

# **Development of Algorithms and Strategies for Monitoring Chlorophyll and Primary Productivity in Coastal Ocean, Estuarine and Inland Water Ecosystems**

Semi-Annual Technical Report  
July 15, 1998

Janet W. Campbell  
University of New Hampshire  
Durham, New Hampshire 03824

## **Summary**

This is the semi-annual progress report for the period January through June 1998 for the Execution Phase of my MODIS Instrument Team investigator project. The objectives of this work are:

Establish a protocol for developing regional or site-specific bio-optical algorithms for coastal "case 2" waters.

Demonstrate the protocol by developing algorithms for two coastal seas: the Gulf of Maine/Mid-Atlantic region and the Yellow Sea/East China Sea region.

Prescribe a protocol for "stitching together" local or site-specific algorithms.

Develop a strategy for monitoring coastal oceans, estuaries, and inland waters.

In this report, I will describe progress toward the first three objectives. This report reflects the efforts of a research team consisting of myself, an assistant research scientist (Timothy Moore), a research associate (Karen Garrison), and three graduate student assistants (Hui Feng, S. Gaudreau and K. Jacobs). During the time period covered by this report, I have been working at NASA Headquarters as program manager for Ocean Biology and Biogeochemistry. Therefore, my direct involvement in this research has been limited, although I maintain contact with and continue to supervise the work of the team at UNH.

## **Chlorophyll Algorithm Protocol Development**

Broadly stated, the goal of a bio-optical algorithm is to estimate a suite of water constituents, including chlorophyll, which affect the optical properties of the water. Other constituents to be estimated include colored dissolved organic matter (CDOM), organic detritus, and suspended sediments. Our goal is to define a protocol for parameterizing regional bio-optical algorithms and a technique for 'stitching together' retrievals from different regional algorithms. Chlorophyll derived by the bio-optical algorithm will be used to estimate primary productivity. Progress toward the goal of parameterizing primary productivity algorithms will be described later.

### **Task #1: Define a generalized mathematical framework for bio-optical algorithms**

The framework for bio-optical algorithms will be a radiance model that predicts upwelling spectral radiance as a function of the inherent optical properties of the water (Table 1).

Table 1. Equations used to derive  $L_{wN}$  from  $b_b$  and  $a$  (left column) and the inversion equations used to derive  $b_b/a$  from  $L_{wN}$  (right column). Eqs. I-1 to I-3 predict the normalized water-leaving radiance given  $a$  and  $b_b$ , whereas eqns. I-4 to I-6 are the basis for analytical algorithms used to derive in-water optical properties given water-leaving spectral radiance measurements.

Given the inherent optical properties,  $a$  and  $b_b$ , we define  $X$  as follows:

$$X \equiv \frac{b_b}{a+b_b} \quad (\text{I-1})$$

where  $a$  and  $b_b$  are the effective absorption and back-scatter coefficients within the upper optical depth.

Based on results of Gordon (1986), the remote-sensing reflectance is accurately represented as:

$$R_{rs} = 0.0949 X + 0.0794 X^2 \quad (\text{I-2})$$

for solar zenith angles  $\theta_0 > 20^\circ$ .

According to the "Semianalytic Radiance Model" of Gordon et al., (1988), the normalized water-leaving radiance can be modeled as:

$$L_{wN} = \frac{(1-\rho)(1-\rho')F_0R_{rs}}{m^2(1-rQR_{rs})} \quad (\text{I-3})$$

where the symbols are defined previously (see text). In this expression, the term  $(1-rR)$  which appears in the paper by Gordon et al., (1988), (their equation 1), is replaced by the term  $(1-rQR_{rs})$  where  $Q$  is an estimate of  $R/R_{rs}$ .  $Q$  does not need to be particularly accurate since  $(1-rR)$  only varies from about 0.92 to 1.0, and is sometimes ignored.

Given normalized water-leaving radiance,  $L_{wN}$ , equation (I-3) is inverted to obtain the remote-sensing reflectance:

$$R_{rs} = \frac{L_{wN}/F_0}{M + rQ*L_{wN}/F_0} \quad (\text{I-4})$$

where  $M = (1-\rho)(1-\rho')/m^2$ . Note that both  $L_{wN}$  and  $F_0$  depend on wavelength, whereas the other terms in (I-4) do not.

Equation (I-2) is a quadratic equation with two roots. The only positive root is:

$$X = \frac{-0.0949 + \sqrt{0.0090 + 0.3176 R_{rs}}}{0.1588} \quad (\text{I-5})$$

Since  $a \gg b_b$  in most Case 1 waters, equation (I-1) is often approximated by  $X \approx b_b/a$ . However, this approximation is unnecessary, since the ratio of  $b_b$  to  $a$  is easily computed as:

$$\frac{b_b}{a} = \frac{X}{1-X} \quad (\text{I-6})$$

Thus, beginning with normalized water-leaving radiances in the visible-range ocean bands of MODIS, the ratio of back-scattering,  $b_b(\lambda_i)$ , to absorption  $a(\lambda_i)$  is derived for each band. Models of inherent optical properties are then parameterized to express the dependence of  $b_b$  and  $a$  on in-water constituents (e.g., CHL, TSS, CDOM, etc.).

The inherent optical properties, in turn, are parameterized as functions of in-water constituents. The bio-optical algorithm is the inverse of the radiance model in that it predicts in-water constituent concentrations given upwelling spectral radiance.

Table 1 summarizes the equations for deriving the normalized water-leaving radiance as a function of inherent optical properties (forward direction of the radiance model) and the equations for inverting the model to derive the ratio  $X = b_b/a$ . These equations are based on the semi-analytic radiance model of Gordon et al. (1988) which is the basis for our work and that of others (Carder et al., 1997; Garver and Siegel, 1997; Hoge and Lyon; 1996).

This approach is flexible enough to be used in a variety of environments including the open ocean (case 1 waters) as well as coastal, estuarine and inland regions (case 2 waters). Flexibility is derived by allowing for locally parameterized inherent optical property (IOP) sub-models. Specifically, we require parameterizations of the effective upper-water absorption and backscattering coefficients:

$$a(\lambda) = a_w(\lambda) + a_{ph}(\lambda) + a_{cdm}(\lambda) \quad (1)$$

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda) \quad (2)$$

where  $a_w(\lambda)$  and  $b_{bw}(\lambda)$  are known properties of pure water (or seawater);  $a_{ph}(\lambda)$  is the absorption coefficient of phytoplankton;  $a_{cdm}(\lambda)$  is the absorption coefficient of colored dissolved and particulate organic matter; and  $b_{bp}(\lambda)$  is the backscattering coefficient of particulate matter in the upper water column. This is illustrated in Table 2.

<b>Inherent Optical Property Submodel</b>	<b>Retrieval Variables</b>
<p><i>Backscattering coefficient of particles:</i></p> $b_{bp}(\lambda) = b_{bp}(550) \left( \frac{550}{\lambda} \right)^Y$	<p><b><math>b_{bp}(550)</math></b> -- related to TSM  <b><math>Y</math></b> -- size distribution of particles</p>
<p><i>Absorption coefficient of phytoplankton:</i></p> $a_{ph}(\lambda) = a_{ph}^*(\lambda) \text{ CHL}$	<p><b>CHL</b> -- chlorophyll concentration</p>
<p><i>Absorption coefficient of CDOM and detritus:</i></p> $a_{cdm}(\lambda) = a_{cdm}(400) e^{-S(\lambda-400)}$	<p><b><math>a_{cdm}(400)</math></b> -- CDOM and detritus concentrations</p>

The absorption and backscattering coefficients will require locally parameterized submodels to account for the effects of materials found in each body of water. Specifically, regional algorithms will differ in how they parameterize the absorption and backscattering coefficients associated with phytoplankton cells, detritus, CDOM, sediments and other mineral particles. Algorithms based on this framework will differ in two ways. First, the submodels associated with each constituent (e.g., phytoplankton, detritus, CDOM, sediment, etc.) may differ in their mathematical form, and second, parameters may differ among algorithms using the same mathematical form.

Task #2: Prescribe a protocol for parameterizing algorithms

Once the mathematical relationships have been established, the next step is the estimation of parameters. Each radiance model has an associated set of variables and parameters. The distinction between parameters and variables is often blurred because parameters tend to vary spatially and temporally. However, the distinction is clear in the context of an algorithm. Any terms associated with the radiance model (factors, exponents, slopes, intercepts, etc.) that are estimated using data and subsequently held constant when the algorithm is applied are considered *parameters*. Terms in the model that are either input to or output from the algorithm on a pixel by pixel basis are considered *variables*.

For example, the absorption coefficient for CDOM and detritus is usually modeled as

$$a_{\text{cdm}}(\lambda) = a_{\text{cdm}}(\lambda_0) \exp[-S(\lambda - \lambda_0)] \quad (3)$$

In this expression,  $a_{\text{cdm}}(\lambda_0)$  is a variable that is used as an index of the CDOM plus detritus concentrations. Most algorithms seek to estimate  $a_{\text{cdm}}(\lambda_0)$  and assume a constant value for  $S$ , in which case,  $S$  is a parameter. On the other hand, if  $S$  is estimated as part of the algorithm then  $S$  is also a variable which would vary from pixel to pixel.

The estimation of parameters requires concurrent data on constituent concentrations, inherent optical properties, and remote-sensing reflectance at  $n$  stations. The latter is derived from measurements of upwelling radiance and downwelling irradiance. Given measured remote-sensing reflectances ( $R_{\text{rs}}$ ) and constituent concentrations ( $C$ ), the unknowns are the  $m$  elements of the parameter vector,  $\theta$ .

Our objective is to define a statistical procedure for determining the parameter values,  $\theta$ , which minimize errors in  $C$ . Errors may be defined either as simple differences between the retrieved and measured values ( $\Delta$ ) or as differences in log-transformed values ( $\Delta'$ ). For either definition, the goal is to minimize the mean squared error:

$$\text{MSE} = \frac{1}{n} \sum_{i=0}^n \Delta_i^2 \quad (4)$$

defined here in terms of simple difference errors.

The ordinary least-squares procedure is to solve a system of  $m$  equations:

$$\frac{\delta \text{MSE}}{\delta \theta_j} = \frac{2}{n} \sum_{i=0}^n \Delta_i \frac{\delta \Delta_i}{\delta \theta_j} = 0 \quad (j = 1, \dots, m) \quad (5)$$

which can be written:

$$\sum_{i=0}^n C_i \frac{\delta \hat{C}_i}{\delta \theta_j} = \sum_{i=0}^n \hat{C}_i \frac{\delta \hat{C}_i}{\delta \theta_j} \quad (j = 1, \dots, m) \quad (6)$$

Because of the nonlinear nature of the radiance model, we do not have an explicit expression for  $\hat{C}_i$  as a function of the input variables and parameters. However, numerical solutions may be obtained for the above system of equations.

Hui Feng (as part of his Ph.D. thesis) has investigated optimization procedures for parameterizing a simplified reflectance model:

$$R = G \frac{b_{bw} + b_{bp}}{a + b_{bw} + b_{bp}} \quad (7)$$

where  $G$  is a constant. Using the spectral absorption measurements and reflectance data from Tokyo Bay (Kishino et al., 1985), he solved for a "derived"  $b_{bp}$ :

$$b_{bp} = \frac{\frac{R}{G} (a + b_{bw}) - b_{bw}}{(1 - \frac{R}{G})} \quad (8)$$

Hui then parameterized  $b_{bp}$  as a function of total suspended solids [TSS]:

$$b_{bp} = A [\text{TSS}]^B \quad (9)$$

where the parameters  $A$  and  $B$  are spectrally varying. The measured absorption coefficients were parameterized in terms of  $\text{CHL}$  and  $a_{\text{cdm}}(400)$ . Parameters were estimated by two methods: One method minimized the MSE in the IOP submodels, and the other method minimized the MSE in remote-sensing reflectance (the latter required a nonlinear optimization technique). The resulting parameters were similar.

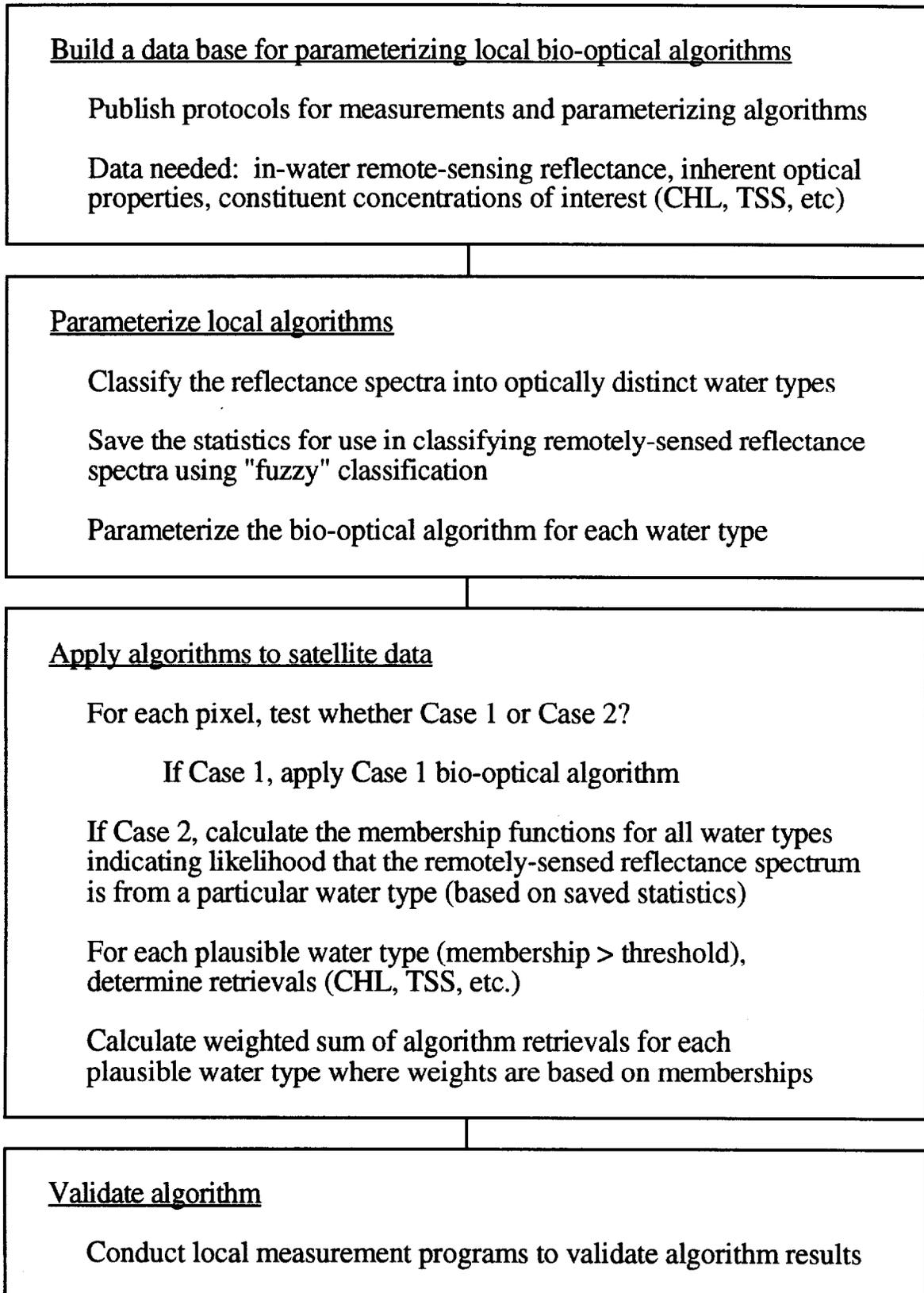
We found that a single reflectance model (based on parameterizations of  $b_{bp}$ ,  $a_{\text{ph}}$  and  $a_{\text{cdm}}$ ) was not adequate to represent the range of variability found in the Tokyo Bay data set. Using an unsupervised classification procedure, the measured reflectance spectra were classified into three distinct water types. Then algorithms were parameterized for each water type. Results were greatly improved. The Tokyo Bay data set is considered to be representative of coastal waters where complex mixtures of different waters types exist. A publication based on these results is in preparation (Feng et al. "Modeling the Spectral Reflectance of Optically Complex Waters: A Demonstration for Tokyo Bay.") These results were presented at the Ocean Sciences Meeting in San Diego in February, 1998.

An error analysis to determine the sensitivity of chlorophyll retrievals to errors in the IOP parameterization was presented at the 1998 Pacific Ocean Remote Sensing Conference (PORSEC '98). See Appendix A. This paper is now being prepared for publication in a refereed journal.

### Task #3: Prescribe a protocol for 'stitching together' algorithm retrievals

In coastal regions, large lakes or estuaries, there is often a mixture of waters of distinct optical properties. Well-defined color fronts mark the boundary between water types. Each

Fig. 1 - Proposed Strategy for "Case 2" Bio-Optical Algorithm Development



water type may require its own algorithm, and thus there is a need to "stitch together" algorithms in regions where two or more water types mix. We are investigating the use of a 'fuzzy logic' approach to this task.

The overall strategy (Figure 1) begins with the development of a data base for parameterizing coastal ocean algorithms. The data needed include: in-water remote-sensing reflectance, inherent optical properties, and constituent concentrations of interest. We have begun to collect data for the Gulf of Maine and Mid-Atlantic region. Timothy Moore and Hui Feng have been working to develop and demonstrate the procedures outlined in Fig. 1. An unsupervised classification procedure was applied to remote-sensing reflectance data from the Gulf of Maine and Mid-Atlantic region, and 5 distinct water classes were identified.

After classifying the in-situ reflectance data into 5 classes, we derived the mean  $\mathbf{M}_i$  and covariance matrix  $\Sigma_i$  for remote sensing reflectance within each water class  $i$ ,  $i = 1, \dots, 5$ . These statistics are used to compute the class membership functions for any measured reflectance spectrum,  $\mathbf{R}$ , as follows.

Let  $\mathbf{R}$  be a measured remote-sensing reflectance vector (dropping the 'rs' subscript). Then

$$Z^2 = (\mathbf{R} - \mathbf{M}_i)^t \Sigma_i^{-1} (\mathbf{R} - \mathbf{M}_i) \quad (10)$$

is the squared Mahalanobis distance between the measured  $\mathbf{R}$  and the  $i$ th class mean  $\mathbf{M}_i$ . If  $\mathbf{R}$  belongs to class  $i$ , then  $Z^2$  has a  $\chi^2$  distribution with  $n$  degrees of freedom (where  $n$  is the dimension of  $\mathbf{R}$ ). As a measure of the likelihood that  $\mathbf{R}$  is drawn from class  $i$ , we compute the probability:

$$P_i = 1 - F_n(Z^2) \quad (11)$$

where  $F_n(Z^2)$  is the cumulative  $\chi^2$  distribution with  $n$  degrees of freedom. For even values of  $n$ ,  $P_i$  has the exact solution:

$$P_i = \exp(-Z^2) \sum_{i=0}^{m} \frac{(Z^2/2)^i}{i!} \quad (12)$$

where  $m = n/2 - 1$ . Images of  $P_i$  ( $i = 1, \dots, 5$ ) for the 5 water classes from SeaWiFS data for Gulf of Maine and Mid-Atlantic region were produced.

We will set a threshold for  $P_i$  such that any class with a probability above this threshold is considered "plausible." Then, for each pixel, the retrieved variables will be a weighted sum of plausible retrievals, where weights are based on the  $P_i$  values.

In the future, we believe that a large globally distributed data base can be classified using similar techniques to identify water types. The algorithms parameterized for each water type may be used -- regardless of where that type was located in the original in-situ data base. I am collaborating with Dr. Sinjae Yoo, of the Korean Ocean Research and Development Institute (KORDI) to promote the development of this globally distributed data base. We have proposed that this become a pilot project of the Coastal Global Ocean Observing System (C-GOOS). (Dr. Yoo is a member of the C-GOOS panel. See Appendix B).

My graduate student (Ken Jacobs) is developing an algorithm for the Great Bay Estuary in New Hampshire. Ken has obtained the relevant data (optical, biological and chemical) to characterize optical properties of waters mixing in this estuary. We will thus have the data to test 'fuzzy logic' concepts in these local waters. In addition, we are collaborating with Dr. Yoo to test 'fuzzy logic' concepts on algorithms in the Yellow Sea and East China Sea. Dr. Yoo will provide in-water biological, chemical and optical (MER and PRR) data.

#### Publications/Presentations:

Moore, T. S., H. Feng, and J. W. Campbell. Fuzzy classification of remote sensing reflectance data for merging bio-optical algorithms in coastal regions. A poster on this was presented at the Ocean Sciences Meeting in San Diego in February 1998, and a poster demonstrating the technique for the Gulf of Maine/Mid-Atlantic region will be presented at the Ocean Optics meeting in Nov. 1998.

### **Primary Productivity Algorithm Protocol Development**

MODIS will have two types of primary productivity algorithms (Esaias, 1996, ATBD). There will be an algorithm that predicts *annual* primary productivity (units:  $\text{g C m}^{-2} \text{y}^{-1}$ ) from the average annual surface chlorophyll concentration. This algorithm is based on an empirically-determined linear relationship and will be applied to a running mean chlorophyll updated 'weekly' (every 8 days). The algorithm will only be applied in areas determined to be "high variance" regions where there are seasonal blooms. These regions account for most (~70%) of the export production (i.e., phytoplankton carbon exported from the euphotic zone to the deep ocean). In low-variance regions, surface chlorophyll (gross production) tends to be tightly coupled with grazers (secondary production) and thus the surface chlorophyll concentration is relatively constant and does not reflect variation in primary production. The annual variance in chlorophyll will be determined from the satellite-derived chlorophyll and used to decide whether to apply the algorithm at a particular location.

A second type of algorithm predicts *daily* primary productivity (units:  $\text{g C m}^{-2} \text{d}^{-1}$ ). Although there have been many candidate algorithms proposed and described (see Behrenfeld and Falkowski, 1997), there are two that will be implemented with MODIS. One computes daily integral productivity in the upper mixed layer, and the other computes productivity in the euphotic zone. Both of these algorithms will be applied to level-3 'weekly' chlorophyll data although they could be applied on a daily basis to level-2 chlorophyll data. The weekly values will be summed to estimate annual primary productivity, and results from the two short-term algorithms will be compared with that of the annual primary productivity algorithm.

#### Task #1: Establish a protocol for evaluating primary productivity algorithms

A protocol for evaluating primary productivity algorithms was established and applied to algorithms in a round-robin experiment (see task #2). Algorithm results were compared with  $^{14}\text{C}$ -based estimates at stations distributed over a range of ocean environments. This protocol was based on errors defined in two ways: the difference between the algorithm and measured ( $^{14}\text{C}$ ) estimate ( $\Delta$ ), and the difference in log-transformed estimates ( $\Delta'$ ). The latter is a measure of relative error. For each definition of error, two measures of performance were determined: the mean error or bias, and the root-mean-square error (RMSE).

Neither of the two measures of error ( $\Delta$  and  $\Delta'$ ) is entirely satisfactory. The linear measure,  $\Delta$ , is overly sensitive to algorithm performance in high-productivity regions and insensitive in low-productivity regions. Daily primary productivity ranges over three orders of magnitude globally, and the global distribution is probably lognormal or a mixture of lognormal distributions (i.e., one having several modes). Low-productivity regions are predominant both on a areal basis and on the basis of global carbon fluxes. Although the data set used for evaluating algorithms included stations from low and high productivity areas, its distribution did not reflect the distribution found globally. The linear error  $\Delta'$  was disproportionately affected by the high-productivity stations.

The  $\Delta'$  errors tended to be symmetrically distributed about their mean ( $M'$ ) and approximately normally distributed. Assuming an underlying normal distribution, 68% of the  $\Delta'$  values would lie within one standard deviation of the mean. Thus, we used the mean and standard deviation of  $\Delta'$  to define a "one-sigma" range for  $\delta$ :

$$\delta_{\min} = (10^{M'-S'} - 1) \tag{13}$$

$$\delta_{\max} = (10^{M'+S'} - 1)$$

where

$$\delta = (10^{\Delta'} - 1) \tag{14}$$

Most algorithms tested in the round robin were within a factor  $F = 2.3$  of the measured productivity, where  $F$  was defined by:

$$F = \max\left[\frac{1}{1+\delta_{\min}}, 1+\delta_{\max}\right] \tag{15}$$

**Task #2: Conduct an evaluation of candidate algorithms**

The second Primary Productivity Algorithm Round Robin (PPARR-2) was completed in the summer of 1997. A manuscript describing the results was drafted and distributed to 23 co-authors in September 1997 (just before my taking the position at NASA Headquarters). Co-authors returned comments and corrections and a second draft was distributed in January 1998. However, since that time there has been no progress on completing this task. The graduate student who was working on this (S. Gaudreau) left UNH after the fall semester 1997. I have not had time to devote to this work since January.

## UNCERTAINTY ANALYSIS FOR RETRIEVAL OF CHLOROPHYLL CONCENTRATION FROM OCEAN COLOR: A SIMULATION STUDY

Hui Feng, Janet W. Campbell, and Timothy S. Moore  
Ocean Process Analysis Laboratory  
Institute for the Study of Earth, Oceans and Space  
University of New Hampshire, Durham, NH 03824, USA  
( Phone: 603-862-0690; Fax: 603-862-0243; e-mail: feng@jerlov.sr.unh.edu )

### Abstract

This work presents a general approach to quantifying retrieval errors in chlorophyll concentration induced by uncertainty in the underlying model parameterization. The chlorophyll retrieval is obtained by inverting an ocean color model with nonlinear inherent optical property (IOP) submodels. Here we demonstrate and quantify how uncertainty in the IOP submodel parameterization influences the accuracy of the chlorophyll concentration retrieval at different chlorophyll concentration levels.

### Background

One of the main objectives in ocean color remote sensing is to determine in-water constituent concentrations. Constituents of interest include phytoplankton chlorophyll, colored dissolved organic matter (CDOM), and suspended sediments. Techniques for concentration retrieval have evolved from empirical towards analytical (model-based) algorithms for the last two decades. Analytical algorithms usually resort to an inversion technique applied to a parameterized ocean color model. Currently, several such inversion techniques have been proposed. Hoge and Lyon (1996) applied a semi-analytical radiance model (Gordon et al., 1988) with IOP submodels to the retrieval of three in-water variables. For each constituent under consideration, its IOP was modeled as the product of the IOP at a reference wavelength multiplied by a spectral shape function. The spectral shape functions were fixed and independent of in-water constituent concentrations. Thus, the IOPs at reference wavelengths could be retrieved by a linear system inversion.

Garver and Seigel (1997) presented an inverse model to retrieve chlorophyll concentrations. In their model, the chlorophyll-specific absorption coefficient was a nonlinear function of the chlorophyll concentration, and thus a non-linear optimization technique was adopted to invert their model. Campbell et al. (1997) used a radiance model configured with nonlinear IOPs to retrieve chlorophyll concentration, gelbstoff (CDOM) absorption, and a variable associated with total particle backscattering. Their model can be inverted using a nonlinear optimization method. The model configurations and inverse techniques essentially differ in the three models mentioned above. For any inverse model, it is necessary to quantify the potential sources of uncertainty in an inverse solution. Hoge and Lyon (1996) showed that their inverse solution is very sensitive to model parameters. The goal of this work is to characterize retrieval errors resulting from IOP parameterization uncertainties for the model of Campbell et al. (1997) using normalized water-leaving radiances in the first five SeaWiFS bands (412, 443, 490, 510, and 550 nm). We focus strictly on chlorophyll retrieval errors in this paper.

### Methodology

#### General Consideration

Without losing generality, a forward ocean color model may be expressed as

$$L(\lambda) = f(C, \Theta(\lambda)) \quad (1)$$

where  $L(\lambda)$  is an ocean color measurement at wavelength  $\lambda$  (e.g. water-leaving radiance), and  $f$  is a model configuration (or function) relating  $L(\lambda)$  to the in-water concentration vector  $C$  to be retrieved through a model parameter vector  $\Theta(\lambda)$ . Distinct model configurations  $f$  will possess different parameter vectors. For example, parameters related to IOPs will depend on the constituent-specific IOP submodels used. In the case of the chlorophyll-specific absorption coefficient, for example, the models of Carder et al. (1991) and Bricaud et al. (1995) would have different parameters.

The retrieval of  $C$  based on (1) can be written:

$$C = f^{-1}(L(\lambda), \Theta(\lambda)) \quad (2)$$

where  $f^{-1}$  represents an inversion of the model  $f$ . This might be an explicit expression, as in the case of a model that is linear with respect to the in-water concentrations (e.g., Hoge and Lyon, 1996), or the symbol  $f^{-1}$  might denote an inversion technique if the inversion of (1) does not yield an explicit solution.

Generally speaking, there exist four potential candidates responsible for accuracy in retrieving the concentration vector  $C$ . The first candidate source of error is the model  $f$  itself which links an apparent optical property,  $L(\lambda)$ , to inherent optical properties. Exact solutions to radiative transfer equations are highly complex (Zaneveld, 1995), and are not amenable to inverse solutions. Almost all forward models are approximations to the more complex equations. The second source of error is the inherent variability in constituent-specific IOPs. The IOP submodels and their associated model parameter vector  $\Theta(\lambda)$  only approximate the actual constituent-specific IOPs. One of the main objectives in this work is to quantify the retrieval errors caused by the parameterization of constituent IOPs. A third source of error might be the inversion scheme itself although this can generally be controlled by setting convergence criteria. Finally, errors in the measurements  $L(\lambda)$  caused by sensor calibration errors, atmospheric correction errors, etc., can affect accuracy of the retrievals. In this paper, we consider only the errors resulting from the parameterization of constituent-specific IOP submodels. Analysis of other error sources will be the subject of future work.

### Radiance Model

Normalized water-leaving radiance is related to remote sensing reflectance  $R_{rs}(\lambda)$  by the relationship

$$L_w(\lambda) = \frac{MF_0 R_{rs}(\lambda)}{1 - rQR_{rs}(\lambda)} \quad (3)$$

where  $r$ ,  $M$ ,  $F_0$  and  $Q$  are spectral constants (Table 1);  $R_{rs}(\lambda)$  is directly related to IOPs by

$$R_{rs}(\lambda) = 0.0949 X(\lambda) + 0.0794 X(\lambda)^2 \quad (4)$$

(Gordon et al., 1975, 1988) where

$$X(\lambda) = \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \quad (5)$$

The absorption and backscattering coefficients are modeled by  $a(\lambda) = a_w(\lambda) + a_\phi(\lambda) + a_{gd}(\lambda)$  and  $b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda)$ , respectively, where subscripts  $w$ ,  $\phi$ ,  $g$ ,  $d$ , and  $p$  refer to pure water, phytoplankton, gelbstoff (CDOM), detritus, and particles, respectively.

The absorption of detritus decreases exponentially with increasing wavelength in a manner similar to that of gelbstoff (Carder et al., 1991). For simplification, we combine their absorption coefficients into a single term,  $a_{gd}(\lambda) = a_{gd}(375) \exp(-S(\lambda - 375))$ . The absorption coefficient of gelbstoff and detritus at 375 nm,  $a_{gd}(375)$ , is used as a measure of the gelbstoff and detritus concentration. The spectral shape parameter for gelbstoff and detritus absorption,  $S$ , varies between 0.011 and 0.021 with a mean of 0.0145 (Bukada et al., 1995). The phytoplankton absorption coefficient  $a_\phi(\lambda) = A_c(\lambda) \text{Chl}^{B_c(\lambda)}$  is based on the model of Bricaud et al. (1995) which gives

the chlorophyll-specific absorption coefficient,  $a_\phi^*(\lambda)$ , as a function of the chlorophyll concentration,  $\text{Chl}$ . Using over 800 globally-distributed observations of absorption spectra and chlorophyll concentration, Bricaud et al (1995) fit lines to  $\log(a_\phi^*(\lambda))$  vs.  $\log(\text{Chl})$ . The model used for our analysis involves two parameters,  $A_c(\lambda)$  and  $B_c(\lambda) = B_c^*(\lambda) + 1$ , where  $A_c(\lambda)$  and  $B_c^*(\lambda)$  are from Bricaud et al. (1995).

The particle backscattering submodel is  $b_{bp}(\lambda) = b_0 A_b(\lambda) \text{Chl}^{B_b(\lambda)}$  and is based on the parameterization scheme of Gordon et al. (1988). This expression involves two variables,  $b_0$  and  $\text{Chl}$ , being used to describe variations in particle backscattering. The variable  $b_0$  is associated with variability in total particle scattering which was found empirically to be a function of chlorophyll given by  $b(\lambda) = b_0 \text{Chl}^{0.62}$ . The particle backscattering probability  $b_b(\lambda)/b(\lambda)$  was parameterized as a power-law function of chlorophyll (Gordon et al. 1988), and its parameters,  $A_b(\lambda)$  and  $B_b^*(\lambda)$  were obtained by linear fits on log-log plots. The exponent in the particle backscattering submodel was, thus,  $B_b(\lambda) = B_b^*(\lambda) + 0.62$ . Variation in the amplitude of particle backscattering is associated with the properties of particles, such as their size distribution and composition (i.e. refraction index).

The model parameter vector,  $\Theta(\lambda) = [S, A_c(\lambda), B_c(\lambda), A_b(\lambda), B_b(\lambda)]$ , fully defines the IOP submodels for each wavelength. Constant model parameters used in the inversion algorithm are given in Table 1. The in-water concentration vector associated with this model is  $C = [a_{gd}(375), \text{Chl}, b_0]$ .

### Inversion method

Given a measured radiance,  $L_w(\lambda)$ , and assumed values for  $r$ ,  $M$ ,  $F_0$  and  $Q$  (in Table 1), equation (3) can be solved for  $R_{rs}(\lambda)$ , and equation (4) can be solved for  $X(\lambda)$ . Thus, a value of  $X(\lambda)$  is derived from the measured radiance,  $L_w(\lambda)$ . A non-linear optimization technique is then required to solve equation (5) for the constituent vector,  $C$ . The Levenberg-Marquardt algorithm was selected for this purpose. This algorithm, which is now widely used in ocean color inversions (Bukada et al., 1991; Roesler and Perry, 1995; Garver and Seigel, 1997; Feng et al., 1997), is similar to the Gauss-Newton algorithm with a modification to quicken its convergence. Its theoretical base has been described in detail by Press et al. (1992). Bukada et al. (1995) gave an excellent review of its potential applications in water color interpretation.

### Simulations

Uncertainties in three constituent-specific IOPs were simulated as follows:

$$\begin{aligned} S' &= S + \delta_S \\ \log(a_\phi'(\lambda)) &= \log(a_\phi(\lambda)) + \delta_\phi(\lambda) \\ \log(b_{bp}'(\lambda)) &= \log(b_{bp}(\lambda)) + \delta_{bp}(\lambda) \end{aligned} \quad (6)$$

where  $\delta_S$ ,  $\delta_\phi(\lambda)$ , and  $\delta_{bp}(\lambda)$  are normally distributed pseudo-random errors with zero means, and standard deviations equal to 0.0015, 0.04, and 0.04, respectively. These represent 10% RMS errors in the IOP submodels. Two complete sets of simulations were carried out: one in which errors were uncorrelated, and another in which  $\delta_\phi(\lambda_1) = \dots = \delta_\phi(\lambda_5)$ , and  $\delta_{bp}(\lambda_1) = \dots = \delta_{bp}(\lambda_5)$ , but where  $\delta_\phi(\lambda_i)$  and  $\delta_{bp}(\lambda_i)$  were uncorrelated. We refer to these as the "Independent Error" and "Equal Error" simulations, respectively. In future work, we will attempt to estimate the error covariance matrix.

To determine whether retrieval accuracy is dependent on the chlorophyll level, we conducted simulations for three levels representing low ( $\text{Chl} = 0.1 \text{ mg m}^{-3}$ ), medium ( $\text{Chl} = 1.0 \text{ mg m}^{-3}$ ) and high ( $\text{Chl} = 10 \text{ mg m}^{-3}$ ) chlorophyll concentrations. For each chlorophyll level, we simulated a random sample of  $n = 200$   $a_{gd}(375)$  and  $b_0$  values. The distribution of  $a_{gd}(375)$  depended on Chl as follows:  $\log a_{gd}(375) = 0.47909 \log(\text{Chl}) - 0.75657 + \delta_{ag}$  where  $\delta_{ag}$  was normally distributed with zero mean and a standard deviation of 0.1649. This relationship was derived from in-situ measurements of Chl and  $a_{gd}$ . The distribution of  $b_0$  was assumed to be normal with a mean of 0.3 and a standard deviation of 0.07, and  $b_0$  was independent of  $a_{gd}(375)$ .

For each chlorophyll level, the following steps were taken:

**Step One:** The radiance model was run in the forward direction and forced by the ensemble of 200 in-water concentration vectors,  $C = [a_{gd}(375), \text{Chl}, b_0]$ , using the constant model parameter vector (Table 1) to generate a set of 200  $L_w(\lambda)$  vectors. Each  $L_w(\lambda)$  vector was then inverted to obtain the concentration vector  $C'$  which was then compared with  $C$ . The purpose of this step was to estimate errors due to the inversion method, since the parameter vector was assumed to be exact.

**Step Two:** The model was run forward again with a perturbed model parameter vector  $\Theta'(\lambda)$  as defined by Eq.(6) to produce another simulated data set of 200  $L_w(\lambda)$  vectors, and these were then inverted to obtain  $C'$  as in Step One. We perturbed each IOP separately to isolate the impact of individual IOPs, and then we added errors to all three IOPs to see their combined effect,

**Step Three:** In general, differences between vectors  $C'$  and  $C$  would be used to define retrieval errors caused by perturbations in the model parameter vector. In this paper, we concentrate only on the accuracy of the chlorophyll retrieval. The RMS error in both Chl and  $\log(\text{Chl})$  were used as measures of error. Specifically, from the sample of 200 retrievals, we calculated simple difference errors:  $\Delta = \text{Chl}' - \text{Chl}$ , and log difference errors  $\Delta_{\log} = \log \text{Chl}' - \log \text{Chl}$ , which are related to relative errors. For each type of error, two statistics were obtained: the mean error,  $M$  and  $M_{\log}$  and mean-square error,  $MSE$  and  $MSE_{\log}$ .

**Steps Two and Three** were repeated 100 times, each time with a new random set of parameter perturbations, applied simultaneously and individually. The statistics on  $M$ ,  $M_{\log}$ ,  $MSE$  and  $MSE_{\log}$  were accumulated. Then, the root-mean-square errors,  $RMSE$  and  $RMSE_{\log}$  were computed as the square roots of  $MSE$  and  $MSE_{\log}$ . Resulting statistics were thus based on  $N = 20,000$  random retrieval errors (100 simulations involving an ensemble of 200  $C$  vectors).

## Results and Discussion

### Effect of Inversion Scheme (Step One)

We found no errors resulting from the inversion scheme in Step One. That is, the Levenberg-Marquardt algorithm was able to retrieve each of the 200  $C$  vectors to an arbitrary level of accuracy controlled by the convergence criteria.

### Effect of IOP Uncertainties (Steps Two and Three)

Results for all the simulations are shown in Table 2, where the units of  $M$  and  $RMSE$  are  $mg\ Chl\ m^{-3}$  and the units of  $M_{log}$  and  $RMSE_{log}$  are decades of log. Results for the "Independent Error" simulations in plain font are followed by the "Equal Error" results in bold font. In all simulations, the "Equal Error" retrievals had much less error than the "Independent Error" retrievals. This was not surprising. "Equal Error" perturbations result in shifts of the IOP spectra upward or downward without changing their spectral shapes. The resulting radiances also tend to vary with minimal changes in spectral shape, and thus band ratios, for example, remain stable. In the "Independent Error" case, however, spectral shapes were altered significantly, and this resulted in much larger errors in the chlorophyll retrievals.

We were surprised at first by the fact that "Equal Error" perturbations in  $b_{bp}(\lambda)$  produced no errors in chlorophyll retrievals. In examining the retrievals of  $a_{gd}(375)$  and  $b_0$ , we found that only  $b_0$  retrievals had errors, and in fact its errors were precisely equal to the  $\delta_{bp}(\lambda)$  perturbation. The effect of a nonspectral perturbation in  $b_{bp}(\lambda)$  is equivalent to a perturbation in  $b_0$ . In Step One we found that perturbations in  $a_{gd}(375)$  and  $b_0$  did not affect our ability to retrieve Chl, in the absence of other IOP perturbations. This result is quite significant, because in the case of "Independent Errors" the  $b_{bp}(\lambda)$  uncertainty was the largest source of error (Table 2). *This suggests that the accuracy in chlorophyll retrievals can be improved significantly if the spectral shape of the particle backscattering coefficient is known, but that the absolute level of  $b_{bp}(\lambda)$  (whether shifted upward or downward) does not affect chlorophyll retrievals.*

The effect of uncertainty in  $a_{gd}(\lambda)$  and  $a_{\phi}(\lambda)$  depended on the chlorophyll level, with a general tendency of  $a_{\phi}(\lambda)$  becoming more critical as Chl increased. Comparing low-chlorophyll (Table 2a) and medium-chlorophyll (Table 2b) results, we see that errors tended to be proportional to the chlorophyll level. For example, in the "Independent Error" case, the combined  $M$  ranged from 10% to 12%, and the  $RMSE$  was between 63% and 70% of the chlorophyll level. The tendency for errors to be proportional to Chl is also indicated by the consistency in  $M_{log}$  and  $RMSE_{log}$  between Tables 2a and 2b, since the log-difference  $\Delta_{log}$  reflects a "relative" error.

In the high-chlorophyll "Independent Error" case (Table 2c), the combined  $RMSE$  represented only a 48% error, but the  $RMSE_{log}$  increased from 0.30 to 0.42 between medium and high chlorophyll cases. Further examination of the retrievals revealed a number of anomalously low chlorophyll retrievals ( $< 1\ mg\ m^{-3}$ ) when the true chlorophyll was high ( $10\ mg\ m^{-3}$ ). We have been unable to account for this, as we have not seen any pattern in either the  $a_{gd}(375)$ ,  $b_0$  variations, nor in the perturbations of the IOPs that is consistently related to these low chlorophyll values. We are continuing to examine this question.

The most significant finding was the improvement in retrieval accuracy that resulted from "Equal Errors" compared with "Independent Errors." Although the inherent variability in IOPs cannot be controlled, we believe it is important to model their spectral shape as accurately as possible. Knowledge of the spectral shape is critical, particularly in the case of the particle backscattering coefficient. Shifts in the IOP spectra upward or downward had little effect on chlorophyll retrieval accuracy, whereas random independent perturbations to the spectral IOP values resulted in very large errors ( $RMSE$  values as high as 70%).

## Summary

A general approach is presented to qualifying retrieval errors of in-water concentrations. The simulations focus on demonstrating how retrieval errors in chlorophyll concentration are affected by uncertainties of inherent optical property (IOP) submodels in an underlying radiance model. Two complete sets of simulations, which were designed and conducted, represent two extreme cases between which "real" cases may occur. The results from "Equal Errors" and "Independent Errors" are significantly different. It is suggested that precise determinations of spectral shapes of IOP submodels is important in chlorophyll retrieval.

## References

- Bukada, R.P, J.H. Jerome, K. Y. Kondratyev and D.V. Pozdnyakov, Optical properties and remote sensing of inland and coastal water, CRC Press, 1995.
- Bricaud, A, M. Babin, A. Morel, and H. Claustre, Variability in the chlorophyll-specific coefficients of natural phytoplankton: analysis and parameterization, 100, C7, 13,321-13,332, 1995.
- Campbell, W.J, .T. Moore, and H. Feng, Phytoplankton backscattering properties derived from satellite ocean color data, Abstract, ISLO97, Santa Fe, NM, 1997.
- Carder, K.L., S.K. Hawes, K.A. Baker, R.C. Smith, R.G. Steward, and B.G. Mitchell, Reflectance model for quantifying chlorophyll a in the presence of productivity degradation products, J.Geophys. Res., 96, 20,599-20,611,1991.
- Feng, H., J.W. Campbell, and T. Moore, Parameterization of water color model from in-situ measurements from Tokyo Bay, EOS Trans. AGU, 79(3), 1998.
- Garver, S.A. and D.A. Seigel, Global Applications of the UASB non-linear inherent optical property inversion model, ( A manuscript)
- Gordon, H.R., O.B. Brown, R.H.Evans, J.W.Brown, R.C. Smith, K.S. Baker, and D.K. Clark, A semianalytical radiance model of ocean color, J. Geophys. Res., 93, 10,909-10,924, 1988.
- Hoge, F. E. and P.E. Lyon, Satellite retrieval of inherent optical properties by linear matrix inversion of oceanic radiance model: an analysis of model and radiance measurement error. J. Geophys. Res., 100(C7), 16631-16648, 1996.
- Press, W.H. and S.A. Teukolsy, W.T. Vetterling and B.P. Flannery, Numerical Recipes in C: The Art of Scientific Computing, Cambridge Press, 2nd edition, 994pp, 1992.
- Zaneveld, R., A theoretical derivation of the dependence of the remotely sensed reflectance of the ocean on the inherent optical properties, J.Geophys. Res., 100, C7, 13,135-13,142,1995.

**Table 1.** Model parameters assumed to be constant in inverting the radiance model. Parameters are listed in column 1, units in column 2, and values for each spectral band in columns 3-7.

Parameters	Units	412nm	443nm	490nm	510nm	555nm
$rQ(\lambda)$	None	1.92	1.92	1.92	1.92	1.92
$M(\lambda)$	None	0.5375	0.5373	0.5398	0.5390	0.5390
$F_0(\lambda)$	Mw/cm <sup>2</sup> /μm	171.7	189.2	194.4	187.5	185.9
$A_c(\lambda)$	m <sup>-1</sup>	0.0313	0.0393	0.0274	0.0180	0.0071
$B_c(\lambda)$	None	0.7270	0.6600	0.6390	0.7400	0.9660
$A_b(\lambda)$	None	0.0111	0.0100	0.0104	0.0108	0.0109
$B_b(\lambda)$	None	0.2390	0.2250	0.2850	0.3190	0.3620
S	nm <sup>-1</sup>	0.0145	0.0145	0.0145	0.0145	0.0145

**Table 2.** Average error statistics: M, RMSE,  $M_{log}$ , and  $RMSE_{log}$  after perturbing each IOP submodel separately, and after perturbing all three IOP submodels ("combined") for spectrally-independent case and spectrally-equal case(bold numbers). The averages of M, RMSE,  $M_{log}$ , and  $RMSE_{log}$  are shown here for the 100 simulations where steps 2 and 3 are repeated.**Table 2a.** Case of low chlorophyll (Chl = 0.1 mg m<sup>-3</sup>).

Submodel	M	RMSE	$M_{log}$	$RMSE_{log}$
$a_{gd}(\lambda)$	0.0029/ <b>0.0029</b>	0.0276/ <b>0.0276</b>	-0.0022/ <b>-0.0022</b>	0.1137/ <b>0.1137</b>
$a_{\phi}(\lambda)$	0.0014/ <b>0.0009</b>	0.0206/ <b>0.0119</b>	-0.0027/ <b>0.0011</b>	0.0887/ <b>0.0505</b>
$b_{bp}(\lambda)$	0.0059/ <b>0.0000</b>	0.0500/ <b>0.0000</b>	-0.0273/ <b>0.0000</b>	0.2317/ <b>0.0000</b>
combined	0.0102/ <b>0.0039</b>	0.0629/ <b>0.0301</b>	-0.0330/ <b>-0.0011</b>	0.2817/ <b>0.1250</b>

**Table 2b.** Case of medium chlorophyll (Chl = 1.0 mg m<sup>-3</sup>).

Submodel	M	RMSE	$M_{log}$	$RMSE_{log}$
$a_{gd}(\lambda)$	0.0239/ <b>0.0239</b>	0.2581/ <b>0.2581</b>	-0.0032/ <b>-0.0032</b>	0.1094/ <b>0.1094</b>
$a_{\phi}(\lambda)$	0.0128/ <b>0.0118</b>	0.2081/ <b>0.1387</b>	-0.0038/ <b>0.0011</b>	0.0906/ <b>0.0589</b>
$b_{bp}(\lambda)$	0.0828/ <b>0.0000</b>	0.5768/ <b>0.0000</b>	-0.0273/ <b>0.0000</b>	0.2543/ <b>0.0000</b>
combined	0.1207/ <b>0.0356</b>	0.6991/ <b>0.2938</b>	-0.0352/ <b>-0.0022</b>	0.3022/ <b>0.1254</b>

**Table 2c.** Case of high chlorophyll (Chl = 10 mg m<sup>-3</sup>).

Submodel	M	RMSE	$M_{log}$	$RMSE_{log}$
$a_{gd}(\lambda)$	-0.1680/ <b>-0.1680</b>	0.7008/ <b>0.7008</b>	-0.0087/ <b>-0.0087</b>	0.0378/ <b>0.0378</b>
$a_{\phi}(\lambda)$	-0.4090/ <b>-0.1744</b>	2.5222/ <b>0.8445</b>	-0.0384/ <b>-0.0094</b>	0.1552/ <b>0.0414</b>
$b_{bp}(\lambda)$	-1.0078/ <b>0.0108</b>	4.0424/ <b>0.0158</b>	-0.1208/ <b>0.0005</b>	0.3474/ <b>0.0007</b>
combined	-1.0772/ <b>-0.2949</b>	4.7737/ <b>1.2029</b>	-0.1549/ <b>-0.0174</b>	0.4244/ <b>0.0697</b>

IN-SITU DATA NEEDED TO SUPPORT THE DEVELOPMENT OF  
ALGORITHMS FOR OCEAN COLOR REMOTE SENSING IN COASTAL  
REGIONS: A PROPOSED PILOT PROJECT FOR C-GOOS

Sinjaee Yoo and Janet Campbell

August 12, 1998

## INTRODUCTION

The ultimate goal is to develop a network of laboratories supplying the in-situ data needed to parameterize optical properties of coastal waters. This information will become the basis for algorithms applied to satellite ocean color data as part of the Coastal Global Ocean Observing System (C-GOOS). This capability will facilitate the operational use of remote sensing as part of a continuing, long-term monitoring program of the coastal ocean environment.

A C-GOOS pilot project is proposed to initiate the development of this network, and to demonstrate the use of in-situ data for constructing site-specific algorithms for coastal regions. It is proposed to conduct the pilot project in two coastal regions: (1) the Yellow and East China Seas Region where turbid waters from the Huang He (Yellow) and Yangtze rivers mix with oceanic waters west of the Korean peninsula; and (2) the Chesapeake Bay Region where programs of coastal monitoring involving remote sensing have recently been established. Techniques for making the relevant in-situ measurements will be developed and protocols carefully documented. The data will be used to parametrize reflectance models which will become the basis for the regional algorithms. Laboratories in each region will be invited to participate in collecting data that will be put into a shared data base. The network and data base will be expanded in future years to include coastal regions worldwide.

## BACKGROUND

Ocean color remote sensing techniques have been successfully applied to estimate algal pigment concentrations in the open ocean. Optical properties of these so-called "Case 1 waters" are relatively well-characterized because the only materials affecting ocean color are the phytoplankton cells or their decay products. Optical properties of coastal and estuarine waters are generally more complex. In these "Case 2 waters," mixtures of organic and inorganic materials affect the color of the water, and this complexity calls for more sophisticated algorithms for sorting out the various constituents found in proximity to land.

Present or near-future ocean color sensors (Table 1) should be capable of quantifying up to two additional constituents, in addition to algal pigments, which are found in coastal regions. Spectral bands have been selected to quantify concentrations of phytoplankton chlorophyll (CHL), total suspended sediment (TSS), and colored dissolved organic matter (CDOM). However, the optical properties of these materials (particularly TSS) are not universal. Site-specific algorithms must be developed to account for the unique optical characteristics of the materials found in each region.

In many respects, the coastal oceans remain as a frontier for future ocean color technology development. Present-day sensors listed in Table 1 lack the spatial and temporal resolutions

needed to resolve space scales of order 100 m associated with nearshore and estuarine features, and time scales dictated by the diurnal and semi-diurnal tides. Furthermore, improved spectral resolution may be needed to distinguish more complex mixtures in "Case 2" waters. Two emerging technologies are aimed at solving these limitations.

A hyperspectral imager, the Coastal Ocean Imaging Spectrometer (COIS), will be flown on the U.S. Navy's NEMO satellite to be launched in 2000. This sensor will collect data only at about 50 specific coastal sites. Its spatial resolution will be 30-60 meters depending on the mode of operation, and typical scenes will be about 30 km by 300 km. The sensor will have 200 2-nm-wide spectral bands covering the visible and near-infrared spectral region.

To address the need for high temporal resolution, a "Special Events Imager" is proposed to be flown on a geostationary satellite. The present design calls for 10-12 spectral bands with 300 m resolution. The area viewed in any image would be about 300 km x 300 km. Images of an event can be refreshed every 10 minutes allowing high temporal resolution of rapidly changing conditions. Coastal ocean applications would include oil spills, river plumes, storm flooding, hurricanes, and support of process-oriented field work.

## CASE 2 ALGORITHMS

A general framework for developing coastal or "Case 2" algorithms is described by Campbell and Yoo (1998). This framework can serve as the basis for the C-GOOS pilot project. The algorithms described are called "semi-analytic algorithms" because they are based on principles of radiative transfer (hence analytic), but involve empirical parameterizations. Radiative transfer theory has been used to derive a robust (though approximate) relationship between the "remote sensing reflectance" (derived from satellite measurements) and the water's inherent optical properties (IOPs). Inherent optical properties (absorption and backscattering coefficients) are related to the water constituents through empirically derived relationships based on in-situ data.

Semi-analytic algorithms invert the reflectance model to derive constituent concentrations (chlorophyll, CDOM, and total suspended sediment). The spectral absorption and backscattering coefficients may also be derived in an intermediate step. In the following section, we describe the data needed to parameterize semi-analytic algorithms.

## IN-SITU DATA: RETRIEVED VARIABLES

The goal of a semi-analytic algorithm is to "retrieve" variable properties of the water related to the constituents which affect water color. These properties are called "retrieval variables." The remote-sensing reflectance is related to materials in the upper optical depth (or secchi depth). Water samples collected and profiles measured should provide an adequate representation of this layer. The number of samples and the depth of profiles will depend on the vertical structure or stratification.

There can be some flexibility in deciding what in-water chemical or biological properties are to be retrieved by the algorithm. The only requirement is that the material have a distinctive effect on the color of the water. In some cases, the retrieval variables will depend on the application or the analytical methods available to measure them. Here we describe three substances which we categorize heuristically according to color: "blue absorbing" CDOM, "green" chlorophyll, and "brown" sediment.

Colored dissolved organic matter (CDOM). Also known as "yellow substance" or "gelbstoff", CDOM includes any dissolved material affecting the water color after particles are removed by filtration. The actual chemical composition of CDOM is poorly defined

(often it includes tannins and lignins leached from the soil but may also include materials excreted by the phytoplankton cells). Although its chemical makeup is not well-defined, the optical properties of CDOM are relatively stable. Rather than use a mass concentration, CDOM is usually measured in terms of its absorption at a reference wavelength,  $a_g(\lambda_0)$ . Thus it is determined by optical measurements as described below.

Chlorophyll concentration (CHL). Chlorophyll-a in extracted pigments measured either by HPLC, fluorometry, or spectrophotometry. Details of the method used should be specified.

Total suspended matter (TSM). The dry weight (mass) of material collected on a filter. The same pore size filter should be used as that used for the optical measurements (e.g., particle absorption).

Total suspended sediment (TSS) The dry weight (mass) left on the same filter after combustion to remove organic matter.

The retrieval variable will be either TSM or TSS. Compared with TSM (which includes phytoplankton and organic detritus), the inorganic TSS should have an optical signature that is more distinguishable from chlorophyll, and hence might be a better retrieval variable. However, in some regions, mineral particles are coated with an organic film, in which case optical properties may be better correlated with TSM.

The variables listed above represent the "retrieved" constituents and hence are a minimum set of measurements needed for algorithm development. Other standard oceanographic measurements such as sea surface temperature and salinity should also be included in the data base if available. These properties may prove useful in deciding between algorithms, or there may be indirect relationships (e.g., between CDOM and salinity) which can be exploited.

Properties that have a direct effect on the optics should be measured and included in the data if possible (though not essential). These include: particle size spectra, or size-fractionated chlorophyll and TSM; dominant phytoplankton species and/or accessory pigments. Other information such as nutrients, tidal stage, and primary productivity, may prove useful in subsequent applications after the retrieved variables are obtained from remote sensing measurements.

## IN-SITU OPTICAL MEASUREMENTS

All of the optical properties are spectral measurements. That is, they are made at discrete spectral wavelengths or within spectral bands. Measurements made with high spectral resolution can be integrated to simulate the response of a variety of sensors. The disadvantage is that the required instruments tend to be expensive, and there is more data to deal with initially. Less expensive instruments are available which simulate the spectral bands of specific sensors. Whether one begins with full spectral data or data collected in discrete bands, the optical data used for algorithm development should be representative of the spectral response of the ocean color sensor to which the algorithm will be applied.

Upwelling spectral radiance. Profile of upwelling spectral radiance measured below the water surface within the upper optical depth.

Downwelling spectral irradiance. Profile of downwelling spectral irradiance measured at the same depths and same wavelengths as the upwelling radiance.

Particle absorption. Particles collected on a filter (same pore size as used for TSM) and measured with a spectrophotometer. Unless the samples are analyzed immediately, the water should be filtered and samples preserved by freezing (e.g., in liquid N<sub>2</sub>).

Pigment absorption. After measuring the particle absorption, pigments are extracted and the remaining material reanalyzed with the spectrophotometer. The difference is the pigment absorption.

CDOM absorption. This is the absorption measured on the filtrate in a cuvette as referenced to distilled water. The spectrum should follow a negative exponential:  $a_g(\lambda) = a_g(\lambda_0) \exp[-S(\lambda - \lambda_0)]$  where S is a positive constant usually between 0.01 and 0.02 nm<sup>-1</sup> and  $\lambda_0$  is a reference wavelength (usually  $\leq 400$  nm). The concentration of CDOM is indexed by  $a_g(\lambda_0)$ , which is the retrieval variable "CDOM."

Particle backscattering. Whereas measurements of the total scattering coefficient (b) derived from transmissometer measurements are relatively common, measurements of the backscattering coefficient ( $b_b$ ) are rare. Campbell and Yoo (1998) describe a method for deriving an "estimated" backscattering coefficient using the other optical properties described above, when backscattering is missing, but ideally the data set should include the spectral backscattering coefficient. This is obtained from a light scattering meter which measures light scattered at discrete angles. A volume scattering function fitted to the scattering data is then integrated over angles  $\geq 90$  degrees to obtain the backscattering coefficient. New instruments are now becoming available for making the appropriate scattering measurements.

With the above suite of data, semi-analytical algorithms can be parameterized as described in Campbell and Yoo (1998).

## PROTOCOL TESTING AND DOCUMENTATION

Standard protocols do not exist for many of the measurements listed above. The laboratories participating in the pilot project will test various methods and decide upon a prescribed set of protocols. These will be fully documented and reviewed by members of NASA's SIMBIOS Science Team before being published. The trade-off between the need for expensive optical instruments versus simple standard techniques will be considered. Our goal will be to develop methods and protocols that can be readily implemented with a minimum cost by many laboratories worldwide.

## TRAINING

Once protocols are tested and documented, we propose that a series of training courses be offered to instruct scientists and technicians in making the measurements, and in applying their in-situ data to parameterize algorithms for their own region. In exchange for training funded through IOC, participants will be required to provide data to the C-GOOS data base.

Table 1. Ocean Color Sensors

Sensor	Satellite	Country	Dates	Spatial Resolution	# Bands in Vis-NIR	Comments
CZCS	Nimbus-7	USA	Nov. '78 - Jul. '86	825 m	5	Proof-of-concept instrument
MOS	IRS P3	Germany/India	Mar. '96 -	523 m	13	Requires ground receiving station
OCTS	ADEOS-1	Japan	Aug. '96 - Jun. '97	700 m	8	Plus 4 thermal IR bands for SST
SeaWiFS	SeaStar	USA	Aug. '97 -	1100 m	8	Full-resolution (LAC) data requires ground receiving station
OCI	ROCSAT-1	Taiwan	Dec. '98 -	800	6	A 7th band is redundant 555-nm band; 35-deg. inclined orbit.
OCM	IRS P4 (Oceansat-1)	India	Late '98 -	360 m	8	Scanning multifrequency microwave to provide SST
MODIS	EOS AM-1 EOS PM-1	USA	1999 - 2000-	1000 m	9	Plus 27 other bands for land, atmosphere, and ocean SST
OSMI	KOMPSAT-1	S. Korea	1999 -		8	Selectable bands
MERIS	ENVISAT	ESA	2000 -	250 m LAC 1000 m GAC	15	LAC data reception requires receiving station
GLI	ADEOS-2	Japan	2000 -		9	Very similar to MODIS, also has 27 additional bands for other appl.
C-OCTS	HY-1	China	2000 -	1100 m	10	Plus 2 thermal IR bands for SST