New approaches for surface water quality estimation in Lake Erken, Sweden, using remotely sensed hyperspectral data

HAMED HAMID MUHAMMED
School of Technology and Health STH
Royal Institute of Technology KTH
Alfred Nobels Alle 10, SE-141 52 Huddinge, Stockholm
SWEDEN
Hamed.muhammed@sth.kth.se http://www.sth.kth.se/hamed

Abstract: - This work demonstrates the efficiency of using linear statistical modelling for estimation of concentrations of various substances in lake water using remotely sensed multi- and hyperspectral images together with extensive field measurements collected over Lake Erken in Sweden. A linear relationship was assumed between image data and the corresponding field measurements, and the transformation coefficients were estimated using the least squares method. The resulting coefficients were used to transform new image data into the corresponding substance concentrations. Estimation errors were computed and concentration maps were generated for chlorophyll-*a* and phaeophytine-*a*, suspended particulate organic matter SPOM, suspended particulate inorganic matter SPIM, as well as total suspended particulate matter SPM (SPOM+SPIM). Good correlation was obtained between estimated and measured values. Backward elimination was performed to find the most useful spectral bands for the case study of this work. Descriptive spectral signatures, describing the impact of underlying processes on the spectral characteristics of water, were generated, analysed and also used to predict the corresponding water quality parameters in image data, with the same estimation accuracy as the linear statistical model. Feature vector based analysis FVBA was also employed to generate transformation coefficients that could be used to estimate water quality parameters from image data, also, with the same accuracy as the previous methods. Finally, the impact of performing atmospheric correction was investigated, in addition to applying linear statistical modelling for the purpose of combined atmospheric correction and ground reflectance estimation.

Key-Words: - Remote sensing, Statistical modelling, Water quality, Descriptive spectral signatures

1 Introduction

Fortunately, inland waters (e.g. lakes and water reservoirs) with different water types and qualities show different spectral characteristics that can be remotely sensed, e.g. using satellite- , airborne, ship-mounted or handheld spectroscopy systems. Unique spectral signatures (i.e. spectra) can be generated, assuming that the spectral resolution of acquired data (number and quality of spectral bands) is high enough. Multispectral data has up to 10 spectral bands, while more than 10 bands were recorded in hyperspectral data. An important advantage of satellite- and airborne remote sensing techniques is that it can provide information over the whole region of interest (e.g. a large lake) in a single data set.

Remotely sensed multi- and hyperspectral image data were employed in numerous studies (Lindell et al. 1999 [29]; McCluney 1975 [30]; Philipson et al. 2003 [36]; Östlund et al. 2001 [34]; Strömbeck et al. 2003 [41]; Dekker et al 1993 [7]; Bukata et al. 1991 [3], 1995 [4]; Kutser 1997 [24]; Giardino et al. 2001 [11]; Koponen et al. 2002 [23]; Lathrop et al. 1986 [25], 1990 [26], 1991 [27]; Lillesand et al. 1983 [28]; Olmanson et al. 2000 [33]; Baruah et al. 2000 [2]; Cairns et al. 1997 [5]) which have shown that it was possible to observe optically active substances in inland waters. This is possible because these substances affect and change the spectral properties of water.

Numerous studies addressed various observation, detection, prediction, quantification and analysis of chlorophyll-*a* and phaeophytine-*a* in the investigated water regions (Strong 1974 [42]; Philipson et al. 2003 [36]; Östlund et al. 2001 [34]; Han et al. 1997 [17]; Mittenzewy et al. 1992 [31]; Koponen et al. 2001 [22]; Schalles et al. 1998 [39]; Giardino et al. 2001 [11]; Baruah et al. 2000 [2]; Ibrahim et al. 2004 [20]). Suspended particulate matter SPM, both organic and inorganic, was targeted in many studies utilizing remotely sensed spectral data (Schiebe et al. 1975 [40]; Han 1997

[16]; Curran et al. 1988 [6]; Baruah et al. 2000 [2]; Tassan 1994 [44]; Doerffer et al. 1989 [8]). Dissolved organic matter DOM (or yellow substance) was also studied and retrieved from this type of data (Tassan 1988 [43], 1994 [44]; Davies-Colley et al. 1990), in addition to other interesting water quality parameters that were remotely predicted, such as turbidity, secchi disk depth and surface temperature (Giardino et al. 2001 [11]; Lidell et al. 1985; McCluney 1975 [30]).

Various techniques (including analytical, semi-analytical/semi-empirical and empirical approaches) were developed to estimate water quality parameters from remotely sensed multi- and hyperspectral data. Among the analytical and semi-analytical approaches one can for instance find inverse modelling (Morel 1988 [32]; Sathyendranath et al. 1989; Doerffer 1994 [9]) and bio-optical modelling (Ammenberg et al. 2002 [1]; Strömbeck et al. 2003 [41]). Empirical approaches include linear simple regression (Pena-Martinez et al. 2003 [35]), linear multivariate regression (Koponen et al 2001 [22]) and neural network algorithms (Zhang et al. 2003 [47]; Baruah et al 2000 [2]), as well as statistical approaches, such as principal component analysis PCA (Ibrahim et al 2004 [20]), characteristic vector analysis (Fischer et al. 1986 [10]; Zalloum et al. 1992 [46]; Roger 1996 [38]) and feature vector based analysis FVBA (Hamid Muhammed et al. 2001 [13]).

In this paper, FVBA as well as linear statistical models are developed to predict surface water quality, to generate descriptive spectral signatures for different water quality parameters, and to convert raw upwelling radiance data into ground reflectance data. Furthermore, backward elimination is employed to find the most useful spectral bands for each water quality parameter of interest.

2 Area of Investigation

Lake Erken (Fig. 1a) which is located about 75 kilometers north-east of Stockholm, Sweden, is a moderately eutrophic lake with a surface area of 24 km^2, a mean depth of 9 m, and a maximum depth of 21 m. the lake has a drainage basin of 140 km^2 and a residence time of 7.4 years. The summer thermal stratification usually begins between May and June, and disappears in August or early September. Between mid July and early August, a bloom of the colonial cyanobacterium, Gloeotrichia echinulata, regularly occurs. At such times Gloeotrichia echinulata can account for up to 50% of chlorophyll-*a* in the lake, and this algae rapidly migrates to the surface under periods of low vertical mixing.

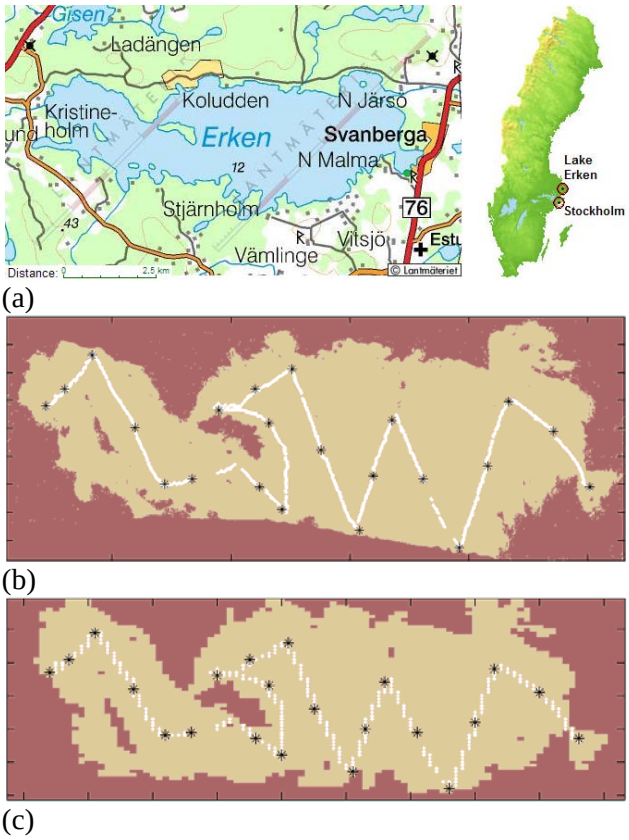

Fig. 1. (a) Map of the research area of Lake Erken, north-east of Stockholm, Sweden. (b) Spatial-mode mosaic-image. (c) Spectral-mode mosaic-image, with overlaid white trace for the continuous measurements and black stars showing the sampling stations' positions in the lake.

In August 1997, there was a bloom of Gloeotrichia echinulata with colonies clearly accumulating at the surface. The bloom was not evenly distributed and higher concentrations were found at the western end of the lake where surface chlorophyll concentrations could exceed 50 μg/l.

3 Image and Ground Truth Data

Hyperspectral images were acquired, by using the Compact Airborne Spectrographic Imager [1] (CASI), over Lake Erken (Fig. 1a) in Sweden, during a CASI campaign in August 1997. Altitude of operation was 3000 m (10,000 ft).

[1] CASI is produced by: ITRES Research Limited, Calgary, Canada.

At approximately the same time of recording these image data, water samples were collected (approximately 30 cm below water surface at 22 sampling stations with known coordinates measured by using a dGPS unit) and analysed in laboratory to measure the concentration of combined chlorophyll-*a* + phaeophytine-*a* (referred to as "Chl-a" in this paper), suspended particulate organic matter, SPOM, suspended particulate inorganic matter, SPIM, and suspended particulate matter, SPM, in these samples. Table 1 presents the ranges of these measured values. A Dual GER 1500 spectroradiometer [2] was used to measure the up- and downwelling radiance above the lake surface at all sampling stations, and to finally compute the reflectance spectra at these stations. The measured spectra had 513 spectral bands in the wavelength range 400 – 900 nm.

Table 1. Ground truth measurements.

		Chl-a & Phaeo-a µg/l	SPM mg/l	SPOM mg/l	SPIM mg/l
22 Sampling stations measurements	Min	2.9	1.7	1.3	0
	Max	50.6	7.3	6.9	1
1722 Continuous measurements	Min	1.9	1.25	1.2	0
	Max	75.5	11.11	10.3	0.8

In addition to that, a dataset of 1722 continuous field measurements were acquired. This dataset included concentrations of Chl-a, SPM, SPOM and SPIM, sampled with 2 second interval. The measurements were performed on a boat which scanned the lake in a zigzag manner, as shown in Figures (1b) and (1c). These figures present spatial-mode respectively spectral-mode mosaic-images generated by putting together spatially neighbouring CASI-scan lines. The white trace in Figures (1b) and (1c) illustrates the continuous measurements while the sampling stations are marked with black stars. The ranges of the values of these measurements are listed in Table 1.

Two spatial mode images, referred to as "Image 1" and "Image 2", were acquired on August 6th and 7th, respectively, during the CASI campaign in 1997. Each of these images had 14 spectral bands and 850×2650 pixels, where each pixel covered 4×4 meters of the lake surface. Table 2 presents the band specifications of CASI spatial mode. On the other hand, only one spectral-mode mosaic-image, referred to as "Image 3", was obtained. It had 288 spectral bands and 61×110 pixels, where each pixel covered about 57×105 meters of the lake surface. These bands covered an approximate wavelength range of 400-900 nm.

Table 2. Wavelength band-set definition for the CASI in spatial mode.

Band No	Start wavelength [nm]	End wavelength [nm]
B1	403.5	415.6
B2	436.5	446.9
B3	483.7	494.2
B4	504.8	515.3
B5	545.3	554.2
B6	555.9	564.8
B7	614.5	625.2
B8	659.0	669.8
B9	676.9	684.1
B10	700.2	709.1
B11	750.3	755.7
B12	761.1	766.5
B13	768.3	780.8
B14	859.9	868.9

4 Image Pre-processing

The CASI data were geometrically corrected and radiometrically calibrated at delivery [3]. These data were then atmospherically corrected by using the 6S-code (Vermote et al. 1997 [45]), which compensated for the atmospheric effects and converted the data (which represent upwelling radiance at the sensor) into ground reflectance. Edge correction was employed to compensate for the across-track variations in brightness (considerably higher brightness-values were observed near the edges of the CASI-scan line).

Thereafter, water areas in the imaged regions were identified and extracted by the following method. The mean images were computed for the

[2] The GER 1500 spectroradiometer is produced by: Geophysical & Environmental Research Corperation, New York, USA, while the dual GER was built at Uppsala University.

[3] The CASI data were provided by Borstad Associates Ltd., which was also the CASI operator.

three hyperspectral images: “Image 1”, “Image 2” and “Image 3”. Finally, global thresholding was applied to each of these mean images to find and then extract the water-region hyperspectral pixels from the corresponding images.

5 Methodology

The aim of the methods presented in this section is to generate transformation coefficients that can be used for surface water quality parameter retrieval from multi- and hyperspectral image data.

Let $\boldsymbol{S}$ be the spectral-data matrix containing spectral vectors on its columns (i.e. each column represents a measured spectrum – e.g. a hyperspectral pixel from a CASI image). The transformation into water quality parameters (forming column vector $\boldsymbol{p}$) is performed as follows

$$\boldsymbol{p}^{\mathrm{T}} = a + b\,\boldsymbol{t}^{\mathrm{T}}\,\boldsymbol{S} \qquad (1)$$

where $(\bullet)^{\mathrm{T}}$ denotes matrix transpose, a and b are scalars, while $\boldsymbol{t}$ is a transformation column vector computed by one of the approaches presented in this section (Linear statistical modelling, descriptive spectral signatures, or feature vector based analysis FVBA). Vector $\boldsymbol{t}$ is estimated by using a training data set consisting of measurements performed on water samples to obtain spectral vectors (arranged as columns in a matrix $\boldsymbol{S}_t$) and the corresponding water quality parameter values (forming a column vector $\boldsymbol{p}_t$).

Thereafter, linear regression, between $\boldsymbol{p}_t$ and the result of computing $(\boldsymbol{t}^{\mathrm{T}}\ \boldsymbol{S}_{t\text{-}img})$, is employed to estimate the values of the scalars a and b to be used in eq. (1), where $\boldsymbol{S}_{t\text{-}img}$ is the spectral matrix, extracted from the image, which corresponds to $\boldsymbol{S}_t$. How each of the three approaches estimates the vector $\boldsymbol{t}$ is explained in the following subsections.

5.1 Linear statistical modelling

In this approach, vector $\boldsymbol{t}$ is estimated by solving the linear system of equations

$$\boldsymbol{p}_t^{\mathrm{T}} = \boldsymbol{t}^{\mathrm{T}}\ \boldsymbol{S}_t \qquad (2)$$

The least squares method can, for instance, be used to estimate $\boldsymbol{t}$. Applying this approach is equivalent to using linear multivariate regression.

5.2 Descriptive spectral signatures

In Hamid Muhammed et al. (2003) [14] and Hamid Muhammed (2005) [15], these signatures were computed and employed to analyse the impact of the parameter of interest on the spectral properties of the target object.

The training spectral data $\boldsymbol{S}_t$ were, at first, normalised into zero-mean and unit-variance data (i.e. whitened), by employing two iterative normalisation approaches, where a number of alternating pixel-wise (Pw) and band-wise (Bw) whitening operations were performed. In Pw-whitening, each multi- or hyperspectral pixel vector was whitened, while each spectral image band was whitened when Bw-whitening was performed. In the first iterative approach, a series of alternating Pw- and Bw-whitening operations, beginning and ending with Pw-operations, were performed. On the other hand, the second iterative approach started with Bw-whitening and ended with a Pw-operation.

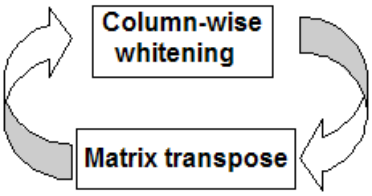

Fig. 2. Iterative normalisation.

Figure 2 illustrates this iterative procedure, where spectral data are inserted as columns or rows (in the case of performing the first respectively the second approach) in matrix $\boldsymbol{S}_t$ on which a series of alternating column-wise whitening and matrix transposing operations are performed. In each approach, a unique stationary result (matrices $\boldsymbol{S}_{t1}$ and $\boldsymbol{S}_{t2}$ result from the first and the second approach, respectively) is achieved after a limited number of iterations.

Linear relationship is assumed between vector $\boldsymbol{p}_t$ and the resulting $\boldsymbol{S}_{t1}$ or $\boldsymbol{S}_{t2}$ matrices from the first or the second normalisation approach, respectively. The systems of linear equations can be written as follows

$$\boldsymbol{S}_{t1}^{\mathrm{T}}\ \boldsymbol{t}_1 = \boldsymbol{p}_t \qquad (3)$$
$$\boldsymbol{S}_{t2}^{\mathrm{T}}\ \boldsymbol{t}_2 = \boldsymbol{p}_t \qquad (4)$$

where the vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$, that can be computed by using the least squares method, function as spectral signatures describing the changes in spectral characteristics with respect to the parameters of interest, which are, in this paper, a number of water quality parameters. A descriptive spectral signature pair, $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$, should be computed for each of the parameters of interest.

The required $\boldsymbol{t}$ vector to be used in eq. (1) can simply be either $\boldsymbol{t}_1$ or $\boldsymbol{t}_2$, but since each of these

vectors describes a different part of the dynamics of the spectral properties with respect to the considered parameter, it should be more efficient if a combination of t_1 and t_2 was used. Good candidates for t can be (t_1 + t_2) or (t_1 * t_2), where + and * are element wise addition and multiplication, respectively.

5.3 Feature vector based analysis (FVBA)

A recently developed technique called Feature-Vector Based Analysis (FVBA) was used for mapping the relative concentrations of the prime or desired substances in inland waters (Hamid Muhammed et al. 2001 [13]). This technique was based on using linear transformation methods, such as Principal Component Analysis (PCA) (Jolliffe 1986 [21]) and Independent Component Analysis (ICA) (Hyvärinen 1999 [18]; Hyvärinen and Oja 2000 [19]), to obtain a number of pairs of basis vectors and the projections of the hyperspectral data on these vectors.

These results were considered as *Component-Feature Vector* pairs, as the FVBA-task took over. In the case of hyperspectral image data, the *Feature Vectors* represent spectral signatures while the *Components* are maps showing the correlation between the corresponding feature vectors and hyperspectral image data. Thereafter, linear combinations of the spectral feature vectors were taken to construct spectral signatures that were as close as possible to selected spectral signatures, obtained by laboratory measurements, of the desired substances (such as the absorption spectra of chlorophyll-a and chlorophyll-b, shown in Fig. 3, depicted from Robinson 1983 [37]).

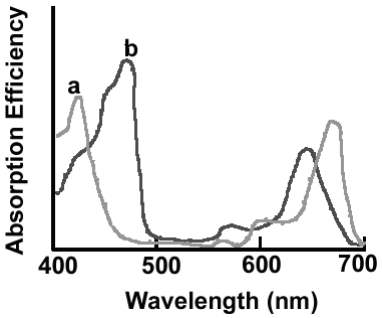

Fig. 3. Absorption spectra of chlorophyll-a and chlorophyll-b. Reproduced from Robinson (1983) [37].

Projecting the hyperspectral image data on these spectral signatures (after shifting the values to get rid of negative elements) produced maps representing the relative concentration maps of the desired substances in the area of investigation. This means that these spectral signatures represent good estimates for the required vector t in eq. (1).

In this work, ICA was used to produce the required *Component-Feature Vector* pairs, since it could produce more easy-to-use feature vectors (according to Hamid Muhammed et al. 2001 [13]), and the computationally efficient FastICA algorithm was selected among many other available ICA algorithms. The FastICA algorithm was downloaded from the Neural Networks Research Centre, Laboratory of Computer and Information Science (CIS), Helsinki University of Technology, Finland.

The desired linear combinations of the feature vectors were found by employing a genetic optimisation algorithm, which simply tried a number of linear combinations chosen according to a mechanism imitating biological genetic evolution (Goldberg 1989 [12]). The used coefficients were limited to the set {-1, 0, +1} leading to considerable simplification of the genetic algorithm, in addition to simulating a more realistic approach. The linear combination, with the least deviation from the target spectrum, was selected to be used as the desired vector t.

5.4 Backward Elimination

This method was applied to Image 3 that had 288 narrow spectral bands, in order to find out which of these bands could be excluded and still get the same or maybe better parameter-prediction results, when using the methods presented previously. The method started by including all narrow bands in the used model and then removing those bands that were not useful, one by one. Two band removal approaches were used. The first one started from band 1, continued towards band 288, and then went back to band 1. The second one started from band 288, towards band 1, and went back to band 288.

After reviewing the retained narrow spectral bands, further simplifications were performed so that broad spectral bands were formed of nearby narrow bands in wavelength regions, where the majority of the narrow bands were retained.

By this way, two respectively three broad spectral bands were obtained (by taking the mean values of each narrow-bands group) when using the first and the second backward elimination approaches, respectively, as illustrated by Fig. 4, where broad bands *a* and *d* resulted from the first backward elimination approach, while broad bands *a*, *b* and *c* were obtained when applying the second approach. Broad band *a* was computed as the mean value of narrow bands 66-108 (covering the wavelengths 517-591.5 nm), broad band *b* was the average of narrow bands 203-206 (761.7-766.4 nm),

broad band *c* combined narrow bands 229-266 (807.6-874.4 nm), and broad band *d* covered narrow bands 232-288 (813.5-914.6 nm).

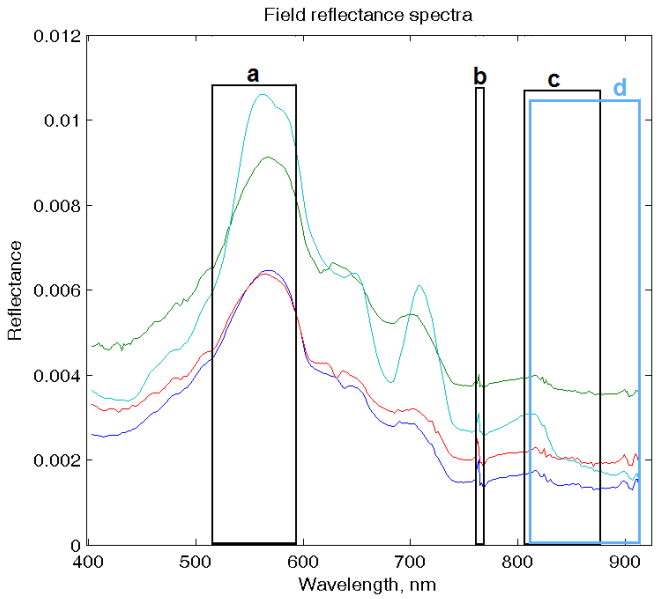

Fig. 4. The backward elimination results when using the first approach (producing broad bands *a* and *d*), and the second approach (producing broad bands *a*, *b* and *c*). Four randomly selected field reflectance spectra were also plotted for comparison.

Approximately the same performance was achieved when using the *a*+*d* or *a*+*b*+*c* broad band sets, leading to the conclusion that the *a*+*d* broad band set was a better and more cost-effective choice to use, since band *a* represented a green band while band *d* represented a near infrared band. Band *b*, on the other hand, was much narrower than *a* and *d* and more expensive and difficult to achieve physically, by using an optical filter.

Note: *Tables 3-14 are presented in* ***Appendix A*** *and Figures 5-22 are presented in* ***Appendix B****. These tables and figures are needed to be able to understand the subject and reproduce the results achieved in this work easily.*

6 Experimental Results and discussion

Extensive experiments were performed to evaluate the performance of the approaches proposed in this work, using three CASI images of Lake Erken (Image 1 and Image 2 with 14 spectral bands and 850×2650 pixels, and Image 3 with 288 bands and 61×110 pixels), a set of measurements from 22 sampling stations (of 513-bands reflectance spectra and the corresponding Chl-a, SPM, SPOM and SPIM concentrations), as well as a set of continuous measurements of the concentrations of Chl-a, SPM, SPOM and SPIM.

The more accurate measurements, at the 22 sampling stations, were used as training data in all experiments with the exception that field reflectance spectra (referred to as FRS) were replaced by image (reflectance) spectra (referred to as IS or IRS) extracted from the image-pixels that correspond to the sampling stations in the image of the lake. Note that Image 1 and Image 2 (both have 14 bands) were atmospherically corrected and consequently contain ground-reflectance spectra (referred to as IRS), while two versions of Image 3 (288 bands) were available, one before and one after atmospheric correction (referred to as Image 3-Raw and Image 3-AtmCorr, respectively), providing upwelling radiance (referred to as image spectra IS) and ground reflectance (referred to as image reflectance spectra IRS), respectively.

Reduced band set images, consisting of 14 and 2 bands, were generated from each of the two versions of Image 3. The simulated 14-bands images (referred to as Image 3-Raw-14-bands and Image 3-AtmCorr-14-bands) followed the same band settings listed in Table 2, while the 2-bands images (referred to as Image 3-Raw-2-bands and Image 3-AtmCorr-2-bands) were constructed according to the first backward elimination approach (consisting of broad bands *a* and *d*, as explained in Fig. 4).

The left column of Figure 5 shows four graphs each of which presents four field reflectance spectra (FRS), measured with the Dual GER 1500 spectroradiometer, then resampled into 288 spectral bands according to the band specifications of Image 3. These spectra correspond to various levels of concentrations of Chl-a, SPM, SPOM and SPIM in the graphs on rows 1, 2, 3 and 4, respectively. On each row in the right column, the corresponding simulated 14-bands FRS, generated according to the band settings found in Table 2, are presented.

Figure 6 presents graphs with four image reflectance spectra (IRS) corresponding to various concentrations of Chl-a, SPM, SPOM and SPIM, on rows 1, 2, 3 and 4, respectively. IRS extracted from Image 1 and Image 2 are presented in the left and right columns, respectively.

Figures 7 and 8 have the same structure as Fig. 5, but show image spectra (IS) respectively image reflectance spectra (IRS), extracted from Image 3-Raw and Image 3-Raw-14-bands, respectively, Image 3-AtmCorr and Image 3-AtmCorr-14-bands.

To start with, estimation results from all approaches were compared with corresponding field measurements and common shifts between these values were found, despite the small time lag

between field measurements and CASI image data, which was, in the case of Image 1, less than 15 minutes in the western part of the lake where the measurements started. The most likely reason for this observation could simply be bad geometric correction of image data due to large errors in the GPS measurements. Hence, the different CASI-scan lines were badly aligned when building a mosaic image covering the whole lake. Distortions in the correspondence between estimated and measured values were clearly observed when the boat went forth and back between the two main CASI-scan lines forming Image 1; e.g., roughly, in the case of Image 1 (see Figure 15), samples 1-100, 340-560, 720-1722 belong to the scan line in the middle, while the rest belong to the upper scan line. The bottom scan line (in Image 1) also contributes to some distortion as some samples belong to it.

Therefore, simple matching was performed between estimated and measured values, to achieve more fair comparison results. For example, in the case of Image 1, the matching was performed by roughly discarding the first 100 estimated and measured values, shifting estimated values 400-1722 to the right with respect to the measured values, and discarding the measured values around 400. The graphs, in the left and right columns of Fig. 11, show comparisons between estimated and measured values, before and after matching, respectively.

First of all, the concentrations of Chl-a, SPM, SPOM and SPIM were estimated using spectral band 10 (in the 14-bands images), as suggested by Östlund et al. (2001) [34]. Table 3 presents the results when applying this approach to the 14-bands images: Image 1, Image 2, Image 3-AtmCorr-14-bands and Image 3-Raw-14-bands. This approach is referred to as the band-10 approach.

The estimation results of the band-10 approach were much better than those of the bio-optical modelling approach presented by Ammenberg et al. (2002) [1] and Philipson et al. (2003) [36]. Note that the correlation coefficient R, not the coefficient of determination R^2, was used as an evaluation measure in these two papers. Applying this approach to CASI data acquired over Lake Malaren (near Stockholm, Sweden) resulted in R^2=78% and R^2=69% in the case of estimating Chl-a and SPIM, respectively. The used model was derived using a large amount of historical measurements concerning Lake Malaren. The results were much worse when applying this model to the Lake Erken's 14-bands images, 1 and 2, used in this work (R^2=44% when estimating Chl-a).

Tables 4, 5 and 6 present the estimation results of applying the linear statistical method to all available images: Image 1, Image 2, Image 3-AtmCorr and Image 3-Raw, Image 3-AtmCorr-14-bands and Image 3-Raw-14-bands, Image 3-AtmCorr-2-bands as well as Image 3-Raw-2-bands. The results of Image 1 were slightly better, while the other images gave slightly worse results, than the band-10 approach, except Image 2 which gave approximately the same results in both approaches. These results also show the clear impact of the spatial resolution on the estimation accuracy; the finer resolution, the better. It is also shown that the image Image 3-Raw-2-bands, generated by the backward elimination procedure, gave much better results than those of the original one. On the other hand, the results are approximately the same when using Image 3-AtmCorr, Image 3-AtmCorr-14-bands or Image 3-AtmCorr-2-bands.

Figures 9 and 10 present the descriptive spectral signatures pairs (when considering the field concentrations of Chl-a, SPM, SPOM and SPIM) obtained when using the simulated 14-bands and the 288-bands field reflectance spectra, respectively. These signatures describe the variations in the spectral characteristics with respect to each of the water quality parameters (i.e. Chl-a, SPM, SPOM or SPIM). The descriptive spectral signatures resulting from the second normalisation approach (Bw...Pw), presented in the left columns of Figs. 9 and 10, show if there is a relative decrease or increase in the amplitude at each band, while the signatures generated by the first normalisation approach (Pw...Pw), presented in the right columns of Figs. 9 and 10, describe how fast the variations are at each band (e.g. a steeper or sharper peak, or a deeper trough).

In the cases of Chl-a, SPM and SPOM, the descriptive spectral signatures pairs were similar, describing overall increase in reflectance in the wavelength intervals 490-670 nm and 700-735 nm. Also, steeper peaks appear in the wavelength intervals 540-670 nm and 690-845 nm, while deeper troughs can be found in the wavelength intervals 400-530 nm and 670-690 nm. In the case of SPIM, overall reflectance increase can be observed in the wavelength intervals 500-530 nm and 560-690 nm, a deeper trough at wavelengths below 530 nm, and steeper peaks in the wavelength intervals 550-690 nm and 760-775 nm.

Table 7 gives a brief presentation of the estimation errors (the root mean square errors, RMSE, in addition to the percentage results with errors larger than certain levels), while Figures 12, 13 and 14 present histograms of the estimation

errors when applying the linear statistical method to Image 1, Image 3-AtmCorr-2-bands and Image 3-AtmCorr-288-bands, respectively (using these images gave the best results in the case of each band set). Investigating Table 7 and Figures 12, 13 and 14, shows that the RMSE can be considered as a good error measure in this study (the majority of the results are below the RMSE value).

Figures 15, 16 and 17, show the resulting concentration maps of Chl-a, SPM, SPOM and SPIM in Image 1, Image 2 and Image 3, respectively. High correlations are observed between the concentrations of Chl-a, SPM, SPOM and SPIM in these images. The concentration map of Chl-a is similar to the corresponding map presented by Östlund et al. (2001) [34], using the band-10 approach.

Figure 18 shows the sums and products of the descriptive spectral signatures presented in Figs. 9 and 10, in the case of utilising the 14-bands and the 288-bands field reflectance spectra.

Tables 8, 9 and 10 present the estimation results of correlating the sums of the descriptive spectral signatures with all available images, while Tables 11, 12 and 13 present the estimation results when correlating the products of the descriptive spectral signatures with the images. In both cases, the results are slightly better than when using the linear statistical method. It can also be noticed that the estimation accuracy is not sensitive to the number of bands when using the various variants of Image 3 (the Raw and the AtmCorr cases).

Figure 19 presents 10 spectral signatures obtained by applying linear ICA to Image 1. Figure 20 shows a spectral reflectance signature for chlorophyll-*a*, obtained as a linear combination of the 10 spectral signatures of Fig. 19, to produce a spectral signature that is as close as possible to the absorption spectra of chlorophyll-a presented in Fig. 3. Table 14 presents the results of performing FVBA on the 14-bands images: Image 1, Image 2, Image 3-AtmCorr-14-bands and Image 3-Raw-14-bands. The results were also similar to those of the previous approaches.

A final observation that can be noted here is that raw CASI data produced results of approximately the same quality as the corresponding atmospherically corrected data (which gave slightly better results). This indicates that a sort of atmospheric correction was implicitly performed by the computations of the linear model.

Figure 21 shows comparison between three selected raw (upwelling radiance) spectra and the corresponding atmospherically corrected reflectance spectra when using the 6S-code, and a linear transformation results, where the transformation coefficients were estimated using a linear system of equations

$$\boldsymbol{S_{fld}} = \boldsymbol{S_{img}}\,\boldsymbol{T} \qquad (5)$$

where $\boldsymbol{S_{fld}}$ and $\boldsymbol{S_{img}}$ are matrices consisting of field reflectance spectra and the corresponding image (upwelling radiance) spectra. The transformation matrix $\boldsymbol{T}$ was estimated, then used to correct additional image spectra. It can easily be noticed that the resulted spectra are as smooth as the field reflectance spectra.

7 Summary and Conclusions

In this work, three statistical methods were developed for the prediction of the concentrations of Chl-a, SPM, SPOM and SPIM from remotely sensed multi- and hyperspectral images. A linear statistical model was employed to estimate transformation coefficients that were used to convert image data into concentration values of Chl-a, SPM, SPOM and SPIM. Two normalisation approaches were employed to produce descriptive spectral signatures pairs for each water quality parameter.

These signature pairs were useful for revealing the impact of the corresponding parameters on the spectral characteristics of the water. Furthermore, these signature pairs were used to estimate the concentrations of Chl-a, SPM, SPOM and SPIM from image data. FVBA was also used to estimate a Chl-a spectral signature that could be used to estimate the concentrations of Chl-a in the lake. The estimation accuracy was approximately the same for the three approaches. Concentration maps of Chl-a, SPM, SPOM and SPIM in Image 1, Image 2 and Image 3, were produced.

High correlations were observed between the concentrations of Chl-a, SPM, SPOM and SPIM in these images. Backward elimination was employed to find out the most useful spectral bands for these water quality parameters. This method showed that using a green band (centred at 556 nm) and a near infrared band (starting at 813.5 nm) produced the best estimation results for Chl-a, SPM, SPOM and SPIM. Finally, the linear statistical model was employed to convert raw CASI data (upwelling radiance) into ground reflectance data.

Analysis of the erroneous correspondence observed between estimated and measured values, leads to the conclusion that going forth and back between the different CASI-scan lines which form a mosaic-image, gives raise to considerable errors that affects the evaluation of the investigated methods.

Hence, imperfect geometric correction and scan-lines alignment are the main error sources to suspect.

In addition to that, artefacts between merged CASI-scan lines are clearly visible in the resulting concentration maps in Figures 15, 16 and 17; probably due to bad edge correction, resulting in darker scan-line edges. This can also be a side effect of imperfect radiometric calibration and/or incomplete atmospheric correction.

In order to further investigate the 1722 continuous field measurements, a special type of 2D-interpolation (by using Pseudo-Dirichlet tessellations) was performed. Visual inspection of the interpolation-like results presented in Fig. 22, reveals unexpected artefacts which can be explained by the impact of direct sunlight on these measurements (whether the measuring instrument was turned towards or away from direct sunlight).

However, the results are very promising and show good correlation between measured and estimated surface water quality parameters, despite the artefacts and shortcomings of the used data, which certainly decrease the observed performance of the employed methods when processing these data; cf. Hamid Muhammed et al. 2003 [14] and Hamid Muhammed 2005 [15], where much better results were obtained by applying methods, similar to those of the current work, to artefact-free hyperspectral data.

The linear statistical method and the approach using descriptive spectral signatures are more computationally efficient than FVBA. These approaches can be applied to other remote sensing applications utilising multi- and hyperspectral data acquired by satellite- and airborne systems. Site- and sensor independent statistical method can be achieved if a one-to-one relation can be achieved between field reflectance spectra and the corresponding corrected image spectra. This requires more efficient atmospheric correction and radiometric calibration methods. It is also possible to build a number of statistical models for various waters and use descriptive spectral signatures to be able to identify the most suitable model to be used in each case.

Acknowledgements:
This work was financed by a research grant from the Swedish National Space Board (SNSB). The author would like to thank Prof. Ewert Bengtsson and Prof. Tommy Lindell (both are working at Centre for Image analysis at Uppsala University, Sweden) for providing the datasets and also for fruitful discussions concerning this work.

References:

[1] AMMENBERG, P., FLINK, P., PIERSON, D.C., LINDELL, T., STRÖMBECK, N. Bio-Optical Modeling Combined with Remote Sensing to Assess Water Quality. International Journal of Remote Sensing, Vol.23(8), 2002, pp.1621-1638.

[2] BARUAH, P.J., OKI, K., NISHIMURA, H. A Neural Network Model for estimating Surface Chlorophyll and Sediment Content at the Lake Kasumi Gaura of Japan. Proc. of Asian Conference on Remote Sensing (ACRS 2000), 2000.

[3] BUKATA, R.P., JEROME, J.H., KONDRATYEV, K.Y.A., POZDNYAKOV, D.V. Satellite monitoring of optically-active components of inland waters: an essential input to regional climate change impact studies. J. Great Lakes Res., Vol.17(4), 1991, pp.470-478.

[4] BUKATA, R.P., JEROME, J.H., KONDRATYEV, K.Y.A., POZDNYAKOV, D.V. *Optical Properties and Remote Sensing of Inland and Coastal Waters*, CRC Press Inc., Florida, 1995, 362 p.

[5] CAIRNS, S.H., DICKSON, K.L., ATKINSON, S.F. An examination of measuring selected water quality trophic indicators with SPOT satellite HRV data, Photogramm. Eng. Remote Sens., Vol.63, 1997, pp.263-265.

[6] CURRAN, P.J., NOVO, E.M.M. The relationship between suspended sediment concentration and remotely sensed spectral radiance: a review. J. Coastal Research. Vol.4, 1988, pp.351-368.

[7] DEKKER, A.G., PETERS, S.W.M. The use of the Thematic Mapper for the analysis of eutrophic lakes: a case study in the Netherlands. Int. J. Remote Sens., Vol.14, 1993, pp.799-821.

[8] DOERFFER, R., FISCHER, J., STOSSEL, M., BROCKMANN, C. Analysis of Thematic Mapper data for studying the suspended matter distribution in the coastal area of the Germany Bight (North Sea). Remote Sens. Environ. Vol.28, 1989, pp.61-73.

[9] DOERFFER R. FISCHER J. Concentrations of chlorophyll suspended matter and gelbstoff in case II waters derived from satellite coastal zone color scanner data with inverse modeling methods. J. Geophys. Res., Vol.99(C4), 1994, pp.7457-7466.

[10] FISCHER J., DOERFFER R., GRASSL H. Factor analysis of multisprectral radiance over coastal and open ocean water based on radiative transfer calculations. AO, Vol.25(3), 1986, pp.448-456.

[11] GIARDINO, C., PEPE, M., BRIVIO, P.A., GHEZZI, P., ZILIOLI, E. Detecting chlorophyll,

secchi-disk depth and surface temperature in a sub-alpine lake using Landsat imagery: Science of the Total Environment, Vol.268, 2001, pp.19–29.

[12] GOLDBERG, D.E. *Genetic algorithms in search, optimisation and machine learning*. Addison & Wesley, 1989.

[13] HAMID MUHAMMED, H., AMMENBERG, P., BENGTSSON, E. Using Feature-Vector Based Analysis, Based on Principal Component Analysis and Independent Component Analysis, for Analyzing Hyperspectral Images. Proc. of International Conference on Image Analysis and Processing (ICIAP 2001), IEEE Computer Society Press, 2001, pp.309-315.

[14] HAMID MUHAMMED H., LARSOLLE A. Feature Vector Based Analysis of Hyperspectral Crop Reflectance Data for Discrimination and Quantification of Fungal Disease Severity in Wheat, Biosystems Engineering, Silsoe Research Institute, Elsevier Ltd., UK, Vol.86(2), 2003, pp.125-134.

[15] HAMID MUHAMMED H. Hyperspectral Crop Reflectance Data for characterising and estimating Fungal Disease Severity in Wheat, Biosystems Engineering, Silsoe Research Institute, Elsevier Ltd., UK, Vol.91(1), 2005, pp.9-20.

[16] HAN, L. Spectral Reflectance with Varying Suspended Sediment Concentrations in Clear and Algal-Laden Waters, Photogrammetric Eng. and Remote Sensing, Vol.63(6), 1997, pp.701-705.

[17] HAN, L., RUNDQUIST, D.C. Comparison of NIR/RED Ratio and First Derivative on Reflectance in Estimating Algal-Chlorophyll-a Concentration: A Case Study in a Turbid Reservoir, Remote Sensing Environment, Vol.62, 1997, pp.253-261.

[18] HYVÄRINEN, A. Survey on independent component analysis. Neural Computing Surveys, Vol.2, 1999, pp.94-128.

[19] HYVÄRINEN, A., OJA, E. Independent Component Analysis: Algorithms and Applications. Neural Networks, Vol.13(4-5), 2000, pp.411-430.

[20] IBRAHIM, A.N., MURAKAMI, M., MABUCHI, Y. Monitoring Algae with Different Levels of Turbidity in the Ishizuchi Storm Water Reservoir in Kochi Prefecture, Japan. Proc. of International Conference on Hydrology and Water Resources in Asia Pacific Region (APHW 2004), 2004.

[21] JOLLIFFE, I.T. *Principal Component Analysis*. Springer Verlag, 1986.

[22] KOPONEN, S., PULLIAINEN, J., SERVOMAA, H., ZHANG, Y., HALLIKAINEN, M., KALLIO, K., VEPSÄLÄINEN, J., PYHÄLAHTI, T., HANNONEN, T. Ananlysis on the feasibility of multisource remote sensing observations for chl-a monitoring in Finnish lakes. The Science of the Total Environment, Vol.268, 2001, pp.95-106.

[23] KOPONEN, S., PULLIAINEN, J., KALLIO, K., HALLIKAINEN, M. Lake water quality classification with airborne hyperspectral spectrometer and simulated MERIS data. Remote Sens. Environ., Vol.79, 2002, pp.51-59.

[24] KUTSER, T. *Estimation of Water Quality in Turbid Inland and Coastal Waters by Passive Optical Remote Sensing*, Doctor thesis, University of Tartu, Estonia, 1997.

[25] LATHROP, R.G.Jr., LILLESAND, T.M. Use of Thematic Mapper data to assess water quality in Green Bay and central Lake Michigan. Photogramm. Eng. Remote Sens., Vol.52, 1986, pp.671-680.

[26] LATHROP, R.G.Jr., VANDE CASTLE, J.R., LILLESAND, T.M. Monitoring river plume transport and mesoscale circulation in Green Bay, Lake Michigan, through satellite remote sensing; Journal of Great Lakes Research, 16(43), 1990, pp.471-478.

[27] LATHROP, R.G.Jr., LILLESAND, T.M., YANDELL, B.S. Testing the utility simple multi-date Thematic Mapper calibration algorithms for monitoring turbid inland waters. Int. J. Remote Sens. Vol.12, 1991, pp.2045-2063.

[28] LILLESAND, T.M., JOHNSON, W.L., DEUELL, R.L., LINDSTROM, O.M., MIESNER, D.E. Use of Landsat data to predict trophic status of Minnesota lakes. Photogramm. Eng. Remote Sens. Vol.49, 1983, pp.219-229.

[29] LINDELL, T., PIERSON, D., PREMAZZI, G., ZILIOLI, E. *Manual for monitoring European lakes using remote sensing techniques*, (Luxembourg: Official Publications of the European Communities), EUR 18665 EN, 1999, 161 p.

[30] McCLUNEY, W.R. *Remote Measurement of "turbidity" and other Water Quality Parameters*. NASA/Goddard Space Flight Center, Greenbelt, MD, 1975.

[31] MITTENZEWY, K.H., GITELSON, A.A., KONDRATYEV, K.Y. Determination of Chlorophyll-a of Inland Waters on the Basis Of Spectral Reflectance. Limnological Engineering, Vol.37, 1992, pp.147-149.

[32] MOREL, A. In-water and remote measurements of color. Boundary-layer Meteorol., Vol.18, 1980, pp.177-201.

[33] OLMANSON, L.G., Bresonick, P.L., Kloiber, S.M., Bauer, M.E., Day, E.E. *Lake Water*

Clarity Assessment of Minnesota's 10,000 Lakes: A Comprehensive View from Space, Water Resources Center, University of Minnesota, 2000.

[34] ÖSTLUND, C., FLINK, P., STRÖMBECK, N., PIERSON, D., LINDELL, T. Mapping Of The Water Quality Of Lake Erken, Sweden, From Imaging Spectrometry And Landsat Thematic Mapper, The Science of the Total Environment, Vol.268, pp.139-154.

[35] PENA-MARTINEZ, R., DOMINGUEZ-GOMEZ, J.A., RUIZ-VERDU, A. Mapping of Photosynthetic Pigments in Spanish Reservoirs. Proceedings of MERIS user Workshop, ESA-ESRIN, Frascati, Italy, 2003.

[36] PHILIPSON, P., PIERSON, D., LINDELL, T. Evaluation of Swedish lake water quality modeling from remote sensing. Proc. of SPIE Remote Sensing for Environmental Monitoring, GIS applications and Geology II, edited by Manfred Ehlers (SPIE, Bellingham, WA), Vol.4886, 2003, pp.561-570.

[37] ROBINSON, I.S. Satellite observations of ocean colour. Phil. Trans. Roy. Soc. London, A309, 1983, pp.415-432.

[38] ROGER R.E. Principal components transform with simple automatic noise adjustment IJRS, Vol.17(14), 1996, pp.2719-2728.

[39] SCHALLES, J.F., GITELSON, A.A., YACOBI, Y.Z., KROENKE, A.E. Estimation of chlorophyll a from time series measurements of high spectral resolution reflectance in an eutrophic lake. Journal of Phycology, Vol.34(2), 1998, 383-390.

[40] SCHIEBE, F.R., RITCHIE, J.C. *Color Measurements and Suspended Sediments in North Mississippi Reservoirs*, Remote Sensing of Earth Resources, Vol. IV, The University of Tennessee Space Institute, Tullahoma, Tennessee, 1975.

[41] STRÖMBECK, N., CANDIANI, G., GIARDINO, C. Water quality monitoring of Lake Garda using multi-temporal MERIS data. The Meris user workshop – 2003, 2003.

[42] STRONG, A.E. Remote sensing of algal blooms by aircraft and satellite in Lake Erie and Utah Lake, Remote Sens. Environ., Vol.3, 1974, pp.99-107.

[43] TASSAN, S. Evaluation of the potential of the Thematic Mapper for marine application. Int. J. Remote Sens., Vol.8, 1987, pp.1455-1478.

[44] TASSAN, S. Local algorithms using SeaWiFS data for the retrieval of phytoplankton, pigments, suspended sediment, and yellow substance in coastal waters. Appl. Opt., Vol.33, 1994, pp.2369-2378.

[45] VERMOTE, E.F., TANRÉ, D., DEUZÉ, J.L., HERMAN, M. AND MORCRETTE, J.J. Second simulation of the satellite signal in the solar spectrum, 6S, An Overview. *IEEE Transactions on Geoscience and Remote Sensing*, Vol.35(3), 1997, pp.675-686.

[46] ZALLOUM. O.E., O' MONGAIN J. Walsh S. Danaher and L. Stapleton Dye concentration estimation by remotely-sensed spectral radiometry. International Journal of Remote Sensing, Vol.14, 1992, pp.2285-2300.

[47] ZHANG, Y., KOPONEN, S., PULLIAINEN, J., HALLIKAINEN, M. Application of empirical neural networks to chlorophyll-a estimation in coastal waters using optosensors. *IEEE Sensors Journal*, Vol.3(4), 2003, pp.376-382.

Appendix A. Tables 3-14

Table 3. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when using band 10.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)	SPM R^2, % (RMSE, mg/l)	SPOM R^2, % (RMSE, mg/l)	SPIM R^2, % (RMSE, mg/l)
Image 1	Before matching	73.12 (8.0)	73.12 (0.9)	73.08 (0.8)	68.84 (0.09)
	After matching	85.52 (6.7)	85.52 (0.7)	85.5 (0.7)	82.17 (0.08)
Image 2	Before matching	72.08 (8.2)	72.07 (0.9)	72.01 (0.9)	67.78 (0.1)
	After matching	83.69 (7.1)	83.68 (0.7)	83.66 (0.7)	80.42 (0.1)
Image 3-AtmCorr (14 bands)	Before matching	46.57 (9.7)	46.58 (1.2)	46.55 (1.1)	44.24 (0.1)
	After matching	77.72 (6.7)	77.72 (0.8)	77.71 (0.7)	73.44 (0.07)
Image 3-Raw (14 bands)	Before matching	45.78 (9.8)	45.79 (1.2)	45.76 (1.1)	43.38 (0.1)
	After matching	78.78 (6.6)	78.79 (0.7)	78.8 (0.7)	74.17 (0.07)

Table 4. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when applying the linear statistical model to Image 1 & Image 2.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)		SPM R^2, % (RMSE, mg/l)		SPOM R^2, % (RMSE, mg/l)		SPIM R^2, % (RMSE, mg/l)	
		Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2
Using FRS	Before matching	71.98 (8.3)	68.14 (9.8)	71.89 (1.0)	67.96 (1.2)	71.83 (0.9)	67.88 (1.1)	67.44 (0.3)	63.74 (0.3)
	After matching	86.36 (6.7)	82.27 (8.4)	86.29 (0.8)	82.11 (1.0)	86.24 (0.7)	82.07 (0.9)	82.16 (0.3)	79.38 (0.3)
Using IRS	Before matching	71.93 (8.0)	70.15 (8.9)	71.82 (1.0)	70.06 (1.1)	71.75 (0.9)	69.96 (1.0)	67.42 (0.3)	65.97 (0.3)
	After matching	86.13 (6.3)	83.16 (7.5)	86.04 (0.8)	82.99 (0.9)	85.98 (0.7)	82.89 (0.8)	82.09 (0.3)	80.18 (0.3)

Table 5. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when applying the linear statistical model to Image 3-Raw.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)			SPM R^2, % (RMSE, mg/l)			SPOM R^2, % (RMSE, mg/l)			SPIM R^2, % (RMSE, mg/l)		
		288	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*
Using FRS	Before matching	50.0 (10.6)	50.21 (10.5)	51.78 (10.3)	49.96 (1.3)	50.16 (1.3)	51.79 (1.2)	49.93 (1.3)	50.14 (1.2)	51.76 (1.2)	46.99 (0.3)	47.22 (0.3)	48.66 (0.3)
	After matching	72.87 (8.0)	72.96 (7.9)	74.86 (7.6)	72.70 (0.9)	72.78 (0.9)	74.87 (0.9)	72.69 (1.0)	72.76 (0.9)	74.86 (0.9)	67.65 (0.3)	67.69 (0.3)	69.70 (0.3)
Using IS	Before matching	47.75 (11.0)	48.11 (10.8)	51.68 (10.3)	47.58 (1.3)	47.93 (1.3)	51.69 (1.2)	47.55 (1.3)	47.91 (1.3)	51.66 (1.2)	44.90 (0.3)	45.23 (0.3)	48.59 (0.3)
	After matching	67.91 (8.8)	68.12 (8.6)	74.67 (7.7)	67.56 (1.0)	67.77 (1.0)	74.69 (1.0)	67.53 (1.0)	67.74 (1.0)	74.68 (0.9)	62.77 (0.3)	62.92 (0.3)	69.52 (0.3)

Table 6. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when applying the linear statistical model to Image 3-AtmCorr.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)			SPM R^2, % (RMSE, mg/l)			SPOM R^2, % (RMSE, mg/l)			SPIM R^2, % (RMSE, mg/l)		
		288	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*
Using FRS	Before matching	50.72 (10.2)	50.91 (10.3)	52.14 (10.1)	50.73 (1.2)	50.91 (1.2)	52.15 (1.2)	50.69 (1.2)	50.87 (1.2)	52.12 (1.2)	47.76 (0.3)	47.97 (0.3)	49.08 (0.3)
	After matching	74.28 (7.5)	74.09 (7.7)	74.9 (7.4)	74.21 (0.9)	74.0 (0.9)	74.91 (0.9)	74.18 (0.9)	73.97 (0.9)	74.88 (0.9)	69.39 (0.3)	69.2 (0.3)	69.91 (0.3)
Using IRS	Before matching	50.69 (10.1)	51.16 (10.3)	52.04 (9.9)	50.69 (1.2)	51.18 (1.2)	52.04 (1.2)	50.66 (1.2)	51.14 (1.2)	52.01 (1.2)	47.45 (0.3)	48.18 (0.3)	48.75 (0.3)
	After matching	73.94 (7.5)	74.13 (7.6)	74.94 (7.3)	73.83 (0.9)	74.06 (0.9)	74.95 (0.8)	73.81 (0.9)	74.03 (0.9)	74.93 (0.9)	68.88 (0.3)	69.30 (0.3)	69.82 (0.2)

Table 7. Estimation errors when applying the linear statistical model.

	Chl-a & Phaeo-a			SPM			SPOM			SPIM		
	Percentage results with error			*Percentage results with error*			*Percentage results with error*			*Percentage results with error*		
	> 5 µg/l	*> 10 µg/l*	*> 15 µg/l*	*> 0.5 mg/l*	*> 1 mg/l*	*> 1.5 mg/l*	*> 0.5 mg/l*	*> 1 mg/l*	*> 1.5 mg/l*	*> 0.2 mg/l*	*> 0.4 mg/l*	*> 0.6 mg/l*
Image 1	42.3%	12.4%	3.7%	52.7%	22.6%	6.3%	46.7%	11.5%	4.2%	48.0%	21.2%	3.6%
	RMSE = 6.7 µg/l			RMSE = 0.78 mg/l			RMSE = 0.71 mg/l			RMSE = 0.34 mg/l		
Image 3-AtmCorr (2 bands)	52.7%	15.4%	2.7%	55.2%	25.7%	7.0%	65.3%	25.4%	8.6%	37.5%	17.6%	0.9%
	RMSE = 7.3 µg/l			RMSE = 0.83 mg/l			RMSE = 0.94 mg/l			RMSE = 0.31 mg/l		
Image 3-AtmCorr (288 bands)	50.2%	19.8%	3.6%	53.5%	24.6%	9.5%	63.6%	23.8%	7.3%	40.1%	14.5%	1.0%
	RMSE = 7.5 µg/l			RMSE = 0.89 mg/l			RMSE = 0.91 mg/l			RMSE = 0.3 mg/l		

Table 8. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when correlating the sum of the descriptive spectral signatures with Image 1 & Image 2.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)		SPM R^2, % (RMSE, mg/l)		SPOM R^2, % (RMSE, mg/l)		SPIM R^2, % (RMSE, mg/l)	
		Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2
Using FRS	Before matching	73.6 (6.1)	70.7 (11.0)	73.6 (1.4)	70.7 (1.9)	73.6 (1.3)	70.6 (1.8)	68.4 (0.13)	65.7 (0.21)
	After matching	87.6 (4.9)	84.6 (10.4)	87.6 (1.4)	84.6 (1.9)	87.5 (1.2)	84.4 (1.7)	83.3 (0.12)	81.3 (0.21)
Using IRS	Before matching	73.3 (7.2)	70.4 (7.4)	73.2 (0.8)	70.3 (0.9)	73.1 (0.7)	70.1 (0.8)	68.1 (0.12)	64.0 (0.14)
	After matching	87.5 (5.8)	84.4 (6.1)	87.5 (0.6)	84.3 (0.6)	87.5 (0.5)	84.2 (0.6)	83.0 (0.10)	80.6 (0.13)

Table 9. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when correlating the sum of the descriptive spectral signatures with Image 3-Raw.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)			SPM R^2, % (RMSE, mg/l)			SPOM R^2, % (RMSE, mg/l)			SPIM R^2, % (RMSE, mg/l)		
		288	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*
Using FRS	Before matching	49.7 (14.9)	50.4 (13.3)	51.1 (9.7)	49.8 (1.7)	50.4 (1.5)	51.1 (1.2)	49.7 (1.5)	50.3 (1.4)	51.2 (1.1)	45.8 (0.15)	47.1 (0.13)	35.0 (0.10)
	After matching	74.7 (14.8)	75.6 (13.0)	73.8 (9.0)	74.8 (1.7)	75.6 (1.5)	73.8 (1.1)	74.7 (1.5)	75.5 (1.4)	73.9 (1.1)	68.5 (0.15)	70.2 (0.13)	52.9 (0.10)
Using IS	Before matching	50.2 (9.4)	50.7 (9.4)	51.6 (9.3)	49.8 (1.1)	50.5 (1.1)	51.1 (1.1)	49.8 (1.0)	50.5 (1.0)	51.1 (1.0)	45.7 (0.13)	46.6 (0.12)	46.7 (0.13)
	After matching	74.8 (7.2)	75.8 (7.1)	74.6 (7.2)	74.0 (0.8)	75.5 (0.8)	73.8 (0.8)	74.0 (0.8)	75.6 (0.7)	73.8 (0.8)	67.7 (0.10)	68.0 (0.10)	67.9 (0.10)

Table 10. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when correlating the sum of the descriptive spectral signatures with Image 3-AtmCorr.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)			SPM R^2, % (RMSE, mg/l)			SPOM R^2, % (RMSE, mg/l)			SPIM R^2, % (RMSE, mg/l)		
		288	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*
Using FRS	Before matching	49.6 (13.9)	50.4 (11.9)	51.5 (10.4)	49.7 (1.5)	50.5 (1.3)	51.5 (1.2)	49.8 (1.4)	50.5 (1.2)	51.5 (1.1)	45.0 (0.14)	46.9 (0.13)	31.0 (0.21)
	After matching	74.6 (12.9)	75.2 (10.5)	73.4 (7.9)	74.8 (1.3)	75.3 (1.1)	73.4 (0.9)	74.9 (1.2)	75.3 (1.0)	73.5 (0.8)	67.6 (0.13)	69.6 (0.11)	47.9 (0.19)
Using IRS	Before matching	49.6 (9.4)	50.6 (9.3)	51.5 (9.2)	49.1 (1.1)	50.4 (1.1)	51.5 (1.1)	49.3 (1.0)	50.5 (1.0)	51.5 (1.0)	43.2 (0.14)	45.4 (0.13)	48.5 (0.11)
	After matching	74.4 (7.2)	75.3 (7.1)	73.4 (7.1)	73.5 (0.8)	75.0 (0.8)	73.4 (0.8)	73.8 (0.8)	75.1 (0.7)	73.4 (0.8)	65.5 (0.11)	68.1 (0.11)	68.3 (0.08)

Table 11. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when correlating the product of the descriptive spectral signatures with Image 1 & Image 2.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)		SPM R^2, % (RMSE, mg/l)		SPOM R^2, % (RMSE, mg/l)		SPIM R^2, % (RMSE, mg/l)	
		Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2
Using FRS	Before matching	73.5 (6.0)	71.6 (10.0)	73.5 (1.4)	71.6 (1.9)	73.6 (1.4)	71.6 (1.8)	68.5 (0.07)	66.4 (0.12)
	After matching	87.1 (4.7)	84.7 (9.3)	87.1 (1.4)	84.7 (1.8)	87.1 (1.4)	84.7 (1.8)	83.2 (0.05)	81.5 (0.12)
Using IRS	Before matching	73.1 (7.3)	71.2 (7.5)	73.0 (0.8)	71.0 (0.9)	72.8 (0.8)	70.8 (0.8)	66.5 (0.12)	59.7 (0.15)
	After matching	87.0 (5.9)	84.4 (6.1)	87.0 (0.6)	84.2 (0.6)	86.9 (0.6)	84.0 (0.6)	81.9 (0.11)	78.1 (0.13)

Table 12. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when correlating the product of the descriptive spectral signatures with Image 3-Raw.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)			SPM R^2, % (RMSE, mg/l)			SPOM R^2, % (RMSE, mg/l)			SPIM R^2, % (RMSE, mg/l)		
		288	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*
Using FRS	Before matching	50.3 (9.4)	50.5 (9.3)	51.1 (18.5)	50.3 (2.1)	50.5 (1.8)	51.2 (1.2)	50.2 (2.0)	50.4 (1.7)	51.2 (1.1)	46.7 (0.10)	47.6 (0.13)	35.0 (0.17)
	After matching	75.6 (8.9)	76.5 (8.6)	73.8 (17.9)	76.6 (2.1)	76.5 (1.7)	73.8 (1.1)	76.6 (2.0)	76.6 (1.7)	73.9 (1.0)	70.1 (0.10)	70.6 (0.12)	52.9 (0.16)
Using IS	Before matching	50.8 (9.4)	50.8 (9.4)	51.6 (9.3)	50.7 (1.1)	51.0 (1.1)	51.1 (1.1)	50.7 (1.0)	50.9 (1.0)	51.1 (1.0)	45.9 (0.13)	46.0 (0.13)	46.7 (0.13)
	After matching	75.9 (7.1)	76.4 (7.0)	74.6 (7.2)	75.5 (0.8)	76.5 (0.8)	73.8 (0.8)	75.5 (0.7)	76.5 (0.7)	73.8 (0.8)	68.5 (0.10)	69.0 (0.10)	67.9 (0.10)

Table 13. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when correlating the product of the descriptive spectral signatures with Image 3-AtmCorr.

		Chl-a & Phaeo-a R^2, % (RMSE, µg/l)			SPM R^2, % (RMSE, mg/l)			SPOM R^2, % (RMSE, mg/l)			SPIM R^2, % (RMSE, mg/l)		
		288	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*	*288*	*14*	*2*
Using FRS	Before matching	50.2 (9.1)	50.7 (8.5)	51.5 (16.1)	50.2 (1.7)	50.6 (1.4)	51.5 (1.3)	50.0 (1.7)	50.5 (1.4)	51.5 (1.1)	46.0 (0.11)	47.5 (0.11)	31.0 (0.27)
	After matching	76.2 (7.0)	76.1 (5.7)	73.4 (14.5)	76.3 (1.6)	76.1 (1.2)	73.4 (1.0)	76.2 (1.6)	76.2 (1.2)	73.5 (0.8)	69.1 (0.08)	70.4 (0.08)	47.9 (0.25)
Using IRS	Before matching	50.6 (9.3)	50.9 (9.3)	51.5 (9.2)	50.2 (1.1)	51.1 (1.1)	51.5 (1.1)	50.3 (1.0)	51.0 (1.0)	51.5 (1.0)	41.8 (0.15)	42.8 (0.14)	48.5 (0.11)
	After matching	76.1 (7.0)	76.1 (6.9)	73.4 (7.1)	75.4 (0.8)	76.4 (0.8)	73.4 (0.8)	75.6 (0.7)	76.3 (0.7)	73.4 (0.8)	64.1 (0.12)	65.3 (0.12)	68.3 (0.08)

Table 14. Coefficient of determination R^2 and the root-mean square estimation-error RMSE when using FVBA.

	Estimation of chlorophyll-a concentrations R^2, %, (RMSE, μg/l)			
	Image 1	Image 2	Image 3-AtmCorr (14 bands)	Image 3-Raw (14 bands)
Before matching	72.6 (7.7)	68.7 (7.3)	49.9 (9.1)	49.3 (9.1)
After matching	87.3 (7.0)	84.1 (5.9)	75.5 (7.0)	75.4 (7.1)

Appendix B. Figures 5-22

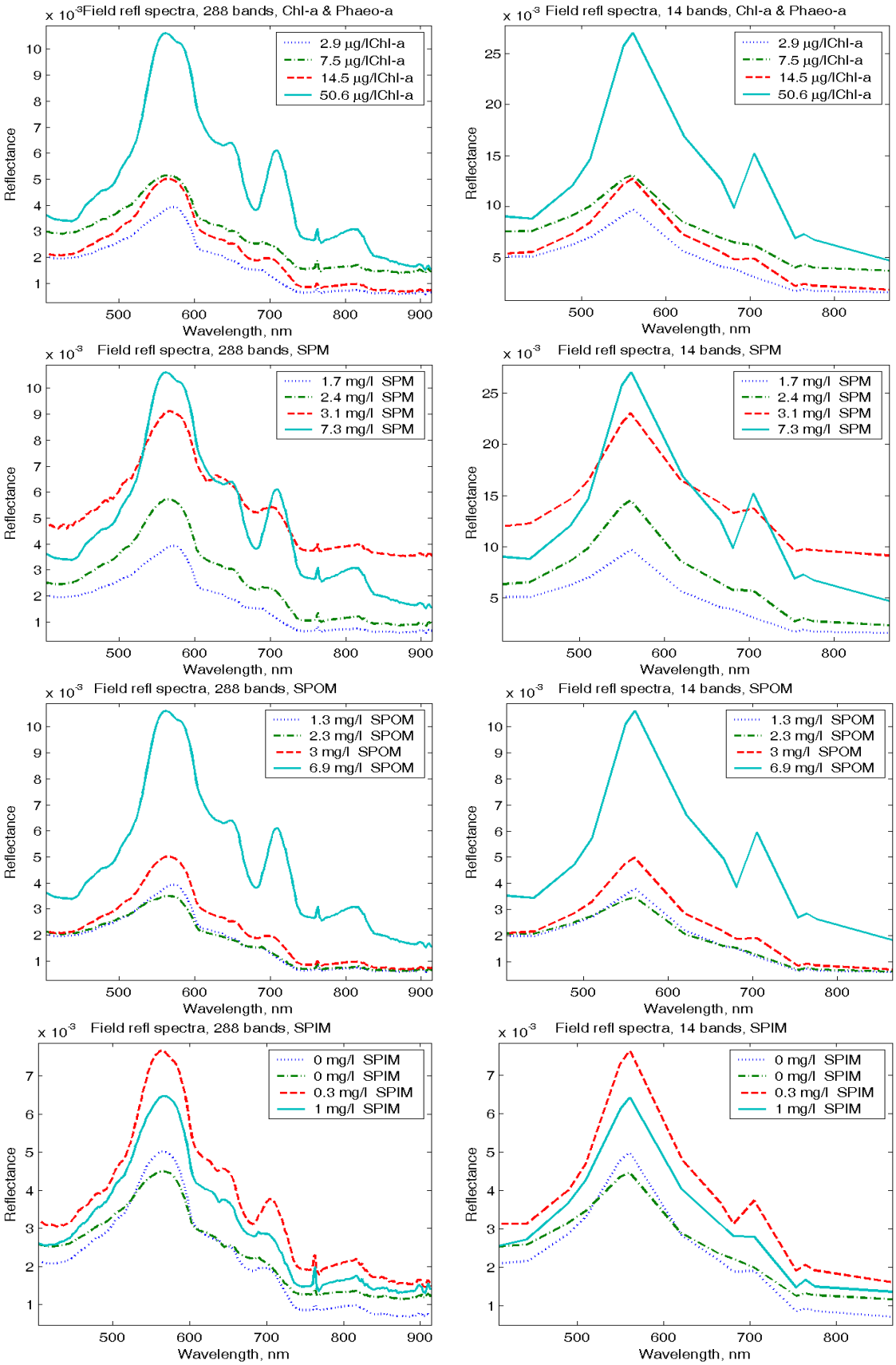

Fig. 5. Field reflectance spectra (FRS) (left column), and simulated 14-bands FRS (right column), corresponding to various concentrations of Chl-a, SPM, SPOM and SPIM, on rows 1, 2, 3 and 4, respectively.

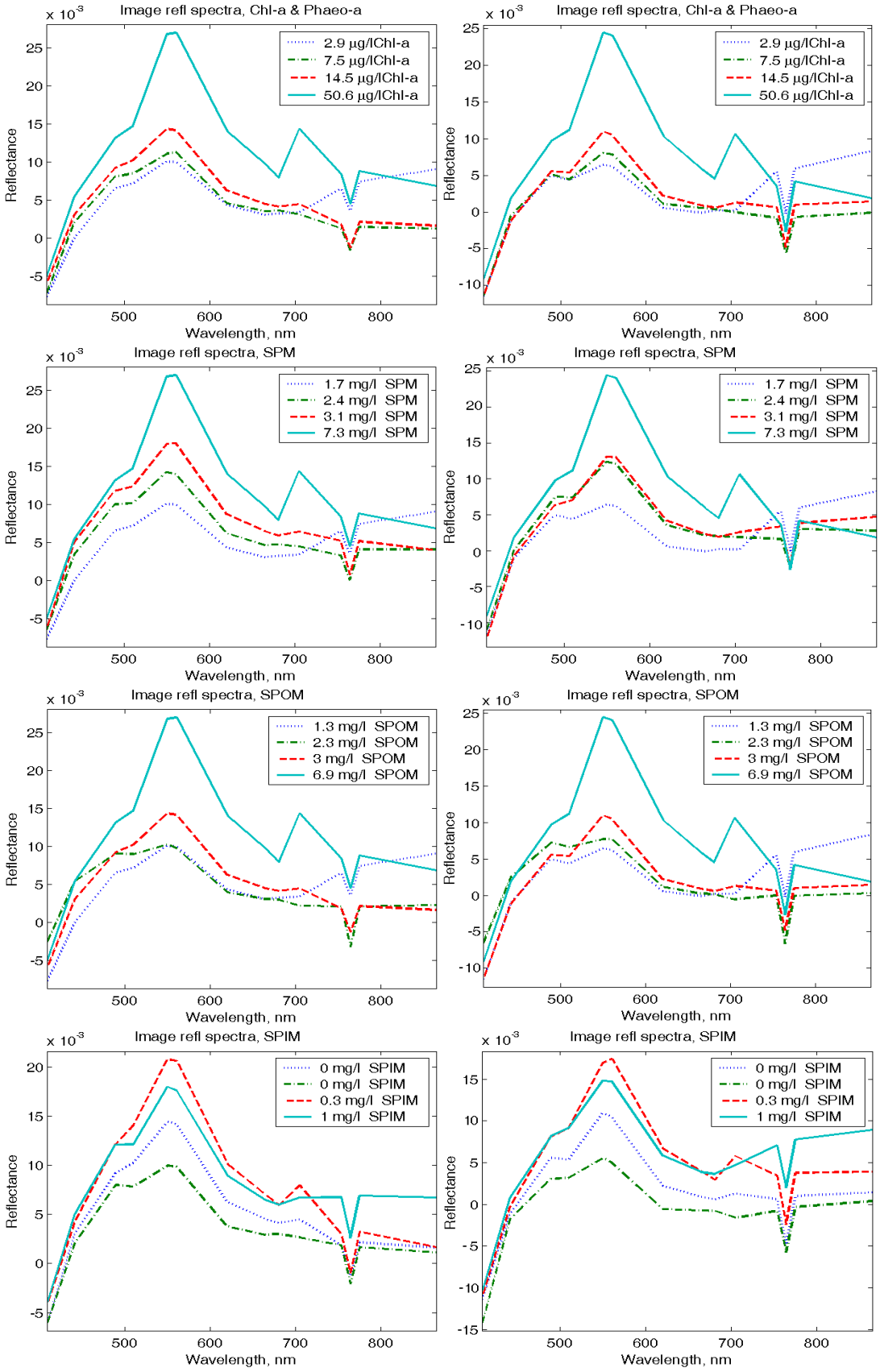

Fig. 6. Graphs with four image reflectance spectra (IRS) corresponding to various concentrations of Chl-a, SPM, SPOM and SPIM, on rows 1, 2, 3 and 4, respectively. IRS extracted from Image 1 and Image 2 are presented in the left and right columns, respectively.

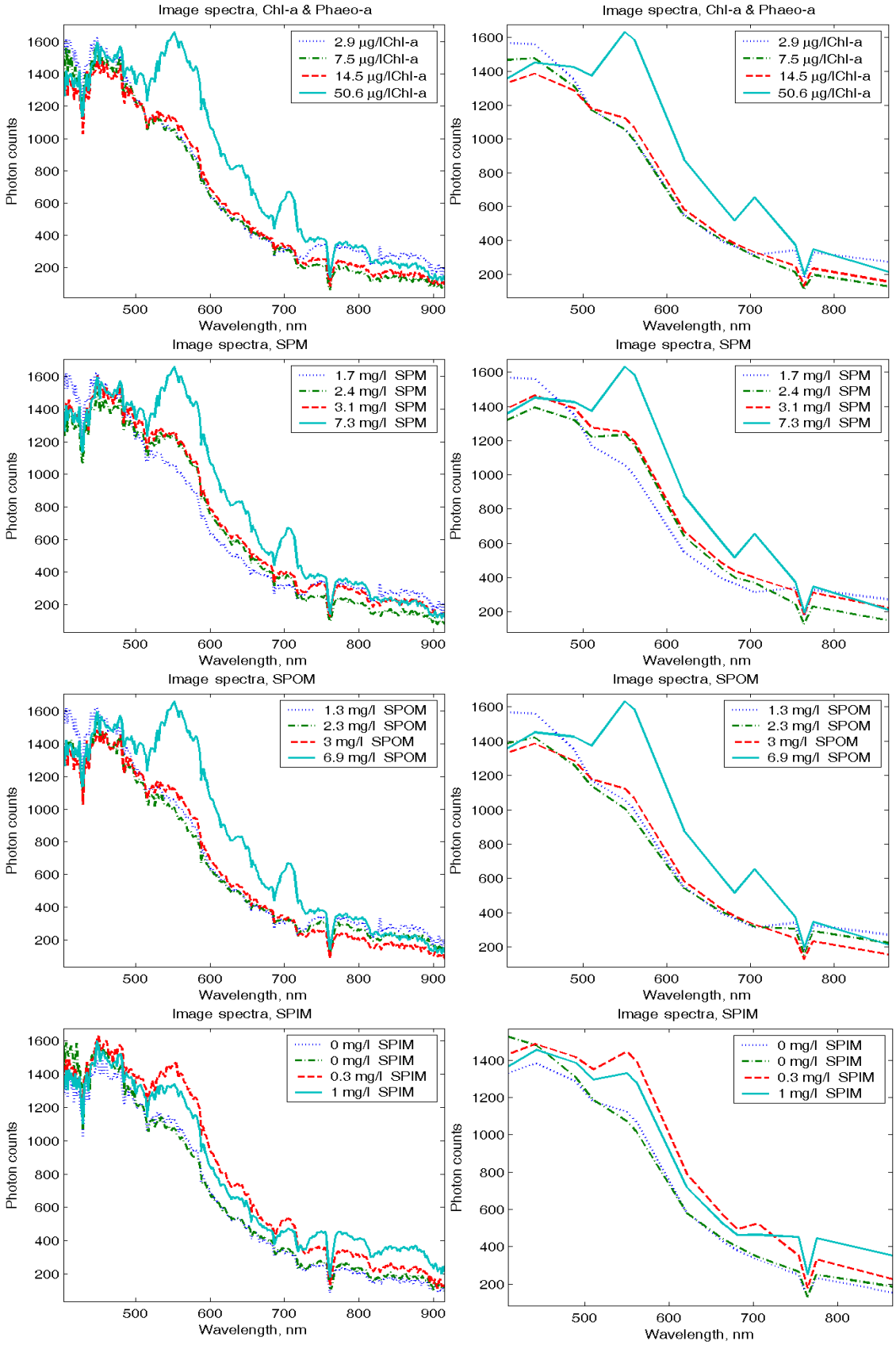

Fig. 7. Graphs with four image spectra (IS) corresponding to various concentrations of Chl-a, SPM, SPOM and SPIM, on rows 1, 2, 3 and 4, respectively. IS extracted from Image 3-Raw and Image 3-Raw-14-bands are presented in the left and right columns, respectively.

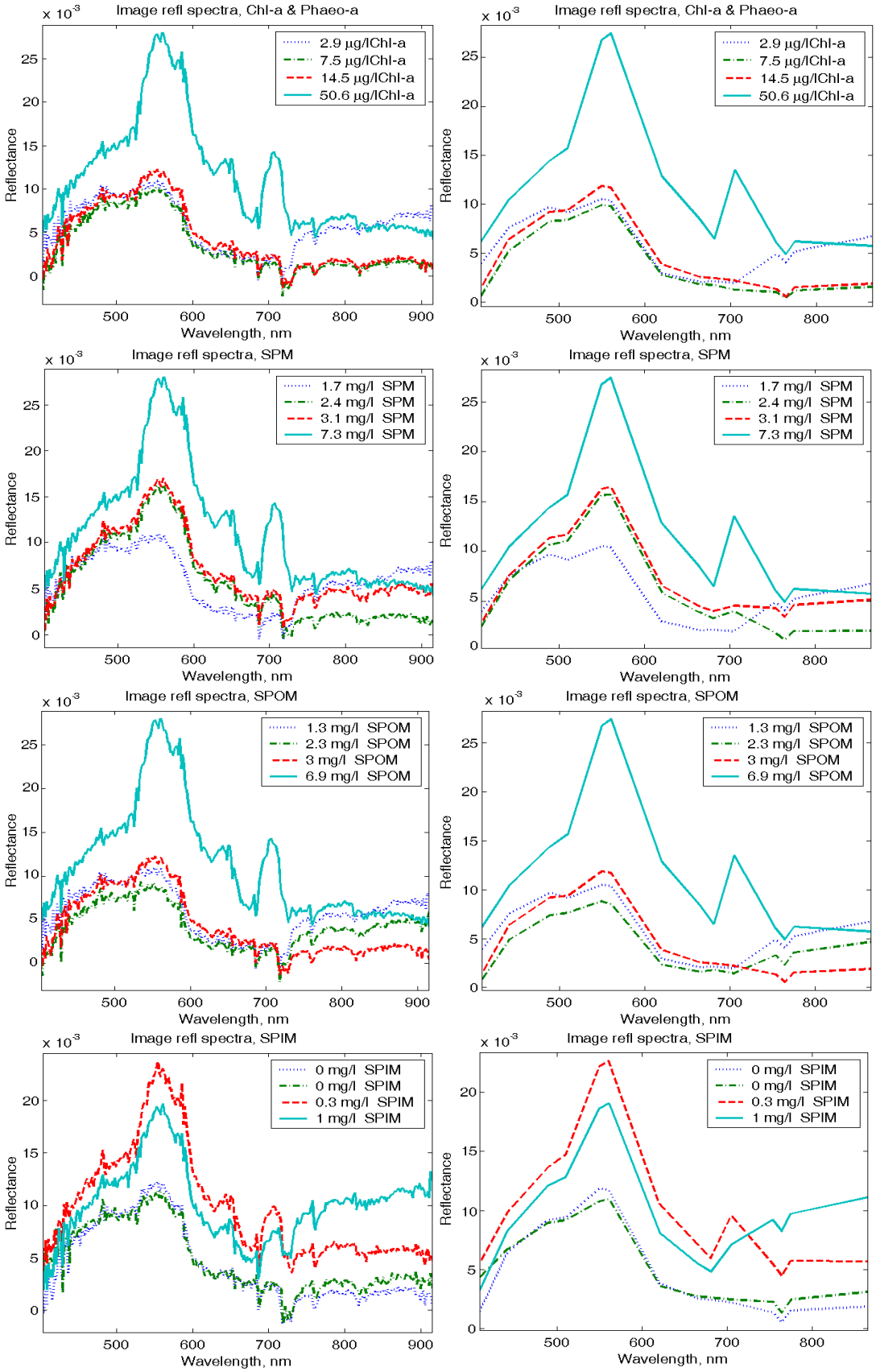

Fig. 8. Graphs with four image reflectance spectra (IRS) corresponding to various concentrations of Chl-a, SPM, SPOM and SPIM, on rows 1, 2, 3 and 4, respectively. IRS extracted from Image 3-AtmCorr and Image 3-AtmCorr-14-bands are presented in the left and right columns, respectively.

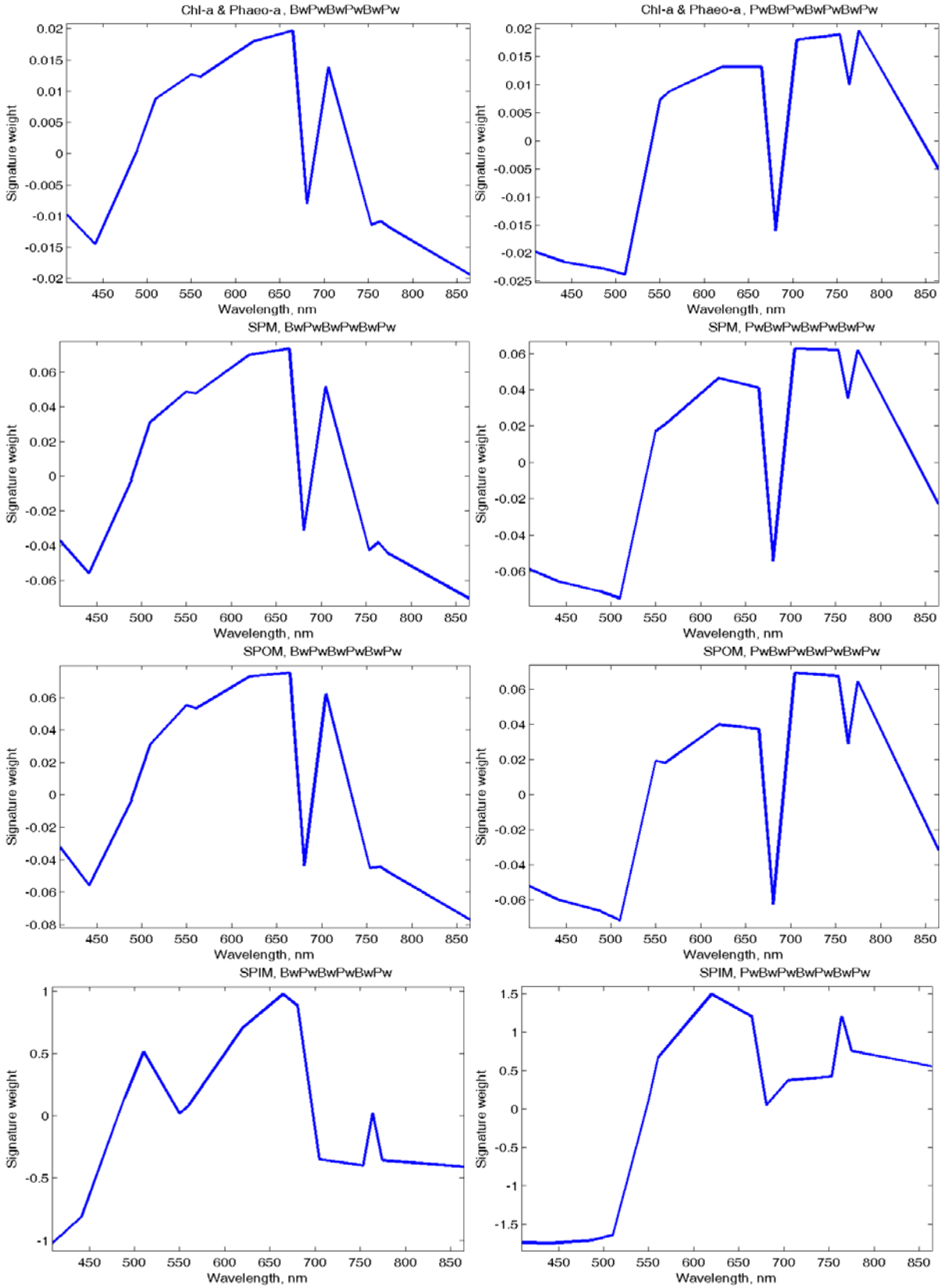

Fig. 9. Descriptive spectral signatures pairs for the concentrations of Chl-a, SPM, SPOM and SPIM (presented on rows 1, 2, 3 and 4, respectively), obtained when considering field concentrations and using simulated 14-bands field reflectance spectra.

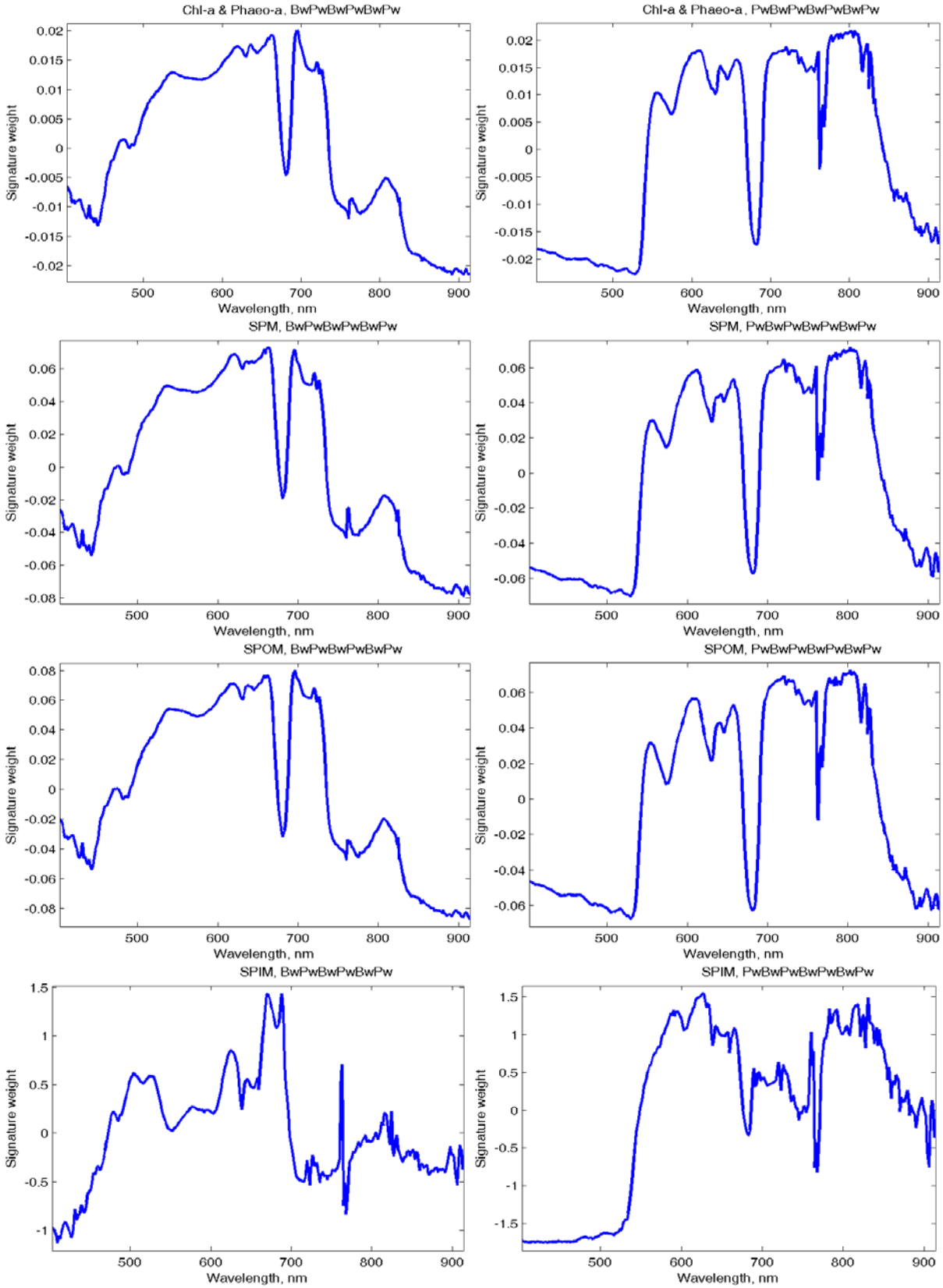

Fig. 10. Descriptive spectral signatures pairs for the concentrations of Chl-a, SPM, SPOM and SPIM (presented on rows 1, 2, 3 and 4, respectively), obtained when considering field concentrations and using simulated 288-bands field reflectance spectra.

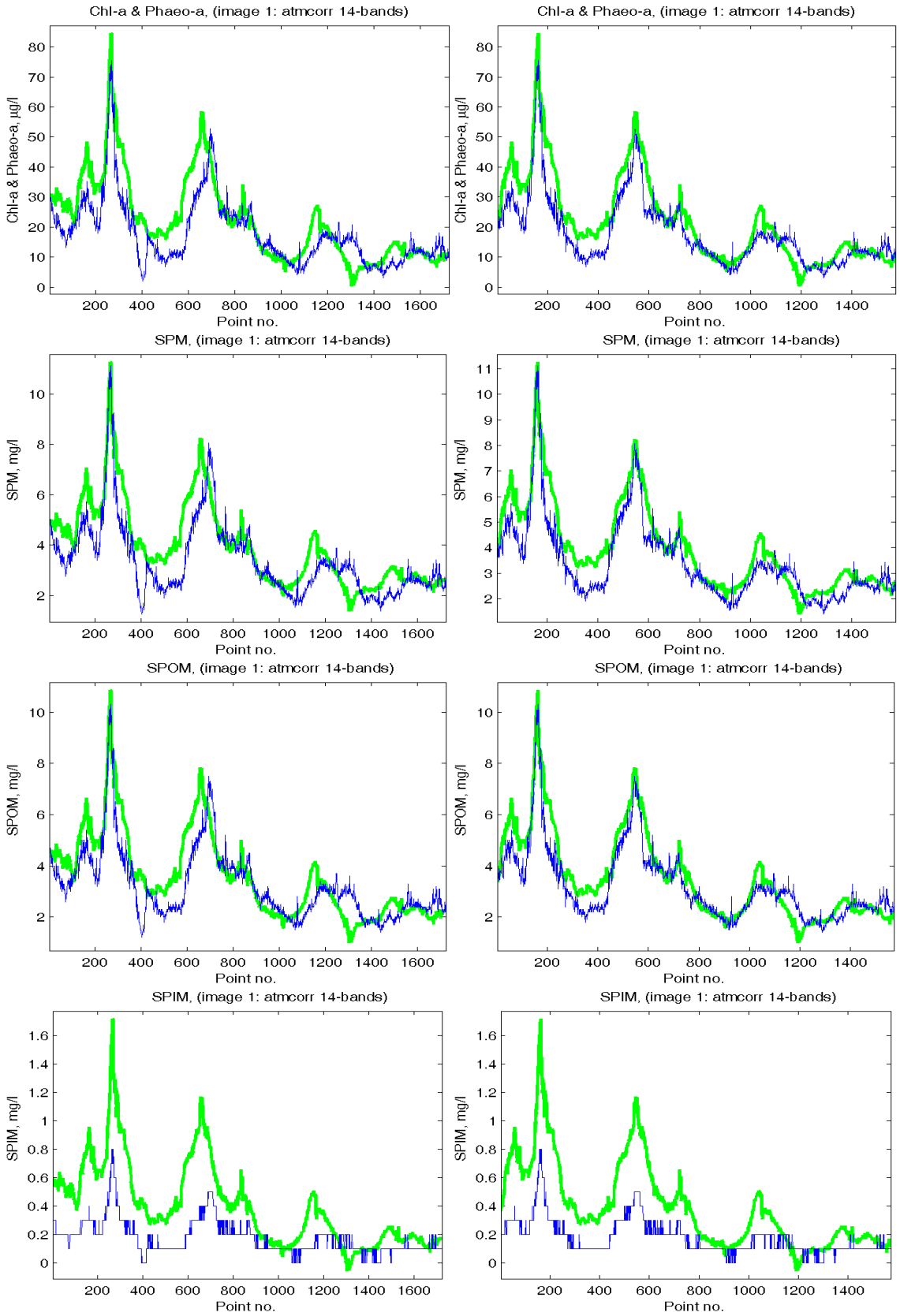

Fig. 11. Comparison between estimated (the green curve) and measured (the blue curve) surface water quality parameters, (left column) before matching, and (right column) after matching. The parameters of interest are Chl-a, SPM, SPOM and SPIM, for which results are presented on rows 1, 2, 3 and 4, respectively.

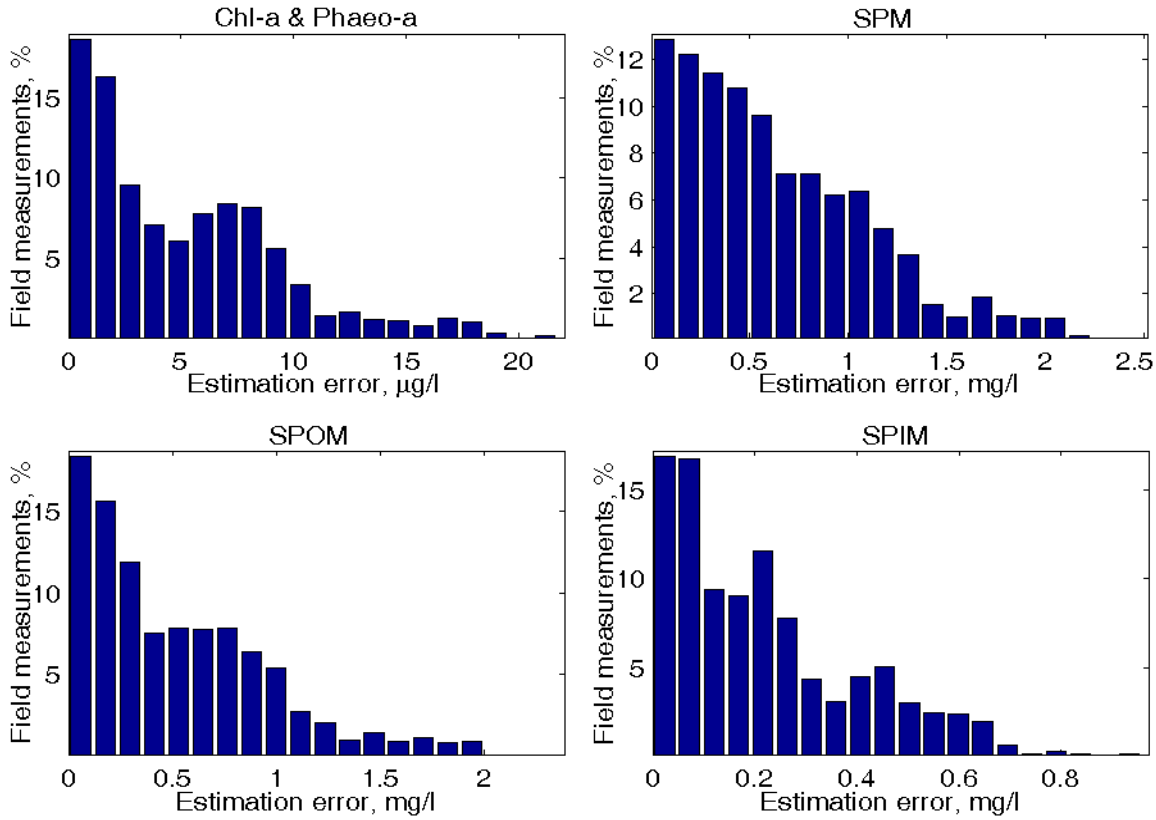

Fig. 12. Histograms of the estimation errors when applying the linear statistical method to Image 1.

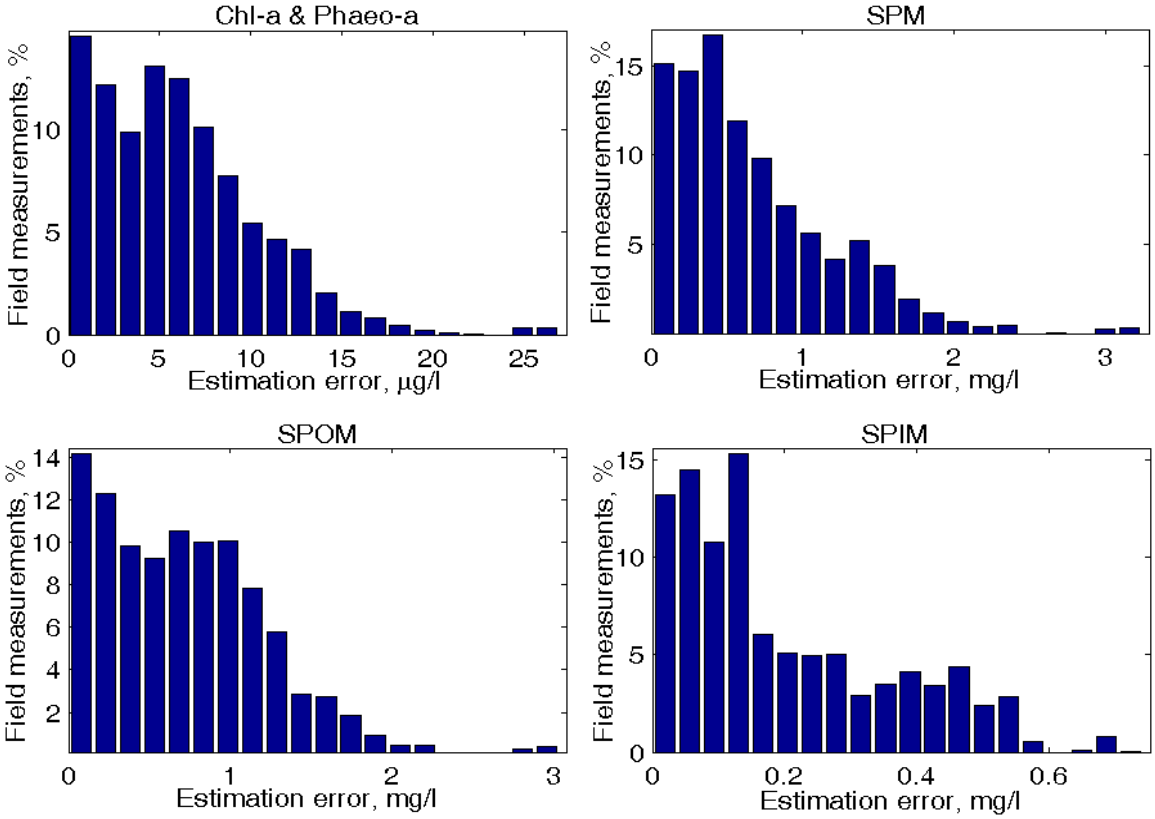

Fig. 13. Histograms of the estimation errors when applying the linear statistical method to Image 3-AtmCorr-2-bands.

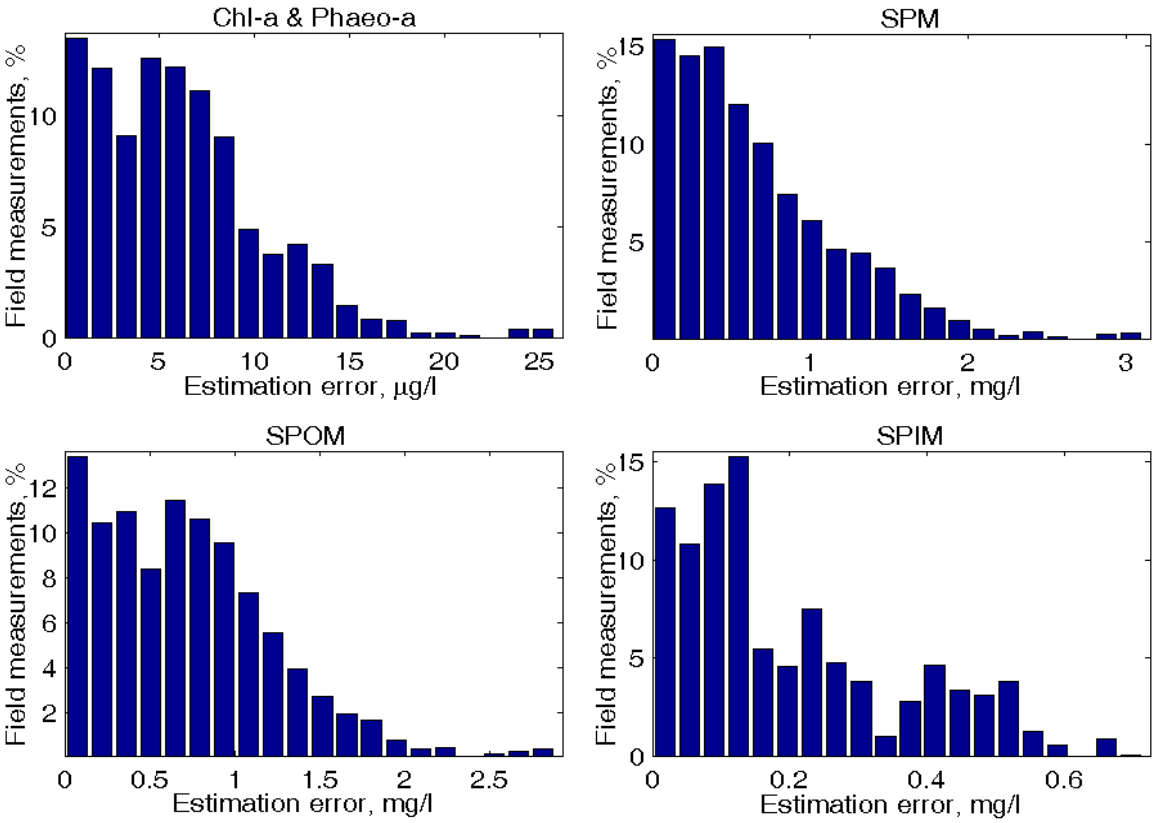

Fig. 14. Histograms of the estimation errors when applying the linear statistical method to Image 3-AtmCorr-288-bands.

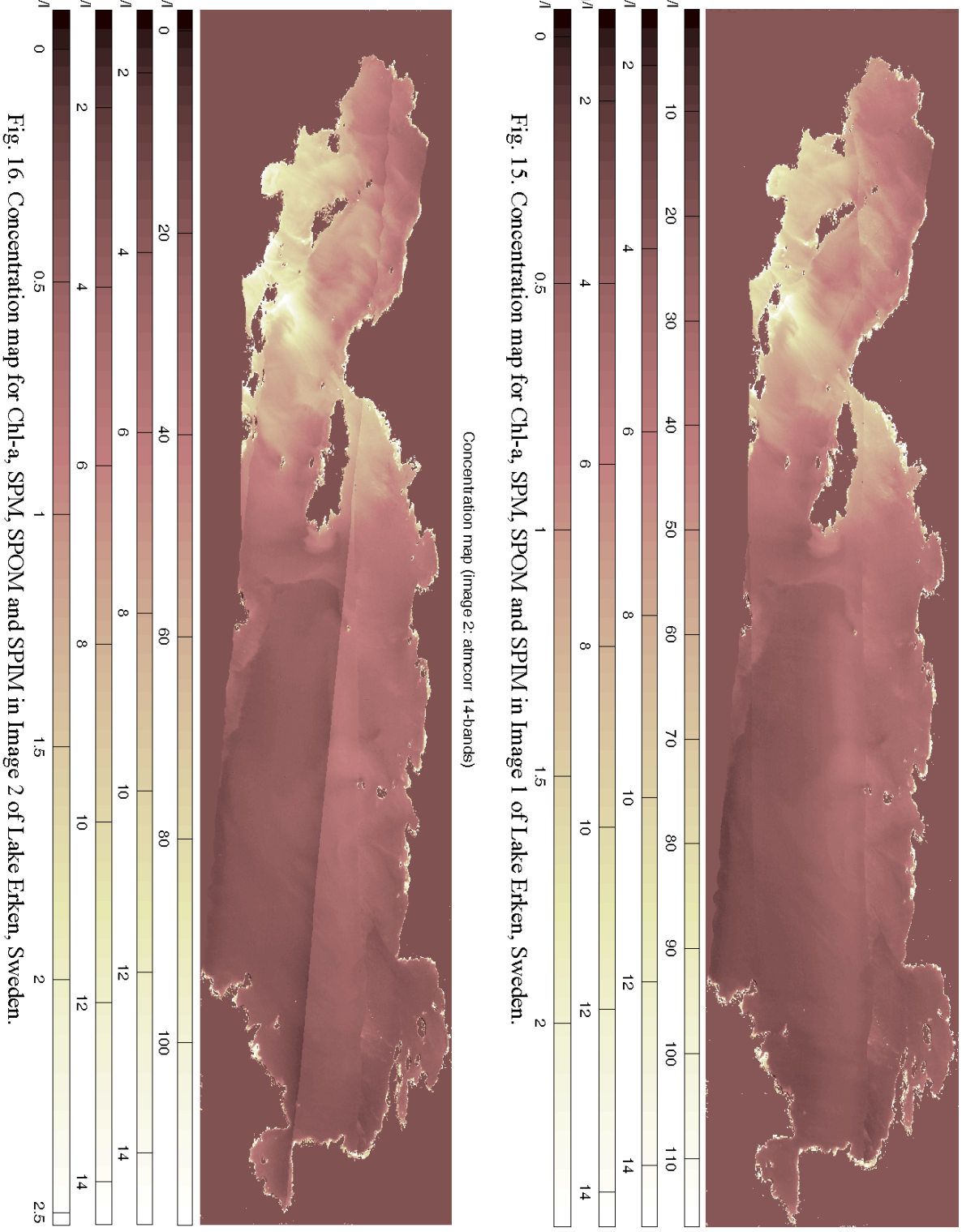

Fig. 15. Concentration map for Chl-a, SPM, SPOM and SPIM in Image 1 of Lake Erken, Sweden.

Fig. 16. Concentration map for Chl-a, SPM, SPOM and SPIM in Image 2 of Lake Erken, Sweden.

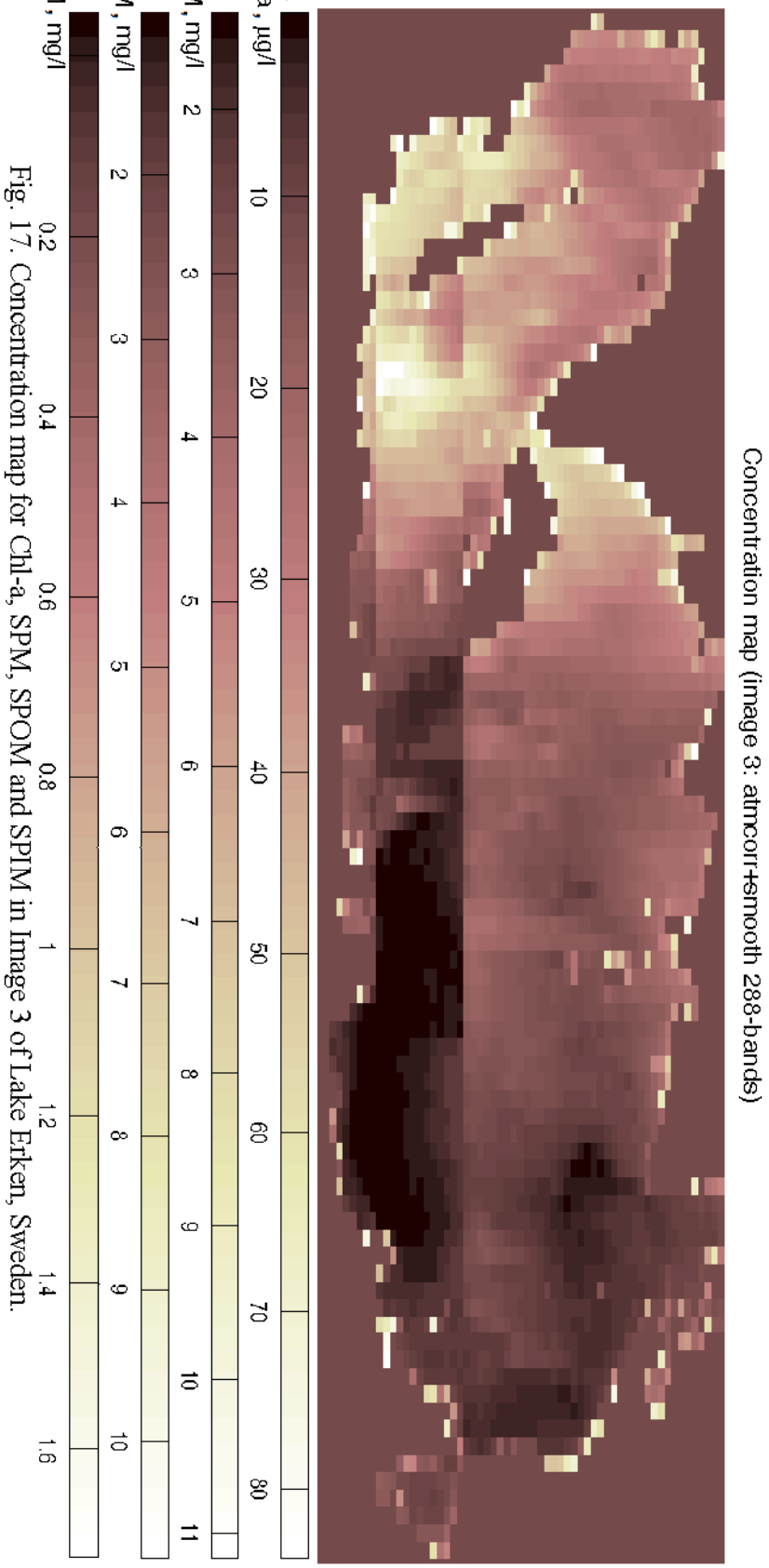

Fig. 17. Concentration map for Chl-a, SPM, SPOM and SPIM in Image 3 of Lake Erken, Sweden.

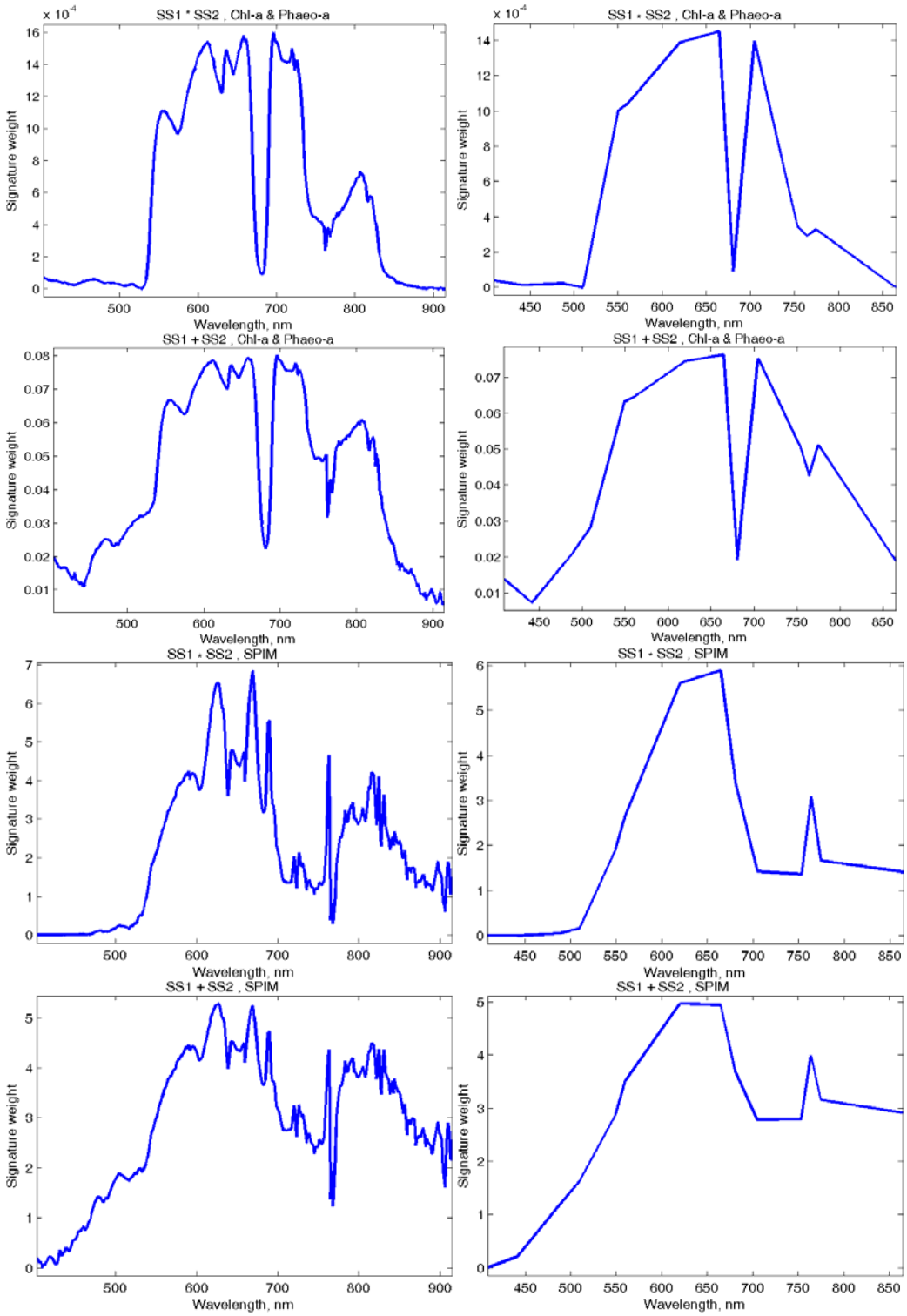

Fig. 18. The products (row 1) and the sums (row 2) of the descriptive spectral signatures in the case of utilising the 288-bands (left column) and the 14-bands (right column) field reflectance spectra.

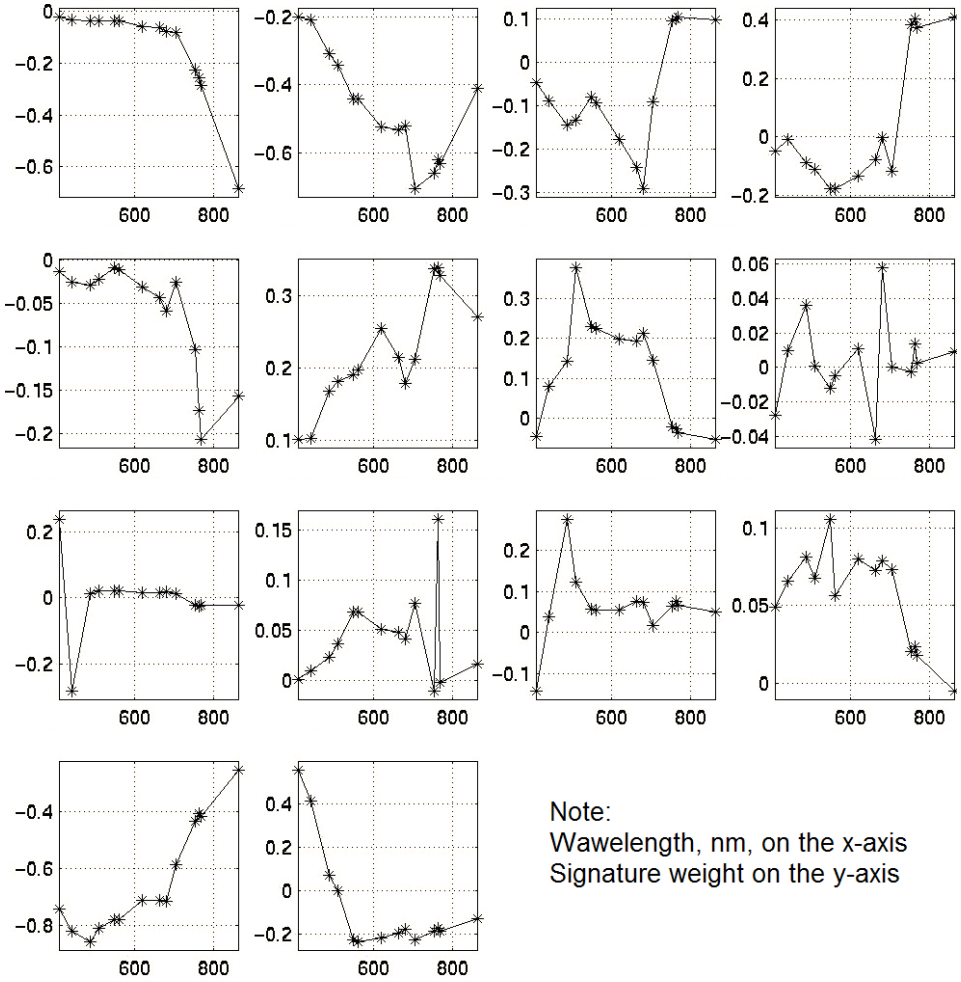

Fig. 19. Spectral signatures of applying linear ICA to Image 1.

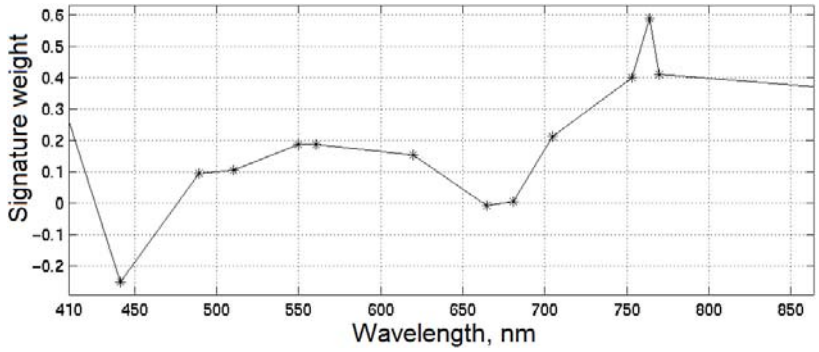

Fig. 20. Spectral signature for chlorophyll-*a*, obtained as a linear combination of the spectral signatures in fig. 19.

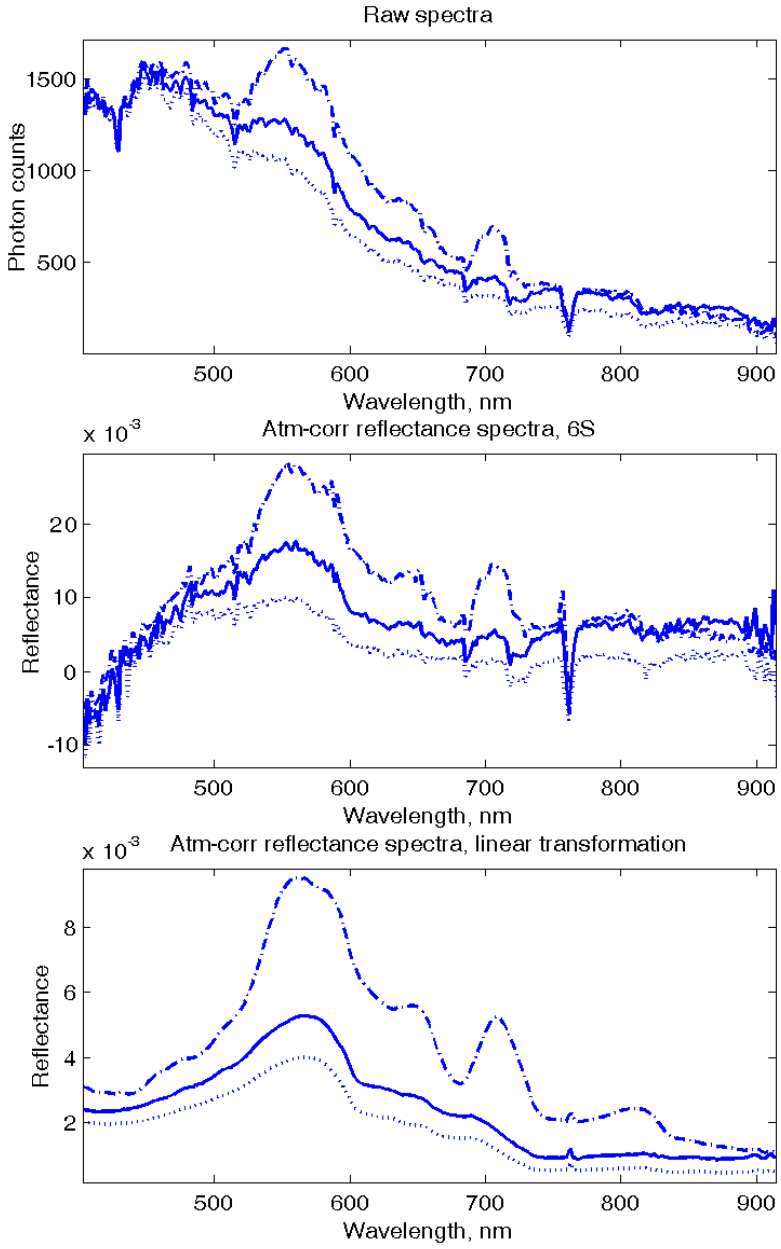

Fig. 21. Comparison between (row 1) three selected raw (upwelling radiance) spectra, and the corresponding atmospherically corrected reflectance spectra (row 2) when using the 6S-code, and (row 3) linear transformation.

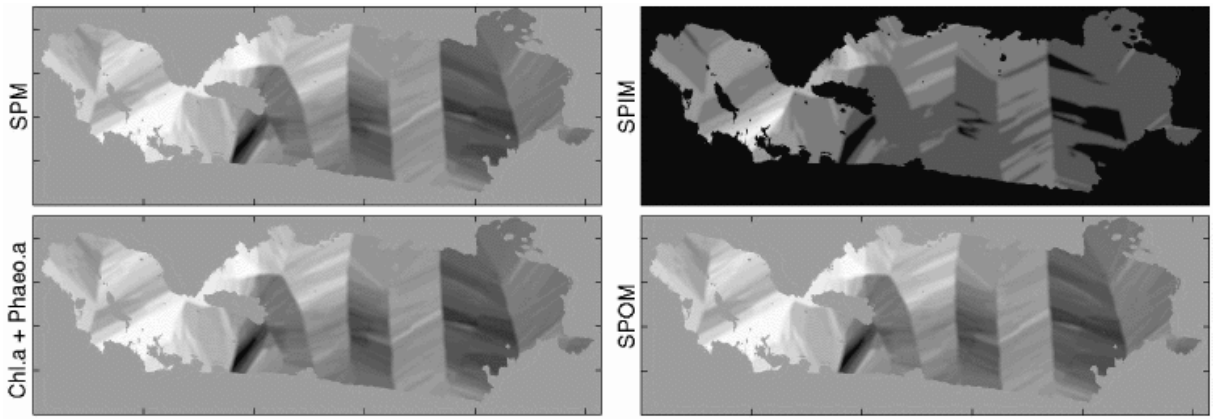

Fig. 22. Concentration maps for Chl-a, SPM, SPOM and SPIM in Lake Erken, generated by 2D-interpolation of the 1722 continuous field measurements.