Ocean Colour Climate Change Initiative (OC_CCI) - Phase One

Uncertainty Characterisation Document (UCD)

Ref: D2.9

Date: 05/12/11

Issue: 1.2

For: ESA / ESRIN

Ref: AO-1/6207/09/I-LG

PML Plymouth Marine Laboratory

VEGA Helmholtz-Zentrum Geesthacht Centre for Materials and Coastal Research BROCKMANN CONSULT JRC EUROPEAN COMMISSION HYGEOS NERSC

This Page is Intentionally Blank

SIGNATURES AND COPYRIGHT

Project : **Ocean Colour Climate Change Initiative (OC_CCI) - Phase One**

Document Title: **Uncertainty Characterisation Document (UCD)**

Reference : **D2.9**

Issued : **05 December 2011**

Issue : **1.2**

Client : **ESA / ESRIN**

Authored :

[Dr Shubha Sathyendranath - PML] and [Dr Roland Doerffer – HZG]

Reviewed :

[Dr Bob Brewin – PML]

Approved :

[Dr Shubha Sathyendranath - PML]

Address : Plymouth Marine Laboratory (PML)

Prospect Place

The Hoe

Plymouth PL1 2DH

Document Change Log

Issue/ Revision	Date	Comment
0.1	01/08/11	Initial Template
1.0	01/11/11	First Issue
1.1	30/11/11	Updated for final issue
1.2	05/12/11	Updated Authored section to acknowledge HZG technical input

This Page is Intentionally Blank

TABLE OF CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

This Page is Intentionally Blank

1. INTRODUCTION

In this document, the methods used for identifying errors and uncertainties in the products of Ocean-Colour Climate Change Initiative (OC-CCI) are detailed. In selecting the metrics for quantifying errors and uncertainties, we have been guided by the following requirements:

1. The need to use metrics that have been commonly used in the literature, to facilitate comparisons with other work;

2. The need to separate random and systematic components of errors;

3. The requirement to map errors and uncertainties on a pixel-by-pixel basis;

4. The need to serve the modelling community as well as the Earth Observation community; and

5. The need to select metrics that are appropriate for the algorithms.

The algorithm suites that have been considered for atmospheric correction and for retrieval of atmospheric properties may be categorised broadly into two types: empirical and model-based. Certain methods of error characterisation are limited to model-based methods.

Bearing in mind that the in-water measurements of the properties are also prone to measurement errors (for example, associated with collection of water samples; handling of the water samples; and instrument errors), the term “uncertainties” appears to be more suitable to describe the differences between ocean-colour retrievals and in-water measurements than “errors”. Interestingly, satellite observations are free from errors associated with handling of samples. However, for consistency with other CCI’s, and in agreement with the recommendations of the CCI drafting team on errors and uncertainties at the first CCI Colocation meeting in ESRIN in September 2010, we use the term error to express the difference between an observation and the true value of the measured quantity. In this context, we take the true value to be to be the in situ observations, limitations of in situ observations notwithstanding.

When comparing satellite-based measurements directly with in situ observations, we also have to be alert to the differences between satellite and in situ observations that may be attributable to the mismatch in the spatial scales of observations: they are typically of the order of hundreds or thousands of square km for satellite data, and of the order of one

square metre or less for in situ observations. It is common practice examine variability in the satellite data around the central pixel matched with in situ data, to assess or minimise differences arising from pixelisation of satellite data and spatial variability in the data (Hu et al. 2001; Bailey and Werdell, 2006). The assumption here is that if inter-pixel variability is low, then sub-pixel variability would also be low. Issues related to mismatch in sampling time also has to be recognised: typically, differences of a few hours of are admitted, implicitly assuming that the field did not evolve significantly during this interval.

One might even argue that, at the scale of a satellite pixel, the satellite observation is the "truth", and that the real challenge is to devise in situ methods that would faithfully replicate the satellite observations at the appropriate time and space scales. But, regardless of the validity of such a position, it would side-step the true object of the document, which is to record the differences that a user might anticipate between in situ and satellite observations. For the reasons noted, we refer to these differences as errors, and use statistical methods of error analysis to quantify the errors.

A considerable amount of work has already been done in the field, notably by NASA, to establish errors associated with the NASA algorithm in use today. This document has made extensive use of this previous work (Campbell and O'Reilley, 2006) and builds on it, and has been developed in close collaboration with colleagues from NASA.

A suite of error metrics are provided, to meet a variety of requirements. The error estimates discussed here may be grouped into two classes: those required by the user community (especially modellers), and additional uncertainty estimates required for evaluation of algorithm performance by the Earth Observation community. Many of the metrics of error used here are among the quantitative criteria used for the selection of algorithms based on inter-comparisons. Error estimates are treated in the first part of the technical note, and uncertainties in the second part.

Because this document is being prepared concurrently with the algorithm selection process, we have opted to illustrate some of the methods proposed here using the NASA algorithm, which is extensively used today. But they are readily applied to any algorithm that might be selected, with the proviso that some of the uncertainty estimates for algorithm performance evaluation, such as goodness of fit, can only be applied to model-based algorithms and not to purely empirical algorithms.

The OC-CCI Team has proposed that the algorithm comparisons be repeated periodically, especially when there is a significant increase in the in situ data that could be used for the comparison, or when there are significant improvements in the algorithms available for retrieval of ocean-colour products. Whenever a significant improvement to

product is achieved, a re-processing would have to be undertaken. This will ensure that the OC-CCI products are systematically updated as more relevant information becomes available. The corresponding updated error and uncertainty estimates would also have to be made available to the community.

2. DEFINITION OF TERMS

The terms used in this document area consistent with the recommendations of the drafting team on errors and uncertainties from the first CCI co-location meeting, as has been captured in the draft ECD of the SST-CCI. In the interest of brevity, these definitions are not repeated here.

3. GLOBAL MEASURES OF ERRORS

The first and simplest set of uncertainties emerges from the validation exercise, in which products retrieved from ocean-colour data were compared with in-water measurements. A suite of metrics are offered here, since no single metric covers all the attributes of uncertainties that have to be characterised.

Root Mean Square Error: This is a commonly-used (IOCCG, 2006; Bailey and Werdell 2006) metric for measuring errors, and it is often recommended for model – observation inter-comparisons in Earth Sciences (Fox 1981; Willmott et al. 1985; Murphy 1988; Taylor xx?). This was one of the metrics on which the selection of algorithms was based, in the OC-CCI algorithm inter-comparisons (Refer to PVP and PVASR). The RMSE is defined as:

$$RMSE = \Delta_{RMS} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{N}} \quad (1)$$

where Δ_{RMS} is the RMSE, x_j are the individual observations and y_j are the corresponding estimates from remote sensing, and N is the total number of matched observations.

Note that by squaring the differences between observations and estimates, the possibility is avoided that negative errors get cancelled against positive errors. The interpretation of RMSE is facilitated if the data (both observations and estimates) have normal distributions. In fact, it has been demonstrated that bio-optical variables such as chlorophyll-a approach a normal distribution when the data have been long-transformed (Campbell 1995). Furthermore, ocean-colour products span a range of some four decades (from 0.01 to 100 mg m^{-3}). The ideal algorithm would work across the whole range equally well; on the other hand, an estimate of error that relies on squared values would be biased towards errors at the high-end of the range. For both these reasons, the error estimates for bio-optical properties are based on log-transformed data (see also Bailey and Werdell 2006; Campbell and O'Reilley 2006; Moore et al. 2009). The RMS error is related to the correlation coefficient, the bias and the variance in the data (Murphy, 1988; Taylor xx?). Of these, the bias is a measure of accuracy, and the variance or its square root, the standard deviation, is a measure of precision.

Bias and Accuracy: The bias is defined as the residual offset that remains when positive and negative errors are cancelled against each other. It is often taken as a measure of accuracy. The bias is computed as:

$$\text{Bias} = m = \frac{\sum_{i=1}^{N}(x_i - y_i)}{N} = \bar{x} - \bar{y} \quad (2)$$

where *m* is the bias, and the mean values of *x* and *y* are indicated by the over-bar.

Standard Deviations and Precision: The bias and the RMSE can be used to derive the standard deviation of the error, which is a measure of the precision of the algorithm. The standard deviation may be expressed as (Fox 1981):

$$s = \sqrt{\frac{N(\Delta_{RMS}^2 - m^2)}{N-1}} \quad , \qquad (3)$$

where the standard deviation, or precision, is denoted *s*.

Mean Absolute Error: The root mean square error has been criticised as being a less than ideal measure of error (Willmott and Matsuura 2005), as it depends not just in the magnitude of the errors, but also on their variance and on their number. On the other hand, the Mean Absolute Error (MAE or Δ_{MA}), defined as:

$$MAE = \Delta_{MA} = \frac{\sum_{i=1}^{N} |x_i - y_i|}{N} \quad , \qquad (4)$$

depends only on the magnitudes of the errors themselves, and may therefore be preferred as a metric for judging the performance of algorithms.

From Validation to Global Error Estimates: The error estimates based on the data used for validation and inter-comparison of algorithms may not be representative of the global error in the product, unless the errors are uniform across the full range of values of the product, or the frequency distribution of the data used in the validation mirrors that in the global data. In fact, it has been shown (Campbell and O'Reilley 2006) that these conditions are not met, and it is confirmed here for NOMAD (NASA bio-Optical Marine Algorithm Dataset) used for validation of the in-water algorithms.

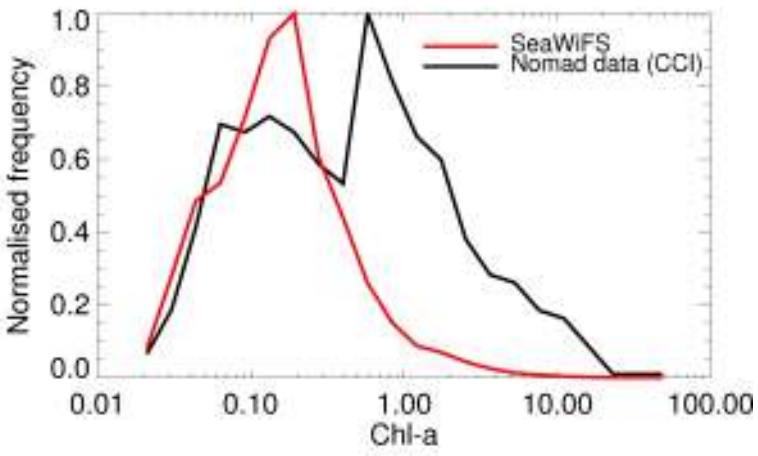


Figure 3-1: Frequency distribution of Chl-a

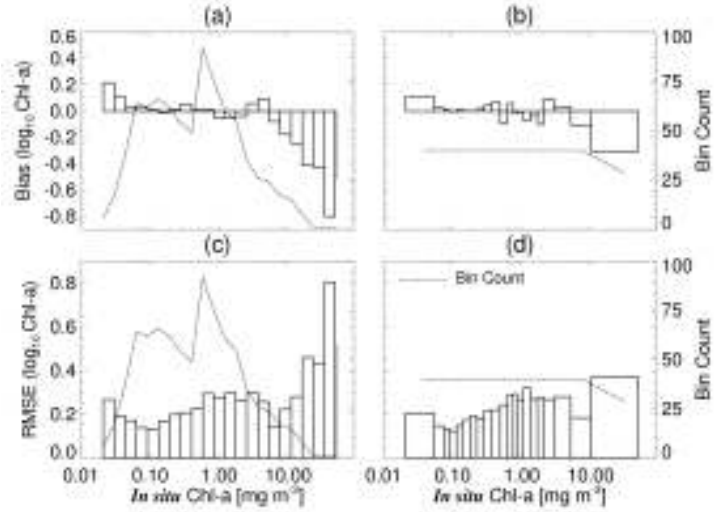


Figure 3-2: RMSE and Bias of Chl-a

Figure 3-1 shows the frequency distribution of chlorophyll-a in the global climatology derived from SeaWiFS, and in the NOMAD dataset used for the validation, both normalised to unity at the maximum frequency. The NOMAD data set is relatively richer in high-chlorophyll values than the satellite-derived frequency distribution. In Figure 3-2, the RMSE and the bias are plotted, after binning them into a number of class intervals. The number of observations in each bin is also shown. In the left-hand panel, the bins are of equal size in log-chlorophyll units, and it emerges that the number of observations in each bin is highly variable, with the bins at the tails of the distribution being poorly represented. The errors vary with the class interval, with the bias changing sign from low-chlorophyll to high-chlorophyll values. To minimise the impact on the statistics arising from the variable number of observations per bin, the analysis was repeated with variable bin size, but holding the number of samples per bin fixed. The results are shown in the right-hand panel of Figure 3-2. It is seen that the variability across bins is reduced when the number of samples is held constant. The size of the bins grows large towards extreme ends of the distribution, drawing attention to areas where more observations are needed. Gregg et al. (2009) also pointed out the need for more samples in extremely-low-chlorophyll environments. Note that this is not a trivial problem: those extreme values remain poorly sampled because they are rarely observed in nature by in situ techniques. Interestingly, the uncertainties are lowest at chlorophyll values corresponding to the highest frequency in the global chlorophyll distribution (Figure 3-1).

The Implementation Plan of GCOS has identified error requirements for ocean-colour products, notably chlorophyll-a. To establish whether this goal is achieved at the global level, weighted error

statistics can be computed, using the binned results, and the frequency of chlorophyll-a in the same bins, estimated from global satellite data:

$$(RMSE)_w = (\Delta_{RMSE})_w = \sqrt{\frac{\sum_{j=1}^{M} f_j (\Delta_{RMS})_j^2}{\sum_{j=1}^{M} f_j}} \quad , \qquad (5)$$

where the subscript *w* indicates weighted values, M is the number of bins, f_j is the frequency in the j^{th} bin and and $(\Delta_{RMS})_j$ is the corresponding RMSE (see also Gregg et al. 2009). The weighted RMSE and bias established in this manner are shown in Table 3-1 for the NASA chlorophyll algorithm (OC4.v6) along with the un-weighted error metrics from the round-robin inter-comparison. The accuracy and precision (as represented by the bias and standard deviation respectively) are within the GCOS requirements, and meet many of the user requirements identified in the OC-CCI User Requirements Document. Because the errors are lower for those intervals that are most represented in the global ocean, the weighted RMSE is lower than the un-weighted case. The bias changes sign between the two cases, but both values remain close to zero.

Table 3-1: Globally representative error statistics for the OC4v6 algorithm using a CCI subset of NOMAD data.

	Unweighted	Weighted
RMSE	0.237	0.191
Bias	-0.010	0.017

Unweighted statistics assume the distribution of data in NOMAD is representative of the global ocean. Weighted statistics are calculated by averaging the RMSE and bias over all bins and weighting the bin according to the frequency in each log(Chl) bin as determined by the 1997-2010 SeaWiFS climatology. Results are based on variable bin sizes that maintain the same number of samples per bin, except in the last bin.

Towards Pixel-by-Pixel Error Estimates

It was evident from the user consultation (see OC-CCI User Requirements Document) that there is a need for error estimates on a pixel-by-pixel basis. The global error estimates presented above do not meet this requirement for various applications, especially for assimilation of ocean-colour data into models, and for validation of model results. We demonstrate below two methods for achieving this goal. The first one follows directly from the binned statistics presented above, the other is based on optical classification of the ocean waters.

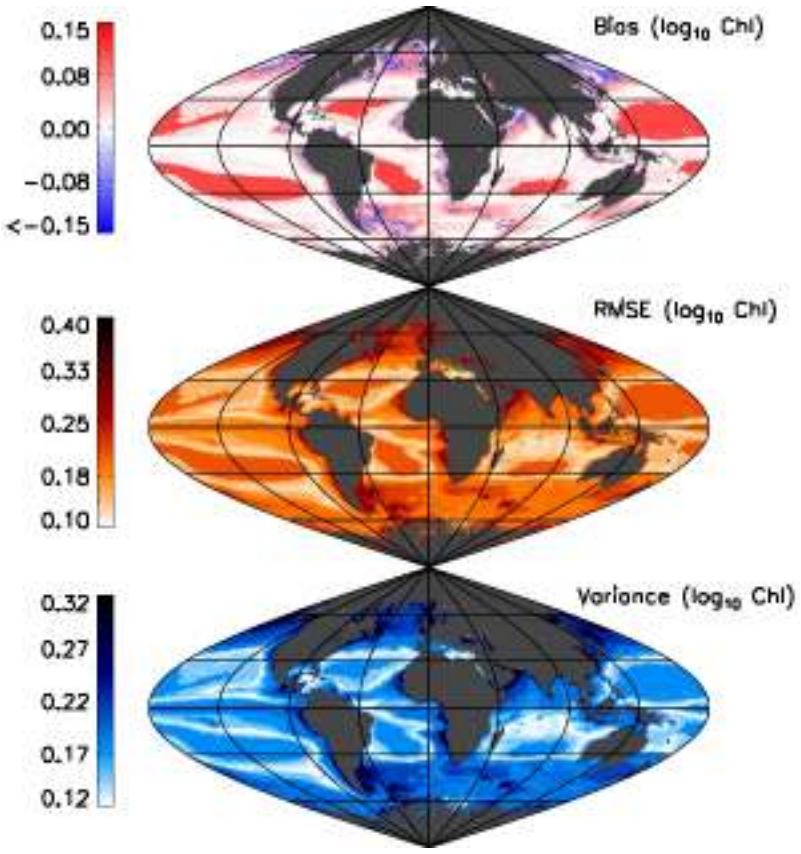


Figure 3-3: Pixel-by-pixel error statistics at the global scale, based on the multi-year SeaWiFS average.

Error Estimates from Binned Statistics: To extend the error estimates to meet this requirement, we propose to follow two different approaches: The first one is to compute the error statistics in narrow bins along the range of the variable (see Fig. 3-1), such that errors can be specified for each of the bins (e.g. Campbell and O'Reilley, 2006). These binned statistics also allow us to estimate global errors by weighing each class interval of the bins by the frequency of occurrence of the class interval in global images (Campbell and O'Reilley, 2006), as done above.

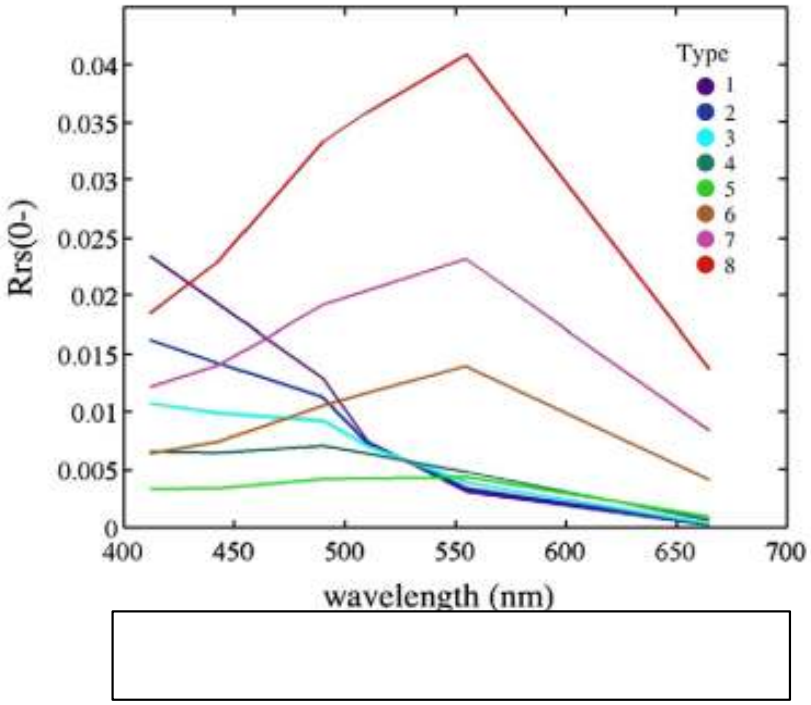
In this approach the binned statistics are used to assign error estimates according to the pixel value of the ocean-colour product: the assigned errors are those of the corresponding bins. The approach is illustrated for chlorophyll-a mapped using the NASA algorithm (Figure 3-3) applied to SeaWiFS climatology.

These error assignments are made on the basis of a single product, for example, the chlorophyll-a concentration, as in the example shown in Figure 3-3. A more generalised approach would take into account the effect on errors, of change in the optical water type. Such an approach would include the collective effect of variations in all optically-significant substances. This is the approach followed in the method discussed in the next session.

Error Estimates on the Basis of Optical Classification of Oceanic Waters:

In the second approach, we propose to use error estimates based on optical classification of the waters, following the method proposed by Moore et al. (2009).

Figure 3-4: From Moore et al. (2009). Average remote-sensing reflectance spectra of the eight water classes.



Here, a large database of water-leaving radiances is used to classify the waters into 8 water classes. A fuzzy c-means clustering algorithm was used by Moore et al. (2009) to identify eight optical water types. Mean remote-sensing reflectance spectra for the eight water types are shown

in Figure 3-4. In the clustering method used, the optimal number of clusters is selected on the basis of validity functions. The shapes of the spectra indicate that classes 6 – 8 are clearly

Table 3-2: Error statistics for SeaWiFS for each of the water types in chlorophyll retrieval, determined using a validation dataset of 1576 data points. Clasification of water types is based on Fuzzy c-means algorithm (results from Moore et al. 2009).

Optical water type	Average relative error (%)	RMS log error	Bias log error	# points	Average Chl	Min Chl	Max Chl
1	35	0.302	0.087	35	0.075	0.030	0.89
2	53	0.260	−0.059	173	0.257	0.024	0.77
3	35	0.216	0.029	103	0.361	0.078	2.60
4	73	0.283	−0.083	235	2.45	0.05	8.38
5	77	0.273	−0.063	371	2.27	0.147	19.85
6	93	0.319	−0.152	540	7.05	0.14	129.3
7	95	0.314	−0.174	114	1.05	0.056	22.70
8	110	0.589	0.138	5	9.97	0.65	3.40
All	77	0.292	−0.094	1576	3.52	0.147	23.48

influenced by scattering substances other than phytoplankton. An additional water type has been recently identified by these authors as being waters dominated by coccolithophores in bloom conditions (Mark Dowell, personal communication). Adding this additional class eliminates open-ocean pixels that emerged as unclassified according to the original work.

After assigning each of the validation spectra to one of the optical water types, error statistics were computed by Moore et al. (2009) for chlorophyll retrieval, as shown in Table 3-2. Each pixel in a satellite image can then be classified according to its membership (ranging from 0 to 1) to each of the water types, and the errors to the pixels are assigned according to membership.

These error statistics are what the modellers in our team are interested in (based on feedback from Stephanie D, Stefano Ciavatta, Takafumi Hirata, Corinne Lequérré and Laurent Bertino). However, ocean-colour algorithm developers and evaluators need additional criteria for evaluating model performance. These measures of uncertainties and errors are discussed in Section 4 of the Report

4. SOURCES OF ERRORS

As with any other observation technique, remote sensing data have errors. In case of ocean colour remote sensing the errors start with the primary data, i.e. the top of atmosphere spectral radiances. The error increases in the form of a cascade with subsequent levels of derived data, starting with the water reflectance after correcting the influence of the atmosphere, the inherent optical properties (IOPs), the geophysical variables, which are derived from IOPs and more complex estimates such as primary production. If one considers the manifold factors, which determine the TOA radiances and the few properties, which can be derived from the radiance, then it is clear that we have to expect a variable and sometimes large range of errors in the products (Fig. 4.1). The errors can be determined on a statistical basis by comparing a large number of remotely-sensed data with in situ observations.

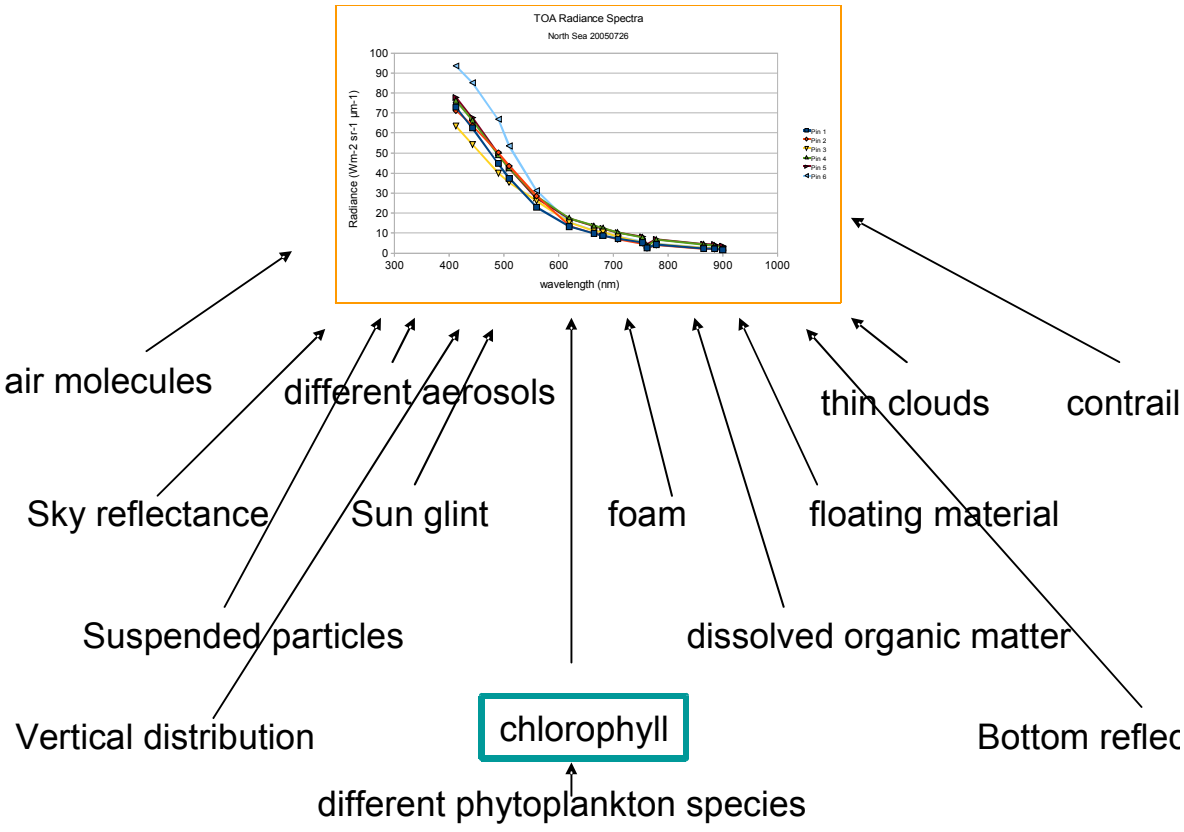


Figure 4-1: ***Factors, which determine top of atmosphere radiance spectra***

Each step in the cascading algorithms and error propagations has to be analysed to determine mean error levels. Errors can be determined and clustered by this method for various water types and atmospheric conditions. Such analyses help pin-point the biggest sources of overall errors, such that attempts at error reduction can focus on those aspects which might lead to maximum benefit.

5. Uncertainties in Algorithms

Whereas the general statistical approach and the error analysis are described in Section 3 of the document, this section covers the pixel-by-pixel method for estimating uncertainties arising from the algorithm itself. Every algorithm has an inherent capability to resolve the retrieved parameters to within a certain precision, and high-performance algorithms would be able to achieve parameter retrievel with high precision. Hence, for algorithm evaluation, it is useful to understand, and compare uncertainties in parameter retrievals.

An overall measure of performance of algorithms based on modelled radiance or reflectance spectra is the chi-square or analagous quantity, which gives a measure of the goodness of fit when the modelled spectrum is compared with the observed spectrum. This is a very useful quantity, but it does not give an indication of the uncertainties in retrieval of individual parameters. Note that the chi-square estimates are not possible for regression algorithms that do not have an underlying model. Nor is the chi-square meaningful when a simplified model is used to solve for the unknown parameters algebraically.

Furthermore, we note that the algorithms should be flagged as being out-of-scope, if a model-based method is unable to reproduce an observed spectrum within a pre-determined value of the goodness of fitt. Therefore, determination of uncertainty envisaged here comprises two parts: (1) detection of cases (radiance or reflectance spectra), which are out of scope of the forward model and (2) the uncertainty in the variables to be derived for the cases, which are sufficiently within the scope of the algorithm. The core and most critical part in designing an algorithm is the definition of the optical models of the atmosphere for the atmospheric correction part and of the water and its constituents for the in water algorithm part. These models describe the spectral optical properties including the number and nature of components, which determine the properties, and their ranges and co-variances. Since the amount of information, which is hidden in the TOA radiance spectra - our primary and only data - is limited, the number of components in the optical models should be on one hand as small as possible, but should also be capable of describing the variability of the TOA or water leaving radiances with sufficient accuracy. Thus, for these models components have to be defined, which function as proxies for a larger number of constituents in atmosphere and water with similar spectral properties (Fig. 5.1).

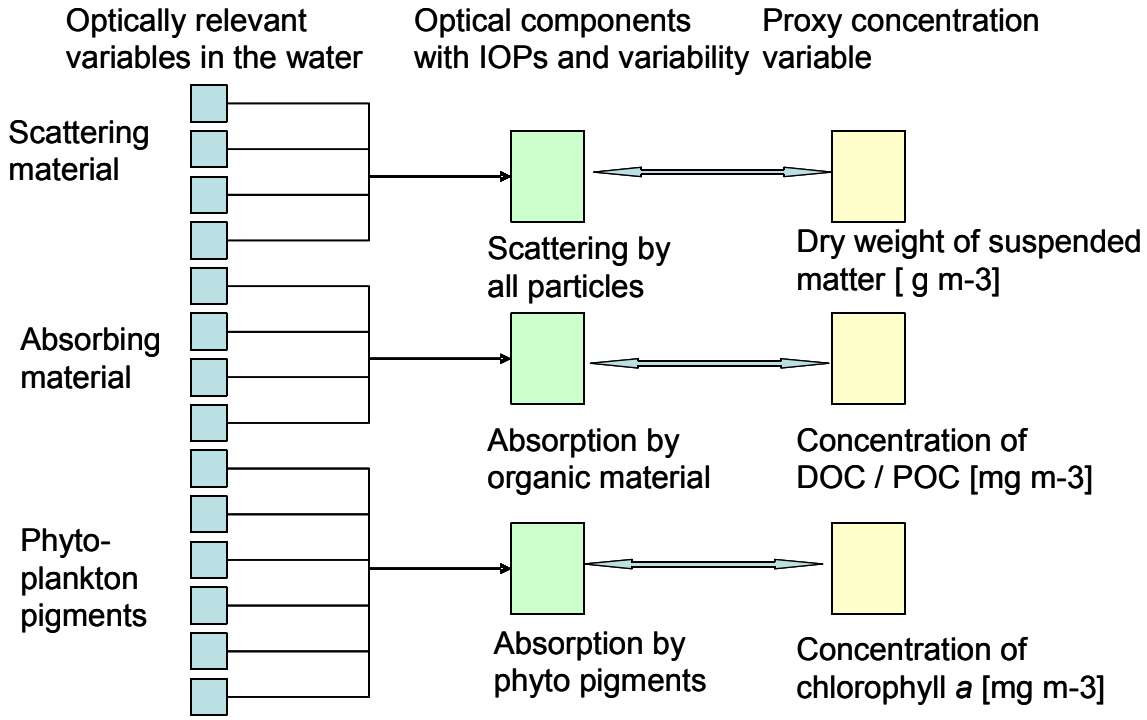


Figure 5-1: Uncertainties due to the bio-optical model

This section describes a procedure to detect top of atmosphere (or water leaving) radiance reflectance spectra that are out of scope of the algorithm, and a procedure to determine the uncertainty ranges of the variables to be retrieved, i.e. water-leaving radiance reflectances, inherent optical properties, attenuation coefficient for downwelling irradianc and the chlorophyll concentration.

Based on the optical models the relationships between inherent optical properties and/or concentrations on one hand and the water leaving or TOA radiances on the other hand have to be modelled in the form of empirical regressions or radiative transfer models. These relationships are then the basis for any algorithm. Two types of algorithms are possible: (1) inverse algorithm, which determine the independent variable from the dependent variable. This can be a simple band-ratio algorithm, which determines the chlorophyll concentration from a ratio of water reflectances at two wavelengths, or a more complex multiple non-linear regression technique such as artificial neural networks. (2) The other type of algorithm uses a forward model, which determines the dependent variable from the independent, and which is used iteratively to get the best fit between the measured and modelled dependent variables, which are the radiance or reflectance spectra. Here, we describe the second approach, which allows the determination of the uncertainty range for each variable derived. The procedure to detect out of scope spectra is an auto-associative neural network. The uncertainty ranges are computed from the Hessian

Matrix, which is part of an optimization procedure with constraints and forward neural networks, which have been trained using simulated reflectance spectra.

Overview:

The detection of uncertainties consists of two procedures: (1) an auto-associative neural network to detect spectra, which are out of scope of the algorithms, and (2) the determination of the range of each independent variable (concentrations, IOPs), within which the variable can be varied, without changing the reflectance spectra outside its uncertainty range. This range is then the uncertainty of the variable to be retrieved.

Detection of out of scope spectra

The detection of out of scope spectra is performed with auto-associative neural networks (autoNN), which are based on the same training data set used for the retrieval algorithms (Fig. 5.2). An autoNN is designed and trained to reproduce the input variables as outputs, i.e. input and output variables are identical in the training run. In our case these variables are the reflectances at different wavelengths. Furthermore, one of the hidden layers has a number of neurons which is as small as possible but as large as necessary to reproduce the input variables with sufficient accuracy. This number represents the information content of the training data set i.e. the number of components, which are necessary to describe the variability of the reflectance spectra of the training data set.

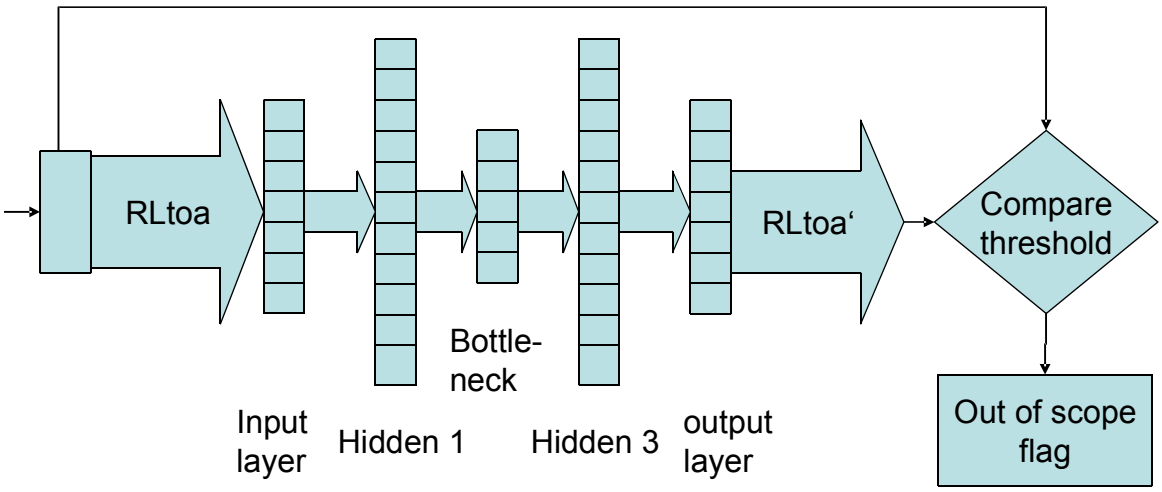


Figure 5-2: ***auto-associative neural network to detect out of scope RLtoa spectra***

The autoNN is created and applied only to TOSA reflectance spectra. TOSA stands for top of standard atmosphere, which is defined by a surface pressure of 1013.2 hPa and an ozone content of zero or 350 DU (model dependent). It is trained for a variety of aerosols and fresnel reflectances including sun glint and for the variability of in water constituents, which has been used to train the in water neural network. The output radiance reflectance spectrum (RL_tosa_nn) of this nn is compared to the input reflectance spectrum (RL_tosa_ms), which is the measured spectrum. The deviation between both are calculated as the root mean square relative to the measured spectrum.

formula rel_rmsdev=sum(u) ((rl_tosa_nn(i)-rl_tosa_ms(i))/rl_tosa_ms(i))**2

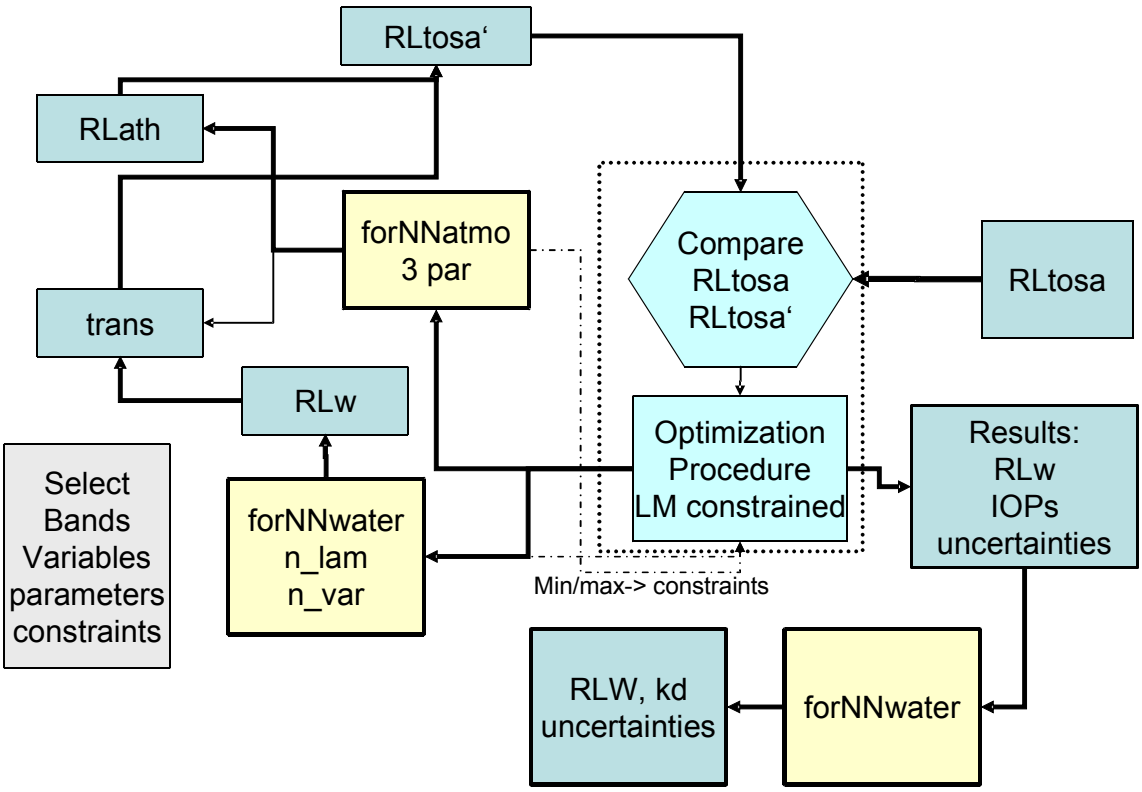
The deviation should not exceed 0.03, which is the uncertainty in the calibration of MERIS, other wise the out_of_scope flag is raised.

6. DETERMINATION OF UNCERTAINTIES

All valid spectra, which are in scope of the algorithm, will be used for retrieving the water leaving radiance reflectances, the attenuation coefficients of the downelling irradiances, the IOPs and the concentration of chlorophyll. For each of these variables the uncertainty range will be computed.

General approach to determine uncertainty ranges

Uncertainties can be computed when using a forward model to compute water leaving or top of atmosphere radiance from IOPs within an optimization loop, in which the sum of squared deviations between the measured and the modelled spectrum per band is minimized. In the case of OC-CCI two forward neural nets are used, one for computing the water leaving radiances from IOPs as input, the other for computing the top of atmosphere path radiance and transmittances. The uncertainties are derived from the Hessian Matrix, which is the second partial derivatives of the independent variables (i.e. IOPs or concentrations) with respect to the target variable, which is the root mean square deviation between measured spectrum and the spectrum produced with the forward model. In our case the forward model is the neural network, which is trained with simulated reflectances. The Hessian matrix can be computed independently or as part of the optimization procedure, which is used to get the best fit between the measured and simulated spectrum by varying the IOPs or concentrations. In the case of the Levenberg-Marquard optimization procedure, which is used in this study, a quasi Hessian Matrix is computed from the Jacobian Matrix (first order derivatives), from which, in turn, the co-variance matrix is derived. Finally, the standard deviations can be extracted per variable from its squared diagonal. Depending on the type of variable, for which the uncertainty shall be computed, two different approaches have been implemented. (1) For the independent variables, such a IOPs or concentrations, the uncertainty can be directly taken from the co-variance matrix of the inversion procedure. (2) For apparent optical properties, i.e. water leaving reflectance or kd the uncertainty ranges have to be computed using the forward model again after the uncertainties of the IOPs / concentrations are known.



Uncertainties of atmospheric parameters and water IOPs and concentrations

The uncertainties are taken from the co-variance matrix. The squared diagonal delivers the standard deviations for each variable. In case of the atmosphere the uncertainties are computed for three variables, which are used as input to the forward neural network, i.e. aerosol optical thickness at 550 nm, the angstrom coefficient and the wind speed together with the sun and observation angles. Output of the NN are the path radiances and up- and downwelling transmittances for all bands used in OC-CCI (MERIS, MODIS, SeaWIFS, CZCS). For the in water part the forward neural network computes water leaving radiance reflectances from the 4 IOPs, i.e. absorption coefficients of gelbstoff, particulate matter and phytoplankton pigments, and scattering coefficient of suspended matter. Further control parameters are the sun and observations angles. Thus, the uncertainties are calculated either for all 4 IOPs or only for those

of importance for a certain region. In the latter case only the selected IOPs are kept as variables in the optimization loop, while the others are set to parameters with a fixed value. Thus, altogether 7 variables are varied in the optimization loop.

Uncertainties of water leaving radiance reflectances

Although the water leaving radiance reflectances are derived from TOA radiance reflectances, they are not independent variables, but depend on the water constituents as the toa radiances. To overcome this problem, the uncertainty ranges of all IOPs, which determine the Rlws, are first computed and then used in the forward model to compute the range of Rlw of all spectral bands with the forward model. Three spectra are computed: (1) the reflectance, (2) the lower reflectance and (3) an upper reflectance. For the lower reflectance the standard deviation of all absorption coefficients is added to the mean and of all scattering coefficients are subtracted from the mean and vice versa for the upper bound. This scheme is illustrated in Fig. 6.1).

Uncertainties of downwelling irradiance coefficients

The computation of the downelling irradiance attenuation coefficients (kd) is similar to that of the water leaving radiance, since kd is not an independent variable in the model. Thus the uncertainty range will be computed after the uncertainties of all IOPs are known. The upper and lower standard deviations of the IOPs are used to compute the upper and lower values of kd and thus its uncertainty range with the forward NN.

Constraints

The neural networks, which are used here as a proxy of the forward radiative transfer model, are trained with real forward model for a limited range of input variables. Thus, an optimization procedure must be used, where constraints can be set. The upper and lower limits for each input variable are taken from the header of the neural network. The constraints can be further restricted to a smaller range if the range of IOPs are known for the area under request.

Example

The following example with simulated toa reflectances and errors illustrates the various steps and intermediate products, which can be used to determine the uncertainty range in water leaving radiance reflectance spectrum and water IOPs. It demonstrates also the sensitivity of the system to only very small errors in TOA reflectances. Forward models for this test are 2 neural networks,

which have been trained using radiative transfer simulations. The model of the atmosphere used here has been prepared by R. Santer. It is based on AERONET data and an SOS code, which includes polarisation. The water model is Hydrolight 5.1 (Mobley). For training of the NN about 300 000 cases have been simulated. The achieved errors (comparing model with NN results) are in the order of 10-6. For the following demonstration case we have selected the following parameters:

Parameter	Value
sun zenith angle	45 deg
viewing nadir angle	32 deg
azimuth difference	50 deg
water temperature	15 deg
Salinity	35 psu
Aerosol optical thickness at 550 nm	0.2
Angstrom coefficient	1.3
Wind Speed	3.7 m s-1
absorption coefficient of particulate matter at 443 nm	0.1 m-1
absorption coefficient of yellow substance at 443 nm	0.6 m-1
absorption coefficient of phytoplankton pigments at 443 nm	0.2 m-1
scattering coefficient of particulate matter at 443 nm	2.1 m-1

Note: no atmospheric gaseous absorption was included in this study.

The corresponding water leaving reflectance and top of atmosphere reflectances are presented in Fig. 6.2.

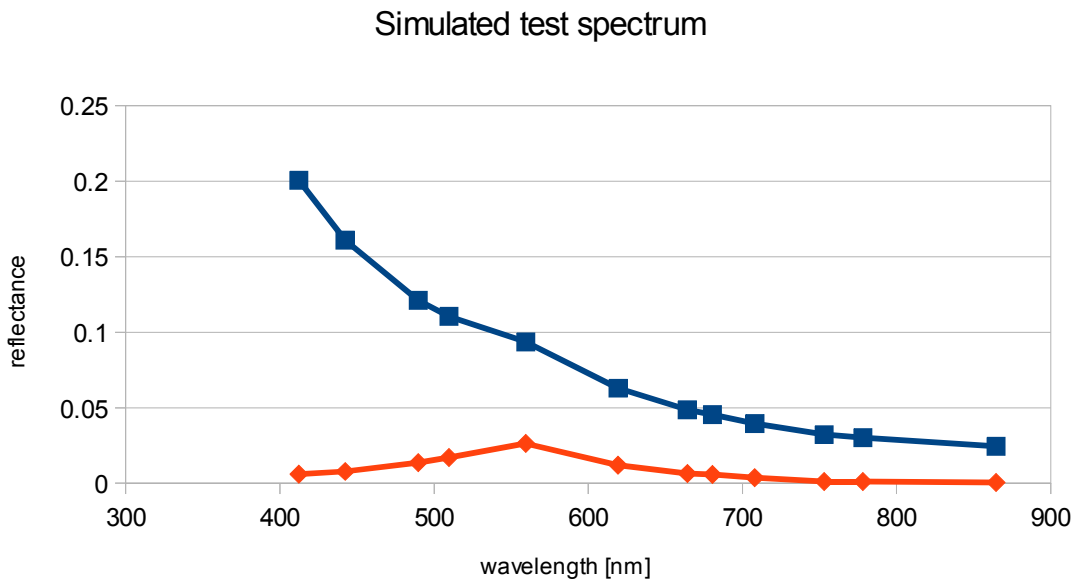


Figure 6-2: Simulated test case: water leaving (red) and top of atmosphere reflectance (blue)

First test assuming no errors in the toa reflectance spectrum

In the first test the simulated (“true”) toa reflectance is used without any noise or assumed systematic errors. Using the neural networks as forward models and a constrained Leveberg-Marquard optimization procedure we fit the toa reflectance spectrum to the simulated (“true”) one. The result is shown in fig. 6.3. The deviations between the true and the retrieved reflectances are < <0.1 %.

wavelength	simulated Rtoa	retrieved Rtoa	Deviation %
412.3	0.2006	0.2006	-0.001
442.3	0.161	0.161	0
489.7	0.1212	0.1212	0
509.6	0.1105	0.1105	-0.001
559.5	0.09357	0.09357	-0.001
619.4	0.06293	0.06293	0.003
664.3	0.04872	0.04872	0.006
680.6	0.04539	0.04539	0.003
708.1	0.03953	0.03953	0.006
753.1	0.03224	0.03224	0.012
778.2	0.03021	0.03022	0.011
864.6	0.02453	0.02452	-0.016

Figure 6-3: Comparison between the reflectances of the simulated ("true") and the retrieved spectrum, deviations are in %

The deviation between the corresponding simulated (“true”) and retrieved water reflectances is shown in Fig. 6.4. Also here the deviation is < < 0.1%.

wavelength	Rw retrieved	Rw true	Deviation	Deviation %
412.3	0.006007	0.006009	-0.000002799	-0.04658
442.3	0.007831	0.007831	-1.214E-017	-1.551E-013
489.7	0.01356	0.01356	0	0
509.6	0.01708	0.01708	6.939E-018	4.062E-014
559.5	0.0263	0.0263	-6.939E-018	-2.638E-014
619.4	0.01195	0.01195	-0.000004121	-0.03447
664.3	0.006522	0.006524	-0.000002225	-0.0341
680.6	0.005772	0.005776	-0.000003936	-0.06814
708.1	0.003702	0.003705	-0.000002599	-0.07014
753.1	0.001123	0.001123	-4.072E-007	-0.03625
778.2	0.001192	0.001192	-8.558E-007	-0.07177
864.6	0.0004886	0.0004889	-3.532E-007	-0.07224

Figure 6-4: Water reflectances, retrieved and "true" with deviation and deviation in %

	AOT	Angstrom	wind	a_part	a_y	a_pig
“true”	0.2	1.3	3.7	0.1	0.6	0.2
retrieved	0.2001	1.305	3.739	0.1111	0.5988	0.2005

Figure 6-5: "true" and retrieved parameters: aerosol optical thickness, angstrom coeffcient, wind speed and the 4 IOPs

Also the retrieved parameters (s. Fig. 6.5) are very close to those that have been used for the simulations.

Now we determine the uncertainties for this case. Although the simulated TOA spectrum, which is used as a “measured spectrum”, is error free and the retrieved TOA spectrum is nearly error free (<<0.1 %), the independent variables, which we have derived iteratively by fitting the toa spectrum, have an uncertainty range. This range is retrieved from the covariance matrix of the 7 variables, which are modified in the iteration.

varaible	AOT	Angstrom	wind	a_part	a_y	a_pig	b_part
AOT	0.000000004	3.135E-007	0.000002772	-0.00003335	3.302E-007	-1.077E-006	-1.916E-007
Angstrom	3.135E-007	0.00002703	0.0002337	-0.002843	0.00002835	-0.00009095	-0.00001581
wind	0.000002772	0.0002337	0.002116	-0.02786	0.0002752	-0.0008898	-0.0001552
a_part	-0.00003335	-0.002843	-0.02786	0.5875	-0.005972	0.01924	0.003241
a_y	3.302E-007	0.00002835	0.0002752	-0.005972	0.00006136	-0.0001972	-0.00003299
a_pig	-0.000001077	-0.00009095	-0.0008898	0.01924	-0.0001972	0.00064	0.0001077
b_part	-1.916E-007	-0.00001581	-0.0001552	0.003241	-0.00003299	0.0001077	0.0000183

Figure 6-6: Co-variance matrix of the 7 variables after successful fit of R_toa

The square roots of the co-variance matrix are the standard deviations of the 7 variables. Note that AOT, angstrom and wind are given on the linear scale, the 4 IOPs on the log scale.

aot	ang	wind	apart	agelb	apig	btot
6.324e-05	5.199e-03	4.600e-02	7.665e-01	7.833e-03	2.530e-02	4.278e-03

Figure 6-7: Standard deviation of the 7 variables

When the 4 IOPs are transformed to the linear scale, the following ranges are the result:

variable	- 1 stdev	mean	+1 stedev	TRUE
AOT	2.000E-001	2.001E-001	2.002E-001	0.2
angstrom	1.300E+000	1.305E+000	1.310E+000	1.3
wind	3.693E+000	3.739E+000	3.785E+000	3.7
a_part	5.161E-002	1.111E-001	2.391E-001	0.1
a_y	5.941E-001	5.988E-001	6.035E-001	0.6
a_pig	1.955E-001	2.005E+000	2.057E-001	0.2
b_part	2.090E+000	2.099E+000	2.108E+000	2.1

Figure 6-8: Uncertainty range of the retrieved variables

The variables have different uncertainty ranges. Obviously a_part has the largest range of +-50%, while all others have only a small range. It means that the effect a_part on R_toa is small.

Next step is to determine the uncertainty range of the water reflectances, which are not independent variables and not the variables, which have to be fitted.

Thus, the uncertainty ranges have to be determined from the uncertainty ranges of the IOPs. To determine the lower limit for 1 standard deviation, the standard deviations of those IOPs have been added, which have an decreasing effect on reflectances, i.e. the absorption coefficients and vice versa. For the upper boundary the absorption coefficients are subtracted and the others are added. Then the water reflectance are determined again for the lower and upper boundaries using the water forward neural network.

The results are given in Fig. 6.9. Although the fit of the toa reflectance is nearly perfect, we get uncertainties for the water reflectances up to 2.5%.

wavelength	Rw	Rw_min	Rw_max	difference	uncertainty %
412.3	0.006007	0.005844	0.006154	0.0001551	2.582
442.3	0.007831	0.007638	0.008012	0.0001869	2.386
489.7	0.01356	0.01324	0.01387	0.0003135	2.312
509.6	0.01708	0.01668	0.01745	0.0003846	2.251
559.5	0.0263	0.02578	0.02679	0.0005061	1.924
619.4	0.01195	0.01181	0.01209	0.00014	1.172
664.3	0.006522	0.006465	0.006589	0.00006233	0.9556
680.6	0.005772	0.005725	0.005831	0.00005317	0.9212
708.1	0.003702	0.003672	0.003733	0.00003053	0.8246
753.1	0.001123	0.001114	0.001132	0.000008751	0.7793
778.2	0.001192	0.001183	0.001201	0.000009414	0.7901
864.6	0.0004886	0.0004845	0.0004927	0.000004061	0.8311

Figure 6-9: Uncertainty range for the water leaving reflectances

7. GUIDELINES FOR USING THE PRODUCTS

This section is to be developed as the algorithm selection progresses.

8. CONCLUSION

This is a preliminary report outlining the current thinking in OC-CCI on error characterisation, it will be updated as the work on the products advances.

9. BIBLIOGRAPHY

Bailey, SW, Werdell, PJ, 2006. A multi-sensor approach for the on-orbit validation of ocean color satellite data products. Remote Sensing of Environment. 102: 12–23. doi:10.1016/j.rse.2006.01.015

Campbell, JW (1995) The log-normal distribution as a model for bio-optical variability in the sea. J. Geophys. Res. 100 (C7): 13237–13254.

Janet W. Campbell & John E. O'Reilly. 2006. Metrics for Quantifying the Uncertainty in a Chlorophyll Algorithm: Explicit equations and examples using the OC4.v4 algorithm and NOMAD data.

Fox, DG (1981) Judging air quality model performance. Bull. Amer. Meteorol. Soc. 62 (5): 599–609.

Gregg, WW, Casey, N, O'Reilly, JE, Esaias, WE (2007) An empirical approach to ocean color data: reducing bias and the need for post-launch radiometric re-calibration. Remote Sensing of Environment, 113: 1598–1612. doi:10.1016/j.rse.2009.03.005

Hu, C, Carder, KL, Muller-Karger, FE (2001) How precise are SeaWiFS ocean color estimates? Implications of digitization-noise errors. Remote Sensing of Environment, 76, 239−249.

IOCCG (2006). *Remote Sensing of Inherent Optical Properties: Fundamentals, tests of algorithms, and applications. Lee, Z.-P. (ed.),* Reports of the International Ocean-Colour Coordinating Group, No. 5, IOCCG, Dartmouth, Canada.

Moore, TS, Campbell, JW, Dowell, MD (2009) A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. Remote Sensing of Environment. 113: 2424–22430.

Murphy, AH (1988) Skill scores based on the mean square error and their relationships to the correlation coefficient. Monthly Weather Review. 116: 2417–2424.

Werdell, PJ, Bailey, SW (2005) An improved bio-optical data set for ocean color algorithm development and satellite data product validation. Remote Sensing of Environ., 98: 122−140.
Willmott, CJ, Ackleson, SG, Robert, ED, Feddema, JJ, Klink, KM, Legates, DR, O'Donnell, J, Rowe, CM (1985) Statistics for the evaluation and comparison of models. J. Geophys. Res. 90 (C5): 8995–9005.

End of Document