SOPhyE Satellite Data Processing Technical Report Series: 1. Ocean Colour Satellite Intercalibration

Stephanie Clay, Emmanuel Devred

Fisheries and Oceans Canada
Science Branch, Maritimes Region
Bedford Institute of Oceanography
P.O. Box 1006
1 Challenger Drive
Dartmouth, NS B2Y 4A2

2023

**Canadian Technical Report of
Fisheries and Aquatic Sciences 3560**

Fisheries and Oceans Canada   Pêches et Océans Canada

Canada

Canadian Technical Report of Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Rapport technique canadien des sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. Il n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques.*

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report of
Fisheries and Aquatic Sciences 3560

2023

SOPHYE SATELLITE DATA PROCESSING TECHNICAL REPORT SERIES: 1. OCEAN COLOUR SATELLITE INTERCALIBRATION

by

Stephanie Clay[1] and Emmanuel Devred[1]

[1]Bedford Institute of Oceanography
Fisheries and Oceans Canada, 1 Challenger Drive
Dartmouth, Nova Scotia, B2Y 4A2, Canada

Correct citation for this publication:

Clay, S. and Devred, E. 2023. SOPhyE Satellite Data Processing Technical Report Series: 1.
Ocean Colour Satellite Intercalibration. Can. Tech. Rep. Fish. Aquat. Sci. 3560: vi + 42 p.

CONTENTS

ABSTRACT

Clay, S. and Devred, E. 2023. SOPhyE Satellite Data Processing Technical Report Series: 1. Ocean Colour Satellite Intercalibration. Can. Tech. Rep. Fish. Aquat. Sci. 3560: vi + 42 p.

Though satellite ocean colour data has massively increased the scope of ocean observations in the past few decades, differences in waveband placement and their spectral response function between satellite sensors lead to inconsistencies between datasets that prevent users from taking advantage of the combined satellite data repository in environmental studies. In this report we examine the spatio-temporal differences between remote-sensing reflectance and chlorophyll-a concentration products derived from three NASA ocean colour sensors (i.e., MODIS-Aqua, SeaWiFS, and VIIRS-SNPP). We then present a simple method to align SeaWiFS and VIIRS reflectances to MODIS, and derive chlorophyll-a concentration from the modelled bands. These chlorophyll-a concentration estimates are compared to in situ matchups and the effects on spring phytoplankton bloom metrics are examined. The "alignment" method can theoretically be applied to any set of ocean colour sensors in order to generate a common set of wavebands, which could then be merged into a single multi-sensor product for higher spatial and temporal coverage suitable for climate studies.

RÉSUMÉ

Clay, S. and Devred, E. 2023. SOPhyE Satellite Data Processing Technical Report Series: 1. Ocean Colour Satellite Intercalibration. Can. Tech. Rep. Fish. Aquat. Sci. 3560: vi + 42 p.

Bien que les données satellitaires sur la couleur des océans aient considérablement élargi la portée des observations océaniques au cours des dernières décennies, les différences dans le placement des longueur d'ondes et leur fonction de réponse spectrale entre les capteurs satellitaires entraînent des incohérences entre les ensembles de données qui empêchent les utilisateurs de tirer parti des données satellitaires combinées dans le cadre d'études environnementales. Dans ce rapport, nous examinons les différences spatio-temporelles entre les produits de télédétection, i.e., la réflectance et la concentration en chlorophylle-a, dérivés de trois capteurs de couleur de l'eau de la NASA (MODIS-Aqua, SeaWiFS et VIIRS-SNPP). Nous présentons ensuite une méthode simple pour aligner les réflectances de SeaWiFS et de VIIRS sur celle de MODIS, et pour dériver la concentration de chlorophylle-a à partir des bandes modélisées. Ces estimations de concentration de chlorophylle-a sont comparées à des mesures in situ et les effets sur les indices de la floraison printanière du phytoplancton sont examinés. La méthode d'"alignement" peut théoriquement être appliquée à n'importe quel ensemble de capteurs de couleurs de l'eau afin de générer un ensemble commun de longueur d'ondes, qui pourraient ensuite être fusionnées en un seul produit multi-capteurs pour une couverture spatiale et temporelle plus élevée convenant aux études climatiques.

1   Introduction

Satellite remote-sensing has been widely used in recent decades to collect medium-resolution and high-frequency ocean colour data over the entire globe. This technology provides a synoptic view of the oceans for use in long-term studies, but is not without its drawbacks. There is a frequent need for validation of algorithms with relatively sparse corresponding in situ data matchups, and gaps in the dataset from a variety of causes such as low solar zenith angle (particularly at high latitudes), cloud and ice cover, sun glint, or swath width (i.e. viewing area) of the sensor which limits the frequency of revisit (https://oceancolor.gsfc.nasa.gov/resources/atbd/ocl2flags).

Temporal and spatial gaps in the satellite data repository have always been a significant problem. Various techniques have been tested in order to estimate missing values of ocean colour variables, which are important as they are included in the set of Essential Climate Variables (ECV) used to describe the Earth's climate (Belward et al. 2016). Some focus on filling spatial gaps in sequences of images, such as mathematical methods like DINEOF (Data Interpolating Empirical Orthogonal Functions) that estimate missing data based on the surrounding regions and time series (Alvera-Azcárate et al. 2009). Other methods of dealing with spatial data gaps include sacrificing temporal frequency by compositing images over a selected time period to increase spatial coverage.

As multiple ocean colour satellites have operated concurrently, combining imagery into multisensor datasets can improve spatial and temporal coverage. In this report, we use three common ocean colour satellites (MODIS, SeaWiFS, and VIIRS) and investigate a method to ensure consistency between sensor Remote Sensing Reflectance data (*Rrs*) so that MODIS may be used in conjunction with the other sensors. Inconsistencies between sensors can arise from various sources in the field such as different solar and sensor viewing angles, atmospheric paths, and satellite pass times (Barnes et al. 2021). Moreover, spatial resolution varies between sensor which affects the final aggregated pixel value, and each sensor has its own unique spectral response, which is defined as the number of photons detected by the sensor, varying across the visible spectrum with increased sensitivity around certain wavelengths. In this study we focused on differences between similar bands but did not characterize the effects of other potential sources of discontinuities.

Previous attempts have used different models to estimate *Rrs* from one sensor at wavebands from another sensor, with varying degrees of complexity and success. Notably, the Ocean Colour Climate Change Initiative (OC-CCI) (Sathyendranath et al. 2019) relies on a semi-analytical model that calculates the inherent optical properties (IOPs) of the surface water using the existing *Rrs* from a sensor. The model can then be inverted and the IOPs used to estimate *Rrs* at different wavebands, specifically the bands from the chosen reference sensor. This is computationally expensive, as the spectrum from each individual pixel must be uniquely modelled in order to retrieve the IOPs and *Rrs* at other wavelengths. Another simpler model was proposed in (Wang et al. 2020), relying on a pre-calculated factor to convert radiances or reflectances from one sensor band to another used in specific ocean variables. However, this requires sufficient in situ radiance measurements to constrain the conversion factor, and was only tested to shift band values to nearby wavelengths.

This report details a method of using simple multilinear regression to model *Rrs* to match bands

between sensors. This method is empirical and requires spatially and temporally overlapping sensor data in order to collect coincident pixels for the training set. The objectives of this report are 1) to investigate the differences between MODIS, SeaWiFS, and VIIRS reflectances, 2) model all visible MODIS bands using data from the other sensors, and 3) evaluate model performance and possible sources of error in relation to the true MODIS data. Chlorophyll-a (*Chl-a*) products are derived from the band-modelled data and used to determine the band-modelling effect on in situ sample matchups and spring phytoplankton bloom metrics.

Calculation of the multilinear regression coefficients and resulting SeaWiFS (VIIRS)-modelled MODIS datasets is simple and relatively fast, and the coefficients can be easily tuned to a specific region. In this case, our region of interest is the Northwest Atlantic (*NWA*, 42°- 95°W, 39°- 82°N), an area of particular importance for Canadian fisheries management (Therriault et al. 1998), climate studies (Azetsu-Scott et al. 2008), and ecosystem studies (Trzcinski et al. 2013). The region is surveyed on the bi-annual Atlantic Zone Monitoring Program (AZMP) and annual Atlantic Zone Off-Shelf Monitoring Program (AZOMP) cruises, which collect in situ biogeochemical variables primarily along the Scotian Shelf and in the Labrador Sea to monitor the state of the ocean (Therriault et al. 1998). This low frequency sampling means that short-term changes are often missed, for example the spring and fall phytoplankton blooms (see Figure 2 in Ringuette et al. 2022). Satellite sensors, with their high sampling frequency and wide spatial coverage, can capture more events that are often missed by in situ sampling. However, they are susceptible to cloud and ice cover as well as lack of data in the more northern areas due to low solar zenith angle, and could benefit greatly from well-tested gap-filling techniques. The modelled MODIS bands using the method in this report, coupled with true MODIS data, will provide a longer consistent satellite time series that will help fill large temporal and spatial data gaps and be more compatible with studies of long-term ecosystem changes.

2 Data and methods

2.1 Sensor data

For simplicity and computational efficiency but without compromising the scientific soundness of the study, level-3 binned (L3b) satellite *Rrs* and *Chl-a* datasets were used. At this stage, NASA's default atmospheric correction algorithm has already been applied, and the data have been filtered using a common set of quality-control flags, before being composited into daily images and projected into approximately equal-area bins at 4.64km-resolution on a common irregular grid (IOCCG 2004). Changes in the *Rrs* and *Chl-a* fields over small time scales (within the same day) could account for some of the differences of these parameters between satellites on a given day as they cross the equator at different times, however these were not quantified in the current report. Data for the 2018 reprocessing of MODIS-Aqua (2003-present) and VIIRS-SNPP (2012-present) were downloaded from NASA's Ocean Biology Processing Group (OBPG; https://oceancolor.gsfc.nasa.gov) and subset to the *NWA* (Figure 1).
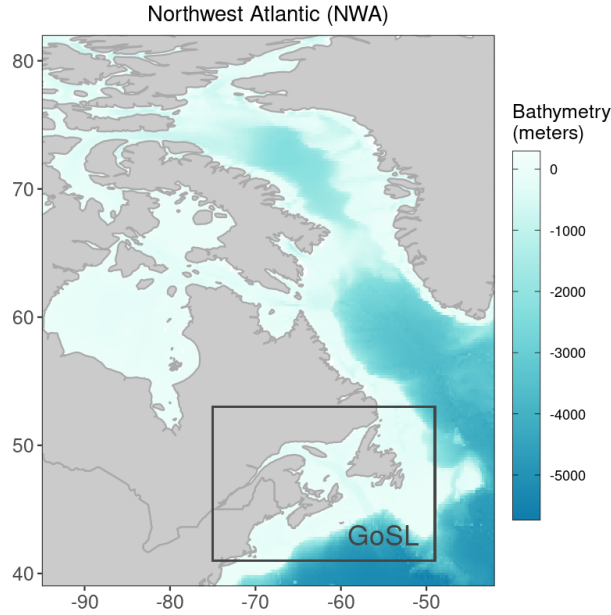
Figure 1. Map of the region of interest for the current study. The black box is drawn around the Gulf of Saint Lawrence subregion (*GoSL*, 49° - 75° W, 41° - 53° N) used exclusively for EOF *Chl-a*, described below in this section.

Similar SeaWiFS (1997-2010) L3b data is stored on NASA's servers only in 9km-resolution. In order to remain consistent with other sensors in this study, we downloaded SeaWiFS MLAC (Merged Local Area Coverage at 1.1 km resolution) level-2 files (individual satellite passes before spatial and temporal binning), and used the *l2bin* function from NASA's *OCSSW* software (https://oceancolor.gsfc.nasa.gov/docs/ocssw) to generate the L3b files at 4km-resolution. MODIS was selected as the baseline sensor for calibration of the other two sensors, as it overlaps the time series of both (Figure 2). VIIRS has the widest swath at 3000 km, while MODIS is the narrowest at 2330 km, and SeaWiFS is 2800 km wide.
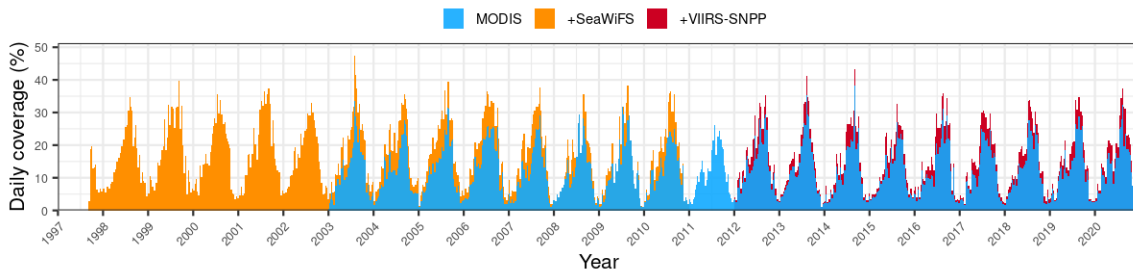

Figure 2. Daily percent coverage of MODIS (blue), plus the added percent coverage from SeaWiFS (orange) and VIIRS-SNPP (red) across the *NWA*, for the shortest visible *Rrs* band extracted from each sensor (410nm for VIIRS, 412nm for SeaWiFS and MODIS).

NASA's quality control flags filter out pixels with potential navigation, atmospheric correction, or *Chl-a* algorithm issues, as well as pixels that are saturated or have very low water-leaving

radiance, sun glint, or stray light from other sources. Pixels contaminated by aerosols, coccolithophore blooms, cloud, or ice are also removed by the flags, and values outside the range of 0-25000 sr^{-1} (for *Rrs*) and 0.001-100 mg m^{-3} (for *Chl-a*) were filtered out after downloading. The maximum allowed solar zenith angle was set to 75 degrees, limiting the amount of data that could be retrieved farther north and in the winter. Sensor viewing angles were limited to 60 degrees, but sensor-specific geometry could still account for differences in sensor values, which are not evaluated here.

Valid coincident pixels between each pair of sensors were collected from each daily image within the *NWA*, which contains a total of 295425 marine pixels in the level-3 binned images from NASA. The number of valid pixels varied between *Rrs* and *Chl-a* datasets depending on whether the *Chl-a* model could successfully retrieve a data value for a given pixel. We selected four *Chl-a* algorithms to use for the final sensor comparison, all of which use different combinations of *Rrs* bands as input:

1. NASA's globally-tuned empirical algorithm OCI, which is a combination of a band ratio model OCx (O'Reilly et al. 1998) and the Hu model (Hu et al. 2012) at lower concentrations,

2. The regionally-tuned band ratio model POLY4 (Clay et al. 2019),

3. The regionally-tuned semi-analytical algorithm GSM_GS (Maritorena et al. 2002; Clay et al. 2019), and

4. The empirical algorithm EOF (Laliberté et al. 2018), which has been trained for use only in the *GoSL*, and was therefore only applied to this subregion in the sensor comparison.

For the MODIS-SeaWiFS comparison and band modelling, datasets overlapped from 2003-2010, with significant sections of missing data in the time series in the last three years (see Figure 2, where the additional daily percent coverage from SeaWiFS drops, particularly in 2008 and 2009). The MODIS and VIIRS time series have overlapped since 2012. The number of valid coincident pixels across the time series between two sensors varied significantly across the map, concentrated more in coastal regions (Figure 3). High northern latitudes have very little overlap between SeaWiFS and MODIS, compared to VIIRS and MODIS. The lack of SeaWiFS data in later years of operation, particularly in Baffin Bay, results from the loss of a large subset of the full-resolution MLAC data between 2005 and 2010 due to storage issues, detailed here: https://forum.earthdata.nasa.gov/viewtopic.php?t=2261
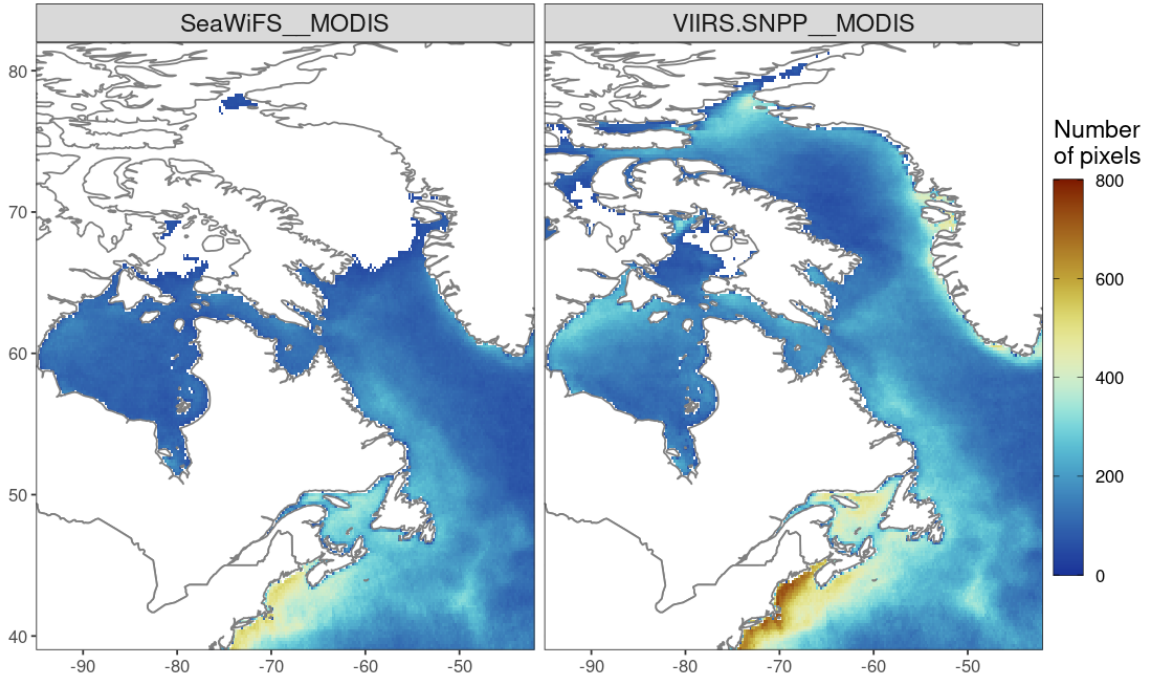
Figure 3. Map of the number of matching pixels between the two sensor pairs (SeaWiFS vs MODIS and VIIRS vs MODIS) across the time series for a given spatial point, using the *Rrs412* MODIS band and the closest corresponding bands from SeaWiFS and VIIRS.

For the band model, *Rrs* data were extracted from each pixel for all sensors and all visible bands: for MODIS, 412, 443, 469, 488, 531, 547, 555, 645, 667, and 678, all in nanometres; for SeaWiFS, 412, 443, 490, 510, 555, and 670nm; and for VIIRS, 410, 443, 486, 551, and 671nm. In tests where outliers were removed prior to modelling, an outlier was defined as a value in the 1st or 99th percentile of the ratio of SeaWiFS (or VIIRS) to MODIS, calculated separately for each variable. After removing unmatched, invalid, or outlying pixels, tens of millions of pixels still remained for sensor comparison and modelling purposes.

2.2   Modelling MODIS bands using SeaWiFS and VIIRS bands

The set of coincident pixels between SeaWiFS (or VIIRS) and MODIS were split into 80:20% training:test subsets. The training pixels were used to model each visible MODIS *Rrs* band as a multilinear regression of the *Rrs* from all the visible bands in either SeaWiFS or VIIRS:

$$R_m(i) = a_i R_s(412) + b_i R_s(443) + c_i R_s(490) + d_i R_s(510) + e_i R_s(555) + f_i R_s(670) \tag{1}$$

$$R_m(i) = n_i R_v(410) + o_i R_v(443) + p_i R_v(486) + q_i R_v(551) + r_i R_v(671) \tag{2}$$

R is shorthand for the *Rrs* of a particular sensor and band, subscripts *m*, *s*, and *v* denote MODIS, SeaWiFS, and VIIRS, respectively, and coefficients *a-f* (for SeaWiFS-modelled MODIS) and *n-r* (for VIIRS-modelled MODIS) are trained by the model and are specific to each MODIS band *i* that is being modelled.

5

We explored the effect of removing outliers or log10-transforming the *Rrs* prior to modelling. Note that zero-values were removed prior to logging the *Rrs*. If the data were not logged, negative predicted *Rrs* could potentially be generated, but were retained because small negative *Rrs* (i.e. >= -0.001) are allowed in the calculation of some *Chl-a* algorithms. To test the performance of the band models, the modelled MODIS from a given sensor (i.e. SeaWiFS or VIIRS) of the test set were linearly regressed against the true MODIS data, obtaining summary statistics for comparison. A simple correlation analysis was performed using the *cor.test()* function in R (R Core Team 2022) on the percent difference between true and modelled MODIS data and four variables that could potentially affect model accuracy: latitude, bathymetry, day of year, and year (i.e. the full time series of overlapping data).

The optimal model was defined as the model with statistics closest to the ideal value (i.e. slope = 1, intercept = 0, R^2 = 1, RMSE = 0). The resulting coefficients were used to create the modelled *Rrs* files for SeaWiFS and VIIRS at MODIS-specific wavebands. The new *Rrs* were then used in combination with the four *Chl-a* algorithms (OCI, POLY4, GSM_GS, and EOF, with MODIS-specific coefficients) to make the corresponding *Chl-a* datasets. SeaWiFS-modelled MODIS datasets and VIIRS-modelled MODIS datasets will be referred to as *S-MODIS* and *V-MODIS*, respectively.

2.3   Comparing modelled and true MODIS data to in situ data

To further evaluate the band models' performance, we conducted a linear regression of the logged MODIS, S-MODIS, and V-MODIS *Chl-a* against logged HPLC (High Performance Liquid Chromatography) *Chl-a* from in situ sample matchups within 10 metres of the surface between 2003 and 2020 (Devine et al. 2014; DFO 2023). Correlation tests were performed on the percent error between both S (or V)-MODIS and in situ data, and true MODIS and in situ data, against the same four variables as before (latitude, bathymetry, day of year, and year), as well as the distance between the center of the satellite pixel and the in situ sample location (for the matchups, the pixel nearest the in situ sample was selected). These tests were repeated using each of the four satellite *Chl-a* models (OCI, POLY4, GSM_GS, and EOF).

2.4   Calculating phytoplankton bloom metrics using S-MODIS and V-MODIS

The *Chl-a* datasets were formatted to be used in the PhytoFit application (Clay et al. 2021), which has the capability to model the annual Gaussian curve of the phytoplankton spring bloom given sufficient satellite data for a region. The app calculates metrics to summarize changes in the bloom patterns over time: the timing of the start and maximum concentration of the bloom, the duration, the amplitude (i.e. the maximum concentration), and the area under the Gaussian curve (i.e. the total phytoplankton mass over the bloom period). This bloom modelling procedure was run for the entire available time series for each original sensor (SeaWiFS, VIIRS, and MODIS) as well as S-MODIS and V-MODIS. This was repeated for several AZMP and AZOMP boxes of interest in the Labrador Sea, Gulf of Saint Lawrence (GoSL), and on the Scotian Shelf (SS, see Figure 4). For these tests, only the POLY4 *Chl-a* algorithm was selected to compute the spring bloom metrics (Clay et al. 2019), as it has has been tuned for use specifically in the *NWA* and found to outperform other common *Chl-a* models (i.e. OCI, GSM) in this region.
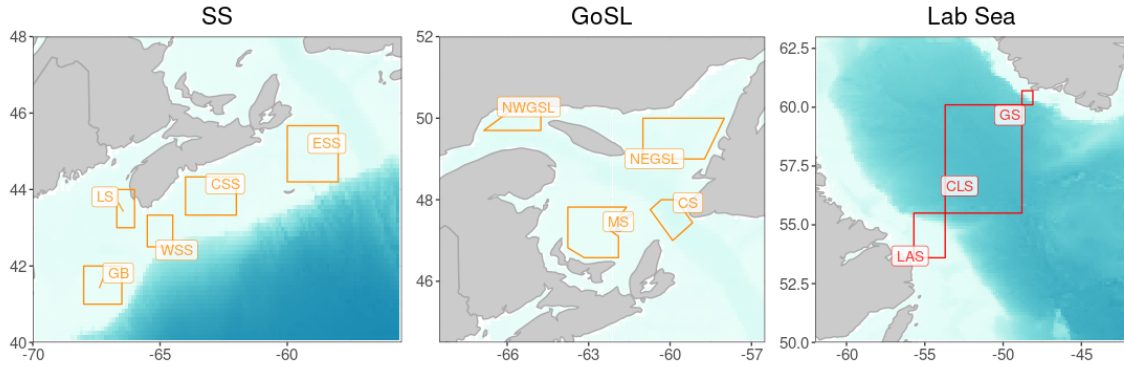
6

Figure 4. Map of the boxes that were used in the bloom fits for SS, GoSL, and the Labrador Sea. Orange boxes = "South", red boxes = "North".

The input parameters used for the bloom fits were part of two standard sets of parameters (for "North" and "South" boxes, divided by 52.5° N) that were determined in Devred et. al. (in preparation) to retrieve the highest number of valid bloom fits across the time series for each of the boxes in the region using MODIS data. A subset of several years for various boxes was fit by manually adjusting the range of days used in the model, as the standard set of parameters did not provide a valid fit. Bloom metrics were collected for each sensor in the different boxes, and presented in boxplots to examine the overall change in the distribution of metrics depending on sensor.

2.5   Differences between sensor data

For the sensor comparison, a subset of five *Rrs* bands in each sensor were used (see Table 1), as well as the four *Chl-a* algorithms listed in section 2.1.

Table 1. Nominal wavelengths (nm) of the five sensor bands used in the comparison for the three sensors of interest.

Band	MODIS	SeaWiFS	VIIRS
B1	412	412	410
B2	443	443	443
B3	488	490	486
B4	547	555	551
B5	667	670	671

The five bands used in the comparison differ slightly from sensor to sensor in terms of bands width and center, and Spectral Response Function (SRF, Figure 5), particularly in the "blue" and "green" part of the spectrum (B3-B4). These bands were selected for the comparison as their central wavelengths were most similar of all available bands.
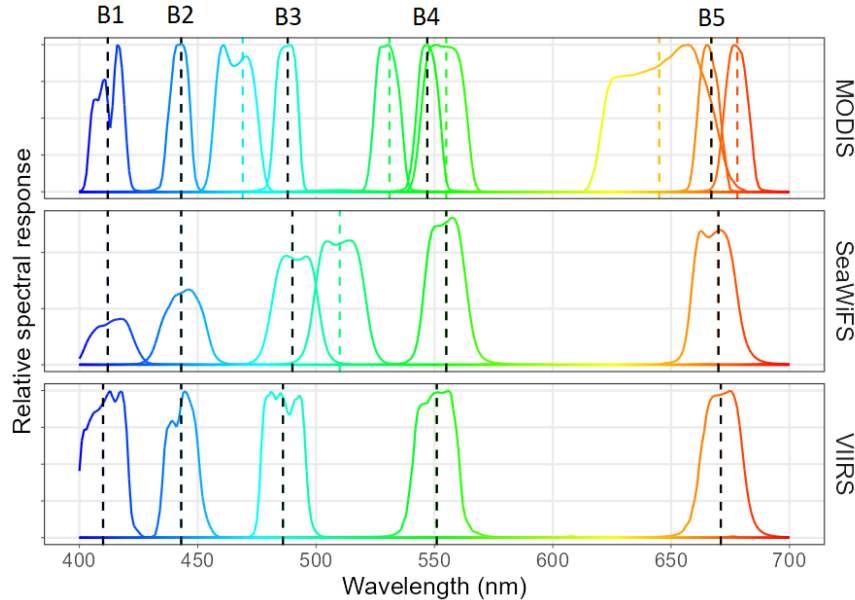
Figure 5. Relative SRFs for each of the sensors (https://oceancolor.gsfc.nasa.gov/resources/docs/rsr_tables), modelling their spectral response across the visible spectrum. Colours indicate the approximate colour of the band for a given wavelength. Coloured vertical dashed lines indicate the nominal wavelength of the band, while black vertical dashed lines indicate the nominal wavelengths of the five most similar bands that are used in the sensor comparison.

Boxplots were constructed to show the distribution of the percent difference between SeaWiFS (VIIRS) and MODIS for each variable of interest over the entire time series and region, before investigating the differences as a function of time or space, described in the following sections.

2.5.1 Temporal statistics

We removed outliers and log10-transformed *Chl-a* to more evenly distribute low and high concentrations, as *Chl-a* typically follows a lognormal distribution (Campbell 1995), meaning that the data are more likely to be normally distributed after the transformation. Subsequently, we summarized the differences across the time series between SeaWiFS (or VIIRS) and MODIS for each of the five most similar *Rrs* bands after removing outliers, as well as the log-transformed *Chl-a*. For a given pixel, all the SeaWiFS (or VIIRS) data across the time series was regressed against coincident MODIS data, leaving that pixel with a set of summary values: intercept, slope, R^2, p-value, and number of observations in the regression, as well as the median ratio of (unlogged) SeaWiFS (VIIRS) to MODIS. To test for spatial patterns in sensor differences, the pixel-wise values of the aforementioned statistics were mapped per variable, excluding pixels with fewer than 50 observations across time (see Figure 6). This comparison was repeated for S-MODIS and V-MODIS against true MODIS, to determine the effect of the band model on the patterns seen in previous sensor comparisons.
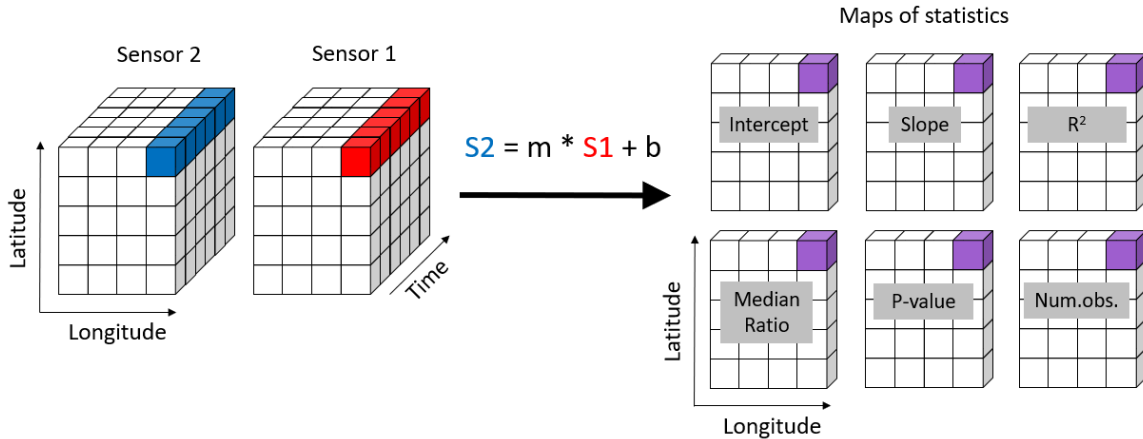
Figure 6. Diagram of the method used to summarize differences between sensors across time: Regressing all matching pixels across time (i.e. blue pixels for sensor 2 against red for sensor 1), giving that pixel a set of regression statistics (purple). After doing this for every pixel, the result was a map for each statistic.

2.5.2   Spatial statistics

Similar to the temporal statistics, we summarized the sensor differences across the region. For a given daily satellite image, SeaWiFS (VIIRS) data was regressed against MODIS data across space. The same statistics as described in Section 2.5.1 were then plotted across time in order to gauge temporal patterns in the sensor differences. Again, this was repeated for S-MODIS and V-MODIS against true MODIS.
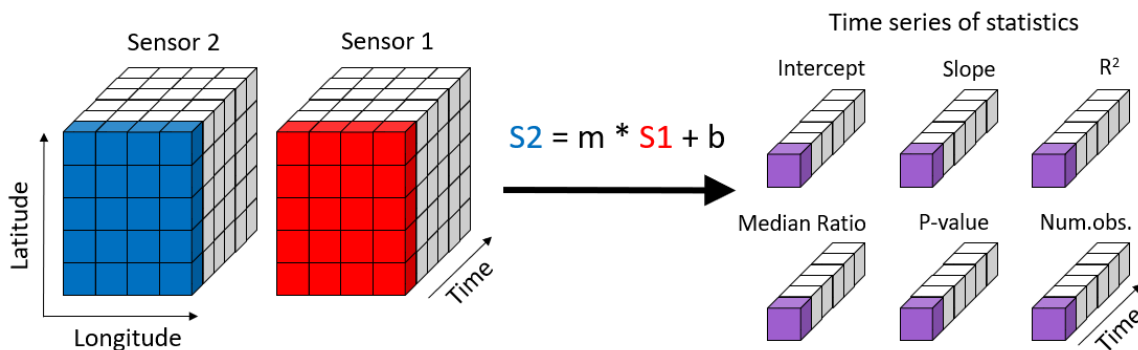

Figure 7. Diagram of the method used to summarize differences between sensors across space: Regressing all matching pixels over the map (i.e. blue against red), condensing that layer (i.e. the L3b daily image) into a single set of regression statistics (purple). This was repeated for each image, so the statistics could then be plotted as a function of time.

9

3 Results

3.1 Band model performance and improved coverage

Removing outliers or logging the *Rrs* did not improve the accuracy of the band model (see Figure A.1, which shows that the linear model that includes outliers produces the best statistics from the regression of modelled *Rrs* against true *Rrs*). Instead, it only decreased the size of our dataset by approximately 7.2% for SeaWiFS and 3.9% for VIIRS. Therefore, S-MODIS and V-MODIS datasets were generated using the band model trained with linear *Rrs* data that included outliers.

There was good alignment between the modelled and true *Rrs* in the test dataset for both S-MODIS and V-MODIS, which generally improved as wavelength increased (see Tables A.2 and A.4). Slopes were generally around one, bias around zero, and R^2 increased with wavelength except for the longest 2-3 modelled bands (645, 667, and 678nm) where there was a slight decrease. The p-values of the estimates for each waveband in the models were all statistically significant (<0.01), indicating that each band provided useful information in the model and could not be removed. *Rrs* and *Chl-a* uncertainties can vary with water composition, which can change significantly depending on time and location and negatively impact band model performance. Nevertheless, correlations between model percent error and latitude, bathymetry, day of year, or year were all negligible, with the exception of a slightly negative correlation with day of year for S-MODIS "green" bands (469-555nm).

Table 2 shows the increase in average daily percent coverage as a result of using MODIS data supplemented by SeaWiFS and VIIRS instead of one sensor alone. Coverage varies widely by season, with the highest increase from single sensor (MODIS) to Multisensor coverage in fall (+5.76% for MODIS/SeaWiFS, +3.54% for MODIS/VIIRS). Moreover, on 99% of the days when one sensor was unable to retrieve any data, the second sensor had an average percent coverage of 9.1%, up to a maximum of 34.9%.

Table 2. Average daily percent coverage in the NWA per sensor and season, as well as the average percent coverage of the two sensors combined (Multisensor). Seasons are defined as March-May (Spring), June-August (Summer), September-November (Fall), and December-February (Winter). Split by years 2003-2010 (MODIS-SeaWiFS overlap) and 2012-2020 (MODIS-VIIRS overlap). Coverage increase is the difference between Multisensor and MODIS coverage.

Years	Season	MODIS	SeaWiFS	VIIRS	Multisensor	Coverage increase
2003-2010	Spring	5.80	6.69		8.68	2.88
	Summer	11.63	12.64		17.38	5.75
	Fall	14.81	14.91		20.57	5.76
	Winter	2.30	2.5		3.68	1.38
2012-2020	Spring	5.31		5.38	6.74	1.43
	Summer	11.82		12.11	15.29	3.47
	Fall	14.70		15.06	18.24	3.54
	Winter	2.07		2.18	2.84	0.77

Figure 8 focuses on two randomly-selected days from the overlapping SeaWiFS-MODIS and VIIRS-MODIS time periods as an example of the spatial distribution of coverage between the individual sensor images and the combined image. There are significant regions of overlap between sensor data, highlighting the importance of ensuring consistency between datasets. The zoomed-in area of Figure 8 shows a region of overlapping data with discrepancies in the bands, easily visible due to the straight edge in the MODIS dataset. Despite these differences, the overall spatial patterns between sensors are very similar.
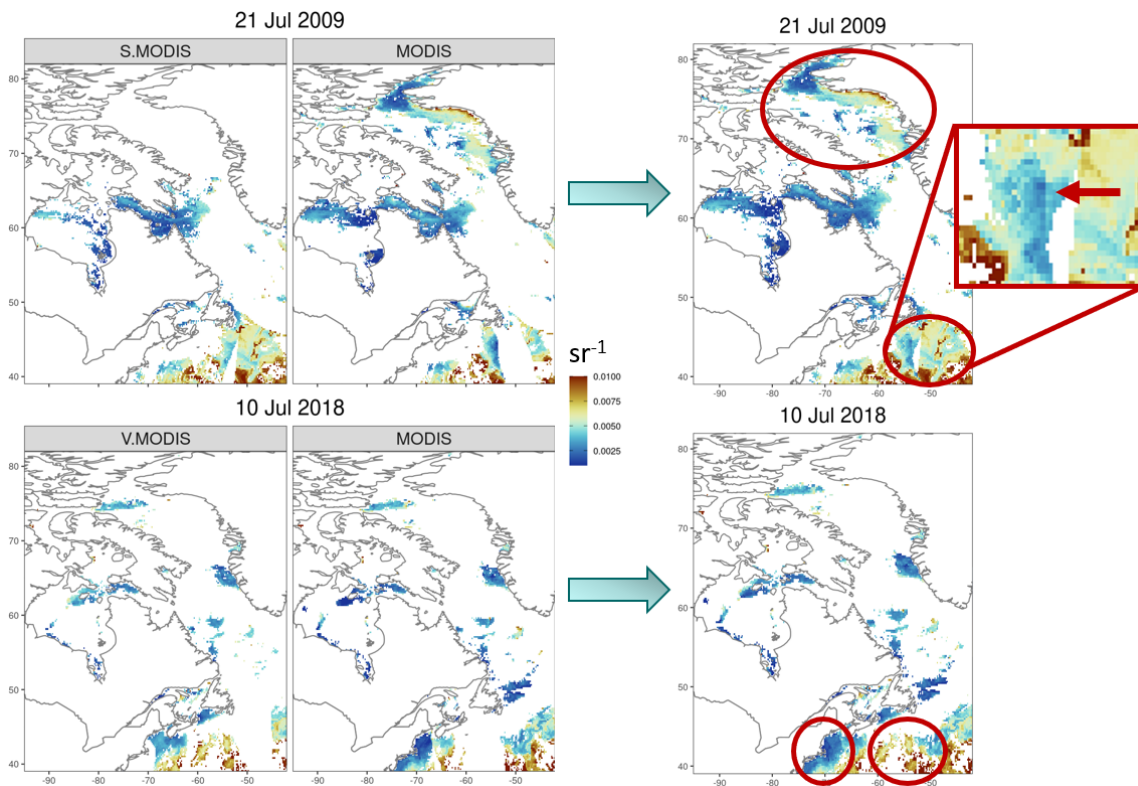

Figure 8. Daily *Rrs_412* maps from the individual sensors (left column) and the combined sensors (right), showing the increase in coverage for two separate days. Top row: S-MODIS and true MODIS; Bottom row: V-MODIS and true MODIS. Overlapping pixels were averaged. Red circles highlight sections of the combined image that were only present in one of the individual sensor images. The zoomed-in area to the right shows the visible line at the edge of the MODIS field overlapping the SeaWiFS field.

3.2   Band model impact on in situ matchups and spring bloom metrics

S-MODIS *Chl-a* in situ matchups had consistently lower regression slopes than the true MODIS matchups. This was paired with a higher and significant ($p<0.01$) correlation between percent difference (between satellite and in situ) and day of year than the true values for all but EOF *Chl-a* (Figure 9, and Appendix B). This indicates that S-MODIS *Chl-a* is less accurate during the fall bloom, moreso than true MODIS *Chl-a*. For POLY4 and GSM_GS *Chl-a*, the correlation between true MODIS/in situ percent difference and bathymetry was significant ($p<0.05$) and

11

higher (0.27 and 0.24, respectively) than that of S-MODIS. The dissimilarities suggest that these *Chl-a* algorithms are negatively impacted by shallow water, but the effect does not carry over to S-MODIS. In contrast, there were significant and positive correlations (>0.16) between bathymetry and both the true MODIS/in situ and V-MODIS/in situ percent differences using both POLY4 and GSM_GS *Chl-a*.
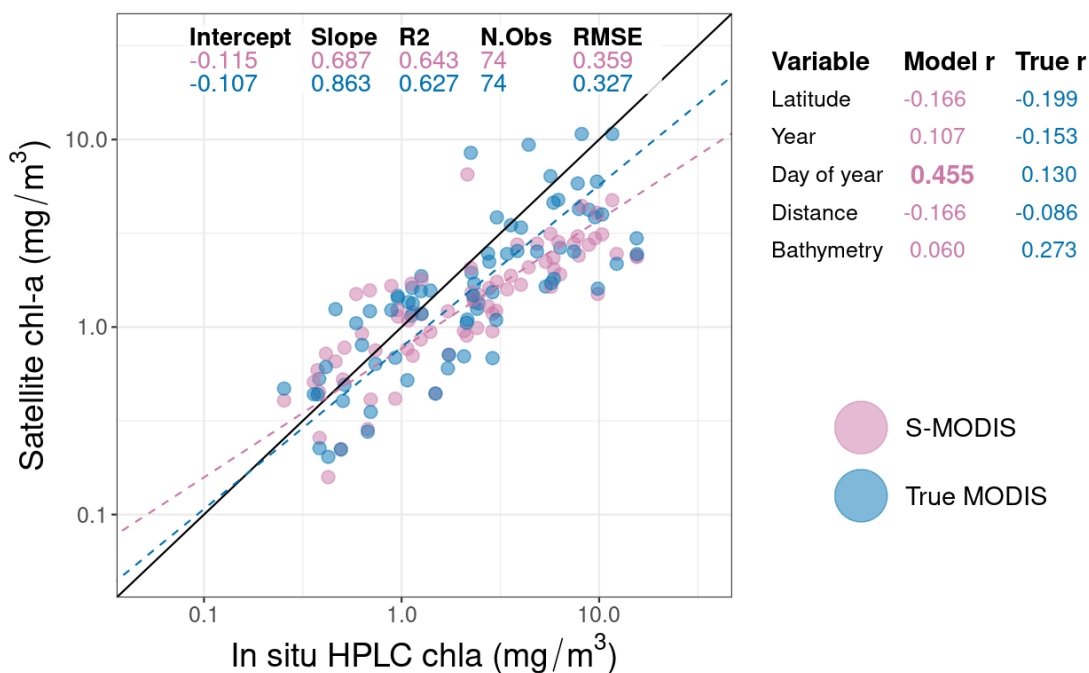

Intercept	Slope	R2	N.Obs	RMSE
-0.115	0.687	0.643	74	0.359
-0.107	0.863	0.627	74	0.327

Variable	Model r	True r
Latitude	-0.166	-0.199
Year	0.107	-0.153
Day of year	**0.455**	0.130
Distance	-0.166	-0.086
Bathymetry	0.060	0.273

Figure 9. Left: Matchups between S-MODIS (or true MODIS) POLY4 Chl-a and in situ samples, with regression statistics at the top of the plot. Right: Table of Pearson's correlation coefficients (*r*) of the percent difference between sensor pixel *Chl-a* and in situ *Chl-a* with each of the five variables in the first column. Only one coefficient was statistically significant (p<0.01), indicated in bold. Points are colored by sensor (S-MODIS in pink, true MODIS in blue).

V-MODIS *Chl-a* and true MODIS *Chl-a* had very similar relationships to in situ *Chl-a* with respect to their regression coefficients, with the exception of EOF *Chl-a*. V-MODIS EOF *Chl-a* tended to underestimate higher concentrations moreso than MODIS EOF *Chl-a*, leading to a lower regression slope against in situ *Chl-a* (Figure B.3). Unlike S-MODIS which had a higher correlation between percent difference and day of year than the true MODIS, the correlations between percent difference and the tested variables in V-MODIS were very similar to those in true MODIS. The number of in situ matchups from the overlapping SeaWiFS/MODIS or VIIRS/MODIS time series ranged from 66-74 (S-MODIS) and 216-234 (V-MODIS), which could have potentially affected the statistics. Figure B.1 shows the overlap of sample locations and satellite pixels across the time series, displaying the higher number of potential matchup pixels from V-MODIS.

The timings of the start and peak concentration of the bloom were generally consistent between sensors (Figure 10). Small differences in timing and concentration were magnified in the amplitude and magnitude (the total production over the bloom period), making these metrics

more sensitive to changes in sensor. In particular, bloom curve amplitudes calculated using S-MODIS were lower than the amplitudes calculated using either SeaWiFS or MODIS data, with a smaller spread of values.
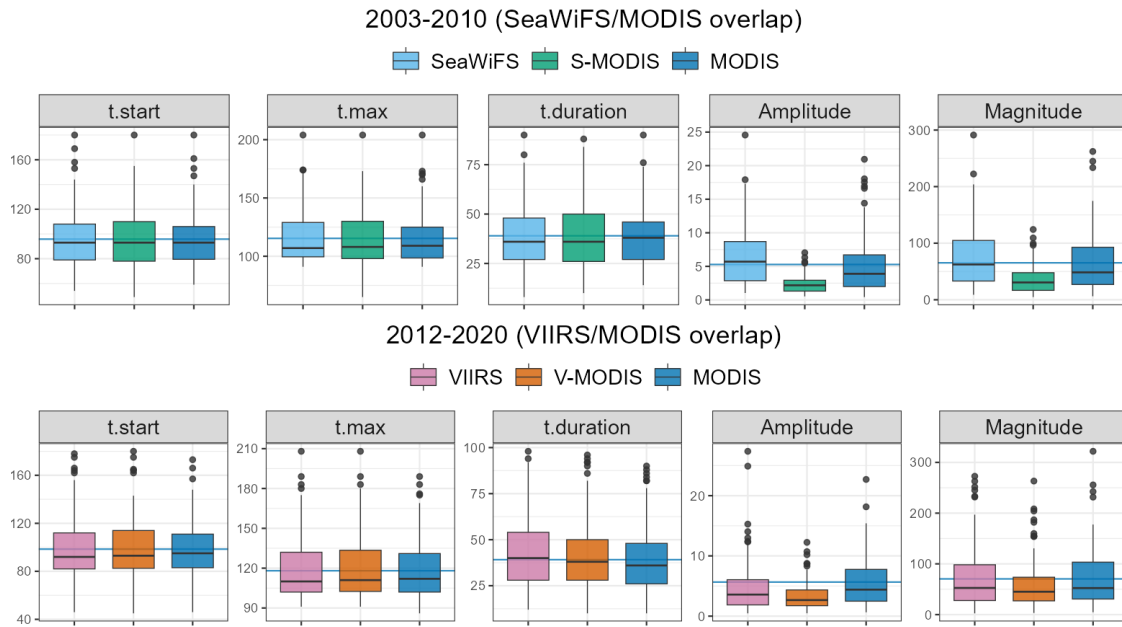

Figure 10. Boxplots of the phytoplankton spring bloom metrics calculated separately using SeaWiFS, MODIS, VIIRS, S-MODIS, and V-MODIS. The following bloom metrics were retrieved from the SS, GoSL, and Labrador Sea boxes: the day of year of the start of the bloom (*t.start*), day of maximum concentration (*t.max*), duration of the bloom (*t.duration*), the amplitude, i.e. peak *Chl-a* concentration within the bloom period (*Amplitude*), and the magnitude, i.e. the area under the daily *Chl-a* data points (*Magnitude*). Bloom fits were inspected manually for each year, box, and sensor, to remove poor fits before plotting. The only years/boxes used in the plot are those with valid fits for both SeaWiFS and MODIS, or VIIRS and MODIS. Blue horizontal line indicates the mean value of each metric from MODIS. Black horizontal line within each box indicates the median of that group.

The pattern of lower bloom amplitudes from S-MODIS data is observed in each of the three main regions (SS, GoSL, and the Labrador Sea, see Figure 11), but most prominent in SS and GoSL. This coincides with regions of lower slopes seen in the S-MODIS POLY4 maps of Figure C.2, suggesting that the peak *Chl-a* concentrations (i.e. those observed during spring blooms) were lower in the S-MODIS data than in true MODIS, and lower than the original SeaWiFS as well, which had very high slopes.
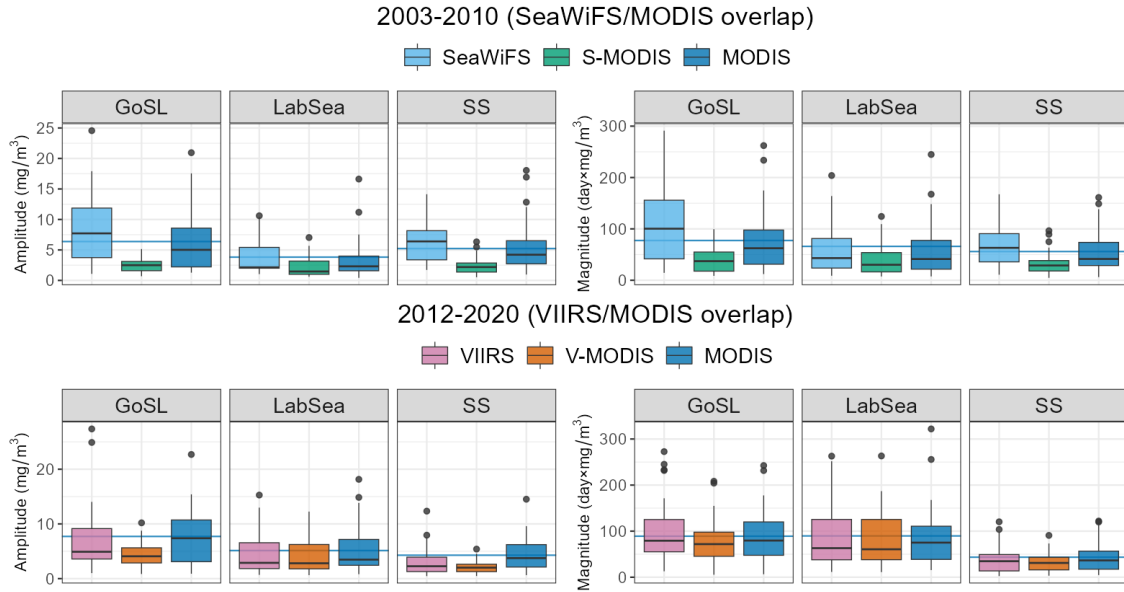
13

Figure 11. Amplitude and Magnitude bloom metrics from Figure 10, split by region. Blue horizontal line indicates the regional mean value of each metric from MODIS. Black horizontal lines within each box indicate the median value of that group.

3.3 Differences between sensors before and after alignment

The percent difference between SeaWiFS and MODIS varied widely, with the largest range of differences in the *Rrs* occurring at the shortest (blue) and longest (red) wavebands, and the *Chl-a* variables (particularly the band ratio algorithms OCI and POLY4, see Figure 12). VIIRS followed a similar pattern with smaller overall percent differences in the mid-range visible bands. On average, both SeaWiFS and VIIRS provided *Rrs* and *Chl-a* lower than MODIS. The median percent differences between S (or V)-MODIS and true MODIS were closer to zero across all variables, and each variable had approximately the same or reduced range of differences. Outliers (90th percentile) in the shortest *Rrs* band were more negatively affected by the model. This is consistent with the findings of larger overall error in the shortest band (https://oceancolor.gsfc.nasa.gov/data/reprocessing/r2018/aqua) and may result from imperfect atmospheric correction (Mobley et al. 2016).
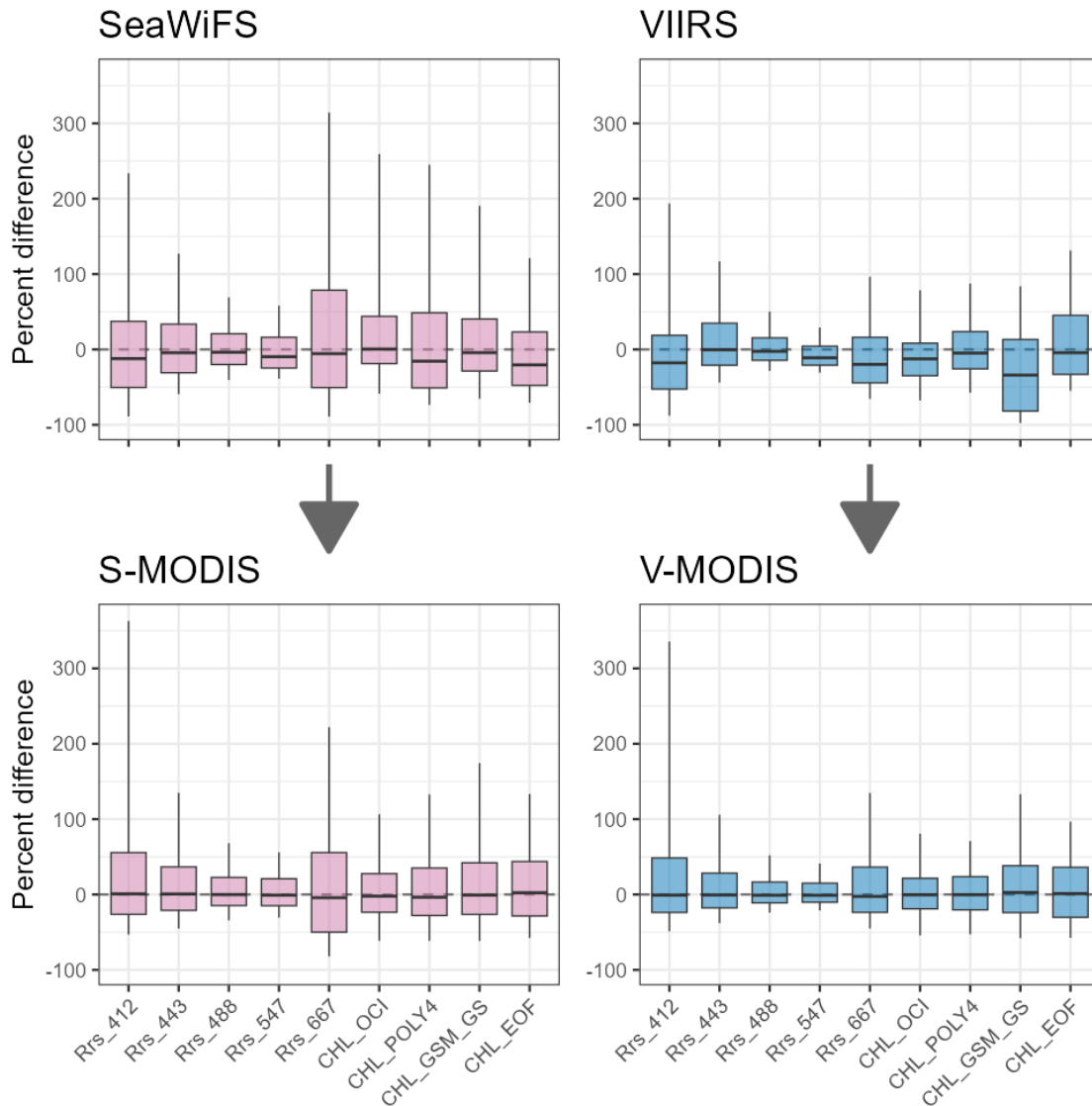
Figure 12. Boxplot of the percent difference between SeaWiFS (or VIIRS) and MODIS data for each variable, before and after applying the band model as indicated by the arrows. *Rrs* labels on the x-axis are the bands used in MODIS, S-MODIS, and V-MODIS (corresponding original bands from SeawiFS and VIIRS are given in Table 1). Boxes show the values between the 10th and 90th percentiles, and whiskers extend to the 1st and 99th percentiles (i.e. all the data used in the temporal and spatial statistics, and in the tested band models where outliers were removed). Horizontal lines in each box indicate the median of that group, and a dashed horizontal line is plotted at zero for reference.

Summary statistics of each pixel across the time series revealed noticeable patterns across the *NWA* (Appendix C). The similarity between sensors was typically linked to physical features in the ocean. Shelves and basins (e.g. Labrador Basin, Grand Banks, Gulf of Saint Lawrence) were often easily distinguishable in the maps (Appendix C), and usually had poorer alignment between sensors. The pixel-by-pixel linear regressions resulted in a vast majority (>99.99%) of pixels in

the *NWA* with statistically significant relationships between sensors.

Though SeaWiFS and VIIRS *Rrs* were found to be typically lower than or equal to MODIS (see *Rrs* medianRatio, Figures C.1 and C.3), particularly in shorter wavebands, the longest band (*Rrs667*) had large regression slopes in the Labrador Sea (Figure 13. Slopes were more variable pixel-to-pixel from the SeaWiFS/MODIS regression, but typically >1.2 in that region, whereas VIIRS/MODIS slopes were more consistent across space, with slopes ranging from 1.1-1.2 in the Labrador Basin. This shows that the red band reflectances in SeaWiFS and VIIRS typically had a greater range of values than that of MODIS in that region).
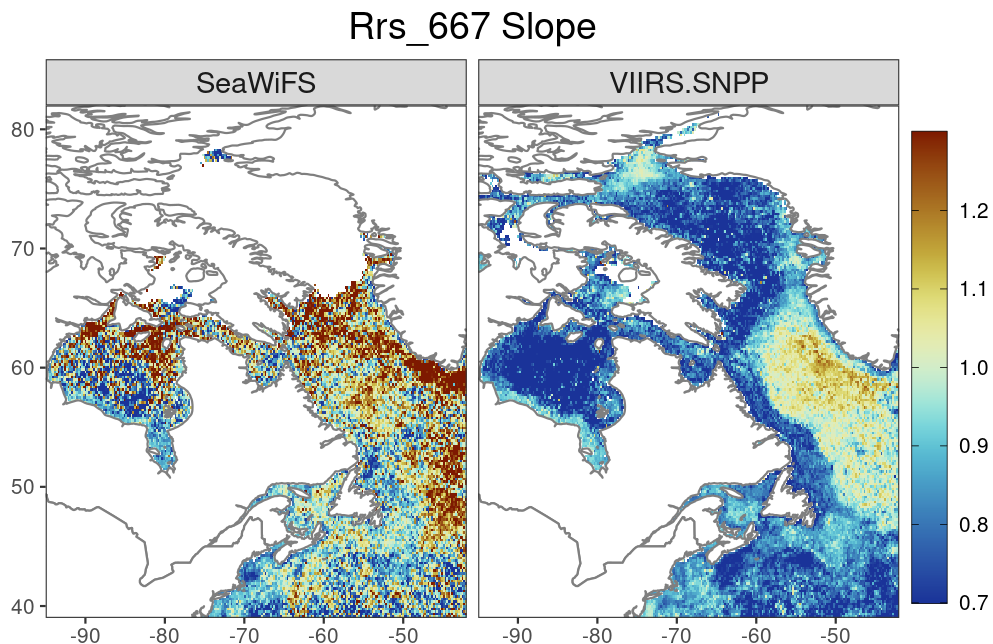

Figure 13. Pixel-by-pixel regression slopes of SeaWiFS (left) or VIIRS (right) *Rrs667* against MODIS across their respective overlapping time series, using the method described in Figure 6.

"Green" wavebands (i.e. ~488-547nm) generally had the best agreement in both SeaWiFS/MODIS and VIIRS/MODIS comparisons. Regions of low slope and high intercept reveal that SeaWiFS (VIIRS) reflectances are often overestimated relative to low MODIS reflectances and underestimated relative to high MODIS reflectances.

When looking at the relationships between *Chl-a* models from SeaWiFS and MODIS, EOF *Chl-a* was consistently lower in SeaWiFS than in MODIS, with low intercepts, slopes, and median ratios across the map. The OCI, POLY4, and GSM_GS algorithms were all largely uniform across the study region, with the exception of the Nova Scotia coastline, Gulf of Saint Lawrence, and the eastern side of Hudson Bay (Figure C.2). These areas have low slopes (<1) and high intercepts (>0), which indicates a lower gradient of SeaWiFS *Chl-a* concentrations than MODIS. Over the remainder of the region, OCI and GSM_GS are close in value between SeaWiFS and MODIS, yet the pattern diverges in the POLY4 *Chl-a* map which had consistently higher slopes (>1.1) paired with low median sensor ratios (<1). This suggests that while POLY4 *Chl-a* is generally

lower in SeaWiFS than in MODIS, the gradient of concentrations is steeper. Figure 14 shows the overall improvement in POLY4 *Chl-a* alignment between S-MODIS and MODIS. However, areas of low slope still remain, and have expanded over the Scotian Shelf.
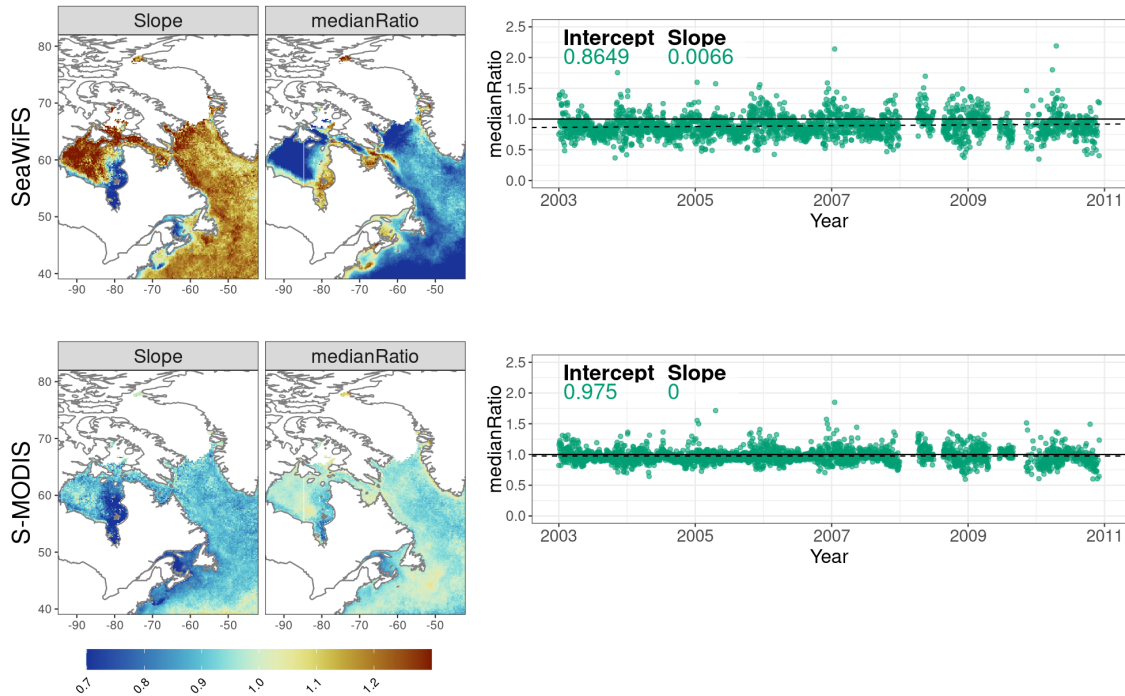

Figure 14. Top row: Map of the pixel-by-pixel regression slopes and median ratios of SeaWiFS against MODIS POLY4 *Chl-a*, and the time series of the daily median ratio of SeaWiFS:MODIS POLY4 *Chl-a*. Dashed line in the time series shows the regression of median ratio over time, with regression coefficients given in the plot. Bottom row: S-MODIS *Chl-a* against true MODIS *Chl-a*.

Both OCI and POLY4 VIIRS were more similar to MODIS than SeaWiFS, but GSM_GS *Chl-a* varied widely between the two sensors depending on physical ocean features, with low slopes along the coastal regions and high slopes in the Labrador Sea and Baffin Bay, as seen in the *Rrs667* VIIRS-MODIS map (Figure C.4).

Band model overcompensation resulted in higher median ratios in *Rrs* around the Maritimes (the Gulf of Saint Lawrence, Newfoundland, and Nova Scotia), meaning S (or V)-MODIS data values were generally too high in this area. This was particularly noticeable in the S-MODIS dataset, but less so in V-MODIS. Slopes of *Rrs* were lowered and intercepts raised, indicating underestimation of larger S(V)-MODIS reflectances relative to true MODIS and overestimation of smaller reflectances. This effect was more prominent in shorter wavebands and coastal regions, with the exception of the Labrador Sea which retained a patch of high slopes in both S-MODIS and V-MODIS in the *Rrs667* band. This bias propagated to *Chl-a*, though less so for VIIRS. The high slopes in VIIRS vs MODIS *Rrs667* translated to high GSM_GS *Chl-a* slopes in the original VIIRS-MODIS comparison, but this was corrected in V-MODIS GSM_GS (Figure 15).
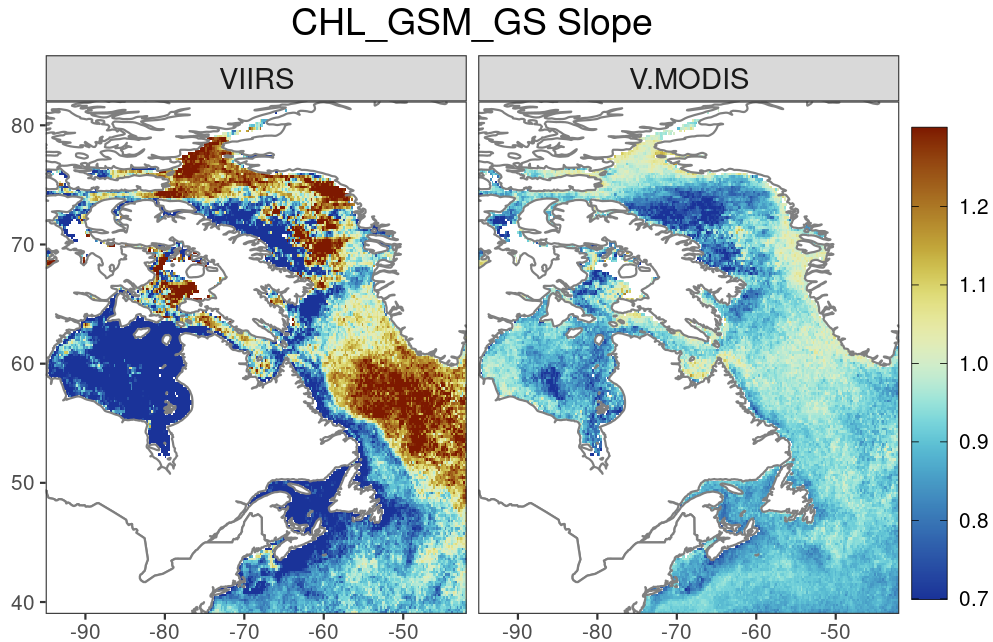
CHL_GSM_GS Slope

Figure 15. Left: Pixel-by-pixel regression slopes of VIIRS against MODIS GSM_GS *Chl-a*. Right: Using V-MODIS instead of VIIRS.

Sensor differences were found to be generally stable over time (see Appendix D), with a seasonal pattern of larger discrepancies and lower R^2 in the winter months when there were fewer data points available across the *NWA* due to low sun angle and cloud or ice cover (not shown). The exception is a pattern of divergence between some VIIRS and MODIS variables in recent years (see "medianRatio" plots from Figures D.3 and D.4). The daily median ratio of VIIRS to MODIS for *Rrs443* was found to be increasing over time, more noticeably than the change in the green-orange wavebands (see Figure 16), while the ratio in the shortest waveband has been decreasing. As the OCI *Chl-a* algorithm is allowed to switch between multiple potential blue-to-green band ratios and POLY4 does not use the 443nm band, neither were affected by this pattern. However, the median VIIRS:MODIS ratio of both GSM_GS and EOF diverged in recent years as both these algorithms use all wavebands and were therefore affected by the change. This shows that VIIRS GSM_GS has reduced to a small fraction of MODIS GSM_GS, and conversely, VIIRS EOF is now twice as high as MODIS EOF. The time series of *Chl-a* intercepts mirrored these patterns, but the slopes were relatively unaffected, suggesting an overall increase in the VIIRS/MODIS EOF ratio and decrease in the GSM_GS ratio across all *Chl-a* concentrations. The most notable improvement from the band model was in the removal of this divergence between VIIRS and MODIS (in particular, see the bottom panel of Figure 16, where EOF and GSM_GS *Chl-a* have realigned with MODIS during the last few years of operation).
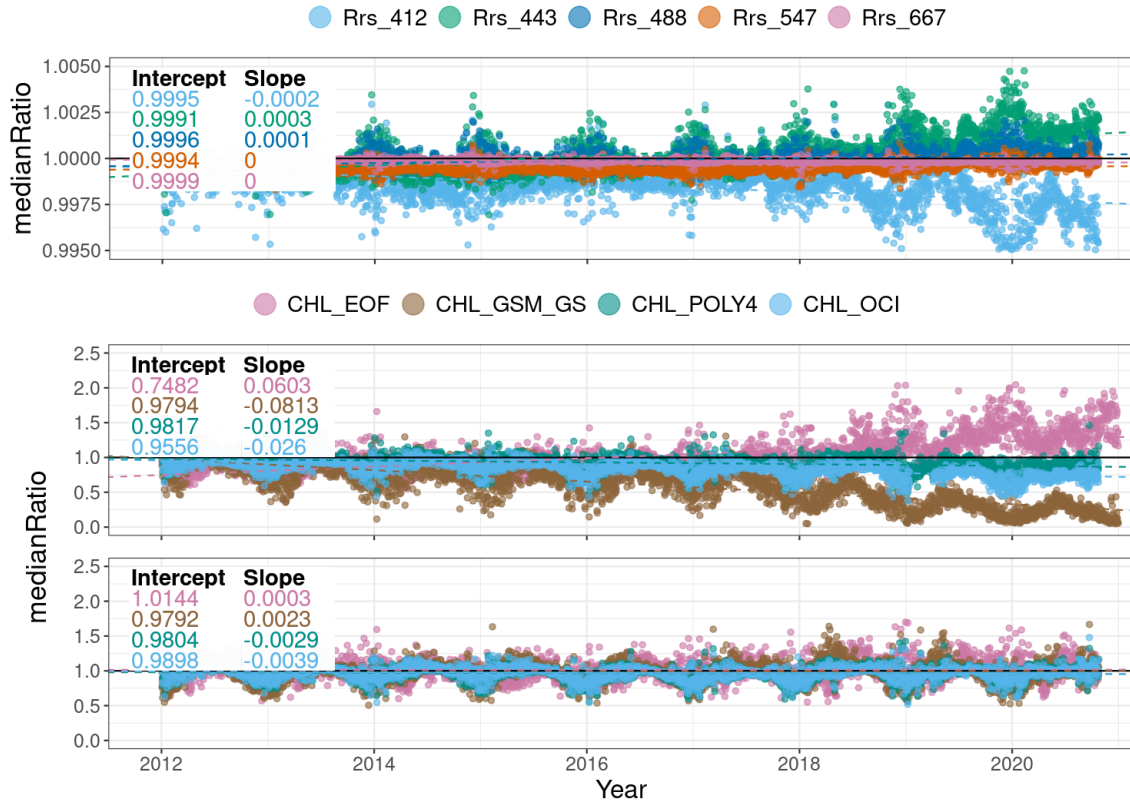
Figure 16. Top panel: Time series of median VIIRS:MODIS *Rrs* for each daily image. Middle panel: VIIRS:MODIS POLY4 *Chl-a*. Bottom panel: Median ratio of V-MODIS to true MODIS POLY4 *Chl-a*. These were created using the method described in Figure 7.

Shorter *Rrs* bands were more affected by seasonal oscillations in sensor differences. Gaps in the SeaWiFS time series were apparent in its last three years of operation, particularly in 2008-2009, but there were no large changes in the differences between SeaWiFS and MODIS (see Figure D.2). POLY4 *Chl-a* was found to have a consistently higher daily SeaWiFS-to-MODIS regression slopes and lower median SeaWiFS:MODIS ratios across the time series, similar to the patterns observed across the SeaWiFS:MODIS POLY4 *Chl-a* maps, which indicates a larger range of SeaWiFS *Chl-a* with lower values overall. EOF and GSM_GS *Chl-a* sensor ratios oscillate with season but follow opposite patterns, with the GSM_GS ratio highest in the winter months, coinciding with the opposite peaks and troughs in the seasonal patterns of the blue band *Rrs* sensor ratios. The VIIRS/MODIS EOF *Chl-a* time series behaved similarly to SeaWiFS/MODIS EOF *Chl-a*, but GSM_GS exhibited the opposite pattern to SeaWiFS/MODIS GSM_GS with its lowest median ratios in the winter months, in closer alignment with EOF *Chl-a*.

Overall the spatial and temporal comparison metrics showed better agreement between S(V)-MODIS and true MODIS compared to SeaWiFS/VIIRS and MODIS, with varying differences on smaller spatial and temporal scales dependent on *Rrs* band and *Chl-a* model (see Appendices C and D). These findings are in agreement with previous studies which found sensor differences to vary regionally and seasonally (Melin 2010; Barnes et al. 2021), with larger differences in coastal areas and higher latitudes such as the *NWA* (Djavidnia et al. 2010).

4   Discussion

4.1   Sensor differences and model performance

Although the band model brought both VIIRS and SeaWiFS into closer alignment with MODIS overall, with slightly better agreement using the VIIRS data, there is still significant spatial and temporal variation between the three sensors and the variables of interest. Shelves, coastlines, and the Gulf of Saint Lawrence are still easily distinguishable in many of the maps due to larger differences observed between the sensors (Appendix C). This suggests that splitting the band model into different regions, perhaps based on physical features or bathymetry, might reduce the differences. Seasonal oscillations in the time series of sensor ratios demonstrate that splitting the band model by season might also be beneficial (see Appendix D). However, this raises the question of how to blend the models across the boundaries of different regions or seasons, and would require further testing to ensure that it does not significantly increase the error between the modelled sensor and in situ validation samples.

Figure 8 provides a quick view of the combined data from two different sensors, demonstrating that while the overall patterns are consistent and coverage has improved, blending the remaining inconsistencies in the data fields will be a challenge. This is particularly problematic in regions where datasets overlap and one or more have a straight edge (for instance, at the edge of a satellite swath, or a region flagged by high sun glint). Another study uncovered this issue in the OC-CCI product as well, showing discontinuities between satellite scan lines that had propagated to *Chl-a* derived from the merged band products (Garnesson et al. 2019). Although the OC-CCI product is more complex than the band model described in this report, and applies bias correction after band modelling, they also use a simple average to blend overlapping reflectance data from multiple sensors. This contrast could potentially be reduced by experimenting with a similar bias-correction technique, as well as different blending methods, particularly at the edges of overlapping data fields.

High sensor zenith angles (>40 degrees) can cause noticeable variation in *Rrs* and *Chl-a* within a single sensor, with additional discrepancies between sensors at lower zenith angles as well (Barnes and Hu 2016). A potential improvement could be made in the methodology to collect coincident satellite pixels to use in model training by limiting them to Simultaneous Nadir Overpass (SNO) pixels (Cao et al. 2004), which can reduce the error attributed to differences in atmospheric path, solar and sensor geometry, and short-term time differences. However, this would require the use of level-2 satellite imagery (i.e. individual passes) to restrict the difference in satellite pass times, and could significantly limit the number of matching pixels. A compromise could be made by simply introducing stricter limits on geometry and timing than those used in the current study (Barnes et al. 2021). The final product will still contain residual errors due to varying sensor angles, which should be characterized in future work.

4.2   Sensor-specific derived metrics and in situ validation

The comparison of modelled and true MODIS *Chl-a* to in situ *Chl-a* revealed that V-MODIS is typically representative of the true MODIS data with respect to in situ matchups for each of

the *Chl-a* algorithms, with a slight increase in error for the EOF algorithm. This means that the use of MODIS *Chl-a* coefficients is appropriate for the V-MODIS *Rrs* (Figure B.2). S-MODIS, on the other hand, had an increased bias, resulting in a larger underestimation of high *Chl-a* concentrations than with the true MODIS *Chl-a* data against the in situ *Chl-a* (Figure B.3). For MODIS and in situ matchups, the percent error correlated with day of year as *Chl-a* estimate error increases during the fall months, but the correlation doubled in the case of S-MODIS and in situ percent error. The increase in error when using S-MODIS suggests the band model needs improvement, however the low number of in situ matchups in the S-MODIS comparison (ranging from 66-74) could account for a degree of error as opposed to the V-MODIS comparison, which had 216-234 matchups depending on *Chl-a* algorithm (see Figure B.1). In addition, the in situ matchups to the S-MODIS data were primarily located in areas of low slopes (Figures C.2 and B.1), with higher slopes located farther south, which explains the negative bias in the matchups.

Phytoplankton bloom timing was found to be very similar between each sensor, and as a result, the modelled MODIS data did not significantly change these metrics. However, small changes in the timing propagated to larger changes in the duration, and lower modelled MODIS *Chl-a* slopes, particularly in S-MODIS, resulted in lower bloom amplitudes and subsequently lower magnitudes (Figure 10). The significant decrease in amplitude in S-MODIS was a result of the alignment of POLY4 *Chl-a* between sensors (reduction of the high slopes in the original sensor comparison to MODIS, Figure C.2), which was used in the bloom metric calculations. Differences between the true MODIS metrics and S(V)-MODIS metrics could also be attributed to differences in their spatio-temporal coverage. However, a brief analysis of the differences in percent coverage in relation to bloom amplitudes showed that this was not an issue.

4.3   Recent divergence of VIIRS and MODIS

The divergence of the blue bands between MODIS and VIIRS beginning in 2018 was found to have a minor effect on the band ratio algorithms OCI and POLY4 since they do not use the shortest blue band, but the effect on EOF and GSM_GS was significant as they make use of all bands and are therefore affected by these patterns. However, as the EOF and GSM_GS models use the bands in different ways (i.e. GSM_GS attempts to fit the spectral shape to a polynomial, while EOF uses each of the bands in a Principal Component Analysis), they experienced opposite effects. The divergence of the bands also correlated with an increase in striping in VIIRS GSM_GS *Chl-a* images, which was not present in MODIS and fainter in the corresponding VIIRS OCI, POLY4, and EOF *Chl-a*. Figure 17 reveals that the striping effect (the diagonal patterns in the middle row) was reduced (but not completely removed) in the V-MODIS dataset.
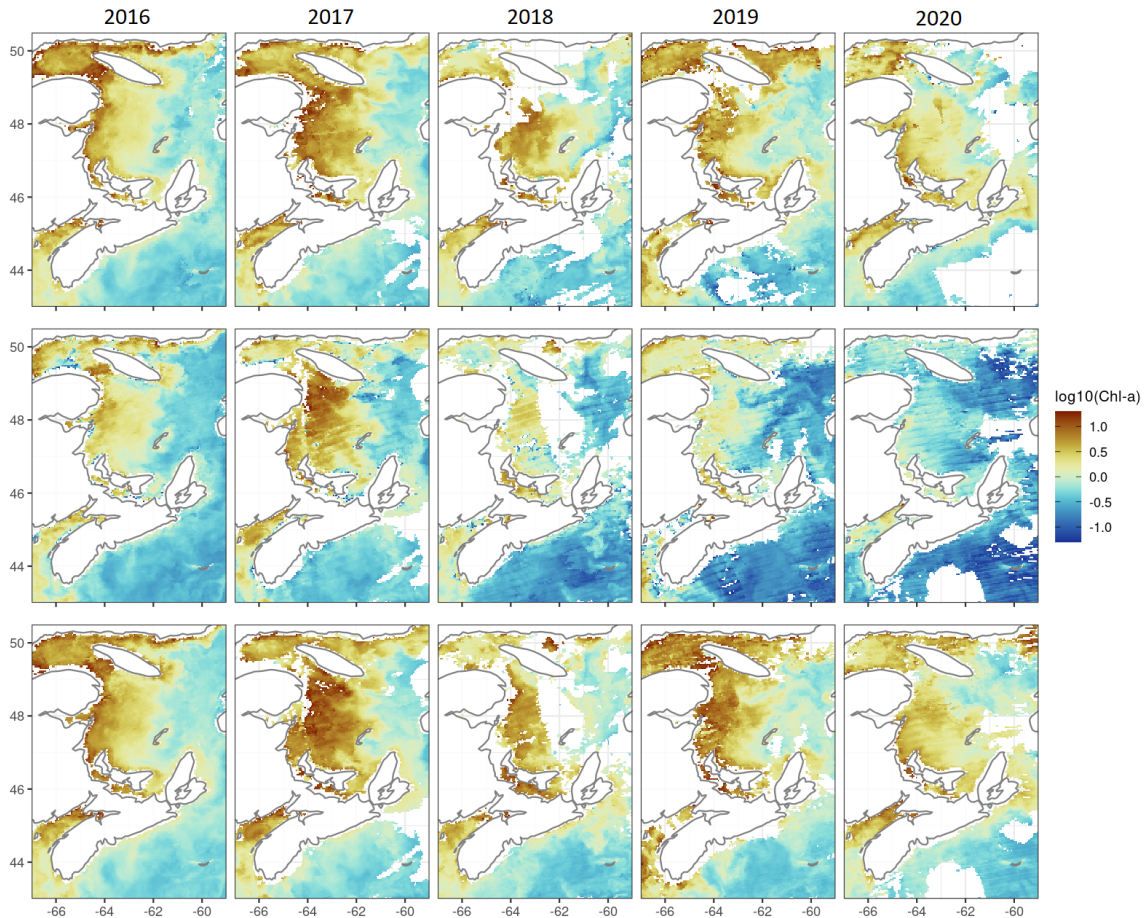
Figure 17. 8-day MODIS (top row), VIIRS (middle row), and V-MODIS (bottom row) log-transformed GSM_GS *Chl-a* images for the period from June 18 - June 25, from 2016-2020.

VIIRS has experienced gradual sensor degradation since its launch due to tungsten contamination on the mirrors, though this has mainly affected wavebands longer than the visible bands (Blonski and Cao 2013). Striping, particularly in shorter wavebands (Choi et al. 2022), is a known issue with a few potential causes; namely differences in the calibration, performance, or solar-sensor geometry of individual sensor detectors in the multi-detector sensor array, and differences in the reflectances on either side of the rotating mirror that directs light into the sensor (Mikelsons et al. 2014). Previous reprocessings of VIIRS at NASA OBPG have reduced striping and removed sensor artifacts. The most recent reprocessing (R2022.0, https://oceancolor.gsfc.nasa.gov/data/reprocessing/r2022/snpp), is expected to have helped correct the observed sensor artifacts in recent years that resulted in the striping and divergence of the two shortest bands relative to MODIS.

As POLY4 is the best-performing *Chl-a* model in the *NWA* (Clay et al. 2019), it is used in regular operations (e.g. AZMP and AZOMP annual bloom metrics), and therefore this divergence pattern has minimal negative effect on multisensor time series. However, EOF can be a better choice of *Chl-a* model in smaller regions using a training set composed of samples with a limited range of water properties, such as in the Gulf of Saint Lawrence (Laliberté et al. 2018). For this reason,

it might be preferable for use in studies confined to smaller spatial scales. GSM_GS has its own benefits over the empirical *Chl-a* models, as it offers a breakdown of the inherent optical properties of the water and is therefore required for studies investigating properties beyond *Chl-a* concentration. The *Chl-a* model must be selected depending on the study area and required target variables, with caution to the possible effects of this divergence between sensors if a multisensor time series will be used.

4.4   Conclusion

Overall, this band modelling method had the most positive results using VIIRS data (V-MODIS), as the SeaWiFS model (S-MODIS) tended to overcompensate depending on region. Internal consistency between sensors can offer information about long-term patterns, underlining the importance of correcting alignment between sensors. S-MODIS could benefit from further investigation and perhaps splitting the model by region or season and developing an approach to blend these submodels, for example applying sensor-specific weights. Both models could potentially be improved by implementing a bias-correction technique similar to OC-CCI, as well as limiting the pixels used in model training between different satellites to a smaller range of time differences, and solar and sensor zenith angles.

The method to tune the coefficients is relatively simple and quick to process, so they can be reoptimized after regular reprocessing of the raw satellite data by space agencies. MODIS-Aqua is nearing its end of life, so further investigation will be required into the use of other sensors as a baseline for sensor alignment. Additional sensors (e.g. OLCI-A and OLCI-B onboard the Sentinel-3A and 3B satellites) are a part of the regular processing chain at NASA OBPG, and will be added to the time series using the same multilinear regression band modelling method. The S-MODIS and V-MODIS bands will be merged with true MODIS bands to create a long time series of sensor data to improve spatial and temporal coverage. *Chl-a* will be derived from the final merged product and compared to ongoing in situ matchups collected during regular sampling cruises, as well as existing multi-sensor products such as those from GlobColour (`https://www.globcolour.info`) and OC-CCI. The resulting high-coverage long time series satellite product, tuned specifically for the *NWA*, will be key in examining climate signals and other long-term trends relating to fisheries stocks and phytoplankton bloom indices reported by the AZMP and AZOMP, as well monitoring existing MPAs.

4.5   Acknowledgements

5   References

Alvera-Azcárate, A., Barth, A., Sirjacobs, D., and Beckers, J.-M. 2009. Enhancing temporal correlations in EOF expansions for the reconstruction of missing data using DINEOF. Ocean Science (OS) 5.

Azetsu-Scott, K., Head, E., Harrison, G., Hendry, R., Li, W., Loder, J., Vezina, A., Wright, D., Yashayaev, I., and Yeats, P. 2008. Climate change in the northwest atlantic. pp. 4–11.

Barnes, B.B., and Hu, C. 2016. Dependence of satellite ocean color data products on viewing angles: A comparison between SeaWiFS, MODIS, and VIIRS. Remote Sens. Environ. 175: 120–129. Elsevier Inc.

Barnes, B.B., Hu, C., Bailey, S.W., Pahlevan, N., and Franz, B.A. 2021. Cross-calibration of MODIS and VIIRS long near infrared bands for ocean color science and applications. Remote Sens. Environ. 260(August 2020). Elsevier Inc.

Belward, A., Bourassa, M., Dowell, M., briggs, stephen, Dolman, H.(A.J.)., Holmlund, K., Husband, R., Quegan, S., Simmons, A., Sloyan, B., Stammer, D., Steffen, K., Tanhua, T., Verstraete, M., Richter, C., Eggleston, S., Hill, K., Tassone, C., Mora, B., and Saatchi, S. 2016, December. The global observing system for climate: Implementation needs.

Blonski, S., and Cao, C. 2013. Monitoring and predicting rate of VIIRS sensitivity degradation from telescope contamination by tungsten oxide. Sensors Syst. Sp. Appl. VI 8739(May): 87390D.

Campbell, J. 1995. The lognormal distribution as a model for bio-optical variability in the sea. Journal of Geophysical Research Atmospheres 100.

Cao, C., Weinreb, M., and Xu, H. 2004. Predicting simultaneous nadir overpasses among polar-orbiting meteorological satellites for the intersatellite calibration of radiometers. J. Atmos. Ocean. Technol. 21(4): 537–542.

Choi, T., Cao, C., Shao, X., and Wang, W. 2022. S-NPP VIIRS Lunar Calibrations over 10 Years in Reflective Solar Bands (RSB). Remote Sens. 14(14).

Clay, S., Layton, C., and Devred, E. 2021. PhytoFit. `https://github.com/BIO-RSG/PhytoFit`.

Clay, S., Peña, A., DeTracey, B., and Devred, E. 2019. Evaluation of satellite-based algorithms to retrieve chlorophyll-a concentration in the canadian atlantic and pacific oceans. Remote Sens. 11(22).

Devine, L., Kennedy, M.K., St. Pierre, I., Lafleur, C., Ouellet, M., and Bond, S. 2014. BioChem: the Fisheries and Oceans Canada Database for Biological and Chemical Data. Can. Tech. Rep. Fish. Aquat. Sci. 3073: iv + 40.

DFO. 2023. BioChem: Database of biological and chemical oceanographic data. `http://www.dfo-mpo.gc.ca/science/data-donnees/biochem/index-eng.html`.

Djavidnia, S., Mélin, F., and Hoepffner, N. 2010. Comparison of global ocean colour data records. Ocean Sci. 6(1): 61–76.

Garnesson, P., Mangin, A., Fanton d'Andon, O., Demaria, J., and Bretagnon, M. 2019. The CMEMS GlobColour chlorophyll a product based on satellite observation: Multi-sensor merging and flagging strategies. Ocean Science 15(3): 819–830.

Hu, C., Lee, Z., and Franz, B. 2012. Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. Journal of Geophysical Research 117: C01011.

IOCCG. 2004. Guide to the creation and use of ocean-colour, level-3, binned data products. *Edited by*D. Antoine. IOCCG, Dartmouth, Canada.

Laliberté, J., Larouche, P., Devred, E., and Craig, S. 2018. Chlorophyll-a concentration retrieval in the optically complex waters of the St. Lawrence Estuary and Gulf using principal component analysis. Remote Sensing 10(2).

Maritorena, S., Siegel, D.A., and Peterson, A.R. 2002. Optimization of a semi-analytical ocean color model for global-scale applications. Appl. Opt 41: 2705–2713.

Melin, F. 2010. Global distribution of the random uncertainty associated with satellite-derived chl a. Geoscience and Remote Sensing Letters, IEEE 7: 220–224.

Mikelsons, K., Wang, M., Jiang, L., and Bouali, M. 2014. Destriping algorithm for improved satellite-derived ocean color product imagery. Opt. Express 22(23): 28058–28070. Optica Publishing Group.

Mobley, C.D., Werdell, J., Franz, B., Ahmad, Z., and Bailey, S. 2016. Atmospheric correction for satellite ocean color radiometry. A Tutor. Doc. NASA Ocean Biol. Process. Gr. (June): 1–73.

O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., and McClain, C. 1998. Ocean color chlorophyll algorithm for SeaWiFS. J. Geophys. Res. 103: 24, 937–24, 953.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ringuette, M., Devred, E., Azetsu-Scott, K., Head, E., Punshon, S., Casault, B., and Clay, S. 2022. Optical, Chemical, and Biological Oceanographic Conditions in the Labrador Sea between 2014 and 2018. DFO Can. Sci. Advis. Sec. Res. Doc. 2022(021): v+38p.

Sathyendranath, S., Brewin, B., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., Cipollini, P., Couto, A., Dingle, J., Doerffer, R., Donlon, C., Dowell, M., Farman, A., Grant, M., Groom, S., Horseman, A., Jackson, T., Krasemann, H., Lavender, S., and Platt, T. 2019. An ocean-colour time series for use in climate studies: The experience of the ocean-colour climate change initiative (OC-CCI). Sensors 19: 4285.

Therriault, J.-C., Petrie, B., P. Pépin, J.G., Gregory, D., Helbig, J., Herman, A., Lefaivre, D., Mitchell, M., Pelchat, B., RUNGE, J., and SAMEOTO., D. 1998. Proposal for a northwest atlantic zonal monitoring program. *In* Can. Tech. Rep. Hydrogr. Ocean sci. DFO. p. vii+57.

Trzcinski, M., Devred, E., Platt, T., and Sathyendranath, S. 2013. Variation in ocean colour may help predict cod and haddock recruitment. Marine Ecology Progress Series 491: 187–197.

Wang, M., Jiang, L., Son, S., Liu, X., and Voss, K. 2020. Deriving consistent ocean biological and biogeochemical products from multiple satellite ocean color sensors. Optics Express 28.

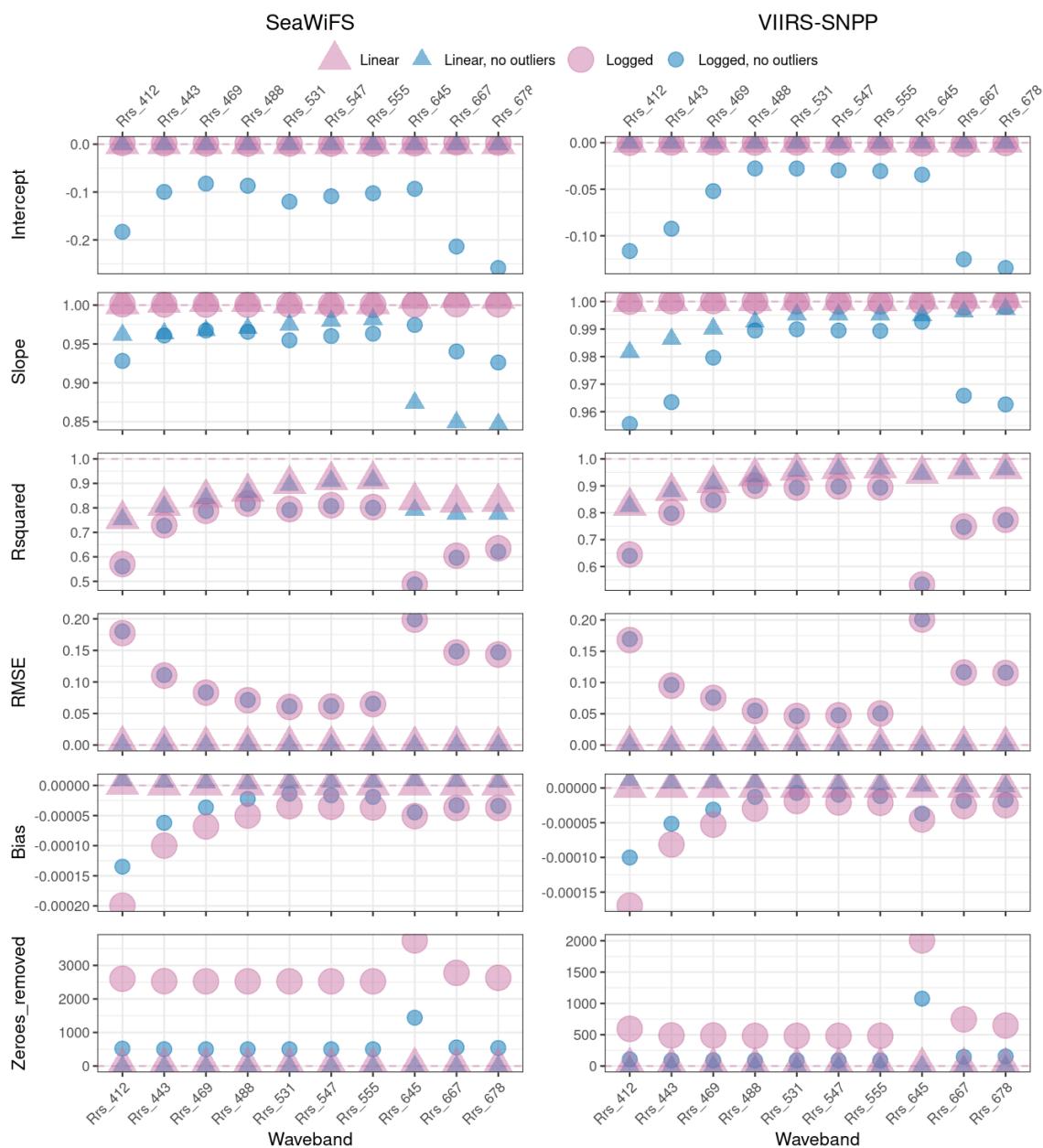# APPENDIX A   Band model evaluation

Figure A.1. Statistics of the linear regression of S (or V)-MODIS vs true MODIS data, using the test set. Note the "ideal" value for each statistic is marked with a horizontal dashed red line for easier comparison. P-value plots are not presented because they were all found to be statistically significant (<0.01). Bias is the average difference between true and modelled values. "Zeroes_removed" is the number of zero-values (or in some cases, negative values) that had to be removed from the training and test sets before log10-transforming values for the logged fits.
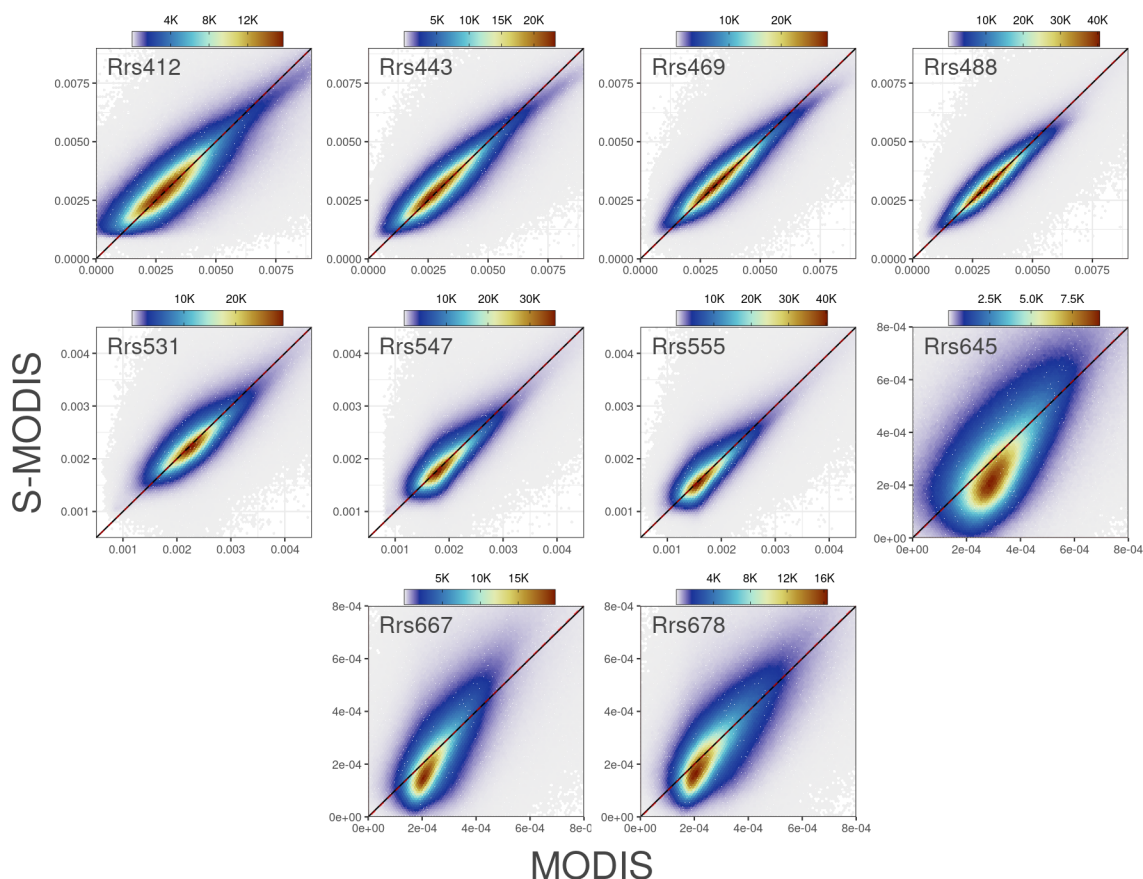
A.1 Band models using SeaWiFS data

Figure A.2. Density plots of the S-MODIS bands from the test set against the corresponding true MODIS bands. "K" on the legend axis represents x1000. Dashed red lines in the plots indicate the linear relationship between modelled and true sensor values. MODIS waveband is indicated at the top left of each plot. Note these models were performed without removing outliers or logging the input data.

Table A.1. Table of optimal coefficients derived using SeaWiFS Rrs bands (*) to model each MODIS band (+).

	Intercept	Rrs412*	Rrs443*	Rrs490*	Rrs510*	Rrs555*	Rrs670*
Rrs412+	0.001	0.398	0.216	1.086	-1.215	0.240	-0.035
Rrs443+	0.001	0.066	0.197	1.391	-1.048	0.159	-0.039
Rrs469+	0.001	-0.094	0.129	1.549	-0.902	0.174	-0.055
Rrs488+	0.000	-0.175	0.022	1.487	-0.628	0.214	-0.115
Rrs531+	0.000	-0.115	-0.129	0.566	-0.089	0.741	-0.175
Rrs547+	0.000	-0.105	-0.116	0.449	-0.231	0.953	-0.116
Rrs555+	0.000	-0.095	-0.098	0.397	-0.314	0.998	-0.074

Continued on next page ...

28

	Intercept	Rrs412*	Rrs443*	Rrs490*	Rrs510*	Rrs555*	Rrs670*
Rrs645+	0.000	-0.061	0.006	0.234	-0.462	0.475	0.526
Rrs667+	0.000	-0.040	0.002	0.195	-0.395	0.379	0.507
Rrs678+	0.000	-0.020	-0.006	0.141	-0.364	0.381	0.492

Table A.2. Table of statistics calculated for S-MODIS against MODIS data. RrsLT0 gives the percentage of data points in the test set for which the model predicted negative Rrs values. Regression intercepts (and bias) were all within 1e-4 (1e-6) from zero. Pearson's correlation coefficients were derived by testing percent error and latitude ($r_{latitude}$), bathymetry ($r_{bathymetry}$), day of year (r_{DOY}), and year (r_{year}). * = p < 0.01.

| Band	Slope	R^2	RMSE	RrsLT0	$r_{latitude}$	$r_{bathymetry}$	r_{DOY}	r_{year}
Rrs412	1.001	0.755	0.001	0.006	0	0.001	-0.001	0
Rrs443	1.003	0.808	0.001	0.006	-0.016*	0.055*	-0.091*	-0.015*
Rrs469	1.004	0.842	0.001	0.006	-0.024*	0.072*	-0.17*	0.049*
Rrs488	1.004	0.865	0.001	0.009	-0.009*	0.084*	-0.163*	-0.012*
Rrs531	1.001	0.897	0.000	0.009	-0.01*	0.064*	-0.214*	-0.01*
Rrs547	1.001	0.913	0.000	0.010	-0.009*	0.076*	-0.172*	0.011*
Rrs555	1.001	0.915	0.000	0.010	-0.026*	0.062*	-0.197*	-0.03*
Rrs645	1.008	0.829	0.000	0.946	-0.004*	-0.004*	0.001*	0
Rrs667	1.009	0.821	0.000	0.584	-0.001	0.001*	0	0
Rrs678	1.008	0.824	0.000	0.342	0	0.001*	0	0

A.2 Band models using VIIRS data
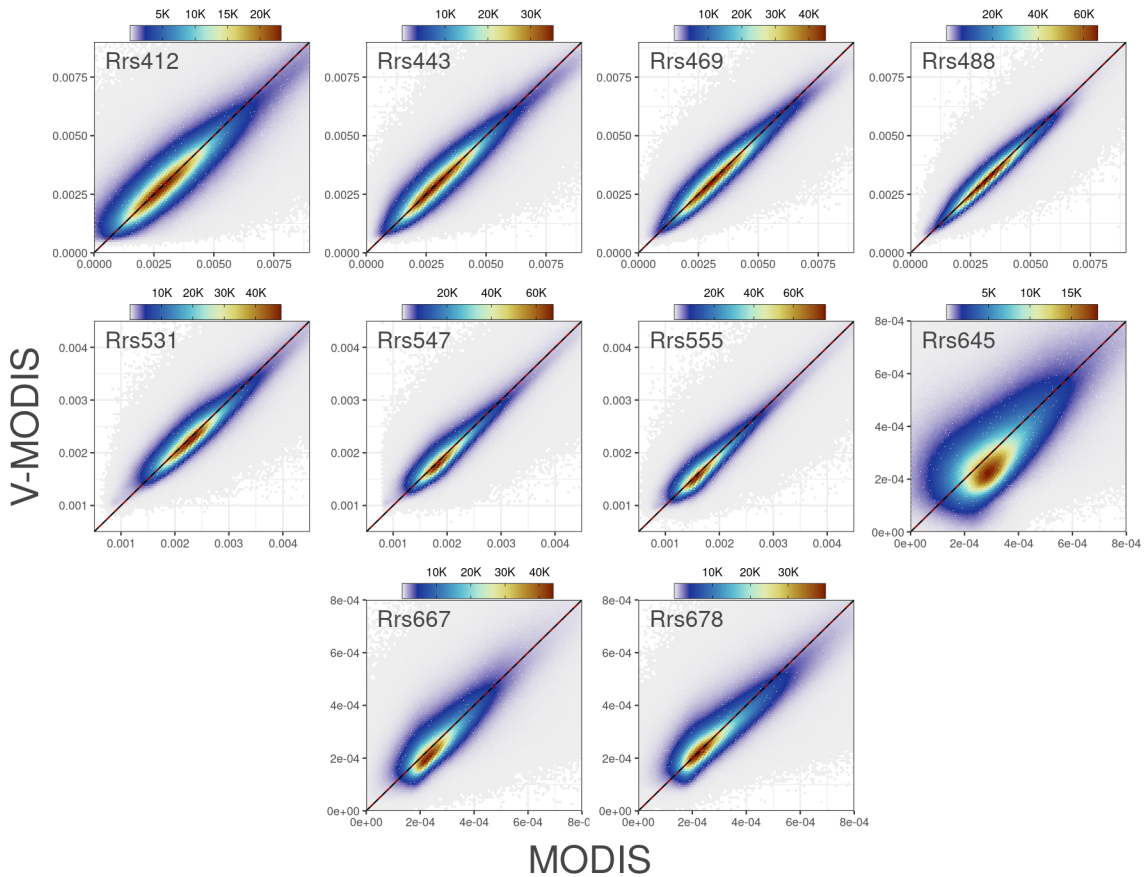
Figure A.3. Density plots of the V-MODIS bands from the test set against the corresponding true MODIS bands. "K" on the legend axis represents x1000. Dashed red lines in the plots indicate the linear relationship between modelled and true sensor values. MODIS waveband is indicated at the top left of each plot. Note these models were performed without removing outliers or logging the input data.

Table A.3. Table of optimal coefficients derived using VIIRS Rrs bands (*) to model each MODIS band (+).

	Intercept	Rrs410*	Rrs443*	Rrs486*	Rrs551*	Rrs671*
Rrs412+	0	0.439	0.835	-0.304	-0.006	0.045
Rrs443+	0	0.369	-0.036	0.797	-0.299	0.211
Rrs469+	0	0.270	-0.387	1.275	-0.293	0.216
Rrs488+	0	0.094	-0.360	1.280	-0.068	0.042
Rrs531+	0	0.024	-0.231	0.413	0.843	-0.177
Rrs547+	0	0.026	-0.173	0.184	0.966	-0.037
Rrs555+	0	0.040	-0.162	0.101	0.943	0.059

Continued on next page ...

	Intercept	Rrs410*	Rrs443*	Rrs486*	Rrs551*	Rrs671*
Rrs645+	0	-0.034	0.028	-0.014	0.064	1.079
Rrs667+	0	-0.035	0.051	-0.020	0.007	1.027
Rrs678+	0	-0.022	0.059	-0.071	0.035	0.989

Table A.4. Table of statistics calculated for V-MODIS against MODIS data. RrsLT0 gives the percentage of data points in the test set for which the model predicted negative Rrs values. Regression intercepts and bias were all within 1e-6 from zero. Pearson's correlation coefficients were derived by testing percent error and latitude ($r_{latitude}$), bathymetry ($r_{bathymetry}$), day of year (r_{DOY}), and year (r_{year}). * = $p < 0.01$.

Band	Slope	R^2	RMSE	RrsLT0	$r_{latitude}$	$r_{bathymetry}$	r_{DOY}	r_{year}
Rrs412	1.000	0.825	0.001	0.000	-0.001	0.001*	0	0
Rrs443	1.000	0.879	0.001	0.000	-0.005*	0.006*	-0.008*	0.011*
Rrs469	1.000	0.909	0.001	0.000	0	0	0	0
Rrs488	1.000	0.936	0.000	0.000	-0.073*	0.001	-0.032*	0.005*
Rrs531	1.000	0.956	0.000	0.000	-0.117*	-0.044*	-0.107*	-0.013*
Rrs547	1.000	0.963	0.000	0.000	-0.086*	-0.006*	-0.045*	-0.016*
Rrs555	1.000	0.964	0.000	0.000	-0.081*	0.002*	-0.026*	0.014*
Rrs645	1.001	0.944	0.000	0.064	-0.004*	-0.004*	0.001*	0.001*
Rrs667	1.001	0.961	0.000	0.000	-0.001*	0.001	0	0
Rrs678	1.001	0.961	0.000	0.000	-0.001	0.001	0	0

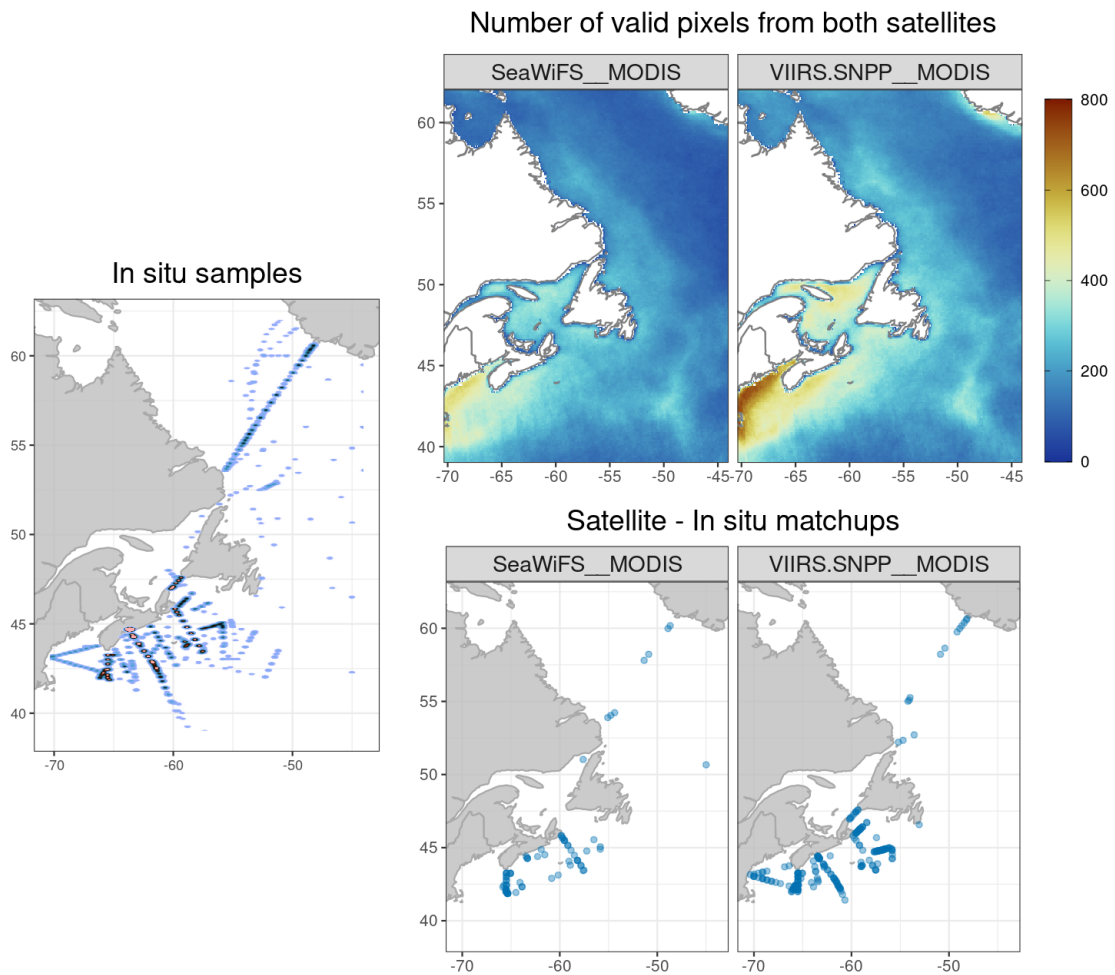# APPENDIX B   Modelled MODIS satellite data and in situ chla

Figure B.1. Left: Heat map of in situ HPLC *Chl-a* samples used in the comparison between satellite and in situ data (2003-2020). Top right: Number of matching points between SeaWiFS (or VIIRS) and MODIS across the time series overlapping the two sensors. Bottom right: The resulting matchups between in situ samples, MODIS, and SeaWiFS (or VIIRS).
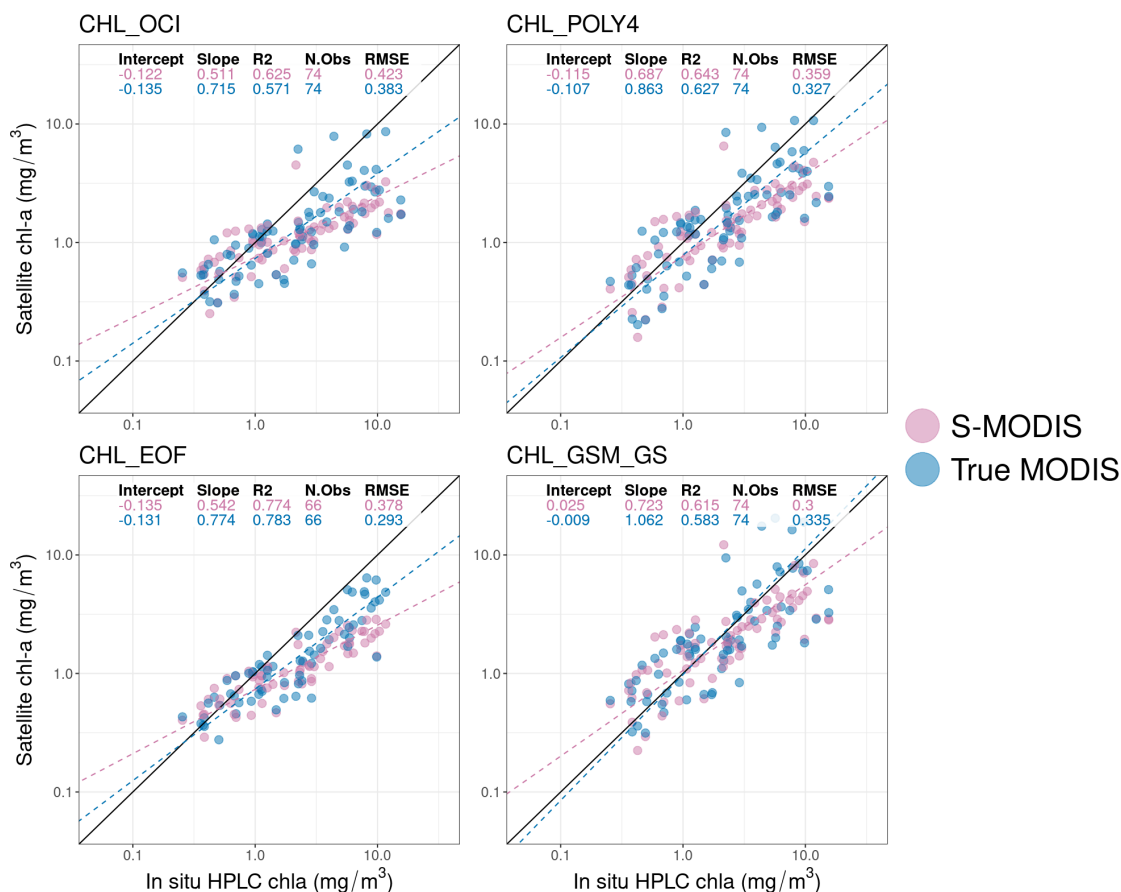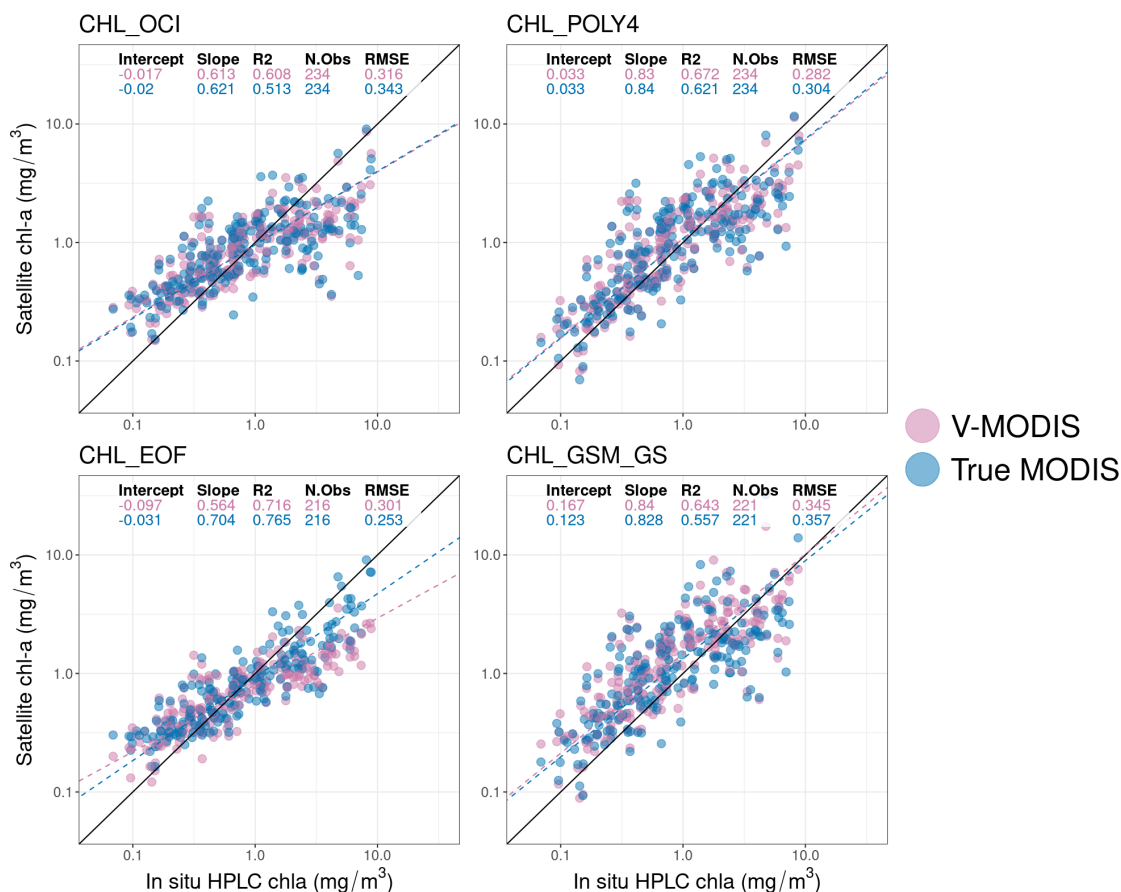
Figure B.2. Results of the regression between true MODIS (or S-MODIS) and in situ *Chl-a* using four different models (OCI, POLY4, EOF, and GSM_GS). Two extreme satellite GSM_GS points are outside the range of the y-axis (83.45 and 90.02 mg/m^3).

Table B.1. Pearson's correlation coefficients between percent error (satellite vs in situ chl-a), and five variables. 'Distance' is between pixel and in situ sample location. True MODIS chl-a and S-MODIS chl-a are regressed against in situ chl-a separately. * = p < 0.01

Satellite data	Algorithm	Latitude	Year	Day of year	Distance	Bathymetry
True MODIS	OCI	-0.161	-0.126	0.249	-0.086	0.148
	POLY4	-0.199	-0.153	0.130	-0.086	0.273
	GSMGS	-0.165	-0.297	-0.049	0.005	0.244
	EOF	0.150	0.066	0.504*	-0.241	0.094
S-MODIS	OCI	-0.140	0.101	0.523*	-0.154	-0.028
	POLY4	-0.166	0.107	0.455*	-0.166	0.060
	GSMGS	-0.188	0.126	0.377*	-0.127	0.049
	EOF	0.237	0.214	0.531*	-0.176	-0.041

B.2 VIIRS

Figure B.3. Results of the regression between true MODIS (or V-MODIS) and in situ *Chl-a* using four different models (OCI, POLY4, EOF, and GSM_GS).

Table B.2. Pearson's correlation coefficients between percent error (satellite vs in situ chl-a), and five variables. 'Distance' is between pixel and in situ sample location. True MODIS chl-a and V-MODIS chl-a are regressed against in situ chl-a separately. * = p < 0.01

Satellite data	Algorithm	Latitude	Year	Day of year	Distance	Bathymetry
True MODIS	OCI	-0.034	0.118	0.470*	-0.037	0.088
	POLY4	0.061	0.123	0.329*	-0.057	0.199*
	GSMGS	0.050	0.109	0.337*	-0.047	0.204*
	EOF	0.061	0.083	0.412*	0.019	-0.039
V-MODIS	OCI	-0.057	0.158	0.471*	-0.007	0.049
	POLY4	0.039	0.130	0.335*	-0.020	0.164
	GSMGS	0.009	0.205*	0.307*	-0.056	0.174*
	EOF	0.049	0.031	0.540*	-0.024	0.013

APPENDIX C   Temporal statistics
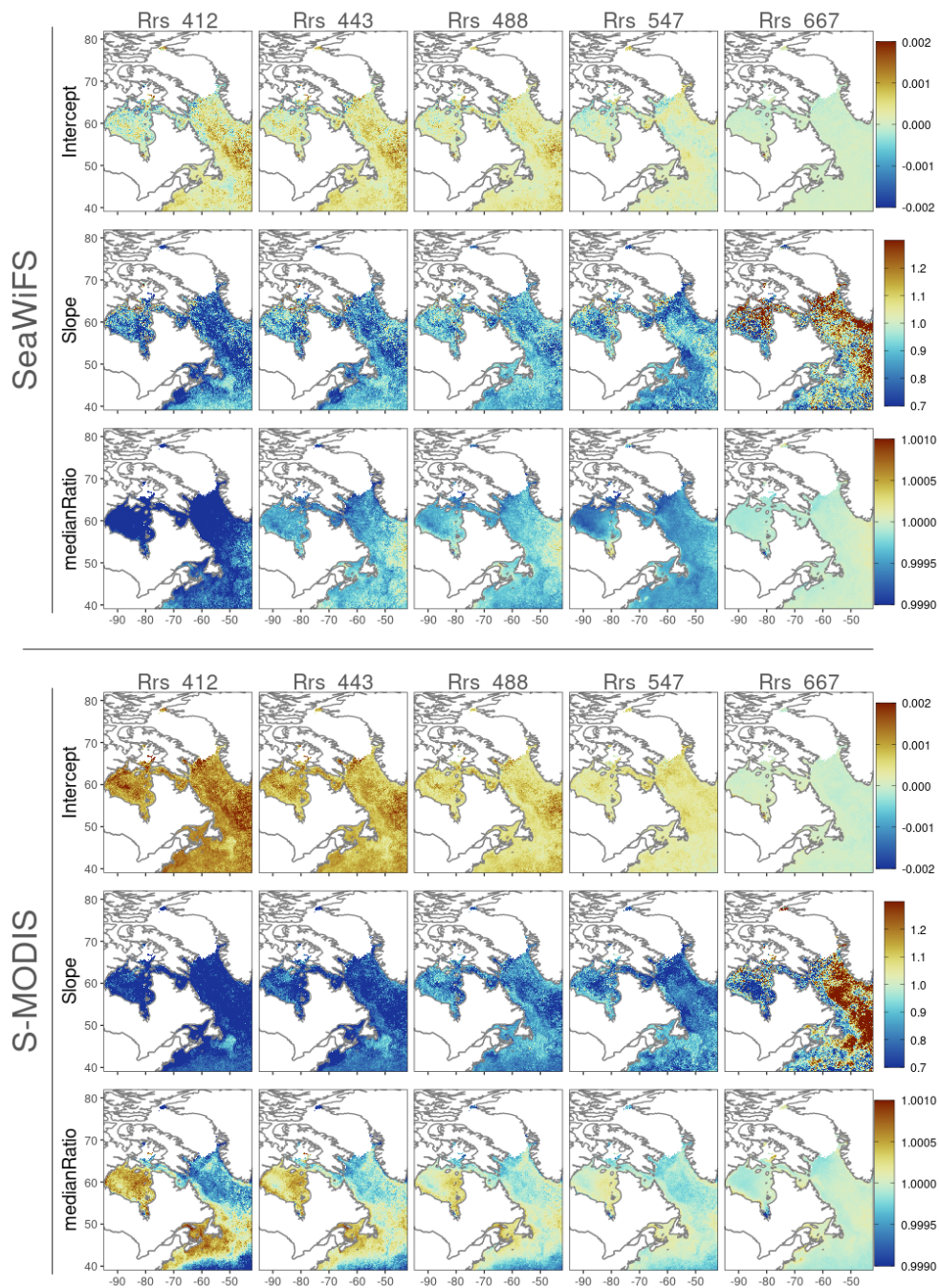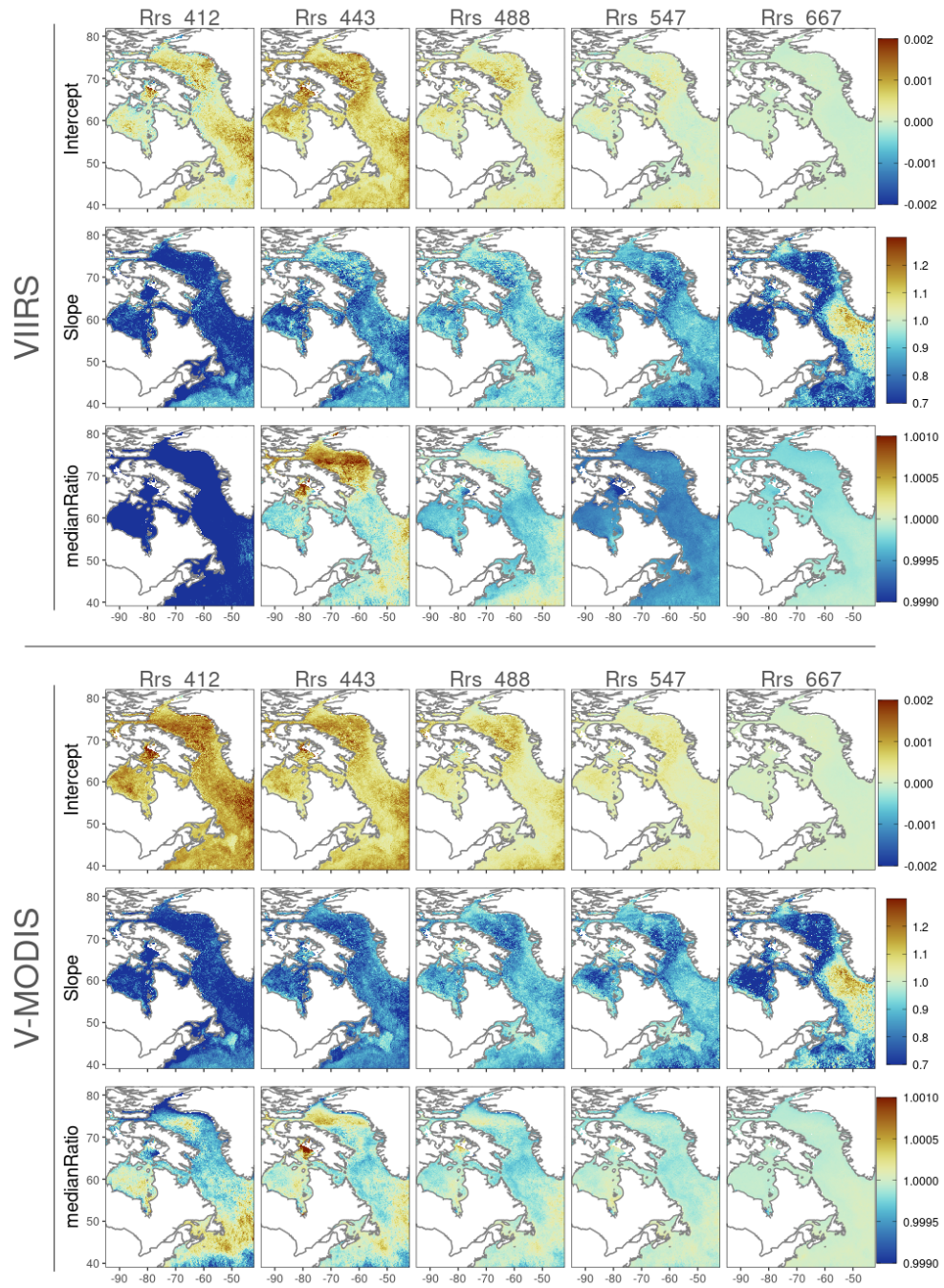

Figure C.1. Maps created from regressing SeaWiFS *Rrs* against MODIS *Rrs*, before (top three rows) and after (bottom three rows) applying the band model, using the method described in Figure 6. Pixels with fewer than 50 observations were removed. Before calculating the statistics, outliers were removed (pixels where the ratio of sensors was in the 1st or 99th percentiles). The median of the sensor ratio (i.e. SeaWiFS:MODIS) across time was also calculated. For S-MODIS to MODIS, the *Rrs* compared between sensors are from the same wavebands.
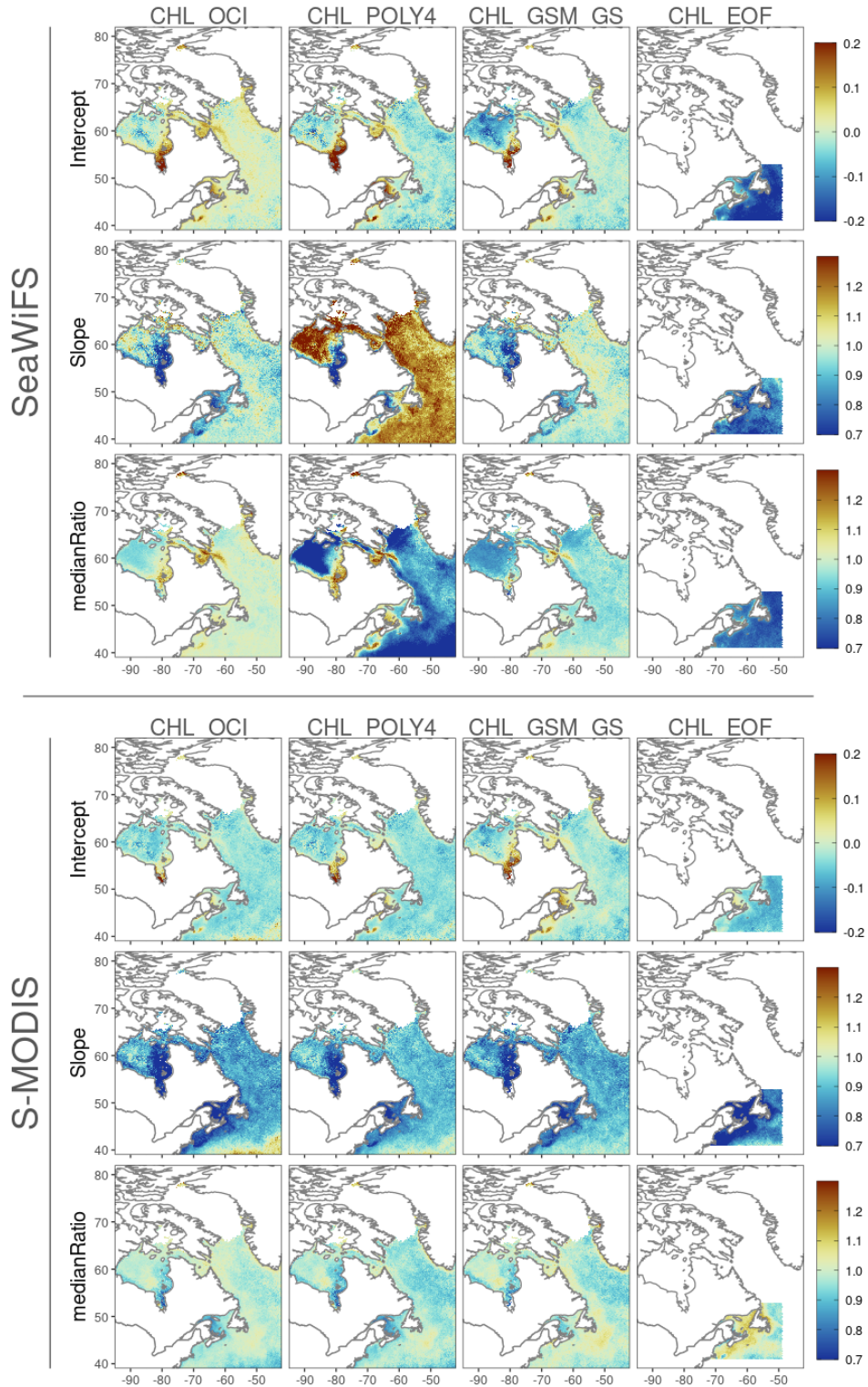
Figure C.2. Maps created from regressing log10(SeaWiFS *Chl-a*) against log10(MODIS *Chl-a*), before (top three rows) and after (bottom three rows) applying the band model, using the method described in Figure 6. Pixels with fewer than 50 observations were removed from the maps. Before calculating the statistics, outliers were removed (pixels where the ratio of sensors was in the 1st or 99th percentiles), and *Chl-a* was log10-transformed. The median of the sensor ratio (i.e. SeaWiFS:MODIS) across time was also calculated (using unlogged data).

Figure C.3. Maps created from regressing VIIRS *Rrs* against MODIS *Rrs*, before (top three rows) and after (bottom three rows) applying the band model. See Figure C.1 caption for details.
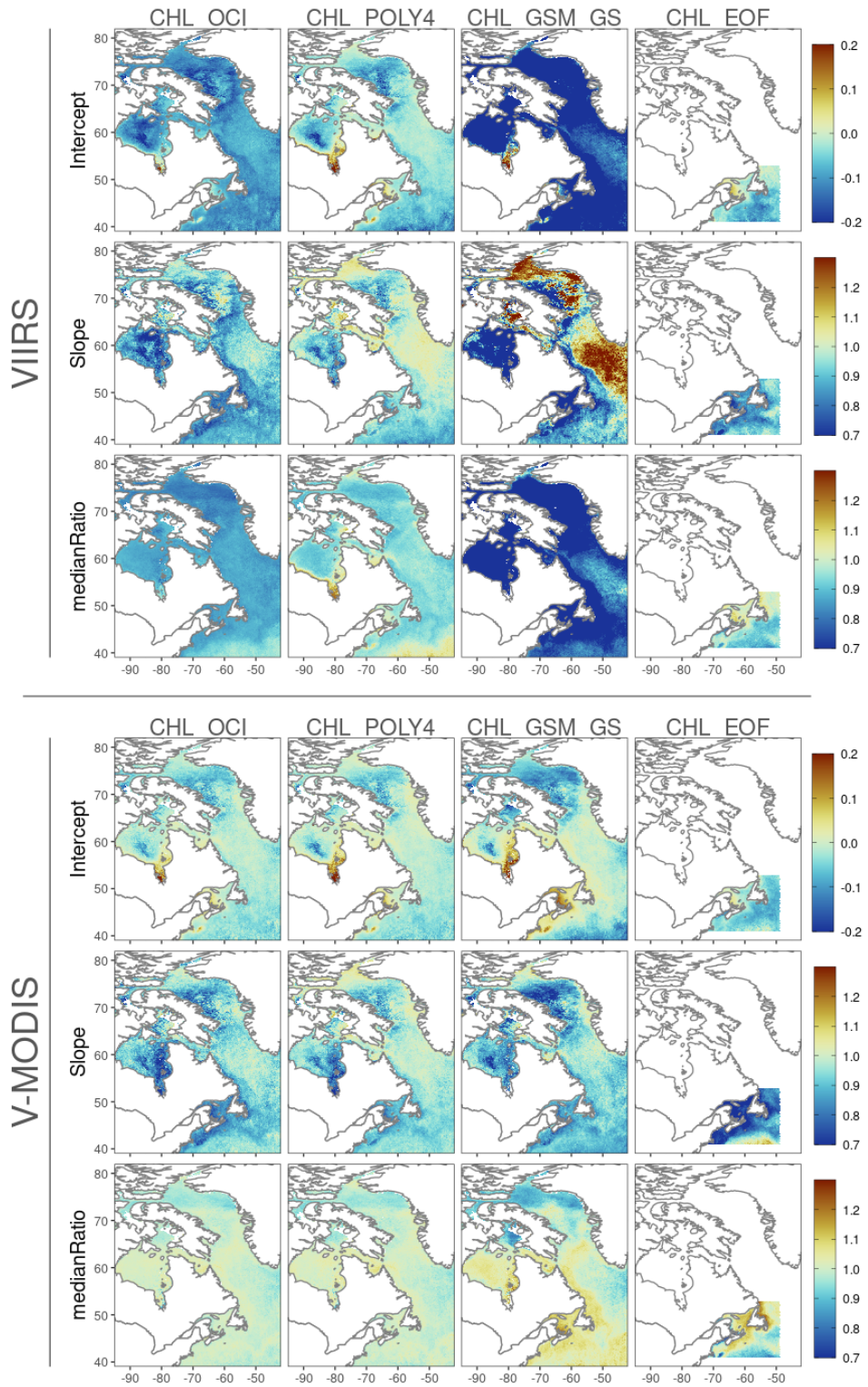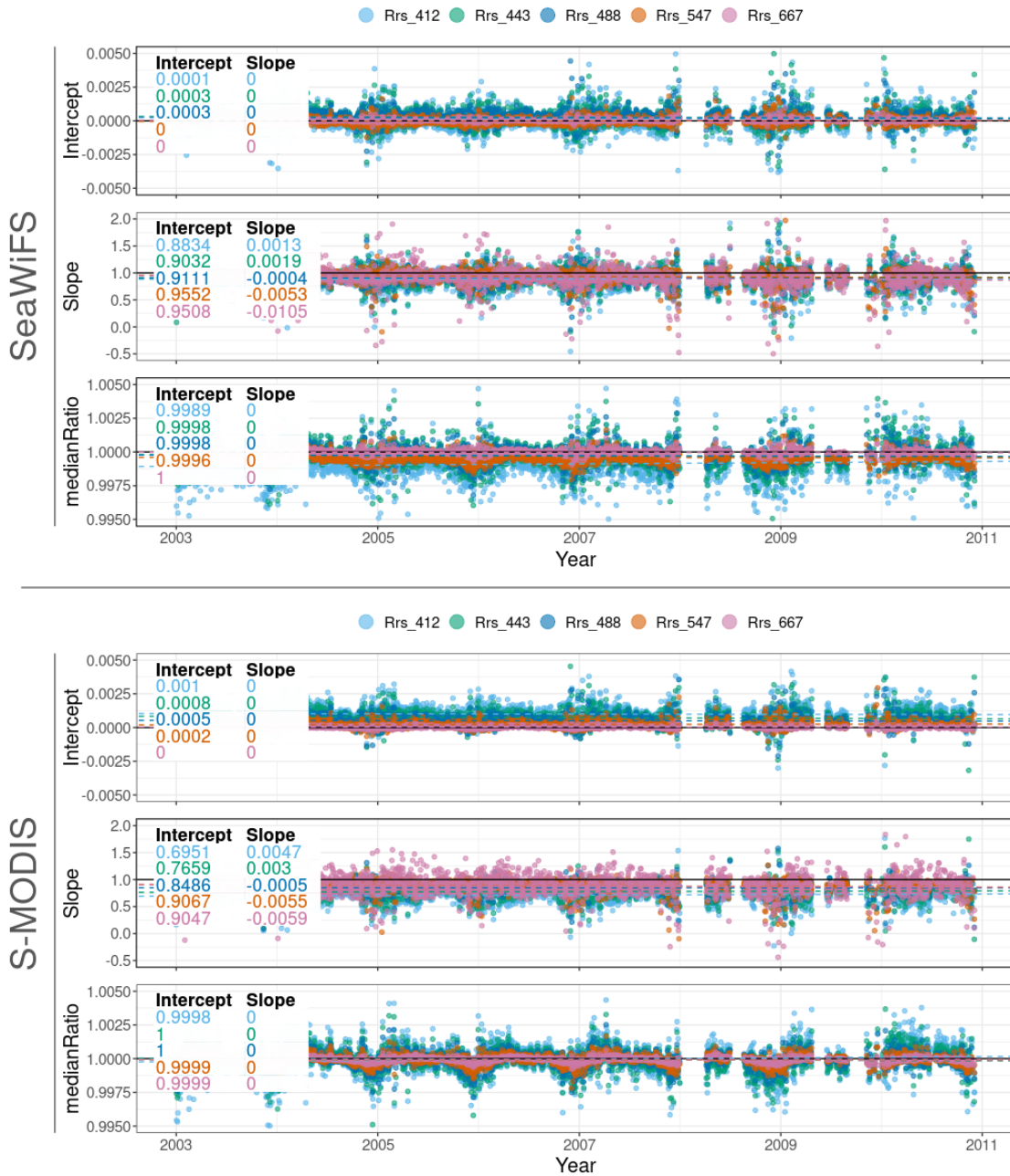
Figure C.4. Maps created from regressing log10(VIIRS *Chl-a*) against log10(MODIS *Chl-a*), before (top three rows) and after (bottom three rows) applying the band model. See Figure C.2 caption for details.

Figure D.1. Time series created from regressing daily SeaWiFS *Rrs* against MODIS *Rrs*, before (top three rows) and after (bottom three rows) applying the band model, using the method in Figure 7. Before calculating statistics, outliers (pixels where the ratio of sensors was in the 1st or 99th percentiles) and daily images with fewer than 50 observations were removed. The daily median ratio SeaWiFS:MODIS across the *NWA* was also calculated. Ideal statistics (e.g. slope=0, ratio=1) are marked with a black horizontal line. P-values have been omitted as >99.4% of the daily linear regressions across the times series were significant (p<0.01).
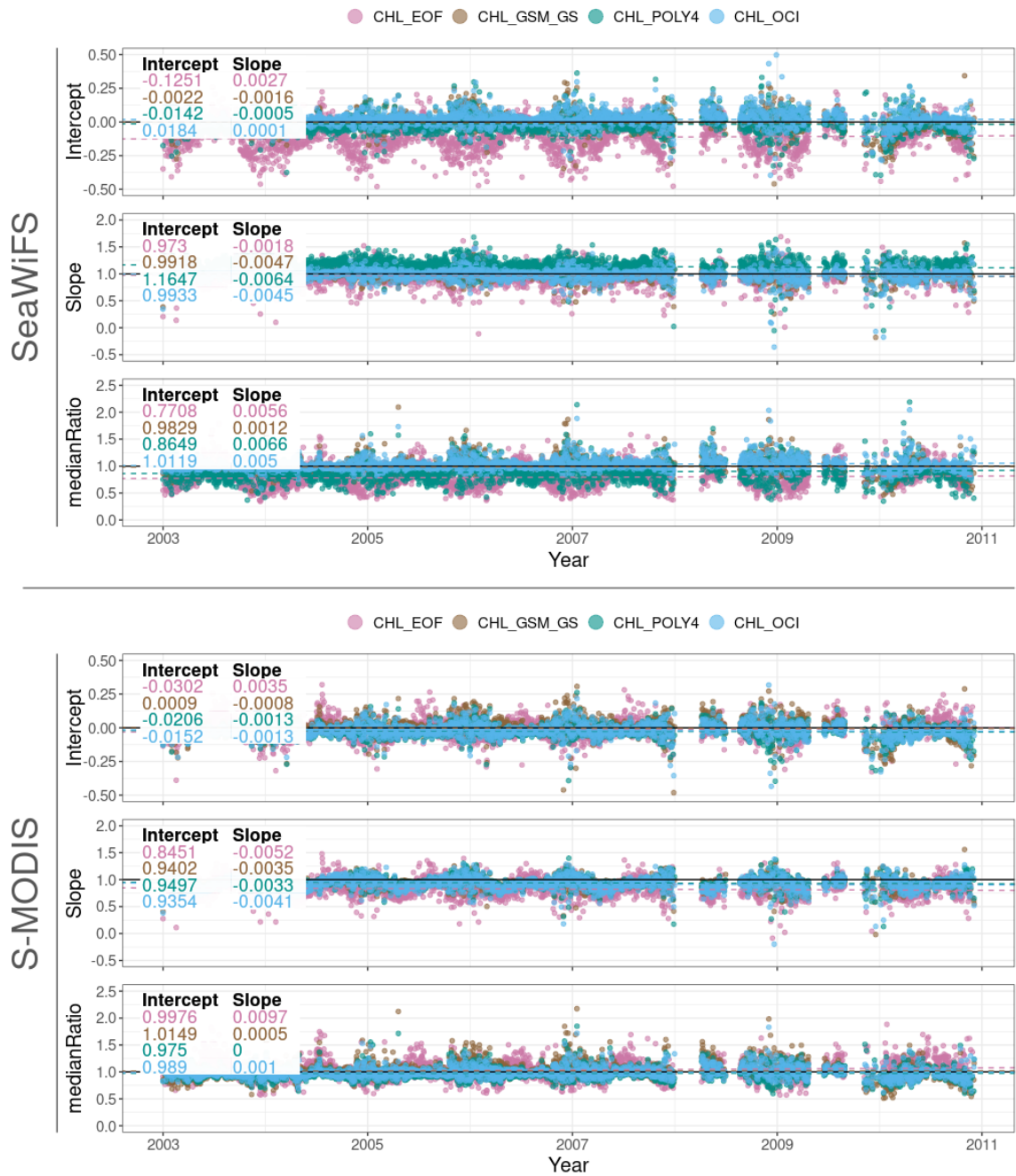
Figure D.2. Time series created from regressing daily log10(SeaWiFS *Chl-a*) against log10(MODIS *Chl-a*), before (top three rows) and after (bottom three rows) applying the band model, using the method described in Figure 7. Before calculating statistics, outliers (pixels where the ratio of sensors was in the 1st or 99th percentiles) and daily images with fewer than 50 observations were removed, and *Chl-a* was log10-transformed. The daily median ratio SeaWiFS:MODIS across the *NWA* was also calculated (using unlogged data). Ideal statistics (e.g. slope=0, ratio=1) are marked with a black horizontal line. P-values have been omitted as >99.7% of the daily linear regressions across the times series were statistically significant (p<0.01).

Figure D.3. Time series created from regressing daily VIIRS *Rrs* against MODIS *Rrs*, before (top three rows) and after (bottom three rows) applying the band model. See Figure D.1 caption for details. P-values have been omitted as >99.8% of the daily linear regressions across the times series were statistically significant (p<0.01).
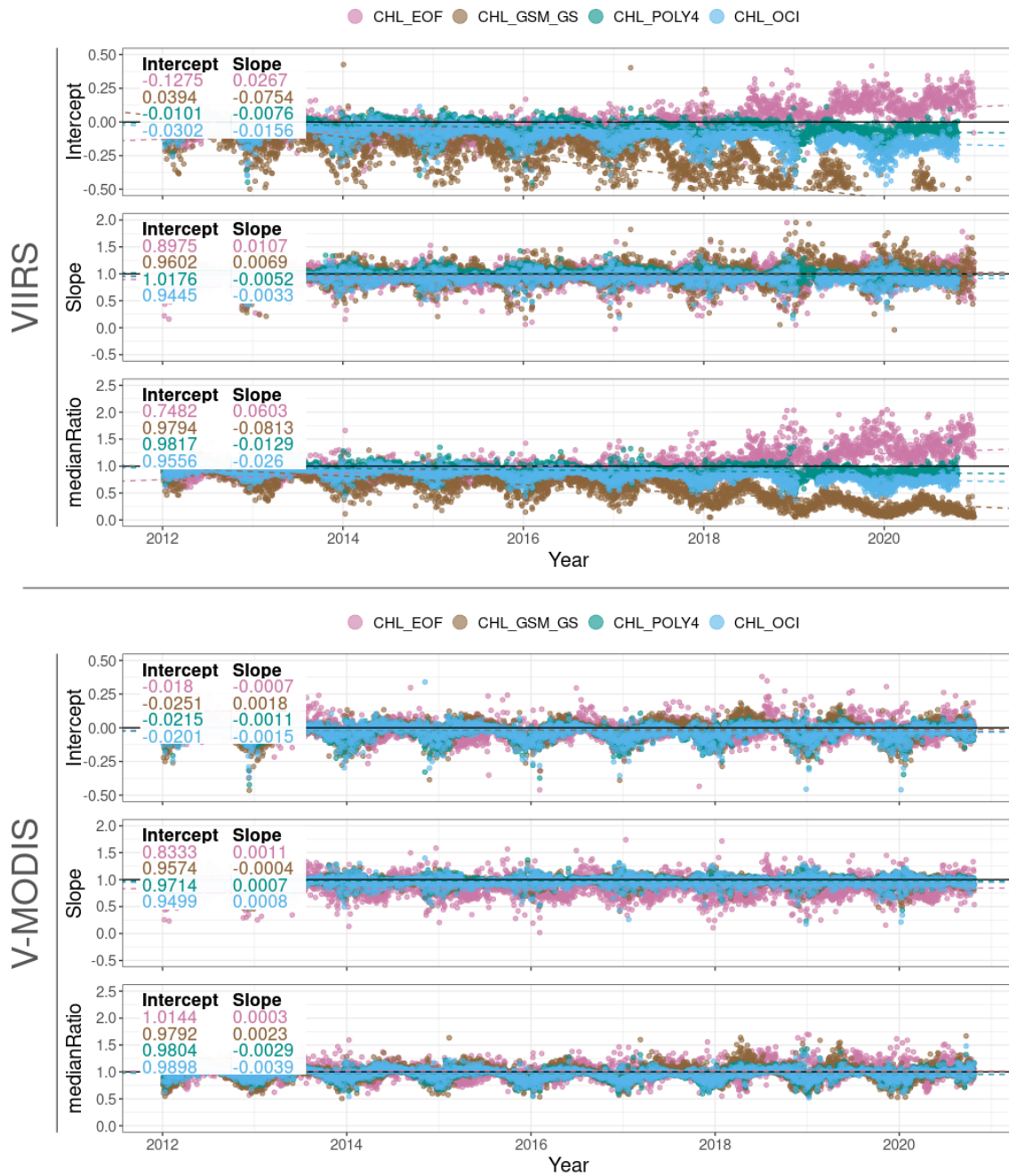
Figure D.4. Time series created from regressing daily log10(VIIRS *Chl-a*) against log10(MODIS *Chl-a*), before (top three rows) and after (bottom three rows) applying the band model. See Figure D.2 caption for details. P-values have been omitted as >99.9% of the daily linear regressions across the times series were statistically significant (p<0.01).