JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

Vol. 47, No. 1 AMERICAN WATER RESOURCES ASSOCIATION February 2011

INCORPORATING SOCIAL CONTEXT VARIABLES INTO PAIRED WATERSHED DESIGNS TO TEST NONPOINT SOURCE PROGRAM EFFECTIVENESS[1]

Linda Stalker Prokopy, Z. Asligül Göçmen, Jing Gao, Shorna Broussard Allred, Joseph E. Bonnell, Kenneth Genskow, Alicia Molloy, and Rebecca Power[2]

ABSTRACT: In a traditional paired watershed study, watersheds are selected to be as similar as possible so that conclusions may be drawn about the performance of Best Management Practices. We have extended the paired watershed concept to examine the effectiveness of watershed management programs by adding comparative criteria for social characteristics. For four different 8 or 11/12 digit hydrologic unit code (HUC) watersheds in the Midwest, we have piloted a systematic method for selecting paired subwatersheds. First, we developed a list of 11 key variables. Next, a factor analysis was conducted to determine the underlying structure of the 11 input variables. Finally, in each of the four watersheds, potential paired subwatersheds (all 14 digit HUCs) were selected using the factors in a cluster analysis. Informal interviews were then held with key informants in each watershed to provide qualitative assessments of criteria that could impact the comparability of the subwatersheds. This method for selecting paired watersheds should be helpful for other researchers to test the effectiveness of watershed management programs focused on behavior change.

(KEY TERMS: social indicators; watershed planning; outreach and education; nonpoint source pollution.)

Prokopy, Linda Stalker, Z. Asligül Göçmen, Jing Gao, Shorna Broussard Allred, Joseph E. Bonnell, Kenneth Genskow, Alicia Molloy, and Rebecca Power, 2011. Incorporating Social Context Variables Into Paired Watershed Designs to Test Nonpoint Source Program Effectiveness. *Journal of the American Water Resources Association* (JAWRA) 47(1):196-202. DOI: 10.1111/j.1752-1688.2010.00508.x

INTRODUCTION

Since Clausen and Spooner (1993) wrote about the use of paired watersheds to test the effectiveness of Best Management Practices (BMPs) at reducing nonpoint source (NPS) pollution, over one hundred published and countless unpublished studies have used this method. This approach consists of having at least two different watersheds: one serving as a control watershed, and one serving as a treatment watershed where some sort of intervention is applied. Data are collected in both watersheds at a minimum of two points in time: a calibration period when everything

[1]Paper No. JAWRA-10-0026-N of the *Journal of the American Water Resources Association* (JAWRA). Received March 1, 2010; accepted October 13, 2010. **Discussions are open until six months from print publication**.

[2]Respectively, Associate Professor, Department of Forestry and Natural Resources, Purdue University, 195 Marsteller Street, West Lafayette, Indiana 47907; Assistant Professor, Department of Urban and Regional Planning, University of Wisconsin-Madison, Madison, Wisconsin; Graduate Student, University of Wisconsin-Madison, Madison, Wisconsin; Associate Professor, Department of Natural Resources, Cornell University, Ithaca, New York; Program Director for Watershed Management, Ohio State University Extension, Columbus, Ohio; Assistant Professor, Department of Urban and Regional Planning, University of Wisconsin-Madison, Madison, Wisconsin; Watershed Planning Specialist, Department of Forestry and Natural Resources, Purdue University, West Lafayette, Indiana; Co-Director, Great Lakes Regional Water Program, Environmental Resources Center, University of Wisconsin-Extension, Madison, Wisconsin (E-Mail/Prokopy: lprokopy@purdue.edu).

is constant and a treatment period when an intervention is put in place in one of the watersheds. This basic approach is intended to measure the effectiveness of watershed management programs, including education and outreach, technical assistance efforts, and financial assistance. While some of these studies have been inconclusive, others have demonstrated impact or the potential for change to be shown using this approach (see, e.g., Fiener and Auerswald, 2003; Dietz *et al.*, 2004; King *et al.*, 2008). This paper presents a method that we developed in our work in three Midwestern states to extend earlier paired watershed work to include measures of social context. We conclude by discussing the limitations and applications of such an approach.

Clausen and Spooner (1993) give four basic criteria for choosing which watersheds to study in a paired watershed approach. They state that watersheds should: (1) be similar in size, slope, location, soils, and land cover, (2) be small enough to obtain uniform treatment over the entire watershed, (3) have a stable channel and cross section for discharge monitoring, and should not leak at the outlet, and (4) be in the same land cover for a number of years prior to the study so that they are at a steady state. Many of the studies since Clausen and Spooner have looked at aspects of the four basic criteria such as *size* (Burton, 1997; Andreassian, 2004; Dietz *et al.*, 2004), *slope* (Dietz *et al.*, 2004; Mol and Ouboter, 2004), *location* (Church, 1999; Sovell *et al.*, 2000; Andreassian, 2004; Dietz *et al.*, 2004), *soil types* (Schilling, 2002), *land cover* (Meals and Hopkins, 2002), and *stableness of the channel* (Andreassian, 2004) to determine comparability and similarities among watersheds. Additional criteria considered have included *land use* (Sovell *et al.*, 2000; Andreassian, 2004), *precipitation* (Spooner *et al.*, 1995, Schilling, 2002), *runoff patterns* (Spooner *et al.*, 1995; Burton, 1997), and *similar geomorphic units* (Sovell *et al.*, 2000; Andreassian, 2004).

Some studies have also included social aspects of the watershed in addition to the biophysical influences to determine comparability including volunteer activity (Dietz *et al.*, 2004), potential willingness of landowners to enroll in conservation design (King *et al.*, 2008), and fertilizer usage behavior (Schilling, 2002). However, the extension of the paired watersheds methodology to study the efficacy of educational interventions in watershed management is limited. Dietz *et al.*'s (2004) study of the effectiveness of an educational intervention in a residential neighborhood in Connecticut is the only study we are familiar with that used a paired watershed approach to assess changes in human behavior due to an intervention. In Dietz *et al.*, both the treatment and control watershed were in communities with active volunteer groups; active volunteer groups was the only social factor used to pair the watersheds. In the treatment watershed, homeowners were educated about NPS pollution through seminars. Based upon baseline and follow-up survey data, they found no significant changes in behaviors in the treatment watershed. More studies like Dietz *et al.*'s are needed to understand what interventions motivate the adoption of BMPs. It is our hope that an improved paired watershed design as presented in this paper will help inform this work and help understand how local context influences adoption decisions.

This work began as part of an effort to measure short- and medium-term impacts of state and federal programs to address NPS water pollution. One way that accountability can be demonstrated is through the use of social indicators. Social indicators in this context are measures of awareness, attitudes, constraints, and behaviors of target audiences whose behavior can increase or reduce NPS pollution. If watershed management efforts can show that they are changing social indicators then there is evidence that these changes will lead to increased adoption of BMPs over time (Prokopy *et al.*, 2008) and some evidence that adoption of BMPs will lead to improvements in environmental conditions (see e.g., Fiener and Auerswald, 2003). The Social Indicators Planning and Evaluation System, currently undergoing pilot tests in the Great Lakes Region, consists of collecting baseline data about a target audience and using that data to inform social outcomes and educational plans (Prokopy *et al.*, 2009; Genskow and Prokopy, 2010). It further consists of conducting surveys following implementation to see if measurable changes in indicator values have occurred. We developed the paired watershed selection process to determine if using baseline social indicator data would enable conservation promoters to develop education and outreach programs that work better than either: (1) no education and outreach or (2) "business as usual" education and outreach which is not informed by detailed knowledge about the target audience.

Our work takes place in four watersheds: the La Moine watershed, an 8-digit HUC in western Illinois; the Clifty Creek watershed, an 11-digit HUC in south-central Indiana; the Sandusky River-Tiffin watershed, a 12-digit HUC in northwestern Ohio; and the Upper Scioto River watershed, a 12-digit HUC in central Ohio. Approximately, 60% of the 2,140 square miles in the La Moine River watershed, 92% of the land in the 205-square-mile Clifty Creek watershed, 84% of the 117-square-mile Sandusky River-Tiffin watershed, and 72% of the 718 square miles in the Upper Scioto River watershed is in agricultural production. The challenge was to select comparable watersheds to measure the effect of the targeted educational activities in the study watersheds.

METHODS

We developed a list of 11 criteria that we could use to compare the social and environmental fabric of the watersheds in question (see Table 1). All of these data are readily available and the variables used were generated from spatial data for smaller units of analysis (e.g., census block groups, a grid size of 30 m by 30 m) and aggregated at the subwatershed level using GIS. In the instances where the subwatershed boundary did not follow the boundary of the variable of interest such as the census block group geography, we conducted an area-weighted analysis to divide the census block group. This division allowed a reasonable estimation of, for example, total population, the number of individuals with certain socio-demographic characteristics, and the number of housing units recently constructed within each subwatershed.

It quickly became apparent that the 11 measurements were correlated to each other – sometimes highly correlated (e.g., population density and land cover development). Therefore, to get independent parameters for the cluster analysis, we needed to execute a factor analysis first. A factor analysis uncovers the underlying, abstract factors present in a set of measures (Kim and Mueller, 1978; Dunteman, 1989). We then used Principal Component Analysis with Varimax rotation to extract factors that were sufficiently independent from one another (Netemeyer *et al.*, 2003). To create one unified factor structure, all 111 subwatersheds from the 4 watersheds in the three states were used. Nine factors which explained 96.5% of the variance were extracted; these nine factors had Eigenvalues greater than or near 1.00, in other words suggesting that each factor extracts at least as much as or close to the equivalent of the original variables.

We then calculated the scores on the nine factors, using their readings in the 11 measurements and Anderson-Rubin method, for each of the 111 subwatersheds. These factor scores were used in conducting a cluster analysis for each of the four study watersheds. In doing so, we were able to see how a watershed's subwatersheds statistically clustered together. We used Ward's method to combine clusters and the squared Euclidean distance to measure the intervals between clusters (Aldenderfer and Blashfield, 1984). This method minimizes the variance within each cluster and maximizes the variance between clusters. This process generated a dendrogram for each watershed; the dendrogram from the Clifty Creek watershed is shown in Figure 1.

In the Clifty Creek watershed, there are 16 14-digit HUCs. On the left-hand side of the dendrogram, the case number axis is numbered from 1 to 16 and represents the 14-digit subwatersheds. On the top of the dendrogram, the distance cluster scale ranges from 0 to 25. Moving from left to right, the closer to the origin (0) on the scale that the subwatersheds meet, the greater the similarity between the subwatersheds. For example, 7, 8, 10, 11, and 12 meet at approximately 4 on the scale, showing that they are quite similar. On the other hand, they do not meet with any other subwatersheds until about 12, illustrating that beyond those five subwatersheds, they are not too comparable to the other subwatersheds within the Clifty Creek watershed. All 16 subwatersheds meet at 25. We selected groupings of subwatersheds that had similar subwatersheds – indicated by clustering together as early on the scale as possible for further comparability analysis. In other words, none of the subwatersheds in a group was very similar to the others in that group, but the groups all together displayed similar physical and social characteristics.

Following the quantitative selection of subwatersheds, we collected additional contextual data from key informants (such as watershed coordinators, Soil and Water Conservation District personnel and

TABLE 1. Initial Variables Used to Select Subwatersheds.

Variable	Source
Percent population graduated from high school	2000 Census (Block group geography)
Average household income	2000 Census (Block group geography)
Percent new construction post-1990	2000 Census (Block group geography)
Population change 1990-2000 (%)	2000 Census (Block group geography)
Population density in 2000 (persons/sq km)	2000 Census (Block group geography)
Land cover: percent developed	National Land Cover Dataset 2001
Land cover: percent agriculture	National Land Cover Dataset 2001
Percentage of area with slope ≤2%	USGS National Elevation Dataset, Digital Elevation Model (DEM) publication date 1999 (metadata)
Natural stream density [ft/acre (.122 meters/hectare)]	National Hydrography Dataset (NHD) from USGS publication date 1999 (metadata)
Areas of prime farmland (%)	SSURGO 2.2 data (Soil Survey Geographic Database) 2006
Areas of not prime farmland (%)	SSURGO 2.2 data (Soil Survey Geographic Database) 2006

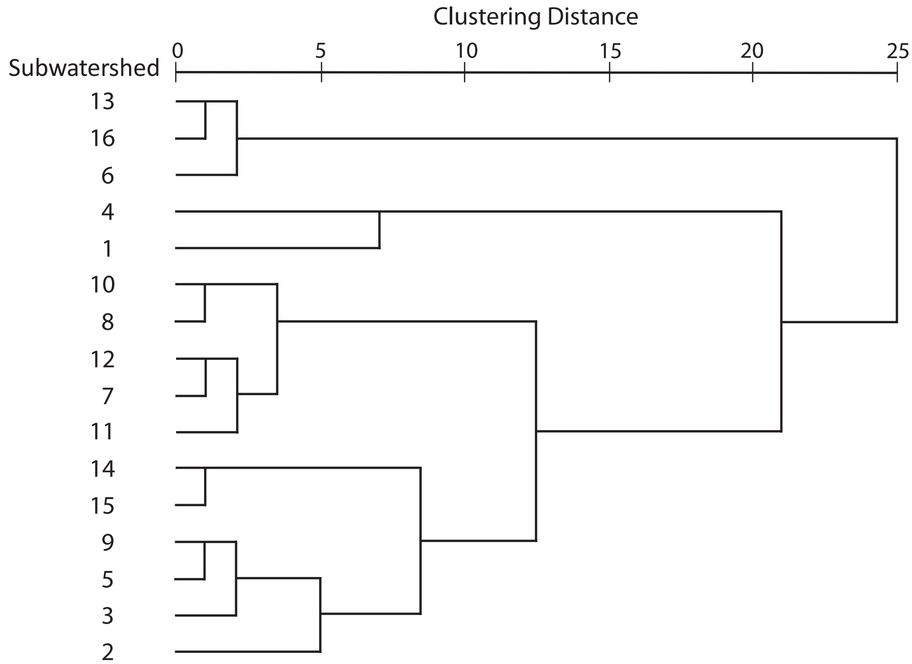

FIGURE 1. Dendrogram Using Ward Method for Clifty Creek Watershed, Indiana.

Natural Resource Conservation District personnel) within the watersheds to further allow comparability of watersheds. In the watersheds we were working in, none of this contextual information was available in quantitative form by subwatershed. Instead of trying to construct new data layers, we felt it was sufficient to speak to local experts. The information we collected included:

- Which subwatersheds were a priority for the watershed group.
- What was the current level of funding for work in the subwatersheds.
- Which water quality impairments have been identified.
- What conservation efforts/BMPs were currently being used on the ground in the watersheds.
- What conservation/BMPs were the groups actively promoting.
- What kinds of educational and outreach programs had the projects tried in the past.
- What barriers to adoption of BMPs were present in the watershed.
- The amount of tile drainage (qualitative assessment).
- The average cost per acre of agricultural land.
- What kind of unique social features were part of this watershed (e.g., presence of Amish communities, prevalence of equestrian operations).
- The approximate number of farmers that live in each watershed.

In each watershed, we selected groups of subwatersheds that: (1) were similar to each other on important social and environmental statistics, (2) were compatible with local watershed plans and capacity to conduct outreach, and (3) had an adequate number of agricultural producers living within each watershed to aid in having enough social data to actually be able to measure change after implementing the education and outreach program. From these groupings, we assigned one subwatershed to each of the three different experimental types (treatment, business as usual, control). We tried to assign the treatment subwatersheds based on geographic continuity wherever possible to allow ease of reaching people through the outreach activities. Figure 2 shows the three groupings that were ultimately chosen for the study in the Clifty Creek watershed based on the quantitative and qualitative data gathered.

DISCUSSION

The method outlined above can be modified as necessary to enable unique research needs to be met. In our case we wanted three different types of subwatersheds (treatment, control, business as usual) in each of our watershed areas and we wanted a sufficient number of producers in each subwatershed to have a sufficient audience size for program delivery and statistical

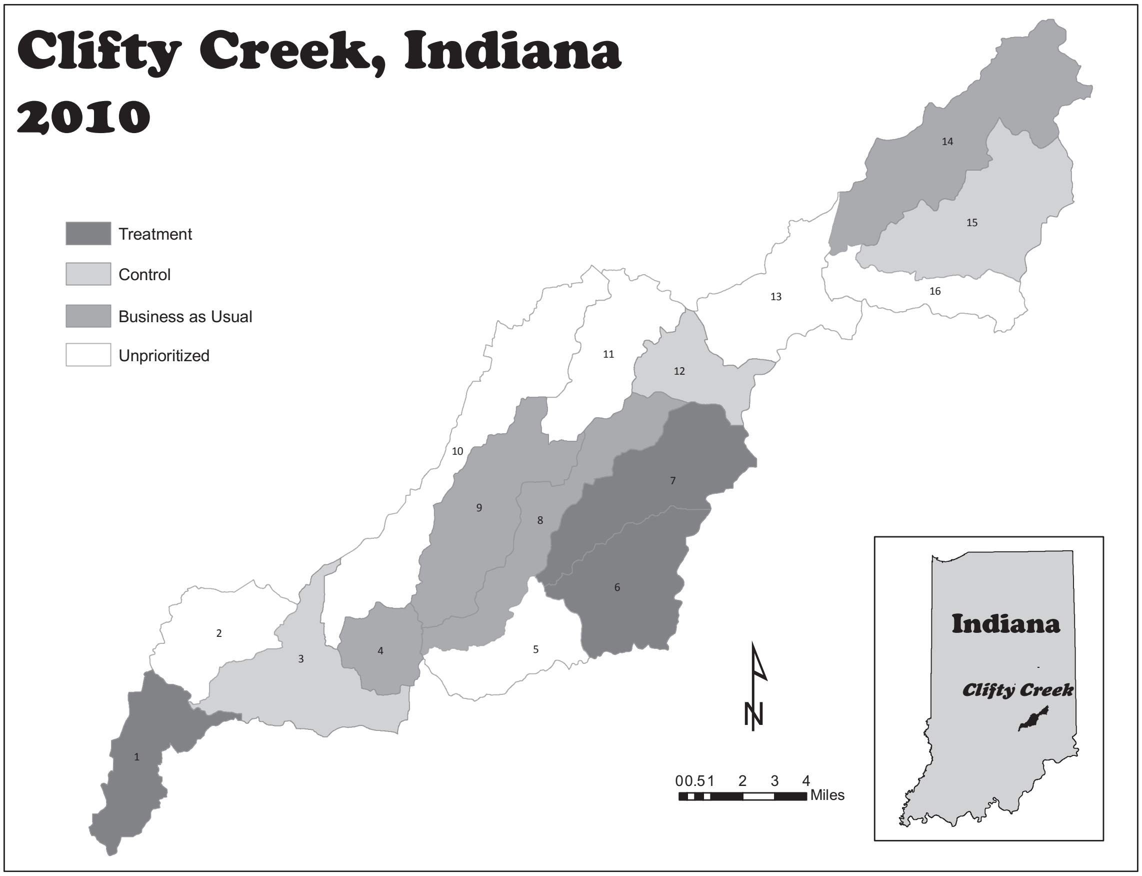

FIGURE 2. Final Groupings of Clifty Creek Watershed, Based on Quantitative and Qualitative Analysis.

significance. Different research studies will likely have different needs. It is essential, however, for any study on the impacts of watershed management programs to consider having treatment and control audiences. In watershed work, it makes sense to compare these audiences along watershed boundaries using a paired watershed methodology. This allows both social and environmental conditions to be measured over time.

The advantages of this approach are similar to those outlined by Clausen and Spooner in 1993 for biophysical studies and include: (1) external influences such as broad educational campaigns or fish kills are statistically controlled, (2) awareness, attitudinal and behavioral changes can be attributed to a treatment, (3) having a control watershed eliminates the need to measure all components causing change, (4) watersheds need not be identical, (5) study can be completed in shorter time frame than trend studies, and (6) cause-effect relationships can be indicated. Table 2 shows the comparison between the advantages of using paired watersheds for a social study compared to a biophysical study.

TABLE 2. Advantages of Paired Watershed Studies.

Biophysical Study (from Clausen and Spooner, 1993)	**Social Study**
1. *Climate and hydrological differences over years* are statistically controlled	1. *External influences such as broad educational campaigns or fish kills* are statistically controlled
2. Can attribute *water quality* changes to a treatment	2. Can attribute *awareness*, *attitudinal and behavioral* changes to a treatment
3. Control watershed eliminates need to measure all components causing change	3. Control watershed eliminates need to measure all components causing change
4. Watersheds need not be identical	4. Watersheds need not be identical
5. Study can be completed in shorter time frame than trend studies	5. Study can be completed in shorter time frame than trend studies
6. Cause-effect relationships can be indicated	6. Cause-effect relationships can be indicated

There are also several disadvantages or limitations that are unique to using paired watersheds for social study. First and foremost, it is not possible to ensure that control subwatersheds do not receive any education and outreach or general offers of technical and program support. In a best-case scenario, all members of a particular target audience in all subwatersheds (treatment and control) will receive the same basic levels of support which would then be supplemented in a treatment subwatershed by the intervention to be tested. Second, unlike implementing a BMP which one can guarantee stays in the watershed it's placed in, it is not possible to ensure that broader outreach interventions stay in their intended watershed. For example, an especially well-designed brochure or similar educational material may be passed from farmer to farmer and cross watershed lines. Similarly, even if the intended audience for a field day is only farmers in the treatment subwatershed, it is not possible to prohibit others from attending. Third, much social data can be difficult to acquire at a watershed scale (e.g., census data is available only in blocks which do not correspond to watershed boundaries). This requires extensive data manipulation to determine comparability of watersheds.

There are many additional ways that our method could be enhanced. For example, in our case we did not match subwatersheds based on baseline survey data although we did check the survey data qualitatively to determine there were no large differences such as awareness levels or use of BMPs between respondents in the different subwatersheds. Matching subwatersheds based upon baseline levels of awareness, attitudes, and behaviors would be an excellent way to demonstrate measurable change.

Finally, the method proposed in this paper has some methodological contributions that are beneficial for both social and biophysical studies. The use of factor and cluster analysis based upon readily available GIS data enabled an objective selection of paired watersheds. This method works for large watersheds which are not typically considered in paired watershed studies; the purely qualitative approach often used to pair watersheds may not work for such large watersheds.

CONCLUSION

To improve the quality of our nation's waterways, efforts at changing behaviors that affect watersheds need to be evaluated more thoroughly. Education and outreach methods for encouraging adoption of water quality BMPs have been criticized, in part because of a lack of scientifically rigorous research and evaluation to demonstrate impact. Many evaluation efforts rely on post-intervention, self-reports of knowledge, awareness, and behavior change. Even the exceptional evaluation that utilizes pre- and post-intervention measurements cannot account for external influences on measured changes. The method for paired watershed selection outlined in this article can be used to support scientific assessment of educational and other program management interventions while allowing researchers or evaluators to draw conclusions about causal relationships between interventions and observed changes. To strengthen the validity of paired watershed studies, it is imperative that researchers demonstrate comparability of the paired watersheds. Four primary characteristics have been suggested for evaluating comparability of paired watersheds for biophysical studies. We have recommended additional criteria for evaluating comparability of paired watersheds for social science research. Studies that use this approach could test the benefits of any number of different interventions, such as local BMP auctions, market-based trading, peer learning, demonstrations, etc., on different types of audiences. Finally, the use of this approach will enable interdisciplinary research that evaluates both the effectiveness of education on BMP adoption and BMP adoption on improvements in water quality.

ACKNOWLEDGMENTS

Funding for this study was provided by a USDA Cooperative State Research, Education, and Extension Service Section 406 Integrated Water Quality grant (agreement number 2006-51130-03701). Any opinions, findings, conclusion, or recommendations expressed in this article are solely those of the authors and do not necessarily reflect the views of the USDA. We want to thank Joe Campbell and Aaron Thompson for their input and assistance in developing this methodology. We also want to thank our watershed partners for their help in selecting subwatersheds – Heather Siesel and Christine Goldstein at the Clifty Creek Watershed Project, Jeff Boeckler at the Illinois Department of Natural Resources, Ed Miller at the Delaware Soil and Water Conservation District, and Cynthia Brookes of the Sandusky River Watershed Coalition.

LITERATURE CITED

Aldenderfer, M.S. and R.K. Blashfield, 1984. Cluster Analysis. Sage Publications, Thousand Oaks, California, Quantitative Applications in the Social Sciences Series No. 44.

Andreassian, V., 2004. Waters and Forests: From Historical Controversy to Scientific Debate. Journal of Hydrology (Amsterdam) 291:1-27.

Burton, T.A., 1997. Effects of Basin-Scale Timber Harvest on Water Yield and Peak Streamflow. Journal of the American Water Resources Association 33:1187-1196.

Church, M.R., 1999. The Bear Brook Watershed Manipulation Project: Watershed Science in a Policy Perspective. Environmental Monitoring and Assessment 55:1-5.

Clausen, J.C. and J. Spooner, 1993. Paired Watershed Study Design. Office of Water, U.S. Environmental Protection Agency, Washington, D.C., EPA 841-F-93-009, 8 pp.

Dietz, M.E., J.C. Clausen, and K.K. Filchak, 2004. Education and Changes in Residential Nonpoint Source Pollution. Environmental Management 34:684-690.

Dunteman, G.H., 1989. Principal Components Analysis. Series in Quantitative Applications for the Social Sciences, Sage University Paper #69, Newbury Park, California, 96 pp.

Fiener, P. and K. Auerswald, 2003. Concept and Effects of a Multi-Purpose Grassed Waterway. Soil Use and Management 19: 65-72.

Genskow, K. and L. Prokopy, 2010. Lessons Learned in Developing Social Indicators for Regional Water Quality Management. Society and Natural Resources 23(1):83-91.

Kim, J. and C.W. Mueller, 1978. Factor Analysis: Statistical Methods and Practical Issues. Series in Quantitative Applications for the Social Sciences, Sage University Paper #14, Newbury Park, California, 88 pp.

King, K.W., P.C. Smiley, Jr., B.J. Baker, and N.R. Fausey, 2008. Validation of Paired Watersheds for Assessing Conservation Practices in the Upper Big Walnut Creek Watershed, Ohio. Journal of Soil and Water Conservation 63(6):380-395.

Meals, D.W. and R.B. Hopkins, 2002. Phosphorus Reductions Following Riparian Restoration in Two Agricultural Watersheds in Vermont, USA. Water Science and Technology 45:51-60.

Mol, J.H. and P.E. Ouboter, 2004. Downstream Effects of Erosion From Small-Scale Gold Mining on the Instream Habitat and Fish Community of a Small Neotropical Rainforest Stream. Conservation Biology 18:201-214.

Netemeyer, R.G., W.O. Bearden, and S. Sharma, 2003. Scaling Procedures: Issues and Applications. Sage Publications, Thousand Oaks, California, 206 pp.

Prokopy, L., K. Floress, D. Klotthor-Weinkauf, and A. Baumgart-Getz, 2008. Determinants of Agricultural BMP Adoption: Evidence From the Literature. Journal of Soil and Water Conservation 63(5):300-311.

Prokopy, L., K. Genskow, J. Asher, A. Baumgart-Getz, J.E. Bonnell, S. Broussard, C. Curtis, K. Floress, K. McDermaid, R. Power, and D. Woods, 2009. Designing a Regional System of Social Indicators to Evaluate Nonpoint Source Water Projects. Journal of Extension 42(2). http://www.joe.org/joe/2009april/a1.php.

Schilling, K.E., 2002. Chemical Transport From Paired Agricultural and Restored Prairie Watersheds. Journal of Environmental Quality 31:1184-1193.

Sovell, L.A., B. Vondracek, J.A. Frost, and K.G. Mumford, 2000. Impacts of Rotational Grazing and Riparian Buffers on Physiochemical and Biological Characteristics of Southeastern Minnesota, USA, Streams. Environmental Management 26:629-641.

Spooner, J., D.E. Line, S.W. Coffey, D.L. Osmond, and J.A. Gale, 1995. Linking Water Quality Trends With Land Treatment Trends: The Rural Clean Water Program Experience. http://www.water.ncsu.edu/watershedss/info/brochures/ten.html, *accessed* November 2010.