Application of the support vector regression method for turbidity assessment with MODIS on a shallow coral reef lagoon (Voh-Koné-Pouembout, New Caledonia)

Guillaume Wattelez, Cecile Dupouy, Jérôme Lefevre, Sylvain Ouillon, Jean-Michel Fernandez, Farid Juillot

To cite this version:

HAL Id: ird-01725040
https://ird.hal.science/ird-01725040
Submitted on 7 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Article

Application of the Support Vector Regression Method for Turbidity Assessment with MODIS on a Shallow Coral Reef Lagoon (Voh-Koné-Pouembout, New Caledonia)

Guillaume Wattelez 1,2,*, Cécile Dupouy 1,2, Jérôme Lefèvre 1,2,3, Sylvain Ouillon 3, Jean-Michel Fernandez 4 and Farid Juillot 2,5

1 Aix Marseille Univ, Université de Toulon, CNRS, IRD, MIO UM 110, 13288 Marseille, France
2 Institut de Recherche pour le Développement (IRD), BP A5 98848 Nouméa CEDEX, New Caledonia; cecile.dupouy@ird.fr (C.D.); jerome.lefevre@ird.fr (J.L.); farid.juillot@ird.fr (F.J.)
3 LEGOS, Université de Toulouse, IRD, CNES, CNRS, UPS, 14 avenue Edouard Belin, 31400 Toulouse, France; sylvain.ouillon@legos.obs-mip.fr
4 Analytical and Environmental Laboratory (AEL), IRD-Nouméa, BP A5, 98800 Nouméa, New Caledonia; jmfernandez@ael-environnement.nc
5 Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie (IMPMC), UMR IRD 206, UMR CNRS 7590, MNHN, Université Pierre et Marie Curie, Campus Jussieu, 75005 Paris, France

* Correspondence: guillaume.wattelez@univ-nc.nc; Tel.: +687-290-591

Received: 23 June 2017; Accepted: 19 September 2017; Published: 27 September 2017

Abstract: Particle transport by erosion from ultramafic lands in pristine tropical lagoons is a crucial problem, especially for the benthic and pelagic biodiversity associated with coral reefs. Satellite imagery is useful for assessing particle transport from land to sea. However, in the oligotrophic and shallow waters of tropical lagoons, the bottom reflection of downwelling light usually hampers the use of classical optical algorithms. In order to address this issue, a Support Vector Regression (SVR) model was developed and tested. The proposed application concerns the lagoon of New Caledonia—the second longest continuous coral reef in the world—which is frequently exposed to river plumes from ultramafic watersheds. The SVR model is based on a large training sample of in-situ turbidity values representative of the annual variability in the Voh-Koné-Pouembout lagoon (Western Coast of New Caledonia) during the 2014–2015 period and on coincident satellite reflectance values from MODerate Resolution Imaging Spectroradiometer (MODIS). It was trained with reflectance and two other explanatory parameters—bathymetry and bottom colour. This approach significantly improved the model’s capacity for retrieving the in-situ turbidity range from MODIS images, as compared with algorithms dedicated to deep oligotrophic or turbid waters, which were shown to be inadequate. This SVR model is applicable to the whole shallow lagoon waters from the Western Coast of New Caledonia and it is now ready to be tested over other oligotrophic shallow lagoon waters worldwide.

Keywords: turbidity; remote-sensing; MODerate Resolution Imaging Spectroradiometer (MODIS); Support Vector Regression (SVR); oligotrophic lagoon; bathymetry; reflectance; seabed colour; coral reef; New Caledonia

1. Introduction

In numerous tropical Pacific islands, the clarity of coastal lagoon waters is an essential parameter allowing the development of massive coral reefs and numerous benthic living species of prime importance for ecology and for fishing. This richness is essentially due to the oligotrophy of
surrounding oceanic waters as it is found along the Great Barrier Reef and many other Pacific Islands. In these areas, human forcing on river-derived inputs of sediments, nutrients and organic matter to the coastal oceans can have negative effects on marine biogeochemical cycles and biodiversity [1–3]. It is therefore important to monitor these inputs at high spatial and time scales in order to estimate both their temporal and spatial fluctuations and to anticipate their possible influences on marine biota from ocean colour remote sensing [4].

With about one third of its terrestrial surface (8000 km\(^2\)) covered with ultramafic rocks, about 85% of endemic terrestrial plants and trees species, 24 tree species of the 70 identified in mangrove ecosystems worldwide, about 2800 species of marine molluscs and with the second longest continuous coral reef in the world [5–8], New Caledonia is one of the tropical and intertropical areas most concerned with the potential impacts of anthropogenic forcings at continental margins on marine biodiversity [9]. The main island (Grande Terre) of this small archipelago is characterized by a large occurrence of ultramafic rocks (i.e., peridotites) as the geological setting [10]. Strong weathering of these rocks upon tropical climate lead to deep lateritic covers that are enriched in trace metals like nickel or cobalt [11–14]. Natural geological and climatic events have then made the lateritic covers on ultramafic rocks a very important economical resource for New Caledonia [15]. However, these pedogeological formations are subject to acute erosion [16] and this natural process is significantly enhanced by mining activities [17–19]. Due to the shoreline location of these covers, eroded lateritic materials are directly transported to the coastal ecosystems, as evidenced by remote sensing or sea measurements [20–22].

Remote sensing provides efficient tools for monitoring sediment transport at high spatial and temporal scales since it offers a synoptic and instantaneous field view of the total suspended matter (TSM) concentration (e.g., [23–29]). Hu et al. [30] first determined a single band algorithm using the MODerate Resolution Imaging Spectroradiometer (MODIS) 645 nm-reflectance for mapping the turbidity in the Tampa Bay (Florida, USA). More recently, Nechad et al. [31] and Novoa et al. [32] proposed single band algorithms using channels 520 to 885 nm based on equations of the radiative transfer. Dogliotti et al. [33] developed a general algorithm designed to map turbidity concentrations from 2 to 1000 FNU, with a switching band algorithm that uses the red 645 nm band for low turbidity values (i.e., lower than 15 FNU) and the Near Infrared (NIR) 859 nm band for high turbidity values (up to 1000 FNU). Other attempts for estimating turbidity from MODIS NASA algorithms have been proposed on the basis of either MODIS-645 nm reflectance [34,35] or the ratios of the MODIS reflectance at 645 nm over 667 nm [25,36–38]. More recently, supervised methods based on classification of spectrally-enhanced quasi-true colour MODIS images have also been proposed by Álvarez-Romero et al. [34] for mapping river plumes in the Great Barrier Reef (Australia).

The only algorithm available at the moment for the oligotrophic waters of the New Caledonian lagoon is the one developed by Ouillon et al. [39]. However, this algorithm that relies on polynomial and exponential models using in-situ reflectance channels over deep waters or turbid waters is not suitable for the oligotrophic and shallow waters (shallower than 5 m) of the Western lagoon of New Caledonia. This is probably because the effect of bottom reflectance over the coral reefs ecosystems for the retrieval of water quality parameters such as chlorophyll-\(a\) concentration ([chl-\(a\)]) [40] or turbidity [41] is particularly strong in this context. Indeed, such a contribution of both bathymetry and bottom colour in oligotrophic waters has already been shown to impact the detection of [chl-\(a\)] from MERIS reflectance [42,43] and AVNIR2/MODIS reflectance [44]. In a similar context, analytical algorithms using 8 “pure” bottom end-members pointed to the same conclusion [45]. However, it has also been shown that the remote sensing reflectance signal shows no significant contamination (\(R_{\text{corr}} < 0.0005\)) from bottom reflectance for water depths larger than 17 m for MODIS images with the brightest reflectance (i.e., white sands and corals such as those found at some places along the Great Barrier Reef [46]).

For about two decades, supervised learning based on neural networks or support vector machines (SVM) has been largely used to estimate oceanic parameters [47–49]. Zhan et al. [50]
The lagoon is particularly concerned by the enhanced inputs of sediments due to mining activities in the lagoon and open ocean waters of New Caledonia. A similar approach already gave interesting improvements for [chl-a] assessment in the lagoon and open ocean waters of New Caledonia [53]. Due to the possible strong influence of bathymetry and bottom colour, our support vector regression (SVR) model considers not only reflectance channels but also these two physical parameters as independent variables. Comparison of the results of our approach with published algorithms for estimation of turbidity from in-situ reflectance channels [39] or from MODIS images [33] emphasizes the potential of our model at retrieving the in-situ turbidity in shallow oligotrophic waters.

2. Materials and Methods

2.1. Study Area

New Caledonia is a South Pacific archipelago located between longitudes 162° and 169° E and latitudes 19° and 23° S. The study area—the Voh-Koné-Pouembout (VKP) lagoon in the Northern Province of New Caledonia—extends from 164.5° to 164.9° E and from 20.89° to 21.22° S (Figure 1). This lagoon is particularly concerned by the enhanced inputs of sediments due to mining activities since the Koniambo Nickel SAS (KNS) company started mining nickel at the Koniambo regolith in 2013. During the arrangement of the Koniambo regolith for mining vehicle access, as well as the construction of the nickel pyrometallurgical plant including large dredging in the lagoon, the whole area was monitored in order to assess the possible environmental impacts, especially on fish, coral reefs and marine vegetation [8]. Although it can be controlled by both river discharge and resuspension [18], turbidity was defined as an indicator for assessing water quality in the lagoon [54].

![Figure 1](image-url)  
**Figure 1.** Visited stations at the Voh-Koné-Pouembout (VKP) lagoon (Google Maps Terrain overlay). Points colours correspond to bathymetry, i.e., ◆: 0–10 m depth; ○: 10–20 m depth; ◼: 20–30 m depth; □: > 30 m depth.
The three main bays of the VKP lagoon are the Chasseloup Bay, the Vavouto Bay and the Katavili Bay (Figure 1). The watersheds contributing to water discharge in these bays are principally drained by the Voh, the Taléa/Coco, the Confiance and the Koné rivers.

The major fraction of the shoreline at this area is made of mangrove, a very productive ecosystem that protects the coast from erosion, acts as a refuge for marine biodiversity and potentially contributes to CO₂ fixation [55]. The sea bottom is made of mud (red to grey), sand (white to grey), fringing or reticulated coral reefs (white) or vegetation as sea grass or algae (grey) and it is delimited by a barrier reef [8].

The bathymetry of the VKP lagoon (Figure 2a) was extracted from the official database of New Caledonia administration [56]. A double-check was achieved by determining at each station the maximal depth recorded by the Conductivity Temperature Depth (CTD) probe (between 5 and 20 profiles per station performed in 2014 and 2015, Figure 2b). Following this protocol, the maximal water depth is 63 m. About 75% of the stations show a depth lower than 13 m, and more than 90% show a depth lower than 30 m (Figure 2b). The deepest waters are located in channels at and around passes to the open ocean while 70% of the lagoon shows a water depth lower than 5 m (Figure 2a).

![Figure 2](image.png)

**Figure 2.** (a) Map of the bathymetry (in m) at the Voh-Koné-Pouembout (VKP) lagoon. Points colours correspond to bottom colour, i.e., ○: white bottom; ■: grey bottom; ●: brown bottom; black areas correspond to land and those near the barrier reef are emerged reefs. (b) Histogram of the measured bathymetry on the visited stations. (c) Histogram of the in-situ turbidity values measured along CTD profiles.
2.2. Data

2.2.1. Field Measurements

In-situ turbidity values were collected by Analytical Environmental Laboratory (AEL) during a two-year survey (2014–2015) performed for the KNS company in an environmental monitoring context. During this survey, 76 stations were monitored with a SeaBird 19+ CTD probe (Bellevue, WA, USA) that provided measurements at several depths for different parameters including turbidity (in NTU) and fluorescence measured with an ECOFLNTU (from WetLabs, Philomath, OR, USA), pH, oxygen concentration, and salinity according to protocols described in [18]. The bottom colour at each station was estimated from visual in-situ observations [57].

In-situ turbidity values along profiles were generally low, but could exceptionally exceed 10.0 NTU offshore (an exceptional value of 24.0 NTU was recorded on 21 June 2014). More than 70% of turbidity values were below 1.0 NTU and more than 85% were below 2.0 NTU (Figure 2c). These turbidity values are typical of the New Caledonia lagoon [18,39], as well as of the Great Barrier Reef, depending on rain intensity [26,34]. In order to provide in-situ turbidity values representative of the CTD profile, we used the median value of all filtered values over a 10 m depth rather than taking the median values of the 3 first meters as in [39]. Such a calculation aimed at taking into account the variations of turbidity along the water column [18,39].

2.2.2. Satellite Data

MODIS Aqua images were processed from the level 1A to level 2 by creating 250 m resolution data as in Bailey and Werdell [58] for all MODIS data over New Caledonia [59]. Atmospheric corrections were made by default (SeaDAS, but a specific flag was applied composed of 6 SeaDAS flags Land, Cloud, High Sun Glint, Stray Light, High TOA Radiance and Atmospheric Correction Failure adapted to shallow coastal lagoons [59]). As a result, only a few pixels with negative reflectance values in the NIR (1240 nm) were found and subsequently eliminated from the coincidence research. Moreover, the match-ups with an $R_{\text{rs}}$ (1240) value above 0.001 Sr$^{-1}$ were discarded because such values in infrared channels were considered as indicative of wrong atmospheric corrections or of the presence of emerged reefs within the pixel.

The product of this MODIS database is marine remote sensing reflectance ($R_{\text{rs}}$) available at 14 channels: 412, 443, 469, 488, 531, 547, 555, 645, 667, 678, 748, 859, 869 and 1240 nm. The processing provides also non-phytoplankton absorption coefficients ($a_{\text{dp}}$), particulate backscattering coefficients ($b_{\text{bp}}$) at 7 channels: 412, 433, 488, 531, 547, 555 and 667 nm. Turbidity assessment according to Ouillon et al. [39] (see Equation (1) below) and Dogliotti et al. [33] were calculated as outputs of the Level2-imagery processing, as well as [chl-a] by Wattelez et al. [53].

\[
\text{TURB3} = \begin{cases} 
90.647 \left( \frac{R_{\text{rs}}(620)}{R_{\text{rs}}(510)} \right)^{0.594} & \text{if Turb} < 1 \text{ FTU} \\
\text{Turb} & \text{if Turb} \geq 1 \text{ FTU}
\end{cases}
\]

with Turb = $-6204217 R_{\text{rs}}(681)^3 + 179652 R_{\text{rs}}(681)^2 + 36.49 R_{\text{rs}}(681) + 0.452$

2.2.3. Match-Ups

494 match-ups from MODIS Aqua images were selected using a 0.01° square (about 1 × 1 km$^2$) centred on the visited station and in a 2-day temporal window [53,58]. Table 1 summarizes numbers of match-ups according to the campaigns periods and lists the corresponding MODIS files.
Table 1. Campaigns periods, number of match-ups per period and corresponding MODIS files.

<table>
<thead>
<tr>
<th>Period</th>
<th>Number of CTD Stations</th>
<th>Number of Match-Ups</th>
<th>MODIS Files</th>
</tr>
</thead>
<tbody>
<tr>
<td>22–23/4/2014</td>
<td>31</td>
<td>30</td>
<td>A2014111025500 A2014113024000</td>
</tr>
<tr>
<td>19–29/5/2014</td>
<td>59</td>
<td>53</td>
<td>A2014138023500 A2014139032000 A2014141030500 A2014145024000 A2014147023000</td>
</tr>
<tr>
<td>24–27/6/2014</td>
<td>43</td>
<td>40</td>
<td>A2014173030500 A2014175025500 A2014177024000</td>
</tr>
<tr>
<td>28–30/7/2014</td>
<td>35</td>
<td>29</td>
<td>A2014206021000 A2014207025500 A2014209024000 A2014211023000 A2014229021500</td>
</tr>
<tr>
<td>18–20/8/2014</td>
<td>8</td>
<td>7</td>
<td>A2014227023000 A2014229021500</td>
</tr>
<tr>
<td>28–30/10/2014</td>
<td>28</td>
<td>27</td>
<td>A2014301030500 A2014302021000</td>
</tr>
<tr>
<td>21–23/1/2015</td>
<td>34</td>
<td>34</td>
<td>A2015021032500 A2015024021500</td>
</tr>
<tr>
<td>19–27/3/2015</td>
<td>33</td>
<td>32</td>
<td>A2015079022000 A2015085032500 A2015086023000</td>
</tr>
<tr>
<td>28–30/4/2015</td>
<td>24</td>
<td>19</td>
<td>A2015116024000 A2015117032500 A2015121030000</td>
</tr>
<tr>
<td>6–7/5/2015</td>
<td>18</td>
<td>17</td>
<td>A2015128030500</td>
</tr>
<tr>
<td>22–30/6/2015</td>
<td>36</td>
<td>35</td>
<td>A2015173023500 A2015174032000 A2015175022500 A2015176030500 A2015180024000</td>
</tr>
<tr>
<td>22–27/7/2015</td>
<td>17</td>
<td>17</td>
<td>A2015205023500 A2015207022000</td>
</tr>
<tr>
<td>17–26/8/2015</td>
<td>35</td>
<td>26</td>
<td>A2015228024000 A2015234020500 A2015236033000</td>
</tr>
<tr>
<td>28/9/2015–1/10/2015</td>
<td>47</td>
<td>45</td>
<td>A2015272030500 A2015273021000 A2015274025000</td>
</tr>
<tr>
<td>26–30/10/2015</td>
<td>41</td>
<td>40</td>
<td>A2015299024500 A2015300033000 A2015302031500</td>
</tr>
<tr>
<td>16–20/11/2015</td>
<td>39</td>
<td>39</td>
<td>A2015325032500</td>
</tr>
<tr>
<td>22/12/2015</td>
<td>6</td>
<td>4</td>
<td>A2015357032500</td>
</tr>
</tbody>
</table>

Satellite values were assigned according to three different methods as preconized for the research of coincident pixels [58], i.e., with the closest neighbour method (CL), the weighted mean method (WMM) and the filtered mean method (FMM). This approach has already been successfully used for lagoon waters of New Caledonia in Dupouy et al. [60] and Wattelez et al. [53].
2.3. Creation of the Support Vector Regression (SVR) Model

2.3.1. Sampling

Support Vector Regression (SVR) Models are built with a learning sample and then tested with a randomly selected test sample. In our study, the learning sample was constructed with 70% of the data and the test sample contained the remaining 30% of the data. This method is necessary to check the algorithm effectiveness without an overtraining effect. Each model was created and tested ten times (by using ten randomly selected samples). This process allowed selecting a model well fitted on average (and not on a particular random selection).

2.3.2. Indicators

Several indices were computed in order to compare the different models. These indices were the mean normalized bias (MNB), the mean normalized absolute error (MNAE), the mean absolute error (MAE) and the root mean square error (RMSE) with the following respective mathematical expressions:

\[
MNB (y) = \frac{1}{n} \sum_{i=1}^{n} \frac{x_i - y_i}{x_i}
\]

\[
MNAE (y) = \frac{1}{n} \sum_{i=1}^{n} \frac{|x_i - y_i|}{x_i}
\]

\[
MAE (y) = \frac{1}{n} \sum_{i=1}^{n} |x_i - y_i|
\]

\[
RMSE (y) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2}
\]

where \(n\) is the number of observations, \(x_i\) is the \(i^{th}\) in-situ observation, \(y_i\) is the \(i^{th}\) remote sensing assessment.

As a model was created and tested ten times, each index was computed ten times. Indicators of differences of two models were compared thanks to a paired Student’s \(t\)-test. Different models were also compared using the values range, the coefficient of determination (\(R^2\)) in both linear and log-regression modes.

2.3.3. Support Vector Regression

In this study, the SVR was built with the following parameters: in-situ turbidity as the explained variable, and remote sensing parameters as the explanatory variables. We first performed tests with all remote sensing parameters (i.e., \(R_{rs}\) in the visible spectra from 412 to 678 nm, and \(a_{dg}\) and \(b_{bp}\)) as explanatory variables. Considering the low turbidity values of our in-situ dataset, we decided in a first approach not to select the MODIS NIR channels (i.e., 748, 859, 869 and 1240 nm) available in the products data set. This option was chosen on the basis of previous studies [31,33,39] which suggested that these channels should not bring information in the low turbidity values range of our study. The first SVR model was therefore deliberately based on visible channels, and the bathymetry and bottom colour were added as explanatory variables to test if these physical parameters bring some significant information. However, to check this assumption, we re-integrated the NIR channels in a second step in order to check the capacity of these channels at improving our SVR model in the case of oligotrophic shallow waters.

The SVR was implemented by using the “svm” function of the R package “e1071” [61]. This function uses an epsilon-regression with a radial kernel whose \(\gamma\) parameter is equal to \(\frac{1}{m}\) where \(m\) is the number of explanatory variables, \(\varepsilon = 0.1\) for the insensitive-loss function and a cost parameter \(C = 1\) is used in the Lagrange formulation.
2.3.4. Algorithm Steps

The forward stepwise approach was used by successively adding the optical parameters, one by one. First, each optical parameter was tested as an explanatory variable in the model containing only one explanatory variable. The one giving the best results according to the aforementioned indicators was retained. Then, the selected model was expanded with a second optical parameter as an explanatory variable, selected according to values of the indicators, and so on.

At each step and for each optical parameter, 10 models were built with 10 different random learning samples and then tested with the 10 corresponding test samples, giving 10 different values for the indicators. If the added parameter did not bring information statistically significant, the last model was kept as the best model.

The last steps were the following: testing the bathymetry and the bottom colour significance, one by one, then removing some parameters in the model and testing again if results were significantly different. For a model, 10 values of an indicator were available (recall that there are 6 indicators, i.e., MNB, MNAE, MAE, RMSE, $R^2$ and log $R^2$). Then, comparison of the different models one to another by successive paired $t$-tests on the series of indicators enabled checking the significance of the indicators of differences. During this latter procedure, the $H_0$ hypothesis was “There is no significant difference between the two tested models”, whereas the $H_1$ hypothesis was “Indicators computed from the model using an additional parameter are better than the others”. We considered that the final model was the one providing the best indicators results and using the lowest number of parameters.

2.4. Interpolated Maps for In-situ Values

The resulting model of this study must be compared to other usual models and to in-situ values. Maps are a useful tool to clearly perceive spatial structures induced by models. But in-situ data are punctual, that is why they were interpolated before mapping.

In an aim of building a map of the interpolated data and comparing with a model applied on a MODIS image, ordinary kriging was implemented on the corresponding MODIS 250m-satellite grid. Turbidity values on all the stations were used to get an empirical variogram from which a variogram model (exponential, Gaussian or else according to the spatial variation structure) was designed with the “fit.variogram” function of the gstat R package. Then, the “krige” function was applied on data with the fitted variogram model. Finally, the kriging output was mapped with the colour scale used for the satellite data mapping.

3. Results

3.1. Evaluation of the SVR Model at Visible Wavelengths

At each step of our SVR approach, $R_s$ values generally provided better results than $R_s$ ratios. Other optical parameters such as $a_{dg}$ and $b_{bp}$ did not add information so they were discarded from the explanatory variables.

Our approach converged with a 3-parameters model that includes $R_s$ (555), $R_s$ (645) and $R_s$ (667) as optical parameters (Optical Model, O.M.), with bathymetry (B) and bottom colour (C) added as explanatory variables. The indicators values computed with the corresponding assessments are shown in Table 2 C.B.O.M. part. The first $t$-test ($t$-test 1) aimed at checking the significance of adding (B) to the (O.M.) to yield a (B.O.M.), whereas the second one ($t$-test 2) aimed at checking the significance of adding (C) to the (B.O.M.) to yield a (C.B.O.M.).
Table 2. Minimum, mean and maximum values for the series of indicators computed on each test sample with C.B.O.M. and C.B.NIR.M. (see Section 3.3). The t-tests p-values indicate the statistical significance of improvements between models. The t-test 1: O.M. vs. B.O.M.; the t-test 2: B.O.M. vs. C.B.O.M.; t-test 3: C.B.O.M. vs. C.B.NIR.M.

<table>
<thead>
<tr>
<th>Indicators</th>
<th>C.B.O.M.</th>
<th>C.B.NIR.M.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min.</td>
<td>Mean</td>
</tr>
<tr>
<td>MNB</td>
<td>−0.104</td>
<td>0.050</td>
</tr>
<tr>
<td>MNAE</td>
<td>0.211</td>
<td>0.233</td>
</tr>
<tr>
<td>MAE</td>
<td>0.094</td>
<td>0.139</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.126</td>
<td>0.235</td>
</tr>
<tr>
<td>R²</td>
<td>0.426</td>
<td>0.494</td>
</tr>
<tr>
<td>R² (log)</td>
<td>0.431</td>
<td>0.539</td>
</tr>
<tr>
<td>In-situ values</td>
<td>0.216</td>
<td>0.549</td>
</tr>
<tr>
<td>Assessment values</td>
<td>0.230</td>
<td>0.513</td>
</tr>
</tbody>
</table>

* t-test conditions are verified in this case; * t-test conditions were not verified in this case so a Wilcoxon paired rank test was applied.

The t-test 1 results clearly showed that (B) brings significant information in remotely-sensed assessment of turbidity in the VKP lagoon. Indeed, MNAE, MAE and RMSE computed with the (B.O.M.) were significantly lower than those computed with the (O.M.) and all the p-values were below 0.05 with a H₁ hypothesis being “The indicator is significantly below for the (B.O.M.)” (Table 2). Similarly, the R² and the log-R² computed with the (B.O.M.) were significantly larger than those computed with the (O.M.), the H₁ hypothesis being “The indicator is significantly above for the (B.O.M.)”. Similar results were obtained for the t-test 2, which indicated that adding (C) as an explanatory variable to the (B.O.M.) significantly improved the remotely-sensed assessments (Table 2).

Comparison of the in-situ turbidity with the remote-sensed turbidity assessed by our different SVR models showed that adding (B) and then (C) to the (O.M.) significantly improved the quality of the model (Figure 3), especially above brown bottoms. Despite this improvement, the highest turbidity values (i.e., around 2 NTU in this study), remain underestimated with our SVR model, even if both bathymetry (B.O.M., Figure 3b) and bottom colour (C.B.O.M., Figure 3c) parameters are used.
Water shows then that C.B.O.M. is more suited than a general model in the lagoon of the VKP area.

Figure 4 shows that the density of errors by tested on satellite data over oligotrophic shallow waters.

Log-linear regressions between in-situ turbidity and remote sensing turbidity assessed by the different optical SVR models. The red line is the first bisector. (a) O.M. Optical Model; (b) B.O.M. Optical Model + Bathymetry; (c) C.B.O.M. Optical Model + Bathymetry + bottom Colour. Points colours correspond to bottom colour, i.e., white: white bottom; grey: grey bottom; brown bottom.

3.2. Comparison with Other Models

This SVR model was then compared with the already existing algorithms of Ouillon et al. [39] (hereafter referred as O2008, set for New Caledonia waters, based on in-situ reflectance data) and Dogliotti et al. [33] (hereafter referred as D2015, based on MODIS images), that have not yet been tested on satellite data over oligotrophic shallow waters. Figure 4 shows that the density of errors is close to 0 with C.B.O.M., which is not the case for the O2008 and D2015 models. This comparison shows then that C.B.O.M. is more suited than a general model in the lagoon of the VKP area.

Figure 4. Error density distribution on the 10 test samples obtained with the different SVR models (i.e., C.B.O.M., B.O.M., and O.M.) and with the O2008 and D2015 models.
3.3. Using NIR Channels as Explanatory Variables

Testing the use of NIR channels (from 700 to 869 nm) through the three $R_{s}$ (859), $R_{s}$ (488)/$R_{s}$ (555) and $R_{s}$ (667)/$R_{s}$ (678) parameters for another SVR model including both bathymetry and bottom colour yields a new model (i.e., C.B.NIR.M.) that improved the quality of turbidity assessment (Figure 5). This improvement can be evaluated by comparison of Figures 3b and 5, which shows that the highest turbidity values retrieved with C.B.NIR.M. are above those retrieved with C.B.O.M. and then slightly better fit the in-situ values.

Figure 5. Log-linear regression between in-situ turbidity and remote sensing turbidity retrieved with C.B.NIR.M. The red line is the first bisector. Only the high turbidity values (corresponding to brown-bottom stations) are not well retrieved by the SVR model. Points colours correspond to bottom colour, i.e., ○: white bottom; ●: grey bottom; ▾: brown bottom.

The statistical improvement of C.B.NIR.M. compared to C.B.O.M. is confirmed in Table 2 that compares the various indicators for both models with t-test 3 which aimed at checking a significant improvement in using C.B.NIR.M. instead of C.B.O.M. Table 2 indicates that RMSE (0.220 for C.B.NIR.M. and 0.235 for C.B.O.M. in mean), $R^{2}$ (0.553 for C.B.NIR.M. against 0.494 for C.B.O.M. in mean) and log-$R^{2}$ (0.590 for C.B.NIR.M. against 0.539 for C.B.O.M. in mean) are significantly improved with C.B.NIR.M. with t-tests $p$-values < 0.05. Considering the NIR channels as explanatory variables enables in particular better retrieval of the remote sensing turbidity values on brown bottom, the in-situ values of which are high (between 1.5 and 2 NTU).

Despite this improvement, the highest in-situ turbidity values are not well fitted. With a maximum value of 1.23 NTU, C.B.NIR.M. seems unable to assess high turbidity values compared to C.B.O.M., which provides a maximum value of 1.29 NTU. Both models yield slightly underestimated values.

3.4. Application to MODIS Images

The B.O.M. was applied on MODIS images despite the efficiency of C.B.O.M. at retrieving the in-situ turbidity as the full C.B.O.M. model could not be applied on MODIS images since the bottom colour (C) is not known at all the image pixels. Figure 6a–c shows respectively the in-situ and the remotely-sensed turbidity assessed by O2008 and the B.O.M. on the VKP lagoon area for 21 April 2014 (low turbidity period). The in-situ map (Figure 6a) was made from a kriging interpolation based on a Gaussian variogram model. This map shows a coastal enhancement of turbidity up to 1.5 NTU (164.8° lon, -21.1° lat), while stations in the middle part of the lagoon (164.6° lon, -21.0° lat) and over barrier reefs show a moderate turbidity of 0.3 NTU. High turbidity values shown near the barrier reef
(164.55° lon, −21.00° lat) are probably due to localised coral resuspension. Figure 6b (O2008) shows that high turbidity values were located in shallow waters near the coast (164.78° lon, −21.12° lat) as well as on shallow reef flats (164.75° lon, −21.15° lat) due to bottom effect. Figure 5c shows that with the B.O.M., some pixels with turbidity above the mean level (164.75° lon, −21.15° lat) were retrieved in shallow waters though generally high measured turbidity values were not retrieved everywhere.

**Figure 6.** Turbidity in the VKP lagoon area on 21 April 2014. (a) in-situ turbidity values (in NTU) interpolated by ordinary kriging and their histogram, as measured with the CTD; (b) Map and histogram of turbidity (in FTU) retrieved from the MODIS image with the O2008 model; (c) Map and histogram of the turbidity values (in NTU) retrieved from the MODIS image with our B.O.M. Black areas correspond to MODIS land mask and grey areas correspond to deep ocean. Points colours correspond to bottom colour, i.e., ○: white bottom; ●: grey bottom; ■: brown bottom. On maps (b) and (c) the white areas correspond to flagged pixels.

Similarly, Figure 7a–c show in-situ and retrieved turbidity values with O2008 and B.O.M. on the VKP area for the MODIS image captured on 24 June 2014. The in-situ map (Figure 7a) was made from
a kriging interpolation based on an exponential variogram model. For this day, in-situ values highlight a coastal enhancement too (164.68° lon, −20.98° lat). With B.O.M., assessed values around the barrier reef were quite low (<1.5 NTU, Figure 6c) whereas values assessed by O2008 were usually high (up to 5 NTU).

![Figure 6](https://example.com/figure6.png)

**Figure 6.** (a) in-situ turbidity values (in NTU) interpolated by ordinary kriging and their histogram, as measured with the CTD; (b) Map and histogram of turbidity (in FTU) retrieved from the MODIS image with the O2008 model; (c) Map and histogram of the turbidity values (in NTU) retrieved from the MODIS image with our B.O.M. Black areas correspond to MODIS land mask and grey areas correspond to deep ocean. Points colours correspond to bottom colour, i.e., ◯: white bottom; ⊙: grey bottom; ●: brown bottom. On maps (b) and (c) the white areas correspond to flagged pixels.

Finally, both Figures 6c and 7c on the two selected MODIS images of April and June confirm the capacity of the B.O.M. at retrieving the in-situ turbidity of the oligotrophic shallow waters of the West Coast of New Caledonia as high values do not appear on reefs.

![Figure 7](https://example.com/figure7.png)

**Figure 7.** Turbidity in the VKP lagoon area on 24 June 2014. (a) in-situ turbidity values (in NTU) interpolated by ordinary kriging and their histogram, as measured with the CTD; (b) Map and histogram of turbidity (in FTU) retrieved from the MODIS image with the O2008 model; (c) Map and histogram of the turbidity values (in NTU) retrieved from the MODIS image with our B.O.M. Black areas correspond to MODIS land mask and grey areas correspond to deep ocean. Points colours correspond to bottom colour, i.e., ◯: white bottom; ⊙: grey bottom; ●: brown bottom. On maps (b) and (c) the white areas correspond to flagged pixels.

Finally, both Figures 6c and 7c on the two selected MODIS images of April and June confirm the capacity of the B.O.M. at retrieving the in-situ turbidity of the oligotrophic shallow waters of the West Coast of New Caledonia as high values do not appear on reefs.
Tables 3 and 4 exhibit the main quantile values for the turbidity assessments by different models for the days of 21 April 2014 and 24 June 2014 respectively. As expected, the ranges of turbidity values retrieved by B.O.M. are strongly reduced and their range (from 0.1 to 1.40 NTU, Figure 3b) more accurate than those retrieved by the O2008 model (Figures 6 and 7). These results indicate that, despite its difficulty at assessing the highest turbidity values, the present SVR model is more suited than the O2008 model that overestimated turbidity with more than 25% for pixels with a turbidity value above 5 FTU.

Table 3. Main quantile values of turbidity estimations by the O2008 model, the B.O.M. and the B.NIR.M. (see Sections 3.1 and 3.3) in the Voh-Koné-Pouembout lagoon area on 21 April 2014.

<table>
<thead>
<tr>
<th>Model</th>
<th>Min.</th>
<th>1st Decile</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>9th Decile</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>O2008 (FTU)</td>
<td>0.0500</td>
<td>0.1360</td>
<td>0.4418</td>
<td>1.8898</td>
<td>5.6038</td>
<td>11.4643</td>
<td>23.4488</td>
</tr>
<tr>
<td>B.O.M. (NTU)</td>
<td>0.1036</td>
<td>0.4084</td>
<td>0.5195</td>
<td>0.6615</td>
<td>0.7768</td>
<td>0.8885</td>
<td>1.3938</td>
</tr>
<tr>
<td>B.NIR.M. (NTU)</td>
<td>0.2294</td>
<td>0.3880</td>
<td>0.4503</td>
<td>0.6029</td>
<td>0.7264</td>
<td>0.8191</td>
<td>1.3116</td>
</tr>
</tbody>
</table>

Table 4. Main quantile values of turbidity estimations by the O2008 model, the B.O.M. and the B.NIR.M. (see Sections 3.1 and 3.3) in the Voh-Koné-Pouembout lagoon area on 24 June 2014.

<table>
<thead>
<tr>
<th>Model</th>
<th>Min.</th>
<th>1st Decile</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>9th Decile</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>O2008 (FTU)</td>
<td>0.0500</td>
<td>0.05</td>
<td>0.1209</td>
<td>0.7162</td>
<td>3.2655</td>
<td>6.8045</td>
<td>30.0500</td>
</tr>
<tr>
<td>B.O.M. (NTU)</td>
<td>0.1190</td>
<td>0.3935</td>
<td>0.4862</td>
<td>0.6373</td>
<td>0.7611</td>
<td>0.8964</td>
<td>1.4139</td>
</tr>
<tr>
<td>B.NIR.M. (NTU)</td>
<td>0.2232</td>
<td>0.4104</td>
<td>0.5125</td>
<td>0.6465</td>
<td>0.7914</td>
<td>0.8904</td>
<td>1.3017</td>
</tr>
</tbody>
</table>

4. Discussion

4.1. Validity of the SVR for Turbidity or SPM Estimation, and Comparison to Previous Algorithms

The SVR model was tested using in-situ turbidity with a SeaBird 19+ CTD calibrated during the period 2014–2015. A relationship can be inferred between turbidity and Suspended Particulate Matter (SPM) [62,63] but this relation is highly dependent on the area explored and on the season considered as well [64]. Consequently, the present SVR model can be used to estimate suspended matter concentration, but only on the VKP region where this regression was set.

The SVR optical model uses 3 channels in the visible, i.e., $R_{555}$, $R_{645}$ and $R_{667}$. Many algorithms have used the 667 nm wavelength (see for instance [26,33]). It has been shown that the sensitivity of one-band algorithms depends on both wavelength and turbidity range, with reflectance at shorter wavelengths more sensitive to low turbidity and reflectance at longer wavelengths more sensitive to high turbidity [31,39,65]. Even if these proposed algorithms show performance with low mean relative errors on a large turbidity range (e.g., with a 20% RMSE for turbidity ranging from 1 to 1000 FNU [33]), they were not developed for oligotrophic shallow tropical waters.

The O2008 algorithm developed over the Southern Lagoon of New Caledonia is designed for oligotrophic waters. However, it is restricted to water depth > 14 m or to water with turbidity > 1 FTU and depth > 10.5 m. In its present state, it can then not be applied successfully to the shallow parts of the lagoon where the bottom influence is not negligible. Indeed, in that context, the upwelling light emerging from the sea surface is affected by the bottom reflectance that is in fact composed of two terms ($R_{\text{water}}$ and $R_{\text{bottom}}$). The measured $R_{\text{ns}}$ can thus no more be considered to infer the inversion algorithm because its value over shallow waters is higher than the $R_{\text{water}}$ value. This mismatch will yield an overestimation of the turbidity retrieved from the O2008 or D2015 (for the same reason) algorithms compared to that retrieved using the SVR method.

As shown on Figure 8, the strongest differences between the O2008 model and our B.O.M. are on white bottom stations and on very shallow brown bottom stations. This figure also shows that the difference in retrieved turbidity between a generic algorithm established for oligotrophic and deep waters (O2008 model) and a SVR model (B.O.M.) decreases with increasing water depth (Figure 8).
The addition of a NIR channel (C.B.NIR.M.) slightly improved the model performance, in particular over the brown bottoms. The smaller penetration depth of near infra-red light as compared to the visible bands (due to the high absorption of light by water molecules, see for example References [66–68]) makes the upwelling radiance much more dependent on water and dissolved or suspended matter properties than on bottom colour. That is likely the reason why the performance of the C.B.NIR.M model did not depend anymore or very little on the bottom colour (see Figure 5).

Petus et al. [25] found that the MODIS-Aqua band at 859 nm was not sensitive enough to detect turbidity variations between 0.01 and 10 NTU in the Adour River. Nevertheless, as this study context is widely different from ours (oligotrophic waters), good results provided by C.B.NIR.M. may encourage us to test in the future another model using the visible and near-infrared bands but without considering anymore the bottom colour as a possible explanatory variable. However, the C.B.NIR.M. failed more than the C.B.O.M. at retrieving the highest values of turbidity with underestimation of the highest values by 49% and 46.5%, respectively (Table 2). This performance can be explained by reasons that are discussed in the following part, such as a temporal window of 48 hours between satellite overpass and field measurements, and that can be considered in the next applications of this method. Another reason of discrepancy may be considered in future tests as the penetration depth of light is very low in the NIR. Considering turbidity averaged over 10 m for model training could be reduced to a smaller depth below the surface in next applications.

4.2. Other Possible Improvements of Our Modelling Approach

4.2.1. Vertical Heterogeneity of In-situ Turbidity Profiles

Examination of the in-situ turbidity vertical profiles indicates that the coastal VKP lagoon area is rather mixed as is the South-Western lagoon [18,69]. We then decided to link the ocean colour on each pixel to the median value of the in-situ turbidity value from 0 to 10 m depth. However, the ocean colour on a pixel actually depends on turbidity values in the water column weighted by an exponential function of the depth of measurement [66–68] and, in shallow waters, on the bottom colour [70–72]. It could then be interesting to depict more precisely the function that links the surface value of in-situ turbidity to the vertical profile and to the bottom colour by considering separately the SPM and the seabed contributions to this parameter.
4.2.2. Turbidity Values Distribution

Figures 3 and 5 highlighted that both C.B.O.M. and C.B.NIR.M. underestimated high in-situ turbidity values. This limitation is considered to be linked to the asymmetric distributions of in-situ turbidity values in our learning data set (Figures 6a and 7a) and it should then be improved when a sufficient number of high turbidity values is available for the training of the SVR model (i.e., along a greater sampling period including significant climatic events). Since it is usually difficult to obtain images with plumes and high turbidity right after a significant climatic event because of the cloud coverage, an alternative way would be to train the SVR model with in-situ reflectance values rather than with satellite values.

Moreover, customizing the parameters of our SVR model, including the kernel function and the $C$ cost parameter, could improve the results by providing a distribution of retrieved values closer to the distribution of the in-situ values.

4.2.3. Match-Up Research Procedure

A limitation of our approach during the match-up process in the lagoon waters from the VKP area in New Caledonia is the use of methods mainly developed for the open ocean [58], where sea floor and bathymetry do not influence the ocean colour and where spatial and temporal changes in biogeochemical parameters are quite low. However, in lagoon waters, many localized and transient phenomena, such as upwelling, river inputs and resuspension due to wind bursts, can influence in-situ turbidity values at short distance and time scales. A future improvement of our approach could then consist in reducing both the temporal and spatial windows that are used during the match-up process. However, such a reduction would imply a reduced number of coincidences, which would raise the issue of data representativeness and significance to create a robust model.

4.2.4. Model Conception

Generic models designed to remotely assess turbidity values are generally customized with in-situ turbidity values and in-situ reflectance values. By this way, resulting models have only to be fitted according to remote reflectance sensors. Since the current SVR model was customized with in-situ turbidity values and remote reflectance values from Aqua-MODIS, it is highly dependent from the MODIS sensors. Nevertheless, the methodology developed with this SVR model should be easily used with other sensors providing a sufficient training dataset in coincidence with in-situ data is available.

An alternative could be to develop a SVR model from in-situ turbidity and $R_{rs}$ for any sensor obtained from in-situ hyperspectral $R_{rs}$ values and the spectral sensitivity of the given sensor. Such an opportunity would overcome the need of many sets of match-up for different sensors.

4.2.5. Spectral Classification of MODIS Pixels in the VKP Lagoon

Figure 9 shows the MODIS reflectance spectra obtained on all coincident pixels of the VKP lagoon area for the 2014–2015 period that was used to construct our SVR model. Grey bottom pixels’ reflectance spectra show little variability according to turbidity range. The highest $R_{rs}$ values were linked to the brown bottom pixels, and to white bottom pixels with turbidity values below 1 NTU. On white bottom pixels, the higher reflectance values for lower turbidity values clearly show the prevalence of the bottom effect in the optical signal. Although the current data set was too small to strongly support this statement, this effect seems to be opposite on brown bottom pixels. This latter point illustrates the difficulty to remotely assess turbidity in oligotrophic shallow waters and it then emphasizes the interest of the SVR method.
Figure 9. Spectra of MODIS reflectance on all pixels used for the construction of our SVR model and classified by both turbidity values (i.e., Upper row: Turb < 0.5 NTU, Middle row: 0.5 < Turb < 1, Lower row: Turb > 1) and bottom colour (i.e., Left: Brown bottom, Middle: Grey bottom, Right: White bottom).

4.2.6. Including the Bottom Colour

In the present SVR model, the optical signal of the bottom is integrated in the model conception and it is weighted by the bathymetry and the bottom colour. Some bathymetric patterns slightly appear on the derived turbidity maps (e.g., Figures 6c and 7c as compared to Figure 2) but a generalized integration of the bottom colour in the application should bring more accurate results. The main remaining challenge of this approach lies in constructing a SVR model that could include the bottom colour at each pixel. To reach this goal, a first approach would consist in collecting this variable by in-situ observations at the largest possible number of stations. However, this option appears difficult to apply at the lagoon scale. Another option would be to extrapolate a bottom colour from geological maps. An alternative approach could be to retrieve the bottom colour from an extremely clear image, where the water column is considered null, by using the Lyzenga’s method from ocean colour reflectance [70,73,74], and then operate a spectral classification to associate each pixel of the image to a bottom colour as successfully applied in other areas of New Caledonia [42,43]. Such a method will be applied on the VKP lagoon area in order to evaluate its improvements toward the C.B.O.M. for turbidity retrieving in oligotrophic shallow waters. Figures 5 and 6 show an application of the B.O.M. SVR on two particular days for MODIS. We may also use the Sentinel 2 or Sentinel 3 data in order to produce synoptic maps for studying the temporal variations on the whole area.

5. Conclusions

This paper introduced an empirical algorithm—based on the SVR method—for assessing turbidity values in oligotrophic shallow waters from MODIS images. This algorithm was tested on the oligotrophic shallow waters of the West Coast of New Caledonia, but it may be applicable to other
similar areas. The optical explanatory variables included in this SVR model were selected according to statistical considerations, and improving results of turbidity assessments given by generic algorithms which are not adapted to shallow oligotrophic waters. Since bathymetry and bottom colour showed to widely influence the remotely-sensed optical signal, both parameters were introduced as explanatory parameters in the SVR model. Despite the complex optical character of the waters at the VKP lagoon area studied, this latter approach significantly improved the capacity of our SVR model at retrieving the in-situ turbidity data in these oligotrophic shallow waters.

Since the SVR model introduced in this paper is based on a limited set of in-situ data with low turbidity values, extending the range of these data should improve its accuracy for higher turbidity values. Considering that this method is widely applicable on hyperspectral and multispectral remote sensors, such as Sentinels, it should allow a better monitoring of coastal shallow waters in coral reefs.

Acknowledgments: This work was financially supported by the Centre de Recherche pour le Nickel et son environnement (www.cnrt.nc/) under the project DYNAMINE, “Dynamique des metaux de la mine au lagon en Nouvelle-Caledonie” project (CSF N° 3PS2013-CNRT.IRD/DYNAMINE) and by French Institute of Research for the Development (IRD—www.ird.fr/). The PC-cluster used for the ocean colour applications was funded through an IRD/IFREMER collaboration and by the IRD SPIRALES Valhysat program. We gratefully acknowledge the NASA Ocean Biology Processing Group (OBPG) for making MODIS ocean-colour imagery and products available. We thank Koniambo Nickel SAS (KNS) for allowing access to the set of in-situ turbidity measurements. Finally, we especially thank Morgan Mangeas (IRD, UMR 228 ESPACE-DEV) for his advices about SVR modelling.

Author Contributions: G.W. and C.D. designed the study. J.-M.F. performed the in-situ measurements of turbidity. J.L. developed the MODIS database Valhysat software. G.W. developed the SVR model and adapted it to the site. G.W., C.D. and F.J. performed the statistics. C.D., G.W., F.J. and S.O. analysed the results and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References


36. Miller, R.L.; McKee, B.A. Using MODIS Terra 250 m imagery to map concentrations of total suspended matter in coastal waters. Remote Sens. Environ. 2004, 93, 259–266. [CrossRef]


71. Tolk, B.L.; Han, L.; Rundquist, D.C. The impact of bottom brightness on spectral reflectance of suspended sediments. *Int. J. Remote Sens.* **2000**, *21*, 2259–2268. [CrossRef]

