CSIRO

SAMBUCA

Semi-Analytical Model for Bathymetry, Un-mixing, and Concentration Assessment

Magnus Wettle and Vittorio Ernesto Brando

CSIRO Land and Water Science Report 22/06

July 2006

Important Disclaimer:

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

ISSN: 1883-4563

SAMBUCA

Semi-Analytical Model for Bathymetry, Un-mixing, and Concentration Assessment

Magnus Wettle and Vittorio E. Brando

Environmental Remote Sensing Group

CSIRO Land and Water

CSIRO Land and Water Science Report 22/06

July 2006

Acknowledgements

We thank The University of Hull for partially funding the initial stages of SAMBUCA development, and Arnold Dekker of CSIRO Land and Water for contributing to the conceptual development of SAMBUCA.

Executive Summary

SAMBUCA is optimization-driven analytical model for estimating water column constituent concentrations, water column depth, and benthic substrate composition from remote sensing data. Based on an approach conceptualized by Lee et al. (1998, 1999, 2001), SAMBUCA was implemented in an IDL/ENVI environment. Following extensive testing and application of the approach to different data sets, several important modifications and improvements were made to the original Lee et al. (1998, 1999, 2001) approach. Notably, the optical parameterization of both the water column and the benthic substratum – fundamental properties of the model – were re-defined. Furthermore, features such as accessing spectral libraries of optical properties, allowing for more than one substratum to contribute to the remote sensing signal, and tracking the optical depth of a given system, were added. Comparison of output from SAMBUCA and the industry standard HYDROLIGHT model yield satisfactory results. The aim of this technical report is to give a comprehensive description of SAMBUCA.

Table of Contents

1. Introduction

1.1. Background

The bulk of past aquatic remote sensing work, particularly in optically shallow applications, has been image-based. Examples include exploring statistics within the imagery, developing band ratios and indices in an effort to circumvent the difficulties posed by bathymetry and water colour, and incorporating site specific data into supervised classifications. These approaches have the advantage of often being relatively straightforward to apply, but offer several limitations including that they tend to be site specific, sensor specific, and/or time specific (Kutser et al. 2003).

An alternative to these approaches is a physics-based, radiative transfer method. Advantages to such methods include: the application of one algorithm to time series of remote sensing data, the ability to simulate the relevance of various sensors to a particular application, and a reduction in the amount of field and laboratory-based measurements (Dekker et al., 2001). A drawback to the physics-based approach is that it is relatively complex in its application.

In the case of coral remote sensing from satellite sensors, a physics-based method would allow for temporal and spatial variations in the reef water optical properties and be able to cope with spatially varying bathymetry. Such a method would significantly reduce the requirement for field or aerial survey approaches for mapping bleaching.

Recent reviews on the status of mapping coral reefs demonstrate the need for a mapping procedure able to work across a variety of depths, water column types and substrates. Several special journal issues on this subject (e.g. Limnology and Oceanography (2003); Coral Reefs (2004)) cover a variety of mapping approaches, all highlighting problems of lack of repeatable techniques in variable water depth, colour and substrate type. Previous attempts to solve this problem have extracted depth, water column constituents or substrate, but not all three (Goodman and Ustin, 2003; Hedley and Mumby, 2003).

1.2. Introducing SAMBUCA

To this end, SAMBUCA (Semi-Analytical Model for Bathymetry, Un-mixing, and Concentration Assessment) was developed. It is based on the approach conceptualised by Lee et al. (1998, 1999, 2001) for retrieving environmental variables from remote sensing data using an analytical model and optimisation routine. This approach was chosen based on the need for a physics-based analytical solution for retrieving environmental variables on a pixel-by-pixel basis.

The advantages and justifications of this approach have already been introduced. The analytical nature of the model offers a relatively fast solution of the underwater radiative transfer equation. This is of particular advantage when the data set is comprised of hyperspectral remote sensing imagery.

The physics-based nature of the model offers an objective and repeatable algorithm for extracting environmental information from remote sensing data. The majority of coral reef applications of remote sensing in the literature have been empirical in nature, and/or required a degree of user-input and interpretation. Such approaches are inherently flawed for applications such as monitoring change; an objective and repeatable methodology is clearly preferable.

A physics-based method significantly reduces the need for field work. In the case of SAMBUCA, once the model is parameterised with field data, additional fieldwork enhances the application but is not necessary.

SAMBUCA can be applied to any digital spectral sensor data with little extra effort. This limited need of field data and portability across data sets means for example that remote sensing data archives can be explored. Such access to baseline environmental conditions can be important in change monitoring applications.

The Lee et al. (1998, 1999, 2001) approach was implemented in IDL for evaluation purposes. From this, SAMBUCA was developed through a series of modifications. The evolution of SAMBUCA was guided both by the need to align the model parameterisation with our field data protocols as well as to provide enhanced functionality. An example of the latter is the ability of SAMBUCA to perform sub-pixel un-mixing given a spectral library of substrate reflectance data. The following sections present the details of the semi-analytical model, including its modifications, functionality, and implementation.

2. The model

The essence of the approach lies in expressing the subsurface remote sensing reflectance R_{rs} as a function of a set of relevant variables. This modelled remote sensing reflectance, $R_{rs}^{modelled}$, is then compared to the measured remote sensing reflectance, $R_{rs}^{measured}$. The set of variables that minimises the difference between these two spectra is retained as the result of the minimisation. These variables are then used to estimate the environmental variables being sought.

At the core of the model lies an analytical expression for subsurface remote sensing reflectance, R_{rs}, which expresses R_{rs} as a combination of the subsurface remote sensing reflectance due to the water column scattering R_{rs}^{C} and the subsurface reflectance due to the bottom reflectance R_{rs}^{B} (in an optically shallow system) (Equation 1).

$$R_{rs} = R_{rs}{}^{C} + R_{rs}{}^{B} \qquad (1)$$

This is the same general form as that proposed by Maritorena et al. (1994). The analytical approximation for Equation 1 is written as

$$R_{rs}{}^{\text{modelled}} = R_{rs}{}^{DW}\left(1-e\right)^{-\left(\frac{1}{\cos\theta_{subsun}}+\frac{D_{UC}}{\cos\theta_{subview}}\right)\kappa H} + \frac{1}{\pi}\rho(\lambda)e^{-\left(\frac{1}{\cos\theta_{subsun}}+\frac{D_{UB}}{\cos\theta_{subview}}\right)\kappa H} \qquad (2)$$

where R_{rs}^{DW} is the remote sensing reflectance associated with an infinitely deep water column, and D_{UC} and D_{UB} are the so-called optical elongation factors for the signal originating from the water column and bottom substrate, respectively. $\cos\theta_{subsun}$ is the subsurface solar zenith angle, and $\cos\theta_{subview}$ is the subsurface viewing angle. ρ denotes the irradiance reflectance of the bottom substrate, and κ is the attenuation coefficient, defined as

$$\kappa = a + b_b \qquad (3)$$

In Lee et al. (1998, 1999, 2001), the remote sensing reflectance from the infinitely deep water column is defined by

$$R_{rs}^{DW} = (0.084 + 0.17u)u \tag{4}$$

and the optical elongation factors for scattered photons from the water column and the bottom substrate are defined according to

$$D_{UC} = 1.03(1 + 2.4u)^{0.5} \tag{5}$$

$$D_{UB} = 1.04(1 + 5.4u)^{0.5} \tag{6}$$

where

$$u = \frac{b_b}{(a + b_b)} \tag{7}$$

Finally, the inherent optical properties, the absorption and backscattering coefficients, are defined by

$$a(\lambda) = a_w(\lambda) + a_\varphi(\lambda) + a_g(\lambda) \tag{8}$$

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bp}(\lambda) \tag{9}$$

Here, the total absorption coefficient is considered as the sum of the absorption coefficient of water (a_w), phytoplankton (a_φ), and gelbstoff (a_g). The total backscattering coefficient is the sum of the backscattering coefficient for pure water (b_{bw}) and the backscattering coefficient of suspended particles (b_{bp}) in the water column.

In summary, for a fixed set of viewing and illumination angles, the subsurface remote sensing reflectance has been expressed as a function of the absorption coefficient a, the backscattering coefficient b_b, bottom substrate albedo ρ and depth H (Equation 10. (It should be kept in mind that apart from depth, these entities are spectral in nature, and not scalar variables.)

$$R_{rs}^{\text{modelled}} = f(a(\lambda), b_b(\lambda), \rho(\lambda), H) \tag{10}$$

The next section will set the two coefficients a and b_b as functions of optically active constituents in the water column and describe ρ in terms of the optical properties of the bottom substrate, thereby linking the remote sensing signal to environmental variables.

3. Parameterisation of the model

3.1.1. Parameterisation of total absorption

In the model proposed by Lee et al. (1998, 1999, 2001), the total absorption coefficient is written according to Equation 8. This does not explicitly take into account absorption due to in-organic particles in the water, and the definition for gelbstoff, which relates to a_g, according to Lee et al. (1998, 1999, 2001), is not clear. Therefore, for consistency with the nomenclature and the formulation adopted in Brando & Dekker (2003), the expression for the total absorption coefficient has been re-written as

$$a_{TOT} = a_{WATER} + a_{PHY} + a_{CDOM} + a_{TR} \tag{11}$$

where a_{CDOM} is the absorption coefficient due to CDOM (coloured dissolved organic matter), and a_{TR} is the absorption coefficient due to tripton (the non-algal particulate component in the water).

The absorption for pure water, a_{WATER} is taken from Pope and Fry (1997).

The absorption due to phytoplankton was defined by Lee et al. (1998, 1999, 2001) as

$$a_{\varphi}(\lambda) = \left(a_0(\lambda) + a_1(\lambda)\ln(P)\right)P \tag{12}$$

where a_0 and a_1 were determined through regressions using extensive field data (Lee, 1994), and P is a single variable (note that a_0 and a_1 are spectra). This allows for the shape and magnitude of a_{φ} to be controlled by just varying P.

However, a_0 and a_1 were determined based on field data from the Florida Keys, and may not apply globally. Furthermore, it could be argued that the variable P does not directly express a useful measure of an environmental variable.

Instead, SAMBUCA defines the absorption due to phytoplankton as

$$a_{PHY} = C_{CHL} \cdot a^*_{PHY} \tag{13}$$

Here, C_{CHL} is the concentration of chlorophyll, and a^*_{PHY} is the specific absorption of phytoplankton (the absorption of pigments normalised to chlorophyll concentration). The latter can be thought of as the wavelength-specific shape of the pigment absorption, which is then scaled according C_{CHL}. Given a^*_{PHY} the total absorption coefficient due to phytoplankton in the water column is now controlled by the single variable C_{CHL} instead.

Both Equation 12 and Equation 13 describe the same general behaviour: a wavelength-dependant shape multiplied by a scalar. However, Equation 13 enables the direct retrieval of chlorophyll concentration, as well as taking advantage of the availability of field data for estimation of a^*_{PHY}.

The absorption not due to pure water or phytoplankton was attributed to gelbstoff by Lee et al. (1998, 1999, 2001), and expressed as

$$a_g(\lambda) = Ge^{(-S(\lambda-\lambda_0))} \tag{14}$$

This is an exponential function, whose slope is wavelength-dependent and is controlled by the value of S. This spectral shape is multiplied by the variable G. As with Equation 12, G is the absorption due to gelbstoff at the reference wavelength λ_0.

Within SAMBUCA, the absorption not due to pure water or phytoplankton is instead attributed to both CDOM (coloured dissolved organic matter) and tripton (the non-algal component of suspended solids). The absorption coefficient due to CDOM is written as

$$a_{\mathrm{CDOM}}(\lambda) = C_{\mathrm{CDOM}} \cdot a^*_{\mathrm{CDOM}}(\lambda_0)^{(-S_C(\lambda-\lambda_0))} \tag{15}$$

where C_{CDOM} is the concentration of CDOM defined such that $a^*_{\mathrm{CDOM}}(\lambda_0)$ is set equal to 1. S_C is the slope of the CDOM absorption; its value determines the shape of the exponential absorption curve. Values for S_C can be determined through laboratory analysis of field samples.

The absorption due to tripton is similarly written as

$$a_{\mathrm{TR}}(\lambda) = C_{TR} \cdot a^*_{\mathrm{TR}}(\lambda_0)^{(-S_{TR}(\lambda-\lambda_0))} \tag{16}$$

where C_{TR} is the concentration of tripton in the water column and $a^*_{\mathrm{TR}}(\lambda_0)$ and S_{TR} are again shape-determining and sample-dependent.

3.1.2. Parameterisation of total backscatter

Lee et al. (1998, 1999, 2001) define the total backscatter coefficient through Equation 9, with the backscatter coefficient of pure water b_{bW} taken from the literature.

The second term in Equation 9 is defined as

$$b_{bp}(\lambda) = X\left(\frac{\lambda_0}{\lambda}\right)^Y \tag{17}$$

where the shape of the backscattering curve is defined by the slope Y, and multiplied by the variable X. As with the above parameterisations of the original Lee et al. (1998, 1999, 2001) model, the scalar variable X is the value of the backscattering coefficient at a reference wavelength λ_0.

The particulate backscattering term has been expanded into two terms within SAMBUCA; a backscattering coefficient for phytoplankton particles b_{bPHY} and a backscattering coefficient for tripton b_{bTR}, for consistency with the nomenclature and the formulation adopted in Brando & Dekker (2003).

$$b_b(\lambda) = b_{bw}(\lambda) + b_{bPHY}(\lambda) + b_{bTR}(\lambda) \tag{18}$$

The two last terms in Equation 18 are defined by

$$b_{bPHY}(\lambda) = C_{PHY} * X_{PHY}\left(\frac{\lambda_0}{\lambda}\right)^Y \tag{19}$$

and

$$b_{bTR}(\lambda) = C_{TR} * X_{TR}\left(\frac{\lambda_0}{\lambda}\right)^Y \tag{20}$$

Again, the slope of these functions is described by the factor Y, and C_{PHY} and C_{TR} represent the concentrations of phytoplankton and tripton in the water, respectively. The original X term from the Lee et al. (1998, 1999, 2001) model relates to the SAMBUCA parameterisation through

$$X = C_{TR}X_{TR} + C_{PHY}X_{PHY} \tag{21}$$

3.1.3. Parameterisation of bottom albedo

In the original Lee et al. (1998, 1999, 2001) parameterisation, one substrate reflectance spectra was used to represent the optical properties of the bottom substrate through

$$\rho = B \cdot \rho_{\lambda 0} \tag{22}$$

Where $\rho_{\lambda 0}$ is the reflectance spectra of the bottom substrate normalised at wavelength λ_0 and, and B is a scalar. The variable B can be thought of as a 'denormalisation' factor that allows for the same substrate to have varying magnitudes (but the same shape) of reflectance. Given a reflectance spectra, the contribution from bottom substrate is therefore controlled by the single variable B.

This approach was applied to waters in Tampa Bay (Florida), where sand was the omnipresent bottom substrate. It was further developed to first make a rough assessment of bottom albedo, and to then use either a seagrass or a sand spectra for $\rho_{\lambda 0}$ (Lee et al., 2001).

Coral reef environments present higher spatial heterogeneity in bottom substrate composition. The bottom substrate parameterisation for SAMBUCA was evolved in order to account for this, through

$$\rho(\lambda) = q_1\rho_1 + q_2\rho_2 + ... + q_n\rho_n \qquad \text{where} \qquad q_1 + q_2 + ...q_n = 1 \tag{23}$$

where n is the number of substrate spectra within the pixel, and q_n is the proportion of substrate n within the pixel. Typically, no more than 3 substrates are allowed within each pixel (n = 3). In the case of two substrates - the most commonly used parameterisation for this work - Equation 23 can be re-written as

$$\rho(\lambda) = q_{ij}\rho_i + (1 - q_{ij})\rho_j \tag{24}$$

Given ρ_i and ρ_j, ρ is now governed by the single variable q_{ij}, which represents the proportion of substrate *i* to substrate *j*.

However, even if each pixel is allowed to contain e.g. two substrates, this does not solve the problem of taking into account the high diversity of bottom types encountered in a coral reef environment. Implementing Equation 24 with a library of bottom substrate reflectance spectra was seen as the solution. Hence, SAMBUCA cycles through all the possible combinations of e.g. a pair of spectra taken from a spectral library, retaining the substrates (ρ_i and ρ_j) and the proportions of each (q_{ij} and 1 – q_{ij}) that give the best solution. Note that the identification of which two substrates allowed for the best solution is implicit in the retrieval of q_{ij}. The limitation to this spectral library approach is the factorial increase in processing time with the number of spectra in the library.

3.1.4. The final parameterisation

The ability to express the wavelength dependant absorption of e.g. phytoplankton with a single variable P (Equation 12) is central to the minimisation approach developed by Lee et al. (1998, 1999, 2001). The same principle is seen in Equation 14, Equation 17, and Equation 23, where G, X, and B are scalar variables that govern the equations, respectively. This allows for the implementation of an optimisation routine, which will be discussed in further detail in the next section.

Given the appropriate spectra a_0 and a_1 for expressing phytoplankton absorption (Equation 12, setting the subsurface solar zenith angle and the subsurface viewing angle (Equation 2), and inserting Equation 12, Equation 14, Equation 17, and Equation 23 into Equation 2, the subsurface remote sensing reflectance can be constructed as a function of seven non-wavelength dependent variables:

$$R_{rs}(\lambda)^{\text{modelled}} = f\left(P, G, X, q1, H, S, Y\right) \tag{25}$$

where H represents depth (Equation 2). If the slope S of the function for the absorption of gelbstoff and the slope Y of the function for the backscattering coefficient for particulate matter (Equation 14 and Equation 17 respectively) are known, Equation 25 reduces to

$$R_{rs}(\lambda)^{\text{modelled}} = f\left(P, G, X, q1, H\right) \tag{26}$$

which is a common implementation of the Lee et al. (1998, 1999, 2001) approach.

In a similar manner, SAMBUCA expresses the modelled subsurface remote sensing reflectance as a function of simple variables. Inserting Equation 13, Equation 15, Equation 19, Equation 20, and Equation 24 into Equation 2, and fixing $\theta_{subview}$ and θ_{subsun} (Equation 2), R_{rs} can be expressed as a function of the following set of variables

$$R_{rs}(\lambda)^{\text{modelled}} = f\left(C_{CHL}, C_{CDOM}, C_{TR}, X_{CHL}, X_{TR}, \rho_i, \rho_j, q_{ij}, H, S_{CDOM}, S_{TR}, a^*_{\text{TR}}(\lambda_0), a^*_{\text{PHY}}, Y\right) \tag{27}$$

The SIOPs (specific inherent optical properties), or shape-determining factors, in the SAMBUCA parameterisation (Equation 27) can be determined through field work and laboratory analysis (S_{CDOM}, S_{TR}, a^*_{TR}, a^*_{PHY} and Y). Given a set of substrate reflectance spectra, Equation 25 can then be reduced to

$$R_{rs}(\lambda)^{\text{modelled}} = f\left(C_{CHL}, C_{CDOM}, C_{TR}, q_{ij}, H\right) \tag{28}$$

The subsurface remote sensing reflectance is now a function of the concentration of phytoplankton, CDOM, and tripton, as well as the proportional composition of e.g. two bottom substrates, and depth. In other words, the signal measured by a remote sensor is expressed as an analytical function of a set of relevant environmental variables.

4. Implementation of the optimisation routine

For a given set of values for e.g. P, G, X, B, and H, Lee et al. (1998, 1999, 2001) compute the difference between the measured (e.g. field or image data) subsurface remote sensing reflectance spectra $R_{rs}^{measured}$ and the modelled subsurface remote sensing reflectance spectra $R_{rs}^{modelled}$ (Equation 29).

$$\text{error} = f(R_{rs}^{measured}, R_{rs}^{modelled}) \tag{29}$$

Using an optimisation routine, this error is minimised by varying the - in this example - five variables. The lowest error achieved is retained, together with the values used for each variable. Thus, this technique can allow for the retrieval of a set of variables, which relate to environmental parameters, from a single $R_{rs}^{measured}$ measurement. Inherent in the application of this approach is the need to control the underlying functions with simple variables, which the preceding section on the parameterisation has outlined.

SAMBUCA uses a simplex algorithm, often referred to as the 'amoeba' algorithm, to perform the optimisation (Press et al., 2003). Boundary conditions are set for the variables being sought, effectively determining the parameter space being explored. These ranges have been found to aide the minimisation by confining the search to realistic values. Furthermore, suitable starting values for the variables being sought can have a marked effect on the success of the minimisation as this can help prevent the simplex algorithm from finding local solutions in local minima. The fixing of these ranges and starting points is done to the best extent possible, ideally based on local knowledge and/or data from the study area.

In theory, the retrieval of a set of variables from one measurement is possible provided the system is over-determined. In the case of remote sensing applications, this generally means data with more spectral bands than the number of variables being sought. However, reliability will increase with more spectral bands.

5. Applications of SAMBUCA

The complete parameterisation for SAMBUCA is given by Equation 27, where up to eleven variables determine the resulting $R_{rs}^{modelled}$. This theoretically requires data with at least twelve wavelength channels, although in practice several more (in the order of 3-5 more) have been found to be preferable. Allowing for a relatively large number of degrees of freedom increases the chance of incorrectly retrieving one or more variables; the effects of one parameter are more likely to be compensated for, or included in, one or more other variables. As a simple example, SAMBUCA may overestimate backscattering in the water column in order to compensate for a shallow, bright bottom substrate. It is therefore recommended to fix (or to limit the range of) as many of the variables in Equation 27 as possible, within the limitations and scope of the application. The choice of which parameters that are to be fixed is dictated by the type of application and the availability of field data.

SAMBUCA can be applied to hyperspectral image data, where each pixel in the image corresponds to one $R_{rs}^{measured}$ spectrum. For coral reef remote sensing work, this involves optically shallow systems. Therefore, the depth H and the bottom substrate parameter q1 need to be solved for. Assuming that the concentrations of optically active constituents in the water column exhibit horizontal gradients within the scene, C_{CHL}, C_{CDOM}, and C_{TR} need to be given the freedom to vary as well. The remaining variables in Equation 27 are the SIOPs, which are responsible for determining the shapes of the various absorption and backscattering coefficients. Provided that fieldwork data and/or knowledge of the site enable the fixing of these SIOPs, Equation 28 can be used as a parameterisation.

Alternatively, for an optically deep system, the depth can be set to infinite, obviating the need to define the optical properties of the bottom substrate (no substrate spectral library and no q1 variable to solve for). This can be of advantage if the SIOPs of the water need to be estimated. An example of such a parameterisation would be

$$R_{rs}^{modelled} = f(C_{PHY}, C_{CDOM}, C_{TR}, S_C, S_{TR}, Y) \quad (30)$$

Here, the slopes of the absorption and backscattering functions S_C, S_{TR} and Y are sought, as well the three concentrations for chlorophyll, CDOM, and tripton.

For an optically shallow system where the SIOPs are not known, one approach could be to first apply Equation 30 parameterisation to nearby optically deep waters and retrieve a set of SIOPs. Assuming that these SIOPs do not differ too much with the optically shallow areas, SAMBUCA can then be run with the Equation 28 parameterisation using the SIOPs previously determined. Although certain assumptions are necessary, this limits the amount of degrees of freedom during each of the SAMBUCA runs.

SAMBUCA can also be used in simulation mode, applicable to for example sensitivity studies. The model can be run in 'forward' mode in order to generate a set of $R_{rs}^{modelled}$ spectra. For a given set of simulations, one or more variables can be varied, and the effect on $R_{rs}^{modelled}$ can be evaluated. As an example, using a fixed set of concentrations and SIOPs, $R_{rs}^{modelled}$ can be calculated for a substrate comprised of increasingly bleached coral through a range of water depths. The resulting subsurface remote sensing reflectance spectra can then used to evaluate the ability of various remote sensors to detect this changes in bleaching at different water depths.

The $R_{rs}^{modelled}$ spectra (from SAMBUCA or e.g. Hydrolight) can also be fed to SAMBUCA in 'inversion' mode, and the success of the parameter retrieval can be assessed. This allows for the exploration of parameter retrieval limits in controlled conditions.

6. A simulated SAMBUCA run

Figure 1 shows an example of the graphical output during a SAMBUCA run in simulation mode. The original value for each of the five parameters being retrieved is known. These known values are depicted by a magenta line across each of the five graphs. Each parameter is varied by SAMBUCA in an attempt to minimised the error in Equation 29, and each graph displays the attempted input values (y-axis) with a white line through time (x-axis). In this example, the concentration of chlorophyll CHL (top graph) was over-estimated by SAMBUCA, whereas e.g. the values for depth H (bottom graph) and the concentration of tripton TR were correctly retrieved (Figure 1).

Figure 2 is taken from the same example. The first graph in Figure 2 displays $R_{rs}^{measured}$ as a white dashed line, and the various attempts at matching $R_{rs}^{modelled}$ are shown by a serious of thinner coloured lines, progressing in time from dark green, through blue, to red in colour. These $R_{rs}^{modelled}$ spectra are the result of the variations in the five parameters shown in Figure 1. As these parameters are varied, the IOPs change (Equation 11 and Equation 18). The variations in the 'intermediate' IOPs are displayed in the middle (absorption a) and right (backscattering bb) graphs of Figure 2.

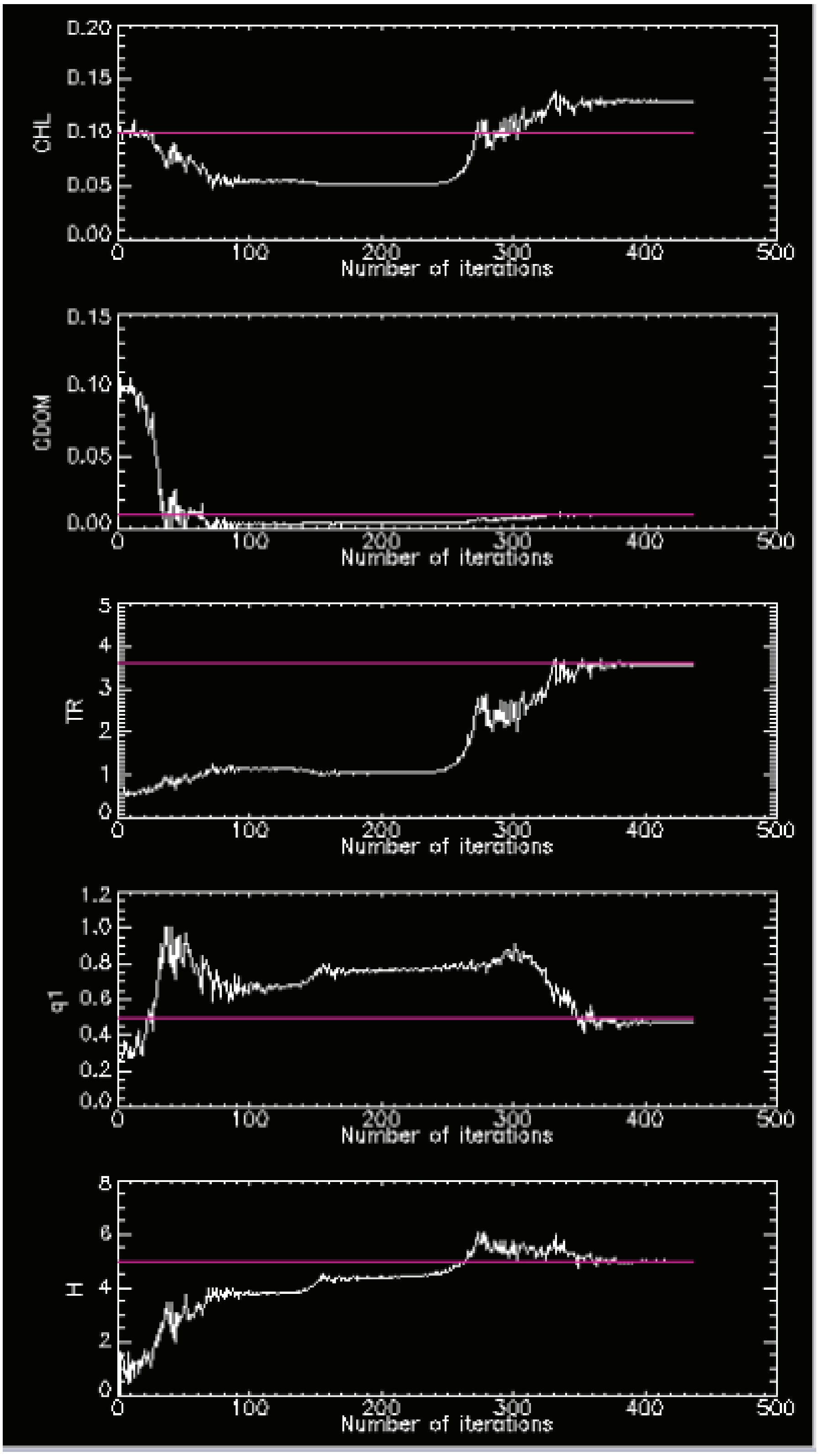

Figure 1: Graphical output of an example SAMBUCA run for a given pair of substrate spectra. Magenta lines depict value for each parameter used to create $R_{rs}^{measured}$, and the white plots depict the path taken by each variable throughout the minimisation process.

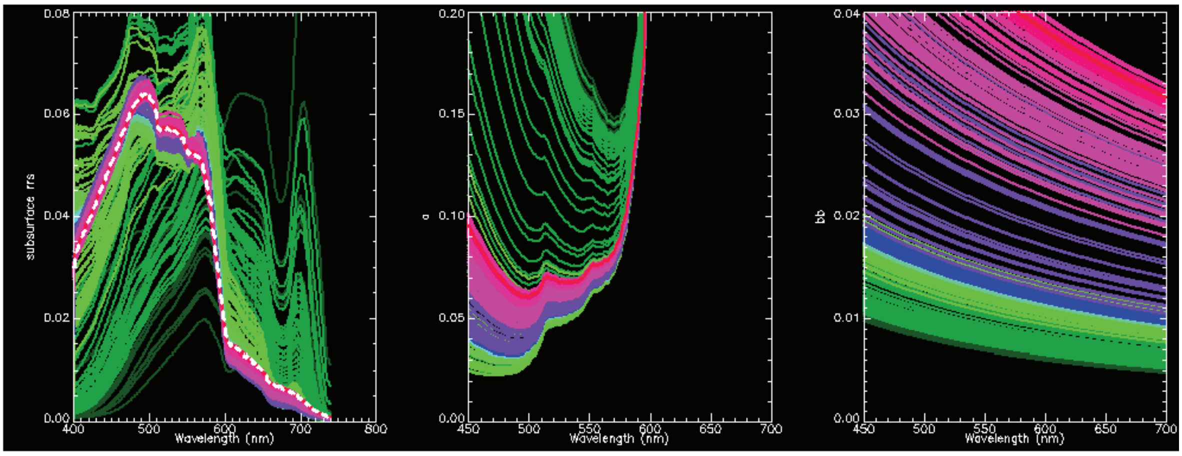

Figure 2: Graphical output of an example SAMBUCA run. The graph to the left shows $R_{rs}^{measured}$ as a white dashed line, and the various attempts at a matching $R_{rs}^{modelled}$ is depicted by a serious of thinner coloured lines, progressing in time from dark green, through blue, to red in colour. The middle and right graphs display the corresponding variations in the absorption coefficient a and backscattering bb coefficient as the optimisation progresses.

If SAMBUCA is being applied to remote sensing imagery, the output from SAMBUCA is an image of each of the variables retrieved. In addition, the error for the best fit between measured and modelled R_{rs} (Equation 29) is retained and can be output in map form. This provides a reliability index for the parameter retrieval, on a pixel-by-pixel basis.

7. Spectral matching

The error computed in Equation 29 is a measure of the goodness-of-fit between measured and modelled spectra. It is the fluctuation of this value that drives the optimisation, and it is therefore central to the applicability of SAMBUCA. Lee et al. (1999, 2001) calculate this through minor variations of the following general equation for spectral albedo matching:

$$fval = \frac{\sqrt{\sum (r_{rsmeasured} - r_{rs\,modelled})^2}}{r_{rsmeasured}} \qquad (31)$$

An alternative matching technique that has been explored for SAMBUCA is the Spectral Angle Mapper (SAM) approach found within the ENVI/IDL image processing software. This technique focuses on the spectral shape, thereby dampening the importance of the spectral albedo. The magnitude of the subsurface remote sensing reflectance is strongly affected by depth, and the original Lee et al. (1998, 1999, 2001) matching approach would seem more suitable. However, experience has shown that using the SAM matching approach works better in some applications. This is attributed to the fact that variations in depth will affect the spectral shape as well, and that SAM matching is therefore still sensitive to depth variations. Also, it is believed that certain noise in image data, e.g. sun-glint or wave-focusing effects (which are more or less wavelength-independent), is more prone to cause errors in parameter retrieval when the magnitude of the spectra also plays a role in the error-assessment.

The SAM approach is based on expressing a spectrum as a vector in n-dimensional space, where n is the number of wavelength bands, and the distance from the origin of the vector is defined by the magnitude of (in this case) reflectance for a particular waveband. Two spectra can then be compared based on the angle between their corresponding vectors. The implementation of this in SAMBUCA is contained in Equation 32.

$$alphaval = Arc\cos\left(\frac{r_{rsmeasured} \cdot r_{rs\,\mathrm{mod}\,elled}}{\sqrt{\sum(r_{rsmeasured})^2}\sqrt{\sum(r_{rs\,\mathrm{mod}\,elled})^2}}\right) \tag{32}$$

fval and alphaval have are used to refer to the errors calculated by Equation 31 and Equation 32 respectively.

Based on experience and referring to Dennison et al. (2004) it is suggested that a function using a combination of *fval* and *alphaval* type error is the most suitable for driving a spectral-matching optimisation algorithm.

*alphafval = alphaval * fval* (33)

Recent applications of SAMBUCA have exclusively relied on alphafval as a spectral closure criteria. As mentioned in the previous section, maps of spectral fit discrepancy are one of the outputs from SAMBUCA.

8. Additional features and modifications

8.1.1. Accounting for more than one optical domain

SAMBUCA has also been configured to allow for several sets of SIOPs within one scene. It is increasingly being recognized that, particularly in coastal waters, horizontal gradients in SIOPs can exist (Dekker et al., 2004). These differing optical domains can be included in the parameterisation, and SAMBUCA retains not only the variables that provided the best fit, but also the set of SIOPs that was used. This is essentially the same approach used for incorporating a spectral reflectance library within SAMBUCA, and does not affect the number of degrees of freedom for the retrieval. As with the spectral library, the drawback to this feature is the increase in processing time.

8.1.2. Tracking the optical closure

For each successful spectral match, SAMBUCA retains the simulated spectrum ($R_{rs}^{modelled}$ in Equation 29). This modelled spectra can then be compared to the original image data, revealing the features of the optical closure between model and data.

If SAMBUCA is being applied to imagery, SAMBUCA outputs both the simulated image, and a difference image of the modelled and original data.

8.1.3. Identifying optically deep water

It was found that the error assessment of SAMBUCA (e.g. alphafval) was not sufficient as a measure of the reliability of variable retrieval. SAMBUCA may be achieving an acceptable spectral closure (low alphafval), but in optically deep water, the depth and substrate composition retrieved should not be relied on: the bottom substrate is by definition not contributing to the solution (SAMBUCA may be randomly choosing a depth and substrate composition).

As part of the reliability assessment of SAMBUCA results, it was therefore necessary to identify where SAMBUCA encountered an optically deep system. This done by comparing the water column component $R_{rs}{}^{C}$ of the solution (Equation 1) with. Where the difference between these is less than a set threshold (for one or more channels), the data is labelled as optically deep and/or the retrieved substrate composition and depth are labelled as unreliable. This threshold can be determined empirically, or based on e.g. the environmental noise in the data.

An alternative to directly labelling water identified as optically, an index of optical depth (IOD) was defined as

$$IOD = (R_{rs}^{modeled} - R_{rs}^{DW}) / NE\Delta R_E$$

where

$R_{rs}^{modeled}$ is taken from Equation 2
R_{rs}^{DW} is the deep water component of R_{rs} (Equations 2 & 4)
$NE\Delta R_E$ is the environmental noise equivalent of the system (Dekker, 1993; Brando and Dekker, 2003; Wettle et al, 2004).

The index of optical depth (IOD) quantifies the contribution of the substrate to remote sensing signal for a given sensor as it uses the noise equivalent difference in reflectance $NE\Delta R(0-)_E$ as a scaling factor. In other words, decreasing values indicate that the system is converging to optically deep. IOD allows identifying three classes of waters:

“optically deep waters”, where no signal from the substrate is visible (i.e. measurable);
“quasi-optically deep waters”, where the contribution from the substrate is small,
“optically shallow waters” where the signal from the substrate is visible

IOD indicates how close the difference between modelled R_{rs} and deep water R_{rs}^{DW} is to the noise inherent in the data.

Where the data is in image form, SAMBUCA outputs a map of the difference between $R_{rs}^{modelled}$ and $R_{rs}{}^{C}$, as well as a map of the IOD. Such a map, together with the error map output, provides an enhanced means of assessing the reliability of the variable retrieval.

For an application focusing on e.g. the retrieval of water column concentrations, the inverse reasoning applies: the closer the modelled spectrum is to being defined by an optically deep system, the more reliable the variable retrieval.

9. SAMBUCA vs. Hydrolight

9.1. Introduction

Hydrolight (Mobley, 1994) is considered the industry standard in underwater light climate modelling software. It is an exact numerical solution to the radiative transfer problem in water, based on Mobley (1994), and has been extensively tested and applied. This model cannot be directly inverted, as it is an exact numerical forward model, and it is furthermore computationally expensive. The model within SAMBUCA on the other hand is a semi-analytical solution to the underwater light field problem and therefore allows for a relatively fast solution of the RT equations. The semi-analytical nature of SAMBUCA lies in the formulation of the remote sensing reflectance associated with an infinitely deep water column (R_{rs}^{DW} in Equation 4) and the optical elongation factors (D_{UC} and D_{UB} in Equations 5 and 6) estimated by (Lee et al., 1998, 1999, 2001). Indeed, the parameterisation of these shape factors were determined through regressions against Hydrolight simulations. For these simulations, Hydrolight was parameterised with field data from the Florida Keys(Lee et al., 1998, 1999, 2001). It was therefore decided to compare SAMBUCA and Hydrolight output when using a different set of model parameters, in order to verify that the analytical approximations in Equation 4. 5 and 6 still hold.

9.1.1. Methodology

SAMBUCA and Hydrolight were parameterised using field data from Heron Island coral reef waters and benthic substrates reported in Table 1. The field data collection and parameterisation are described in detail in Wettle (2005). The parameterisation for the two models was identical.

Table 1. Optical water property parameterisation of SAMBUCA. Values are derived from sets of measurements directly above coral beds and within the lagoon (coral water) or in adjacent open ocean waters (open water). Please see Equation 1 for a description of each variable. For the work reported here, only the coral water parameterisation was used. Where a single value is used, it is listed in the max column.

	coral water		**open water**	
	min	*max*	*min*	*max*
C_{CHL} (μg/L)	0.00	0.25	0.20	0.80
C_{TR} (mg/L)	0.50	4.00	0.00	1.25
C_{CDOM} (conc)		0.003		0.004
S_C		-0.0183		-0.0168
S_{TR}		-0.0101		-0.0098
a^*_{TR} *(550)*		0.0017		0.0043
X_{PHY}		0.0008		0.0016
X_{TR}		0.0118		0.0225
Y		1.178		0.878

The models were compared both for an optically shallow and an optically deep system. For the optically shallow modelling, reflectance spectra for three common types of coral reef benthic substrates were used: coral, sediment and green algae. For each simulation, two substrates were mixed in equal proportions and placed at a depth of one, three, or five metres. All possible combinations of substrates and depths were simulated. The results of the simulations are expressed in remote sensing reflectance R_{rs} (Lu/Ed) just below the waters surface. The output from SAMBUCA and Hydrolight for each combination of parameters were compared. For the optically deep system, a substrate mix of coral and sediment was placed at a water depth of 50 m, and the output from the two models compared.

9.1.2. Results and Discussion

The results for the optically shallow simulations are shown in Figure 3. For each set of parameters (which includes water column properties, depth, and substrate composition) the resulting subsurface remote sensing spectra from both models are coded in the same colour. The Hydrolight output is differentiated by a black outline on its graphical symbol.

The first conclusion that can be drawn is that there is no difference between the two models for the coral and sediment mixture and the green algae and sediment mixture at both 1 and 3 meters. For green algae and sediment, and coral and sediment mixtures at 5 meters depth, Figure 3 reveals a slightly higher values for SAMBUCA, most noticeable around the 480nm to 510nm region. The SAMBUCA and Hydrolight values differ at most by approximately 0.0011 in this region, or 0.11% in subsurface remote sensing reflectance. Assuming a Q factor of 4.00, this equates to a 0.44% difference in R(0-).

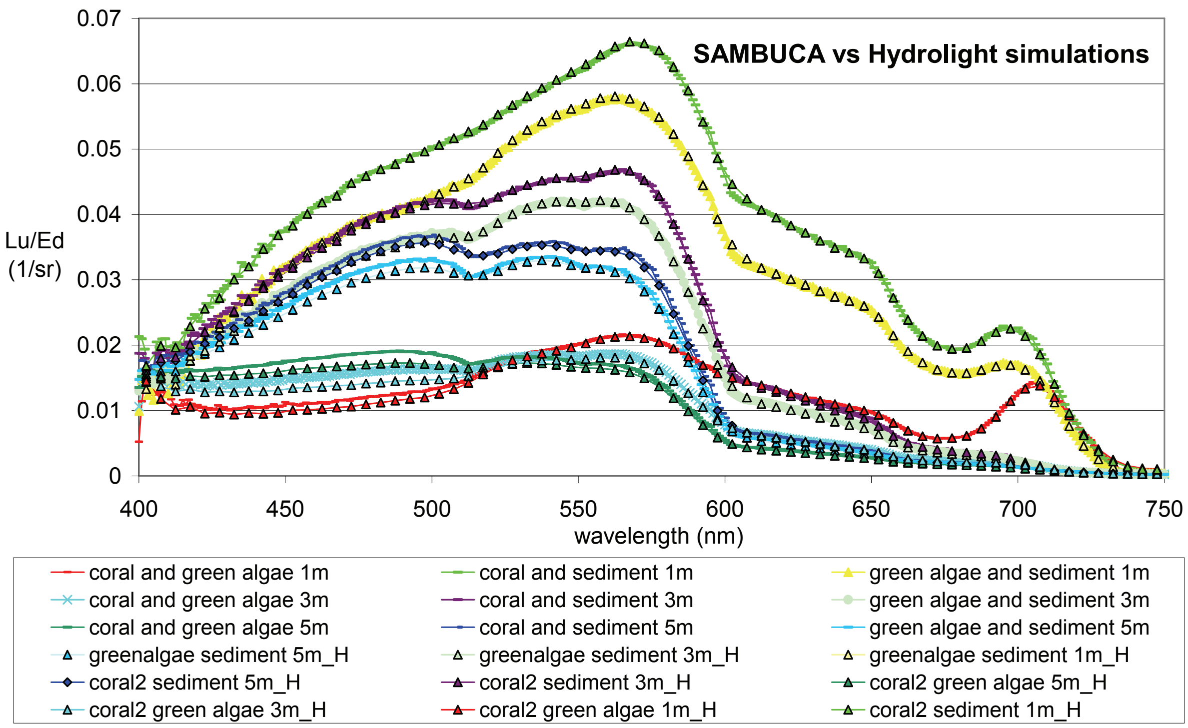

Figure 3: Subsurface remote sensing reflectance spectra from SAMBUCA and Hydrolight simulations of an optically shallow coral reef system. Each pair of benthic substrates are mixed in equal proportions, and placed at 1, 3 and 5 meters depth. The resulting spectra from a given set of parameters is colour-coded the same for both models, whereas the Hydrolight output has a black outline on its graph symbol.

The greatest absolute difference in R_{rs} is seen for the 5 metre deep coral and green algae simulation, which is approximately 0.0016, or 0.64% difference in R(0-), around the 495nm region. This is the darkest substrate mixture at the greatest depth. The difference between model outputs for this substrate mixture decreases with depth; the maximum difference between the 1 metre deep coral and algae simulations is never more than approximately 0.001.

With regards to the three spectra resulting from the dark coral and algae mixture, it is worth noting the decrease in reflectance magnitude with the decrease in depth. In other words, as this benthic substrate mixture is placed deeper in the water column, the remote sensing signal is perceived as brighter. This is perhaps somewhat counter-intuitive, but is caused by the relative brightness of the optically deep water column. In this case, the albedo of the coral and green algae mixture is lower than that of the optically deep water column (given this specific set of optical water properties). Therefore, the magnitude of the remote sensing signal increases with depth for this particular substrate composition. Note that in empirical/traditional approaches this would not be interpreted correctly.

This is corroborated by the optically deep spectral simulations in Figure 4, where the highest value of approximately 0.026 in R_{rs} terms (at 490nm) is higher than the highest values for the three shallow coral and green algae simulations (approximately 0.019 in R_{rs} terms near 490nm).

The deep water simulation features (Figure 4) the same slightly higher values for SAMBUCA, particularly below 500nm region. At most, the difference between the two spectra is 0.0012 in absolute R_{rs} terms, which equates to 0.48% in R(0-) terms.

From this small set of simulations, the greatest discrepancy in R(0-) terms between SAMBUCA and Hydrolight is 0.64%. This was achieved by running both models using a substrate mixture of coral and green algae at 5 metres depth. This figure can be compared to the environmental noise equivalent value of remote sensing data. For example, Brando and Dekker (2003) established that their Hyperion data over Moreton Bay, Queensland, Australia, contained an environmental noise equivalent in remote sensing reflectance terms of between 0.3 - 0.4% , depending on wavelength. The maximum discrepancy between the SAMBUCA and Hydrolight simulations presented here is therefore less than two environmental noise equivalent levels as reported for Hyperion in the literature, and only for a limited number of the sensor's spectral bands. With the exception of two of the simulated pairs, the difference between SAMBUCA and Hydrolight is within one Hyperion environmental noise equivalent envelope.

In Wettle (2005) the environmental noise equivalent for selected MERIS image data is established. For the wavelength interval where the highest discrepancy (0.64%) is noted between SAMBUCA and Hydrolight (around 490nm) the MERIS environmental noise equivalent in remote sensing reflectance terms is approximately 0.18%. In other words, there is a maximum of approximately three environmental noise equivalent levels of MERIS data between these SAMBUCA and Hydrolight simulations. This occurs for one spectral band of the MERIS sensor. For the remaining spectral bands the maximum difference between these SAMBUCA and Hydrolight simulations is less than three environmental noise equivalent levels.

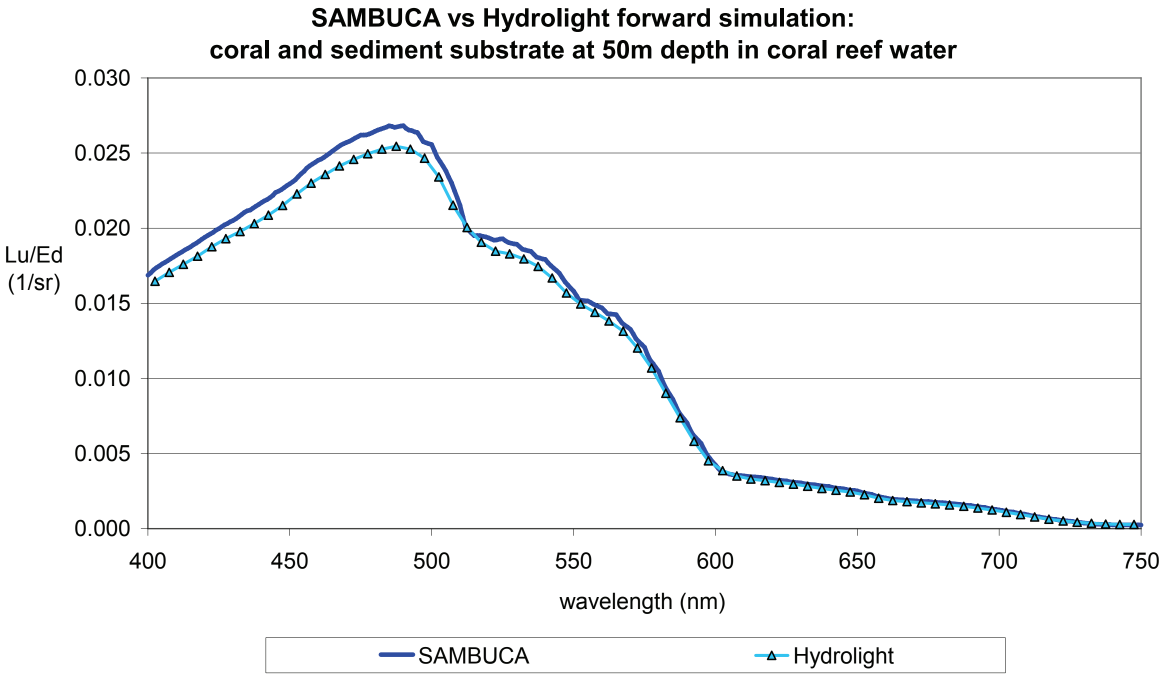

Figure 4: Subsurface remote sensing reflectance spectra from SAMBUCA and Hydrolight simulations of an optically deep coral reef system. A mixture of coral and sediment substrate is placed at 50 meters depth. The SAMBUCA output is colour-coded dark blue.

As a rule of thumb, a signal within two levels is not considered to be measurable. Therefore, only one set of simulations produced a difference between SAMBUCA and Hydrolight that is theoretically measurable by the MERIS sensor (three environmental noise levels). Furthermore, this difference is theoretically only able to be detected on one of the MERIS spectral channels.

9.1.3. Conclusions and recommendations for SAMBUCA vs. Hydrolight

These simulations were intended as a verification that the coral water parameterisation of SAMBUCA is still valid for the shape factor approximations derived from the Hydrolight work done by Lee et al. (1998, 1999, 2001). For this optical system, SAMBUCA provided a semi-analytical solution to the underwater light field which was in good agreement with the exact numerical solution offered by Hydrolight. Discrepancies between the two models were centred on the 500nm region, where SAMBUCA consistently over-estimated the subsurface reflectance. Overall, these discrepancies were considered acceptable, particularly in the context of sensor environmental noise characteristics such as that of Hyperion and MERIS.

Although small, discrepancies between the two models were revealed. This suggests that approximations used to derive analytical models from exact numerical solutions should routinely be tested and evaluated within the context of the optical climate and sensor specifications of the application.

For the purposes of this project, SAMBUCA's agreement with the HYDROLIGHT solutions is considered acceptable. However, if SAMBUCA is to be routinely applied in a specific geographic region, and in particular if it features a differing optical climate than the coral environments discussed here (e.g. temperate, coastal system), it would be advisable to re-establish the shape factor approximations (Equation 4,5 and 6) established by Lee et al. (1998, 1999, 2001) for waters in the Bahamas. This would be done through regressions on a large set of Hydrolight simulations, using optical water quality parameters measured for the region/sites in question.

An additional example of when this could be warranted would be the integration of SAMBUCA into an operational, site-specific, monitoring system. An example would be monitoring coral bleaching for the Great Barrier Reef; the analytical approximations within SAMBUCA's RT equations could be re-established for optimal alignment with the optical climate(s) of the GBR lagoon.

10. Conclusions

SAMBUCA is a quantitative and sophisticated technique for retrieving environmental variables from remote sensing data. It uses an analytical expression for subsurface remote sensing reflectance R_{rs} to estimate a set of variables from a single R_{rs} measurement. This is done through an optimisation process, also known as a predictor-corrector scheme.

The advantages of SAMBUCA include:

• Objective and repeatable

• Limited need for fieldwork

• Easily applicable across sensor/data types

• Easily adaptable for simulation work

• Ability to cope with variations in water column optical quality and depth

SAMBUCA is based on the approach developed by Lee et al. (1998, 1999, 2001), but features several modifications:

• The water column parameterisation is expressed terms of concentrations of optically active constituents in the water column.

• The water column parameterisation defines the optically active constituents as being composed of pure water, chlorophyll, CDOM, and tripton.

• The substrate parameterisation allows for linear un-mixing of several (typically two) bottom type reflectance spectra within each pixel, retaining the proportion of each substrate estimated to be present.

• The substrate parameterisation can access a library of reflectance spectra from which all combinations of bottom substrate types are investigated.

• The SIOP parameterisation can access several sets of SIOPs from which the set that provided the best spectral match is retained.

• The approximation of a modelled spectrum to an optically deep spectrum is tracked

In principle, SAMBUCA can retrieve concentrations of optically active constituents in the water column, bottom substrate composition (from a substrate spectral library), depth, and select an appropriate set of SIOPs on a pixel-by-pixel basis.

SAMBUCA can be run in simulation mode (e.g. sensitivity studies) as well as applied mode (e.g. on hyperspectral imagery). The set of environmental variables to be retrieved is configurable, and is determined by the nature of the application as well as the dimensionality of the data set.

The success of the variable retrieval can be improved by maximizing the ratio of wavelength channels to number of variables being sought, as well as by fixing the ranges and starting points of the variables in question.

SAMBUCA simulations were found to be in good agreement with Hydrolight output, confirming that the approximations inherent to the analytical nature of SAMBUCA were valid for the coral reef optical system being studied in those specific simulations.

11. Recommendations and future work

SAMBUCA results are highly sensitive to the substrate reflectance spectral library and water optical parameterization used as input. Care should be taken in the selection of these, and trial runs using a limited set of spectra are recommended. Based on these initial results, the input parameters may need to be refined or altered.

The atmospheric correction of the data ingested by SAMBUCA is furthermore critical to the SAMBUCA output. Where data is not available for validation of atmospheric correction, it can be advantageous to perform SAMBUCA runs on a limited set of measured (atmospherically corrected) spectra in order to asses the closure of SAMBUCA. Towards the red and near infra-red region of the spectrum, SAMBUCA modelled spectra rapidly decrease towards zero water leaving reflectance, and changes in the SAMBUCA parameterisation have little effect on the modelled spectrum. Hence, if spectral closure in this region of the spectrum is relatively poor, it may be an indication that the atmospheric correction needs to be re-visited.

Where the substrate reflectance library used by SAMBUCA and/or the data set being processed is large processing times may be prohibitively high (e.g. a high spatial resolution Quickbird scene was estimated to require in the order of 100-200 processing days). One means of reducing processing time is to apply biological constraints on the substrate reflectance spectra. For example, certain types of benthic substrates may be known to exist only at certain depths in the water column. Such constraints would reduce the number of substrate permutations attempted by SAMBUCA for any given depth interval. Furthermore, data volume reduction techniques are currently being explored, whereby data variability is reduced based on the environmental noise inherent to the system (an unsupervised classification variant), effectively reducing the number of spectra that require processing by SAMBUCA.

The SAMBUCA code would benefit from a re-design of the parameterisation input environment. Currently under consideration is a plan to enable the user to define all optical water quality parameters and substrate reflectance libraries in a spreadsheet template.

References

Brando, V. E. and Dekker, A. G. (2003). Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality. IEEE Trans.Geosci.Remote Sens., 41:1378–1387.

CoralReefs (2004). Special issue of coral reefs, on remote sensing of coral reefs. Coral Reefs, 23.

Dekker, A. G. and Peters, S. W. M., (1993), The use of the Thematic Mapper for the analysis of eutrophic lakes: A case study in The Netherlands, *International Journal of Remote Sensing*, **14**: 799-822.

Dekker, A. G., Brando, V. E., Anstee, J. M., Pinnel, N., Kutser, T., Hoogenboom, H. J., Peters, S. W. M., Pasterkamp, R., Vos, R. J., Olbert, C., and Malthus, T. J. (2001). Imaging spectrometry of water, volume IV. Kluwer Academic Publishers, Dordrecht.

Dekker, A. G., Brando, V. E., Oubelkheir, K., and Wettle, M. (2004). When freshwater meets ocean water: How variable siops affect remote sensing products of estuaries, bays and coastal seas. Ocean Optics XVII, Freemantle, Australia.

Dennison, P. E; Halligan, K. Q.; Roberts, D. A..(2004) A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. Remote Sensing of Environment. 93: pp. 359-367

Goodman, J. and Ustin, S. (2003). Airborne hyperspectral analysis of coral ecosystems in the hawaiian islands. 30th International Symposium on Remote Sensing of Environment, Hawaii, pages 10–14.

Hedley, J. and Mumby, P. (2003). A remote sensing method for resolving depth and subpixel composition of aquatic benthos. Limnology and Oceanography, 48:480– 488.

Kutser, T., Dekker, A. G., and Skirving, W. (2003). Modelling spectral discrimination of great barrier reef benthic communities by remote sensing instruments. Limn. and Oceanogr., 48(1):497–510.

Lee, Z. P. (1994). Visible-infrared remote sensing model and applications for ocean waters. PhD thesis, University of South Florida.

Lee, Z. P., Carder, K. L., Mobley, C. D., Steward, R. G., and Patch, J. S. (1998). Hyperspectral remote sensing for shallow waters. 1. a semianalytical model. Applied Optics, 37(27):6329–6338.

Lee, Z., Carder, K. L., Mobley, C. D., Steward, R. G., and Patch, J. F. (1999). Hyperspectral remote sensing for shallow waters: 2. deriving bottom depths and water properties by optimization. Applied Optics, 38(18):3831–3843.

Lee, Z., Kendall, L. C., Chen, R. F., and Peacock, T. G. (2001). Properties of the water column and bottom derived from airborne visible imaging spectrometer (aviris) data. Journal of Geophysical Research, 106(C6):11639–1165

Limnology and Oceanography (2003). Special issue on Light in shallow water. Limnology and Oceanography, 48 (1, part 2).

Maritorena, S., Morel, A., and Gentili, B. (1994). Di_use reflectance of oceanic shallow waters: influence of water depth and bottom albedo. Limnol.Oceanogr., 39(7):1689–1703.

Mobley, C. D. (1994). Light and water; Radiative transfer in natural waters. Academic Press, London.

Pope, R. M. and Fry, E. S. (1997). Absorption spectrum (380-700 nm) of pure water. ii. integrating cavity measurements. Appl.Opt., 36(33):8710–8723.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (2003) *Numerical Recipes in C: The Art of Scientific Computing* 2nd Edition, Cambridge University Press, 994 pp.

Wettle, M., Brando, V.E., Dekker, A.G. (2004). A methodology for retrieval of noise equivalent spectra applied to four Hyperion scenes of the same tropical coral reef. *Remote Sensing of Environment*, **93,** 188- 197.

Wettle, M. (2005). Monitoring coral reef bleaching from space – a feasibility study using a physics-based remote sensing approach. PhD thesis, University of Hull.