

## DETERMINATION OF WATER QUALITY PARAMETERS IN INDIAN PONDS USING REMOTE SENSING METHODS

*Mathias Kneubühler<sup>1</sup>, Christian Gemperli<sup>1</sup>, Daniel Schläpfer<sup>1</sup>, Rainer Zah<sup>2</sup> and Klaus Itten<sup>1</sup>*

1. Remote Sensing Laboratories (RSL), University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland, kneub@geo.unizh.ch
2. Swiss Federal Laboratories for Materials Testing and Research (EMPA), Lerchenfeldstrasse 5, 9014 St. Gallen, Switzerland

### ABSTRACT

This paper presents a water quality study performed on five small domestic ponds and one bayou of the Ganga river in the State of West Bengal, India. The concept of eutrophication is used to describe water quality by using chlorophyll as an indicator for the presence of algae in the water.

During a field campaign in May 2003, reflectance spectra were measured 1m above the water surface using a handheld spectroradiometer. Concentration of *Total Chlorophyll Content* (TCHL) and other limnological standard parameters such as *Secchi Disk Depth* (SDD), *nutrient content*, *Total Organic Content* (TOC) and *Dissolved Organic Content* (DOC) were determined using laboratory analysis and a colorimetric testkit. A comparison of several existing semi-empirical algorithms to determine chlorophyll content was made by applying them to the spectra and chlorophyll measurements collected in-situ. *Simple Reflectance Band Ratio* and *Continuum Interpolated Band Ratio* (CIBR) were tested at several wavelengths. Another three tested algorithms rely on the reflectance peak at 700 nm whose shape and position depend strongly upon chlorophyll concentration: The *Area* and *Magnitude of the Peak above a Baseline* as well as the *Position of the Peak* show a linear relationship to CHL-a concentration from laboratory spectrophotometric measurements. All algorithms proved to be of value. Best results were obtained by using the algorithms *Peak Magnitude above a Baseline* and *CIBR*, which yielded  $R^2$  around 0.97 and mean deviations around 20 µg/l.

The algorithms were further applied to spectra convolved to the bands of the hyperspectral sensors APEX, HyMap and Hyperion. A comparison of the spectral characteristics of the chosen sensors was made in regard of their suitability for water quality determination. The differences on the outcome of the sensors were small, but the results indicate that an intermediate band width of around 10 nm seems appropriate. All sensors showed an adequate band positioning for the algorithms based on the peak near 700 nm. Using *Simple Band Ratios* and *CIBR*, APEX and Hyperion perform slightly better than HyMap.

### 1. INTRODUCTION

The scope of this study is to investigate the potential of algorithms for monitoring water quality parameters in West Bengal ponds, many of them being used for jute retting. The concept of eutrophication was chosen to describe water quality, using chlorophyll as indicator for the presence of algae in the water. A semi-empirical approach was used to determine the parameter. The algorithms chosen from the literature had to perform well in other studies on water bodies with comparable limnological characteristics.

The aims of the study are:

- the assessment of the trophic state of the ponds,
- the evaluation of semi-empirical algorithms for determining CHL-a content using spectral reflectance data collected with a handheld field-spectroradiometer,

- the application of these algorithms to the spectra convolved to the bands of selected sensors,
- the evaluation of the most suitable sensors and applied algorithms for determining CHL-a content in the study area.

### **1.1. Water Quality**

There is no single definition of water quality, as it depends on the respective processes and on the intended use of the water. In this work the focus is laid on the process of eutrophication. Trophic state is an absolute scale that describes the biological condition of a water body and does not imply a normative statement per se. It has to be stressed that trophic state is not the same thing as water quality but trophic state certainly is one aspect of water quality.

#### **Eutrophication**

The term 'eutrophication' means nutrient enrichment of lakes and describes the processes in water caused through over-enrichment by nutrients. Too much nutrient input causes a chain of events that may have undesirable effects on lakes, e.g. growth of algae and macrophytes. The enhanced rates of decomposition and attendant consumption of oxygen lead finally to oxygen depletion in the hypolimnion, decreasing the number of species. Eutrophication can occur under natural conditions, but often the process is induced by anthropological factors, such as fertilizers and sewage [4]. Organic matter introduced to surface waters (e.g. jute), leads to enhanced rates of decomposition and attendant consumption of oxygen.

## **2. DATA ACQUISITION BASELINE**

### **2.1. Study Area**

The study area is situated in the state of West Bengal (Nadia District), India. Agriculture plays a pivotal role in the state's income, and nearly three out of four persons in the state are directly or indirectly involved in agriculture. The state accounted for 66.5 % of the country's jute production including mesta in 1993-1994.

The summer months are from March to June. The monsoon season lasts from June to September and brings heavy rain. The monsoon brings respite to the parched plains but they often cause floods and landslides. The winter months are from October to February. Summer temperatures range from 24°C to 40°C and winter temperatures from 7°C to 26°C. The yearly average rainfall is around 175 cm.

Nadia District is located in the low-land basin of the Ganga river northeast of the state's capital, Kolkata. As it can be seen in Fig.1 (left), the rural area is scattered with small ponds and bayous (oxbow lakes) of the Hooghly river which crosses the district in the south. Their sizes vary between a few square meters to several square kilometers. Many of them are man made, serving as water reservoirs for various purposes, such as washing, bathing, irrigation, fishing and jute retting.

Between May 23, and June 6, 2003, a field campaign was organized in the above mentioned area. The aims of the campaign were to collect physical-chemical parameters and spectral data on jute fields and water bodies of this region. As there was very little knowledge about the water bodies and the general circumstances prior to the campaign, the selection of the water bodies had to be made during the campaign. The following criteria were considered:

- All ponds must be suitable for jute retting, being a major cause for eutrophication.
- Minimum size must be 150 m x 100 m.
- The water bodies must be easily accessible.
- The permission of the owner must be obtained.

A broad range of estimated reflectance spectra should be in the samples of the selected water bodies (estimation through water color and turbidity).

All ponds have neither an in- nor an outlet and are only fed by rainwater and by periodical flooding during monsoon season. It can be assumed that the limnological cycle varies strongly during the year. In autumn, the heavy monsoon floods fill the ponds and large amounts of nutrients are probably washed into the ponds from the neighboring fields. At this time, the water bodies have been described to be very clear by the local farmers. In spring, with rising temperatures, bio-activity of the ponds rise and evaporation and human use causes the water level to fall several meters.

## **2.2. Field Measurement Data**

### **2.2.1. Spectral Data**

The radiometric characterization of the selected ponds was determined by using a FieldSpec Pro FR Spectoradiometer from Analytical Spectral Devices (ASD) [1]. Reflectance measurements were made at 400 to 2500 nm at a spectral resolution of 1nm, with a fore-optic FOV of 8°. Spectra were recorded using a white reference panel (spectralon) to obtain absolute reflectance values. The spectral measurements took place between 10 am and 14 pm. The sky conditions were for the most parts hazy with some high-level clouds, as it is typical before the arrival of the monsoon. Spectral data was taken from 6 ponds, four of them were covered with 1-2 transects using a boat (P1, P2, P5, P6). At two ponds (P3, P4), measurements were made only at the shoreline. At each pond, approximately 150 single spectra were recorded (see Fig.1).

### **2.2.2. Water Constituents Data**

It was the aim of the campaign to get a broad overview of the general state of the water bodies, upon which several limnological parameters were determined. The limnological sampling was performed on the same day as spectral characterization of the water. In all ponds, one water sample was taken at a location in the middle of the water body; in bigger ponds, an additional sample was taken nearer to the shoreline. Due to the shallowness of the ponds and low SDD values, mixed samples were taken from the surface to SDD (Integrated SDD). Storage of the samples determined in India was done according to the procedure described by Lindell et al. [16]. Immediately after sampling and during transport, all samples were stored cool using ice. Within 4 hours, they were either brought to the laboratory and deep frozen until parameter determination or analyzed in the field using a colorimetric testkit.

*Fig. 1: Left: Landsat ETM panchromatic image of the study site, acquired on April 22, 2002; right: averaged spectra of each water body. The numbers indicate its CHL concentration in  $\mu\text{g/L}$ . The concentration in brackets was determined one week before and was not included into analysis.*

The determination of TCHL, DOC and TOC was done by the Department of Agricultural Chemicals, University of Bidhan Chandra Krishi Viswavidyalaya, Mohanpur. It was intended to use the standard method according ISO 10260 [16] for determination of CHL-a, but as the laboratory was not used to this method, Annon's method was applied. Chlorophyll content of the samples was

assessed by spectrophotometric determination of the acetone extract of the sample at 645 and 663 nm. Dissolved organic carbon was determined by using the dichromate oxidation method of Vance et al. Total organic carbon was determined using the Nelson Somer's method. Secchi Disk Depth was recorded using the procedure described by Lindell et al. [16].

Hence, all of the ponds were very shallow at the time of the campaign. Maximum measured depth did not exceed 1m in all of the ponds. Large amounts of yellow substances and inorganic solids are suspended in the water leading to a very poor visibility. SDD did not exceed 0.12 m in the domestic ponds, in P5 it was approx. 0.9 m (Fig.1). In the domestic ponds, (all ponds, excluding P5) measured concentrations of TCHL were extremely high, most with contents between 300 and 520 µg/L (see Fig.1, right). Such high concentrations normally characterize a bloom situation, which should be indicated by scum on the surface. This could only be observed in P3.

A concentration of 0 µg/L was determined for both samples of P5, but analysis of the spectra show that there must be at least small contents of chlorophyll. Either laboratory analysis failed or the concentrations were below the detection limit. Regarding these contradictions, there rests an uncertainty about the accuracy of the chlorophyll determination at the laboratory.

Given the fact that the vast majority of algae in a water body are cyanobacteria that only contain CHL-a [16], the TCHL concentration measured in the laboratory is assumed to represent the concentration of CHL-a, solely.

A variety of further limnological parameters (e.g., nutrient content, organic carbon, pH, total hardness, heavy metals, water temperature), that were also determined during the field campaign are not presented in this paper.

### **3. METHODOLOGY OF PARAMETER RETRIEVAL**

#### **3.1. Methods used in this study**

A wide range of different algorithms and models can be found in literature to estimate chlorophyll content in Case II waters. The semi-analytical algorithms applied in this work are all based on the use of reflectance data in the red and NIR range of the electromagnetic spectrum, since other portions of the spectrum are not suitable for chlorophyll estimation in Case II waters [6][10].

Since there was no knowledge about the general state of the water bodies, appropriate models had to be selected after the campaign. Semi-empirical models were chosen as an intermediate stage between the complex physical approaches and empirical approaches. Since chemical-physical in-situ data indicated a hypertrophic state of most of the sampled water bodies, the focus had to be laid on models which were already successfully tested in previous studies on water bodies with very high CHL content.

However, very few studies can be found in literature which have dealt with hyper-eutrophic water bodies with CHL concentrations over 100 µg/L. The only study which used semi-empirical algorithms other than simple band ratios, was done by Gitelson. He applied several algorithms, which are all related to the peak near 700 nm. All of them were successfully used to estimate CHL concentrations in water bodies of different trophic states, e.g. in wastewater ponds, where chlorophytes algae dominated with extremely high CHL-a concentrations ranging from 70 to 520 µg/L [8][9][10][19].

Simple band ratios were chosen because they are robust and easily applicable. In addition, they had already been applied in numerous studies on water bodies of all trophic states [6][10][13][14][16][20].

Finally the Continuum Interleaved Band Ratio (CIBR) was selected because it was used in a previous study to estimate chlorophyll concentration in inland water [14]. However, in contrary to the Indian ponds, the CHL concentration in that study from a Swiss lake was very low. Thus, it is necessary to verify if the algorithm could be applied on water bodies with high concentrations.

### 3.2. Chlorophyll Determination Algorithms

#### Simple Band Ratio

Band ratios determine the amount of absorption by the use of a measurement channel, which is divided through a reference channel.

$$BR = \frac{R(\lambda_m)}{R(\lambda_r)}$$

where  $R$  is the Reflectance,  $m$  indicates the measurement channel and  $r$  the reference channel.

The reflectance in the measurement channel should only be influenced by the absorption of the parameter. The reflectance of the reference channel should not be influenced by any kind of water constituents [14].

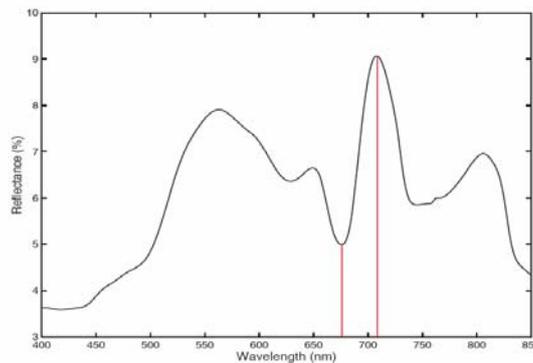


Fig. 2: Simple Band Ratio Algorithm.

Various authors such as Dekker [6], Koponen et al. [13] and Yacobi et al. [22] concluded that for the retrieval of CHL-a concentration, a ratio of channels centered at about 675 and 705 nm is appropriate in several lake types ranging from oligotrophic to hypertrophic state (Fig.2). As the band combination depends on the chlorophyll content range of the water bodies and therefore on the position of the reflectance peak, the used wavelength of the algorithm have to be slightly modified in most cases [12].

#### Continuum Interpolated Band Ratio (CIBR)

The Continuum Interpolated Band Ratio uses two reference bands instead of one. The advantage of the method over a simple band ratio is that the unequal influence of scattering by other water constituents on the reference and the absorption band is eliminated.

$$CIBR = \frac{R(\lambda_m)}{c_1 R(\lambda_{r1}) + c_2 R(\lambda_{r2})}$$

where  $R$  is the reflectance at a given wavelength  $\lambda$ , index  $m$  indicates the absorption band.  $r_1$  and  $r_2$  indicate the two reference channels, given through

$$c_1 = \frac{(\lambda_1 - \lambda_m)}{(\lambda_{r2} - \lambda_{r2})}, c_2 = \frac{(\lambda_m - \lambda_{r1})}{(\lambda_{r2} - \lambda_{r1})}$$

By using linear interpolation between two reference bands at each side of the absorption band, a reference value with 'normalized influence of scattering' is simulated on the absorption band. The absorption value is then divided by a reference value similarly affected by scattering, diminishing the influence of scattering on the ratio [14]. This method, originally applied by Bruegge et al. [3] in order to estimate water vapor in the atmosphere, was used for the first time by Kurer [14] for determination of CHL-a concentration.

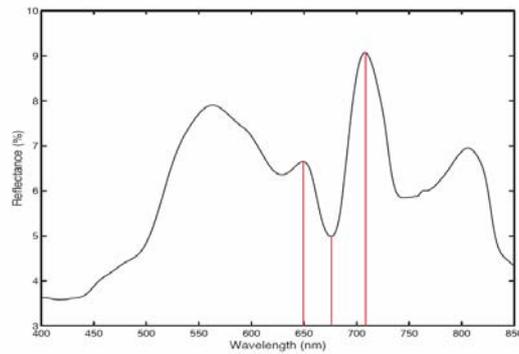


Fig. 3: Continuum Interpolated Band Ratio (CIBR).

### Area above a Baseline

Gitelson et al. [10] proposed an algorithm based on the area delimited by the reflectance curve and a baseline drawn from 670 to 750 nm (Fig.4):

$$A = \int_a^b f(x) - g(x) dx ,$$

with  $[a,b]$  as delimiters on the reflectance curve and  $f, g$  on  $[a,b]$  with  $f(x) \geq g(x) \geq 0$ .

The developed algorithm was tested and validated by Gitelson et al. [10] in several water bodies ranging from oligotrophic to hypertrophic state. They found linear relationships between the area above the base line and CHL concentration with high  $R^2$  around 0.9. The algorithms were also used to determine CHL content in wastewater with CHL concentrations ranging between 70-520  $\mu\text{g/L}$ , which is comparable to the content of the ponds in West Bengal. They yielded correlation coefficients  $R^2$  of 0.88, with a standard error of 59  $\mu\text{g/L}$  [10].

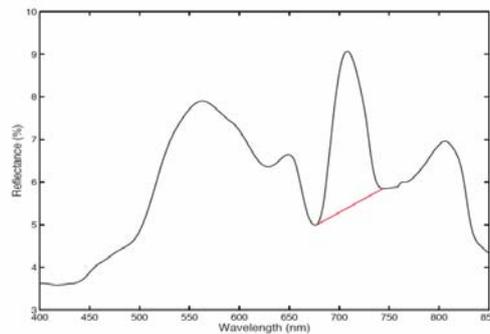


Fig. 4: Base Line Algorithm.

Generally, the trough at 670 nm shows minimum sensitivity to CHL concentration. According to studies by Gitelson et al. [8], Dekker [6], Yacobi et al. [22], the reflectance at this wavelength primarily depends on concentration of non-organic suspended matter for a CHL-a concentration higher than 20  $\text{mg/m}^3$ . Reflectance beyond 750 nm is insensitive to algal pigments and the variation of  $R_{750}$  is comparatively small because of strong water absorption in the NIR range. Therefore, the slope of the baseline between 670 and 750 nm depends primarily on scattering by water constituents, except phytoplankton; variations in concentration of non-organic and non-pigmented organic suspended matter changes the slope of the baseline, but the influence on the height and area of the 700 nm-peak above the base line remains small [10]. Thiemann et al. [20], who used the same algorithm on German lakes, asserted that the base line also serves as a correction for a higher reflectance signal due to wave action and differing angles between water surface and the spectrometer.

### Peak Magnitude above a Baseline

The height of the mentioned peak above a baseline was related to CHL-a concentration for the first time by Neville and Gower [17]:

$$M_{Peak} = f(\lambda_m) - g(\lambda_m) ,$$

where  $f(\lambda_m)$  is the spectrum curve and  $g(\lambda_m)$  the baseline;  $(\lambda_m)$  defines the wavelength of the maximum reflectance of the peak.

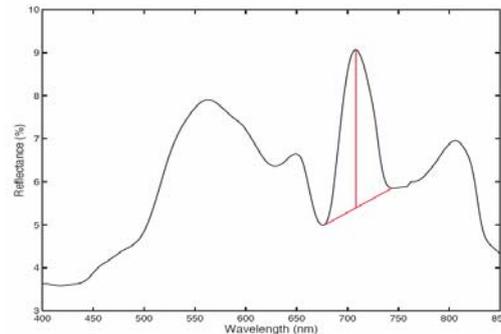


Fig. 5: Peak Magnitude above a Baseline Algorithm by Gitelson.

This algorithm correlates with chlorophyll concentration via the link between chlorophyll and algal biomass. However, it is not fully clear how the peak is correlated to phytoplankton concentration. According to Gitelson [10], the magnitude of the peak depends on scattering by all suspended matter and thus increases with increase of phytoplankton biomass. An increase of the algal biomass than leads to an increase of the active cell surface and thus to more scattering and higher reflectance values.

### Position of the Peak near 700 nm

The spectral position of the mentioned peak on the axis representing the wavelength is closely related to the chlorophyll content of the water body. Increasing absorption by chlorophyll leads to an offset of cell scattering at progressively higher wavelengths and the position of the peak shifts toward longer wavelengths [10][22].

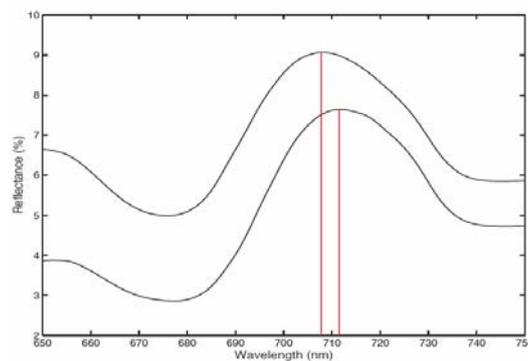


Fig. 6: Peak Position Algorithm by Gitelson.

This shift, measured in nm, can be used as precise indicator for chlorophyll content. Unfortunately, the algorithm requires very narrow bands at a short sampling interval due to the short spectral range where the shift of the peak takes place. In an extensive study Gitelson [7] obtained  $R^2$  values of over 0.9 and estimation errors of less than 2 nm for all water bodies.

### 3.3. Statistical methods

Semi-empirical algorithms use physical knowledge for the determination of the model but determine certain coefficients by means of statistics. Mostly the model is linked to the in-situ data by

means of regression analysis to derive the empirical coefficients. This way, the model is calibrated to the spatial or seasonal characteristics of the study area. In the literature, the Coefficient of Determination and the Standard Error is mostly taken to describe the quality of the regression. Within these measures, the residues become squared, which weights outlier stronger. As very few samples were available in this study, one has to act with caution with these measures. As a better measure, the mean deviation of the residues is taken instead. Nevertheless, the Coefficient of Determination and the Standard Error are calculated for better comparison with other studies in this field.

The Coefficient of Determination ( $R^2$ ) expresses how much of the variation (percent of the total sum of squares) is explained by the regression equation. Therefore, the regression sum of squares (explained deviation) is divided by the total sum of squares (total deviation) [2]:

$$R^2 = \frac{eS_y^2}{S_y^2} ,$$

where  $eS_y^2$  is the variance explained through the regression and  $S_y^2$  is the total variance.

The Standard Error of the Estimate (SSE) is a measure of the accuracy of predictions made by a regression line. It is defined as the square root of the average squared deviation from the regression line,

$$g_{est} = \sqrt{\frac{\sum (y - y^{\circ})^2}{N}} ,$$

where  $y$  are the measured and  $y'$  the predicted values,  $N$  is the number of pairs of  $(x,y)$  points.

As it can be seen in the above equation, the more observations you have and the less variable the raw data is, the smaller is SSE and the more accurate is the estimate [11][2].

The Mean Deviation ( $MD_a$ ) is defined as the mean of the absolute deviations of a set of data divided by the data's mean. For a sample size  $N$ , the mean deviation is defined by:

$$MD_a = \frac{1}{N} \sum_{i=1}^N |y_i - y_{i,e}| ,$$

where  $y_i$  are the measured values and  $y_{i,e}$  are the estimated values by the regression equation.

To express the Mean Deviation as a Percentage ( $MD_{\%}$ ), the deviations are divided by the sum of the estimated values:

$$MD_{\%} = \frac{1}{N} \frac{\sum_{i=1}^N |y_i - y_{i,e}|}{\sum_{i=1}^N |y_{i,e}|} .$$

## 4. RESULTS OF PARAMETER DETERMINATION

### 4.1. Simple Band Ratios

Because the exact positions of any two channels can be variable, they have to be adjusted for each wavelength region [15]. As a consequence, several different band ratios were tested. The ratios were first applied to the raw, un-normalized spectra gathered with the handheld spectroradiometer using a linear least square fit. Even though the  $R^2$  indicates a good correlation, analysis of variance (ANOVA) showed that the results were not statistically significant at a confidence level of 5%.

Applying the algorithm to the spectra normalized at 900 nm, the correlation coefficients improved but were still not significant at a confidence level of 5% (Tab.1, left).

By using a logarithmic curve, the fit between the variables improved. An ANOVA test showed significant results for all used band combinations at a significance level of 5%. The best result was again obtained using the ratio [702/672]. The  $R^2$  for that ratio is now above 90 per cent and the absolute error below 40  $\mu\text{g/L}$ .

Table 1: Results for simple band ratios, left: linear fit, normalized at 900nm, right: logarithmic fit.

Wavelength	$R^2$	Std error	$MD_R$	$MD_a$	Wavelength	$R^2$	Std error	$MD_R$	$MD_a$
705/678	0.83	101.5	0.25	70.44	711/681	0.94	57.7	0.14	39.95
711/681	0.81	106.6	0.26	74.55	711/665	0.96	49.9	0.12	34.88
711/665	0.85	95.7	0.23	67.46	709/678	0.95	56.8	0.13	38.77
709/678	0.81	105.6	0.26	73.54	705/687	0.94	57.6	0.14	38.84
702/672	0.87	89.1	0.21	61.73	702/672	0.96	49.7	0.11	33.01

The equation which produced the best results can finally be written as:

$$C_{chl} = 452.88 * \ln(R_{[702/672]}) + 4.5279$$

#### 4.2. Continuum Interpolated Band Ratio

Kurer [14] yielded the best results using AVIRIS channels 26,29,31 with center wavelengths situated at 648, 677 and 697 nm to calculate the CIBR. The correlation with this band combination was statistically not significant even though  $R^2$  was quite high.

The water body studied by Kurer showed a CHL-a content ranging between 0 to 5  $\mu\text{g/L}$ . As the spectral features of chlorophyll absorption between 650 and 750 nm shift with increasing chlorophyll absorption towards longer wavelengths, this band combination did not seem suitable for the Indian water bodies with very high chlorophyll contents.

Modification of the band combination to position the absorption band on the absorption maxima and the reference bands on the shoulders of the trough (Fig.3) resulted in different wavelength combinations shown in Tab.2. The performance of the algorithms improved significantly. The wavelength combination 651,675,713 nm yielded best results on radiometric in-situ data with a correlation coefficient of  $R^2= 0.97$  and an absolute deviation of 21  $\mu\text{g/L}$ . Regarding the very broad range of 0 to 460  $\mu\text{g/L}$  CHL this result can be regarded as very good.

Table 2: Results for CIBR algorithm applied to FieldSpec Pro FR spectra.

Wavelength	$R^2$	Std error	$MD_R$	$MD_a$
655/678/705	0.97	42.7	0.09	24.87
655/675/708	0.97	42.6	0.09	24.98
651/682/713	0.97	38.9	0.09	25.28
651/675/713	0.97	35.9	0.07	20.89
650/675/708	0.97	36.5	0.08	21.73
637/667/687 (Kurer)	0.67	140.4	0.29	84.6

The equation which produced the best results can finally be written as:

$$C_{chl} = -947.63 * CIBR_{[651/675/713]} + 897.54$$

### 4.3. Area above a Baseline

The area under the curve was calculated by means of an integral function. The first calculation was done using fixed wavelengths as delimiters at 680 and 740 nm. The algorithm was then slightly modified, using different wavelengths as delimiters than those described in literature. The disadvantage of the approach with fixed delimiters is that it does not take into account the shift of the peak of each spectra (Tab.5, left). Especially for a data-set with water bodies ranging from low to high CHL content the delimiters should have to be fitted to the shape of each spectrum.

Tab. 3: Results for Area above a Baseline Algorithm.

Wavelength	$R^2$	Std error	$MD_{\%}$	$MD_{\sigma}$
fixed: 680/740	0.932	63.87	0.13	37
not fixed	0.933	63.2	0.13	36

### 4.4. Peak Magnitude above a Baseline

The height of the peak above a baseline drawn from 680/740 nm showed very good results when using a linear curve fit. The results are similar to those of the CIBR, with an absolute deviation error of 23  $\mu\text{g/L}$  and a relative error of 8% (Tab.4).

Tab. 4: Results for Peak Magnitude above a Baseline Algorithm.

Wavelength	$R^2$	Std error	$MD_{\%}$	$MD_{\sigma}$
fixed 680/740	0.973	40.3	0.08	23

### 4.5. Position of Peak near 700 nm

This method applied onto radiometric in-situ data showed the best results of all algorithms used with  $R^2$  values of 0.97 and a relative error ( $MD_{\%}$ ) smaller than 6%. However, the algorithm has some significant disadvantages: As it can be seen in Fig.5, the shift of the peak occurs in a small range of around 20 nm width of the wavelength range. This implies that

this method requires data from a sensor with many bands at a very short sampling interval. Sensors with a higher spectral resolution might only have two or three bands in this wavelength range. The problem could be solved by an interpolation using a polynomial function. But several presumptions regarding the shape of the curve would have to be done leading to a lost in accuracy.

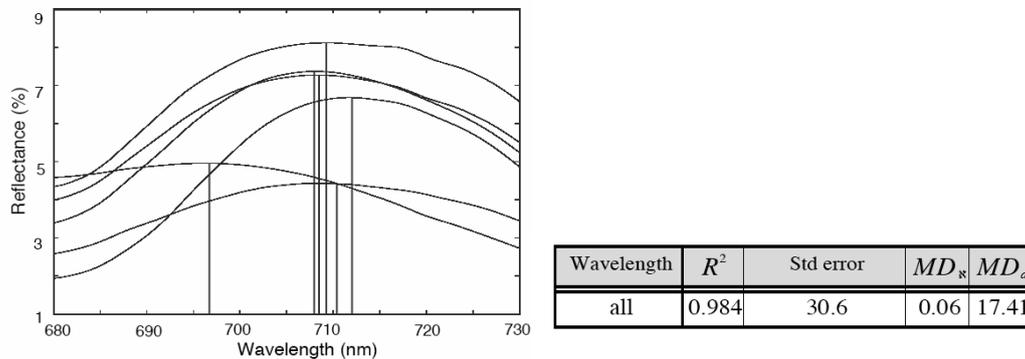
even for a spectral resolution of 1 nm the peak position parameter could have only 20 different values. This means, that for an inversion of the algorithm, only 20 possible values of concentration could be assigned to each pixel. In water bodies with high ranges of CHL, application of this algorithm is therefore not feasible without polynomial interpolation.

Due to these significant disadvantages, the algorithm was not further applied to the convolved spectra.

The equation which produced the best results can finally be written as:

$$C_{chl} = \frac{(Peakpos_{[nm]} - 696.19)}{35.007}$$

Tab. 5: Range of the position of the peaks near 700 nm on the wavelength axis (left), and results for Position of Peak near 700 nm Algorithm (right).



#### 4.6. Application of Algorithms to Sensor-specific Convolved Spectra

In order to determine the suitability of existing (Hyperion [21] and HyMap [5]) and future (APEX [18]) imaging spectrometers, the algorithms which performed best on ASD FieldSpec Pro FR spectra were applied to convolved spectra of the three sensors with regard to their center wavelengths and FWHM's. If there was no center wavelength at the desired position the next adjacent band was taken.

For the best performing *Simple Band Ratio* at the wavelengths 702/672, no significant difference between Fieldspec, APEX and Hyperion can be detected. The difference between the results is less than 2%. HyMap does not have the bands at this wavelength but adjacent bands performed similarly as well. It is surprising that the difference between the results is rather small. It must be taken into consideration, that the width of the bands varies from 1 nm (Fieldspec) to 10 nm (Hyperion) and 15 nm (HyMap).

*CIBR* shows similar results: There is almost no difference between Fieldspec, APEX and Hyperion, whereas HyMap does not have its bands right positioned for this algorithm. Again, the differences are surprisingly small.

When applying the *Area above a Baseline Algorithm* to the convolved spectra, wider bands performed clearly better. The outcome of Hyperion and HyMap with the delimiters at the same wavelengths are clearly better than the result from unconvolved FieldSpec data. It seems that the more the shape of the curve is generalized through less bands, the better the outcome (see Fig. 7).

The results of the *Peak Magnitude above a Baseline Algorithm* have the same tendency but less clearly. Again, wider bands performed better.

Tab. 6: Results for Simple Band Ratios, logarithmic fit (left), and for CIBR, linear fit (right).

Bands from	Wvl	$R^2$	Std error	$MD_R$	$MD_d$
Fieldspec Pro FR	702/672	0.96	49.7	0.11	33.01
Fieldspec Pro FR	709/678	0.95	56.8	0.13	38.77
Fieldspec Pro FR	711/665	0.96	49.9	0.12	34.88
Fieldspec Pro FR	711/681	0.94	57.7	0.14	39.95
Fieldspec Pro FR	705/678	0.94	57.6	0.14	38.84
HyMap	711/678	0.95	54.6	0.13	38.12
HyMap	711/665	0.95	52.1	0.13	36.36
HyMap	696/681	0.95	54.56	0.13	36.91
Hyperion	702/672	0.96	48.9	0.12	33.09
Hyperion	712/682	0.95	56.3	0.14	39.35
APEX	702/672	0.96	49.97	0.12	33.31
APEX	706/678	0.94	57.6	0.14	39.10
APEX	710/678	0.95	56.3	0.13	38.62
APEX	713/682	0.95	55.5	0.13	38.78
Fieldspec Pro FR	637/667/687	0.67	140.4	0.29	84.6
Fieldspec Pro FR	650/675/708	0.97	36.5	0.08	21.73
Fieldspec Pro FR	651/682/713	0.97	38.9	0.09	25.28
Fieldspec Pro FR	651/675/713	0.97	35.9	0.07	20.89
Fieldspec Pro FR	655/678/705	0.97	42.7	0.09	24.87
Fieldspec Pro FR	655/675/708	0.97	42.6	0.09	24.98
HyMap	650/681/711	0.97	39.4	0.09	26.28
Hyperion	651/682/712	0.97	40.9	0.09	27.1
Hyperion	651-672-712	0.98	34.1	0.07	20.88
APEX	651/682/713	0.97	38.0	0.09	24.74
APEX	651-675-713	0.98	35.3	0.07	20.31
	706/678	0.94	57.6	0.14	39.10

Tab. 7: Results for Area above a Baseline Algorithm (left), and for Peak Magnitude above a Baseline Algorithm (right); f. denotes “fixed”, nf. Denotes “not fixed”.

Bands from	Wvl	$R^2$	Std error	$MD_n$	$MD_a$	Bands from	Wvl	$R^2$	Std error	$MD_n$	$MD_a$
Fieldspec Pro FR	f. 680/740	0.93	63.9	0.13	37	Fieldspec Pro FR	f. 680/740	0.97	40.3	0.08	23
Fieldspec Pro FR	n. f. 680/740	0.93	63.2	0.13	36						
	f. 682/743	0.95	54.8	0.13	38		682/743	0.97	39.2	0.08	24
Hyperion	f. 682/743	0.96	47.6	0.1	28	Hyperion	f. 682/743	0.988	27.3	0.06	17
APEX	n. f. 682/743	0.95	55.9	0.11	33	APEX	f. 682/741	0.974	39.6	0.08	23
APEX	f. 678/741	0.92	69	0.14	39						
HyMap	f. 681/742	0.96	46.2	0.1	28	HyMap	f. 681/742	0.984	30.75	0.07	19

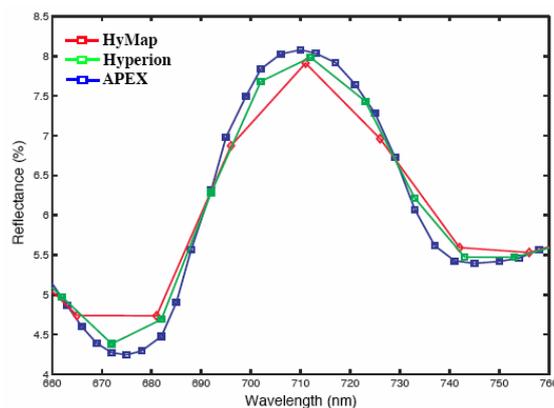


Fig. 7: Spectral Data convolved to the bands of the sensors under investigation. Dots denote band center wavelengths.

Tab.8 shows the position of the bands which are required for a good outcome of the algorithms. The central wavelength of the band is indicated:

Tab. 8: Positions of required bands and their availability for the sensors under investigation.

Sensor	Simple band ratio	CIBR	Area above baseline	Magnitude above baseline
Needed band position (nm)	702, 672	651, 672, 712	681, 742	681, 712, 742
Fieldspec	Y	Y	Y	Y
APEX	Y	Y	Y	Y
Hyperion	Y	Y	Y	Y
HyMap	N	N	Y	Y

Regarding the comparison of the selected sensors, the following conclusions can be drawn:

- If the bands are rightly positioned, a better spectral resolution does not increase the quality of the results significantly for spectral bandwidths below 10 nm. Quite the contrary has been observed, in most cases wider bands performed slightly better.
- There was no significant difference between choosing the bands in a fixed mode or bands adjusted to each spectrum individually (see Chapter 5.3).
- For the algorithms *Peak Magnitude above a Baseline* and *Area above a Baseline* all sensors have their bands at the needed wavelength position (Tab. 8). HyMap does not have the optimal configuration for *Simple Band Ratios* and *CIBR* for the observed hypertrophic Indian waters, even though the differences are small.

It has to be noted that the original bandwidth of the sensors was taken. APEX has the advantage in that it is fully programmable. Bands can be convolved and the width of each band can be cus-

tomized for the desired purpose. Hence, APEX has all the capabilities of HyMap and Hyperion together with its new features, including better spectral resolution.

## 5. CONCLUSIONS

All algorithms applied proved to be of value. Given the large range of observed TCHL concentrations (0–460 µg/L), the yielded  $R^2$  were high and mean deviations low. However, certain presumptions had to be made (e.g., use of TCHL instead of CHL-a in the modelling) and there still remain limitations concerning the used input data. In general, the measured nutrient levels seem too low in regard to CHL concentration. Visually, no algal scum, indicating a bloom, could be observed. Nonetheless, spectral analysis and comparison with similar spectra [9] indicate that quite high concentrations of CHL can be expected, although they were visually not clearly detectable. Macrophytes density supports the assumption of high bio-activity. Therefore, all domestic ponds would be classified as hypertrophic. The trophic state of P5 is not classifiable.

It has to be kept in mind, that the *trophic state classification system* was developed for larger lakes in the Western hemisphere. In these lakes, the limnological cycle is completely different in terms of seasonality, discharge and lack of monsoon. The classification system, thus, may not be applicable to small and shallow ponds in the tropics.

Moreover, this study represents a point-in-time measurement being performed at the end of the dry season. At this time, several meters of the water column had already evaporated, hence the concentration rises. Regarding the cycle of these ponds, a regular monitoring over different seasons would be necessary to draw further conclusions.

Concerning the outcome of the comparison of selected possible hyperspectral sensors for water quality determination, it can be mentioned that the differences were small. An intermediate bandwidth of 10 nm seems appropriate for application of the selected algorithms. All sensors showed an adequate band positioning for the algorithms based on the peak near 700 nm. Using *Simple Band Ratios* and *CIBR*, APEX and Hyperion performed slightly better than HyMap.

To verify whether the chosen algorithms are suitable for an operative monitoring of ponds in the study area, further verification data is indispensable. Implementation of an operative monitoring of water quality in the region remains challenging in regard of seasonal difficulties during the monsoon season and the spatial resolution of current spaceborne sensors.

## 6. ACKNOWLEDGMENTS

We would like to thank the staff of Bidhan Chandra Krishi Viswavidyalaya University, Mohanpur, for the cooperation and infrastructure. Special thanks go to Prof. Dr. Hussein and Dr. Saha.

## 7. References

1. Analytical Spectral Devices (ASD), 1999: *Technical Guide*. 3rd Ed. Section 5-7.
2. Bohley, P., 1989: *Statistik – Einfuehrendes Lehrbuch fuer Wirtschafts- und Sozialwissenschaftler*. 3. Ed., Muenchen, Wien: Oldenbourg-Verlag.
3. Bruegge C.J., Conel J.E., Margolis J.S., Green R.O., Toon G., Carrere V., Holm R.G. and Hoover G., 1990: *In-situ atmospheric water-vapor retrieval in support of AVIRIS validation*. SPIE Vol. 1298 *Imaging Spectroscopy of the Terrestrial Environment*, pp. 150 – 163.
4. Carlson, R.E. and Simpson, J., 1996: *A coordinator's guide to volunteer lake monitoring methods*. North American Lake Management Society. pp. 96.
5. Cocks, T., R., Jensen, R., Stewart, A., Wilson, I. and T. Sheilds, T. 1998: *The HyMap Airborne Hyperspectral Sensor: The system, calibration and performance*. Proc. 1st EARSEL Workshop on Imaging Spectroscopy, Zurich, October 1998, pp. 37-42.

6. Dekker, A. G., 1993: *Detection of optical water quality parameters for eutrophic waters by high resolution remote sensing*. PhD thesis, Free University, Amsterdam.
7. Gitelson, A., 1993: *The nature of the peak near 700 nm on the radiance spectra and its application for remote estimation of phytoplankton pigments in inland waters*. *Optical Engineering and Remote Sensing*, SPIE 1971, pp. 170-179.
8. Gitelson, A., Garbuzov, G., Szilagyi, F., Mittenzwey, K-H., Karnieli A. and A. Kaiser, 1993: *Quantitative remote sensing methods for real-time monitoring of inland water quality*. *International Journal of Remote Sensing*, 14, pp. 1269-1295.
9. Gitelson, A.A., Stark, R. and I. Dor, 1997: *Quantitative near-surface remote sensing of wastewater quality in oxidation ponds and reservoirs: A case study of the Naan system*. *Water Environment Research*, Vol. 69, No. 7, pp. 1263-1271.
10. Gitelson, A.A., Yacobi, Y.Z., Rundquist, D. C., Stark, R., Han, L., and D. Etzion, 2000: *Remote estimation of chlorophyll concentration in productive waters: Principals, algorithm development and validation*. NWQMC Conference Proceedings 2000. Available at: [http://www.nwqmc.org/2000proceeding/table\\_of\\_contents.htm](http://www.nwqmc.org/2000proceeding/table_of_contents.htm) [Access: March 2005].
11. Hyperstat Online: <http://www.ruf.rice.edu/~lane/hyperstat/A120567.html> [Access: March 2005].
12. Kallio, K., Kutser, T., Hannonen, T., Koponen, S., Pulliainen, J., Vepsäläinen, J. and T. Pyhälä, 2001: *Retrieval of water quality by airborne imaging spectrometer in various lake types at different seasons*. *The Science of the Total Environment*, Vol. 268, 1-3, pp. 59-77.
13. Koponen, S., Pulliainen, J., Kallio, K., M. Hallikainen, 2002: *Lake water quality classification with airborne hyperspectral spectrometer and simulated MERIS data*. *Remote Sensing of Environment* 79, pp. 51-59.
14. Kurer, U., 1994: *Bestimmung der Chlorophyllkonzentration im Zugersee anhand von AVIRIS-Bildspektrometriedaten*. MSc. Thesis. Dept. of Geography, University of Zuerich.
15. Kutser, T., Herlevi, A., Kallio, K. and H. Arst, 2001: *A hyperspectral model for interpretation of passive optical remote sensing data from turbid lakes*. *The Science of the Total Environment* 268, pp. 47-58.
16. Lindell, T., Pierson D., Premazzi, G. and E. Zilioli, 1999: *Manual for monitoring European lakes using remote sensing techniques*. Report EUR 18665 EN, Luxembourg: European Commission.
17. Neville, R.A. and J.F.R. Gower, 1977: *Passive remote sensing of phytoplankton via chlorophyll-a fluorescence*. *Journal of Geophysical Research* 82: 3487-3493, pp. 709-722.
18. Nieke, J., Itten, K.I., Debruyne, W. and the APEX Team, 2005: *Status of the Airborne Imaging Spectrometer APEX*. Presented at the 4<sup>th</sup> EARSeL Workshop on Imaging Spectroscopy, Warsaw, April 27-29, 2005.
19. Oron, G. and A. Gitelson, 1996: *Real-time quality monitoring by remote sensing of contaminated water bodies: Waste Stabilization Pond Effluent*. *Water Research*, Vol. 30 (12), 3106-3114.
20. Thiemann, S. and H. Kaufmann, 2002: *Lake water quality monitoring using hyperspectral airborne data- a semi-empirical multisensor and multitemporal approach for the Mecklenburg Lake District, Germany*. *Remote Sensing of Environment* 81, pp. 228-237.
21. USGS, EO-1 Website. <http://eo1.gsfc.nasa.gov/Technology/Hyperion.html> [Access: March 2005].
22. Yacobi, Y. Z., Gitelson, A. and M. Mayo, 1995: *Remote sensing of chlorophyll in Lake Kinneret using high spectral resolution radiometer and Landsat TM: Spectral features of reflectance and algorithm development*. *Journal of Plankton Research* 17, pp. 2155-2173.

15. Kutser, T., Herlevi, A., Kallio, K. and H. Arst, 2001: *A hyperspectral model for interpretation of passive optical remote sensing data from turbid lakes*. The Science of the Total Environment 268, pp. 47-58.
16. Lindell, T., Pierson D., Premazzi, G. and E. Zilioli, 1999: *Manual for monitoring European lakes using remote sensing techniques*. Report EUR 18665 EN, Luxembourg: European Commission.
17. Neville, R.A. and J.F.R. Gower, 1977: *Passive remote sensing of phytoplankton via chlorophyll-a fluorescence*. Journal of Geophysical Research 82: 3487-3493, pp. 709-722.
18. Nieke, J., Itten, K.I., Debruyn, W. and the APEX Team, 2005: *Status of the Airborne Imaging Spectrometer APEX*. Presented at the 4<sup>th</sup> EARSeL Workshop on Imaging Spectroscopy, Warsaw, April 27-29, 2005.
19. Oron, G. and A. Gitelson, 1996: *Real-time quality monitoring by remote sensing of contaminated water bodies: Waste Stabilization Pond Effluent*. Water Research, Vol. 30 (12), 3106-3114.
20. Thiemann, S. and H. Kaufmann, 2002: *Lake water quality monitoring using hyperspectral airborne data- a semi-empirical multisensor and multitemporal approach for the Mecklenburg Lake District, Germany*. Remote Sensing of Environment 81, pp. 228-237.
21. USGS, EO-1 Website. <http://eo1.gsfc.nasa.gov/Technology/Hyperion.html> [Access: March 2005].
22. Yacobi, Y. Z., Gitelson, A. and M. Mayo, 1995: *Remote sensing of chlorophyll in Lake Kinneret using high spectral resolution radiometer and Landsat TM: Spectral features of reflectance and algorithm development*. Journal of Plankton Research 17, pp. 2155-2173.