# Use of a Neurovariational Inversion for Retrieving Oceanic and Atmospheric Constituents from Ocean Color Imagery: A Feasibility Study

Cédric Jamet, Sylvie Thiria, Cyril Moulin, Michel Crépon

# Use of a Neurovariational Inversion for Retrieving Oceanic and Atmospheric Constituents from Ocean Color Imagery: A Feasibility Study

C. JAMET

*LODyC, Paris, and ACRIst, Sophia-Antipolis, France*

S. THIRIA

*LODyC, Paris, France*

C. MOULIN

*Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France*

M. CREPON

*LODyC, Paris, France*

(Manuscript received 9 October 2003, in final form 23 April 2004)

ABSTRACT

This paper presents a neurovariational method for inverting satellite ocean-color signals. The method is based on a combination of neural networks and classical variational inversion. The radiative transfer equations are modeled by neural networks whose inputs are the oceanic and atmospheric parameters, and outputs the top of the atmosphere reflectance at several wavelengths. The procedure consists in minimizing a quadratic cost function that is the distance between the satellite-observed reflectance and the computed neural-network reflectance, the control parameters being the oceanic and atmospheric parameters.

First, a feasibility experiment using synthetic data is presented to show that chlorophyll-*a* can be retrieved with an error of 19.7% when the atmospheric parameters are known exactly. Then both atmospheric and oceanic parameters are relaxed. A first guess for the atmospheric parameters was provided by a direct inverse neural network whose inputs are at near-infrared wavelengths. Sensitivity experiments showed that these parameters can be retrieved with an adequate accuracy.

An inversion of a composite SeaWiFS image is presented. Optical thickness and chlorophyll-*a* both give coherent spatial structures when a background term is added to the cost function. Finally, chlorophyll-*a* retrievals are compared with SeaWiFS product through in situ data. It shows a better estimation of the chlorophyll-*a* with the neurovariational inversion for the oligotrophic regions.

## 1. Introduction

Since the launch in 1978 of *CZCS*, which was the first satellite dedicated to the study of the ocean color, several ocean-color sensors were put into orbit to observe it (MODIS, SeaWiFS, POLDER, MERIS). We know that the solar energy backscattered by the ocean–atmosphere system is mainly due to the aerosols in the atmosphere (up to 90%) and to a lesser extent to the chlorophyll-*a* concentration (chl-*a* hereafter) and the phaeophytin *a* in the ocean (especially for case 1 waters). The latter parameters allow us to make a quantitative assessment of oceanic primary production and its role in the global carbon cycle, which are critical environmental and scientific issues. Knowledge of primary production is necessary to calculate new production, to derive the effect of biological processes on the partial pressure of carbon dioxide ($CO_2$), and therefore to better understand how phytoplankton carbon fixation affects the net $CO_2$ flux across the air–sea interface.

For case 1 waters, which constitute more than 90% of the World Ocean, optical properties depend on phytoplankton pigments and their covarying degradation products. The upward radiance containing the information on the phytoplanktonic pigments, the so-called water-leaving reflectance, is transmitted to the top of the atmosphere (TOA). This radiance represents at most

*Corresponding author address:* Prof. Sylvie Thiria, LODyC, tour 45-55, 5eme étage, case courrier 100, Université Pierre et Marie Curie, 4, place Jussieu, 75252 Paris, Cedex 05, France.
E-mail: thiria@lodyc.jussieu.fr

10% of the total signal received by the satellite sensor in the blue and less in the green wavelength. Therefore, the water-leaving radiance must be extracted from the measured radiance—a process referred to as atmospheric correction.

The major problem in atmospheric correction is to estimate the aerosol contribution to the measured TOA radiance. Aerosols are highly variable in space and time as well as in composition; therefore, their physical–chemical characteristics need to be determined for each pixel of an image. Current methods rely on near-infrared (NIR) bands to estimate the aerosol contribution at visible wavelengths (Gordon and Wang 1994a). However, these algorithms generally fail when aerosols absorb radiation (Gordon 1997).

New algorithms have recently been developed to account for aerosol absorption by simultaneously retrieving aerosol parameters (APs) and the phytoplanktonic pigments (Gordon et al. 1997; Chomko and Gordon 1998). The principle of the latter method is to use the NIR bands to retrieve some of the APs {viz., the size-distribution parameter $\nu$ and the aerosol optical thickness $\tau(865)$ at 865 nm [$\tau(865)$]}, assuming that the water-leaving reflectance $\rho_w$ is negligible. Then the oceanic constituents (OCs), that is, chl-*a* and a particulate scattering parameter, $b^0$, and two APs, the aerosol refractive index ($m = m_r - im_i$, which is independent of the wavelength) are determined by minimizing a cost function defined as the distance between the computed radiance at the TOA and the observed radiance at different wavelengths in the visible.

The Chomko and Gordon (1998) procedure uses lookup tables (LUTs) to compute the TOA radiance, that is, the atmospheric and oceanic radiative transfer equations (RTEs). The result is that the LUTs are incorporated into the minimization procedure of a dedicated cost function. Since these LUTs are discrete, the computation of this cost function and its gradient, which are essential ingredients of the minimization procedure, might be difficult. Moreover, the interpolation between contiguous values of the LUTs may introduce error into the retrieved parameters. We expect that the use of neural networks, especially the multilayer perceptrons (MLPs), which are universal approximators of any nonlinear continuous function (Bishop 1995), can replace the LUT and improve the accuracy of the retrieval of APs and OCs since MLPs yield continuous and differentiable functions. The second advantage is the straigthforward calculation of the MLP's gradient in a variational procedure.

In this paper, we propose an extension of the Chomko and Gordon method. The LUTs are replaced by MLPs. The main idea of the paper is to investigate the flexibility of a variational inversion where the transfer radiative equations are represented by neural networks and its suitability for operational inversion. This methodology is denoted neurovariational inversion (NVI). This work is just a feasibility study; we restrain the aerosol models to weakly and moderate absorbing aerosols. The physics is less difficult than for strongly absorbing aerosols. We want to show the robustness of our method and the advantages.

The paper is structured as follows: description of the RTE of the ocean–atmosphere system and the aerosol and ocean models used in the inversion, based on the formalism of Chomko and Gordon, is given in section 2. Section 3 describes the synthetic dataset used to calibrate the MLPs and validate the method. The theory of the MLP is briefly described in section 4 with the model of the RTE with MLPs; the NVI procedure and the influence of the first guess on the retrieved OC for weakly absorbing aerosols using simulated data are presented, and a first qualitative inversion of a week of SeaWiFS ocean-color imagery is detailed in section 5. This section ends with a validation of the NVI chlorophyll-*a* retrievals with in situ data.

## 2. Oceanic and atmospheric radiative transfer models

In this section we present the RTE used for the atmosphere–ocean system.

### *a. Radiative transfer equations*

We dealt with the reflectance $\rho$ rather than the radiance $L$, which is defined as $\rho = \pi L/F_0 \cos\theta_0$, where $F_0$ is the extraterrestrial solar irradiance and $\theta_0$ the solar zenith angle. The reason is that $\rho$ is dimensionless; besides, it is possible to more accurately calibrate future sensors in reflectance than in radiance (Gordon and Wang 1994a). After correction for oxygen and water-vapor absorption effects and removal of the whitecap perturbation (Gordon and Wang 1994b), the reflectance at the top of the atmosphere received by the satellite $\rho_t(\lambda)$ comprises of four terms (Gordon and Wang 1994b; Gordon 1997): the (pure) Rayleigh scattering contribution $\rho_r(\lambda)$, the (pure) aerosol scattering contribution $\rho_a(\lambda)$, the contribution of the interaction between the aerosol and air molecules $\rho_{\mathrm{ra}}(\lambda)$ (Deschamps et al. 1983), and the water-leaving reflectance $\rho_w(\lambda)$, that is,

$$\rho_t(\lambda) = \rho_r(\lambda) + \rho_a(\lambda) + \rho_{\mathrm{ra}}(\lambda) + t(\lambda)\rho_w(\lambda), \qquad (1)$$

where $t(\lambda)$ is the diffuse transmittance of the atmosphere (Gordon 1997).

The (pure) Rayleigh contribution $\rho_r$ can be computed precisely from an estimate of the surface atmospheric pressure (Gordon et al. 1988a). In fact, we obtain a corrected TOA reflectance $\rho_{\mathrm{cor}}$, that is,

$$\rho_{\mathrm{cor}}(\lambda) = \rho_t(\lambda) - \rho_r(\lambda) = \rho_A(\lambda) + t(\lambda)\rho_w(\lambda), \qquad (2)$$

where $\rho_A(\lambda) = \rho_a(\lambda) + \rho_{\mathrm{ra}}(\lambda)$. We now present the RTE modeling of the three terms of Eq. (2), that is, $\rho_A$, $\rho_w$, and $t$.

*b. The atmospheric radiative transfer model*

We use the atmospheric RTE proposed by Chomko and Gordon (1998). Its parameters are the sun-viewing geometry (the solar zenith angle $\theta_s$, the sensor zenith angle $\theta_v$, the difference between the solar and sensor azimuth angle $d\phi_v$), and the AP, that is, the aerosol optical thickness $\tau(865)$, a wavelength-independent complex refractive index $m = m_r - im_r$ and a particle size distribution parameter $\nu$, which is a parameter of the Junge power law (Junge 1958) that models the distribution of particle size. The Junge power law is of the form

$$\frac{dN}{dD} = \begin{cases} K, & D_0 < \mathrm{D} \le D_1 \\ K\left[\dfrac{D_1}{D}\right]^{\nu+1}, & D_1 < \mathrm{D} \le D_2 \\ 0, & \mathrm{D} > D_2 \end{cases},$$

where $dN$ is the number of particles per unit volume with diameters varying from $D$ to $D + dD$. Chomko and Gordon selected $D_0 = 0.06$ $\mu$m, $D_1 = 0.20$ $\mu$m, and $D_2 = 20$ $\mu$m; $K$ is a constant.

In contrast to classical bimodal lognormal aerosol size distributions, which require several parameters to describe the modal diameters and standard deviation of both modes, and their relative concentrations (Shettle and Fenn 1979), the simplicity of the Junge power-law distribution is attractive in that only one parameter, $\nu$, is required. Chomko and Gordon have shown that the principal advantage of the use of the Junge power law is that realistic aerosol models are not required for the extraction of ocean properties. Unfortunately, by abandoning realistic models in favor of the simple Junge power law, we may lose some ability to retrieve meaningful aerosol properties. So in this paper we focus on the retrieval of OC, especially the pigment concentration $C$.

The refractive index is combined with this size distribution to compute the aerosol optical properties—the phase function, optical thickness $\tau_a(\lambda)$, and single-scattering albedo $\omega_0$—following Mie theory. The aerosol optical properties are then used to generate the reflectance $\rho_A$ as a function of $\nu$, $m$, and $\tau(865)$ for the different sun-viewing geometries.

*c. The diffuse transmittance model*

The diffuse transmittances are computed separately by assuming that the in-water upwelling radiance distribution just beneath the surface is uniform (Yang and Gordon 1997). The diffuse transmittance is suitable for the water-leaving reflectance $\rho_w$ and whitecap radiance, as they have near-uniform angular distribution (Gordon 1997). It is defined as the water-leaving reflectance $\rho_w$ in a particular viewing direction $(\theta_s, \theta_v)$ "transmitted" to the top of the atmosphere. The diffuse transmittance accounts for the direct loss from $\rho_w(\theta_s, \theta_v)$ due to absorption and scattering within the atmosphere, as the gain in radiance in the direction $(\theta_s, \theta_v)$ due to the scattering of $\rho_w(\theta_s, \theta_v)$; $t(\theta_s, \theta_v, \lambda)$ is approximated by (Gordon et al. 1983)

$$t(\theta_s, \theta_v, \lambda) = \exp\left[-\left(\frac{\tau_r(\lambda)}{2} + \tau_{Oz}\right)\left(\frac{1}{\nu_v}\right)\right] t_a(\theta_v, \lambda), \tag{3}$$

where

$$t_a(\theta_v, \lambda) = \exp\left[-\frac{[1 - \omega_a(\lambda)F_a(\nu_v, \lambda)]\tau_a(\lambda)}{\nu_v}\right], \tag{4}$$

and $\nu_v = \cos\theta_v$, $F_a(\nu_v, \lambda)$ is the aerosol phase function, $\tau_r(\lambda)$ is the Rayleigh aerosol optical thickness, $\tau_{Oz}$ is the ozone aerosol optical thickness, and $\omega_a$ is the aerosol single-scattering albedo.

*d. The oceanic radiative transfer model*

The spectral marine reflectance $\rho_w$ defined as the ratio of upwelled to downwelled irradiance just below the surface, is a function of the pigment concentration $C$ and the scattering parameter $b^0$ using the biooptical model of Gordon et al. (1988b).

## 3. The synthetic dataset

The NVI was tuned using a synthetic dataset (NLUT), which was computed using the previous RTEs. In the present study we only dealt with the five visible SeaWiFS wavelengths: $\lambda_i(nm) = (412, 443, 490, 510, 555)$ because $C$ can mainly be retrieved in these wavebands for case 1 waters.

The synthetic $\rho_A$ and $t$ were computed for various sun-viewing geometries $(\theta_s, \theta_v, d\phi_v)$ and APs by solving the scalar RTE for a two-layer atmosphere system with aerosols confined to the lower layer. The sampling of the AP and sun-viewing geometries for the atmosphere gave us a wide range of realistic configurations. The values of the sun-viewing geometry cover most of the range of the corresponding SeaWiFS values. The synthetic $\rho_w(\lambda)$ was computed by using the biooptical model of Gordon et al. (1988b). The sampling of the AP $[\nu, m_r, m_i$ and $\tau(865)]$ for weakly absorbing aerosols and OC (the scattering coefficient $b^0$ and $C$) is given in Table 1.

At the end of the data simulation we obtained three distinct NLUTs: NLUT-A for $\rho_A$, NLUT-T for $t$, and NLUT-O for $\rho_w$. From each NLUT we extracted three different datasets: the learning set, the validation set, and the test set. For each NLUT, these different subsets were constituted by picking at random, in the NLUT, the spectra with the related values of the parameters. So each set presents the same statistical properties. The $\rho_{cor}$ and $t$ learning, validation, and test database have 20 000 different patterns and the $\rho_w$ learning and validation database have 500 patterns, and the test set only 224 patterns. These datasets allow us to model Eq. (2)

TABLE 1. Range of each parameter used in the LUT.

| Parameters | Range | Steps |
|---|---|---|
| Aerosols optical properties | | |
| $\nu$ | 2–4.5 | 0.5 |
| $m_r$ | 1.33; 1.5 | |
| $m_i$ | 0–0.003 | log |
| $\tau(865)$ | 0.05–0.35 | 0.1 |
| Geometry angles | | |
| $\theta_s$ | 20–50 | 3 |
| $\theta_v$ | 0–60 | 3 |
| $d\phi_v$ | 0–360 | 36 |
| Oceanic constituents | | |
| $b^0$ | 0.12–0.45 | 0.03 |
| $C$ (mg m$^{-3}$) | 0.03–3 | log |

with three specific MLPs dedicated to the modeling of $\rho_A$, $\rho_w$, and $t$, respectively. These MLPs are suitable for modeling a large variety of multidimensional and nonlinear functions (Bishop 1995) such as those encountered in geophysics (Thiria et al. 1993). The three independent MLPs used in NVI were calibrated using the learning set described above.

## 4. Neuromodeling of the RTE

The main idea of NVI is to model the RTE (direct model) with MLPs in the minimization procedure, instead of using discrete LUTs. A brief description of the MLPs used in the variational inversion is given here. The MLP is a universal approximator of continuous functions (Cybenko 1989; Hornik et al. 1989; Pinkus 1999). We also give an overview of the principles of the MLP used for nonlinear regression.

### *a. MLP presentation*

A neuron is an elementary transfer function that provides an output $s$ when an input $A$ is applied. An MLP is a set of interconnected neurons. Each neuron receives from and sends signals to only the neurons to which it is connected. Thanks to this association of elementary tasks, an MLP is able to solve complicated problems. The specificity of an MLP depends on the topology of the neurons (number of layers, numbers of neurons in each layer) and on the connection weights $w_{ij}$ from neuron $j$ to neuron $i$. The MLP architecture has one layer receiving the input, one layer broadcasting the output, and one or more intermediate layers (the hidden layers). All of the networks we used in the present study were fully connected.

The $w_{ij}$ were computed by a calibration process, the so-called learning process, using the learning set. We minimized a cost function defined most of the time as a quadratic difference between the desired and the computed output, the $w_{ij}$ values being the control variables. For this we used a conjugate-gradient technique that is an iterative optimization method adapted to MLP, the so-called backpropagation gradient (Bishop 1995). When the calibration is done, the MLP model does algebraic operations only, leading to fast computation. The optimal architecture of each MLP is obtained by doing successive tests in which the number of neurons and hidden layers is increasing, using the validation set. We chose the architecture that combined a minimal error for a minimal number of neurons. The optimal architectures are presented in Table 2.

According to Eq. (2) we used three MLPs to model $\rho_{\rm cor}$: MLP-A to model $\rho_A$, MLP-T to model $t$, and MLP-O to model $\rho_w$; combining the three outputs of the three MLPs allowed us to compute the value of $\rho_{\rm cor}$ as displayed in Fig. 1.

### *b. Test on MLPs modeling the RTE*

At the end of the learning phase the three MLPs were thus capable of computing $\rho_A$, $t$, $\rho_w$, and consequently $\rho_{\rm cor,}$ according to Eq. (2), at any given visible wavelength and for the given sun-viewing geometry, AP, and OC. The MLP statistical performances are discussed below.

Figure 2 displays the scatterplot (in terms of a datum line) of the calculated MLP-T versus the desired $t$ for the test set Test-T. This figure demonstrates that MLP-T has nearly perfectly modeled the diffuse transmittance, since almost all of the data are on the diagonal. MLP-T approximates the transmittance to a very high accuracy of about $2.4 \times 10^{-3}$ in terms of rmse and of 0.5% for the relative rmse.

Figure 3 displays the scatterplot (in term of datum line) of the calculated MLP-O versus the desired $\rho_w$ for the test set Test-O. All of the data are on the diagonal, the rmse was $3.69 \times 10^{-5}$, and the relative rmse was 1%. The MLP-O model can thus approximate the water-leaving reflectance ($r^2 \simeq 1$) to a high accuracy with no bias.

Figure 4 displays the scatterplot (in terms of a datum line) of the calculated MLP-A versus the desired $\rho_A$ for the test set Test-A. We observed a quite large scattering of the desired values of $\rho_A$ between 0 and 0.1 and above 0.35. The values of $\rho_A$ less than 0.1 corresponded to weak values of $\tau(865)$ [$\tau(865) \leq 0.08$] and to weakly absorbing aerosols for which the atmospheric correction is easy to apply. In the case of moderately absorb-

TABLE 2. Architecture of each MLP.

| | MLP-A | MLP-T | MLP-O |
|---|---|---|---|
| Inputs | $\nu$, m, $\tau(865)$, $\theta_s$, $\theta_v$, $d\phi_v$, $\lambda$ | $\nu$, m, $\tau(865)$, $\theta_s$, $\theta_v$, $\lambda$ | $b^0$, $C$, $\lambda$ |
| Number of neurons (first hidden layer) | 20 | 10 | 10 |
| Number of neurons (second hidden layer) | 20 | 10 | 10 |
| Output | $\rho_A$ | $t$ | $\rho_w$ |

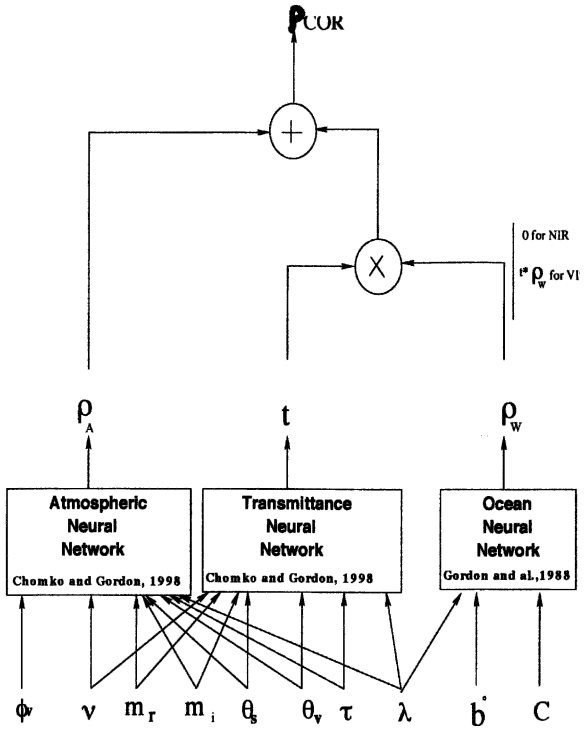


FIG. 1. MLP direct model in the visible (the desired parameters are circled).

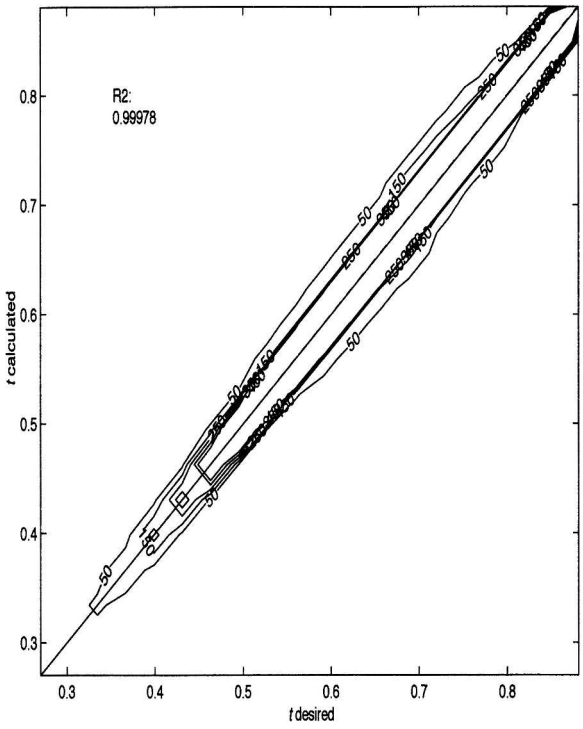


FIG. 2. Scatterplot (in terms of a datum line) of MLP-calculated vs desired transmittance $t$: $R^2$ is the correlation coefficient. The solid line corresponds to the slope l line and the dotted line to the regression line.

ing aerosols ($m_i \leq 0.002$) the values of $\rho_A$ were very close to the diagonal. From the scatterplot presented in Fig. 4 and as the relative rmse was 6.9%, it was concluded that $\rho_A$ is relatively difficult to learn with an MLP. However, these results are acceptable for this application since the rmse is equivalent to the measurement error due to the radiometric noise for most of the ocean-color sensors and relatively large for weak values of $\rho_A$, for which the effect of the atmosphere on $\rho_{\text{cor}}$ is small.

In conclusion, the performance of the MLPs leads to very satisfactory results. The three MLPs modeled the three signals with a high accuracy, especially for the diffuse transmittance $t$ and the water-leaving reflectance $\rho_w$. For the three MLPs, the correlation coefficient was almost 1, meaning that there was no bias in the model. The MLPs have therefore correctly learned the oceanic and atmospheric RTE.

The accuracy of $\rho_{\text{cor}}$, which is the parameter that is actually measured by the sensor, can be estimated by combining the three MLPs. We tested the performances achieved in computing $\rho_{\text{cor}}$ by picking at random 1008 different cases in the initial NLUT, varying AP, the geometry, and OC. In the following, Test-1008 refers to this independent set of spectra. The Test-1008 set included only weakly absorbing aerosols ($m_i \leq 0.003$). This set of aerosols is somewhat similar to the data used in the standard atmospheric correction algorithm of SeaWiFS (Shettle and Fenn 1979). It is therefore a pertinent example to test the robustness and the accuracy of our inversion method. The rmse was $1.56 \times 10^{-3}$. In terms of relative rmse, the score is of 3.63% and the model is weakly biased, $r^2$ being equal to 99.93%. With respect to these results we conclude that we have obtained an accurate model of $\rho_{\text{cor}}$ with the combination of the three MLPs.

## 5. The neurovariational inversion

### *a. The methodology*

The method uses the full spectrum. The goal of the variational inversion is to retrieve the AP and OC by minimizing a cost function $J$ defined in the observation space. In the present study, which is a feasibility study, we used a simple quadratic cost function of the form

$$J = \frac{1}{2}\sum_{i=1}^{n} [\rho_{\text{cor}}^{\text{obs}}(\lambda_i) - \rho_{\text{cor}}^{\text{cal}}(\lambda_i)]^2, \tag{5}$$

where $\rho_{\text{cor}}^{\text{obs}}$ is the satellite observation (the desired value), $\rho_{\text{cor}}^{\text{cal}}$ is the observation computed by using the multimodular MLP modeling Eq. (2), and $n$ stands for the number of wavelengths taken into account.

The minimization is done by a gradient-descent method which is an iterative process. In this paper we used a second-order gradient method, which showed its competitiveness in many problems. We defined the

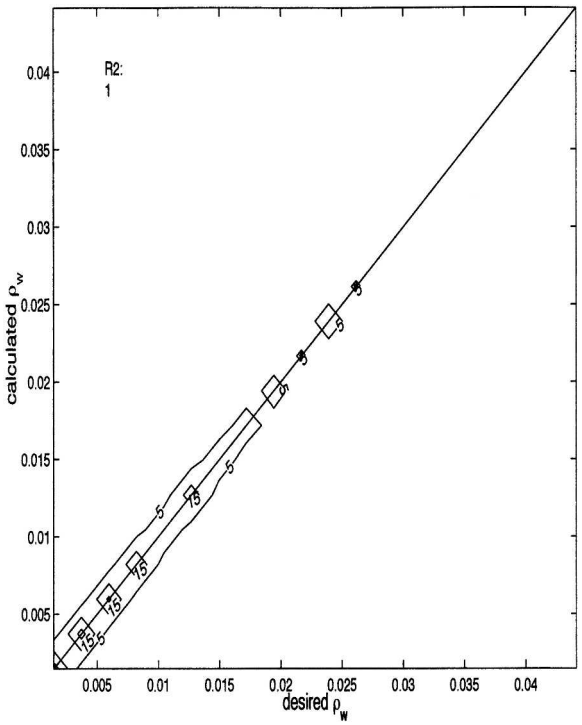


FIG. 3. As in Fig. 2 but for MLP-calculated vs desired $\rho_w$.

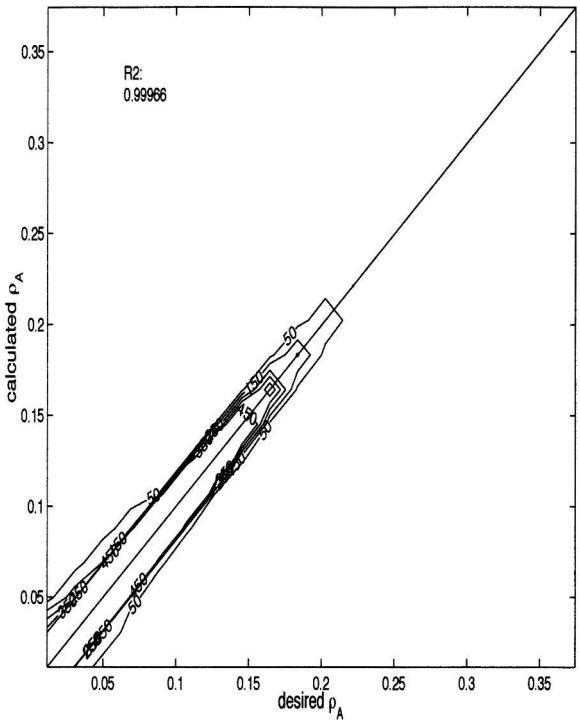


FIG. 4. As in Fig. 2 but for MLP-calculated vs desired $\rho_A$.

control vector $\mathbf{x}$ whose components are the physical parameters we wanted to retrieve (AP and OC). At each iteration we modified the control variable according to the scheme until $J$ reached a minimum value, in the first approximation:

$$x_{t+1} = \mathbf{x}_t - \epsilon \cdot \nabla J, \tag{6}$$

where $\nabla J$ is the gradient with respect to the control variable and $\epsilon$, the adaptative step gradient. The procedure was initialized by choosing a particular spectrum, the so-called first guess. In the following we will show that this initial value has a very strong effect on the retrieval.

The computation of the partial derivatives of the cost function $J$ implies the computation of the partial derivative of the direct model, which associates the physical parameters with the observations. The major advantage of this approach is that the calculation of the gradients of an MLP by an adjoint technique is straightforward, thanks to the procedure used to compute the weights of the MLP (Bishop 1995). In the present work, these direct models are the multimodular MLPs that model the RTE. The minimization of the cost function $J$ implies $n$ computed and observed reflectance values at $n$ different $\lambda$ values (wavelengths). The architecture of the minimization process is given in Fig. 5.

At the present stage of this study, we only control OC, that is, $C$ and $b^0$ and one AP, the imagery refractive index, $m_i$. Hence we used $\rho_{\text{cor}}$ at five visible wavelengths. In the following we present first the results obtained performing some academic inversion using the NLUT: we tested the robustness of the NVI following the first guess. We focused our work on the retrieval of $C$, which is the main desired parameter. Then we applied the method to SeaWiFS imagery in which the other atmospheric parameters, the size distribution parameter $\nu$, the real refractive index $m_r$, and the aerosol optical thickness at 865-nm $\tau(865)$ were determined during a preliminary phase, described in the appendix, by using an MLP direct inverse model in the NIR.

### b. Numerical results of the NVI

Section 5b(1) presents some preliminary results, in order to show the behavior of the NVI method with synthetic data. In particular, since the computation of $\rho_{\text{cor}}$ is highly nonlinear with respect to the control variables (the OC in the present case), the minimization required some preprocessing. The principal result was that the first guess for the three APs have to be carefully chosen. In section 5b(2) we discuss Sea WiFS images, the first guess of the required AP being computed with accuracy by using the IR spectrum (see the appendix).

#### 1) RETRIEVAL OF THE OCEAN PARAMETERS (OCS)

We tested the NVI using synthetic $\rho_{\text{cor}}$ given by the Test-1008 set. The goal of this study was to show that

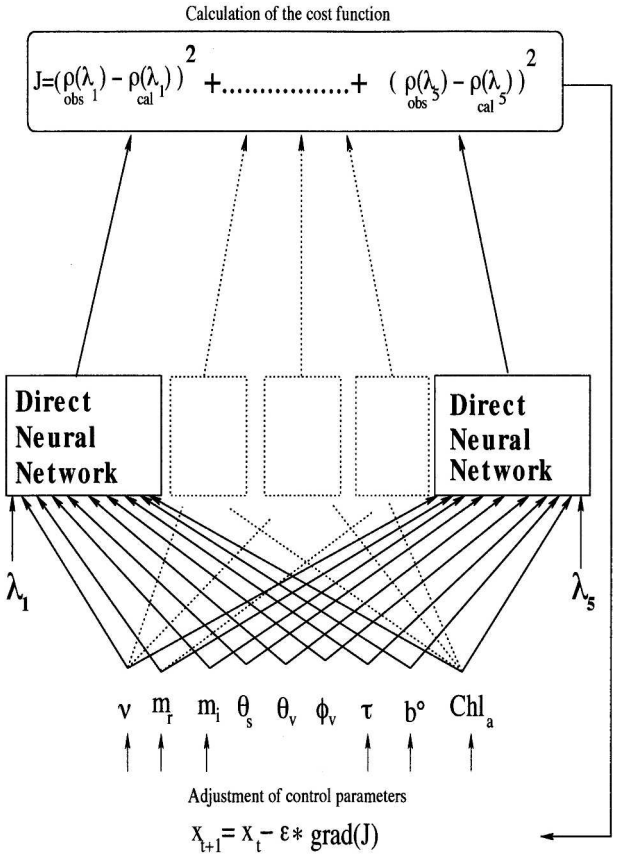


FIG. 5. Neurovariational method in the visible spectrum.

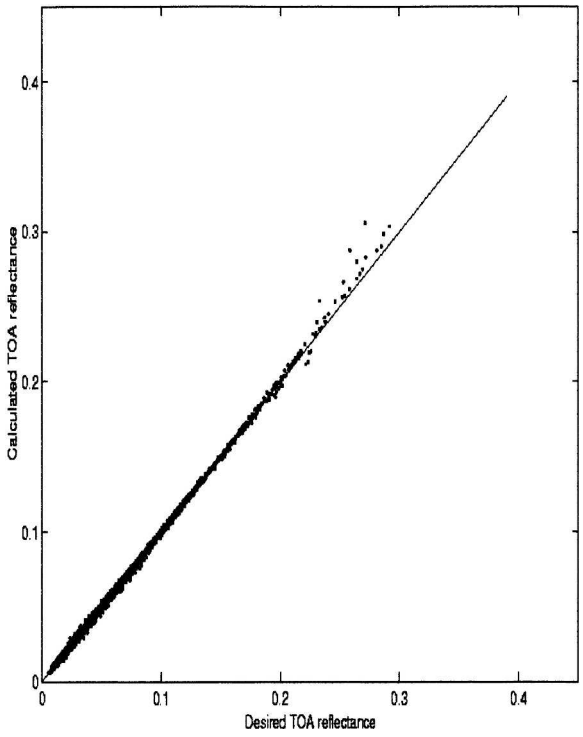


FIG. 6. Scatterplot of retrieved vs desired $\rho_{cor}$ for weakly absorbing aerosols $m_i \leq 0.003$ with $[\nu, m_r, \tau(865)]$ known exactly.

the NVI is able to generate an efficient atmospheric correction when the APs are known, allowing us to retrieve the OCs with accuracy. We therefore ran an experiment in which the AP $\nu$, $m_r$, and $\tau(865)$ are known exactly and fixed, the control variables being the OCs, namely $C$ and $b^0$ and the imagery refractive index $m_i$. In all of the experiments presented below, the first guess value of $C$ was set at 0.2 mg $m^{-3}$, which corresponds to an average value for case 1 waters, especially in the Mediterranean Sea, and $b^0$ was set at 0.285 ($m_i$ was set at 0.002).

Figure 6 shows the scatterplot of the MLP-computed versus desired $\rho_{cor}$ (MLP-computed $\rho_{cor}$ is the $\rho_{cor}$ computed with the multimodular MLP). Almost all the data were on the line of slope 1, meaning that $\rho_{cor}$ is computed with a good accuracy. The rmse and relative rmse for the computed $\rho_{cor}$, $\rho_A$, and $t^*\rho_w$ have been calculated for each visible wavelength (not shown). The rmse for $\rho_{cor}$ was $4.91 \times 10^{-4}$ for the entire dataset, and it was below the theoretical requirement for the SeaWiFS sensors ($\simeq 10^{-3}$). This comparison between the computed $\rho_{cor}$ and the desired $\rho_{cor}$ gives an idea of the effectiveness of the minimization method (the relative rmse was 0.88%). The $\rho_{cor}$ can be retrieved with a high accuracy for every wavelength, the rmse being lower than $10^{-3}$, and the maximum relative rmse is equal to 1.28% for $\lambda = 555$ nm.

The theoretical requirement for the SeaWiFS sensors was 1–2 ($\times 10^{-3}$) at 443 nm. The conclusion of this work is that the variational inversion with an adjoint technique performs well, each component of $\rho_{cor}$ being retrieved with a good accuracy for each wavelength.

Numerical performances for the retrieval of the ocean parameters OC are given in Table 3. Clearly, perfect knowledge of AP allows good retrieval of the $C$ content. The $C$ can be retrieved with a relative rmse of 19.7% on the entire test dataset when we know $\nu$, $m_r$, and $\tau(865)$. Figure 7 displays the scatterplot of the mean retrieved versus desired $C$. We could note that the weak values of $C$, that is, $\leq 0.6$ mg.$m^{-3}$, were retrieved with

TABLE 3. Rmse and relative rmse computed for Test-1008 of the retrieved parameters.

| | $\rho_{cor}$ | $m_i$ | $b$ | $C$ |
|---|---|---|---|---|
| Rmse | $4.91 \times 10^{-4}$ | $7.90 \times 10^{-4}$ | $7.51 \times 10^{-2}$ | $5.29 \times 10^{-2}$ |
| Relative rmse (%) | 0.88 | $\infty$ | 20.7 | 19.7 |

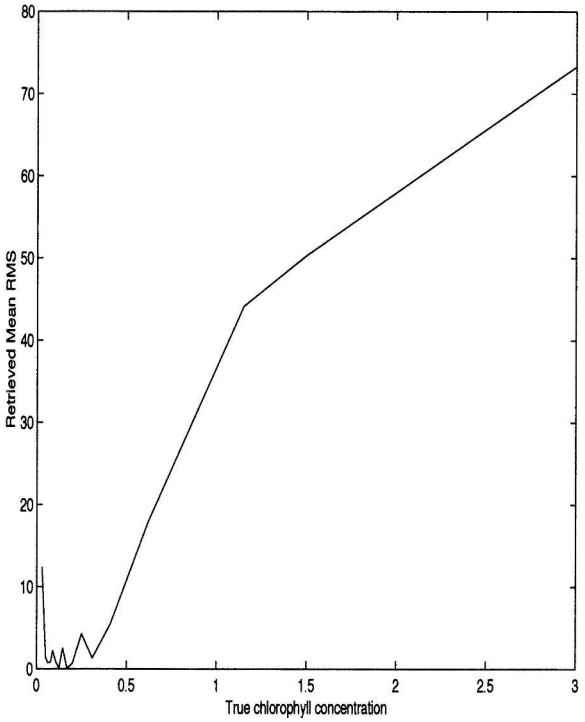


FIG. 7. Variation of the mean relative rmse vs the desired $C$.

pretty good accuracy, given that there is almost no scattering and the maximum bias in term of relative rmse was equal to 17.95%. On the other hand, for strong values of $C$, we observed an underestimation, in particular for $C = 3.0$ mg m$^{-3}$, the bias being equal to 77%. This value was very high, and it was likely that the interpolation for these high values of $C$ failed because of the weakness of the sampling of the learning dataset used to calibrate the MLP-O (four values of $C$ between 0.62 and 3 mg m$^{-3}$). Another explanation might be that $\rho_w$ was very low for high values of $C$. So, the contribution of $\rho_w$ to $\rho_{cor}$ was also weak, and the signal may be difficult to retrieve for these cases. For the total dataset, the relative rmse was 19.7% versus 35% required for SeaWiFS.

We present now additional results (Table 4) for different APs and OCs in Fig. 8. Figure 8 shows the spectra of the desired and MLP-computed $\rho_{cor}$, $\rho_A$, and $t\rho_w$. We analyzed the desired $\rho_{cor}$ with respect to the $\rho_{cor}$ computed with the fixed AP and retrieved OC for different $C$ and $\tau(865)$ (Fig. 8). For these different cases, $\rho_{cor}$ was well retrieved whatever the configurations. In terms of rmse, it was below that required for SeaWiFS, that is, $10^{-3}$. The same conclusion could be reached for each component of $\rho_{cor}$. The $\rho_{cor}$ and $t\rho_w$ were computed with the NVI retrieval values, and we noted that the accuracy on each signal was very high. This could mean that the NVI retrieval values were obtained with a pretty good accuracy too. When $\rho_A$ was the major contribution to $\rho_{cor}$ (see Figs. 8c,e), the inversion method worked well and could retrieve $C$ with a high accuracy. When the contribution of $\rho_A$ and $t\rho_w$ were equivalent (Fig. 8b), the accuracy of the retrieved parameters was also high.

TABLE 4. Configurations of aerosol optical properties and sun-viewing geometries corresponding to Figs. 8a–f for Test-1008.

| Fig. | $\nu$ | $m_r$ | $m_i$ | $\tau(865)$ | $\theta_s$ | $\theta_v$ | $d\phi_v$ |
|---|---|---|---|---|---|---|---|
| 8a | 2 | 1.33 | 0 | 0.35 | 47 | 51 | 72 |
| 8b | 2 | 1.33 | 0.001 | 0.15 | 32 | 18 | 288 |
| 8c | 3.5 | 1.33 | 0.001 | 0.35 | 44 | 15 | 72 |
| 8d | 2.5 | 1.50 | 0.003 | 0.35 | 20 | 24 | 72 |
| 8e | 3 | 1.50 | 0.003 | 0.05 | 20 | 0 | 0 |
| 8f | 3.5 | 1.50 | 0.001 | 0.35 | 44 | 15 | 72 |

Assuming that the AP $\nu$, $m_r$, and $\tau(865)$ are known, we can conclude that the NVI retrieves the $C$ with a good accuracy, even when its first-guess value is far from the true value and when $\tau(865)$ is high. A mean accuracy of 19.7% is considered an adequate solution.

### 2) RETRIEVAL OF THE ATMOSPHERIC (APS) AND OCEAN PARAMETERS (OCS)

In this section, we retrieve $\nu$, $m_r$, and $\tau(865)$ by processing the information embedded in the NIR bands (Gordon and Wang 1994a; Chomko and Gordon 1998) before handling the NVI. We performed such an inversion with dedicated MLPs working as a classifier. These MLPs gave the probability of $\nu$ and $\tau(865)$ having a given value depending on the three NIR-$\rho_{cor}$ channels and the three observation angles (see the appendix). This classification gives three [$\nu$, $\tau(865)$] pairs having different values with different probabilities depending on the value of $m_r$. As the values of [$\nu$, $m_r$, and $\tau(865)$] are retrieved with a certain error, we decided to allow these parameters to vary during the NVI within a certain range from their initial values. In the following, we present only the results for $m_r = 1.50$ and for the first couple (we have obtained the same results for another value of $m_r = 1.33$), and this discussion focused on $C$.

Figure 9 shows the scatterplot of the calculated $\rho_{cor}$ computed using the parameters obtained from the NVI versus the desired $\rho_{cor}$ for $m_r = 1.50$; $\rho_{cor}$ was relatively well retrieved. However the first guess for the AP given by the MLP classifiers using the NIR observations corresponded to the most probable ($\nu$, $m_r$) pair. This pair corresponded to the true pair in 85% of the cases. So for 15% of the solutions the first guess differed from the true values. The retrieved AP also depended on the value of $m_r$. If the chosen value differs from the true value, the first guess can also be far from the desired value. Nevertheless, in most cases, $\rho_{cor}$ can be retrieved with good accuracy. The error of the retrieved $\rho_{cor}$ and AP and OC are summarized in the Table 5. For the entire dataset, the rmse for the $\rho_{cor}$ was $4.70 \times 10^{-3}$ and the relative rmse was 7.32%. For the retrieved $C$, the relative rmse was 51.80%. In the preceding subsection,

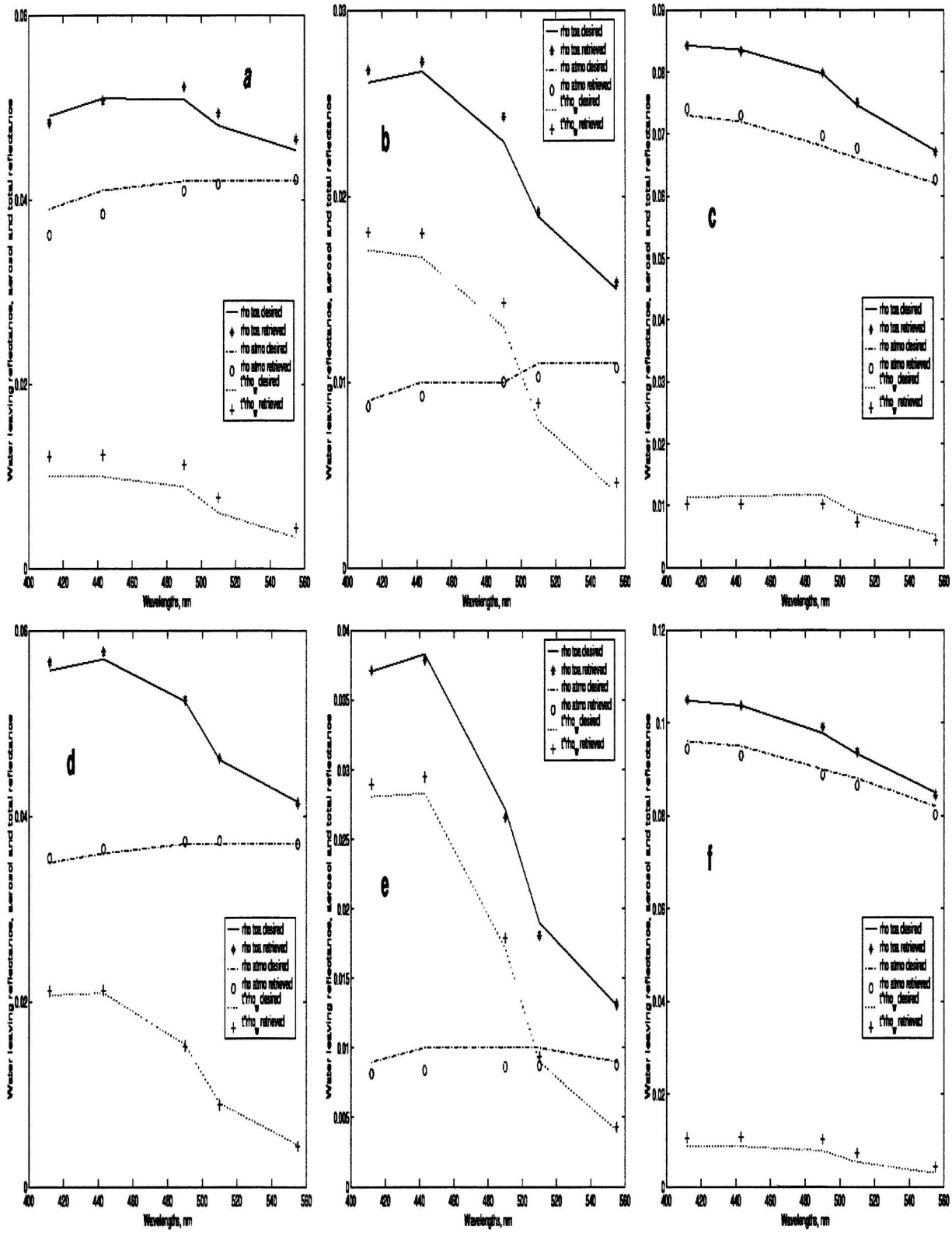


FIG. 8. Comparison between the desired (curves) $\rho_{cor}$ and the computed (dots) $\rho_{cor}$ with the retrieved atmospheric and oceanic parameters in the case of weakly absorbing aerosols for different geometries, type of aerosols, and $C$.

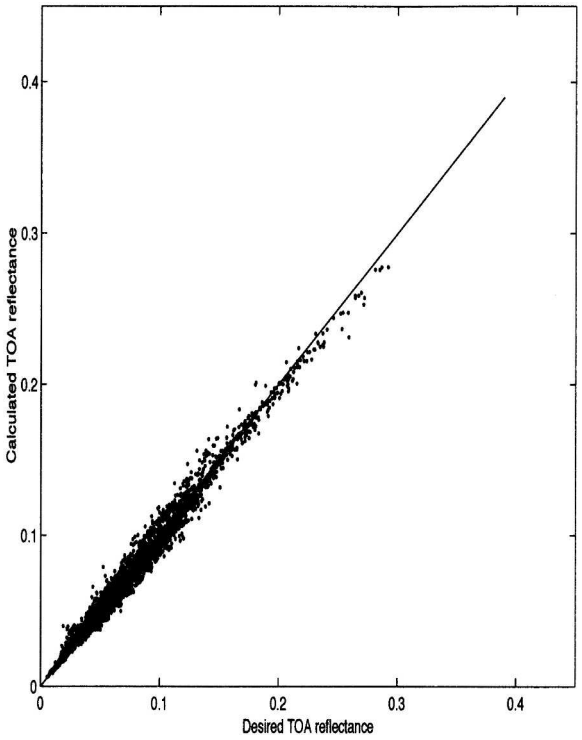


FIG. 9. Scatterplot of MLP-calculated vs desired $\rho_{cor}$ with a first guess.

we showed that we are able to retrieve $C$ with an error of 19.7% if we have a perfect knowledge of the atmospheric parameters $\nu$, $m_r$, and $\tau$ (865). In the present experiment, we estimated these parameters using the information contained in the observed spectrum in NIR $\rho_{cor}$. The performance of the NVI was degraded due to the fact that the first guess of the AP was estimated with some noise. We now explore the variation of the error following the desired $C$ in order to validate our inversion.

Figure 10 presents the variation of the mean retrieved $C$ with respect to each desired values of $C$. The mean retrieved $C$ values were close to the desired $C$ when $C$ was $\leq 0.6$ mg m$^{-3}$; that is, the bias is weak. For this range, the bias varied between 0.25% and 86.08%, leading to an overall mean bias of 26.8%. Therefore, for these values of $C$, the NVI gave accurate retrievals. For $C$ greater than 0.6 mg m$^{-3}$, the mean relative rmse was higher: it varied between 32.79% and 80.14%. We also found the same behavior of the NVI algorithm for strong $C$.

These good performances show that the minimization procedure is effective. With the retrieved AP and OC, we were able to model $\rho_{cor}$ with good accuracy. We have tested the NVI for weakly absorbing aerosols, considering that we could obtain a first guess for $\nu$, $m_r$, and $\tau$(865) with an MLP functioning in a classifier mode dealing with the three NIR bands. Using this first guess, we retrieved $C$ with a global error of 51.80%. The performance seems quite good for low values of $C$, which corresponds to the main situation in the ocean.

### 3) INVERSION OF A SERIES OF SEAWIFS IMAGES

In this section, we present a first inversion of a series of SeaWiFS images. We used a composite SeaWiFS image produced by averaging a series of SeaWiFS images of the Mediterranean Sea during 8–14 April 1999. Comparisons were made between the image processed by the Sea-WiFS algorithm (Gordon and Wang 1994a; McClain et al. 2000) and the NVI algorithm. We decided to show a composite image to show the efficiency of the NVI algorithm over several following days. Then a first validation is presented with a comparison of the NVI and SeaWiFS retrieval chl-$a$ in relation to in situ data obtained during the Productivity of Oceanic Pelagic Systems (PROSOPE) campaign.

The AP $\nu$, $m_r$, and $\tau$ (865) and the $C$ and $b^0$ were the control variables of the NVI procedure. In this part, we have compared the chlorophyll-$a$ concentration instead of the pigment concentration $C$, which is related to chl-$a$ by the following formula:

$$C = 1.34 \text{ chl-}a^{0.983}. \tag{7}$$

The first-guess value for chl-$a$ was equal to 0.2 mg m$^{-3}$ for each pixel of the image. This value represents an average value for the Mediterranean Sea. The value of this parameter has no effect on the quality of the final retrieved value. The first-guess values of the AP $\nu$, $m_r$, and $\tau$(865) were provided by the MLP inversion described in the appendix. As in the previous experiment, only the most probable pair of ($\nu$, $\tau$) corresponding to $m_r = 1.33$ was chosen. To avoid unrealistic values for the atmospheric parameters, we added a penalty term to the cost function in the form of a restoring to a background value for the AP. In the present experiment, these background values were the first-guess values of the AP. The cost function is therefore of the form

$$J = \alpha \frac{1}{2} \sum_{i=1}^{n} [\rho_{cor}^{obs}(\lambda_i) - \rho_{cor}^{cal}(\lambda_i)]^2 + \beta \sum_{i=1}^{n} (x_{cal} - x_{fg})^2, \tag{8}$$

TABLE 5. Rmse and relative rmse of the retrieved parameters for $m_r = 1.50$.

| | $\rho_{cor}$ | $\nu$ | $m_r$ | $m_i$ | $\tau$(865) | $b^0$ | $C$ |
|---|---|---|---|---|---|---|---|
| Rmse | $4.70 \times 10^{-3}$ | $2.25 \times 10^{-1}$ | $1.01 \times 10^{-1}$ | $1.54 \times 10^{-3}$ | $0.5 \times 10^{-2}$ | $1.3 \times 10^{-1}$ | 0.69 |
| Relative rmse (%) | 7.32 | 5.24 | 6.06 | $\infty$ | 14.72 | 45.81 | 51.80 |

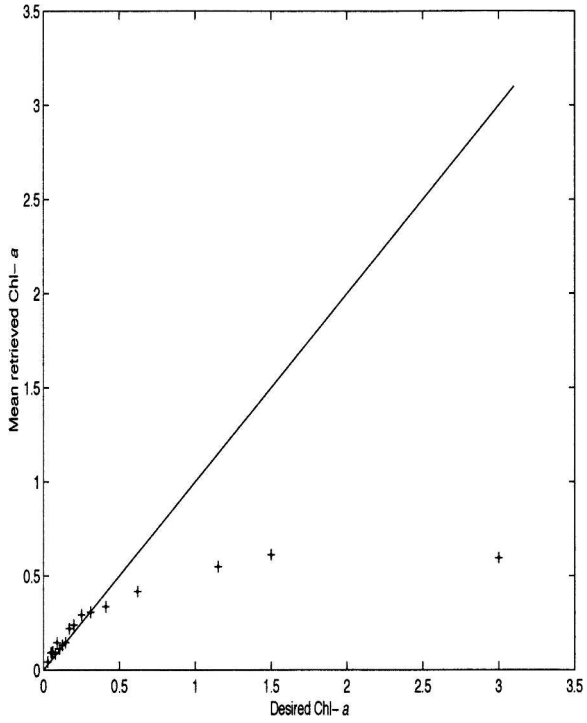


FIG. 10. Case of a first guess: scatterplot of the mean retrieved vs desired $C$.

where $x_{cal}$ stands for the parameters calculated at each iteration and $x_{fg}$ stands for the parameters obtained with the NIR inversion named first guess; $\alpha$ and $\beta$ are weights.

*(i) Spatial pattern comparison*

We present here the retrieved values of $\tau(865)$ (Fig. 11) and chl-*a* (Fig. 12). At the top of these figures, the SeaWiFS-retrieved quantities are shown, and the NVI ones are at the bottom.

The first analysis consisted in a comparison of the spatial extent of $\tau(865)$ (Fig. 11) and chl-*a* (Fig. 12). It could be a first indication of the efficiency of the atmospheric correction process. In the case of the NVI method, we noted that there was no correspondence between $\tau(865)$ and chl-*a* spatial extent, especially in the zone where the aerosol concentration is pretty strong, that is, $>0.2$, in particular near the Lybian coast. The SeaWiFS-$\tau(865)$ and NVI-$\tau$ (865) map presented a close similarity except near the Lybian coast. In this zone, the strong $\tau(865)$ spatial extent was not retrieved with the SeaWiFS algorithm. The reason was that it corresponded to a dust aerosol event for 10 April (not presented), and the SeaWiFS algorithms do not take into account these aerosol models. When looking at the wind maps provided by Meteo-France for this date we noted that the wind was blowing from the Sahara Desert to the sea. This wind direction caused the dust event. As the NVI algorithm does not flag these aerosol models, we were able to observe the spatial pattern of the dust event contrary to SeaWiFS algorithm. The NVI method is able to retrieve the high values of $\tau(865)$ associated with the dust event. But due to the quite low values of $\tau(865)$ in the NLUT used to train the MLP for the inversion, we were not able to retrieve values of $\tau(865)$ greater than 0.35. Despite these strong values of $\tau(865)$, the NVI retrieval values of chl-*a* seemed not influenced by these aerosols. The NVI was able to deal with these aerosols, thanks to the variation of $m_i$, the related absorption parameter.

We then analyzed the retrieved SeaWiFS chl-*a* and NVI chl-*a* shown in Fig. 12. We can first note that there was good similarity between both images in terms of spatial and form extent, in particular the raising of the chl-*a* in the Strait of Gibraltar (Alboran Sea), in the Gulf of Lion, along the south coast of Sicily, and in the Black Sea was retrieved by the NVI. Note that the results in the Black Sea were not detailed due to the sea proprieties (case 2 waters). The NVI chl-*a* evolution follows those of SeaWiFS, whatever the region. However, we note that globally the intensity of NVI chl-*a* was weaker than those of SeaWiFS chl-*a*, in particular for the strong values. That was the case in the Gulf of Lion and to a lesser extent in the Alboran Sea. For these two regions, there was a strong variation of the chl-*a*. NVI was able to follow the variation, in particular for the Alboran Sea. The bloom extent in the Gulf of Lion was very well retrieved, only the intensity was different. For the strong chl-*a*, the NVI algorithm seemed not able to retrieve them with a correct accuracy. One reason might be the weakness of the bio–optical algorithm, those of Gordon et al. (1988b) concerning values of chl-$a \geq 1$ mg $m^{-3}$. For the other part of the Mediterranean Sea, the chl-*a* varied weakly and there were no blooms; thus we observed good concordance of the spatial extent and the intensities. But, if mean values of a small region were calculated (not shown), we noted that the NVI chl-*a* were, mostly, weaker than the SeaWiFS chl-*a*. Previous studies (Bricaud et al. 2002; D'Ortenzio et al. 2002) have compared the SeaWiFS chl-*a* with in situ data in the western and central Mediterranean Sea and all authors have reached the same conclusion. Even with a local biooptical algorithm, the SeaWiFS chl-*a* were strongly overestimated. This overestimation was due to the atmospheric correction algorithm. In our case, in all oligotrophic parts of the Mediterranean Sea, that is, chl-$a \leq$ 0.2 mg $m^{-3}$, the NVI chl-*a* were weaker than those of SeaWiFS, and so more realistic and closest to the in situ data. We could therefore believe that this improvement came from the atmospheric correction of the NVI algorithm. This belief could be assessed with a comparison of a few in situ data, presented in the following paragraph.

*(ii) Validation with in situ data*

A first validation has been done with in situ data collected during the PROSOPE campaign. The PROSOPE

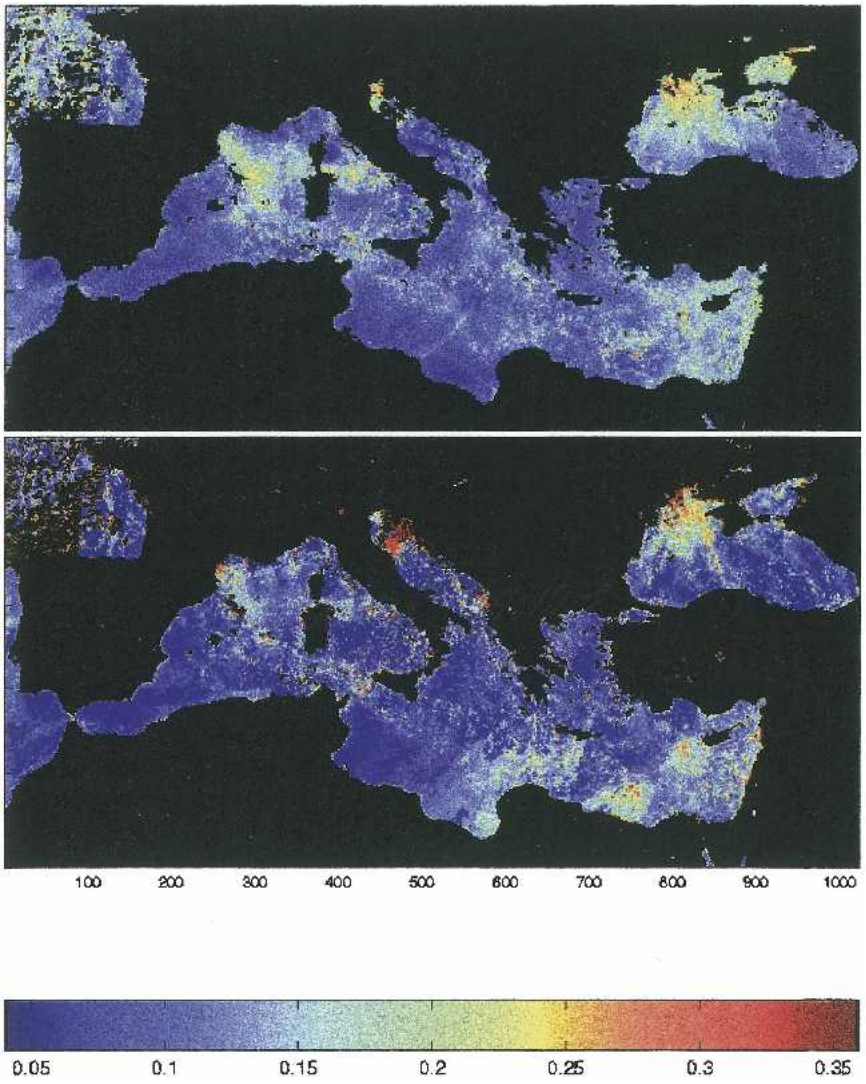


FIG. 11. Optical thickness $\tau(865)$ retrieved for the second week of Apr 1999 with (top) the SeaWiFS algorithm and (bottom) the NVI.

cruise took place from 4 September (Agadir, Morocco) to 4 October (Toulon, France) 1999 aboard the RV *Thalassa*. The cruise path corresponds to the western part of the Mediterranean Sea. Only seven in situ data were available during this period for the validation due to the cloud cover and satellite daily time of passage. The in situ data were collected for several depths but only data obtained between 0 and 20 m were taken into account. The mean values and the variance were calculated in this interval for each day. The comparison of in situ measurements with satellite retrievals was made by extracting chl-*a* computed from both SeaWiFS and NVI algorithms over a grid of 3 by 3 pixels centered on the position of the research vessel. The mean and standard deviation were calculated on this grid.

Figure 13 presents the retrieval values of SeaWiFS and NVI chl-*a* versus the PROSOPE in situ data. The diagonal line represents the slope 1 line. For moderate and strong values of chl-*a*, only one measurement in situ data was available. For this data the NVI value was weaker than the SeaWiFS value. The NVI retrieval showed an underestimate of chl-*a*, particularly for bloom events (Fig. 13). The underestimation could be due to the bio–optical model chosen for this study. This was related to the underestimation observed previously for the entire Mediterranean Sea.

For oligotrophic waters, the NVI algorithm allowed one to retrieve chl-*a* with better accuracy than those obtained with SeaWiFS algorithm. There was a net improvement of the retrieval values with the NVI algorithm for these waters. Bricaud et al. (2002) have compared the SeaWiFS chl-*a* product to in situ data, among them the PROSOPE data, and concluded that even with a regional biooptical model there was still an

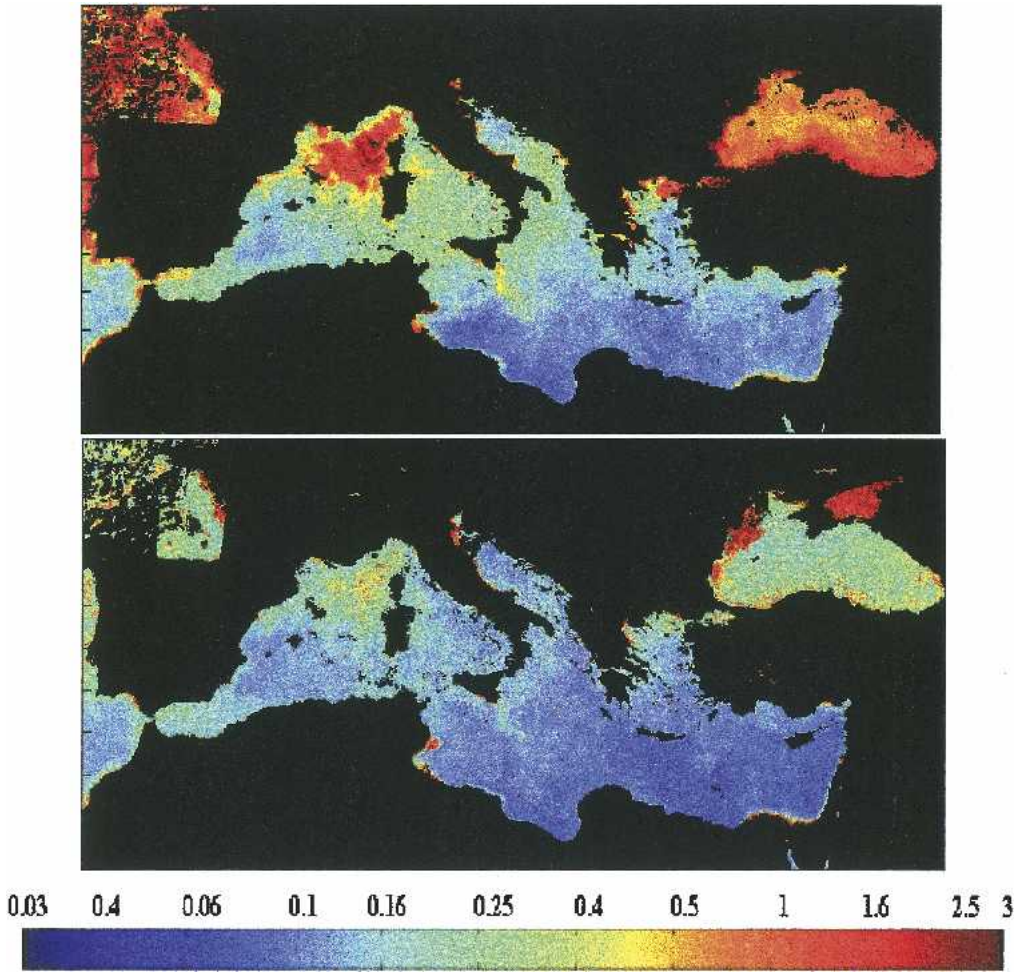


FIG. 12. Chl-*a* retrieved for the second week of Apr 1999 with (top) the SeaWiFS algorithm and (bottom) the NVI.

overestimation and that this gap was due to the atmospheric correction algorithm. In our case, there was no reason that the Gordon et al. biooptical model has been different from the OC4V4 SeaWiFS biooptical model (O'Reilly et al. 1998), the two biooptical models not being calibrated in the Mediterranean Sea. We could think that the improvement in the retrieval chl-*a* values by the NVI algorithm in the oligotrophic waters for this dataset came from the improvement of the atmospheric correction algorithm, which was the goal of the NVI method.

## 6. Conclusions

We have developed a new method to process ocean-color images provided by satellite sensors. The retrieval of ocean constituents from these measurements is difficult because of the presence of the atmosphere, which represents at least 90% of the reflectance received by the sensor.

The neurovariational inversion combines the advantages of neural networks and the variational estimation. The variational method consists in inverting a direct model (a radiative transfer model in the present case) to retrieve the aerosol and oceanic parameters by minimizing the distance between the observed reflectance at the top of the atmosphere and the calculated one at several wavelengths, the control parameters being the AP and the OC. The direct model is represented by MLPs, which are universal approximators and represent continuous and derivable functions. Their derivatives and consequently the adjoint of the direct model are easy to compute. The accuracy of the MLPs modeling the RTE is better than 7% in terms of relative rmse.

First we studied the feasibility of the method by inverting a synthetic dataset for which the AP are given. The first-guess value for $C$ is 0.2 mg $m^{-3}$, whatever its actual value. This first-guess value might be far from the true value and can be considered as a severe con-

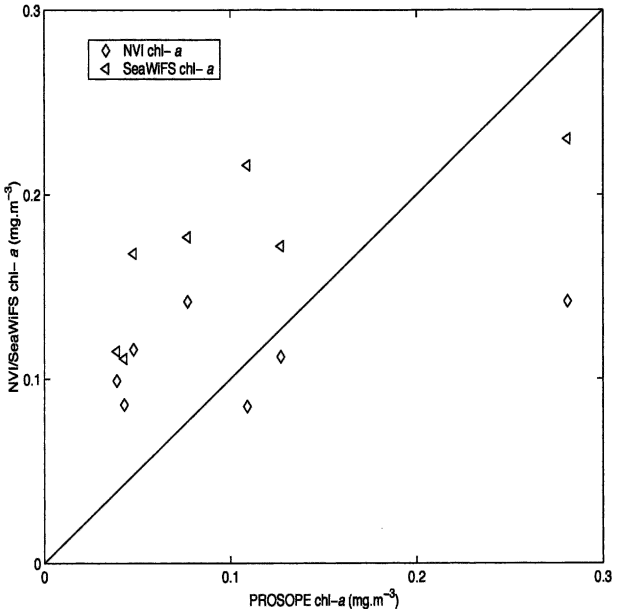


FIG. 13. Scatterplot of the NVI and SeaWiFS retrieved chl-*a* vs PROSOPE in situ data.

dition for the inversion. The results showed that we can retrieve $C$ with an error of 19.7% if we know exactly $\nu$, $m_r$, and $\tau(865)$, proving the good functioning of the NVI.

Second, still using the synthetic dataset, we estimated the AP $\nu$, $m_r$, and $\tau(865)$ with specific MLPs working in classifier mode. These MLPs process the TOA NIR information. The estimated AP were introduced as first guesses in the NVI. The first-guess value for $C$ was still 0.2 mg m$^{-3}$. In fact, during the minimization only $C$ was active. The AP moved slightly around their first-guess values, probably due to the good quality of the MLP inversion. Using these first-guess values for the AP, $C$ can be retrieved with a relative error of 51.8% within a range of [0.03, 3 mg cm$^{-3}$], but for $C \leq 0.6$ mg cm$^{-3}$, the retrieval error was under 30%, showing the efficacity of the method.

Finally, we have processed a composite SeaWiFS image for the Mediterranean Sea during the period 8–14 April 1999. The AP $\nu$, $m_r$, and $\tau(865)$ are still estimated with specific MLPs working in a classifier mode. In order to constrain the AP given by the NVI, a penalty term was added to the cost function in the form of a quadratic restoring it to the first guess (background in the variational optimization). The first-guess value for chl-*a* is fixed at 0.2 mg m$^{-3}$, which represents an average value for the Mediterranean Sea but this value had no influence on the quality of the retrievals. The NVI retrieves the same spatial pattern of $\tau(865)$ and chl-*a* as that provided by the SeaWiFS algorithms, but the quantitative values are somewhat different, especially for the bloom events. The NVI chl-*a* values are always weaker than those provided by the SeaWiFS algorithm. This was assessed with a final validation with in situ data for the western Mediterranean. Actually the chl-*a* values computed by the SeaWiFS algorithm are always larger than those provided by the in situ measurements in the Mediterranean Sea (Claustre et al. 2002; Bricaud et al. 2002) in the oligotrophic sea, and this overestimation is due to a lack in the atmospheric correction process. Thus the NVI algorithm would give more realistic values for chl-*a* in oligotrophics conditions, and it validated the method developed to improve the atmospheric correction process.

The NVI methodology might be improved by increasing the range of variation of the AP during the calibration of the different MLPs. In the present paper we focused on the feasibility of the NVI method. It offers the possibility of taking into account the absorbing aerosols in a rational manner. This is under investigation. Furthermore, the NVI method also offers the possibility of inverting an image globally and not pixel by pixel; the global inversion allows us to take into account the spatial context, which should dramatically improve the quality of the inversion, at least theoretically.

*Acknowledgments.* This work was supported by the European Community NAOC Programme (Contract EVG1-CT-2000-00034) and the firm ACRIst. We thank Dr. H.R. Gordon of the University of Miami for providing the database and the PROSOPE project for the use of their data.

## APPENDIX

### Classification with MLP

MLP can be used as a classifier (Badran et al. 1991; Thiria et al. 1993; Bishop 1995): the desired answer $y_k$ is now the vector indexing the class $c_i$ of the input $x_k$. The problem is to classify the input pattern $x_k$ in $p$ different classes. The response $y_k$ was the form $y_k = [y_{1k}, \ldots, y_{pk}]$ with $y_{ik}=+1$ if $x_k$ belongs to class $w_i$ and $y_{ik} = 0$ for $j \neq i$.

With the different classes obtained, it is possible to calculate the probability of each class. For a class $c_i$ corresponding to $x_i$, the probability is

$$P(c_i) = \frac{e^{ci}}{\sum_{k=1}^{n} e^{ck}}, \tag{A1}$$

where $n$ is the number of classes. The classifier gives the probability of each class being the true class.

As we need a first guess for the neurovariational inversion in the visible bands, we thus decided to perform a first inversion by processing the red and near-infrared SeaWiFS bands (670, 765, 865 nm), where the sea does not influence the TOA signal ($\rho_w$ can be neglected for case 1 waters). This enables us to get a reasonable first

guess for two key aerosol parameters [the aerosol size distribution $\nu$ and the aerosol optical thickness $\tau(865)$ at 865 nm] following a specific value of $m_r$ because the spectral variations of $\rho_A$ in the NIR depend mostly on these three parameters. The objective of this first step is therefore to find the best combination of $\nu$, $\tau(865)$, and $m_r$ that best matches the measured $\rho_{\text{cor}}$. We trained four dedicated MLP classifiers, one for $\nu$ and another for $\tau(865)$ for two specific values of $m_r$ (1.33 and 1.50). These MLPs are displayed in Fig. A1. We used six classes of $\nu$ for the $\nu$ classifiers and four classes of $\tau(865)$ for the $\tau(865)$ classifiers. Each classifier gives the probability of each class to be the true class. To retrieve the best combination of $[\nu, m_r, \tau(865)]$, the calculation of the cross-probabilities between each combination of $\nu$ and $\tau(865)$ is done for each $m_r$ value. This allows us to retrieve the three best pairs of $[\nu, \tau(865)]$ together with their associated probability level. The six best solutions {three pairs $[\nu_i, \tau(865)_i]$ for each value of $m_r$} are then retained to be used as first guesses for the second part of the inversion described below.

Each MLP classifier has two hidden layers of ten neurons each. The inputs are the three NIR reflectances, $\rho_{\text{cor}}$, and the sun-viewing geometries of the measurement ($\theta_s$, $\theta_v$, and $d\phi_v$).

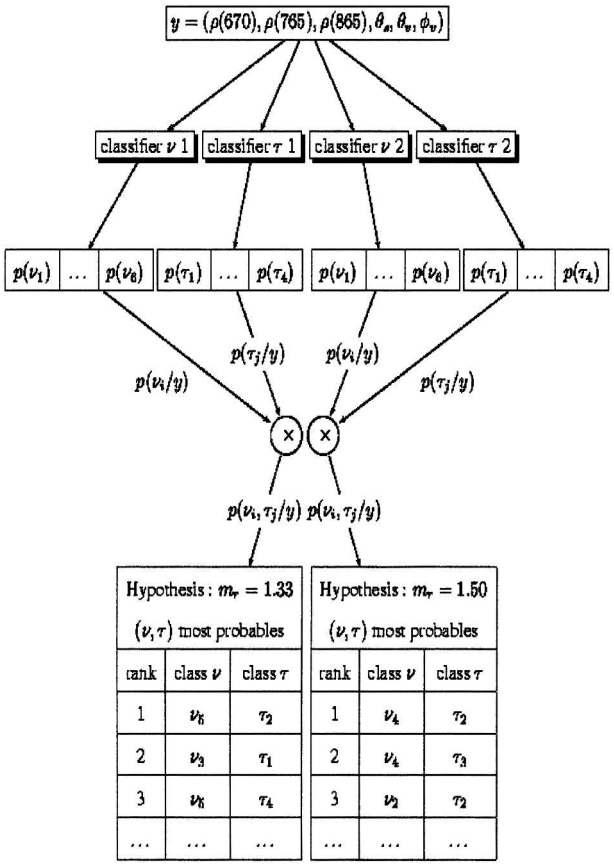


FIG. A1. Diagram of the MLP classification in the NIR.

TABLE A1. Cross-probability of the couple $[\nu, \tau(865)]$.

| | $m_r$ | |
|---|---|---|
| | 1.33 | 1.50 |
| $[\nu, \tau(865)]$true = $[\nu, \tau(865)]$calc 1 | 85.01% | 89.21% |
| $[\nu, \tau(865)]$true = $[\nu, \tau(865)]$calc 2 | 12.47% | 9.63% |
| $[\nu, \tau(865)]$true = $[\nu, \tau(865)]$calc 3 | 1.38% | 0.73% |

The calculation of the cross-probabilities between each combination of $\nu$ and $\tau(865)$ (24 in total) makes possible the determination of the most probable pair $[\nu, \tau(865)]$. Table A1 shows the probability of the three first couples. The probability to obtain the true couple in the three first couples is 98.86% for $m_r = 1.33$ and 99.57% for $m_r = 1.50$. When considering only the first best couple, we retrieve the true couple $[\nu, \tau(865)]$ with a probability of 85.01% for $m_r = 1.33$ and 89.21% for $m_r = 1.50$.

The performance of the direct inversion with MLP in the NIR is very satisfactory and leads to a good accuracy of the retrieval parameter $\nu$ and $\tau(865)$. These results therefore provide a satisfactory first guess for the inversion in the visible.

## REFERENCES


Badran, F., S. Thiria, and M. Crepon, 1991: Wind ambiguity removal by the use of neural networks techniques. *J. Geophys. Res.,* **96C,** 20 521–20 529.

Bishop, C. M., 1995: *Neural Networks for Pattern Recognition.* Oxford University Press, 482 pp.

Bricaud, A., E. Bosc, and D. Antoine, 2002: Algal biomass and sea surface temperature in the Mediterranean basin intercomparison of data from various satellite sensors, and implications for primary production estimates. *Remote Sens. Environ.,* **81,** 163–178.

Chomko, R., and H. R. Gordon, 1998: Atmospheric correction of ocean color imagery: Use of a Junge power-law size distribution with variable refractive index to handle aerosol absorption. *Appl. Opt.,* **37,** 5560–5572.

Claustre, H., and Coauthors, 2002: Is desert dust making oligotrophic water greener? *Geophys. Res. Lett.,* **29,** 1469, doi: 10.1029/2001GL014056.

Cybenko, G., 1989: Approximation by superposition of a sigmoidal function. *Math. Control Signal Syst.,* **2,** 303–313.

Deschamps, P. Y., M. Herman, and D. Tanre, 1983: Modeling of the atmospheric effects and its application of the remote sensing of ocean color. *Appl. Opt.,* **22,** 3751–3758.

D'Ortenzio, F., S. Marullo, M. Ragni, M. Ribera d'Alcalā, and R. Santoleri, 2002: Validation of empirical SeaWiFS algorithms for chlorophyll-a retrieval in the Mediterranean Sea: A case study for oligotrophic seas. *Remote Sens. Environ.,* **82,** 79–94.

Gordon, H. R., 1997: Atmospheric correction of ocean color imagery in the Earth Observing System era. *J. Geophys. Res.,* **102,** 17 081–17 106.

——, and M. Wang, 1994a: Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: A preliminary algorithm. *Appl. Opt.,* **33,** 443–452.

——, and ——, 1994b: Influence of oceanic whitecaps on atmospheric correction of ocean-color sensors. *Appl. Opt.,* **33,** 7754–7763.

——, 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 between ship

determinations and Coastal Zone Color Scanner. *Appl. Opt.,* **22,** 20–36.

——, J. W. Brown, and R. H. Evans, 1988a: Exact Rayleigh scattering calculations for use of the Nimbus 7 coastal zone color scanner. *Appl. Opt.,* **26,** 2111–2122.

——, O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker, and D. K. Clark, 1988b: A semi-analytic radiance model of ocean color. *J. Geophys. Res.,* **93,** 10 909–10 924.

——, T. Du, and T. Zhang, 1997: Remote sensing of ocean color imagery and aerosol properties: Resolving the issue of aerosol absorption. *Appl. Opt.,* **36,** 8670–8684.

Hornik, K., M. Stinchcomb, and F. E. Muller-Karger, 1989: Multilayer feedforward networks are universal approximators. *Neural Networks,* **2,** 359–366.

Junge, C. E., 1958: *Atmospheric Chemistry Advances in Geophysics,* Vol. 4, Academic Press, 1–108.

McClain, C. R., E. J. Ainsworth, R. A. Barnes, R. E. Eplee Jr., F. S. Patt, W. D. Robinson, M. Wang, and S. W. Bailey, 2000: SeaWiFS postlaunch calibration and validation analyses, Part 1. NASA Tech. Memo 206982, 85 pp.

O'Reilly, J. R., and Coauthors, 1998: Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.,* **103,** 24 937–24 953.

Pinkus, A., 1999: Approximation theory of the MLP model in neural networks, *Acta Numer.,* 143–195.

Shettle, E. P., and R. W. Fenn, 1979: Models of the atmospheric aerosols and their optical properties. Air Force Geophysical Laboratory (AFGL), Rep. AFGL-TR-79-0214, Hanscomb Airforce Base, MA.

Thiria, S., C. Mejia, F. Badran, and M. Crepon, 1993: A neural network approach for modelling transfer functions: Application for wind retrieval from spaceborne scatterometer data. *J. Geophys. Res.,* **98,** 22 827–22 841.

Yang, H., and H. R. Gordon, 1997: Remote sensing of ocean color: Assessment of the water-leaving radiance bidirectional effects on the atmospheric diffuse transmittance. *Appl. Opt.,* **36,** 7887–7897.