How Well Does Electrical Conductivity Predict Stream Impairment?

Paul F. Ziemkiewicz
West Virginia Water Research Institute, West Virginia University, Morgantown, West Virginia

Steven Northway
Faculty of Forestry, University of British Columbia, Vancouver, British Columbia

ABSTRACT

Section 303 of the Federal Clean Water Act (CWA) mandates that states develop a process for defining stream impairment. Criteria can be 'numeric,' i.e., water quality parameters, or 'narrative,' which generally involves more qualitative and biological indices such as the West Virginia Stream Condition Index (WVSCI). Recent guidance from USEPA has encouraged the use of an electrical conductivity (EC) based criteria (300 µS/cm) to estimate WVSCI impairment. Despite the enormous regulatory and economic repercussions of this approach, little work has been done to determine the quality of the resulting estimates. This paper reports our efforts to critically evaluate the quality of EC as a predictor of WVSCI scores. We present a case study using West Virginia stream segments to identify error rates as a function of various EC criteria. We found that an EC based regression model, while highly statistically significant, only accounted for 26 percent of the variation of WVSCI scores. This highlights the difference between a statistically significant model and a model that can actually predict a specific outcome with confidence. No evidence was found that an EC value less than 300 µS/cm is a reliable predictor of WVSCI values that meet the West Virginia narrative criteria for stream impairment. This implies that factors other than EC play a significant role in determining WVSCI ranking and that efforts to lower a stream's EC alone will not reliably improve its condition.

INTRODUCTION

Under section 303 of the Federal Clean Water Act, states are obliged to develop and periodically update a list of impaired steam segments. States may employ water quality based numerical standards or biologically based narrative standards to determine impairment. Impaired stream segments may be subject to an improvement program which generally consists of a determination of the causal pollutant(s), a determination of pollutant loadings and sources and, finally, an implementation plan to improve the stream to the extent that it will no longer be impaired. Implementation may consist of restricting loads of the causal pollutants from permitted discharges or denying permits altogether. This may cause significant economic impact on discharge permit holders as their existing treatment systems may require reconfiguration costing tens of millions of dollars. The expectation is that removal of the putative pollutant will correct the impairment. The purpose of this study is to estimate the probability that a change in EC will result in an improvement in the WVSCI score for a stream segment.

West Virginia's impairment level is ideally considered to be a WVSCI score of 68. However, accounting for sampling error, the State's working criterion for impairment is a WVSCI score of 60.6: "For purposes of Federal Clean Water Act compliance under section 303d a West Virginia Stream Condition Index (WVSCI) score of 68 is considered the threshold for impairment. In

recognition of the sampling error (7.4) when using the recommended procedure, the State classifies streams impaired at a WVSCI score of 60.6" (WVDEP 2012). Both criteria are employed in this study as appropriate.

The WVSCI based impairment definition leaves the challenge of developing individual discharge standards under section 402 of the CWA such that the additional discharge will not cause the stream to become impaired. There have been attempts to solve this problem by using water quality parameters such as electrical conductivity (EC), dissolved solids or specific ions such as sulfate to predict stream condition index (SCI) scores. Most attempts to correlate water quality with biological assessments, however, rely on EC since it is easy for field samplers to collect and involves no laboratory analytical cost. As a result, large databases containing both WVSCI scores and EC values are available for many stream segments. Recent guidance from USEPA has encouraged the use of an EC criterion (300 µS/cm) to estimate stream impairment (USEPA 2010). In this sense, EC becomes surrogate for WVSCI.

Despite the enormous regulatory and economic repercussions of this approach, little work has been done to determine the quality of the resulting estimates. Quality has two components in this context: the statistical significance of the prediction model and the practical implications of the error range around the predicted values.

Much of the discussion surrounding the effects of surface mining on stream condition revolves around the relationship between EC and WVSCI. It is understood that the same EC value can result from a wide variety of chemistries and that not all combinations yield the same stream benthic effects. Elphick et al. (2011) found a strong negative interaction between hardness and sulfate concentration, reporting that the interaction of sulfate toxicity and water hardness resulted in separate values for soft (10–40 mg/L), moderately hard (80–100 mg/L) and hard water (160–250 mg/L). The resulting values were 129, 644 and 725 mg/L sulfate, respectively, following the SSD approach, and 75, 625 and 675 mg/L sulfate, following the safety factor approach. Dissolved salt concentrations and EC in Appalachian mining district streams are invariably highest during low stream flow periods (Petty et al. 2010), since mines provide significant base flow.

Bernhardt et al. (2012) concluded that EC was a good predictor of WVSCI, suggesting an EC impairment threshold of 308 µS/cm. Their reported upper and lower confidence intervals (CI) were asymmetrical about the mean at 365 and 245 µS/cm at 95 percent probability, or about 22 percent of the mean value. They made no mention of the classification error rate that would have resulted from the application of their model.

METHODOLOGY

This paper reports on our efforts to critically evaluate the quality of EC as a predictor of WVSCI scores. We present a case study using West Virginia stream segments to identify error rates as a function of various EC criteria. The first objective of this study was to identify whether a correlation exists between the WVSCI and EC. Secondly, we identified a best-fit model for predicting WVSCI as a function EC. Third, we identified the precision with which EC predicts WVSCI.

We used a data set consisting of 222 samples containing both EC and WVSCI data for West Virginia stream segments for this study. It was collected by the staff of West Virginia Department of Environmental Protection (WVDEP) as part of the State's Watershed Assessment Program. It will be referred to as the WVDEP data set. This data set was also the basis of a study by Bernhardt et al. (2012). The data were a subset of a larger database collected by WVDEP and winnowed by the agency in an attempt to remove sampling stations with confounding factors. The sampling stations included pristine streams through unmined, but developed, watersheds to watersheds with increasing mining intensities. The data set was collected by WVDEP staff between 1996 and 2008 during the dry period of the West Virginia hydrologic year, April through October.

We used two key parameters for our evaluation: (1) WVSCI, which classifies stream insect families into tolerant and intolerant categories

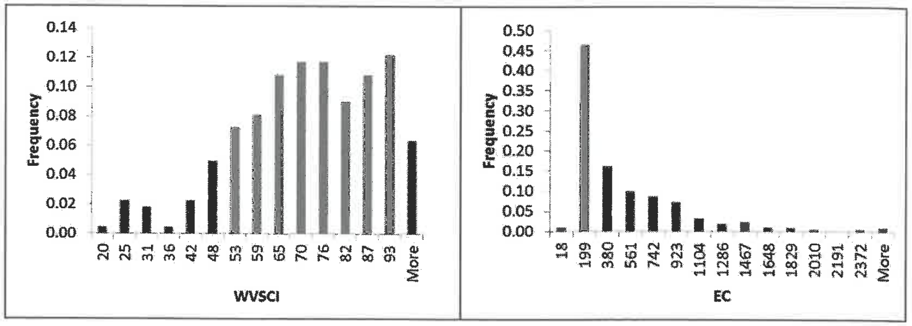

Figure 1. Sample distribution as a function of WVSCI score and EC values. Sample distribution within the WVDEP data set was skewed toward high WVSCI scores and low EC values.

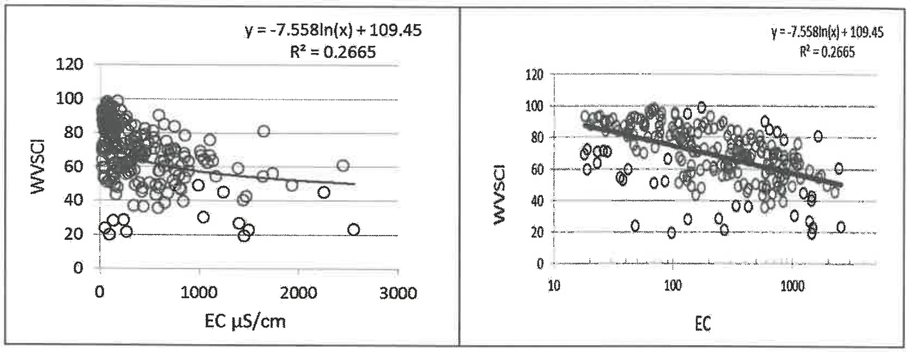

Figure 2. The WVDEP data set was graphed on an non-logarithmic (left) and logarithmic (right) x axis. The black line is the prediction curve given by the model in the upper right corner.

to infer adverse changes in the population due to pollution (Gerritsen et al. 2000); and (2) EC, which is a field based parameter identifying the ability of water to carry an electrical charge. EC increases as ionic concentrations increase. An attempt to identify the relationship between a chemical parameter like EC and WVSCI necessarily infers the behavior of watershed scale processes from limited number of samples. This study attempts to develop a model for predicting a WVSCI outcome based on stream EC. It then estimates the precision with which that prediction can be made, given the inherent variability within the sample and general population.

RESULTS AND DISCUSSION

The observations were not randomly distributed across the range of EC values. Almost two-thirds of samples in the WVDEP data set were in the EC range of 0–399 µS/cm, 21 percent were between 400 and 799 and the remaining 15 percent were distributed between 800 and 2,600 µS/cm. High WVSCI scores and low EC values constituted roughly half of the samples. This results in estimates of higher precision for the low EC ranges and lower precision in the upper EC ranges (Figure 1).

Figure 2 illustrates the relationship between WVSCI and EC in the data set. The two graphs

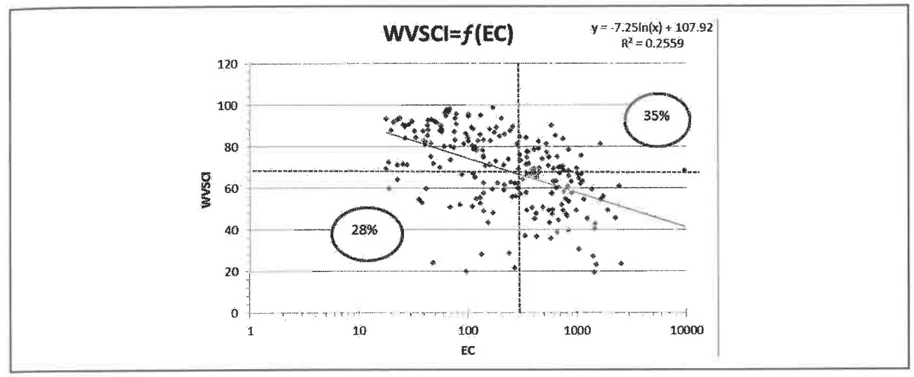

Figure 3. Graph of WVSCI versus EC including mooted thresholds. The vertical red line is an EC value of 300 µS/cm and the horizontal red line is a WVSCI score of 68. The numbers in the red circles represent the proportion of samples that would be misclassified using the thresholds.

in the figure differ only in that the graph on the right utilizes a logarithmic scale for EC. The logarithmic transformation helps to linearize the relationship between WVSCI and EC. The logarithmic transformation also has the advantage of clarifying what is happening at lower EC values. The model yields a weak correlation coefficient of only 0.266.

Model-Free Analysis

Figure 3 overlays the data set with mooted thresholds of 68 for WVSCI and 300 for EC. The upper left quadrant contains samples that would be classified as unimpaired by both (WVSCI and EC) classifications and the lower right quadrant contains samples that would be classified as impaired by both classifications. The upper right and lower left quadrants contain samples where the two measures disagree about impairment classification.

The error rate for a given EC criteria can be calculated by simply plotting the observed values and counting the number of observations which fall outside the predicted responses. For example, if an EC threshold of 300 µS/cm was a perfect predictor of WVSCI impairment, then all WVSCI values would be in either the upper left or lower right quadrants. WVSCI values in the lower left and upper right quadrants represent misclassification. The error rate is a summation of all observations that fall outside the predicted levels at a given EC criterion. Figure 3 indicates that at EC=300 µS/cm (approximately, per Bernhardt et al. 2012) 28 percent of observations with EC<300 also had WVSCI<68 while 35 percent of observations at EC>300 µS/cm had WVSCI scores greater than 68. This constitutes an empirically derived error rate for the WVDEP data set of about 30 percent.

Discretized Analysis

In order to minimize assumptions, we calculated means and confidence intervals for WVSCI scores separately, for separate EC ranges. Table 1 presents the results for this discretized analysis of the relationship between WVSCI and EC. The table presents WVSCI mean values and 95 percent confidence intervals calculated to determine significant differences among means and from WVSCI criteria. Means that are not significantly different from each other are followed by the same letter. Means that are significantly different from the WVSCI criteria, such as 60.6 and 68, are followed by 'yes'; the plus or minus sign indicates whether the WVSCI score is significantly greater or less than the WVSCI criteria. Singular values, with zero degrees of freedom, are indicated by N/A. Only two class mean WVSCI values could be distinguished from the remaining eighteen EC ranges: 0–99 and 1400–1499. One

Table 1. Model-free analysis of WVSCI vs. EC including a test* to discriminate mean values among EC ranges

EC Range	WVSCI LCL	WVSCI Mean	WVSCI UCL	n	Significant Difference	Different from 60.6?	Different from 68?
0–99	76.46	80.44	84.43	64	a	Yes+	Yes+
100–199	68.15	73.25	78.35	41	ab	Yes+	No
200–299	56.66	65.41	74.16	19	b	No	No
300–399	61.47	67.26	73.06	19	b	Yes+	No
400–499	56.62	64.36	72.10	14	b	No	No
500–599	51.10	62.37	73.65	11	b	No	No
600–699	54.32	62.26	70.19	12	b	No	No
700–799	52.83	61.69	70.55	9	b	No	No
800–899	50.80	59.31	67.83	9	b	No	Yes–
900–999	N/A	49.33	N/A	1	N/A	N/A	N/A
1000–1099	41.06	58.81	76.55	5	ab	No	No
1100–1199	52.95	65.17	77.38	4	ab	No	No
1200–1299	N/A	45.22	N/A	1	N/A	N/A	N/A
1300–1399	–101.93	43.27	188.46	2	ab	No	No
1400–1499	15.12	31.46	47.80	4	c	Yes–	Yes–
1500–1699	–52.47	67.92	188.30	2	ab	No	No
1700–1999	22.25	52.77	83.28	2	ab	No	No
2000–2599	4.93	43.30	81.68	3	ab	No	No
Total				222			

* Mean tests were carried out using Student's t-test (Ostle 1963).

EC range (1400–1499) fell below the WVSCI criteria of 60.6 and that two EC ranges, (800–899 and 1400–1499) fell below the 68 WVSCI criteria. The latter EC range was followed by ranges that were not significantly different from either WVSCI criterion.

This analysis confirms the general tendency of higher EC to be associated with lower WVSCI. However, the sample size is not sufficient for this approach to demonstrate that the means differ from each other. Most importantly, the inherent lack of precision within the data does not allow a determination of the critical WVSCI values except at extreme levels of EC.

Regression Model of WVSCI vs. EC

An attempt was made to develop regression models to predict WVSCI from EC. None of the models were particularly successful. Table 2 summarizes the performance of five curve types ranked in order of decreasing correlation coefficient (R^2). The highest R^2 was 0.266, indicating that the model only explains 27 percent of the population variation. The poor R^2 results from

Table 2. WVSCI vs. EC regression models* ranked according to correlation coefficient

n = 222 Curve Type	y = WVSCI, x = EC Model	R^2
Logarithmic	y=-7.558ln(x)+109.45	0.2665
Polynomial(2°)	y=8E-06x2+0.0334x+79.454	0.2622
Linear	y=0.0192x+76.784	0.2372
Exponential	y=75.265e-3E-04x	0.2147
Power	y=125.5x-0.118	0.2110

*All models reported in this study were derived using the Microsoft Excel Regression package.

the large amount of residual variation in WVSCI not explained by EC (Petty et al. 2010) combined with sampling error. We consider this a weak model.

Figure 4 presents the logarithmic regression model between EC and WVSCI. The central solid (red) line is the model. The closest solid lines to the model line (light blue and light green) are the 95 percent confidence interval for the model. The outer solid lines (dark blue and dark green) are the confidence intervals for the population

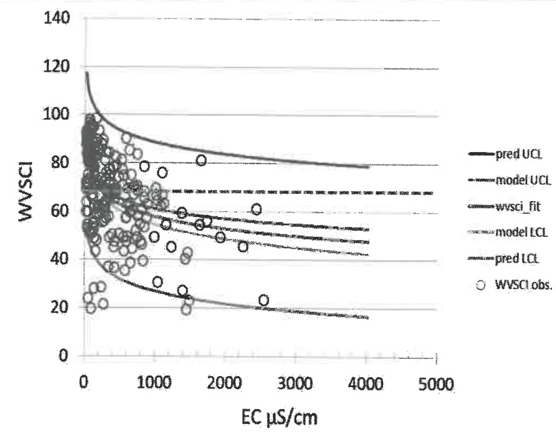

Figure 4. Log-linear model between EC and WVSCI with confidence intervals

(Draper and Smith 1981). The dashed line is the WVSCI criterion of 68.

The model curve and the confidence interval curves for the model in Figure 4 are very similar to the results reported by Bernhardt et al. (2012). Consistent with Bernhardt et al. (2012), a EC value of approximately 300 µS/cm is associated with a model predicted mean WVSCI score of 68. Critically though, Bernhardt et al. (2012) did not report confidence intervals for the sample population. Figure 4 clearly illustrates that observations are readily possible on either side of the WVSC=68 criterion at all levels of EC. This illustrates that no level of EC is suited to predict the divide between WVSCI impairment.

A further distinction is important to make between model and population confidence intervals. A larger sample size will result in smaller confidence intervals around the model, but it will not affect the size of the confidence interval around the population. The latter determines the ability of the model to reliably predict a WVSCI score based on the stream's EC.

Regression Model of WVSCI Impairment vs. EC

WVSCI scores were classified within EC ranges that were arbitrarily set at 100 µS/cm. This allowed a mean and CI to be generated for each EC range. The resulting distribution was determined to be normal and was used in further regression analyses. Using the central points of the EC ranges as the independent variable, regression models were developed to estimate the probability that, at a given EC value, WVSCI scores would be above or below two stream condition criteria: 60.6 (unimpaired gray zone) and 68 (unimpaired, good). The coefficient of correlation (R^2) was used to select the model that best fit the data.

Strong models were developed with the power models providing the highest R^2 values for both WVSCI criteria (Table 3). The results indicate that the probability that a WVSCI score would fall below the 60.6 criteria 50 percent of the time would occur at an EC of 969, while the probability of falling below 68 50 percent of the time would occur at an EC of 389 µS/cm.

Figure 5 illustrates the resulting curves for the two criteria models. The colored horizontal lines represent probabilities of 0.25, 0.50 and 0.75. The lower curve (blue) represents the prediction curve for the 60.6 WVSCI criteria while the upper curve (red) represents the prediction curve for the WVSCI criteria of 68. The polygons are the observed probabilities.

This analysis is again consistent with previous analyses in finding that a CE value of approximately 300 µS/cm is associated with a 50 percent

Table 3. Regression models were developed to estimate the EC at which WVSCI scores would fall below two criteria

n=222	WVSCI Criterion > 60.6		p = WVSCI <		
Curve Type	Model	R^2	0.25	0.50	0.75
Power	$y=0.0255x^{0.4328}$	0.7221	195	969	2472
Polynomial (2°)	$y=0.00000002x^2+0.0003x+0.2$	0.6748	128	907	1620
Linear	$y=0.0003x+0.1991$	0.6735	170	1003	1836
Logarithmic	$y=0.1517\ln(x)-0.5251$	0.5853	166	860	4471
Exponential	$y=0.223e^{0.0007x}$	0.5607	163	1153	1733

n=222	WVSCI Criterion > 68.0		p = WVSCI <		
Curve Type	Model	R^2	0.25	0.50	0.75
Power	$y=0.0321x^{0.4605}$	0.8983	86	389	937
Logarithmic	$y=0.2026\ln(x)-0.6545$	0.8941	87	298	1025
Polynomial (2°)	y=-0.0000002x2+0.0008x+0.2	0.8599	21	368	810
Linear	$y=0.0004x+0.3502$	0.7609	<50	375	1000
Exponential	$y=0.3404e^{0.0007x}$	0.5534	<50	549	1129

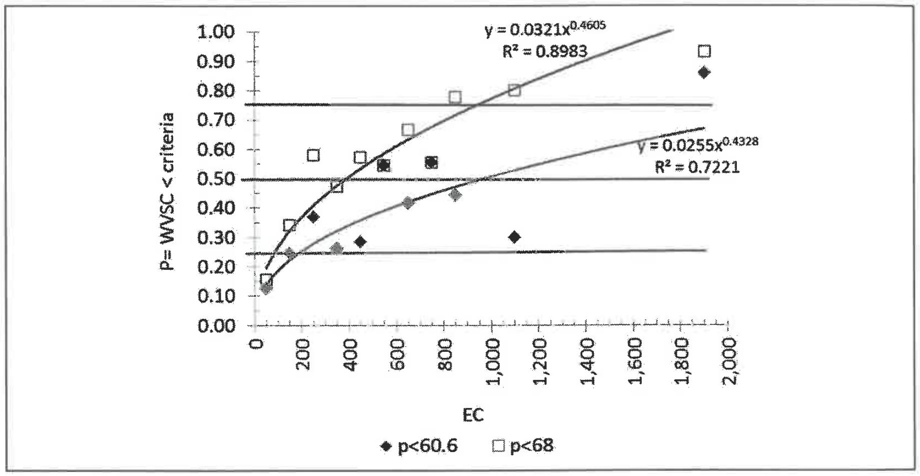

Figure 5. The probability of achieving a WVSCI score below the impairment criteria at a given EC value. The brown, green, and purple horizontal lines represent probabilities of 0.25, 0.50 and 0.75, respectively.

probability of finding a WVSCI score greater than 68. It is also consistent with previous analyses in finding that there is a very high probability of finding WVSCI scores on either side of 68 at a very wide range of EC values. This again illustrates that no level of EC is suited to predict the divide between WVSCI impairment.

Logistic Regression

Logistic regression offers an alternative approach for estimating probability in binomial populations. Figure 6 summarizes the results indicating the probabilities of exceeding a WVSCI score of 68 at different EC values. The vertical line labeled 0.79 (red) represents the average EC value of WVDEP's reference (undisturbed) sites, 68.3. At

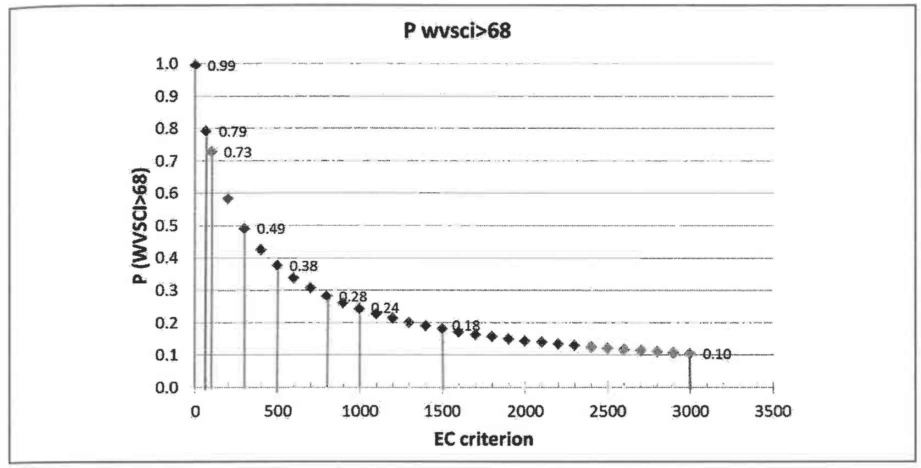

Figure 6. Logistic regression indicating the probability that WVSCI will exceed 68) at a given EC value. The vertical red line is the mean EC value of WVDEP's reference sites.

this EC, about 21 percent of sites would not meet the WVSCI criterion of 68.

This analysis is again consistent with previous analyses in finding that an EC value of approximately 300 µS/cm is associated with a 49 percent probability of finding a WVSCI score greater than 68. It is also consistent with previous analysis in finding that there is a very probability of finding WVSCI scores on either side of 68 at a very wide range of values for EC. This again illustrates that no level of EC is suited to predict the divide between WVSCI impairment.

CONCLUSIONS

The first objective of this study was to identify whether a correlation exists between the WVSCI and EC. We found a statistically significant relationship between EC and WVSCI. The second objective was to identify the best model for predicting WVSCI as a function EC: a log-linear model was the most successful model form. The third objective was to identify the precision with which EC predicts WVSCI. We found that there is very little meaningful predictability for individual observations.

It is important to note that the narrow confidence intervals presented by Bernhardt et al. (2012) only indicate the variability to be expected if the model was refit to a similar but independent data set, and not the ability of the model to predict individual observations. The population confidence intervals illustrate that a WVSCI=68 can be attained at a very wide range of ECs (perhaps encompassing the whole range of the ECs found in the data). This conclusion is supported by either simple inspection of the data or by the prediction confidence bounds for the log-linear model or through the variety of other analysis techniques presented herein.

No evidence was found that an EC value less than 300 µS/cm is a reliable predictor of WVSCI values that meet the West Virginia narrative criteria for stream impairment. This implies that factors other than EC play a significant role in determining WVSCI ranking and that efforts to lower a stream's EC alone will not reliably improve its stream condition.

ACKOWLEDGMENT

This study was sponsored by the Appalachian Research Initiative for Environmental Science (ARIES). The views, opinions and recommendations expressed herein are solely those of the authors and do not imply any endorsement

by ARIES employees, other ARIES-affiliated researchers or industrial members. Information about ARIES can be found at http://www.energy.vt.edu/ARIES.

REFERENCES

Bernhardt, E.S., Lutz, B.D., King, R.S., Fay, J.P., Carter, C.E., Helton, A.M., Campagna, D., and Amos, J. 2012. How many mountains can we mine? Assessing the regional degradation of Central Appalachian rivers by surface coal mining. *Environ. Sci. Technol.* 46: 8115–8122.

Draper, N.R. and Smith, H. 1981. *Applied Regression Analysis,* 2nd ed. New York: Wiley.

Elphick, J.R., Davies, M., Gilron, G., Canaria, E.C., Lo, B., and Bailey, H.C. 2011. An aquatic toxicological evaluation of sulfate: The case for considering hardness as a modifying factor in setting water quality guidelines. *Environmental Toxicology and Chemistry.* 30(1):247–253.

Gerritsen, J., Burton, J., and Barbour, M.T. 2000. *A Stream Condition Index for West Virginia Wadeable Streams.* Prepared for U.S. EPA Region 3 Environmental Services Division and U.S. EPA Office of Science and Technology Office of Water. Available online at http://www.dep.wv.gov/wwe/watershed/bio_fish/documents/wvsci.pdf.

Petty, J.T., Fulton, J.B., Strager, M.P., Merovich, G.T., Stiles, J.M. and Ziemkiewicz, P.F. 2010. Landscape indicators and thresholds of stream ecological impairment in an intensively mined Appalachian watershed. *J. N. Am. Benthol. Soc.* 29(4):1292–1309.

Ostle, B. 1963. *Statistics in Research,* 2nd ed. Ames: Iowa State University Press.

U.S. EPA (Environmental Protection Agency). 2010. A Field-based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams. Office of Research and Development, National Center for Environmental Assessment, Washington, DC. EPA/600/R-10/023A.

West Virginia Department of Environmental Protection (WVDEP). 2012. *Benthic Macroinvertebrate and Fish Information.* http://www.dep.wv.gov/WWE/watershed/bio_fish/Pages/Bio_Fish.aspx.