On the influence of landscape-scale factors on stream ecosystems and macroinvertebrate assemblages in the southeastern USA: an examination of alternative statistical methods and case studies.

by

Brad Patrick Schneid

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 1, 2015

Approved by

John W. Feminella, Chair, Professor and Chair, Department of Biological Sciences
Christopher Anderson, Co-chair, Associate Professor, School of Forestry and Wildlife Sciences
Ash Abebe, Associate Professor, Department of Mathematics and Statistics
Brian Helms, Assistant Research Professor, Department of Biological Sciences

Abstract

The influence of land cover on stream ecosystems has increasingly been the focus of research in a variety of fields, including ecology, hydrology, and engineering. Human-altered land cover can increase overland storm runnoff and be a source of nutrients, chemicals and sediment to streams, which can negatively affect biota. Urbanization has been recognized as a particularly influential form of land cover, and low levels of urban development (e.g., suburban land) has become the focus of an increasing number of studies, as it is predicted to be a more prevalent form of land-cover change over the next century. The coastal plains of the southeastern US are a relatively understudied region; however, some research has indicated that these lowland streams may be influenced by land cover to a lesser degree than those from more frequently studied and higher gradient landscapes. Current trends and predictions in land-cover change and human population growth suggest that streams and rivers will become increasingly influenced by human activity. Thus, research is warranted to further investigate the influence of low levels of urban development on streams and how this influence may vary regionally.

Studies on the influence of land cover on stream ecosystems are complicated by multiple issues, including the following: 1) experimental manipulation of whole watersheds is generally impractical, therefore most studies have been observational, 2) land cover is generally expressed as proportions; thus, land-cover classes are likely to be correlated, 3) whole-watershed replication is frequently low relative to the number of variables considered, thus constraining statistical analyses, 4) land cover is typically not the direct cause of biological degradation, and 5) real-world data are not ideal and anomalies (outliers) are generally present and not always

realized by the investigator. Some of the above issues are problematic when using traditional statistical methods, including ordinary least-squares (OLS) regression. Small sample sizes, correlated predictor variables, and outliers separately and in combination contribute to unreliable estimation and prediction by OLS regression. Whereas no true solution to these problems exists, most ecologists often do not consider alternatives that may, to some degree, address or alleviate these analytical challenges.

Overall, my dissertation provides important information regarding appropriate statistical choices for analyzing real-world data which usually should not be assumed to conform to assumptional requirements of traditional methods. Nowhere are these issues more evident than in land-cover or ecological studies in general where small sample sizes are commonplace, predictor variables are frequently highly correlated, and outliers are likely present. In addition, my dissertation contributes to research on the influence of land cover on stream ecosystems in the understudied southeastern coastal plains, suggesting that impervious surface cover $\leq 11\%$ likely influences hydrology and physicochemistry of streams. This information is particularly important because low-levels of urban development are predicted to be the most prevalent of land-cover change along the Gulf Coast in the foreseeable future. Last, my dissertation importantly highlights specific differences in macroinvertebrate taxonomic and trait composition that exist in lowland coastal plains versus those in highland regions. Macroinvertebrates are the most frequently used taxonomic group in stream biomonitoring, and my research reinforces the body of literature suggesting that regional bioassessment metrics are needed to accurately identify impairment specific to each region.

Acknowledgements

I must thank my parents from the bottom of my heart. Joseph and Sandra Schneid provided a loving, caring environment to grow up in and instilled me with an important moral base that has proven invaluable. My parents encouraged and supported my inquisitiveness and ever-changing interests as I attempted to find myself. I also thank my entire family, but especially my older brother Matt for helping me and attempting on multiple occasions to steer me towards success and my younger brother Drew for being one of my best friends and helping me in more ways than he will ever know.

I thank Drs. Jack Feminella, Christopher Anderson, Ash Abebe and Brian Helms for their guidance, support, and understanding as committee members. Specifically, I would like to thank my co-advisors, Jack and Chris, for inviting me to join their labs and providing many years of sound advice and generous support. I thank Ash for also being an an excellent M.P.S. advisor and for truly fostering my interest in statistics (and for bringing me to Zimbabwe!). I also must thank Brian for his mentorship, endless help in the field and lab, and for being a good friend.

I thank Dr. Latif Kalin for serving as my outside reader and whose Advanced Forest Hydrology course made me 1st realize that I actually liked mathematics. I also must thank Dr. Nedret Billor for being an excellent teacher; it is your fault that I became obsessed with statistics. I must sincerely thank Stephen Sefick for joining the Feminella lab and being there to laugh with and answer my seemingly endless questions about computers, statistics and ecology. I would also like to thank Molli Newman, Emily Hartfield-Kirk, Brian Lowe, Susan Riethel Colvin,

Flynt Barksdale, Diane Alix and many others for their friendship and help in the field and lab. Last, but not least, I would also like to thank Dr. Alexis M. Janosik for her friendship, loving support, and understanding; meeting you made graduate school worth the trouble.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction to the influence of land cover on streams and associated analytical issues.

1.1 Introduction

Biodiversity loss is of great global concern, with freshwater ecosystems experiencing rates of loss close to that of tropical forests (Ricciardi and Rasmussen 1999; Dudgeon et al. 2006). Despite occupying <1 % of the Earth's surface, ~10% of all formally described species inhabit freshwaters, and 90% of those (~10,000 species) are invertebrates (Strayer and Dudgeon 2010). The southeastern United States is disproportionately rich in native freshwater invertebrate species. The state of Alabama alone contains 40, 43, and 60% of native U.S. aquatic insect, gill-breathing snail, and mussel species, respectively (Lydeard and Mayden 1995; Meyer et al. 2007). Information regarding the status of most freshwater species is generally inadequate; however, losses for freshwater species are estimated to be as high as 4% per decade (Dudgeon et al. 2006).

Exploitation, pollution, flow modification, habitat degradation, and invasive species have been identified as the major threats to freshwater biodiversity, each of which is linked to anthropogenic activity at the landscape (or finer) scale (Dudgeon et al. 2006). Human population expansion created a need to alter landscapes to provide additional resources; currently at 7.25 billion, the global human population is ~35-55 times greater than pre-agriculture populations (Cincotta 2011). Landscape conversion for agricultural and urban use composes roughly one-third of the Earth's surface (Cincotta and Gorenflo 2011). The size of the human population is expected to increase by 33% in the next 30 y (Alig et al. 2004; Cincotta and Gorenflo 2011)

thereby likely causing additional significant landscape alteration (Cincotta 2011). Urbanization (including urban sprawl) is perhaps the most pervasive form of land-use/land-cover (LULC) change in the southeastern US (O'Driscoll et al. 2010), with this region predicted to have the greatest regional growth in population and increase in land development in coming decades (Xian et al. 2012).

Urbanization is a particularly influential form of LULC characterized by high levels of impervious surface cover (ISC), which directly alters catchment hydrology and increases runoff velocity and volume (Brabec 2002; Brown et al. 2009). The "urban stream syndrome" has been coined in reference to a suite of consistently observed alterations to stream ecosystems in urbanized watersheds (Walsh et al. 2005). Findings from recent studies suggest that the relative magnitude to which urban LULC influences hydrology, sediment export, physical habitat and aquatic organisms decreases along a gradient of decreasing topographical relief (Appalachians > Piedmont > coastal plains) (Utz et al. 2009, 2011; Utz and Hilderbrand 2011). Additional research is warranted, as few studies have taken place in the coastal plains relative to other regions in the Southeast (O'Driscoll et al. 2010; Nagy et al. 2011).

Declines in diversity and richness and shifts to more tolerant benthic taxa are generally associated with LULC change, but exactly why this shift occurs is poorly understood. Utz et al. (2009) and Utz and Hilderbrand (2011) reported inter-regional differences in macroinvertebrate sensitivity to urban stressors and inter-regional variation in recolonization to disturbance. In their study, more rapid recolonization occurred in the coastal plains relative to the Piedmont, with the authors hypothesizing that coastal plains macroinvertebrate assemblages consist of more resistant and/or resilient taxa better adapted to such disturbance (Utz and Hilderbrand 2011). Ecological information regarding biological traits (e.g., functional, morphological) purportedly offers

benefits over purely taxonomic-based analyses (Culp et al. 2010), but it is not clear how macroinvertebrate assemblages in coastal plains streams differ from those highland (upland) streams in the Southeast in terms of traits that may offer resistance or resilience to LULC-related stressors.

To date, most urban studies have focused on high levels of urban development (Cunningham et al. 2009; Chadwick et al. 2011), although more recent attention has been given to understand abiotic and biotic responses to lower level, suburbanization/exurban development (Hansen et al. 2005; Burcher and Benfield 2006). Watershed impervious cover of $\leq$10% has been shown to alter stream hydrologic and physicochemical conditions and lead to declines in species richness (Brabec 2002; Morse et al. 2003; Nagy et al. 2011; Nagy et al. 2012), although some studies have shown responses of assemblages to much lower urbanization levels (4.4% ISC; Wenger et al. 2009). Research on the potential impacts of low-level urban development is generally limited, so additional research is necessary to determine generalities in macroinvertebrate response.

Coastal areas worldwide are under increasing pressures from to human population growth and associated land development (Nagy et al. 2011). In the past 2 decades, roughly half of the urban LULC change along the US Gulf of Mexico has occurred within 50 km of the coast; with the dominant LULC change from the Florida panhandle to Louisiana has been low-intensity development (e.g., suburban, urban-sprawl; Xian et al. 2012). The SE region has led, and is predicted to lead US regions in developed LULC and population growth into 2030, and low-density development is predicted to continue expanding along the coast of the Gulf of Mexico (Wear and Greis 2002; Alig et al. 2004; White et al. 2008; Xian et al. 2012); thus, there is an

urgent need to predict and understand impacts of low-density development on stream ecosystems in such rapidly growing coastal watersheds.

Identification of mechanisms by which LULC affects stream ecosystems has become a priority, and is of obvious importance for prescriptive management (Allan 2004). Ordinary least-squares (OLS) regression has traditionally been used to develop predictive or descriptive models elucidating potential functional relationships between response and predictor variables (Mac Nally 2000, 2002). In studies on the influence of LULC on stream ecosystems, land-cover data is generally summarized as proportions, and is therefore inherently collinear, leading to inflated estimation variance and complicating interpretation of statistical analyses (Allan 2004; King et al. 2005). Practitioners have been taught to address collinearity by manually removing one or more variables (Whittingham et al. 2006; Smith et al. 2009) or use automated selection techniques (e.g., stepwise regression), which have been criticized and may produce spurious results (Murtaugh 2009). Small sample sizes (n) are also common to many land-cover studies and ecological research in general (Bissonette 1999; Van Sickle 2003). Estimated coefficient variance can be inflated by collinearity, but it also increases with decreasing n (Speed 1994; O'Brien 2007); thus, recommendations have been made on the number of observations relative to predictors (e.g., ratio of 10:1; Speed 1994; Maxwell 2000). Several alternative estimation methods (to standard OLS) offer a trade-off such that they might be biased, but with smaller estimation variance and mean squared error; however, most are generally not considered by ecologists (Dahlgren 2010).

Outliers are an unfortunate reality for those dealing with real-world data; outliers often are notoriously difficult to detect, especially when dimensions of the data become large (Rocke and Woodruff 1996). Alternative statistical methods, including partial least-squares (PLS), have

been suggested for consideration in ecological research, but these methods (like OLS) are sensitive to outliers (Carrascal et al. 2009; Dahlgren 2010). Currently implemented robust PLS methods assume outliers occur across entire rows (Møller et al. 2005); thus, these methods work to identify and downweigh influences of whole observations. In ecological data, outliers may not be confined to existing as entire observations and can occur randomly within a dataset (Møller et al. 2005; Rousseeuw et al. 2006); thus, PLS methods that are robust to randomly positioned data outliers are needed.

The broad goals of my dissertation are to examine alternative statistical methods (to OLS regression) for use in LULC studies, and to focus on pertinent issues regarding LULC change and stream ecosystems in the SE, with an emphasis on the coastal plains. The primary objectives of my dissertation are: 1) to assess the utility of alternative regression methods for the analysis of highly collinear data (e.g., LULC classes) in small sample size situations; 2) to assess the performance of a simple modification to an alternative regression method for robust estimation when random outliers are present; 3) to determine if empirical evidence suggests that low-levels of urban development in coastal watersheds influences stream hydrology, physicochemistry and/or benthic macroinvertebrate assemblages; and 4) to determine if empirical evidence suggests that coastal plains benthic assemblages are "better" adapted to urban related stressors than assemblages in higher-gradient regions

The 1st chapter of my dissertation provides a brief review of the issues addressed in subsequent chapters. In the 2nd chapter, Monte-Carlo simulations are used to compare OLS regression and related model selection procedures (e.g., stepwise regression) with alternative regression methods that included shrinkage methods (Lasso) and latent variable methods (partial least-squares (PLS)). My results suggest that PLS may offer more reliable coefficient estimates

and more accurately identify important predictors than OLS when sample sizes are extremely small (< 15) and predictor variables are highly correlated.

In my 3rd chapter, I examined whether empirical evidence suggested that low-levels of urban development influences stream hydrology, physicochemistry and/or benthic macroinvertebrate assemblages of small coastal streams in southern Alabama, USA. My results suggested that such development may lead to increased stormwater flashiness and spate frequency, as well as higher median water temperatures, concentrations of total suspended solids, and total nitrogen concentrations. Suprsingly, development did not appear to influence macroinvertebrate richness or sensitivity metrics, which may be more influenced by natural gradients in organic matter and hydrologic permanence.

In my 4th chapter, I examined the efficacy of a simple modification to the PLS algorithm using rank-based cross-products for regression estimation and prediction when outliers exist randomly in a dataset. My results indicated that rank-based PLS outperforms standard PLS when outliers are present and is also highly efficient compared to standard PLS when outliers are absent. In addition, rank-based PLS also outperformed existing "robust" PLS algorithms when outliers were placed randomly throughout datasets, as the existing robust algorithms assume outliers exist across individual observations (i.e. within rows of data).

In my final and 5th chapter, I examined the idea that macroinvertebrate assemblages in the southeastern US coastal plains are more resistant and/or resilient to stressors associated with watershed urbanization than assemblages in adjacent highland regions. This idea stemmed from empirical evidence suggesting that the magnitude of response by macroinvertebrates in coastal plains streams was lower than those in highland regions. My results suggested that coastal plains

assemblages may be more resistant to sedimentation and more tolerant of low dissolved oxygen conditions and organic pollution than highland assemblagess.

1.2 References

Alig, R. J., J. D. Kline, and M. Lichtenstein. 2004. Urbanization on the US landscape: looking ahead in the 21st century. Landscape and Urban Planning 69:219-234.

Allan, J. D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annual Review of Ecology, Evolution, and Systematics 35:257-284.

Bissonette, J. A. 1999. Small sample size problems in wildlife ecology: a contingent analytical approach. Wildlife Biology 5:65-71.

Brabec, E. 2002. Impervious surfaces and water quality: a review of current literature and its implications for watershed planning. Journal of Planning Literature 16:499-514.

Brown, L., T. Cuffney, J. Coles, F. Fitzpatrick, G. McMahon, J. Steuer, A. Bell, and J. May. 2009. Urban streams across the USA: lessons learned from studies in 9 metropolitan areas. Journal of the North American Benthological Society 28:1051-1069.

Burcher, C. L. and E. Benfield. 2006. Physical and biological responses of streams to suburbanization of historically agricultural watersheds. Journal of the North American Benthological Society 25:356-369.

Carrascal, L. M., I. Galván, and O. Gordo. 2009. Partial least squares regression as an alternative to current regression methods used in ecology. Oikos 118:681-690.

Chadwick, M. A., J. E. Thiele, A. D. Huryn, A. C. Benke, and D. R. Dobberfuhl. 2011. Effects of urbanization on macroinvertebrates in tributaries of the St. Johns River, Florida, USA. Urban Ecosystems:1-19.

Cincotta, R. P. 2011. The Biological Diversity that Is Humanly Possible: Three Models Relevant to Human Population's Relationship with Native Species. Pages 61-72 *in* R. P. Cincotta

and L. Gorenflo, editors. Human Population: Its influences on biodiversity. Springer, Berlin.

Cincotta, R. P. and L. Gorenflo. 2011. Introduction: Influences of Human Population on Biological Diversity. Pages 1-9 *in* R. P. Cincotta and L. Gorenflo, editors. Human Population: Its influences on biodiversity. Springer, Berlin.

Culp, J. M., D. G. Armanini, M. J. Dunbar, J. M. Orlofske, N. L. R. Poff, A. I. Pollard, A. G. Yates, and G. C. Hose. 2010. Incorporating traits in aquatic biomonitoring to enhance causal diagnosis and prediction. Integrated environmental assessment and management 7:187-197.

Cunningham, M. A., C. M. O'Reilly, K. M. Menking, D. P. Gillikin, K. C. Smith, C. M. Foley, S. L. Belli, A. M. Pregnall, M. A. Schlessman, and P. Batur. 2009. The suburban stream syndrome: evaluating land use and stream impairments in the suburbs. Physical Geography 30:269-284.

Dahlgren, J. P. 2010. Alternative regression methods are not considered in Murtaugh (2009) or by ecologists in general. Ecology Letters 13:E7-E9.

Dudgeon, D., A. H. Arthington, M. O. Gessner, Z. I. Kawabata, D. J. Knowler, C. Lévêque, R. J. Naiman, A. H. Prieur-Richard, D. Soto, and M. L. J. Stiassny. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews of the Cambridge Philosophical Society 81:163-182.

Hansen, A. J., R. L. Knight, J. M. Marzluff, S. Powell, K. Brown, P. H. Gude, and K. Jones. 2005. Effects of exurban development on biodiversity: patterns, mechanisms, and research needs. Ecological Applications 15:1893-1905.

King, R. S., M. E. Baker, D. F. Whigham, D. E. Weller, T. E. Jordan, P. F. Kazyak, and M. K. Hurd. 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15:137-153.

Lydeard, C. and R. L. Mayden. 1995. A diverse and endangered aquatic ecosystem of the Southeast United States. Conservation Biology 9:800-805.

Mac Nally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between - and reconciliation of - 'predictive' and 'explanatory' models. Biodiversity and Conservation 9:655-671.

Mac Nally, R. 2002. Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables. Biodiversity and Conservation 11:1397-1401.

Maxwell, S. E. 2000. Sample size and multiple regression analysis. Psychological Methods 5:434-458.

Meyer, J. L., D. L. Strayer, J. B. Wallace, S. L. Eggert, G. S. Helfman, and N. E. Leonard. 2007. The Contribution of Headwater Streams to Biodiversity in River Networks. Journal of the American Water Resources Association 43:86-103.

Møller, S. F., J. von Frese, and R. Bro. 2005. Robust methods for multivariate data analysis. Journal of Chemometrics 19:549-563.

Morse, C. C., A. D. Huryn, and C. Cronan. 2003. Impervious surface area as a predictor of the effects of urbanization on stream insect communities in Maine, USA. Environmental Monitoring and Assessment 89:95-127.

Murtaugh, P. A. 2009. Performance of several variable selection methods applied to real ecological data. Ecology Letters 12:1061-1068.

Nagy, R. C., B. G. Lockaby, B. Helms, L. Kalin, and D. Stoeckel. 2011. Water Resources and Land Use and Cover in a Humid Region: The Southeastern United States. J. Environ. Qual 40:867-878.

Nagy, R. C., B. G. Lockaby, L. Kalin, and C. Anderson. 2012. Effects of urbanization on stream hydrology and water quality: the Florida Gulf Coast. Hydrological Processes 26:2019–2030.

O'Brien, R. M. 2007. A caution regarding rules of thumb for variance inflation factors. Quality & Quantity 41:673-690.

O'Driscoll, M., S. Clinton, A. Jefferson, A. Manda, and S. McMillan. 2010. Urbanization Effects on Watershed Hydrology and In-Stream Processes in the Southern United States. Water 2:605-648.

Ricciardi, A. and J. B. Rasmussen. 1999. Extinction rates of North American freshwater fauna. Conservation Biology 13:1220-1222.

Rocke, D.M. and D.L. Woodruff. 1996. Identification of outliers in multivariate data. Journal of the American Statistical Association 91:1047-1061.

Rousseeuw, P. J., M. Debruyne, S. Engelen, and M. Hubert. 2006. Robustness and outlier detection in chemometrics. Critical reviews in analytical chemistry 36:221-242.

Smith, A. C., N. Koper, C. M. Francis, and L. Fahrig. 2009. Confronting collinearity: comparing methods for disentangling the effects of habitat loss and fragmentation. Landscape Ecology 24:1271-1285.

Speed, R. 1994. Regression type techniques and small samples: A guide to good practice. Journal of Marketing Management 10:89-104.

Strayer, D. L. and D. Dudgeon. 2010. Freshwater biodiversity conservation: recent progress and future challenges. Journal of the North American Benthological Society 29:344-358.

Utz, R., K. Eshleman, and R. Hilderbrand. 2011. Variation in physicochemical responses to urbanization in streams between two Mid-Atlantic physiographic regions. Ecological Applications 21:402-415.

Utz, R. M. and R. H. Hilderbrand. 2011. Interregional variation in urbanization-induced geomorphic change and macroinvertebrate habitat colonization in headwater streams. Journal of the North American Benthological Society 30:25-37.

Utz, R. M., R. H. Hilderbrand, and D. M. Boward. 2009. Identifying regional differences in threshold responses of aquatic invertebrates to land cover gradients. ecological indicators 9:556-567.

Van Sickle, J. 2003. Analyzing correlations between stream and watershed attributes. Journal of the American Water Resources Association 39:717-726.

Walsh, C., A. Roy, J. Feminella, P. Cottingham, P. Groffman, and R. Morgan. 2005. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24:706-723.

Wear, D. N. and J. G. Greis. 2002. Southern forest resource assessment: summary of findings. Journal of Forestry 100:6-14.

Wenger, S. J., A. H. Roy, C. R. Jackson, E. S. Bernhardt, T. L. Carter, S. Filoso, C. A. Gibson, W. C. Hession, S. S. Kaushal, and E. MartÃ-. 2009. Twenty-six key research questions in urban stream ecology: an assessment of the state of the science. Journal of the North American Benthological Society 28:1080-1098.

White, E. M., A. T. Morzillo, and R. J. Alig. 2008. Past and projected rural land conversion in the US at state, regional, and national levels. Landscape and Urban Planning 89:37-48.

Whittingham, M. J., P. A. Stephens, R. B. Bradbury, and R. P. Freckleton. 2006. Why do we still use stepwise modelling in ecology and behaviour? Journal of Animal Ecology 75:1182-1189.

Xian, G., C. Homer, B. Bunde, P. Danielson, J. Dewitz, J. Fry, and R. Pu. 2012. Quantifying urban land cover change between 2001 and 2006 in the Gulf of Mexico region. Geocarto International 1:1-19.

Chapter 2. Small sample sizes, collinear predictors and linear modeling: a simulation study comparing alternative methods and a case study on landscape-stream ecosystem research.

2.1 Introduction

Deforestation and the general alteration of watersheds can greatly influence stream ecosystems (Allan 2004; Walsh et al. 2005). Forested lands, specifically those directly adjacent to stream channels (riparian forests), often act as material sinks and natural physical filters, mechanistically linked to in-stream sediment conditions and nutrient concentrations (Naiman & Décamps 1997). Conversely, human-altered landscapes (e.g., forest conversion to agriculture or urban uses) act as sources of various types of pollutants (e.g., nutrients, solids) and can impede water infiltration, influencing stream hydrology (e.g., stormflow magnitude; Allan 2004; Walsh et al. 2005). Anthropogenic land-cover also is thought to indirectly affect stream biota through altered hydrology, physicochemical conditions and habitat (Burcher, Valett & Benfield 2007). In studies on the influence of land-cover on stream ecosystems, land-cover data is generally summarized as proportions of entire watersheds; as a result, several correlated land-cover categories can be associated with a response variable, complicating interpretation of statistical analyses (Allan 2004; King et al. 2005).

Cost-effective and accurate monitoring tools for predicting stream water-quality from remotely sensed (geographic information systems) land-cover data should have great appeal for land management. The identification of mechanisms by which land-cover likely affects stream ecosystems has become a priority, and is of obvious importance for prescriptive management

(Allan 2004). Historically, ordinary least-squares (OLS) regression has been used to develop predictive or descriptive models elucidating potential functional relationships between response (y) and predictor (x) variables (Mac Nally 2000, 2002; Morrice et al. 2008). Typically, true causal mechanisms are never known; however, it is important that statistical modeling be based on theoretical/proposed causal relationships to provide optimally informative models and better predictive performance given new data (Mac Nally, 2000; Dormann et al. 2013).

Simple bivariate regression/correlations can initially measure the strength of relationships and potential importance of a predictor variables; however, these simple analyses likely produce spurious results when collinearity is present (Van Sickle 2003; King et al. 2005). The goal of model/variable selection and an important part of statistical inference is to identify the best approximating model (Buckland, Burnham & Augustin 1997). All estimation methods produce biased parameter estimates (long-run average not equal to true value) when functionally related predictor variables (hereafter "functional") are omitted; this approach is unfortunate because most practitioners address collinearity by removing one or more variables (Whittingham et al. 2006; Smith et al. 2009; Dormann et al. 2013). In these situations, non-functional collinear predictor variables included in the final model may be allocated explanatory power attributable to omitted functional variables and falsely classified as "important" by the investigator (Mac Nally 2002). Use of partial regression/correlation coefficients can statistically control for the effect of additional predictor variables (Graham 2003; King et al. 2005), and in such cases collinearity is not generally a problem as long as all important functional variables are included (Smith et al. 2009; Dormann et al. 2013). Unfortunately practitioners cannot know if all the key functional variables have been included; therefore, if one or more correlated predictors are

thought to represent different processes (i.e. not mechanistically redundant) they should remain in the analysis (Smith et al. 2009).

Small sample sizes (n) are an unfortunate reality of many land-cover studies and ecological research in general (Bissonette 1999; Van Sickle 2003). Regardless of n and in all but the most extreme cases of collinearity, OLS regression models containing all influential variables produce unbiased, conditional (partial) parameter estimates (Farrar & Glauber 1967; Freckleton 2002; Smith et al. 2009). Coefficient estimation variance can be inflated by collinearity, but it also increases with decreasing n (Speed 1994; O'Brien 2007). Small n should impose limits on the number of predictors used in model building (Dormann et al. 2013) and recommendations of the size of final models ($\geq$ 5:1 or 10:1, observations to predictors; Speed 1994; Maxwell 2000) have left some to rule out use of multiple regression with small n all together (McFarland & Hauck 1999). Interestingly, decreasing n does not increase chances of falsely classifying a predictor as important (false positive rate; set by α) in a multiple regression when all functional variables are included. Rather, decreasing n decreases the power to detect important relationships when they exist; thus, the main concern with small sample sizes is overly conservative testing and not spurious results (Speed 1994).

Automated selection techniques (e.g., stepwise regression) have been criticized because they inherently remove thought from the model building process and may promote data dredging (Murtaugh 2009). An examination of all possible regression model subsets has also been criticized for possibly promoting dredging (Anderson, Burnham & Thompson 2000); however, a well-defined set of candidate models or theoretically linked variables is of extreme importance for the model selection process (Burnham & Anderson 2002). When an *a priori* variable set is defined, analysis of all model subsets ("all subsets") can be a defendable method for examining

the relative effects of a set of predictors, generally forms the basis for model selection or multi-model inference based on information theory (e.g., Akaike's Information Criterion (AIC)), and has been recommended for descriptive regression analyses and to avoid problems associated with collinear datasets (e.g., excluding variables) (Anderson, Burnham & Thompson 2000, Mac Nally 2000; Graham 2003; Whittingham et al. 2006).

Realistically, there is no perfect solution for problems associated with collinear data (Dormann et al. 2013), but there is much interest and research on the use of alternative regression methods that address collinearity (Wold et al. 1984; Tibshirani 1996; Graham 2003; Chong & Jun 2005; Zou & Hastie 2005; Grömping 2007). Methods to avoid the negative effects of collinearity include model averaging with AIC, parameter estimate shrinkage methods (e.g., the least absolute shrinkage and selection operator (Lasso)), and latent variable modeling with partial least squares (PLS; Wold et al. 1984; Tibshirani 1996; Anderson et al. 2000; Dahlgren 2010). These alternative estimation methods (to standard OLS) offer a trade-off such that they might be biased, but with much smaller estimation variance (Fig. 1; Full model vs. PLS for example); however, most are generally not considered by ecologists (Dahlgren 2010).

AIC quantifies the strength of evidence and was inspired by information theory as an estimate of the relative information lost by a model when approximating truth (Anderson, Burnham & Thompson 2000; Burnham, Anderson & Huyvaert 2011). AIC allows for the comparative ranking/weighting of a series of models representing alternative hypotheses related to a single response variable and has become commonly used for model selection in ecological studies (Anderson, Burnham & Thompson 2000; Johnson & Omland 2004). AIC is calculated as the sum of model fit (likelihood deviance) plus a model size (complexity) penalty:

$$\mathrm{AIC} = -2\log_e(L(\hat{\theta}|\mathrm{data})) + 2\mathrm{K}, \quad [1]$$

where $\log_e(L(\hat{\theta}|\text{data}))$ is the maximized log-likelihood value over the unknown parameters (θ) of size K. AIC can also be easily calculated for OLS, as maximized log-likelihood is proportional to the residual sum of squares (Anderson, Burnham & Thompson 2000). AIC has been modified for finite/small sample size (n) situations (AICc) and is given by AIC + 2K(K+1)/(n-K-1) (Hurvich & Tsai 1989). AIC allows users to select the single "best" model (model with minimum AIC) or comparatively rank models using differences in AIC values (ΔAIC_i = AIC_i – min(AIC)). The importance of each model can be indicated by AIC model weights, which can be used for averaging across models and are calculated for each of R models as:

$$w_i = \exp(-0.5\Delta AIC_i) \Big/ \sum_1^R \exp(-0.5\Delta AIC_i). \quad [2]$$

Multi-model averaging (MMA) and inference across several models can better incorporate the uncertainty of model selection into the model building process (Buckland et al. 1997). MMA can be accomplished in two ways: 1) the "natural method", where averaging occurs over all variable subsets, but ignores models where a given variable is excluded and 2) the "zero" method, where excluded variables in each model are assigned a zero effect ($\hat{b}$ = 0) (Grueber et al. 2011; Symonds & Moussalli 2011). The zero method tends to shrink coefficient estimates and has been suggested to reduce bias related to model selection uncertainty or collinearity (Burnham et al. 2011; Grueber et al. 2011; Symonds & Moussalli 2011). Variable importance also can be assessed with AIC by summing model weights across all models for each variable separately (Burnham & Anderson, 2002).

"Alternative methods" to OLS (and associated selection criteria) generally are not used in ecological studies (Dahlgren 2010); therefore we provide a brief comparative summary. Ridge regression (although not considered herein) is a method that penalizes the least-squares solution to counteract overestimation bias from collinearity by "shrinking" some coefficients close to (but

not) zero; therefore, it does not produce parsimonious models (Zou & Hastie 2005). Lasso (Tibshirani 1996) and the "elastic net" (EN; Zou & Hastie 2005) are related to ridge regression; both also penalize the least-squares solution but can shrink some coefficients to exactly zero, thus performing simultaneous variable selection. For comparison, OLS, Ridge, Lasso and EN aim to find a set of slope values (β; an "argument", hence argmin below) that minimize functions of the model residuals ($y - \mathbf{X}\beta$):

$$\hat{\beta}_{OLS} = \underset{\beta}{\operatorname{argmin}}(\|y - \mathbf{X}\beta\|_2^2), \quad [3]$$

$$\hat{\beta}_{Ridge} = \underset{\beta}{\operatorname{argmin}}(\|y - \mathbf{X}\beta\|_2^2 + \lambda\|\beta\|_2^2), \quad [4]$$

$$\hat{\beta}_{Lasso} = \underset{\beta}{\operatorname{argmin}}(\|y - \mathbf{X}\beta\|_2^2 + \lambda\|\beta\|_1), \text{ and} \quad [5]$$

$$\hat{\beta}_{EN} = \underset{\beta}{\operatorname{argmin}}(\|y - \mathbf{X}\beta\|_2^2 + \lambda_2\|\beta\|_2^2 + \lambda_1\|\beta\|_1), \quad [6]$$

where $y_{(n, 1)}$ is a centered response vector, $\mathbf{X}_{(n, p)}$ is a scaled/centered matrix of predictor variables (scaling/centering not necessary for OLS), $\|v\|_1 = \sum_i |v_i|$ is the L_1 (Taxicab) norm, $\|v\|_2 = (\sum_i |v_i^2|)^{1/2}$ is the L_2 (Euclidean) norm, and λ, λ_1, and λ_2 are tuning parameters. While ridge regression does not discard any variables, Lasso can indiscriminately discard all but one variable if it belongs to a group of highly correlated variables; EN incorporates both penalty types and its solution lies intermediate to Ridge and Lasso, thereby discarding variables but capturing 'all the big fish' (Zou & Hastie 2005; Grömping 2009). Lastly, shrinkage methods have been shown to outperform subset methods in small sample size situations (Dahlgren 2010).

Partial least-squares regression (PLS; Wold et al. 1984) is a multivariate method, related to principal components analysis (PCA), which has gained popularity in biological research because of its ability to handle large collinear datasets with n << p (Boulesteix & Strimmer 2007). The basic assumption of PLS is that the system under investigation can be described by a

few underlying and unmeasured latent variables (Rosipal & Krämer 2006); PLS directly addresses collinear structure in a dataset through linear transformation of **X** to latent variables (projections) maximally related to variation in y (Wold, Sjöström & Eriksson 2001). The similarities between PCA and PLS can be seen in the construction of their latent variable components (PCs and Ts respectively); the kth components of PCA and PLS are obtained by finding their respective weight vectors (w; "loadings"):

$$PCA: w_k = \underset{\|w\|_2=1}{\operatorname{argmax}}(var(\boldsymbol{X}w)), \text{ and} \quad [7]$$

$$PLS: w_k = \underset{\|w\|_2=1}{\operatorname{argmax}}(cov(\boldsymbol{X}w, y)), \quad [8]$$

which are constrained to have length equal = 1 (unit vectors). Components (scores) of PCA and PLS are required to be orthogonal and are calculated as $\mathbf{X}w_k$ (Boulesteix and Strimmer 2007). PCs sequentially decrease in the total variation of **X** explained, whereas PLS's T components sequentially decrease with regards to their covariation with y; thus PLS seeks latent variables maximally related to y. Regression coefficients corresponding to the original **X** variables can be calculated as:

$$\hat{\beta}_{PLS} = \boldsymbol{W}Q', \quad [9]$$

where $\boldsymbol{W}$ is the matrix of PLS loadings and Q' is a transposed vector of crossproducts between **T** (matrix of PLS scores) and y (Boulesteix and Strimmer 2007). A measure known as "variable importance in projection" (VIP) which can be calculated for each of j variables as:

$$VIP_j = \sqrt{p \sum_{k=1}^{K} \left[SS_k \left(w_{kj} / \|w_k\|_2^2\right)\right] / \sum_{k=1}^{K} (SS_k)} \quad [10]$$

where p = number of columns of **X**, k = kth latent variable out of K latent variables retained, and SS_k is the sums of squares explained by the kth component (Mehmood et al. 2012). Typically $VIP_j > 1$ indicates that the jth variable should be retained in model selection (Mehmood et al. 2012). PLS has recently been suggested for use in ecological (Carrascal, Galván & Gordo 2009) and

land-cover studies (Shi et al. 2013; Zhang et al. 2010) and has been shown to provide reasonable results with small sample sizes and large p (Carrascal et al. 2009). Last, PLS is an ordination method, so graphics can be used to display interrelationships between **X** and *y* (Abdi 2010).

Alternative methods and their potential benefits may not be well known to ecologists and therefore are generally not considered for use. Simulation studies have addressed performance of some methods in prediction (Dormann et al. 2013), coefficient estimation (Smith et al. 2009), and/or the ability to classify predictors as important/unimportant (Carrascal et al. 2009; Chong & Jun 2005) in collinear situations; however, no simulation studies have addressed all three with a focus on small sample situations. We provide brief introduction to some alternatives to OLS regression and compared these methods using the three performance criteria mentioned above in small sample size and collinear settings.

2.2 Methods

2.2.1 Simulation Study

The models/methods compared in this study are summarized in Table 1. Simulation settings may be chosen specifically or arbitrarily, but should not be expected to represent reality; in this simulation study, we attempted to create variables similar to land-cover type data or proportional data in general. We assumed that relationships were linear, or were made linear through transformations (e.g., log-linear), data generally conformed to assumptions of linear modeling (Montgomery et al. 2001), data were examined for common problems (e.g., no influential observations) and a relatively small set of variables were preselected based on theoretical/mechanistic assumptions (i.e. our simulation did not include large p).

We created simulations that spanned a range of sample sizes (n = 10 to 25, by 5) and, for brevity and clarity, we refrained from creating a large number of simulation settings. We created

an initial simulation setting of p = 3 (no. of **X** variables) to demonstrate the concept of omitted variable bias, and used a second simulation setting with p = 7 for the remainder of the study. For both settings, we created p-1variables using a multivariate normal random number generator from the R package "mvtnorm" (Genz et al. 2014), and set inter-predictor correlation levels to be moderately strong (simulation average $|r| \cong 0.70$) (Table 2). From these variables, we adjusted the first two (X_1 and X_2) so that their minimum values were set to 0; these two variables had a maximum ≅ 0.30 and summed to less than one (e.g., similar to land-cover proportions). We then created a new variable (X_3; to mimic a forest-like variable with initial 100% coverage) as $X_3 = 1 - (X_1 + X_2 + U[0, 0.40])$, where U follows a uniform distribution defined by its lower and upper limits. For the p=7 simulation, the remaining variables (dubbed X_4-X_7) were simply considered to be other correlated variables of theoretical interest (e.g., population density) with respect to the response (y).

For simplicity, we created y to be a function of only 2 predictors for both simulation settings (p = 3 & 7) as $y = X_1 b_1 + X_3 b_3 + \epsilon$, where $\epsilon \sim N(0, 0.1)$; we refer to X_1 and X_3 as "functional predictors" herein. We selected population slope values as $b_1 = 1$ and b_3 -0.50 and the remaining b_i (elements in vector β) are set equal 0 and referred to as "non-functional". To assess predictive performance, for every iteration we created a second independent set of "test" data under identical conditions. Simulations settings were iterated 1000 times per setting. Collinearity amongst simulated predictors was assessed with bivariate Pearson's product-moment correlations and variance inflation factors (VIF) calculated for each predictor (Montgomery et al., 2001). We compared linear estimation methods based on several performance criteria:

1. Parameter (slope) estimation
    - $\text{Bias}_{\hat{\beta}} = \text{mean}(\hat{\beta}) - \beta$

- $Variance_{\hat{\beta}} = \frac{1}{m}\sum_{i=1}^{m}\left(\hat{\beta}_i - mean(\hat{\beta})\right)^2$, where m is the number os simulation iterations
- $MSE_{\hat{\beta}} = variance_{\hat{\beta}} + bias_{\hat{\beta}}^2$

2. Classification of predictors (important or not) based on bootstrapped confidence intervals (containing 0 or not) and model specific criteria (Table 1).
    - True positive rate (TPR; power, sensitivity) = % iterations each functional predictor was correctly classified as important
    - False positive rate (FPR; type I error, 1-specificity) = % iterations each non-functional predictor was incorrectly classified as important
    - G = geometric mean of TPR for both functional predictors and 1-FPR for all non-functional predictors each iteration (Chong & Jun 2005)
3. Prediction of y with test data
    - $RMSE_{\hat{y}} = \sqrt{\frac{1}{n.test}\sum_{i=1}^{n.test}(\hat{y}_i - y.test_i)^2}$

We ran the full OLS model contained all predictor variables using the lm function in the "stats" package in R, and AICc was calculated using the function in "MuMIn" (Bartoń 2013). We ran stepwise selection with AICc using a modified version of the "stepAIC" function (both directions) from the "MASS" package in R (Venables & Ripley 2002); this modified algorithm has been used in ecological research (Batáry et al. 2014; Dainese 2011) and the code is available online (http://wwwuser.gwdg.de/~cscherb1/stepwise.txt). We performed multimodel averaging (MMA) using AICc weights to calculate weighted averages for each coefficient over the models the variable is appears in (natural method: MMA.n) or models not containing a variable are given

a zero prior to averaging (zero method: MMA.z) (Burnham & Anderson 2002; Grueber et al. 2011).

We performed regression with Lasso using the "lars" function in the R package with the same name, whereas we conducted EN using the "enet" function in R package "elasticnet" (Zou and Hastie 2012). Final models were chosen for Lasso such that they optimized the Lasso shrinkage parameter (λ) and for EN based on optimization across a grid of both parameters (λ_2, λ_1; Zou & Hastie 2005). Tuning parameters for Lasso and EN were selected using AICc, calculated using model RSS and "effective degrees of freedom" (Zou, Hastie & Tibshirani 2007; Rocha & Yu 2008). Cross-validation generally is used to select tuning parameters (Li, Morris & Martin 2002); however, in preliminary simulation runs, we noted that AICc consistently led to generally more favorable results for EN (and PLS) and was less computationally expensive. PLS was executed using the "pls.regression" function in the R package "plsgenomics" (Boulesteix et al. 2012). Final models for PLS were chosen such that they optimized the number of latent variables retained (k) using AICc, calculated using model RSS and "effective degrees of freedom" (Li, Morris & Martin 2002). PLS models were created using standardized X variables, as PLS is sensitive to the variation among **X**, because latent variables are created that maximize the cov(y, **X**); coefficients were unstandardized for examination of estimation bias and variability.

In practice, inference after model selection usually is based on the selected model, with uncertainty of model selection being rarely incorporated (Buckland, Burnham & Augustin 1997). Using the bootstrap, a proper method for estimating confidence intervals (CIs) would be to apply the model selection method to each resample (Buckland et al. 1997). To directly compare ability of each method to accurately classify predictors, we bootstrapped model coefficients (resampled

observations; 1000 iterations) and calculated CIs. While other CI methods exist, we used the percentile method as it is easy to calculate and more stable than the standard method because it does not rely on an observed statistic as a central point estimate (Efron 2013). With hard threshold selection methods (e.g., shrinkage and stepwise methods), some coefficients are forced to zero; thus there is generally some bootstrap density at zero, even when the majority of the mass lay elsewhere (Tibshirani 2011, as example, see Fig. 6). As a result, we used narrower than traditional bootstrap confidence intervals (75%, 85%), with 85% being suggested and used with other model selection methods (Arnold 2010; Tibshirani & Taylor 2011).

We also classified predictors according to "method-specific importance criteria" (MSIC, Table 1) that would traditionally be used with each method; for example, p-values were used for OLS models with thresholds of ≤ 0.05 and 0.10. For MMA with AICc, thresholds of summed weights (≥ 0.50 & 0.70) were used (Burnham & Anderson 2002), and variable importance in projection (VIP) was used for PLS (≥ 1.0 & 1.05 thresholds; Chong & Jun 2005; Mehmood et al. 2012). For stepwise selection, the inclusion/exclusion of x-variables from the "best model" was used to indicate variable importance, along with non-zero/zero coefficients for Lasso and EN. A secondary threshold for these methods was the magnitude of their respective non-zero coefficients when standardized (≥ 0.05 threshold).

2.2.2 Case Study

We included data from a study on the influence of land-cover on nitrate concentrations in small wadeable streams along the coast of the Gulf of Mexico. In this unpublished study, thirteen non-tidally influenced stream reaches were selected, centered on the town of Foley, Baldwin County, Alabama, USA (30.4056° N, 87.6815° W). ArcGIS and ArcHydro (Environmental Research Systems Institute, Inc., Redland, California) were used to classify land-cover and

delineate sub-catchments upstream of sampling reaches. A few land-cover categories that could be directly linked to in-stream conditions were considered and included % ISC, % agriculture (Ag), and % riparian forest buffer (FB, 100 m width; Allan 2004; Burcher, Valett & Benfield 2007). The selected stream sites varied in terms of other characteristics including watershed area, stream order, and median stream discharge ranged from 112-2481 ha, 1st-2nd order, and 0.04-0.61 m^3s^{-1} at these sites (respectively). However, these watershed size factors were not related to nitrate and not considered further for this simple demonstration. On approximately ten dates over a 1.5 year period, base flow water samples were collected from each site. Concentrations of nitrate ($[N0_3^-]$; mg L^{-1}) were determined by an independent lab using standard procedures (Rice & Association 2012) and we used median $[N0_3^-]$ as our response variable.

A "global validation" of adherence to model assumptions was performed (H_0: assumptions of normality, homoscedasticity, uncorrelatedness and normality of residuals all hold) (Peña & Slate 2006) with the R package gvlma (Peña & Slate 2014), along with standard graphical analyses (e.g., QQ-plots) of residuals (Montgomery et al. 2001). Collinearity between land-cover categories was assessed with simple correlations and variance inflation factors (VIF). To simplify discussion, we refer to slope coefficient estimates as "significant" (different from zero) if their confidence interval (90 %, unless stated otherwise) does not contain zero. All analyses were performed in R-language (R Core Team 2013) and utilized the base packages and packages previously mentioned.

2.3 Results

2.3.1 Omitted-variable bias

Our simulation study with three X variables (n = 15 and 20) showed strong correlation structure between predictor variables ($|r| > 0.67$; nearly identical to those in Table 2) and average

VIF values were 3.59, 3.61, and 5.35, respectively. Models that included both functional variables (1 and 3) were virtually unbiased (simulation average $\hat{b} \cong$ true b) for all slope coefficients, including for the non-functional variable (e.g., $\hat{b}_2 \cong 0$) in the full model. Large differences in average $\hat{b}$s were observed between either OLS models 1 or 3 and any other model, as functional variable(s) were excluded and retained variables are bias (note arrows, Table 3). For example, X_1 estimates (n=15) varied from 0.996 in the full model (1) to 1.507 in the model that excluded X_3 (note: $\hat{b}_3$ = -0.506 for model 1; Table 3). Further, simulation false positive rates were higher than designated by α (0.10 in this example) in cases where functional variables where excluded (model 6), and low true positive rates (statistical power) was observed in the more parameterized true (3) and full (1) models. The correct model (3) had the smallest average AICc value; however, this model was chosen as "best" (ΔAICc = 0) only 17- 42% percent of the time, depending on n (n = 15 or 20; Table 3). Model 3 had a ΔAICc < 2 or < 4 a much greater proportion of the time (44-74% and 94-96%, respectively) (Table 3).

2.3.2 Simulation Study

Strongly correlated data were simulated and average |r| ranged from 0.57 to 0.81 (Table 2). The full OLS model was virtually unbiased (mean value $\cong$ true value) in coefficient estimation, irrespective of sample size (Fig. 1), but showed increasing variability and MSE with decreasing n (Table 4). The high variability associated with the full model estimation of $\hat{b}_1$ and $\hat{b}_2$ corresponded to the high simulation average VIF values for those variables (Table 2).

Stepwise selection was biased for $\hat{b}_1$ (note more extreme median-bias for $\hat{b}_1$ and $\hat{b}_3$) with the smallest sample sizes (n=10 & 15; Fig. 1) and had relatively large variability and total MSE about coefficient estimates (Table 4). Stepwise was generally unbiased and with relatively low variability and MSE for non-functional predictors irrespective of sample size (Fig. 1, Table 4).

The natural model averaging (MMA.n) method with AICc was biased, usually overestimating coefficients b_1 and b_2 (Fig. 1), and had relatively large total MSE (Table 4). Estimation bias as shrinkage (underestimation) was observed for zero method MMA.z as well as for Lasso and EN regarding $\hat{b}_1$ and $\hat{b}_3$ (Fig. 1; note direction of bias for $\hat{b}_3$ mentioned in caption). In addition, MMA.z, Lasso and EN showed relatively low variation and MSE about non-functional coefficient estimates and total MSE (Fig. 1, Table 4). PLS also exhibited shrinkage regarding coefficient estimation for $\hat{b}_1$ and $\hat{b}_3$, and was the only alternative method to consistently (albeit only slightly) overestimate b_2 corresponding to the non-functional "land-cover" proportion variable. PLS generally had low variation about non-functional coefficient estimates (Fig. 1), although each were slightly biased contributing to a large total MSE relative to MMA.z, Lasso and EN; however, this difference decreased with increasing sample size (Table 4).

The full OLS model had low FPR (type I error $\cong$ α) for non-functional predictors (X_2 & X_4 - X_7), but also low TPR (power) for functional predictors (for X_1 & X_3); and thus also poorly classified all predictors at each iteration (G; Fig. 2). TPR for stepwise and AICc (weights) was also relatively low (< 80%) using method specific importance criteria (MSIC), and while AICc attained low FPR, stepwise FPR was roughly 20% for sample sizes > 10 and the median G value for both methods was below 80% (Fig. 2). Lasso, EN and PLS had MSIC TPR $\cong$ 80% or greater for n > 10. Lasso and EN had relatively high (~ 20%) FPR when either $|\hat{b}| > 0$ or standardized $|\hat{b}|$ > 0.05 was considered (Fig. 2). PLS also had relatively high (~ 20%) FPR with VIP > 1; however, rates were much lower when VIP > 1.05 was considered (Fig. 2). Relative to the other methods, PLS, Lasso, and EN showed much greater ability to correctly identify predictors (each iteration) as seen by median values for G with n > 10 (Fig. 2).

Overall, correct classification of predictors was less impressive with the use of bootstrap CIs, which yielded much lower TPR than that from method-specific criteria (with n < 25; Fig. 3). All methods performed generally well at correctly identifying non-functional predictors (low FPR) with the conservative confidence level considered (85%; Fig. 3). Selection using stepwise performed poorly with each sample size, whereas the remaining methods improved markedly with increasing n (Fig. 3).

Predictive accuracy of MMA.z, Lasso, EN and PLS was similar and generally slightly higher than the true value of the error term (SD = 0.10). The full model and natural averaging MMA.n had much larger and variable $RMSE_{\hat{Y}}$ relative to the other methods; however, differences between methods and overall prediction variability were more pronounced with decreasing n (Fig. 4).

2.3.3 Case Study

A reliable assessment of model assumptions is difficult with small sample sizes like that of this study (n = 13; Speed 1994); that said, global validation (*p*-value = 0.94) and graphical analyses of residuals did not indicate large deviations from model assumptions (e.g., normality) or a need for corrective transformations. We expected the least-squares assumption of uncorrelated predictors (Montgomery et al. 2001; Speed 1994) to be violated. Significant negative correlations were observed between % FB and both ISC (r = -0.60, p = 0.032) and Ag (r = -0.53, p = 0.065). A significant correlation was not detected between % Ag and ISC (r = -0.073, p = 0.81) and variance inflation factors for the three land use categories were 3.13 for % FB, 2.03 for Ag and 2.27 for ISC.

All possible subsets of OLS models, including the intercept only model (OLS 8) were calculated along with RSS, R^2 and AICc values, and prediction RSS (PRESS; Table 5). In this

model set (Table 5), nitrate concentration ($[N0_3^-]$) was negatively related with % FB and positively with Ag when each was considered alone (OLS models 5 and 6, respectively); however, partial regression coefficients for Ag were small and insignificant (90% level) in models that included both categories (OLS models 1 and 4). Slope estimates for FB were significant and had similar magnitudes for all OLS models (Table 5, Fig. 5); in addition, FB had a summed AICc weight of 0.98. Model fit statistics indicate that FB alone (OLS model 5) explained a large amount of the total variation in $[N0_3^-]$ ($R^2 = 0.643$); model 5 also had 2nd smallest PRESS and 2nd highest AICc model weight (Table 5). Ag had a low summed AICc weight (0.16) and estimates for Ag were small when % FB was included (OLS 1 and 4, Table 5), but increased roughly proportionally when FB was removed (OLS 2 and 6). The FB and ISC model (OLS 3) had the smallest AICc value, an AICc model weight of 0.457, the highest adjusted R^2 (0.701), smallest RSS (2.003) and PRESS (3.598, Table 5). Percent ISC, on the other hand, had a summed AICc weight of 0.49; however, slope estimates for ISC changed signs between OLS models (Table 5), and the ΔAICc of OLS 8 relative to the FB only model (OLS 5) was very small (0.33).

Each of the alternative models explained large, and roughly equivalent amounts of the variation in $[N0_3^-]$ (minimum $R^2 = 0.648$); of these, PLS had a PRESS statistic comparable to the 2 best OLS models (Table 5). Stepwise selection, model averaging, coefficient shrinkage, and latent variable estimation methods provided coefficient estimates for % FB of same sign and similar magnitude as the OLS models (Table 5, bottom section). In addition, PLS VIP was 1.31 and AICc summed weights were 0.98 for % FB. These models differed mainly with regards to coefficient estimation for % Ag and ISC. Lasso, EN, PLS and MMA.n provided similar estimates for % Ag, which corresponded closely to that of OLS model 4 that included only FB

and Ag. Conversely, MMA.z and stepwise selection estimated the coefficient for Ag to be very small (0.003) or exactly zero (Table 5). CIs for Ag estimates all contained 0 (Fig. 5) and Ag had a very small summed AICc weight (0.16); however, Ag had high PLS VIP values (1.11) and non-zero estimates from Lasso/EN. As mentioned, ISC had a summed AICc weight of 0.49; however, ISC had a PLS VIP value of only 0.23 and coefficients were shrunk to exactly zero by Lasso and EN.

2.4 Discussion

2.4.1 Simulation Study

Few have acknowledged the limitations created by collinear land-cover percentages on estimation and inference (King et al. 2005). The primary goals of this study were to 1) demonstrate problems associated with small n and collinear predictors and 2) to examine the relative performance of OLS alternatives for linear modeling. Although applicable to observational studies in all subfields of ecology, our results are perhaps most relevant to landscape-level studies where collinearity is inevitable and adequate replication is impractical, if not impossible.

Our results (p =3) provided a simple visual demonstration of how omitting variables can lead to estimation bias in retained variables, as well as those omitted (which are estimated as 0). In collinear situations, omitting variables can also lead to spurious results as false positive rates are inflated. The selection of a single best model (with AIC) is considered bad practice (Anderson & Burnham 2002), and our simulation supported this, as AICc infrequently choose the correct model.

As expected, our simulation results (p = 7) indicate high estimation variability and extremely low power for OLS when n is small and parameter number is relatively large, which

suggests the need for model/variable selection in these situations. As alternatives to more traditional OLS selection methods, shrinkage and latent variable methods appear to offer considerable benefits over traditional OLS, or OLS guided by stepwise selection or MMA with AICc. These alternative methods, while biased in terms of coefficient estimation, yielded a relatively lower variance about estimates, predict relatively well given new data, and better identified functional and non-functional predictors (using method-specific criteria) at an apparently much more acceptable rate when n is very small.

Multimodel averaging (MMA.z), shrinkage methods and PLS each provided lower levels of variation and MSE about coefficient estimates, as well as smaller and less variable RMSE of prediction relative to the other methods examined (Figs. 2 & 4, Table 4). Of these, Lasso and EN had smallest MSE for coefficient estimation separately for each coefficient, and in total (Table 4; but see MSE for *b*3 with PLS & n > 15). Shrinkage and latent variable methods also had much higher method specific TPR than the remaining methods. PLS consistently allocated some small effect (overestimation) to X_2 (the non-functional proportional variable); however, PLS was the only method to have both high TPR (> 80%), low FPR (type I error << 20%) and therefore a high proportion of correctly identified all predictors each iteration (G; Fig 2). The superior ability of PLS in identifying functional variables relative to OLS with small sample sizes has been noted by others (Carrascal et al. 2009). Use of PLS with a cutoff slightly higher than the VIP > 1 rule has been suggested, while others have suggested lower (0.70) values (Eriksson, 1999, Chong & Jun 2005; Zhang et al. 2010). Our simulation suggests using VIP > 1 (e.g., 1.05) may be more appropriate, especially with extremely small sample sizes (e.g., n=10) to better balance true and false positive rates; lower VIP thresholds may lead to unacceptably high FPR ($\geq$ 20%).

Bootstrapped CIs incorporate the uncertainty associated with model selection (Buckland, Burnham & Augustin 1997; Efron 2013), and as such, offered a much more pessimistic view of rates of correctly classifying predictor variables in this simulation. Because at least some of the bootstrapped density for any coefficient might occur at exactly 0, simple examination the CI for the presence of 0 may be misleading as a CI endpoint may be exactly 0 (Fig 6). At present, there is no general consensus on how to best calculate CIs for model coefficients in model selection scenarios, or how and what to resample (residuals, observations, etc., see Hall, Lee & Park 2009 and Efron 2014).

2.4.2 Case Study

The case study provided an example analysis of a small and collinear dataset where results varied according to the model/method chosen. Many of the methods/models fit the data similarly and relatively well (most $R^2 > 0.65$); cross-validation (PRESS) suggested that OLS models 3 and 5 and the PLS model may predict well given new data. Each method/model estimated a negative slope of similar magnitude (range: -0.028, -0.048) between the $[NO_3^-]$ and the percentage of riparian zone (100 m width) as forest buffer (% FB). Lasso and EN slope estimates were generally comparable and each allotted exactly zero coefficient estimate for % ISC, while stepwise estimated Ag to be zero (by dropping variable). Inconsistencies with the sign and magnitude of estimated ISC coefficients were observed across OLS models and also between the observed and bootstrap median value for PLS (Table 4). Model averaging and PLS latent variable model allotted some effect to each predictor, albeit not significant.

2.4.3 Concluding Remarks

Even the best modeling efforts are crude approximations of the true underlying system, likely built using mere proxy variables for true (underlying) causal/functional variables.

Simulations allow for comparative studies where truth (causality) is known; however, simulations may provide an unrealistic set of conditions (Murtaugh 2009). Causation cannot be determined in observational/correlative studies; nevertheless, determining causal links (direct or indirect) typically is an unstated goal of many correlative land-cover studies. Unfortunately, we must make decisions on the inclusion of only a subset variables in a system, as it is impractical/impossible to include all functional ("causal") variables in an analysis, which unavoidably leads to estimation bias (Clarke 2005). There may be some consolation in knowing if a method is capable (under simulation settings) of correctly classifying predictor variables, yielding minimally biased coefficient estimates with low variability, and generating useful predictive models given new data.

Traditional null hypothesis testing is based on the unlikely and uninformative null hypothesis (H_0) of no effect, which provides no meaningful information (estimation of magnitude and precision) for management or planning (Anderson, Burnham and Thompson 2000). Alpha (α) level is arbitrarily chosen (traditionally 0.05) and sets the upper-limit for the FPR (type I error) when all functional variables are included and a threshold for classification of a predictor as having either no effect or a "statistically significant" effect. All things equal, and if H_0 is actually true, setting α smaller (as opposed to larger; e.g., 0.10 vs. 0.05) leads to more conservative testing and lower FPR. However, if we assume to be living not "under the null" (if H_0 were truly false), then setting α small leads only to lower TPR (power). In studies on the influence of land-cover on stream ecosystems, the ecological consequences of false positives (claiming an increase in some land-cover proportion has an undesirable effect when it doesn't) may not be as great as type II errors (1-power, false negatives: not making the claim when it

does) when the results of research is used to guide landscape management decisions (Johnson, 1999; Peterman 1990).

In our case study, PLS seems to have offered reasonable coefficient estimates given the estimates from the remaining models. An effect of exactly zero may not be realistic of real-world data, and in our simulation, PLS consistently allocated some small effect (non-zero estimate) to X_2 (the non-functional proportional variable). PLS does not produce partial regression coefficients, but instead uses the correlation structure of predictors to estimate latent structures maximally related to the response. While this might be considered an unattractive attribute of PLS, it might also be viewed as strength in observational studies where randomization does not occur and causal attribution is impossible and our predictors (e.g., land-cover proportions) are simple proxies for underlying and unmeasured processes anyway. PLS also offers some unique advantages relative to the other methods, including fewer and more realistic underlying assumptions compared to OLS (Wold et al. 2001), useful graphical representations of relationships within the data (Abdi 2010), and generally acts to shrink coefficients (Rosipal & Krämer 2006) offering conservative estimates of effect size in the face of collinearity.

We do not suggest or promote the use of small sample sizes in landscape-level studies, although, we realize this situation is a common and unfortunate reality in this field. Based on the results of this simulation study, we offer the following recommendations:

- Develop a well thought-out list of potential functionally related predictor variables for your study *a priori*;
- Avoid small sample sizes where possible; although potentially cost prohibitive, increasing n by a small number (e.x. n = 10 to 15) can lead to substantial improvements in terms of power, type-I-error, predictive performance and may be investment worthy;

- Avoid reporting of bivariate relationships in land-cover studies, as they can be highly misleading due to collinearity and an increase in type I errors;
- Consider use of alternative regression methods to OLS, especially PLS or shrinkage methods as they may better balance estimation bias and variability;
- Consider use of multiple/complementary methods/criteria (Mac Nally 2002; Morrice et al. 2008; Nathans, Oswald & Nimon 2012) when selecting between competing models;
- Be realistic regarding inference from observational data as it should be very limited, especially in small sample and collinear settings.

2.5 References

Abdi H (2010) Partial least squares regression and projection on latent structure regression (PLS Regression). Wiley Interdisciplinary Reviews: Computational Statistics 2:97-106.

Allan JD (2004) Landscapes and riverscapes: the influence of land use on stream ecosystems. Annual Review of Ecology, Evolution, and Systematics 35:257-284.

Anderson DR, Burnham KP (2002) Avoiding pitfalls when using information-theoretic methods. Journal of Wildlife Management 66:912-918.

Anderson DR, Burnham KP, Thompson WL (2000) Null hypothesis testing: problems, prevalence, and an alternative. Journal of wildlife management 64:912-923.

Arnold TW (2010) Uninformative parameters and model selection using Akaike's information criterion. Journal of Wildlife Management 74:1175-1178.

Batáry P, Fronczek S, Normann C, Scherber C, Tscharntke T (2014) How do edge effect and tree species diversity change bird diversity and avian nest survival in Germany's largest deciduous forest? Forest Ecology and Management 319:44-50.

Bartoń K (2013) MuMIn: Multi-model inference. R package version 1.9.13. Available at: http://CRAN.R-project.org/package=MuMIn

Bissonette JA (1999) Small sample size problems in wildlife ecology: a contingent analytical approach. Wildlife Biology 5:65-71.

Boulesteix A-L, Lambert-Lacroix S, Peyre J, Strimmer K (2012) plsgenomics: PLS analyses for genomics. R package version 1.2-6. Available at: http://CRAN.Rproject.org/package=plsgenomics.

Boulesteix A-L, Strimmer K (2007) Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. Briefings in Bioinformatics 8:32-44.

Buckland ST, Burnham KP, Augustin NH (1997) Model selection: an integral part of inference. Biometrics 53:603-618.

Burcher CL, Valett HM, Benfield EF (2007) The land-cover cascade: relationships coupling land and water. Ecology 88:228-242.

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag, New York.

Burnham KP, Anderson DR, Huyvaert KP (2011) AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. Behavioral Ecology and Sociobiology 65:23-35.

Carrascal LM, Galván I, Gordo O (2009) Partial least squares regression as an alternative to current regression methods used in ecology. Oikos 118:681-690.

Chong I-G, Jun C-H (2005) Performance of some variable selection methods when multicollinearity is present. Chemometrics and Intelligent Laboratory Systems 78:103-112.

Clarke KA (2005) The phantom menace: omitted variable bias in econometric research. Conflict Management and Peace Science 22:341-352.

Dahlgren JP (2010) Alternative regression methods are not considered in Murtaugh (2009) or by ecologists in general. Ecology Letters 13:E7-E9.

Dainese M (2011) Impact of land use intensity and temperature on the reproductive performance of Dactylis glomerata populations in the southeastern Alps. Plant Ecology 212:651-661.

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemuller T, McClean C, Osborne P, Reineking B, Schröder

B, Skidmore A, Zurell D, Lautenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:027-046.

Peña EA, Slate EH (2006) Global validation of linear model assumptions. Journal of the American Statistical Association 101:341-354.

Peña EA, Slate EH (2014) gvlma: Global Validation of Linear Models Assumptions. R package version 1.0.0.2. Available at: http://CRAN.R-project.org/package=gvlma

Efron B (2014) Estimation and accuracy after model selection. Journal of the American Statistical Association 109:991-1007.

Eriksson L (1999) Introduction to multi-and megavariate data analysis using projection methods (PCA & PLS). Umetrics AB, Umea, Sweden.

Farrar DE, Glauber RR (1967) Multicollinearity in regression analysis: the problem revisited. Review of Economics and Statistics 49:92-107.

Freckleton RP (2002) On the misuse of residuals in ecology: regression of residuals vs. multiple regression. Journal of Animal Ecology 71:542-545.

Genz A, Bretz F, Miwa T, Xuefei Mi, Leisch F, Scheipl F, Hothorn T (2014) mvtnorm: Multivariate Normal and t Distributions. R package version 0.9-9997. Available at: http://CRAN.R-project.org/package=mvtnorm

Graham MH (2003) Confronting multicollinearity in ecological multiple regression. Ecology 84:2809-2815.

Grömping U (2007) Estimators of relative importance in linear regression based on variance decomposition. American Statistician 61:139-147.

Grömping U (2009) Variable importance assessment in regression: linear regression versus random forest. American Statistician 63.308-319.

Grueber C, Nakagawa S, Laws R, Jamieson I (2011) Multimodel inference in ecology and evolution: challenges and solutions. Journal of Evolutionary Biology 24:699-711.

Hall P, Lee ER, Park BU (2009) Bootstrap-based penalty choice for the lasso, achieving oracle performance. Statistica Sinica 19:449-471.

Hastie T, Efron B (2013) lars: Least Angle Regression, Lasso and Forward Stagewise. R package version 1.2. Available at: http://CRAN.R-project.org/package=lars

Hesterberg T, Choi NH, Meier L, Fraley C (2008) Least angle and the l1 penalized regression: a review. Statistics Surveys 2:61-93.

Hurvich CM, Tsai C-L (1989) Regression and time series model selection in small samples. Biometrika 76:297-307.

Johnson DH (1999) The insignificance of statistical significance testing. The Journal of wildlife management, 63:763-772.

Johnson DH, Omland KS (2004) Model selection in ecology and evolution. Trends in Ecology & Evolution 19, 101-108.

Kelley K, Lai K (2012) MBESS: MBESS. R package version 3.3.3. Available at: http://CRAN.R-project.org/package=MBESS

King RS, Baker ME, Whigham DF, Weller DE, Jordan TE, Kazyak PF, Hurd MK (2005) Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15:137-153.

Lumley T (2009) Leaps: Regression Subset Selection. R Package Version 2.9. Available at: http://CRAN.R-project.org/package=leaps

Li B, Morris J, Martin EB (2002) Model selection for partial least squares regression. Chemometrics and Intelligent Laboratory Systems 64:79-89.

Mac Nally R (2000) Regression and model-building in conservation biology, biogeography and ecology: the distinction between – and reconciliation of – "predictive" and "explanatory" models. Biodiversity & Conservation 9:655-671.

Mac Nally R (2002) Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables. Biodiversity & Conservation 11:1397-1401.

Maxwell SE (2000) Sample size and multiple regression analysis. Psychological Methods 5:434-458.

McFarland A, Hauck LM (1999) Relating agricultural land uses to in-stream stormwater quality. Journal of Environmental Quality 28:836-844.

Mehmood T, Liland KH, Snipen L, Sæbø S (2012) A review of variable selection methods in partial least squares regression. Chemometrics and Intelligent Laboratory Systems 118:62-69.

Montgomery DC, Peck EA, Vining GG, Vining J (2001) Introduction to linear regression analysis: Wiley, New York.

Morrice JA, Danz NP, Regal RR, Kelly JR, Niemi GJ, Reavie ED, Hollenhorst T, Axler RP, Trebitz AS, Cotter AM (2008) Human influences on water quality in Great Lakes coastal wetlands. Environmental Management 41:347-357.

Murtaugh PA (2009) Performance of several variable selection methods applied to real ecological data. Ecology Letters 12:1061-1068.

Naiman RJ, Décamps H (1997) The ecology of interfaces: riparian zones. Annual Review of Ecology and Systematics 28:621-658.

Nathans LL, Oswald FL, Nimon K (2012) Interpreting multiple linear regression: a guidebook of variable importance. Practical Assessment, Research & Evaluation 17:1-19.

O'Brien RM (2007) A caution regarding rules of thumb for variance inflation factors. Quality & Quantity 41:673-690.

Peña EA, Slate EH (2006) Global validation of linear model assumptions. Journal of the American Statistical Association 101:341-354.

Peterman RM (1990) Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47:2-15.

R Development Core Team (2013) R version 3.0.1: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rice EW, Association, A.P.H. (2012) Standard methods for the examination of water and wastewater. American Public Health Association Washington, DC.

Rocha GV, Yu B (2008) Greedy and relaxed approximations to model selection: a simulation study. In Festschrift in Honor of Jorma Rissannen on the Occasion of his 75th Birthday, ser. TICSP Series, P. Grunwald, P. Myllymaki, I. Tabus, M. Weinberger, and B. Yu, Eds., Tampere International Center for Signal Processing, 38:63–80.

Rosipal R, Krämer N (2006) Overview and recent advances in partial least squares. In Subspace, Latent Structure and Feature Selection, pp. 34-51. Springer-Verlag, Berlin.

Shi ZH, Ai L, Li X, Huang XD, Wu GL, Liao W (2013) Partial least-squares regression for linking land-cover patterns to soil erosion and sediment yield in watersheds. Journal of Hydrology 498:165-176.

Smith AC, Koper N, Francis CM, Fahrig L (2009) Confronting collinearity: comparing methods for disentangling the effects of habitat loss and fragmentation. Landscape Ecology 24:1271-1285.

Speed R (1994) Regression type techniques and small samples: a guide to good practice. Journal of Marketing Management 10:89-104.

Symonds MRE, Moussalli A (2011) A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behavioral Ecology and Sociobiology 65:13-21.

Tibshirani RJ (1996) Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58:267-288.

Tibshirani RJ (2011) Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73:273-282.

Tibshirani RJ, Taylor J (2011) The solution path of the generalized lasso. Annals of Statistics 39:1335-1371.

Van Sickle J (2003) Analyzing correlations between stream and watershed attributes. Journal of the American Water Resources Association 39:717-726.

Venables WN, Ripley BD (2002) Modern applied statistics with S. Springer, New York.

Walsh CJ, Roy AH, Feminella JW, Cottingham PD, Groffman PM, Morgan RP (2005) The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24:706-723.

Whittingham MJ, Stephens PA, Bradbury RB, Freckleton RP (2006) Why do we still use stepwise modelling in ecology and behaviour? Journal of Animal Ecology 75:1182-1189.

Wold S, Ruhe A, Wold H, Dunn WJ (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM Journal on Scientific and Statistical Computing 5:735-743.

Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems 58:109-130.

Zhang Y, Dudgeon D, Cheng D, Thoe W, Fok L, Wang Z, Lee JH (2010) Impacts of land use and water quality on macroinvertebrate communities in the Pearl River drainage basin, China. Hydrobiologia 652:71-88.

Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67:301-320.

Zou H, Hastie T, Tibshirani R (2007) On the "degrees of freedom" of the Lasso. Annals of Statistics 35:2173-2192.

Zou H, Hastie T (2012) elasticnet: Elastic-Net for Sparse Estimation and Sparse PCA. R package version 1.1. Available at: http://CRAN.R-project.org/package=elasticnetLiterature Cited

Table 1. Methods and method-specific criteria used to indicate each variable as important (functional) or unimportant (non-functional) in explaining Y. Methods included ordinary least-squares (OLS), OLS model selection based on stepwise selection using finite sample corrected Akaike information criterion (AICc), multi-model averaging (MMA) and alternative methods. Importance criteria for MMA was based on summed AICc weights and thus common to both natural (MMA.n) and zero (MMA.z) averaging methods. Thresholds are given/shown in the order of more-, followed by less-conservative. See text for further information.

Model Name	**Model Description and Selection Type**	**Method-Specific Criteria and Thresholds**
Full Model	OLS w/ all **X**: no selection	$\hat{b}_i$ p-value < 0.10 & 0.15
Stepwise	Stepwise w/ AICc: Single best model	Standardized $\lvert\hat{b}_i\rvert$ > 0 & 0.05
MMA.n	Natural weighted averaging with AICc	Σ(AICc wt) > 0.70 & 0.50
MMA.z	Zero weighted averaging: let NA = 0	Σ(AICc wt) > 0.70 & 0.50
Lasso	Least absolute shrinkage and selection	Standardized $\lvert\hat{b}_i\rvert$ > 0 & 0.05
EN	Elastic Net: shrinkage and selection	Standardized $\lvert\hat{b}_i\rvert$ > 0 & 0.05
PLS	Partial Least-Squares: latent variable	PLS-VIP > 1.05 & 1.00

Table 2. Average pairwise Pearson product-moment correlation estimates (above diagonal) and correlation estimate variances (below diagonal) for Y and X variables simulated in this study (p = 7 and n = 20). Simulation average variance inflation factors ($\overline{VIF}$) for X variables are given in the bottom row. Correlation average estimates and variance for the p = 3 simulation are virtually identical to those given for Y, X_1, X_2, and X_3 below; $\overline{VIF}$ values for p = 3 were 3.61, 3.59, and 5.35, respectively.

	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y	-----	0.81	0.68	-0.82	0.65	-0.57	0.66	-0.57
X_1	0.01	-----	0.69	-0.81	0.79	-0.58	0.79	-0.59
X_2	0.02	0.02	-----	-0.81	0.59	-0.79	0.59	-0.79
X_3	0.01	0.01	0.01	-----	-0.66	0.66	-0.66	0.66
X_4	0.02	0.01	0.03	0.02	-----	-0.69	0.68	-0.69
X_5	0.03	0.03	0.01	0.02	0.02	-----	-0.69	0.69
X_6	0.02	0.01	0.02	0.02	0.02	0.02	-----	-0.69
X_7	0.03	0.03	0.01	0.02	0.02	0.02	0.02	-----
$\overline{VIF}$	-----	29.2	29.1	7.4	15.8	16.2	16.1	15.9

Table 3. Simulation results (P = 3) for coefficient estimation across 2 levels of sample size (15 and 20). Individual OLS models are represented as rows; boldfaced row indicates the true model (b_1 = 1.00, b_2 = 0.00, b_3 = -0.50). Average estimates ($\overline{\hat{b}}$) are provided and arrows (↑↓) indicate direction of bias (> |0.01|) from true *b* values. True positive rate (TPR; power) or false positive rate (FPR; type I error) is provided in parentheses; α = 0.10 for all calculations. Average values for ΔAIC_c ($\overline{\Delta}$) are provided for each model, along with the % of simulation runs where $\overline{\Delta}$ was ≤ specified thresholds.

N	Model	$\overline{\hat{b}}_1$(Power)	$\overline{\hat{b}}_2$(FPR)	$\overline{\hat{b}}_3$(Power)	$\overline{\Delta}$ (%Δ=0, %Δ <2, %Δ <4)
	1	0.996 (0.52)	-0.003 (0.10)	-0.506 (0.41)	5.56 (0.01, 0.04, 0.18)
	2	------	↓-0.101 (0.10)	↓-0.854 (0.84)	5.82 (0.03, 0.09, 0.50)
	3	**0.998 (0.54)**	**------**	**-0.506 (0.56)**	**2.24 (0.17, 0.44, 0.94)**
15	4	↑1.507 (0.88)	↑0.493 (0.24)	------	4.98 (0.07, 0.16, 0.57)
	5	------	------	↓-0.895 (0.99)	3.19 (0.36, 0.52, 0.67)
	6	------	↑1.549 (0.93)	------	11.20 (0.02, 0.06, 0.11)
	7	↑1.849 (0.99)	------	------	3.62 (0.34, 0.49, 0.63)
	1	0.992 (0.71)	-0.018 (0.10)	-0.503 (0.51)	3.50 (0.03, 0.12, 0.71)
	2	------	↓-0.175 (0.12)	↓-0.940 (0.97)	6.98 (0.03, 0.12, 0.39)
	3	**0.997 (0.74)**	**------**	**-0.496 (0.73)**	**1.22 (0.42, 0.74, 0.96)**
20	4	↑1.504 (0.99)	↑0.485 (0.41)	------	4.69 (0.10, 0.25, 0.56)
	5	------	------	↓-0.877 (0.99)	5.16 (0.21, 0.52, 0.65)
	6	------	↑1.239 (0.87)	------	18.08 (0.00, 0.01, 0.04)
	7	↑1.751 (0.99)	------	------	5.05 (0.21, 0.38, 0.53)

Table 4. Coefficient estimation mean squared-error (MSE = variance + bias^2) for each method. Table values represent MSE values at sample sizes 10, 15, 20, 25, respectively for $b_1 - b_4$ and total across all coefficients. Coefficients $b_5 - b_7$ behaved similarly to those for b_4 (Fig 2) and were omitted to save space.

Method	b_1	b_2	b_3	b_4	Total
Full Model	17.88, 3.67, 1.85, 1.20	19.04, 3.40, 1.91, 1.29	0.99, 0.21, 0.11, 0.08	1.33, 0.26, 0.13, 0.09	41.95, 8.04, 4.28, 2.85
Stepwise	1.22, 1.20, 0.92, 0.72	0.81, 0.83, 0.63, 0.44	0.28, 0.21, 0.16, 0.12	0.08, 0.06, 0.04 0.03	2.50, 2.42, 1.84, 1.39
MMA.n	0.82, 0.55, 0.41, 0.29	1.13, 0.62, 0.47, 0.34	0.17, 0.12, 0.08, 0.06	0.12, 0.06, 0.04, 0.03	2.44, 1.45, 1.09, 0.78
MMA.z	0.68, 0.55, 0.48, 0.39	0.21, 0.18, 0.15, 0.11	0.16, 0.13, 0.10, 0.09	0.02, 0.02, 0.01, 0.01	1.11, 0.91, 0.76, 0.61
Lasso	0.577, 0.40, 0.32, 0.26	0.11, 0.10, 0.09, 0.07	0.14, 0.09, 0.07, 0.05	0.01, 0.01, 0.01, 0.01	0.86, 0.63, 0.50, 0.41
EN	0.55, 0.37, 0.29, 0.23	0.14, 0.11, 0.11, 0.08	0.13, 0.08, 0.06, 0.05	0.02, 0.02, 0.01, 0.01	0.84, 0.59, 0.49, 0.38
PLS	0.72, 0.42, 0.26, 0.20	0.79, 0.40, 0.27, 0.27	0.15, 0.09, 0.05, 0.04	0.07, 0.05, 0.03, 0.02	1.90, 1.03, 0.67, 0.58

Table 5. Top: OLS estimates and standard errors (SE, in parentheses) of regression slopes for $[NO_3^-]$ (mg L^{-1}) and all model subsets of three land-cover classes, including the intercept only (null) model (OLS 8, where RSS = TSS). Bold indicates 90% confidence interval (CI, as 1.645*SE) didn't contain 0. Observed residual sum of squares (RSS), unadjusted R^2 (as 1 – (RSS/TSS)), adjusted R^2, ΔAIC_c, $AICc_{wt}$ (weights), and leave-one-out predicted residual sum of squares (PRESS) are provided for each model (except OLS 8). Bottom: Stepwise, AICc multi-model averages, Lasso, Elastic Net (EN) and PLS regression slopes. Bold indicates 90% bootstrap CI (see Fig. 6) didn't contain 0; † denotes 85% bootstrap CI did not contain zero (for top of table as well, all other bootstrap CIs agreed with frequentist CIs). Coefficient estimates ($\widehat{\boldsymbol{b}}$) following a comma (right side) are bootstrap median values ("bagged") of the corresponding left hand observed values.

Models	$\widehat{b}_{\%\mathbf{FB}}$	$\widehat{b}_{\%\mathbf{Ag}}$	$\widehat{b}_{\%\mathbf{ISC}}$	RSS	$\mathbf{R^2}$	$\mathbf{R^2_{adj}}$	ΔAICc	$AICc_{wt}$	PRESS
OLS 1	**-0.046 (0.014)**	0.004 (0.017)	-0.101 (0.068)	1.988	0.753	0.670	5.480	0.029	4.698
OLS 2	-----	**0.045 (0.017)**	0.070 (0.064)	4.523	0.437	0.325	10.590	0.002	6.986
OLS 3	**-0.048 (0.009)**	-----	**-0.110 (0.053)**	2.003	0.751	0.701	0	0.457	3.598
OLS 4	**-0.031 (0.010)**	0.018† (0.015)	-----	2.479	0.692	0.630	2.772	0.114	4.167
OLS 5	**-0.037 (0.008)**	-----	-----	2.867	0.643	0.611	0.330	0.387	3.629
OLS 6	-----	**0.043 (0.017)**	-----	5.064	0.370	0.313	7.725	0.010	7.003
OLS 7	-----	-----	0.058 (0.080)	7.671	0.046	-0.041	13.125	0.001	10.441
OLS 8	-----	-----	-----	8.040	0	-----	10.269	0.002	-----
Stepwise	-0.048, -0.045	0, 0	-0.110, -0.080	2.003	0.751	-----	-----	-----	7.374
MMA.n	**-0.042**, -0.042	0.018, 0.018	-0.109, -0.090	2.826	0.648	-----	-----	-----	4.269
MMA.z	**-0.041**, -0.039	0.003, 0.004	-0.053, -0.041	2.176	0.729	-----	-----	-----	4.685
Lasso	**-0.028**, -0.031	0.013, 0.012	0, 0	2.636	0.672	-----	-----	-----	5.408
EN	-0.028†, -0.034	0.016, 0.015	0, 0	2.541	0.683	-----	-----	-----	5.268
PLS	**-0.030**, -0.035	0.016, 0.017	0.013, -0.020	2.625	0.674	-----	-----	-----	3.766

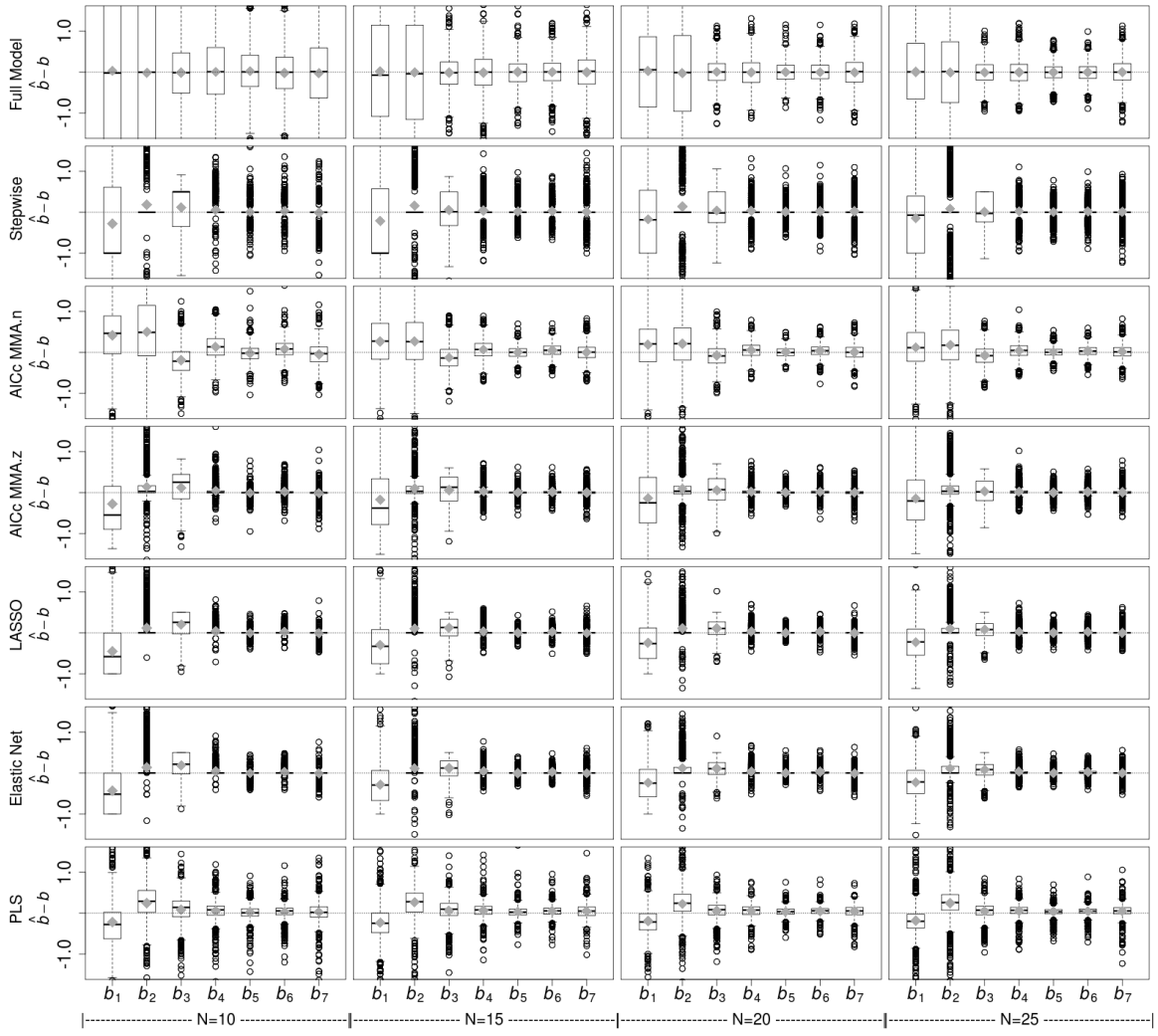

Figure 1. Box plots of the difference between estimated beta and true beta ($\hat{b} - b$) for each iteration and analytical method across a range of sample sizes from 10 (left column of panels) to 25 (right column of panels). Light gray diamonds show mean value and horizontal lines show zero line (unbiased = 0). For all coefficients except b_3, the sign of the y-axis indicates the direction of bias; the true value for b_3 is negative, thus the opposite is true for this coefficient. Plot y-range was restricted to allow for assessment of most data variability; consequently, observations fall outside of plot margins.

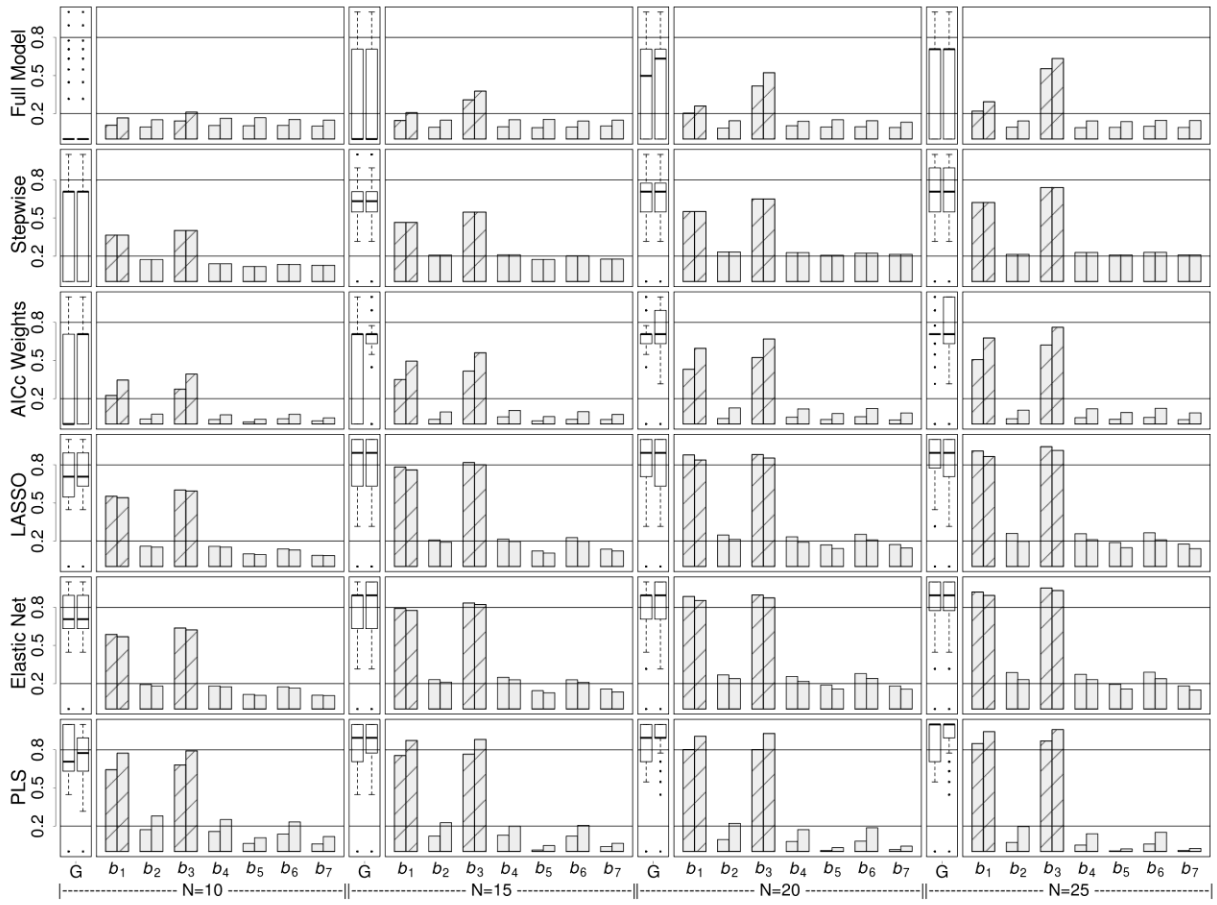

Figure 2. Box plots of geometric mean of % correctly classified variables (functional, or not) each simulation iteration according to method specific importance criteria (MSIC; see text and Table 1). Bar plots show the frequency that each variable was identified as important, true positive rate for functional predictors (striped bars: b_1, b_3) or else false positive rate (b_2, b_4, b_5, b_6, b_7) according to the MSIC. Multiple bars indicate values for more and less conservative cutoff levels: p-values: 0.10, 0.15; AICc wts: 0.70, 0.50; PLS VIP: 1.05, 1.00; Stepwise/Lasso/EN: $\hat{b} \neq 0$ and standardized $\hat{b} > 0.05$.

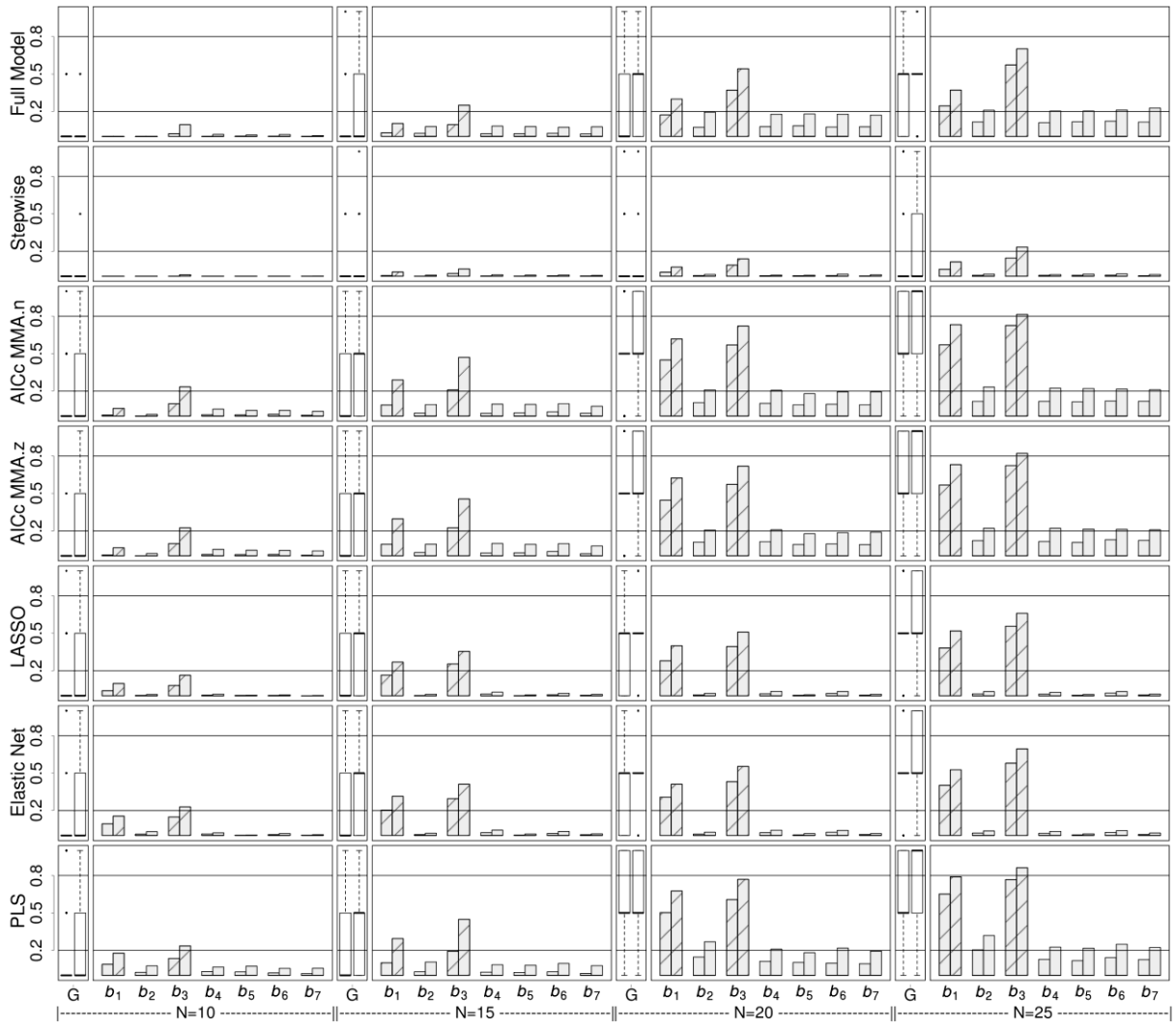

Figure 3. Box plots of geometric mean of % correctly classified variables (functional and non-functional) each simulation iteration. Bar plots show the frequency that each variable was identified as important [power (striped bars: b_1, b_3) or false positive rates (b_2, b_4, b_5, b_6, b_7)] according to bootstrapped CI (containing 0 or not). Side-by-side boxplots and barplots show results for 2 confidence levels: 85% and 75%. Note: true model contains X1 & X3 only.

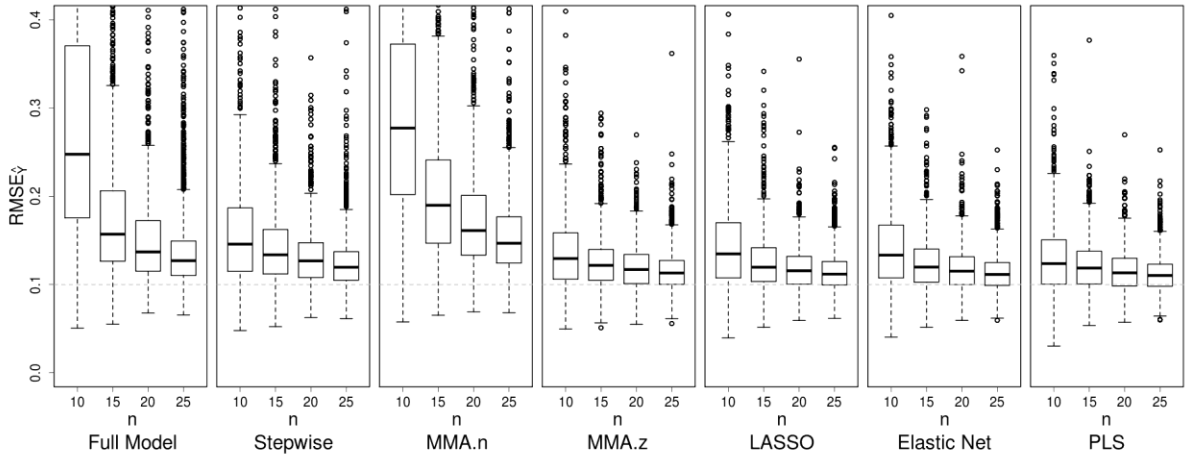

Figure 4. Box plots of prediction accuracy (as root mean squared error, $RMSE_{\hat{Y}}$) on test data set for each analytical method across sample sizes (n; see Table 1 for summary of methods). Plot y-range is restricted to allow for detailed comparison; consequently, some observations fall outside of plot margin. Grey horizontal line indicates standard deviation of error term in data creation.

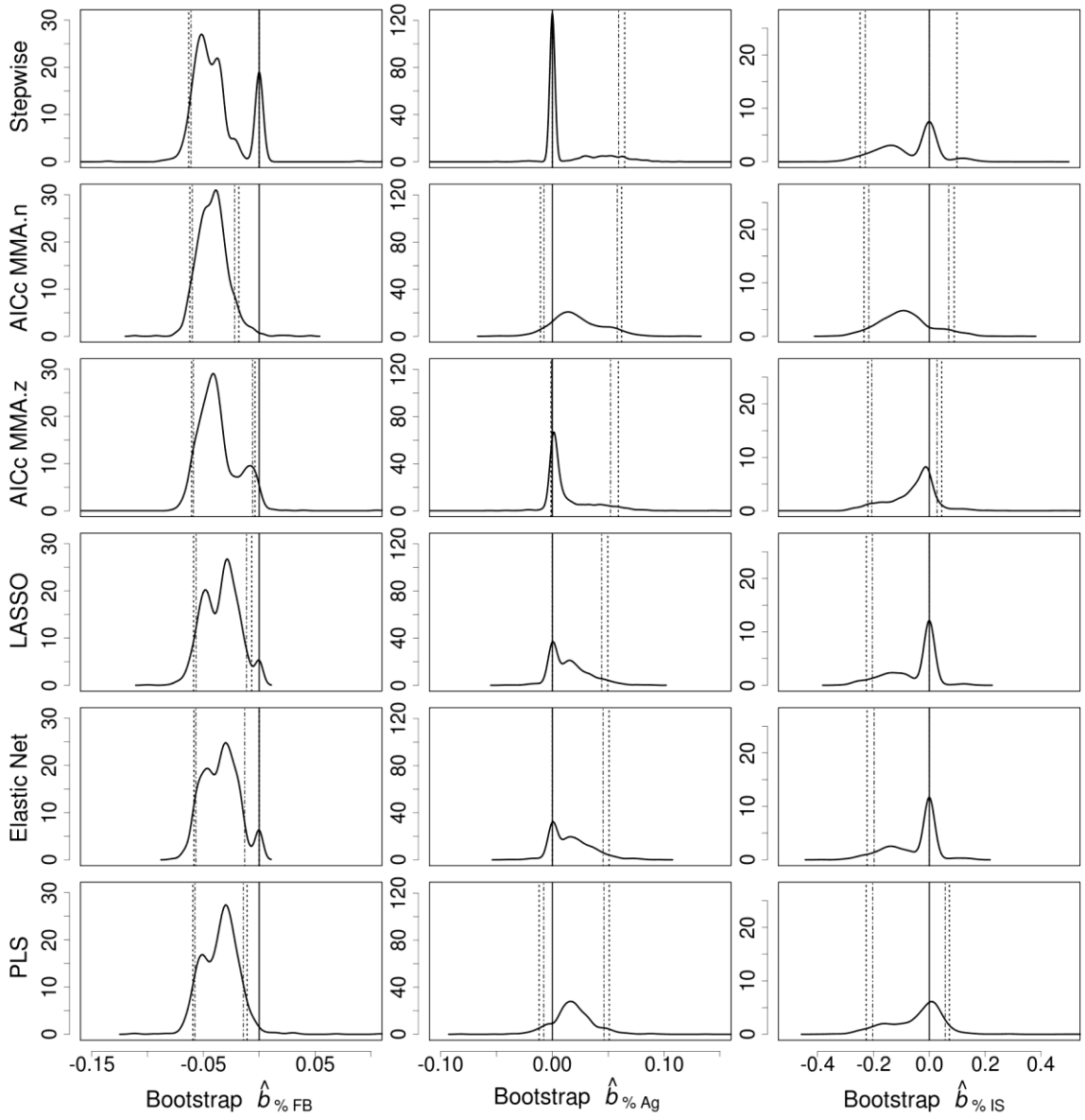

Figure 5. Smoothed density curves of (resampled) bootstrap slope estimates ($\hat{b}$) for the stream nitrate concentration case study. Solid vertical line indicates 0 values for estimated slope values (x-axis). Dotted lines show 90% and dotted/dashed lines show 85% bootstrap confidence intervals (percentile method). Additional details can be found in Table 4.

Chapter 3. Influence of low-intensity watershed development on small coastal Alabama streams: an analysis using partial least-squares

3.1 Introduction

Deforestation from human activity within watersheds can greatly affect stream ecosystems (Allan, 2004; Walsh et al., 2005). Forested lands, specifically riparian forests bordering stream channels, are thought to act as material sinks and natural filters, mechanistically linked to in-stream sediment and nutrient conditions (Naiman & Décamps, 1997). Changes in these physicochemical conditions occur in both agricultural and urbanized watersheds, as both act as sources of nutrients, sediment, and other pollutants (Paul and Meyer 2001; Allan 2004; Walsh et al. 2005; O'Driscoll et al. 2010; Nagy et al. 2011). Urbanization is a particularly influential land-use/cover (LULC) type that is characterized by high levels of impervious surface cover (ISC), which can directly alter watershed hydrology (Brabec, 2002; Brown et al., 2009). ISC is generally thought to increase the flashiness (rate of change), frequency, and magnitude of flood events and decrease flood duration (Rose and Peters 2001; Brown et al. 2009; O'Driscoll et al. 2010).

Stream benthic macroinvertebrates are a diverse group of organisms that exhibit a wide range of environmental tolerances and are frequently used as indicators of stream ecosystem health (Bonada et al. 2006; Kenney et al. 2009). The ubiquitous and generally sedentary nature of benthic macroinvertebrates in freshwaters underlies their use as indicators of local water quality (Bonada et al. 2006). Benthic richness, diversity and the proportion of sensitive taxa are typically observed to be lower in urbanized streams compared to forested sites (Norris & Thoms,

1999; Paul &Meyer, 2001; Walsh et al., 2005; Wenger et al., 2009).

Urbanization impacts on streams are generally higher and more detectable at a much lower areal proportions than agricultural impacts (Allan 2004). Generalized urbanization thresholds (above which impacts are detectable, e.g., ISC <10%) have been suggested (Paul & Meyer 2001), but these purported thresholds likely misrepresent what is a continuum of impacts on stream ecosystems, even at low urbanization (Brabec 2002; Shuster et al. 2005). To date, most urban impact studies have focused primarily on relatively high levels of development, whereas, low-density urban development may be more spatially abundant and encouraged by LULC policy (Cunningham et al. 2009; Chadwick et al. 2011). Recently, lower levels of urban development have gained in research attention as ISC of ≤ 10% has been linked with altered stream hydrology and physicochemistry, and reduced species richness (Burcher and Benfield 2006; Lussier et al. 2008; Cunningham et al. 2009; Nagy et al., 2012). A few studies have detected altered macroinvertebrate assemblage responses to much lower urbanization levels (4.4% ISC, Wenger et al. 2009); although much additional work is needed in more diverse types of systems before generalities can be formed regarding instream impacts.

It has been suggested that the magnitude of hydro-geomorphic response to watershed urbanization is comparatively less in lowland coastal plains streams than in higher-gradient upland regions (e.g., Piedmont; Nagy et al. 2011; Utz et al. 2011). In the southeastern US, virtually all of the coastal plains land area has been altered by human activity over the last 2 centuries, with more recent change stemming from forested/agricultural land converted to urban land (Smock & Gilinsky, 1992; O'Driscoll et al., 2010; Nagy et al., 2011). In areas with historical agriculture, low-density urbanization may not be great enough to produce detectable changes in hydrology, geomorphology, or physicochemistry (Burcher & Benfield, 2006). Coastal

plains watersheds are typically low-gradient with sandy well-drained soils, and thus have lower runoff-rainfall ratios, particle entrainment, and hydrogeomorphic alteration than higher-gradient upland watersheds (Feeley 1992; Nagy et al. 2011; Nagy et al. 2012). Thus, it might be predicted that low-levels of development would not lead to detectable changes in low-gradient coastal streams, especially where historical agricultural may have already affected these systems (Nagy et al. 2012).

Coastal areas worldwide are under increasing pressures from human population growth and associated land development (Nagy et al. 2011). Roughly half of the urban LULC change along the US Gulf of Mexico in the past 2 decades has occurred within 50 km of the coast, with the dominant LULC change from the Florida panhandle to Louisiana being low-intensity development (e.g., suburban, urban-sprawl; White et al. 2008; Xian et al. 2012). Furthermore, the Southeast is predicted to continue leading the US in population growth developed into 2030, with low-density development continuing along the coast of the Gulf of Mexico (Wear and Greis 2002; Alig et al. 2004; White et al. 2008; Xian et al. 2012); thus, there is an urgent need to understand and predict impacts of low-density development on stream ecosystems in this rapidly changing region.

Identifying potential mechanisms of LULC impacts on stream ecosystems is critically important for prescriptive management, and modeling effects based on theoretically derived causal relationships that should improve predictive performance (Allan, 2004; Mac Nally, 2000). Collinearity among proportionally based LULC categories is inherently problematic and can lead to inflated variance about coefficient estimates (Montgomery et al. 2001). Bivariate analyses based on these data likely lead to spurious results (Graham, 2003; King et al., 2005; Varanka & Luoto, 2012), and commonly used variable selection methods (e.g., stepwise) have been shown

to frequently incorrectly identify predictor variables as important (Hegyi and Garamszegi 2011; see Chapter 1). Alternative regression methods have been shown to perform well with small sample sizes and in collinear situations at correctly identifying important variables and providing reasonable (low bias, low variance) coefficient estimates (Dahlgren, 2010; see Chapter 1).

Low-density development along the Gulf of Mexico is common and predicted to become more prevalent. Due to the extensive history of agriculture in the coastal plains, it is possible that low levels of ISC have little influence on stream hydro-geomorphology, physicochemistry or benthic assemblages. The goal of the current study is to determine if empirical evidence suggests that low-density development influences the hydrology, geomorphology, physicochemistry and/or macroinvertebrate assemblages of streams draining small coastal watersheds.

We collected data from 13 stream sites in and adjacent to a small coastal Alabama town that ranged in ISC from only 1 to 11% so that our analyses would not be influenced by more urbanized sites. We used alternative regression methods that have been shown to outperform traditional regression and variable selection methods in identifying important predictors with small n and collinear data.

3.2 Methods

3.2.1 Study Area

We studied 13 sandy-bottom channel, non-tidal (salinity range: 0.02 - 0.04), wadeable streams (1^{st} to 3^{rd} order) spanning a gradient of low ISC ($\leq$ 11%) within or adjacent to the Wolf Bay Basin, southern Baldwin County, Alabama, USA (Fig. 1). The SE coastal plains (CPL) are a relatively understudied region that differs from its highland counterparts with regards to several potentially important physical/climatological characteristics. The CPL has higher annual precipitation levels and rainfall:runoff ratios, shorter recurrence intervals for bankfull events

(0.19 to 1.0/y) and CPL streams typically have more unstable beds than upland regions (Sweet and Geratz 2003; Hardison et al. 2009; Metcalf 2009; Nagy et al. 2011). The stream sites were located in close proximity to one another, not connected to major drainage basins and drained directly into Perdido Bay, Wolf Bay, or Weeks Bay. At each stream site, we established a 100-m study reach chosen to be ≥ 50-m upstream of the nearest stream-road intersection, the typical access point, to minimize the proximate effects (e.g., sediment inputs) of roads,. Study reaches included run-pool sequences and were typically dominated by sandy substrate and/or organic detritus.

For LULC characterization, we delineated sub-watersheds 10-m digital elevation maps (DEM; source: USGS) using ArcHydro (Environmental Research Systems Institute, Inc., Redland, California) and contained all land area upstream of sampling location. We quantified LULC with and 0.15-m resolution aerial photographs (2009) of Baldwin County; LULC classification was based on manual digitization of aerial photographs in ArcGIS (ESRI, Inc., Redland, California). We chose LULC categories that could be theoretically directly linked to instream conditions: % ISC (e.g., buildings, paved roads), % agricultural cover (Ag), and% riparian forest buffer (FB). While the effectiveness of a FB width likely varies according to landform (e.g., slope) and with the goals of the buffer (Osborne and Kovacic 1993), we chose a commonly used FB width of 100-m laterally from the stream (Allan, 2004; Burcher, Valett & Benfield, 2007).

3.2.2 Instream Sampling and Response Variables

Stream hydrology.- We calculated a suite of hydrologic variables using stream stage and discharge data. We used Solnist pressure transducers (Levelogger Gold, model 3001) installed in each study reach to quantify water stage (height above fixed datum) and water temperature at

15-min intervals; barometers (Solnist Barologger Gold) were installed to adjust stage data for atmospheric pressure variation across the study area (period of record: March 2008 to March 2009). We estimated discharge (Q) at various stage levels during the study period (~ 1.5 yrs) using the velocity-area method (Raghunath 2007); velocity was measured using a Marsh-McBirney Model 2000 Flo-Mate (Marsh-McBirney, Inc.). We developed rating curves to convert stage to Q for each study stream using a combination of observed and estimated Q values. We estimated Q for high-flow events that prohibited instream measurements in subsections (Arcement and Schneider, 1989); in-channel flows were estimated with an equation derived specifically for sandy-bottomed streams (Sefick et al., 2015), overbank flows were estimated with Manning's equation (roughness coefficient = 0.15, Arcement and Schneider, 1989).

We quantified several aspects of stream-channel and floodplain geomorphology with three cross-sectional surveys per reach at each stream in close proximity to the stage recorders. We made width and depth measurements in the approximate bankfull channel in the field and derived average floodplain slopes perpendicular to the active channel from field measurements and 10m DEM maps using GIS. We determined bankfull depth (BF_d) and width (BF_w) as the elevation corresponding to the minimum of the ratio of width to depth (Pickup & Warner, 1976; Copeland et al. 2000). Hydrologic extremes may be more biologically/ecologically relevant than characterizing average conditions, so we also calculated cross-sectional dimensions (Max width and Max depth) at maximum flood stage from 1-dimensional models created from time-series stage data and data from cross-sectional surveys.

We selected a single representative hydrologic metric from each important flow category (i.e., frequency, duration, intensity; Clausen & Biggs, 2000), rather than several metrics/category, to reduce the likelihood of spurious relationships between hydrology and

LULC and instream biotic/habitat variables. We characterized relative high flow magnitude with unit-area Q, standardized by watershed area (Konrad et al., 2005). Event frequency and duration are generally characterized relative to set thresholds; while thresholds between 3 and 9 times the median have been used (Monk et al. 2006; Schoonover et al. 2006). Frequency and duration metrics varying in only thresholds are generally highly correlated (Clausen & Biggs, 2000); thus we characterized event frequency and duration as the frequency and number of h Q exceeded 7x the median discharge (Q fre7 and Q dur7, respectively). For intensity, many studies have used the Richard-Baker flashiness index (RBI; Baker et al., 2007), calculated as the sum of the absolute value of daily changes in Q divided by total Q over the period of record. In our work, we observed that storm flood duration frequently was <<24 h (see also Phillips & Scatena, 2010). As a result, we used a stage-based (*sensu* McMahon et al., 2003) flashiness metric calculated as the 97.5^{th} percentile of positive changes in stage (ΔSTG; rising limb) to be a proxy of maximum flashiness (less sensitive to erratic fluctuations in 15-min data). The frequency of stage increases in absolute terms is likely more ecologically relevant than multiples of site-specific medians (e.g., Q fre7); therefore, we also calculated frequency of stage increases ≥ 1 m (STG fre1.0) to assess its relationship with benthic metrics. In addition, we characterized the contribution of ground-water Q (baseflow) to total Q with a baseflow index (BFI = baseflow/total Q); alternatively, 1-BFI can be thought of as a flood-flow index (Clausen & Biggs, 2000). Baseflow was estimated using a standard automated algorithm (three pass recursive digital filter; Nathan & McMahon, 1990).

Stream physicochemistry.- We collected water samples to characterize streamwater physicochemical conditions at each study site (8-10 dates/site). Grab samples were collected during low-flow periods (no rain ≥ 3 days prior) in acid-washed and deionized water rinsed

polypropylene bottles, which were rinsed again with stream water at the sites prior to sample collection. Grab samples were carefully collected from mid-channel at the most downstream portion of the study reach prior to additional field work to prevent the risk of stream disturbance (Lurry and Kolbe, 2000). Samples were kept on ice and refrigerated until analyzed (< 24h). Concentrations of total N (TN; mg L^{-1}) and P (TP; mg L^{-1}) were determined at an independent lab by persulfate digestion using standard procedures (Rice & Association, 2012). Dissolved organic carbon (DOC) concentrations were determined with a Shimadzu TOC/TN analyzer (Shimadzu Scientific Instruments, Columbia, MD). We quantified total suspended solids (TSS; mg L^{-1}) using the volumetric filtration method (Wallace et al. 2006), whereas we determined stream water pH, dissolved oxygen (DO; mg L^{-1}) and specific conductivity (SPC; µS cm^{-1}) *in-situ* with a handheld multi-probe sonde (YSI, Yellow Springs, OH).). We focused on a select few physicochemical variables frequently associated with LULC disturbance (i.e. TP, SPC) and those that may be biologically relevant (i.e. pH, DO); some additional variables were quantified, but were redundant and excluded from analyses (i.e. $NO_3 \cong$ TN, TDS $\cong$ SPC).

We quantified benthic organic matter (BOM) at 10 spots along each stream reach using a 2.5 x 5 cm PVC corer. Core samples were dried in pre-weighed crucibles at 50°C for 48 h before combustion at 550°C for 3 h. Dry and post-combustion weights were recorded and BOM was reported as percent ash-free dry mass (Steinman et al. 2006). In addition, we quantified wetted widths and depths at 10 cross sections over the 100m stream reach during benthic sampling. At each cross section, large woody debris (LWD; wood > 2.5 cm in diameter) was estimated using a modified line-intercept method (Lamberti and Gregory 2006), and expressed as % of total wetted surface channel area.

Biotic sampling.- We sampled benthic macroinvertebrates at each site in the fall (Oct 2008) and spring (March 2009) using a Surber sampler (250 μm mesh, 0.093m^2 per quadrat) in 3 randomly selected run habitats (3 quadrats per run, total area sampled per site = 0.84 m^2) along each stream reach. We sampled run habitats because they were common among streams and typically contained a large proportion of exposed woody debris, which should provide relatively stable substrate and should host a high diversity of organisms (Benke et al. 1985). Invertebrate samples were preserved in 95% EtOH, transported to the laboratory and stored in at 4 C until processed, where we used a 2-phase method (Feminella 1996). Samples were coarse picked with the unaided eye for ≥30 min to remove large (> 2mm) organisms; remaining material was volumetrically subsampled from a homogenized, 1000 mL suspension. Several 50 mL aliquot subsamples (≥ 3 aliquots per sample) were removed from the total suspension and picked at random microscopically until a minimum total of 300 individuals were removed; this method has been found to reduce processing time and yield low within-sample variation (coefficient of variation < 10-15%, Feminella 1996).

We identified invertebrates, mostly aquatic insects, to the lowest practical taxonomic level (usually genus; Epler 2001, Merritt and Cummins 2007). We identified oligochaetes to the order level and excluded small meiofauna that were not reliably sampled or the focus of our study (arachnida, cladocera, ostracoda, and copepoda). We measured individuals for body length to the nearest mm. Enumerations from subsamples were extrapolated according to the fraction of the sample examined to estimate totals for the entire samples.

We focused on commonly used benthic metrics and those generally negatively associated with LULC change as response variables. Among the metrics calculated were numerical density (ind. m^{-2}), genus richness, diversity (as Shannon's *H'*), Pielou's evenness, percent

Ephemeroptera, Plecoptera and Trichoptera (%EPT) excluding purported tolerant families Hydropsychidae and Baetidae (% EPT -H/B; Merriam et al. 2011, but see Chang et al. 2014), and abundance weighted average pollution tolerance values (PTV; U.S. Environmental Protection Agency 2013) (Table 1). We noticed large differences in invertebrate densities during the benthic sorting process and calculated rarefied versions of all benthic metrics to remove the effect of sample sizes on benthic metrics (Gotelli & Colwell, 2001). Rarefaction was achieved by resampling 100 individuals from each site 1000 times without replacement, all metrics were calculated and stored for each resample and mean values were used as rarefied metrics (Walker, Poos & Jackson, 2008).

We included additional benthic metrics that were not a major focus of the study to help describe the benthic assemblages, as coastal streams are relatively understudied. We included compositional metrics describing proportions of dominant groups (e.g., % Chironomidae) and some based on invertebrate traits (Table 1). Multivoltinism, small body size, high fecundity, high dispersal abilities, streamlined/fusiform body shape, fast flow preference, and the ability to diapause were chosen as these may provide resistance or resilience to frequent flood disturbance or extreme temperatures and drought (Townsend and Hildrew, 1994). Trait-based metrics used information derived from the USGS trait database at the genus level (Vieira et al. 2006). Trait information was unavailable for a few taxa. The abundance of traits within an assemblage has been shown to be accurately estimated by family or genus level trait information (Dolédec et al. 2000), therefore we supplemented respective family-level traits where needed. We included small body size (< 2.5 mm) trait based on individual length measurements from this study.

3.2.3 Statistical Analysis

Partial least-squares.- We described relationships between LULC, potential in-stream stressors and commonly used macroinvertebrate assemblage metrics using partial least-squares (PLS). In traditional regression analyses, small sample sizes and inherent collinearity among proportionally based LULC categories are problematic, as both can lead to inflated variance about coefficient estimates (Graham, 2003; King et al., 2005; Varanka & Luoto, 2012). In contrast, the latent-variable method PLS uses the correlation structure of predictor variables (**X**) to construct new uncorrelated/orthogonal variables (latent components) that are linear combinations of **X** most related to the response variable(s) and shows comparatively lower estimation variance (Wold et al., 1984; see Chapter 1). In addition, PLS assumes that the system under investigation is a function of a few unmeasured "latent" variables and assumes very little about the observed data or residual distribution (Rosipal and Krämer, 2006). This property of PLS may be appropriate in observational studies, where true causal relationships are rarely known and collinear LULC variables may simply be proxies for ultimate causal factors.

Univariate response.- We determined the number of PLS components by choosing the minimum AICc, calculated with model residual sum-of-squares (RSS) and the number of PLS components retained (Li, Morris & Martin, 2002). With PLS, "variable importance in projection" (VIP) scores are used to indicate variables that contribute greatly to the model (Mehmood et al., 2012). A recent simulation study suggested that PLS performs relatively well in small sample and collinear situations at providing reasonable coefficient estimates and correctly identifying important predictors when VIP > 1 or slightly greater (e.g., 1.05) is used as selection criteria (see Chapter 1).

For each abiotic response variable, we incorporated the uncertainty associated with component determination into confidence interval calculations by bootstrapping the selection

process (Buckland, Burnham & Augustin, 1997). We used the percentile method for confidence interval (CI) estimation and provided bootstrap averaged ("bagged") coefficient estimates that may more accurately estimate coefficients from model selection (Bühlmann and Yu 2002, Efron, 2014). PLS models were created using standardized **X** variables, as PLS is sensitive to the variation among **X** because latent variables are created that maximize the covariance between predictors and response variables. Regression coefficients were unstandardized for models used to assess relationships between LULC, hydrologic and physicochemical variables, as LULC classes are on the same scale (%).

To analyze relationships between benthic metrics and stressors associated with LULC change, we incorporated a relatively large number of predictor variables (compared to sample size) that may potentially influence benthic assemblages (e.g., STG fre1.0, maximum flood cross-sectional dimensions). Prior to PLS model building, we used principal component analysis (PCA) to examine expected redundancy in benthic diversity/sensitivity metrics; if high levels of redundancy existed, we considered the use of principle axes as responses instead of the original variables. For each biotic response variable, we first conservatively reduced the predictor variable set by omitting variables with VIP scores < 0.70, and refit the PLS model with this reduced variable set (Mehmood et al., 2012; Wold, Sjöström & Eriksson, 2001). We incorporated the uncertainty associated with both variable reduction and component determination into confidence interval calculations by bootstrapping the selection process as mentioned above. Standardized coefficients were reported for ease of comparison between **X** variables which were measured on different scales.

Multivariate response.- We graphically summarized relationships between multiple benthic metrics and environmental variables with multi-response PLS models. The predictive

performance of multi-response PLS can be poor relative to that for single-response, especially when predictors are unrelated (Garthwaite 1994). Multi-response PLS, however, is mathematically similar to canonical correlation analysis (CCA; Rosipal and Krämer 2006, He et al. 2015) and can be useful for visual assessment of relationships between two sets of variables. Initial models were run and final models incorporated only variables with VIP > 0.7 in the initial run. For brevity, we provide only graphical output for final models. We created 3 separate multi-response models, one to describe diversity and sensitivity metrics of central interest to this study, as well as a compositional model and functional/trait model.

We investigated the degree of collinearity between LULC categories using simple correlation analysis and variance inflation factors (VIF; Montgomery et al. 2001). Prior to analyses, we transformed ($\log_{10}$ or square-root) some variables (indicated in text) to alleviate potentially influential observations and to improve data spread and linearity of relationships (Zuur, Ieno & Elphick, 2010). To better balance the probability of type I and II errors with small sample sizes we used a confidence level of 90% for tests and confidence intervals used herein (Peterman, 1990; Toft & Shea, 1983). All analyses were performed in R-language (R Core Team, 2013) and utilized base packages as well as the packages plsdepot (Sanchez 2012) and vegan (Oksanen et al. 2015).

3.3. Results

3.3.1 Abiotic variables

LULC classification.- In this study, ISC was found in high concentrations around the town of Foley and was more dispersed in suburban neighborhoods and larger single home lots; ISC consisted mainly of roads, buildings, and parking lots. Agriculture was widespread and mainly consisted of turf grass and pasture. Riparian forests were generally dense and composed

of overstories dominated by *Magnolia virginiana* (sweet bay), *Liriodendron tulipifera* (tulip poplar), *Quercus nigra* (water oak), and *Liquidambar styraciflua* (sweetgum). Riparian understories included *Acer rubrum* (red maple) the invasive *Triadica seberifera* (popcorn tree) and various shrubs and ground plants. Study watersheds spanned gradients of low ISC (1.5 - 11%), agriculture (15 - 53%; mainly turf grass farms and pasture), and riparian forest cover (28 - 96%; Table 2). An assessment of LULC classification accuracy based on a random selection of 25% of the 21,000 total LULC polygons created indicated 96% accuracy.

The percentage of riparian forest in a 100 m buffer (% FB) was negatively correlated with % ISC ($r = -0.60$, $p = 0.032$) and % Ag ($r = -0.53$, $p = 0.065$) in the study watersheds, but no correlation was detected between % ISC and Ag ($r = -0.07$, $p = 0.81$). VIF values for the above 3 LULC variables as single predictors in OLS framework were 2.27 (for % ISC), 2.028 (% Ag), and 3.124 (% FB), indicating some inflation about OLS coefficient estimation likely would exist due to collinearity.

Abiotic response variables.- Most of the abiotic variables examined were related to at least one of the LULC classes and the direction of these relationships followed theoretical expectations (Table 3). TN was significantly negatively related to FB, and positively to Ag; 90% CIs for these variables did not contain zero and VIP values suggested that both were important predictors of TN as both showed VIP > 1.0 (Table 3). TP, TSS, SPC, ΔSTG, and BF_w were significantly negatively related with % FB and positively with % ISC (Table 3). Q fre7 and median water temperature were positively related to % ISC (alone), whereas BFI was negatively related to % ISC. Streamwater pH was not significantly related to any of the LULC classes (CI contained zero); however, VIP scores were > 1 for % FB and ISC, indicating some importance in explaining pH values within these streams. DOC was significantly positively related to % FB

alone, according to CIs; however, the VIP score for % ISC was > 1.0 suggesting some importance in predicting DOC. Inspection of scatterplots and PLS CIs and VIP suggested that Qmax, duration of Q dur7x, BF_h, and DO were unrelated to the LULC categories examined (data not shown). Lastly, bagged (average bootstrapped) coefficients, which can be more accurate than "observed" values following model selection, were similar to coefficients derived from original data indicating general stability in the PLS component selection process .

3.3.2 Biotic variables

Benthic macroinvertebrates.- Benthic samples included ~100,000 individuals within 92 genera and 51 families. Taxa were most frequently from the dipteran family Chironomidae (51%); the top 5 of those were *Polypedilum* (14.5%), *Thienemanniella* (5.8%), *Corynoneura* (5.6%), *Rheotanytarsus* (5%) and *Tanytarsus* (3.6%) species (spp). Other prevalent groups were *Simulium* spp. (Diptera, 14%), non-insects *Lirceus* spp. (Isopoda, 11%) and *Gammarus* spp. (Amphipoda, 3%), and members of the order Trichoptera (*Chimarra* 1.1% and *Cheumatopsyche* spp. 1.6%). Additional taxa contributed < 1% each to overall benthic count data across sites and dates (Taxa are listed by site in Appendix 1). Mean invertebrate N varied greatly between sites (1,500- 11,000 m^{-2}, Table 4). Rarefied metrics also varied greatly across sites; S varied from 10 to 43, % EPT varied from < 1 to 20% and % Chironomidae from 5 to 90 % (Table 4). Seasonal differences in benthic metrics were not an immediate interest of this study; however, prior to regression analyses we used paired t-tests to determine if differences in benthic metrics existed between sample dates. Within-site differences between sample dates were not observed for any benthic metrics (*p*-value range: 0.25 - 0.61). As a result, we used cross-season average values for metrics in subsequent analyses.

Correlation based PCA was used to examine redundancy in invertebrate N, richness, *H'*, evenness, and % EPT -H/B. As expected, a large proportion of the variation in these 6 metrics was explained by the 1st two PC axes (70.93 and 10.39% respectively). Richness, *H'*, evenness, % EPT-H/B and PTV were loaded high on PC1 indicating (expected) high redundancy in diversity and tolerance metrics. Also as expected, N loaded mainly on PC2, indicating it was uncorrelated with the rarefied diversity/sensitivity metrics.

We examined density and richness in detail and considered richness as a proxy for the highly redundant group of diversity and sensitivity metrics. The model for (log-transformed) N explained 81% of its total variation; model loadings and regression coefficients indicated general associations between N and these predictors (Table 5). % FB and BOM had significant positive relationships with N and % ISC, TSS, ΔSTG, STG fre1.0, and maximum width and depth had negative associations (Table 5). The model created for S explained 74% of its variation. Model results indicated that DOC, BOM and maximum temperature were significantly negatively related to S, while maximum water depth was positively related to S (Table 5). For both models, several variables had relative large VIP values (≥ 0.9) but had CIs that were bound on one end by zero (Table 5). Bootstrapped CIs that incorporate the model selection process may be bound on one end by 0 because at least some resampled models estimated variable coefficients to be exactly zero (generally by exclusion). The interpretation of importance of these variables should be made with caution (see Chapter 1).

A multi-response PLS model described interrelationships between invertebrate N, *H'*, richness, evenness, PTV and % EPT-H/B and environmental variables (Fig. 2). Model results also indicated redundancy in rarefied diversity/tolerance metrics that were unrelated to N as determined by PCA (Fig. 2). In agreement with the univariate models, richness (*H'*, *J'*, % EPT-

H/B and PTV) was correlated with a gradient of DOC, BOM and maximum water temperature and maximum water depth, while N was negatively associated with % ISC, stage-based flashiness and frequency and maximum depth and width (Fig. 2).

Proportions of abundant families and orders showed some relationships with environmental variables (Fig. 3). The proportion of individuals belonging to the families Hydropsychidae and Baetidae were related with % ISC and associated hydrologic stressor gradient (e.g., ΔSTG; Fig. 3). % EPT –H/B, Simuliidae, Dipterans excluding Chironomids and Simuliids (Dipt – C/S), and non-insects generally followed the FB, organic matter, maximum temperature and depth gradient previously described (Fig. 3). The proportion of small bodied and/or streamlined individuals was positively related with ISC and associated hydrologic stressor gradient (Fig. 4). Rheophiles, multivoltine individuals, burrowers, sprawlers, and individuals with high fecundity and/or adult dispersal abilities were positively associated with a gradient of maximum Q and water depth and negatively related to maximum water temperature. The percentage of individuals as clingers and/or those capable of diapause were positively associated with maximum temperature, while individuals with high larval dispersal abilities were positively associated with BFI and negatively with ISC (Fig. 4).

3.4 Discussion

The primary goal of this correlative study was to determine if empirical evidence suggests that low-density development (i.e., low % ISC) influences aspects of stream hydrology, geomorphology, and physicochemistry of small coastal streams. In addition, we aimed to describe relationships between LULC, potential environmental stressors, and common benthic macroinvertebrate metrics typically negatively affected by LULC change (e.g., richness). Our results suggested that the low-density development observed in this study may be influential over

stream hydrology, as storm-event flashiness, frequency and the contribution of stormflow to total flow increased across a gradient of % ISC. In addition, there was evidence that low-density development may lead to an increase in bankfull width, TP, TSS and ionic concentrations, as well as median water temperature. Lastly, our results suggest that low-density development and associated hydrologic stressors may reduce invertebrate densities and that these developed sites host proportionally more potentially tolerant members of the EPT group (Baetidae and Hydropsychidae) than less developed sites in this study.

Schneid (Chapter 2) demonstrated that PLS models can result in relatively acceptable low levels of variance and bias, even in small sample and highly collinear situations ($r > 0.7$, VIF >> 10) and more frequently correctly indicated important predictors than selection methods based on OLS regression (e.g., stepwise). Collinearity often is a problem with LULC studies (Allan 2004). In the current study, correlations among LULC categories were minimal ($r = 0.53 – 0.6$); however, variance inflation factors indicated that some inflation in the variance (maximum VIF = 3.1) about regression coefficient estimates would be present if OLS were used. Alternative regression methods, including PLS, have generally not been considered by ecologists (Dahlgren 2010); however, we note that several recent studies have used PLS regression in analyses of the effects of LULC on stream systems (Zhang et al. 2010, Shi et al. 2013, Yan et al. 2013). PLS may be more appropriate for LULC-stream studies where predictor collinearity is inherent, sample sizes are generally low, and because LULC classes should be considered only proxy variables for underlying causal mechanisms in most cases.

ISC directly reduces infiltration and increases watershed runoff volume and velocity, thus potentially directly influencing flow regime and channel morphology; further, ISC is theoretically linked to the urban heat-island effect and direct thermal pollution from overland

runoff (Wenger et al., 2009). Our results suggest that the low-density development observed in this study (≤ 11% ISC) may increase median water temperatures, storm-event flashiness, spate frequency, bankfull width, and a decrease in the contribution of baseflow to overall stream flow (conversely indicating large volume of runoff). A recent study in nearby Apalachicola (FL) also examined sites along a gradient of low ISC (< 15%) and similarly found positive correlations between stream flashiness, median water temperatures and % ISC, although no relationship was observed for baseflow (as BFI) in that study (Nagy et al. 2012). Decreased stream baseflow is not a consistently observed trend associated with urban areas (Walsh et al. 2005, Roy et al. 2009), and urban stream discharge may be positively influenced by leaky sewers/pipes or negatively influenced by lowered water tables due to a combination of channel incision and reduced infiltration (Groffman et al. 2003).

Interestingly, bankfull depth, maximum stormflow magnitude and duration were not associated with % ISC or other LULC categories considered. The downcutting of stream channels has not been universally observed across the US, but has been associated with urbanization in coastal areas of North Carolina with similar bed composition (Hardison et al. 2009, O'Driscoll et al. 2010). It is possible that sediment input from ongoing construction activities in these watersheds were greater than export associated with storm-event scouring. Noticeable shifting of bed material and migration of large "slugs" of sand were observed in the more urban sites over the 1.5-y period of this study (BPS, personal observation). Stormflow magnitude and duration have also been generally associated with urbanization across the US, (Poff et al. 2006, Brown et al. 2009) and Nagy et al. (2010) found a significant relationship existed between maximum storm magnitude and low levels of % ISC in coastal Florida.

Urban areas and agricultural lands can also contribute to nutrients, sediments and dissolved solids (SPC) in streams through the use of fertilizers and the disturbance of ground cover (Allan 2004). In our study, agriculture was positively related to TN concentrations and ISC positively related to TP, TSS, and SPC. Riparian forests are thought to act as physical filters for sediment, and nutrients as well as being the primary source of organic matter (DOC) in small streams (Naiman & Décamps, 1997). TN, TP, TSS, SPC, pH and bankfull width were negatively related with the %FB of in the 100 m buffer; therefore, riparian forest may act to reduce the overland inputs of nutrients and solids from agricultural and urban LULC in these watersheds/streams. These results support the beneficial view of riparian buffers and their use as best management practice tools to maintain good water-quality in low-gradient coastal areas. Riparian buffers have been shown to drastically reduce sediments from agricultural lands, and are generally considered to be sinks for sediment, sediment-bound P and soluble nutrients (Gregory et al. 1991, Naiman and Décamps 1997). In a meta-analysis, Mayer et al. (2007) showed that wide riparian buffers (> 50 m) were more consistent in nitrogen removal that narrower buffers; however, it has also been shown that riparian zones may become less effective as nitrogen (NO_3^-) sinks as they become more urbanized (Groffman et al. 2002).

ISC was not significantly related to macroinvertebrate density or sensitivity metrics; however, density was negatively associated with %ISC and hydrologic flashiness and frequency. It is generally accepted that LULC indirectly affects biota through intermediate causal factors (e.g., hydrology) (Burcher, Valett & Benfield, 2007); therefore, it may be expected that LULC categories would exhibit weaker associations than potentially direct sources of influence. Interestingly, density was negatively related to TSS, maximum flood width and depth and positively related to % FB and benthic organic matter as well; taken together, these results may

indicate a scouring influence from increased hydrologic disturbance and/or that of resource availability/abundance.

Benthic richness (as well as diversity and sensitivity metrics) were negatively related with maximum stream water temperature, but positively associated with maximum water depth. Chadwick et al. (2011) found a weak but positive relationship between high urban LULC and taxonomic richness of benthic invertebrates of north Florida streams and concluded this observation may be because of greater hydrologic permanence in urban streams (Chadwick et al. 2011). Hydrologic permanence in urban streams might be beneficial for some organisms, as higher flows may increase available habitat, alleviate water chemistry problems through dilution, reduce water temperatures, and/or increase dissolved oxygen concentrations and lead to an overall higher richness/diversity relative to non-urban streams that experience intermittency in summer months (Walsh et al. 2005, Roy et al. 2009). Hydrologic permanence might be a factor in our sites as well, as observed maximum water temperatures (15 min data) may be indicative of shallow stagnant water and potential stream drying during summer low-flow (personal observation).

Subtle compositional trends were observed in more developed streams, as the families Hydropsychidae and Baetidae positively associated with %ISC and hydrologic flashiness/frequency stressor gradient. While variation in tolerance within these two families exists, these groups have been generally assumed to be more tolerant than other EPT taxa (but see Chang et al. 2014). The traits small and streamlined body shape was also associated with this same ISC/hydrology gradient and it was noted that the more developed streams in this study were quite abundant in small bodied, early instar baetid and hydropsychid individuals, as well as small (streamlined) heptageniid mayflies (BPS, personal observation). We also observed that the

traits rheophily, along with high fecundity and high adult dispersal were generally associated with maximum flood water depth, maximum discharge and TSS. These observations correspond to predictions from the habitat templet concept applied to river systems (Townsend and Hildrew 1994), in that more hydrologically disturbed sites should be host to a greater abundance of small, streamlined individuals that prefer fast flowing water, and/or those with a disproportionate ability to recolonize (high fecundity/dispersal) following extreme disturbance events. In addition, the percentage of individuals capable of diapause (e.g., desiccation-resistant eggs) was interestingly positively associated with maximum temperature and sites that likely experienced extreme summer drawdown if not complete drying in some portions of study reaches (BPS, personal observation). Most of these sites were dominated by Simulium spp., some of which produce desiccation-resistant eggs and have been associated with temporary habitats (Bogan et al. 2013). Maximum temperature was negatively associated with richness, thus supporting the idea that extreme temperatures and stream drying might be highly influential over benthic diversity in these streams.

The current study demonstrated an influence of ISC $\leq$ 11% on hydrology and physicochemistry, but only subtle trends with benthic invertebrates and developed land. In contrast, ISC cover $\leq$ 10% has been shown to influence stream hydrology and physicochemistry in some coastal areas (Schiff and Benoit 2007; Lussier et al. 2008; Cunningham et al. 2009; Nagy et al. 2012) and negatively impact sensitive taxa in at least a few studies (Walsh et al. 2007, King et al. 2011). ISC that is directly connected to streams through drainage infrastructure is recognized as being a likely more influential to stream systems than whole-watershed ISC (Wenger et al. 2009). We note that in this study whole-watershed % ISC and % ISC within 100 m of the stream channel were highly correlated ($r = 0.96$, $p << 0.001$). Because ISC located

immediately adjacent to these streams is proportional to that at the whole watershed, we would expect that the general trends observed in this study would be similar those that we would find if we did have information on ISC connectivity.

3.5 Conclusions

Virtually all of the southeastern coastal plains has been altered by human activity and has been historically dominated agriculture and more recently converted to urbanized and suburbanized landscapes (Nagy et al. 2011). The effects of urbanization are often reduced/confounded when agricultural lands are converted (Wenger et al. 2009); however, our study suggests that ISC $\leq$ 11% likely influences nutrient, sediment and ionic concentrations in these streams as well as channel width, hydrologic flashiness, storm event frequency, and baseflow contributions to total stream flow.

Burcher and Benfield (2006) found no relationships between suburban land cover and macroinvertebrate metrics (e.g., taxa richness), the authors concluded that it is likely these sites did not exceed the lower ISC threshold to induce a measureable effect, but noted slight compositional differences using multivariate ordination (Burcher and Benfield 2006). In the current study, the data suggest that ISC and associated stressors (e.g., hydrologic flashiness) may influence invertebrate densities; however, rarefied benthic diversity and sensitivity metrics were not detectably associated with ISC. We conclude that ISC levels in this study were not large enough to lead to a detectable effect of ISC on these diversity and sensitivity/tolerance metrics and that underlying gradients of stream flow permanence, temperature extremes, oxygen levels, organic matter, and potential stream drying (supported by % diapauses trait) may play an important role in the observed variation. Interestingly, more subtle compositional differences in purportedly tolerant families (Baetidae and Hydropsychidae) were observed along a gradient of

ISC and related hydro-geomorphic stressors, as well as relationships in agreement with theoretical considerations from the habitat templet concept that likely confer resistance/resilience to hydrologic disturbance (Southwood 1977; Townsend and Hildrew 1994).

3.6 References

Alig, R.J., Kline, J.D., and M. Lichtenstein. 2004. Urbanization on the US landscape: looking ahead in the 21st century. Landscape and Urban Planning 69: 219-234.

Allan, J.D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annual Review of Ecology, Evolution, and Systematics 35:257-284.

Arcement Jr, G.J. and V. Schneider. 1989. Guide for Selecting Manning's Roughness Coefficients for Natural Channels and Flood Plains. US Geological Survey Water-Supply Paper 2339. U.S. Government Printing Office, Washington, DC, USA. p.1-38.

Baker, D.B., Richards, R.P., Loftus, T.T., and J.W. Kramer. 2007. A new flashiness index: Characteristics and applications to midwestern streams. Journal of the American Water Resources Association 40:503-522.

Benke, A.C., Henry III, R.L., Gillespie, D.M., and R.J. Hunter. 1985. Importance of snag habitat for animal production in southeastern streams. Fisheries 10:8-13.

Bogan, M. T., K. S. Boersma, and D. A. Lytle. 2013. Flow intermittency alters longitudinal patterns of invertebrate diversity and assemblage composition in an arid-land stream network. Freshwater Biology 58:1016-1028.

Bonada, N., N. Prat, V. H. Resh, and B. Statzner. 2006. Developments in aquatic insect biomonitoring: a comparative analysis of recent approaches. Annual Review of Entomology 51:495- 523.

Brabec, E. 2002. Impervious surfaces and water quality: a review of current literature and its implications for watershed planning. Journal of Planning Literature 16:499-514.

Brown, L., Cuffney, T., Coles, J., Fitzpatrick, F., McMahon, G., Steuer, J., Bell, A., and J. May. 2009. Urban streams across the USA: lessons learned from studies in 9 metropolitan areas. Journal of the North American Benthological Society 28:1051-1069.

Buckland, S.T., Burnham, K.P., and N.H. Augustin. 1997. Model selection: an integral part of inference. Biometrics, 53:603-618.

Bühlmann, P. and B. Yu. 2002. Analyzing bagging. Annals of Statistics 30:927-961.

Burcher, C., Valett, H., and E. Benfield. 2007. The land-cover cascade: relationships coupling land and water. Ecology 88:228-242.

Burcher, C.L., and E. Benfield. 2006. Physical and biological responses of streams to suburbanization of historically agricultural watersheds. Journal of the North American Benthological Society 25: 356-369.

Chadwick, M.A., Thiele, J.E., Huryn, A.D., Benke, A.C., and D.R. Dobberfuhl. 2011. Effects of urbanization on macroinvertebrates in tributaries of the St. Johns River, Florida, USA. Urban Ecosystems 15:1-19.

Chang, F.-H., J. E. Lawrence, B. Rios-Touma, and V. H. Resh. 2014. Tolerance values of benthic macroinvertebrates for stream biomonitoring: assessment of assumptions underlying scoring systems worldwide. Environmental Monitoring and Assessment 186:2135-2149.

Clausen, B., and B. Biggs. 2000. Flow variables for ecological studies in temperate streams: groupings based on covariance. Journal of Hydrology 237:184-197.

Copeland, R.R., Biedenharn, D.S. and J.C. Fischenhein. 2000. Channel forming discharge. US Army Corps Eng. Rep., ERDC/CHL CHETN-VIII-5, 1-11.

Cunningham, M.A., O'Reilly, C.M., Menking, K.M., Gillikin, D.P., Smith, K.C., Foley, C.M., Belli, S.L., Pregnall, A.M., Schlessman, M.A., and P. Batur. 2009. The suburban stream syndrome: evaluating land use and stream impairments in the suburbs. Physical geography 30:269-284.

Dahlgren, J.P. 2010. Alternative regression methods are not considered in Murtaugh (2009) or by ecologists in general. Ecology Letters 13:E7-E9.

Dolédec, S., J. M. Olivier, and B. Statzner. 2000. Accurate description of the abundance of taxa and their biological traits in stream invertebrate communities: effects of taxonomic and spatial resolution. Archiv für Hydrobiologie 148:25-43.

Efron, B., 2014. Estimation and accuracy after model selection. Journal of the American Statistical Association 109:991-1007.

Feeley, J.D. 1992. Medium-low-gradient streams of the Gulf Coastal Plain. In: Hackney, C.T., Adams, S.M., Martin, W.H. (Eds.), Biodiversity of the southeastern United States--Aquatic Communities. Wiley and Sons, Inc, New York, pp. 233-269.

Feminella, J.W. 1996. Comparison of benthic macroinvertebrate assemblages in small streams along a gradient of flow permanence. Journal of the North American Benthological Society 4:651-669.

González, I., Lê Cao, K-A., Davis, M.J., and S. Déjean. 2012. Visualising associations between paired 'omics' data sets. BioData mining 5:1-23.

Gotelli, N.J., and R.K Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 4:379-391.

Hegyi, G., L.Z. Garamszegi. 2011. Using information theory as a substitute for stepwise regression in ecology and behavior. Behavioral Ecology and Sociobiology 65:69-76.

Garthwaite, P. H. 1994. An interpretation of partial least squares. Journal of the American Statistical Association 89:122-127.

Graham, M.H. 2003. Confronting multicollinearity in ecological multiple regression. Ecology 84:2809-2815.

Gregory, S. V., F. J. Swanson, W. A. McKee, and K. W. Cummins. 1991. An ecosystem perspective of riparian zones. BioScience 41:540-551.

Groffman, P. M., N. J. Boulware, W. C. Zipperer, R. V. Pouyat, L. E. Band, and M. F. Colosimo. 2002. Soil nitrogen cycle processes in urban riparian zones. Environmental science & technology 36:4547-4552.

Groffman, P. M., D. J. Bain, L. E. Band, K. T. Belt, G. S. Brush, J. M. Grove, R. V. Pouyat, I. C. Yesilonis, and W. C. Zipperer. 2003. Down by the riverside: urban riparian ecology. Frontiers in Ecology and the Environment 1:315-321.

Hardison, E.C., O'Driscoll, M.A., DeLoatch, J.P., Howard, R.J., and M.M. Brinson. 2009. Urban Land Use, Channel Incision, and Water Table Decline Along Coastal Plain Streams, North Carolina. Journal of the Journal of the American Water Resources Association 45:1032-1046.

Kenney, M. A., A. E. Sutton-Grier, R. F. Smith, and S. E. Gresens. 2009. Benthic macroinvertebrates as indicators of water quality: the intersection of science and policy. Terrestrial Arthropod Reviews 2:99-128.

King, R.S., Baker, M.E., Whigham, D.F., Weller, D.E., Jordan, T.E., Kazyak, P.F., Hurd, M.K., 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15, 137-153.

King, R. S., M. E. Baker, P. F. Kazyak, and D. E. Weller. 2011. How novel is too novel? Stream community thresholds at exceptionally low levels of catchment urbanization. Ecological Applications 21:1659-1678.

Konrad, C.P., Booth, D.B., Brown, L., Gray, R., Hughes, R., Meador, M. 2005. Hydrologic changes in urban streams and their ecological significance. American Fisheries Society Symposium 47: 157-177.

Lamberti, G. A. and S. V. Gregory. 2006. CPOM transport, retention, and measurement. Pages 273-289 in F. R. Hauer and G. A. Lamberti, editors.Methods in stream ecology. Academic Press.

Lammert, M. and J.D. Allan.1999. Assessing biotic integrity of streams: effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental management 23:257-270.

Lurry, D.L. and C.M. Kolbe. 2000. Interagency field manual for the collection of water-quality data. U.S. Geological Survey. Open-File Report 00-213.

Li, B., Morris, J., Martin, E.B. 2002. Model selection for partial least squares regression. Chemometrics and Intelligent Laboratory Systems 64:79-89.

Lussier, S.M., da Silva, S.N., Charpentier, M., Heltshe, J.F., Cormier, S.M., Klemm, D.J., Chintala, M., Jayaraman, S., 2008. The influence of suburban land use on habitat and biotic integrity of coastal Rhode Island streams. Environmental Monitoring and Assessment 139:119-136.

Mac Nally, R., 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between - and reconciliation of – ‘predictive’ and ‘explanatory’ models. Biodiversity & Conservation 9:655-671.

Mac Nally, R., 2002. Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables. Biodiversity & Conservation 11:1397-1401.

Mac Nally, R., Walsh, C.J., 2004. Hierarchical partitioning public-domain software. Biodiversity and conservation 13:659-660.

Mayer, P. M., S. K. Reynolds, M. D. McCutchen, and T. J. Canfield. 2007. Meta-analysis of nitrogen removal in riparian buffers. Journal of Environmental Quality 36:1172-1180.

McMahon, G., Bales, J.D., Coles, J.F., Giddings, E.M.P., Zappia, H., 2003. Use of stage data to characterize hydrologic conditions in an urbanizing environment. Journal of the American Water Resources Association 39:1529-1546.

Mehmood, T., Liland, K.H., Snipen, L., Sæbø, S. 2012. A review of variable selection methods in partial least squares regression. Chemometrics and Intelligent Laboratory Systems 118:62-69.

Merritt, R.W., Cummins, K.W., Berg, M.B. 2008. An introduction to the aquatic insects of North America. Kendall Hunt.

Metcalf, C., 2009. Alabama Riparian Reference Reach and Regional Curve Study. US Fish and Wildlife Service unpublished report prepared for the Alabama Department of Environmental Management, Panama City, Florida, 34 pp. http://www.fws.gov/panamacity/, accessed April.

Montgomery, D. C., E. A. Peck, G. G. Vining, and J. Vining. 2001. Introduction to linear regression analysis. Wiley New York.

Morrice, J.A., Danz, N.P., Regal, R.R., Kelly, J.R., Niemi, G.J., Reavie, E.D., Hollenhorst, T., Axler, R.P., Trebitz, A.S., Cotter, A.M. 2008. Human influences on water quality in Great Lakes coastal wetlands. Environmental management 41:347-357.

Nagy, R.C., Lockaby, B.G., Helms, B., Kalin, L., Stoeckel, D. 2011. Water Resources and Land Use and Cover in a Humid Region: The Southeastern United States. J. Environ. Qual 40:867-878.

Nagy, R.C., Lockaby, B.G., Kalin, L., Anderson, C. 2012. Effects of urbanization on stream hydrology and water quality: the Florida Gulf Coast. Hydrological Processes 26:2019–2030.

Naiman, R.J., Décamps, H. 1997. The ecology of interfaces: riparian zones. Annual Review of Ecology and Systematics 28:621-658.

Nathan, R., McMahon, T. 1990. Evaluation of automated techniques for base flow and recession analyses. Water Resources Research 26:1465-1473.

Norris, R.H., Thoms, M.C. 1999. What is river health? Freshwater Biology 41, 197-209.

O'Driscoll, M., Clinton, S., Jefferson, A., Manda, A., McMillan, S., 2010. Urbanization Effects on Watershed Hydrology and In-Stream Processes in the Southern United States. Water 2:605-648.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., M., Stevens, M.H., Wagner, H. 2015. vegan: Community Ecology Package. R package version 2.2-1. http://CRAN.R-project.org/package=vegan

Paul, M.J., Meyer, J.L. 2001. Streams in the urban landscape. Annual Reviewof Ecolological Systems 32:333-365.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47:2-15.

Phillips, C., Scatena, F. 2010. Flashiness Indices for Urban and Rural Streams in Puerto Rico. American Water Resource Association. American Water Resource Association Summer Specialty Conference, Sam Juan, Puerto Rico.

Pickup, G., Warner, R. 1976. Effects of hydrologic regime on magnitude and frequency of dominant discharge. Journal of Hydrology 29:51-75.

Poff, N., Bledsoe, B., Cuhaciyan, C. 2006. Hydrologic variation with land use across the contiguous United States: geomorphic and ecological consequences for stream ecosystems. Geomorphology 79:264-285.

R Core Team, 2013. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Raghunath, H.M., 2007. Hydrology: principles, analysis, design. New Age International.

Rice, E.W., Association, A.P.H. 2012. Standard methods for the examination of water and wastewater. American Public Health Association Washington, DC.

Rose, S., Peters, N.E. 2001. Effects of urbanization on streamflow in the Atlanta area (Georgia, USA): a comparative hydrological approach. Hydrological Processes 15:1441-1457.

Rosipal, R., Krämer, N. 2006. Overview and recent advances in partial least squares. Subspace, Latent Structure and Feature Selection. Springer, pp. 34-51.

Roy, A. H., A. L. Dybas, K. M. Fritz, and H. R. Lubbers. 2009. Urbanization affects the extent and hydrologic permanence of headwater streams in a midwestern US metropolitan area. Journal of the North American Benthological Society 28:911-928.

Sanchez, G. 2012. plsdepot: Partial Least Squares (PLS) Data Analysis Methods. R package version 0.1.17. http://CRAN.R-project.org/package=plsdepot

Schiff, R., and G. Benoit. 2007. Effects of Impervious Cover at Multiple Spatial Scales on Coastal Watershed Streams. Journal of the Journal of the American Water Resources Association 43:712-730.

Schoonover, J.E., Lockaby, B.G. and B. S. Helms. 2006. Impacts of land cover on stream hydrology in the West Georgia Piedmont, USA. Journal of Environmental Quality 35:2123-2131.

Sefick, S.A., Kalin, L., Kosniki, E., Schneid, B.P., Jarrell, M.S., Anderson, C.J., Paller, M.H. Feminella, J.W. 2015. Empirical Estimation of Stream Discharge Using Channel Geometry in Low-Gradient, Sand-Bed Streams of the Southeastern Plains. Journal of the Journal of the American Water Resources Association 1-12. DOI: 10.1111/jawr.12278

Shi, Z., L. Ai, X. Li, X. Huang, G. Wu, and W. Liao. 2013. Partial least-squares regression for linking land-cover patterns to soil erosion and sediment yield in watersheds. Journal of Hydrology 498:165-176.

Shuster, W., Bonta, J., Thurston, H., Warnemuende, E., Smith, D. 2005. Impacts of impervious surface on watershed hydrology: A review. Urban Water Journal 2:263-275.

Smock, L.A., Gilinsky, E. 1992. Coastal Plain blackwater streams. In: Hackney, C.T., Adams, S.M., Martin, W.H. (Eds.), Biodiversity of the southeastern United States--Aquatic Communities. Wiley and Sons, Inc., New York.

Southwood, T. 1977. Habitat, the templet for ecological strategies? The Journal of Animal Ecology 46:337-365.

Steinman, A. D., G. A. Lamberti, and P. R. Leavitt. 2006. Biomass and pigments of benthic algae. Pages 357-380 in F. R. Hauer and G. A. Lamberti, editors. Methods in stream ecology.Academic Press.

Sweet, W., Geratz, J. 2003. Bankfull hydraulic geometry relationships and recurrence intervals for north carolina's coastal plain. Journal of the Journal of the American Water Resources Association 39:861-871.

Thomas, H., Nisbet, T. 2007. An assessment of the impact of floodplain woodland on flood flows. Water and Environment Journal 21:114-126.

Toft, C.A., Shea, P.J. 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. American Naturalist 122:618-625.

Townsend, C.R., Hildrew, A.G. 1994. Species traits in relation to a habitat templet for river systems. Freshwater Biology 31:265-275.

Utz, R.M., Hilderbrand, R.H. 2011. Interregional variation in urbanization-induced geomorphic change and macroinvertebrate habitat colonization in headwater streams. Journal of the North American Benthological Society 30:25-37.

U.S. Environmental Protection Agency. 2013. National rivers and streams assessment 2008-2009: technical report (DRAFT). U.S. Environmental Protection Agency, Office of Wetlands, Oceans and Watersheds Office of Research and Development. Washington, DC, p 127.http://water.epa.gov/type/rsl/monitoring/riverssurvey/upload/NRSA0809_ Technical Report_130325_Web.pdf. Accessed 28 April 2015.

Varanka, S., Luoto, M., 2012. Environmental determinants of water quality in boreal rivers based on partitioning methods. River Research and Applications 28:1034-1046.

Vieira, N.K.M., Poff, N.L., Carlisle, D.M., Moulton, S.R., Koski, M.L., Kondratieff, B.C. 2006. A database of lotic invertebrate traits for North America. US Geological Survey Data Series 187:1–15.

Walker, S.C., Poos, M.S., Jackson, D.A. 2008. Functional rarefaction: estimating functional diversity from field data. Oikos 117:286-296.

Wallace JB, Grumbaugh JW, Whiles MR. 1993. Influences of coarse woody debris on stream habitats, invertebrate diversity. In Biodiversity and Coarse Woody Debris in Southern Forests. Proceedings of the Workshop on Coarse Woody Debris in Southern Forests: Effects on Biodiversity. Athens, GA. October 18–20, 1993. USDA Forest Service. Southern Research Station General Technical Report SE-94:119–129.

Walsh, C.R., A., Feminella, J.W., Cottingham, P., Groffman, P., Morgan, R. 2005. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24:706-723.

Walsh, C. J., K. A. Waller, J. Gehling, and R. M. Nally. 2007. Riverine invertebrate assemblages are degraded more by catchment urbanisation than by riparian deforestation. Freshwater Biology 52:574-587.

Wear, D.N., Greis, J.G. 2002. Southern forest resource assessment: summary of findings. Journal of Forestry 100:6-14.

Wenger, S. J., A. H. Roy, C. R. Jackson, E. S. Bernhardt, T. L. Carter, S. Filoso, C. A. Gibson, W. C. Hession, S. S. Kaushal, E. Martí, J. L. Meyer, M. A. Palmer, M. J. Paul, A. H. Purcell, A. Ramirez, A. D. Rosemond, K. A. Schofield, E. B. Sudduth, and C. J. Walsh. Twenty-six key research questions in urban stream ecology: an assessment of the state of the science. Journal of the North American Benthological Society 28:1080-1098.

White, E.M., Morzillo, A.T., Alig, R.J. 2008. Past and projected rural land conversion in the US at state, regional, and national levels. Landscape and Urban Planning 89:37-48.

Wold, S., Ruhe, A., Wold, H., Dunn, I., W.J. 1984. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM Journal on Scientific and Statistical Computing 5:735-743.

Wold, S., Sjöström, M., Eriksson, L. 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems 58:109-130.

Xian, G., Homer, C., Bunde, B., Danielson, P., Dewitz, J., Fry, J., Pu, R. 2012. Quantifying urban land cover change between 2001 and 2006 in the Gulf of Mexico region. Geocarto International 1:1-19.

Yan, B., N. Fang, P. Zhang, and Z. Shi. 2013. Impacts of land use change on watershed streamflow and sediment yield: an assessment using hydrologic modelling and partial least squares regression. Journal of Hydrology 484:26-37.

Zhang, Y., D. Dudgeon, D. Cheng, W. Thoe, L. Fok, Z. Wang, and J. H. Lee. 2010. Impacts of land use and water quality on macroinvertebrate communities in the Pearl River drainage basin, China. Hydrobiologia 652:71-88.

Zuur, A.F., Ieno, E.N., Elphick, C.S. 2010. A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution 1:3-14.

Table 1. Abbreviations and descriptions of benthic metrics used in this study.

Measure	Metric	Definition
Abundance	Numerical density (N)	No. individuals per m^2
Diversity	Richness (S)	No. of taxa
	Shannon's diversity (*H'*)	= $-\sum_{i=1}^{S} p_i \ln p_i$; where p_i is proportion of ith taxa
	Pielou's evenness (*J'*)	= H'/lnS
	Pollution tolerance (PTV)	= $\frac{1}{N}\sum_{i=1}^{S} p_i \, PTV_i$
Composition	% EPT	% as Ephemeroptera, Plecoptera, Trichoptera
	% EPT-H/B	% as EPT minus Hydropsychidae and Baetidae
	% Hydropsychidae	% as family Hydropsychidae
	% Baetidae	% as family Baetidae
	% Chironomidae	% as family Chironomidae
	% Simuliidae	% as family Simuliidae
	% Diptera–C/S	% as order Diptera minus Chironomidae and Simuliidae
	% Colepotera	% as order Colepotera
	% Odonata	% as order Odonata
	% Non-insect	% as non-insect taxa
Habits	Burrower	Burrows in fine sediments
	Sprawler	Lives on plants or surface of fine sediments
	Clinger	Clings to stable substrates
Traits	Multivoltine	> 1 generation per year
	Small	Body length < 2mm
	Streamlined	Body shape resists drag/flow forces
	Rheophile	Prefers fast flowing water
	High adult dispersal	Flying up to 100 km
	High larval dispersal	Crawling up to 100 m
	High fecundity	> 10,000 eggs
	Diapause	Ability to diapause (dormancy)

Table 2. Basic watershed attributes including watershed impervious surface cover (ISC), agriculture (Ag), riparian forest buffer (FB), watershed area, and stream order (SO). Median values and standard deviations (in parentheses) for are given for observed stream discharge, salinity (SAL), pH, specific conductivity (SPC), dissolved oxygen (DO), total nitrogen (TN) and total phosphorus (TP).

Site	ISC (%)	Ag (%)	FB (%)	Area (ha)	SO	Discharge (m^3 s^{-1})	SAL (ppt)	pH (unitless)	SPC (µS cm^{-1})	DO (mg/L)	TN (mg/L)	TP (mg/L)
BON12	8.7	49.4	33.8	2481	2	0.58 (1.76)	0.04(0.01)	5.62 (0.20)	75.0 (8.9)	7.32 (0.9)	2.37 (0.6)	0.02 (0.09)
FPR29	5.1	45	75.1	253	1	0.05 (0.11)	0.03(0.01)	4.86 (0.34)	58.0 (11.0)	6.82 (2.2)	0.66 (0.2)	0.01 (0.01)
FPR30	2.2	15.3	95.5	222	1	0.04 (0.10)	0.02(0.00)	4.45 (0.35)	49.0 (7.8)	5.07 (1.7)	0.4 (0.2)	0.01 (0.01)
GUM13	5.7	28.7	53.6	655	1	0.06 (0.73)	0.03(0.01)	6.11 (0.36)	69.0 (14.1)	8.34 (2.0)	1.05 (0.5)	0.02 (0.01)
HMK33	3.7	22.2	74.2	621	1	0.04 (0.32)	0.02(0.01)	4.90 (0.29)	49.0 (11.3)	7.99 (2.0)	0.58 (0.2)	0.01 (0.00)
MAG65	9.5	38.5	27.8	2306	2	0.54 (1.87)	0.03(0.01)	5.48 (0.21)	66.0 (8.5)	7.45 (0.6)	2.45 (0.6)	0.01 (0.04)
MFL08	4.4	45.1	49.6	854	2	0.18 (0.65)	0.04(0.01)	5.78 (0.28)	74.0 (11.6)	7.42 (1.7)	2.57 (0.8)	0.03 (0.03)
MFL83	2.6	52.9	50.7	112	1	0.05 (0.06)	0.04(0.00)	5.63 (0.14)	86.0 (7.5)	6.86 (2.0)	2.19 (0.2)	0.01 (0.02)
PLM20	1.9	36.8	56.5	466	1	0.11 (0.11)	0.03(0.00)	5.5 (0.17)	63.0 (7.1)	6.89 (1.4)	1.77 (0.3)	0.01 (0.01)
SAN06	4.3	36	64.6	1493	3	0.61 (1.11)	0.02(0.01)	5.68 (0.33)	54.0 (6.2)	8.32 (0.9)	1.58 (0.4)	0.01 (0.02)
SAN7E	6.8	31.4	48.3	202	1	0.04 (0.35)	0.03(0.00)	5.52 (0.19)	61.0 (8.5)	8.03 (1.0)	2.08 (0.4)	0.02 (0.06)
SAN7W	1.5	43.1	55.7	699	1	0.06 (0.42)	0.03(0.01)	5.33 (0.23)	58.0 (10.9)	8.96 (0.9)	2.88 (0.9)	0.02 (0.06)
WLF01	10.9	22	50.2	845	2	0.17 (1.00)	0.04(0.01)	6.10 (0.21)	88.5 (16.7)	6.44 (2.1)	1.92 (1.2)	0.58 (0.29)

Table 3. Summary table for PLS model coefficient estimates for each LULC category ($\hat{b}_{LULC}$) and model intercepts ($\hat{b}_0$). Bootstrap aggregated (bagged) average coefficients are in parenthesis, and bootstrap 90% confidence intervals (CI) are provided below each coefficient estimate. R^2 ($cor(y,\hat{y})^2$) is provided in the far right column for each model. Boldfaced values indicate CI (for slopes only) does not contain 0. † = PLS-VIP ≥ 1.0. Coefficients were un-standardized, and some multiplied by 100 (*100) to improve readability. Transformations are noted in the left-hand column.

Response		$\hat{\boldsymbol{b}}_0$	$\hat{\boldsymbol{b}}_{FB}$	$\hat{\boldsymbol{b}}_{Ag}$	$\hat{\boldsymbol{b}}_{ISC}$	R^2
TN	$\hat{b}$	2.80 (2.71)	**-0.031 (-0.030)**	**0.017 (0.020)**	0.013 (-0.005)	0.674
	CI	0.92, 4.71	**-0.062, -0.009**†	**0.001, 0.049**†	-0.173, 0.108	---
log(TP)	$\hat{b}$	-3.47 (-3.74)	**-0.017 (-0.014)**	0.004 (0.004)	**0.030 (0.035)**	0.355
	CI	-4.74, -2.56	**-0.031, -0.003**†	-0.006, 0.017	**0.001, 0.111**†	---
log(TSS)	$\hat{b}$	2.13 (1.95)	**-0.018 (-0.016)**	-0.001 (-0.003)	**0.042 (0.054)**	0.398
	CI	0.81, 3.40	**-0.034, -0.002**†	-0.026, 0.014	**0.003, 0.141**†	---
DOC	$\hat{b}$	4.195 (6.43)	**0.012 (0.078)**	-0.029 (0.036)	-0.054 (-0.049)	0.481
	CI	-0.637, 10.96	**0.015, 0.155**†	-0.303, 0.473	-0.119, 0.017†	---
SPC	$\hat{b}$	81.07 (73.56)	**-0.458 (-0.365)**	0.175 (0.199)	**0.771 (0.975)**	0.424
	CI	46.82, 107.31	**-0.803, -0.096**†	-0.096, 0.632	**0.014, 2.503**†	---
Temp.	$\hat{b}$	20.26 (20.04)	-0.011 (-0.007)	-0.000 (-0.003)	**0.048 (0.060)**	0.443
	CI	18.87, 21.43	-0.024, 0.006	-0.024, 0.011	**0.006, 0.167**†	---
sqrt(pH)	$\hat{b}$	239.7 (241.5)	-0.151 (-0.214)	0.0176 (0.010)	0.303 (0.567)	0.503
*100	CI	224.5, 269.1	-0.705, 0.017†	-0.245, 0.270	-0.204, 2.102†	---
ΔSTG	$\hat{b}$	3.229 (3.379)	**-0.022 (-0.028)**	-0.003 (-0.009)	**0.059 (0.128)**	0.607
*100	CI	1.604, 6.419	**-0.071, -0.001**†	-0.046, 0.020	**0.008, 0.314**†	---
Q freq7	$\hat{b}$	18.02 (11.62)	0.038 (0.103)	-0.162 (-0.104)	**1.419 (1.534)**	0.300
	CI	-34.565, 39.078	-0.249, 0.579	-0.519, 0.318	**0.175, 3.975**	---
BFI	$\hat{b}$	69.73 (73.73)	0.162 (0.094)	0.105 (0.131)	**-0.921 (-1.164)**	0.268
*100	CI	53.01, 91.68	-0.130, 0.370	-0.113, 0.509	**-3.015, -0.167**†	---
BF_w	$\hat{b}$	13.574 (10.782)	**-0.176 (-0.121)**	0.011 (-0.001)	**0.395 (0.445)**	0.575
	CI	3.234, 20.671	**-0.246, -0.029**†	-0.137, 0.107	**0.108, 0.989**†	---

Table 4. Mean benthic density and diversity/sensitivity metric values for the 13 study streams. Abbreviations are provided in Table 1.

Site	N	S	*H'*	*J'*	% EPT	% Chironomidae	PTV
BON12	1,622	35	2.79	0.78	11.63	64.55	4.87
FPR29	11,159	22	1.99	0.64	3.33	16.82	4.88
FPR30	7,116	10	1.03	0.43	0.14	5.40	7.23
GUM13	2,247	29	2.48	0.74	16.25	66.59	4.90
HMK33	9,785	20	1.58	0.53	1.43	14.35	5.80
MAG65	1,556	31	2.52	0.73	13.23	63.78	5.18
MFL08	6,168	24	1.47	0.46	0.91	89.69	5.24
MFL83	9,138	28	2.38	0.71	9.67	57.90	5.13
PLM20	6,726	37	2.84	0.78	10.38	44.52	4.59
SAN06	2,286	38	2.88	0.79	11.90	53.55	4.69
SAN7E	1,501	30	2.03	0.60	2.86	65.59	5.51
SAN7W	2,486	43	2.91	0.77	9.10	56.60	4.65
WLF01	2,293	28	2.53	0.76	19.44	64.19	4.93

Table 5. Summary results from PLS regression models explaining invertebrate density (N) and rarefied genus richness. Variable importance in projection scores (VIP), PLS model loadings, standardized regression coefficients ($\hat{\beta}$), and 90% confidence intervals (CIs) are provided for each model. Only variables included in final PLS models are shown, and each had initial model VIP > 0.7. Boldface rows indicate 90% CI for regression coefficients did not contain zero, and "---" indicates variables not retained in final model. R^2 for the model describing density was 0.81 and R^2 was 0.74 for the model describing richness.

	Log_{10}(N)				Richness			
Variables	VIP	Loadings	$\hat{\beta}$ (bagged $\hat{\beta}$)	$\hat{\beta}$ CIs	VIP	Loadings	$\hat{\beta}$ (bagged $\hat{\beta}$)	$\hat{\beta}$ CIs
FB (%)	**1.05**	**+0.264**	**0.087 (0.075)**	**0.021, 0.106**	1.03	-0.284	-0.100 (-0.070)	-0.116, 0
Ag (%)	---	---	---	---	0.86	+0.237	0.083 (0.069)	0, 0.135
ISC (%)	**1.01**	**-0.252**	**-0.083 (-0.076)**	**-0.111, -0.034**	---	---	---	---
DOC (mg/L)	0.99	+0.248	0.082 (0.082)	0, 0.166	**1.37**	**-0.379**	**-0.133 (-0.134)**	**-0.227, -0.083**
DO (mg/L)	0.71	-0.178	-0.058 (-0.054)	-0.115, 0	1.09	+0.303	0.106 (0.093)	0, 0.196
log(BOM) (%)	**1.34**	**+0.334**	**0.110 (0.108)**	**0.076, 0.159**	**1.19**	**-0.332**	**-0.116 (-0.095)**	**-0.134, -0.0411**
log(TSS)	**1.33**	**-0.333**	**-0.110 (-0.115)**	**-0.187, -0.073**	0.77	+0.213	0.075 (0.063)	0, 0.125
sqrt(pH)	0.89	-0.222	-0.073 (-0.060)	-0.093, 0	0.96	+0.266	0.093 (0.064)	0, 0.115
LWD	---	---	---	---	0.86	-0.238	-0.083 (-0.074)	-0.154, 0
Temp. max	0.54	+0.134	0.044 (0.028)	-0.025, 0.078	**1.32**	**-0.366**	**-0.128 (-0.119)**	**-0.184, -0.072**
ΔSTG	**1.00**	**-0.251**	**-0.083 (-0.073)**	**-0.107, -0.001**	---	---	---	---
Stg fre1.0	**1.13**	**-0.282**	**-0.093 (-0.085)**	**-0.117, -0.036**	0.43	+0.118	0.041 (0.017)	-0.047, 0.062
Q fre7	---	---	---	---	0.75	-0.207	-0.072 (-0.080)	-0.163, 0
Q dur7	0.57	+0.141	0.047 (0.038)	-0.061, 0.113	0.78	-0.216	-0.076 (-0.080)	-0.163, 0
Max depth	**1.38**	**-0.345**	**-0.114 (-0.115)**	**-0.182, -0.080**	**1.17**	**+0.325**	**0.114 (0.103)**	**0.059, 0.158**
Max width	**1.22**	**-0.306**	**-0.101 (-0.098)**	**-0.135, -0.066**	---	---	---	---

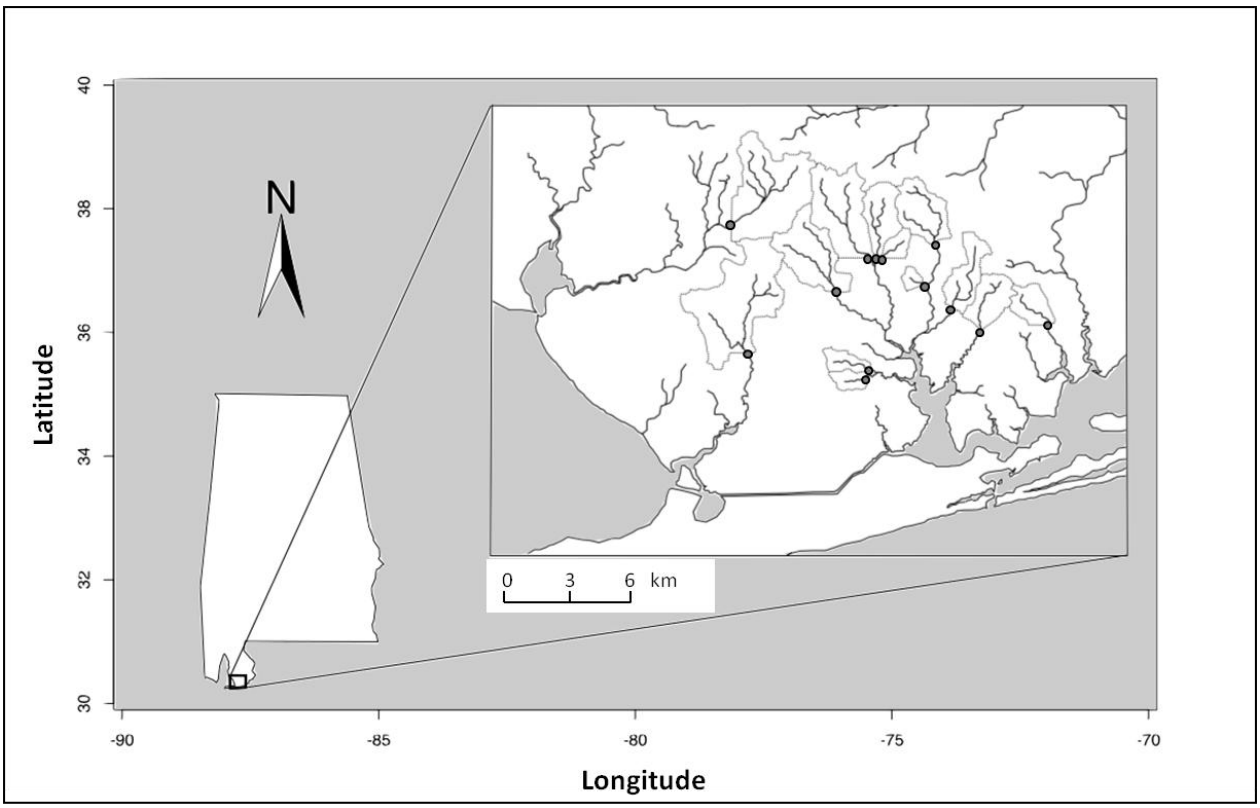

Figure 1. Study site locations in and adjacent to the Wolf Bay Basin, Baldwin County, Alabama, USA. Light grey lines indicate watershed boundaries; circles indicate approximate sample locations.

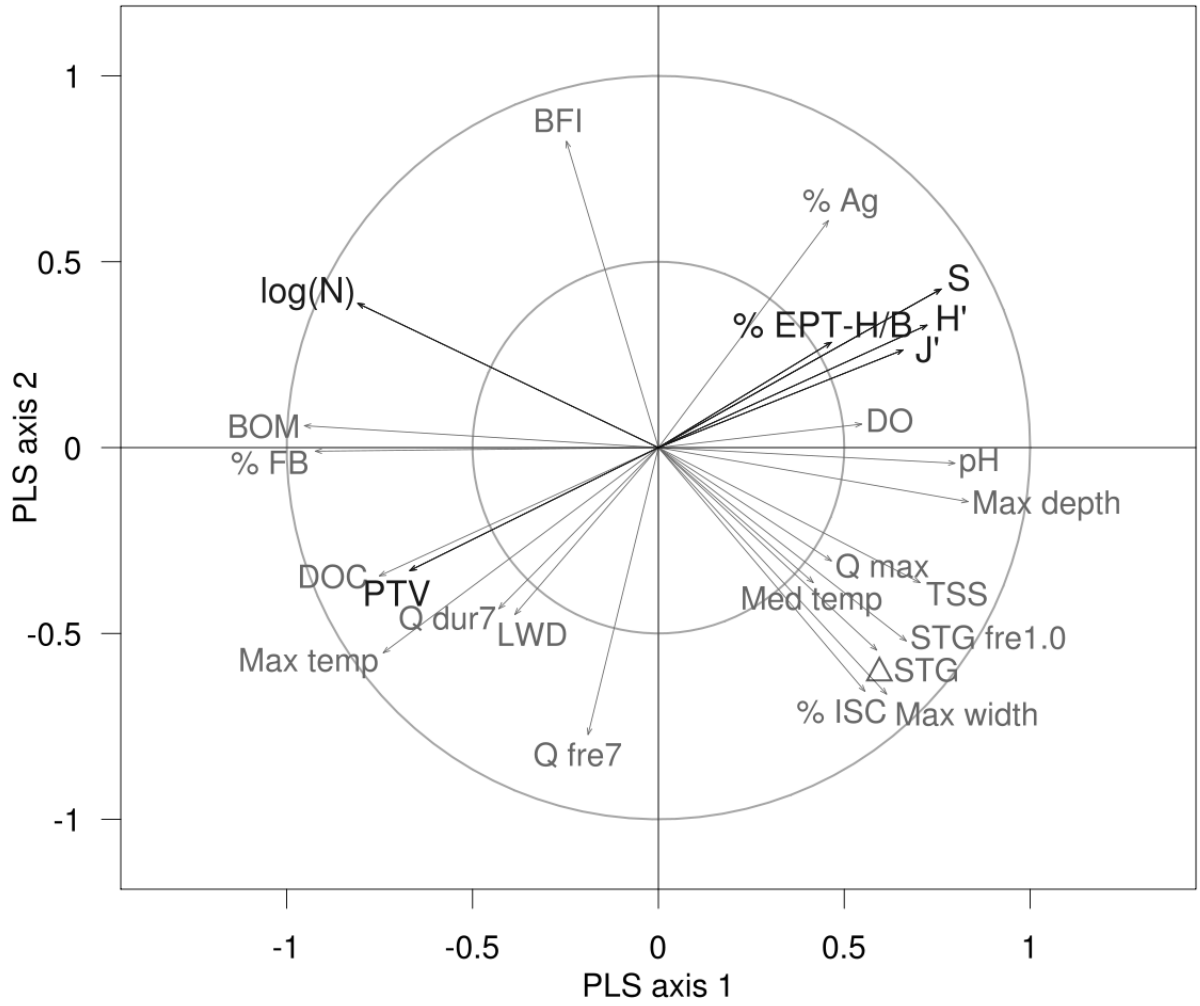

Figure 2. Multi-response PLS plot illustrating relationships between environmental predictor variables (gray), benthic metrics (black) and PLS axes. Graphic depicts correlations (as angles) between predictors, responses, and PLS axes, vector length represents the strength of the relationship with ordination axes, and the circles represent thresholds of correlation (r = 0.50 and 1.0). Variables with initial PLS VIP < 0.70 were excluded from final model. Benthic metric abbreviations and descriptions are described in Table 1. Environmental variables included the frequency and duration discharge (Q) was above 7 times the median (Q fre7 and Q dur7), frequency stage increased by ≥ 1m, maximum flood width (Max width), benthic organic matter (BOM), dissolved organic C (DOC), base flow index (BFI), and large woody debris (LWD).

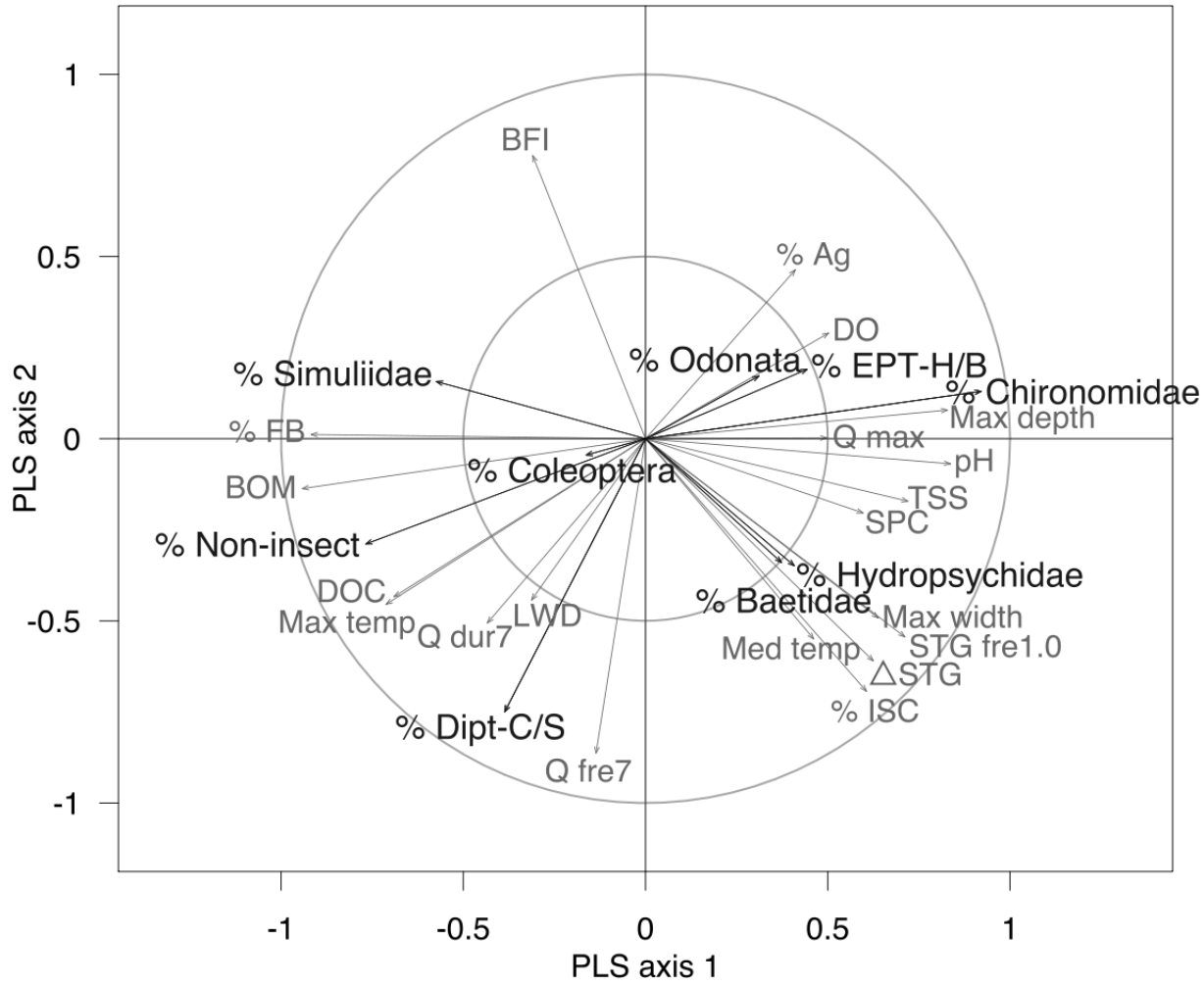

Figure 3. Multi-response PLS graphic illustrating relationships between predictor variables (gray), compositional responses (black) and PLS ordination axes. See Figure 2 for more details and Table 1 for information on compositional metrics. Environmental variables included the frequency and duration discharge (Q) was above 7 times the median (Q fre7 and Q dur7), frequency stage increased by ≥ 1m, maximum flood width (Max width), benthic organic matter (BOM), dissolved organic C (DOC), large woody debris (LWD), total suspended sediment (TSS) and base flow index (BFI).

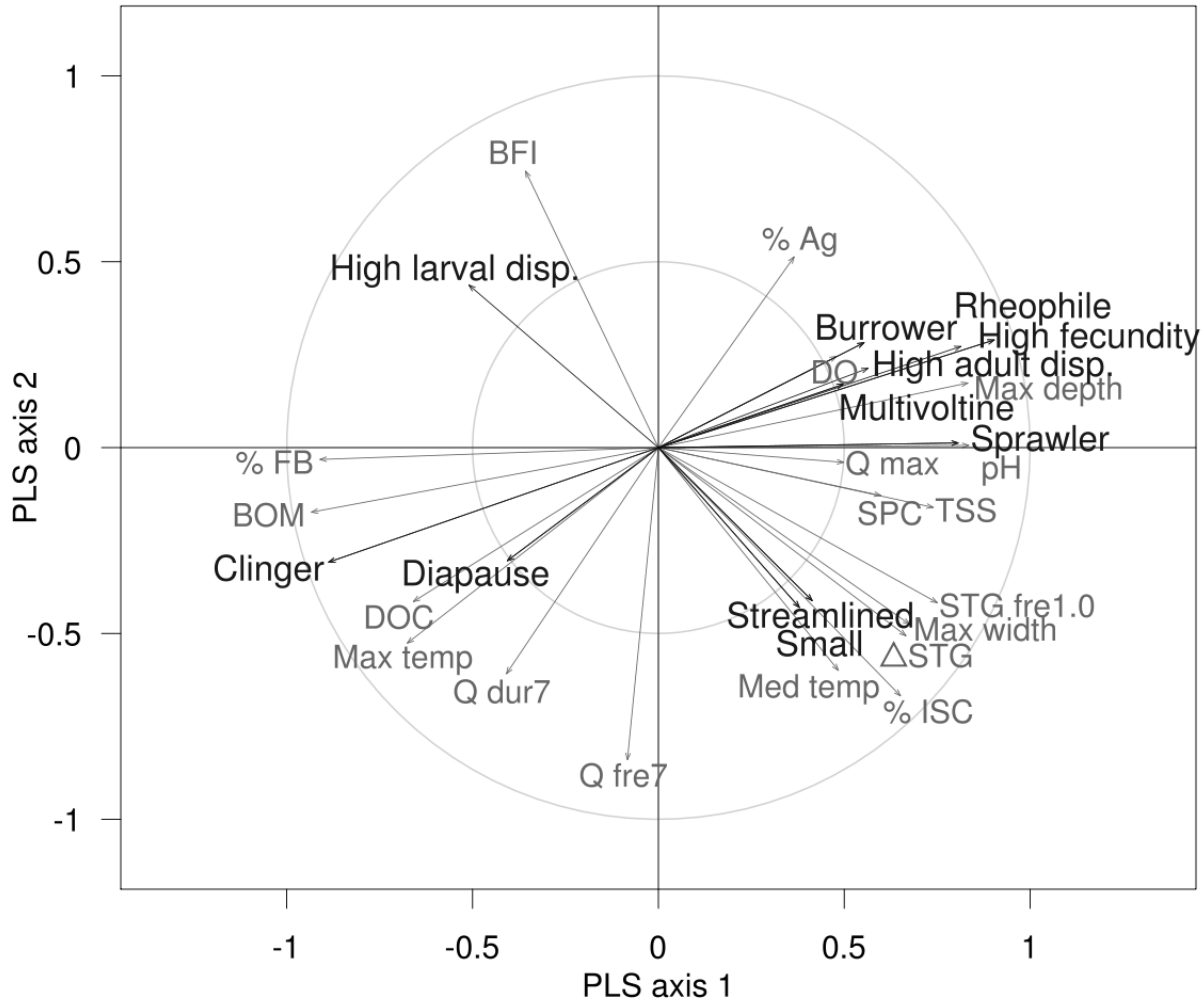

Figure 4. Multi-response PLS graphic illustrating relationships between predictor variables (gray), trait responses (**Y**, black) and PLS ordination axes. See Figure 2 for more details and Table 1 for information on trait metrics. Environmental variables included the frequency and duration discharge (Q) was above 7 times the median (Q fre7 and Q dur7), frequency stage increased by ≥ 1m, maximum flood width (Max width), benthic organic matter (BOM), dissolved organic C (DOC), base flow index (BFI), and large woody debris (LWD).

Chapter 4. On the Robustness of PLS with Simple SIMPLS Modifications

4.1 Introduction

Data are becoming increasingly larger in dimension in the biological sciences, especially in fields of genomic research (Boulesteix and Strimmer 2007). In ecology, researchers also frequently consider large numbers of explanatory/predictor variables (p) as potentially important; this will likely increase with advances in data accumulating technology (e.g., remotes sensing, automated data loggers) (Mac Nally 2000). Ecological datasets can be quite large (p) relative to the number of observations (n; $n < p$) (Mac Nally 2000) and are typically collinear in nature as well (Graham 2003). Traditional methods (e.g., least-squares (OLS) regression) cannot handle cases where $n < p$, and perform poorly when **X**-variables are highly correlated, therefore, alternative statistical methods have been suggested for consideration in ecological research (Carrascal et al. 2009; Dahlgren 2010).

Generally, methods to deal with problems of $n < p$ and collinearity do so by variable selection and/or dimension-reduction (Graham 2003; Carrascal et al. 2009). Variable selection methods (e.g., forward stepwise) can be used to find a subset ($< n$) of predictors that optimize some criteria based on model fit (e.g., RSS); however, these methods have been criticized for several reasons, including potentially large estimation bias when predictors are omitted (Whittingham et al. 2006). Principal components analysis offers reduction in the dimensionality of a collinear predictor matrix (**X)** through the creation of an orthogonal set of variables (linear

combinations of **X)** that can be sequentially regressed against a response variable(s); however, the relationship between the response (**Y**) and **X** is not incorporated in the dimension reduction step and does not guarantee the best predictive model (Garthwaite 1994; Carrascal et al. 2009). Partial least squares (PLS) reduces dimensionality in **X** with respect to the variance in **Y**, and is likely preferential over PCA-regression when the goal is prediction (Carrascal et al. 2009).

PLS exists as two general algorithms: 1) NIPALS (Wold 1966; Wold et al. 1984) and 2) SIMPLS (de Jong 1993; described below); the latter is computationally more efficient (faster), especially when dimensions are large. PLS finds a set (size k; $k \leq p$) of orthonormal vectors $\mathbf{T}_k$ ($\mathbf{T}_k=\mathbf{X}w_k$), where weights ($w_k$) are chosen to maximize covariance between **Y** and $\mathbf{T}_k$. **Y**is then regressed on $\mathbf{T}_k$ and conventional regression coefficients (relating **X** to **Y**) are back-calculated (de Jong 1993). PLS solutions are biased; however, under non-ideal conditions (e.g., small n, collinearity), PLS can offer lower variability (with only small bias) about parameter estimates and more accurate predictions relative to least-squares (OLS; see Chapter 1). PLS has few assumptions other than the underlying system is actually a function of underlying and unmeasured "latent" variables (Wold et al. 2001). PLS does not assume independence (and uncorrelatedness) in **X**, and unlike OLS, which assumes **X** was designed (fixed **X** values) and measured w/o error, PLS tolerates noise in X (Wold et al. 2001).

PLS has obvious appeal, but relies on estimates of location (mean) and scale (variance) and is therefore sensitive to outliers and otherwise non-normally distributed data (e.g., wide tails, skewness) (Kruger et al. 2008). Outliers are generally thought of as unusual observations (a.k.a. cases, entire rows, objects: $x_{i.}$) that do not conform to the patterns shown by, or are distant from, the majority of the data (Møller et al. 2005). Almost every proposed robust method (e.g., multiple regression, PCA) in the literature was created assuming outliers occur across entire rows

(Møller et al. 2005); thus, these methods work to identify and downweigh the influence of whole observations. While unusual data may be correct and valid data, outliers exist in real data almost as a rule due to (among others) machine error, calibration issues, and copying/recording mistakes (Møller et al. 2005; Rousseeuw et al. 2006). Therefore, outliers are not confined to existing as entire observations and can occur as variables (columns: $x_{\cdot j}$) or as individual elements within a dataset (x_{ij}) (Møller et al. 2005; Rousseeuw et al. 2006).

The first attempt at robust PLS consisted of the replacement of one or more OLS solutions in the nonlinear iterative partial least-squares (NIPALS) algorithm with a robust alternative (e.g., Theil-Sen), and offered local or global-levels of robustness, but at potentially high computational costs and low relative efficiency (Wakeling and Macfie 1992; Møller et al. 2005). PLS can be made resistant to outliers in at least several other ways and many methods have been proposed (reviewed/mentioned in Gil and Romera 1998, Møller et al. 2005, and Kruger et al. 2008). Iterative reweighed PLS algorithms use either simple regression residuals to reweigh "internally" (OLS residuals within NIPALS) or "externally" (PLS residuals); but residual reweighing only protects against vertical outliers (in *y*-space) (Møller et al. 2005). The statistically inspired modification of PLS (SIMPLS) algorithm (de Jong 1993) is computationally efficient and can be made robust by the replacement of the initial covariance estimate with a robust version (Møller et al. 2005).

The covariance matrix describes the primary dimensions of the data as an n-dimensional ellipsoid, whose volume is given by its determinant, directions of principal axes given by eigenvectors and axes lengths by corresponding eigenvalues (Johnson and Wichern 2007). The minimum covariance-determinant (MCD) is a robust covariance alternative that searches subsamples of the empirical observations (data rows; e.g., of size 0.75n) for the subset with the

minimum ellipsoid volume, which ideally describes covariation in the majority of the data and is representative of the population (Rousseeuw et al. 2006). Most robust multivariate covariance matrices, however, including MCD, cannot be used in situations where $n < p$ (Gil and Romera 1998; Kruger et al. 2008; Rousseeuw et al. 2006).

A few robust PLS algorithms (RSIMPLS and PRM) are readily available and have been shown to perform relatively well in simulation studies (Hubert and Branden 2003; Serneels et al. 2005). RSIMPLS (Hubert and Branden 2003) estimates robust PLS scores (**T**) using SIMPLS and a robust covariance matrix estimated by a low-dimensional projection of the data [**X**,**Y**] via a robust PCA algorithm (ROPCA; based on projection pursuit and MCD). Outlier information is obtained in the ROPCA step and used to robustify the regression step (**Y** on **T**) (Hubert and Branden 2003). Partial robust M-estimator (PRM) is an external iteratively reweighted M-estimator that uses the geometric mean of continuous weights (on [0, 1]) based on 1) residuals and 2) leverage scores to provide resistance to outliers in **Y**- and/or **X**-space (Serneels et al. 2005). RoPLS is an external reweighing algorithm that uses outlier detection (e.g., BACON) to identify potential outlying observations and iteratively assigns weights to reduce the influence of outlying observations on the PLS solution (Turkmen 2008).

While the abovementioned and additional "complicated" robust PLS algorithms exist (González et al. 2009; Kruger et al. 2008; Møller et al. 2005), plugging robust estimators into classical/standard algorithms may adequately reduce the influence of outliers (Daszykowski et al. 2007). The most simple robust covariance estimators are likely rank based (e.g., bivariate Spearman's correlation), and while rank-based alternatives are not a novel idea, this simple approach was not discussed in some recent reviews on robust multivariate methods or in most robust PLS simulation studies (Daszykowski et al. 2007; Gil and Romera 1998; Møller et al.

2005; Rousseeuw et al. 2006; Serneels et al. 2005). In this study, we used simulations to 1) determine if the simple replacement of covariance estimates with robust estimates based on Spearman's or Kendall's rank correlation estimates result in outlier-resistant PLS, and 2) to compare the performance of rank-based PLS to existing robust PLS algorithms.

4.2 METHODS

4.2.1 PLS algorithms

PLS was developed by Herman Wold using the NIPALS algorithm he originally developed as an alternative method for principal components analysis (PCA) (Wold 1966; Wold et al. 1984). Similarities between PCA and PLS can be seen in the construction of their respective weight vectors (w_k, aka: loadings) and orthogonal components ($t_k = \mathbf{X}w_k$, aka: scores) (Boulesteix and Strimmer 2007):

$$PCA: w_k = \underset{\|w\|_2 = 1}{\operatorname{argmax}}(var(\mathbf{X}w_k)), \text{ and} \quad [1]$$

$$PLS: w_k = \underset{\|w\|_2 = 1}{\operatorname{argmax}}(cov(\mathbf{X}w_k, \mathbf{Y})). \quad [2]$$

PCA components sequentially decrease in the total variation of **X** each explains; PLS's **X** components sequentially decrease with regards to their covariation with, or ability to explain variance in, **Y**.

NIPALS uses an iterative series of OLS solutions (power iteration) to calculate the largest left and right eigenvectors (PLS weights) of the cross-products **X**'**Y** matrix (denoted **S** hereafter; as it is $\propto$ covariance matrix) (Wakeling and Macfie, 1992). Successive PLS component weights are calculated with updated **X** and **Y** matrices ("deflated" using regression residuals) orthogonal to the components created previously (Wakeling and Macfie, 1992). The (next) largest eigenvectors are calculated, the data are deflated, and so on; thus, NIPALS quickly becomes computationally expensive with high dimensional data (Wakeling and Macfie 1992; Gil

and Romera 1998). The SIMPLS algorithm is computationally much faster than NIPALS due to the use of singular value decomposition (SVD) to obtain eigenvalues directly from **S** and sequential deflation of **S** (not **X**) before each calculation additional PLS components (Table 1) (de Jong 1993). Due to deflation differences, the two algorithms produce slightly different results when **Y** is multivariate, but identical results in the case of univariate **Y** (y) (de Jong 1993).

4.2.2 Rank-based SIMPLS

We created two SIMPLS alternatives by replacing the initial cross-product between **Y** and **X** in the SIMPLS algorithm (Table 1) with cross-products based on Kendall's and Spearman's pairwise correlation coefficients (Visuri et al. 2000). If x and y are centered, then:

1) Pearson's product-moment correlation (PPMC), r_{xy}(x,y) = x'y$[(sd(x)sd(y)(n-1))]^{-1}$; where sd(x) is the standard deviation of x.
2) Spearman's rank correlation, rho_{xy}(x,y) = r_{xy}(rank$_c$(x), rank$_c$(y)); PPMC as described above with input as centered ranks (rank$_c$) of x and y.
3) Kendall's rank correlation, tau_{xy}(x,y) = $\left(\#C_p - \#D_p\right)\left(0.5\mathrm{N}(\mathrm{N}-1)\right)^{-1}$; where C_p and D_p are the number of concordant (agreeing on ranking; e.g., $x_i > x_j$ and $y_i > y_j$) and discordant (disagreeing; e.g., $x_i > x_j$ and $y_i < y_j$) pairs, based on rank order agreement.

Kendall's rank correlation (*tau*) and Spearman's rank correlation (*rho*) quantify different population-level values than PPMC, but comparable values can be obtained at the normal model using angular transformations (Moran 1948; Visuri et al. 2000; Croux and Dehon 2010; Boudt et al. 2012). Robust cross-products using transformed rank correlations can be calculated as:

$$\mathbf{S}_{tau} = \sin[(1/2)\pi \bullet tau_{xy}(\mathbf{X},\mathbf{Y})] \bullet sd(\mathbf{X}) \bullet sd(\mathbf{Y}) \bullet (n-1) \quad [3]$$

$$\mathbf{S}_{rho} = 2\sin[(1/6)\pi \bullet rho_{xy}(\mathbf{X},\mathbf{Y})] \bullet sd(\mathbf{X}) \bullet sd(\mathbf{Y}) \bullet (n-1) \quad [4]$$

4.2.3 Simulation details

Data generation

We generated *y* to be a function of 4 of 60 **X** variables ($y = -1.25X_1 + 1.25X_2 + 1.25X_3 - 1.25X_4 + N(0, 1)$), where $\mathbf{X}_{(40,60)}$ follows a multivariate normal distribution, with a mean vector of 0s and correlation structure derived from an Environmental Protection Agency (EPA) dataset (Paulsen et al. 2008). These EPA data included 52 highly collinear variables (Fig. 1; the last eight in **X** were uncorrelated random variables) characterizing various measures of land-cover, water chemistry, stream hydrology and geomorphology from 40 stream/watershed sites from the Piedmont ecoregion of the eastern US (Paulsen et al. 2008).

Simulated variables $X_1, \ldots, X_4$ corresponded to percent watershed as agriculture (Ag, X_1), mean substrate diameter (mm, X_2), mean stream width (m, X_3), and stream water ammonium (mg/L, X_4) concentration in the real EPA dataset. The sign of population slope (β) values for our simulated response (*y*) were chosen so that X_1 (Ag) and X_4 (NH_4^+) negatively, and X_2 (substrate size) and X_3 (stream width) positively influenced *y*.

We created clean training and test datasets as described above. A percentage of outlying data points (ε; 0 – 40%) were added to training data in either *y*, **X**, or [*y*, **X**]. Outliers placed in **X**, or [*y*, **X**] were done so as either 1) in n•ε entire rows (observations/cases) or 2) n•ε•p elements were randomly replaced (independent of row/column). In both cases, the same total number of elements were randomly replaced, and each at a distance of 5 $\pm$ N(0, 0.1) from the data edge (min or max).

4.2.4 PLS methods

For comparison, we included several available robust PLS algorithms: 1) RSIMPLS (Hubert and Branden 2003) available in the Matlab Libra library, 2) PRM (Hubert and Branden 2003) in R package "chemometrics", and 3) RoPLS (Turkmen 2008), Matlab code was obtained

from the author. We adapted MATLAB code for use in Octave; then the package "RccpOctave" allowed us to use Octave/MATLAB functions in R-based simulations.

The choice of the number of latent variables to retain (k) in practice is generally determined with cross-validation (CV) or by the optimization of some function of the model residual sum of squares (Wold et al. 2001). With every iteration and method, we used 2/3^{rds} random subset CV (20 iterations) on the training set and determined optimal choice of k as the model size with minimum mean RMSE. Robust-CV is suggested to improve the choice of k when outliers are present; one such way is to remove observations with large residuals before calculating RMSE (Hubert and Branden 2003).

In preliminary runs with no upper limit, most chose k < 8; to reduce computational time, we allowed CV to consider the 1st eight model sizes only (k = 1 to 8, of 60), forming an nx8 matrix of residuals. The BACON algorithm (Billor et al. 2000) was then used to identify unusually large residuals (observations) to discard prior to the calculation of RMSE for that CV iteration (Turkmen 2008). We retained simulation information for both k chosen with CV and robust-CV; each simulation setting was iterated 500 times.

4.2.5 Performance criteria

We compared standard PLS (as SIMPLS), RSIMPLS, PRM, RoPLS and rank-based SIMPLS algorithms using the following performance criteria based on coefficient estimation and prediction:

1) Mean square error (MSE) of slopes estimation ($\hat{\beta}$) = $\text{MSE}_{\hat{\beta}} = \frac{1}{P}\sum_{i=1}^{P}(\hat{\beta}_i - \beta_{\text{pop}})^2$, where p = number of predictors

2) MSE of prediction ($\hat{y}$) on test data = $\text{MSE}_{\hat{y}} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_{\text{test}\,i}\right)^2$, where n = number of observations in test set

3) Relative efficiency (RE) of an alternative method, relative to the standard, in terms of coefficient estimation ($\hat{\beta}$) or prediction ($\hat{y}$), where: $\mathrm{RE}_{\mathrm{alternative}} = \overline{\mathrm{MSE}}_{\mathrm{standard}} \Big/ \overline{\mathrm{MSE}}_{\mathrm{alternative}}$

4.2.6 Real data example

We focus on univariate response (*y*) for this simulation study; however, rank-based PLS methods should behave similar with multivariate (**Y**; PLS2**)**, relative to other methods, whether it is uni- or multivariate. As real data example, we compared model components (T, $\hat{\beta}$) from standard PLS2 model and rank-based PLS2 models with highly correlated responses. These data are from 20 small-sized streams (< 4th order) in the Piedmont region of Virginia and were collected as a part of the EPA's wadeable stream assessment program (Paulsen et al. 2008). Stream water total nitrogen (TN) and phosphorus (TP) were used as the response matrix (**Y**), and 19 predictor variables (**X**) included watershed size, several land-use and geomorphic variables, other chemical variables (etc.). We fit models to these data, then introduced a small amount of outliers (5%) and compared changes in model estimates graphically and with sum-of-squared differences (SS) relative to the standard PLS2 solution.

4.3 Results

4.3.1 No outliers

Average variance of correlations between simulated predictors (without outliers) was very small (0.022), thus correlation structure was consistently near to the specified values (in Fig. 1). In our simulation, both Spearman and Kendall rank-based PLS algorithms showed relative efficiencies (RE; to standard PLS) of coefficient estimation and prediction around 95-96% at the normal model (Fig 2; top panel). RoPLS and PRM had the highest efficiencies (97-100%) and RSIMPLS had the lowest (89-94%). Between-method trends were similar with

regards to RMSE of both coefficient estimation and prediction; although the former was more variable (Fig. 2; top panel).

4.3.2 Vertical outliers

Note that with vertical outliers (in *y*), the placement of outliers (rows or random) is irrelevant and identical results should be expected. PRM performed quite well relative to the other methods when outliers were placed only in *y* (Figs. 3 & 5, left panels). RSIMPLS also predicted well for contamination between 2.5- 25%, corresponding to the (25%) default setting for the proportion of outlying observations to resist. Rank-based PLS did predict better (lower $RMSE_{\hat{y}}$) than standard PLS starting with low proportion of outliers placed across observation in each setting ($\cong$ 5% in *y* or [*y*, **X**]; Figure 3). Robust CV did not improve performance for any method with outliers only in *y* (Fig. 3, bottom left panel).

4.3.3 Outliers placed across observations/rows

Every method had a lower median RMSE of coefficient estimation and prediction compared to standard PLS with 10% outlier contamination across entire rows in [*y*,**X**] and RSIMPLS and PRM predicted markedly better than the other methods in terms of prediction (Fig. 2; middle panel). RSIMPLS and PRM predicted well across a range of contamination percentages when outliers were placed in [*y*,**X**] (Fig. 3). Rank-based PLS did predict better (lower $RMSE_{\hat{y}}$) than standard PLS starting with low percentages of outliers placed across observation in each setting ($\cong$ 5% in *y* or [*y*, **X**]; Figure 3). Rank-based PLS generally performed similar to the other robust methods (roughly equivalent RMSE) when outliers occurred in **X** or [*y*, **X**], and performed better when contamination exceeded 25%. RSIMPLS, while having poor RE (at 0% contamination) resisted outliers well between 0-25% (default setting to resist 25%

contamination). PRM did not outperform rank-based or standard PLS when outliers were placed only in **X** (Fig. 3).

When outliers where placed in **X** or [*y*,**X**], but not *y* alone, choosing optimal k with robust CV (training set) improved test set predictions for all methods, including standard SIMPLS (Fig. 3). Across all methods, more PLS components were retained each iteration for robust CV than standard CV ($k_{rob} - k_{std}$ ranged from + 2 to 6) when outlying rows are identified and eliminated (Fig. 4).

4.3.4 Outliers placed randomly throughout

Rank-based PLS estimated linear model coefficients and predicted markedly better than the remaining methods with 10% outlier contamination placed randomly as individual elements (Fig. 2; bottom panel). Rank-based PLS also predicted better across a range (starting ~ 5%) of contamination percentages, whereas RSIMPLS predicted worse and PRM predicted either worse than or roughly equivalent to standard PLS (Fig 5). Robust CV did not improve, or otherwise noticeably change test set predictions for any methods choice of number of components retained (Fig. 5).

Although these results are not provided, the same trend of diverging RMSE for rank-based PLS relative to the other methods (as in Fig. 5) can be seen with a set 10% contamination, but by varying the distance of outliers; this divergence began with outliers as few as 1 units distance from the data edge (prior results used a set distance of 5 units +/- N(0, 0.1)).

4.3.4 Real data example

Rank-based multivariate **Y** PLS (PLS2) estimated very similar 1st component (T_1) scores and slope vectors ($\hat{b}_1, \hat{b}_2$) to standard PLS2 with these data (Fig. 6, top 2 panel rows). Five percent of the data elements (in [**Y**, **X**]; 21 total) were randomly replaced with outliers and

models were fit using standard PLS and the rank-based PLS methods. Standard PLS estimates for T_1 were highly influenced by at least two data points and coefficient estimates were greatly different than those calculated with the original data (Fig.6, middle row). Conversly, rank-based PLS model estimates for these data with outliers were much more reasonable and more similar to standard PLS estimates without outlier contamination (Fig.6, bottom 2 rows).

4.4 Discussion and Conclusions

Little research has focused on multivariate methods to handle outliers as individual elements (Møller et al. 2005) and none for robust PLS. In this simulation, rank-based PLS algorithms were 95-96% efficient, relative to standard PLS with no outlier contamination, in terms of regression coefficient estimation and prediction (Fig. 2). These simulation results also indicated that rank-based PLS generally outperformed standard PLS in terms of both coefficient estimation and prediction when outliers are present either y, $\mathbf{X}$, or [y, $\mathbf{X}$] across rows or randomly placed in individual elements. Lastly, the positive attributes highlighted for univariate y PLS (PLS1) in this simulation also appear to extend to multivariate $\mathbf{Y}$ PLS (PLS2) as well (Fig. 6).

Not surprisingly, the observation-based robustification of PLS (i.e. RSIMPLS, PRM) can work well when outliers are nicely placed across observations. Although RSIMPLS and PRM were highly efficient relative to standard PLS in this simulation without outliers, these methods showed no resistance to outliers when they were placed randomly throughout the data. In addition, the robust CV procedure used in this simulation also works based on removing influential observations and only improved predictions when outliers were placed across observations. These observation-based strategies/methods might be ideal for situations where entire observations are susceptible to mislabeling, contamination, or machine error; however,

observation-based approaches should not be expected to impart robustness against randomly placed outliers.

Interestingly, the ways to robustify PLS can be thought of as modular. For example, we incorporated rank cross-products into the PRM algorithm and it performed slightly better than rank-based PLS did in the current study with randomly placed outliers (similar to Fig. 5; results not included), but this rank-PRM performed intermediate to rank-based PLS and PRM when outliers were placed across rows or just in *y* (see Fig. 3). These results suggest that it would be fruitful to examine the performance of different combinations of these "modular parts" for the robustification of the SIMPLS algorithm to be used with various data types.

While an examination of multivariate **Y** PLS2 was not a major component of this study, the real data example provided in our study indicated that rank-based PLS2 can produce results similar to standard PLS with no (additional) contamination and provided more reasonable results when the data were contaminated with outliers. PLS2 is mathematically similar to several methods commonly used in ecological research, including canonical correlation analysis (CCA; Borga et al. 1997) and co-inertia analysis (Dolédec and Chessel 1994, Dray et al. 2003). Therefore, rank-based PLS could be used in a similar way as CCA or co-inertia have been used in ecology or robust rank-based modifications for these methods could similarly and easily be achieved.

Lastly, additional research should examine the efficacy of rank-based PLS followed by linear discriminate analysis for binary or multi-class classification. This two-step classification process has gained much attention in recent years for use with high-dimensional data, such as gene expression data for the classification of disease. Preliminary results (not provided) using rank-based PLS algorithms have indicated that Spearman-based PLS-LDA may provide some

outlier resistance in high-dimensional classification settings and interestingly out-performed standard PLS-LDA when outliers were absent from the simulated data.

4.5 References

Billor, N, Hadi, AS, Velleman, PF. 2000. BACON: blocked adaptive computationally efficient outlier nominators. Computational Statistics & Data Analysis 34:279-298.

Borga, M., Landelius, T., Knutsson, H. 1997. A unified approach to PCA, PLS, MLR and CCA. Report LiTH-ISY-R-1992, ISY, SE-581 83 Linkoping, Sweden.

Boudt, K, Cornelissen, J, Croux, C. 2012. The Gaussian rank correlation estimator: robustness properties. Statistics and Computing 22:471-483.

Boulesteix, A-L, Strimmer K. 2007. Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. Briefings in bioinformatics 8:32-44.

Carrascal, LM, Galván, I, Gordo, O. 2009. Partial least squares regression as an alternative to current regression methods used in ecology. Oikos 118:681-690.

Croux, C, Dehon, C. 2010. Influence functions of the Spearman and Kendall correlation measures. Statistical Methods & Applications 19:497-515.

Dahlgren, JP. 2010. Alternative regression methods are not considered in Murtaugh. 2009. or by ecologists in general. Ecology Letters 13:E7-E9.

Daszykowski, M, Kaczmarek, K, Vander Heyden, Y, Walczak, B. 2007. Robust statistics in data analysis - a review: basic concepts. Chemometrics and Intelligent Laboratory Systems 85:203-219.

de Jong, S. 1993. SIMPLS: an alternative approach to partial least squares regression. Chemometrics and Intelligent Laboratory Systems 18:251-263.

Dolédec, S. and D. Chessel. 1994. Co-inertia analysis: an alternative method for studying species-environment relationships. Freshwater Biology 31:277-294.

Dray, S., D. Chessel, and J. Thioulouse. 2003. Co-inertia analysis and the linking of ecological data tables. Ecology 84:3078-3089.

Garthwaite, PH. 1994. An interpretation of partial least squares. Journal of the American Statistical Association 89:122-127.

Gil, JA, Romera, R. 1998. On robust partial least squares (PLS) methods. Journal of Chemometrics 12:365-378.

González, J, Pena, D, Romera, R. 2009. A robust partial least squares regression method with applications. Journal of Chemometrics 23:78-90.

Graham, MH. 2003. Confronting multicollinearity in ecological multiple regression. Ecology 84:2809-2815.

Hubert, M, Branden, KV. 2003. Robust methods for partial least squares regression. Journal of Chemometrics 17:537-549.

Johnson, RA, Wichern, DW. 2007. Applied multivariate statistical analysis Vol 4. 6th Edn. Prentice hall Englewood Cliffs, NJ

Kruger, U, Zhou, Y, Wang, X, Rooney, D, Thompson, J. 2008. Robust partial least squares regression: Part I, algorithmic developments. Journal of Chemometrics 22:1-13.

Mac Nally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between - and reconciliation of - 'predictive' and 'explanatory' models. Biodiversity & Conservation 9:655-671.

Møller, SF, von Frese, J, Bro, R. 2005. Robust methods for multivariate data analysis. Journal of Chemometrics 19:549-563.

Moran, P. 1948. Rank correlation and product-moment correlation. Biometrika 35:203-206.

Nguyen, T.S., Rojo, J. 2009. Dimension reduction of microarray gene expression data: the accelerated failure time model. Journal of Bioinformatics and Computational Biology 7:939-954.

Paulsen, S.G., Hawkins, C.P., Sickle, J.V., Yuan, L.L., Holdsworth, S.M. 2008. An invitation to apply national survey data to ecological research. Journal of the North American Benthological Society 27:1017-1018.

Rosipal, R., Krämer, N. 2006. Overview and recent advances in partial least squares. In: Subspace, Latent Structure and Feature Selection. Springer, pp 34-51.

Rousseeuw, P.J., Debruyne, M., Engelen, S., Hubert, M. 2006. Robustness and outlier detection in chemometrics. Critical reviews in analytical chemistry 36:221-242.

Serneels, S., Croux, C., Filzmoser, P., Van Espen, P.J. 2005. Partial robust M-regression. Chemometrics and Intelligent Laboratory Systems 79:55-64.

Turkmen, A. 2008. Robust partial least squares for regression and classification. Dissertation; Auburn University, AL, USA.

Visuri, S., Koivunen, V., Oja, H. 2000. Sign and rank covariance matrices. Journal of Statistical Planning and Inference 91:557-575.

Wakeling, I., Macfie, H. 1992. A robust PLS procedure Journal of Chemometrics 6:189-198.

Whittingham, M.J., Stephens, P.A., Bradbury, R.B., Freckleton, R.P. 2006. Why do we still use stepwise modelling in ecology and behaviour? Journal of Animal Ecology 75:1182-1189.

Wold, H. 1966. Estimation of principal components and related models by iterative least squares. Multivariate analysis 1:391-420.

Wold, S., Ruhe, A., Wold, H., Dunn, W.J. 1984. The collinearity problem in linear regression. The partial least squares. PLS. approach to generalized inverses. Journal on Scientific and Statistical Computing 5:735-743.

Wold S., Sjöström M., Eriksson L. 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems 58:109-130.

Table 1. Simplified SIMPLS algorithm (de Jong 1993) for PLS regression with univariate y.

Step	Algorithm	Description
1	$\mathbf{S} = \mathbf{X}'\mathbf{Y}$	Cross-product of mean centered **X** and **Y**
2	For i = 1, …, K	
3	q = $V_{SVD(S)}$ [,1]	$V_{SVD(S)}$[,1] = 1st column of right singular value of SVD(**S**); q = 1 when **Y** is univariate
3	w = **S**q	**X**-component weights
4	t = **X**w	**X**-component scores
5	t = t(t't)$^{-1/2}$	Normalize scores
6	w = w(t't)$^{-1/2}$	Adapt weights for normalized t
7	p = **X**'t	**X**-component loadings
8	q = **Y**'t	**Y**-component loadings
10	u = **Y**q	**Y**-component scores
9	v = p	Initialize orthogonal loadings
11	If i > 1	
12	v = v – **VV**'p	Make current loadings orthogonal to previous
13	v = v(v'v)$^{-1/2}$	Normalize orthogonal loadings
12	u = u – **TT**'u	Make v orthogonal to previous loadings
13	End	
14	**S** = **S** – vv'**S**	Deflate S with respect to orthogonalized loadings
15	$\mathbf{M}_{ab} \leftarrow \begin{bmatrix} m_{11} & \dots & m_{1b} \\ \dots & \dots & \dots \\ m_{1a} & \dots & m_{ab} \end{bmatrix}$	Store w, t, p, q and v vectors as columns in respective uppercase & bold-face matrices
16	End	
17	**β** = **WQ**'	Calculate regression coefficients

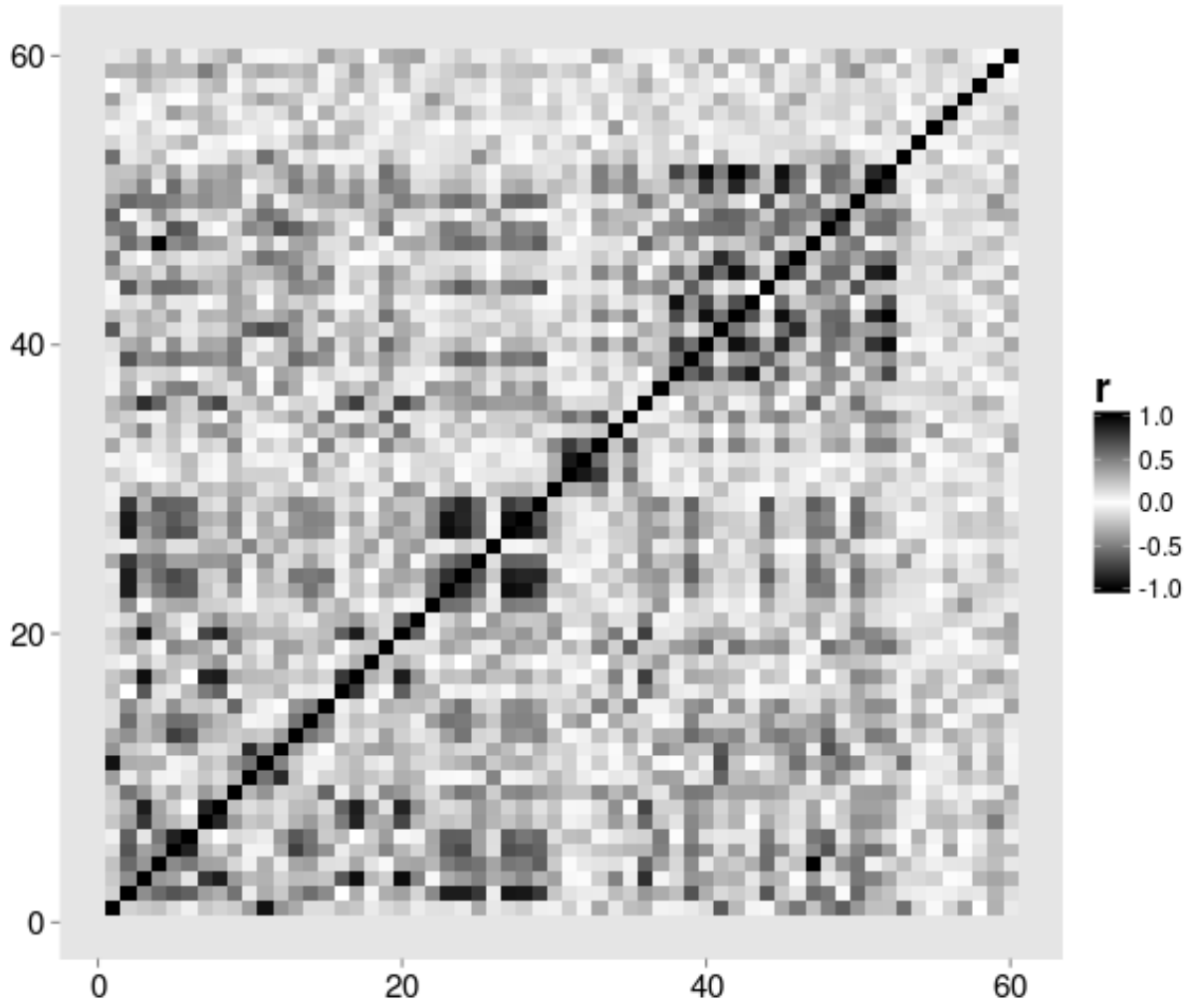

Figure 1. Black and white heat map showing correlation structure from US EPA dataset; correlation structure was used to simulate collinear data. Dark grey to black colors indicate high-levels to perfect correlation (either positive or negative). X variables 1-4 are functionally related to *y* in this study.

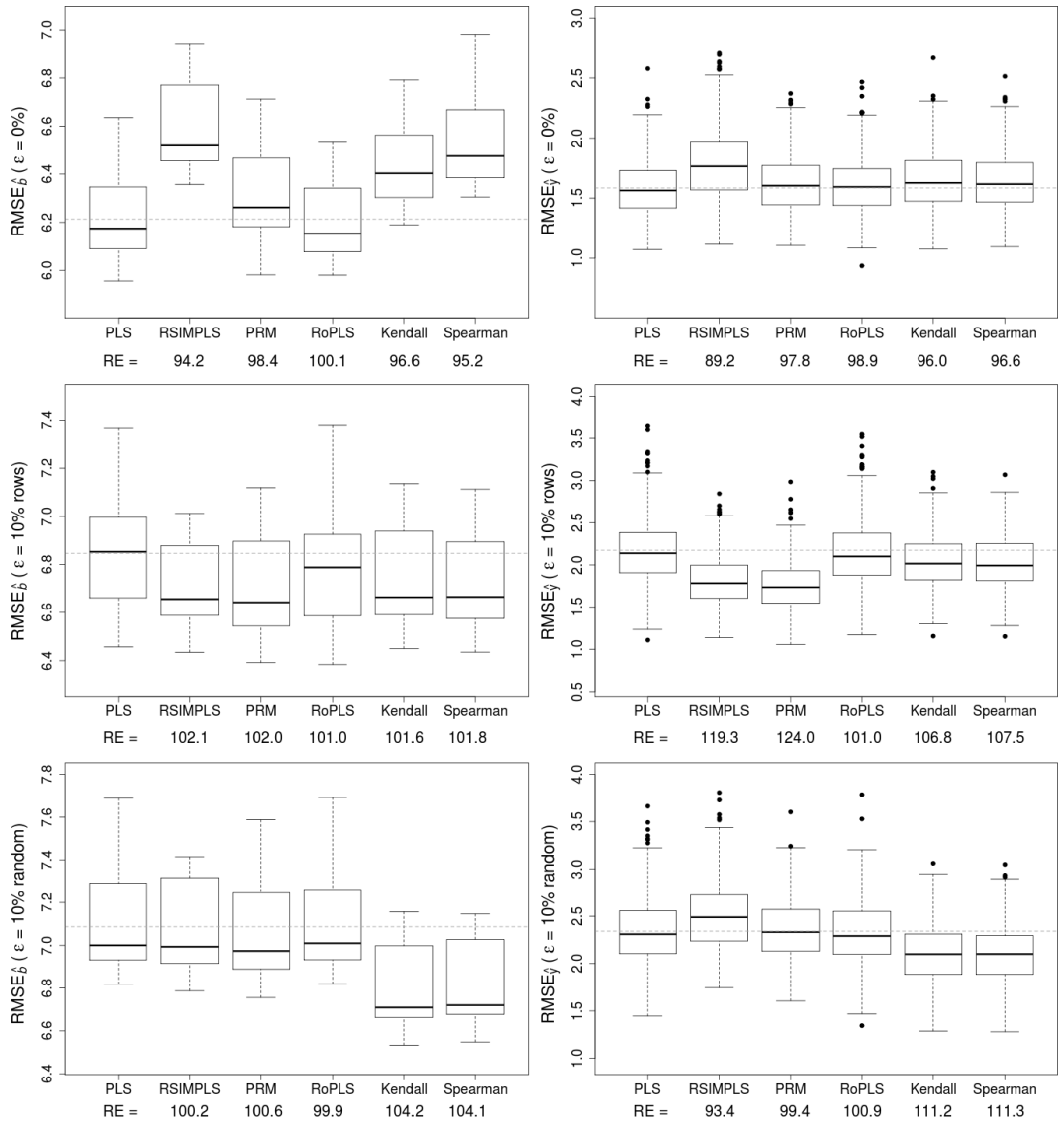

Figure 2. Boxplots of root mean square error (RMSE) with no outliers (top panels), 10% outliers in entire rows (middle panels) and 10% randomly placed outliers (bottom panels). The left column is coefficient estimation $\text{RMSE}_{\hat{b}}$, and the right column is prediction $\text{RMSE}_{\hat{y}}$. Relative efficiencies of alternative methods to standard PLS ($\text{RE} = \overline{\text{RMSE}}_{\text{PLS}} \,/\, \overline{\text{RMSE}}_{\text{alt}}$) are provided below respective boxplots. Horizontal dashed line shows mean RMSE for standard PLS (for comparison).

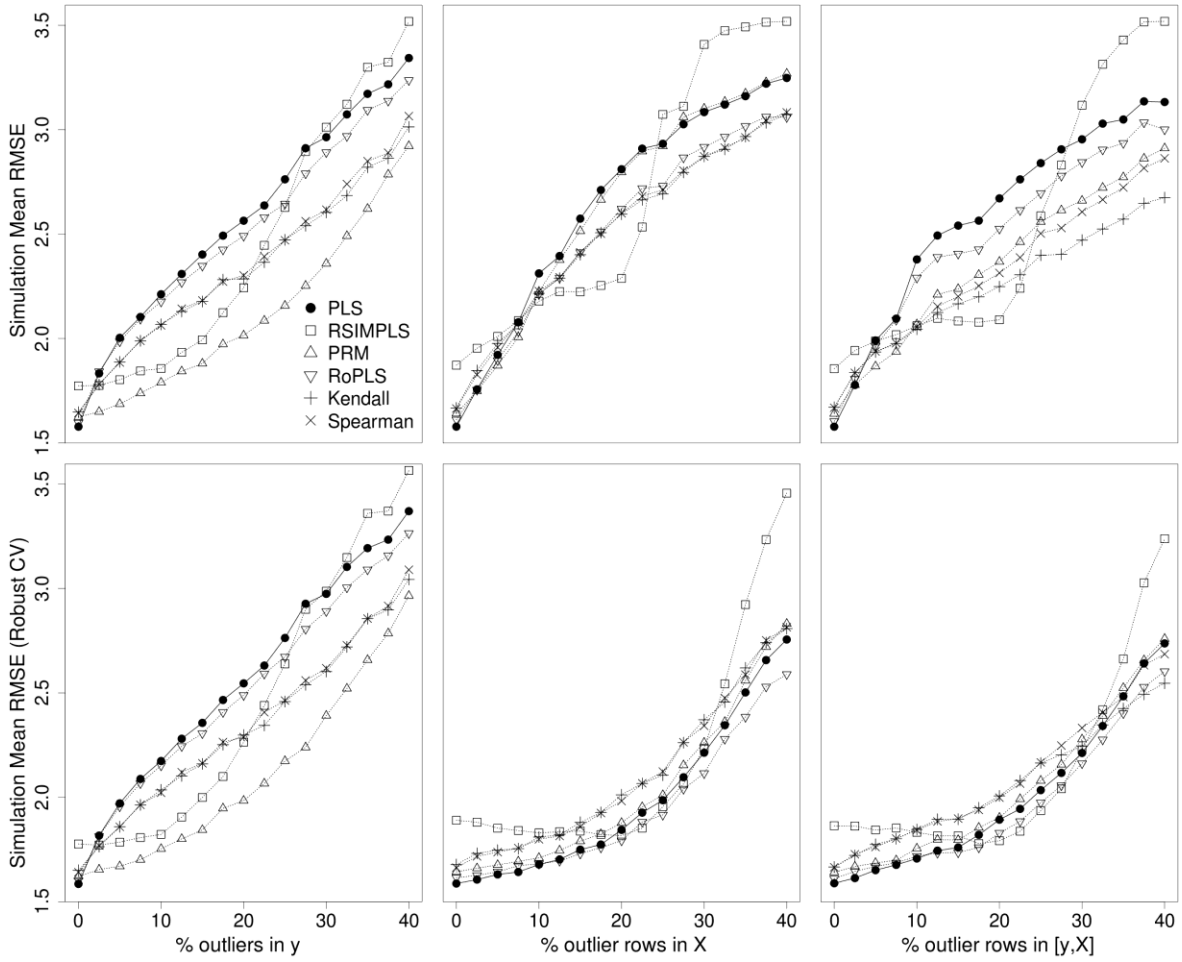

Figure 3. Simulation mean $RMSE_{\hat{y}}$ of prediction for each method with outliers in *y*, X, and [*y*, **X**] spaces; the x-axes indicate the proportion of outlying cases (entire rows). Outliers placed at 5 units past (+/-) data max/min; test data contain no outliers. Top graphs indicate mean $RMSE_{\hat{y}}$ values with model components chosen with standard CV, while lower graphs depict values using robust CV.

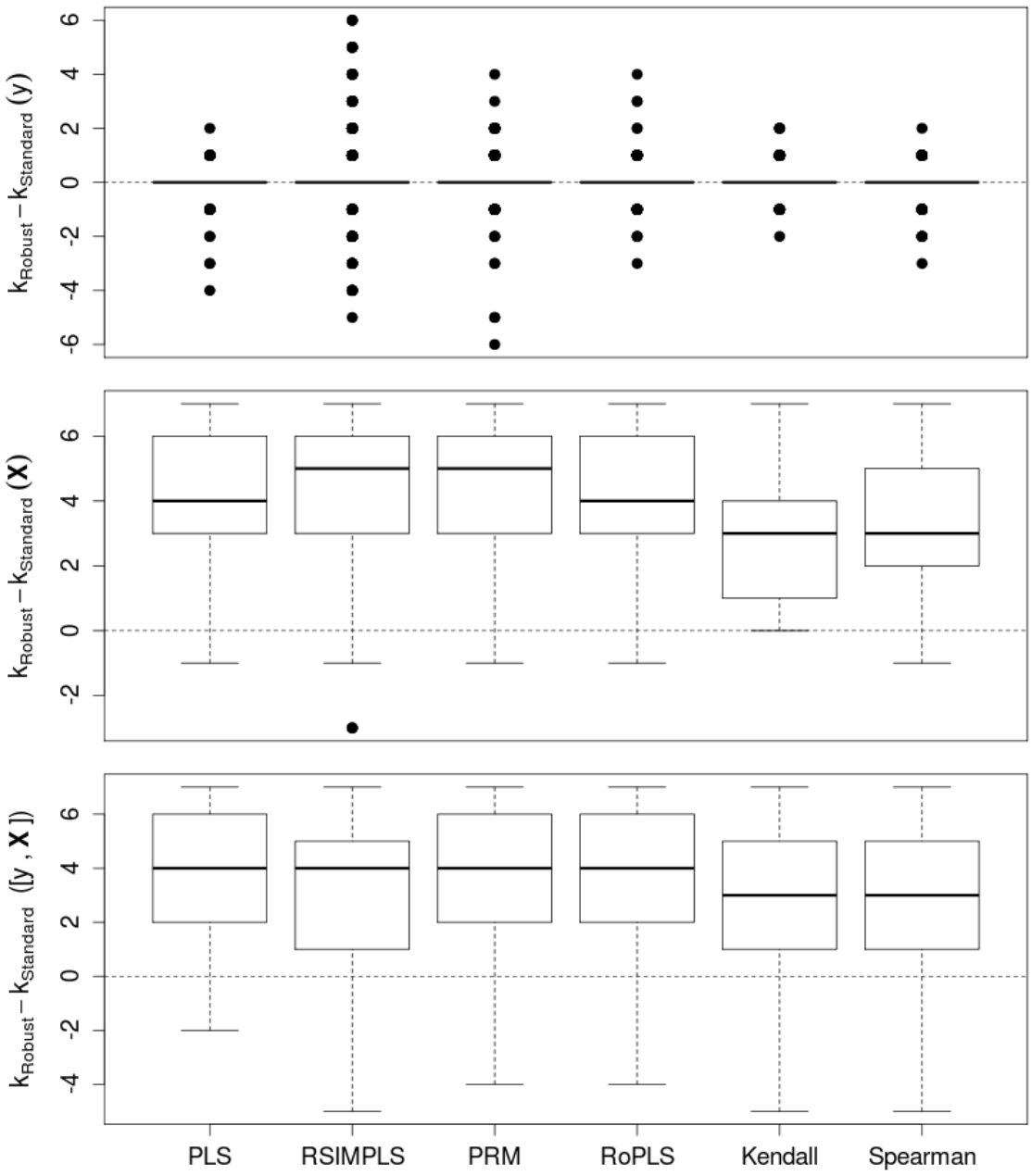

Figure 4. Difference in choice of k for robust cross validation (CV) versus standard CV (each iteration) for each method with 20% outliers in *y*, X, and [*y*, **X**] spaces. Zero indicates robust and standard CV methods chose the same value for k; positive values indicate robust CV chose larger k.

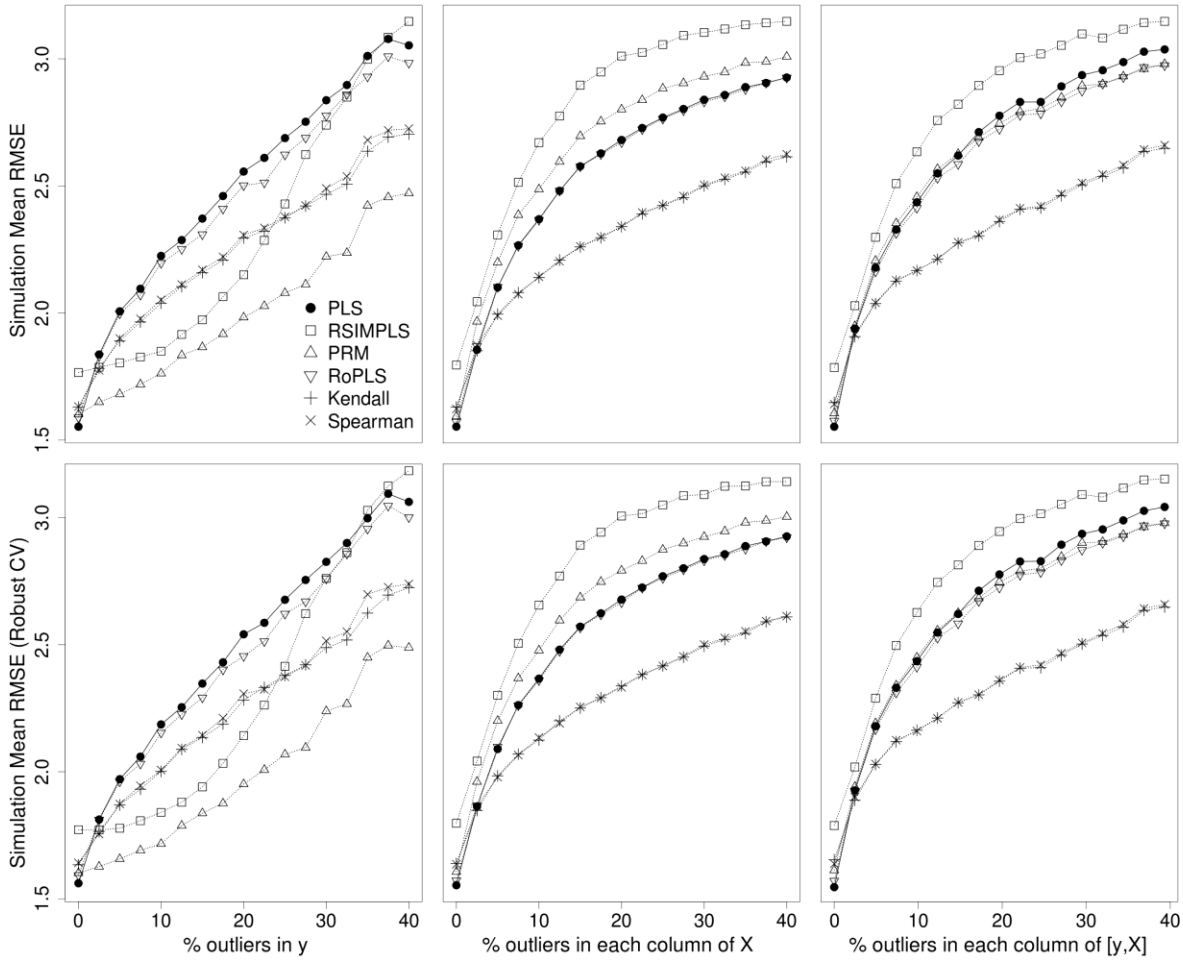

Figure 5. Simulation mean $RMSE_{\hat{y}}$ of prediction for each method with outliers in *y'*, X, and [*y*, **X**] spaces; the x-axes indicate the average proportion of outliers randomly dispersed within each column. Outliers placed at 5 units past (+/-) data max/min; test data contain no outliers. Top graphs indicate mean $RMSE_{\hat{y}}$ values with model components chosen with standard CV, while lower graphs depict values using robust CV.

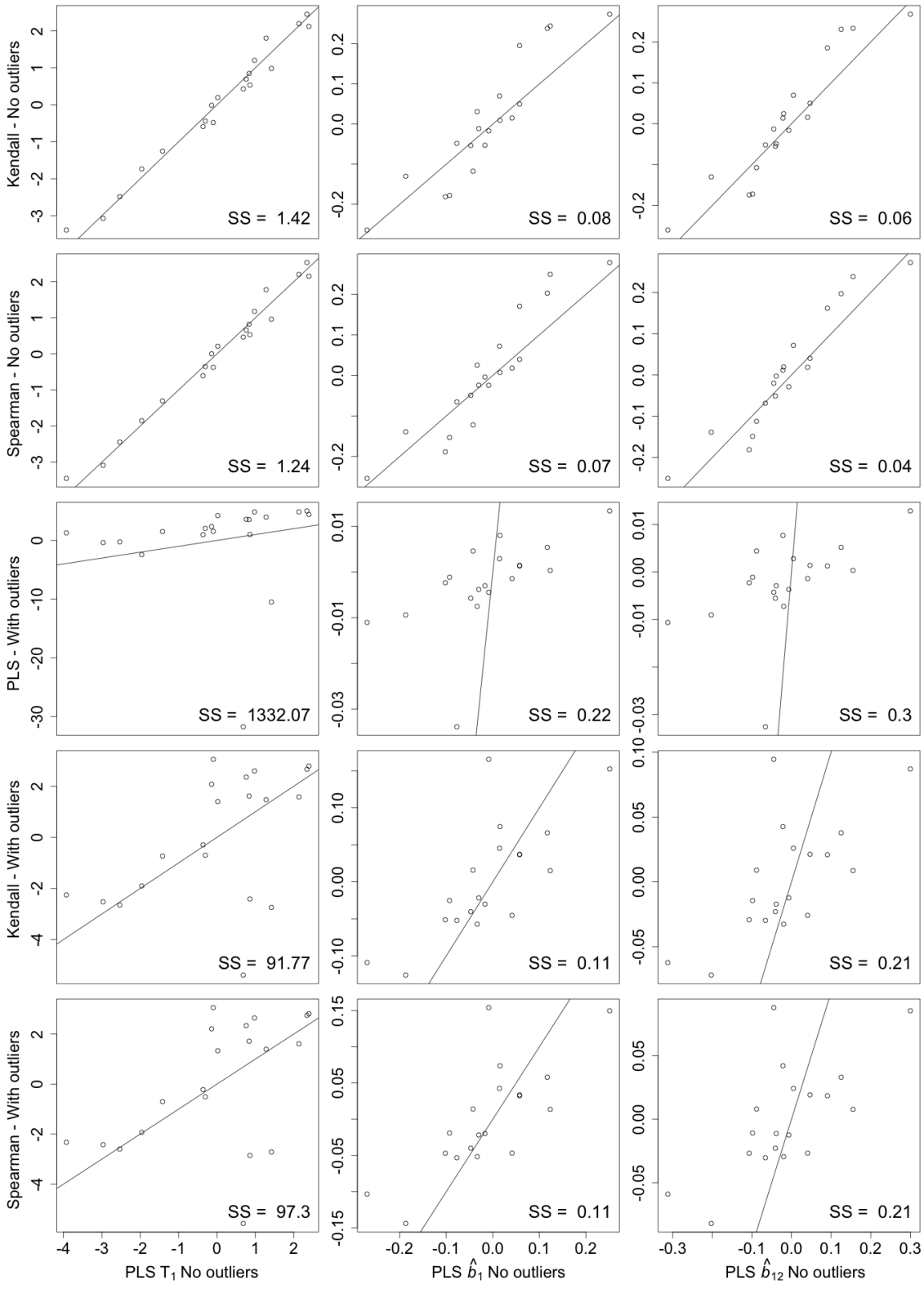
Kendall - No outliers
SS = 1.42
SS = 0.08
SS = 0.06
Spearman - No outliers
SS = 1.24
SS = 0.07
SS = 0.04
PLS - With outliers
SS = 1332.07
SS = 0.22
SS = 0.3
Kendall - With outliers
SS = 91.77
SS = 0.11
SS = 0.21
Spearman - With outliers
SS = 97.3
SS = 0.11
SS = 0.21
PLS T1 No outliers
PLS b̂1 No outliers
PLS b̂12 No outliers

Figure 6. Real data example, prediction of multivariate Y: total nitrogen and phosphorus. Left column: 1st PLS component (T) scores, middle column: coefficient estimation for total nitrogen, right column: coefficient estimation for total phosphorus. X-axis is standard PLS estimates for all panels. Top 2 rows are Kendall and Spearman estimates with original data (no outliers). Bottom 3 rows are standard PLS, Kendall and Spearman estimates for the data with added outliers. 1:1 line of perfect agreement (solid line) to standard PLS estimates without outliers (x-axes) are provided in each plot. Sum-of-squared differences (SS) between standard PLS model fir to original data and other models is provided in each panel.

Chapter 5. Benthic macroinvertebrate assemblages across ecoregions in the southeastern USA: are lowland assemblages composed of taxa inherently more resistant or resilient to land-cover stressors than those in highland regions?

5.1 Introduction

Relative to forested sites, agricultural and urban-dominated streams have been consistently associated with altered in-stream hydrology, chemical composition, physical habitat, thermal regimes, and loss of many sensitive aquatic species (Paul and Meyer 2001, Allan 2004, Walsh et al**.** 2005, Strayer 2006, Feld and Hering 2007). The southeastern US (SE) has experienced substantial deforestation and changes in land-use/cover (LULC) over the last 2 centuries (Feeley 1992, Mulholland and Lenat 1992, Smock and Gilinsky 1992, Turner et al. 2003). Agriculture dominated much of the SE in the 1800 and 1900s; this period was followed by natural and commercial reforestation of a large amount of agricultural land (Feeley 1992, Turner et al. 2003). Much of the current LULC conversion in the SE occurs as urban development; population growth and urban LULC change in this region are predicted to be among the highest across US regions over the next quarter century (Alig et al. 2004, White et al. 2008, Nagy et al. 2011).

Urbanization can be the source of both direct and indirect stressors on macroinvertebrate assemblages (Paul and Meyer 2001, Walsh et al. 2005). Fine sediment accumulation in stream beds occurs in the initial phases of urbanization when sediment delivery is high from disturbed ground cover and construction activities, filling interstitial benthic habitat (Paul and Meyer 2001). Increasing levels of impervious surface cover (ISC) alters watershed hydrology and can

lead to increased flood magnitude, frequency, and flashiness (Rose and Peters 2001, Walsh et al. 2005, Brown et al. 2009, O'Driscoll et al. 2010). As the % of ISC increases and construction declines, sedimentation is reduced, high flow events become more frequent, which increase erosion of accumulated sediment; greater hydraulic (shear) forces generally lead to channel enlargement and scouring of organic matter, large woody debris, food sources and invertebrates (Cobb et al. 1992, Walsh et al. 2005, Cordova et al. 2008, Wilzbach and Cummins 2008). In addition to hydrogeomorphic stressors to benthic organisms, extreme alterations to thermal regimes can occur in urbanized streams, which can exceed tolerance limits and impact biotic distributions (Paul and Meyer 2001, Nelson and Palmer 2007).

Across the SE, the degree to which urban LULC influences hydrology, sediment transport, thermal regimes and aquatic biota appears to decrease along a gradient of decreasing topographical relief (Utz et al. 2009, Nagy et al. 2011, Utz et al. 2011, Utz and Hilderbrand 2011). The steep topography and resistant underlying bedrock of the Appalachians (APL) greatly contrasts with the low-gradient, sandy terrain of the coastal plains (CPL; Feeley 1992, Wallace et al. 1992). These stark geomorphic contrasts translate to inter-regional differences in infiltration, runoff velocity, and the potential for urban LULC to induce a response in stream hydrology and associated biota (Booth 1990, Elosegi et al. 2010, Nagy et al. 2011, Nagy et al. 2012).

Greater proportions of the sensitive aquatic insect order Ephemeroptera were more negatively affected by watershed urbanization in the Piedmont (PMT) relative to CPL (Utz et al. 2009). Some have suggested that CPL macroinvertebrate assemblages consist of taxa that are relatively tolerant of fine sediments, unstable substrates, higher water temperatures and other urban related stressors (Utz et al. 2009, Utz and Hilderbrand 2011). CPL assemblages recolonized experimental benthic habitats faster than PMT assemblages (Utz and Hilderbrand

2011); regional differences in species traits, like recolonization potential, may help to explain differences in tolerances to environmental stressors. Some traits impart resistance to physical stressors; for example, an organism with hooks may more easily cling to stable substrate and resist extreme flows (Townsend and Hildrew 1994). On the other hand, some traits facilitate resilience, or an ability to rebound in population size following catastrophic disturbance and may include high reproductive output and/or high dispersal capabilities (Townsend and Hildrew 1994). While some difficulties exist in the analysis of traits (e.g., correlations among traits; Culp et al. 2010), the use of traits in addition to taxonomic information may provide insight into mechanistic relationships between species and their environment, and be generalizable for application in different regions (Verberk et al. 2013). While one nationwide (US) study indicated slight differences in trait composition of assemblages in least disturbed CPL and southern APL (which included PMT) streams (Zuellig and Schmidt 2012), a detailed comparson of traits among SE ecoregions is lacking.

Future land development and associated degradation of freshwater resources appear imminent in the SE (Alig et al. 2004, White et al. 2008), thus highlighting the need for research to describe regional variability in biotic responses to LULC conversion. In this study, we quantified macroinvertebrate taxonomic and trait diversity, richness, and composition within and between regions of the SE US. Our primary objective was to determine if benthic macroinvertebrate assemblages in these regions differ in trait composition, specifically those that may provide resistance/resilience to LULC-related disturbances.

5.2 Methods

5.2.1 Data description

We used data from 2 large-scale stream assessment surveys performed by the Environmental Protection Agency (EPA) for the current study. The Wadeable Streams Assessment (WSA, 2004-2005) consisted of > 1000 stream sites and the National Rivers and Streams Assessment (NRSA, 2008-2009) consisted of ~ 2000 river and stream sites that spanned the conterminous US (Herlihy et al. 2008, Paulsen et al. 2008 , U.S. EPA 2013). Site selection for both the WSA and NRSA programs was random and based upon all possible stream sites (Paulsen et al. 2008, Stoddard et al. 2008). Sites were sampled during summer low-flow and used identical (or similar) sampling protocols (e.g., D-net with 500 μm mesh, identical habitat protocols). These databases included macroinvertebrate count data as well as suite of environmental variables and are available online (http://water.epa.gov/type/rsl/monitoring/streamsurvey/web_data.cfm and http://water.epa.gov/type/rsl/monitoring/riverssurvey/index.cfm).

The EPA used aggregate ecoregions for the WSA and NRSA projects (Herlihy et al. 2008); similarly, we established the following aggregations of level III ecoregions for this study: 1) appalachian (APL: Central and S.W. APL, Blue Ridge, Ridge and Valley), 2) piedmont (PMT: Piedmont and N. Piedmont), and coastal plain (CPL: Mid. Atlantic CPL, W. Gulf CPL, S. Central Plains, S. CPL, Atl. Coastal Pine Barrens, S.E. Plains, MS Valley Loess and Alluvial Plains, E. Central TX Plains). We referred to these aggregate ecoregions herein simply as regions and occasionally referred to APL and PMT combined as highlands as comparison with the lowland CPL region. We excluded dry/intermittent stream sites, those that were non-wadeable or with a Strahler stream order > 5, and those missing data.

Stream sites were randolmly chosen; however, we wanted to compare streams that were typical of each region. We used environmental criteria designated by the EPA to classifiy

streams as least-disturbed (Herlihy et al. 2008, U.S. EPA 2013) or highly-disturbed ("trashed"; ATH personal communication), but modified criterion thresholds because they varied slightly between the projects (see Table 2 in Herlihy et al. 2008 and Table B-2 in U.S. EPA 2013). For the current study and for each region, if any site exceeded any of the 90^{th} quantiles for TN, TP, CL, SO_4, turbidity, or % fine substrate (each $\log_{10}$ transformed) or if it failed an inorganic acidity criterion (ANC $\leq$ 0 ueq/L and DOC < 5 mg/L) then we classified it as a highly-disturbed (H) site. Few sites met the least-disturbed criteria using 10^{th} quantiles for each criteron, thus we classified the remaining sites as least-moderately (M) disturbed.

We used a suite of environmental variables, including spatial location, elevation, slope, stream order, riparian cover, channel geometry (wetted and bankfull), substrate composition, large woody debris volume, and various water chemical parameters (e.g., TN, TP, DOC, pH; see Appendix 1). Only candidate stream sites with a complete set of benthic and physicochemical data were retained for the analysis. The WSA dataset included a summary of watershed land-cover from the National Land Cover Dataset (NLCD); however, the NRSA did not. Thus, we derived local land-cover within a 1000 m radius of sample sites for all sites from the NLCD (2001 coverage for WSA, 2006 for NRSA).

Benthic metrics were available from both the WSA and NRSA datasets, although we recalculated metrics after standardizing benthic count data for consistency between datasets. For example, we grouped genera that were difficult to differentiate when small (e.g., *Orthocladius* and *Cricotopus* spp. as a group, all baetids at family level; see Stribling et al. 2008). In addition, the NRSA benthic count matrix contained some individuals resolved at the family level; thus, we distributed family among genera according to relative proportions of genera within that family at each site (Cuffney et al. 2007). Lastly, we summarized some taxa at the family (e.g., mollusks)

or order level (e.g., oligochaetes) and excluded some non-insect groups from analyses (e.g., mites).

We used information for 32 macroinvertebrate trait states derived from the USGS macroinvertebrate trait database (Vieira et al. 2006; see Appendix 2). Traits represented a series of life history attributes (body size, shape, mode of respiration, etc.), behavioral habits (e.g., clinger) and ecological habitat preferences (e.g., velocity, oxygen). Invertebrates were usually identified to genus and trait data was appropriately summarized at this level. Genera unrepresented for all states within a trait were supplemented with family-level values to maximize information (Sokol et al. 2011). Categorical trait records were converted to binary trait states where necessary; this binary information was then summed for each taxon and the state with the highest representation was given a value of 1 (else 0) for that taxon (Vieira et al. 2006). If 2+ trait states shared highest representation, each state was equally represented (e.g., 0.5, 0.5) within a trait category (Vieira et al. 2006). We calculated abundance weighted averaged trait values for each site and trait category as $\sum_{i=1}^{S} p_i x_i$, where p_i is the proportion of the ith species (or other taxonomic level), and x_i is its corresponding trait state value (Vandewalle et al. 2010). We focused mainly on trait states (hereafter "traits") that allow for *a priori* predictions, as they may provide resistance (RST) or resilience (RSL) to environmental stressors associated with LULC conversion (Table 1). RST traits, including morphological attachment and/or streamlined body shape may lead to disproportionate survival following high flow disturbance events. Other traits may offer RST to sedimentation or elevated temperatures associated with LULC conversion. In addition, strong dispersal abilities, high reproductive output and other RSL traits may allow for faster or disproportionately greater recolonization following disturbance (Townsend and Hildrew 1994).

Taxonomic differences likely exist between regions; therefore, we also calculated taxonomic richness (S) and Shannon's diversity (*H'*) at the genus and family levels and richness and percentages of taxa in the purportedly sensitive/intolerant aquatic insect orders Ephemeroptera, Plecoptera, and Trichoptera (EPT). We also calculated metrics based on specific taxonomic groups, including % Chironomidae, Odonata, Coleoptera and the percent of non-insect invertebrates (e.g., amphipods, isopods). We calculated assemblage-averaged pollution tolerance values (PTV), which were derived for the WSA and NRSA projects based on combined regional lists (U.S. EPA 2013). A simultaneous set of rarefied taxonomic- and trait-based metrics was created by repeatedly and randomly drawing 100 individuals without replacement from each site (reported as mean of 100 iterations) to minimize the influence of initial sample densities and rare taxa (Walker et al. 2008, Bêche and Statzner 2009).

5.2.2 Statistical Analysis

Because the CPL covers a much larger land-area, we were concerned with the possibility that CPL assemblages would be more dissimilar to each other than assemblages in either the APL or PMT regions. We used average pairwise Bray-Curtis dissimilarities (Bray and Curtis 1957) to assess within- and between-region taxonomic differences at the family and generic levels.

We were specifically interested if 1) M sites differed in central tendencies of benthic metrics (e.g., diversity, trait values) between regions, and 2) if differences existed between M and H sites within each region. We assessed differences in medians of metric distributions graphically using boxplots and statistically with 1-way Kruskal-Wallis (Hollander and Wolfe 1999) with groupings of region-M/H (e.g., APL M vs. APL H). Following a significant global test ($p < 0.05$), planned comparisons between regions (M sites only), and between M and H

within each region were assessed using individual Wilcoxon rank sum tests (Hollander and Wolfe 1999) with adjusted p-values using the sequential Holm-Bonferroni method (Holm 1979) to control family-wise error rates (Ruxton and Beauchamp 2008). Non-parametric rank sum tests have high relative efficiency (~ 95%) to parametric t-tests (Hollander anfd Wolfe 1999) and were chosen because some of the distributions were observed to be skewed in boxplots (Fig. 4).

For descriptive purposes and to informally assess the distinctiveness of regions in our study, we used principal component analysis (PCA) by performing singular value decomposition on Pearson product-moment transformed ($\sin[tau0.5\pi]$) Kendall (tau) correlation matrices (Moran 1948, Visuri et al. 2000, Syms 2008). We used PCA on a set of environmental variables and on assemblage-averaged trait values to graphically display sites in distance biplots in both environmental and separately in trait spaces. We used circle of correlation plots to aid in interpretation of PCA models, which show the correlation between the original variables and the derived component axes (González et al. 2012). Angles (θ) between variable vectors and ordination axes represent their correlations (correlation = $\cos\theta$), such that oppositely positioned vectors represent negative correlations and 90° angles represent orthogonality (Syms 2008, González et al. 2012). Vectors positioned closer to the outer circle (radius 1) have strong relationships with the displayed axes and can be directly interpreted, whereas those closer to the origin are weakly related and interpretation should be made with caution (Syms 2008, González et al. 2012). Circles centered at the origin with radii of 0.5 and 1 are included to help assess the contribution/importance of each variable to the displayed axes (Abdi and Williams 2010).

We used indicator species analysis (ISA; Dufrêne and Legendre 1997) to determine what taxa at the genus level were indicative of M sites within each region. Indicator values (IndV) for genus i and site group j are calculated as a function of both the relative abundance of genus i in

site group j, and the number of occurrences (presence/absence) of genus i in site group j (De Cáceres and Legendre 2009). Taxa with large IndV for a site group indicate that they are generally ubiquitous to those sites and occur in relatively high proportions. In addition, we calculated several resistance (RST) metrics for flow, sediment, and thermal stressors and a general resilience (RSL) metric. We created RST/RSL metrics as combinations of 2 traits picked specifically for each stressor. For each metric, a taxon was assigned a value of 1 if it was characterized by at least one of the specified traits (else 0). RST traits for flow stressors (FLO RST) included streamlined/flattened body shape and flow adaptations (e.g., hooks, silk). RST traits for sedimentation (SED RST) included burrowing habit and tolerance of silty/turbid water. RST traits for elevated water temperature (TMP RST) included tolerance of high temperatures and tolerance of low dissolved oxygen (DO). General RSL traits (GEN RSL) were high fecundity and multivoltinism (> 1 generation/y).

We assessed relationships between selected traits and environmental variables with partial least squares (PLS) regression, which displays much lower estimation variance than ordinary least-squares regression with collinear data (Dahlgren, 2010; Chapter 1). We used rank-based (Kendall) PLS, which provides similar results to standard PLS (relative efficiency ≅ 95%), while also being robust to random outliers or entire outlying observations (Schneid, Chapter 4). The number of components to retain was chosen as that having the smallest average RSS from 100 repeated random data hold-out (2/3) cross-validation. We used PLS as a constrained ordination to display relationships between traits and environmental variables, and displayed distance biplots and circle of correlation plots (González et al. 2012). We also calculated PLS variable importance in projection (VIP) scores as an indication of importance of each predictor variable to the overall model (Mehmood et al. 2012). Our goal was not to build the best

predicting model, and to reduce circularity, we excluded region-indicative but potentially important variables (e.g., slope, geographic coordinates) from descriptive models.

We used the nonparametric permutational MANOVA to test for significant among region differences in multivariate trait centroids (Anderson 2001). If regions were different in terms of traits, we assessed distinctiveness of benthic trait composition within each region by examining ability of traits to classify region membership using PLS followed by linear discriminate analysis (PLS-LDA). This 2-step procedure allows for highly correlated explanatory variables to be used with LDA by guiding dimension reduction prior to classification such that derived PLS latent variables maximally explain group membership (Barker and Rayens 2003). We created an initial PLS model with region indicator variables as response variables and a suite of traits (Appendix 2) used as explanatory variables. LDA then used derived orthogonal PLS variables to create classification rules that best discriminated regional groups. Last, we used cross-validation to determine the number of PLS components required to maximize classification ability.

Some environmental variables were log or square-root transformed to improve skewness (Appendix 1). We conducted all analyses in R (R Core Team 2015). Permutational MANOVA and Bray-Curtis dissimilarties were calculated using the package vegan (Oksanen et al. 2015), and ISA was performed with the indicspecies (De Caceres and Legendre 2009) package. We modified PCA and PLS algorithms found in vegan (Oksanen et al. 2015) and plsgenomics (Boulesteix et al. 2012) packages to incorporate rank-based correlation and covariance matrices.

5.3 Results

5.3.1 General site information

We compared randomly selected sites across 3 SE ecoregions and between 2 general disturbance categories (M/H) within each region. After pre-selection (e.g., dry sites removed),

127 APL (64 M, 63 H), 119 PMT (69 M, 50 H) and 154 CPL (92 M, 62 H) sites were retained (Fig. 1). M/H classes were created for each region based on several abiotic variables (filters, see methods), thus H sites tended to have greater values of filters than M sites (Fig. 2). Trends of general increasing in TP, TN, turbidity and DOC also were observed from highlands (APL and PMT) to lowlands (CPL) in M sites (Fig. 2).

Stream sites were similar in terms of average values for land-cover and stream size. Means (SD) of % land-cover in a 1000m radius of sample sites were 0.64 (0.27) for forest (included wetlands and grasslands), 0.21 (0.22) for agriculture (Ag), and 0.14 (0.17) for urban (and developed land). Sites were constrained to those < 5th order; the median order was 2 within each of the regions. Mean channel dimensions were 2.20-m (1.28) for stream widths and 2.80-m (1.13) for stream depths.

Results from a PCA of 37 environmental variables indicated that the 1st principal component (PC1) explained 20.72% of the total variation and described general gradients in substrate size and flow and PC2 (17.31%) a gradient of LULC disturbance. PC1 was most heavily loaded by variables describing substrate size (-), TSS (+) and turbidity (+), stream velocity (-), relative bed stability (RBS, +; Fig. 3). In contrast, water chemistry variables (-) and % forest cover (+) were most highly loaded on PC2. Biplots of site order along PC1 and PC2 indicated general separation between CPL and APL or PMT sites in this reduced environmental space (Fig. 3). The 3rd PC axis (not shown) explained an additional 11% of variation in environmental variables and showed similar shift between CPL and highland (APL and PMT) sites. M/H class information was not implicitly used in this PCA; however, the shift in the distributions of H sites observed within each region relative to M sites (Fig. 3) should be expected as it included variables used to create M/H classes.

5.3.2 Patterns for moderately disturbed (M) sites

Although the CPL region covers a much larger area of the SE US than either the APL or PMT, median values of within-region Bray-Curtis dissimilarities for M sites were only sightly higher for the CPL than for the APL or PMT with family-level or genus-level taxonomic resolution (≅ 5-6% higher; Table 2). Among-region dissimilarities were not much higher than within-region dissimilarities, with APL vs. PMT being more similar than either APL vs. CPL or PMT vs. CPL for family and genus-level resolution respectively (Table 2). The highest dissimilarity was between APL and CPL at the genus level (0.805, Table 2).

Within M sites, CPL differed from highlands (APL or PMT) for many metrics (Fig. 4, shaded boxes). CPL had significantly lower median values for rarefied genus and family-level richness and Shannon's diversity (*H'*) than APL or PMT, which were similar to each other in richness values (Fig. 4). CPL also had significantly lower EPT richness, proportions of EPT, and lower Coleopterans than either APL or PMT (Fig. 4). Conversely, CPL had higher % non-insects (e.g., Isopoda, Amphipoda, Oligochaeta) than APL or PMT (Fig. 4). CPL also had higher median proportion of Chironomidae (38%) than APL (27%); however, PMT (35%) was not statistically distinguishable from either APL or CPL (Fig. 4). Other major taxonomic groups, including the orders Odonata and Diptera (excluding Chironomidae) were low in numbers and similar between regions within M sites (Fig. 4).

ISA indentified taxa that were both highly abundant and ubiquitous at sites within each SE region (Table 3). The top 12 and 25 taxa in each region (largest IndV) were examined; taxa in the top 12 and 25 had *p*-values ≤ 0.001 and 0.040, respectively. The number of genera (462) considered resulted in general non-significance after adjustment for multiple testing (19 total had adjusted $p \leq 0.05$); however, we proceeded with this analysis as our goal was simply to describe

traits of taxa with highest affinities to each region. Individuals of the family Chironomidae (Diptera) composed 48% of the top 25 TolV taxa in the CPL compared to only 32% for PMT and 20 % for APL. Conversely, EPT taxa composed only 8% of the top 25 CPL taxa compared to 36% for PMT and 52% for APL. We examined the proportion of taxa with resistance/resilience traits among these top indicator taxa. The proportion of individuals with traits imparting resistance to flow disturbances (FLO RST) was higher in the APL and PMT (84, 76%) than in the CPL (52%; Table 3). In contrast, trends of increasing resistance or resilience were observed from highland to lowland sites (APL < PMT < CPL) for sediment resistance (SED RST: 48, 64, 92%), high temperature/low DO resistance (TDO RST: 44, 60, 92%) and resilience by high reproductive (REP RSL: 60, 80, 96%) or dispersal potential (REP RSL: 92, 96, 100%; Table 3).

Permutational MANOVA indicated multivariate differences existed between at least 2 study regions based on a suite of traits ($p < 0.001$; Appendix 2). Using 3 retained PLS components, a PLS-LDA model correctly classified CPL sites 86% of the time (average of 100 ¾ hold-out runs), whereas APL and PMT regions were only correctly classified 55 and 30% of the time, respectively. These results indicate that CPL has a relatively distinctive trait structure because of the disproportionately higher ability to distinguish CPL sites from highland (APL and PMT) sites based on trait data. The 1st discriminant function (LDA) explained 94% of between-class variance. The 1st PLS component (input to LDA) had a much larger weight (-0.46) associated with the 1st discriminant than either PLS component 2 (0.28) or 3 (-0.22). Graphical results from the PLS step (not provided) showed relatively large separation along the 1st model axis, with most CPL sites negatively positioned and highland sites positioned positively on PLS 1. According to loading magnitudes and VIP values (> 1 for PLS axis 1), the traits streamlined-flattened body shape (PLS1 loading = 0.19), flow adaptation (0.20), rheophily (0.32), clinger

(0.30) and burrower (-0.25) habits, large (0.32) and small (-0.29) substrate size preferences, multivoltinism (-0.17), high larval dispersal ability (-0.23), and low O_2 tolerance (-0.20) contributed the most to discrimination of lowlands (CPL) from highlands.

Several univariate differences existed in assemblage-averaged trait values within M sites among regions (shaded boxes, Fig. 5). Most strikingly, regional median values for streamlined body shape, flow adaptations, % clingers and % rheophiles decreased from highlands (APL and PMT) to lowlands (CPL). Conversely, median values for % burrowers, low O_2 tolerance, organic pollution tolerance, and multivoltinism increased from highlands (APL) to lowlands (CPL) (Fig. 5). Median proportions of individuals with generally decreased from highlands to lowlands, whereas other traits (e.g., high temperature tolerance) were similar across regions (Fig. 5). For EPT taxa, CPL sites had a greater median proportion of individuals that were small bodied, multivoltine, and low O_2 and pollution tolerant than highland regions (details excluded for brevity); although CPL showed had much lower values for % EPT and EPT richness (13%, 3, respectively) than PMT (29%, 7) or APL (34%, 8).

5.3.3 Moderately (M) vs. highly (H) disturbed sites

Many differences existed between M and H sites within each region. PMT and CPL had lower genus- and family-level richness and diversity (*H'*) in H compared to M sites; similar trends occured in APL sites (Fig. 4) but differences were not detectable. M sites showed higher EPT richness on average compared to H sites for each region (Fig. 4). Significantly lower % EPT also was observed in H sites relative to M sites in PMT; a similar, but non-significant trend occured in APL, whereas there was no difference in CPL (Fig. 4). The proportion of macroinvertebrate taxa as non-insects was higher in H vs. M sites in the highlands (APL and

PMT), but not in the lowland CPL. Suprisingly, additional taxonomic differences were not observed between M and H sites (e.g., % Dipteran, Fig. 4).

A few traits displayed differences between M and H sites (Fig. 5). Proportion of clingers was lower in H than M sites in the APL and PMT, whereas proportions of burrowers were generally higher in H than M sites in the APL and CPL. Similar, but weak trends occurred for clingers in CPL and burrowers in PMT, but differences were not detectable. Median proportion of individuals with morphological adaptations for high flow was lower in H vs. M sites in PMT and CPL, but not in APL (Fig. 5). In addition, there was a general trend for lower median proportion of streamlined-fusiform individuals in H relative to M sites the highlands (Fig. 5, APL difference not detected). Conversely, hydrodynamic body shapes were more prevalent in H relative to M sites in CPL, which also showed more sprawlers in H than M (Fig. 5). Last, H sites showed higher median assemblage-level pollution tolerance values than M sites in each of region (Fig. 5).

A PCA of 31 traits was performed to informally assess distinctness of each region in multiple assemblage-level traits, as well as associations between trait compositions. The first 2 PC axes explained 53% of the total variation on traits. A distance biplot of sites in trait space showed distinct separation between highland and CPL sites along the major axis of variation (PC1, Fig 6). M/H class information was not incorporated into this PCA model of traits, although distributions of H sites were slightly shifted to the positive end of PC1 and negative end of PC2 for each region (similar to that seen on the abiotic PCA, Fig. 6). PC1 was most highly positively loaded (eigenvectors) by low O_2 tolerance (0.37), preference for small substrate (0.34), and burrowing taxa (0.27) and negatively loaded by clinger taxa (-0.35), ability to cement/adhere eggs (-0.25) and/or to diapause (-0.23), preferences for fast water (-0.27), riffle habitat (-0.25),

and/or large substrate (-0.34) (Fig. 6). In contrast, PC2 was most highly positively loaded by small-bodied taxa (0.32), taxa with a propensity to drift (0.26), and collector gatherer taxa (0.28), and negatively loaded by taxa with long adult life (-0.30), that are scrapers (-0.30), are large (-0.29) and/or hardshelled (-0.40; Fig. 6).

A multi-response PLS regression model of resistance and resilience (RST/RSL) trait combinations and site averaged pollution tolerance values was performed to describe relationships between these traits and environmental variables in streams across the SE and in both impact designations (M/H). An initial PLS model was performed and variables were excluded in the final model with VIP < 1.0. Two PLS components were retained and these axes explained between 13 and 48 % of the variability of each response variable (Fig. 7). Traits were explained mainly by the 1st PLS axis, which was highly loaded by environmental predictors describing sediment size, water velocity and bed stability (Fig. 7). TOL VAL, SED RST, DSP RSL and TDO RST were each positively correlated with PLS 1, which was most highly loaded by TSS, turbidity, % fines, % slow water, DOC and TP. FLO RST was diametrically opposed on PLS1, which was which was most associated with % fast water, larger and more stable substrate (RBS) and erodible substrate diameter (Dmm). Site positions in a PLS distance plot indicate general separation between CPL and highland regions, and also a slight shift between M and H sites for each region in the same directions in reduced space (positive in PLS1 and PLS2 directions, Fig. 7). This M-H sift corresponded with axes generally describing increased fine sediments, TSS and turbidity (PLS1) and decreased local % forest and increased channel size (BKFw, INCh) and concentrations of several chemical constituents (e.g., SPC, pH, Fig. 7).

5.4 Discussion

Benthic macroinvertebrate assemblages have been shown to respond to watershed urbanization to a lesser degree in the CPL relative to highland regions of the SE and it has been suggested that lowland assemblages are composed of more resistant (RST) and/or resilient (RSL) taxa compared to highland regions (Utz et al. 2009, Nagy et al. 2011, Utz et al. 2011, Utz and Hilderbrand 2011). We examined data from 2 large scale EPA data sets to determine the degree to which lowland CPL region differs from highland APL and PMT regions in invertebrate traits, which have been useful in characterizing stream conditions (Horrigan and Baird 2008, Pollard and Yuan 2010). Our study provided supporting evidence for the general distinctiveness of CPL assemblages in relatively undisturbed sites relative to highlands in the composition of multiple traits and indicated higher prevalence of resistance traits that may be beneficial in coping with adverse conditions (e.g., increased fine sediments, decreased DO). In addition, multivoltinism (> 1 generation/y) showed trends of increasing proportion in assemblages from highlands to lowlands (APL < PMT < CPL), which may be an important factor for recolonization rates following disturbance. In general, our results are in agreement with Utz and Hilderbrand (2011) and Utz et al. (2011), who suggested that CPL assemblages may be more adapted for hyporheic habitat and shifting (unstable) sediments and generally less sensitive to stressors associated with land cover change.

Harsh environmental conditions can "filter" poorly adapted species from any given site (Poff 1997); thus, it seems logical to assume that trait filtering has led to regionally distinct trait composition among SE regions as a result of strong contrasts in environmental conditions. Our results strongly suggest distinct trait differences between the CPL and highland regions that appear to coincide with contrasting environmental conditions observed across these regions. Highland (APL and PMT) streams tended to have fast-flowing (likely well-oxygenated) waters

and were dominated by relatively large, stable bed particles, and assemblages were correspondingly dominated by rheophilic taxa with adaptations and habits that allow them to anchor to stable substrates and persist in fast-flow. Alternatively, lowland CPL streams were predominantly slow-moving (likely oxygen-depleted) with small and unstable substrate, and assemblages dominated by taxa with an affinity for small substrates, a burrowing lifestyle, and tolerance of low dissolved oxygen levels.

Previous studies have suggested that magnitudes of hydrologic alterations associated with urbanization (e.g., flood frequency, magnitude, and flashiness) are greater in high-gradient than lower-gradient streams (Brown et al. 2009; Utz et al. 2011). Such geographic contrasts in hydrologic response to urbanization may translate to interregional differences in influence on abiotic and biotic components of receiving streams. In SE highlands, urban sites often show a higher proportion of fine sediments relative to non-urban streams in some studies (Walters et al. 2003, Freeman and Schorr 2004); however, the opposite pattern also has been observed (Helms et al. 2009, Utz and Hilderbrand 2011). Utz and Hilderbrand (2011) found that CPL urban streams were not different in sediment size (measured at ≥ D50) than rural CPL streams and particles were more mobile in CPL streams than PMT streams in both rural and urban streams. The conflicting results regarding fine sediments and urbanization in highland streams may be from differences in the age of the urban development across sites and studies, as sediment delivery slows and erosional forces are thought to dominate at some point after construction activities have ceased (Nagy et al. 2011).

The fact that CPL benthic assemblages consisted of more sediment-RST taxa, and highland assemblagess with more flow-disturbance RST taxa should not be suprising. However, the lower magnitude of benthic responses to urbanization observed in CPL relative to highland

regions is likely because of the prevalence of natural adaptations to survive in unstable, sandy streams. Urban-related increases in fine sediments are likely more important in highlands relative to CPLs, where dominant sediment sizes are naturally small (Utz and Hilderbrand 2011). Highland taxa evolved to inhabit streams with large and relatively stable substrate and may generally lack traits that facilitate survival of sediment accumulation associated with agriculture and (initial phases) of watershed urbanization. In addition, the magnitude of hydrologic change is lower in the CPL along a gradient of LULC change (Nagy et al. 2011), and a lack of flow RST traits in CPL assemblages may be inconsequential for taxa in this region.

Within each region, there were biotic differences characterizing M and H sites; most notably were generally lower richness/diversity, fewer clingers but more burrowers, and higher tolerance values in H vs. M sites (Figs. 4 and 5). As expected, a general shift was seen between M and H sites in reduced PC space for environmental variables (Fig. 3); interestingly, this same pattern occurred separately for multiple assemblage-averaged traits (Fig. 6). The M-H shifts observed in the environmental ordination (PCA) and constrained ordination of covariance in trait categories and environmental variables (PLS) was similarly observed in an unconstrained (no environmental input) trait ordination. These shifts are expected in environmental-based ordinations as some of these data (e.g., % fines, turbidity, TP, CL) were used in creation of the M/H classes and resulting gradients. Interestingly, the unconstrained PCA trait model indicated a similar M-H shift along an axis describing a gradient of clingers/rheophiles to burrowing/low O_2 tolerant taxa (Fig. 6). H sites were generally positioned further toward directions of increased sedimentation (environmental variables, Fig. 3) or burrowing-dominated assemblages (traits, Fig. 6) along their respective major axes (1^{st} PCs), or shifted toward greater land-cover disturbance

and elevated chemistries (environmental variables) or toward the direction of larger, more armored individuals with long adult lifespan that can exit the water (2^{nd} PCs).

The use of M and H classes are far from ideal, as they represent multiple LULC stressors - not a single stressor gradient - and also lacked detailed hydrologic data. Historical agriculture and channel alterations were so widespread in the Southeast that unimpacted streams may not exist, and current conditions may be influenced by historical land use (Maloney et al. 2008, Hardison et al. 2009). Likely for this reason, least-disturbed sites were relatively uncommon these probabilistic surveys (randomly sampled sites; Stoddard et al. 2008). Our M/H designations were based on the WSA/NRSA reference screening process, and were intended to indicate and separate highly disturbed H sites from less disturbed M sites. Detailed LULC data were not derived for these data as a whole. However it is important to point out that local (1000m) estimates of LULC % for WSA sites were highly significantly, but weakly correlated with entire watershed values ($r \cong 0.40$); however, mean local % Urb values were consistently higher than that for whole watershed estimates; therefore, we are somewhat confident that our M sites were relatively low in urban LULC (generally < 20% summed NLCD developed classes). Because least-disturbed sites are generally uncommon, it could be argued that this M classification may be more representative of a "pre-urban state" within these regions.

Use of taxonomic-based indices (e.g., richness, diversity, or composition) may not be generalizable across regions in respect to stressor gradients due to spatial turnover in species; however, traits common within a taxonomic group (e.g., benthic invertebrates in general) may serve as more consistent indicators of anthropogenic impact (Dolédec et al. 1999, Charvet et al. 2000, Bonada et al. 2006, Culp et al. 2010). Individuals with adaptive "resistance" traits (e.g., small body size, streamlined morphology) may disproportionately survive disturbance events, or

those with “resilience” traits (e.g., short generation times, large clutch sizes, strong dispersal ability) may re-colonize faster or in disproportionately higher numbers following disturbance (Townsend and Hildrew 1994); thus, some traits may have the added benefit of being potentially predictable due to mechanistic-based relationships. Trait-based analyses can be complicated by several factors; for example, species are characterized by a suite of non-independent traits that may be highly interrelated (confounding interpretation) and/or linked to its evolutionary history (taxonomic and trait variation coincide) (Townsend and Hildrew 1994, Poff et al. 2006). Some traits, including adult size, voltinism, and habit are considered to be less dependent on taxonomy and responsive to local selection, therefore potentially useful for biomonitoring purposes (Poff et al. 2006, Horrigan and Baird 2008).

Across the US, the invertebrate habit of clinging to stable substrates was a more consistent and mechanistically linked indicator of sedimentation than traditional indicators based on sensitive taxa (EPT; Pollard and Yuan 2010). The proportion of a trait-state represented within an assemblage is directly linked to the proportion of other states within that trait group (e.g., clinger and burrower habits), but these states are also potentially influenced by one or more of the same stressors. For example, fine sediment deposition can bury large/stable bed particles and lead to a loss of individuals who cling to stable substrates during extreme flow events; conversely, fine sediments provide suitable habitat for borrowing individuals (Pollard and Yuan 2010, Monaghan and Soares 2012). In our study, graphical representation of assemblage-averaged trait values in reduced multivariate space (PCA) allowed for an assessment of trait correlations in relation to the major axes of variation in those traits. Correlations naturally existed in these data between traits that may be considered effective resistance traits for dealing

with flow-related disturbances and those likely to be beneficial in high sediment environments are problematic in terms of the use of these traits in a mechanistic (causal) framework.

It is also possible that a single effective trait may ensure individual success in an extreme environment without the need for additional traits; and, unless an assemblage collectively solves environmental problems in a related way, detection of clear trends between assemblage-level traits and environmental conditions will be unlikely (Townsend and Hildrew 1994). These complications combined with the correlated nature of traits suggest that additional research into whether trait combinations, tailored multimetrics (Monaghan and Soares 2012) or metrics based on trait syndromes/strategies (Verberk et al. 2008) more consistently respond to known environmental gradients, and are therefore better suited for bioassessment purposes than individual traits. Our results suggest CPL assemblages may be well equipped to deal with certain stressors associated with LULC change (e.g., sedimentation), although the questions remain of what aspects of CPL assemblage structure, or what summary metrics are consistent indicators of urban LULC change in this lowland region.

Suprisingly, we did not observe any differences in the relative proportions of high temperature tolerant taxa among regions in these data (Fig. 4). Differences in urban water temperature responses have been observed between regions in the SE (Utz et al. 2011). Cooler, higher-gradient, streams may be affected to a greater degree by temperature increases associated with urbanization, compared to relatively benign heat additions to naturally warmer low-gradient coastal plains streams (Utz et al. 2011). In addition, biota of higher-gradient streams may be adapted to natural cool-water conditions, making them relatively more susceptible to LULC shifts than warm-water coastal plains species (Utz et al. 2011). While it is possible that differences in (assemblage-averaged) warm water tolerances do not exist, it may be that the

currently available trait data on thermal tolerances is not sufficient at the resolution we used (genus level, supplemented with family level). For example, few taxa were represented by high temperature tolerance (i.e. "warm eurythermal") in a recent study in the western US on climate change that included warm streams (Poff et al. 2010).

The ecoregion concept was created to organize ecologically homogenous areas based on land form, soils, climate and vegetation, and provides the basis for hypotheses regarding how environmental variables affect biota (Gerritsen et al. 2000, Hawkins and Norris 2000, Johnson and Host 2010). The aggregate ecoregions in this study represented an environmental gradient from highlands to lowlands in terms of important environmental variables such as substrate size and flow regime (Fig. 3), both of which are known to be important in stream settings (Clausen and Biggs 1997, Jones et al. 2011). Our findings are in general agreement with Feminella (2000) who examined SE ecoregions and concluded that lowland streams were taxonomically distinct from highland streams, with the CPL region having lower levels of richness and diversity as well as lower proportions and richness of EPT taxa than the highland APL and PMT regions (Fig. 4). The general similarity observed between APL and PMT in this study agrees with Feminella (2000) who concluded that while the PMT is considered transitional zone between highland and lowland regions, both biologically and environmentally, it was more taxonomically similar to more mountainous Appalachians (APL) than lowland regions, which was unexpected as it is a subregion of Coastal Plains (CPL).

The size of the human population is expected to increase by 33% in the next 30 y (Alig et al. 2004, Cincotta and Gorenflo 2011), thereby likely leading to additional and substantial landscape alteration (Cincotta 2011). Urbanization (including urban sprawl) is one of the more pervasive forms of land-use/land-cover (hereafter LULC) change in the Southeast US

(O'Driscoll et al. 2010, Nagy et al. 2011), and this region is predicted to have the greatest regional growth in population and increase in land development in coming decades (Xian et al. 2012). Our findings emphasize the need for region-specific metrics and models for assessing the influence of land-cover on stream ecosystems, which may reduce variation attributable to natural differences in environmental conditions and reliance on biotic taxonomy on large scales and improve overall model performance and interpretability (Gerritsen et al. 2000). A more detailed analysis of data from these regions that includes watershed land-cover and estimates basic stream hydrology (baseflow magnitude) is warranted. Specifically, future research should focus on how these regions vary in terms of the relative influence of important environmental variables on individual traits, and whether trait-combinations may be better suited as indicators of anthropogenic disturbance in the SE.

5.5 References

Abdi, H. and L. J. Williams. 2010. Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics 2:433-459.

Alig, R. J., J. D. Kline, and M. Lichtenstein. 2004. Urbanization on the US landscape: looking ahead in the 21st century. Landscape and Urban Planning 69:219-234.

Allan, J. D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annual Review of Ecology, Evolution, and Systematics 35:257-284.

Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecology 26:32-46.

Barker, M. and W. Rayens. 2003. Partial least squares for discrimination. Chemometrics 17:166-173.

Bêche, L. A. and B. Statzner. 2009. Richness gradients of stream invertebrates across the USA: taxonomy-and trait-based approaches. Biodiversity and Conservation 18:3909-3930.

Booth, D. B. 1990. Stream channel incision following drainage basin urbanization. American Water Resources Association 26:407-417.

Boulesteix, A-L, Lambert-Lacroix, S., Peyre, J., and K. Strimmer. 2012. plsgenomics: PLS analyses for genomics. R package version 1.2-6. Available at: http://CRAN.R-project.org/package=plsgenomics.

Bray, J. R. and J. T. Curtis. 1957. An ordination of upland forest communities of southern Wisconsin. Ecological Monographs 27:325-349.

Brown, L., T. Cuffney, J. Coles, F. Fitzpatrick, G. McMahon, J. Steuer, A. Bell, and J. May. 2009. Urban streams across the USA: lessons learned from studies in 9 metropolitan areas. Journal of the North American Benthological Society 28:1051-1069.

Carter, T., C. R. Jackson, A. Rosemond, C. Pringle, D. Radcliffe, W. Tollner, J. Maerz, D. Leigh, and A. Trice. 2009. Beyond the urban gradient: barriers and opportunities for timely studies of urbanization effects on aquatic ecosystems. Journal of the North American Benthological Society 28:1038-1050.

Cincotta, R. P. 2011. The biological diversity that is humanly possible: three models relevant to human population's relationship with native species. Pages 61-72 in R. P. Cincotta and L. Gorenflo, editors. Human Population: Its influences on biodiversity. Springer, Berlin.

Cincotta, R. P. and L. Gorenflo. 2011. Introduction: Influences of Human Population on Biological Diversity. Pages 1-9 in R. P. Cincotta and L. Gorenflo, editors. Human Population: Its influences on biodiversity. Springer, Berlin.

Clausen, B. and B. Biggs. 1997. Relationships between benthic biota and hydrological indices in New Zealand streams. Freshwater Biology 38:327-342.

Cobb, D., T. Galloway, and J. Flannagan. 1992. Effects of discharge and substrate stability on density and species composition of stream insects. Canadian Journal of Fisheries and Aquatic Sciences 49:1788-1795.

Cordova, J. M., E. J. Rosi-Marshall, J. L. Tank, and G. A. Lamberti. 2008. Coarse particulate organic matter transport in low-gradient streams of the Upper Peninsula of Michigan. Journal of the North American Benthological Society 27:760-771.

Cuffney, T., M. Bilger, and A. Haigler. 2007. Ambiguous taxa: effects on the characterization and interpretation of invertebrate assemblages. Journal of the North American Benthological Society 26:286-307.

Culp, J. M., D. G. Armanini, M. J. Dunbar, J. M. Orlofske, N. L. R. Poff, A. I. Pollard, A. G. Yates, and G. C. Hose. 2010. Incorporating traits in aquatic biomonitoring to enhance

causal diagnosis and prediction. Integrated environmental assessment and management 7:187-197.

Dahlgren, J. P. 2010. Alternative regression methods are not considered in Murtaugh (2009) or by ecologists in general. Ecology Letters 13:E7-E9.

De Cáceres, M. and P. Legendre. 2009. Associations between species and groups of sites: indices and statistical inference. Ecology 90:3566-3574.

Dufrêne, M. and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs 67:345-366.

Elosegi, A., J. Díez, and M. Mutz. 2010. Effects of hydromorphological integrity on biodiversity and functioning of river ecosystems. Hydrobiologia 1:199-215.

Feeley, J. D. 1992. Medium-low-gradient streams of the Gulf Coastal Plain. Pages 233-269 in C. T. Hackney, S. M. Adams, and W. H. Martin, editors. Biodiversity of the southeastern United States--Aquatic Communities. Wiley and Sons, Inc, New York.

Feminella, J. W. 2000. Correspondence between stream macroinvertebrate assemblages and 4 ecoregions of the southeastern USA. Journal of the North American Benthological Society 19:442-461.

Freeman, P. L. and M. S. Schorr. 2004. Influence of watershed urbanization on fine sediment and macroinvertebrate assemblage characteristics in Tennessee Ridge and Valley Streams. Freshwater Ecology 19:353-362.

Gerritsen, J., M. T. Barbour, and K. King. 2000. Apples, oranges, and ecoregions: on determining pattern in aquatic assemblages. Journal of the North American Benthological Society 19:487-496.

González, I., K.-A. Lê Cao, M. J. Davis, and S. Déjean. 2012. Visualising associations between paired 'omics' data sets. BioData mining 5:1-23.

Hardison, E. C., M. A. O'Driscoll, J. P. DeLoatch, R. J. Howard, and M. M. Brinson. 2009. Urban Land Use, Channel Incision, and Water Table Decline Along Coastal Plain Streams, North Carolina. American Water Resources Association 45:1032-1046.

Hawkins, C. P. and R. H. Norris. 2000. Performance of different landscape classifications for aquatic bioassessments: introduction to the series. Journal of the North American Benthological Society 19:367-369.

Helms, B. S., J. E. Schoonover, and J. W. Feminella. 2009. Seasonal variability of landuse impacts on macroinvertebrate assemblages in streams of western Georgia, USA. Journal of the North American Benthological Society 28:991-1006.

Herlihy, A. T., S. G. Paulsen, J. V. Sickle, J. L. Stoddard, C. P. Hawkins, and L. L. Yuan. 2008. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. Journal of the North American Benthological Society 27:860-877.

Hollander, M. and D. A. Wolfe. 1999. Nonparametric statistical methods.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6:65-70.

Horrigan, N. and D. J. Baird. 2008. Trait patterns of aquatic insects across gradients of flow-related factors: a multivariate analysis of Canadian national data. Canadian Journal of Fisheries and Aquatic Sciences 65:670-680.

Johnson, L. B. and G. E. Host. 2010. Recent developments in landscape approaches for the study of aquatic ecosystems. Journal of the North American Benthological Society 29:41-66.

Jones, J., J. Murphy, A. Collins, D. Sear, P. Naden, and P. Armitage. 2011. The impact of fine sediment on macro-invertebrates. River Research and Applications 28:1055-1071.

Maloney, K. O., J. W. Feminella, R. M. Mitchell, S. A. Miller, P. J. Mulholland, and J. N. Houser. 2008. Landuse legacies and small streams: identifying relationships between historical land use and contemporary stream conditions. Journal of the North American Benthological Society 27:280-294.

Monaghan, K. A. and A. M. Soares. 2012. Bringing new knowledge to an old problem: Building a biotic index from lotic macroinvertebrate traits. Ecological Indicators 20:213-220.

Moran, P. 1948. Rank correlation and product-moment correlation. Biometrika 35:203-206.

Mulholland, P. J. and D. R. Lenat. 1992. Streams of the southeastern Piedmont, Atlantic drainage. Pages 193-232 in C. T. Hackney, S. M. Adams, and W. H. Martin, editors. Biodiversity of the southeastern United States: aquatic communities. John Wiley & Sons, Inc.

Nagy, R. C., B. G. Lockaby, B. Helms, L. Kalin, and D. Stoeckel. 2011. Water Resources and Land Use and Cover in a Humid Region: The Southeastern United States. Environmental Quality 40:867-878.

Nagy, R. C., B. G. Lockaby, L. Kalin, and C. Anderson. 2012. Effects of urbanization on stream hydrology and water quality: the Florida Gulf Coast. Hydrological Processes 26:2019–2030.

Nelson, K. C. and M. A. Palmer. 2007. Stream temperature surges under urbanization and climate change: data, models, and responses. American Water Resources Association 43:440-452.

O'Driscoll, M., S. Clinton, A. Jefferson, A. Manda, and S. McMillan. 2010. Urbanization effects on watershed hydrology and in-stream processes in the southern United States. Water 2:605-648.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., M., Stevens, M.H., Wagner, H. 2015. vegan: Community Ecology Package. R package version 2.2-1. http://CRAN.R-project.org/package=vegan

Paul, M. J. and J. L. Meyer. 2001. Streams in the urban landscape. Annual Review Ecological Systems 32:333-365.

Paulsen, S. G., A. Mayio, D. V. Peck, J. L. Stoddard, E. Tarquinio, S. M. Holdsworth, J. Van Sickle, L. L. Yuan, C. P. Hawkins, and A. T. Herlihy. 2008. Condition of stream ecosystems in the US: an overview of the first national assessment. Journal of the North American Benthological Society 27:812-821.

Poff, N., M. Pyne, B. Bledsoe, C. Cuhaciyan, and D. Carlisle. 2010. Developing linkages between species traits and multiscaled environmental variation to explore vulnerability of stream benthic communities to climate change. Journal of the North American Benthological Society 29:1441-1458.

Poff, N. L. R. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16:391-409.

Pollard, A. and L. Yuan. 2010. Assessing the consistency of response metrics of the invertebrate benthos: a comparison of trait- and identity-based measures. Freshwater Biology 55:1420-1429.

Rose, S. and N. E. Peters. 2001. Effects of urbanization on streamflow in the Atlanta area (Georgia, USA): a comparative hydrological approach. Hydrological Processes 15:1441-1457.

Ruxton, G. D. and G. Beauchamp. 2008. Time for some a priori thinking about post hoc testing. Behavioral Ecology 19:690-693.

Smock, L. A. and E. Gilinsky. 1992. Coastal Plain blackwater streams.in C. T. Hackney, S. M. Adams, and W. H. Martin, editors. Biodiversity of the southeastern United States--Aquatic Communities. Wiley and Sons, Inc., New York.

Sokol, E. R., E. Benfield, L. K. Belden, and H. M. Valett. 2011. The assembly of ecological communities inferred from taxonomic and functional composition. The American Naturalist 177:630-644.

Stoddard, J. L., A. T. Herlihy, D. V. Peck, R. M. Hughes, T. R. Whittier, and E. Tarquinio. 2008. A process for creating multimetric indices for large-scale aquatic surveys. Journal of the North American Benthological Society 27:878-891.

Strayer, D. L. 2006. Challenges for freshwater invertebrate conservation. Journal of the North American Benthological Society 25:271-287.

Stribling, J. B., K. L. Pavlik, S. M. Holdsworth, and E. W. Leppo. 2008. Data quality, performance, and uncertainty in taxonomic identification for biological assessments. Journal of the North American Benthological Society 27:906-919.

Syms, C. 2008. Encyclopedia of Ecology. Pages 2940-2949 in S. E. Jorgensen and B. Fath, editors. Elsevier, Oxford.

Townsend, C. R. and A. G. Hildrew. 1994. Species traits in relation to a habitat templet for river systems. Freshwater Biology 31:265-275.

Turner, M. G., S. M. Pearson, P. Bolstad, and D. N. Wear. 2003. Effects of land-cover change on spatial pattern of forest communities in the Southern Appalachian Mountains (USA). Landscape Ecology 18:449-464.

U.S. Environmental Protection Agency. 2013. National rivers and streams assessment 2008-2009: technical report (DRAFT). U.S. Environmental Protection Agency, Office of Wetlands, Oceans and Watersheds Office of Research and Development. Washington, DC, p 127.http://water.epa.gov/type/rsl/monitoring/riverssurvey/upload/NRSA0809_ Technical Report_130325_Web.pdf. Accessed 28 April 2015.

Utz, R., K. Eshleman, and R. Hilderbrand. 2011. Variation in physicochemical responses to urbanization in streams between two Mid-Atlantic physiographic regions. Ecological Applications 21:402-415.

Utz, R. M. and R. H. Hilderbrand. 2011. Interregional variation in urbanization-induced geomorphic change and macroinvertebrate habitat colonization in headwater streams. Journal of the North American Benthological Society 30:25-37.

Utz, R. M., R. H. Hilderbrand, and D. M. Boward. 2009. Identifying regional differences in threshold responses of aquatic invertebrates to land cover gradients. Ecological Indicators 9:556-567.

Vandewalle, M., F. De Bello, M. P. Berg, T. Bolger, S. Dolédec, F. Dubs, C. K. Feld, R. Harrington, P. A. Harrison, and S. Lavorel. 2010. Functional traits as indicators of biodiversity response to land use changes across ecosystems and organisms. Biodiversity and Conservation 19:2921-2947.

Verberk, W. C. E. P., H. Siepel, and H. Esselink. 2008. Life history strategies in freshwater macroinvertebrates. Freshwater Biology 53:1722-1738.

Vieira, N. K. M., N. L. Poff, D. M. Carlisle, S. R. Moulton, M. L. Koski, and B. C. Kondratieff. 2006. A database of lotic invertebrate traits for North America. US Geological Survey Data Series 187:1–15.

Visuri, S., V. Koivunen, and H. Oja. 2000. Sign and rank covariance matrices. Journal of Statistical Planning and Inference 91:557-575.

Walker, S. C., M. S. Poos, and D. A. Jackson. 2008. Functional rarefaction: estimating functional diversity from field data. Oikos 117:286-296.

Wallace, J. B., J. R. Webster, and R. L. Lowe. 1992. High-gradient streams of the Appalachians. Pages 133-191 in C. T. Hackney, S. M. Adams, and W. H. Martin, editors. Biodiversity of the southeastern United States--Aquatic Communities. Wiley and Sons, Inc, New York.

Walsh, C., A. Roy, J. Feminella, P. Cottingham, P. Groffman, and R. Morgan. 2005. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24:706-723.

Walters, D., D. Leigh, and A. Bearden. 2003. Urbanization, sedimentation, and the homogenization of fish assemblages in the Etowah River Basin, USA. Hydrobiologia 494:5-10.

White, E. M., A. T. Morzillo, and R. J. Alig. 2008. Past and projected rural land conversion in the US at state, regional, and national levels. Landscape and Urban Planning 89:37-48.

Wilzbach, M. A. and K. W. Cummins. 2008. Rivers and streams: physical setting and adapted biota. Pages 3095-3106 in S. E. Jorgensen and B. D. Fath, editors. Encyclopedia of Ecology. Elsevier, Oxford.

Xian, G., C. Homer, B. Bunde, P. Danielson, J. Dewitz, J. Fry, and R. Pu. 2012. Quantifying urban land cover change between 2001 and 2006 in the Gulf of Mexico region. Geocarto International 1:1-19.

Zuellig, R. and T. Schmidt. 2012. Characterizing invertebrate traits in wadeable streams of the contiguous US: differences among ecoregions and land uses. Freshwater Science 31:1042-1056.

Table 1. Resistance (RST) and resilience (RSL) traits for several impact types associated with land-cover change. Predicted associations (↑ = increase, ↓ = decrease) between trait states and impact type and rationale for expected relationships are provided. †Small body size is listed as RSL trait due to association with r-strategy (to maximize population growth rate), but has been regarded as flow RST trait as well.

Impact type	RST /RSL	Trait state & prediction (↑/↓)	Rationale	Source
↑ Flow	RST	Streamlined-flat shape ↑	Resistance to high flow events	1, 3, 4
↑ Flow	RST	Attachment (e.g., hooks) ↑	Holdfast ability in high flow	1, 2, 3
↑ Flow	RST	Clinger habit ↑	Clings to stable substrates	5
↑ Flow	RST	Rheophily preference ↑	Fast water habitat tolerant	1
↑ Sediment	RST	Burrower habit ↑	Prefers fine sediments	3, 4, 6
↑ Sediment	RST	Silt/turbidity tolerance ↑	Tolerant of suspended sediments	1
↓ O_2	RST	Atm. Breather ↑	Independent of O_2 concentrations	3
↓ O_2	RST	Low O_2 tolerance ↑	Resistant to low dissolved O_2	1
↑ Temps	RST	High T tolerance ↑	Ability to withstand high T	1, 6
General	RSL	Multivoltine ↑	Short generation time	1, 2, 3
General	RSL	†Small body size ↑	Linked with short life span	1
General	RSL	High fecundity ↑	Many offspring for recolonization	1, 2, 3
General	RSL	High dispersal ↑	High recolonization potential	1, 2

Sources: 1 = Townsend and Hildrew (1994), 2 = Resh et al. (1994), 3 = Statzner et al. (2005), 4 = Statzner and Bêche (2010), 5 = Merritt et al. (2008), 6=Monaghanans and Soares (2012).

Table 2. Median (SD) values for pairwise Bray-Curtis dissimilarities within least-moderately disturbed (M) sites at the family and genus levels; comparisons are within (e.g., APL-APL) and between (e.g., APL-PMT) aggregate ecoregions: Appalachain (APL), Piedmont (PMT) and coastal plains (CPL). Standard deviations for dissimilarities are in parentheses. Taxa represented in less than 5% of the sites were excluded.

Level	APL-APL	PMT-PMT	CPL-CPL	APL-PMT	APL-CPL	PMT-CPL
Family	0.602	0.556	0.618	0.587	0.685	0.636
	(0.112)	(0.106)	(0.113)	(0.106)	(0.112)	(0.120)
Genus	0.730	0.681	0.733	0.716	0.805	0.758
	(0.106)	(0.097)	(0.105)	(0.099)	(0.101)	(0.107)

Table 3. The top indicator taxa and values (IndV) for moderately disturbed (M) sites in the three SE ecoregions (Reg.): Appalachain (APL), Piedmont (PMT) and coastal plains (CPL). Taxa were ordered for each Reg. by decreasing IndV statistics. Binary indicators for traits that may provide resistance (RST) or resilience (RSL) to extreme flow events (FLO), sedimentation (SED). A value of 1 is given if that taxa was represented by at least one trait in each category. The proportion of taxa (top 25 IndV) represented for each trait category is given in the terminal row for each Reg. To save space, "ae" was removed from Family names and Order names abbreviated: Amph = Amphipoda, Clpt = Coleoptera, Dptr = Diptera, Ephm = Ephemeroptera, Hplt = Haplotaxida, Ispd = Isopoda, Ntng = Neotaenioglossa, Plcp = Plecoptera, Trch = Trichoptera, Vnrd = Veneroida.

Reg.	IndV	Order	Family	Taxa	FLO RST	SED RST	TDO RST	REP RSL	DSP RSL
APL	0.628	Ephm	Baetid	Baetidae	1	0	0	1	0
APL	0.606	Clpt	Elmid	Optioservus	1	0	0	1	1
APL	0.586	Dptr	Chironomid	Parametriocnemus	0	1	1	1	1
APL	0.556	Dptr	Chironomid	Micropsectra	0	1	1	1	1
APL	0.534	Clpt	Elmid	Oulimnius	1	0	0	1	1
APL	0.528	Dptr	Chironomid	Eukiefferiella	0	0	0	1	1
APL	0.520	Ephm	Ephemerellid	Eurylophella	1	0	0	1	1
APL	0.514	Dptr	Chironomid	Cricotopus	1	0	1	1	1
APL	0.509	Odnt	Gomphid	Gomphidae	1	1	1	0	0
APL	0.505	Trch	Hydropsychid	Diplectrona	1	0	0	1	1
APL	0.485	Ephm	Hepageniid	Leucrocuta	1	1	1	1	1
APL	0.479	Mglp	Corydalid	Nigronia	1	1	0	0	1
APL	---	---	---	---	84%	48%	44%	60%	92%
PMT	0.758	Trch	Hydropsychid	Cheumatopsyche	1	0	0	1	1
PMT	0.577	Dptr	Chironomid	Microtendipes	1	1	1	1	1
PMT	0.549	Trch	Hydropsychid	Hydropsyche	1	1	1	0	1
PMT	0.524	Dptr	Chironomid	Cladotanytarsus	1	1	1	1	1
PMT	0.482	Clpt	Elmid	Dubiraphia	1	1	0	1	1
PMT	0.480	Ephm	Baetiscid	Baetisca	1	1	1	0	1
PMT	0.470	Clpt	Elmid	Macronychus	1	0	0	1	1
PMT	0.465	Trch	Hydroptilid	Hydroptila	1	0	0	1	1
PMT	0.447	Dptr	Empidid	Hemerodromia	0	1	1	1	1
PMT	0.442	Dptr	Chironomid	Corynoneura	0	1	1	1	1
PMT	0.420	Vnrd	Corbiculid	Corbiculidae	1	1	1	1	1
PMT	0.408	Dptr	Chironomid	Parakiefferiella	0	1	1	1	1
PMT	---	---	---	---	76%	64%	60%	80%	96%
CPL	0.636	Dptr	Chironomid	Dicrotendipes	0	1	1	1	1
CPL	0.610	Dptr	Chironomid	Ablabesmyia	0	1	1	1	1
CPL	0.575	Dptr	Ceratopogonid	Ceratopogonidae	0	1	1	1	1
CPL	0.575	Ntng	Hydrobiid	Hydrobiidae	1	1	0	1	1
CPL	0.564	Dptr	Chironomid	Cryptochironomus	0	1	1	1	1
CPL	0.561	Hplt	Naidid	Naididae	0	1	1	1	1
CPL	0.550	Ephm	Caenid	Caenis	1	1	1	1	1
CPL	0.517	Dptr	Chironomid	Chironomus	1	1	1	0	1
CPL	0.486	Trch	Leptocerid	Oecetis	1	0	1	1	1
CPL	0.478	Dptr	Chironomid	Labrundinia	0	1	1	1	1
CPL	0.477	Amph	Hyalellid	Hyalella	1	1	1	1	1
CPL	0.475	Dptr	Chironomid	Clinotanypus	0	1	1	1	1
CPL	---	---	---	---	52%	92%	92%	96%	100%

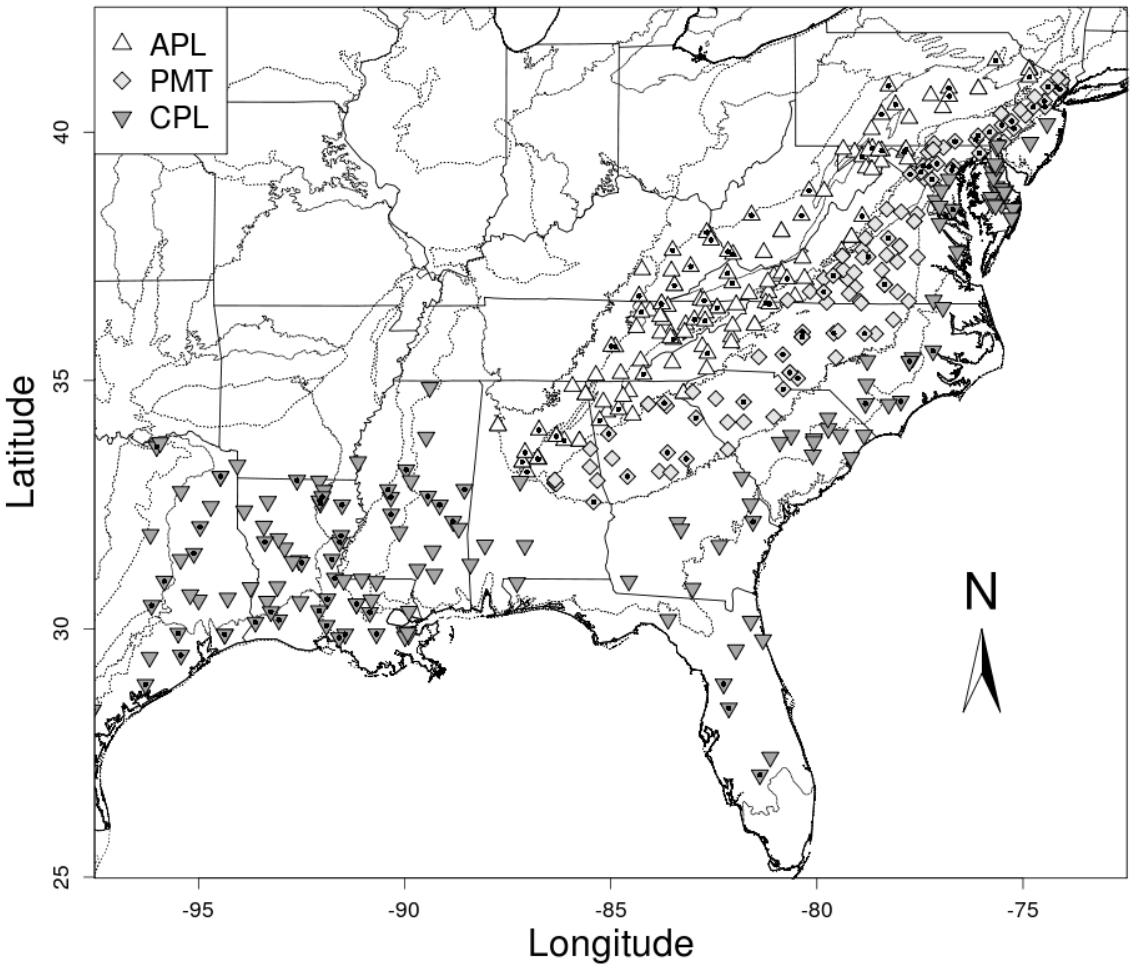

Figure 1. Map of 400 study sites located in the Appalachain (APL), Piedmont (PMT) and coastal plains (CPL) aggregate ecoregions. Symbols denote site locations and region designation. Highly disturbed (H) sites are marked by black dots (●), moderately disturbed (M) sites lack dots. Solid lines indicate state boundaries and dotted lines indicate Level III Ecoregion boundaries.

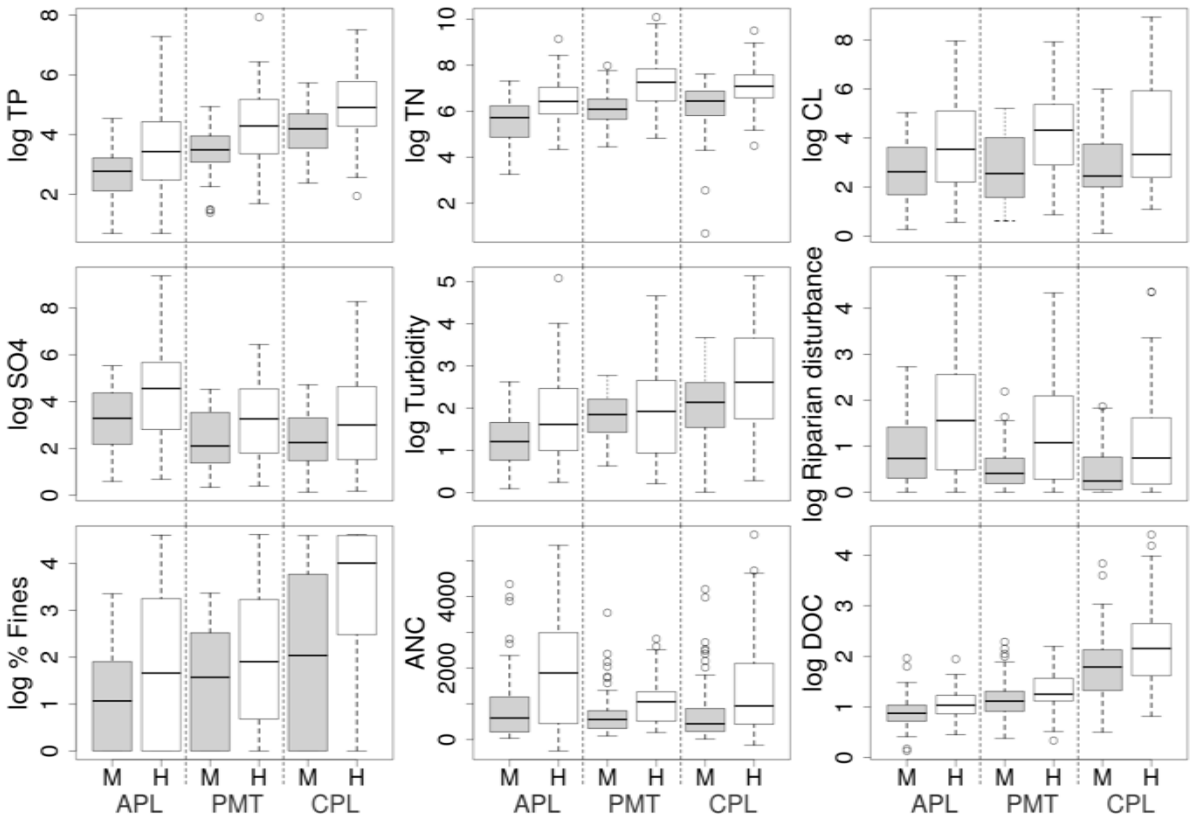

Figure 2. Boxplots of environmental filter variables used to create M/H impact classes (modified from Herlihy et al. (2008), see methods). Separate boxplots are provided for Appalachain (APL), Piedmont (PMT) and coastal plains (CPL) regions and moderately (M) and highly-disturbed (H) sites within each region. Dark lines show median values, boxes cover the interquartile range. TP = total phosphorus, TN = total nitrogen, Fines = fine sediment, ANC = acid neutralizing capacity, DOC = dissolved organic carbon.

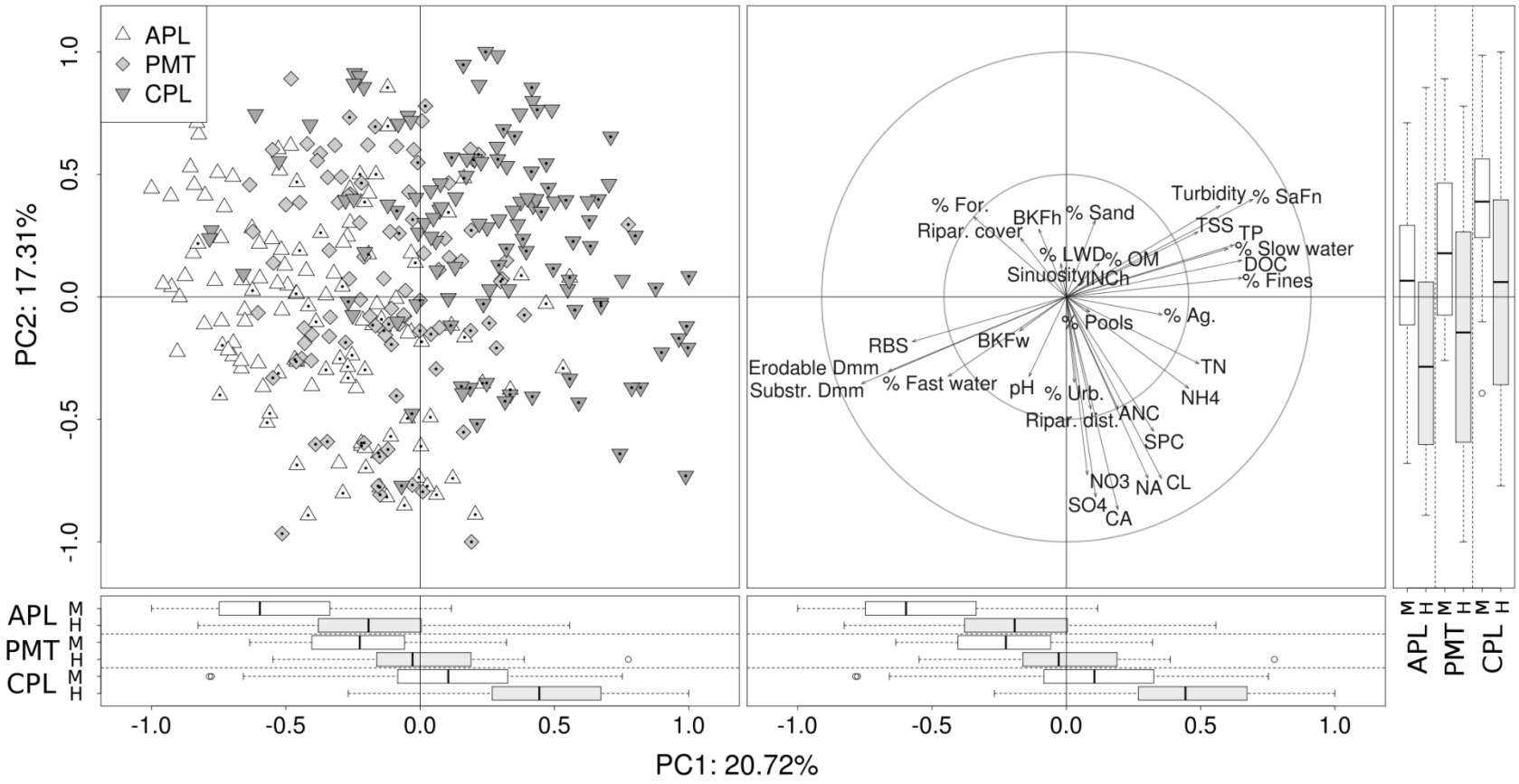

Figure 3. PCA on the (rank-based) covariance matrix of 37 environmental variables. Top-left panel shows sites in environmental space for the 1st two PC axes. Top-middle panel shows circle of correlations between variables and PC axes. Identical boxplots along the bottom row show distribution of site positions in each region and impact designation along PC1; boxplots in the far right column show site positions along PC2. Sites are shown in reduced space, shapes denote region designation (see legend). Highly-disturbed (H) sites are marked additionally by black dots (●), moderately-disturbed (M) sites lack these dots. See Appendix 5.1 for variable abbreviations, transformations, and descriptions.

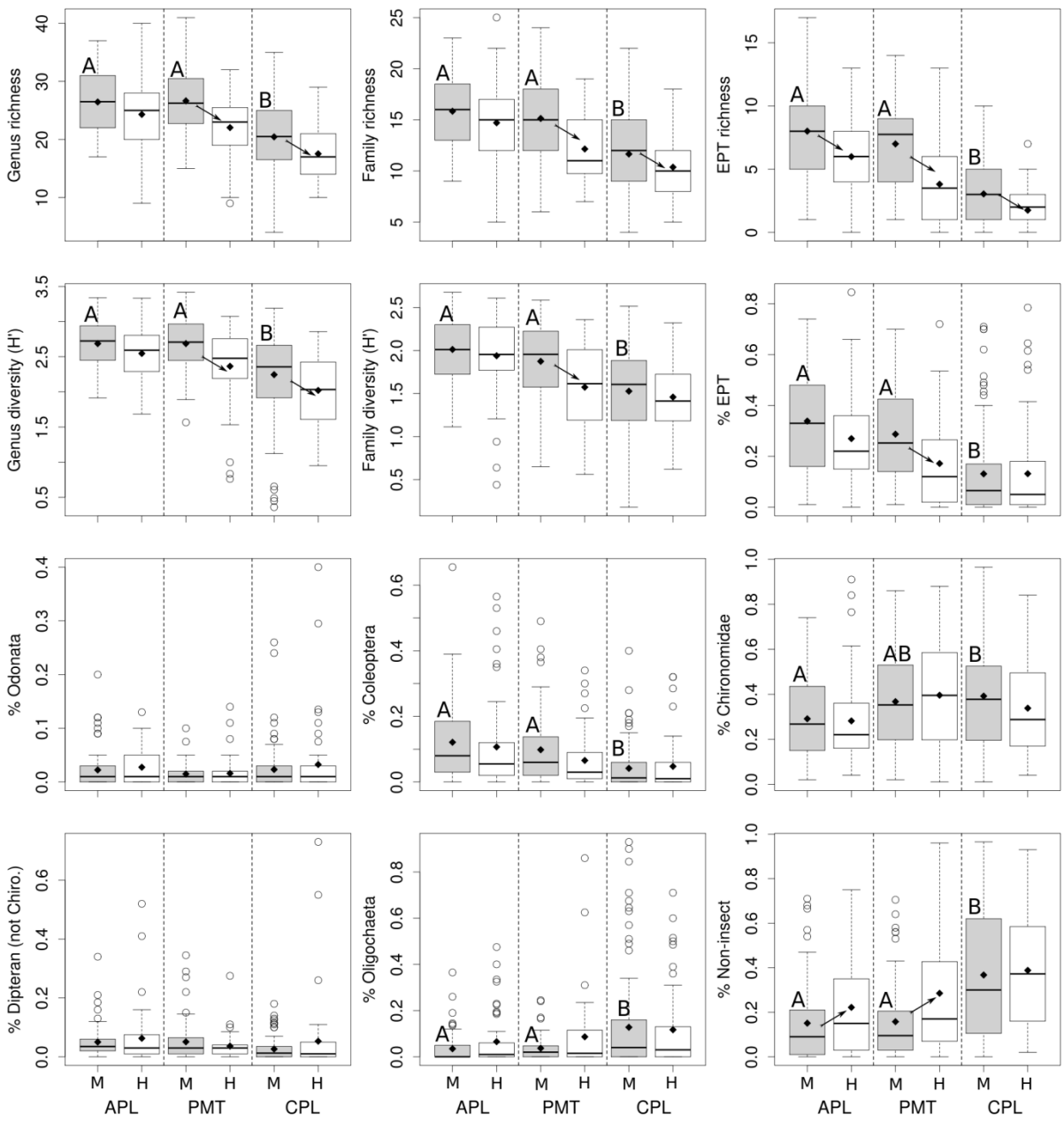

Figure 4. Boxplots of rarefied benthic metrics: richness/diversity at the levels of family and genus, richness/proportions of taxonomic groups including Ephemeroptera, Plecoptera and Trichoptera (EPT). Separate boxplots are provided for each aggregate region (see Fig. 1 for abbreviations) and moderately (M) and highly-disturbed (H) sites within each region. Dark lines show median values, boxes cover the interquartile range and small black dots show mean values. For M sites only, significant differences in regional means are indicated by A, B, C groupings. Arrows between grey- and white-filled boxes show significant differences between the M and H groups within a region. An absence of letters or arrows indicates no differences between medians (see methods for details) of groups.

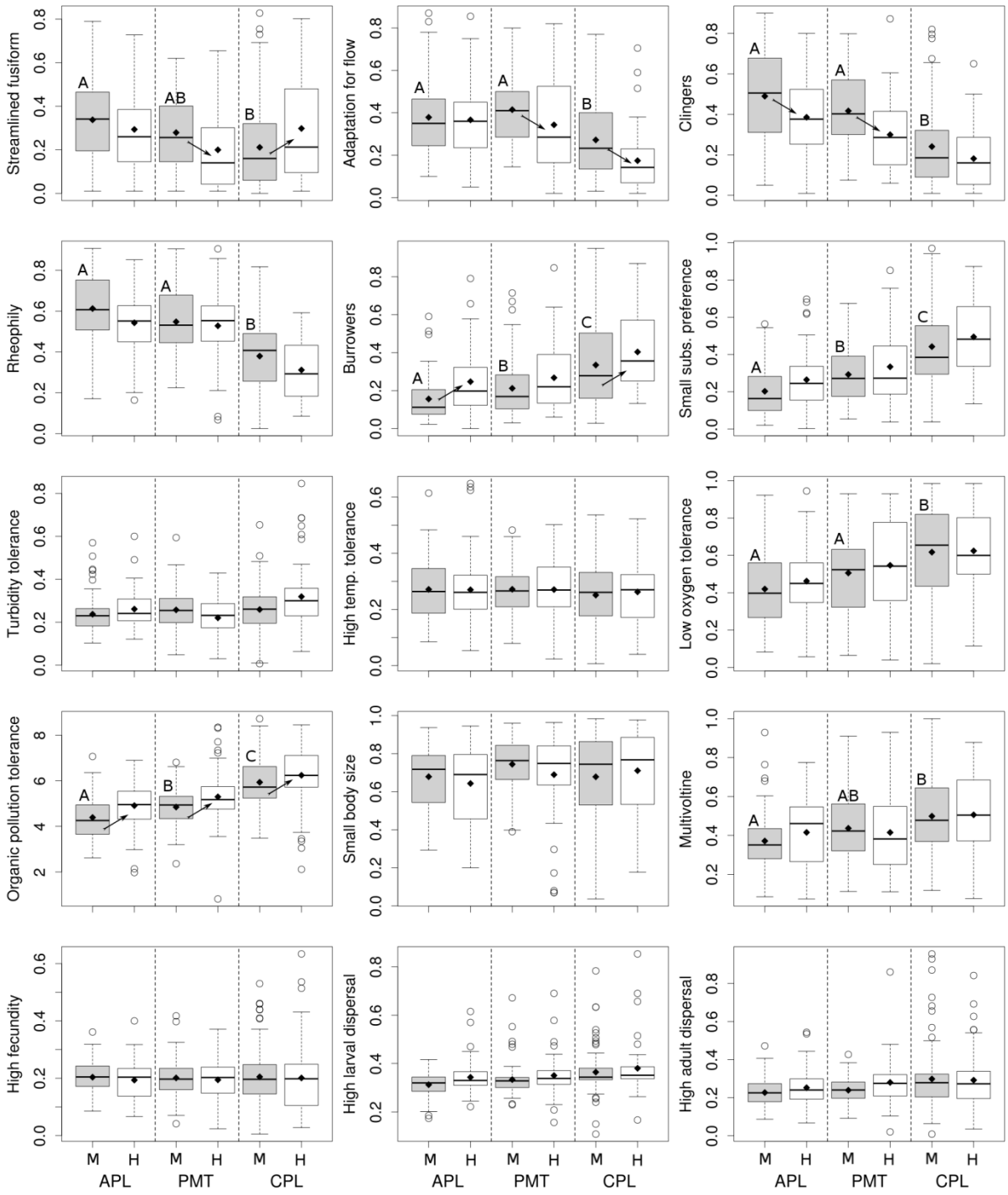

Figure 5. Boxplots of select trait states for each aggregate region and disturbance designation (M/H, see Figs. 1 and 4 for additional details and abbreviations).

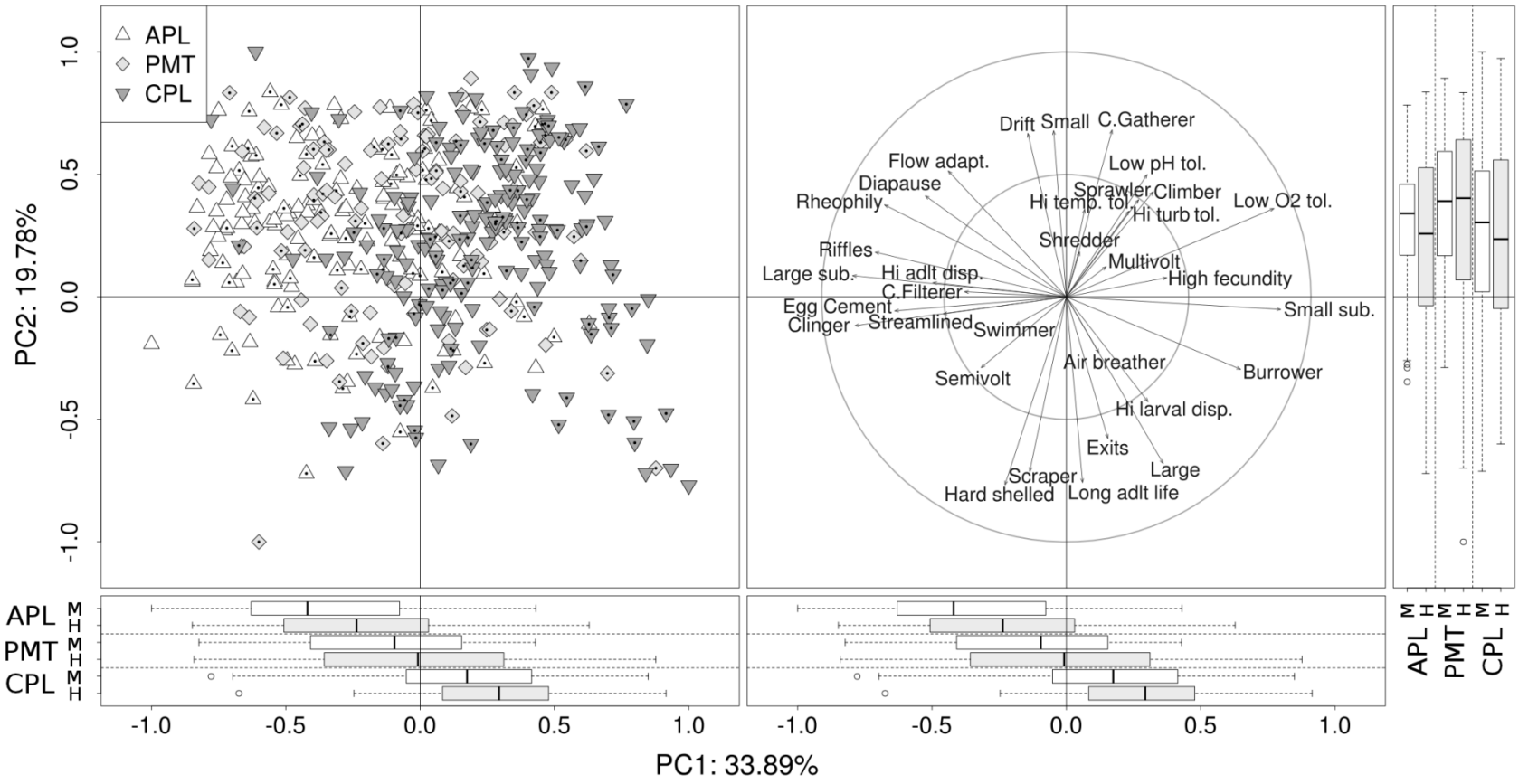

Figure 6. Major axes of variability in assemblage averaged trait values determined by principal component analysis; variability explained by each axis is provided next to axis label. Top-left panel shows sites in trait space for the 1st two principal components (PC) axes. See Fig. 1 for abbreviations and Fig. 3 for details regarding site symbols, boxplots, and circle of correlations plot. Additional traits not outlined in Table 1 are listed in Appendix 5.2.

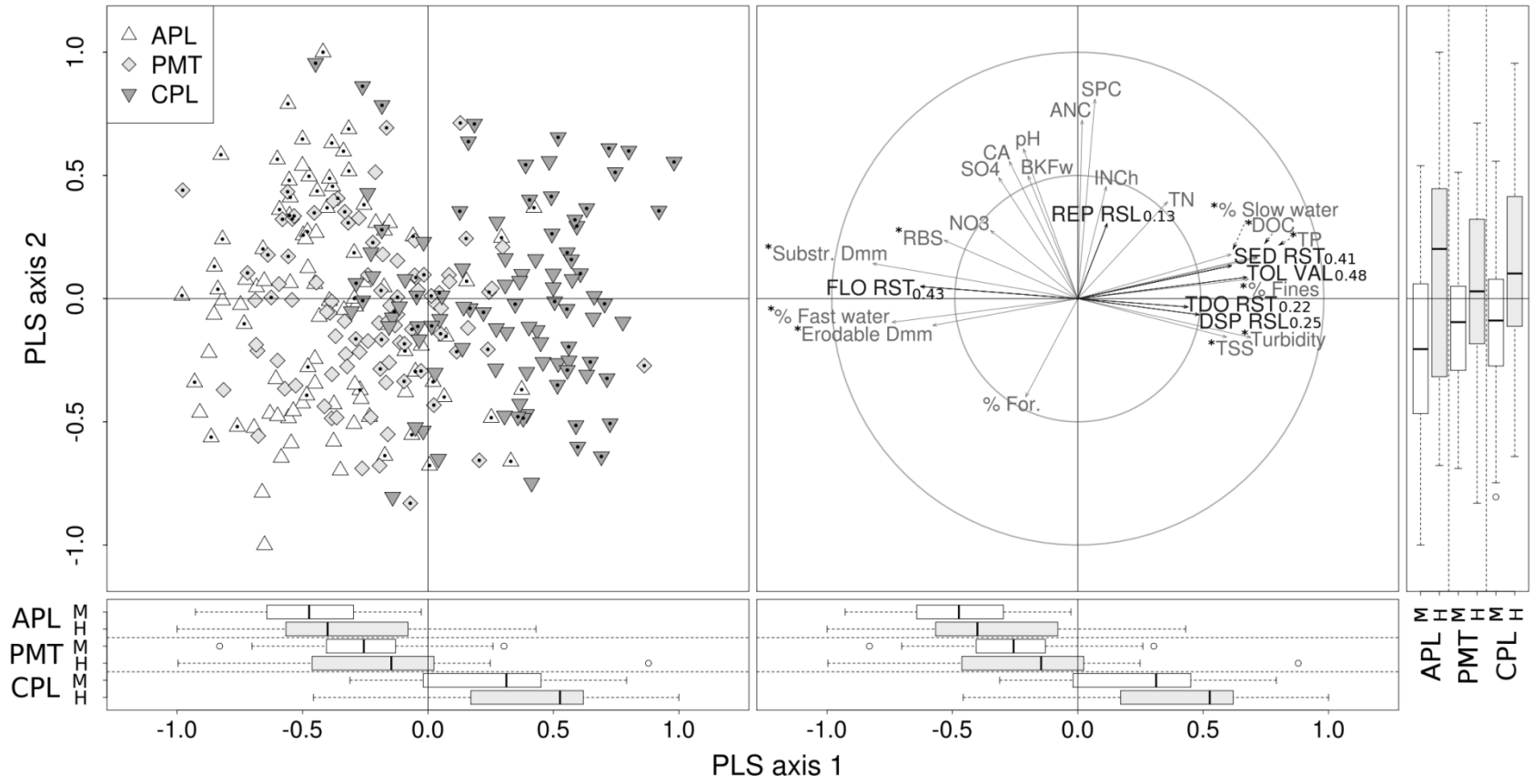

Figure 7. Partial least-squares (PLS) model of resistance (RST) and resilience (RSL) trait values (**Y**s) and environmental variables (**X**s; see Fig. 2 and Appendix 1). **X** variables with VIP < 1.0 in initial model were excluded from final model. Top-left panel shows site positions in reduced space that describe axes of greatest correlation between **Y** and **X**. See Fig. 1 for abbreviations and Fig. 3 for details regarding site symbols, boxplots, and circle of correlations plot. RST traits were for flow (FLO), sediment (SED), temperature/dissolved O_2 (TDO), and RSL traits were related to reproductive output (REP) and dispersal potential (DSP). * indicates environmental variables with final model VIP > 1.0. Final model R^2 (as $\text{cor}(\mathbf{Y}, \widehat{\mathbf{Y}})^2$) are printed next to variable name in the circle of correlation plot. See Appendix 5.1 for variable abbreviations, transformations, and descriptions.

Appendix 3.1. Taxa list for 13 sites in/near Foley AL, Wolf Bay. Site number corresponds to order in Table 1 (PP = private property): 1) BON12 = Bon Secour at AL Hwy 12 ,2) FPR29 = Foley preserve, Graham bayou creek, north crossing, 3) FPR30 = south crossing, 4) GUM13 = Gum branch (PP), 5) HMK33 = Hammock creek (PP), 6) MAG65 = Magnolia river at Hwy 65, 7) MFL08 = Miflin creek at Hwy 98, 8) MFL83 = Miflin creek tributary (PP), 9) PLM20 = Palmetto creek at Hwy 20, 10) SAN06 = Sandy creek at Hwy 98, 11) SAN 7E = Sandy east tributary at Hwy 98, 12) SAN 7E = Sandy west, 13) WLF01 = Wolf creek at Doc McDuffy Rd.

Order	Family	Taxon	Site presence
Amph	Gammaridae	Gammarus	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Amph	Hyalellidae	Hyalella	1, 9, 13
Clpt	Dytiscidae	Dytiscidae	4, 7, 9
Clpt	Elmidae	Ancyronyx	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Clpt	Elmidae	Dubiraphia	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 13
Clpt	Elmidae	Gonielmis	6, 10
Clpt	Elmidae	Macronychus	1, 4
Clpt	Elmidae	Microcylloepus	1, 2, 4, 8, 9, 10, 12, 13
Clpt	Elmidae	Promoresia	10
Clpt	Elmidae	Stenelmis	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Clpt	Gyrinidae	Dineutus	7, 9, 11, 12
Clpt	Helodidae	Scirtes	1, 13
Clpt	Hydrophilidae	Berosus	2, 5
Clpt	Psephenidae	Ectopria	6, 12
Dcpd	Cambaridae	Procambarus	1, 2, 3, 4, 5, 6, 7, 9, 10, 12
Dptr	Ceratopogonidae	Atrichopogon	7, 10, 12
Dptr	Ceratopogonidae	Probezzia	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Ablabesmyia	1, 2, 4, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Apedilum	6, 9
Dptr	Chironomidae	Chironomus	7
Dptr	Chironomidae	Clinotanypus	2, 5, 9
Dptr	Chironomidae	Corynoneura	1, 2, 4, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Cricotopus	1, 2, 4, 6, 7, 8, 10, 11, 12, 13
Dptr	Chironomidae	Cryptochironomus	1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Demicryptochironomus	1, 11
Dptr	Chironomidae	Harnischia	10, 11, 12
Dptr	Chironomidae	Labrundinia	4, 7, 8
Dptr	Chironomidae	Larsia	1, 4, 5, 9, 10, 11, 12, 13
Dptr	Chironomidae	Microtendipes	6, 9, 10
Dptr	Chironomidae	Nanocladius	1, 7
Dptr	Chironomidae	Paracladopelma	1, 6, 7, 9, 10, 11, 12, 13
Dptr	Chironomidae	Parakiefferiella	10
Dptr	Chironomidae	Parametriocnemus	1, 2, 4, 6, 7, 8, 9, 10, 11, 12
Dptr	Chironomidae	Paratendipes	10
Dptr	Chironomidae	Polypedilum	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13

Dptr	Chironomidae	Pseudorthocladius	1, 6, 7, 9, 10, 12
Dptr	Chironomidae	Pseudosmittia	1
Dptr	Chironomidae	Rheocricotopus	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Rheosmittia	1, 5, 6, 10, 11, 12, 13
Dptr	Chironomidae	Rheotanytarsus	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Robackia	11, 12
Dptr	Chironomidae	Saetheria	4, 7, 11, 12
Dptr	Chironomidae	Stelechomyia	1, 6, 10, 12
Dptr	Chironomidae	Stempellina	10
Dptr	Chironomidae	Stempellinella	1, 2, 10
Dptr	Chironomidae	Stenochironomus	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Tanytarsus	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Thienemanniella	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Thienemannimyia	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Tribelos	1, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Chironomidae	Tvetenia	1, 4, 6, 10
Dptr	Chironomidae	Xylotopus	1, 6, 7, 9, 10, 11, 12, 13
Dptr	Empididae	Hemerodromia	1, 6, 7, 8, 9, 10, 11, 12
Dptr	Psychodidae	Pericoma	7
Dptr	Simuliidae	Simulium	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Dptr	Tabanidae	Tabanus	1, 10, 11, 12
Dptr	Tipulidae	Hexatoma	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
Dptr	Tipulidae	Pilaria	2
Dptr	Tipulidae	Tipula	1, 5, 6, 7, 12
Ephm	Baetidae	Baetidae	4, 6, 7, 9, 10, 11, 12, 13
Ephm	Caenidae	Caenis	1, 6, 7, 9, 10, 11, 12, 13
Ephm	Heptageniidae	Stenonema	1, 4, 6, 7, 8, 9, 10, 12, 13
Ephm	Leptophlebiidae	Paraleptophlebia	9, 12
Ispd	Asellidae	Lirceus	1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 13
Mglp	Corydalidae	Nigronia	2, 11
Mglp	Sialidae	Sialis	9
Odnt	Aeshnidae	Boyeria	2, 4, 5, 7, 8, 9, 10, 11, 12
Odnt	Calopterygidae	Calopteryx	2, 4, 5, 7, 8, 9, 10, 11, 12, 13
Odnt	Coenagrionidae	Argia	4, 5, 7, 9, 11, 12, 13
Odnt	Coenagrionidae	Enallagma	4, 5, 8, 9, 12
Odnt	Cordulegastridae	Cordulegaster	8
Odnt	Corduliidae	Didymops	9, 12
Odnt	Corduliidae	Epitheca	2
Odnt	Corduliidae	Neurocordulia	2, 5, 9, 11, 12
Odnt	Gomphidae	Gomphus	4, 7, 8, 9, 12
Odnt	Gomphidae	Progomphus	1, 4, 6, 7, 8, 10, 11, 12, 13
Odnt	Libellulidae	Libellulidae	5, 12
Odnt	Libellulidae	Perithemis	1, 2, 3, 8, 9, 12
Plcp	Leuctridae	Leuctra	1, 9, 11, 12

Plcp	Perlidae	Perlidae	4, 6, 10, 11, 12
Trch	Calamoceratidae	Anisocentropus	10
Trch	Hydropsychidae	Cheumatopsyche	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
Trch	Hydropsychidae	Hydropsyche	1, 2, 4, 8, 9, 10, 11, 12, 13
Trch	Hydroptilidae	Hydroptila	2, 4, 5, 6, 9, 10, 12
Trch	Hydroptilidae	Oxyethira	2, 4, 5, 8, 12, 13
Trch	Leptoceridae	Oecetis	1, 4, 7, 9, 10
Trch	Leptoceridae	Setodes	2, 8
Trch	Limnephilidae	Pycnopsyche	2, 5
Trch	Philopotamidae	Chimarra	1, 4, 6, 8, 9, 11, 12, 13
Trch	Polycentropodidae	Cyrnellus	5, 6
Trch	Polycentropodidae	Polycentropus	9

Appendix 5.1 List of environmental variables used and transformations to improve distributional symmetry (reduce skewness).

Variable	Original variable abbreviation (if different) and description (if necessary)	Transformation (additional)
TP	PTL, Total phosphorous	$\log_{10}$
TN	NTL, Total nitrogen	$\log_{10}$
NO3	Nitrate	$\log_{10}$
NA	NA.	$\log_{10}$
NH4	Ammonium	$\log_{10}$
DOC	Dissolved Organic Carbon	$\log_{10}$
CA	Calcium	$\log_{10}$
pH		none
SO4	Sulfate	$\log_{10}$
ANC	Acid neutralizing capacity	$\log_{10}$
SPC	COND, Conductivity	$\log_{10}$
TSS	Total suspended solids	$\log_{10}$
Turbidity	TURB	$\log_{10}$
Ripar. cover	XC, riparian cover	$\log_{10}$
Ripar. dist.	W1_HALL, riparian disturbance	none
INCh	XINC_H, Incised height	$\log_{10}$
BKFw	XBKF_W, Bankfull width	$\log_{10}$
BKFh	XBKF_H, Bankfull height	$\log_{10}$
Slope	XSLOPE	$\log_{10}$
% Fines	PCT_FN	none
% SaFn	PCT_SAFN, % sand and fines	$\log_{10}$
% Fast water	PCT_FAST	$\log_{10}$
% Slow water	PCT_SLOW	none
% Pools	PCT_POOL	none
Substr. Dmm	LSUB_DMM, log mean substrate diameter	none
Erodable Dmm	LSUB_DMM, log erodible substrate diameter	none
RBS	LRBS_BW5, log relative bed stability	none
LWD	XFC_LWD, % large woody debris	Rank transformed

Appendix 5.2 List of species traits derived from USGS trait database.

Trait state	Original trait name, trait state
Streamlined	Body_shape, Streamlined / fusiform
Small	Max_body_size, Small (length < 9 mm)
Large	Max_body_size, Large (length > 16 mm)
Hard shelled	Armor, Hard shelled
Flow adapt.	Morph_adapt_suckers, Morph_adapt_friction, Morph_adapt_hooks, Morph_adapt_silk, Morph_adapt_ballast, and Morph_adapt_hairy
Clinger	Habit_prim, Clinger
Burrower	Habit_prim, Burrower
Climber	Habit_prim, Climber
Sprawler	Habit_prim, Sprawler
Swimmer	Habit_prim, Swimmer
C. Filterer	Feed_mode_prim, Collector-filterer
C. Gatherer	Feed_mode_prim, Collector-gatherer
Scraper	Feed_mode_prim, Scraper/grazer
Shredder	Feed_mode_prim, Shredder
Semivolt	Voltinism, < 1 Generation per year
Multivolt	Voltinism, > 1 Generation per year
Diapause	Diapause, Yes
Rheophily	Current_fast_lam, and Current_fast_turb
High fecundity	Fecundity, > 10,000 eggs
Long adlt life	Adult_lifespan, Months
Hi adlt disp.	Adult_disp, 100 km or less
Hi larval disp.	Larval_disp, 11-100 m
Drift	Drift_early and Drift_late, Strong (active / often)
Riffles	Lateral_preference, Lat_riffle
Large sub.	Microhab_gravel, Microhab_rocks, and Microhab_boulder
Small sub.	Microhab_sand and Microhab_silt
Low pH tol.	pH_acidic
Hi temp. tol.	Thermal_pref, Hot euthermal
Hi turb tol.	Turbidity, Silted/murky water
Lo O2 tol.	O2_low
Air breather	Resp_early and Resp_late, Atmospheric breathers
Exits	Exit_temporarily, Yes
Egg cement	Eggs_cement, Yes