How ambiguous is the inverse problem of ocean color in coastal waters?

Michael Defoin-Platel, Malik Chami

How ambiguous is the inverse problem of ocean color in coastal waters?

Michael Defoin-Platel[1,2] and Malik Chami[1]

Received 28 July 2006; revised 4 October 2006; accepted 12 October 2006; published 6 March 2007.

[1] The inverse problem of ocean color consists in deriving the inherent optical properties (IOP) of marine particles from a reflectance spectrum measured at the sea surface. Such a problem is ill-posed or ambiguous because of the nonuniqueness of the solution; that is, several combinations of IOP values can lead to a unique reflectance spectrum. Currently, great efforts are made in the development of inverse methods to accurately retrieve the IOPs. However, many fewer studies have been devoted to the analysis of the ambiguities, which affect yet the error on the IOPs retrieval. In this paper, the ambiguities related to the ocean color problem in coastal waters are characterized and their implications for inverse modeling are studied. A synthetic data set is created on the basis of radiative transfer modeling. The simulations are constrained using in situ observations and statistical rules to make the data set realistic. The ambiguity rate of remote sensing reflectance (Rrs) spectra is around 90%, thus meaning that the ocean color problem is extremely ambiguous. The influence of the ambiguities on the IOPs retrieval is evaluated. It is demonstrated that the error that is ascribed to the occurrence of ambiguity is equal to the dispersion of all the plausible IOPs solutions. The ambiguity error made on the total absorption coefficient is shown to be greater in highly absorbing water mass. On the other hand, the ambiguity error made on the total backscattering coefficient is higher in turbid scattering waters. Finally, different strategies to reduce the effects of ambiguities are discussed.

Citation: Defoin-Platel, M., and M. Chami (2007), How ambiguous is the inverse problem of ocean color in coastal waters?, *J. Geophys. Res.*, *112*, C03004, doi:10.1029/2006JC003847.

1. Introduction

[2] The inverse problem of ocean color consists in determining biogeochemical parameters such as the chlorophyll *a* concentration from the upwelling radiance spectrum. Because the inherent optical properties (IOPs) of the particles, namely the scattering, backscattering and absorption coefficients, are at the center of measured water leaving radiance and water constituents, the inverse problem is often examined as a two steps process: the derivation of IOPs from the radiance, and then biogeochemical parameters from the IOPs. Such an IOPs-based inversion maximizes the information gained from remote sensing. A number of methods for the IOPs retrieval of each water components, namely phytoplankton (Chl), nonalgal particles (NAP), inorganic material (sed) and colored dissolved organic matter (CDOM) from remotely sensed data have been proposed. Empirical approaches consist in establishing statistical relationships between the marine reflectance and the desired optical parameters. This method is fairly reliable when the IOPs are mainly ascribed to phytoplankton and associated products as observed in the open ocean [*Gordon et al.*, 1983; *Morel*, 1988; *O'Reilly et al.*, 1988]. However, it can be highly inaccurate in coastal zones owing to the prevailing contribution to the light field from inorganic and/or dissolved material. In the latter case, semianalytical algorithms, which are based on radiative transfer simulations, are used to approximate the relationship between the IOPs and the reflectance. Equation 1 provides an example of relationships that is often used as a basis of many inversion algorithms:

$$\mathrm{Rrs} = L_u/E_d = g\frac{\mathrm{b_{btot}}}{\mathrm{a_{tot}}}, \tag{1}$$

where Rrs is the subsurface remote sensing reflectance, defined by the ratio of nadir-viewed upwelling radiance L_u (in W m^{-2} sr^{-1}) to downwelling irradiance E_d (in W m^{-2}) just beneath the sea surface, $\mathrm{b_{btot}}$ is the total backscattering coefficient (in m^{-1}), $\mathrm{a_{tot}}$ is the total absorption coefficient (in m^{-1}) and g is a proportionality factor (sr^{-1}).

[3] The major difficulty of solving the inverse problem of ocean color rests on the fact that the relationship between the Rrs and the IOPs of each water component is mathematically not a bijection. Because of their additive property,

[1]Laboratoire d'Océanographie de Villefranche sur Mer, Unité Mixte de Recherche CNRS 7093, Université Pierre et Marie Curie-Paris, Villefranche sur Mer, France.
[2]Also at Laboratoire d'Informatique Signaux et Systèmes de Sophia Antipolis, UNSA-CNRS 2000, Sophia Antipolis, France.

0148-0227/07/2006JC003847

IOPs of each water component combined in different ways can lead to identical sum of IOPs and thus to similar reflectance values. Furthermore, a given value of the ratio b_{btot}/a_{tot} can be obtained from different values of b_{btot} and a_{tot}, thus also generating a similar reflectance value as inferred by equation 1. Because the solutions are not unique, the inverse problem of ocean color is said ambiguous or ill-posed. Therefore, fundamentally, both the inversion algorithm and the presence of ambiguities are sources of error in the retrieved IOPs. Note that the error due to the occurrence of ambiguities also corresponds to the minimum error that can be made when inverting the reflectance, whatever the algorithm employed. Most of the current studies focus on the development of original inversion algorithms [*Carder et al.*, 1999; *Chami and Robilliard*, 2002; *Garver and Siegel*, 1997; *Hoge and Lyon*, 1996; *Lee et al.*, 2002, 1996; *Maritorena et al.*, 2002; *Roesler and Boss*, 2003; *Roesler and Perry*, 1995; *Schiller and Doerffer*, 2005]. These studies contribute to significantly document and reduce the fraction of the error that is ascribed to the methodology. The variability of the error related to the methodology was recently studied in the framework of a rigorous intercomparison exercise organized by the *International Ocean Colour Coordinating Group* [2006]. However, much less efforts have been devoted to examine the influence of the ambiguities caused by the non bijectivity of the inversion problem on the error made in the retrieved IOPs. It is the purpose of this paper to deeply analyze the nonuniqueness of the solution of the inversion problem and the resulting implications in terms of accuracy of IOPs retrieval. The objective is thus to quantify the fraction of the error that is ascribed to the physical part of the inversion problem rather than that due to the algorithmical part. Our analysis relies on a synthetic data set that is representative of optically complex waters such as those found in coastal areas. The paper is organized as follows. First, the radiative transfer simulations are described and the synthetic data set is analyzed. Second, formal mathematical definitions of the ambiguities are provided and the distribution of the nonuniqueness of the solution within the data set is characterized. Then the impact of the ambiguities on the accuracy of the retrieved optical parameters is evaluated. Finally, several strategies that might be relevant to tackle the ambiguity problem are discussed.

2. Simulation and Analysis of the Synthetic Data Set

[4] A great part controlling the success of an algorithm developed to address the inverse problem of ocean color in coastal waters is its ability to correctly represent real-world conditions. One way of doing consists in using empirical data set. However, inverse models based on in situ measurements strongly depend on both the quantity and the quality of the available data. It is also often very difficult to gather enough measurements, especially in coastal waters, to properly educate an inversion algorithm. Furthermore, owing to the huge variability in the optical properties of particles in coastal waters, empirical-based methods may lead to produce overspecialized models that cannot be efficiently applied on unseen data, i.e., data that are not used to build the inverse model. Another way to make an inversion algorithm able to represent real-world conditions is to use a simulated data set constrained by observations. In this latter case, a great variety of optical conditions can be taken into account in the data set. Furthermore, a huge number of data can be simulated such that the algorithm can be reliable, from a statistical point of view. A synthetic data set has often been used to describe and validate methods suitable for data inversion [*Gross et al.*, 2000; *Schiller and Doerffer*, 1999]. Currently, a large set of efficient inverse methods, all having their own specificities, is available. Further significant improvements in the solving of the inverse problem of ocean color in coastal waters will come firstly from the quality of data creation process and then from a deeper knowledge of the statistical properties of the data set, in particular of the (often underestimated) question of ambiguities. Afterward, we think that, with a well known characterization of the data set, the judicious choice of appropriate inverse methods will be greatly facilitated. However, most of the current studies have rarely gone to great efforts to rigorously design and analyze the simulated data set. Recently, the working group of IOCCG proposed a simulated data set inspired by observations [*International Ocean Colour Coordinating Group*, 2006] but a thorough analysis of the data set itself is still missing. Here a synthetic data set is generated following the same logic as that proposed by IOCCG. However, the originality of our approach is to focus not only on the physical but also on the statistical constraints to ensure a generation of a synthetic data set which is realistic with respect with observations and reliable for algorithmic development purposes.

2.1. Forward Model

[5] The simulations were performed using the OSOA radiative transfer model [*Chami et al.*, 2001]. The OSOA model solves the vector radiative transfer equation for the coupled atmosphere-ocean system using the successive orders of scattering method. Given a set of inherent optical properties in the water column, the OSOA model outputs the angular distribution of the radiance field. The originality of OSOA model, when comparing to other widely used models such as Hydrolight [*Mobley*, 1989], is to account for the polarization state of light in the water mass. Note that owing to that interesting feature, it was not relevant in this study to use the synthetic data set generated in the framework of the IOCCG working group, which is based on Hydrolight simulations.

[6] We report here on the inputs parameters used in the OSOA model. Standard atmosphere with tropospheric aerosols T70 [*Shettle and Fenn*, 1979] having an optical depth of 0.2 at 555 nm (i.e., horizontal visibility of 23 km) was used to simulate the incoming solar light. The solar zenith angle was set up to 30°. The nadir-viewed remote sensing reflectance was simulated just beneath the surface. The wavelengths used to compute the Rrs were: 412 nm, 443 nm, 490 nm, 510 nm, 555 nm, 620 nm and 665 nm. The set of wavelengths is referred to as $L = \{412, 443, 490, 510, 555, 620, 665\}$. The inherent optical properties required for the simulation are the absorption coefficients, the scattering coefficients and the phase function of the particles. For computation of the IOPs, a five components seawater model is considered. The five components are pure seawater, phytoplankton and their covarying particles (Chl),

Table 1. Range of Variations of the Optical Parametersat 443 nm[a]

Data	Mean	Sigma	Min	Max
a_{chl}	1.43E-01	1.50E-01	6.12E-03	9.04E-01
a_{nap}	1.39E-01	2.23E-01	2.05E-03	1.73E+00
a_{CDOM}	5.01E-01	2.88E-01	9.27E-05	1.00E+00
b_{chl}	1.87E+00	3.14E+00	1.81E-05	2.08E+01
b_{sed}	1.85E+00	3.07E+00	1.04E-04	2.06E+01
bb_{chl}	1.73E-02	3.58E-02	1.00E-08	3.94E-01
bb_{sed}	4.64E-02	7.89E-02	1.00E-08	6.59E-01
a_p	2.83E-01	3.04E-01	8.56E-03	2.57E+00
b_p	3.72E+00	5.08E+00	7.59E-02	2.10E+01
bb_p	6.37E-02	9.62E-02	2.40E-04	6.64E-01
a_{tot}	7.91E-01	4.16E-01	2.07E-02	3.15E+00
b_{tot}	3.72E+00	5.08E+00	8.07E-02	2.10E+01
bb_{tot}	6.61E-02	9.62E-02	2.60E-03	6.66E-01
Rrs	1.40E-02	1.24E-02	3.30E-04	1.28E-01

[a]Sigma is the standard deviation. Read, for example, 6.12E-03 as 6.12×10^{-3}.

colored dissolved organic matter (CDOM), nonalgal particles (NAP) and inorganic particles (sed). In the notation, IOP_v is a vector composed by each component of the IOPs, such that $IOP_v = \{a_{CDOM}, a_{NAP}, a_{chl}, b_{chl}, b_{sed}, b_{bchl}, b_{b_{sed}}\}$. In this water model, the total absorption coefficient a_{tot} is the sum of contributions from all components. Since CDOM is supposed to be nonscattering material, the total scattering coefficient b_{tot} and backscattering coefficient b_{btot} are the sum of all other components. We note also a_p, b_p and b_{bp}, respectively the absorption, scattering and backscattering coefficients of particles. The phase functions of particles are computed using Mie theory and different refractive indices for the phytoplankton and inorganic particles. Standard ranges of variations in the refractive index are used, namely [1.05, 1.10] and [1.15, 1.20] for phytoplankton and inorganic matter respectively. The size distribution of the particles is modeled as a Junge power law. The Junge exponent is considered as a parameter varying in the range [3, 5]. Note that the variations of the refractive indices and the Junge exponent mean that many directional effects of particles are included in the computations, which is consistent with the variability observed in the phase functions in coastal areas [*Chami et al.*, 2005, 2006]. Such important features are not necessarily accounted for in synthetic data sets in which a single particle phase function is used [*Gross et al.*, 2000; *Schiller and Doerffer*, 1999]. A number of $N = 10000$ reflectance spectra are simulated to cover a large variety of optical coastal waters conditions. For convenience, the synthetic data set is hereafter referred to as DPC_{K2}.

2.2. Random but Constrained Generation of Data

[7] To feed the OSOA model with several thousands sets of IOP vectors a random assignment of each IOP value is performed. Nevertheless two different types of constraints have to be taken into account.

[8] 1. The ''realistic constraints'' arise from the optical and biogeochemical properties of the oceanic constituents. Typically, these constraints consist in the selection of appropriate ranges of variation for each IOP, ranges of covariation between IOPs but also appropriate spectral dependencies for each IOP. The realistic constraints have to be respected to generate synthetic spectra similar to natural ones. In this study, we will mainly use the set of COASTlooc measurements [*Babin et al.*, 2003a, 2003b] to satisfy these realistic constraints. The COASTlooc data, which were acquired in many European coastal waters (387 data in total), are highly appropriate inasmuch as they cover a wide variety of optical conditions, ranging from CDOM dominated waters (such as the Baltic Sea) to mineral dominated waters (such as the English Channel).

[9] 2. The ''statistical constraints'' require that, in a ideal data set, all the variables within the samples should be independent and identically distributed [*Vapnik*, 1995]. The notation *iid* is hereafter used to specify a data set which is independent and identically distributed. Nevertheless, it is clear that this ideal case cannot be reached owing to the realistic constraints above mentioned. For example, the independence of variables is in our case meaningless because of the spectral variations properties of the IOPs. The distribution of values within any data sets needs to be also investigated carefully since it could greatly influence the performance of the resulting inverse models. The inverse models built on uniformly distributed data usually report uniformly distributed errors so that the overall quality of these models can be systematically predicted with precision. However the distribution of in situ collected data may not be uniform. This could be problematic for a user that intends to invert them using models build on uniformly distributed data. Therefore in the case of measurements demonstrating overrepresented values, i.e., a nonuniform distribution, it is clear that the inverse models should be overspecialized on these prevailing values even though this overspecialization leads to introduce differently distributed (i.e., not *iid*) values within the data sets.

2.3. A Two-Step Creation Process

[10] The procedure that is used to create the synthetic data set includes two steps. First, the IOP values at 443 nm are fixed for the N samples of DPC_{K2} accounting for the constraints above described. Second, the IOP values are determined at other wavelengths.

2.3.1. First Step: Setting of the IOP Values at 443 nm

[11] We start with the random assignment of the particulate scattering coefficient $b_p(443)$ so that each value is drawn in a specific range of variation based on COASTlooc measurements at 443 nm (Table 1). The contribution to $b_p(443)$ from $b_{chl}(443)$ (i.e., b_{chl}/b_p) is uniformly drawn in the range [0, 1], thus allowing the determination of both $b_{chl}(443)$ and $b_{sed}(443)$. The backscattering coefficient of phytoplankton b_{bchl} and inorganic matter b_{bsed}, are calculated by integrating their respective phase function in the backward hemisphere and multiplying the result of the integration by their respective scattering coefficient. Note that the values of $b_{bchl}(443)$ and $b_{bsed}(443)$ are not independent.

[12] To account for the covariations of the particulate scattering coefficient b_p with the particulate absorption properties, namely the covariations between (1) b_p and a_{chl} and (2) b_p and a_{NAP}, the values of $a_{chl}(443)$ and $a_{NAP}(443)$ are drawn into two specific ranges both computed from the set of COASTlooc measurements. These two specific ranges are, respectively, the set of covariation bounds between $b_p(443)$ and $a_{chl}(443)$ and the set of covariation bounds between $b_p(443)$ and $a_{NAP}(443)$. We call set of

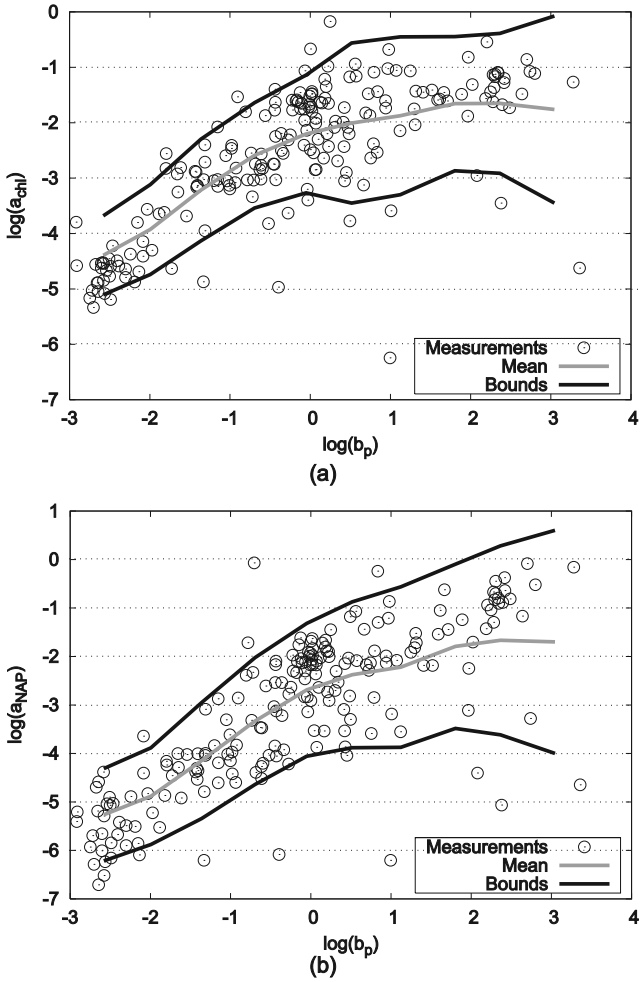

Figure 1. Examples of IOP covariations bounds at 443 nm: (a) between a_{chl} and b_p and (b) between a_{NAP} and b_p. The bounds were determined using the COASTlooc database. Units of the *x* axis and *y* axis are m^{-1}.

covariation bounds, the set of extreme values (lower and upper) giving the limits of an envelope in which data covariations are considered to be realistic. Accounting for the dispersion of COASTlooc measurements (Figure 1) where low values prevail, the two sets of covariation bounds are calculated using a logarithmic scaling. The upper and lower covariation bounds are centered on local mean values of the measurements. Note that the distance between the bounds depends on the local dispersions of the measurements. Because no covariation was found between a_{CDOM} and the other IOPs, each value of $a_{CDOM}(443)$ is uniformly drawn in a range of variation motivated by the related literature [*Babin et al.*, 2003b].

[13] After this first step, all of the synthetic IOP values at 443 nm are fully specified. The effective ranges of variation of the IOPs at 443 nm used in this study are reported in Table 1.

2.3.2. Second Step: Setting of the IOP Values at Other Wavelengths

[14] The spectral values of nonalgal particles and CDOM absorption are calculated using the bio-optical models developed by *Bricaud et al.* [1981] and *Babin et al.* [2003b] which report an exponential decrease of the absorption coefficient with wavelength (equation (2)). The spectral values of b_{chl} and b_{sed} are obtained using a power law function (equation (2)).

[15] For all λ in L we have

$$a_{NAP}(\lambda) = a_{NAP}(443)e^{-SaNAP(\lambda-443)} \quad , \tag{2a}$$

$$a_{CDOM}(\lambda) = a_{CDOM}(443)e^{-S_{aCDOM}(\lambda-443)} \quad , \tag{2b}$$

$$b_{chl}(\lambda) = b_{chl}(443)\left(\frac{\lambda}{443}\right)^{-S_{bchl}} , \tag{2c}$$

$$b_{sed}(\lambda) = b_{sed}(443)\left(\frac{\lambda}{443}\right)^{-S_{bsed}} . \tag{2d}$$

[16] The spectral slopes Sa_{NAP}, Sa_{CDOM}, Sb_{chl} and Sb_{sed} are uniformly drawn in specific ranges of variation [*Babin et al.*, 2003b] (Table 2). Figure 2a shows examples of synthetic b_{chl} spectra.

[17] The spectral slope of the scattering coefficient Sb_p is evaluated using a nonlinear regression, assuming that the spectral variation of b_p also follows a power law function. It is interesting to note that the derived values of Sb_p can be accurately approximated (the correlation coefficient is 0.98) using the following formulation:

$$Sb_p \approx \frac{b_{chl}(443)\cdot Sb_{chl} + b_{sed}(443)\cdot Sb_{sed}}{b_{chl}(443) + b_{sed}(443)}. \tag{3}$$

[18] The computation of a_{chl} at other wavelengths than 443 nm is somewhat more specific. Contrary to a_{NAP} and a_{CDOM} for which parametric models of their spectral variations are available, there is no standard law describing the spectral dependencies of a_{chl} in coastal waters. The spectral shape of a_{chl} is thus determined by averaging the three absorption spectra taken in the COASTlOOC database for which the measured value of $a_{chl}(443)$ is the nearest to the randomly selected value of $a_{chl}(443)$. The advantage of such way of doing is to account for the relationship that exists between the spectral shape of phytoplankton absorption coefficient and the magnitude of $a_{chl}(443)$. Note that this latter relationship is due to the packaging effect within phytoplankton cells. Therefore the data set is in accordance

Table 2. Range of Variations of the Spectral Slopes Used in the Optical Models[a]

Data	Mean	Sigma	Min	Max
S_{aNAP}	1.30E-02	1.16E-03	1.10E-02	1.50E-02
S_{aCDOM}	1.70E-02	1.70E-02	1.50E-02	1.90E-02
S_{bp}	1.00E+00	5.76E-01	5.23E-05	2.00E+00
S_{bchl}	1.01E+00	5.79E-01	5.23E-05	2.00E+00
S_{bsed}	1.00E+00	5.79E-01	9.38E-05	2.00E+00

[a]Sigma is the standard deviation. Read, for example, 1.30E-02 as 1.3×10^{-2}.

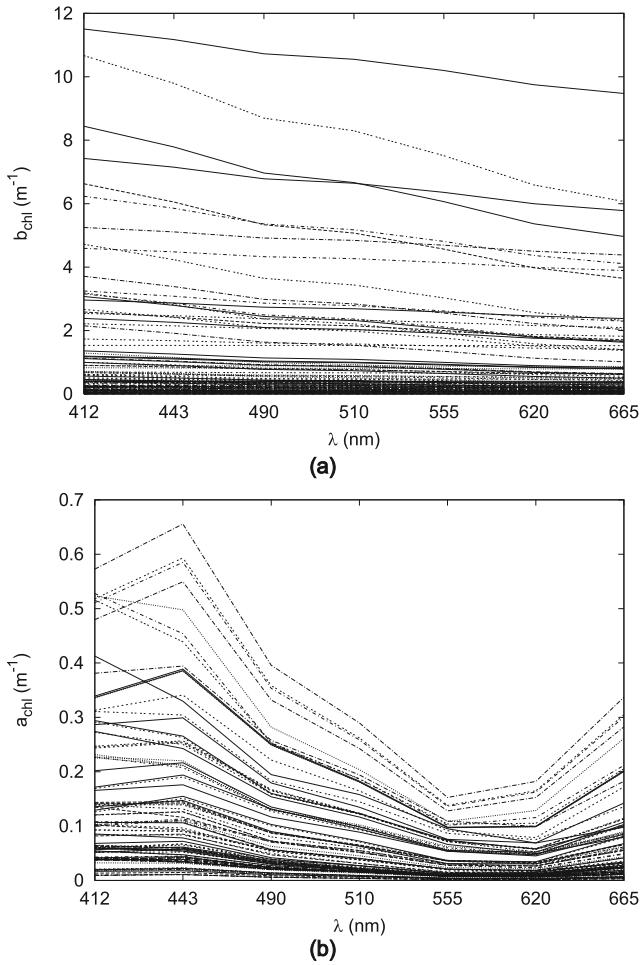

Figure 2. Examples of spectral dependencies of (a) b_{chl} and (b) a_{chl}, as computed using optical models and random spectral slopes. The *x* axis is the wavelength (nm), and the *y* axis is b_{chl} or a_{chl} (m^{-1}).

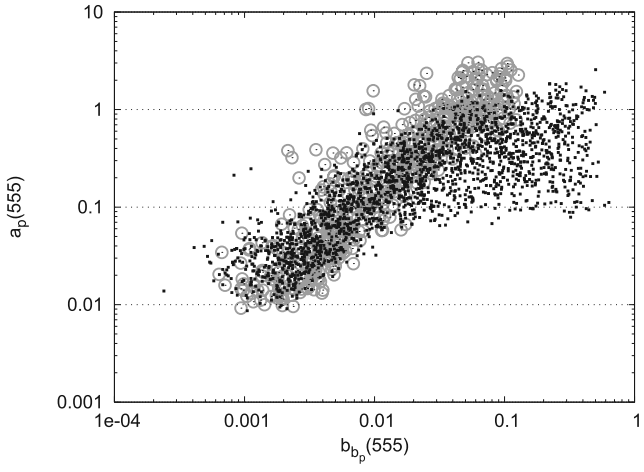

Figure 3. Covariations between the particulate backscattering coefficient b_{bp} (m^{-1}) and the particulate absorption coefficient a_p (m^{-1}) at 555 nm found in the IOCCG data set (gray circles) and in the DPC_{K2} data set (black dots).

Table 3. Correlation Between the IOPs at 443 nm Found in the Synthetic Data Set DPC_{K2}

IOP	a_{chl}	a_{nap}	a_{CDOM}	b_{chl}	b_{sed}	bb_{chl}	bb_{sed}
a_{chl}	100.00						
a_{nap}	9.06	100.00					
a_{CDOM}	0.03	0.01	100.00				
b_{chl}	12.09	17.76	0.01	100.00			
b_{sed}	11.60	16.85	0.00	11.04	100.00		
bb_{chl}	6.64	9.77	0.00	59.41	6.38	100.00	
bb_{sed}	10.61	15.40	0.00	9.91	90.56	10.80	100.00

with the biological properties of phytoplankton. The ratio, $A(\lambda)/A(443)$ where $A(\lambda)$ is the average absorption coefficient calculated over the three COASTlooc spectra, thus determines the spectral shape of a_{chl}. The spectral values of $a_{chl}(\lambda)$ are then calculated as follows:

$$a_{chl}(\lambda) = \frac{A(\lambda)}{A(443)} a_{chl}(443). \quad (4)$$

Figure 2b shows examples of synthetic $a_{chl}(443)$ spectra.

[19] Once the second step is achieved, a set of synthetic vectors of IOP_v is created and is introduced in the OSOA forward model to generate subsequent Rrs spectra. Finally the data set DPC_{K2} contains $N = 10000$ couples of (Rrs, IOP_v).

2.4. Data Set Analysis

[20] As we have already mentioned (section 2.2) our primarily objective when generating the DPC_{K2} data set is to produce realistic IOP spectra. Even though the data set creation process relies on truth measurements, a comparison of DPC_{K2} with the IOCCG data set [*IOCCG*, 2006] is performed to verify that DPC_{K2} is not overspecialized on the COASTlooc database. In particular the covariations that exist between IOPs are studied. Figure 3 shows that the covariation between b_{bp} and a_p at 555 nm found in both data set are in good agreement, thus demonstrating the consistency of the DPC_{K2} data set with the literature.

[21] In order to gain in efficiency in inverse problem solving, the different variables within the data set should be *iid*, which can obviously not be obtained when dealing with the inverse problem of ocean color in coastal waters. Nevertheless, it is important to gain insight into the relationships that may exist within the data set between the different variables. It is well known that two variables are statistically independent if they are uncorrelated. To study the correlations within DPC_{K2}, the percent of variance each IOP at 443 nm have in common with the others is calculated using the coefficient of determination R^2 (Table 3). Table 3 shows that most of the variables exhibit very low correlation with the others, which means that the statistical constraint is fairly verified. Because the backscattering coefficients are fractions of the scattering coefficients, it is not surprising to observe high coefficients of determination between b_{bchl} and b_{chl} as well as between b_{bsed} and $b_{sed.}$ Since the Rrs spectra within DPC_{K2} were generated using realistic inputs, their spectral shapes are not random and so the Rrs values at different wavelengths cannot be independent. Thus we have also investigated how the Rrs values are correlated with

Table 4. Correlation Between the Remote Sensing Reflectance Rrs at Different Wavelengths Found in the Synthetic Data Set DPC_{K2}

λ	412	443	490	510	555	620	665
412	100.00						
443	96.41	100.00					
490	88.24	96.83	100.00				
510	84.37	93.97	99.40	100.00			
555	73.96	85.16	94.59	97.49	100.00		
620	53.85	66.78	78.72	83.34	91.07	100.00	
665	46.63	60.31	72.40	76.72	84.29	98.06	100.00

regard to the wavelength. The coefficients of determination R^2 reported Table 4 indicate that the Rrs values are highly correlated when considering two neighboring wavelengths. Therefore an inverse model will not necessarily gain in efficiency from the Rrs measured at two neighboring wavelengths. Most of the useful information is provided by one of these Rrs only. At the opposite, the correlation between Rrs(412) and Rrs(665) is low, thus meaning that the information brought by the full Rrs spectra will be considered as relevant by an inverse algorithm.

[22] Some of the data distributions within DPC_{K2} are, by construction, well known, in particular the values at 443 nm. In this latter case, two different distributions were used depending on the optical parameter that is drawn. The standard Uniform distribution was used to draw the $a_{CDOM}(443)$ values whose empirical distribution is still poorly documented. The uniform distribution, noted U(a, b), is characterized by the following probability density function:

$$\phi_U(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases} . \quad (5)$$

[23] A log-uniform distribution, which corresponds to an uniform draw in a logarithmic scale, was used to set the values of $a_{chl}(443)$, $a_{NAP}(443)$ and $b_p(443)$.The log-uniform distribution, noted LU(a, b) is described with the following probability density function:

$$\phi_{LU}(x) = \begin{cases} \dfrac{1}{\log b - \log a} \times \dfrac{1}{x} & a < x < b \\ 0 & \text{otherwise} \end{cases} . \quad (6)$$

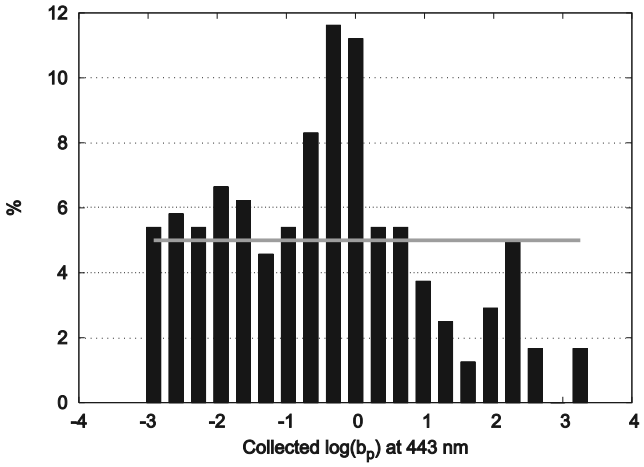

Figure 4. Distribution of the $\log(b_p)$ values (m^{-1}) taken from the COASTlooc database (in black) and the theoretical uniform distribution (in gray) at 443 nm.

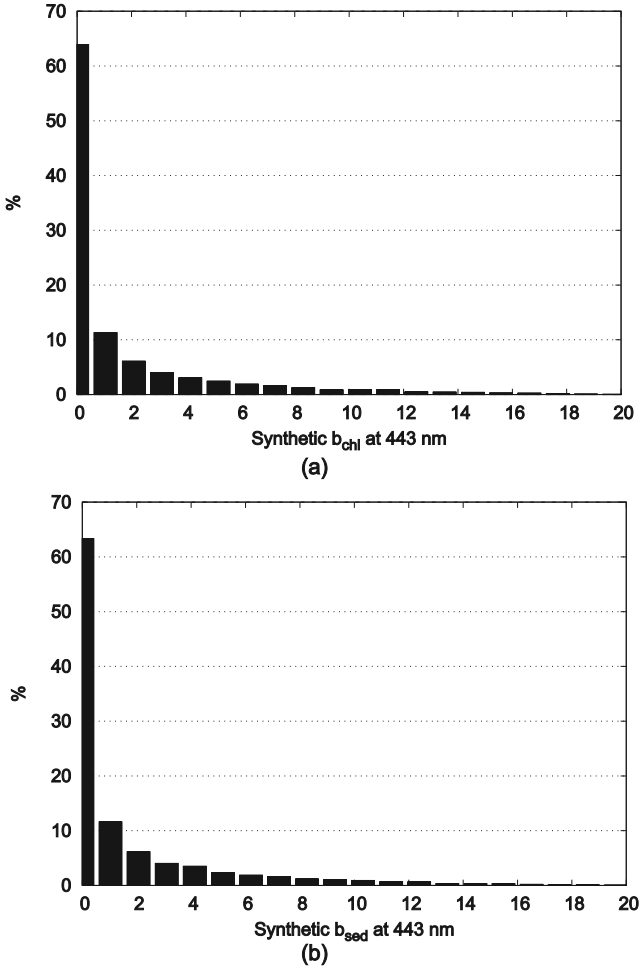

Figure 5. Distributions of synthetic (a) b_{chl} and (b) b_{sed} values at 443 nm within DPC_{K2}. Units of the x axis are m^{-1}.

[24] Despite it is often admitted that the log-normal distribution is suitable to model the ocean bio-optical variability [*Campbell*, 1995], this original log-uniform distribution (equation (6)) was selected because the distribution of the weak values of a_{chl}, a_{NAP} and b_p in the COASTlooc database compares reasonably well with the Uniform distribution in the log-space (Figure 4). The distributions of $b_{chl}(443)$ and $b_{sed}(443)$ are much more complex (see Figures 5a and 5b). They depend on the distribution of $b_p(443)$ and correspond to the product of both distributions LU($a = \min(b_p)$, $b = \max(b_p)$) and U($a = 0$, $b = 1$). The subsequent probability density functions are reported hereafter:

$$\phi_{LU.U}(x) \approx \begin{cases} \dfrac{1}{\log b - \log a} \times \left(\dfrac{-1}{b} + \dfrac{1}{x}\right) & a < x < b \\ 0 & \text{otherwise} \end{cases} . \quad (7)$$

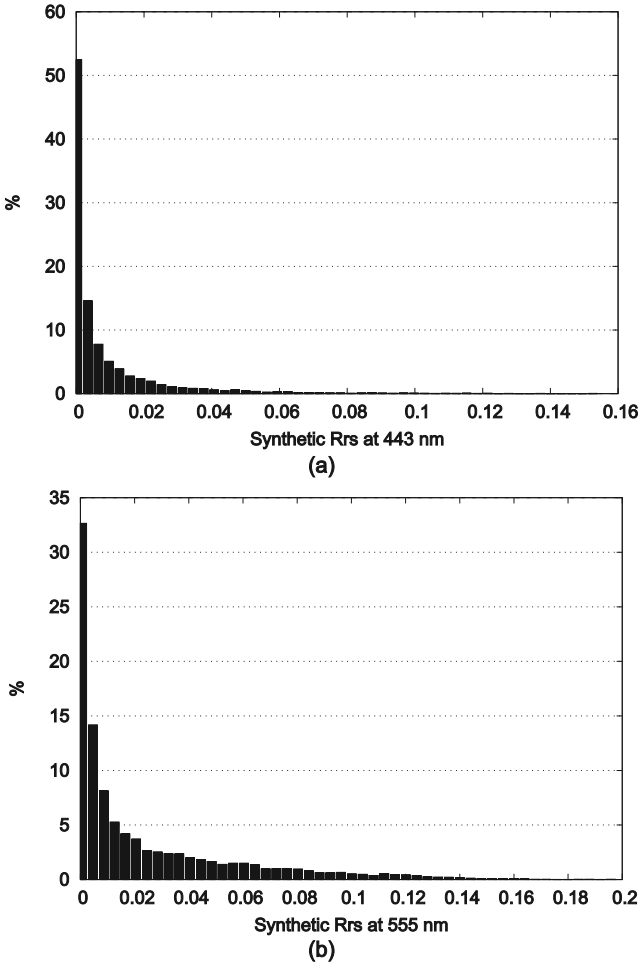

Figure 6. Distributions of the synthetic Rrs (sr^{-1}) (a) at 443 and (b) 555 nm within DPC_{K2}.

[25] Similarly, since $b_{bchl}(443)$ and $b_{bsed}(443)$ are fractions of respectively $b_{chl}(443)$ and $b_{sed}(443)$, they report similar distributions as b_{chl} and b_{sed}. It is important to note that similar shapes of distributions are observed at other wavelengths than 443 nm.

[26] Since the Rrs spectra are the outputs of the forward model, the distributions of the Rrs values are a priori unknown, for all the wavelengths of L. We found that the distribution of the Rrs values at 443 nm (Figure 6a) corresponds, in first approximation, to a lognormal distribution LN(μ, σ) having the following probability density function:

$$\phi_{LN}(x) = \begin{cases} \dfrac{e^{-\frac{1}{2}\left(\frac{\log x - T}{\sigma}\right)^2}}{x\sigma\sqrt{2\pi}} & 0 < x \\ 0 & \text{otherwise} \end{cases}. \quad (8)$$

[27] This hypothesis has been verified using a nonlinear regression method to fit the parameters T and σ. We found that $T \approx 7.10$, $\sigma \approx 6.92$. With such settings, the theoretical distributions of Rrs values at 443 nm showed very good agreements ($r = 0.98$) with the synthetic distribution plotted in Figure 6a. Similar agreements have been obtained at other wavelengths, see for example Figure 6b, but with other settings of the parameters T and σ. This indicates that even if the Rrs values are not rigorously identically distributed in the data set, they report comparable distributions each other and so the best possible performances of inverse models might be expected.

3. Ambiguity and Ill-Posedness of the Inverse Problem

[28] The inverse problem of ocean color in coastal waters is very complex mainly because this is an ill-posed problem with many ambiguities. Intuitively we understand that the ambiguous samples of a data set are that for which several combinations of IOP values correspond to one "unique" Rrs spectrum. As an example, Figure 7 shows the variations in the Rrs at 555 nm as a function of the total absorption, scattering and backscattering coefficient. It is clearly observed that identical values of Rrs are obtained for different combinations of total IOPs. It is obvious that this situation is problematic for inverse models since they may not be able to discriminate several similar Rrs spectra and so to output the desired IOPs. Therefore one of the major issues of the inverse problem of ocean color in coastal waters consists in a deep characterization of the ambiguities.

3.1. Relationship Between Ambiguity and Ill-Posedness

[29] According to *Hadamard* [1902] a well-posed inverse problem must satisfy three criteria: (1) A solution always exists (for any data), (2) the relationship between inputs and outputs is continuous and (3) the solution is unique.

[30] Regarding the inverse problem of ocean color in coastal waters, the first criterion is always guaranteed because the forward model is based on the well-posed Radiative Transfer Equation (RTE) and thus, each synthetic Rrs spectrum of the data set is associated with an IOP spectrum. To check whether the second criterion is satisfied, it is rigorously necessary to demonstrate that the relationship between the Rrs and the IOPs is continuous. Practically, such a demonstration is intractable. However, it is mathematically well known that a stable problem is systematically continuous. Note that a stable inverse problem can be

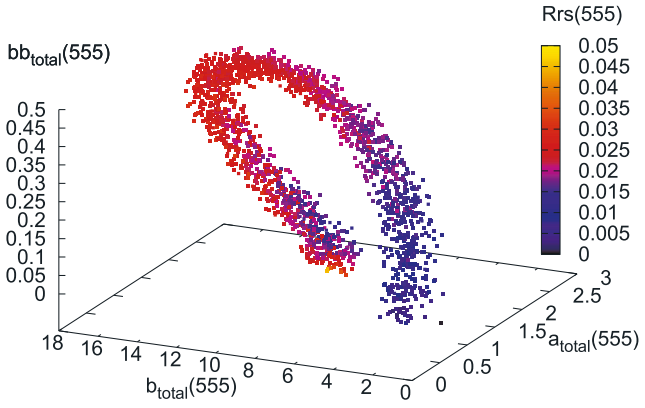

Figure 7. Example of ambiguities in the inverse problem of ocean color in coastal waters. To a unique Rrs(555) value may correspond several triplets ($a_{tot}(555)$, $b_{tot}(555)$, $bb_{tot}(555)$). The units are: sr^{-1} for Rrs and m^{-1} for a_{tot}, b_{tot} and b_{btot}.

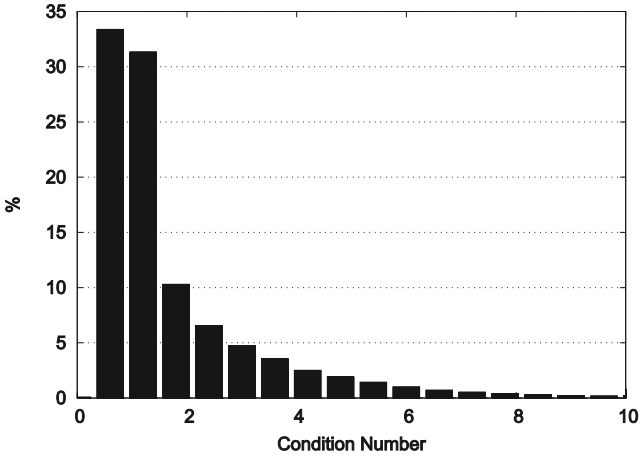

Figure 8. Distribution of the condition number within DPC_{K2}.

defined such that weak variations of the inputs (i.e., Rrs spectra) induce weak variations of the outputs (i.e., IOP_v). Therefore the stability of the inverse problem of ocean color is investigated to test the second criterion of well-posedness. The so-called condition number (*CN*) (equation (9)) is then computed to establish the stability of the solutions of the inverse problem. Note that *CN* values that are much greater than 1 (typically $CN > 10$) indicate an unstable problem. The equation 9 gives a good empirical approximation of *CN* for a couple $(Rrs_1, IOP_{v1}) \in DPC_{K2}$.

$$CN(Rrs_1, IOP_{v1}) \approx \max_{\lambda \in \mathrm{L}} \left(\frac{\left| \frac{IOP_{v1}(\lambda) - \overline{IOP}(\lambda)}{IOP_{v1}(\lambda)} \right|}{\left| \frac{Rrs_1(\lambda) - \overline{Rrs}(\lambda)}{Rrs_1(\lambda)} \right|} \right). \tag{9}$$

[31] In equation (9), $\overline{IOP}$ and $\overline{Rrs}$ are the mean of IOP_v sets and Rrs spectra respectively. In Figure 8, the distribution of *CN* for all the samples is reported. We see that for the vast majority of the data, the *CN* values are smaller than 10. The inverse problem is most of the time stable and thus, most of the time continuous. Hence only the presence of ambiguities and so only the "nonuniqueness" of the solution, which is the third criterion, is an obstacle to the well-posedness of the inverse problem of ocean color in coastal waters. These results show that a correct handling of the ambiguities should significantly improve the performances of any inverse model since the problem will tend to be well-posed.

3.2. Formal Definitions of Ambiguity

[32] In order to perform a thorough analysis of the ambiguities present within the data set, a formalism defining the ambiguities is required.

3.2.1. Spectral Distance

[33] An objective measure of similarity is needed to determine "uniqueness" of several Rrs spectra. For that purpose, we define the spectral distance δ_{Rrs} between two Rrs spectra Rrs_1 and Rrs_2,

$$\delta_{Rrs}(Rrs_1, Rrs_2) \approx \max_{\lambda \in \mathrm{L}} \left(\left| \frac{Rrs_1(\lambda) - Rrs_2(\lambda)}{\frac{1}{2}(Rrs_1(\lambda) + Rrs_2(\lambda))} \right| \right). \tag{10}$$

[34] Equation (10) corresponds to the maximum of the absolute difference between Rrs_1 and Rrs_2 normalized to the average over these two spectra and computed for all wavelengths in L. We note that this measure is symmetric and varies, in this study, in the range [0, 2]. Similarly, we define the spectral distance δ_{IOP} between two sets of IOP vector IOP_{v1} and IOP_{v2},

$$\delta_{IOP}(IOP_{v1}, IOP_{v2}) \approx \max_{p \in IOP} \cdot \left(\max_{\lambda \in \mathrm{L}} \left(\left| \frac{IOP_{v1,p}(\lambda) - IOP_{v2,p}(\lambda)}{\frac{1}{2}\left(IOP_{v1,p}(\lambda) + IOP_{v2,p}(\lambda)\right)} \right| \right) \right). \tag{11}$$

[35] In equation (11), p is one element of the IOP vector (i.e., the absorption, scattering or backscattering coefficient of one of the seawater component). It is also possible to compute a spectral distance for each element of the IOP vector. This means that the maximum value in equation 11 is only searched over the set of wavelengths and not over the entire set of IOP_v. The spectral distance defined for each element of the IOP vector is referred using the notation of this element. As an example, the b_{chl} spectral distance between IOP_{v1} and IOP_{v2} is referred to as $\delta_{\mathrm{bchl}}(\mathrm{IOP_{v1}},\mathrm{IOP_{v2}})$.

3.2.2. Spectral Neighborhood

[36] We define the set of spectral neighborhood of one Rrs spectrum Rrs_1 within the DPC_{K2} data set as

$$N_{DPCK2,\varepsilon}(Rrs_1) = \{Rrs_2 | (Rrs_2, IOP_{v2}) \in DPC_{K2}, \delta_{Rrs}(Rrs_1, Rrs_2) < \varepsilon\}, \tag{12}$$

with ε a threshold value, typically the uncertainty in the Rrs measurements. In this paper ε is fixed to 5%, which is consistent with the accuracy provided by the current commercially available instruments. For clarity, DPC_{K2} and ε will be further omitted in the notation. Note that the definition given in equation (12) and other forthcoming equations could be used for any other data set than DPC_{K2}. By construction all the elements of the set N (Rrs_1) are said to be similar as Rrs_1 and so they will be considered as unique by any inversion algorithm.

3.2.3. Inversion Set

[37] We define the inversion set of a given Rrs spectrum Rrs_1 as

$$\mathrm{S}(Rrs_1) = \{IOP_{v2} | (Rrs_2, IOP_{v2}) \in DPC_{K2}, Rrs_2 \in \mathrm{N}(Rrs_1)\}. \tag{13}$$

[38] The inversion set S (Rrs_1) contains all the IOP vectors IOP_v for which the forward OSOA model outputs Rrs spectra that are considered similar as Rrs_1 accordingly to the definition of the spectral neighborhood of Rrs_1 (equation (12)).

3.2.4. Spectral Ambiguity Test

[39] A spectrum Rrs_1 within the data set is considered as ambiguous if the spectral distance between the IOP vectors that belong to the inversion set S(Rrs_1) is significant

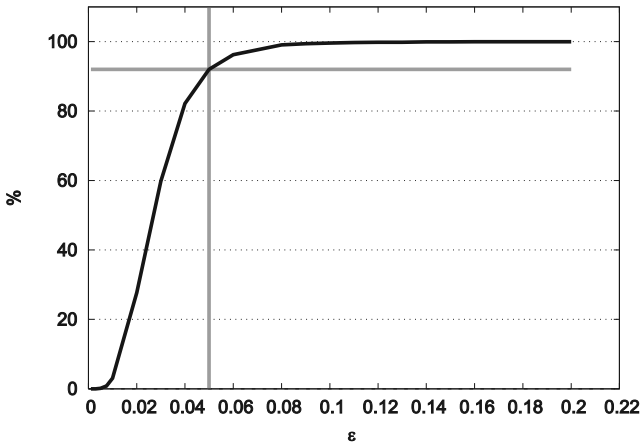

Figure 9. Percentage of Rrs spectra having at least one neighbor as a function of ε (in %).

(typically greater than ε). To test whether Rrs_1 is ambiguous, the following function is used:

$$TA_{IOP}(Rrs_1) = \begin{cases} true & \exists IOP_{v2}, IOP_{v3} \in S(Rrs_1) | \delta_{\mathrm{IOP}}(IOP_{v2}, IOP_{v3}) > \varepsilon \\ false & \text{otherwise} \end{cases}. \quad (14)$$

[40] Therefore Rrs_1 is considered as an ambiguous spectrum when the predicate $\mathrm{TA_{IOP}}(Rrs_1)$, which is called the spectral ambiguity of Rrs_1, is true. It is also possible to test for the spectral ambiguity of Rrs_1 relatively to each of the elements of the IOP vector. As an example, the test for the b_{chl} spectral ambiguity is referred as $\mathrm{TA}_{b_{chl}}(Rrs_1)$.

3.3. Analysis of the Ambiguities Within the Data Set

3.3.1. Relevance of the Spectral Ambiguity Test

[41] The practical evaluation of ambiguities within any data set is not straightforward. The spectral ambiguity test $\mathrm{TA_{IOP}}(Rrs_1)$ related to a given spectrum Rrs_1 is trustworthy only if the inversion set S (Rrs_1), and thus the spectral neighborhood $\mathrm{N}(Rrs_1)$, are not empty sets. The number of elements in the sets of spectral neighborhood, hereafter referred to as # ($\mathrm{N}(Rrs_1)$), depends on the value of ε. The percentage of the trustworthy samples of DPC_{K2}, i.e., the samples for which $\#(\mathrm{N}(Rrs_1)) \geq 2$, is plotted as a function of ε in Figure 9. When ε is greater than 15%, each sample of DPC_{K2} has at least one spectral neighbor. When ε is equal to 5%, which is the value adopted in this study, the proportion of trustworthy samples is very high, around 92%, thus making reliable our analysis of the ambiguities within the DPC_{K2} data set.

[42] We define the ambiguity rate $\mathrm{R_{IOP}}$ as the proportion of ambiguous samples within DPC_{K2},

$$\mathrm{R_{IOP}} = \frac{1}{N} \sum_{(Rrs_1, IOP_{v1}) \in DPC_{K2}} \begin{cases} 1 & \text{if} \quad \mathrm{TA_{IOP}}(Rrs_1) \text{is true} \\ 0 & \text{otherwise} \end{cases}. \quad (15)$$

[43] In equation (15), N is the number of element of DPC_{K2}. The ambiguity rate could be very helpful when trying to compare the ambiguities within different data sets and thus to evaluate the impact of strategies to tackle the ambiguity problem, which will be discussed later. The ambiguity rates within the data set DPC_{K2} is calculated for each element of the IOP vector and results are reported in Table 5. The ambiguity rates show fairly similar values (around 92%). Therefore almost all the trustworthy samples of DPC_{K2} have been considered to be ambiguous.

3.3.2. Ambiguity Distance

[44] Because the spectral distance between the IOP vectors corresponding to ambiguous Rrs samples may be highly variable (i.e., all the samples are not equally ambiguous), it is interesting to evaluate the maximum difference that exists between the IOP vectors belonging to the inversion sets. In this objective, we define the ambiguity distance Δ_{IOP} of a given Rrs spectrum Rrs_1 as follows:

$$\Delta_{IOP}(Rrs_1) = \max\{\delta_{IOP}(IOP_{v1}, IOP_{v2}) | IOP_{v1}, IOP_{v2} \in S(Rrs_1)\}. \quad (16)$$

We note that it is also possible to compute an ambiguity distance for each element of the IOP vector.

[45] Figure 10a shows three similar Rrs spectra. One of them (solid lines) is taken as the reference spectrum Rrs_1 and the two other spectra were selected within the spectral neighborhood set $\mathrm{N}(Rrs_1)$. The corresponding b_{chl} spectra that belong to the inversion set $\mathrm{S}(Rrs_1)$ are plotted in Figure 10b. It is observed that highly different values of b_{chl} can lead to identical Rrs values. In this example, the maximum bias between the b_{chl} spectra (i.e., the ambiguity distance $\Delta_{bchl}(Rrs_1)$) reaches 172.8%. Figure 10 clearly illustrates the necessity of quantifying the ambiguity distance for inverse modeling purpose, as it will be discussed later.

[46] We have computed the mean ambiguity distance, noted $\bar{\Delta}_{\mathrm{IOP}}$, over the 10000 samples of the DPC_{K2} data set. The mean ambiguity distance has to be understood as the average discrepancy between two sets of IOP_v that lead to similar Rrs spectra. The results are reported in Table 5. The mean ambiguity distance found for each IOP shows high values. However, a strong variability is observed depending on the inherent optical property that is studied. As expected from its definition, $\bar{\Delta}_{\mathrm{IOP}}$ exhibits the highest value. We observe also that despite a considerable difference between the ranges of variation of b_{chl} and b_{bchl} (Table 1), their mean ambiguity distance are almost equal. A similar remark can

Table 5. Ambiguity Rate $\mathrm{R}(DPC_{K2})$ and Mean Ambiguity Distance $\bar{\Delta}$ (DPC_{K2}) Within DPC_{K2}

Data	$R(DPC_{K2})$, %	$\bar{\Delta}$ (DPC_{K2}), %
IOP	92.00	165.76
a_{chl}	91.63	114.22
a_{nap}	91.96	129.50
a_{CDOM}	91.38	86.30
b_{chl}	92.00	151.58
b_{sed}	91.97	124.06
bb_{chl}	91.97	153.04
bb_{sed}	91.92	119.11

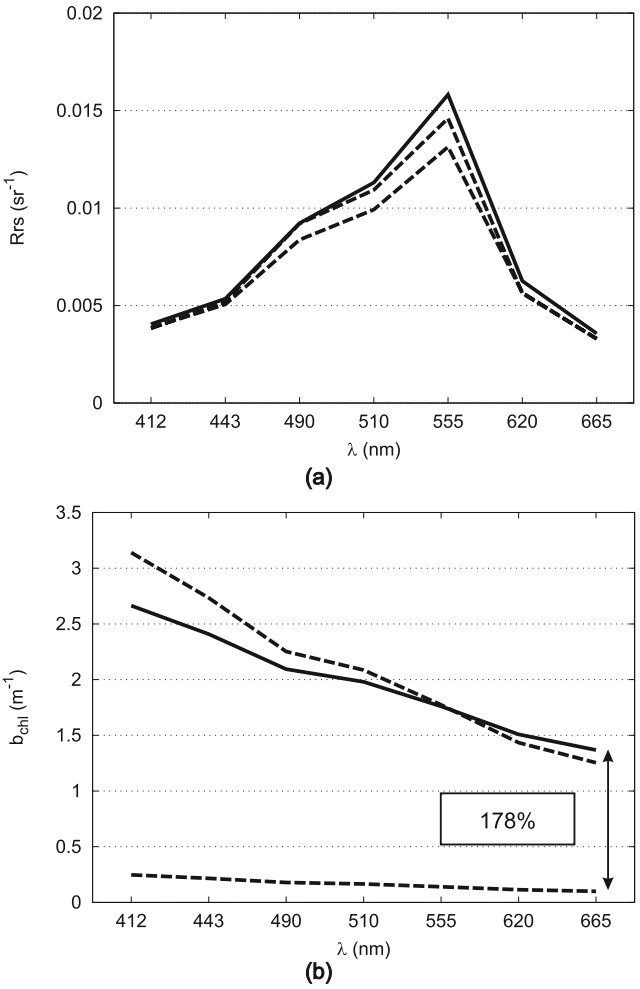

Figure 10. (a) a set of spectral neighborhood $N(Rrs_1)$ containing three Rrs spectra (sr^{-1}) with Rrs_1 as the solid line and (b) the corresponding inversion set $S(Rrs_1)$ containing three b_{chl} spectra (m^{-1}) with IOP_1 as the solid line. The *x* axis is the wavelength (nm).

be made with regard to b_{sed} and b_{bsed}. This is not surprising since the backscattering coefficient is strongly correlated with the scattering coefficient. The effects of the uncertainties in the particulate phase function on the ambiguity were studied through the variations in the backscattering ratio (i.e., b_{bp}/b_p). The ambiguity distance $\Delta b_{bp}/b_p$ was related to Δ_{IOP}. A good positive correlation was obtained ($R^2 = 0.80$). This means that most of the ambiguity found in the IOP (nearly 80%) can be ascribed to uncertainties in the phase function. This result is physically very consistent. The *Rrs* is highly sensitive to the phase function of the particles (the directional effects of the particles are contained in the parameter "g" in equation (1)) [*Mobley et al.*, 2002]. Therefore, given similar Rrs spectra, a significant variation of the phase function (i.e., variations of "g" in equation (1)) induces an increase of the number of the combinations between the IOPs (for example a_{tot}, b_{btot} in equation (1)), thus leading to an increase in the ambiguity of the IOPs parameters.

[47] We also computed the distributions of the ambiguity distances for all the samples of DPC_{K2}. Two examples are presented Figure 11. The Δa_{CDOM} distribution is nearly uniform thus meaning that there is an equivalent amount of very ambiguous or slightly ambiguous samples. On the contrary the ambiguity distance Δ_{bchl} shows a weak variance and is clearly not equally distributed. Note that the first class of the histograms in Figure 11 corresponds to the unambiguous samples (with $\Delta \leq 5\%$) and holds around 8% of the samples of DPC_{K2}, which is consistent with the value of 92% found for the ambiguity rate (Table 5).

[48] Since the ambiguity distance Δ_{IOP} is not equally distributed within DPC_{K2}, it can be very useful for a final user to gain a priori knowledge about Δ_{IOP} before the inversion of a given Rrs spectrum. Therefore we propose to classify the Rrs spectra of the data set according to their ambiguity distance Δ_{IOP}. We first define the class *U* for unambiguous samples ($0\% \leq \Delta_{IOP} < 5\%$), the class *M* for fairly ambiguous samples ($5\% \leq \Delta_{IOP} < 100\%$) and class *H* for highly ambiguous samples ($100\% \leq \Delta_{IOP}$). Such a classification allows the computation of the probability $P(Rrs_1 \in C)$ that a given Rrs spectrum Rrs_1 belongs to a given class C (i.e., classes U, M or H). For example, we found that $P(Rrs_1 \in U) \approx 8\%$, $P(Rrs_1 \in M) \approx 52\%$ and $P(Rrs_1 \in H) \approx 40\%$ regarding Δa_{CDOM} whereas $P(Rrs_1 \in U) \approx 8\%$, $P(Rrs_1 \in M) \approx 8\%$ and $P(Rrs_1 \in H) \approx 84\%$ regarding Δ_{bchl}.

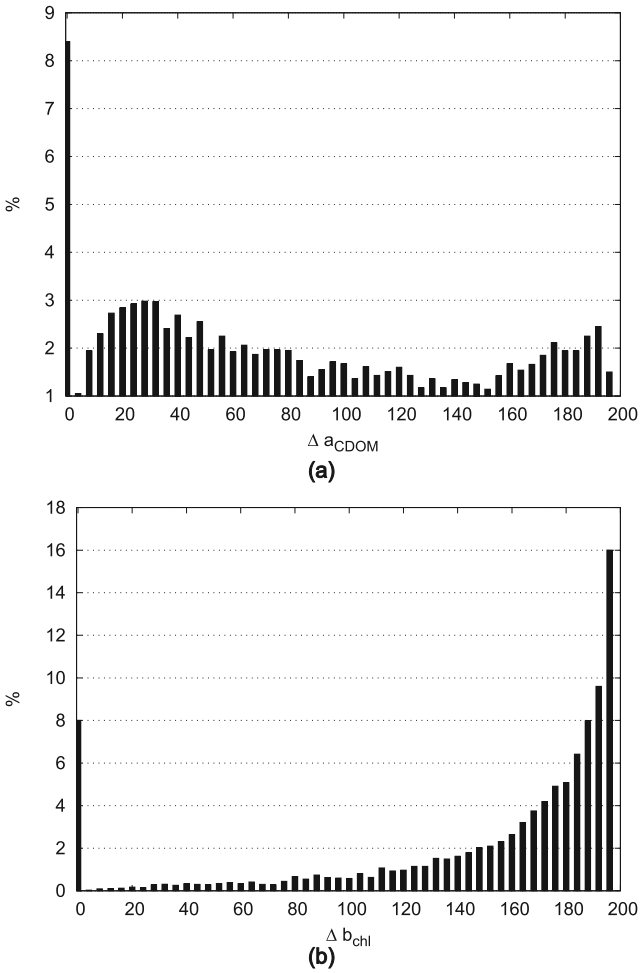

Figure 11. Distributions of ambiguity distances (dimensionless) within DPC_{K2}: (a) Δa_{CDOM} and (b) Δb_{chl}.

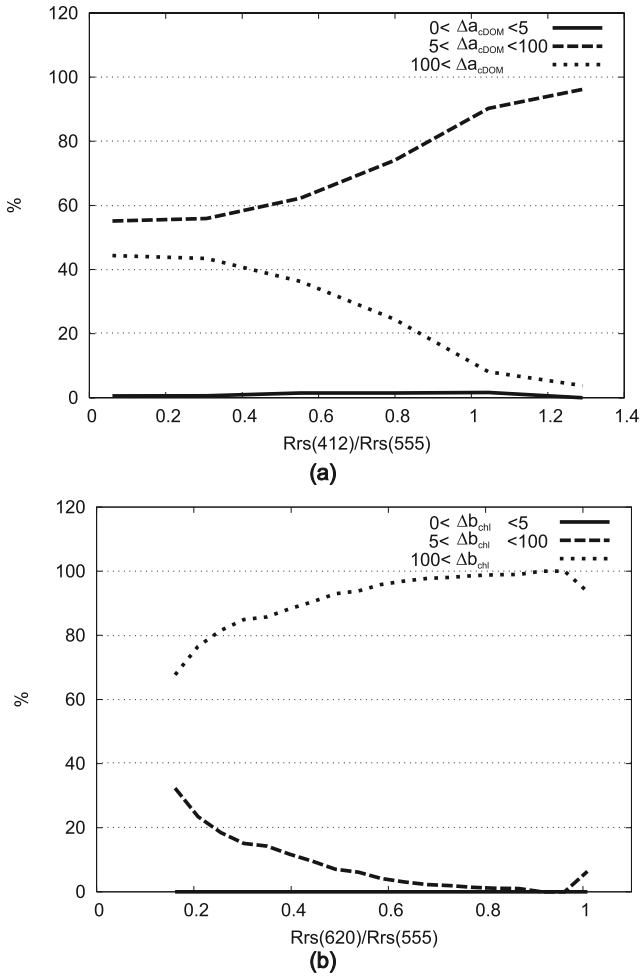

Figure 12. Classifications of ambiguity distances (a) Δa_{CDOM} as a function of the ratio Rrs(412)/Rrs(555) and (b) Δb_{chl} as a function of the ratio Rrs(620)/Rrs(555).

[49] In Figure 12, we plotted these probabilities according to band reflectance ratio. The spectral ratio Rrs_1 (412)/Rrs_1 (555) and Rrs_1 (620)/Rrs_1 (555) are used for the a_{CDOM} and the b_{chl} coefficient respectively. Actually, Figure 12 shows the conditional probabilities $P(Rrs_1 \in C|Rrs_1$ (412)/Rrs_1 (555)) and $P(Rrs_1 \in C|Rrs_1(620)/Rrs_1$ (555)). When performing the inversion of a given Rrs spectrum, we know a priori that the spectral distance between a_{CDOM} spectra in the inversion set is highly sensitive to the reflectance ratio $Rrs_1(412)/Rrs_1(555)$. In particular, the different a_{CDOM} spectra to be retrieved will be probably not so dissimilar from each others when the values of the ratio $Rrs_1(412)/Rrs_1(555)$ are high. On the contrary, the ambiguity distance between $\Delta_{b_{chl}}$ spectra is systematically high and fairly independent on the reflectance ratio $Rrs_1(620)/Rrs_1$ (555). This means that the retrieval of b_{chl} spectra from the reflectance is expected to be more difficult than that of a_{CDOM} spectra.

4. Impact of Ambiguities on Inverse Modeling

[50] We investigate now the influence of ambiguities on the performances of inversion algorithms. Most of the inversion algorithm uses least squares minimization to converge toward good solutions. It has been shown that, in presence of nonuniqueness of the solution, these types of inverse methods tend to output the average of the desired parameters [*Bishop*, 1994]. Thus these models approximate the conditional mean $E(y|x)$, where y is the desired parameter and x is the input of the inverse model. When addressing the inversion of the ambiguous ocean color problem, the conditional mean $E(\mathrm{IOP}(\lambda)|\mathrm{Rrs}(\lambda))$ will be approximated by most of the inversion algorithm. This situation is often dramatically problematic since the answer $E(\mathrm{IOP}(\lambda)|\mathrm{Rrs}(\lambda))$ can be, in the better case, only one element of the set of all the possible solutions but worst it can also be outside this set (if this latter is not a convex set). Thus, even when inverse models are based on realistic data, they could report outputs far from physically plausible solutions.

[51] To illustrate this, we present in Figure 13 the inversion of the spectral neighborhood $N(Rrs_1)$ which was previously reported in Figure 10. A Multi-Layer Perceptron (MLP), also known as a Multi-Layer Feed-Forward Artificial Neural Network, which was trained on the DPC_{K2} data set was used to perform the inversion. We can see the three b_{chl} spectra of the inversion set $S_{b_{chl}}(Rrs_1)$ that should be retrieved and the three b_{chl} spectra that are actually predicted by the MLP. The three retrieved b_{chl} spectra from the inverse model are very similar each other but far from the desired spectra. This is not surprising since these retrieved b_{chl} spectra also correspond roughly to the average spectrum over the set of the three desired b_{chl} spectra accordingly to theory [*Bishop*, 1994].

[52] Thus, in the general case, the performance of an inverse model when trying to retrieve an inversion set $S(Rrs_1)$ from a neighborhood $N(Rrs_1)$ depends on the dispersion of the spectra within the set $S(Rrs_1)$. We note D_{Rrs_1} the restricted data set that is composed only of the n pairs (Rrs,IOP_v), where Rrs belong to the set $N(Rrs_1)$ and IOP_v belongs to the set $S(Rrs_1)$. We note also $E(S(Rrs_1))$ and $\sigma(S(Rrs_1))$ respectively the mean spectrum and the standard deviation of $S(Rrs_1)$. We note m the best inverse model that can be constructed on D_{Rrs_1} such that $m(Rrs_1) = E(S(Rrs_1))$. If the error function used to ensure the convergence of the algorithm is for example the root mean square error RMSE, then we have

$$\begin{aligned}
\mathrm{RMSE}[m, D_{Rrs1}] &= \sqrt{\frac{1}{n}\sum_{(Rrs_2, IOP_{v2})\in D_{Rrs_1}} (IOP_{v2} - m(Rrs_2))^2} \\
&\approx \sqrt{\frac{1}{n}\sum_{(Rrs_2, IOP_{v2})\in D_{Rrs_1}} (IOP_{v2} - m(Rrs_1))^2} \\
&\approx \sqrt{\frac{1}{n}\sum_{(Rrs_2, IOP_{v2})\in D_{Rrs_1}} (IOP_{v2} - E(\mathrm{S}(Rrs_1)))^2} \\
&\approx \sqrt{\frac{1}{n}\sum_{(Rrs_2, IOP_{v2})\in D_{Rrs_1}} (IOP_{v2} - E(IOP_{v2}))^2} \\
&= \sigma(\mathrm{S}(Rrs_1)). \qquad (17)
\end{aligned}$$

[53] Therefore, on the basis of equation (17), any inversion method using the standard RMSE to ensure the

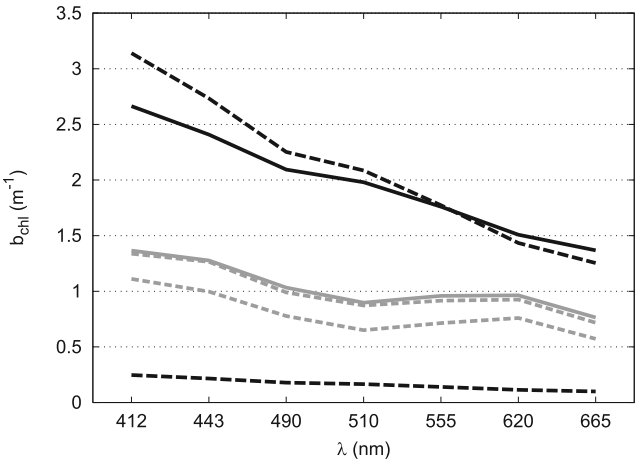

Figure 13. Comparison between an expected b_{chl} spectrum (black lines) and b_{chl} spectra (gray lines) predicted using a Multi-Layer Perceptron. The x axis is the wavelength (nm^{-1}), and the y axis is b_{chl} (m^{-1}).

convergence of the algorithm within the restricted data set D_{Rrs_1} cannot report an error smaller than $\sigma(\mathrm{S}(Rrs_1))$. Thus the dispersion $\sigma(\mathrm{S}(Rrs_1))$ of the IOP spectra within an inversion set is called the minimum inverse problem error (MIPE). Note that the unit of the MIPE is m^{-1} when retrieving inherent optical properties of marine particles.

[54] Since the ambiguities are not equally distributed within the synthetic data set DPC_{K2} and since the inversion sets hold different numbers of spectra with different distributions of their IOP values, it is only possible to provide an estimate of the MIPE for the entire data set DPC_{K2},

$$\mathrm{MIPE}(DPC_{K2}) = \mathrm{RMSE}[m, DPC_{K2}] \approx \sqrt{\frac{1}{N}\sum_{(Rrs_1, IOP_{v1})\in D_{K2}} \frac{1}{n}\sigma(\mathrm{S}(Rrs_1))^2}. \quad (18)$$

[55] Figure 14 shows the spectral variations of the estimated MIPE computed for a_{NAP} and for b_{chl}. We also plotted the spectral variations of the training RMSE corresponding to the inversion of Rrs using a MLP inversion algorithm. As expected, the MIPE and the training error made by the MLP algorithm shows very similar spectral shape for both IOPs. It is interesting to note that the magnitude of the MIPE and the MLP errors made on a_{NAP} and b_{chl} retrieval significantly differ, thus highlighting that the MLP algorithm may not be the best one. The previous results indicate that the MIPE gives a reliable approximation of the performance of an inverse model on DPC_{K2} and more generally on any ambiguous data set. Hence the MIPE can also be used to identify the most ambiguous regions within a data set. In order to build two-dimensional maps of MIPE, we divided the whole data set DPC_{K2} into a regular grid of 20 × 20 subsets according either to the Rrs or IOP values. Note that the number of elements differs within subsets, since Rrs or IOP values are not equally distributed, and some of them can even be empty. For each subset that hold a significant number of items (>10), a local MIPE has been computed. Figure 15 shows maps of MIPE regarding the total absorption coefficient a_{tot} (Figure 15a) and the total backscattering coefficient b_{btot} (Figure 15b). Each of these IOP is plotted as a function of Rrs. The wavelengths 443 and 620 nm are used for a_{tot} and b_{btot} respectively. The MIPE obtained regarding a_{tot} shows the highest values when the Rrs(443) is low and when the absorption coefficient is high. It is observed that the minimum error in the a_{tot} retrieval is within the range [±0.1 m^{-1}, ±0.2 m^{-1}] when the remote sensing reflectance is smaller than 0.02 sr^{-1}. When the particles are strongly absorbing, the Rrs values are so low that it is very difficult to clearly identify the contribution to Rrs from a_{tot}. As a result, the determination of a_{tot} is made more difficult. A similar reason led *O'Reilly et al.* [1988] to propose the use of the reflectance at 490 nm or 510 nm instead of the reflectance at 443 nm in the standard ocean color algorithms such as those used for SeaWiFS to correctly retrieve the biomass concentration in the case of turbid open ocean waters. On the other hand, the MIPE significantly decreases as the reflectance increases. In this latter case, the contribution to Rrs(443) from a_{tot} is weak. The problem due to the

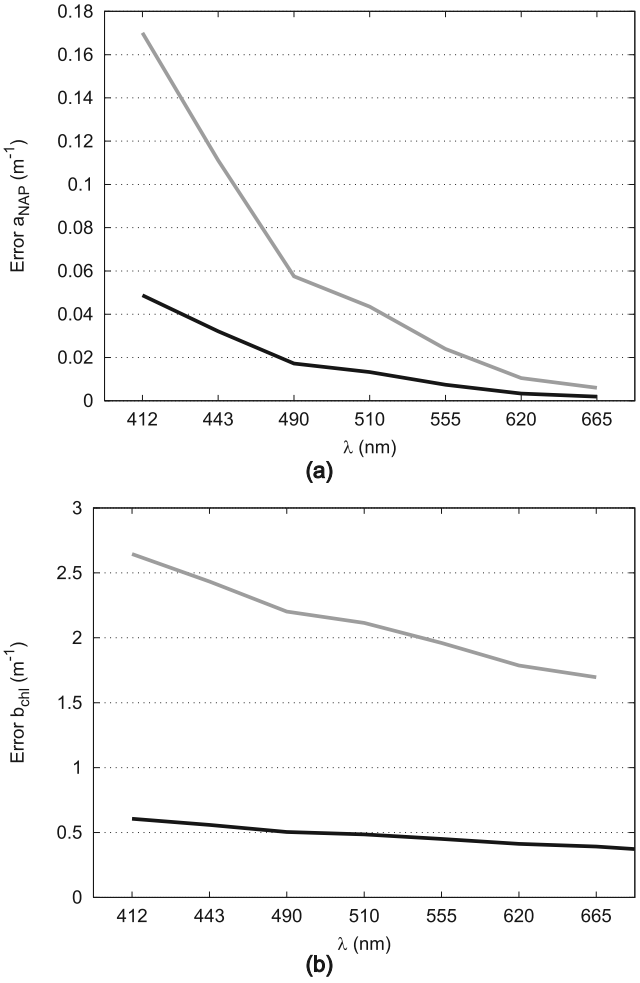

Figure 14. Spectral variations of the minimum inverse problem error of DPC_{K2} (in black) and the root-mean-square error (in gray) as provided by a Multi-Layer Perceptron trained on DPC_{K2} (a) for a_{NAP} (m^{-1}) and (b) for b_{chl} (m^{-1}). The x axis is the wavelength (nm).

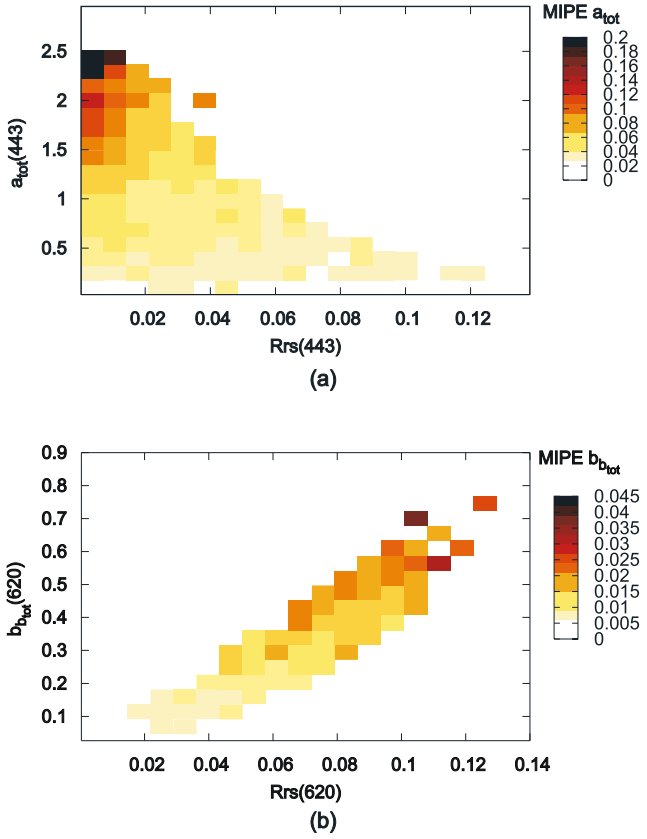

Figure 15. Maps of minimum inverse problem error (m^{-1}) expressed in terms of IOPs versus Rrs: (a) for a_{tot} (443) (m^{-1}) and (b) for b_{btot} (620) (m^{-1}).

low signal vanishes and the retrieval of a_{tot} is improved. The MIPE obtained for b_{btot} at 620 nm (Figure 15b) shows some interesting features. The wavelength 620 nm is appropriate to evaluate the scattering properties of the particles in coastal waters since the particulate absorption is strongly attenuated. The MIPE linearly increases with both the reflectance and backscattering coefficient. The highest values of MIPE point out that b_{btot} cannot be retrieved with a precision better than ±0.045 m^{-1} in turbid waters. The magnitude of the reflectance can be so high in turbid waters that a saturation of the signal occurs. As a result, it is difficult to distinguish the different values of the scattering coefficient from the reflectance, thus leading to high values of MIPE. Because the phenomenon of saturation of the reflectance is commonly observed in highly turbid waters, several authors has already investigated the use of the reflectance in the near infrared to get a measured signal more sensitive to the suspended matter [*Doxaran et al.*, 2002; *Ruddick et al.*, 2006]. Here we have illustrated through two different examples that our results are consistent with theory and observations.

[56] We propose also a second kind of map in Figure 16 in which the values of the MIPE for a_{tot} and b_{btot} are reported along with Rrs(443) and Rrs(555). This type of representation can be relevant for final users whose objectives are to inverse an in situ measured Rrs spectrum. Figure 16a shows that the minimum inversion error for a_{tot} is the greatest when the ratio Rrs(443)/Rrs(555) is low. An opposite trend is observed regarding b_{btot} (Figure 16b); high reflectance ratio leads to significant MIPE. These results are consistent with those obtained in Figure 15.

5. Discussion

[57] In this paper, we showed that the inverse problem of ocean color in coastal waters is ill-posed. We described its ambiguous aspect and how the presence of ambiguities influences the performance of the inverse models. In this section, we discuss three different strategies that should be investigated to tackle the ambiguity problem.

5.1. Integration Strategy

[58] The integration strategy is concerned with the set of methods able to integrate the ambiguities into the inverse models. To overcome the nonuniqueness of the solution, the methods based on the integration strategy provide a global unique solution, even though this global solution consists in a distribution or a complex and structured collection of all the plausible solutions.

[59] The first idea of the integration strategy is to ensure that the output of an inverse model (i.e., the retrieved IOP) is at least one of the plausible solutions instead of predicting the conditional mean $E(\mathrm{IOP}(\lambda)|\mathrm{Rrs}(\lambda))$. The underlying theory to achieve that is called learning with a distal teacher [*Jordan and Rumelhart*, 1992]. Such a theory has already been applied in the framework of the ocean color inverse problem [*Doerffer and Schiller*, 2000] or more generally in

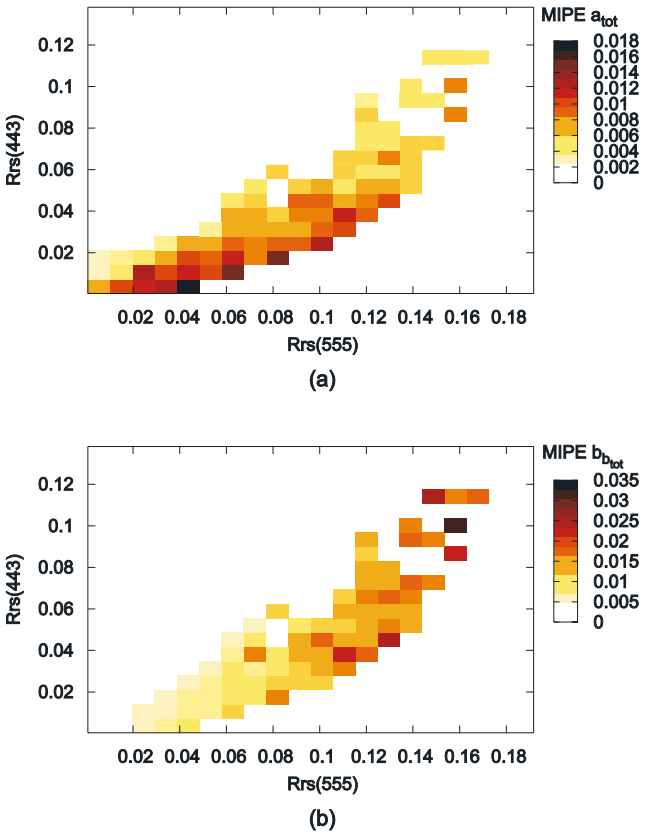

Figure 16. Maps of minimum inverse problem error (m^{-1}) expressed in term of Rrs: (a) for a_{tot}(443) (m^{-1}) and (b) for b_{btot} (443) (m^{-1}).

the work of *Schiller* [2003] to detect dramatic behaviors of models. It is worth noticing that the choice of this unique plausible solution among all the others is a critical issue that can be tackled using prior information; see for example the estimation of canopy variables in the work of *Combal et al.* [2002].

[60] A second idea of the integration strategy is to give an estimate of the distribution of the plausible solutions, for example, with the probability density function. *Gouveia and Scales* [1998] showed that ambiguities existing in inverse problems can be directly handled using the Bayesian approach [*Tarantola*, 1987] which operates on data distribution. Therefore *Gouveia and Scales* [1998] concluded that there is no ill-posed inverse problem from a Bayesian viewpoint. Unfortunately, Bayesian-based statistical approaches, such as the Mixture Density Network [*Bishop*, 1994] or Hierarchical Mixture of Experts [*Jordan and Jacobs*, 1994] cannot be systematically employed because they require both the construction of prior distributions of data and huge computational costs. Nevertheless interesting results have already been obtained for various real-world remote sensing inverse problems, for example in wind vectors retrieval [*Cornford et al.*, 1999], ocean tomography inversion [*Stéphan et al.*, 1996], and atmospheric prediction of carbon monoxide [*Hadji-Lazaro et al.*, 1999].

[61] In order to integrate the nonuniqueness of the solution into the inverse models, a third alternative in the integration strategy is to make use of a discrete representation, for example, to output discrete lists of solutions for each IOP. To increase the level of confidence in the retrieved parameters, it is often possible to associate some probabilities of appearance to each of the proposed solution. In this context, *Richaume et al.* [2000] developed a technique based on the discretization of the distributions using histograms. Such a technique might be applied regarding the ocean color inverse problem.

5.2. Divide and Conquer Strategy

[62] The Divide and Conquer strategy, which is widely used in the computer science research field, can be applied to tackle the ambiguity problem. This may lead to transform a huge complex and strongly ambiguous inverse problem into numerous small unambiguous ones. We discern at least two ways of dividing the global problem.

[63] The first way concerns the concept of transductive learning (local or personalized modeling) as opposed to inductive learning (global modeling); see *Kasabov* [2007] for a complete review. In the local modeling approach, a set of local models are created from data, each representing a subspace (cluster) of the problem space whereas a model is created for each input of the problem space in the personalized modeling approach. Different transductive learning techniques have already proved their efficiency in the Divide and Conquer strategy and should be used to eliminate the ambiguities, see for example the Polyhedral Mixture of Experts [*Karniel et al.*, 1998] or the Evolving Connectionist System [*Futschik et al.*, 2003].

[64] The second way is much more concerned with physics and consists in the specialization of inverse models to particular restricted subproblems. Practically, inverse models specifically designed for given geographic areas and/or for given seasons and/or for particular coastal waters types, i.e., CDOM or mineral dominated waters, might be satisfactorily applied to overcome the problem of ambiguities.

5.3. Enrichment Strategy

[65] The enrichment strategy proposes to drastically reduce the effects of ambiguities using ancillary data. We have identified at least three potential sources of ancillary information.

[66] In situ measurements obtained during field experiments or using buoys moorings provide a significant amount of statistical ancillary information. This information can be processed offline, during a long time period, to increase the global knowledge of the experts, for example to inform about seasonal or geographical variations of IOPs or about prior distributions that can be critical in the Bayesian approach [*Tarantola*, 1987]. The global knowledge of the experts can then be used in real-time to resolve ambiguities of a given inversion or beforehand to improve the synthetic data set creation process. The in situ collected data can also be used directly to adjust the inverse model or to select the most probable answer. This latter case can be envisaged only if the Rrs measurements of the database were collected at the same time period and same area as the Rrs spectrum to invert. The most likely solution of the inverse model is then the retrieved IOP that best matches the measured IOP of the database. Unfortunately, match-up events do no occur so often and cannot be integrated routinely into an inversion algorithm. They can only be used as a validation step.

[67] The second source of ancillary data relies on the use of additional physical information than that commonly measured in ocean color field experiment. In particular, the directional variations and the polarization of the reflectance might be relevant physical constraints that should contribute to significantly reduce the influence of the ambiguities on the retrieved solutions. Currently, the satellite sensor PARASOL (CNES) is the only one that is able to measure the directionality and polarization of the reflectance. However, the exploitation of the data provided by this sensor still remains to be achieved. Note that in situ instruments devoted to the measurements of the directional and polarized subsurface reflectance are not fully operational yet and are still being developed [*Souaidia and Voss*, 2006]. Hyperspectral measurements might also be used as relevant information to reduce the number of solutions of the inverse problem. Previous studies [*Doxaran et al.*, 2002; *Ruddick et al.*, 2006] showed that hyperspectral data can significantly improve the determination of the optical properties of particles in coastal waters. Nevertheless, most of the results were obtained based on ship measurements and great efforts should still be put in the development of hyperspectral satellite sensors to increase the spatial coverage of coastal zones. Satellite multisensor approaches could also be highly promising to tackle the ambiguity problem inasmuch as any pixel can be characterized by various physical information such as wind and sea surface temperature. The exploitation of data provided by satellite platforms such as the ENVISAT platform (European Space Agency) should be encouraged to reach the objectives above mentioned.

[68] The last source of potential information that has to be investigated is the local context of the measurement. Indeed, it is most of the time clear that the different variables describing the physical phenomenon to study, in our case

the set of all IOPs, are temporally and/or spatially correlated. Thus, when inverting a given highly ambiguous spectrum for which multiple answers are produced by inverse models, it is promising to explore the spatial and/or temporal neighborhood of the Rrs spectrum to invert to find the nearest nonambiguous measurement. Then the ambiguity problem can be solved locally with a close-to-close approach. As an example, when computing the value of a pixel in a satellite image, the number of solutions might be significantly reduced accounting for the information provided by neighboring pixels.

6. Summary

[69] The nonuniqueness of the solution of the inverse problem of ocean color, so-called ambiguity problem, was characterized for coastal waters applications. First, great efforts were devoted to the creation of a synthetic data set representative of real-world conditions. In particular, the simulations were constrained by both observations and statistical rules, thus making the data set reliable for inverse modeling purposes. Second, mathematical definitions were proposed to quantify the presence of ambiguities. Note that this is the first time, to our knowledge, that such formulations of the ambiguity problem are provided. The rate of ambiguity of the remote sensing reflectance spectra within the data set was high, around 90%. The influence of the ambiguities on inverse modeling was studied. It was demonstrated that the fraction of the error that is ascribed to the ambiguities, so-called minimum inverse problem error, is equal to the dispersion of the plausible solutions of the problem. Practically, the ambiguity error is estimated calculating the standard deviation over all the IOPs that lead to one given unique Rrs spectrum. On the basis of this important result, it was shown that the minimum error made on the total absorption coefficient is the highest when the reflectance values in the blue are weak. On the other hand, a greater error is made regarding the total backscattering coefficient when the reflectance level is high. Several strategies are finally discussed to tackle the ambiguity problem. Future efforts should be put in the application of these strategies to significantly enhance the performance of the inverse modeling in optically complex waters.

Notation

λ	wavelength, nm.
θ_s	solar zenith angle, degrees.
θ_v	viewing zenith angle, degrees.
a_{tot}	total absorption coefficient, m^{-1}.
b_{tot}	total scattering coefficient, m^{-1}.
b_{btot}	total backscattering coefficient, m^{-1}.
a_p	particulate absorption coefficient, m^{-1}.
b_p	particulate scattering coefficient, m^{-1}.
b_p	particulate backscattering coefficient, m^{-1}.
a_{chl}	absorption coefficient of chlorophyll *a*, m^{-1}.
a_{NAP}	absorption coefficient of non algal particles, m^{-1}.
a_{CDOM}	absorption coefficient of colored dissolved organic matter, m^{-1}.
b_{chl}	scattering coefficient of chlorophyll *a,* m^{-1}.
b_{sed}	scattering coefficient of sediment, m^{-1}.
b_{bchl}	backscattering coefficient of chlorophyll *a,* m^{-1}.
b_{bsed}	backscattering coefficient of sediment, m^{-1}.
iid	independent and identically distributed.
CNES	Centre National d'Etudes Spatiales.
ESA	European Space Agency.
IOP	inherent optical property.
MIPE	minimum inverse problem error.
OSOA	Ordres Successifs Ocean Atmosphere.
PARASOL	Polarization and Anisotropy of Reflectances for Atmospheric Sciences Coupled with Observations from a Lidar.
RMSE	root mean square error.
Rrs	subsurface remote sensing reflectance.
RTE	radiative transfer equation.
R^2	coefficient of determination.

[70] **Acknowledgments.** This work was funded by the Centre National d'Etudes Spatiales (CNES, France). We are grateful to Marcel Babin for providing the COASTLOOC database. We also would like to thank S. Verel for valuable discussion.

References

Babin, M., A. Morel, V. Fournier-Sicre, F. Fell, and D. Stramski (2003a), Light scattering properties of marine particles in coastal and open ocean waters as related to the particle mass concentration, *Limnol. Oceanogr.*, *48*, 843–859.

Babin, M., D. Stramski, G. M. Ferrari, H. Claustre, A. Bricaud, G. Obolensky, and N. Hoepffner (2003b), Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe, *J. Geophys. Res.*, *108*(C7), 3211, doi:10.1029/2001JC000882.

Bishop, C. M. (1994), Mixture density networks, *Tech. Rep. NCGR/4288*, Neural Comput. Res. Group, Aston Univ., Birmingham, U. K.

Bricaud, A., A. Morel, and L. Prieur (1981), Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains, *Limnol. Oceanogr.*, *26*, 43–53.

Campbell, J. W. (1995), The lognormal distribution as a model for bio-optical variability in the sea, *J. Geophys. Res.*, *100*, 237–254.

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.*, *104*, 5403–5421.

Chami, M., and D. Robilliard (2002), Inversion of oceanic constituents in case I and II waters with genetic programming algorithms, *Appl. Opt.*, *41*, 260–274.

Chami, M., R. Santer, and E. Dilligeard (2001), Radiative transfer model for the computation of radiance and polarization in an ocean-atmosphere system: Polarization properties of suspended matter for remote sensing, *Appl. Opt.*, *40*, 2398–2416.

Chami, M., E. B. Shybanov, T. Y. Churilova, G. A. Khomenko, M. E.-G. Lee, O. V. Martynov, G. A. Berseneva, and G. K. Korotaev (2005), Optical properties of the particles in the Crimea coastal waters (Black Sea), *J. Geophys. Res.*, *110*, C11020, doi:10.1029/2005JC003008.

Chami, M., E. Marken, J. J. Stamnes, G. Khomenko, and G. Korotaev (2006), Variability of the relationship between the particulate backscattering coefficient and the volume scattering function measured at fixed angles, *J. Geophys. Res.*, *111*, C05013, doi:10.1029/2005JC003230.

Combal, B., F. Baret, M. Weiss, A. Trubuil, D. Macé, A. Pragnère, R. Myneni, Y. Knyazikhin, and L. Wang (2002), Retrieval of canopy biophysical variables from bidirectional reflectance using prior information to solve the ill-posed inverse problem, *Remote Sens. Environ.*, *84*, 1–15.

Cornford, D., I. T. Nabney, and C. K. I. Williams (1999), Bayesian inference for wind field retrieval, *Neurocomput. Lett.*, *26–27*, 1013–1018.

Doerffer, R., and H. Schiller (2000), Neural network for retrieval of concentrations of water constituents with the possibility of detecting exceptional out of scope spectra, paper presented at IGARSS 2000, Inst. of Electr. and Electron. Eng., Honolulu, Hawaii.

Doxaran, D., J. M. Froidefond, and P. Castaing (2002), A reflectance band ratio used to estimate suspended matter concentrations in coastal sediment-dominated waters, *Int. J. Remote Sens.*, *23*, 5079–5085.

Futschik, M. E., A. Reeve, and N. Kasabov (2003), Evolving connectionist systems for knowledge discovery from gene expression data of cancer tissue, *Artif. Intell. Med.*, *28*, 165–189.

Garver, S. A., and D. Siegel (1997), Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: 1. Time series from the Sargasso Sea, *J. Geophys. Res.*, *102*, 607–625.

Gordon, H. R., D. K. Clark, J. W. Brown, O. B. Brown, R. H. Evans, and W. W. Broenkow (1983), Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison of ship determinations and CZCS estimates, *Appl. Opt.*, *22*, 20–36.

Gouveia, W., and J. A. Scales (1998), Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis, *J. Geophys. Res.*, *103*, 2759–2779.

Gross, L., S. Thiria, R. Frouin, and B. G. Mitchell (2000), Artificial neural networks for modelling the transfer function between marine reflectances and phytoplankton pigment concentration, *J. Geophys. Res.*, *105*, 3483–3495.

Hadamard, J. (1902), Sur les problemes aux derivees parielies et leur signification physique, *Princeton Univ. Bull.*, *13*, 49–52.

Hadji-Lazaro, J., C. Clerbaux, and S. Thiria (1999), An inversion algorithm using neural networks to retrieve atmospheric CO total columns from high-resolution nadir radiances, *J. Geophys. Res.*, *104*, 23,841–23,854.

Hoge, F. E., and P. E. Lyon (1996), Satellite retrieval of inherent optical properties by linear matrix inversion of oceanic radiance models: An analysis of model and radiance measurement errors, *J. Geophys. Res.*, *101*, 631–648.

International Ocean Colour Coordinating Group (2006), Remote sensing of inherent optical properties: Fundamentals, tests of algorithms, and applications, in *Reports of the International Ocean Colour Coordinating Group*, vol. 5, edited by Z. P. Lee, pp. 1–126, Dartmouth, N. S., Canada.

Jordan, M. I., and R. A. Jacobs (1994), Hierarchical mixtures of experts and the EM algorithm, *Neural Comput.*, *6*, 181–214.

Jordan, M. I., and D. E. Rumelhart (1992), Forward models: Supervised learning with a distal teacher, *Cognitive Sci.*, *16*, 307–354.

Karniel, A., R. Meir, and G. Inbar (1998), Polyhedral mixture of linear experts for many-to-one mapping inversion, *in ESANN '98 Meeting*, pp. 155–160, Eur. Symp. on Art. Neural Networks, Brussels, Belgium.

Kasabov, N. (2007), Modelling and profile discovery in bioinformatics: Global, local and personalised approach, *Pattern Recognition Lett.*, in press.

Lee, Z. P., K. L. Carder, T. G. Peacock, C. O. Davis, and J. L. Mueller (1996), Method to derive ocean absorption coefficients from remote-sensing reflectance, *Appl. Opt.*, *35*, 453–462.

Lee, Z. P., K. L. Carder, and R. Arnone (2002), Deriving inherent optical properties from water color: A multi- band quasi-analytical algorithm for optically deep waters, *Appl. Opt.*, *41*, 755–772.

Maritorena, S., D. A. Siegel, and A. R. Peterson (2002), Optimization of a semianalytical ocean color model for global-scale applications, *Appl. Opt.*, *41*, 5–14.

Mobley, C. D. (1989), A numerical model for the computation of radiance distributions in natural waters with wind roughened surfaces, *Limnol. Oceanogr.*, *34*, 1473–1483.

Mobley, C. D., L. K. Sundman, and E. Boss (2002), Phase function effects on oceanic light fields, *Appl. Opt.*, *41*, 1035–1050.

Morel, A. (1988), Optical modeling of the upper ocean in relation to its biogenous matter content (Case I waters), *J. Geophys. Res.*, *93*, 749–768.

O'Reilly, J. E., S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru, and C. McClain (1988), Ocean color algorithms for SeaWiFS, *J. Geophys. Res.*, *103*, 937–953.

Richaume, P., F. Badran, M. Crepon, C. Mejía, H. Roquet, and S. Thiria (2000), Neural network wind retrieval from ERS-1 scatterometer data, *J. Geophys. Res.*, *105*, 8737–8751.

Roesler, C. S., and E. Boss (2003), Spectral beam attenuation coefficient retrieved from ocean color inversion, *Geophys. Res. Lett.*, *30*(9), 1468, doi:10.1029/2002GL016185.

Roesler, C. S., and M. J. Perry (1995), In situ phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance, *J. Geophys. Res.*, *100*, 279–294.

Ruddick, K. G., V. D. Cauwer, J. Y. Park, and G. Moore (2006), Seaborne measurements of near infrared water leaving reflectance: The similarity spectrum for turbid waters, *Limnol. Oceanogr.*, *51*, 1167–1179.

Schiller, H. (2003), Neural net architectures for scope check and monitoring, paper presented at International Symposium on Computational Intelligence for Measurement Systems and Applications, Inst. of Electr. and Electron. Eng., New York.

Schiller, H., and R. Doerffer (1999), Neural network for emulation of an inverse model-operational derivation of case II water properties from MERIS data, *Int. J. Remote Sens.*, *20*, 1735–1746.

Schiller, H., and R. Doerffer (2005), Improved determination of coastal water constituent concentrations from MERIS data, *IEEE Trans. Geosci. Remote Sens.*, *43*, 585–591.

Shettle, E. P., and R. W. Fenn (1979), Models for the aerosols of the lower atmosphere and the effect of humidity variations on their optical properties, *Environ. Res. Pap.*, *676*, 1–31.

Souaidia, N., and K. J. Voss (2006), POLRADS: Polarization upwelling radiance distribution measurements, paper presented at the International Ocean Optics Conference XVIII, Oceanogr. Soc., Montreal, Quebec, Canada.

Stéphan, Y., S. Thiria, and F. Badran (1996), Application of multilayered neural networks to ocean tomography inversions, *Inverse Probl. Eng.*, *1*, 181–304.

Tarantola, A. (1987), *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier, New York.

Vapnik, V. N. (1995), *The Nature of Statistical Learning Theory*, Springer, New York.

M. Chami and M. Defoin-Platel, Université Pierre et Marie Curie-Paris, Laboratoire d'Océanographie de Villefranche sur Mer, Unité Mixte de Recherche CNRS 7093, Quai de la Darse, BP08, F-06230 Villefranche sur Mer, France. (dpm@obs-vlfr.fr)