MEETC2: Ocean color atmospheric corrections in coastal complex waters using a Bayesian latent class model and potential for the incoming sentinel 3 - OLCI mission

Bertrand Saulquin, Ronan Fablet, Ludovic Bourg, Grégoire Mercier, Odile Fanton d'Andon

▶ To cite this version:

Bertrand Saulquin, Ronan Fablet, Ludovic Bourg, Grégoire Mercier, Odile Fanton d'Andon. MEETC2: Ocean color atmospheric corrections in coastal complex waters using a Bayesian latent class model and potential for the incoming sentinel 3 - OLCI mission. Remote Sensing of Environment, 2016, 172, pp.39 - 49. ⟨10.1016/j.rse.2015.10.035⟩. ⟨hal-01611644⟩

HAL Id: hal-01611644

https://hal.science/hal-01611644v1

Submitted on 6 Oct 2017

MEETC2: Ocean Color Atmospheric corrections in coastal complex waters using a Bayesian latent class model and potential for the incoming Sentinel 3 - OLCI mission.

Bertrand Saulquin (1, 2), Ronan Fablet (2, 3), Ludovic Bourg (1), Grégoire Mercier (2, 3), Odile Fanton d'Andon (1).

(1) ACRI-ST, Sophia-Antipolis, 260 route du Pin Montard, BP 234 06904 Sophia-Antipolis, France

(2) Institut Mines-Telecom, Télécom Bretagne; UMR CNRS 3192 Lab-STICC, Technopôle Brest Iroise CS 83818, 29238 Brest, France

(3) Université Européenne de Bretagne, 35000 Rennes, France

Corresponding author: bertrand.saulquin@acri-st.fr

Submission date: 10/06/2015

Abstract

From top-of-atmosphere (TOA) observations, atmospheric correction for ocean color inversion aims at distinguishing atmosphere and water contributions. From a methodological point of view, our approach relies on a Bayesian inference using Gaussian Mixture Model prior distributions on reference spectra of aerosol and water reflectance. A reference spectrum for the aerosol characterizes the specific signature of the aerosols on the observed aerosol reflectance. A reference spectrum for the water characterizes the specific signature of chlorophyll-a, suspended particulate matters and colored dissolved organic matters on the observed sea surface reflectance. In our Bayesian inversion scheme, prior distributions of the marine and aerosol variables are corrected using the observed values of covariates, typically acquisition geometry acquisition conditions and pre-estimates of the aerosol and water reflectance in the near-infrared part of the spectrum, to optimize the 25 random initializations for our MEETC2 algorithm.

We evaluate our estimates of the sea surface reflectance from the MERIS TOA observations. Using the MERMAID radiometric in-situ dataset, we obtain significant improvements in the estimation of the sea surface reflectance, especially for the 412, 442, 490 and 510 nm bands, compared with the standard ESA MEGS algorithm and the a state-of-the-art neural network approach (C2R). The mean gain value on the relative error for the 13 bands between 412 and 885 nm is of 57% compared with MEGS algorithm and 10% compared with the C2R. The water leaving reflectances are used in Ocean Color for the estimation of the chl-a concentration, the colored dissolved organic matters absorption and the suspended particulate matters concentration underlying the potential of such approach to improve the standard level 2 products in coastal areas. We further discuss the potential of MEETC2 for the incoming OLCI / Sentinel 3 mission that should be launched in 2015.

Index term: 1) Atmospheric corrections in complex coastal waters. 2) Ocean color. 3) Bayesian inversion. 4) Gaussian Mixture Model.

1 Introduction

The inversion of Ocean Color signal in coastal areas from top-of-atmosphere (TOA) measurements remains a scientific challenge. This is a crucial point for the ocean color community as many governmental policies such as the European Water Framework directive (WFD) rely on estimation of coastal water quality, itself possibly derived from space-based ocean-color measurements (http://ec.europa.eu/environment/water/water-framework/index_en.html). Hence, ocean color inversion is certainly among highest priority research topics for ocean–color community. Different aspects may explain the difficulties encountered in this inversion process. Firstly, the contribution of suspended matters to the reflectance in the near infrared (700-900 nm) is an issue as many algorithms expect these reflectances to be null. This assumption is called the black pixel hypothesis and relies on the strong natural absorption of the water in this domain (Antoine et al. 2006, be Gordon and Wang, 1994). Secondly, bio-optical modelling, i.e. the estimation of the water-leaving reflectance from the Inherent Optical Properties (IOPs, namely the absorption and backscattering of the sea water constituents) in complex coastal waters is also challenging. Despite accurate physical models exist for open clear waters that cover 85% of the oceans, their derivation for coastal waters is more complex (Maritorena et al., 2002; IOCCG Report 3&5, 2000). Lastly, aerosol and water reflectance spectra may show important correlation in the near infrared, a spectral domain typically used by the standard algorithms to distinguish the two contributions.

As a consequence, available operational standard level-2 reflectance products may perform poorly in coastal areas, and consequently these products are often flagged as anomalous values for such areas (MERIS DPM, 2005). The result for end users is typically that very few observations are available in coastal areas if the standard quality standards are applied. For available pixels in coastal turbid waters, reflectances in the blue and green bands are often underestimated and may involve physically-meaningless negative values (Jamet et al., 2011; Goyens et al., 2013). Obviously Park et al.(2004) show this strongly affects relevance of level-2 products for the end users, which typically use water-reflectance spectra as inputs to estimate the chlorophyll-a and the suspended particulate matter concentrations (SPM, Doxaran et al. 2002), or the vertical light attenuation (Saulquin et al. 2013; Morel et al. 2006, Wang et al., 2009; Jamet et al., 2012).

Over the last fifteen years, many coastal water algorithms have been developed to address user's needs for reliable water-reflectance data in coastal areas. Among them, the Schiller & Doerffer

(1999) MERIS Case2-Regional (C2R) based on a non-linear machine learning model, namely a Neural Networks (NN, Krasnopolsky & Schiller, 2003; Schiller & Doerffer, 1999), estimates water reflectance over turbid areas. The learning paradigm relies on the calibration of a non-linear model to relate the available satellite-derived observations to the geophysical quantity of interest from a training dataset. This training dataset typically consists of a collection of in-situ measurements along with the satellite-derived measurements. This learning-based strategy may suffer from two major drawbacks: weak geophysical/biological interpretability of this 'black-box' model and the assumption on the representativity of the training dataset. They may restrict the applicability of the model to a specific region and questions its validity with respect to the generally unknown variability of the atmospheric and water conditions.

Here, we develop a Bayesian latent class approach to address these limitations. To our knowledge, Bayesian model mixtures have been seldom explored for ocean color inversion (Frouin & Pelletier, 2014). The key feature of our model is the inversion of water and atmospheric signals from TOA observations using a multi-hypothesis setting. Rather than considering a single model, linear or not, we develop a Bayesian framework where the priors stated as mixture of models. Mixture models are trained both for water and aerosol contributions and lead to the identification of the reference spectrum families characterized by their mean spectrum and the associated covariance matrices. This training phase exploits in-situ data or radiative transfer simulations in the atmosphere and the water (Barker et al., 2008; Berk, 1999, Deuzé et al. 1989). Contrary to the machine learning approaches (NN, or Support Vector Regressions, SVR; Burges, 1998), the identified a priori distributions of the water and aerosol variables are directly linked to interpretable reference water or atmospheric spectra.

Our inversion scheme, referred to hereafter as MEETC2, is applied here to the estimation of water reflectances in complex waters from the MEdium Resolution Imaging Spectrometer (MERIS) TOA observations. Nevertheless the methodology is generic and may be directly applied to other sensors such as the incoming OLCI sensor embedded onto the sentinel 3 platform. Model calibration and validation involve the MEris MAtchup In-situ Database (MERMAID) radiometric in-situ dataset (Barker et al. 2008). Quantitative comparisons with the standard MEGS v8 (Antoine et al. 2006) and the MERIS C2R Neural Network outputs clearly demonstrate the relevance of our approach.

2 Review of the standard Ocean Color inversion method

2.1 Atmospheric correction principles

Ocean-color sensor measures at TOA the upwelling radiance (L_u) in $mW.m^{-2}.sr^{-1}$ backscattered by the ocean-atmosphere system. This radiance originates from photons scattered by air molecules and/or aerosols, which may also have been reflected directly at the sea surface (glint effect, Cox & Munk, 1954a&b), and may potentially have penetrated into the ocean. The measured TOA reflectance (ρ_{TOA}) is the ratio between the upwelling radiance L_u and the downwelling irradiance E_d, i.e. L_u integrated over the solid angle [0;2π]. The water reflectance contribution measured at TOA, i.e. transmitted through the atmosphere, is generally lower than 20% of the signal (Robinson et al. 2004). Due to this low signal/noise ratio, the unmixing of the atmospheric contribution from the water one reveals particularly complex. This inversion proves even more complex in coastal areas where the spectral water contribution may be closer to the aerosol contribution.

The traditional signal decomposition expresses measured TOA reflectance for each wavelength λ as a sum of elementary contributions:

$$\rho_{gc}(\lambda) = \rho_{Ray}(\lambda) + \rho_{aer}(\lambda) + t_d(\lambda).\rho_w(\lambda) + \rho_{coupl}(\lambda) + \varepsilon(\lambda) \quad (1)$$

where ρ_{GC} is the observed TOA reflectance ρ_{TOA} corrected from the glint and gaseous absorption, ρ_{Ray} (known) the reflectance of a purely molecular atmosphere (no aerosol),

ρ_{aer} (unknown) the reflectance of the aerosols, ρ_{coupl} (unknown) the coupling between air and aerosol molecules, t_d (unknown) the diffuse transmittance of the atmosphere, ρ_w (unknown) the water reflectance which is here the main quantity of interest to be estimated. ε is considered a white noise process. Here we consider a classical multivariate normal distribution (MVN) noise with null mean and spectral covariance matrix Σ_{ϵ}.

We consider here the Rayleigh-corrected reflectance variable $\rho_{RC}(\lambda)$ (Antoine et al., 2005, Santer et al. 1999, Gordon and Wang, Gordon 1997):

$$\rho_{RC}(\lambda) = \rho_{gc}(\lambda) - \rho_{Ray}(\lambda) = \rho_{aer}(\lambda) + t_d(\lambda).\rho_w(\lambda) + \rho_{coupl}(\lambda) + \varepsilon(\lambda) \quad (2)$$

The diffuse transmittance t_d is the product of both air molecules and aerosol particles scattering:

$$t_d(\lambda) = e^{-[0.5.\tau_{ray}(\lambda)+(1-w_a(\lambda).F_a(\lambda)).\tau_a(\lambda))].M} \quad (3)$$

where $\tau_{ray}(\lambda)$ is the Rayleigh optical thickness, $\tau_a(\lambda)$ is the aerosol optical thickness, M the air mass factor, w_a the aerosol single scattering albedo, F_a the forward probability scattering. τ_a is linked with the estimated aerosol reflectance for primary scattering (MERIS DPM, 2005):

$$\rho_{aer}(\lambda) = \frac{P(\lambda).w_a(\lambda)}{4(\cos(\Theta s) + \cos(\Theta v))}(1 - e^{-\tau a(\lambda).M}) \quad (4)$$

where $P(\lambda).Wa(\lambda)$ is the aerosol phase function times the single scattering albedo for the current scattering angle, Θs and Θv are respectively the sun and the view zenith angles. For a fixed geometry, aerosols contributions in the NIR are often assumed to follow an exponential decay (Gordon & Wang, 1994):

$$\rho_{aer}(\lambda) = \rho_{aer}(\lambda_0)e^{c(\lambda-\lambda_0)} \quad (5)$$

where $\lambda_0 = 865$ nm and c is the exponential decay of the aerosol spectrum, i.e. representative of the aerosol type. Though relevant in the NIR domain, the assumption of an exponential decay appears too restrictive in the 400-700 nm range where multiple scattering between aerosol and air molecules may become significant. Following Steinmetz et al. (2011), a polynomial model is considered to provide a more general model of aerosol contributions. Using our training dataset (cf § 4) a polynomial of order 3 was found as relevant to estimate the aerosol contributions:

$$\rho_{aer}(\lambda) = \rho_{aer}(\lambda_0) + a_1(\lambda - \lambda_0) + a_2(\lambda - \lambda_0)^2 + a_3(\lambda - \lambda_0)^3 \quad (6)$$

2.1.1 The MERIS standard processing atmospheric correction scheme

In the standard Level 2 processing of MERIS, the following four-step scheme is applied to estimate the water-leaving reflectances (Antoine & Morel, 2005):

1/ The signal is corrected from absorbing gaseous such as ozone, oxygen, water vapor and nitrogen dioxide.

2/ The estimated contribution of suspended matter particles in the NIR is removed from TOA observations after single scattering transmittance through the atmosphere. This step is known as the Bright Pixel Atmospheric Correction (BPAC) and detailed in the next section.

3/ A mixture of two aerosol models among 34 (for MERIS) is estimated from the values of ratio $\rho_{path} = \rho_{gc} / \rho_{ray}$ (Eq.2) at 779 and 865 nm, leading to the estimation of both the aerosol reflectance $\rho_{aer}(\lambda)$ and the multiple scattering transmittance $t_d(\lambda)$ (Eq.3).

4/ Water reflectance contribution is estimated by subtracting the estimated aerosol contribution from $\rho_{aer}(\lambda)$ using Eq. 2.

The Bright Pixel Atmospheric Correction (BPAC)

Whereas, in open ocean waters, one can exploit null contribution of water reflectance in the near infrared (NIR) range to infer aerosol contributions, no such simple inversion scheme applies in coastal waters, which are characterized by a non-null contribution in this domain (Ruddick et al., 2005&2006) and possible correlations between atmospheric and water spectra. This is a major issue to be dealt with in the atmospheric corrections in coastal waters. BPAC is an iterative algorithm to correct the TOA signal from the estimated contribution of turbidity. It aims at removing the water contribution, caused by suspended matters, of the TOA observed reflectance (Moore et al., 1999). This step is essential in the standard MERIS level 2 processing as the estimation of the aerosols is performed using the NIR bands under the assumption $\rho_w(NIR) = 0$. Moore proposed for MERIS a two steps algorithm which iterates: the estimation of $\rho_{aer}(709, 865)$, c and $\rho_{aer}(779)$ using $\rho_{path}(779, 865)$, then, using the estimated residuals $\hat{\rho}_w$ in the NIR from Eq.2 and a parametric model, the estimation of the SPM concentration and related $\hat{\rho}_w$ at TOA. This converging algorithm suffers from important drawbacks for very turbid waters. In such areas, the considered water model does not allow retrieving high concentrations of SPM (Goyens et al. 2013). It typically leads to an over-correction of the blue water-reflectance, i.e. an underestimation of ρ_w at 412 and 442 nm with the standard Level 2 processing, and may resort to geophysically-meaningless negative reflectance values.

3 Method

3.1 Spectral reference signatures of the sea water using Non-Negative Matrix Factorization

Given the spectral overlap of water and aerosol contributions especially in coastal areas, inversion of (Eq.1) requires some prior knowledge on water contributions. We propose here to determine from the training dataset a parametric spectral representation of water contributions. We use here a Non-Negative Matrix Factorization (NNMF) with projected gradients (Lin, 2007). Similarly to PCA, it relies on an additive decomposition on a basis learnt from the data. In contrast to PCA, it does not involve orthogonality constraints but imposes non-negativity for both the basis function and the projection coefficients. NNMF is among the most popular approaches in multispectral and hyperspectral remote sensing as a mean to unmix contributions issued from various sources in a sensed environment (Jia & Qian, 2009). Formally, NNMF leads to the following parametric representation of a given water spectrum $\rho_w\,(\lambda)$:

$$\rho_w\,(\lambda) = W(\lambda, n) * h(n) \tag{7}$$

where W(λ, n)>0 spectral reference signature basis identified by NNMF using the training data, and h(n)>0 refer to the coordinates of the spectrum $\rho_w\,(\lambda)$ in the decomposition space. It may be noticed that NNMF decomposition could also be replaced here by a bio-optical model. Nevertheless, to our knowledge none of this model is today performant enough to estimate, in coastal areas, the water leaving reflectance spectrum from the water's constituents. The NNMF decomposition, by imposing non-negativity of both the coordinates and the reference spectral signatures appropriately constrains our inversion to converge toward physically-realistic solutions (cf § 4.3.1) conversely to the standard Level-2 processing (ESA and NASA). In our case, four spectral reference signatures were needed to address the training in-situ spectrum variability (cf 4.2).

3.2 Bayesian setting

From Eq.1, the variables to be estimated are $X_w = \{h_i\}$, i.e. the coordinates of ρ_w in the basis W (Eq. 7), and $X_a = \{a_i\}$ i.e. the polynomial coefficients of the aerosol models (Eq.6). Conversely to a standard least square estimation framework, we develop a Bayesian setting. It relies not only on the likelihood of the residuals $\delta\rho_{RC\,=}\,\hat{\rho}_{RC} - \rho_{RC}$ but also on the prior distributions of X_a and X_w.

Formally, we consider the Maximum A Posteriori estimation (MAP, Harold & Sorenson, 1980) which aims at maximizing the posterior probability $P(\mathrm{X_a}, \mathrm{X_w} | \rho_{RC}, \varphi)$:

$$P(\mathrm{X_a}, \mathrm{X_w} | \rho_{RC}, \varphi) \;\alpha\; P(\rho_{RC} | \mathrm{X_a}, \mathrm{X_w}, \varphi) \,.\, P(\mathrm{X_a}, \mathrm{X_w} | \, \varphi)$$

We suppose here that $\mathrm{x_a}$ *and* $\mathrm{x_w}$ are independent i.e.: (8)

$$P(\mathrm{X_a}, \mathrm{X_w} | \rho_{RC}, \varphi) \;\alpha\; P(\rho_{RC} | \mathrm{X_a}, \mathrm{X_w}, \varphi) \,.\, P(\mathrm{X_a} | \, \varphi) . P(\, \mathrm{X_w} | \, \varphi)$$

The first term $P(\rho_{RC} | \mathrm{x_a}, \mathrm{x_w}, \varphi)$ is the likelihood of the observation model (Eq. 1) with respect to variables $\mathrm{X_a}$, $\mathrm{X_w}$ and φ. φ is here a vector of covariates composed of the observation geometry and pre-estimates of the water and aerosol contributions in the NIR performed in the bright pixel estimation step (BPE, § 4.3). $P(\, \mathrm{X_a} | \, \varphi)$ and $P(\, \mathrm{X_w} | \, \varphi)$ refer to the priors on $\mathrm{X_a}$ and $\mathrm{X_w}$ variables given the covariates.

In the proposed Bayesian framework, $\mathrm{P}(\rho_{RC} \,|\, \mathrm{X_a}, \mathrm{X_w}, \varphi\,)$ is modeled with a multivariate normal distribution with a null mean vector and full covariance matrix Σ_{ϵ}. As detailed in the next sections, $\mathrm{P}(\mathrm{X_a} | \, \varphi)$ and $\mathrm{P}(\mathrm{X_w} | \, \varphi)$ a priori distributions are modeled using a mixture of MVN distributions, namely a Gaussian Mixture Models (GMM, Reynolds, 1995). The MAP criterion cost function is finally expressed using the log likelihood:

$$C = -\log(P(\mathrm{X_a}, \mathrm{X_w} | \rho_{RC}, \varphi)) \tag{9}$$

Covariates and non-homogeneous prior distributions

Covariates are here geophysical parameters significantly correlated with the variables of interest. From a physical point of view, the observed shape of aerosol reflectance spectrum $\rho_{aer}(\lambda)$, i.e. $\mathrm{a_i}$ coefficients of Eq.6, is correlated with the variables which describe geometry of acquisition conditions (Θs, the sun zenith angle, Θv, the view zenith angle, and δψ, the delta azimuth), and the variables which describe the aerosol type and quantity, $\rho_{aer}(865)$ and c (Eq. 5), estimated using the NIR part of the spectrum (cf § 4.3). To characterize the correlation between variables and possible covariates, we use a linear discriminant analysis (McLachlan et al., 2004) and the training dataset.

The selection of the significant contributors led to consider $\varphi_a = \{\rho_{aer}(865), \mathrm{c}, \theta\mathrm{v}, \theta\mathrm{s}\}$ for the aerosol variable X_a and $\varphi_w = \{\rho_w(780), \mathrm{c}, \theta\mathrm{v}, \theta\mathrm{s}, \delta\psi\}$ for the water variable X_w. To derive the priors of X_a and X_w, we model, using the training dataset and the Expectation Maximization algorithm (Dempster, 1977), the joint distributions $P(X_w, \varphi_w)$ and $P(X_a, \varphi_a)$ as GMMs:

$$P(X_w, \varphi_w) = \sum_{\mathrm{Zw}=i} \mathit{\Lambda}_i \, g_{\Sigma_{\{X_w,\varphi_w\}_i}}\left(\{X_w, \varphi_w\} - \mu_{\{X_w,\varphi_w\}_i}\right)$$

$$P(X_a, \varphi_a) = \sum_{\mathrm{Za}=j} \mathit{\Lambda}_j \, g_{\Sigma_{\{\mathrm{X_a},\varphi_a\}_j}}\left(\{X_a, \varphi_a\} - \mu_{\{\mathrm{X_a},\varphi_a\}_j}\right) \tag{10}$$

We use subscript i (resp. j) for the water-specific (resp. aerosol-specific) GMM. $\mathit{\Lambda}_i$ is the prior probability of mode i in the GMMs. It refers to the probability of the hidden state variable Z_w to be in mode i, $P(Z_w=i)$ (resp. $P(Z_{a=j})$). $g_{\Sigma_{\{X_w,\varphi_w\}_i}}$ (resp. $g_{\Sigma_{\{X_a,\varphi_a\}_j}}$) is a zero-mean MVN distributions with covariance matrice $\Sigma_{\{X_w,\varphi_w\}_i}$ (resp. $\Sigma_{\{X_a,\varphi_a\}_j}$) and mean vector $\mu_{\{X_w,\varphi_w\}_i}$ (resp. $\mu_{\{X_a,\varphi_a\}_j}$) for the joint variables $\{X_w, \varphi_w\}$ (resp. $\{X_a, \varphi_a\}$) for mode i (resp. j).

We update, given the covariate's values, the a priori distributions of Eq. (10) to obtain the a priori conditional distributions of Eq.8:

$$P(X_w|\varphi_w) = \sum_{\mathrm{Zw}=i} \mathit{\Lambda}_{i|\varphi_w} \, g_{\Sigma_{X_w|\varphi_{w,i}}}\left(X_w - \mu_{X_w|\varphi_{w,i}}\right)$$

$$P(X_a|\varphi_a) = \sum_{\mathrm{Za}=j} \mathit{\Lambda}_{j|\varphi_a} \, g_{\Sigma_{\mathrm{X_a}|\varphi_{a,j}}}\left(\mathrm{X_a} - \mu_{\mathrm{X_a}|\varphi_{a,j}}\right) \tag{11}$$

These non-homogeneous priors on X_a and X_w involve conditional means, covariances and priors given the covariates and the initial GMM model estimated onto the joint variables. For instance, for mode i of the aerosol prior, the conditional parameters are given by (Petersen, 2008):

$$\mu_{X_a|\varphi_a,j} = \mathrm{E}(X_a|\varphi_a, Z_a = j) = \mu_{X_a,\mathrm{j}} + \Sigma_{X_a,\varphi_a,j}.\Sigma^{-1}_{\varphi_a,j}\ .(\varphi_a - \mu_{\varphi_a,\mathrm{j}})$$

$$\Sigma_{X_a|\varphi_a,j} = \Sigma_{X_a,j} - \Sigma_{X_a,\varphi_a,j}.\Sigma^{-1}_{\varphi_a,j}.\Sigma_{\varphi_a,X_a,j} \tag{12}$$

$$\Lambda_{X_a|\varphi_a,j} = \Lambda_j * P(X_a|\varphi_a, Z_a = j) \,/ \sum_l \Lambda_j * P(X_a|\varphi_a, Z_a = l)$$

3.3 Functional scheme

Figure 1 summarizes the functional scheme for the proposed Bayesian inversion given the calibrated model parameters, i.e. means and covariance matrices for the GMM models and the MVN distribution of the residuals, $\hat{\rho}_{RC}(\lambda) - \rho_{RC}(\lambda)$, of Eq. (1). In the first step, the Bright Pixel Estimation (BPE) is based on the water similarity spectrum (Ruddick et al., 2005) to estimate $\rho_w(780)$, $\rho_{aer}(865)$ and c the slope of the aerosol. An iterative convergent algorithm is used. Given the estimated covariates, we update in step 2 the GMM for X_a and X_w conditionally to the covariates (Eq.(12). Step 1 & 3 involve gradient descent based inversions and a taylor series of Eq. (4). As the MAP criterion may not be a concave criterion, the initialization of the gradient descent, step 3, is a key issue as gradient-based maximization may converge toward local minima. We proceed as follows: 25 aerosol parameters are randomly generated using the updated distributions. X_w initialization is performed using the estimated $\hat{\rho}_w(780)$, X_a initialization and Eq. 2. Overall, among the 25 computations, we select in step 4 the solution corresponding to the highest value of the MAP criterion (Eq.9).

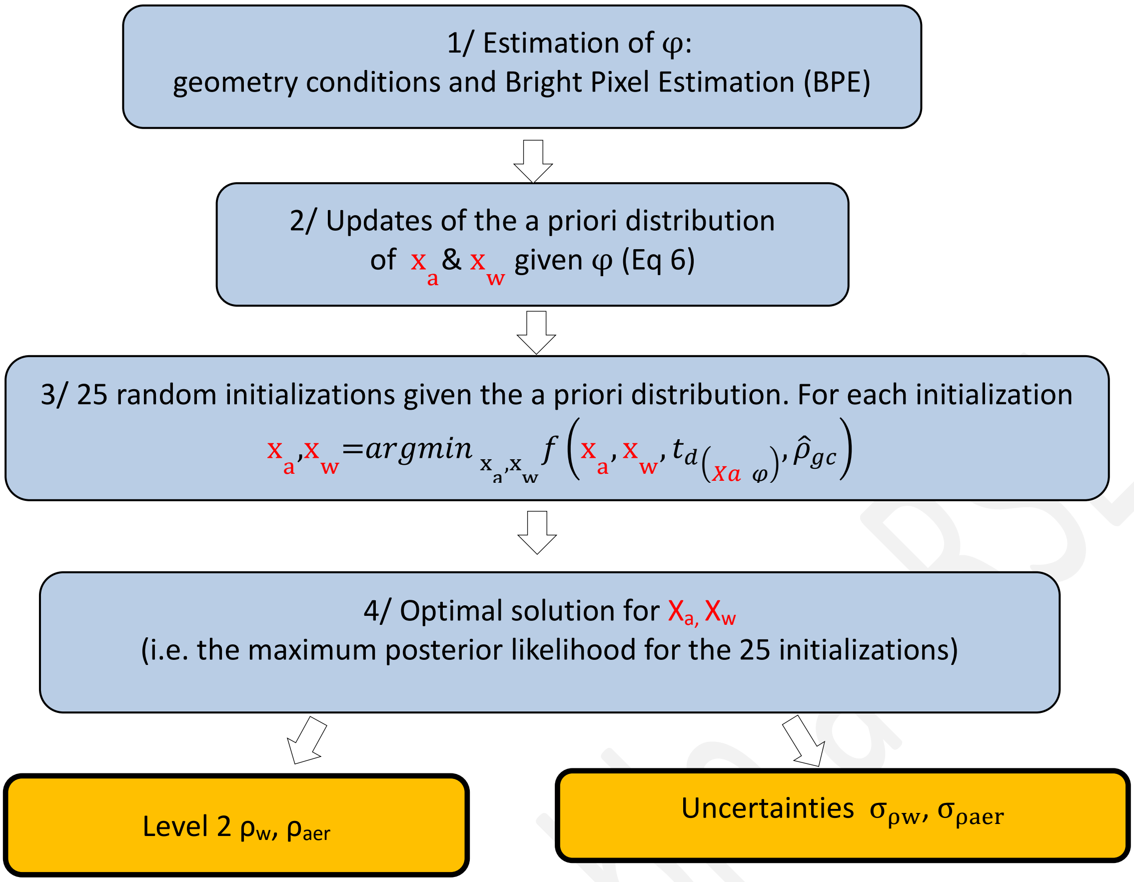


Figure 1: operational scheme for the atmospheric correction MEETC2 bayesian inversion

4 NUMERICAL EXPERIMENTS

To validate the proposed methodology, the 5976 radiometric in-situ profiles have been randomly splitted into two sets of equal size: a training dataset and a validation dataset. Model parameters are estimated using the training dataset. The optimal number of clusters, k, used in the GMM to estimate X_a and X_w a priori Probability Density Functions (PDF), is determined using the Bayes Information Criterion (BIC) (Bhat & Kumar, 2010) and the explained variance criterion (Saulquin et al., 2014). Validation is performed with the validation dataset, using scatter plots between estimated and in-situ $\rho_w(\lambda)$, histograms, and related regression statistics. We evaluate type-II regression statistics, i.e. a regression model that accounts for uncertainties for both y and x as the in-situ measurements also involves uncertainties (Laws, 1997).

4.1 The in-situ MERMAID dataset

The MERMAID (http://hermes.acri.fr/mermaid/home/home.php) in-situ matchup database is a comprehensive dataset that gathers in-situ measurements of water leaving radiances, IOPs, and MERIS TOA reflectances measured at the same location. Many sites are available and among them, the most known are the NASA bio-Optical Marine Algorithm Dataset (NOMAD, Werdell & Bailey, 2005), the “BOUée pour l'acquiSition d'une Série Optique à Long termE” (BOUSSOLE, Antoine et al., 2006) mooring program, the Aerosol Robotic Network (AERONET; Zibordi et al. 2009) stations, the Helgoland transect (Petersen et al. 2008) that provides a full dataset of radiometric in-situ measurements in the Baltic Sea complex waters, and the MUMM Trios dataset (Ruddick et al.,2006). Our initial dataset gathers 5976 matchups (without glint) measured at the MERIS wavelengths: 412.5, 442.5, 490, 510, 560, 630, 665, 681, 708, 753.75, 778.75, 865 and 885 nm. For each in-situ measurement of Figure 2 (left), we use the corresponding 3 by 3 MERIS pixels (Bailey & Werdell, 2006).

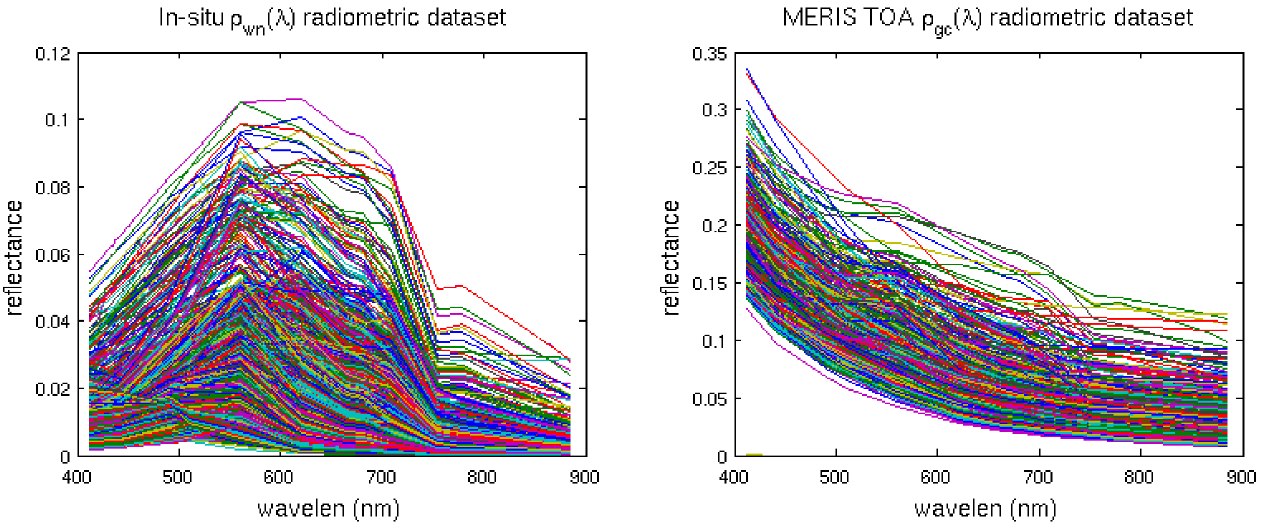


Figure 2: Left, the 5976 in-situ water reflectance spectra in complex waters. Right, the corresponding (matchups) ρ_{GC} (TOA) observed from the MERIS sensor.

4.2 Calibrated priors

A 25-mode mixture model (cf § 3.3) was selected to model the joint distribution of $\{X_{a,} \varphi_{a}\}$. Figure 3 shows the 25 aerosol modes reconstructed from the GMM centers for $\rho_{aer}(865) = 0.01$.We remind that the PDF of $\{X_{a,} \varphi_{a}\}$ involves a full covariance matrix $\Sigma 0_{Xai}$ for each mode that is accounted for in the maximization of Eq. 8.

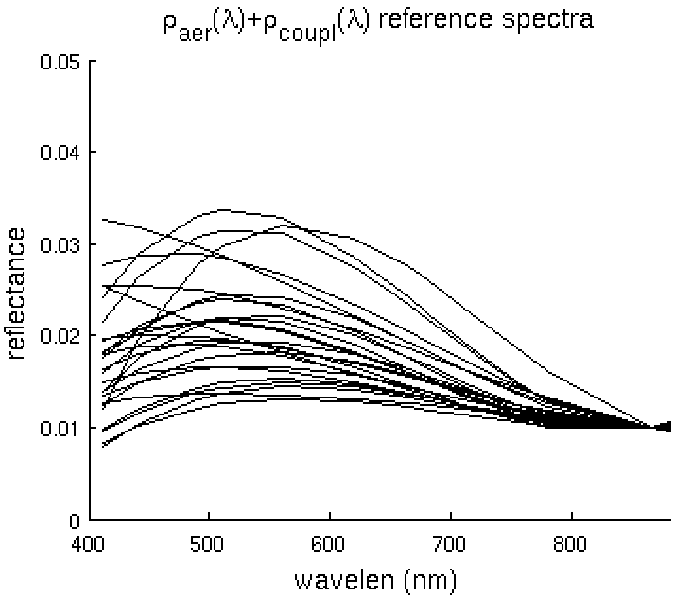


Figure 3: The 35 aerosol modes reconstructed from the GMM and Eq 6.

From the NNMF applied to the in-situ water spectra, a four reference spectral signatures, W(λ) was needed to reconstruct 99% of the variance of the in-situ spectra training dataset (Eq.7, Figure 4a). The NNMF reference spectral signatures characterize the influence of the optically active constituents of the water column onto the observed water leaving reflectance spectra:

- Reference n°1 (dark blue) is a typical spectral signature observed in presence of SPM (Doxaran et al., 2002; Bricaud et al. 1998).
- Reference n°2 (green) highlights the spectral signature of CDOM absorption with its typical decrease toward the blue.

- Reference n°3 (red) is the typical spectral signature of the chl-a, i.e. absorption in the blue and the resulting observed in the green (560 nm).
- Reference n°4 (light blue) is the spectral signature of the pure water (Pope & Fry, 1997).

A 35-mode GMM is optimal to fit the prior distribution of $\{X_w, \varphi_w\}$. Figure 4 shows the 35 water reflectance reference spectra. Similarly to the prior distribution of aerosol contributions, a full covariance matrix $\Sigma 0_{Xwi}$ is estimated for each mode i.

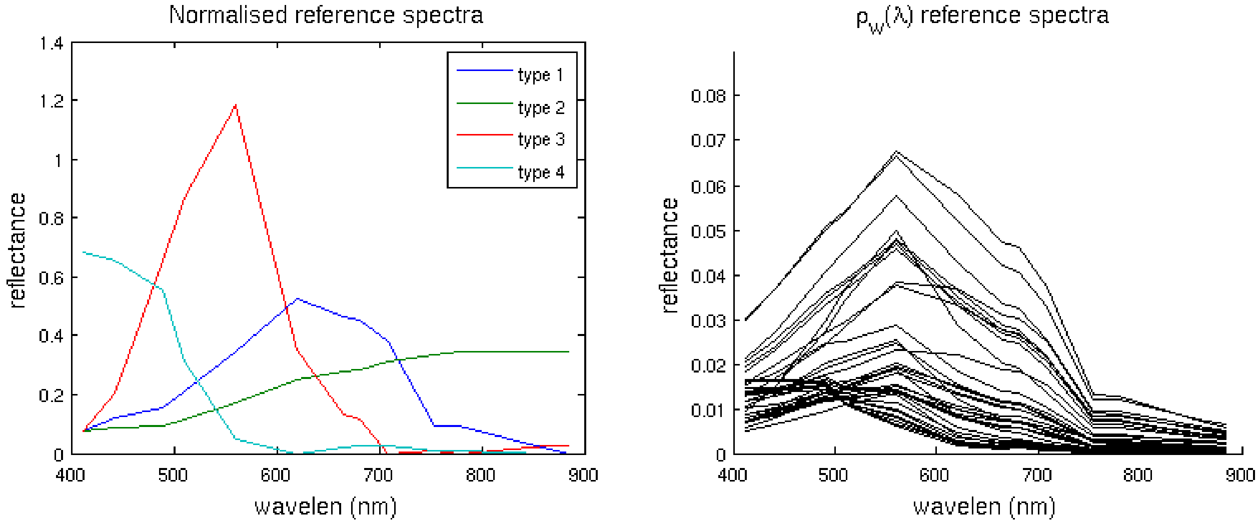


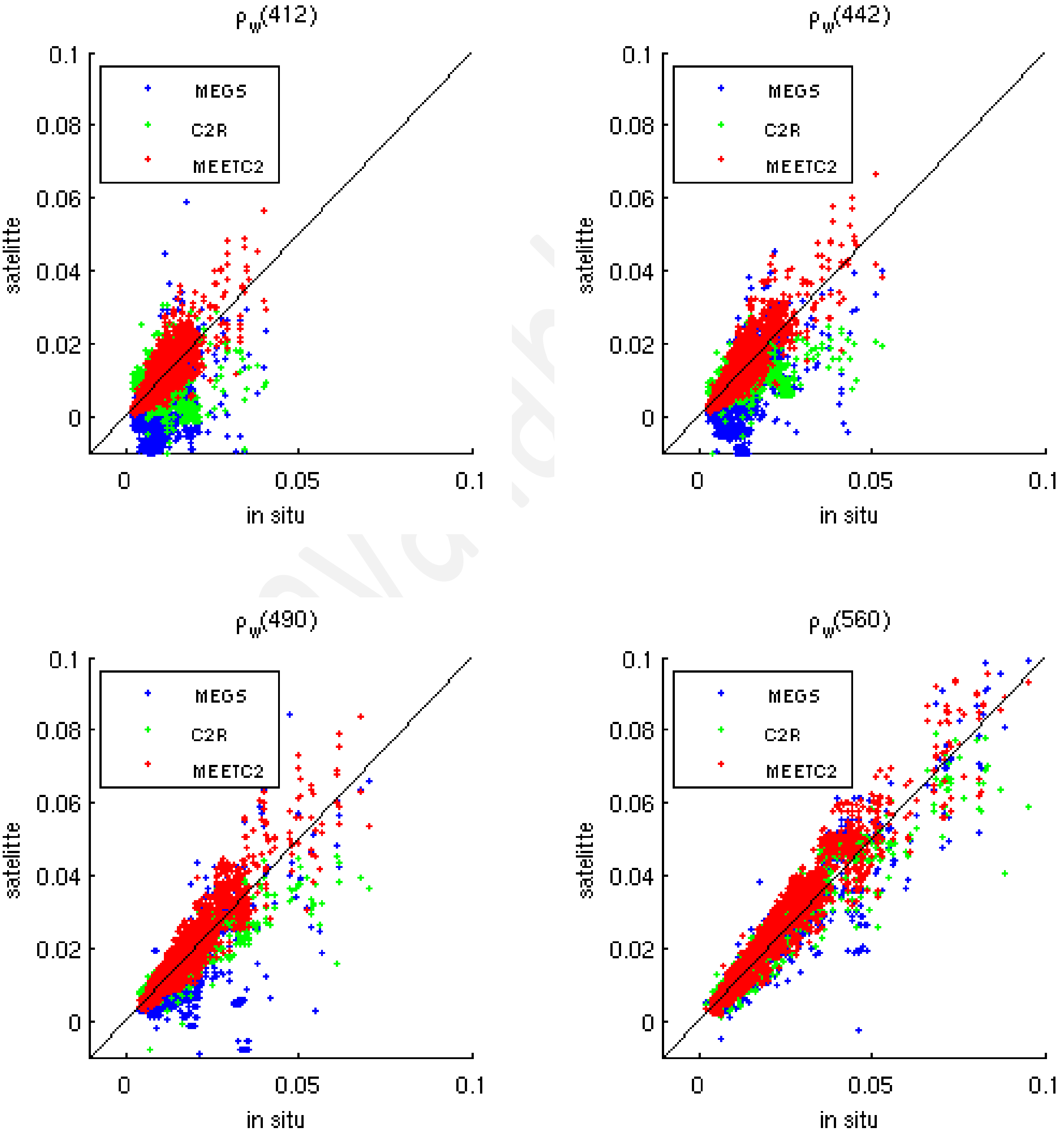
Figure 4: Left, the reference spectral signature basis, W(λ), estimated using NNMF with projected gradients. Right, the 35 reference water models reconstructed using the GMM centers and Eq 7.

4.3 *Ocean color inversion results*

4.3.1 *Inversion performance for the Mermaid dataset*

We perform a quantitative evaluation of the performance of the proposed Bayesian inversion model, MEETC2, for the Mermaid dataset and coastal waters. For the validation dataset, i.e. the half of the 5976 spectra, we analyze for each wavelength the estimated water reflectances $\hat{\rho}_w$ against in-situ measurements (Figure 5, red). In addition to the proposed Bayesian inversion, we also report on Figure 5 the inversion performed with MEGS v8 (blue), and C2R (green). Table 1 summarizes the corresponding statistical results for the 13 wavelengths.

On this validation dataset, MEETC2 clearly outperforms MEGS and C2R at bands 412, 442, 490 and 510 nm in term of mean-bias, mean absolute error, slope, R² coefficient and σ. From 620 to 885 nm MEETC2 slightly outperforms the two other models. Overall, the gain on the relative absolute error over the 12 bands is of 57% compared with MEGS and 10% compared with C2R.



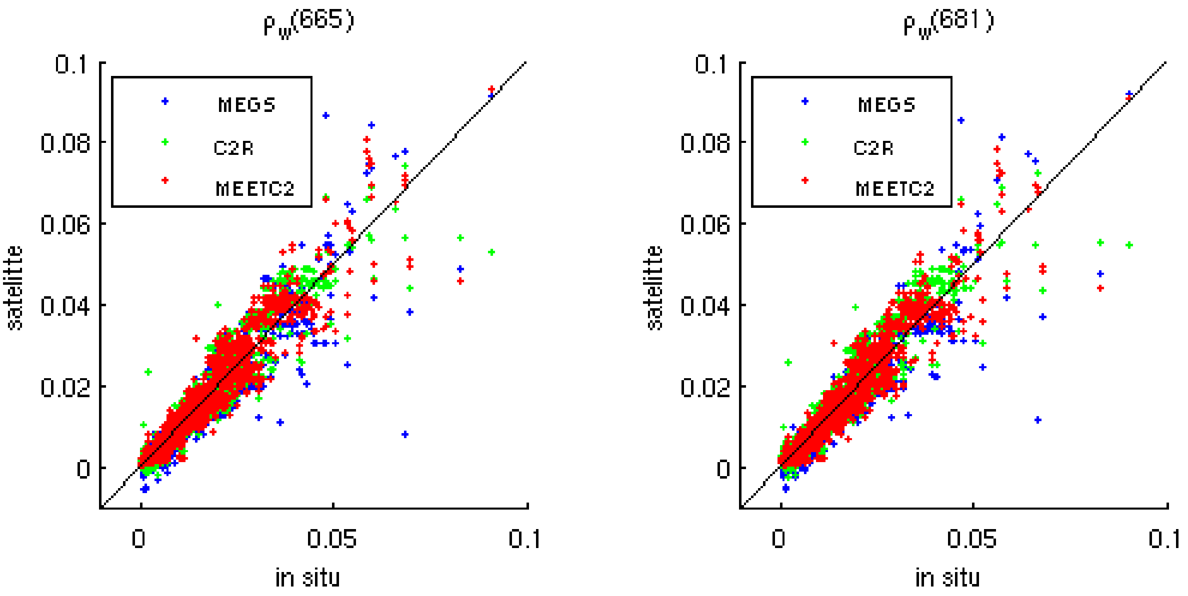


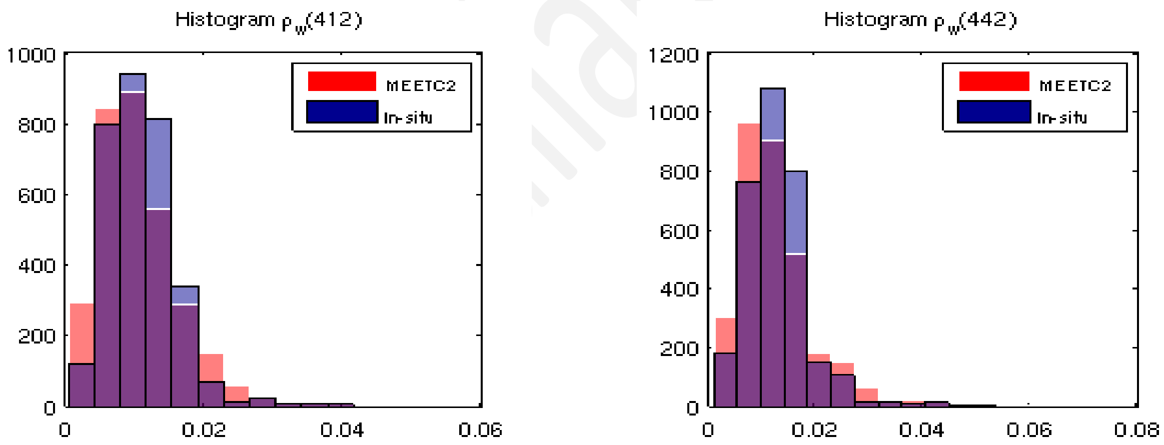
Figure 5: comparisons between the estimated $\hat{\rho}_w$ at 412, 442, 490, 560, 665 and 681 nm using MEETC2 vs in-situ (red), MEGS 8 vs in-situ (blue) and C2R (NN) vs in-situ (green).

Table 1: Statistical analyses of the estimated water reflectances vs. in-situ data for the proposed Bayesian model (MEETC2), the standard MEGS processor and the neural-net-based algorithm C2R. For each wavelength, we report the mean error (bias), the slope of the regression with the in situ data, the associated R^2 score and standard deviation, σ. We report in bold the algorithms which provided the best performance.

λ(nm)		Mean error	Slope	R^2 (Pearson)	σ
412.5	MEETC2	**-0.0004**	**1.14**	**0.70**	**0.0039**
	MEGS	-0.0083	0.44	0.16	0.0149
	C2R	-0.0023	0.42	0.15	0.0075
442.5	MEETC2	**-0.0002**	**1.13**	**0.75**	**0.0041**
	MEGS	-0.0060	0.39	0.38	0.0128
	C2R	-0.0031	0.31	0.60	0.0060
490	MEETC2	**0.0004**	**0.97**	**0.77**	**0.0049**
	MEGS	-0.0033	0.92	0.76	0.0098
	C2R	-0.0022	0.77	0.57	0.0046
510	MEETC2	**0.0007**	**0.96**	**0.85**	**0.0046**
	MEGS	-0.0020	0.91	0.65	0.0085
	C2R	-0.0013	0.72	0.78	0.0040
560	MEETC2	**0.0007**	**1.04**	0.88	**0.0049**
	MEGS	-0.0007	0.95	0.81	0.0056
	C2R	-0.0007	0.88	**0.90**	0.0050
620	MEETC2	**0.0006**	**1.00**	**0.93**	**0.0038**
	MEGS	-0.0014	1.05	0.85	0.0050
	C2R	-0.0012	0.97	0.90	0.0039
665	MEETC2	**1.1809e-03**	**0.97**	**0.88**	**0.0033**

	MEGS	-0.8881-03	1.07	0.85	0.0043
	C2R	0.2650-03	1.02	**0.88**	**0.0033**
681	MEETC2	**0.0657e-03**	**0.99**	**0.92**	**0.0033**
	MEGS	-0.6257e-03	1.06	0.85	0.0041
	C2R	0.5037e-03	1.02	0.89	**0.0033**
708	MEETC2	**-0.0136e-03**	**0.94**	**0.87**	**0.0030**
	MEGS	-0.6400e-03	1.13	0.83	0.0039
	C2R	0.7668e-03	1.10	**0.87**	0.0037
753	MEETC2	**0.2656 e-03**	**0.90**	**0.90**	**0.0014**
	MEGS	-0.4222 e-03	1.35	0.78	0.0027
778	MEETC2	**0.2418 e-04**	**0.90**	**0.89**	**0.0014**
	MEGS	-0.4084e-04	1.21	0.74	0.0029
865	MEETC2	**0.3793 e-04**	0.94	**0.88**	**0.0009**
	MEGS	-0.2259e-04	**1.02**	0.75	0.0017
885	MEETC2	**0.2794 e-04**	0.95	**0.88**	**0.0007**
	MEGS	-0.2269e-04	**1.02**	**0.88**	0.0014

We further analyze the extent to which we recover realistic water reflectances from the proposed Bayesian inversion MEETC2. To this end, we compared for each wavelength, the distribution of in-situ measurements to the MEETC2 estimates. Figure 6 shows a global agreement between the distributions of $\hat{\rho}_w(\lambda)$, compared to the reference in-situ distributions, for wavelengths 412, 442, 510 and 560 nm.



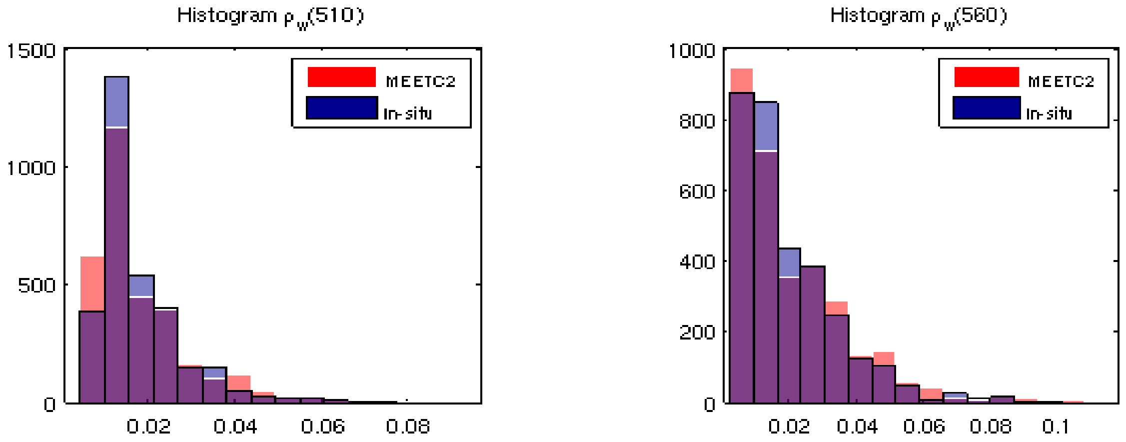
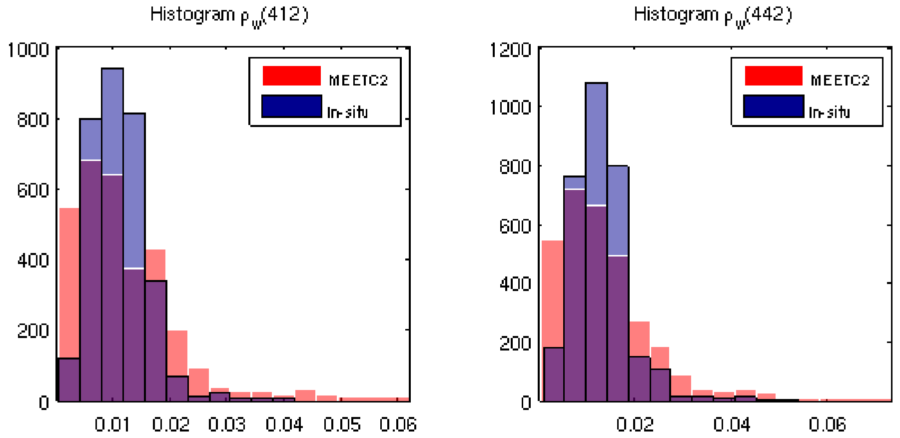


Figure 6: Comparison of the distributions of the estimated water reflectances $\hat{\rho}_w$ at 412, 442, 510, 560 nm for in-situ measurements (blue) and the proposed inversion (MEETC2 model, red).

To illustrate the added value of the introduction of priors on both water and aerosol spectra, we implement model (Eq. 8) without priors on X_a and X_w, i.e. the cost function of Eq. 9. In this case, the cost function in the inversion is directly comparable with the one of a Generalized Least Square Model (GLS) with error covariance matrix Σ_ϵ. Figure 7 shows the corresponding results obtained, using the same validation dataset. We clearly see in Figure 7a smoothing effects for bands 412, 442, 510 and 560 nm on the estimated distributions of $\hat{\rho}_w$. The resulting bias with the in-situ is lower using the MAP estimator and a priori knowledge (Figure 6).



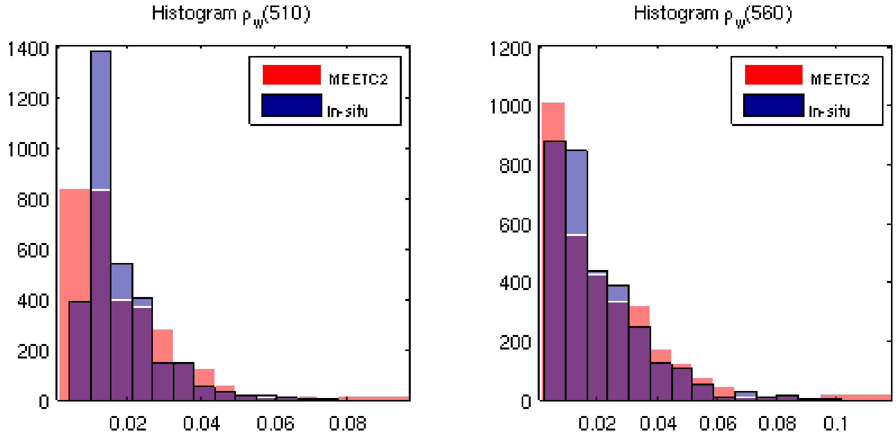


Figure 7: Comparison of the distributions of the estimated water reflectances $\hat{\rho}_w$ at 412, 442, 510, 560 using a cost function without a priori for the inversion (Eq. 9) vs in-situ. In that case, the MAP criterion reduces to the Maximum Likelihood criterion.

4.3.2 Example of estimated water reflectance on a very turbid area

Figure 8 shows the estimated $\hat{\rho}_w$, using the 20090322 MERIS Full Resolution (FR) level 1 observations over the French La Gironde's estuary, using the three algorithms. At springtime in this area, a bloom occurs leading to high chl-a concentrations (typically of magnitude from 5 to 15 mg.m^{-3}). At the same time, the seasonal river outflow involves high, SPM concentrations and CDOM absorption (Doxaran et al. 2009). In the same manner, Figure 9 shows the estimated $\hat{\rho}_w$, using the 20040208 MERIS Full Resolution (FR) level 1 observations over the French La Seine's estuary.

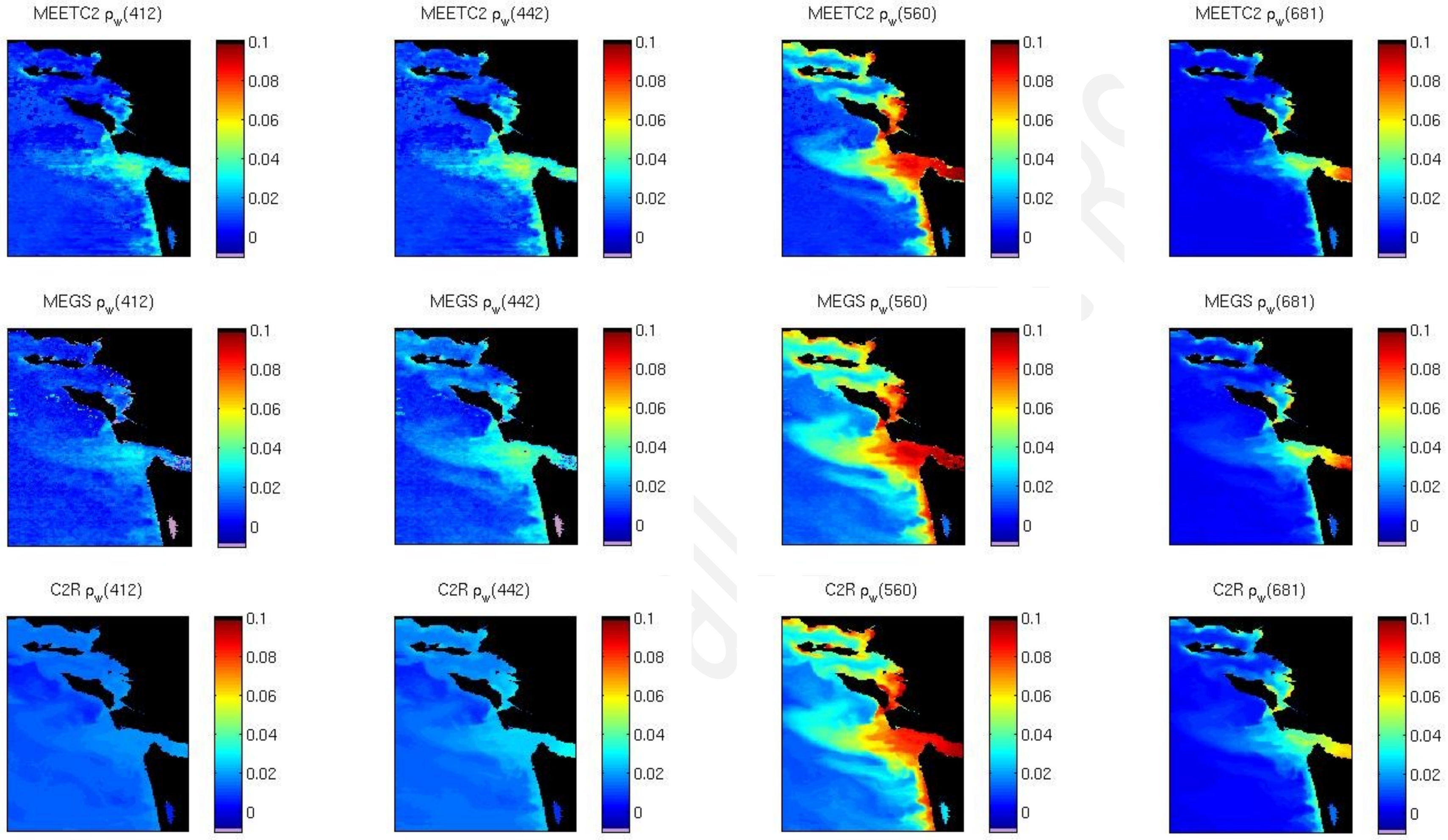


Figure 8: Estimated $\rho_w(412, 442, 560, 680\ )$ (left to right) from the MERIS FR Level 1 image of the 20090322 over the French river La Gironde's estuary. Top, MEETC2 retrievals, middle, MEGS v8 and bottom C2R retrievals.

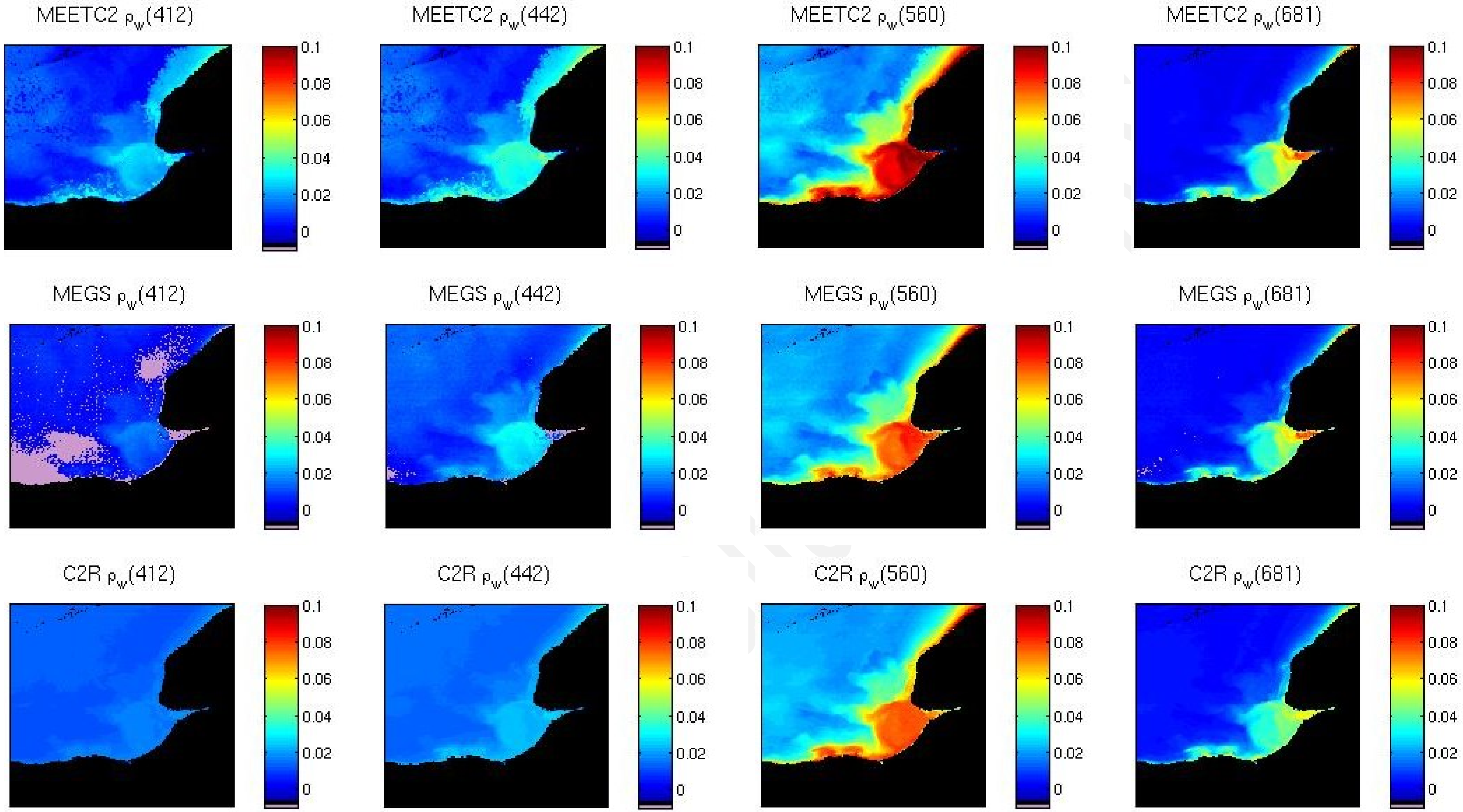
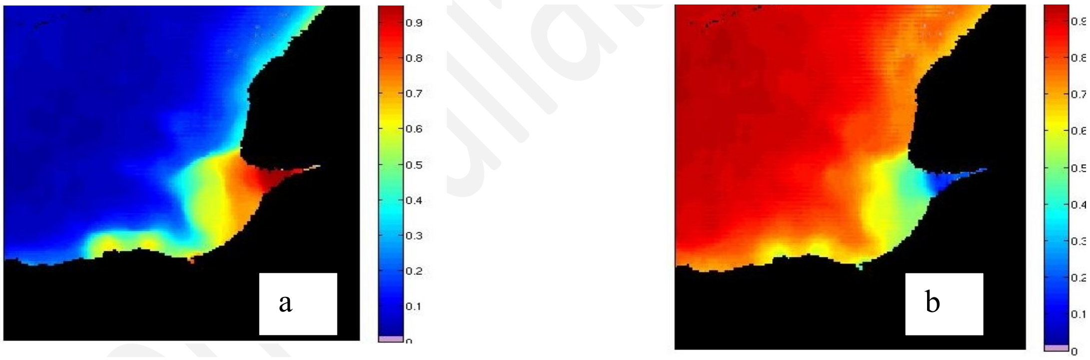


Figure 9: Estimated $\rho_w(412, 442, 560, 680\ )$ (left to right) from the MERIS FR Level 1 image of the 20040208 over the French river La Seine's estuary. Top, MEETC2 retrievals, middle, MEGS v8 and bottom C2R retrievals. In pink are highlighted negative reflectances.

4.3.3 Estimated water types associated with the MEETC2 inversion

The NNMF reference spectral signatures in Figure 4 (left) characterize the influence of the optically active constituents of the water onto the observed water leaving reflectance spectra. This reference basis may be used to validate indirectly, using oceanographic knowledge, the spatial coherency of the estimated $\hat{\rho}_w$. Figure 10 depicts the projection coefficients of the estimated MEETC2 $\hat{\rho}_w$ onto the reference spectral signatures. Figure 10c depicts the presence of chl-a over the all area as expected for this spring period and region. We observe the typical clear contrasted situation in an estuary between waters whose spectral shape is mainly constrained by SPM (Figure 10a) and clearer waters (Figure 10d) in the oceanic part of the estuary. This spatial consistency of the distribution of the water types from the estimated $\hat{\rho}_w$, relatively to our knowledge of the seasonal behavior in this area, contributes to validate the shapes of our estimated spectra.

Type 1: SPM dominated ρ_w Type 2: Cdom dominated ρ_w



Type 3: Chl-a dominated ρ_w Type 4: pure water dominated ρ_w

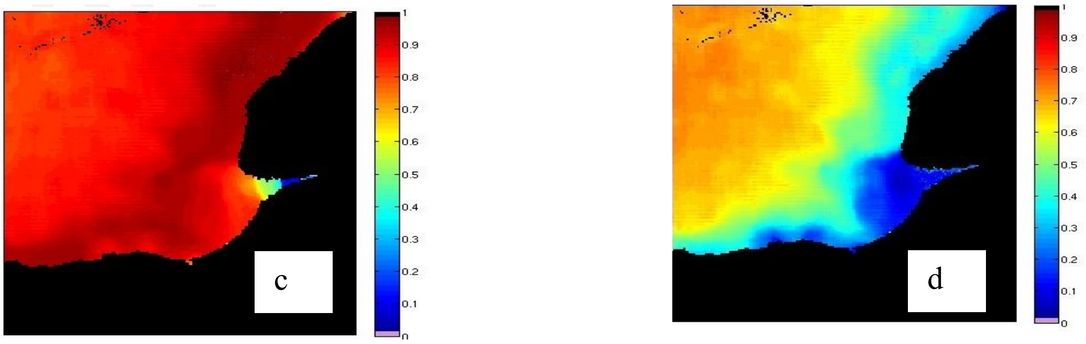


Figure 10: projection coordinates of the MEETC2 estimated $\hat{\rho}_w$ in the four NNMF reference spectral signature basis.

5 Discussion

A significant improvement of ocean color inversion in coastal waters.

Retrieving reliable Ocean Color reflectances from space in coastal areas remains a major challenge for a number of operational and scientific issues, including for instance the delivery of reliable satellite-derived products in coastal areas for the space agencies, bio-optical and biological modeling, as well as environmental monitoring policies such as the WFD. Using the MERMAID satellite/in-situ collocated observation database, a Bayesian latent class model was shown to significantly enhance the inversion of water reflectances for complex waters compared to the standard MEGS inversion scheme and the C2R, a Neural Network trained using similar in-situ data (Schiller & Doerffer, 1999).

The improvements were especially noticeable for the 412, 442, 490 and 510 nm bands, which are used in Ocean Color for the estimation of the chl-a concentration, CDOM absorption and light attenuation underlying the potential of such approach to improve the standard level 2 products in coastal areas. An additional important feature of the proposed inversion, based onto the Non Negative Matrix Factorization water model, is strictly positive estimates of the water leaving reflectances in coastal areas. Meaningless negative estimates, as observed in the standard MEGS products are not anymore possible.

The complexity of the inversion is particularly stressed by the number of needed hidden models, respectively 25 for coastal aerosol reflectances and 35 for water reflectances, to address the spectral variability of both water and atmospheric contributions in such areas, and to unmix the possibly correlated aerosol and water spectra.

A physically-interpretable modelling framework

Conversely to Neural Network, the modes retrieved by the Gaussian Mixture Models correspond to identifiable aerosols, such as identified in the MERIS and the OLCI reference aerosol database, and water types. The fact that we explicitly distinguish parametric representations of aerosol and water spectra makes also easier the independent calibration of the models and our Bayesian model may benefit in a much simpler manner for newly collected and/or simulated dataset to improve each prior distribution independently. This is regarded as a key property for future operational applications with respect to ongoing advances in radiative transfer modelling, in-situ monitoring and future satellite missions.

Operational potential in the framework of the ocean sensor of upcoming Sentinel 3 platform

The incoming OLCI Ocean Color sensor, embedded on the Sentinel 3 platform, should succeed the MERIS sensor in 2015. The available spectral bands will be close to the MERIS ones. Beyond genericity of our Bayesian framework, we thus expect the considered parameterization, especially the NNMF-based representation, the GMM-based priors and the covariance models, to be directly transferable to the future OLCI observations. Our ongoing work addresses the development of an operational product based on the proposed Bayesian mode that will be freely distributed in the Odesa software (http://www.odesa-info.eu/info/). The dependency of both aerosol and water prior distributions to the observation geometry conditions will be addressed soon using radiative transfer simulations such as the Successive Order Scattering radiative transfer code (Deuzé, 1989) and Hydrolight (Mobley, 1998) to cover the full possible range of observation conditions.

From a modeling perspective, additional developments appear of interest, especially new covariates, e.g. humidity and wind conditions to further constrain the prior distributions of the water and aerosol variables. Parallelized implementation is also under investigation, as, conversely to existing MEGS and C2R processors, our optimization is computationally more demanding than these as it relies on quasi-randomized initializations for the atmospheric initial model, i.e. multiple initializations given the observed geometry conditions and per-estimates in the near infrared. Optimal and noiseless results will be obtained with increased number of random initializations to converge towards the ‘true’ solution. This random initialization issue and the associated computing cost, is classic for genetic algorithms (Davis, 1991) and the new generation of satellite products such as the Soil Moisture Ocean Salinity (SMOS) product (Font et al., 2010).

6 References

Aiken, J., & Moore, G. (2000). Case 2 (S) Bright pixel atmospheric correction. *MERIS ATBD*, *2*, 6-6.

Antoine, D., Chami, M., Claustre, H., d'Ortenzio, F., Morel, A., Becu, G., Gentilli, B., Louis, F., Ras, J., Roussier, E., Scott, A., Tailliez, D., Hooker, S. B., Guevel, P., Deste, J.-F., Dempsey, C. & Adams, D. (Eds.) 2006. BOUSSOLE: A Joint CNRS-INSU, ESA, CNES, and NASA Ocean Color Calibration and Validation Activity NASA Technical Memorandum No. 2006-214147. NASA/GSFC, Greenbelt, MD.

Antoine, D., & Morel, A. (2005). MERIS ATBD 2.7: Atmospheric correction of the MERIS observations over ocean case 1 waters.

Antoine, D., & Morel, A. (1999). A multiple scattering algorithm for atmospheric correction of remotely sensed ocean colour (MERIS instrument): principle and implementation for atmospheres carrying various aerosols including absorbing ones. *International Journal of Remote Sensing*, *20*(9), 1875-1916.

Barker, K., Mazeran, C., Lerebourg, C., Bouvet, M., Antoine, D., Ondrusek, M., Zibordi,, Lavender, S. (2008), MERMAID : The MEris MAtchup In-situ Database,proceedings of the 2nd MERIS/(A)ATSR User Workshop, ESA/ESRIN, Italy, September 2008.

Bailey, S. W., & Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, *102*(1), 12-23.

Berk, A., Anderson, G. P., Bernstein, L. S., Acharya, P. K., Dothe, H., Matthew, M. W., ... & Hoke, M. L. (1999, October). MODTRAN4 radiative transfer modeling for atmospheric correction. In *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation* (pp. 348-353). International Society for Optics and Photonics.Bhat, H. S., & Kumar, N. (2010). On the derivation of the Bayesian Information Criterion. School of Natural Sciences, University of California.

Bricaud, A., Morel, A., Babin, M., Allali, K., & Claustre, H. (1998). Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models. *Journal of Geophysical Research: Oceans (1978–2012)*, *103*(C13), 31033-31044.

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, *2*(2), 121-157.

Cox, C. and W. H. Munk (1954a): Statistics of the sea surface derived from sun glitter. J. Mar. Res., 13, 198–227

Cox, C., & Munk, W. (1954). Measurement of the roughness of the sea surface from photographs of the sun's glitter. *JOSA*, *44*(11), 838-850.

Davis, L. (Ed.). (1991). *Handbook of genetic algorithms* (Vol. 115). New York: Van Nostrand Reinhold.

Dempster, A.P.; Laird, N.M.; Rubin, D.B. (1977). "Maximum Likelihood from Incomplete Data via the EM Algorithm". Journal of the Royal Statistical Society, Series B 39 (1): 1–38. JSTOR 2984875. MR 0501537.

Deuzé J.L, M. Herman, and R. Santer, « Fourier series expansion of the transfer equation in the atmosphere-ocean system », J. Quant. Spectrosc. Radiat. Transfer, vol. 41, no. 6, pp. 483-494, 1989.

Doxaran D., J.M. Froidefond, P. Castaing and M. Babin, (2009). Dynamics of the turbidity maximum zone in a macrotidal estuary (the Gironde, France): Observations from field and MODIS satellite data. Estuarine, Coastal and Shelf Science 81, 321–332. (Get PDF Reprint)

Doxaran, D., Froidefond, J. M., Lavender, S., & Castaing, P. (2002). Spectral signature of highly turbid waters: Application with SPOT data to quantify suspended particulate matter concentrations. *Remote sensing of Environment*, *81*(1), 149-161.

Font, J., Camps, A., Borges, A., Martín-Neira, M., Boutin, J., Reul, N., Mecklenburg, S. (2010). SMOS: The challenging sea surface salinity measurement from space. *Proceedings of the IEEE*, *98*(5), 649-665.

Frouin, R., & Pelletier, B. (2014). Bayesian Methodology for ocean color remote sensing.

Gordon, H. R. and M. Wang (1994). “Influence of oceanic whitecaps on atmospheric correction of ocean-color sensors.” Applied Optics 33(33): 7754-7763.

Gordon, H. R., Du, T., & Zhang, T. (1997). Atmospheric correction of ocean color sensors: analysis of the effects of residual instrument polarization sensitivity. *Applied optics*, *36*(27), 6938-6948.

Goyens, C., Jamet, C., & Schroeder, T. (2013). Evaluation of four atmospheric correction algorithms for MODIS-Aqua images over contrasted coastal waters. *Remote Sensing of Environment*, *131*, 63-75.

Harold W. Sorenson, (1980) "Parameter Estimation: Principles and Problems", Marcel Dekker.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A. & Smirnov, A. (1998). AERONET-A federated instrument network and data archive for aerosol characterization. *Remote sensing of environment*, *66*(1), 1-16.

IOCCG (2000): Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters. Edited by Shubha Sathyendranath, pp. 140.

Krasnopolsky, V. M., & Schiller, H. (2003). Some neural network applications in environmental sciences. Part I: forward and inverse problems in geophysical remote measurements. Neural Networks, 16(3), 321-334.

Jamet, C., Loisel, H., & Dessailly, D. (2012). Retrieval of the spectral diffuse attenuation coefficient Kd (λ) in open and coastal ocean waters using a neural network inversion. *Journal of Geophysical Research: Oceans (1978–2012)*, *117*(C10).

Jamet, C., Loisel, H., Kuchinke, C. P., Ruddick, K., Zibordi, G., & Feng, H. (2011). Comparison of three SeaWiFS atmospheric correction algorithms for turbid waters using AERONET-OC measurements. *Remote Sensing of Environment*, *115*(8), 1955-1965.

Jia, S., & Qian, Y. (2009). Constrained nonnegative matrix factorization for hyperspectral unmixing. *Geoscience and Remote Sensing, IEEE Transactions on*, *47*(1), 161-173

Laws, E. A. (1997). *Mathematical methods for oceanographers: An introduction*. John Wiley & Sons.

Levenberg, K. (1944). A Method for the Solution of Certain Non-Linear Problems in Least Squares. Quarterly of Applied Mathematics 2: 164–168.

Lin, C. J. (2007). Projected gradient methods for nonnegative matrix factorization. *Neural computation*, *19*(10), 2756-2779.

Maritorena, S., Siegel, D. A., & Peterson, A. R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Applied Optics*, *41*(15), 2705-2714.

McLachlan, G. J. (2004). Discriminant Analysis and Statistical Pattern Recognition. Wiley Interscience.

MERIS Level 2 Detailed Processing Model, Doc. no PO-TN-MEL-GS-0006, issue 7, revision 2, June 2005.

Mobley, C. D. (1998). *Hydrolight 4.0 Users Guide*. SEQUOIA SCIENTIFIC INC MERCER ISLAND WA.

Moore, G. F., Aiken, J., Lavender, S. (1999), The atmospheric correction scheme of water colour and the quantitative retrieval of suspended particulate matter in Case II waters : application to MERIS, International Journal of Remote Sensing, 20, 1713–1733.

Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., & Franz, B. A. (2007). Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. *Remote Sensing of Environment*, *111*(1), 69-88.

Morel, A., Claustre, H., Antoine, D., & Gentili, B. (2007). Natural variability of bio-optical properties in Case 1 waters: attenuation and reflectance within the visible and near-UV spectral domains, as observed in South Pacific and Mediterranean waters. *Biogeosciences Discussions*, *4*(4), 2147-2178.

Morel, A., Gentili, B., Chami, M., & Ras, J. (2006). Bio-optical properties of high chlorophyll Case 1 waters, and of yellow substance-dominated Case 2 waters. Deep-Sea Research, 53, 1439 − 1459.

Morel, A., and B. Gentili (1996), Diffuse reflectance of oceanic waters 3, Implication of bidirectionality for the remote-sensing problem, Applied Optics, 35, 4850–4862.

Park, Y., De Cauwer, V., Nechad, B., & Ruddick, K. (2004). Validation of MERIS water products for Belgian coastal waters: 2002-2003.

Petersen, W., Wehde, H., Krasemann, H., Colijn, F., & Schroeder, F. (2008). FerryBox and MERIS–Assessment of coastal and shelf sea ecosystems by combining in situ and remotely sensed data. *Estuarine, Coastal and Shelf Science*, *77*(2), 296-307.

Petersen, K. B., & Pedersen, M. S. (2008). The matrix cookbook. *Technical University of Denmark*, 7-15.

Pope, R. M., & Fry, E. S. (1997). Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements. *Applied optics*, *36*(33), 810-8723.

Reynolds, D. A. (1995). Speaker identification and verification using Gaussian mixture speaker models. *Speech communication*, *17*(1), 91-108.

Robinson, I. S. (2004). Measuring the oceans from space: the principles and methods of satellite oceanography. Springer Science & Business Media.

Ruddick K., De Cauwer V., Park Y. & Moore G. (2006). Seaborne measurements of near infrarewater-leaving reflectance: The similarity spectrum for turbid waters. Limnology and Oceanography, Vol. 51(2), pp. 1157–1179

Ruddick, K. G., De Cauwer, V., Van Mol, B. (2005) Use of the near infrared similarity reflectance spectrum for the quality control of remote sensing data, in Remote Sensing of the Coastal Oceanic Environment, edited by Frouin, Robert J., Babin, Marcel, Sathyendranath, Shubha. Proceedings of the SPIE, Volume 5885, pp. 1–12.

Santer, R., Carrere, V., Dubuisson, P., & Roger, J. C. (1999). Atmospheric correction over land for MERIS. *International Journal of Remote Sensing*, *20*(9), 1819-1840.

Saulquin, B., Hamdi, A., Gohin, F., Populus, J., Mangin, A., & d'Andon, O. F. (2013). Estimation of the diffuse attenuation coefficient KdPAR using MERIS and application to seabed habitat mapping. *Remote Sensing of Environment*, *128*, 224-233.

Schiller, H., & Doerffer, R. (1999). Neural network for emulation of an inverse model operational derivation of Case II water properties from MERIS data. International Journal of Remote Sensing, 20(9), 1735-1746.

Steinmetz, F., Deschamps, P. Y., & Ramon, D. (2011). Atmospheric correction in presence of sun glint: application to MERIS. *Optics express*, *19*(10), 9783-9800.

Saulquin, B., Fablet, R., Ailliot, P., Mercier, G., Doxaran, D., Mangin, A., ... & Hembise, O. (2015). Characterization of Time-Varying Regimes in Remote Sensing Time Series: Application to the Forecasting of Satellite-Derived Suspended Matter Concentrations. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, *8*(1), 406-417.

Wang, M., Son, S., & Harding, L. W. (2009). Retrieval of diffuse attenuation coefficient in the Chesapeake Bay and turbid ocean regions for satellite ocean color applications. *Journal of Geophysical Research: Oceans (1978–2012)*, *114*(C10).

Werdell, P. J., & Bailey, S. W. (2005). An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, *98*(1), 122-140.

Zibordi, G., Mélin, F., Berthon, J. F., Holben, B., Slutsker, I., Giles, D., ... & Seppälä, J. (2009). AERONET-OC: a network for the validation of ocean color primary products. *Journal of Atmospheric and Oceanic Technology*, *26*(8), 1634-1651.

7 List of caption

Figure 1: no caption

Fig2 left: In-situ $\rho_w\,(\lambda)$ radiometric dataset

Fig2 right: MERIS TOA $\rho_{gc}\,(\lambda)$ radiometric dataset

Fig 3: $\rho_{aer}\,(\lambda) + \rho_{coupl}(\lambda)$ coastal reference spectra

Fig4 left: Reference spectral signatures

Fig4 right: $\rho_w\,(\lambda)$ coastal reference spectra

Fig5: $\rho_w(412), \rho_w(442),$

$\rho_w(490), \rho_w(560),$

$\rho_w(665), \rho_w(681)$

Fig6: Histogram $\rho_w(412),$ Histogram $\rho_w(442),$

Histogram $\rho_w(510)$, Histogram $\rho_w(560)$

Fig7: Histogram $\rho_w(412)$, Histogram $\rho_w(442)$,
Histogram $\rho_w(510)$, Histogram $\rho_w(560)$

Fig 8: MEETC2 $\rho_w(412)$, MEETC2 $\rho_w(442)$, MEETC2 $\rho_w(560)$, MEETC2 $\rho_w(681)$
MEGS $\rho_w(412)$, MEGS $\rho_w(442)$, MEGS $\rho_w(560)$, MEGS $\rho_w(681)$
C2R $\rho_w(412)$, C2R $\rho_w(442)$, C2R $\rho_w(560)$, C2R $\rho_w(681)$

Fig 9: MEETC2 $\rho_w(412)$, MEETC2 $\rho_w(442)$, MEETC2 $\rho_w(560)$, MEETC2 $\rho_w(681)$
MEGS $\rho_w(412)$, MEGS $\rho_w(442)$, MEGS $\rho_w(560)$, MEGS $\rho_w(681)$
C2R $\rho_w(412)$, C2R $\rho_w(442)$, C2R $\rho_w(560)$, C2R $\rho_w(681)$

Fig 10a: Type 1: SPM dominated ρ_w

Fig 10b: Type 2: Cdom dominated ρ_w

Fig 10c: Type 3: Chl-a dominated ρ_w

Fig 10d: Type 4: pure water dominated ρ_w