Hyperspectral Remote Sensing of Water Quality Parameters for Large Rivers in the Ohio River Basin

Naseer A. Shafique, Florence Fulk, Bradley C. Autrey, Joseph Flotemersch

Abstract

Optical indicators of water quality have the potential of enhancing the abilities of resource managers to monitor water bodies in a timely and cost-effective manner. However, the degree to which optical indicators are useful may depend on their applicability to data collected from multiple water bodies. In 1999, a Compact Airborne Spectrographic Imager (CASI) was flown over the relatively shallow Great Miami River (GMR), in Southwest Ohio, collecting hyperspectral bands of data. Concurrently, water quality samples and hand-held spectrometer data were collected directly from the river. Using correlations between the ground-truth data and combinations of spectral bands from the remotely sensed data, spectral indices were developed which could be used to estimate chlorophyll *a*, turbidity and phosphorus. In 2001, a similar study was conducted in which a CASI was flown over a portion of the Ohio River while ground-truth data were collected. These data were analyzed and tested against the spectral indices developed during the 1999 study. The GMR's spectral index for chlorophyll *a* was applicable to the Ohio River data. However, slightly refined spectral indices for turbidity and phosphorus were required in this new environmental setting. This study demonstrates the ubiquitous application of the chlorophyll *a* spectral index while revealing the limited reliability of the turbidity and phosphorous spectral indices. Although differences between the dynamics of the two rivers may have made these spectral indices incompatible, with further refinement they may yet prove to be useful tools that can be modified for use in other rivers to detect potential water quality problems.

Keywords: remote sensing, water quality, chlorophyll *a*, nutrients

Introduction

Eutrophication diminishes water quality by promoting the excessive growth of algae, and increasing suspended organic material. When degraded, unpleasant odors and tastes can result from the excessive amounts of algae. Furthermore, microorganisms associated with eutrophication may pose health risks to consumers. It is important for water resource managers to find the most efficient way to diagnose the condition of drinking water sources. This may be especially demanding when assessing water quality damage in large rivers when field measurements may be time consuming, costly, and limited logistically. Increases in water quality parameters such as chlorophyll *a*, turbidity, total suspended solids (TSS), and nutrients are symptomatic of eutrophic conditions. Concentrations of these parameters can provide insight on the extent of eutrophication and the potential impact on aquatic biota and overall water quality. It would be advantageous to resource manager to be able to detect eutrophic conditions using multiple sites in a river without relying on field measurements.

Shafique is a Remote Sensing Specialist at SoBran, Inc. and a Contractor of the U.S. Environmental Protection Agency, Cincinnati, OH 45268. E-mail: Shafique.Naseer@epa.gov. Fulk is a Statistician, Autrey is a Research Biologist, and Flotemersch is a Research Ecologist for the U.S. Environmental Protection Agency, Ecological Exposure Research Division, National Exposure Research Laboratory, Cincinnati, OH 45268.

Environmental researchers have been making efforts to monitor, simulate and control eutrophication for more than two decades. Various mathematical models have been developed and applied to rivers, lakes and estuaries (Lung 1986, Thomann and Mueller 1987, Kuo and Wu 1991, Kuo et al. 1994). All water quality models simulate increases in eutrophication based on the initial condition of the water body and therefore demand comprehensive water quality sampling programs. However, the conventional measurement of water quality requires *in situ* sampling and expensive and time-consuming laboratory work. Due to these limitations, the sampling effort often does not represent the condition of an entire water body. Therefore, the difficulty of overall and successive water quality sampling becomes a barrier to water quality monitoring and forecasting.

Remote sensing could overcome these constraints by providing an alternative means of water quality monitoring over a greater range of temporal and spatial scales (Shafique et al. 2001). Remote sensing is the science of measuring the properties of objects by measuring the amount of radiation they absorb, emit, or reflect at various wavelengths along the electromagnetic spectrum. Optical water quality research has a broad scope for developing environmental indicators that are useful in assessing, quantifying and monitoring instream water quality. Measurable parameters for optical water quality includes the attenuation coefficient (K_D) of photosynthetic active radiation (PAR), turbidity, concentrations of algal chlorophyll, suspended sediment, and dissolved organic matter. More fundamentally, the absorption and scattering of light by components of a stream's water column provide basic information from which relationships with other water quality indicators (such as water clarity from Secchi disk readings) can be derived (Jupp et al. 1994a, Dekker 1997). Although a fairly new method, the development of spectral indices can be a useful and easy tool for the diagnosis of eutrophic conditions by water resource managers.

Remote sensing techniques for monitoring coastal and inland waters have been under development since the early 1980's. The tools used to develop these techniques have ranged from an empirically-based method for producing qualitative water quality maps to semi-empirical techniques and analytical methods for producing quantitative water quality maps (Dekker 1997). Several investigators (e.g., Dekker 1993, Gitelson, et al. 1993, Jupp et al. 1994a, Jupp et al. 1994b) have developed empirical regression formulas for the prediction of lake water quality parameters from spectrometer data by employing spectral ratios, typically reflectance ratios, as the independent variables. The predicted water quality parameters have included chlorophyll *a* concentrations, suspended matter concentrations and turbidity.

This investigation also focuses on the prediction of chlorophyll *a* concentrations, turbidity and total phosphorus concentrations by applying spectral indices developed from spectral data collected by spectro-radiometers as independent variables. However, this study is unique in that it has developed a regression formula for lotic systems, specifically using large rivers as a model. Spectral indices are transformations of reflectance values at specific wavelengths that minimally correspond to a field tested concentration of the parameter of interest and minimize the effects of other optically active constituents. The method used made correlations using simultaneous collected remote data, field spectrometer data, and field collected water quality data to demonstrate the feasibility of remote sensing techniques for water quality monitoring in large rivers. This paper also discusses the application of the optical properties of water for multiple river systems. This includes the determination of the optical properties of some water quality parameters and the development of spectral indices using hyperspectral airborne data for the shallower Great Miami River (GMR), and the transfer of these spectral indices to the larger and deeper Ohio River.

Materials and Methodology

The investigation utilized compact airborne spectrographic imager (CASI) data from approximately 60 river miles of the GMR (Shafique et al. 2001), and 80 river miles of the Ohio River. Based on the analyses of preliminary field spectrometer data, 19 appropriate spectral bands with 5-nm spectral resolution were selected and programmed into the CASI unit. Data were collected during four field efforts, carried out under similar conditions, during 1999 and 2001. While the hyperspectral data were being collected by the airborne CASI, *in situ* water samples were collected and a field spectrometer was used to collect spectral data directly from the river.

The field, laboratory and remotely sensed data were analyzed in a systematic manner. First, the spectral library database was developed and used to establish the variability and/or stability of absorption and scattering coefficients in the GMR (Shafique et al. 2001) and the Ohio River. Single spectral bands, ratios of spectral bands, and combinations of multiple bands were then used to develop linear regression equations. The semi-empirical models were developed in Excel (Microsoft Software, v. 2000) spreadsheets, and the imagery was analyzed using the ENVI (3.6) image processing software. First, in order to represent a homogeneous unit in the imagery, the water area was masked. The unsupervised image classification technique, K Mean, was used to cluster imagery into spectrally similar categories. This classification technique was used over another method, supervised classification, because the identifications made by the latter are made on the basis of human sight, which is limited to visible wavelength range (Vincent 1997). Scatter plots were created between spectrally classified image and ground-truthing data, based on their linear trends; simple linear regressions were used to determine the relationships between single and combinations of bands and water quality parameters. Based on the image pixels of the locations from which ground-truth data were collected, equations were developed for particular water quality parameters. Then, the entire image was converted into a water quality map using the predictive equations. Some of the ground-truth data were not used to develop the equations, but were instead used to validate the predictive quality of the equations. Using this semi-empirical approach, separate equations were developed for each water quality parameter. This approach was the primary means used to analyze the images collected for this study.

The analytical approach of spectral image analysis used the spectral library that was developed from the 1999 GMR study and the 2001 Ohio River study (Shafique et al. 2001). The reflectance/radiative transfer model used in that analytical approach quantifies and simulates the individual contributions of the water constituents to the reflectance measured by the remote sensors. The development of a reflectance/radiative transfer model depends on how well the specific absorption and scattering coefficients are determined for the various constituents. Once stable specific coefficients were established, the reflectance/radiative transfer model could be used to mathematically convert airborne imagery into water quality maps with limited use of ground-truth data. The use of the ground-truth data could be limited because a physical understanding of the interactions between the various water constituents and water reflectance was incorporated into the radiative transfer equations. The success of the analytical approach depends on the successful optical characterization of the water body and potential contributing sources such as industrial wastewater discharges. Once such a characterization is made, available optical water quality toolkits can be adapted to a particular study region. One of these prototype toolkits has been developed by Dekker (1997).

Models

All bands were tested for relationships with water quality parameter until it was found which bands and parameters correlated with the highest certainty. Scatter plots showed that linear models using the ratio of wavelengths 705/675 nm and the logarithmic ratio of wavelengths 554/675 nm can describe chlorophyll *a* and total phosphorus, respectively. Logarithmic transformation is useful in cases, such as this, where it is necessary to stress the difference between scores in a manner that is proportional to their ratio rather than in terms of their absolute difference. The band readings that represented the difference of 740 nm from 710 nm correlated best to turbidity. The r-values and R^2 for each of these are above 0.7 and 0.5 respectively, therefore, indicating the ability to provide good linear models for these water quality parameters (Figure 1). Based on the linear relationship with water quality parameters, the spectral indices were then transferred to the following mathematical models to calculate the concentrations of the respective water quality parameters.

Chlorophyll a = 48.849 * (705/675 nm) - 34.876 (1)

TP = 0.1081 * log (554/675 nm) - 0.0371 (2)

Turbidity = 186.59 * (710 – 740 nm) + 8.5516 (3)

Results and Discussion

The results of this research can be divided into three general sections. The first section describes the preliminary data that were collected using the field spectrometer and the results of the water laboratory analyses. This data was used to explore the feasibility of using hyperspectral data for the identification and discrimination of in-stream features via the GMR and Ohio River. The second section describes the development of spectral indices using water constituents. It includes the A preliminary radiance/reflectance transfer model is presented that can be used to interpret remote sensing imagery in the same environment where the specific absorption and scattering coefficients are known. The third section addresses the correlations made between the remotely sensed imagery and the consequent estimations made for water quality parameters using the field spectrometer data. Through atmospheric and water column correction, the instrument calibrates the remotely-sensed data to the conditions that were present at the water's surface at the time of CASI data acquisition. Both semi-empirical and analytical models were applied to mathematically convert the hyperspectral imagery into water quality maps.

Relationships found using water quality and field spectrometer data

The correlation between various water quality parameters was calculated using a Pearson's correlation test. The most significant relationship observed was between the concentration of chlorophyll *a* plus pheophytin and the concentration of dissolved oxygen (DO). As is often observed in eutrophic systems, DO is negatively correlated (r = -0.81) with chlorophyll *a* plus pheophytin. Overall, dissolved oxygen had a negative correlation with other measured parameters.

There was a weak relationship between chlorophyll *a* concentration and water depth, and there was no significant correlation between reflectance at any wavelength and water depth. The water samples for turbidity data and the water samples for the chlorophyll *a* analysis were collected in close spatial and temporal proximity. Although it is assumed that the composition of the water did not vary significantly, it is recognized that an unaccounted error may be introduced when trying to establish relationships for these parameters.

Development of spectral indices

Spectral indices are simple arithmetic expressions of a combination of spectral bands that help reduce or eliminate some differences in viewing geometry and atmospheric conditions between measurements. Due to the large number of bands measured with the field spectrometer data, a sub-set of bands at 5 nm intervals was selected to develop the spectral indices. The bands that carry the most information about water quality parameters were selected by qualitatively analyzing field spectral plots of actual measurements. Generally, bands that show peaks and troughs (i.e., where the reflectance spectral curve changes slope) were selected.
Correlation values were calculated between ground-truth spectral data, water quality data from laboratory analyses, and the spectral indices developed from exploratory single bands, ratios of bands, differences between bands, and/or combinations of differences and ratios. The values from the indices with the highest correlation values with the ground-truth and laboratory data were used to produce scatter plots and calculate the R^2 values.

Once the bands were selected, three types of indices were developed. These were the difference, ratio, and combination of ratio and difference indices (Figure 1) originally developed from the GMR (Shafique et al. 2001). Table 1 shows spectral indices for the GMR and Ohio, River. It is noteworthy that the same index for chlorophyll *a* (i.e., 705/675 nm) worked equally well for both rivers. However, the other two indices (i.e. turbidity and total phosphorus) required slight modifications in the spectral band selections. Likely the modification was needed due to differences of suspended sediments and their effects on reflectance in the two rivers.

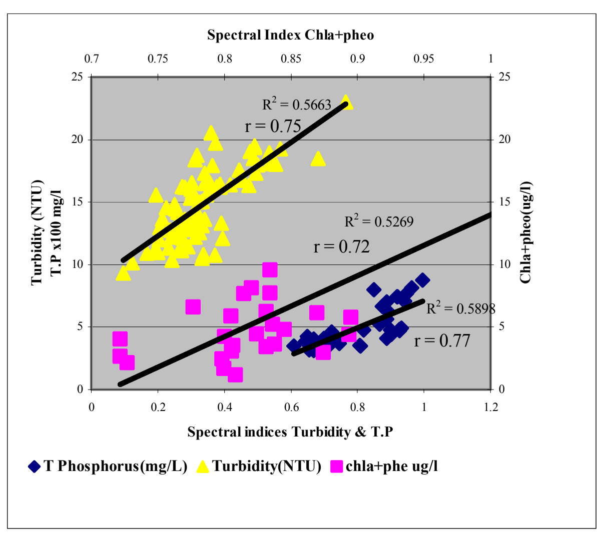

Figure 6. Correlation between water quality parameters and spectral indices. Total phosphorus values were multiplied by 100 to fit in the scale.

Table 4. Correlation coefficients for water quality parameters (WQP) and spectral indices (SI) for the two rivers.

WQP	SI	GMR	Ohio River
Turbidity	675-700	0.79	0.20
Turbidity	710-740	-0.33	0.75
Chlorophyll-*a*	705/675	0.71	0.72
T.P	(554/740)-(620-740)	0.66	0.34
T.P	log(554/675)	0.29	0.77

Correlations between water quality parameters and spectral data

Generally, the correlations between the ground-truth data and spectral parameters were stronger with the ratio indices and weaker with the individual bands. The difference and ratio indices seem to identify with a particular group of ground-truth data. For example, the difference index performed better for parameters such as turbidity, TSS and secchi depth, while the ratio index correlated better with chlorophyll *a* and pheophytin parameters. These differences are likely attributed to the reduction of turbulent effects in the water body when the ratio index was used, while the difference index reduces some of the atmospheric effects.

Spectral bands with wavelengths at 670, 675, 700, 705 and 740 nm appear to dominate the spectral indices. Of these, difference indices involving 675, 700 and 740-nm wavelengths provide information about turbidity while the ratio and combination indices using 672, 675, 700 and 705-nm wavelengths provide information about algal parameters. Taking into account all of the chlorophyll parameters, the concentration of chlorophyll *a* plus pheophytin correlated best with the same spectral parameters for the GMR and Ohio River. This can be explained by chlorophyll *a* and pheophytin tending to reflect and absorb light energy similarly. Logarithmic ratios of spectral bands in the green and red to near infrared region of the spectrum (i.e. 554 and 675-nm wavelengths) showed good correlation for TP.

Using the relationships between chlorophyll *a* and a ratio index (i.e., 705/675-nm wavelengths) from the CASI image, the ratio's 16 (K-means) spectral classes, converted into chlorophyll *a* concentration levels, ranged from 1.64 µg/l to 38.5 µg/l (Figure 2). The broad chlorophyll *a* levels are indicated on the atmospherically and radiometrically corrected CASI images acquired in September and November 2001 (Figure 2). The overlay of these classes on a true color image revealed plumes of higher chlorophyll *a* concentrations at some of the confluences of tributaries with the Ohio River. For example, the relatively high concentration of chlorophyll *a* in the Licking River can be seen at the confluences with the Ohio River. Similarly, classes of turbidity and TP were overlaid on the CASI image (Figure 2).

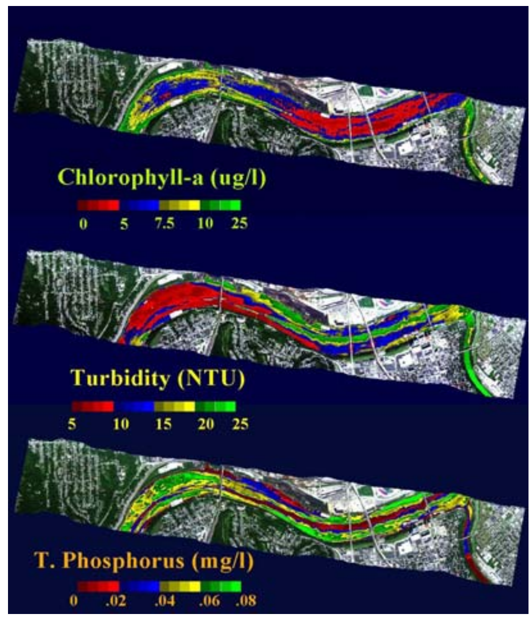

Figure 2. Water quality parameters map of Ohio River at the confluence of Licking River near Cincinnati, OH. Higher chlorophyll *a* and turbidity levels are found at the confluence. R^2 values of scatter plots of observed and estimated parameters are above 0.9 indicating high accuracy.

Conclusions

This study demonstrates that the hyperspectral remote sensing technique can be a useful tool for monitoring the distributions of chlorophyll *a* concentrations in large rivers. In this study, the wavelengths of 675 nm and 705 nm from the CASI data were found to be the most suitable wavelengths for predicting chlorophyll *a* concentrations. Correlation analysis between remotely sensed data and chlorophyll *a* data has indicated the possibility of mapping chlorophyll *a* concentrations accurately. The strong correlations of reflectance ratios corresponding to these wavelengths with field spectrometer data were used in the development of equations and constants for the estimated chlorophyll *a* concentrations in the Markland Pool.

The results show that it is also feasible to estimate the relative chlorophyll *a* levels of large rivers when ground-truth data is not routinely available. This is essential for operational applications in large rivers where the total number of *in situ* water quality observations only cover a small fraction of the river for a limited time. Moreover, the results indicate that chlorophyll *a* has a unique spectral signature and it is possible to estimate chlorophyll *a* concentrations for any inland water body with the chlorophyll *a* spectral index.

The methods developed and analyzed in this paper used CASI, but it is predicted that the same information can be revealed using hyperspectral data acquired by a satellite such as the Hyperion satellite. Although the spatial resolution of the data collected by the Hyperion satellite is only 30 m, in contrast to the 2-m resolution of the CASI data, the use of the satellite may be more cost effective and as reliable because the Hyperion spectrometer includes channels with the same wavelength bands employed here for chlorophyll *a* retrieval. However, it is still necessary to compare the results of data collected from *in situ*, CASI, and Hyperion studies to investigate the reliability of the data. Future research will provide information about whether satellite data can be substituted for field-collected data to determine water quality parameters such as chlorophyll *a*, nutrients, and turbidity. In the future, Hyperion remote sensing data may prove to be the preferable method for the detection of eutrophic water quality indicators over large areas of water.

Acknowledgments

The authors appreciate the efforts of Alicia Shelton of SoBran Environmental, Inc. in editing this manuscript, and the reviews and comments provided by Michael Troyer and Michael Moeykens of the U.S. EPA.

References

Dekker, A.G. 1993. Detection of Optical Water Parameters for Eutrophic Lakes by High Resolution Remote Sensing. Ph.D Thesis. Free University, Amsterdam.

Dekker, A.G. 1997. Operational tools for remote sensing of water quality: A prototype tool kit. Vrije Universiteit Amsterdam, Institute for Environmental Studies. BCRS Report 96-18.

Gitelson, A., G. Garbuzov, F. Szilagyi, K.-H. Mittenzwey, K. Karnieli, and A. Kaiser. 1993. Quantitative remote sensing methods for real-time monitoring of inland waters quality. International Journal of Remote Sensing 14:1269-1295.

Jupp, D.L., J.T.O. Kirk, and G.P. Harris. 1994a. Detection, identification and mapping of cyanobacteria using remote sensing to measure the optical quality of turbid inland waters. Australian Journal of Freshwater Research 45:801-828.

Jupp, D.L., G. Byrne, E. McDonald, J. Anstee, G. Harris, and J. Smith, 1994b. Monitoring optical water quality of the Hawkesbury River using airborne remote sensing. CSIRO Division of Water Resources, Consultancy Report No. 94/9.

Kuo, J.T., and J.H. Wu. 1991. A nutrient model for a lake with time-variable volumes. Water Science Technology 24(6):133-139.

Kuo, J. T., J.H. Wu, and W.S. Chu. 1994. Water quality simulation of Te-Chi reservoir using two-dimensional models. Water Science Technology 30(2):63-72.

Lung, W.S. 1986. Assessing phosphorus control in the James River Basin. Journal of Environmental Engineering 112(1):44-60.

Shafique, N.A., B.C. Autrey, F. Fulk, and S.M. Cormier. 2001. Hyperspectral narrow wavebands selection for optimizing water quality monitoring on the Great Miami River, Ohio. Journal of Spatial Hydrology 1(1):1-22.

Thomann, R.V., and J.A. Mueller. 1987. Principles of Surface Water Quality Modeling and Control. Harper & Row Publishers, New York.

Vincent, K.R., 1997. Fundamentals of Geological and Environmental Remote Sensing. Prentice-Hall, Upper Saddle River, NJ.