UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE CIENCIAS MARINAS
INSTITUTO DE INVESTIGACIONES OCEANOLÓGICAS

Parametrización de algoritmos de color del océano y de concentración de clorofila *a* y estudio de su variabilidad utilizando datos satelitales e *in situ* registrados frente a la península de Baja California

T E S I S

QUE PARA CUBRIR PARCIALMENTE LOS REQUISITOS NECESARIOS PARA OBTENER EL GRADO DE

DOCTOR EN CIENCIAS EN OCEANOGRAFÍA COSTERA

PRESENTA

PATRICIA ALVARADO GRAEF

ENSENADA, BAJA CALIFORNIA, MÉXICO, SEPTIEMBRE DE 2020

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE CIENCIAS MARINAS

Parametrización de algoritmos de color del océano y de concentración de clorofila *a* y estudio de su variabilidad utilizando datos satelitales e *in situ* registrados frente a la península de Baja California

INSTITUTO DE INVESTIGACIONES OCEANOLÓGICAS

T E S I S

QUE PARA CUBRIR PARCIALMENTE LOS REQUISITOS NECESARIOS PARA OBTENER EL GRADO DE

DOCTOR EN CIENCIAS EN OCEANOGRAFÍA COSTERA

PRESENTA

PATRICIA ALVARADO GRAEF

APROBADA POR:

DRA. BEATRIZ MARTÍN ATIENZA

DIRECTORA DE TESIS

DR. REGINALDO DURAZO ARVIZU

SINODAL

DR. RAMÓN SOSA ÁVALOS

SINODAL

DR. RAFAEL HERNÁNDEZ WALLS

SINODAL

DR. FRANCISCO DELGADILLO HINOJOSA

SINODAL

RESUMEN

El color del océano está dado por los diferentes componentes del agua de mar. Datos de sensores remotos de color del océano se usan frecuentemente para inferir las Propiedades Ópticas Inherentes (POIs) del agua de mar y la concentración de clorofila-*a* (Chl-*a*) utilizando algoritmos matemáticos. Estos algoritmos funcionar mejor cuando se regionalizan con datos locales y, por lo tanto, responden a las propiedades particulares de una región. En este trabajo, se utilizaron datos tomados en la porción sur del Sistema de la Corriente de California (SCC) frente a la Península de Baja California para probar algunos de los algoritmos existentes y, posteriormente, establecer nuevos parámetros y/o nuevos algoritmos que incluyen las características propias del área, así, brindando herramientas que dan mejores resultados para la región.

En primer lugar, utilizamos datos *in situ* y satelitales de 1999 para probar el algoritmo de inversión de POIs de Garver y Siegel (GSM). Se propusieron parámetros para las diferentes estaciones del año. Esto produce una mejora en las predicciones de las POIs a comparación del modelo de Garver y Siegel. En segundo lugar, varios algoritmos de Chl-*a* se probaron con datos de 1998 a 2016. Se propusieron adaptaciones a los algoritmos obteniendo el mejor resultado cuando se consideraron las condiciones climáticas (El Niño/La Niña).

El monitoreo de variables oceanográficas permite el análisis del estado actual de los ecosistemas oceánicos y, por lo tanto, es una herramienta invaluable para su manejo. Se requieren series de tiempo largas para comprender la variabilidad de las propiedades bio-ópticas y sus consecuencias. Los algoritmos regionales funcionales que dan resultados adecuados y probados se pueden utilizar para construir series de tiempo largas de POIs y de Chl-*a*. Estos resultados dan información invaluable en, entre otras, la ecología del fitoplancton, el ciclo del carbono, cambio climático y productividad oceánica.

Abstract

Ocean color is determined by the different components of seawater. Remotely sensed ocean color data are commonly used to infer the Inherent Optical Properties (IOPs) of seawater and concentration of Chlorophyll-*a* (Chl-*a*) using mathematical algorithms. These algorithms work better when local data is used to regionalize them, therefore, responding to the properties of a particular area. In this paper, we use data from the southern part of the California Current System (CCS) off the Baja California Peninsula to test some of the existing algorithms and, then, establish new parameters and/or new algorithms that comprise the unique characteristics of the area providing tools that yield better results in the region.

First, we use *in situ* and satellite data from 1999 to test the Garver and Siegel IOPs Inversion Model (GSM). Season sensitive parameters for a new algorithm are proposed. These improve on the results yielded with the Garver and Siegel Model. Second, several Chl-*a* algorithms were tested using data ranging from 1998 to 2016. Adaptations to one of the algorithms were proposed and the best results were obtained when the algorithm was adapted considering climatic conditions (El Niño/La Niña).

Monitoring oceanographic variables allows the analysis of the state of the oceanic ecosystem and, therefore, is invaluable for their management. Long time-series of these variables are required to understand the variability of bio-optical properties and its consequences. Functional regional algorithms that yield adequate and proven results permit the construction of long term-series of IOPs and Chl-*a*. The results obtained with regional algorithms will provide invaluable information on, among others, phytoplankton ecology, the carbon cycle, climate change and oceanic productivity.

Table of Contents

INTRODUCTION

Constant monitoring of the ocean provides information to describe and understand the state of the ocean, its changes and its management. There have been many efforts worldwide to obtain samples from the ocean. For example, in 1949 the *California Cooperative Oceanic Fisheries Investigations* (CalCOFI, https://calcofi.org/) data collecting began. Its main goal was to study the sardine population. Since then, CalCOFI has conducted quarterly collecting of oceanographic data in front of California. In 1997, the *Investigaciones Mexicanas de la Corriente de California* (IMECOCAL, https://imecocal.cicese.mx/wp/) project began with sampling in front of the Baja California peninsula. Even with all these efforts, the spatial and temporal distribution of the data is sparse. Collecting data requires a lot of time, work and money. Because of this, remote sensing has arisen as a way of creating long time series of regularly spaced data. Remote sensing coupled and verified with *in situ* data has proven to be a very efficient way of obtaining information in order to study phenomena at long spatial and temporal scale.

The distribution and variability of phytoplankton in the ocean are important as these organisms are the basis of the trophic web of the ocean and play a very important role in the carbon cycle. Ocean color, dependent on the different components of seawater, has been used to estimate the concentration of Chlorophyll-*a* (Chl-*a*), the main photosynthetic pigment, as well as primary productivity [1], [2]. Ocean color has been described for more than three centuries though it has only been measured since 1890 [3]. By the end of the XIX century, the Forel-Ule scale for the classification of natural waters was introduced as the first remote devise to quantify the color of the ocean [4]. By the end of the 1970's, data of the color of the ocean as a proxy of the phytoplankton biomass started to be collected by remote sensors. The Coastal Zone Color Scanner (CZCS) was the first sensor developed. It was put into orbit in 1978 by NASA (National Aeronautic and Space Administration), [5]. Since then, sensors have been evolving but each satellite sensor still has different optical bands, sensitivities, noise patterns and overpass times [6] as well as different uncertainties due to calibration and algorithms applied [7]. By 2019 a 22-year record or ocean-color has been comprised, but it includes many sensors with different characteristics that have operated at different times [8]. It has been made clear that a merger of data is necessary to improve the coverage of satellite data and provide consistent time series [8], [9], [10]. Operational satellite sensor constellations, such as Copernicus, provide ocean color data as elements of integrated, sustained observing system [8]. Even with the refinement of the sensors, much work is still needed to decrease the uncertainties associated to the data, for instance, the evaluation and adaptation of algorithms to particular regions.

The relationship between the remote sensed data and the components of seawater is modeled through mathematical algorithms. These can be empirical: based on the statistics of the data; and semianalytic: based on theoretical knowledge with empirical parameters [11]. Also, they can use an atmospheric correction or relate a remotely sensed data directly to the optical property of interest [12]. Furthermore, algorithms can be based on radiative transfer simulations (following the incident photon flux) or quasi-single scattering approximation (simplified analytical model that assumes most scattering is done in the forward direction) [12]. Ocean color algorithms are based on the relationship between the downwelling irradiance ($E_d(\lambda)$), that is the radiant flux received by a surface per unit area and the upwelling radiance ($L_u(\lambda)$) that is, the radiant flux emitted, reflected or transmitted by a given surface, per unit solid angle per unit projected area. The ratio of upwelling radiance to downwelling irradiance is called reflectance ($R(\lambda)$) and generally describes the relationship between the quality and quantity of light that arrives at the ocean and the light that leaves the ocean after being modified by it. In other words, reflectance is a measure of the light that exits the ocean normalized by the incident irradiance.

In this work we analyze ocean color algorithms for the IMECOCAL region in front of the Baja California peninsula. These algorithms use remotely sensed reflectance and *in situ* data. In the first chapter, a new semi-analytic algorithm that relates the Inherent Optical Properties (IOPs) ratio to reflectance is proposed. Given the unique properties observed, different parameters were obtained for each season with better results than when a global algorithm is applied to the data. To test the algorithm, Chl-*a* data was used as a proxy. The results obtained show that the algorithm works well for the area. This first chapter uses a single source for satellite data and focuses on the relationship between the in water properties (IOPs) with remotely sensed data (reflectance).

On the second chapter, various Chl-*a* empirical algorithms were tested. Using the algorithm that gave the best fit for the area, new parameters were obtained. The regional algorithm differed from the global on the most eutrophic and oligotrophic waters. Afterward, new parameters were sought by separating the database according to (a) dynamic boundaries in the area; (b) bio-optical properties and (c) climatic condition (El Niño/La Niña) with the latter giving the best results. This chapter uses merged satellite data and focuses on producing the best Chl-*a* estimation.

Both chapters contribute with functional algorithms that yield better results for the IMECOCAL region. Along this research, it is constantly shown that the IMECOCAL area has different bio-optical properties than those of the global dataset and, more importantly, than those of the CalCOFI region. Thus, we provide a tool that can be used to create consistent time series of the bio-optical properties which, in turn, can be used to study the effect of climate change (*i.e.* the decrease in phytoplankton biomass due to global warming and its accompanying community composition shift).

References

[1] IOCCG, (1998). Minimum Requirements for an Operational, Ocean-Colour Sensor for the Open Ocean. *Reports of the International Ocean-Colour Coordinating Group,* No. 1. IOCCG, Dartmouth, Canada.

[2] IOCCG, (2006). Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. Lee, Z.-P. (ed.), *Reports of the International Ocean-Colour Coordinating Group*, No. 5, IOCCG, Dartmouth, Canada.

[3] Wermand, M. R. and H. J. van der Woerd, (2010). Ocean colour changes in the North Pacific since 1930. *Journal of the European Optical Society*-Rapid Publications 5:100,145.

[4] Wermand, M. R. and H. J. van der Woerd, (2010) Spectral Analysis of the Forel-Ule Ocean Colour Comparator Scale. *Journal of the European Optical Society*-Rapid Publications 5:100,145

[5] IOCCG, (2000). Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters. *Reports of the International Ocean-Colour Coordinating Group*, No. 3, IOCCG, Dartmouth, Canada.

[6] Maritorena, S. and D. A. Siegel, (2005). Consistent merging of satellite ocean color data sets using a bio-optical model. *Remote Sensing Environ*. 94, 429–440.

[7] McClain, C. R., (2009). A decade of satellite ocean color observations. *Annu.Rev.Mar. Sci.* 1, 19–42.

[8] Groom, S. , S Sathyendranath, Y. Ban, S. Bernard, R. Brewin, V. Brotas, C. Brockmann, P. Chauhan, J. L Choi, A. Chuprin, S. Ciavatta, P. Cipollini, C. Donlon, B. Franz B, X. He, T. Hirata, T. Jackson, M. Kampel, H. Krasemann, S. Lavender, S. Pardo-Martinez , F. Mélin, T. Platt, R. Santoleri, J. Skakala, B. Schaeffer B, M. Smith, F. Steinmetz, A. Valente M. and Wang, (2019) Satellite Ocean Colour: Current Status and Future Perspective. *Front. Mar. Sci*. 6:485. doi: 10.3389/fmars.2019.00485

[9] Kahru, M., R. M. Kudela, C. R. Anderson, and B. G. Mitchell. (2015). Optimized merger of ocean chlorophyll algorithms of MODIS-Aqua and VIIRS. *IEEE Geoscience and Remote Sensing Letters*, 12(11), 2282-2285. doi: 10.1109/lgrs.2015.2470250

[10] IOCCG, (2007). Ocean-color data merging. In: Gregg, W.(Ed.), *Report of the International Ocean-Color Coordinating Group*, No. 6 , IOCCG, Dartmouth, Canada.

[11] O'Reilly, J. E., S. Maritorena, B. G. Mitchell, D. A. ,Siegel, K. L.,Carder, S. A. Garver, M. Kahru, C. R. McClain. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *J.Geophys.Res.*103, 24937–24953.

[12] Werdell, P. J. , L. I. W. McKinaa, E. Boss, S. G. Ackleson, S. E. Craig, W. W. Gregg, Z. Lee, S. Maritorena, C. S. Roesler, C. S. Rousseaux, D. Stramski, J. M. Sullivan, M. S. Twardowski, M. Tzortziou, X. Zhang, (2018). An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing. *Progress in Oceanography*, vol.160, pp.186-212. doi: 10.1016/j.pocean.2018.01.001.

Chapter I. An Ocean Color Algorithm Based on Power Functions to Retrieve Inherent Optical Properties from Remotely Sensed Data off the Baja California Peninsula, México [1]

Abstract— Ocean color is determined by the different components of seawater. Satellite ocean color data are commonly used to infer Inherent Optical Properties (IOPs) of seawater and concentration of chlorophyll *a* (Chl-*a*). Many of the algorithms that relate satellite data with the components of seawater use global data to determine their parameters. Regionalizing the algorithms is recommended to improve the retrieval of IOPs from satellite data. In this paper, we use *in situ* and satellite data from the southern part of the California Current System (CCS) off the Baja California Peninsula during 1999 to test the Garver and Siegel IOPs Inversion Model (GSM). The strong seasonal variability observed in the data was not reproduced by this algorithm. Thus, we propose a new regional algorithm whose order and coefficient depend on seasonal variability. This algorithm includes the linear form of the GSM model but allows changes in two parameters according to the season. These two parameters are the magnitude and spectral shape of the relationship of remote sensing reflectance with the ratio of the backscattering coefficient and the sum of backscattering and absorption coefficients. In this area both parameters increase gradually from January to October. The new algorithm improves the retrieval of the IOPs ratio from a $\boldsymbol{r^2}$ below 0.5 with the GSM model to a $\boldsymbol{r^2}$ than ranges from 0.87 in April and October to 0.93 and 0.94 in January and August, respectively. Furthermore, Chl-*a* data was used as a proxy to evaluate the new algorithm. Absorption data retrieved from remotely sensed data for 1999, 2008 and 2011 relate well to the *in situ* Chl-*a* data.

Index Terms— Remote sensing, algorithms, color, water, sea measurements, satellites.

INTRODUCTION

The optical properties of seawater, responsible for its color, can be divided into inherent and apparent optical properties (IOPs and AOPs, respectively, [1], [2]). Two IOPs are the absorption coefficient, $a(\lambda)$, and the scattering coefficient, $b(\lambda)$. They differ from AOPs in that the first depend only on the characteristics of the medium and not in the structure of the light field. The IOPs determine the quantity and spectral quality of water leaving radiance which is measured by the remote sensor. Hence, the water leaving radiance is related to the processes that occur in the uppermost part of the water column. This relates the IOPs with an AOP, the remote sensing reflectance, $r_{rs}(\lambda)$, defined as the ratio between the water leaving radiance, $L_w(\lambda)$, and downwelling irradiance, $E_d(\lambda)$, all of them referring to just beneath the surface of the water,

$$r_{rs}(\lambda) = \frac{L_w(\lambda)}{E_d(\lambda)} \tag{1}$$

Remote sensors have allowed scientists to study the properties and phenomena of the ocean at a synoptic scale as well as to generate long time series. The use of this type of sensors

[1] P. Alvarado-Graef, B. Martín-Atienza, R. Sosa-Ávalos and R. Durazo, "An Ocean Color Algorithm Based on Power Functions to Retrieve Inherent Optical Properties From Remotely Sensed Data Off the Baja California Peninsula, Mexico," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 3, pp. 1868-1876, March 2020, doi: 10.1109/TGRS.2019.2949946.

has been centered in physical properties such as temperature, sea level and biological properties like the concentration of chlorophyll *a* (Chl-*a*), which is indicative of the abundance of phytoplankton in seawater [3]. The study of ocean color at a global and/or regional scale contributes to the understanding of phytoplankton ecology and oceanic productivity [3]. The goal of the ocean color remote sensors is to obtain measurements at temporal and spatial scales unattainable by the common sampling methods, such as oceanographic cruises, and with them further the understanding of processes such as the oceanic response to climate change, the health of fisheries, and the carbon cycle among others [4].

The authors in [4] describe two kinds of bio-optical algorithms to retrieve IOPs from satellite measurements: (a) those that apply an atmospheric correction and are a function of reflectance and (b) those that relate at-sensor radiance directly to the IOPs. Furthermore, the relationship between remotely-sensed ocean color data and the components of seawater can be described by radiative transfer simulations and quasi-single scattering approximation. The latter approximates remote sensing reflectance as a function of spectral IOPs [4] and use empirical functions for certain parts of the model [5]. This technique has resulted in an improvement in the understanding of the relationship between remote sensing reflectance and backscattering to absorption ratio [4], [6], [7]. This is done under the assumption that the upwelling light field depends mainly upon the absorption considering, thusly, that most scattering is forward scattering [4]. These models can be used to determine IOPs based on measurements of satellite reflectance, $r_{rs}(\lambda)$, providing valuable information about biogeochemical properties [8], [4]. In [9], the authors recognize that one of the main issues in using remote sensors is to extrapolate *in situ* data to larger areas and longer time records. To solve this, one approach has been to divide the ocean into regions with similar bio-optical characteristics [10]. Thus, given that the empirical parts of the models depend on season, region and oceanographic conditions [11], the best results of these algorithms arise when they are adapted to a given area [12].

The California Current System (CCS) is the eastern limb of the North Pacific gyre [13] and it presents the usual characteristics of an eastern boundary current such as coastal upwelling, eddies and fronts. All these features impact the ecosystem through transport and mixing of water and, therefore, a variation in nutrient concentration [14]. The southern border of the CCS has been described as a transitional zone where subarctic waters and subtropical waters meet. Its location, along the Baja California Peninsula's coasts, varies seasonally [15], producing strong seasonal variability in most of its oceanographic characteristics. In this region, there is a predominant equatorward surface flow, with a poleward subsurface countercurrent adjacent to the shelf slope. The convergence of different water masses and their interactions determines the conditions that modulate biological diversity and chemical variability [18], [19]. The IMECOCAL (spanish acronym for Mexican Research Program of the California Current) program focuses on research in this transition zone. The main objective of the IMECOCAL studies is to obtain realistic predictions of the response of the pelagic ecosystem to global and regional climatic changes and anthropogenic activities such as fisheries.

In this paper, we use *in situ* data of absorption coefficient from the IMECOCAL database and satellite data of remote sensing reflectance from NASA (http://oceancolor.gsfc.nasa.gov/) to obtain a better prediction of the IOPs in the form of the quotient between backscattering coefficient, $b_b(\lambda)$, and sum of absorption coefficient and backscattering coefficient, $a(\lambda) + b_b(\lambda)$, as

$$X(\lambda) = \frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)}, \qquad (2)$$

This, in turn, can be used to study the temporal variability of the oceanographic properties of this area, contributing to the study the biological processes in the Southern California Current System (SCCS) including the phytoplankton biomass. In consequence, this will further the knowledge of the base of the ecosystem in the area providing a better understanding of the way the ecosystem functions.

The main objective of this research is to use an algorithm that relates remote sensing reflectance data with the quotient of IOPs, as shown in equation (2), for the prevailing conditions of the SCCS off the Baja California Peninsula during 1999. This algorithm should account for the strong seasonal variability observed in the oceanographic parameters. In order to achieve this, we first tested the Garver and Siegel IOPs (GSM) inversion model proposed by [16] and optimized by [17]. This algorithm has been successfully used in many other oceanic regions around the world. However, when used for this transition zone it does not reproduce the observations adequately for all seasons. Consequently, the algorithm needs to be regionalized. The use of new parameters for the algorithm does not yield the spectral shape expected for the IOPs quotient. Hence, we propose a new algorithm with two free parameters that can describe the seasonal variability within this transition zone. The strength of this new algorithm is that it includes the linear form of the GSM model, while also allowing for a change in magnitude and spectral shape of the relationship between remote sensing reflectance and the IOPs. The new model is inverted and the spectrum of IOPs quotient is retrieved with favorable results. Given the lack of absorption coefficient data to validate the algorithm, Chl-a data was used as a proxy. Both the spectral shape of the IOPs quotient as well as the relationship between absorption coefficient and Chl-a validate the results of the new algorithm.

METHODS AND DATA

The ocean color algorithm tested in this paper is the quasi-single scattering approximation IOPs inversion model of Garver and Siegel [16] and Maritorena *et al.* [17]. This algorithm assumes that the relationship between $R_{rs}(\lambda)$ (measured above the surface of the water) and the IOPs ratio is well known and stable, and can be described as

$$R_{rs}(\lambda) = 0.54 \sum_{i=1}^{2} l_i \, [X(\lambda)]^i \tag{3}$$

The coefficients $l_1 = 0.0949$ sr^{-1} and $l_2 = 0.0794$ sr^{-1} of the quadratic equation (3) were proposed by [20] and represent geometrical factors [8], [7], sea surface properties and the shape of the volume scattering function [4]. These coefficients are commonly used [16], [17], [21] and [4] among many others. To achieve better results for the area of interest, these two parameters may be modified. The coefficient 0.54 is a factor that relates the satellite-measured quantity, $R_{rs}(\lambda)$, with the *in situ* value measured just beneath the ocean surface, $r_{rs}(\lambda)$ (equation 1) and it relates to the transmissibility of the air-water interface [11].

In the quotient of equation (2), both the absorption and backscattering coefficients are decomposed into additive components as

$$a(\lambda) = a_w(\lambda) + a_p(\lambda) + a_g(\lambda) \tag{4}$$

where

$$a_p(\lambda) = a_\phi(\lambda) + a_d(\lambda)$$

and

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda). \tag{5}$$

The subscripts *w*, *p, d,* ϕ and *g* refer to the absorption or backscattering due to water, particulate matter, detritus (de-pigmented particulate matter), phytoplankton, and colored dissolved materials, respectively.

According to several studies, [22], [23], [24] and [25], the spectral shape of the absorption coefficients due to detritus, $a_d(\lambda)$, and colored dissolved materials, $a_g(\lambda)$, can be described together, with the same single exponential function

$$a_{dg}(\lambda) = a_{dg}(\lambda_0)exp[S(\lambda - \lambda_0)], \quad (6)$$

where $a_{dg}(\lambda_0)$ is the magnitude of this coefficient at a reference wavelength $\lambda_0 = 443\ nm$ and S is the exponential decay constant. The value of S can have regional values in order to achieve the best results. It has been reported that its value is close to -0.020 nm^{-1} [8], [26], [27].

The spectral shape of the backscattering coefficient due to the total particulate materials is frequently modeled as a function of wavelength as λ^{-n} [16], [28], [29]. In [16], the authors tested the sensitivity of their algorithm using values for *n* of 0, 1 and 2 and concluded that the choice of *n* has little effect on the results. Therefore, and given that we have no backscattering data, we used the model proposed by Carder *et al.* [27], namely,

$$b_{bp}(\lambda) = A\left(\frac{551}{\lambda}\right)^B, \quad (7)$$

where A is the magnitude of the backscattering coefficient of the particulate materials at a wavelength of 551 nm; B determines the spectral shape of said coefficient and its value is related to the size of the particles that cause the backscattering [16]. A and B are calculated by the empirical relationships as a function of $R_{rs}(\lambda)$, as shown in equations (8) and (9)

$$A = A_0 + A_1\, R_{rs}(551) \quad (8)$$

and

$$B = B_0 + B_1 \frac{R_{rs}(443)}{R_{rs}(488)}, \quad (9)$$

where A_0, A_1, B_0 and B_1 are empirically derived values [27]. Though it would be ideal to obtain these values for the region, no backscattering data are available and a model has to be applied. The model described by [27] allows the backscattering coefficient to be calculated based upon reflectances more than a model that only depends on wavelength.

Therefore, to regionalize the IOPs inversion model described above, we use *in situ* and satellite data. The *in situ* data are the spectral absorption coefficients. These were obtained by collecting two liters of seawater at depths corresponding to the following percentages of surface irradiance: 100%, 50%, 30%, 20%, 10% and 1% during four IMECOCAL cruises conducted in

1999: winter (January), spring (April), summer (August) and autumn (October) (Fig. 1 and Table I). On each sample, seawater was filtered with Whatman GF/F glass fiber filters using positive pressure and immediately frozen in liquid nitrogen for later analysis. Total particles spectra collected by the filter were measured in a Varian Cary 1E UV-Visible spectrophotometer according to the method described by [30], [31] and [32]. These procedures were used to determine the absorption coefficient due to particles, $a_p(\lambda)$, and to detritus, $a_d(\lambda)$. Absorption spectra were corrected for the light pathlength amplification caused by light scattering in the filter with the equations for GF/F filters recommended by [31]. The absorption coefficient of phytoplankton, $a_\phi(\lambda)$, was obtained by subtracting $a_d(\lambda)$ from $a_p(\lambda)$: $a_\phi(\lambda) = a_p(\lambda) - a_d(\lambda)$. The absorption data was described in more detail in [36] and [38]. Afterwards, we calculated the arithmetical average of the data collected at 100% and 50% of the surface irradiance.

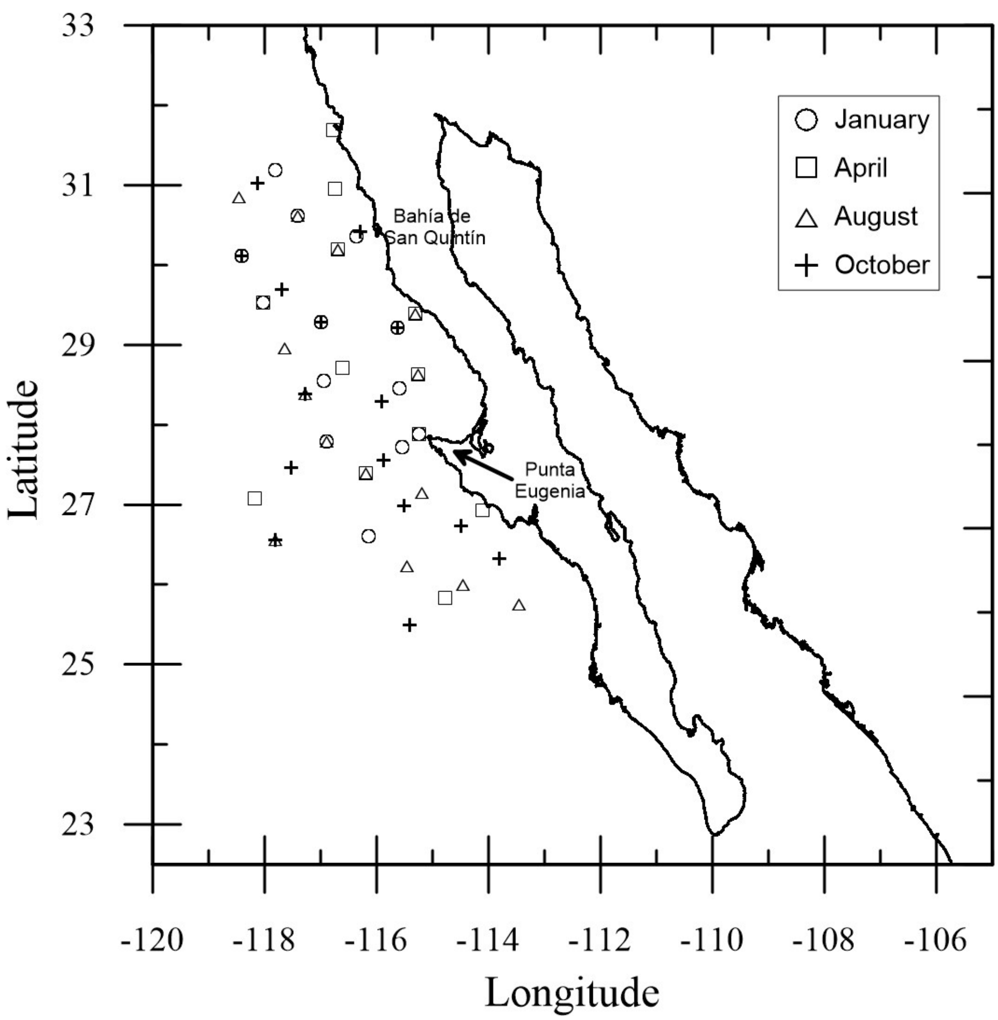

Fig. 1. Station maps for IMECOCAL surveys conducted during 1999 in January, April, August and October.

TABLE I

DATES AND NUMBER OF STATIONS ON EACH OF THE FOUR SURVEYS OF IMECOCAL 1999

Survey	Dates	Number of stations
January	January 14thto February 2nd	13
April	March 30th to April 17th	12
August	August 8thto August 22nd	14
October	October 3rd to October 22nd	15

For Chl-a, water was collected in 5 L Niskin bottles coupled in a rosette. Silicone bands were used for closure. For the samples of 1999, water was collected at to the following percentages of surface irradiance: 100%, 50%, 30%, 20%, 10% and 1% whereas for 2008 and 2011 water was collected at 1 m, 10 m, 20 m, 50 m, 100 m, 150 m and 200 m of depth. Two liters of water at each depth were filtered through Whatman GF/F at a pressure below 150 mm Hg. The filters were then placed in Histoprep capsules and frozen in liquid nitrogen until their analysis. Chlorophyll was extracted with acetone at 90% and at 4°C in darkness for 24 hours as recommended by [39]. Chl-a concentration (mg m^{-3}) was determined by the fluorometric method [40], [41], with a Turner Designs A10 fluorometer for 1999 samples and a Turner Designs trilogy fluorometer for 2008 and 2011. Both fluorometers were calibrated with pure Chl-a (Sigma). In this paper, only the surface data of chlorophyll concentration were used. Water was collected from the network of stations for IMECOCAL cruises (Fig. 1). More information on the chlorophyll data can be seen in [42].

Regarding satellite data, level 3 monthly averages of remotely-sensed reflectance, $R_{rs}(\lambda)$, at six SeaWIFS bands (412, 443, 490, 510, 555 and 670 nm), were obtained from NASA´s Ocean Color Web (http://oceancolor.gsfc.nasa.gov/) with a 9 km ground resolution.

RESULTS

A. Backscattering and absorption coefficients

To obtain the backscattering coefficient (5), we used the $b_{bw}(\lambda)$ values published by [33] and [34], while $b_{bp}(\lambda)$ was modeled according to equation (7) where parameters A and B were calculated. These were obtained for each IMECOCAL survey of 1999 using satellite data. Afterwards, an arithmetical average and the standard deviation of A and B were obtained for each cruise. The results are shown in Table II.

TABLE II
MAGNITUDE (A) AND SPECTRAL SHAPE (B) OF THE BACKSCATTERING COEFFICIENT OF THE PARTICULATE MATERIALS, $b_{bp}(\lambda)$, FOR EACH IMECOCAL 1999 CRUISE ACCORDING TO THE EMPIRICAL MODEL PRESENTED BY [29]

Survey	$\boldsymbol{A}\ (\boldsymbol{m^{-1}})$	$\boldsymbol{B}$
January	0.0008 ± 0.0006	1.9 ± 0.4
April	0.0018 ± 0.0008	2.0 ± 0.4
August	0.0017 ± 0.0007	2.3 ± 0.4
October	0.0013 ± 0.0004	2.2 ± 0.4

Garver and Siegel [16] found that the values of A vary between 0.001 and 0.003 m^{-1}. In our study area, the arithmetical average of A varies within $0.0008 < A < 0.0018$ m^{-1} for all surveys. The magnitude (A) was lowest in January and highest in April. It was only during January that the value of A falls slightly below the range presented by these authors.

Parameter B commonly varies between 0 and 2 with higher values corresponding to small particles with sizes between 0.2 and 0.5 μm [16]. Accordingly, B is expected to be greater than 2 in oligotrophic regions [27], where backscattering is due to small particles [35] with low Chl-a. In areas with Chl-a concentrations higher than 10 mg m^{-3}, B is usually 0 as a result of the presence of larger phytoplankton cells which could have packaging pigments. In [36] the authors observed values of Chl-a in the range of 0.42 to 1.50 mg m^{-3} for the 1999 IMECOCAL surveys. Therefore, values of B in these surveys are expected to be close to 2. B values calculated in this paper vary within $1.9 < B < 2.2$ (See Table II). Regarding the seasonal variation, B was lowest during January and April and highest in August and October. B values of August and October are related to the presence of small cells such as flagellates that were dominant during these surveys. This type of organism is typical of warm water. In April, B is lower due to the increase in Chl-a concentration and the presence of larger phytoplankton cells, such diatoms and dinoflagellates [9].

The spectra of the absorption coefficients, $a_{\phi}(\lambda)$ and $a_d(\lambda)$, the remote sensing reflectance, $R_{rs}(\lambda)$, and the quotient $X(\lambda) = \frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)}$, for each survey are presented in Fig. 2. The magnitude of $a_{\phi}(\lambda)$ was largest in April. In August and October, locations closely related to regions of upwelling, as San Quintín and Punta Eugenia (see Fig. 1), stood out because of high $a_{\phi}(\lambda)$ magnitudes. Moreover, $a_d(\lambda)$ was highest in October due to the presence of Santa Ana winds [9], and it showed little variation among the other three surveys. Parameter S was obtained from the $a_d(\lambda)$ *in situ* data with the following results: 0.0108, 0.0109, 0.0116 and 0.0162 nm^{-1} for January, April, August and October, respectively. The highest variability of $R_{rs}(\lambda)$ was always in short wavelengths (412 and 443 nm). During April, we observed the highest values of $a_{\phi}(\lambda)$. Also, the low $a_{\phi}(\lambda)$ values observed during August and October are related to the expected seasonal increase in surface temperature of seawater and, thus, the presence of smaller phytoplankton characteristic of tropical zones. Likewise, except for one station, August showed the lowest values of $a_{\phi}(\lambda)$ and the highest of $R_{rs}(\lambda)$. During April, the lower values of $R_{rs}(\lambda)$ are related to an increase in Chl-a and larger phytoplankton cells, therefore, an increase in $a_{\phi}(\lambda)$. In contrast, higher values of $R_{rs}(\lambda)$ are a result of the decrease in Chl-a and the $a(\lambda)$, accordingly. Finally, $X(\lambda)$ had its minimal values during October and its maximum ones in January. The variability of $X(\lambda)$ also decreased from January to October. The highest $a_{\phi}(\lambda)$ was observed in April and the highest values of $R_{rs}(\lambda)$ in August (Fig. 2) which concurs with the results presented by [38].

As stated by [37], the highest phytoplankton biomass in the four cruises in 1999 was during April, the season of intense coastal upwelling associated with relatively high concentrations of Chl-a and colder surface waters of increased salinity [38]. In [38], the authors found the lowest primary productivity during August.

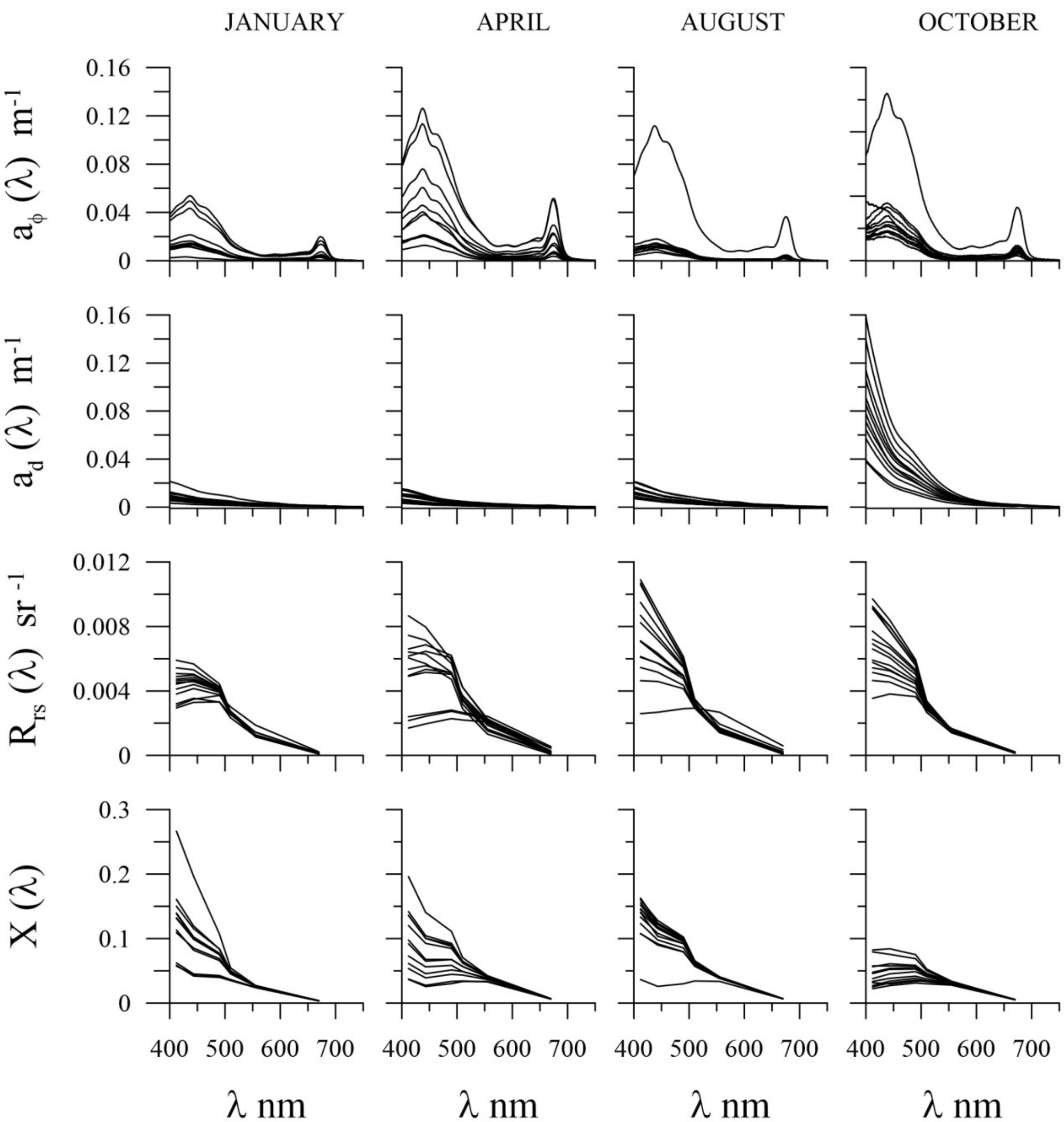

Fig. 2. Spectra of the absorption coefficient for: (a) phytoplankton, (b) detritus, (c) remote sensing reflectance, and (d) $X(\lambda)$ retrieved in equations (2). Each solid line in the graphs shows each station sampled during the IMECOCAL 1999 cruises (Fig. 1).

In October, aeolian transport from arid areas to the ocean due to easterly Santa Ana winds was reported by [38] and this was associated to high values of $a_p(\lambda)$ and $a_d(\lambda)$ because of the presence of non-pigmented particles. These authors also concluded that, even when the bio-optical properties of the ocean changed because of the transport of particles by strong land-to-sea winds, it is possible that the estimation of Chl-a using remote sensors in the IMECOCAL region was not substantially altered. Nevertheless, a difference on the scattering and attenuation is expected. The model used in our paper for A and B [27] is based on ratios of remote sensing reflectance. The effect of the Santa Ana Winds is not obvious in our results (Table II). However, the effect is evident in $a_d(\lambda)$ and X(λ) as shown in Fig. 2. This suggests that X(λ) is sensitive to variations in $a_d(\lambda)$ that cannot be detected with $R_{rs}(\lambda)$ only. Therefore, it is possible that the backscattering model (7), based on ratios of remote sensing reflectances, may not be appropriate under these conditions.

B. Relationship between IOPs and remote sensing reflectance

The functional relationship between $X(\lambda)$ and $R_{rs}(\lambda)$ is examined in Fig. 3 for each IMECOCAL 1999 survey. Each solid line represents the data for one station, in which $X(\lambda)$ decreases, generally, with the wavelength. The data are compared using both linear and quadratic algorithms with $l_1 = 0.0949\ sr^{-1}$and $l_2 = 0.0794\ sr^{-1}$ in equation (3), as stated by [16]. Values of $X(\lambda) < 0.05$ pertain to wavelengths $\lambda \geq 510\ nm$. For wavelengths $\lambda \geq 510\ nm$ the absorption by pure water is the main component of the absorption coefficient [33]: 80% or higher of the total absorption. Given this and that the absorption by pure water is constant in time, the spatial and temporal variability of the absorption coefficient for $\lambda \geq 510\ nm$ is minimal. The seasonal variability becomes important for smaller wavelengths. In January, the GSM [16] algorithm underestimates $X(\lambda)$ while in October it overestimates $X(\lambda)$. The differences observed suggest that a better fit might be found with parameters adjusted to account for the seasonal variability within the IMECOCAL region. The parameters for the algorithm were adjusted to the region. However, when the algorithm was inverted, the spectral shape of the IOPs quotient was not adequately reproduced.

There is also variability between stations, that is, between each solid line (Fig. 3): larger variability in $X(\lambda)$ is observed along coastal stations with a smaller variability in oceanic stations. In general, it is expected that for wavelengths between 412 nm and 670 nm, $b_b(\lambda) \ll a(\lambda)$. Thus, variation in $X(\lambda)$ is mainly a function of the absorption coefficient. Given equation (2), the value of $X(\lambda)$ will be small when the value of $a(\lambda)$ is large. Therefore, in stations where Chl-a is high, $a(\lambda)$ will also be high and $X(\lambda)$ will be low. Chl-a increases towards the coast and, in consequence, $X(\lambda)$ decreases towards the coast. For wavelengths 555 and 670 nm, the absorption coefficient is dominated by the absorption of water and, hence, there is almost no variation in $X(\lambda)$.

There is a different mathematical relationship between $X(\lambda)$ and $R_{rs}(\lambda)$ for each survey. These relationships vary in magnitude and in shape. Thus, a power function was proposed for each 1999 survey

$$R_{rs}(\lambda) = e^{b}X^{m}(\lambda) , \tag{10}$$

where the parameters m and b were obtained using a log-log linear model

$$ln(R_{rs}(\lambda)) = mln\big(X(\lambda)\big) + b , \tag{11}$$

Many IOPs are modeled using a magnitude and a spectral shape *e.g.* [16] and [27]. For instance, in (5) the spectral shape is an exponential function that determines the way $b_{bp}(\lambda)$ varies with the wavelength. In our model, e^{b} is the magnitude of $R_{rs}(\lambda)$ while m is the shape and represents how $R_{rs}(\lambda)$ varies with $X(\lambda)$. The GSM [16] linear model corresponds to an m of 1 and b of –2.97 considering this power function. The advantage of using this new function is that it allows a non-integer power as well as a scaling factor or coefficient. In Table III the results for each survey are shown, as well as the associated errors with the fit. The value of b ranges from –2.9 to –0.8 from January to October, respectively. Given that e^{b} is the magnitude, its value is lowest for January and highest for October.

As seen in Fig. 3, the shape of the function varies seasonally. In our model, the power m varies from almost 1 in January to 1.5 in October (Table III). This implies that during January the relationship between $X(\lambda)$ and $R_{rs}(\lambda)$ is almost linear. However, during October, $R_{rs}(\lambda)$ is neither a linear nor a quadratic function of $X(\lambda)$. Also, $a_{d}(\lambda)$ is high in October due to the presence of Santa Ana winds [9], the value of $X(\lambda)$ for every wavelength which, in turn, alters the shape of the function.

TABLE III
PROPOSED IOP MODELS FOR EACH SURVEY OF THE 1999 IMECOCAL WHERE r^2 IS THE SQUARE OF PEARSON´S CORRELATION COEFFICIENT. **Δm** AND **Δb** ARE THE UNCERTAINTIES ASSOCIATED TO THE STANDARD DEVIATION OF THE DATA.

	Model Parameters				
Survey	**m**	**Δm**	**b**	**Δb**	r^2
January	1.03	0.03	-2.9	0.1	0.93
April	1.19	0.06	-2.3	0.2	0.87
August	1.23	0.04	-2.3	0.1	0.94
October	1.50	0.07	-0.8	0.3	0.87

Clear waters are usually waters of high $R_{rs}(\lambda)$ at short wavelengths, that is, blue waters. However, in Fig 2, we can see that high values of $R_{rs}(\lambda)$ do not correspond to waters with low values of absorption coefficients. The clearer waters in our surveys, with the lowest values of $a_{\phi}(\lambda)$ and $a_d(\lambda)$, are for January. During this season, the m parameter (Table III), that represents the shape of the function, has a value close to 1, denoting a linear relationship between $R_{rs}(\lambda)$ and the IOPs ratio, $X(\lambda)$. However, in the GSM [16] algorithm, in clear water with high values of $R_{rs}(\lambda)$ it is the quadratic term that expresses the shape of the function. We find that the higher values of m are for August and October (Table III), where both surveys had high values of $R_{rs}(\lambda)$.

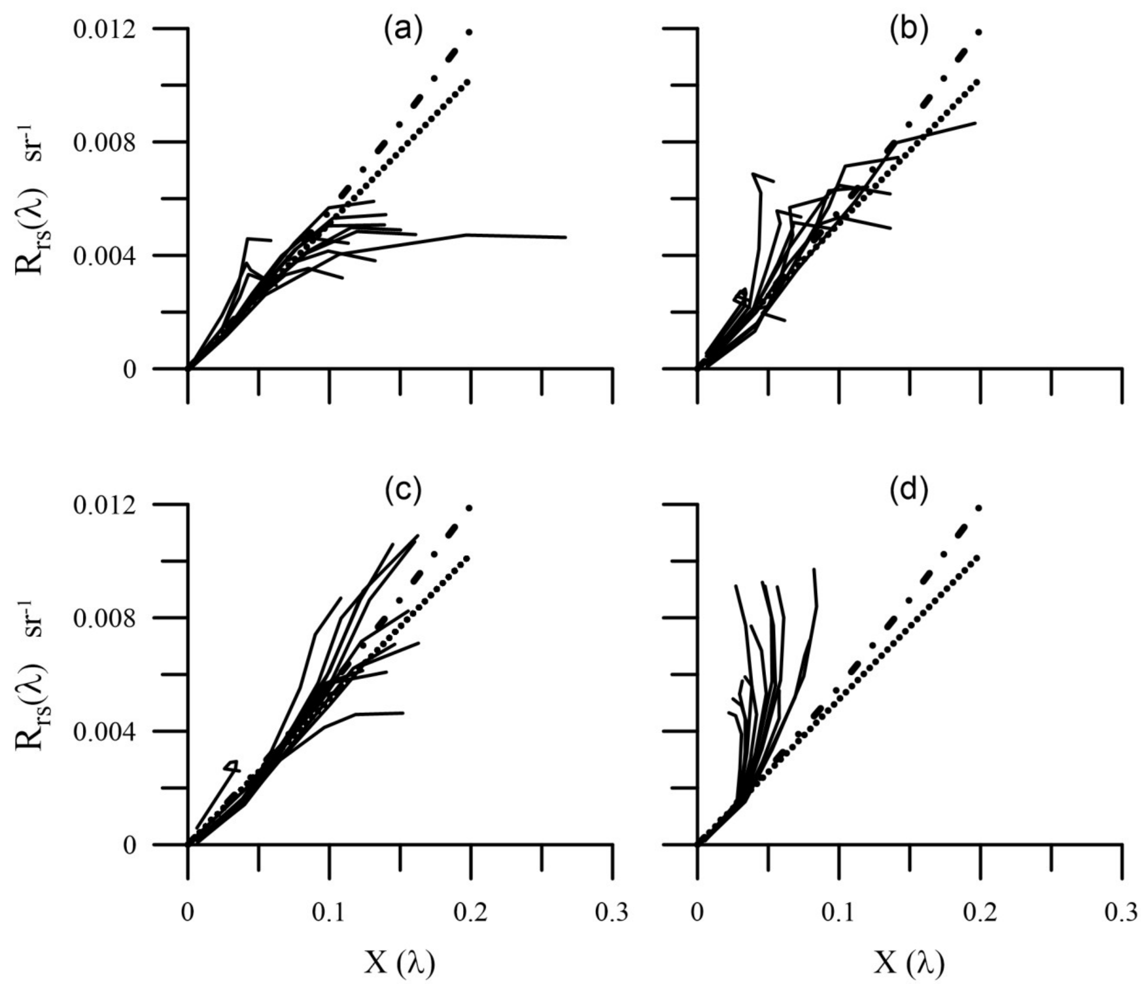

Fig. 3. Relationship between $X(\lambda)$, and remote sensing reflectance, $R_{rs}(\lambda)$, for each survey of IMECOCAL 1999: (a) January, (b) April, (c) August, and (d) October. The results are compared with those obtained using the GSM model [16]. In each graph, the solid line represents the data for one station; the dotted line and the dash-dotted line correspond to the linear and quadratic GSM model, respectively.

In order to understand the effect of the parameters m and b, an assessment of the sensitivity of the algorithm was made (Fig. 4). For this, $X(\lambda)$ values were calculated from $R_{rs}(\lambda)$ data varying the parameters m and b. First (Fig.4a), parameter m was varied and b was kept constant. Second, (Fig.4b) parameter b was varied while m was kept constant. In both instances, the parameter that was modified changed in approximately ±1 standard deviation considering the results shown in Table III. As expected, the highest fluctuations happened in the short wavelengths. When b was kept constant (Fig. 4a), a change of approximately 1 standard deviation in m, the spectral shape, produced a variation of $X(412\,nm)$ of 0.06. An increase in m resulted in an increase of $X(\lambda)$, that is associated with a decrease in the total absorption coefficient. When m was kept constant (Fig. 4b) the fluctuation on approximately 1 standard deviation in b, the magnitude, produced a change of ~0.03 to ~0.05 in $X(412\,nm)$. An increase in b resulted in a decrease in the magnitude of $X(\lambda)$, that is associated with an increase of the total absorption coefficient.

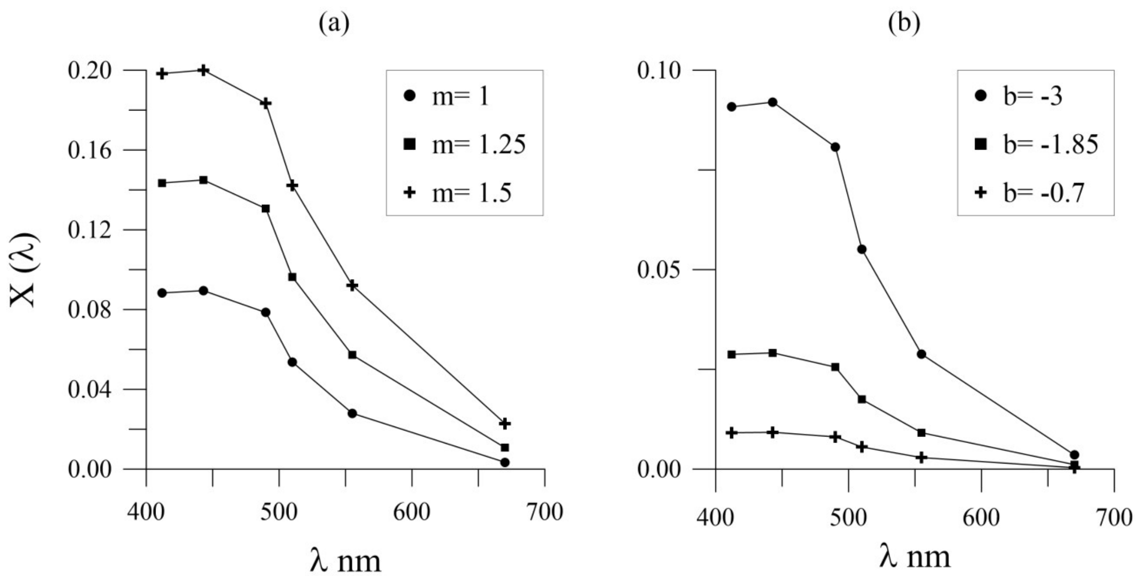

Fig. 4. Sensitivity of the proposed algorithm to variations in the two parameters m and b. In (a) we keep b=-2.97 and vary m while in (b) m=1 and we vary b.

C. IOPs inversion models

In order to validate the new algorithm, the standard deviation and the mean of the $X(\lambda)$ spectra were obtained for each survey. The mean plus and minus the standard deviation are plotted. This represents the normal variability observed in the data. With the satellite data, the new algorithm was inverted to obtain de spectra of $X(\lambda)$. The inversion, with its corresponding error (Table III) plotted as error bars, falls within the expected values of $X(\lambda)$ and follows the expected shape (Fig. 5). Therefore, the inversion adequately reproduces the shape of (λ) . In Fig. 5c, there is one result that falls out of the expected variability and it corresponds to a station of high Chla-a and $a_{\phi}(\lambda)$.

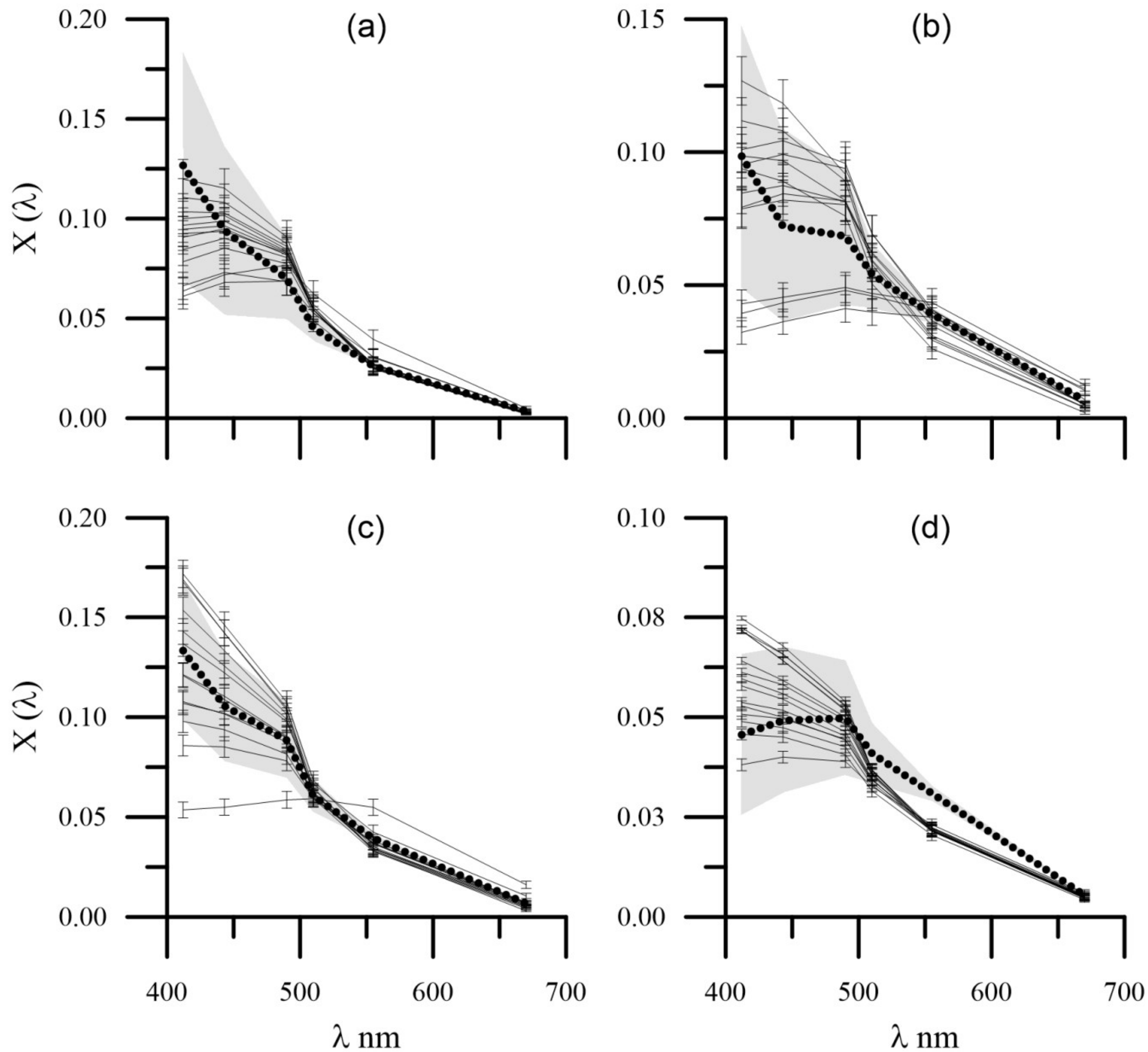

Fig. 5. Inversion of the new IOPs model and its error from (Table III) as compared to the mean (circles) $\pm$ standard deviation of the data (shaded) for (a) January, (b) April, (c) August and (d) October.

In Fig. 6 we compare the observed values of $X(\lambda)$, that is, the *in situ* data, with the results of the linear GSM algorithm as well as with the results of the power function proposed. As shown in Fig. 6, the GSM model overestimates the data and the square of Pearson's correlation coefficient is of 0.4599. This same coefficient is of 0.8175 for the proposed power function. The increase in this coefficient indicates an improvement in the retrieval of $X(\lambda)$ with the power function. Tiwari and Shanmugan [43] state that many IOPs inversion models are only applicable in clear oceanic waters and provide no information in the longer wavelengths because the absorption coefficient within this wavelength domain is controlled by pure seawater. Thus, the dispersion is mainly found for shorter wavelengths as expected.

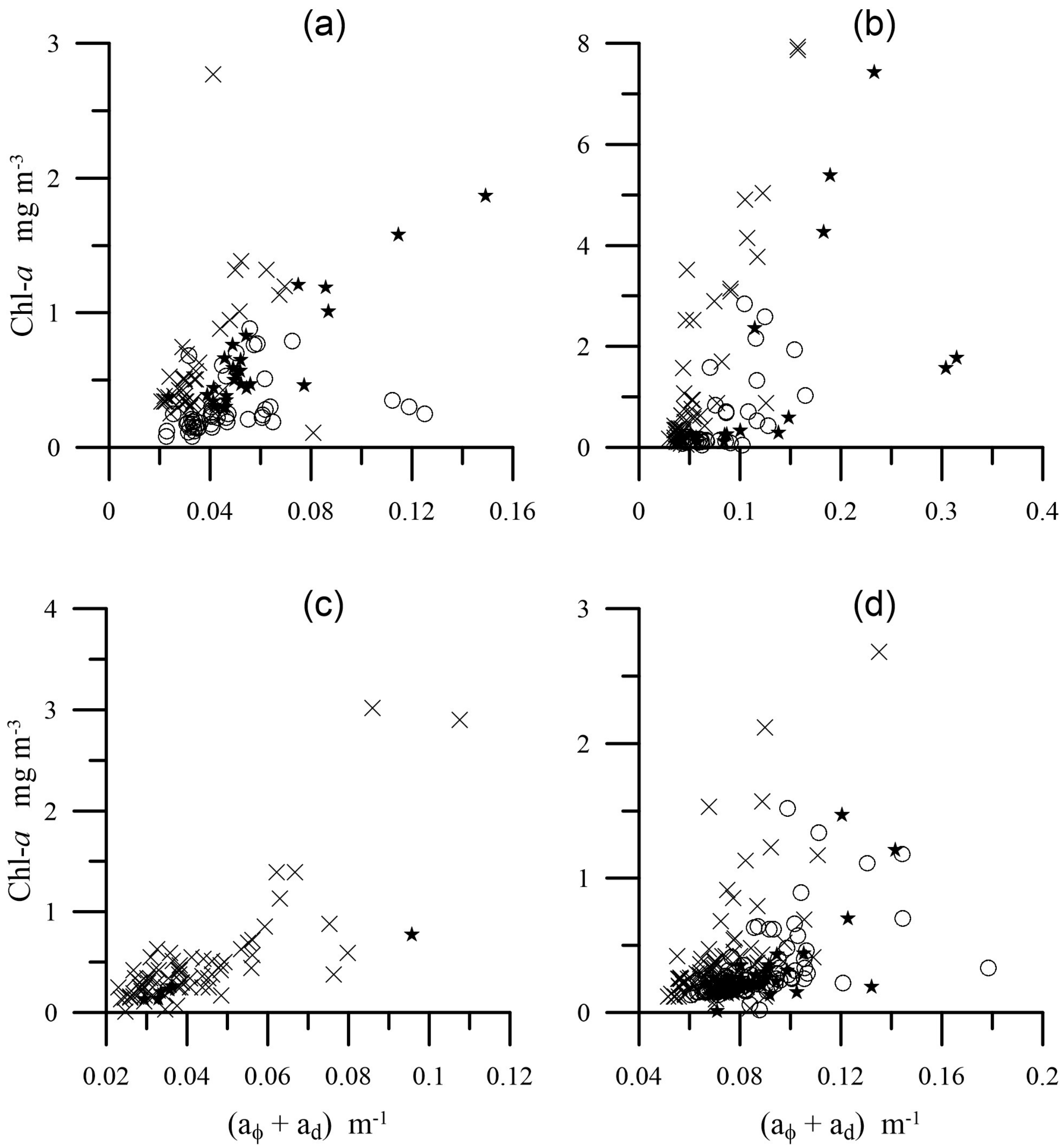

Fig. 6. Chlorophyll a concentration data versus the absorption coefficient of the particulate matter, $a_p(\lambda)$ at 443 nm for (a) January, (b) April, (c) August and (d) October. $a_p(\lambda)$ for 1999 (cross) are the *in situ* data whereas the 2008 (stars) and 2011 (circles) absorption coefficients come from the retrieval of $X(\lambda)$ using the algorithm proposed in this paper. For visibility purposes, there are different scales on the axis of the figures for each survey.

In the interest of validating the algorithm and considering that Chl-a is related to the IOPs, the IMECOCAL Chl-a database was used. This database was crosschecked with the Southern Oscillation Index (SOI) to ensure that the data used to validate the algorithm has the comparable climatic properties. During three years, 1999, 2008 and 2011 there is data for most of the months in the study (January, April, August and October) with comparable SOI values. For the 1999 surveys, the *in situ* data of the absorption coefficient of particulate matter, $a_p(\lambda)$, was plotted against *in situ* data of Chl-a (Fig. 6). $R_{rs}(\lambda)$ data was used to obtain $X(\lambda)$ for the 2008 and 2011 surveys using the algorithm proposed in this paper. The backscattering coefficient was modeled using the parameters stated Table II and the total absorption coefficient was calculated. Then, the absorption due to water was subtracted in order to derive the absorption coefficient of the particulate matter, $a_p(\lambda)$. The modeled $a_p(\lambda)$ for the 2008 and 2011 surveys was plotted against Chl-a too. The results from the inversion of our model produce the same the general pattern observed from the *in situ* data of. It is important to state that for 2011 there is no data in August. There are some locations where a high absorption coefficient does not correspond to high Chl-a. These points come from stations close to the coast where other processes can be at play. Generally, as $a_p(\lambda)$ increases at 443 nm, so does the Chl-a. The results of the absorption coefficient for 2008 and 2011 yield the expected results.

CONCLUSIONS

The inherent optical properties of seawater on the western coast of Baja California display variations of such magnitude that their appropriate calculation from remote sensing reflectances depends on the use of regionally and seasonally adapted algorithms or parameters. A known algorithm was tested and regionalized. However, the seasonal variability could not be reproduced for the region. The relationship between the remote sensing reflectance and the IOPs quotient, $X(\lambda)$, is close to linear only in January, a survey during which the variation of absorption and backscattering coefficients, remote sensing reflectance and $X(\lambda)$ are minimum. In the three other surveys, the relationship between the remote sensing reflectance and $X(\lambda)$ is not linear. Therefore, a new model that gives a better representation of the seasonal variation is proposed. This new algorithm consists of a power function with different parameters for each season. The spectra of the $X(\lambda)$ obtained by inverting the model adequately reproduces what is expected of this IOPs, with maximum dispersion for short wavelengths. Remote sensing reflectance and *in situ* Chl-a for January, April, August and October of 2008 and 2011 is used to validate the new algorithm. The model was inverted using the satellite data to obtain the absorption coefficient of particulate matter. When this modelled data is plotted versus the *in situ* Chl-a the expected pattern is reproduced. The new algorithm has two parameters for each season: a spectral shape, m, and a magnitude, b. The value of m has more influence in the resulting spectrum of $X(\lambda)$. The results presented can be used to generate time series of IOPs and their quotient $X(\lambda)$ that will give insight into their temporal and spatial variability within the southern border of the CCS.

References

[1] Preisendorfer, R. W., (1961). Applications of radiation transfer theory to light measurements in the sea. *Union of Geodetic Geophysical Institute Monograph*, vol. 10, pp. 11-30.

[2] Preisendorfer, R. W., (1976). Hydrologic Optics. *National Oceanic and Atmospheric Administration, U. S. Department of Commerce*, ed., 6 volumes.

[3] IOCCG, (2000). Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters. Sathyendranath, S. (ed.), *Reports of the International Ocean-Colour Coordinating Group*, No. 3, IOCCG, Dartmouth, Canada.

[4] Werdell, P. J. , L. I. W. McKinaa, E. Boss, S. G. Ackleson, S. E. Craig, W. W. Gregg, Z. Lee, S. Maritorena, C. S. Roesler, C. S. Rousseaux, D. Stramski, J. M. Sullivan, M. S. Twardowski, M. Tzortziou, X. Zhang, (2018). An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing. *Progress in Oceanography*, vol.160, pp.186-212. doi:10.1016/j.pocean.2018.01.001.

[5] Ogashawara, I., (2015). Terminology and classification of bio-optical algorithms, *Remote Sens. Lett.*, vol.6, no.8, pp.613-617. doi: 10.1080/2150704X.2015.1066523

[6] Morel, A., and L. Prieur, (1977). Analysis of variations in ocean color. *Limnol. Oceanogr.*, vol. 22, no. 4, pp. 709-722. doi:10.4319/lo.1977.22.4.0709

[7] Carder, K. L., R. G. Steward, J. H. Paul, and G. A. Vargo, (1986). Relationships between chlorophyll and ocean color constituents as they affect remote-sensing reflectance models. *Limnol. Oceanogr.*, vol. 31, no. 2, pp. 403-413. doi: 10.4319/lo.1986.31.2.0403.

[8] IOCCG, (2006). Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications. Lee, Z. P. (ed.), *Reports of the International Ocean-Colour Coordinating Group*, No. 5, IOCCG, Dartmouth, Canada.

[9] Sosa-Ávalos, R., R. Durazo, B. G. Mitchell, J. Cepeda-Morales, and G. Gaxiola-Castro, (2017). Phytoplankton photosynthetic parameters off Baja California: A tool to estimate primary production with remote sensing data. *Ciencias Marinas*, vol. 43, no. 3, pp. 157–172. doi: 10.7773/cm.v43i3.2746.

[10] Sathyendranath, S., A. Longhurst, C. M. Caverhill, and T. Platt, (1995). Regionally and seasonally differenciated primary production in the North Atlantic. *Deep-Sea Res. Part I.: Oceanographic Research Papers*, vol. 42, no. 10, pp. 1773-1802. doi: 10.1016/0967-0637(95)00059-F.

[11] Martin, S., (2004). An introduction to ocean remote sensing. *Cambridge University Press*.

[12] Bricaud, A., M. Babin, A. Morel, and H. Claustre. (1995). Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parametrization. *J. Geophys. Res.*, vol. 100, no. C7, pp. 13321-13332, doi: 10.1029/95JC00463.

[13] Lynn, R. J. and J. J. Simpson, (1987). The California Current system: The seasonal variability of its physical characteristics, *J. Geophys. Res. Oceans,* vol. 92, no. C12, pp. 12947-12966.

[14] Kurian, J., F. Colas, X. Capet, J. C. McWilliams, and D. B. Chelton, (2011). Eddy properties in the California Current System, *J. Geophys. Res.*, vo. 116, no. C08027. doi:10.1029/2010JC006895.

[15] Durazo, R., (2015). Seasonality of the transitional region of the California Current System off Baja California, *J.Geophys. Res. Oceans*, vol. 120, pp. 1173–1196. doi:10.1002/2014JC010405.

[16] Garver, S. A. and D. A. Siegel, (1997). Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation 1. Time series from the Sargasso Sea, *J. Geophys.Res.*, vol. 102, no. C8, pp. 18607-18625. http://dx.doi.org/10.1029/96JC03243.

[17] Maritorena, S., D. A. Siegel, and A. R. Peterson, (2002). Optimization of a semianalytical ocean color model for global-scale applications, *Appl. Opt.*, vol. 41, no. 15, pp. 2705-2714. doi: 10.1364/AO.41.002705.

[18] Gaxiola-Castro, G., R. Durazo, B. Lavaniegos, M. E. De La Cruz-Orozco, E. Millán-Núñez, L. Soto-Mardones, and J. Cepeda-Morales, (2008). Pelagic ecosystem response to interanual variability off Baja California, *Ciencias Marinas*, vol. 34, no. 2, pp. 263–270.ISSN: 0185-3880.

[19] Durazo, R., (2009). Climate and upper ocean variability off Baja California, Mexico: 1997–2008, *Progress in Oceanography*, vol. 83, no. 1–4, pp. 361-368. doi: 10.1016/j.pocean.2009.07.043.

[20] Gordon, H. R., O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker, and D. K. Clarck, (1988). A Semianalytic Radiance Model of Ocean Color, *J. Geophys. Res.*, vol. 93, no. D9, pp. 10909-10924. doi: 10.1029/JD093iD09p10909.

[21] Lee, Z. P. , K.L. Carder and K.P. Du. (2004). Effects of molecular and particle scatterings on model parameters for remote-sensing reflectance. *Appl. Opt.,* vol. 43, pp. 4957-4964.

[22] Carder, K. L., S. K. Hawes, K. A. Baker, R. C. Smith, R. G. Steward, and B. G. Mitchell, (1991). Reflectance model for quantifying chlorophyll a in the presence of productivity degradation products, *J. Geophys. Res.*, vol. 96, No. C11, pp. 20599-20611. doi: 10.1029/91JC02117.

[23] Siegel, D. A. and A. F. Michaels, (1996). Quantification of non-algal light attenuation in the Sargasso Sea: Implications for biogeochemistry and remote sensing, *Deep-Sea Res. Part II: Topical Studies in Oceanography*, vol. 43, no. 2–3, pp. 321-345. doi: 10.1016/0967-0645(96)00088-4.

[24] N. B. Nelson, D. A. Siegel, and A. F. Michaels, Seasonal dynamics of colored dissolved material in the Sargasso Sea, *Deep-Sea Res. Part I.: Oceanographic Research Papers*, vol. 45, No.6, pp. 931-957, Jan. 1998, doi: 10.1016/S0967-0637(97)00106-4.

[25] Nelson, N. B. and D. A. Siegel, (2002). Chromophoric DOM in the open ocean. In: Hansell, D.A., Carlson, C.A. (Eds.), *Biogeochemistry of Marine Dissolved Organic Matter*. Academic Press, San Diego, CA, pp. 547–578. ISSN: 0967-0637.

[26] Roesler, C. S. and M. J. Perry, (1989). Modeling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters, *Limnol.Oceanogr.*, vol. 34, no. 8, pp. 1510-l523. doi: 10.4319/lo.1989.34.8.1510.

[27] Carder, K. L., F. R. Chen, Z. P. Lee, S. K. Hawes, and D. Kamykowski, (1999). Semianalytic Moderate-Resolution Imaging Spectrometer Algorithms for Chlorophyll *a* and Absorption with Bio-Optical Domains Based on Nitrate-Depletion Temperatures, *J. Geophys.Res.*, vol. 104, no. C3, pp. 5403-5421. doi: 10.1029/1998JC900082.

[28] Morel, A., (1988). Optical modeling of the upper ocean in relation to its biogenous matter content (Case I waters), *J. Geophys. Res.*, vol.93, no. C9, pp. 10749-10768. doi: 10.1029/JC093iC09p10749.

[29] Roesler, C. S. and M. J. Perry, (1995). *In situ* phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance, *J. Geophys. Res.*, vol. 100, no. C7, pp. 13279-13294. doi: 10.1029/95JC00455.

[30] Kishino, M., N. Takahashi, N. Okami and S.Ichimura, (1985). Estimation of the spectral absorption coefficients of phytoplankton in the sea, *Bulletin of Marine Science*, vol. 37, pp. 634-642. ISSN 0007-4977.

[31] Mitchell, B. G., (1990). Algorithms for determining the absorption coefficient of aquatic particulates using the quantitative filter technique (QFT), *Ocean Optics*, vol. X, pp. 137-148. doi: 10.1117/12.21440.

[32] Mitchell, B. G., M. Kahru, J. Wieland, M. Stramska, (2003). Determination of spectral absorption coefficients of particles, dissolved material and phytoplankton from discrete samples, in *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation*. Mueller, J. L., Fargion, G. S., McClain, C.R. (eds.). NASA Technical Memorandum, Revision 4, vol. 4, pp. 39-64.

[33] Smith, R. C. and K. S. Baker, (1981). Optical properties of the clearest natural waters (200-800 nm), *Appl. Opt.*, vol. 20, no. 2, pp. 177-184. doi: 10.1364/AO.20.000177.

[34] Pope, R. M. and E. S. Fry, (1997). Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements, Appl. Opt., vol. 36, no. 33, pp. 8710-8723. doi: 10.1364/AO.36.008710.

[35] Gordon, H. R. and A. Y. Morel, (1983). Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery: A Review, 114 pp., Springer-Verlag, New York. doi: 10.1007/978-1-4684-6280-7.

[36] Sosa-Ávalos, R., G. Gaxiola-Castro, B. G. Mitchell and J. Cepeda-Morales, (2010). Coeficiente específico de absorción de la luz por fitoplancton, in G. Gaxiola-Castro and R. Durazo (eds) *Diez años de investigaciones mexicanas de la Corriente de California*, INE/CICESE/UABC/SEMARNAT. 501 pp.

[37] Durazo, R. and T. Baumgartner. (2002). Evolution of oceanographic conditions off Baja California: 1997–1999, *Progr. Oceanogr.*, vol. 54, pp. 7–31. doi: 10.1016/S0079-6611(02)00041-1

[38] Sosa-Ávalos, R., G. Gaxiola-Castro, R. Durazo and B. G. Mitchell, (2005). Effect of Santa Ana winds on bio-optical properties off Baja California, *Ciencias Marinas*, vol. 31, no. 2, pp.339-348. ISSN: 0185-3880.

[39] Venrick, E. L. and T. L. Hayward, (1984). Determining chlorophyll on the 1984 CalCOFI surveys *CalCOFI Rep.* 25 pp. 74-79.

[40] Holm Hansen, O., C. J. Lorenzen, R. W. Holm, J. D. H. Strickland, (1965). Fluorometric determination of chlorophyll, *J. Cons Int. Explor. Mer.*, vol. 30. Pp 3-15.

[41] Yentsch, C. S., D. W. Menzel, (1963). A method for the determination of phytoplankton, chlorophyll and phaeophytin by fluorescence, *Deep-Sea Res.*, vol. 10, pp 221-231.

[42] De la Cruz-Orozco, M. E., E. Gómez-Ocampo, L. E. Miranda-Bojórquez, J. Cepeda-Morales, R. Durazo, B. E. Lavaniegos, T. L. Espinosa-Carreón, R. Sosa-Avalos, E. Aguirre-Hernández, G. Gaxiola-Castro, (2017). Phytoplankton biomass and production off the Baja California Peninsula: 1997 – 2016. *Ciencias Marinas*, vol. 43, no. 4, pp 217-228.

[43] Tiwari, S. P. and P. Shanmugan, (2012). An Optical Model for the Remote-Sensing of Absorption Coefficients of Phytoplankton in Oceanic/Coastal Waters, *Advances in Remote Sensing*, vol. 1, pp. 19-34. doi: 10.4236/ars.2012.12003.

CHAPTER II. EVALUATION OF OCEAN COLOR ALGORITHMS TO RETRIEVE CHLOROPHYLL-A CONCENTRATION IN THE MEXICAN PACIFIC OCEAN OFF THE BAJA CALIFORNIA PENINSULA, MÉXICO

Abstract— Ocean color data is used to infer the concentration of Chlorophyll–*a* (Chl-*a*), a proxy of phytoplankton biomass. Mathematical algorithms relate satellite data of ocean color with the surface Chl-*a* concentration. These mathematical tools work best when they are adapted to the unique bio-optical properties of a particular oceanic province. Ocean color algorithms should also consider that there are significant differences between datasets derived from different sensors. Common solutions are to provide different parameters for each sensor or use merged satellite data. In this paper we use satellite data from the Copernicus merged product suite and *in situ* data from the southernmost part of the California Current System to test two widely used global algorithms, OCx and CI, and a regional algorithm, CalCOFI2. The OCx algorithm gave the best result, therefore, it was regionalized and, again, tested. The database was, then, separated according to (a) dynamic boundaries in the area; (b) bio-optical properties and (c) climatic condition (El Niño/La Niña). Regional algorithm were obtained and tested for each partition. The Chl-*a* retrievals for each model were tested and compared. The best result was for the climatic condition separation allowing for the construction of consistent regionally adapted time-series and demonstrating the importance of El Niño/La Niña events on the bio-optical properties of the area.

Index Terms— Remote sensing, algorithms, color, Chlorophyll-*a*.

INTRODUCTION

The state of the oceanic ecosystems and their management are determined through the monitoring of oceanographic variables. Long time series and adequate spatial resolution are required in order to understand these complex ecosystems. While ship-based monitoring provides the best results, the frequency and spatial coverage of it are poor whilst the cost for obtaining them is high. Therefore, satellite derived data gives a better temporal coverage, albeit limited to surface variables. While the spatial resolution of the remote data is coarser, it covers most of the ocean surface. Each satellite sensor has different optical bands, sensitivities, noise patterns and overpass times [1]. Satellite data, as every other measurement, has uncertainties associated with instrument calibration and the processing algorithms [2]. There are significant differences between the data acquired from different sensors [3], [4], therefore, the parameters of the algorithms applied to the data are particular to each sensor [5]. For example, for data in the California Current, ocean color algorithms were proposed for a suite of satellite sensors (Ocean Color and Temperature Scanner , OCTS, Sea-viewing Wide Field-of-view Sensor, SeaWiFS, Medium Resolution Imaging Spectrometer, MERIS, Moderate Resolution Imaging Spectroradiometer, MODIS, and Visible Infrared Imaging Radiometer Suite, VIIRS) by [3], [6].

By 2018 a 21-year record of ocean-color has been comprised. However, while this dataset has contributed in our understanding of biogeochemical processes [4] it is made of data from individual sensors each with different characteristics that have operated at different times [8]. New parameters are required for each sensor. Therefore, in order to improve the coverage of satellite data and provide consistent time series, merging of data from different sensors is needed [3], [7], [8] which, in turn, requires updating the parameters of the algorithms. For example, [4] derived 65 OC algorithms for 25 satellite ocean color sensors. Operational satellite sensor constellations, such as Copernicus, provide ocean color data as elements of an integrated, sustained observing system [8] that has better global coverage and will allow the study of phenomena of larger timescale than what an individual sensor provides [4]. [9], among others, have suggested decadal-scale fluctuations in phytoplankton community, however, the length of the dataset cannot yet resolve for these long-term trends. Therefore, it is of dire importance to obtain a consistent long-term dataset that allows this type of analysis.

Bio-optical algorithms have been developed to estimate surface Chlorophyll-*a* concentration (Chl-*a*) from satellite data. These algorithms are usually empirical, deriving their parameters from the data [10]. The Level-3 Chl-*a* images available for downloading in the NASA Ocean Color Web Site employ an OCx algorithm with different parameters for different sensors, [5] merged with the color index algorithm (CI) [11]. However, these global algorithms are not regionally optimal [6]. For example, they underestimate the values of Chl-*a* in the California Current area [6], [12] and overestimate them in the Northern Bering Sea and Chukchi Sea [13], [14]. Therefore, it is important to have an evaluation of the performance of the global algorithms at a regional level, even more so, considering merged satellite data. This will help in building sturdy, long-term data series and providing better information for the study and management of ocean ecosystems.

The eastern limb of the North Pacific gyre is the California Current System (CCS) [15]. As an eastern boundary current, it presents coastal upwelling, eddies and fronts. The mixture and the upwelling of nutrients determine the characteristics of the ecosystem [16], [17]. [17] and [19] describe the southernmost part of the CCS, in front of the Baja California Peninsula, as a transitional zone where cold water from the subarctic meets warmer water coming from the tropics and subtropics. The location of this transitional zone varies seasonally [17] and with this, the state of the oceanic ecosystem. Therefore, circulation and water masses convergence and interaction modulate the biological and chemical processes [20], [21]. Previous studies

([17] and [21]) have shown that in this area there are two distinct provinces separated at around 28° N, at Punta Eugenia (Fig. 1). North of Punta Eugenia there are subarctic waters throughout most of the year, with year-round upwelling most intense during spring and early summer [19]. In the southern region coastal upwelling occurs mainly in spring and summer, while the tropical and subtropical influence is limited to summer and fall. Also, there are two cyclonic gyres: north and south of Punta Eugenia [17]. In the vicinity of 32° N, water from the California Current (CC) converges with an intrusion of water from the Central North Pacific [18] thus, the CC turns toward the shore and part of it turns north to incorporate into the Southern California Eddy and the Southern California Countercurrent while the remainder turns southward along the Baja California coast [22]. Therefore, this convergence forms a persistent feature called the Ensenada Front, an ecological transition zone [18] marked by a strong gradient in Chl-*a* concentration. In the Ensenada Front the eutrophic waters of the CC end abruptly possibly due to the subduction of this pigment and nutrient rich waters to the southwest [22].

The California Current has been studied for many years. In 1949 the California Cooperative Oceanic Fisheries Investigations (CalCOFI) began making surveys in front of the California coast in order to investigate the ecological aspects of the sardine population collapse. The objective of the CalCOFI program has now become the study of the marine environment off the coast of California as well as its management through analysis of diverse oceanographical variables, which include chlorophyll-*a* concentration. Apart from NASA's global algorithms, several algorithms have been developed for the CC area [10], [23] and [24].

Since 1997, off Baja California, the variability of the ecosystem in this transitional area has been studied through the IMECOCAL (Spanish acronym for Mexican Research Program of the California Current). IMECOCAL activities include *in situ* sampling in open ocean waters in front of the Baja California Peninsula. Therefore, there is a database of 22 years of *in situ* data that, coupled with satellite information, can be used to evaluate and regionalize ocean color algorithms to retrieve Chl-*a* concentrations.

This complex transitional ecosystem is also affected by interannual variations, such as El Niño – Southern Oscillation (ENSO). A warm El Niño event is usually associated with weaker alongshore winds, as well as with the poleward transport of warmer, nutrient depleted water. A cold La Niña event brings stronger winds and colder, nutrient rich water. In general, a decrease in Chl-*a* is expected during a warm event. In waters off the Baja California Peninsula coastal upwelling determines the variability of the phytoplankton community [25]. Therefore, when there is a disruption in these pattern, *i.e.* during El Niño/La Niña events, biomass and community composition experience a shift. During a warm event there is a decrease in diatom abundance and an increase in the small cell communities whereas cold events favor diatom growth [26], [27]. [28] analyzed the extent of eutrophic and mesotrophic areas of the California Current during ENSO events and observed important differences between the waters off Baja California to those off California. [28] noted a reduction of the eutrophic and mesotrophic areas off the Southern California Bight associated to a decrease in upwelling; also, there was an increase in the extent of the mesotrophic areas off Baja California extending up to 700 km offshore, thus, not likely due to upwelling but, possibly to cyanobacteria. During the El Niño event of 1997-1998 [29] noted that in the region South of Punta Eugenia, tropical and subtropical waters dominated the region, while North of Punta Eugenia there was a coastal poleward flow that displaced the core of the California Current offshore. Much more work is needed in order to understand the processes within this ecosystem. Algorithms need to be tested for the area in order to evaluate their performance and adapt them to its distinct characteristics. Using regionally adapted algorithms to merged satellite data will allow the construction of long time series that will, in turn, shed light on these processes. When the local processes are factored into the algorithm parameters, the results yielded by the algorithm can provide more detailed ecological information mainly for the more eutrophic and oligotrophic waters.

In this work, the performance of two global/operational algorithms (OCx, CI) and an algorithm proposed for the CCS [10] is evaluated using *in situ* measurements of Chl-*a* off the Baja California Peninsula and Copernicus merged satellite data. The performance of Chl-*a* algorithms in this area for this satellite data has not yet been tested. This is important because Copernicus is a satellite conste llation that provides increased coverage and time series length compared to single-sensor data streams due to the fact that it is a merged sensor data record. [9] have suggested decadal-scale fluctuations linked to climate forcing, but the length of this record is insufficient to resolve longer-term trends.

After testing the algorithms a regionalization for the area and the satellite dataset is proposed and tested. Seeking further improvement, the dataset was separated and analyzed according to: (a) hydrographic properties that determine dynamic boundaries in the area, (b) bio-optical properties (blue/green reflectance ratio), and (c) climatic events (ENSO). Algorithms were proposed and tested for each classification. The results gathered permit the selection of the model that best represent the area using this particular satellite dataset all the while explaining some of the unique bio-optical properties of the region. A well assured retrieval of Chl-*a* from a consistent cross-instrumental dataset will allow to further the insights into biogeochemical processes as well as decadal and climate change.

MATERIALS AND METHODS

The *in situ* Chl-*a* belong to the database collected from the IMECOCAL region (Fig. 1) ranging from January 1998 to April 2016 (Table 1). For Chl-*a*, water was collected from the network of stations (Fig. 1) in 5 L Niskin bottles coupled in a rosette. Silicone bands and O-rings were used for closure. The water samples were collected from 1 m, 10 m, 20 m, 50 m, 100 m, 150 m and 200 m of depth. Two liters of water at each depth were filtered through Whatman GF/F Filters at a pressure below 150 mm Hg. The filters were then placed in Histoprep capsules and frozen in liquid nitrogen until their analysis. Chlorophyll was extracted with acetone at 90% and at 4°C in darkness for 24 hours as recommended by [30]. Chl-*a* concentration (mg m^{-3}) was determined by the fluorometric method [31], [32], with a Turner Designs A10 fluorimeter for the 1998 to 2005 samples and a Turner Designs trilogy fluorimeter for 2005 to 2016. Both fluorimeters were calibrated with pure Chl-*a* (Sigma). In this paper, the mean Chl-*a* between the surface and 10 m depth data was used. More information on the chlorophyll data can be seen in [33].

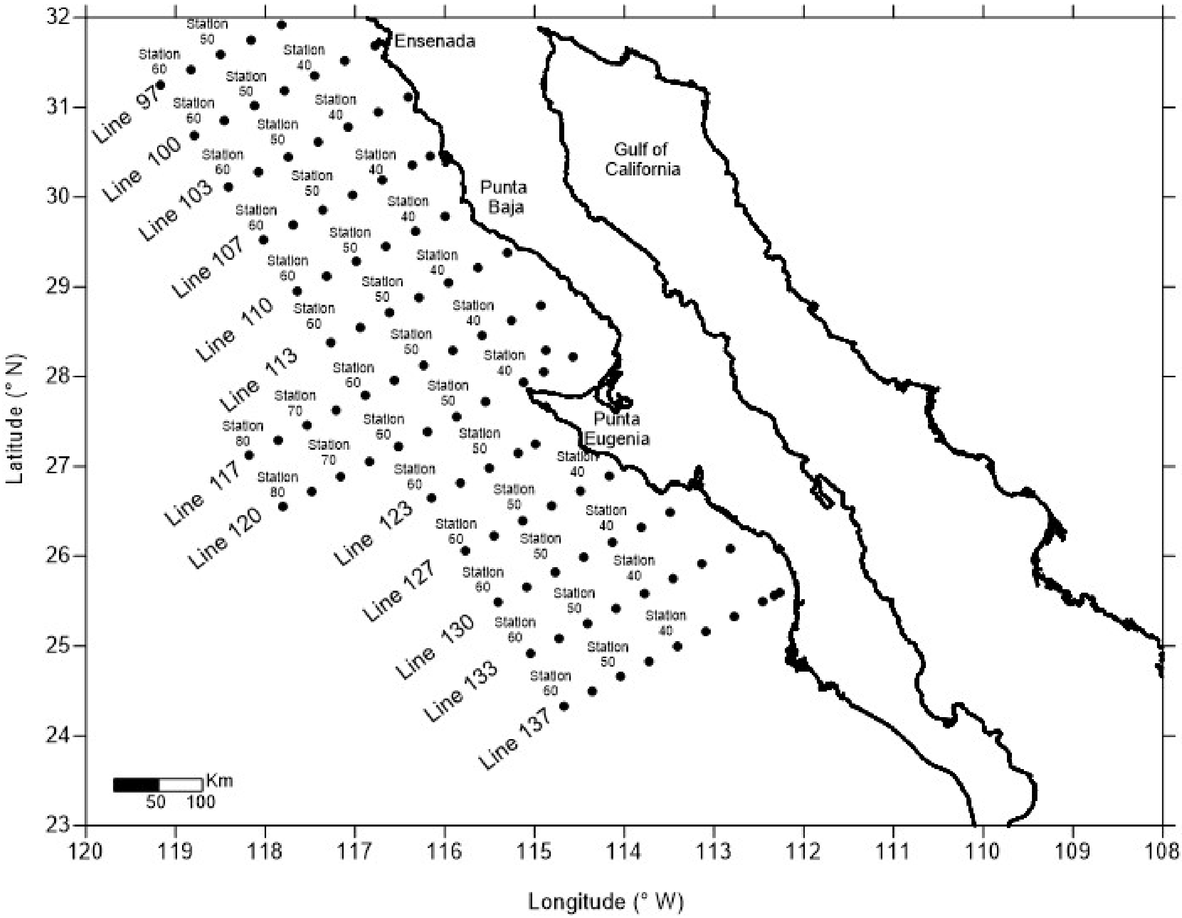

Fig 1. General grid of IMECOCAL stations where *in situ* samples were collected. The line number increases from 97 toward the South on line 137. The station number increases with distance from the coast.

TABLE I

DATES OF IMECOCAL CAMPAIGN WHERE Chl-a DATA WAS OBTAINED.

Year	January	February	March	April	May	June	July	August	September	October	November	December
1998	■	■					■		■	■		
1999	■		■	■				■		■		
2000	■	■		■			■			■		
2001	■	■		■		■	■			■		
2002	■	■		■	■		■	■		■	■	
2003	■	■		■			■			■		
2004	■	■		■	■		■			■		
2005	■	■		■	■		■	■	■			
2006		■		■	■		■					
2007	■	■		■	■			■	■			
2008	■	■		■			■	■		■		
2009				■						■	■	
2010			■	■								
2011	■	■		■	■		■			■		
2012	■	■	■									
2013		■			■	■						
2014							■	■	■			
2015									■			
2016	■			■								

Regarding satellite data, level 3 monthly averages between January of 1998 and May of 2016 of remote sensing reflectance $Rrs(\lambda)$ for 412 nm, 443 nm, 490 nm, 555 nm and 670 nm merged from different sensors were obtained from Copernicus (https://www.copernicus.eu/). The spatial resolution of each composite image is 4 km.

There is a relationship between the ocean color data measured with the satellites and the quantity and quality of the various constituents of the particular water body. Bio-optical algorithms estimate the near-surface Chl-a concentration from satellite data. These algorithms, developed since the 1970´s, use mostly empirical equations obtained by statistical regression of *in situ* Chl-a data and remotely-sensed variables [10]. Among the many algorithms, the most commonly used by the NASA Ocean Color web site are the O´Reilly band ratio (OCx) and the Hu Color index algorithm (CI). The OCx algorithm is a polynomial relationship between the $\log_{10}$ of a Rrs ratio and the $\log_{10}$ of Chl-a, whereas the CI algorithm is a three-band reflectance difference model. The NASA Chl-a estimation available for download is calculated by both algorithms. CI is used for low Chl-a retrievals, OCx for high Chl-a retrievals and a weighted approach for the values in between. In this paper, we use the OCx [5], [34] and CI [11] as the global algorithms. Furthermore, regional algorithms have been applied in order to obtain a better estimate given the particular properties of an area. [23] proposed various algorithms for the CalCOFI region. One of these, the CalCOFI 2 algorithm is used in this paper as a regional algorithm. All of these are described subsequently.

The OCx algorithm is a fourth order polynomial relationship between F, the logarithm of the blue/green ratio of remotely-sensed reflectances, R

$$F = log_{10}[R] , \tag{1}$$

where $= \frac{max(Rrs(443,490))}{Rrs(555)}$, and the logarithm of the Chl-*a* concentration, as follows

$$log_{10}(\text{Chl-}a)= a_0 + \sum_{i=1}^{4} a_i F^i, \tag{2}$$

where the coefficients of this function are sensor-specific and were obtained using *in situ* global data by [5] and [34]. In this work, we use the results obtained by using the SeaWiFs parameters $(a_0 = 0.3272,\ a_1 = -2.9940,\ a_2 = 2.7218, a_3 = -1.2259, a_4 = -0.5683)$ to obtain the Chl-*a* concentration with the OCx algorithm. These parameters were selected after testing with the algorithm parameters for SeaWiFs, MERIS, MODIS and OCTS [34].

The CI algorithm [11] is a three-band model that defines an index using reflectance on the blue, green and red bands (443, 555 and 670 nm, respectively)

$$CI = Rrs(555) - \left[Rrs(443) + \frac{555-443}{670-443}\left(Rrs(670) - Rrs(443)\right)\right], \tag{3}$$

which is related to the Chl-*a* concentration through the equation

$$log_{10}(\text{Chl-}a)= -0.409 + 191.6590\ CI \tag{4}$$

The CI algorithm is used when retrievals are below 0.15 mg m^{-3}, while the OCx algorithm is used for retrievals higher than 0.2 mg m^{-3}. A weighted approach of both is used for the values in between [34].

Alongside the OCx and CI global algorithms, there are working algorithms for the California Current System. Proposed specifically for the CCS, CalCOFI 2 band linear algorithm [23] uses a blue/green ratio of reflectance, R, (Eq. 1) in an exponential function

$$\text{Chl-}a = 10^{0.444R-2.431} \tag{5}$$

The regionalization of the algorithms was made using a least squares regression [36]. Algorithm evaluation was done using the coefficient of determination (R^2), the adjusted coefficient of determination (R_a^2), the sum of square error (SSE):

$$SSE = \sum_{i=1}^{N} (\log_{10}(\text{Chl-}a_{\text{ modeled}})_i - \log_{10}(\text{Chl-}a_{\text{ observed}})_i)^2 \tag{6}$$

the root mean squared logarithmic error (RMSLE):

$$RMSLE = \sqrt{\frac{1}{N} SSE} \tag{7}$$

and the Akaike information criterion (AIC):

$$AIC = -2ln(L) + 2k \tag{8}$$

where L is the log-likelihood, that is, it is a measure of how likely it is to obtain the observed data from a model, and k is the number of parameters in the model [35].

RESULTS

Performance of OCx, CI and CalCOFI 2 algorithms

The two standard global algorithms OCx (Eq. 2) and CI (Eq. 4) and one adapted for the CCS (CalCOFI 2, Eq. 5) were evaluated for waters from the IMECOCAL region (Fig. 1). These algorithms were applied to all the reflectance data and the logarithm of the Chl-*a* concentration was obtained (Fig. 2). Two of the three algorithms tested use F (Eq. 2), therefore, the plot is shown in terms of F (Fig. 2). The CI algorithm produces more dispersion in the Chl-*a* estimates than the OCx algorithm and CalCOFI 2 because the CI index depends on reflectance on the blue, green and red bands (Eq. 3).

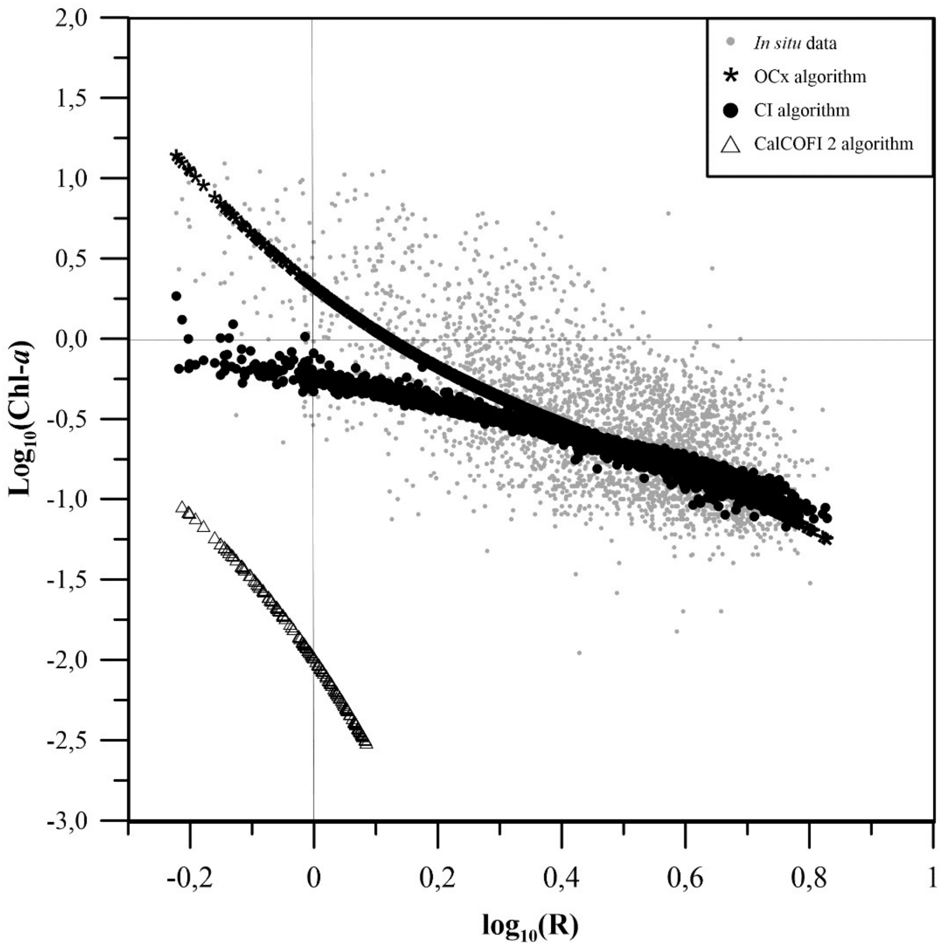

Fig. 2. Results from applying CalCOFI2 (triangle), CI (circle) and OCx (asterisk) algorithms to remote sensing reflectance data ($F=\log_{10}(R)$, Eq. 2) off Baja California compared to the $\log_{10}$ of *in situ* Chl-*a* data (grey circles).

As the water becomes greener, F diminishes its value and yields greater Chl-a concentrations. When the algorithms were applied to the IMECOCAL data (Fig. 2), the main divergence in the algorithms corresponds to values of Chl-a higher than or equal to 0.3 mg m^{-3} (log_{10}(Chl-a) ~ -0.5), and a ratio of blue/green reflectance of ≤ 2.5 ($F \leq 0.4$). For these values of the reflectance ratio OCx produces higher values of Chl-a than CI. The difference between CI and OCx retrievals for waters with large R values is small, yielding very similar results for the area though slightly higher retrievals for the CI algorithm. Regarding the CalCOFI 2 algorithm, it underestimates the Chl-a concentration for the entire domain of the data. The underestimation increases with the value of R, the blue/green reflectance ratio. For the IMECOCAL area, OCx is the algorithm that yields the best results from a qualitative point of view. The errors will be presented at the end of the Results section. Based on the algorithm results, OCx is used as the basic algorithm for further comparisons, discarding the CI and CalCOFI 2 algorithms.

Regional and OCx algorithms

As seen in Fig. 2, the OCx algorithm is a good fit for the area, although a better fit is expected when adapted to the specific characteristics of the area. Therefore, using all the *in situ* data a polynomial fit was made between F and the *in situ* Chl-a IMECOCAL data. After testing polynomials of orders one to five, the fourth order one was chosen based on the results of errors (R^2, R_a^2, SSE, RMSLE and AIC). The fourth order regional algorithm that best fit the data is

$$log(\text{Chl-}a) = 0.1746 - 1.9952\text{F} + 1.9992F^2 - 4.1958F^3 + 3.3837F^4 \qquad (9)$$

The comparison between *in situ* Chl-a data, OCx (global) and the regional algorithm (Eq. 9) is shown in Fig. 3. The regional algorithm yields values of Chl-a concentration that are similar to the global algorithm, mainly within the $1.5<R<4.5$ ($0.18<F<0.65$) range of blue/green reflectance ratio. For values of $R>4.5$ ($F>0.65$), the regional algorithm provides larger estimates of Chl-a, about 1.8 mg m^{-3} higher than the global algorithm which is even larger than the CI retrievals (Fig. 2). The regional algorithm gives a lower estimate of Chl-a for ratios $R < 1.5$ ($F<0.18$).

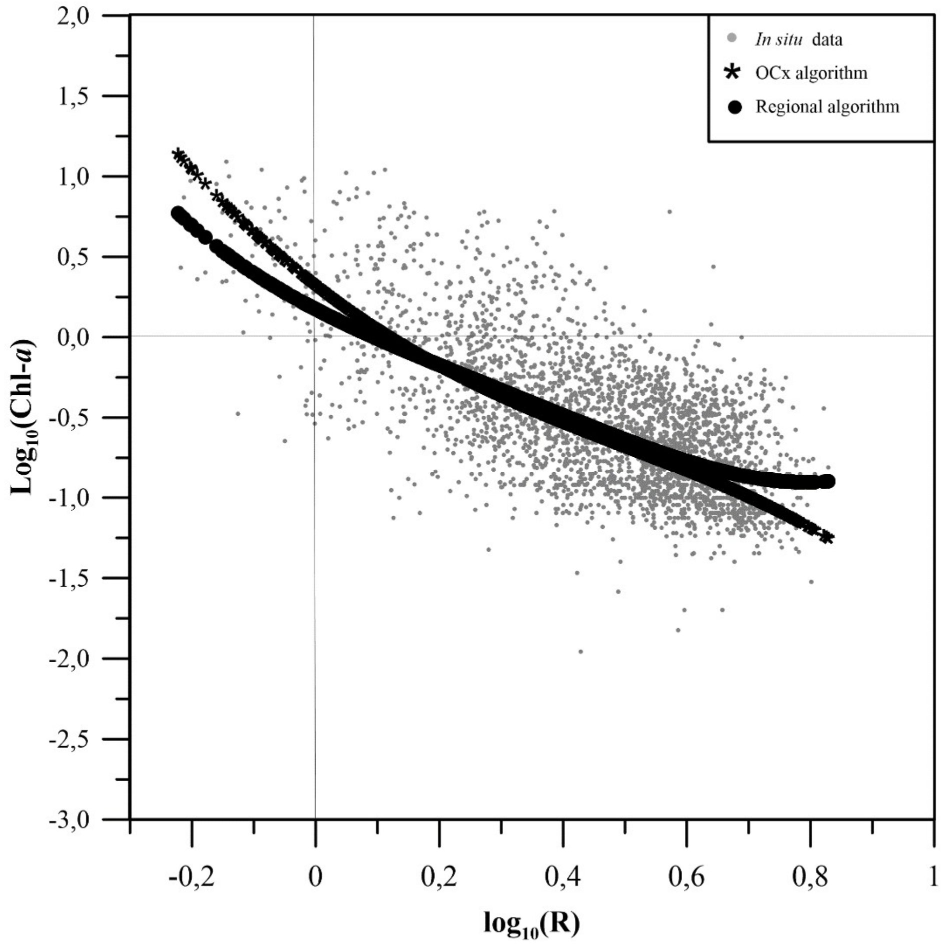

Fig. 3. Regional (circles, Eq. 9) and OCx (triangle) algorithm retrievals applied to remotely sensed reflectance data ($F=\log_{10}(R)$, Eq. 2) off Baja California compared to the $\log_{10}$ of *in situ* Chl-a data.

Separating the dataset

Dynamic properties

The Ensenada Front separates the CC region [28], [22] in a northern mesotrophic area which has been studied by the CalCOFI project, and a southern oligotrophic area studied by the IMECOCAL project. As shown in Fig. 2, the retrieval of Chl-*a* using the CalCOFI 2 algorithm demonstrates the need for different algorithms for this areas. Therefore, a dynamic boundary such as the Ensenada Front separates regions with markedly different bio-optical properties. This suggests that other dynamic boundaries could have the same effect. [17] showed that there is a dynamic boundary within the IMECOCAL region, in front of Punta Eugenia. This boundary interrupts the hydrodynamic interconnection between waters North and South of Punta Eugenia, and [19] suggested the separation is due to the change of sign in the wind stress curl at that latitude. There are two distinct provinces, with cyclonic background circulation to the north and south [17]. This interruption in circulation has consequences in the biological properties of the area. The Northern province is characterized as a subarctic domain while the Southern varies between subarctic and tropical-subtropical domains. Given that the Ensenada Front produces very important differences is the bio-optical properties, it is likely that the boundary in front of Punta Eugenia does the same. Therefore, the dataset was separated to determine if these differences would yield an improvement in the Chl-*a* retrievals. The Southern province includes lines 123 to 137 and the Northern province includes lines 97 to 120, that is, North of Punta Eugenia (Fig. 1). A difference in the distribution or the values of R or Chl-*a* was expected. However, the dispersion of the data is very similar and the polynomials fitted (not shown) were almost the same and very similar to the regional algorithm proposed. Variance in Chl-*a* over the Northern province was of 1.2583 $(mg\ m^{-3})^2$ while in the Southern province was 1.1840 $(mg\ m^{-3})^2$, with no statistical difference between the areas.

Bio-optical properties

As with the dynamic boundaries described before, proximity to the coast is expected to have an effect on the Chl-*a* concentration due to the outflow from the continent, matter transported by the wind, mixing due to shoaling and the anthropogenic effect, among others. Given that the station number is related to nearness to the coast (increases towards the ocean; Fig. 1), the station number was plotted against the green/blue ratio, R^{-1}. Here a pattern could be seen (Fig 4). When the green/blue ratio was higher than or equal to 1, implying that the green reflectance is the same or higher than the blue, the stations were close to the coast, and they probably contain more Chl-*a*. This suggests that the data could be separated using a bio-optical measurement. The dataset was separated in three categories: (a) coastal waters with a green/blue ratio of reflectance of 1 and higher, (b) transitional waters with green/blue ratio between 0.5 and 1, and (c) oceanic waters with green/blue ratio of 0.5 and lower. When the green/blue ratio decreases, the Chl-*a* concentration also decreases.

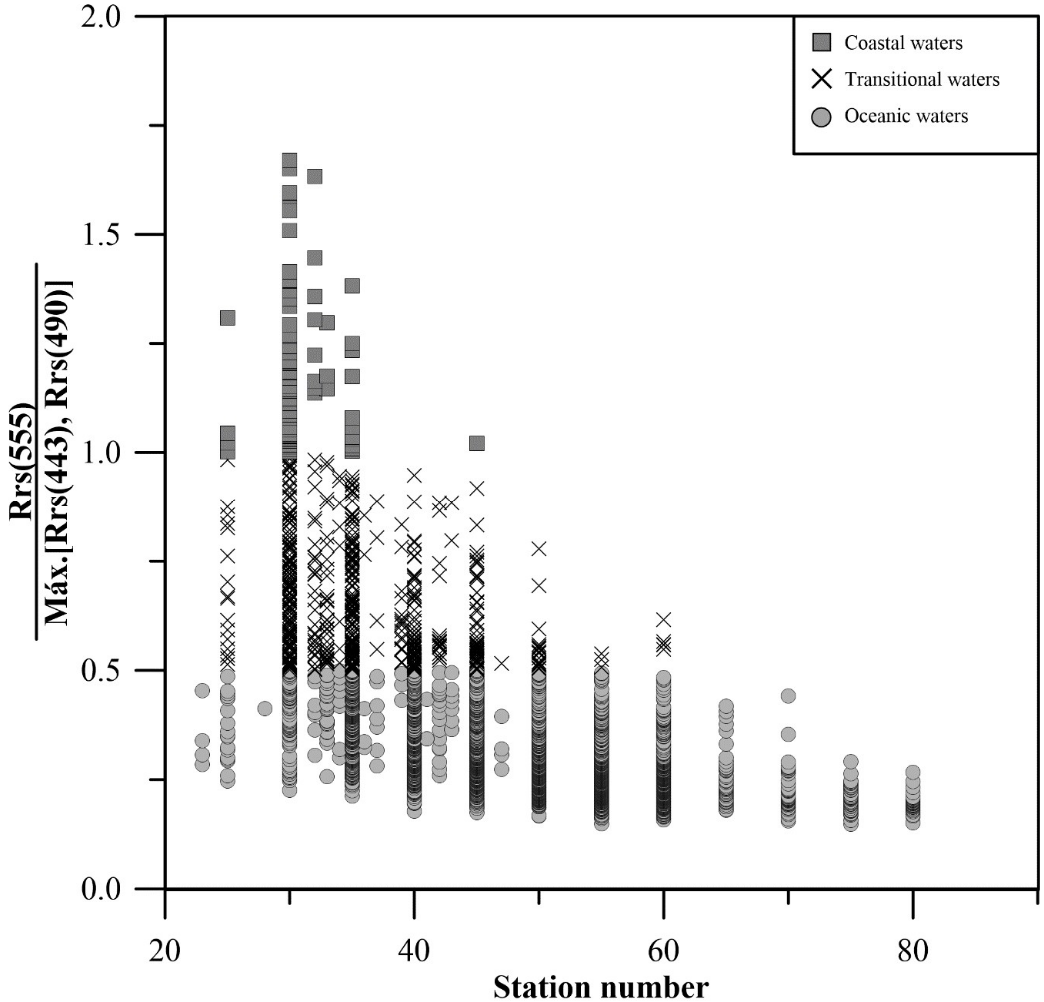

Fig. 4. Variation of green/blue reflectance ratio (R^{-1}) with the station number (Fig. 1) that increases with the distance from the coast. The stations were classified as: coastal waters with a green/blue ratio of reflectance of 1 and higher (squares, 105 stations), transitional waters with green/blue ratio between 0.5 and 1 (crosses, 653 stations), and oceanic waters with green/blue ratio of 0.5 and under (circles, 2607 stations).

A polynomial algorithm was fitted for each section (Fig. 5). The order of the algorithm selected was based on the results of errors (R^2, R_a^2, SSE, RMSLE and AIC). Therefore, the resulting algorithms have different orders: for coastal behavior a linear fit was best (Eq. 10) while for transitional (Eq. 11) and oceanic (Eq. 12) behaviors a second order polynomial gave the best fit.

$$log_{10}(\text{Chl-}a) = 0.2138 - 2.6481F \quad (10)$$

$$log_{10}(\text{Chl-}a) = 0.2501 - 1.7957F - 0.4325F^2 \quad (11)$$

$$log_{10}(\text{Chl-}a) = 0.2786 - 2.1925F + 0.8474F^2 \quad (12)$$

These algorithm´s results follow the general pattern of the data, very similar to the regional algorithm. The functions that represent the data according to their bio-optical characteristics have the same properties as the regionalized function: lower Chl-*a* prediction than OCx for coastal waters (small R, large R^{-1}) and higher Chl-*a* predictions than OCx for oceanic waters (small R^{-1}). The regional algorithm has the advantage of being one function for all the R range whereas the algorithms for coastal, transitional and oceanic behaviors have the advantage of a smaller order, therefore, an easier inversion and less chance of overfitting. It also allows different bio-optical characteristics within the area.

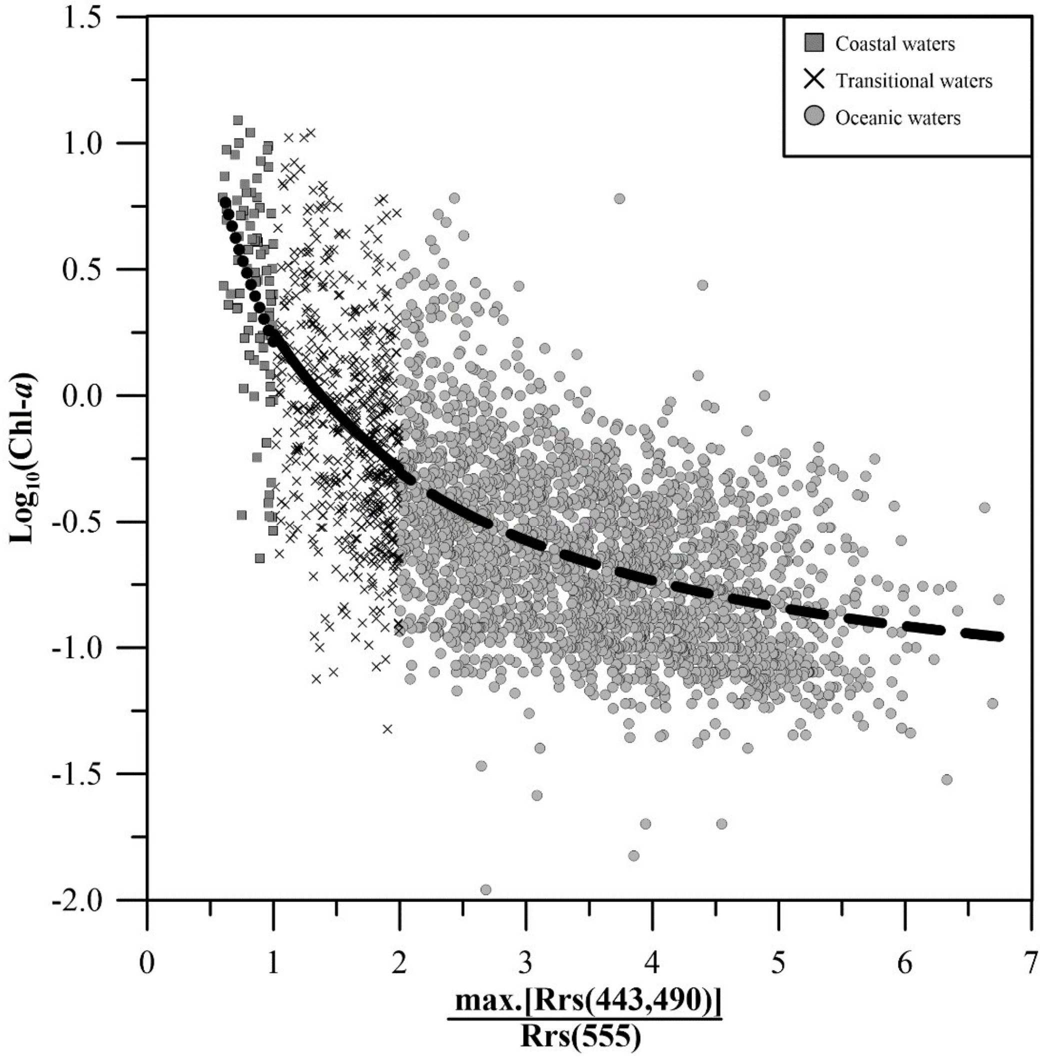

Fig. 5. Retrievals from three algorithms developed considering bio-optical properties in the area compared to the $\log_{10}$ of *in situ* Chl-*a* data.

Climatic events

Besides the dynamic and bio-optical properties of the region, large scale climatic events should be considered to determine how they influence the properties in the region. [28] observed differences between data during ENSO events in the CCS. [21] detected differences in the extent of subarctic waters or tropical and subtropical waters during cold and warm events. This shift in temperature is generally associated to changes in the circulation and in the biological and chemical characteristics. Therefore, an analysis of the differences in the data during these events was made. The data was classified as El Niño, La Niña and normal. Three different indexes were used to classify the data:

- the Southern Oscillation Index (SOI):

https://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/,

- the Oceanic Niño Index (ONI):

https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php,

- the multivariate ESNO index (MEI)

https://www.esrl.noaa.gov/psd/enso/mei/

The data was sorted as El Niño or La Niña when any of the three indexes identified the month and year as warm or cold event, respectively. If the three indexes registered no cold or warm event for that month and year, the data was considered as normal. Once classified, a polynomial function was fitted for each condition (Fig. 6).

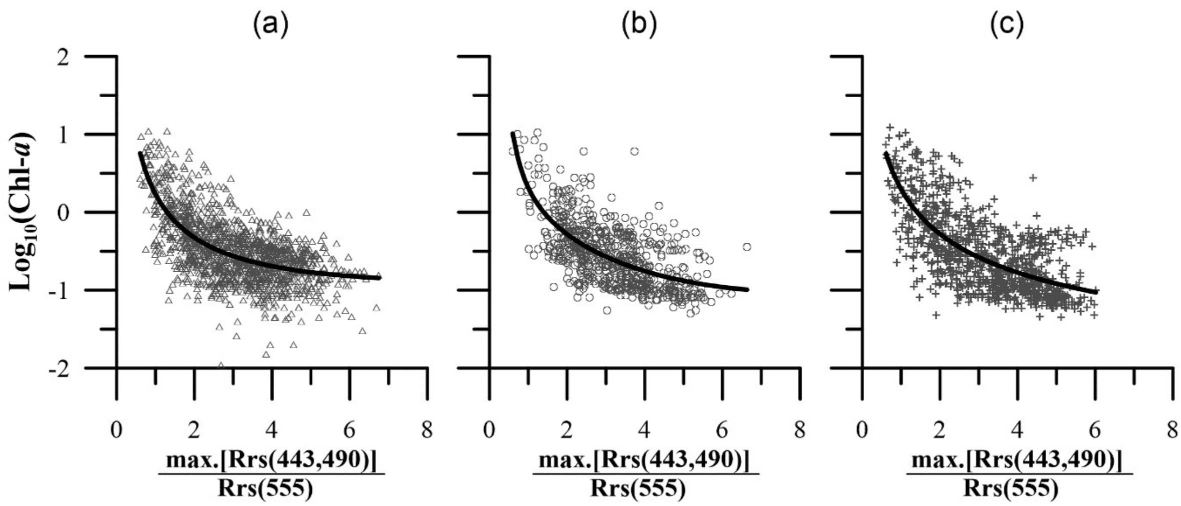

Fig. 6. Algorithm retrievals for data classified as: (a) La Niña; (b) El Niño; (c) normal compared to the $\log_{10}$ of *in situ* Chl-*a* data.

The analysis for interannual warm and cold events in figure 6 demonstrated that most data have high R values and low Chl-*a* during El Niño. Therefore, during El Niño events, waters tend to be more oligotrophic, showing the expected shift towards smaller cells [27]. Also, different behaviors can be seen with the simplest algorithm for the normal (Eq. 13) and La Niña (Eq. 14) data and the more complex for El Niño events (Eq. 15). As before, the selection of the order of the polynomial function was based on the results of errors (R^2, R_a^2, SSE, RMSLE and AIC).

$$log_{10}(\text{Chl-}a) = 0.2962 - 2.0437F + 0.4425F^2 \quad (13)$$

$$log_{10}(\text{Chl-}a) = 0.2337 - 2.1695F + 1.0492F^2 \quad (14)$$

$$log_{10}(\text{Chl} - a) = 0.3141 - 2.4323F + 2.2698F^2 - 3.1653F^3 + 2.0039F^4 \quad (15)$$

Errors

The errors were described in Eq. 6 through 8 and the results are shown in Table 2. The combination of all of this aids in deciding which algorithm fits the data best. In Table 2, the CalCOFI 2 algorithm has the biggest error confirming that this algorithm does not work in the waters off Baja California. All the other algorithms have results of the same order. OCx has better results than CI as seen by the SSE. The smallest value of SSE is for the regional algorithm that considered the climatic condition (El Niño/La Niña). The difference between the coefficient of determination and its adjusted value depends on the degrees of freedom and is reflected in the third decimal number. The RMSLE and the SSE are lowest for the algorithm with the climatic condition. The AIC evaluates the model's fit on the training data, and adds a penalty term for the complexity of the model (number of parameters in the model). The AIC scores are only useful in comparison with other AIC scores for the same dataset. The lower score is considered the best balance of model fit. The lowest AIC score is also for the climatic condition data. The regional algorithm improves from the global algorithms tested even with the same model complexity. However, the scores decrease their value when the dataset is separated considering bio-optical properties and climatic conditions.

TABLE II

COEFFICIENT OF DETERMINATION (R^2), ADJUSTED COEFFICIENT OF DETERMINATION (R_a^2), SUM OF SQUARES ERROR (SSE), ROOT MEAN SQUARED LOGARITHMIC ERROR (RMSLE) AND AKAIKE INFORMATION CRITERION (AIC) FOR EACH OF THE TESTED AND PROPOSED ALGORITHMS.

Error	**CalCOFI 2**	**CI**	**OCx**	**Regional**	**Bio-optical properties**	**Climatic condition**
SSE	156,141.8	450.60	364.03	340.33	338.39	331.52
R^2	- 254.05	0.2640	0.4054	0.4441	0.4473	0.4585
R_a^2	- 254.13	0.2637	0.4046	0.4434	0.4465	0.4579
AIC	25,949.82	8,195.45	7,553.75	7,349.37	7,332.03	7,267.75
RMSLE	7.1715	0.3853	0.3463	0.3348	0.3339	0.3304

DISCUSSION

Global algorithms were tested for the waters off the Baja California Peninsula. The CI and OCx algorithms are very similar for waters with of Chl-*a* lower than 0.3 mg m^{-3} ($\log_{10}$(Chl-*a*) ~ - 0.5) and a ratio of blue/green reflectance of R ≥ 2.5 (F≥0.4) (Fig. 2). CI is considered better for retrievals are below 0.15 mg m^{-3} while OCx works better for higher Chl-*a* values (Chl-*a* ≥0.2 mg m^{-3}) [34]. The natural dispersion of the data is much larger than the difference between the CI and OCx algorithms for Chl-*a* <0.3 mg m^{-3}. For this reason, OCx was considered a better fit for the area and was chosen for the regionalization. The regional algorithm proposed has a better fit than the OCx for the area (Table 2). When comparing the data (Fig. 3) we see that the algorithms differ in the extreme conditions of oligotrophic waters (large R or F, low Chl-*a*) and eutrophic waters (low R or F, high Chl-*a*). The difference between algorithms suggests that water in the IMECOCAL area has higher pigment concentration in oligotrophic waters than the global database with which the OCx parameters were obtained.When there are warm waters in front of the Baja California Peninsula, these are dominated by small cells (cyanobacteria) typical of tropical and subtropical waters [27], [28] which could account for this difference. According to [5], for global data the OCx has a good fit in eutrophic areas because it yields higher Chl-*a* values. However, the IMECOCAL area has data with lower Chl-*a* for small R values which make the regional algorithm yield lower Chl-*a* results. This implies that there are lower pigment concentrations in the eutrophic zone in the IMECOCAL area than those in the global database.

The regional algorithm takes into consideration the specific oceanographical and optical properties of the area and there is an improvement in the prediction for the eutrophic and oligotrophic areas. When this separation was made the resulting models had the least complexity (Eq. 10 through 12). For coastal waters the best fit was a linear model (Fig. 5). As stated before, the best overall fit was when the database was separated considering climatic conditions. El Niño condition has a more complex distribution of data than La Niña and normal conditions. The shift in the phytoplankton community is reflected in the algorithms proving to be more important that the dynamic boundary found close to Punta Eugenia.

CONCLUSIONS

The Chl-*a* satellite images are one of the most used products in oceanography [36]. However, the global algorithms applied are biased because their parameters are obtained with data from all around the globe and for a particular sensor. Therefore, it is important to assess their efficiency in order to determine if they are regionally appropriate before its use [36]. When applied to the IMECOCAL area with the Copernicus satellite data, the two global algorithms (OCx and CI) yielded results within the range of Chl-*a* values expected. Because it includes information of reflectance at 670 nm [11], the CI algorithm produces Chl-*a* values with more dispersion than the OCx. However, for the IMECOCAL area the difference between the results of the CI and OCx algorithms is so slight that the change in algorithm does not warrant improvement in the Chl-*a* estimates.

For the CCS, in front of the California, the global algorithms underestimate the Chl-*a* concentration by a factor of 5 [3], [6]. However, for the IMECOCAL area these algorithms retrieved Chl-*a* of the same order of magnitude as the *in situ* data. The CalCOFI 2 algorithm, on the contrary, drastically underestimates the Chl-*a* concentration in the IMECOCAL area, suggesting that there are significantly different bio-optical properties between these sections of the CC. The fact that the Ensenada Front works as a dynamic boundary that greatly alters the bio-optical properties in different sections of the CCS suggests that the boundary at around 28° N [17] that interrupts ecological connectivity would do the same. Nonetheless, the data are not significantly different showing that this dynamic boundary does not have a perceptible effect in the bio-optical properties of seawater in the area. The differences observed in the algorithms could be due, not only to the bio-optical differences in the data per se but to the differences in the merged satellite data. Therefore, it is important to establish accurate parameters for the specific merged satellite data.

The regional algorithm differs from the global OCx mainly in oligotrophic waters, with higher Chl-*a* values that could be due to the presence of small cells characteristic of tropical and subtropical waters. The database was then classified according to the green/blue ratio of reflectances (R^{-1}) which allowed a distinction between coastal, transitional and oceanic waters, each with a polynomial function fitted to it. Though the differences between the regional algorithm and this set of algorithms are small, the classification in coastal, transitional and oceanic waters allows a better parameterization for different properties as with the combined use of CI and OCx by NASA. The weighed approached used for CI and OCx is based on Chl-*a* value, which requires a retrieval for the determination of the adequate algorithm. The use of this set allows a distinction of properties based of a value of R, that is, the classification of the data is done before the retrieval.

During El Niño events, Chl-*a* was lower than during the transitional and La Niña conditions. When analyzing the interannual variations associated with ENSO, [28] found an increase in the mesotrophic area in the IMECOCAL area during these warm events. This coincides with the fact that for the IMECOCAL area, the dispersion in R values is less during an El Niño. [37] found that during the cold event of 2008 the productivity rates increased. This coincides with the high Chl-*a* observed in the IMECOCAL data during La Niña and normal conditions than during an El Niño.

[36] stated that regionalizing the algorithms diminishes the bias, but uncertainties are still important because of the natural variation of the optical properties of phytoplankton, the presence of other optically active components, and the atmospheric correction. These authors found that removing the 443 nm band improved the performance of the OCx algorithm. There is still much work to do in order to give a more accurate estimate of Chl-*a* from satellite data.

However, the evaluation of the algorithms using merged satellite data gives a deeper understanding of the area and a better option for constructing long time-series that can aid in the comprehension of the processes in the area. With the Copernicus satellite data, the climatic conditions algorithm allows the construction of more consistent time series for the IMECOCAL region.

References

[1] Maritorena, S., D. A. Siegel, , (2205). Consistent merging of satellite ocean color data sets using a bio-optical model. *Remote Sensing Environ*. 94, 429–440. doi: 10.1016/j.rse.2004.08.014

[2] McClain, C. R., (2009). A decade of satellite ocean color observations. *Annu.Rev.Mar. Sci*. 1, 19–42. doi: 10.1146/annurev.marine.010908.163650

[3] Kahru, M., R. M, Kudela, C. R. Anderson, C. R., B. G. Mitchell, (2015). Optimized merger of ocean chlorophyll algorithms of MODIS-Aqua and VIIRS. *Ieee Geoscience and Remote Sensing Letters*. 12(11), 2282-2285. doi: 10.1109/lgrs.2015.2470250

[4] O´Reilly, J. E. and P. J. Werdell, (2019). Chlorophyll algorithms for ocean color sensors – OC4, OC5 & OC6. *Remote Sensing of Environment.* 229, 32-47. doi:10.1016/j.rse.2019.04.021

[5] O'Reilly, J. E., S. Maritorena,, M. C. O'Brien, D. A. Siegel, D. Toole, D. Menzies, , R. C. Smith, , J. L. Mueller, B. G. Mitchell, M. Kahru, F. P. Chavez, P. Strutton, G. F. Cota, S. B. Hooker, C. R. McClain, K. L. Carder, F. Müller-Karger, L. Harding, A. Magnuson, D. Phinney, G. F. Moore, J. Aiken, K. R. Arrigo,R. Letelier, M. Culver, , (2000). SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3. Hooker, S. B. and E. R. Firestone, (Eds.), NASATech. Memo. 2000- 206892. NASA Goddard Space Flight Center 11, 49 pp.

[6] Kahru, M., R. M. Kudela, M. Manzano-Sarabia, B. G. Mitchell, (2012). Trends in the surface chlorophyll of the California Current: Merging data from multiple ocean color satellites. *Deep-Sea Research II.* 77–80:89–98. doi: 10.1016/j.dsr2.2012.04.007

[7] IOCCG, (2007). Ocean-color data merging. In: Gregg, W. (Ed.), Report of the International Ocean-Color Coordinating Group, IOCCG, Dartmouth, Canada, No.6, 68 pp.

[8] Groom, S., S. Sathyendranath, Y. Ban, S. Bernard, R. Brewin,V. Brotas, C. Brockmann, P. Chauhan, J. L. Choi, A. Chuprin, S. Ciavatta, P. Cipollini, C. Donlon, B. Franz, X. He, T. Hirata, T. Jackson, M. Kampel, H. Krasemann, S. Lavender, S. Pardo-Martinez, F. Mélin, T. Platt, R. Santoleri, J. Skakala, B. Schaeffer, M. Smith, F. Steinmetz, A. Valente, M. Wang, (2019). Satellite Ocean Colour: Current Status and Future Perspective. *Front. Mar. Sci*. 6:485. doi: 10.3389/fmars.2019.00485

[9] Boyce, D.G., M. L. Lewis, B. Worm, B, (2010). Global phytoplankton decline over the past century. *Nature*, 466, 591–596. doi: 10.1038/nature09268.

[10] O'Reilly, J. E., S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru, C. R. McClain, (1998). Ocean color chlorophyll algorithms for SeaWiFS. *J.Geophys.Res.* 103, 24937–24953. doi: 10.1029/98JC02160

[11] Hu, C., Z. Lee, B. Franz, B., (2012). Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *J. Geophys. Res.* 117, C01011, doi: 10.1029/2011JC007395.

[12] Kahru, M., R. M. Kudela, C. R. Anderson, M. Manzano-Sarabia, B- G- Mitchell, (2014). Evaluation of satellite retrievals of ocean chlorophyll-a in the California Current. *Remote Sensing,* 6:9, 8524-8540. doi: 10.3390/rs6098524

[13] Lee, S.H., J. Ryu, J. Park, D. Lee, J. Kwon, J. Zhao, S. Son, (2018). Improved Chlorophyll-a Algorithm for the Satellite Ocean Color Data in the Northern Bering Sea and Southern Chukchi Sea Ocean *Sci. J.,* 53(3), 475-485. doi: 10.1007/s12601-018-0011-5

[14] Mustapha, S. B., S. Bélanger, P. Larouche, P., (2012). Evaluation of ocean color algorithms in the southeastern Beaufort Sea, Canadian Arctic: New parameterization using SeaWiFS, MODIS, and MERIS spectral bands, *Can. J. Remote Sens.*, 38:5, 535-556, doi: 10.5589/m12-045

[15] Lynn, R.J. and J. J. Simpson, (1987). The California Current system: The seasonal variability of its physical characteristics, *J. Geophys. Res. Oceans,* 92: C12, 12947-12966. doi: 10.1029/JC092iC12p12947

[16] Kurian, J., F. Colas, X. Capet, J. C. McWilliams, D. B. Chelton, (2011). Eddy properties in the California Current System, *J. Geophys. Res.,* 116, C08027. doi: 10.1029/2010JC006895.

[17] Durazo, R., A. M. Ramírez Manguilar, L. E. Miranda, L. A. Soto Mardones,(2010). Climatología de variables oceanográficas. In: GaxiolaCastro G, Durazo R (eds.), Dinámica del ecosistema pelágico frente a Baja California 1997-2007. SEMARNAT, CICESE, UABC, pp .25-58.

[18] Venrick, E. L., (2000). Summer in the Ensenada Front: The distribution of phytoplankton species, July 1985 and September 1988. *Journal of Plankton Research*, 22:5, 813–841. doi: 10.1093/plankt/22.5.813

[19] Durazo, R., (2015). Seasonality of the transitional region of the California Current System off Baja California, *J.Geophys. Res. Oceans*, 120,1173–1196. doi: 10.1002/2014JC010405.

[20] Gaxiola-Castro, G., R. Durazo, B. Lavaniegos, M. E. De La Cruz-Orozco, E. Millán-Núñez, L. Soto-Mardones, J. Cepeda-Morales, (2008).. Pelagic ecosystem response to interanual variability off Baja California, *Ciencias Marinas*, 34:2, 263–270. doi: 10.7773/cm.v34i2.1413

[21] Durazo, R., (2009). Climate and upper ocean variability off Baja California, Mexico: 1997–2008, *Progress in Oceanography,* 83:1–4, 361-368, doi:10.1016/j.pocean.2009.07.043.

[22] Haury, L. R., E. L. Venrick, C. L. Fey, J. A. McGowan, P. P. Niiler, (1993). The Ensenada Front. *CalCOFI Rep.,* 34: 69-88.

[23] Mitchell, B.G. and M. Kahru, (1998). Algortihms for SeaWiFs standard products developed with CalCOFI bio-optical data. *CalCOFI Rep*., 39, 133-147.

[24] Kahru, M., B. G. Mitchell, (1999). Empirical chlorophyll algorithm and preliminary SeaWiFS validation for the California Current, *International Journal of Remote Sensing,* 20:17, 3423-3429, doi: 10.1080/014311699211453.

[25] Gaxiola, G.G., J. Cepeda–Morales, S. Nájera–Mártinez, T. L. Espinosa–Carreón, M. E. De la Cruz–Orozco, R. Sosa–Avalos, E. Aguirre–Hernández, J. P. Cantú–Ontiveros, (2010). Biomasa y producción del fitoplancton. Durazo, R., Ramírez Manguilar, A.M., Miranda, L.E., Soto Mardones, L.A. Climatología de variables oceanográficas. In: Gaxiola Castro G, Durazo R (eds.), Dinámica del ecosistema pelágico frente a Baja California 1997-2007. SEMARNAT, CICESE, UABC, pp. 59–86.

[26] Moran, X. A., A. Lopez–Urrutia, A. Calvo–Diaz, W. K. W. Li, (2010). Increasing importance of small phytoplankton in a warmer ocean. *Glob. Change Biol.* 16, 1137–1144. doi: 10.1111/ /j.1365-2486.2009.01960.x

[27] González-Silvera, A., E. Santamaría-del-Ángel, V. Camacho-Ibar, J. López-Calderón, J. Santander-Cruz, A. Mercado-Santana, A., (2020). The Effect of Cold and Warm Anomalies on Phytoplankton Pigment Composition in Waters off the Northern Baja California Peninsula (México): 2007–2016. *J. Mar. Sci. Eng.*, 8, 533, doi:10.3390/jmse8070533

[28] Kahru, M. and B. G. Mitchell,(2000). Influence of the 1997-98 El Niño on the surface chlorophyll in the California Current. *Geophys. Res. Letters*. 27(18): 2937-2940.

[29] Durazo, R. and T. Baumgartner, (2002). Evolution of oceanographic conditions off Baja California: 1997–1999. *Progr. Oceanogr*., 54: 7–31.

[30] Venrick, E. L. and T. L. Hayward, (1984) Determining chlorophyll on the 1984 CalCOFI surveys CalCOFI, Rep. 25, 74-79.

[31] Holm Hansen, O., C. J. Lorenzen, R. W. Holm, J. D. H. Strickland, (1965). Fluorometric determination of chlorophyll. *J. Cons Int. Explor. Mer.,* 1965, 30, 3-15.

[32] Yentsch, C. S. and D. W. Menzel, D. W., (1963). A method for the determination of phytoplankton, chlorophyll and phaeophytin by fluorescence, *Deep-Sea Res.,* 10, 221-231.

[33] De la Cruz-Orozco, M. E., E. Gómez-Ocampo, L. E. Miranda-Bojórquez, J. Cepeda-Morales, R. Durazo, B. E. Lavaniegos, T. L. Espinosa-Carreón, R. Sosa-Avalos, E. Aguirre-Hernández, G. Gaxiola-Castro, (2017). Phytoplankton biomass and production off the Baja California Peninsula: 1997 – 2016. *Ciencias Marinas*, 43, 217-228. doi: 10.7773/cm.v43i4.2793

[34] NASA. (2010). SeaWiFS Project online. Seattle (WA, USA): NASA. Available online: https://oceancolor.gsfc.nasa.gov/SeaWiFS/ (accessed on 09/10/2019).

[35] Burnham, K. P. and D. R. Anderson, (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd Edition, Springer-Verlag, U.S.A.,488 pp.

[36] Clay, S., A. Peña, B. DeTracey, E. Devred, (2019). Evaluation of Satellite-Based Algorithms to Retrieve Chlorophyll-a Concentration in the Canadian Atlantic and Pacific Oceans. *Remote Sens.*, 11:2609. doi: 10.3390/RS11222609

[37] Gomez-Ocampo, E., G. Gaxiola-Castro,R. Durazo, E. Beier, (2017). Effects of the 2013-2016 warm anomalies on the California Current phytoplankton. *Deep-Sea Res. II,* 151, 64-76. doi: 10.1016/j.dsr2.2017.01.005.

CONCLUSIONS

Ocean color algorithms relate remotely sensed data with optical properties or biogeochemical variables. Due to the vastness of the ocean and the fact that collecting data from it is very expensive in time and resources, remotely sensed data has given scientists a unique opportunity to study phenomena with better resolution in time and covering areas that would be difficult to reach. However, the best results from the algorithms come when they are adapted to the particular characteristics of an area. Therefore, in order to improve the knowledge of optical properties and biogeochemical variables, algorithms need to be tested and regionalized.

Ocean color data records have been registered for a couple of decades with many different sensors. This has made it difficult to compare between results as well to provide consistency within the dataset across satellite instruments [1]. Therefore, changes that span a decade or longer cannot be adequately studied with an individual sensor. For instance, [2] studied the variation of phytoplankton biomass over the last century finding that there is a decline in most oceans. Thus, there has been an effort to improve the single instrument results [3] through the merging of the individual sensors and through the future use of sensors of similar design with a planned replacement strategy [4]. Combining the data increases the coverage but it requires instrument consistency, compatibility of the atmospheric correction method and consistency of the algorithms applied [1]. Satellite constellations like Copernicus ensure the provision of integrated, sustained and consistent ocean color datasets [4] building a three decade long time series that will allow researchers to discriminate between climate change and natural variability [4].

In this research two kinds of remotely sensed data were used. For the first chapter, an individual sensor was used while the second chapter, the satellite data was from a merged product. The first paramount difference between these is the wavelengths available. The individual sensor included data at 510 nm. This wavelength is not available for the merged data. As expected, there was more invalid data with the single sensor dataset than with the merged data.

Two types of variable retrievals where tested and improved through this research: (1) the optical properties ratio through a semianalytical quasi-single scattering approximation algorithm with atmospheric correction, and (2) concentration of a photosynthetic pigment (Chl-*a*) through an empirical quasi-single scattering approximation algorithm with atmospheric correction. For both of these retrievals, working algorithms were tested and, because of the results obtained, they were adapted to the IMECOCAL area. These algorithms improved the retrieval of IOPs or Chl-*a*. Throughout this research, it is shown that seasonal variability or climatic variability due to events such as El Niño or La Niña were of dire importance in modeling the area. Therefore, the algorithms proposed differ according to these conditions adequately representing the dynamic of the area.

Notwithstanding, when global algorithms are applied, retrievals are made and then used to classify the data. This classification may result in having to apply another algorithm or a combination of them. With the algorithms proposed throughout this paper, data can be sorted before the algorithm is applied.

As more data become available, it can be used in the algorithms to continue to both test and adapt them. Particularly, the merged satellite product has to be tested for the IOPs algorithms. In further research, the inversion of these functions will be done thorough different methods in order to identify which provides the best result. Also, a functional relationship between IOPs and Chl-*a* is still pending.

References

[1] O´Reilly, J. E. and P. J. Werdell, (2019). Chlorophyll algorithms for ocean color sensors – OC4, OC5 & OC6. *Remote Sensing of Environment.* 229, 32-47. doi:10.1016/j.rse.2019.04.021

[2] Boyce, D.G., M. L. Lewis, B. Worm, (2010). Global phytoplankton decline over the past century. *Nature*, 466, 591–596. doi: 10.1038/nature09268.

[3] IOCCG, (2007). Ocean-color data merging. In: Gregg, W.(Ed.), *Report of the International Ocean-Color Coordinating Group*, No. 6 , IOCCG ,Dartmouth, Canada.

[4] Groom, S. , S Sathyendranath, Y. Ban, S. Bernard, R. Brewin, V. Brotas, C. Brockmann, P. Chauhan, J. L Choi, A. Chuprin, S. Ciavatta, P. Cipollini, C. Donlon, B. Franz B, X. He, T. Hirata, T. Jackson, M. Kampel, H. Krasemann, S. Lavender, S. Pardo-Martinez , F. Mélin, T. Platt, R. Santoleri, J. Skakala, B. Schaeffer B, M. Smith, F. Steinmetz, A. Valente M. and Wang, (2019) Satellite Ocean Colour: Current Status and Future Perspective. *Front. Mar. Sci*. 6:485. doi: 10.3389/fmars.2019.00485

Appendix A. A Brief Comparison of the Optical Properties in the California Current System in front of California and Baja California

In the vicinity of 32° N, there is a persistent feature called the Ensenada Front, that has been recognized as an ecological transition zone characterized by a strong gradient in Chl-*a*. CalCOFI and IMECOCAL have been collecting data pertaining to the CC oceanic environment for many years. It is important to compare data from both areas in order to establish if there are significant differences between them of if they can be treated as the same area. Thus, CalCOFI data from to campaigns ranging between January 2000 through October 2008 were compared to IMECOCAL data collected between January and October 1999. Both databases are comprised of the absorption coefficient due to detritus and particulate matter. With these *in situ* data, the inherent optical properties ratio $X(\lambda) = \frac{b_b(\lambda)}{b_b(\lambda)+a(\lambda)}$ was obtained.

In Fig. A the $X(\lambda)$ spectrum is shown for the CalCOFI and the IMECOCAL data. In this figure we can see that the range of $X(\lambda)$ is smaller for IMECOCAL area, approximately, between 0 and 0.17, while in the CalCOFI region $X(\lambda)$ varies between 0 and 0.38, implying less variability in the absorption in the area. This, in turn, shows that the IOPs are different. Therefore there are bio-optical differences in the area.

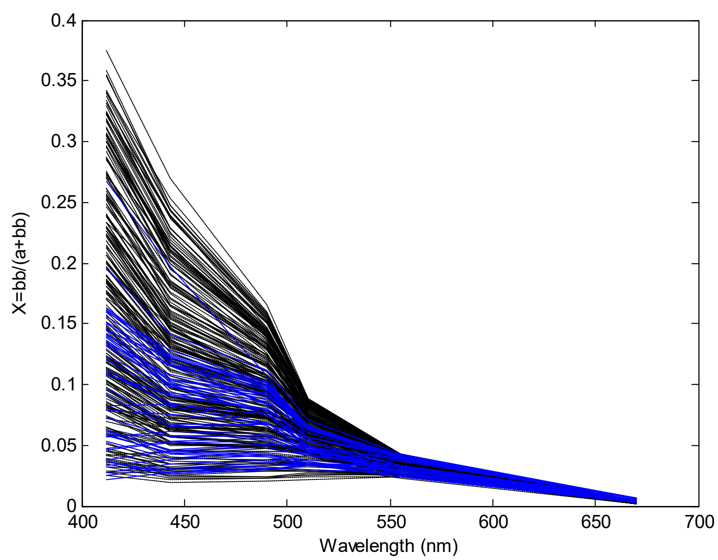

Figure A1. $X(\lambda) = \frac{b_b(\lambda)}{b_b(\lambda)+a(\lambda)}$ *spectrum for the CalCOFI (black) data and the IMECOCAL data (blue).*

Appendix B. Sensibility Analysis in the Quotient of Inherent Optical Properties

In order to analyze the effect of the backscattering coefficient on the retrieval of $X(\lambda) = \frac{b_b(\lambda)}{b_b(\lambda)+a(\lambda)}$, we calculated this quotient for 50 different values of magnitude of the backscattering coefficient (A in Eq. 7, Chapter I). These values were selected within the range stated in the manuscript, $0.0008 < \mathrm{A} < 0.0013$ m^{-1}. The $X(\lambda)$ calculated with these values of A are shown as red dots in Figure B(a). These results fall within the expected range of $X(\lambda)$ for the IMECOCAL area (black lines). The influence of the variation of A in the retrieval of $X(\lambda)$ is minimal. To serve as comparison, the same calculation was done but now varying the magnitude absorption coefficient of the particulate matter as shown in Figure B(b). This figure clearly demonstrates that the influence of the absorption coefficient is much greater than that of the backscattering coefficient for the retrieval of $X(\lambda)$. Considering that the influence of the $b_{bp}(\lambda)$ is minimal in $X(\lambda)$ we have used the same model for oceanic and coastal regions, allowing only seasonal variations.

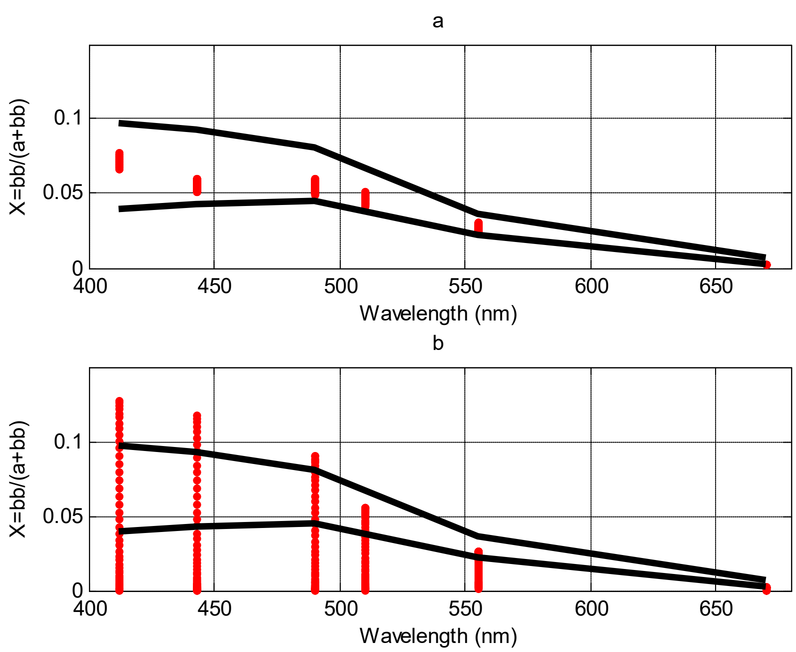

Figure B1. *Variation of the $X(\lambda)$ spectrum for 50 different values (a) of A and (b) of the magnitude of the absoprtion coefficient of the particulate matter $a_p(\lambda)$. The black lines represent the range within which $X(\lambda)$ is expected for the IMECOCAL area.*