DUE Coastcolour

Consensus Case 2 Regional Algorithm Protocols

Deliverable DEL-26

Issue 1.0

Author: R. Doerffer, C. Brockmann

Distribution: S. Pinnock, ESA Esrin
CoastColour Team

Revisions:

Issue	Date	Subject	Author
0.9	20130630	Draft	Doerffer
1.0	20140331	Update Section numbering	Brockmann

The Coastcolour Team

BROCKMANN CONSULT

With

MUMM

GKSS FORSCHUNGSZENTRUM in der HELMHOLTZ GEMEINSCHAFT

ADRINORD

PML Plymouth Marine Laboratory

Centro de Oceanografia University of Lisbon

And the Consultant Team

- Dr. Yu-Hwan Ahn (KORDI)
- Dr. Jim Gower (DFO)
- Dr. Mark Dowell (JRC)
- Dr. Stewart Bernard (CSIR)
- Dr. Zhongping Lee (U. Mississippi)
- Dr. Bryan Franz (NASA)
- Dr. Thomas Schroeder and Dr. Arnold Dekker (CSIRO)

Scope of this document

This document summarizes the experience which we gained from the Coastcolour project regarding the development of algorithms for retrieving water optical properties and concentrations of water constituents from top of atmosphere reflectance spectra over different types of coastal waters, in particular from the full resolution data of the Medium Resolution Imaging Spectrometer MERIS.

Based on this experience the document will provide best practice protocols for defining regional algorithms for a specific area and will provide:

- An overview of the individual steps required to defining a regional ocean colour algorithm and documenting existing software and tools that can be used to undertake such a task.
- Minimum requirements for in-situ data for defining a regional algorithm.
- Methods for the definition and parameterisation of a reflectance model (forward model) for a specific regional of interest, including techniques for atmospheric correction and modelling marine optics.
- Approaches to solving the inverse problem, including a comparison of different methods.

The document includes a series of synthesized protocols, which document agreed approaches for defining regional empirical and semi-analytical Case 2 algorithms, including:

- description of existing regional and class-based approaches and their related uncertainties;
- methods of atmospheric correction;
- water constituent, IOP and AOP product types;
- approaches for deriving new regional algorithms;
- characterisation of regional bio-optical water types;
- EO and in-situ data requirements and available tools;
- relevant bibliography.

Furthermore, a simple users' guide is included to allow non-specialist ocean colour users to quickly assess regional characteristics and likely uncertainty value classes.

The Consensus Case 2 Regional Protocols will also be published on the Coastcolour web portal.

Table of Content

1 Introduction

Remote sensing of optically complex water is a difficult task. We have to solve an under determined system, when we try to derive a few variables from top of atmosphere (TOA) spectra of radiances or reflectances, which are determined by a large number of independent variables.

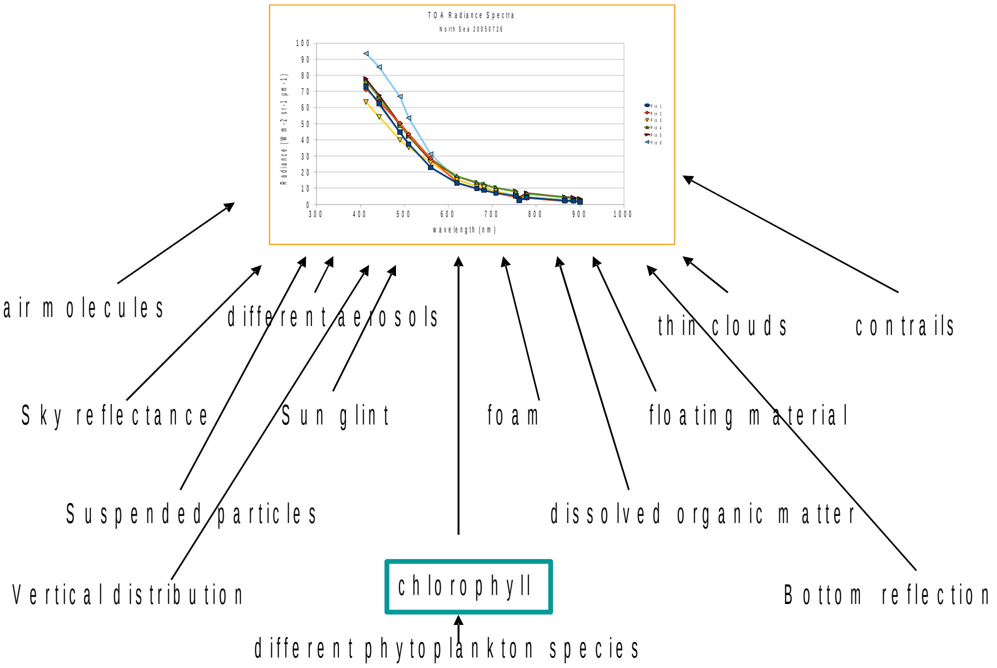

Fig. 1: Factors which determine the top of atmosphere (TOA) radiance spectra over coastal waters, also when only the concentration of chlorophyll shall be determined. On the other hand the information content of the TOA radiance spectra is much lower than the factors which determine it.

Fig. 1 illustrates this problem. Even if we want to determine only the concentration of phytoplankton chlorophyll our signal is influenced by a manifold of factors. On the other hand, the information content of the TOA reflectance spectra, i.e. the number of parameters, which can be retrieved from such spectra, is limited.

This issue sets the frame and the limitation for our attempt to design procedures and algorithms for retrieving optical properties and the concentration of constituents of coastal waters with its variety of different phytoplankton species, suspended particles and organic compounds, all of which scatter or absorb the sun light, which we use as information carrier.

On the other hand a large number of publications have shown examples of successful applications of optical remote sensing even in coastal zones with high concentrations of water constituents. An overview is given in the IOCCG report No. 3, Sathyendranath (2000).

Thus, in this report we have to analyse, under which conditions remote sensing of coastal waters is possible with sufficient accuracy, which procedures and algorithms can be used and how uncertainties and errors can be detected, quantified and flagged in the products. This has to include the correction of the influence of the atmosphere to determine water reflectance spectra, as well as the retrieval of optical parameters and concentrations from the water reflectance spectra. We have to expect that there is no unique procedure, which could fulfil our requirements for all types of coastal waters and all conditions.

2 The problem of Case 2 water remote sensing

The design of an ocean colour remote sensing algorithm is rather simple if only one component is present in the water, which changes the water leaving radiance. Then the concentration of this component can be determined in principle from the radiance or reflectances at one spectral band on which this component has a strong effect by absorption or scattering. An alternative is to use 2 spectral bands, of which one band is strongly affected by absorption and scattering and the other less. Water types which optical variations can be described by only one component has been defined by Morel & Prieur (1977) as case 1 water.

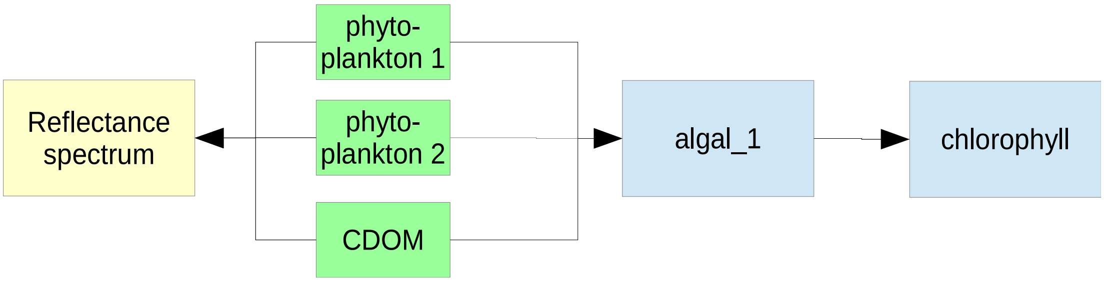

Fig. 2: case 1 water bio-optical model: relationship between various water constituents, its proxy algal_1 (MERIS) and the chlorophyll concentration, and the reflectance spectrum

This type of water is present mainly in the open ocean, where the optical properties are determined by the absorption by phytoplankton pigments and where all other water constituents have a minor optical effect and are correlated with phytoplankton. Fig. 2 shows the principle of the relationship between phytoplankton chlorophyll and the reflectance ratio.

The inversion, i.e. the determination of this component from the reflectance spectrum, can then be based on a regression between two variables: on one hand the reflectance at one band or the ratio of reflectances at 2 different bands on the other hand the concentration . The regression can be based on a linear relationship between the logarithm of the two variables or by any other function such as a polynomial. The coefficients of this function can be determined by measured chlorophyll and reflectances or by simulations using a radiative transfer model. Fig. 3 shows this relationship for measurements.

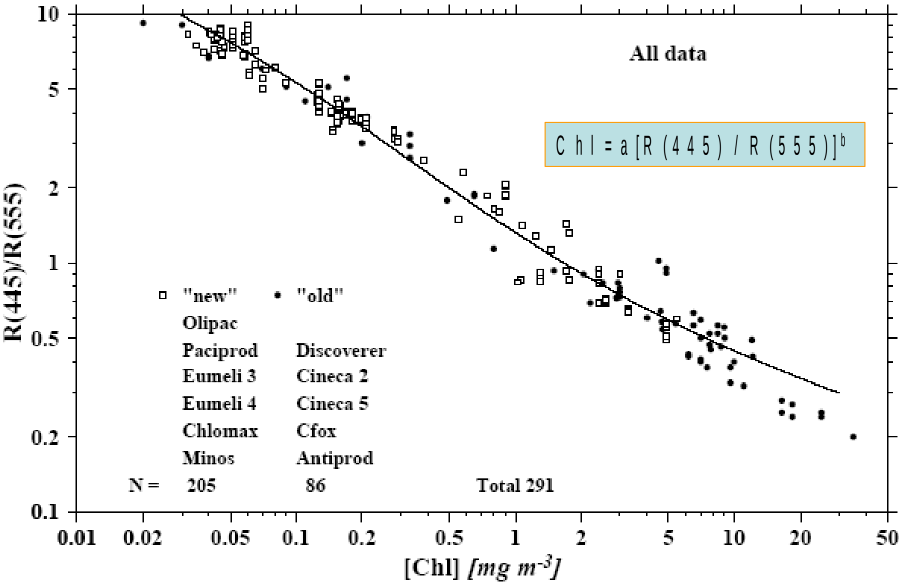

Fig. 3: Scatterplot of the chlorophyll concentration and the reflectance ratio from various cruises, in Morel / Antoine MERIS Case 1 water ATBD

All other water types, for which one component is not sufficient to describe the optical variability, are called case 2 water. Then a simple band ratio algorithm is not sufficient and other procedures have to be designed to determine multiple components from the reflectance spectrum, which is the main topic of this report.

However, the number of independent components in the water, which determine the reflectance can be larger than the number of components, we are able to derive. It depends on the concentration and its specific absorption or scattering coefficients (i.e. absorption or scattering coefficient per concentration unit, which are also variable and often unknown) to what extend a water constituent contributes to the reflectance spectrum. Furthermore, some of the water constituents have similar absorption and scattering properties, which makes it difficult to separate them from a mixture of these substances and from a reflectance spectrum. Because of these issues it is necessary to combine different but optically similar water constituents to one optical component. This means that we have to reduce the multitude of water constituents to a few optical components, the absorption and scattering spectra of which should then clearly be distinguishable. Fig. 4 illustrates this relationship in a bio-optical model.

But even with such a reduced bio-optical model, it depends on the mixture of these components, if we are able to retrieve them all. In cases when one component is dominant, e.g. waters with high concentrations of suspended matter, the effect of the other components on the reflectance spectrum, e.g. phytoplankton chlorophyll, may be so small, that it is impossible to determine the concentration with sufficient accuracy. This is the masking effect of a dominant component (Fig. 20).

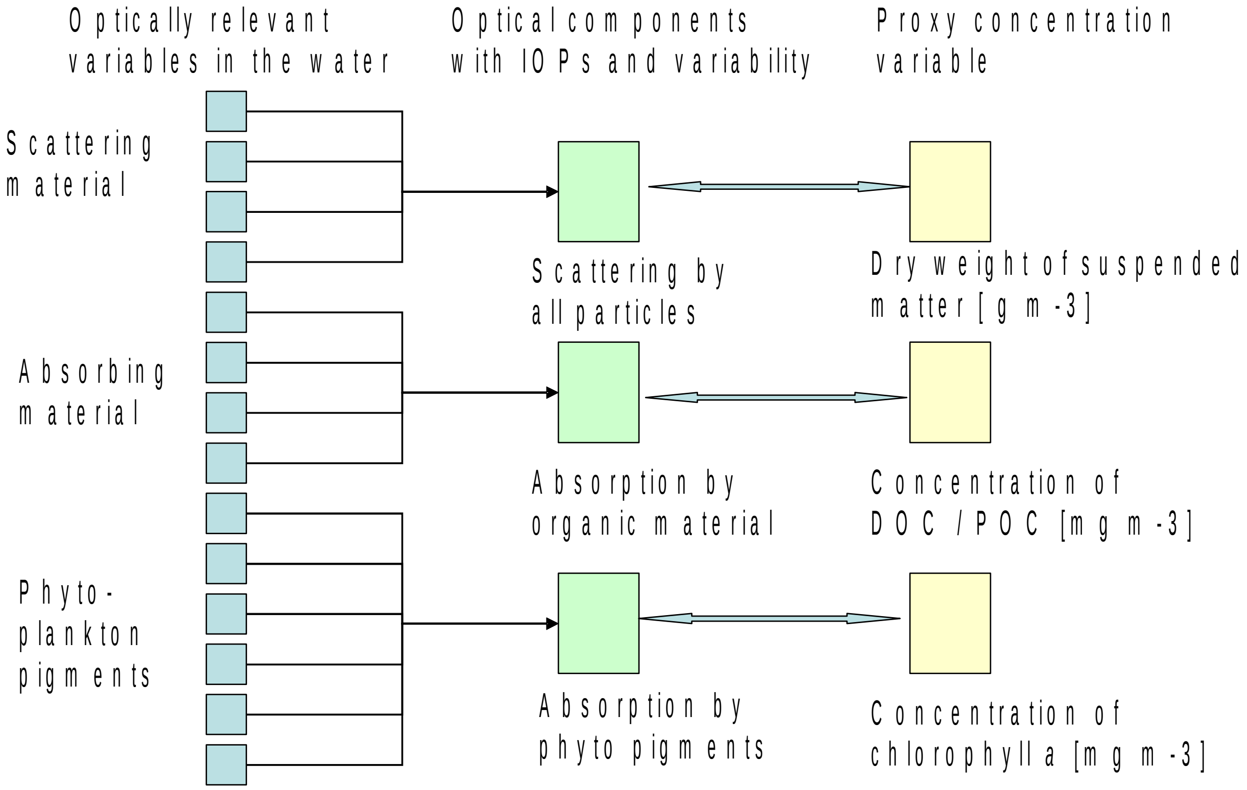

Fig. 4: Case 2 water bio-optical model: relationship between water constituents, their representation as optical components and their proxy concentration variables

Another issue is the saturation effect. The reflectance spectrum is always determined by the optical properties of pure water, which are assumed to be constant, and the optical properties of a water constituent, which change with concentration. With increasing concentration the properties of the constituent become dominant and the overall absorption and scattering of the mixture of water and constituents approaches the property of the water constituent and finally become constant. Then a further increase in concentration will not change the reflectance. Since the optical properties of pure water change with wavelength with a strong increase in absorption at wavelength > 700 nm, also this saturation effect is wavelength dependent. With increasing concentration of suspended particles the reflectance levels off in the blue spectral range at much lower concentrations than in the red spectral range (Fig.5).

Other issues of coastal water remote sensing are reflection by the sea bottom in the case of shallow water and the more complex atmospheric correction due to the higher variability of aerosols compared to the open ocean atmosphere and the adjacency effect of land.

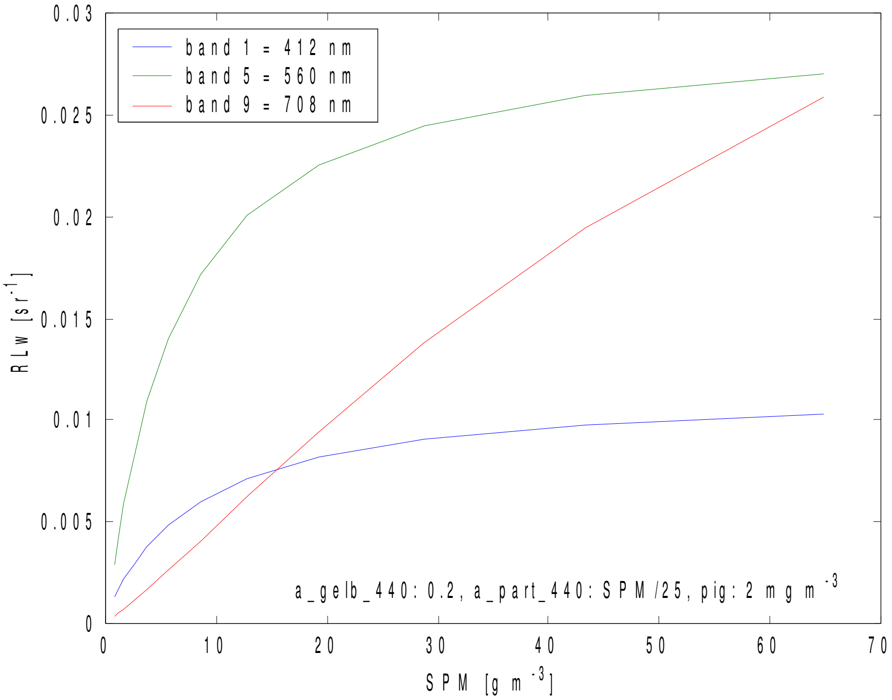

Fig. 5: Saturation effect: with increasing concentration of suspended matter (SPM) the reflectances at 3 spectral bands of MERIS approach the reflectances of the pure SPM, but dependent on the pure water absorption coefficient, which increases with wavelength

3 Approaches and existing algorithms

As discussed above, coastal water remote sensing involves a number of problems and restrictions so that one cannot expect to solve all issues with one algorithm. The inversion of water reflectances can be solved with different approaches, some of them are based on empirical relationships between observations, and others are based on a bio-optical model and the simulation of reflectance spectra. In all cases algorithms have a limited scope with respect to the water type, concentration range, the bio-optical model used for the algorithm development and conditions like solar angle, sea state, and atmosphere, which has to be respected. Conditions, which are out of scope of the respective algorithm, have to be detected and indicated by flags. Also the uncertainty of the results has to be quantified.

A special task is the development of algorithms for the correction of the influence of the atmosphere above coastal water. Finally procedures are necessary to test and validate the algorithms.

3.1 Empirical algorithms

3.1.1 Band ratio algorithms

The most common empirical algorithms are based on ratios of the reflectances at different spectral bands, called band ratio algorithms. One typical representative is the OC4 algorithm used by the Ocean Colour Group of NASA to produce maps of the chlorophyll concentration from SeaWIFS and MODIS data (O'Reilly et al., 1998).

For the linear regression

$$\log<|\mathrm{Chl\ a}> = a + b\cdot\log<R443/R555>$$

the coefficients a and b have to be determined from field measurements of chlorophyll and corresponding reflectances. From the scatter of the data also the uncertainty can be determined.

Since the relationship between the logarithm of the chlorophyll concentration and the logarithm of the green/blue band ratio is not sufficiently linear over the full concentration range of 0.03 – 30 mg m^{-3} 3 band ratios with different spectral bands are used.

The logarithm of the maximum of the remote sensing reflectance (R_{rs}) ratio is then used to compute the logarithm of the chlorophyll concentration C_a with a 4th order polynomial.

R_{max} =Maximum of [R_{rs}-ratio(443/555, 490/555, 510/555)]

$R_L = \log_{10}(R_{max})$

$\mathrm{Log}_{10}(C_a) = 0.366 + 3.067\,R_L + 1.930\,R_L^2 + 0.649\,R_L^3 - 1.532\,R_L^4$

A band ratio algorithm could also be used to determine the concentration of other water constituents and optical properties, e.g. suspended matter, by using bands in the red spectral range. Further relationships can be determined between a band ratio and absorption and scattering coefficients, the irradiance attenuation coefficient k_d, the Secchi disc depth, and the water depth.

Since all optical active water substances influence all spectral bands in the visible spectral range, simple band ratios are applicable only, when the used bands are affected by only one constituent or by one component, which may consists of more but co-varying substances.

3.1.2 Line height algorithms

Phytoplankton pigments do not show only absorption bands but also fluorescence, which is induced by the absorbed sun light. It can be determined in the form of a line height algorithm. This type of algorithm determines the absorption or the fluorescence, which occur in relatively small spectral bands in the red and NIR spectral range. For this purpose a base line is computed from the neighbouring spectral bands, which are not affected by this absorption or fluorescence band. This type of algorithm was first developed by Neville and Gower (1977) to determine the sun light induced fluorescence of phytoplankton chlorophyll in the form of the Fluorescence Line Height (FLH) and to relate this to the chlorophyll concentration. In fact this technique was one of the design drivers for MERIS with its fluorescence band at 681 nm.

The general line height algorithm is based on radiances L at 3 spectral bands, λ_1. λ_2, λ_3. The center band, λ_2, is located at the wavelength with maximum absorption or scattering / fluorescence. The two other bands are located at wavelengths around λ_2, where the effect is low.

$$\mathrm{LH} = L_2 - L_1 - (L_3 - L_1)\,(\lambda_2 - \lambda_1) / (\lambda_3 - \lambda_1)$$

For the FLH for MERIS the 3 spectral bands are: $\lambda_1 = 665$ nm, $\lambda_2 = 681$ nm, $\lambda_3 = 708$ nm.

For the FLH Gower(2003) recommends a correction coefficient $k = 1.005$ to reduce the effect of thin clouds.

$$\mathrm{FLH} = L_2 - k\,[L_1 + (L_3 - L_1)\,(\lambda_2 - \lambda_1) / (\lambda_3 - \lambda_1)]$$

The FLH works for chlorophyll concentrations > 0.5 mg m^{-3}, but one has to consider that the fluorescence efficiency, i.e. the emitted energy per unit of chlorophyll, is variable due to physiological effects of phytoplankton but also due to other absorbing material in the water, which might reduce the excitation in the blue spectral range, so that the conversion from FLH to chlorophyll concentration has to be validated with ground samples. Nevertheless, FLH is a signal, which is very specific to chlorophyll. Another advantage is that in can be used also for shallow waters, where the penetration depth of light is limited by the high absorption of pure water in the red spectral range, and it can be used without atmospheric correction.

When the chlorophyll concentration increases to very high levels, part of the fluorescence light is quenched due to the broadening absorption band of chlorophyll around 670 nm. In these extreme cases, the reflection spectrum shows a maximum in the MERIS band at 708 nm and a minimum in the bands at 665 and 681 nm. This feature has been used to define the Maximum Chlorophyll Index (MCI). It is formulated, similar to FLH, as:

$$\mathrm{MCI} = L_2 - k\,[L_1 + (L_3 - L_1)\,(\lambda_2 - \lambda_1) / (\lambda_3 - \lambda_1)]$$

with the 3 spectral bands for MERIS: $\lambda_1 = 681$ nm, $\lambda_2 = 708$ nm, $\lambda_3 = 753$ nm.

The MCI has been successfully applied to detect and map the distribution of floating cyanobaceteria (Fig 6 and Fig. 7,) such as present in the Baltic Sea during summer months, or floating Sargassum macro algae in the Sargasso Sea of the mid Atlantic ocean (Gower & King, 2011). Thus, the MCI is an important algorithms also for Coastcolour for cases with very high chlorophyll concentrations.

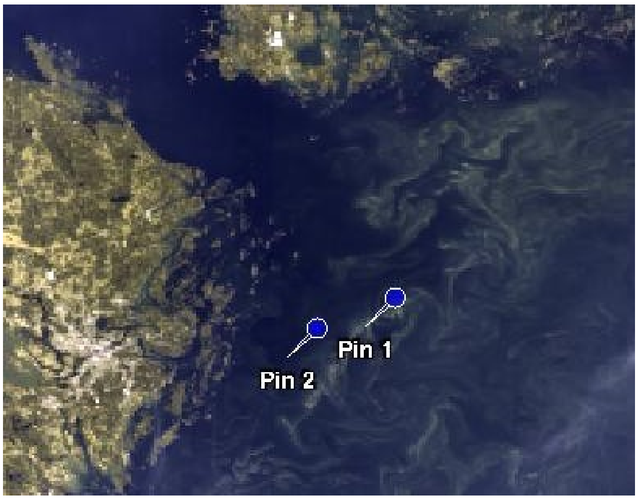

Fig. 6: Baltic Sea MERIS L1b RGB image of 20070806, Pin 1 inside, Pin 2 outside cynobacteria bloom

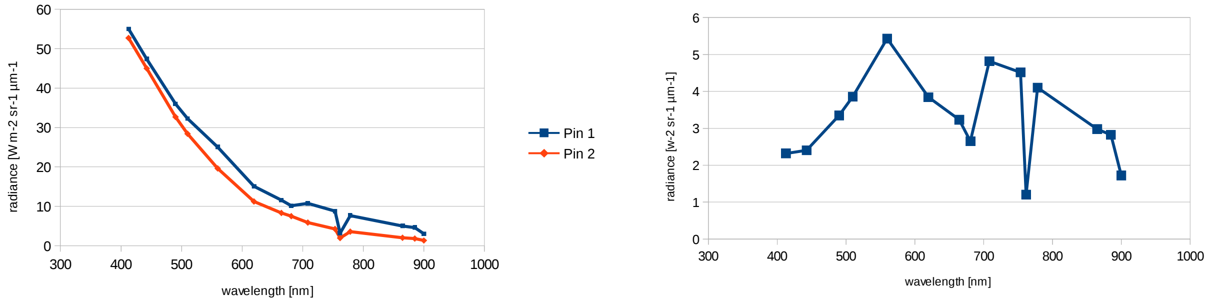

Fig. 7: Baltic Sea radiance spectra (MERIS L1b 20070806), left: radiance spectra inside (Pin 1) and outside (Pin 2) of a cyanobacteria bloom, right: difference between the two spectra

3.1.3 Neural Networks trained with measured reflectances and concentrations

A special kind of empirical methods for determining multiple variables from reflectance spectra are artificial neural networks (NN, Fig. 9). In principle a NN, as it used here, is a regression method, but instead of using pairs of 2 variables it allows the use of multiple variables on both sides of the equation. A NN is constructed of a multiple of neurons, which are non-linear functions, each of which can receive multiple inputs and have one output. These neurons are interconnected and ordered in the form of layers with one input and one output layer and 1 or more hidden layer(s) in between.

When such a network is constructed, the coefficients of each neuron have to be determine by an optimization process (“training of a NN”) of known pairs of data so that the difference between the values of the output layer and the corresponding true data are minimized for a large number of cases. Input of such a NN can be the reflectances at a number of spectral bands and output the IOPs or concetrations of water constituents, or the irradiance attenuation coefficients, kd, or the Secchi disc depth. Such a NN can be trained with measured concentrations or IOPs and corresponding reflectance spectra. However, in many cases the data set for training will comprise only a few samples, which will not be evenly spaced over the full range of concentrations. A NN, which is based on such a limited data set will work very good for these data, but, due to the many gaps in the training data set, its interpolation capability will be low and its universal use will be very limited. Examples of such NN can be found in Morozov et al., 2010, d'Alimonte & Zibordi, 2003.

The alternative is a NN, which is trained with simulated data (s. below).

3.1.4 Linear combinations of reflectance bands

This type of algorithm is also an empirical multiband algorithm. It is based on the Principle Component Analysis (PCA), Doerffer & Murphy (1989), Doerffer et al. (1989), . The initial step is the co-variance matrix of of all reflectance bands of interest of a scene. A high co-variance between the bands indicate that only a few factors determine the variability of the reflectance spectra, the number of which can be estimated by computed the eigenvectors and eigenvalues. Based on the selected most significant eigenvectors weights for a linear combination of the normalized reflectances can computed and used to convert the normalized reflectances into factors, which then again represent the most significant factors, which determine the spectral variability. With the help of in situ data these factors can be interpreted in the form of IOPs or concentrations of water constituents.

The result and success of this method depends very much on the data set. As it is the case for all statistical methods, interesting features in the water might get lost, if they do not contribute to the co-variance matrix because of their low frequency. It can be useful to apply this method to the logarithm of the reflectances to consider the non-linear nature between IOPs and reflectances.

The method can also be applied to simulated reflectance data. In this case the variability of the reflectance spectra can be controlled and since the underlying IOPs or concentrations are known they can be related much better to the eigenfactors / linear combinations s. Fischer et al., (1986) and Krawczyk et al. (2007) .

The method is similar to the neural network which are trained with simulated data except that a NN has the advantage to better describe the non-linear relationship between IOPs and reflectances.

3.2 Semi-analytical algorithms

Semi-analytical algorithms use a simplified model to compute reflectances from IOPs, which are in most cases the absorption a and the backscattering coefficients b_b.

$$Rrs = Lw(0+)/Ed(0+) = g\frac{b_b}{a}$$

Rrs is the water leaving radiance reflectance, also called remote sensing reflectance, Lw is the water leaving radiance just above the sea surface (without the specularly reflected radiance) and Ed the downwelling irradiance just above the surface. They are called semi-analytical because the empirical coefficients summarized under *g* of this type of models have to be determined by observational data sets. The models can be used in forward mode together with an optimization procedure to fit the measured reflectance spectrum (s. below) or in a decomposition mode.

3.2.1 Semi-analytical decomposition algorithm (QAA)

The Quasi-Analytical Algorithm (QAA) was developed by Lee et al. (2002) to derive inherent optical properties of optically deep waters. It uses a number of empirical coefficients. Here the description is partly quoted from IOCCG Report 5 (Lee et al., 2006).

QAA separates the inversion process into two consecutive sections. The first section is the derivation of coefficients of total absorption, *a*, and backscattering, b_b. The second section decomposes the total absorption coefficient from the first section into its major components, i.e. absorption by phytoplankton pigments and gelbstoff.

QAA starts with the calculation of *a* at a reference wavelength (λ_0, 555 or 640 nm), with the assumption that remote-sensing reflectance at this wavelength is well measured from a remote-sensing platform.

The total absorption coefficient at λ_0 is expressed as

$$a(\lambda_0) = a_w(\lambda_0) + \Delta a(\lambda_0),$$

where $a_w(\lambda_0)$ is the contribution from water molecules (Pope and Fry, 1997), and $\Delta a(\lambda_0)$ represents the contribution from dissolved and suspended constituents. For this $a(\lambda_0)$, errors in its estimation are limited as long as $a_w(\lambda_0)$ makes up a big portion (at least one third of the total).

Lee et al. (2002) proposed two λ_0 for dealing with IOP inversion: 555 nm for oceanic and most coastal waters and 640 nm for waters with high absorption coefficients ($a(440) > \sim 0.5\ \mathrm{m}^{-1}$).

When 555 nm is selected as λ_0 , *a*(555) is estimated using the Morel-Maritorena approach (Morel and Maritorena, 2001). First $k_d(555)$, the downwelling irradiance attenuation coefficient, is estimated with ρ = log(max(Rrs(440, 490, 510))/Rrs(555)),

$$k_d(555) = 0.0605 + 10^{-1.163-1.969p+1.239p^2+0.417p^3-0.984p^4} ,$$

and then

$$a(555) = \frac{0.9kd(555)(1 \quad 6.8Rrs(555))}{1+15.3Rrs(555)}$$

The subsurface remote sensing reflectance, $r_{rs}(\lambda)$, is calculated from $R_{rs}(\lambda)$ through

$$r_{rs}(\lambda) = R_{rs}(\lambda)/(0.52 + 1.7R_{rs}(\lambda)),$$

where 0.52 and 1.7 are empirical values derived from data simulated by Hydrolight (Lee et al., 1999). $r_{rs}(\lambda)$ can be modelled as a polynomial function of $b_b/(a + b_b)$ (Gordon et al., 1988; Lee et al., 1998a), then $b_b/(a + b_b)$ (represented as symbol *u*) can be calculated algebraically from $r_{rs}(\lambda)$ (Hoge and Lyon, 1996; Lee et al., 2002),

$$u(\lambda) \equiv \frac{b_b(\lambda)}{b_{(\lambda)} + a(\lambda)} = \frac{-0.0895 + \sqrt{0.008 + 0.499\, r_{rs}(\lambda)}}{0.249}$$

The spectral $b_b(\lambda)$ is modelled with the widely used expression (Smith and Baker, 1981; Gordon and Morel, 1983),

$$b_b(\lambda)=b_{bw}(\lambda)+b_{bp}(\lambda_0)\left(\frac{\lambda_0}{\lambda}\right)^y$$

where b_{bw} and b_{bp} are the backscattering coefficients of pure seawater and suspended particles, respectively. Values of $b_{bw}(\lambda)$ are provided in Morel (1974). When $a(\lambda_0)$, $u(\lambda_0)$, and $b_{bw}(\lambda_0)$ are known, $b_{bp}(\lambda_0)$ can be easily derived. The values of $b_b(\lambda)$ at other wavelengths are then calculated when the wavelength exponent (Y) is estimated from Lee et al. (2002)

$$y=2.2\left(1-1.2\exp\left(-0.9\frac{r_{rs}(440)}{r_{rs}(555)}\right)\right)$$

Finally, applying $b_b(\lambda)$ back to $u(\lambda)$ (derived from $r_{rs}(\lambda)$), the total absorption coefficient at wavelength λ, $a(\lambda)$, is calculated algebraically,

$$a(\lambda)=\frac{(1-u(\lambda))b_b(\lambda)}{u(\lambda)}$$

The decompositon of $a(\lambda)$ into the 2 components $a_{pig}(440)$ and $a_{dg}(440)$ uses the spectral absorption ratio of these two components with

$$\zeta = a_{pig}(410) / a_{pig}(440) \quad \text{and} \quad \xi = a_{dg}(410) / a_{dg}(440)$$

The value of ζ is estimated using the spectral ratio of $r_{rs}(440) / r_{rs}(555)$ based on the field data (Lee et al., 1998b):

$$\zeta=a_{pig}(410)/a_{pig}(440)=0.71+\frac{0.06}{0.8+r_{rs}(440)/r_{rs}(555)}$$

The value of ξ is calculated after the spectral slope S (used to describe the spectral shape of a_{dg}()) is selected (0.015 nm^{-1} is used here):

$$\xi=a_{dg}(410)/a_{dg}(440)=\exp(S(440-410))$$

When the values of $a(410)$, $a(440)$, and are known, $a_{pig}(440)$ and $a_{dg}(440)$ are calculated algebraically,

$$a_{dg}(440)=\frac{(a(410)-\zeta a(440))}{\xi-\zeta}-\frac{(a_w(410)-\zeta a_w(440))}{\xi-\zeta},$$

and

$$a_{pig}(440)=a(440)-a_{dg}(440)-a_w(440)$$

3.3 Inversion of a forward model

3.3.1 Linear Matrix Inversion

The simple reflectance model developed by Gordon et al. (1988) can be also used to construct a matrix, which includes the absorption coefficients of phytoplankton, a_{pig}, and of Coloured Organic Matter (COM), i.e. of detritus + gelbstoff (a_{dg}), of the backscattering coefficient of all particles, b_{bp}, of the constant absorption and backscattering coefficient of pure sea-water, a_w and b_w respectively, and the water reflectance, r_{rs}, for a set of wavelengths. This matrix can be inverted pixel by pixel to get the 3 IOPs a_{pig}, a_{dg}, b_{bp} from the reflectance r_{rs} at 3 wavelength.

The technique is described in IOCCG report No. 5 by Lyon & Hoge (2006) and has been developed by Hoge & Lyon, 1996. It uses the 3 spectral bands at 412, 490 and 555 nm.

The reflectance is modelled by 2 terms:

$$r_{rs}=g1\left(\frac{b_b}{b_b+a}\right)+g2\left(\frac{b_b}{b_b+a}\right)^2$$

if one sets again:

$$u\equiv\frac{b_b}{b_b+a} \text{ and } v\equiv 1-\frac{1}{u}$$

then by re-arranging one can construct:

$$a_{pig}(\lambda)+a_{dg}(\lambda)+b_{bp}(\lambda)v(\lambda)=-a_w(\lambda)-b_{bw}(\lambda)v(\lambda)$$

When the 3 spectral bands are used and all IOPs are normalized with the respective band at 412 nm, one can set up a matrix, which can be inverted to derive the IOPs from rrs:

$$\begin{bmatrix} 1 & 1 & v(\lambda_r) \\ a^o_{pig}(\lambda_2) & a^o_{dg}(\lambda_2) & b^o_{bp}(\lambda_2) \\ a^o_{pig}(\lambda_3) & a^o_{dg}(\lambda_3) & b^o_{bp}(\lambda_3) \end{bmatrix} \begin{bmatrix} a_{pig}(\lambda_r) \\ a_{dg}(\lambda_r) \\ b_{bp}(\lambda_r) \end{bmatrix} = -\begin{bmatrix} a_w(\lambda_r)+b_{bw}(\lambda_r)v(\lambda_r) \\ a_w(\lambda_2)+b_{bw}(\lambda_2)v(\lambda_2) \\ a_w(\lambda_3)+b_{bw}(\lambda_3)v(\lambda_3) \end{bmatrix}$$

where a^o and b^o are the normalized IOPs with respect to the reference wavelength λr. For the inversion an LU decomposition technique was used (Hoge & Lyon, 1999).

3.3.2 Reflectance spectrum fitting

This type of algorithms fits a modelled reflectance spectrum to a measured spectrum by modification of the IOPs or concentrations iteratively using an optimization technique.

To start the iteration a first guess of the IOPs is used, which should be as close as possible to the solution. The success is determined by computing the root means square (RMS) difference between the measured and modelled reflectances of all or a selected number of spectral bands. The iteration ends when the RMS difference is smaller than a pre-defined value, when its change from loop to loop is smaller than a pre-defined threshold, or when a pre-defined number of iterations is reached. The IOPs at the end are then used as the solution and the remaining RMS difference as a quality index. During the descend side valleys with local RMS minima can be reached, which leads to multiple solutions / ambiguities.

In principle any model can be used, which relates the reflectances to the IOPs. A complex radiative transfer model is much too slow to be applied to full scenes. Thus, simplified models such as $r = b_b/(a+b_b)$ or the two-flow model of Joseph (1950) and Doerffer(1979) are used. Such a method was successfully applied to data of the Coastal Zone Colour Scanner (CZCS) by Doerffer & Fischer (1994).

An alternative is a fast neural network with the IOPs as input and reflectances as output. Such a NN can be trained either with measured or simulated data. For the simulation a complex be-directional radiative transfer model can be used with the advantage that also the solar and observation angles can be input to the NN and any parametrisation of the angular dependence, as it is necessary for simply reflectance models, can be avoided. Training with measured data is not appropriate in most cases because the number of observations and the frequency distribution with respect to the IOPs and angles is not sufficient to train a NN with sufficient generalisation power.

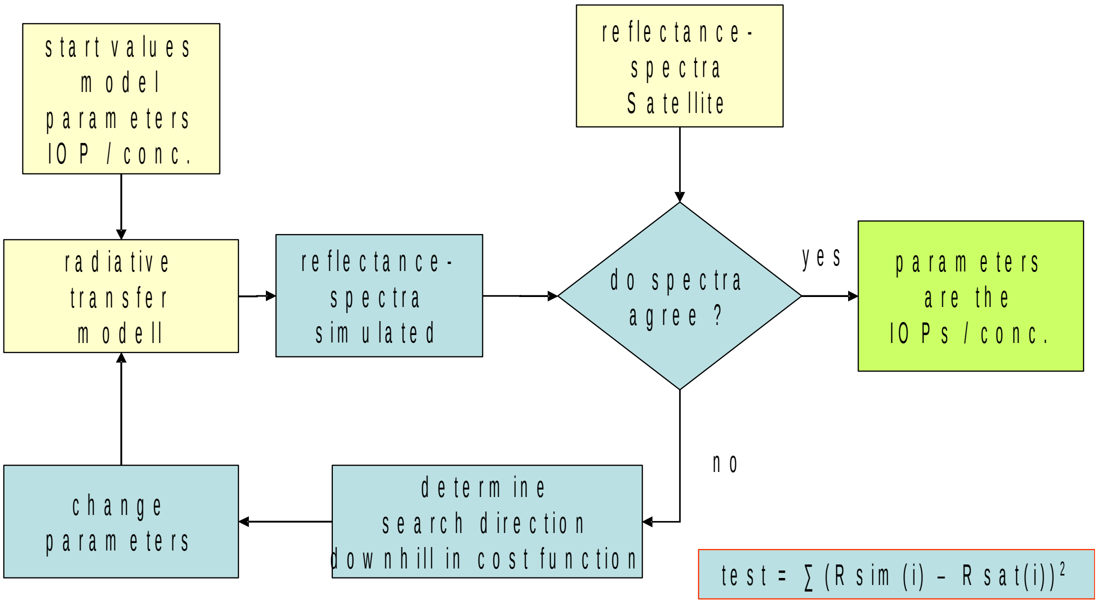

Fig. 8: Scheme of inverse modeling using a forward model and an optimization procedure

3.4 Table look-up procedures

Simulations of reflectances using a radiative transfer model can be ordered in the form of a look-up table. Search in such a table for a reflectance spectrum which best fits to the measured spectrum also leads to the IOPs or concentrations, which are associated with the reflectance spectrum. Such a technique has been applied by Lesser and Mobley (2007) to coral reefs. Speed and success of this method depends on the number of inputs and the structure of the table as well as the search strategy.

3.5 Inversion by an inverse neural network

Neural networks can be designed and trained also in the inverse mode, i.e. the dependent reflectances are the input and the IOPs are the output. Training can be based on measured IOPs and corresponding reflectance spectra or with simulated data. In most cases measured data are not appropriate, because their range and frequency distribution are not sufficient. NNs, which are trained with observations, are in most cases only valid for these data.

$$y_l = s\left(-d_l + \sum_{k=1}^{3} w_{kl} \cdot s\left(-c_k + \sum_{j=1}^{5} v_{jk} \cdot s\left(-b_j + \sum_{i=1}^{4} u_{ij} x_i\right)\right)\right)$$

During the training process, the coefficients of all neurons have to be adjusted so that the root mean square difference between the output of the NN (x_i) and the corresponding variables of the model (y'_i) becomes minimum.

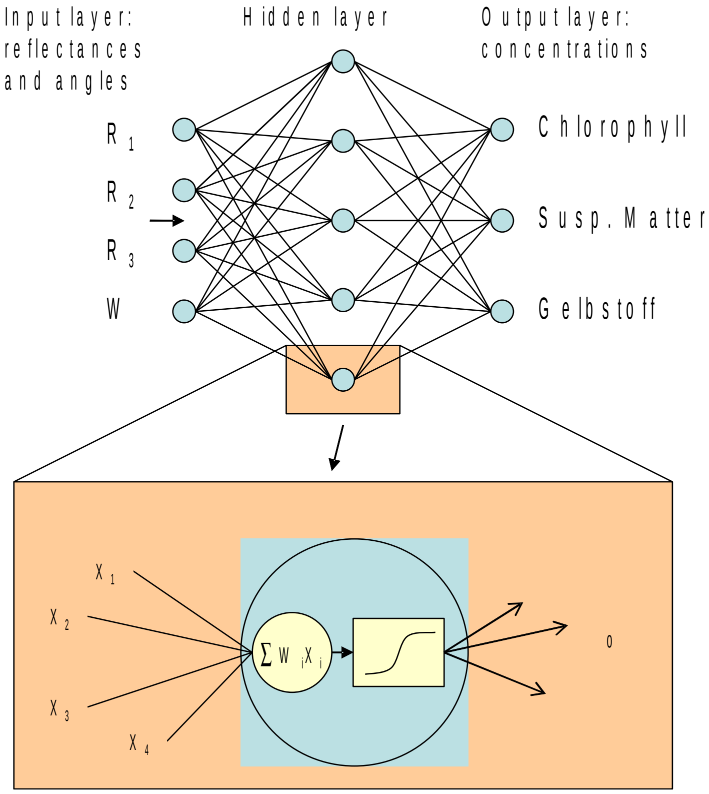

Fig. 9: Principle structure of a neural network and a single neuron

Inverse neural networks were used as the core algorithm of the Coastcolour project for the atmospheric correction as well as the retrieval of IOPs and kd. They are based on simulated training data (Schiller & Doerffer, 1993, 1997, Doerffer & Schiller, 2007). The results of the IOCCG test is described in Doerffer & Schiller, 2006.

4 Characterisation of the area and water types

As mentioned above, it is hardly possible to develop an universal algorithm, which would work in all types of water and would produce data with the same quality. Thus, for the development of an algorithm it is necessary as a first step to define its anticipated scope with respect to the water type, concentration range, atmospheric conditions etc. For all algorithms, which are based on IOP-reflectance models, as it is the case for optically complex waters, one has to define and set up a bio-optical model. This model must comprise

- Dominant water constituents
- Concentration ranges
- Concentration matrix -> any dominant components
- Co-variances
- Occurence of exceptional events (blooms), floating material
- Water depth

Furthermore it has to be investigated if any specific atmospheric properties, which may occur close to the coast, are important for the atmospheric correction procedure, this may include:

- Industrial smoke like soot
- Desert dust
- Biomass burning
- Volcanoe smoke

4.1 Definition of the bio-optical model

A large number of different water constituents are present in particular in shallow soft bottom coastal waters with the influence of estuaries. Since it is impossible to determine all of these substances independently because of their optical similarity, groups of substances with similar optical properties have to be merged and represented by an optical component. This critical step is illustrated in Fig. 4.

The spectral absorption and scattering coefficients of the optical components can be complemented with the variability of their constituents. In the case of a neural network algorithm the variability can then be incorporated in the simulated training data set.

Finally the relationship between the IOPs and the mass concentrations, e.g. chlorophyll, DOC, or suspended matter dry weight, has to be determined from in situ measurements and water samples.

A critical aspect is the co-variation between the constituents, which form an optical component, and also between the optical components. In the extreme case, the co-variation is so high that one component is sufficient to describe the variability of the reflectance spectra, as it is the case – by definition – in case 1 water. But also the optical dominance of a component may explain that only 1 component is sufficient to describe the variability of the reflectance spectra for a certain area. In case 1 water it is the absorption by phytoplankton pigments, in turbid coastal water it may be the scattering and absorption by suspended matter.

4.2 Out of scope conditions

Most of the bio-optical models are defined for certain standard conditions of the global ocean or a specific coastal area. However, exceptions of these standard conditions have to be taken into account either by flagging these pixels or by an alternative algorithm.

In some areas, like the Baltic Sea, floating cyanobacteria occur at the surface during calm and sunny summer periods. These events mask the underlying water body and require totally different algorithms for the atmospheric correction as well as for the retrieval of IOPs or water constituents. Furthermore, a volume related concentration quantity may then no longer be applicable, but a coverage related dimension. Other events may be red tides formed by *Noctiluca* plankton or harmful algal blooms (HABS). If it is known that such exceptional events occur, dedicated algorithms to detect these cases should be included in the procedure.

Another issue of coastal water remote sensing, which may lead to errors, is the reflection by the sea floor in shallow water. Then it is difficult if not impossible to separate the influence of water constituents from the effect of the sea bottom. Areas with a water depth less than the maximum signal depth z90max, i.e. the depth from which 90% of the upward directed radiance at the surface comes from, should be masked out using bathymetric maps. An alternative is to use only spectral bands in the red part of the spectrum, as e.g. the fluorescence line height, for which z90max is limited by pure water absorption to about 2 meters.

Finally it has to be checked if stratification in the area of research within z90max may have an effect on the retrieval of water constituents since for most algorithm a homogeneous vertical distribution is assumed.

5 Step-by-step definition of the algorithm

This chapter provides a step by step overview of all items which have to be considered for optical remote sensing of coastal waters. A scheme of the steps for the atmospheric correction is shown in Fig. 10.

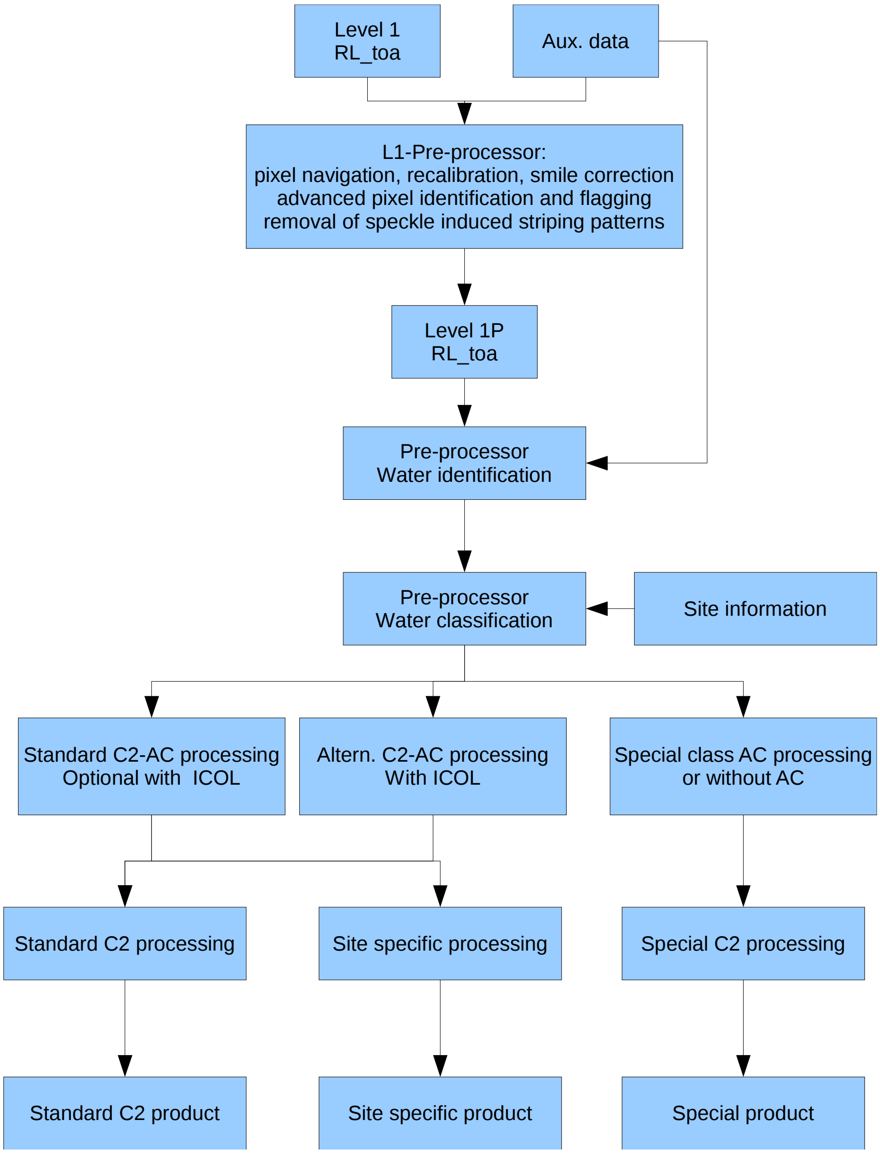

Fig. 10: Steps and decisions for the atmospheric correction of coastal water L1b spectra

5.1 Step 1: Decision about the type of water and the corresponding type of algorithm

First of all the optical type of coastal water has to be determined.

5.1.1 Clear coastal water

For rocky shores with oligotrophic deep water, which is located far from the influence of rivers and estuaries, standard products of the phytoplankton chlorophyll concentration, which are provided by the space agencies and which are produced with the case 1 algorithms, may be totally appropriate. But the atmospheric correction may cause a problem for areas close to a coast with dense vegetation due to the adjacency effect. In this case a special atmospheric correction procedure has to be applied to L1b data. One of these procedures (Santer and Schmechtig, 2000) is implemented in the BEAM processor. A more simple approach was developed by Sterckx et al. (2011).

5.1.2 Clear water with bottom reflection

In the case of oligotrophic but shallow water the reflection of the bottom has to be considered (Joseph, 1950, Maritorena et al., 1994). Then these areas have to be masked out or a special algorithm, which takes the bottom reflection into account, has to be used. This might be of interest for the generation or update of bathymetric maps, mapping of benthic vegetation or coral reefs. A special algorithm for mapping the sea floor in particular coral reefs under oligotrophic clear water conditons, which is based on look-up tables, has been developed by Lesser and Mobley (2007). Other multispectral procedures are described in e.g. Mumby et al., 1998, Lee et al., 1999, and Hochberg & Atkinson, 2000.

5.1.3 Waters with high concentrations of suspended and organic matter

Coastal waters with shallow soft sediment bottoms and estuaries are optically the most complex water types for remote sensing. Waters with high concentrations of suspended matter have an intense reflectance in all spectral bands, while the reflectance of waters with high concentrations of organic matter can be very low and even less than the uncertainty range of the atmospheric correction also in the blue part of the spectrum. This can cause a large error in retrieving the water reflectance from TOA reflectance. Thus, this type of water requires both a special algorithm for the atmospheric correction and one for the retrieval of IOPs. Also a procedure to determine the uncertainty range for each product is necessary, which depends on the mixture of water constituents. This uncertainty range is variable and can be very large.

In the case of turbid water with low absorption by organic constituents the standard MERIS L2 data are appropriate, which are produced with an atmospheric correction procedure, which includes already an extension for turbid water.

Standard MERIS products of water with constituents with high absorption coefficients show often negative reflectances in the blue and even in the green spectral range. These are caused when the extrapolated path radiance, which is determined from the NIR spectral bands, is slightly overestimated and then subtracted from the TOA reflectance. In this case an atmospheric correction procedure is necessary which uses the full spectrum, such as the neural network procedure, which was developed for Coastcolour.

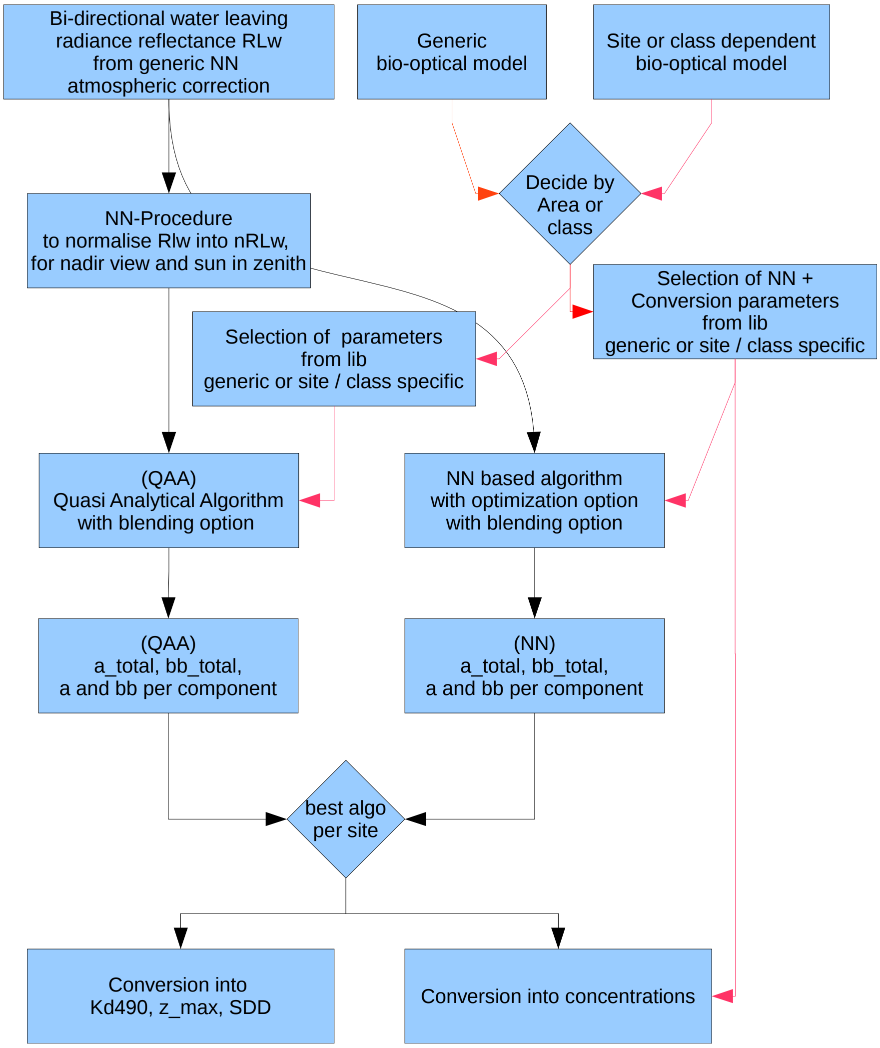

Fig. 11: Scheme and steps for algorithms to determine IOPs or concentrations from water reflectance spectra (after atmospheric correction)

For the retrieval of IOPs in both cases a multi-parameter retrieval algorithm has to be used, which can be a multi-band reflectance ratio algorithm, a model based optimization best fit algorithm, a linear matrix inversion or an inverse neural network.

For Coastcolour the semi-analytical decomposition algorithm and the inverse neural network algorithm were implemented (Fig. 11).

5.1.4 *Water with exceptional plankton blooms*

Red tides, coccolithophoride and cyanobaceteria blooms have totally different optical properties than the standard phytoplankton, i.e. diatoms and dino-flagellates. Intense blooms of these types can easily be identified by visual inspection of contrast stretched RGB L1b images. If present, the area where these blooms occur, have to flagged and masked out, or a dedicated algorithm has to be used.

One option for extreme blooms are the FLH and the MCI algorithms, which are implemented in BEAM. They provide only relative concentration changes and have to be calibrated by in situ samples.

For floating cyanobacteria the MCI is the most appropriated algorithm. It can be directly applied to L1b data without atmospheric correction. Another option to quantify and map cyanobacteria blooms is the spectral unmixing procedure, which is also implemented in BEAM. For its application one has to determine a L1b spectrum at a pixel with a maximum concentration and another spectrum outside the cyanobacteria bloom. The algorithm computes the mixing ratio of both spectra, which is a measure for the coverage with floating cyanobacteria.

5.2 Guideline questions

As a guideline for the step of algorithm selection one has to ask the following questions:

- Which type of water do I have to deal with?
    - Case 1 type of water with 1 dominant component
    - complex water with many and varying components
    - water with bottom reflection
    - water with exceptional plankton blooms
- Which type of atmospheric correction is appropriate for my type of water?
    - can I rely on the water reflectances of the standard L2 or Coastcolour products
    - or do I have to develop my own AC or adapt an existing one
    - or can I work directly with L1b data in the case of the FLH or MCI algorithm.
- Is the reflection by the sea bottom an issue or even the task?
    - Which of the parameters is most variable: IOPs of water constituents, bottom reflection, or water depth?
    - Can I fix one these variables by using other data, such as a bathymetric map, or are some of the variables sufficiently constant over the research area, such as the water optical properties or the reflection of the sea bottom?
    - Which type of algorithm can I apply, which includes all three types of variables: IOPs of water, water depth, and bottom reflection (e.g. look up table, best fit technique with a forward model)?
    - Can I avoid sea bottom reflection by using the FLH or MCI algorithm or any other algorithm, which uses only the red and NIR spectral bands?
- Does the bio-optical model of the standard product or the Coastcolour product agree with the optical properties of my research area?
- Do I have to adapt the relationship between IOPs and the mass concentration of water constituents as implemented in the standard or Coastcolour product?
- Are their any parameters or conditions out of scope of the algorithm:
    - bio-optical model or concentration range
    - aerosol type and optical depth
    - solar or observation angle
    - wind speed → white caps
    - water depth

  - vertical distribution of water constituents?
- Are any test procedures and corresponding data available: flags, uncertainty indices?
- Do the data agree with already existing information?

5.3 Step 2: Atmospheric correction

It has been mentioned in several of the chapters above that the correction of the influence of the atmosphere is the most critical task in particular in the case of complex coastal waters (s. Fig. 12 - 14). The atmospheric correction is the procedure to determine the water leaving radiance of reflectance from the top of atmosphere radiance or reflectance. Thus, it includes not only the path radiance of and transmittance through the atmosphere but also the specular reflectance at the sea surface, i.e. the sky and sun glint. The success and accuracy depends mostly on the ratio of the TOA radiance and the water leaving radiance, on the presence of thin and sub-pixel clouds and on sun glint.

There are much less atmospheric correction procedures available than retrieval algorithms for IOPs or water constituents, and they are more difficult to implement. For the standard user of remote sensing data it means in most cases to accept the L2 reflectances as provided by the satellite operator or a project like Coastcolour or to use available processors such as BEAM or ODESA in the case of MERIS or SeaDAS in the case of SeaWiFS or MODIS and modify the corresponding parameters of these processors if necessary.

The following list provides a guide of items which should be checked:

- Check for cloud flagging, extend threshold if necessary.
- Check for cloud shadows, which may lead to artefacts in the final L2 product.
- Check sun glint and foam (wind).
- Check TOA RGB image for doubtful pixels /artefacts.
- Check for negative reflectances of strange reflectance spectra.
- Determine which type of AC is required, depending on type of water:
  - turbid water
  - water with exceptional high turbidity like some estuaries
  - water with high concentrations of absorbing material
  - water with exceptional plankton blooms, in particular floating cyano-bacteria.
- Check for AERONET data and compare with the optical thickness of the satellite data.
- Check for data of other satellite instruments and compare.

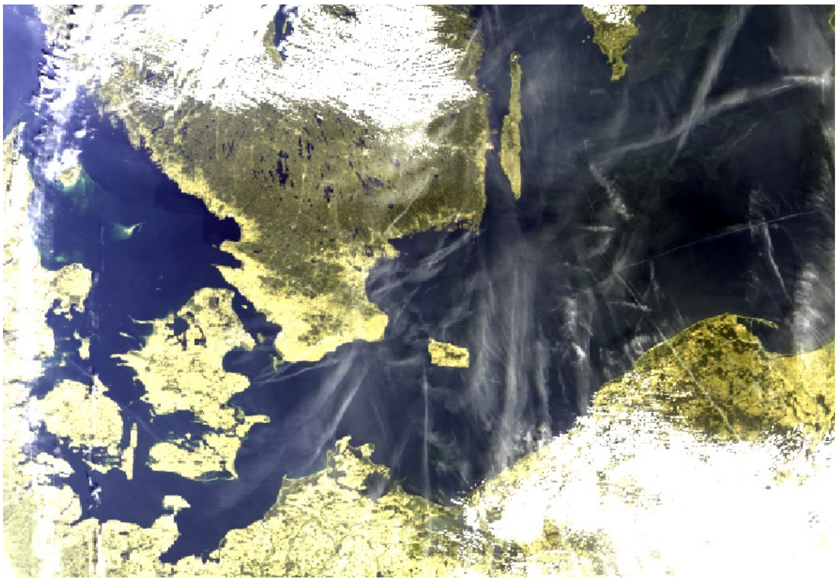

Fig. 12: A typical scene of the Baltic Sea (MERIS Rtoa, RGB, 20070806) with complex clouds and contrails and cyanobacteria blooms

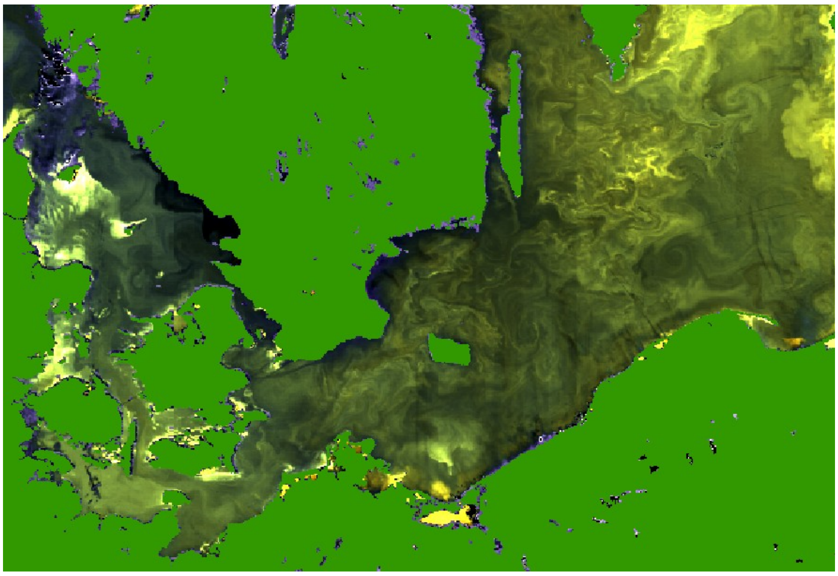

Fig. 13: The corresponding RGB image of Rwater (after atmospheric correction)

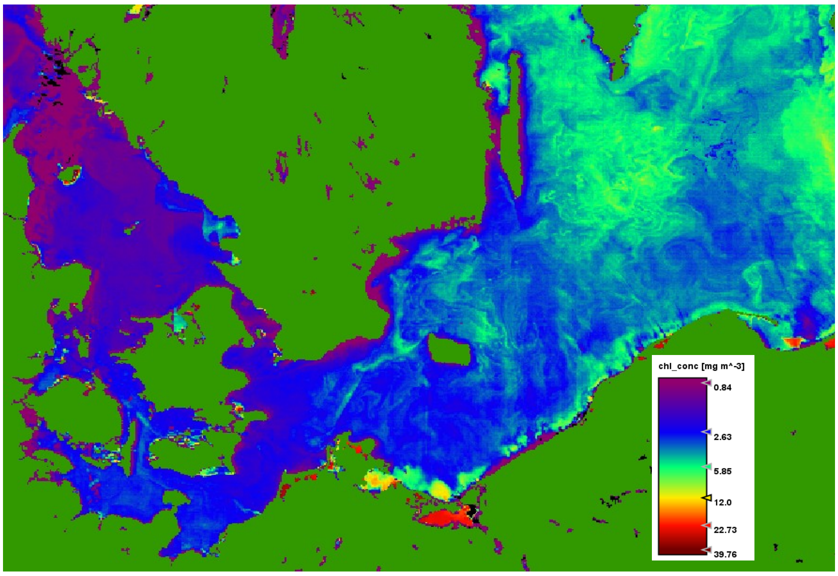

Fig. 14: The corresponding image of the chlorophyll distribution

5.4 Step 3: Bio-optical model and scope of an algorithm

Any ocean colour remote sensing algorithm has a certain scope, which is determined by the bio-optical model, the empirical relationship between e.g. chlorophyll concentration and the reflectance ratio, properties of the atmosphere or the solar and observation angles. If the algorithm is applied to a reflectance spectrum, which is based on conditions outside the definition, than this reflectance spectrum is out of scope of the algorithm. If the deviation is small, than the implied error may also be small and negligible. But large or critical deviations may lead to totally erroneous results and artefacts, e.g. a shadow of a contrail my lead to linear structure of higher chlorophyll concentration (s. Fig. 12-14).

For this reason it is necessary to check if out-of-scope conditions are present. To some extend this is performed already by co-algorithms, which indicate such a condition by flagging the corresponding pixels. But not all possible out-of-scope conditions are recognized automatically. In this case it is worth to look into the ATBD of the algorithms and check the data product using the following items:

- Check if the retrieved concentrations or IOPs are outside or very close to the boundaries of the algorithms
- Respect existing flags
- Check the conditions of atmosphere and water:
    - solar and observation angle,
    - haze and sea fog (from the optical thickness product),
    - wind (foam, glint, from the aux data),
    - floating material on the water (high reflectance in NIR spectral bands),
    - sun glint or slicks within sun glint areas.
- Re-define existing or create new flags, e.g. by setting or modifying thresholds in BEAM.

- Check if the water reflectance spectrum differs from the spectra, which are typical for the water type (e.g. in Fig. 15).
- Check for negative reflectances.
- Reproduce the spectrum in question with the retrieved IOPs and a forward model and compare: is the deviation above a certain threshold?

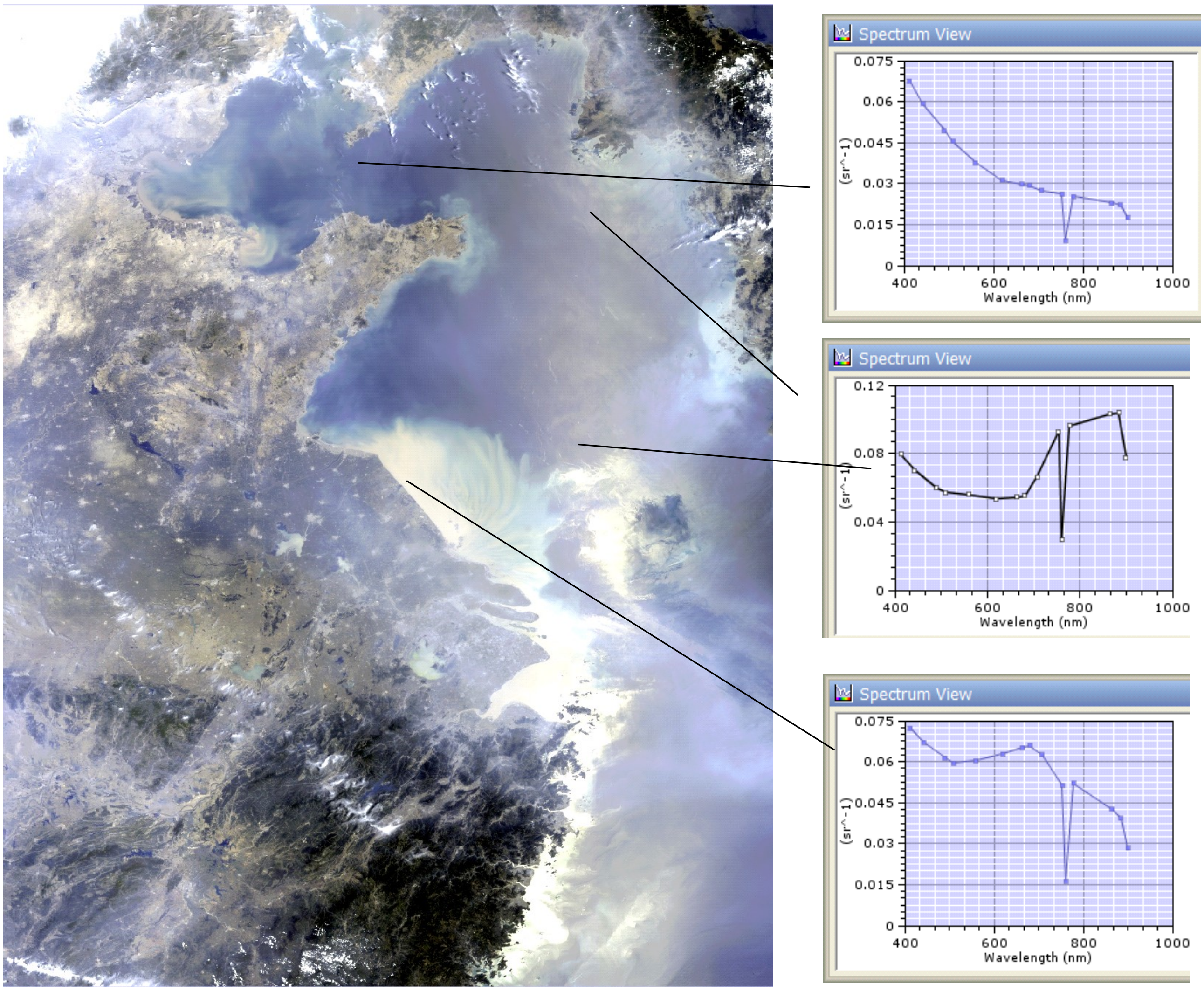

Fig. 15: MERIS scene of the Yellow Sea with TOA radiance reflectance spectra; the lower 2 spectra indicate out of scope conditons

If the scope of the algorithm is defined by a large data set, as it is the case for the training of a neural network or a PCA, then the out-of-scope conditions can be tested using an auto-associative neural network (aaNN). An aaNN is designed with the same data as input and output, e.g. for testing TOA reflectances. Furthermore, the aaNN has a bottleneck hidden layer; the number of neurons of this layer has just to be sufficient to reproduce the TOA reflectances with the output of the training data set, which was used for training of the atmospheric correction NN. This number indicates also the number of independent information, which can be derived statistically from the TOA reflectances of the training data set (Fig. 16.

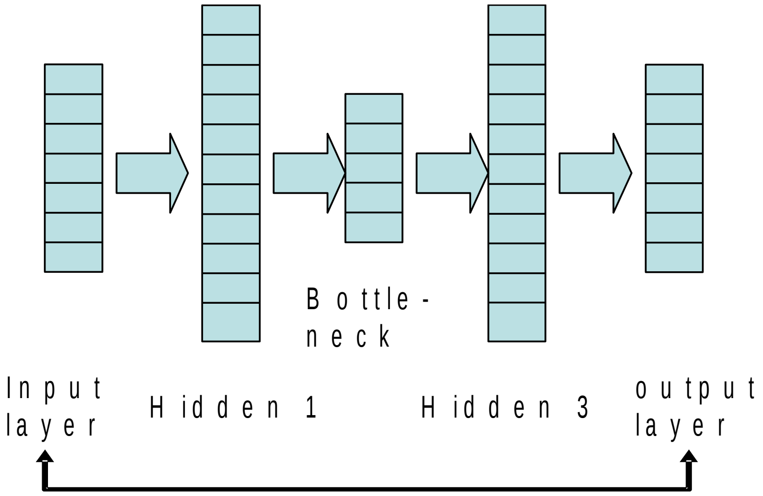

Fig. 16: Principle scheme of an auto-associative neural network aaNN

When such an aaNN is fed with a real TOA reflectance spectrum, which is out of scope of the training data set, then the deviation between the measured spectrum and the output of the aaNN increases.

An example from the Yellow Sea is given in Fig. 17 and 18. Here the coastal area with the extremely high concentration of suspended matter north of the Yangtse mouth is out of scope, while the sun glint area more to the east is still in scope, both as expected from the parameters used for simulation of the training data set.

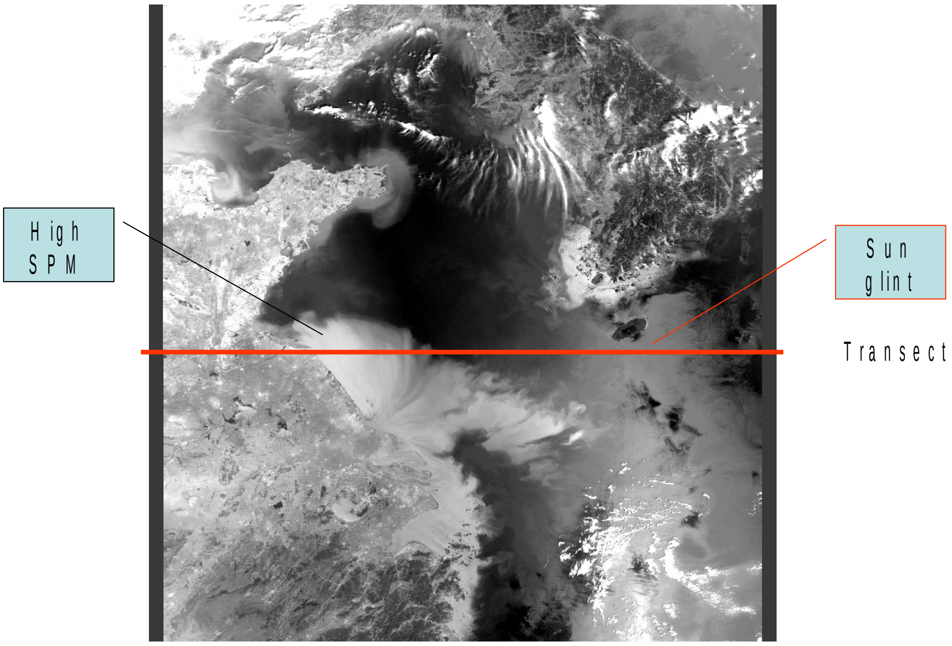

Fig. 17: MERIS L1b scene of the Yellow Sea for testing the aaNN along the transect

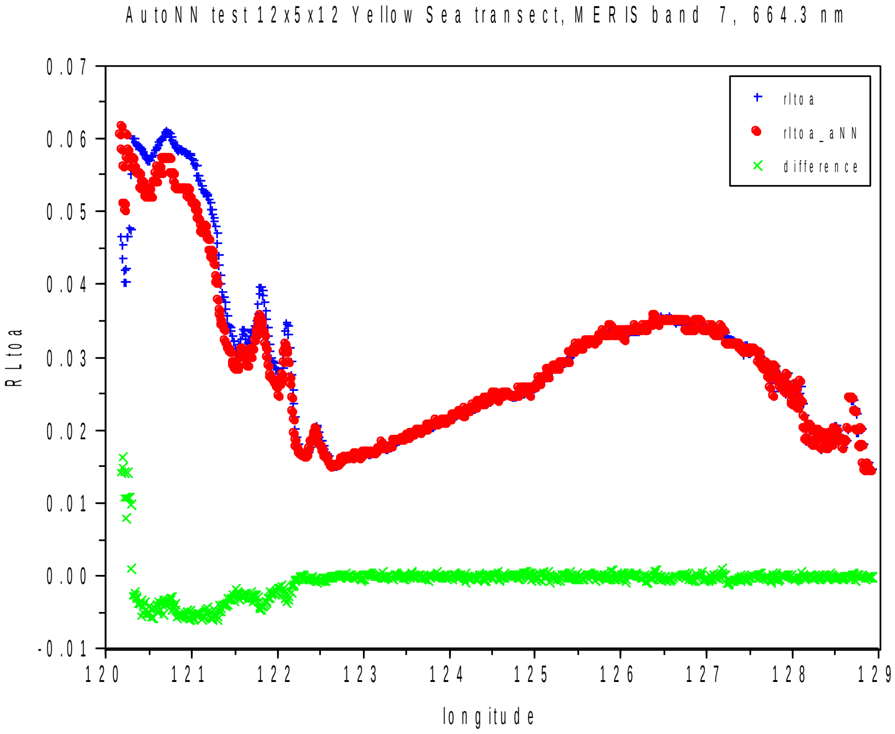

Fig. 18: TOA reflectance (blue) and output of the aaNN (red) with the difference (green) along the transect

6 Uncertainties

All ocean colour products have a certain range of uncertainty. It is partly due to the nature of a complex type of water within one or a tenth of a square kilometre of water surface, partly caused by the principle limitation due to the limited information content of a reflectance spectrum at top of atmosphere and partly caused by the limitation of the algorithm, which has only a limited scope. The uncertainty may change from pixel to pixel due to changes in the composition of the water constituents, due to changes in the relationship between IOPs and mass concentration, or due to other conditions, which are out of scope.

One major source are uncertainties in the atmospheric correction, i.e. the L2 water reflectances. But also the in situ data, which are used as the benchmark to check the products, have their own uncertainty range or are not directly comparable with the remote sensing product, because of the size of the water volume, time difference or method to determine the quantity. One example for the latter aspect is the method to separate the different fractions from the water sample, such as the absorption by phytoplankton pigments, which depends on the bleaching procedure, or by the definition of the dissolved fraction, which depends on the pore size of the filter.

Altogether, an uncertainty in the range of a factor of 2, when compared to in situ observations, can easily occur and is not necessarily an error.

For any application these uncertainties have to be quantified. This is not an easy task, because some of the causes, which are mentioned above, are known in principle but not actual for the pixel or area. Thus, here are some hints about how to estimate uncertainties:

- Look for possible sources of uncertainties:
    - When compared to in situ data the uncertainties of the in situ data have to be taken into account.
    - Determine the spatial and temporal variability of the relevant data of your area to assess the deviations between in situ and remote sensing data.
    - Determine the relationship between concentrations in different depths and the surface concentration.
    - Determine the optical properties of water constituents and compare with the bio-optical model, which was used for the algorithms (s. ATBD).
    - Determine if any of the water constituents is dominating the water reflectance so that the contribution of other substances, which shall be derived from the spectrum, is minor.
    - Check if at very high concentrations the reflectances in the blue – green spectral range are close to saturation.

For the quantification of uncertainties different approaches are possible:

- It can be based on the analysis of the relationship between algorithm output and in situ reference data
    - as the mean uncertainty of the whole scene or even time series of scenes, in particular for case 1 water,

  - as the mean uncertainty per concentration interval: for this purpose the full concentration or IOP range has to be divided in useful intervals and the uncertainty determined per interval,
  - as a mean uncertainty for each optical water type: then the area has to be classified and the uncertainty has to be determined for each type if sufficient in situ data per type are available,
  - if the uncertainty has been determined per concentration interval or water type, then the product has to be analysed accordingly and the corresponding uncertainty has to be transferred pixel by pixel using a look-up table,
- Alternatively, if no or not sufficient in situ data are available, the uncertainty can be estimated by computation pixel by pixel for each of the products
  - from second partial derivatives using a forward model: in this case the sensitivity of each component with respect to the reflectance spectrum for the actual concentrations of a pixel is determined,
  - for the semi-analytical decomposition algorithm (s. QAA algorithm) the uncertainty can be computed using a standard error propagation procedure,
  - or by the ensemble method: compare results from different algorithms and / or sensors.

7 Testing of algorithms

Before algorithms are applied to remote sensing data, a number of tests can be performed to check if the algorithm performs with the expected quality. It is also useful to participate in inter-comparison (“round robin”) exercises, which are organised in the context of various projects for selecting algorithms, s. Coastcolour intercomparison.

7.1 Test with simulated data

Simulated reflectances based on a full bi-directional radiative transfer model such as Hydrolight (Mobley, 1994) or Monte Carlo photon tracing model can be used to generate water leaving radiance reflectances for testing the algorithm. In these cases all parameters including the IOPs of water constituents are under control and the expected results are known. Furthermore various sensitivity studies can be performed:

- influence of changes of the spectral IOPs,
- effect of inhomogeneous vertical distribution,
- saturation effect,
- effect of the dominance of one of the water constituents,
- effect of solar and observation angles for the same concentrations, etc.

Also the sensitivity of the algorithm due to errors in the reflectance spectra can be tested by loading known errors on the reflectance spectra (single bands or the whole spectrum).

All tests can be performed with systematic or random changes. To identify cases of deviations in a large simulated data set, it is convenient to plot all cases in linear order for visual inspection or to filter the data with deviations above a certain threshold.

These types of tests can also be expanded to TOA reflectances in the same way if a model for the atmosphere is available.

Fig. 19 shows a test of a neural network with reflectances, which were simulated with the same bio-optical model as the training data. In the turbid water (left) chlorophyll can only be retrieved with sufficient accuracy down to a concentration of about 2 mg m^{-3}, while in the water with less suspended matter (right) the range of chlorophyll includes concentrations down to 0.1 mg m^{-3}.

The reason for this changing uncertainties due to the presence of suspended matter is visible in Fig. 20. At a SPM concentration of 100 g m^{-3} the impact of a change in chlorophyll concentration from 5 to 10 mg m^{-3} has only a negligible effect on the reflectance spectrum in the blue-green spectral range. Only the red / NIR range can be used her to detect changes in chlorophyll concentration.

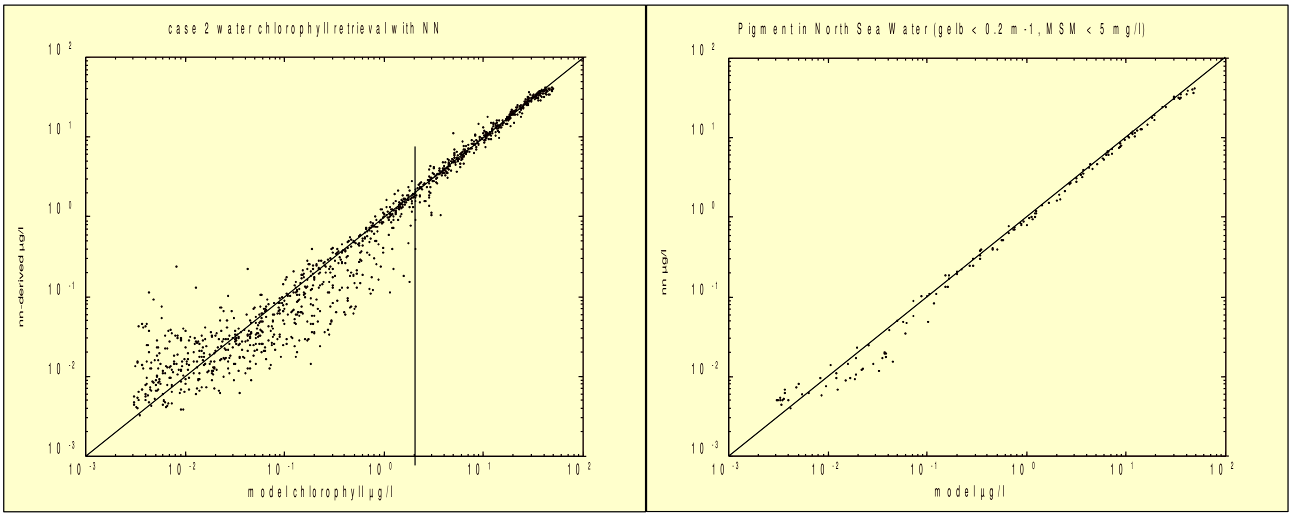

Fig. 19: Test of a NN for the retrieval of the chlorophyll concentration with simulated data: left with the full range of suspended matter and gelbstoff concentrations, right with limited gelbstoff absorption of a442 < 0.2 m^{-1} *and SPM < 5 g* m^{-3}*. The vertical line on the left side marks the upper concentration range of chlorophyll > 2 mg* m^{-3}*, which can be derived in turbid water, while in the clearer water the useful concentration range goes down to < 0.1 mg* m^{-3}*.*

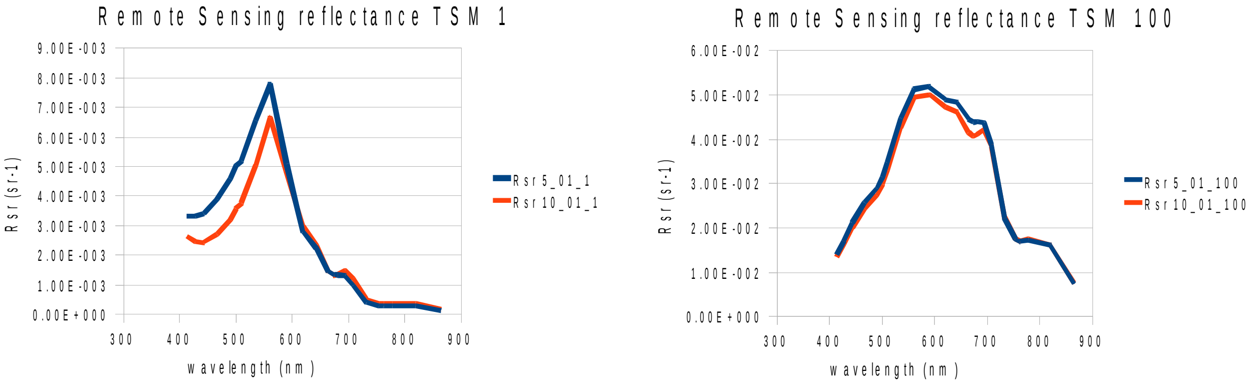

Fig. 20: Simulated water leaving radiance reflectances for two chlorophyll concentrations of 5 (blue) and 10 (red) mg m-3 : left in the presence of 1 g m^{-3} *SPM and right of 100 g* m^{-3} *SPM, the gelbstoff concentration is* a_{ys}*(443 nm) of 0.1* m^{-1}.

7.2 Tests with measured data

Several test data sets are available, which include simultaneously sampled concentrations of water constituents, IOPs, water reflectance spectra and, in the case of match-up data sets, also the corresponding TOA reflectances of satellite sensors. Algorithms for determining IOPs or concentrations can also be tested with in sit reflectance measurements (Fig. 21 right, in this case a neural network was used as the forward model iteratively with a Levenberg-Marquard optimization procedure).

Most popular and well maintained and quality tested are the MERMAID data set of ESA/ACRI and the NOMAD and SeaBASS data set of NASA Goddard. Note: to publish results using these data set the authorship of the data have to be respected according to the rules of the corresponding archive. Furthermore, these data sets are not error free, are collected under various conditions and have uncertainties (Fig. 21, left), which have to be taken into account. Also the limitation in the geographical distribution, optical types of waters, and solar and observation angles have to be considered.

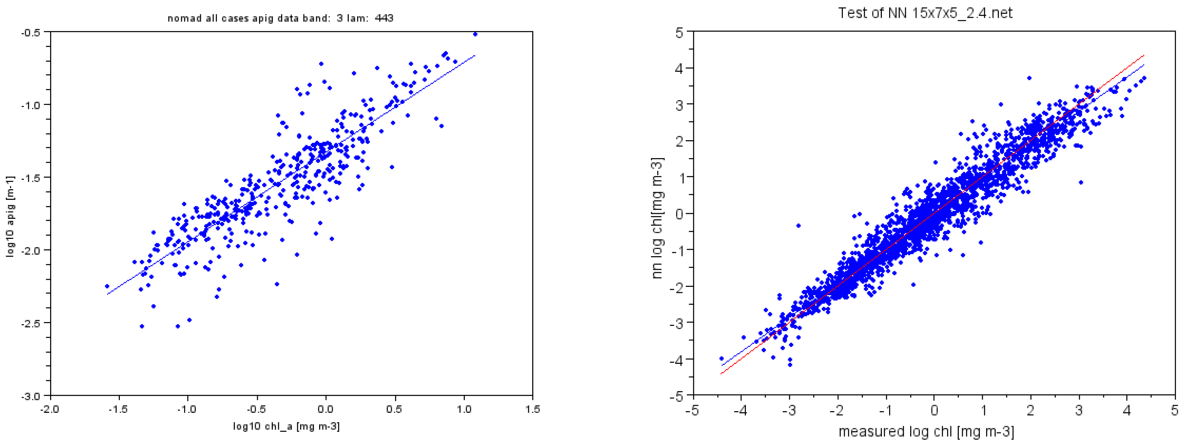

Fig. 21: Test with NOMAD data. Left: relationship between chlorophyll concentration (log10) and absorption of phytoplankton pigments a_{pig}*(443 nm) from water samples. Right: relationship between the chlorophyll concentration (logn) determined from samples (x-axis) and with the neural network from the corresponding water reflectance spectra (y-axis).*

7.3 Tests with only satellite data

Some useful tests of the algorithms can be performed without in situ data.

- Check of reflectance spectra: do they look “reasonable”, i.e. do the spectra of clear ocean water or turbid coastal water have the typical shapes? Check for negative reflectances.
- Check for artefacts: are there any strange structures in the water, which could be caused by the atmospheric correction, like structures in the vicinity of clouds or cloud shadows, or structures, which seem to cross the coast.
- Check transects: Do transects from the coast to the open sea show the expected gradient in concentration?
- Compare data of consecutive days: in mid latitudes part of the image is observed at consecutive days but under different angles; do they show similar results under stable conditions?
- Check annual time series: do they show the expected seasonal changes in chlorophyll concentration; are there any effects visible, which may depend on the solar angle?
- Compare results from alternative algorithms and from other missions: are there any deviations, do they dependent on water type, observation angle, atmospheric conditions?

It is also convenient to focus on one or more diagnostic sites or transects of which the concentrations and optical properties as well as conditions of the atmosphere are monitored (e.g. by an AERONET sun photometer station).

8 Validation of products

The validation of products is a permanent effort, because conditions and optical properties or the composition of water constituents may change with season or inter-annually. Some of the methods, which have been described already in the chapters above, apply also to validation. Important is to respect available flags and not to include flagged data in the quality assessment. Also check for artefacts and doubtful pixels, which are not flagged, such as cloud shadows and pixels at the rim of clouds.

Here are some more advices:

If no match up data are available for validation, compare the product with existing knowledge, e.g. the range of concentration, seasonal dependence, gradient from coast to the open sea, concentration in estuaries.

Check for research and monitoring programs, which sample data occasionally or on a regular basis: ferrybox programs, monitoring by environmental agencies, research projects, automatic stations. When comparing these data with the RS products it is necessary to look for details of the data and the protocols of the methods applied. Ferryboxes e.g. use an inlet for the continuous water stream, which can be located a few meters below the surface. Chlorophyll is determined in most cases with a continuous flow fluorometer, the results of which may differ from the analysis of samples per HPLC. If transects are compared, only a few data point may satisfy the match up criteria of ± 30 minutes around satellite overpass, while other data points my come from water bodies, which may have drifted away from the position when the sample was taken due to tidal currents.

If own measurements are planned for validation it is important to set set up a strategy to get as many match ups as possible. Since research vessels are extremely expensive to run and are available only for a short period, it may be better to look for ships of opportunity. These can be already planned research cruises, which would accept a small validation team. It can be regular monitoring cruises or helicopter flights by monitoring agencies, or ferry boats. Ferry boats have the advantage that they run daily in most cases so that one can join them on short notice under appropriate weather conditions. Disadvantage is that they do not stop: water samples can be taken only from the surface and no instruments can be put into the water. Above water reflectance measurements have to be taken from the bow of the ship, which is the only position without foam from the bow wash.

Other opportunities for validation are provided by oil rigs, platforms for research and wind farms, light houses, research buoys, poles, and jetties.

Measurements should be performed and samples processed according to standard protocols for match up analysis. These are provided by the validation teams of the respective space agencies. It is necessary to follow the definitions, which were used for the bio-optical models of the algorithms. Chlorophyll concentration can be defined for fluorometric or HPLC analysis and with or without degradation products. In some coastal waters the concentration of degradation products of phytoplankton pigments can even surmount the concentration of “living” chlorophyll, although the absorption spectrum and the effect on the reflectance spectrum might be very similar and not distinguishable by remote sensing.

Since the validation is only valid for the season and the conditions during sampling of match up data, it is important to sample all seasons of interest.

9 Summary, Conclusions and outlook

Optical remote sensing of coastal waters is a difficult task. The success depends on the type of water and the corresponding algorithms for the atmospheric correction, for the retrieval of IOPs and the concentrations of water constituents or for the determination of the spectral irradiance attenuation coefficient k_d.

While for clear coastal waters of case 1 water type simple reflectance band ratio algorithms are sufficient to determine the chlorophyll concentration, other types of water with bottom reflection and a manifold of different scattering and absorbing substances require complex inversion procedures and can impose problems in the retrieval and large uncertainties.

Due to the large variability of coastal waters there is no unique ideal type of algorithm. Thus, it makes sense to generate a portfolio of different algorithms, in which each of it has a dedicated area of application.

For Coastcolour we have selected and applied (1) an inverse neural network for the atmospheric correction, which can handle also turbid and absorbing types of water up to very extreme cases, (2) for the retrieval of IOPs a neural network and the quasi-semi-analytical algorithm (QAA) and (3) for floating cyanobacteria the MCI algorithm. But there are also other types of procedures, which are appropriate, such as the linear matrix inversion, the reflectance fitting algorithms using a forward reflectance model and an optimization techniques, and table look-up procedures.

For all model based approaches the definition of the bio-optical model is the most critical task. It defines the number and nature of optical components, its spectral optical properties and ranges.

For simulated data, which are used for the training of neural networks or for setting up look-up tables, also the frequency distribution and the co-variances of the components are important.

Since all algorithms have limitations it is necessary to complement the algorithms with co-algorithms for determining spectra, which are out of scope, and for determining the uncertainties of the products.

Limitations and large uncertainties of coastal water remote sensing products have to be expected, when the component of interest is dominated by the other components in the water, so that its effect on the reflectance spectrum is small or even negligible. Also at very high concentrations the reflectance approaches a constant level (“saturation effect”) so that further changes in the concentration cause only a very small change of the reflectance spectrum.

Although the full resolution products of MERIS have provided important information about the distribution and dynamics of phytoplankton, suspended matter and organic compounds in many coastal waters, it is always necessary to check the products for any artefacts, carefully respect the flags and in case of failures or large uncertainties test alternative algorithms.

What did we learn from Coastcolour with respect to algorithms for future coastal water missions?

- The processor for coastal waters should include a suite of algorithms, each of which has been adapted to and optimized for different optical types of coastal waters.
- As a first step the optical type should be identified from the top of atmosphere reflectance so that also the type of atmospheric correction can be selected including procedures, which can

be performed without an atmospheric correction such as the FLI or MCI.

- It is important to detect conditions and respective spectra which are out of scope of the algorithms.
- Since the uncertainty of coastal water products can be extremely variable and can change from pixel to pixel it is necessary to complement the processor with procedures, which allow the determination of the degree of uncertainty on a pixel by pixel basis and provide this information for each product.

10 References

D'Alimonte, D. Zibordi, G.: Phytoplankton determination in an optically complex coastal region using a multilayer perceptron neural network, Geoscience and Remote Sensing, IEEE Transactions on (Volume:41 , Issue: 12) 2003

Doerffer R, Fischer J (1994) Concentrations of chlorophyll, suspended matter, and gelbstoff in case II waters de-rived from satellite coastal zone color scanner data with inverse modelling methods. Journal Geophysical Research 99:7457-7466

Doerffer R. and Schiller, H. (2007). The MERIS Case 2 water algorithm, , International Journal of Remote Sensing, 28, (3-4): 517-535.

Doerffer, R. (1979) Untersuchungen über die Verteilung oberflächennaher Substanzen im Elbe-Aestuar mit Hilfe von Fernmeßverfahren. Arch Hydrobiol Suppl. 43 Elbe-Aestuar 4 2/3:119-224

Doerffer, R. and Murphy, D.: Factor analysis and classification of remotely sensed data for monitoring tidal flats, Helgoländer Meeresuntersuchungen, 1989, Volume 43, Issue 3-4, pp 275-293

Doerffer, R., Fischer, J., Stössel, M., Brockmann, C., Graßl, H.: Analysis of thematic mapper data for studying the suspended matter distribution in the coastal area of the German Bight (North Sea), Remote Sensing of Environment, Volume 28, 1989, Pages 61–73

Doerffer, R., Schiller, H.: The MERIS Neural Network Algorithm, IOCCG Report Number 5, 2006, 43-47

Fischer J, Doerffer R, Graßl H (1986) Factor analysis of multispectral radiances over coastal and open water based on radiative transfer calculations. Applied Optics 25:448-456

Gordon, H.R., and Morel, A. (1983). Remote assessment of ocean color for interpretation of satellite visible imagery: A review. Springer-Verlag, New York.

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

Gower, J. F. R. & S. A. King, 2011: Distribution of floating Sargassum in the Gulf of Mexico and the Atlantic Ocean mapped using MERIS, International Journal of Remote Sensing, 32:7, 1917-1929

Gower, J.F.R., Doerffer, R. and Borstad, G.A., 1999, Interpretation of the 685 nm peak in water-leaving radiance spectra in terms of fluorescence, absorption and scattering, and its observation by MERIS. International Journal of Remote Sensing, 9, pp. 1771–1786.

Hochberg, E.J., Atkinson, M.J.: Spectral discrimination of coral reef benthic , Coral Reefs, 2000, 19, 164-171

Hoge, F.E., and Lyon, P.E. (1996). Satellite retrieval of inherent optical properties by linear matrix inversion of oceanic radiance models: an analysis of model and radiance measurement errors. J. Geophys. Res. 101: 16631-16648.

Hoge, F.E., and Lyon, P.E. (1999). Spectral parameters of inherent optical property models: Methods for satellite retrieval by matrix inversion of an oceanic radiance model. Appl. Opt. 38: 1657-1662.

Joseph, J.: Untersuchungen über Ober- und Unterlichtmessungen im Meere und über ihren Zusammenhang mit Durchsichtigkeitsmessungen Deutsche Hydrographische Zeitschrift (1950) 3: 324-335 , September 01, 1950

Krawczyk, H., Neumann, A., Gerasch, B. and Walzel, T. (2007) 'Regional products for the Baltic Sea using MERIS data', International Journal of Remote Sensing, 28:3, 593 – 608

Lee, Z. P., Carder, K., Arnone, R.: The Quasi-Analytical Algorithm, IOCCG Report Number 5, 2006, 73-80

Lee, Z. P., K. L. Carder, C. D. Mobley, R. G. Steward, and J. S. Patch, 1999. Hyperspectral remote sensing for shallow waters: 2. Deriving depths and optical properties by optimization. Applied Optics, 38(18), 3831-3843

Lee, Z.P. (ed.) Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications, IOCCG Report Number 5, 2006, 126 pp.

Lee, Z.P., Carder, K.L., and Arnone, R. (2002). Deriving inherent optical properties from water color: A multi-band quasi-analytical algorithm for optically deep waters. Appl. Opt. 41: 5755-5772.

Lee, Z.P., Carder, K.L., Mobley, C.D., Steward, R.G., and Patch, J.S. (1998a). Hyperspectral remote sensing for shallow waters. 1. A semianalytical model. Appl. Opt. 37: 6329-6338.

Lee, Z.P., Carder, K.L., Mobley, C.D., Steward, R.G., and Patch, J.S. (1999). Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. Appl. Opt. 38: 3831-3843.

Lee, Z.P., Carder, K.L., Steward, R.G., Peacock, T.G., Davis, C.O., and Patch, J.S. (1998b). An empirical algorithm for light absorption by ocean water based on color. J. Geophys. Res. 103: 27967-27978.

Lesser, M. P. and C. D. Mobley, 2007. Bathymetry, water optical properties, and benthic classification of coral reefs using hyperspectral remote sensing imagery. Coral Reefs, 26, 819-829

Lyon, P. & Hoge, F. (2006): The Linear Matrix Inversion Algorithm, chapter 7 in IOCCG report No. 5, 49-56

Maritorena, S., Morel, A., Gentili, B. (1994): Diffuse reflectance of oceanic shallow waters: influence of wagter depth and bottom albedo, Limn. Ocean., 39, 1689-1703

Mobley, C.: , 1994: Light and water: radiative transfer in natural waters, Cademic Press, San Diego, 592 pp.

Morel, A. (1974). Optical properties of pure water and pure sea water. In: Optical Aspects of Oceanography. Jerlov, N.G., and Nielsen, E. S. (eds). Academic Press, New York, p. 1-24.

Morel, A., and L. Prieur, 1977: Analysis of variations in ocean color. Limnol. Oceanogr., 22(4), 709-722.

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

Morozov, E., Korosov, A., Pozdnyakov, D., Pettersson, L., Sychev, V.: A new area-specific bio-optical algorithm for the Bay of Biscay and assessment of its potential for SeaWiFS and MODIS/Aqua data merging, International Journal of Remote Sensing, 2010; 31(24):6541-6565.

Mumby, P.J., Green, E.P., Clark, C.D., Edwards, A.J. (1998): Digital analysis of mulstispectral airborne imagery of coral reefs, Coral Reefs, 17, 59-69

Neville, R. A. and Gower, J.F.R. 1977: Passive remote sensing of phytoplankton via chlorophyll a fluorescence, Journal of Geophysical Research, 82, 3487-3493

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel,D. A., Carder, K.L., Garver, S.A., Kahru, M., McClain, C.: Ocean color chlorophyll algorithms for SeaWiFS, Journal of Geophysical Research: Oceans (1978–2012), Volume 103, Issue C11, pages 24937–24953, 1998

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

Santer, R., Schmechtig, C.: Adjacency Effects on Water Surfaces: Primary Scattering Approximation and Sensitivity Study, Applied Optics, Vol. 39, Issue 3, pp. 361-375 (2000)

Sathyendranath, S. (ed.) Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters, IOCCG Report Number 3, 2000, 139 pp.

Schiller H, Doerffer R (1993) Fast computational scheme for inverse modelling of multi-spectral radiances: application for remote sensing of the ocean. Applied Optics 32:3280-3285

Schiller, Helmut & Roland Doerffer (1999) Neural network for emulation of an inverse model operational derivation of Case II water properties, International Journal of Remote Sensing, 9, 1735-1746

Smith, R.C., and Baker, K.S. (1981). Optical properties of the clearest natural waters. Appl. Opt. 20: 177-184.

Sterckx, S., Knaepsa, E. & Ruddick, K: Detection and correction of adjacency effects in hyperspectral airborne data of coastal and inland waters: the use of the near infrared similarity spectrum, International Journal of Remote Sensing, Volume 32, Issue 21, 2011, pp 6479-6505

11 Symbols and acronyms

Acronym	Explanation
a	Absorption coefficient
AC	Atmospheric correction
algal_1	L2 product of MERIS, expressed in units of phytoplankton chlorophyll [mg m^{-3}]
ATBD	Algorithm theoretical basis document
bb	Backscattering coefficient
BEAM	
Case 1 / 2	Optical water type, the variability of which can be described with one component, mainly phytoplankton pigment, while for case 2 water > 1 components are necessary.
CDOM	Coloured dissolved organic matter, or chromophoric dissolved organic matter
DOC / POC	Dissolved organic carbon, particulate organic carbon
Ed	Downwelling irradiance
FLH	Fluorescence line height
HAB	Harmful algal bloom
HPLC	High performance liquid chromatography
IOCCG	International Ocean Colour Coordination Group
IOP	Inherent optical property
kd	Attenuation coefficient of the downwelling irradiance
L, Lw	Radiance, water-leaving radiance
L1b	Level 1b data, data of an instrument, which is corrected and calibrated in the original physical units and the first level, which is provided to the user
L2	Level 2 data, derived data from L1b, in the case of ocean colour after the atmospheric correction
MCI	Maximum chlorophyll index
MERIS	Medium resolution imaging spectrometer, on board ENVISAT, 2002-2012
NIR	Near infrared
NN, aaNN	Neural network, auto-associative NN
OC4	Ocean colour algorithm version 4, used for SeaWiFS and MODIS

Acronym	Explanation
PCA	Principle component analysis
QAA	Quasi-Analytical Algorithm
RMS	Root mean square deviation
Rrs	Remote sensing reflectance, is the water-leaving radiance reflectance
Secchi disc depth	Depth at which a white disc dissappear in the ocean, named after the sientist Antonio Secchi, who first applied this instrument
SPM	Suspended particulate matter
TOA	Top of Atmosphere
z90max	Signal depth, depth from which 90% of the reflected sun light at the surface of the ocean stems from