Chapter 9

Optimization Of Ocean Color Algorithms: Application To Satellite And In Situ Data Merging.

Stéphane Maritorena and David A. Siegel

Institute for Computational Earth System Science, UCSB, Santa Barbara, CA 93106-3060, USA

André Morel

Laboratoire de Physique et Chimie Marines, Villefranche Sur Mer, France

9.1 INTRODUCTION

The objective of our program is to develop and validate a procedure for ocean color data merging which is one of the major goals of the SIMBIOS project (McClain et al., 1995). The need for a merging capability is dictated by the fact that since the launch of MODIS on the Terra platform and over the next decade, several global ocean color missions from various space agencies are or will be operational simultaneously. The apparent redundancy in simultaneous ocean color missions can actually be exploited to various benefits. The most obvious benefit is improved coverage (Gregg et al., 1998; Gregg & Woodward, 1998). The patchy and uneven daily coverage from any single sensor can be improved by using a combination of sensors. Beside improved coverage of the global ocean the merging of ocean color data should also result in new, improved, more diverse and better data products with lower uncertainties. Ultimately, ocean color data merging should result in the development of a unified, scientific quality, ocean color time series, from SeaWiFS to NPOESS and beyond.

Various approaches can be used for ocean color data merging and several have been tested within the frame of the SIMBIOS program (see e.g. Kwiatkowska & Fargion, 2003, Franz et al., 2003). As part of the SIMBIOS Program, we have developed a merging method for ocean color data. Conversely to other methods our approach does not combine end-products like the subsurface chlorophyll concentration (chl) from different sensors to generate a unified product. Instead, our procedure uses the normalized water-leaving radiances ($L_{wN}(\lambda)$) from single or multiple sensors and uses them in the inversion of a semi-analytical ocean color model that allows the retrieval of several ocean color variables simultaneously. Beside ensuring simultaneity and consistency of the retrievals (all products are derived from a single algorithm), this model-based approach has various benefits over techniques that blend end-products (e.g. chlorophyll): 1) it works with single or multiple data sources regardless of their specific bands, 2) it exploits band redundancies and band differences, 3) it accounts for uncertainties in the $L_{wN}(\lambda)$ data and, 4) it provides uncertainty estimates for the retrieved variables.

9.2 RESEARCH ACTIVITIES

Development of an ocean color database and algorithm development

Over the past 3 years, we have assembled a large comprehensive in situ ocean color data set that contains inherent (IOP) and apparent (AOP) optical properties as well as chlorophyll a concentration data from various locations. This database is designed for ocean color algorithm development and is well suited for semi-analytical algorithm development in particular. Since it contains both IOPs and AOPs, this database is better suited for semi-analytical algorithms development than data sets like the one used during SeaBAM (O'Reilly et al., 1998) which only contained chl and remote sensing reflectance data. Most of the data included in the database come from the NASA SIMBIOS SeaBASS archive but several investigators have provided data sets or subsets directly to us. Various quality control (QC) procedures have been developed (Fargion & McClain, 2003) to identify corrupted data, outliers or specific bio-optical situations. The database contains chlorophyll a concentration, diffuse attenuation coefficients, L_{wN} (λ), particulate backscattering and component absorption (phytoplankton, detrital, dissolved). Most of the absorption data are hyperspectral. The current status of the IOP/AOP data set is described in Table 9.1.

Table 9.1: Status of the AOP/IOP data set. The first number in each cell indicates the number of stations for which data are available. The numbers in parentheses indicate the number of available wavelengths

EXPERIMENT	**Chl a**	**Kd**	**Rrs**	**bb**	**ad**	**ag**	**ap**
AAOT	94	-	35 (5)	-	-	-	-
Aerosols Indoex	53	3367 (12)	5309 (6)	-	70 (225)	70 (200)	73 (225)
Ace-Asia	116	696 (18)	903 (18)	780 (6)	48 (225)	69 (200)	48 (225)
AMLR	268	1060 (18)	1302 (18)	-	101 (225)	91 (200)	101 (225)
Atlantic Meridional Transect	281	281 (5)	281 (5)	-	-	-	-
Bermuda Bio-Optics Project	406	4024 (16)	6505 (12)	-	77 (520)	77 (520)	77 (520)
Arc00	159	-	7262 (5)	-	-	-	-
BIOCOMPLEXITY	300	-	6505 (13)	-	74 (500)	74 (315)	74 (500)
CALCOFI	430	938 (12)	1064 (23)	-	159 (225)	130 (200)	143 (225)
CaTS	11	-	16 (5)	-	-	-	-
CoJet	53	-	74 (15)	-	-	-	-
CSC	23	-	23 (5)	-	-	-	-
DOER	10	-	10 (5)	-	-	-	-
ECOFRONT	37	3142 (6)	2168 (7)	-	16 (225)	16 (200)	16 (225)
Ecology of Harmful Algal Blooms	429	-	246 (200)	-	422 (200)	398 (200)	422 (200)
EPIC	152	1988 (5)	3179 (5)	-	-	-	-
JGOFS-EQPAC	125	1785-	1737 (5)	-	-	-	-
Globec Biomapper	45	-	-	-	43 (200)	43 (200)	43 (200)
GOCAL	NA	-	1328 (7)	28 (6)	-	-	-
NASA-Gulf of Maine	148	-	-	-	150 (200)	137 (200)	150 (200)
NOAA-Gulf of Maine	316	-	-	-	-	-	-
JGOFS-SIO	89	1671 (12)	1875 (12)	-	23 (225)	23 (225)	23 (225)
JGOFS-Arabian Sea	2693	3085 (5)	3083 (5)	-	-	-	-
Kieber Photochemistry 2002	427	1791 (5)	1772 (5)	-	-	-	-
Lab2000	51	-	3191 (5)	-	-	-	-
Lab1997	204	-	-	-	-	-	-
Lab1996	42	-	113 (12)	-	122 (500)	74 (500)	122 (500)
LMER-TIES	237	-	76 (7)	-	85 (250)	85 (200)	85 (250)
MBARI	34	-	34 (5)	-	-	-	-
MOCE	13	-	13 (5)	-	-	-	-
North Atlantic Bloom Experiment	110	609 (5)	431 (5)	-	-	-	-
Ocean Research Consortium	17	-	18 (10)	14 (6)	16 (470)	16 (470)	16 (470)
Plumes and Blooms	786	7702 (7)	10330 (12)	3867 (7)	500 (13)	431 (13)	505 (13)
ROAVERRS	162	-	162 (5)	-	-	-	-
Sea of Japan	54	593 (12)	777 (12)	666 (6)	51 (225)	51 (200)	51 (225)
Scotia Prince Ferry	7420	-	60 (5)	7147 (3)	-	6591 (9)	6478 (9)
Tongue of the Ocean	88	-	86 (200)	16 (6)	95 (200)	95 (225)	95 (200)
World Ocean Circulation Experim	68	-	68 (5)	-	-	-	-
Total Measurements	**15951**	**32732**	**60036**	**12518**	**2052**	**8471**	**8522**

The IOP/AOP database is designed in part to help the development of our merging model. The transformation of the model into a fully hyperspectral mode (which would make it usable right away with any sensor) and improved parameterizations of some of the components of the model are the major modifications we are trying to implement. A preliminary version of the hyperspectral model has been developed and optimized using data from the AOP/IOP database. Although it shows some good overall results for all three retrieved variables this preliminary hyperspectral version does not always perform well at the extremes of the chlorophyll range (either in very clear or very rich waters). This new version of the model still requires some work and a more conservative, step-by-step approach is now used for its development.

Satellite Ocean Color Data Merging

Our approach for ocean color data merging is based on the inversion of a semi-analytical model that relates L_{wN} to the backscattering and absorption coefficients (Eq. 9.1) as described in Gordon et al. (1988).

$$L_{wN}(\lambda) = \frac{t\,F_0(\lambda)}{n_w^2} \sum_{i=1}^{2} g_i \left(\frac{b_{bw}(\lambda) + b_{bp}(\lambda)}{b_{bw}(\lambda) + b_{bp}(\lambda) + a_w(\lambda) + a_{ph}(\lambda) + a_{cdm}(\lambda)} \right)^i \qquad (9.1)$$

Each of the non-water components in a and b_b is expressed as a known shape function with an unknown magnitude:

$$a_{ph}(\lambda) = \text{Chl}\ a_{ph}^*(\lambda) \qquad (9.2)$$
$$a_{cdm}(\lambda) = a_{cdm}(443)\ \exp(-S(\lambda-443) \qquad (9.3)$$
$$b_{bp}(\lambda) = b_{bp}(443)\ (\lambda/443)^{-\eta} \qquad (9.4)$$

where t, n_w, g_i, $F_o(\lambda)$, $a_w(\lambda)$ and $b_{bw}(\lambda)$ at taken from the literature whereas η, S, $a_{ph}^*(\lambda)$ were determined by "tuning" the model against a large in situ data set (Maritorena et al., 2002). A non-linear least-square fitting technique is used to solve for the unknowns (chl, $a_{cdm}(443)$ and $b_{bp}(443)$) from $L_{wN}(\lambda)$ data at 4 or more wavelengths. The model also provides uncertainty estimates for each of the retrievals using a linear approximation to the calculation of non-linear regression inference regions (Bates & Watts, 1988). The model, hereafter referred to as the GSM01 model, is fully described in Garver & Siegel (1997) and Maritorena et al. (2002).

Since the model retrievals are generated using a curve-fitting technique that minimizes the least-squares differences between the measured and modeled $L_{wN}(\lambda)$, it is straightforward to use the GSM01 model with multiple data sets like in data merging. When multiple data sets are used (e.g. two or more sensors have L_{wN} measurements over a given pixel), the $L_{wN}(\lambda)$ data from all available sources are concatenated (as are the relevant wavebands information) allowing the curve-fitting step to be conducted with more data points than with a single source. When data sources have different bands, the fitting procedure also benefits from an increased spectral resolution. A key aspect of the procedure is that there is no transformation or averaging of the input L_{wN} data, they are used "as is" in the curve-fitting technique. A schematic of the input and output products of the merging model is presented in figure 9.1.

Our model-based approach for ocean color data merging was first tested using SeaWiFS and MOS data and results were presented during the SIMBIOS science Team meeting in Baltimore (Jan. 15-17, 2002). Since then, we have successfully used SeaWiFS and MODIS data (from both the Terra and Aqua platforms). We have tested our merging approach using daily level-3 data from SeaWiFS (reprocessing # 4, 9 km) and MODIS (collection # 4, 4.6 km) for 18 dates between December 4, 2000 and March 22, 2003. Only the MODIS "best quality" data (i.e. quality 0) were used during these tests. Since the SeaWiFS and MODIS $L_{wN}(\lambda)$ data products have different spatial resolution, it is necessary to first adapt the MODIS data to the SeaWiFS resolution by averaging four 4.6 km bins into a 9 km bin and to have the 2 data sources set to a common binned grid. The data are processed between 65 degrees North and 65 degrees South. We have focused most of our effort on the period for which MODIS Terra collection # 4 products are validated (11/1/2000-3/19/2002). Some data outside this time window were also used to illustrate SeaWiFS/Terra/Aqua data merging and to assess improvement in coverage when merging 3 different ocean color data sources.

9.3 RESEARCH RESULTS

The aim of our SIMBIOS work is to demonstrate the feasibility of an ocean color data merging procedure based on a semi-analytical mode that uses $L_{wN}(\lambda)$ data from one or more sources. It is out of the scope of this report to document the accuracy of the model retrievals with in situ or satellite data (but see Maritorena et al., 2002; Siegel et al., 2002 and Siegel et al., 2003. Examples of global maps of chlorophyll a concentration, $a_{cdm}(443)$ and $b_{bp}(443)$ generated by the GSM01 merging model with SeaWiFS and MODIS-Terra $L_{wN}(\lambda)$ data are presented in figure 9.2. Global maps of chl, $a_{cdm}(443)$ and the $b_{bp}(443)$ images generated by the GSM01 merging model using the Terra and SeaWiFS data show very good consistency overall (Maritorena et al., 2003). In general, the retrieved fields do not show discontinuities when the model switched from an area with a single data source to an area where both SeaWiFS and MODIS $L_{wN}(\lambda)$ data were used. This reflects the generally good agreement between the $L_{wN}(\lambda)$ data from both sources and the robustness of the model. The level of agreement between the two sensors also has an influence on the estimated uncertainties of the derived products. When considering the pixels that are covered by both sensors, a very large majority of them show reduced uncertainties in the merged products compared to those generated from a single data source. Figure 9.3 shows the frequency distribution of the ratio of the Chl uncertainties using either SeaWIFS or MODIS L_{wN} data alone over the Chl uncertainties when both sources are used. Overall, the uncertainties tend to decrease when multiple sources are used and this is true for all 3 products generated by the GSM01 model. This decrease in the uncertainties of the derived products is observed for all products and at all the dates we have processed. In the worst case, 70% of the pixels showed lower uncertainties in the merged products. Uncertainties are generally higher in the merged products when L_{wN} data do not agree well between the data sources.

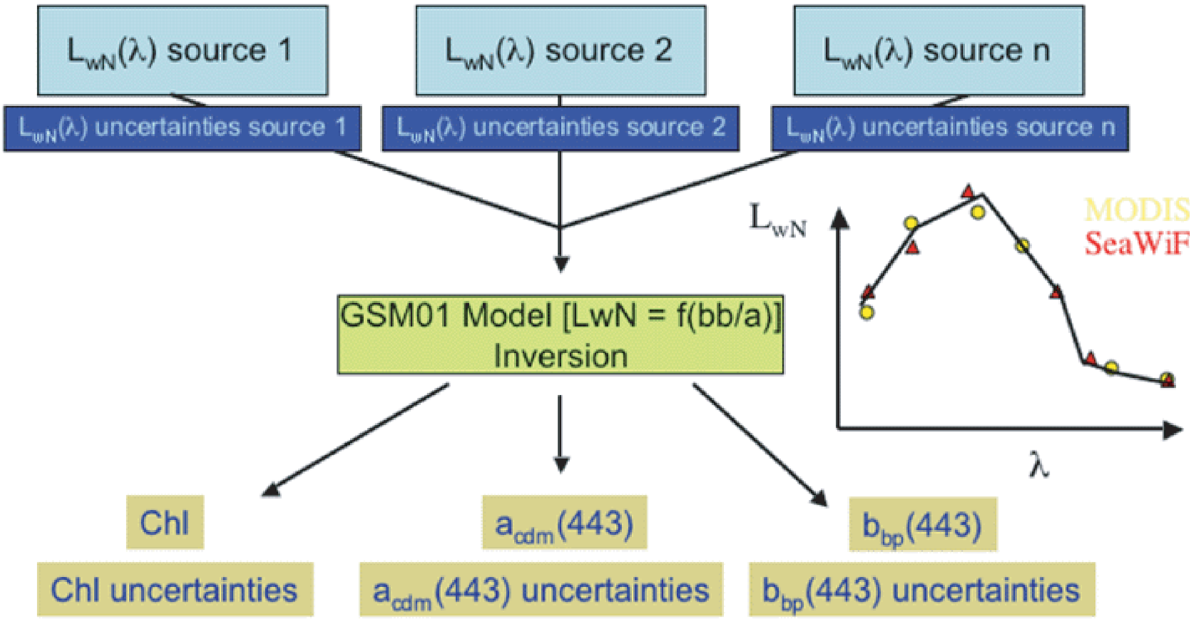

Figure 9.1: Schematic of the input and output data of the semi-analytical ocean color merging model.

This is illustrated in figure 9.4 where $L_{wN}(443)$ from both sensors are compared when uncertainties improved in the merged product (merged/unmerged uncertainty ratio of 0.5 or less) and when they got worse (uncertainty ratio > 1). When the uncertainties are lower in the merged product, the agreement between both sets of $L_{wN}(\lambda)$ is generally very good with most of the data points on or very close to the 1:1 line (figure 9.4 shows the 443 data but this is true for all bands). When the uncertainties were higher in the merged images, the $L_{wN}(\lambda)$ data showed some clear differences between SeaWiFS and MODIS. Areas where the products uncertainties are actually higher in the merged image are frequently next to gaps caused by sun glint, gaps between the swaths or clouds. They also appear to be mostly located in the south hemisphere. Further analyses are needed to assess what causes these features.

Improvement in coverage is obvious. The daily surface area effectively covered by any individual sensor depends upon various factors such as the sensor's technical and orbital characteristics, sun glint, cloud cover and season. The increased coverage that results from the use of multiple data sources is illustrated in figure 9.5 for the 18 dates we have used. Daily coverage jumps from ~12-15% of the ocean surface (in the 65N to 65S range) when SeaWiFS is used alone to ~25% when it is used with MODIS-Terra. When MODIS-Aqua data are used in the merging process along with SeaWiFS and Terra the daily percentage coverage reaches ~30-35% to the ocean surface. These numbers agree well with those derived from a theoretical analysis prior to SeaWiFS and MODIS launches (Gregg and Woodward, 1998) as well as with those obtained by the SIMBIOS Project with their Level-3 merged chlorophyll product (Franz et al.).

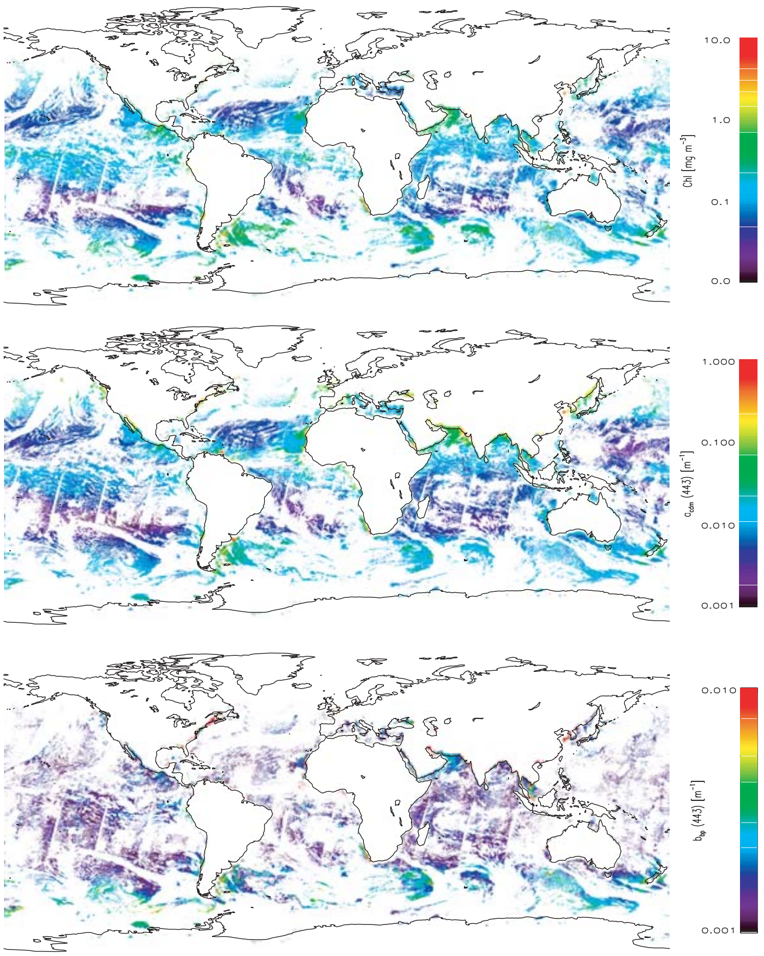

Figure 9.2: Example of daily images (December 4, 2000) of chl (upper panel), a_{cdm} (443) (center panel) and b_{bp}(443) (lower panel) generated by the GSM01 merging model using daily level-3 $L_{wN}(\lambda)$ data from SeaWiFS and MODIS-Terra.

The absence of marked discontinuities in the mapped product and the lower products uncertainties are two important results. At this point, it should be mentioned that some of the features of our merging approach cannot be used to their full efficiency mostly because the MODIS data are not fully characterized or stabilized. For example, the direct use of all available L_{wN} data in the fitting procedure assumes they agree well and have similar or close uncertainty levels and that none of the data sources contains noticeable bias. This may not be always true. Although not used in the results presented here, the merging model has the ability of weighting each individual $L_{wN}(\lambda)$ data to insure that the best observations are given a higher weight in the fitting procedure that leads to the derivation of the retrievals. Uncertainties ($\sigma_i(\lambda_j)$) of input $L_{wN}(\lambda)$ can be accounted for in the the least squares minimization (LSM) procedure as

$$\mathrm{LSM} = \sum_{i=1}^{N_{sat}} \sum_{j=1}^{N_{\lambda i}} \left[\frac{L_{wN\text{-}i}(\lambda_j)_{mod.} - L_{wN\text{-}i}(\lambda_j)_{meas.}}{\sigma_i(\lambda_j)} \right]^2 \qquad (9.5)$$

where N_{sat} is the number of data source and $N\lambda$ is the number of bands for each source. This has not yet been used mostly because the uncertainties associated with the MODIS bands of Terra and Aqua cannot yet be fully assessed in time and space. This requires matchup analyses from a large and diverse set of in situ and satellite data. These analyses are available for SeaWiFS but more matchup points are needed to complete the analysis for the MODIS data. It is also necessary to have some knowledge of the uncertainties variability in space and time. Once the characterization of Terra and Aqua is detailed enough, it should be possible to implement uncertainty weighting functions. A consistent BRDF correction scheme for the sensors involved in the data merging would also represent an improvement and upcoming reprocessings of SeaWiFS and MODIS data should take care of that particular aspect.

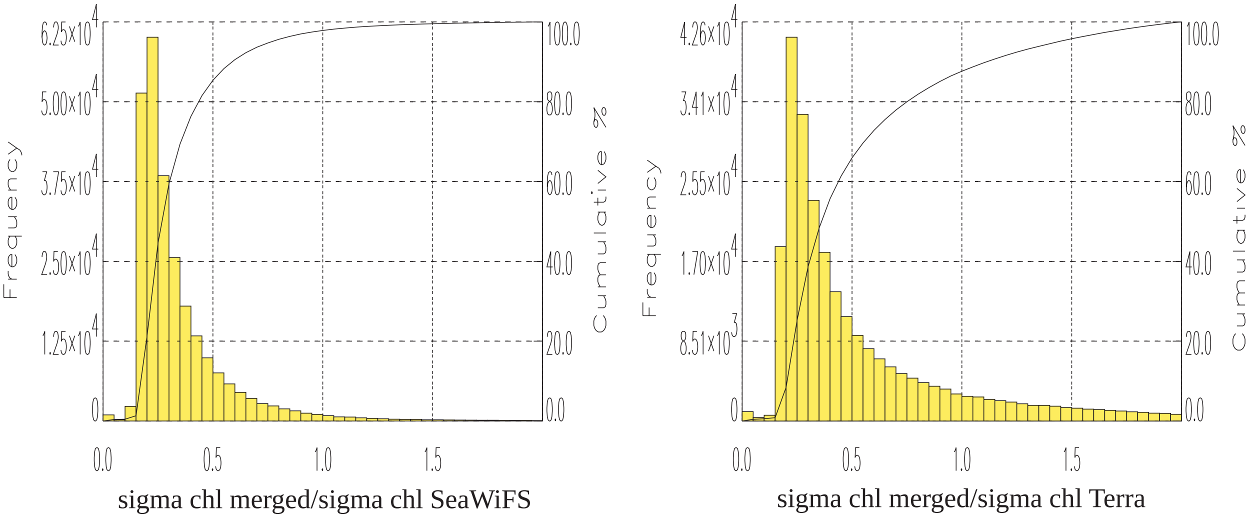

Figure 9.3: Frequency distribution of the ratio of the chl uncertainties using either SeaWiFS or MODIS-Terra L_{wN} data alone (December 4, 2000) over the chl uncertainties when both sources are used (for the same pixels). Ratios < 1 show an improvement in the uncertainties of the merged products.

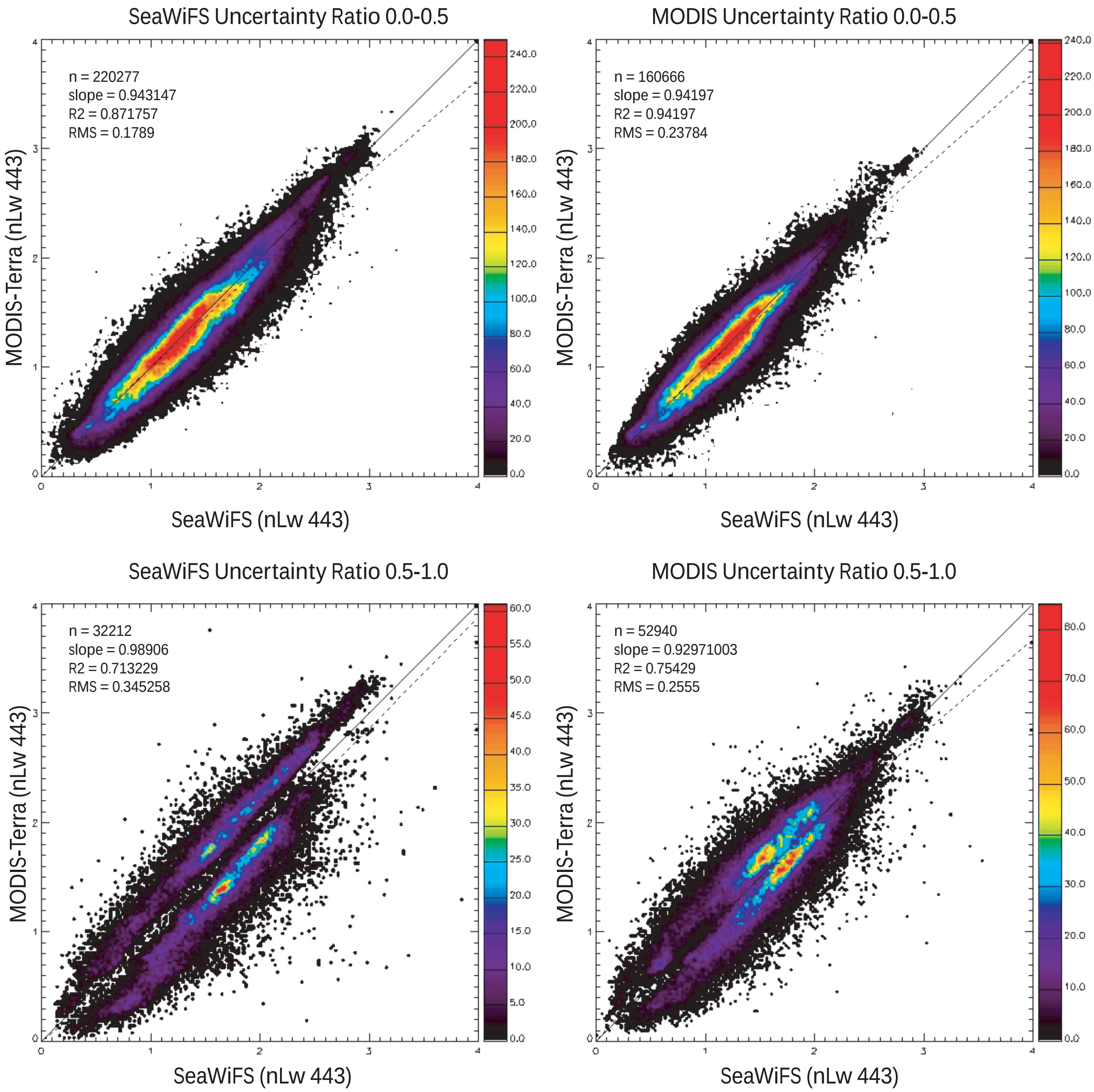

Figure 9.4: Comparison of L_{wN}(443) data from SeaWiFS and MODIS-Terra data when uncertainties improved in the merged chl product (merged/unmerged uncertainty ratio of 0.5 or less, upper panels) and when they got worse (uncertainty ratio > 1, lower panels).

In this project, we have demonstrated that ocean color data merging can effectively be conducted using L_{wN} data and the inversion of a semi-analytical algorithm. The method can be applied straightforwardly to any suite of ocean color sensors. Beside the feasibility aspect, the improvement in daily coverage and the lower uncertainties in the merged products are two important results of our work. Some refinements (e.g. weighting functions based on the uncertainty levels of the input $L_{wN}(\lambda)$ data, BRDF correction) can be added to the current approach in the future when some of the satellite data will be more mature

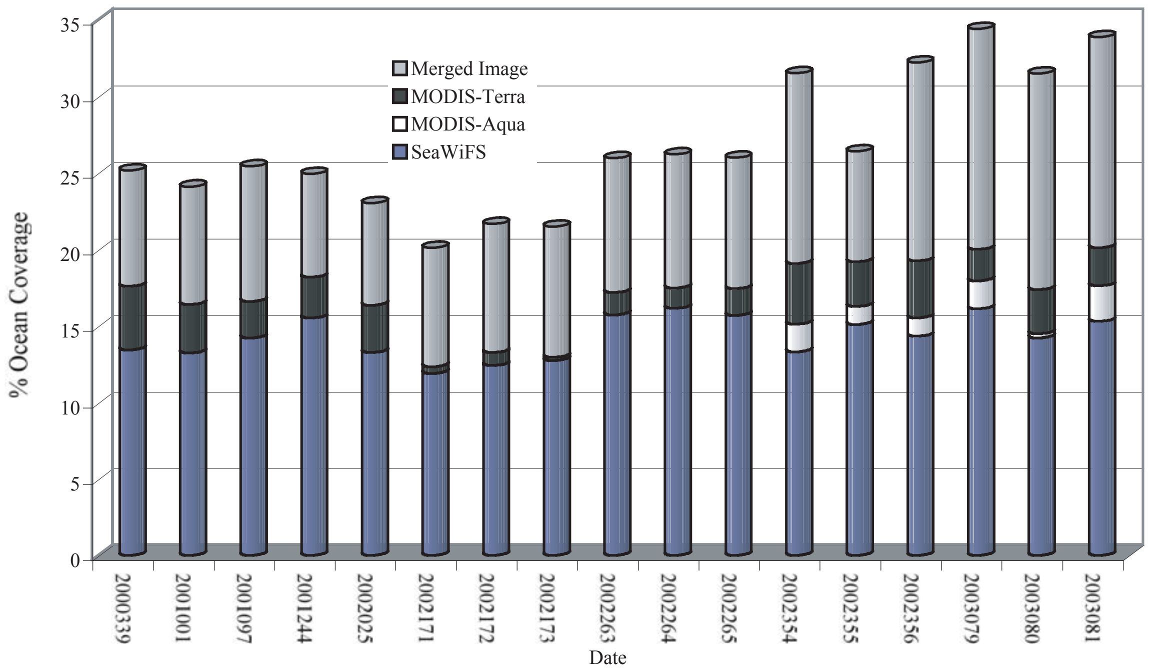

Figure 9.5: Daily coverage resulting from the merging of SeaWiFS, MODIS-Terra and MODIS-Aqua for the 18 dates used in this study. The coverage is computed as the percentage of the total ocean area (between 65 degrees North and 65 degrees South).

REFERENCES

Bates, D.M., and D.G. Watts, 1988: *Nonlinear Regression Analysis and its Applications*. Wily, 365 pp.Garver, S.A., and D.A. Siegel, 1997: Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: I. Time series from the Sargasso Sea. *J. Geophys. Res.*, **102**, 18,607-18,625.

Fargion G.S. and C.R. McClain, 2003: SIMBIOS Project Annual Report 2002, *NASA Technical Memorandum*, **TM-2003-211622**, NASA/GSFC, Greenbelt, Maryland, p157.

Franz B.A., J.G. Wilding and J.M. Gales, 2003: Operational Merging of MODIS and SeaWiFS Ocean Color Products at Level-3. In "MODIS Validation, Data Merger and Other Activities by the SIMBIOS Project:2002-2003", Fargion G.S. and C.R. McClain, NASA Technical Memorandum, TM-2003-, NASA/GSFC, Greenbelt, Maryland, (In Press).

Garver, S.A., and D.A. Siegel, 1997: Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: I. Time series from the Sargasso Sea. *J. Geophys. Res.*, 102, 18,607-18,625.

Gordon, H.R., O.B. Brown, R.H. Evans, J.W. Brown, R.C. Smith, K.S. Baker, and D.K. Clark. 1988: A semi-analytic radiance model of ocean color. *J. Geophys. Res.,* **93**, 10,909-10,924.

Gregg, W.W., W.E. Esaias, G.C. Feldman, R. Frouin, S.B. Hooker, C.R. McClain and R.H. Woodward. 1998: Coverage opportunities for global ocean color in a multimission era. IEEE Transactions On Geoscience And Remote Sensing, **36** :1620-1627.

Gregg, W.W. and R.H. Woodward. 1998: Improvements in coverage frequency of ocean color: Combining data from SeaWiFS and MODIS. IEEE Transactions On Geoscience And Remote Sensing, **36** :1350-1353.

Kwiatkowska, E.J. & G.S. Fargion, 2003. Merger of ocean color data from multiple satellite missions within the SIMBIOS project. Ocean remote sensing and applications, R. Frouin, Y. Yuan and H. Kawamura [Eds.], Proceedings of SPIE, Vol. **4892**, 168-182.

McClain, C.R., W. Esaias, G. Feldman, R. Frouin, W. Gregg and S. Hooker. 2002: The Proposal for the NASA Sensor Intercalibration and Merger for Biological and Interdisciplinary Oceanic Studies (SIMBIOS) Program, 1995. NASA Technical Memorandum, *TM-2002-210008*, NASA/GSFC, Greenbelt, Maryland, p54.

Maritorena S., D.A. Siegel and A. Peterson. 2002: Optimization of a Semi-Analytical Ocean Color Model for Global Scale Applications. *Applied Optics*. **41**(15): 2705-2714.

Maritorena, S., D.A. Siegel and D.B. Court, 2003: Ocean Color Data Merging Using a Semi-Analytical Algorithm. In preparation for Remote sensing of Environment.

O'Reilly J. E., S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru and C. R. McClain. 1998. Ocean Color Chlorophyll Algorithms for SeaWiFS. *J. Geophys. Res.*, **103**(C11): 24,937-24,953.

Siegel, D.A., S. Maritorena, N.B. Nelson, D.A. Hansell & M. Lorenzi-Kayser. 2002. Global Distribution and Dynamics of Colored Dissolved and Detrital Organic Materials. *Journal of Geophysical Research*. **107**(C12): 10.1029/2001JC000965.

Siegel, D.A., S. Maritorena, N.B. Nelson, 2003: Independence and interdependencies of ocean optical properties viewed using SeaWiFS. In preparation for *Journal of Geophysical Research*.

This Research was Supported by the NASA Contract # 00196

Publications

Behrenfeld, M.J., E. Boss, D.A. Siegel and D.M. Shea, 2003: Global remote sensing of phytoplankton physiology. Submitted to *Science*.

Chomko, R.M., H.R. Gordon, S. Maritorena and D.A. Siegel, 2003: Simultaneous determination of oceanic and atmospheric parameters for ocean color imagery by spectral optimization: a validation. *Remote Sensing of the Environment*, **84**, 208-220.

Maritorena S., D.A. Siegel & A. Peterson. 2002: Optimization of a Semi-Analytical Ocean Color Model for Global Scale Applications. *Applied Optics*. 41(15): 2705-2714.

Maritorena, S., D.A. Siegel and D.B. Court, 2003: Ocean Color Data Merging Using a Semi-Analytical Algorithm. In preparation for Remote sensing of Environment.

Morel A. and S. Maritorena. 2001: Bio-optical properties of oceanic waters: a reappraisal. *J. Geophys. Res.*, 106(C4) : 7163-7180.

Siegel, D.A., S. Maritorena, N. B. Nelson, D.A. Hansell and M. Lorenzi-Kayser, 2002: Global ocean distribution and dynamics of colored dissolved and detrital organic materials. *Journal of Geophysical Research*, **107**, 3228, DOI: 10.1029/2001JC000965.

Siegel, D.A., S. Maritorena, N.B. Nelson, 2003: Independence and interdependencies of ocean optical properties viewed using SeaWiFS. For *Journal of Geophysical Research*.