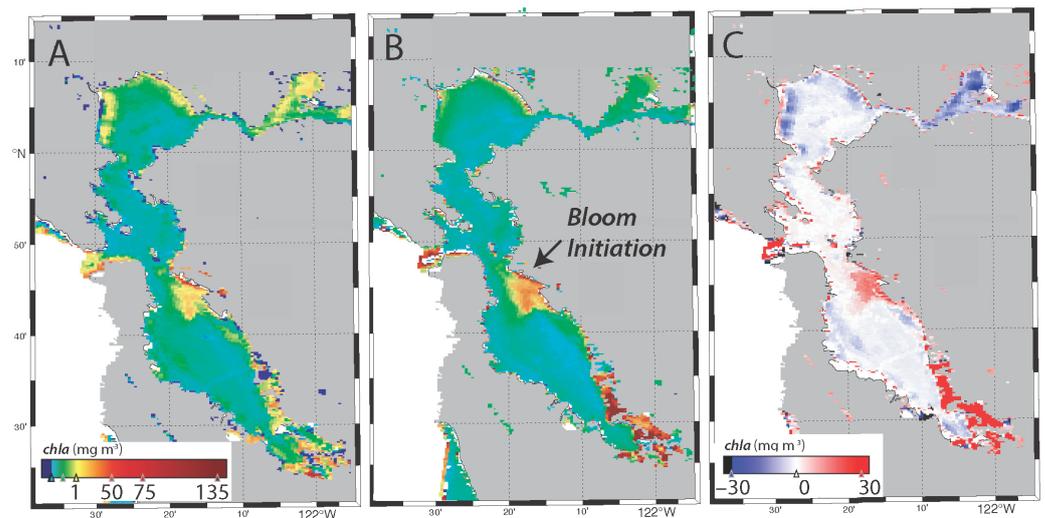


Figure 7. RE22 *chl a* (A) for August 7, 2022, plotted in log *chl a*, RE-SFB (B), and RE-SFB minus RE22 (C) plotted in a linear *chl a* scale. RE-SFB (C) reduces the *chl a* in North Bay, while providing greater delineation (e.g., nearshore intensification) of the initiation site. The bloom was first reported near the “bloom initiation” (B) site in July 2022. The very high *chl a* in the far south is associated with the South Bay salt ponds.

4. Discussion

Development of an accurate *chl a* algorithm for San Francisco Bay is a critical requirement for event response to unusual bloom events, such as the *H. akashiwo* bloom of summer 2022, and routine monitoring of non-bloom conditions. More broadly, *chl a* is commonly used to assess potential impairment of the system [2], and satellite imagery is routinely used for monitoring and assessment of other large estuaries (e.g., [36–39]). Given the optical complexity of estuaries, it is perhaps not surprising that red-edge algorithms generally outperform MBR algorithms such as OC4Me, which are optimized for the global oceans [18]. One caution in using the red-edge family of algorithms is that they perform poorly below *chl a* concentrations of $\sim 1 \text{ mg m}^{-3}$ *chl a* [21]. The lowest observed *chl a* value used for this analysis was 2.73 mg m^{-3} , so no potential matchups were removed, but the RE-SFB algorithm does show increasing bias at low *chl a* levels (Figure 6).

A second potential bias in RE-SFB is that red-edge algorithms are sensitive to the specific *chl a* absorption term (a^*_{ph}). Gilerson 2010 [20] and Wynne et al. 2023 [21] used a fixed exponential term that was generally optimized for coastal organisms, while RE-SFB adjusted that term specifically for the *H. akashiwo* bloom. This may result in non-optimal performance if large blooms of other organisms occur in the future (e.g., *Akashiwo sanguinea*, [12]), although published absorption spectra suggest that *A. sanguinea* exhibits similar $a^*_{ph}(670)$ as *H. akashiwo* [40]. This potential bias is somewhat mitigated with the use of a two-step algorithm, as the values for *chl a* $< 30 \text{ mg m}^{-3}$ are not optimized to a specific organism.

A major difference between OC4Me, C2RCC, and the red-edge algorithms was the atmospheric correction. The standard ESA processing for OC4Me uses a neural network for the atmospheric correction but then switches to the standard MBR algorithm to derive *chl a* (there is also a neural network-derived *chl a* product which was not evaluated). C2RCC is different in that both the atmospheric correction and derived *chl a* product are based on neural networks. Past evaluations of atmospheric correction schemes have demonstrated C2RCC to perform well in optically complex coastal waters [41], as well as in San Francisco Bay [42]. It was therefore surprising that it performed relatively poorly in this analysis with the worst MAE values of the algorithms tested. All of the red-edge algorithms used a variation of the black pixel assumption [29], which subtracts the radiance from the darkest nearby pixel or subtracts the radiance at a presumed zero water-leaving radiance band in

the near-infrared from the other bands. As noted by Wynne et al. 2022 [21], this atmospheric correction scheme tends to retrieve the most pixels in complex coastal and estuarine waters and is fairly insensitive to residual aerosol error in the red and near-infrared bands because ratios of spectrally close bands are used. This is evident in the variable number of successful matchups, with both OC4Me and C2RCC exhibiting failed matchups due to atmospheric correction issues (Table 1). A significant advantage for an operational algorithm is that this atmospheric correction is fairly simple compared to many other algorithms, reducing the processing burden and therefore speeding up data production. More computationally intense processing associated with typical Level 2 ocean color products often results in delays for operational products.

While RE-SFB was developed and tested with a relatively small dataset and two years of data, it is based on the successful application of red-edge algorithms to complex coastal waters. Future work could examine the sensitivity of the a^*_{ph} exponential term (Equation (3)) and/or addition of more matchup data. However, the results presented are robust with a reasonably sized matchup dataset ($n = 425$) across a wide dynamic range of *chl*_a concentrations. The final algorithm provides typical accuracy of $\pm \sim 46$ – 47% based on MAE. While this certainly leaves room for improvement, it is comparable to other operational algorithms in both the Chesapeake Bay (MAE for RE22 is 1.57; [21]), and for analysis of global datasets using standard algorithms (MAE of 1.62–2.05 for eutrophic waters; [35]).

5. Conclusions

This analysis demonstrates the efficacy of red-edge algorithms for the detection of extreme bloom (*chl*_a) events in optically complex coastal waters. Tuning for San Francisco Bay improves performance compared to previous versions [20,21] and benefits from a simplified atmospheric correction scheme, allowing for rapid data processing and dissemination. The use of OLCI provides a reasonable compromise between satellite return rate (\sim daily) and spatial resolution (300 m) for San Francisco Bay.

While higher resolution sensors, such as the Operational Land Imager aboard Landsat-8/9 [43] and the Multi Spectral Instrument (MSI) aboard Sentinel-2, provide greatly improved spatial resolution [43,44], the return rate (~ 7 – 14 days) and lack of appropriate bands for OLCI preclude operational use for events such as the *H. akashiwo* bloom, which evolved over a period of days. Wynne et al. [21] argued for the inclusion of RE10 or similar algorithms as a standard product for the operational NOAA CoastWatch program; this analysis provides further evidence that the broad application of red-edge algorithms to complex coastal waters would greatly benefit end-users.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs16061103/s1>, Figure S1: discrete vs. underway *chl*_a; Figure S2: map showing underway mapping stations used for matchups; Table S1: satellite imagery used in this analysis; Table S2: discrete matchup data; Table S3: High resolution in-water 300 m median binned matchup *chl*_a. Reference [45] is cited in the Supplemental Material.

Author Contributions: Conceptualization, R.M.K.; methodology, R.M.K., D.B.S., B.A.B. and L.S.; software, R.M.K. and L.S.; validation, R.M.K., D.B.S. and B.A.B.; formal analysis, R.M.K.; investigation, R.M.K.; resources, D.B.S., E.T.R. and B.A.B.; data curation, D.B.S., E.T.R. and B.A.B. writing—original draft preparation, R.M.K.; writing—review and editing, R.M.K., D.B.S., E.T.R., K.B.-G., L.S. and B.A.B.; visualization, R.M.K.; supervision, D.B.S.; project administration, D.B.S.; funding acquisition, R.M.K., D.B.S., K.B.-G. and B.A.B. All authors have read and agreed to the published version of the manuscript.

Funding: Partial funding was provided by the San Francisco Bay Nutrient Management Strategy, NOAA Award NA23NOS4780288, U.S. Geological Survey Cooperative Matching Funds, and NASA Award 80NSSC21K1395.

Data Availability Statement: Satellite and validation data are publicly available as cited in the relevant sections.

Acknowledgments: We thank Luis Solorzano and Damon Tighe for early documentation of the event, Hal MacLean (East Bay Parks Recreation District), Janai Southworth, and Jenny Jacox for first identifying the bloom as *H. akashiwo*, and (BayKeeper members) Ian Wren, Jon Rosenfeld, Jon Rosenfeld, Aundi Mevoli, and Julia Dowell for tracking of the event. The *H. akashiwo* culture used for a*ph measurements was kindly provided by Holly Bowers, San Jose State University. We thank the USGS Bay Water Quality Monitoring group, including Jim Cloern, Tara Schraga, Erica Nejad, Taylor Eddy, and Captain Joel Fritsch, for access to the USGS water quality data. We also thank the following individuals, who are among the team that made the USGS/SFEI mapping surveys in the South Bay possible: (USGS) Katy O'Donnell, Ayelet Delascagigas, Dylan Burau, and Crystal Sturgeon; (SFEI) Dan Killam and Ariella Chelsky. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

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

References

1. Conomos, T.J.; Smith, R.E.; Gartner, J.W. Environmental Setting of San Francisco Bay. *Hydrobiologia* **1985**, *129*, 1–12. [\[CrossRef\]](#)
2. Sutula, M.; Kudela, R.; Hagy, J.D.; Harding, L.W.; Senn, D.; Cloern, J.E.; Bricker, S.; Berg, G.M.; Beck, M. Novel Analyses of Long-Term Data Provide a Scientific Basis for Chlorophyll-a Thresholds in San Francisco Bay. *Estuar. Coast. Shelf Sci.* **2017**, *197*, 107–118. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Kudela, R.; Howard, M.; Monismith, S.; Paerl, H. Status, Trends, and Drivers of Harmful Algal Blooms Along the Freshwater-to-Marine Gradient in the San Francisco Bay–Delta System. *San Franc. Estuary Watershed Sci.* **2023**, *20*, 6. [\[CrossRef\]](#)
4. Boesch, D.F.; Brinsfield, R.B.; Magnien, R.E. Chesapeake Bay Eutrophication: Scientific Understanding, Ecosystem Restoration, and Challenges for Agriculture. *J. Environ. Qual.* **2001**, *30*, 303–320. [\[CrossRef\]](#)
5. Bricker, S.B.; Longstaff, B.; Dennison, W.; Jones, A.; Boicourt, K.; Wicks, C.; Woerner, J. Effects of Nutrient Enrichment in the Nation's Estuaries: A Decade of Change. *Harmful Algae* **2008**, *8*, 21–32. [\[CrossRef\]](#)
6. Nixon, S.W. Coastal Marine Eutrophication: A Definition, Social Causes, and Future Concerns. *Ophelia* **1995**, *41*, 199–219. [\[CrossRef\]](#)
7. Diaz, R.J.; Rosenberg, R. Spreading Dead Zones and Consequences for Marine Ecosystems. *Science* **2008**, *321*, 926–929. [\[CrossRef\]](#)
8. Heisler, J.; Glibert, P.M.; Burkholder, J.M.; Anderson, D.M.; Cochlan, W.; Dennison, W.C.; Dortch, Q.; Gobler, C.J.; Heil, C.A.; Humphries, E.; et al. Eutrophication and Harmful Algal Blooms: A Scientific Consensus. *Harmful Algae* **2008**, *8*, 3–13. [\[CrossRef\]](#)
9. Cloern, J. Does the Benthos Control Phytoplankton Biomass in South San Francisco Bay? *Mar. Ecol. Prog. Ser.* **1982**, *9*, 191–202. [\[CrossRef\]](#)
10. Cloern, J.E. Phytoplankton Bloom Dynamics in Coastal Ecosystems: A Review with Some General Lessons from Sustained Investigation of San Francisco Bay, California. *Rev. Geophys.* **1996**, *34*, 127–168. [\[CrossRef\]](#)
11. Jessup, D.A.; Miller, M.A.; Ryan, J.P.; Nevins, H.M.; Kerkering, H.A.; Mekebr, A.; Crane, D.B.; Johnson, T.A.; Kudela, R.M. Mass Stranding of Marine Birds Caused by a Surfactant-Producing Red Tide. *PLoS ONE* **2009**, *4*, e4550. [\[CrossRef\]](#)
12. Cloern, J.E.; Schraga, T.S.; Lopez, C.B.; Knowles, N.; Grover Labiosa, R.; Dugdale, R. Climate Anomalies Generate an Exceptional Dinoflagellate Bloom in San Francisco Bay. *Geophys. Res. Lett.* **2005**, *32*, L14608. [\[CrossRef\]](#)
13. Flores-Leñero, A.; Vargas-Torres, V.; Paredes-Mella, J.; Norambuena, L.; Fuenzalida, G.; Lee-Chang, K.; Mardones, J.I. Heterosigma Akashiwo in Patagonian Fjords: Genetics, Growth, Pigment Signature and Role of PUFA and ROS in Ichthyotoxicity. *Toxins* **2022**, *14*, 577. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Herndon, J.; Cochlan, W.P. Nitrogen Utilization by the Raphidophyte *Heterosigma akashiwo*: Growth and Uptake Kinetics in Laboratory Cultures. *Harmful Algae* **2007**, *6*, 260–270. [\[CrossRef\]](#)
15. Herndon, J. *Nitrogen Uptake by the Raphidophyte Heterosigma akashiwo: A Laboratory and Field Study*; M.A., San Francisco State University: San Francisco, CA, USA, 2003.
16. Fichot, C.G.; Downing, B.D.; Bergamaschi, B.A.; Windham-Myers, L.; Marvin-DiPasquale, M.; Thompson, D.R.; Gierach, M.M. High-Resolution Remote Sensing of Water Quality in the San Francisco Bay–Delta Estuary. *Environ. Sci. Technol.* **2016**, *50*, 573–583. [\[CrossRef\]](#)
17. Pahlevan, N.; Smith, B.; Binding, C.; Gurlin, D.; Li, L.; Bresciani, M.; Giardino, C. Hyperspectral Retrievals of Phytoplankton Absorption and Chlorophyll-a in Inland and Nearshore Coastal Waters. *Remote Sens. Environ.* **2021**, *253*, 112200. [\[CrossRef\]](#)
18. O'Reilly, J.E.; Werdell, P.J. Chlorophyll Algorithms for Ocean Color Sensors—OC4, OC5 & OC6. *Remote Sens. Environ.* **2019**, *229*, 32–47. [\[CrossRef\]](#)
19. Moses, W.J.; Gitelson, A.A.; Berdnikov, S.; Saprygin, V.; Povazhnyi, V. Operational MERIS-Based NIR-Red Algorithms for Estimating Chlorophyll-a Concentrations in Coastal Waters—The Azov Sea Case Study. *Remote Sens. Environ.* **2012**, *121*, 118–124. [\[CrossRef\]](#)
20. Gilerson, A.A.; Gitelson, A.A.; Zhou, J.; Gurlin, D.; Moses, W.; Ioannou, I.; Ahmed, S.A. Algorithms for Remote Estimation of Chlorophyll-a in Coastal and Inland Waters Using Red and near Infrared Bands. *Opt. Express* **2010**, *18*, 24109. [\[CrossRef\]](#)

21. Wynne, T.T.; Tomlinson, M.C.; Briggs, T.O.; Mishra, S.; Meredith, A.; Vogel, R.L.; Stumpf, R.P. Evaluating the Efficacy of Five Chlorophyll-a Algorithms in Chesapeake Bay (USA) for Operational Monitoring and Assessment. *J. Mar. Sci. Eng.* **2022**, *10*, 1104. [[CrossRef](#)]
22. Tran, M.D.; Vantrepotte, V.; Loisel, H.; Oliveira, E.N.; Tran, K.T.; Jorge, D.; Mériaux, X.; Paranhos, R. Band Ratios Combination for Estimating Chlorophyll-a from Sentinel-2 and Sentinel-3 in Coastal Waters. *Remote Sens.* **2023**, *15*, 1653. [[CrossRef](#)]
23. Wolny, J.L.; Tomlinson, M.C.; Schollaert Uz, S.; Egerton, T.A.; McKay, J.R.; Meredith, A.; Reece, K.S.; Scott, G.P.; Stumpf, R.P. Current and Future Remote Sensing of Harmful Algal Blooms in the Chesapeake Bay to Support the Shellfish Industry. *Front. Mar. Sci.* **2020**, *7*, 337. [[CrossRef](#)]
24. Rodríguez-Benito, C.V.; Navarro, G.; Caballero, I. Using Copernicus Sentinel-2 and Sentinel-3 Data to Monitor Harmful Algal Blooms in Southern Chile during the COVID-19 Lockdown. *Mar. Pollut. Bull.* **2020**, *161*, 111722. [[CrossRef](#)]
25. Jordan, C.; Cusack, C.; Tomlinson, M.C.; Meredith, A.; McGeedy, R.; Salas, R.; Gregory, C.; Croot, P.L. Using the Red Band Difference Algorithm to Detect and Monitor a *Karenia* Spp. Bloom Off the South Coast of Ireland, June 2019. *Front. Mar. Sci.* **2021**, *8*, 638889. [[CrossRef](#)]
26. Kahru, M.; Anderson, C.; Barton, A.D.; Carter, M.L.; Catlett, D.; Send, U.; Sosik, H.M.; Weiss, E.L.; Mitchell, B.G. Satellite Detection of Dinoflagellate Blooms off California by UV Reflectance Ratios. *Elem. Sci. Anthr.* **2021**, *9*, 00157. [[CrossRef](#)]
27. Windle, A.E.; Evers-King, H.; Loveday, B.R.; Ondrusek, M.; Silsbe, G.M. Evaluating Atmospheric Correction Algorithms Applied to OLCI Sentinel-3 Data of Chesapeake Bay Waters. *Remote Sens.* **2022**, *14*, 1881. [[CrossRef](#)]
28. Brockmann, C.; Doerffer, R.; Peters, M.; Stelzer, K.; Embacher, S.; Ruescas, A. Evolution of the C2RCC Neural Network for Sentinel 2 and 3 for the Retrieval of Ocean Color Products in Normal and Extreme Optically Complex Waters. In Proceedings of the Living Planet Symposium 2016, Prague, Czech Republic, 9 May 2016; European Space Agency Special Publication: Prague, Czech Republic, 2016; Volume ESA SP, pp. 1–6.
29. Stumpf, R.P.; Tyler, M.A. Satellite Detection of Bloom and Pigment Distributions in Estuaries. *Remote Sens. Environ.* **1988**, *24*, 385–404. [[CrossRef](#)]
30. Schraga, T.S.; Cloern, J.E. Water Quality Measurements in San Francisco Bay by the U.S. Geological Survey, 1969–2015. *Sci. Data* **2017**, *4*, 170098. [[CrossRef](#)]
31. Richardson, E.T.; O'Donnell, K.; Soto Perez, J.; Sturgeon, C.L.; Brinkman, J.; Delascagigas, A.; Nakatsuka, K.; Uebner, M.Q.; Jaegge, A.; Dellwo, P.; et al. *Assessing Spatial Variability of Nutrients, Phytoplankton and Related Water-Quality Constituents in the San Francisco Bay, California: 2021–2022 High-Resolution Mapping Surveys*; U.S. Geological Survey data release; U.S. Geological Survey: Reston, WV, USA, 2024. [[CrossRef](#)]
32. Schraga, T.; Nejad, E.S.; Martin, C.A.; Cloern, J.E. *USGS Measurements of Water Quality in San Francisco Bay (CA), 2016–2021 (Ver. 4.0, March 2023)*; U.S. Geological Survey Data Release; U.S. Geological Survey: Reston, WV, USA, 2018. [[CrossRef](#)]
33. Morel, A.; Huot, Y.; Gentili, B.; Werdell, P.J.; Hooker, S.B.; Franz, B.A. Examining the Consistency of Products Derived from Various Ocean Color Sensors in Open Ocean (Case 1) Waters in the Perspective of a Multi-Sensor Approach. *Remote Sens. Environ.* **2007**, *111*, 69–88. [[CrossRef](#)]
34. Stramski, D.; Reynolds, R.A.; Kaczmarek, S.; Uitz, J.; Zheng, G. Correction of Pathlength Amplification in the Filter-Pad Technique for Measurements of Particulate Absorption Coefficient in the Visible Spectral Region. *Appl. Opt.* **2015**, *54*, 6763. [[CrossRef](#)]
35. Seegers, B.N.; Stumpf, R.P.; Schaeffer, B.A.; Loftin, K.A.; Werdell, P.J. Performance Metrics for the Assessment of Satellite Data Products: An Ocean Color Case Study. *Opt. Express* **2018**, *26*, 7404. [[CrossRef](#)] [[PubMed](#)]
36. Le, C.; Hu, C.; English, D.; Cannizzaro, J.; Kovach, C. Climate-Driven Chlorophyll-a Changes in a Turbid Estuary: Observations from Satellites and Implications for Management. *Remote Sens. Environ.* **2013**, *130*, 11–24. [[CrossRef](#)]
37. Chen, J.; Pan, D.; Liu, M.; Mao, Z.; Zhu, Q.; Chen, N.; Zhang, X.; Tao, B. Relationships between Long-Term Trend of Satellite-Derived Chlorophyll-a and Hypoxia Off the Changjiang Estuary. *Estuaries Coasts* **2017**, *40*, 1055–1065. [[CrossRef](#)]
38. Keith, D.J. Satellite Remote Sensing of Chlorophyll a in Support of Nutrient Management in the Neuse and Tar–Pamlico River (North Carolina) Estuaries. *Remote Sens. Environ.* **2014**, *153*, 61–78. [[CrossRef](#)]
39. Conmy, R.N.; Schaeffer, B.A.; Schubauer-Berigan, J.; Aukamp, J.; Duffy, A.; Lehrter, J.C.; Greene, R.M. Characterizing Light Attenuation within Northwest Florida Estuaries: Implications for RESTORE Act Water Quality Monitoring. *Mar. Pollut. Bull.* **2017**, *114*, 995–1006. [[CrossRef](#)]
40. Kim, Y.; Yoo, S.; Son, Y.B. Optical Discrimination of Harmful *Cochlodinium polykrikoides* Blooms in Korean Coastal Waters. *Opt. Express* **2016**, *24*, A1471. [[CrossRef](#)] [[PubMed](#)]
41. Mognane, M.; Jamet, C.; Loisel, H.; Vantrepotte, V.; Mériaux, X.; Cauvin, A. Evaluation of Five Atmospheric Correction Algorithms over French Optically-Complex Waters for the Sentinel-3A OLCI Ocean Color Sensor. *Remote Sens.* **2019**, *11*, 668. [[CrossRef](#)]
42. Taylor, N.C.; Kudela, R.M. Spatial Variability of Suspended Sediments in San Francisco Bay, California. *Remote Sens.* **2021**, *13*, 4625. [[CrossRef](#)]
43. Franz, B.A.; Bailey, S.W.; Kuring, N.; Werdell, P.J. Ocean Color Measurements with the Operational Land Imager on Landsat-8: Implementation and Evaluation in SeaDAS. *J. Appl. Remote Sens.* **2015**, *9*, 096070. [[CrossRef](#)]

-
44. Bramich, J.; Bolch, C.J.S.; Fischer, A. Improved Red-Edge Chlorophyll-a Detection for Sentinel 2. *Ecol. Indic.* **2021**, *120*, 106876. [[CrossRef](#)]
 45. U.S. Geological Survey. USGS Water Data for the Nation: U.S. Geological Survey National. Water Information System Database. 2024. Available online: <https://waterdata.usgs.gov/nwis> (accessed on 8 March 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.