Research Project co-founded by the
European Commission
Research Directorate-General
6th Framework Programme
FP6-2002-Space-1-GMES
Ocean and Marine Applications
Contract No. AIP3-CT-2003-502885

MERSEA – IP

Marine EnviRonment and Security for
the European Area - Integrated Project

Report on Deliverable D.2.3.11

Assessment of Global Ocean Colour Products against In-situ Datasets

Marco Clerici[1], Frédéric Mélin and Nicolas Hoepffner

European Commission - Joint Research Centre
Institute for Environment & Sustainability
Global Environment Monitoring Unit
TP 272 - Via E. Fermi 2749
Ispra (VA), 21027
Italy

Ref: MERSEA-WP02-JRC-STR-002-01A.pdf

April 2008

EUR 23357 EN

[1]Corresponding author email address: marco.clerici@jrc.it

The mission of the Institute for Environment and Sustainability is to provide scientific and technical support to the European Union's policies for protecting the environment and the EU Strategy for Sustainable Development.

European Commission
Directorate-General Joint Research Centre
Institute for Environment and Sustainability

Contact information
Address: Marco Clerici, EC - JRC, IES, GEM, TP-272, Via E. Fermi 2749, Ispra (VA), 21027, Italy
E-mail: marco.clerici@jrc.it
Tel.: +39 0332 785727
Fax: +39 0332 789034

http://ies.jrc.cec.eu.int
http://www.jrc.ec.europa.eu

EUR 23357 EN
ISSN 1018-5593

Luxembourg: Office for Official Publications of the European Communities

Printed in Italy

Contents

1 Introduction **1**

2 Chlorophyll in situ data **3**
2.1 NODC Data Set 3
2.2 SeaBASS Data Set 5
2.3 JODC Data Set 7
2.4 Statistical Analysis 9

3 Satellite and in situ data comparison **10**
3.1 Introduction 10
3.2 SeaWiFS Validation Set 10
3.3 MODIS Validation Set 13

4 Regional analysis **16**
4.1 Data set merging 16
4.2 Regional Analysis 18
4.2.1 SeaWiFS 22
4.2.2 MODIS 26
4.3 Coastal areas and open ocean 27

5 Conclusions **29**

Acknowledgements **30**

List of Acronyms **31**

Document Change Record

Author	Modification	Issue	Date
M. Clerici	First Draft	0.1	January 2008
M. Clerici	Final Isue	1.0	April 2008

Table 1: Document Change Record

Abstract

Ocean colour from satellite has given over the last two decades another dimension to ecosystem studies and marine biology, providing key information on the timing and spatial distribution of phytoplankton blooms, and the magnitude of primary production. Remote observations of ocean colour from space represent therefore a major tool directly related to the marine biogeochemical distributions and associated processes.

One of the goals of the European GMES Integrated Project MERSEA is to provide an accurate and consistent stream of ocean colour data, by exploiting the products made available in a number of individual missions launched by various space agencies. In this context, validation exercises, done via the direct comparison of satellite derived quantities with in situ measurements, represents a critical component in establishing the accuracy of the remotely-sensed data.

In this study we present a validation of chlorophyll-a concentration derived from SeaWiFS and MODIS sensors, against in situ measurements retrieved from three different datasets (NODC, SeaBASS, JODC). The results of this comparison are well in line with previous analysis conducted on SeaWiFS, both from the point of view of the global statistics than for most of the regional results, and the uncertainties are lower than the value of 0.35 often considered as the objective for chlorophyll-a distributions. The SeaWiFS global average of RMS difference (for log-transformed values) shows an uncertainty of 0.29, while it is is slightly higher for MODIS (0.31), a difference likely partly due to a smaller statistical basis. The agreement is better for open ocean regions (RMSD reduced to 0.26 and 0.27 for SeaWiFS and MODIS respectively) than for coastal areas.

An important objective of this work, that goes beyond the scope of the present report, was to develop the validation procedure and protocols for further analyses regularly reviewing validation results to take into account successive reprocessing and other sensors, as well as including additional in situ data sets.

Chapter 1

Introduction

A continuous long time series of bio-optical and geophysical variables derived from ocean colour satellite data enables oceanographers to monitor and model changes in coastal and open ocean biological systems that might occur naturally or as a direct result of climate change and human activities. One of the goals of the GMES Integrated Project MERSEA is to provide an accurate and consistent stream of ocean colour data at a resolution and format compatible with operational forecasting of the marine environment. For example, only ocean colour radiometry can provide the high-resolution view of spatial structures, the wide area coverage and the repeated sampling over time to capture the detailed evolution of a phytoplankton population that is required to constrain the model parameters.

Space-borne ocean colour instruments measure the spectrum of sunlight reflected from ocean waters at selected visible and near-infrared wavebands. In turn, the radiance spectra are used to estimate geophysical parameters, such as the surface concentration of the phytoplankton pigment chlorophyll-a, via the application of bio-optical algorithms. The performance of these algorithms depends on several factors: i) the characteristics of the sensor (radiometric resolution, signal to noise ratio; ii) our ability to correct the remotely-sensed signal for the atmospheric effects; and iii) our ability to parameterize the water leaving radiance as function of the optical properties of the water, themselves reflecting a particular structure and biogeochemical composition of the water column. The integrity of the output products is optimized through proper calibration of the remote sensing system, i.e. sensor and algorithms. Accordingly, global accuracy goals often announced by space agencies for ocean colour sensors are to retrieve water leaving radiance in open ocean waters at an error not exceeding 5% and 35% for chlorophyll$a$ concentration (Hooker et al. 1992).

These targets are difficult to achieve, because of the limitations on the calibration accuracy that can be reached for space optical sensors, and the difficulty in properly representing the aerosol components and the optical properties of the sea water constituents. Validation exercises thus remain a critical component in establishing the uncertainties in retrieving the geophysical products, assessing their scientific utility, and identifying conditions for which their reliability is suspect. Validation activities are done via the direct comparison of remotely-sensed data with coincident in situ measurements. The objective of the work presented in this report is to evaluate on the global and regional scale the quality of the satellite data base on surface chlorophyll$a$ concentration as assembled within MERSEA. In situ data are taken from large oceanographic data base publicly accessible (namely NODC, SeaBASS and JODC).

Another goal of this work is to put in place the capability of conducting validation studies at large scale for Chl$a$ in the future, in order to monitor the performance of the satellite systems in producing quality records of this Essential Climate Variable (ECV). This is one facet of the various approaches to assess the quality of the ocean colour data record. In that respect it is a complement to the validation of the atmospheric correction schemes using aerosol and radiometric in situ values (e.g., Mélin et al. 2007, Zibordi, G. and Mélin, F. and Berthon, J.-F. 2006), the comprehensive validation of apparent and inherent optical properties for a limited number of sites (e.g., Mélin et al.

2007a) or the inter-comparison of products obtained from different sensors (Djavidnia et al. 2006, Mélin and Zibordi 2003).

The present validation exercise is focusing on the Sea-viewing Wide Field-of-view Sensor (SeaWiFS, Hooker et al. 1992) and the Moderate Resolution Imaging Spectroradiometer (MODIS, Salomonson, V.V. and Barnes, W.L. and Maymon, P.W. and Montgomery, H.E. and Ostrow, H. 1989).

Chapter 2

Chlorophyll in situ data

2.1 NODC Data Set

The U.S. National Oceanographic Data Center (NODC *) is a national repository and dissemination facility for global oceanographic in situ and remote sensing data. The downloaded data set contains, for the period 1997 to 2004, 6526 Chl*a* measurement stations. When the condition of $depth_0$ (i.e. the measurement depth closest to the water surface) less than 5 m is applied, 6077 measurements are identified, distributed as shown in Fig. 2.1.

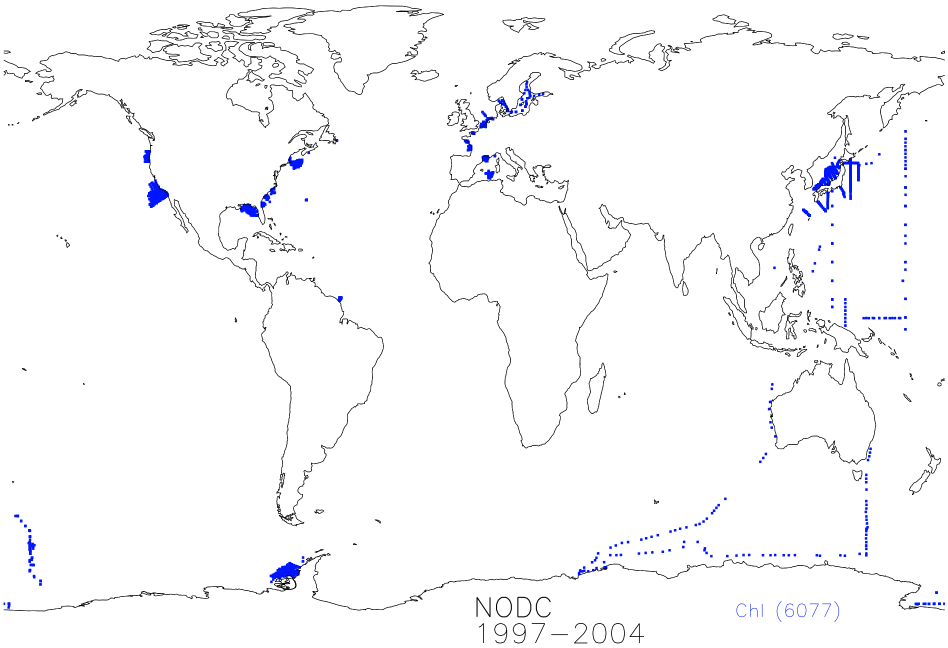

Figure 2.1: NODC data set: location of measurement stations.

As shown in Table 2.1, the availability of NODC data set measurements is concentrated over the period 1997 to 1999 (70% of the total), and therefore this data base is mainly useful for SeaWiFS Chl*a* product assessment.

The Chl*a* measurements can be either at a single depth (1445 cases) or along a vertical profile, containing two or more measurements (4632 cases). In the former case, the so-called remote sensing (or satellite) pigment concentration C_{sat} is simply assumed as C_{depth_0}, while in the latter it is calculated as the integral over depth of the profile values $C(z)$ weighted by a factor $g(z)$, function of the

*http://www.nodc.noaa.gov/

Year	1997	1998	1999	2000	2001	2002	2003	2004
N.meas.	765	1862	1678	369	392	501	245	265

Table 2.1: Number of NODC measurements per year.

diffuse attenuation coefficient for down-welling irradiance E_d (Gordon and Clark 1980):

$$C_{sat} = \frac{\int_0^{z_{90}} g(z) \cdot C(z) \cdot dz}{\int_0^{z_{90}} g(z) \cdot dz} \tag{2.1}$$

where

$$g(z) = exp(-2 \int_0^z K_d(z') \cdot dz') \tag{2.2}$$

z_{90} is the penetration depth taken to be the depth where E_d is equal to $E_d(0) \cdot e^{-1}$, or $1/K_d$ (Gordon and Clark 1980; Smith 1981). The diffuse attenuation coefficient K_d is here estimated at 500 nm from the Chl$a$ concentration, assuming a Case-1 water model (Morel and Maritorena 2001), as:

$$K_d(z, 500) = 0.02188 + 0.06579 \cdot C(z)^{0.6888} \tag{2.3}$$

The derived C_{sat} values are compared with C_{depth_0} in fig. 2.1, showing a good agreement (the coefficient of determination r^2 is 0.992 and the bias† is -0.0076). There are a few cases, easily recognizable in the scatter plot, where C_{sat} is much higher then C_{depth_0}, due to a significant Chl$a$ maximum at depth higher than that associated with C_{depth_0}. In 11 cases, not shown in the plot, Chl$a$ traces have been been found in deep water, while the measurement at the surface is zero.

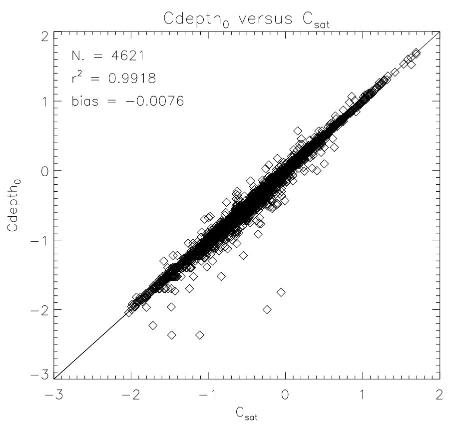

Figure 2.2: NODC dataset - C$_{sat}$ derived from vertical integration vs. C_{depth_0}

†Bias for log-normal distributions is computed as average of log(C_{depth_0}) - log(C_{sat}), see Djavidnia et al. (2006, sec. 3.3.6)

2.2 SeaBASS Data Set

The SeaWiFS Bio-optical Archive and Storage System (SeaBASS, ‡), originally developed by the SeaWiFS Project (Werdell, P.J. and Bailey, S.W. and Fargion, G.S. and Pietras, C. and Knobelspiesse, K.D. and Feldman, G.C. and McClain, C.R. 2003) is now maintained by the NASA Ocean Biology Processing Group (OBPG). It contains radiometric (optical) and phytoplankton pigment data, as well as atmospheric data, provided by various institutions (80 contributors), that are suitable for activities of validation or algorithm development (e.g., Werdell, P.J., Bailey, S.W. and Pietras, C. and Knobelspiesse, K.D. and Feldman, G.C. and McClain, C.R. 2005; Bailey and Werdell 2006). The data set contains Chl$a$ estimates derived with both optical (fluorimetry/spectrophotometry) and High Performance Liquid Chromatography (HPLC) methods. Assessments of the discrepancies that can arise from the different methods, for instance performed within the SIMBIOS Project Chlorophyll Round Robin Activities (Van Heukelem et al. 2002), show significant differences. Therefore, in the following analysis, we distinguish two SeaBASS data sets, according to the Chl$a$ measurement methods (FLUO and HPLC).

In the retrieved data set, there are overall 25,256 measurement stations: 24,517 of them are considered, with $depth_0$ less than 5 m. Chl$_{FLUO}$ is measured at 24,382 stations and Chl$_{HPLC}$ at 2,961 stations (see also Table 2.2).

Dataset	Single meas.	Vertical Profile	Total
FLUO	18,849	5,533	24,382
HPLC	2,175	786	2,961

Table 2.2: Number of SeaBASS measurement stations - FLUO and HPLC

The location of the SeaBASS stations is shown in Figure 2.2, which highlights a better coverage of open ocean regions, with respect to the NODC data set.

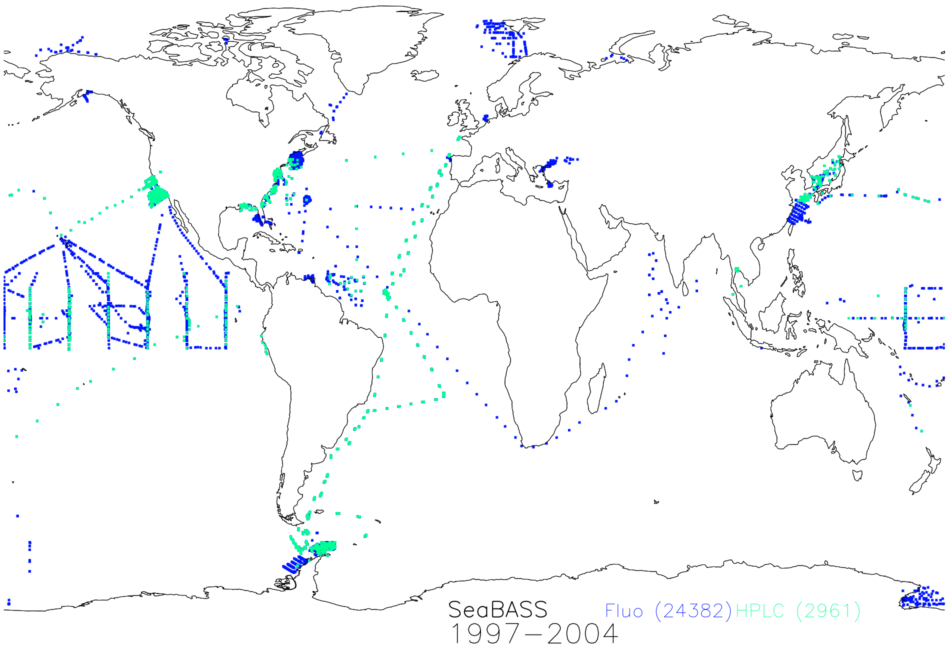

Figure 2.3: SeaBASS data set: location of measurement stations.

‡http://seabass.gsfc.nasa.gov/

Year	1997	1998	1999	2000	2001	2002	2003	2004
N.FLUO	3498	7085	4935	4209	2625	1408	606	16
N.HPLC	70	178	278	293	892	809	441	0

Table 2.3: Number of SeaBASS measurements per year.

For the two data sets (FLUO and HPLC) C_{sat} is computed as in 2.1, and shows a very high coefficient of determination vs. C_{depth_0} (higher than 0.98 in both cases), while the bias of logarithmically transformed measurement values amounts to -0.030 and -0.048, respectively.

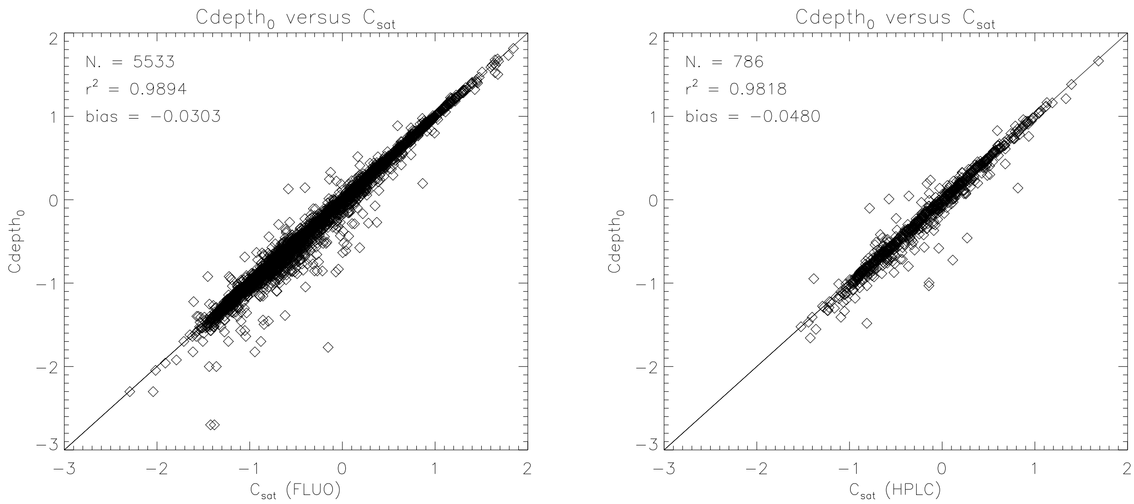

Figure 2.4: SeaBASS data set (FLUO and HPLC) - C$_{sat}$ derived from vertical integration vs. C_{depth_0}

2.3 JODC Data Set

Japan Oceanographic Data Center (JODC, [§]) is acting as the marine data bank of Japan. In addition JODC has been carrying out international services as the National Oceanographic Data Center of Japan under the framework of International Oceanographic Data and Information Exchange(IODE) promoted by IOC. Therefore, some measurement stations, originated by Japanese cruises, are also present in the U.S. NODC data set: these commonalities will be considered in merging the different data set (see 4.1). The data set retrieved from JODC in the period 1997 to 2005, contains 5,548 measurements; 5,411 of them are selected, as $depth_0$ is less than 5 m.

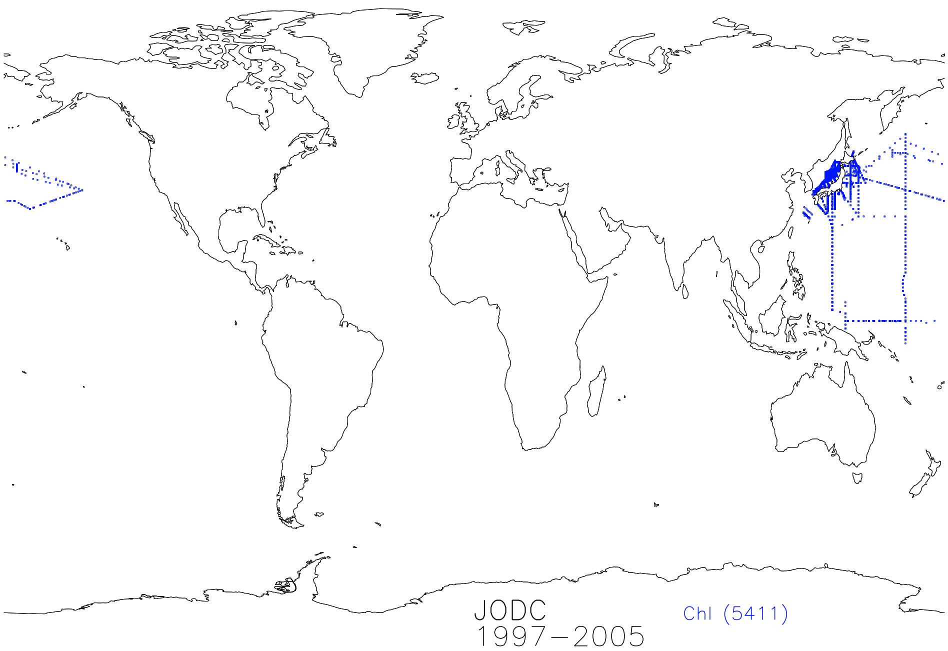

Figure 2.5: JODC dataset: location of measurements stations.

These measurements are almost equally distributed in the period 1997 to 2005 (Table 2.4), and a significant number spans the MODIS operational period (almost 2400 in the years 2002 to 2005). They are mostly located in the western Pacific Ocean (Figure 2.5).

Year	1997	1998	1999	2000	2001	2002	2003	2004	2005
N.meas.	374	660	638	674	678	674	576	500	637

Table 2.4: Number of JODC measurements per year.

[§]http://www.jodc.go.jp/

C_{sat} is computed as for the previous data sets, and the relationship with C_{depth_0} is shown in Figure 2.6. Once again there is, as expected, a very high correlation between the two variables (r^2 is 0.998) and the log-computed bias is only -0.0135. JODC Chl*a* measurements are represented in the data set text files as float numbers with two digits of decimals; this discretization is visible in the log-scale scatter plot at low Chl*a* values (namely, 0.01 to 0.03 mg m^{-3}).

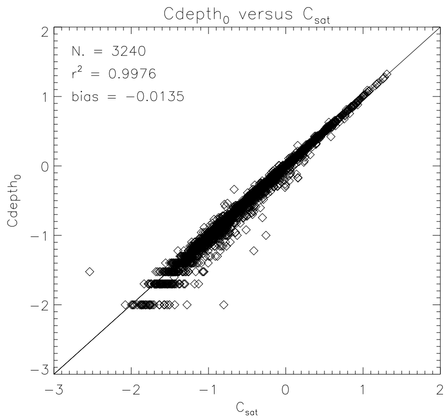

Figure 2.6: JODC data set - C$_{sat}$ derived from vertical integration vs. C_{depth_0}

2.4 Statistical Analysis

Chlorophyll-a concentration and other bio-optical water properties generally present, at least on the global scale, a log-normal distribution (Campbell 1995). This largely used assumption is verified on the available in situ data sets, described in the previous sections, by computing shape parameters of the distributions, namely skewness and kurtosis (see Table 2.5). Note that kurtosis is computed subtracting from the raw kurtosis the kurtosis of the Gaussian distribution (3.0).

Dataset	Skew. Chl	Skew. log(Chl)	Kurt. Chl	Kurt. log(Chl)
NODC	8.12	0.160	109.	0.113
SeaBASS-FLUO	6.43	-0.134	62.1	-0.329
SeaBASS-HPLC	4.93	0.198	4.93	0.198
JODC	6.42	0.090	58.9	0.110

Table 2.5: Skewness and kurtosis of the original and log-transformed data sets.

The original Chl$a$ distributions exhibit a positive value of the skewness (between 4.9 and 8.1), due to the typical right tail of the log-normal distribution, which also causes a positive value of the kurtosis (pointed or peaked). The logarithmically transformed distributions present values very close to 0, i.e. the expected value of skewness and kurtosis for normal curves.

An extensive description of the different statistical indicators used in the framework of the comparison of ocean colour products can be found in Djavidnia et al. (2006, p.14). In the present report, the differences between Chl$a$ derived from remote sensing (Chl_S) and in situ (Chl_I) are assessed using the root mean square relative difference (Δ), the bias (δ) and the coefficient of determination (r^2), defined as:

$$\Delta = \sqrt{\frac{1}{N}\sum (x-y)^2} \tag{2.4}$$

$$\delta = \frac{1}{N}\sum (x-y) \tag{2.5}$$

$$r^2 = \left(\frac{N\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - \left(\sum x\right)^2}\sqrt{n\sum y^2 - \left(\sum y\right)^2}} \right)^2 \tag{2.6}$$

where x = $\log(\mathrm{Chl}_S)$ and y = $\log(\mathrm{Chl}_I)$. The percentage of cases where Chl_S is greater than Chl_I is also computed, and displayed in the scatter plots as *Up*. Note that in the scatter plot figures Δ is represented multiplied by 100, in order to make it more readable (value always below the unit).

Chapter 3

Satellite and in situ data comparison

3.1 Introduction

At this stage of the analysis, Chl$a$ estimates from remote sensing (Chl$_S$) and in situ (Chl$_I$) are compared on the global scale, separately for each in situ data set. Chl$_S$ is extracted from SeaWiFS and MODIS L3 daily images (obtained after the latest NASA reprocessing), and associated with Chl$_I$, into a match-up. When multiple in situ measurements occur for the same day and at the same location (i.e., within the same satellite grid point), the average value is computed and a single match-up is retained.

3.2 SeaWiFS Validation Set

After considering coincident and collocated NODC and SeaWiFS data, 654 match-ups are found. Most of them are in the northwest Atlantic coastal regions (Gulf of Maine and Caribbean Province*), European regional seas (Mediterranean, North Sea and Baltic Sea), in the Pacific Ocean, close to Japan and North America West Coast (including the California upwelling coastal province) (see Figure 3.1a). With respect to the other data sets (see Table 3.1 for a summary) NODC exhibits an average bias δ relatively low (+0.025), while the Δ is higher (0.309) and the coefficient of determination lower (0.655). As it can be seen in Figure 3.1b), for low values of Chl$a$ concentration, the satellite estimates can be much higher than the in situ measurements.

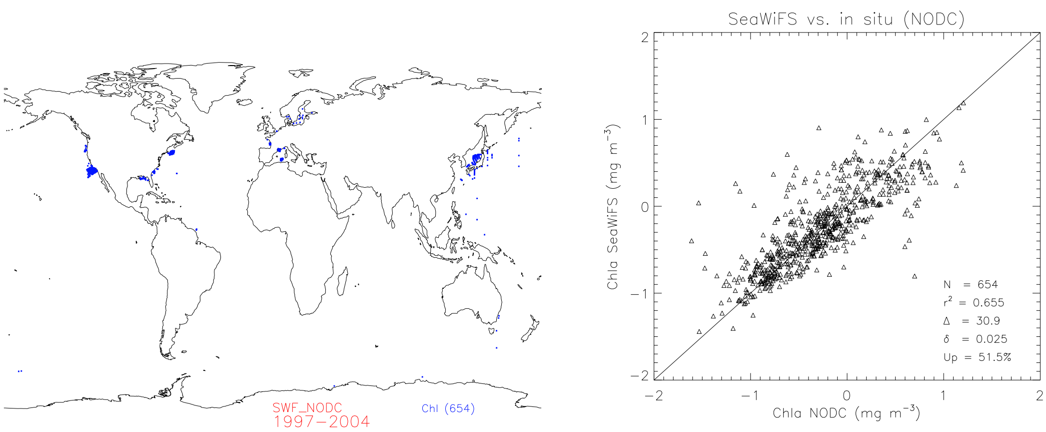

Figure 3.1: SeaWiFS-NODC match-ups: position (a) and scatter plot (b)

*Ecological provinces are defined in Section 4

The very large number of observations contained in the SeaBASS data set and their geographic distribution, yields a better coverage of the open ocean areas when we consider the SeaWiFS-SeaBASS match-ups (see Figure 3.2). Out of the 2417 FLUO and 321 HPLC collocated measurements, many are situated in the Central Pacific (Pacific Equatorial Divergence Province) and in the Austral Polar Province, close to the Antarctic Peninsula. Satellite derived Chl$a$ over-estimates the in situ measurements (δ is +0.073 for Chl_{FLUO} and +0.068 for Chl_{HPLC}), the latter measures having a higher Δ (0.296 vs. 0.270) and a lower coefficient of determination (0.75 vs. 0.78).

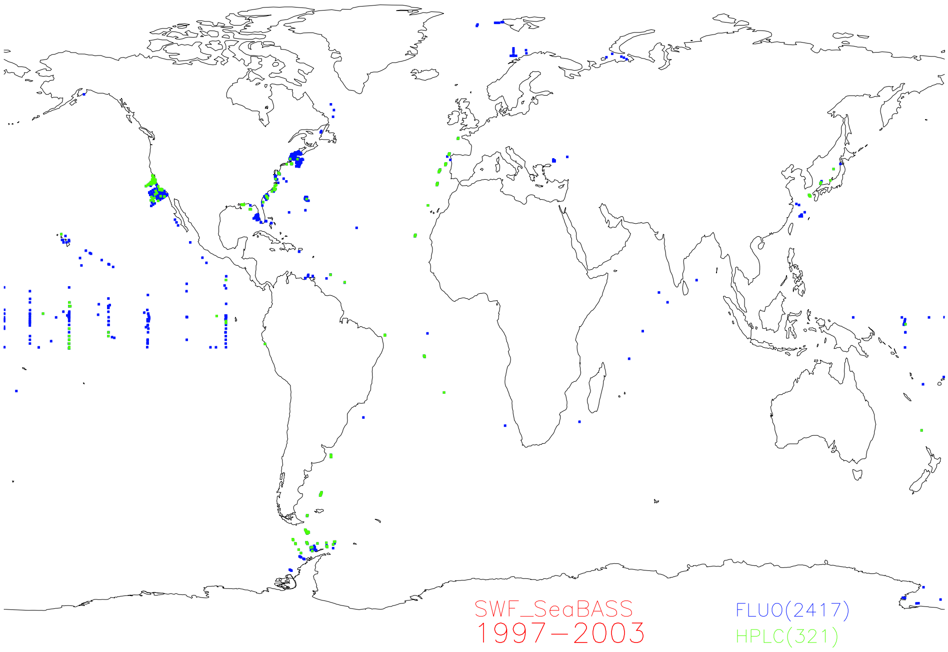

Figure 3.2: SeaWiFS-SeaBASS match-ups: position

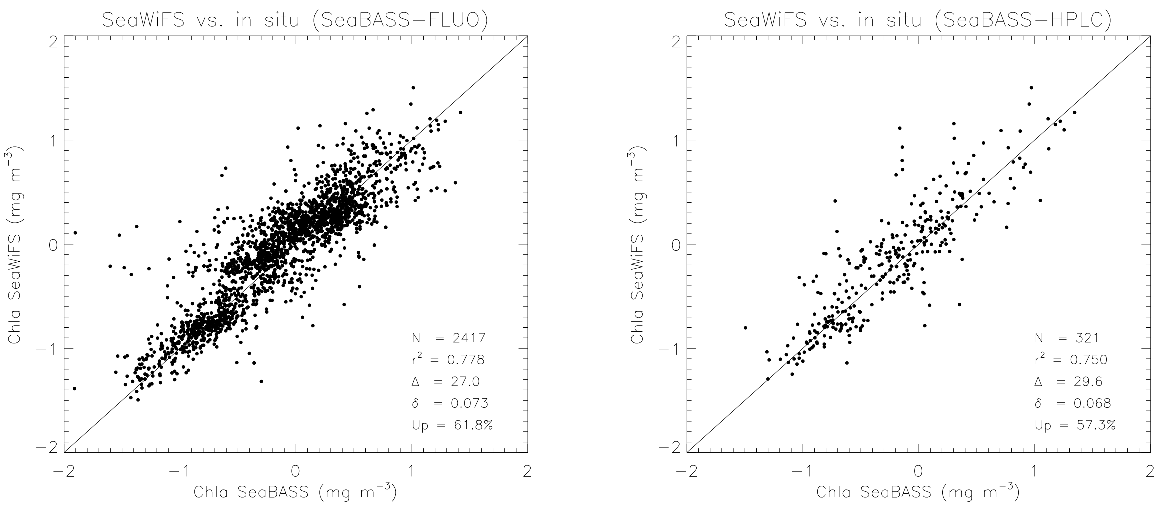

Figure 3.3: SeaWiFS-SeaBASS match-ups: scatter plot of FLUO (a) and HPLC (b) measures

Only 536 match-ups are found by associating SeaWiFS derived Chl$a$ with JODC observations (see Figure 3.4a), which are in the North Pacific Tropical Provinces, and particularly along the Japanese coasts (KURO, NPTW and WARM provinces and Japan Sea). On average Chl$_I$ is, again, lower than SeaWiFS estimates (δ is +0.059); the coefficient of determination is lower than for SeaBASS (0.686), and Δ is higher, 0.313.

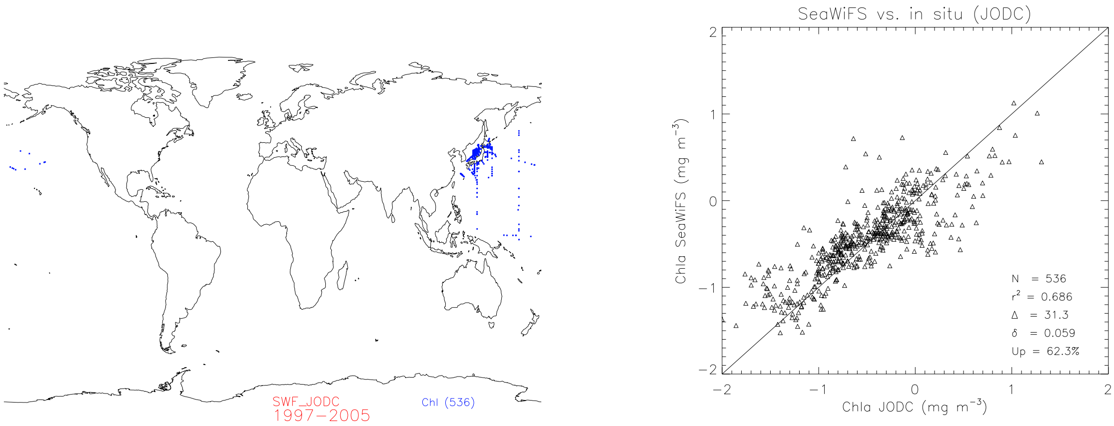

Figure 3.4: SeaWiFS-JODC match-ups: position (a) and scatter plot (b)

Table 3.1 summarizes the comparison of SeaWiFS Chl$a$ derived values and in situ measurements, coming from the different data sets, in terms of number of match-ups, RMSD (Δ), bias (δ) and coefficient of determination (refer to paragraph 2.4 for the definition of these variables).

Dataset	N.	Δ	δ	r^2
NODC	654	0.309	0.025	0.655
SeaBASS-FLUO	2417	0.270	0.073	0.778
SeaBASS-HPLC	321	0.296	0.068	0.750
JODC	536	0.313	0.059	0.686

Table 3.1: SeaWiFS matchup statistics.

3.3 MODIS Validation Set

MODIS L3 Chl*a* products are available since mid 2002, providing therefore a limited overlap with the NODC data set. Only 82 match-ups are identified, all located in the California Coastal Province - CCAL (see Figure 3.3). Chl*a* values are mainly in the range 0.1 to 1.0 mg m^{-3}, and the agreement between in situ and remote sensing values is very good, with a δ as low as -0.021, a limited RMSD (Δ=0.182) and a high determination coefficient (r^2=0.805).

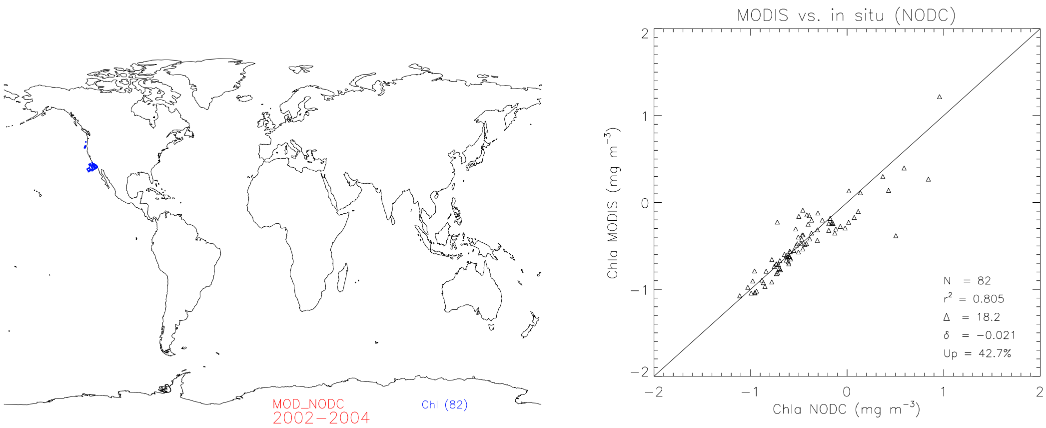

Figure 3.5: MODIS-NODC match-ups: position (a) and scatter plot (b)

A total of 257 match-ups is extracted from the SeaBASS data set (146 FLUO and 111 HPLC records), and most of them are related to measurement campaigns on the western American Coast, within the CCAL province. The MODIS product generally overestimates Chl*a* concentration with respect to in situ values, both for FLUO and HPLC (in 63% and 76% of the cases), leading to a significant bias (δ_{FLUO}=+0.111 and δ_{HPLC}=+0.185). The dispersion of the match-up points is also more evident for the HPLC than for FLUO methodology, both in terms of Δ (0.349 vs. 0.309) and r^2 (0.674 vs. 0.783).

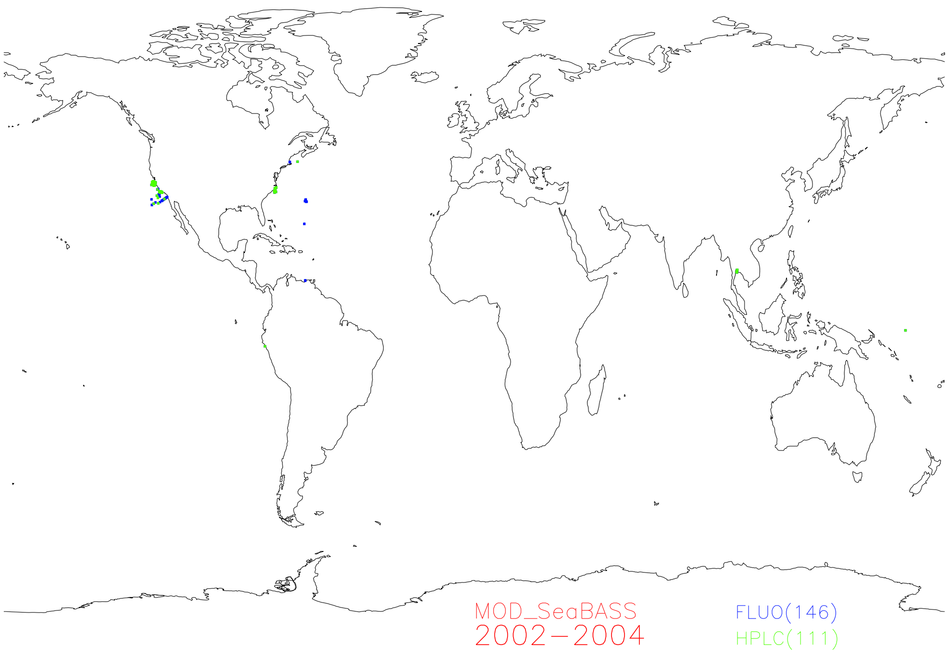

Figure 3.6: MODIS-SeaBASS match-ups: position

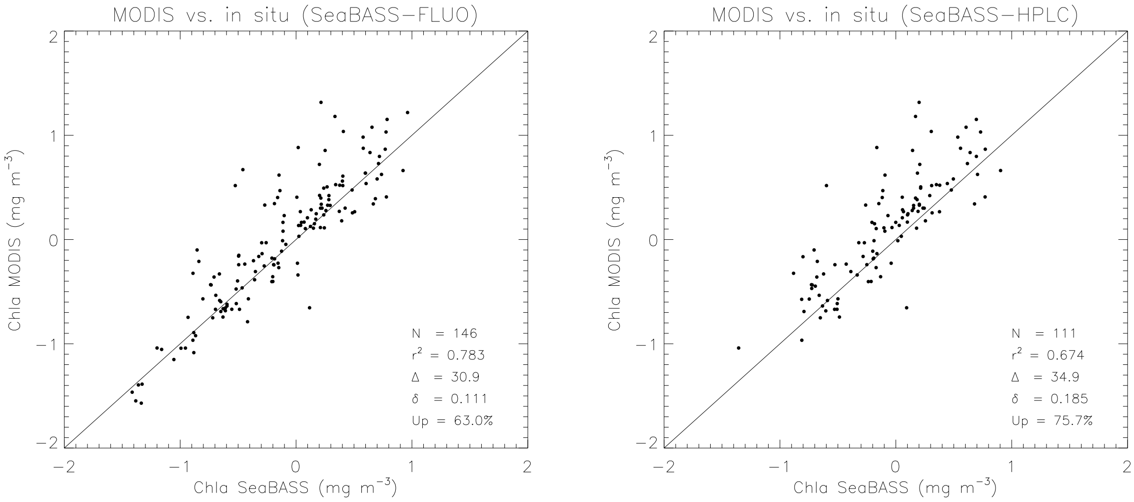

Figure 3.7: MODIS-SeaBASS match-ups: scatter plot for FLUO (a) and HPLC (b) measurements

A similar result is found by comparing MODIS Chl$a$ with JODC measurements (see Figure 3.3a and b). In situ Chl$a$ has a large range of variation (from 0.01 to more than 10. mg m^{-3}), and is on average over-estimated by MODIS ($\delta = +0.057$).

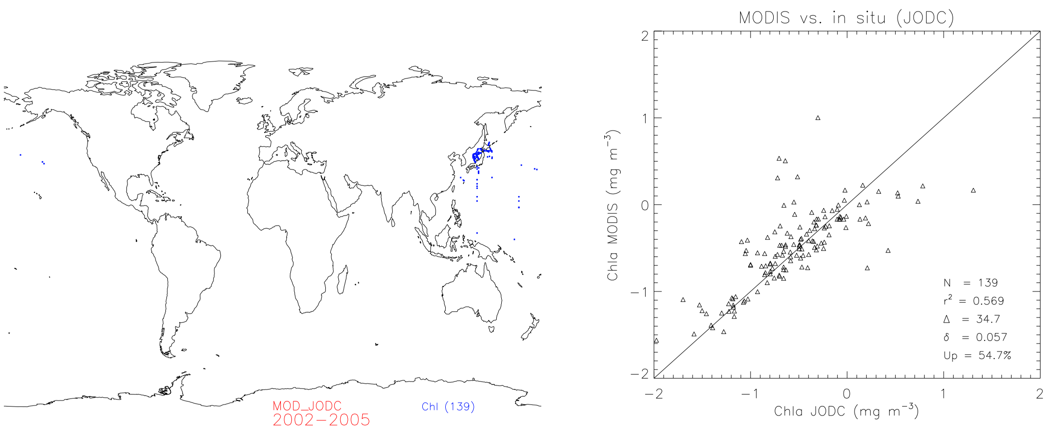

Figure 3.8: MODIS-JODC match-ups: position (a) and scatter plot (b)

Table 3.2 represents the statistics of the MODIS match-ups, and can be compared with Table 3.1.

Dataset	N. match-ups	RMSD	bias (log)	r^2
NODC	82	0.182	-0.021	0.805
SeaBASS-FLUO	146	0.309	0.111	0.783
SeaBASS-HPLC	111	0.349	0.185	0.674
JODC	139	0.347	0.057	0.569

Table 3.2: MODIS match-up statistics.

Chapter 4

Regional analysis

4.1 Data set merging

NODC, SeaBASS and JODC match-ups are merged into two series, for SeaWiFS and MODIS products, which do not contain duplicated measurements, i.e. observations present in both NODC and JODC data sets (see paragraph 2.3). As done previously for the single data sets, location and statistics of the match-ups are analysed (see Figure 4.1 and 4.2). The SeaWiFS 'merged' validation set contains 3778 match-ups, showing an average overestimate of Chl_S vs. Chl_I in about 60% of the cases ($\delta = +0.06$). The average Δ is found equal to 0.286.

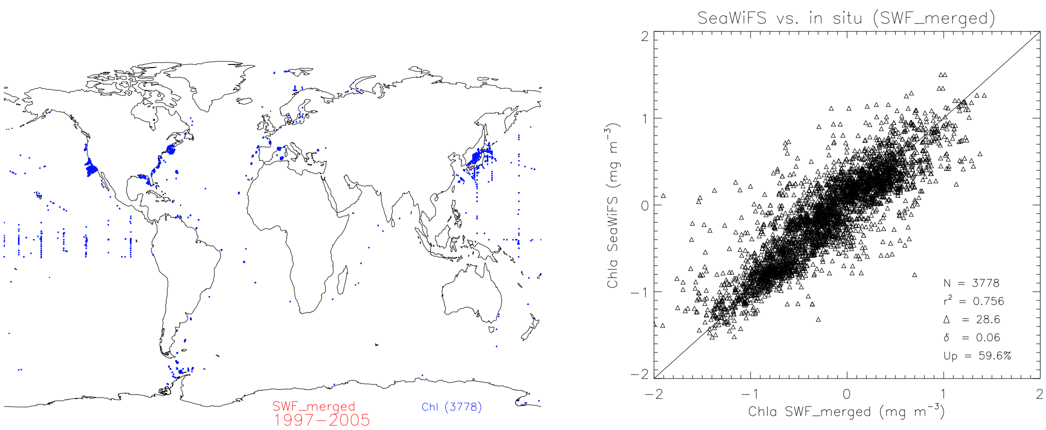

Figure 4.1: SeaWiFS match-ups - merged data set (NODC/SeaBASS/JODC): position (a) and scatter plot (b)

Figure 4.2 shows the equivalent results for MODIS, obtained with 478 match-up points: the statistics on this smaller data set are slightly degraded, with larger bias (δ equal to +0.09) and Δ (0.313).

The merged data sets are used in the following as a basis for assessing the agreement of Chl_S and Chl_I on regional scale, and depending on the basin bottom depth.

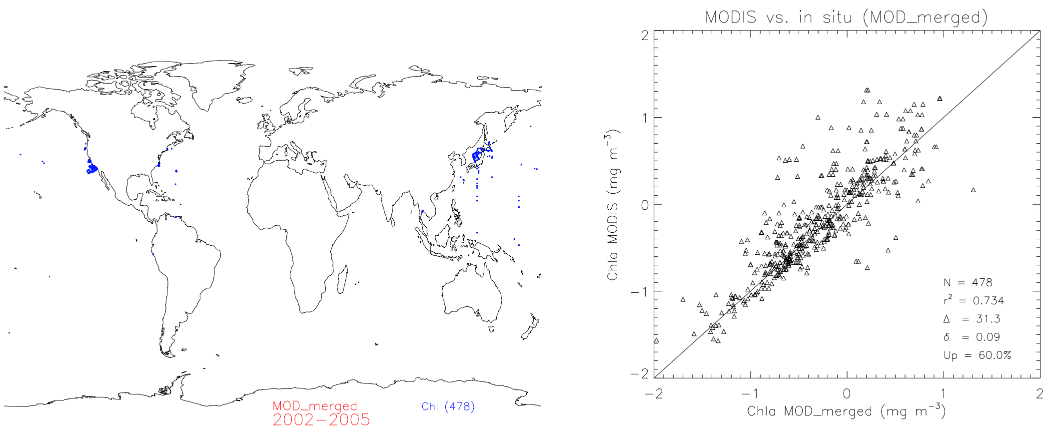

Figure 4.2: MODIS match-ups - merged data set (NODC/SeaBASS/JODC): position (a) and scatter plot (b)

4.2 Regional Analysis

The match-ups described in Section 4.1 (i.e. after merging of the in situ data sets) are identified and gathered according to the oceanographic provinces proposed by Longhurst (Longhurst 1998). Statistics of δ and Δ are computed for provinces having at least 8 match-ups, and reported in Table 4.1 and 4.2. The Longhurst provinces that exhibit a high number of match-ups (at least 50) are analysed in more detail (see maps and scatter plots in Figure 4.6 to 4.14 for SeaWiFS and Figure 4.15 to 4.16 for MODIS). General results are provided by Figures 4.4 and 4.5, and Tables 4.1 and 4.2.

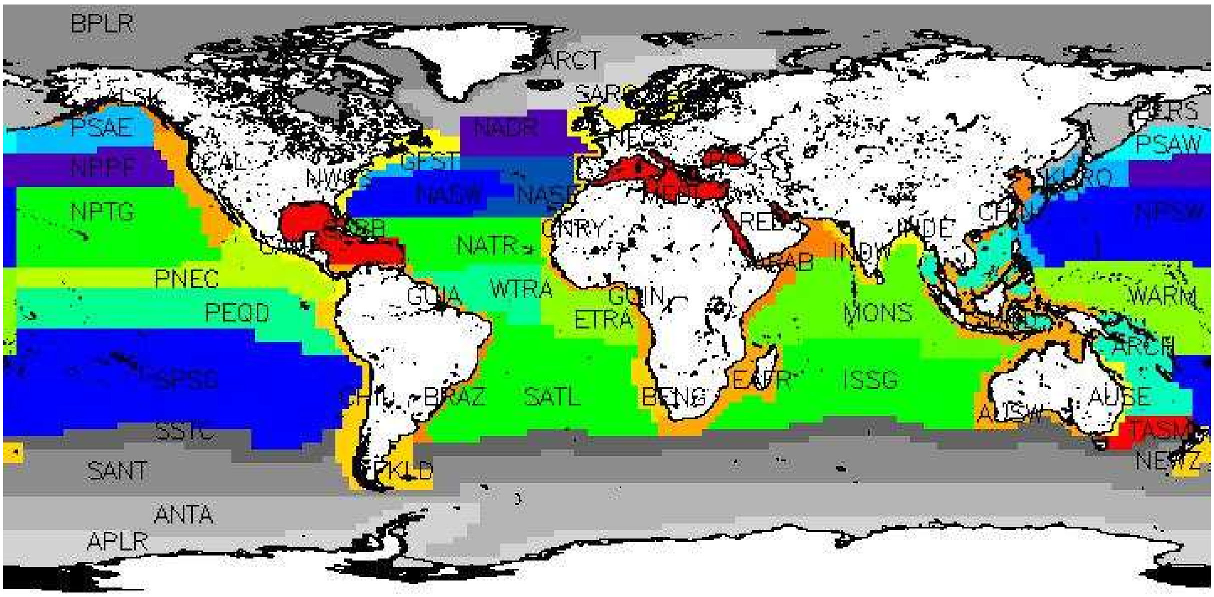

Figure 4.3: Longhurst oceanographic Provinces

Province	N.	δ	Δ	r^2	Up
Arctic Ocean					
ARCT	9	0.41	0.518	0.156	88.90
SARC	11	-0.13	0.166	0.718	9.10
Pacific Ocean					
PSAW	9	-0.04	0.148	0.221	44.40
NPPF	8	-0.05	0.116	0.928	37.50
KURO	459	0.05	0.300	0.595	61.20
NPTW	57	0.17	0.337	0.163	68.40
NPTE	25	-0.01	0.154	0.271	48.00
WARM	47	-0.03	0.152	0.742	40.40
PNEC	14	0.04	0.130	0.799	50.00
PEQD	146	-0.02	0.125	0.605	43.20
SPSG	11	-0.23	0.421	0.226	18.20
CCAL	744	0.01	0.233	0.787	47.60
Atlantic Ocean					
NWCS	1607	0.10	0.302	0.543	67.80
NECS	33	0.41	0.459	0.710	0.00
NASW	51	0.08	0.208	0.656	64.70
NASE	48	0.11	0.171	0.845	79.20
NATR	14	0.05	0.145	0.906	64.30
CARB	151	0.16	0.288	0.735	77.50
GUIA	25	0.31	0.443	0.758	88.00
SATL	12	0.01	0.089	0.637	66.70
BRAZ	15	0.27	0.311	0.307	0.00
Southern Ocean					
SANT	19	-0.10	0.145	0.820	26.30
ANTA	35	-0.17	0.261	0.802	22.90
APLR	55	-0.28	0.433	0.482	16.40
FKLD	20	-0.23	0.288	0.900	0.00
Regional Seas					
BALT	12	0.07	0.325	0.046	83.30
MEDI	52	-0.05	0.256	0.651	40.40
CHIN	12	0.03	0.249	0.657	41.70

Table 4.1: SeaWiFS regional analysis.

Province	N.	δ	Δ	r^2	Up
Pacific Ocean					
KURO	111	0.05	0.376	0.347	55.00
NPTW	14	0.09	0.217	0.329	64.30
CCAL	280	0.07	0.244	0.780	61.10
Atlantic Ocean					
NWCS	34	0.28	0.520	0.608	73.50
NASW	9	-0.12	0.178	0.831	11.10
Regional Seas					
SUND	8	0.60	0.695	0.732	0.00

Table 4.2: MODIS regional Analysis.

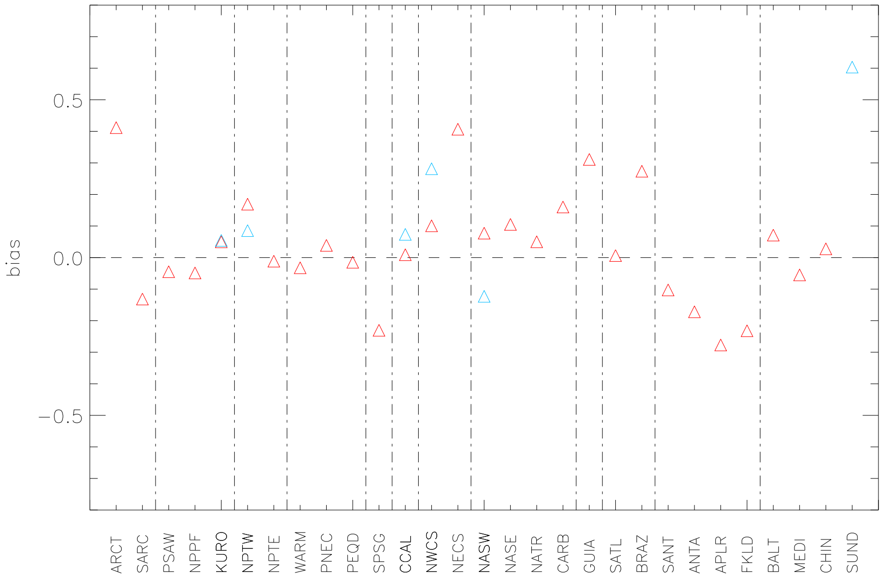

Figure 4.4: Regional analysis of bias (δ) for SeaWiFS (red) and MODIS (blue)

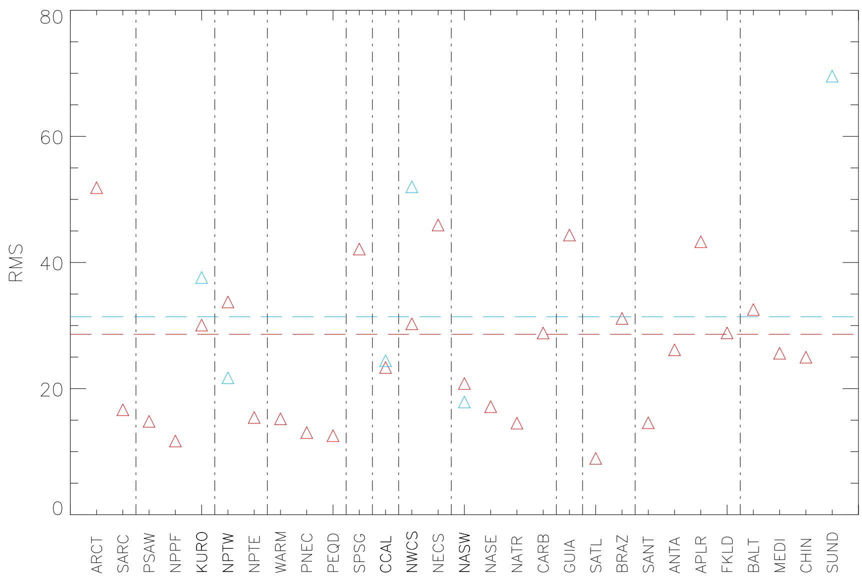

Figure 4.5: Regional analysis of RMSD (Δ) for SeaWiFS (red) and MODIS (blue)

4.2.1 SeaWiFS

Figures 4.4 and 4.5 and Table 4.1 give the validation results by province. Not surprisingly, the regional statistics show a variety of results with respect to the global values. Overall, the most represented regions correspond to the north American coastal waters, included in the provinces NWCS and CARB on the Atlantic side (Figures 4.10 and 4.12), and CCAL (Figure 4.9) on the Pacific coast. The latter show validation statistics that are better than the global average (δ and Δ equal to +0.01 and 0.23, respectively). The small set available in the Sargasso Sea (NASW) provides encouraging results, with Δ of 0.21 (Figure 4.11), as do the results of other subtropical/tropical Atlantic waters (NASE and NATR). With the support of the JODC, the KURO province has a large validation set (N=459, Figure 4.6), with statistics close to the global average (δ and Δ equal to +0.05 and 0.30, respectively). The Equatorial Pacific, as represented by NPTW, NPTE, WARM, PNEC and PEQD, is associated with a relatively large number of match-ups, especially in terms of open ocean waters (Figures 4.7, 4.8). NPTW aside, the discrepancies are found lower than the global averages (Δ of approximately 0.14). Southern Ocean waters (FKLD, SANT, ANTA, APLR) show varying results (see also Figure 4.13), the main feature being the under-estimate by the SeaWiFS products (δ between -0.10 and -0.28). Statistics for some marginal seas indicate rather large discrepancies, for instance for NECS, BALT or GUIA. In general, this discussion is completely coherent with the results presented by Gregg and Casey (2004). The major exception is for the Mediterranean basin, that show here satisfactory statistics. This likely results from a different data set, located in the Ligurian Sea and close to the northwest African coasts. Larger discrepancies with a positive bias have been found by various studies (e.g., Gregg and Casey (2004), Mélin et al. (2007b), Volpe, G. and Santoleri, R. and Vellucci, V. and Ribera d'Acalá, M. and Marullo, S., and D'Ortenzio, F. (2007)).

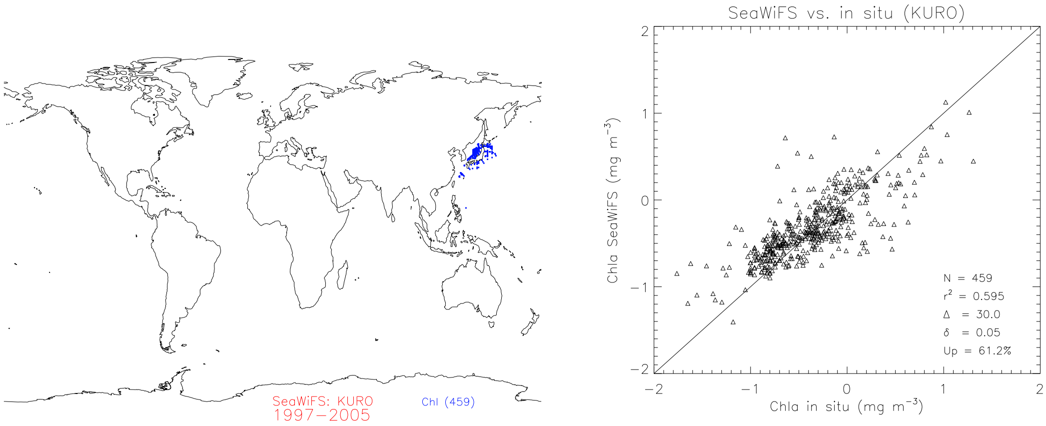

Figure 4.6: SeaWiFS match-ups for KURO province: position (a) and scatter plot (b)

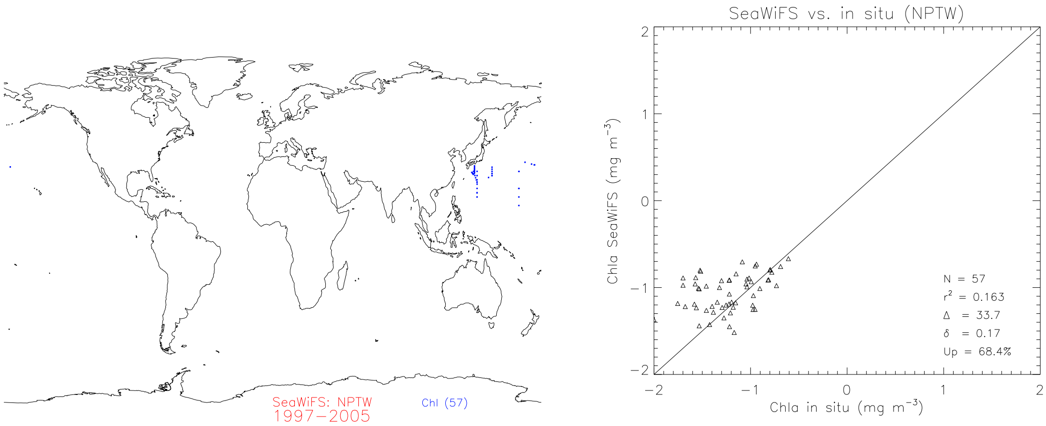

Figure 4.7: SeaWiFS match-ups for NPTW province: position (a) and scatter plot (b)

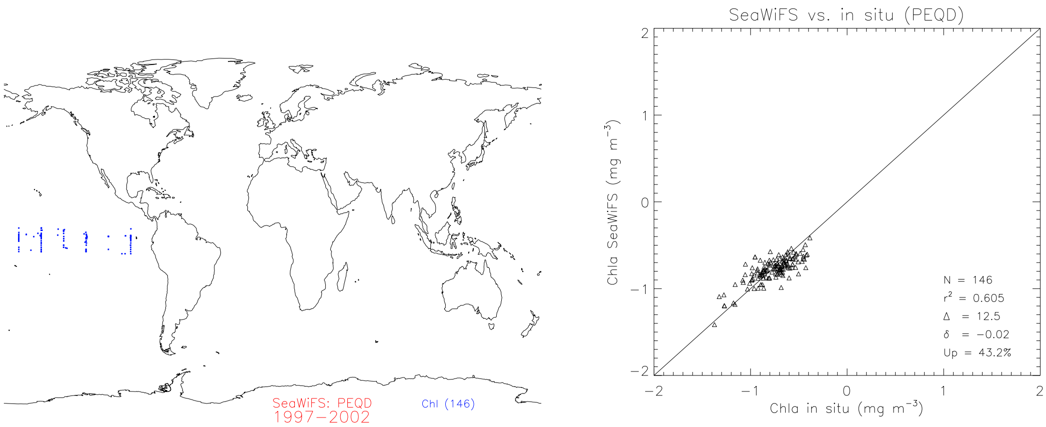

Figure 4.8: SeaWiFS match-ups for PEQD province: position (a) and scatter plot (b)

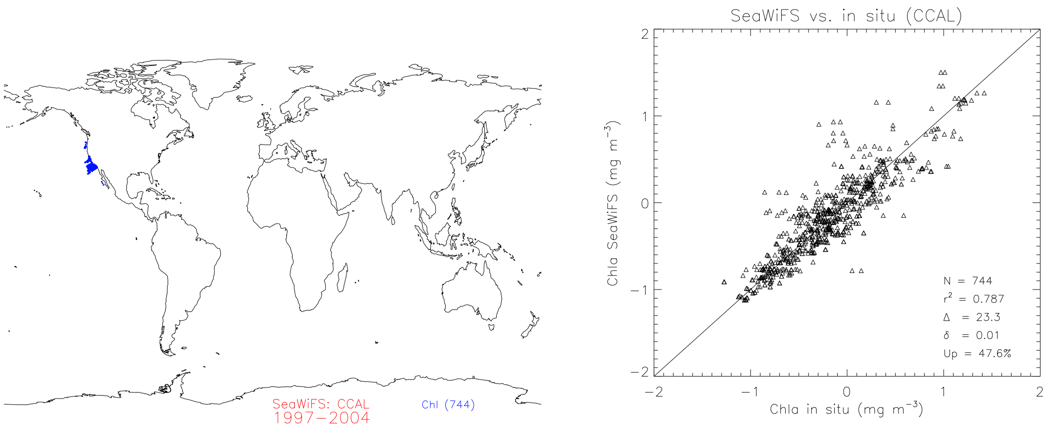

Figure 4.9: SeaWiFS match-ups for CCAL province: position (a) and scatter plot (b)

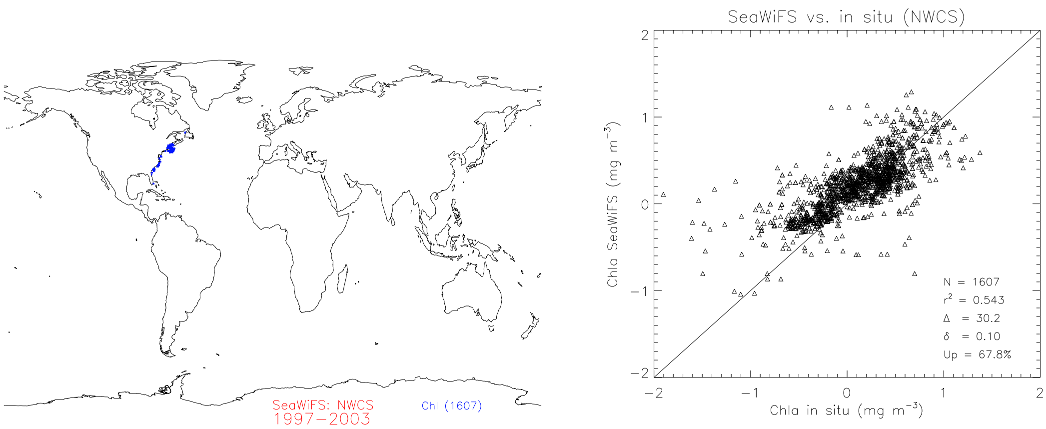

Figure 4.10: SeaWiFS match-ups for NWCS province: position (a) and scatter plot (b)

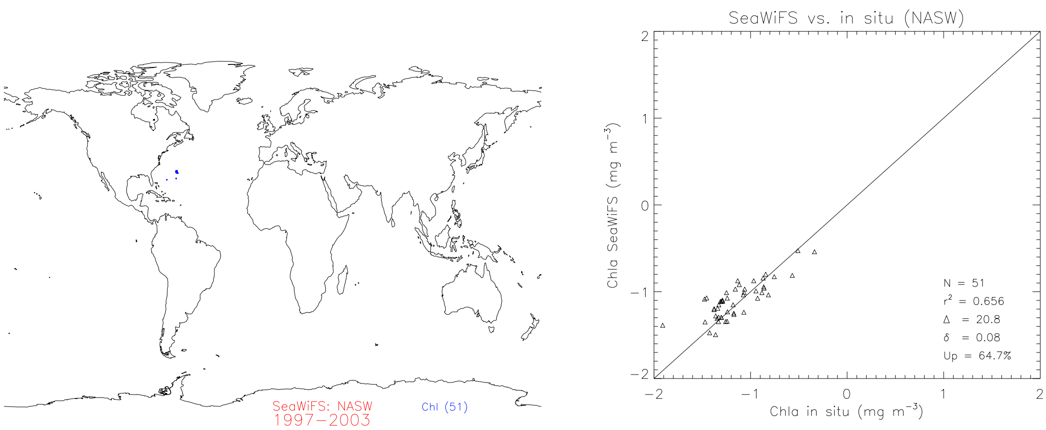

Figure 4.11: SeaWiFS match-ups for NASW province: position (a) and scatter plot (b)

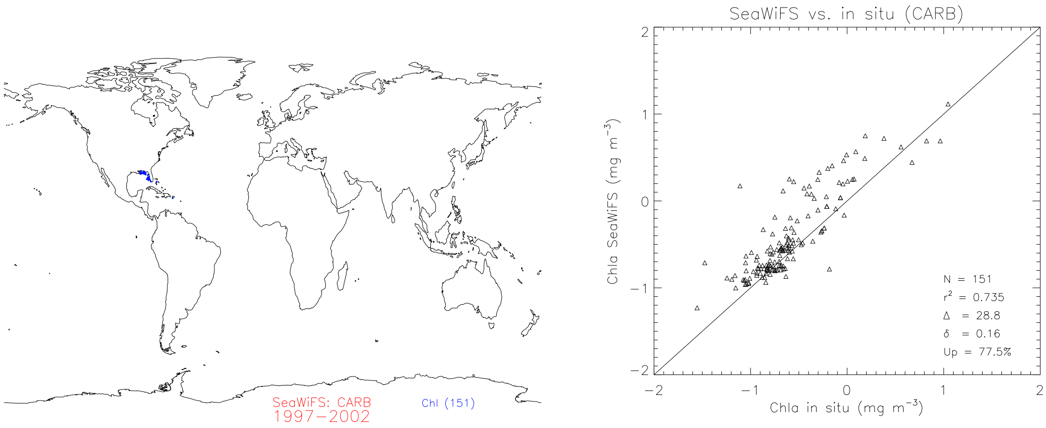

Figure 4.12: SeaWiFS match-ups for CARB province: position (a) and scatter plot (b)

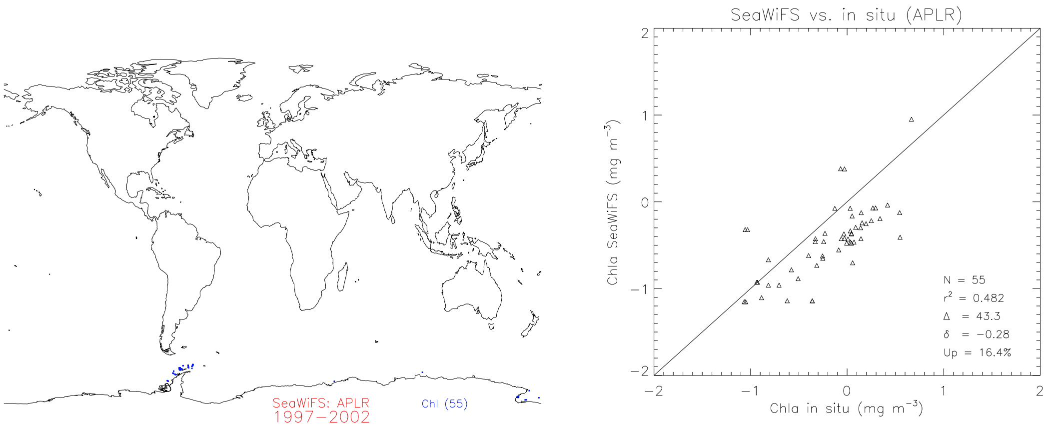

Figure 4.13: SeaWiFS match-ups for APLR province: position (a) and scatter plot (b)

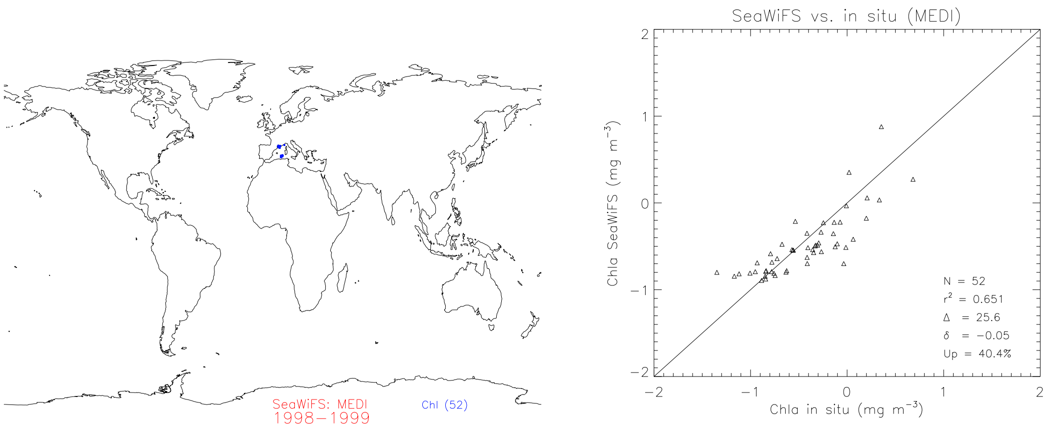

Figure 4.14: SeaWiFS match-ups for MEDI province: position (a) and scatter plot (b)

4.2.2 MODIS

The match-ups available for MODIS (Table 4.2) are essentially in the KURO and CCAL provinces (Figures 4.15 and 4.16). The agreement with in situ data is found closer for CCAL (Δ of 0.24) even though the bias δ is slightly higher (+0.07).

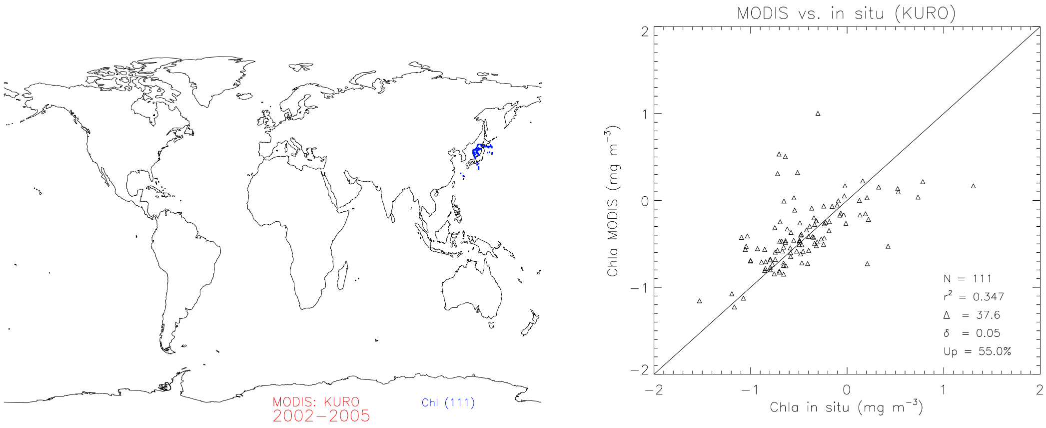

Figure 4.15: MODIS match-ups for KURO province: position (a) and scatter plot (b)

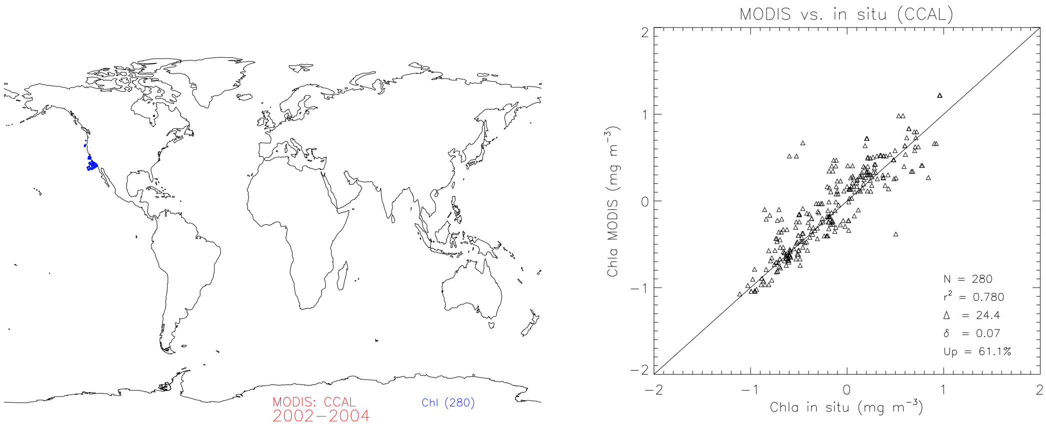

Figure 4.16: MODIS match-ups for CCAL province: position (a) and scatter plot (b)

4.3 Coastal areas and open ocean

In this section we consider the comparison of Chl$a$ satellite/in situ on the basis of a partition coastal/shelf waters and open ocean, using a threshold on bathymetry. Match-ups located in open ocean (bottom depth larger than 200 m) show, for both satellites, a better agreement than the ones in coastal areas (see Figures 4.17 to 4.20). The SeaWiFS data set contains 2139 match-ups in open ocean, with a reduced bias (+0.03 vs. +0.11), a higher determination coefficient (0.751 vs. 0.613) and a significantly reduced RMSD (0.265 vs. 0.311).

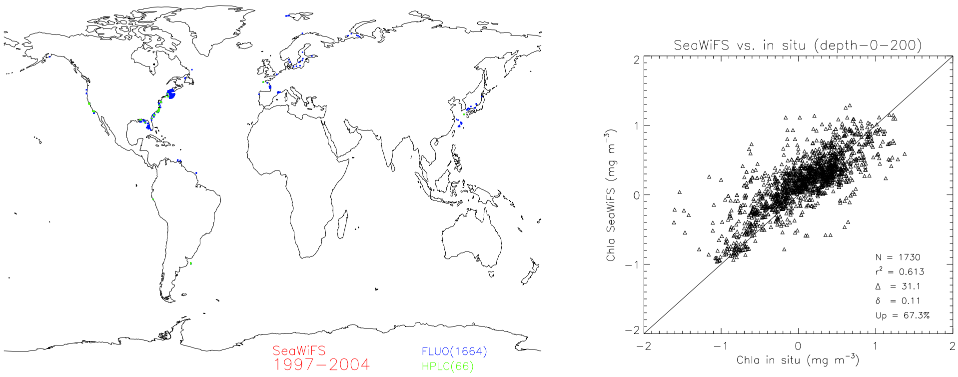

Figure 4.17: SeaWiFS match-ups - depth < 200: position (a) and scatter plot (b)

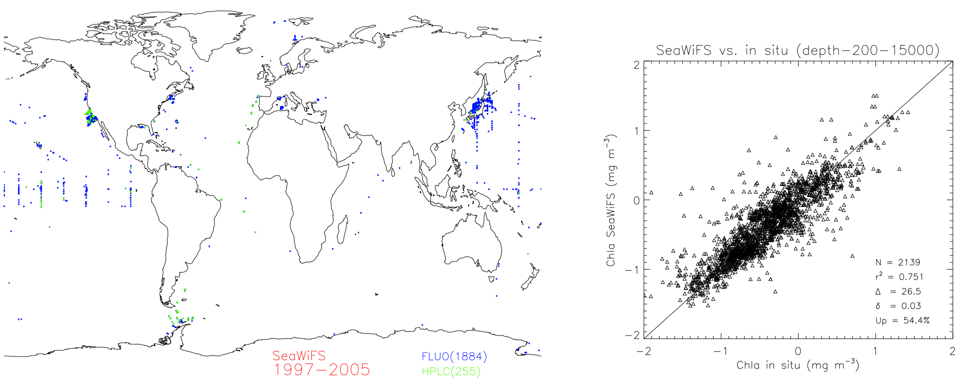

Figure 4.18: SeaWiFS match-ups - depth > 200: position (a) and scatter plot (b)

MODIS match-ups confirm a better result for open ocean areas (N=412) than for coastal regions (N=66) (see Figures 4.17 and 4.18). A particularly poor agreement is found in the latter case, with a bias as high as 0.25, RMSD of 0.491 and a determination coefficient of only 0.452.

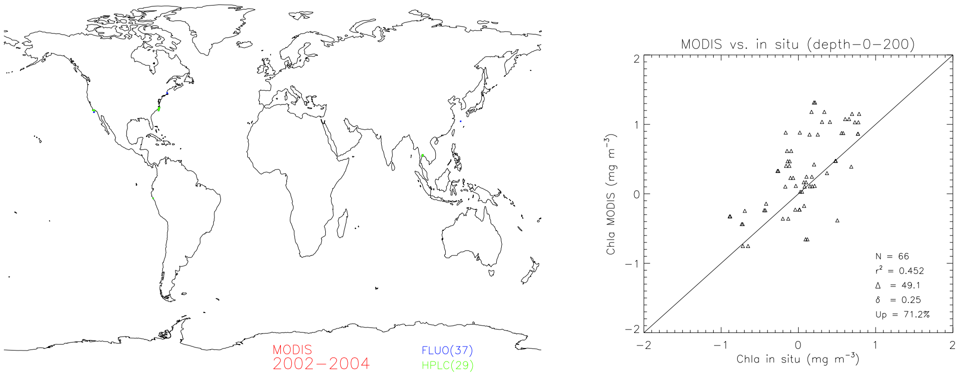

Figure 4.19: MODIS match-ups - depth < 200: position (a) and scatter plot (b)

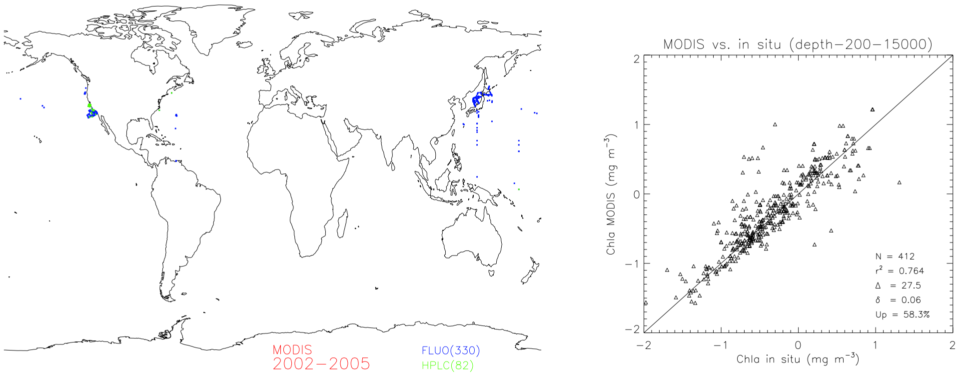

Figure 4.20: MODIS match-ups - depth > 200: position (a) and scatter plot (b)

Chapter 5

Conclusions

The conclusions of this study can be summarized in a few key points. First, they appear completely coherent with a previous study by Gregg and Casey (2004), both from the point of view of the global statistics than for most of the regional results. Secondly, the SeaWiFS global average of RMS difference (for log-transformed values) shows an uncertainty of 0.29, lower than the value of 0.35 often considered as the objective for Chl$a$ distributions. Δ for MODIS is slightly higher (0.31), a difference likely partly due to a smaller statistical basis. An important objective of this work, that goes beyond the scope of the present report, was to develop the validation procedure and protocols for further analyses regularly reviewing validation results to take into account successive reprocessing and other sensors, as well as including additional in situ data sets.

Acknowledgements

This work was supported by the MERSEA (Marine Environment and Security for the European Area) EU-FP6 Integrated Project, in the framework of the Ocean and Marine services element of GMES (Global Monitoring for Environment and Security).

The authors would like to acknowledge the Ocean Biology Processing Group and the Distributed Active Archive Center (Code 902) at the Goddard Space Flight Center of the U.S. National Aeronautics and Space Administration for the production and the distribution of the SeaWiFS and MODIS-Aqua data, and all accompanying software and documentation.

The in-situ 'NODC' dataset is produced and distributed by the National Oceanographic Data Center (NODC), operated by the NOAA. SeaBASS in-situ measurements are prepared and supplied by the NASA Ocean Biology Processing Group (OBPG). 'JODC' dataset is retrieved from the Japan Oceanographic Data Center, in the Hydrographic and Oceanographic Department, Japan Coast Guard.

List of Acronyms

DAAC: Distributed Active Archive Centre
ESA: European Space Agency

GSFC: Goddard Space Flight Centre

JRC: Joint Research Centre (European Commission)

MBR: Maximum Band Ratio

MODIS: MOderate Resolution Imaging Spectrometer

NASA: National Aeronautics and Space Administration

SeaWiFS: Sea-viewing Wide Field-of-view Sensor

List of Figures

2.1 NODC data set: location of measurement stations. . . . 3
2.2 NODC dataset - C$_{sat}$ derived from vertical integration vs. C_{depth_0} . . . 4
2.3 SeaBASS data set: location of measurement stations. . . . 5
2.4 SeaBASS data set (FLUO and HPLC) - C$_{sat}$ derived from vertical integration vs. C_{depth_0} 6
2.5 JODC dataset: location of measurements stations. . . . 7
2.6 JODC data set - C$_{sat}$ derived from vertical integration vs. C_{depth_0} . . . 8

3.1 SeaWiFS-NODC match-ups: position (a) and scatter plot (b) . . . 10
3.2 SeaWiFS-SeaBASS match-ups: position . . . 11
3.3 SeaWiFS-SeaBASS match-ups: scatter plot of FLUO (a) and HPLC (b) measures . . 11
3.4 SeaWiFS-JODC match-ups: position (a) and scatter plot (b) . . . 12
3.5 MODIS-NODC match-ups: position (a) and scatter plot (b) . . . 13
3.6 MODIS-SeaBASS match-ups: position . . . 14
3.7 MODIS-SeaBASS match-ups: scatter plot for FLUO (a) and HPLC (b) measurements 15
3.8 MODIS-JODC match-ups: position (a) and scatter plot (b) . . . 15

4.1 SeaWiFS match-ups - merged data set (NODC/SeaBASS/JODC): position (a) and scatter plot (b) . . . 16
4.2 MODIS match-ups - merged data set (NODC/SeaBASS/JODC): position (a) and scatter plot (b) . . . 17
4.3 Longhurst oceanographic Provinces . . . 18
4.4 Regional analysis of bias (δ) for SeaWiFS (red) and MODIS (blue) . . . 21
4.5 Regional analysis of RMSD (Δ) for SeaWiFS (red) and MODIS (blue) . . . 21
4.6 SeaWiFS match-ups for KURO province: position (a) and scatter plot (b) . . . 23
4.7 SeaWiFS match-ups for NPTW province: position (a) and scatter plot (b) . . . 23
4.8 SeaWiFS match-ups for PEQD province: position (a) and scatter plot (b) . . . 23
4.9 SeaWiFS match-ups for CCAL province: position (a) and scatter plot (b) . . . 24
4.10 SeaWiFS match-ups for NWCS province: position (a) and scatter plot (b) . . . 24
4.11 SeaWiFS match-ups for NASW province: position (a) and scatter plot (b) . . . 24
4.12 SeaWiFS match-ups for CARB province: position (a) and scatter plot (b) . . . 25
4.13 SeaWiFS match-ups for APLR province: position (a) and scatter plot (b) . . . 25
4.14 SeaWiFS match-ups for MEDI province: position (a) and scatter plot (b) . . . 25
4.15 MODIS match-ups for KURO province: position (a) and scatter plot (b) . . . 26
4.16 MODIS match-ups for CCAL province: position (a) and scatter plot (b) . . . 26
4.17 SeaWiFS match-ups - depth $<$ 200: position (a) and scatter plot (b) . . . 27
4.18 SeaWiFS match-ups - depth $>$ 200: position (a) and scatter plot (b) . . . 27
4.19 MODIS match-ups - depth $<$ 200: position (a) and scatter plot (b) . . . 28
4.20 MODIS match-ups - depth $>$ 200: position (a) and scatter plot (b) . . . 28

List of Tables

1 Document Change Record 4

2.1 Number of NODC measurements per year. 4
2.2 Number of SeaBASS measurement stations - FLUO and HPLC 5
2.3 Number of SeaBASS measurements per year. 6
2.4 Number of JODC measurements per year. 7
2.5 Skewness and kurtosis of the original and log-transformed data sets. 9

3.1 SeaWiFS matchup statistics. 12
3.2 MODIS match-up statistics. 15

4.1 SeaWiFS regional analysis. 19
4.2 MODIS regional Analysis. 20

References

Bailey, S. and P. Werdell (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *RSE 102*, 12–23.

Campbell, J. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *JGR 100*, 13237–13254.

Djavidnia, S., F. Melín, and N. Hoepffner (2006). Analysis of MultiSensor Global and Regional Ocean Colour Products. Technical Report EUR 22238 EN, JRC.

Gordon, H. and D. Clark (1980). Remote sensing optical properties of stratified ocean: an improved interpretation. *AO 19*, 3428–3430.

Gregg, W. and N. Casey (2004). Global and regional evaluation of the seawifs chlorophyll data set. *RSE 93*, 463–479.

Hooker, S., W. Esaias, G. Feldman, W. Gregg, and C. McClain (1992). An overview of seawifs and ocean color. Technical Report 1992-104566, vol. 1, NASA.

Longhurst, A. (1998). *Ecological Geography of the Sea*. San Diego: Academic Press.

Mélin, F. and G. Zibordi (2003). An optically-based technique for producing merged spectra of water leaving radiances from ocean color. *Appl. Opt., 41*, 548–558.

Mélin, F., G. Zibordi, and J.-F. Berthon (2007a). Assessment of satellite ocean color products at a coastal site. remote sens. environ. *Remote Sens. Environ., 192–215*, 436–450.

Mélin, F., G. Zibordi, and J.-F. Berthon (2007b). Assessment of seawifs atmospheric and marine products for the northern adriatic sea. *IEEE Trans. Geosci. Remote Sens., 46*, 3856–3869.

Mélin, F., G. Zibordi, and S. Djavidnia (2007). Development and validation of a technique for merging satellite derived aerosol optical depth from seawifs and modis. *Remote Sens. Environ., 108*, 436–450.

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

Salomonson, V.V. and Barnes, W.L. and Maymon, P.W. and Montgomery, H.E. and Ostrow, H. (1989). MODIS: Advanced facility instrument for studies of the Earth as a system. *IEEE Trans. Geosci. Remote Sens., 27*, 145–152.

Smith, R. (1981). Remote sensing and depth distribution of ocean chlorophyll. *Mar. Ecol. Prog. Ser., 5*, 359–361.

Van Heukelem, L., C. Thomas, and P. Glibert (2002). Sources Of Variability In Chlorophyll Analysis By Fluorometry And High Performance Liquid Chromatography In a SIMBIOS Inter-Calibration Exercise. Technical Report TM-2002-02338-0, NASA.

Volpe, G. and Santoleri, R. and Vellucci, V. and Ribera d'Acalá, M. and Marullo, S., and D'Ortenzio, F. (2007). The colour of the Mediterranean Sea: Global versus regional bio-optical algorithms evaluation and implication for satellite chlorophyll estimates. *Remote Sens. Environ., 107*, 625–638.

Werdell, P.J. and Bailey, S.W. and Fargion, G.S. and Pietras, C. and Knobelspiesse, K.D. and Feldman, G.C. and McClain, C.R. (2003). Unique data repository facilitates ocean color satellite validation. *EOS Trans. Amer. Geophys. Union, 84*, 377.

Werdell, P.J., Bailey, S.W. and Pietras, C. and Knobelspiesse, K.D. and Feldman, G.C. and McClain, C.R. (2005). An improved in situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sens. Environ., 98*, 122–140.

Zibordi, G. and Mélin, F. and Berthon, J.-F. (2006). Comparison of SeaWiFS, MODIS and MERIS radiometric products at a coastal site. *Geophys. Res. Lett., 33*, 122–140.

European Commission

EUR 23357 EN - DG Joint Research Centre, Institute for Environment and Sustainability
Title: Assessment of Global Ocean Colour Products against In-situ Datasets
Authors: M.Clerici, F. Mélin, N. Hoepffner
Luxembourg: Office for Official Publications of the European Communities
2008 - 44 pp. - 21 x 29.7 cm
EUR - Scientific and Technical Research series; ISSN 1018-5593

Abstract

Ocean colour from satellite has given over the last two decades another dimension to ecosystem studies and marine biology, providing key information on the timing and spatial distribution of phytoplankton blooms, and the magnitude of primary production. Remote observations of ocean colour from space represent therefore a major tool directly related to the marine biogeochemical distributions and associated processes. Presently, the ocean colour community relies on a number of individual missions launched by various space agencies, and the differences resulting from this variety of missions need to be addressed in order to create consistent long term data records.

One of the goals of the European GMES Integrated Project MERSEA is to provide an accurate and consistent stream of ocean colour data. In this study we present an extensive comparative analysis of standard ocean colour products obtained from the latest operational global ocean colour sensors (SeaWiFS, MODIS-Aqua, MERIS, GLI, POLDER-2, OCTS and POLDER-1), on both global and regional scales. We analysed and compared monthly mean chlorophyll-a sea surface concentration (Chl$a$) and normalized water leaving radiance (nL_W) between July 2002 and June 2005.

For Chl$a$ on a global scale, SeaWiFS yields the highest geometric mean value whilst MERIS has the lowest values with percentage root mean square differences (Δ) between the two sensors in the order of 17% and a negligible percentage bias (δ). MODIS-Aqua and SeaWiFS are in better agreement with an average global Δ of approximately 13% and with an average δ of +6% with SeaWiFS higher than MODIS-Aqua. On a regional scale, the analysis undertaken on Longhurst's biogeographical provinces show a rather close agreement between SeaWiFS and MODIS-Aqua Chl$a$, with SeaWiFS slightly higher, and larger differences between these products and MERIS Chl$a$. Larger differences are also noticed for GLI and POLDER-2 in the short period where these products are available.

For the nL_W products, on a global scale MODIS-Aqua has the lowest signal at the blue end of the spectrum (412, 443 and 490 nm), with Δ's showing no seasonality and on average equal to 20%, 20% and 13% respectively with respect to SeaWiFS nL_W. MERIS yields the highest mean values at 412 nm with Δ's reaching 50% when compared with the equivalent product from SeaWiFS or MODIS-Aqua. Regional analysis shows diverse results with rather strong seasonalities for the different wavelengths.

An important outcome of the study is that the results of the inter-comparison analysis are variable with season and area, and that globally averaged statistics are not necessarily applicable on a regional basis. Thus the use of coincident ocean colour products at regional and seasonal scales needs to take this variability into account.