Papers

Simple statistical formulas for estimating biogeochemical properties of suspended particulate matter in the southern Baltic Sea potentially useful for optical remote sensing applications^{*} doi:10.5697/oc.56-1.007 OCEANOLOGIA, 56 (1), 2014. pp. 7–39.

> © Copyright by Polish Academy of Sciences, Institute of Oceanology, 2014.

KEYWORDS Biogeochemical properties of suspended particulate matter Light absorption and backscattering coefficients Remote-sensing reflectance

Sławomir B. Woźniak

Institute of Oceanology, Polish Academy of Sciences, Powstańców Warszawy 55, 81–712 Sopot, Poland;

e-mail: woznjr@iopan.gda.pl

Received 6 June 2013, revised 8 October 2013, accepted 16 October 2013.

Abstract

Simple statistical formulas for estimating various biogeochemical properties of suspended particulate matter in the southern Baltic Sea are presented in this paper. These include formulas for estimating mass concentrations of suspended particulate matter (SPM), particulate organic matter (POM), particulate organic carbon (POC) and total chlorophyll a (Chl a). Two different approaches have been adopted. The first approach was to use the available empirical material (the results of field measurements and laboratory analyses of discrete water samples) and find statistical formulas for estimating the biogeochemical properties of suspended particulate matter from those of inherent optical properties (IOPs), which are

^{*} Financial support for this research was provided by Statutory Research Programme No. I.1 at the Institute of Oceanology, Polish Academy of Sciences, by research grant No. NN306 2838 33 awarded to Sławomir B. Woźniak by the Polish Ministry of Science and Higher Education, and by the 'SatBałtyk' project funded by the European Union through the European Regional Development Fund (contract No. POIG.01.01.02-22-011/09 entitled 'The Satellite Monitoring of the Baltic Sea Environment').

The complete text of the paper is available at http://www.iopan.gda.pl/oceanologia/

potentially retrievable from remote sensing measurements. The second approach was to find formulas that would enable biogeochemical properties of suspended particulate matter to be estimated directly from spectral values of the remotesensing reflectance R_{rs} . The latter was based on statistical analyses of a synthetic data set of R_{rs} obtained from numerical simulations of radiative transfer for which the available empirical material on seawater IOPs and biogeochemistry served as input data. Among the empirical formulas based on seawater IOPs that could be used as a step in two-stage remote sensing algorithms (the other step is estimating certain IOPs from reflectance), the best error statistics are found for estimates of SPM and POM from the particulate backscattering coefficient b_{bp} in the blue region of light wavelengths (443 nm), and for estimates of POC and Chl a from the coefficient of light absorption by the sum of all non-water (i.e. suspended and dissolved) constituents of seawater a_n , in the blue (443 nm) and green (555 nm) parts of the spectrum respectively. For the semi-empirical formulas under consideration, which could serve as starting points in the development of local onestage (direct) remote sensing algorithms, the best error statistics are found when SPM, POM and POC are estimated from the same blue-to-red band reflectance ratio $(R_{rs}(490)/R_{rs}(645))$ (with estimated SPM reaching a better precision than estimated POM and POC), and when Chl a is estimated from the green-to-red band ratio $(R_{rs}(555)/R_{rs}(645))$.

1. Introduction

As a result of the enormous technological advances of recent decades, remote observations of ocean colour have become an extensively used research tool in contemporary oceanography. By ocean colour we mean spectra of the optical quantity known as remote-sensing reflectance (for definitions of this and other optical quantities used here, see e.g. Mobley (1994)). Observations of ocean colour (remote-sensing reflectance spectrum) can be made from a number of satellites orbiting the Earth (detailed information on the development of optical remote methods for marine exploration can be found, for instance, in a series of reports by the International Ocean-Colour Coordinating Group (see IOCCG (2012)) or earlier reports cited there). Under favourable meteorological conditions (clear skies), satellite measurements allow scientists to obtain very large spatial and temporal scales of observations. This was not achievable with the traditional direct oceanographic methods of investigations conducted either by means of in situ measurements of the physical and chemical properties of seawater or by laboratory analyses of discrete water samples. But the ability to fully utilize the results of remote observations in routine environmental monitoring requires a profound understanding of a chain of complicated relations. Firstly, we need to know how the presence of dissolved and suspended constituents of seawater, possessing different properties and occurring in different concentrations, influences its inherent

optical properties (IOPs), e.g. spectra of the light absorption and backscattering coefficients of seawater. And secondly, we require knowledge of how these IOPs, in certain ambient light field conditions, affect the formation of different apparent optical properties (AOPs), one of which is the spectrum of remote-sensing reflectance. In addition, this chain of relations – biogeochemistry of water constituents vs. seawater IOPs and vs. seawater AOPs – is usually much more complicated in oceanic shelf and coastal regions and in semi-enclosed and enclosed seas (generally belonging to Case 2 waters according to the optical classification introduced by Morel and Prieur (1977)) than in open regions of global oceans (generally belonging to Case 1).

The Baltic Sea is an example of a marine basin classified as Case 2 that possesses a very high degree of optical complexity. In this semi-enclosed and relatively shallow sea we may find a variety of optically significant dissolved and suspended substances of both allo- and autogenic origin, and their concentrations may be uncorrelated with one another. Different aspects of light interaction with Baltic Sea waters have been studied for more than half a century (see e.g. Dera & Woźniak (2010), and the extensive list of works cited there). A lot has been done within that discipline, and the last few years have witnessed an intensification in the development of remote optical methods for Baltic Sea monitoring, among other things, as a result of new multi-institutional scientific projects like the 'SatBałtyk' project conducted in Poland (see e.g. Woźniak et al. (2011a,b)). But in addition to the pursuit of precise and often very complicated descriptions of the relations between biogeochemistry and seawater optics, a difficult task that continues to be of practical importance, there is a need to develop relatively simple yet efficient methods enabling the retrieval of information about the marine environment from remote sensing optical measurements made with currently available devices, technology and knowledge. The search for such simple but effective methods is the subject of the present work. In one of our earlier papers (see S. B. Woźniak et al. (2011)), we reported on the very large variability in relationships between the biogeochemical quantities characterising suspended matter and the in situ measured optical properties of seawater in the southern Baltic Sea on the basis of field measurements and laboratory analyses of data collected in 2006–2009. In the present paper it is intended to take fuller advantage of that earlier empirical material. At the same time, however, this paper also has to be a step towards broadening the practical applicability of remote sensing optical methods for estimating concentrations and properties of suspended particulate matter occurring in the subsurface waters of the southern Baltic Sea. That earlier paper (S. B. Woźniak et al. 2011) gave the first few examples of simple statistical

formulas. But those examples allowed only a rough estimate of certain biogeochemical properties of suspended particulate matter based on optical properties to be obtained as a result of direct in situ measurements. In the current work the search for simple statistical relationships continues, but this time it focuses on making use of remotely sensed optical properties. That is why the main objective of this work has been formulated as follows: to develop (and assess the precision of) a set of simple statistical formulas for estimating the biogeochemical properties of suspended particulate matter in the southern Baltic Sea based on different optical properties of seawater, which are observable/retrievable with current optical remote sensing techniques.

2. Material and methods

Statistical formulas for estimating various biogeochemical properties of suspended particulate matter in the southern Baltic Sea are developed in this paper. They include formulas for estimating mass concentrations of suspended particulate matter (SPM), particulate organic matter (POM), particulate organic carbon (POC) and total chlorophyll a (Chl a). These formulas have to be based on different optical properties of seawater potentially observable or retrievable with remote sensing techniques. Owing to the nature of the empirical dataset available (composed of data on



Figure 1. Block diagram illustrating two different approaches adopted when developing statistical formulas for estimating biogeochemical properties of suspended particulate matter

biogeochemical properties of suspended matter and data on inherent optical properties (IOPs) such as backscattering, absorption and attenuation coefficients of seawater), two different approaches were adopted when developing these formulas (see the block diagram in Figure 1). The first approach was to use only the empirical data available and thus to find strictly empirical statistical relationships for estimating biogeochemical properties of suspended matter from IOPs, which are potentially retrievable from remote sensing measurements of ocean colour (i.e. the backscattering and absorption coefficients of light by seawater at certain light wavelengths). In the second approach, theoretical radiative transfer modelling was additionally incorporated, which enabled the existing empirical dataset to be supplemented with modelled spectra of remote-sensing reflectance. Based on the extended dataset, including both empirical and modelling results, another set of statistical formulas, of a semi-empirical nature, were then found. This enabled the biogeochemical properties of suspended particulate matter to be estimated directly from remote-sensing reflectance values at certain light wavelengths or from reflectance ratios. The methodological details of these procedures are given below.

2.1. Empirical dataset

The empirical dataset on the biogeochemical properties and IOPs of surface seawater available for the purpose of the current work is mostly a selection from the results of field measurements and laboratory analyses of discrete water samples already described in an earlier work (S. B. Woźniak et al. 2011). In the current work, therefore, where appropriate, the empirical methods used are described only briefly; the interested reader will find comprehensive information on the subject in that earlier paper.

The empirical data utilised in this work were gathered at 294 sampling stations during 16 short cruises on board r/v 'Oceania' between August 2006 and September 2009. The study area covered the open waters of the southern Baltic Sea as well as the coastal regions of the Gulf of Gdańsk and the Szczecin Lagoon (the area located roughly between $12^{\circ}38'E$ and $19^{\circ}30'E$, and $53^{\circ}42'N$ and $55^{\circ}38'N$, see Figure 2). At each station the seawater IOPs were measured in situ in the surface layer of seawater (in practice, depending on the sea state, the depth of this layer varied between 1 to 1.5 m), and water samples from that layer were also collected with 20 L Niskin bottles for the laboratory analysis of different biogeochemical properties of suspended matter. The Secchi depth at the sampling stations varied from 1 m to 12 m.

The biogeochemical properties of suspended matter in the surface water samples were characterised in terms of suspended particulate matter concentration (SPM) [g m⁻³] using a standard gravimetric technique,



Figure 2. Location of all the sampling stations (denoted by black and grey dots). The subset of stations for which the radiative transfer numerical simulations were performed is denoted by grey dots

the particulate organic matter concentration (POM) [g m⁻³] using the loss on ignition technique, the particulate organic carbon concentration (POC) [g m⁻³] using a high temperature combustion technique, and the total concentration of chlorophyll *a* (Chl *a*) [mg m⁻³] (defined as the sum of chlorophyll *a*, allomer and epimer, chlorophyllide *a* and phaeophytin *a*) with aid of high performance liquid chromatography (HPLC) (as already mentioned, for more methodological details, see an earlier work by S.B. Woźniak et al. (2011)).

For the purposes of the current work the IOPs of surface sea waters were characterised with data from two standard optical instruments performing in situ measurements in the surface layer of seawater. The first instrument used was a spectral backscattering meter (HydroScat-4; HOBI Labs). This measured values of the volume scattering function at an angle centred at 140° and at four light wavelengths – 420, 488, 550 and 620 nm. These raw values were then used to estimate the backscattering coefficients of light in seawater b_b [m⁻¹] at these four wavelengths, according to the method described in Maffione & Dana (1997) and in Dana & Maffione (2002). A correction for the incomplete recovery of backscattered light in highly attenuating waters (the so-called sigma-correction) was applied in accordance with the instrument User's Manual (HOBI Labs 2008), using data on light absorption and attenuation coefficients measured with another optical instrument. To obtain the backscattering coefficients of particles b_{bp} [m⁻¹], the theoretical backscattering coefficients of pure water were subtracted (according to Morel (1974)). The second optical instrument was a spectral absorption-attenuation meter (AC-9; WET Labs). Equipped

with a 25 cm optical path length, this instrument measured the light absorption and attenuation coefficients of all the non-water (i.e. suspended and dissolved) constituents of seawater, $a_n \, [m^{-1}]$ and $c_n \, [m^{-1}]$ respectively, at nine light wavelengths (412, 440, 488, 510, 532, 555, 650, 676 and 715 nm). Corrections for in situ temperature and salinity effects on the optical properties of water were applied according to Pegau et al. (1997). A correction for the incomplete recovery of the scattered light in the absorption tube of the AC-9 instrument was applied according to Zaneveld et al. (1994) (the so-called proportional method, according to which the measured values for the longest light wavelength (715 nm in the case of our instrument) are assumed to be caused entirely by the unwanted scattering error effect, and the corrected value of absorption at this band was assumed to be 0). At this point, the reader should note an important methodological difference between the current work and the paper of S.B. Woźniak et al. (2011) mentioned earlier. In that paper the light absorption properties of suspended particles and coloured dissolved organic matter were characterised separately, not in situ, but based on measurements of discrete seawater samples performed in a land-based laboratory using a bench-top spectrophotometer. In the current work only the in situ measured (with the AC-9 instrument) total absorption coefficient of all suspended and dissolved non-water constituents of seawater a_n is taken into consideration. It is relatively easy to measure the latter optical coefficient during oceanographic campaigns, so data on coefficient a_n are often present in different oceanographic datasets used for the calibration and validation of remote sensing algorithms.

2.2. Derivation of statistical formulas based on IOPs

To derive statistical formulas for the biogeochemical properties of suspended particulate matter as functions of selected seawater IOPs, described here briefly as empirical formulas, the spectral backscattering coefficients of particles b_{bp} and the absorption coefficients of all the nonwater constituents of seawater a_n at two wavelengths (443 and 555 nm) were taken into account. Values of $b_{bp}(443)$, $b_{bp}(555)$ and $a_n(443)$ were interpolated from those measured in situ with the two above-mentioned optical instruments (to be precise: linear interpolations were performed between the values of $\log(b_{bp}(\lambda))$ and $\log(\lambda)$ and also between the values of $\log(a_n(\lambda))$ and λ). The values of $a_n(555)$ were taken as measured. The values of b_{bp} and a_n (we recall that the latter coefficient is the sum of the absorption coefficient of phytoplankton (a_{ph}) and the absorption coefficient of dissolved and detrital material (a_{dq})) at the 'blue' wavelength

14 S.B. Woźniak

of 443 nm, were, at the moment when this text was written, among the socalled evaluation products accessible with the global Level 3 data browser on the NASA Ocean Color Web page (http://oceancolor.gsfc.nasa.gov) for MODIS Aqua and SeaWiFS sensors (with different variants of these coefficients calculated according to either the Garver-Siegel-Maritorena algorithm (Maritorena et al. (2002), Maritorena & Siegel (2005), Maritorena et al. (2010), the Quasi-Analytical Algorithm (Lee et al. (2002), or the Generalized IOP (GIOP) model (Franz & Werdell (2010)). Since it is well known that optical properties of Baltic Sea waters are often dominated by the presence of relatively high concentrations of coloured dissolved organic matter (CDOM) (with exponential absorption coefficient spectra) (see e.g. Kowalczuk (1999) or Kowalczuk et al. (2005)), it is highly likely that in order to obtain reliable results, the retrieval of IOPs (particularly b_{bp}) at light wavelengths longer than 443 nm may be necessary for at least some of the potential environmental situations encountered in the Baltic Sea. That is why it was decided to analyse the additional 'green' wavelength of 555 nm here. The 555 nm band was available to the SeaWiFS sensor when that was operational, and it is still available to the MODIS Aqua sensor. This means that, at least theoretically, coefficients b_{bp} and a_n for that particular band are potentially retrievable from archival and current satellite mission data. The 555 nm band was also used by Stramski et al. (2008) when those authors were developing their two-step empirical algorithm for POC, and some of the results they obtained will be used here for comparison. Statistical analyses of that empirical material were then performed, and the best-fit power functions approximately representing relationships between the biogeochemical properties of suspended particulate matter and seawater IOPs were found with use of the least square linear regression method applied to the log-transformed variables. It is also worth noting that in the case of every statistical formula presented later, all the available empirical data were used for working it out, which is why the number of samples in the case of each empirical formula is different (from 122 to 253 samples).

In addition to the presentation of IOP-based relationships for the two satellite light wavelengths of 443 and 555 nm, the statistical analyses are supplemented with examples of analogous relationships but determined at the optimal bands chosen from among those original light wavelengths for which the HydroScat-4 and AC-9 instruments performed in situ measurements.

2.3. Derivation of statistical formulas based on remote-sensing reflectance

To derive statistical formulas for biogeochemical properties of suspended matter as functions of remote-sensing reflectance values, the available dataset has to be extended with the aid of radiative transfer modelling. It has been common practice in much optical modelling work that the average values of the constituent-specific optical coefficient multiplied by the assumed concentrations of these constituents give the modelled absolute values of these optical coefficients, which are then used as further inputs for radiative transfer modelling. But because the very large variability of constituent-specific optical coefficients of suspended matter in the southern Baltic Sea were documented in an earlier work by S.B. Woźniak et al. (2011), it was decided not to use averaged values as the modelling input. Instead, a different approach to the problem is taken: in each separate modelling case the real, measured optical coefficients (i.e. the values of the coefficients $a_n(\lambda)$, $c_n(\lambda)$ and $b_{bp}(\lambda)$ are used as modelling input and the corresponding and actually measured values of biogeochemical properties are also used in the subsequent statistical analyses. From the available empirical material a subset of 83 cases was selected (see the stations denoted by grey dots in Figure 2), which consists of only those cases for which all the biogeochemical properties of the relevant particulate matter (i.e. concentrations of SPM, POM, POC and Chl a) and all the seawater IOPs (i.e. values of $a_n(\lambda)$, $c_n(\lambda)$ and $b_b(\lambda)$) required for further modelling were measured at the same time. For this particular data subset, the hypothetical spectra of the remote-sensing reflectance R_{rs} [sr⁻¹] were then determined on the basis of radiative transfer numerical simulations. The Hydrolight-Ecolight 5.0 (Sequoia Scientific, Inc.) model was applied with a set of simplifying assumptions. The modelled hypothetical water bodies were chosen to be infinitely deep, and all the IOPs of the modelled waters were chosen to be constant with depth. This assumption is obviously a significant simplification, but it most likely represents quite well a common situation in the Baltic Sea, where the relatively shallow subsurface layer of water penetrated by sunlight is mixed as a result of wave action and turbulence caused by surface wind stress. Another simplification was the assumption that no inelastic scattering (no Raman scattering, or chlorophyll or CDOM fluorescence) and no internal sources (no bioluminescence) were taken into account. For each modelled case the total absorption coefficient $a(\lambda)$ of the hypothetical water body was taken to be the sum of the absorption coefficient of pure water $a_w(\lambda)$ and the spectrally inter- and extrapolated absorption coefficients $a_n(\lambda)$ measured with the AC-9 instrument, whereas the total scattering coefficient $b(\lambda)$ was taken to be the sum of scattering

coefficients of pure water $b_w(\lambda)$ and scattering coefficient of particles $b_p(\lambda)$ represented by the difference between measured coefficients $c_n(\lambda)$ and $a_n(\lambda)$. The values of $a_w(\lambda)$ and $b_w(\lambda)$ representing pure water were taken from Pope & Fry (1997), Sogandares & Fry (1997), Smith & Baker (1981) and Morel (1974). The backscattering coefficients of water $b_h(\lambda)$ were obtained as a result of the spectral inter- and extrapolation of values measured with the HydroScat-4 instrument. The Fournier-Forand scattering phase functions were also used in the modelling (Fournier & Forand (1994)), and these functions were selected on the basis of the ratio of $b_b(\lambda)/b(\lambda)$. For simplification, the sea surface state was modelled with an assumed low wind speed of 1 m s^{-1} . Clear sky model conditions and a constant solar zenith angle of 30° were also assumed for all cases. With all these assumptions the remote-sensing reflectances just above the sea surface $R_{rs}(\lambda)$ were then modelled for all 83 cases within the spectral range from 400 to 750 nm and with a spectral resolution of 5 nm. However, of these modelled (synthetic) spectra only the values of $R_{rs}(\lambda)$ at five bands (445, 490, 555, 645 and 665 nm) were chosen for further examination (by way of example). The reader should note at this point that the selection of these spectral bands should be treated purely as a demonstration: they are intended to represent in a simplified manner different parts of the visible light spectrum (445 and 490 nm bands represent the indigo/blue region, 555 nm the green region, and 645 and 665 nm the red region). This selection was performed in consideration of the existing spectral bands of the MODIS Aqua instrument currently used by the oceanographic community (note that the so-called level 2 products from that satellite sensor include values of $R_{rs}(\lambda)$ at 443, 488, 555, 645 and 667 nm; see e.g. the documentation available at http://oceancolor.gsfc.nasa.gov). At the same time, when choosing R_{rs} spectral bands for further analyses, it was also important to choose them relatively close to the bands present in the input data for radiative transfer modelling, especially close to the bands of coefficient a_n (we recall that the closest a_n coefficient input bands were 440, 488, 555, 650 and 676 nm). As in the case of the empirical formulas described earlier, statistical analyses of the combined empirical and modelled material were performed. The best-fit power functions representing the relationships between the biogeochemical properties of suspended matter and the remote-sensing reflectances at chosen wavelengths or reflectance ratios were found. As in previous analyses, this was done with the aid of least square linear regression applied to the log-transformed variables. Owing to the adopted method of input data selection the number of samples in the case of each semi-empirical formula presented does not exceed 83.

3. Results and discussion

3.1. Statistical formulas based on IOPs

The empirical formulas found as a result of the analyses are shown in Table 1. This contains 16 best-fit power functions approximating different variants of the relationship between one of the biogeochemical quantities (SPM, POM, POC or Chl *a*) and the backscattering coefficients of particles $b_{bp}(\lambda)$ or absorption coefficient $a_n(\lambda)$ at light wavelengths of either 443 or 555 nm. The quality of these best-fit functions may be assessed with the aid of different statistical parameters, also presented in Table 1. These

Table 1. The best-fit power functions $(y = C_1 x^{C_2})$ between biogeochemical properties of suspended particulate matter and backscattering coefficients of particles b_{bp} or absorption coefficients of all the non-water constituents of seawater at the two light wavelengths (443 and 555 nm) commonly used in optical remote sensing. The coefficients of determination r^2 between the log-transformed variables, mean normalised bias (MNB^{*}), normalised root mean square error (NRMSE^{**}), standard error factor (X^{***}) and number of samples (n) are also given for each fitted function

Relationship	C_1	C_2	r^2	MNB [%]	NRMSE [%]	X	n
SPM vs. $b_{bp}(443)$	60.2	0.827	0.78	6.3	38.5	1.43	154
SPM vs. $b_{bp}(555)$	61.1	0.779	0.77	6.7	40.5	1.44	154
SPM vs. $a_n(443)$	3.25	1.12	0.79	9.5	50.3	1.53	233
SPM vs. $a_n(555)$	13.5	0.876	0.72	12.5	58.9	1.63	233
POM vs. $b_{bp}(443)$	37.6	0.774	0.72	8.8	52.9	1.48	154
POM vs. $b_{bp}(555)$	36.8	0.721	0.7	9.4	55.1	1.5	154
POM vs. $a_n(443)$	2.48	1.04	0.75	10.2	54.3	1.54	233
POM vs. $a_n(555)$	9.37	0.817	0.7	12	56.8	1.61	233
POC vs. $b_{bp}(443)$	13.9	0.779	0.59	15	70.7	1.66	122
POC vs. $b_{bp}(555)$	14.9	0.769	0.61	13.9	65.9	1.65	122
POC vs. $a_n(443)$	0.766	0.971	0.71	12.2	63.2	1.59	162
POC vs. $a_n(555)$	2.74	0.758	0.67	13.3	62.6	1.64	162
Chl a vs. $b_{bp}(443)$	303	0.944	0.64	16.7	75.2	1.74	182
Chl <i>a</i> vs. $b_{bp}(555)$	272	0.864	0.58	20.3	87.7	1.81	182
Chl a vs. $a_n(443)$	10.1	1.17	0.76	13.1	71.7	1.59	253
Chl a vs. $a_n(555)$	50.7	0.975	0.79	10.6	56.8	1.54	253

*MNB = $\frac{1}{n} \sum_{i=1}^{n} \left(\frac{P_i - O_i}{O_i}\right)$, where: P_i – predicted value, O_i – measured value; **NRMSE = $\left(\frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{P_i - O_i}{O_i} - \text{MNB}\right)^2\right)^{\frac{1}{2}}$; *** $X = 10^{\sigma_{\log}}$, where $\sigma_{\log} = \left(\frac{1}{n-1} \sum_{i=1}^{n} \left(\log(P_i/O_i) - \frac{1}{n} \sum_{i=1}^{n} \log(P_i/O_i)\right)^2\right)^{\frac{1}{2}}$.



Figure 3. Examples of empirical relationships between biogeochemical properties of suspended particulate matter and seawater IOPs offering the best statistical fit results from among all the formulas given in Table 1: a) SPM vs. $b_{bp}(443)$; b) POM vs. $b_{bp}(443)$; c) POC vs. $a_n(443)$ and d) Chl *a* vs. $a_n(555)$. The black lines represent the best-fit power functions; the formulas are given in each panel. For error statistics, see Table 1

statistical parameters are as follows: the coefficient of determination r^2 calculated for the log-transformed variables, the mean normalised bias (MNB) and normalised root mean square error (NRMSE) representing the systematic and statistical errors of the so-called linear statistics, and the standard error factor X representing the statistical error of the so-called logarithmic statistics (see the footnote to Table 1 for definitions of these statistical parameters). Note also that the systematic errors of the logarithmic statistics are not listed there as they are always equal to 0 (this is because the presented best-fit power functions were found using least square linear regression applied to log-transformed variables). As can be seen, the statistical parameters listed in Table 1 vary significantly between

the different best-fit formulas. For example, the coefficients of determination r^2 vary between 0.58 and 0.79, while the standard error factors X vary between 1.43 and 1.81. The best error statistics of all the different potential estimation formulas are obtained for the relationship between SPM and b_{bp} at the blue light wavelength of 443 nm (see Table 1 and also Figure 3a):

$$SPM = 60.2(b_{bp}(443))^{0.827}.$$
(1)

This particular formula has, among other statistical parameters, the lowest standard error factor X of 1.43. At the same time, a similar formula representing the relationship between SPM and $b_{bp}(555)$ (see line 2 in Table 1) has only slightly inferior statistical parameters (e.g. in this case the standard error factor X is 1.44). On the other hand, when the bestfit formulas for SPM as a function of a_n are considered (see lines 3 and 4 in Table 1), distinctly worse standard error factors are obtained (i.e. 1.53 and 1.63, for formulas based on $a_n(443)$ and $a_n(555)$ respectively). That is why the formulas based on coefficients b_{bp} , like the formula given by equation (1), rather than other formulas based on coefficient a_n , are suggested as being the best candidates for estimating SPM for the southern Baltic Sea. A similar criterion (i.e. achieving the lowest possible standard error factor X) is also used below when the potential best candidate formulas for estimating other biogeochemical quantities like POM, POC and Chl a concentrations are suggested from among the variants given in Table 1.

For estimating POM, as in the case of SPM, the value of $b_{bp}(443)$ also seems to be the most appropriate from the statistical point of view. The following statistical formula is suggested (see Table 1 and Figure 3b):

$$POM = 37.6(b_{bp}(443))^{0.774}.$$
(2)

The standard error factor X in this case is 1.48, which is not much higher than in the case of the SPM formula given by equation (1). Please note at this point, that for the southern Baltic Sea samples taken into consideration in this work, the variation in the ratio between POM and SPM concentrations was rather limited. As reported by the author earlier (see S. B. Woźniak et al. (2011)), the average value of POM/SPM for southern Baltic samples was about 0.8 and the appropriate coefficient of variation (CV, defined as the ratio of the standard deviation to the average value and expressed as a percentage) of that ratio was only about 22%. This means that in most cases the composition of suspended matter encountered in the southern Baltic is dominated by organic matter. This fact may explain and justify the existence of similarly strong statistical relationships between SPM and b_{bp} , and between POM and b_{bp} .

With regard to the estimation of POC concentrations, it turns out that the statistical results are slightly better when coefficients a_n rather than b_{bp} are used. The following formula for the blue light wavelength of 443 nm (see Table 1 and Figure 3c) gave the best statistical results:

$$POC = 0.766(a_n(443))^{0.971}.$$
(3)

However, the standard error factor X in this case is 1.59, a distinctly higher value than in the case of formulas (1) and (2). Thus it is expected that the quality of the estimates of POC concentrations with formula (3) would in most cases be inferior to that for SPM or POM.

Finally, for estimating Chl a the best statistical results are obtained for the following formula based on coefficient a_n at the green light wavelength of 555 nm (see Table 1 and Figure 3d):

$$\operatorname{Chl} a = 50.7(a_n(555))^{0.975}.$$
(4)

This formula has a standard error factor X of 1.54 (note that this time the other formula based on coefficient a_n at the blue band of 443 nm has a higher standard error factor of 1.59).

All four simplified empirical formulas presented above (equations (1) to (4)) are put forward as the best candidates from among the 16 different statistical formulas listed in Table 1. Obviously, these four formulas offer a potential accuracy that is rather limited and far from perfect – the corresponding standard error factors X lie between 1.43 and 1.59 - soeveryone interested in the potential application of these formulas has to be aware of this. But, at the same time, while assessing the potential usefulness of these and similar simplified formulas, one has to bear in mind that a very large variability exists among the relationships between biogeochemical properties and seawater IOPs in the southern Baltic Sea (as reported earlier in S. B. Woźniak et al. (2011)). In view of this, and also taking into account the fact that concentrations of SPM, POM, POC and Chl a in the southern Baltic may change within a range covering about two orders of magnitude or more, the accuracy offered by the statistical formulas presented here still seems quite reasonable. Additionally, one has to remember that the overall accuracy of procedures or algorithms making use of these simplified statistical relations should be accessed simultaneously when they are combined with other required estimation steps, such as the estimation of coefficients $b_{bp}(\lambda)$ or $a_n(\lambda)$ from remote sensing measurements. In reality it may turn out that formulas among those presented in Table 1 other than the four examples suggested above may ultimately offer the better combined accuracy of estimation.

If one wishes to compare the statistical formulas presented here with similar results from the literature, there is unfortunately not much of a choice. Nevertheless, in some cases at least, the ranges of variations between the optical and biogeochemical properties of suspended particulate matter in the southern Baltic represented by these nonlinear relationships may be compared with the average values and standard deviations of constituent-specific optical coefficients given in the literature by different authors for relatively close light wavelengths and for different marine basins (unfortunately not for the Baltic Sea). For example, the nonlinear relationship obtained in this work between SPM and $b_{bp}(555)$ (which takes the form: SPM = $61.1(b_{bp}(555))^{0.779}$, and is characterised, as we recall, by the standard error factor X = 1.44, see line 2 in Table 1) was obtained on the basis of data for which, if we calculate the average value of the mass-specific backscattering coefficient $b_{bp}^*(555)$ (i.e. coefficient $b_{bp}(555)$ normalised to SPM values), it takes the value of $0.0065(\pm 0.0030)$ m² g⁻¹. The literature value of the mass-specific backscattering coefficient at the relatively close wavelength of 532 nm given by Loisel et al. (2009) (a work cited after Neukermans et al. (2012)) for coastal waters of Cayenne (French Guyana), is very similar – according to these authors $b_{bp}^*(532) = 0.0065(\pm 0.0025)$ $m^2 g^{-1}$. At the same time, according to other results published by Martinez-Vicente et al. (2010) for the western English Channel, the average value of $b_{bp}^*(532)$ may also be distinctly smaller (the average value given by these authors is $0.0034(\pm 0.0008)$ m² g⁻¹). The other relationship that can be indirectly and roughly compared with the literature results is the relationship between Chl a and $b_{bp}(443)$. The formula obtained in this work (which takes the form $\operatorname{Chl} a = 303(b_{bp}(443))^{0.944}$ and is characterised, as we recall, by a relatively high standard error factor X = 1.74 – see line 13 in Table 1) was obtained on the basis of data for which, if we calculate the average value of the chlorophyll-specific backscattering coefficient $b_{bp}^{*(\text{Chl}\,a)}(443)$ (i.e. coefficient $b_{bp}(443)$ normalised to Chl *a* values), it takes the value of $0.0030(\pm 0.0019)$ m² mg⁻¹. When we compare the latter with the literature value of the average chlorophyll-specific backscattering coefficient at the relatively close wavelength of 470 nm given by McKee & Cunningham (2006) for Irish Sea waters (i.e. with the value of $b_{bp}^{*(\text{Chl}a)}(470) = 0.0050(\pm 0.0009) \text{ m}^2 \text{ mg}^{-1}$, the differences are obvious. Such a comparison may suggest that the average efficiency of light backscattering (in the blue part of the spectrum) per unit concentration of chlorophyll a for Baltic Sea suspended matter is about 40% less than for Irish Sea waters.

The only statistical formula from Table 1 that can be compared with literature results in a straightforward way is the formula for estimating POC as a function of $b_{bp}(555)$. This formula, which has only a slightly less attractive standard error factor (X = 1.65) than the formula (3) suggested earlier, takes the following form (see Figure 4):

$$POC = 14.9(b_{bp}(555))^{0.769}.$$
(5)

It can be directly compared with the two linear relationships given by (2008) for the eastern South Pacific and the eastern Stramski et al. Atlantic Oceans (one variant representing all the data of Stramski et al. is $POC = 70.851 b_{bp}(555) - 0.009088$, while another variant for which these authors excluded Chilean upwelling data is $POC = 53.607b_{bp}(555) + 0.002468)$ and also with the linear relationship given by Loisel et al. (2001) for the Mediterranean Sea (POC = $37.75 b_{bp}(555) + 0.0013$) (see the additional dotted and dashed lines in Figure 4). As can be seen for low values of $b_{bp}(555)$, of about 0.005 m⁻¹, both oceanic formulas according to Stramski et al. (2008) would produce estimated average results in relative agreement with those given by formula (5), but for $b_{bp}(555)$ values larger by about one order of magnitude (i.e. values of about 0.05 m^{-1}) there would be a distinct overestimation of POC concentration when compared to the results obtained with the Baltic Sea formula. The linear formula according to Loisel et al. (2001) obtained for the Mediterranean Sea generally stands in



Figure 4. Empirical relationship between POC and $b_{bp}(555)$ for southern Baltic data. The black line represents the best-fit power function obtained in this work (the formula is given on the Figure; see also line 10 in Table 1). Additionally, the two dotted lines represent two variants of the statistical linear relationships given by Stramski et al. (2008) for eastern South Pacific and eastern Atlantic Ocean data, and the dashed line represents the statistical linear relationships given by Loisel et al. (2001) for the Mediterranean Sea (see text for details)

better agreement with formula (5) for the range of $b_{bp}(555)$ values registered in the Baltic Sea, but obviously there are also differences for the low and high values of $b_{bp}(555)$ as a result of the nonlinearity of formula (5).

The above presentation of IOP-based relationships for the two satellite light wavelengths of 443 and 555 nm can be supplemented with examples of similar relationships but determined at the optimal bands chosen directly from among the available empirical material. Let us recall that in the case of the backscattering coefficient b_{bp} the following four wavelengths were available: 420, 488, 550 and 620 nm, whereas for the absorption coefficient a_n it was the following eight wavelengths: 412, 440, 488, 510, 532, 555, 650 and 676 nm. Four examples of optimal wavelength relationships, one for each biogeochemical quantity, are given in Table 2. In the case of SPM and POC estimates, the best results are achieved when values of b_{bp} are used for the wavelength 420 nm (see lines 1 and 2 in Table 2). But the statistical parameters characterising these two new relationships are very similar to those given for the two formulas presented earlier (equations (1) and (2)) which make use of approximated values of $b_{bp}(443)$. No significant improvement is achieved in these two cases (compare the statistical parameters shown in Table 2 and Table 1). A small but noticeable improvement can be found for the statistical relationship between POC and $a_n(488)$ (see line 3 in Table 2, and Figure 5a):

$$POC = 1.35(a_n(488))^{0.923}.$$
(6)

In this case, when we compare it to the equation (3) presented earlier, there is a decrease in the standard error factor X from 1.59 to 1.55. But the largest possible improvement in favour of a formula making use of the optimal wavelength is obtainable (and this is also in agreement with common physical intuition) for a formula for estimating Chl *a* based on values of

Table 2. Examples of the best-fit power functions $(y = C_1 x^{C_2})$ between biogeochemical properties of suspended particulate matter and backscattering coefficients of particles $b_{bp}(\lambda)$ or absorption coefficients of all the non-water constituents of seawater $a_n(\lambda)$ at the 'optimal' light wavelengths chosen from all the wavelengths available from the in situ measurement data (see text for details). The values of the statistical parameters (as in Table 1) are also given for each fitted function

Relationship	C_1	C_2	r^2	MNB [%]	NRMSE $[\%]$	X	n
SPM vs. $b_{bp}(420)$	57.3	0.83	0.78	6.2	37.6	1.43	154
POM vs. $b_{bp}(420)$	36.6	0.781	0.73	8.5	51.7	1.47	154
POC vs. $a_n(488)$	1.35	0.923	0.74	11	59.8	1.55	162
Chl a vs. $a_n(676)$	45.6	0.854	0.9	4.8	34.5	1.35	253



Figure 5. Examples of empirical relationships between the biogeochemical properties of suspended particulate matter and in situ measured seawater IOPs, for which a noticeable improvement of statistical parameters is observed when compared to the example relationships presented in Figure 3 (see text for details): a) POC vs. $a_n(488)$; and b) Chl *a* vs. $a_n(676)$. The black lines represent the best-fit power functions; formulas are given in each panel. For error statistics, see Table 2

 $a_n(676)$, i.e. values at the red peak of that pigment absorption spectrum (see line 4 in Table 2, and Figure 5b):

$$Chl a = 45.6(a_n(676))^{0.854}.$$
(7)

In this case, when we compare the standard error factor X to equation (4) presented earlier, the improvement in its value is the largest (i.e. the value of X decreases from 1.54 to 1.35). But the values of all the statistical parameters obtained in that particular case have to be treated with extra caution. The values of coefficient $a_n(676)$ measured with the AC-9 instrument are spectrally located close to the 715 nm band, at which, according to the absorption measurement correction methodology applied in this work (the so-called proportional method, see the Methods section), the whole of the measured signal was assumed to have been caused by light scattering, and was consequently subtracted to make $a_n(715)$ equal to 0. Although this methodology has been widely used by many oceanographers, it is known to be an imperfect simplification (see e.g. the discussion in the paper by McKee et al. (2008)). In situations where the assumption that absorption by particles in water of the 715 nm band is negligible does not hold, the resultant corrected absorption coefficients a_n could be encumbered with a certain error, especially for bands lying spectrally close to the band used for correction. As a result of this, the corrected values of $a_n(676)$ in our case may represent the height of the 676 nm absorption peak above the true but unknown value of absorption at 715 nm rather than the real

absolute value of absorption at 676 nm. The other fact which should also be taken into account, and is obviously not analysed here, is that apart from the supposed statistical attractiveness of the Chl *a* vs. $a_n(676)$ relationship, the retrieval of $a_n(676)$ coefficients with remote sensing techniques might be rather difficult to achieve in practice.

3.2. Statistical formulas based on remote-sensing reflectance

Now let us move to the additional category of statistical formulas based on reflectance (semi-empirical formulas). Figure 6 presents all 83 modelled (synthetic) spectra of the remote-sensing reflectance $R_{rs}(\lambda)$ obtained in this work, with the five selected spectral bands of 445, 490, 555, 645 and 665 nm marked by the grey dashed lines. The absolute values of reflectance or different reflectance ratios at these selected bands were



Figure 6. Modelled spectra of remote-sensing reflectance. The vertical dashed lines highlight the spectral bands chosen for use in further statistical analyses

Table 3. The best-fit power functions $(y = C_1 x^{C_2})$ between biogeochemical properties of suspended particulate matter and modelled values of remote-sensing reflectances $R_{rs}(\lambda)$. The values of the statistical parameters (as in Table 1) are also given for each fitted function

Relationship	C_1	C_2	r^2	MNB [%]	NRMSE $[\%]$	X	n
SPM vs. $R_{rs}(645)$	865	0.891	0.75	5.9	34.5	1.43	83
SPM vs. $R_{rs}(665)$	1150	0.889	0.73	6.4	35.8	1.45	83
POM vs. $R_{rs}(645)$	319	0.776	0.63	8.6	44.7	1.52	83
POM vs. $R_{rs}(665)$	397	0.77	0.6	9.1	45.9	1.54	83
POC vs. $R_{rs}(645)$	143	0.831	0.51	18.2	76.5	1.77	83

Table 4. The best-fit power functions $(y = C_1 x^{C_2})$ between biogeochemical properties of suspended particulate matter and modelled values of reflectance ratios $R_{rs}(\lambda_i)/R_{rs}(\lambda_j)$. The values of the statistical parameters (as in Table 1) are also given for each fitted function

Relationship	C_1	C_2	r^2	MNB [%]	NRMSE $[\%]$	X	n
SPM vs. $R_{rs}(445)/R_{rs}(645)$	2.32	-1.06	0.85	3.8	28.4	1.32	83
SPM vs. $R_{rs}(445)/R_{rs}(665)$	3.34	-1.07	0.83	4.3	30.3	1.34	83
SPM vs. $R_{rs}(490)/R_{rs}(645)$	3.85	-1.1	0.86	3.6	28.8	1.3	83
SPM vs. $R_{rs}(490)/R_{rs}(665)$	5.7	-1.11	0.85	3.9	30.1	1.31	83
SPM vs. $R_{rs}(555)/R_{rs}(645)$	11.9	-1.57	0.74	6.4	38.1	1.44	83
SPM vs. $R_{rs}(555)/R_{rs}(665)$	21.4	-1.61	0.73	6.9	39.4	1.46	83
SPM vs. $R_{rs}(490)/R_{rs}(555)$	0.613	-2.11	0.67	7.9	40	1.51	83
POM vs. $R_{rs}(445)/R_{rs}(645)$	1.86	-0.97	0.78	4.8	31.5	1.37	83
POM vs. $R_{rs}(445)/R_{rs}(665)$	2.6	-0.973	0.76	5.3	33.2	1.4	83
POM vs. $R_{rs}(490)/R_{rs}(645)$	3.01	-1.03	0.83	4	29.8	1.32	83
POM vs. $R_{rs}(490)/R_{rs}(665)$	4.33	-1.04	0.82	4.3	31.3	1.34	83
POM vs. $R_{rs}(555)/R_{rs}(645)$	8.68	-1.48	0.72	6.4	37.9	1.43	83
POM vs. $R_{rs}(555)/R_{rs}(665)$	15	-1.5	0.7	7	39.6	1.46	83
POM vs. $R_{rs}(490)/R_{rs}(555)$	0.542	-1.96	0.64	7.3	35.4	1.51	83
POC vs. $R_{rs}(445)/R_{rs}(645)$	0.581	-1.06	0.65	13.3	66.8	1.62	83
POC vs. $R_{rs}(445)/R_{rs}(665)$	0.835	-1.06	0.63	13.8	67.6	1.64	83
POC vs. $R_{rs}(490)/R_{rs}(645)$	0.988	-1.13	0.71	11.7	63.4	1.56	83
POC vs. $R_{rs}(490)/R_{rs}(665)$	1.48	-1.14	0.69	12	63.8	1.6	83
POC vs. $R_{rs}(555)/R_{rs}(645)$	3.13	-1.62	0.6	15	69.5	1.67	83
POC vs. $R_{rs}(555)/R_{rs}(665)$	5.69	-1.65	0.58	15.5	69.9	1.69	83
POC vs. $R_{rs}(490)/R_{rs}(555)$	0.148	-2.18	0.55	15.4	69.8	1.73	83
Chl <i>a</i> vs. $R_{rs}(445)/R_{rs}(645)$	8.45	-0.973	0.58	13.5	57.9	1.68	82
Chl <i>a</i> vs. $R_{rs}(445)/R_{rs}(665)$	11.8	-0.969	0.55	14.1	58.5	1.7	82
Chl <i>a</i> vs. $R_{rs}(490)/R_{rs}(645)$	14.4	-1.11	0.71	9.7	49.2	1.54	82
Chl <i>a</i> vs. $R_{rs}(490)/R_{rs}(665)$	21.3	-1.12	0.69	10.2	50	1.56	82
Chl <i>a</i> vs. $R_{rs}(555)/R_{rs}(645)$	58.8	-1.81	0.79	6.6	39.1	1.44	82
Chl <i>a</i> vs. $R_{rs}(555)/R_{rs}(665)$	115	-1.84	0.77	7.5	41.8	1.47	82

the subject of subsequent statistical analyses. Of the many different variants of best-fit power functions approximating relationships between the biogeochemical properties of particulate matter and remote-sensing reflectance or reflectance ratios, only those for which the appropriate coefficient of determination r^2 between the log-transformed variables were > 0.5 are presented here (see Tables 3 and 4). It turned out only five of the statistical relationships making use of absolute values of $R_{rs}(\lambda)$ (one band formulas) fulfilled the above criterion (see Table 3). These five formulas represent the statistical relationships only between SPM, POM and POC concentrations and R_{rs} in the red bands of 645 and 665 nm. No relationship between Chl *a* and the absolute value of R_{rs} at any analysed band was found



Figure 7. An example of a semi-empirical relationship between SPM and $R_{rs}(645)$. The black line represents the best-fit power function; the formula is given in the Figure. For error statistics, see Table 3

satisfactory. Of all the variants presented in Table 3 the best-fit function, which has the lowest standard error factor X of 1.43, is the one representing the SPM vs. $R_{rs}(645)$ relationship (see Figure 7). It takes the following form:

$$SPM = 865(R_{rs}(645))^{0.891}.$$
(8)

Note that for the similar relationship in the other red band of 665 nm, the standard error factor X is only slightly worse and is equal to 1.45 (see the second line in Table 3). For the other biogeochemical properties of suspended matter, i.e. for POM and POC concentrations, the respective standard error factors X are evidently larger (at 1.52 and 1.77; see the third and fifth lines in Table 3).

Distinctly better statistical results are achieved when the next group of semi-empirical formulas is considered. Within the group of formulas based on different reflectance ratios many more of the best-fit power functions fulfilled the criterion of $r^2 > 0.5$. Table 4 lists 27 different variants of statistical relationships. Among them are formulas using blue-to-red, greento-red and blue-to-green reflectance ratios. However, we may infer from the values of the statistical parameters presented in Table 4 that the best results from the statistical point of view are to be expected when the SPM, POM and POC concentrations are estimated from the same blue-to-red band reflectance ratio (i.e. ratio of $R_{rs}(490)/R_{rs}(645)$). The following three formulas were found (see Figure 8a, b and c):

$$SPM = 3.85 (R_{rs}(490)/R_{rs}(645))^{-1.1},$$
(9)

$$POM = 3.01 (R_{rs}(490) / R_{rs}(645))^{-1.03},$$
(10)

$$POC = 0.988 (R_{rs}(490) / R_{rs}(645))^{-1.13}.$$
 (11)

The respective values of the standard error factor X for these three formulas are 1.30, 1.32 and 1.56. Note that these three formulas using another blueto-red ratio of $R_{rs}(490)/R_{rs}(665)$ give similar and only slightly inferior results in terms of statistical parameters (see Table 4). The fact that statistical analyses suggest using the same reflectance ratio for estimating



Figure 8. Examples of semi-empirical relationships between the biogeochemical properties of suspended particulate matter and the remote-sensing reflectance ratios offering the best statistical fit results from among all formulas given in Table 2: a) SPM vs. $R_{rs}(490)/R_{rs}(645)$; b) POM vs. $R_{rs}(490)/R_{rs}(645)$; c) POC vs. $R_{rs}(490)/R_{rs}(645)$; d) Chl *a* vs. $R_{rs}(555)/R_{rs}(645)$. The black lines represent the best-fit power functions; the relevant formulas are given in each panel. For error statistics, see Table 3

SPM, POM and POC (but with a different precision) is worth commenting on. It suggests that in the case of the southern Baltic Sea the SPM concentration seems to be the one biogeochemical quantity most strongly linked to the reflectance ratio. Other quantities, i.e. POM and especially the POC concentration, then seem to be linked rather indirectly to this particular reflectance ratio, thanks to its partial covariation with SPM. This is not surprising since, as already mentioned in an earlier section, the suspended particle populations encountered in the southern Baltic Sea consist primarily of organic matter and a partial covariation between SPM, POM and POC exists (see also S.B. Woźniak et al. (2011)). In case of the 83 southern Baltic samples chosen here as input for radiative transfer modelling, the calculated average POM/SPM and POC/SPM ratios are respectively equal to 0.84 and 0.27, and the corresponding coefficients of variation are relatively small (18% and 35%). In view of this, the fact that we can find three different statistical formulas like formulas (9), (10) and (11) using the same reflectance ratio seems to be justified.

Instead, for estimating the Chl *a* concentration, a different reflectance ratio from the statistical point of view seems to offer the best results. The following formula making use of the green-to-red band ratio was found (see Figure 8d):

$$Chl a = 58.8 (R_{rs}(555)/R_{rs}(645))^{-1.81}.$$
(12)

The standard error factor X in this case is equal to 1.44. Note that a similar formula making use of another red wavelength, i.e. the formula based on the $R_{rs}(555)/R_{rs}(665)$ ratio, offers quite similar and only slightly less attractive statistical parameters (see the last line in Table 4). Note also that unlike the formulas for estimating SPM, POM, and POC, there is no formula using the blue-to-green ratio among the six 6 variants for estimating Chl *a*. Such a formula is not listed in Table 4 because, as mentioned already, different variants of relationships that are statistically too weak, i.e. do not fulfil the criterion of $r^2 > 0.5$, are not presented.

The latter four semi-empirical formulas (equations (9)-(12)) are put forward here as the best candidates from among all the different semiempirical formulas listed in Tables 3 and 4. But let us emphasise once more, that all the semi-empirical formulas presented here are much simplified, based as they are on hypothetical modelled remote-sensing reflectance spectra obtained with many simplifying assumptions. Moreover, it should not be forgotten that these formulas were determined on the basis of analyses of only five bands in the visible light spectrum (i.e. 445, 490, 555, 645 and 665 nm). Regardless of this evident limitation, it seems to be a significant and meaningful result that the formulas found here to be

30 S.B. Woźniak

the most effective clearly demonstrate a potential for retrieving information on suspended matter in the Baltic Sea using the red part of the remotesensing reflectance spectrum. This particular result is in agreement with the conclusion reported much earlier by Siegel and his collaborators (see e.g. Siegel et al. (1994) and the list of works cited there). Those authors showed that for the case of the Baltic Sea one could achieve a distinct improvement in the accuracy of remote sensing algorithms for estimating suspended matter, chlorophyll, and also yellow substance and euphotic depth, with the use of red wavelengths in the reflectance ratios. They proposed various algorithms for the different satellite instruments of that time (i.e. for CZCS, SeaWiFS and (planned at that time) the MERIS instrument) using, among others, the 670 and 710 nm bands in the red part of the light spectrum. Nevertheless, the possibility of using red band reflectance has also been reported for different coastal environments, especially for determining the suspended matter mass concentration. For example, Ahn et al. (2001) suggested using reflectance values in the 625 nm band as optimal for SPM concentration retrieval in coastal regions of the Korean peninsula (the equation they suggested was $\text{SPM} = 647.8(R_{rs}(625))^{0.86}$). The possibility of estimating SPM using Band 1 of the MODIS sensor was also discussed in a few other papers (we recall that MODIS Band 1 is a relatively broad spectral band (620–670 nm), with a nominal centre wavelength of 645 nm, originally not designed for ocean colour applications but rather for detecting land/cloud/aerosols boundaries, and providing data with a spatial resolution of up to 250 m, see e.g. the documentation available at http://oceancolor.gsfc.nasa.gov). Linear relationships for SPM as functions of values obtained for that band were given by Miller & McKee (2004) for data from selected coastal environments of the Gulf of Mexico. by Rodriguez-Guzman & Gilbes-Santaella (2009) for a tropical open bay in western Puerto Rico, and by Wang et al. (2012) for the Bohai Sea of China. In another work, Wong et al. (2007) pointed out the possibility of using different combinations of MODIS sensor bands (among which there was also a Band 1) for data from coastal regions of Hong Kong. But in case of the Baltic Sea data analysed here, the formula (9) using a blue-tored ratio $(R_{rs}(490)/R_{rs}(645))$ seems to be more effective than formula (8), which is based on the absolute reflectance value in the red band $(R_{rs}(645))$.

With regard to the remaining semi-empirical formulas presented in Table 4, the one which could be directly compared with results known from the literature is the formula for estimating POC concentration based on the blue-to-green reflectance ratio $(R_{rs}(490)/R_{rs}(555))$ (see line 21 in Table 4)

$$POC = 0.148 (R_{rs}(490) / R_{rs}(555))^{-2.18}.$$
(13)

This particular formula may be compared with the formula presented in the previously cited paper by Stramski et al. (2008) on relationships between POC and optical properties in the eastern South Pacific and eastern Atlantic Oceans. The authors of that work gave two very similar variants of the POC vs. $R_{rs}(490)/R_{rs}(555)$ relationships, one of which (relating to all the data in Stramski et al., i.e. including the Chilean upwelling stations) took the following form: POC = $0.3083(R_{rs}(490)/R_{rs}(555))^{-1.639}$. The latter formula is plotted in Figure 9 together with formula (13). Such a comparison shows clearly that the formula describing the average oceanic relationship has a less steep slope (compare the constants C_2 : -2.18 with -1.639). As a consequence of that within the range of minimal blue-to-green reflectance values in the analysed Baltic Sea dataset (values of about 0.4) both formulas would predict similar POC concentrations, but within the range of maximum blue-to-green values (here ca. 0.9) the POC concentrations predicted according to the oceanic formula would be about twice as high as those estimated with formula (13). However, while performing such a comparison it has to be borne in mind that formula (13)does not offer very attractive values of statistical parameters: among other things, the standard error factor X is equal to 1.74, which is much higher



Figure 9. Semi-empirical relationship between POC and $R_{rs}(490)/R_{rs}(555)$ for southern Baltic data. The black line represents the best-fit power function obtained in this work (the formula is given in the Figure caption; see also line 15 in Table 3). The dotted line represents the statistical relationship given by Stramski et al. (2008) for eastern South Pacific and eastern Atlantic Ocean data (see text for details)

than the value of X of 1.56 obtained with formula (11), which makes use of the blue-to-red ratio.

With regard to formulas for estimating Chl a, the fact that no single band formula was found to be acceptable for estimating that pigment concentration for the Baltic Sea data analysed here (no such formula is presented in Table 3) is in agreement with one of the conclusions suggested by Bukata et al. (1995), namely, that a reliable estimate of chlorophyll concentration in waters other than Case 1 (other than open ocean regions) most likely cannot result from a single wavelength reflectance relationship. The other important fact is that among the reflectance ratio formulas found here to be acceptable for estimating the Chl a concentration in the southern Baltic Sea (see the last six lines in Table 4) there is also no formula using the classic blue-to-green ratio that would resemble any of the standard remote sensing algorithms commonly used for Case 1 waters. This is in agreement with earlier studies documenting the generally poor performance of standard Chl a satellite algorithms when they were applied to the Baltic Sea environment (see e.g. Darecki & Stramski (2004)). But it has to be pointed out that the few positive observations/arguments presented above are only qualitative in their nature. If one wished to make a quantitative comparison with known literature reflectance ratio based formulas for Chl a given for the Baltic Sea waters, additional tests would have to be performed, and the modelled reflectance data at wavelengths other than those previously selected would have to be exploited. One of the known literature formulas for estimating Chl a obtained for the Baltic Sea environment is the one given by Siegel et al. (1994). It uses the greento-red reflectance ratio (but at wavelengths slightly shifted compared to the wavelengths already analysed in this work) and takes the following form: $\operatorname{Chl} a = 31.05 (R_{rs}(510)/R_{rs}(670))^{-2.115}$. If we used the modelled reflectance spectra obtained in this work, the equivalent formula would take the form $\text{Chl} a = 32.3 (R_{rs}(510)/R_{rs}(670))^{-1.24}$ $(n = 82; r^2 = 0.7; X = 1.54).$ As can be seen in Figure 10a, these two last formulas would agree only in the ranges of the relatively low values of the $R_{rs}(510)/R_{rs}(670)$ ratio (which corresponds to Chl *a* concentrations of the order of 10 mg m⁻³ and higher). For high values of that green-to-red reflectance ratio, the latter formula would predict Chl a values several times higher than the one given by Siegel et al. (1994). The other formula known from the literature is the one from the paper by Darecki et al. (2005). It uses the green-toorange ratio of $R_{rs}(550)/R_{rs}(590)$ and after simple transformation takes the form $\text{Chl} a = 5.47 (R_{rs}(550)/R_{rs}(590))^{-4.681}$. Based on the modelling results obtained in the present work, the equivalent formula using the same reflectance ratio would be $Chl a = 30(R_{rs}(550)/R_{rs}(590))^{-3.33}$ (n = 82;



Figure 10. Two additional examples of the semi-empirical relationship between Chl *a* and reflectance ratios for southern Baltic data: a) Chl *a* vs. $R_{rs}(510)/R_{rs}(670)$; and b) Chl *a* vs. $R_{rs}(550)/R_{rs}(590)$. The black lines represent the best-fit power functions based on modelling results obtained in this work (the formulas are given in the Figure captions). The dotted lines represent the statistical relationships between Chl *a* and respective reflectance ratio given by Siegel et al. (1994) and Darecki et al. (2005) for their Baltic Sea data (see text for details)

 $r^2 = 0.76$; X = 1.48). Figure 10b shows that these last two formulas would exhibit distinct differences. Both formulas are relatively steep functions of the green-to-orange reflectance ratio but for the same values of this, the predicted ranges of Chl *a* would differ by about one order of magnitude. However, in view of the results of the latter comparison, it has to be emphasised that the 590 nm reflectance band taken for that additional test lies relatively far from the modelling input data on the light absorption coefficient $a_n(\lambda)$ (we recall that the nearest a_n input data bands were at 555 and 650 nm). As a consequence, the modelled values of R_{rs} at 590 nm band should be treated with a relatively low level of confidence. Nevertheless, the last two additional quantitative comparisons of the relationships between Chl *a* and different colour ratios should warn the potential user that all the results of the simplified modelling performed here, and in effect, all the semi-empirical (reflectance-based) formulas presented in this work, should be treated as qualitative rather than quantitative.

Finally, let us comment on the comparison of all the statistical parameters obtained here for different variants of both empirical (see Tables 1 and 2) and semi-empirical formulas (see Tables 3 and 4). Although these statistical parameters should not be directly and indiscriminately compared (the datasets utilised to determine the empirical and semiempirical formulas were not the same; the numbers of samples were

34 S.B. Woźniak

distinctly lower for the reflectance-based formulas), a general comparison of their values may suggest that using the R_{rs} ratio instead of IOPs directly for estimating the biogeochemical quantities of suspended matter seems to give better results (compare, for example, the values of the standard error X for formulas (1) to (4) with the corresponding values for formulas (9) to (12)). One might ask whether there is an inconsistency in the fact that the R_{rs} spectra were actually created with the measured IOPs (so theoretically the input was the same)? The answer to such a question is that there is none. The remote sensing reflectance may carry more implicit information on seawater IOPs and, as a consequence, more information on seawater biogeochemistry than a single wavelength value of a particular seawater IOP. It is well known that the remote sensing reflectance is approximately proportional to the ratio of backscattering coefficient of water to the sum of absorption and backscattering of water $(b_b/(a+b_b))$ (see e.g. Gordon et al. (1975)). Therefore R_{rs} implicitly combines information on both the backscattering and absorption properties of seawater. Using the reflectance spectral ratio in statistical analyses means that, on the one hand, the information on the absolute values of R_{rs} is lost, but that on the other, the information from two different wavelengths on seawater backscattering and absorption properties are combined. The simple statistical approach under favourable conditions (i.e. if the proper spectral bands are chosen) may benefit from this.

4. Summary

It is important to stress once again that all the results presented in this work represent a strongly simplified statistical illustration of the complicated relationships between the biogeochemical properties of particulate matter suspended in seawater and its optical properties. But the main aim of applying such a simplified methodology was to make full use of the available empirical material and to try to find a simple and practical, yet acceptably efficient methods for retrieving information from the remote sensing of the optically complicated southern Baltic Sea waters. The examples of empirical formulas (equations (1) to (4) and the others in Table 1), though encumbered by significant statistical errors, can be used to make rough estimates of the biogeochemical properties of suspended particulate matter and can thus also play a role in the derivation of local remote sensing algorithms for the region of southern Baltic Sea. These IOPbased formulas can already (or after small modifications) be used as one step in two-stage remote sensing algorithms (the other step is to estimate certain IOPs, either b_{bp} or a_n , directly from remote-sensing reflectance). The analyses performed here showed that the best error statistics were

found for estimates of SPM and POM from the particulate backscattering coefficient b_{bp} in the blue region of light wavelengths (443 nm), and for estimates of POC and Chl a from the coefficient of light absorption by the sum of all non-water (i.e. suspended and dissolved) constituents of seawater a_n in the blue (443 nm) and green (555 nm) parts of the spectrum respectively. At the same time the additional semi-empirical reflectance based formulas presented here (equations (8) to (12) and the others in Tables 3 and 4), owing to their even more simplified modelling nature, ought to be treated as qualitative examples, suggesting the possibility of using the red part of the of remote-sensing reflectance spectrum for estimating the biogeochemical properties of suspended matter in the environmental conditions of the Baltic Sea. These analyses have shown that the best error statistics are found when SPM, POM and POC are estimated from the same blue-to-red band reflectance ratio $(R_{rs}(490)/R_{rs}(645))$ (with the estimated SPM achieving a better precision than that of POM or POC), and when Chl a is estimated from the green-to-red band ratio $(R_{rs}(555)/R_{rs}(645))$. In spite of the much simplified nature of the semi-empirical formulas presented here, they are potentially good starting points for the derivation of new direct (one-stage) remote sensing algorithms for the southern Baltic Sea. Obviously, all the example formulas presented in this work (both empirical and semi-empirical) should be treated with the necessary caution. Anyone who wishes to apply these formulas has to bear in mind the significant errors which are inevitable, given their simplified statistical nature. It is also important to note that the potential applicability of all these formulas cannot be assessed merely by comparing the standard error factors or the values of other statistical parameters presented in this work. The statistical parameters reported here should be treated merely as a initial guideline in the search for different possible approaches in the development of new algorithms for the remote sensing of the southern Baltic Sea marine environment. The accuracy of estimations of the biogeochemical properties of suspended particulate matter using these (and similar) formulas should be carefully tested, preferably on an independent and sufficiently large When evaluating overall accuracy, one should also take into data set. consideration the effective precision of other potential steps involved, such as the estimation of seawater IOPs, which would then serve as proxies for the biogeochemical properties of suspended matter in a new hypothetical two-stage algorithm.

Acknowledgements

The author thanks his colleagues from IO PAS – Justyna Meler, Barbara Lednicka, Agnieszka Zdun and Joanna Stoń-Egiert – for their help in

collecting the empirical material, Sławomir Sagan for his assistance with the AC-9 instrument measurements and Dorota Burska from the Institute of Oceanography, University of Gdańsk, Poland, for her analysis of the samples for particulate organic carbon.

References

- Ahn Y.-H., Moon J.-E., Gallegos S., 2001, Development of suspended particulate matter algorithms for ocean color remote sensing, Korean J. Remote Sens., 17 (4), 285–295.
- Bukata R. P., Jerome J. H., Kondratyev K. Ya., Pozdnyakov D. V., 1995, Optical properties and remote sensing of inland and coastal waters, CRC Press, Boca Raton, 362 pp.
- Dana D. R., Maffione R. A., 2002, Determining the backward scattering coefficient with fixed-angle backscattering sensors – revisited, Ocean Optics XVI Conf., November 18–22, Santa Fe, New Mexico, 9 pp.
- Darecki M., Kaczmarek S., Olszewski J., 2005, SeaWiFS ocean colour chlorophyll algorithms for the southern Baltic Sea, Int. J. Remote Sens., 26 (2), 247–260, http://dx.doi.org/10.1080/01431160410001720298.
- Darecki M., Stramski D., 2004, An evaluation of MODIS and SeaWiFS biooptical algorithms in the Baltic Sea, Remote Sens. Environ., 89(3), 326–350, http://dx.doi.org/10.1016/j.rse.2003.10.012.
- Dera J., Woźniak B., 2010, Solar radiation in the Baltic Sea, Oceanologia, 52(4), 533–582, http://dx.doi.org/10.5697/oc.52-4.533.
- Fournier G. R., Forand J. L., 1994, Analytic phase function for ocean water, Ocean Optics XII, Proc. SPIE Int. Soc. Opt. Eng., 2258, 194–201.
- Franz B. A., Werdell P. J., 2010, A generalized framework for modeling of inherent optical properties in remote sensing applications, Proc. Ocean Optics 2010, Anchorage, Alaska, USA, 27 September–1 October 2010, 13 pp.
- Gordon H. R., Brown O. B., Jacobs M. M., 1975, Computed relationships between inherent and apparent optical properties of a flat, homogeneous ocean, Appl. Optics, 14 (2), 417–427, http://dx.doi.org/10.1364/AO.14.000417.
- HOBI Labs (Hydro-Optics, Biology & Instrumentation Laboratories, Inc.), 2008, HydroScat-4 spectral backscattering sensor, USER'S MANUAL, Rev. 4., June 15, 2008, 65 pp.
- IOCCG, 2012, Ocean-colour observations from a geostationary orbit, [in:] Reports of the International Ocean-Colour Coordinating Group No. 12, D. Antoine (ed.), IOCCG, Dartmouth, 103 pp.
- Kowalczuk P., 1999, Seasonal variability of yellow substance absorption in the surface layer of the Baltic Sea, J. Geophys. Res., 104(12), 30047–30058, http://dx.doi.org/10.1029/1999JC900198.
- Kowalczuk P., Olszewski J., Darecki M., Kaczmarek S., 2005, Empirical relationships between coloured dissolved organic matter (CDOM) absorption

and apparent optical properties in Baltic Sea waters, Int. J. Remote Sens., 26 (2), 345–370, http://dx.doi.org/10.1080/01431160410001720270.

- Lee Z. P., Carder K. L., Arnone R., 2002, Deriving inherent optical properties from water color: A multi-band quasi-analytical algorithm for optically deep waters, Appl. Optics, 41(27), 5755–5772, http://dx.doi.org/10.1364/AO.41.005755.
- Loisel H., Bosc E., Stramski D., Oubelkheir K., Deschamps P.Y., 2001, Seasonal variability of the backscattering coefficient in the Mediterranean Sea based on Satellite SeaWiFS imagery, Geophys. Res. Lett., 28 (22), 4203–4206, http: //dx.doi.org/10.1029/2001GL013863.
- Loisel H., Meriaux X., Poteau A., Artigas L. F., Lubac B., Gardel A., Cafflaud J., Lesourd S., 2009, Analyze of the inherent optical properties of French Guiana coastal waters for remote sensing applications, J. Coast. Res., 56 (SI), 1532 -1536.
- Maffione R. A., Dana D. R., 1997, Instruments and methods for measuring the backward-scattering coefficient of ocean waters, Appl. Optics, 36 (24), 6057 –6067, http://dx.doi.org/10.1364/AO.36.006057.
- Maritorena S., Hembise Fanton d'Andon O., Mangin A., Siegel D. A., 2010, Merged satellite ocean color data products using a bio-optical model: characteristics, benefits and issues, Remote Sens. Environ., 114 (8), 1791–1804, http://dx.doi. org/10.1016/j.rse.2010.04.002.
- Maritorena S., Siegel D.A., 2005, Consistent merging of satellite ocean color data sets using a bio-optical model, Remote Sens. Environ., 94(4), 429–440, http://dx.doi.org/10.1016/j.rse.2004.08.014.
- Maritorena S., Siegel D. A., Peterson A., 2002, Optimization of a semi-analytical ocean color model for global scale applications, Appl. Optics, 41 (15), 2705 –2714, http://dx.doi.org/10.1364/AO.41.002705.
- Martinez-Vicente V., Land P. E., Tilstone G. H., Widdicombe C., Fishwick J. R., 2010, Particulate scattering and backscattering related to water constituents and seasonal changes in the Western English Channel, J. Plankton Res., 32 (5), 603–619, http://dx.doi.org/10.1093/plankt/fbq013.
- McKee D., Cunningham A., 2006, Identification and characterization of two optical water types in the Irish Sea from in situ inherent optical properties and seawater constituents, Estuar. Coast. Shelf Sci., 68 (1–2), 305–316, http://dx.doi.org/10.1016/j.ecss.2006.02.010.
- McKee D., Piskozub J., Brown I., 2008, Scattering error corrections for in situ absorption and attenuation measurements, Opt. Express, 16 (24), 19480 -19482, http://dx.doi.org/10.1364/OE.16.019480.
- Miller R. L., McKee B. A., 2004, Using MODIS Terra 250 m imagery to map concentrations of total suspended matter in coastal waters, Remote Sens. Environ., 93 (1-2), 259-266, http://dx.doi.org/10.1016/j.rse.2004.07.012.
- Mobley C.D., 1994, *Light and water; radiative transfer in natural waters*, Acad. Press, San Diego, 592 pp.

- Morel A., 1974, Optical properties of pure water and pure sea water, [in:] Optical aspects of oceanography, N.G. Jerlov & E. Steemann Nielsen (eds.), Acad. Press, London, 1–24.
- Morel A., Prieur L., 1977, *Analysis of variations in ocean color*, Limnol. Oceanogr., 22 (4), 709–722, http://dx.doi.org/10.4319/lo.1977.22.4.0709.
- Neukermans G., Loisel H., Meriaux X., Astoreca R., McKee D., 2012, In situ variability of mass-specific beam attenuation and backscattering of marine particles with respect to particle size, density, and composition, Limnol. Oceanogr., 57(1), 124–144, http://dx.doi.org/10.4319/lo.2012.57.1.0124.
- Pegau W. S., Gray D., Zaneveld J. R. V., 1997, Absorption and attenuation of visible and near-infrared light in water: Dependence on temperature and salinity, Appl. Optics, 36 (24), 6035–6046, http://dx.doi.org/10.1364/AO.36.006035.
- Pope R. M., Fry E. S., 1997, Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements, Appl. Optics, 36 (33), 8710–8723, http: //dx.doi.org/10.1364/AO.36.008710.
- Rodriguez-Guzman V., Gilbes-Santaella F., 2009, Using MODIS 250 m imagery to estimate total suspended sediment in a tropical open bay, Int. J. Syst. Appl. Eng. & Dev., 3 (1), 36–44.
- Siegel H., Gerth M., Beckert M., 1994, The variation of optical properties in the Baltic Sea and algorithms for the application of remote sensing data, Ocean Optics XII, Proc. SPIE Int. Soc. Opt. Eng., 2258, 894–905.
- Sogandares F. M., Fry E. S., 1997, Absorption spectrum (340-640 nm) of pure water. I. Photothermal measurements, Appl. Optics, 36 (33), 8699-8709, http: //dx.doi.org/10.1364/AO.36.008699.
- Smith R. C., Baker K. S., 1981, Optical properties of the clearest natural waters (200-800 nm), Appl. Optics, 20 (2), 177–184, http://dx.doi.org/10.1364/AO. 20.000177.
- Stramski D., Reynolds R. A., Babin M., Kaczmarek S., Lewis M. R., Röttgers R., Sciandra A., Stramska M., Twardowski M. S., Franz B. A., Claustre H., 2008, Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic Oceans, Biogeosciences, 5(1), 171–201, http://dx.doi.org/10.5194/bg-5-171-2008.
- Wang L., Zhao D., Yang J., Chen Y., 2012, Retrieval of total suspended matter from MODIS 250 m imagery in the Bohai Sea of China, J. Oceanogr., 68 (5), 719–725, http://dx.doi.org/10.1007/s10872-012-0129-5.
- Wong M. S., Lee K. H., Kim Y. J., Nichol J. E., Li Z., Emerson N., 2007, Modeling of suspended solids and sea surface salinity in Hong Kong using Aqua/MODIS satellite images, Korean J. Remote Sens., 23 (3), 161–169.
- Woźniak B., Bradtke K., Darecki M., Dera J., Dzierzbicka L., Ficek D., Furmańczyk K., Kowalewski M., Krężel A., Majchrowski R., Ostrowska M., Paszkuta M., Stoń-Egiert J., Stramska M., Zapadka T., 2011a, SatBaltyk A Baltic environmental satellite remote sensing system an ongoing Project in Poland. Part 1: Assumptions, scope and operating range, Oceanologia, 53 (4), 897–924, http://dx.doi.org/10.5697/oc.53-4.897.

- Woźniak B., Bradtke K., Darecki M., Dera J., Dudzińska-Nowak J., Dzierzbicka L., Ficek D., Furmańczyk K., Kowalewski M., Krężel A., Majchrowski R., Ostrowska M., Paszkuta M., Stoń-Egiert J., Stramska M., Zapadka T., 2011b, SatBaltyk – A Baltic environmental satellite remote sensing system – an ongoing Project in Poland. Part 2: Practical applicability and preliminary results, Oceanologia, 53 (4), 925–958, http://dx.doi.org/10.5697/oc.53-4.925.
- Woźniak S.B., Meler J., Lednicka B., Zdun A., Stoń-Egiert J., 2011, Inherent optical properties of suspended particulate matter in the southern Baltic Sea, Oceanologia, 53 (3), 691–729, http://dx.doi.org/10.5697/oc.53-3.691.
- Zaneveld J. R. V., Kitchen J. C., Moore C., 1994, The scattering error correction of reflecting-tube absorption meters, Ocean Optics XII, Proc. SPIE Int. Soc. Opt. Eng., 2258, 44–55.