ABSTRACT

MODELING DISSOLVED ORGANIC CARBON (DOC) IN SUBALPINE AND ALPINE LAKES WITH GIS AND REMOTE SENSING

by Neil Thomas Winn

We use remote sensing and geographic information system (GIS) tools to develop simple predictive models to define relationships between watershed variables known to influence lake DOC concentrations and lake water color in the Absaroka-Beartooth Wilderness in Montana and Wyoming, USA. Variables examined include watershed area, topography, and vegetation cover. The resulting GIS model predicts DOC concentrations at the lake watershed scale with a high degree of accuracy ($R^2 = 0.92; p < 0.001$) by including only two variables: vegetation cover (representing sites of organic carbon fixation) and areas of low slope (0-5%) within the watershed (wetland sites of DOC production). Modeling with Advanced Land Imager satellite remote sensing data provided a somewhat weaker relationship between water color and DOC concentrations ($R^2 = 0.672; p < 0.001$). We compare model predictions to each other to determine success of DOC modeling methods ($R = 0.761; p < 0.001$).
Modeling Dissolved Organic Carbon (DOC) in Subalpine and Alpine Lakes With GIS and Remote Sensing

An internship report
Submitted to the Faculty of Miami University in partial fulfillment of the requirements for the degree of Master of Arts
Department of Geography

by
Neil Thomas Winn
Miami University
Oxford, Ohio
2008

Advisor__________________________________________
Robbyn J. F. Abbitt

Reader__________________________________________
Craig E. Williamson

Reader__________________________________________
William H. Renwick

Reader__________________________________________
Mary C. Henry
# Table of Contents

Introduction ......................................................................................................................... 1

Methods ................................................................................................................................. 4

- Study Area ....................................................................................................................... 4
- Field Sampling .................................................................................................................. 4
- ALI Reflectance Model ..................................................................................................... 5
- GIS Watershed Model ....................................................................................................... 6
  - Watershed Delineation ................................................................................................. 6
  - Land Cover ....................................................................................................................... 7
  - Hydrologic DOC Enhancement ..................................................................................... 8
  - Complete GIS Model .................................................................................................... 9
  - Model Comparison ........................................................................................................ 10

Results .................................................................................................................................. 11

- ALI Reflectance Model ..................................................................................................... 11
- GIS Watershed Model ....................................................................................................... 11
- Model Comparison ............................................................................................................ 12

Discussion ............................................................................................................................ 13

- ALI Reflectance Model ..................................................................................................... 13
- GIS Watershed Model ....................................................................................................... 14
  - Vegetation Classification .............................................................................................. 14
  - Low Slope ....................................................................................................................... 14
  - Complete GIS Model .................................................................................................... 14

- Model Comparison ............................................................................................................ 15
- Conclusion ......................................................................................................................... 16

References ............................................................................................................................. 22
List of Figures

Figure 1. Location map and comparison of land cover classification schemes. .................................................................17

Figure 2. Regression of ALI blue band minimum reflectance of lake water against DOC and CDOM .................................................................18

Figure 3. Regression of CDOM against multiple variables. .........................19

Figure 4. Regression of DOC against watershed variables and the complete model .................................................................................................................20

Figure 5. Correlation of independent model results ........................................21

Appendix 1. Regression statistics for unused watershed variables against DOC ...........................................................................................................23
Acknowledgements

Funds for this research provided by Miami University and the National Science Foundation DEB 0639901 grant were greatly appreciated.

I would like to thank my advisor, Robbyn Abbitt, for having an open door policy and fixing all my GIS issues. I also would like to thank my boss, Craig Williamson, for this research opportunity. Further, I would like to thank Bill Renwick, Kevin Rose, and Mary Henry for all their assistance with this project.

In addition, I appreciate all of the help with field sampling from Tiffany Tisler along with Jasmine Saros and her field crew from The University of Maine (Misa Saros, Erin Wilcox, Chelsea Lucas, and William Gray).

Finally, I would like to thank the Miami University Geography graduate students for promoting a healthy learning environment with a barbecue grill.
Introduction

Dissolved organic carbon (DOC) plays a fundamental role in aquatic ecosystem structure and function (Carpenter et al. 1998, Williamson 1999, Cole et al. 2006, Pace et al. 2007). Because a large portion of aquatic DOC is terrestrially derived, DOC concentrations provide valuable information about the surrounding terrestrial landscape as well as atmospheric processes (Williamson et al. in press). In several regions of North America and Europe long-term trends of increasing DOC concentrations have been reported (Striegl et al. 2005, Evans et al. 2006, Monteith et al. 2007). In contrast, decreases in riverine export of DOC have been observed in other regions such as the Yukon River Basin (Striegl et al. 2005).

Attenuation of ultraviolet and visible light in small oligotrophic to mesotrophic glacial lakes is heavily dependent on chromophoric dissolved organic matter (CDOM), the colored component of DOC (Morris et al. 1995; Fee et al. 1996; Gunn et al. 2001), although in high elevation low DOC lakes (< 0.3 mg/L DOC) phytoplankton may play a major role in attenuating ultraviolet radiation (Laurion et al. 2000; Sommaruga and Augustin 2006). In most temperate and boreal landscapes the concentrations of DOC in inland waters are regulated by a wide variety of watershed characteristics including the quantity and type of vegetation, watershed slope, and particularly the extent and nature of wetlands (Engstrom 1987; David and Vance 1991; Frost et al. 2006; Rae et al. 2001; Rice 2002; Williamson et al. 2001; Canham et al. 2002; Xenopoulos et al. 2003). No single set of watershed characteristics can be used to predict DOC concentrations across broad geographic regions (Xenopoulos et al. 2003). DOC concentrations can also be influenced by temperature fluctuations (Cooper, Thoss and Watson 2007; Harrison et al. 2008; Hudson, Dillon and Somers 2003; Striegl et al. 2005) and changes in atmospheric SO$_4$ deposition (DeWit et al. 2007; Evans et al. 2006; Evans et al. 2005; Monteith et al. 2007). Hydrologic characteristics in turn determine how much of that carbon will be exported to downstream sites (Boyer et al. 1997a; Boyer et al. 1997b; Inamdar, Christopher and Mitchell 2004; Ogawa et al. 2006; Worral et al. 2002). Remote sensing and GIS provide valuable tools to assess these watershed variables across the landscape to help us better understand current regional trends of complex and changing DOC concentrations.

The two most fundamental processes leading to DOC accumulation within a watershed are carbon fixation by vegetation (primarily terrestrial in most cases) and slow decomposition of dead organic carbon leading to the accumulation of DOC rather than remineralization of fixed
carbon to CO$_2$. In lower elevation landscapes most areas of a given watershed will contribute to both carbon fixation and DOC generation to different degrees. Alternatively, high elevation watersheds are generally comprised of regions with strongly contrasting roles in terms of their potential contributions to organic carbon fixation and DOC generation. For example, a substantial portion of alpine landscapes are often covered in rock and ice. In addition, steep, well-drained, and well-aerated slopes provide conditions that accelerate decomposition and generate minimal DOC, while low-slope regions with inundated soils and low flow rates create conditions that combine with low temperatures to slow decomposition rates and favor DOC generation. Alpine watersheds thus lend themselves well to quantification of contrasting landscape types with GIS and remote sensing tools that can estimate potential contributions to DOC pools and fluxes.

In lower elevation well-forested regions such as the Adirondacks of New York State spatially explicit models have identified wetlands, forests, roads, flow accumulation, and flow-path distances as important predictors of DOC in lakes with a fair degree of accuracy ($R^2 = 0.546$) (Canham et al. 2004). Inclusion of “cryptic wetlands”, regions of the landscape with low slope that may have relatively inundated soils but no surficially visible wetland habitat, improved prediction of DOC export in the Algoma Highlands of central Ontario ($R^2 = 0.85-0.88$) (Creed et al. 2004). Here we incorporate the concept of cryptic wetlands to assess the ability of two major landscape categories to predict DOC concentrations in high elevation lakes in the Absaroka-Beartooth Wilderness region of Montana-Wyoming, USA. We found that DOC concentrations in this region of highly contrasting landscapes can be predicted with a high degree of accuracy with just two major and often overlapping landscape categories: (1) regions that generate fixed carbon, defined as all areas with any type of vegetation, and (2) regions likely to generate DOC from this fixed carbon, including wetlands and cryptic wetlands defined as all areas with low slope (0-5%).

A variety of researchers have used remotely sensed data to assess variations in water color (Hirtle and Rencz 2003; Nelson et al 2003; Kutser et al 2005a; Kutser et al 2005b; Witte et al 1982). This is important to our research because highly oligotrophic or dystrophic lakes display differences in water color (CDOM) as a function of DOC concentration. Commonly available and widely used satellite imagery, such as Landsat data, is sufficient only to identify broadly classified differences in lake water color (Hirtle and Rencz 2003; Kutser et al 2005a;
Nelson et al 2003). Kutser et al (2005a) compared remote sensing data sources with CDOM measurements in lakes of Sweden and Finland. Data from the Advanced Land Imager (ALI) sensor onboard the EO-1 satellite provided the best measurement of CDOM because of the 16-bit radiometric resolution. Less sensitive data sources, such as Landsat (8-bit) and Ikonos (11-bit), introduced noise into relationships with increasing DOC concentrations. For the current study ALI data were obtained from the USGS EROS data center.

Here we compare the predicted DOC concentrations of remote sensing of lake water color using ALI data with the watershed-based GIS model in order to determine success of independent predictive methods. This study is part of a larger effort to understand nitrogen deposition, climate change and DOC impacts in the Absaroka-Beartooth Wilderness in Montana and Wyoming (Saros et al 2003; Doyle, Saros and Williamson 2005; Cooke et al 2006).
Methods

Our research methods involve creation of two independent predictive DOC models. The first model relates remotely sensed water color to DOC and CDOM. Due to lake size and shape limitations, the 19 study lakes are reduced to 15 usable lakes. The model is applied to 58 Absaroka-Beartooth Wilderness lakes that are large enough for measuring lake water color. The second model includes measuring watershed variables in relation to DOC and CDOM. All 19 study lakes are usable within this model. This watershed model is applied to 354 Absaroka-Beartooth Wilderness lakes. Remaining lakes were eliminated due to watershed errors within the DEM. Finally, we compared output DOC concentrations from 58 lakes common to both models.

Study Area

The Absaroka-Beartooth Wilderness is located just northeast of Yellowstone National Park and is part of the Gallatin and Custer National Forests (Figure 1A). The study area, a small (700 km$^2$) section of the Absaroka-Beartooth Wilderness, is home to >2,000 lakes according to the 1:24000 USGS National Hydrographic Dataset (2007). Elevation within the study area ranges from 1900 to 3900 meters above sea level with the tree line at about 3100 meters (USGS National Elevation Dataset 2007). Land cover ranges from bare rock and perennial ice/snow to coniferous forests at lower elevations. The ice-free growing season varies depending on elevation and topography, but is generally limited to the summer months. Runoff within the study site is typically generated by summer snowmelt and controlled by the highly variable topography. Overall, the topographic ruggedness and elevation of the Beartooth Plateau provide a study area with low human impact and oligotrophic lakes having generally low DOC concentrations (< 300 mg/L).

Field Sampling

We performed in situ sampling of lake DOC concentrations in 20 lakes in the Absaroka-Beartooth Wilderness from July 2 to July 13, 2007. We selected lakes to represent a wide range of DOC concentrations and also by their proximity to accessible field sites that are under ongoing study. We retrieved water samples from within the mixed layer, between 0.5-3m deep.
within the pelagic region from a small rubber raft when possible or otherwise from the littoral area of the lake. We then filtered water samples through a 0.7 µm Whatman GF/F filter and stored them in 40ml glass bottles for DOC concentration measurement. All samples remained cold and dark prior to shipment to Miami University where they were analyzed two weeks post-sampling. Final analyses of the 20 lake samples identified obvious contamination of the Ouzel Lake sample, which was removed from the set.

**ALI Reflectance Model**

We obtained ALI data from the USGS EROS data center comprising an image captured July 2, 2008 at 17:48 Greenwich Mean Time (10:48 Mountain Standard Time) with 0% cloud cover over the study area. Advanced Land Imager (ALI) has a spatial resolution of 30 meters, limiting the size of usable study lakes. Shoreline interference of lake pixels introduces error into lake water reflectance by increasing reflected light to the sensor (Lodhi and Rundquist 2001; Nelson et al 2002). Identifying lake pelagic regions is difficult because bathymetric data are not available for all study lakes. We removed the outermost lake border pixels in order to reduce shoreline reflectance error. Elimination of the two outermost border pixels provided the best correlation of CDOM absorbance with ALI reflectance in bands within the visible spectrum. Four sample study lakes were eliminated by border reduction leaving 15 study lakes for building a statistical model. The remaining >2,000 lakes within the Absaroka-Beartooth Wilderness study area were also reduced by border pixel reduction in order to apply the predictive statistical model. Of the >2,000 lakes, 58 were large enough for application of DOC and CDOM predictions.

Regression of lake pixel values within visible bands against CDOM spectral absorbance and DOC concentrations provided the best relationship using the blue ALI band (450-515 nm). Initially, we calculated the mean of pixel values within the “two-border defined” pelagic region of the lake. However, pelagic regions within lakes are difficult to identify from lake shoreline reduction because of the common irregular shape of Absaroka-Beartooth Wilderness lakes. We then calculated minimum reflectance to avoid assumptions of centrally distributed lake pelagic regions. Minimum reflectance values rely on single pixel values to define the darkest section of the lake. Though single pixels are not as reliable as multiple pixel averages, this method is most
applicable in defining pelagic regions for the small, irregularly shaped lakes of the Absaroka-Beartooth Wilderness.

Reflectance values in relation to water color have a nonlinear trend because of the relationship of light attenuation and depth in the water column (Williamson and Zagarese 2003). Attenuation of light exponentially increases as a function of CDOM with depth in the water column. Therefore, lower CDOM (DOC) allows deeper penetration of light waves before reflectance to the satellite sensor. Since light attenuation has a nonlinear relationship to depth in the water column reflectance must also show a nonlinear trend. We expressed relationships between the log transform of DOC concentrations and CDOM spectral absorbance with extracted reflectance values with linear regression models using SPSS statistical software. We then applied the best-fit regression models to other lakes to predict DOC concentrations and CDOM spectral absorbance.

GIS Watershed Model

Watershed Delineation

We delineated watersheds for study area lakes using ESRI ArcGIS 9.2 (ArcInfo) software and USGS 1/3 arc second (10 meter) and 1 arc second (30 meter) National Elevation Dataset digital elevation models (DEM). ArcHydro Tools, an extension for the ArcGIS environment, delineated watersheds in the study area. Initially, we used the 10 and 30 meter DEMs, but the 30 meter DEM incorrectly delineated some of the 19 study lake watersheds. For this research, all NHD lakes within the study area boundary were input for watershed delineation using the 10 meter DEM. Output data included 460 watersheds out of >2,000 original input NHD lakes. The failure to produce watersheds for many lakes is likely due to the accuracy of DEM data considering watershed size and location of NHD lakes in relation to accumulating runoff sites. Further quality assessment of output watersheds resulted in 353 usable watersheds. Criteria for removal through quality assessment included sites where watersheds were smaller than lakes and obviously misshapen watersheds caused by errors within the DEM.
Land Cover

Fixation of carbon to soils depends on vegetation cover, type, and location (Bukaveckas and Robbins-Forbes 2000; Frost et al 2006; Johnson et al 2006; Rae et al 2001). Landscape vegetation data are a common output of remotely sensed data. Unfortunately, publicly available National Land Cover Dataset (NLCD) data are insufficient for identifying sources of terrestrial carbon for this research. The most recent NLCD was produced in 2001 by the USEPA and metadata reported no independent accuracy assessment on land cover data in NLCD region three where the Absaroka-Beartooth Wilderness is located. Our independent accuracy assessment of 2001 NLCD data within the study area resulted in an overall accuracy of 32.67% in seven classes (Water, Ice/Snow, Rock, Forest, Shrubland, Grassland, and Wetland). To overcome the problem of inaccurate land cover data we produced an alternative land cover map from the ALI data obtained in July 2007.

Unsupervised classification of ALI data provided a sufficient source of vegetation cover data for the study area. Vegetation indices used in land cover map creation include NDVI (Normalized Difference Vegetation Index), TVI (Transformed Vegetation Index), EVI (Enhanced Vegetation Index), and the Tasseled Cap (Kauth-Thomas Transform). NDVI, TVI, and EVI rely on the near infrared band (775-805nm), associated with vegetation biomass, and other visible band reflectance values for calculation of indices that identify vegetation cover. The Tasseled Cap index outputs six calculated bands where the primary bands correspond to brightness (soil brightness), greenness (vegetation), and wetness (moisture content) (Jensen 2005, 313-314). The calculation of additional bands using a combination of spectral reflectance bands can better identify cover types for unsupervised classification especially considering fixation of carbon to soil.

To determine variability between bands for land cover classification, we performed a principal components analysis (PCA) on all ALI bands (minus thermal bands) and vegetation indices. The top six PCA bands accounted for nearly all variability between image bands and vegetation indices. Unsupervised classification of the top six PCA bands into 30 classes resulted in a reclassification output of six cover types (water, grassland, forest, shrubland, rock, perennial ice/snow) (Jensen 2005, 296-301). An independent accuracy assessment of 300 stratified random points in these classes revealed an overall classification accuracy of 72.27%. Therefore, the ALI-derived land cover is more suitable than NLCD for identifying carbon fixation sources.
The NLCD has a liberal identification of grassland in the region where ALI-derived land cover classification is more conservative (Figure 1). Some areas identified as rock in the ALI-derived land cover may have patchy grasses, but it is unlikely that soil beneath these patches is optimal for storing carbon.

We further reclassified ALI-derived land cover by addressing the importance of carbon fixation. We combined areas with rock or ice assumed to not play an important role in DOC production. Vegetation type influences DOC differently (Rae et al. 2001), but the 19 study lake watersheds are not adequately representative of all vegetation classes within the region. Most lake watersheds are limited to early successional vegetation, such as grasses, and the influence of late successional vegetation, such as forest, is not sufficiently representative throughout the 19 study watersheds. In order to facilitate statistical vegetation data in this model, we combined all vegetation classes into a single class. After combination of vegetation classes, our independent accuracy assessment improved to an overall classification accuracy of 82.42%. This further increases the statistical strength in predicting DOC by limiting model inputs and reducing multiple collinearity from a limited number of field samples.

**Hydrologic DOC Enhancement**

Hydrologic processes, as functions of watershed geomorphology, control how much and how often stored carbon is exported from soils. Snowmelt-dominated watersheds, for example, depend on increased runoff produced asynchronously for the movement of DOC from terrestrial sources to downstream sites (Boyer et al. 2000; Boyer et al. 1997). Ultimately, the intensity of runoff events, induced by snowmelt and precipitation, determines export to downstream sites (Boyer et al. 1997; Boyer et al. 2000; Cooper, Thoss and Watson 2007; Inamdar, Christopher and Mitchell 2004; Pace and Cole 2002; Striegl et al. 2005; Worral et al. 2002). In this study, field sampling occurred in July 2007 and, therefore, avoided peak snowmelt times in early summer where major differences in runoff would have occurred. For this model, we assume runoff is synchronous throughout the study area.

Researchers identify wetlands as hydrologically critical in enhancing downstream DOC concentrations (Bukaveckas and Robbins-Forbes 2000; Creed et al. 2003; David and Vance 1991; Freeman et al. 2001; Worral et al. 2002). In addition to above-ground wetlands, topographic characterization identifies sites of hydrologic importance to DOC export (Creed et al. 2003;
Ogawa et al 2006). Available wetland data for the Absaroka-Beartooth Wilderness region, which identify only above-ground wetlands, may be incomplete in identifying sites of DOC enhancement. Available wetland data sources involve the 1:24000 National Hydrography Dataset wetlands, which are drawn from aerial photography and the National Land Cover Dataset (NLCD), which is based on Landsat satellite imagery. Additional wetland data from the National Wetlands Inventory (NWI) are not available within the Absaroka-Beartooth Wilderness region. Commonly, these datasets are insufficient in identifying wetlands beneath the canopy and, most important, insufficient in identifying topographically important sites of DOC export.

Creed et al (2003) identified sites of enhanced DOC using topographic watershed characterization. Inundated soils below the surface create anoxic conditions, which are optimal for DOC enhancement (Creed et al 2003; Ogawa et al 2006; Worral et al 2002). Further, areas with low slope may enhance DOC export by flushing upper soil levels with rising groundwater. Additional flushing of DOC comes from subsurface flow, which can reach deep into soil layers. We used similar topographic characterization of a digital elevation model in order to address hydrologic DOC enhancement. We calculated proportionate slope data for the study area using the same 10 meter DEM obtained for watershed delineation. We determined optimal DOC enhancement zones by calculating median percent slope values within the NHD wetlands, which resulted in 5% slope. We then reclassified the percent slope grid to identify only areas ranging from 0-5% slope. We further limited low slope areas by vegetation cover. In other words, we removed areas with low slope coinciding with rock or perennial ice/snow cover from the low slope grid. Rock outcroppings and perennial snow/ice have no soil structure beneath and, therefore, are incapable of storing significant amounts of DOC for future export regardless of slope. Final analysis involved the linear regression of proportionate 0-5% slope areas within lake watersheds against lake DOC concentrations.

Complete GIS Model

Two major types of controls of lake DOC concentrations occur within lake watersheds: vegetation and low slope. Vegetation cover represents the sites of organic carbon production while low slope areas represent sites of transformation of fixed organic carbon to DOC. This latter variable includes both conspicuous and cryptic wetlands. We selected these input variables by predictive power among many landscape variables (Appendix 1). The predictive DOC model
uses multiple regression of proportionate lake watershed vegetation and area with 0-5% vegetated slope within watersheds against field-measured lake DOC concentrations. In addition, we regressed landscape variables against CDOM spectral absorbance at 320nm (Williamson et al 1999), 420nm (Kutser et al 2005), 440nm (Cuthbert and del Giorgio 1992), and 720nm as a control group. CDOM relationships with watershed variables were less significant and, therefore, were not used for prediction. We used the results of the watershed DOC model to predict DOC in the 354 delineated lake watersheds.

**Model Comparison**

Separate DOC models resulted in different numbers of predicted lakes. The resulting 58 predicted DOC lakes from the ALI blue band reflectance model were also present within the 354 output DOC predictions from the GIS watershed model. We compared those 58 lakes with a simple Pearson correlation using SPSS statistical software to determine consistency between predictions.
Results

Development and completion of independent predictive models for lake DOC concentrations resulted in statistical significance in both cases. The GIS watershed model, using 19 field samples, provides the best relationship to lake DOC concentrations whereas the ALI blue band reflectance model, using 15 field samples, is less statistically significant. Final comparison of 58 lake DOC predictions displays a degree of consistency and, therefore, a degree of reliability in both methods.

ALI Reflectance Model

Extracted blue band minimum reflectance values result in greatest similarity between the log transform of 15 lake DOC concentrations and CDOM spectral absorbance at 440nm in comparison to other absorbance wavelengths tested (Figure 2). Relationships between other spectral absorbance wavelengths (320nm, 420nm, and 720nm) had less statistically significant relationships.

GIS Watershed Model

Watershed vegetation, identified using unsupervised classification of ALI data, and 0-5% slope areas, identified using a 10m resolution DEM, provide a very good relationship with DOC and CDOM (Figure 4). Additionally, collinearity statistics of the complete DOC model, concerning two input variables, result in a variance inflation factor of 5.360, which is well below the multiple collinearity threshold of 15 (SPSS 2006).
Model Comparison

Comparison of 58 DOC concentrations predicted by the ALI blue band minimum reflectance model and the GIS landscape variable model results in a similar relationship (Figure 5A). Removal of Hatchet Lake, a statistical outlier, improves the relationship between predicted DOC concentrations (Figure 5B). The scatter plot trend follows the 1:1 line closely in both cases concluding that both lake DOC concentration prediction methods provide similar insight with particular advantages in each method.
Discussion

Our research shows that the strongly contrasting landscape of subalpine and alpine environments provides optimal conditions for identifying watershed variables that influence DOC in lakes via simple and broadly assumptive classes: vegetation presence or absence and 0-5% slope. Compared to similar work by Canham et al (2004) in lower elevation systems, our research provides a higher R² value with fewer independent variables. Additionally, the independent ALI blue band reflectance model provides a comparatively lower R² value but displays some degree of prediction when results are compared to those from the GIS watershed model. Ultimately, comparison of the output results from the two predictive models displays consistency between model predictions.

ALI Reflectance Model

Initially, the small range of DOC concentrations (0.4 – 2.79 mg/L), compared to Kutser et al (2005ab) (6 – 13 mg/L), requires increased sensitivity in order to differentiate water color between lakes. Furthermore, the small number of samples (15) increases the potential influence of single outlying samples. This model ignores influence from residence time and in lake physical processes which drive material distribution in lakes (Ambrosetti, Barbanti and Sala 2003). For example, spectral absorbance (CDOM) is primarily a function of DOC concentration but relationships decrease with instances of photodegradation (Pace and Cole 2002; Williamson and Zagarese 2003) and variability in terrestrial origin of DOC (David and Vance 1991; Rae et al 2001). For this study, impact of photodegradation was minimized with a two-week sampling period (Kutser et al 2005a). Finally, reliance on single pixels for minimum reflectance values is less desirable than averages of numerous cell values, but the size of lakes in the study area inhibits the use of multiple pixels in defining lake pelagic regions. Nonetheless, the output model has considerable predictive power for application to 58 Absaroka-Beartooth Wilderness lakes; many more than conventional field sampling is able to provide.
GIS Watershed Model

Vegetation Classification

Land cover as vegetated or unvegetated classes significantly correlates with lake DOC concentrations and may be attributed to the strongly contrasting land cover within subalpine and alpine lake watersheds. ALI-derived land cover shows that late successional vegetation, such as forested areas, is not well represented throughout study area watersheds. Therefore, late successional vegetation, alone, cannot be included in our model. Consolidation of all vegetation classes into one class is less accurate, but increases the strength of the relationship with DOC in this highly variable landscape. The relationship reveals a possible nonlinear trend, but these values are essentially grouped into opposing categories of high and low DOC concentrations, and a linear trend is most applicable (Figure 4A). If a nonlinear trend is present, more sample points with well represented distribution are necessary in order to describe it. Ultimately, these groups represent the contrasting extremes of vegetation cover in subalpine and alpine environments.

Low Slope

Flat areas in watersheds provided a better relationship with lake DOC concentrations than vegetation. Functionally, the opposing groups of high and low values seen in the correlation of watershed vegetation and lake DOC concentrations are less apparent. A large number of low DOC concentrations are clustered within the model but remain limited in variation from the trend (Figure 5B). These data suggest that 0-5% slope areas do control DOC export within watersheds. Above-ground wetlands, identified by visual remote sensing, show a similar relationship, but are not are not well represented in the 19 study watersheds (Appendix 1). Ground referencing is necessary to identify the hydrologic importance of 0-5% slope sites (i.e. canopy-hidden wetlands or subsurface wetlands).

Complete GIS Model

Multiple regression of proportionate vegetation and low slope in 19 lake watersheds results in a very strong relationship with lake DOC concentrations. The final scatter plot of actual DOC and landscape predicted DOC displays some clustering at lower DOC concentrations but less than 0.5 mg/L variation from the trend (Figure 4C). The greatest limitation to this model
is the number of field samples, but a reduction of model inputs to two (vegetation and low slope), limits reduction of the adjusted $R^2$ resulting in a value of 0.91.

A more detailed and spatially explicit model may be needed in order to explain more variation in DOC concentrations among Absaroka-Beartooth Wilderness Lakes. For example, the sampled lakes with the two highest DOC concentrations have substantially different watersheds and both lakes vary similarly on opposite sides of the trend (Figure 4C). Kersey Lake has ~50% of its watershed covered in woody vegetation and is predicted low by the GIS model while the Chain Lakes are mostly grassland and predicted high within the GIS model. These values are consistent with previous research on differences in lake transparency as a function of grasslands versus forests (Rae et al 2001). Therefore, a more detailed model of Absaroka-Beartooth Wilderness lakes should include a comparison of vegetation type within low slope areas. A spatially explicit model may also include higher-resolution spatial data. In some cases, lake watersheds were represented with less than 100 30m x 30m pixels. Therefore, higher-resolution spatial data, detailed vegetation data, and more physically sampled lakes with late successional vegetation are necessary for more complete representation.

**Model Comparison**

Comparison of lake DOC predictions in the Absaroka-Beartooth Wilderness from completely independent sources indicates that these methods are consistent and reliable in subalpine and alpine ecosystems. General error potentially derives from the ALI blue band minimum reflectance model producing a majority of its scatter in higher DOC concentrations. However, all error cannot be attributed to the statistically weaker ALI blue band reflectance model. For example, an extreme outlier in the model, Hatchet Lake, is most likely an error within the GIS watershed model (Figure 5). This watershed has a high proportion of vegetation and 0-5% slope to watershed area but what is not accounted for in the model is lake size in relation to watershed size. In most cases, Absaroka-Beartooth Wilderness lakes make up a minimal proportion of their watersheds, while Hatchet Lake comprises nearly $\frac{1}{2}$ of its watershed (Figure 5C). This implies that any amount of DOC reaching the lake, as measured in concentration, is potentially diluted by the large amount of water in the lake. Lake size in relation to watershed size was not included in the model because it was not a statistically significant predictor of DOC within the 19 field samples ($R^2 = 0.027$).
Conclusion

Our models show strong relationships with lake DOC concentrations involving a limited number of field samples and inputs of simple and broadly assumptive variables. In addition, comparison of independent model predictions results in significant consistency. More detailed geospatial research may result in a better relationship with lake DOC concentrations but may also show that such a contrasting landscape only requires such simple and broadly assumptive inputs.

Results from our research are practical in that they can inform future in situ research within predicted study lakes and apply common geographic principles such as land cover change. Due to the consistency of independent model predictions, individual lakes can be identified for more exhaustive research. For example, if researchers need to study lakes with particularly high or low DOC concentrations then our results can help in identifying appropriate lakes. It is important, though, that assumptions within models, such as the lack of any explicit effects from lake size in relation to watershed size are considered. Variability between lake DOC concentrations may also change over time as a function of land cover change. Vegetation change, potentially due to climate change, invasive species introduction, or anthropogenic nitrogen deposition for examples, has potential for measurement from remote sensing sources. If incorporated into our GIS watershed model, temporal trends in DOC can be estimated. In addition, measurement of vegetation change may be extended into the past as remotely sensed satellite imagery has been available for a few decades. Overall, GIS and remote sensing tools show great potential to inform future DOC research and measure broad-scale changes especially in the strongly contrasting landscape of the Absaroka-Beartooth Wilderness.
Figure 1. Location map and comparison of land cover classification schemes.

Land cover mapping is difficult in areas with high topographic relief. The study area, within the ALI scene in map A, is located in very high topographic relief. Maps B and C display differences in USEPA NLCD and land cover created from ALI data collected for this research. USEPA NLCD data is liberal in identification of vegetation throughout the study area. The more conservative ALI-derived land cover better indicates land cover with the purpose of identifying sources of DOC. Areas of no data exist outside of study watershed boundaries or within areas identified separately due to shadowing from topographic relief.
Figure 2. Regression of ALI blue band minimum reflectance of lake water against DOC and CDOM

ALI minimum blue band reflectance correlates highest with the log transform of lake DOC concentrations. Relationships are similar between plot A and plot B likely due to the commonality between DOC concentration and spectral absorbance at 440 nm light compared to other wavelengths (A320 $R^2 = 0.544$, $p = <0.001$; A420 $R^2 = 0.548$, $p = <0.001$; A720 $R^2 = 0.313$, $p = 0.13$). The relationship of DOC and reflectance values is used to predict lake DOC concentrations.
Figure 3. Regression of CDOM against multiple variables.

Regression statistics for the relationships among spectral absorbance, dissolved organic carbon (DOC), proportion of the watershed in vegetation according to the ALI data (ALIV), and proportion of the watershed in low (0-5%) slope (WSLS).

DOC provides the strongest relationship with spectral absorbance at short wavelengths as is also demonstrated by watershed predictors of DOC (vegetation and low slope).

<table>
<thead>
<tr>
<th>λ (nm)</th>
<th>Slope</th>
<th>Y-intercept</th>
<th>R²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>320</td>
<td>0.132</td>
<td>0.029</td>
<td>0.949</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>420</td>
<td>0.026</td>
<td>0.041</td>
<td>0.886</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>440</td>
<td>0.02</td>
<td>0.041</td>
<td>0.864</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>720</td>
<td>0.002</td>
<td>0.088</td>
<td>0.281</td>
<td>0.02</td>
</tr>
</tbody>
</table>

DOC regression:

\[ R^2 = 0.949 \]
\[ p = <0.001 \]
\[ A320 = \text{DOC} (0.132)+0.029 \]

ALI-Vegetation regression:

\[ R^2 = 0.879 \]
\[ p = <0.001 \]
\[ A320 = \text{ALIV} (0.371)+0.013 \]

0-5% Slope regression:

\[ R^2 = 0.879 \]
\[ p = <0.001 \]
\[ A320 = \text{WSLS} (1.769)+0.095 \]
Figure 4. Regression of DOC against watershed variables and the complete model.

The complete GIS watershed variable model (C) of lake DOC concentrations includes inputs of proportionate vegetation (A) and areas with low slope (B) within each watershed. Both methods are independently good predictors of lake DOC concentrations but when combined represent introduction and enhancement of upstream DOC. Ultimately, the final watershed model shows a strong prediction of lake DOC concentrations. The final equation in chart C includes: 

\[ c = a(0.971)+b(9.161)+0.276 \]

Clustering in low DOC concentrations still results in limited variation from the trend.
Figure 5. Correlation of independent model results.

Correlation of 58 output results from both the GIS watershed model and ALI blue band reflectance model produced similar results (A). Most of the predicted DOC concentration variability is within higher concentrations likely resulting from the errors within the ALI-reflectance model. The major outlier, Hatchet Lake, has mismatched values potentially deriving from the GIS watershed model (B). This model did not include lake size in comparison to watershed size and, therefore, dilution of DOC may be a factor (C).
References


SPSS 15.0 for Windows. (2006). Copyright SPSS Inc.


Appendix 1. Regression statistics for unused watershed variables against DOC. These independent variables were not included in the GIS watershed model, but were initially calculated. All variables were calculated as proportions of lake watersheds and regressed against lake DOC concentrations.

<table>
<thead>
<tr>
<th></th>
<th>Slope</th>
<th>Y-intercept</th>
<th>R²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a. NLCD Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLCD Wetland</td>
<td>69.54</td>
<td>0.886</td>
<td>0.495</td>
<td>0.001</td>
</tr>
<tr>
<td>NLCD Grass</td>
<td>-3.358</td>
<td>2.516</td>
<td>0.711</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NLCD Shrub</td>
<td>2.73</td>
<td>0.119</td>
<td>0.433</td>
<td>0.002</td>
</tr>
<tr>
<td>NLCD Forest</td>
<td>4.079</td>
<td>0.727</td>
<td>0.411</td>
<td>0.003</td>
</tr>
<tr>
<td>NLCD Rock</td>
<td>-9.214</td>
<td>1.605</td>
<td>0.508</td>
<td>0.001</td>
</tr>
<tr>
<td>NLCD Woody Vegetation</td>
<td>3.274</td>
<td>-0.344</td>
<td>0.85</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NLCD Impervious Surface</td>
<td>14.846</td>
<td>0.983</td>
<td>0.124</td>
<td>0.139</td>
</tr>
<tr>
<td>NLCD Ice and Snow</td>
<td>-25.442</td>
<td>1.358</td>
<td>0.268</td>
<td>0.023</td>
</tr>
<tr>
<td>NLCD Total Vegetation</td>
<td>4.856</td>
<td>-3.121</td>
<td>0.232</td>
<td>0.037</td>
</tr>
<tr>
<td><strong>b. ALI Land Cover Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALI Grass</td>
<td>2.302</td>
<td>0.423</td>
<td>0.469</td>
<td>0.001</td>
</tr>
<tr>
<td>ALI Shrub</td>
<td>-0.321</td>
<td>1.091</td>
<td>0</td>
<td>0.943</td>
</tr>
<tr>
<td>ALI Forest</td>
<td>4.246</td>
<td>0.787</td>
<td>0.459</td>
<td>0.001</td>
</tr>
<tr>
<td>ALI Rock and Snow</td>
<td>-2.54</td>
<td>2.346</td>
<td>0.764</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ALI Total Vegetation</td>
<td>2.704</td>
<td>-0.066</td>
<td>0.841</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*a.* Proportion of NLCD defined land cover types in lake watersheds  
*b.* Proportion of ALI classified land cover types in lake watersheds
### Appendix 1. (Continued)

<table>
<thead>
<tr>
<th>Slope</th>
<th>Y-intercept</th>
<th>$R^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>c. Low Slope in Proximity to Lakes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30m Slope 500m</td>
<td>6.259</td>
<td>0.559</td>
<td>0.752</td>
</tr>
<tr>
<td>30m Slope 1500m</td>
<td>12.098</td>
<td>0.52</td>
<td>0.886</td>
</tr>
<tr>
<td>30m Slope 2000m</td>
<td>7.967</td>
<td>0.551</td>
<td>0.877</td>
</tr>
<tr>
<td>30m Slope Total Watershed</td>
<td>12.138</td>
<td>0.516</td>
<td>0.886</td>
</tr>
<tr>
<td>10m Slope 500m</td>
<td>7.091</td>
<td>0.541</td>
<td>0.806</td>
</tr>
<tr>
<td>10m Slope 1500m</td>
<td>8.17</td>
<td>0.536</td>
<td>0.876</td>
</tr>
<tr>
<td>10m Slope 2000m</td>
<td>8.771</td>
<td>0.541</td>
<td>0.888</td>
</tr>
<tr>
<td>10m Slope Total Watershed</td>
<td>13.334</td>
<td>0.507</td>
<td>0.9</td>
</tr>
</tbody>
</table>

**Miscellaneous Variables**

<table>
<thead>
<tr>
<th></th>
<th>Slope</th>
<th>Y-intercept</th>
<th>$R^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>d. Lake Size/ Wshed Size</strong></td>
<td>-2.451</td>
<td>1.173</td>
<td>0.027</td>
<td>0.497</td>
</tr>
<tr>
<td><strong>e. NHD Wetland</strong></td>
<td>51.678</td>
<td>0.675</td>
<td>0.913</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>f. Slope &gt;45°</strong></td>
<td>-0.625</td>
<td>1.16</td>
<td>0.074</td>
<td>0.259</td>
</tr>
<tr>
<td><strong>g. Wshed Perimeter/Area</strong></td>
<td>-27.805</td>
<td>1.694</td>
<td>0.094</td>
<td>0.203</td>
</tr>
<tr>
<td><strong>h. Watershed Lakes</strong></td>
<td>6.303</td>
<td>0.937</td>
<td>0.08</td>
<td>0.242</td>
</tr>
<tr>
<td><strong>i. Shaded</strong></td>
<td>-3.379</td>
<td>1.438</td>
<td>0.269</td>
<td>0.023</td>
</tr>
</tbody>
</table>

- **c.** Proportion of defined distance uphill from lake (500m, 1500m, 2000, total watershed) with 0-5% slope. Represented for both 10m and 30m DEMs.
- **d.** Lake area divided by watershed area
- **e.** Proportion of lake watersheds with NHD defined wetlands
- **f.** Proportion of lake watershed with slope > 45°
- **g.** Watershed perimeter length divided by watershed area
- **h.** Total lake area within individual lake watersheds
- **i.** Proportion of watershed mostly shaded from direct sunlight