Determination of water quality parameters using imaging spectrometry
(case study for the Sajó floodplain, Hungary)

Ulanbek Turdukulov
March, 2003

Determination of water quality parameters using imaging spectrometry (case study for the Sajó floodplain, Hungary)

by

Ulanbek Turdukulov

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in

Geo-information Science and Earth Observation,

Environmental System Analyses and Management specialization

Degree Assessment Board

Prof. A.M.J. Meijerink (Chairman) WRS Department, ITC
Dr. Ir. D.C.M. Augustijn, (External Examiner) CTM, UT Twente
Dr. Z. Vekerdy (first supervisor) WRES Department, ITC
Dr. Ir. C. Mannaerts (member) WRES Department, ITC
Prof. M. Hale (member) ESA Department, ITC

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Acknowledgements

I am highly grateful to the following organizations and persons who made a valuable contribution to this academic accomplishment:

WRES department of ITC, the Netherlands, for facilitating this course; Netherlands Fellowship Program for the pleasure and possibility to do this research;

Special thanks to my external supervisor, Dr. Steef Peters, VU, for his guidance and support. The completion of this study is mainly attributed to his invaluable contribution. Thanks to Dr. Chris Mannaerts, WRES, ITC as he always was ready to respond to all academic and personal needs.

Thanks to my first supervisor Dr. Zoltan Vekerdy for initiating and coordinating the HYSENSE project, for the attitude and support during the research, and especially for giving an excellent introduction of Hungary.

ITC staff, especially to Dr. Boudewijn de Smeth, Barbara Cassentini, Harald van der Werff and Jelle Ferwerda for social and academic endeavours.

Thanks to the HYSENSE 2002 project group, especially to:

- DLR team- Dr. Andreas Muller, Ing. Stefanie Holzwarth for coordinating image processing and acquisition, Dr. Peter Gege and Dr. Sabine Thiemann for sharing experience in the field of this research;
- JRC team - Dr. Stefan Sommer and Thomas Kemper for the cooperation during the fieldwork campaign;
- MÁFI group-Dr. Károly Brezsnyánszki, Dr. Péter Kardeván and Mr. László Roth, for the coordination of the HySens project and their kind help in getting all the necessary data;
- Dr.Tibor Zelenka, Geological Survey of Hungary, for his full support and help in the fieldwork;
- Dr. Péter Bakonyi, Dr. Ferenc László and Dr. János Szekeres and Mrs. Márta Spitzer, VITUKI, for the hydrological data.

Special thanks to the staff of Miskólc EPA laboratory – Dr. Oszkár Balázs and Móre Mélinda for the examination of samples and detailed introduction to the study area.

Special heartfelt thanks to Marleen Noomen. She was the centre of all inspiration, confidence, never-ending support and encouragement during my stay and study.

Enschede, March 2003

Ulanbek Turdukulov

Abstract

This MSc study was undertaken as a part of the HYSENS 2002 project, to investigate the use of hyperspectral remote sensing in environmental studies. The test site was chosen due to a history of strong industrial pollution on the floodplain, distributed by the Sajó River, Hungary.

The main objective of the study was the quantification of water quality parameters using remote sensing. The water quality parameters were: Total Suspended Matter (TSM), Inorganic Suspended Matter (ISM), Organic Suspended Matter (OSM), Chlorophyll-a (CHL) and Turbidity, which can indicate both an environmental state of the water bodies (CHL) and indirectly show the distribution of pollutants (pollutants attach to the fine particles present in the water).

The method is based on imaging spectrometry (using the sensors ROSIS and DAIS on board of the aircraft of the German Aerospace Centre, DLR) coupled with field spectroscopy. Field spectrometry measurements were up-scaled to the lower spatial and spectral resolution airborne sensors (ROSIS, DAIS) and widely used satellite sensor (Landsat TM). Furthermore, empirical and semi-analytical methods were applied to each sensor in order to retrieve the concentrations of the water constituents.

The results of the analysis indicate that for CHL an empirical algorithm, based on band ratio NIR/Red (704.46 nm/672.17 nm) shows high correlation. According to the laboratory data, Inorganic suspended matter (ISM) was highly correlated with TSM. Therefore a two-step procedure is proposed to estimate ISM from spectra: first, TSM is estimated using spectral data, and then the ISM is determined from the TSM estimations. To quantify TSM, an inversion of a simplified bio-optical model in the NIR region (where the influence of other water quality parameters is negligible) was applied. The modelling results indicated that the specific inherent optical properties of the Sajó River were changing with time due to flood retreat. Proper modelling required the separation of the backscattering by large and small suspended particles. In lack of direct measurements, the suspended particle size was calculated as a function of the river velocity. Turbidity was identified using an empirical algorithm based on the reflectance at 695.64 nm.

Furthermore, the developed algorithms were applied to the atmospherically corrected ROSIS and DAIS images to show spatial distribution of water quality parameters. At this stage, limitations of the use of remote sensing for quantification of studied water quality parameters have been also discussed like proper atmospheric correction and filtering of the images. The results show: i) the possibilities of using hyperspectral remote sensing (ROSIS and DAIS sensors) for operational water quality monitoring, ii) the methods for quantification of water quality parameters using remote sensing data, and iii) the possibilities of using the spectral characteristics of the Landsat TM sensor for water quality monitoring.

Table of content

Acknowledgements iv
Abstract v
Table of content vi
List of figures viii
List of tables x
List of frequently used abbreviations xi
Chapter 1. Introduction 1
1.1. Problem definition 1
1.2. Study objectives 3
1.3. Project background and study area description 3
1.3.1. Study area 4
1.3.2. Source of pollutants 4
1.3.3. Hydrological and sediment transport characteristics of River Sajó and its floodplain 5
1.4. Thesis structure 6
Chapter 2. Data and analysis methods 7
2.1. General setup 7
2.2. Data description 7
2.2.1. Field data collection 7
2.2.2. Description of dataset collected from the Dutch lakes 8
2.3. Flight campaign and sensors description. 8
2.4. Description of the methods 9
2.4.1. Empirical approach 10
2.4.2. Analytical approach 10
2.4.3. Air-water interface correction algorithm 12
2.4.4. Limitations of the methods 14
Chapter 3. Results and discussions on using statistical approach 15
3.1. Statistical relations between the water quality parameters 15
3.2. Statistical analysis of the CHL content and corresponding spectra 16
3.2.1. Spectral signatures of waters with high CHL content 16
3.2.2. Linear regression between reflectance and CHL 17
3.2.3. First derivatives and CHL 18
3.2.4. Band ratio NIR/Red and CHL. 19
3.2.5. Discussion of results on using statistical approach to retrieve CHL 20
3.2.6. Testing and final conclusion on using statistical algorithm for CHL 20
3.3. Statistical analysis of the TSM and corresponding spectra 22

3.3.1. Spectral signatures of waters with high TSM concentrations 22
3.3.2. Results and discussion on using statistical approach for TSM quantification 23
3.4. Statistical analysis of turbidity and corresponding spectra 25
Chapter 4. Results and discussion on using semi-analytical approach 27
4.1. Bio-optical modelling of TSM 27
4.1.1. Identification of the spectral region where influence of other parameters than TSM is minimal. 27
4.1.2. Preliminary modelling 29
4.1.3. Results of bio-optical modelling of TSM 31
4.1.4. Discussion of the results 32
4.1.5. Justification of ratio of large to small particles concentration by calculating sediment transport capacity of Sajó River using Yalin equation 34
4.2. Modelling Chlorophyll-a using a semi-analytical solution of the band ratio NIR/Red 36
4.2.1. Algorithm description 36
4.2.2. Result and discussion on bio-optical modelling of CHL 37
4.3. Error analysis and sensitivity of the models 40
Chapter 5. Application of the developed algorithms to the hyperspectral images 43
5.1. Comments on using remote sensing data in studying the water bodies. 43
5.2. Examination of the images 43
5.3. Image processing 45
5.4. Analysis of the resulted maps 46
5.4.1. CHL maps 47
5.4.2. Map of TSM 47
5.4.3. Turbidity maps 47
5.4.4. ISM and OSM maps 47
5.5. Concluding remarks on the image processing 47
Chapter 6. Conclusions and recommendations 49
6.1. Conclusions and recommendations on applying the empirical approach 49
6.2. Conclusions and recommendations on applying bio-optical modelling 50
6.3. Conclusions and recommendations on applying atmospheric and air-water corrections .. 50
6.4. Conclusions on the processing remote sensing data 51
References 52
Appendix A. Standards and procedure for the laboratory analysis 55
Appendix B. Senosors and Spectra 59
Appendix C. Computation of sediment transport capacity using Yalin equation (using water level measured at Sajólad gauging station on 17/08/2002) 63
Appendix D. Maps 64

List of figures

Figure 1.1 Location of the study area (Source: CIA World Factbook, 2001); FCC: [Red-10.3 µm, Green –2.µm, Blue-0.921 µm] 5
Figure 2.1 Central wavelengths of the sensors used in the research 9
Figure 2.2 Methods for interpreting remote sensing data (adopted with modifications from Krijgsman, 1994). 9
Figure 2.3 Mean specific absorption (left) and backscattering (right) coefficients of the Dutch lakes (Adopted from BIOPTI 1.0, Erin Hogenboom,1995) 12
Figure 2.4 Effect of waves on the spectral measurements at the sampling point L2 on 17th August. 13
Figure 2.5 Examples of R_{rs} and corresponding R(0-) spectra 14
Figure 3.1 Correlation between TSM and ISM for the Hungarian dataset 16
Figure 3.2 Measured TSM and ISM concentrations from the Dutch lakes with the trend line plotted using regression equation established in the Hungarian dataset (see Figure 3.1) 16
Figure 3.3 Reflectance spectra of water bodies with the highest chlorophyll content 17
Figure 3.4 Regression coefficients between the WQ parameters and reflectance (Sensor: Spectrometer) 17
Figure 3.5 First derivatives at 684 nm vs CHL for the ROSIS sensor 18
Figure 3.6 Band ratio versus CHL for the ROSIS sensor 19
Figure 3.7 Band ratio 693/675 vs CHL for the DAIS sensor 19
Figure 3.8 Band ratio 834/661 vs CHL for the Landsat TM sensor 20
Figure 3.9 Band ratios vs. CHL concentrations for the Dutch lakes with plotted regression lines: .. 21
Figure 3.10 Modelled vs. Observed for the Hungarian dataset (applying to the ROSIS sensor's wavelengths) 22
Figure 3.11 Result of simulating subsurface volume reflectance by changing TSM from 5mg/l (run1) to 185mg/l(run 10) with 20mg/l step (done by using BIOPTI 1.0 software, Erin Hogenboom,VU, 1995) 23
Figure 3.12 Regression lines and coefficients between reflectance at 686 nm and TSM concentrations for the Hungarian dataset (ROSIS sensor) 24
Figure 3.13 Regression line between TSM and reflectance at 712 nm for the Dutch data set 24
Figure 3.14 Regression line between Turbidity and Reflectance at 695.64 nm for the Hungarian dataset (Sensor: Spectrometer) 26
Figure 3.15 Turbidity vs. first derivatives at775.92 nm (Sensor: Spectrometer) 26
Figure 4.1 Grain-size distribution of sediment sampled at the different discharges (Q - discharge) in central Belgium (Adopted from Steegen et al., 1998). 29
Figure 4.2 Regression line between the observed and modelled TSM concentrations at 762 nm using resampled to the ROSIS sensors reflectance 31

Figure 4.3 Regression line between the observed TSM and the modelled TSM concentrations at 747 nm using resampled to the DAIS sensor reflectance 32
Figure 4.4 Regression line between the observed and modelled at 834 nm TSM concentrations using resampled to the LandsatTM sensor reflectance 32
Figure 4.5 Regression coefficients between the modelled and observed TSM for different sensors 33
Figure 4.6 Regression graph between the observed and modelled at 832 nm using Gege's formula TSM concentrations (for the DAIS sensor reflectance) 33
Figure 4.7 Regression coefficients between the observed and modelled TSM using the Gege's formula 34
Figure 4.8 Bio-optically modelled CHL vs. observed CHL (for the ROSIS sensor) 37
Figure 4.9 Bio-optically modelled CHL vs. observed CHL for the Landsat TM sensor 38
Figure 4.10 Sensors sensitivity to the region 650-710 nm with the spectrum of the sampling point L1_2008 38
Figure 4.11 Band ratio vs CHL for the MERIS sensor 39
Figure 5.1The ROSIS image's spectra from 17th Aug (solid lines) in comparison with the field spectra from 18th Aug (dashed lines) 44
Figure 5.2 The DAIS image's spectra (solid lines) in comparison with the field spectra (dashed lines) 44
Figure 5.3 Results of applying empirical line correction algorithm (dotted lines are the DAIS image spectra, solid –field spectra) 45
Figure 5.4 Raw image (utmost left), image with applied Lee smoothing filter, kernel size [3*3] (in the middle) and image with the applied smoothing (convolution low pass) filter, kernel size [5*5]. 46
Figure A.1 Map of the sampling sites 58
Figure B.1 Field spectra collected with the spectrometer GER 3700 60
Figure B.2 Field spectra resampled to the wavelengths of the ROSIS sensor 61
Figure B.3 Field spectra resampled to the wavelengths of the DAIS sensor 61
Figure B.4 Field spectra resampled to the wavelengths of the Landsat TM sensor 62
Figure B.5 Reflectance spectra collected from the Dutch lakes 62
Figure D.1 CHL maps 64
Figure D.2 Turbidity maps 65
Figure D.3 TSM map 66
Figure D.4 ISM and OSM maps 67

List of tables

Table 3.1 Correlation coefficients between WQ parameters 15
Table 3.2 Summary of the regression analysis between CHL and single band reflectance (indicated wavelengths show where the highest regression coefficients were observed) 17
Table 3.3 Summary results of the regression analysis for the first derivatives and CHL..... 18
Table 3.4 Summary of the regression analysis between band ratios and CHL..... 19
Table 3.5 Summary of results on use of the statistical approach to retrieve TSM concentrations.... 23
Table 3.6 Summary of results using statistical approach for the turbidity determination 25
Table 4.1 Assumption values of ratio of large to small particle's concentration in the Sajó River during the flood retreat. 30
Table 4.2 Input parameters for TSM modelling (resampled to the DAIS sensor) 31
Table 4.3 Input parameters to the Yallin equation 35
Table 4.4 Result of sediment transport capacity calculation in comparison with the assumption of the bio-optical modelling of TSM 36
Table 4.5 Comparison of the regression coefficients and equations between the observed and modelled CHL using 2 approaches 40
Table 4.6 Summary of the standard errors (N=15) between the modelled and observed WQ parameters for the proposed algorithms 41
Table 4.7 Sensitivity of the bio-optical model for the sampling point R2_2008 (example of the ROSIS sensor) 41
Table 4.8 Sensitivity of the bio-optical model for the sampling point L3_2008 (example of the ROSIS sensor) 42
Table 5.1 Description of the algorithms applied to the hyperspectral images 46
Table A.1 Water sampling results and descriptive statistics..... 56
Table A.2 Result of ANOVA analysis for the laboratory samples 57
Table B.1 Specifications of Spectrometer GER3700 (Source: http://www.ger.com/3700.html) 59
Table B.2 ROSIS sensor's specifications (Sources: DLR, Martin Habermeyer, personal communication (2002) and http://www.op.dlr.de/ne-oe/fo/rosis/home.html) 60
Table B.3 DAIS sensor's specifications (Source: http://www.op.dlr.de/dais/dais.htm) 60

List of frequently used abbreviations

BIOPTI	Bio-optical model for Inland waters
CDOM	Coloured Dissolved Organic Matter
CHL	Chlorophyll-a
DAIS	Digital Airborne Imaging Spectrometer
DLR	German Aerospace Centre
ENVI	Environment for Visualizing Images software
FOV	Field of view
ISM	Inorganic Suspended Matter
MERIS	Medium Resolution Imaging Spectrometer
NIR	Near Infra Red
NTU	Nephelometric Turbidity Unit
OSM	Organic Suspended Matter
RIZA	Institute of Water Pollution Control, the Nether-lands
ROSIS	Reflective Optics System Imaging Spectrometer
SIOP	Specific Inherent Optical Properties
TSM	Total Suspended Matter
VITUKI	Water Research Centre, Hungary
WQ	Water Quality

Chapter 1. Introduction

1.1. Problem definition

With the advent of industrialisation and the increasing population, the range of requirements for the water has increased together with greater demands for higher water quality. In parallel with the water use for the variety of human activities (drinking and personal hygiene, fisheries, agriculture, industry, transport and recreation), since ancient times water has been considered as the most suitable medium to clean, disperse, transport and dispose wastes.

Increasing dispose of wastes in the water bodies means a great potential for the environmental damage and emphasizes the need to monitor, protect and manage water resources.

Traditionally, the water quality analysis has involved directly sampling areas in question. Conventional measurements of water quality require *in situ* sampling and expensive and time-consuming laboratory work. Due to these limitations, the sampling size often cannot be large enough to cover the entire water body. Therefore the difficulty of synoptic and successive water quality sampling becomes a barrier to water quality monitoring and forecasting (Shafique et al., 2001).

Remote sensing offers the possibility of covering a large spatial area with a high temporal frequency. It also provides a spatial distribution of the constituents, which direct sampling cannot economically accomplish. Spatial distributions provide deeper insight into many of the hydrologic and biological processes that are directly affected by the concentrations of water constituents. These water constituents can also play a role to calibrate and validate two- and three-dimensional hydrodynamic and ecological models (Krijgsman, 1994). Examples of the later can be erosion models, sediment transport models, global climate change models, etc.

During the last decade the quantitative estimation of water constituents by remote sensing became a more and more common task due to several reasons. Firstly, there exist a number of established techniques for the retrieval of substance concentrations based on empirical, semi-empirical or analytical (Dekker, 1993; Doerffer and Fischer, 1994; Schaale et al., 1999) methods for coastal or inland waters. Secondly, the atmospheric and geometric corrections which have to be applied to remote sensing data are becoming more and more precise, and are close to be operational today (Schaale et al., 1999).

Jaquet et al.(1994) included the list of water quality descriptors, which they felt had the potential to be estimated by remote sensing:

- Directly measurable physical descriptors such as temperature, colour, transparency and turbidity;
- Estimable physical descriptors such as concentration of suspended minerals and suspended solids;

- Estimable chemical descriptors such as Dissolved Organic Carbon (DOC) and Coloured Dissolved Organic Matter (CDOM);
- Estimable hydrobiological descriptors such as water levels and aquatic vegetation species;
- Estimable hydrobiological descriptor such as Chlorophyll-a.

In this research, due to time and finance constrains, attention was given to the water quality descriptors which can indicate both an environmental state of the water bodies and indirectly show the distribution of pollutants such as:

- Turbidity;
- Total Suspended Matter;
- Inorganic Suspended Matter;
- Organic Suspended Matter;
- Chlorophyll-a.

Turbidity is a unit of measurement quantifying the degree to which light travelling through a water column is scattered by the suspended organic (including algae) and inorganic particles. Thus, indirectly it points on the presence of all other studied parameters. Turbidity is commonly measured in Nephelometric Turbidity Units (NTU).

Total Suspended Matter (TSM) includes all suspended material smaller than 150 μm but larger than 0.45 μm. The parts of it that can be burnt in the oven (550°C, 24 hours) consist of organic matter and the rest is called **Inorganic Suspended Matter** (ISM). Units of measurements for both TSM and ISM are milligrams per liter (mg/l) or ppm (part per million). Thus, by defining TSS and ISM, one can calculate **Organic Suspended Matter** (OSM): OSM (mg/l) = TSM - ISM

Chlorophyll-a. Concentration of a photosynthetic pigment Chlorophyll-a (CHL) shows primary production and trophic state of a water body, e.g. algae growth that forms a food chain for higher organisms. Commonly used unit of measurements is μg/l.

There are several negative effects of turbid waters to the potential users (Shafique et al. 2001):

Health effect: The suspended particles may be composed of organic and/or inorganic constituents. Because inorganic particles may attach heavy metals and pesticides, and organic particulate may harbor pathogenic microorganisms, turbid conditions may be dangerous for health.

Industrial effect: Turbid water may not be suitable for use in industrial processes. An abundance of suspended solids may obstruct or scour pipes and machinery.

Recreational effect: Highly turbid waters may be hazardous to the welfare of swimmers and boaters. Turbidity may help to conceal potentially dangerous obstructions such as boulders and logs. Also, the organic constituents of turbid waters may harbor high concentrations of pathogenic bacteria, viruses and protozoan.

Environmental effect: The array of turbidity-induced effects that can occur in a water body may change the composition of an aquatic community:

- First, turbidity caused by a large volume of suspended sediment will reduce the extent to which light can penetrate the water column, thereby suppressing the photosynthetic activity of phytoplankton, algae, and macrophytes, especially those farther below the surface.
- If high turbidity is largely the result of high algae content (CHL), light penetration will be limited and primary production will, therefore, be restricted to the uppermost strata of the water column. Excess turbidity leads to fewer photosynthetic organisms available to serve as food sources for many invertebrates. As a result, the overall numbers of invertebrates may decline, which may lead to a decline in the fish population.
- If turbidity is largely due to excess nutrients (like the in case of runoff from agricultural areas), Dissolved Oxygen (DO) depletion may occur in the water body. The available excess nutrients will increase the rate at which microorganism break down detritus, a process that requires DO. In addition, excess nutrients may result in increased algae growth. Although the algae's photosynthetic processes produce DO during the day, these algae also respire at night, a process that consumes DO. Large declines in fish communities are often the result of extensive DO depletion;
- Finally, heavy metals and pesticides attached to the suspended particles in turbid waters may lead to the extinction of species in the aquatic community and consequently have impact on the whole ecosystem of a floodplain.

Many of those impacts were observed in the study area. The Sajó River basin was one of the most industrialized regions in the past socialistic history of Hungary. Consequently pollution has been a serious problem for many years. The main pollutants were heavy metals. Although in recent years some of the factories were closed down, short-term peak contamination as well as shifts to new chemicals can be observed continuously (RIZA and VITUKI, 1994). Thus, the region is still in need of a good monitoring system for water quality, as the river is the main medium of pollution transport. It is particularly important in relevance to the Hungarian entrance to the European Union.

1.2. Study objectives

The main objective of the research is *the quantification of water quality parameters (TSM, ISM, OSM, Turbidity, CHL) using remote sensing, particularly imaging spectrometry.*
In order to achieve the main objective, the following specific objectives were set:

- To develop model(s) to process remote sensing data for the monitoring water quality of Sajó river floodplain;
- To apply established model(s) to the images;
- To explore possibilities and limitations of water quality monitoring using the established model(s) with lower spectral and spatial resolution sensors.

1.3. Project background and study area description

Study was undertaken as a part of ITC contribution to the HYSENS project aiming to show the use of hyperspectral remote sensing in environmental studies. The project group was formed by two Hungarian institutions (MÁFI and VITUKI), by the ITC from the Netherlands and by the European

Commission's Joint Research Centre (JRC) in Italy. Their aim was to develop applications of hyper-spectral remote sensing using three Hungarian test areas, which were polluted by mining and industrial sources at different levels. A flight campaign was completed over the selected areas by DLR (German Aerospace Centre), who provided the pre-processing of the airborne images too. One of the Hungarian test sites was a reach of the Sajó River and its floodplain.

1.3.1. Study area

River Sajó has a catchment area of 12 707 km^2, 8 494 km^2 of which is situated in Slovakia. Its two important tributaries are the Bódva and Hernád rivers. Length of Sajó River is 223 km. The study area is located along a lower reach of it; the flight line is about 10.7 km long including a confluence with Hernád River (Figure 1.1). The site was chosen due to a history of high industrial pollution of the river upstream of the study area, which resulted in the deposition of pollutants on the studied floddplain. The main pollutants were heavy metals. A study undertaken jointly by RIZA (Institute of Water Pollution Control, The Netherlands) and VITUKI (Water Research Centre, Hungary) in Sajó River showed that several heavy metals including mercury, lead, cadmium and iron exceeded Hungarian requirements (RIZA and VITUKI, 1994).

In addition to the rivers, there are several lakes in the study area. Most of them were result of gravel extraction and gradual filling up of groundwater. A reconnaissance survey revealed that lakes are deep (up to 15 meters) but not stratified and relatively clean in terms of TSM and turbidity. However, a few of them are prone to eutrophication with high CHL (Table A.1, Appendix A). Besides of a functionality of being the gravel extraction sites, most abundant lakes serve as recreational sites.

1.3.2. Source of pollutants

River Sajó and its tributaries receive discharges from a number of point and diffuse sources including municipal and industrial wastewater discharges (Miskolc Waste Water Treatment Plant (WWTP), Saslog Chemical Plant, Borsod Chemical Plant, Kazinabracika WWT and Ozd WWT), unsewered communities, run-off from agricultural land, and licensed and unlicensed waste disposal sites. Beside the man made pollution sources, there are natural sources of the river's pollution namely soil and bank erosion - both valleys of Sajó and Hernád rivers consist of easily erodible soil material (RIZA and VITUKI, 1994).

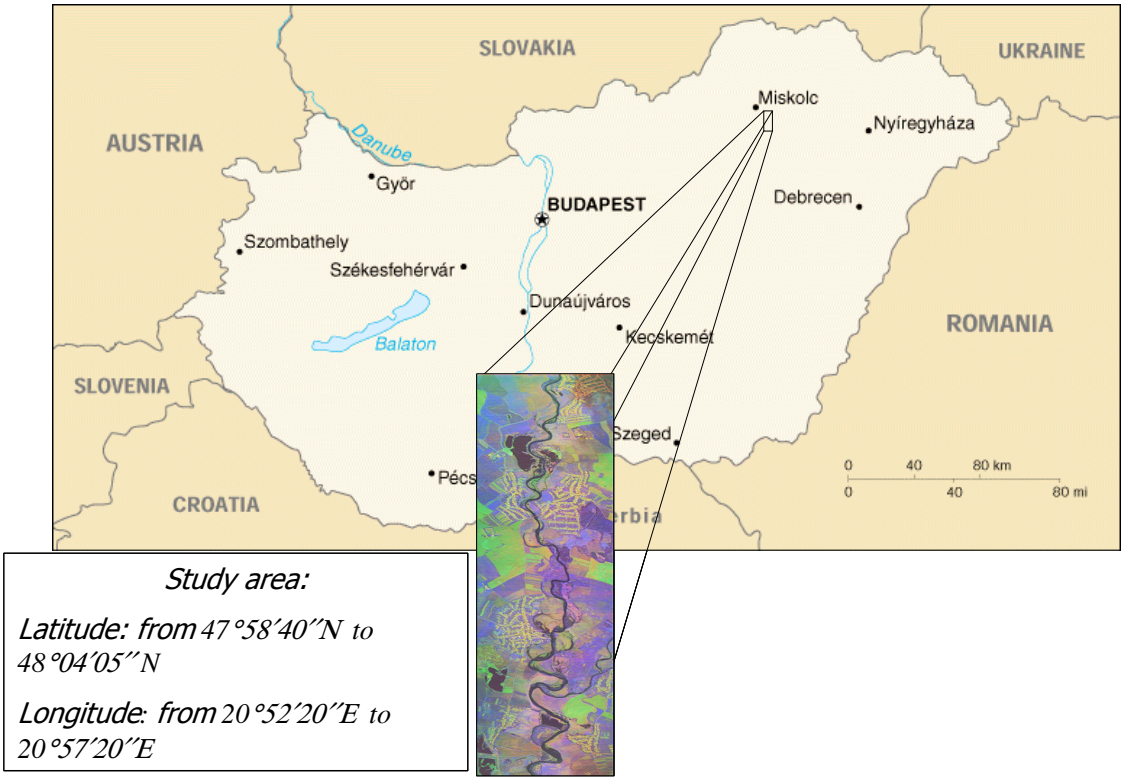


Figure 1.1 Location of the study area (Source: CIA World Factbook, 2001); FCC: [Red-10.3 µm, Green –2.µm, Blue-0.921 µm]

1.3.3. Hydrological and sediment transport characteristics of River Sajó and its floodplain

According to the topographic map, River Sajó in the study area has an average flow velocity between 0.6 and 0.9 m/s and width of about 40 meters (depending on the water level). The difference between minimum and maximum water levels at a given gauge station varies from 2.3 to 4.3 m along the river. Mean monthly water levels in April and May are higher due to the usual spring rains than the water levels in March due to snowmelt floods (RIZA and VITUKI, 1994).

Due to the fact that valleys of Sajó and Hernád River consist of easily erodible materials and because of the relatively high slopes of the river (40-50 cm/km) they carry a high amount of suspended sediments. Especially at high flows they can attain up to several thousands mg/l. The suspended sediment has an average diameter of 0.04 mm. About 30% of the suspended sediment is falling into the fine sand category (between 0.05 and 0.1 mm) and about 10% is finer than 0.006 mm. (RIZA and VITUKI, 1994). Similar results were found by Yun Liu (2003), who did the particle's size distribution test of samples taken on 19th August from the fresh sediment traps left after the flood: about 18.5 % of sediments consists of silt (particles size 0.002 - 0.05 mm – 8.9%) and clay (particle size < 0.002 mm – 9.6%).

Since precipitation greatly affects the river's discharge, consequently its flow velocity and sediment transport capacity, it is necessary to mention the weather condition during the field campaign. July and August are to be the hottest months (RIZA and VITUKI, 1994), but at the moment of our field-work (30th July-27th August 2002) there were about 2 weeks of heavy rains. It resulted in unusually high water levels (measurements of water level taken at the Sajólad station – 276 cm on 15th August during the highest flow and 178 cm on 17th during the flood retreat).

1.4. Thesis structure

According to the developed methodology and chronology of the research and in view of clarity to the reader, the present thesis has been structured into five further chapters:

Chapter 2. Describes the materials and approaches used for the methodology development of modelling water quality parameters using remote sensing. Two approaches are used to develop such method being referred to *the statistical* and *the semi-analytical* approach (bio-optical modelling);

Chapter 3. Application of the statistical approach;

Chapter 4. Application of the bio-optical modelling;

Chapter 5. Gives discussion and results of applying models to the images;

Chapter 6. Summarizes overall results and limitations of the use of remote sensing for quantification of studied water quality parameters.

Chapter 2. Data and analysis methods

This chapter describes a general approach followed in this research. It also gives an overview of data available for the developing models including field spectra collection and flight campaign data. Additionally, this chapter provides an overview of the methods used for the interpretation of the field data

2.1. General setup

Since the beginning of the project campaign it was not certain whenever we would obtain the images. Reasons were: short time "window" for the flight and changing weather conditions. Even if the flight would take place (which fortunately happened), DLR would require at least 5 months to process the images, which was considered to be on the edge for their use in the MSc thesis. All this made us be independent from the flight and carry out field spectrometry along with instantaneous water sampling. General approach was to develop methodology for water quality monitoring based on the collected WQ parameters and corresponding field spectra. In case the images would be given in time, the developed methodology would be applied to the images. Thus, below follows the available data as well as sensors descriptions.

2.2. Data description

2.2.1. Field data collection

In situ spectrometry measurements were carried out from 20 to 25 August 2002 along a reach of the Sajó River, Hungary. All the measurements were performed on a flat water surface and in sunny condition. The following measurements were made:

- Upwelling radiance spectra of the water bodies with the sensor (spectrometer GER 3700, fiber optics with 23^0 FOV, with about 1.5 nm interval) viewing the water surface vertically (30 cm above the water surface). Nine measurements were taken from each site consecutively, which were later averaged to minimize random effects
- Downwelling radiance spectrum is measured when the sensor views a reference panel (Barium sulphate plate with approximately 100% reflectance, 20 cm above the panel);
- Reflectance spectra of the water body (ratio of the upwelling radiance of the water and that of a reference panel);
- Water samples were taken at 0.2 m depth in order to analyse them for water quality parameters.

Appendix A describes in details the procedures and methods, which the laboratory of Environmental Protection Agency in Miskolc followed to analyze water samples. It also gives a map with the location of the sampling sites.

The *in situ* measurements have been done in 15 points (weather conditions and the limited availability of the spectrometer restricted the number of the possible measurements.). Analysis results show large variances and standard deviations because of a few high values (examples are point R4_2008 for TSM and L4_2508 for CHL, Appendix A, Table A.1). This was due to the difficulty of finding water bodies with gradual changes in water quality parameters. The same reason limited the number of measurements to 15.

The weakness of having only 15 data to make a feasible relationship made us look for the similar spectrometry measurements on other water bodies. The next part of the thesis describes data collected from the Dutch lakes by the Institute of Environmental Studies (IVM) of Free University Amsterdam (VU) with a kind permission to use them in this research.

2.2.2. Description of dataset collected from the Dutch lakes

The dataset consists of spectra taken from the Dutch lakes with spectrometer PR650-1 having range from 380 to 780 nm with 4 nm intervals (Appendix B, Figure B.5). The number of measurements taken in cloudless condition was 34. Along with the spectra, the following WQ parameters were measured: TSM, ISM, OSM, CHL, Secchi Disc depth, CDOM. Unfortunately, turbidity values were not measured on Dutch lakes.

The dataset from the Dutch lakes was used to evaluate the performance of the models based on the Hungarian dataset. In order not to confuse the reader, from now on spectra from Dutch lakes will be named as “Dutch dataset” and field spectra collected in Sajó river floodplain will be called “Hungarian dataset”.

2.3. Flight campaign and sensors description.

Flights took place on 17th and 18th August 2002. The main sensors onboard of the aircraft were ROSIS and DAIS. ROSIS (Reflective Optics System Imaging Spectrometer) is a compact airborne imaging spectrometer that has been primarily developed for oceanographic and limnology applications, but land monitoring has been aspired as well. Table B.2 (Appendix B) illustrates the main specifications of the ROSIS sensor used during the flight in Hungary[1].

The second sensor onboard, DAIS (Digital Airborne Imaging Spectrometer, DAIS 7915) had the characteristics described in Table B.3 (Appendix B).

The collected field spectra, having range a from 350 to 950 nm and interval of about 1.5 nm, were resampled to the ROSIS and DAIS sensors using Gaussian distribution, full-width-half-maximum [FHWM] and central wavelength values of above-mentioned sensors. Spectral convolution was done in ENVI 3.5 using spectral tools option. As one of the specific objectives is to use lower spectral resolution sensor, field spectra were resampled to the wavelength of a Landsat TM sensor too. Spectral range considered here was from 400 nm to 900 nm due to a gradually increasing noise effect outside of the given range. For this range, Figure 2.1 gives descriptive information about the spectral

[1] ROSIS configuration might change from flight to flight

resolution of sensors used in this research and Appendix B gives the graphs of resampled spectra for each sensor.

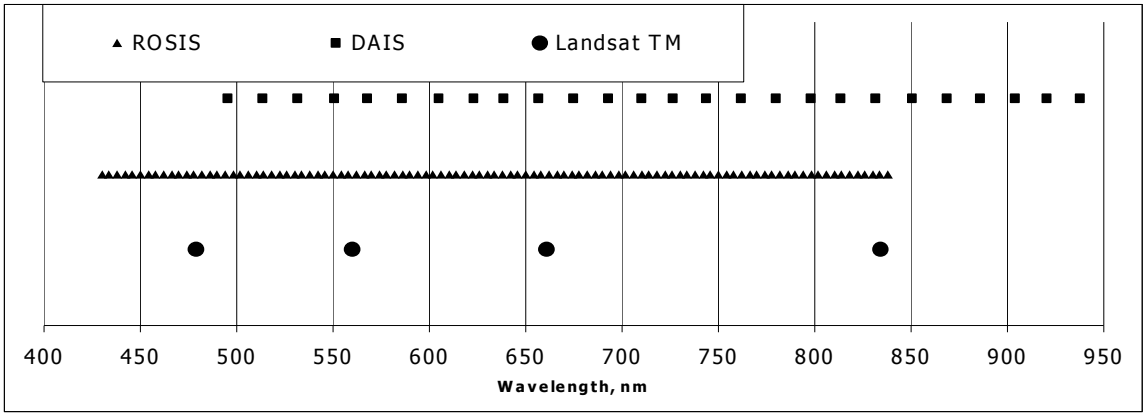


Figure 2.1 Central wavelengths of the sensors used in the research

2.4. Description of the methods

Comprehensive research on quantification of water quality parameters and reflectance (or radiance) spectra has started since the early 70's. Generally, there are two different approaches to estimate the concentrations of WQ parameters from the remote sensing reflectance namely, empirical (also called statistical), and analytical approach (also called bio-optical modelling, Figure 2.2, Krijgsman, 1994).

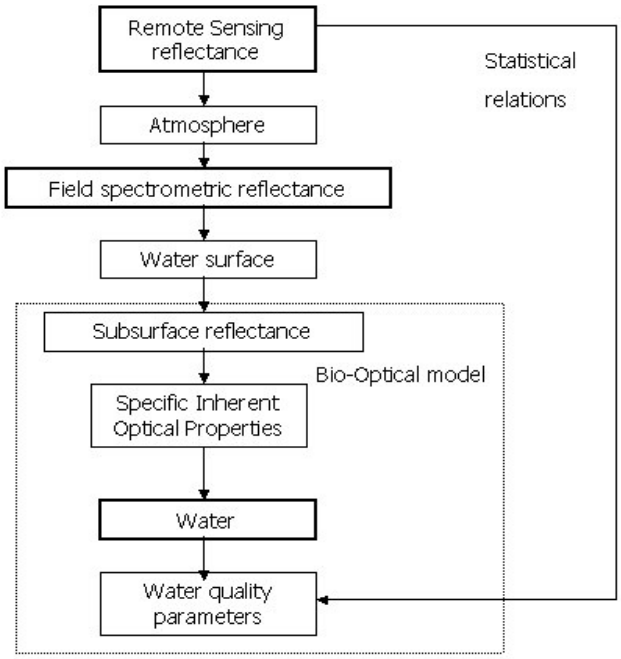


Figure 2.2 Methods for interpreting remote sensing data (adopted with modifications from Krijgsman, 1994).

2.4.1. Empirical approach

The empirical approach is based on the calculation of a statistical relation between the water constituent concentrations and reflectance (or radiance). The advantage is that the empirical algorithms are easy to use and they are straightforward. Disadvantages are: spurious results may occur while using this method, because a causal relationship does not necessarily exist between the parameters studied (Hogenboom and Dekker, 1999) and results of empirical algorithm always need *in situ* data because illumination, surface water, atmospheric conditions and subsequently underwater conditions may change between different remote sensing missions. However, spectrometry measurements performed at the flat water surface and sunny conditions along with direct water sampling are free from some of those limitations (Figure 2.2).

Literature review on empirical algorithm for estimating water quality parameters shows vast variety of algorithms used. They start from a simple linear regression between reflectance and water constituent concentrations to non-linear multiple regressions between combination of band ratio(s) and the concentrations. It is not a scope of this research to describe all the statistical algorithms used so far. Interested readers are pointed to, for example, Dekker (1993) and De Haan et al. (1999).

There is a relatively new trend in the statistical approach for the water quality determination to use the derivatives of the measured spectra. Along with the band ratio, this method is considered to be the way to separate spectral effects of different water constituents (Han and Runquist, 1997; Fraser, 1998; Lahet et al., 2001).

However, bearing in mind the time limitation for this research, only commonly used statistical methods for the quantification of water quality parameters were chosen such as:

- Regression between reflectance and water quality parameters;
- Regression between first derivative of spectra and water quality parameters;
- Regression between band ratios and water quality parameters.

Due to the limitations of statistical approach, currently there is an ongoing shift from using empirical algorithms to semi-analytical and analytical methods.

2.4.2. Analytical approach

An analytical approach (also called bio-optical modelling) for water quality retrieval relates the subsurface irradiance reflectance (or *subsurface volume reflectance*, or simply *volume reflectance,* Bukata et al., 1995) to the water constituent concentrations. Reason to use volume reflectance is that subsurface irradiance reflectance is nearly independent of atmospheric properties and is almost entirely determined by the optical properties of the water and its constituents (Figure 2.2).

Several models for coastal and inland waters were investigated by Gordon et al.(1975). They are similar to a solution of the radiative transfer equation: volume reflectance is expressed as a function of absorption and backscattering coefficients of the water constituents. The main differences in bio-optical models were in including the water constituents contributing to the subsurface reflectance. In this research, model of Dekker (1993) was used since it was already successfully applied to the

Dutch dataset. The water constituents in the model are: CDOM, CHL and TSM. Commonly, the model looks like[2]:

$$R(0-) = f\frac{b_b}{a+b_b}$$

$$b_b = b_w * B_w + b^*_{tsm} * B_{tsm} * TSM + b^*_{chl} * B_{chl} * CHL$$

$$a = a_w + a^*_{tsm} * TSM + a^*_{cdom} * CDOM + a^*_{chl} * CHL$$

Equation 2.1

Where,

R(0-) is the volume reflectance;
f is a proportionality factor related to the illumination condition and viewing geometry;
b_b is the total backscattering coefficient;
b^*_{tsm}, and b^*_{chl} are the specific scattering coefficients of TSM and CHL respectively;
B_w, B_{tsm}, B_{chl} are the probabilities that light will backscatter back to the sensor from a given water constituent;
a is the total absorption coefficient ;
a^*_{tsm}, a^*_{cdom}, a^*_{chl} are specific absorption coefficients of TSM, CDOM and CHL respectively;
a_w , b_w are absorption and scattering coefficients of pure water;
TSM, CDOM and *CHL* are concentrations of water constituents: TSM, CDOM and CHL respectively.

Equation 2.1 uses specific absorption and backscattering coefficients of the water constituents, commonly known as Specific Inherent Optical Properties (SIOPs). Essentially, it is absorption and backscattering per unit of water constituents (Figure 2.3), which do change from one type of water body to another. For example, specific backscattering and absorption of TSM depend on the particles present in the studied water and often they are not the same as in another, absorption and backscattering of CHL depends on the type of algae present in the studied waters and so on.

Advantage of using the analytical approach is that once the optical properties of studied water bodies are identified, the model could be applied to any remote sensing scenery irrespective to the time of its acquisition. A disadvantage is: the model uses various input parameters, which often are not available.

[2] Multiplication of specific scattering coefficient and backscattering probability called specific backscattering coefficient. For example, the specific backscattering coefficient of TSM, $b_b{}^*_{_tsm} = b^*_{tsm} * B_{tsm}$

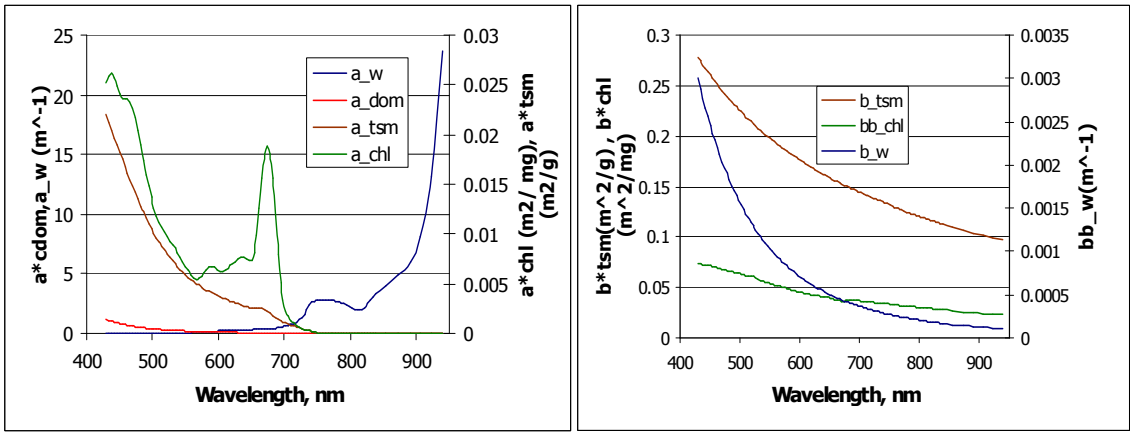


Figure 2.3 Mean specific absorption (left) and backscattering (right) coefficients of the Dutch lakes (Adopted from BIOPTI 1.0, Erin Hogenboom,1995)

In the case of the Hungarian dataset, we did not have any knowledge of the local SIOPs. Due to the many unknown parameters required by the model, some simplifications and assumptions had to be made (the results are discussed in detail in Chapter 4). Thus, author feels more comfortable to name the applied approach *semi-analytical*. Due to the fact that the bio-optical modelling requires air-water corrected spectra, the next part of the methods description will deal with an algorithm for air-water interface correction.

2.4.3. Air-water interface correction algorithm

Upon delivery, images would be atmospherically corrected. However, for the modelling purpose, we need to transfer remote sensing reflectance above water surface to subsurface reflectance (below the water surface). Given transformation is usually done by corrections for the air-water interface.

Morel &Gentili (1993) used Equation 2.2 to convert remote sensing reflectance into subsurface reflectance:

$$R_{rs} = \frac{(1-\rho)*(1-\rho')*R(0-)}{1-r*R(0-)*n^2*Q} + R_{surf}$$

Equation 2.2

Solving Equation 2.2 for subsurface reflectance lead to:

$$R(0-) = \frac{R_{rs} - R_{surf}}{\frac{(1-\rho)*(1-\rho')}{n^2} + r*Q*(R_{rs} - R_{surf})}$$

Equation 2.3

Where, R_{rs} is the remote sensing reflectance;

$R(0-)$ is the volume reflectance;

R_{surf} is a specular reflectance from the surface of the water body;

Q is a ratio of upwelling irradiance to upwelling radiance (5 sr^{-1});

ρ is an internal Fresnel reflectance (0.03);

ρ ' is an air-water Fresnel reflection at the interface (0.54);

n- refractive index of water (1.34) ;

r-water-air reflection (0.54).

Equation 2.2 is used, for example, by Lee et al (1998) for calculating R_{rs} from simulated measured *R(0-)* spectra. Gege (2001) implemented this formula for calculating remote sensing reflectance in the WASI 2.0 (Water Colour Simulator) software. Default values indicated in the brackets of the key of Equation 2.2 are taken from the WASI 2.0 (Gege, 2001).

Upwelling spectra measured above the water can be affected by sun glint from the water surface (specular reflectance, R_{surf}) due to waves or foam (Figure 2.4). In order to detect R_{surf} in measured spectra, we followed the assumption that the light absorption by pure water is predominant in NIR (970-1000 nm) and the water-leaving radiance in that region is zero (Ouillon et al., 1997). However, Doxaran et al. (2002) showed that this assumption does not hold in highly turbid waters and they proposed averaging of the successive measurements. On the other hand, Han and Rundquist (1998) observed that the averaged reflectance spectra are still influenced by the specular reflectance. Thus, considering the TSM range found in Hungary we assume that the waters are not yet highly turbid and we follow the assumption of Ouillon et al (1997).

Considering the gradually increasing noise in the collected spectra, we found minimum reflectance at 930 nm (Figure B.1, Appendix B). The given value is used to remove specular reflectance and wave effects (assuming that they are wavelength independent) by subtracting it from the whole spectra.

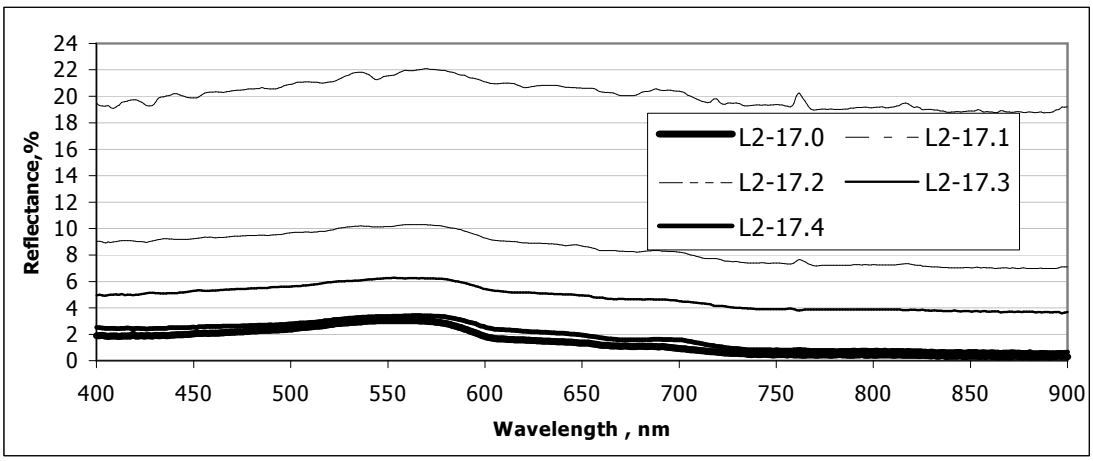


Figure 2.4 Effect of waves on the spectral measurements at the sampling point L2 on 17th August

After normalizing spectra, the air-water correction formula was applied. Figure 2.5, which presents two examples of air-water interface correction application, shows that reflectance of corrected spectra are generally higher than the remote sensing reflectance.

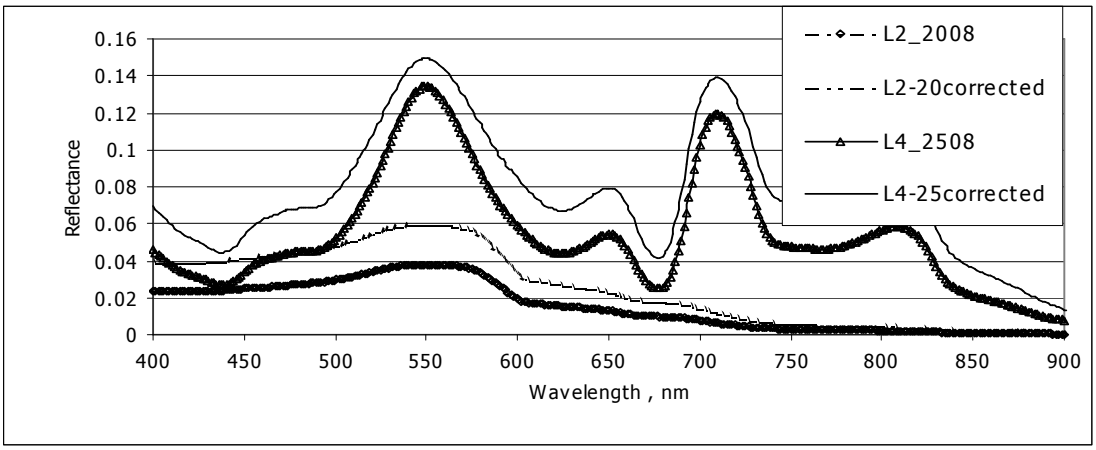


Figure 2.5 Examples of R_{rs} and corresponding R(0-) spectra

2.4.4. Limitations of the methods

It should be noted that there are several limitations in both approaches. Common limitation for both of them is that remote sensing data generally receives information from the top layer of the water body (depending on how deep and "clean" the water is). A commonly used measure of a depth to which the light penetrates and reflects back to the sensor is the Secchi Disc depth. Consequently, quantification of water quality parameters is restricted to that depth only. Water bodies with larger depth then the Secchi Disc depth are considered to be optically deep.

As it was pointed out already, there might be a reflectance from the bottom of the water body that is generally higher than the signal from the optically deep parts (depending on the bottom type). Though presently there are several models available to eliminate signal from the bottom or to use it in a shallow water bathymetry mapping (Maritorena et al., 1994), in this research we made it sure that the spectra were taken from optically deep waters, and in the image processing we assume that the bottom of the water body does not influence the remote sensing signal (this results in some errors at shallow waters).

There are specific limitations for the bio-optical modeling due to its main assumption: modeled water quality parameters are distributed evenly within the water column, which is not always true, especially in the dynamic flowing systems as rivers.

Some limitations of the statistical approach have been already mentioned before. It is important to add here that statistical models (as well as simplified analytical models) are generally applicable only to the range they were defined.

Chapter 3. Results and discussions on using statistical approach

This chapter provides an overview of applying the statistical approach. It was arranged in the way to present results and discussions for each studied water constituent. The chapter starts with the description of statistical relationships between the water quality parameters themselves.

3.1. Statistical relations between the water quality parameters

First, a correlation analysis was performed between the measured in Hungary water quality (WQ) parameters. Very high correlations were obtained between TSM and ISM, Turbidity and OSM (r =0.99 and r =0.93 respectively, Table 3.1)

Table 3.1 Correlation coefficients between WQ parameters

	TSM, mg/l	*ISM, mg/l*	*OSM, mg/l*	*Turbidity*	*CHL, mg/l*	*TSM+Chla*
TSM, mg/l	1					
ISM, mg/l	0.992	1				
OSM, mg/l	0.886	0.818	1			
Turbidity	0.892	0.846	0.931	1		
CHL,mg/l	0.182	0.107	0.425	0.518	1	
TSM+Chla, mg/l	0.742	0.686	0.837	0.905	0.794	1

Due to the strong correlation between TSM and ISM (Figure 3.1), a two-step procedure is applied to estimate ISM from the spectra: first, TSM will be estimated using spectral data, and then ISM can be estimated from those of TSM estimations.

To check the relation between TSM and ISM on the Dutch dataset, the given regression line (ISM=0.801*TSM –2.243, Figure 3.1) was plotted on a graph of measured TSM and ISM concentrations from the Dutch lakes. Figure 3.2 shows that in general the established algorithm can not be applied for defining ISM from TSM in the Dutch dataset, although some points do follow the established trend line. It leads to the conclusion that the TSM of the Dutch data set had two distinct sources: ISM dominated lakes (which follow trend line plotted from the Hungary data set) and OSM (particularly CHL) dominated lakes (a distinct group below the trend line, Figure 3.2).

Thus, keeping in mind the established relation between TSM and ISM for the Hungarian dataset, the following statistical (and semi-analytical) analysis is based on the next parameters:

- CHL;
- TSM;
- Turbidity.

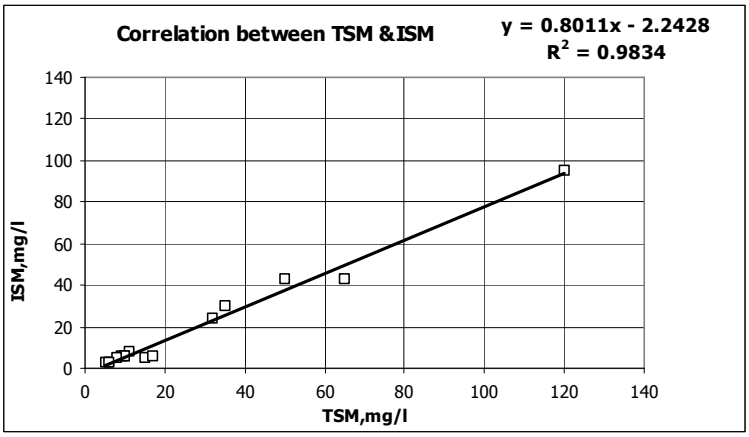


Figure 3.1 Correlation between TSM and ISM for the Hungarian dataset

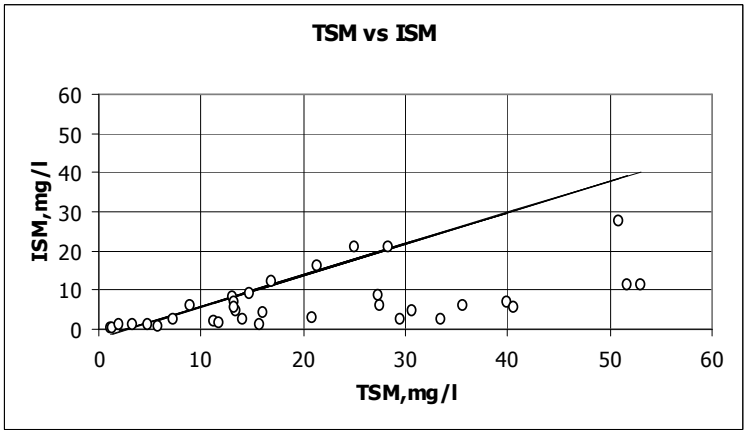


Figure 3.2 Measured TSM and ISM concentrations from the Dutch lakes with the trend line plotted using regression equation established in the Hungarian dataset (see Figure 3.1)

3.2. Statistical analysis of the CHL content and corresponding spectra

3.2.1. Spectral signatures of waters with high CHL content

Four characteristic spectral features associated with the CHL content were observed (Figure 3.3 and Appendix B):

1. Low reflectance between 400 and 500 nm due to pronounced absorption by algae (CHL) and CDOM (Han and Rundquist, 1997 and Figure 2.3);
2. Reflectance about 550 nm from algae biomass, coupled with strong backscattering from inorganic suspended sediments (Han and Rundquist, 1997 and Figure 2.3)
3. Minimum reflectance at 675 nm caused by Chlorophyll-a absorption (Dekker, 1993 and Figure 2.3)
4. Prominent reflectance maximum at about 705 nm due to minimum in absorption of all components and thus mainly scattering (Thiemann and Kaufmann, 2002).

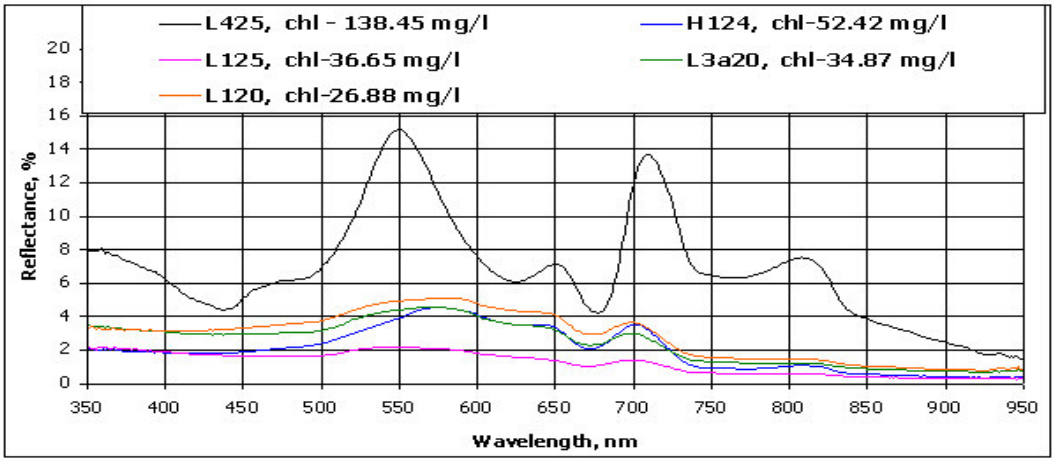


Figure 3.3 Reflectance spectra of water bodies with the highest chlorophyll content

3.2.2. Linear regression between reflectance and CHL [3]

Regression analysis between the measured CHL and spectra (for the Hungarian data set) shows that the highest regression coefficient for CHL (r^2=0.56, N=15) is at 716.22 nm (Figure 3.4, in case of the field spectra), close to the second reflectance peak. Similar results were obtained for the other sensors (Table 3.2).

Table 3.2 Summary of the regression analysis between CHL and single band reflectance (indicated wavelengths show where the highest regression coefficients were observed)

Sensor	*Wavelength, nm*	*r^2*	*Slope*	*Intercept*	*St. error*
Field Spectrometer (FS),%	716.22	0.56	8.53	-3.87	23.74
ROSIS,%	718	0.56	8.81	-3.65	23.81
DAIS,%	729	0.54	11.38	-3.91	24.44
Landsat TM,%	834	0.58	18.27	4.66	23.32

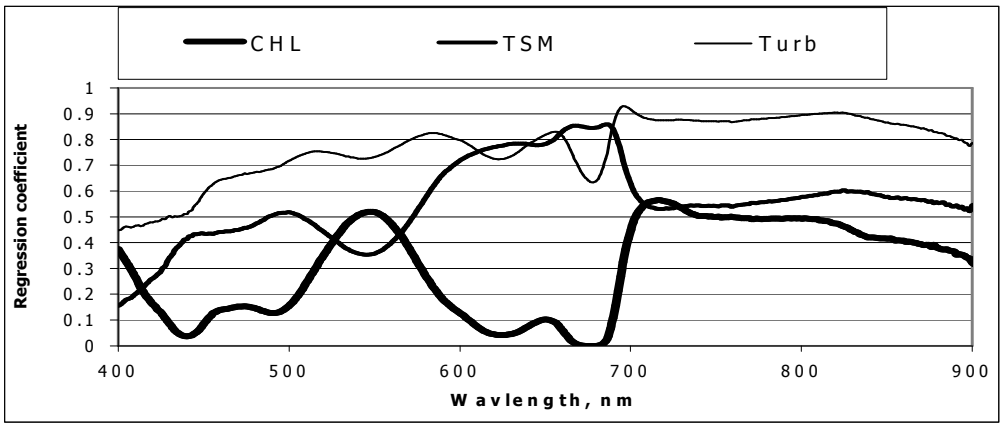


Figure 3.4 Regression coefficients between the WQ parameters and reflectance (Sensor: Spectrometer)

[3] In all cases wavelength indicates spectral location where the highest regression coefficient was obtained. Also, reflectance values used in the statistical analysis were in %.

3.2.3. First derivatives and CHL

First derivatives (or local gradients) were calculated using the following equation:

Equation 3.1

$$DR = \frac{R(\lambda 2) - R(\lambda 1)}{\lambda 2 - \lambda 1}$$

Where,

DR is the first derivative at location [(λ2+λ1)/2] nm;
R (λ2) and *R (λ2)* are reflectance in wavelength λ2 and λ1 respectively (λ2>λ1);
λ2 -λ1 is bandwidth.

The graph of first derivatives showed "noisy" character in case of spectra collected with the spectrometer. Given inconsistency was found by many observers (Tsai and Philpot, 1998) and a commonly used measure for it is smoothing. There are still disputes in literature about the smoothing technique: smoothing assumes a loss of information due to averaging. Assuming that the same loss would be due to resampling to a coarser resolution sensor, in this section only reflectance spectra resampled to the ROSIS, DAIS and Landsat TM sensors were used for the derivative analysis.

Linear regression between CHL and first derivatives showed that the highest correlation coefficient (r^2=0.93) is determined by the gradient from the second CHL absorption band to second reflectance peak, at 684 nm (Table 3.3).

Table 3.3 Summary results of the regression analysis for the first derivatives and CHL

Sensor	*Wavelength, nm*	*r^2*	*Slope*	*Intercept*	*St. error*
ROSIS	684	0.93	601.709	4.775	9.7
DAIS	686.5	0.92	578.561	15.636	10.0
Landsat TM	610.9	0.01	-280.015	26.142	35.9

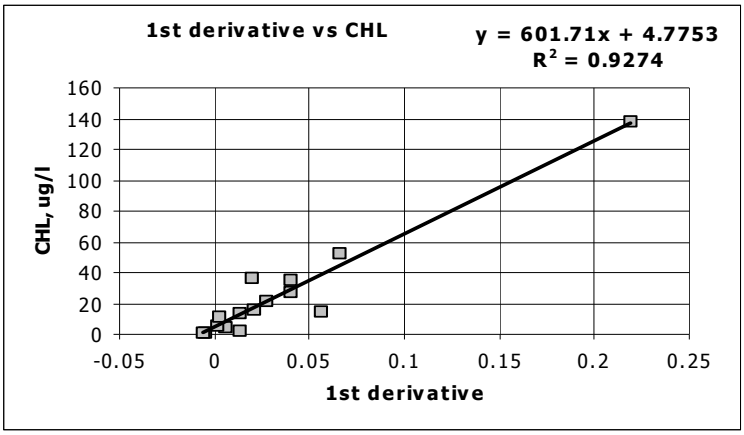


Figure 3.5 First derivatives at 684 nm vs CHL for the ROSIS sensor

3.2.4. Band ratio NIR/Red and CHL.

Band ratio is a widely used technique for determination of CHL content in the water, since it is relatively insensitive to illumination conditions and to the presence of suspended sediments (Lahet et al., 2001). Summary of the regression analysis using band ratios are presented in Table 3.4 and Figure 3.6 to 3.8.

Table 3.4 Summary of the regression analysis between band ratios and CHL

Sensor	*Wavelength, nm*	*r^2*	*Slope*	*Intercept*	*St. error*
Spectrometer	704.46/672.17	0.98	63.096	-50.321	5.44
ROSIS	706/670	0.98	66.594	-51.816	5.47
DAIS	693/675	0.97	200.47	-184.83	6.5
Landsat TM	834/661	0.73	159.758	-7.078	18.64

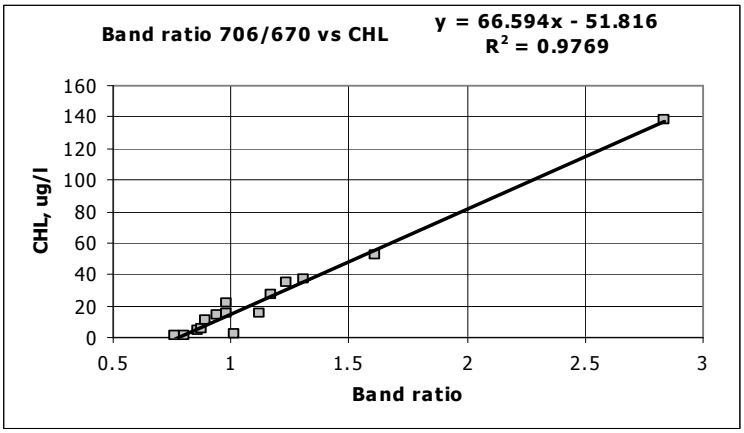


Figure 3.6 Band ratio versus CHL for the ROSIS sensor

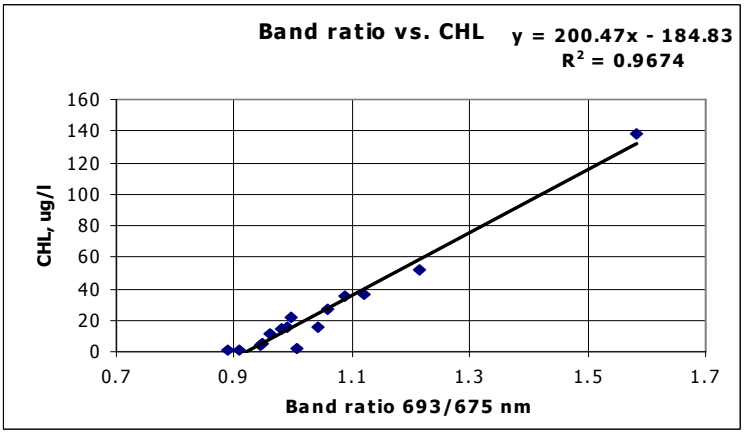


Figure 3.7 Band ratio 693/675 vs CHL for the DAIS sensor

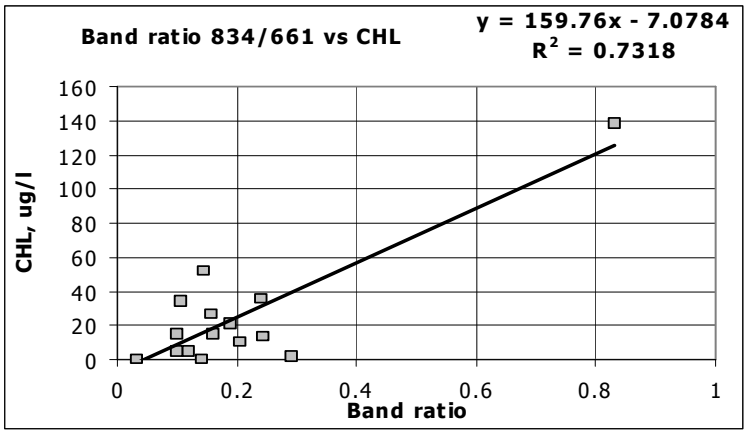


Figure 3.8 Band ratio 834/661 vs CHL for the Landsat TM sensor

3.2.5. Discussion of results on using statistical approach to retrieve CHL

Chlorophyll-a shows two diagnostic absorption bands centered at around 440 and 675 nm. The reflectance peaks at around 550 and 705 nm rise with increasing chlorophyll concentration. It seems intuitively logical that the slope of the line at about 690 nm would be important for measuring chlorophyll concentration because of its "spectral location" between 675 nm red absorption maximum and 705 nm NIR reflection maximum. In any case, this indicator seems to allow one to "disaggregate" the composite spectral response from the surface waters to yield a separate chlorophyll signal.

As a result, the band ratio algorithms perform better than the first derivative and the single band reflectance algorithms (Table 3.2 to 3.4). However, with decreasing spectral sensitivity of the sensor (see Appendix B) in the region of 670-705 nm, accuracy of retrieving CHL also decreases, which tendency is indicated by both methods: first derivative and the band ratio (Table 3.3 and Table 3.4). Rather high regression coefficient for the Landsat TM (r^2=0.73) is due to the presence of an outlier (Figure 3.8), without it the regression coefficient reduces drastically (r^2=0.01). This shows that statistical algorithms to quantify CHL are generally not applicable to the Landsat TM, because of its coarse spectral resolution (Figure 4.10).

Summarizing results, our conclusion is: for a given range (1-140 µg/l) for the Hungarian dataset CHL content is best to determine using band ratio NIR/Red algorithm. Proposed equations are[4]:

- For the spectrometer's wavelength: $CHL = 63.096*(R704.46/R672.17) - 50.321$
- For the ROSIS sensor: $CHL = 66.594*(R706/R670) - 51.816$
- For the DAIS sensor: $CHL = 200.47*(R693/R675) - 184.83$
- For the Landsat TM: $CHL = 159.76*(R834/R661) - 7.0784$

3.2.6. Testing and final conclusion on using statistical algorithm for CHL

To evaluate a performance of the proposed algorithms, they were applied to the Dutch dataset. Overall number of measurements was 34. However after removing measurements on lakes with higher

[4] R indicates reflectance

CHL concentration (>160 mg/l) than the range found in the Hungarian dataset, the number reduced to 29.

The spectra of the Dutch dataset looked quite similar to the ones of the ROSIS sensor (both had 4 nm interval), but with 2 nm displacement. Since for the spectrometer the proposed rationing wavelengths are 704.46 and 672.17 nm, the choice of the rationing wavelengths for the Dutch set was for 704 and 672 nm (compare to the ROSIS wavelengths 706/670).

Figure 3.9 presents the results of the band ratio (704/672) algorithm from the Dutch dataset with line plotted using regression equation of the ROSIS sensor: *CHL = 66.594*(R704/R672) – 51.816*. However, other proposed regression equations were also tested, but the best suited was found to be the proposed equation of the ROSIS sensor.

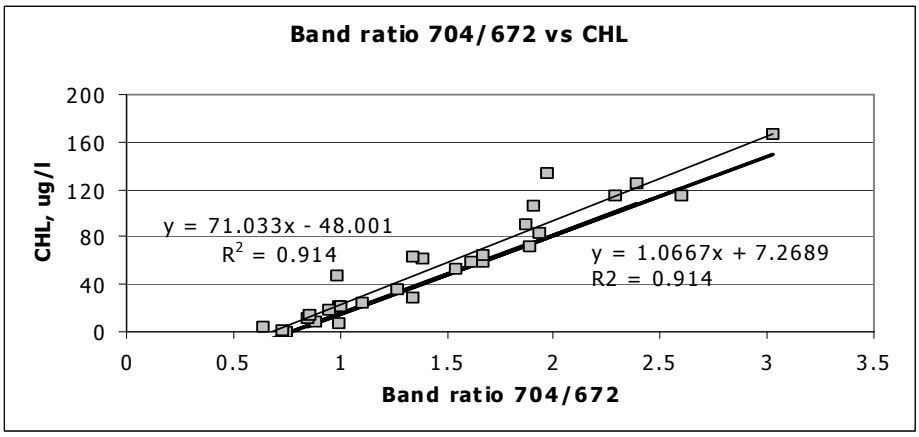


Figure 3.9 Band ratios vs. CHL concentrations for the Dutch lakes with plotted regression lines:

- *Thin one is the trend line from the dataset itself with an equation on the top of the line;*
- *Thick one is from the proposed algorithm of the ROSIS sensor with an equation between the modelled and observed CHL (below the line).*

Regression analysis between the modelled (using the equation for the ROSIS sensor) and the observed CHL concentrations in the Dutch dataset showed a standard error of about 13 µg/l versus 6 µg/l for the Hungarian dataset.

For comparison, first derivative method was also tested on the Dutch dataset. However, the results are lower than that of band ratio (N=29, r^2=0.78) and the spectral location where the highest regression coefficient was found, shifted towards longer wavelengths (694 nm versus 684 nm in the Hungarian dataset, Table 3.3).

The ANOVA test performed on both trend lines of Figure 3.9 showed absence of a significant variation between two trend lines (N=58,α =0.05, p=0.37). Additionally, the developed models are similar to the ones found by other researches, for example, in Thiemann and Kaufmann, (2002): *CHL = 73.59*BR (R705nm/R678nm)-52.9*1.

Thus to conclude: to retrieve CHL concentrations, preference among the statistical models is given to the band ratio algorithm as it performed well on both datasets.

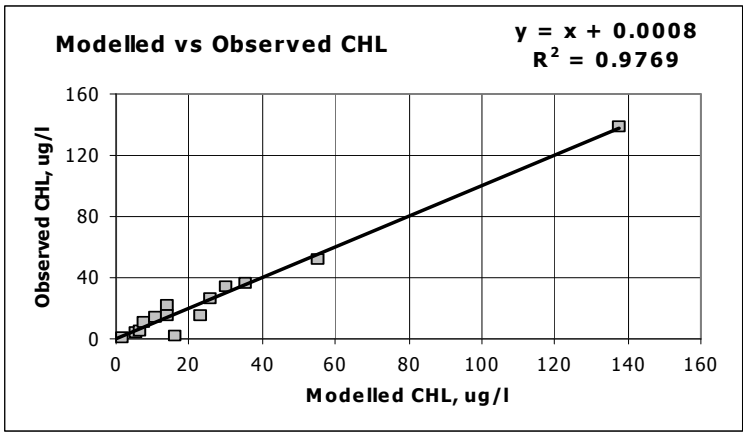


Figure 3.10 Modelled vs. Observed for the Hungarian dataset (applying to the ROSIS sensor's wavelengths)

3.3. Statistical analysis of the TSM and corresponding spectra

Previous researches on correlation between TSM and corresponding spectra show a vast variety of methods applied, as well as absence of any single agreement on which statistical algorithm and on which specific band(s) combinations to use. Essentially, it is due to the influence of TSM on the spectra.

3.3.1. Spectral signatures of waters with high TSM concentrations

Principally, an occurrence of only TSM in the water increases reflectance spectra (Figure 3.11), but the presence of other constituents, such as CDOM, CHL and the water itself with their absorption and backscattering properties (Figure 2.3), makes the quantification of TSM a difficult task. The main information about TSM is backscattering, since its absorption has the same shape as CDOM (though much lower in the magnitude).

There is a small window (around 680-720 nm) between the absorption of CHL and the absorption of water (Figure 2.3) where the backscattering of TSM can be "visible" using statistical approach. Many statistical algorithms exist for estimating TSM (mostly simple linear regressions between reflectance on a single wavelength and TSM concentrations) in that spectral region. Such algorithms were also was found in both of our datasets, as the next section illustrates it.

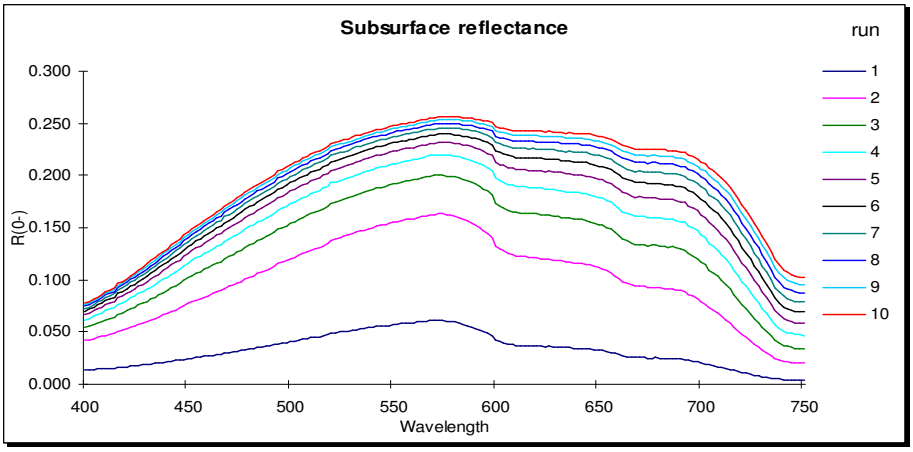


Figure 3.11 Result of simulating subsurface volume reflectance by changing TSM from 5mg/l (run1) to 185mg/l(run 10) with 20mg/l step (done by using BIOPTI 1.0 software, Erin Hogenboom,VU, 1995)

3.3.2. Results and discussion on using statistical approach for TSM quantification

Summary of the results on the use of statistical approach (Table 3.5) indicates an agreement with the previous researches although the correlation found between TSM and the reflectance at 686 nm is more logarithmic than linear (Figure 3.12).

Spectrometer	*Regression function*	*Wavelength, nm*	r^2	*Slope*	*Intercept*	*St. error*
Reflectance,%	LN (TSM)	686.83	0.91	0.50	1.05	0.30
First derivative	TSM	770.02	0.82	3699.13	24.07	13.89
Band ratio	LN (TSM)	810.66/536.15	0.76	6.48	0.58	0.49
ROSIS						
Reflectance,%	LN (TSM)	686	0.91	0.50	1.06	0.30
First derivative	TSM	804	0.87	3993.47	10.27	15.84
Band ratio	LN (TSM)	810/538	0.76	6.60	0.56	0.49
DIAS						
Reflectance,%	LN (TSM)	675	0.91	0.50	1.07	0.30
First derivative	TSM	771	0.69	262.72	2.85	0.56
Band ratio	LN (TSM)	798/532	0.73	6.71	0.52	0.52
Landsat TM						
Reflectance,%	LN (TSM)	661	0.84	0.4	0.94	0.39
First derivative	TSM	747.8	0.037	-12.996	27.89	32.02
Band ratio	LN (TSM)	834/561	0.65	0.094	-0.12	0.07

Table 3.5 Summary of results on use of the statistical approach to retrieve TSM concentrations

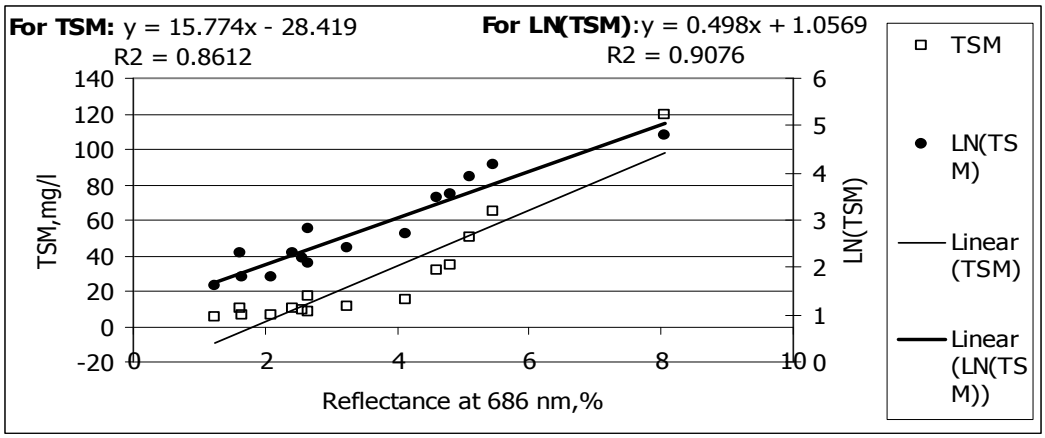


Figure 3.12 Regression lines and coefficients between reflectance at 686 nm and TSM concentrations for the Hungarian dataset (ROSIS sensor)

Nevertheless, the test of all the above-mentioned algorithms on the Dutch datasets did not yield similar results. Similar to the relation on Figure 3.12, the regression coefficient of the Dutch set was highest at 712 nm and the relationship was linear (Figure 3.13) in contrast to the Hungarian dataset

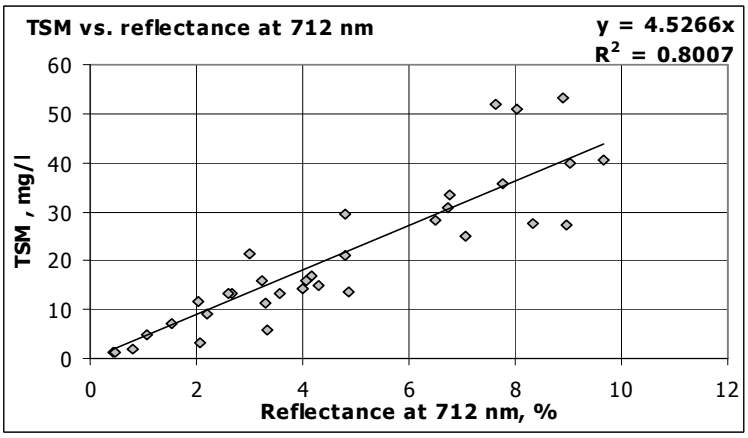


Figure 3.13 Regression line between TSM and reflectance at 712 nm for the Dutch data set

Therefore, it is rather difficult to predict the exact wavelength where the highest correlation can occur. The algorithm is also affected by the number of measurements and by the origin of TSM present in the studied waters: TSM consisting mainly of phytoplankton cells will have more absorption at 680 nm and subsequently there would be a shift of algorithm towards longer wavelengths, as in the case of the Dutch dataset.

Knowing that in the study area we had two different sources of TSM: phytoplankton dominated TSM in the lakes and ISM dominated in the rivers, we made an attempt to separate statistical analysis for the rivers and lakes. However, the wavelengths where the highest regression coefficients found once more did not agree with the Dutch dataset (694 nm for the rivers and 730 nm for the lakes).

Since the main information about TSM is its backscattering, using band ratios or first derivative is not very helpful: backscattering of TSM is gradually decaying along with the backscattering of other constituents. There is one situation where this is not a problem, and that is when the TSM mainly

consists of phytoplankton cells. In this case CHL and TSM are strongly correlated and then band ratios are (indirect) indicators of TSM (Steef Peters, personal communication, 2002). Thus in the result, no commonly applicable (for the given 2 datasets) statistical algorithm for TSM quantification was found, although high correlation was found in each individual data set.

3.4. Statistical analysis of turbidity and corresponding spectra

Turbidity is a measure of the transparency of the water body. As was mentioned before, unfortunately, turbidity values were not recorded in the Dutch dataset.

This is a controversial parameter: on the one hand, turbidity is relatively easy to measure, on the other hand, turbidity is a rather loose term - it is a lumped consequence of the inorganic (ISM) and /or organic materials (OSM, CHL and CDOM) presence in the water column (Bukata et al. 1995). That is why – perhaps – there was not much attention paid to this parameter. Additionally, the bio-optical modelling does not include this parameter. Therefore, our estimation of turbidity entirely depends on the statistical algorithm developed on the Hungarian dataset.

In our case, turbidity was due to presence of both inorganic and organic materials: correlation analysis carried out between water constituents showed high correlations between turbidity and OSM, and between turbidity and the sum of (TSM +CHL) (Table 3.1).

Statistical analysis of turbidity and corresponding spectra (Table 3.6, Figure 3.14 and 3.15) showed highest correlations of reflectance at 695.64 nm and first derivative of 775.92 nm. The high slopes and the low derivative values make the application of the derivative algorithm rather uncertain (Table 3.6). In addition, both algorithms are biased by the high values (Figure 3.14 and 3.15). Nevertheless, both algorithms were applied to the images and compared with 3 field measurements of turbidity carried out on 17th of August during the first flight.

Table 3.6 Summary of results using statistical approach for the turbidity determination

Spectrometer	*Wave length, nm*	*r^2*	*Slope*	*Intercept*	*St. error*
Reflectance,%	695.64	0.93	6.20	-9.79	4.34
First derivative	775.92	0.97	1997.48	7.72	2.98
Band Ratio	823.99/418.77	0.90	48.76	-11.29	5.16
ROSIS					
Reflectance,%	698	0.93	5.72	-8.20	4.44
first derivative	772	0.94	3330.85	11.87	4.10
Band Ratio	826/430	0.89	50.40	-10.89	5.50
DIAS					
Reflectance,%	693	0.93	6.81	-10.6	4.08
First der	789	0.93	2492.2	7.25	4.31
BR	802/798	0.85	68.88	-15.96	6.31
Landsat TM					
Reflectance,%	834	0.85	10.08	3.17	6.24
First derivative	747.8	0.02	-5.16	15.13	16.17
Band ratio	834/479	0.87	73.87	-1.57	5.91

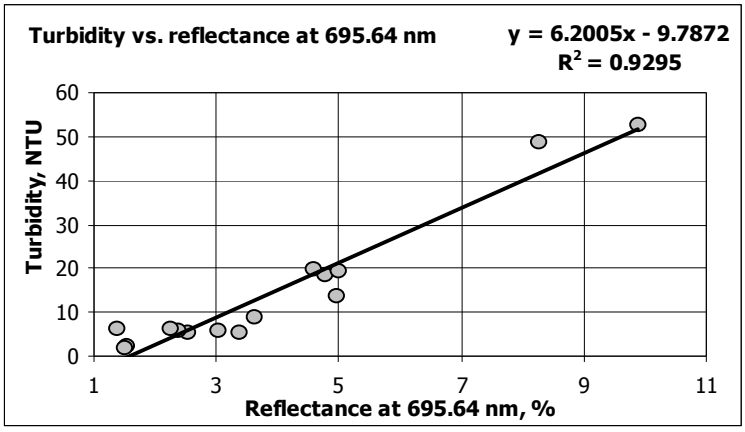


Figure 3.14 Regression line between Turbidity and Reflectance at 695.64 nm for the Hungarian dataset (Sensor: Spectrometer)

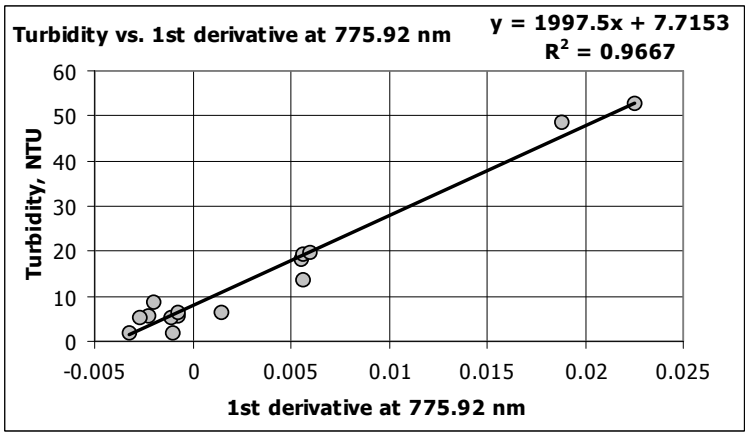


Figure 3.15 Turbidity vs. first derivatives at775.92 nm (Sensor: Spectrometer)

Chapter 4. Results and discussion on using semi-analytical approach

As described in the previous chapter, no universally applicable statistical algorithm was found for the TSM mapping of both data sets. Knowing the capabilities of (semi) analytical methods to minimize *in situ* requirements for the water quality monitoring and to utilize archive images, an attempt has been made to apply the bio-optical model for the quantification of TSM and for the CHL estimation using bio-optically modelled band ratios. This chapter starts with the bio-optical modelling of TSM.

4.1. Bio-optical modelling of TSM

Due to many unknown parameters for the bio-optical modelling, we had to make some simplifications. Unknown parameters include (see Equation 2.1, Chapter 2): SIOPs of the studied water bodies, proportionality factor *f*, and backscattering probabilities of water constituents. The SIOPs of the water itself are explored quite well (Bukata et al.1995) and a use of some common values of the proportionality factor and backscattering probabilities can be found in the literature (Gordon et al., 1975, Bukata et al. 1995). However, the values used by other researchers for absorption and backscattering properties of CHL and TSM were varying by a magnitude of 10. In addition to the absence of knowledge of the SIOP we did not have information on the amount of CDOM present in our samples. Altogether, this made us to find a spectral region, where the influence of other parameters compared to TSM would be minimal.

4.1.1. Identification of the spectral region where influence of other parameters than TSM is minimal.

Bukata et al. (1995) have conducted research on the impact of water constituents on different regions of spectrum. Their results indicate that volume reflectance in NIR region is mostly influenced by the backscattering of TSM, if the TSM concentrations are more than 0.1 mg/l. Babin and Stramski (2002) have proved the research of Bukata et al. by measuring absorption properties of different suspensions. They concluded that absorption properties of other constituents than water in NIR region are negligible.

Thus by assuming that NIR region of volume reflectance spectra is mostly affected by backscattering properties only, from the bio-optical modelling of NIR region, Equation 2.1 (See Chapter 2) reduces to:

$$R(0-) = f\frac{b_b}{a+b_b}$$

$$b_b = b_w * B_w + b^*_{tsm} * B_{tsm} * TSM + b^*_{chl} * B_{chl} * CHL$$

$$a = a_w;$$

Equation 4.1

where,

R(0-) is the volume reflectance at Nth band in NIR region;
f is a proportionality factor related to the illumination condition and viewing geometry;
b_b is the total backscattering coefficient;
a is the total absorption coefficient ;
a_w , b_w are absorption and scattering coefficients of pure water;
TSM and *CHL* are the concentrations of TSM and CHL;
B_w, B_{tsm}, B_{chl} are the probabilities that light will backscatter back to the sensor from a given water constituent;
b^*_{tsm} *and* b^*_{chl} is the specific scattering coefficients of TSM and CHL respectively;

Thus, the main assumptions were:

- Absorption in the NIR is only due to the absorption of water
- Backscattering in NIR is due to the backscattering of water, CHL and TSM

Substituting as:

$b_{btsm} = b^*_{tsm} * B_{tsm} * TSM$, where b_{b_tsm} is a total backscattering by TSM only,
$b_{b_chl} = b^*_{chl} * B_{chl} * CHL$, where b_{b_chl} is a total backscattering by CHL only,

and solving Equation 4.1 for C_{tsm} in the NIR region leads to the following expression:

$$b_{b_tsm} = \frac{f * b_w * R(0-) + f * b_{b_chl} - R(0-) * b_{b_chl} - R(0-) * (a_w + b_w * B_w)}{R(0-) - f}$$

and

$$TSM = \frac{b_{b_tsm}}{b^*_{tsm} * B_{tsm}}$$

Equation 4.2

To find NIR region where the influence of other parameters but TSM would be negligible, results of statistical analysis were used, particularly regression coefficients between TSM concentrations and corresponding spectra (Figure 3.4, see Chapter 3). The values indicate the highest regression coefficients near to 682 nm (red region) and from 710 nm onwards (NIR region). The modelling was performed on both regions, though the region at 682 nm is still influenced by other WQ constituents than TSM (a^*_{chl} and a^*_{cdom}), which was proved, later by modelling (Figure 4.5).

In this work, *f* is assumed to be constant and equal to 0.33, which is a good approximation for clear sky condition and high solar elevation (Gordon et al., 1975). Values for backscattering probabilities of pure water, TSM and CHL are used from Bukata et al (1995): B_w=0.5, B_{tsm} =0.08, B_{chl} =0.016. Specific scattering coefficients of TSM and CHL are taken from BIOTPI 1.0 software (Erin Hogenboom,VU,1995).

4.1.2. Preliminary modelling

Initially modelling was performed using the mean specific backscattering coefficients of the Dutch lakes (Erin Hogenboom,1995). Results indicated poor correlation between modelled and observed TSM concentrations for the whole Hungarian data set. However, an analysis of the obtained results showed that overall regression coefficients were low due to the modelled TSM from the rivers only (sampling points R4_2008, R2_2408, R2-3_2408, R1_2508 and H1_2408).

Further, by trying to find out the "best fit" specific backscattering coefficients for River Sajó revealed that the backscattering of TSM is changing with time (it required 3 different sets of backscattering coefficients for the rivers on 20th, 24th and 25th August). During the flight on 17th August, the Sajó River had a higher flow velocity than on the other sampling dates, because the river stage was dropping after a flood. The particle size of the suspended matter is proportional to the flow velocity and consequently to the discharge (example in Figure 4.1, Steegen et al., 1998). This leads to the conclusion that the river's TSM backscattering is influenced by the particle size of the suspended matter. Given conclusion is proved by many authors, who used Mie theory (Van de Hulst, 1957) to model specific backscattering of TSM depending on particle's diameter.

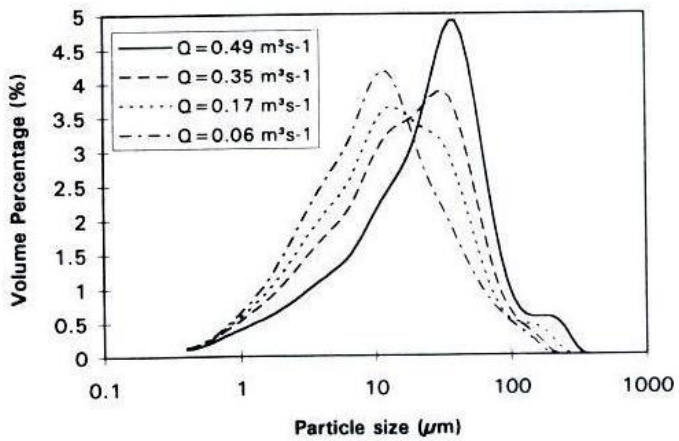


Figure 4.1 Grain-size distribution of sediment sampled at the different discharges (Q - discharge) in central Belgium (Adopted from Steegen et al., 1998).

Literature review showed that Gege (2001) used two different sets of backscattering coefficients for modelling TSM in Lake Constance: backscattering by large and small particles where the first one is wavelength independent. To define the terms large and small, Mie theory (Van de Hulst, 1957) suggests that backscattering is wavelength independent if particles are "much larger" than the wave-

length. Assuming roughly that "much larger" means a factor of 10, than particles sizes should be above 0.01 mm for the spectral region around 1000 nm (NIR). Thus, introducing specific backscattering by large and small particles leads Equation 4.2 to the following (Gege, 2001):

$$TSM_{large} = \frac{b_{b_tsm}}{b^*_{b_tsm_large} + \frac{b^*_{b_tsm_small}}{r}}$$

$$TSM_{small} = \frac{TSM_{large}}{r}$$

Equation 4.3

where,

$b^*_{b_tsm_large}$ is a specific backscattering by large particles of TSM;
$b^*_{b_tsm_small}$ is a specific backscattering by small TSM particles;
r is a ratio of large particle's concentration to small particle's concentration;
TSM_{large} *and* TSM_{small} are the concentrations of large and small TSM particles respectively.

A study undertaken jointly by RIZA and VITUKI (1994) illustrated that at the average diameter of suspended sediment in Sajó River is 0.04 mm. About 30% of the suspended sediment is falling into the fine sand category (between 0.05 and 0.1 mm) and about 10% is finer than 0.006 mm. Assuming the exponential function of flood retrieving from 17th August (high flow) to 25th August (low flow) with an average of 27% of small particles (finer than 0.01 mm), following *r* values were used for further modelling:

Date	*Ratio of large to small particle's concentration, r*	*Percentage of Large particle's concentration*	*Percentage of Small particle's concentration*
17	4.79	83	17
18	4.19	81	19
19	3.66	79	21
20	3.20	76	24
21	2.796	74	26
22	2.44	71	29
23	2.13	68	32
24	1.86	65	35
25	1.63	62	38
Average		*73%*	*27%*

Table 4.1 Assumption values of ratio of large to small particle's concentration in the Sajó River during the flood retreat.

According to Yun Liu (2003), who performed the particle's size distribution test of samples taken on 19th August from the fresh sediment traps left after the flood in the study area, about 18.5 % of sediments consists of fine materials with the following distribution: silt (particles size 0.002 - 0.05 mm) –

8.9% and clay (particle size < 0.002 mm) – 9.6%. Although silt includes particles greater than 0.01 mm, the result of the test indirectly supports the values of ratio large to small particles used in the modelling.

Table 4.2 shows input parameters (besides the remote sensing reflectance corrected for air-water interface) used for the TSM modelling using DAIS sensor's properties:

Wavelength, nm	a_w	b_w	$c_w=(a_w+b_w*B_w)$	$b^*_{tsm_small}$	$b^*_{tsm_large}$	b^*_{chl}	Constants:	
747	2.498243	0.000277	2.498382	0.13247	0.01	0.033556	B_{tsm}	0.08
766	2.709976	0.000248	2.7101	0.127604	0.01	0.03207	B_w	0.5
783	2.564723	0.000226	2.564836	0.123799	0.01	0.031035	B_{chl}	0.016
802	2.22497	0.000203	2.225072	0.11987	0.01	0.029988	f	0.33
819	2.369144	0.000186	2.369237	0.116552	0.01	0.029086		
837	3.221024	0.000169	3.221109	0.113193	0.01	0.028157		
854	4.042807	0.000155	4.042885	0.110183	0.01	0.027309		
873	4.921592	0.000141	4.921663	0.106978	0.01	0.026405		
890	6.023928	0.00013	6.023993	0.104243	0.01	0.025753		

Table 4.2 Input parameters for TSM modelling (resampled to the DAIS sensor)

4.1.3. Results of bio-optical modelling of TSM

Thus, by using the above estimated ratio of concentration of large to small particles, mean specific scattering by small particles from the Dutch lakes, specific scattering by large particles from Lake Constance and mean specific backscattering of CHL from the Dutch lakes, the following results were obtained from modelling of TSM:

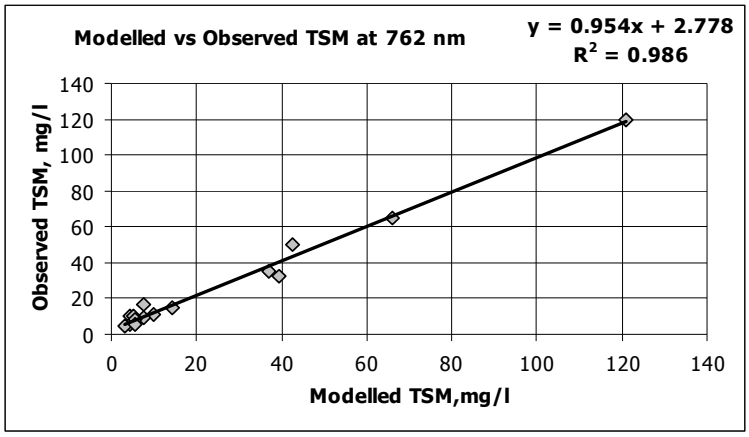


Figure 4.2 Regression line between the observed and modelled TSM concentrations at 762 nm using resampled to the ROSIS sensors reflectance

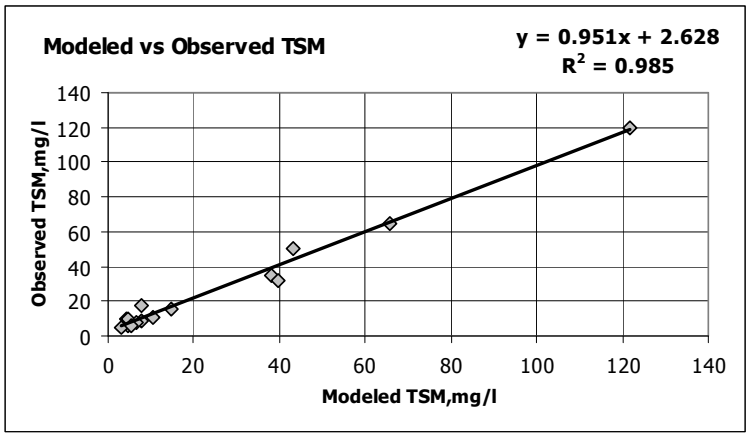


Figure 4.3 Regression line between the observed TSM and the modelled TSM concentrations at 747 nm using resampled to the DAIS sensor reflectance

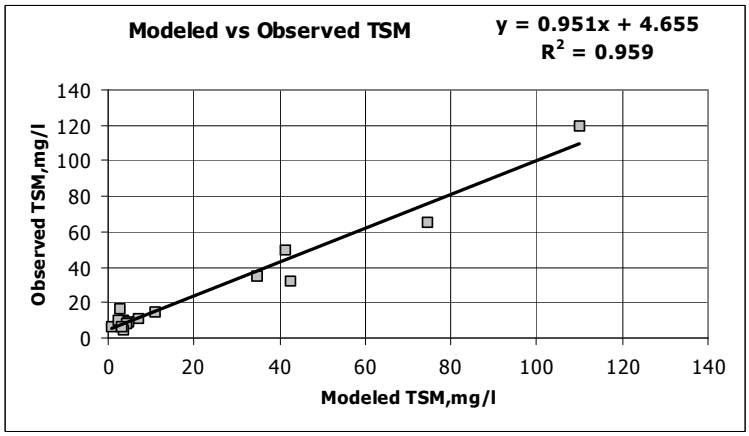


Figure 4.4 Regression line between the observed and modelled at 834 nm TSM concentrations using resampled to the LandsatTM sensor reflectance

4.1.4. Discussion of the results

In general the results indicate that the modelled values are very close to the observed ones. Regression equations for the 3 sensors indicate that slope is nearly 1 and intercept varies between 3 and 5 (2.78, 2.63 and 4.66 for the ROSIS, DAIS and Landsat TM sensors respectively) which indicates that the model slightly underestimates TSM concentrations for about 4 mg/l. Regression coefficients (r^2) between the modelled and observed TSM in the range of 700 to 900 nm show similar values within the sensors with the highest at 762 nm (for the ROSIS sensor) and 766 nm (for the DAIS sensor) (Figure 4.5). Standard errors between the observed and modelled concentrations of TSM greatly varied between the sensors: 4.2mg/l for the ROSIS and DAIS and 7.2mg/l for the Landsat TM sensor.

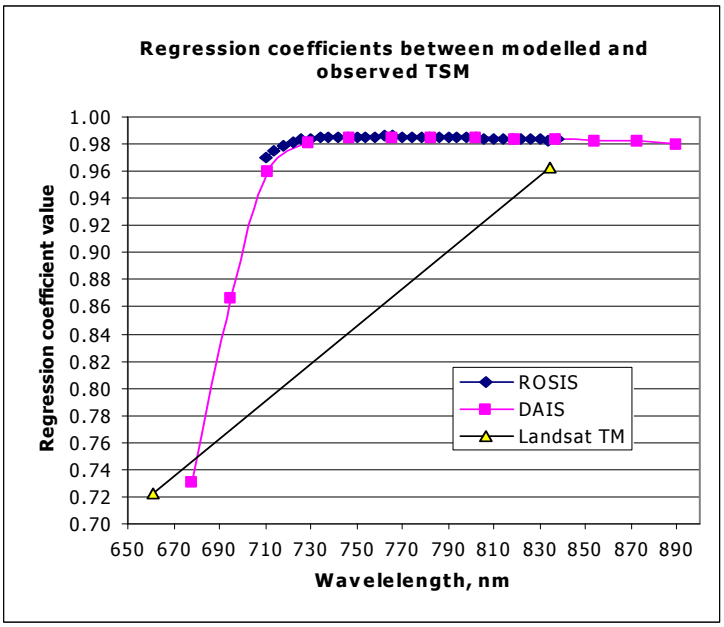


Figure 4.5 Regression coefficients between the modelled and observed TSM for different sensors

A formula, similar to the Equation 4.2, was used by Gege (2001) to model TSM concentrations in Lake Constance. Difference is only in application of the specific backscattering of CHL - he assumed that in the NIR region backscattering is chlorophyll independent (b_{b_chl} = 0). Assuming that, he obtained the following formula:

$$b_{b_tsm} = \frac{f * b_w * R(0-) - R(0-) * (a_w + b_w * B_w)}{R(0-) - f} \quad \textit{Equation 4.4}$$

The results obtained using Equation 4.4 are almost identical to the previously described ones, with slightly higher slope and intercept values obtained from the formula used by Gege (2001) and slightly lower regression coefficients (Figure 4.6). Also, the highest regression coefficients shift towards longer wavelength when the Gege's formula was used (832 nm, Figure 4.7).

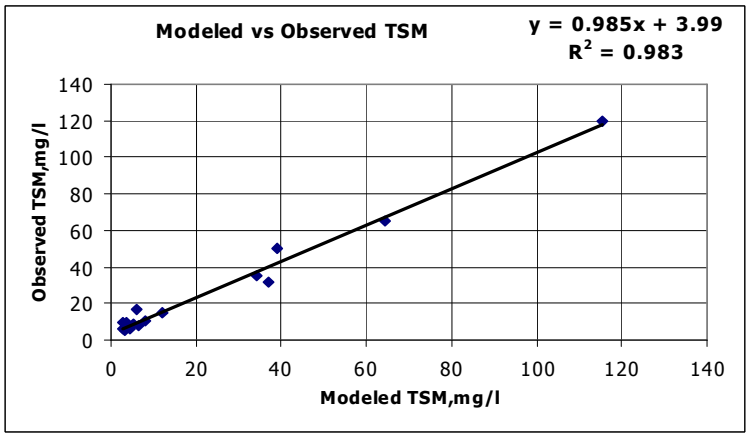


Figure 4.6 Regression graph between the observed and modelled at 832 nm using Gege's formula TSM concentrations (for the DAIS sensor reflectance)

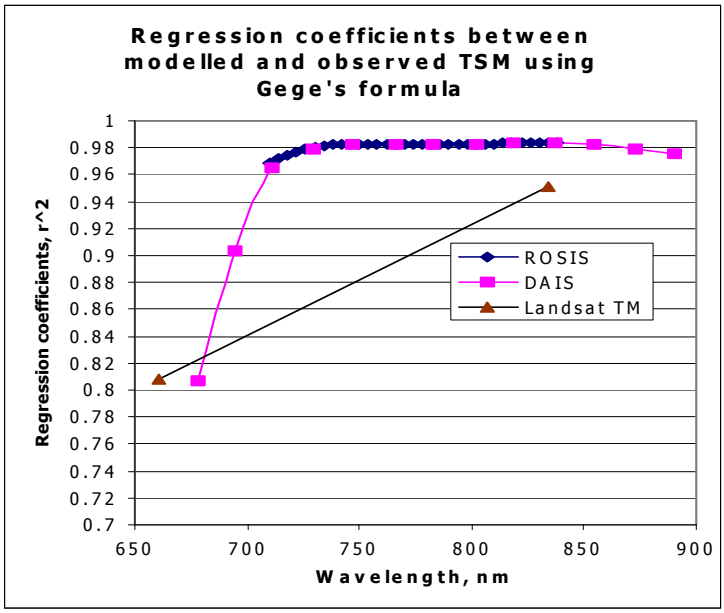


Figure 4.7 Regression coefficients between the observed and modelled TSM using the Gege's formula

Thus, the results indicate that *CHL backscattering does not affect very much the TSM estimate in the NIR region*. Therefore, it is possible to model TSM concentrations in the absence of CHL concentration values and chlorophyll backscattering coefficients. Although, the introduction of CHL backscattering improved results of TSM retrieval for the lakes in the study area. Yet taking into consideration the advantages of Gege's formula *we propose it for the application to the hyperspectral images*.

Besides other parameters, the ratio of large to small particle's concentration is very crucial for the modelling of TSM with the given data set (during the flood retreat). Although it seems to be intuitively right to assume that the amount of large particles is higher during the high flow and lower during the low flow, the exact value still has to be determined by modelling river's transport capacity using different flow velocities.

Also, for the modelling purpose, mean backscattering coefficients from Dutch lakes and backscattering by large TSM particles from Lake Constance were used. In order to improve the model, backscattering coefficients could be later determined by direct measurements in the field.

4.1.5. Justification of ratio of large to small particles concentration by calculating sediment transport capacity of Sajó River using Yalin equation.

The purpose of a sediment transport capacity calculation is to separate the amount of sediment having particle size less than 0.01 and more than 0.01 mm. The calculation was limited to the 17th August since we had only one observation of the water level of the river Sajó.

A spreadsheet for maximum sediment concentration calculation using Yalin equation is given in Appendix C. Below, only a brief introduction to the theory is described. For the details please refer to Yalin (1963) or to Mannaerts (2002).

Several sediment transport capacity equations exist to describe sediment load and concentration as a function of hydraulic channel and sediment parameters. Alonso et al. (1981) evaluated nine sediment transport equations and concluded that the Yalin equation provided reliable estimates for transport capacity for shallow overland flow and streamflow. Also, the Yalin equation included, for example, in the KINEROS model (USDA, 1990). The total load formula of Yalin is given as:

$$T_{sc} = S_g * g * \rho_w * d * u * 0.635 * (\frac{\tau_*}{\tau} - 1) * [1 - \frac{\ln(1+\sigma)}{\sigma}],$$

where, $a = 2.45 * S_g^{-0.4} * \sqrt{\tau_{cr}}$; $\delta = [\frac{\tau_*}{\tau_{cr}} - 1]$; $\sigma = a * \delta$

Equation 4.5

And finally sediment concentration in the flow can be written as:

$$C = \frac{T_{sc}}{10^{-6} * \rho_w * u * h}$$

Equation 4.6

where,

C, sediment concentrations by dry mass [ppm or mg/l];
ρ_w , specific weight of water [kg/m3];
Sg, specific gravity of sediment [dimensionless];
u, depth averaged velocity of water flow [m/s];
h, mean flow depth [m];
T_{sc}, sediment transport rate in per unit width per unit time [kg/(m*s)];
d, particles diameter [mm];
τ*, dimensionless shear;
τ, shear stress of flow exerted per unit area [N/m2]
τ_{cr}, critical (dimensionless) shear read from Shield's curve

Input parameters to the equation, as well as sources are presented in Table 4.5

Input parameter	*Source*	*Value*
Flow rate Q, m3/s on 17th Aug	VITUKI (Water Research Centre, Hungary)	78
Manning's coefficient	Table value	0.022
Friction slope S, m/m	From the topographic map	0.001
Particle diameter d, mm		0.075 (for large) and 0.007 (for small particles)
Particle specific gravity, S_g	Density of sediment/density of water	2.65
Channel width W, m	VITUKI	38
Density of sediment, kg/m3	Table value	2650
Density of water, kg/m3	Table value	1000
Temperature of water, o C	From the DAIS image (band80)	14

Table 4.3 Input parameters to the Yallin equation

Knowing that Yalin formula shows strange behaviour if applied to the complete particle size range (Mannaerts, 2002), only two particle groups were considered:

- Large particles (>0.01 mm) with an average diameter of 0.075 mm.
- Small particles (<0.01 mm) with an average diameter of 0.007 mm.

Average diameters for these two groups were derived from the joint research of RIZA and VITUKI (RIZA and VITUKI, 1994) and after considering a test done by Yun Liu (2003).

The results of applying the Yalin equation are shown in Table 4.4. The results indicate that calculated flow velocity is quite realistic for the high flow (1.59 m/s, see Appendix C): average flow velocity of Sajó river derived from a topographic map is 0.8 m/s. Accordingly, the ratio of large to small particles concentration assumed in bio-optical modelling is very close to the one which was calculated using Yalin equation (Table 4.4).

Maximum sediment concentration of *large* particles, mg/l	1703.99
Maximum sediment concentration of *small* particles, mg/l	332.98
Ratio of large to small particles concentration	5.12
Compare: *ratio used in bio-optical modelling of TSM on 17/08/2002*	4.79

Table 4.4 Result of sediment transport capacity calculation in comparison with the assumption of the bio-optical modelling of TSM

4.2. Modelling Chlorophyll-a using a semi-analytical solution of the band ratio NIR/Red

The purpose of this part of the research is to make an attempt to improve the regression equation found in the statistical analysis between band ratio NIR/Red and CHL by modelling the ratio.

4.2.1. Algorithm description

The algorithm presented here was first applied by Gons (1999) to model CHL concentrations in lakes in China, in the Netherlands and in Belgium. Applying Equation 2.1 to the band ratio 704/672 with the following assumptions:

- Backscattering is wavelength and CHL independent;
- At 672 nm: absorption of other water constituents than CHL and water is negligible;
- At 704 nm: absorption of other water constituents than water is insignificant in comparison to water absorption,

Gons reduced Equation 2.1 to the following ratio:

$$BR(\frac{704nm}{672nm}) = \frac{a_w^{(672)} + a_{chl}^{*(672)} * CHL + b_b}{a_w^{(704)} + b_b}$$

Equation 4.7

and rearranging Equation 4.7 to:

$$CHL = \frac{BR * (a_w^{(704)} + b_b) - a_w^{(672)} - b_b)}{a_{chl}^{*(672)}}$$

Equation 4.8

where,

CHL is the concentration of Chlorophyll-a in µg/l;
BR- the value of the band ratio;
$a_w^{(704)}$ is absorption of water at 704 nm;
b_b is wavelength and CHL independent total backscattering;
$a_{chl}^{*(672)}$ is specific Chlorophyll-a absorption at 672 nm.

The algorithm was applied to air-water corrected spectra. To model the band ratios, only those having the highest regression coefficients found in the statistical analysis part were used. Thus, for the ROSIS sensor it was a ratio of 706/670 nm and for the DAIS sensor it was 693/675 nm. Values for the absorption of water at the rationed wavelengths are taken from BIOPTI 1.0 (Erin Hogenboom,1995). As suggested by Gons (1999), the total backscattering coefficient is assumed to be CHL independent and equal to wavelength independent backscattering of TSM (b_{b_tsm}) as previously defined in the TSM modelling part of the research.

4.2.2. Result and discussion on bio-optical modelling of CHL

In general, the results (Figure 4.8 and 4.9) indicate that the regression coefficients of the modelled band ratios are less than that from the statistical analysis (compare to the Table 3.4) and generally the modelled band ratio underestimates the CHL concentrations. Presence of an outlier greatly affects regression coefficients: by removing it from Figure 4.8 for example, regression coefficient reduces to 0.87.

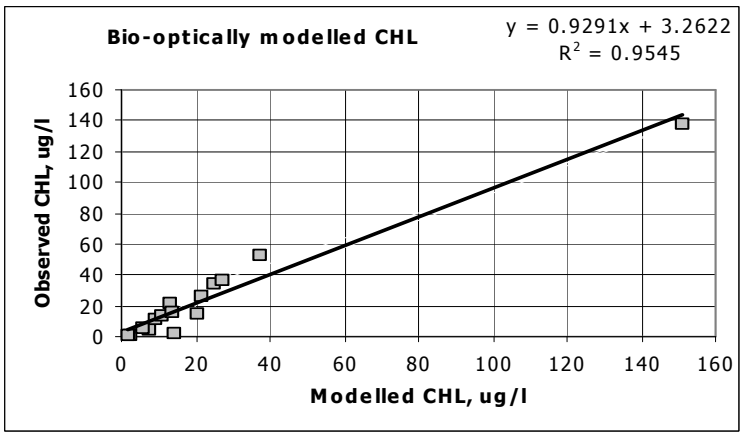


Figure 4.8 Bio-optically modelled CHL vs. observed CHL (for the ROSIS sensor)

The point distribution in Figure 4.9 shows that the algorithm to estimate CHL is generally not applicable to Landsat TM, because of its coarse spectral resolution. This led to the conclusion that *both*

band ratio algorithms (the statistical approach and the bio-optical modelling) are very sensitive to the precise detection of the CHL absorption peak at 675 nm and the following reflectance peak at 705 nm (Figure 4.10). Example of a space-borne sensor, which can trace those peaks, is MERIS (Medium Resolution Imaging Spectrometer) on board of the Envisat Earth Observation Satellite. Considering its coarse spatial resolution (300 meter), the satellite could be used only for the CHL monitoring in the lakes in the study area. Figure 4.11 gives the results of empirical algorithm for band ratio of the MERIS sensor.

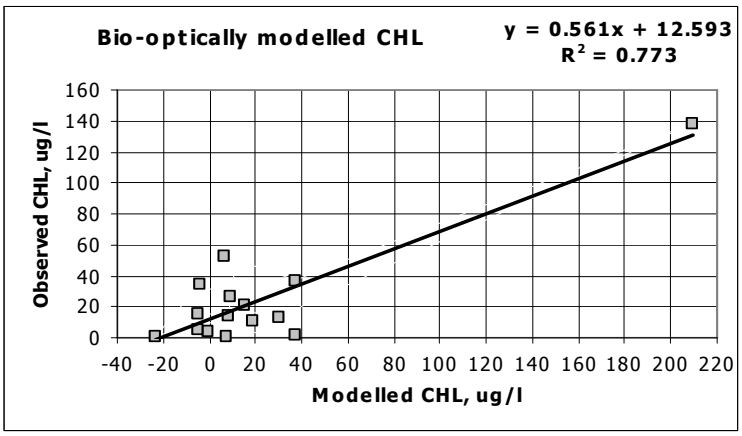


Figure 4.9 Bio-optically modelled CHL vs. observed CHL for the Landsat TM sensor

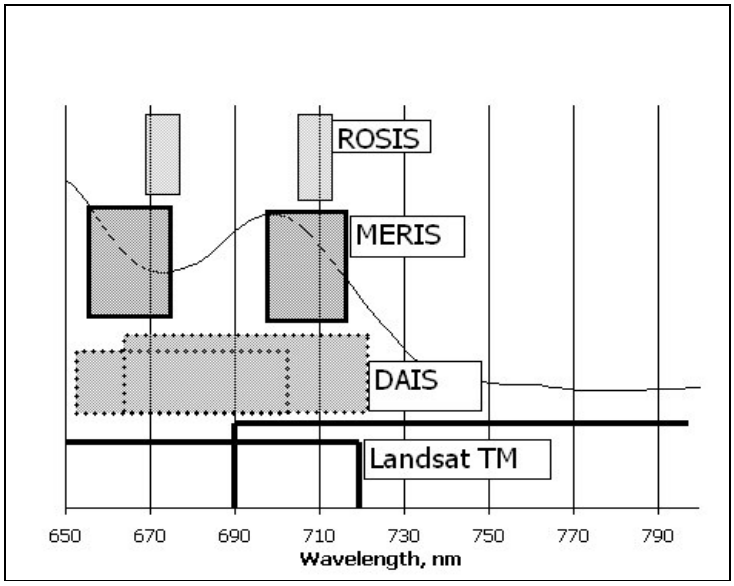


Figure 4.10 Sensors sensitivity to the region 650-710 nm with the spectrum of the sampling point L1_2008

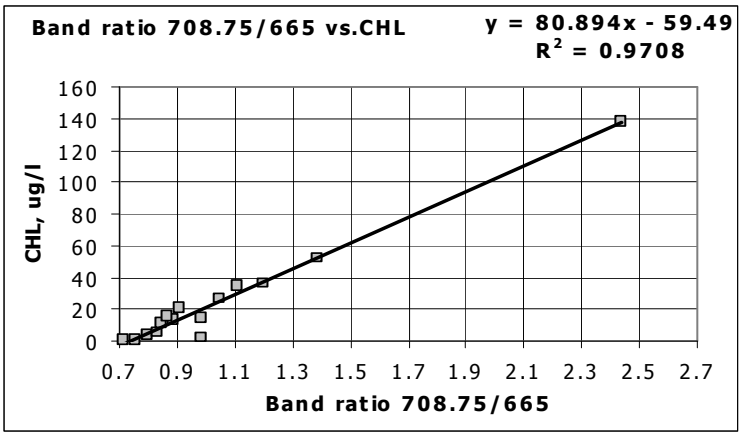


Figure 4.11 Band ratio vs CHL for the MERIS sensor

De Haan et al.(1999) analysed a strategy for algorithm development and showed the link between modelling NIR/Red ratio algorithm and statistical solution for Chlorophyll-a determination. In brief, the regression equation found in the statistical analysis between CHL and band ratio looks like:

$$CHL = BR * a - b$$

Equation 4.9

where, *a* and *b* are the regression coefficients. At the same time, Equation 4.7could be rearranged in the way to represent Equation 4.9 as:

$$CHL = BR * \frac{(a_w^{(706)} + b_b)}{a_{chl}^{*(670)}} - \frac{a_w^{(670)} - b_b)}{a_{chl}^{*(670)}}$$

Equation 4.10

where,

regression coefficient $a = \frac{(a_w^{(706)} + b_b)}{a_{chl}^{*(670)}}$ and regression coefficient $b = -\frac{a_w^{(670)} - b_b)}{a_{chl}^{*(670)}}$

Though Equation 4.10 is more flexible (it has a variable b_b which is changing from one sampling point to the another), in our case, the empirical approach between the band ratio and CHL gave slightly higher regression coefficients than the bio-optical modelling of the ratio using the Gons's formula (r^2= 0.977 and 0.955 for the ROSIS sensor respectively). This fact could be explained by assumptions made in the Equation 4.7, namely that *not all the WQ parameters contributing to the band ratio value have been included in the* Equation 4.7, particularly in regards to the absorption of CDOM and the backscattering of CHL.

Table 4.7 summarizes the regression equations obtained by statistical analysis and bio-optical modelling. It shows that results obtained from the statistical analysis are closer to the observed values than those from the bio-optical modelling of the band ratio.

	Regression coefficient, r^2	*Slope*	*Intercept*	*St. error*
Statistical analysis:				
ROSIS	0.98	1	0	5.47
DAIS	0.97	1	0.01	6.5
Landsat TM	0.73	2.16	-3.66	18.64
Bio-optical modelling:				
ROSIS	0.95	0.92	3.26	7.68
DAIS	0.96	2.39	-1.80	7.11
Landsat TM	0.77	0.56	12.59	17.15

Table 4.5 Comparison of the regression coefficients and equations between the observed and modelled CHL using 2 approaches

Unfortunately, we could not improve the semi-analytical solution of band ratio algorithm by introducing more WQ parameters due to the absence of data on the amount of CDOM and specific absorption properties of CDOM of the studied water bodies.

Nevertheless, the overall results indicate that the Gons formula could be used for the monitoring of CHL since it is an essentially simplified "physical" model. Gons (1999) proposed to use of his algorithm for the waters with higher CHL concentrations than we observed in the study area (from 1 to 215 mg/l). In addition, the appliance of the band ratio makes the model to be insensitive to the proportionality factor, *f,* and consequently to the illumination conditions. All above makes the algorithm independent from *in situ* calibration data, which is always needed for the statistical approach. In this research *we propose both methods: the bio-optical modelling and the empirical approach of the band ratio for a visual comparison.*

4.3. Error analysis and sensitivity of the models

The bio-optical modelling has a good potential to become a fully operational method for remote sensing of water quality. The aim of the operation is to provide the users with reliable and accurate information on water quality derived from remote sensing data. To achieve this aim, an error analysis is required (Ambarwulan, 2002).

Main sources of errors in the models are due to:

- The simplifications made in the models;
- Model parameters which are assumed to be constants such as: the proportionality factor, *f*, the backscattering probabilities;
- Errors in sampling, in taking spectral measurements and laboratory errors;
- But mainly: due to the accuracy of retrieving the subsurface volume reflectance *R (0-)* and the SIOPs

This part of the thesis summarises the standard errors obtained as a result of regression analysis between the modelled and observed concentrations of WQ parameters for the proposed algorithms.

Table 4.6.indicates that the errors are generally not large, possibly due to the low number of measurements (N=15).

Water quality parameter	*Sensor*	*Standard error of the statistical approach*	*Standard error of the bio-optical modelling*
CHL (µg/l)	ROSIS	5.47	7.68
	DAIS	6.5	7.11
	Landsat TM	Not proposed	Not proposed
TSM (mg/l)	ROSIS	Not proposed	4.2
	DAIS	Not proposed	4.2
	Landsat TM	Not proposed	7.2
Turbidity (NTU)	ROSIS	4.4	N/A
	DAIS	4.1	N/A
	Landsat TM	6.2	N/A

Table 4.6 Summary of the standard errors (N=15) between the modelled and observed WQ parameters for the proposed algorithms

Generally speaking, the sensitivity of the statistical algorithms depends only on the remote sensing reflectance. The empirical as well as bio-optical algorithm of CHL determination using band ratio indicated sensitivity of the algorithms to the precise detection of the CHL absorption peak at 675 and reflectance peak at 705 nm, which are generally narrow (about 40 nm in total, Figure 2.3 and Figure 4.10).

More complicated is the sensitivity of the bio-optical model used for determination of TSM concentrations (Table 4.7 and 4.8): the model is sensitive to almost all the parameters, except for the backscattering by large TSM particles. But most of all it is very susceptible to the decrease of the proportionality factor, f, to the increase of the TSM backscattering probability, B_{tsm} (in case of the river) and to the increase of backscattering by small particles in case of lakes. Point L3_2008 illustrated low sensitivity to the 50 % change of subsurface reflectance due to the low value of the reflectance from the lake (Table 4.8).

Input parameters	+50%change in input parameter results output to change:	-50%change in input parameter results output to change:
Proportionality factor, f	-40%	+207%
Ratio of large to small particles concentration, r	+28%	-48%
Specific scattering by small particles, $b^*_{tsm_small}$	-29%	+68%
Specific scattering by large particles, $b^*_{tsm_large}$	-9%	+10%
Subsurface reflectance, $R(0-)$	+68%	-57%
Backscattering probability of TSM, B_{tsm}	+103%	-32

Table 4.7 Sensitivity of the bio-optical model for the sampling point R2_2008 (example of the ROSIS sensor)

Input parameters	+50%change in input parameter results output to change:	-50%change in input parameter results output to change:
Proportionality factor, *f*	+87%	-71%
Specific scattering by small particles, $b^*_{tsm_small}$	-33%	+100%
Subsurface reflectance, *R(0-)*	-34%	+21%
Backscattering probability of TSM, B_{tsm}	+29%	-13%

Table 4.8 Sensitivity of the bio-optical model for the sampling point L3_2008 (example of the ROSIS sensor)

Though some models suggest to use field measurements to calculate f and B_{tsm} (Dekker et al., 2001), normally these parameters are alternatively defined by an optimisation (matching of simulated *R(0-)* using local SIOPs to measured *R(0-)* in the same location). In this research they were assumed to be constant, though B_{tsm} , the backscattering probability, obviously should be different for 2 dynamically different water systems such as lakes and rivers.

In addition, a direct measurement of the SIOPs in the study area will confidently increase the accuracy of the bio-optical-modelling for both TSM and CHL. Furthermore, a study of a relation between the flow velocity and particle size and subsequently backscattering properties of the particles will greatly contribute to the understanding of the river's reflectance properties and later the bio-optical modelling of TSM in the rivers.

Chapter 5. Application of the developed algorithms to the hyperspectral images

5.1. Comments on using remote sensing data in studying the water bodies.

One of the principal differences between image processing of land surfaces and that of water is the much smaller range of variability of signal from water bodies. A sensor such as the Coastal Zone Colour Scanner, which is specifically designed for the sea use, copes with the small variability of subsurface reflection of solar light by having a high sensitivity within the expected range of water-viewing radiances. As a consequence, it often saturates over land. When using other sensors such as Landsat, NOAA, AVHRR or Meteosat, it often occurs that the variability that is being studied is not much larger that the sensor resolution, or the digitisation interval (Robinson, 1994).

In case of the observed system noise, a compromise could be achieved by applying smoothing filters. The size of the filter, as well as the type, has to be chosen carefully in order to preserve spatial variability and eliminate sensor's noise.

Thus, for the proper application of the WQ algorithms, images and sensors have to be carefully examined. In our case, the radiometric characteristic of the ROSIS sensor was especially suitable for the water applications (see Chapter 2), whereas the DAIS sensor was mainly designed for land monitoring.

Flights took place on 17th and 18th August. Weather conditions on 18th August were more appropriate for the water quality monitoring (absence of clouds and wind). Regrettably, due to an internal error in the performance of the ROSIS sensor, only the atmospheric correction (apart from the geometric correction) of the ROSIS image from 17th August was performed. For the same reason, the ROSIS sensor failed to acquire an image on 18th August. Regarding the DAIS sensor, it performed well in both days and for the analysis we used the atmospherically and geometrically corrected DAIS image from 18th August.

5.2. Examination of the images

Atmospheric corrections of the images were done by DLR using the ATCOR4 software. Comparison of ROSIS image's spectra and field spectra on the sampling points L3_1808 and R2-3_1808 exposed some differences (Figure 5.1). Possible source of differences of the ROSIS image are:

- In shorter wavelengths (< 500 nm): the radiometric resolution of the sensor and low reflectance from the water bodies resulted in "noisy" character of the spectra in the blue region (the signal-

to-noise ratio of ROSIS data is low for channels below 500 nm, Rudolf Richter, personal communication, 2003);

- In the longer wavelengths, it is a combined effect of both: the sensor properties and the applied atmospheric correction algorithm, which was calibrated for land targets. This later resulted in higher reflectance values in NIR region and few artificial reflectance peaks;
- Difference in dates of measurements (the image was acquired on 17th August, whereas the field spectra were collected on 18th);

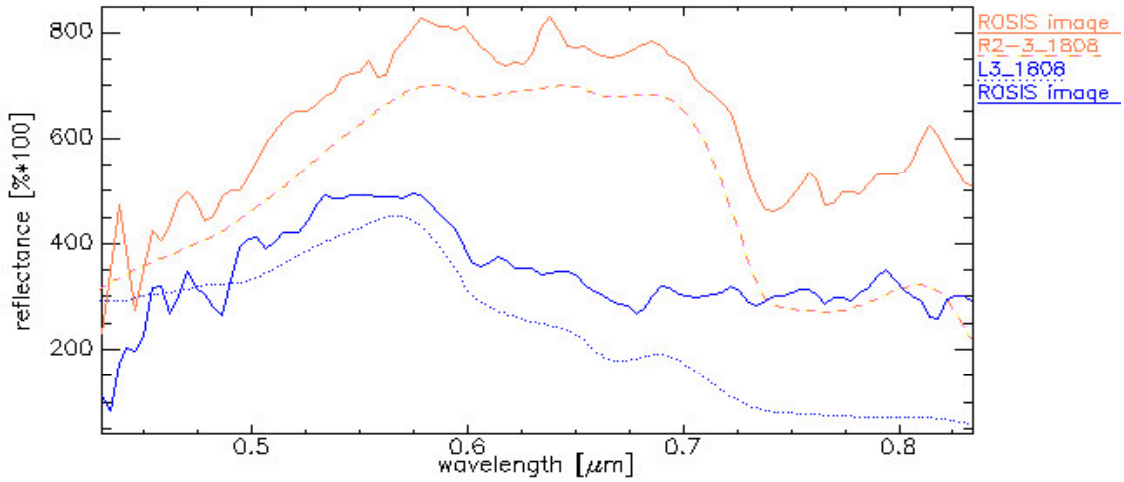


Figure 5.1The ROSIS image's spectra from 17th Aug (solid lines) in comparison with the field spectra from 18th Aug (dashed lines)

But even larger differences, especially in the NIR region, were found between the DAIS image and the field spectra (Figure 5.2). As the image and the field spectra are from the same date, possible sources of error are mainly the sensor sensitivity and the applied atmospheric corrections, which were calibrated to meet requirements for the land monitoring.

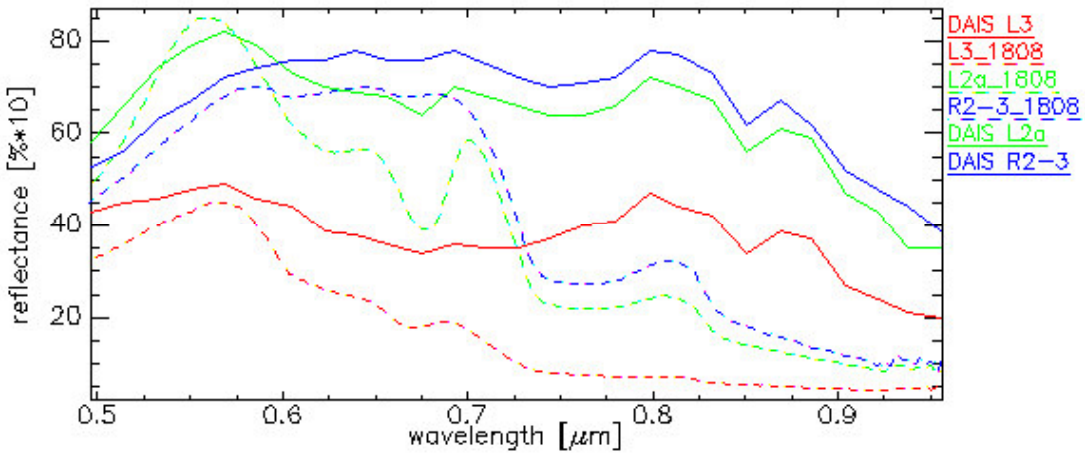


Figure 5.2 The DAIS image's spectra (solid lines) in comparison with the field spectra (dashed lines)

Thus, to match the DAIS image spectra to the observed field spectra, an empirical line calibration algorithm was applied to the masked (see in Section 5.3) image using ENVI. Basically this is a linear transformation; the software calculates a gain and an offset for the image spectra based on field spec-

tra and applies them to bands of the image. Three field spectra were used for the transformation, and the results are shown in Figure 5.3.

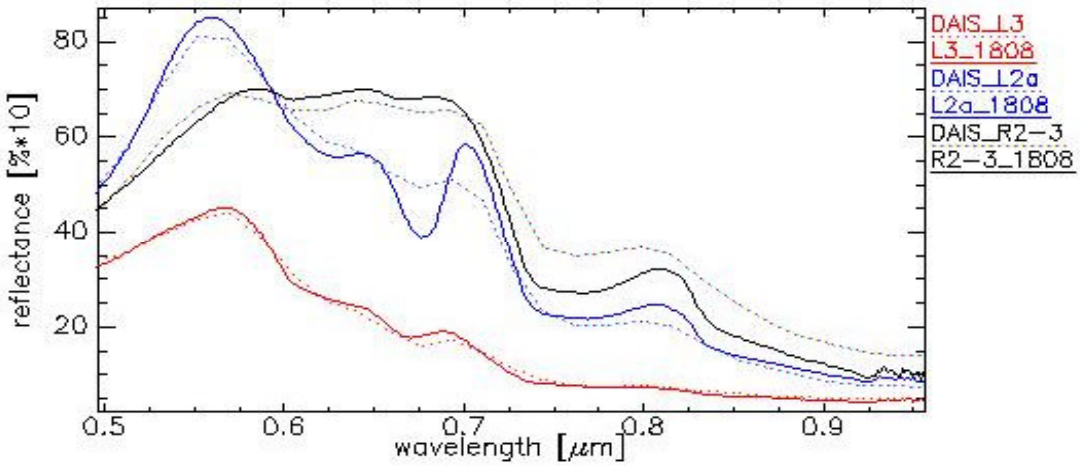


Figure 5.3 Results of applying empirical line correction algorithm (dotted lines are the DAIS image spectra, solid –field spectra).

5.3. Image processing

In contrast to the ROSIS image, the DAIS image had system noise in the form of stripes (Figure 5.4). Several filters were applied in order to remove stripes and to preserve spatial variability at the same time. In our opinion, the best result was achieved by applying a smoothing (low pass convolution) filter with a kernel size of [5*5]. Figure 5.4 shows results of applying two filters on the band 20 (832 nm) of the DAIS image

The same filter was also applied to the ROSIS image with an aim to reduce the spectral noise (e.g. the fake variances in Figure 5.1). The resulting spectra looked smoother though few artificial peaks remained. Among them is a reflectance peak at O_2 absorption band (760 nm) that made the application of the first derivative algorithm for the turbidity determination impossible.

Masking the water bodies was done in ENVI using image slicing of band 24 (904 nm) of the DAIS image. Satisfactory results for masking the water bodies were achieved using a band range from 1(to eliminate area with no image) to 190 [units: reflectance%*10].

The description of the algorithms and procedure for estimating water quality parameters were introduced in the Chapters 3 and 4. The results of applying those steps to the images are presented in Appendix D. Table 5.1 reviews algorithms applied to the images. The ROSIS image was not used for the bio-optical modelling, since to obtain the subsurface volume reflectance we required the reflectance at 930 nm (the ROSIS image had a spectral range up to 838 nm only and was not geometrically corrected for incorporating it with the DAIS's reflectance at 938 nm, see Chapter 2). For the TSM determination only the Gege's formula was applied since it is relatively insensitive to the amount of CHL.

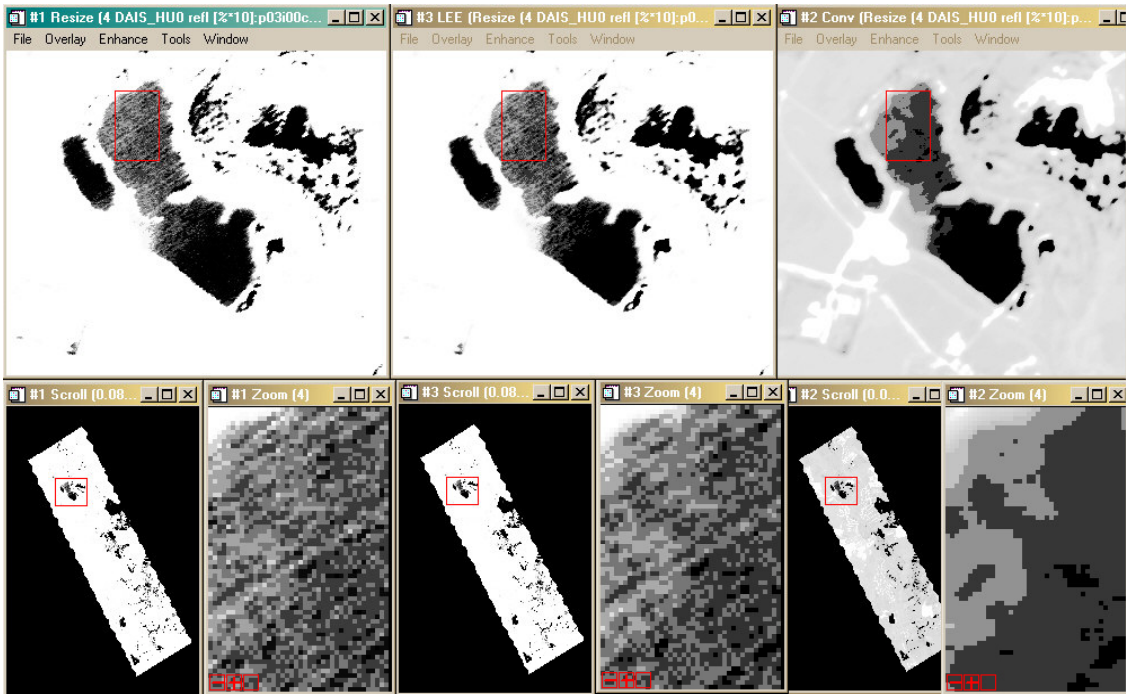


*Figure 5.4 Raw image (utmost left), image with applied Lee smoothing filter, kernel size [3*3] (in the middle) and image with the applied smoothing (convolution low pass) filter, kernel size [5*5].*

Table 5.1 Description of the algorithms applied to the hyperspectral images

Water quality parameter	*Sensor*	*Statistical approach based on*	*Bio-optical modelling based on:*
CHL (µg/l)	ROSIS	Band ratio 706/670 nm	N/An
	DAIS	Band ratio 693/675 nm	The Gons equation for 693/675 nm
TSM (mg/l)	ROSIS	Not proposed	N/A
	DAIS	Not proposed	The Gege formula for 832 nm
Turbidity (NTU)	ROSIS	Reflectance at 698 nm	N/A
	DAIS	Reflectance at 695 nm and first derivative at 792.5 nm	N/A

5.4. Analysis of the resulted maps

The water quality parameter maps of the Sajó floodplain are presented in Appendix D. All the discussions in this section below refer to these maps.

In comparing the maps of the ROSIS and the maps of the DAIS sensors, it has to be noted, that the data acquisition took place with one day difference. As it was mentioned before, in optically shallow

waters, the bottom reflectance has an effect on the mapped concentrations. This and the mixed-pixel effect is the reason of some higher constituent concentrations along the shores of the lakes.

5.4.1. CHL maps

Essentially all the three maps gave the same order of magnitude of the CHL distribution. The empirical approach applied to the DAIS image gave slightly higher CHL values for the small lakes and the river Sajó than two other algorithms. Only a little variation exists between the river and the lakes: this is the result of the previous flood, which washed through the lakes on the floodplain either directly, or via the groundwater.

5.4.2. Map of TSM

It shows that the river Sajó and Hernád had higher values of TSM than the surrounding lakes. As the DAIS image was from 18th August, the TSM concentrations in the river are generally higher than those that were observed on 20th and later (see Table A.1, Appendix A). It consents with the higher flow velocity and the capacity to carry more sediment during the high flow on 18th August. Also the map shows higher TSM concentrations for the lake L1 and the small ponds next to L3, which agrees with the fact that there was a direct surface inflow from the river into these water bodies during the flood. The TSM distribution shows higher values at the shores of rivers and lakes possibly due to the bottom reflectance or/and due to the applied smoothing filter.

5.4.3. Turbidity maps

The first derivative algorithm yielded higher Turbidity values than empirical algorithms employing reflectance at 698 nm (for the ROSIS sensor) and 693 nm (for the DAIS). Reasons as they were mentioned already, are: high coefficients in the regression equation of the first derivative and low first derivative values. The ROSIS image did not contain sampling points L2 and L2a, although point L2 is more or less similar to the point L3 and L2a is similar to the L3a (see Table A.1, Appendix A). The field turbidity measurements carried out during the flight on 17th August shows consistency with the empirical algorithm for the ROSIS sensor: sampling points R2 – 35.6 NTU, L2 -2.7 NTU and L2a-28.1 NTU.

5.4.4. ISM and OSM maps

The ISM map was produced according to the correlation described in the Chapter 3, whereas OSM was calculated by formula: *OSM=TSM-ISM*. Map shows that the ISM content of the Lake L3 is slightly lower than the OSM concentrations. Conversely, the TSM content of the river Sajó mainly consists of the inorganic particles.

5.5. Concluding remarks on the image processing

Assuming that all the parameters of the bio-optical modelling are properly defined and algorithms are well established, then the main source of errors in the modelling would be the retrieval of the subsurface volume reflectance, *R (0-)*, which greatly depends on the applied atmospheric correction. The

examination of the images showed the necessity of a precise correction for the atmosphere over the water bodies.

Random effects (variable spectra) could be seen in all the images, so it was necessary to apply filtering to remove this noise. As a function of the spatial resolution, this remains a main limitation in the utilization of the satellite data for most of the rivers and small lakes. For example, with a smoothing filter of [5*5] pixels to obtain an “undisturbed”(free from shore and bottom) reflectance will require at least a 15 pixels-wide river (450 meters for the Landsat sensor).

However, a solution could be found in developing a filter, which would take into account the signal distribution within the water bodies and at the same time would preserve pixels with waters less wide than the filter size.

Apart from the spatial resolution, we assumed that radiometric sensitivity of the sensors is the same as the spectrometer. In practice, the radiometric resolution of the Landsat TM for example, is much lower, especially in the NIR region.

Thus, although the research showed the possibility of utilizing the spectral reflectance of the Landsat TM sensor for the TSM estimations in the study area, it is practically non applicable for Sajó river.

Chapter 6. Conclusions and recommendations

The main objective of the study was the quantification of the WQ parameters using imaging spectrometry. Methods were developed and tested on the Sajó River floodplain, Hungary. The objective has been achieved by using a systematic approach through the establishment of the empirical and semi-analytical relations between the field spectrometry reflectance, resampled to the wavelengths of the ROSIS and DIAS sensors, and the water constituent concentrations. The models have been successfully applied to the hyperspectral images and WQ maps have been derived.

For the investigation of the possibilities of WQ monitoring using lower resolution satellite, the Landsat TM's spectral characteristics were found to be suitable to a limited extent for the monitoring of TSM, ISM, OSM and turbidity. For the monitoring CHL in the lakes in the study area the MERIS sensor was found to be more suitable.

A detailed view on the conclusions and recommendations is given in the following sections.

6.1. Conclusions and recommendations on applying the empirical approach

The empirical approach, particularly the proposed band ratio algorithms for the modelling CHL was found to be sufficiently accurate. The developed statistical relationships have been tested on both data sets (total N=44). The regression equation found in this research is similar to the ones found by other researches.

The empirical approach for the turbidity mapping was based only on 15 measurements and was found to be less accurate then the empirical estimates of CHL. However, the resulting maps matched fairly well with the field turbidity measurements. To formulate a more consistent algorithm we recommend making more measurements with preferably a wider range of turbidity values.

The statistical approach based on reflectance for the estimation of TSM yielded highest correlation at different spectral locations in each individual dataset. The algorithm is also affected by the number of measurements and by the origin of TSM present in the studied waters; therefore it is difficult to predict the exact wavelength where the highest correlation can occur using this approach.

Among the proposed empirical algorithms, the first derivative approach found to be less accurate in application to the images due to its usually low derivative values and high slope coefficients. Nevertheless, it would be interesting to see the effect of generalization the wavelength intervals on the performance of the first derivative method.

The statistical analysis based on the field spectroscopy and instantaneous water sampling will produce more consistent regression equations than the statistical approach based on the image spectra. Reasons are: time limitation for the water sampling in the time of the satellite (or aircraft) overpass and more importantly - a precision of the atmospheric correction.

6.2. Conclusions and recommendations on applying bio-optical modelling

Generally, the use of other SIOPs than the ones determined from the studied water bodies can be a source of errors. However, in the absence of direct data from the Hungarian sites, we used SIOPs from the Dutch lakes. The results of bio-optical modelling of TSM indicate similarity between the SIOPs of the lakes in Holland and the lakes in the study area.

The water of the Sajó River, due to its high TSM values with a wide range of grain sizes, required a different approach – the separation between the large and small suspended particles.

Although it is intuitively right to assume that the river's TSM backscattering depends on the particle size (which is a function of the flow velocity) and we proved it by the modelling the river's transport capacity; essentially our assumption was based on only 5 measurements in the rivers. It can be concluded that direct measurements of SIOPs in the study area will increase the accuracy of the bio-optical-modelling for both TSM and CHL. Furthermore, a study of the relation between the flow velocity and particle size and subsequently the backscattering properties of the particles will greatly contribute to the understanding of the river's reflectance properties and later to the better bio-optical modelling of TSM and CHL in the rivers.

6.3. Conclusions and recommendations on applying atmospheric and air-water corrections

Once a bio-optical algorithm is established, errors in retrieving WQ constituents will mainly depend on the atmospheric corrections applied to the remote sensing data and on the air-water interface correction algorithm. While the air-water interface correction applied here gave generally acceptable results (proved by the bio-optical modelling), it has a disadvantage that the water-leaving radiance had to be normalized. For that purpose we used a subtraction of the reflectance at 930 nm from all bands of the spectrum. This NIR sensor wavelength is often not available on the satellites. A use of other air-water correction algorithms and a relative comparison with the applied one would raise the overall accuracy of the bio-optical modelling.

As it was shown in our case, most of the lands monitoring focused atmospheric corrections are not suitable for the water targets. In that sense, reference spectra directly measured on water bodies and the application of special atmospheric correction methods, developed for water bodies provide only good results.

6.4. Conclusions on the processing remote sensing data

Remote sensing data can provide useful information for water quality monitoring. However, depending on the level of WQ parameters quantification, the sensor selection procedure should take into consideration the radiometric sensitivity of the sensors. Later, each remote sensing image has to be carefully examined in respect to the signal received by the sensor from the water body. In the presence of noise, filter(s) has to be applied to the image. However, filtering will yield better results if applied to the water bodies only, eliminating signal from the shore. For the future research, we recommend a development of a selective filter that would take into account only the signal distribution within the water bodies and at the same time it would preserve pixels with waters smaller than the filter size.

References

Alonso, C.V., Neibling, W.H. and Foster, G.R. (1981) *Estimating sediment transport capacity in watershed modelling* Transactions of the American Society of Agricultural Engineers 24:1211-1220,1226.

Ambarwulan, W. (2002) *Mapping of TSM concentrations from SPOT and Landsat TM Satellite Images for Inegrated Coastal Zone Management in Teluk Banten, Indonesia*. MSc Thesis, ITC, The Netherlands, 130p.

Babin, M. and D. Stramski (2002). "*Light absorption by aquatic particles in the near-infrared spectral region.*" Limnology and Oceanography 47(3): 911-915.

Bukata, R. P., Jerome, J. H., Kondratyev, K.Y. and Pozdnyakov, D.V. (1995). *Optical properties and Remote Sensing of Inland and Coastal Waters*, CRC Press.

De Haan, J. F., Kokke J. M. M., Dekker, A. G. and Rijkeboer, M.. (1999). *Remote Sensing Algorithm Development: toolkit for water quality continued*. Delft, Report of the Netherlands Remote Sensing Board (BCRS).

Dekker, A. G. (1993). *Detection of optical water quality parameters for euthrophic waters by high resolution remote sensing*, PhD thesis, Free University Amsterdam.

Dekker, A. G., Brando V. E., Anstee, J.M. and Pinnel, N. (2001). *Chapter 11. Imaging spectrometry of water*. In *Imaging spectrometry*, edited by F. D. Van Der Meer and S.M. De Jong , KLUWER publishers.

Doerffer, R. and J. Fischer (1994). "*Concentration of chlorophyll, suspended matter, gelbstoff in case 2 waters derived from SCZS with inverse modelling methods.*" Journal of Geophysical Research. 99(C4): 7457-7466.

Doxaran, D., . Froidefond, J.M., Lavender, S. and Castaing, P. (2002). "*Spectral signature of highly turbid waters. Application with SPOT data to quantify suspended particulate matter concentrations*." Remote Sensing of Environment 81: 149-161.

Fraser, R. N. (1998). "*Hyperspectral remote sensing of turbidity and Chlorophyll-a among Nebraska Sand Hills lakes.*" International journal of remote sensing 19(8): 1579 -1589.

Gege, P. (2001). *A software tool for simulation and analysis of optical in-situ spectra*. 4th Berlin Workshop on Ocean remote sensing, Berlin, Germany.

Gons, H. J. (1999). "*Optical teledetection of Chlorophyll-a in Turbid Inlands Waters.*" Environmental Science and Technology 33: 1127-1132.

Gordon, H. R., Brown, O. B. and Jacobs, M.M. (1975). "*Computed relationship between the inherent and apparent optical properties of a flat homogeneous ocean*." Applied Optics 14(2): 417-427.

Han, L. and D. C. Rundquist (1997). "*Comparison of NIR/Red Ratio and First Derivative of Reflectance in estimating Algal-Chlorophyll Concentration: A Case Study in a Turbid Reservoir*." Remote Sensing of Environment 62: 253-261.

Han, L. and D. C. Rundquist (1997). "*The impact of a wind-roughened water surface on remote measurements of turbidity*." Remote Sensing of Environment 19(1): 195-201.

Hogenboom, H. J. and A. G. Dekker (1999). Report: *InveRSion: assessment of water composition from spectral reflectance. A feasibility study to the use of matrix inversion method*. Delft, Report of the Netherlands Remote Sensing Board (BCRS) 98-15.

Jaquet , J. M., Strauffacher, M., Lehmann,A. and Nakayama M. (1994). *Water quality assessment and management by GIS and remote sensing: The GISWAQ project,* Report of a WMO Regional Workshop, Vienna, Austria, Tech. Rep. In Hydrology and Water Resources 42: 249-268

Koponen, S., Pulliainen, J., Kallio, K. and Hallikainen, M. (2002). "*Lake water quality classification with airborne hyperspectral spectrometer and simulated MERIS data.*" Remote Sensing of Environment 79(51-59).

Krijgsman, J. (1994). *Optical propersties of Water Quality parameters. Interpretation of Reflectance Spectra*, PhD thesis, Delft University of Technology: 200p.

Lahet, F., Ouillon, S. and Forget, P. (2001). "*Colour classification of coastal waters of the Ebro river plume from spectral reflectance*." International journal of remote sensing vol. 22(no. 9): 1639–1664.

Lee, Z. P., Carder, K. L., Mobley, C. D., Steward, R. G. and Patch, J. S. (1998). "*Hyperspectral remote sensing for shallow waters. I. A semi-analytical model*." Applied Optics 37: 6329-6338.

Maritorena ,S., Morel, A. and Gendtili, B.(1994) *Diffuse reflectance of of oceanic shallow waters: influence of water depth and bottom albedo*, Limnology and Oceanography 39: 1689-1703

Mannaerts, C.(2002) *Lecture notes on Module 9 WREM2.ESAM:Modelling water erosion, sediment and chemical transport*, unpublished

Morel, A. and B. Gentili (1993). "*Diffuse reflectance of oceanic waters: II. Bidirectional aspects*" Applied Optics 32: 6864-6879.

Ouillon, S., P. Forget, Froidefond, J.M. and Naudin, J.J.. (1997). "*Estimating suspended matter concentrations from SPOT data and from field measurements in the Rhone river plume.*" Marine Technology Society Journal 31(2): 15-20.

RIZA and VITUKI (1994). Report: *Ecological Rehabilitation of Floodplains of Rivers*, Institute for Water Pollution Control, the Netherlands, Water Research Centre, Hungary.

Robinson, I. S. (1994). *Satellite Oceanography. An introduction of Oceanographers and Remote Sensing scientists*, Praxis publishing Ltd.

Schaale, M., Olbert, C. and Fischer, J.. (1999). *Routine water quality monitoring of all Berlin lakes*. Proceedings of the Fourth International Airborne Remote Sensing Conference and Exhibition / 21st Canadian Symposium on Remote Sensing,, Ottawa, Ontario, Canada.

Shafique, N. A., B. C. Autrey, Fulk, F. and Cormier, S. M.. (2001). "*The Selection of Narrow Wavebands for Optimising Water Quality Monitoring on the Great Miami River, Ohio using Hyperspectral Remote Sensor Data.*" Journal of Spatial Hydrology 1(1).

Steegen, A.,Covers, G., Beuselink, L., Nactergaele, J., et al. (1998) *Variation in sediment yield from an agricultural drainage basin in central Belgium* Proceedings of the International Symposium on Modelling Soil Erosion, Sediment Transport and closely related Hydrological Processes, Vienna, Austria

Thiemann, S. and H. Kaufmann (2002). "*Lake water monitoring using hyperspectral airborne data-a semi-empirical multisensor and multitemporal approach for the Mecklenburg Lake District, Germany*." Remote Sensing of Environment 81: 228-237.

Tsai F. and Philpot W.(1998*) "Derivative Analysis of Hyperspectral Data."* Remote Sensing of Environment 66: 41-51

Van de Hulst, H. C. (1957). *Light scattering by small particles*. New York, Dover.

Yallin, Y. S. (1963). "*An expression for bed-load transportation.*" Journal of the Hydraulic Division, Proceedings of the American Society of Civil Engineers 89(HY3): 221-250.

Yun Liu (2003) *Possibilities of Assessing Heavy Metal Contamination in the Sajó River Flood Plains (Hungary) using Reflectance Spectroscopy* MSc Thesis, ITC, The Netherlands, 61p.

Appendix A. Standards and procedure for the laboratory analysis

This documents contains manual for laboratory measurements that were done at Miskolc Environmental Protection Agency laboratory, Hungary. It also gives a map with the location of the sampling points.

Concentration of TSM (mg/l)

The method is based on the ISO (International Standardization Organization). The samples were filtered on Whatman GF/F (25 mm ∅, 0.45 µm pore size) filtered, dried (105^0C) and weighted.

Equipment: Filtration device (25 mm∅); exsiccator (for storage of filters at dry place); oven (105^0C), microbalance.

Method:

- Filtration of samples using Whatman filter and 100 ml of sample. Add 100 ml of tap water on top of the filter to wash away any present salt;
- Put filter in a large petridish in the oven 105^0C for 2 hours;
- After 2 hours weight the filter on the microbalance

Concentration of ISM (mg/l)

Equipment: Filtration device (25 mm∅); exsiccator (for storage of filters at dry place); oven (550^0C), microbalance.

Method:

- The same as in the TSM, but after filtering put the filter in the ashing oven (550^0C) for 24 hours;
- Weight the filter on the microbalance

Concentration of OSM (mg/l)= TSM-ISM

Concentration of CHL (µg/l)

The method is based on the ISO 10260: 1992. The pigments are extracted using 90% ethanol at 78^0C.The concentration is determined spectrophotometrically (device –UNICAM UV-VIS) by measuring the extinction coefficients at 665 and 750 nm before and after acidification of the sample.

Equipment: Filtration device (0.45 µm pore size), water bath (with shake possibility and set at 78^0C), spectrophotometer UNICAM UV-VIS

Method:

- Filtering the same as in case of the TSM determination but with 500ml of sample;
- Transfer the folded filter into a large black-coated reagent tube. Add 50 ml ethanol 90% to the tube containing filter;
- Heat the tube for 8 min at 78^0C with shake possibility;
- Cool down to the room temperature;

- If needed, filtrate content of the reagent tube over 0.45 mm pore size;
- Put the filtered sample in the spectrophotometer UNICAM UV-VIS.

Turbidity determination

Turbidity measurements were done according to the manual of a portable turbidimeter HACH (Model 2100P).

Table A.1 Water sampling results and descriptive statistics

		TSM, mg/l	*ISM, mg/l*	*OSM, mg/l*	*Turb, NTU*	*CHL, µg/l*	*Depth, m*	*Secchi[5] disc depth, m*	*Time of measurement*	*Sun zenith angle,⁰*
1	*L1_2008*	11	8	3	8.87	26.88	5.2	1.1	11.08	54.72754
2	*L2_2008*	10	6	4	2.07	1.11	8.05	2.5	12.58	42.33215
3	*L2a_2008*	15	5	10	18.30	15.18	5.7	0.9	13.30	38.17602
4	*L3_2008*	6	3	3	1.89	0.89	5.7	2.2	14.50	35.34220
5	*L3a_2008*	17	6	11	5.77	34.87	5.15	0.9	15.28	36.82086
6	*R4_2008*	120	95	25	48.50	1.78	1.5	0.6	11.72	49.07362
7	*GR_2208*	9	6	3	5.19	4.66	8.4	1.4	15.90	39.66444
8	*Lbc_2308*	10	6	4	5.66	5.55	4.3	1.5	15.12	36.30181
9	*R2_2408*	35	30	5	13.60	21.77	0.9	0.7	11.90	47.49806
10	*R3-2_2408*	50	43	7	19.20	11.33	0.9	0.7	14.47	35.28796
11	*H1_2408*	8	5	3	5.16	52.42	1.1	0.7	11.08	54.72818
12	*L1_2508*	5	3	2	6.30	36.65	5.5	1.2	14.03	35.68152
13	*L4_2508*	65	43	22	52.60	138.45	1.9	0.6	13.55	37.06619
14	*R1_2508*	32	24	8	19.80	14.07	2	0.7	10.90	56.43816
15	*L5_2508*	6	3	3	6.30	15.55	2.3	0.9	14.35	35.30460
	Mean	26.60	19.07	7.53	14.61	25.411	3.91	1.11	11.70	49.19723
	Median	11.00	6.00	4.00	6.30	15.18	4.3	0.9		
	STDV	31.439	25.398	7.060	15.768	34.685	2.56	0.58		
	Sample Variance	988.400	645.067	49.838	248.629	1203.04	6.55	0.33		
	Minimum	5	3	2	1.89	0.89	0.9	0.6		
	Maximum	120	95	25	52.6	138.45	8.4	1.5		

ANOVA two factors without replication analysis for the laboratory results (α=0.05, N=105)

SUMMARY	*Count*	*Sum*	*Average*	*Variance*
L1_2008	7	64.05	9.15	72.90063
L2_2008	7	33.73	4.818571	11.00651
L2a_2008	7	70.08	10.01143	41.17025
L3_2008	7	22.68	3.24	3.701833

[5] Secchi disc depth represents relative information, because of absence of the standard instrument; only a self-made, smaller than standard one was used during the fieldwork

L3a_2008	7	80.69	11.52714	132.0695
R4_2008	7	292.38	41.76857	2365.956
GR_2208	7	37.65	5.378571	7.444748
Lbc_2308	7	37.01	5.287143	6.650157
R2_2408	7	106.97	15.28143	195.9231
R3-2_2408	7	132.13	18.87571	400.3768
H1_2408	7	75.38	10.76857	343.7412
L1_2508	7	59.65	8.521429	157.3832
L4_2508	7	323.55	46.22143	2257.032
R1_2508	7	100.57	14.36714	135.5986
L5_2508	7	37.05	5.292857	24.24869
TSM, mg/l	15	399	26.6	988.4
ISM, mg/l	15	286	19.06667	645.0667
OSM, mg/l	15	113	7.533333	49.8381
Turb, NTU	15	219.21	14.614	248.6292
CHL, mg/l	15	381.16	25.41067	1203.044
Depth, m	15	58.6	3.906667	6.549595
Secchi[1] disc depth, m	15	16.6	1.106667	0.334952

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F observed*	*P-value*	*F crit*
Rows	16428.88	14	1173.491	3.577043	0.000127	1.811298
Columns	9374.018	6	1562.336	4.762322	0.000319	2.208552
Error	27557.2	84	328.0619			
Total	53360.09	104				

Table A.2 Result of ANOVA analysis for the laboratory samples

Analysis of Variance (ANOVA) test checks between-group-variability and within-group-variability. Null hypothesis asserts that there is no variability between group and within the group, or in other words (with given level of significance 95% (e.g.100-α), null hypothesis asserts that samples from different locations and different WQ parameters belongs to the same population:

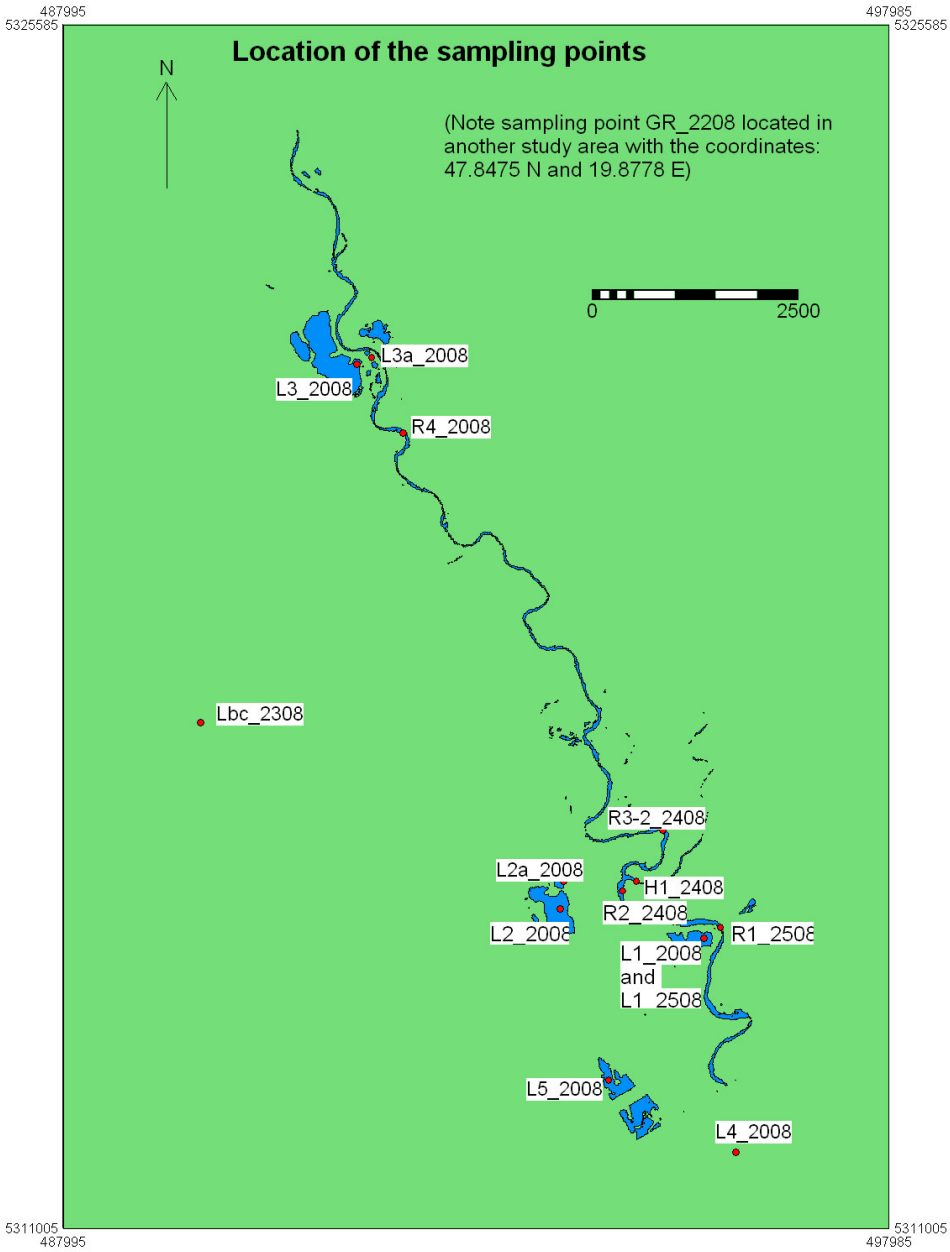
$$H_0=\mu_{loc1}= \mu_{loc2}=\mu_{loc3}=\mu_{loc4} \text{ and } H_0=\mu_{wq1}= \mu_{wq2}=\mu_{wq3}=\mu_{wq4}$$

Whereas alternative hypothesis asserts that samples are from separate and distinct populations:

$$H_a\neq\mu_{loc1} \neq\mu_{loc2}\neq\mu_{loc3}\neq\mu_{loc4} \text{ and } H_a\neq\mu_{wq1} \neq\mu_{wq2}\neq\mu_{wq3}\neq\mu_{wq4}$$

Since F observed is more than F critical for both sampling points and water constituents; P-value is lower than α=0.05, we can conclude that with 95% confidence level that there is a significant variation within measurements at different locations and within different water constituents concentrations.

Figure A.1 Map of the sampling sites



Appendix B. Senosors and Spectra

Sensors characteristics

Spectrometer GER3700

Table B.1 Specifications of Spectrometer GER3700 (Source: http://www.ger.com/3700.html)

Spectral Range	350 nm to 2500 nm
Channels	704
Linear Arrays	(1) 512 Si (1) 128 PbS (1) 64 PbS
Bandwidth 1.5 nm: 6.5 nm:	350 nm to 1050 nm 1050nm to 1900nm
Scan Time	50 ms and up (selectable)
FOV	3° standard, 10° option 23° option with fiber optic
Head Size	320 mm x 335 mm x 320 mm 12.5" x 13.75" x 4.5"
Weight	6.3 kg, 13.8 lbs.
Battery Type	(1) 12 VDC
Battery Life	4 hours
Digitization	16 bit
Wavelength Repeatability	+0.1 nm
Noise Equivalent Radiance	0.5 s integration time 400 nm: 2.2 x 10^{-9}W · cm^{-2} ·nm^{-1}· sr^{-1} 700 nm: 1.6 x 10^{-9}W · cm^{-2} ·nm^{-1}· sr^{-1} 900 nm: 6.2 x 10^{-9} W · cm^{-2} ·nm^{-1} · sr^{-1} 1500 nm: 1.1 x 10^{-8} W · cm^{-2} ·nm^{-1} · sr^{-1}
Maximum Radiance Levels	13 ms integration time 700 nm: 1.5 x 10^{-4}W · cm^{-2} ·nm^{-1}· sr^{-1}
Radiometric Calibration Accurancy (Traceable to NIST)	400nm : ± 5% 700nm : ± 4% 1000nm : ± 5%
Dark Current Correction UV-NIR-SWIR	automatic
Spectrum Averaging	selectable
Humidity	10 to 90% RH (non-condensing)
Temperature	-10° to 50°C
Sighting	Laser

ROSIS sensor

Table B.2 ROSIS sensor's specifications (Sources: DLR, Martin Habermeyer, personal communication (2002) and http://www.op.dlr.de/ne-oe/fo/rosis/home.html)

System	Push broom imaging spectrometer
Detector	CCD array (Thomson THX 7895)
FOV	16°
IFOV	0.56 mrad
Spectral range	430 - 850 nm
Spectral half width	4 nm
Radiometric encoding	14 bit
Pixel size	3 meters

DAIS sensor

Table B.3 DAIS sensor's specifications (Source: http://www.op.dlr.de/dais/dais.htm)

Detector	Si
FOV	0.894 rad (±26 degrees)
IFOV	3.3 mrad, (0.189 degrees)
Spectral range: VIS/NIR[6]	400 - 1000 nm
Number of bands: VIR/NIR[3]	32
Spectral half width	About 17.6 (depending on wavelength)
Radiometric encoding	15 bit
Pixel size	5 meters

Spectra

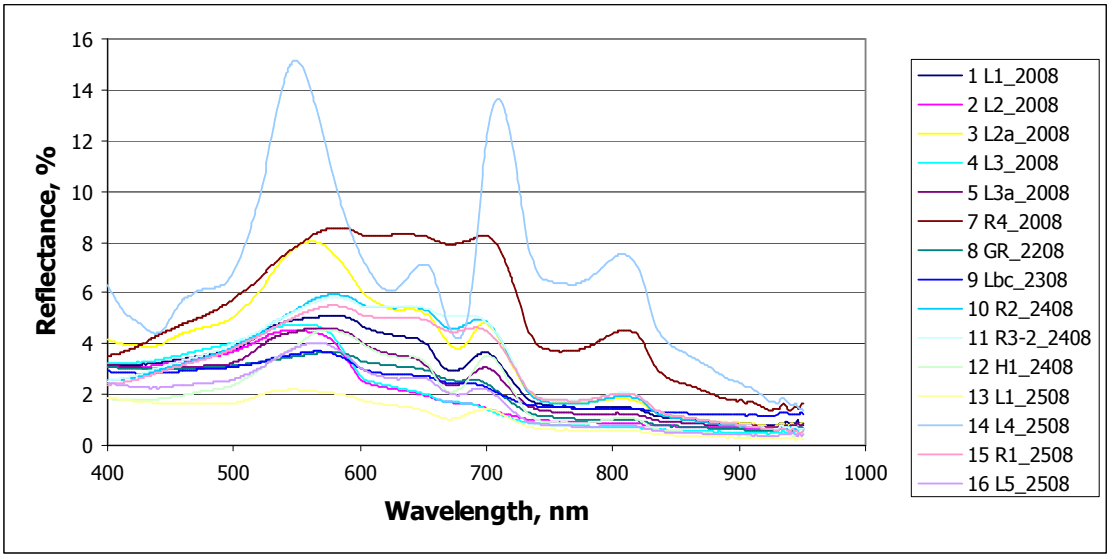


Figure B.1 Field spectra collected with the spectrometer GER 3700

[6]In the spectral range, only VIS/NIR range considered here (suitable for WQ monitoring). In fact, it has much wider spectral range – up to 12600 nm in 79 bands

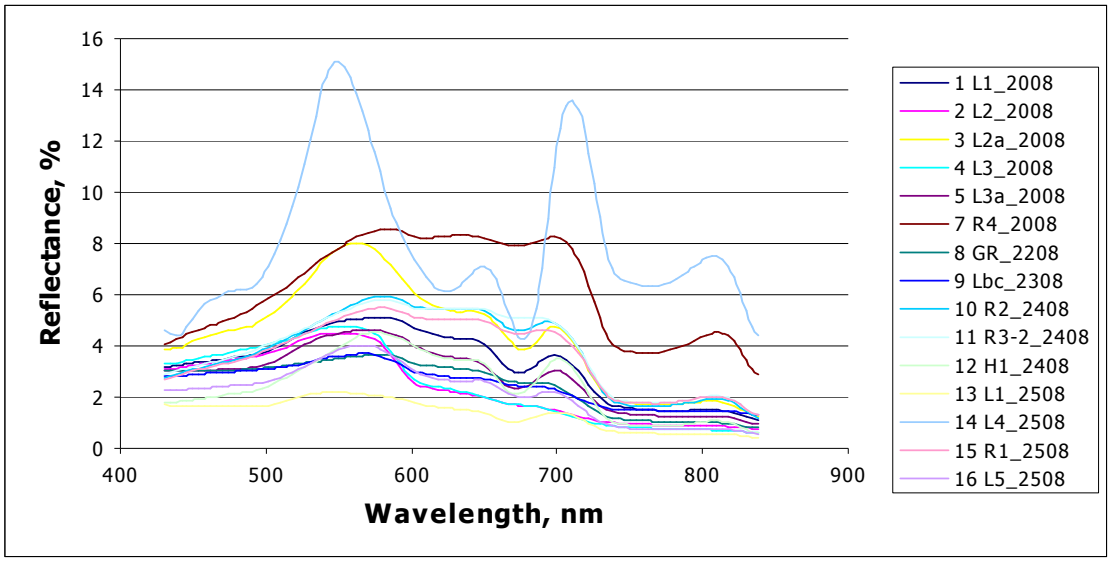


Figure B.2 Field spectra resampled to the wavelengths of the ROSIS sensor

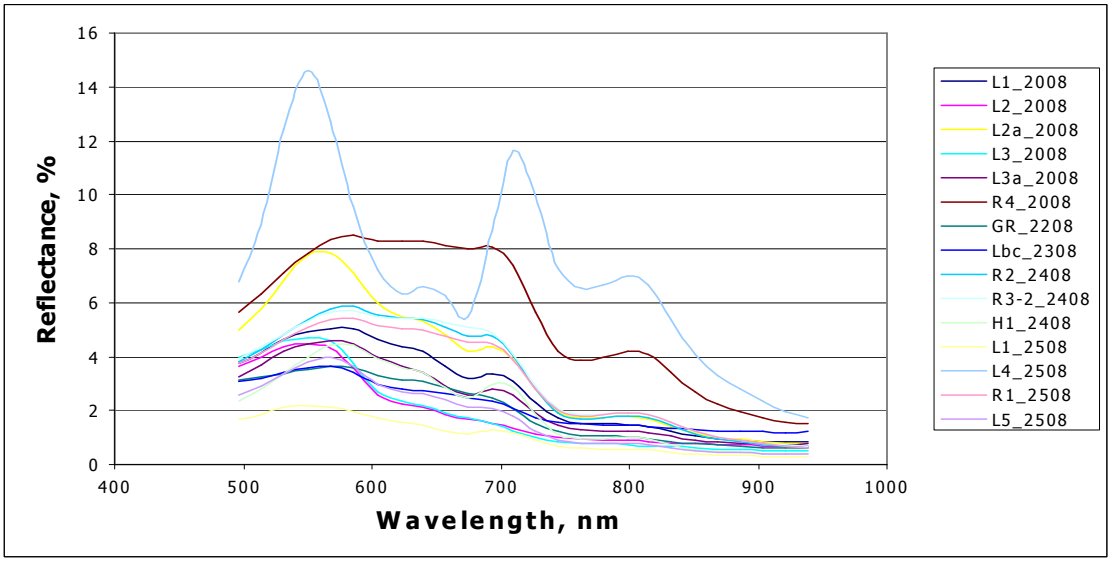


Figure B.3 Field spectra resampled to the wavelengths of the DAIS sensor

ANOVA two factors without replication analysis for the field spectrometry measurements (α=0.05, N=6592)

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F observ*	*P-value*	*F crit*
Rows	12802.214	411	31.14894	38.593559	0	1.1219372
Columns	11682.093	15	778.80618	964.9414	0	1.6680062
Error	4975.7841	6165	0.8071021			
Total	29460.091	6591				

As in the case of samples analysis, F observed is far higher than F critical and P value is 0 for both rows and columns, which means there is a significant variation within the field spectrometry measurements at different locations and within reflectance values at different wavelengths.

Landsat TM

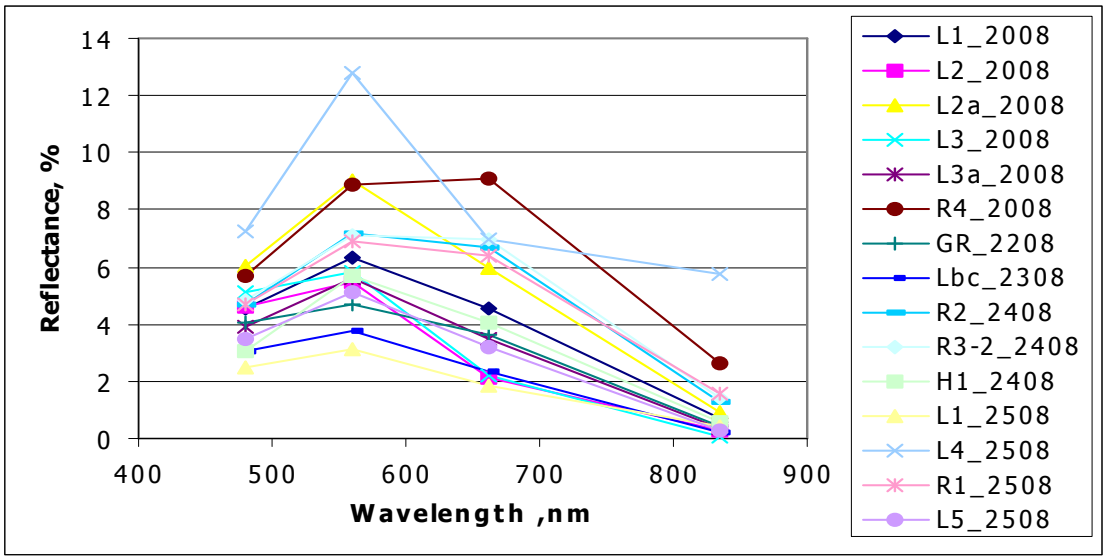


Figure B.4 Field spectra resampled to the wavelengths of the Landsat TM sensor

Spectra from the Dutch lakes (The Dutch data set)

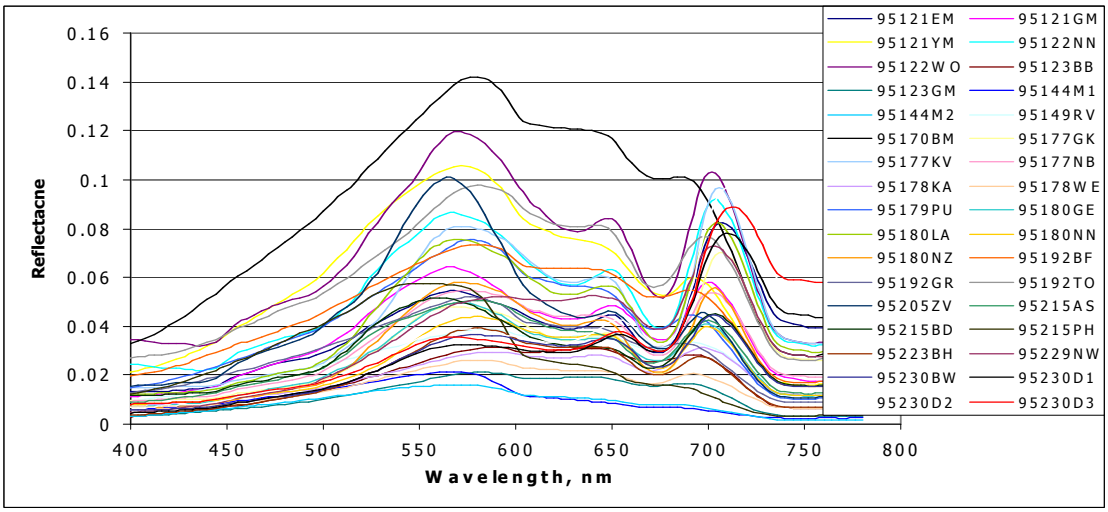


Figure B.5 Reflectance spectra collected from the Dutch lakes

Appendix C. Computation of sediment transport capacity using Yalin equation (using water level measured at Sajólad gauging station on 17/08/2002)

Small particle's concentration calculation

Input	Flow rate Q, m3/s	78
	Manning's n	0.022
	Friction slope S, m/m	0.001
	Particle diameter d, mm	0.007
	Particle specific gravity Sg	2.65
	Channel width W, m	38
	Density of sediment, kg/m3	2650
	Density of water, kg/m3	1000
	Temperature of water , o C	14
Output		
Kinematic viscisity (f (temp.)		1.18E-06
Flow depth D,m		1.24764
Hydraulic radius R,m		1.170762
Flow shear Á, N/m2		11.48517
Shear velocity U*, m/s		0.107169
Dimensionless shear T*		101.3646
Boundary Reynold Re(b)		0.637655
LN (Tcr) -read Shiel'd curve		-1.66301
Critical shear Tcr		0.189567
a	from Yallin equation	0.722352
d	from Yallin equation	533.7172
s	from Yallin equation	385.5317
Psc factor	from Yallin equation	333.6736
Sediment trans. capacity Qsc per unit width, kg/m*s		0.663337
Transport capacity of channel Tc, kg/s		25.20682
Flow velocity from Manning eq. u, m/s		1.596699
Maximum sediment concentration C sm, kg/m3		0.332983

Large particle's concentration calculation

Input	Flow rate Q, m3/s	78
	Manning's n	0.022
	Friction slope S, m/m	0.001
	Particle diameter d, mm	0.075
	Particle specific gravity Sg	2.65
	Channel width W, m	38
	Density of sediment, kg/m3	2650
	Density of water, kg/m3	1000
	Temperature of water , o C	14
Output		
Kinematic viscisity (f (temp.)		1.18E-06
Flow depth D,m		1.24764
Hydraulic radius R,m		1.170762
Flow shear Á, N/m2		11.48517
Shear velocity U*, m/s		0.107169
Dimensionless shear T*		9.4607
Boundary Reynold Re(b)		6.832017
LN (Tcr) -read Shiel'd curve		-3.33686
Critical shear Tcr		0.035548
a	from Yallin equation	0.312807
d	from Yallin equation	265.1368
s	from Yallin equation	82.93674
Psc factor	from Yallin equation	159.3688
Sediment trans. capacity Qsc per unit width, kg/m*s		3.394527
Transport capacity of channel Tc, kg/s		128.992
Flow velocity from Manning eq. u, m/s		1.596699
Maximum sediment concentration C sm, kg/m3		1.70399

Maximum sediment concentration of large particles , kg/m3	1.70399
Maximum sediment concentration of small particles, kg/m3	0.332983
Ratio of large particle's concentration to small	5.117346
Compare: ratio used in bio-optical modelling of TSM on 17/08/2002:	*4.79*

Appendix D. Maps

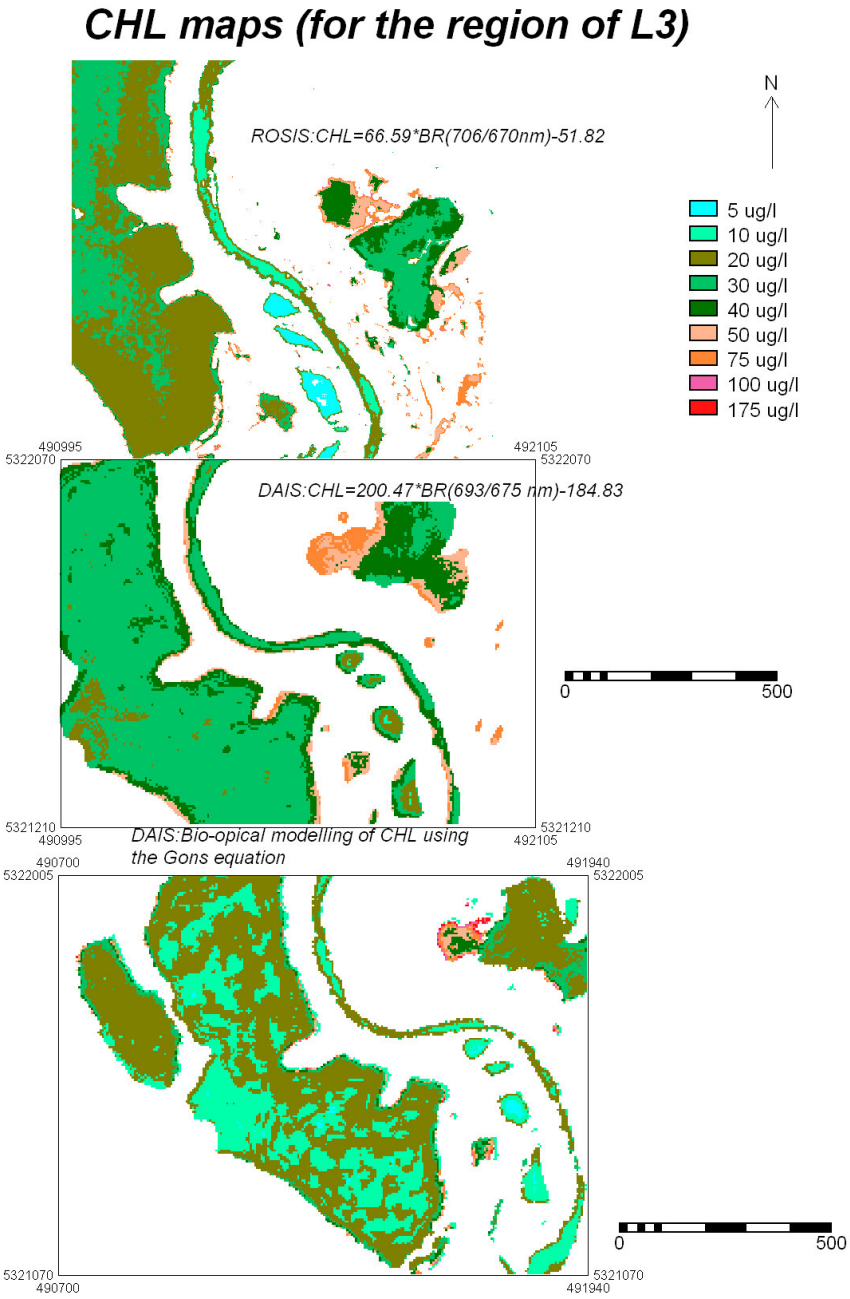


Figure D.1 CHL maps

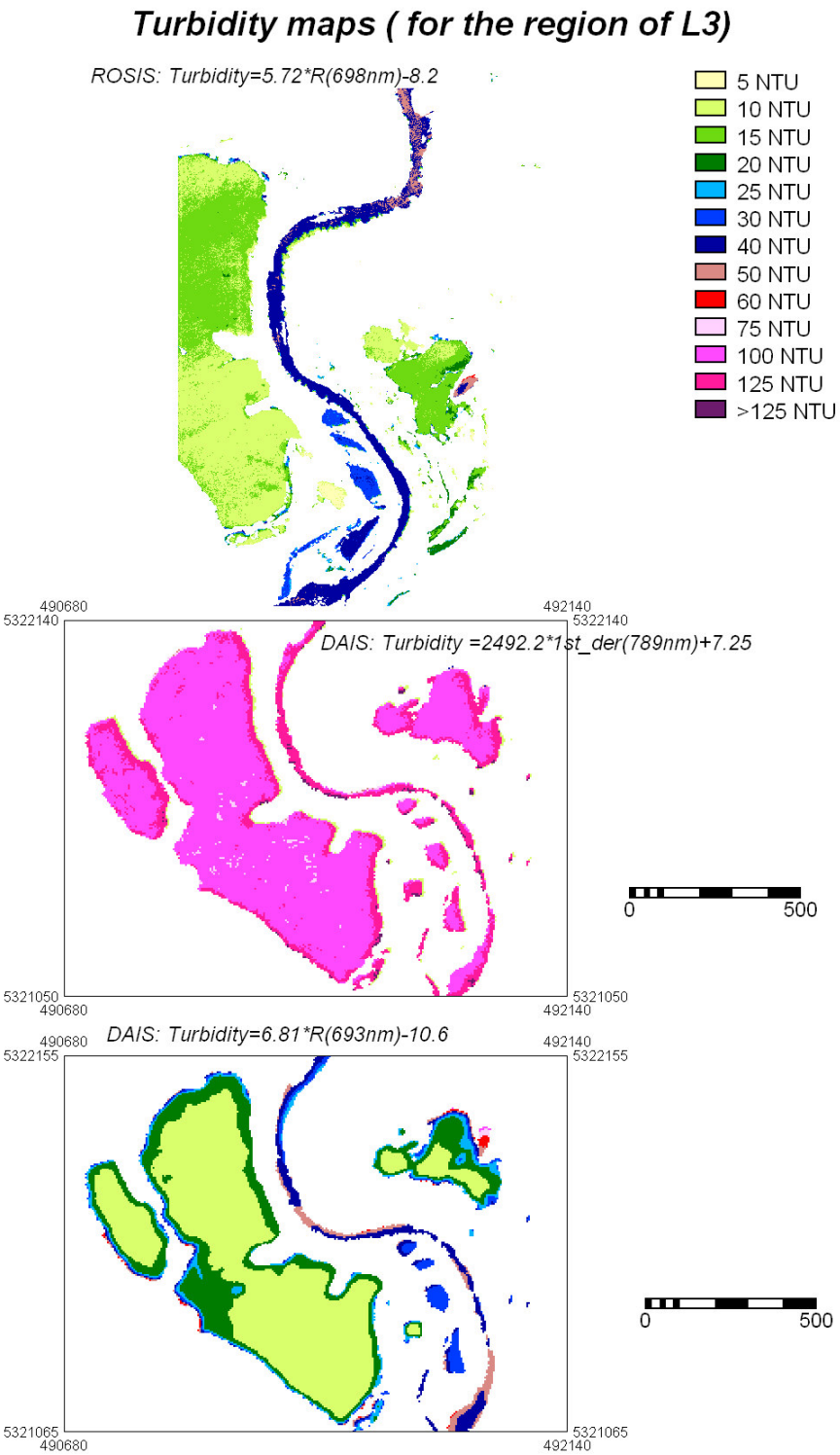


Figure D.2 Turbidity maps

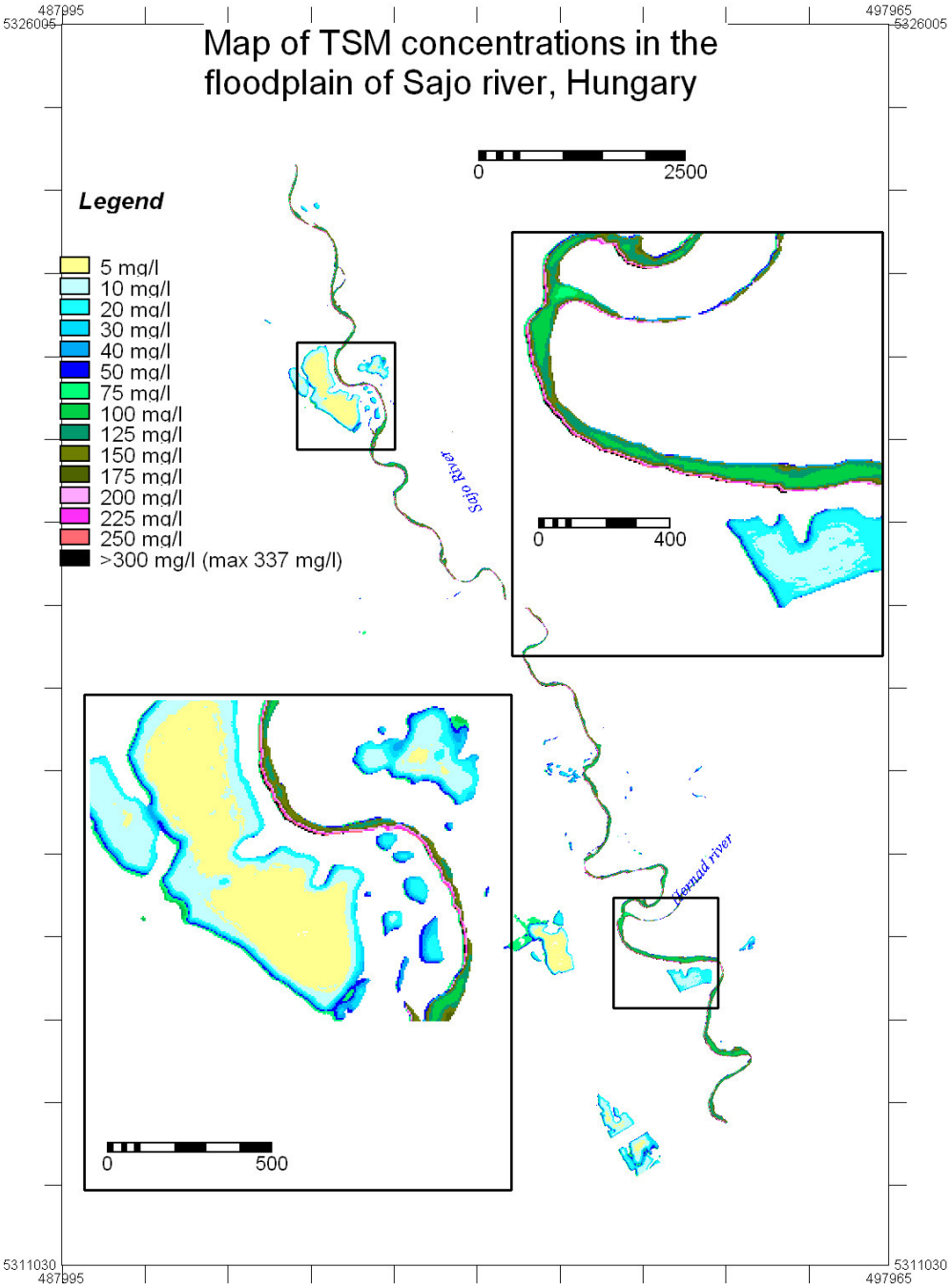


Figure D.3 TSM map

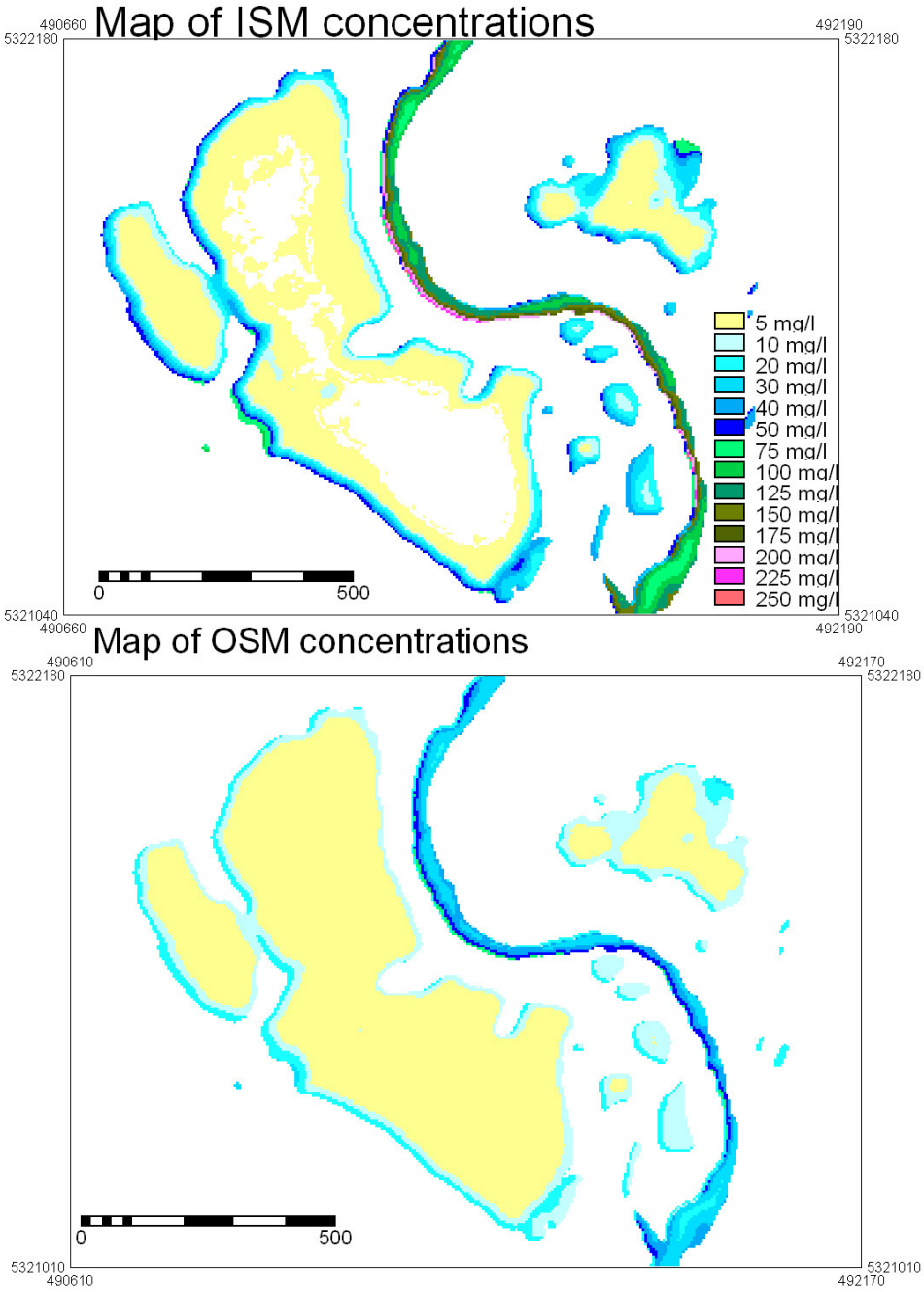


Figure D.4 ISM and OSM maps