WASHINGTON STATE
DEPARTMENT OF
ECOLOGY

Factors Affecting Waters with a High pH

Statewide Analysis

February 2002

Publication No. 02-03-005
printed on recycled paper

This report is available on the Department of Ecology home page on the World Wide Web at http://www.ecy.wa.gov/biblio/0203005.html

For additional copies of this publication, please contact:

Department of Ecology Publications Distributions Office
Address: PO Box 47600, Olympia WA 98504-7600
E-mail: ecypub@ecy.wa.gov
Phone: (360) 407-7472

Refer to Publication Number 02-03-005

Factors Affecting Waters with a High pH

Statewide Analysis

by
Steve Butkus

Environmental Assessment Program
Olympia, Washington 98504-7710

February 2002

Publication No. 02-03-005
printed on recycled paper

Table of Contents

List of Figures and Tables

Figures

Tables

Abstract

An analysis was conducted to evaluate factors that may result in violations of the state's water quality standard for pH. Washington State has a number of waters that are impaired with high pH. These impairments are likely caused by the introduction of excessive nutrients due to human activities. As a result of measurements made that show the pH criterion is exceeded, 26 streams are included on the Washington 1998 Section 303(d) list due to nonpoint sources alone. Section 303(d) of the federal Clean Water Act mandates that the state establish Total Maximum Daily Loads (TMDLs) for surface waters that do not meet standards after application of technology-based pollution controls. The analysis is being conducted to provide information for future TMDL development. The analysis conducted in this report uses various statistical tests to evaluate the relationships between pH and other measured water quality variables. A multivariate empirical model that predicts pH from temperature, flow, and phosphorus was developed. Although the model could be constructed using surrogate variables significantly related ($p < 0.001$) to pH, only 17% of the variation in the data was accounted for. There are other factors that influence pH not considered in this analysis. Therefore, because of the large unexplained variance, the empirical model formulated is not recommended for application in a TMDL.

Introduction

The process of photosynthesis by aquatic plants and algae can affect stream pH. During the Spring and Summer growing period, photosynthesis consumes carbon dioxide that is dissolved in these streams in the form of carbonic acid. The consumption of the carbon dioxide shifts the chemical equilibrium of the stream carbonate buffering system. The result of the photosynthesis liberates hydrogen ions which in turn raises the pH. Excessive photosynthesis can result in the pH of a stream being increased above the criterion set for protection of aquatic life.

Aquatic plant and algal growth depends on the nutrients of phosphorus and nitrogen for growth. The concentrations of these nutrients limit the amount of growth that can occur. Human activities often increase the loading of nutrients to surface waters causing excessive productive growth in a process called eutrophication. Increased cultural eutrophication upsets the ecological balance of surface waters. One of the impacts of excessive nutrient loading is increased pH level that can affect aquatic life.

Washington State has a number of waters that are impaired by high pH. These impairments are likely caused by the introduction of excessive nutrients due to human activities. Although some of these nutrients may be introduced as a result of wastewater discharges, the largest number of impacted waters appears to be caused by nonpoint sources of pollution. This analysis was conducted to evaluate factors that may result in violations of the state's water quality standard for pH.

Applicable Criteria

Within The State of Washington, water quality standards are published pursuant to Chapter 90.48 of the Revised Code of Washington (RCW). Authority to adopt rules, regulations, and standards as are necessary to protect the environment is vested with the Department of Ecology. Under the federal Clean Water Act, the EPA Regional Administrator must approve the water quality standards adopted by the State (Section 303(c)(3)). Through adoption of these water quality standards, Washington has designated certain characteristic uses to be protected and the criteria necessary to protect these uses [Washington Administrative Code (WAC), Chapter 173-201A). These standards were last adopted in November 1997.

The characteristic uses designated for protection in streams are as follows:

"Characteristic uses. Characteristic uses shall include, but not be limited to, the following:

(iii) Fish and shellfish:

Salmonid migration, rearing, spawning, and harvesting.

Other fish migration, rearing, spawning, and harvesting.

Clam and mussel rearing, spawning, and harvesting.

Crayfish rearing, spawning, and harvesting.

(iv) Wildlife habitat.

(v) Recreation (primary contact recreation, sport fishing, boating, and aesthetic enjoyment).

[WAC 173-201A-030(1)&(2)]

The water quality standards describe pH criteria for the protection of characteristic uses. The water quality limited streams for which this analysis applies are either Class AA or Class A. These classes have the same pH criterion to protect the characteristic uses

"pH should be within the range of 6.5 to 8.5 (freshwater) and 7.0 to 8.5 (marine water) with a human caused variation of within the range of 0.5 units."

[WAC 173-201A-030(1)(c)(v)] &
[WAC 173-201A-030(2)(c)(v)]

In cases where natural background conditions exceed a standard, the water quality standards state the following:

"Whenever the natural conditions of said waters are of a lower quality than the criteria assigned, the natural conditions shall constitute the water quality criteria."

[WAC 173-201A-070(2)]

Water Quality Impairments

As a result of measurements made that show the pH criterion is exceeded, 26 streams are included on the Washington 1998 Section 303(d) list due to nonpoint sources alone (Table 1). Streams listed for pH downstream of permitted point sources discharging phosphorus are not included in Table 1.

Table 1. Streams on the 1998 Section 303(d) list for High pH due to Nonpoint Sources

WRIA	Stream Name	pH Measurement Location			Maximum pH Measured
		Township	Range	Section	
8	Sammamish River	25N	05E	11	9.1
23	Scatter Creek	15N	03W	08	9.0
23	Skookumchuck River	14N	02W	07	8.8
28	Burnt Bridge Creek	02N	01E	38	9.0
28	Lacamas Creek	01N	03E	44	9.4
28	Lacamas Creek	02N	03E	10	8.8
32	Mill Creek	07N	36E	23	9.4
34	Rock Creek	18N	39E	05	9.4
34	Pine Creek	20N	42E	28	9.1
36	Esquatzel Coulee	10N	30E	08	8.9
36	Scooteney Wasteway	14N	30E	01	9.1
37	Moxee (Birchfield) Drain	13N	19E	16	10.7
38	Naches River	13N	18E	12	9.1
41	Winchester Wasteway	18N	27E	32	9.0
41	Frenchman Hills Wasteway	17N	27E	09	8.7
41	Lind Coulee	18N	29E	35	8.6
41	Sand Hollow Creek	17N	23E	27	8.8
47	Mitchell Creek	29N	21E	34	8.6
52	O'Brien Creek, S.F.	36N	33E	26	8.8
55	Little Spokane River	27N	43E	32	8.7
55	Deadman Creek	27N	43E	33	8.7
59	Mill Creek	36N	39E	31	8.7
59	Chewelah Creek, S.F.	33N	41E	23	8.8
60	Pierre Creek	40N	37E	32	8.8
61	Deep Creek	40N	40E	33	8.7
61	Smackout Creek	38N	41E	03	9.0

Section 303(d) of the federal Clean Water Act mandates that the State establish Total Maximum Daily Loads (TMDLs) for surface waters that do not meet standards after application of technology-based pollution controls. The U.S. Environmental Protection Agency (EPA) has promulgated new regulations (40 CFR 130) and developed guidance (EPA, 1991) for establishing TMDLs. The pace of TMDL development by the State of Washington is based on a schedule agreed upon in the recent settlement of a lawsuit (U.S. District Court, 1998).

Under the Clean Water Act, every state has its own water quality standards designed to protect, restore, and preserve water quality. Water quality standards consist of designated uses, such as cold water biota and drinking water supply, and criteria, usually numeric criteria, to achieve those uses. When a lake, river or stream fails to meet water quality standards after application of required technology-based controls, the Clean Water Act requires that the state place the water body on a list of "impaired" water bodies and to prepare an analysis called a Total Maximum Daily Load (TMDL).

The goal of a TMDL is to ensure the impaired water will attain water quality standards. A TMDL includes a written, quantitative assessment of water quality problems and of the pollutant sources causing the problem. The TMDL determines the amount of a given pollutant that can be discharged to the water body and still meet standards, the loading capacity, and allocates that load among the various sources. If the pollutant comes from a point source, that facility's share of the loading capacity is called a wasteload allocation. If it comes from a diffuse source such as a farm, that facility's share is called a load allocation. The TMDL must also consider seasonal variations and include a margin of safety that takes into account any lack of knowledge about the causes of the water quality problem or its loading capacity. The sum of the individual allocations and the margin of safety must be equal to or less than the loading capacity.

This analysis is being conducted to provide information for future TMDL development.

Analysis of Existing Data

The analysis conducted in this report uses various statistical tests to evaluate the relationships between pH and other measured water quality variables. The following is a list of the statistical tests used and their purpose, in order of their application.

Generation of Descriptive Statistics - The existing raw data of several variables were reduced to provide various distribution statistics. These statistics were derived to look at seasonal differences. The data reduction was conducted provide a consistent sample size and variance.

Testing for Normality - The seasonal descriptive statistics were tested for use in parametric tests. Both log transformed and non-transformed data were tested. The Kolmogrov-Smirnov goodness of fit test was applied but the results were rejected because of the high sample sizes. The standard error of both kurtosis and skewness were assessed to show normality and support the use of parametric statistical tests in the subsequent analyses.

Single Regression Analysis - The seasonal descriptive statistics of pH as the dependant variable were regressed against the seasonal descriptive statistics of each other possible independent variable for which data were available. The single regression results were used to cull variables that have no significant influence on pH from use in the subsequent analyses.

Analysis of Variance - A factorial design one-way analysis of variance was applied to evaluate differences between ecoregions and nutrient ratios. These tests were used to guide the stratification of variables into groups for the subsequent analyses.

Cluster Analysis - Standardized variables were clustered into groups using the hierarchical Ward method applied to evaluate differences between ecoregions. These tests were used to guide the stratification of variables into groups for the subsequent analyses.

Discriminant Analysis - A stepwise discriminant analysis was conducted on the results of the cluster analysis to evaluate which of the variables had the most influence on differentiating the identified clusters. These tests were used to guide the stratification of variables into groups for the subsequent analyses.

Multicollinearity Tests - The independent variables were assessed for inter-correlation to avoid ill-conditioning during the subsequent multiple regression analysis. Variables found to have significant multicollinearity were culled from use in the subsequent analyses.

Ordination Analysis - A principal components analysis was conducted to help explore which variables show important empirical relationships to pH in multidimensional space. Variables found not explain much variance were culled from use in the subsequent multivariate analysis.

Multiple Regression Analysis - A stepwise multiple linear regression was conducted with the variables selected as a result of the analyses conducted. The approach culled other variables resulting in a final multivariate model.

Data Reduction

Water quality data collected monthly between 10/1991 and 9/1998 for Ecology's Ambient Monitoring Program were used for this analysis. These data represented 222 collection stations. Sixty-two of these are "core" stations which have all 7 years of data. The remaining 160 stations have data for one or sometimes two complete years. Only verified data meeting a high level of quality assurance were used (Hallock and Ehinger, 2000). Estimated values below the quantitation limit were used. The detection limit values was used for those samples recorded as below the limit. The following water quality variables were used in the analysis: Temperature, flow, conductivity, dissolved oxygen, pH, Total Nitrogen (persulfate digestion), Total Inorganic Nitrogen (sum of Ammonia-N and Nitrate+Nitrite-N), Total Phosphorus, Ortho Phosphorus (soluble reactive P), Turbidity, and Total Suspended Solids.

These data were reduced to descriptive statistic variables for the analysis. Data were compiled into 3 separate seasonal periods to test the differences between them: the annual period (January - December), the Spring/Summer period (April - September), and the Fall/Winter period (October - March). The descriptive statistics derived for each station and water quality variable in these periods were sample size (count), maximum value, 90th percentile value, median, 10^{th} percentile value, and the minimum value (Appendix A).

In this analysis, the change in pH due to excessive productivity had to be assessed. Since few diel studies are available that actually measure this change, the distribution of pH data were used instead. Sampling the ambient stations generally occur at different times of the day. Sometimes a station is sampled early in the day where the pH does not reflect the algal productivity. Other times the station is sampled in the afternoon with the greatest productivity reflecting the highest pH values. It is assumed that the pH due to productivity can be estimated by comparing the range of pH values sampled at any one station.

Data on pH collected in throughout the state show no definite pattern of seasonal variation. Data collected by Ecology between October 1991 and September 1998 were compiled and descriptive statistics generated. Descriptive statistics are shown for pH measurements made on a monthly basis (Table 2). There is not much variation between the months with just a slight increase in overall pH values in the summer period likely due to higher algal productivity.

Table 2. Seasonal Variation of Statewide pH Descriptive Statistics

Month	Descriptive Statistics of Measured pH Values					
	Number of Measurements	Minimum	Median	Mean	Maximum	Range
January	506	6.3	7.4	7.5	8.8	2.5
February	537	6.3	7.5	7.6	8.8	2.5
March	555	6.4	7.6	7.7	9.0	2.6
April	572	6.5	7.7	7.8	9.1	2.6
May	573	6.6	7.7	7.7	9.7	3.1
June	576	6.3	7.7	7.7	9.2	2.9
July	553	6.4	7.8	7.8	9.6	3.2
August	554	6.7	7.8	7.9	9.5	2.8
September	568	6.7	7.8	7.8	9.7	3.0
October	653	6.5	7.7	7.8	9.4	2.9
November	645	6.2	7.6	7.7	9.4	3.2
December	565	6.4	7.5	7.6	9.5	3.1

Three separate pH ranges were derived for each station so that the best dependant variable for regression analysis could be assessed. In all three cases, the 90th percentile of the Spring/Summer period was used as the high pH value measured to represent algal productivity. The maximum pH values were not used to eliminate outliers that may be caused by the inaccurate measurement of pH sometimes encountered. Three different lower range pH values were derived as possibly representing conditions without significant productivity: the Fall/Winter period median pH, The Spring/Summer period 10th percentile pH, and the Spring/Summer period median pH.

Since the reduced data were to be used in parametric statistics, the assumption that the distributions were normal was tested for each variable. The data were base 10 logarithm and square root transformed for testing distribution normality. The Kolmogrov-Smirnov goodness of fit test was used to assess the distributions of both the untransformed and transformed data for normality (Zar, 1984). None of the distributions were determined to be significantly normal, regardless of the data transformation. The Kolmogrov-Smirnov test is likely not appropriate for these data since the critical values become diminishingly small at the high sample sizes used in this analysis.

Instead of using the Kolmogrov-Smirnov test for assessing distributions, the values of kurtosis (level of peakedness) and skewness (level of symmetry) and their respective standard error were used to determine whether data transformation would provide a more normal distribution. Based on this analysis, it was determined that no transformation was needed for temperature and dissolved oxygen data. The two dependant variables using the lower pH values of the Winter/Fall median and Spring/Summer 10th percentile best fit the square root transformation. All other variables best fit the base 10 logarithm transformation. These best fit distributions were used in all subsequent parametric analyses.

Variable Selection for Regression Analysis

To begin examining the influential relationships that control productivity-induced pH changes, regressions were derived with each of the descriptive statistics of the water quality variables for all three seasonal periods examined. The adjusted coefficient of determination (R^2) and the associated probability from F ratio of the analysis of variance table were compiled for each of these regressions (Appendix B). Relationships showing a significantly high confidence (at 95% and 99%) were identified for possible further use in other regression models. An analysis of the residuals was not conducted since this effort was only used to cull factors of importance for further investigation. From these single regressions, several interesting patterns and relationships were observed. From these patterns, a single dependant pH variable and several candidate independent variables were selected.

The potential dependent variable of the difference between the 90th percentile Spring/Summer pH and the 10th percentile Spring/Summer pH showed interesting relationships with the potential independent variables. Low winter temperatures and high conductivity were found to show a good relationship with this delta pH, most likely due to geographic patterns. This emphasizes a need to stratify the analysis according to regional characteristics. Several of the variables are not necessarily independent and represent a spurious correlation. High dissolved oxygen, high turbidity, and high suspended solids were correlated with delta pH. However, these high values are also likely products of high productivity. As such, they should not be used to predict pH changes since they are not causing the observed effect, but are a consequence of it.

The potential independent variable of the difference between the 90th percentile Spring/Summer pH and the Median Spring/Summer pH showed some additional relationships with the potential independent variables. Both the maximum values of total phosphorus and total nitrogen showed significant correlation's with the delta pH variable. This is expected because more nutrients will cause higher productivity affecting diel pH changes. Also, the minimum dissolved N:P ratio and minimum dissolved nitrogen correlated well indicating that the analysis may have to be stratified to account for the limiting nutrient.

The potential dependent variable of the difference between the 90th percentile Spring/Summer pH and the median Fall/Winter pH showed the most useful relationships with potential independent variables. For many of the variables, the correlation's were significant for both the extreme values (maximum or minimum) but also the 90th or 10th percentile values. This indicates that the data are more predictive since outlier values are not necessarily driving the regression results. There are also some additional independent variables that show significant correlation's to the delta pH variable. High temperatures are associated with high productivity by likely influencing growth kinetics. Low flows affect stream dynamics and mixing processes. Dissolved nutrients are more available for immediate algal growth than particulate forms of the total and directly affect productivity.

As a result of the inspection of these single regression results, specific variables were chosen to conduct more detailed analyses in developing a predictive empirical model. Due to the larger number of non-spurious correlation's, the difference between the 90th percentile Spring/Summer pH and the median Fall/Winter pH was chosen as the dependant variable. The 90th percentile

Spring/Summer pH represents the measurements related to the highest productivity. The median Fall/Winter pH represents the baseline pH values that are not affected by high productivity. The difference between these pH values represents the change in pH that is likely between due to enhanced productivity. Independent variables chosen for further empirical model testing were maximum total phosphorus, maximum ortho-phosphorus, maximum total nitrogen, maximum dissolved nitrogen, minimum dissolved N:P ratio, high temperature, and low flow for the annual period.

Stratification of Data for Analysis

During the investigation of variables for selection in regression analysis, patterns in the data suggested that relationships may be different depending on the regions of the state and nutrient limiting algal production. To assess whether these factors are significant, a factorial design one-way analysis of variance (ANOVA) was applied to each of the transformed variables selected for regression analysis. Two different groups were tested. For regional differences, all stations were stratified into the 8 ecoregions in Washington State defined by Omernik and Gallant (1986). For nutrient limitation, the stations were stratified based on the dissolved nitrogen to dissolved phosphorus ratio of 10 as defined by Thomann and Mueller (1987). This ratio is the point where nitrogen takes over the limitation of algal growth from phosphorus.

The ANOVA test was applied to all ecoregions combined for each variable. The results showed that there is a significant difference between the ecoregions for each variable except for flow To help distinguish which ecoregions are different from the others, an ANOVA using the Tukey method of pairwise means comparisons, was applied. Results show significant differences ($p < 0.05$) are generally observed between the Eastern Cascades Slopes and Foothills Ecoregion, and each of the Coast Range, Puget Lowland, and Cascades Ecoregions. (Table 3).

The ANOVA test was also applied to stations based on phosphorus or nitrogen algal growth limitation. The Bonferroni pairwise mean comparison method was used since the number of comparisons is small (e.g. either nitrogen limited or phosphorus limited). The results showed a significant difference between the stations for each of the three variables: total nitrogen, total phosphorus, and ortho-phosphorus.

These ANOVA tests verify that there are distinct strata within the data variables that may show different relationships in a regression analysis. In order to begin to stratify these variables into groups that can be regressed together, a cluster analysis was conducted of the median values of each ecoregion (Appendix C). Cluster analysis is used to create groupings of observations on the basis of their similarity as represented by a set of multivariate data.

Table 3. ANOVA Pairwise Comparison of Ecoregions

Transformed Variable	Significant Difference between Ecoregions
Temperature	**Eastern Cascades Slopes and Foothills** and the **Coast Range**, Eastern Cascades Slopes and Foothills and the Puget Lowland, Eastern Cascades Slopes and Foothills and the Cascades, **Columbia Basin** and the **Cascades**.
Flow	No significant differences between ecoregions found.
Delta pH	Eastern Cascades Slopes and Foothills and the Cascades.
Total Nitrogen	Eastern Cascades Slopes and Foothills and the Coast Range, Eastern Cascades Slopes and Foothills and the Cascades.
Total Inorganic Nitrogen	Eastern Cascades Slopes and Foothills and the Coast Range, Eastern Cascades Slopes and Foothills and the Cascades.
Total Phosphorus	Eastern Cascades Slopes and Foothills and the Coast Range, Eastern Cascades Slopes and Foothills and the Puget Lowland.
Ortho-Phosphorus	Eastern Cascades Slopes and Foothills and the Coast Range, Eastern Cascades Slopes and Foothills and the Puget Lowland, Eastern Cascades Slopes and Foothills and the Cascades, Columbia Basin and the Coast Range.
Dissolved N:P Ratio	Eastern Cascades Slopes and Foothills and the Puget Lowland.

All the variables that showed a significant difference between ecoregions were used in the cluster analysis (e.g. excluded flow). One should not generally apply formal statistical tests to the results of a cluster analysis. The major problem is that the data would be used to both form the clusters and also test their significance. Ideally, one would want the data groupings beforehand. A problem arises by affecting the degrees of freedom for the significance testing by using the data to specify the hypothesis of cluster groupings (Gaugush, 1986). In this report, the cluster analysis is used only to investigate inter-relationships for further regression study.

First, the variables must be standardized to remove the influence of the units of measurement from the results of the analysis. These standardized variables are unitless, so any linear change in the units will not affect the results. Since some of the variables likely covary (e.g. dissolved N and total N), the transformed, standardized variables were clustered using the hierarchical Ward method which resembles the centroid linkage approach but adjusts for covariance. The Pearson product-moment correlation was the distance measure used for clustering since the data are continuous. Other clustering approaches were also investigated, but not selected based on the resulting distance pairings being much smaller showing less differentiation between groups.

The results show 2 very distinct clusters based on essentially a westside/eastside split of the state (Figure 1). These results make sense in that these regions experience different climatic patterns: The westside of the state is affected by a marine climate where the eastside has a more continental climate. The next split of clusters differentiates the eastside and westside clusters based more on elevation. The lower elevation mountains are grouped with adjacent lowlands separated by distinct groups of the higher elevation mountains. Based on this analysis, four clusters of ecoregions seem to make the most sense from the grouping of the variables: (1) Puget Lowlands, Coast Range, and the Willamette Valley; (2) The Cascades; (3) The Eastern Cascades Slopes, and Northern Rockies; (4) Columbia Basin and the Blue Mountains.

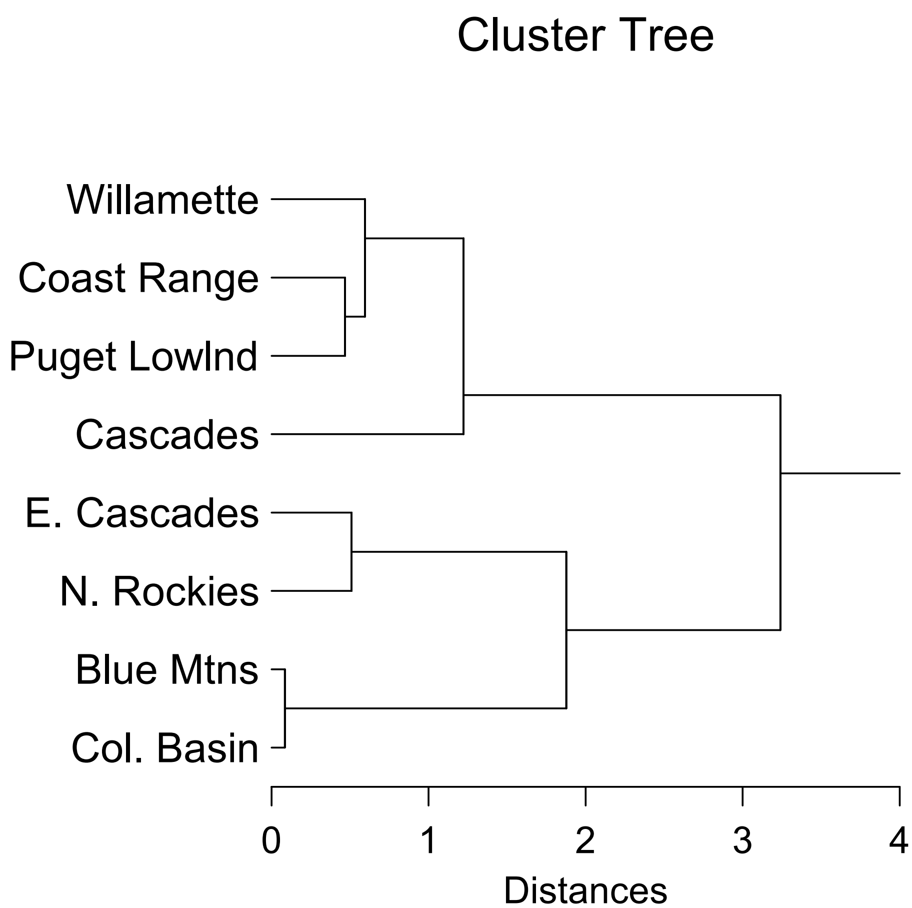

Figure 1. Hierarchical Cluster of Ecoregions Using Regression Variables

In distinguishing the 4 clusters, all variables that showed some significant difference in ANOVA test were used. Some of these variables undoubtedly have a greater influence on the resulting cluster grouping than others. A discriminate analysis was conducted to help identify which of the variables had the most influence on differentiating the 4 clusters. With group membership identified, discriminate analysis can be used to define a linear function of variables that may be used to predict group membership. Conducting a stepwise discriminate analysis, those variables that most influence this group membership can be identified. This information would be useful by helping prepare a strategy for the model development by regression analysis.

The clustered, transformed, standardized variables were subjected to stepwise discriminate analysis. The effect of specific variables on cluster separation can be shown by plotting the canonical variable scores with the confidence ellipse centered on the centroid of each group. Canonical variables are formed in the discriminate analysis as a linear combination on variables that best discriminate among the groups. Additional canonical variates are orthogonal to each other and represent separate, combinations of the variables that distinguish cluster groups that are uncorrelated to each the other canonical variates.

The results of the discriminate analysis show that cluster differences are most influenced by temperature and ortho-phosphorus. A plot of the canonical variables against each other show the separation between the ecoregion clusters caused by temperature and ortho-phosphorus (Figure 2). The eigenvalues show that the first two canonical variables account for 76% of the dispersion between clusters. The next two variables that also influence the cluster grouping are total phosphorus and the dissolved nitrogen to phosphorus ratio. Adding these variables accounts for 97% of the dispersion between clusters. Although the F-statistic shows a significant difference between clusters with all four variables, the visual representation of the canonical scores plot does not show much separation (Figure 3). The F-statistic for the remaining variables does not show significant influence on the separation of the ecoregions clusters. This means that the four variables most responsible for differences between the clusters based on grouping ecoregions are temperature, ortho-phosphorus, total phosphorus, and nutrient limitation.

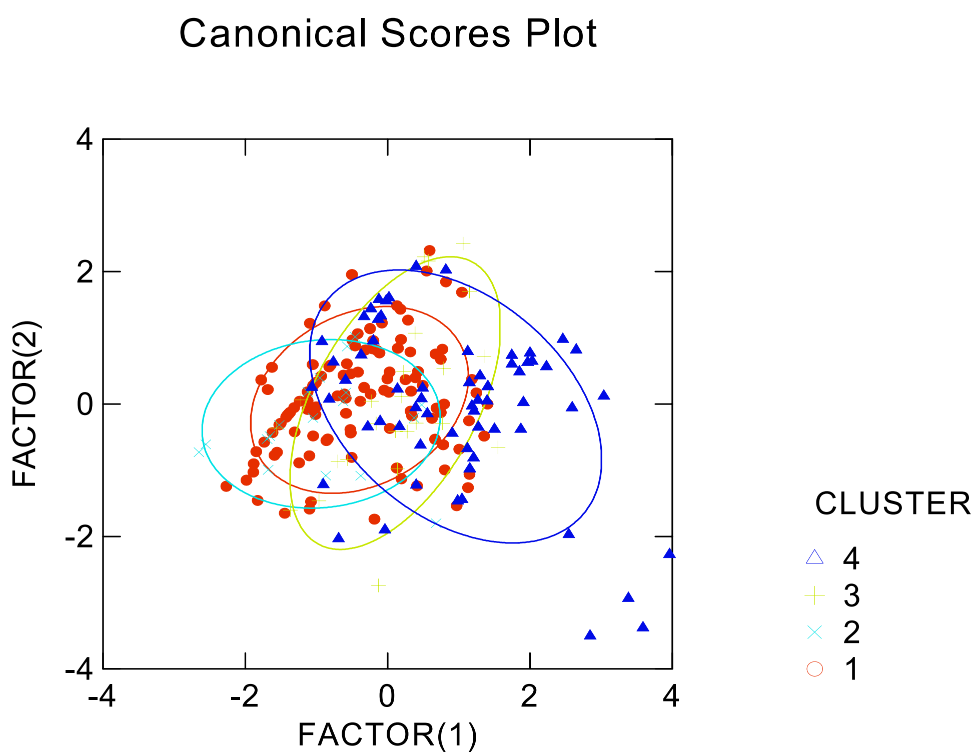

Figure 2. Canonical Variates of Ecoregion Clusters with the Two Most Influentail Variables

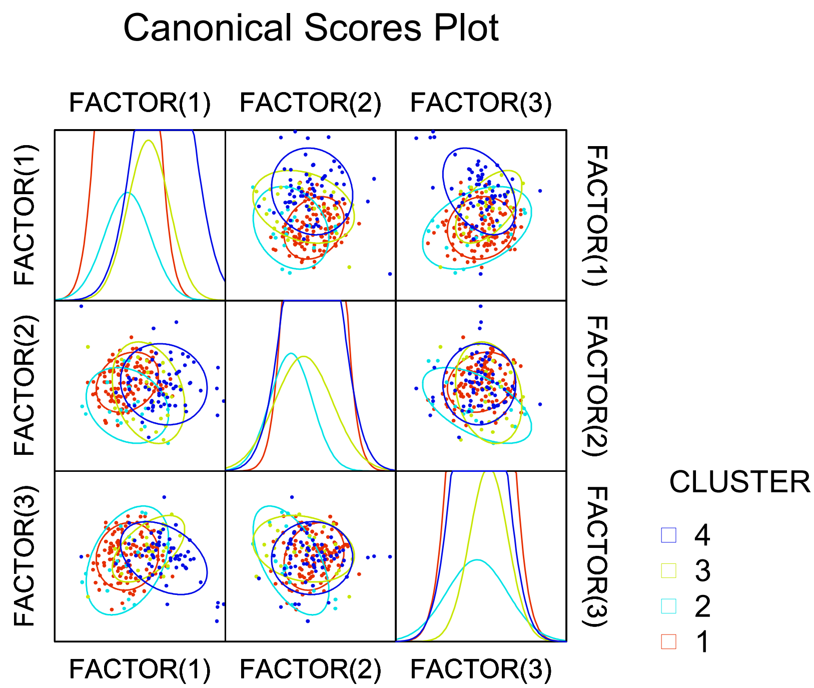

Figure 3. Canonical Variates of Ecoregion Clusters with the Four Most Influentail Variables

Development of an Empirical Model

With some understanding of the influential factors and relationships, the variables can be used to develop a predictive model using regression analysis. The first step was to build a multiple linear regression model using all of the transformed variables. The results tested the robustness of the least-squares sets to sources of ill-conditioning. The conditioning problem of greatest concern was multicollinearity which occurs when independent variables are highly inter-correlated. Multicollinearity occurs when tolerance values are very small and when condition indexes exceed 15. When there are high correlations among independent variables, the estimates of the regression coefficients can become unstable. Tolerance is a measure of this condition. Tolerance is 1 minus the multiple correlation between a predictor and the remaining predictors in the model. The condition indices are the square roots of the ratios of the largest eigenvalue to each successive eigenvalue The condition index is an indicator of the redundancy of the data set.

Results of a multiple linear regression model using all the variables show a notable problem with multicollinearity. Tolerance values for total nitrogen and total inorganic nitrogen are very low (Table 4). The condition indices from the two smallest eigenvalues are 20.8 and 27.4, respectively. These results suggest that some of the variables should be removed because they are inflating the standard errors and F statistics through inter-correlation.

Table 4. Multiple Linear Regression Results Using All Variables

Variable Effect	Coefficient	Tolerance	Probability
Constant	0.331	-	0.003
Temperature	0.019	0.669	<0.001
Flow	-0.040	0.659	0.002
Total Nitrogen	-0.088	0.058	0.434
Total Inorganic Nitrogen	0.034	0.054	0.745
Total Phosphorus	0.031	0.669	0.322
Ortho-Phosphorus	-0.016	0.267	0.731
Dissolved N:P Ratio	0.005	0.309	0.912

One way to observe which variables are correlated is by use of Pearson correlation coefficients and a scatter plot matrix. Apply these tests to the independent variables show that some level of correlation is observed between most of the nutrients. The correlation matrix showing the Pearson coefficients indicate that total nitrogen and total inorganic nitrogen are the most correlated (Table 5). This observation supports the observation of low tolerance values from the multiple linear regression. Total nitrogen and total inorganic nitrogen are also highly correlated with ortho-phosphorus and dissolved N:P ratio. These correlations are particularly evident in the scatter plot matrix showing the 75% gaussian bivariate distribution ellispe (Figure 4).

Table 5. Pearson Correlation Matrix Using All Independent Variables

	Temp.	Flow	Total Nitrogen	Total Inorganic Nitrogen	Total Phosphorus	Ortho-Phosphorus	Dissolved N:P Ratio
Temperature	1.000	0.070	0.377	0.296	0.364	0.379	0.075
Flow	0.070	1.000	-0.475	-0.455	-0.101	-0.251	-0.253
Total Nitrogen	0.377	-0.475	1.000	0.963	0.456	0.696	0.604
Total Inorganic Nitrogen	0.296	-0.455	0.963	1.000	0.443	0.674	0.662
Total Phosphorus	0.364	-0.101	0.456	0.443	1.000	0.521	0.212
Ortho-Phosphorus	0.379	-0.251	0.696	0.674	0.521	1.000	0.091
Dissolved N:P Ratio	0.075	-0.253	0.604	0.662	0.212	0.091	1.000

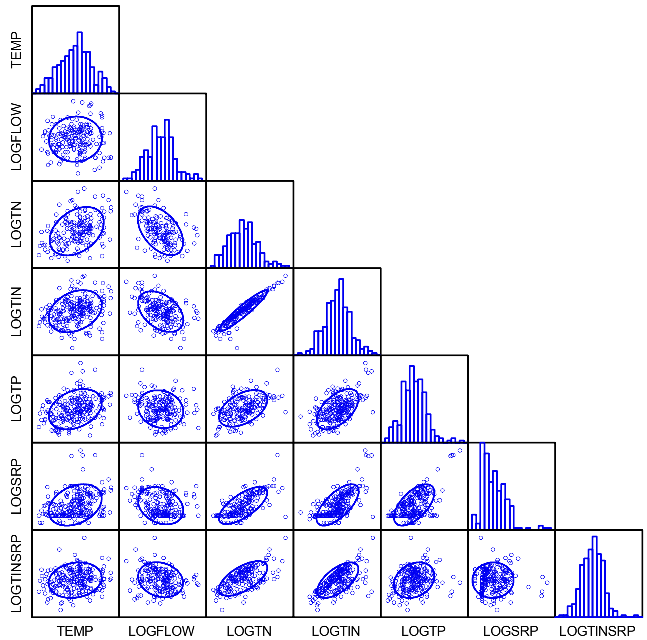

Figure 4. Scatter Plot Matrix of the Regression Variables

These analyses suggest that the multicolinearity is due to the variables representing the total nitrogen and the inorganic nitrogen forms. One way to deal with this problem is to drop these nitrogen variables from the regression model development. The approach would produce a model that would not have a nitrogen term, even for waters that are considered nitrogen-limited. However, it is generally accepted that phosphorus should be the nutrient controlled even for water that is nitrogen-limited (Welch, 1980). Nitrogen-limited waters usually contain cyanobacteria that can fix the atmospheric nitrogen into forms available for algal growth. Controlling nitrogen inputs to these systems is generally unsuccessful. The recommended management approach is to reduce phosphorus to levels that convert the waters to phosphorus limitation. This removes the factor of atmospheric loading that is uncontrollable. Therefore, developing the regression model without the nitrogen terms should not affect the ultimate goal setting for management.

After excluding the total nitrogen and total inorganic nitrogen terms, the remaining variables were analyzed using ordination techniques. These techniques can help define which factors are important in empirical relationships. In ordination, the stations are arranged in relation to one or more coordinate axes such that their relative position provides maximum information about their similarities. Conceptually, ordination can be visualized as placing stations within a variable hyper-space, where there is a single dimension for each variable. The aim of ordination is to

simplify and condense large data sets in the hope that relationships will emerge.
The most common ordination technique is principal components analysis (PCA). PCA is a method of partitioning a variable matrix into a set of orthogonal axes or components. Each PCA axis corresponds to the eigenvalue of the matrix. The eigenvalue is the variance accounted for by that axis. Even though PCA can be used to formally test hypotheses of relationships, the method is typically used just to look for a smaller number of true structural factors or dimensions.

Applying PCA to the remaining variables indicate that there are four distinct factors in the multidimentional variable space. The first two principal components explain 40% and 22% of the variance, respectively. The total phosphorus and ortho-phosphorus appear in the same variable space (Figure 5). Examination of the component loadings shows the nearly same influence of these two variables. One method of further elucidating interpretable factors is through axis rotation to reduce the influence of the large component loadings. Using the most commonly used varimax rotation did not change the effect of spacing the factors appreciably (Figure 6). Therefore, it appears that the four main factors best explain the variable hyper-space.

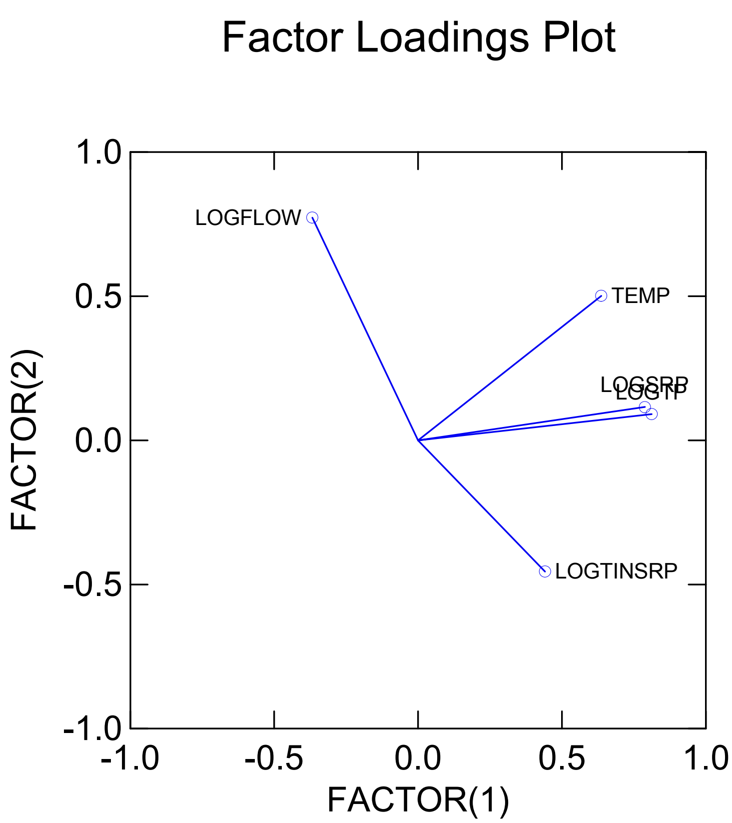

Figure 5. Unrotated Principal Component Factor Loadings

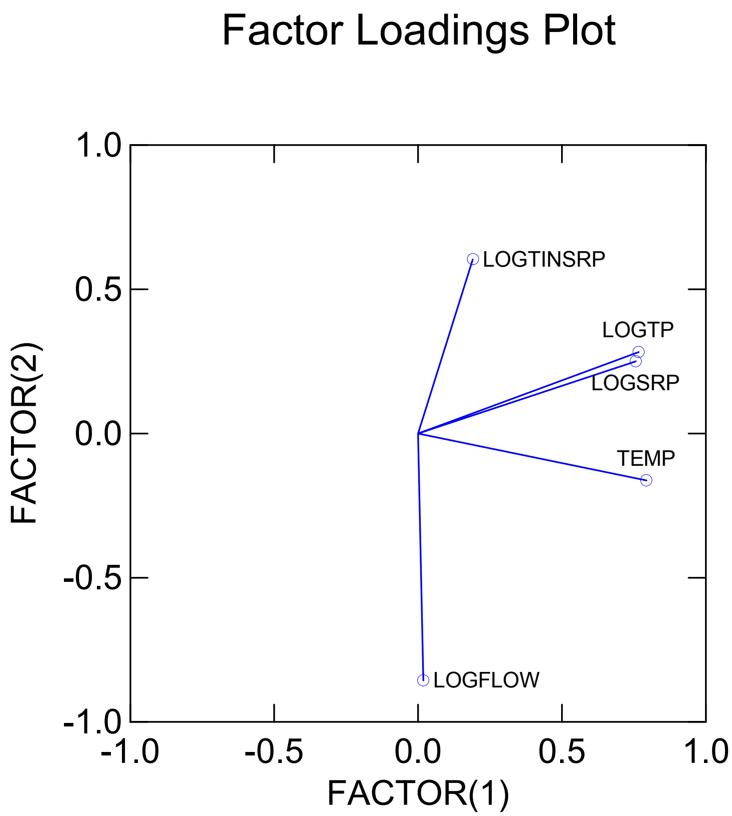

Figure 6. Principal Component Factor Loadings with Varimax Axis Rotation

The remaining 5 variables were then subjected to stepwise multiple linear regression analysis. This process removes the variables that do not significantly contribute to the model performance. Both the Dissolved N:P ratio and the ortho-phosphorus were removed from the model. The final model resulted in a linear combination of maximum temperature, minimum flow, and maximum total phosphorus as best explaining the delta pH caused by algal productivity. The final model derived from the data is highly significant ($p < 0.001$) and explains about 17% of the variation in the data: $(\text{delta pH})^{1/2} = 0.365 + 0.019 * \text{Temperature} - 0.035 * \log(\text{Flow}) + 0.032 * \log(\text{Total P})$

The final model was tested to determine whether stratifying the data by different ecoregions or region clusters would improve predictive performance (Appendix D). The model developed from the Eastern Cascades ecoregion data showed a high explained variance that was not significant. The residuals of this model showed heteroscedasticity with 4 stations exerting a large undue leverage on the results. The model developed from the Columbia Basin ecoregions seemed to develop the best predictive model with a significant explained variance of 35%. This relationship was also evident in all ecoregion clusters where the Columbia Basin data were included. However, the best overall predictive model for all terms was without regional stratification by combining the data from all ecoregions into one set. The probability of each term was lowest for the unstratified model and the residuals plot indicated good homoscedasticity. (Figure 7).

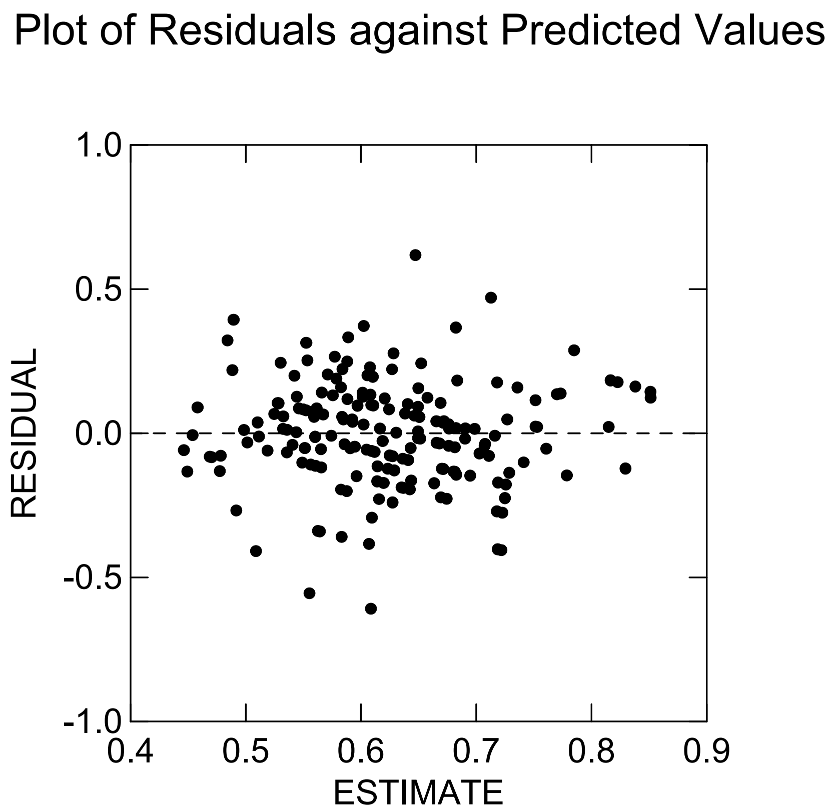

Figure 7. Resdiual Distribution of Final Empirical Model

The effect of nutrient limitation was also tested on the final model. Stations that indicated nitrogen limitation for algal productivity were removed from the data set. A multiple regression model derived using the same three independent variables produced nearly the same results. Explained variance was marginally higher at 18% and the coefficients were nearly the same. Therefore, the separation of stations based on nutrient limitation appears to be an unnecessary step in applying the model.

Implications for TMDL Development

One possible use of the empirical model is to estimate the loading capacity of each of the streams listed for high pH. Identification of the loading capacity is an important step in developing TMDLs. The loading capacity provides a reference for calculating the amount of pollutant reduction needed to bring a water into compliance with water quality standards. By definition, a TMDL is the sum of the allocations. An allocation is defined as the portion of a receiving water's loading capacity that is assigned to particular source. EPA defines the loading capacity as "the greatest amount of loading that a water can receive without violating water quality standards."

Determination of the loading capacity requires defining the critical conditions when pollutant loading has the greatest impact on attaining water quality standards. For this analysis, the critical conditions are the annual maximum temperature and minimum flow measured from the available data at each station location which formed the basis for the segment listing. If these data were not available, the median value for the ecoregion that the segment belongs was used (Appendix C). The empirical model was then applied using the data compiled to calculate the total phosphorus concentration required to achieve the pH standard (Table E).

The statute requires that a margin of safety be identified to account for uncertainty when establishing a TMDL. The margin of safety can be explicit in the form of an allocation, or implicit by the of use of conservative assumptions in the analysis. Uncertainty can be derived for the empirical model developed to predict high pH due to enhanced algal productivity. In addition, the model uses the extreme values for flow and temperature measured at any particular station. These factors provide an inherent margin of safety over uncertainty as required by the statute and are listed below:

1. Used the minimum annual flows measured at each listed segment for determining the total phosphorus concentration required to meet the pH standard.
2. Used the maximum annual temperature measured at each listed segment for determining the total phosphorus concentration required to meet the pH standard.

In order to determine the daily total phosphorus load needed to achieve the pH criterion, flow conditions at each segment must be derived. Annual flow statistics should be used for extrapolation to the daily load since the model uses annual total phosphorus. The annual mean flows for each watershed of the listed segments were estimated by use of a nearby gauged stream based on watershed proportion. The daily total phosphorus loading capacity was then determined based on these flows (Table 6).

Table 6. Total Phosphorus Daily Load Estimated to Meet the pH Criterion

WRIA	Stream Name	Segment Location (TRS)	Annual Mean Daily Flow (cfs)	Total Phosphorus Load (kg/day)
8	Sammamish River	25N-05E-11	214.4	$1.4*10^{+7}$
23	Scatter Creek	15N-03W-08	68.1	$2.5*10^{+12}$
23	Skookumchuck River	14N-02W-07	261.8	$2.9*10^{+15}$
28	Burnt Bridge Creek	02N-01E-38	55.0	$2.3*10^{+16}$
28	Lacamas Creek	01N-03E-44	209.6	$6.2*10^{+15}$
28	Lacamas Creek	02N-03E-10	60.0	$7.3*10^{+17}$
32	Mill Creek	07N-36E-23	60.8	$2.1*10^{-2}$
34	Rock Creek	18N-39E-05	562.3	$1.2*10^{-11}$
34	Pine Creek	20N-42E-28	24.4	$1.6*10^{-5}$
36	Esquatzel Coulee	10N-30E-08	6.9	$1.3*10^{-11}$
36	Scooteney Wasteway	14N-30E-01	1.6	$1.1*10^{-9}$
37	Moxee (Birchfield) Drain	13N-19E-16	3.3	$6.9*10^{-14}$
38	Naches River	13N-18E-12	2169.9	$2.4*10^{+1}$
41	Winchester Wasteway	18N-27E-32	3.7	$6.5*10^{-9}$
41	Frenchman Hills Wasteway	17N-27E-09	4.1	$2.9*10^{-22}$
41	Lind Coulee	18N-29E-35	2.2	$3.4*10^{-12}$
41	Sand Hollow Creek	17N-23E-27	1.9	0
47	Mitchell Creek	29N-21E-34	5.4	$1.1*10^{+10}$
52	O'Brien Creek, S.F.	36N-33E-26	3.1	$2.5*10^{+1}$
55	Little Spokane River	27N-43E-32	269.4	3.8
55	Deadman Creek	27N-43E-33	36.6	$1.3*10^{-11}$
59	Mill Creek	36N-39E-31	56.0	$6.1*10^{+3}$
59	Chewelah Creek, S.F.	33N-41E-23	10.1	$2.1*10^{-1}$
60	Pierre Creek	40N-37E-32	17.4	$2.8*10^{+3}$
61	Deep Creek	40N-40E-33	56.7	$2.9*10^{-9}$
61	Smackout Creek	38N-41E-03	7.6	3.6

The predicted total phosphorus loading values are exceedingly variable. In fact, since the Sand Hollow Creek winter/fall median pH is above the pH criterion, there is no loading capacity estimated by the empirical model for total phosphorus. These highly variable total phosphorus loads are a result of the large unexplained variance of the empirical model of 83%. There are simply other factors that influence the pH that were not considered in this analysis. For example, the amount of riparian shading likely has an influence on lowering the productivity and the resulting pH in streams with higher phosphorus. Because of the large unexplained variance, the empirical model formulated is not suitable for application in a TMDL.

References Cited

Cassidy, K.M., 1997. Land Cover of Washington State: Description and Management. Volume 1 in Washington State Gap Analysis Project Final Report. Washington Cooperative Fish and Wildlife Research Unit, University of Washington, Seattle. WA.

EPA, 1991. Guidance for Water Quality-based Decisions: The TMDL Process. EPA 440/4-91-001. U.S. Environmnetal Protection Agency, Washington, DC.

Gaugush, R.F., 1986. Statistical Methods for Reservoir Water Quality. Report E-86-2. Waterways Experiment Station, U.S. Army Corps of Engineers, Vicksburg, MS.

Hallock, D. and W. Ehinger, 2000. River and Stream Ambient Monitoring Report for Water Year 2000. Publication No. 00-03-035. Washington Department of Ecology, Olympia, WA

Omernik, J.M. and A.L. Gallant, 1986. Ecoregions of the Pacific Northwest. EPA/600/3-86/033. Environmental Research Laboratory, U.S. Envrionmental Protection Agency, Corvallis, OR.

Reckhow, K.H., Beaulac, M.N. and J.T. Simpson, 1980. Modeling Phosphorus Loading and Lake response Under Uncertainty: A Manual and Compilation of Export Coefficients. U.S. Environmental Protection Agency, EPA 440/5-80-011, Washington, DC.

Thomann, R.V. and J.A. Mueller, 1987. Principles of Surface Water Quality Modeling and Control. Harper Collins Publishers, New York, NY.

U.S. District Court, 1998. Consent Decree for Northwest Environmnetal Advocates and Northwest Environmental Defense Center v. Carol Browner, Administrator, U.S. EPA. Case No. C91-427. United States District Court for the Western District of Washington. January 15, 1998.

Welch, E.B, 1980. Ecological Effects of Wastewater. Cambridge University Press, Cambridge, UK.

Zar, J.H., 1984. Biostatistical Analysis. Prentice-Hall, Inc. New Jersey.

APPENDIX A

Descriptive Statistics of the Independent Variables

Table A1. Descriptive Statistics of the Minimum Values of the Independent Variables from the Annual Period (January - December)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.010	0.033	1.500	0.132
Ortho P (mg/L)	222	0.005	0.009	0.019	0.845	0.076
Total N (mg/L)	173	0.010	0.139	0.323	3.600	0.478
Inorganic N (mg/L)	222	0.010	0.048	0.219	4.600	0.550
Total N:P	173	0.1	2.9	5.1	40.0	6.0
Dissolved N:P	220	0.1	4.0	9.9	442	30.7
Temperature (°C)	222	-1.2	2.5	2.6	8.6	2.2
Dissolved Oxygen (mg/L)	222	0.4	9.0	8.7	11.3	1.7
Flow (cfs)	186	0.1	81	2,230	96,100	10,992
Turbidity (NTU)	222	0.02	0.9	1.3	11.0	1.5
Total Suspended Solids (mg/L)	220	1	1	2	12	2
Conductivity (umho/cm)	222	16	54	84	465	75

Table A2. Descriptive Statistics of the 10th Percentile Values of the Independent Variables from the Annual Period (January - December)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.014	0.043	1.719	0.163
Ortho P (mg/L)	221	0.005	0.010	0.022	0.868	0.083
Total N (mg/L)	173	0.010	0.197	0.433	3.620	0.591
Inorganic N (mg/L)	222	0.020	0.102	0.324	6.729	0.724
Total N:P	173	0.3	5.4	7.8	81.7	8.6
Dissolved N:P	220	0.1	7.1	14.9	600	41.8
Temperature (°C)	222	0.4	4.6	4.3	9.0	1.9
Dissolved Oxygen (mg/L)	222	0.6	9.6	9.3	11.5	1.5
Flow (cfs)	188	0.2	1.5	3,276	96,790	14,647
Turbidity (NTU)	222	0.2	1.3	1.9	14.0	2.0
Total Suspended Solids (mg/L)	220	1	2	3	17	3
Conductivity (umho/cm)	222	20	66	101	515	89

Table A3. Descriptive Statistics of the Median Values of the Independent Variables from the Annual Period (January - December)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.029	0.071	3.070	0.261
Ortho P (mg/L)	221	0.005	0.010	0.045	2.560	0.223
Total N (mg/L)	173	0.031	0.337	0.703	6.620	0.978
Inorganic N (mg/L)	222	0.020	0.226	0.595	11.620	1.208
Total N:P	173	1.3	11.4	15.4	132	14.9
Dissolved N:P	220	1.3	18.8	32.7	1294	90.4
Temperature (°C)	222	2.6	9.4	9.4	14.5	1.9
Dissolved Oxygen (mg/L)	222	3.8	11.2	11.0	12.4	1.1
Flow (cfs)	188	2.3	361	5,660	156,850	23,412
Turbidity (NTU)	222	0.3	2.9	5.1	40.0	6.2
Total Suspended Solids (mg/L)	220	1	5	9	83	11
Conductivity (umho/cm)	222	35	146	211	1236	181

Table A4. Descriptive Statistics of the 90th Percentile Values of the Independent Variables from the Annual Period (January - December)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.061	0.131	4.028	0.374
Ortho P (mg/L)	221	0.005	0.013	0.073	3.084	0.306
Total N (mg/L)	173	0.072	0.623	1.178	16.400	1.829
Inorganic N (mg/L)	222	0.020	0.488	1.051	16.610	1.992
Total N:P	173	3.3	22.7	31.8	459	41.4
Dissolved N:P	220	2.2	38.2	70.0	3322	227.4
Temperature (°C)	222	7.7	16.1	16.3	23.6	3.1
Dissolved Oxygen (mg/L)	222	5.5	12.5	12.4	14.9	1.1
Flow (cfs)	188	9	1,530	10,683	297,970	38,092
Turbidity (NTU)	222	0.7	12.1	40.0	2800	198.3
Total Suspended Solids (mg/L)	220	2	21	67	2240	191
Conductivity (umho/cm)	222	34	119	174	829	146

Table A5. Descriptive Statistics of the Maximum Values of the Independent Variables from the Annual Period (January - December)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.118	0.297	9.940	0.861
Ortho P (mg/L)	221	0.005	0.023	0.141	7.120	0.690
Total N (mg/L)	173	0.113	0.826	1.512	16.500	2.159
Inorganic N (mg/L)	222	0.020	0.681	1.300	17.110	2.197
Total N:P	173	3.4	32.2	48.2	500	56.4
Dissolved N:P	220	2.5	52.3	96.3	3422	248.6
Temperature (°C)	222	10.3	19.1	18.7	27.0	3.7
Dissolved Oxygen (mg/L)	222	6.1	13.1	13.2	24.2	1.5
Flow (cfs)	188	10	2,170	18,912	560,000	61,382
Turbidity (NTU)	222	1.2	28.0	277	9500	1079
Total Suspended Solids (mg/L)	220	2	55	342	9,200	947
Conductivity (umho/cm)	222	35	147	211	1236	181

Table A6. Descriptive Statistics of the Minimum Values of the Independent Variables from the Spring/Summer Period (April - September)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.010	0.043	2.070	0.196
Ortho P (mg/L)	221	0.005	0.010	0.020	0.845	0.077
Total N (mg/L)	173	0.010	0.153	0.354	3.600	0.519
Inorganic N (mg/L)	220	0.010	0.051	0.242	6.379	0.641
Total N:P	173	0.1	3.3	5.9	40.0	6.9
Dissolved N:P	220	0.01	4.2	11.0	442	31.3
Temperature (°C)	222	2.1	7.5	7.7	16.3	2.4
Dissolved Oxygen (mg/L)	222	0.4	9.0	8.7	11.3	1.6
Flow (cfs)	186	0.1	93	2,434	104,700	12,127
Turbidity (NTU)	222	0.1	1.0	1.8	14.0	2.2
Total Suspended Solids (mg/L)	220	1	2	3	26	4
Conductivity (umho/cm)	222	17	64	97	548	97

Table A7. Descriptive Statistics of the 10th Percentile Values of the Independent Variables from the Spring/Summer Period (April - September)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.015	0.050	2.535	0.220
Ortho P (mg/L)	221	0.005	0.010	0.023	0.845	0.081
Total N (mg/L)	173	0.010	0.178	0.417	3.608	0.591
Inorganic N (mg/L)	220	0.015	0.084	0.312	8.130	0.775
Total N:P	173	0.2	4.6	7.6	56.9	8.1
Dissolved N:P	220	0.1	6.0	14.7	505	37.6
Temperature (°C)	222	3.9	9.3	9.3	16.4	2.3
Dissolved Oxygen (mg/L)	222	0.4	11.3	9.4	9.1	1.5
Flow (cfs)	187	0.2	125	3,535	119,700	16,131
Turbidity (NTU)	222	0.3	1.3	2.3	17.0	2.6
Total Suspended Solids (mg/L)	220	1	3	4	27	5
Conductivity (umho/cm)	222	19	71	107	550	93

Table A8. Descriptive Statistics of the Median Values of the Independent Variables from the Spring/Summer Period (April - September)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.027	0.071	3.210	0.273
Ortho P (mg/L)	221	0.005	0.010	0.045	2.610	0.229
Total N (mg/L)	173	0.020	0.266	0.573	4.930	0.820
Inorganic N (mg/L)	220	0.020	0.150	0.456	11.450	1.062
Total N:P	173	0.9	9.5	13.7	109	14.6
Dissolved N:P	220	0.3	13.2	24.8	666	53.7
Temperature (°C)	222	7.3	13.6	13.9	20.8	2.6
Dissolved Oxygen (mg/L)	222	2.1	10.1	10.0	11.8	1.2
Flow (cfs)	188	2	286	6,391	184,000	26,401
Turbidity (NTU)	222	0.3	2.4	4.9	44.0	6.4
Total Suspended Solids (mg/L)	220	1	4	10	93	13
Conductivity (umho/cm)	222	25	94	134	595	107

Table A9. Descriptive Statistics of the 90th Percentile Values of the Independent Variables from the Spring/Summer Period (April - September)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.050	0.114	3.740	0.338
Ortho P (mg/L)	221	0.005	0.013	0.062	2.945	0.276
Total N (mg/L)	173	0.038	0.413	0.814	9.120	1.182
Inorganic N (mg/L)	220	0.020	0.253	0.683	13.620	1.427
Total N:P	173	1.3	15.5	24.4	245	27.4
Dissolved N:P	220	1.5	24.2	42.9	930	80.1
Temperature (°C)	222	9.1	17.5	17.3	25.1	3.2
Dissolved Oxygen (mg/L)	222	3.9	11.4	11.3	13.6	1.1
Flow (cfs)	188	3	772	11,166	327,000	42,306
Turbidity (NTU)	222	0.5	6.4	19.5	1027	71.5
Total Suspended Solids (mg/L)	220	1	12	38	1220	97
Conductivity (umho/cm)	222	28	113	164	690	125

Table A10. Descriptive Statistics of the Maximum Values of the Independent Variables from the Spring/Summer Period (April - September)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.085	0.173	4.100	0.411
Ortho P (mg/L)	221	0.005	0.017	0.073	3.100	0.314
Total N (mg/L)	173	0.038	0.573	0.948	9.410	1.254
Inorganic N (mg/L)	220	0.020	0.328	.0813	13.900	1.525
Total N:P	173	1.4	20.6	32.7	311	36.0
Dissolved N:P	220	1.7	33.4	60.5	1062	114.0
Temperature (°C)	222	10.3	19.0	18.7	27.0	3.7
Dissolved Oxygen (mg/L)	222	4.6	11.7	11.8	15.8	1.2
Flow (cfs)	188	3	1,118	16,527	560,000	61,466
Turbidity (NTU)	222	0.5	12.5	106.2	8000	598.3
Total Suspended Solids (mg/L)	220	1	21	130	3740	384
Conductivity (umho/cm)	222	30	132	179	757	136

Table A11. Descriptive Statistics of the Minimum Values of the Independent Variables from the Fall/Winter Period (October - March)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.011	0.037	1.500	0.133
Ortho P (mg/L)	220	0.005	0.010	0.025	1.210	0.109
Total N (mg/L)	173	0.010	0.246	0.523	6.620	0.797
Inorganic N (mg/L)	222	0.010	0.140	0.408	7.840	0.874
Total N:P	173	0.1	5.3	7.9	81.7	9.4
Dissolved N:P	220	0.4	7.9	20.1	1294	87.9
Temperature (°C)	222	-1.2	2.5	2.6	8.6	2.2
Dissolved Oxygen (mg/L)	222	0.5	10.4	10.1	12.4	1.6
Flow (cfs)	187	0.2	129	2,562	96,500	11,616
Turbidity (NTU)	222	0.02	1.2	2.1	90.0	6.2
Total Suspended Solids (mg/L)	220	1	2	3	40	4
Conductivity (umho/cm)	222	16	62	97	491	89

Table A12. Descriptive Statistics of the 10th Percentile Values of the Independent Variables from the Fall/Winter Period (October - March)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.017	0.044	1.585	0.147
Ortho P (mg/L)	220	0.005	0.010	0.028	1.245	0.118
Total N (mg/L)	173	0.016	0.318	0.646	11.060	1.109
Inorganic N (mg/L)	222	0.020	0.211	0.535	10.790	1.165
Total N:P	173	0.6	7.5	10.5	133	12.5
Dissolved N:P	220	0.8	13.8	29.2	2158	145.5
Temperature (°C)	222	0.1	3.8	3.6	8.7	1.9
Dissolved Oxygen (mg/L)	222	1.8	10.8	10.6	12.5	1.4
Flow (cfs)	187	0.9	247	3,591	101,900	14,899
Turbidity (NTU)	222	0.2	1.7	3.0	102.0	7.2
Total Suspended Solids (mg/L)	220	1	3	4	48	5
Conductivity (umho/cm)	222	20	67	111	547	102

Table A13. Descriptive Statistics of the Median Values of the Independent Variables from the Fall/Winter Period (October - March)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.033	0.072	2.640	0.235
Ortho P (mg/L)	220	0.005	0.010	0.045	2.370	0.206
Total N (mg/L)	173	0.043	0.480	0.935	15.700	1.562
Inorganic N (mg/L)	222	0.020	0.372	0.823	16.110	1.656
Total N:P	173	2.2	14.0	19.6	352	28.9
Dissolved N:P	220	1.3	24.6	50.5	3222	217
Temperature (°C)	222	1.2	5.7	5.6	10.5	1.8
Dissolved Oxygen (mg/L)	222	4.1	11.9	11.8	13.8	1.2
Flow (cfs)	187	2.5	565	5,614	143,000	21,714
Turbidity (NTU)	222	0.3	3.7	15.5	1900	127.4
Total Suspended Solids (mg/L)	220	1	6	16	588	48
Conductivity (umho/cm)	222	22	84	139	750	131

Table A14. Descriptive Statistics of the 90th Percentile Values of the Independent Variables from the Fall/Winter Period (October - March)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.062	0.160	5.670	0.532
Ortho P (mg/L)	220	0.005	0.015	0.097	4.736	0.476
Total N (mg/L)	173	0.077	0.709	1.278	16.450	1.955
Inorganic N (mg/L)	222	0.020	0.563	1.124	16.860	2.053
Total N:P	173	3.0	25.4	33.5	479	42.0
Dissolved N:P	220	2.4	40.7	72.5	3372	229.8
Temperature (°C)	222	2.6	9.3	9.5	16.9	2.3
Dissolved Oxygen (mg/L)	222	5.8	12.6	12.6	15.0	1.2
Flow (cfs)	187	4	1,345	9,400	262,000	32,704
Turbidity (NTU)	222	0.5	13.0	59.9	4180	304.3
Total Suspended Solids (mg/L)	220	2	22	98	4526	346
Conductivity (umho/cm)	222	25	105	167	850	148

Table A15. Descriptive Statistics of the Maximum Values of the Independent Variables from the Fall/Winter Period (October - March)

Variables	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Total P (mg/L)	222	0.010	0.096	0.268	9.940	0.863
Ortho P (mg/L)	220	0.005	0.021	0.138	7.120	0.692
Total N (mg/L)	173	0.091	0.822	1.484	16.500	2.164
Inorganic N (mg/L)	222	0.020	0.654	1.282	17.110	2.197
Total N:P	173	3.3	30.7	44.5	500	55.1
Dissolved N:P	220	2.5	46.5	89.7	3422	243.4
Temperature (°C)	222	3.1	10.8	10.9	18.3	2.6
Dissolved Oxygen (mg/L)	222	6.1	13.0	13.1	24.2	1.5
Flow (cfs)	187	6	1885	13,242	330,000	38,888
Turbidity (NTU)	222	0.5	22.0	226.7	9500	938.1
Total Suspended Solids (mg/L)	220	2	42	302	9200	909
Conductivity (umho/cm)	222	30	128	192	1053	168

APPENDIX B

Regression Results of Variables

Table B1. Regression Results of Variables on the Difference between the 90th Percentile Spring/Summer and the Median Fall/Winter pH

Transfomed Variables	Descriptive Statistic for Annual Period (January-December)									
	Minimum		10th Percentile		Median		90Th Percentile		Maximum	
	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability
Log Total P	**2.8%**	**0.008**	**2.9%**	**0.007**	**3.1%**	**0.005**	**3.8%**	**0.002**	**5.3%**	**<0.001**
Log Ortho P	0.6%	0.136	**1.4%**	**0.047**	**2.4%**	**0.013**	**2.8%**	**0.008**	**3.1%**	**0.005**
Log Total N	0%	0.589	0%	0.399	0.8%	0.130	**1.7%**	**0.050**	**3.0%**	**0.013**
Log Inorganic N	0%	0.374	0%	0.855	1.0%	0.073	**3.1%**	**0.005**	**3.9%**	**0.002**
Log Total N:P	0%	0.728	0%	0.660	0%	0.474	0%	0.675	0%	0.837
Log Dissolved N:P	**2.6%**	**0.010**	**1.7%**	**0.032**	0%	0.706	0%	0.345	0.3%	0.191
Temperature	0%	0.394	0%	0.683	0.2%	0.232	**10.2%**	**<0.001**	**11.4%**	**<0.001**
Oxygen	0.8%	0.102	0%	0.368	0%	0.667	0%	0.876	0.1%	0.253
Log Flow	**3.1%**	**0.009**	**1.8%**	**0.037**	0.4%	0.195	0.1%	0.288	0%	0.752
Log Turbidity	0.1%	0.265	1.3%	0.051	**1.9%**	**0.022**	**2.0%**	**0.021**	**4.5%**	**0.001**
Log Total Suspended Solids	0%	0.658	0.5%	0.154	0%	0.410	0%	0.441	**2.2%**	**0.016**
Log Conductivity	0%	0.718	0%	0.643	0%	0.813	0%	0.597	0.3%	0.206

Confidence in the Relationship => 95% in **Bold**, 99% in **Bold Underline**

Table B2. Regression Results of Variables on the Difference between the 90th Percentile Spring/Summer and the Median Fall/Winter pH

Transfomed Variables	Descriptive Statistic for Spring/Summer Period (April-September)									
	Minimum		10th Percentile		Median		90Th Percentile		Maximum	
	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability
Log Total P	**4.1%**	**0.045**	**1.8%**	**0.026**	**2.0%**	**0.020**	<u>**2.8%**</u>	<u>**0.007**</u>	<u>**3.1%**</u>	<u>**0.005**</u>
Log Ortho P	0.8%	0.093	1.4%	0.047	**1.9%**	**0.025**	**2.2%**	**0.016**	<u>**2.6%**</u>	<u>**0.010**</u>
Log Total N	0%	0.556	0%	0.467	0.1%	0.276	1.3%	0.076	**1.9%**	**0.038**
Log Inorganic N	0%	0.441	0%	0.703	0%	0.660	0.4%	0.165	1.3%	0.051
Log Total N:P	0%	0.829	0%	0.520	0%	0.620	0%	0.933	0%	0.785
Log Dissolved N:P	<u>**2.6%**</u>	<u>**0.010**</u>	**2.2%**	**0.017**	0.4%	0.172	0%	0.929	0%	0.725
Temperature	0.8%	0.093	**1.4%**	**0.043**	<u>**8.8%**</u>	<u>**<0.001**</u>	<u>**11.0%**</u>	<u>**<0.001**</u>	<u>**11.4%**</u>	<u>**<0.001**</u>
Oxygen	0.5%	0.154	0%	0.316	0%	0.722	<u>**2.7%**</u>	<u>**0.008**</u>	<u>**6.4%**</u>	<u>**<0.001**</u>
Log Flow	<u>**3.8%**</u>	<u>**0.004**</u>	<u>**3.3%**</u>	<u>**0.008**</u>	**1.9%**	**0.034**	0.3%	0.226	0.1%	0.291
Log Turbidity	0%	0.910	0%	0.896	0%	0.919	0%	0.560	0.9%	0.080
Log Total Suspended Solids	0%	0.850	0%	0.913	0%	0.647	0%	0.994	0.3%	0.212
Log Conductivity	0%	0.492	0%	0.479	0%	0.408	0%	0.295	0.3%	0.206

Confidence in the Relationship => 95% in **Bold**, 99% in <u>**Bold Underline**</u>

Table B3. Regression Results of Variables on the Difference between the 90th Percentile Spring/Summer and the Median Fall/Winter pH

Transfomed Variables	Descriptive Statistic for Fall/Winter Period (October-March)									
	Minimum		10th Percentile		Median		90Th Percentile		Maximum	
	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability
Log Total P	**<u>3.4%</u>**	**<u>0.003</u>**	**<u>3.8%</u>**	**<u>0.002</u>**	**<u>3.8%</u>**	**<u>0.002</u>**	**<u>6.5%</u>**	**<u><0.001</u>**	**<u>5.9%</u>**	**<u><0.001</u>**
Log Ortho P	0.2%	0.246	0.9%	0.092	**2.4%**	**0.012**	**<u>3.7%</u>**	**<u>0.002</u>**	**<u>3.4%</u>**	**<u>0.004</u>**
Log Total N	0%	0.368	0.2%	0.235	1.2%	0.084	**2.1%**	**0.033**	**2.8%**	**0.015**
Log Inorganic N	0.3%	0.206	0.6%	0.135	**2.4%**	**0.013**	**<u>3.6%</u>**	**<u>0.003</u>**	**<u>4.4%</u>**	**<u>0.001</u>**
Log Total N:P	0%	0.667	0%	0.465	0%	0.538	0%	0.650	0%	0.772
Log Dissolved N:P	0%	0.638	0%	0.556	0%	0.696	0.1%	0.277	0.7%	0.117
Temperature	0%	0.423	0%	0.982	0%	0.546	0%	0.306	0.1%	0.251
Oxygen	0%	0.467	0%	0.676	0%	0.372	0%	0.784	0%	0.992
Log Flow	**2.0%**	**0.031**	0.7%	0.132	0%	0.307	0%	0.931	0%	0.556
Log Turbidity	**<u>3.9%</u>**	**<u>0.002</u>**	**<u>5.9%</u>**	**<u><0.001</u>**	**<u>5.5%</u>**	**<u><0.001</u>**	**<u>4.7%</u>**	**<u>0.001</u>**	**<u>6.2%</u>**	**<u><0.001</u>**
Log Total Suspended Solids	0.1%	0.279	**1.6%**	**0.037**	1.1%	0.070	**1.5%**	**0.041**	**<u>3.2%</u>**	**<u>0.005</u>**
Log Conductivity	0%	0.670	0%	0.865	0%	0.865	0%	0.665	0%	0.302

Confidence in the Relationship => 95% in **Bold**, 99% in **<u>Bold Underline</u>**

Table B4. Regression Results of Variables on the Difference between the 90th Percentile and the 10th Percentile Spring/Summer pH

Transfomed Variables	Descriptive Statistic for Annual Period (January-December)									
	Minimum		10th Percentile		Median		90Th Percentile		Maximum	
	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability
Log Total P	0%	0.558	0%	0.986	0%	0.723	0.5%	0.142	**<u>3.6%</u>**	**<u>0.003</u>**
Log Ortho P	0.2%	0.242	0%	0.623	0%	0.910	0%	0.308	0.3%	0.203
Log Total N	0.9%	0.107	0%	0.552	0%	0.885	0%	0.981	0%	0.401
Log Inorganic N	**<u>3.9%</u>**	**<u>0.002</u>**	**1.7%**	**0.029**	0%	0.299	0%	0.780	0%	0.805
Log Total N:P	**<u>3.8%</u>**	**<u>0.006</u>**	**1.9%**	**0.040**	0.3%	0.229	0%	0.311	0%	0.678
Log Dissolved N:P	**<u>5.5%</u>**	**<u><0.001</u>**	**<u>3.6%</u>**	**<u>0.003</u>**	1.0%	0.077	0%	0.365	0%	0.894
Temperature	**<u>12.7%</u>**	**<u><0.001</u>**	**<u>8.4%</u>**	**<u><0.001</u>**	0.6%	0.120	**<u>4.0%</u>**	**<u>0.002</u>**	**<u>4.2%</u>**	**<u>0.001</u>**
Oxygen	0%	0.571	0%	0.600	**<u>4.0%</u>**	**<u>0.002</u>**	**<u>8.8%</u>**	**<u><0.001</u>**	**<u>12.3%</u>**	**<u><0.001</u>**
Log Flow	0%	0.752	0%	0.810	0%	0.478	0.8%	0.120	**1.8%**	**0.039**
Log Turbidity	**1.5%**	**0.038**	0%	0.653	0%	0.496	0.3%	0.188	**<u>4.7%</u>**	**<u>0.001</u>**
Log Total Suspended Solids	0.3%	0.201	0%	0.699	0%	0.771	0%	0.439	**<u>3.7%</u>**	**<u>0.002</u>**
Log Conductivity	0%	0.714	0%	0.880	0%	0.412	0.2%	0.226	**1.6%**	**0.031**

Confidence in the Relationship => 95% in **Bold**, 99% in **<u>Bold Underline</u>**

Table B5. Regression Results of Variables on the Difference between the 90th Percentile and the 10th Percentile Spring/Summer pH

Transfomed Variables	Descriptive Statistic for Spring/Summer Period (April-September)									
	Minimum		10th Percentile		Median		90Th Percentile		Maximum	
	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability
Log Total P	0%	0.393	0%	0.808	0%	0.614	**1.6%**	**0.032**	**5.4%**	**<0.001**
Log Ortho P	0.2%	0.227	0%	0.505	0%	0.962	0%	0.332	0.9%	0.090
Log Total N	0.2%	0.239	0%	0.637	0%	0.856	0%	0.435	1.2%	0.082
Log Inorganic N	**4.2%**	**0.001**	**2.3%**	**0.014**	0.8%	0.093	0%	0.492	0%	0.706
Log Total N:P	**4.4%**	**0.003**	**2.3%**	**0.027**	0%	0.337	0%	0.817	0%	0.404
Log Dissolved N:P	**5.6%**	**<0.001**	**3.9%**	**0.002**	1.3%	0.051	0%	0.390	0%	0.767
Temperature	**5.8%**	**<0.001**	**4.3%**	**0.001**	0.9%	0.085	**3.1%**	**0.005**	**4.2%**	**0.001**
Oxygen	0%	0.509	0%	0.798	**1.9%**	**0.022**	**12.2%**	**<0.001**	**18.8%**	**<0.001**
Log Flow	0%	0.991	0%	0.684	0.4%	0.194	**2.6%**	**0.016**	**3.2%**	**0.008**
Log Turbidity	0.3%	0.202	0%	0.757	0.2%	0.239	**2.7%**	**0.008**	**7.9%**	**<0.001**
Log Total Suspended Solids	0%	0.342	0%	0.789	0.4%	0.178	**2.4%**	**0.013**	**8.4%**	**<0.001**
Log Conductivity	0.1%	0.288	0%	0.596	0%	0.805	0%	0.330	0.9%	0.087

Confidence in the Relationship => 95% in **Bold**, 99% in **Bold Underline**

Table B6. Regression Results of Variables on the Difference between the 90th Percentile and the 10th Percentile Spring/Summer pH

Transfomed Variables	Descriptive Statistic for Fall/Winter Period (October-March)									
	Minimum		10th Percentile		Median		90Th Percentile		Maximum	
	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability
Log Total P	0%	0.717	0%	0.967	0%	0.843	0%	0.305	0.9%	0.079
Log Ortho P	0.1%	0.268	0%	0.587	0%	0.676	0%	0.347	0.2%	0.222
Log Total N	1.5%	0.058	0.3%	0.216	0%	0.693	0%	0.905	0%	0.646
Log Inorganic N	**2.0%**	**0.019**	1.2%	0.053	0%	0.467	0%	0.728	0%	0.792
Log Total N:P	**<u>3.4%</u>**	**<u>0.008</u>**	**2.5%**	**0.021**	0.1%	0.273	0%	0.316	0%	0.702
Log Dissolved N:P	**<u>3.5%</u>**	**<u>0.003</u>**	**<u>3.0%</u>**	**<u>0.006</u>**	0.8%	0.096	0%	0.392	0%	0.926
Temperature	**<u>12.6%</u>**	**<u><0.001</u>**	**<u>10.1%</u>**	**<u><0.001</u>**	**<u>4.6%</u>**	**<u>0.001</u>**	**1.9%**	**0.021**	**<u>3.7%</u>**	**<u>0.002</u>**
Oxygen	0%	0.741	0.2%	0.231	**<u>3.3%</u>**	**<u>0.004</u>**	**<u>7.9%</u>**	**<u><0.001</u>**	**<u>7.8%</u>**	**<u><0.001</u>**
Log Flow	0.1%	0.261	05	0.600	0%	0.799	0%	0.458	0.3%	0.216
Log Turbidity	1.2%	0.055	0%	0.342	0%	0.603	0%	0.843	**2.0%**	**0.021**
Log Total Suspended Solids	0.1%	0.277	0%	0.713	0%	0.364	0%	0.735	1.0%	0.071
Log Conductivity	0%	0.778	0%	0.422	0.2%	0.218	0.7%	0.111	**2.1%**	**0.018**

Confidence in the Relationship => 95% in **Bold**, 99% in **<u>Bold Underline</u>**

Table B7. Regression Results of Variables on the Difference between the 90th Percentile and the Median Spring/Summer pH

Transfomed Variables	Descriptive Statistic for Annual Period (January - December)									
	Minimum		10th Percentile		Median		90Th Percentile		Maximum	
	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability
Log Total P	0%	0.546	0%	0.870	0%	0.734	0.1%	0.266	**2.1%**	**0.018**
Log Ortho P	0.1%	0.271	0%	0.720	0%	0.679	0.2%	0.241	0.5%	0.150
Log Total N	0%	0.955	0%	0.434	0.3%	0.211	0.2%	0.247	1.7%	0.051
Log Inorganic N	**2.3%**	**0.015**	0.6%	0.130	0%	0.720	0%	0.794	0%	0.454
Log Total N:P	0%	0.360	0%	0.570	0%	0.997	0%	0.929	0.3%	0.225
Log Dissolved N:P	<u>**2.9%**</u>	<u>**0.007**</u>	**1.7%**	**0.032**	0.1%	0.291	0%	0.603	0%	0.912
Temperature	<u>**6.8%**</u>	<u>**<0.001**</u>	<u>**5.2%**</u>	<u>**<0.001**</u>	0%	0.696	<u>**4.7%**</u>	<u>**0.001**</u>	<u>**4.5%**</u>	<u>**0.001**</u>
Oxygen	0.8%	0.102	0%	0.542	**1.8%**	**0.027**	<u>**5.5%**</u>	<u>**<0.001**</u>	<u>**9.4%**</u>	<u>**<0.001**</u>
Log Flow	0%	0.681	0%	0.790	4.0%	0.590	0.2%	0.248	0.7%	0.139
Log Turbidity	0.2%	0.223	0%	0.624	0.6%	0.133	0.9%	0.082	<u>**5.1%**</u>	<u>**<0.001**</u>
Log Total Suspended Solids	0%	0.682	0.5%	0.159	0.5%	0.149	0.3%	0.211	<u>**4.1%**</u>	<u>**0.002**</u>
Log Conductivity	0%	0.346	0.4%	0.176	1.2%	0.054	**1.8%**	**0.026**	<u>**4.1%**</u>	<u>**0.001**</u>

Confidence in the Relationship => 95% in **Bold**, 99% in <u>**Bold Underline**</u>

Table B8. Regression Results of Variables on the Difference between the 90th Percentile and the Median Spring/Summer pH

Transfomed Variables	Descriptive Statistic for Spring/Summer Period (April-September)									
	Minimum		10th Percentile		Median		90Th Percentile		Maximum	
	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability
Log Total P	0.4%	0.178	0%	0.451	0%	0.843	0.7%	0.112	**<u>3.0%</u>**	**<u>0.006</u>**
Log Ortho P	0.3%	0.208	0%	0.511	0%	0.995	0%	0.339	0.7%	0.117
Log Total N	0%	0.848	0%	0.392	0.5%	0.184	1.3%	0.071	**<u>3.3%</u>**	**<u>0.010</u>**
Log Inorganic N	2.1%	0.020	1.0%	0.079	0.2%	0.231	0%	0.744	0%	0.615
Log Total N:P	0.1%	0.299	0%	0.531	0%	0.888	0%	0.325	1.3%	0.073
Log Dissolved N:P	**<u>2.9%</u>**	**<u>0.007</u>**	**1.9%**	**0.022**	0.5%	0.141	0%	0.663	0%	0.675
Temperature	1.1%	0.064	0.6%	0.129	1.3%	0.052	**<u>3.9%</u>**	**<u>0.002</u>**	**<u>4.5%</u>**	**<u>0.001</u>**
Oxygen	0.8%	0.099	0%	0.322	0.4%	0.170	**<u>6.9%</u>**	**<u><0.001</u>**	**<u>11.9%</u>**	**<u><0.001</u>**
Log Flow	0%	0.976	0%	0.551	0.1%	0.271	**1.7%**	**0.044**	**1.6%**	**0.045**
Log Turbidity	0%	0.638	0%	0.642	0.6%	0.126	**2.4%**	**0.012**	**<u>6.5%</u>**	**<u><0.001</u>**
Log Total Suspended Solids	0%	0.784	0.5%	0.153	1.1%	0.066	**<u>2.7%</u>**	**<u>0.008</u>**	**<u>8.1%</u>**	**<u><0.001</u>**
Log Conductivity	0%	0.586	0%	0.376	0.6%	0.136	1.3%	0.053	**2.5%**	**0.011**

Confidence in the Relationship => 95% in **Bold**, 99% in **<u>Bold Underline</u>**

Table B9. Regression Results of Variables on the Difference between the 90th Percentile and the Median Spring/Summer pH

Transfomed Variables	Descriptive Statistic for Fall/Winter Period (October-March)									
	Minimum		10th Percentile		Median		90Th Percentile		Maximum	
	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability	Percent Explained Varience	Probability
Log Total P	0%	0.964	0%	0.793	0%	0.663	0%	0.363	0.5%	0.155
Log Ortho P	0%	0.496	0%	0.921	0%	0.327	0.3%	0.207	0.4%	0.162
Log Total N	0%	0.955	0%	0.587	0%	0.323	0.1%	0.294	0.9%	0.111
Log Inorganic N	0.4%	0.173	0%	0.413	0%	0.912	0%	0.800	0%	0.411
Log Total N:P	0%	0.365	0%	0.435	0%	0.814	0%	0.898	0%	0.301
Log Dissolved N:P	0.9%	0.082	0.7%	0.119	0%	0.334	0%	0.574	0%	0.886
Temperature	**<u>6.9%</u>**	**<u><0.001</u>**	**<u>5.3%</u>**	**<u><0.001</u>**	**<u>3.7%</u>**	**<u>0.002</u>**	0%	0.627	1.2%	0.057
Oxygen	0%	0.980	0%	0.497	1.7%	0.031	**<u>4.5%</u>**	**<u>0.001</u>**	**<u>5.2%</u>**	**<u><0.001</u>**
Log Flow	0%	0.342	0%	0.514	0%	0.532	0%	0.756	0%	0.517
Log Turbidity	0%	0.516	0%	0.707	0%	0.441	0%	0.394	**<u>2.6%</u>**	**<u>0.010</u>**
Log Total Suspended Solids	0%	0.651	0%	0.307	0%	0.651	0%	0.845	**1.5%**	**0.042**
Log Conductivity	0.4%	0.180	0.9%	0.085	**1.7%**	**0.028**	**2.4%**	**0.013**	**<u>4.5%</u>**	**<u>0.001</u>**

Confidence in the Relationship => 95% in **Bold**, 99% in **<u>Bold Underline</u>**

APPENDIX C

Descriptive Statistics of the Variables for Regression Testing Stratified by Ecoregion

Table C1. Descriptive Statistics of the Variables for Regression Testing Stratified by Ecoregion

Ecoregion Name	Delta pH (standard units)					
	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Coast Range	22	0.15	0.41	0.41	0.61	0.15
Puget Lowland	84	0.01	0.40	0.44	1.60	0.26
Willamette Valley	8	0.15	0.50	0.54	1.35	0.36
Cascades	18	0	0.25	0.27	0.70	0.18
Eastern Cascades Slopes and Foothills	6	0.20	0.35	0.39	0.65	0.18
Columbia Basin	62	0.05	0.40	0.52	1.40	0.30
Northern Rockies	18	0	0.35	0.37	0.080	0.19
Blue Mountains	1	0.05	0.05	0.05	0.05	not available

Table C2. Descriptive Statistics of the Variables for Regression Testing Stratified by Ecoregion

Ecoregion Name	**Maximum Temperature (°C)**					
	Number of Stations	**Minimum**	**Median**	**Mean**	**Maximum**	**Standard Deviation**
Coast Range	22	10.3	17.1	17.2	24.5	3.3
Puget Lowland	86	10.9	18.2	17.8	25.3	3.2
Willamette Valley	8	14.2	19.3	19.5	25.0	3.6
Cascades	18	10.5	16.0	15.7	19.8	2.9
Eastern Cascades Slopes and Foothills	6	11.7	19.9	17.0	20.3	3.1
Columbia Basin	63	12.4	21.2	21.0	27.0	3.2
Northern Rockies	18	12.2	19.6	19.6	26.7	4.5
Blue Mountains	1	20.0	20.0	20.0	20.0	not available

Table C3. Descriptive Statistics of the Variables for Regression Testing Stratified by Ecoregion

Ecoregion Name	Annual Minimum Flow (cfs)					
	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Coast Range	20	4	38	127	554	180
Puget Lowland	72	1	123	325	4,340	697
Willamette Valley	5	1	35	258	1,200	527
Cascades	14	1	101	691	4,910	1,372
Eastern Cascades Slopes and Foothills	5	1	19	15,545	77,400	34,578
Columbia Basin	53	0	109	5,207	96,100	17,311
Northern Rockies	16	1	46	1,472	11,400	3,220
Blue Mountains	1	743	743	743	743	not available

Table C4. Descriptive Statistics of the Variables for Regression Testing Stratified by Ecoregion

Ecoregion Name	**Minimum Dissolved N:P Ratio**					
	Number of Stations	**Minimum**	**Median**	**Mean**	**Maximum**	**Standard Deviation**
Coast Range	22	11	52	74	198	60
Puget Lowland	86	20	79	96	616	79
Willamette Valley	8	28	81	87	204	57
Cascades	18	2	23	68	372	106
Eastern Cascades Slopes and Foothills	6	4	12	189	1062	428
Columbia Basin	61	6	34	115	3422	439
Northern Rockies	18	5	31	67	275	88
Blue Mountains	1	23	23	23	23	not available

Table C5. Descriptive Statistics of the Variables for Regression Testing Stratified by Ecoregion

Ecoregion Name	**Maximum Total Phosphorus (mg/L))**					
	Number of Stations	**Minimum**	**Median**	**Mean**	**Maximum**	**Standard Deviation**
Coast Range	22	0.011	0.079	0.103	0.334	0.097
Puget Lowland	86	0.013	0.118	0.184	0.872	0.187
Willamette Valley	8	0.020	0.065	0.098	0.308	0.096
Cascades	18	0.012	0.085	0.277	1.450	0.432
Eastern Cascades Slopes and Foothills	6	0.031	0.080	0.157	0.457	0.170
Columbia Basin	63	0.010	0.192	0.618	9.940	1.544
Northern Rockies	18	0.029	0.093	0.121	0.300	0.090
Blue Mountains	1	0.097	0.097	0.097	0.097	not available

Table C6. Descriptive Statistics of the Variables for Regression Testing Stratified by Ecoregion

Ecoregion Name	**Maximum Ortho-Phosphorus (mg/L))**					
	Number of Stations	**Minimum**	**Median**	**Mean**	**Maximum**	**Standard Deviation**
Coast Range	22	0.005	0.010	0.014	0.091	0.017
Puget Lowland	86	0.005	0.020	0.035	0.203	0.039
Willamette Valley	8	0.010	0.018	0.038	0.171	0.054
Cascades	18	0.005	0.012	0.028	0.190	0.043
Eastern Cascades Slopes and Foothills	6	0.024	0.029	0.032	0.052	0.010
Columbia Basin	62	0.007	0.085	0.418	7.120	1.268
Northern Rockies	18	0.010	0.034	0.050	0.182	0.052
Blue Mountains	1	0.043	0.043	0.043	0.043	not available

Table C7. Descriptive Statistics of the Variables for Regression Testing Stratified by Ecoregion

Ecoregion Name	Maximum Total Nitrogen (mg/L))					
	Number of Stations	Minimum	Median	Mean	Maximum	Standard Deviation
Coast Range	19	0.113	0.409	0.522	1.616	0.414
Puget Lowland	71	0.229	0.974	1.223	6.930	1.097
Willamette Valley	4	0.717	1.069	1.411	2.790	0.964
Cascades	10	0.158	0.351	0.954	4.450	1.377
Eastern Cascades Slopes and Foothills	4	0.245	0.585	1.721	5.470	2.505
Columbia Basin	49	0.135	1.510	2.589	16.500	3.459
Northern Rockies	16	0.157	0.846	0.993	3.820	0.905
Blue Mountains	0	not available	not available	not available	not available	not available

Table C8. Descriptive Statistics of the Variables for Regression Testing Stratified by Ecoregion

Ecoregion Name	**Maximum Total Inorganic Nitrogen (mg/L))**					
	Number of Stations	**Minimum**	**Median**	**Mean**	**Maximum**	**Standard Deviation**
Coast Range	22	0.094	0.394	0.558	1.649	0.474
Puget Lowland	86	0.179	0.748	0.930	4.940	0.782
Willamette Valley	8	0.276	0.818	0.986	2.170	0.641
Cascades	18	0.052	0.172	0.644	4.760	1.128
Eastern Cascades Slopes and Foothills	6	0.064	0.217	1.097	5.450	2.139
Columbia Basin	63	0.020	0.982	2.479	17.110	3.647
Northern Rockies	18	0.047	0.500	0.769	3.323	0.826
Blue Mountains	1	0.229	0.229	0.229	0.229	not available

APPENDIX D

Multiple Regression Results of Data Stratified by Ecoregions and Regional Clusters

Table D. Multiple Regression Results of Data Stratified by Ecoregions and Regional Clusters

Ecoregions	Adjusted Coefficient Of Determination	Probability					Resdiual Case Warnings
		Overall	Constant	Temperature	Flow	Total Phosphorus	
Coast Range	0.241	0.061	0.874	0.022	0.155	0.868	None
Puget Lowland	0.056	0.074	0.025	0.018	0.813	0.527	1 outlier
Willamette Valley	0.149	0.566	0.515	0.308	0.699	0.490	4 leverages
Cascades	0.000	0.757	0.844	0.326	0.555	0.820	1 outlier
Eastern Cascades Slopes and Foothills	0.844	0.250	0.403	0.148	0.182	0.160	4 leverages
Columbia Basin	0.345	<0.001	0.006	0.085	<0.001	0.387	None
Northern Rockies	0.008	0.409	0.063	0.634	0.495	0.304	2 outliers
Blue Mountains	1.000	0	0	0	0	0	None
Cluster = Coast Range + Puget Lowland + Willamette	0.094	0.007	0.024	0.001	0.834	0.711	1 outlier
Cluster = E. Cascades + Northern Rockies	0.000	0.517	0.061	0.337	0.719	0.508	1 outlier
Cluster = Columbia Basin + Blue Mountains	0.352	<0.001	0.008	0.085	<0.001	0.334	None
Cluster = Coast Range + Puget Lowland + Willamette + Cascades	0.102	0.002	0.073	<0.001	0.756	0.878	2 outliers
Cluster = E. Cascades + Northern Rockies + Columbia Basin + Blue Mountains	0.277	<0.001	0.006	0.003	0.001	0.169	1 outlier
All Ecoregions	0.168	<0.001	<0.001	<0.001	0.005	0.305	None

APPENDIX E

Final Regression Model Variables for 1998 Section 303(d) Listed Segments for High pH due to Nonpoint Sources

Table E. Final Regression Model Variables for 1998 Section 303(d) Listed Segments for High pH due to Nonpoint Sources.

WRIA	Stream Name	Segment Location (TRS)	Median Fall/Winter pH	Delta pH To Meet Standard	Minimum Flow (cfs)	Maximum Temperature (°C)	Total Phosphorus To Meet Standard (mg/L)
8	Sammamish River	25N-05E-11	7.6	0.9	18.0	25.6	0.001
23	Scatter Creek	15N-03W-08	7.3*	1.2	0.6	20.9	0.756
23	Skookumchuck River	14N-02W-07	7.3*	1.2	60.3	20.4	1.822
28	Burnt Bridge Creek	02N-01E-38	7.0	1.5	1.3**	21.5	13.627
28	Lacamas Creek	01N-03E-44	7.0	1.5	5.0**	24.5	2.151
28	Lacamas Creek	02N-03E-10	6.9	1.6	1.4**	21.2	45.414
32	Mill Creek	07N-36E-23	8.0	0.5	0.2	23.2	<0.001
34	Rock Creek	18N-39E-05	8.4	0.1	7.1**	22.7	<0.001
34	Pine Creek	20N-42E-28	8.2	0.3	1.0	20.7	<0.001
36	Esquatzel Coulee	10N-30E-08	8.4	0.1	27.0	20.5	<0.001
36	Scooteney Wasteway	14N-30E-01	8.2	0.3	5.0	27.0	<0.001
37	Moxee (Birchfield) Drain	13N-19E-16	8.3	0.2	3.0	28.9	<0.001
38	Naches River	13N-18E-12	8.1	0.4	270.0	22.5	<0.001

Table E Continued. Final Regression Model Variables for 1998 Section 303(d) Listed Segments for High pH due to Nonpoint Sources.

WRIA	Stream Name	Segment Location (TRS)	Median Fall/Winter pH	Delta pH To Meet Standard	Minimum Flow (cfs)	Maximum Temperature (°C)	Total Phosphorus To Meet Standard (mg/L)
41	Winchester Wasteway	18N-27E-32	8.3	0.2	76.0	23.2	<0.001
41	Frenchman Hills Wasteway	17N-27E-09	8.5	0	123.0	22.6	<0.001
41	Lind Coulee	18N-29E-35	8.4	0.1	57.0	21.2	<0.001
41	Sand Hollow Creek	17N-23E-27	8.6	0	10.0	21.4	<0.001
47	Mitchell Creek	29N-21E-34	7.6*	0.9	0.8	15.5	0.764
52	O'Brien Creek, S.F.	36N-33E-26	8.1*	0.4	2.7	14.0	0.001
55	Little Spokane River	27N-43E-32	8.1	0.4	58.2**	21.1	<0.001
55	Deadman Creek	27N-43E-33	8.4	0.1	7.9**	20.7	<0.001
59	Mill Creek	36N-39E-31	8.1*	0.4	5.1	12.6	0.003
59	Chewelah Creek, S.F.	33N-41E-23	8.1*	0.4	0.5	17.0	<0.001
60	Pierre Creek	40N-37E-32	8.1*	0.4	1.0	11.0	0.006
61	Deep Creek	40N-40E-33	8.4	0.1	1.6	15.8	<0.001
61	Smackout Creek	38N-41E-03	8.1*	0.4	0.7	15.0	<0.001

* Used the median Fall/Winter pH measured for the segment's ecoregion
** Used flows adjusted from a nearby gaged stream based on watershed proportion

APPENDIX F

Land Use Distribution for Watersheds of the 1998 Section 303(d) Listed Segments for High pH due to Nonpoint Sources.

Table F1. Land Use Distribution for Watersheds of the 1998 Section 303(d) Listed Segments for High pH due to Nonpoint Sources.

WRIA	Stream Name	Segment Location (TRS)	Watershed Land Area (ha)	Agriculture (%)	Developed (%)	Schrubland (%)	Forest (%)
8	Sammamish River	25N-05E-11	28,978	0	48	0	52
23	Scatter Creek	15N-03W-08	12,252	3	2	44	51
23	Skookumchuck River	14N-02W-07	47,216	5	1	25	69
28	Burnt Bridge Creek	02N-01E-38	3,634	23	74	0	4
28	Lacamas Creek	01N-03E-44	13,852	51	2	8	40
28	Lacamas Creek	02N-03E-10	3,966	5	0	17	78
32	Mill Creek	07N-36E-23	9,154	24	1	17	57
34	Rock Creek	18N-39E-05	180,429	83	0	12	5
34	Pine Creek	20N-42E-28	78,305	95	0	3	2
36	Esquatzel Coulee	10N-30E-08	4,458	92	8	0	0
36	Scooteney Wasteway	14N-30E-01	1,043	97	3	0	0
37	Moxee (Birchfield) Drain	13N-19E-16	1,969	100	0	0	0
38	Naches River	13N-18E-12	281,571	6	0	26	68

Table F1 Continued… Land Use Distribution for Watersheds of the 1998 Section 303(d) Listed Segments for High pH due to Nonpoint Sources.

WRIA	Stream Name	Segment Location (TRS)	Watershed Land Area (ha)	Agriculture (%)	Developed (%)	Schrubland (%)	Forest (%)
41	Winchester Wasteway	18N-27E-32	24,230	71	0	29	0
41	Frenchman Hills Wasteway	17N-27E-09	26,524	65	0	35	0
41	Lind Coulee	18N-29E-35	14,094	93	0	7	0
41	Sand Hollow Creek	17N-23E-27	12,372	70	0	30	0
47	Mitchell Creek	29N-21E-34	3,077	0	0	97	3
52	O'Brien Creek, S.F.	36N-33E-26	3,969	0	0	67	33
55	Little Spokane River	27N-43E-32	167,040	50	2	2	45
55	Deadman Creek	27N-43E-33	22,668	44	7	0	49
59	Mill Creek	36N-39E-31	35,195	7	0	18	76
59	Chewelah Creek, S.F.	33N-41E-23	6,757	0	0	27	73
60	Pierre Creek	40N-37E-32	5,926	0	0	43	57
61	Deep Creek	40N-40E-33	47,976	5	0	18	78
61	Smackout Creek	38N-41E-03	3,904	3	0	27	70

Table F2. Estimated Total Phosphorus Load Distribution for Watersheds of the 1998 Section 303(d) Listed Segments for High pH due to Nonpoint Sources.

WRIA	Stream Name	Segment Location (TRS)	Total Phosphorus Load (kg/day)	Agriculture (%)	Developed (%)	Schrubland (%)	Forest (%)
8	Sammamish River	25N-05E-11	50.3	0	83	0	17
23	Scatter Creek	15N-03W-08	16.9	4	3	72	21
23	Skookumchuck River	14N-02W-07	51.8	11	2	51	36
28	Burnt Bridge Creek	02N-01E-38	10.4	22	78	0	1
28	Lacamas Creek	01N-03E-44	38.9	84	2	6	8
28	Lacamas Creek	02N-03E-10	3.7	14	0	40	47
32	Mill Creek	07N-36E-23	35.5	66	2	18	15
34	Rock Creek	18N-39E-05	971.5	95	0	5	1
34	Pine Creek	20N-42E-28	461.8	98	0	2	0
36	Esquatzel Coulee	10N-30E-08	26.3	96	4	0	0
36	Scooteney Wasteway	14N-30E-01	6.3	99	0	1	0
37	Moxee (Birchfield) Drain	13N-19E-16	12.1	100	0	0	0
38	Naches River	13N-18E-12	377.7	28	0	43	28

Table F2 Continued… Estimated Total Phosphorus Load Distribution for Watersheds of the 1998 Section 303(d) Listed Segments for High pH due to Nonpoint Sources.

WRIA	Stream Name	Segment Location (TRS)	Total Phosphorus Load (kg/day)	Agriculture (%)	Developed (%)	Schrubland (%)	Forest (%)
41	Winchester Wasteway	18N-27E-32	120.8	87	0	13	0
41	Frenchman Hills Wasteway	17N-27E-09	126.1	84	0	16	0
41	Lind Coulee	18N-29E-35	81.9	97	0	3	0
41	Sand Hollow Creek	17N-23E-27	61.6	87	0	13	0
47	Mitchell Creek	29N-21E-34	6.7	0	0	99	1
52	O'Brien Creek, S.F.	36N-33E-26	6.7	0	0	89	11
55	Little Spokane River	27N-43E-32	367.7	83	3	2	12
55	Deadman Creek	27N-43E-33	64.8	83	8	0	10
59	Mill Creek	36N-39E-31	34.6	17	0	40	44
59	Chewelah Creek, S.F.	33N-41E-23	6.9	0	0	60	40
60	Pierre Creek	40N-37E-32	7.6	0	0	75	25
61	Deep Creek	40N-40E-33	45.3	12	0	41	47
61	Smackout Creek	38N-41E-03	4.2	7	0	57	37