Chapter 5. MODIS Semi-Analytic Algorithm for IOP

K. L. Carder, J. P. Cannizzaro, F. R. Chen, Z. P. Lee

5.1 Introduction

The Moderate-Resolution Imaging Spectrometer (MODIS) semi-analytic algorithm (Carder_MODIS here after) (Carder et al., 1999; Carder et al., 2004) derives chlorophyll-a concentrations and inherent optical properties (IOP) (phytoplankton absorption coefficients, $a_{ph}(\lambda)$; combined detrital and gelbstoff absorption coefficients, $a_{dg}(\lambda)$; and particle backscattering coefficients, $b_{bp}(\lambda)$) from remote-sensing reflectance spectrum, $R_{rs}(\lambda)$. This algorithm is composed with an algebraic portion and an empirical portion. The algebraic portion is for waters with low absorption (mostly oceanic waters) while the empirical portion is for waters with high absorption (mostly coastal waters). The main characters of this algorithm include that it responds to the large global variability observed in 1) chlorophyll-specific phytoplankton absorption coefficients, $a^*_{ph}(\_)$ and 2) gelbstoff-to-phytoplankton absorption ratios. This algorithm utilizes differences between measured sea-surface temperatures and known nitrate-depletion temperatures (NDT) (Kamykowski & Zentara, 1986; Kamykowski, 1987) to select the most appropriate $a^*_{ph}(\_)$ for a given bio-optical domain. The algorithm was first developed and evaluated using high-light, tropical/subtropical and summer temperate field data (Carder et al., 1999) and later expanded to include parameters appropriate for low-light, polar data (Carder et al., 2004).

5.2 Algorithm Description

5.2.1 Remote-sensing reflectance model

By making several approximations, the $R_{rs}(\_)$ used in Carder_MODIS algorithm is simplified to (Carder et al., 1999)

$$R_{rs}(\lambda) \approx constant \frac{b_b(\lambda)}{a(\lambda)} \tag{5.1}$$

where the "constant" is unchanging with respect to wavelength and solar zenith angle. The value of the constant is not relevant to the algorithm since, as will be shown later, the algorithm (for absorption and chlorophyll-a concentration) uses spectral ratios of $R_{rs}(\_)$ and the constant term factors out.

Further, both $b_b(\_)$ and $a(\_)$ are partitioned into several separate terms. Each term is described empirically and is written in a general fashion as a function of variables and empirically derived parameters. Since sea surface temperatures were not provided in the IOCCG data sets, the unpackaged parameters regarding $a^*_{ph}(\_)$ derived from high-light, tropical/subtropical and summer temperate waters were employed (Carder et al., 1999) (Table 1). While $a^*_{ph}(\_)$ is extremely important for deriving chlorophyll-a concentrations accurately, retrievals of $a_{ph}(\lambda)$ and $a_{dg}(\lambda)$ are less sensitive to differences in $a^*_{ph}(\_)$.

5.2.2 Backscattering coefficients

The total backscattering coefficient, $b_b(\_)$, can be expanded as

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda) \tag{5.2}$$

where the subscripts "w" and "p" refer to water and particles, respectively. $b_{bw}(\_)$ is constant and well known (Morel, 1974).

Particulate backscattering, $b_{bp}(\_)$, is modeled as (Carder et al., 1999)

$$b_{bp}(\lambda) = X\left[\frac{551}{\lambda}\right]^{Y} \tag{5.3}$$

where X is the magnitude of particulate backscattering at 551nm, $b_{bp}(551)$, and Y describes the spectral shape of the particle backscattering spectra. Values for X and Y were determined empirically by model inversion (Lee et al., 1994) and are described as

$$X = X_0 + X_1 R_{rs}(551) \tag{5.4}$$

$$Y = Y_0 + Y_1 \frac{R_{rs}(443)}{R_{rs}(488)} \tag{5.5}$$

where $X_{0,1}$ and $Y_{0,1}$ are empirically derived constants (Table 1).

When absorption due to water molecules does not dominate the total absorption coefficient at 551 nm, algorithms that utilize wavelengths longer than 551 nm that take advantage of the large inflection in the pure water absorption spectra between 570-610nm (Pope & Fry, 1997) are required. Using measurements of $R_{rs}(\lambda)$ and $b_{bp}(\lambda)$ collected from the west Florida shelf (WFS), the equation

$$b_{bp}(551) = 10^{(0.933 - 0.134\log(R_{rs}(551)) + 1.029\log(R_{rs}(667)))} - 0.000966 \tag{5.6}$$

(n=154, R^2=0.96, RMSE = 0.160) was derived for MODIS-like wavelengths. This function was used when Carder_MODIS algorithm was applied to the IOCCG synthetic data set. Since remote-sensing data with wavelengths longer than 555 nm were not available for the IOCCG *in situ* data set, however, $b_{bp}(551)$ values were estimated for these data using Eq. 5.4.

5.2.3 Absorption coefficients

The total absorption coefficient, $a(\lambda)$, can be expanded as

$$a(\lambda) = a_w(\lambda) + a_{ph}(\lambda) + a_{dg}(\lambda) \tag{5.7}$$

where the subscripts "w", "ph," and "dg" refer to water, phytoplankton, and the combination of detritus and gelbstoff, respectively. Here, $a_w(\_)$ is taken from Pope and Fry (1997).

The shape of the $a_{ph}(\_)$ spectrum for a given water mass changes due to the pigment-package effect and changes in pigment composition. For the MODIS wavebands centered at _ = 412, 443, 488, and 551nm, a hyperbolic tangent function was chosen to empirically model the ratio of $a_{ph}(\_)/a_{ph}(675)$ in order to ensure that this ratio approaches an asymptote at very high or very low values of $a_{ph}(675)$ (Carder et al., 1999)

$$a_{ph}(\lambda) = a_0(\lambda)\exp\left[a_1(\lambda)\tanh\left[a_2(\lambda)\ln\left(a_{ph}(675)/a_3(\lambda)\right)\right]\right] * a_{ph}(675) \tag{5.8}$$

The parameters $a_{0\text{-}3}(\_)$ are provided in Table 1.

The cumulative effects of detritus and gelbstoff absorption, $a_{dg}(\lambda)$, are expressed as

$$a_{dg}(\lambda) = a_{dg}(400)\exp^{-S(\lambda-400)} \tag{5.9}$$

where S is the spectral slope, and a value of 0.0225 nm^{-1} provided optimal retrieval results for the Carder_MODIS algorithm to calculate chlorophyll-a concentrations (Carder et al., 1999). It is larger than the mean ocean value of about 0.015 nm^{-1}, likely compensating in part for uncertainties in other parts of the model.

5.2.4 Model inversion

Via Eqs. 5.1-5.9, $R_{rs}(\_)$ is reduced to a function of three unknowns ("constant" term, $a_{ph}(675)$, and $a_{dg}(400)$) along with model constants for $X_{0,1}$, $Y_{0,1}$, $a_{0\text{-}3}(\_)$, and S (Table 1). To

solve for the values of the two desired unknowns ($a_{ph}(675)$ and $a_{dg}(400)$), spectral ratios of 412:443 and 443:551 for $R_{rs}(\_)$ as shown

$$\frac{R_{rs}(412)}{R_{rs}(443)} = \frac{b_b(412)}{b_b(443)}\frac{a(443)}{a(412)}, \qquad \frac{R_{rs}(443)}{R_{rs}(551)} = \frac{b_b(443)}{b_b(551)}\frac{a(551)}{a(443)}. \tag{5.10}$$

provided the best separation of the two absorption contributions. The equations can be solved algebraically to provide values for $a_{ph}(675)$ and $a_{dg}(400)$. Details on the computational method of solving these equations are discussed in Carder et al. (1999).

5.2.5 Semi-analytical IOP's for other MODIS wavebands

Phytoplankton absorption coefficients for MODIS wavebands (412, 443, 488, and 551nm) are calculated by inserting the modeled $a_{ph}(675)$ value and empirical factors from Table 1 into Eq. 5.8. Combined detrital and gelbstoff absorption coefficients for MODIS wavebands are calculated by inserting the modeled $a_{dg}(400)$ value and an S value of 0.0225 nm^{-1} into Eq. 5.9. Total absorption coefficients at 412, 443, 488 and 551nm are then calculated using Eq. 5.7 by combining $a_{ph}(\_)$ and $a_{dg}(\_)$ with $a_w(\_)$ (Pope & Fry, 1997).

5.2.6 Empirical portion of Carder_MODIS

For waters with high concentrations of gelbstoff and chlorophyll, $R_{rs}(412)$ and $R_{rs}(443)$ values are small, and therefore the above semi-analytical approach cannot perform properly due to low signal-to-noise ratios. Thus the semi-analytic approach is designed to return values only when modeled $a_{ph}(675)$ values are less than 0.025 m^{-1}, which is equivalent to a chlorophyll concentration of about 1.5 mg m^{-3}. Otherwise, the following empirical algorithms derived from west Florida shelf (1999-2001) and Bayboro Harbor (St. Petersburg, Fla.) field data (n=319) are used.

For $a_{ph}(443)$, there is

$$a_{ph}(443)_{emp} = 10^{\left(-1.164-1.2095\rho_{35}-1.566\rho_{35}^{\,2}-1.708\rho_{45}+19.502\rho_{45}^{\,2}\right)} \tag{5.11}$$

where $\__{ij}$ is the log-transformed ratio of $R_{rs}(\lambda_i)$ to $R_{rs}(\lambda_j)$ and the subscripts i and j are wavebands #1-6 that represent MODIS wavebands 412, 443, 488, 531, 551, and 667nm, respectively. Since this equation requires the MODIS $R_{rs}(531)$ waveband and the SeaWiFS waveband $R_{rs}(510)$ was provided instead with the IOCCG *in situ* data, a modified SeaWiFS algorithm was also developed

$$a_{ph}(443)_{emp} = 10^{\left(-1.189-1.133\rho_{35}-2.151\rho_{35}^{\,2}-0.775\rho_{45S}+7.592\rho_{45S}^{\,2}\right)} \tag{5.12}$$

and applied to the IOCCG *in situ* data set. Here ρ_{45S} is equal to $\log(R_{rs}(510)/R_{rs}(555))$.

The empirical algorithm for $a_{dg}(443)$ is

$$a_{dg}(443)_{emp} = 10^{\left(-1.144-0.738\rho_{15}-1.386\rho_{15}^{\,2}-0.644\rho_{25}+2.451\rho_{25}^{\,2}\right)} \tag{5.13}$$

and was applied to the IOCCG *in situ* data set. Since adding a ρ_{65} term reduced the RMS error by 40% for calculating $a_{dg}(443)$ for the WFS and Bayboro Harbor data, the derived equation

$$a_{dg}(443)_{emp} = 10^{\left(0.043-0.185\rho_{25}-1.081\rho_{35}+1.234\rho_{65}\right)} \tag{5.14}$$

was applied to the synthetic data set where $R_{rs}(670)$ data (considered equal to $R_{rs}(667)$) were available.

Empirical retrievals of $a(\lambda)$ at 412, 443, and 488nm also improved for the WFS and Bayboro Harbor data set when a red reflectance waveband was included. Thus, the empirical expression applied to the synthetic data set takes the form

$$a(\lambda)_{emp} = 10^{(c_0(\lambda)+c_1(\lambda)\log(R_{rs}(443))+c_2(\lambda)\log(R_{rs}(488))+c_3(\lambda)\log(R_{rs}(667)))} \quad (5.15)$$

where $c_{0\text{-}3}(\lambda)$ are empirically derived parameters (Table 2). Note that while reflectance ratios are used to calculate $a_{ph}(443)_{emp}$ and $a_{dg}(443)_{emp}$, reflectance values are used to calculate $a(\lambda)_{emp}$ in Eq. 5.15. For the IOCCG *in situ* data set that does not have a red reflectance waveband, an empirical expression similar to that of Lee et al. (1998)

$$a(\lambda)_{emp} = 10^{(t_0(\lambda)+t_1(\lambda)\rho_{25}+t_2(\lambda)\rho_{25}^2+t_3(\lambda)\rho_{35}+t_4(\lambda)\rho_{35}^2)} \quad (5.16)$$

is used, with $t_{0\text{-}4}(\lambda)$ empirically derived parameters (Table 3).

5.2.7 Blending semi-analytic and empirical IOP values

In order to provide a smooth transition in modeled IOP values when the algorithm switches from the semi-analytical to the empirical method, a weighted average of the modeled values returned by both algorithms is used near the transition border (Carder et al., 1999). When the semi-analytical portion returns an $a_{ph}(675)$ value between 0.015 and 0.025 m^{-1}, IOP values are calculated as

$$\text{IOP} = w(\text{IOP})_{sa} + (1 - w)(\text{IOP})_{emp} \quad (5.17)$$

where $(IOP)_{sa}$ is the semi-analytically-derived value, $(IOP)_{emp}$ is the empirically derived value, and w is the weighting factor equal to $[0.025\text{-}a_{ph}(675)]/0.015$. Semi-analytical and empirical IOP values are used when modeled $a_{ph}(675)$ values are less than 0.015 m^{-1} and greater than 0.025 m^{-1}, respectively. Note that this transition range can vary with pigment packaging (e.g. see Carder et al. 2004).

5.3 Algorithm performance with the IOCCG data sets

The Carder_MODIS algorithm requires $R_{rs}(\lambda)$ data at a minimum of five wavebands: 412, 443, 488, 531, and 551nm. Further inclusion of the $R_{rs}(667)$ waveband improves retrievals of $a_{dg}(443)_{emp}$ (Eq. 5.14) and $a(\lambda)_{emp}$ (Eq. 5.15) values. Since the synthetic $R_{rs}(\lambda)$ data was generated in 10nm increments from 400-800nm, reflectance values at 410, 440, 490, 530, 550, and 670nm were considered similar enough to the MODIS wavebands and were input into the algorithm. The IOCCG *in situ* data set contains $R_{rs}(\lambda)$ data at 412, 443, 490, 510, and 555nm which were input into the equations that do not require a 667nm waveband.

5.3.1 Synthetic data set

Using the Carder_MODIS algorithm, the inherent optical properties $a(412)$, $a(443)$, $a(488)$, $a_{ph}(443)$, $a_{dg}(443)$, and $b_{bp}(551)$ were derived from the synthetic $R_{rs}(\lambda)$ data (Figure 1). Statistical analyses were performed on log-transformed data and include the slope, y-offset, and correlation coefficient (R^2) of the best-fit linear regression function and the root-mean-square (RMSE) error (Table 4a).

Particulate backscattering coefficients at 550nm retrieved using Eq. 5.6 from $R_{rs}(550)$ and $R_{rs}(670)$ are very accurate (RMSE = 4.2%). Total absorption coefficients at 412, 443, and 488 nm were also retrieved accurately for the synthetic data set with RMSE errors equal to 7.1, 5.9, and 6.5%, respectively. RMSE values for $a_{ph}(443)$ (14.1%) and $a_{dg}(443)$, (13.5%) are slightly

more than double the error calculated for $a(443)$ since phytoplankton and detritus/gelbstoff exhibit overlapping absorption spectra making it difficult to separate them.

5.3.2 *In situ* data set

The results of the Carder_MODIS algorithm when applied to the IOCCG *in situ* data set were not as good as the results observed for the synthetic data set because errors in field R_{rs} and IOP data, not present in the synthetic data, are significant in the *in situ* data.

Total absorption coefficients at 412, 443, and 488nm were derived from the *in situ* $R_{rs}(\lambda)$ data yielding RMSE errors equal to 19.7, 20.4, and 20.6%, respectively (Figure 2, Table 4b). Errors for $a_{ph}(443)$ and $a_{dg}(443)$ were only slightly higher than $a(443)$ and were 19.5% and 27.9%, respectively. While the semi-analytic $a_{ph}(443)$ values derived from synthetic $R_{rs}(\lambda)$ data were overestimated, values derived from the *in situ* $R_{rs}(\lambda)$ data were more centered about the one-to-one line. This may indicate that perhaps the underlying $a_{ph}(\lambda)$ functions used to generate the synthetic data for oligotrophic waters are not quite representative of the distribution of the naturally occurring $a_{ph}(\lambda)$ data, or at least Eq. 5.8 is more consistent with the $a_{ph}(\lambda)$ functionality of the *in-situ* data set than with that of the synthetic data set.

Large errors that occur in empirically derived $a(412)$, $a(443)$, and $a_{ph}(443)$ values and appear as linear horizonal rows of data in Figure 2 at ~0.23, 0.15, and 0.07 m^{-1}, respectively, can be traced to a single investigator for a large multi-year, coastal data set. Removal of these points would improve the performance of the empirical portion of our algorithm. Furthermore, empirical retrievals of $a(\lambda)$ and $a_{dg}(440)$ may also improve for this data set if $R_{rs}(\lambda)$ data were available for wavelengths longer than 555nm.

5.4 Conclusions

The Carder_MODIS algorithm (Carder et al., 1999) calculated $b_{bp}(550)$ and $a(\lambda)$ values very accurately for the synthetic data set. Values for $a_{ph}(443)$ and $a_{dg}(443)$ were calculated less accurately because phytoplankton and detritus/gelbstoff exhibit overlapping absorption making it more difficult to separate them using $R_{rs}(\lambda)$. Retrieval errors tripled for $a(\lambda)$ and doubled for $a_{dg}(443)$ when the algorithm was applied to the *in situ* data set as compared to the synthetic data set. The fact that the partitioned values fell within the same error range as the total-absorption values suggests that much of the error imputed to the algorithms for the *in situ* data set may be attributable to errors or inconsistencies among the measured data sets, whereas the synthetic data set had no measurement noise.

IOP retrieval errors calculated for the *in situ* data set may improve if $R_{rs}(667)$ data were available. Significant error reductions were observed for empirically derived backscattering and total absorption coefficients when red reflectance data were used for our high-absorption Florida data set and for the synthetic data set. Note, however, that while $R_{rs}(667)$ can be used for "perfect" synthetic data, accurate measurements of $R_{rs}(667)$ from space are much more subject to error due to smaller signal-to-noise ratios. A waveband near 610-620 nm would perhaps be a better compromise than use of 667nm for satellites. The Medium Resolution Imaging Spectrometer (MERIS) has such a waveband at 620 nm.

Finally, the expansion of available global data sets in the past 10 years and the broad range of data synthesized in the numerical data set have provided examples of how various older algorithms may be improved, and we are grateful for being included in this challenging algorithm intercomparison.

References

Bricaud, A., Babin, M., & Morel, A. (1995), Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization. *J. Geophys. Res.*, 100, 13,321-13,332.

Carder, K.L., Chen, F.R., Cannizzaro, J.P., Campbell, J.W., & Mitchell, B.G. (2004), Performance of the MODIS semi-analytical ocean color algorithm for chlorophyll-*a*. *Adv. Space Res.*, 33, 1152-1159.

Carder, K.L., Chen, F.R., Lee, Z.P., Hawes, S.K., & Kamykowski, D. (1999), Semi-analytic Moderate-Resolution Imaging Spectrometer algorithms for chlorophyll *a* and absorption with bio-optical domains based on nitrate-depletion temperatures. *J. Geophys. Res.*, 104, 5403-5422.

Kamykowski, D. (1987), A preliminary biophysical model of the relationship between temperature and plant nutrients in the upper ocean. *Deep-Sea Res.*, 34, 1067-1079.

Kamykowski, D. & Zentara, S.-J. (1986), Predicting plant nutrient concentrations from temperature and sigma-t in the upper kilometer of the world ocean. *Deep-Sea Res.*, 33, 89-105.

Lee, Z., Carder, K.L., Hawes, S.K., Steward, R.G., Peacock, T.G., et al. (1994), Model for the interpretation of hyperspectral remote-sensing reflectance. *Appl. Opt.*, 33, 5721-5732.

Lee, Z.P., Carder, K.L., Steward, R.G., Peacock, T.G., Davis, C.O., et al. (1998), An empirical algorithm for light absorption by ocean water based on color. *J. Geophys. Res.*, 103, 27,967-27,978.

Morel, A. (1974), Optical properties of pure water and pure sea water. In Jerlov, N.G., & Nielsen, E.S. (Ed.), *Optical Aspects of Oceanography* (pp. 1-24). London: Academic Press.

Pope, R. & Fry, E. (1997), Absorption spectrum (380-700nm) of pure waters: II. Integrating cavity measurements. *Appl. Opt.*, 36, 8710-8723.

Table 1. Parameters for the MODIS semi-analytical algorithm for regions without packaged pigments.

λ	a_0	a_1	a_2	a_3	X_0	X_1	Y_0	Y_1	S
412	2.20	0.75	-0.5	0.0112	-0.00182	2.058	-1.13	2.57	0.0225
443	3.59	0.80							
488	2.27	0.59							
551	0.42	-0.22							

Table 2. Wavelength-dependent parameters for the high-absorption empirical $a(\lambda)$ algorithm (Eq. 5.15) that requires $R_{rs}(670)$.

	$c_0(\lambda)$	$c_1(\lambda)$	$c_2(\lambda)$	$c_3(\lambda)$
$a(412)$	-0.349	-1.041	0.171	0.754
$a(443)$	-0.166	0.068	-1.284	1.077
$a(488)$	-0.167	0.478	-1.639	1.075

Table 3. Wavelength-dependent parameters for the high-absorption empirical $a(\lambda)$ algorithm (Eq. 5.16) that does not require $R_{rs}(670)$.

	$t_0(\lambda)$	$t_1(\lambda)$	$t_2(\lambda)$	$t_3(\lambda)$	$t_4(\lambda)$
$a(412)$	-0.640	-0.718	-0.650	-1.365	2.369
$a(443)$	-0.837	-0.860	-0.791	-1.162	2.855
$a(488)$	-0.947	-0.343	-0.721	-1.633	2.741

Table 4a. $RMSE_{log10}$ and linear-regression results of the synthetic data set (θ^0=30°). R_{rs}(_) values at 410, 440, 490, 530, 550, and 670 nm were used as inputs.

	Offset (in log10)	Slope (in log10)	R^2	N	$RMSE_{log10}$
a(410)	0.015	0.990	0.990	500	0.071
a(440)	0.030	1.030	0.993	500	0.059
a(490)	0.079	1.082	0.993	500	0.065
a_{ph}(440)	-0.046	0.908	0.963	500	0.141
a_{dg}(440)	0.083	1.098	0.978	500	0.135
b_{bp}(550)	-0.012	0.998	0.995	500	0.042

Table 4b: $RMSE_{log10}$ and linear-regression results of the *in situ* data set. R_{rs}(_) values at 412, 443, 490, 510, and 555 were used as inputs.

	Offset (in log10)	Slope (in log10)	R^2	N	$RMSE_{log10}$
a(412)	0.098	1.066	0.826	656	0.197
a(443)	0.030	1.111	0.831	656	0.205
a(488)	0.131	1.173	0.789	656	0.206
a_{ph}(443)	-0.052	0.986	0.827	656	0.195
a_{dg}(443)	-0.041	1.082	0.771	656	0.279

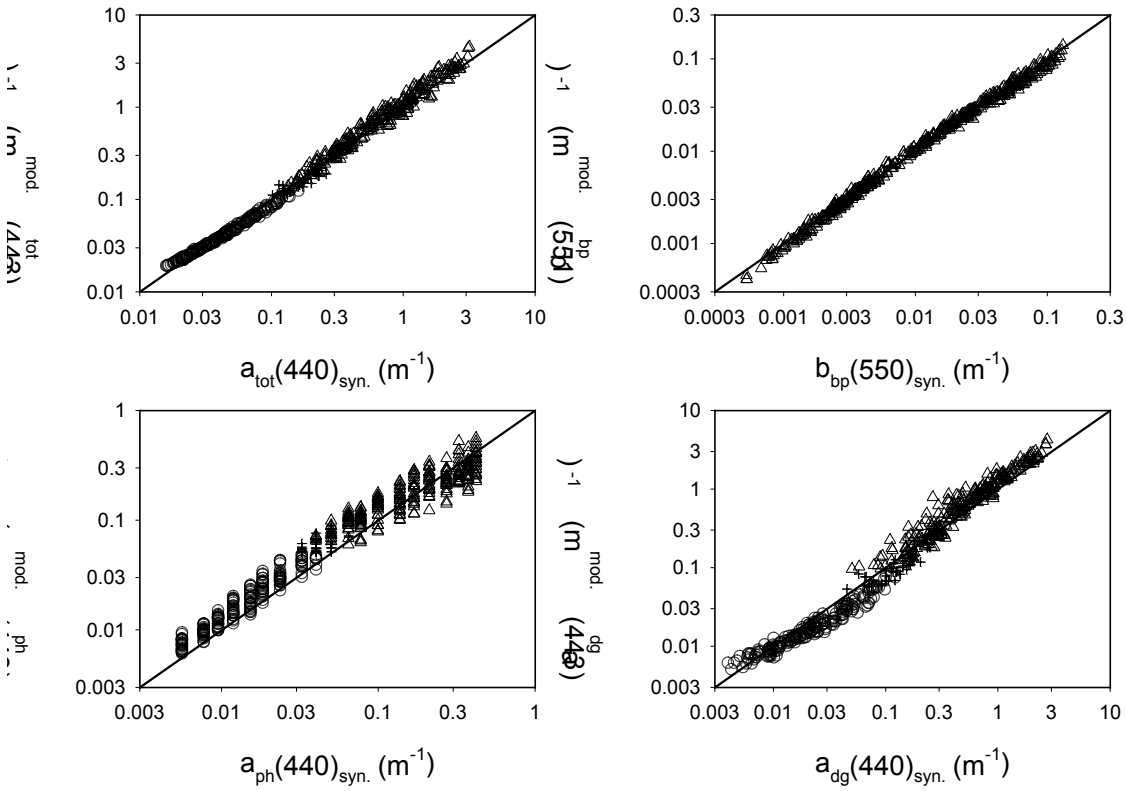

Figure 1. Relationships between known and retrieved IOPs using the Carder_MODIS algorithm (synthetic data set), with $R_{rs}(\_)$ at 410, 440, 490, 530, 550, and 670 nm used as inputs. Symbols: semi-analytic (o), blended (+), and empirical (-).

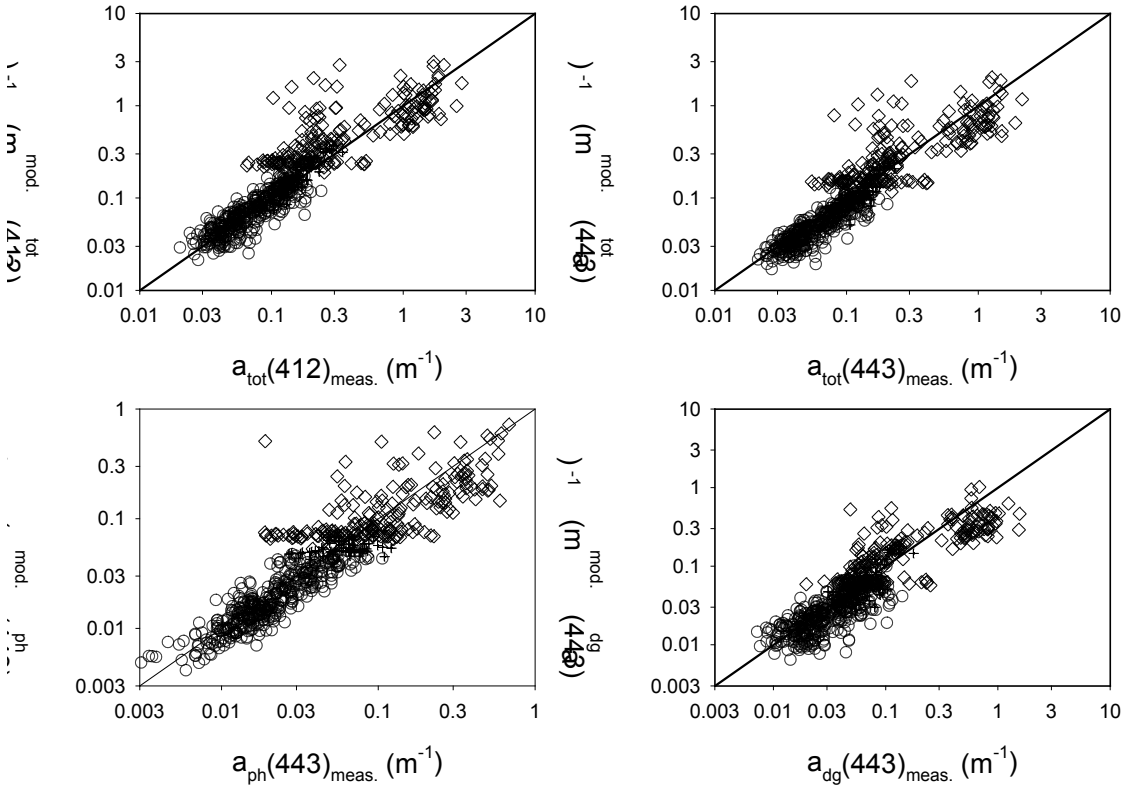

Figure 2. Relationships between measured and retrieved IOPs using the Carder_MODIS algorithm (*in situ* data set), with $R_{rs}(\_)$ at 412, 443, 490, 510, and 555 nm used as inputs. Symbols: semi-analytic (o), blended (+), and empirical (-).