



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Comparação de algoritmos para a estimativa da concentração de clorofila-*a* por satélite nas regiões Norte e Sudeste da margem continental brasileira

Elias Rosenberg Krausz

**Relatório final de pesquisa de iniciação
científica orientada pelo Prof. Dr. Milton
Kampel**

INPE
São José dos Campos
2019

Agradecimentos

À minha querida Juliana, por todo amor e companheirismo desse mundo e também pelo apoio em todos os momentos possíveis. À minha família pela melhor educação e carinho. Aos meus grandes amigos da vida por estarem sempre comigo. Ao Mateus Chuqui, pela paciência e suporte ao longo de todo o projeto. Ao meu orientador Prof. Dr. Milton Kampel pela oportunidade de realizar essa pesquisa e desenvolver meu trabalho acadêmico.

Resumo

As propriedades óticas da água do mar permitem a estimativa de sua composição bioquímica, quando analisadas objetivamente por meio de medidas de reflectância. Espectroradiômetros acoplados à satélites que orbitam a Terra permitem que a radiação refletida pelo oceano seja medida em diferentes bandas de comprimentos de onda. Uma série de algoritmos para os diferentes sensores possibilitam a estimativa das concentrações de clorofila-a, que representa um índice de biomassa do fitoplâncton, na superfície, sendo variável seu desempenho em diferentes regiões do planeta devido a presença de detritos e material dissolvido que interagem com a luz nas mesmas faixas de comprimento de onda que a clorofila e ao espalhamento da luz na atmosfera. Inúmeras bases de dados estão disponíveis contendo dados de clorofila-a já processados, no entanto é necessária a verificação de sua performance com dados *in situ*. O presente trabalho visa a avaliação do desempenho dos algoritmos padrões da Agência Espacial Europeia e da Agência Espacial dos Estados Unidos, além dos algoritmos de redes neurais (NN), *Ocean Color* (OCx) e OC4 para imagens obtidas pelos sensores *Ocean and Land Colour Instrument* (OLCI) e *AQUA MODerate Resolution Imaging Spectroradiometer* (Aqua MODIS) nas regiões Norte e Sudeste da margem continental brasileira de forma comparativa aos dados obtidos *in situ*. As áreas de estudo compreendem uma estação fixa no Litoral Norte do estado de São Paulo, as regiões da elevação de Rio Grande e da Cadeia Vitória-Trindade e a área de influência da pluma estuarina do Rio Amazonas, na costa Norte do Brasil, sendo possível a existência de importantes diferenças regionais entre as mesmas. Para a avaliação da performance dos algoritmos foram utilizados os parâmetros estatísticos como: coeficiente de correlação de *Pearson*, coeficiente angular, coeficiente linear, erro médio quadrático e bias. Os resultados indicam um melhor desempenho do algoritmo OCx aplicado à janelas de 1x1 pixels para o sensor AquaMODIS em comparação com o algoritmo *standard* e às janelas espaciais de 3x3 pixels.

Palavras chave: Cor do oceano, clorofila, fitoplâncton, oceanografia por satélites

Comparison of algorithms for the estimation of concentrations of chlorophyll-*a* by satellite in the North and Southeast regions of the Brazilian continental margin

Abstract

The optical properties of sea water allow the estimation of its biochemical composition, when objectively analyzed by means of reflectance measurements. Spectroradiometers coupled to Earth-orbiting satellites allow the radiation reflected by the ocean to be measured in different bands of wavelengths. A series of algorithms for the different sensors make it possible to estimate the concentrations of chlorophyll-*a*, which represents a biomass index of the phytoplankton, on the surface. Its performance in different regions of the planet varies due to the presence of debris and dissolved material interacting with light in the same wavelength bands as chlorophyll and scattering of light into the atmosphere. Numerous databases are available containing chlorophyll-*a* data already processed, however it is necessary to verify performance with data *in situ*. The aim of this work is to evaluate the performance of the standard algorithms of the European Space Agency and the United States Space Agency, as well as the neural network (NN), Ocean Color (OCx) and OC4 for the sensors Ocean and Land Colour Instrument (OLCI) and AQUA MODerate Resolution Imaging Spectroradiometer (Aqua MODIS) in the North and Southeast regions of the Brazilian continental margin in a comparative way to the data obtained *in situ*. The study areas comprise a fixed coast station in the North coast of the state of São Paulo, the Rio Grande elevation and the Vitória-Trindade chain, and the estuarine plume area of the Amazon River on the North coast of Brazil, with important regional differences between them. For the evaluation of the performance of the algorithms were used the statistical parameters as: Pearson correlation coefficient, angular coefficient, linear coefficient, mean square error and bias.

Keywords: Ocean color, chlorophyll, phytoplankton, satellite oceanography

Sumário

1	Introdução	1
1.1	Objetivos	2
1.2	Área de Estudo	3
2	Material e Métodos	5
2.1	Algoritmos para a estimativa de clorofila	6
2.2	Avaliações estatísticas	7
3	Resultados	8
3.1	AquaMODIS	8
3.2	OLCI	10
4	Discussão	14
4.1	AquaMODIS	14
4.2	OLCI	14
5	Considerações finais	16
6	Referências	17

Lista de Figuras

1	Esquema representativo do funcionamento de sensores acoplados à satélites para a obtenção de dados de cor do oceano (Elias Rosenberg Krausz, 2019)	1
2	A Plataforma Continental Brasileira e a subdivisão utilizada por Ciotti (2010). Os gradientes de cor mostram o desvio padrão da média anual de concentração de clorofila em mg.m^{-3} , em escala logarítmica, para dados combinados de 2000 à 2005 obtidos pelo sensor SeaWIFS. A linha representa a isóbata de 200m que define o limite da plataforma continental. Fonte: Ciotti (2010)	3
3	Mapa das estações de coleta de dados <i>in situ</i> de clorofila dos projetos Recifes Amazônicos, Ilhas II, Antares-Ubatuba e Marine E-TECH utilizados neste trabalho	6
4	Regressões lineares (log-log) entre dados <i>in situ</i> e estimativas por satélite da concentração de clorofila utilizando o algoritmo padrão (a) e o algoritmo OCx (b) e janelas de 1x1 pixel para o sensor AquaMODIS, em laranja. Linha 1:1 em verde. Os matchups estão representados por cor de acordo com o projeto ou cruzeiro, da seguinte forma: Antares Ubatuba - Vermelho, Ilhas Oceânicas - Verde, Recifes Amazônicos - Azul.	8
5	Regressões lineares (log-log) entre dados <i>in situ</i> e estimativas por satélite da concentração de clorofila utilizando o algoritmo Standard (a) e o algoritmo OCX (b) e média de janelas de 3x3 pixels para o sensor Aqua MODIS, em laranja. Linha 1:1 em verde. Os matchups estão representados por cor de acordo com o projeto ou cruzeiro, da seguinte forma: Antares Ubatuba - Vermelho, Ilhas Oceânicas - Verde, Recifes Amazônicos - Azul	9
6	Regressões lineares (log-log) entre dados <i>in situ</i> e estimativas por satélite da concentração de clorofila utilizando o algoritmo Standard (a) e o algoritmo OCX (b) e mediana de janelas de 3x3 pixels para o sensor Aqua MODIS, em laranja. Linha 1:1 em verde. Os matchups estão representados por cor de acordo com o projeto ou cruzeiro, da seguinte forma: Antares Ubatuba - Vermelho, Ilhas Oceânicas - Verde, Recifes Amazônicos - Azul	10
7	Regressões lineares (log-log) entre dados <i>in situ</i> e estimativas por satélite da concentração de clorofila utilizando o algoritmo de redes neurais (NN) (a) e o algoritmo OC4 (b) e janelas de 1x1 pixel para o sensor OLCI, em azul. Linha 1:1 em laranja	11
8	Regressões lineares (log-log) entre dados <i>in situ</i> e estimativas por satélite da concentração de clorofila utilizando o algoritmo de redes neurais (NN) (a) e o algoritmo OC4 (b) e média de janelas de 3x3 pixels para o sensor OLCI, em azul. Linha 1:1 em laranja . . .	11
9	Regressões lineares (log-log) entre dados <i>in situ</i> e estimativas por satélite da concentração de clorofila utilizando o algoritmo de redes neurais (NN) (a) e o algoritmo OC4 (b) e mediana de janelas de 3x3 pixels para o sensor OLCI, em azul. Linha 1:1 em laranja .	12

Lista de Tabelas

- 1 Estações com matchups válidos para um ou mais sensores utilizados no estudo 8
- 2 Avaliação estatística da performance dos algoritmos do sensor AquaMODIS para a margem continental brasileira. Janela se refere à janela espacial em pixels. N é o número de pontos nas imagens de satélite para os quais foram feitas as comparações entre dados *in situ* e as estimativas do algoritmo. R^2 se refere ao coeficiente de correlação de Pearson. RMSE é a raiz do erro quadrático médio. C.A. é o coeficiente angular e C.L. é o coeficiente linear da regressão. Os valores considerados mais consistentes para cada parâmetro estão em negrito 9
- 3 Avaliação estatística da performance dos algoritmos do sensor OLCI para a margem continental brasileira. Janela se refere à janela espacial em pixels. N é o número de pontos nas imagens de satélite para os quais foram feitas as comparações entre dados *in situ* e as estimativas do algoritmo. R^2 se refere ao coeficiente de correlação de Pearson. RMSE é a raiz do erro quadrático médio. C.A. é o coeficiente angular e C.L. é o coeficiente linear da regressão. Os valores considerados mais consistentes para cada parâmetro estão em negrito 13

1 Introdução

O oceano é responsável diretamente pela manutenção do clima e da vida no planeta Terra, constituindo um compartimento fundamental para o ciclo do carbono, mais especificamente no que diz respeito aos processos biogeoquímicos relacionados ao fitoplâncton, que é a base da cadeia trófica marinha (1). Os principais serviços desempenhados pelo fitoplâncton são reguladores, como o sequestro de carbono atmosférico e sua sedimentação em águas profundas, e de produção, como a liberação de gás oxigênio pela fotossíntese, incluindo cerca de 50% do oxigênio atmosférico e a ciclagem de nutrientes (2). O monitoramento e o estudo do fitoplâncton em escala global podem ser realizados por meio do sensoriamento remoto orbital. Isso é possível porque parte da radiação solar que penetra a água retorna à superfície por espalhamento e é detectada por sensores presentes em satélites que orbitam o planeta (3), conforme representado na figura 1.

A cor do oceano, que é produto de suas propriedades óticas, permite a avaliação da concentração de clorofila-*a* a partir das relações entre as reflectâncias da superfície do mar em diferentes comprimentos de onda, sendo possível a utilização de diferentes algoritmos para essa finalidade. Entretanto, os algoritmos que possibilitam tal avaliação também devem levar em conta fatores como a interação da luz com aerossóis presentes na atmosfera, a presença de material particulado, matéria orgânica cromófora no oceano e as características bio-ópticas da comunidade planctônica, de modo que as estimativas de concentração de clorofila-*a* sejam mais precisas (4). De acordo com suas propriedades intrínsecas a água do mar pode ser classificada como de casos 1 ou 2 (3). Regiões oceânicas e oligotróficas apresentam tipicamente águas do caso 1, em que há baixas concentrações de materiais em suspensão e substâncias cromóforas dissolvidas, ou seja, as propriedades óticas aparentes dependem majoritariamente do fitoplâncton presente no meio. Já em zonas costeiras a interferência de tais fatores tende a ser mais intensa, configurando águas do caso 2. Esses fatores interagem com a luz na mesma banda do espectro que a clorofila-*a*, o que torna necessária a utilização de algoritmos regionais.

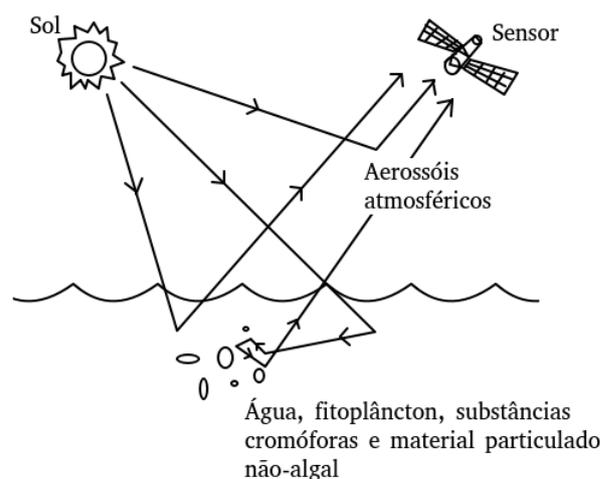


Figura 1: Esquema representativo do funcionamento de sensores acoplados à satélites para a obtenção de dados de cor do oceano (Elias Rosenberg Krausz, 2019)

1.1 Objetivos

Com isso, este estudo tem como objetivo avaliar a performance dos algoritmos de clorofila-*a* dos sensores *Ocean and Land Colour Instrument* (OLCI) e *MODerate Resolution Imaging Spectroradiometer* (Aqua MODIS) nas regiões Norte e Sudeste da margem continental brasileira de forma comparativa aos dados obtidos *in situ*. Os algoritmos a serem avaliados são os padrões das agências espaciais responsáveis pelos satélites (Agência Espacial Europeia no caso do OLCI e Agência Espacial dos Estados Unidos no caso do Aqua MODIS), além dos algoritmos de redes neurais (NN), o *ocean color* (OCX) e o *ocean color 4* (OC4).

O sensor MODIS foi lançado a bordo do satélites *Aqua* e *Terra* pela NASA em 2002 e sua operação deve ter fim nos próximos anos (5). Já o OLCI, lançado pela ESA a bordo dos satélites Sentinel-3A e Sentinel-3B no ano de 2016, ainda necessita o aprimoramento de seus produtos à medida que a comunidade científica internacional realiza a validação dos mesmos. No momento em que esse estudo foi realizado, a operação simultânea dos dois sensores, proporciona uma interessante avaliação comparativa de seus produtos para a estimativa de clorofila-*a*.

A margem continental brasileira é uma província em que há baixa disponibilidade de dados *in situ* de concentração de clorofila, o que torna a avaliação dos algoritmos um desafio. Pesquisas a respeito da biomassa fitoplancônica e produção primária na margem continental brasileira ainda são razoavelmente esparsas, embora alguns padrões temporais e espaciais já tenham sido observados (6). No entanto, o banco de dados de campo utilizado para o presente trabalho reúne estações que abrangem 4 das 7 grandes áreas da plataforma continental brasileira conforme a subdivisão dessa província proposta por Ciotti (2010) representada na figura 2, sendo elas as áreas 1, 2, 5 e 7. De tal forma, o estudo engloba regiões com características bastante diversas no que diz respeito às propriedades bio-óticas da água, o que torna possível variações na performance regional dos algoritmos.

1.2 Área de Estudo

A plataforma continental brasileira é uma província que ocupa mais de 40° de latitude, apresentando diferenças regionais significantes em função principalmente de sua extensão, da influência da circulação oceânica e do aporte continental (7). A largura da plataforma também é bastante variável ao longo de sua extensão, sendo máxima na região amazônica, onde atinge até 330km e mínima em sua porção leste, com apenas 10km. Em relação aos fatores e processos que influenciam diretamente as concentrações de nutrientes e a produtividade primária, é notável um padrão de variação meridional ao longo da plataforma (6). Esse estudo está centrado nas regiões Norte e Sudeste da plataforma continental brasileira, incluindo uma estação costeira fixa na região de Ubatuba, no litoral Norte de São Paulo.

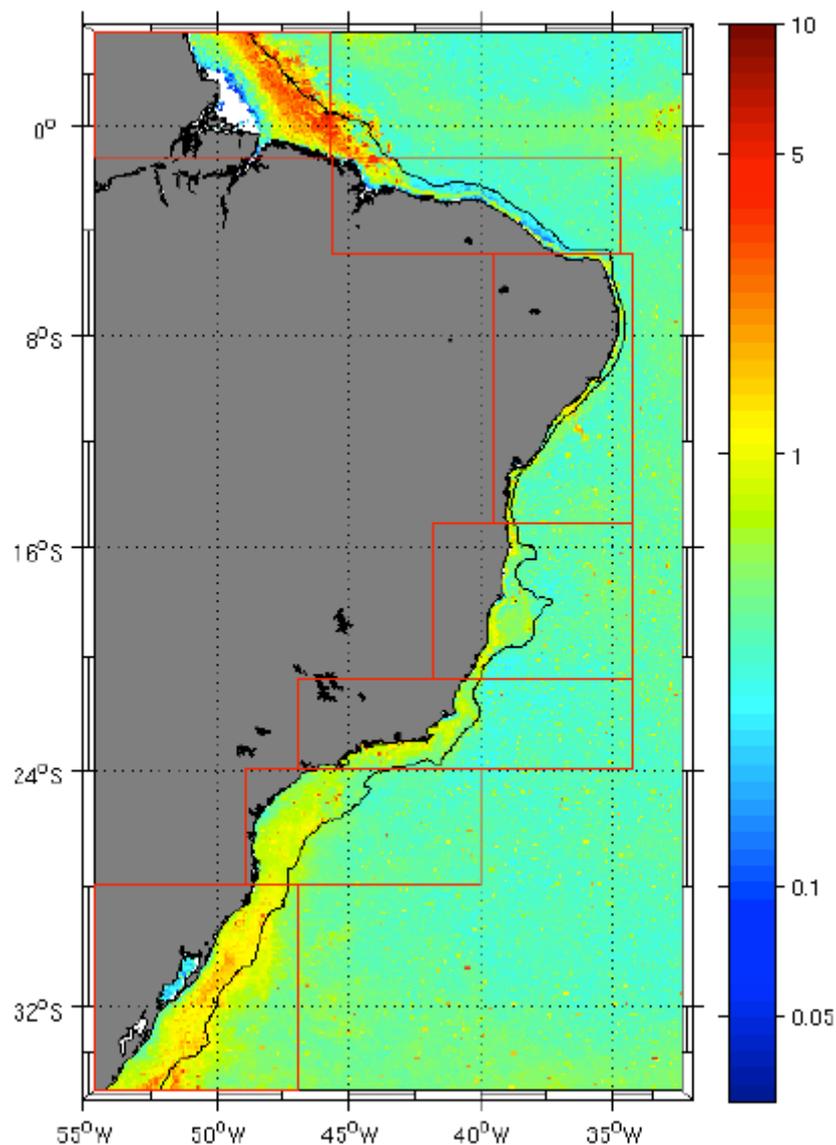


Figura 2: A Plataforma Continental Brasileira e a subdivisão utilizada por Ciotti (2010). Os gradientes de cor mostram o desvio padrão da média anual de concentração de chlorofila em mg.m^{-3} , em escala logarítmica, para dados combinados de 2000 à 2005 obtidos pelo sensor SeaWiFS. A linha representa a isóbata de 200m que define o limite da plataforma continental. Fonte: Ciotti (2010)

A região Norte da plataforma continental, influenciada pela Corrente Norte do Brasil, apresenta um importante aporte de nutrientes por meio do Rio Amazonas, especialmente à Noroeste de sua foz. A região é fertilizada pelos nutrientes dissolvidos e detritos orgânicos provenientes da descarga fluvial, abastecendo a produção primária da comunidade fitoplanctônica, constituída principalmente por diatomáceas (7). Devido a sua alta turbidez, suas águas são classificadas como do caso 2 e apresentam um padrão de variação sazonal de acordo com o regime de chuvas na região amazônica.

A região Sudeste da plataforma, entretanto, possui características bastante diferentes. De caráter oligotrófico e permanentemente estratificado, apresenta fontes de nutrientes predominantemente oceânicas, especialmente por meio da intrusão da Água Central do Atlântico Sul (ACAS) em função de processos físicos de mesoescala (7). Sua produção primária é dominada pelo microfitoplâncton e suas águas são bastante oligotróficas. A região da estação "Antares Ubatuba", apesar de estar dentro da área compreendida como Plataforma Continental Sudeste, tem a influência continental destacada quando comparada às estações dos cruzeiros "Marine E-Tech" e "Ilhas Oceânicas", devido à proximidade do continente e menor profundidade da coluna d'água, apresentando águas do caso 2 (8).

2 Material e Métodos

As imagens de satélite em nível de processamento L2, em alta resolução espacial (*L2 full resolution*) foram obtidas por meio dos portais “Copernicus Online Data Access” da “EUMETSAT” (*European Organisation for the Exploitation of Meteorological Satellites*) e “NASA Ocean Color” e processadas por meio dos softwares “SNAP” e “Seadas”. Foram utilizadas janelas espaciais de 1x1 e 3x3 pixels para os diferentes produtos centradas nas coordenadas geográficas das estações de coletas *in situ*, sendo que no caso das janelas de 3x3 pixels foram avaliadas separadamente a média e a mediana dos valores de concentração obtidos. Já os dados *in situ* serão reunidos em um banco de dados consistente de algumas regiões distintas dos cruzeiros “Marine E-Tech”, “Ilhas Oceânicas”, “Recifes Amazônicos”, e “Antares-Ubatuba”, cujas estações estão representadas no mapa da figura 2. A resolução temporal utilizada entre as medidas *in situ* e os dados de satélite é de no máximo 24 horas. Considerando-se os matchups de acordo com a resolução temporal, o banco de dados conta com 172 pontos para os quais os algoritmos foram aplicados em pelo menos um dos sensores avaliados, conforme representado na tabela 1.

O critério de escolha para as áreas de estudo se baseia na disponibilidade dos dados *in situ* e na amostragem de regiões que apresentem características diversas em relação às propriedades bio-óticas. Os cruzeiros “Ilhas 2” e “Marine E-Tech” compreendem áreas oceânicas e oligotróficas, sob baixa influência continental, caracterizadas pela presença de águas do caso 1. Além disso, a Elevação de Rio Grande, onde foram realizadas as coletas do cruzeiro “Marine E-Tech” é uma região estratégica no cenário político, sendo uma nova fronteira da pesquisa científica no país. A área de influência da pluma estuarina do rio Amazonas, amostrada nas estações do projeto “Recifes Amazônicos”, representa uma região extrema no que diz respeito à matéria orgânica dissolvida, sendo um exemplo típico de águas do caso 2. Já a região da estação “Antares-Ubatuba” diferencia-se por estar sob forte influência continental, embora apresente características oligotróficas. A longa série temporal realizada nessa estação possibilita estudos sobre a variação temporal da produção primária e garante grande consistência para a validação dos algoritmos na região.

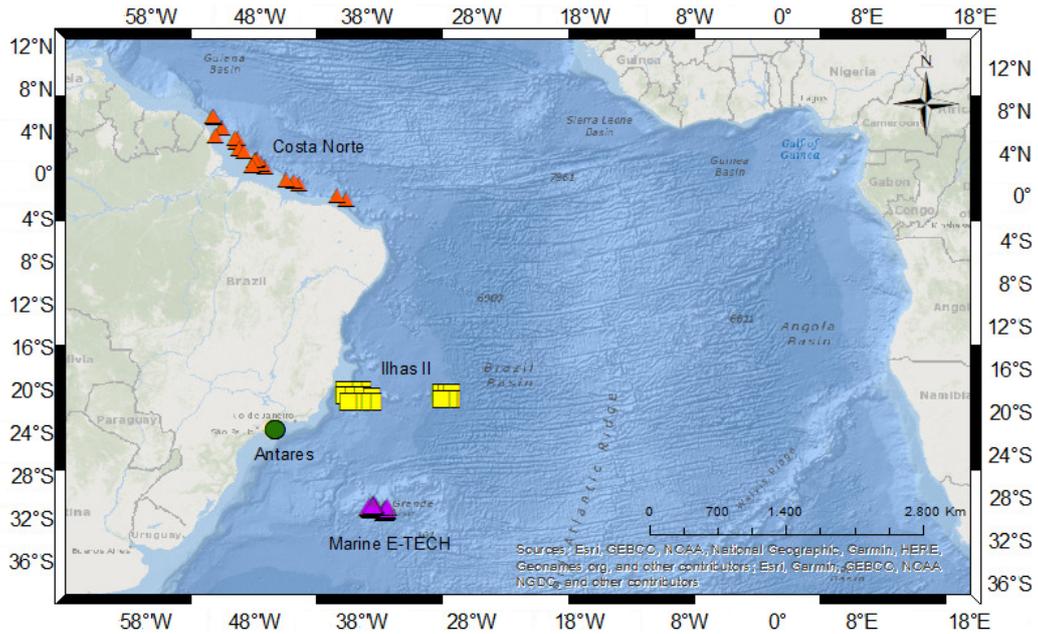


Figura 3: Mapa das estações de coleta de dados *in situ* de clorofila dos projetos Recifes Amazônicos, Ilhas II, Antares-Ubatuba e Marine E-TECH utilizados neste trabalho

2.1 Algoritmos para a estimativa de clorofila

O algoritmo padrão do sensor AquaMODIS fornece a estimativa da concentração de clorofila-a na superfície em $\text{mg}\cdot\text{m}^{-3}$, calculada usando uma relação empírica derivada de medições *in situ* das razões de bandas azul e verde do espectro visível das refletâncias de sensoriamento remoto *in situ*. A implementação depende da disponibilidade de três ou mais bandas de sensor abrangendo o regime espectral de 443-547 nm para o AquaMODIS, sendo aplicável a todos os sensores atuais de cor do oceano. A implementação do atual algoritmo *standard* de clorofila utiliza o algoritmo padrão de razão de banda OC3/OC4 (OCx) fundido com o índice de cor (CI) estabelecido por Hu et al.(2012). Esse refinamento é restrito a águas do caso 1, reduzindo o erro nas estimativas de clorofila para regiões oligotróficas. Para valores de concentração de clorofila de até $0.15\text{mg}\cdot\text{m}^{-3}$ é utilizado o algoritmo CI, já o OCx é aplicado para concentrações de clorofila superiores à $0.2\text{mg}\cdot\text{m}^{-3}$. No intervalo entre os dois valores é realizada uma aproximação entre os dois algoritmos.

O algoritmo CI opera utilizando diferenças das reflectâncias em 3 bandas, aplicando a diferença entre a reflectância na banda do verde e uma referência definida linearmente entre as reflectâncias nas bandas do azul e do verde, da seguinte forma (9):

$$CI = R_{rs}(\lambda_{verde}) [R_{rs}(\lambda_{azul}) + (\lambda_{verde} \lambda_{azul}) / (\lambda_{vermelho} \lambda_{azul}) (R_{rs}(\lambda_{vermelho}) R_{rs}(\lambda_{azul}))]$$

Em que R_{rs} representa reflectância e λ_{azul} , λ_{verde} e $\lambda_{vermelho}$ se referem respectivamente aos comprimentos de onda específicos de 443, 555 e 670nm, respectivamente para o sensor AquaMODIS.

O algoritmo OCx é uma relação polinomial de quarta ordem entre razões de reflectância e clorofila-a, de modo que:

$$\log_{10}(clo - a) = a_0 + a_i \sum_{i=1}^4 \left(\log_{10} \frac{R_{rs}(\lambda_{azul})}{R_{rs}(\lambda_{verde})} \right)^i \quad (1)$$

Onde $R_{rs}\lambda_{azul}$ é o maior valor entre os diferentes R_{rs} utilizados e, para o sensor AquaMODIS, os coeficientes de a_0 à a_4 são respectivamente: 0.2424, -2.7423, 1.8017, 0.0015, -1.2280.

2.2 Avaliações estatísticas

Os gráficos e análises estatísticas foram produzidos por meio de rotinas em Python. Para a avaliação da performance dos algoritmos serão utilizados os parâmetros: coeficiente de correlação de *Pearson*, coeficiente angular, coeficiente linear, erro médio quadrático e bias (10). A incerteza é caracterizada pelo erro médio (bias) e raiz do erro quadrático médio (RMSE) (11). Menores valores de tais índices indicam menor incerteza. O coeficiente angular e a correlação de *Pearson* são medidas da distribuição dos dados, enquanto o coeficiente linear determina o ponto em que a linha de tendência intercepta o eixo y. Por fim, espera-se definir quais algoritmos estimam melhor a clorofila ao longo da costa brasileira.

3 Resultados

3.1 AquaMODIS

Tabela 1: Estações com matchups válidos para um ou mais sensores utilizados no estudo

Cruzeiro/Projeto	Ano	Número de estações
Antares Ubatuba	2006-2018	130
Ilhas Oceânicas	2017	14
Recifes Amazônicos	2017	20
Marine E-Tech	2018	8

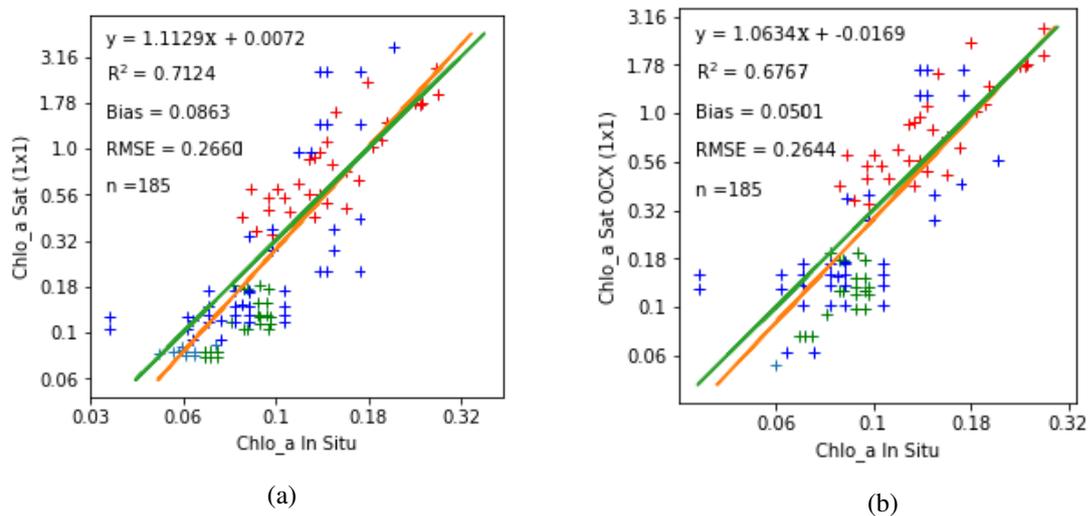


Figura 4: Regressões lineares (log-log) entre dados *in situ* e estimativas por satélite da concentração de clorofila utilizando o algoritmo padrão (a) e o algoritmo OCx (b) e janelas de 1x1 pixel para o sensor AquaMODIS, em laranja. Linha 1:1 em verde. Os matchups estão representados por cor de acordo com o projeto ou cruzeiro, da seguinte forma: Antares Ubatuba - Vermelho, Ilhas Oceânicas - Verde, Recifes Amazônicos - Azul.

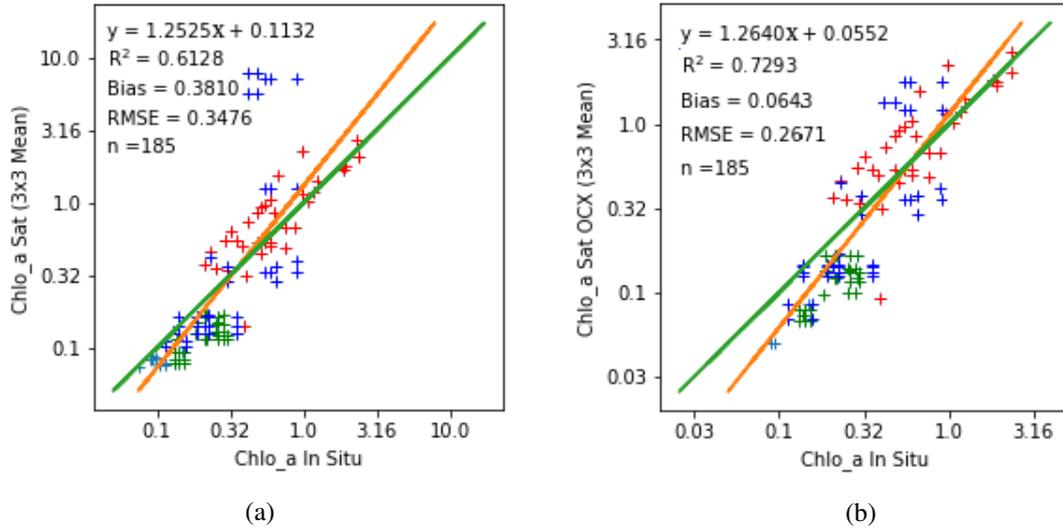
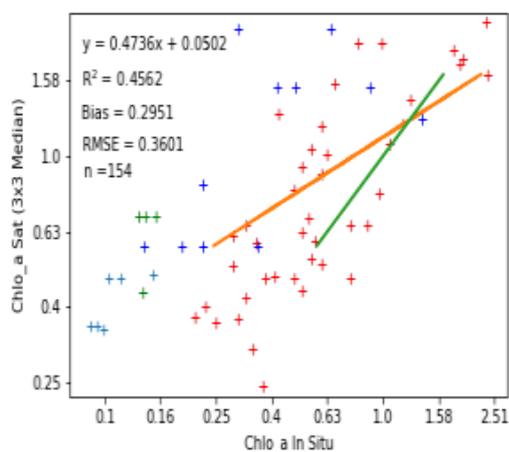


Figura 5: Regressões lineares (log-log) entre dados *in situ* e estimativas por satélite da concentração de clorofila utilizando o algoritmo Standard (a) e o algoritmo OCX (b) e média de janelas de 3x3 pixels para o sensor Aqua MODIS, em laranja. Linha 1:1 em verde. Os matchups estão representados por cor de acordo com o projeto ou cruzeiro, da seguinte forma: Antares Ubatuba - Vermelho, Ilhas Oceânicas - Verde, Recifes Amazônicos - Azul

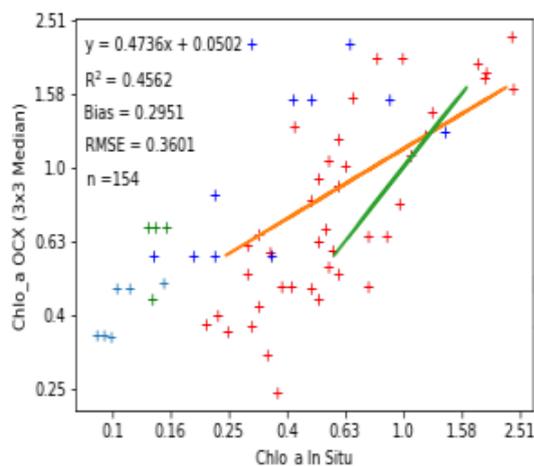
A partir da combinação entre os dados *in situ* e as estimativas de satélite obtidas por meio do sensor AquaMODIS foram produzidas as figuras abaixo. Os pontos considerados na produção dos gráficos e das análises estatísticas estão diferenciados por cor de acordo com o cruzeiro por meio do qual os dados *in situ* foram adquiridos, conforme descrito na tabela 1. As janelas espaciais utilizadas foram de 1x1 e 3x3 pixels, considerando-se a média entre os 9 pixels dessa janela. As linhas cor de laranja descrevem as regressões lineares logaritmizadas para cada conjunto de dados. As linhas verdes representam a relação 1:1 para fins comparativos. A tabela 2 contém os parâmetros estatísticos utilizados para a avaliação da performance dos diferentes algoritmos para cada sensor e janela espacial utilizada.

Tabela 2: Avaliação estatística da performance dos algoritmos do sensor AquaMODIS para a margem continental brasileira. Janela se refere à janela espacial em pixels. N é o número de pontos nas imagens de satélite para os quais foram feitas as comparações entre dados *in situ* e as estimativas do algoritmo. R^2 se refere ao coeficiente de correlação de Pearson. RMSE é a raiz do erro quadrático médio. C.A. é o coeficiente angular e C.L. é o coeficiente linear da regressão. Os valores considerados mais consistentes para cada parâmetro estão em negrito

Algoritmo	Janela	N	R^2	Bias	RMSE	C.A.	C.L.
Standard	1x1	185	0.7124	0.0863	0.2660	1.1129	0.0072
OCX	1x1	185	0.6767	0.0501	0.2644	1.0634	-0.0169
Standard	3x3 (mean)	185	0.6128	0.3810	0.3467	1.2525	0.1132
OCX	3x3 (mean)	185	0.7293	0.0643	0.2671	1.2640	0.0552
Standard	3x3 (median)	154	0.4562	0.2951	0.3601	0.4736	0.0502
OCX	3x3 (median)	154	0.4562	0.2951	0.3601	0.4736	0.0502



(a)

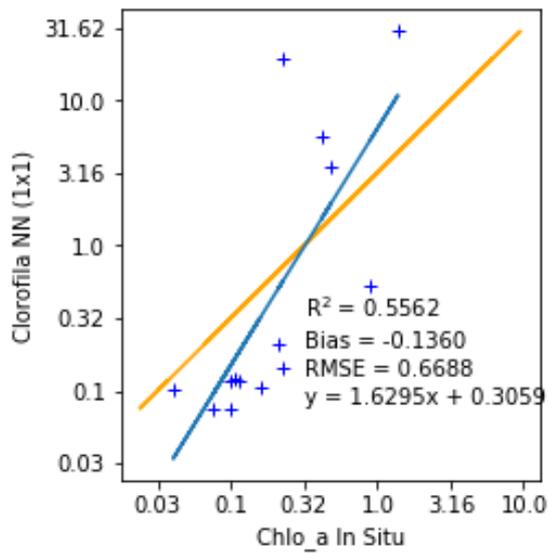


(b)

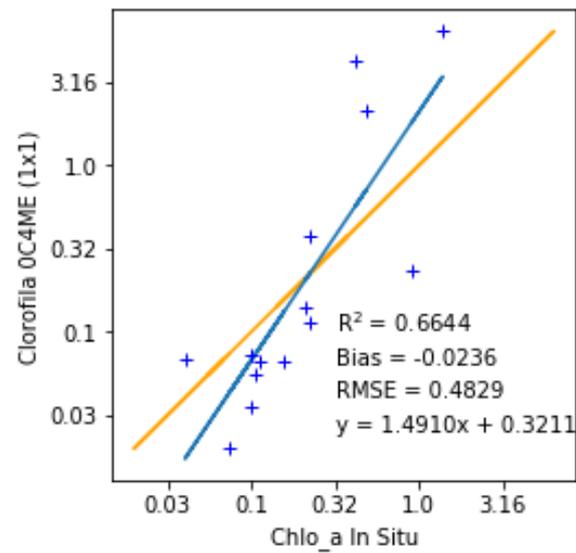
Figura 6: Regressões lineares (log-log) entre dados *in situ* e estimativas por satélite da concentração de clorofila utilizando o algoritmo Standard (a) e o algoritmo OCX (b) e mediana de janelas de 3x3 pixels para o sensor Aqua MODIS, em laranja. Linha 1:1 em verde. Os matchups estão representados por cor de acordo com o projeto ou cruzeiro, da seguinte forma: Antares Ubatuba - Vermelho, Ilhas Oceânicas - Verde, Recifes Amazônicos - Azul

3.2 OLCI

A partir da combinação entre os dados *in situ* e as estimativas de satélite obtidas por meio do sensor OLCI foram produzidas as figuras abaixo. Os 15 pontos considerados na produção dos gráficos e das análises estatísticas são provenientes dos cruzeiros "Recifes Amazônicos" e "Marine E-Tech". As janelas espaciais utilizadas foram de 1x1 e 3x3 pixels, considerando-se a média entre os 9 pixels dessa janela. As linhas azuis descrevem as regressões lineares logaritimizadas para cada conjunto de dados. As linhas cor de laranja representam a relação 1:1 para fins comparativos. A tabela 3 contém os parâmetros estatísticos utilizados para a avaliação da performance dos diferentes algoritmos para cada sensor e janela espacial utilizada.

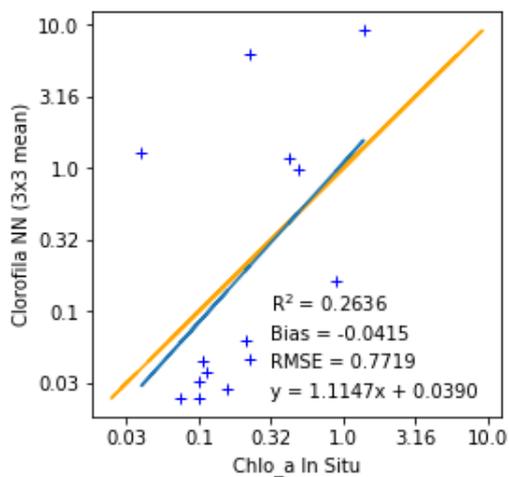


(a)

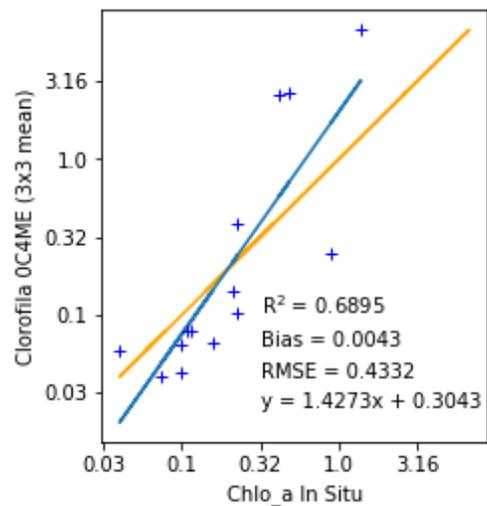


(b)

Figura 7: Regressões lineares (log-log) entre dados *in situ* e estimativas por satélite da concentração de clorofila utilizando o algoritmo de redes neurais (NN) (a) e o algoritmo OC4 (b) e janelas de 1x1 pixel para o sensor OLCI, em azul. Linha 1:1 em laranja

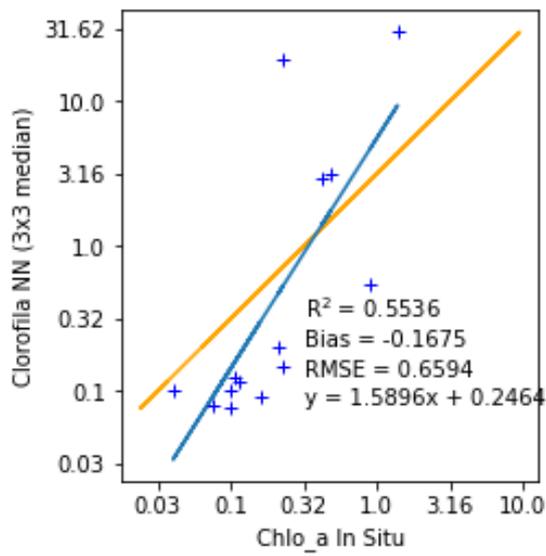


(a)

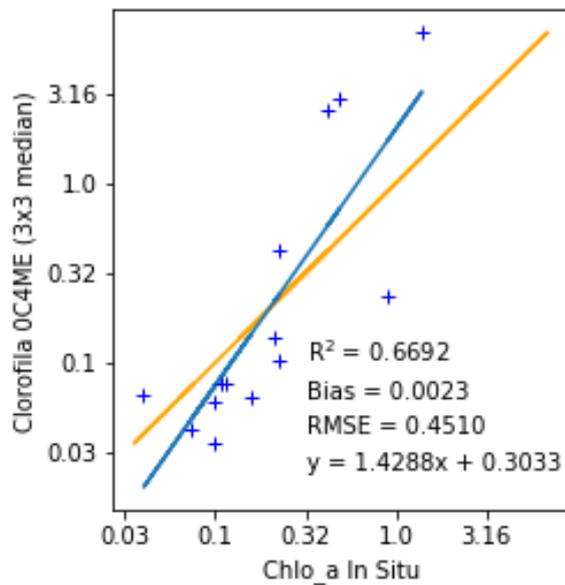


(b)

Figura 8: Regressões lineares (log-log) entre dados *in situ* e estimativas por satélite da concentração de clorofila utilizando o algoritmo de redes neurais (NN) (a) e o algoritmo OC4 (b) e média de janelas de 3x3 pixels para o sensor OLCI, em azul. Linha 1:1 em laranja



(a)



(b)

Figura 9: Regressões lineares (log-log) entre dados *in situ* e estimativas por satélite da concentração de clorofila utilizando o algoritmo de redes neurais (NN) (a) e o algoritmo OC4 (b) e mediana de janelas de 3x3 pixels para o sensor OLCI, em azul. Linha 1:1 em laranja

Tabela 3: Avaliação estatística da performance dos algoritmos do sensor OLCI para a margem continental brasileira. Janela se refere à janela espacial em pixels. N é o número de pontos nas imagens de satélite para os quais foram feitas as comparações entre dados *in situ* e as estimativas do algoritmo. R^2 se refere ao coeficiente de correlação de Pearson. RMSE é a raiz do erro quadrático médio. C.A. é o coeficiente angular e C.L. é o coeficiente linear da regressão. Os valores considerados mais consistentes para cada parâmetro estão em negrito

Algoritmo	Janela	N	R^2	Bias	RMSE	C.A.	C.L
NN	1x1	15	0.5562	-0.1360	0.6688	1.6295	0.3059
OC4	1x1	15	0.6644	-0.0236	0.4829	1.4910	0.3211
NN	3x3 (mean)	15	0.2636	-0.0415	0.7719	1.1147	0.0390
OC4	3x3 (mean)	15	0.6895	0.0043	0.4332	1.4273	0.3043
NN	3x3 (median)	15	0.5536	-0.1675	0.6594	1.5896	0.2464
OCX	3x3 (median)	15	0.6692	0.0023	0.4510	1.4288	0.3033

4 Discussão

Com base nos resultados estatísticos obtidos para o conjunto total de dados avaliado nesse estudo é possível propor algumas inferências à respeito do desempenho dos algoritmos Standard e OCX aplicados em janelas amostrais de 1x1 e média e mediana de 3x3 pixels para o sensor AquaMODIS na área de estudo. Em relação aos resultados referentes aos algoritmos NN e OC4 aplicados em janelas amostrais de 1x1 e média e mediana de 3x3 pixels para o sensor OLCI, o pequeno número amostral é limitante para que seja possível uma avaliação de desempenho, no entanto algumas tendências podem ser observadas a partir dos mesmos. O coeficiente angular da regressão logaritimizada indica a inclinação da reta, sendo uma forma de visualizar a correlação direta entre os valores *in situ* e os estimados pelo algoritmo à medida que se aproxima de 1. O coeficiente linear, por sua vez, equivale ao ponto em que a reta cruza o eixo y, podendo ser útil para avaliar se o algoritmo subestima ou superestima a concentração de clorofila em relação aos dados medidos *in situ*. Valores de coeficiente linear mais próximos de 0 indicam uma maior precisão do produto. O RMSE e o Bias são ferramentas para calcular a incerteza das regressões obtidas, de modo que valores menores desses parâmetros indiquem maior consistência do conjunto de dados. O coeficiente de correlação de *Pearson*, por sua vez, também é um parâmetro estatístico que calcula a correlação entre as duas variáveis avaliadas, sendo que essa correlação é considerada forte conceitualmente a partir de valores maiores que 0.7 e moderada para valores entre 0.5 e 0.7.

4.1 AquaMODIS

O algoritmo OCX em janela de 1x1 pixels apresenta os valores mais consistentes para 3 dos 5 parâmetros estatísticos avaliados, sendo eles o coeficiente angular da regressão, Bias e RMSE. O coeficiente de correlação de *Pearson* indicou uma correlação forte apenas para o algoritmo standard em janela de 1x1 pixels e para o OCX em média de janela de 3x3 pixels, sendo máximo para o último em 0.7293. Em relação ao coeficiente linear, o valor mais próximo de zero foi 0.0072, para o algoritmo standard em janela de 1x1 pixels.

Os gráficos produzidos utilizando média e mediana de janelas de 3x3 pixels assim como os parâmetros estatísticos calculados para esse conjunto de dados apresentaram resultados idênticos por razões ainda não compreendidas até a produção desse documento. Tais dados ainda devem ser reavaliados e reprocessados de modo que se possa compreender melhor os resultados obtidos.

4.2 OLCI

Destacam-se nos resultados obtidos os algoritmos NN e OC4 utilizando média de janelas de 3x3 pixels para o sensor OLCI. Ambos apresentaram os valores mais consistentes para 2 dos 5 parâmetros estatísticos avaliados. Em função do pequeno número amostral, o coeficiente de correlação de *Pearson* obtido foi moderado para ambos os algoritmos em janelas de 1x1 pixels e mediana de janelas de 3x3 pixels, assim como para o OC4 na média de janela de 3x3 pixels. O R^2 foi fraco para o algoritmo NN aplicado na média de janela de 3x3 pixels, embora seu coeficiente angular tenha sido o mais próximo de 1.0 entre os observados para o sensor OLCI, com inclinação igual a 1.1147. O coeficiente linear para o algoritmo NN em média de janelas de 3x3 pixels também foi o mais próximo de zero: sua reta cruza o eixo x em $y = 0.0390$. O algoritmo OC4 aplicado na média de janelas de 3x3 pixels, por sua vez, se destaca pelo maior valor de coeficiente de correlação de *Pearson* (0.6895) e pelo menor RMSE (0.4332).

De maneira geral, o algoritmo OC4 apresentou melhores resultados do que o NN, considerando-se o conjunto dos dados avaliados nos três formatos distintos.

5 Considerações finais

A validação e comparação do desempenho de algoritmos para a estimativa da concentração de clorofila-*a* por satélites na margem continental brasileira é um trabalho extenso e que necessita de constante atualização a medida que novos sensores e algoritmos são desenvolvidos. O presente estudo contribui com um pequeno acréscimo ao conhecimento nessa área, embora seus resultados sejam pouco consistentes em função principalmente do pequeno número amostral avaliado para parte das regiões e algoritmos avaliados. A continuidade e aprofundamento do estudo se dará por meio do trabalho final de graduação do aluno Elias Rosenberg Krausz, do IO-USP, que será concluído até o final desse ano. É planejada também a avaliação do desempenho dos algoritmos de forma local separadamente para a região Norte da margem continental brasileira e para as regiões oceânicas e costeiras da região Sudeste da plataforma, assim como a comparação de algoritmos do sensor *Visible Infrared Imaging Radiometer Suite* (VIIRS). Além disso, é de interesse do estudo a elaboração de uma longa série temporal da concentração de clorofila-*a* na estação "Antares Ubatuba", aplicando os algoritmos que apresentarem melhor desempenho para a região, demonstrando a importância da validação dos algoritmos de clorofila-*a* regionalmente para pesquisa à respeito da produção primária do fitoplâncton marinho.

6 Referências

Referências

1. De Groot RS, Wilson MA, Boumans RM. A typology for the classification, description and valuation of ecosystem functions, goods and services. *Ecological economics*. 2002;41(3):393–408.
2. Field CB, Behrenfeld MJ, Randerson JT, Falkowski P. Primary production of the biosphere: integrating terrestrial and oceanic components. *science*. 1998;281(5374):237–240.
3. Kirk JT. *Light and photosynthesis in aquatic ecosystems*. Cambridge university press; 1994.
4. Gohin F, Stanev E. Annual cycles of chlorophyll-a, non-algal suspended particulate matter, and turbidity observed from space and in-situ in coastal waters. *Ocean Science*. 2011;7(5).
5. Li F, Zhang X, Kondragunta S, Csiszar I. Comparison of fire radiative power estimates from VIIRS and MODIS observations. *Journal of Geophysical Research: Atmospheres*. 2018;123(9):4545–4563.
6. Ciotti AM, Garcia CAE, Jorge DSF. Temporal and meridional variability of Satellite-estimates of surface chlorophyll concentration over the Brazilian continental shelf. 2010;.
7. Castro BM, Brandini FP, Pires-Vanin AMS, Miranda LB. Multidisciplinary oceanographic processes on the Western Atlantic continental shelf between 4 N and 34 S. *The sea*. 2006;11:209–251.
8. Kampel M, Gaeta SA, Lorenzetti JA, Pompeu M, Rudorff FM, Frouin RJ. Bio-optical variability in coastal waters of southeast Brazil. In: *Coastal Ocean Remote Sensing*. vol. 6680. International Society for Optics and Photonics; 2007. p. 668000.
9. Hu C, Lee Z, Franz B. Chlorophyll algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *Journal of Geophysical Research: Oceans*. 2012;117(C1).
10. Brewin RJ, Raitsos DE, Dall’Olmo G, Zarokanellos N, Jackson T, Racault MF, et al. Regional ocean-colour chlorophyll algorithms for the Red Sea. *Remote Sensing of Environment*. 2015;165:64–85.
11. Campbell JW, O’Reilly JE. Metrics for Quantifying the Uncertainty in a Chlorophyll Algorithm: Explicit equations and examples using the OC4. v4 algorithm and NOMAD data. In: *Ocean color bio-optical algorithm mini workshop*. Vol. vol. 4; 2006. p. 1–15.