Remote Sensing of Environment 114 (2010) 1791–1804

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues

Stéphane Maritorena [a,*], Odile Hembise Fanton d'Andon [b], Antoine Mangin [b], David A. Siegel [a,c]

[a] Institute for Computational Earth System Science, University of California, Santa Barbara, Santa Barbara, CA 93106-3060, USA
[b] ACRI-ST, 260, route du Pin Montard — B.P. 234, 06904 Sophia Antipolis Cedex, France
[c] Department of Geography, University of California, Santa Barbara, Santa Barbara, CA 93106-4060, USA

ARTICLE INFO

Article history:
Received 28 August 2009
Received in revised form 2 February 2010
Accepted 3 April 2010

Keywords:
Ocean color
Merged data
SeaWiFS
MODIS-AQUA
MERIS
Semi-analytical model
Uncertainties
Time-series
Climate Data Records

ABSTRACT

The characteristics and benefits of ocean color merged data sets created using a semi-analytical model and the normalized water-leaving radiance observations from the SeaWiFS, MODIS-AQUA and MERIS ocean color missions are presented. Merged data products are coalesced from multiple mission observations into a single data product with better spatial and temporal coverage than the individual missions. Using the data from SeaWiFS, MODIS-AQUA and MERIS for the 2002–2009 time period, the average daily coverage of a merged product is ~25% of the world ocean which is nearly twice that of any single mission's observations. The frequency at which a particular area is sampled from space is also greatly improved in merged data as some areas can be sampled as frequently as 64% of the time (in days). The merged data presented here are validated through matchup analyses and by comparing them to the data sets obtained from individual missions. Further, a complete error budget for the final merged data products was developed which accounts for uncertainty associated with input water-leaving radiances and provides uncertainty levels for the output products (i.e. the chlorophyll concentration, the combined dissolved and detrital absorption coefficient and the particulate backscattering coefficient). These merged products and their uncertainties at each pixel were developed within the NASA REASON/MEaSUREs and ESA GlobColour projects and are available to the scientific community. Our approach has many benefits for the creation of unified Climate Data Records from satellite ocean color observations.

1. Introduction

Remote sensing of the Earth system has been an invaluable tool to study and monitor the biosphere and its components (land, ocean or atmosphere) and over the years a wide variety of physical or biogeochemical variables have been collected by various Earth observing satellite sensors. Passive-microwave and thermal infrared (e.g. AVHRR) satellite data are collected over the oceans since about 3 decades to monitor sea surface temperature (see e.g. Reynolds et al., 2007; Parkinson & Cavalieri, 2008) while others like altimetry or ocean color are more recent and observations have been made on a global scale for more than a decade (e.g. Merrifield et al., 2008; McClain, 2009). For each of these variables, multiple sensors with similar characteristics have contributed to the development of the satellite time-series through a succession of missions designed to provide one or more specific products. The planning and timing of satellite missions are generally designed to allow some overlap between successive generations of a sensor, which permits sensor intercomparisons and allows biases across missions to be quantified and accounted for (e.g. Castro et al., 2008).

Presently, there are several global satellite sensors orbiting the Earth and sampling the color of the oceans (e.g. GeoEye's SeaWiFS, NASA's MODIS on the TERRA and AQUA platforms and ESA's MERIS on Envisat). It seems logical to attempt to merge the data sets from these missions to create unified data products. This would result in sets of single synthetic products with greater coverage of the world's oceans on shorter time scales instead of several, possibly divergent, data sets. If the data products are well characterized and consistent among missions, their unification as a single synthetic product is a way to build a univocal time-series of a particular product over the lifetimes of several missions. It is also a way to develop consistent Essential Climate Variables time-series (GCOS, 2003) or Climate Data Records (NRC, 2004; Loeb et al., 2009; Siegel et al., in preparation) and international planning for a virtual constellation of ocean color satellite missions which would satisfy these needs has recently begun (e.g., IOCCG newsletter Sept. 2008, http://www.ioccg.org/news/Sept2008/news.html). Such a unification can be conducted by merging the data from the different sensors.

* Corresponding author.
E-mail address: stephane@icess.ucsb.edu (S. Maritorena).

0034-4257/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2010.04.002

In the case of ocean color, several data merging efforts have taken place in recent years (Gregg and Conkright, 2001; Kwiatkowska & Fargion, 2002; Maritorena & Siegel, 2005; Pottier et al., 2006; IOCCG, 2007) under the NASA SIMBIOS and REASoN projects or the ESA GlobColour program. This merger can be performed on final data products, literally compositing observations from different missions into a single data set, or by intermixing observations of the spectral water-leaving radiance, NLw(λ), from different sources in a bio-optical model to derive merged products (e.g., IOCCG, 2007). Both approaches have been used and merged data sets are now available to the scientific community. For example, the NASA Ocean Biology Processing Group (OBPG) distributes a merged Chlorophyll (CHL) product from SeaWiFS and MODIS-AQUA CHL data while fields of CHL and Inherent Optical Property (IOP) products from the inversion of merged NLw(λ) of SeaWiFS, MODIS-AQUA and MERIS are available through the UCSB-NASA Ocean Color MEaSUREs (formerly REASoN) and the ESA GlobColour projects (see Table 1 for URLs). This paper will focus on the latter approach where combined NLw(λ) spectra from different sensors are used in a bio-optical model (Maritorena & Siegel, 2005) to simultaneously derive several "merged" ocean color products as well as uncertainty determinations for each of them. Maritorena and Siegel (2005) presented only a demonstration of this approach using a limited number of scenes (18) from SeaWiFS and MODIS-TERRA. Since then, data from other ocean color sensors (MODIS-AQUA and MERIS) have become available while issues with the MODIS-TERRA data made that sensor unreliable for climate-scale ocean color studies (Franz et al., 2008).

Here, the model-based approach for ocean color data merging described in Maritorena and Siegel (2005) is applied to the SeaWiFS, MODIS-AQUA (hereafter referred to as "AQUA") and MERIS NLw(λ) data to create long time-series of merged data produced under the NASA Ocean Color MEaSUREs and the ESA GlobColour projects. In the present paper, some of the characteristics and benefits of these merged data sets are documented for the 2002–2009 period for which the three sensors are simultaneously operational. We provide coverage estimates to demonstrate the benefits of merging and present a validation of the final merged products. An error budget is also presented where uncertainties in the input NLw(λ) from each sensor and model errors are quantified and taken into account to generate uncertainty estimates for each output product and at each pixel of an image. Outstanding issues associated with sensors data or inter-sensor biases are also discussed as well as how data merging could be the path toward building consistent Climate Data Records.

2. Methods

The bio-optical model-based merging procedure and its associated features are described in Maritorena et al. (2002) and Maritorena and Siegel (2005) and its main traits are briefly summarized below. One of the major features of the approach is that it combines the normalized water-leaving radiances from different sensor data sets. Over each particular pixel of a geographical grid common to SeaWiFS, AQUA and MERIS, the NLw(λ) spectra from the available sensors at that pixel are selected and combined in a single, multi-source, spectrum which is then used in the inversion of the GSM01 semi-analytical ocean color model (Maritorena et al., 2002). The model inversion results in the simultaneous retrieval of multiple ocean color related variables, namely the sub-surface chlorophyll-a concentration, CHL, the combined absorption coefficient of the particulate and dissolved organic material at 443 nm, CDM, and the particulate backscattering coefficient at 443 nm, BBP, as well as uncertainty levels and covariance matrices of the retrievals. Only the bands in the visible are used in the inversion and, depending on which sensors collected data for a particular bin, the NLw(λ) spectrum that enters the model is made from 6 to 19 spectral data points. When more than one sensor is available, the resulting radiance spectrum will consist of a mix of spectral bands that are either unique, coming from only one sensor, or replicated for bands that are common to several sensors (Fig. 1). In other words, the resulting NLw(λ) combined spectrum has improved spectral resolution and contains replicated measurements for some bands as illustrated in Fig. 1. While both the MEaSUREs and GlobColour projects use the GSM01 model to generate merged products, there are some differences in data processing and data handling which result in differences between the two sets of products, mostly in eutrophic waters and coastal areas. The differences between the MEaSUREs and GlobColour projects are summarized in Table 1.

Merged products were generated using the GSM01 model from SeaWiFS (version 5.2), AQUA(v1.1) and MERIS (v2.0/Q) daily NLw(λ) fields for the time period common to the three sensors (July 2002–2009). The complete archives of SeaWiFS, AQUA and MERIS were also individually processed using the GSM01 model in order to compare merged and individual mission data sets. The three main variables (CHL, CDM and BBP) are produced along with their uncertainty estimates. In addition, metrics of the quality of the retrievals, like the χ^2, the covariance matrix of the retrievals and residuals (spectral differences between observed and modeled radiances re-built from the retrieved variables) are also available. The products are also time-

Table 1
Differences between the MEaSUREs and GlobColour data processing using the GSM01 model. For GlobColour, the data are processed from level-2, 1 km resolution NLw(λ) data that include the red bands and output products are at a 4.6 km resolution. In the MEaSUREs project the MODIS-AQUA and MERIS level-3 binned NLw(λ) data are first converted to 9 km data to match the resolution of SeaWiFS which is also the resolution of the output products. Note that the MEaSUREs project initially merged the SeaWiFS and AQUA data only but it now includes the MERIS data. The NASA Ocean Biology Processing Group CHL merged product is available at http://oceancolor.gsfc.nasa.gov/cgi/l3. Some of the NASA OBPG and UCSB GSM products are also available through the NASA Giovanni system: http://reason.gsfc.nasa.gov/Giovanni/.

	MEaSUREs	GlobColour
Sensors	SeaWiFS MODIS-AQUA (MERIS)	SeaWiFS MODIS-AQUA MERIS
Input data	Level 3 binned	Level 2
Bands used	SeaWiFS: 412, 443, 490, 510, 555 AQUA: 412, 443, 488, 531, 551 MERIS: 412, 443, 490, 510, 560	SeaWiFS: 412, 443, 490, 510, 555, 670 AQUA: 412, 443, 488, 531, 551, 667 MERIS: 412, 443, 490, 510, 560, 620, 665
Sensor specific weighting of NLw(λ)	None	Derived from matchup statistics (see text)
Spatial resolution of output products	9 km	4.6 km
Products uncertainties and QC products	Confidence intervals (linear approximation of nonlinear regression inference region; Bates & Watts,1988)	Standard deviation (diagonal elements of the retrievals covariance matrix). Available on request: χ^2, residuals, covariance matrices
Data file format	HDF	NetCDF
Data access	http://wiki.icess.ucsb.edu/measures/	http://www.globcolour.info/ http://hermes.acri.fr/

SeaWiFS	412	443	490	510			555		670
AQUA	412	443	488		531	547			667
MERIS	412	443	490	510			560	620	665
MERGED	412	443	490	510	531	547	557	620	667
	(3)	(3)	(3)	(2)	(1)	(1)	(2)	(1)	(3)

Fig. 1. Bands of the individual sensor used in the merging and resulting combination. Numbers in parentheses indicate the number of times a band is repeated in the concatenated NLw(λ) array.

composited in 4-day, weekly (8-day) and monthly data. These products are available as global mapped or binned files (see Table 1 for details and how to access to the merged and individual mission data products).

3. Results

3.1. Example of data products

Fig. 2 shows some examples of daily imagery for NLw(443) from SeaWiFS (Fig. 2A), AQUA (Fig. 2B) and MERIS (Fig. 2C) as well as for the daily GSM CHL merged products (Fig. 2D) for August 18, 2005. Fig. 2A, B and C illustrates the differences in daily sampling between the three sensors in terms of sun glint and their swaths width and orientation. SeaWiFS provides the best coverage mostly because of its ability to tilt to avoid sun glint and because of its large swath. Data gaps caused by sun glint are visible inside the AQUA swaths (Fig. 2B) while another noticeable characteristic is MERIS's narrow swath (Fig. 2C) relative to SeaWiFS or AQUA. The CHL image (Fig. 2D) that results from the merging of the NLw(λ) data from the three sensors using GSM01 shows the obvious improvement in daily coverage compared to the individual missions.

For daily data, the combinations of data sources can be mapped for each bin (Fig. 3A, for August 18, 2009) to show which sensor or assemblage of sensors were available over each bin, globally, on a particular date. For daily images, a majority of the bins comes from individual sensors as, on average over the 2002–2009 period, 31% (standard deviation: 4.6%) of the valid bins of a product come from SeaWiFS only data, 22% (standard deviation: 4.5%) come from AQUA only data and 17% (standard deviation: 3.7%) come from MERIS only data. The SeaWiFS–AQUA combination makes about 17% (standard deviation: 2.9%) of a product daily bins while each of the other combinations (SeaWiFS–MERIS, MERIS–AQUA and SeaWiFS–AQUA–MERIS) represent less than 5.1% of the bins on average.

The multi-day imagery (4-day, 8-day, and monthly) for CHL and the other two daily GSM products (CDM and BBP) are also presented as examples (Fig. 3B to F). The composite imagery shows the significant gain in coverage with multi-day merged data (see below for more details). The relationships between the daily CHL, CDM and BBP products generated by the GSM01 model and comparisons with the SeaWiFS OC4v4 (O'Reilly et al., 2000) CHL are described in Siegel et al., 2005a,b. Typically, the CHL and CDM merged products values span 2 or 3 orders of magnitudes and they generally do not show discontinuities between areas with different data sources. The range of BBP values is more constrained (1.5 order of magnitude) and the BBP imagery is more affected by noise than the other products as will be discussed later in the manuscript.

3.2. Spatial coverage

Increased daily coverage of the world ocean is the most obvious benefit of the merging of data from multiple sensors. The average daily global coverage by SeaWiFS, AQUA and MERIS and for their various possible combinations for the 2002–2009 period is presented in Fig. 4 where the 100% ocean coverage is determined from the sum of the deep and shallow water masks in the SeaDAS software (http://oceancolor.gsfc.nasa.gov/seadas/). Each single sensor covers, on average, from about 8% to ~15% of the world ocean daily (MERIS: 7.77% with a standard deviation of 1.05%; AQUA: 11.81% with a standard deviation of 1.03%; SeaWiFS: 14.58% with a standard deviation of 1.03%). When sensors are combined, daily coverage increases from about 17% (AQUA–MERIS) to ~25% (SeaWiFS–AQUA–MERIS) (right side of Table 2). Coverage shows some seasonal variation with peaks in the boreal spring (March–May) and early fall (Sept–Oct) and lower values in summer (June–July) and winter (Dec–

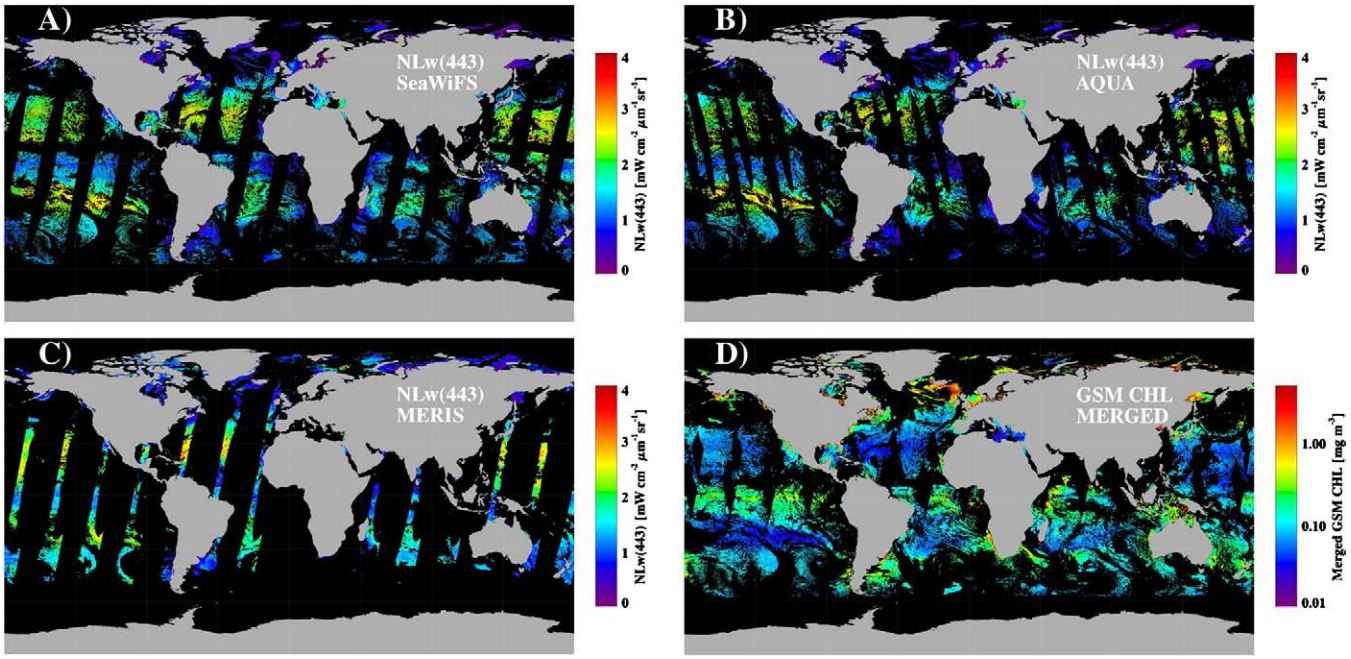

Fig. 2. Daily maps (August 18, 2005) of the NLw(443) from A) SeaWiFS, B) MODIS-AQUA and C) MERIS. D) Daily map of the merged CHL derived after the inversion in the GSM01 model of the combined NLw(λ) from the three sensors.

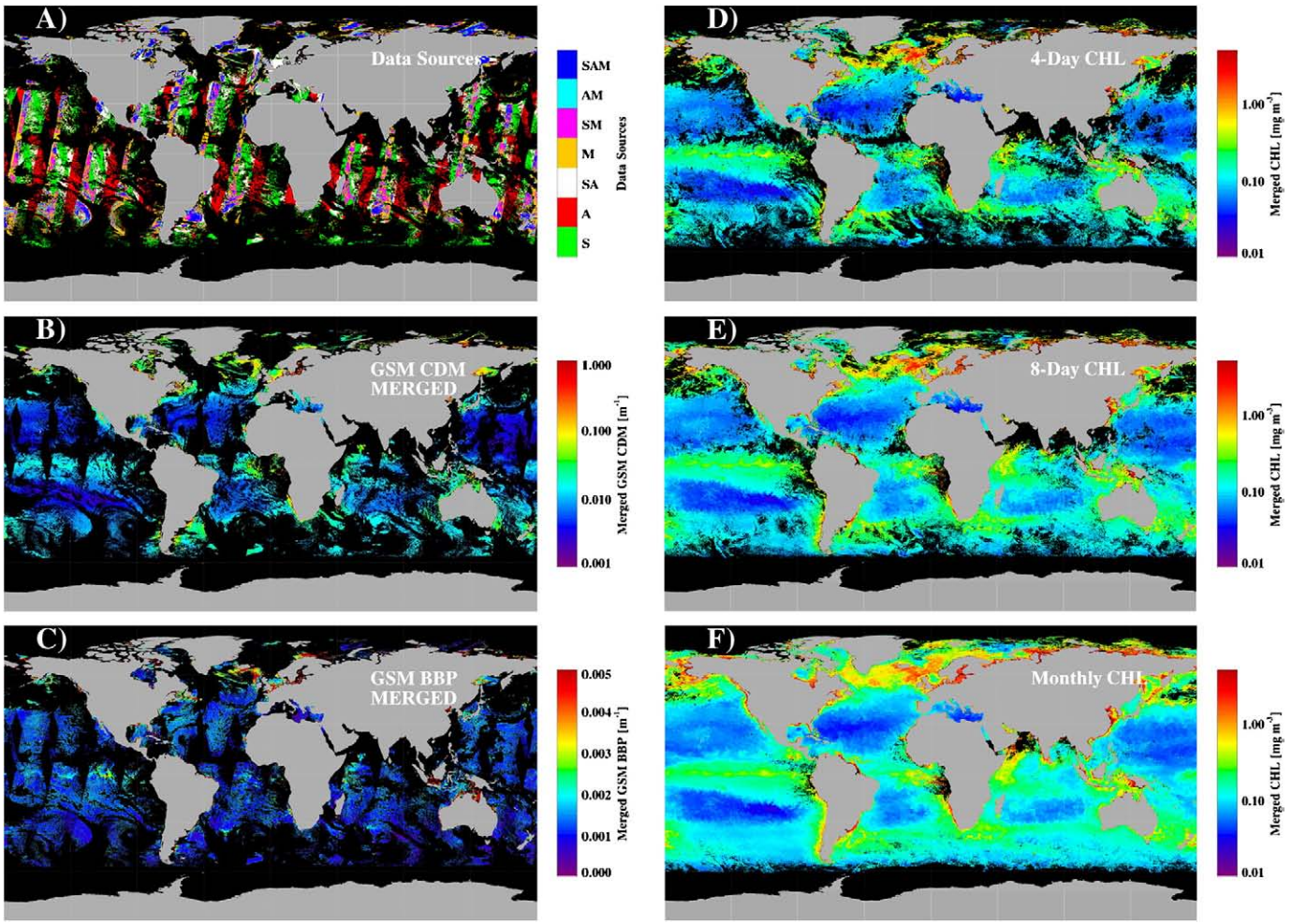

Fig. 3. A) Map of the data sources used to develop the merged image. Data sources are indicated by the first letter of the sensors (S: SeaWiFS, A: AQUA, M: MERIS) or by a combination of letters when more than one source is present (SA: SeaWiFS + AQUA, SM: SeaWiFS + MERIS, AM: AQUA + MERIS and SAM: SeaWiFS + AQUA + MERIS, B) daily merged CDM product, C) daily BBP product, D) 4-day merged CHL composite, E) 8-day merged CHL composite and, F) monthly merged CHL composite. Data are for August 18, 2005.

Jan). The improvement in coverage is most beneficial for daily data as the difference in coverage between individual sensors and merged data decreases with composite imagery (Fig. 5, left side of Table 2). For daily data, the improvement in coverage between a single sensor and the merged data (calculated as 100*(nbins_merged − nbins_sensor)/nbins_sensor) is 72.9%, 113.5% and 224.6% for SeaWiFS, AQUA and MERIS, respectively. These numbers drop to 5.7%, 10.4% and 19.5% for monthly composite imagery. The combination of sensors improves coverage significantly even for composite imagery over long periods of time but the biggest improvements are clearly seen for daily and

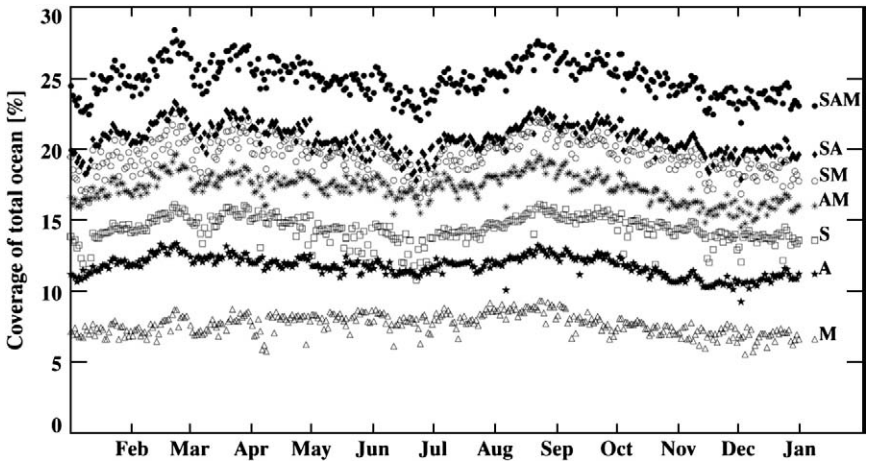

Fig. 4. Average daily global coverage for the 2002–2009 period for SeaWiFS (white squares), AQUA (stars) and MERIS (triangles) and the four possible combinations of these sensors. As in Fig. 3a, the combinations of sensors are identified by the letters associated with each individual sensor.

Table 2
Improvement in global coverage in merged data compared to individual missions (2002–2009).

	Improvement (%) over single mission			Average percentage cover of the ocean (standard deviation)			
	SeaWiFS	MODIS-AQUA	MERIS	SeaWiFS	MODIS-AQUA	MERIS	MERGED
Daily	72.98	113.55	224.58	14.58 (1.03)	11.81 (1.03)	7.77 (1.05)	25.22 (1.88)
4-Day	36.90	68.03	126.27	41.62 (4.67)	33.91 (2.97)	25.18 (4.97)	56.98 (6.24)
8-Day	22.76	41.32	80.23	59.06 (6.72)	51.31 (4.52)	40.23 (6.90)	72.50 (7.59)
Monthly	5.71	10.39	19.55	82.95 (9.79)	79.43 (4.75)	73.34 (9.47)	87.69 (8.87)

composite imagery over a few days. Beside surface area coverage improvement, another important aspect of coverage is the frequency at which a particular area is sampled with data merging relative to individual missions.

3.3. Temporal coverage (=sampling frequency)

The combination of data from multiple sensors also allows for improved temporal coverage because it increases the frequency at which a particular bin can be observed. The global sampling frequency for the individual sensors and when they are merged is illustrated in Fig. 6. The frequency of coverage (expressed in number of days of valid observations over the period July 2002–April 2009) is influenced by the orbital and field-of-view characteristics of each sensor as well as by clouds, sun glint and atmospheric correction failures (caused e.g. by dust or absorbing aerosols). The general distribution of the frequency of coverage is similar for the 3 sensors and shows a strong asymmetry between the northern and southern hemispheres and some differences between the different oceans. The central Pacific Ocean south of the equator, the western part of the tropical South Atlantic and Indian, the Mediterranean Sea, central Atlantic, the West coasts of North and Central America, South Africa, Madagascar and Australia are the areas with most frequent coverage. Conversely, high latitudes (50° and up), the eastern part of the north equatorial Pacific, the tropical eastern Atlantic and off the Peru–North Chile coast are much less often sampled.

For the July 2002 to April 2009 period, the median value of the coverage frequency by SeaWiFS is 14.45% which is equivalent to ~53 days/year on average. One fourth of the bins get data more than 19% of the time (~69 days/year) and another 25% gets data less than 8% of the time (29 days/year). At best, a bin is sampled 45.37% of the time (~158 days/year) when using SeaWiFS data alone. For AQUA, the median sampling frequency is 10.94% of the time (~40 days/year), 25% of the bins get data less than 6.4% of the time (~23 days/year) and the 25% of the bins that are the most frequently sampled contains data at least 16.25% of the time (~59 days/year) and the area with the best sampling gets data 47.84% of the time (~175 days/year). AQUA achieves a better maximum coverage value than SeaWiFS mostly because SeaWiFS had several relatively long data gaps since 2008 and also because the transformation of the AQUA data to the 9 km resolution improves its overall coverage slightly (because the 9 km bins can sometimes be made out of less than four 4.5 km bins). For MERIS, the median sampling frequency value is 7.08% (~26 days/year), 25% of the bins get data less than 4.22% of the time (~15 days/year) and 25% of the bins are sampled more often than 9.73% of the time (~36 days/year). Maximum sampling frequency for MERIS is 23.46% of the time (~85 days/year). The overall lower numbers for MERIS are mostly a consequence of its narrower swath. When the data of the 3 sensors are merged, the median value is 22.45% (~82 days/year) and the low and high 25% of the bins limits are 12.72% (~46 days/year) and 30.95% (~113 days/year), respectively. For the merged data, the maximum frequency of sampling is 64.06% of the time (233 days/year). This illustrates the huge benefits in spatial and temporal coverage when using merged data. It also emphasizes the fact that in most areas of the 9 km grid, composite imagery from any particular mission is generally made out of very few data, even with monthly composites.

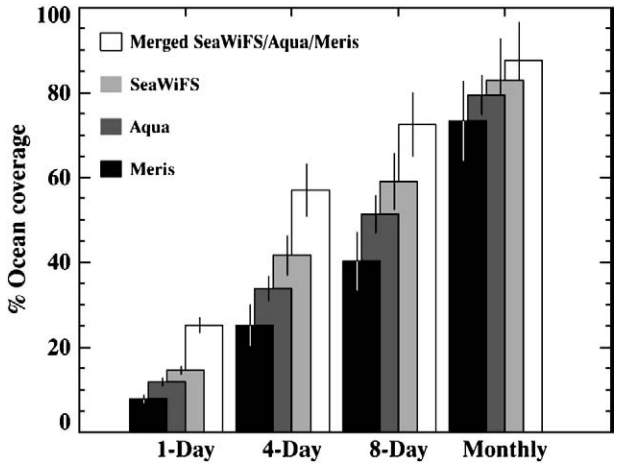

Fig. 5. Average and standard deviation of coverage for daily and multiple-day composite imagery for SeaWiFS, AQUA and MERIS and when the three sensors are merged for the 2002–2009 period.

SeaWiFS is the most frequent contributor to the total number of days of data available in each individual bin. Typically, SeaWiFS consistently contributes 40% to 50% of the time (Fig. 7) and this contribution is relatively homogeneous in the 45°N–45°S region. SeaWiFS contribution drops below 40% at high latitudes and on the Eastern side of South America, South Africa and Australia. AQUA contributes 30–40% of the time relatively uniformly globally except in the equatorial region where its contribution falls to 20–30% because of loss of data caused by sun glint. AQUA is also an important contributor in areas where SeaWiFS is a lesser contributor: West South America and West Southern Africa. Because of its narrower swath, MERIS contributes less frequently than the other two sensors. Typically, MERIS collects data 20 to 30% of the time except in the gyres and in the South hemisphere between the subtropical and Antarctic convergence zones where its contribution drops below 20%. Interestingly, MERIS is a stronger contributor at high latitudes and in the equatorial region where it compensates well the relative lack of AQUA data. These differences and complementarities in coverage result from the sensors' technical and orbital characteristics but in the case of MERIS, some of the differences are also likely to come from differences in data processing (e.g. cloud masking and atmospheric correction).

3.4. Error budget and product uncertainties

The model-based approach for data merging described here also enables the assessment of uncertainties associated with the different components of the procedure. The non linear fitting method in the GSM01 model searches the set of retrievals that best minimizes the mean square difference between the modeled and measured $NLw(\lambda)$. The fitting technique can also account for uncertainties in the measured $NLw(\lambda)$ and in the model by using them as weights in

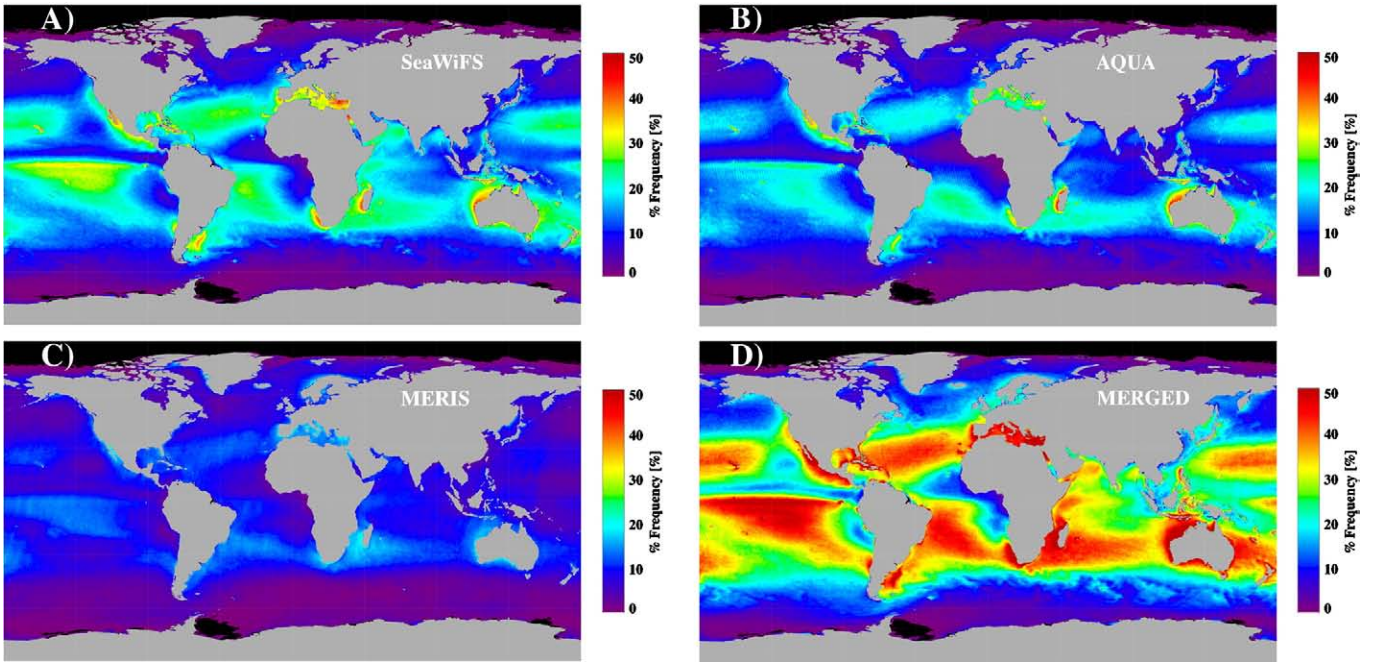

Fig. 6. Geographic distribution of the frequency of coverage (in % of total number of valid days) for the 2002–2009 period for A) SeaWiFS, B) MERIS, C) AQUA and D) when the three sensors are merged.

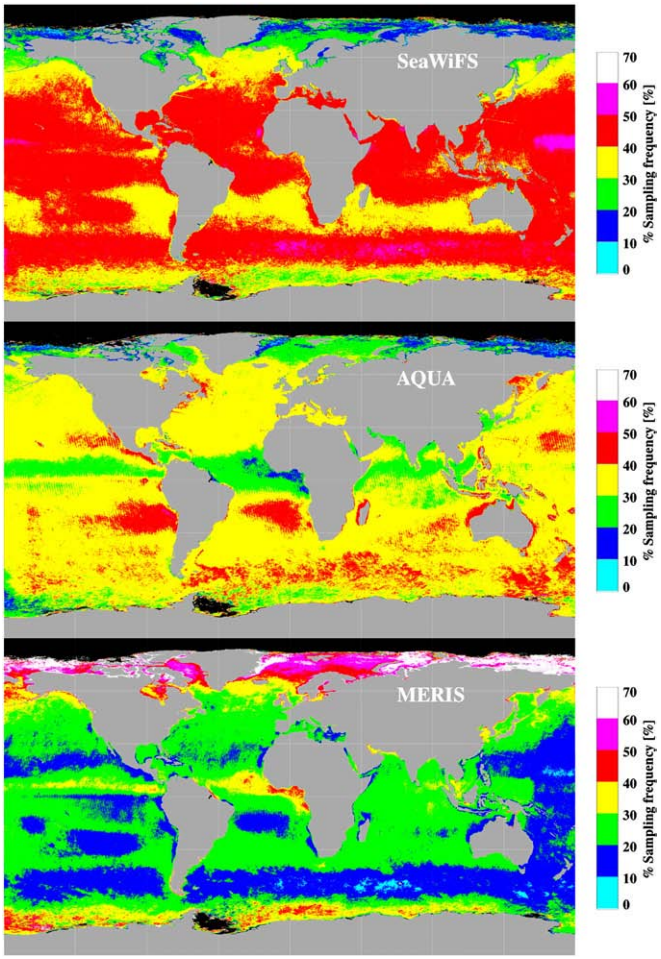

Fig. 7. Geographic distribution of the frequency at which each particular sensor contributes to the imagery (2002–2009).

the fitting procedure. When the inversion is performed using several sensors (N_{sensor}), each with several wavebands ($N_{\lambda j}$), the normalized least-squares minimization function (χ^2) can be expressed as

$$\chi^2 = \frac{1}{(Nb\_obs - Nb\_param)} \sum_{j=1}^{N_{sensor}} \sum_{i=1}^{N_{\lambda j}} \frac{(NLw(\mathrm{gsm},i,j) - NLw(\mathrm{sensor},i,j))^2}{\sigma^2_{\mathrm{sensor}}(i,j) + \sigma^2_{\mathrm{gsm}}(i,j)} \quad (1)$$

where Nb_obs is the total number of combined bands, Nb_param is the number of retrieved parameters, NLw(gsm, i, j) is the modeled normalized water-leaving radiance for band i of sensor j, NLw(sensor, i, j) is the measured normalized water-leaving radiance in band i of sensor j, $\sigma_{\mathrm{sensor}}(i,j)$ is the uncertainty associated with the NLw of band i of sensor j and $\sigma_{\mathrm{gsm}}(i,j)$ is the uncertainty due to the GSM model for band i of sensor j. Note that the sum of $\sigma^2_{\mathrm{sensor}}(i,j)$ and $\sigma^2_{\mathrm{gsm}}(i,j)$ corresponds to the diagonal terms of the full covariance matrix of errors (i.e. the sum of covariances of the measured and modeled radiances).

Assuming the errors on the input NLw(λ), σ_{sensor}, and from the bio-optical model, σ_{gsm}, can be reliably characterized and that the errors in the satellite observations are unbiased, it is then possible to propagate this information through the retrieval procedure and to derive uncertainty estimates for each of the output products on a pixel-by-pixel basis. It should be mentioned that possible (but unknown) systematic errors in the satellite data or in the model are not taken into account in the present Gaussian approach. The present work should be viewed as a practical, first attempt at deriving a reasonable error budget for ocean color data and products.

3.4.1. Uncertainties of the satellite water-leaving radiance spectra, σ_{sensor}

Statistical figures from comparisons between satellite and in situ NLw(λ) observations were used to determine the expected level of error for each band. Using the original NOMAD data set (Werdell & Bailey, 2005) plus data from the NASA SeaBASS archive (Werdell & Bailey, 2002), the BOUSSOLE buoy (Antoine et al., 2008b) and above-water radiometric measurements (Hooker et al., 2004; Zibordi et al.,

2006b,c), the normalized water-leaving radiances from each of the sensors were compared to coincident in situ data for each of the bands in the visible through regression statistics. These comparisons are presented in Table 3. Overall, the agreement between the sensors' measurements and in situ data is good as indicated by the high R^2 or slope values. The agreement is generally slightly poorer for the blue bands (412 and 443 nm) as these bands can sometimes be affected by imperfect atmospheric correction. The 510 nm band in MERIS also shows some discrepancies compared to in situ data. The spectral shapes of these NLw(λ) uncertainties are similar for the 3 sensors and show a general decrease of the RMS from the blue to the red bands. For the inversion, the uncertainty level of each band i of each sensor j, $\sigma_{\text{sensor}}(i,j)$, is arbitrarily set to half the value of the RMS shown in Table 3 (Fanton d'Andon et al., 2008; Mélin et al., 2009). It should be noted that this only impacts the products uncertainty estimates derived during the inversion but has no influence on the values of the product retrievals.

3.4.2. Model errors, $\sigma_{gsm}(i,j)$

To evaluate the errors caused by the GSM01 model the input radiance spectra from the NOMADv1 data set were compared to the GSM01 radiance spectra that correspond to the best fit obtained during the inversion (or, in other words, the reconstructed radiance spectra when the model is used in the forward mode and produces a radiance spectrum given a set of Chl, CDM and BBP values). This was performed in two steps. At first, inversions are performed without any weighting of the NLw(λ) data. Then, the RMS/2 values from the comparison of the measured and modeled NLw(λ) spectra are used to weight the NLw(λ) in a new set of inversions which will produce the final estimates of the model error. Additional sets of inversions are not required as the spectral RMS errors remain stable. The statistical comparisons between the input in situ NLw(λ) and the GSM spectra reconstructed from the retrievals of the inversion (forward mode) are presented in Table 4. Like for the NLw(λ) errors, the model error ($\sigma_{\text{gsm},i}$) for each band is set to the corresponding RMS/2 value.

3.4.3. Uncertainties of the output products, $\sigma_{prod}(x)$

Currently ocean color satellite products from any space agency are distributed without estimates of their uncertainties. While matchup analyses can provide an instantaneous global estimate of the level of error one can expect on a particular product, they do not say much about the spatial and temporal distribution of the errors. A pixel-by-pixel uncertainty estimate for any particular satellite-derived product is very important for modelers or in data assimilation. It is also a requirement of most space agencies but these uncertainty estimates are seldom available. With model-based approaches, it is possible to evaluate the uncertainties of the retrieved products. When the GSM01 model was developed, the confidence interval of the retrievals was estimated by linear approximation of nonlinear regression inference region (Bates & Watts, 1988). Similarly when a Levenberg–Marquardt technique is used for the least-squares minimization, the square-root of diagonal elements of the covariance matrix provide the standard deviation associated with each of the retrieved parameters (Press et al., 1992) as is done in the GlobColour project (see also Van Der Woerd & Pasterkamp, 2008 or Wang et al., 2005 for an alternative approach). The calculation of the χ^2 or the residuals between observed and GSM01 modeled NLw(λ) are metrics available to assess the quality of the inversion and of the error budget.

It is possible to assess how accurate some components of our error budget are by comparing modeled values to actual in situ measurements. For example, using the satellite radiance data, NLw(sensor, i,j), that have a coincident in situ measurement in the NOMADv2 data set, it is possible to compare NLw(sensor, i,j) to the spectra reconstructed by GSM01 from the retrievals, NLw(gsm, i,j). The mean and standard deviation of the residuals (with residuals = NLw(sensor, i, j) − GSM reconstructed NLw(gsm, i, j)) between the input and modeled NLw spectra are presented in Fig. 8 for each of the sensors used in the present study. The range of spectral values of the NLw(sensor, i, j) input uncertainties (Table 3) are also plotted for each sensor (white squares). The figure shows that the input uncertainties on NLw(λ) are consistent with the residuals which is a desirable feature and validates, to a large degree, the NLw(λ) input uncertainties analysis and the use of RMS/2 to quantify the uncertainties in the radiances and model. A bias is however observed at 443 nm for the three sensors and, to a lesser extent, for SeaWiFS data at 490 nm. This consistent bias among the three sensors suggests a possible weakness in the model at 443 nm.

Using the NOMADv2 in situ data set, it is also possible to assess whether or not the uncertainty estimates for the output products are accurate. This can be done by comparing the estimated uncertainties of product x, $\sigma_{\text{prod}}(x)$, from the model to the actual errors between the model retrieved variables and the in situ measurements. If for a particular product x the uncertainty estimates $\sigma_{\text{prod}}(x)$ are reliable (in the sense that they are truly representative of its standard deviation), the actual error should be lower than $\sigma_{\text{prod}}(x)$ in about 68% of the cases or, similarly, the distribution of true observed errors normalized by the estimated errors ((retrieved x − in situ x)/$\sigma_{\text{prod}}(x)$) should follow a standard

Table 3
Characterization of uncertainties on NLw(λ) for the three sensors from statistical analyses between satellite and in situ data.

Band	Sensor	N	Slope	Intercept	R^2	Mean ratio	Median ratio	Mean %diff	Median %diff	Bias	RMS	In_situ range		Satellite_range	
412	MERIS	210	0.766	0.375	0.475	1.688	1.275	77.220	31.742	0.206	0.376	0.025	2.278	0.13	2.501
443	MERIS	299	0.785	0.311	0.625	1.357	1.150	44.125	19.552	0.131	0.296	0.047	2.220	0.177	2.485
490	MERIS	204	0.799	0.216	0.774	1.159	1.098	24.668	14.831	0.061	0.183	0.108	2.454	0.207	2.254
510	MERIS	129	0.577	0.326	0.532	1.16	1.126	23.263	17.092	0.069	0.160	0.275	2.324	0.362	2.075
555	MERIS	314	0.766	0.119	0.877	1.027	0.972	20.146	16.734	−0.050	0.224	0.167	3.573	0.211	2.960
620	MERIS	8	1.123	0.031	0.649	1.767	1.534	76.719	53.431	0.038	0.041	0.032	0.085	0.061	0.126
670	MERIS	154	0.739	0.038	0.888	2.119	1.179	124.632	31.606	0.008	0.049	0.005	0.754	0.028	0.645
412	AQUA	363	0.664	0.125	0.525	0.908	0.836	30.577	22.191	−0.114	0.262	0.078	2.475	0.104	2.390
443	AQUA	530	0.768	0.142	0.663	1.001	0.958	21.923	15.575	−0.038	0.202	0.064	2.220	0.075	2.074
490	AQUA	297	0.857	0.088	0.794	0.986	0.978	14.927	10.730	−0.023	0.160	0.138	2.153	0.099	2.034
531	AQUA	41	0.771	0.080	0.774	0.962	0.958	21.589	16.604	−0.064	0.196	0.322	2.097	0.224	2.009
555	AQUA	550	0.902	0.044	0.888	1.006	0.973	14.785	9.437	−0.019	0.128	0.050	2.327	0.161	2.429
670	AQUA	327	0.896	0.001	0.792	1.271	0.938	54.890	30.059	−0.008	0.043	0.005	0.588	0.016	0.877
412	SeaWiFS	440	0.991	−0.020	0.835	1.008	0.929	29.843	19.914	−0.028	0.239	0.058	2.834	0.133	2.766
443	SeaWiFS	647	0.929	0.071	0.801	1.044	0.992	22.080	14.682	0.008	0.200	0.100	2.815	0.091	2.352
490	SeaWiFS	516	0.889	0.067	0.779	0.978	0.962	15.973	12.616	−0.029	0.168	0.176	2.330	0.163	2.272
510	SeaWiFS	377	0.87	0.063	0.787	0.982	0.964	15.272	12.191	−0.022	0.130	0.195	2.218	0.219	2.383
555	SeaWiFS	686	0.904	0.032	0.864	0.995	0.954	18.396	15.129	−0.024	0.143	0.167	2.492	0.172	2.486
670	SeaWiFS	210	0.845	0.010	0.878	1.365	0.941	65.316	32.437	−0.007	0.046	0.003	0.819	0.002	0.748

Table 4
Statistical comparison between the input in situ NLw(λ) and the GSM spectra reconstructed from the retrievals of the inversion (forward mode). For each band, the RMS/2 value is used as the model error.

Band	N	RMS	Bias	RMs rel	Bias rel	Mean ratio	Mean(%) diff	Median(%) diff	Min insitu	Max insitu	Min mod	Max mod	R^2	Intercept	Slope
412	260	0.015	−0.011	0.037	−0.015	0.969	3.089	1.179	0.075	2.420	0.055	2.366	0.995	−0.024	1.015
443	260	0.121	−0.079	0.063	−0.046	0.903	9.703	7.349	0.095	3.203	0.090	2.669	0.982	0.023	0.893
490	260	0.119	0.094	0.045	0.039	1.095	9.472	8.564	0.177	4.375	0.205	4.915	0.991	−0.089	1.168
510	260	0.078	−0.060	0.052	−0.039	0.918	9.015	9.726	0.218	4.437	0.237	4.375	0.980	−0.086	1.028
531	260	0.056	−0.042	0.048	−0.037	0.922	8.414	9.075	0.273	4.331	0.272	4.249	0.991	−0.067	1.030
555	260	0.044	−0.031	0.045	−0.031	0.935	7.567	6.169	0.202	4.202	0.191	4.092	0.991	−0.012	0.974

centered normal distribution. The estimated and actual errors for the three GSM retrievals using the NOMAD data set are compared in Fig. 9. These plots show a usually conservative overestimation of the GSM01 errors for CHL (upper left panel) and BBP (lower left panel) as the model error estimates are generally higher than the actual errors. The situation is different for CDM (upper right panel) where the modeled error is frequently lower than the actual error. This is confirmed by the probability density distribution of the reduced residuals (=(retrieved x − in situ x)/$\sigma_{\mathrm{prod}}(x)$; lower right plot) where CHL and BBP are very close to the expected normal distribution while CDM departs from it. The symmetrical distribution observed for the three GSM01 retrievals indicates that the inversion does not introduce a bias (or a very small one in the case of CDM).

An example of an uncertainty map for the merged CHL on February 2, 2009 is presented in Fig. 10A and uncertainty maps associated with each of the GSM products are available from the websites listed in Table 1. The merged CHL uncertainty map for February 2, 2009 shows a wide range of variation with complex and uneven patterns that generally depend on which sensor or combination of sensors were available at each particular bin and how close the modeled and satellite NLw(λ) spectra are (see Maritorena & Siegel, 2005). Uncertainties for the multi-day (composite) imagery can be calculated as

$$\sigma_{\mathrm{multidays}} = \sqrt{\frac{1}{\sum_{i=1}^{ndays} \frac{1}{\sigma_i^2}}} \quad (2)$$

under the assumption that the error is independent from one day to the next. Consequently the multi-days error estimates decrease as the number of days of data increases and the uncertainties for the composite imagery (Fig. 10B) are generally much smoother and lower than the daily values.

3.5. Matchups

Matchup analyses (Bailey & Werdell, 2006) are the most common means to assess and validate satellite ocean color data, although this approach may be questionable considering the scarce measurements available for some products. Matchup analyses consist in comparisons between satellite retrievals and in situ measurements collected in close spatial and temporal windows. For validation of single mission data, matchup analyses are generally performed with satellite data at the highest possible resolution in order to compare the pinpoint in situ measurement with a satellite retrieval that represents the smallest possible surface area. Typically, these matchup comparisons are performed using level-2 data at a resolution close to 1 km with a time constraint of having both the satellite and in situ data collected within 3 to 6 h from each other (Bailey & Werdell, 2006).

In the case of the data presented here, it is necessary to relax the spatial and temporal constraints to maximize the number of points available for comparisons (see discussion in IOCCG, 2007). Here, 9 km binned satellite GSM retrievals from the merging of the SeaWiFS, AQUA and MERIS data are compared to in situ measurements of CHL, CDM and BBP made out of the NOMAD data set (Werdell & Bailey, 2005) for the 1997–2003 time period and additional data from the SeaBASS archive (Werdell & Bailey, 2005) for the 2003–2007 period. In the present analysis, a matchup is considered valid when on a particular day, the latitude and longitude of an in situ measurement fall within the boundaries of a valid (i.e. non-zero) satellite 9 km bin (equal-area sinusoidal grid) for the same date. The satellite component of a matchups is limited to the bin identified by the date and its geographical coordinates (=center bin latitude and longitude) and, because of the coarse 9 km used here, it is not extended to surrounding bins as is usually done for higher resolution single mission matchups (Bailey & Werdell, 2006).

Matchups for the three GSM products derived from the merging of SeaWiFS, AQUA and MERIS data are presented in Fig. 11. For each matchup point, there are seven possible scenarios: the satellite

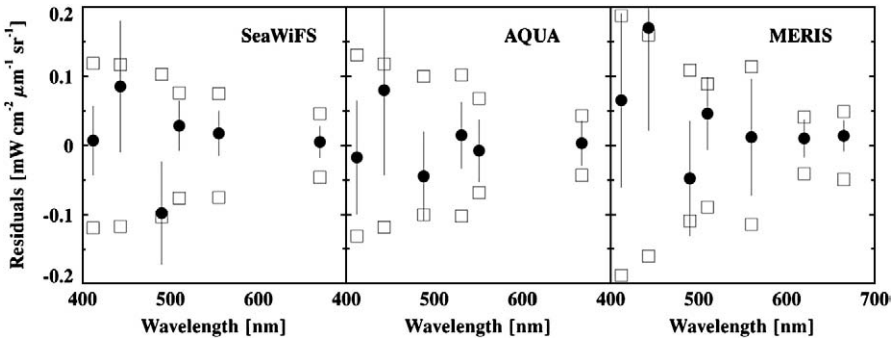

Fig. 8. Mean and standard deviation of the spectral residuals between the NLw(sensor, i, j) spectra and the "best fit" spectra from the inversion of the GSM01 model (residuals = NLw(sensor, i, j) − GSM reconstructed NLw(gsm, i, j)) for the SeaWiFS (left), MODIS-AQUA (center) and MERIS (right) sensors. The range of uncertainties of the input NLw(λ) as determined from the matchup analyses and model error analyses are also plotted (white squares). The NLw(sensor, i, j) data used in the figure have a coincident in situ measurement in the NOMADv2 data set.

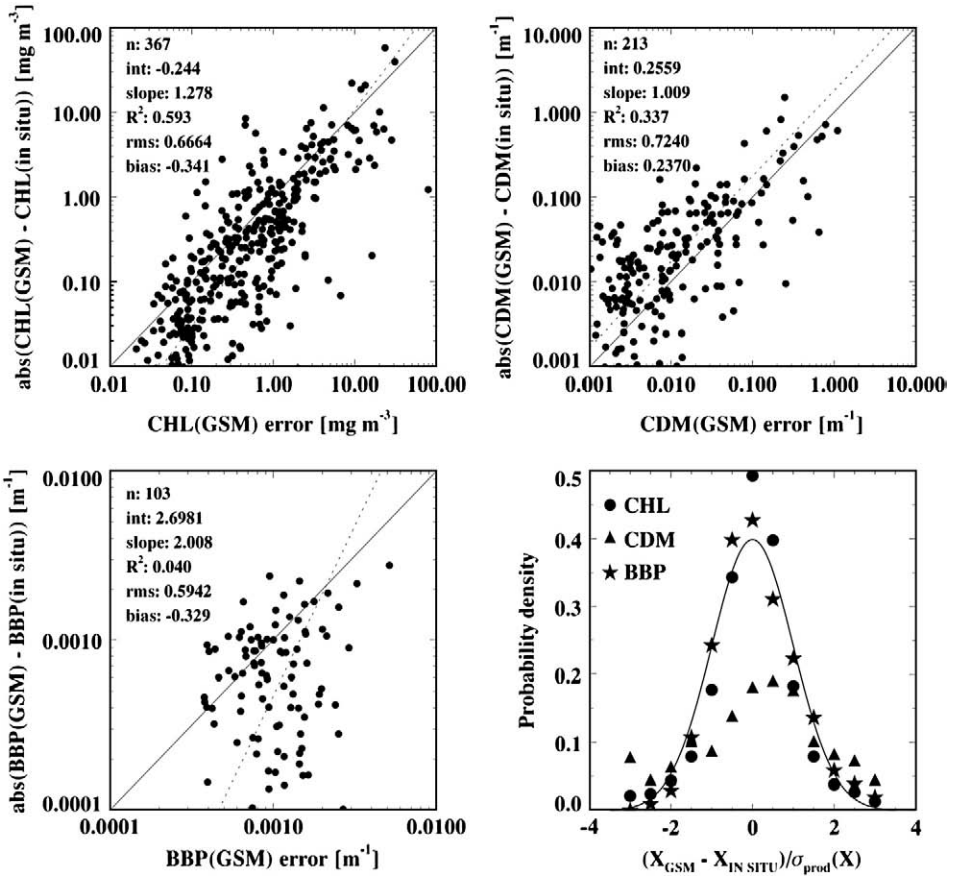

Fig. 9. Comparisons between the predicted and actual uncertainties in the GSM01 products (upper left: CHL; upper right: CDM; lower left: BBP). The predicted uncertainties $\sigma_{prod}(x)$ are outputs from the least-squares minimization on the radiance spectra (see text). Here, we used NLw(sensor, i, j) data that have a coincident in situ measurement in the NOMADv2 data set. If the predicted uncertainties are accurate, about 2/3 of the data points should be below the 1:1 line. The centered variables ((retrieved x − in situ x)/$\sigma_{prod}(x)$ − lower right panel) show a normal distribution for CHL (circles) and BBP (stars) while the CDM (triangles) distribution departs from normal (curve).

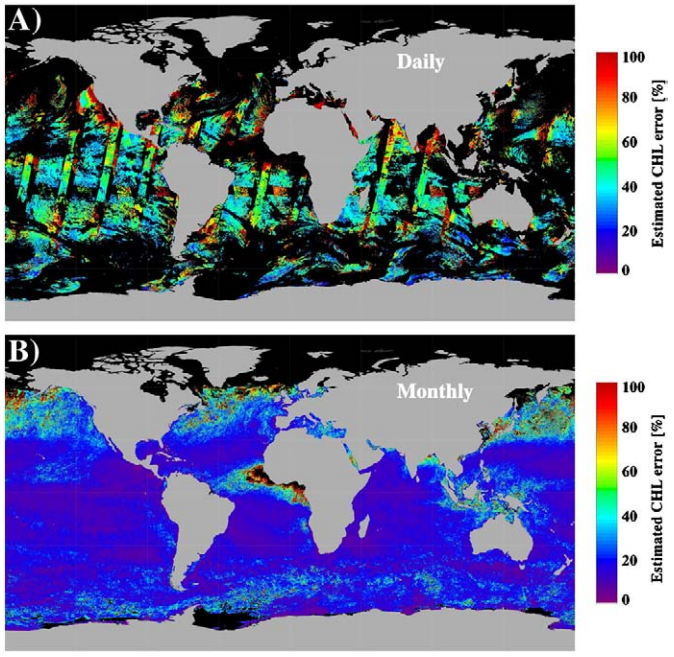

Fig. 10. Maps of the estimated error in the CHL product. A) Estimated CHL error (as a % of the retrieved CHL value) for a daily image (February 2, 2009) and, B) estimated CHL error for a monthly image (February 2009).

retrieval comes from the data of a single sensor, i.e. SeaWiFS only or AQUA only or MERIS only, or it can come from one of the four possible combinations of these sensors (SeaWiFS + AQUA, SeaWiFS + MERIS, AQUA + MERIS or SeaWiFS + AQUA + MERIS, see e.g. Fig. 4). In Fig. 11, each situation is illustrated by a particular symbol, consistent with previous figures, and the overall regression statistics between in situ and satellite estimates are also included (Type II regression on log10-transformed data). The regression statistics between the in situ and satellite-derived CHL estimates shows a good agreement between the two sets with the slope of the regression for the merged CHL close to one (0.945). However the GSM retrievals seem to underestimate in situ CHL for some stations in meso- and eutrophic waters (CHL>1 mg m^{-3}). In the matchups with the retrievals from the individual sensors only (i.e. not merged), the underestimation of CHL is more systematic with AQUA than with SeaWiFS or MERIS (not shown). This remains true when the NASA or ESA operational CHL algorithms are used as they show a systematic difference in CHL between AQUA and the other sensors (see also Siegel et al., in preparation; NASA OBPG at http://oceancolor.gsfc.nasa.gov/ANALYSIS/PROCTEST/ and discussion below). For the three sensors and for the merged data, the negative CHL bias is driven, to some extent, by some very coastal (e.g. West Florida Shelf, Monterey Bay, Coastal Benguela, and Chesapeake Bay) or high latitude (operantar cruise) stations for which the GSM model may not be well suited (Maritorena et al., 2002). An additional issue in these waters comes from frequently inaccurate atmospheric correction that often results in water-leaving radiance in the blue to be too low which in turn affects the quality of the GSM retrievals.

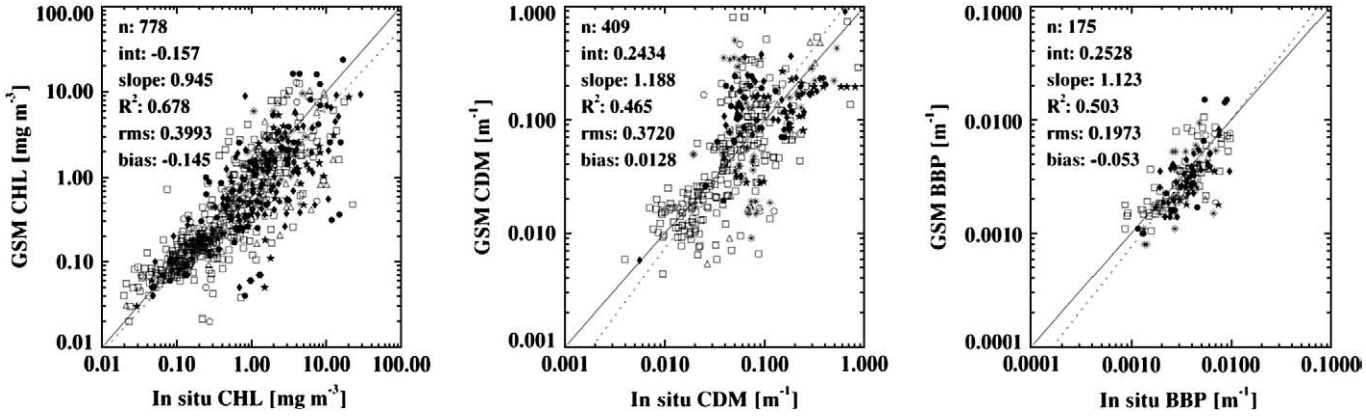

Fig. 11. Matchup statistics for the three GSM merged products, CHL (left), CDM (center) and BBP (right). The symbol of each matchup point indicates which satellite data sources were used for that point (the symbol coding is the same as in Fig. 4; squares: SeaWiFS only, stars: AQUA only, triangles: MERIS only, asterisks: AQUA + MERIS, white circles: SeaWiFS + MERIS, diamonds: SeaWiFS + AQUA, black circles: SeaWiFS + AQUA + MERIS).

Matchups for CDM and BBP are also plotted in Fig. 11 for the same time period (1997–2007). The CDM matchups show that the satellite retrievals are generally good over the two orders of magnitude covered by the available in situ data. A group of outliers that strongly overestimates CDM are stations from the "Plumes and Blooms" program in the Santa Barbara Channel where atmospheric correction is known to be a major issue (Bailey & Werdell, 2006). When these particular stations are removed from the analysis, the regression statistics between the in situ and satellite-derived CDM are generally very good even in waters with very high CDM (not shown). The CDM retrievals with the MERIS data alone show a lot of dispersion and the slope of the regression is relatively high (1.58). Except for the high CDM estimates mentioned above, the CDM retrievals from MERIS generally underestimate the in situ data. This is likely related to the fact that the MERIS NLw(412) values are generally high compared to those of AQUA or SeaWiFS (Zibordi et al., 2006a; Antoine et al., 2008a) which would tend to generate lower CDM estimates.

Because accurate backscattering measurements are difficult to achieve and are not yet accurately routinely performed, available validation data for particulate backscattering are much scarcer than for CHL or CDM and they cover a limited range of values. The backscattering matchups are good overall even though the great majority of in situ data are from coastal areas. The BBP retrievals are best for the merged data set and for SeaWiFS alone while retrievals using both AQUA alone or MERIS alone show higher slope and dispersion. Overall, the matchups with the merged data are good, particularly considering that many coastal observations are included in the in situ data set used in the present study.

3.6. Time-series of the GSM products from the individual sensors and from the merged radiances

The comparison of the individual sensors time-series with the time-series from the merged data is another way to assess and validate the GSM products and to check their consistency. Fig. 12 shows the time-series of the monthly global geometric mean of the GSM CHL, CDM and BBP retrievals for the individual sensors and for the merged data between 50°N and 50°S for bins where the bottom depth is deeper than 1000 m (in order to discard polar and coastal waters). For CHL (Fig. 12, upper panel) the agreement between the four sets of products (from SeaWiFS only, from AQUA only, MERIS only and from the merged radiances) is generally good between 2002 and 2006 with a slight seasonality characterized by peaks of CHL during early boreal spring and end of summer (although the two peaks do not appear every year). During that time period, the AQUA retrievals are slightly lower than the estimates from the other sources while the highest CHL values are generally derived when using the MERIS data alone. From 2006 and on, the AQUA CHL departs significantly from that of the other sources and shows a steady decrease with time while the CHL estimates from SeaWiFS or MERIS are stable. The AQUA CHL estimates decrease by 16% between the 2002–2005 and 2006–2009 periods (0.104 and 0.087 mg m^{-3}, respectively on average). The shift associated with the AQUA data is also clearly seen on the other products.

For CDM, the AQUA time-series tends to show slightly lower values than SeaWiFS for the 2002–2006 time period but they get higher than SeaWiFS around 2006 and at the end of the time-series (Fig. 12, middle panel). The MERIS CDM estimates are consistently lower than for the other sensors. For MERIS, the generally higher CHL and lower CDM values are again likely to be a consequence of the higher radiance values in the blue in the MERIS data relative to the other sensors. For each of the primary data sources, the CDM estimates show an increasing trend starting between 2006 and 2007.

Clear differences also exist for the BBP product. The SeaWiFS and MERIS BBP estimates are generally very close and are significantly higher than those of AQUA (Fig. 12, lower panel). Again, the AQUA data show an increasing trend that matches in time the changes observed in CHL and CDM whereas the BBP values from SeaWiFS and MERIS do not show an increase in BBP.

It is out of the scope of this paper to present an analysis of possible long-term trends in ocean color products; however for these analyses to be performed in the future, it is necessary to ensure that the observed variations of a particular product are not artifacts caused by instrumental or observational issues and it is hoped that significant improvements will come from reprocessing efforts that are on-going for all three sensors. That said, there are clear differences between the GSM retrievals from the different sensors. These differences and their possible causes are discussed and analyzed in more details below.

3.7. Discussion and conclusion

The discrepancies between the GSM products derived from different data sources can be explained, to a large extent, by differences in the radiometric data from each sensor. Differences in the radiometry can have various origins, from calibration to atmospheric correction, to sensitivity drift or data processing, for example. Because of its fitting technique and the possibility to weigh the different bands of each sensor in it, the merging procedure described here can handle a significant amount of noise in the NLw(λ) data. In any case, the retrievals are affected by these differences and important biases between data sets will often result in inaccurate or erroneous retrievals.

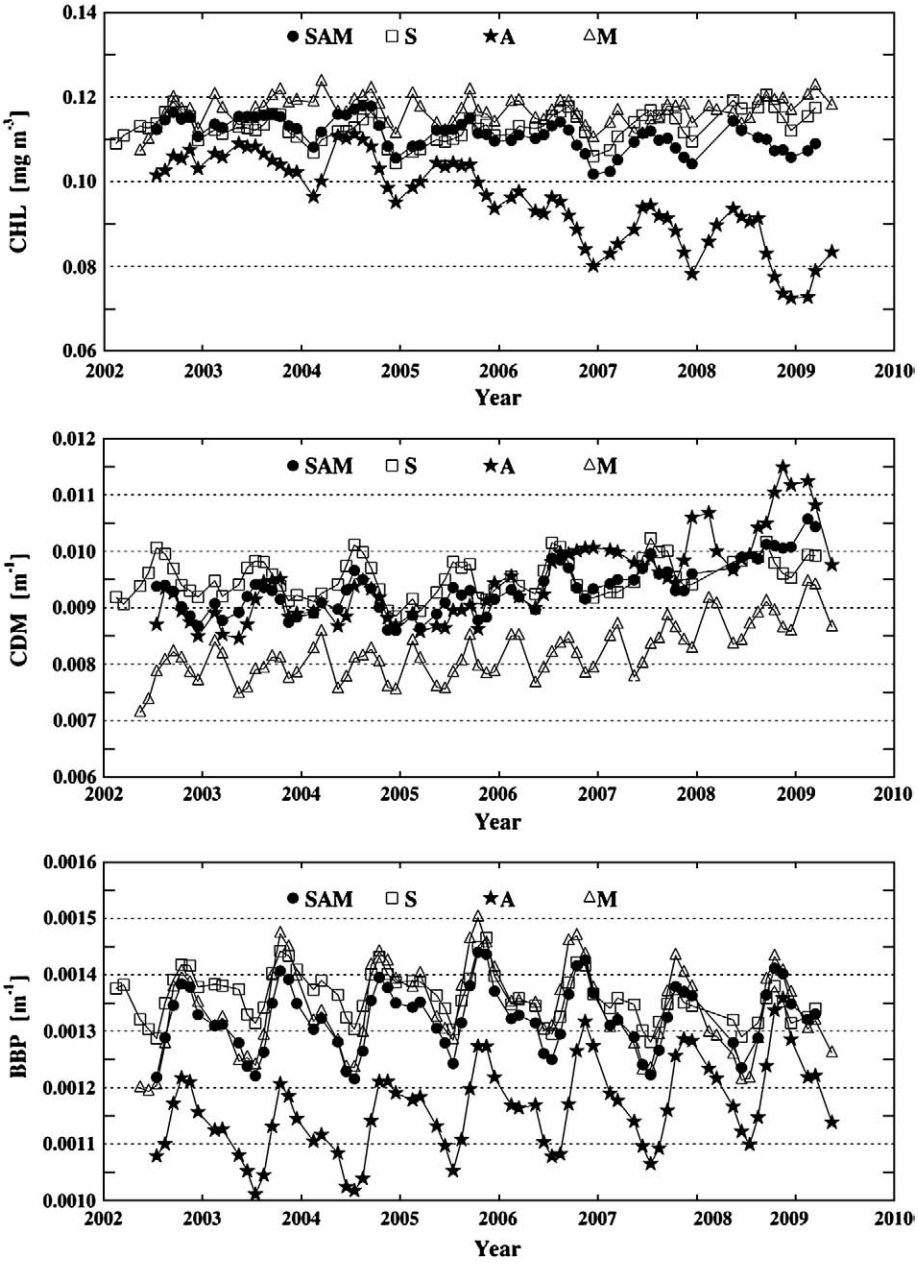

Fig. 12. Time-series of the global, 50°N–50°S, deep-water ($Z>1000$ m) geometric mean for CHL (upper plot), CDM (middle plot) and BBP (lower plot) for SeaWiFS only (squares), AQUA only (stars), MERIS only (triangles) and for the merged products (circles).

In the plots presented in Fig. 12, each of the GSM products shows a time-series that departs from those of other sensors. In some cases, a consistent difference is observed throughout the whole time-series (e.g. MERIS CDM or AQUA BBP) while in other cases a change appears sometime during the time-series (e.g. AQUA CHL or CDM). The drop in the global CHL average from AQUA starts approximately at the end of 2005 and the increase in the CDM and BBP values starts at the same time which suggests that these changes in the retrievals have the same cause. Because all three products are affected at the same time and since neither the SeaWiFS nor the MERIS time-series shows a similar decrease in CHL, it seems legitimate to suspect that an issue with one or more of the AQUA bands is responsible for the variations observed in the AQUA GSM products from 2006 and on. The global geometric mean of NLw(λ) corresponding to the GSM products time-series (50°N–50°S, depth>1000 m) are presented in Fig. 13 and they clearly show a decrease in the average NLw(412) since around the beginning of 2006. The decrease is rather important considering this is an average over a very large area. The average NLw(412) value for the 2002–2005 period is 1.56 mW cm^{-2} μm^{-1} sr^{-1} (standard deviation = 0.04) and is 1.39 mW cm^{-2} μm^{-1} sr^{-1} (standard deviation = 0.11) in 2006–2009. The other bands appear stable except maybe for 443 nm which also seems to decrease slightly at the end of the time-series. Surprisingly, the effect of the decrease in the 412 channel is more pronounced in the CHL retrievals than in CDM. The AQUA CDM time-series shows that the AQUA values were close and generally lower than those of SeaWiFS during the first ~3.5 years of the mission but tend to get higher than SeaWiFS once the 412 nm channel drift has started. An increase in the CDM retrievals with decreasing NLw(412) makes sense since the CDM absorption is at its highest at that channel. The big effect of the NLw (412) decrease on CHL is likely to be related to the parameterization of $a_{ph}(412)$ in GSM which has a very low coefficient (see Maritorena et al., 2002) and is thus sensitive to variations of the radiance in that waveband. The BBP values are also affected as they show an increase that starts towards the end of 2005. This issue with the AQUA NLw(412)

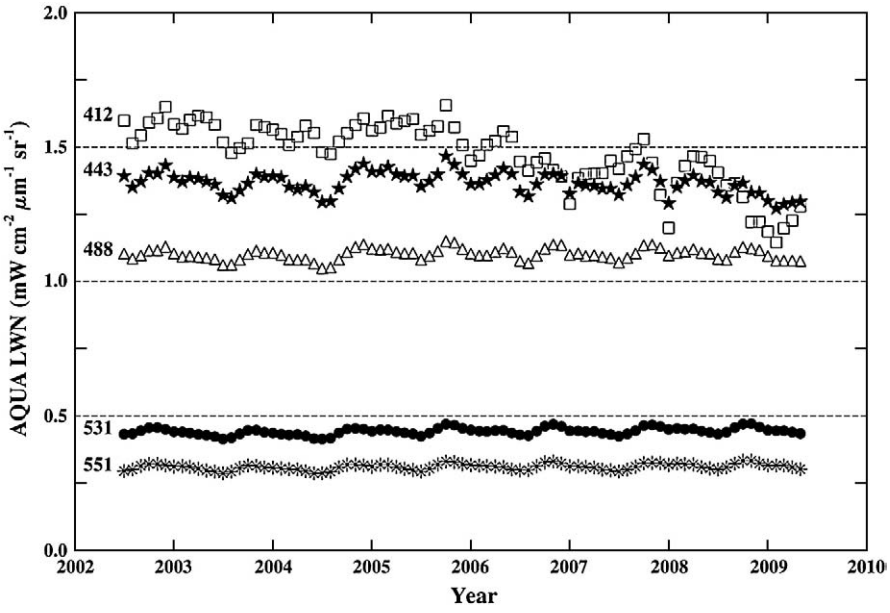

Fig. 13. Times-series of the global, 50°N–50°S, deep-water (Z>1000 m) geometric mean of AQUA NLw(λ) for the 2002–2009 time period.

does not influence the CHL estimates from the NASA OC3M operational maximum band-ratio algorithm (O'Reilly et al., 2000) as it does not use that particular band.

Other issues can be observed on the BBP retrievals, in the SeaWiFS and MERIS data in particular. For SeaWiFS, the BBP images frequently show some isolated very high BBP values in areas where the surrounding bins have lower and consistent values. These high BBP retrievals are generally observed in the vicinity of clouds. These high BBP values correspond to pixels with very high radiance compared to their neighbor pixels and are known as the "speckling" issue in SeaWiFS data (Hu et al., 2000, 2001). Speckling has several possible origins: incorrect atmospheric correction, incomplete cloud masking, digitalization errors for the near-infrared wavebands and the lack of filtering in the binning process that creates binned GAC data from higher spatial resolution LAC data. The inversion of the GSM model using these high radiance spectra results in BBP values that are overestimated (Fig. 14). The other two GSM products are much less or not affected by this issue as isolated high values are infrequent and of much lower magnitude for CHL and CDM. The stronger effect of speckling on BBP may be the result of the spectral characteristics of the erroneous radiances if the enhancement is more important in the green to red bands relative to the blue bands. In practice, these high radiance bins could be filtered (e.g. by not considering bins within some distance of a cloud) and not used in the GSM inversion but this could reduce significantly the number of useable bins for the merging. This can also be filtered during the LAC to GAC binning process and such a procedure is considered for the next SeaWiFS reprocessing (B. Franz, personal communication).

High BBP values are also observed in some swaths associated with MERIS data. In that case, the swath is divided in two with its Eastern side appearing as a stripe of high BBP values. The same pattern is present in the radiances and is clearly noticeable for the long wavelength bands (≥560 nm) where radiances are higher than on the other side of the swath (Fig. 14). This issue results from a combination of factors: camera effect (different cameras see the Western and Eastern parts of a swath), imperfect calibrations and inter-calibrations among the cameras and thermal dependency of the dark signal correction scheme used for MERIS. The problem seems restricted to some areas and is not present throughout the MERIS mission. A fix is currently being investigated (L. Bourg, personal communication).

These issues do not appear on the AQUA data and do not seem to affect the other products. BBP being the variable with the smallest values (one or more order of magnitudes smaller than CDM) of the three retrieved variables it is more sensitive to noise in the radiance data than CHL or CDM.

The data issues discussed above emphasize that for semi-analytical models that use all possible visible spectral bands of a sensor, like GSM, the quality of the spectral data entering the model is key to derive accurate retrievals even if uncertainties among the bands are taken into account exist as is the case in this study. It should also be mentioned that the upcoming reprocessing of the three sensors used here are likely to correct or at least alleviate many of the issues described above.

Ocean color products constructed from multiple sensor data have various potential benefits for the research and applications communities working on biogeochemistry or the marine biosphere in general. Well characterized and quality controlled merged data sets eliminate the headaches that result from the existence of multiple versions of the same product from different missions simultaneously operational. With merged data sets, data users have access to a unified product with features and characteristics that none of the individual missions have. Here, we have documented the benefits and characteristics of ocean color merged data products derived from the inversion of the GSM01 semi-analytical model. This model allows the merging of satellite data at the normalized water-leaving radiance level and the simultaneous retrieval of several ocean color products. Merged products obviously have better spatial and temporal coverage than what individual missions can provide. Using the data from the three current operational global ocean color missions, namely SeaWiFS, AQUA and MERIS, the global daily coverage of a merged product is about 25% of the world ocean, about twice as much as what the single mission with the best coverage, SeaWiFS, provides. Similarly, the frequency at which a particular area is sampled from space is generally greatly improved when merged data are used. The magnitude of the improvement in sampling frequency is not homogeneous throughout the world ocean. High latitude regions remain poorly sampled even when combining the data from three sensors. Conversely, the inter-tropical zone is the most frequently sampled except for the equatorial region because of sun-glint issues with non-tilting sensors. By combining the data from SeaWiFS, AQUA and MERIS some areas are sampled as frequently as 64% of the time (233 days/year on average).

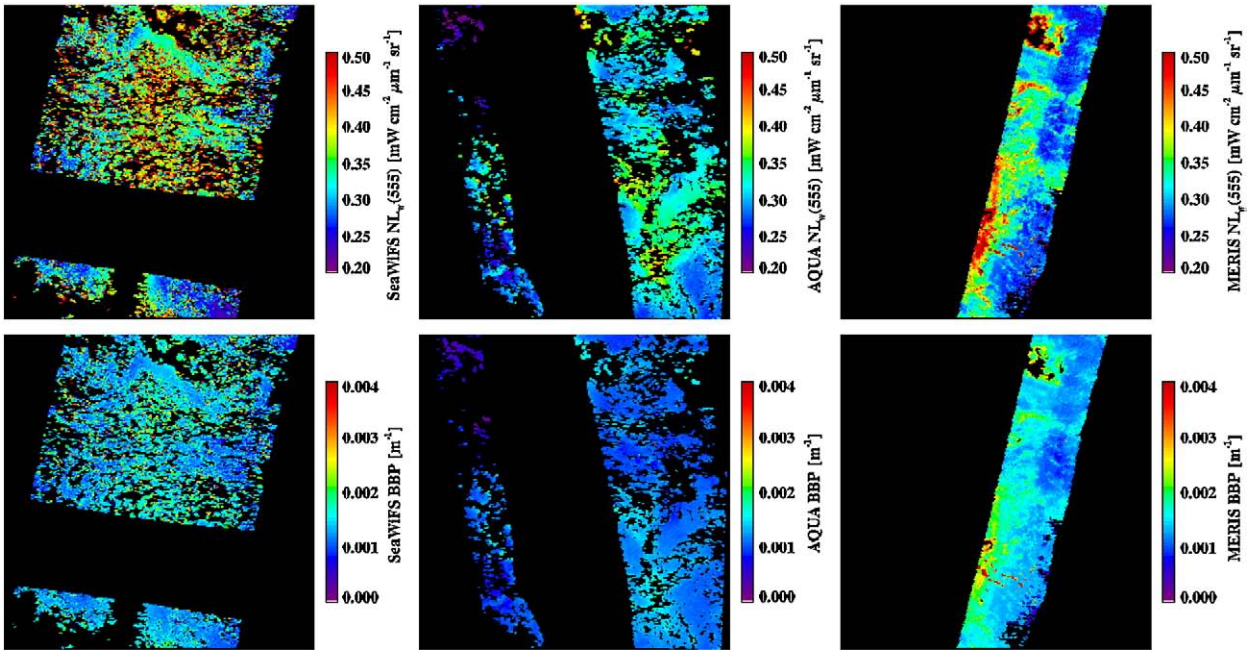

Fig. 14. Top panels: maps of NLw(555) from SeaWiFS (left), AQUA (center) and MERIS (right) and corresponding BBP maps (lower panels) for the 0°–20°S/155°W–135°W region on January 20, 2003. High NLw(555) are present around cloud gaps in SeaWiFS data and these high NLw(λ) spectra result in high BBP retrievals. Generally, AQUA does not show these high radiance spectra and consequently, the AQUA imagery has much smoother BBP fields than SeaWiFS. For MERIS, some swaths show a clear difference in the NLw(λ) between their left and right sides which seems to indicate an issue with the cameras and their intercalibration (see text).

The merged data presented here have been validated through matchup analyses and by comparing them to the data sets obtained from individual missions. Overall, the three main products derived from the GSM01 model agree very well with in situ data except for some particular coastal waters where atmospheric correction is a known issue which limits the accuracy of the NLw(λ) spectra entering the model. Disagreements between products derived for the single sensors are traced to disagreement in the radiometric data and each of the sensors processed here has its own set of issues, most of which should be taken care of during future reprocessings. This stresses the fact that the constant monitoring of data quality of individual sensors, intercomparisons of sensors and bias analysis will remain an absolute and crucial requirement for ocean color in the future.

Another important feature of the merged data presented here is that all the GSM01-derived products have documented uncertainties associated to them on a pixel-by-pixel basis. To our knowledge, the MEaSUREs and GlobColour data sets are the first and only examples of distributed ocean color data products with uncertainties associated with each pixel of an image. The uncertainty maps are data products in their own rights and they provide information that should be useful to modelers using ocean color data assimilation (e.g. Nerger & Gregg, 2007; Gregg et al., 2009). Moreover, a complete error budget has been developed from the input radiances to the model and, ultimately, the products. This is also the first time this has been done for satellite ocean color data. Comparisons of the NLw(λ) from each sensor with in situ data were used to assign spectral weights to the radiances of each sensor in the inversion of the GSM01 model. The errors caused by the model itself were also assessed by comparing in situ NLw(λ) spectra to those corresponding to the model inversion best fits. These error estimates were taken into account in assigning the uncertainty weights of the different wavebands of each sensor. It should be noted that the analysis of the uncertainties in the NLw(λ) was conducted before the AQUA NLw (412) drift started. However, the uncertainty analysis should ideally be performed regularly, in parallel with what is done with matchup data. One of the limitations of the error budget presented here is that a constant uncertainty value is set for each band of each sensor while it is likely that these uncertainties vary seasonally and geographically, and trends exist over time. Future ocean color missions like OLCI on-board Sentinel 3 will go further by including an error budget analysis that will allow propagating the uncertainties on NLw(λ) at pixel level from TOA measured radiance errors.

The MEaSUREs and GlobColour merged data sets are available to the research community (Table 1) and a growing number of users have shown interest in using these data for their research work. It is very likely that more merged data sets will emerge in the future as there is a growing need for the development of long-term time-series of ocean color data. In particular, the quantification of changes in the ocean biosphere over several decades will require data from multiple sensors to be assembled together as it is very unlikely that a single sensor or even a single family of sensors (e.g. the VIIRS sensors on NPP and NPOESS, or the OLCIs on Sentinel-3A and 3B) can provide all the necessary data. Climate Data Records (NRC, 2004), CDRs, are satellite data sets covering a long enough period of time to allow the determination of climate variability and change. Because CDRs need to be long enough to allow discrimination between trends and oceanic decadal-scale cycles, they can only be developed from multiple satellite missions. In addition, a CDR needs to be made out of data that are consistent throughout the time-series and since it is likely that data from different sensors will overlap at some point in the time-series, the use of a GSM-like ocean color model will satisfy both the consistency requirement (in terms of data product derivation) and the merging capability to unify data sets during periods of overlap. The development of ocean color CDRs will face many difficulties one of which is how to deal with possible biases among sensors (see Siegel et al., in preparation; Djavidnia et al., 2009). With the ability of the GSM model approach to apply statistical weight to the NLw(λ), it is possible, to some extent, to correct for possible biases in the data from different sensors. However as shown in the present study, the quality of the radiometric data is a crucial element to ensure accurate retrievals from a multispectral semi-analytical model and thus the ability to monitor the stability and accuracy of the radiance data is essential to the development of future ocean color CDRs.

In the foreseeable future of satellite ocean color, it is probable that there will be periods of time when data will come from multiple

missions flying simultaneously, while at other times there will be only a single instrument available. The model-based approach used here can handle all these situations while ensuring consistency in products derivation throughout the whole time-series. It has the additional benefit of being able to generate ocean color products from multiple sources of data as coherent wholes, with documented uncertainties. These features and the fact that there is ample room for improvement in both ocean color semi-analytical models and merging procedures make model-based approaches obvious candidates for the creation of ocean color Climate Data Records in the future.

Acknowledgments

The work presented here was funded by NASA under the SIMBIOS, REASoN and MEaSUREs programs and by the ESA DUE GlobColour program with continuation support from the EC FP7/2007–2013 program under grant agreement no. 218812 (MyOcean). The SeaWiFS and MODIS-AQUA data were available thanks to NASA and OrbView while ESA provided the MERIS data. The NASA OBPG group is also thankfully acknowledged for his constant support and help and for providing access to the NOMAD data set and the SeaBASS archive. The authors also wish to thank Manuela Lorenzi-Kayser, Pete Peterson, Erik Fields, Jim Frew, Michael Colee, Peter Slaughter, Julien Demaria, Gilbert Barrot, Jacques Daniel and the GlobColour Team for their invaluable help at various stages of this work.

References

Antoine, D., d'Ortenzio, F., Hooker, S. B., Bécu, G., Gentili, B., Tailliez, D., et al. (2008a). Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean Q4 Sea (BOUSSOLE project). *Journal of Geophysical Research*, *113*, C07013. doi:10.1029/2007JC004472

Antoine, D., Guevel, P., Desté, J. -F., Bécu, G., Louis, F., Scott, A. J., et al. (2008b). The «BOUSSOLE» buoy: A new transparent-to-swell taut mooring dedicated to marine optics: Design, tests and performance at sea. *Journal of Atmospheric and Oceanic Technology*, *25*, 968−989.

Bailey, S. W., & Werdell, P. J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, *102*, 12−23.

Bates, D. M., & Watts, D. G. (1988). Nonlinear regression analysis and its applications. Wiley.

Castro, S. L., Wick, G. A., Jackson, D. L., & Emery, W. J. (2008). Error characterization of infrared and microwave satellite sea surface temperature products for merging and analysis. *Journal of Geophysical Research*, *113*, C03010.

Djavidnia, S., Mélin, F., & Hoepffner, N. (2009). Comparative analysis of the multi-sensor global ocean colour data record. *Ocean Science – Discussion*, *6*, 1611−1653.

Fanton d'Andon, O., Antoine, D., Mangin, A., Maritorena, S., Durand, D., Pradhan, Y., et al. (2008). Ocean colour sensors characterisation and expected error estimates of ocean colour merged products from GlobColour. *Proceedings ocean optics XIX, Barga, Italy, October 6–10* (pp. 51).

Franz, B. A., Kwiatkowska, E. J., Meister, G., & McClain, C. (2008). Moderate Resolution Imaging Spectroradiometer on Terra: Limitations for ocean color applications. *Journal of Applied Remote Sensing*, *2*, 023525.

GCOS, (2003). The second report on the adequacy of the global observing systems for climate in support of the UNFCCC. GCOS-82 (WMO/TD No. 1143), Bonn, Germany, April 2003. (http://www.wmo.int/pages/prog/gcos/Publications/gcos-82_2AR.pdf).

Gregg, W. W., & Conkright, M. E. (2001). Global seasonal climatologies of ocean chlorophyll: Blending *in situ* and satellite data for the coastal zone color scanner era. *Journal of Geophysical Research*, *106*(C2), 2499−2516.

Gregg, W. W., Friedrichs, M. A. M., Robinson, A. R., Rose, K. A., Schlitzer, R., Thompson, K. R., et al. (2009). Skill assessment in ocean biological data assimilation. *Journal of Marine Systems*, *76*(1–2), 16−33.

Hooker, S. B., Zibordi, G., Berthon, J. -F., & Brown, J. W. (2004). Above-water radiometry in shallow coastal waters. *Applied Optics*, *43*(21), 4254−4268.

Hu, C., Carder, K. L., & Muller-Karger, F. E. (2000). Atmospheric correction of SeaWiFS imagery: Assessment of the use of alternative bands. *Applied Optics*, *39*, 3573−3581.

Hu, C., Carder, K. L., & Muller-Karger, F. E. (2001). How precise are SeaWiFS ocean color estimates? Implications of digitization-noise errors. *Remote Sensing of Environment*, *76*(2), 239−249.

IOCCG, Ocean-Colour Data Merging. (2007). In W. Gregg (Ed.), *Report of the International Ocean-Colour Coordinating Group, No. 6, IOCCG, Dartmouth, Canada*.

Kwiatkowska, E. J., & Fargion, G. S. (2002). Merger of ocean color data from multiple satellite missions within the SIMBIOS project. *Proc. SPIE symp. — Remote sensing of the atmosphere, ocean, environment, and spaceOcean remote sensing and applications, Hangzhou, China, vol. 4892*. (pp. 168−182) Oct.

Loeb, N. G., Wielicki, B. A., Wong, T., & Parker, P. A. (2009). Impact of data gaps on satellite broadband radiation records. *Journal of Geophysical Research*, *114*, D11109. doi:10.1029/2008JD011183

Maritorena, S., & Siegel, D. A. (2005). Consistent merging of satellite ocean color data sets using a bio-optical model. *Remote Sensing of Environment*, *94*(4), 429−440.

Maritorena, S., Siegel, D. A., & Peterson, A. (2002). Optimization of a semi-analytical ocean color model for global scale applications. *Applied Optics*, *41*(15), 2705−2714.

McClain, C. R. (2009). A decade of satellite ocean color observations. *Annual Review of Marine Science*, *1*, 19−42.

Mélin, F., Zibordi, G., & Djavidnia, S. (2009). Merged series of normalized water leaving radiances obtained from multiple satellite missions for the Mediterranean Sea. *Advances in Space Research*, *43*(3), 423−437.

Merrifield, M. A., Gill, S., Mitchum, G. T., & Woodworth, P. L. (2008). Sea level variations, in state of the climate in 2007. In D. H. Levinson, & J. H. Lawrimore (Eds.), *Bulletin of the American Meteorological Society*, *89*. (pp. S107−S109).

National Research Council (NRC). (2004). *Climate data records from environmental satellites.*Washington, DC: National Academies Press 116 pp.

Nerger, L., & Gregg, W. W. (2007). Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter. *Journal of Marine Systems*, *68*(1–2), 237−254.

O'Reilly, J.E., & 24 Coauthors, 2000: SeaWiFS postlaunch calibration and validation analyses, part 3. NASA tech. memo. 2000-206892, Vol. 11, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, 49 pp.

Parkinson, C. L., & Cavalieri, D. J. (2008). Arctic sea ice variability and trends, 1979–2006. *Journal of Geophysical Research*, *113*(C7), C07003.

Pottier, C., Garçon, V., Larnicol, G., Sudre, J., Schaeffer, P., & Le Traon, P. -Y. (2006). Merging SeaWiFS and MODIS/AQUA ocean color data in North and Equatorial Atlantic using weighted averaging and objective analysis. *IEEE Transactions On Geoscience and Remote Sensing*, *44*(11), 3436−3451 Part 2.

Press, W. H., Teukolsky, S. A., Vettering, W. T., & Flannery, B. P. (1992). Numerical recipes in C: The art of scientific computing, 2nd edition. Cambridge University Press 994 pp.

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, *20*(22), 5473−5496.

Siegel, D. A., Maritorena, S., Nelson, N. B., & Behrenfeld, M. J. (2005a). Independence and interdependencies among global ocean color properties: Reassessing the bio-optical assumption. *Journal of Geophysical Research*, *110*, C07011. doi:10.1029/2004JC002527

Siegel, D. A., Maritorena, S., Nelson, N. B., Behrenfeld, M. J., & McClain, C. R. (2005b). Colored dissolved organic matter and its influence on the satellite-based characterization of the ocean biosphere. *Journal of Geophysical Research*, *32*, L20605. doi:10.1029/2005GL024310

Siegel, D., McClain C.R., Behrenfeld, M.J., Fanton d'Andon Hembise, O., Antoine, D., Maritorena, S., Bailey, S., Franz, B., and others (in preparation). Challenges facing the creation of satellite ocean color Climate Data Records.

Van Der Woerd, H. J., & Pasterkamp, R. (2008). HYDROPT: A fast and flexible method to retrieve chlorophyll-a from multispectral satellite observations of optically complex coastal waters. *Remote Sensing of Environment*, *112*, 1795−1807.

Wang, P., Boss, E., & Roesler, C. (2005). Uncertainties of inherent optical properties obtained from semi-analytical inversions of ocean color. *Applied Optics*, *44*, 4074−4085.

Werdell, P. J., & Bailey, S. W. (2002). *The SeaWiFS Bio-optical Archive and Storage System (SeaBASS): Current architecture and implementation, NASA/TM-2002-211617*. Greenbelt, MD: NASA Goddard Space Flight Center 45 pp.

Werdell, P. J., & Bailey, S. W. (2005). An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, *98*, 122−140.

Zibordi, G., Mélin, F., & Berthon, J. -F. (2006a). Comparison of SeaWiFS, MODIS and MERIS radiometric products at a coastal site. *Geophysical Research Letters*, *33*, L06617. doi:10.1029/2006GL025778

Zibordi, G., Mélin, F., & Berthon, J. -F. (2006b). A time-series of above-water radiometric measurements for coastal water monitoring and remote sensing product validation. *IEEE Geoscience and Remote Sensing Letters*, *3*(1), 120−124.

Zibordi, G., Strombeck, N., Mélin, F., & Berthon, J. -F. (2006c). Tower-based radiometric observations at a coastal site in the Baltic Proper. *Estuarine, Coastal and Shelf Science*, *69*(3–4), 649−654.