Using Spectral Descriptive Signatures for Industrial Plume Detection

HAMED HAMID MUHAMMED [a], ABDOLAMIR KARBALAIE [a],
MOHAMMAD MEHDI MONTAZERI [b]

[a] Medical Technology Department
School of Technology and Health STH
Royal Institute of Technology KTH
SE-100 44, Stockholm
SWEDEN

[b] Department of Mathematics
Khomeini Shahr Branch, Islamic Azad University
Khomeini Shahr, Isfahan
IRAN

Email: hamed.muhammed@sth.kth.se abdolamir.karbalaie@sth.kth.se
URL: http://www.sth.kth.se/hamed/ Tel: +4687904855

Abstract: - This paper presents a novel approach for anomaly detection base on computing and utilizing descriptive spectral signatures. The goal of the work is to distinguish between contaminated and normal water areas within a region of investigation. A site-independent approach was developed by considering descriptive spectral signatures characterising normal sweat lake water as reference spectral features. Thereafter, it was possible to detect and determine the distribution of industrial outlet plumes which usually have spectral characteristics that deviate from the surrounding unaffected normal waters. The method was evaluated on airborne hyperspectral remotely-sensed image-data acquired over the region of Norrsundet, Sweden. In this region, areas of different water types were found, such as riverine sweet water, coastal salt seawater, as well as waste water discharged from paper-pulp industries. The work aimed at identifying these types of waters and their distributions. The needed reference descriptive spectral signatures of uncontaminated normal water were generated by using a dataset consisting of laboratory measurements of chlorophyll-*a* and phaeophytine-*a* concentrations and the corresponding field reflectance spectra collected at 22 sampling stations in Lake Erken, Sweden. The final results, showing the locations and distributions of contaminated and normal water areas, are in full agreement with field observations in the investigated region.

Key-Words: - Industrial plume detection, Remote sensing, Chlorophyll-*a*, Phaeophytine-*a*, Descriptive spectral signatures.

1 Introduction

The European Water Framework Directive (WFD; 2000/60/EC) requires that all member states should monitor all of their aquatic ecosystems. Therefore, the WFD requires new design methods to monitor all polluting substances discharged into the aquatic environment. Pollutants resulting from the paper and pulp industries are considered as serious polluting substances. The environmental effects of such effluents from the paper and pulp industries have been investigated by many researchers, such as Wilander et al. [54]; Hansson [20]; Kautsky [27]; Ekstrand [11]; Welch et al. [53]; Chambers et al. [6]; Culp et al. [9].

Serious physiological changes and disturbances have been seen in aquatic ecosystems exposed to pulp and paper mill effluent (PPME). These changes can produce numerous adverse impacts.

Serious side effects can be seen on fish populations and plants that live around the PPME [12, 38].

Graveline et al. [14] explored the potential uses and constrains of the screening methods, in respond to the new requirements of the WFD, for different hydrological and environmental conditions. Collection of samples and long-term monitoring of water areas, are often used to detect industrial contaminations by analyzing the spatial and temporal dynamics within water areas [52]. However, this type of analysis is time consuming and requires high costs.

On the other hand, remote sensing can be used to study a large area on Earth covering the whole region of interest. Multi- or hyperspectral image data, acquired by air-borne or satellite-borne imaging systems, are usually used to (visually or automatically) inspect and assess the water quality and detect areas with abnormal or contaminated waters. Remotely sensed multi- and hyperspectral data have been widely used to estimate major surface water quality variables such as chlorophyll-a, turbidity, suspended sediment concentration, Secchi disk depth, surface water temperature, wave height, and sea surface roughness, etc [2, 3, 4, 5, 10, 18, 19 13, 23, 24, 28, 33, 34, 35, 36, 37, 39, 42, 48, 56, 57, 58].

However, apart from linear regression and correlation studies, as proposed by Philipson et al. [41] and Jaruskova and Liska [26], rather little effort has been done in order to detect industrial plumes and contaminations in water bodies using remotely sensed multi- or hyperspectral data. Philipson et al. [41] proposed the use of the spectral angle mapper (SAM) for plume detection while Jaruskova and Liska [26] used non-parametrical methods, such as the locally weighted scatterplot smoothing, and also used parametric methods, such as linear regression for estimation of a trend of time series observations. Valent et al. [49] used ARMA models (autoregressive moving average models) to represent the linear time series models class. They also used SETAR models (self-exciting threshold autoregressive models) and MSW models (Markov switching models) to represent the nonlinear time series models class with multiple regimes. They found that the relative accuracy improvement of SETAR and MSW models when compared to ARMA models was high when using multiple-regimes modeling.

On the other hand, anomaly detection became increasingly promising and important when introducing the use of hyperspectral imagers which can resolve the spectral characteristics of many material substances more accurately than multispectral imagers. It is, for example, easier to identify a variety of natural and man-made material and to differentiate between them, by employing hyperspectral images comprising hundreds of contiguous bands [44]. Almost all anomaly detection methods attempt to locate anything that looks different, spatially or spectrally, from its surroundings. And this is what our new method does by using a rather different approach, as described in the next section.

However, automatic real-time anomaly detection is becoming increasingly important for many application fields. But the huge amount of data acquired by hyperspectral imagery, calls for utilizing parallel computing techniques. A cost-effective solution to achieve this goal is to implement the parallel algorithms on graphics processing units (GPUs). Paz and Plaza, [40], implemented several GPU-based anomaly and target detection algorithms for exploitation of hyperspectral data.

2 Spectral Anomaly Detection

In spectral anomaly detection algorithms, the pixels in the hyperspectral image (representing certain materials in the imaged scene) that have significantly different spectral signatures from their neighbouring background-clutter pixels are identified as spectral anomalies. Spectral anomaly detection algorithms [7, 30, 45, 46, 55] could also use spectral signatures to detect anomalies embedded within a background clutter with a very low signal-to-noise ratio SNR. In spectral anomaly detectors, no prior knowledge of the target spectral signature is utilized or assumed.

Two interesting anomaly detectors will be discussed in this work. One was developed by Reed and Yu [43] to detect targets whose signatures are distinct from their surroundings [1, 47]. This approach is referred to as the RX detector (RXD). The other one (which is proposed and developed in Harsanyi [21] and Harsanyi et al. [22]) was designed to detect targets with low probabilities in an unknown image scene. This approach is referred to as the low probability detector (LPD). Interestingly, both approaches operate as a matched

filter, but differ in two aspects. Firstly, the RXD uses the pixel currently being processed as the matched signal, while the LPD makes use of the unity vector. Secondly, the RXD uses the sample covariance matrix to take into account the sample spectral correlation, while the LPD makes use of the sample correlation matrix. A uniform target detector (UTD) is obtained when employing sample covariance matrix in LPD. The UTD can be used for background detection, making it efficient to use in combination with RXD to remove the background and detect the anomalies, as suggested by Ashton et al. [1]. This combined approach is referred to as RXD-UTD and can be mathematically presented and explained in matrix form as follows.

The RXD is specified by

$$\delta_{RXD}(\boldsymbol{s}) = (\boldsymbol{s}-\boldsymbol{\mu})^T \boldsymbol{C}^{-1} (\boldsymbol{s}-\boldsymbol{\mu}) \quad (1)$$

where $\boldsymbol{s}$ is a sample spectrum, $\boldsymbol{\mu}$ is global sample mean, and $\boldsymbol{C}$ is the sample covariance matrix of the image. $(\bullet)^T$ and $(\bullet)^{-1}$ denotes matrix transpose and inverse, respectively.

While LPD is defined as

$$\delta_{LPD}(\boldsymbol{s}) = (\boldsymbol{1}-\boldsymbol{\mu})^T \boldsymbol{R}^{-1} (\boldsymbol{s}-\boldsymbol{\mu}) \quad (2)$$

where $\boldsymbol{1}$ is unity vector, and $\boldsymbol{R}$ is the sample correlation matrix.

Consequently, the UTD becomes

$$\delta_{UTD}(\boldsymbol{s}) = (\boldsymbol{1}-\boldsymbol{\mu})^T \boldsymbol{C}^{-1} (\boldsymbol{s}-\boldsymbol{\mu}) \quad (3)$$

And the RXD-UTD becomes

$$\delta_{RXD\text{-}UTD}(\boldsymbol{s}) = (\boldsymbol{s}-\boldsymbol{1})^T \boldsymbol{C}^{-1} (\boldsymbol{s}-\boldsymbol{\mu}) \quad (4)$$

Chang et al. [7] suggested replacing $\boldsymbol{C}$ with $\boldsymbol{R}$ and $(\boldsymbol{s}-\boldsymbol{\mu})$ with $\boldsymbol{s}$ in Eqs. (1), (3) and (4) in order to enhance the performance of those detectors, since the new versions could account for both the first-order and second-order statistics. The same authors also presented and discussed (in Chiang et al. [8]) a number of efficient distance measures for anomaly classification, and derived a number of matched-filter-based target discrimination measures. The most interesting measure was the correlation matched-filter measure (RMFM) given as follows

$$\rho(\boldsymbol{s}_i, \boldsymbol{s}_j) = \boldsymbol{s}_i^T \boldsymbol{R}^{-1} \boldsymbol{s}_j \quad (5)$$

where $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ are two target spectra, and the larger the matched value ρ is, the more likely that the two targets, $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$, belong to the same class.

3 Descriptive Spectral Signatures

In Hamid Muhammed et al. [16], Larsolle et al. [31, 32] and Hamid Muhammed [17, 18, 19], descriptive spectral signatures, characterising the effect of the parameter of interest on the spectral properties of the target object, were extracted by using relatively small training data sets consisting of parameter measurements and the corresponding multi- or hyperspectral data (e.g. measured spectra).

The training spectral data were at first normalised into zero-mean and unit-variance data by using two iterative normalisation approaches. This type of normalisation is called whitening and the processed data is called whitened data. Explanation about the meaning of whitening of a dataset, how to apply it and how to make use of it can be found for example in Hyvärinen et al. [25] and Van Etten [50].

In our work, dataset normalisation is performed by using 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.

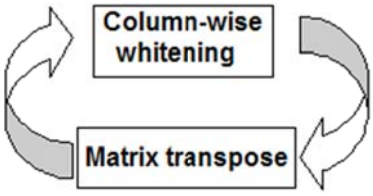

Fig. 1. Iterative normalisation.

Fig. 1 illustrates the iterative procedure, where spectral data are inserted as columns or rows (in the case of performing the first respectively the second approach) in a matrix $\boldsymbol{S}$ on which a series of alternating column-wise whitening and matrix

transposing operations are performed. In each approach a unique stationary result is achieved after a limited number of iterations.

Let matrices $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$ denote the resulting matrices from the first and the second approach, respectively. If a linear relationship is assumed between the measured parameters vector $\boldsymbol{p}$ and each of the resulting $\boldsymbol{S_1}$ or $\boldsymbol{S_2}$ matrices, then these two systems of linear equations can be written as follows

$$\boldsymbol{S_1}^{\mathrm{T}}\ \boldsymbol{t_1} = \boldsymbol{p} \quad (6)$$
$$\boldsymbol{S_2}^{\mathrm{T}}\ \boldsymbol{t_2} = \boldsymbol{p} \quad (7)$$

where $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$ are transformation vectors that can be computed by using the least squares method as follows

$$\boldsymbol{t_1} = \boldsymbol{S_1}\ (\boldsymbol{S_1}^{\mathrm{T}}\ \boldsymbol{S_1})^{-1}\ \boldsymbol{p} \quad (8)$$
$$\boldsymbol{t_2} = \boldsymbol{S_2}\ (\boldsymbol{S_2}^{\mathrm{T}}\ \boldsymbol{S_2})^{-1}\ \boldsymbol{p} \quad (9)$$

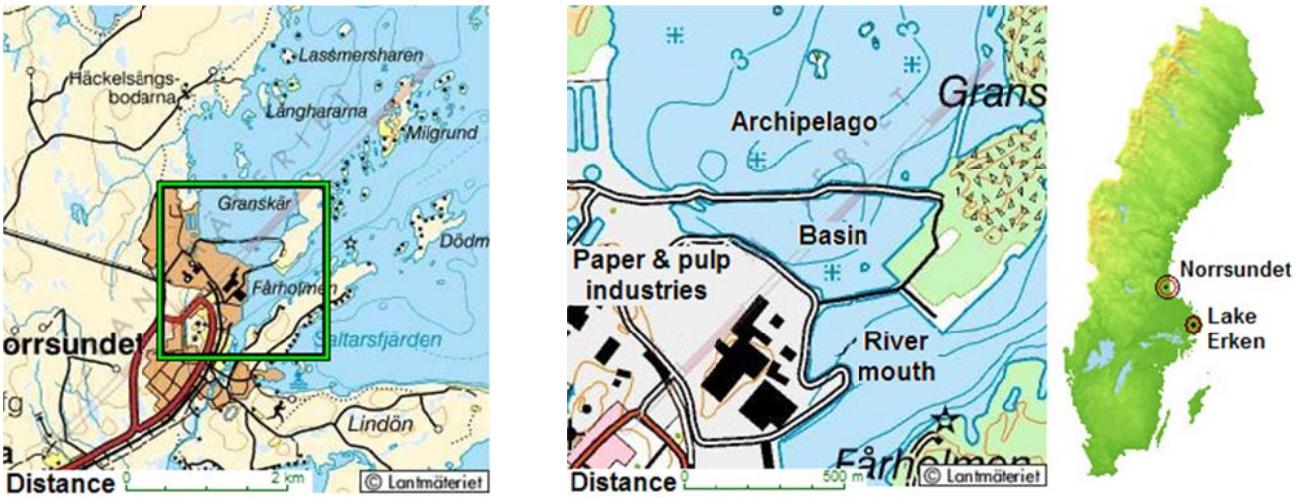

Fig. 2. Map over the region of Norrsundet, Sweden.

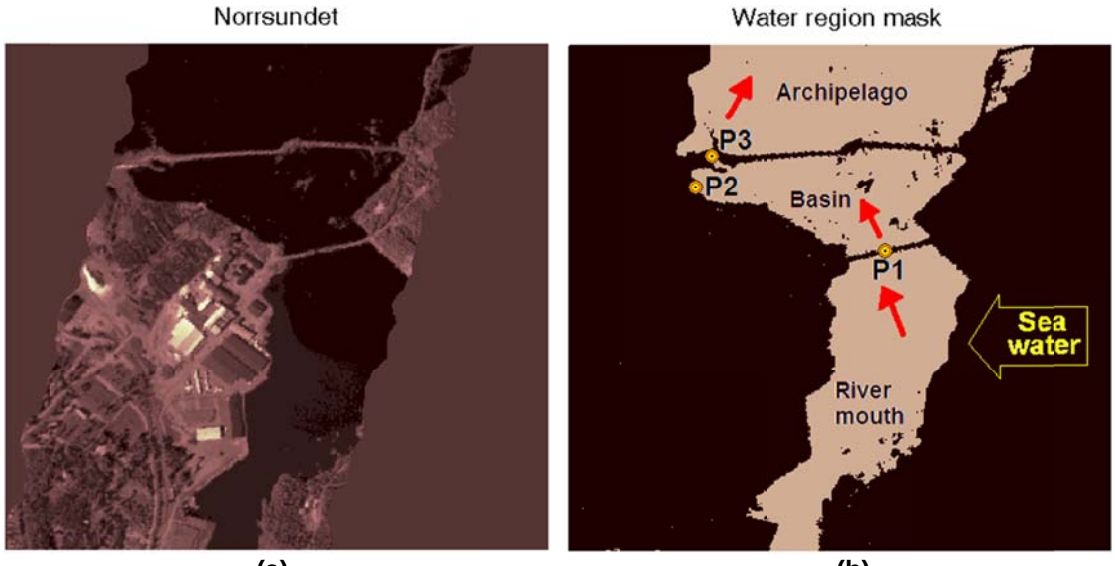

Fig. 3. (a) The mean 400×400-pixels sub-image covering the area of investigation. (b) The corresponding water-region mask with the marking points P1, P2 and P3 for the passage between the basin and river mouth, the waste-water discharge point into the basin, and the outlet point from the basin to the archipelago.

The two resulting transformation vectors, $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$, function as spectral signatures describing the changes in spectral characteristics with respect to the parameters of interest. The deduction of these equations is explained in Appendix A. In Hamid Muhammed et al. [16] , Larsolle et al. [31, 32] and Hamid Muhammed [17, 18], these signatures reveal the effect of increased disease severity on the spectral properties of wheat plants, while in Hamid Muhammed [18, 19], the signatures explain the impact of various water quality parameters on the spectral characteristics of lake water. A descriptive spectral signature pair, $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$, should be computed for each parameter of the set of parameters of interest.

4 Anomaly Detection Using Descriptive Spectral Signatures

The resulting descriptive spectral signature pairs, $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$, can be used to analyse new multi- or hyperspectral data (forming a spectral matrix $\boldsymbol{S_N}$) with respect to the desired parameters. After normalising the spectral matrix $\boldsymbol{S_N}$ according to the two approaches presented above, and producing $\boldsymbol{S_{N1}}$ and $\boldsymbol{S_{N2}}$ when using the first and the second approach, the appropriateness of the new data can be estimated as follows

$$\boldsymbol{a_1} = \boldsymbol{S_{N1}}^{\mathrm{T}}\ \boldsymbol{t_1} \quad (10)$$
$$\boldsymbol{a_2} = \boldsymbol{S_{N2}}^{\mathrm{T}}\ \boldsymbol{t_2} \quad (11)$$

where $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ are two estimates of the matching or correspondence between new and training data, with respect to the used $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$ pair, which correspond to a certain parameter vector $\boldsymbol{p}$.

Equations (10) and (11) can be written as

$$\boldsymbol{a_1} = \boldsymbol{S_{N1}}^{\mathrm{T}}\ \boldsymbol{S_1}\ (\boldsymbol{S_1}^{\mathrm{T}}\ \boldsymbol{S_1})^{-1}\ \boldsymbol{p} \quad (12)$$
$$\boldsymbol{a_2} = \boldsymbol{S_{N2}}^{\mathrm{T}}\ \boldsymbol{S_2}\ (\boldsymbol{S_2}^{\mathrm{T}}\ \boldsymbol{S_2})^{-1}\ \boldsymbol{p} \quad (13)$$

Noting that the terms $(\boldsymbol{S_1}^{\mathrm{T}}\ \boldsymbol{S_1})^{-1}$ and $(\boldsymbol{S_2}^{\mathrm{T}}\ \boldsymbol{S_2})^{-1}$ represent the correlation matrices of the multi- or hyperspectral data samples, it can easily be seen that equations (12) and (13) compute the RMFM correlation measure, presented in eq. (5), between the training parameter measurements vector $\boldsymbol{p}$ and $(\boldsymbol{S_{N1}}^{\mathrm{T}}\ \boldsymbol{S_1})$ or $(\boldsymbol{S_{N2}}^{\mathrm{T}}\ \boldsymbol{S_2})$ which represent the correlations between the training normalised spectral matrices, $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$, and the new normalised spectral matrices, $\boldsymbol{S_{N1}}$ and $\boldsymbol{S_{N2}}$.

This means that high $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ values should be obtained for target spectra, from matrix $\boldsymbol{S_N}$, belonging to the same class as the spectra in matrix $\boldsymbol{S}$. With other words, it can be said that the classes of matrices $\boldsymbol{S_N}$ and $\boldsymbol{S}$ are close to each other. Consequently, low $\boldsymbol{a_1}$ and/or $\boldsymbol{a_2}$ values indicate that the corresponding spectra can be classified as anomalous when compared to the spectra in matrix $\boldsymbol{S}$. Or that the classes of matrices $\boldsymbol{S_N}$ and $\boldsymbol{S}$ are different from each other and not that close when compared to the first case of high $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ values.

5 Image and Ground Truth Data

Remotely sensed hyperspectral image data has been acquired, by using the Compact Airborne Spectrographic Imager (CASI), over the region of Norrsundet (Fig. 2) in Sweden, during a CASI-campaign in August 1997. Norrsundet is located 30 kilometers north of the city of Gavle. The waters in this region were affected by an outlet from the small Hamrangean river and the waste water outlet from the Norrsundet paper-pulp industry, as shown in Fig. 2.

A CASI spatial-mode image, covering this area of investigation, was acquired on August 5th, 1997. This CASI image had 10 spectral bands, a spatial resolution of 4×4 meters, and a size of 400×400 pixels. Fig. 3a shows a gray-scale image where each pixel value is the mean value of the corresponding 10 pixel values of the CASI image. Table 1 presents the band settings of the CASI spatial-mode which are similar to that of the MERIS sensor on Envisat.

Table 1. Wavelength band-set definition for the CASI in spatial mode.

Band No	Start wavwlength [nm]	End wavwlength [nm]
1	403.5	415.6
2	436.5	446.9
3	483.7	494.2
4	504.8	515.3
5	545.3	554.2
6	614.5	625.2
7	659.0	669.8
8	676.9	684.1
9	700.2	709.1
10	750.3	755.7

Another dataset was acuired one day later, on August 6th, 1997, but at another site, namely Lake Erken, located about 180 kilometers south-east of Norrsundet. At this site, water samples were collected from 22 sampling stations and analysed in laboratory to measure the concentration of chlorophyll-*a* and phaeophytine-*a* in these samples.

The measured concentrations varied between 2.9 – 50.6 μg/l. In addition to that, a handheld Dual GER 1500 spectroradiometer was used to measure the up- and down-welling radiance above the lake surface at all sampling stations, and finally compute the reflectance spectra at these stations. The measured spectra had 513 spectral bands in the wavelength range 400 – 900 nm.

6 Area of Investigation

Fig. 3a presents the mean image covering the investigated area, while Fig. 3b shows the corresponding water-region mask with the marking points P1, P2 and P3. P1 marks the passage between the river mouth and the basin, into which concentrated waste-water from the paper-pulp industry is discharged at the point marked P2, while the outlet point from the basin to the archipelago is marked P3, through which waste water is mechanically forced, causing a compensation water flow through P1 from the river mouth into the basin. The thin (red) arrows indicate the direction of water flow in the area of investigation, in addition to the thick (yellow) arrow labelled with 'Sea water' indicating the sea water input.

7 Image Pre-Processing

The CASI data were geometrically corrected and radiometrically calibrated at delivery. These data were then atmospherically corrected by using the 6S-code [51], which compensated for the atmospheric effects and converted the data (representing upwelling radiance at the sensor) into ground reflectance.

Thereafter, the mean image was computed for the 10-bands image (Fig. 3a), and global thresholding was applied to the mean image to identify and extract the water-region hyperspectral pixels from the 10-bands image (Fig. 3b).

8 Experimental Results and Discussion

The site of Norrsundet has been investigated previously by Philipson et al. [41] by employing the spectral angle mapper (SAM) to classify the same CASI image data used in this paper. Absorbance measurements of concentrated outlet water have shown clearly different spectral characteristics when compared to natural inland waters.

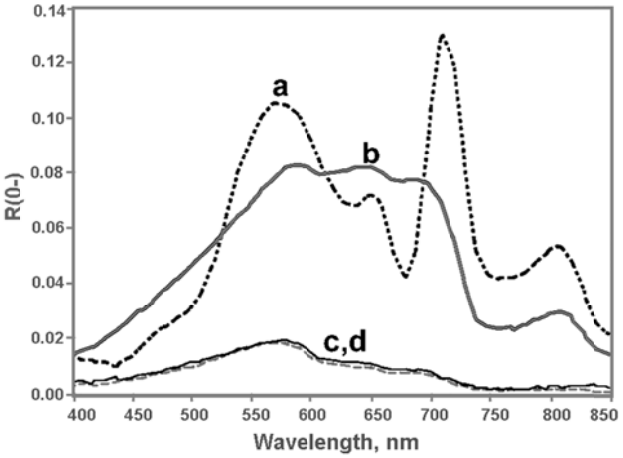

Fig. 4. Reflectance spectra of three major water types: (a) with high phytoplankton and SPM concentrations, (b) with low phytoplankton but high SPM concentrations,

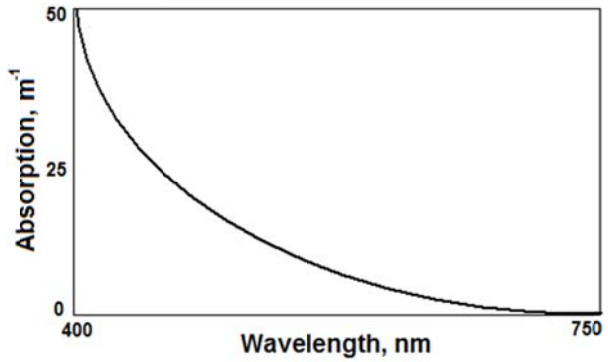

Fig. 5. Absorption spectrum of concentrated outlet water. Reproduced from Philipson et al. 2005.

Fig. 4, which is reproduced from Dekker [10], shows typical reflectance spectra of three major water types. In this figure, spectral differences can easily be observed between water samples with high phytoplankton and SPM concentrations, other samples with low phytoplankton but high SPM concentrations, and water samples with low phytoplankton and SPM concentrations.

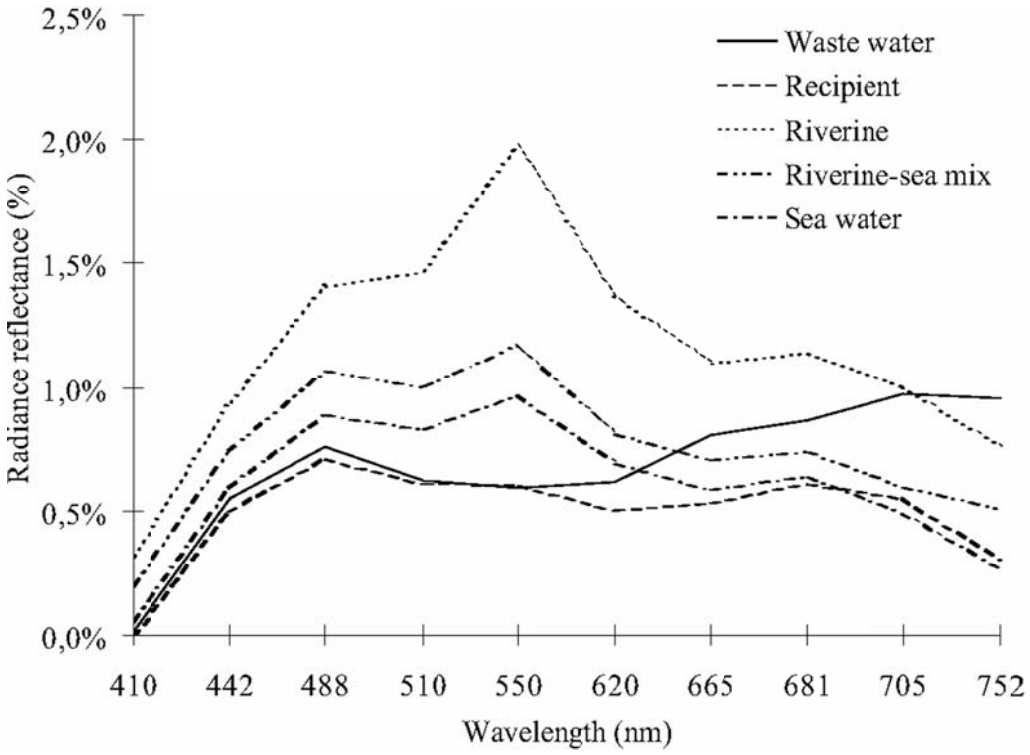

Fig. 6. Reference CASI-image spectra for SAM classification. Reproduced from Philipson et al. 2005.

Furthermore, comparing the spectral characteristics of these water types with those of the concentrated outlet water shows more significant spectral differences. Fig. 5, which is reproduced from Philipson et al. [41], shows the absorption spectrum of concentrated outlet water. In this figure, it is clear that the reflectance increases with the wavelength.

In Philipson et al. [41], the investigation of this CASI image also showed clear differences in the spectral characteristics captured by CASI pixels located in five different areas in the investigated region. Fig. 6, which is also reproduced from Philipson et al. [41], shows the five reference reflectance spectra that were chosen from the CASI image to be used for SAM classification.

The conclusion that can be drawn from these observations is that spectral differences do exist between spectra corresponding to different water types. To be able to efficiently reveal these differences, it is important to perform both band- and pixel-wise comparisons. Band-wise comparison (denoted by Bw) reveals amplitude differences between different spectra, while within spectrum variations are revealed when performing pixel-wise comparison (denoted by Pw).

This motivates to use the two iterative normalisation approaches described previously in this work and illustrated in Fig. 1.

These normalisation approaches (using series of "Pw, Bw, .. Pw" or "Bw, Pw, .. Pw" normalisations), were applied to the field reflectance spectra measured on water samples from Lake Erken. Fig. 7 shows four reflectance spectra of water samples with various concentrations of chlorophyll-*a* and phaeophytine-*a*, collected from four different sampling stations in the lake.

The resulting $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$ matrices (containing normalised field spectra) from the first and the second normalisation approaches, respectively, were employed in the systems of linear equations in equations (6) and (7). In these equations, the parameter vector $\boldsymbol{p}$ consists of the measured concentrations of chlorophyll-*a* and phaeophytine-*a*, which correspond to the measured field spectra. The transformation vectors $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$ were computed by using the least squares method as described by equations (8) and (9). These vectors (presented in Fig. 8) represent descriptive spectral signatures characterising the effect of increased concentrations

of chlorophyll-*a* and phaeophytine-*a* on the spectral properties of the water.

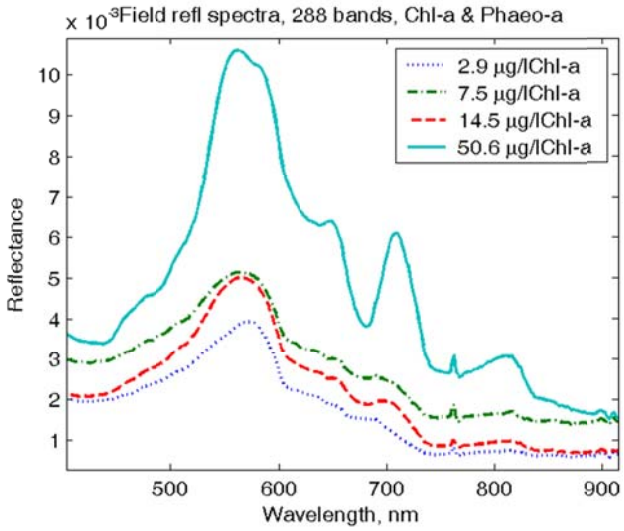

Fig. 7. Reflectance spectra of water samples with various concentrations of chlorophyll-*a* and phaeophytine-*a*, collected from four sampling stations in Lake Erken.

Thereafter, the CASI image data were also normalised by using exactly the same two normalisation approaches used previously. By this way, the two matrices $\boldsymbol{S_{N1}}$ and $\boldsymbol{S_{N2}}$, containing normalised image spectra, were obtained. The resulting matrices were used in equations (10) and (11) to estimate $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ which can be interpreted as two-dimensional (2D) maps showing the correspondence between $\boldsymbol{S_1}$ and $\boldsymbol{S_{N1}}$ and between $\boldsymbol{S_2}$ and $\boldsymbol{S_{N2}}$, respectively. Higher values in $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ can be associated with high correspondence between $\boldsymbol{S_i}$ and $\boldsymbol{S_{Ni}}$ (where i is 1 or 2). Fig. 9 presents the resulting $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ 2D maps as well as the resulting summation 2D-map ($\boldsymbol{a_1}$ + $\boldsymbol{a_2}$) in addition to the resulting element-wise-multiplication 2D-map ($\boldsymbol{a_1}$ - $\boldsymbol{\mu_1}$) × ($\boldsymbol{a_2}$ - $\boldsymbol{\mu_2}$), where $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2}$ are the mean values of $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$, respectively. In all 2D maps, in Fig. 9, lower values are shown with darker colours.

Poor correspondence between $\boldsymbol{S_i}$ and $\boldsymbol{S_{Ni}}$ (where i is 1 or 2) can be detected by utilizing both of the summation 2D-map and the multiplication 2D-map. This task can be performed by finding the low values in the summation 2D-map (corresponding to low values in both $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$) and also by identifying the negative values in the multiplication 2D-map, corresponding to high positive $\boldsymbol{a_1}$ values and low negative $\boldsymbol{a_2}$ values, or low negative $\boldsymbol{a_1}$ values and high positive $\boldsymbol{a_2}$ values. For efficient visualisation, the background (which is empty and not included in the computations) of the figure presenting the multiplication 2D-map was set to zero, making the negative values darker than the background, while the positive values are brighter.

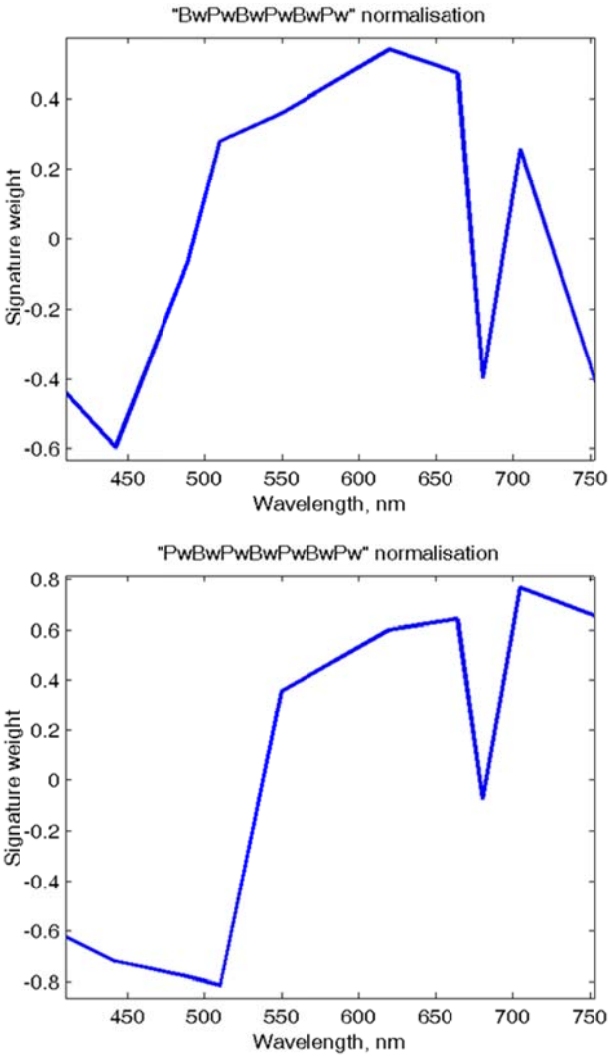

Fig. 8. Descriptive spectral signatures.

Fig. 10 presents a 2D-map generated by classifying and segmenting the image of the area of investigation by using the $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ 2D-maps. In the result shown in Fig. 10, class 1 represents pixels corresponding to high $\boldsymbol{a_1}$ values but low $\boldsymbol{a_2}$ values, class 2 corresponds to low values in both of the $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ maps, class 3 corresponds to low $\boldsymbol{a_1}$ values but high $\boldsymbol{a_2}$ values, and finally, class 4 corresponds to high values in both of the $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ maps. The mean values of $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ were used as a threshold to determine if a value was considered as high or low in the corresponding maps.

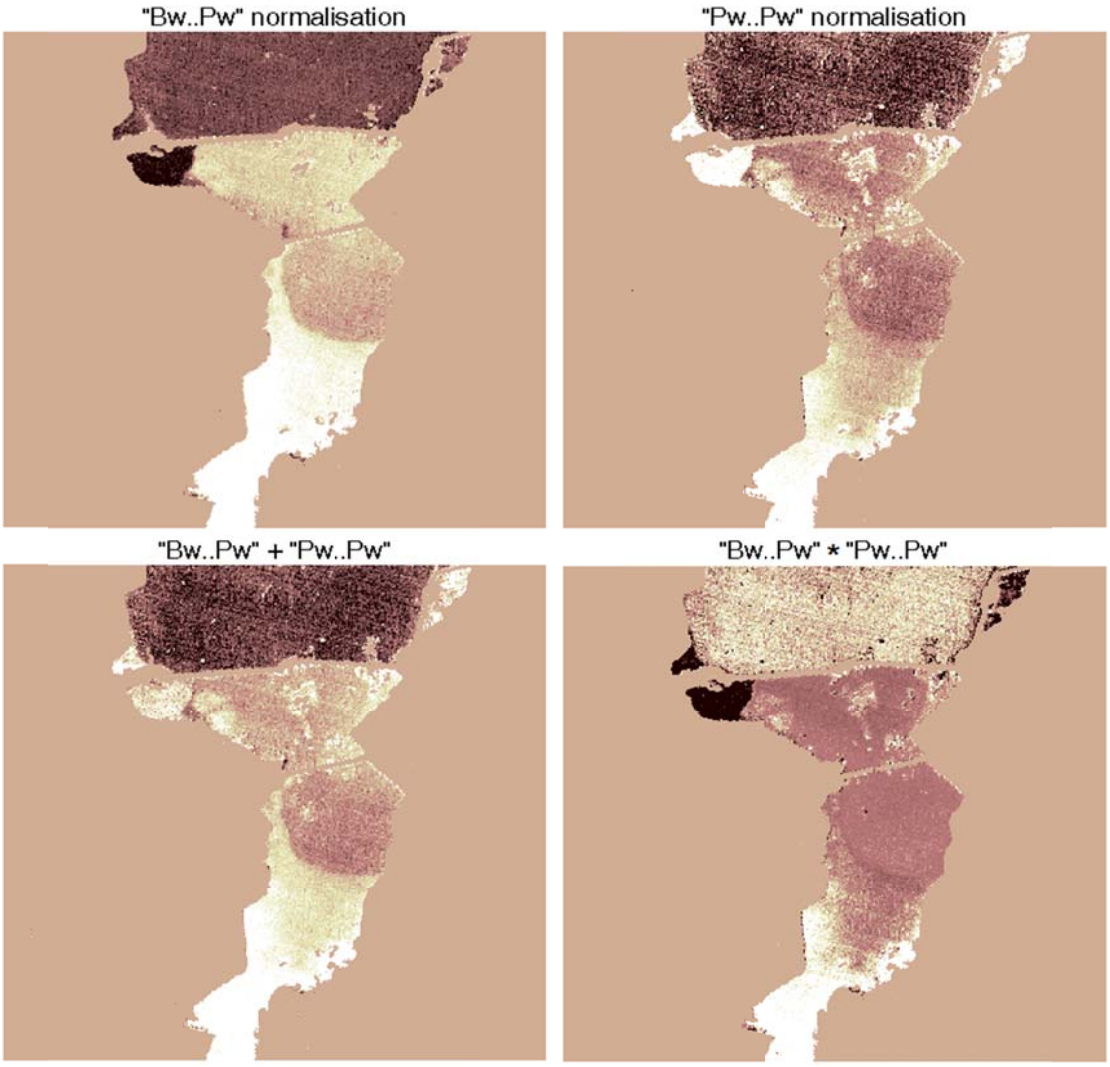

Fig. 9. The resulting a_1 (upper right), a_2 (upper left), the sum ($a_1 + a_2$) (down left), and the element-wise multiplication ($a_1 \times a_2$) (down right).

Obviously, class 4 represents water unaffected by the industrial waste water, since there good correspondence between this water type and the uncontaminated water in Lake Erken. On the other hand, classes 1, 2 and 3 represent contaminated waters. Comparison with the map of the Norrsundet region in Fig. 3 results in the important observations that class-1 waters are mainly found near the waste-water discharge point P2 into the basin, and the outlet P3 point from the basin to the recipient (the archipelago), class-3 waters are found in the basin and also in the river mouth region near the passage between the basin and river mouth (it can be clearly seen how contaminated waters flow southwards through the passage into the river mouth), while class-2 waters are mainly found in the recipient (with coastal seawater) and also observed in the area where a mixture between riverine and coastal seawater is found. Furthermore, in this figure, the sea water injection can be clearly seen, in addition to the sharp boundary where riverine water meets coastal seawater.

9 Summary and Conclusions

This work presents a novel site- and sensor independent approach for water quality assessment and differentiation between contaminated and normal waters within an industrial region. For this purpose, a pair of descriptive spectral signatures characterising uncontaminated lake water were computed and used as a reference expressing normal water spectral characteristics. By employing a new anomaly detection technique which utilizes these descriptive spectral signatures,

water bodies affected by industrial waste water having spectral characteristics that usually deviate from the surrounding unaffected water can easily be detected.

The new approach was tested on airborne hyperspectral remote sensing data acquired over the region of Norrsundet, Sweden, where paper-pulp industries are located. On the other hand, normal-water descriptive spectral signatures were derived by using another data set consisting of laboratory measurements of chlorophyll-*a* and phaeophytine-*a* concentrations and the corresponding field reflectance spectra. These data were collected from Lake Erken, Sweden, about 180 km south-east of the region of Norrsundet.

Systems of linear equations were assumed and used to describe the relationship between spectral data and water quality measures. Solving this kind of equations when using the Lake Erken dataset resulted in two descriptive spectral signatures.

Each one of these two descriptive spectral signatures was used to produce a 2D map showing the correspondence between the dataset of Lake Erken (which represents uncontaminated waters) and the CASI image data of the region of Norrsundet.

High correspondence values in both of the two maps indicate that the corresponding water is of the same type as the water in Lake-Erken. Low correspondence values in one or both of the two maps indicate anomalous water type (different, in some sense, from the water in Lake-Erken). Classification using these simple rules produced a 4-classes map that could be used not only to detect and map the waste water, but also to analyse the dynamics and types of the water system in the region of Norrsundet, where the waste water functioned as an efficient tracer.

Acknowledgements
This work was financed by a research grant from the Swedish National Space Board (SNSB). Dr. Hamed hamid Muhammed would like to thank his former supervisors, Prof. Ewert Bengtsson and Ass. Prof. Tommy Lindell for their support and for providing the datasets used in this work.

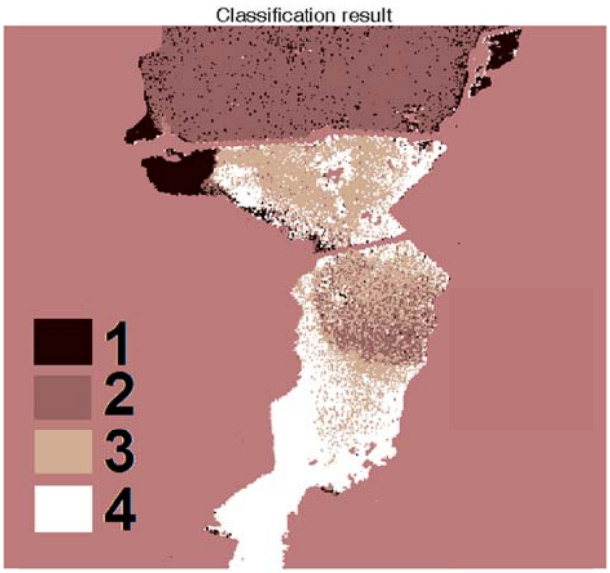

Fig. 10. Classification result showing four types of water.

References:

[1] E.A. Ashton, A. Schaum, Algorithms for the detection of sub-pixel targets in multispectral imagery, *Photogram. Eng. Remote Sens*, 1998, pp.723–731.

[2] P.J. Baruah, K. Oki, H. Nishimura, 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*, 2000.

[3] R.P. Bukata, J.H. Jerome,Y.A. Kondratyev K,V.D. Pozdnyakov, 1991: Satellite monitoring of optically-active components of inland waters: an essential input to regional climate change impact studies, *J. Great Lakes Res*., Vol.17, No.4, 1991, pp. 470-478.

[4] R.P. Bukata, J.H. Jerome, K.Y.A. Kondratyev, D.V. Pozdnyakov, *Optical Properties and Remote Sensing of Inland and Coastal Waters*, CRC Press Inc., Florida, 1995.

[5] S.H. Cairns, K.L. Dickson, S.F. Atkinson, 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] P.A. Chambers, A.R. Dale, G.J. Scrimgeour, M.L. Bothwell, Nutrient enrichment of northern rivers in response to pulp mill and municipal discharges, *Journal of Aquatic*

Ecosystem Stress and Recovery, Vol.8, No.1, 2000, pp. 53-66.

[7] C.I. Chang, S.S. Chiang, Anomaly detection and classification for hyperspectral imagery, *IEEE Trans. Geosci. Remote Sensing*, Vol.40, No.6, 2002, pp1314–1325.

[8] S.S. Chiang, C.I. Chang, Discrimination measures for target classification, *In Proc. IEEE IGARSS*, Sydney, Australia, 2001.

[9] J.M. Culp, C.L. Podemski, K.J. Cash, Interactive effects of nutrients and contaminants from pulp mill effluents on riverine benthos, *J. Aquat. Ecosyst. Stress Recovery*, Vol.8, No.1, 2000, pp. 67-75.

[10] A.G. Dekker, *detection of optical water quality parameters for eutrophic waters by high resolution remote sensing*, Ph.D Thesis. Free University, Amsterdam, 1993, pp. 78.

[11] S. Ekstrand, *Determination of algal production in an industrial waste water plume using Landsat TM satellite imagery*, Report, IVL Swedish Environmental Research Institute, Stockholm, Sweden, 1998.

[12] C.A. Flinders, G.W. Minshall, R.L. Ragsdale, Tj. Hall, Patterns of macro invertebrate assemblages in a long-term watershed-scale study to address the effects of pulp and paper mill discharges in four US receiving streams, *Integr Environ Assess Manag*., Vol.5, No.2, 2009,248–258.

[13] C. Giardino, M. Pepe, P.A. Brivio, P. Ghezzi, E. Zilioli, 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.

[14] N. Graveline, L. Maton, J. Rinaudo, H. Lukge, E. Interwies, J. Rouillard, P. Strosser, K. Palkaniete, D. Taverne, An operational perspective on potential uses and constraints of emerging tools for monitoring water quality, *Trends Analyt., Chem.* Vol.29, 2010, pp. 378–384.

[15] H. Hamid Muhammed, Using Hyperspectral Reflectance Data for Discrimination Between Healthy and Diseased Plants, and Determination of Damage-Level in Diseased Plants, *31st Applied Imagery Pattern Recognition Workshop*, Washington DC, USA. IEEE Computer Society Press. 2002.

[16] H. Hamid Muhammed, A. Larsolle, Feature Vector Based Analysis of Hyperspectral Crop Reflectance Data for Discrimination and Quantification of Fungal Disease Severity in Wheat. *Biosystems Engineering*, Vol. 86, No. 2, 2003, pp. 125-134.

[17] H. Hamid Muhammed, Hyperspectral Crop Reflectance Data for characterising and estimating Fungal Disease Severity in Wheat, *Biosystems Engineering*, Vol.91, No.1, 2005, pp. 9-20.

[18] H. Hamid Muhammed, *Hyperspectral Image Generation, Processing and Analysis*, PhD thesis. Centre for Image Analysis, Uppsala University, Sweden, Acta Universitatis Upsaliensis, 2005.

[19] H. Hamid Muhammed, New approaches for surface water quality estimation in Lake Erken, Sweden, using remotely sensed hyperspectral data, *WSEAS Transactions on Environment and Development*, Vol.7, No.10, 2011, pp. 285-314.

[20] S. Hansson, Effects of pulp and paper mill effluents on coastal fish communities in the Gulf of Bothnia, *Baltic Sea. Ambio*, Vol.16, No.6, 1987, pp. 344-348.

[21] J.C. Harsanyi, *Detection and classification of subpixel spectral signatures in hyperspectral image sequences*, Ph.D. dissertation. Dept. Elect. Eng., Univ. Maryland-Baltimore County, Baltimore, MD, 1993.

[22] J.C. Harsanyi, W. Farrand, C.I. Chang, Detection of subpixel spectral signatures in hyperspectral image sequences, in *Proc. Amer. Soc. Photogram. Remote Sens. Reno*, NV, 1994, pp. 236–247.

[23] P. D. Hunter, A.N Tyler, L. Carvalho, G.A. CoddA., S.C. Maberly, Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes, *Remote Sensing of Environment*, Vol.114, No.11, 2010, pp. 2705-2718.

[24] C. Hu, Z. Lee, R. Ma, K. Yu, D. Li, S. Shang, MODIS observations of cyanobacteria blooms in Taihu Lake. China. Marine Science Faculty Publications, 2010, pp. 40.

[25] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and

applications, *Neural Networks*, Vol.13, No.4–5, 2000, pp. 411-430.

[26] D. Jaruskova, I. Liska, Statistical analysis of trends in organic pollution and pollution by nutrients at selected Danube river stations, *J. Environ. Monit.*, Vol.13, 2011, pp. 1435-1445.

[27] H. Kautsky, The impact of pulp-mill effluents on phytobentic communities in the Baltic Sea, *Ambio.*, Vol.12, No.4, 1992, pp. 308-311.

[28] S. Koponen, J. Pulliainen, H. Servomaa, Y. Zhang, M. Hallikainen, K. Kallio, J. Vepsäläinen, T. Pyhälahti, T. Hannonen, 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.

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

[30] H. Kwon, S.Z. Der, N.M. Nasrabadi, Adaptive anomaly detection using subspace separation for hyperspectral images, *Optical Engineering*, Vol.42, No.11, 2003, pp. 3342–3351.

[31] A. Larsolle, H. Hamid Muhammed, Measuring crop status using multivariate analysis of hyperspectral field reflectance with application on disease severity and amount of plant density. Precision Agriculture '05, edited by *J.V. Stafford, Proceedings of 5th European Conference on Precision Agriculture (5ECPA)*, Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden. 2005, pp. 217-225.

[32] A. Larsolle, H. Hamid Muhammed, Measuring crop status using multivariate analysis of hyperspectral field reflectance with application to disease severity and plant density, *Precision Agriculture*, Vol.8, No.1-2, 2007, pp. 37-47.

[33] R.G. Jr. Lathrop, T.M. Lillesand, 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.

[34] R.G. Jr. Lathrop, J.R. Vande Castle, T.M. Lillesand, Monitoring river plume transport and mesoscale circulation in Green Bay. Lake Michigan, through satellite remote sensing, *Journal of Great Lakes Research*, Vol.16, No.43, 1990, pp. 471-478.

[35] R.G.Jr. Lathrop, T.M. Lillesand, B.S. Yandell, 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.

[36] T.M. Lillesand, W.L. Johnson, R.L. Deuell, O.M. Lindstrom, D.E. Miesner, Use of Landsat data to predict trophic status of Minnesota lakes, *Photogramm. Eng. Remote Sens.*, Vol.49, 1983, pp. 219-229.

[37] C.R. Mcclain, M.L. Cleave, G.C. Feldman, W.W. Gregg, S.B. Hooker, Science quality SeaWiFS data for global biosphere research. *Sea Tech.*, Vol.39, 1998, pp.10-16.

[38] F. Mower, K. R. Munkittrick, M. E. Mcmaster, R. J. Van Beneden, Response of white sucker (Catostomus commersoni) to pulp and paper mill effluent in the Androscoggin River, Maine, USA. *Environmental Toxicology and Chemistry*, Vol.30, 2011, pp. 142–153.

[39] L.G. Olmanson, P.L. Bresonick, S.M. Kloiber, M.E. Bauer, E.E. Day, *Lake Water Clarity Assessment of Minnesota's 10,000 Lakes: A Comprehensive View from Space*, Water Resources Center, University of Minnesota, 2000.

[40] A. Paz, A. Plaza, Gpu implementation of target and anomaly detection algorithms for remotely sensed hyperspectral image analysis. SPIE Optics and Photonics, *Satellite Data Compression, Communication, and Processing Conference*, 2010.

[41] P. Philipson, M. Liljeberg, T. Lindell, Industrial plume detection in hyperspectral remote sensing data. *International Journal of Remote Sensing*, Vol.26, No.2, 2005, pp. 295-313.

[42] K. Randolph, J. Wilson, L. Tedesco, L. Li, D. L. Pascual, E. Soyeux, Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments chlorophyll and phycocyanin. *Remote Sensing of Environment*, Vol.112, 2008, pp. 4009−4019.

[43] I.S. Reed, X. Yu, Adaptive multiple-band CFAR detection of an optical pattern with

unknown spectral distribution. *IEEE Trans. Acoust. Speech, Signal Processing*, Vol.38, 1990, pp. 1760–1770.

[44] G. Shaw, H. Burke, Spectral imaging for remote sensing, *Lincoln Lab. J.*, Vol.14, No.1, 2003, pp. 3–28.

[45] D.W.J. Stein, Stochastic compositional models applied to subpixel analysis of hyperspectral imagery. *In Proc. SPIE 4480*, 2001, pp. 49–56.

[46] D.W.J. Stein, S.G. Beaven, L.E. Hoff, E.M Winter, A.P. Schaum, A.D. Stocker, Anomaly detection from hyperspectral imagery. *IEEE Signal Processing Mag.*, Vol.19, 2002, pp. 58–69.

[47] C.M. Stellman, G.G. Hazel, F. Bucholtz, J.V. Michalowicz, A. Stocker, W. Scaaf, Real-time hyperspectral detection and cuing, *Opt. Eng.*, Vol.39, 2000, pp. 1928–1935.

[48] J. Tao, Z. Zhang, W. Yu,: Monitoring Taihu water quality by using high resolution satellite image. Geoinformatics, *19th International Conference on*, 2011, pp.1-4.

[49] P. Valent, N.J.K. Howden, J. Szolgay, M. Komornikova, Analysis of nitrate concentrations using nonlinear time series models. *J. of Hydrology and Hydromechanics*, Vol.59, No.3, 2011, pp. 157-170.

[50] W.C. Van Etten, *In Introduction to Random Signals and Noise*. John Wiley & Sons, Ltd, Chichester, UK, 2006.

[51] E.F. Vermote, D. Tanré, J.L. Deuzé, M. Herman, J.J. Morcrette, Second simulation of the satellite signal in the solar spectrum, 6S, An Overview. *IEEE Transactions on Geoscience and Remote Sensing*, Vol.35, No.3, 1997, pp. 675-686.

[52] L. Volaufová, J. Langhammer, Specific pollution of surface water and sediments in the Klabava River basin. *J. Hydrol. Hydromech.*, Vol 55, No 2, 2007, pp 122–134.

[53] E.B. Welch, T. Lindell, *Ecological effects of waste water*. Cambridge University Press, Revised edition, 1992.

[54] A. Wilander, H. Kvarnäs, T. Lindell, A modified fluorometric method for measurement of lignin sulfonates and its in situ application in natural waters, *Water Research*, Vol.8, No.12, 1974, pp. 1037-1045

[55] X. Yu, I.S. Reed, A.D. Stocker, Comparative performance analysis of adaptive multispectral detectors. *IEEE Trans. Signal Process*, Vol.41, No.8, 1993, pp. 2639–2656.

[56] Y. Zhang, *Surface water quality estimation using remote sensing in the gulf of Finland and the Finnish archipelago sea*, Ph.D. thesis, Helsinki University of Technology, Laboratory of Space Technology, Finland, 2005.

[57] Y. Zhang, J. Pulliainen, S. Koponen, M. Hallikainen, Application of an empirical neural network to surface water quality estimation in the Gulf of Finland using combined optical data and microwave data. *Remote Sensing of Environment*, Vol.81, No.2-3, 2002, pp. 327-336.

[58] Y. Zhang, L. Feng, J. Li, L. Luo ,Y. Yin ,M. Liu,Y. Li, Seasonal–spatial variation and remote sensing of phytoplankton absorption in Lake Taihu, a large eutrophic and shallow lake in China. *J. Plankton Res.*, Vol.32, No.7, 2010, 1023-1037.