Potomac Watershed Priority Lands Strategy:
Conserving lands to benefit drinking water quality

By

Emily Weidner
Dean Urban, Advisor

Master’s Project submitted in partial fulfillment of the requirements for the Master of Environmental Management degree in the Nicholas School of the Environment of Duke University

May 2009
# Table of Contents

List of Tables and Figures ........................................ ii

Abstract ................................................................ 1

Introduction .......................................................... 2

Study Area: The Potomac Watershed ......................... 5

Water Quality Model: Assessing the effect of land use
   Approach .......................................................... 5
   Data .................................................................. 7
   Landscape Characteristics ..................................... 9
   Statistical Analysis ............................................. 13
   Regression Model & Water Quality Model ............... 14

Multi-Criteria Prioritization Model
   Approach .......................................................... 16
   Intake Importance Areas ..................................... 17
   Development Vulnerability ................................... 19
   Ecological Value Areas ....................................... 19
   Final Model as Decision-Support Tool ................... 19

Conclusions & Recommendations .............................. 20

Acknowledgements .................................................. 23

References ................................................................ 24

Tables and Figures .................................................. 27
List of Tables & Figures

Table 1. C-factor values for different NLCD land cover types. 27

Table 2. Correlations between forest variables. 27

Table 3. Correlations between urban variables. 27

Table 4. R-squared results from a series of regressions. 28

Table 5. Water quality regression results. 28

Table 6. Elasticities for final water quality regression. 28

Figure 1. Land use in the Potomac watershed. 29

Figure 2. MBSS monitoring points and physiographic regions. 30

Figure 3. Inverse distance weighted schematic. 31

Figure 4. Buffer capacity schematic. 31

Figure 5. Soil erosion model. 32

Figure 6. Water quality model. 33

Figure 7. Water quality model, close-up. 34

Figure 8. Contaminant distance-decay function. 35

Figure 9. Intake priority area, all intakes equal. 36

Figure 10. Intake priority area, intakes weighted by population served. 37

Figure 11. Intakes and their overlapping catchments 38

Figure 12. Chesapeake Bay Program’s development vulnerability model. 39

Figure 13. Chesapeake Bay Program’s “hubs and corridors” ecological model. 40

Figure 14. Final conservation priority model. 41

Figure 15. Final conservation priority model, close-up. 42

Figure 16. Conservation priority model with current protected areas. 43

Figure 17. Conservation priority model, available forest land for conservation. 44
Abstract

Rapid development in the Potomac watershed — the conversion of forests to agricultural, suburban, and urban land — threatens water quality. Similarly, strategic land conservation can protect water quality. Inspired by examples of water purification through land conservation, the EPA Region 3 and the Potomac River Basin Source Water Protection Partnership (PRBSWPP) aim to prioritize areas of the Potomac watershed for conservation. To work toward this goal, I analyzed two questions: (1) How do land areas in the Potomac Watershed support stream water quality?; and (2) In a larger context, how should land conservation be prioritized to protect water quality?

To assess the relationship between land use and water quality, I created a regression model to correlate land characteristics including land use composition, land use pattern, and hydrological connectivity, with water quality. The final regression shows that buffer capacity (i.e. the average percentage of downstream forest area) has the largest impact on water quality, followed by urban saturation (i.e. average percentage of downstream urban area), and two estimates of soil loss and erodibility. I mapped the output of this regression analysis.

To identify priority lands for conservation, I developed a multi-object decision analysis (MODA) tool. I used the weighted averaging approach to combine a land parcel’s water quality protection value, water intake protection value, and ecological value, along with its vulnerability to future development. This resulted in a map showing areas of higher and lower conservation priority, which can be used to allocate funds for conservation, update local zoning to designate strategically located natural areas, assist developers in minimizing their environmental impact, and strengthen coalitions in developing a common understanding of the multiple benefits of land conservation.
I. Introduction

The Potomac watershed is home to increasingly widespread development and fragmentation of its landscape, and new and intensifying contamination of its waterways. This type of land use change, along with the fragmentation that goes with it, threatens both aquatic health and clean drinking water that the Potomac’s growing population depends on.

Natural area conservation has long been assumed to play an important role in safeguarding water resources and providing supplies of clean water. For example, riparian buffers are frequently cited for their ability to filter contaminants before reaching the waterways. Similarly, it is firmly established that development and land use changes cause changes in the water quality and related aquatic habitats. Agricultural areas are associated with nitrogen, pesticides, and increases in erosion and sedimentation, while urban areas are associated with volatile organic compounds and other human-made chemicals. Pasture land has been associated with high levels of pathogens such as cryptosporidium and e. coli. With rain these contaminants often bind to sediments and are carried by overland flow through the landscape to the rivers or they infiltrate through the soil to the groundwater. Increased levels of contaminants in waterways has lead to fish kills, restrictions on the recreational use of public water bodies and waterways, and even collapse of aquatic communities.

Drinking water is also affected by this process. The contaminants that reach the rivers threaten the safety of our drinking water by making it more difficult for source water treatment companies to easily and effectively treat river water. If water treatment companies are not prepared for the type or quantity of contaminant, residents may be faced with lowered drinking water quality or even temporary lack of service. Over time, continued increases in contamination necessitates newer and more sophisticated filtration technology which costs the water treatment companies millions, a cost that is likely passed on to the local residents.
Historically, a waterworks engineering approach has been employed to deal with increases in contamination of the intake water. The fundamental assumption of this approach is that as new contaminants are introduced into the waterways, people are able to devise new treatments to maintain high drinking water standards. Conversely, a land conservation approach is an alternate approach to maintaining water quality that acknowledges land use’s profound effect on water quality. While the engineering approach seeks to remove contaminants from drinking water, the land conservation approach seeks to avoid contaminants getting to the waterway by harnessing the natural filtration capabilities of forests and wetlands.

Though most drinking water providers still use the engineering approach, New York City has been the posterchild of the land conservation approach. With the 1989 promulgation of the Surface Water Treatment Rule, New York City was faced with a decision to pay billions for a new filtration plant or hundreds of millions in watershed protection and conservation. Needless to say, New York City chose to forgo the filtration project and instead rely on watershed protection to maintain high quality drinking water for New York City residents. To date it has protected more than 375,000 acres through purchase and conservation easements, representing more than 35% of the watershed (NYC-DEP, 2007; and Murphy 1995) and it remains in compliance with the drinking water quality standards put forth in the Surface Water Treatment Rule.

Inspired by the success of this program, source water advocates in other regions are exploring conservation as a viable method for protecting drinking water quality. In 2007, the Schuylkill Action Network (SAN) Land Protection Collaborative produced the Schuylkill Watershed Priority Lands Strategy. This strategy, a joint effort by the Natural Lands Trust, Philadelphia Water Department, and the Delaware Valley Regional Planning Commission, identified areas within the Schuylkill Watershed that are the most important to preserve for both ecological and drinking water source protection. The idea was that this could be used “to direct inappropriate uses away from high priority resource areas” to promote a healthy aquatic ecosystem and clean drinking water (Schuylkill Watershed Priority Lands Strategy, 2009). Regional strategies for land conservation for
drinking water quality such as this are especially important in areas with multiple drinking water providers and competing land uses.

Like the Schuylkill Action Network, advocates for the Potomac Watershed have a similar vision of source water protection, but are challenged by coming to an agreement about which lands should be prioritized for conservation to help achieve this vision. In the Potomac Watershed, there are many advocates for source water protection—dozens of water suppliers, four states’ Department of Natural Resources and Department of Environmental Protection, the Environmental Protection Agency (EPA) Region 3 Drinking Water Branch, and innumerable local land trusts, conservation organizations, and watershed protection groups. But with these multitudes of advocates, having a clear message and single unified plan for conservation is a challenge. The Potomac River Basin Source Water Protection Partnership, a voluntary association of water suppliers and government agencies, seeks to unite those focused on protecting drinking water sources in the Potomac Basin (PRBSWPP, 2009). And with the leadership of EPA Region 3 Drinking Water Branch they have proposed the formation of a Potomac Watershed Priority Lands Strategy to identify lands for priority conservation for drinking water quality. This paper represents this priority lands conservation strategy.

The goal of the Potomac Watershed Priority Lands Strategy is to identify which areas of the Potomac watershed should be prioritized for conservation to benefit drinking water quality. To work toward this goal, this paper is structured around two main questions: (1) How do land areas in the Potomac Watershed contribute to stream water quality? (2) How should land conservation be prioritized for drinking water given land’s contribution to water quality?

After a brief description of the study area, I discuss the approach, methodology, and results of a Water Quality Model that explains visually how land areas in the Potomac Watershed contribute to water quality. Next, I discuss additional criteria considered when prioritizing lands for conservation, and I suggest a Prioritization Model that serves
as a decision-support tool. The paper concludes with general recommendations for the tool’s use.

II. Study Area

The Potomac Watershed spans four states—Pennsylvania, Maryland, Virginia, and West Virginia—plus the District of Columbia. This region is one of the fastest developing areas in the country. Associated land cover changes have resulted in a variety of land cover patterns across the landscape which present an opportunity to investigate the interrelationships between land cover changes and ecological responses.

The mouth of the watershed begins in the eastern lowlands and is part of the Coastal Plains physiographic region. Moving westward, the watershed includes the Piedmont, Blue Ridge Mountains, Valley and Ridge, and Appalachian Plateaus physiographic regions. Most of the population is found in the DC metro area and surrounding suburbs. There is a belt of agricultural and pastoral lands running through the center of the watershed from the northeast to the southwest. The western part of the watershed is largely defined by rural and mainly forested lands. (See Figure 1.)

III. Water Quality Model: Assessing the effect of land use

Approach

In a time of continued land use change and fragmentation, and worsening water quality in the Potomac River basin, it becomes critical to evaluate how changes in land use affect drinking water quality. Only after understanding how land use change and patterns affects this ecosystem service can conservation lands be effectively prioritized.

A weighted layer-based approach created in a geographic information system (GIS) is the conventional approach for creating models, in this case maps, of land areas with water
quality importance. This approach uses two or more component map layers, weights them based on their relative importance, and adds them together, to get a composite model of the complex importance of areas. For example, to get a water quality model that shows land’s contribution to water quality, there may be components such as forest area, grassland area, or impervious surface that could be included in the model.

In some models, weights are determined by the model creator or expert consensus (for example, Schuylkill Watershed Priority Lands Strategy). In others, where data is available, a regression analysis is used to determine the appropriate weights of each of the components. Although using user-defined weights is methodologically simpler, using regression analysis to determine the relative importance of each component gives more accurate and precise estimates of weights. It allows the model to be calibrated to real water quality monitoring data, and takes human biases out of the equation.

It is important to note, however, that the regression approach can only be used where there is data on the variable being modeled. For instance, when modeling water quality, water quality data is needed to use the regression approach. For a variable such as “importance to Y land trust” there is no data giving a numerical value to “importance to Y land trust,” and therefore the user-defined weights would be appropriate to use. For this paper, I use the regression approach in determining the appropriate weights for the land contribution to water quality model. The benthic index of biological integrity (B-IBI) from Maryland Biological Stream Survey data acts as a surrogate for water quality in the analysis.

The literature is rich with studies correlating land use to water quality using regression analysis, however few of them consider how the pattern of land use across the landscape affect aquatic habitat and water quality (for example: Mehaffey, 2005; Strayer, 2003). Among the patterns considered are inverse distance weighted land cover (King, 2005; Baker, 2006), riparian buffer widths (Baker, 2006; Baker, 2001; Jones, 2001), and others (Alberti, 2007). Other studies have focused on identifying the relationship between land use and water quality at multiple scales (Roth, 1996; Silva, 2001; Lammert, 1999).
Having a better understanding of how land use patterns affect water quality can help conservation organizations determine which lands, if conserved, may have more water quality value than others. I hypothesize that in addition to the composition of a landscape, the pattern and hydrological connectivity also play a large role in determining the water quality.

To test this hypothesis, and to create a calibrated and accurate model of land’s contribution to water quality, I first compiled a set of landscape metrics for each of the monitoring point catchments. Then I used regression analysis to determine the relative importance of each of the landscape metrics on water quality. Next, using the outcomes of the regression, landscape metrics were weighted according to their relative importance to water quality to determine which areas of the Potomac Watershed most contribute to good water quality.

Data

Monitoring Data and their Catchments

Water quality data for 564 sites across the Maryland portion of the Potomac watershed were obtained from the Maryland Biological Stream Survey (MBSS). This dataset included benthic and fish index of biological integrity (IBIs), water chemistry, and in-stream habitat data collected between 2000 and 2004. Murcurio et al (1999) provide details of sampling methods. Water chemistry data varies among seasons, days, and even hours, and therefore the benthic IBI was used as a surrogate for water quality. In the event that a site was sampled more than once in the four years, only data from 2002 (or 2001 if no 2002 data existed) was used in this study.

Because the Potomac watershed spans five states, the challenges of finding compatible water quality datasets were immense. Though the MBSS dataset does not evenly cover the Potomac Watershed, it does provide data for a cross-section of the physiographic regions in the watershed (Figure 2). In this way it is representative of the entire
watershed. Additionally, the MBSS dataset has many different types of water quality, has a great density of points, and was collected with consistent measuring techniques.

**Geographic Data Sources**

Land cover data was derived from the 2001 National Land Cover Database (NLCD) (USGS, 2007b). To emphasize where roads fragmented the landscape, local streets data (Tele Atlas North America, Inc. and ESRI, 2003) were overlain on the NLCD layer as “developed, open space” as most of the visible roads in the NLCD were coded. Similarly, river data (USGS, 2007a) were overlain on the NLCD layer to assure flow continuity of all rivers and streams.

Elevation information was obtained from the USGS National Elevation Dataset (NED) Digital Elevation Model (DEM) (USGS, 1999). The streams from the modified NLCD were “burned” into the elevation layer to digitally excavate stream channels. When conducting flow analyses this serves to improve accuracy of flow directions and flow paths. Catchment boundaries for each of the monitoring sites were created using the “watershed” function in ArcGIS using the elevation layer. Watershed size ranged from 10 hectares to 71,377 hectares.

Soils data for the Potomac watershed were derived from the U.S. General Soil Mal (STATSGO2) database from the USDA’s Natural Resources Conservation Service (NRCS, 2006) using the Soil Data Viewer (NRCS, 2007). Soil type varies within small regions; therefore each map unit was comprised of one or more component of soil. The soil type assigned to each map unit is equal to the soil type of the dominant condition within the map unit. The soil data used for this paper represents the surface layer of soil and considers the whole soil and not just the fine-Earth fraction of less than 2 mm. Although using SSURGO soils data would have been preferable because of its finer spatial scale, STATSGO2 was used because, unlike SSURGO, it covered the entire Potomac Watershed.
All data layers were transformed into a consistent raster format with 30 meter square pixels and were projected into North American Datum (NAD) 1983 Albers projection. All landscape metric parameters were created using ArcGIS 9.3 (ESRI, 2009) and the spatial analyst extension.

**Landscape Metrics**

Each of the monitoring catchments in the Potomac Watershed was quantified by using landscape composition, pattern, and hydrologic connectivity metrics. Percent of each land use type within a watershed was used as the most basic measure of landscape composition, while the inverse distance weighted (IDW) technique was used to measure composition by weighting areas closer to the streams more than areas farther away from the streams. I used three measures of buffering capacity that depend on the patterning of the landscape. Two measures, the unconstrained buffer width, and flow path buffer, consider the ability of only riparian forests (forest or wetland areas contiguous to the stream) to buffer, or filter, contaminants and sediments. The other pattern metric, buffer capacity, considers the buffering abilities of all forest or wetland areas in the watershed. Metrics that help estimate soil loss such as soil erodibility, slope, and runoff were used to estimate functional hydrological processes across the landscape. The number of stream-road intersections provided a measure of impervious connectivity to the stream system.

**Percentage of each land cover type in the watershed**

The percent of each land cover type is a straightforward measure of purely composition of the landscape. It is calculated as,

\[ P_{wi} = 100 \times \frac{A_{wi}}{A_w} \]  

(1)

where \( A_w \) is total area of watershed \( w \), \( A_{wi} \) is the area of class \( i \) pixels in watershed \( w \), and \( P_{wi} \) is the percent of class \( i \) land cover in the watershed, \( w \). This was calculated for combined forest and wetland areas, urban areas, agricultural areas, and impervious areas.
Inverse Distance Weighted (IDW) of each land cover type

Inverse distance weighted (IDW) land cover is similar to the simple percent of watershed measure above, but it considers distance to stream as an important weighting factor. That is, the closer a pixel of land is to the stream, the more it will count. First, a binary raster is created that assigns values of 1 to presence of land cover type \( j \), and 0 to presence of all other land cover types. Next, a flow distance to stream raster is created. When the land use binary raster (composed of featured land use type with a value of 1 and all other land use types with a value of 0) is divided by the distance raster, the inverse distance weighted raster is formed (Figure 3). To standardize this measure across many watersheds, the sum of the pixels in this IDW layer is divided by the sum of the pixels in another IDW layer that has a land use weight value of 1 for every cell. The final IDW value for each watershed is represented by,

\[
IDW_i = 100 \times \frac{\sum (f_{ij}/d_j)}{\sum (1/d_j)},
\]

Where \( d_j \) is the distance from the pixel \( j \) to the stream and \( f_{ij} \) is a weight that takes value of 0 when a pixel \( j \) is not in class \( i \) and a value of 1 when a pixel \( j \) is in class \( i \). The distance from the pixel \( j \) to the stream can be calculated in several ways. In this paper I use the flow path distance as opposed to Euclidean distance. Euclidean distance measures the straight line distance from the pixel \( j \) to the nearest stream pixel. Flow path distance measures the distance from the pixel \( j \) along the downstream flow path to the first stream cell that pixel \( j \)’s runoff would encounter. IDW land use was calculated for combined forest and wetland areas, urban areas, agricultural areas, and impervious areas.

Buffer Capacity

The measure of total buffer capacity represents how well each pixel is buffered with respect to the maximum possible buffering of that pixel. The average is calculated for each watershed and is scaled so that values of 100 represent watersheds that are buffered maximally, and 0 represents the absence of any buffering. Figure 4 illustrates the effect of a forest cell being located close or far from the stream and its buffering effect on the watershed. The buffering capacity value for each watershed is represented by,

\[
B = 100 \times \frac{\sum (b_{ij}/t_j)}{N},
\]
where $b_{ij}$ is the number of land use class $i$ pixels in the flow path of pixel $j$, $t_j$ is the total number of pixels in the flow path of pixel $j$, and $N$ is the number of pixels in the watershed. A variable, urban saturation, is the same metric, but uses urban land cover cells instead of forest or wetland cells.

*Unconstrained Buffer Distance*

The unconstrained buffer distance measures the average riparian buffer width. The riparian buffer is defined as any forest that is contiguous to the stream. The width is measured with a Euclidian distance, or straight line distance from the stream to the edge of the riparian buffer. See Baker et al (2006) for detailed methods.

*Flow Path Buffer*

The flow path buffer measures the distance through riparian buffer which runoff from agricultural or urban land runs through to get to the stream (assuming no infiltration). This measures how well urban and agriculture areas are filtered by riparian areas. See Baker et al (2006) for detailed methods.

*Road/Stream Crossings*

Roads have a significant effect on water quality (Trombulak, 2000), and the number of road-stream intersections provides important information about the connectivity of the roads to the water network. Road-stream intersection intensity is measured as,

$$I_w = \frac{R_w}{S_w},$$

where $I_w$ is the road-stream intensity for watershed $w$, $R_w$ is the number of road-stream intersections in a watershed $w$, and $S_w$ is the stream length in a given watershed $w$.

*Soil Erodibility Metrics*

Soil erosion has a significant impact on water quality. Contaminants often bind to larger sediment particles and are washed into the waterways with overland flow from a rainfall event.
The universal Soil-Loss Equation (USLE) (Wischmeier and Smith, 1961) and the Revised Universal Soil Loss Equation (RUSLE) have been commonly used to estimate the average annual rate of soil loss by sheet and rill erosion in small watersheds. The basic equation is made up of six factors as follows below:

\[ A = R \times K \times LS \times CP , \]  

where \( A \) is estimated watershed average soil loss in tons per acre per year, \( R \) is rainfall-runoff erosivity factor, \( K \) is soil erodibility factor, \( LS \) is slope length and steepness factor, and \( CP \) is cover-management and support practice factor.

In this paper, this watershed-scale equation is adapted to a pixel by pixel estimate of soil erosion. When assuming spatially constant rainfall, the equation can be adapted to:

\[ E = K \times S \times C , \]  

where \( E \) is pixel soil erosion index, \( K \) is soil erodibility factor (K-factor), \( S \) is slope factor (S-factor), and \( C \) is land cover runoff factor (C-factor).

The soil erodibility factor, \( K \), is represented by the soil K-factor and indicates the susceptibility of a soil to sheet and rill erosion by water. The estimates are based primarily on percentage of silt, sand, and organic matter and on soil structure and saturated hydraulic conductivity (Ksat). Values of \( K \) range from 0.02 to 0.69. All other factors being equal, the higher the value, the more susceptible the soil is to sheet and rill erosion by water (Michigan State University, 2009).

The slope factor, \( S \), replaces the watershed-scale slope length and steepness factor of the USLE. It is represented by a rescaled measure of slope. Zero values were rescaled to a value of 1.

\[ S_i = \text{rescaling coefficient} (s_i) , \]  

where \( S \) is the slope factor, and \( s_i \) is the slope degree of a given pixel.

The land cover runoff factor, \( C \), like the cover-management and support practice factor (CP) of RUSLE, represents the vegetation’s role in reducing runoff and associated
surface erosion. These values have been adapted from the relative values of runoff curve numbers of different land cover types. The assigned C-factor values are found in Table 1. Developed land has the highest value representing large amounts of runoff, while forest and wetlands have the lowest values representing less runoff and surface erosion.

Each factor was rescaled to be within the 1 to 100 range, and each factor and each combination of factors was calculated for each pixel. Figure 5 illustrates the soil erosion model. In addition to the average value per watershed, the inverse distance weighted was calculated. The method for doing this is similar to the method described earlier for binary data. In this case, the final IDW value for each watershed is represented by,

\[ IDW = 100 \times \frac{\sum (f_j / d_j)}{\sum (m / d_j)}, \]

where \( d_j \) is the distance from the pixel \( j \) to the stream and \( f_{jm} \) is the value of pixel \( j \) from a weight layer with maximum value of \( m \). In this paper, I ran an IDW with weight layers representing slope, k-factor, c-factor, and all combinations of these three factors.

**Statistical Analysis**

With values of each of the landscape metrics for each of the monitoring point catchments, and corresponding water quality data for each watershed, simple statistical analyses helped reveal the relative importance of the variables. First, correlation analyses were used to determine how the variables were related to each other and if there were any redundancies. Next, a simple regression with benthic IBI as the dependant variable and the landscape metrics as the independent variables was conducted. Several variables were log transformed to improve normality. Correlations and regressions were run for several subsections of the data to determine if different physiographic region had significantly different results. Using the final regression, elasticities of each non-log-transformed regression independent variable were calculated according to the following equation,

\[ \text{Elasticity} = \beta_x \times \frac{\mu_d}{\mu_x} \]

where \( \beta_x \) is the regression coefficient for independent variable \( x \), \( \mu_d \) is the mean value of the dependant variable, and \( \mu_x \) is the mean value of the independent variable \( x \).

Elasticities describe a percent change in the dependant variable from a one percent
increase in each independent variable. In this way elasticities provide a measure of the relative importance of each significant independent variable.

The final step in creating a water quality importance map, includes multiplying the explanatory variables times their corresponding elasticity to create an index of land’s contribution to water quality.

\[
\text{Index} = E_1(x_1) + E_2(x_2) + E_3(x_3) + \ldots + E_n(x_n), \quad (10)
\]

where E is elasticity values for each equally scaled explanatory variable x. In this way, the regression tells us which variables are significant and it determine the weights for each variable based solely on the data and not preconceived ideas of which variables most influence water quality.

**Regression Model & Final Water Quality Model**

All variables associated with forest cover were correlated at a significant level (*Table 2*). Additionally, all variables associated with urban or impervious cover were highly correlated (*Table 3*). Teasing out the effects of correlated variables is extremely difficult. Using this dataset, it proved impossible. When two highly correlated variables were included in a regression, both variables became insignificant. To determine which of the correlated variables explained most of the variability of the data, I ran a series of regressions with each of the correlated variables by themselves. The buffer capacity was the most significant forest cover variable, while the IDW impervious was the most significant urban/impervious cover variable followed by urban saturation (*Table 4*).

When including both the buffer capacity and the IDW impervious variables in a regression together, the IDW impervious lost its significance. When replacing IDW impervious with urban saturation in a regression with buffer capacity, both variables became significant. Therefore these two variables were included in the final model. This method strikes a balance between correlated variable bias and omitted variable bias. It is the simplest and most straight-forward way of dealing with highly correlated variables in regressions, though there have been other more complicated, but effective methods (Van Sickle, 2003).
The fact that buffer capacity showed the strongest relationship with water quality suggests that both the amount of forest and its relative location to the stream are important. Further, it makes the claim that the buffering ability of riparian forests is important, but the buffering ability of non-contiguous upland forests also has a noticeable role in water quality.

The regression that explained most of the variability of the data using buffer capacity and urban saturation is summarized as,

\[
IBI = \beta_0 + \beta_1 (BC) + \beta_2 (US) + \beta_3 (KSC) + \beta_4 (SC) + \varepsilon,
\]

where BC is buffer capacity, US is urban saturation, KSC is K-factor times slope (S-factor) times C-factor, SC is slope (S-factor) times C-factor, \(\beta\) is the regression coefficient for each variable, and \(\varepsilon\) is the error term. Table 5 summarizes the results. Dummy variables for the different physiographic regions were initially included in the regression, but were found to be insignificant.

The elasticities of this regression demonstrate that buffer capacity is the most influential variable. Its positive sign demonstrates that increased buffer capacity is related to better water quality. The remaining variables, urban saturation, K*S*C, and S*C, all have negative relationships with water quality. Increased urban saturation, increased K*S*C, and increased S*C, are all related to worse water quality (Table 6).

Using these regression elasticities, landscape variables were weighted according to their relative importance to water quality to create a map layer that shows the relative water quality importance of any given point. By using the elasticities as weights, instead of user-defined weights, user bias was excluded from the model and the data were allowed to speak for themselves.

The final water quality map (Figure 6 and Figure 7) ranks lands by their relative importance to water quality. Values range from 0 to 100 where value of 100 represents high positive contribution to good water quality, while lower values contribute less.
positively to water quality. Less developed lands such as forests and wetlands tend to have higher values as do areas near the streams and rivers. This confirms the initial hypotheses that landscape pattern is important in determining water quality.

This water quality model serves as a criterion for conservation and can be readily incorporated into a prioritization model with other criteria such as development vulnerability, ecological value, and intake importance areas.

IV. Multi-Criteria Prioritization Model

Approach

Conservation planners have varying reasons for conserving. For example a local land trust may highly value historical and scenic characteristics of a property, while a New York City’s drinking water team may place high value on the water quality value of a land. Yet another group may value lands based on how much biodiversity would be preserved. All these groups would likely prioritize lands that are more vulnerable to rapid development.

Given these differences in value, how can conservation lands be prioritized? And what tool can land conservationists use to prioritize lands for conservation based on their unique set of conservation values and preferences?

In this paper I create a multi-objective decision analysis (MODA) tool, a type of multi-criteria decision analysis (MCDA) tool, that has an undefined number of alternatives that are ranked along a continuous spectrum of better and worse. There are numerous types of decision rules used in MODA studies, although the most well-known and most frequently used approach, and the one used here, is the weighted summation approach with Boolean operations. (Malczewski, 2006)

In this approach land prioritization criteria are ranked and weighted according to the goals and values of the decision-maker to create the final product. In this way,
conservation prioritization schemes are dependent upon the values and judgments of those designing the conservation. Allowing the decision-maker to assign the weight values also increases transparency in how decisions are made.

In this paper, I focus the prioritization explicitly on drinking water quality but incorporate four major components which were identified by EPA as important criteria (Rick Rogers, EPA Region 3 Drinking Water Chief, personal communication, June 2008): 1) land contribution to water quality (described in the previous section), 2) drinking water intake importance areas, 3) development vulnerability, and 4) ecological value. Combining these components in a multi-criteria prioritization model, areas of conservation priority are revealed. The water quality model was described in the previous section, and the other components in the model are described below.

**Intake Importance Areas**

Sediments and contaminants in water pose major difficulties for drinking water suppliers. It can force introduction of a new and costly technology to remove the contaminants, or it can place stress on existing systems. Alternatively, if the contaminants cannot be removed, the public can be faced with worsened drinking water quality. In extreme cases the water services can be temporarily shut down which poses a great risk to public health, and risk to other water-dependent services such as fire fighting. Unsurprisingly, keeping the level of sediment and contaminants low near the drinking water intakes is a priority for clean drinking water.

As water contaminants move through the streams and rivers, they are affected by three primary processes: dilution, decay, and deposition. There have been many different types of nutrient transfer models that try to explain these processes (Reckow et al, 1989). To represent the basic relationship of distance and nutrient levels, I used a nutrient exponential decay function,

\[ Y = e^{-0.01 \times d}, \quad (12) \]
where Y is relative nutrient level and d is distance (km) to intake. In standard decay functions, time is the independent predictor variable, but in this case distance replaces time under the assumption that as the water flows downstream time elapses at a constant rate. This is an over-simplification, but is used for this paper because more precise measurements require extensive data on stream width, depth, and flow amounts along many points of the water system. The equation is not specified for one particular contaminant; instead, it represents a generalized version of the standard nutrient decay function. The model estimates that at 69 km, 50% of the contaminants remain in the waterway; at 300 km, 5% of the contaminants remain in the waterway. Though not ideal, the equation presented above provides a reasonable relative scale of the distance-decay of contaminants in the water. This can be used to represent a scale of intake importance areas.

In addition to distance from intakes, the size of the population served by each intake varies greatly. A risk-based approach may prioritize both distance to intakes and population served by intakes. Two intake priority models were created: one based solely on distance to intakes, and another based on distance to intake and population size served by each intake.

The first model treats all intakes equally in effort to lower the purification cost burden equally to all intakes. The second takes a health risk-based approach considering the human impact of conserving a space. The distance-only intake priority area model can be represented by,

\[ I = \sum (e^{-0.01 * f_k}) , \]

where I is the relative importance to intakes, and \( f_k \) is the flow length to intake \( k \). (See Figure 9). The distance-population intake priority area model can be represented by,

\[ I = \sum (p_k * e^{-0.01 * f_k}) , \]

where \( p_k \) is the population served by intake \( k \). (See Figure 10). Data for the surface water intake locations and populations served were obtained by EPA Region 3 Safe Drinking Water Information System (SDWIS) dataset (US EPA Region 3, 2008). Only the 65 intakes that serve 1,000 or more people were included in the analysis (Figure 11).
Development Vulnerability

Another important component of the conservation prioritization model is vulnerability to development. As part of their Resource Lands Assessment, The Chesapeake Bay Program (CBP) developed a model of development vulnerability in the Chesapeake Bay. It shows the “relative potential risk of future land conversion to urban uses” (CBP, 2008c). The model assesses land on its suitability for development and proximity to growth “hot spots.” The development vulnerability GIS model was acquired from the Chesapeake Bay Program (2008d, data; and 2008c, methods; See Figure 12).

Ecological Value

Though their primary focus is on drinking water quality, the EPA drinking water team specified ecological value as an important component of their future conservation efforts (Rick Rogers, EPA Region 3 Drinking Water Chief, personal communication, June 2008). Indeed, this is a great way to partner with a variety of organizations in conservation efforts. I used the Chesapeake Bay Program’s (CBP) ecological value model (CBP, 2008a).

CBP’s ecological value model identifies the most important remaining habitats in the Chesapeake Bay. The model uses a “hubs and corridors” approach which seeks to link large contiguous patches of ecologically important areas. The “hubs,” or habitat patches, were prioritized by ecoregion based on a set of ecological parameters including rare species presence, rarity and population viability; vegetation and vertebrate richness; habitat area, condition, and diversity; intactness and remoteness; connectivity potential; and the nature of the surrounding landscape. (Weber, 2004, CBP, 2008b; See Figure 13).

Final Model as Decision-Support Tool

To create a combined index of conservation priority, I consider the multiple relevant criteria. The model uses the weighted summation approach, and serves as a multi-object
decision analysis (MODA) tool where the output conservation priority is on a continuous scale.

The basic conservation priority model is represented by,

\[
\text{Cons. Priority} = W_1 \text{ (water quality)} + W_2 \text{ (intake importance)} + W_3 \text{ (development vulnerability)} + W_4 \text{ (ecological value)},
\]  

(15)

where \( W \) is the weight factor assigned to each model component, assumed to be 1 if otherwise not specified.

For demonstration purposes, I used equal weights for each component; however, the conservation priority model above should be adapted to the appropriate organizational value system. After the components are combined (Figure 14 and Figure 15), the final conservation priority map is complete. The output values are on a continuous scale from low priority to high priority for conservation. In this example, high priority areas are spread out throughout the watershed, but more concentrated in the western part of the watershed. When overlaying current protected areas (Figure 16), it becomes clear that current protected areas are doing a good job of being located in high priority areas. When only considering non-protected forested areas in intake watersheds (Figure 17) a majority of lands are excluded, and priority action areas become easier to identify.

The model defined above implies tradeoffs of water quality, intake protection, and ecological threats, and so decision-makers will have to carefully weigh their preferences for the different model components. The outcome of such critical and careful thought will be the definition of the weight factors of the different model components. Such a process brings transparency to the prioritization scheme, and helps decision-makers pin down their exact value system.

\section{V. Conclusions & Recommendations}

In a climate of multiple stakeholders working toward similar, yet uncoordinated goals of source water protection, EPA Region 3 Drinking Water Branch proposed the Potomac
Watershed Priority Lands Strategy. The goal of this strategy was to identify lands for priority conservation for drinking water quality. In the first section of this paper, I modeled how land contributes to water quality. Next, I created a multi-object decision support tool that combines four prioritization criteria—water quality, intake protection, ecological threats, and development vulnerability—to create a final prioritization model. This flexible tool considers the unique preferences and values of the decision-maker, fosters coalition building, and can be used in multiple ways to leverage real and lasting change. In this way it can uniquely adjust to the values of the end-user, and help unite the different groups with similar source water protection goals.

Ultimately deciding which areas to prioritize for conservation is a question of value judgments and preferences. In this way the outcome is adaptable to the diverse needs of different groups. When deciding the weight factors, decision-makers are encouraged to experiment with multiple scenarios. This allows decision-makers to see the implications of their preferences, and in this way helps to clarify how those preferences and value systems are structured. This feedback approach also engenders confidence in the results and allows for hands-on understanding of the model and its assumptions.

In addition to serving as an adaptable tool for different end users, it can be used to build consensus among coalitions of stakeholders with similar objectives. To come to a consensus about the weight values of the model components, the coalition will have to find shared values and inevitably small compromises will be made. For the source water protection in the Potomac, the Potomac River Basin Source Water Protection Partnership (PRBSWPP) represents the coalition of diverse stakeholders, and would be an ideal setting to discuss shared values of source water protection and to define the weight values together.

The transparency of weighting system allows diverse parties to see how different components are valued. For example, the drinking water providers in the PRBSWPP may be attracted to the model because of its emphasis on drinking water and the consideration of contamination levels specifically at the surface water intakes. The
Department of Natural Resources, Department of Environmental Protection, and Environmental Protection Agency may be attracted to the model because of the inclusion of ecological values and aquatic habitat in general in addition to the intake locations. The process of agreeing on weight values can help get buy-in from diverse parties, which may lead to more generous contributions to help fund the conservation projects that this model prioritizes.

Once the weight system is defined, the model is complete and priority conservation areas are easily seen. The model can be used in a number of ways by a number of different parties including:

- Choosing which areas to prioritize and fund conservation;
- Rationale for funding for land acquisitions or conservation easements;
- Updating county or municipal zoning to incorporate strategically located natural areas;
- Consulting with developers to develop the least impacting site plan scenarios;
- Building common understanding as to the benefits of conservation.

In addition, land parcel boundaries can be overlain on each property, and comparing the average priority value of each parcel serves as a means to compare relative priority of a set of parcels.

My final recommendations to EPA Region 3 Drinking Water Branch are:

1. Present this tool to the PRBSWPP as a means to start a discussion on coordinated source water protection in priority areas.
2. Facilitate discussion on how the different weight factors should be weighed to unite the multiple stakeholders, and to get buy-in from all participants.
3. Beginning with the ideas listed in the previous paragraph, discuss in what ways this prioritization model would be most effective, and create a master plan strategy for how to carry out these activities.
4. Bring in the support of GIS professionals to facilitate map-making, and scenario building in all phases of this process.
Acknowledgements

This work was supported by the Environmental Protection Agency Region 3, Drinking Water Branch through the National Network of Environmental Management Studies (NNEMS) program. Maryland Department of Natural Resources graciously provided the Maryland Biological Stream Survey data.

Sincere thanks to all who helped me along the way. Rick Rogers, Chuck Kanetsky, KR Young, Vicky Binetti, and Ellen Schmitt at the EPA gave support, guidance, and inspiration for the project. At the Nicholas School of the Environment, Dean Urban was the primary academic advisor, and John Fay provided numerous insightful bits of GIS advice. Many thanks to the uncountable number of friends and family offering nuggets of guidance, inspiration, motivation, and encouragement all along the way.
References


### Tables & Figures

**Table 1.** C-factor values for different NLCD land cover types.

<table>
<thead>
<tr>
<th>NLCD Land Cover Type</th>
<th>C factor value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>1</td>
</tr>
<tr>
<td>Developed, Open Space</td>
<td>70</td>
</tr>
<tr>
<td>Developed, Low intensity</td>
<td>76</td>
</tr>
<tr>
<td>Developed, Medium Intensity</td>
<td>85</td>
</tr>
<tr>
<td>Developed, High Intensity</td>
<td>92</td>
</tr>
<tr>
<td>Bare rock/sand/clay</td>
<td>85</td>
</tr>
<tr>
<td>Deciduous Forest</td>
<td>55</td>
</tr>
<tr>
<td>Evergreen Forest</td>
<td>55</td>
</tr>
<tr>
<td>Mixed Forest</td>
<td>55</td>
</tr>
<tr>
<td>Pasture/Hay</td>
<td>69</td>
</tr>
<tr>
<td>Row Crops</td>
<td>75</td>
</tr>
<tr>
<td>Woody Wetlands</td>
<td>40</td>
</tr>
<tr>
<td>Emergent Herbaceous Wetlands</td>
<td>40</td>
</tr>
</tbody>
</table>

**Table 2.** Correlations between forest variables. All correlations were highly significant (p<0.0001).

<table>
<thead>
<tr>
<th></th>
<th>% Forest/Wetland</th>
<th>IDW Forest/Wetland</th>
<th>Buffer Capacity</th>
<th>Euclidian Buffer Distance</th>
<th>LN Euclidian Buffer Distance</th>
<th>Flowpath Buffer</th>
<th>LN Flowpath Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Forest/Wetland</td>
<td>1</td>
<td>0.979</td>
<td>0.967</td>
<td>0.597</td>
<td>0.858</td>
<td>0.214</td>
<td>0.564</td>
</tr>
<tr>
<td>IDW Forest/Wetland</td>
<td></td>
<td>1</td>
<td>0.985</td>
<td>0.611</td>
<td>0.898</td>
<td>0.219</td>
<td>0.568</td>
</tr>
<tr>
<td>Buffer Capacity</td>
<td></td>
<td></td>
<td>1</td>
<td>0.591</td>
<td>0.881</td>
<td>0.22</td>
<td>0.573</td>
</tr>
<tr>
<td>Unconstrained Buffer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LN Euclidian Buffer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td></td>
<td>0.597</td>
<td>0.811</td>
<td>0.6871</td>
<td>0.449</td>
<td>0.516</td>
<td>0.681</td>
</tr>
<tr>
<td>Flowpath Buffer</td>
<td>0.214</td>
<td>0.219</td>
<td>0.22</td>
<td>0.188</td>
<td>0.303</td>
<td>0.303</td>
<td>1</td>
</tr>
<tr>
<td>LN Flowpath Buffer</td>
<td>0.564</td>
<td>0.568</td>
<td>0.573</td>
<td>0.449</td>
<td>0.516</td>
<td>0.681</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 3.** Correlations between urban variables. All correlations were highly significant (p<0.0001).

<table>
<thead>
<tr>
<th></th>
<th>% Urban</th>
<th>IDW Urban</th>
<th>Urban Saturation</th>
<th>Mean Impervious</th>
<th>IDW Impervious</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Urban</td>
<td>1</td>
<td>0.968</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDW Urban</td>
<td>0.968</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban Saturation</td>
<td>0.954</td>
<td>0.974</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Impervious</td>
<td>0.959</td>
<td>0.922</td>
<td>0.899</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IDW Impervious</td>
<td>0.936</td>
<td>0.931</td>
<td>0.896</td>
<td>0.978</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 4. R-squared results from a series of regressions on B-IBI with each of the correlated variables by themselves to determine which of the correlated variables explained most of the variability of the data.

<table>
<thead>
<tr>
<th>Forest Variables</th>
<th>R-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Forest/Wetland</td>
<td>0.2190</td>
</tr>
<tr>
<td>IDW Forest/Wetland</td>
<td>0.2358</td>
</tr>
<tr>
<td>Buffer Capacity</td>
<td>0.2637</td>
</tr>
<tr>
<td>Euclidian Buffer Distance</td>
<td>0.0771</td>
</tr>
<tr>
<td>LN Euclidian Buffer Distance</td>
<td>0.1553</td>
</tr>
<tr>
<td>Flowpath Buffer</td>
<td>0.0076</td>
</tr>
<tr>
<td>LN Flowpath Buffer</td>
<td>0.0838</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Urban Variables</th>
<th>R-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Urban</td>
<td>0.1530</td>
</tr>
<tr>
<td>IDW Urban</td>
<td>0.1718</td>
</tr>
<tr>
<td>Urban Saturation</td>
<td>0.1761</td>
</tr>
<tr>
<td>Mean Impervious</td>
<td>0.1208</td>
</tr>
<tr>
<td>IDW Impervious</td>
<td>0.2358</td>
</tr>
</tbody>
</table>

Table 5. Final water quality regression results. R-squared is 0.3153.

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>St. Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer Capacity</td>
<td>0.0156971</td>
<td>0.0017456</td>
<td>0.000</td>
</tr>
<tr>
<td>Urban Saturation</td>
<td>-0.0143129</td>
<td>0.0027168</td>
<td>0.000</td>
</tr>
<tr>
<td>K<em>S</em>C</td>
<td>-0.0000138</td>
<td>0.0000551</td>
<td>0.087</td>
</tr>
<tr>
<td>S*C</td>
<td>-0.0002156</td>
<td>0.0000563</td>
<td>0.000</td>
</tr>
<tr>
<td>constant</td>
<td>2.982008</td>
<td>0.466416</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 6. Elasticities for final water quality regression.

<table>
<thead>
<tr>
<th></th>
<th>Elasticity</th>
<th>St. Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer Capacity</td>
<td>0.3130365</td>
<td>0.03511</td>
<td>0.000</td>
</tr>
<tr>
<td>Urban Saturation</td>
<td>-0.0853064</td>
<td>0.01624</td>
<td>0.000</td>
</tr>
<tr>
<td>K<em>S</em>C</td>
<td>-0.0137249</td>
<td>0.05556</td>
<td>0.087</td>
</tr>
<tr>
<td>S*C</td>
<td>-0.3739041</td>
<td>0.09779</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Figure 1. Land use in Potomac watershed
Figure 2. MBSS stream monitoring points spanning the five physiographic regions
Given Land Cover

Figure 3. Inverse distance weighted schematic. The sum of the pixels in the top right raster is multiplied by 100 and divided by the sum of the pixels in the bottom right raster to get the final IDW value for the watershed.

Figure 4. Buffer capacity is calculated for each pixel in a watershed.
Figure 5. Soil erosion model comprised of K-factor * Slope * Land Use.
Figure 6. Water quality model. Green areas represent high positive contribution to water quality, while red areas contribute less positively to water quality.
Figure 7. Water quality model with zoom in. Green areas represent high positive contribution to water quality, while red areas contribute less positively to water quality.
Figure 8. Decay function, $Y = 100 \times e^{(-0.01 \times d)}$, where $d$ = distance (km) and $Y$ = relative nutrient level.
Figure 9. Intake priority area, all intakes equal.
Figure 10. Intake priority area, intakes weighted by population served.
Figure 11. The intakes in the Potomac Watershed and their overlapping catchments.
Figure 12. Chesapeake Bay program’s development vulnerability model.
Figure 13. Chesapeake Bay program’s “hubs and corridors” approach to ecological model.
Figure 16. Final Conservation Priority Model with current protected areas.
Figure 17. Final Conservation Priority Model showing available forest land for conservation.