Linking upstream mining to downstream water quality: Mountaintop mining in West Virginia

By

Catherine Carter

Date: ______________

Approved:

____________________
Dr. Emily S. Bernhardt, Advisor

Masters 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

April 30, 2010
Abstract
Mountaintop mining valley fill (MTM/VF) coal mining is currently the dominant form of land use change in the central Appalachians. MTM/VF activities level mountains, remove forests and forest soils, bury headwater streams and generate substantial amounts of acid and alkaline mine drainage. Numerous case studies have documented elevated concentrations of sulfate and trace metal and metalloids with known toxicity in surface waters downstream from MTM/VF activity, yet no comprehensive effort has been made to link landscape scale mining activity and water quality. Here, I used newly obtained remote sensing data of surface mining activity delineated from 1976 to 2005 to estimate the extent of MTM/VF impact on downstream surface water quality in the Coal and Guyandotte river basins of WV. Hydrologic connectivity between mining and water quality was estimated using an inverse distance weighting technique in GIS (ESRI, Inc.). The findings show significant biogeochemical alterations, including streamwater conductivity and sulfate concentrations, even when small amounts of surface mining (<5%) are observed. Results provide the first comprehensive analysis of the cumulative impact of mining activity in these watersheds on water quality and demonstrate the need for further investigation involving strategic water quality sampling with the ultimate goal of developing an empirical basis on which to form regulations governing MTM/VF throughout the central Appalachians.

Introduction
The central Appalachian ecoregion, stretching north and south beyond the state borders of West Virginia (WV), is comprised almost entirely of forested landscape. However, during the period from 1973 to 2000, forest cover in the Central Appalachians declined from 86.6% to 83.3%, primarily as a result of increasing surface mining in the ecoregion (Saylor 2008). The increase in the intensity and spatial extent of mining activity was spurred throughout the 1990s by an increased demand for WV’s highly efficient, low-sulfur coal (Davis & Duffy 2009, Fox 1999). Ironically, this demand was witnessed after the 1990 amendment to the Clean Air Act (Fox 1999). With advances in technology and rising natural gas and oil prices, surface mining in WV jumped from 10% to nearly a third of the state’s total coal production; the majority of this increase in surface mining is due to an increase in mountaintop mining coupled with valley filling (MTM/VF; Davis & Duffy 2009).

While the coal extracted from these low-sulfur coal seams produces less atmospheric pollution upon combustion, the MTM/VF methods for extracting coal from the mountains of WV are accompanied by severe ecological and societal costs (e.g. Fox 1999, Sams & Beer 2000, Lemly 2008, Davis & Duffy 2009, Palmer et al. 2010, among others). Throughout the MTM/VF process, forests are cleared, topsoil removed, and upper portions of mountains lost as explosives
are used to gain access to underlying veins of coal. Excess rock and soil (overburden) left after coal extraction is then collected and dumped into surrounding valleys, burying nearby streams (Fox 1999, Davis & Duffy 2009, Palmer et al. 2010). In addition to the destruction of filled streams, MTM/VF also converts high elevation hardwood forests to reclaimed grasslands. Beyond their local impacts, MTM/VF practices have a variety of landscape scale impacts, among them are reports of reduced infiltration capacity of soils at mined sites leading to higher threats of downstream flooding and ground and surface water contamination through the release of toxic solutes and ions (Fox 1999, Sams & Beer 2000, Lemly 2008, Davis & Duffy 2009, Townsend et al. 2009, Palmer et al. 2010). Attempts to reclaim the impacts of MTM/VF through reforestation, planting grasslands, or the creation or enhancement of streams have not been reported as successful methods for mitigation as there is strong evidence that reclaimed mine sites remain fundamentally altered from their pre-mining state (Townsend et al. 2009, Palmer et al. 2010). Despite laws designed to regulate environmental effects from coal mining such as the Surface Mining Control and Reclamation Act of 1977 (SMCRA, Sams & Beer 2000), examination of the known influences of MTM/VF on the environment has led to the designation of WV as an environmental sacrifice zone, where low-sulfur coal is provided at the expense of WV’s “rapid degradation of its own environment” (Fox 1999).

Perhaps the most well understood impact of MTM/VF activities are their effects on downstream water quality. When iron-sulfide minerals (pyrite) found in and adjacent to coal seams are exposed to oxygen and water through MTM/VF practices, sulfuric acid ($H_2SO_4$) is produced resulting in acid mine drainage (AMD). Secondary reactions of sulfuric acid with carbonate rocks in mining overburden can buffer acidity ($H^+$) through the release of high concentrations of bicarbonate ions and base cations such as calcium and magnesium. Sulfuric acid weathering of parent material also leads to the release of aluminum, manganese, zinc, and selenium into streams draining upstream valley fills (Sams & Beer 2000, Lemly 2008, Palmer et al. 2010). Sulfate ($SO_4^{2-}$) has been identified as an indicator of the presence of mine drainage in streams (Sams & Beer 2000) as well as being closely linked to levels of mining activity in watersheds (Paybins et al. 2000, Palmer et al. 2010). Throughout WV, elevated concentrations of $SO_4^{2-}$ are accompanied by elevated concentrations of trace metal and metalloids with known toxicity (Al, Mn and Se; Palmer et al. 2010). In the same streams, macroinvertebrate communities tend to become dominated by more pollution tolerant taxa and to lose sensitive
species with higher concentrations of SO$_4^{2-}$ (Palmer et al. 2010). Stream health metrics derived from this macroinvertebrate community composition data (the WV Stream Condition Index, WVSCI) have recently been shown to be an important correlate of incidences of a variety of human cancer cases in WV (Hitt & Hendryx 2010). Not only do these relationships support the use of SO$_4^{2-}$ as an indicator of the extent of current mining activity in watersheds, but elevated SO$_4^{2-}$ loading to streamwaters has been shown to persist long after mining activity ceases (Sams & Beer 2000), suggesting that SO$_4^{2-}$ concentrations serve as an indicator of both past and present mining activity.

Selenium (Se) is a metalloid contaminant of particular concern in the mining impacted areas of WV as it is an element that occurs at high concentrations in coal, which is weathered and released to solution during AMD reactions. Se becomes more soluble at a higher pH and is thus of particular concern in alkaline mine drainages prevalent in much of the Central Appalachians. Se has the ability to bioaccumulate and has been identified as being significantly toxic to aquatic life (Lemly 2008, Conley et al. 2009, WVDEP 2010). The direct toxicity and reproductive impacts of Se on fish have been thoroughly investigated in the Mud River ecosystem in southwestern WV (Lemly 2008, WVDEP 2010). Reported findings indicate the Mud River (including the Upper Mud Reservoir) is at the highest hazard rating based on Se concentrations found in water and fish tissues compared to known toxic thresholds. Se concentrations for the waters of the Mud River and Reservoir were reported as being 1.5 to 4.5 times the U.S. Environmental Protection Agency (EPA) freshwater criterion established to protect aquatic life (5ug/L, Lemly 2008). These findings support the threat of a potential collapse of fish populations throughout the Mud River ecosystem as a response to elevated Se concentrations in stream and reservoir waters resulting from upstream MTM/VF (Lemly 2008, WVDEP 2010). In a 2009 study performed by Conley et al., a laboratory life cycle assay was conducted to assess the impacts of realistic scenarios of dietary exposure of Se to the mayfly (Centroptilum triangulifer). The authors found that grazing mayflies fed periphyton that had been exposed to Se concentrations ranging from 5 to 20 µg L$^{-1}$ accumulated body burdens of Se greater than 3ug g$^{-1}$, which is the recommended threshold suggested for the protection of fish and wildlife (Lemly 1993). Additional findings in this study refuted the argument that aquatic insects are not impaired by exposure to Se as a decrease in fecundity and a reduction in growth (adult body mass) associated with elevated Se body burdens were observed (Conley et al. 2009).
Studies such as these demonstrate the magnitude of influence Se contamination in surface waters draining MTM/VF sites has on the health of biota which depend on clean water for survival.

With the exception of a single USGS study linking basinwide coal production to sulfate concentrations in several major rivers (Sams and Beer 2000), there has been no attempt to directly link streamwater SO\textsubscript{4}\textsuperscript{2-} concentrations to the extent of upstream watershed mining. Building such predictive statistical relationships between mining activity in a watershed, streamwater SO\textsubscript{4}\textsuperscript{2-} concentrations, and ultimately environmental degradation are important for understanding the relationship between mining practices and observed ecological impacts. Establishing the strength, direction, and shape of the relationship between mining activity and SO\textsubscript{4}\textsuperscript{2-} would provide an empirical basis on which to base regulations about the downstream and cumulative impacts of mining activity within watersheds. In addition, the development of spatially-explicit models to be used as predictions of stream ecosystem condition relative to mining locations would be beneficial for prioritizing locations for future water quality analyses. A model of this type would indicate the extent of mining activity as well as areas in watersheds that most strongly influence downstream water quality due to mining (Van Sickle & Johnson 2008).

In order to spatially represent the direct link between land-use/landcover and downstream water quality, several studies have implemented an inverse distance weighting (IDW) technique performed using hydrologic tools provided by geographic information system (GIS) technology (ESRI, Inc., Hunsaker & Levine 1995, Comeleo et al. 1996, Soranno et al. 1996, King et al. 2005, Van Sickle & Johnson 2008). Models developed using IDW assume that the influence of a particular land-use/landcover on in-stream water quality increases with proximity to streams or stream sampling locations (King et al. 2005). Two measurements have been previously performed for determining the proximity of a land-use/landcover type to stream sampling locations – 1) linear Euclidean distance measurements where distances were divided into bins, and 2) flow-length measurements that accounted for the differences between overland flow distance and in-stream flow distance (Comeleo et al. 1996, King et al. 2005, Van Sickle & Johnson 2008).

Through the implementation of a customized IDW technique, developed as a combination of the aforementioned linear and flow-length methods, the objectives of my study are 1) to determine the relationship between mining activity and downstream water quality, and
2) to determine the downstream extent of contamination in streams draining MTM/VF. This analysis uses $\text{SO}_4^{2-}$ as an indicator relating mining activity to environmental degradation. Both linear and flow-length methods were combined for this study to develop an analysis that is fairly easy to interpret, but also considers the varying impacts between mining activity local to streams (within or close to streams) to those performed at greater hydrologic distances from stream sampling sites.

The IDW method has never before been implemented for linking upstream mining in WV to downstream water quality impacts. An additional novelty of this study is its use of classified satellite imagery that provides delineations of mining activity spanning the period 1976 to 2005. Mining activity delineated across these 4 decades was provided by SkyTruth, Inc., a nonprofit organization in WV (501(c)(3)) devoted to promoting environmental awareness and protection with remote sensing and digital mapping technology (http://www.skytruth.org/). Previous statistical analyses have relied upon permit boundaries provided by the WVDEP to estimate mining influence (Sams & Beers 2000, Palmer et al. 2010). Permit boundaries, while updated regularly, may misrepresent actual areas devoted to active mining (or previously active mining), often exaggerating levels of mining activity within a watershed since not all of the permitted area is impacted. Permit boundaries also provide little information about the spatial configuration and exact location of intense mining activity.

**Methods**

**Study Area** – The Coal and Upper Guyandotte River Basins, located in southwest WV, were selected as the focus for this study because they contain areas of high density surface mining activity relative to the rest of the state (Figure 1). The Mud River Watershed, a portion of the Lower Guyandotte river basin, was also included in this study because it had only one surface mine located in its headwaters and provided a unique analysis of the downstream distance impacted by MTM/VF that could not have been performed in watersheds containing a myriad of surface mining locations. An additional reason for selecting the Mud River Watershed was that recent analyses have demonstrated concern for the level of contamination found in surface waters downstream from the one mining site (Lemly 2008, WVDEP 2010).

Landcover composition within the Coal and Upper Guyandotte River Basins and the Mud River Watershed were determined through the use of two datasets – 1) the National Landcover Dataset 2001 (NLCD generated by the Multi-Resolution Land Characteristics Consortium), and
2) surface mining activity delineated by SkyTruth, Inc. Landcover classifications were performed by SkyTruth, Inc. through a decision tree analysis using Landsat satellite imagery to isolate areas of mining activity to serve as an account of historical occurrences of mountaintop mining from 1976 to 2005. Throughout this analysis, ‘surface mining activity’ included all mining activity delineated by SkyTruth, Inc. from 1976 to 2005, including valley fills. Based on these available data, landcover within the study area was found to be predominantly forested (~80%) while development, including low, medium, and high intensity, covered only 2% of the landscape. Surface mining activities occupied approximately 8% of the land area within the Coal and Upper Guyandotte River Basins, roughly 2.5% of the Mud River Watershed, and close to 7% when the two areas were combined.

Water Quality – Water quality data used throughout this analysis was obtained from the West Virginia Department of Environmental Protection’s (WVDEP) water quality database with stream sampling locations scattered throughout the entire state of WV having collection dates ranging from 1996 to 2008. It is important to note that assessing the environmental impacts of surface mining activity was not the WVDEP’s primary purpose for creating this database. Samples collected within the Coal and Upper Guyandotte River Basins and the Mud River Watershed were selected for use in this analysis and were limited to only include sampling locations where surface water SO$_4^{2-}$ concentrations were recorded. If two samples were collected from an identical sampling location, only the sample having the greatest SO$_4^{2-}$ concentration was retained for this analysis; the dataset thus refined provided information on water quality for 309 stream sampling locations (Figure 2). While SO$_4^{2-}$ was the primary water quality parameter of interest for this study, I also investigated the relationship between watershed mining activity and surface water Se concentrations (mg L$^{-1}$), conductivity (µmhos cm$^{-1}$), chloride (Cl$^-$; mg L$^{-1}$) and macroinvertebrate metrics compiled using either family or genus level taxonomic resolution. While my interest in determining the relationship of surface mining to Se was due to the known impacts of Se on biota living in waters drained from mining locations (Lemly 2008, Conley et al. 2009, WVDEP 2010), my interest in studying conductivity and Cl$^-$ was to determine if SO$_4^{2-}$ released during mining was the primary source of elevated conductivity levels in surface waters. My decision to include macroinvertebrate metrics in this analysis was to assess the relationship between surface mining activity and an indicator of biological response.
\textit{Geospatial Analysis} – Geospatial data layers utilized throughout this analysis included elevation, flow accumulation, and flow direction grids at a 30 meter resolution obtained from the National Hydrography Dataset Plus (NHD Plus, http://www.horizon-systems.com/nhdplus/). A stream network was manually generated by setting a flow accumulation threshold of 1,000 cells and designating all cells having a flow accumulation count greater than 1,000 as streams. This method for generating streams created a stream network that closely agreed with the 1:24,000-scale NHD Plus flowlines for this region (see Appendix, Figure 1); however, the stream layer I generated had slightly more tributaries. These tributaries, obtained by selecting the appropriate flow accumulation threshold, were desired because water sampling locations were found to coincide with most of the smaller tributaries that would have otherwise been excluded had I used the NHD Plus flowlines dataset for my analysis. A riparian corridor extending 30m (Euclidean distance) away on each side of the streams and throughout the entire stream network was created to later be used in the IDW process.

Sample locations served as watershed pour points throughout this analysis as one watershed was delineated from the NHD Plus flow accumulation grid and individually inspected for each sample location (309 samples were included). Some of the higher order (3rd – 5th order, Strahler stream order) streams had more than one sample location throughout their reach. In order to avoid creating nested watersheds, all 309 watersheds were delineated in an iterative process. While having several sample locations on the same stream reach created dependencies between sample points, this attribute did allow for the assessment of downstream distance impacted by surface mining activities. Once watersheds were delineated, the 30 m riparian corridor and flow direction grid were masked to each of the 309 watershed boundaries.

The IDW process was initiated by determining the flow-length distance for every cell within each of the individually created watershed boundaries. Flow-length distance was measured by tracing the path of steepest descent between grid cells across the flow direction raster and was implemented in GIS (ESRI, Inc.; King \textit{et al.} 2005). To capture the direct impact of surface mining activity within or directly adjacent to stream channels, flow-length to a sample location was first estimated for each cell within the 30m riparian corridor for each watershed individually. These distances were then divided into unequal-interval classes and new distance values were applied to all cells within each distance class (Table 1). Assigning new values to the cells of each distance class allowed for mining sites within or directly adjacent to stream
locations to receive higher weights throughout the IDW process when compared to mining sites on land that were not located within a riparian corridor. The decision to assign new values to riparian corridor cells followed the assumption that \( \text{SO}_4^{2-} \) released within streams or riparian corridors had a greater influence on \( \text{SO}_4^{2-} \) concentrations measured at each sample location than did \( \text{SO}_4^{2-} \) released at mining sites with greater hydrologic distances from sample locations. Flow-length distance was then estimated for each cell within the entire drainage area for every sample location to account for overland flow distance. Overland flow distance values for each cell located within the 30m riparian corridor were replaced with the newly assigned flow distance values generated for the cells of the 30m riparian corridor. All flow distance values for each watershed were then divided into unequal-interval distance classes based on each cell’s distance from sampling location (Table 2; process outlined in Figure 3). Flow distance classes were selected to match those that were utilized throughout the IDW technique reported by King et al. (2005) who divided distances into 8 classes: 0-100m, 101-250m, 251-500m, 501-1,000m, 1,001-2,000m, 2,001-5,000m, 5,001-10,000m, and >10,000m. Surface mining area within each distance class was calculated for each watershed and received a weight based on its proximity to the sample location – activities in lower distance classes (closer to a sampling location) received a higher weight than those that were in higher distance classes. The highest distance within each range defining the distance classes was used to represent all surface mining cells within the range in the final IDW calculation (e.g. 501-1000m distance class was assigned a distance of 1000).

The equation for calculating inverse distance-weighted percentage surface mining activity in a watershed is that which was used by King et al. (2005) and is as follows:

\[
\% \text{ IDW surface mining activity in watershed} = 100 \times \frac{\sum_{i=1}^{C} n_m \cdot W_c}{\sum_{i=1}^{C} n \cdot W_c}
\]

where \( C \) is the number of distance classes, \( n_m \) is the number of surface mining cells within distance class \( i \), \( W_c \) is the inverse-distance weight for distance class \( i \) where \( d^{-1} \) = the maximum distance between a cell in distance class \( i \) and the sample location (Table 2). The \% IDW surface mining activity calculated for each watershed was then compared to the water quality parameters recorded for each corresponding sample location. Due to the inconsistencies between the dates for surface mining data (1976 – 2005) and water quality data (1996 – 2008), some sample locations were found to have low \( \text{SO}_4^{2-} \) concentrations (< 300 mg L\(^{-1}\)) while \% IDW surface mining activity I determined for their watershed was high (> 10%). For many of these locations,
I found that water quality samples were collected prior to the onset of surface mining in the corresponding watershed. For all sample locations found to fit this scenario, % IDW surface mining activity was changed to zero (See Appendix, Table 1).

In order to assess whether $\text{SO}_4^{2-}$, conductivity, $\text{Cl}^-$, or macroinvertebrate metric levels associated with the level of mining in a watershed were statistically significant when compared to levels for samples associated with unmined watersheds (background levels), a series of hypothesis tests were performed. Welch two sample $t$-tests were performed in R (v.2.9.1; R Development Core Team 2009) under the assumptions of normality and independency between and within samples. Log transformations were performed on all measured concentrations for $\text{SO}_4^{2-}$, conductivity, and $\text{Cl}^-$ in order to fit the normality assumption; however, a log transformation was not entirely effective for all sample distributions, but proved to be the best overall transformation for these data based on the Shapiro-Wilk normality test (Shapiro & Wilk 1965, Shapiro et al. 1968). Original macroinvertebrate metric values were retained for this analysis. The water quality data collected by the WVDEP is known to exhibit dependency because sample locations were located within the same stream network (e.g. along the main stem in the Mud River Watershed) and a measured concentration at a downstream location is dependent upon a measured concentration at an upstream location. While the $t$-test is known for its robustness to the normality assumption, a more robust analysis could have been performed if watersheds were non-overlapping, in other words, if only one water quality sample was collected from each tributary analyzed; however, this would have dismissed the opportunity to perform an analysis to assess the downstream distance impacted by mining activity.

The downstream distance impacted by surface mining activity was measured for both the Mud and Coal Rivers. As previously mentioned, the Mud River had only one surface mining site within the headwaters of its watershed. Sample dates along the Mud River were either 1998 or 2003 and, for the most part, alternated by date from mining site to pour point of the watershed. This unique sampling pattern allowed for the downstream extent of impact from one mining site to be investigated for both 1998 and 2003, and allowed for a comparison of the increase in downstream surface water $\text{SO}_4^{2-}$ measured during this time period as a response to the increase in surface mining activity from 1995 to 2005. Increase in surface mining activity for the Mud River Watershed could not be determined for 1998 to 2003 because mining activity was delineated on the 5th year of every decade from 1976 to 2005; therefore, the time range used to assess the
amount of mining activity growth was from 1995 to 2005. Water quality sampling along the Coal River only allowed for a single year assessment for 1997 instead of a temporal trend analysis such as that which was performed for the Mud River. This analysis was performed as a preliminary investigation of the cumulative impacts of surface mining on downstream water quality and serves as a general comparison between the level of surface mining activity in a watershed and the associated level of downstream impact. A more robust temporal trend analysis accounting for changes in discharge throughout the stream network within the bracketed time period as well as repeat monitoring of water quality at consistent sample locations would allow for a direct comparison between surface mining activity in a watershed and cumulative downstream impacts.

Results

Inverse Distance Weighting - The resulting % IDW surface mining activity for all 310 watersheds were divided into 5 bins: <5%, 5-10%, 10-20%, 20-40%, and >40%. The distribution for measured water quality parameters ($SO_4^{2-}$, Se, conductivity, macroinvertebrate indices) was plotted for each of the surface mining bins as a box and whisker plot (e.g. Figure 4). For box and whisker plots, the central box shows the interquartile range—the spread of concentrations from the 25th to the 75th percentile of the distribution. The solid line in the center shows the median (50th percentile). Vertical lines on either side of the box (the “whiskers”) show the complete range, from the minimum to the maximum observed value excluding outliers. The distribution for water quality parameters measured at all sample locations having 0% IDW surface mining activity within their watersheds was plotted separate from the box and whisker plots in order to highlight recorded values for samples that were not impacted by surface mining activity (e.g. Figure 4). All individually inspected sample locations found to be influenced by underground mining or reclamation were plotted separate from any unexplained outliers (e.g. Figure 4).

$Sulfate$ – A positive relationship was found between % IDW surface mining activity in a watershed and surface water $SO_4^{2-}$ concentration; however, there were a few interesting outliers that required further investigation (Figure 5). Outliers representing samples with recorded $SO_4^{2-}$ concentrations >300 mg L$^{-1}$ for watersheds containing less than 5% IDW surface mining activity were individually inspected to see if elevated $SO_4^{2-}$ concentrations could have resulted from mining activity that was not accounted for in the SkyTruth, Inc. surface mining data, such as underground mining or upstream processing plant influences. Underground mining limits and
coal processing plant locations were obtained from the WVDEP (http://gis.wvdep.org/data/omr.html). Each were displayed in GIS (ESRI, Inc.) and a visual inspection was performed for each of the aforementioned outliers to assess whether elevated \( \text{SO}_4^{2-} \) concentrations appeared to be influenced by underground mining activity or by processing plant locations within the watershed. Six of the outliers were found to be associated with these activities (Figure 5; See Appendix, Table 2). Outliers having recorded \( \text{SO}_4^{2-} \) concentrations <300mg L\(^{-1}\) for watersheds containing >10% IDW surface mining activity were also inspected. Eight of these outliers were found to have reclaimed mining sites within their watersheds that were included in the calculation for %IDW surface mining activity (See Appendix, Table 1).

After individually inspecting each outlier and performing t-tests, unmined watersheds were found to have \( \text{SO}_4^{2-} \) concentrations that varied between 53 and 71 mg L\(^{-1}\) (log scale: 3.3 – 3.6; 95% confidence interval (CI)) and an estimated mean value of 58 ±7 mgSO\(_4^{2-}\) L\(^{-1}\) (log scale: 3.4 ±0.09). All %IDW surface mining activity categories (<5%, 5-10%, 10-20%, 20-40%, and >40%) were estimated to have mean \( \text{SO}_4^{2-} \) concentrations that were significantly higher than the mean \( \text{SO}_4^{2-} \) concentration estimated for sample locations corresponding to 0% IDW surface mining activity (p-values all < 0.0001; See Table 3 for t-test results). In addition to being statistically significant from background levels of \( \text{SO}_4^{2-} \), the magnitude of difference in means between unmined samples and samples in all other categories of mining activity increased from approximately 3 times greater for samples impacted by < 5% surface mining to approximately 12 times greater for samples impacted by > 40% surface mining activity (Table 3). In other words, samples impacted by < 5% surface mining activity had a mean \( \text{SO}_4^{2-} \) concentration found to be approximately 3 times greater than the \( \text{SO}_4^{2-} \) concentration reported for samples in unmined watersheds, and samples impacted by > 40% surface mining activity had a mean \( \text{SO}_4^{2-} \) concentration found to be approximately 12 times greater than the mean \( \text{SO}_4^{2-} \) concentration found for samples in unmined watersheds.

Selenium – Surface water Se concentrations were also found to have a positive relationship with the extent of surface mining activity in a watershed (Figure 5). While this relationship is apparent, sample sizes for Se in areas of greatest density of surface mining activity were very low (~ 3 samples, Figure 5) and additional sampling is required to further explore the relationship between Se and the extent of surface mining activity in a watershed.
**Conductivity** – The relationship between the extent of surface mining activity in a watershed and conductivity was as expected, a positive relationship (Figure 6). Unmined watersheds were found to have conductivity levels varying between 258 and 325 μS cm⁻¹ (log scale: 5.1 – 5.3; 95% CI) and an estimated mean value of 275 ±25 μS cm⁻¹ (log scale: 5.2 ±0.08). All %IDW surface mining activity categories (<5%, 5-10%, 10-20%, 20-40%, and >40%) were estimated to have mean conductivity levels that were significantly higher than the mean conductivity estimated for sample locations corresponding to 0% IDW surface mining activity (p-values all < 0.0001; Table 3). In addition to being statistically significant from background levels of conductivity, the magnitude of difference in means between unmined samples and samples in all other categories of mining activity increased from approximately 2 times greater for samples impacted by < 5% surface mining to approximately 6 times greater for samples impacted by > 40% surface mining activity (Table 3). In other words, samples impacted by < 5% surface mining activity had a mean conductivity found to be approximately 2 times greater than the conductivity reported for samples in unmined watersheds, and samples impacted by > 40% surface mining activity had a mean conductivity found to be approximately 6 times greater than the mean conductivity found for samples in unmined watersheds.

Conductivity to SO₄²⁻ ratios were calculated for each sample location to assess whether SO₄²⁻ was a dominant contributor to elevated conductivity levels for samples greatly influenced by surface mining activity (Figure 7). Computing this ratio allowed for an interesting comparison between conductivity constituents at low levels of mining to those at high levels of mining. All samples influenced by surface mining activity, divided by mining category (<5%, 5-10%, 10-20%, 20-40%, and >40%), were found to have conductivity: SO₄²⁻ ratio values significantly lower than samples in unmined watersheds (p-values < 0.01; Table 3 and Figure 7). Lower ratios indicate that SO₄²⁻ is a greater contributor to measured conductivity. Mean ratio values were farthest below ratios calculated for unmined watersheds for samples associated with > 40% surface mining activity (approximately 2.4 times lower; Table 3). In comparison, the mean Cl⁻ concentration (mg L⁻¹) recorded for samples associated with > 40% surface mining activity was not found to be significantly different from the mean Cl⁻ concentration estimated for samples in unmined watersheds (p-value = 0.19; Table 3 and Figure 8). However, it is important to note that the sample size for Cl⁻ concentrations measured at samples having > 40% surface mining activity...
was only 10 samples; this is well below the commonly recommended sample size of 30 when performing hypothesis tests.

**Macroinvertebrate metrics** – The family level WVSCI and the genus level GLIMPSS macroinvertebrate metrics were plotted against % surface mining activity in a watershed. For both of these metrics, high values are more desirable than low values as they serve as indicators of ecological integrity where high values represent high biodiversity among macroinvertebrate communities including taxa that are intolerant to pollution. WVSCI values for unmined watersheds were found to range between 66 and 70 (95% CI) with an estimated mean of 67 ±1.5. All %IDW surface mining activity categories (<5%, 5-10%, 10-20%, 20-40%, and >40%) were estimated to have mean WVSCI values that were significantly lower than the mean WVSCI value estimated for sample locations corresponding to 0% IDW surface mining activity (p-values all < 0.05; Table 4). GLIMPSS values for unmined watersheds were found to range between 52 and 57 (95% CI) with an estimated mean of 53 ±1.9. Similar to WVSCI, all %IDW surface mining activity categories were estimated to have mean GLIMPSS values that were significantly lower than the mean GLIMPSS value estimated for sample locations corresponding to 0% IDW surface mining activity (p-values all < 0.001; Table 4). A comparison of the family based WVSCI to the genus based GLIMPSS (Figure 9) demonstrated that genus level identification was more sensitive to mining activity in watersheds when compared to family level identification; however, there was a general decline in metric value for both family and genus based metrics (Figure 9) and the lower values for these metrics indicate a reduction in ecosystem integrity.

**Downstream distance impacted by surface mining activity** – There were few data available to assess the longitudinal and cumulative impacts of mining activity. Multiple sites were sampled on the Mud River in both 1998 and 2003, allowing a comparison of the longitudinal patterns in streamwater SO$_4^{2-}$ concentrations downstream of the Hobet mine. The average SO$_4^{2-}$ concentrations of streams in Coal and Guyandotte watersheds without surface mining was 58 ± 7 mg L$^{-1}$. In 1998 the WV DEP site on the Mud River closest to the Hobet mine measured SO$_4^{2-}$ concentrations of 430 mg L$^{-1}$, and dilution by unmined tributaries lowered the Mud River SO$_4^{2-}$ to 33 mg L$^{-1}$ approximately 111 km farther downstream. Between 1995 and 2005 the size of the Hobet mine nearly doubled, increasing from 1,400 to 2,700 hectares as delineated by SkyTruth, Inc. (Figure 10). In samples collected in 2003, WVDEP found that sulfate concentrations in the
stream were higher, not only immediately below the mine (529mg L$^{-1}$) but also more than 112km downstream of the mine, where streamwater concentrations were 4x higher than in 1998 (121 mg L$^{-1}$, Figure 10) and well above the concentrations measured in unmined streams of the area. While the increase in sulfate concentrations is consistent with the increase in mining, no data are available to determine the extent to which this temporal pattern could also be due to differences in streamflow between the two sampling dates.

Multiple samples were collected along both branches of the Coal River in 1997, and demonstrate that in this basin with many individual surface mines SO$_4^{2-}$ concentrations generally increase throughout the river basin as a response to several surface mining sources draining to the Coal River (Figure 11). There were no data available to assess whether streamwater SO$_4^{2-}$ concentrations in this watershed have changed through time as a result of the increasing spatial extent of surface mining.

Discussion

Implications of Study – One primary conclusion is supported by the results of this study. There is a positive relationship between the extent of surface mining activity in a watershed and the conductivity and sulfate concentrations of downstream surface waters. I found that even small amounts of mining activity (<5%) generally lead to significantly higher streamwater conductivity and SO$_4^{2-}$ concentrations, both of which have been shown to be strong correlates of biotic integrity (U.S. EPA 2010, Palmer et al. 2010). I did not see a strong breakpoint in the data that would suggest a critical threshold beyond which mining activities begin to have more severe impacts on aquatic ecosystems. These conclusions support the conclusions of Palmer et al. (2010) which linked SO$_4^{2-}$ concentrations to known surface water contaminants and impaired macroinvertebrate communities throughout WV, and inferred that these chemical and biological responses resulted from mining activities. My analyses provide a spatially explicit, hydrologic link between the extent of mining activity in a watershed, observed water quality impacts, and ecosystem health.

Future Data Collection – Inconsistencies between the two primary datasets included in this study greatly limited our ability to analyze the shape of the relationship between mining activity and water quality and biotic integrity. Surface mining activity was reported from 1976 to 2005; however, water quality data was collected from 1996 to 2008. The only overlapping time range between these datasets was from 1996 to 2005. As previously mentioned, inconsistencies
between dates of surface mining activity and sample collection resulted in unexpected outliers requiring adjustments upon justification after visually inspecting each of the outliers. Alternatively, the maps I generated can and should underpin future water quality sampling efforts in the region aimed explicitly at linking mining activity to chemical and biological attributes. While ideal but challenging, repeat monitoring at sample locations each year would provide the basis for which a robust temporal trend analysis could be performed. To accomplish this goal, sample locations should first be strategically targeted in order to capture the impacts unique to mining. For example, many of the water quality parameters discussed throughout this study had very few samples recorded in watersheds where mining was in its greatest density, which is where sampling is most sorely needed. Also, most of the interesting outliers found in this study proved to be downstream from reclaimed sites, suggesting that reclamation of surface mining activity could potentially lead to improved water quality; however, more in depth spatial sampling of reclaimed mines is needed in order to address this hypothesis. Second, repeat monitoring of sampling locations should be conducted to develop a basis for understanding trends over time in response to changes in mining activity. The analysis performed in this study as an assessment of the downstream distance impacted by surface mining activity from 1998 to 2003 was only possible for the Mud River. No other rivers within the study area were found to have repeated sampling dates that would allow a similar comparison.

Inverse Distance Weighting – The IDW technique developed throughout this study was a combination of linear and flow-length distance weightings where distance classes were created to account for the differences between in-stream flow and overland flow where both types of flow were simply assumed to have a linear affect on surface water SO$_4^{2-}$ concentrations. A more sophisticated distance weighting technique, perhaps what was presented by Van Sickle and Johnson (2008) which determined the unique influence of each cell on water quality recorded at a downstream sample location, would greatly improve the outcome of this analysis. The Van Sickle and Johnson (2008) study did not require dividing measured distances into distance classes and would resolve issues of applying different weights to adjacent cells classified as surface mining activity that potentially had equal influence on downstream water quality. If the method for determining distance classes was retained for this analysis, extending the riparian corridor to capture the influence of surface mining, specifically valley fills, on ephemeral headwater streams would potentially reduce the amount of sample locations reported as having
elevated SO$_4^{2-}$ concentrations and low percentages of surface mining in their watershed. Capturing surface mining activity in these ephemeral headwater streams would allow higher weights to be applied to this activity and would provide a more accurate relationship between elevated SO$_4^{2-}$ concentrations and extent of mining in a watershed.

Since the aforementioned recommended method of repeat sampling would require time and resources, in the meantime, I recommend making adjustments to the geospatial analysis that was performed throughout this study to account for temporal differences in the onset of mining activity within each watershed as well as applying varying weights to mining activity based on the known year of reclamation. Instead of combining all surface mining activity into one group, activity could be divided into the year it was first classified and correlated only with samples collected after the onset of this activity. Dividing the mining activity data in this way would utilize the complete power of the surface mining dataset obtained from SkyTruth, Inc. which provides surface mining delineations across 4 decades. Incorporating the year of mining activity into the analysis would also allow for reduction in the apparent inconsistencies between the two datasets that were observed throughout this study.

Conductivity – The U.S. EPA recently established a new guidance setting an in-stream maximum level of conductivity to 500 $\mu$S cm$^{-1}$ for Central Appalachian streams (U.S. EPA 2010). The goal of this guidance was to prevent significant and irreversible damage to Appalachian watersheds and is expected to significantly reduce the amount of permitted valley fills. Conductivity is the measure of the ability of water to pass an electrical current and serves as an indication of the presence of dissolved solids (such as SO$_4^{2-}$ or Cl$^-$) in streamwater. Based on the results from this analysis, I found that conductivity in unmined streams is very rarely above the EPA recommended level of 500 $\mu$S cm$^{-1}$, and that virtually all streams with >5% mining activity in their watersheds have conductivities above this value. We found that 18 of the 134 streams in the “unmined” category had conductivities higher than 500 $\mu$S cm$^{-1}$ and that 41 of these streams had conductivities over 300 $\mu$S cm$^{-1}$, suggesting that there are other important sources of conductivity in southern WV streams. These streams typically had a Conductivity:SO$_4^{2-}$ ratio well above what we measured in mining impacted streams – suggesting that in these streams Cl$^-$ is a dominant source of ionic strength. Conductivity is more closely associated with SO$_4^{2-}$ levels in areas receiving the greatest impact from mining than with Cl$^-$ (See Appendix, Figure 2). While SO$_4^{2-}$ may be a more precise indicator of the extent of impact coming from
active surface mines, conductivity can serve as an umbrella parameter accounting for several constituents released during the mining process. Potentially, a two-tiered approach could be applied when analyzing stream water quality where conductivity measurements, which are quicker and cheaper than SO$_4^{2-}$ measurements, could be performed to target areas of concern. Then, for samples where high conductivity is measured, subsequent SO$_4^{2-}$ measurements could be performed to determine whether surface mining or other activities (e.g. road salting) are the dominant drivers of stream conductivity. This two tiered analysis would provide important guidance for managing and mitigating elevated conductivity problems in Central Appalachian streams.

**Acknowledgements**

I thank Dr. Emily Bernhardt in the Biology Department at Duke University for providing helpful comments and exceptional guidance throughout the duration of this study. I also thank Professor of Landscape Ecology, Dr. Dean Urban, Geospatial Analysis Instructor, John Fay, and graduate student, Kayleigh Somers in the Nicholas School of the Environment at Duke University for providing additional guidance. Data of surface mining activity in southwestern WV was provided by SkyTruth, Inc. (http://www.skytruth.org/); I thank SkyTruth, Inc. for offering this data. Funding for this study was provided by the Stanback Foundation.
Tables and Figures:

Figure 1: Overview of study area location within WV. The Coal and Guyandotte River Basins were selected for this study because they contained the greatest density of surface mining activity relative to the rest of the state. The highest density of mining (based on a visual inspection) is in Boone, Logan, and Mingo counties within WV.
Figure 2: Water quality sampling locations within the study area (Coal and Guyandotte River Basins, WV). Lower order streams are not displayed, but coincide with the sampling locations. \(\text{SO}_4^{2-}\) concentration is coded by color, with darker colors representing elevated \(\text{SO}_4^{2-}\). Notice that elevated \(\text{SO}_4^{2-}\) concentrations in surface waters appear to be spatially related to surface mining locations. Surface mines in this map include valley fill locations.
Table 1: Distance class divisions used to assign new distance values to cells within the 30m riparian corridor for each watershed.

<table>
<thead>
<tr>
<th>Distance class divisions</th>
<th>New distance values assigned to every cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 1,000</td>
<td>100</td>
</tr>
<tr>
<td>1,001 - 3,000</td>
<td>200</td>
</tr>
<tr>
<td>3,001 - 7,000</td>
<td>400</td>
</tr>
<tr>
<td>7,001 - 15,000</td>
<td>800</td>
</tr>
<tr>
<td>15,001 - 31,000</td>
<td>1,600</td>
</tr>
<tr>
<td>&gt; 31,000</td>
<td>3,200</td>
</tr>
</tbody>
</table>

Table 2: Distance class divisions used to assign weights to surface mining activity within each watershed and the corresponding weights applied throughout the IDW calculation.

<table>
<thead>
<tr>
<th>Distance class divisions</th>
<th>Inverse-distance weight applied to surface mining activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 100</td>
<td>1/100</td>
</tr>
<tr>
<td>101 - 250</td>
<td>1/250</td>
</tr>
<tr>
<td>251 - 500</td>
<td>1/500</td>
</tr>
<tr>
<td>501 - 1,000</td>
<td>1/1,000</td>
</tr>
<tr>
<td>1,001 - 2,000</td>
<td>1/2,000</td>
</tr>
<tr>
<td>2,001 - 5,000</td>
<td>1/5,000</td>
</tr>
<tr>
<td>5,001 - 10,000</td>
<td>1/10,000</td>
</tr>
<tr>
<td>&gt; 10,000</td>
<td>1/20,000</td>
</tr>
</tbody>
</table>
Figure 3: Process for creating distance classes to represent in-stream and overland flow-length for each watershed. First, flow-length to the sample location was calculated for each cell within the 30m riparian buffer (A.) and for each cell within the entire watershed (B.) for each of the delineated watersheds. Flow-length values for the riparian corridor were divided into distance classes and assigned new values (Table 1). Overland flow distance values for each cell located within the 30m riparian corridor were replaced with these newly assigned flow distance values. All cells for each watershed were then divided into unequal-interval distance classes based on each cell’s overland flow-length or newly assigned distance value (C.). Total area of surface mining activity was then calculated for each resulting distance class (D.) and inverse distance weights (Table 2) were applied to determine the % IDW surface mining activity within each watershed. The distance classes for distances ranging from 5,001 – 10,000m or >10,000m were not included in this figure because this watershed did not contain flow distances within these ranges.
Figure 4: Surface water $\text{SO}_4^{2-}$ concentrations plotted against the extent of surface mining activity within each watershed. There is an apparent positive relationship between these two variables, but some questionable outliers do exist. Outliers represented as open circles are those that were found to be influenced by underground mining activity within their watershed. Underground mining was not accounted for in this analysis as only surface mining activity was delineated by SkyTruth, Inc. Outliers represented as shaded triangles were found to be downstream of surface mining activity that had been reclaimed prior to either 1985 or 1995. The distribution for $\text{SO}_4^{2-}$ concentrations for all sampling locations having 0% mining in their watershed is plotted as a solid line (the mean of the distribution) encompassed by two dotted lines (the upper and lower bounds of the 95% confidence interval). Based on the observed narrow spread of $\text{SO}_4^{2-}$ at background levels (0% mining), it is apparent that surface mining activity greatly influences surface water $\text{SO}_4^{2-}$ concentrations.
Table 3: Hypothesis test results comparing water quality parameters between background levels (0% mining) and all other levels of mining included in this study. Exp = exponential function.

**Log (Sulfate) ~ Extent of Mining**

<table>
<thead>
<tr>
<th>Mining</th>
<th>95% Confidence Interval</th>
<th>Difference in Means</th>
<th>Exp of Difference</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper Bound</td>
<td>Lower Bound</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0% vs. &lt; 5%</td>
<td>-0.87</td>
<td>-1.41</td>
<td>1.14</td>
<td>3.11</td>
</tr>
<tr>
<td>0% vs. 5-10%</td>
<td>-1.57</td>
<td>-2.17</td>
<td>1.87</td>
<td>6.49</td>
</tr>
<tr>
<td>0% vs. 10-20%</td>
<td>-1.61</td>
<td>-2.35</td>
<td>1.98</td>
<td>7.25</td>
</tr>
<tr>
<td>0% vs. 20-40%</td>
<td>-2.23</td>
<td>-2.99</td>
<td>2.61</td>
<td>13.5</td>
</tr>
<tr>
<td>0% vs. &gt;40%</td>
<td>-1.79</td>
<td>-3.21</td>
<td>2.5</td>
<td>12.22</td>
</tr>
</tbody>
</table>

**Log (Conductivity) ~ Extent of Mining**

<table>
<thead>
<tr>
<th>Mining</th>
<th>95% Confidence Interval</th>
<th>Difference in Means</th>
<th>Exp of Difference</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper Bound</td>
<td>Lower Bound</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0% vs. &lt; 5%</td>
<td>-0.67</td>
<td>-1.09</td>
<td>0.88</td>
<td>2.41</td>
</tr>
<tr>
<td>0% vs. 5-10%</td>
<td>-1.08</td>
<td>-1.5</td>
<td>1.29</td>
<td>3.64</td>
</tr>
<tr>
<td>0% vs. 10-20%</td>
<td>-1.12</td>
<td>-1.7</td>
<td>1.41</td>
<td>4.1</td>
</tr>
<tr>
<td>0% vs. 20-40%</td>
<td>-1.43</td>
<td>-2.06</td>
<td>1.74</td>
<td>5.71</td>
</tr>
<tr>
<td>0% vs. &gt;40%</td>
<td>-1.46</td>
<td>-2.23</td>
<td>1.85</td>
<td>6.33</td>
</tr>
</tbody>
</table>

**Log (Conductivity:Sulfate Ratio) ~ Extent of Mining**

<table>
<thead>
<tr>
<th>Mining</th>
<th>95% Confidence Interval</th>
<th>Difference in Means*</th>
<th>Exp of Difference*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper Bound</td>
<td>Lower Bound</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0% vs. &lt; 5%</td>
<td>0.41</td>
<td>0.11</td>
<td>0.26</td>
<td>1.29</td>
</tr>
<tr>
<td>0% vs. 5-10%</td>
<td>0.77</td>
<td>0.38</td>
<td>0.57</td>
<td>1.77</td>
</tr>
<tr>
<td>0% vs. 10-20%</td>
<td>0.77</td>
<td>0.37</td>
<td>0.57</td>
<td>1.77</td>
</tr>
<tr>
<td>0% vs. 20-40%</td>
<td>1.05</td>
<td>0.67</td>
<td>0.86</td>
<td>2.37</td>
</tr>
<tr>
<td>0% vs. &gt;40%</td>
<td>1.11</td>
<td>0.67</td>
<td>0.89</td>
<td>2.43</td>
</tr>
</tbody>
</table>

**Log (Chloride) ~ Extent of Mining**

<table>
<thead>
<tr>
<th>Mining</th>
<th>95% Confidence Interval</th>
<th>Difference in Means</th>
<th>Exp of Difference</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper Bound</td>
<td>Lower Bound</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0% vs. &lt; 5%</td>
<td>-0.26</td>
<td>-1</td>
<td>0.61</td>
<td>1.84</td>
</tr>
<tr>
<td>0% vs. 5-10%</td>
<td>-0.2</td>
<td>-0.96</td>
<td>0.58</td>
<td>1.78</td>
</tr>
<tr>
<td>0% vs. 10-20%</td>
<td>-0.3</td>
<td>-1.18</td>
<td>0.74</td>
<td>2.1</td>
</tr>
<tr>
<td>0% vs. 20-40%</td>
<td>-0.05</td>
<td>-1.54</td>
<td>0.8</td>
<td>2.22</td>
</tr>
<tr>
<td>0% vs. &gt;40%</td>
<td>0.51</td>
<td>-2.23</td>
<td>0.86</td>
<td>2.36</td>
</tr>
</tbody>
</table>

*For conductivity:sulfate ratios, the mean value for 0% surface mining was greater than means estimated for all other mining categories.
Figure 5: Se concentration plotted against the extent of surface mining activity in each watershed included in this study. There is a positive relationship between the two variables, but more sampling for Se is needed in areas where mining is in its greatest density.
Figure 6: Conductivity plotted against the extent of surface mining activity in each watershed included in this study. A positive relationship does exist between these two variables. Outliers represented as open circles are those that were found to be influenced by underground mining activity within their watershed. Underground mining was not accounted for in this analysis as only surface mining activity was delineated by SkyTruth, Inc. Shaded triangles represent samples that were found to be downstream of surface mining activity that had been reclaimed prior to either 1985 or 1995. The distribution for conductivity for all sampling locations with 0% mining in their watershed is plotted as a solid line (the mean of the distribution) encompassed by two dotted lines (the upper and lower bounds of the 95% confidence interval).
**Figure 7**: Conductivity: $\text{SO}_4^{2-}$ ratio plotted against the extent of surface mining activity in each watershed included in this study. For areas where surface mining is in its greatest density (either 20-40% or >40%), $\text{SO}_4^{2-}$ and conductivity are highly correlated resulting in very low ratio values implying that $\text{SO}_4^{2-}$ could be the dominant constituent to conductivity. In areas where the extent of upstream surface mining activity is low (either 0% or <5%), it appears as if there may be sources unrelated to mining that contribute to conductivity. Open circles represent samples that were found to be influenced by underground mining activity within their watershed. Underground mining was not accounted for in this analysis as only surface mining activity was delineated by SkyTruth, Inc. Shaded triangles represent samples that were found to be downstream of surface mining activity that had been reclaimed prior to either 1985 or 1995. The conductivity: $\text{SO}_4^{2-}$ ratio distribution for all sampling locations with 0% mining in their watershed is plotted as a solid line (the mean of the distribution) encompassed by two dotted lines (the upper and lower bounds of the 95% confidence interval).
Figure 8: Conductivity (us cm$^{-1}$), SO$_4^{2-}$ (mg L$^{-1}$) and Cl$^-$ (mg L$^{-1}$) plotted against the extent of surface mining activity upstream from sampling locations. Elevated conductivity values associated with a greater density of mining within a watershed appear to be a result of increased SO$_4^{2-}$ concentrations at these sampling locations when compared to Cl$^-$. All sites with underground mining and <5% surface mining were removed from the datasets prior to constructing these plots. All sites that had greater than 10% surface mining in their watershed and found to have been reclaimed before 1995 were also removed.
Table 4: Hypothesis test results for macroinvertebrate metric values for both the family based WVSCI and the genus based GLIMPSS. Tests were performed to compare metric values for unmined watersheds to those impacted by mining.

### WVSCI ~ Extent of Mining

<table>
<thead>
<tr>
<th>Mining</th>
<th>95% Confidence Interval</th>
<th>Difference in Means*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upper Bound</td>
<td>Lower Bound</td>
<td></td>
</tr>
<tr>
<td>0% vs. &lt; 5%</td>
<td>9.69</td>
<td>0.44</td>
<td>5.07</td>
</tr>
<tr>
<td>0% vs. 5 - 10%</td>
<td>15.57</td>
<td>5.57</td>
<td>10.57</td>
</tr>
<tr>
<td>0% vs. 10 - 20%</td>
<td>14.42</td>
<td>2.33</td>
<td>8.37</td>
</tr>
<tr>
<td>0% vs. 20-40%</td>
<td>17.47</td>
<td>1.05</td>
<td>16.42</td>
</tr>
<tr>
<td>0% vs. &gt; 40%</td>
<td>28.27</td>
<td>5.78</td>
<td>22.53</td>
</tr>
</tbody>
</table>

### GLIMPSS ~ Extent of Mining

<table>
<thead>
<tr>
<th>Mining</th>
<th>95% Confidence Interval</th>
<th>Difference in Means*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upper Bound</td>
<td>Lower Bound</td>
<td></td>
</tr>
<tr>
<td>0% vs. &lt; 5%</td>
<td>14.37</td>
<td>3.78</td>
<td>9.07</td>
</tr>
<tr>
<td>0% vs. 5 - 10%</td>
<td>23.91</td>
<td>14.45</td>
<td>19.18</td>
</tr>
<tr>
<td>0% vs. 10 - 20%</td>
<td>21.17</td>
<td>7.27</td>
<td>14.21</td>
</tr>
<tr>
<td>0% vs. 20-40%</td>
<td>26.80</td>
<td>13.00</td>
<td>19.89</td>
</tr>
<tr>
<td>0% vs. &gt; 40%</td>
<td>36.62</td>
<td>14.84</td>
<td>21.78</td>
</tr>
</tbody>
</table>

*Estimated mean metric values for unmined watersheds (0% mining) were consistently greater than metric values recorded for all watersheds impacted by surface mining activity.
Figure 9: Both the genus level (A., GLIMPSS) and family level (B., WVSCI) macroinvertebrate metrics recorded for each sample location are plotted against the extent of surface mining activity in each watershed. While an apparent negative relationship exists between these variables, it is evident that the genus level index is more sensitive to surface mining activity. Outliers represented as open circles are those that were found to be influenced by underground mining activity within their watershed. Underground mining was not accounted for in this analysis as only surface mining activity was delineated by SkyTruth, Inc. Shaded triangles represent samples that were found to be downstream of surface mining activity that had been reclaimed prior to either 1985 or 1995. The distribution for GLIMPSS (A.) or WVSCI (B.) for all sampling locations with 0% mining in their watershed is plotted as a solid line (the mean of the distribution) encompassed by two dotted lines (the upper and lower bounds of the 95% confidence interval).
Figure 10: Sample locations along the Mud River for both 1998 and 2003. Notice the increase in surface mining activity from 1995 to 2005 within the watershed nearly doubles. \( \text{SO}_4^{2-} \) calculations were found to decline as downstream distance from mining site increased; however, \( \text{SO}_4^{2-} \) at these sampling locations increased from 1998 to 2003 as a general response to a doubling in surface mining activity during this time period.
Figure 11: Sample locations along the Coal River for 1997 with corresponding SO$_4^{2-}$ concentrations. Surface mining activity is displayed in two stages: 1) activity up to 1995, occurring before the water quality samples were collected (black), and 2) new activity in 2005, after water quality samples were collected (grey). Downstream distance impacted within the Coal River Basin is extensive; the farthest downstream sample location had a SO$_4^{2-}$ concentration recorded as 353mg L$^{-1}$; this concentration exceeds the mean SO$_4^{2-}$ concentration of 57.55 mg L$^{-1}$ found throughout this study to be a background level (resulting from samples with 0% surface mining in their watershed). The influence of surface mining activity and dilutions is apparent in the varying SO$_4^{2-}$ concentrations recorded throughout this river basin. (See Appendix for enlarged copy of figure).
References


Appendix

Figure 1: Comparison of generated stream network to the NHD Plus 1:24,000-scale flowlines. The stream network generated for this analysis contains a larger quantity of smaller tributaries when compared to the NHD Plus network. These tributaries were retained for the analysis because several water quality sample locations were found to coincide with these streams.
Figure 2: A SO$_4^{2-}$:Cl$^-$ ratio comparison was performed between samples having > 300 µS cm$^{-1}$ and < 10% surface mining activity in their watersheds to samples having > 300 µS cm$^{-1}$ and >10% surface mining activity in their watersheds. The objective for this comparison was to determine which constituent (either SO$_4^{2-}$ or Cl$^-$) is the dominant contributor for elevated conductivity levels at both low and high levels of mining. A hypothesis test (two sample t-test) was performed on log transformed SO$_4^{2-}$:Cl$^-$ ratios based on the assumptions of normality and independency. The assumptions were not entirely met with these data due to sampling patterns throughout the stream network. However, based on this comparison, SO$_4^{2-}$:Cl$^-$ ratios were found to be significantly lower for samples with < 10% surface mining influence than they were for samples with >10% surface mining influence when only samples having conductivity values >300 µS cm$^{-1}$ are considered (p-value = 0.0008). Based on this result, it can be concluded that elevated conductivity levels at areas that are minimally influenced by mining (<10% surface mining) are dominated by Cl$^-$ relative to SO$_4^{2-}$. 
Table 1: Sample locations labeled by sampling station identification number having \( \text{SO}_4^{2-} \) concentrations < 300 mg L\(^{-1}\) and %IDW surface mining activity >10%. For sample locations listed as “Mining date was 2005, after sample was collected,” % IDW surface mining activity was adjusted to zero after justification based on visual inspection of the mining activity within each watershed. Sample locations listed as being reclaimed were plotted individually in order to assess the distribution of \( \text{SO}_4^{2-} \) concentrations following reclamation.

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Sample Date</th>
<th>Sulfate (mg L(^{-1}))</th>
<th>% IDW surface mining activity</th>
<th>Description of watershed based on visual inspection</th>
</tr>
</thead>
<tbody>
<tr>
<td>7367</td>
<td>5/9/05</td>
<td>35.00</td>
<td>10.20</td>
<td>Mining date was 1976 and was reclaimed before 1985</td>
</tr>
<tr>
<td>5247</td>
<td>9/5/00</td>
<td>7.09</td>
<td>14.36</td>
<td>Mining date was 1976 and was reclaimed before 1985</td>
</tr>
<tr>
<td>5044</td>
<td>9/6/00</td>
<td>105.08</td>
<td>16.25</td>
<td>Mining date was 1976 and was reclaimed before 1985</td>
</tr>
<tr>
<td>5043</td>
<td>9/6/00</td>
<td>140.20</td>
<td>40.55</td>
<td>Mining date was 1976 and was reclaimed before 1985</td>
</tr>
<tr>
<td>1326</td>
<td>5/3/07</td>
<td>238.00</td>
<td>33.37</td>
<td>Mining date was 1976 and was reclaimed before 1985</td>
</tr>
<tr>
<td>5030</td>
<td>5/4/04</td>
<td>31.90</td>
<td>12.44</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>982</td>
<td>10/8/97</td>
<td>59.00</td>
<td>14.83</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>5021</td>
<td>8/29/00</td>
<td>76.94</td>
<td>18.10</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>1287</td>
<td>9/23/97</td>
<td>140.00</td>
<td>19.34</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>1127</td>
<td>10/6/97</td>
<td>120.00</td>
<td>20.39</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>1055</td>
<td>9/18/97</td>
<td>160.00</td>
<td>21.93</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>1057</td>
<td>9/18/97</td>
<td>140.00</td>
<td>22.70</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>5193</td>
<td>5/27/03</td>
<td>102.00</td>
<td>31.56</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>5025</td>
<td>9/6/00</td>
<td>188.63</td>
<td>41.11</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>927</td>
<td>9/25/97</td>
<td>170.00</td>
<td>57.11</td>
<td>Mining date was 2005, after the sample was collected</td>
</tr>
<tr>
<td>1128</td>
<td>4/24/03</td>
<td>176.00</td>
<td>10.93</td>
<td>Mining dates were 1976 and 1985; 1976 mining reclaimed before 1985; 1985 mining reclaimed before 1995</td>
</tr>
<tr>
<td>5038</td>
<td>9/7/00</td>
<td>150.00</td>
<td>46.98</td>
<td>Mining dates were 1976 or 1985; 1976 mining reclaimed before 1985; 1985 mining reclaimed before 1995</td>
</tr>
<tr>
<td>5265</td>
<td>8/29/00</td>
<td>220.91</td>
<td>12.77</td>
<td>Mining dates were 1985 or earlier; 1976 mining reclaimed before 1995; 1985 mining reclaimed before 1995</td>
</tr>
</tbody>
</table>
Table 2: Sample locations labeled by sampling station identification number having $\text{SO}_4^{2-}$ concentrations $> 300 \text{ mg L}^{-1}$ and %IDW surface mining activity $< 5\%$. For all sample locations listed, visual inspections were performed to assess whether high $\text{SO}_4^{2-}$ concentrations could be a result of underground mining activity, processing plants, and/or injection sites within the watershed. Also, some watersheds contained valley fills that did not receive weights that corresponded to the 30m riparian corridor because the stream layer did not account for some of the ephemeral headwater streams filled by these valley fills.

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Sample Date</th>
<th>Sulfate (mg L$^{-1}$)</th>
<th>%IDW surface mining activity</th>
<th>Description of watershed based on visual inspection</th>
</tr>
</thead>
<tbody>
<tr>
<td>5069</td>
<td>9/11/00</td>
<td>380.00</td>
<td>1.52</td>
<td>Valley fill present that did not receive weight applied to streams during IDW; one preparation plant upstream (expires in 2013); two reprocessing plants upstream (expired in 1991 and 1992)</td>
</tr>
<tr>
<td>5268</td>
<td>8/30/00</td>
<td>473.65</td>
<td>1.72</td>
<td>Underground mining activity present</td>
</tr>
<tr>
<td>5028</td>
<td>8/29/00</td>
<td>309.44</td>
<td>2.55</td>
<td>Underground mining activity present</td>
</tr>
<tr>
<td>5230</td>
<td>8/29/00</td>
<td>390.00</td>
<td>3.83</td>
<td>Valley fill present that did not receive weight applied to streams during IDW</td>
</tr>
<tr>
<td>5177</td>
<td>8/23/00</td>
<td>410.00</td>
<td>4.26</td>
<td>Underground mining activity present; valley fill present that did not receive weight applied to streams during IDW; reprocessing plant upstream (expired in 1996)</td>
</tr>
<tr>
<td>5222</td>
<td>8/31/00</td>
<td>320.00</td>
<td>4.57</td>
<td>Underground mining activity present; two underground mining injection sites upstream</td>
</tr>
</tbody>
</table>
Figure: Enlarged copy of Figure 11 from report.
Models created using ESRI’s ArcGIS (v.9.3) model builder:

Model 1: Flow length estimations and distance class divisions performed in ArcGIS (ESRI, Inc.)
Model 2: An iterative process was performed using this model to delineate a unique watershed for each sampling location.

Selected sample points one at a time to perform flow length steps in IDW_FlowLength tool.
Python Script: For flow-length and IDW analysis in GIS

For Model 1:
# =================================================================================
# IDW_FlowLength_scpt.py
# Created on: Wed Apr 21 2010 03:49:03 PM
# (generated by ArcGIS/ModelBuilder)
# Usage: IDW_FlowLength_scpt <Landcover_Area_dbf> <Combined_Flow_Reclass>
# <Combined_Corridor_and_Watershed_Flow> <WQ_Sample>
# =================================================================================

# Import system modules
import sys, string, os, arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create()

# Check out any necessary licenses
gp.CheckOutExtension("spatial")

# Load required toolboxes...
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Spatial Analyst Tools.tbx")

# Set the Geoprocessing environment...
gp.XYResolution = ""
gp.scratchWorkspace = "F:\MountaintopMining\scratch"
gp.MTolerance = ""
gp.randomGenerator = "0 ACM599"
gp.outputCoordinateSystem = "PROJCS[NAD_1983_Albers',GEOGCS[GCS_North_American_1983',DATUM[D_North_American_1983',SPHEROID[GRS_1980',6378137.0,298.257222101'],PRIMEM[Greenwich',0.0],UNIT[Degree',0.0174532925199433'],PROJECTION[Albers'],PARAMETER[False_Easting',0.0],PARAMETER[False_Northing',0.0],PARAMETER[Central_Meridian','-96.0'],PARAMETER[Standard_Parallel_1',29.5],PARAMETER[Standard_Parallel_2',45.5],PARAMETER[Latitude_Of_Origin',23.0],UNIT[Meter',1.0],VERTCS[Unknown VCS from ArcInfo Workstation',VDATUM[Unknown'],PARAMETER[Vertical_Shift',0.0],PARAMETER[Directi on',1.0],UNIT[User Defined Unit',0.01']]"
gp.snapRaster = "F:\MountaintopMining\scratch\elev_5_c"
gp.outputZFlag = "Same As Input"
gp.qualifiedFieldNames = "true"
gp.extent = "DEFAULT"
gp.XYTolerance = ""
gp.cellSize = "30"
gp.outputZValue = ""
gp.outputMFlag = "Same As Input"
gp.geographicTransformations = ""
gp.ZResolution = ""
gp.mask = "F:\MountaintopMining\Scratch\WS_Coal_UPGuy_alb.shp"
gp.workspace = "F:\MountaintopMining\DATA"
gp.MResolution = ""
gp.ZTolerance = ""

# Script arguments...
Landcover_Area_dbf = sys.argv[1]
if Landcover_Area_dbf == '#':
    Landcover_Area_dbf = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_Area\27306_area2.dbf"  # provide a default value if unspecified

Combined_Flow_Reclass = sys.argv[2]
if Combined_Flow_Reclass == '#':
    Combined_Flow_Reclass = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_ws\27306_fl_rc"  # provide a default value if unspecified

Combined_Corridor_and_Watershed_Flow = sys.argv[3]
if Combined_Corridor_and_Watershed_Flow == '#':
    Combined_Corridor_and_Watershed_Flow = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_ws\27306_fl_con"  # provide a default value if unspecified

WQ_Sample = sys.argv[4]
if WQ_Sample == '#':
    WQ_Sample = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_ras\27306.shp"  # provide a default value if unspecified

# Local variables...
Sample_Watershed = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_ws\27306_ws"
Flow_Direction__1_ = "F:\MountaintopMining\Scratch\Coal_UPGuy\fdr_cug"
Watershed_Flow_Direction = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_ws\27306_fdr"
Flow_Direction__3_ = "F:\MountaintopMining\Scratch\Coal_UPGuy\fdr_cug"
Watershed_Flow_Length = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_ws\27306_fl"
Riparian_Corridor_for_Sample_Watershed = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_ws\27306_strm2"
Corridor_Flow_Direction = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_ws\27306_strmfdr"
Flow_Direction__2_ = "F:\MountaintopMining\Scratch\Coal_UPGuy\fdr_cug"
Corridor_Flow_Length = "F:\MountaintopMining\Scratch\Coal_UPGuy\Sulf_ws\27306_strmfl"
Corridor_Flow_Reclass__1_ = "F:\MountaintopMining\Scratch\Coal_UpGuy\Sulf_ws\27306_stfl_rc"
Corridor_Flow_Reclass__2_ = "F:\MountaintopMining\Scratch\Coal_UpGuy\Sulf_ws\27306_stfl_1"
Landcover = "F:\MountaintopMining\Scratch\Coal_UpGuy\vf_cug_6"
Riparian_Corridor__30m_ = "F:\MountaintopMining\Scratch\strm_buff_r30"

# Process: Watershed...
gp.Watershed_sa(Flow_Direction__1_, WQ_Sample, Sample_Watershed, "sample")

# Process: Extract by Mask (1)...
gp.ExtractByMask_sa(Riparian_Corridor__30m_, Sample_Watershed, Riparian_Corridor_for_Sample_Watershed)

# Process: Extract by Mask (2)...
gp.ExtractByMask_sa(Flow_Direction__2_, Riparian_Corridor_for_Sample_Watershed, Corridor_Flow_DIR)

# Process: Flow Length (1)...
gp.FlowLength_sa(Corridor_Flow_DIR, Corridor_Flow_LENGTH, "DOWNSTREAM", "")

# Process: Reclassify (1)...
gp.Reclassify_sa(Corridor_Flow_LENGTH, "Value", "0 1000 100;1000 3000 200;3000 7000 400;7000 15000 800;15000 31000 1600;31000 148690.578125 3200;NODATA 0",
Corridor_Flow_Reclass__1_, "DATA")

# Process: Extract by Mask (3)...
gp.ExtractByMask_sa(Corridor_Flow_Reclass__1_, Sample_Watershed, Corridor_Flow_Reclass__2_)

# Process: Extract by Mask (4)...
gp.ExtractByMask_sa(Flow_Direction__3_, Sample_Watershed, Watershed_Flow_DIR)

# Process: Flow Length (2)...
gp.FlowLength_sa(Watershed_Flow_DIR, Watershed_Flow_LENGTH, "DOWNSTREAM", "")

# Process: Con...
gp.Con_sa(Corridor_Flow_Reclass__2_, Corridor_Flow_Reclass__2_, Combined_Corridor_and_Watershed_FLOW, Watershed_Flow_LENGTH, ""VALUE" = 100 OR "VALUE" = 200 OR "VALUE" = 400 OR "VALUE" =800 OR "VALUE" = 1600 OR "VALUE" = 3200")

# Process: Reclassify (2)...
gp.Reclassify_sa(Combined_Corridor_and_Watershed_Flow, "Value", "0 100 100 250 250;250 500 500;500 1000 1000;1000 2000 2000;2000 5000 5000;5000 10000 10000;10000 150076.90625 10001", Combined_Flow_Reclass, "DATA")

# Process: Tabulate Area...
gp.TabulateArea_sa(Combined_Flow_Reclass, "VALUE", Landcover, "VALUE", Landcover_Area_dbf, "30")

For Model 2:

# IDW_FlowLength_Iteration_scpt.py
# Created on: Wed Apr 21 2010 03:55:31 PM
# (generated by ArcGIS/ModelBuilder)

# Import system modules
import sys, string, os, arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create()

# Check out any necessary licenses
gp.CheckOutExtension("spatial")

# Load required toolboxes...
gp.AddToolbox("F:/MountaintopMining/MTR_Mining.tbx")
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Analysis Tools.tbx")

# Local variables...
SelPoint = "in_memory\SelPoint"
WQ_Sample_File = "F:\MountaintopMining\Scratch\Coal_UpGuy\sulf_pnt_CUG.shp"
Table_N__dbf = "F:\MountaintopMining\Scratch\Coal_UpGuy\Sulf_Area\Table%N%.dbf"
WS__N_ = "F:\MountaintopMining\Scratch\Coal_UpGuy\Sulf_ws\WS_%N%"
WSfl__N_ = "F:\MountaintopMining\Scratch\Coal_UpGuy\Sulf_ws\WSfl_%N%"

# Process: Select...
gp.Select_analysis(WQ_Sample_File, SelPoint, ";id\" = %N% + 1")

# Process: IDW_FlowLength...
gp.toolbox = "F:/MountaintopMining/MTR_Mining.tbx";
gp.flowpathcoaluguy22(Table_N__dbf, WS__N_, WSfl__N_, SelPoint)