Evaluating water quality monitoring with hyperspectral imagery

Adam Goodenough

Follow this and additional works at: http://scholarworks.rit.edu/theses

Recommended Citation

This Thesis is brought to you for free and open access by the Thesis/Dissertation Collections at RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact rit scholar works@rit.edu.
Evaluating Water Quality Monitoring 
with Hyperspectral Imagery

Final Report

Adam Goodenough
Center for Imaging Science
Rochester Institute of Technology
May 2001

Table of Contents
Evaluating Water Quality Monitoring with Hyperspectral Imagery

Adam Goodenough

Table of Contents

Abstract

Copyright

Acknowledgments

Introduction

Background and Theory

- Modeling radiative transfer in water
- Constructing and using the lookup table
- Empirical line method calibration in ENVI

Methods

- Overview of methods
- Tools used
- Water model: Atmospheric input
- Matching measured spectra with spectra in the look up table

Results

- Verification of the water model
- Concentrations derived from hyperspectral imagery

Discussion

Conclusions

References

Title Page
Abstract

Water quality is an important indicator of the health of an environmental system. Traditionally, water quality analysis has involved directly sampling areas in question. This method is difficult, if not impractical, to apply to large areas where data needs to be taken frequently. 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. The drawback of remote sensing data, however, is that these physical quantities must be derived with little or no physical sampling of the water. Fortunately, there are many indicators of water quality, one of which is the spectral character of the light leaving the water surface. This project attempts to monitor water quality through the use of radiative transfer models that simulate water optical properties for varying levels of constituents. These simulated signals are then matched against those observed by an airborne instrument. The thrust of this project is to determine the accuracy and sensitivity of the water quality extraction algorithms to key input parameters by comparing against ground truth. Details of the techniques will be presented along with an analysis of the errors.
Title: Evaluating Water Quality Monitoring with Hyperspectral Imagery

Authors: Adam A. Goodenough, Rolando V. Raqueno, Jason A. Hamel

Project Advisor: Dr. John R. Schott

SIMG 503 Instructor: Anthony Vodacek

Table of Contents
Evaluating Water Quality Monitoring with Hyperspectral Imagery

Adam Goodenough

Acknowledgment

Many thanks to Dr. John Schott for being my project advisor and Jason Hamel for doing most of the look up table generation.

I would especially like to thank Rolando Raqueno for letting me work on this joint project and being patient with me as I went through the learning process.

Table of Contents
Evaluating Water Quality Monitoring with Hyperspectral Imagery

Adam Goodenough

Introduction

The objective of this study was to produce concentration maps of various water quality indicators that are found in the Great Lakes for environmental purposes. The focus of this particular work was to examine and validate the water and atmospheric models. This report, however, attempts to address most of the important aspects of the joint project as well as the final results in addition to the primary focus. Please see http://www.cis.rit.edu/research/dirs/research/hywater for additional treatment of the topics covered here as well as more information on the project itself.

Background and Theory

Modeling radiative transfer in water

The optical properties of water can be broken down into two categories. Inherent optical properties (IOP's) are properties that depend solely upon the medium through which the light is traveling (i.e. they are independent of the spatial distribution of the light). Inherent optical properties include the absorption coefficient, the scattering coefficient, the volume scattering function, the index of refraction, the beam attenuation coefficient, and the single-scattering albedo. Apparent optical properties (AOP's) include properties that are dependent on both the medium and the distribution of the light. Examples of apparent optical properties include the diffuse attenuation coefficients and the irradiance reflectance.

This project deals mainly with the relationship between a wide variety of inherent optical properties and the apparent optical property named spectral remote-sensing reflectance ($R_{rs}$). The spectral remote-sensing reflectance measures how much of the downwelled irradiance that is incident on the water's surface is returned to the surface in a specific direction. This differs from the spectral irradiance reflectance ($R$) in that the spectral irradiance reflectance is simply a ratio of the upwelled irradiance to the downwelled irradiance. The equation for the spectral remote-sensing reflectance is given in equation 1.

$$R_{rs} = \frac{L(\theta, \varphi, \lambda)}{E_d(\sigma = a, \lambda)}$$

(Equation 1)

$R_{rs}$ is defined in terms of the direction of the water leaving radiance relative to the surface ($\theta, \varphi$) for each wavelength ($\lambda$). The water leaving radiance ($L$) is a function of the water leaving direction ($\theta, \varphi$) and the wavelength ($\lambda$). The depth ($z$) is set to the air just above the surface of the water ($a$). Similarly, the downwelled irradiance ($E_d$) is evaluated just above the surface ($z=a$) for each wavelength ($\lambda$).

The beam attenuation coefficient, $c(\lambda)$, is an inherent optical property that describes how much of the radiant energy traveling through a medium is lost through the processes of absorption, $a(\lambda)$, and scattering, $b(\lambda)$, per unit distance. This is shown mathematically in equation 2.

$$c(\lambda) = a(\lambda) + b(\lambda)$$

(Equation 2)

The absorption is defined as the ratio of the amount of flux that is absorbed to the total radiant flux, as given in equation 3.

$$a(\lambda) = \frac{[\Phi(\xi, \lambda)]_{abs}}{\Phi(\xi, \lambda) d\xi}$$

(Equation 3)

This equation gives the fraction of flux absorbed over an infinitesimal distance ($d\xi$) that will be evaluated over a particular distance ($r$) and wavelength ($\lambda$). Similarly, the scattering coefficient, $b(\lambda)$, is defined in equation 4.

$$b(\lambda) = \frac{[\Phi(\xi, \lambda)]_{scat}}{\Phi(\xi, \lambda) d\xi}$$

(Equation 4)

Again, this evaluates the fraction of radiant flux that is scattered over a distance ($r$) for a particular wavelength ($\lambda$) ($1$).

The total absorption and scattering of a medium can be described as the sum of the absorption and scattering coefficients of the individual components of the medium. Thus, if the medium is said to have four components (denoted ABCase2 by Hydrolight 4.1), four sets of absorption...
and scattering coefficients describe the water. Examples of the absorption and scattering coefficient curves of chlorophyll in the visible spectrum are shown in figures 1 and 2, respectively.

The four components of this model (ABCase2) are the water itself, chlorophyll (CHL), total suspended solids (TSS), and colored dissolved organic matter (CDOM). The chlorophyll component consists of living organisms such as zooplankton, phytoplankton, bacterioplankton, and algal fungi. Total suspended solids consist of suspended minerals and other inorganic matter. The category of colored dissolved organic matter (also called CDOM or gelbstoff) is comprised of dissolved yellow matter that does not scatter.

Although the scattering coefficient describes the fraction of light scattered out of the incident beam, it does not provide any information as to where the light goes. In pure water, light would be scattered in a manner consistent with Rayleigh scattering with approximately equal amounts being scattered backwards as forwards. However, as soon as there are particles introduced into the water, the shape of the scattering changes drastically so that the majority of the radiant flux is scattered in the forward direction. Fournier and Forand have described the shape of the scattering (the scattering phase function) as a function of two input parameters. The first parameter, \( n \), is related to the refractive index of the scattering particles in the water. The second parameter, \( \nu \), is related to the size distribution of the particles. Examples of phase functions for different parameters \((\nu, n)\) are shown in figure 3 on a logarithmic scale.

In modeling the radiative transfer of light in water, it is also necessary to model how much light reaches the surface of the water from the atmosphere (the downwelled irradiance). This project will examine two approaches of defining the atmosphere. The first involves using a simplistic model named RADTRAN that is built into Hydrolight 4.1. The second entails an involved process of modeling the atmosphere using MODTRAN. The success of either method can be determined by comparing the total downwelled irradiance calculated by the model with actual measurements of the irradiance made at the scene locale.

Constructing and using the lookup table

The radiative transfer model (Hydrolight 4.1) can be run for a number of different concentrations of the constituents (CHL, TSS, and CDOM) to construct a lookup table of reflectance spectra. This three-dimensional lookup table (LUT) can then be used to predict the component concentrations in a scene from the spectra of a hyperspectral image in reflectance space. Such a lookup table is shown in figure 4.
Thus, each of the reflectance spectra in the lookup table is indexed by the concentrations of the constituents. An algorithm can be used to find a pairing between the image reflectance curve and a lookup table curve that has the least difference between the two. Such a measure of goodness of fit is squared error as given in equation 5.

\[ SE = \sum (R_{\text{obs}}(\lambda) - R_{\text{m}}(\lambda))^2 \]  

(Equation 5)

The squared error (SE) is simply the sum of the squared differences between the observed reflectance \( R_{\text{obs}} \) from the image and the modeled reflectance \( R_{\text{m}} \) from the lookup table for each wavelength. Spectra that fall between members of the lookup table can also be estimated by interpolating with surrounding spectra (11).

**Empirical line method calibration in ENVI**

ELM calibration in ENVI takes pairs of data spectra and field spectra that are given as inputs. For every band within these spectra, ENVI extracts pairs of reference values (from the field spectra) and corresponding digital counts (from the data spectra) and performs a linear regression through the data points. Figure 5 is an example of this with 4 different pairs of spectra.

![Figure 5](https://www.cis.rit.edu/research/thesis/bs/2001/goodenough/th_body.html#I.0)

Usually it is only necessary to have two pairs—in which case the linear regression is simply a line connecting the two points. If, for some reason, ENVI is only given one pair of spectra, it uses the origin (a 0 digital count and a 0 reference value) as the second point for the regression (this is not recommended). Essentially, this allows the relationship between the digital counts and the reference values for each band to be defined by two constants: the slope and the intercept of the regression line. ENVI produces a text file with a ".cff" extension when the ELM calibration is done that is simply a file containing the slopes and intercepts for each band. An example of the output file is shown in Figure 6.

![Figure 6](https://www.cis.rit.edu/research/thesis/bs/2001/goodenough/th_body.html#I.0)

For every pixel in every band, ENVI uses the relationships given by these coefficients to convert a digital count to a reflectance—resulting in the ELM calibrated image. ENVI refers to the slope curve as Solar Irradiance and the intercept curve as Path Radiance. The final output image is in reflectance space rather than radiance space.
Methods

Overview of methods

There are two main directions that this project must be addressed from. The first involves inverting hyperspectral imagery acquired from two airborne sensors. The first sensor, AVIRIS, has 224 bands from 0.4-2.45 microns. The second sensor, MISI, has 72 bands ranging from 440 nm to 1020 nm. Once the hyperspectral images are in reflectance space, the accuracy of the calibration can be determined by comparison with reflectance spectra measured (with a spectrometer from Analytical Spectral Devices, Inc.) while the sensor was overhead. The second half of this approach involves using Hydrolight 4.1 to model the reflectance spectra from water containing particular component concentrations. A large number of runs are made for a variety of component concentrations and a lookup table is formed. This allows measured spectra to be matched to the lookup table resulting in predicted concentrations (inverting the water model). The accuracy of the lookup table and the algorithm that matches it to observed spectra can be measured by attempting to match collected spectra (ASD) of targets with known constituent concentrations. An overview of this process is shown in figure 7.

Tools used

Figure 8 identifies the actual tools used for each part of the project. MISI and AVIRIS supply the source imagery. Digital counts in the hyperspectral images are converted from radiance space at the sensor to reflectance space at the surface of the water using the empirical line method transformation discussed previously. MODTRAN is used to model the atmospheric conditions at the time of interest. Specifically, MODTRAN is used to create the spatial and spectral characteristics of the downwelled irradiance used as input for Hydrolight. Finally, Hydrolight 4.1 is used to model the radiative transfer of light in water.
Water model: Atmospheric input

In addition to the key inputs of inherent optical properties and scattering phase functions that are described in *Modeling Radiative Transfer in Water* above, the input downwelled irradiance to Hydrolight is supplied by MODTRAN. This input consists of the spectral characteristic and amount of downwelled irradiance coming from discrete quadrants defined for the sky. For this project these quadrants were approximately 5 degrees (azimuth angle) by 5 degrees (zenith angle) in size. An example of MODTRAN's output is shown in figure 9. The left image shows the spatial profile of the sky with the sun at 30 degrees measured from zenith (the quadrant including the sun is not used when scaling the rest of the sky between black and white). The diagram on the right shows the total downwelled irradiance from the sky as well as the diffuse downwelled irradiance (not including the direct irradiance from the sun). The calculated total irradiance is comparable to measured values. An example of measured total downwelled irradiance from the ASD website ([http://www.asdi.com/apps/arm.html](http://www.asdi.com/apps/arm.html)) is shown for comparison in figure 10 (units difference between microns and nanometers accounts for the scaling difference of a 1000).
Matching measured spectra with spectra in the look up table

Spectra of interest must be matched to the look up table in order to get predicted concentrations. This matching process is done using the AMOEBA algorithm implemented in IDL. Not only does the matching process find the closest match between the input spectra and the spectra contained in the look up table, the process also interpolates the spectra in the look up table to find spectra corresponding to component concentrations not included in the look up table. This means that results are not limited to the discrete number of concentrations input to the look up table. It was found that a simple, full spectral minimal error matching did not produce consistent nor accurate matches. This may be due to the fact that the three parameters (chlorophyll, total suspended solids, and colored dissolved organic matter) do not affect the optical properties of the water independently. Therefore it was decided to weight the spectrum differently for each constituent based on empirical evidence that a correlation existed between a particular portion of the spectrum and an accurate prediction of the component's concentration. Although many different weightings were tried, figure 11 is a representative example of the type and location of the weightings. Concentrations are color coded to denote to the particular weighting that they were derived from.
Results

Verification of the water model

The validity of the water model Hydrolight 4.1 was established using measured data. The measured data consisted of measured water spectra as well as actual measured constituent concentrations (chlorophyll, total suspended solids, and colored dissolved organic matter) for that water sample. The measured water spectra were inputted into the inverted water model (look up table) to find predicted constituent concentrations. These predicted concentrations were then compared with the actual measured concentrations. Results from this analysis for four data points is shown in figure 12 (a perfect match would fall along the diagonal line).
Concentrations derived from hyperspectral imagery

After verifying that the inverted water model was working properly, the hyperspectral images underwent the empirical line method transformation. The same process of finding matches in the look up table was applied to the hyperspectral imagery (now in reflectance space at the surface of the water). Error analysis is done by again comparing areas with known (measured) constituent concentrations with the predicted values. The results of this process for the three constituents are shown in figures 14, 15, and 16 while figure 13 identifies the areas of interest.

![AVIRIS (Color Infrared)](image)

Figure 13: Areas of interest

![Chlorophyll Map](image)

Figure 14: Chlorophyll concentrations and error analysis

RMS = 11.6 mg/m³
18% of [CHL] range
Discussion

The inverted water model seems to be accurately predicting the constituent concentrations within a reasonable degree of error (no
minimum error requirements were established as the necessary prediction accuracy varies greatly with the particular project the data might be used for. However, confidence in the validation of the model is limited by the fact that only four data points were available for analysis (water samples for which both the spectral reflectance and constituent concentrations were known). The model needs to be validated further with more data. Additionally, the effects of light reflecting off the bottom of the lake (in shallow regions) was not taken into account. This seems to be a necessary addition in regions of the lake where water is both shallow and clear and the spectral characteristics of the lake floor might affect the light that returns to the sensor. The process of inputting hyperspectral images into the same procedure also seems to give promising results. However, there is a significant phenomenon occurring in the center of the lake where concentrations of constituents (particularly CDOM) is inordinately high. This is likely due to the presence of sun glint (where the sun is reflecting directly off the water's surface in the direction of the sensor) which boosts the reflectance values and therefore the concentrations as well. This effect must be studied in greater depth. Additionally, the linear transformation between the radiance values at the sensor and reflectance values at the surface of the water is not the ideal way to convert the data. Optimally, the light could be propagated down through the atmosphere using known climate and atmospheric data for the time that the imagery was taken. The empirical line method is also highly dependent on the choice of reference points, which adds another source of error to the process.

Conclusions

Despite the fact that there are many inputs and steps to the process, the inverted water model (look up table) does seem to be able to accurately predict concentrations of constituents over a large spatial area. The entire process (with the exception of the empirical line method transformation and the choice of weightings) is based on fundamental physical principles. This suggests that the methods shown here can be applied to other hyperspectral images with similar results.
Evaluating Water Quality Monitoring
with Hyperspectral Imagery

Adam Goodenough

References


Table of Contents