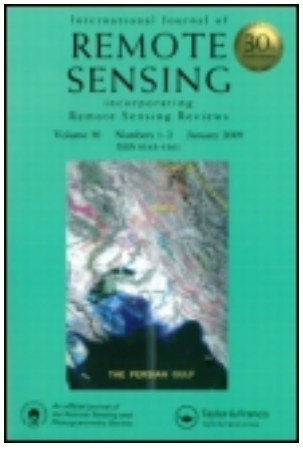

International Journal of Remote Sensing

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tres20

An advanced algorithm for operational retrieval of water quality from satellite data in the visible

D. Pozdnyakov Corresponding author [a], A. Korosov [a], H. Grassl [b] & L. Pettersson [c]

[a] Nansen International Environmental and Remote Sensing Centre, 26/28 Bolshaya Monetnaya, 197101, St Petersburg, Russia

[b] Max-Planck Institute for Meteorology, Bundesstr. 53, 20146, Hamburg, Germany

[c] Nansen Environmental and Remote Sensing Centre, Thormølhensgate 47, N-5006, Bergen, Norway

Available online: 22 Feb 2007

To cite this article: D. Pozdnyakov Corresponding author, A. Korosov, H. Grassl & L. Pettersson (2005): An advanced algorithm for operational retrieval of water quality from satellite data in the visible, International Journal of Remote Sensing, 26:12, 2669-2687

To link to this article: http://dx.doi.org/10.1080/01431160500044697

International Journal of Remote Sensing
Vol. 26, No. 12, 20 June 2005, 2669–2687

An advanced algorithm for operational retrieval of water quality from satellite data in the visible

D. POZDNYAKOV*†, A. KOROSOV†, H. GRASSL‡ and L. PETTERSSON§

†Nansen International Environmental and Remote Sensing Centre, 26/28 Bolshaya Monetnaya, 197101, St Petersburg, Russia

‡Max-Planck Institute for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany

§Nansen Environmental and Remote Sensing Centre, Thormølhensgate 47, N-5006 Bergen, Norway

(*Received 6 December 2004; in final form 13 December 2004*)

We present a new operational algorithm for the retrieval of water quality from optical remote sensing data for both clear and turbid waters. It contains an array of neural networks providing input for the Levenberg–Marquardt multivariate optimization procedure as the final retrieval tool. With a given accuracy threshold, the developed algorithm is sufficiently robust to data with noise up to 15% for certain hydro-optical conditions. To avoid inadequate retrieval results, the algorithm identifies and eventually discards the pixels with inadequate atmospheric correction and/or water optical properties incompatible with the applied hydro-optical model. This procedure also identifies coccolith expressions. Examples of practical applications of the developed algorithm are given.

1. Introduction and the study objective outline

Satellite remote sensing of the world oceans became increasingly important over the recent decades in light of both the problem of global climate change and frequent deterioration of the aquatic ecology status driven by the ever-increasing needs of growing populations for drinking water as well as fish and other sea food.

Only *visible* radiation appreciably penetrates into the water column (e.g. Jerlov 1976). Travelling through the aquatic medium, photons are subject of a variety of interactions with the encountered molecules, which often leads to essential alterations of the spectral composition of the upwelling radiative flux (e.g. Pozdnyakov and Grassl 2003). Thus, the backscattered flux emerging from beneath the water surface contains the information about the optical properties of the water column. The water colour is a convolution of all interactions of photons with the aquatic medium. Therefore, satellite optical sensors are in principle able to observe in-water parameters and associated processes (Kondratyev *et al.* 1999).

In the case of open ocean waters, the algorithms for the retrieval of the desired water quality parameter(s) are rather unsophisticated semi-analytical ones (e.g. the OC4 algorithm currently used by the National Aeronautics and Space Administration (NASA) (O'Reilly *et al.* 1998, Ackleson 2001)). This is possible due to the optical simplicity of such waters, referred to as case I waters by Morel and Prieur (1977). Open ocean water is generally only loaded with indigenous

*Corresponding author. Email: dmitry.pozdnyakov@niersc.spb.ru

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160500044697

phytoplankton and their *retinue*, i.e. accompanying and co-varying products of their life cycles as well as some microscopic organisms such as flagellates, bacteria and viruses, which are also indigenous to off-shore/mid-oceanic waters (Sathyendranath 2000). Therefore, spectral variations in upwelling radiance must be strictly proportional to the concentration of phytoplankton, or its proxy—the concentration of chlorophyll (*chl*) contained in algal cells. In this case a relationship can be established between the concentration of *chl* and the emerging radiance at two or more wavelengths normalized by the incident irradiance at the same wavelengths.

However, waters in the coastal zone as well as inland waters, referred to as case II waters by Morel and Prieur (1977), are characterized by optical properties that are influenced not only by the indigenous phytoplankton and the substances originating from the phytoplankton's life cycle, but also by other matters independent of phytoplankton (denoted herein as *chl*), notably inorganic/terrigenous particulate matter in suspension (*sm*) and dissolved organics (*doc*). Their content in the water column is often abundant enough to compete with phytoplankton in influencing the resultant optical properties of case II waters, thus rendering such waters optically very complex.

When observing case II waters from above, it is impossible to retrieve the concentration of a single component (e.g. *chl*) without also inferring simultaneously the content of the other major water constituents influencing water colour, the so-called colour producing agents (CPAs) encompassing first and foremost *chl*, *sm* and *doc* (Bukata *et al.* 1995). These constituents are traditionally associated with Water Quality Parameters (WQP), although more substances contribute to these (Anonymous 1992). Thus, case II waters complicate significantly the retrieval of CPAs by remote sensing.

Another complicating factor is the inaccurate atmospheric correction (Land and Haigh 1997, Ruddick *et al.* 2000). The major challenge here is an accurate assessment of the path radiance originating from photon interactions with the atmospheric aerosol, especially with the aerosol in the lower troposphere.

The atmospheric aerosol over case I waters is generally homogeneous spatially and chemically both in the upper and lower troposphere. It is mostly marine particulate matter, and its influence on light transfer is largely by scattering. Its optical properties are established fairly well and, hence, can be adequately assessed in the atmospheric correction procedure. The inaccuracy of water constituent retrievals arising from atmospheric correction errors over case I waters is about 5% (for references see Pozdnyakov *et al.* 2000a).

In the atmosphere over case II waters the aerosol composition originates from a wealth of distributed and point sources, often anthropogenic. This results in a pronounced spatial and compositional inhomogeneity of the aerosol. In addition, the aerosol over such areas is no longer purely scattering but also absorbing. This complicates significantly the atmospheric correction (for references see Pozdnyakov *et al.* 2000b).

In general, the error in atmospheric correction can be distributed either uniformly or normally and its level can depend on and be independent of wavelength λ. The scarce reports indicate that the error in upwelling radiance arising from imperfect atmospheric correction over case II waters can be as high as 15% in the short wavelength region diminishing nearly linearly with λ.

Earlier we reported (Pozdnyakov and Lyaskovsky 1999) that under conditions of low noise input and availability of an adequate hydro-optical model the

Levenberg–Marquardt (L–M) multivariate optimization procedure can successfully be used for the retrieval of CPAs in case II waters. However, the procedure performance is rather slow, which prevents using it for operational processing of satellite images. Neural networks (NN) used as a water constituent retrieval tool prove to be much faster, although they are inferior to the L–M procedure in terms of the attainable accuracy.

In light of the above, it seemed worthwhile for us to comparatively analyse the retrieval accuracy of the L–M and NN algorithms for signals with noise. Further, exploiting the advantages of both techniques, we try to develop a new retrieval tool meeting the requirements of both accuracy and high operability given a noisy input.

We begin with a concise description of the basic principles of both techniques. A more detailed one can be found elsewhere (e.g. Kondratyev *et al.* 1990, Press *et al.* 1992, Atkinson and Tatnall 1997).

2. Multivariate optimization procedure

The subsurface remote sensing reflectance, $R_{rsw}(\lambda, \boldsymbol{C}, a, b_b)$ is the upwelling spectral radiance just beneath the water–air interface, $L(-0, \lambda)$ normalized by the downwelling spectral irradiance, $E(-0, \lambda)$ at the same level (Jerome *et al.* 1996). It depends on both $C = \sum C_i$, the concentration vector encompassing all the major CPAs and the specific[i] absorption a and backscattering b_b coefficients of the coexisting water constituents. If S_j is the measured/retrieved subsurface remote sensing reflectance at a wavelength λ_j, then the residual between S_j and R_{rswj} can be computed by one of the following ways:

$$g_j = \left(S_j - R_{rswj}\right)/S_j \tag{1}$$

$$g_j = \left(S_j - R_{rswj}\right)/R_{rswj} \tag{2}$$

$$g_j = \left(S_j - R_{rswj}\right) \tag{3}$$

The multidimensional least-square solution using all wavelengths is found by minimizing the squares of the residuals:

$$f(\boldsymbol{C}) = \sum_j g_j^2(\boldsymbol{C}) \tag{4}$$

The absolute minimum of $f(\boldsymbol{C})$ can be found with the Levenberg–Marquardt finite difference algorithm (Levenberg 1944, Marquardt 1963), which assures a rapid convergence of the iterative procedure. The following iterative expression is used to this goal:

$$C_{k+1} = C_k + \zeta_k \left(F_k^t F_k + \mu_k D_k\right)^{-1} F_k^t \left(1 - \frac{R_{rsw}(C_k)}{S_k}\right) \tag{5}$$

where k=iteration step, $D_k = diag\left(F_k^t F_k\right)$ = diagonal matrix consisting of the elements $F_k^t F_k$, with $F(\boldsymbol{C}) = \left|\frac{\partial R_{rswj}}{\partial C_j}\right|$ = matrix composed of $n \times m$ elements (n=number of wavelengths, at which the measurements/retrievals S_j have been conducted,

m=dimension of the concentration vector $\boldsymbol{C}$), $F^t(\boldsymbol{C})$=transposed matrix, μ_k=direction of minimization, ζ_k=length of the minimization step.

The actual length ζ_k can be chosen easily and effectively through applying the method of reduction, which consists in the following: at a certain iteration step k a deliberately high value of ζ_k is chosen (e.g. 50). From equation (5) a new concentration vector $\boldsymbol{C}_{k+1}$ is determined, for which the inequality $f(\boldsymbol{C}_{k+1})<f(\boldsymbol{C}_k)$ is tested. If it is not satisfied, ζ_k is reduced by a factor of two, and a new round of search is initiated. This iterative procedure is continued until the above inequality is satisfied, i.e. until the newly determined concentration vector proves to be closer to the minimum of the function of residuals $f(\boldsymbol{C})$ as compared to the concentration vector established in the previous iteration step.

In the course of the iterative descent to the global minimum of the function of residuals, the above procedure could end in a *local* minimum. If the determined $\boldsymbol{C}_k$ happens to be on the downhill and the length of the iteration step ζ is sufficiently high, then the next concentration vector $\boldsymbol{C}_{k+1}$ might occur on the downhill in direction to a deeper minimum. In this case the process of search will move to a new point provided the inequality $f(\boldsymbol{C}_{k+1})<f(\boldsymbol{C}_k)$ is satisfied. Thus, a search with a variable length ζ can, under favourable conditions, evade a local *minor* minimum and end either in a *deeper* or even in the *global* minimum.

However, equally possible is the situation that the concentration vector $\boldsymbol{C}_{k+1}$ is *uphill* of a certain local minimum of $f(\boldsymbol{C})$. In this case the search ends in this local minimum, which corresponds to an inadequate solution of the inverse problem. Naturally, the probability for the latter option increases with measurement/retrieval errors.

To avoid the above difficulty, it is advisable to initiate the iteration procedure for an array of initial guess values, $\boldsymbol{C}_0$. Afterwards the deepest minimum of $f(\boldsymbol{C})$ is selected. The number $\boldsymbol{N}$ of the initial vectors $\boldsymbol{C}_0$ should not be excessively high because the computation time for the inverse problem solution increases proportionally with $\boldsymbol{N}$.

But the use of an array of initial vectors $\boldsymbol{C}_0$ does not guarantee that the iterative procedure be converging, or/and the eventually established concentration vector $\boldsymbol{C}$ be realistic. To overcome this problem, *a priori* limits can be imposed upon each of CPA concentrations, C_i:

$$C_{i\,\min} \le C_i \le C_{i\,\max} \tag{6}$$

where i is the ith constituent of the aquatic medium.

3. Neural networks (NN) algorithm

There are many different types of neural networks but one of the most commonly used neural networks in remote sensing is the Multi-Layer Perceptron (MLP) (Atkinson and Tatnall 1997). The MLP generally consists of three layers: the input layer neurons are the elements of a vector, which might consist of radiances at certain wavelengths; the second layer is the internal or 'hidden' layer; in the third layer, the number of neurons equals the number of parameters to be determined. Each neuron in the network is connected to all neurons in both the preceding and subsequent layer with an associated weight.

The input signals are transferred to the neurons in the next layer in a feed-forward manner. As the signal propagates from neuron to neuron, it is modified by the

appropriate connection weight. The receiving neuron sums the weighted signals from all neurons in the previous layer. The total input that a single neuron receives is weighted in the following way:

$$net_j = \sum \omega_{ji} o_i \tag{7}$$

where ω_{ji} is the weight between neuron i of the preceding layer and neuron j in the receiving layer, and o_i is the output from neuron i of the preceding layer. The output from a given neuron j is then obtained from:

$$o_j = f\left(net_j\right) \tag{8}$$

The function f is usually a nonlinear sigmoid function. It is applied to the weighted sum of inputs before the signal reaches the next layer. When the signal reaches the output layer, the network output is produced. The created network should be trained so that it can generalize and predict outputs from inputs that it has not processed before. A training pattern is fed into the neural network and the signals are forwarded. After that, the network output is compared to the true output, the error is computed and back-propagated through the network. As a result, the connection weights are modified following the generalized rule:

$$\Delta\omega_{ji}(n+1) = \eta\left(\delta_j o_i\right) + \alpha\Delta\omega_{ji}(n) \tag{9}$$

where η is the learning rate parameter, δ_j is an index of the error change rate, α is the momentum parameter. The training is conducted until the output error reaches a desired level. The trained neural network is then tested against some verification data to assess the network performance.

4. Assessment of robustness of the L–M and NN procedures against measurement errors

The L–M procedure in the MatLab medium has been realized by A. Lyaskovsky (Pozdnyakov and Lyaskovsky 1999) using the recipes given by Press *et al.* (1992). An improved code for the L–M procedure used in the present study was developed by one of us (A.K.) using the C++ language.

Subsurface remote sensing reflectance $R_{rsw}(\lambda)$ was calculated using an empirical equation derived from numerical simulations (Jerome *et al.* 1996) that relate $R_{rsw}(\lambda)$ to the inherent hydro-optical characteristics, here the absorption and backscattering coefficients a and b_b. The reasons for such a choice of the relationship in question are given elsewhere (Pozdnyakov and Grassl 2003).

Being additive by nature, the coefficients a and b_b are a sum of respective contributions (a_i, b_{bj}) from the major CPAs. Calculations of the resultant bulk coefficients for the simulated aquatic medium were performed by using the hydro-optical model suggested by Kondratyev *et al.* (1990). The model is a collection of tabulated spectral values of cross-sections of absorption (a^*) and backscattering (b_b^*) for the CPAs *chl*, *sm* and *doc* within the spectral region 400–700 nm, as obtained for case II waters with the trophic status varying from oligotrophic to eutrophic. The values of $R_{rsw}(\lambda)$ were calculated for a very wide range of concentrations of *chl*, *sm* and *doc* (0–70 μg l^{-1}; 0–30 mg l^{-1} and 0–30 mgC l^{-1}, respectively).

Up to 400 first guess concentration vectors $\boldsymbol{C}_0$ were randomly selected for the L–M procedure. About 360 000 spectra of R_{rsw} were generated, and the spectral values of S_i were simulated by contaminating the calculated spectral values R_{rsw} with

'noise'. The following measurement error levels were considered: 0, 5, 10 and 15%. For each error level, the error distribution was assumed to be either normal or uniform. In addition, the numerical experiments were conducted for two options: the error was assumed to be dependent on or independent of the wavelength λ.

As mentioned above, the value of the residual can be calculated by several ways (equations (1), (2) and (3)). When options 1 and 2 are used, the L–M procedure seeks a solution that 'pulls' R_{rsw} to S_i as closely as possible, whereby the closest R_{rsw} can be either totally above or below S_i spectral values. Option 3 might result in spectral values of R_{rsw} located above as well as below the spectral values of S_i, the principle of least squares being fulfilled.

A comparative analysis of the results of numerical simulations conducted for the three options has shown that for the error level chosen here in the range 0–15% all three yield nearly identical results. Even at a 15% error level the absolute and normalized differences between S_i and $R_{rsw\ i}$ across the entire visible spectrum are still insignificant and hence the L–M procedure assures very similar retrieval results. However, this is not the case when the L–M procedure is applied to data from satellite sounding.

The criterion for the termination of the iterative procedure is $f(\boldsymbol{C}) \leqslant 10^{-5}\boldsymbol{C}$ for options 1 and 2, and $f(\boldsymbol{C}) \leqslant 10^{-5}\,\mathrm{sr}^{-1}$ for option 3.

Regarding the NN procedure, we used the software developed at the University of Stuttgart, Institute for Parallel and Distributed Performance Systems (SNNS 1995). The number of neurons in the input layer equals the number of spectral channels of the satellite sensor (6, 7 or 9 in the case of Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS) and Medium Resolution Image Spectrometer (MERIS), respectively). By trial and error the number of neurons in two hidden layers for each option of the CPA retrieval accuracy was fixed to 30. The output layer contains only three neurons corresponding to the desired CPA concentrations (*chl*, *sm* and *doc*).

The numerical experiments with the NNs were conducted similarly to the ones for the L–M case, i.e. through simulation of spectral values of subsurface remote sensing reflectance, R_{rsw} on the basis of both the parametric expression for R_{rsw} as a function of $\boldsymbol{a}$ and $\boldsymbol{b_b}$ given by Jerome *et al.* (1996) and the hydro-optical model suggested by Kondratyev *et al.* (1990). The same levels of error were applied (i.e. 0, 5, 10 and 15%) and the same assumptions were made as to the nature of error distribution and its dependence on wavelength. More than 8000 spectra of R_{rsw} were generated including the verification pool, for which 2000 R_{rsw} spectra were reserved. The training continued until the rms error equalled 0.00049, requiring about 10 000 training cycles.

Figure 1(*a*), (*b*) and (*c*) illustrates the results of a comparison for the L–M and NN procedures. The tested ranges of concentrations of *chl*, *sm* and *doc* corresponded to variations of these CPAs characteristic of both large inland waters (such as Lakes Ladoga, Onega and the Great North American Lakes) and many coastal waters in mid- and high latitudes (Petrova 1990). The error level was first set to zero. In conformity with our previous report (Pozdnyakov and Lyaskovsky 1999), the L–M procedure, as opposed to NN emulators, is capable of reaching (at least under noiseless conditions) a significantly higher accuracy of retrievals of *chl*, *sm* and *doc* in a wide range of concentrations. For further discussion, it is important to underline that the NN procedure designed and trained for a wide range of concentrations of *chl*, *sm* and *doc* (C_{chl}, C_{sm}, C_{doc}) provides the least accurate results for waters with low CPA concentrations.

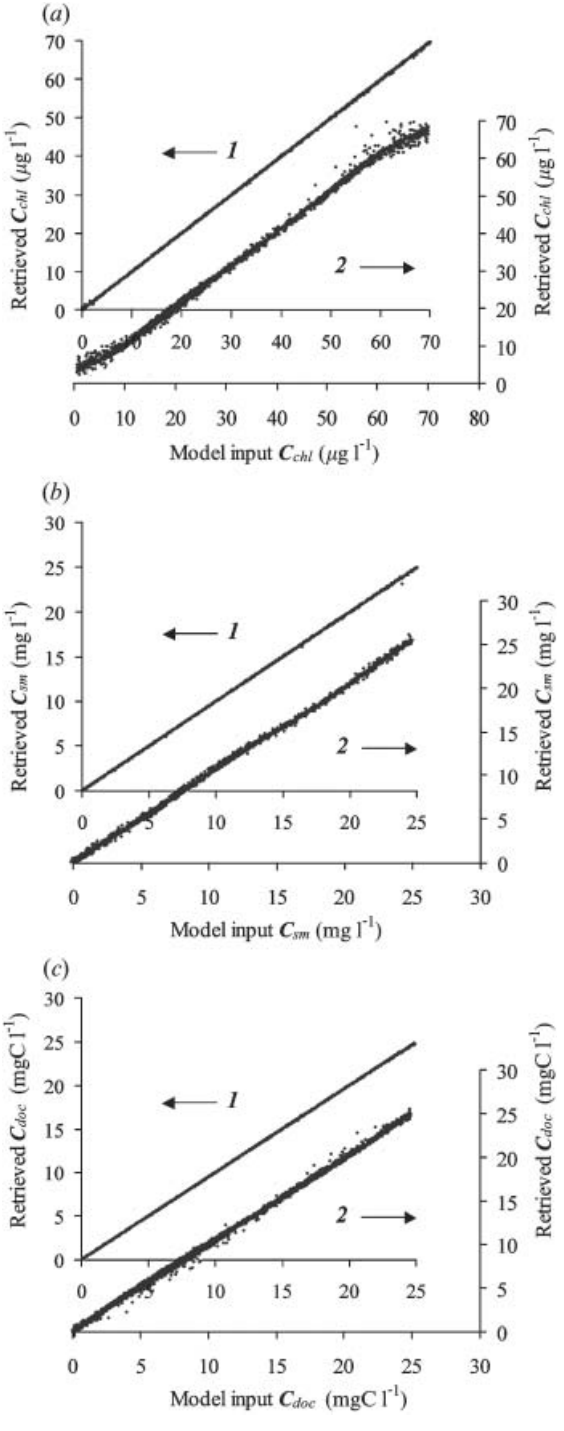

Figure 1. Comparison of performance efficiency of the L–M (***1***) and NN (***2***) procedures for the retrieval of *chl* (*a*), *sm* (*b*), and *doc* (*c*) concentrations without measurement errors.

The above impediment can be largely allayed if some 'specialized' NN emulators designed and trained specifically for the case of low CPA concentrations are employed. Such a specialized NN is trained for a range of small concentrations of a sole CPA, whereas the other CPAs are allowed to vary within the wide limits indicated above.

The expediency of such an approach is illustrated in table 1 providing the comparative data on the mean relative and mean absolute errors of retrieval of C_{chl}, C_{sm}, C_{doc} when a 'broad-band' (B-B) and 'narrow-band' (N-B) NN is applied.

Relative errors were calculated as the difference between retrieved (C_i*) and model concentrations (C_i**) for C_{chl}, C_{sm}, C_{doc} normalized by C_i**. Accordingly, when determining absolute errors no normalization was applied. The values given are averaged over 2000 concentration vectors used in the numerical experiment.

Our simulations indicate that for measurement errors up to 15%, the L–M procedure is more stable than the NN. It was also found that both the error level and its distribution as well as its dependence on wavelength strongly influence the attainable retrieval accuracy. The latter is also a function of the CPA concentration vector *per se*.

Assuming that, at least, for case II waters, the *chl* retrieval errors specified in table 2 for a number of concentration ranges are acceptable, the simulations indicate that these accuracy requirements can only be fulfilled for specific hydro-optical conditions, i.e. the concentrations of *chl*, *sm* and *doc* have to fall into specific concentration ranges.

Our simulations also indicate that the concentration ranges shrink with increasing measurement error. The most stringent limitations arise in the case of normally distributed and λ-dependent errors (the spectral dependence was assumed to be linearly decreasing with λ (Land and Haigh 1997, Ruddick *et al.* 2000)).

Table 3 illustrates the results of our simulations for a normally distributed noise in the model input data. The most stringent requirements to hydro-optical conditions arise in the case of spectrally dependent errors. It is noteworthy that, unlike L–M, the NN procedure is incapable of reaching the accuracy predefined in table 2 in all of the CPA concentration ranges considered in this study in the case of spectrally dependent errors.

Table 1. Accuracy of retrieval of C_{chl}, C_{sm}, C_{doc} utilizing a 'broad-band' NN (B-B NN) or a 'narrow-band' NN (N-B NN) for the reduced concentration range 0–5 ($\mu g\,l^{-1}$, $mg\,l^{-1}$ and $mg\,C\,l^{-1}$, respectively). Zero error level is assumed for the model input.

	Error			
	Mean relative, %		Mean absolute	
CPA	B-B NN	N-B NN	B-B NN	N-B NN
chl	±151	±34	$\pm 0.6\,\mu g\,l^{-1}$	$\pm 0.07\,\mu g\,l^{-1}$
sm	±102	±35	$\pm 0.2\,mg\,l^{-1}$	$\pm 0.05\,mg\,l^{-1}$
doc	±293	±36	$\pm 0.2\,mg\,C\,l^{-1}$	$\pm 0.03\,mg\,C\,l^{-1}$

Table 2. Acceptable errors of C_{chl} retrievals for a number of *chl* concentration ranges.

C_{chl} range, $\mu g\,l^{-1}$	0–1	1–2	2–5	5–10	10–20	20–50
Error, %	±50	±30	±30	±30	±20	±20

Table 3. Ranges of C_{chl}, C_{sm}, C_{doc}, for which the *chl* retrieval accuracy specified in table 2 is reached.

Procedure	Measurement error	C_{sm}, mg l^{-1} / C_{doc}, mgC l^{-1}	0–0.5	0.5–1	1–5	5–10	10–20
L-M	λ-independent	0–2	0–50	0–50	0–50	0–50	0–1 2–50
		2–5	0–50	1–50	0–50	0–50	0–50
		5–10	0–50	0–50	0–50	0–50	0–50
		10–20	0–50	1–50	2–50	0–50	2–50
	λ-dependent	0–2	0–5	0–5	1–2	1–2	1–2
		2–5	–	–	0–2	0–2	1–2
		5–10	–	–	–	0–1	0–1
		10–20	–	–	–	–	–
NN	λ-independent	0–2	1–20	0–50	0–50	0–50	2–10
		2–5	0–50	1–50	0–50	0–50	2–10
		5–10	0–2 5–50	0–1 5–50	0–50	0–50	2–10
		10–20	5–50	5–50	0–50	0–50	2–5
	λ-dependent	0–2	–	–	–	–	–
		2–5	–	–	–	–	–
		5–10	–	–	–	–	–
		10–20	–	–	–	–	–

However, the dependence of measured errors on wavelength is inherent in satellite images subjected to the atmospheric correction procedures, as its base is the extrapolation of the aerosol path radiance determined in the near-infrared to the blue spectral region (Land and Haigh 1997, Ruddick *et al*. 2000, for references see also Pozdnyakov *et al*. 2000a, b). It can, nevertheless, be expected that atmospheric correction approaches based on other concepts, such as the inversion method (Doerffer and Fischer 1994), or the methods of statistical regularization (Doubovik *et al*. 1994), multivariate optimization, other neural networks, etc., would result in λ-independent errors. This would broaden considerably the range of hydro-optical conditions, in which the desired accuracy of *chl* retrieval can be reached.

Thus, our numerical experiments indicate that the L–M procedure is capable of solving the CPA retrieval task more successfully than the NN emulator. The only 'bottleneck' of the L–M approach consists in its relatively slow convergence conditioned by the necessity to proceed with a large number of initial concentration vectors $\boldsymbol{C}_0$ (see above)—the drawback, which is not inherent in NNs.

In light of the above, a combination of the L–M and NN procedures might therefore lead to a fast-operating bio-optical algorithm, which would provide a high accuracy of CPA retrievals under conditions of comparably large measurement errors.

In fact, our comparison of the speed of the B-B L–M procedure *per se* and a two-step scheme involving a B-B NN emulator before the B-B L–M (table 4) explicitly indicates that due to an essential narrowing of the range of initial concentrations $\boldsymbol{C}_0$ (attained in the first step by the B-B NN emulator), the L–M algorithm in the second step solves the inverse problem far more rapidly.

The speed of the L–M procedure without a preceding NN narrowing the $\boldsymbol{C_0}$ range significantly depends upon the concrete hydro-optical situation: as the concentration of *sm* decreases, the time required for the solution of the inverse problem (i.e. for

Table 4. A comparison of performance efficiency (in s/pixel) for the B-B L–M procedure and the coupled NN→L–M sequence at a λ-independent 15% measurement error.

Procedure	L–M		NN→L–M		LM/(NN→L–M)	
Noise distribution	uniform	normal	uniform	normal	uniform	normal
$C_{sm}<5\,\mathrm{mg\,l^{-1}}$	0.030	0.011	0.003	0.002	9.3	5.0
$C_{sm}\subset(5\text{–}20)\ \mathrm{mg\,l^{-1}}$	0.121	0.042	0.003	0.002	43.6	21

the retrieval of the CPA concentration vector) also shortens. Understandably, the actual shortening depends entirely upon the power of the computer used. Contrarily, in the case of the above sequence of coupled (NN→L–M) procedures the dependence of the computation time on C_{sm} reduces to nought: it takes only several thousandths of a second to retrieve one CPA concentration vector (or otherwise, to process one pixel in a satellite image).

5. A combined advanced operational algorithm

The principal scheme/flow-diagram of an advanced operational algorithm based upon the coupling between NN and L–M procedures is given in figure 2. Having been subjected to a preliminary processing, also incorporating the atmospheric correction, the satellite image is further treated by calculating for each pixel the corresponding subsurface remote sensing reflectance spectrum $R_{rsw}(\lambda)$ (Pozdnyakov *et al*. 2003a). First, the normalized upwelling radiance ($L_w(\lambda)$) is achieved on the basis of ancillary data, namely instantaneous extraterrestrial solar flux ($F_0(\lambda)$) and Sun zenith angle θ_0.

For each pixel, $R_{rsw}(\lambda)$ serves as input for the B-B NN procedure (NN-1). If one of the CPA concentrations as determined by the B-B NN emulator falls in the concentration range 0–5 (in respective units), a corresponding NB-NN (NN-2, NN-3 or NN-4, see figure 2) is started.

In the next step, the CPA concentration vector $\boldsymbol{C_0}$ thus determined is used to get an array of initial concentrations $\boldsymbol{C}_0$. We have used the range from 0.7 to 1.3 C remembering that the maximal measurement error was set to 15%.

From the created pool of initial concentrations $\boldsymbol{C}_0$, starting concentrations C_i for each CPA are randomly taken as input for the L–M procedure crowning the retrieval process.

If available for the waterbody under investigation, minimum and maximum concentrations of *chl*, *sm* and *doc* are imposed on the pursued multivariate search of the optimal CPA concentration vector (this option is not explicitly shown in figure 2).

At the exit of the coupled NN→L–M sequence the final concentration values of *chl*, *sm* and *doc* emerge. The above retrieval algorithm functions only if a hydro-optical model appropriate for the water body amenable to remote sensing as well as a parametric relationship (e.g. Jerome *et al*. 1996) relating $R_{rsw}(\lambda)$ and the water bulk absorption and backscattering coefficients $\boldsymbol{a}$ and $\boldsymbol{b_b}$ are available.

6. Quality assurance

Our multistage algorithm has two quality assessment units. The first eliminates the pixels suffering from imprecise atmospheric correction. It manifests itself in (a)

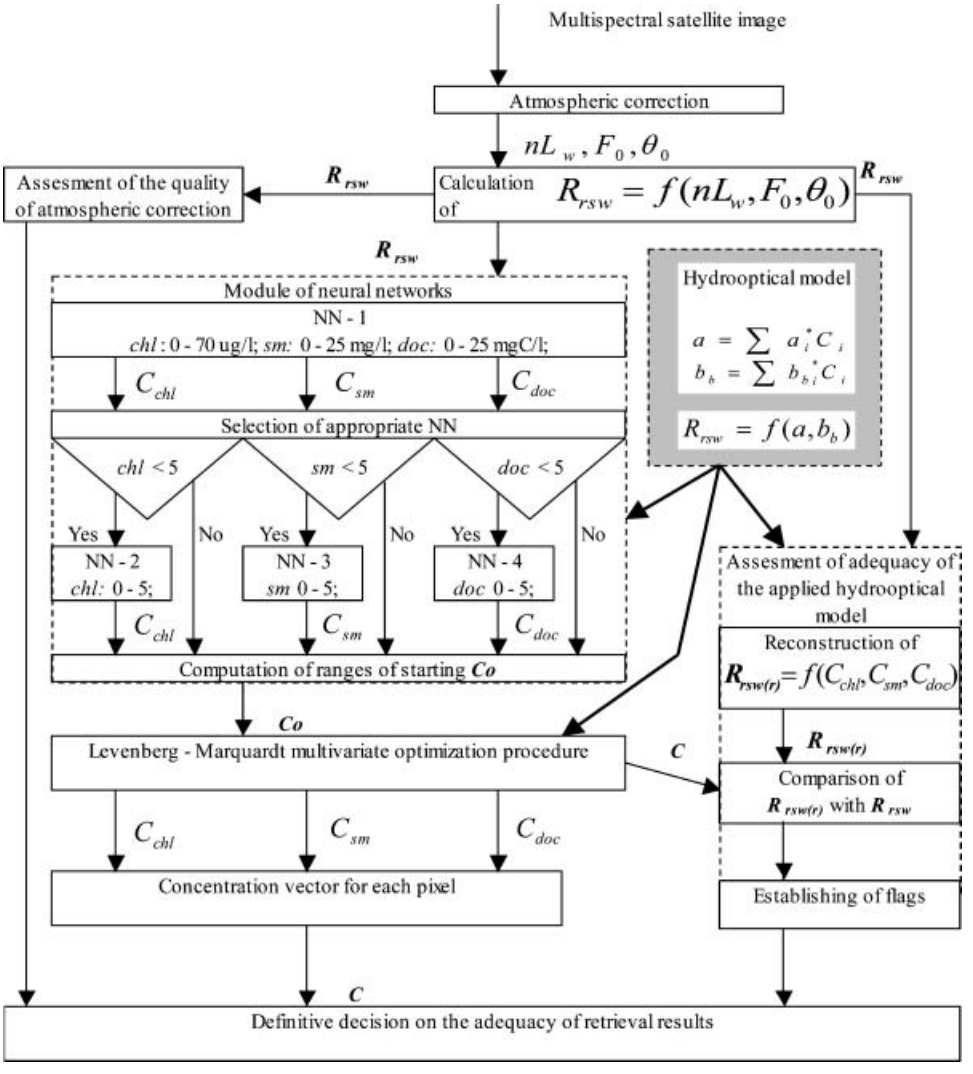

Figure 2. A flow-through diagram for an advanced operational algorithm for the retrieval of CPA concentrations in case II waters.

negative values of R_{rsw} at short wavelengths (400–450 nm) due to an *overestimation* of the path radiance, and (b) enhanced values of R_{rsw} in the blue part of the spectrum (channels 1 and 2 in the case of SeaWiFS) followed by a *dip* in either the second or third channel(s) due to an *underestimation* of the path radiance.

In comparison with case I waters, case II waters are generally loaded with appreciable amounts of *sm*. As a result, the R_{rsw} spectrum peaks at about 560 nm (channel 5 of SeaWiFS), on both sides of which it gradually decreases.

Therefore, we discard pixels with negative R_{rsw} values in the blue. Next we calculate the gradient of the processed R_{rsw} spectrum starting from the first channel (i.e. the shortest wavelength) and moving to the red portion of the spectrum. The pixel is preserved only if the corresponding spectrum exhibits a stable positive gradient at $\lambda < \sim 560$ nm (it is allowed to be just non-negative at $\lambda \leqslant 450$ nm),

followed by a negative gradient for $\lambda > \sim 560$ nm. The pixel is discarded, if along with the long-wave maximum in the R_{rsw} spectrum there is a dip at shorter wavelengths. However, this mask does not discard pixels with the R_{rsw} spectra characteristic of clear case I waters, should they be present in the image. Identification of case I and case II waters is performed through jointly analysing both the envelope of R_{rsw} spectra and their mean spectral value. The latter is known (e.g. Bukata *et al.* 1985) to be distinctly different for clear and turbid waters, and a corresponding threshold is used to this end.

The second quality assurance (mask) is intended to detect pixels corresponding to such areas of the waterbody under surveillance whose hydro-optical properties significantly differ from those covered by the applied hydro-optical model. It might be due to either the area-specific composition of *doc*, and/or *sm* or the presence of some particular phytoplankton species or their residuals (e.g. 'red-tide' algae, coccoliths, etc.), which go beyond the scope of the employed hydro-optical model. Also, based on the literature review we recently discussed elsewhere (Pozdnyakov and Grassl 2003) the possible role of air bubbles in altering water-leaving spectral radiances/water colour due to enhanced backscattering has to be taken into account. The air bubbles are reported to originate in water due to wind action, but also because of gas emissions caused by phytoplankton and possibly bacterioplankton life cycles.

This second mask/flag operates as follows: as soon as the definitive concentration vector $\boldsymbol{C}$ is established for the processed pixel, it is further used for the reconstruction of the corresponding R_{rsw} spectrum using the adopted both hydro-optical model and parametric relationship $R_{rsw}=f(\boldsymbol{a}, \boldsymbol{b_b})$. The reconstructed R_{rsw} spectrum ($R_{rsw(r)}$ in figure 2) is then compared with the initial $R_{rsw}(\lambda)$ spectrum retrieved from the satellite data (R_{rsw} in figure 2). If the mean spectral value of R_{rsw} is significantly higher than $R_{rsw(r)}$, then the pixel can either be discarded or attributed to a coccoliths expression. Coccoliths are known (e.g. Smyth *et al.* 2002) to appear *on* or *just below* the water surface as whitish and hence relatively bright spots surrounded on all sides by dark water. Therefore, the coccolith option comes into action provided the retrieved R_{rsw} spectrum, in addition to its high values for the visible wavelengths, exhibits only a *weak* spectral variance ($\sim\lambda^{-\zeta}$, $1<\zeta<1.35$: see e.g. Balch *et al.* 1996, Smyth *et al.* 2002). The appropriate limits for departures of R_{rsw} (λ) from its mean spectral value are duly imposed.

Application of both quality assurance stages helps to eliminate from the retrieved spatial distributions of CPA the pixels/areas with erroneous concentrations of either *chl* or *sm* and *doc*. This inevitably results in the appearance of lacunas in the retrieved distributions. However, application of the well-known SURFER code (SURFER Brochure 2004) facilitates the visual interpretation of such images by presenting them as a sequence of grey-scaled areas, the borders of which correspond to concentration isolines of the relevant CPA (see below).

When describing the L–M procedure, we mentioned that the use of normalized or absolute differences (equations (1)–(3)) between S_i and R_{rsw} to *simulated data* yields very similar retrieval results. However, this is not the case when the L–M procedure is applied to *life data* from satellite sounding. Indeed, unlike smooth simulated R_{rsw} data, life spectra of R_{rsw} frequently are characterized by dips and/or steep increases/decreases, the nature of these features being variable and often reside in inaccurate atmospheric correction, but not only. The choice between normalized and absolute difference options becomes consequential, when the R_{rsw} spectrum exhibits very low/

nearly zero values in the blue part of the spectrum (e.g. the first one to two SeaWiFS channels), and rather high values in the middle of the visible spectral range. The problem resides in the applied hydro-optical models: at any CPA concentration vector $\boldsymbol{C}$, very low values of water volume reflectance (or remote sensing reflectance) at short wavelengths, λ are not followed by high values of this quantity at longer λ.

When the *relative* difference is applied, the L–M procedure tends to create an R_{rsw}, spectrum, which fits most closely the R_{rsw} spectrum in the first channels, through minimizing the greatest residuals. Understandably (equations (1) and (2)), these are in the blue, given very low values of R_{rsw} spectrum in the first channels. As a result, the L–M procedure ends up in yielding abnormally high values of C_{doc} as the only CPA encompassed by the hydro-optical model, which is capable to crucially reduce the water column reflectance at short wavelengths. Understandably, the retrieved concentrations of *chl* and *sm* prove in this case to be also conspicuously incorrect.

When the absolute difference is applied (equation (3)), the L–M procedure, as a least-square method, tends to equally minimize the residuals at all λ in all channels, and it ends up with creating an $R_{rsw(r)}$, spectrum, which, in comparison with R_{rsw}, is relatively enhanced in the blue, and subdued in the middle of the visible. Conceivably, this is also conducive to the retrieval of unrealistic CPA concentrations. We found that the problem can be successfully resolved if the relative difference is applied to the discussed type of R_{rsw} spectrum, provided the first one or two channels are discarded.

This option is also accommodated in the described algorithm through introducing a respective threshold for R_{rsw} values in the short wavelength region of the spectrum.

7. Verification of the developed algorithm in different hydro-optical environments

In the course of implementation of a number of projects ('WHITESEA'/ INCO-COPERNICUS, 'FINGULF'/INTAS, 'MAQREL'/TASIS, 'MICHIGAN'/ ALTARUM-NIERSC) some small-scale field campaigns have been conducted in the White Sea (The Onega River Bay), the eastern Gulf of Finland, eastern margins of Lake Michigan, and southern Lake Ladoga. *In situ* measurements of *chl*, *sm* and *doc* were carried out synchronously or quasi-synchronously with the SeaWiFS overflights (Pozdnyakov and Grassl 2003, Pozdnyakov *et al*. 2003b).

Tables 5–7 illustrate the comparison of CPA concentrations measured *in situ* and retrieved from SeaWiFS data. When processing SeaWiFS data we used SeaDAS but

Table 5. Comparison of the CPA concentrations measured *in situ* and retrieved from SeaWiFS data for the White Sea, Onega Bay, July 2001.

Date of *in situ* measurements	C_{chl} (*in situ*) ($\mu g l^{-1}$)	C_{doc} (*in situ*) ($mgC l^{-1}$)	C_{sm} (*in situ*) ($mg l^{-1}$)	Date of SeaWiFS overpass	C_{chl} (SeaDAS) ($\mu g l^{-1}$)	C_{chl} (AA) ($\mu g l^{-1}$)	C_{doc} (AA) ($mgC l^{-1}$)	C_{sm} (AA) ($mg l^{-1}$)
10 July 2001	1.6	8.3	0.25	10.07.01	3.7	1.3	6.5	0.8
10 July 2001	1.1	5.8	0.25	10.07.01	3.0	1.2	5.5	0.7
10 July 2001	1.1	5.9	0.65	10.07.01	3.2	1.1	4.5	1.1
10 July 2001	1.5	5.8	0.70	10.07.01	4.0	1.5	4.0	1.0
10 July 2001	1.8	6.3	0.30	10.07.01	4.2	1.6	4.0	1.0
10 July 2001	1.8	5.3	0.10	10.07.01	5.4	1.7	3.9	0.9

Table 6. Comparison of CPA concentrations measured *in situ* and retrieved from SeaWiFS data for the Baltic Sea, eastern Gulf of Finland, August 2002.

Date of *in situ* measurements	C_{chl} (*in situ*) ($\mu g l^{-1}$)	C_{doc} (*in situ*) ($mgC l^{-1}$)	C_{sm} (*in situ*) ($mg l^{-1}$)	Date of SeaWiFS overpass	C_{chl} (SeaDAS) ($\mu g l^{-1}$)	C_{chl} (AA) ($\mu g l^{-1}$)	C_{doc} (AA) ($mgC l^{-1}$)	C_{sm} (AA) ($mg l^{-1}$)
20 August	4.23	10.5	0.8	20.08	10.1	15.0	14.2	0.5
20 August	9.00	12.0	1.1	20.08	8.7	12.0	14.8	2.7
20 August	17.89	13.5	4.5	20.08	8.1	12.1	14.7	3.2
20 August	15.18	14.3	3.1	20.08	8.6	13.5	14.9	1.1
20 August	7.32	9.0	0.9	20.08	9.4	18.6	14.7	1.9
20 August	13.12	8.6	0.9	20.08	7.9	18.0	14.8	2.3
20 August	9.97	8.3	1.0	20.08	7.8	18.0	7.3	1.1
20 August	7.31	8.3	1.2	20.08	6.8	8.1	9.0	1.5
20 August	7.58	7.9	0.8	20.08	8.0	10.0	8.1	1.3
20 August	7.10	7.9	0.7	20.08	7.3	19.0	8.3	1.2
20 August	11.47	25.5	0.8	20.08	5.4	15.0	14.0	1.2

Table 7. Comparison of CPA concentrations measured *in situ* and retrieved from SeaWiFS data for the eastern coast of Lake Michigan, July 2003; TSS—total suspended matter, TOC—total organic carbon.

Station number and coordinates			*In situ* measurements			Retrieved CPA concentrations		
Station number	° N	° W	C_{chl} ($\mu g l^{-1}$)	C_{TSS} ($mg l^{-1}$)	C_{TOC} ($mgC l^{-1}$)	C_{chl} ($\mu g l^{-1}$)	C_{sm} ($mg l^{-1}$)	C_{doc} ($mgC l^{-1}$)
1	42.6706	86.2491	0.72	2.42	24.63	1.11	10.17	23.86
2	42.6662	86.2437	1.20	1.34	22.19	1.85	9.81	24.37
3	42.6668	86.2547	0.45	4.80	6.96	1.85	9.81	7.47
4	42.6671	86.2849	0.50	4.38	15.96	0.46	2.58	10.12
5	42.6699	86.2889	0.46	1.65	37.73	0.44	3.91	24.37
6	42.6700	86.2779	0.53	9.38	4.98	0.50	4.43	7.32
7	42.6617	86.2772	0.63	1.36	8.45	0.57	2.76	7.89
8	42.6614	86.2890	0.55	2.68	18.63	0.42	2.55	11.14
9	42.6791	86.3284	0.54	1.45	7.00	0.20	1.30	4.80
10	42.6833	86.3312	1.56	2.99	9.78	0.13	1.53	5.58
11	42.6832	86.3166	0.63	5.33	9.55	0.08	1.53	5.71
12	42.6749	86.3205	0.71	3.95	7.71	0.12	1.19	4.51
13	42.6761	86.3319	0.51	0.72	24.07	0.21	1.36	15.08

the embedded atmospheric correction code was replaced by the MUMM code developed by Ruddick *et al.* (2000). That was necessary due to the nearly complete inability of the SeaDAS atmospheric correction to grapple with case II waters. The *in situ* measured CPA concentrations are compared in tables 5–7 with the concentrations retrieved both by the SeaDAS bio-optical algorithm and our advanced algorithm (AA) described above.

In analysing tables 5–7, it should be taken into account that each *in situ* measurement relates to a single point (measuring less than 1 m by 1 m) on the water surface, whereas the satellite signal, from which the CPA concerning vector was retrieved, refers to a pixel measuring about 1 km by 1 km. Although the pixel area encompasses the point at which the *in situ* measurement has been performed, this fact nevertheless implies that the retrieved CPA concentration vector is a surface-averaged value for the processed pixel. Understandably, in the case of pronounced

spatial heterogeneity of CPA distributions the departures between the *in situ* measurements and retrieved CPAs are inevitable. In addition, for instance the Lake Michigan truth data incorporates, instead of *doc* and *sm*, TOC and TSS values. This also is bound to result in departures between the retrieved and *in situ* measured quantities. In some cases, there was lack of synchronism between shipborne measurements and satellite overflights. In light of the above, it can be assumed that the shipborne and spaceborne date given in tables 5–7 generally compare quite satisfactorily. Moreover, bearing in mind that the employed hydro-optical models have been specifically developed only for Lake Ladoga and Ontario, and further tentatively used for, respectively, the White Sea, the eastern Gulf of Finland and Lake Michigan, it would hardly be reasonable to expect a closer correspondence between the *in situ* and retrieved data. Tables 5–7 are only intended to illustrate that, given adequate input, our advanced algorithm is capable of successfully dealing with optically complex/case II waters and of providing data simultaneously on three CPAs—*chl*, *sm* and *doc*, whereas SeaDAS (strictly applicable only to case I waters) provides exclusively C*chl*. A poor performance of SeaDAS when applied to case II waters is explicitly indicated in tables 5 and 6: *chl* concentrations retrieved by the SeaDAS algorithm (although improved by using the more adequate atmospheric correction code MUMM) are largely inconsistent with the shipborne measurements of this water constituent.

The performance of the SURFER software package in converting a CPA spatial distribution onto a grey-scaled image is illustrated in figure 3 for Lake Ladoga for the SeaWiFS image of 17 July 2003. The retrieved concentrations of CPAs are in complete compliance with the historical data from Lake Ladoga. In mid-summer, the concentration of *chl* in Lake Ladoga, which is a mesotrophic waterbody, may vary within a wide range 1–25 $\mu g l^{-1}$. The central and northern parts of Lake Ladoga are deep and remain rather cold even in July. As a result, the *chl* concentrations are lower than they are in shallow/well-warmed areas. This is especially applicable to the southern zone.

The lake water is also rich in *doc*: in nearshore regions C_{doc} can be as high as 10–13 $mgC l^{-1}$ (Petrova 1990). But nearshore waters of Lake Ladoga also contain in moderate amounts terrigenous *sm* brought in with river discharge. The largest rivers discharging into the lake are located in the south-eastern area (the lower right-hand part of figure 3), but there is also inflow along the western shore (left-hand side of figure 3).

The central/deep regions of Lake Ladoga generally contain low amounts of *sm* unless temporarily persistent winds and/or meandering currents transport suspended mineral matter offshore (Anonymous 1992). All these features can be found in the CPA distributions given in figure 3, and this can be considered as a qualitative substantiation of the retrieved spatial distributions of *chl*, *sm* and *doc* in Lake Ladoga.

8. Concluding remarks

Our new, more advanced algorithm for the retrieval of natural water quality can generally be utilized for both clear and turbid waters, but it is intended to be used preferably for processing satellite images over case II waters, which present a more serious challenge due to their enhanced optical complexity. For a successful performance of the developed algorithm an accurate atmospheric correction and an

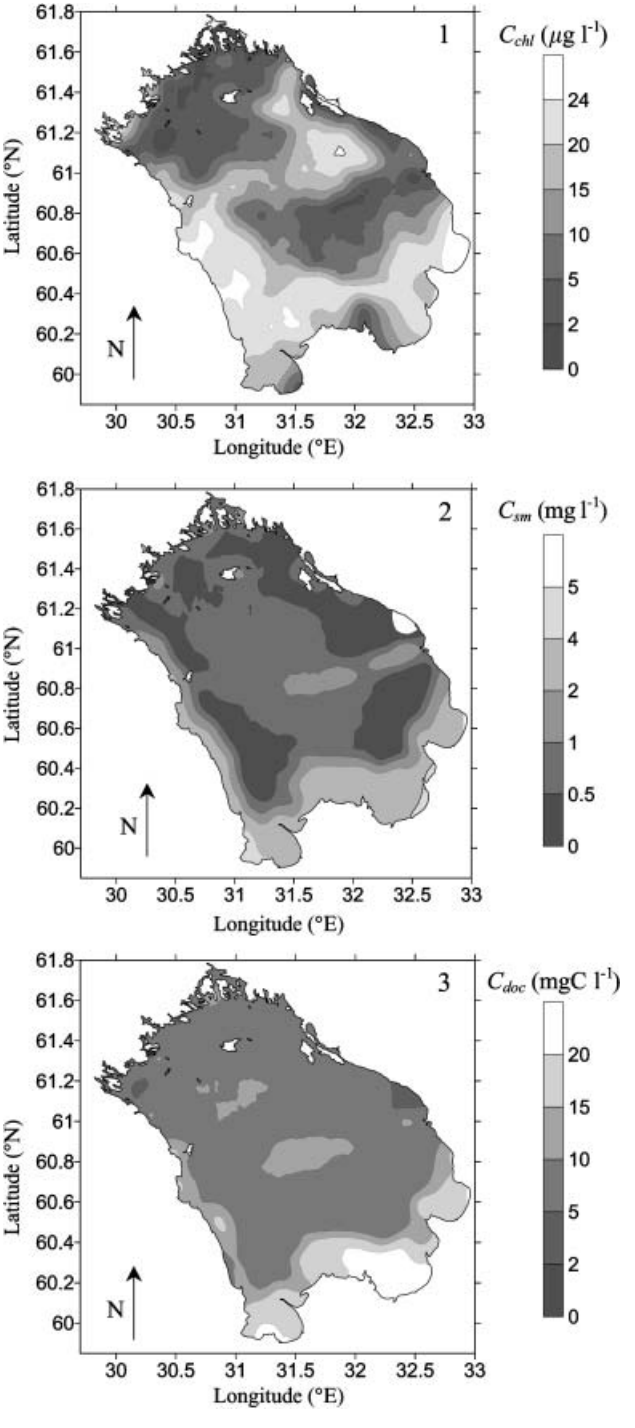

Figure 3. Distribution of *chl* (1), *sm* (2) and *doc* (3) concentrations in Lake Ladoga on 17 July 2003.

adequate hydro-optical model of the waters amenable to remote sensing are required.

Both requirements are often hard to be met. The major problem with atmospheric correction resides in the atmospheric aerosol, whose composition, optical properties and distribution are highly variable in space and time, especially over land areas and coastal zones.

The availability of adequate hydro-optical models for the remotely sensed waterbodies is hampered by many reasons. Phytoplankton communities are complex and their composition is highly variable throughout vegetation seasons. They are very sensitive to the availability of nutrients, light climate and water temperature—the factors which are subject to large spatial/temporal variations themselves.

Due to erosion of the soils in the watershed, suspended minerals are also area-specific (Pozdnyakov and Grassl 2003). In addition, weather conditions (dry or rainy periods) control not only the amounts of *sm* brought in, but also their microphysical properties (e.g. size distribution). The same applies to *doc*: the ratio between allochtonic and autochtonic *doc* is significantly dependent upon weather conditions.

On top, the existing hydro-optical models take account of only three major CPAs, *chl*, *sm* and *doc*, thus ignoring other matters such as bacterioplankton, zooplankton, detritus, and air bubbles. It was also shown (Pozdnyakov and Grassl 2003) that some transspectral processes (fluorescence of *doc* and *chl* in the first place) could be important in forming bulk optical properties of the waters under study. This implies that the existing hydro-optical models should be extended to include such parameters as the fluorescence yield of *chl* and *doc*.

Finally, air bubbles originating from wind action but also produced by phytoplankton and, possibly, bacterioplankton present an unaccounted factor capable of contributing to the bulk backscattering of a water column. Although there are several reports on the optical properties of air bubbles (for references see e.g. Pozdnyakov and Grassl 2003), to our knowledge, no appropriate models have been suggested so far to include air bubbles into the general scheme of photon interactions with the aquatic medium. In addition, bursting of air bubbles just above the water surface (Rossodivita and Andreussi 1999) can constitute an unaccounted contribution to the light scattering in the atmosphere, which might be consequential for a precise assessment of the path radiance, and hence accurate atmospheric correction.

Naturally, the neglect of the above factors has to result in retrieval errors, the importance of which is not necessarily high, but still needs to be considered in concrete circumstances.

Ideally, individual hydro-optical models are required for each waterbody under satellite surveillance, and, so much so, for every vegetation season. This can, probably, be reached for large waterbodies, but certainly not for small ones. In reality, the developed hydro-optical models are tentatively applied to neighbouring waterbodies assuming that they might be akin in terms of their optical properties, which is not necessarily the case.

The quality assurance units complementing the developed algorithm are intended to allay to a certain degree the problems arising from both imperfect atmospheric correction and inadequacy of hydro-optical models. However, their sole function

consists in *eliminating* the wrong information rather than *overcoming* the essence of the problem.

Notwithstanding the above impediments, it could be advocated that the new algorithm can become a reliable basis for further improvement of remote sensing of optically complex natural waters.

Acknowledgments

This work was carried out under the projects supported by INTAS (Project 03-51-4494 'Synergy'), INCO-COPERNUCUS (Project FP6-003605 'EcoMon'). The authors would like to thank Professor N. Filatov (Northern Water Problems Institute, Petrozavodsk, Russia) and Dr R. Schuchman (Altarum, Ann Arbor, MI, USA) for logistic support and fruitful discussions, as well as the Russian R/V 'Ecolog' and Western Michigan University scientists and technicians for water sampling and laboratory analyses. We also extend our gratitude to Mr Are Folkestad (Nansen Environmental and Remote Sensing Centre, Bergen, Norway) for his valuable assistance in training some of the neural networks used in the present study as well as taking part in fruitful discussions.

References

ACKLESON, S., 2001, Ocean optics research at the start of the 21st century. *Oceanography*, **22**, pp. 5–8.

ANONYMOUS, 1992, *Lake Ladoga* (St Petersburg: Nauka Publ.).

ATKINSON, P.M. and TATNALL, A.R.T., 1997, Neural networks in remote sensing. *International Journal of Remote Sensing*, **18**, pp. 699–709.

BALCH, W.M., KILPATRICK, K.A., HOLLIGAN, P., HARBOUR, D. and FERNANDEZ, E., 1996, The 1991 coccolithophore bloom in the central north Atlantic. 2. Relating optics to coccolith concentration. *Limnology and Oceanography*, **41**, pp. 1684–1696.

BUKATA, R.P., JEROME, J.H. and BRUTON, J.E., 1985, Application of direct measurements of optical parameters to the estimation of lake water quality indicators. Environment Canada, Inland Waters Directorate Scientific Series, No. 140.

BUKATA, R.P., JEROME, J.H., KONDRATYEV, K.YA. and POZDNYAKOV, D.V., 1995, *Optical Properties and Remote Sensing of Inland and Coastal Waters* (Boca Raton: CRC Press).

DOERFFER, R. and FISCHER, J., 1994, Concentrations of chlorophyll, suspended matter, and gelbstoff in Case II waters derived from satellite coastal zone color scanner data with inverse modeling methods. *Journal of Geophysical Research*, **99**, pp. 7457–7466.

DOUBOVICK, O.V., OSTCHEPKOV, S.L. and LAPENOCK, T.V., 1994, Iteration/regularization method of solution to nonlinear inverse problems and its application to interpreting water leaving radiance spectra. *Proceedings of the Russian Academy of Sciences: Ocean and Atmospheric Physics*, **30**, pp. 106–113 (in Russian).

JERLOV, N.G., 1976, *Marine Optics*, Elsevier Oceanography Series 14 (Amsterdam: Elsevier).

JEROME, J.H., BUKATA, R.P. and MILLER, J.R., 1996, Remote sensing reflectance and its relationship to optical properties of natural water. *International Journal of Remote Sensing*, **17**, pp. 43–52.

KONDRATYEV, K.YA., POZDNYAKOV, D.V. and ISAKOV, V.JU., 1990, *Radiation–hydrooptical Experiments in Lakes* (Leningrad: Nauka Publ.) (in Russian).

KONDRATYEV, K.YA., FILATOV, N.N., JOHANNESSEN, O.M., MELENTYEV, V.V., POZDNYAKOV, D.V., RYANZHIN, S.V., SHALINA, E.V. and TIKHOMIROV, A.I., 1999, Chapter 3, Modern passive and active optical and microwave remote sensing: advanced feasibilities for application in contemporary limnological stidies. In: *Limnology and Remote Sensing: a contemporary approach*, K.Ya. Kondratyev and N.N. Filatov (Eds), pp. 169–305 (Springer-Praxis: Chichester).

LAND, P.E. and HAIGH, J.D., 1997, Atmospheric correction over case 2 waters using an iterative fitting algorithm including relative humidity. *International Society for Optical Engineering*, **2963**, pp. 609–612 .

LEVENBERG, K., 1944, A method for the solution of certain non-linear problems in least squares. *Quantitative and Applied Mathematic*, **2**, pp. 164–168.

MARQUARDT, D.W., 1963, An algorithm for least-squares estimation of non-linear parameters. *Journal of the International Society of Applied Mathematic*, **11**(2), pp. 36–48.

MOREL, A. and PRIEUR, L., 1977, Analysis of variations in ocean colour. *Limnology and Oceanography*, **22**, pp. 709–722.

O'REILLY, J.E., MARITORENA, S., MITCHELL, B.G., SIEGEL, D.A., CARDER, K.L., GARVER, S.A., KAHRU, M. and MCCLAIN, C., 1998, Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, **103**, pp. 2493–2495.

PETROVA, N.A., 1990, *Phytoplankton Successions with the Anthropogenic Eutrophication of Large Lakes* (Leningrad: Nauka Publ.) (in Russian).

POZDNYAKOV, D. and GRASSL, H., 2003, *Colour of Inland and Coastal Waters: A Methodology for its Interpretation* (Chichester: Springer-Praxis).

POZDNYAKOV, D. and LYASKOVSKY, A., 1999, A comparative analysis of some retrieval algorithms for case II waters. *Earth Observations and Remote Sensing*, **1**, pp. 70–78 (in Russian).

POZDNYAKOV, D., BAKAN, S. and GRASSL, H., 2000a, Atmospheric correction of colour images of case I waters—a review. Report No. 308, pp. 1–33, Max-Planck Institute for Meteorology, Germany.

POZDNYAKOV, D., BAKAN, S. and GRASSL, H., 2000b, Atmospheric correction of colour images of case II waters—a review. Report No. 308, pp. 34–53, Max-Planck Institute for Meteorology, Germany.

POZDNYAKOV, D.V., KOROSOV, A.A. and PETTERSSON, L.H., 2003a, Advanced bio-optical algorithm for case II waters: underpinning methodology, facility assessment, validation and realization exemplifications. *The 30th International Symposium on Remote Sensing of Environment*, 10–14 November 2003, Hawaii, Honolulu (CD-ROM edition), Tucson, US, International Centre for Remote Sensing of Environment.

POZDNYAKOV, D.V., PETTERSSON, L.H., JOHANNESSEN, O.M., LYASKOVSKY, A.V., FILATOV, N.N. and BOBYLEV, L.P., 2003b, SeaWiFS maps water quality parameters of the White Sea. *International Journal of Remote Sensing*, **24**, pp. 3–6.

PRESS, W.H., TEUKOLSKY, S.A., VETTERING, W.T. and FLANNERY, B.P., 1992, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn (New York: Cambridge University Press).

ROSSODIVITA, A. and ANDREUSSI, P., 1999, Spray production by air bubbles bursting on a water surface. *Journal of Geophysical Research*, **104**, pp. 30 059–30 066.

RUDDICK, K.G., OVIDO, F. and RIJKEBOER, M., 2000, Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Applied Optics*, **39**, pp. 897–912.

SATHYENDRANATH, S. (Ed.), 2000, Remote sensing of ocean colour in coastal, and other optically complex waters. Report of the International Ocean-colour Coordinating Group No. 3, IOCCG, Canada.

SMYTH, T.J., MOORE, G.F., GROOM, S.B., LAND, P.E. and TYRRELL, T., 2002, Optical modelling and measurements of a coccolithophore bloom. *Applied Optics*, **41**, pp. 7679–7688.

SNNS (Stuttgart Neural Network Simulator), 1995, User manual, version 3.1. University of Stuttgart, Institute for Parallel and Distributed Performance Systems, ftp://ftp.informatik.uni-stuttgart.de

SURFER Brochure, 2004, ftp://goldensoftwareftp-dl:golden@goldensoftwareftp.com/S8_Brochure.pdf