Systematical underestimation of MODIS global chlorophyll-a concentration estimation algorithm associating with scale effect

Jun Chen, Wenting Quan & Hui Sheng

1The key laboratory of marine hydrocarbon resources and geology, Qingdao Institute of Marine Geology, Qingdao 266071, China.
2Shanxi Remote Sensing Information Center Agriculture, 710 000, China
3College of Earth Resources and Information, China University of Petroleum, Qingdao 266 200, China

1[E-mail: cjun@cgs.cn]

Received 31 May 2012; revised 16 December 2012

A framework for characterizing the scale effects while using ocean chlorophyll 4 algorithm of MODIS (OC3M) to estimate chlorophyll-a concentration (chl\(a\)) from Moderate Resolution Imaging Spectroradiometer (MODIS) imageries is presented. An eight neighborhood algorithm is used to estimate the distribution variance of chl\(a\) concentration, \(D(R_{3M})\), from MODIS imageries. Based on the analysis for global data series of monthly composites of MODIS Aqua satellite data from 2008 to 2010 at 4 km resolution, the chl\(a\) concentration estimated by OC3M algorithm would be smaller than in situ measurements in theory. This is a even if the atmospheric effects has been accurately removed owing to the scale effects. Relative scale error has characterizes associating with spatial and time dependents. Largely, scale error mostly distributed in the coastal waters, which is closely related with human society. In this study, \(D(R_{3M})\) is estimated very coarsely, which would lead to the underestimation of scale errors of global chl\(a\) data. As a result, how to accurately estimate \(D(R_{3M})\) may be possible areas for scale effect study in future.

[Keywords: MODIS, OC3M algorithm, global oceans, Chlorophyll-a concentration, Scale effect]

Introduction

The use of satellite imagery has been proven beneficial for deriving biophysical variables in optically-deep waters dominated by phytoplankton at temporal and spatial scales, which is difficult to be attained with direct field measurements. Chl\(a\) concentration has typically been the main biophysical variable derived from ocean-colour imagery. Modern ocean color sensor, such as MODIS on the Aqua spacecraft is widely assumed to produce global representations of chl\(a\) concentration\(^2\). MODIS is typically 2-3 times more sensitive than Sea-viewing Wide Field-of-view Sensor (Sea WiFS), which is turn approximately twice as sensitive as Coastal Zone Color Scanner (CZCS)\(^3\), and shorter revisit time than Landsat Thematic Mapper (TM), which makes the MODIS be an ideal tool to assess chl\(a\) concentration\(^4\). However, the MODIS sensor only captures data at a spatial resolution of ~1 km, \(i.e.,\) simply using one value to represent the chl\(a\) concentration of 1 km\(^2\) water area for each pixel. In non-homogeneously distributing instance, that chl\(a\) concentration is equal to the average concentration of ~1 km\(^2\) water area of a pixel. However, that instance is unreality in practical application of remote sensing owing to scale effects\(^5\), even when the atmospheric effects and other environmental influence are accurately removed from MODIS imagery.

In general, the chl\(a\) concentration estimation algorithm is usually constructed and regressed from fieldworks within 0.1 m\(^2\) water areas (observed with 5 solid angle and 0.5 m above the water surface)\(^6\). advised that it may produce scale problem when used the laws, principles and algorithms deduced from non-homogeneously situations to the homogeneously remote sensing pixels. For example, the water quality concentration retrieved from MODIS imagery may be not equal to the realistically average concentration of pixel in theory\(^7,8\) revealed that it would underestimate the suspended sediment concentration from remote sensing imagery owing to the scale effects. Furthermore, the scale effects may make some physical phenomenon become more complicated, \(e.g.,\) the systematical underestimation of MODIS’s global chl\(a\) concentration products is discussed in this paper.

Present study is an attempt to discuss the scale effects of ocean chlorophyll 4 algorithm of MODIS (OC3M), which is the standard MODIS global chl\(a\) concentration estimation algorithm, and then proves that the scale problem can cause the systematical
underestimation of OC3M algorithm. Finally, the study would use the eight neighborhood algorithm (EN) advised by to coarsely estimate the distributing variance of chla concentration in a MODIS pixel, and then uses this variance to calculate the scale error of OC3M algorithm for each MODIS imagery pixel.

Materials and Methods

Data used

A global series of monthly composites of MODIS Aqua satellite data from 2008 to 2010 at 4 km resolution were obtained from the NASA ocean color data archive (oceancolor.gsfc.nasa.nov). Products required for this study includes the reflectance fields at 443, 489 and 551 nm, which are the basically inputting parameters of OC3M algorithm. The scale effects associated with chla distribution can be derived as long as the method is used operationally.

Scale effects of OC3M algorithm

The empirical OC3M algorithm is extended from the OC4 and OC2 algorithms developed for the SeaWiFS sensor but adapted to the spectral bands of MODIS. Algorithm is statistically derived based on chla concentration images produced by NASA’s Ocean Color Group using MODIS imagery are based on the OC3M algorithm, defined as:

\[
[chla]_{p} = f(R_{3M}) = 10^{(0.283 - 2.753R_{3M} + 1.457R_{3M}^{2} + 0.659R_{3M}^{3} - 1.403R_{3M}^{4})} \quad \ldots (1a)
\]

\[
R_{3M} = \log_{10}\frac{\text{max}[R_{s}(443 \text{ nm}), R_{s}(488 \text{ nm})]}{R_{s}(551 \text{ nm})} \quad \ldots (1b)
\]

\[
R_{s}(\lambda) = \frac{L_{w}(\lambda)}{E_{s}(\lambda)} \quad \ldots (1c)
\]

Where, max() specifies the greater of the two values. Above-water remote sensing reflectance, \(R_s(\lambda)\), is defined in Eq. (1c) with \(L_w(\lambda)\) representing water-leaving radiance and \(E_s(\lambda)\) representing above-water incident irradiance \([Chla]_p\).

\(p\) is the chla concentration estimated at pixel’s scale. The first-order and second-order derivate of OC3M algorithm are:

\[
f'(R_{3M}) = 2.3026 f(R_{3M}) \quad \ldots (2)
\]

\[
-2.753 + 2.914R_{3M} + 1.977R_{3M}^{2} - 5.612R_{3M}^{3} \quad \ldots (3a)
\]

\[
g(R_{3M}) = 4.817 - 13.13R_{3M} - 0.418R_{3M}^{2} + 36.61R_{3M}^{3} \quad \ldots (3b)
\]

\[-28.797R_{3M}^{2} - 22.190R_{3M}^{3} + 31.495R_{3M}^{4} \quad \ldots (3c)
\]

If considers the homogeneous distribution of \(R_{3M}\) in MODIS pixel, the chla of MODIS pixel observed at final scale, \([chla]_f\), can be denoted:

\[
[chla]_f = \frac{\int f(R_{3M})dA}{A} \quad \ldots (4)
\]

In Eq. (4), we assume that the MODIS pixel can be divided into several non-homogeneous sub-pixels, whose area is \(dA\). In each sub-pixel, the scale is final enough so that the \(R_{3M}\) non-homogeneously distributes in sub-pixel. Extending Eq. (4) by Tayler series and discarding the items with more than second-order derivate, yielding Eq. (5) by Taylor series and discarding the items with more than second-order derivate, yielding Eq. (5).

\[
[chla]_f = [chla]_p + f(R_{0})E(R_{3M} - R_{0}) + 0.5f''(R_{0})D(R_{3M}) \quad \ldots (5a)
\]

\[
R_{0} = \frac{1}{A} \int R_{3M}dA \quad \ldots (5b)
\]

\[
E(R_{3M} - R_{0}) = \frac{1}{A} \int (R_{3M} - R_{0})dA = 0 \quad \ldots (5c)
\]

Where, \(R_{0}\) represents the average \(R_{3M}\) of a MODIS pixel, and \(D(R_{3M})\) is the corresponding variance. And then, the formula for estimating scale variance is shown:

\[
\epsilon_r = \frac{[chla]_f - [chla]_p}{[chla]_p} = 0.5f''(R_{0})D(R_{3}) \quad \ldots (6)
\]

And the corresponding relative error is:

\[
\epsilon_r = \frac{[chla]_f - [chla]_p}{[chla]_p} = \frac{f''(R_{0})D(R_{3})}{2[chla]_p} \quad \ldots (7)
\]

EN algorithm

Actually, the homogeneous pixel can be divided into several non-homogeneous sub-pixels as long as the scale of sub-pixel is final enough. At this scale, the chla concentration non-homogeneously distributes in each sub-pixel. However, the desirable sub-pixels
are not presently possible because the requisite comprehensive oceanic optical data sets are not available. Such data sets must contain simultaneous measurements of the continuum changes of \( R_{3M} \) and chl \( a \) concentration of the water column in each MODIS pixel (Mobley et al. 1993). Consequently, it is impossible to accurately estimate or measure the \( D(R_s) \) in theory or practical. However, the \( D(R_s) \) is the key parameter to assess the scale effects of OC3M algorithm. In order to better understand the algorithms’ scale effects, the EN algorithm advised by\(^5\) is utilized to approximately estimate the \( D(R_s) \) in MODIS’s pixel. The EN algorithm mainly uses the \( D(R_s) \) in 3×3 window to coarsely instead for the \( D(R_{3M}) \) in MODIS’s pixel.

\[
E(R_s) = \frac{1}{9} \sum_{i,j} R_{3M,i,j} \quad \ldots \ (8a)
\]

\[
D(R_{3M}) = D(R_s) = \frac{1}{9} \sum_{i,j} \left[ R_{3M,i,j} - E(R_s) \right]^2 \quad \ldots \ (8b)
\]

Where, \( R_{3M,i,j} \) is the \( R_{3M} \) value of MODIS’s pixel associated with 3×3 window. Space dose not permit a full reveal of EN algorithm here, instead, the reader is referred to our previous study\(^5,9\). Obviously, the approximately substitution of Eq. (8) is greatly coarse, but it is very useful for us to comprehend the scale effects of OC3M algorithm on chl \( a \) concentration estimation, because there is not available tool or method that can be used to accurately estimate \( D(R_{3M}) \) from MODIS imagery. It may be underestimation of \( D(R_{3M}) \) using Eq. (8), because the \( R_{3M} \) is approximately equal to the area-weighted of all non-homogeneous sub-pixels in MODIS pixel.

Results and Discussions

The distributing probability of second-order derives of OC3M algorithm

The series of seasonal composites of MODIS Aqua satellite data collected from 2005 to 2010, including 108 imageries associating with the global ocean, are used to statistic the distributing probability of \( R_{3M} \) in global ocean. Although the global chl \( a \) products derived from NASA’s ocean color satellite programs have a nominal uncertainty of \( \pm 35\% \) when uses the OC3M algorithm to estimate the global chl \( a \) concentration from the Level-3 MODIS oceanic reflectance products. However, in order to prove the scale effects on NASA’s chl \( a \) concentration product, we would not carry out any data processing operations on the NASA’s Level-3 MODIS oceanic reflectance products to improve the influences of atmosphere.

The dataset about the \( R_{3M} \) is constructed according to these 108 imageries, and then the histogram is calculated based on this dataset. Fig. 1 shows the distributing probability of \( R_{3M} \) in global ocean. According to Fig. 1, it is found that the \( R_{3M} \) in global ocean varies from -0.3 to 1.1, and the probability associating with \( R_{3M}>0 \) is much larger than probability associating with \( R_{3M}<0 \). Distributing probability of second-order derives of OC3M algorithm in oceanic waters can be estimated by Eq. (3) as long as known the distributing probability of \( R_{3M} \). Fig. 2 shows the distributing probability of second-order derives of OC3M algorithm when the \( R_{3M} \) distributes as shown in Fig. 1. According to Fig. 2, it is found that second-order derives of OC3M algorithm varies from 3.8 to 45.53 and the corresponding expected value is 9.96.

![Fig.1— Distributing probability of \( R_{3M} \) in global ocean](image1)

![Fig.2— Distributing probability of second-order derives of OC3M algorithm in global ocean](image2)
comprehensive oceanic optical data sets are not available. In this study, we use the NE algorithm as shown in Eq. (8) to estimate the $D(R_3)$ from MODIS imageries. Although this approach is quite coarse, it is help for understanding the scale effects on the MODIS standard global chl concentration algorithm when the more accurately method is unavailable.

Fig. 3 shows the distributing probability of $D(R_3)$ in global ocean computed from MODIS imagery collected from 2008 to 2010. Fig. 3 shows that the $D(R_3)$ varies from 0 to 0.035, and the corresponding expected value is 0.0013.

The chla concentration estimated by OC3M algorithm is usually larger than the minimum concentration and smaller than the maximum concentration of a remote sensing pixel. As a result, the actual chla concentration may be more non-homogeneous than the distribution situations in 3x3 window, which would lead to the underestimation of $D(R_{SM})$, and then result in the underestimation of scale error of OC3M algorithm in global ocean. As a result, the expected value of $D(R_3)$ in global ocean may be larger than 0.0013.

**Systematical underestimation of OC3M algorithm**

According to Eq. (6), the chla concentration estimated by OC3M algorithm from MODIS imageries includes two items: chla concentration estimated at final scale and scale error. Relative scale error as shown in Eq. (7) is depended on two parameters: second-order derives of OC3M algorithm and the $D(R_3)$. The $D(R_3)$ is statistically non-negativity. Hence, whether the overestimation or underestimation of chla concentration is depended on the values of second-order derives of OC3M algorithm. As discussed in previous, the second-order derives of OC3M algorithm changes from 3.8 to 45.53, which is non-negativity, and then the relative scale error in Eq. (7) is also non-negativity. Consequently, the chla concentration from MODIS imageries is smaller than the chla concentration estimated at final scale, such as in situ measurements, so we can get conclusion that the OC3M algorithm would underestimate the chla concentration from MODIS imageries owing to the scale effects in theory. In this study, we find that the average uncertainties of OC3M algorithm associating with scale effect is ~1.29% (expected values of $D(R_{SM})$ and second-order derives of OC3M algorithm are 0.0013 and 9.96, respectively) when underestimation of $D(R_{SM})$ is not taken into account. Accordingly, it can be known that there is at least 1.29% systematical underestimation of OC3M algorithm in estimating chla concentration from the global ocean.

**Timely and spatial changes of scale effects of OC3M algorithm**

The scale error shown in Eq. (7) can be estimated while the $D(R_3)$ and the second-order derives of OC3M algorithm are known. Fig. 4 shows the scale error of OC3M algorithm in global ocean computed from the global data series of monthly composites of MODIS Aqua satellite data from 2008 to 2010 at 4 km resolution. According to Fig. 4, it is found that uncertainty associated with scale effects of OC3M algorithm varies from 0% to 40%. There are ~90% global oceans whose relative scale error is smaller than 1%, and ~2% global oceans whose relative scale error is larger than 5%. However, there are still 0.5% regions, whose scale error is larger than 10%.

The scale effects of OC3M algorithm is obviously timely dependent. From January to December, the relative scale error increases, and then decreases. Relative scale error in summer is largest and is smallest in winter. Especially in northern hemisphere, these time dependent characterizes is much more obvious. Additionally, the scale error has characterizes associating with spatial dependent. Scale error in northern hemisphere is larger than that in southern hemisphere, and the scale error is relative larger in the coastal waters around the Asia than that in other oceanic waters. Largely scale error mostly distributed in the coastal waters. By comparison with the water types classified by associated with global oceans, the scale error is relative smaller in the Case I waters (<1%), and relative larger in the Case II waters (>5%), owing to the more non-homogeneous distribution of chla concentration in Case II waters.

![Fig.3— Distributing probability of $D(R_3)$ in global ocean](image-url)
Conclusions

Scale error is very important in oceanic remote sensing e.g., the scale effect makes OC3M algorithm ~1.29% underestimate chl$\alpha$ concentration in global ocean from MODIS sensors; the validation dataset should distribute in the pixel owing to the scale effects; as the development of oceanic color remote sensing, it has to consider the scale effects, if we want to obtain the chl$\alpha$ concentration with low estimation uncertainties. Actually, theoretical study is usually earlier than application study. Therefore, scale effects study is very meaningful for oceanic color remote sensing in presents and future.

A framework for characterizing the scale effects of NASA’s global chl$\alpha$ products estimated by OC3M algorithm to. Study results show that there are ~90% global oceans whose relative scale error is smaller than 1%, and ~2% global oceans whose relative scale error is larger than 5%. There are still 0.5% regions, whose scale error is larger than 10% scale error in northern hemisphere is larger than that in southern hemisphere, and the scale error is relative larger in the coastal waters around the Asia than that in other oceanic waters. By comparison, the scale error in case II waters is larger than it in case II waters.
The EN algorithm is greatly coarse in estimating $D(R_{3M})$, but it is very useful for us to comprehend the scale effects of OC3M algorithm on chl $\alpha$ concentration estimation, because there is not available tool or method that can be used to accurately estimate $D(R_{3M})$ from MODIS imagery. However, the $D(R_{3M})$ may be underestimated using EN algorithm, which leads to underestimation of scale errors of OC3M algorithm. In order to clearly understand the scale effects on MODIS global chl $\alpha$ concentration products, it is very necessary for developing an effective $D(R_{3M})$ estimation algorithm in future.

Acknowledgement

This study is supported by the open fund of Key Laboratory of Mar Hydrocarbon Resources and Environmental Geology (MRE201109).

References