Journal of *Marine Science and Engineering*

MDPI

Article

Regional Algorithm of Quantitative Assessment of Cyanobacteria Blooms in the Eastern Part of the Gulf of Finland Using Satellite Ocean Color Data

Svetlana Vazyulya [1], Oleg Kopelevich [1], Inna Sahling [1], Ekaterina Kochetkova [2], Evgenia Lange [1], Alexander Khrapko [1], Tatyana Eremina [2] and Dmitry Glukhovets [1,3,*]

[1] Shirshov Institute of Oceanology of the Russian Academy of Sciences, 117997 Moscow, Russia; vasyulya.sv@ocean.ru (S.V.)
[2] Russian State Hydrometeorological University, 195196 St. Petersburg, Russia
[3] Moscow Institute of Physics and Technology, 141700 Dolgoprudny, Russia
* Correspondence: glukhovets@ocean.ru; Tel.: +7-(499)-124-75-83

Abstract: Summer blooms of potentially harmful cyanobacteria are common in the Baltic Sea. Under clear sky conditions, the cyanobacterial blooms are easily detectable from space. We propose a new regional algorithm for cyanobacteria biomass estimation from satellite data in the eastern part of the Gulf of Finland, developed on the basis of field measurements in July–August 2012–2014. The multi-regression equation defines the cyanobacteria biomass as a function of the particle backscattering coefficient and chlorophyll concentration. The use of this equation provides the best performance in comparison to the linear one, which is reflected in both R^2 and RMSE values (0.61 and 272 mg m^{-3} respectively). Unlike other algorithms, which determine only the cyanobacteria bloom area in the Baltic Sea, our algorithm allows the determination of both a bloom area and its intensity. Considering the algorithm errors, the bloom detection threshold has been shifted from the 200 mg m^{-3} determined by biologists to 300 mg m^{-3}. Based on data from the MODIS-Aqua satellite ocean color scanner, the spatial and temporal variability of cyanobacterial blooms in this region from 2003 to 2022 was analyzed. Significant interannual variability of cyanobacteria biomass was revealed in the central part of the studied region, with minimum values in 2014 and maximum in 2004. The record bloom during the studied period occurred in July 2004 (the average cyanobacteria biomass was 780 mg m^{-3}). The weakest blooms were observed in 2009, 2010, and 2014, when both in July and August, the bloom areas did not exceed 30% of the study region.

Keywords: cyanobacteria bloom; Baltic Sea; MODIS; particle backscattering coefficient; chlorophyll concentration

Citation: Vazyulya, S.; Kopelevich, O.; Sahling, I.; Kochetkova, E.; Lange, E.; Khrapko, A.; Eremina, T.; Glukhovets, D. Regional Algorithm of Quantitative Assessment of Cyanobacteria Blooms in the Eastern Part of the Gulf of Finland Using Satellite Ocean Color Data. *J. Mar. Sci. Eng.* **2023**, *11*, 1746. https://doi.org/10.3390/jmse11091746

Academic Editor: Juan José Dorantes-Aranda

Received: 31 July 2023
Revised: 29 August 2023
Accepted: 2 September 2023
Published: 5 September 2023

1. Introduction

Massive blooms of cyanobacteria (also known as blue-green algae) affect the Baltic Sea almost every summer. However, there is growing evidence that they are likely to expand even further in the coming decades due to ongoing eutrophication, rising atmospheric CO_2 concentrations, and global warming [1]. On cloudless days, satellite sensors capture these blooms well [2–12].

Cyanobacteria biomass (Bcyan) is one of the main indicators for the Baltic Sea eutrophication assessment [13]. The phytoplankton monitoring of the HELCOM COMBINE (Cooperative Monitoring in the Baltic Marine Environment) program includes in situ Bcyan measurements carried out during the summer months (June–August) for the entire Baltic Sea. An annual report about cyanobacteria biomass is published on the HELCOM website in the Baltic Sea Environment Fact Sheet section [14]. Note that during the period 1990–2019, throughout the Baltics, the highest values of Bcyan were observed in the Gulf of Finland. In 2004, Bcyan in the Gulf of Finland broke the record and reached 7470 mg m^{-3}.

Satellite data along with in situ data are actively used to monitor the cyanobacteria blooms in the Baltic Sea. Since 2002, from June to August (or longer when blooms continue), maps of cyanobacteria blooms in the Baltic Sea appear daily on the website of the Swedish Meteorological and Hydrological Institute (https://www.smhi.se/en/weather/sweden-weather/the-algae-situation, accessed on 4 September 2023), and an annual report has been published on the HELCOM website [15] since 2004. However, the algorithms for the operational assessment of cyanobacterial blooms based on satellite data determine only its areas [4–7], whereas the number of bloom days or pixels marked determines the intensity of bloom.

There are several methods for estimating phycocyanin concentration (PC) as an indicator of cyanobacterial blooms [11,12,16,17], which can be applied to satellite ocean color data such as MERIS (Envisat) or OLCI (Sentinel-3). MERIS has several spectral bands in the red and near-infrared that allow spectral shape algorithms to target strong blooms [18,19]. Forms of these algorithms for MERIS data include Fluorescence Line Height (FLH) [20], the Maximum Chlorophyll Index (MCI) [18], and the Cyanobacteria Index (CI) [19]. MODIS has a band similar to MERIS, centered at 678 nm, which suggests the possible application of an equivalent algorithm to calculate the CI [21]. The analysis of CI for MODIS data showed that the CI_{MODIS} identifies the same blooms detected with CI_{MERIS}, with similar spatial patterns [22]. Hu developed a novel ocean color index, the Floating Algae Index (FAI), and used it to detect floating algae in open ocean using medium-resolution (250 and 500 m) data from MODIS [23]. The PC and MCI algorithms have been tested to monitor cyanobacterial blooms in the Baltic Sea [2,3,11,12].

Here, we present a regional algorithm for estimation of the cyanobacteria biomass based on satellite data in the eastern part of the Gulf of Finland, developed on the basis of in situ measurements in July–August 2012–2014. The novelty of our algorithm lies in the ability to determine not only a bloom area, but also its intensity. The article has two goals: to evaluate the accuracy of the new regional algorithm and to study the interannual variability of Bcyan in selected regions using satellite data and the developed algorithm. The MODIS-Aqua satellite ocean color scanner data made it possible to analyze the spatial and temporal variability of Bcyan in this region from 2003 to 2022.

2. Materials and Methods

2.1. In Situ Data

Three summer expeditions (from late July to early August 2012–2014) aboard the yacht Centaurus II in the eastern part of the Gulf of Finland between the islands of Kotlin and Gogland provided the in situ data for the study. The in situ dataset includes direct determinations of the total biomass of phytoplankton and cyanobacteria in the surface layer, the concentration of chlorophyll a (Chl), as well as the subsurface radiance reflectance $\rho(\lambda)$ in the visual spectral range. To develop the algorithm, data from 28 stations were used, where the Bcyan value exceeded 20 mg m^{-3} (Figure 1). At some stations near the Neva Bay, the biomass of cyanobacteria exceeded 1000 mg m^{-3}.

Phytoplankton biomass and chlorophyll samples, collected with a Ruttner water sampler, represent integral values in the euphotic zone, defined as a tripled depth of the Secchi disk. The phytoplankton samples (0.5 L) were fixed in Lugol's iodine solution with an admixture of formaldehyde and acetic acid. Chlorophyll concentration was measured by filtering 500–1000 mL of water through a membrane filter with a pore diameter of about 0.8 μm. The concentration of chlorophyll a was determined in acetone extraction by the spectrophotometric technique recommended by UNESCO [24].

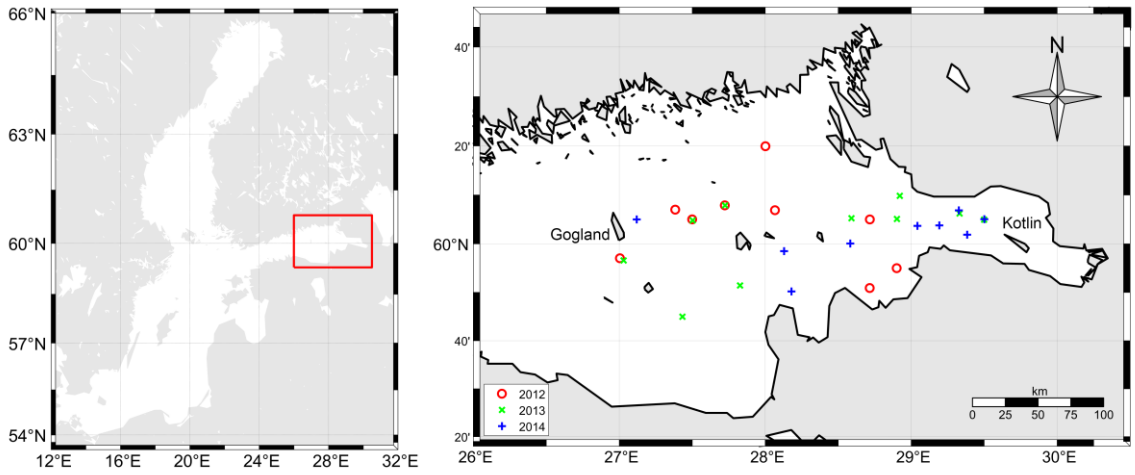

Figure 1. Location of stations in the eastern part of the Gulf of Finland.

A floating spectroradiometer was used to measure the subsurface spectral radiance reflectance ρ(λ) [25]. The spectral measurement range is 390–700 nm, the band resolution is 2 nm. The measurement accuracy is about 5%. The spectra of ρ(λ) allow the use of regional algorithms to calculate such bio-optical parameters as particle backscattering coefficient, chlorophyll, and suspended matter concentrations; in addition, these data give us an opportunity to estimate atmospheric correction errors [26].

2.2. Satellite Data

We used the MODIS-Aqua satellite spectroradiometer data available on the NASA website (http://oceancolor.gsfc.nasa.gov, accessed on 4 September 2023). For the evaluation of cyanobacteria blooms, the values of the spectral remote sensing reflectance $R_{rs}(\lambda)$ with a spatial resolution of 1 km (Level L2) were used. The spectral subsurface radiance reflectance ρ(λ), introduced above, is related to $R_{rs}(\lambda)$ by the simple formula (Equation (25) in [27]).

The MODIS-Aqua data used for the spatial and temporal variability of cyanobacteria biomass cover almost the same region as the in situ measurements: it is bounded by 26° E in the west and by the line of a storm-surge barrier (dam) in the east. The monthly average (July and August) and seasonal distribution of cyanobacteria biomass for 2003–2016 for the selected region was calculated. The calculations were limited only to July and August, as these months are characterized by intense blooms, while the June blooms are irregular and much weaker. The resulting monthly and seasonal Level L3 distributions were obtained by averaging the Level 2 distribution over a given of 2 × 2 km bin and filtering out the pixels flagged as LAND, CLDICE, and STRAYLIGHT, the latter flag allows the elimination of errors related to the edge of the cloud and the proximity of the coast.

2.3. Accuracy Assessment Metrics

The statistical metrics for the Bcyan models were the coefficient of determination R^2, the root-mean-square error RMSE, the coefficient of variation CV. RMSE and CV were calculated using the formulas:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i^{mod} - y_i^{meas}\right)^2}, \quad (1)$$

$$\mathrm{CV} = 100\% \times \mathrm{RMSE} / \left(\frac{1}{N}\sum_{i=1}^{N} y_i^{meas}\right). \quad (2)$$

where y_i^{meas} is the i^{th} measured and y_i^{mod} is i^{th} modelled value. We also calculated the Ratio and median percent difference (MPD) for measured and modelled values:

$$\text{Ratio} = \text{median}\left(\frac{y_i^{mod}}{y_i^{meas}}\right), \quad (3)$$

$$\text{MPD} = \text{median}\left(100\% \cdot \left|\frac{y_i^{mod} - y_i^{meas}}{y_i^{meas}}\right|\right). \quad (4)$$

2.4. Development of a Regional Algorithm for Estimating the Cyanobacteria Biomass

Kahru et al. proposed an algorithm [6,7] for the estimation of blooming areas in the Baltic Sea. It is based on the use of satellite R_{rs} values (667 nm channel for MODIS-Aqua), where a pixel is labelled with a "blooming" mark when the R_{rs} value exceeds the given threshold. Figure 2 presents examples of $R_{rs}(\lambda)$ spectra obtained using the formula [27] from the $\rho(\lambda)$ spectra measured with the floating spectroradiometer in 2014. Indeed, for cyanobacterial bloom stations, where Bcyan exceeded 200 mg m^{-3}, the $R_{rs}(667)$ values exceed the threshold value of 0.0012; moreover, the $R_{rs}(667)$ values are commensurate with Bcyan values. There are features of the $R_{rs}(\lambda)$ spectra, previously detected from the data of 2012 and 2013 [26] that are especially evident for the 1F5 station, notable for the record Bcyan value of 2065 mg m^{-3}. The R_{rs} maximum near 650 nm between the two minima at 620 and 680 nm is presumably associated with phycocyanin fluorescence at 650 nm, while a sharp minimum at 620 nm is related to the phycocyanin absorption maximum. Note that phycocyanin pigment makes it possible to identify the presence of cyanobacteria [28].

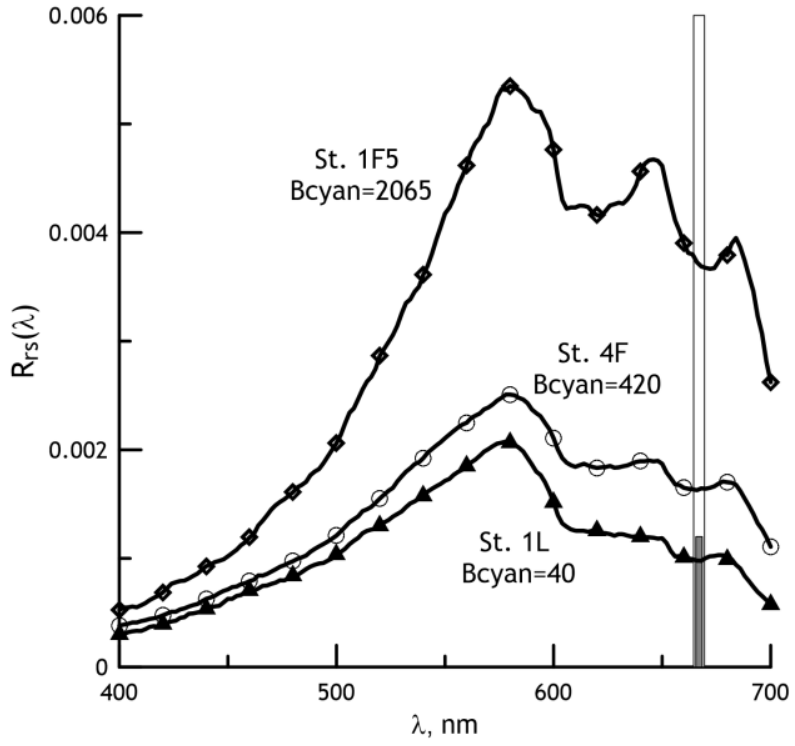

Figure 2. Examples of the spectral remote sensing reflectance $R_{rs}(\lambda)$ derived from the $\rho(\lambda)$ spectra measured with the floating spectroradiometer in 2014. The grey bar shows the threshold $R_{rs}(667)$ value for determining the cyanobacteria bloom according to the algorithm [7].

The main disadvantage of the Kahru algorithm is that it determines only the bloom area and it does not allow the estimation of the intensity of bloom. The proposed algorithm for estimating the biomass of cyanobacteria from satellite data makes it possible to

determine both the area and the intensity of bloom. Two bio-optical products were used to create this algorithm; the algorithms for determining which satellite data were previously developed by us. Instead of the $R_{rs}(667)$ values, we used the particle backscattering coefficient b_{bp}, which is already "free" from the absorption influence. Another option is to use the chlorophyll *a* concentration value; according to [29], the correlation coefficient *R* between the average values of Bcyan and Chl, calculated from field measurements of 2004–2011, is 0.81.

Table 1 presents the results of testing the above options, as well as the use of the multi-regression equation to determine Bcyan from b_{bp} and Chl. In these calculations, the in situ measured Chl were used, and the b_{bp} values were calculated from the in situ spectra of subsurface radiance reflectance ρ(λ); a brief description of the b_{bp} calculation algorithm is available on the website (http://optics.ocean.ru, accessed on 4 September 2023) and in more detail, see [30]. The multi-regression equation provides the best results with the highest R^2 values and the smallest deviations between the modelled and measured Bcyan values; therefore, it was chosen for all subsequent Bcyan calculations. The workflow of the algorithm is presented in Figure 3.

Table 1. Comparison of algorithms for cyanobacteria biomass calculations.

Algorithm	R^2	RMSE, mg m^{-3}	CV, %	Ratio	MPD, %
b_{bp} based: Bcyan = 64.6 b_{bp} 10^3 − 263	0.55	282	75	1.5	60
Chl based: Bcyan = 92 Chl + 6	0.50	304	81	1.1	52
Multi-regression: Bcyan = 45 b_{bp} 10^3 + 38.5 Chl − 227	0.61	272	72	1.4	54

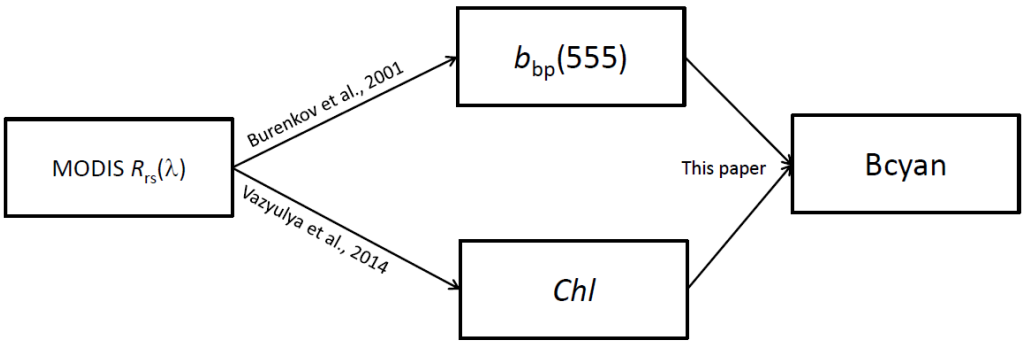

Figure 3. The workflow of the developed algorithm, which uses algorithms [26,30].

Figure 4a shows a comparison of the Bcyan values calculated from the field data (Bcyan model) and measured (Bcyan measured). The regression algorithm provides satisfactory agreement with the measured Bcyan values. The mean values for the measured and model values of Bcyan are in close agreement and equal to 378 mg m^{-3}. The Ratio for all 28 stations is 1.4. However, one can see in Figure 4a that overestimation of Bcyan occurs mainly for stations where Bcyan values were less than 200 mg m^{-3}, where there was no bloom [31]. Indeed, for 17 stations, where Bcyan > 200 mg m^{-3}, the Ratio is 0.9, and for 11 stations with Bcyan < 200 mg m^{-3}, it is already 2.4.

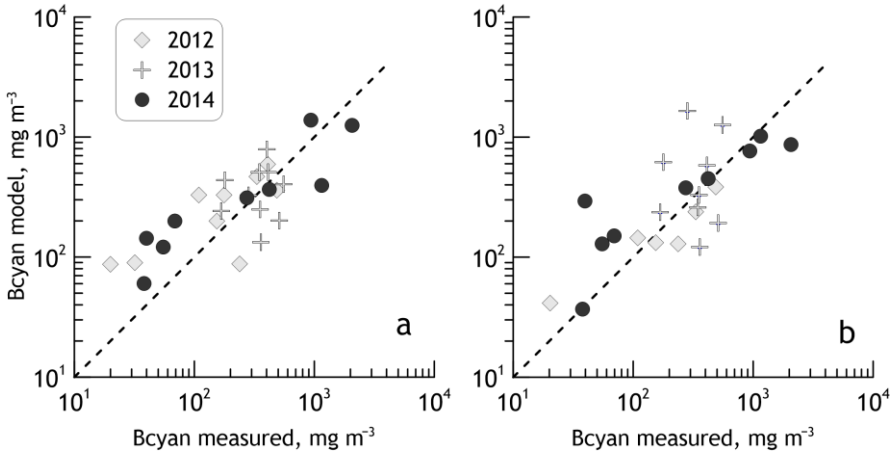

Figure 4. Comparison of the Bcyan values (mg m^{-3}): calculated (Bcyan model) versus measured (Bcyan measured) in 2012–2014: (**a**) with field data; (**b**) with satellite data. The dashed line shows the perfect agreement.

3. Results

3.1. Validation of the Algorithm with Satellite Data

To validate the algorithm on satellite data, the average distributions of Bcyan for the periods of the expedition (26–30 July 2012; 24 July–2 August 2013; 25 July–2 August 2014) were generated using the multiple regression formula. The concentration of chlorophyll was calculated using a regional algorithm [26]. Thus, it was possible to obtain Bcyan estimates for 24 stations from MODIS-Aqua data. The comparison of Bcyan values measured and calculated from satellite data is shown in Figure 4b. As expected, the correspondence between measured Bcyan and model Bcyan using satellite data is worse than for the field dataset (Figure 4a). However, the resulting discrepancy is not very large. The coefficient of determination R^2 is 0.53, the RMSE is 427 mg m^{-3}, the averages for the measured and model Bcyan are 398 and 434 mg m^{-3}, respectively. The Ratio for all 24 stations is 0.87. At the same time, as in the case of field dataset (Figure 4a), the Ratio is 0.75 for 15 stations with an evident bloom, and it is equal to 2.1 for nine stations without bloom.

When satellite data were used, deterioration in the correspondence between the model Bcyan and the measured ones was expected. Most likely it is due to the averaging. In our calculation, the satellite data were averaged both over space (the size of a pixel of the input L2 level is about 1 km, and at the L3 level, it equals to 2 km) and over time (about 1 week during expeditions). Time averaging was necessary due to the lack of appropriate satellite data for field measurements under cloudy conditions. Only in 2014, for almost all stations (eight from nine), L2 satellite data was available, while the discrepancy between satellite and field measurements time did not exceed 6 h. As a result, it was possible to compare the correspondence indicators between the model Bcyan calculated for different input dataset and measured Bcyan (Table 2). The first set of input data (#1 in situ) was obtained from the field measurements (R_{rs} and Chl); the same data were used to derive the algorithm. The second dataset (#2 MODIS L2, 1 px) was generated from the MODIS-Aqua L2 data using the nearest pixel to the station. The third dataset (#3 MODIS L2, 9 px) was also based on the MODIS-Aqua L2 data, but averaging was carried out for the nine nearest pixels. The fourth dataset (#4 MODIS L3) was generated with the MODIS-Aqua L3 data averaged during the expedition (25 July–2 August 2014) as a mean of 5–7 satellite overpasses; the size of bin was 2 km.

Table 2. The parameters of correspondence between the model and measured Bcyan for 8 stations in 2014. The model Bcyan were computed for the different input datasets (see text for details).

Input Dataset	<model> *, mg m^{-3}	<model>/<measured>	R^2	RMSE, mg m^{-3}	Ratio	Range of 'mod/meas'
#1 in situ	522	0.83	0.59	428	1.31	0.34–3.61
#2 MODIS L2, 1 px	527	0.84	0.66	393	1.53	0.53–3.41
#3 MODIS L2, 9 px	564	0.90	0.75	332	1.45	0.58–4.18
#4 MODIS L3	506	0.81	0.56	443	1.22	0.42–7.40

*—<model> is average value of Bcyan model; <model>/<measured> is the ratio of the mean values of Bcyan model and measured; R^2 is coefficient of determination; RMSE is root-mean-square error; Ratio is median value of the ratio of Bcyan model to Bcyan measured; 'mod/meas' is ratio of Bcyan model to measured.

Table 2 shows that all four input datasets yield very similar results. Based on the coefficient of determination, the RMSE, and the ratio of the mean values <model>/<measured>, the best match between the model and measured values is with the MODIS-Aqua L2 data with averaging over the nine nearest pixels (#3 MODIS L2, 9 px). It is even better than using the field dataset (#1 in situ). If judging by the Ratio value, then the best match is for the MODIS-Aqua L3 dataset (#4 MODIS L3). The other statistical metrics of the dataset (#4 MODIS L3) differ insignificantly from those for the considered datasets, and they seem to be close to the statistical parameters of the field data (#1 in situ). This suggests that time-averaged satellite data can be used for both the algorithm validation and investigation of the spatial and temporal variability of Bcyan. Note that Table 2 only represents results for eight stations and that the field data, both the measured Bcyan and input values of R_{rs} and Chl in the dataset (#1 in situ), contain measurement errors.

Figure 5 presents another attempt to validate the algorithm of Bcyan estimation from satellite data. The average field values of Bcyan, measured during summer expeditions (late July–early August) in the eastern part of the Gulf of Finland in 2004–2011 [29] are compared with values of Bcyan calculated from satellite data. To better match the field data, the MODIS-Aqua estimates were generated over the expedition period (about one week), and the region for calculating the mean Bcyan was reduced (it is bounded in the west by the 27° E, which passes through the island of Gogland). The in situ average values were obtained from measurements for about 20 stations, while satellite estimations of Bcyan in the given region were calculated by averaging over approximately 2000 pixels. This leads to a significant difference between the measured and estimated-from-satellite data Bcyan values. Nevertheless, Figure 5 shows that the estimates from satellite data confirm the cyanobacteria biomass reduction in 2009–2011 compared to 2004–2007, which can be related to the invasion of polychaetes *Marenzelleria arctia* [29]. But according to satellite estimations, the decrease in Bcyan was not as sharp (only about three times) as in the field data (about seven times).

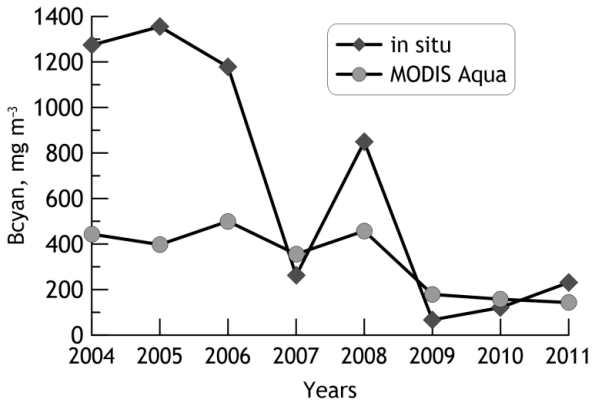

Figure 5. Comparison of the average Bcyan values (mg m^{-3}) in the eastern part of the Gulf of Finland estimated from MODIS Aqua data and measured in situ during summer expeditions [29].

3.2. Threshold for Determining the Cyanobacteria Bloom

According to [31], cyanobacteria vital activity may be considered as "blooms" at a biomass concentration of about 200 mg m^{-3} in the mixed upper layer of the water. According to the Kahru algorithm for MODIS-Aqua [6,7], a pixel receives a "blooming" mark if $R_{rs}(667) > 0.0012$. An example of bloom areas comparison on 25 July 2014, according to the Kahru and our algorithms, is shown in Figure 6. Apparently, the locations of the bloom regions are similar for both algorithms, but there are negligible differences, particularly near the shore. Using the biologists' bloom threshold value for field data (Bcyan = 200 mg m^{-3}), the bloom area by our satellite algorithm appears much larger than by the Kahru algorithm, 4380 versus 2574 km^2, respectively. Considering the errors in the satellite Bcyan algorithm, the Bcyan bloom absence threshold should be shifted to a value of Bcyan = 300 mg m^{-3}. In this case, the resulting bloom area estimated by our algorithm is 2579 km^2, and the bloom areas estimated by both algorithms almost coincide. Further in our study (Section 3.3), value Bcyan = 300 mg m^{-3} has been used as a threshold for determining the cyanobacteria bloom area from satellite data.

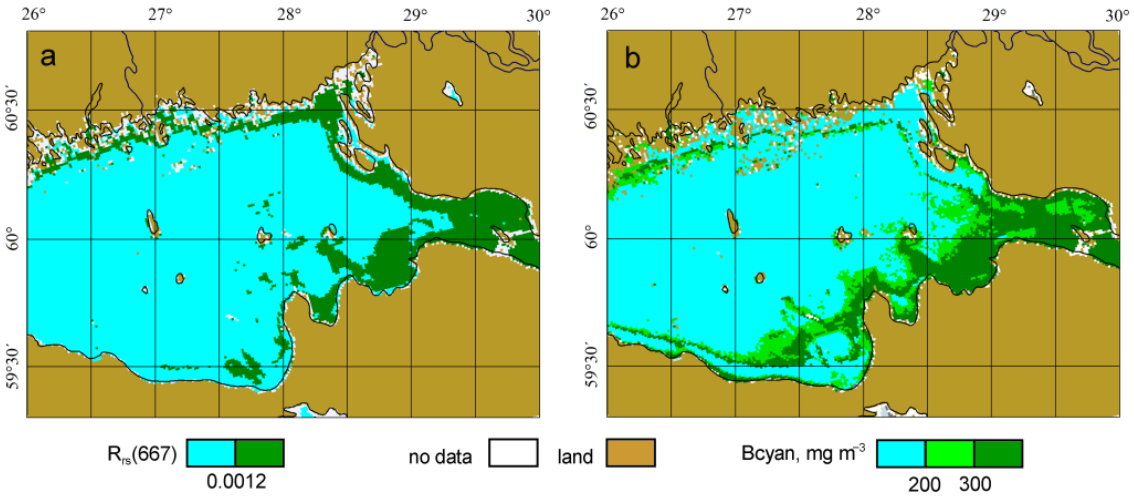

Figure 6. Comparison of bloom areas on 25 July 2014, according to Kahru algorithm (**a**) and value of Bcyan (**b**).

3.3. Interannual Changes of the Cyanobacteria Blooms' Characteristics

To study the spatial and temporal variability of cyanobacteria biomass in the eastern part of the Gulf of Finland, the mean distributions of cyanobacteria biomass in 2003–2022 for months of the bloom season (July–August) were calculated using the MODIS-Aqua data with the help of the developed algorithm. The maps of mean monthly distributions are available in the SIO RAS electronic bio-optical Atlas (http://optics.ocean.ru, accessed on 4 September 2023).

Figure 7 shows the variability of the monthly mean values of Bcyan in the study region in 2003–2022. As one can see, almost all the years when obvious blooms were observed, the Bcyan value was higher in July than in August. Only in 2015 and 2020 did the average Bcyan values in August significantly exceeded those in July. The record bloom during the studied period occurred in July 2004 (<Bcyan> = 780 mg m^{-3}). Figure 7 also agrees well with the conclusion about a significant reduction in the blue-green algae biomass in 2009–2011 compared to 2004–2008 [29].

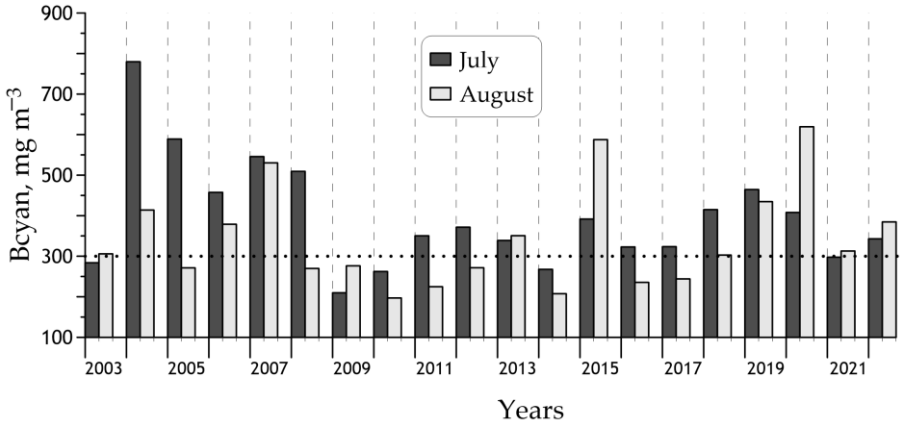

Figure 7. Variability of the monthly mean values of Bcyan (mg m^{-3}) in the eastern part of the Gulf of Finland in 2003–2022. The horizontal dotted line is the threshold value for cyanobacteria blooms.

The analysis of the interannual variability of cyanobacterial bloom areas (Figure 8) is based on the monthly distributions of Bcyan. The assessment of these areas shows that the maximum cyanobacteria bloom in the eastern part of the Gulf of Finland usually happens in July, similarly to findings based on analysis of mean values of Bcyan. In July 2004, 2005, and 2008, the bloom areas occupied more than 90% of the region area; in August 2015 and 2020, they occupied89 and 92%, respectively. The weakest blooms were observed in 2009, 2010, and 2014, when both in July and August, the bloom areas did not exceed 30% of the study region.

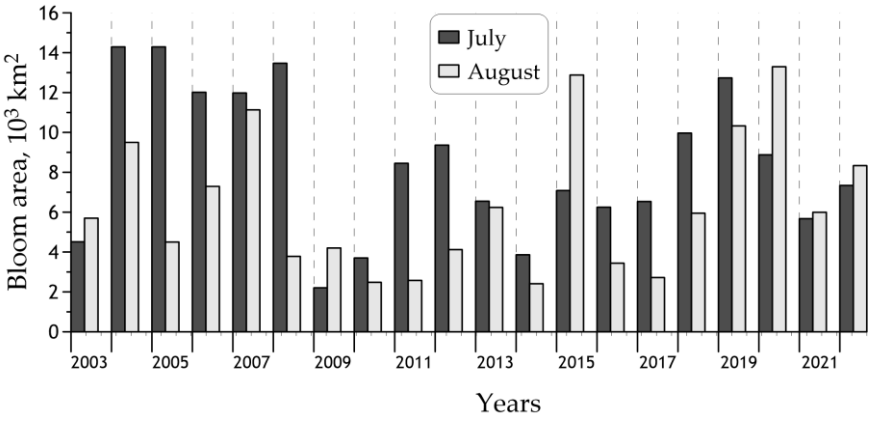

Figure 8. Variability of the monthly mean values of cyanobacteria bloom area (10^3 km^2) in the eastern part of the Gulf of Finland in 2003–2022. Value Bcyan = 300 mg m^{-3} is used as a threshold for determining the cyanobacteria bloom area.

Figure 9 presents maps of mean distributions of Bcyan for the bloom season (July–August) from 2003 to 2022, while Table 3 gives the statistical parameters. Almost every year, maximum values of Bcyan (over 900 mg m^{-3}) occur in the estuary of the Neva River, in the Kaporskaya and Luga Bays. There is considerable interannual variance in Bcyan in the central part of the studied region from the minimum values in 2014 to the maximum in 2004. The most intense and extensive blooms fell on 2004–2008, 2015, and 2019–2020: the region's average biomass exceeded 400 mg m^{-3}, with bloom areas occupying more than 75% of the study region. The record bloom was observed in 2004: average <Bcyan> = 598 mg m^{-3}; 100% of the area was occupied by the bloom; 43% of the study region was occupied by intensive bloom with Bcyan > 600 mg m^{-3}. The weakest blooms were in 2009, 2010, and 2014: the average biomass in the region did not exceed 250 mg m^{-3}, while the bloom areas occupied less than 25% of the region.

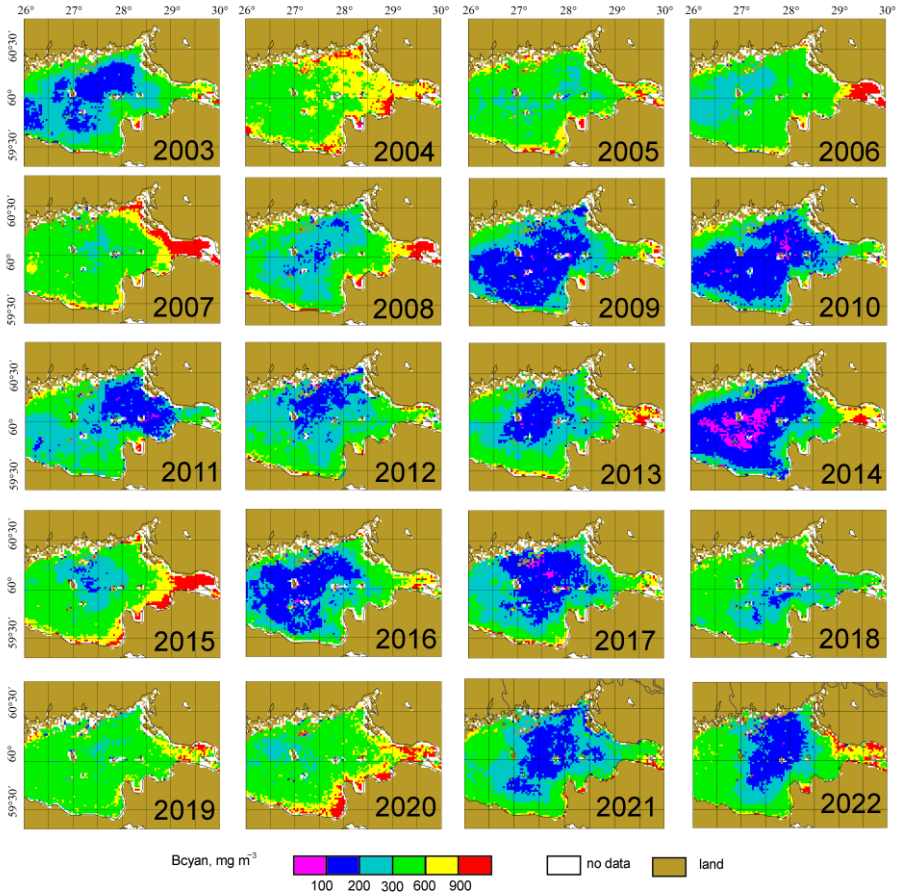

Figure 9. Mean distributions of Bcyan for bloom season (July–August) in the eastern part of the Gulf of Finland in 2003–2022.

Table 3. Characteristics of mean distributions of Bcyan for bloom season (July–August) in the eastern Gulf of Finland for 2003–2022: the region-average cyanobacteria biomasses <Bcyan> and standard deviations; bloom areas with Bcyan exceeding 300 and 600 mg m^{-3}; the numbers in parentheses indicate the proportion of blooms in relation to the area of the region.

Year	<Bcyan>, mg m^{-3}	Bloom Area, 10^3 km^2	
		>300 mg m^{-3}	>600 mg m^{-3}
2003	297 ± 148	5.1 (35%)	0.7 (5%)
2004	598 ± 205	14.5 (100%)	6.2 (43%)
2005	434 ± 182	12.6 (87%)	2.1 (15%)
2006	422 ± 219	11.2 (77%)	1.6 (11%)
2007	543 ± 308	13.6 (94%)	3.3 (22%)
2008	401 ± 265	8.2 (56%)	1.9 (13%)
2009	247 ± 150	3.0 (21%)	0.5 (3%)
2010	231 ± 135	2.7 (19%)	0.4 (3%)
2011	289 ± 150	5.2 (36%)	0.4 (3%)
2012	325 ± 181	6.2 (43%)	0.8 (5%)
2013	347 ± 209	7.1 (49%)	1.4 (9%)
2014	240 ± 205	3.0 (20%)	0.9 (6%)
2015	498 ± 313	11.3 (78%)	3.3 (23%)
2016	283 ± 173	4.4 (30%)	0.8 (6%)
2017	289 ± 204	4.3 (30%)	0.9 (6%)
2018	360 ± 121	9.9 (68%)	0.4 (3%)
2019	451 ± 163	13.6 (94%)	1.5 (10%)
2020	518 ± 256	13.1 (91%)	3.7 (26%)
2021	302 ± 142	6.1 (42%)	0.6 (4%)
2022	365 ± 217	8.4 (58%)	1.6 (11%)
2003–2022	**372 ± 107**	**8.2 (56%)**	**1.6 (11%)**

The interannual variability of the region average cyanobacteria biomass <Bcyan>, shown in Table 3, is in good agreement with the results of in situ measurements in the Gulf of Finland performed during the HELCOM COMBINE program [14]. Considering the apparently expected discrepancy between satellite estimates and in situ measured values of cyanobacteria biomass, and only a partial overlap for data acquisition areas (in situ HELCOM COMBINE measurements were carried out mainly in the western part of the Gulf of Finland), the obtained coincidence of the average <Bcyan> values over the region is especially remarkable.

Figure 2 from the HELCOM Baltic Sea Environment Fact Sheet 2020 [14] maintains that for the period 2003–2019, the maximum values of Bcyan were observed in 2004, 2015, and 2018. The region average values of cyanobacteria biomass <Bcyan> calculated using satellite data (Table 3) and in situ measurements [14] are very close as well: in 2004, the satellite value of <Bcyan> is 598 against in situ estimates about 625 mg m^{-3}, while in 2015, these estimates are about 500 and 490 mg m^{-3}, respectively.

For years with weak cyanobacteria bloom, the coincidence of satellite and in situ assessments of <Bcyan> turned out worse. According to in situ data, the minimum cyanobacteria biomass was observed in 2014, when <Bcyan> were about 140 mg m^{-3}, versus ~240 mg m^{-3} from satellite data. At the same time, according to satellite estimates, 2014 is one of the three weakest blooming years. The greatest difference between satellite estimates of <Bcyan> and in situ values occurred in 2007: 543 versus 200 mg m^{-3}. A possible explanation of the disagreement is that 2007 intensive bloom and high Bcyan values were observed in the northeastern part of the study region (Figure 9), outside the in situ data coverage. Additionally, the authors of [29] suggest that the introduction of *Marenzelleria arctia* polychaetes caused a decrease in <Bcyan> starting 2009. Since the polychaetes introduction in the western part of the region occurred earlier than in the eastern part, the decrease in <Bcyan> came into effect earlier.

The satellite and in situ average Bcyan values for the period 2003–2019 are in good agreement: 368 and 330 mg m^{-3}, respectively.

4. Discussion

We present the regional algorithm for cyanobacteria biomass assessment in the eastern part of the Gulf of Finland using satellite data. The algorithm was developed on the basis of field measurements in July–August 2012–2014. According to the measurement data, the cyanobacteria biomass correlates with two bio-optical characteristics of water: the chlorophyll a concentration (the coefficient of determination R^2 = 0.50) and the particle backscattering coefficient (R^2 = 0.50). The multi-regression equation with these two parameters provides the best correlation (R^2 = 0.61) and the smallest Bcyan determination errors (RMSE is 272 mg m^{-3}, median percent difference is 54%). At the same time, for stations with cyanobacteria blooms, the relative errors are much smaller.

The proposed algorithm makes it possible to determine both the area of cyanobacteria bloom and its intensity. Considering the algorithm errors, the threshold value for bloom detection was shifted from 200 mg m^{-3} determined by biologists [31] to 300 mg m^{-3}.

The MODIS-Aqua ocean color data gave us an opportunity to analyze the spatial and temporal variability of cyanobacteria blooms in the eastern part of the Gulf of Finland from 2003 to 2022. Almost every year, the maximum values of Bcyan (more than 900 mg m^{-3}) are observed in the estuary of the Neva river, in the Kaporskaya and Luga bays. In the central part of the region under study, there is a significant interannual scatter of Bcyan from the minimum values in 2014 to the maximum in 2004. The bloom record was observed in 2004: average <Bcyan> = 598 mg m^{-3}; 100% of the area was occupied by the bloom; 43%—intensive bloom with Bcyan > 600 mg m^{-3}. The interannual variability of the region-averaged cyanobacteria biomass <Bcyan> is in good agreement with in situ measurements in the Gulf of Finland [14] obtained within the HELCOM COMBINE program.

The proposed algorithm makes it possible to use satellite data from all known ocean color scanners, including SeaWiFS, VIIRS, MODIS, and OLCI. The algorithm can be used for

real-time monitoring of cyanobacteria blooms in the region with a spatial resolution up to 300 m, according to OLCI data, as well as to analyze interannual variability since 1998, when SeaWiFS data became available. There are recent applications of remote sensing data for the Baltic Sea. For instance, there are emerging opportunities to utilize these data for time-series monitoring of shorelines, classification of coastal regions, and wetland monitoring in the Baltic Sea [32–34]. Note that this algorithm can be applied to AERONET-OC R_{rs} data [35], which are free from atmospheric correction errors. In addition, the developed algorithm will allow in the future not only the monitoring of harmful algal blooms in the study area, but the exploration of the relationship between bloom characteristics and marine hydrological factors [36].

5. Conclusions

A regional algorithm for estimating the biomass of cyanobacteria and their bloom area in the eastern part of the Gulf of Finland in the Baltic Sea has been developed and validated. It is based on field bio-optical measurements and direct species composition determinations carried out in July–August 2012–2014 in the eastern part of the Gulf of Finland. The algorithm uses the particle backscattering coefficient and chlorophyll concentration as input parameters. The Bcyan values are calculated according to a multi-regression equation, which provides better performance than a linear one, resulting in both R^2 and *RMSE* values (0.61 and 272 mg m^{-3}, respectively). Unlike other algorithms [6,7], which determine only the bloom area, our algorithm also allows the estimation of its intensity. Note that the location of bloom areas is the same for both types of algorithms, with minor differences near the shore. Taking into account the errors in the satellite Bcyan algorithm, we shifted the bloom threshold to Bcyan = 300 mg m^{-3}. According to the results of the analysis of the processed MODIS ocean color data for the period 2002–2022, significant interannual variability of cyanobacteria biomass in this region was revealed. For almost all the years when blooms were observed, the Bcyan value was higher in July than in August. The record bloom during the study period occurred in July 2004 (the average Bcyan value was equal to 780 mg m^{-3}). The weakest blooms were observed in 2009, 2010, and 2014, when both in July and August, the bloom areas did not exceed 30% of the study region. These results are in good agreement with the in situ measurements in the Gulf of Finland obtained within the framework of the HELCOM COMBINE program [14]. It is shown that the agreement between direct measurements and satellite data is higher for cases of more intense blooms. Maps of mean monthly distributions are available in the electronic bio-optical Atlas of the SIO RAS (http://optics.ocean.ru, accessed on 4 September 2023) [37].

Author Contributions: Conceptualization, S.V. and O.K.; methodology, S.V., O.K. and E.K.; software, I.S.; validation, S.V.; formal analysis, S.V. and O.K.; investigation, S.V., E.K., I.S., E.L., A.K. and T.E.; writing, S.V., E.K. and D.G.; visualization, S.V. and E.K.; project administration, O.K.; funding acquisition, D.G. All authors have read and agreed to the published version of the manuscript.

Funding: Shipboard data retrieval was carried out as part of the state assignment of SIO RAS (theme No. FMWE-2021-0001). Regional algorithm development and validation were funded by the Russian Science Foundation (research project 21-77-10059). Analysis of the spatial and temporal variability of cyanobacteria blooms was sponsored under the project funded by the Russian Hydrometeorological Service (contract agreement # 169-15-2023-002).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on reasonable request from the corresponding author.

Acknowledgments: This work is dedicated to the memory of Oleg Kopelevich, our teacher and inspirator, deceased at the end of 2020.

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

References

1. Huisman, J.; Codd, G.A.; Paerl, H.W.; Ibelings, B.W.; Verspagen, J.M.H.; Visser, P.M. Cyanobacterial blooms. *Nat. Rev. Microbiol.* **2018**, *16*, 471–483. [CrossRef] [PubMed]
2. Gower, J.; King, S. New results from a global survey using MERIS MCI. In Proceedings of the 2nd MERIS/(A)ATSR User Workshop, Frascati, Italy, 22–26 September 2008.
3. Gower, J.; King, S.; Goncalves, P. Global monitoring of plankton blooms using MERIS MCI. *Int. J. Remote Sens.* **2008**, *29*, 6209–6216. [CrossRef]
4. Hansson, M.; Hakansson, B. The Baltic Algae Watch System—A remote sensing application for monitoring cyanobacterial blooms in the Baltic Sea. *J. Appl. Remote Sens.* **2007**, *1*, 011507. [CrossRef]
5. Hansson, M.; Pemberton, P.; Håkansson, B.; Reinart, A.; Alikas, K. Operational Nowcasting of Algal Blooms in the Baltic Sea Using MERIS and MODIS. In Proceedings of the ESA Living Planet Symposium, Bergen, Norway, 28 June–2 July 2010; Special Publication SP-686.
6. Kahru, M.; Savchuk, O.P.; Elmgren, R. Satellite measurements of cyanobacterial bloom frequency in the Baltic Sea: Interannual and spatial variability. *Mar. Ecol. Prog. Ser.* **2007**, *343*, 15–23. [CrossRef]
7. Kahru, M.; Elmgren, R. Multidecadal time series of satellite-detected accumulations of cyanobacteria in the Baltic Sea. *Biogeosciences* **2014**, *11*, 3619–3633. [CrossRef]
8. Kahru, M.; Elmgren, R.; Kaiser, J.; Wasmund, N.; Savchuk, O. Cyanobacterial blooms in the Baltic Sea: Correlations with environmental factors. *Harmful Algae* **2020**, *92*, 101739. [CrossRef]
9. Kutser, T.; Metsamaa, L.; Strömbeck, N.; Vahtmäe, E. Monitoring cyanobacterial blooms by satellite remote sensing. *Estuar. Coast. Shelf Sci.* **2006**, *67*, 303–312. [CrossRef]
10. Reinart, A.; Kutser, T. Comparison of different satellite sensors in detecting cyanobacterial bloom events in the Baltic Sea. *Remote Sens. Environ.* **2006**, *102*, 74–85. [CrossRef]
11. Riha, S.; Krawczyk, H. Development of a remote sensing algorithm for cyanobacterial phycocyanin pigment in the Baltic Sea using neural network approach. In Proceedings of the Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions, Prague, Czech Republic, 7 October 2011. [CrossRef]
12. Woźniak, M.; Bradtke, K.M.; Darecki, M.; Krężel, A. Empirical Model for Phycocyanin Concentration Estimation as an Indicator of Cyanobacterial Bloom in the Optically Complex Coastal Waters of the Baltic Sea. *Remote Sens.* **2016**, *8*, 212. [CrossRef]
13. HELCOM Thematic Assessment of Eutrophication 2011–2016. Baltic Sea Environment Proceedings No. 156. Available online: http://www.helcom.fi/baltic-sea-trends/holistic-assessments/state-of-the-baltic-sea-2018/reports-and-materials (accessed on 27 July 2022).
14. Kownacka, J.; Busch, S.; Göbel, J.; Gromisz, S.; Hällfors, H.; Höglander, H.; Huseby, S.; Jaanus, A.; Jakobsen, H.H.; Johansen, M.; et al. Cyanobacteria biomass, 1990–2019. HELCOM Baltic Sea Environment Fact Sheets 2020. Available online: https://helcom.fi/wp-content/uploads/2020/09/BSEFS-Cyanobacteria-biomass-1990-2019-1.pdf (accessed on 27 July 2022).
15. Öberg, J. Cyanobacteria Blooms in the Baltic Sea. HELCOM Baltic Sea Environment Fact Sheets 2017. Available online: https://helcom.fi/wp-content/uploads/2020/06/BSEFS-Cyanobacteria-blooms-in-the-Baltic-Sea.pdf (accessed on 27 July 2022).
16. Simis, S.G.H.; Peters, S.W.M.; Gons, H.J. Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnol. Oceanogr.* **2005**, *50*, 237–245. [CrossRef]
17. Soja-Woźniak, M.; Craig, S.E.; Kratzer, S.; Wojtasiewicz, B.; Darecki, M.; Jones, C.T. A Novel Statistical Approach for Ocean Colour Estimation of Inherent Optical Properties and Cyanobacteria Abundance in Optically Complex Waters. *Remote Sens.* **2017**, *9*, 343. [CrossRef]
18. Gower, J.; King, S.; Borstad, G.; Brown, L. Detection of intense plankton blooms using the 709 nm band of the MERIS imaging spectrometer. *Int. J. Remote Sens.* **2005**, *26*, 2005–2012. [CrossRef]
19. Wynne, T.T.; Stumpf, R.P.; Tomlinson, M.C.; Warner, R.A.; Tester, P.A.; Dyble, J.; Fahnenstiel, G.L. Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes. *Int. J. Remote Sens.* **2008**, *29*, 3665–3672. [CrossRef]
20. Gower, J.F.R.; Doerffer, R.; Borstad, G.A. Interpretation of the 685 nm peak in water-leaving radiance spectra in terms of fluorescence, absorption and scattering, and its observation by MERIS. *Int. J. Remote Sens.* **1999**, *20*, 1771–1786. [CrossRef]
21. Wynne, T.T.; Stumpf, R.P.; Tomlinson, M.C.; Dyble, J. Characterizing a cyanobacterial bloom in Western Lake Erie using satellite imagery and meteorological data. *Limnol. Oceanogr.* **2010**, *55*, 2025–2036. [CrossRef]
22. Wynne, T.; Stumpf, R.; Briggs, T. Comparing MODIS and MERIS spectral shapes for cyanobacterial bloom detection. *Int. J. Remote Sens.* **2013**, *34*, 6668–6678. [CrossRef]
23. Hu, C. A novel ocean color index to detect floating algae in the global oceans. *Remote Sens. Environ.* **2009**, *113*, 2118–2129. [CrossRef]
24. Report of SCOR-UNESCO Working Group 17 on Determination of Photosynthetic Pigments. In *Determination of Photosynthetic Pigments in Sea-Water*; UNESCO: Paris, France, 1966; pp. 9–16.
25. Artemiev, V.A.; Burenkov, V.I.; Vortman, M.I.; Grigoriev, A.V.; Kopelevich, O.V.; Khrapko, A.N. Sea-truth measurements of ocean color: A new floating spectroradiometer and its metrology. *Oceanology* **2000**, *40*, 139–145.
26. Vazyulya, S.; Khrapko, A.; Kopelevich, O.; Burenkov, V.; Eremina, T.; Isaev, A. Regional algorithms for the estimation of chlorophyll and suspended matter concentration in the Gulf of Finland from MODIS-Aqua satellite data. *Oceanologia* **2014**, *56*, 737–756. [CrossRef]

27. Lee, Z.; Carder, K.L.; Mobley, C.D.; Steward, R.G.; Patch, J.S. Hyperspectral remote sensing for shallow waters I A semianalytical model. *Appl. Opt.* **1998**, *37*, 6329–6338. [CrossRef]
28. Seppälä, J.; Ylöstalo, P.; Kaitala, S.; Hällfors, S.; Raateoja, M.; Maunula, P. Ship-of-opportunity based phycocyanin fluorescence monitoring of the filamentous cyanobacteria bloom dynamics in the Baltic Sea. *Estuar. Coast. Shelf Sci.* **2007**, *73*, 489–500. [CrossRef]
29. Maximov, A.A.; Eremina, T.R.; Lange, E.K.; Litvinchuk, L.F.; Maximova, O.B. Regime shift in the ecosystem of the eastern Gulf of Finland caused by the invasion of the polychaete Marenzelleria arctia. *Oceanology* **2014**, *54*, 46–53. [CrossRef]
30. Burenkov, V.I.; Ershova, S.V.; Kopelevich, O.V.; Sheberstov, S.V.; Shevchenko, V.P. An estimate of the distribution of sus-pended matter in the Barents Sea waters on the basis of the SeaWiFS satellite ocean color scanner. *Oceanology* **2001**, *41*, 622–628.
31. Wasmund, N. Occurrence of cyanobacterial blooms in the baltic sea in relation to environmental conditions. *Int. Rev. Hydrobiol.* **1997**, *82*, 169–184. [CrossRef]
32. Tiede, J.; Jordan, C.; Moghimi, A.; Schlurmann, T. Long-term shoreline changes at large spatial scales at the Baltic Sea: Remote-sensing based assessment and potential drivers. *Front. Mar. Sci.* **2023**, *10*, 1207524. [CrossRef]
33. Cazzaniga, I.; Zibordi, G.; Mélin, F. Spectral features of ocean colour radiometric products in the presence of cyanobacteria blooms in the Baltic Sea. *Remote Sens. Environ.* **2023**, *287*, 113464. [CrossRef]
34. Tilstone, G.H.; Pardo, S.; Simis, S.G.H.; Qin, P.; Selmes, N.; Dessailly, D.; Kwiatkowska, E. Consistency between Satellite Ocean Colour Products under High Coloured Dissolved Organic Matter Absorption in the Baltic Sea. *Remote Sens.* **2021**, *14*, 89. [CrossRef]
35. Qin, P.; Simis, S.G.; Tilstone, G.H. Radiometric validation of atmospheric correction for MERIS in the Baltic Sea based on continuous observations from ships and AERONET-OC. *Remote Sens. Environ.* **2017**, *200*, 263–280. [CrossRef]
36. Chen, C.; Liang, J.; Yang, G.; Sun, W. Spatio-temporal distribution of harmful algal blooms and their correlations with marine hydrological elements in offshore areas, China. *Ocean Coast. Manag.* **2023**, *238*, 106554. [CrossRef]
37. Kopelevich, O.V.; Sahling, I.V.; Vazyulya, S.V.; Glukhovets, D.I.; Sheberstov, S.V.; Burenkov, V.I.; Karalli, P.G.; Yushmanova, A.V. *Bio-Optical Characteristics of the Seas, Surrounding the Western Part of Russia, from Data of the Satellite Ocean Color Scanners of 1998–2017*; VASh FORMAT, OOO: Moscow, Russia, 2018.