An Investigation of the Relationships Between Stream Benthic Macroinvertebrate Assemblage Conditions and Their Stressors

Laurie Frondorf

Thesis Submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science

in

Biological Systems Engineering

Dr. Saied Mostaghimi, Chair
Dr. Theo A. Dillaha, III
Dr. J. Reese Voshell
Dr. John V. Perumpral, Department Head

May 7, 2001
Blacksburg, Virginia

Keywords: Benthic Macroinvertebrate, Stressors, Sediment, RBP, TMDL

An Investigation of the Relationships Between Stream Benthic Macroinvertebrate Assemblage Conditions and Their Stressors

Laurie Frondorf

ABSTRACT

Agriculture, urbanization, and human activities, if not managed carefully, can expose a water body to environmental degradation, decreased water quality, and ultimately impaired benthic macroinvertebrate assemblage conditions. In streams where the benthic macroinvertebrates are impaired, the stream itself will not be meeting the water quality standards set forth in the Clean Water Act. As a result, the goal of this study was to establish relationships between benthic macroinvertebrates and their stressors so that stressor levels that would not adversely impact the benthic macroinvertebrates could be determined. Stressors such as sediment, habitat, water quality, landuse, watershed characteristics, and livestock numbers impact the benthic macroinvertebrate assemblage conditions. Since sediment is recognized as the Nation's leading pollutant and since the benthic macroinvertebrates live in the sediment on a stream bottom, this study placed emphasis upon the investigation of sediment as a primary stressor to the benthic macroinvertebrates. The specific objectives of this study were to develop relationships between the benthic macroinvertebrates and sediment and other stressors for Virginia streams, to evaluate the accuracy of the stressor/benthos relationships, and to discuss the implications of the study results for development of benthic TMDLs.

A procedure to determine the relationships between stressors and benthic macroinvertebrate assemblage conditions was developed. Existing data on sediment, habitat, water quality, landuse, watershed characteristics, livestock numbers, and benthic macroinvertebrate assemblage conditions were compiled for 34 stations with 105 samples collected from the fall of 1996 to the fall of 1998. The 34 stations were located within 13 counties in Virginia (Rockbridge, Rockingham, Augusta, Frederick, Shenandoah, Page, Loudoun, Fairfax, Prince William, Fauquier, Culpeper, Rappahannock, and Madison) and in watersheds dominated by agricultural, urban, and forested landuses. Virginia currently uses the Rapid Bioassessment Protocol (RBP) method in its Biological Assessment Program. The RBP compares habitat and biological measures of the benthic macroinvertebrates to reference conditions using individual metrics. VADEQ's Biomonitoring Database, together with Ambient Water Quality Monitoring reports, GIS data layers, and VADCR's Hydrologic Unit Animal Census Database provided all of the necessary information for the stressor variables and benthic macroinvertebrate conditions. Accordingly, the stressor/benthos relationships were evaluated using statistical analyses procedures such as forward, backward, and stepwise multiple regression techniques; correlation analysis; principal component analysis; and r-square analysis. The statistical results indicated that sediment alone cannot be used to assess the benthic macroinvertebrate assemblage conditions. Other stressors such as dissolved oxygen, flow, % urban land, total suspended solids, temperature, stream velocity, substrate, hardness and alkalinity greatly impact the benthic macroinvertebrate assemblage conditions. The study results also indicate that the individual metrics within the RBP procedure are just as critical as the final RBP values in describing the benthic macroinvertebrate assemblage conditions.

Upon completing the steps needed to develop stressor/benthos relationships, the validity of the relationships were verified for their application to other streams in Virginia. Validation was completed using 10 stations with 29 samples from the fall of 1996 to the fall of 1998. The 10 stations were located within 8 counties in Virginia (Bedford, Montgomery, Pulaski, Giles, Botetourt, Albemarle, Orange, and Culpeper) and in watersheds dominated by agricultural, urban, and forested landuses to correspond with the stations used to develop the stressor/benthos relationships.

The implications of the relationships with regard to TMDLs were also studied using total suspended solids (TSS) loadings, turbidity levels, and embeddedness levels as the stressors of concern. The results for all benthic stations within Virginia showed that moderately impaired streams generally need to reduce the amount of embeddedness by 11, reduce turbidity levels by 5 FTU (57%), and reduce TSS values by 7 mg/L (68%) to meet a threshold value that would no longer adversely impact the benthic macroinvertebrates. Similarly, for the severely impaired stations throughout Virginia to meet threshold values, embeddedness amounts need to be reduced by 22, turbidity reduced by 57 FTU (93%), and TSS reduced by 74 mg/L (96%).

This study was important since the proposed stressor/benthos relationships can provide policymakers with a useful tool to determine stressor thresholds that will not adversely impact the benthic macroinvertebrate assemblage conditions for use in developing benthic TMDLs in Virginia. The stressor/benthos relationships could also be used to determine the impact of certain activities or stressors on the benthic macroinvertebrates assemblage conditions in a given stream.

ACKNOWLEDGMENTS

I would like to express my gratitude to the following individuals for their support throughout this project:

I would like to sincerely thank my advisor Dr. Saied Mostaghimi for his continuous confidence, advice, patience, and for his encouragement that I even pursue my master's degree.

Dr. Dillaha and Dr. Voshell for taking the time to serve on my committee and for their valuable insight and suggestions.

Dr. Perumpral for the opportunity to continue my education.

The Virginia Department of Conservation and Recreation and the Biological Systems Engineering Department for providing funds.

Phillip McClellan for proposing the study as well as for his help and advice.

Karl Huber of the Virginia Department of Conservation and Recreation and Gabriel Darkwah of the Virginia Department of Environmental Quality for their cooperation and assistance in data collection.

Dr. Eric Smith of the Statistical Consulting Center and Kevin Brannan for assistance with the statistical analyses procedures.

Jennifer Miller, Jan Carr, and Jeff Wynn in the Watershed Assessment Lab who provided endless help with contacts, data collection, and computer glitches.

All of my friends and fellow graduate students for making the time enjoyable.

Most of all, I would like to thank my parents, Kirk and Diane, and boyfriend, David. Their love, support, and constant encouragement made this all possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1.0 INTRODUCTION

1.1 Background

Over 40% of assessed waters in the United States do not meet the water quality standards set forth by states, territories, and authorized parties (USEPA, 2000a). The impaired waters account for 300,000 miles of rivers and shorelines and approximately 5 million acres of lakes. With 218 million people living within 10 miles of the impaired waters, citizen organizations have demanded changes through legal actions (USEPA, 2000a; Richman, 1997; Pelley, 1998).

The Clean Water Act (CWA) requires that water bodies be managed to maintain and restore their biotic integrity. Under section 303(d) of the CWA, states, territories, and authorized parties are required to develop lists of impaired waters (USEPA, 2000a). The law requires the jurisdictions to establish priority rankings for waters on the lists and develop Total Maximum Daily Loads (TMDLs). The TMDL plans should be approved by the Environmental Protection Agency (EPA) prior to their implementation.

The Virginia Department of Environmental Quality’s 1998 303(d) List identified 883 water bodies as impaired. The listed waters together contained 1002 impairments originating from 43 different sources. Of the 1002 impairments, benthic degradation accounted for 111 impairments, making it the fourth leading cause of impairments in the commonwealth. Only the Virginia Department of Health shellfish restriction, NH_3-N, and fecal coliform impairments ranked above benthic impaired waters in Virginia (USEPA, 2000b). In addition, the CWA states, "it is the objective of the Act to restore and maintain the chemical, physical, and biological integrity of the Nation's waters" (Novotny and Olem, 1994). Of the three characteristics, biological integrity may be the most important since organisms not only integrate the full range of environmental influences (chemical, physical, and biological), but complete their life cycles in the water and, as such, are continuous “monitors” of environmental quality (Richards and Host, 1994). Therefore, evaluation of benthic conditions and development of benthic TMDLs are of great importance to meeting current water quality goals in Virginia.

Traditional measures or performance based standards of water quality, such as levels of dissolved oxygen or concentrations of toxic contaminants in water, are indirect ways to determine the health of a water body. Conclusions about expected effects on aquatic life may be inferred from the performance-based standards; however, the biological responses in the stream cannot be directly studied. By assessing the invertebrate communities and comparing the results to those found in pollution-free areas, it is possible to determine whether or not pollution is causing ecological effects, such as the loss of sensitive groups of organisms (Pawlak, 1999).

Biological impairment of the benthic community can be indicated by the absence of pollution sensitive macroinvertebrates (Barbour et al., 1999). Since the benthic macroinvertebrate community is used as an indicator of the overall health of a water body, the streams' macroinvertebrate data do not demonstrate which specific pollutants are causing the benthic impairments. Of the impaired streams on the 1998 list in Virginia, 44%, 11%, and 9% were classified as originating from nonpoint sources, agriculture, and urban areas, respectively (USEPA, 2000b).

Due to urbanization and agricultural activities, the amount of impervious area and/or amount of disturbed land increases. Land disturbances directly influence the magnitude of stormwater runoff, and ultimately increasing the amount of stream flow from surface runoff rather than from baseflow or groundwater (Richards and Host, 1994; Booth and Jackson, 1997). The results are higher and more frequent high flow events and lower low flow or even no flow during dry weather conditions. Higher flow rates result in an increase in sediment losses from disturbed areas and in bank erosion and channel scouring. Consequently, the hydrological changes alter the habitat and geometry of the streams and increase the amount of sediment pollution (Knighton, 1984).

With sediment being recognized as the major pollutant of United States waters, excessive sediment loads are thought to be a major contributor to the decline of a stream's benthic community (USEPA, 1990). Sediment affects the benthic community by altering water movement, food quality, and interstitial spacing (Minshall, 1984). Fine sediment decreases diversity since the suspended solids absorb heat from sunlight, causing a temperature increase and ultimately a reduction in dissolved oxygen (Murphy, 2000; MIDEQ, 2000). Sediment also reduces habitat, gradually decreasing the standing crop, taxa richness, and diversity without a drastic change in overall taxonomic composition (Lenat et al., 1981).

In addition to sediment pollution, increased concentrations of pollutants in the runoff affect the benthic community. Pollutants of concern include nutrients, toxics, and suspended materials. Temperature and dissolved oxygen effects may also be present due to both runoff and loss of riparian vegetation (Hem, 1992; Watzin and McIntosh, 1999). The geomorphic, hydrologic, and water quality parameters affecting the invertebrate community make isolation of a single agent difficult (Allan, 1995). Thus, aquatic resources may be under stress posed by a multitude of practices within a watershed. Overall, the biological community provides an ideal response indicator serving as a pertinent measure for water quality goals.

Even though biological integrity and ultimately water quality conditions can be assessed with benthic macroinvertebrates, it is not currently possible for Virginia to develop a TMDL for a stream listed for benthic impairment since the exact causes of benthic impairment and the level of pollution reduction required to restore the stream's benthic community is unknown. This study evaluated the relationship between the benthic community, as quantified by the state's Rapid Bioassessment Protocols (RBPs), and simulated sediment loads on a watershed basis. Since the benthic community can be impacted by numerous pollutants, additional variables such as landuse, watershed characteristics, habitat, livestock numbers, and water quality parameters were also included in the analyses. Even though the primary focus was the sediment indicators, the additional variables were incorporated to improve the accuracy of the results. The overall intention was, therefore, to develop sediment and other stressors thresholds for restoring a stream's biological integrity. Successful completion of this study would assist Virginia in developing TMDLs for streams with benthic impairments.

1.2 Goal and Objectives

The goal of the study was to establish relationships between stream benthic macroinvertebrate assemblage conditions and their stressors in Virginia. The specific objectives were to:

1.) Develop relationships between the RBP indices (or individual RBP metrics) and stressors such as sediment loads, landuse, habitat, etc while placing emphasis upon the sediment loadings as a primary stressor.

2.) Evaluate the accuracy of the stressor/benthos relationships using data from selected watersheds in Virginia.

3.) Discuss the implications of the study results for benthic TMDLs. Specifically, investigate stressor thresholds that would not adversely impact the benthic macroinvertebrate assemblage conditions.

1.3 Hypothesis

HYPOTHESIS

Sediment loading, or related stressors, can be used to assess benthic macroinvertebrate assemblage conditions in a stream.

NULL HYPOTHESIS

Sediment loading, or related stressors, can **not** be used to assess benthic macroinvertebrate assemblage conditions in a stream.

CHAPTER 2.0 LITERATURE REVIEW

In order to determine the level of sediment and other stressors that will not adversely impact the benthic macroinvertebrate community, a review of literature concerning water quality and benthic macroinvertebrates was conducted. First, the progression of water quality goals and regulations and the origin of the current issue of TMDLs are reviewed. Second, an overview of benthic macroinvertebrates, specifically the impact of stressors upon the benthic macroinvertebrates and the history of benthic macroinvertebrates as monitors of stream quality, are presented. Third, EPA’s Rapid Bioassessment Protocol is reviewed to determine the protocol’s effectiveness in evaluating a stream’s quality using habitat, physiochemical, and biological assessments. Finally, a discussion of sediment modeling, specifically within GIS, is outlined since the relationship between sediment and invertebrates is of primary concern in this study.

2.1 Water Quality

An increase in environmental awareness in recent years has led to new perceptions of water quality. The Water Pollution Control Act Amendments in 1972, or Clean Water Act (CWA), marked one of the most far-reaching environmental legislative acts to solve environmental problems (Novotny and Olem, 1994; Copeland, 1999). Pollution, defined as any artificial or human induced change to the environment, originates from either a point source or nonpoint source (Hun, 1998). During the years following the enactment of the CWA, the majority of the cleanup focused upon sewage and industrial wastewater discharge point sources and their impact on drinking water. However, in 1987 amendments to the CWA authorized measures to address nonpoint source (NPS) pollution, such as stormwater runoff from farmlands, forests, construction sites, and urban areas (Copeland, 1999). NPS pollution is reported to be the main reason that approximately 40 percent of our surveyed rivers, lakes, and estuaries are not clean enough to meet basic uses such as swimming or fishing, making NPS pollution the Nation’s largest source of current water quality problems (USEPA, 1999).

NPS pollution is caused by rainfall or snowmelt moving over and through the ground. As the runoff moves, the water picks up and carries away natural and human-made pollutants, finally depositing them into lakes, rivers, wetlands, coastal waters, and even our underground sources of drinking water (USEPA, 1999). Virginia's Nonpoint Source Pollution Management Program states NPS pollution is (DSWC, 1989):

> "*…caused by diffuse sources that are not regulated as point sources and normally is associated with agricultural, silvicultural and urban runoff, runoff from construction activities, etc. Such pollution results in the human-made or human-induced alteration of the chemical, physical, biological, and radiological integrity of water. In practical terms, nonpoint source pollution does not result from a discharge at a specific, single location (such as a single pipe) but generally results from land runoff, precipitation, atmospheric deposition, or percolation*."

NPSs may include a variety of industries, practices, and regions, such as agriculture, natural sources, mining, hydrologic and habitat modification, urban, land disposal, construction, and silviculture. However, agriculture has become the most targeted industry for control of NPS pollution The latest *National Water Quality Inventory* indicates that agriculture is the leading contributor to water quality impairments, degrading 60 percent of the impaired river miles and half of the impaired lake acreage surveyed by states, territories, and tribes (USEPA, 1999). The most common NPS pollutant from agricultural areas is sediment (USEPA, 1992a; USEPA, 1999).

Sediment is recognized as the major pollutant of United States waters due to the immense amounts produced by human activities. In a USEPA inventory of stream water quality, siltation was identified as the most important cause of river and stream pollution (USEPA, 1990). In fact, the magnitude of siltation was 50% greater than the next most important cause (nutrients). However, sediment pollution is not just a recent concern. For instance, Oschwald (1972), in a review of a U.S. Department of Agriculture Report to the President (USDA 1969), stated, "*sediment originating from soil erosion has been called the major pollutant of surface waters in terms of quantity involved*." Ritchie (1972) began his review on the impact of sediment on fish with a similar statement: "*Quantitatively, sediment is the greatest single pollutant in the nation's water*."

Landuse activities greatly impact the amount of sediment deposition in a stream. Basnyat et al. (1999) found that forests and grasslands located adjacent to streams provide the best water quality while agricultural and urban areas have negative impacts on water quality. With landuse in mind, the concern over erosion and sediment deposition naturally centers upon the loss of agricultural soils, decreased water-retention capacity of forest lands, increased flood frequency, and rapid filling of reservoirs. The less obvious effect of sedimentation is in small streams where the degraded water quality and habitat reduce the diversity and productivity of biotic communities (Waters, 1995; Richards and Host, 1994).

Even though water quality has no unique definition, Novotny and Olem (1994) describe water quality as a reflection upon "*…the composition of water as affected by natural causes and man's cultural activities, expressed in terms of measurable quantities and related to intended water use.*" In conjunction with adequate water quality, the CWA requires that the "*…chemical, physical and biological integrity of the Nation's waters*" be preserved, and that human activities do not degrade existing beneficial uses (Novotny and Olem, 1994). In order to demonstrate compliance with the CWA, each state must monitor the water quality of the aquatic resources. Since a large number of streams must be monitored with limited resources, the monitoring and establishment of pollutant reduction procedures must be carried out in a manner that provides useful information with minimal effort. The Total Maximum Daily Load process has aided the advancing efforts to reduce NPS pollution, while also changing the focus of the water quality efforts "from a clean water program based primarily on technology-based controls to water-quality-based controls implemented on a watershed basis" (Perciasepe, 1997).

2.2 Total Maximum Daily Loads

Total Maximum Daily Loads (TMDLs) became a requirement of the Federal Clean Water Act (CWA) in 1972 under section 303(d). However, TMDL procedures did not begin to take shape until environmental groups began monitoring EPA's progress toward development of TMDLs in 1992. At that time, the environmental groups began challenging EPA about their procedures, arguing that not all evaluations are complete or legally and scientifically sufficient to restore a contaminated water body to EPA standard (Richman, 1997; Pelley, 1998). To date, 30 legal actions in 38 states have been filed (USEPA, 2000a). With the added distress of lawsuits and legal action by the environmental groups, EPA formed a Federal Advisory Committee Act

(FACA) in 1996 in an effort to speed up the Nation's progress toward achieving water quality standards and enhancing the TMDL program.

The FACA, composed of 20 individuals with diverse backgrounds in agriculture, forestry, environmental advocacy, industry, and government, issued recommendations in 1998 (MDE, 2001). These recommendations helped to guide the development of proposed changes to TMDL regulations. The changes were published on July 13, 2000 as FACA's final rule. However, Congress has prohibited EPA from spending FY2000 and FY2001 money to implement the new rule (USEPA, 2000a). Consequently, the current rule remains in effect until 30 days after Congress permits EPA to implement the new rule.

The current rules are those issued in 1985 and amended in 1992 (40 CFR Part 130, section 130.7). These regulations require states to identify the waters that do not meet water quality standards and submit the list to EPA for its approval. The Federal regulations require states to consider all existing and readily available water quality-related data and information in developing the 303(d) list.

A TMDL establishes the maximum allowable pollutant loading for a water body to meet Water Quality Standards (WQS). TMDLs are defined as "… *a tool for implementing State water quality standards and are based on the relationship between pollution sources and in-stream water quality conditions*" (USEPA, 1991). A TMDL addresses a single pollutant or stressor for each water body. However, a TMDL is the sum of the allowed pollutant loads for point sources, non-point sources, projected growth, and a margin of safety as shown in Equation 2.1 (Al-Smadi et al., 2000).

$$\mathrm{TMDL} = \mathrm{LC} = \mathrm{WLA} + \mathrm{LA} + \mathrm{MOS} \tag{2.1}$$

where:
TMDL = Total maximum daily load,
LC = Loading capacity,
WLA = Waste load allocation from point sources,
LA = Load allocation from nonpoint sources and natural conditions,
MOS = Margin of safety.

TMDLs are expressed in either mass per time, toxicity, or other appropriate measures that relate to a State's water quality standard. TMDLs must contain a margin of safety and a consideration of seasonal variations (USEPA, 2000a).

A water body requires TMDLs when the technology-based effluent limitations required by the CWA; the stringent effluent limitations required by either State or local authority; or, other pollution control requirements required by local, State, or Federal authority are not stringent enough to meet the applicable WQSs (MDE, 2001). To help minimize confusion, EPA distinguished between pollutants and pollution. A pollutant is the addition of a substance to a water body while pollution is any artificial or human induced change to the environment. Consequently, EPA will only use the applicable term pollutants to set TMDLs and to determine when an impaired water body should be removed from the 303(d) list (Hun, 1998).

After a TMDL has been approved, States, territories, and authorized parties should develop schedules for implementing TMDLs, generally within 8-13 years of being listed (USEPA, 2000a). State and local water quality management plans should be updated and control measures should be implemented. For point sources, the National Pollution Discharge Elimination System (NPDES) Permit is issued for water quality-based limits. The nonpoint source controls are established through BMPs, technical assistance, financial assistance, education, training, technology transfer, and demonstration projects.

In summary, the TMDL process is an important step in broadening public participation and providing a mechanism for the development of innovative approaches for attaining and maintaining water quality standards when implemented on a local scale (USEPA, 1991). Thus, each individual state has the flexibility to implement new and innovative approaches to the identification and control of nonpoint sources of pollution due to the USEPA's broad definition of the TMDL process.

2.3 Benthic Macroinvertebrates

Benthic macroinvertebrates are stream-inhabiting organisms, easily viewed with the naked eye, that spend at least part of their lives living in or on the stream bottom. The name benthic macroinvertebrate derives from the fact that they are bottom dwelling (benthic), large enough to see (macro), and small organisms without backbones (invertebrates) (DEP, 2000). Since the invertebrates inhabit the stream bottom, any modification of the streambed by deposited sediment will most likely have a profound effect upon the benthic community. In addition, each species of macroinvertebrate has a specific tolerance for water conditions. Due to their close interactions within a water body, the benthic macroinvertebrate community of freshwater systems is a good indicator of ecosystem health (Clarke and Scruton, 1997*)*.

Benthic macroinvertebrates are often used as indicators of stream health because they occur in almost all streams, are easily collected, exhibit great diversity and are sensitive to pollution (DEP, 2000). Benthic macroinvertebrates display a wide range of responses to pollution, while not all benthic macroinvertebrates have the same sensitivity to pollution. One important macroinvertebrate community indicator is the EPT richness, or the total number of Ephemeroptera (mayfly), Plecoptera (stonefly), and Trichoptera (caddisfly) taxa in a sample. The EPT richness summarizes taxa richness within those insect taxa considered sensitive to water quality degradation. An increasing EPT richness value correlates with increasing water quality (Rothrock et al., 1998). Moreover, the EPT richness value relates the pollution sensitive taxa to the water quality they reside in.

2.3.1 Influence of Sediment on Benthic Communities

Fine sediment, specifically total suspended solids (TSS), affects the habitat quality of a substrate by altering water movement, food quality, oxygen availability, and interstitial spacing (Minshall, 1984). TSS includes all particles suspended in water which will not pass through a filter. Sediment and silt are often used to refer to suspended solids. As the levels of TSS increase, a water body begins to lose its ability to support a diversified aquatic life. Suspended solids absorb heat from sunlight, which increases water temperature and subsequently decreases levels of dissolved oxygen (Murphy, 2000; MIDEQ, 2000). These changes can dramatically impact aquatic life while the TSS itself can smother the eggs of aquatic insects. Most people consider water with a TSS concentration of less than 20 mg/L to be clear. Water with TSS levels between 40 and 80 mg/L appear cloudy and TSS over 150 mg/L appear dirty (MIDEQ, 2000).

Lenat et al. (1981) hypothesized two mechanisms of habitat degradation due to sediment, and termed them ‘habitat reduction’ and ‘habitat change.’ Habitat reduction was defined as a decrease in interstitial space due to small amounts of excess fine sediment. The characteristic response of habitat reduction is a gradual decrease in standing crop, taxa richness, and diversity, without a drastic change in overall taxonomic composition (Lenat et al., 1981). Organisms most affected by habitat reduction are those whose method of attachment or food acquisition requires that they inhabit rock surfaces. Most EPT taxa fall into this category (Waters, 1995).

The habitat change response as a result of sedimentation is related to differential substrate preference among benthic invertebrate species (e.g., Minshall, 1984). Large amounts of sediment effectively change coarse-bedded substrates to sand. The altered substrate produces a corresponding change in benthic community structure. The most common change is a shift from EPT taxa to burrowing forms capable of rapid colonization, such as oligichaetes and chironomids (Waters, 1995).

Numerous studies have evaluated the influence of sediment on benthic communities (Bjornn et al., 1974; Lenat et al, 1981; Richards and Bacon, 1994; Zanetell and Peckarsky, 1996; Shaver et al., 1997; Wood and Armitage, 1997; and Chanat, 1998). The effect of fine sediment on invertebrate standing stock, although sometimes dramatic, can be inconsistent and difficult to detect. Bjornn et al. (1974) reported a negative correlation between embeddedness and Shannon-Weaver diversity in natural streams in Idaho. They also detected decreased diversity in riffles to which sediment had been added. Bjornn et al. (1977) found negative correlations between embeddedness and invertebrate density in natural streams in Idaho, but reported that adding fine sediment to natural riffles had little effect on the invertebrate density. Lenat et al. (1981) reported decreased taxa richness in areas affected by fine sediment, and provided evidence that the effect was due to decreased habitable area per unit sample area. From these examples, the selective nature among the invertebrate community in response to habitat reduction and the observed shifts in community structure due to habitat change can be detected. The benthic macroinvertebrates’ selective nature and shifts suggest that metrics other than biomass might serve as better, different, or complementary indicators of sediment pollution (Chanat, 1998).

2.3.2 Benthic Macroinvertebrates for Bioassessment

Multimetric indices derived from biological data are increasingly used to measure the ecological health of streams. The indices consist of a collection of metrics that summarize information from population, community, and ecosystems levels into a single number through bioassessment. Bioassessment is a monitoring technique intended to characterize the overall health of a water body. A water body's health is determined by gathering multiple measures of biological data, converting the data into a single numeric index, then comparing the index with an index developed for a reference condition. Reference conditions are established by characterizing the biology and water quality of reference sites with unimpacted water bodies (Pawlak, 1999).

To a varying degree, water quality, sediment quality, and biodiversity are intimately related (Chapman, 1990; Burton and Scott, 1992; Nelson et al., 1992; Rosenberg and Resh, 1993; USEPA, 1992b). Consequently, a convergence towards the integral study of hydrology and ecology takes shape in watershed management. Biometrics can be used to offer assimilative indication of water quality, to measure overall system's health, and to directly measure valued ecological components of a system placed under widespread management (Burton and Scott, 1992; Chapman et al., 1992; Chapman, 1995; Rosenberg and Resh, 1993). While benthic algae and fish are used in many stream assessments, the benthic macroinvertebrates are the most commonly used taxonomic group because they live in close association with the substrate. Southerland and Stribling (1995) reported that more than 85% of State water quality agencies in the United States used some form of multimetric biocriteria to monitor their aquatic resources. Ninety percent of those programs used benthic macroinvertebrates.

Advantages to using benthic macroinvertebrates for monitoring stream quality include: a) macroinvertebrates are good indicators of localized conditions; b) they can integrate the effects of short-term environmental variations since sensitive life stages will respond quickly to stress, while the community responds slowly; c) degraded conditions can be detected by an experienced biologist with a quick preliminary examination of the assemblage; d) assemblages are made up of a range of trophic levels and pollution tolerances; e) sampling is easy and requires minimal equipment and people; f) macroinvertebrates are abundant in most streams, unlike fish; and, g) most State water quality agencies focus on macroinvertebrates over fish for biosurveys (Barbour et al., 1999).

2.4 Development of Ecological Assessment Procedures

The endeavor to characterize stream health based upon aquatic communities led to the development of thresholds and ultimately multimetric indices. Karr (1981) published a set of assessment thresholds for fish communities in the streams of Illinois and Indiana. The assessment is known today as the Index of Biotic Integrity (IBI). The IBI was based upon discrete measurements of 12 different metrics of a fish community (Karr et al., 1986). When combined, the metrics provided an overall assessment of the condition of the community. To arrive at the final assessment, each metric was calculated based upon a sample community at a test site. The metric values were assigned a numerical score based upon differences with reference conditions. An overall assessment score could then be calculated by summing the 12 metric scores. The final score would designate the community's integrity as excellent, good, fair, poor, very poor, or no fish.

The IBI led to the development of multimetric indices for benthic macroinvertebrate. In 1986, the Ohio Environmental Protection Agency developed the Invertebrate Community Index (ICI) (OEPA, 1989). ICI embraced the multimetric approach described in the IBI using a composite of 10 structural benthic macroinvertebrate community metrics. The ICI score is favored by ecologists for its lumped, integral simplicity (USEPA, 1996a). In 1989, the US Environmental Protection Agency (USEPA) published a multimetric index using benthic macroinvertebrates and fish. The USEPA called the index Rapid Bioassessment Protocols (RBPs) for use across the nation (Plafkin et al, 1989). Kerans and Karr (1994) then developed a Benthic Index of Biotic Integrity (B-IBI) for the river of Tennessee Valley. The most recent set of guidelines published in conjunction with the USEPA is the guidebook for Rapid Bioassessment Protocols for use in wadeable streams and rivers using periphyton, benthic macroinvertebrates, and fish by Barbour et al. (1999). Of course, numerous other indices have been developed for specific areas. The above-mentioned are the most widely used and known benthic indices.

Multimetric indices are preferred in most cases since the usage of multiple metrics can increase the probability of an accurate diagnosis (Fore et al., 1996). For instance, the ICI includes 10 metrics based primarily upon organism tolerance, while the RBP uses a total of 8 metrics with some being based upon functional groups (Watzin and McIntosh, 1999). In addition to using multiple metrics, the accuracy of the assessment depends upon the assumption that in the absence of impairments, a test site and a reference site have similar benthic macroinvertebrate communities.

To overcome high expenses and time requirements, many multimetric biocriteria use a standardized qualitative approach to make assessments. Conventional assessments often rely on quantitative sampling methods that require replicate samples and detailed statistical analysis. As a result, processing of the quantitative samples is slow and tedious work. Rapid bioassessment programs were introduced with the standardized qualitative sampling procedures to minimize expenses and time. However, the rapid bioassessment approach is primarily a screening tool and was not designed to completely replace traditional quantitative sampling methods. After having been developed by several State water quality agencies, rapid assessment was integrated by the United States Environmental Protection Agency (Plafkin et al., 1989).

2.5 Rapid Bioassessment Protocols

Rapid assessment was developed in the mid-1980's with both federal and state monitoring programs embracing it in the United States. Rapid Bioassessment Protocols (RBPs) use cost-saving techniques that allow many more stream sites to be assessed for a given time and budget. RBPs offer three advantages to water quality assessments (Hannaford and Resh, 1995). First, RBPs broaden the analysis from just water quality to habitat quality. Second, RBPs allow for more study cases to be summarized and for management decisions to be implemented in as little as five working days. The last advantage of RBPs is that the metrics are sensitive to changes that occur in benthic macroinvertebrate community composition when pollution is present. On the other hand, information regarding spatial variability is lost. Consequently, there is no quantitative quality assurance or quality control (QA/QC) when a single sample is taken or a single habitat assessment is done at a test site (Hannaford and Resh, 1995).

The original RBPs (Plafkin et al., 1989) were developed in two phases. The first phase centered on the development and refinement of the benthic macroinvertebrate protocols. The second phase involved the addition of protocols for fish assemblage assessment. Three benthic protocols were developed from procedures in use by various state water quality agencies, RBP I, II, and III. The fish assessment was split into two protocols, RBP IV and V. Many State water quality agencies employ trained and experienced benthic pathologists allowing for the accumulation of considerable background data on macroinvertebrates. However, water quality standards, legislative mandate, and public opinion are directly related to the status of a water body as a fishery source (Barbour et al., 1999).

RBP I was based upon benthic assessment approaches used by Missouri and Michigan Departments of Natural Resources. RBP II was based upon the approach used by the Virginia State Water Control Board in the late 1980s. RBP II was more time intensive, incorporating field sampling and family-level taxonomy. The third protocol, RPB III, was the most rigorous one and incorporated methods from the North Carolina Division of Environmental Management and the New York Department of Environmental Conservation. RBP IV, the first fish protocol, used a questionnaire approach, while RBP V was based upon the IBI and the Gammon's Index of Well Being (IWB).

Refinements to the original RBPs (Barbour et al., 1999) have occurred from regional testing and adaptation by state agency biologists and researchers. One key addition is the usage of periphyton assemblages to monitor water quality. As in the benthic and fish assemblages, integration of structural and functional characteristics provides the best means of assessing impairments. Even though periphyton have not been widely incorporated in monitoring programs, the periphyton represent the primary producer level and exhibit a different range of sensitivities that have ultimately led to their inclusion as an ample monitoring assemblage.

The purpose of biological assessment is to characterize the status of water resources and to monitor trends in the condition of biological communities. Fundamental to all rapid bioassessment methods is the classification of streams so that comparisons can be made between reference areas and areas of concern. In addition, rapid assessment approaches are included to expedite both assessment of water quality and any subsequent management decision. A series of steps can be used in developing and conducting rapid assessment programs. First, classification is done to determine stream class designation and representative sites for each stream class. Second, the biota and habitat must be surveyed. Third, biota is characterized through determination of measures. Fourth, calibration procedures are conducted to differentiate between impaired and non-impaired sites. Fifth, the index is developed. Lastly, biological criteria are developed from the indices (Resh et al., 1995).

Analytical methods described in the RBP involve calculating a percent similarity to the reference site for each metric. Each metric is then assigned an impairment threshold. If the test site's value is lower than the reference's, it is considered impaired. Metrics for the test sites are categorized as no impact, slight impact, moderate impact, or severe impact (Hannaford and Resh, 1995). Virginia's Biological Assessment Program supports the usage of RBP methods since the impairment categories directly correspond with the CWA goal categories of fully supporting, partially supporting, and non-supporting.

2.6 Virginia's Biological Assessment Program

The Virginia Department of Environmental Quality (VADEQ) has been conducting biological assessments since 1978. VADEQ's Biological Monitoring Program utilizes benthic macroinvertebrates to determine overall water quality; however, VADEQ did not adopt the use of U.S. EPA's Rapid Bioassessment Protocol II until 1990. The program is composed of approximately 400 monitoring stations that are examined twice a year, once in the spring and once in the fall. The RBP II method is preferred since it aids in the determination of whether or not waters are being maintained to meet the water quality standards set forth within the CWA. Each sampled station is assigned an impairment classification of non-impaired, slightly, moderately, or severely impaired, while only moderately and severely impaired stations are placed on the 303(d) list as a stream with a benthic impairment. Currently, Virginia still utilizes the RBP method for biological assessments. However, the state is looking to diversify its assessment procedures while still encompassing the fundamental RBP practices of incorporating habitat and metric evaluations in a timely and efficient manner.

2.7 Benthic Rapid Bioassessment Protocols

Three tactics account for the minimal benthic analysis efforts in rapid bioassessments. First, one large sample consisting of two to several collections is taken. Second, a standardized subsampling procedure is used which reduces the number of organisms processed and provides consistency. Third, the results of surveys can be summarized in ways that can be understood by non-specialists.

The assessment results categorize the benthic community as non-impaired, slightly, moderately, or severely impaired. A non-impaired water has a benthic community comparable to other undisturbed streams within the region. The community should be characterized by a maximum taxa richness, balanced taxa groups, and a good representation of intolerant individuals. Slightly impacted waters have small deviations from the reference site in terms of the taxa present. Moderately impaired waters have reduced macroinvertebrates richness. The taxa composition changes result in reduced community balance and ultimately absent taxa. Severely impaired waters have a dramatic change in the benthic community. Macroinvertebrates are dominated by only a few tolerant taxa that are abundant (NJDEP, 2000).

Benthic RBPs are primarily a biosurvey technique with composite biological samples collected for multimetric analysis. However, habitat assessment is also an integral component of the final evaluation of impairment. All of the habitat parameters evaluated are related to the overall aquatic life and are a potential source of limitation to the aquatic biota.

2.7.1 Habitat Assessment and Physiochemical Parameters

Geomorphic, hydrologic, and water quality parameters affecting the invertebrate community can be related, making isolation of the effect of a single agent difficult (Allan, 1995). Consequently, habitat classifications should be considered when assessing aquatic ecology. An understanding of the relationships between habitat parameters is important for two reasons. First, this understanding helps avoid making misleading conclusions about the importance of a single agent. Second, it aids interpretation of quantitative analyses; parameter estimates in statistical models which incorporate correlated predictors can have highly inflated standard errors (Ott, 1993).

The physical relationships between width, depth, velocity, discharge, stream order, gradient, median bed material particle size, downstream distance, and drainage area are well documented (e.g., Dunne and Leopold, 1978; Knighton, 1984; Ritter et al., 1995). These characteristics, in turn, influence physiochemical properties such as temperature (e.g., Ward, 1985), dissolved oxygen (e.g., Hem, 1992), dissolved solids (e.g., Webb and Walling, 1992), and pH (e.g., Hem, 1992).

At the reach scale, variations in channel slope correspond with variations in cross-sectional area, bed materials, and sediment transport capacity (Frissell et al., 1986). Aspect and riparian shading affect available light, which is important for the growth of algae and aquatic macrophytes (Allan, 1995).

At the microhabitat scale, substrate type is closely related to local hydraulic forces such as depth, velocity, and slope. These factors, along with substrate composition, determine the conditions under which bed material is entrained and deposited (e.g., Knighton, 1984).

The EPA developed RBPs I, II, and III in conjunction with benthic macroinvertebrate sampling. The RBPs advocate an integrated assessment, comparing habitat and biological measures with empirically defined reference conditions (Plafkin et al., 1989; Karr et al., 1986). The habitat scoring and field survey methods used for RBP II are listed in Table 2.1.

The habitat characteristics are all assessed on a visual basis. Habitat evaluations are first made on instream habitat, followed by channel morphology, bank structural features and riparian vegetation. All parameters are evaluated and rated on a numerical scale of 0 to 20 for each sampling reach. The ratings are totaled and compared to a reference condition to provide a final habitat ranking. Scores increase as habitat quality increases.

Table 2.1 The habitat characteristics and field survey metrics incorporated in RBP II (Adapted from Plafkin et al., 1989; Barbour et al., 1999)

EPA RBP Qualitative Field Habitat Scoring Indices	**EPA RBP Associated Field Survey Metrics**
Instream: 1. Substrate & Available Cover 2. Embeddedness 3. Flow/Velocity	***Instream:*** 1. % Physical Substrate: Bedrock, Boulder, Cobble, Gravel, Sand, Silt, Clay 2. % Organic Substrate: Detritus (CPOM), Muck-Mud (FPOM), Marl (Shell) 3. Run Velocity 4. Land Use 5. Local NPS Pollution 6. Sediment Odors 7. Sediment Oils 8. Sediment Deposits 9. Anaerobic Conditions
Physical: 4. Channel Alteration 5. Bottom Scouring & Deposition 6. Pool/Riffle or Run/Bend Ratio	***Physical:*** 10. Average Stream Width 11. Average Stream Depth: Pool, Run, Riffle 12. High Water Mark 13. Dam Present 14. Channelization
Structural: 7. Bank Stability 8. Bank Vegetation 9. Streamside Cover	***Structural:*** 15. Relative Canopy Cover 16. Erosion

The field survey metrics represent the physical characteristics of the stream habitat. Physical characterization includes documentation of substrate and instream features, stream characterization and watershed features, and riparian vegetation. The information collected supplements the biological surveys. The physical characterization provides insight to the ability of the stream to support a healthy aquatic community, and to the presence of stressors to the stream ecosystem.

2.7.2 Biological Assessment

In addition to the habitat and physical field parameters, a biosurvey component is incorporated into the RBP process. The collection procedure for RBP II focuses on a riffle/run sample and a coarse particulate organic matter (CPOM) sample. The riffle/run habitat is the most productive habitat available in stream systems and includes many pollution sensitive taxa of the Scraper and filtering Collector Functional Feeding Groups. The CPOM sample provides a measure of effects of the Shredders. After the samples are sorted and the organisms are identified, the raw benthic data are used to calculate metrics. Plafkin et al. (1989) lists the following eight metrics as the measurements used in the RBP II assessment procedure to characterize the biological condition:

1.) Taxa Richness

Measures the total number of families present. Generally increases with increasing water quality, habitat diversity, and habitat suitability.

2.) Modified Family Biotic Index

The index was developed to detect organic pollution and is based on the original species-level index. Tolerance values range from 0 to 10 for families and increase as water quality decreases. The formula for calculating the index is:

$$\text{FBI} = \sum \frac{x_i t_i}{n} \qquad [2.2]$$

where: x_i = number of individuals within a taxon,
t_i = tolerance value of a taxon, and
n = total number of organisms in the sample.

3.) Ratio of Scrapers and Filtering Collectors

The ratio reflects the riffle/run community foodbase. It is an indication of the periphyton community composition, availability of suspended Fine Particulate Organic Material, and availability of attachment sites for filtering.

4.) Ratio of EPT and Chironomidae Abundances

The indicator groups Ephemeroptera, Plecoptera, Trichoptera, and Chironomidae are measures of community balance. An even distribution among the four groups with substantial representation of the sensitive EPT groups reflects a good biotic condition.

5.) Percent Contribution of Dominant Family

This is the abundance of the numerically dominant taxon relative to the rest of the population as an indication of community balance at the family-level. A community dominated by relatively few families would indicate environmental stress.

6.) EPT Index

The EPT Index value is the total number of distinct taxa within the groups Ephemeroptera, Plecoptera, and Trichoptera. The index value summarizes the taxa richness within the insect groups that are generally considered pollution sensitive. The EPT Index generally increases with increasing water quality.

7.) Community Similarity Indices

Community Similarity Indices are used in situations where a reference community exists. The indices are designed to be used with either species level identifications or higher taxonomic levels. Three possible indices are:

i.) Community Loss Index – measures the loss of benthic taxa between a reference station and the station of comparison

ii.) Jaccard Coefficient of Community Similarity – measures the degree of similarity in taxonomic composition between two stations in terms of taxon presence or absence

iii.) Pinkham and Pearson Community Similarity Index – incorporates abundance and compositional information

8.) Ratio of Shredder Functional Feeding Group and Total Number of Individuals Collected

Using the CPOM sample, the abundance of the Shredder Functional Group relative to the abundance of all other Functional Groups allows for the evaluation of the potential impairment. Shredders are good indicators of landuses and toxic effects when the toxicants involved are readily absorbed to the CPOM.

The calculated metrics are compared to values derived from a reference site. Each metric is assigned a score according to the comparability of calculated and reference values. Scores for the eight metrics are then summed and compared to the total metric score for the reference station. The percent comparison between the total scores provides a final evaluation of biological condition. When the final value falls within an intermediate range, habitat assessment and physical characterization may aid in the evaluation process.

Even though the RBP aids in relating the water quality to the biological conditions, certain watershed characteristics such as soil type, topography of the landscape, climate, land use patterns, and types of farming practices and associated conservation efforts cannot be directly characterized in the bioassessment procedure (Watzin and McIntosh, 1999). Accordingly, the emergence of Geographical Information Systems (GIS) has enhanced the ability to examine agricultural pollution at the landscape level.

2.8 Geographical Information Systems

Geographical Information Systems (GIS) are popular tools in watershed management applications, along with increased focuses in areas of science and engineering. Cartographers began to adopt computer techniques in the 1960s with considerable investments in the development and application of computer-assisted cartography taking shape by the late 1970s. After twenty years of technical development, GIS have become a worldwide phenomenon. In 1995, it was estimated that 93,000 sites worldwide had installed GIS technology (Burrough and McDonnel, 1998). These technical solutions are utilized under various circumstances such as for soil science, rural and urban planning, utility networks, remote sensing, hydrologic modeling, and many other spatially oriented applications.

GIS are defined by Burrough and McDonnel (1998) as "*a powerful set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world*". In short, GIS are computer-based tools that allow for the input, data management, manipulation and analysis, and output of georeferenced data. The spatial analysis capabilities distinguish the GIS from other graphics oriented systems like computer aided design and drafting (Aronoff, 1993). Despite their spatial analysis capabilities, GIS lack decision support systems. The GIS represents a powerful tool to gain knowledge required to make the decision (Cowen, 1998). The main advantage in using GIS is that it allows for the development of numerous different modeling approaches due to the substantial time saving and data handling capabilities (Fisher, 1989).

2.8.1 Sediment Modeling Within a GIS Environment

The Universal Soil Loss Equation (USLE) is generally used to predict the amount of sediment loss from individual fields. However, water quality analysis requires knowledge on the amount of sediment that impacts the stream. Consequently, a delivery ratio is often applied to the soil loss estimates either at a field level or for an entire watershed. By using the erosion rates from the USLE and a delivery ratio value based on landuse and topographic characteristics surrounding the cell, the final sediment yield value can be predicted. Predicting sediment yield based upon a load calculated at a field or cell requires a model which can incorporate the on-site load with off-site characteristics to determine the stream impact (Kleene, 1995).

The Universal Soil Loss Equation (USLE) originated from a method devised from 1945 to 1965 which estimated soil losses based on statistical analyses of field data from small plots located in many states. The method is useful in estimating sediment losses and for predicting nonpoint source sediment losses in pollution control programs (Schwab et al., 1993). Utilizing input parameters for landuse, topography, soil characteristics, and rainfall intensity, the amount of soil loss per unit area due to erosion by rain is calculated. To arrive at the annual soil loss value for a homogenous field or cell area, the equation is as follows:

$$A = R\ K\ L\ S\ C\ P \qquad [2.3]$$

where:
A = annual soil loss (Mg/ha/yr or tons/acre/yr),
R = rainfall and runoff erosivity index (Mg/ha/yr or tons/acre/yr),
K = soil erodibility factor,
L = slope length factor,
S = slope steepness factor,
C = cover management factor, and
P = conservation practice factor.

The usage of a delivery ratio with the gross sediment loss estimated by the USLE takes into account deposition of sediment during travel through the watershed. The delivery ratio is the ratio of the actual sediment yield from a watershed to the gross erosion load occurring on that watershed. Numerous methods have been proposed to estimate sediment delivery ratio such as those presented by Roehl in 1962 (Barfield et al., 1983) and by Williams and Berndt (1972).

However, Shanholtz (1988) developed a delivery ratio that can be used within the VirGIS database (Kleene, 1995).

2.8.2 Water Quality Index

The Virginia Geographic Information System (VirGIS) project began in 1985 as part of an agricultural nonpoint source pollution control program for the cleanup of the Chesapeake Bay. The Information Support Systems Laboratory (ISSL) at Virginia Tech was contracted by the Department of Conservation and Recreation, Virginia Division of Soil and Water Conservation to develop VirGIS. The VirGIS system includes an extensive array of software for cartographic analysis, data entry, analysis, and display and a digital geographic database covering over fifty counties. Reports were organized by county to describe the VirGIS procedures (Shanholtz et al., 1990).

The Water Quality Index was an index developed for rankings agricultural nonpoint source pollution. The USLE with a delivery ratio was chosen as the most appropriate modeling technique due to the availability of data, computational efficiency, and its acceptance and use as a planning tool by the USDA in Virginia. GIS technology was selected for creating, storing, analyzing, and displaying information. The elevation, soils, landuse, and hydrography layers were used to derive a spatially varying filter to represent the ratio of sediment delivered from a given field location to the nearest downstream channel. Consequently, the Water Quality Index emerged as a method of predicting the potential impact to a surface water body from nonpoint sources of pollution, based upon the following sediment yield equation:

$$A_{Di} = A_i * DR_i \qquad [2.4]$$

where:
A_D = potential sediment delivered to a stream (tons/acre/yr),
A = annual sediment loss from the USLE equation (tons/acre/yr),
DR = delivery ratio, and
i = cell count.

The delivery ratio was approximated using a first order function that included the influence of cover, and the steepness and length of the flow path (Shanholtz, 1988).

$$DR_i = e^{-kdS_f} \qquad [2.5]$$

where: k = coefficient that varies with land cover,
d = distance to stream, and
S_f = slope function.

By assuming DR as 1 when d equals zero and DR as 0.184, 0.055, and 0.009 for cropland, pastureland, and nonagricultural land, respectively, at a distance of 1312 feet, the following parameters were derived for Equation 2.5.

$k_1 = -0.4233$ (cropland)
$k_2 = -0.71$ (pastureland)
$k_3 = -1.1842$ (nonagricultural land/wooded)

The slope function was based on a relationship developed by Heatwole, et al. (1987).

$$S_f = e^{-n(S + S_0)} + S_{fmin} \qquad [2.6]$$

where: $n = 16.1$,
S = slope,
$S_0 = 0.057$,
$S_{fmin} = 0.60$.

The potential sediment delivered to the water body was derived based upon soils, landuse, surface water, elevations, watersheds, county boundaries, and SWCD boundaries GIS data layers. It should also be noted that Equation 2.4, the sediment yield equation, applies only to agricultural areas and that the Equations 2.4 and 2.6 have never been validated with field data.

2.9 Literature Review Summary

Water quality standards are implemented in order to reduce negative impacts of pollution on receiving waters. When environmental groups became concerned about the effectiveness of water quality monitoring procedures, Total Maximum Daily Loads (TMDLs) were introduced as a method to achieve the water quality goals set forth in the Clean Water Act.

In Virginia, benthic impairments rank fourth among all water impairments in the state. Since geomorphic, hydrologic, and water quality parameters all affect the invertebrate community, it is hard to pinpoint a single pollutant that is responsible for the benthic impairments. The development of relationships between stressors and benthic macroinvertebrate assemblage conditions would aid in the development of benthic TMDLs that could restore a stream's biological integrity.

A review of the literature on benthic macroinvertebrates indicated that macroinvertebrates can be used to quantify the effects of pollutants on water quality. Benthic macroinvertebrates are sensitive to a wide range of variables within a watershed from sediment to landuse, habitat, watershed characteristics, and water quality parameters. Pollution affects the invertebrate community by altering water movement, habitat, food quality, and oxygen availability. Accordingly, biological monitoring investigates the effects of pollution upon invertebrates by incorporating diverse metrics that are sensitive to changes in the invertebrate community in the presence of pollution. Aquatic monitoring was further enhanced with the development of the rapid bioassessment procedures that allowed assessments to be less expensive and time consuming. As a result, EPA's Rapid Bioassessment Protocol (RBP), a rapid bioassessment approach, was embraced both federally and statewide since the program not only saved time and money, but incorporated habitat, water quality, and biological monitoring all in one program.

Even though biological monitoring relates water quality to biological conditions, certain watershed characteristics cannot be directly characterized in the bioassessment procedure. To overcome this, GIS can be used to examine pollution at the landscape level. Information on topography, landuse patterns, and sediment loads can all be quantified within a GIS environment. The combined efforts between biological assessment and GIS characterization will allow for the development of relationships between benthic stressors and the benthic community.

CHAPTER 3.0 METHODS

To develop relationships between benthic macroinvertebrate assemblages and sediment or other stressors, four steps were completed. First, existing data were compiled, prepared and converted to usable forms. Second, data analyses and organization were completed to obtain important parameters for each station. Third, the data collected in the first two steps were evaluated using statistical analysis to develop relationships between the benthic macroinvertebrate community and their possible stressors. Lastly, the relationships were analyzed in terms of their application to the development of benthic TMDLs in Virginia.

3.1 Compilation of Data

Sediment, habitat parameters, water quality parameters, landuse characteristics, watershed characteristics, and livestock numbers are all possible sources of stress for the benthic macroinvertebrates following the stressor identification guidelines proposed by the EPA (USEPA, 2000c). To develop relationships between benthic macroinvertebrates and their stressors, benthic data, water quality data, GIS data, and livestock numbers were gathered and analyzed. The sources for the data along with a description of the type of data are described in detail in the following sections.

3.1.1 Benthic Data

The Virginia Department of Environmental Quality (DEQ) was the primary source for benthic and water quality data. DEQ collects benthic data at approximately 400 sites in Virginia and uses the U.S. EPA's Rapid Bioassessment Protocol II for the assessment of streams (USEPA, 1996b; Plafkin et al., 1989). The benthic data contain habitat variables, water variables, information on the stations, and background information on the biological data for the macroinvertebrates (Darkwah, 2000).

The DEQ office in Richmond provided benthic data from their BIOMON database. The database is compromised of 31 files covering 418 stations from the fall of 1994 to the fall of 1998. The files are accessed using FoxPro; however, if FoxPro software is unavailable, the database files in the format of *.dbf can be opened as spreadsheets in Microsoft Excel. The field files of main interest for this study were STATION, SAMPLE, EPAHAB, and WATER (Darkwah, 2000).

The STATION file contains information on the location of the biological monitoring stations using roads, latitude and longitude, county, DEQ region, etc. In addition to location information, the STATION file lists stream order, predominant landuse type, and a survey reason. The item of main importance from this file was the latitude and longitude values that enabled the incorporation of station points within GIS so as to accurately delineate watershed boundaries. Other notable uses of the STATION file were the stream order information used as an additional stressor variable as well as the predominant landuse and the station monitoring reasons, which aided in the determination of usable stations for this study.

The SAMPLE file provided all of the biological information. The file included date and time of sampling, reference station information, the individual metric calculation values and scores, the final score, and the RBP II impairment classification along with DEQ's impairment assessment and a comment column that would justify DEQ's assessment or provide additional information on the site. For the 1994 to 1998 period, a total of 1,431 samples were taken from the 418 stations. Table 3.1 summarizes the benthic impairments for all samples.

Table 3.1 The impairment classifications for DEQ's benthic samples

IMPAIRMENT CLASSIFICATION	NUMBER OF SAMPLES					TOTAL
	1994	1995	1996	1997	1998	
Could Not Calculate	44	50	60	61	55	270
Non-Impaired	69	98	98	140	114	519
Slightly Impaired	0	24	0	0	15	39
Moderately Impaired	70	112	114	167	109	572
Severely Impaired	2	11	6	8	4	31
TOTAL	185	295	278	376	297	1431

The five impairment classifications shown in Table 3.1 correspond to the RBP II classifications. However, the 'Could Not Calculate' classification was added in by DEQ when the methodology used in the RBP II process failed to arrive at a tangible impairment classification. Numerous reasons may have caused the inconclusive results. For instance, the evaluation may have been inconclusive if sampling occurred during flood or drought conditions. To compensate, DEQ provides a column that indicates what they believe the final impairment classification would be and a second column that provides comments to justify their assigned impairment classification. In this case, if the RBP II procedure produced inconclusive results, DEQ's assigned impairment classification was used.

Overall, most samples in Table 3.1 were either non-impaired or moderately impaired throughout the five-year period shown. Very few samples were slightly or severely impaired. This lack of variety among impairment classification distributions could pose a problem when conducting data analysis. Ideally, values from all classifications need to be evenly incorporated in order to accurately justify the results. Hence, the more diversified the original data are, the more meaningful the statistical results will be (Ott, 1993).

The locations of the original benthic stations as given in DEQ's STATION file are shown in Figure 3.1. The majority of the stations in Figure 3.1 fall within the West Central, Valley, Northern, and Piedmont DEQ regions. Even though numerous benthic stations exist in the Southwest region, DEQ's database does not have latitude and longitude measurements available for all Southwest sampling stations. Consequently, the Southwestern stations with unknown locations are not shown in Figure 3.1. Moreover, the locations in Figure 3.1 were preliminary and adjustments were needed in determining the exact location of the stations and in selecting suitable stations for inclusion in the study.

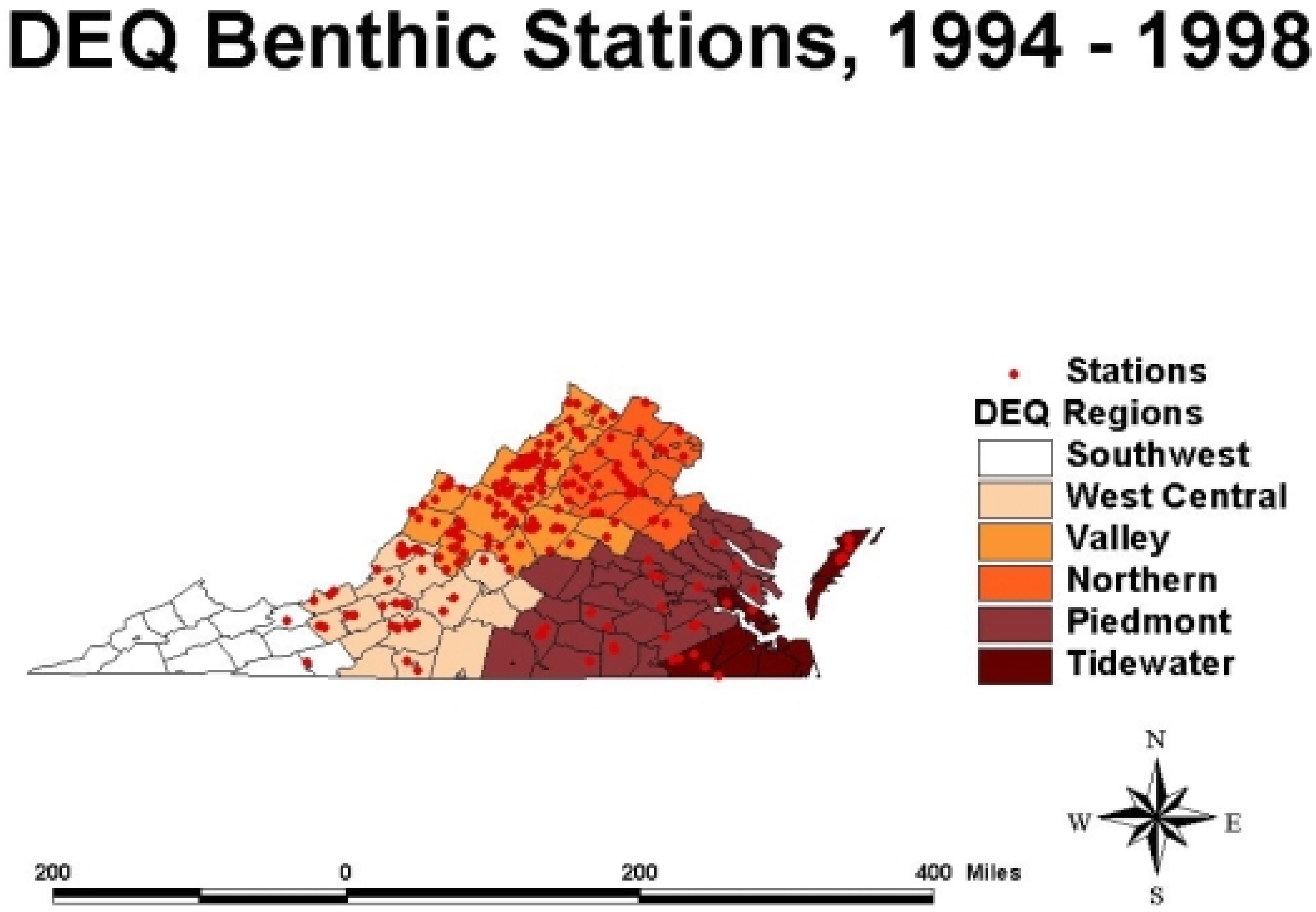

Figure 3.1   The locations of DEQ benthic stations in Virginia (Darkwah, 2000)

The EPAHAB data file provided the habitat variables for all samples based upon the guidelines established by EPA (Plafkin et al., 1989). The habitat variables are visually assessed by the biologist at the time of their visit. The parameters are rated on a numerical scale of 0 to 20 for each sampling reach with scores increasing as habitat quality increases. For instance, optimal conditions receive a score between 20 and 16, suboptimal conditions range from 15 to 11, marginal is between 10 and 6, and poor conditions fall between 5 and 0 (Barbour et al., 1999). The habitat variables that were included as possible variables affecting benthic health were: substrate, cover, embeddedness, flow, velocity, channel alteration, sediment, riffle, bank stability, bank vegetation, graze, and riparian vegetation. A general description for the variables provided by Barbour et al. (1999) are summarized as follows.

- Substrate

  This parameter refers to the bottom substrate that is available to provide habitat and support of aquatic organisms. The presence of rock and gravel in flowing streams is generally considered the most desirable habitat for aquatic life. The bottom substrate is evaluated and rated by observation.

- Cover

  Cover refers to the available cover for habitat and support of aquatic organisms. Logs, tree roots, submerged vegetation, undercut banks, etc. provide the niches required for community support. As variety and abundance of cover decreases, habitat structure becomes monotonous, diversity decreases and the potential for recovery following disturbance decreases.

- Embeddedness

  The degree to which boulders, cobble, or gravel are covered or sunken into silt, sand, or mud of the stream bottom is measured by the embeddedness parameter. This parameter can indicate suitability of the stream substrate as habitat for benthic macroinvertebrates. As rocks become embedded, the surface area available to macroinvertebrates is decreased.

- Flow

  Relates to the ability of a stream to provide and maintain a stable aquatic environment. Flow is the predominant constraint in low flow areas of ≤ 0.15 m^3/s (5 ft^3/s). Flow is a measure of how much the channel is filled with water, especially as related to exposed substrate that would be suitable habitat for aquatic organisms if covered by water. The flow should be estimated for small streams in a straight stretch where banks are parallel and bottom contour is relatively flat.

- Velocity

  Larger streams that have a flow greater than 0.15 m^3/s directly influence the structure of the benthic community through velocity and depth. The best streams will have all 4 patterns present: slow-deep, slow-shallow, fast-deep, and fast-shallow. The general guidelines are 0.5 m depth to separate shallow from deep, and a velocity of 0.3 m/sec to separate fast from slow.

- Channel Alteration

  A measure of large-scale changes in the shape of the stream channel. The severity of watershed and bank erosion can be detected through characterization of the sediment deposits. Sediment bars tend to increase in depth and length with continued watershed disturbance. The presence of deposition on the inside of bends, below channel constrictions, or where stream gradients flatten out also indicates channel alteration. Channelization is undesirable since it decreases stream sinuosity, which increases stream velocity and the potential for scouring.

- Sediment

  This parameter refers to bottom scouring and deposition. Deposition occurs from large-scale movement of sediment. The degree of siltation in pools and riffles is measured by estimating a percentage of a reach that is scoured or silted. Transport of sediment or other particulates may be an indication of large-scale watershed erosion.

- Riffle

  Riffles are assumed to provide more diverse habitat than a straight run or uniform depth stream. The high degree of sinuosity provides for diverse habitat and fauna. The pool to riffle ratio is measured by dividing the average distance between riffles by the average stream width.

- Bank Stability

  Bank stability refers to the potential detachment of soil from the upper or lower stream bank and its potential movement into the stream. Streams with poor banks generally have poor instream habitat. Steep banks are more likely to collapse and suffer from erosion.

- Bank Vegetation

  Measures the amount of cover of the streambank and riparian zone by naturally growing native vegetation. Plant roots, boulders, cobble, and gravel all minimize bank erosion. Vegetation increases erosive resistance of the streambanks, helps control instream scouring, provides nutrients to aquatic species, reduces contamination of runoff water, controls runoff volumes and velocities entering the stream, provides wildlife habitat, and controls extreme fluctuations in water temperature.

- Graze

  Streamside cover consisting of grass is measured along the stream bottom, bank, and top of bank. The grass can provide habitat and cover for the benthic community.

- Riparian Vegetation

  A measure of the streamside cover vegetation width from the edge of the stream bank out through the riparian zone. The vegetative zone provides stream shading, cover, refuge, and controls erosion. Riparian vegetation dominated by shrubs and trees provides the coarse particulate organic matter source. The amount of riparian vegetation is measured on the stream bottom, bank, and top of bank.

The last biomonitoring field file of importance, the WATER file, contained water quality parameters measured by DEQ biologist at the time when benthic conditions were assessed. Water quality information was used to provide insight on the ability of a stream to support a healthy aquatic community, and to detect the presence of chemical and non-chemical stressors to the stream ecosystem (Barbour et al., 1999). DEQ used water quality instruments to sample temperature, dissolved oxygen, conductivity, pH, salt, and chlorine levels. Only the temperature, dissolved oxygen, and pH values from the WATER file were used in this study since these three variables pose the greatest threat to the macroinvertebrate assemblages and are known to be highly accurate when measured in the field (Hem, 1992; Ward, 1985). In addition to temperature, dissolved oxygen, and pH, other parameters such as nutrients and turbidity also impact the macroinvertebrate community. Since the biological data did not investigate all possible water quality parameters, an additional water quality data source, DEQ's Ambient Water Quality Monitoring Reports, was investigated due to its regulatory applicability (water quality standards, VPDES permits, TMDLs, etc.) and for its accessibility.

3.1.2 Water Quality Data

The DEQ's 1999, 1998, and 1997 Ambient Water Quality Monitoring Reports were used to obtain important water quality parameters (VADEQ, 1999; VADEQ, 1998; VADEQ, 1997). Each of the three reports are available for download on the Virginia Department of Environmental Quality's website at: http://www.deq.state.va.us/water/my99rpt.html in addition to being available from the regional DEQ Offices, the Water Quality Assessment Office in Richmond, or the DEQ Chesapeake Bay Program Office in Richmond. These reports provide information to Virginia's citizens on stream, reservoir, and estuary conditions and serve as an important water quality management tool. The data presented cover the majority of the parameters sampled through the Ambient Water Quality Monitoring, Biological Monitoring, and Chesapeake Bay Water Quality Monitoring Programs. The reports provide a summary of the data collected at water quality monitoring stations throughout the Commonwealth.

The mean annual value for chemical variables, along with total suspended solids (TSS) were extracted from the DEQ Ambient Water Quality Monitoring Reports. The variables used from the reports were: dissolved oxygen, temperature, pH, hardness, alkalinity, conductivity, turbidity, and TSS. Total phosphorus and NH_3 levels were included first. However, the nutrient values tended to be the same for the stations used in this study. The phosphorus values were generally 0.1 mg/L, while NH_3 were 0.04 mg/L. Thus, the nutrient variables were eliminated from the data set due to insufficient variability. The dissolved oxygen, temperature, and pH values presented in the reports were not used if they were available from the biological WATER file. Since the WATER file had data corresponding with the exact date and time of biological sampling, the WATER file values were more accurate than the average values presented in the ambient water quality reports.

Upon compiling the water quality data, all water quality variables, besides TSS and turbidity, were included as additional variables that may impact the benthic community. TSS and turbidity were used as possible sediment indicators and were investigated accordingly. It is important to note that all values obtained from the water quality reports were averages for each station over a year. The annual reports summarize data from the previous year's fall up till that year's spring. For instance, the 1997 water quality report covered July 1, 1996 to June 30, 1997. This time period corresponds with the fall 1996 and the spring 1997 benthic data. Thus, each station would have identical chemical values for both the fall of 1996 and the spring of 1997. In addition, the available ambient monitoring water quality reports obtained from DEQ covered a shorter time frame than the benthic data. The annual water quality reports were from 1997, 1998, and 1999, which only correspond with the fall of 1996 through the fall of 1998 benthic data. Consequently, the benthic data from the fall of 1994 to the spring of 1996 did not have corresponding water quality values.

3.1.3 GIS Data Layers

The third important data source, GIS data layers, was needed to completely analyze each station's watershed characteristics. Watershed characteristic parameters such as stream length, watershed size, and percent landuse were collected within GIS and used as potential stressors to the benthic macroinvertebrates. GIS data for each station were obtained from the Virginia Department of Conservation and Recreation (VADCR), the Division of Soil and Water Conservation (DSWC). Data layers for landuse, hydrography, and sediment loading were required. The sediment load file of choice was the Water Quality Index (WQI) within VirGIS. WQI can be used to estimate agricultural sediment loads from selected watersheds using the USLE equation together with a delivery ratio as described in Section 2.8.2. The uncategorized WQI, a file that is not ranked by order of magnitude within a specific geographic unit, was used over the categorized version for its applicability statewide. Since VirGIS was used for the sediment data layer, landuse and hydrography data were also obtained from VirGIS (Shanholtz et al., 1990).

The VirGIS files for landuse, water, and WQI uncategorized were converted to Arc ASCII files (*.aia). An additional detailed landuse file was exported as a GRID (*.e00). All files were placed in the Universal Transverse Mercator (UTM) plane coordinate system NAD27 and compressed with Lempel-Ziv before being sent by VADCR.

Color infrared photography from the Agricultural Stabilization and Conservation Service's (ASCS) National High Altitude Program imagery was used for defining the agricultural landuse. The VirGIS landuse file contains three classifications numbered as follows:

(1) non-agricultural,
(2) cropland, and
(3) pasture.

Non-agricultural land was defined as those areas not classified as crop or pasture and generally included forest, industrial, commercial, residential, cemeteries, bodies of water, roads, etc. Cropland was defined as tilled land, bare, or stripped soil. Pasture was defined as areas with year round cover that are not tilled and include permanent hay, grazed land and idle fields with good cover.

The surface drainage data layer was created by digitizing all surface water features delineated on USGS 7-1/2 minute quadrangle maps. VirGIS software is rasterized to a 1/9 hectare cell size. The VirGIS water file generally contains five classifications numbered as follows:

(0) land;
(1) marsh areas;
(2) linear water features;
(3) aerial water features; and
(4) gravel pits, dug ponds, lagoons, etc.

The exported GRID detailed landuse file, mrlc*.e00, was developed by personnel at the EROS Data Center (EDC, 1997). The Multi-Resolution Land Characteristics Consortium data set was developed to generate a generalized and consistent land cover data layer for EPA Region III. The project was initiated during the summer of 1995 with Version 1 completed in February 1996, Version 2 several months later, and Version 3b was completed in February 1997. The data set file for Version 3b, the file used in this study, contains fifteen different classifications:

(1) water,
(2) low intensity developed,
(3) high intensity residential,
(4) high intensity commercial/industrial,
(5) hay/pasture,
(6) row crops,
(7) other grass (lawns, city parks, golf courses),
(8) evergreen forest,
(9) mixed forest,
(10) deciduous forest,
(11) woody wetland,
(12) emergent herbaceous wetland,
(13) bare (quarries, strip mines, sand pits),
(14) bare (bare rock and sand), and
(15) bare (transitional).

An additional GIS data layer utilized in this study was the roads file. As a means of including additional variables, the road data layer was needed to complete the livestock number calculations. The roads data file was obtained for each county using the U.S Census Tiger 1995 Data (ESRI, 1999). The data layer, provided by ESRI, was downloaded online as the ‘line features, streets’ option.

3.1.4 Livestock Numbers

Animal numbers were included in this study as possible benthic stressors since livestock that enter the stream contribute to bank erosion and ultimately sediment loadings. The objective in estimating animal numbers was to analyze livestock that are provided with stream access such as beef and milk cattle. To complete the livestock number analysis, the data file entitled animrev.dbf was obtained from the Virginia Department of Conservation and Recreation (VADCR, 1996). The file contained livestock estimates for Virginia based upon HUP and FIPS. Beef, milk, hogs, sheep, chicken, broilers, turkey, and horses were all included in the data file. Out of the eight livestock types, generally, only beef and milk cattle are provided with stream access. Accordingly, only the beef and milk livestock numbers were considered in this study.

The initial estimates of the beef and dairy cattle populations in each watershed were made by VADCR in 1996 by averaging the estimates of the populations from the 1987 and 1992 Agricultural Census and disaggregating these numbers to the hydrologic unit or watershed level. The numbers then needed to be determined for each subwatershed. The modifications made to the numbers are based upon the procedures used by Virginia Tech in the preparation of the Big Otter River fecal coliform TMDL (Al-Smadi et al., 2000).

To evaluate if livestock numbers impacted the benthic macroinvertebrate assemblage conditions, only the beef cattle numbers were calculated first. If the beef cattle numbers were found to be consequential to the benthic macroinvertebrate community, the dairy numbers would be evaluated. The beef cattle numbers were first adjusted based upon the amount of pasture acreage difference between the individual subwatersheds and the entire watershed. The resulting numbers were adjusted further based upon the stream access in a subwatershed. It was assumed that in order for the cattle to have stream access, the cattle must be on pasture that is not separated from the stream by other landuse types and must be within 1640 ft (500m) of streams. In addition, when a road passed through a pasture, only the pasture between the stream and road could be used. Pasture separated from a stream by a road would not provide the cattle with direct stream access. Moreover, the beef numbers were adjusted to compensate for both watershed size differences and available stream access.

Upon reviewing the beef cattles’ influence on the benthic macroinvertebrate community, no correlation existed between the two. Accordingly, the investigation of dairy numbers was not completed for this study. Only beef cattle were investigated as a livestock that might negatively impact the benthic macroinvertebrate assemblage conditions.

3.1.5 “Usable” Benthic Data

After collecting all available data, the original benthic data needed to be organized into a subset that contained “usable” data applicable to this study. “Usable” data in this case refer to stations with complete data sets. In order for the stations to have complete data sets, the “usable” benthic data had to meet several criteria. First, the benthic data had to correspond with the water quality data. Since DEQ’s ambient water quality reports covered a shorter time frame than the benthic data, only benthic data corresponding with the water quality reports period from the fall of 1996 to the fall of 1998 could be used in this study. Second, only the sites with a fourth order or smaller stream were investigated to minimize the intricacies of watershed assessment. Generally, as the stream order increases, the watershed assessment process intensifies. For instance, pollutant origin and transport can become difficult to differentiate as the number of pollution sources increases in larger order streams. The third criterion was to include sites with impairments from agricultural and urban activities while also including forested areas. The agricultural and urban areas were included to reflect nonpoint source pollution sources that are responsible for higher levels of sediment loadings. Since agricultural and urban areas were added to constitute impaired sites, heavily forested areas were included to add balance to the impairment classifications. Basnyat et al. (1999) identified forest areas as sinks or transformations zones where pollution is minimized. Thus, forested areas would add non-impaired sites into the “usable” benthic dataset. The last requirement was to attempt to include data from all six DEQ regions. However, regional selection was limited by the distribution of the original benthic data and the availability of GIS data. Both the benthic and VirGIS data were concentrated in the West Central, Valley, and Northern watersheds with limited coverage in the Southwest, Piedmont, and Tidewater. In addition to data availability, studies have found that the furthest reaches of Virginia, the Southwest, Piedmont and Tidewater regions, have distinct differences in benthic compositions when compared to the rest of Virginia due to ecoregion differences (Evans, 1997). Accordingly, only data from the West Central, Valley, and Northern regions were considered in this study. Therefore, the “usable” benthic data set included data from periods corresponding with water quality sampling; sites with fourth order of smaller streams; forested, agricultural, and urban areas; and sites from the West Central, Valley, and Northern regions of Virginia.

The resulting "usable" benthic data set that originated contained 44 stations with a total of 134 samples from the fall of 1996 to the fall of 1998. Table 3.2 summarizes the impairment classifications for the compiled data set.

Table 3.2   The impairment classifications for "usable" benthic samples

IMPAIRMENT CLASSIFICATION	NUMBER OF SAMPLES					TOTAL
	Fall 1996	Spring 1997	Fall 1997	Spring 1998	Fall 1998	
Non-Impaired	7	10	13	9	13	52
Slightly Impaired	0	0	1	0	5	6
Moderately Impaired	14	18	22	7	11	72
Severely Impaired	1	2	0	0	1	4
TOTAL	22	30	36	16	30	134

The results in Table 3.2 demonstrate the same trends as the original database. First, very few samples are slightly and severely impaired. Second, the number of non-impaired and moderately impaired samples are fairly similar.   Streams with moderately impaired benthic conditions are placed on the TMDL list in Virginia. Thus, even though the samples are not evenly distributed throughout the four RBP II classifications, the classification distribution is beneficial for looking at streams that would be on the TMDL list versus streams that are not. Figure 3.2 shows locations of the "usable" stations.

The stations in Figure 3.2 fall within the Northern, Valley, and West Central DEQ regions of Virginia. With the stations clustered within the same general vicinity, ecoregion differences are minimal. The summary table in Table A.1, Appendix A, includes information on the 44 stations such as ecoregions and site characteristics. Overall, it was important to narrow down the benthic database into "usable" stations for this study. Not only was it easier to analyze the attributes of the stations that pertained to the study, but also GIS data layer collection efforts were minimized.

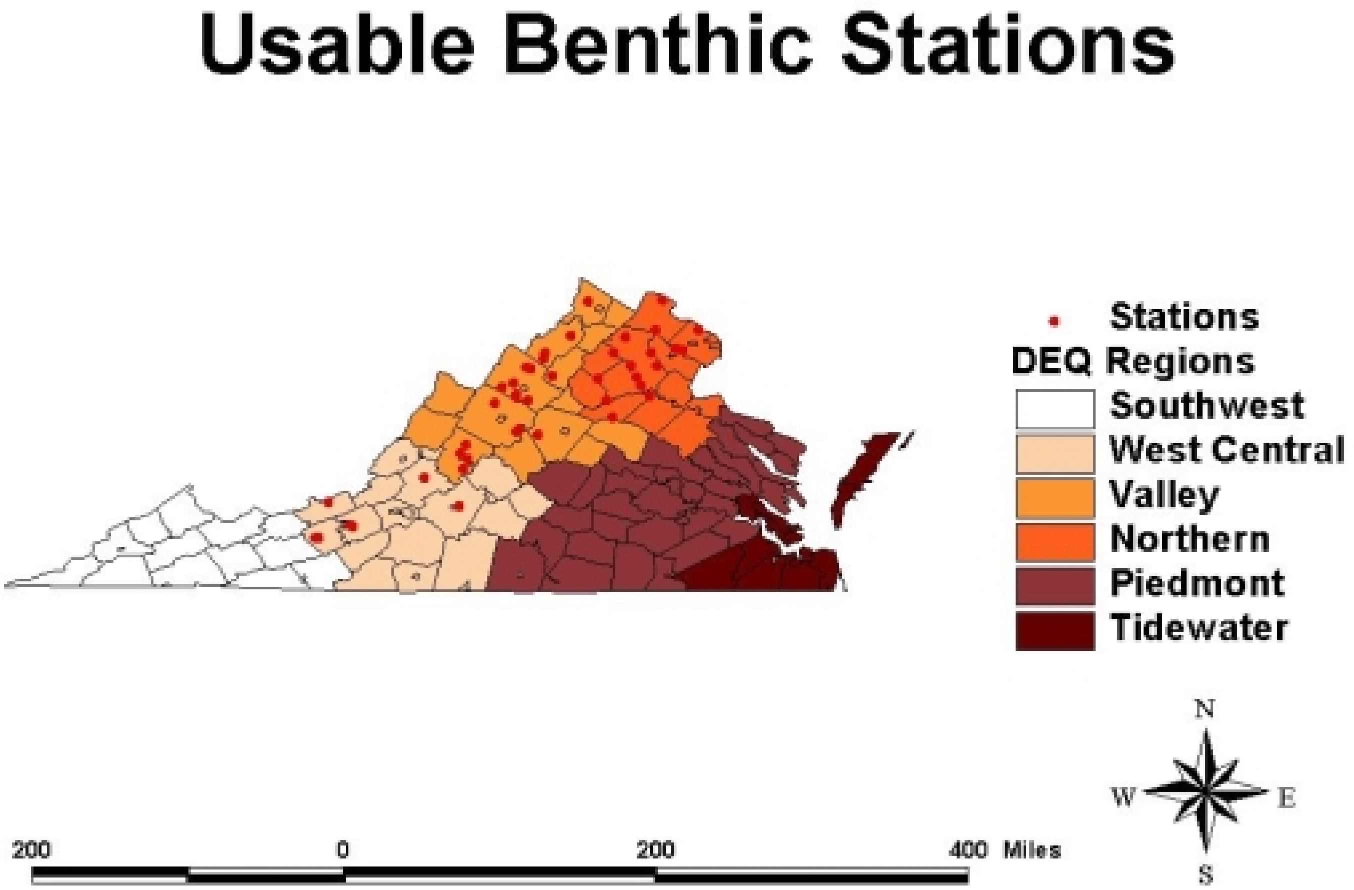

Figure 3.2 The locations of the usable benthic stations in Virginia

3.2 Data Analysis

The intent of this study was to relate benthic macroinvertebrate assemblage conditions to their stressors while placing emphasis upon the effect of sediment loading rates. To achieve this objective, all possible benthic macroinvertebrate stressors needed to be evaluated and the data needed to be organized as stressors or as macroinvertebrate assemblage conditions. This section will review the collection of variables in GIS, the thirty possible independent variables or macroinvertebrate stressors, the six dependent variables or macroinvertebrate assemblage conditions, and the developed data sets used in statistical analyses procedures.

3.2.1 GIS

Within the data collection process, variables such as water quality and habitat were obtained. However, watershed characteristics such as stream length, watershed area, and percent landuse were not readily measurable in the field. Accordingly, ArcVIEW GIS (ESRI, 1998) was used to obtain the additional variables in an efficient and timely manner. To begin analysis in GIS, each station's watershed was delineated based upon its location. From the 44 possible usable stations, 34 stations were used to develop the relationships between benthic macroinvertebrates and stressors. Each of the watersheds containing the 34 stations was delineated. The watershed sizes ranged from 1,175 acres to 77,500 acres with the average delineated watershed measuring 24,100 acres. The delineated watersheds were then investigated one at time to find percent landuse using the Multi-Resolution Land Characteristics Consortium (mrlc) file. The landuses of interest were pasture, cropland, agriculture, forest, and urban. The pasture and cropland areas were combined and classified as agriculture area while the %urban areas corresponded to a % impervious measure. Once the landuses were completed, the total watershed area and stream length were measured and the livestock numbers for each subwatershed were estimated as explained in Section 3.1.4.

One other important data analysis procedure in GIS was the calculation of WQI values. As explained in Section 2.8.2, WQI, a sediment loading value for agricultural areas, is an actual GIS data layer within VirGIS. The data layer indicates the potential sediment that may be delivered to a stream on a cell by cell basis using the USLE equation times a delivery ratio. Accordingly, the total agricultural sediment loss (tons/year) could be estimated for each watershed. To represent the sediment load as tons/acre/year, the tons/year value was divided by the agricultural area in each watershed.

3.2.2 Independent Variables

After completing the data analysis within GIS, all variables used to describe benthic macroinvertebrate assemblage conditions and their stressors were compiled. To relate the stressors to the benthic macroinvertebrate assemblage conditions, the variables were classified as independent and dependent variables. The stressors control or affect the macroinvertebrate assemblage structure. Thus, thirty collected stressors or independent variables were organized into six different categories: (1) sediment parameters, (2) habitat parameters, (3) water quality variables, (4) GIS/landuse parameters, (5) watershed characteristics, and (6) livestock numbers. A complete data list is shown in Table3.3.

Table 3.3 The 30 independent variables investigated in this study

Sediment	Habitat	Water Quality	GIS/Landuse	Watershed Characteristics	Livestock #s
WQI	Substrate	DO	% Pasture	Watershed Area	Beef cattle w/ Stream Access
TSS	Cover	Temperature	% Cropland	Stream Length	
Turbidity	Embeddedness	pH	% Agriculture	Stream Order	
	Flow	Hardness	% Forest		
	Velocity	Alkalinity	% Urban		
	Channel Alteration	Conductivity			
	Sediment Deposition				
	Riffle				
	Bank Stability				
	Bank Vegetation				
	Graze				
	Riparian Vegetation				

The sediment variables investigated were WQI and TSS and turbidity. WQI is a sediment loading variable calculated in GIS using data provided by VADCR whereas TSS and turbidity are water quality variables collected by VADEQ. The 12 habitat variables were all reported in VADEQ's benthic data file EPAHAB. The water quality variables were either collected from VADEQ's benthic data file WATER or VADEQ's Ambient Water Quality Reports. The

GIS/landuse, watershed characteristics, and livestock numbers were based upon data provided by VADCR, except for stream order, which was in VADEQ's benthic data file STATION.

3.2.3 Dependent Variables

In addition to distinguishing the 30 stressor variables as independent variables, the variables that represented macroinvertebrate conditions were organized as the dependent variables. The macroinvertebrate assemblages are assessed using metrics to arrive at a final RBP impairment type. Both the metrics and RBP impairment type are important in describing macroinvertebrate assemblages. Accordingly, the data were organized to first investigate RBP values as a representation of macroinvertebrate assemblage. Then, the individual RBP metrics were singled out one at time to depict the stressors' impact upon the benthic macroinvertebrates.

The individual RBP metrics were investigated along with the RBP classification for two primary reasons. First, benthic assessment procedures have been adapted in recent years to become more universal. The metrics, the building blocks of the final RBP classification, are still being incorporated into the biological assessment procedures; however, the trend has been to eliminate the final RBP impairment classification procedures (Voshell, 2001). In light of these biological assessment changes, the incorporation of metrics within this study will aid in the applicability of this study's findings for future work. Second, the individual metrics help minimize the error introduced from the numerical RBP classifications which were assigned as 0, 1, 2, or 3 for non-impaired to severely impaired, respectively. These assigned whole numbers do not provide a range and make, for instance, all of the severely impaired stations seem identical regardless of their individual characteristics. In addition, with the stations' impairment classifications lacking the two categories of slightly and severely impaired stations, the statistical variability greatly limited the usage of the RBP impairments. Therefore, the inclusion of the individual metrics helped to alleviate these two problems, making this study applicable to future work while adding an extra dimension to the statistical analysis.

When investigating the RBP values, the classification system of non-impaired, slightly, moderately, and severely impaired was restructured into numerical classifications of 0, 1, 2, and 3, respectively. The numerical classifications or dummy variables were assigned so as to allow for statistical analysis. For the metric evaluations, five metrics out of the eight possible were investigated (Voshell, 2001). The five metrics chosen were: taxa richness (TAXA), the modified family biotic index (MFBI), percent contribution of dominant family (DOMPCT), EPT Index (EPTI), and the ratio of the shredders to the total number of organisms (S-T). These five are the most commonly used metrics and they include diversity by investigating richness, tolerance, and feeding measures (Voshell, 2001). The three metrics that were not included were ratio of scrapers to filtering collectors, ratio of EPT to Chironomidae abundance, and community similarity indices. These three metrics were eliminated since they are not as widely used and/or since ratios could result in the division by zero, which is a non-existent number. The S-T ratio was still incorporated as one of the five metrics since it is not really a ratio. The S-T metric is related as a percent since the denominator is based upon the total number of organisms and will never be zero.

Each of the five metrics used in this study had different scales. The TAXA metric is a richness measure that measures the total number of families present. For the “usable” stations (Table A.1, Appendix A), the TAXA values ranged from 5 to 28 and generally increased with improvements in water quality. The MFBI metric is a tolerance measure that reports tolerances from 0 to 10 and increased as water quality becomes degraded due to organic pollution. The MFBI range for the “usable” stations was from 2.75 to 7.95. The DOMPCT metric is a tolerance measure and was reported as a decimal from 0 to 1 with the “usable” stations’ data ranging from 0.139 to 0.962. A community dominated by few families has a low DOMPCT number, which indicates environmental stress. The EPTI metric is a richness measure that reports the number of distinct taxa within the EPT (Ephemeroptera, Plecoptera, and Trichoptera) pollution sensitive groups. The “usable” stations’ EPTI values ranged from 0 to 12 with the higher numbers generally indicating improved water quality. The S-T metric is a feeding measure that is reported as a decimal ranging from 0 to 1 with the values for the “usable” stations only ranging from 0 to 0.236. Since shredders are good indicators of landuses, toxics, and riparian vegetation, as the value increased, water quality degraded.

In summary, the RBP values and metrics within DEQ’s biological assessment database were used as the dependent variables to represent the benthic macroinvertebrate assemblage conditions. However, only five metrics, TAXA, MFBI, DOMPCT, EPTI, and S-T, were analyzed. Thus, the data were organized into six dependent variables that most adequately represented the benthic macroinvertebrate assemblage conditions.

3.2.4 Data Sets

The final step in the data analysis procedures was to organize the data into data sets. The dependent/independent relationships that represented how stressors affect the benthic macroinvertebrate assemblage conditions were formed based upon two different data sets. The initial data set focused upon investigating WQI's potential as a sediment loading variable that may affect the macroinvertebrate assemblages. The final data set investigated TSS, as the sediment indicator, as well as additional variables that may impact macroinvertebrate assemblages.

3.2.4.1 Initial Data Set: Investigating the WQI as a Sediment Indicator

The initial data set that investigated the use of the WQI as a sediment indicator was developed prior to developing the "usable" benthic data. Some of the criteria used to develop the "usable" data set were applicable to this data set. For instance, this initial data set contained stations with 4th order streams or smaller, stations with agricultural or urban activities, and stations that had GIS data available. However, unlike the "usable" data set, the sampling period was not restrained by the water quality data since water quality variables were not investigated. Thus, the WQI analysis was able to include data samples from the fall of 1994 to the fall of 1998 and stations regardless of whether or not water data was available.

The resulting initial data set used to investigate the WQI contained 18 benthic stations with a total of 90 samples from three highly agricultural counties: Rockbridge, Rockingham, and Augusta counties. Table 3.4 summarizes the impairment classifications for the compiled data set. The majority of the impairments were moderately impaired with only a few non-impaired, slightly, and severely impaired stations. In addition, the number of samples was evenly distributed between the sampling periods except the spring of 1998. Therefore, the uneven distribution between impairment classifications could present problems when developing the stressor/benthos relationships.

The description of each of the 18 stations is presented in Table B.1, Appendix B. The summary reveals that the majority of the stations are affected by agricultural nonpoint source pollution. Since the WQI investigates sediment loadings for agricultural areas, it was important to include stations highly influenced by agricultural activities. The specific locations of the 18 stations are shown in Figure 3.3.

In addition to investigating WQI, watershed area, agricultural area, % pasture, % row crop, % grass, % pasture + row crop, % cultivated, % grassland, and stream length were also included as independent variables that could influence the RBP value. Moreover, no water quality data, livestock numbers, or habitat variables, were included and no metric value relationships were investigated.

The statistical analysis results for the investigation of WQI as a sediment indicator led to the investigation of the final data set for three reasons. First, the developed relationships never included WQI. Second, the relationships between RBP and the remaining nine independent variables could not even explain half of the variability in the benthic macroinvertebrate assemblage conditions. Lastly, WQI only investigates sediment yields from agricultural areas. An additional sediment indicator that would be consistent with the WQI procedures was needed for urban and forested areas. Accordingly, the final data set was organized to minimize these drawbacks.

Table 3.4 The impairment classifications for the samples used to evaluate WQI

IMPAIRMENT CLASSIFICATION	NUMBER OF SAMPLES									TOTAL
	Fall 1994	Spring 1995	Fall 1995	Spring 1996	Fall 1996	Spring 1997	Fall 1997	Spring 1998	Fall 1998	
Non-Impaired	0	1	0	0	0	0	0	0	0	1
Slightly Impaired	0	0	5	0	0	0	0	0	6	11
Moderately Impaired	10	12	5	7	7	7	13	0	4	65
Severely Impaired	2	2	4	1	1	3	0	0	0	13
TOTAL	12	15	14	8	8	10	13	0	10	90

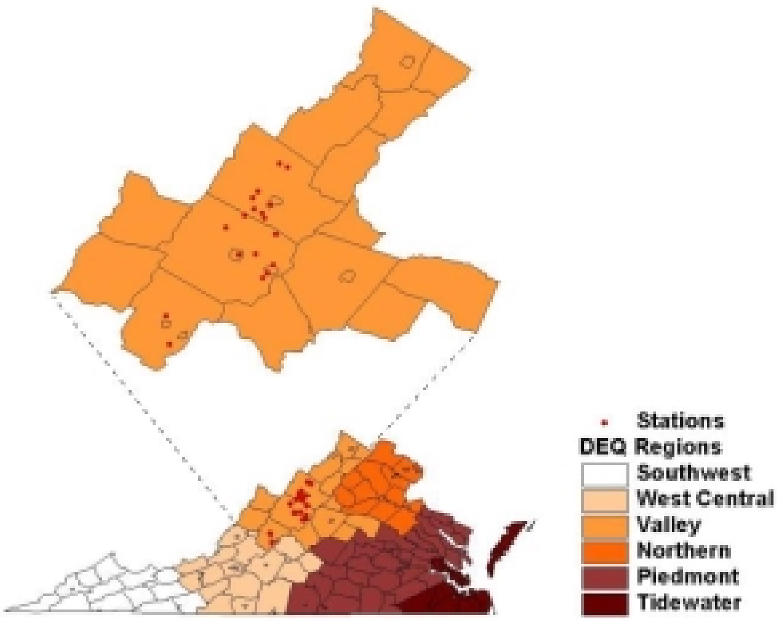

Figure 3.3 The location of the 18 benthic stations used in the WQI analysis

3.2.4.2 Final Data Set: Investigating TSS and Other Stressor Variables

The final data set was used to assess the relationships between TSS as the new sediment indicator as well as numerous independent variables and the benthic macroinvertebrate assemblage conditions. When considering the disadvantages of the initial data set, TSS was chosen as the sediment indicator for this final data set since it applies to all landuse types and is easily measured in the field. Turbidity was added as an additional sediment variable since it is a measurement of the clarity of the water and is caused by sediment particles suspended in the water. In addition to TSS and turbidity, the 27 other independent variables, excluding WQI, listed in Table 3.3, were investigated. Since the few independent variables investigated in the initial data set explained a minimum amount of variability, the independent variable list was expanded in the final data set to account for the vast number of stressors that may impact the macroinvertebrate conditions.

The final data group included 34 out of the 44 possible “usable” stations so that the remaining 10 stations could be used to validate the developed relationships. The 34 stations with a total of 105 samples used to develop the relationships were collected from 13 counties: Rockbridge, Rockingham, Augusta, Frederick, Shenandoah, Page, Loudoun, Fairfax, Prince William, Fauquier, Culpeper, Rappahannock, and Madison. Table 3.5 summarizes the impairment classifications for the compiled data set. The majority of the impairments were either non-impaired or moderately impaired, as was the case for the original “usable” stations. Only a few of the samples were slightly or severely impaired stations. In addition, the number of samples was evenly distributed among the seasons except for the spring of 1998. Therefore, the uneven distribution between impairment classifications could present problems when developing the stressor/benthos relationships.

The 34 stations within the final data set were located within the Valley and Northern VADEQ regions. The reasons for inclusions of these stations in the DEQ monitoring program as well as the geographic locations (ecoregions) of the stations are explained in Table B.3, Appendix B. The locations of these stations are shown in Figure 3.4.

Table 3.5 The impairment classifications for the “final” benthic samples

IMPAIRMENT CLASSIFICATION	NUMBER OF SAMPLES					TOTAL
	Fall 1996	Spring 1997	Fall 1997	Spring 1998	Fall 1998	
Non-Impaired	6	7	10	6	10	39
Slightly Impaired	0	0	0	0	5	5
Moderately Impaired	12	15	18	3	10	58
Severely Impaired	1	2	0	0	0	3
TOTAL	19	24	28	9	25	105

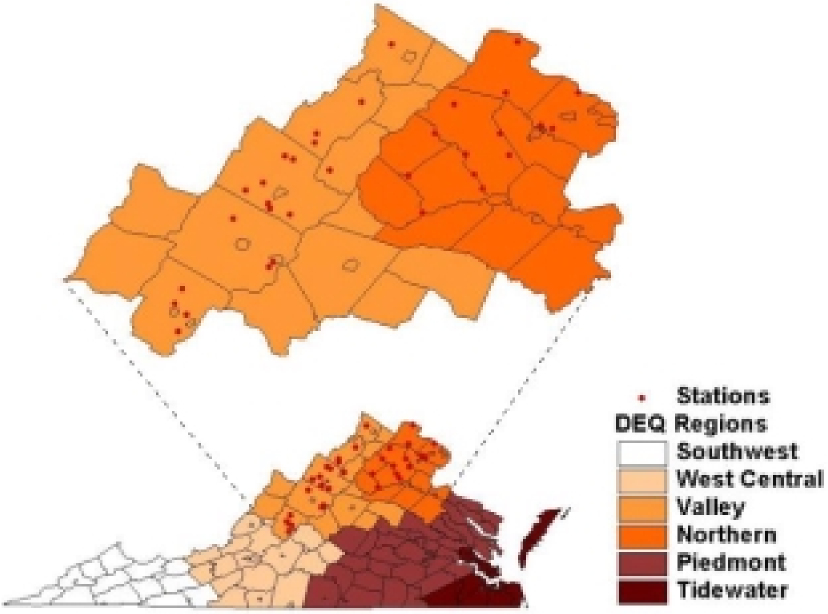

Figure 3.4 The location of the 34 benthic stations used in the final data set

As shown in Figure 3.4, the stations are relatively evenly split between the Valley and Northern regions. The Valley region contains 19 stations with 46 samples whereas the Northern region has 15 stations with 59 samples. The combined 34 stations were used to assess the 29 independent variables, excluding WQI, mentioned in Section 3.2.2 and the six dependent variables discussed in Section 3.2.3. To assess the relationships between the RBP and/or metrics and other stressors, statistical analysis procedures were performed.

3.3 Statistical Analyses

Various statistical procedures were used to evaluate the existence of relationships between the benthic macroinvertebrate assemblage conditions and their stressors. Once all data sets were assembled, statistical analysis procedures were conducted to develop the relationships with multiple regression techniques being the primary statistical procedure. In addition, correlation analysis was used to examine the linkage between the benthic macroinvertebrate assemblage conditions and seasonal, annual, and ecoregional differences. To justify the results of the multiple regression procedures and to make sure no other valid relationships existed, principal component analysis and r-square analysis procedures were conducted. Finally, the accuracy of the developed relationships was tested during the validation processes.

3.3.1 Multiple Regression

Multiple regression was the primary method used to develop relationships between the stressors and the benthic macroinvertebrate assemblage conditions. The general purpose of multiple regression is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable (StatSoft, 2001). The multiple regression model relating a dependent variable, y, to independent variables, x, is written as (Ott, 1993):

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \qquad [3.1]$$

where: β = partial slope,
ε = random error.

By assuming the average value of ε, $E(\varepsilon)$, is equal to zero, the expected value of y is described as (Ott, 1993):

$$E(y) = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \qquad [3.2]$$

The simplest type of multiple regression equation is a first-order model in which case all independent variables are as they appear with no cross-product terms, powers, logs, or other transformations. Aside from the first-order model, the independent variables in Equation 3.2 may also be powers, cross-products, or logs of other independent variables. For example, x_2 might be x_1^2 or x_3 might be equal to x_1x_2.

Accordingly, the independent variables in Table 3.3 were used to include cross-products, logs, and squared variables. The original variables in Table 3.3 were assumed to be in linear form. To account for non-linear interactions, cross-product variables were included if the linear variables exhibited collinearity with one another. Variables that account for overlapping pieces of variability were determined graphically and statistically. If two independent variables were graphed against one another and a linear one-to-one trend was evident, the two variables were exhibiting collinearity. To validate the graphical results, the Pearson correlation coefficients and p-values were obtained. If the correlation coefficient was closer to –1 or 1 than to 0 and if the p-value was less than 0.05, then a correlation existed between the variables in question. Thus, six cross-product terms were obtained: DO x Temperature; TSS x Turbidity; TSS x % Agriculture; Alkalinity x Hardness; TSS x Channel Alteration; and, TSS x Bank Vegetation. In addition to cross-product terms as non-linear variables, log transformations and squared transformations of all 30 independent variables, except WQI, and the 6 new cross-product terms were also included. Moreover, non-linear variables were distinguished as the six cross-product terms, 35 log transformations, and 35 squared transformations.

To develop the relationships between the independent variables, both linear and non-linear forms, and the dependent variables, three procedures were used: backward elimination, forward selection, and stepwise regression procedures. The backward elimination method begins with a complete model that includes all independent variables and eliminates variables, one at a time, until a reasonable candidate regression model is found. A variable is removed from the model if it does not meet the set significance, alpha value, to remain in the equation. Variables furthest away from meeting the significance criteria are removed first (StatSoft, 2001). The backward elimination method is advantageous in that it is the only method out of the three that shows the full model. On the other hand, the backward elimination method is inflexible in re-entering the variables once removed (Ye, 1999).

The forward selection method starts with no variables entered in the model or as $y = \beta_0 + \varepsilon$. The model progressively adds one variable at a time in terms of which variables are best at meeting the significance criteria. Once the variable is entered, the variable cannot be eliminated from the regression equation. The procedures stops when no other variables meet the significance criteria for entry into the model. The forward selection method is advantageous since it is intuitive to enter the most significant variable first. The disadvantage to using this method is that it is inflexible in removing variables (Ye, 1999).

The stepwise regression method also starts with the model $y = \beta_0 + \varepsilon$ and adds or removes variables one at a time. The difference in this method from the forward selection method is that the stepwise method allows for variables to be removed from the model (Ott, 1993). Since the original best single variable might not be one of the best two variables, the original variable may be removed later in the process. The procedure stops once no other variables can be removed or added based upon the significance criteria. The advantage of the stepwise regression method is in its flexibility in removing and re-entering variables. The disadvantage of this method is that it does not guarantee that the final model is the best overall model in describing the variability among the benthic macroinvertebrate assemblage conditions (Ye, 1999). The stepwise regression model can only include the most appropriate combination of variables that enter the model with significant variables.

Since the backward, forward, and stepwise regression procedures each has its own advantages and disadvantages that are equally important, the regression procedure that provided the best overall relationships was singled out and focused upon during analysis. A relationship is considered highly feasible if a high amount of variability can be explained through the inclusion of minimum number of independent variables. In addition, the relationship between the dependent and independent variables should follow the expected pattern of direct (positive) or indirect (negative) association. Thus, these considerations were taken into account when analyzing the relationships depicted through multiple regression techniques for both the initial data set and final data set.

3.3.1.1 Initial Data Set: Investigating the WQI as a Sediment Indicator

The primary objective for the first data set was to detect whether or not the WQI would be an adequate sediment load indicator that accesses the impact of sediment upon the benthic macroinvertebrate assemblage conditions. Since the initial data set was used as a preliminary investigation of WQI, the statistical analysis procedures were limited and only the first-order multiple regression model was investigated. No cross-products or transformations of the independent variables were examined. In addition, the first data set only investigated the RBP values as dependent variables. Thus, the RBP values were related to WQI and 9 additional stressors: watershed area, agricultural area, % pasture, % row crop, % grass, % pasture + row crop, % cultivated, % grassland, and stream length.

To develop the relationship between the RBP values and the possible stressors, backward, forward, and stepwise regression techniques were employed. The level of significance (p-value) for entry and/or removal of variables was 0.5, 0.1, and 0.15 for the forward, backward, and stepwise, respectively. The three regression techniques were used on six different data subsets: “spring”, “fall”, “all”, “spring average”, “fall average”, and “all average”. The six subsets were used to look at RBP and seasonal differences. The “spring” subset contained data from all samples collected during the spring periods, the “fall” subset contained data only from samples collected during the fall periods, the “all” subset included all samples regardless of season, the “spring average” evaluated the average of all independent and dependent variables for each station’s spring samples, the “fall average” evaluated the average of all independent and dependent variables for each station’s fall samples, and the “all average” subset evaluated the average of all independent and dependent variables for all samples for a given station, regardless of which season the variables were collected. Moreover, the “averaged” subsets averaged each station’s samples together so that each station only had one set of variables. Asides from the backward, forward, and stepwise first-order models, no additional statistical procedures were used on the initial data set.

3.3.1.2 Final Data Set: Investigating TSS and Other Stressor Variables

The final data set, as the name implies, was the ultimate data set used in investigating the relationship between stressors and stream benthic macroinvertebrate assemblage conditions. Accordingly, various statistical analysis procedures were explored to justify the developed relationships. The final data set incorporated all independent variables from the six developed categories (sediment, habitat, water quality, GIS/landuse, watershed characteristics, and livestock numbers) in Table 3.3, except WQI since WQI was already investigated using the first data set. In addition to using the 29 independent variables, all six possible dependent variables (RBP values, TAXA, MFBI, DOMPCT, EPTI, and S-T) were examined one at a time. Analysis procedures included the backward, forward, and stepwise regression techniques for not only the first-order or linear models, but also for the non-linear models that incorporated cross-product terms, squared terms, and log-transformed terms. A significance level of 0.05 ($p = 0.05$) was used in all cases so that only the most representative independent variables would be included in the final equations (Smith, 2001).

Linear and non-linear techniques were performed for the six subsets (“spring”, “fall”, “all”, “spring average”, “fall average”, “all average”) used in the initial data set as well as for three annual averages for “1997”, “1998”, and “1999”. The sample periods included in the annual average data subsets were arranged to follow the water quality data sampling periods. For instance, the “1997” subset averaged the variables for the fall of 1996 and the spring of 1997 into one value for each station. Once all linear and non-linear multiple regression analyses were completed for the nine subsets, the regression results were broken down into relationships with the RBP value as the dependent variable and relationships with a metric as the dependent variables so as to compare and contrast the RBP and metric relationships. The developed linear and non-linear RBP and metric relationships were analyzed and compared to determine the most suitable relationships. Relationships were selected as the most suitable based upon the amount of variability they explained, the number of independent variables incorporated in the relationship, whether or not the variable associations (either positive or negative) corresponded with knowledge already known on the subject matter, and the sampling size of the data. A suitable relationship would explain a high amount of variability with minimum independent variables that were correctly associated with the dependent variable and had an appropriate sampling size that would not distort the developed relationships. However, to accurately determine the seasonal, annual, and ecoregional influences upon the benthic macroinvertebrate assemblage conditions, correlation analysis was performed. Likewise, to ensure that the most adequate stressor/benthos relationships had been developed, the principal component analysis and r-square analysis procedures were examined.

3.3.2 Correlation Analysis

The next step in the development of the relationship between stressors and stream benthic macroinvertebrate assemblage conditions was to perform correlation analysis. The multiple regression techniques did not distinguish whether or not the seasonal or annual subsets differed in terms of how the stressors impacted the benthic macroinvertebrate assemblages. Correlation analysis can be used to evaluate whether or not seasonal or annual differences in the RBP values and/or metric values (TAXA, MFBI, DOMPCT, EPTI, S-T) exist. Correlation analysis only looks at the variable being tested (i.e. season or year) and relates how strongly it impacts the variable in question (i.e. RBP value or RBP metrics). In addition to seasonal and annual differences, ecoregional differences were investigated since it has been reported that benthic macroinvertebrate assemblage conditions differ among ecoregions in Virginia (Evans, 1997).

To complete the correlation analysis, the data within the final data set were used. Since seasons, years, and ecoregions are all classified by name, numerical classifications or dummy variables were assigned to the categories. The seasonal, annual, and ecoregional data were then evaluated one at a time for all six dependent variables (RBP values, TAXA, MFBI, EPTI, DOMPCT, S-T) to detect whether or not any notable differences in benthic macroinvertebrate assemblage conditions could be discerned based upon the seasons, years, or ecoregions.

To interpret the results, the correlation coefficient (r-value) and significance value (p-value) determined whether or not the three factors affect the assemblage conditions. When the r-value was closer to 1 or –1 than to zero and when the p-value was less than 0.05, a correlation was present. The presence of a correlation suggests that the factor impacts the benthic macroinvertebrates differently between the types available. For instance, a seasonal correlation would imply that the fall and spring seasons affect the macroinvertebrates differently; an annual correlation would imply that the 1997, 1998, and 1999 years affect the macroinvertebrates differently; and, an ecoregional correlation would suggest that the Central Appalachian Ridges and Valley and Northern Piedmont affect the macroinvertebrates differently.

Correlation analysis is advantageous over multiple regression techniques since it directly evaluates whether or not a relationship exists between two variables. The analysis determines how strongly correlated the two variables are to one another. However, correlation analysis, unlike multiple regression techniques, does not depict a model or equation that relates how the two variables are linked to one another.

3.3.3 Principal Component Analysis & R-square Analysis

After having developed relationships using the multiple regression techniques, principal component analysis and r-square analysis techniques were implemented to ensure that the developed regression relationships from the final data set were the most suitable and representative relationships. These two analysis procedures were chosen since they minimize the weaknesses of the multiple regression techniques. The multiple regression techniques were disadvantageous since excessive independent variables were investigated and since the developed relationship may have explained in excess the magnitude of variability if too many variables were included.

3.3.3.1 Principal Component Analysis

Principal component analysis is helpful in grouping similar variables into one new variable. The method is applied as a data reduction or structure detection method (StatSoft, 2001). For this study, three variable clusters were investigated to form three new variables: Prin1, Prin2, and Prin3. Prin1 primarily included habitat variables as well as hardness, alkalinity, conductivity, and stream order. Prin2 included landuses such as % forest, % pasture, and % cropland as well as turbidity. Prin3 included watershed characteristics and beef cattle numbers. Thus, Prin1 was termed the HABITAT variable, Prin2 was termed the LANDUSE variable, and Prin3 was called the WATERSHED CHARACTERISTIC variable.

The advantage of using principal component analysis was that only three variables and a few additional variables not incorporated in the three new variables were investigated as independent variables. Thus, the number of independent variables was minimized from the original 29 linear independent variables and 105 non-linear independent variables. Even though principal component analysis minimizes the number of variables investigated by grouping similar variables together, the original variables would still have to be measured. Furthermore, the principal component analysis procedure could not be used to minimize the data collection process.

3.3.3.2 R-Square Analysis

The r-square analysis procedure is a form of multiple regression that examines all possible regression models starting with all relationships that incorporate one independent variable at a time, all possible relationships that incorporate two variables at a time, three variables, etc. (StatSoft, 2001). Within the r-square analysis procedure Mallow's C_p and mean standard error (MSE) values were also listed with the R^2 values. The R^2 values were important to show how much variability a relationship explained. The C_p values were important in determining the number of variables that should be included in an equation so as to avoid multicollinearity and an underspecified relationship (Ott, 1993). The MSE values were used to make sure that the amount of error in the developed relationships was minimal. Ideally, the relationship should have a high R^2 value close to one, a Cp value approximately equal to the number of parameters plus one, and a MSE value close to zero (Ott, 1993). Thus, the selected relationship was one that most accurately met all three requirements and did not show much improvement when additional variables were added.

In addition to determining what relationships were ideal based upon the R^2, Cp, and MSE values, the r-square analysis procedure was used to investigate why the TSS sediment variable did not enter into certain relationships. The r-square values of each individual independent variable showed the amount of variability TSS could explain by itself and how many other independent variables could explain more variability than the TSS variable.

Overall, the r-square analysis procedure was advantageous since it could look at any relationship and determine how much variability was explained. In addition, the r-square analysis procedure could use the number of variables and error values to determine an appropriate relationship. The disadvantage of this procedure over the previous multiple regression techniques is that a relationship could include variables that did not meet the 0.05 significance requirement.

3.3.4 Validation

The last step in the statistical analysis procedures was to make sure that the developed relationships could be applied with reasonable accuracy to additional streams in Virginia. To validate the relationships, the remaining 10 “usable” stations that were not used in the development of the relationships were used for validation. The 10 stations with a total of 29 samples from the fall of 1996 to the fall of 1998 have their impairment classifications summarized in Table 3.6 and locations shown in Figure 3.5.

The majority of the impairments for the 29 samples were either non-impaired or moderately impaired as was the case for the original “usable” stations and the stations used to developed the relationships. Only a few of the samples were slightly or severely impaired stations. Therefore, only the uneven distribution between impairment classifications could present problems when developing the stressor/benthos relationships.

The validation stations were chosen to be representative of the stations used to develop the relationships. Figure 3.5 demonstrates that the 10 validation stations were scattered throughout the same three DEQ regions as the “usable” stations. The descriptions of the validation stations in Table B.5, Appendix B, demonstrate that agricultural, urban, and forested areas were included throughout 8 counties (Bedford, Montgomery, Pulaski, Giles, Botetourt, Albemarle, Orange, and Culpeper) and 2 ecoregions (the Central Appalachians Ridges and Valleys and the Northern Piedmont).

Table 3.6 The impairment classifications for the “validation” benthic samples

IMPAIRMENT CLASSIFICATION	NUMBER OF SAMPLES					TOTAL
	Fall 1996	Spring 1997	Fall 1997	Spring 1998	Fall 1998	
Non-Impaired	1	3	3	3	3	13
Slightly Impaired	0	0	1	0	0	1
Moderately Impaired	2	3	4	4	1	14
Severely Impaired	0	0	0	0	1	1
TOTAL	3	6	8	7	5	29

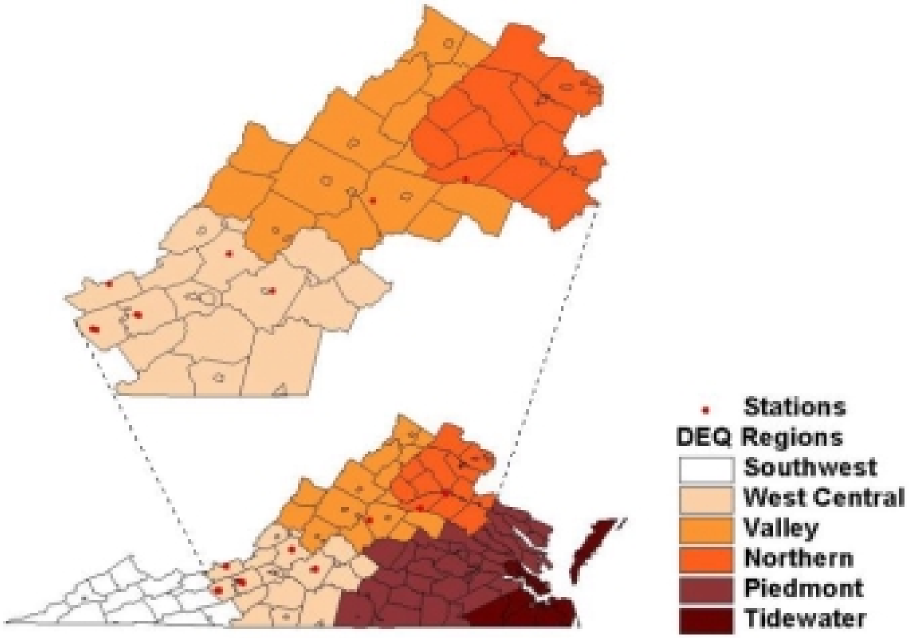

Figure 3.5 The locations of the validation stations

The independent and dependent variables were compiled for the 10 stations with 29 samples and the dependent variables were classified as the observed dependent values. The predicted dependent values were estimated using the developed regression equations by inserting the independent variables into the developed relationships and back calculating for the predicted value. Thus, the benthic macroinvertebrate assemblage conditions had an observed and predicted value for all samples.

The predicted and observed dependent variables were then plotted against each other and a best-fit linear line was drawn to relate the observed and predicted values to one another. Investigations of the r-square value, slope and intercept, and confidence interval were used to validate the relationships. A relationship was proven to be valid when R^2 values for the new predicted/observed line and the original relationship were similar, a predicted/observed line had a slope close to one and an intercept near zero, and when the predicted/observed data had minimal separation distance within the 90% confidence band. In ending, the validation steps were necessary to test the applicability of the relationships and to ensure that conclusions drawn from the relationships would be justifiable.

3.4 Applicability of the Developed Stressor/Benthos Relationships to TMDLs

The last component of this study was to determine the implications of the benthic stressor relationships for the development of benthic TMDLs. TMDLs, in general, focus upon the establishment of the maximum pollutant load a water body can withstand and still meet water quality standards. For benthic TMDLs there is a need to develop stressor threshold values below which the stressor will no longer adversely impact the benthic macroinvertebrate assemblage conditions.

The general procedure for determining thresholds is to use the data from non-impaired samples only. The non-impaired samples of all of Virginia’s biomonitoring stations and the non-impaired samples of only the benthic stations used to develop the relationships were investigated. Both groups were used primarily to verify that the relationships could be applied to other stations regardless of whether or not the station was used to develop the stressor/benthos relationship. If the two sampling groups produced similar threshold values, then applicability of the relationships to other streams in Virginia would be verified. In both cases, each variable's values for the non-impaired samples were averaged. The averaged variable values represented the stressor levels needed to ensure that a water body is not benthically impaired. Thus, when these average values were inserted into the developed stressor/benthos relationship, the threshold value for the stressor of concern could be calculated. The estimated threshold value would indicate that any stressor value exceeding the threshold would result in a stream that is no longer non-impaired.

When a benthic TMDL is being developed for a certain stressor, the stressor variable should be a variable that can be altered through implementation of practices upstream of the impairment site. The sediment variables are good examples of a stressor that can be minimized through implementation of upstream practices such as streambank restoration or vegetated waterways. Accordingly, the sediment variables were used as an example of the implications of the relationships with regard to TMDLs. Furthermore, the sediment variables were investigated to determine the maximum sediment level that would ensure that the benthic macroinvertebrates were not adversely impacted.

In addition to the threshold values, this study investigated the stressor values typically displayed for moderately and severely impaired streams in Virginia. The typical values were evaluated for the data group that included all Virginia biomonitoring stations to determine what the sediment levels currently are for moderately and severely impaired stations in Virginia and how much change is needed to achieve the sediment level necessary for benthic TMDLs. As with the threshold estimation, the average values for all variables except the one in question were determined from the original data and then the stressor in question was calculated using the developed stressor/benthos relationships. The typical value could then be compared to the threshold value to estimate the amount of change needed for the typical Virginia stream to achieve a threshold value that does not adversely impact the benthic macroinvertebrate assemblage conditions.

Moreover, the relationships between the stressors and benthic macroinvertebrate assemblage conditions were investigated for applicability in developing a benthic TMDL. The applicability of the developed stressor/benthos relationships was a beneficial investigation since it demonstrated how policymakers, biologists, or stakeholders can use the stressor/benthos relationships to develop benthic TMDLs and/or determine the impact of certain activities or stressor on the benthic macroinvertebrate assemblage conditions in a stream.

3.5 Chapter Summary

To determine the relationship between stressor and benthic macroinvertebrate assemblage conditions, data needed to be organized and analyzed and statistical analyses procedures needed to be performed. Benthic macroinvertebrate variables, habitat parameters, and water quality parameters, were obtained from VADEQ from their Biological Assessment Database and the Ambient Water Quality Reports. WQI, landuse characteristics, watershed characteristics, and livestock numbers were obtained from the VirGIS database, mrlc files, and, the Hydrologic Unit Animal Census Database. Out of the 418 benthic stations monitored by VADEQ, 44 stations (34 for development and 10 for validation) were used in this study to develop and evaluate the stressor/benthic relationships.

The data variables collected from VADEQ and VADCR, either collected from a database or collected within GIS, were organized for 34 of the 44 applicable stations and statistical analysis procedures were performed to develop the relationships. Statistical analysis procedures such as multiple regression, correlations analysis, principal component analysis, and r-square analysis were used to develop the relationships between stressors (sediment variables, habitat parameters, water quality parameters, landuse characteristics, watershed characteristics, and livestock numbers) and benthic macroinvertebrate assemblage conditions (RBP values, TAXA, MFBI, DOMPCT, EPTI, and S-T). After developing the relationships, validation was performed for the remaining 10 stations using linear regression values such as R^2 values, slopes and intercepts, and confidence intervals. Finally, the implications of the developed relationships with regard to TMDL development plans were investigated. Ultimately, the proposed relationships could be used to determine stressor threshold values that would not adversely impact the benthic macroinvertebrate assemblage conditions.

CHAPTER 4.0 RESULTS AND DISCUSSION

The goal of the study was to establish relationships between stream benthic macroinvertebrate assemblage conditions and their stressors. The specific objectives were to: (1) develop relationships between the RBP indices (or individual metrics) and stressors; (2) evaluate the accuracy of the relationships using data from selected watersheds in Virginia not used during the development procedures; and, (3) discuss the implications of the study results for benthic TMDLs. To achieve these objectives, existing data were organized into three data sets used for: a) investigating WQI as a sediment indicator, b) relating TSS and other stressor variables, and c) validation of the stressor/benthos relationships. The WQI data set was used to assess WQI's potential as a sediment loading indicator; the data set for TSS and other stressor variables was used to develop the stressor/benthos relationships; and, the validation data set was used to test the accuracy of the developed relationships for usage with additional streams in Virginia. Upon completing analysis of the three data sets, the utilization of the relationships for the development of benthic TMDLs was discussed.

4.1 The Data Set Used to Investigate WQI as a Sediment Indicator

The initial data set was organized to assess the potential of using the Water Quality Index (WQI) as a sediment indicator. The WQI is a procedure developed by Shanholtz et al. (1990) that estimates the amount of agricultural sediment delivered to a stream (Section 2.8.2). The goal of the study was to develop relationships for benthic macroinvertebrate assemblage conditions and their stressors. However, emphasis was placed upon sediment indicators as a primary stressor to the benthic macroinvertebrate assemblages since sediment is recognized as the major pollutant of the Nation's waters (USEPA, 1990). With the benthic macroinvertebrates inhabiting in the sediment accumulated on a stream's bottom, the direct interaction of the benthic macroinvertebrates with sediment only increases the likelihood that it could be a significant stressor for the benthic macroinvertebrate assemblages. Accordingly, the initial data set was organized to evaluate the utility of WQI as a sediment indicator.

The Rapid Bioassessment Protocol (RBP) values from the fall of 1994 through the fall of 1998; WQI; watershed area; agricultural area; % pasture; % row crop; % grass; % pasture + row crop;

% cultivated; % grassland; and stream length were obtained for 18 stations in Rockingham, Augusta, and Rockbridge counties. Detailed information on these 18 stations and the collected data are shown in Tables B.1 and B.2, Appendix B.

The majority of the 18 stations were monitored by DEQ to determine the effect of agricultural NPS pollution on the water quality and biota community. In cases were agricultural land was not the dominant landuse in the watershed, forest prevailed with agricultural landuse surrounding the stream. Stations with significant agricultural influence were chosen since WQI focuses upon the potential sediment loads from agricultural areas. Moreover, the additional stressor variables investigated along with the WQI focused primarily upon agricultural landuse while only the RBP values (no metrics) were used to represent the benthic macroinvertebrate assemblage conditions. Table 4.1 presents a summary of the results found for the three counties.

Table 4.1 Summary of the collected data for the data set used to investigate WQI used as a sediment indicator

County	Station	WQI (t/ac/yr)	RBP[a]	Average RBP Values[b]	% Agricultural Area	Watershed Area (ac)
Rockingham	Dry River	32.25	slight/mod	1.50	61.09	10111
	Linville	62.95	mod/slight	1.83	76.24	29858
	Blacks Run	28.03	moderately	2.00	25.65	7496
	Turley Creek	57.57	moderately	2.00	43.00	9259
	Muddy Creek (lf)	73.15	moderately	2.00	51.59	10863
	Muddy Creek (rt)	60.81	mod/sev	2.14	72.71	9238
	Cooks Creek	94.70	sev/mod	2.50	78.14	14500
	Pleasant Run	45.16	sev/mod	2.86	77.73	5343
Augusta	South River 1	45.04	mod/slight/non	1.25	20.69	2251
	Christians Creek	57.55	mod/slight	1.50	62.41	52937
	Mossy Creek	62.18	mod/slight	1.67	78.80	8209
	Middle River	54.13	mod/slight	1.67	79.17	22102
	South River 2	30.11	mod/slight	1.75	22.86	17795
	Back Creek	33.13	moderately	2.00	8.04	26645
	Moffett Creek	57.02	moderately	2.00	29.29	6123
	Lewis Creek	54.24	mod/sev	2.33	42.45	9656
Rockbridge	Kerrs Creek	110.75	mod/slight	1.67	26.57	22178
	Buffalo Creek	99.28	mod/slight	1.75	28.21	77510

[a] All final RBP impairment classifications found throughout a station's samples were listed for the station with the impairment that occurred most often listed first.

[b] The RBP values for each station's samples were averaged based upon the convention of non-impaired = 0; slightly = 1; moderately = 2; severely = 3.

The RBP values shown in Table 4.1 indicate the type of impairment classifications during the sampling period from the fall of 1994 through the fall of 1998. If more than one impairment was noted throughout the sampling period, the classification that occurred most often is indicated first in the RBP column. The average RBP number is the sum of all impairment classifications for a given station divided by that station's total number of samples. The average RBP values are arranged in increasing order for each county. If WQI is a good indicator of the benthic macroinvertebrate assemblage conditions, then WQI values should be increasing as the average RBP values increase. The column indicating percent of agricultural area and watershed area were added to help explain any deviations in the RBP and WQI trends. As shown in Table 4.1, in situations where the WQI value decreased instead of increasing with RBP, the agricultural area comprised only a small percentage of the total landuse in the watershed. For instance, the percent agriculture in the watersheds for both South River 2 and Back Creek in Augusta County accounted for only 23 and 8 percent of the total landuse area, respectively, while these stations' WQI values were less than those estimated for the other stations. Since the WQI values appear to be related to the percent agricultural area in a watershed, other sediment indicators, rather than WQI, may be better suited for watersheds with small agricultural areas.

To verify the relationship between the benthic impairments and WQI, statistical analyses were performed. Chemical data, livestock numbers, or habitat variables were not included as independent variables in these analyses. The independent variables that could influence the RBP value included: WQI; watershed area; agricultural area; % pasture; % row crop; % grass; % pasture + row crop; % cultivated; % grassland; and stream length. The forward, backward, and stepwise regression analyses techniques were used to develop the most suitable linear relationship (Ott, 1993). The level of significance (p-value) for entry and/or removal of variables was 0.5, 0.1, and 0.15 for the forward, backward, and stepwise respectively. All three statistical techniques were conducted for "spring", "fall", and then "all" samples regardless of season. The analysis was also performed on the "average" RBP values. For the "spring average" subset, for instance, all spring RBP values for each station were averaged so that each station had only one spring RBP value. This was repeated for each station's fall samples ("fall average") and then for all samples regardless of the season ("all average"). Each station's entire number of samples as well as the average of samples were investigated since benthic macroinvertebrates integrate the effects of short-term environmental alterations with sensitive life stages responding quickly to stress, while long-terms effects are incorporated through the entire benthic community which responds slowly to stress (Barbour et al., 1999). Thus, all individual samples as well as an average of the samples were included in the analyses to account for all characteristics the benthic macroinvertebrates exemplify.

The statistical results showed the forward procedure provided the highest R^2 values. Accordingly, only the forward procedure was used to determine which seasonal and RBP subset was preferred. Detailed information on the forward procedure results for the WQI subsets are included in Table C.1, Appendix C. A summary of the forward procedure results for the "spring", "spring average", "fall", "fall average", "all", and "all average" subsets is provided in Table 4.2.

Table 4.2 Summary of the forward multiple regression results for the data set with WQI as a sediment indicator

Variable	**Spring**[a]	**Spring Average**[a]	**Fall**[a]	**Fall Average**[a]	**All**[a]	**All Average**[a]
WQI						
Watershed Area		X[b]	X		X	X
Agricultural Area		X	X		X	X
% Pasture		X	X			
% Row Crop						
% Grass						
% Pasture + Row Crop						
% Cultivated						
% Grassland	X	X			X	X
Stream Length		X	X	X		
No. of samples	33	18	57	18	90	18
R-square	0.21	0.44	0.20	0.18	0.20	0.32

[a] One of the six subsets used to group the data as discussed in Section 4.1.

[b] The 'X' implies that the associated variable was included in the final relationship and significantly contributed to the overall R^2 value.

As indicated in Table 4.2, the WQI values never entered into any of relationships for the six possible subsets. Only watershed area, agricultural area, % pasture, % grassland, and stream length were ever included. In addition, no seasonal differences were detected as the relationships for the "spring" and "fall" subsets both only explained about 20 percent of the variability in the benthic macroinvertebrate assemblage conditions. However, the seasons did show differences in terms of what variables were included in their relationships. The "spring" relationship showed improvements in variability (increased R^2 values) when the samples were averaged together in the “spring average” relationship, whereas the “fall average” relationship actually showed a slight decrease in the amount of variability explained from the original “fall” relationship. However, the R^2 value increased for the “spring average” relationship at the cost of including five times as many variables when compared with the “spring” relationship, whereas the “fall average” relationship’s R^2 value decreased slightly when including fewer variables than were in the “fall” relationship. Overall, the relationship for the “spring average” subset proved to be the most inclusive of all relationships in terms of the variables included in the final relationship, and it provided the highest R^2 value of 0.44. Equation 4.1 shows the relationship developed for the “spring average” subset ($R^2 = 0.44$):

$$\begin{aligned} \text{RBP} = 0.68 \;&+\; 2.04\text{E-05 (watershed area)} \;-\\ &6.81\text{E-05 (agricultural area)} \;-\; 0.30\text{ (\% pasture)} \;+\\ &0.32\text{(\% grassland)} \;+\; 3.54\text{E-05 (stream length)} \end{aligned} \tag{4.1}$$

Of the five variables that were related to the RBP values (Equation 4.1), the agricultural area was not correctly associated with the RBP values. The inclusion of agricultural area supports the idea that agricultural areas impact the benthic community; however, the indirect relationship demonstrates that as agriculture areas increase, the benthic community becomes less impaired. The relationship does not correspond with the research findings that suggest agricultural areas could produce NPS pollution and sediment and thus result in decreased water quality and negatively impact the benthic macroinvertebrate assemblage conditions.

Even though the relationship for the “spring average” subset produced the largest R^2 value, the sampling size was only 18. Distortions may occur based upon small sampling sizes (Ott, 1993). Thus, the “all” or “fall” subsets with sampling sizes of 90 and 57, respectively, may actually provide better representations of the results. Of course, both relationships for the “fall” and “all” subsets had R^2 values around 0.20, much less than the R^2 values that resulted for the “spring average” relationship. To further understand if the sampling size does impact the results, additional data would need to be investigated. However, regardless of the sampling size, the results obtained indicated poor correlations between RBP and the 10 variables investigated.

The results also indicted that WQI, as the sediment indicator, was never introduced into any of the relationships. Furthermore, since the WQI approach is only valid for agricultural areas, other procedures for modeling sediment loads from urban and forest locations would have to be used. Thus, alternative methods for representing sediment were explored. A method that would be consistent for all landuse types was desired so as to minimize any error that may be associated with differences among multiple procedures.

4.2 The Data Set Used to Investigate TSS and Other Stressor Variables

Upon reviewing the results of the WQI data set procedures, WQI was dismissed as the sediment indicator and total suspended solids (TSS) was selected instead due to its ease in measurement and applicability to all landuse types. The impact of additional stressor variables on the benthic macroinvertebrate assemblage conditions were also investigated since the results of the WQI data analysis suggested that many other stressors, in addition to sediment, are important in describing the benthic macroinvertebrate assemblage conditions. Accordingly, RBP values and metrics, TSS, habitat, water quality, landuse, watershed characteristics, and livestock numbers from the fall of 1996 to the fall of 1998 were all obtained for 34 stations in Rockingham, Augusta, Rockbridge, Page, Shenandoah, Frederick, Loudoun, Fairfax, Prince William, Fauquier, Culpeper, Rappahannock, and Madison counties. Information on the 34 stations and the collected data are shown in Tables B.3 and B.4, Appendix B.

The majority of the 34 stations in the data set, used to investigate TSS and other stressors, were monitored by DEQ to determine the effect of agricultural and urban NPS pollution on stream water quality and biotic community with the remaining stations being monitored to represent reference conditions or to extend the monitoring of a particular region. In situations where agricultural or urban areas did not dominate the watershed, forest landuse prevailed with agricultural landuse surrounding the stream area.

Unlike the data set used to investigate WQI, this data set investigated both RBP values and five metric values as dependent variables. The five metrics were included in an effort to enhance the relationships between benthic macroinvertebrates and their stressors since the metrics are closely related to and are more representative of the macroinvertebrate assemblage conditions than the

RBP values. Table 4.3 summarizes the results found for the 34 stations with the data organized to assist in the evaluation of trends between the dependent variables and TSS. The data presented in Table 4.3 are the averaged values for all samples collected at each station, regardless of the season. Thus, each station's RBP values, 5 metrics, and TSS values from all sampling times were averaged. The TSS values were sorted in ascending order for ease in analysis.

Table 4.3 Summary of collected data for the data set investigating TSS

STATION	RIVERMILE	RBP VALUE	TAXA	MFBI	DOMPCT	EPTI	S_T	TSS (mg/L)
Toms Brook	TMB000.54	0.00	25.00	4.85	0.28	12.00	0.07	3.00
Cub Run	CUB000.40	1.00	16.50	3.93	0.22	9.00	0.05	3.00
South River (00)	STH000.21	0.50	17.50	4.29	0.29	8.00	0.05	3.00
Goose Creek	GOO044.36	0.00	13.00	3.92	0.33	7.00	0.14	3.00
Hughes River	HUE000.20	0.00	15.60	3.97	0.29	6.80	0.02	3.00
Kerrs Creek	KRR001.54	1.50	15.00	4.08	0.29	6.50	0.02	3.00
Back Creek	BCK000.78	2.00	14.00	5.31	0.35	6.00	0.05	3.00
Popes Head	POE002.00	2.00	11.00	4.54	0.27	4.00	0.03	3.00
Hogue Creek	HOC006.23	1.50	13.50	3.93	0.36	5.00	0.02	3.25
Mill Creek (L)	MIL002.20	2.00	16.75	5.04	0.34	6.50	0.04	3.38
Great Run	GRT001.70	0.80	18.00	4.27	0.30	5.80	0.02	3.40
Hazel River	HAZ032.54	0.00	18.00	4.12	0.23	6.60	0.02	3.80
Beaver Creek	BVR003.60	2.00	17.00	4.14	0.24	9.00	0.09	4.00
Rappahannock River (175)	RPP175.51	0.00	17.00	3.81	0.21	7.00	0.09	4.00
Cedar Run	CER016.46	0.50	18.75	4.57	0.23	6.75	0.02	4.00
Turley Creek	TRL000.02	2.00	12.25	4.88	0.43	6.50	0.03	4.38
Mountain Run	MTN000.59	1.20	19.40	4.91	0.24	5.40	0.04	4.40
South River (27)	STH027.08	1.00	16.00	5.60	0.63	7.00	0.01	4.50
Buffalo Creek	BLD000.22	1.00	11.00	3.70	0.25	5.00	0.00	4.50
Catoctin Creek	CAX004.78	0.00	17.80	3.70	0.22	6.60	0.05	4.60
Difficult Run	DIF000.86	1.20	14.60	4.70	0.28	4.40	0.04	4.60
Moffett Creek	MFT006.24	2.00	14.33	4.75	0.32	6.00	0.11	4.67
South Run	SOT001.44	2.00	18.00	4.92	0.29	4.00	0.02	4.90
Mill Creek (C)	MIC001.00	2.00	14.33	5.44	0.31	4.67	0.03	5.33
Bull Run	BUL010.28	2.00	15.75	5.35	0.24	4.50	0.02	6.00
Robinson River	ROB001.90	0.00	18.20	3.88	0.24	8.00	0.04	6.20
Rappahannock River (147)	RPP147.10	0.00	18.67	3.19	0.25	8.00	0.14	6.67
Holmans Creek	HMN002.09	2.00	11.75	5.01	0.42	5.75	0.03	8.00
Linville	LNV000.71	2.00	15.00	5.96	0.33	5.00	0.01	10.00
Little River	LIV004.78	1.33	14.67	4.40	0.24	4.67	0.03	10.67
Muddy Creek (rt)	MUD005.81	2.00	12.00	6.48	0.38	3.33	0.02	14.33
Hays Creek	HYS001.41	2.00	16.00	4.44	0.24	6.00	0.03	15.50
Pleasant Run	PLE000.08	2.67	7.33	7.49	0.64	0.67	0.00	20.17
Cooks Creek	CKS003.04	2.25	12.50	5.95	0.42	2.75	0.01	29.25

Generally, as the TSS values increase, the RBP, MFBI, and DOMPCT values should also increase while TAXA, EPTI, and S-T values should decrease. However, as the results presented in Table 4.3 demonstrate, those trends are not consistent since the dependent variables (RBP value and the metrics) do not follow their respected ascending or descending order throughout the 34 stations. The inconsistency may be attributed to the narrow range in TSS values with the majority of the values varying from 3.0 and 7.0 mg/L. Yet, when TSS values exceed 10 mg/L, trends start to appear and the values of the dependent variables generally follow their expected trends. For instance, the RBP values for TSS greater than 10 mg/L are relatively high in comparison to the rest of the RBP values. Since RBP should increase with an increase in sediment, the trend may actually be upheld. To help further verify the relationship between the benthic impairments and sediment and other stressor variables, several statistical analyses procedures were performed.

The variables in the data set that investigated TSS and other stressors were sediment indicators, habitat parameters, water quality data, landuse variables, watershed characteristics, and livestock numbers. Thus, a total of 29 variables were used as the independent variables that could influence the RBP value and/or five RBP metrics. The independent variables, listed by category, along with their associated abbreviations are shown in Table 4.4. The variable list presented in Table 4.4 has both TSS and turbidity as sediment indicators. TSS was the original parameter selected to represent sediment for its ease in measurement and applicability to all landuse types. Turbidity was added since it is a measurement of the clarity of the water and is a measure of sediment particles suspended in water. The sediments originate from eroded and/or disturbed soil, which is transported with runoff to downstream water bodies. Even though algae are also associated with turbidity, TSS and turbidity are closely related to one another (Wood and Armitage, 1997). Accordingly, both TSS and turbidity were used as sediment indicators in this study.

Table 4.4 The 29 linear independent stressor variables and their abbreviations organized by stressor type

Sediment	Water Quality	Habitat	GIS/% Landuse	Watershed Characteristics	Livestock Numbers
TSS Turb = Turbidity	Alk = Alkalinity Cond = Conductivity DO = Dissolved oxygen Hard = Hardness pH Temp = Temperature	Alter = Channel Alteration Banks = Bank stability Bankveg = Bank vegetation Cover Embed = Embeddedness Flow Graze Riffles Ripveg = Riparian vegetation Sediment = Deposition Subs = Substrate Vel = Stream Velocity	Ag = Agriculture Cropland Forest Pasture Urban	Length = Stream length Order = Stream order Wshed area	Beef = Beef cattle w/ stream access

Statistical analyses were performed to determine a suitable relationship between the benthic community and possible stressors. Both linear and non-linear relationships were performed. Based upon scientific theory and relationships detected from graphical representations of the data and correlation results, the following cross-product variables were also added: DO x Temperature, Alkalinity x Hardness, TSS x Turbidity, TSS x Channel Alteration, TSS x Bank Vegetation, TSS x % Agricultural Area. Cross-product terms were used in situations where interaction was present between variables since the variables were highly correlated with or dependent upon one another (Ott, 1993). In addition to interactions between variables, an independent variable may not be linearly related to the dependent variable and would need to be transformed to help linearize the data (Ott, 1993). Accordingly, the 29 original variables and 6 cross-product variables were all log transformed and squared. Therefore, the 29 linear variables (Table 4.4) were multiplied or transformed to arrive at 105 non-linear variables (29 linear, 6 cross-products, 35 log transformed, 35 squared). To determine the relationship between the stressors and the benthic macroinvertebrate assemblage conditions, multiple regression techniques, correlation analysis, principal component analysis, and r-square analysis were used to develop and/or explain relationships.

4.2.1 Multiple Regression Techniques

The forward, backward, and stepwise multiple regression techniques were used in SAS to investigate linear and non-linear possibilities (SAS, 1999). A significance level of 0.05 ($p = 0.05$) was used in all cases so that only the most representative independent variables would be included in the final equations (Smith, 2001). All three regression techniques were applied to nine subsets: “fall” data, “spring” data, “all” data regardless of season, an average of each station’s fall data (“fall average”), an average of each station’s spring data (“spring average”), an average of each station’s complete data (“all average”), plus average annual values for each station for “1997”, “1998”, and “1999”. Data from each station’s entire number of samples as well as the average of samples were investigated since benthic macroinvertebrates integrate the effects of short-term environmental alterations with sensitive life stages responding quickly to stress, while long-term effects are incorporated through the entire benthic community which responds slowly (Barbour et al., 1999). The annual data subsets were investigated to determine if the long-term variations could be distinguished through yearly patterns. The annual subsets included samples from the fall of the previous year and the spring of the year mentioned. For instance, the 1997 annual average corresponded with the average values from the fall of 1996 and the spring of 1997. Therefore, six possible dependent scenarios (RBP values, TAXA, MFBI, DOMPCT, EPTI, and S-T) were evaluated, both linearly and non-linearly, for nine subsets ("all", "all average", "fall", "fall average", "spring", "spring average", "1997", "1998", and "1999"), using stepwise, backward, and forward techniques.

The statistical results indicated the stepwise procedure provided the highest R^2 values and was generally more likely to include TSS. Accordingly, only the stepwise procedure was used to determine which dependent variable and seasonal or annual subset was ideal in developing the relationships between the benthic macroinvertebrates and their stressors. The stepwise statistical results are summarized in Table C.2, Appendix C. The results of the statistical analyses are discussed in the following sections for both RBP relationships and metric relationships.

4.2.1.1 RBP Relationships

Eighteen possible relationships that used RBP values to describe the benthic macroinvertebrate assemblage conditions were formulated using the stepwise multiple regression technique. The eighteen relationships consisted of nine linear and nine non-linear relationships to correspond with the nine subsets ("all", "all average", "fall", "fall average", "spring", "spring average", "1997", "1998", and "1999"). Tables 4.5 and 4.6 present the linear and non-linear RBP relationships, respectively. Tables 4.5 and 4.6 provide information on which variables were included in a relationship; whether the variables were positively or negatively associated with the RBP values; the sampling size used to develop the relationship; the R^2 values; and each relationship's R^2 rank from 1-9. A rank of 1 for R^2 referred to the relationship that explained the highest amount of variability. To help clarify the abbreviations used for the variables, refer to Table 4.4.

Table 4.5 The linear RBP results using the stepwise multiple regression technique

	All[a]	All Avg[a]	Fall[a]	Fall Avg[a]	Spring[a]	Spring Avg[a]	1997[a]	1998[a]	1999[a]
Variables Included[b,c]	Hard, -Alk, -Order, Urban	-Vel, -Order, Urban	-Sediment, Hard, -Cropland, Urban	-Embed, -Flow, -pH, Cond, -Beef	-Vel	pH	-Vel, Urban	-Embed	-Riffles, Hard, -Alk, -Ag
No. of samples	105	34	72	33	33	24	26	28	25
R-square	0.51	0.65	0.52	0.73	0.46	0.57	0.56	0.29	0.90
R^2 Rank[d]	7	3	6	2	8	4	5	9	1

[a] One of the nine subsets investigated (all, all average, fall, fall average, spring, spring average, 1997, 1998, and 1999) as discussed in Section 4.2.1.

[b] Refer to Table 4.4 for a list of the abbreviations for the independent variables.

[c] A negative sign implies that the variable has a negative (indirect) association with the RBP values; no sign implies that the variable has a positive (direct) association with the RBP values.

[d] The R^2 values are ranked 1 to 9, high to low respectively.

Table 4.6 The non-linear RBP results using the stepwise multiple regression technique

	All[a]	All Avg[a]	Fall[a]	Fall Avg[a]	Spring[a]	Spring Avg[a]	1997[a]	1998[a]	1999[a]
Variables Included[b,c]	-Cropland, Urban, DO*Temp, Beef, -Sq(vel), -Sq(order)	-Log(flow), Log(do*temp), -Sq(embed)	Hard, -Alk, -Log(pasture), -Sq(embed)	-Log(pH), -Log(order), Log(Urban), -Sq(cover), -Sq(embed), Sq(banks)	Log(urban), -Sq(vel)	Log(pH), -Sq(vel)	Vel, Log(do*temp), -Sq(vel), Sq(banks)	Log(urban), -Sq(embed), -Sq(flow), Sq(do*temp)	-Length, -Log(vel), Log(hard), -Log(pasture)
No. of samples	105	34	72	33	33	24	26	28	25
R-square	0.62	0.64	0.51	0.81	0.57	0.66	0.78	0.65	0.89
R^2 Rank[d]	7	6	9	2	8	4	3	5	1

[a] One of the nine subsets investigated (all, all average, fall, fall average, spring, spring average, 1997, 1998, and 1999) as discussed in Section 4.2.1.

[b] Refer to Table 4.4 for a list of the abbreviations for the independent variables.

[c] A negative sign implies that the variable has a negative (indirect) association with the RBP values; no sign implies that the variable has a positive (direct) association with the RBP values.

[d] The R^2 values are ranked 1 to 9, high to low respectively.

When analyzing the relationships presented in Tables 4.5 and 4.6, it was important to look at the coefficient of determination (R^2), the number of variables, the direct or indirect association of the variables with the RBP values, and the sample size. The R^2 value was the distinguishing factor in deciding which equation was more appropriate since the R^2 value is the proportion of variability in the dependent variable that is accounted for by the independent variables (Ott, 1993). Thus, the closer the R^2 value was to 1 (100%), the more variability the independent variables explained and the stronger the relationship. One problem with using the R^2 criterion for the best-fitting regression equation is that R^2 increases for each independent variable, even when the new independent variable has very little predictive power (Ott, 1993). Therefore, an equation that included five or fewer variables was judged to be reasonable for the data used in this study (Smith, 2001). The variables that were included also needed to be correctly associated with the RBP values with a positive association implying that the variable was detrimental to the benthic macroinvertebrates while a negative association suggested that the variable was beneficial to the benthic macroinvertebrates. For instance, % urban landuse should be positively associated in the equations, suggesting that the amount of imperviousness could lead to benthic impairments. Also, since an optimal habitat was indicated with higher values, habitat variables should have a negative association with the RBP values. The last consideration was to choose a sample size of 30 or more for this study so that the results would not be distorted (Smith, 2001). Thus, an ideal relationship should explain as much variability as possible based upon no more than five independent variables that were all correctly associated with RBP and based upon a sampling size of thirty or more.

Several observations are noteworthy when evaluating the results presented in Tables 4.5 and 4.6. TSS was never included in any of the 18 relationships. One reason why TSS may not have affected the RBP values is due to the fact that RBP values are based upon the aggregation of metrics and habitat conditions against reference conditions (Barbour et al., 1999). Since the RBP values are based upon the combined impact of multiple processes, the intricacies involved in each process are minimized each time one process is combined with another. Thus, the impact of sediment or another stressor may not be as evident when the benthic macroinvertebrate assemblage conditions are depicted with the final RBP value. A variable most closely related to a sediment indicator was embeddedness, which was included in 6 out of the 18 relationships.

Instead of including sediment, most relationships included the following independent variables: % urban, velocity, embeddedness, hardness, stream order, pH, and DO x temperature. In addition, the linear relationships had an R^2 ranging from 0.29 to 0.90 while the non-linear R^2's ranged from 0.51 to 0.89. The non-linear relationships almost always showed improvement from the linear relationships in terms of the amount of variability they explained. Generally, the non-linear relationships explained an additional 10% of the variability while the most additional variability the linear relationships explained over a non-linear relationship was 1%. As shown in Tables 4.5 and 4.6, the relationship for the “average” subsets always showed improvements in R^2 values over their respected original subsets while the seasonal and annual differences were undecipherable. When comparing the “fall” results against the “spring” results, no differences were apparent in the linear or non-linear variability amounts. However, the “spring” subset was able to explain just as much variability as the “fall” subset and included fewer variables than the “fall” subset. The annual results showed differences in the type of variables they included and the amounts of variability explained in the linear relationships, whereas the annual non-linear relationships offset such problems. Moreover, the results inconclusively revealed whether or not seasonal or annual subsets affected the benthic macroinvertebrate assemblage conditions differently. Lastly, the “all average” linear and non-linear relationships (Tables 4.5 and 4.6) proved to be the most suitable linear and non-linear relationships, based upon the criteria of an ideal relationship explaining as much variability as possible with no more than five independent variables that were all correctly associated with RBP and based upon a sampling size of thirty or more.

When implementing the analysis criteria (high explanation of variability, five or fewer variables that are correctly associated, and a sample size of 30 or more) for the linear results in Table 4.5, the “all average” subset was the first relationship to meet the criteria. Evaluating the relationships one at a time based upon the R^2 ranks, the 1999 relationship did not associate agricultural area correctly with RBP and it had a small sampling size of 25. The “fall average” relationship was also inadequate since the number of beef cattle with stream access should be positively associated with RBP. The “all average” relationship, ranked third with an R^2 of 0.65, met all analysis requirements since it was able to explain over half of the variability in the benthic macroinvertebrate assemblage conditions through the velocity, stream order, and percent urban characteristics of 34 samples. To make sure no other relationships were more suitable, the fourth and fifth ranked relationships were also analyzed.

Both the fourth and fifth ranked relationships, the “spring average” and “1997” respectively, met the requirements of including five or fewer variables that were correctly associated with RBP. However, the “spring average” and “1997” relationships had small sampling sizes of 24 and 26, respectively. In addition, the two relationships each explained about 56% of the total variability, a significant drop from the 65% explained by the “all average” relationship. Accordingly, the “all average” relationship was chosen as the most suitable representation for RBP linear relationships. The equation for the linear, “all average”, RBP relationship ($R^2 = 0.65$) is:

$$\text{RBP} = 4.58 \ - \ 0.16\,(\text{velocity}) \ - \ 0.31\,(\text{stream order}) \ + \ 4.21\text{E-02}\,(\%\ \text{urban}) \qquad [4.2]$$

The relationship presented in Equation 4.2 demonstrates that a habitat variable (velocity), watershed characteristic (stream order), and landuse variable (% urban) influenced the RBP values. Sediment, water quality, or livestock variables did not account for any variability in the linear, “all average”, RBP relationship. Furthermore, the results indicated that sites with an optimal stream velocity/depth regime where slow-deep, slow-shallow, fast-deep, and fast-shallow (slow being <0.3 m/s, deep being >0.5m) regimes are all present and sites with a larger stream order were beneficial to the benthic macroinvertebrate assemblage conditions, whereas sites with large amounts of urban area were detrimental to the benthic macroinvertebrate assemblages. A summary of the stepwise procedure’s results is shown in Table C.3, Appendix C. The stepwise results indicated that stream velocity was the variable that explained most of the variability with a R^2 value of 0.48, or roughly 73% of the total variability explained in the relationship presented in Equation 4.2.

Among the non-linear RBP relationships presented in Table 4.6, the “all average” subset was the first relationship to meet all analysis criteria. The “1999” relationship, analyzed first for its 0.89 R^2 value, failed to meet the variable association requirement for pasture and its sampling size of 25 was too small. The “fall average” relationship, with a R^2 value of 0.82, did not meet all analysis criteria since it included too many independent variables (six) and the bank stability variable, which is a habitat parameter, was incorrectly associated with RBP. The third relationship analyzed, for the “1997” subset, had a R^2 value of 0.78 and also did not meet analysis criteria since stream velocity was incorrectly associated and had a small sampling size of 26. The fourth ranked R^2 value of 0.66 was for the “spring average” relationship. The number of variables and variable associations were adequate for the “spring average” relationship; however, the small sampling size of 24 was insufficient. The “1998” relationship with a R^2 value of 0.65 was evaluated next. The “1998” relationship met all requirements except for the small sample size of 28. The sixth relationship for the “all average” subset met all requirements and the 0.64 R^2 value was only 0.01 (1%) less than the 1998’s 0.65 value. Since the “all average” relationship did not sacrifice the amount of variability explained in the “1998” relationship and since the sample size of 34 was more acceptable than the “1998” subset’s 28 samples, the relationship for the “all average” subset was selected as the best non-linear relationship between the stressors and RBP values.

To make sure no other relationships could better explain the stressors and RBP relationship, the remaining possibilities listed in Table 4.6 were examined. The “all” subset’s relationship was insufficient due to too many variables (six) and incorrect association of cropland with the RBP values. The “spring” relationship met all requirements; however, it had a R^2 value that was 7% less than the “all average” relationship. The “fall” relationship also explained a minimal amount of variability when compared to the “all average” and incorrectly associated pasture with the RBP values. Thus, the relationship for the “all average” subset, which explained 64% of the variability, was chosen as the most representative non-linear RBP relationship:

$$\text{RBP} = 7.99 - 7.41\,(\text{Log(flow)}) + 1.91\,(\text{Log(DO x temp)}) - 5.63\text{E-03}\,(\text{embeddedness}^2) \qquad [4.3]$$

The relationship presented in Equation 4.3 demonstrates that two habitat variables (flow and embeddedness) and two water quality variables (DO x temperature) influenced RBP values. Sediment, landuse, watershed characteristics, or livestock variables did not account for any significant variability in the non-linear, “all average”, RBP relationship. When the flow is at an optimal range where the water reaches the base of both lower banks and minimal amounts of channel substrate are exposed and when embeddedness is such that gravel, cobble, and boulder particles are 0-25% surrounded by fine sediment, flow and embeddedness are beneficial to the benthic macroinvertebrate assemblage conditions (Barbour et al, 1999). On the other hand, high values of dissolved oxygen and temperature are detrimental to the benthic macroinvertebrate assemblage conditions. A summary of the stepwise procedure's results is shown in Table C.4, Appendix C. The stepwise results indicated that the squared embeddedness term was the variable that explained most of the variability (0.50) or roughly 77% of the total variability explained by Equation 4.3.

In summary, the most suitable RBP relationships for both linear and non-linear results were provided by the "all average" subset. Each of the linear and non-linear relationships was based upon three variables and the amount of variability explained was around 65%. The only notable difference between the two relationships was the type of independent variables included in the relationship. Each relationship was based upon three different variables that were significant at the 0.05 level with the combined six variables being: velocity, stream order, % urban, flow, DO x temperature, and embeddedness. Overall, the two relationships included variables that could be determined or measured in the field and the relationships explained over half of the variability in the RBP values. However, since the RBP relationships never included TSS, the sediment variable, and in an effort to develop stronger relationships, the individual RBP metrics were examined and used as the dependent variables in the stressor/benthos relationships.

4.2.1.2 Metric Relationships

RBP metrics were used in addition to the RBP values to describe the benthic macroinvertebrate assemblage conditions. Even though the RBP procedures and RBP value classifications are changing in Virginia, the methods used to evaluate the individual metrics are remaining the same. In addition, the individual RBP metrics are more representative of the benthos than the final RBP value since the metrics represent attributes of a targeted aquatic assemblage while the RBP value is based upon the aggregation of the metrics and habitat conditions against reference conditions (Barbour et al., 1999). Thus, the metrics, in addition to the RBP values, were used in the analyses since the individual metrics focus on a specific biological condition one at a time.

The metrics investigated in this study were taxa richness (TAXA), the modified family biotic index (MFBI), the percent dominant family (DOMPCT), the EPT index (EPTI), and the shredders to the total number of organisms (S-T). These five metrics were selected since they are the most commonly used metrics and they include diversity by investigating richness, tolerance, and feeding measures (Voshell, 2001). The TAXA metric is a richness measure that measures the total number of families present. For the "usable" stations (Table A.1, Appendix A), the TAXA values ranged from 5 to 28 and generally increase with improvements in water quality. The MFBI metric is a tolerance measure that reports tolerances from 0 to 10 and increases as water quality becomes degraded due to organic pollution. The MFBI range for the "usable" stations was from 2.75 to 7.95. The DOMPCT metric is a tolerance measure and was reported as a decimal ranging from 0 to 1 with the "usable" stations' data ranging from 0.139 to 0.962. A community dominated by few families has a low DOMPCT number, which indicates environmental stress. The EPTI metric is a richness measure that reports the number of distinct taxa within the EPT (Ephemeroptera, Plecoptera, and Trichoptera) pollution sensitive groups. The "usable" stations' EPTI values ranged from 0 to 12 with the higher numbers generally indicating improved water quality. The S-T metric is a feeding measure that is reported as a decimal ranging from 0 to 1 with values for the "usable" stations only ranging from 0 to 0.236. Since shredders are good indicators of landuse, toxics, and riparian vegetation, as the S-T value increases, water quality is degraded.

An investigation of the 5 metrics (TAXA, MFBI, DOMPCT, EPTI, and S-T) for 9 subsets ("all", "all average", "fall", "fall average", "spring", "spring average", "1997", "1998", and "1999"), both linearly and non-linearly, resulted in 90 stepwise metric relationships (45 linear and 45 non-linear). The linear and non-linear stepwise metric relationships were further analyzed based upon: relationships either included a sediment indicator as a significant independent variable or the relationships explained a high amount of variability for the benthic macroinvertebrate assemblage conditions. It was important to group the analysis into these two categories since the objective of the study was to develop the most suitable relationship (the relationship that explained the most variability), while focusing on the evaluation of the relationship between sediment and the benthic macroinvertebrate assemblages. Therefore, the original 90 relationships were reduced into 17 relationships that included sediment and 18 relationships that explained the most variability.

4.2.1.2.1 Metric Relationships Which Included a Sediment Indicator as an Independent Variable

Summaries of the linear and non-linear stepwise regression results that included a sediment indicator are provided in Tables 4.7 and 4.8. Both TSS and turbidity were considered to be sediment indicator variables. TSS was the original parameter selected to represent sediment for its ease in measurement and applicability to all landuse types. Turbidity was added since it is a measurement of the clarity of the water and is caused by sediment particles suspended in water. The sediments originate from eroded and/or disturbed soil, which is transported with runoff to downstream water bodies. Even though algae are also associated with turbidity, TSS and turbidity are closely related to one another (Wood and Armitage, 1997). Accordingly, both TSS and turbidity were used as sediment indicators in this study.

Tables 4.7 and 4.8 show the variables that were included for the seasonal or annual subsets under a dependent variable as well as the sample size, R^2 values, and the R^2 rankings ranging anywhere from 1 (high value) to 10 (low value). To help clarify the abbreviations used for the variables, refer to Table 4.4.

The same analysis criteria as the RBP relationships were used with the metric results presented in Tables 4.7 and 4.8. The goal was to select relationships which had: a R^2 value close to one (100% variability), a few number of variables ($\leq$ 5) that were associated correctly with the metrics, and a sample size of thirty or more. For EPTI, S-T, and TAXA metrics, a positive association implied that as the variable value increased, the variable was beneficial to the benthic macroinvertebrates and a negative association was detrimental to the benthic macroinvertebrates. The DOMPCT associations were just the opposite: positive implied detrimental while negative suggested beneficial impacts to the benthic conditions. Thus, TSS and turbidity should be negatively associated in EPTI, TAXA, and S-T relationships while positively associated in DOMPCT relationships. Furthermore, since an optimal habitat is indicated with a higher variable value, the habitat variables should have a positive association with EPTI, TAXA, and S-T and a negative association with DOMPCT.

Table 4.7 The stepwise, linear, metric results that included a sediment indicator

	All[a]		All Avg[a]	Fall Avg[a]	Spring[a]			Spring Avg[a]		1997[a]
	TAXA	EPTI	EPTI	EPTI	TAXA	EPTI	S-T	TAXA	S-T	EPTI
Variables Included[b,c]	-TSS, Graze, -DO, Order, -Urban	-TSS, Cover, Subs, Flow, -DO, -Temp, pH, - Wshed area, -Urban	Embed, Flow, -DO, -Temp, -Turb	Riffles, Flow, -Turb	-TSS, Subs, Pasture	-TSS, Subs, Flow, -DO, pH, -Beef, -Urban	-TSS, Sediment, -Ripveg	-TSS, -DO, Turb	-TSS, Sediment, -Ripveg, Turb	-TSS, -DO, Forest
No. of samples	105		34	33	33			24		26
R-square	0.41	0.57	0.66	0.51	0.53	0.77	0.51	0.65	0.73	0.78
R^2 Rank[d]	10	6	4	8	7	2	9	5	3	1

[a] One of the nine subsets investigated (all, all average, fall, fall average, spring, spring average, 1997, 1998, and 1999) as discussed in Section 4.2.1.

[b] Refer to Table 4.4 for a list of the abbreviations for the independent variables.

[c] A negative sign implies that the variable has a negative (indirect) association with the metric values; no sign implies that the variable has a positive (direct) association with the metric values.

[d] The R^2 values are ranked 1 to 10, high to low respectively.

Table 4.8 The stepwise, non-linear, metric results that included a sediment indicator

	All[a]	All Avg[a]	Fall Avg[a]		Spring[a]	Spring Avg[a]	1997[a]
	EPTI	EPTI	DOMPCT	EPTI	TAXA	EPTI	EPTI
Variables Included[b,c]	Flow, DO, -(TSS*ag), Log(subs), -Log(temp), Log(pH), -Log(cond), Sq(cover), -Sq(DO)	-Log(temp), -Log(TSS*turb), Sq(flow), -Sq(DO)	(TSS*ag), -Log(cover), -Log(order), -Sq(wshed area), Sq(order)	-Flow, Turb, Log(riffles), -Log(cond), -Log(wshed area), -Log(DO*temp), -Log(TSS*turb), Log(TSS*ag)	-Sq(TSS*ag)	-Urban, -Sq(beef), -Sq(DO*temp), -Sq(TSS*bankveg)	Forest, -Sq(TSS), -Sq(DO)
No. of samples	105	34	33		33	24	26
R-square	0.66	0.70	0.75	0.98	0.39	0.81	0.78
R^2 Rank[d]	6	5	4	1	7	2	3

[a] One of the nine subsets investigated (all, all average, fall, fall average, spring, spring average, 1997, 1998, and 1999) as discussed in Section 4.2.1.

[b] Refer to Table 4.4 for a list of the abbreviations for the independent variables.

[c] A negative sign implies that the variable has a negative (indirect) association with the metric values; no sign implies that the variable has a positive (direct) association with the metric values.

[d] The R^2 values are ranked 1 to 7, high to low respectively.

As shown in Tables 4.7 and 4.8, both the linear and non-linear results generally had a sediment indicator related to the EPTI or TAXA metric. Both EPTI and TAXA are richness measures used to determine the adequacy of niche space, habitat, and food source (Barbour et al., 1999). Since sediment affects niche space, habitat, and the food source, it was appropriate for the EPTI and TAXA metric to be related to sediment. The ten linear relationships that included a sediment indicator had R^2 values ranging from 0.41 to 0.78 while the R^2 values for the seven non-linear relationships that included a sediment indicator ranged from 0.39 to 0.98. The inclusion of a sediment indicator did not appear to vary by season, as a sediment indicator variable was included in both “fall” and “spring” relationships. However, the effect of sediment upon the benthic macroinvertebrates did appear to vary among the three years studied since sediment was only significant during the 1997. Lastly, besides TSS and turbidity, the relationships generally included dissolved oxygen, flow or some other habitat variable such as substrate or riffles, and % urban landuse.

Among the linear relationships in Table 4.7, the “all average”, EPTI relationship was the first relationship to meet all analysis criteria. The “1997”, EPTI relationship had a small sample size of 26 and sediment was not included in the similar relationships developed for the “1998” and “1999” subsets. The “spring”, EPTI relationship included too many variables (seven). The “spring average”, S-T relationship had a sample size of 24 that was too small and the association of riparian vegetation and turbidity did not match theory. The “all average”, EPTI relationship included five variables with correct association, had an adequate sample size of 34, and had the next highest R^2 value of 0.66. The next two relationships with the highest R^2 values were the “spring average”, TAXA that had a small sample size and wrong association for turbidity and the “all”, EPTI that included nine variables. Since the “all average”, EPTI equation was the first relationship to meet all considerations and since no other relationships meet all criteria, the relationship was selected as the linear sediment model. The equation for the linear, “all average”, EPTI relationship (R^2 = 0.66) is:

$$\begin{aligned}\text{EPTI} = 7.61 \;+\; 0.14\,(\text{embeddedness}) \;+\; 0.58\,(\text{flow}) \;-\\ 0.95\,(\text{DO}) \;-\; 0.20\,(\text{temperature}) \;-\; 0.21\,(\text{turbidity})\end{aligned} \qquad [4.4]$$

The relationship presented in Equation 4.4 demonstrates that turbidity, two additional water quality variables (DO and temperature), and two habitat variables (embeddedness and flow) influenced EPTI. No landuse, watershed characteristics, or livestock variables accounted for any significant variability in the linear, “all average”, EPTI relationship. The associations suggest that higher values of turbidity, temperature, and dissolved oxygen impair the benthic macroinvertebrates. However, when the flow is in optimal range where the water reaches the base of both lower banks and minimal amounts of channel substrate are exposed and when embeddedness is such that gravel, cobble, and boulder particles are 0-25% surrounded by fine sediment, flow and embeddedness are beneficial to the benthic macroinvertebrate assemblage conditions (Barbour et al., 1999). A summary of the stepwise procedure’s results is shown in Table C.5, Appendix C. The stepwise results indicated that turbidity was the variable that explained most of the variability with a R^2 value of 0.35, or roughly 53% of the total variability.

Using the analysis criteria for the non-linear results presented in Table 4.8, the fifth ranked relationship, the “all average”, EPTI, was selected since the relationship had a R^2 value of 0.70, an adequate sample size of 34, and included four variables correctly associated with EPTI. The “fall average”, EPTI relationship included too many variables (eight) and three of the variables (flow, turbidity, and log(TSS x ag)) were associated incorrectly with EPTI. The “spring average”, EPTI relationship had a small sample size of 24. The “1997”, EPTI relationship also had a small sample size of 26 and since no other annual subsets ever included a sediment indicator in their relationships, it was best to explore other data sets. The “fall average”, DOMPCT relationship included two variables that were associated with stream order, one directly related and the other indirectly related. Furthermore, DOMPCT was not a typical benthic community indicator that sediment significantly contributed to.

The only two remaining relationships, “all”, EPTI and “spring”, TAXA, had too many variables, incorrect relationships, and/or a small R^2 value. Consequently, the “all average”, EPTI relationship was chosen as the best non-linear relationship which included a sediment indicator. The equation for the non-linear, “all average”, EPTI relationship (R^2 = 0.70) is:

$$\begin{aligned}\text{EPTI} = {} & 17.55 - 6.31\,(\text{Log(temperature)}) - \\ & 2.67\,(\text{Log(TSS x turbidity)}) + 1.52\text{E-}02\,(\text{flow}^2) - \\ & 5.01\text{E-}02\,(\text{DO}^2)\end{aligned} \qquad [4.5]$$

The relationship presented in Equation 4.5 demonstrates that the log of TSS times turbidity, two additional water quality variables (temperature and DO), and a habitat variable (flow) influenced EPTI. Landuse, watershed characteristics, and livestock variables did not account for any variability in the "all average", EPTI relationship. The associations suggest that higher values of temperature, TSS, turbidity, and dissolved oxygen impair the benthic macroinvertebrates while optimal flow characteristics where the water reaches the base of both lower banks and minimal amounts of channel substrate are exposed improve benthic macroinvertebrate assemblage conditions (Barbour et al, 1999). A summary of the stepwise procedure's results is shown in Table C.6, Appendix C. The stepwise results indicated the log of TSS times turbidity was the variable that explained most of the variability with a R^2 value of 0.37, or roughly 53% of the total variability.

After having analyzed both the linear and non-linear metric results for relationships that included a sediment indicator variable, the "all average", EPTI relationship was chosen in both cases. Both the linear and non-linear relationships included a flow variable, dissolved oxygen, temperature, and turbidity. The linear model also included embeddedness while the non-linear model included TSS. Of the two relationships, the non-linear representation is preferred since it does incorporate TSS, uses one less variable (4 versus 5), and resulted in a slightly higher R^2 value (0.70). Therefore, the selected relationship was the non-linear, stepwise, "all average", EPTI. The equation (Equation 4.5) includes four variables, has a sample size of 34, results in a R^2 value of 0.70, and sediment accounts for roughly half of the variability (53% of the total variability).

4.2.1.2.2 Metric Relationships With the Highest R^2 Values

Another means of analyzing the metric relationships was to consider all variables equally instead of focusing upon relationships which included a sediment indicator as an independent variable. The goal was to arrive at the most suitable equation that demonstrates the relationship between the benthic community and possible stressors. Since nine subsets ("all", "all average", "fall",

"fall average", "spring", "spring average", "1997", "1998", and "1999") were investigated for both linear and non-linear representations for five RBP metrics (TAXA, MFBI, DOMPCT, EPTI, and S-T), a total of 90 relationships were investigated. To determine which relationship best represents the interactions between the benthic macroinvertebrates and stressors, only the metric for each of the nine subsets that produced the highest R^2 value was singled out and analyzed. Thus, a total of 18 relationships, 9 linear and 9 non-linear for each of the nine subsets, were focused upon. Tables 4.9 and 4.10 show the relationships that produced the highest R^2 values for both linear and non-linear representations. Within Tables 4.9 and 4.10, the variables that were included for the seasonal or annual subsets under an individual metric are listed as well as the sampling size, R^2 value, and R^2 ranking. To help clarify the abbreviations used for the variables, refer to Table 4.4.

The results presented in Tables 4.9 and 4.10 were analyzed using the same criteria as those used for the RBP relationships and metrics relationships that included a sediment indicator as an independent variable. The criteria included: a R^2 value close to one (100% variability), few number of variables ($\leq$ 5) that were associated correctly with the metrics, and a sample size of thirty or more. As mentioned previously, for EPTI, S-T, and TAXA, a positive association implied that as the magnitude of the variable increases it would be more beneficial to the benthic macroinvertebrates and a negative association implied that as the magnitude of the variable increases it was detrimental to the benthic macroinvertebrates. The MFBI metric followed the opposite pattern. A positive association implied that an increase in magnitude for a variable was beneficial to the benthic macroinvertebrates and a negative association implied that an increase in magnitude for a variable was detrimental to the benthic macroinvertebrates.

Table 4.9 The stepwise, linear, metric results that produced the highest R^2 values

	All[a]	All Avg[a]	Fall[a]	Fall Avg[a]	Spring[a]	Spring Avg[a]	1997[a]	1998[a]	1999[a]
	MFBI	MFBI	MFBI	MFBI	MFBI	MFBI	S-T	MFBI	TAXA
Variables Included[b,c]	-Vel, DO, -Order, Pasture, Urban	-Flow, Hard, -Alk	-Subs, -Flow, DO, -Forest, Urban	-Flow, Hard, -Alk	-Flow, DO, Temp, Cond, Length, -Cropland	DO, pH, Cond	-DO, -Temp, pH, -Cond, Beef	-Subs, -Flow, DO, Pasture, Urban	Sediment, Temp, Beef, Order
No. of samples	105	34	72	33	33	24	26	28	25
R-square	0.62	0.71	0.61	0.67	0.80	0.83	0.83	0.86	0.80
R^2 Rank[d]	8	6	9	7	5	3	2	1	4

[a] One of the nine subsets investigated (all, all average, fall, fall average, spring, spring average, 1997, 1998, and 1999) as discussed in Section 4.2.1.

[b] Refer to Table 4.4 for a list of the abbreviations for the independent variables.

[c] A negative sign implies that the variable has a negative (indirect) association with the metric values; no sign implies that the variable has a positive (direct) association with the metric values.

[d] The R^2 values are ranked 1 to 9, high to low respectively.

Table 4.10 The stepwise, non-linear, metric results that produced the highest R^2 values

	All[a]	All Avg[a]	Fall[a]	Fall Avg[a]	Spring[a]	Spring Avg[a]	1997[a]	1998[a]	1999[a]
	MFBI	MFBI	MFBI	EPTI	MFBI	MFBI	S-T	MFBI	TAXA
Variables Included[b,c]	DO, Log(urban), -Sq(subs), -Sq(vel), Sq(hard), -Sq(alk), -Sq(alk*hard)	-(Alk*hard), -Log(flow), Sq(hard)	Log(urban), -Sq(subs), Sq(hard), -Sq(alk)	-Flow, Turb, Log(riffles), -Log(cond), -Log(wshed area), -Log(DO*temp), -Log(TSS*turb), Log(TSS*ag)	Cond, -Log(flow), -Sq(cropland), Sq(DO*temp)	Cond, Sq(DO), Sq(pH)	-Cond, -(DO*temp), Log(pH), Sq(beef)	Urban, -Log(subs), -Sq(flow), Sq(DO), Sq(ag)	Beef, Sq(temp), Sq(order)
No. of samples	105	34	72	33	33	24	26	28	25
R-square	0.70	0.74	0.64	0.98	0.76	0.83	0.82	0.88	0.78
R^2 Rank[d]	8	7	9	1	6	3	4	2	5

[a] One of the nine subsets investigated (all, all average, fall, fall average, spring, spring average, 1997, 1998, and 1999) as discussed in Section 4.2.1.

[b] Refer to Table 4.4 for a list of the abbreviations for the independent variables.

[c] A negative sign implies that the variable has a negative (indirect) association with the metric values; no sign implies that the variable has a positive (direct) association with the metric values.

[d] The R^2 values are ranked 1 to 9, high to low respectively.

The results presented in Tables 4.9 and 4.10 generally had the MFBI metric as the dependent variable that produced the highest R^2 value for both the linear and non-linear representations. The MFBI metric is a measure of tolerance developed specifically to detect organic pollution. Since organic pollution is related to various stressors and widely impairs benthic macroinvertebrates, the MFBI metric relationships appropriately produced high R^2 values. S-T, TAXA, and EPTI metrics were also included in the relationships with high R^2 values; however, MFBI accounted for 13 of the 18 relationships that resulted in the highest R^2 value. The R^2 values for the linear relationships ranged from 0.61 to 0.86. The R^2 values for the non-linear relationships were similar with a range from 0.64 to 0.98. For both the linear and non-linear representations, the relationships for the "annual" subsets and the "spring average" subsets tended to produce higher R^2 values than any of the other subsets.

As shown in Table 4.9, the MFBI metric for the "1998" subset was ranked first with the highest R^2 value of 0.86. The relationship included four variables associated correctly with MFBI; however, the relationship was developed from a borderline sample size of 28. Overall the "1998", MFBI relationship met the analysis criteria, except the sampling size was questionable. Accordingly, other relationships were analyzed to evaluate the existence of any other suitable relationships.

The next three relationships with the highest R^2 values were the S-T metric for the "1997" subset ($R^2 = 0.8267$); the MFBI metric for the "spring average" subset ($R^2 = 0.8257$); and, the TAXA metric for the "1999" subset ($R^2 = 0.8033$). The "1997", S-T relationship was inadequate since it had a small sample size of 26 and the beef cattle variable was associated incorrectly with S-T. The "spring average", MFBI relationship associated all three of its variables correctly; however, the sampling size of 24 was inadequate. The "1999", TAXA relationship had a small sampling size of 25 and did not associate the sediment habitat variable and the number of beef cattle correctly with TAXA. Thus, only the "spring average", MFBI and the "1998", MFBI relationships were able to correctly associate all independent variables with MFBI, yet both relationships had questionable sample sizes. Since all other analysis criteria, except sample size, were achieved for both data relationships, the two relationships needed to be compared against one another.

The “1998”, MFBI included 5 variables, had a sample size of 28, and a R^2 of 0.86. The “spring average”, MFBI included 3 variables, had a sample size of 24, and a R^2 of 0.83. Thus, the “1998” subset was preferred for its larger sample size and slightly higher R^2 value compared with the “spring average” subset. On the other hand, the “spring average” subset was preferred since it minimized 2/5 of the variables at the cost of only losing 3% of the variability. In addition, the “spring average” subset is preferred in terms of its future applicability. A seasonal subset has fewer deviations in flow and climatic patterns than an annual subset. Seasons are more likely to exemplify the same trends season after season and can be compared with one another; however, trends are not as continuous on an annual basis. For instance, one year might have a flood or unseasonable cold temperatures that would dramatically impact the benthic macroinvertebrate conditions, making comparisons to other years unacceptable. Consequently, the “spring average”, MFBI relationship is preferred over the “1998”, MFBI relationship when factoring in applicability of the subsets to future uses.

In summary, both the “1998”, MFBI and “spring average”, MFBI relationships were the most adequate linear, metric relationships with the highest R^2 values; however, the relationship developed for the “spring average” subset is preferred since it minimized the number of variables and would be applicable to future seasonal data. On the other hand, the “1998”, MFBI relationship was preferred solely based upon the higher R^2 value and larger sample size. Accordingly, both equations are described in Equations 4.6 and 4.7. The equation for the linear, “spring average”, MFBI metric relationship ($R^2 = 0.83$) is presented in Equation 4.6:

$$\text{MFBI} = -5.36 + 0.21\,(\text{DO}) + 0.92\,(\text{pH}) + 1.32\text{E-}03\,(\text{conductivity}) \qquad [4.6]$$

Equation 4.6 demonstrates that three water quality variables (DO, pH, and conductivity) significantly influenced the MFBI values. Sediment, habitat, watershed characteristics, percent landuse, and livestock numbers did not account for any variability in the linear, “spring average”, MFBI relationship. The associations suggest that higher values of dissolved oxygen, pH, and conductivity impair the benthic macroinvertebrates. A summary of the stepwise procedure’s results is shown in Table C.7, Appendix C. The stepwise results indicated pH was the variable that explained most of the variability with a R^2 value of 0.59, or roughly 72% of the total variability.

The equation for the linear, “1998”, MFBI metric relationship is:

$$\begin{aligned} \text{MFBI} = 5.83 \;-\; 0.16\,(\text{substrate}) \;-\; 0.13\,(\text{flow}) \;+ \\ 0.28\,(\text{DO}) \;+\; 2.87\text{E-02}\,(\%\ \text{pasture}) \;+\; 6.85\text{E-02}\,(\%\ \text{urban}) \end{aligned} \qquad [4.7]$$

Equation 4.7 included two habitat variables (substrate and flow), one water quality variable (DO), and two landuse variables (% pasture and % urban). Watershed characteristics and livestock numbers were not included in the final equation. The variable associations suggest that higher values of dissolved oxygen, pastureland, and urban areas impair the benthic macroinvertebrate assemblage conditions. Optimal substrate characteristics (a mixture of substrate materials with gravel and firm sand with the addition of root mats and submerged vegetation) and optimal flow characteristics (where the water reaches the base of both lower banks and minimal amounts of channel substrate are exposed), improve the benthic macroinvertebrate assemblage conditions (Barbour et al., 1999). A summary of the stepwise procedure’s results is shown in Table C.8, Appendix C. The stepwise results indicated the substrate variable explained most of the variability with a R^2 value of 0.50, or roughly 58% of the total variability.

The non-linear results presented in Table 4.10 were analyzed beginning with the EPTI metric for the “fall average” subset. The “fall average”, EPTI relationship included too many variables (eight) and three of the variables (flow, turbidity, and TSS x agricultural land) were associated incorrectly with EPTI. The next highest R^2 of 0.88 resulted from the “1998”, MFBI relationship. The relationship included five variables that were associated correctly with MFBI; however, the sample size of 28 was questionable. To verify whether or not to accept this relationship despite the borderline sample size, the next three highest R^2 relationships were investigated. The “spring average”, MFBI relationship met all requirements; yet, it too had a small sample size of 24. The “1997”, S-T and the “1999”, TAXA relationships were both disregarded since the number of beef cattle were incorrectly associated with the respected metric. Consequently, the non-linear

"1998", MFBI and the "spring average", MFBI relationships needed to be compared as before during the linear investigations.

The "1998", MFBI included 5 variables, had a sample size of 28, and a R^2 of 0.88. The "spring average", MFBI included 3 variables, had a sample size of 24, and a R^2 of 0.83. Thus, the relationship for "1998" subset was preferred for its larger sample size and slightly higher R^2 value that explained 5% more than the "spring average" subset. On the other hand, the "spring average" subset was preferred since it minimized the variables with minimal loss in the R^2 value. However, the "spring average" subset is preferred in terms of its future applicability. Consequently, the "spring average", MFBI relationship is preferred over the "1998", MFBI relationship when factoring in applicability of the subsets to future uses.

The equation for the non-linear, "spring average", MFBI relationship ($R^2 = 0.83$) is:

$$\text{MFBI} = -0.48 + 1.29\text{E-}03\,(\text{conductivity}) + 9.66\text{E-}03\,(\text{DO}^2) + 5.63\text{E-}02\,(\text{pH}^2) \tag{4.8}$$

Equation 4.8 demonstrates that three water quality variables (conductivity, DO, and pH) influenced the MFBI values. Sediment, habitat, watershed characteristics, percent landuse, and livestock numbers did not account for any significant variability in the non-linear, "spring average", MFBI relationship. The associations suggest that higher values of conductivity, dissolved oxygen, and pH impair the benthic macroinvertebrate assemblage conditions. A summary of the stepwise procedure's results is shown in Table C.9, Appendix C. The stepwise results indicated the squared pH term explained most of the variability with a R^2 value of 0.60, or roughly 72% of the total variability.

The MFBI metric for the "1998" subset was the best relationship when considering R^2 values and sample size differences between the "1998", MFBI relationship and the "spring average", MFBI relationship. The equation for the non-linear, "1998", MFBI relationship is explained by Equation 4.9:

$$\begin{aligned}\text{MFBI} = 9.70 &+ 6.63\text{E-}02\ (\%\ \text{urban}) - 4.75\ (\text{Log(substrate)}) - \\ &4.69\text{E-}03\ (\text{flow}^2) + 1.46\text{E-}02\ (\text{DO}^2) + \\ &2.01\text{E-}04\ ((\%\ \text{agriculture})^2)\end{aligned} \qquad [4.9]$$

Equation 4.9 included two habitat variables (substrate and flow), one water quality variable (DO), and two landuse variables (% urban and % agriculture). No watershed characteristics or livestock numbers were included in the final relationship. The variable associations suggest that higher values of urban area, dissolved oxygen, and agriculture land impair the benthic macroinvertebrate assemblage conditions. Optimal substrate characteristics (a mixture of substrate materials with gravel and firm sand with the addition of root mats and submerged vegetation) and optimal flow characteristics (a flow where the water reaches the base of both lower banks and minimal amounts of channel substrate are exposed), improve the benthic macroinvertebrate assemblage conditions (Barbour et al., 1999). A summary of the stepwise procedure's results is shown in Table C.10, Appendix C. The stepwise results indicated the log of the substrate variable explained most of the variability with a R^2 value of 0.53, or roughly 60% of the total variability.

The relationships for the "1998" and "spring average" subsets provided the best stressor/benthos relationships, both linearly and non-linearly. Both the linear and non-linear "1998" subsets included a form of flow, dissolved oxygen, substrate, % pasture, and % urban land as independent variables into the final equation while the "spring average" linear and non-linear representations both included dissolved oxygen, pH, and conductivity. In addition, both the "1998" and "spring average" subsets produced non-linear relationships with higher R^2 values than the linear relationship. With no differences in the number and type of variables between linear and non-linear relationships and improvements in the non-linear R^2 values, the non-linear relationships were favored for the "1998" and "spring average" subsets. The relationship developed with the "spring average" subset was preferred to the "1998" relationship since the "spring average" relationship is applicable to future uses. Therefore, the best relationship between stressors and stream benthic macroinvertebrate assemblage conditions was the non-linear, stepwise, "spring average", MFBI relationship (Equation 4.8) that included three variables, had a sample size of 24, resulted in a R^2 value of 0.83, and the pH term accounted for roughly 72% of the total variability.

4.2.1.3 Summary of the Multiple Regression Techniques

Relationships between stressors and benthic macroinvertebrate assemblage conditions for the data set that investigated TSS and other stressors were established using the stepwise multiple regression techniques. The stepwise relationships were analyzed separately based upon the RBP values and their metric values. Overall, the metric relationships were favored over the RBP relationships due to their higher R^2 values and the fact that only the individual metrics, but not the RBPs were influenced by sediment. However, the selected RBP relationships still explained 65% of the variability and contained measurable variables as seen with the non-linear, “all average”, RBP relationship that had a R^2 value of 0.64 based upon flow, DO x temperature, and embeddedness (Table 4.6; Equation 4.3).

The metric relationships demonstrated that EPTI and TAXA metrics were generally associated with sediment indicators, the MFBI metric generally explained the most variability, and DOMPCT and S-T metrics were inconsistent and unsuccessful in explaining high degrees of variability. EPTI and TAXA are richness measures that depend upon niche space, habitat, and food sources (Barbour et al., 1999). Since sediment impacts niche space, habitat, and food sources, it was appropriate for the sediment variables to be associated with the EPTI and TAXA metrics. The MFBI metric is a tolerance measure that was developed to specifically detect organic pollution (Plafkin et al., 1989). Since organic pollution is related to various stressors and widely impairs benthic macroinvertebrates, the MFBI metric relationships appropriately produced high R^2 values. The DOMPCT metric is also a tolerance measure; however, the tolerance level is non-specific to the type of stressor (Barbour et al., 1999). Since DOMPCT is non-specific to the type of stressor, the inconsistency in the DOMPCT relationships was most likely the result of the tolerance measure being unable to pinpoint specific stressors. The S-T metric is a feeding measure that investigates the coarse particulate organic matter-based shredder community (Plafkin et al., 1989). Shredders are good indicators of toxic effects and since no toxic stressors were investigated in this study, the S-T relationships showed inconsistent results. Overall, the ideal metric relationship that included a sediment indicator was the non-linear, “all average”, EPTI relationship (R^2 = 0.70; Table 4.8; Equation 4.5) whereas the ideal metric relationship that produces the highest R^2 value was the non-linear, “spring average”, MFBI relationship ($R^2 = 0.83$; Table 4.10; Equation 4.8).

One problem noted with the stepwise regression results was that the stepwise multiple regression technique is useful in determining which independent variables affect the dependent variable; however, it does not clearly differentiate differences among subsets (i.e. seasons and years). Thus, correlation analysis was used to investigate the seasonal and annual impacts on benthic macroinvertebrate assemblage conditions. In addition, principal component analysis and r-square analysis procedures were used to verify that the stepwise technique was the most appropriate statistical procedure for the determination of stressors and benthos relationships. The principal component analysis was used since it groups the independent variables together so that errors between sampling size and number of variables investigated are not of major concern. The r-square analysis is advantageous since it is not just based upon the R^2 values for one selected relationship. The r-square analysis investigates all possible relationships so that the relationship that most adequately explains the variability (R^2) and the error (MSE) and number of variables (C_p) could be selected.

4.2.2 Correlation Analysis

The next step in the development of the relationship between stressors and stream benthic macroinvertebrate assemblage conditions was to perform correlation analyses. Correlation analysis directly tests the strength of a relationship among variables, thus it was performed to determine whether or not seasonal or annual differences existed among the RBP values and/or RBP metrics (TAXA, MFBI, DOMPCT, EPTI, S-T). The multiple regression techniques did not distinguish whether or not the seasonal or annual subsets differed in terms of how the stressors impacted the benthic macroinvertebrate assemblages. Correlation analysis can be used to evaluate if seasonal or annual differences exist in the RBP values and/or metric values. Correlation analysis only looks at the variable being tested (i.e. season or year) and relates how strongly it impacts the variable in question (i.e. RBP value or RBP metrics). In addition to seasonal and annual differences, ecoregional differences were also investigated since it has been reported that benthic macroinvertebrate assemblage conditions differ among ecoregions in Virginia (Evans, 1997).

The data set used to investigate TSS and other stressors (Table B.2, Appendix B) was used to perform the correlation analyses. Since seasons, years, and ecoregions are all classified by name (i.e. fall, spring; 1997, 1998, 1999; Central Appalachian Ridges and Valleys, and Northern Piedmont), numerical classifications or dummy variables were assigned to the categories (i.e. 0,1; 0,1,2; 0,1). When testing seasonal differences, all seasonal samples for a station and the average seasonal samples for a station were investigated to correspond with the analysis subsets of “spring” versus “fall” and “spring average” versus “fall average”. Thus, one data subset was organized to include all spring and all fall samples while the other subset had averaged each station’s dependent variables for the spring samples and then for the fall samples. For both data subsets (all seasonal samples and average seasonal samples), the dummy variables strength in relating to the six dependent variables was tested. Therefore, the correlation analysis would determine if there were significant seasonal differences in the Rapid Bioassessment Protocol (RBP) values, taxa richness (TAXA), modified family biotic index (MFBI), percent dominant family (DOMPCT), EPT Index (EPTI), or the shredders to total number of organisms (S-T). The results for all seasonal samples and averaged seasonal samples are presented in Tables 4.11 and 4.12, respectively.

Table 4.11 Correlation analysis results for the investigation of seasonal influences upon benthic macroinvertebrate assemblage conditions

	RBP VALUE	TAXA	MFBI	DOMPCT	EPTI	S-T
r-value	-0.024	-0.018	-0.077	0.102	-0.146	0.156
p-value	0.811	0.855	0.434	0.298	0.136	0.113

Table 4.12 Correlation analysis results for the investigation of seasonal average influences upon benthic macroinvertebrate assemblage conditions

	RBP VALUE	TAXA	MFBI	DOMPCT	EPTI	S-T
r-value	-0.180	0.026	0.008	0.111	-0.103	0.191
p-value	0.180	0.850	0.950	0.410	0.444	0.155

The correlation results included the r-values (correlation coefficients) and p-values (significance values) for each dependent variable/seasonal relationship. For a correlation or relationship to exist, the r-value has to be closer to –1 or 1 than to zero and the p-value has to be less than 0.05.

When determining whether or not a correlation exists between the variables, the p-value is the ultimate decision factor since the p-value looks directly at how significant the relationship is.

The results presented in Tables 4.11 and 4.12 show that the highest r-value was 0.191 for the, S-T relationship in Table 4.12. A r-value of 0.191 suggests that no linear relationship exists. In addition, the S-T relationship in Table 4.11 has the smallest p-value of 0.113, much greater than the maximum 0.05 value. Thus, the results in Tables 4.11 and 4.12 clearly indicate that no difference existed between the benthic macroinvertebrate assemblage conditions for spring and fall samples.

The correlation analysis results used to determine whether or not benthic macroinvertebrate assemblage conditions differed among the 1997, 1998, and 1999 years are presented in Table 4.13.

Table 4.13 Correlation analysis results for the investigation of annual influences upon benthic macroinvertebrate assemblage conditions

	RBP VALUE	TAXA	MFBI	DOMPCT	EPTI	S-T
r-value	-0.189	0.058	-0.096	0.119	-0.033	-0.016
p-value	0.095	0.610	0.401	0.296	0.770	0.891

The annual results, with the smallest p-value being 0.095, also indicated that no difference existed among the benthic macroinvertebrate assemblage conditions for the years of 1997, 1998, and 1999. Thus, neither annual nor seasonal differences impact the benthic macroinvertebrate assemblage conditions.

For ecoregional differences, the data used in this study were collected from the Central Appalachian Ridges and Valleys ecoregion and Northern Piedmont ecoregion. The correlation analysis results of the ecoregional influences upon benthic macroinvertebrate assemblage conditions are presented in Table 4.14.

Table 4.14 Correlation analysis results for the investigation of ecoregional influences upon benthic macroinvertebrate assemblage conditions

	RBP VALUE	TAXA	MFBI	DOMPCT	EPTI	S-T
r-value	-0.524	0.416	-0.467	-0.466	0.084	0.043
p-value	<0.0001	<0.0001	<0.0001	<0.0001	0.393	0.663

The results presented in Table 4.14 indicate that the benthic macroinvertebrate assemblage conditions do differ between the two ecoregions for the RBP values, TAXA metric, MFBI metric, and DOMPCT metric as shown by p-values of less than 0.05. However, the EPTI and S-T results did not demonstrate that the benthic macroinvertebrate assemblages differed by ecoregion. To determine whether or not the ecoregions influenced the assemblage structure of the benthic macroinvertebrates, the stepwise multiple regression techniques were conducted on the Central Appalachian Ridges & Valleys and the Northern Piedmont data sets. The newly developed linear and non-linear relationships for the two ecoregion data sets were compared against one another and to the respected relationships previously developed with the “all” subset (Table C.2, Appendix C). The results (not shown here) indicated that relationships for both ecoregional data sets were similar and the R^2 values for both matched the R^2 values of the relationship for the “all” subset that did not separate out ecoregional influences. In addition, the same trends previously noted with the stepwise multiple regression techniques were evident: i.e. sediment was associated with EPTI and TAXA while the MFBI metric generally produced the highest R^2 value.

Even though the correlation analysis suggested that the two ecoregions impacted the benthic macroinvertebrate assemblage conditions differently, the multiple regression relationships between the stressor variables and the benthic macroinvertebrate assemblage conditions for the two ecoregions did not differ. Since this study focused upon developing a relationship between stressors and benthos across Virginia, the lack of differences among the ecoregion relationships eliminated the consideration of ecoregional differences any further.

In summary, the correlation analysis was beneficial in determining that the relationships between water quality stressors and stream benthic macroinvertebrate assemblage conditions are not influenced differently by seasons, years, or ecoregions. This determination is important when conducting future statistical analyses. Thus, in the principal component analysis and r-square analysis procedures there was no need to investigate the individual subsets for “spring”, “spring average”, “fall”, “fall average”, “1997”, “1998”, or “1999”. Only the entire data (“all”) and the data that had one set of values per station by averaging each station’s samples together (“all average”) were used for future statistical analyses.

4.2.3 Principal Component Analysis

Principal component analysis (PCA) procedures were used on the “all” and “all average” subsets for the data used to investigate TSS and other stressor variables (Table B.2, Appendix B). PCA was used to verify that the linear and non-linear stepwise multiple regression relationships were the most appropriate representation of how stressors affect the benthic macroinvertebrate assemblage conditions. Since the linear stepwise techniques investigated the 29 independent variables listed in Table 4.4 and the non-linear techniques investigated 105 variations (linear, cross-products, log transformations, and squared transformations) of the linear independent variables, the PCA was used in an effort to minimize the number of independent variables. The minimized number of independent variables would help with any distortions from sample size in addition to the possibility of minimizing the number of independent variables included in the stressor/benthos relationships. Even though the PCA procedure reduces the number of independent variables, data on all original variables would still need to be collected in order to produce variable groups. Accordingly, the PCA procedure effectively reduces the number of independent variables; however, it does not reduce the amount of effort or time needed for data collection.

PCA was used to group the independent variables into three new variables: Prin1, Prin2, and Prin3. Prin1 generally included habitat variables such as sediment, substrate, velocity, riffles, bank stability, bank vegetation, graze, and riparian vegetation as well as hardness, alkalinity, conductivity, and stream order. Since Prin1 consisted primarily of habitat variables, Prin1 was termed the “habitat” group. Prin2 was termed the “landuse” group since % pasture, % cropland, and % forest as well as turbidity were clustered together. The last group, Prin3, was termed the “watershed characteristics” group since it was made up of watershed area, stream length, and beef cattle numbers. For the non-linear analysis, when the linear independent variable was included into a PCA group, the log transformation and squared transformation of the variable were also included into the PCA group. The remaining variables that were not included in Prin1, Prin2, or Prin3 were individually placed in the independent variable list with Prin1, Prin2, and Prin3. Thus, the original 29 linear variables and 105 non-linear variables were reduced significantly with the PCA procedure so that the stepwise multiple regression technique could be applied on the PCA's reduced variable list. The multiple regression relationships developed with the PCA variables could then be compared against the original relationships.

Prior to completing the stepwise multiple regression analyses on PCA's new variable list, the three new variables were investigated separately. The separate investigation of Prin1, Prin2, and Prin3 was conducted to determine which variable(s) (out of the three possible) explained the most variability in the benthic macroinvertebrate assemblage conditions. The variable(s) that explained the most variability were then always included in the models and investigated along with the remaining independent variables to detect the impact of stressor variables upon benthic macroinvertebrate assemblage conditions. The stepwise multiple regression technique was run as explained in Section 4.2.1 with the p-value for entry and removal still equaling 0.05 for the independent variables. However, the p-value did not apply to Prin1, Prin2 and Prin3 since they were automatically included in the relationship. In addition to always including Prin1, Prin2, and/or Prin3 into the relationships, two other alterations were made to the multiple regression procedures used to develop relationships for the PCA's variables: 1) only the "all" and "all average" subsets were investigated following the correlation analysis results and 2) only the dependent variables noted in the RBP and two metric categories (Sections 4.2.1.2.1 and 4.2.1.2.2) in Tables 4.5 through 4.10 were investigated.

Tables 4.15 and 4.16, respectively, present the linear and non-linear stepwise results for the PCA's independent variables. The two tables also include information on the original developed stepwise relationships. Since both procedures developed relationships using stepwise techniques, the new results, 'PCA', are separated from the original results, 'ORIGINAL', with a double line as well as by the 'PCA' and 'ORIGINAL' classifications. To help clarify the abbreviations used in Tables 4.15 and 4.16, refer to Table 4.4.

The PCA stepwise results presented in Tables 4.15 and 4.16 were analyzed using the same criteria as those used for the original stepwise relationships so as to effectively compare the PCA stepwise relationships and the original stepwise relationships. The criteria included: a R^2 value close to one (100% variability), few number of variables (≤ 5) that were associated correctly with the metrics, and a sample size of thirty or more. The goal of the PCA method was to minimize the number of independent variables that were considered during the development of the stressor/benthos relationships. It was thought that the PCA relationships might also result in higher R^2 values.

The stepwise relationships developed from the PCA variables that explained a higher amount of variability than the original relationships were the linear, “all average”, EPTI and linear, “all average”, MFBI relationships. The stepwise technique used on the PCA variables of the linear, “all average”, EPTI relationship was only able to explain 0.05% more variability over the original relationship at the cost of including one additional variable. Similarly, the PCA’s linear, “all average”, MFBI relationship only explained an additional 6.6% variability over the corresponding original stepwise relationship at the cost of adding two additional independent variables. Thus, the PCA relationships were not preferred over the original relationships based on their R^2 values.

It was also important to note which relationships developed from the PCA procedures included fewer variables with minimal loss of variability for two reasons. First, the PCA stepwise relationships that include fewer variables with minimal loss of variability are beneficial in reducing overparameterization. If variables are only able to explain a small percent of variability in the RBP/metric value, the extra variability explained does not justify the inclusion of the variables. Second, since the PCA includes fewer variables and reduces overparameterization, the new PCA relationship could reduce the number of variables down to the required five or fewer.

Table 4.15 The linear, PCA, stepwise relationships versus the original, linear, stepwise relationships

		All[a]				All Avg[a]		
		RBP	TAXA	EPTI	MFBI	RBP	EPTI	MFBI
PCA Stepwise Relationships	PCA Variables Included[b,c]	-Prin1, Urban	Prin1, -DO	Prin1, -TSS, Flow, -DO, -Length, -Urban	-Prin1, Prin2, -Flow, DO, Urban	-Prin1, Urban	Prin1, -Prin2, -TSS, Flow, -DO, -Urban	-Prin1, Prin2, -Flow, DO, Urban
	R^2 (PCA)	0.45	0.33	0.41	0.60	0.57	**0.66**[d]	**0.77**[d]
ORIGINAL Stepwise Relationships[e]	R^2 (Original)	0.51	0.41	0.57	0.62	0.65	0.66	0.71
	Original Variables Included[b,c]	Hard, -Alk, -Order, Urban	-TSS, Graze, -DO, Order, -Urban	-TSS, Cover, Subs, Flow, -DO, -Temp, pH, -Wshed area, -Urban	-Vel, DO, -Order, Pasture, Urban	-Vel, -Order, Urban	Embed, Flow, -DO, -Temp, -Turb	-Flow, Hard, -Alk
		RBP	TAXA	EPTI	MFBI	RBP	EPTI	MFBI
		All[a]				All Avg[a]		

[a] One of the nine subsets originally investigated (all, all average, fall, fall average, spring, spring average, 1997, 1998, and 1999) as discussed in Section 4.2.1.

[b] Refer to Table 4.4 for a list of the abbreviations for the independent variables.

[c] A negative sign implies that the variable has a negative (indirect) association with the RBP/metric values; no sign implies that the variable has a positive (direct) association with the RBP/metric values.

[d] The bolded R^2 (PCA) values imply the PCA stepwise relationship produced a higher R^2 value than the original stepwise relationship.

[e] Selected stepwise relationships previously presented in Tables 4.5 to 4.10.

Table 4.16 The non-linear, PCA, stepwise relationships versus the original, non-linear, stepwise relationships

		All[a]			All Avg[a]		
		RBP	EPTI	MFBI	RBP	EPTI	MFBI
PCA Stepwise Relationships	PCA Variables Included[b,c]	-Prin1, Beef, Urban, DO*temp, Log(forest), Log(urban)	Prin1, -Prin2, -Urban, -Log(wshed area), Sq(flow), -Sq(DO)	-Prin1, Log(urban), Sq(DO)	-Prin1, -Prin2, Urban	Prin1, -Prin2, -Urban, Sq(flow), -Sq(do)	-Prin1, Log(urban)
	R^2 (PCA)	0.56	0.44	0.58	0.59	0.67	0.65
ORIGINAL Stepwise Relatioships[d]	R^2 (Original)	0.62	0.66	0.70	0.64	0.70	0.74
	Original Variables Included[b,c]	-Cropland, Urban, DO*temp, Beef, -Sq(vel), -Sq(order)	Flow, DO, -(TSS*ag), Log(subs), -Log(temp), Log(pH), -Log(cond), Sq(cover), -Sq(DO)	DO, Log(urban), -Sq(subs), -Sq(vel), Sq(hard), -Sq(alk), -Sq(alk*hard)	-Log(flow), Log(DO*temp), -Sq(embed)	-Log(temp), -Log(TSS*turb), Sq(flow), -Sq(DO)	-(Alk*hard), -Log(flow), Sq(hard)
		RBP	EPTI	MFBI	RBP	EPTI	MFBI
		All[a]			All Avg[a]		

[a] One of the nine subsets originally investigated (all, all average, fall, fall average, spring, spring average, 1997, 1998, and 1999) as discussed in Section 4.2.1.

[b] Refer to Table 4.4 for a list of the abbreviations for the independent variables.

[c] A negative sign implies that the variable has a negative (indirect) association with the RBP/metric values; no sign implies that the variable has a positive (direct) association with the RBP/metric values.

[d] Selected stepwise relationships previously presented in Tables 4.5 to 4.10.

The stepwise relationships developed from the PCA methods (Tables 4.15 and 4.16) that included fewer variables with minimal loss of variability were the: linear, “all”, TAXA; non-linear, “all”, MFBI; and linear, “all”, RBP relationships. On average, the elimination of a variable in the new PCA stepwise relationships, reduced the original R^2 values by 3%. Thus, the PCA method is beneficial for 3 out of 13 relationships in reducing the number of parameters with minimal reduction in the R^2 value, which effectively helps minimize overparameterization.

In summary, the PCA procedures were used together with the stepwise multiple regression techniques for 13 relationships. Upon comparing the new stepwise relationships developed from the PCA procedures (Tables 4.15 and 4.16) to the original stepwise relationships (Tables 4.5 to 4.10), the new stepwise relationships did not show much improvement. Only two of the new relationships developed from PCA procedures produced higher R^2 values than the original relationships. However, these two relationships explained an insignificant amount of additional variability when determining how many variables factored into the relationship. In addition, the PCA relationships were able to reduce the number of independent variables at the cost of losing a minimal amount of the R^2 values for only 3 of the 13 investigated relationships. Thus, the PCA procedures were ineffective in improving the original relationships’ R^2 values and the PCA procedures on occasion detected minor changes in how to minimize the inclusion of independent variables.

4.2.4 R-Square Analysis

The other statistical analysis procedure utilized, in addition to the principal component analysis procedure, to ensure the original stepwise relationships were the most appropriate representation of how stressors affect the benthic macroinvertebrate assemblage conditions was the r-square analysis procedure. The advantage of using the r-square analysis procedure is that it evaluates all possible relationships. R-square analysis, unlike the stepwise multiple regression technique, can also incorporate the error of variance and number of variables needed to ensure the relationship is not underspecified or overspecified through the mean squared error (MSE) and Mallow’s C_p values, respectively (Ott, 1993). However, the main disadvantage of the r-square analysis procedure is that the significance level for entry and/or removal of variables cannot be set. Thus, the r-square relationships may include variables that are not significant at the desired 0.05 level.

R-square analysis procedures were used on linear “all” and “all average” subsets only for the dependent variables noted in Tables 4.5, 4.7, and 4.9. The r-square analysis procedure was not run for the non-linear data since SAS was unable to process all relationships for 105 independent variables. Therefore, the r-square analysis procedure evaluated all possible linear relationships for seven evaluations (“all”, RBP; “all”, TAXA; “all”, EPTI; “all”, MFBI; “all average”, RBP; “all average”, EPTI; and “all average”, MFBI) starting with relationships based upon one dependent variable, then two, etc. The r-square relationship chosen as the best-fit model for each of the seven evaluations was based upon the combination of three factors: the relationship should have a R^2 value close to 1, a C_p value close to the number of parameters plus one, and a MSE value as small as possible or as close to zero as possible. If the R^2, C_p, and MSE values did not show enough improvement with the addition of another variable, the procedure was stopped. Table C.11, Appendix C presents the seven relationships chosen from the r-square analysis procedures as well as the results for the original stepwise relationships.

The r-square analysis relationships presented in Table C.11, Appendix C generally did not show improvement over the original stepwise relationships (Tables 4.5, 4.7 and 4.9) in terms of producing a higher R^2 value with minimal variable additions and in terms of including fewer variables with minimal amounts of variability being lost. Moreover, the r-square relationships were used to supplement the original relationships since both the r-square relationships and stepwise relationships produced similar results.

Aside from using the r-square analysis procedure to verify the stepwise relationships, the r-square analysis procedure was also used to help explain why the linear RBP relationships (Table 4.5) never included TSS into any of the relationships. The range of R^2 values for the nine possible subsets (“all”, “all average”, “fall”, “fall average”, “spring”, “spring average”, “1997”, “1998”, and “1999) when TSS alone was related to the RBP values in the r-square analysis procedure were from 0.07 to 0.21 with an average R^2 value of 0.14. Thus, TSS, on average, only accounted for 14% of the variability in the RBP values. In addition, TSS was generally the 16th significant variable out of the 29 possible independent variables listed in Table 4.4. Thus, with 14% variability being explained and 15 variables meeting requirements for inclusion prior to TSS, the chances of the RBP values relating to TSS were minimal.

In summary, the r-square analysis procedure was primarily used to ensure the original stepwise procedures had appropriately related the stressors to the benthic macroinvertebrate assemblage conditions. The r-square analysis results, showed no improvements over the stepwise results, but supported the stepwise relationships. However, in addition to testing the stepwise relationships, the r-square analysis was important in determining that TSS does not impact the RBP values to a high degree since TSS can only account for 14% of variability in the RBP values with more than half of the other independent variables being more significant than TSS.

4.3 Validation

After assessing the stressors impact upon the benthic macroinvertebrate community through multiple regression, correlation analysis, principal component analysis, and r-square analysis procedures, the accuracy of the developed relationships for use with other streams in Virginia were evaluated. The statistical analyses demonstrated that the RBP and metric stepwise multiple regression relationships were both appropriate in describing the stressor/benthos relationships; however, the seasonal and annual relationships were unnecessary. Consequently, only the most representative RBP relationship and metric relationship were selected from the two RBP stepwise relationships (Equations 4.2 and 4.3) and the six metric stepwise relationships (Equations 4.4 through 4.9) as discussed in Section 4.2.1. The RBP relationship selected was the non-linear, “all average” relationship (Equation 4.3) and the metric relationship selected was the non-linear, “all average”, EPTI relationship (Equation 4.5). The two relationships were selected since they were not a seasonal or annual subset; they were both consistent in that they were developed from the non-linear, “all average” subset; they included a sediment indicator; and the stressors were able to explain approximately two-thirds of the variability in the benthic macroinvertebrate assemblage conditions. Thus, the validation data set was used to test the accuracy of the non-linear, “all average”, RBP relationship (Equation 4.3) and non-linear, “all average”, EPTI relationship (Equation 4.5).

The validation data set was developed using data from 10 monitoring stations and included a total of 29 samples. The 10 monitoring stations (Table B.5, Appendix B) contained six agricultural watersheds, two urban watersheds, and two forested watersheds located in 8 counties

(Bedford, Montgomery, Pulaski, Giles, Botetourt, Albemarle, Orange, and Culpeper) throughout the Northern Piedmont and the Central Appalachian Ridges and Valleys. The stations that were used to develop the stressor/benthos relationships (Table B.3, Appendix B) were also located in the Northern Piedmont and the Central Appalachian Ridges and Valleys with the majority located within agricultural areas and a few additional stations from urban and forested areas. Thus, the stations within the validation data set included a good combination of the landuses and ecoregions to adequately represent all types of data used in the development of the stressor/benthos relationships.

The developed relationships that were investigated during the validation process were the non-linear, “all average” RBP and EPTI relationships. Equation 4.3 describes the RBP relationship ($R^2 = 0.64$):

$$\text{RBP} = 7.99 - 7.41\,(\text{Log(flow)}) + 1.91\,(\text{Log(DO x temp)}) - 5.63\text{E-03}\,(\text{embeddedness}^2) \quad [4.3]$$

Equation 4.5 describes the EPTI relationship ($R^2 = 0.70$):

$$\text{EPTI} = 17.55 - 6.31\,(\text{Log(temperature)}) - 2.67\,(\text{Log(TSS x turbidity)}) + 1.52\text{E-02}\,(\text{flow}^2) - 5.01\text{E-02}\,(\text{DO}^2) \quad [4.5]$$

Using the “all average” subset, each station’s samples from the fall of 1996 to the fall of 1998 were averaged. Consequently, the validation data set had a sample size of ten to correspond with the ten monitoring stations. The independent variables or stressors that were included in the two investigated relationships, as shown in Equations 4.3 and 4.5, were: temperature, TSS, turbidity, flow, dissolved oxygen, and embeddedness. Therefore, the validation data set included the 6 independent variables: temperature, TSS, turbidity, flow, dissolved oxygen, and embeddedness; the 2 observed dependent variables: Observed RBP and Observed EPTI; and, the 2 predicted dependent variables: Predicted RBP and Predicted EPTI. All 10 variables for the 10 stations are presented in Table B.6, Appendix B.

To investigate the validity of Equations 4.3 and 4.5, the respected observed and predicted dependent values were compared using a best-fit linear equation that described the observed and predicted relationship. The best-fit equation was examined for its slope, intercept, and R^2 value. In addition, the 90% confidence interval was determined for the best-fit line. For the stepwise relationships to be valid, ideally, the developed best-fit line's slope should equal one, the intercept should equal zero, the R^2 value should equal the stepwise relationship's R^2 value, and the 90% confidence bands should have minimal distance between the upper and lower intervals with the data falling in between the confidence band.

For the RBP relationship, Equation 4.3, the observed and predicted data, best-fit line, and 90% confidence interval were evaluated and graphed as shown in Figure 4.1. The best-fit linear equation between observed RBP values and predicted RBP values that was developed produced a R^2 value of 0.36. Comparing the 0.36 value with Equation 4.3's R^2 value of 0.64, the validation procedure matched more than half, 57%, of the amount of variability explained by the original relationship. The best-fit line also produced an intercept of –0.04, which is exceptionally close to the desired value of zero. The slope value of 0.59, on the other hand, was not as close to its desired value of 1. When investigating the 90% confidence interval, approximately half of the data points were within the 90% confidence interval and the distance between the upper and lower bounds was minimal. Figure 4.1 visually depicts the summarized results for the validation of Equation 4.3.

Overall, the validation results for the non-linear, "all average", RBP relationship support the original relationship depicted in Equation 4.3. The R^2 value, slope, intercept, and 90% confidence interval demonstrated a partial correlation between the observed and predicted RBP values since the results generally explained about half of the desired level. One factor that may have influenced the validation results was the number of samples. The sampling size of 10 for the validation data greatly minimized the effectiveness of the validation procedure. Ideally, a sample size of 17, or half the number of samples used for the original "all average" stepwise procedure, should have been used for validation purposes (Ott, 1993).

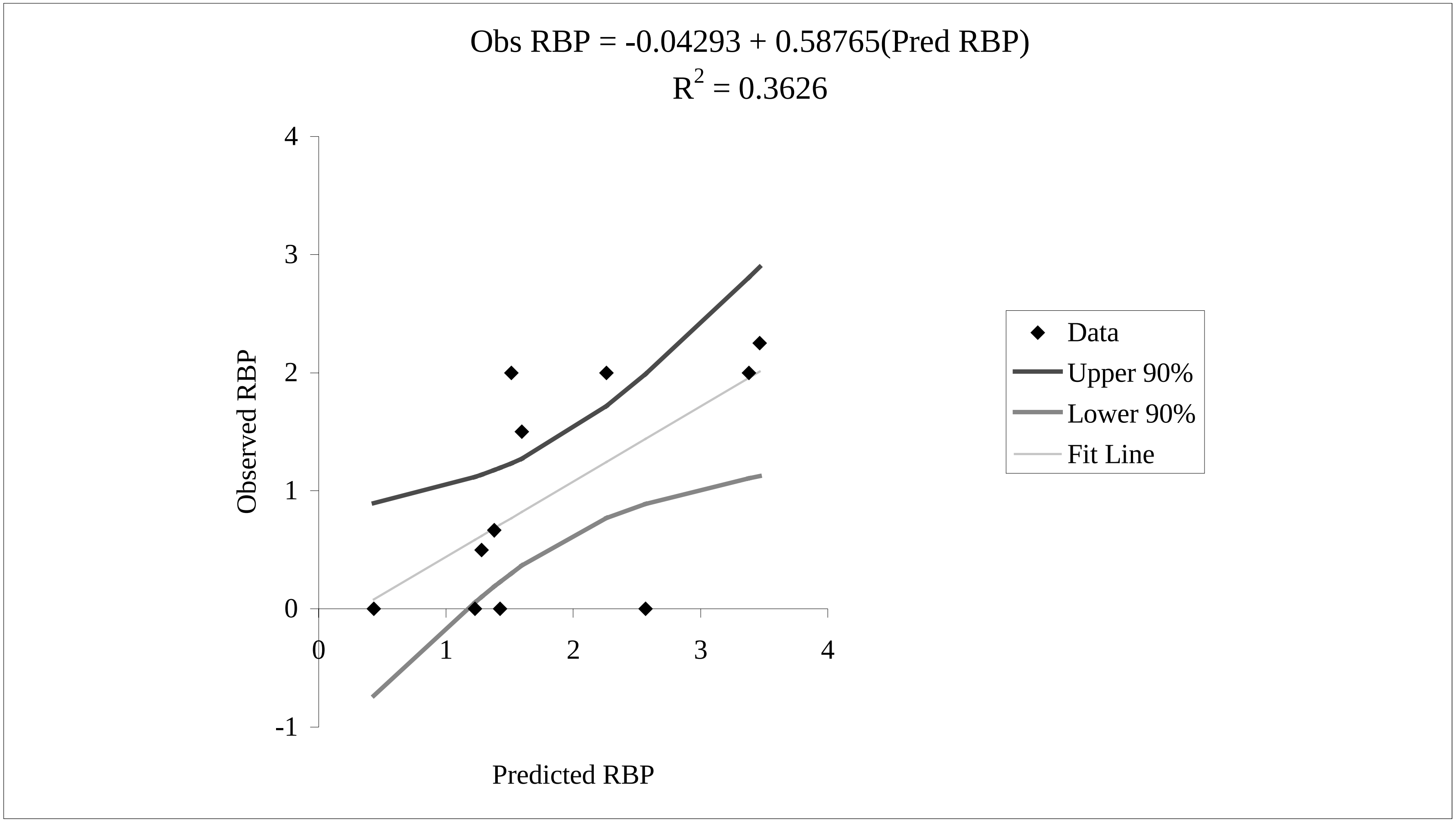

Figure 4.1 Validation results for the developed non-linear, all average, RBP relationship presented in Equation 4.3

The validity of Equation 4.5, the non-linear, “all average”, EPTI relationship, was also investigated and demonstrated similar trends as the RBP validation as shown in Figure 4.2. The best-fit linear equation that was developed produced a R^2 value of 0.29, which explains 42% of the amount of variability when compared to the original relationship’s 0.70 R^2 value. The best-fit line also had an intercept value of 3.08, which was exceptionally higher than the desired value of zero. However, the slope equaled 0.61 and was close to its desired level of 1. When investigating the 90% confidence interval, more than half of the data points fell within the 90% interval and the distance between the upper and lower bounds was minimal. Moreover, as with the RBP validation, the EPTI validation results supported the original relationship; however, the small sample size of 10 for the validation data may have impacted the accuracy of the validation results.

In summary, the validation procedures supported the developed non-linear, “all average” RBP and EPTI relationships. The validation results for both the RBP and EPTI relationships could explain approximately half the desired levels of R^2 values, slopes and/or intercepts, and confidence intervals. In addition, the validation results for the RBP relationship and the EPTI relationship were similar to one another, which helps support the idea that the methodology that goes into the development and validation processes is justified and is sufficiently accurate between the different stressor/benthos relationships. Moreover, the validation results supported the developed relationships representations of how stressors affect the benthic macroinvertebrate assemblage conditions.

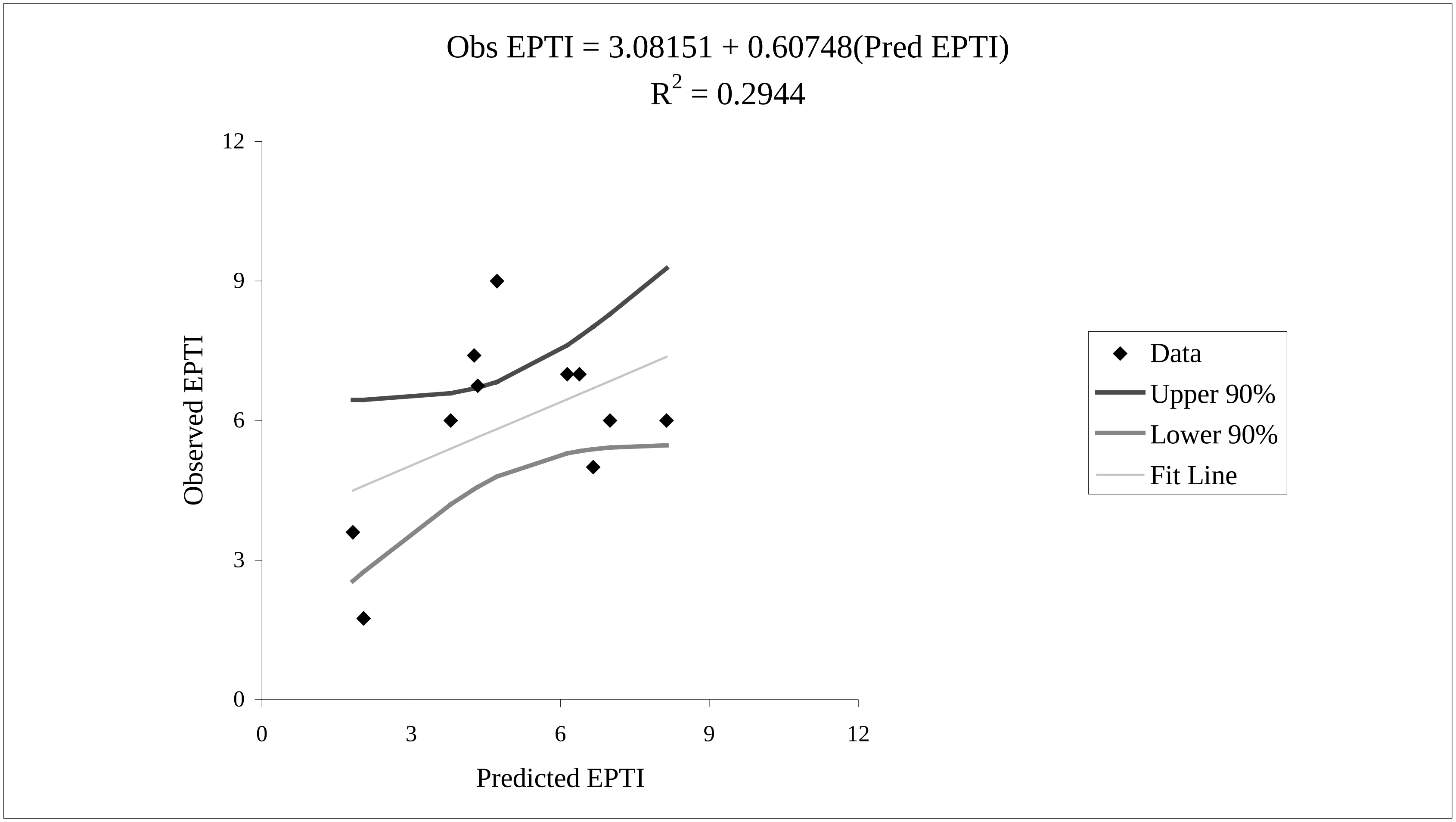

Figure 4.2 Validation results for the developed non-linear, all average, EPTI relationship presented in Equation 4.5

4.4 Applicability of the Developed Stressor/Benthos Relationships to TMDLs

After developing and validating the relationships between stream benthic macroinvertebrate assemblage conditions and their stressors, attempts were made to evaluate the applicability of the relationships to the development of benthic TMDLs. TMDLs, in general, focus upon the establishment of the maximum pollutant load a water body can withstand and still meet water quality standards. For benthic TMDLs there is a need to develop stressor threshold values below which the stressor will no longer adversely impact the benthic macroinvertebrate assemblage conditions.

The two best developed relationships, the non-linear, “all average”, RBP (Equation 4.3) and the non-linear, “all average”, EPTI (Equation 4.5), were used to evaluate the applicability of a stressor/benthos relationship with regard to the development of benthic TMDLs. The stressors or independent variables within the two selected relationships (Equations 4.3 and 4.5) were stream water temperature, flow rate, dissolved oxygen, embeddedness, turbidity, and TSS. All of these variables are measurable parameters and could be controlled to some degree. However, it is difficult to achieve optimal stream water temperature, flow rate, and dissolved oxygen levels. Shading, riparian vegetation, and diversion or re-routing of water can be implemented to achieve optimal stream water temperature, flow rate, and dissolved oxygen levels; yet, flood, drought, and climatic conditions that are beyond human control can also greatly impact the temperature, flow, and dissolved oxygen levels. Thus, it is more practical to control sediment variables such as embeddedness, turbidity, and TSS. Practices such as streambank restoration, vegetated waterways, or conservation tillage can be implemented in the watershed and designed to minimize the sediment loadings to the impaired stream segments. Accordingly, sediment variables (embeddedness, turbidity, and TSS) were investigated to determine the maximum allowable sediment load to the stream before the benthic macroinvertebrates are adversely impacted.

To determine how the developed relationships could be used to estimate a stream’s sediment loading thresholds, analyses were conducted on data from two sampling groups: a group of samples from all Virginia biomonitoring stations and a second set of data from a group of samples that only included the stations used to develop the relationships. The data from the two sampling groups were used primarily to verify that the relationships could indeed be applied to all stations regardless of whether or not the station was used to develop the stressor/benthos relationship. If the two sampling groups produced similar threshold values, then applicability of the relationships to other streams in Virginia would be verified.

For the data that included all Virginia biomonitoring stations, the entire DEQ benthic database of 418 stations (discussed in Section 3.1.1) was analyzed. To estimate the sediment thresholds, data from the non-impaired sites were used to determine the maximum sediment concentration before a water body is impaired. The threshold values were calculated using the averaged values for temperature (14.59 ^{o}C), flow (16), dissolved oxygen (9.55 mg/L), and EPTI (7) of all non-impaired stations. The average values were then used in Equations 4.3 and 4.5 to calculate either embeddedness or TSS x turbidity values for unimpaired sites. Similarly, to estimate the sediment thresholds for the data that included stations used to develop the relationships (Table B.4, Appendix B), data from the non-impaired sites were used. The non-impaired averaged values for temperature (14.83 ^{o}C), flow (18), dissolved oxygen (10.62 mg/L), and EPTI (7) were used in Equations 4.3 and 4.5 to calculate either embeddedness or TSS x turbidity values for unimpaired sites.

The last step completed for both data groups was to separate TSS and turbidity values from the TSS x turbidity parameter. The linear regression equation that describes how TSS and turbidity are related to one another based upon the stations used to develop the relationships was developed and is shown in Equation 4.10 ($R^2 = 0.82$).

$$\text{Turbidity} = 1.49 \quad + \quad 0.77\,(\text{TSS}) \qquad [4.10]$$

The embeddedness, turbidity, and TSS threshold values for the two data groups are presented in Table 4.17. The threshold values demonstrate the ‘best’ conditions or the conditions needed to ensure that the benthic macroinvertebrate assemblage conditions will not be impaired. If a station has an embeddedness value below the threshold value or if the turbidity or TSS value is above the respected turbidity and TSS thresholds, the station’s sediment levels will adversely impact the benthic macroinvertebrate assemblage conditions.

Table 4.17 Summary of the sediment threshold values for all stations in Virginia and for stations used to develop the stressor/benthos relationships

	Threshold Value (All VA Biomonitoring Stations)	Threshold Value (Stations Used to Develop the Stressor/Benthic Relationships)
Embeddedness	21.82	20.79
Turbidity (FTU)	4.01	3.95
TSS (mg/L)	3.26	3.19

The threshold values presented in Table 4.17 are all within measurable ranges except for the embeddedness values. The embeddedness values for both threshold values exceed the 0 to 20 range prescribed for determining habitat conditions. Stations that are assigned a 20 are visually assessed by a biologist and stated to exhibit optimal embeddedness characteristics of having 0-25% of the gravel, cobble, and boulders particles surrounded by fine sediment (Barbour et al., 1999). Therefore, a site that is given an embeddedness score of 20 does not necessarily exemplify the same characteristics as another site with a score of 20. There are degrees with numerical classifications and since habitat assessments are subjective based upon the evaluator's perceptions, an embeddedness value that exceeds 20, even though technically impossible, could still be valid.

When comparing the two data groups against one another, it would be harder to achieve the embeddedness threshold value for the data group that included all Virginia biomonitoring stations while turbidity and TSS levels were stricter for the data group that included stations used to develop the stressor/benthos relationships. Even though the data groups differed in terms of which one had more stringent sediment threshold values, both data groups produced similar sediment threshold values. Accordingly, the approach used in this study to develop sediment thresholds is reliable based upon the similarity between data groups. Since the stressor/benthos relationships are already formulated and can be applied by the user, this approach would be advantageous in minimizing the amounts of effort and time. A major shortcoming of this approach would be if a TMDL plan needed to be developed for a specific stressor that is not included in the developed relationships.

An additional calculation that was performed based upon the developed relationships in Equations 4.3 and 4.5 was the determination of typical values of embeddedness, turbidity, and TSS for moderately and severely impaired samples, the two benthic impairment types that are placed on Virginia's 303(d) List. The typical values were evaluated for the data group that included all Virginia biomonitoring stations to determine the current sediment levels in Virginia streams and evaluate the amount of increase or reduction needed to achieve the sediment level necessary for benthic TMDLs. The moderately impaired samples were averaged to determine the typical values of temperature (15.26 ^{o}C), flow (17), dissolved oxygen (9.68 mg/L), EPTI (5), and RBP (2). The averaged values were then used in Equations 4.3 and 4.5 to calculate the typical values of embeddedness and TSS x turbidity with Equation 4.10 separating TSS and turbidity values. Table 4.18 presents the threshold values and typical values for embeddedness, turbidity, and TSS for moderately impaired samples from all Virginia biomonitoring stations.

Table 4.18 Sediment threshold values and typical values for moderately impaired samples based upon all Virginia biomonitoring stations

	Threshold Value	Moderately Impaired Value	Amount of Change[a]
Embeddedness	21.82	10.88	10.94
Turbidity (FTU)	4.01	9.25	-5.24 (57%)[b]
TSS (mg/L)	3.26	10.06	-6.80 (68%)[b]

[a] The amount of change needed for the moderately impaired sediment value to reach its threshold value.

[b] In parentheses, the percent reduction of the moderately impaired sediment value needed to achieve its threshold value.

The results presented in Table 4.18 demonstrate that moderately impaired stations in Virginia typically have turbidity and TSS values that are more than twice the recommended level while the embeddedness rating is one half of the desired level. Moreover, the turbidity and TSS levels for a moderately impaired stream would typically need to be reduced by 57% and 68%, respectively, to achieve non-impaired benthic conditions.

The estimation of typical sediment values for severely impaired samples was evaluated in a similar manner as the moderately impaired samples. The severely impaired samples were averaged to determine the typical values for temperature (16.43 ^{o}C), flow (17), dissolved oxygen (8.88 mg/L), EPTI (1), and RBP (3). The averaged values were then used in Equations 4.3 and

4.5 to calculate the typical values of embeddedness and TSS x turbidity with Equation 4.10 separating TSS and turbidity values. Table 4.19 presents the threshold values and typical values for embeddedness, turbidity, and TSS for severely impaired samples from all Virginia biomonitoring stations.

Table 4.19 Sediment threshold values and typical values for severely impaired samples based upon all Virginia biomonitoring stations

	Threshold Value	Severely Impaired Value	Amount of Change[a]
Embeddedness	21.82	0	21.82
Turbidity (FTU)	4.01	60.79	-56.78 (93%)[b]
TSS (mg/L)	3.26	76.79	-73.53 (96%)[b]

[a] The amount of change needed for the severely impaired sediment value to reach its threshold value.

[b] In parentheses, the percent reduction of the severely impaired sediment value needed to achieve its threshold value.

The results presented in Table 4.19 demonstrate that the severely impaired station has an embeddedness value of zero, the most un-optimal condition, and it is desired to have a value over 20, optimal conditions, to achieve non-impaired benthic conditions. Similarly, the severely impaired stations in Virginia typically have excessive turbidity and TSS values that would need to be reduced by 93% and 96%, respectively, to achieve non-impaired benthic conditions.

When comparing Tables 4.18 and 4.19, the results presented in Table 4.19 for severely impaired stations showed a reduction in the embeddedness value and increases in the turbidity and TSS values over the moderately impaired station values in Table 4.18. Thus, the results support the idea that as a stream becomes benthically impaired, embeddedness values decrease while turbidity and TSS values increase. Even though, the numerical comparisons between typical and threshold values were presented in Tables 4.18 and 4.19, the degree of effort required to reach the threshold values depends upon which practice is implemented to reduce the sediment levels. Therefore, the comparison is beneficial in estimating the amount of sediment reduction needed, which aids in the determination of the most appropriate, economically feasible, and time effective best management practice.

In summary, the developed stressor/benthos relationships are beneficial in providing policymakers with a useful tool to determine stressor thresholds that will not adversely impact the benthic macroinvertebrate assemblage conditions for use in developing benthic TMDLs in Virginia. The stressor/benthos relationships can also be used to determine the impact of certain activities or stressors on the benthic macroinvertebrate assemblage conditions in a given stream.

CHAPTER 5.0 SUMMARY AND CONCLUSIONS

An investigation was conducted to determine the relationship between stream benthic macroinvertebrate assemblage conditions and their stressors. Stressors known to impact the benthic macroinvertebrates were selected and categorized into six groups: sediment, habitat, water quality, landuse, watershed characteristics, and livestock numbers (Table 3.3, Section 3.2.2). Emphasis was placed upon the sediment or sediment indicator variables as the major contributors to the decline of a stream's benthic macroinvertebrate assemblage conditions since excessive sediment loads alter water movement, food quality, oxygen availability, and interstitial spacing for the benthic macroinvertebrates because the benthos live in and are in direct contact with the stream bottom's substrate (Minshall, 1984 Lenat et al., 1981). Even though the focus of the study was on sediment, the other stressors were also investigated and considered as possible contributors to benthic macroinvertebrate assemblage conditions. The desired outcome in the development of the stressor/benthos relationships was to determine stressor levels or thresholds that would not adversely impact the benthic macroinvertebrate assemblage conditions to aid in the development of benthic TMDLs in Virginia.

The investigation began with the Water Quality Index (WQI) as a sediment indicator and progressed to focus upon total suspended solids (TSS) and other stressor variables. In both cases, data from VADEQ's Biological Assessment Program were used to represent the benthic macroinvertebrate assemblage conditions. VADEQ currently uses the Rapid Bioassessment Protocol (RBP) procedures, which compares habitat and biological conditions to reference conditions using individual metrics to arrive at a final RBP impairment classification. This study used both the individual RBP metrics and RBP impairment values in representing the benthic macroinvertebrate assemblage conditions. The RBP metrics investigated included the metrics that are most commonly used and the ones that incorporate richness, tolerance, and feeding measures. Thus, the five metrics used in the study were taxa richness (TAXA), the modified family biotic index (MFBI), the EPT index (EPTI), the percent dominant family (DOMPCT), and the shredders to the total number of organisms (S-T) (Section 3.2.3)

The WQI, a procedure developed by Shanholtz et al. (1990) that estimates the amount of agricultural sediment delivered to a stream (Section 2.8.2), was evaluated as a sediment indicator. For the WQI investigation, benthic data and landuse data were compiled for 18 stations with a total of 90 samples from the spring and fall periods from the fall of 1994 to the fall of 1998 (Table B.2, Appendix B). The 18 stations (Table B.1, Appendix B) were located within 3 agricultural counties (Rockbridge, Rockingham, and Augusta) with the stations' watersheds dominated by agricultural, urban, or forested landuses. The 10 stressor variables investigated for the WQI investigation were: WQI values, watershed area, agricultural area, % pasture, % row crop, % grass, % pasture + row crop, % cultivated, % grassland, and stream length. The goal was to evaluate the impact of these stressors on the RBP values. All 10 variables and the RBP values were compiled from data collected through the VADEQ's Biological Assessment Program or GIS data layers from VADCR. The stressors and RBP values were related to one another using multiple regression techniques.

For the TSS and other stressors investigations, existing benthic data, water quality data, GIS data, and livestock numbers were gathered and analyzed for 34 stations with a total of 105 samples from the spring and fall periods from the fall of 1996 to fall of 1998 (Table B.4, Appendix B). The 34 stations (Table B.3, Appendix B) were located in 13 counties (Rockbridge, Rockingham, Augusta, Frederick, Shenandoah, Page, Loudoun, Fairfax, Prince William, Fauquier, Culpeper, Rappahannock, and Madison) in VADEQ's Valley and Northern regions in watersheds dominated by agricultural, urban, or forested landuses. For the 34 stations (105 samples), data from VADEQ's Biological Assessment Program, VADEQ's Ambient Water Quality Monitoring Reports, GIS data from VADCR, or VADCR's Hydrologic Unit Animal Census Database were used to formulate the 29 sediment, habitat, water quality, landuse, watershed characteristics and livestock number variables (Table 4.4, Section 4.2). The stressor/benthos relationships were developed from the 29 variables related to the individual RBP metrics and RBP values using multiple regression, correlation analyses, principal component analyses, and r-square analyses.

After developing the stressor/benthos relationships, validation procedures were used to determine whether or not the developed relationships could be applied with reasonable accuracy to other streams in Virginia. For the validation procedures, data were compiled for 10 stations with 29 samples from the fall and spring sampling periods from the fall of 1996 to the fall of

1998 (Table B.6, Appendix B). The 10 stations (Table B.5, Appendix B) were located in 8 counties (Bedford, Montgomery, Pulaski, Giles, Botetourt, Albemarle, Orange, and Culpeper) in VADEQ's Valley and Northern regions in watersheds dominated by agricultural, urban, or forested landuses to correspond with the characteristics of the stations used to develop the stressor/benthos relationships. The stressor/benthos relationships were verified using the best-fit line for the predicted and observed RBP and metric in terms of the best-fit line's r-square value, slope and intercept, and the 90% confidence interval.

The applicability of the stressor/benthos relationships to the development of benthic TMDLs was also evaluated. The evaluation of the applicability of the stressor/benthos relationships was important to determine how policymakers could use the relationships in developing benthic TMDLs. The benthic TMDLs were assessed using sediment threshold values below which sediment would no longer adversely impact the benthic macroinvertebrate assemblage conditions. Typical sediment values for moderately and impaired streams throughout Virginia, or streams placed on Virginia's 303(d) List, were compared to the threshold values to determine the amount of sediment reduction needed to ensure the benthic macroinvertebrate conditions are not impaired.

The investigation of the relationships between the benthic macroinvertebrate assemblage conditions and possible stressors resulted in the following conclusions:

- The WQI parameter could not be used as a reliable sediment indicator since the WQI was not included in any of the stressor/benthos relationships developed in this study. In addition, since the WQI procedure was developed to estimate sediment loadings from agricultural areas, different procedures were needed for urban and forested areas; however, the introduction of other procedures could reduce the overall accuracy of the results due to inconsistency among the procedures.

- TSS values proved to be a better sediment indicator, as compared to WQI; for, TSS values are consistent for all landuse types and TSS values are easily measured in the field. In addition, TSS values were included in several of the developed stressor/benthos relationships.

- The stressor/benthos relationships were formulated using multiple regression procedure. The stepwise procedure presented the best descriptions of the stressor/benthos relationships in terms of the type and number of the variables included in a relationship and the resulting R^2 value. Seasonal, annual, and ecoregional effects did not significantly impact the stream benthic macroinvertebrate assemblage conditions.

- The investigation of WQI and TSS values proved that sediment alone cannot adequately describe the stream benthic conditions. The additional habitat, water quality, landuse, and livestock number stressors are important in describing the benthic macroinvertebrate assemblage conditions. Specifically, dissolved oxygen, flow, % urban land, TSS, temperature, and stream velocity greatly influence the benthic macroinvertebrates.

- When describing the benthic macroinvertebrate assemblage conditions, the RBP values and RBP metrics are equally important. The metrics increase the stressor/benthos relationship's accuracy since the metrics represent attributes of a targeted aquatic assemblage while the RBP values are based upon the aggregation of the metrics and habitat conditions against reference conditions. Even though the metrics provide the best representation of the benthic conditions, the RBP values are still important since Virginia currently uses the final RBP value for assessment of stream water quality conditions and they are need for TMDL development.

- The relationships that related the stressors to the RBP values never included TSS or turbidity as the sediment indicators. The closest sediment variable was embeddedness. The RBP relationships generally included % urban, velocity, embeddedness, hardness, stream order, pH, dissolved oxygen, and temperature. The stressors typically explained 65% of the variability in the RBP values.

- The best RBP/stressor relationship produced the following relationship ($R^2=0.64$):

  $$\text{RBP} = 7.99 - 7.41\,(\text{Log(flow)}) + 1.91\,(\text{Log(DO x temp)}) - 5.63\text{E-03}\,(\text{embeddedness}^2)$$

  with the squared embeddedness term explaining 77% of the total variability in the relationship.

- The relationships that related the stressors to the individual metrics (TAXA, MFBI, EPTI, DOMPCT, and S-T) did include the TSS and turbidity sediment indicators in some relationships. The metric relationships that included a sediment indicator (TSS or turbidity) typically included dissolved oxygen, flow or some other habitat variable such as substrate or riffles. The stressors typically described 70% of the variability in the metric values. In addition, sediment was generally included with the EPTI and TAXA metric relationships. Since the two metrics are richness measures that describe niche space, habitat, and food sources, sediment, which greatly impacts niche space, habitat, and food sources, was appropriately linked to the EPTI and TAXA metrics.

- The DOMPCT and S-T relationships were inconsistent and unsuccessful in explaining high degrees of variability. The DOMPCT metric is a tolerance measure that is non-specific to the type of stressor. Consequently, the inconsistency in the DOMPCT relationship was most likely the result of the tolerance measure being unable to pinpoint specific stressors. The S-T metric is a feeding measure that investigates the coarse particulate organic matter-based shredder community. Shredders are good indicators of landuses and toxic effects and since no toxic stressors were investigated in this study, the S-T relationships showed inconsistent results.

- The most ideal metric/stressor relationship produced the following relationship ($R^2=0.70$):

  $$\begin{aligned}\text{EPTI} = 17.55 &- 6.31\,(\text{Log(temperature)}) - 2.67\,(\text{Log(TSS x turbidity)}) + 1.52\text{E-02}\,(\text{flow}^2)\\ &- 5.01\text{E-02}\,(\text{DO}^2)\end{aligned}$$

  with the Log(TSS x turbidity) term explaining 53% of the total variability in the relationship.

- Validation of the two most representative stressor/benthos relationships proved that the relationships could be applicable to other streams in Virginia.

- When determining the amount of reduction needed for embeddedness, turbidity, and TSS levels to meet threshold values for moderately impaired sites across Virginia, the amount of embeddedness needs to be reduced by 11, turbidity levels reduced by 5 FTU (57%), and TSS values reduced by 7 mg/L (68%). Similarly, the severely impaired sites, need to reduce the amount of embeddedness by 22, reduce turbidity levels by 57 FTU (93%), and reduce TSS values by 74 mg/L (96%) to meet their required sediment threshold values.

The evaluation of the stressor/benthos relationships developed in this study demonstrated that they could be beneficial in providing policymakers with a useful tool to determine stressor thresholds that will not adversely impact the benthic macroinvertebrate assemblage conditions for use in developing benthic TMDLs in Virginia. The threshold values for stations throughout Virginia of 21.82 for embeddedness, 4.01 FTU for turbidity, and 3.26 mg/L for TSS corresponded well with the threshold values for stations used to develop the stressor/benthic relationships of 20.79 for embeddedness, 3.95 FTU for turbidity, and 3.19 mg/L for TSS.

CHAPTER 6.0 RECOMMENDATIONS FOR FURTHER STUDY

The goal of this study was to establish relationships between benthic macroinvertebrate assemblages and sediment or other stressors in Virginia. The specific objectives were to: 1) develop relationships between RBP indices and/or RBP metrics and stressors such as sediment loads, landuse, habitat, etc. while placing emphasis upon sediment as the primary stressor, 2) evaluate the accuracy of the stressor/benthos relationships using data from selected watersheds in Virginia, and 3) discuss the implications of the study with regard to the development of benthic TMDLs.

When relating the benthic macroinvertebrate assemblage conditions to their stressors, the main problem was how to describe the benthic macroinvertebrate assemblage conditions. Since Virginia is currently reviewing its biological monitoring program, the benthic macroinvertebrate conditions needed to be described in a manner that would be applicable to future uses. Similarly, when focusing upon sediment, the main problem in relating sediment to benthic conditions was finding an appropriate sediment indicator. To better understand the impact of sediment on the benthic community, an appropriate model for estimating sediment loadings from all landuse types is needed. In addition, even though biological monitoring does provide useful information on the ability of a stream to support aquatic life, the use of biological data together with water quality data to develop stressor/benthos relationships poses a problem since the benthic and water quality data have different sampling periods (Younos et al., 2000). Thus, the lack of a specific benthic standard, an appropriate sediment indicator, and sampling coordination between biological and water quality monitoring programs are the three main problems affecting the development of stressor/benthos relationships.

It is recommended that future studies on stressor/benthos relationships include the following:

1. Biologist, policymakers, and stakeholders who try to ascertain stressor/benthos relationships need to take note of any changes in the biological monitoring program. Virginia currently uses and this study used the RBP II procedures to define benthic macroinvertebrate assemblage conditions. However, Virginia’s Biological Assessment Program is under review with no set benthic standard yet established. Even though some

processes within the RBP method, specifically the metrics, will still be utilized, a push for changing how the final assessment is categorized is in effect.

2. Efforts to model sediment loadings on a watershed basis should be pursued. This would include a model that is appropriate for diverse landuses and multiple watershed characteristics (location, size, etc.).

3. The selected sediment indicator variable should be associated with deposited sediment instead of suspended sediment in the stream. The benthic macroinvertebrates inhabit the stream bottom and are directly impacted by the deposited materials more so than by the suspended materials.

4. Efforts should be made to match the sampling periods of the water quality data and the biological data with one another. A detailed water quality monitoring program that includes TSS and nutrient concentration as well as flow data is not generally completed at the time of biological samplings. Such coordination will minimize errors involved in the development of stressor/benthos relationships in the future.

5. Future studies should try to incorporate as many sampling sites as possible while investigating all possible stressor variables. The investigation of habitat, water quality, landuse, watershed characteristics, and livestock numbers in addition to sediment stressor variables could enhance the accuracy of the relationships.

APPENDIX A Summary of the “usable” stations

Table A.1 Summary of the stations used to develop and validate the stressor/benthic relationships

STATION	RIVERMILE	COUNTY	ECOREGION	LANDUSE	SURVEY REASON / STATION DESCRIPTION
Back Creek	BCK000.78	Augusta	Central Appalachian Ridges and Valleys	Silvaculture	Monitor agricultural NPS.
Beaver Creek (BRC)	BRC001.88	Orange Co.	Northern Piedmont	Pasture/Hay	Survey and assess agricultural NPS in the upper York River Basin
Beaver Creek (BVR)	BVR003.60	Rockingham	Central Appalachian Ridges and Valleys	Silvaculture	Monitor agricultural NPS.
Buffalo Creek	BLD000.22	Rockbridge	Central Appalachian Ridges and Valleys	Pasture/Hay	This is the major tributary of the Maury River that drains the Karst portion of Rockbridge County.
Bull Run	BUL010.28	Prince William	Northern Piedmont	Suburban	This station is used to monitor impacts from urban NPS, stormwater discharges and construction activities, and agriculture.
Catoctin Creek	CAX004.57	Loudoun	Northern Piedmont	Forest	This station is used as one of the reference sites.
Cedar Run	CER016.46	Fauquier Co.	Northern Piedmont	Pasture/Hay	Expand coverage in this agricultural watershed
Cooks Creek	CKS003.04	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	Track effects of NPS abatement efforts within the watershed.
Crab Creek (04)	CBC004.38	Montgomery	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor agricultural NPS.
Crab Creek (06)	CBC006.35	Montgomery	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor agricultural NPS.
Cub Run	CUB000.40	Page	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor discharge.
Difficult Run	DIF000.86	Fairfax	Northern Piedmont	Urban	This station monitors impacts in an urban watershed. Storm sewer discharges and other nonpoint sources continue to impact this urban watershed.

Table A.1 (continued) Summary of the stations used to develop and validate the stressor/benthic relationships

STATION	RIVERMILE	COUNTY	ECOREGION	LANDUSE	SURVEY REASON / STATION DESCRIPTION
Goose Creek	GOO044.36	Fauquier	Northern Piedmont	Cropland	This station monitors the upper Goose Creek watershed.
Great Run	GRT001.70	Fauquier	Northern Piedmont	Pasture/Hay	This station assesses ambient quality in an agricultural segment.
Hays Creek	HYS001.41	Rockbridge	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor agricultural NPS.
Hazel River	HAZ032.54	Rappahannock	Northern Piedmont	Cropland	This station assesses impacts from agricultural nonpoint sources.
Hogue Creek	HOC006.23	Frederick	Central Appalachian Ridges and Valleys	Silvaculture	Monitor agricultural NPS.
Holmans Creek	HMN002.09	Shenandoah	Central Appalachian Ridges and Valleys	Orchard	Monitor effects of agricultural NPS, and impoundment.
Hughes River	HUE000.20	Rappahannock	Northern Piedmont	Cropland	This station assesses impacts from agricultural nonpoint sources.
James River	JMS326.30	Botetourt	Central Appalachian Ridges and Valleys	Pasture/Hay	To assess the water body.
Kerrs Creek	KRR001.54	Rockbridge	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor the effects of agricultural NPSs.
Linville Creek	LNV000.71	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor the effects of agricultural NPSs; TMDL station. Station is highly affected by channelization and siltation.
Little Otter River	LOR014.75	Bedford	Central Appalachian Ridges and Valleys	Pasture/Hay	Upstream control for Bedford STP.

Table A.1 (continued) Summary of the stations used to develop and validate the stressor/benthic relationships

STATION	RIVERMILE	COUNTY	ECOREGION	LANDUSE	SURVEY REASON / STATION DESCRIPTION
Little River	LIV004.78	Loudoun Co.	Northern Piedmont	Suburban	Expand coverage in the suburban Potomac River watershed. Station impacted by rural development and an upstream impoundment.
Mechums River	MCM018.92	Albemarle	Northern Piedmont	Pasture/Hay	Monitor agricultural NPS.
Mill Creek (C)	MIC001.00	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	TMDL station. Station in middle of heavily used cattle pasture.
Mill Creek (L)	MIL002.20	Shenandoah	Central Appalachian Ridges and Valleys	Pasture/Hay	Tributary to North Fork Shenandoah River.
Moffett Creek	MFT006.24	Augusta	Central Appalachian Ridges and Valleys	Pasture/Hay	Coincident with MAHA station. Station in middle of heavily used cattle pasture.
Mountain Run	MTN000.59	Culpeper	Northern Piedmont	Cropland	This station monitors agricultural and rural nonpoint sources.
Muddy Creek	MUD005.81	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	Station coincident with TMDL and NAWQA studies. Station affected by agricultural NPS and cattle in stream.
Peak Creek (09)	PKC009.29	Pulaski	Central Appalachian Ridges and Valleys	Urban	Downstream of Magnox Inc. and upstream of Downtown East Property.
Peak Creek (11)	PKC011.11	Pulaski	Central Appalachian Ridges and Valleys	Suburban	Upstream control
Pleasant Run	PLE000.08	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	Track agricultural NPS and effects of pollution events.
Popes Head Creek	POE002.00	Fairfax Co.	Northern Piedmont	Urban	Monitoring of urban development and agricultural NPS impacts.

Table A.1 (continued) Summary of the stations used to develop and validate the stressor/benthic relationships

STATION	RIVERMILE	COUNTY	ECOREGION	LANDUSE	SURVEY REASON / STATION DESCRIPTION
Rapidan River	RAP006.53	Culpeper	Northern Piedmont	Forest	This station is used as one of the reference sites.
Rappahannock River (147)	RPP147.10	Culpeper/Fauquier	Northern Piedmont	Cropland	Expand basin coverage and site control above Remington STP
Rappahannock River (175)	RPP175.51	Fauquier	Northern Piedmont	Cropland	To assess impacts in the upper watershed from agricultural NPS.
Robinson River	ROB001.90	Culpeper/Madison	Northern Piedmont	Pasture/Hay	This station is used as one of the reference sites.
South River (00)	STH000.21	Rockbridge	Central Appalachian Ridges and Valleys	Pasture/Hay	Trend station near confluence with the Maury River. Station well buffered by limestone springs; yet, the head waters of this stream may suffer from episodic flushes of acidic run off.
South River (27)	STH027.08	Augusta	Central Appalachian Ridges and Valleys	Suburban	Control upstream of Waynesboro dischargers and urban NPSs.
South Run	SOT001.44	Fauquier	Northern Piedmont	Forest	EPA Core Station. Monitoring Vint Hill STP. Upstream discharge from Lake Brittle.
Stony Creek	SNC000.20	Giles	Central Appalachian Ridges and Valleys	Forest	Downstream of Eastern Ridge Lime.
Toms Brook	TMB000.54	Shenandoah	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor effect of municipal discharge.
Turley Creek	TRL000.02	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	VDCR Watershed Project. Station impacted by agricultural landuse.

APPENDIX B Station summary and the data for each data set investigated

Table B.1 Summary of the stations used to investigate WQI as a sediment indicator

STATION	RIVERMILE	COUNTY	LANDUSE	SURVEY REASON / STATION DESCRIPTION
Back Creek	BCK000.78	Augusta	Silvaculture	Monitor agricultural NPS.
Blacks Run	BLK005.62	Rockingham	Urban	Monitor urban NPSs.
Buffalo Creek	BLD000.22	Rockbridge	Pasture/Hay	This is the major tributary of the Maury River that drains the Karst portion of Rockbridge County.
Christians Creek	CST007.42	Augusta	Pasture/Hay	This station is near a long term AWQM station. Fishersville STP discharges upstream of here.
Cooks Creek	CKS003.04	Rockingham	Pasture/Hay	Track effects of NPS abatement efforts within the watershed.
Dry River	DUR000.11	Rockingham	Pasture/Hay	Monitor effects of agricultural NPSs, Wampler Long Acre, Inc.; TMDL station. Station receives mountain drainage during high flow events.
Kerrs Creek	KRR001.54	Rockbridge	Pasture/Hay	Monitor the effects of Agricultural NPSs.
Lewis Creek	LEW006.95	Augusta	Silvaculture	This urban stream is impacted by NPS flushes from city streets.
Linville	LNV000.71	Rockingham	Pasture/Hay	Monitor the effects of agricultural NPSs; TMDL station. Station is highly affected by channelization and siltation.
Middle River	MDL001.85	Augusta	Cropland	Trend station near confluence with North river.
Moffett Creek	MFT006.24	Augusta	Pasture/Hay	Coincident with MAHA station. Station in middle of heavily used cattle pasture.
Mossy Creek	MSS003.01	Augusta	Silvaculture	Former reference station, very large first order stream, spring fed.
Muddy Creek (lf)	MUD002.10	Rockingham	Pasture/Hay	Survey impact of increased flow from unknown source.
Muddy Creek (rt)	MUD005.81	Rockingham	Pasture/Hay	Station coincident with TMDL and NAWQA studies. Station affected by agricultural NPS and cattle in stream.
Pleasant Run	PLE000.08	Rockingham	Pasture/Hay	Track Agricultural NPS and effects of pollution events.
South River (21)	STH021.72	Augusta	Urban	Monitor the effects of discharges and urban NPSs.
South River (27)	STH027.08	Augusta	Suburban	Control upstream of Waynesboro dischargers and urban NPSs.
Turley Creek	TRL000.02	Rockingham	Pasture/Hay	VDCR Watershed Project. Station impacted by agricultural landuse.

Table B.2 Data collected for the data set used to investigate WQI as a sediment indicator

STATION	RIVERMILE	DATE	RBP	WQI (t/ac/yr)	Wshed Area (ac)	Ag Area (ac)	% Pasture	% Row Crop	% Grass	% Pasture + Row Crop	% Cultivated	% Grassland	Stream Length (m)
Back Creek	BCK000.78	10/19/95	2	33.13	26644.8	2141.5	2.62	5.61	0	8.23	8.23	2.62	22733
Back Creek	BCK000.80	6/6/96	2	33.13	26644.8	2141.5	2.62	5.61	0	8.23	8.23	2.62	22733
Back Creek	BCK000.79	11/5/96	2	33.13	26644.8	2141.5	2.62	5.61	0	8.23	8.23	2.62	22733
Back Creek	BCK000.80	5/5/97	2	33.13	26644.8	2141.5	2.62	5.61	0	8.23	8.23	2.62	22733
Back Creek	BCK000.80	9/18/97	2	33.13	26644.8	2141.5	2.62	5.61	0	8.23	8.23	2.62	22733
Blacks Run	BLK005.62	10/3/94	2	28.03	7496	1922.7	15.03	12.52	1.68	27.54	29.22	16.7	9441
Blacks Run	BLK005.62	5/16/95	2	28.03	7496	1922.7	15.03	12.52	1.68	27.54	29.22	16.7	9441
Blacks Run	BLK005.62	9/27/95	2	28.03	7496	1922.7	15.03	12.52	1.68	27.54	29.22	16.7	9441
Blacks Run	BLK005.62	5/23/96	2	28.03	7496	1922.7	15.03	12.52	1.68	27.54	29.22	16.7	9441
Blacks Run	BLK005.62	10/3/96	2	28.03	7496	1922.7	15.03	12.52	1.68	27.54	29.22	16.7	9441
Blacks Run	BLK005.62	4/30/97	2	28.03	7496	1922.7	15.03	12.52	1.68	27.54	29.22	16.7	9441
Blacks Run	BLK005.62	9/17/97	2	28.03	7496	1922.7	15.03	12.52	1.68	27.54	29.22	16.7	9441
Buffalo Creek	BLD000.22	10/4/94	2	99.28	77509.9	21868.7	9.89	14.86	0	24.75	24.75	9.89	32869
Buffalo Creek	BLD000.22	5/25/95	2	99.28	77509.9	21868.7	9.89	14.86	0	24.75	24.75	9.89	32869
Buffalo Creek	BLD000.23	10/2/97	2	99.28	77509.9	21868.7	9.89	14.86	0	24.75	24.75	9.89	32869
Buffalo Creek	BLD000.24	10/15/98	1	99.28	77509.9	21868.7	9.89	14.86	0	24.75	24.75	9.89	32869
Christians Creek	CST007.42	10/20/94	2	57.55	52936.8	33040.2	27.94	35.2	0	63.14	63.14	27.94	38404
Christians Creek	CST007.44	5/1/95	2	57.55	52936.8	33040.2	27.94	35.2	0	63.14	63.14	27.94	38404
Christians Creek	CST007.43	10/11/95	1	57.55	52936.8	33040.2	27.94	35.2	0	63.14	63.14	27.94	38404
Christians Creek	CST007.44	10/6/98	1	57.55	52936.8	33040.2	27.94	35.2	0	63.14	63.14	27.94	38404
Cooks Creek	CKS003.04	9/26/94	3	94.7	14500.4	11330.7	52.43	24.54	0	76.97	76.97	52.43	17127
Cooks Creek	CKS003.08	5/16/95	3	94.7	14500.4	11330.7	52.43	24.54	0	76.97	76.97	52.43	17127
Cooks Creek	CKS003.05	9/27/95	3	94.7	14500.4	11330.7	52.43	24.54	0	76.97	76.97	52.43	17127
Cooks Creek	CKS003.08	6/3/96	2	94.7	14500.4	11330.7	52.43	24.54	0	76.97	76.97	52.43	17127
Cooks Creek	CKS003.06	10/3/96	2	94.7	14500.4	11330.7	52.43	24.54	0	76.97	76.97	52.43	17127
Cooks Creek	CKS003.08	5/22/97	3	94.7	14500.4	11330.7	52.43	24.54	0	76.97	76.97	52.43	17127
Cooks Creek	CKS003.07	9/17/97	2	94.7	14500.4	11330.7	52.43	24.54	0	76.97	76.97	52.43	17127
Cooks Creek	CKS003.08	10/6/98	2	94.7	14500.4	11330.7	52.43	24.54	0	76.97	76.97	52.43	17127
Dry River	DUR000.11	5/30/95	2	32.25	10111.4	6177.4	41.71	18.3	0	60.01	60.01	41.71	16114
Dry River	DUR000.11	9/28/95	1	32.25	10111.4	6177.4	41.71	18.3	0	60.01	60.01	41.71	16114

Table B.2 (continued) Data collected for the data set used to investigate WQI as a sediment indicator

STATION	RIVERMILE	DATE	RBP	WQI (t/ac/yr)	Wshed Area (ac)	Ag Area (ac)	% Pasture	% Row Crop	% Grass	% Pasture + Row Crop	% Cultivated	% Grassland	Stream Length (m)
Kerrs Creek	KRR001.54	5/25/95	2	110.75	22178.2	5893.5	10.19	13.03	0	23.23	23.23	10.19	14717
Kerrs Creek	KRR001.54	10/2/97	2	110.75	22178.2	5893.5	10.19	13.03	0	23.23	23.23	10.19	14717
Kerrs Creek	KRR001.55	10/15/98	1	110.75	22178.2	5893.5	10.19	13.03	0	23.23	23.23	10.19	14717
Lewis Creek	LEW006.95	10/20/94	2	54.24	9656.2	4099	23.03	21.66	0	44.69	44.69	23.03	9798
Lewis Creek	LEW006.97	5/16/95	2	54.24	9656.2	4099	23.03	21.66	0	44.69	44.69	23.03	9798
Lewis Creek	LEW006.96	10/10/95	3	54.24	9656.2	4099	23.03	21.66	0	44.69	44.69	23.03	9798
Lewis Creek	LEW006.97	6/3/96	2	54.24	9656.2	4099	23.03	21.66	0	44.69	44.69	23.03	9798
Lewis Creek	LEW006.97	5/5/97	3	54.24	9656.2	4099	23.03	21.66	0	44.69	44.69	23.03	9798
Lewis Creek	LEW006.97	9/18/97	2	54.24	9656.2	4099	23.03	21.66	0	44.69	44.69	23.03	9798
Linville	LNV000.71	10/3/94	2	62.95	29858.1	22764.3	43.55	30.81	0	73.83	73.83	43.55	21389
Linville	LNV000.71	5/9/95	2	62.95	29858.1	22764.3	43.55	30.81	0	73.83	73.83	43.55	21389
Linville	LNV000.72	9/28/95	1	62.95	29858.1	22764.3	43.55	30.81	0	73.83	73.83	43.55	21389
Linville	LNV000.72	5/21/96	2	62.95	29858.1	22764.3	43.55	30.81	0	73.83	73.83	43.55	21389
Linville	LNV000.73	9/22/97	2	62.95	29858.1	22764.3	43.55	30.81	0	73.83	73.83	43.55	21389
Linville	LNV000.74	10/23/98	2	62.95	29858.1	22764.3	43.55	30.81	0	73.83	73.83	43.55	21389
Middle River	MDL001.85	10/25/94	2	54.13	22102.1	17497.3	49.78	28.56	0	78.34	78.34	49.78	27722
Middle River	MDL001.86	5/1/95	2	54.13	22102.1	17497.3	49.78	28.56	0	78.34	78.34	49.78	27722
Middle River	MDL001.86	10/3/95	1	54.13	22102.1	17497.3	49.78	28.56	0	78.34	78.34	49.78	27722
Moffett Creek	MFT006.24	10/20/94	2	57.02	6123.3	1793.4	17.63	9.59	0	27.23	27.23	17.63	11015
Moffett Creek	MFT006.24	5/10/95	2	57.02	6123.3	1793.4	17.63	9.59	0	27.23	27.23	17.63	11015
Moffett Creek	MFT006.24	10/10/95	2	57.02	6123.3	1793.4	17.63	9.59	0	27.23	27.23	17.63	11015
Moffett Creek	MFT006.25	5/8/97	2	57.02	6123.3	1793.4	17.63	9.59	0	27.23	27.23	17.63	11015
Moffett Creek	MFT006.24	10/14/97	2	57.02	6123.3	1793.4	17.63	9.59	0	27.23	27.23	17.63	11015
Moffett Creek	MFT006.24	10/6/98	2	57.02	6123.3	1793.4	17.63	9.59	0	27.23	27.23	17.63	11015
Mossy Creek	MSS003.01	5/16/95	2	62.18	8209	6468.9	43.29	31.53	0	74.82	74.82	43.29	12886
Mossy Creek	MSS003.01	10/10/95	1	62.18	8209	6468.9	43.29	31.53	0	74.82	74.82	43.29	12886
Mossy Creek	MSS003.02	10/24/96	2	62.18	8209	6468.9	43.29	31.53	0	74.82	74.82	43.29	12886
Mossy Creek	MSS003.02	5/8/97	2	62.18	8209	6468.9	43.29	31.53	0	74.82	74.82	43.29	12886
Mossy Creek	MSS003.03	10/14/97	2	62.18	8209	6468.9	43.29	31.53	0	74.82	74.82	43.29	12886
Mossy Creek	MSS003.04	10/14/98	1	62.18	8209	6468.9	43.29	31.53	0	74.82	74.82	43.29	12886

Table B.2 (continued) Data collected for the data set used to investigate WQI as a sediment indicator

STATION	RIVERMILE	DATE	RBP	WQI (t/ac/yr)	Wshed Area (ac)	Ag Area (ac)	% Pasture	% Row Crop	% Grass	% Pasture + Row Crop	% Cultivated	% Grassland	Stream Length
Muddy Creek (lf)	MUD002.10	10/22/96	2	73.15	10863.4	5604.7	37.29	12.59	0	49.88	49.88	37.29	13848
Muddy Creek (lf)	MUD002.10	4/30/97	2	73.15	10863.4	5604.7	37.29	12.59	0	49.88	49.88	37.29	13848
Muddy Creek (lf)	MUD002.11	10/1/97	2	73.15	10863.4	5604.7	37.29	12.59	0	49.88	49.88	37.29	13848
Muddy Creek (rt)	MUD005.81	10/3/94	2	60.81	9238.1	6717.4	47.38	24.48	0	71.86	71.86	47.38	10351
Muddy Creek (rt)	MUD005.84	5/16/95	2	60.81	9238.1	6717.4	47.38	24.48	0	71.86	71.86	47.38	10351
Muddy Creek (rt)	MUD005.82	10/30/95	3	60.81	9238.1	6717.4	47.38	24.48	0	71.86	71.86	47.38	10351
Muddy Creek (rt)	MUD005.84	5/23/96	2	60.81	9238.1	6717.4	47.38	24.48	0	71.86	71.86	47.38	10351
Muddy Creek (rt)	MUD005.83	10/22/96	2	60.81	9238.1	6717.4	47.38	24.48	0	71.86	71.86	47.38	10351
Muddy Creek (rt)	MUD005.84	4/30/97	2	60.81	9238.1	6717.4	47.38	24.48	0	71.86	71.86	47.38	10351
Muddy Creek (rt)	MUD005.84	9/23/97	2	60.81	9238.1	6717.4	47.38	24.48	0	71.86	71.86	47.38	10351
Pleasant Run	PLE000.08	10/26/94	3	45.16	5343	4153.1	45.69	30.81	0	76.5	76.5	45.69	10265
Pleasant Run	PLE000.08	5/26/95	3	45.16	5343	4153.1	45.69	30.81	0	76.5	76.5	45.69	10265
Pleasant Run	PLE000.08	9/27/95	3	45.16	5343	4153.1	45.69	30.81	0	76.5	76.5	45.69	10265
Pleasant Run	PLE000.08	6/3/96	3	45.16	5343	4153.1	45.69	30.81	0	76.5	76.6	45.69	10265
Pleasant Run	PLE000.08	10/24/96	3	45.16	5343	4153.1	45.69	30.81	0	76.5	76.5	45.69	10265
Pleasant Run	PLE000.08	4/30/97	3	45.16	5343	4153.1	45.69	30.81	0	76.5	76.5	45.69	10265
Pleasant Run	PLE000.08	9/17/97	2	45.16	5343	4153.1	45.69	30.81	0	76.5	76.5	45.69	10265
South River (21)	STH021.72	10/25/94	2	30.11	17795	4067.5	9.3	15.52	0.8	24.82	25.62	10.1	14962
South River (21)	STH021.72	5/4/95	2	30.11	17795	4067.5	9.3	15.52	0.8	24.82	25.62	10.1	14962
South River (21)	STH021.73	10/19/95	2	30.11	17795	4067.5	9.3	15.52	0.8	24.82	25.62	10.1	14962
South River (21)	STH021.74	10/2/98	1	30.11	17795	4067.5	9.3	15.52	0.8	24.82	25.62	10.1	14962
South River (27)	STH027.08	10/25/94	2	45.04	2250.8	465.6	4.98	17.3	0	22.29	22.29	4.98	5377
South River (27)	STH027.08	5/4/95	0	45.04	2250.8	465.6	4.98	17.3	0	22.29	22.29	4.98	5377
South River (27)	STH027.08	10/19/95	2	45.04	2250.8	465.6	4.98	17.3	0	22.29	22.29	4.98	5377
South River (27)	STH027.08	10/2/98	1	45.04	2250.8	465.6	4.98	17.3	0	22.29	22.29	4.98	5377
Turley Creek	TRL000.02	5/30/96	2	57.57	9259.5	3981.8	18.9	21.4	0	40.3	40.3	18.9	16344
Turley Creek	TRL000.02	10/16/96	2	57.57	9259.5	3981.8	18.9	21.4	0	40.3	40.3	18.9	16344
Turley Creek	TRL000.03	5/29/97	2	57.57	9259.5	3981.8	18.9	21.4	0	40.3	40.3	18.9	16344
Turley Creek	TRL000.03	10/8/97	2	57.57	9259.5	3981.8	18.9	21.4	0	40.3	40.3	18.9	16344
Turley Creek	TRL000.04	10/23/98	2	57.57	9259.5	3981.8	18.9	21.4	0	40.3	40.3	18.9	16344

Table B.3 Summary of the stations used in the data set to investigate TSS and other stressor variables

STATION	RIVERMILE	COUNTY	ECOREGION	LANDUSE	SURVEY REASON / STATION DESCRIPTION
Back Creek	BCK000.78	Augusta	Central Appalachian Ridges and Valleys	Silvaculture	Monitor agricultural NPS.
Beaver Creek (BVR)	BVR003.60	Rockingham	Central Appalachian Ridges and Valleys	Silvaculture	Monitor agricultural NPS.
Buffalo Creek	BLD000.22	Rockbridge	Central Appalachian Ridges and Valleys	Pasture/Hay	This is the major tributary of the Maury River that drains the Karst portion of Rockbridge County.
Bull Run	BUL010.28	Prince William	Northern Piedmont	Suburban	This station is used to monitor impacts from urban NPS, storm water discharges and construction activities, and agriculture.
Catoctin Creek	CAX004.78	Loudoun	Northern Piedmont	Forest	This station is used as one of the reference sites.
Cedar Run	CER016.46	Fauquier Co.	Northern Piedmont	Pasture/Hay	Expand coverage in this agricultural watershed.
Cooks Creek	CKS003.04	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	Track effects of NPS abatement efforts within the watershed.
Cub Run	CUB000.40	Page	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor discharge.
Difficult Run	DIF000.86	Fairfax	Northern Piedmont	Urban	This station monitors impacts in an urban watershed. Storm sewer discharges and other nonpoint sources continue to impact this urban watershed.
Goose Creek	GOO044.36	Fauquier	Northern Piedmont	Cropland	This station monitors the upper Goose Creek watershed.
Great Run	GRT001.70	Fauquier	Northern Piedmont	Pasture/Hay	This station assesses ambient quality in an agricultural segment.
Hays Creek	HYS001.41	Rockbridge	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor agricultural NPS.

Table B.3 (continued) Summary of the stations used in the data set to investigate TSS and other stressor variables

STATION	RIVERMILE	COUNTY	ECOREGION	LANDUSE	SURVEY REASON / STATION DESCRIPTION
Hazel River	HAZ032.54	Culpeper	Northern Piedmont	Cropland	This station assesses impacts from agricultural nonpoint sources.
Hogue Creek	HOC006.23	Frederick	Central Appalachian Ridges and Valleys	Silvaculture	Monitor agricultural NPS.
Holmans Creek	HMN002.09	Shenandoah	Central Appalachian Ridges and Valleys	Orchard	Monitor effects of agricultural NPS, and impoundment.
Hughes River	HUE000.20	Rappahannock	Northern Piedmont	Cropland	This station assesses impacts from agricultural nonpoint sources.
Kerrs Creek	KRR001.54	Rockbridge	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor the effects of agricultural NPSs.
Linville	LNV000.71	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor the effects of agricultural NPSs; TMDL station. Station is highly affected by channelization and siltation.
Little River	LIV004.78	Loudoun Co.	Northern Piedmont	Suburban	Expand coverage in the suburban Potomac River watershed. Station impacted by rural development and an upstream impoundment.
Mill Creek (L)	MIL002.20	Shenandoah	Central Appalachian Ridges and Valleys	Pasture/Hay	Tributary to North Fork Shenandoah River.
Mill Creek (C)	MIC001.00	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	TMDL station. Station in middle of heavily used cattle pasture.
Moffett Creek	MFT006.24	Augusta	Central Appalachian Ridges and Valleys	Pasture/Hay	Coincident with MAHA station. Station in middle of heavily used cattle pasture.
Mountain Run	MTN000.59	Culpeper	Northern Piedmont	Cropland	This station monitors agricultural and rural nonpoint sources.
Muddy Creek (rt)	MUD005.81	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	Station coincident with TMDL and NAWQA studies. Station affected by agricultural NPS and cattle in stream.
Pleasant Run	PLE000.08	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	Track agricultural NPS and effects of pollution events.

Table B.3 (continued) Summary of the stations used in the data set to investigate TSS and other stressor variables

STATION	RIVERMILE	COUNTY	ECOREGION	LANDUSE	SURVEY REASON / STATION DESCRIPTION
Popes Head	POE002.00	Fairfax Co.	Northern Piedmont	Urban	Monitoring of urban development and agricultural NPS impacts.
Rappahannock River (147)	RPP147.10	Culpeper/Fauquier	Northern Piedmont	Cropland	Expand basin coverage and site control above Remington STP
Rappahannock River (175)	RPP175.51	Fauquier	Northern Piedmont	Cropland	To assess impacts in the upper watershed from agricultural NPS.
Robinson River	ROB001.90	Culpeper	Northern Piedmont	Pasture/Hay	This station is used as one of the reference sites.
South River (00)	STH000.21	Rockbridge	Central Appalachian Ridges and Valleys	Pasture/Hay	Trend station near confluence with the Maury River. Station well buffered by limestone springs; yet, the head waters of this stream may suffer from episodic flushes of acidic run off.
South River (27)	STH027.08	Augusta	Central Appalachian Ridges and Valleys	Suburban	Control upstream of Waynesboro dischargers and urban NPSs.
South Run	SOT001.44	Fauquier	Northern Piedmont	Forest	EPA Core Station. Monitoring Vint Hill STP. Upstream discharge from Lake Brittle.
Toms Brook	TMB000.54	Shenandoah	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor effect of municipal discharge.
Turley Creek	TRL000.02	Rockingham	Central Appalachian Ridges and Valleys	Pasture/Hay	VDCR Watershed Project. Station impacted by agricultural landuse.

Table B.4 Data collected for the data set used to investigate TSS and other stressor variables

STATION	RIVERMILE	DATE	RBP	Taxa rich	MFBI	Dompct	EPTI	S_T	TSS (mg/L)	COVER	EMBED	SEDIMENT	SUBSTRATE
Back Creek	BCK000.78	11/5/96	2	15	5.04	0.268	7	0.081	3	11	19	7	17
Back Creek	BCK000.78	5/5/97	2	15	4.94	0.263	9	0.026	3	7	15	9	19
Back Creek	BCK000.78	9/18/97	2	12	5.94	0.505	2	0.028	3	7	17	9	15
Beaver Creek (BVR)	BVR003.60	5/8/97	2	17	4.14	0.235	9	0.088	4	7	13	15	17
Buffalo Creek	BLD000.22	10/15/98	1	11	3.70	0.252	5	0.000	4.5	1	19	19	20
Bull Run	BUL010.28	11/5/96	2	14	5.37	0.213	4	0.000	5	17	13	15	17
Bull Run	BUL010.28	4/9/97	2	16	5.01	0.152	5	0.024	5	19	18	18	18
Bull Run	BUL010.28	8/27/97	2	16	5.81	0.235	4	0.029	7	19	17	18	17
Bull Run	BUL010.28	11/4/98	2	17	5.21	0.363	5	0.014	7	18	16	18	18
Catoctin Creek	CAX004.78	10/24/96	0	15	3.71	0.341	6	0.040	4	19	19	17	19
Catoctin Creek	CAX004.78	4/4/97	0	18	3.76	0.218	6	0.016	4	19	19	19	19
Catoctin Creek	CAX004.78	10/1/97	0	22	3.58	0.169	7	0.108	6	18	18	18	20
Catoctin Creek	CAX004.78	5/26/98	0	17	3.41	0.170	7	0.020	6	18	17	17	18
Catoctin Creek	CAX004.78	11/2/98	0	17	4.06	0.226	7	0.058	3	19	20	19	19
Cedar Run	CER016.46	3/28/97	2	20	4.94	0.216	5	0.020	3	19	18	18	20
Cedar Run	CER016.46	10/1/97	0	19	4.91	0.223	5	0.029	3	17	19	18	19
Cedar Run	CER016.46	4/27/98	0	17	3.83	0.153	10	0.014	3	19	20	19	19
Cedar Run	CER016.46	10/6/98	0	19	4.61	0.310	7	0.000	7	17	18	17	19
Cooks Creek	CKS003.04	10/3/96	2	9	6.07	0.500	2	0.010	26	11	7	9	7
Cooks Creek	CKS003.04	5/22/97	3	9	6.03	0.430	2	0.008	26	13	5	9	11
Cooks Creek	CKS003.04	9/17/97	2	15	6.22	0.288	3	0.018	45	13	5	5	9
Cooks Creek	CKS003.04	10/6/98	2	17	5.49	0.469	4	0.008	20	1	5	18	3
Cub Run	CUB000.40	5/6/97	2	18	4.07	0.228	10	0.065	3.00	13	19	17	19
Cub Run	CUB000.40	9/25/97	0	15	3.78	0.217	8	0.038	3.00	17	17	17	15
Difficult Run	DIF000.86	11/5/96	2	13	4.66	0.243	4	0.057	4	17	19	17	18
Difficult Run	DIF000.86	4/15/97	0	16	4.39	0.250	6	0.060	4	19	19	18	19
Difficult Run	DIF000.86	9/5/97	2	14	4.59	0.202	4	0.016	5	19	19	18	19
Difficult Run	DIF000.86	6/27/98	2	14	5.51	0.490	3	0.034	5	19	18	17	18
Difficult Run	DIF000.86	10/6/98	0	16	4.35	0.206	5	0.028	5	19	19	17	19
Goose Creek	GOO044.36	11/18/96	0	13	3.92	0.328	7	0.136	3	18	18	18	19

Table B.4 (continued) Data collected for the data set used to investigate TSS and other stressor variables

STATION	RIVERMILE	DATE	RBP	Taxa rich	MFBI	Dompct	EPTI	S_T	TSS (mg/L)	COVER	EMBED	SEDIMENT	SUBSTRATE
Great Run	GRT001.70	10/21/96	2	16	4.37	0.234	6	0.016	3	18	18	19	20
Great Run	GRT001.70	3/10/97	0	19	4.36	0.258	6	0.008	3	18	19	17	19
Great Run	GRT001.70	9/17/97	2	19	3.91	0.335	4	0.012	3.5	19	20	19	19
Great Run	GRT001.70	3/16/98	0	18	4.26	0.280	8	0.006	3.5	19	20	18	19
Great Run	GRT001.70	10/27/98	0	18	4.45	0.411	5	0.037	4	19	19	18	19
Hays Creek	HYS001.41	9/24/97	2	16	4.44	0.240	6	0.029	15.5	13	7	11	9
Hazel River	HAZ032.54	10/17/96	0	19	4.34	0.170	6	0.045	3	16	17	17	18
Hazel River	HAZ032.54	4/2/97	0	15	4.36	0.272	6	0.018	3	18	19	18	18
Hazel River	HAZ032.54	10/7/97	0	21	3.92	0.168	7	0.030	3	18	18	18	19
Hazel River	HAZ032.54	3/30/98	0	18	4.05	0.208	7	0.020	3	18	18	16	17
Hazel River	HAZ032.54	10/28/98	0	17	3.91	0.336	7	0.007	7	18	19	16	18
Hogue Creek	HOC006.23	10/7/97	2	15	4.29	0.397	5	0.032	3.00	11	11	9	11
Hogue Creek	HOC006.23	10/9/98	1	12	3.57	0.313	5	0.000	3.5	1	17	13	17
Holmans Creek	HMN002.09	10/16/96	2	13	5.19	0.265	7	0.009	8	15	15	11	15
Holmans Creek	HMN002.09	5/29/97	2	13	4.87	0.359	6	0.026	8	11	13	13	13
Holmans Creek	HMN002.09	10/8/97	2	11	4.85	0.487	5	0.009	8	15	13	13	13
Holmans Creek	HMN002.09	10/27/98	2	10	5.11	0.557	5	0.062	8	1	16	9	19
Hughes River	HUE000.20	10/17/96	0	16	4.37	0.257	6	0.030	3	19	19	19	19
Hughes River	HUE000.20	4/2/97	0	14	3.67	0.349	6	0.018	3	19	19	19	19
Hughes River	HUE000.20	10/7/97	0	19	4.34	0.241	7	0.015	3	19	19	19	19
Hughes River	HUE000.20	3/30/98	0	14	3.94	0.273	8	0.007	3	20	18	18	19
Hughes River	HUE000.20	10/28/98	0	15	3.51	0.343	7	0.017	3	19	19	19	20
Kerrs Creek	KRR001.54	10/2/97	2	14	3.93	0.279	7	0.018	3	17	17	13	17
Kerrs Creek	KRR001.54	10/15/98	1	16	4.23	0.299	6	0.017	3	1	21	20	16
Linville	LNV000.71	9/22/97	2	16	6.37	0.333	6	0.017	5	11	11	7	9
Linville	LNV000.71	10/23/98	2	14	5.54	0.324	4	0.000	15	1	1	1	16
Little River	LIV004.78	4/4/97	2	13	4.65	0.268	4	0.009	20.5	16	18	19	17
Little River	LIV004.78	10/1/97	2	16	4.29	0.205	5	0.034	4	17	18	18	17
Little River	LIV004.78	11/23/98	0	15	4.27	0.240	5	0.040	7.5	19	18	18	18

Table B.4 (continued) Data collected for the data set used to investigate TSS and other stressor variables

STATION	RIVERMILE	DATE	RBP	Taxa rich	MFBI	Dompct	EPTI	S_T	TSS (mg/L)	COVER	EMBED	SEDIMENT	SUBSTRATE
Mill Creek (L)	MIL002.20	10/15/96	2	17	5.15	0.300	6	0.030	3.5	15	17	17	15
Mill Creek (L)	MIL002.20	5/27/97	2	17	5.21	0.349	8	0.008	3.5	15	11	13	15
Mill Creek (L)	MIL002.20	9/23/97	2	19	4.67	0.339	6	0.104	3	17	13	15	15
Mill Creek (L)	MIL002.20	10/20/98	2	14	5.12	0.366	6	0.014	3.5	1	17	18	16
Mill Creek (C)	MIC001.00	10/24/96	2	14	6.05	0.245	5	0.018	6	13	11	15	11
Mill Creek (C)	MIC001.00	5/6/97	2	16	5.37	0.368	5	0.048	6	11	9	7	7
Mill Creek (C)	MIC001.00	10/19/98	2	13	4.89	0.321	4	0.009	4	1	11	11	11
Moffett Creek	MFT006.24	5/8/97	2	18	4.31	0.163	8	0.061	3	19	11	13	15
Moffett Creek	MFT006.24	10/14/97	2	13	4.96	0.377	6	0.019	7.5	17	5	5	15
Moffett Creek	MFT006.24	10/6/98	2	12	4.99	0.431	4	0.236	3.5	1	11	4	7
Mountain Run	MTN000.59	10/29/96	2	17	5.23	0.162	4	0.063	3	19	18	17	19
Mountain Run	MTN000.59	3/10/97	2	13	5.37	0.439	3	0.013	3	19	18	18	19
Mountain Run	MTN000.59	10/2/97	0	28	4.43	0.139	9	0.093	6	19	19	19	19
Mountain Run	MTN000.59	4/2/98	2	19	4.85	0.236	6	0.021	6	19	19	18	20
Mountain Run	MTN000.59	10/20/98	0	20	4.68	0.237	5	0.012	4	19	19	18	19
Muddy Creek (rt)	MUD005.81	10/22/96	2	9	6.82	0.373	2	0.000	11	13	13	15	15
Muddy Creek (rt)	MUD005.81	4/30/97	2	14	5.75	0.323	5	0.048	11	13	9	11	11
Muddy Creek (rt)	MUD005.81	9/23/97	2	13	6.87	0.450	3	0.016	21	11	13	11	13
Pleasant Run	PLE000.08	10/24/96	3	6	7.55	0.602	1	0.000	21	17	7	13	7
Pleasant Run	PLE000.08	4/30/97	3	7	6.97	0.341	0	0.000	21	11	7	11	7
Pleasant Run	PLE000.08	9/17/97	2	9	7.95	0.962	1	0.000	18.5	13	5	17	5
Popes Head	POE002.00	4/15/97	2	8	4.77	0.355	4	0.000	3	17	13	16	15
Popes Head	POE002.00	10/29/97	2	12	4.18	0.230	4	0.082	3	18	15	17	18
Popes Head	POE002.00	12/7/98	2	13	4.67	0.218	4	0.000	3	17	13	16	17
Rappahannock River (147)	RPP147.10	5/12/97	0	18	3.16	0.348	8	0.007	7	21	19	19	21
Rappahannock River (147)	RPP147.10	8/18/97	0	17	2.75	0.193	8	0.193	9	19	19	19	21
Rappahannock River (147)	RPP147.10	9/21/98	0	21	3.65	0.197	8	0.215	4	20	18	18	21
Rappahannock River (175)	RPP175.51	11/26/96	0	17	3.81	0.211	7	0.085	4	18	19	20	19

Table B.4 (continued) Data collected for the data set used to investigate TSS and other stressor variables

STATION	RIVERMILE	DATE	RBP	Taxa rich	MFBI	Dompct	EPTI	S_T	TSS (mg/L)	COVER	EMBED	SEDIMENT	SUBSTRATE
Robinson River	ROB001.90	11/26/96	0	16	3.73	0.248	7	0.064	4	18	18	18	20
Robinson River	ROB001.90	4/7/97	0	16	3.99	0.394	7	0.028	4	17	16	16	16
Robinson River	ROB001.90	10/6/97	0	22	3.88	0.173	9	0.080	10	20	20	19	20
Robinson River	ROB001.90	5/28/98	0	18	3.39	0.188	10	0.009	10	19	18	18	19
Robinson River	ROB001.90	9/17/98	0	19	4.40	0.215	7	0.000	3	19	20	18	20
South River (00)	STH000.21	10/2/97	0	22	4.06	0.217	10	0.078	3	15	19	17	19
South River (00)	STH000.21	10/15/98	1	13	4.52	0.369	6	0.019	3	1	21	20	21
South River (27)	STH027.08	10/2/98	1	16	5.60	0.625	7	0.008	4.5	1	20	20	20
South Run	SOT001.44	10/21/96	2	19	5.13	0.309	4	0.024	6	18	18	19	19
South Run	SOT001.44	3/11/97	2	16	5.00	0.289	4	0.020	6	19	19	19	19
South Run	SOT001.44	9/17/97	2	18	4.62	0.255	4	0.032	4.5	19	19	18	19
South Run	SOT001.44	3/30/98	2	20	4.85	0.268	4	0.000	4.5	19	19	18	19
South Run	SOT001.44	10/21/98	2	17	4.99	0.310	4	0.024	3.5	17	19	18	18
Toms Brook	TMB000.54	10/16/97	0	25	4.85	0.278	12	0.073	3	17	15	15	13
Turley Creek	TRL000.02	10/16/96	2	9	5.56	0.515	4	0.000	3	13	11	9	11
Turley Creek	TRL000.02	5/29/97	2	17	5.01	0.387	9	0.077	3	13	13	15	11
Turley Creek	TRL000.02	10/8/97	2	12	4.64	0.471	6	0.013	8	15	13	15	13
Turley Creek	TRL000.02	10/23/98	2	11	4.30	0.357	7	0.032	3.5	1	16	19	18

Table B.4 (continued) Data collected for the data set used to investigate TSS and other stressor variables

STATION	VELOCITY	ALTER	RIFFLES	FLOW	BANKS	BANK VEG	GRAZE	RIPVEG	DO (mg/L)	TEMP (°C)	pH	Hard (mg/L)	ALK (mg/L)
Back Creek	17	7	9	17	7	7	21	9	11.90	10.70	7.40	14	8
Back Creek	15	7	11	17	7	3	1	1	9.80	18.00	8.40	14	8
Back Creek	9	1	5	11	11	7	11	1	12	8.8	7.4	14.55	10
Beaver Creek (BVR)	15	19	13	21	13	7	7	1	10.80	12.60	7.80	37	30.6
Buffalo Creek	21	17	15	18	14	16	1	13	11.70	11.10	8.50	107.5	138
Bull Run	18	18	13	19	16	17	13	13	12.00	12.90	7.70	98.45	64.3
Bull Run	18	19	15	19	16	17	15	15	13.30	13.60	7.90	98.45	64.3
Bull Run	19	19	15	20	19	18	13	16	9.30	25.70	7.10	102.8	64.85
Bull Run	19	19	16	19	18	17	15	15	9.00	13.90	7.60	124	75.05
Catoctin Creek	18	19	18	19	19	20	19	20	10.80	12.70	6.90	58	33.5
Catoctin Creek	20	20	19	20	18	18	18	18	13.20	16.30	7.30	58	33.5
Catoctin Creek	18	20	18	19	19	19	18	18	11.10	16.60	7.50	57.7	33.7
Catoctin Creek	19	18	18	19	18	18	18	18	1.40	20.60	8.00	57.7	33.7
Catoctin Creek	17	19	18	19	18	19	17	18	8.80	10.80	7.30	68	42.45
Cedar Run	18	18	19	18	17	18	17	18	13.60	14.40	7.70	52	42.8
Cedar Run	17	19	18	20	19	19	16	19	10.80	17.30	7.60	57.4	37.6
Cedar Run	19	19	17	20	19	20	18	18	10.40	16.60	8.30	57.4	37.6
Cedar Run	16	19	15	16	18	18	16	17	10.00	17.20	7.90	68	54.85
Cooks Creek	13	9	5	21	5	5	7	1	10.80	17.60	8.20	297	234.1
Cooks Creek	15	17	11	21	9	9	5	1	12.90	20.90	8.70	297	234.1
Cooks Creek	11	17	9	21	13	9	7	1	10.2	16.4	8	257	228.5
Cooks Creek	9	14	12	21	12	10	1	2	8.80	16.10	8.20	263	233
Cub Run	15	17	19	19	15	13	11	1	9.60	15.50	8.30	32	26.7
Cub Run	15	19	19	19	15	13	3	1	9.60	14.00	8.00	55	49.35
Difficult Run	19	19	18	20	19	17	15	16	11.80	12.60	7.30	38.55	22.8
Difficult Run	19	20	17	19	19	20	15	17	12.70	12.50	8.20	38.55	22.8
Difficult Run	20	20	19	21	19	19	16	18	10.10	17.40	7.70	46.3	27.9
Difficult Run	19	18	18	20	19	19	15	16	8.20	21.90	7.60	46.3	27.9
Difficult Run	19	19	17	19	18	18	16	16	8.50	15.10	6.80	48	32.2
Goose Creek	19	19	18	19	18	18	18	18	12.60	8.10	8.20	38	23.3

Table B.4 (continued)   Data collected for the data set used to investigate TSS and other stressor variables

STATION	VELOCITY	ALTER	RIFFLES	FLOW	BANKS	BANK VEG	GRAZE	RIPVEG	DO (mg/L)	TEMP (°C)	pH	Hard (mg/L)	ALK (mg/L)
Great Run	20	19	18	19	17	17	17	15	11.20	10.90	7.00	46.35	33.9
Great Run	18	17	16	19	16	16	15	15	12.50	11.60	7.30	46.35	33.9
Great Run	18	20	18	20	19	19	17	19	11.60	21.10	7.30	68.2	35.75
Great Run	19	19	18	19	18	15	15	15	11.70	14.40	7.60	68.2	35.75
Great Run	19	18	18	18	19	19	17	18	9.40	11.40	7.00	59	32.25
Hays Creek	11	19	9	21	17	17	19	15	9.00	15.80	8.10	194	194
Hazel River	19	17	19	19	12	16	18	16	11.10	14.10	7.60	13.5	14.6
Hazel River	20	19	18	19	15	16	17	15	13.10	13.10	7.20	13.5	14.6
Hazel River	18	19	19	19	17	17	18	17	11.50	18.30	7.30	15.05	14.4
Hazel River	19	15	16	19	13	15	17	16	11.30	18.00	7.50	15.05	14.4
Hazel River	18	13	17	17	17	17	15	15	10.00	13.50	7.30	16	11.85
Hogue Creek	17	13	13	17	7	7	1	1	8.70	18.60	7.90	105	84.7
Hogue Creek	18	14	13	15	9	1	1	1	10.80	16.00	8.10	92	137
Holmans Creek	11	9	7	21	9	9	11	1	11.40	16.5	7.80	273.5	247.6
Holmans Creek	11	17	9	21	9	9	9	1	10.00	16.20	8.00	273.5	247.6
Holmans Creek	13	15	11	21	13	11	7	1	9.20	17.40	7.70	268	246.5
Holmans Creek	14	12	17	21	11	9	1	3	11.90	10.90	8.20	264	247
Hughes River	20	19	19	20	18	19	19	19	11.30	13.80	7.10	12	11.4
Hughes River	20	19	19	20	19	19	18	18	13.50	12.40	7.60	12	11.4
Hughes River	20	20	19	19	18	18	17	18	11.00	17.90	7.20	15.05	10.85
Hughes River	19	20	19	20	17	18	17	15	11.30	16.50	7.40	15.05	10.85
Hughes River	18	19	19	19	19	20	18	18	9.90	12.50	7.00	16	7.1
Kerrs Creek	17	11	13	21	17	15	19	1	11.1	26	8.3	178	169
Kerrs Creek	19	16	21	17	17	16	1	11	12.40	13.70	8.40	166	155
Linville	11	11	11	21	7	7	7	1	13.00	9.50	8.50	244	224
Linville	11	12	17	19	3	9	1	1	10.20	16.40	8.10	261.5	231.5
Little River	18	19	17	19	18	18	17	20	12.40	14.80	7.70	41.1	25.45
Little River	17	18	17	19	19	19	18	19	10.90	15.90	7.50	39	32.05
Little River	17	18	16	19	18	20	16	19	11.90	5.80	6.90	44	34.9

Table B.4 (continued) Data collected for the data set used to investigate TSS and other stressor variables

STATION	VELOCITY	ALTER	RIFFLES	FLOW	BANKS	BANK VEG	GRAZE	RIPVEG	DO (mg/L)	TEMP (^{o}C)	pH	Hard (mg/L)	ALK (mg/L)
Mill Creek (L)	17	15	19	21	17	17	21	11	13.70	14.30	8.30	203.5	188.5
Mill Creek (L)	19	11	19	21	17	19	17	11	11.10	17.20	8.40	203.5	188.5
Mill Creek (L)	17	19	19	19	17	15	17	13	10.00	15.40	8.00	170	192.5
Mill Creek (L)	11	16	21	18	21	18	1	7	10.40	14.20	8.10	230.5	217.5
Mill Creek (C)	13	17	9	21	11	11	7	1	11.00	14.90	8.30	254.9	222
Mill Creek (C)	15	17	9	19	9	9	5	1	11.20	20.20	8.50	254.9	222
Mill Creek (C)	12	13	4	18	3	1	1	1	10.90	16.90	8.20	234	230
Moffett Creek	13	15	13	21	9	11	9	7	10.50	15.90	8.10	94.1	82.35
Moffett Creek	13	15	13	11	7	5	3	1	10.4	13.05	7.95	135.05	118.65
Moffett Creek	12	16	8	20	5	3	1	1	8.40	16.80	7.80	136	120.5
Mountain Run	18	18	18	19	20	20	17	19	10.20	15.10	7.00	59.9	39.75
Mountain Run	19	19	18	19	19	20	16	18	13.50	11.10	7.90	59.9	39.75
Mountain Run	18	20	19	19	19	19	18	19	11.40	12.30	7.60	46.55	30.1
Mountain Run	19	19	18	19	18	17	17	18	10.80	19.90	7.90	46.55	30.1
Mountain Run	18	19	18	18	20	20	16	17	8.00	14.60	7.30	61.35	53.3
Muddy Creek (rt)	17	15	11	21	11	9	7	1	10.60	15.80	7.80	228	188.7
Muddy Creek (rt)	13	17	11	19	9	9	5	1	13.40	18.20	8.80	228	188.7
Muddy Creek (rt)	11	19	9	17	11	7	7	1	11.05	13.4	8	232	205
Pleasant Run	13	15	9	21	11	7	9	1	12.40	14.50	8.20	282	233
Pleasant Run	11	17	9	21	11	9	7	1	14.50	20.70	8.70	282	233
Pleasant Run	13	19	7	21	17	11	11	1	11.4	15.5	7.9	272.5	237.5
Popes Head	18	18	19	19	19	18	13	13	13.40	14.60	7.90	45	32.4
Popes Head	19	18	19	20	19	18	13	13	11.90	12.10	7.50	23.2	32.1
Popes Head	16	19	15	19	18	18	12	16	10.80	13.80	7.10	80	34.1
Rappahannock River (147)	21	20	17	20	16	19	18	19	10.80	18.00	7.30	23.9	20.8
Rappahannock River (147)	19	20	17	20	19	20	19	19	8.40	23.80	7.60	23	20.4
Rappahannock River (147)	19	20	17	21	17	18	18	18	9.90	26.20	7.50	28	24.2
Rappahannock River (175)	18	20	17	19	18	18	17	19	12.90	4.40	7.50	25.25	20.15

Table B.4 (continued)   Data collected for the data set used to investigate TSS and other stressor variables

STATION	VELOCITY	ALTER	RIFFLES	FLOW	BANKS	BANK VEG	GRAZE	RIPVEG	DO (mg/L)	TEMP (^{o}C)	pH	Hard (mg/L)	ALK (mg/L)
Robinson River	17	18	16	19	18	19	17	18	13.30	4.20	7.20	25	20.6
Robinson River	19	17	17	19	17	17	16	15	10.80	17.80	7.50	25	20.6
Robinson River	19	19	18	21	20	19	19	19	10.90	19.70	7.40	20.1	18
Robinson River	20	20	18	21	19	18	17	16	9.90	18.00	7.90	20.1	18
Robinson River	19	19	18	20	19	19	18	17	7.20	24.40	6.70	20	18.6
South River (00)	17	17	15	21	13	15	17	3	10.6	13.3	8	109.5	100.5
South River (00)	19	16	19	16	18	17	1	17	11.90	12.80	8.50	132.5	123.5
South River (27)	17	19	21	16	14	15	1	10	10.10	16.50	8.40	107	101
South Run	18	20	19	19	21	21	16	19	10.80	11.90	7.30	49.9	34.1
South Run	18	19	20	18	20	20	15	19	12.90	11.40	7.10	49.9	34.1
South Run	16	20	19	16	21	21	17	19	10.20	21.90	7.20	55.15	31.3
South Run	15	18	19	17	20	21	16	19	11.00	21.20	7.90	55.15	31.3
South Run	10	19	12	9	20	20	16	19	7.30	13.70	7.20	110	37.7
Toms Brook	15	17	17	21	17	15	13	9	10.60	10.40	8.00	224.5	217.5
Turley Creek	11	11	13	21	13	11	17	1	11.20	11.40	8.20	246	225.35
Turley Creek	17	13	17	21	9	7	9	1	10.90	13.80	8.30	246	225.35
Turley Creek	13	13	15	17	13	11	11	1	10.3	16.8	8.15	241	220
Turley Creek	16	13	20	17	10	11	1	1	12.00	7.80	8.40	250	233

Table B.4 (continued) Data collected for the data set used to investigate TSS and other stressor variables

STATION	COND (μΩ)	TURB (FTU)	Beef #	Wshed area (ac)	Stream Length (m)	Stream Order	% Pasture	% Cropland	% Pasture + Cropland (AG)	% Forest	% Urban
Back Creek	36.7	2.3	184	26644.81	22733	2	2.62	5.61	8.23	89.44	1.47
Back Creek	36.7	2.3	184	26644.81	22733	2	2.62	5.61	8.23	89.44	1.47
Back Creek	33.1	1.5	184	26644.81	22733	2	2.62	5.61	8.23	89.44	1.47
Beaver Creek (BVR)	88.4	6.1	249	6476.64	9655	2	9.12	4.92	14.03	85.45	0.02
Buffalo Creek	279.5	2.885	1346	77509.88	32869	3	9.89	14.86	24.75	74.11	0.44
Bull Run	362	7	0	1175.07	4400	4	2.38	19.96	22.34	31.14	36.92
Bull Run	362	7	0	1175.07	4400	4	2.38	19.96	22.34	31.14	36.92
Bull Run	385	6.5	0	1175.07	4400	4	2.38	19.96	22.34	31.14	36.92
Bull Run	604	6.335	0	1175.07	4400	4	2.38	19.96	22.34	31.14	36.92
Catoctin Creek	149	5.75	8723	59532.34	41270	4	36.33	29.23	65.56	33.07	0.48
Catoctin Creek	149	5.75	8723	59532.34	41270	4	36.33	29.23	65.56	33.07	0.48
Catoctin Creek	161	6	8723	59532.34	41270	4	36.33	29.23	65.56	33.07	0.48
Catoctin Creek	161	6	8723	59532.34	41270	4	36.33	29.23	65.56	33.07	0.48
Catoctin Creek	175.5	4.26	8723	59532.34	41270	4	36.33	29.23	65.56	33.07	0.48
Cedar Run	148	6.3	3623	65510.12	30045	4	20.53	25.44	45.97	47.16	5.89
Cedar Run	164	9	3623	65510.12	30045	4	20.53	25.44	45.97	47.16	5.89
Cedar Run	164	9	3623	65510.12	30045	4	20.53	25.44	45.97	47.16	5.89
Cedar Run	189.5	3.38	3623	65510.12	30045	4	20.53	25.44	45.97	47.16	5.89
Cooks Creek	571	22	1526	14500.36	17127	3	52.43	24.54	76.97	12.22	10.40
Cooks Creek	571	22	1526	14500.36	17127	3	52.43	24.54	76.97	12.22	10.40
Cooks Creek	537	38	1526	14500.36	17127	3	52.43	24.54	76.97	12.22	10.40
Cooks Creek	505	20.8	1526	14500.36	17127	3	52.43	24.54	76.97	12.22	10.40
Cub Run	74.4	2.9	37	10182.79	14539	2	1.84	1.19	3.03	96.69	0.27
Cub Run	121.5	1	37	10182.79	14539	2	1.84	1.19	3.03	96.69	0.27
Difficult Run	142	4.8	0	37010.72	22435	*4*	6.31	11.98	18.29	52.39	24.80
Difficult Run	142	4.8	0	37010.72	22435	*4*	6.31	11.98	18.29	52.39	24.80
Difficult Run	158	8	0	37010.72	22435	*4*	6.31	11.98	18.29	52.39	24.80
Difficult Run	158	8	0	37010.72	22435	*4*	6.31	11.98	18.29	52.39	24.80
Difficult Run	186	6.99	0	37010.72	22435	4	6.31	11.98	18.29	52.39	24.80
Goose Creek	110	4.1	4716	28557.60	15960	4	16.23	15.67	31.90	66.81	1.01

Table B.4 (continued) Data collected for the data set used to investigate TSS and other stressor variables

STATION	COND (μΩ)	TURB (FTU)	Beef #	Wshed area (ac)	Stream Length (m)	Stream Order	% Pasture	% Cropland	% Pasture + Cropland (AG)	% Forest	% Urban
Great Run	153	3.9	2088	16127.34	22340	4	24.05	23.00	47.06	48.94	3.07
Great Run	153	3.9	2088	16127.34	22340	4	24.05	23.00	47.06	48.94	3.07
Great Run	168.5	4.5	2088	16127.34	22340	4	24.05	23.00	47.06	48.94	3.07
Great Run	168.5	4.5	2088	16127.34	22340	4	24.05	23.00	47.06	48.94	3.07
Great Run	215.5	6.71	2088	16127.34	22340	4	24.05	23.00	47.06	48.94	3.07
Hays Creek	381.5	8	823	15845.93	8950	3	22.88	20.69	43.57	56.20	0.10
Hazel River	42.8	4.1	1240	14936.90	22560	4	8.87	16.44	25.31	74.18	0.17
Hazel River	42.8	4.1	1240	14936.90	22560	4	8.87	16.44	25.31	74.18	0.17
Hazel River	43.65	4	1240	14936.90	22560	4	8.87	16.44	25.31	74.18	0.17
Hazel River	43.65	4	1240	14936.90	22560	4	8.87	16.44	25.31	74.18	0.17
Hazel River	42.65	4.09	1240	14936.90	22560	4	8.87	16.44	25.31	74.18	0.17
Hogue Creek	221	2	691	21176.60	16748	2	5.68	11.70	17.37	80.79	0.79
Hogue Creek	327	3.1	691	21176.60	16748	2	5.68	11.70	17.37	80.79	0.79
Holmans Creek	529.5	6.55	1144	11177.76	15659	2	29.34	38.84	68.19	30.05	0.95
Holmans Creek	529.5	6.55	1144	11177.76	15659	2	29.34	38.84	68.19	30.05	0.95
Holmans Creek	516.5	6.5	1144	11177.76	15659	2	29.34	38.84	68.19	30.05	0.95
Holmans Creek	504	4.88	1144	11177.76	15659	2	29.34	38.84	68.19	30.05	0.95
Hughes River	38.6	4.3	1693	30716.66	22910	4	9.07	16.08	25.15	74.16	0.33
Hughes River	38.6	4.3	1693	30716.66	22910	4	9.07	16.08	25.15	74.16	0.33
Hughes River	37.6	3	1693	30716.66	22910	4	9.07	16.08	25.15	74.16	0.33
Hughes River	37.6	3	1693	30716.66	22910	4	9.07	16.08	25.15	74.16	0.33
Hughes River	36.05	2.695	1693	30716.66	22910	4	9.07	16.08	25.15	74.16	0.33
Kerrs Creek	335	1	633	22178.15	14717	3	10.19	13.03	23.23	75.61	0.86
Kerrs Creek	332	1.13	633	22178.15	14717	3	10.19	13.03	23.23	75.61	0.86
Linville	493	5.9	3427	29858.14	21389	3	43.55	30.28	73.83	23.15	2.40
Linville	520.5	9.5	3427	29858.14	21389	3	43.55	30.28	73.83	23.15	2.40
Little River	100.65	23.15	2643	26718.39	25425	4	34.76	18.40	53.16	45.34	1.42
Little River	108.5	5.5	2643	26718.39	25425	4	34.76	18.40	53.16	45.34	1.42
Little River	125	6.8	2643	26718.39	25425	4	34.76	18.40	53.16	45.34	1.42

Table B.4 (continued) Data collected for the data set used to investigate TSS and other stressor variables

STATION	COND (μΩ)	TURB (FTU)	Beef #	Wshed area (ac)	Stream Length (m)	Stream Order	% Pasture	% Cropland	% Pasture + Cropland (AG)	% Forest	% Urban
Mill Creek (L)	414.5	3.3	1390	25999.07	21168	2	15.14	25.30	40.43	58.45	0.70
Mill Creek (L)	414.5	3.3	1390	25999.07	21168	2	15.14	25.30	40.43	58.45	0.70
Mill Creek (L)	431.5	2	1390	25999.07	21168	2	15.14	25.30	40.43	58.45	0.70
Mill Creek (L)	493	1.48	1390	25999.07	21168	2	15.14	25.30	40.43	58.45	0.70
Mill Creek (C)	495	6.9	547	9686.13	13695	2	44.92	30.87	75.79	18.57	3.34
Mill Creek (C)	495	6.9	547	9686.13	13695	2	44.92	30.87	75.79	18.57	3.34
Mill Creek (C)	470	4.22	547	9686.13	13695	2	44.92	30.87	75.79	18.57	3.34
Moffett Creek	196	3.6	1109	6123.29	11015	2	17.63	9.59	27.23	72.48	0.02
Moffett Creek	259	5	1109	6123.29	11015	2	17.63	9.59	27.23	72.48	0.02
Moffett Creek	258.5	3.715	1109	6123.29	11015	2	17.63	9.59	27.23	72.48	0.02
Mountain Run	176.5	4.9	8163	58363.85	44990	4	29.29	28.36	57.65	36.44	3.93
Mountain Run	176.5	4.9	8163	58363.85	44990	4	29.29	28.36	57.65	36.44	3.93
Mountain Run	153.5	15	8163	58363.85	44990	4	29.29	28.36	57.65	36.44	3.93
Mountain Run	153.5	15	8163	58363.85	44990	4	29.29	28.36	57.65	36.44	3.93
Mountain Run	305	4.12	8163	58363.85	44990	4	29.29	28.36	57.65	36.44	3.93
Muddy Creek (rt)	428	9.9	1841	9238.06	10351	2	47.38	24.48	71.86	26.91	1.05
Muddy Creek (rt)	428	9.9	1841	9238.06	10351	2	47.38	24.48	71.86	26.91	1.05
Muddy Creek (rt)	451	15.5	1841	9238.06	10351	2	47.38	24.48	71.86	26.91	1.05
Pleasant Run	540	15.6	1079	5343.02	10265	1	45.69	30.81	76.50	18.48	4.59
Pleasant Run	540	15.6	1079	5343.02	10265	1	45.69	30.81	76.50	18.48	4.59
Pleasant Run	521.5	14.5	1079	5343.02	10265	1	45.69	30.81	76.50	18.48	4.59
Popes Head	148.5	2.9	0	10678.62	11600	3	3.87	9.10	12.98	70.23	13.81
Popes Head	140.5	7	0	10678.62	11600	3	3.87	9.10	12.98	70.23	13.81
Popes Head	168.5	4.005	0	10678.62	11600	3	3.87	9.10	12.98	70.23	13.81
Rappahannock River (147)	71	7.3	879	8671.39	13560	4	33.70	34.95	68.65	28.34	2.15
Rappahannock River (147)	72.75	9	879	8671.39	13560	4	33.70	34.95	68.65	28.34	2.15
Rappahannock River (147)	89.05	2.84	879	8671.39	13560	4	33.70	34.95	68.65	28.34	2.15
Rappahannock River (175)	75.8	4.1	5772	47416.17	22130	4	14.55	15.98	30.53	68.03	0.57

Table B.4 (continued) Data collected for the data set used to investigate TSS and other stressor variables

STATION	COND (μΩ)	TURB (FTU)	Beef #	Wshed area (ac)	Stream Length (m)	Stream Order	% Pasture	% Cropland	% Pasture + Cropland (AG)	% Forest	% Urban
Robinson River	72	8.8	1941	9381.10	15800	4	21.91	27.42	49.33	47.01	1.53
Robinson River	72	8.8	1941	9381.10	15800	4	21.91	27.42	49.33	47.01	1.53
Robinson River	67.9	8	1941	9381.10	15800	4	21.91	27.42	49.33	47.01	1.53
Robinson River	67.9	8	1941	9381.10	15800	4	21.91	27.42	49.33	47.01	1.53
Robinson River	66	3.8	1941	9381.10	15800	4	21.91	27.42	49.33	47.01	1.53
South River (00)	226	2.5	717	19854.91	13835	4	18.93	19.81	38.75	60.44	0.57
South River (00)	267	1.35	717	19854.91	13835	4	18.93	19.81	38.75	60.44	0.57
South River (27)	226	4.5	41	2250.76	5377	4	4.98	17.30	22.29	67.08	8.65
South Run	151	5.9	4803	37977.96	23325	4	21.44	21.05	42.49	52.83	4.11
South Run	151	5.9	4803	37977.96	23325	4	21.44	21.05	42.49	52.83	4.11
South Run	121	7.5	4803	37977.96	23325	4	21.44	21.05	42.49	52.83	4.11
South Run	121	7.5	4803	37977.96	23325	4	21.44	21.05	42.49	52.83	4.11
South Run	182.5	1.78	4803	37977.96	23325	4	21.44	21.05	42.49	52.83	4.11
Toms Brook	499.5	2	535	10367.56	14265	2	24.08	25.69	49.78	44.42	5.34
Turley Creek	471.5	2.65	1281	9259.48	16344	1	18.90	21.40	40.30	57.05	1.77
Turley Creek	471.5	2.65	1281	9259.48	16344	1	18.90	21.40	40.30	57.05	1.77
Turley Creek	440.5	4	1281	9259.48	16344	1	18.90	21.40	40.30	57.05	1.77
Turley Creek	480.5	1.645	1281	9259.48	16344	1	18.90	21.40	40.30	57.05	1.77

Table B.5 Summary of the stations used in the validation data set

STATION	RIVERMILE	COUNTY	ECOREGION	LANDUSE	SURVEY REASON / STATION DESCRIPTION
Beaver Creek (BRC)	BRC001.88	Orange Co.	Northern Piedmont	Pasture/Hay	Survey and assess agricultural NPS in the upper York River Basin
Crab Creek (04)	CBC004.38	Montgomery	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor agricultural NPS.
Crab Creek (06)	CBC006.35	Montgomery	Central Appalachian Ridges and Valleys	Pasture/Hay	Monitor agricultural NPS.
James River	JMS326.30	Botetourt	Central Appalachian Ridges and Valleys	Pasture/Hay	To assess the water body.
Little Otter River	LOR014.75	Bedford	Central Appalachian Ridges and Valleys	Pasture/Hay	Upstream control for Bedford STP.
Mechums River	MCM018.92	Albemarle	Northern Piedmont	Pasture/Hay	Monitor agricultural NPS.
Peak Creek (09)	PKC009.29	Pulaski	Central Appalachian Ridges and Valleys	Urban	Downstream of Magnox Inc. and upstream of Downtown East Property.
Peak Creek (11)	PKC011.11	Pulaski	Central Appalachian Ridges and Valleys	Suburban	Upstream control
Rapidan River	RAP006.53	Culpeper	Northern Piedmont	Forest	This station is used as one of the reference sites.
Stony Creek	SNC000.20	Giles	Central Appalachian Ridges and Valleys	Forest	Downstream of Eastern Ridge Lime.

Table B.6 Data collected for the validation data set

STATION	RIVERMILE	Obs RBP	Obs EPTI	Pred RBP	Pred EPTI	TSS (mg/L)	EMBED	FLOW	DO (mg/L)	TEMP (^{o}C)	TURB (FTU)
Beaver Creek (BRC)	BRC001.88	0.50	6.75	1.28	4.34	4.00	17.50	17.00	11.48	12.70	8.28
Crab Creek (04)	CBC004.38	2.25	1.75	3.47	2.04	10.38	9.50	13.00	10.43	15.93	7.59
Crab Creek (06)	CBC006.35	2.00	3.60	3.38	1.83	11.20	9.60	13.20	10.78	14.96	7.59
James River	JMS326.30	1.50	7.00	1.60	6.39	5.50	15.50	18.00	9.85	17.30	3.50
Little Otter River	LOR014.33	0.00	6.00	2.57	3.80	10.00	9.50	15.50	9.98	11.15	15.28
Mechums River	MCM018.92	2.00	9.00	1.51	4.73	3.00	14.00	20.00	13.00	13.20	6.20
Peak Creek (09)	PKC009.29	2.00	5.00	2.26	6.67	5.00	10.00	17.50	9.33	14.10	6.00
Peak Creek (11)	PKC011.11	0.67	6.00	1.38	7.01	3.17	16.67	16.67	9.87	12.73	3.97
Rapidan River	RAP006.53	0.00	7.40	1.43	4.27	6.40	17.00	18.80	10.44	21.90	9.46
Stony Creek	SNC000.20	0.00	7.00	0.43	6.14	5.50	20.00	18.00	10.40	11.90	6.50

APPENDIX C Statistical analysis results

Table C.1 Summary of the forward multiple regression printouts in SAS for the data set that investigated WQI as a sediment indicator

ALL			
Step	Variable Entered	Partial R-square	Model R-square
1	% Grassland	0.0915	0.0915
2	Agricultural Area	0.0767	0.1682
3	Watershed Area	0.0355	0.2038
ALL AVERAGE			
Step	Variable Entered	Partial R-square	Model R-square
1	% Grassland	0.1200	0.1200
2	Agricultural Area	0.1279	0.2479
3	Watershed Area	0.0746	0.3224
FALL			
Step	Variable Entered	Partial R-square	Model R-square
1	Stream Length	0.1034	0.1034
2	% Pasture	0.0443	0.1477
3	Agricultural Area	0.0209	0.1686
4	Watershed Area	0.0273	0.1959
FALL AVERAGE			
Step	Variable Entered	Partial R-square	Model R-square
1	Stream Length	0.1784	0.1784
SPRING			
Step	Variable Entered	Partial R-square	Model R-square
1	% Grassland	0.2091	0.2091
SPRING AVERAGE			
Step	Variable Entered	Partial R-square	Model R-square
1	% Grassland	0.2219	0.2219
2	Watershed Area	0.0320	0.2539
3	Agricultural Area	0.1126	0.3665
4	% Pasture	0.0343	0.4007
5	Stream Length	0.0422	0.4430

Table C.2 Summary of stepwise multiple regression relationships for the data set that investigated TSS and other stressor variables

LINEAR			NON-LINEAR	
All (N=105)				
RBP	Hard, -Alk, -Order, Urban	0.5143	-Cropland, Urban, Do*temp, Beef, -Sq vel, -Sq order	0.62
TAXA	-TSS, Graze, -DO, Order, -Urban	0.4079	Log ripveg, Sq flow, -Sq do, - Sq(alk*hard)	0.4568
MFBI	-Vel, DO, -Order, Pasture, Urban	0.6168	DO, Log urban, -Sq subs, - Sq vel, Sq hard, -Sq alk, -Sq(alk*hard)	0.6956
DOMPCT	-Subs, -Flow, Hard	0.3596	-Log subs, -Log wshed area, Sq subs, -Sq flow, -Sq temp, Sq hard	0.5167
EPTI	-TSS, Cover, Subs, Flow, -DO, -Temp, pH, -Wshed area, -Urban	0.5712	Flow, DO, -TSS*ag, Log subs, -Log temp, Log pH, -Log cond, Sq cover, -Sq do	0.656
S-T	Flow, -Cond	0.0818	-Log urban	0.0696
All Avg (N=34)				
RBP	-Vel, -Order, Urban	0.6543	-Log flow, Log(do*temp), -Sq embed	0.6401
TAXA	Riffles, Graze, -DO	0.4586	Riffles, Log graze, -Sq do	0.4804
MFBI	-Flow, Hard, -Alk	0.7056	-Alk*hard, -Log flow, Sq hard	0.7442
DOMPCT	-Flow, Hard, -Wshed area	0.4888	-Log flow, -Log wshed area, Sq sediment, Sq hard, -Sq cond	0.7111
EPTI	Embed, Flow, -DO, -Temp, -Turb	0.6565	-Log temp, -Log(tss*turb), Sq flow, -Sq do	0.7003
S-T	-Cond	0.1392	-Log(do*temp), Sq temp	0.3362
Fall (N=72)				
RBP	-Sediment, Hard, -Cropland, Urban	0.5198	Hard, -Alk, -Log pasture, -Sq embed	0.5135
TAXA	Riffles, Graze, Order	0.4432	Log order, -Sq vel, Sq riffles, Sq graze	0.5065
MFBI	-Subs, -Flow, DO, -Forest, Urban	0.6052	Log urban, -Sq subs, Sq hard, -Sq alk	0.6373
DOMPCT	-Subs, -Order	0.39	-Subs, -Flow, -Log order, Sq subs, Sq cond	0.6034
EPTI	Embed, Riffles, -Urban	0.3413	Embed, Log riffles, -Log urban	0.3695
S-T	Flow, -Cond	0.1689	-Cond, Sq flow	0.1748
Fall Avg (N=33)				
RBP	-Embed, -Flow, -pH, Cond, -Beef	0.7314	-Log pH, -Log order, Log urban, -Sq cover, -Sq embed, Sq banks	0.8107
TAXA	Ripveg	0.3261	Log ripveg, -Sq do	0.4639
MFBI	-Flow, Hard, -Alk	0.6707	-Alk*hard, -Log flow, Sq hard	0.7482
DOMPCT	-Ripveg	0.3267	Tss*ag, -Log cover, -Log order, -Sq wshed area, Sq order	0.748
EPTI	Riffles, Flow, -Turb	0.5148	-Flow, Turb, Log riffles, -Log cond, -Log wshed area, -Log(do*temp), -Log(tss*turb), Log(tss*ag)	0.984
S-T	-Cond	0.1538	-Cond	0.1538

Table C.2 (continued) Summary of stepwise multiple regression relationships for the data set that investigated TSS and other stressor variables

LINEAR			NON-LINEAR	
Spring (N=33)				
RBP	-Vel	0.4553	Log urban, -Sq vel	0.5699
TAXA	-Tss, Subs, Pasture	0.5287	-Sq(tss*ag)	0.3909
MFBI	-Flow, DO, Temp, Cond, Length, -Cropland	0.8018	Cond, -Log flow, -Sq cropland, Sq(do*temp)	0.7641
DOMPCT	Alk	0.1718	Sq(alk*hard)	0.2151
EPTI	-Tss, Subs, Flow, -DO, pH, -Beef, -Urban	0.7724	-Log urban, -Sq(do*temp)	0.4441
S-T	-TSS, Sediment, -Ripveg	0.507	-Ripveg, -Sq(do*temp)	0.5049
Spring Avg (N=24)				
RBP	pH	0.5676	Log pH, -Sq vel	0.6578
TAXA	-Tss, -DO, Turb	0.6451	-Sq(do*temp)	0.4813
MFBI	DO, pH, Cond	0.8257	Cond, Sq do, Sq pH	0.8293
DOMPCT	Alk	0.3046	Log vel, Sq(alk*hard)	0.5399
EPTI	-DO, Forest	0.6409	-Urban, -Sq beef, -Sq(do*temp), -Sq(tss*bankveg)	0.8073
S-T	-Tss, Sediment, -Ripveg, Turb	0.7316	-Ripveg, -Sq(do*temp)	0.6105
1997 (N=26)				
RBP	-Vel, Urban	0.5597	Vel, Log(do*temp), -Sq vel, Sq banks	0.7832
TAXA	Subs, -DO	0.5935	Subs, -DO	0.5935
MFBI	-Subs, -Flow, Cond	0.7943	Cond, -Log subs, -Log flow	0.802
DOMPCT	-Subs, Vel	0.5188	Log vel, -Log order, -Sq alk, Sq(alk*hard)	0.7706
EPTI	-Tss, -DO, Forest	0.778	Forest, -Sq tss, -Sq do	0.7777
S-T	-DO, -Temp, pH, -Cond, Beef	0.8267	-Cond, -Do*temp, Log pH, Sq beef	0.8172
1998 (N=28)				
RBP	-Embed	0.2934	Log urban, -Sq embed, -Sq flow, Sq(do*temp)	0.6461
TAXA	Riffles	0.3168	Log ripveg, Log pH	0.4572
MFBI	-Subs, -Flow, DO, Pasture, Urban	0.8597	Urban, -Log subs, -Sq flow, Sq do, Sq ag	0.8781
DOMPCT	Sediment, -Riffles, -Order	0.7459	Log vel, -Log riffles, -Log order, Sq order	0.8408
EPTI	Riffles, -Urban	0.4261	-Urban, Log riffles	0.4723
S-T	Subs	0.1494	Sq subs	0.1666
1999 (N=25)				
RBP	-Riffles, Hard, -Alk, -Ag	0.8979	-Length, -Log vel, Log hard, -Log pasture	0.8869
TAXA	Sediment, Temp, Beef, Order	0.8033	Beef, Sq temp, Sq order	0.7784
MFBI	-Vel, Urban	0.6055	-Vel, Urban	0.6055
DOMPCT	-Cover	0.297	-Log wshed area, -Sq cover	0.473
EPTI	Riffles, -Cond	0.5067	-Log cond, Sq riffles, Sq(do*temp)	0.6854
S-T	NONE	0	-Log urban, Sq temp	0.3617

Table C.3 Summary of the stepwise multiple regression printout in SAS for the linear, “all average”, RBP relationship developed from the data set used to investigate TSS and other stressors*

Step	Variable Entered	Partial R-square	Model R-square
1	Velocity	0.4798	0.4798
2	% Urban	0.1019	0.5817
3	Stream Order	0.0726	0.6543

* No variables were removed during the stepwise procedure

Table C.4 Summary of the stepwise multiple regression printout in SAS for the non-linear, “all average”, RBP relationship developed from the data set used to investigate TSS and other stressors*

Step	Variable Entered	Partial R-square	Model R-square
1	$(\text{Embeddedness})^2$	0.4955	0.4955
2	Log (Flow)	0.0894	0.5849
3	Log (DOxTemperature)	0.0553	0.6401

* No variables were removed during the stepwise procedure

Table C.5 Summary of the stepwise multiple regression printout in SAS for the linear, “all average”, EPTI relationship developed from the data set used to investigate TSS and other stressors*

Step	Variable Entered	Partial R-square	Model R-square
1	Turbidity	0.3486	0.3486
2	Flow	0.1143	0.4629
3	DO	0.0692	0.5320
4	Temperature	0.0658	0.5978
5	Embeddedness	0.0587	0.6565

* No variables were removed during the stepwise procedure

Table C.6 Summary of the stepwise multiple regression printout in SAS for the non-linear, “all average”, EPTI relationship developed from the data set used to investigate TSS and other stressors*

Step	Variable Entered	Partial R-square	Model R-square
1	Log(TSSxTurbidity)	0.3699	0.3699
2	$(\text{Flow})^2$	0.1381	0.5080
3	$(\text{DO})^2$	0.0983	0.6063
4	Log(Temperature)	0.0940	0.7003

* No variables were removed during the stepwise procedure

Table C.7 Summary of the stepwise multiple regression printout in SAS for the linear, “spring average”, MFBI relationship developed from the data set used to investigate TSS and other stressors*

Step	Variable Entered	Partial R-square	Model R-square
1	pH	0.5923	0.5923
2	DO	0.1932	0.7855
3	Conductivity	0.0402	0.8257

* No variables were removed during the stepwise procedure

Table C.8 Summary of the stepwise multiple regression printout in SAS for the linear, “1998”, MFBI relationship developed from the data set used to investigate TSS and other stressors*

Step	Variable Entered	Partial R-square	Model R-square
1	Substrate	0.5002	0.5002
2	% Urban	0.1067	0.6069
3	DO	0.1142	0.7211
4	% Pasture	0.0579	0.7790
5	Flow	0.0807	0.8597

* No variables were removed during the stepwise procedure

Table C.9 Summary of the stepwise multiple regression printout in SAS for the non-linear, “spring average”, MFBI relationship developed from the data set used to investigate TSS and other stressors*

Step	Variable Entered	Partial R-square	Model R-square
1	$(\mathrm{pH})^2$	0.5977	0.5977
2	$(\mathrm{DO})^2$	0.1935	0.7913
3	Conductivity	0.0381	0.8293

* No variables were removed during the stepwise procedure

Table C.10 Summary of the stepwise multiple regression printout in SAS for the non-linear, “1998”, MFBI relationship developed from the data set used to investigate TSS and other stressors*

Step	Variable Entered	Partial R-square	Model R-square
1	Log(Substrate)	0.5264	0.5264
2	% Urban	0.1016	0.6281
3	$(\mathrm{DO})^2$	0.1110	0.7391
4	$(\mathrm{Flow})^2$	0.0601	0.7992
5	$(\%\ \mathrm{Agriculture})^2$	0.0789	0.8781

* No variables were removed during the stepwise procedure

Table C.11 The R^2 analysis relationships versus the original stepwise multiple regression relationships

		All[a]				All Avg[a]		
		RBP	TAXA	EPTI	MFBI	RBP	EPTI	MFBI
R^2 Analysis Relationships	Variables Included in R^2 Analysis[b,c]	-Vel, -Order, Urban	Riffles, Graze, -DO, Order, -Urban	Cover, Subs, Flow, -DO, -Temp, pH, -Wshed area, -Pasture, -Urban	-Subs, DO, Hard, -Alk	-Vel, -Order, Urban	Flow, -DO, -Turb	-Vel, -Flow, DO, Ag, Urban
	MSE	0.50	8.25	2.12	0.35	0.28	2.15	0.19
	Cp	15.51	10.67	10.61	6.0599	8.64	4.25	7.52
	R^2 (R^2 Analysis)	0.51	**0.41**[d]	**0.57**[d]	0.60	0.65	0.53	**0.80**[d]
Original Stepwise Relationships[e]	R^2 (Stepwise)	0.5143	0.4079	0.5712	0.6168	0.6543	0.6565	0.7056
	Cp	15.78	11.00	10.95	2.92	8.64	0.20	17.87
	MSE	0.50	8.27	2.13	0.33	0.28	1.69	0.26
	Variables Included in Stepwise[b,c]	Hard, -Alk, -Order, Urban	-TSS, Graze, -DO, Order, -Urban	-TSS, Cover, Subs, Flow, -DO, -Temp, pH, -Wshed area, -Urban	-Vel, DO, -Order, Pasture, Urban	-Vel, -Order, Urban	Embed, Flow, -DO, -Temp, -Turb	-Flow, Hard, -Alk
		RBP	TAXA	EPTI	MFBI	RBP	EPTI	MFBI
		All[a]				All Avg[a]		

[a] One of the nine subsets originally investigated (all, all average, fall, fall average, spring, spring average, 1997, 1998, and 1999) as discussed in Section 4.2.1.

[b] Refer to Table 4.4 for a list of the abbreviations for the independent variables.

[c] A negative sign implies that the variable has a negative (indirect) association with the RBP/metric values; no sign implies that the variable has a positive (direct) association with the RBP/metric values.

[d] The bolded R^2 (R^2 analysis) values imply the R^2 analysis stepwise relationship produced a higher R^2 value than the original stepwise relationship.

[e] Selected stepwise relationships previously presented in Tables 4.5, 4.7, and 4.9.

BIBLIOGRAPHY

Allan, J.D. 1995. "Stream Ecology: Structure and Function of Running Waters." Chapman & Hall, London.

Al-Smadi, M., K. Brannan, T.A. Dillaha, III, C. Heatwole, J. Miller, S. Mostaghimi, S. Shah, M.L. Wolfe, and G. Yagow. 2000. Fecal coliform TMDL for Sheep Creek, Elk Creek, Machine Creek, Little Otter River, and Big Otter River in Bedford and Campbell counties, Virginia. Department of Biological Systems Engineering, Blacksburg, VA.

Aronoff, S. 1993. "Geographic Information Systems: A Management Perspective," 3rd ed. WDL Publications, Ontario.

Barfield, B.J., R.C. Warner, and C.T. Haan. 1983. Applied hydrology and sedimentology for disturbed areas. Oklahoma Technical Press, Stillwater, Oklahoma.

Barbour, M.T., J. Gerritsen, B.D. Snyder and J.B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water. Washington D.C.

Basnyat, P., L.D. Teeter, K.M. Flynn, and B.G. Lockaby. 1999. Relationships between landuse characteristics and nonpoint source pollution inputs to coastal estuaries. Environmental Management 23(4): 539-549.

Bjornn, T.C., M.A. Brusven, M. Molnau, F.J. Watts, and R.L. Wallace. 1974. Sediment in streams and its effect on aquatic life. Research Paper B-025-IDA. University of Idaho Water Resources Research Institute. Moscow, Idaho.

Bjornn, T.C., M.A. Brusven, M. Molnau, F.J. Watts, and R.L. Wallace. 1977. Transport of granitic sediment in streams and its effects on insects and fish. Technical Report 17. University of Idaho College of Forestry, Wildlife and Range Sciences. Moscow, Idaho.

Booth, D.B. and C.R. Jackson. 1997. Urbanization of aquatic systems: degradation thresholds, stormwater detection, and the limits of mitigation. Journal of the American Water Resources Association 33(5): 1077-1090.

Burrough, P.A. and R.A. McDonnel. 1998. "Principle of Geographical Information Systems." Oxford University Press. New York.

Burton, G.A., Jr. and K.J. Scott. 1992. Sediment toxicity evaluations: their niche in ecological assessments. Environ. Sci. Technol. 26: 2068-2075.

Chanat, J. 1998. Abiotic Factors Affecting Benthic Macroinvertebrate Biomass and Community Structure in a Fourth-Order Rocky Mountain Watershed. M.S. Thesis, Colorado State University.

Chapman, P.M. 1990. The sediment quality triad approach to determining pollution-induced degradation. Sci. Tox. Environ. 97(8): 815-825.

Chapman, P.M. 1995. Do sediment toxicity tests require field validation? Environ. Toxicol. Chem. 14(9): 1451-1453.

Chapman, P.M., E.A. Power, and G. A. Burton, Jr. 1992. Integrative assessments in aquatic ecosystems. Chpt. 14 in "Sediment Toxicity Assessment," ed. G.A. Burton Jr., pp. 313-340. Lewis Publishers, Chelsea, MI.

Clarke, K.D. and D.A. Scruton. 1997. The benthic community of stream riffles in Newfoundland, Canada and its relationship to selected physical and chemical parameters. Journal of Freshwater Ecology 12(1): 113-121.

Copeland, C. 1999. Clean Water Act: a summary of the law. RL 30030. Congressional Research Service; Environment and Natural Resources Policy Division. Available online by The Committee for the National Institute for the Environment at: http://www.cnie.org/nle/h2o-32.html.

Cowen, D.J. 1998. GIS versus CAD versus DBMS: what are the differences. Photogrammetric Engineering and Remote Sensing (PE&RS) 54(11): 1551-1555.

DEP, Department of Environmental Protection for Montgomery County, MD. 2000. All about stream bugs. Available online at: http://www.co.mo.md.us/dep/Watershed/Biomon/Bugs/bug_main.htm

Darkwah, G.A. 2000. DEQ Biomon 94 fall to 98 fall. Electronic Mail Transmittal: BIOMON 94F-98F.ZIP.

DSWC, Division of Soil and Water Conservation. 1989. "Virginia Nonpoint Source Pollution Management Program." Virginia Department of Conservation and Recreation; Division of Soil and Water Conservation. Richmond, Virginia.

Dunne, T. and L.B. Leopold. 1978. "Water in Environmental Planning." W.H. Freeman & Co., New York.

EDC, Earth Resources Observation Systems (EROS) Data Center. 1997. Multi-Resolution Land Characteristics Consortium: Federal Region III Land Cover Data Set. Sioux Falls, SD.

ESRI, Environmental Systems Research Institute, Inc. 1998. ArcVIEW GIS Version 3.1. Redlands, CA.

ESRI, Environmental Systems Research Institute, Inc. 1999. U.S. Census TIGER (Topologically Integrated Geographic Encoding and Referencing) 1995 Data. Available online at: http://www.esri.com/data/online/tiger/

Evans, S.K. 1997. Studies of Benthic Macroinvertebrates in Western Virginia Streams as Related to the Implementation of Rapid Bioassessment Techniques. Ph.D. thesis, Virginia Polytechnic Institute and State University.

Fisher, G.T. 1989. Geographic information system / watershed model interface. Hydraulic Engineering '89 Proceedings. American Society of Civil Engineers (ASCE). Pp. 851-856.

Fore, L.S., J.R. Karr, and R.W. Wisseman. 1996. Assessing invertebrate responses to Human activities: evaluating alternative approaches. Journal of the North American Benthological Society 15(2): 212-231.

Frissell, C.A., W.J. Liss, C.E. Warren, and M.D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10(2): 199-214.

Hannaford, M.J. and V.H. Resh. 1995. Variability in macroinvertebrate rapid-bioassessment surveys and habitat assessments in a northern California stream. Journal of the North American Benthological Society 14(3): 430-439.

Heatwole, C.D., V.O. Shanholtz, E.R. Yagow, and E.R. Collins, Jr. 1987. Targeting animal waste pollution with Virginia GIS. Paper No. 87-4049. American Society Agricultural Engineers, St. Joseph, MI.

Hem, J.D. 1992. Study and interpretation of the chemical characteristics of natural water. Water-Supply Paper 2254, U.S. Geological Survey, Washington, D.C.

Hun, T. 1998. EPA provides status update at packed technical session. Water Environment & Technology 10(12): 36-37.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6:21-27.

Karr, J.R., K.D. Fausch, P.L. Angermeier, P.R. Yant, and I.J. Schlosser. 1986. Assessing biological integrity in running waters – a method and its rationale. Illinois Natural History Survey. Special Publication 5.

Kerans, B.L. and J.R. Karr. 1994 Development and testing of a benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications 4(4).

Kleene, J.W. 1995. Watershed Nonpoint Source Management System: A Geographic Information System Approach. Ph.D. thesis, Virginia Polytechnic Institute and State University.

Knighton, D. 1984. "Fluvial Forms and Processes." Edward Arnold, London.

Lenat, D.R., D.L. Penrose, and K.W. Eagleson. 1981. Variable effects of sediment addition on stream benthos. Hydrobiologia 79: 187-194.

MDE, Maryland Department of the Environment. 2001. Introduction to Maryland's TMDL program. Available online at: http://www.mde.state.md.us/tmdl/index.html.

MIDEQ, Michigan Department of Environmental Quality. 2000. Surface Water Quality Division Permits Section; Total Suspended Solids. Available online at: http://www.deq.state.mi.us/swq/permits/parameters/tss.htm

Minshall, G.W. 1984. Aquatic insect-substratum relationships. Chpt. 12 in "The Ecology of Aquatic Insects," ed. V.H. Resh and D.M. Rosenberg, pp. 358-400. Praeger, New York.

Murphy, S. 2000. General information on solids. BASINS, Boulder Area Sustainability Information Network. Available online at: http://bcn.boulder.co.us/basin/data/FECAL/info/TSS.html

Nelson, R.L., W.S. Platts, D.P. Larsen, and S.E. Jensen. 1992. Trout distribution and habitat relation to geology and geomorphology in the North Fork Humboldt River drainage, northeastern Nevada. Transactions of the American Fisheries Society 121(4): 405-426.

NJDEP, New Jersey Department of Environmental Protection. 2000. Rapid bioassessment protocol. Available online at: http://www.state.nj.us/dep/watershedmgt/bfbm/rbpinfo.html.

Novotny, V. and H. Olem. 1994. "Water Quality: Prevention, Identification, and Management of Diffuse Pollution." Van Nostrand Reinhold, New York.

OEPA, Ohio Environmental Protection Agency. 1989. Biological criteria for the protection of aquatic life: volume III. Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Division of Water Quality Monitoring and Assessment, Columbus, Ohio.

Oschwald, W.R. 1972. Sediment-water interactions. Journal of Environmental Quality 1: 360-366.

Ott, R.L. 1993. "An Introduction to Statistical Methods and Data Analysis." 4th Edition. Wadsworth, Belmont, CA.

Pawlak, B. 1999. The basics of bioassessment: reference conditions. Virginia Lakes & Watersheds Association Newsletter 48: 8-9.

Pelley, J. 1998. The challenge of watershed cleanup. Environmental Science & Technology 32(15): 364A-365A.

Perciasepe, R. 1997. Memorandum on new policies for establishing and implementing Total Maximum Daily Loads. Former EPA Assistant Administrator. As cited in Virginia Water Central (October 1998). Virginia Water Resources Research Center. Blacksburg, VA.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers. Benthic Macroinvertebrates and Fish. EPA 440-4-89-001. U.S. Environmental Protection Agency; Office of Water Regulations and Standards. Washington, D.C.

Resh, V.H., R.H. Norris, and M.T. Barbour. 1995. Design and implementation of rapid assessment approaches for water resource monitoring using benthic macroinvertebrates. Australian Journal of Ecology 20: 108-121.

Richards, C. and K.L. Bacon. 1994. Influence of fine sediment on macroinvertebrate colonization of surface and hyporheic stream substrates. Great Basin Naturalists 54(2): 106-113.

Richards, C. and G. Host. 1994. Examining land use influences on stream habitats and macroinvertebrates: a GIS approach. Water Resources Bulletin 30(4): 729-739.

Richman, M. 1997. Lawsuits imply states missing the boat on developing lists of impaired waters. Water Environment & Technology 9(1): 18-20.

Ritchie, J.C. 1972. Sediment, fish, and fish habitat. Journal of Soil and Water Conservation 27: 124-125.

Ritter, D.F., R.C. Kocel, and J.R. Miller. 1995. "Process Geomorphology." Wm. C. Brown, Dubuque, IA.

Rosenberg, D.M. and V.R. Resh. 1993. "Freshwater Biomonitoring and Benthic Macroinvertebrates." Chapman and Hall, New York.

Rothrock, J.A., P.K. Barten, and G.L. Ingman. 1998. Land use and aquatic biointegrity in the Blackfoot River watershed, Montana. Journal of the American Water Resources Association 34(3): 565-583.

SAS. 1999. SAS system for Windows Release 8.01. SAS Institute, Inc. Cary, NC.

Schwab, G.O., D.D. Fangmeier, W.J. Elliot, and R.K. Frevert. 1993. "Soil and Water Conservation Engineering," 4th ed. John Wiley & Sons, New York.

Shanholtz, V.O. 1988. Delivery ratio for targeting potentially highly erodible lands. Unpublished Report, Virginia Tech Department of Agricultural Engineering.

Shanholtz, V.O., J.M. Flagg, C. Metz, and C. Desai. 1990. VirGIS: Agricultural pollution potential database for Mountain Castles (Botetourt County) soil and water conservation district. Interim Report ISSL 90-6. Information Support Systems Laboratory, Blacksburg, VA.

Shaver, M.L., J.P. Shannon, K.P. Wilson, P.L. Benenati, and D.W. Blinn. 1997. Effects of suspended sediment and desiccation on the benthic tailwater community in the Colorado River, USA. Hydrobiologia 357: 63-72.

Southerland, M.T. and J.B. Stribling. 1995. Status of biological criteria development and implementation. In "Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making," ed. W.S. Davis and T.P. Simon, pp. 81-96. Lewis Publishers, Boca Raton, FL.

Smith, E.P. 2001. Personal communication. Department of Statistics, Virginia Polytechnic Institute and State University. Blacksburg, Virginia.

StatSoft, Inc. 2001. STATISTICA. Available online at: http://www.statsoft.com/textbook/stathome.html

USEPA, U.S. Environmental Protection Agency. 1990. The Quality of our Nation's Water: A Summary of the 1988 National Water Quality Inventory. EPA Report 440/4-90-005. Washington, DC.

USEPA, U.S. Environmental Protection Agency. 1991. Guidance for Water Quality-Based Decisions: The TMDL Process. Office of Water Regulations and Standards. Washington, D.C.

USEPA, U.S. Environmental Protection Agency. 1992a. National Water Quality Inventory: 1990 Report To Congress. EPA 503/9-92/006. Office of Water. Washington, D.C.

USEPA, U.S. Environmental Protection Agency. 1992b. Framework for Ecological Risk Assessment. EPA Report 630/R-92/001. Risk Assessment Forum. Washington, D.C.

USEPA, U.S. Environmental Protection Agency. 1996a. Biological Criteria: Technical Guidance for Streams and Small Rivers, Revised Ed., EPA 822/B-96-001. Office of Water. Washington, D.C.

USEPA, U.S. Environmental Protection Agency. 1996b. Summary of State Biological Assessment Programs for Streams and Rivers. EPA Report 230-R-96-007. Office of Water. Washington, D.C.

USEPA, U.S. Environmental Protection Agency. 1999. What is NPS Pollution. Available online at: http://www.epa.gov/OWOW/NPS

USEPA, U.S. Environmental Protection Agency. 2000a. Overview of current Total Maximum Daily Load – TMDL – program and regulations. EPA841-F-00-009. Available online at: http://www.epa.gov/owow/tmdl/overviewfs.html

USEPA, U.S. Environmental Protection Agency. 2000b. 1998 Section 303(d) List fact sheet: Virginia. Available online at: http://www.epa.gov/owow/tmdl/states/vafact.html

USEPA, U.S. Environmental Protection Agency. 2000c. Stressor Identification Guidance Document. EPA 822-B-00-025. Office of Water and Office of Research and Development. Washington, D.C.

VADCR, Virginia Department of Conservation & Recreation. 1996. Hydrologic Unit Animal Census Database.

VADEQ, Virginia Department of Environmental Quality. 1997. Virginia Ambient Water Quality Monitoring Report, Monitoring Year 1997. DEQ Technical Bulletin WQA/1997-003. Water Quality Assessments. Richmond, VA.

VADEQ, Virginia Department of Environmental Quality. 1998. Virginia Ambient Water Quality Monitoring Report, Monitoring Year 1998. DEQ Technical Bulletin WQA/1998-003. Water Quality Assessments. Richmond, VA.

VADEQ, Virginia Department of Environmental Quality. 1999. Virginia Ambient Water Quality Monitoring Report, Monitoring Year 1997. DEQ Technical Bulletin WQA/1999-004. Water Quality Assessments. Richmond, VA.

Voshell, J.R. 2001. Personal communication. Department of Entomology, Virginia Polytechnic Institute and State University. Blacksburg, Virginia.

Ward, J.V. 1985. Thermal characteristics of running waters. Hydrobiologia 125: 31-46.

Waters, T.F. 1995. Sediment in streams: sources, biological effects, and control. American Fisheries Society Monograph 7.

Watzin, M.C., and A.W. McIntosh. 1999. Aquatic ecosystems in agricultural landscapes: a review of ecological indicators and achievable ecological outcomes. Journal of Soil and Water Conservation 54(4): 636-644.

Webb, B.W. and D.E. Walling. 1992. Water quality II: chemical characteristics. In “The Rivers Handbook,” ed. P.Calow and G.E. Petts. Blackwell Science Ltd., Oxford.

Williams, J.R. and H.D. Berndt. 1972. Sediment yield computed with universal equation. Journal of Hydraulics Division. American Society of Civil Engineers. Proc. Paper 9426. 98(HY12): 2087-2098.

Wood, P.J. and P.D. Armitage. 1997. Biological effects of fine sediment in the lotic environment. Environmental Management 21(2): 203-217.

Ye, K. 1999. Personal communication. Department of Statistics, Virginia Polytechnic Institute and State University. Blacksburg, Virginia.

Younos, T., M.S. Rosenzweig, and J.L. Walker. 2000. Stream health: relating stream biota to stream water quality. Presented in “NWQMC National Monitoring Conference,” NWQMC, pp. 435-446. Austin, TX.

Zanetell, B.A., and B.L. Peckarsky. 1996. Stoneflies as ecological engineers – hungry predators reduce fine sediments in stream beds. Freshwater Biology 36: 569-577.

VITA

Laurie Frondorf

Laurie Michelle Frondorf, the youngest of three daughters of Kirk and Diane Frondorf, was born in Richmond, Virginia on December 15, 1977. After graduating from Monacan High School in 1995, she earned her B.S. in Biological Systems Engineering at Virginia Polytechnic Institute and State University in 1999. She then continued her education at Virginia Polytechnic Institute and State University by pursuing her M.S in Biological Systems Engineering. Upon completing her M.S. in 2001, she began her career with USInfrastructure, Inc. in Atlanta, Georgia.