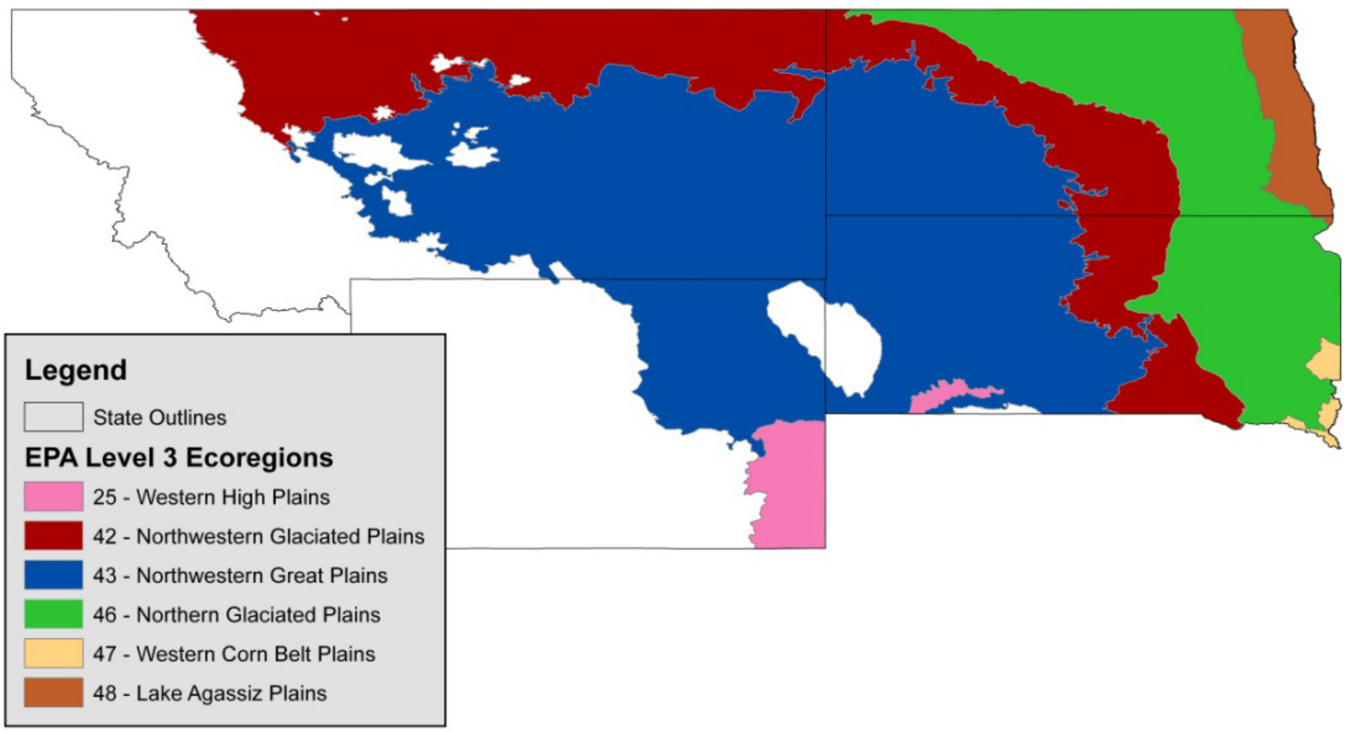
Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and Plain States in Region 8.

United States Environmental Protection Agency



Houston Engineering Inc.

Houston Engineering, Inc.
6901 East Fish Lake Road, Suite 140
Maple Grove, MN 55369
P: 763-493-4522
F: 763-493-5572
www.houstoneng.com

Table of Contents

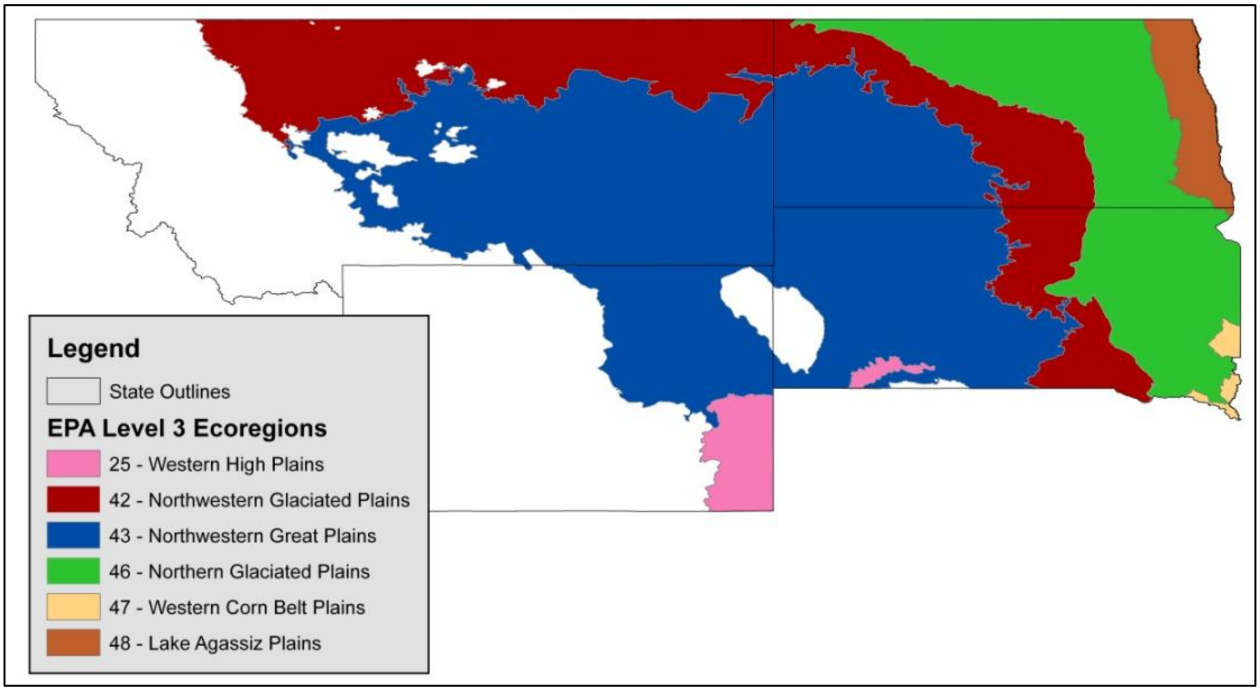
Section I: Executive Summary

Section I: Executive Summary

INTRODUCTION

The following document is comprised of a series of memoranda related to tasks completed under EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8 (i.e. the Nutrient Criteria Project). Work under this contract was performed with the ultimate goal of developing, calibrating, and applying regional models reflective of the watershed nutrient loading to and eutrophication response of the Plains lakes and reservoirs of EPA Region 8 (including portions of North Dakota, South Dakota, Montana, and Wyoming), as shown in **Figure 1**. This work was inspired by efforts completed for the State of North Dakota in 2008 (HEI, 2008), where various classification metrics were tested for their ability to describe eutrophication responses within the State's lakes and reservoirs by developing and applying a regional model based on the classification results. Findings from the North Dakota work suggested that a classification metric based on a water body's surface area, drainage area, and volume may be predictive of the expected eutrophication response of waterbodies in the Plains region. Efforts under this project were to include testing that classification method across the plains portions of EPA Region 8 area.

Figure 1: Nutrient Criteria Project Study Area



Six memoranda were created throughout the course of the EPA Region 8 Nutrient Criteria Project. These memoranda were written, distributed amongst the Project Team (comprised of members from the EPA and each affected Plains State – North Dakota, South Dakota, Wyoming and Montana), discussed, and updated as necessary at critical junctures in the project. The purpose of these memoranda was to update the Project Team on progress made, present incremental results, inform discussion at critical decision points, and prompt feedback from the Project Team on how model development should proceed. The final (most up-to-date) versions of each of those memoranda follow. Combined with this Executive Summary, they comprise the final report for this work.

TASKS AND MEMORANDUM

Following is a list of the tasks to be completed in performance of the EPA Region 8 Nutrient Criteria Project. A brief description of each task is included, along with a summary of the major findings/accomplishments and references to the memorandum that describe in more detail the work that was completed.

Task 1: Project Planning and Organization, Team Designation, and Communications

Efforts under Task 1 of the Nutrient Criteria Project were aimed at developing the Project Team, briefing its members on the work previously completed in North Dakota (the basis for the Region 8 work), and finalizing project timelines and goals. This Task was largely completed during a conference call between HEI and the Project Team on July 15, 2009. In addition to performing the tasks described, the opportunity was taken to make an initial request for data from the states and EPA. A more formal request was created and sent out on July 14, 2009, as discussed under Task 2.

Task 2: Data Requirements and Compilation

Various types of data related to lake and reservoir water quality, geometry, and hydrology and subwatershed characteristics were required for the completion of this project. A list of required (and requested) data was distributed to the Project Team in advance of the July 2009 project kick-off meeting. That list – dated July 14, 2009 – is attached. The request documented both those data that were required for a successful completion of the project and also those data that would provide valuable insight to the study area's lakes and reservoirs. The requested data were to be provided to HEI by the States and EPA. HEI would then use the data to develop master databases of available lake and reservoir parameters for the project's study area.

A description of the data that was provided to and/or collected by HEI is included in the March 5, 2010 memo. Eighteen state and national datasets were used to create two master databases of information on study area lake and reservoir locations, geometries, contributing drainage areas, and the availability of water quality data. **Table 1** summarizes the number of lakes and reservoirs over 10-acres in size that were located across the study area and are included in the master databases.

Table 1: Number of Water Body Locations in the Lakes and Reservoirs Master Databases

<table>
<tr><th></th><th colspan="3">South Dakota</th><th colspan="3">North Dakota</th><th colspan="3">Wyoming</th><th colspan="3">Montana</th></tr>
<tr><td>Lakes</td><td colspan="3">6,739</td><td colspan="3">10,159</td><td colspan="3">293</td><td colspan="3">2,747</td></tr>
<tr><td>Reservoirs</td><td colspan="3">237</td><td colspan="3">306</td><td colspan="3">213</td><td colspan="3">309</td></tr>
<tr><td colspan="13"></td></tr>
<tr><th>EPA Level 3 Ecoregion</th><th colspan="2">25</th><th colspan="2">42</th><th colspan="2">43</th><th colspan="2">46</th><th colspan="2">47</th><th colspan="2">48</th></tr>
<tr><td>Lakes</td><td colspan="2">71</td><td colspan="2">9,267</td><td colspan="2">2,769</td><td colspan="2">7,715</td><td colspan="2">10</td><td colspan="2">106</td></tr>
<tr><td>Reservoirs</td><td colspan="2">62</td><td colspan="2">269</td><td colspan="2">539</td><td colspan="2">166</td><td colspan="2">8</td><td colspan="2">21</td></tr>
</table>

Data on a number of water body characteristics were requested from the Project Team. The most important characteristics for completion of future tasks (including the calculation of the reservoir classification metric and development of the receiving water models), however, include: maximum depth, mean depth, volume, and contributing drainage area. As shown in **Table 2**, the availability of these data was limited. Less than one percent of the 19,938 lakes in the lake master database contain data on either drainage area or volume. All of the lakes that have these data are in North Dakota. The reservoir database is more complete, since the National Inventory of Dams database contains information on these values. Despite the availability of these data, however, major data gaps were still present.

Table 2: Summary of Available Locational, Geometry, and Drainage Area Data across the Study Area

	Location	Surface Area	Maximum Depth	Mean Depth	Volume	Drainage Area	Surface Area, Drainage Area, and Volume
Lakes	19,938	19,938	326	124	97	57	57
Reservoirs	1,065	1,065	241	174	548	377	375

In addition to the lack of data for study area lakes, members of the Project Team expressed concern with the large number of lakes identified in EPA Level 3 Ecoregion 43, particularly in the State of South Dakota. Using data from the National Hydrography Dataset (NHD) to locate lakes and reservoirs in this area may have resulted in an over-estimate due to the presence (and potential misrepresentation in NHD) of a large number of stock and ephemeral ponds in this area. Given these concerns with the available lake data, the Project Team decided to precede forward with the remaining project tasks concentrating only on the study area reservoirs. Further analysis with the lakes was put on hold pending additional information.

To address gaps in the data describing reservoir geometries and contributing drainage areas, regression analysis was used. Linear relationships were developed between available maximum depth-mean depth data combinations, surface area-volume combinations, and surface area-drainage area combinations. Results of these linear regressions were used to fill data gaps in the reservoir master database.

The water quality data made available to HEI by the Project Team is discussed in the September 20, 2010 memorandum. Datasets were provided by North Dakota, South Dakota, and Wyoming detailing observations of total phosphorus (TP), total nitrogen (TN), and chlorophyll-a (Chl-a) concentrations and secchi disk depths (secchi depth) collected within their respective reservoirs from 1977 through 2009 (dates of coverage varied by parameter and by state). [An error was found in the water quality database provided to HEI from the State of Montana, making these data unusable for this project.] Similar to data on reservoir geometry, reservoirs with observed water quality data were limited. Table 3 summarizes the available data, by state, ecoregion, and reservoir classification tier.

Table 3: Summary of Available Water Quality Data across the Study Area

EPA Level 3 Ecoregion	**25**	**42**	**43**	**46**	**47**	**48**
# of reservoirs with water quality data	3	39	74	56	1	5

State	**North Dakota**	**South Dakota**	**Wyoming**
# of reservoirs with water quality data	87	88	3

Reservoir Classification Tier[1]	**I**	**II**	**III**	**IV**
# of reservoirs with water quality data	116	36	14	12

[1] The Reservoir Classification Tier is computed as [(Surface area/Drainage area)*Volume], as discussed in the March 5, 2010 memo and below.

Task 3: Test Classification

The purpose of this task was to test the classification approach developed under the 2008 North Dakota work (HEI, 2008), upon which this project is based. The theory behind the classification approach is that the eutrophication response of reservoirs within each Classification Tier will be distinctly unique from those in other tiers. The reservoir classification in the original work resulted in four lake and reservoir tiers being identified, driven by a metric computed as: [(water body surface area / contributing drainage area) * water body volume]. The validity of this approach (i.e., of the metric) in the Region 8 study area was to be tested using two primary methods: a comparison of notched box plots and an appropriate parametric or nonparametric statistical analysis method. The classification was also to be tested across EPA Level 3 Ecoregions.

The March 5, 2010 memo describes the classification of the study area reservoirs (i.e., those included in the reservoir master database). Results show that the majority of the reservoirs in the study area fall into Classification Tier 1, reflective of reservoirs with small surface areas and volumes and/or large contributing drainage areas. **Table 4** summarizes this result. This finding is similar to what was seen in the North Dakota study that established the approach to classifying waterbodies in this way. With the exception of Wyoming, the spatial distribution (expressed as a density of reservoirs – number per 100 square miles) in each tier is fairly consistent amongst states; it's also relatively consistent amongst ecoregions.

Table 4: Results of Reservoir Classification

Classification Tier	Metric Range [(SA/DA)*Vol] (AF)	Avg Surface Area (acres)	Avg Drainage Area ($miles^2$)	Avg Volume (AF)	Count
I	0 - 7	44	61	330	890
II	7 – 35	268	399	3,219	94
III	35 – 150	1,364	1,022	20,719	37
IV	> 150	25,983	9,569	1,233,901	42

Tests to determine whether or not the reservoir classification approach "works" were based upon whether the observed water quality data in reservoirs of each tier were statistically different from each other. The results of these analyses are presented in the September 20, 2010 memorandum. Using both the qualitative approach of comparing notched box plots and a quantitative approach of employing the non-parametric Kruskal-Wallis statistical test to perform a one-way analysis of variances, results showed that statistical distributions of the individual observations of TP, TN, and Chl-a concentration and secchi depth were not statistically the same between classification tiers. However, further analysis revealed that the distributions were also not consistently statistically-significantly different amongst the tiers. A similar analysis performed on mean annual TP, TN, and Chl-a concentrations and secchi depths revealed even less variation.

Given the results of the Kruskal-Wallis tests, a more detailed statistical test was recommended to gain further insight to the water quality data and determine if confounding factors may assist in separating water quality responses. A two-way analysis of variance test was performed using a general linear model (GLM) to explore the differences amongst individual observations of water quality as a function of classification tier and EPA Level 3 Ecoregion. Results of this work are described in the September 27, 2010 memo. Similar to the results of the Kruskal-Wallis tests, results of the GLM did not provide a clear indication that the reservoir classification technique "worked" for the study area reservoirs (i.e., the statistical distributions of the individual water quality observations were not statistically significantly unique by tier and ecoregion category). Performing the two-way analysis on the mean reservoir values reduces the data variability even more, resulting in less (or no) difference amongst the groups.

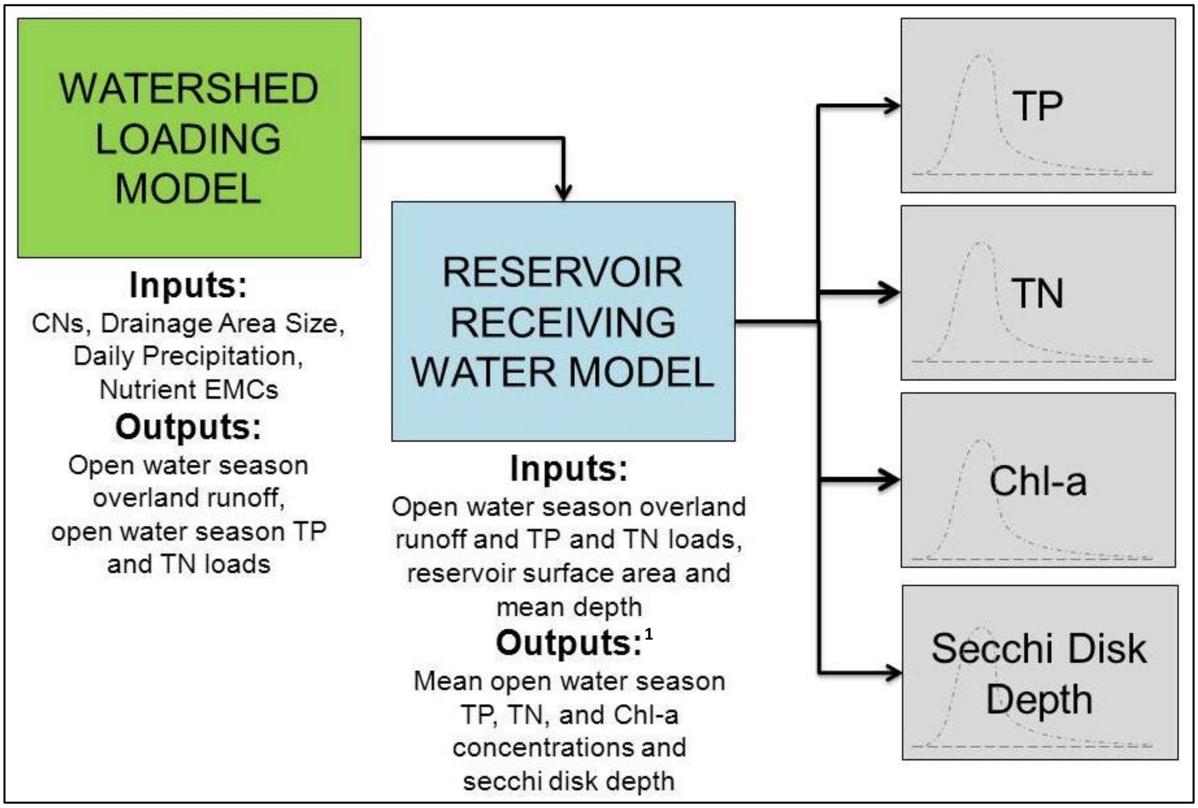
Based on the results of Task 3, creating separate models for the study area reservoirs, by reservoir classification tier, was not warranted. However, the state representatives on the Project Team did feel that, given their intimate understanding of how reservoirs in their states behave, splitting the reservoirs into two larger modeling regions was appropriate. After much discussion, the Project Team decided to create two watershed loading and reservoir receiving water models to address the study area. One model would be created to address reservoirs and land characteristics in Ecoregion 46. The other model would address reservoirs and land characteristics in Ecoregions 42/43. Given the lack of water quality data available in Ecoregions 25, 47, and 48 and their small area of coverage, the Project Team recommended removing those reservoirs from the analysis at this point in time. Additionally, they recommended removing all reservoirs classified as Reservoir Classification Tier 4 reservoirs from the analysis since they consistently appeared to be statistically significantly different than the other tiers (in the Task 3 analyses) and represent the largest reservoirs in the study states. These large reservoirs may have site-specific standards developed for them and would, therefore, not be directly subject to findings of this project's modeling effort.

<u>Task 4: Update Regional Lake and Reservoir Model(s) and Loading Model(s) for the Geographic Region(s)</u>

Task 4 efforts were aimed at updating the previously-developed watershed loading and reservoir receiving water models to address reservoirs in Ecoregion 46 and Ecoregions 42/43, as recommended in Task 3. The original versions of the watershed loading and reservoir receiving water models (created for the North Dakota work) were developed to consider only TP loading from the watershed; due to a lack of data, the models were also only calibrated for watershed hydrology and in-reservoir TP concentrations. For application in the larger EPA Region 8 study area, the Project Team asked that watershed loading model be updated to also compute TN loading into the area's reservoirs. Additionally, given the amount of water quality data available in for this project, the reservoir receiving water models were also to be calibrated for in-reservoir TN and Chl-a concentrations and secchi disk depth.

As described in the April 1, 2011 memo, two regional models were created to simulate the nutrient loading to and eutrophication response within the reservoirs of the study area. The general framework of these models is shown in **Figure 2**, where the watershed loading model computes an open water season (defined as March 1 – November 30) overland runoff, TP loading, and TN loading, based on open water season daily precipitation, and characteristics of the reservoir's contributing drainage area, including: curve numbers, size, and nutrient estimated mean concentrations (EMCs). A reservoir receiving water model (using the CNET model, a spreadsheet version of the Army Corps of Engineer's BATHTUB model (Walker, 1996)) estimates in-reservoir mean open water season TP, TN, and Chl-a concentrations and secchi depths, based on inputs from the watershed loading model and information on the reservoir's geometry, including mean depth and surface area. Modeling is performed stochastically, using a Monte Carlo approach, resulting in a statistical distribution of mean open water season TP, TN, and Chl-a concentrations and secchi depths that are representative of water quality expected to be observed in reservoirs of each of the modeling regions.

Figure 2: General Modeling Framework[1]



[1] Modeling is performed stochastically to create a statistical distribution of mean open water season TP, TN, and Chl-a concentrations and secchi depths.

Statistical distributions for the watershed loading and reservoir receiving water model input variables (i.e., daily precipitation, curve numbers, contributing drainage area sizes, TP and TN EMCs, reservoir mean depths and surface areas) were developed (by region) based on the best data available at the time. The models were calibrated for both watershed hydrology (i.e., unit runoff) and eutrophication response (i.e., in-reservoir mean open water season TP, TN, and Chl-a concentrations and secchi depths). Watershed loadings were not calibrated due to a lack of observed data for use in the process.

Outcomes of the water quality calibration showed that, in general, the empirical water quality relationships available through the CNET/BATHTUB model allow for an accurate estimate of the central tendency of the observed reservoir mean annual water quality data. However, the simulation of the distribution of the mean water quality values (e.g., the 25th and 75th percentiles) is poorer. In a few cases, the simulated water quality distributions show more outliers than the observed values. In other cases, the observed values show a large number of outliers. While a number of explanations for this can be given (as discussed in the April 1, 2011 memo), it is an important discrepancy to note since some methods for setting nutrient criteria rely on these outer percentile and (improper) skewness in the modeled distributions could potentially have an impact on the criteria setting results.

Task 5: Use the Model(s) to Predict Reference/Benchmark Conditions

Once calibrated, the models developed for the reservoirs of the study area can then be used to simulate the impact of various land covers (as a surrogate for loadings) on watershed nutrient loads and the impact that reduced open water season TP and TN loadings from the reservoirs' contributing drainage areas may have on the mean open water season water quality observed within the waters. Simulating these nutrient load reductions and quantifying their impact on open water season water quality within the reservoirs of the study area was the focus of efforts under Task 5. Results of this work are detailed in the April 2, 2011 memo.

Model scenarios were developed to simulate reductions in TP and TN loading to the reservoirs of each modeling region (from their contributing drainage areas). **Table 5** shows an example of the three model scenarios that were performed in Ecoregions 42/43. The Base Conditions model reflects land cover conditions according to the 2001 National Land Cover Dataset (the most recently available at the time of model development). Based on those land use characteristics (and the associated CN and EMC values), a median of 229 kg of TP is delivered to Ecoregions 42/43 reservoirs during the open water season (March 1 – September 30). Scenarios A and B show those loads reduced to 196 and 125 kg, respectively.

Table 5: Ecoregions 42/43 Modeling Scenario Watershed Nutrient Loads

Model Scenario	Median of Simulated Distribution	
	Open Water Season TP Load (kg)	Open Water Season TN Load (kg)
Base Condition	229	1,765
Scenario A	196	1,539
Scenario B	125	1,120

Figure 3 shows the simulated result of reducing the open water season TP and TN loads from the reservoir contributing drainage areas of Ecoregions 42/43 on the in-reservoir mean open water season Chl-a concentrations. **Table 6** shows the results for additional variables, including mean open water

season TP and TN concentrations and secchi depth. Similar results are shown for the Ecoregion 46 model in the April 2, 2011 memo.

Figure 3: Ecoregions 42/43 Simulated Mean Open Water Season Chl-a Concentrations by Model Scenario

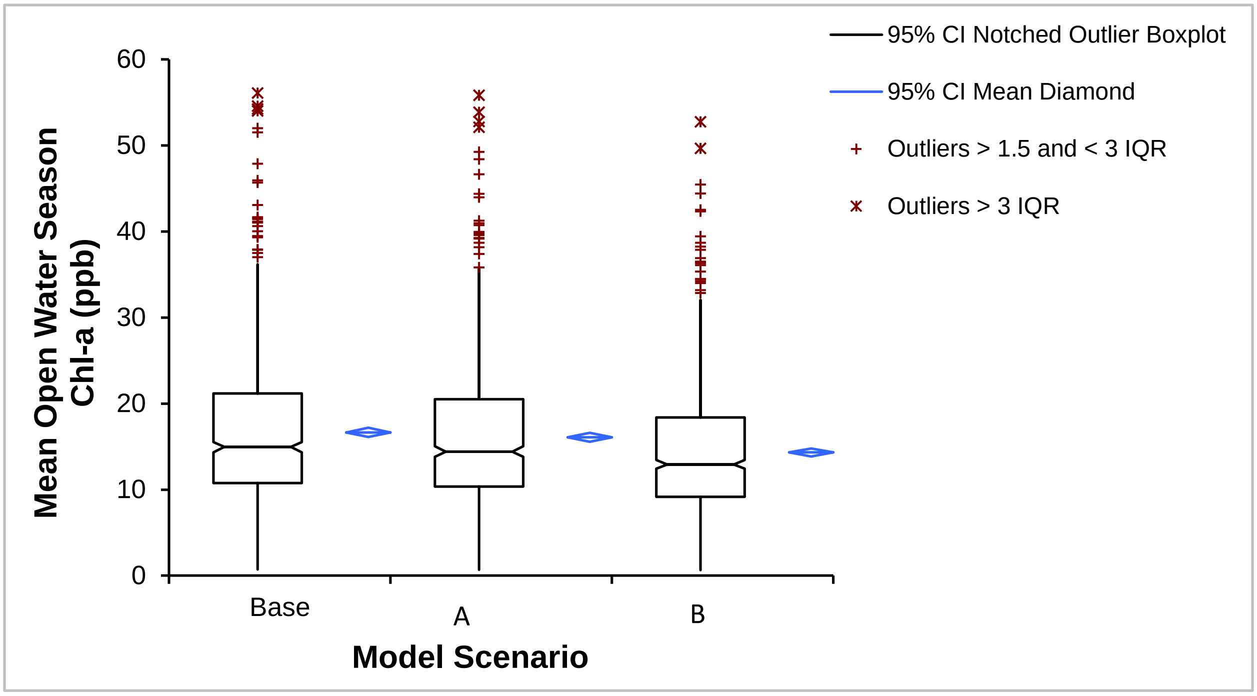


Table 6: Summary of Model Results for Ecoregions 42/43

Model Scenario	Expected (i.e., Median) Mean Open Water Season Values								Expected Nuisance Algal Bloom Frequency[1]
	TP (ppb)	TN (ppb)	Combined Nutrient (ppb)	Chl-a (ppb)	Secchi Depth (m)	Chl-a TSI	TP TSI	Secchi Depth TSI	
Base Condition	135	1,495	82	15.0	1.09	57.2	74.9	58.7	20.3%
Scenario A	126	1,444	78	14.4	1.13	56.8	73.9	58.2	18.3%
Scenario B	104	1,495	68	12.9	1.25	55.7	71.2	56.8	13.1%

[1] Defined as a Chl-a concentration ≥ 20 ppb.

As expected, results of the modeling show that as the nutrient loads into the study area reservoirs reduce, the in-reservoir water quality will improve. The simulated improvements in watershed nutrient loading into the reservoirs, however, are more dramatic than the response seen within the waterbodies themselves. In Ecoregions 42/43, for example, a 40-45% reduction in TP and TN open water season loads (the difference between the Base Conditions and Scenario B) lead to 15-20% improvements in in-reservoir mean open water season concentrations and secchi disk depths. A similar relationship was seen in Ecoregion 46.

Results of this analysis indicate the eutrophication response of the region's reservoirs and provide insight to the Project Team when attempting to make decisions on setting appropriate nutrient criteria for the reservoirs in the Region 8. Further exploration of the reservoir eutrophication response indicates that the reservoirs in the study area may actually be operating under nitrogen limitation. Assuming this is the case, reductions in TN loading will be more effective at improving in-reservoir water quality (related to eutrophication) than reductions in TP.

Similar to the improvement in mean open water season water quality, the frequency of expected nuisance algal blooms (defined by the Project Team as Chl-a concentrations ≥ 20 ppb) is reduced as nutrient loads decrease. Depending on the tolerance of stakeholders/regional citizens for experiencing these types of events, these results may be helpful to the Project Team as they make their decisions on using the model results to inform nutrient criteria standards.

CONCLUSION

The overall goal of this work was to provide information to the EPA and the Project Team for use in establishing nutrient criteria standards for the Plains States of EPA Region 8. The results of analyzing the observed water quality data (Task 3) and developing and applying the regional watershed loading and reservoir receiving water models in Ecoregion 46 and Ecoregions 42/43 (Tasks 4 and 5) provide an appreciation of how water quality varies across the study area and how different nutrient loads delivered to the area's reservoirs may affect the in-reservoir eutrophication response. The Project Team should use these resultants combined with other considerations (such as citizen perceptions of water quality, state and federal regulations, and other policy goals) to inform their decisions as they set forth to set nutrient criteria protective of the water resources of the Region.

REFERENCES

Houston Engineering, Inc. 2008. State of North Dakota Nutrient Criteria Lentic Systems Plan. Prepared for the North Dakota Department of Health Division of Water Quality.

Section II: Project Data Needs

Section II: Project Data Needs

Data Needs
Development of Nutrient Criteria for Lakes and Reservoirs

HEI Project No. 4965-002
EPA Contract No. EP-C-09-001

Lake and Reservoir Morphometry & Mixing Characteristics

- Surface area
- Volume
- Mean depth
- Maximum depth
- Maximum length
- Type of mixing (e.g., dimictic, polymictic)
- Mixed layer depth
- Lake pour points / outlet locations

Additional Data Used to Classify Surface Waters

- National Hydrography Database (NHD) – used as base layer for identifying all lakes and reservoirs
- Other state-wide classifications of lakes and reservoir (e.g., based on recreational use)
- National Wetland Inventory – exclude non-lucustrine systems; identify palustrine system?
- National Dams Database (are there local databases also) – identify reservoir / water control structures

Contributing Drainage Areas

- Subwatershed boundaries (12-digit HUCs) specific to a lake or reservoir – need size
- Soils (SURGO) – for hydrologic soil group
- Land use (National Land Cover) – for developing curve no.
- Annual runoff volume- computed from USGS gage data and curve no.
- National Agricultural Statistical Service database (cropping patterns)
- % developed / imperviousness

Water Quality and Other Data

- Tributary data needed to check watershed model
    - Annual measured runoff volume
    - Loads or yields for nutrients
- In-lake data needed for calibration / verification
    - Annual mean concentrations (e.g., nutrients, chlorophyll-a, clarity, TSI)
    - Episodic "problem" data (e.g., algal bloom frequency)
    - Criteria used for bloom frequency
- Precipitation gage data

Documentation of Data Sources – please identify

- Source of the data (e.g., what agency, how derived, etc)
- Metadata if you have it
- QA process applied to the data

- Documentation / formal reference

Known Data Issues

- State-level water body databases, the relationship to the NHD, and state used naming conventions
- Lack of back data related to lake and reservoir geometry (mean depth to compute volume, used to compute residence time and overflow rate for classification)
- Location of the dam, in the National Dam Database (doesn't always fall at the outlet of a water body)
- Size criteria for wetlands versus lakes; shallow lakes versus deep lakes
- Ability to compute other parameters for classification (e.g., mixing characteristics, fetch, hydraulic residence time, overflow rate)

Section III: Lake/Reservoir Master Databases and Morphometry Data

Section III: Lake/Reservoir Master Databases and Morphometry Data

MEMO

(External Correspondence)

To:	Tina Laidlaw	**From:**	Stephanie Johnson, Ph.D., P.E.
Date:	March 5, 2010	**Through:**	Mark R. Deutschman, Ph.D., P.E.
Cc:	File 4965-002 Dennis McIntyre, GLEC	**Subject:**	Status report on select activities under Tasks 2 & 3 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8

This memo addresses portions of Tasks 2 and 3 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8 (i.e. the Nutrient Criteria project), Under Task 2 of this project, Houston Engineering, Inc. (HEI) created a memorandum describing the lake and reservoir morphometric and water quality data needed to complete the Nutrient Criteria project. That memo, entitled “Data Needs: Development of Nutrient Criteria for Lakes and Reservoirs” was created on July 14, 2009. The requested data was to be provided to HEI by the States and EPA. HEI would then use the data to develop master databases of available lake and reservoir parameters for the project’s study area. This memo addresses the data that has been provided to and/or collected by HEI, the creation of the lake and reservoir databases (including how lakes and reservoirs were defined), the filling of missing morphometric data, and the computation of the reservoir classification metric.

Summary of Data Sources

A combination of state and national data sources were used to gather information on the lake and reservoirs in the study area, which consists of those portions of North Dakota, South Dakota, Wyoming, and Montana that lie within EPA Level 3 ecoregions 25, 42, 43, 46, 47 and 48. These data sources are summarized in **Table 1**.

Table 1: Summary of Data Sources

Product	Agency	Accessed via	File and/or Access Date
National Hydrography Dataset (NHD)	US Geological Survey	http://nhd.usgs.gov/	9/10/2009
National Wetlands Inventory (NWI)	US Fish and Wildlife Service	http://www.fws.gov/wetlands/index.html	9/29/2009
National Lakes Assessment (NLA)	US Environmental Protection Agency	ftp site download	April '09 and Sep '09
National Inventory of Dams (NID)	Army Corps of Engineers	https://nid.usace.army.mil	11/7/2007
List_Stations2.shp[1]	ND Dept of Health and Dept of Game & Fish	HEI project drive	11/7/2007
NDDHLakeData2009.mdb	ND Dept of Health	Joe Gross	April '09
allsdlakes.shp	SD Dept of Environment and Natural Resources	Sean Kruger	8/27/2009
LakeList.pdf	SD Dept of Environment and Natural Resources	Sean Kruger	9/8/2009

Table 1 (cont.): Summary of Data Sources

Product	Agency	Accessed via	File and/or Access Date
SD_LakeReservoir_WQ_Data.xls	SD Dept of Environment and Natural Resources	Sean Kruger	9/28/2009
WyNutrientCriteriaDatabase.mdb	WY Dept of Environmental Quality	Jeremy Zumberge	8/6/2009
Glendo 1st 3_yr Summary.xls	WY Dept of Environmental Quality	Jeremy Zumberge	8/6/2009
Keyhole Vertical Profiles 2002.xls	WY Dept of Environmental Quality	Jeremy Zumberge	8/6/2009
Keyhole Vertical Profiles 2003.xls	WY Dept of Environmental Quality	Jeremy Zumberge	8/6/2009
WY Reservoir Data.xlsx	WY Dept of Environmental Quality	Jeremy Zumberge	8/6/2009
WY_Plains_Lakes.xlsx	WY Dept of Environmental Quality	Jeremy Zumberge	9/24/2009
lake_fwp.shp	MT Fish, Wildlife, & Parks	Mike Suplee	12/9/2009
MeanDepth_Updated.mdb	MT Dept of Natural Resources and Conservation / MT Fish, Wildlife, & Parks	Mike Suplee	12/9/2009
Lake bathymetry maps	SD Dept of Game, Fish, and Parks	www.sdgfp.info/wildlife/Fishing/Lakemaps/Index.htm	Dec. 2009

[1] State data for North Dakota waterbodies were collected and processed during a similar study that preceded this work (HEI, 2008). These data were accessed from their storage location on the HEI server.

All of the datasources listed in **Table 1** are considered secondary datasources per the definition in the "Quality Assurance Project Plan for the Use of Secondary Data" (GLEC, 2009). With the exception of the "LakeList.pdf" and the South Dakota lake bathymetry maps, no quality assurance was performed on the data since it was supplied to HEI directly from the data stewards and assumed to be accurate to their standards. Since "LakeList.pdf" was provided in a file format that was not directly useable, HEI staff manually entered the data into an Excel spreadsheet for use. Staff then QA'd the data by double-checking at least 10 percent of the values transferred and confirming that no mistakes were found. The lake bathymetry maps were hand planimetered to compute volumes for 22 lakes in South Dakota. These volume calculations were also QA'd by checking 10% of the values obtained from the planimeter.

Lake and Reservoir Master Databases

Using the data provided (listed in **Table 1**), HEI set forth to create two master databases; one database to hold information on all of the lakes in the study area and the other database to hold data on the reservoirs. The contents of these databases would then serve as the source of data for the remaining tasks in this project. It is important to note that in work preceding this project (HEI, 2007; 2008), it was determined that waterbodies below 10 acres in surface area were too small to be defined as a lake or reservoir. A similar filter was used in this project and all waterbodies below 10 acres in surface area were immediately removed from the datasets.

The first step to creating the master databases was to determine the difference between a lake and a reservoir, for the purposes of this project. Most of the data sources listed in **Table 1** use a different definition of what constitutes a "lake" or "reservoir". For example, the NHD classifies a lake as "A standing body of water with a predominantly natural shoreline surrounded by land" while the NLA states that a lake is "natural and manmade … greater than 10 acres (4 hectares) in the conterminous U.S., excluding the Great Lakes". To reconcile these differences, lakes and reservoirs were characterized following the procedure discussed in a 2008 report describing similar work done in North Dakota (HEI, 2008). This procedure begins at the state level, using the NHD waterbody layer as a base and adding waterbodies unique to the other datasources (in the order shown in **Table 1**) to build a master database of waterbodies for each state. When investigating duplicate records between state and national sources, in some cases, there was a discrepancy on how a waterbody was defined (as either a lake or a reservoir). In such cases, when one of the databases was not clearly mislabeled (e.g., reservoirs were frequently labeled as lakes in the state database, though upon further investigation a dam was present at one end of the waterbody), the state data was assumed more representative of conditions on the ground and that definition was used.

Table 2 and **Table 3** summarize the results of combining the state and national datasets into two master databases. **Table 2** summarizes the lake master database, showing the number of lakes in the database, both by state and EPA Level 3 ecoregion. **Table 3** follows a similar format, summarizing the results of the reservoir master database. Also shown in these tables (in parentheses), is an indication of the density of lakes and reservoirs in each area, expressed as #/100 sq. miles.

Table 2: Content of Lake Master Database by State and Ecoregion - # and *(#/100 sq.miles)*

	South Dakota	North Dakota	Wyoming	Montana	Total
Ecoregion 25	5 (0.5)	0	66 (1.0)	0	***71 (0.9)***
Ecoregion 42	1,704 (13.8)	5,633 (35.1)	0	1,930 (5.2)	***9,267 (14.2)***
Ecoregion 43	1,362 (3.7)	363 (1.7)	227 (1.2)	817 (1.4)	***2,769 (2.0)***
Ecoregion 46	3,656 (16.6)	4,059 (15.3)	0	0	***7,715 (15.9)***
Ecoregion 47	10 (0.7)	0	0	0	***10 (0.7)***
Ecoregion 48	2 (3.4)	104 (1.5)	0	0	***106 (1.5)***
Total	***6,739 (9.2)***	***10,159 (14.3)***	***293 (1.1)***	***2,747 (2.9)***	***19,938 (7.5)***

Table 3: Content of Reservoir Master Database by State and Ecoregion - # and *(#/100 sq.miles)*

	South Dakota	North Dakota	Wyoming	Montana	Total
Ecoregion 25	3 (0.3)	0	59 (0.9)	0	***62 (0.8)***
Ecoregion 42	47 (0.4)	80 (0.5)	0	142 (0.4)	***269 (0.4)***
Ecoregion 43	127 (0.3)	91 (0.4)	154 (0.8)	167 (0.3)	***539 (0.4)***
Ecoregion 46	52 (0.2)	114 (0.4)	0	0	***166 (0.3)***
Ecoregion 47	8 (0.6)	0	0	0	***8 (0.6)***
Ecoregion 48	0	21 (0.3)	0	0	***21 (0.3)***
Total	***237 (0.3)***	***306 (0.4)***	***213 (0.8)***	***309 (0.3)***	***1,065 (0.4)***

During a October 2009 conference call, staff from the State Agencies indicated some level of discomfort with the number of lakes in the lake master database, with a general feeling that using NHD as the base layer results in an inaccurate (i.e., high) number of lakes in select portions of the study area. This concern seemed particularly true for those areas in ecoregion 43 and especially in South Dakota. Further discussion on this topic is included in an Appendix attached to this memo.

Lake and Reservoir Morphometric Data

In addition to the locations of lakes/reservoirs in the study area, this project also requires information on the morphometric features of these waterbodies. Lake morphometric data was difficult to find. The national datasets have limited information on lakes; for example, the only morphometric data that the NHD and the NWI contain is surface area. The NLA contains information on a few more parameters, such as maximum depth and lake perimeter for a portion of its waterbodies. The main source of morphometric data for the lakes was, therefore, the state-provided databases. Reservoir morphometry was a bit easier to characterize because the NID dataset contains information on reservoir volumes and drainage areas for a portion of its records. The primary dataset for reservoir morphometry was, therefore, the NID followed by state databases. **Table 4** and **Table 5** show the amount of morphometric data available for the lakes and reservoirs, arranged by state. **Table 6** and **Table 7** show the same data, arranged by EPA Level 3 ecoregion. The density of waterbodies with morphomeric data are shown in parentheses.

Table 4: Number of Morphometric Records in the Lake Master Databases - # and *(#/100 sq.miles)*

	South Dakota	North Dakota	Wyoming	Montana
Location	***6,739 (9.2)***	***10,159 (14.3)***	***293 (1.1)***	***2,747 (2.9)***
Maximum Depth	241 (0.3)	74 (0.1)	0	11 (0.01)
Mean Depth	55 (0.1)	65 (0.1)	0	4 (0)
Volume	25 (0.03)	71 (0.1)	0	1 (0)
Drainage Area	0	57 (0.1)	0	0
Surface Area, Drainage Area, and Volume	0	57 (0.1)	0	0

Table 5: Number of Morphometric Records in the Reservoir Master Databases - # and *(#/100 sq.miles)*

	South Dakota	North Dakota	Wyoming	Montana
Location	***237 (0.3)***	***306 (0.4)***	***213 (0.8)***	***309 (0.3)***
Maximum Depth	144 (0.2)	76 (0.1)	0	21 (0.02)
Mean Depth	97 (0.1)	76 (0.1)	0	1 (0)
Volume	155 (0.2)	191 (0.3)	136 (0.5)	66 (0.1)
Drainage Area	133 (0.2)	172 (0.2)	35 (0.1)	37 (0.04)
Surface Area, Drainage Area, and Volume	132 (0.2)	172 (0.2)	34 (0.1)	37 (0.04)

Table 6: Summary of Available Lake Morphometry Data per Ecoregion - # and *(#/100 sq.miles)*

	Ecoregion 25	Ecoregion 42	Ecoregion 43	Ecoregion 46	Ecoregion 47	Ecoregion 48
Location	71 (0.9)	9,267 (14.2)	2,769 (2.0)	7,715 (15.9)	10 (0.7)	106 (1.5)
Maximum Depth	2 (0.03)	82 (0.1)	58 (0.04)	180 (0.4)	2 (0.1)	2 (0.03)
Mean Depth	0	44 (0.1)	3 (0)	75 (0.2)	0	2 (0.03)
Volume	0	41 (0.1)	2 (0)	52 (0.1)	0	2 (0.03)
Drainage Area	0	26 (0.04)	1 (0)	28 (0.1)	0	2 (0.03)
Surface Area, Drainage Area, and Volume	0	26 (0.04)	1 (0)	27 (0.1)	0	2 (0.03)

Table 7: Summary of Available Reservoir Morphometry Data per Ecoregion - # and *(#/100 sq.miles)*

	Ecoregion 25	Ecoregion 42	Ecoregion 43	Ecoregion 46	Ecoregion 47	Ecoregion 48
Location	62 (0.8)	269 (0.4)	539 (0.4)	166 (0.3)	8 (0.6)	21 (0.3)
Maximum Depth	2 (0.03)	50 (0.1)	113 (0.1)	71 (0.2)	2 (0.1)	3 (0.04)
Mean Depth	1 (0.01)	40 (0.1)	70 (0.1)	60 (0.1)	0	3 (0.04)
Volume	28 (0.4)	100 (0.2)	298 (0.2)	107 (0.2)	5 (0.4)	10 (0.1)
Drainage Area	13 (0.2)	71 (0.1)	178 (0.1)	100 (0.2)	6 (0.4)	9 (0.1)
Surface Area, Drainage Area, and Volume	13 (0.2)	70 (0.1)	177 (0.1)	100 (0.2)	5 (0.4)	9 (0.1)

Lake and Reservoir Classification

The nutrient criteria work that was done in North Dakota (HEI, 2008) tested a number of different metrics for classifying lakes and reservoirs for further analysis in developing nutrient criteria. The metric determined most appropriate for that work using the limited available data, was based on the waterbody's surface area, drainage area, and volume (computed as [(surface area/drainage area)*volume]). The scope of this project is to use this same metric and test its application for nutrient criteria development in the Plains region of EPA Region 8. Adequate data on each of these three waterbody parameters must be available.

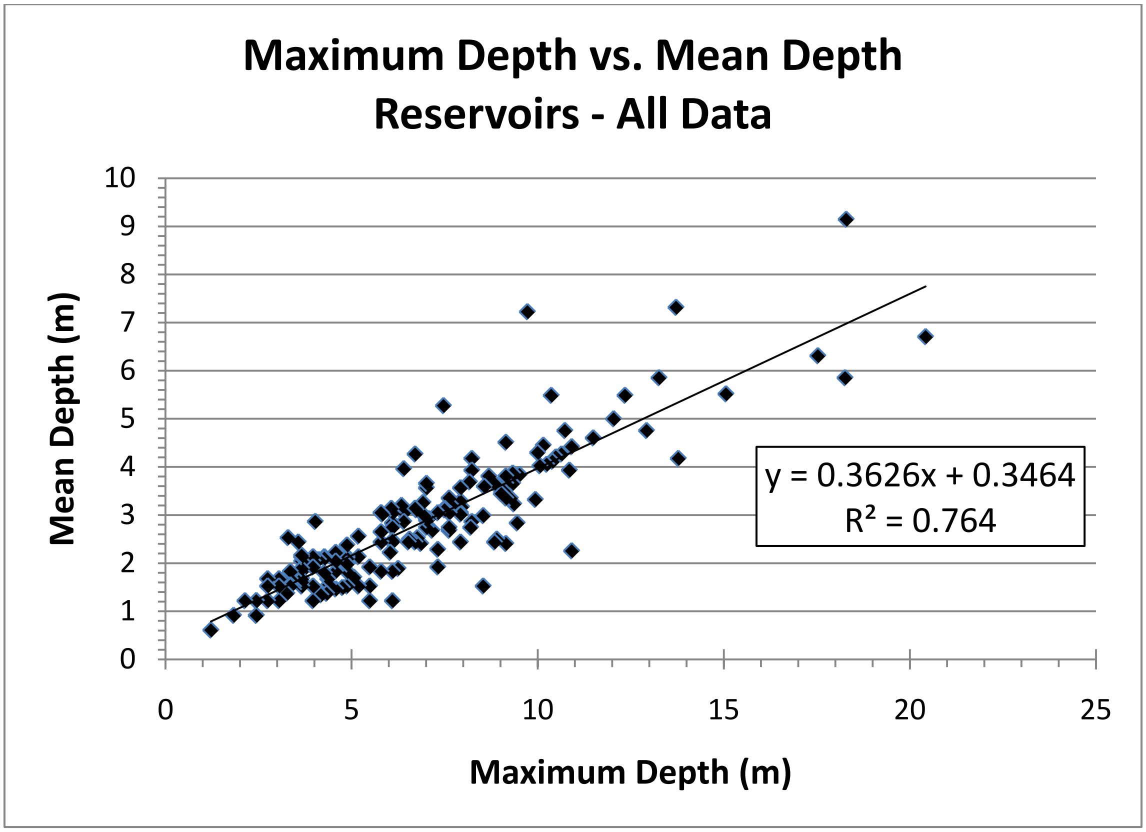
Table 4 through **Table 7** summarize the information that HEI was provided on the parameters needed to classify the lakes and reservoirs using the metric; the bottom row in each table shows the number of lakes/reservoirs that have all of the necessary data. As shown, there are significant data gaps for computing the metric ; less than one percent of the 19,938 lake records, for example, contain data on either drainage area or volume and the vast majority of the lakes that contain these data are in North Dakota. The reservoir database is more complete than the lake database since the NID dataset has drainage area and volumes for many of the reservoirs. Thirty-five percent of the reservoirs have drainage areas associated with them; fifty-one percent have volumes.

This lack of data for computing the lake and reservoir classifications was addressed in a conference call with EPA and State Agency staff in October 2009. At that time, HEI suggested options for filling the data gaps, including field surveys, viewing/digitizing bathymetric maps, GIS analysis, and/or developing regressions between variables. Given the resources available, it was determined that developing regressions is the best approach. However, in the case of the lake data, it is important to note the small number and geographically segregated nature of the data, which could add significant error to the analysis. Given that fact, it was agreed that, in the absence of more complete lake morphometric data, HEI would proceed forward with developing regressions and computing the classification metric for the reservoirs in the master database. Further analysis with the lakes was put on hold pending additional information.

Morphometric data from the reservoir master database was used to create relationships amongst the parameters and to compute the reservoir classification metric. The first step in this process was to create a relationship between maximum and mean depth. One hundred seventy four of the reservoirs in the reservoir master database had values for both maximum and mean depths, while 67 had values for maximum depth but not the mean. Given the ability to compute a reservoir volume from the mean depth (and surface area), it was desirable to estimate these mean depths from the maximum value. It was assumed that computing volumes from mean depths and surface areas would give a more accurate estimate of the actual volume than doing so with a regression equation (described below).

The relationship between the mean and maximum depth was computed by lumping all of the data in the database into a single sample and computing the regression. The regression was then computed for smaller groups of the data, segregating the values on a per state and per ecoregion basis. Results showed that the slopes of the regression lines for the lumped vs. state vs. ecoregion analyses were not statistically significantly different when compared using a 95% confidence interval. It was, therefore, determined that a single relationship could be used to describe all data in the database and the 67 missing mean depth values were computed using the (lumped) regression shown in **Figure 1**. Mean depths values were then multiplied by the surface area values to compute volumes for those reservoirs that lacked volume data, but had a mean depth (either provided in the original datasets or computed via the max depth).

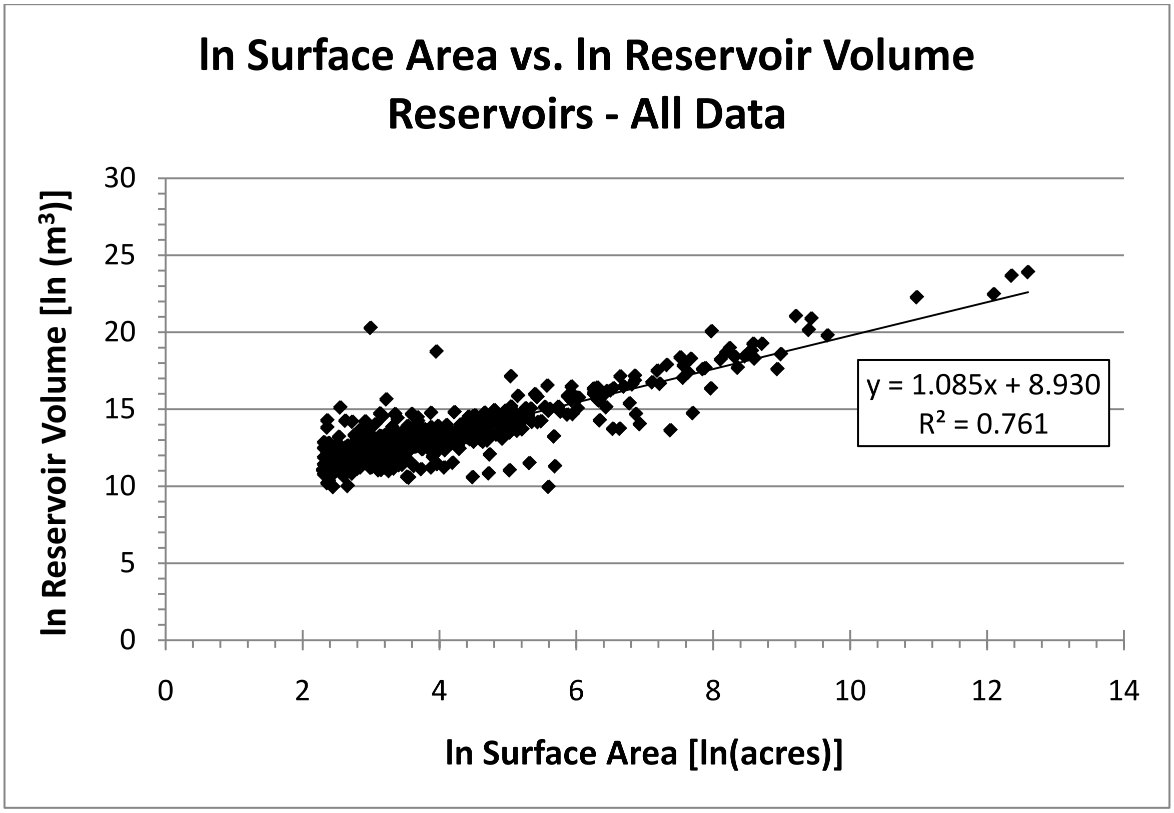
Figure 1: Maximum vs. Mean Depth – Reservoir Master Database



The next two regressions were created to address the relationship between surface area/volume and surface area/drainage area. Because NHD was used as the base layer for creating the Master Databases, all of the reservoirs in the database (except 2) have a surface area value populated. For this reason, surface area was the logical choice to serve as the independent variable for filling data gaps.

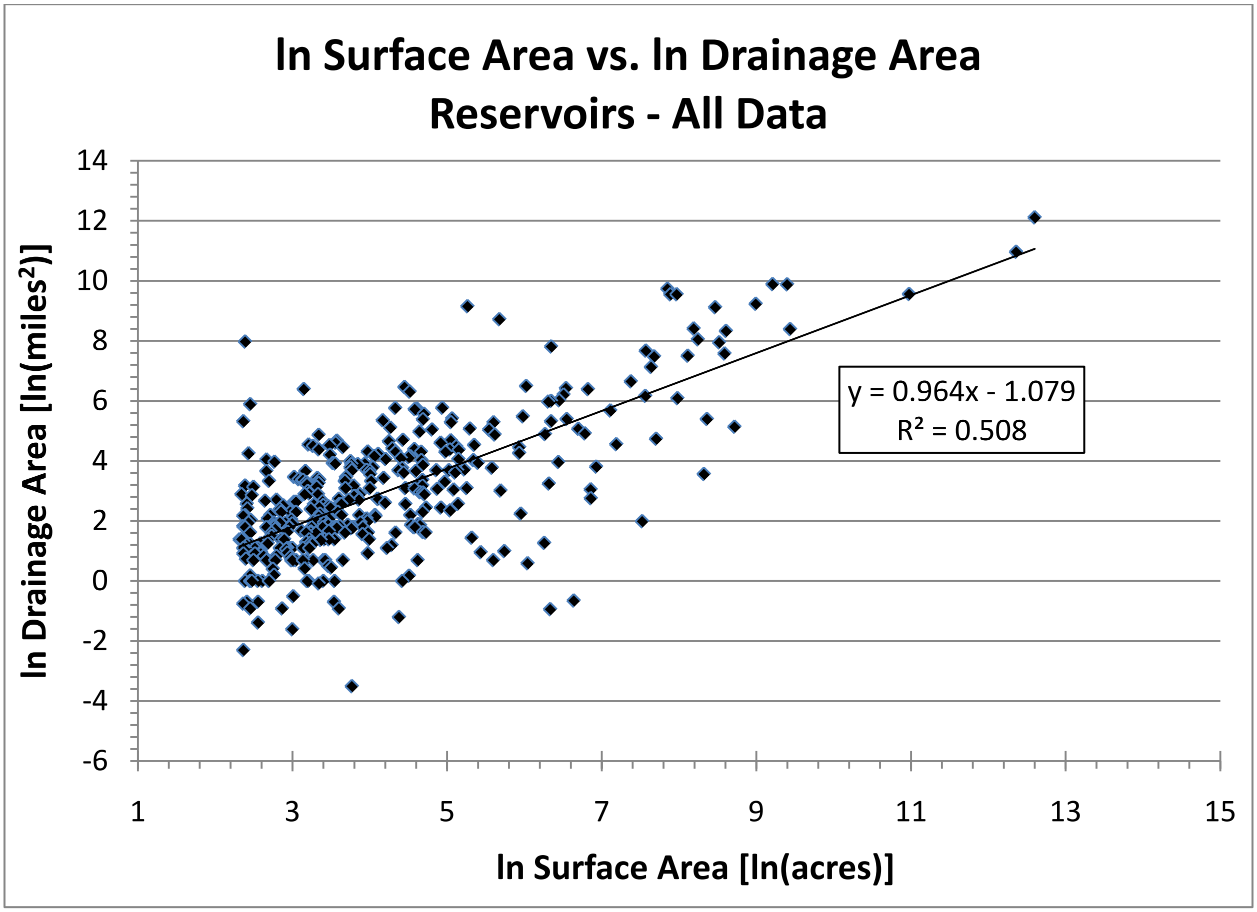
The surface area/volume regression used data from reservoirs that had volume data contained in the original datasets as well as those that had a volume computed from a mean depth, resulting in 596 values used. **Figure 2** shows the results of the regression. Similar to what was done for the maximum vs. mean depth analysis, the surface area vs. volume regression was computed for all data lumped together as well as by state and by ecoregion. The lumped vs. state vs. ecoregion relationships were not statistically significantly different at a 95% confidence interval, so the lumped analysis was retained.

Figure 2: Surface Area vs. Volume – Reservoir Master Database (Natural Log Scale).



Results of the surface area vs. drainage area (lumped) regression are shown in **Figure 3**. Again, results of regressions run on these data segregated by state and by ecoregion were not statistically significantly different than those of the lumped analysis. So the lumped analysis was used for future steps. The number of data points in this analysis was 241.

Figure 3: Surface Area vs. Drainage Area – Reservoir Master Database (Natural Log Scale).



Results of the regressions shown in **Figure 2** and **Figure 3** were used to fill all of the volume and drainage area data gaps in the reservoir master database. Since all necessary data were then populated for each reservoir (with either reported or filled values), the classification metric was then computed. As mentioned, the classification metric used in this project is the one that was determined most effective in the North Dakota study (HEI, 2008) that preceded this work. The metric is computed as [(surface area/drainage area) * volume]. **Table 8** summarizes the results of computing the metrics for the reservoirs in the master database, arranged by tier. Work in the North Dakota study showed that frequency distributions of the metric (when applied to both lake and reservoir data in that study) resulted in four distinct groups (i.e., tiers) of data. Tiers were defined for this work using the same metric ranges reported in the North Dakota study. **Table 9** and **Table 10** show how the numbers of reservoirs in each tier are distributed across the states and EPA Level 3 ecoregions.

Table 8: Results of Reservoir Classification

Classification Tier	Metric Range [(SA/DA)*Vol] (AF)	Avg Surface Area (acres)	Avg Drainage Area (miles^2)	Avg Volume (AF)	Count
I	0 - 7	44	61	330	890
II	7 – 35	268	399	3,219	94
III	35 – 150	1,364	1,022	20,719	37
IV	> 150	25,983	9,569	1,233,901	42

Table 9: Count of Reservoirs in each Classification Tier by Ecoregion- # and *(#/100 sq.miles)*

Classification Tier	Ecoregion						
	25	**42**	**43**	**46**	**47**	**48**	**Total**
I	48 (0.6)	228 (0.4)	465 (0.3)	126 (0.3)	7 (0.5)	16 (0.2)	890
II	8 (0.1)	23 (0.04)	38 (0.03)	20 (0.04)	1 (0.1)	4 (0.1)	94
III	3 (0.04)	5 (0.01)	16 (0.01)	12 (0.02)	0	1 (0.01)	37
IV	2 (0.03)	13 (0.02)	19 (0.01)	8 (0.02)	0	0	42
Total	61 (0.8)	269 (0.4)	538 (0.4)	166 (0.3)	8 (0.6)	21 (0.3)	1,063

Table 10: Count of Reservoirs in each Classification Tier by State- # and *(#/100 sq.miles)*

Classification Tier	State				
	South Dakota	**North Dakota**	**Wyoming**	**Montana**	**Total**
I	196 (0.3)	249 (0.4)	186 (0.7)	259 (0.3)	890
II	28 (0.04)	28 (0.04)	14 (0.1)	24 (0.03)	94
III	5 (0.01)	17 (0.02)	5 (0.02)	10 (0.01)	37
IV	8 (0.01)	12 (0.02)	6 (0.02)	16 (0.02)	42
Total	237 (0.3)	306 (0.4)	211 (0.8)	309 (0.2)	1,063

Note that 1,063 reservoirs were classified with the metric; two of the reservoirs the master database do not have information on surface area so they could not be classified.

Conclusion

Results of the reservoir classification show that a majority of the reservoirs in the study area fall into Tier I of the metric classification and the count of reservoirs in each proceeding tier grow smaller. This finding is similar to what was seen in the North Dakota study that established the approach to classifying waterbodies in this way. With the exception of Wyoming, the density of reservoirs in each tier is fairly consistent amongst states; it's also relatively consistent amongst ecoregions. It should be noted that a lack of reported morphometric data resulted in HEI having to fill data gaps with regression equations. Computing the classification metrics from actual morphometric data for those reservoirs that currently have estimated values in the database, could result in a different metric value and different tier for those waterbodies. Given the amount of data available, this approach is used as an approximation.

The next step in the project is to determine whether water quality expressed as some measure of eutrophication differs among the tiers. A subsequent memorandum will identify the amount of water quality data (i.e., paired total phosphorus, dissolved phosphorus, inorganic nitrogen, chlorophyll-a and secchi depths) available for the reservoirs and present descriptive statistics by reservoir tier. The data will also be evaluate the statistical relationship between total phosphorus and the eutrophication response variables (chlorophyll-a and secchi depth as a measure of water clarity). This information will then be used to calibrate idealized watershed loading and lake response models for each tier to compute the statistical distributions of average annual total phosphorus and the related eutrophication response variables (i.e., chlorophyll-a, secchi depth). Whether the classifications "works" is expected to be based upon whether: 1) the models for each tier can be calibrated; 2) the annual average total P distributions for each tier are largely discrete; and 3) each tier shows a differing distinct response to total phosphorus.

References

Great Lakes Environmental Center. 2009. Quality Assurance Project Plan for the Use of Secondary Data: Contract No. EP-D-09-001.

Houston Engineering, Inc. 2007. State of North Dakota Nutrient Criteria Development Plan. Prepared for the North Dakota Department of Health, Division of Water Quality.

Houston Engineering, Inc. 2008. State of North Dakota Nutrient Criteria Lentic Systems Plan. Prepared for the North Dakota Department of Health, Division of Water Quality.

APPENDIX

Discussion on Results of the Lake Master Database

During discussions with the State Agency staff, there was a general feeling that the number of lakes resulting from HEI's approach largely over-estimates the actual number of lakes that are present in their states. This concern is particularly true for those areas in ecoregion 43 and especially in South Dakota. Staff felt that the root of this concern may be that NHD includes stock ponds on ephermal ponds in their dataset, which results in an inflated number of lakes present in our database. One suggested solution to this problem was that HEI could define a larger minimum surface area for screening lakes, potentially applying the new criteria only to those lakes in ecoregion 43.

Figure 1a shows the distribution of lakes across the entire study area. This plot shows that a large number of lakes are between 10 and 15 acres in size and another large number are between 15 and 30 acres.

Figure 1a: Histogram of Study Area Lakes

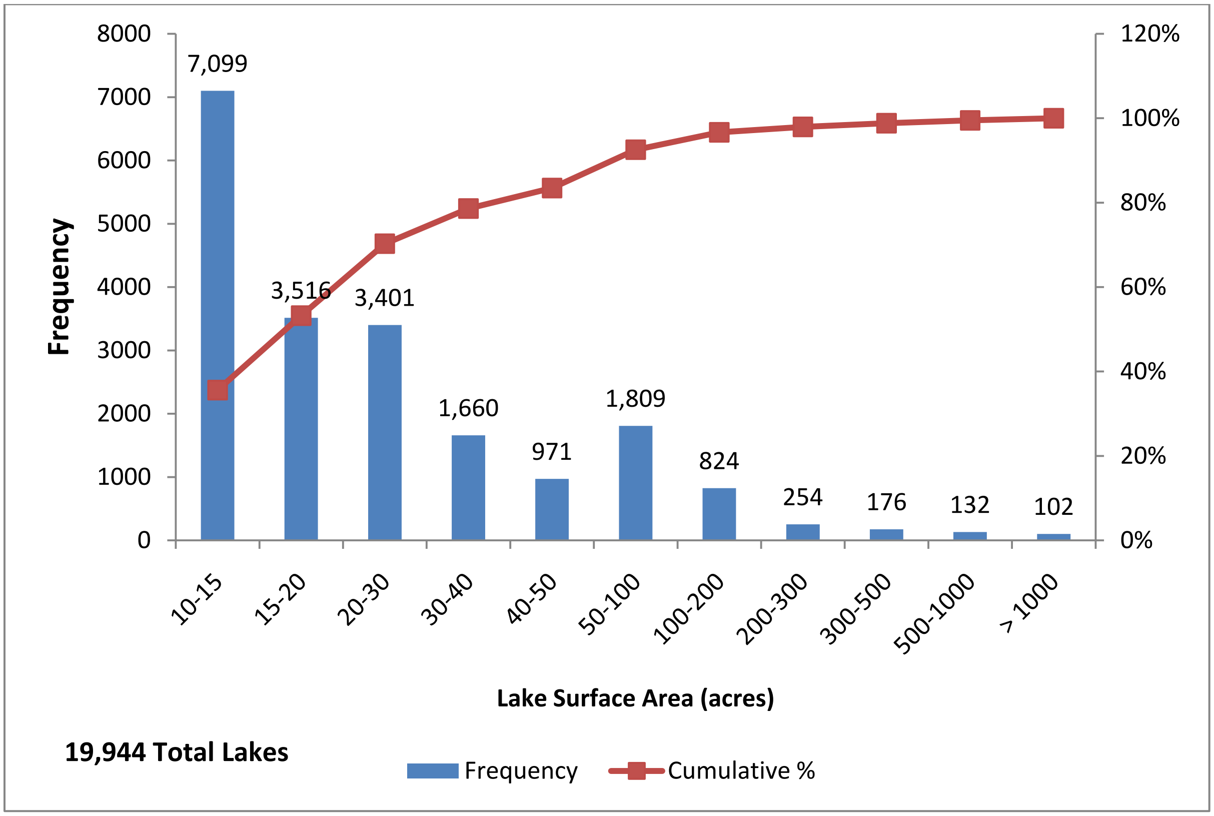
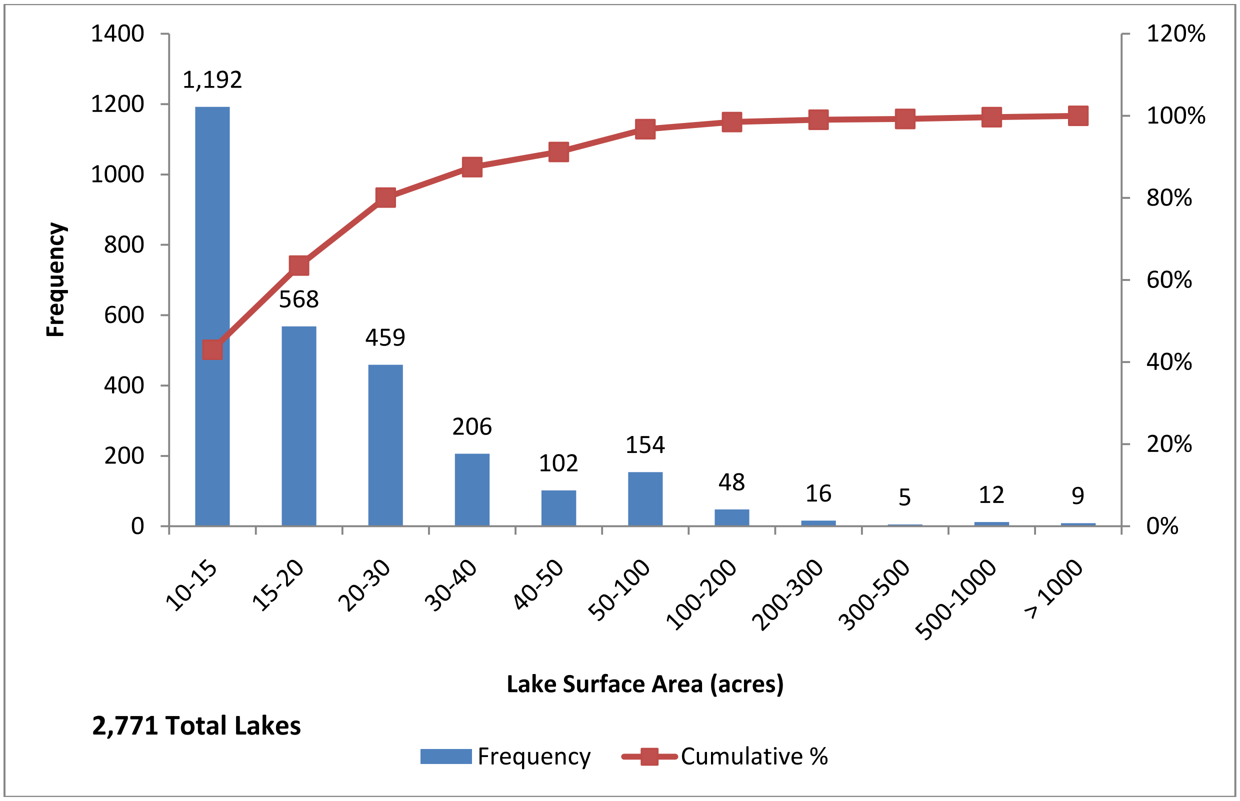


Figure 2a shows the distribution of lake areas in ecoregion 43. This plot shows a similar trend to that seen in **Figure 1a**.

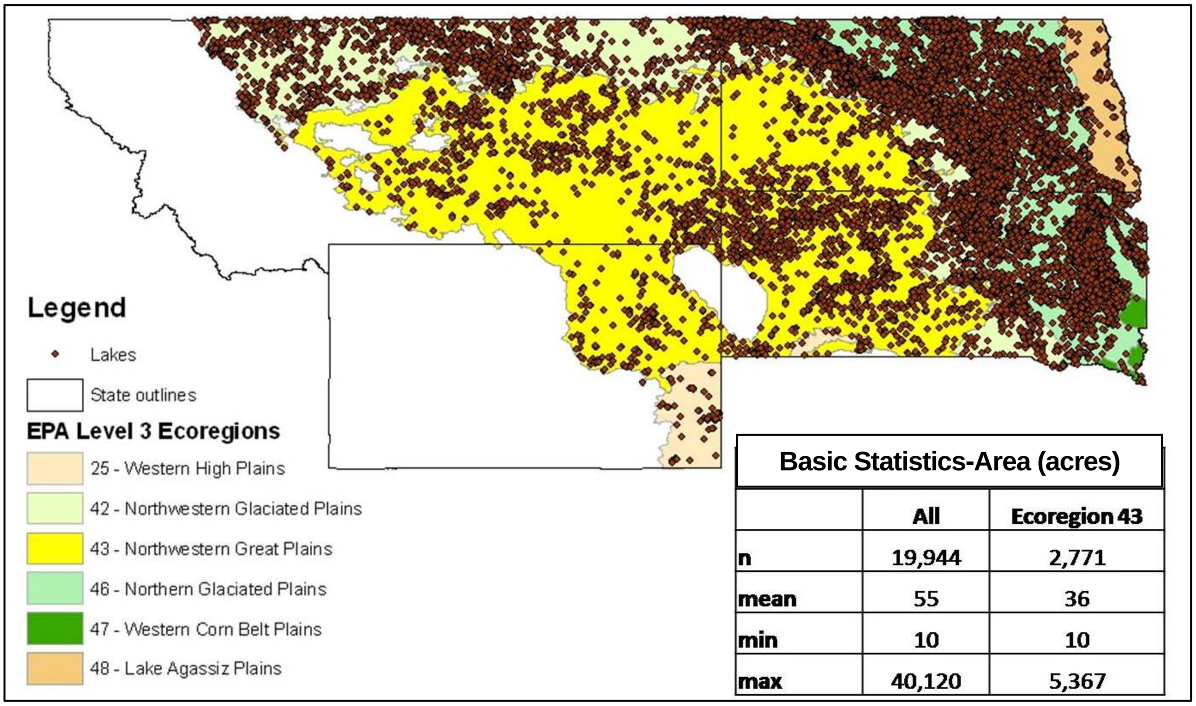
Figure 2a: Histogram of Study Area Lakes in Ecoregion 43



Based on these histograms, it appears that one option for further screening the lakes would be to screen out lakes with surface areas <15 acres. The following figures show what that level of screening would do to the lakes both geospatially and in basic statistics.

Figure 3a shows all of the lakes currently contained in the Lakes Master Database (i.e., lakes >10 acres).

Figure 3a: Contents of the Lakes Master Database (Lakes >10 acres)

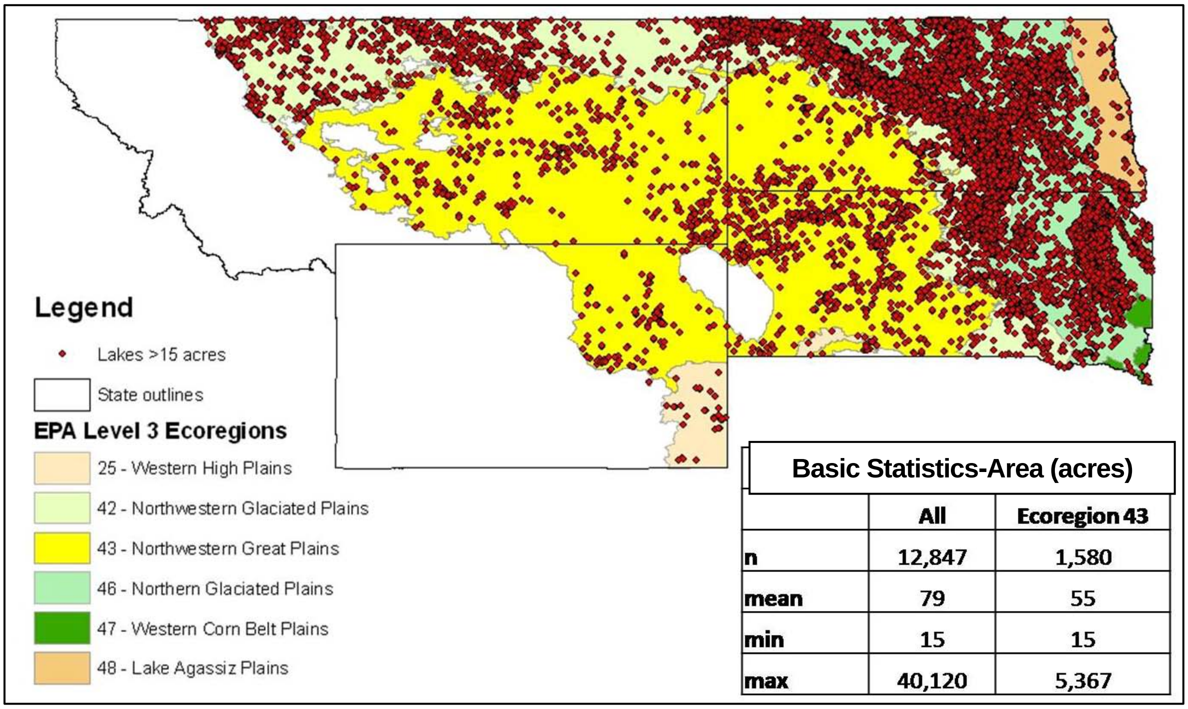


Basic Statistics-Area (acres)

	All	Ecoregion 43
n	19,944	2,771
mean	55	36
min	10	10
max	40,120	5,367

Figure 4a shows the lakes that would remain if the screening criteria were reset at a minimum surface area of 15 acres.

Figure 4a: Contents of the Lakes Master Database using Criteria of Lakes >15 acres

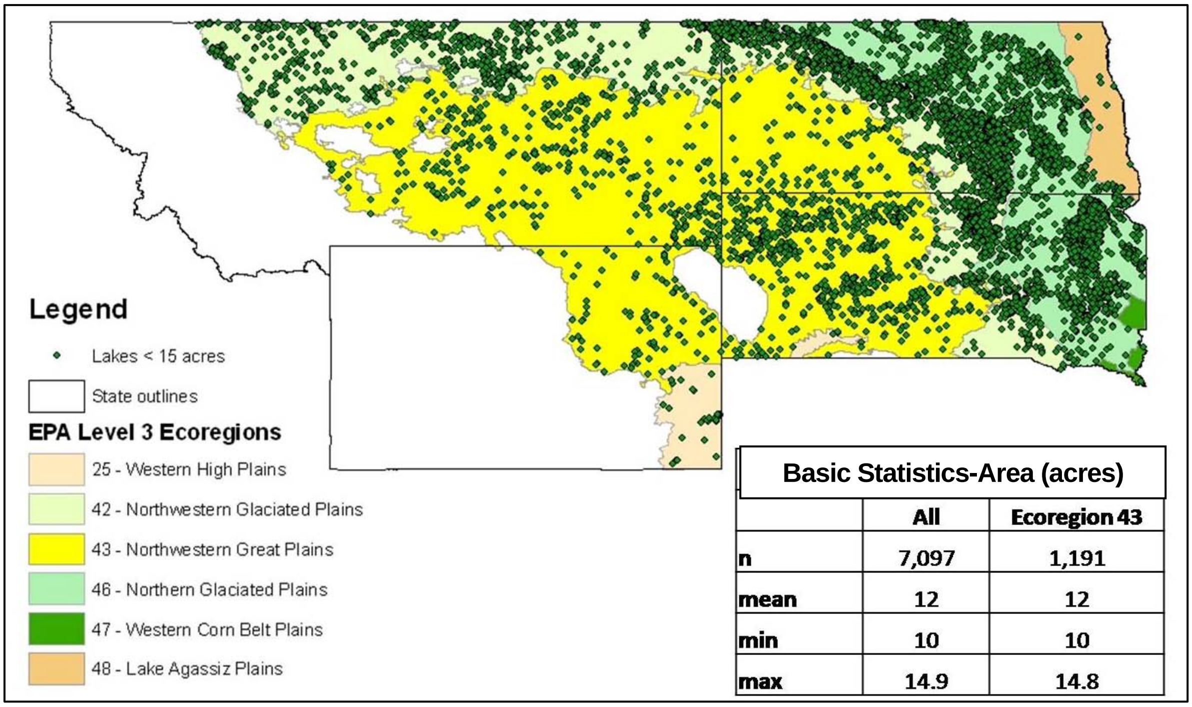


Basic Statistics-Area (acres)

	All	Ecoregion 43
n	12,847	1,580
mean	79	55
min	15	15
max	40,120	5,367

Figure 5a shows the lakes that were screened out by resetting the minimum surface area criteria (i.e., those lakes that are <15 acres).

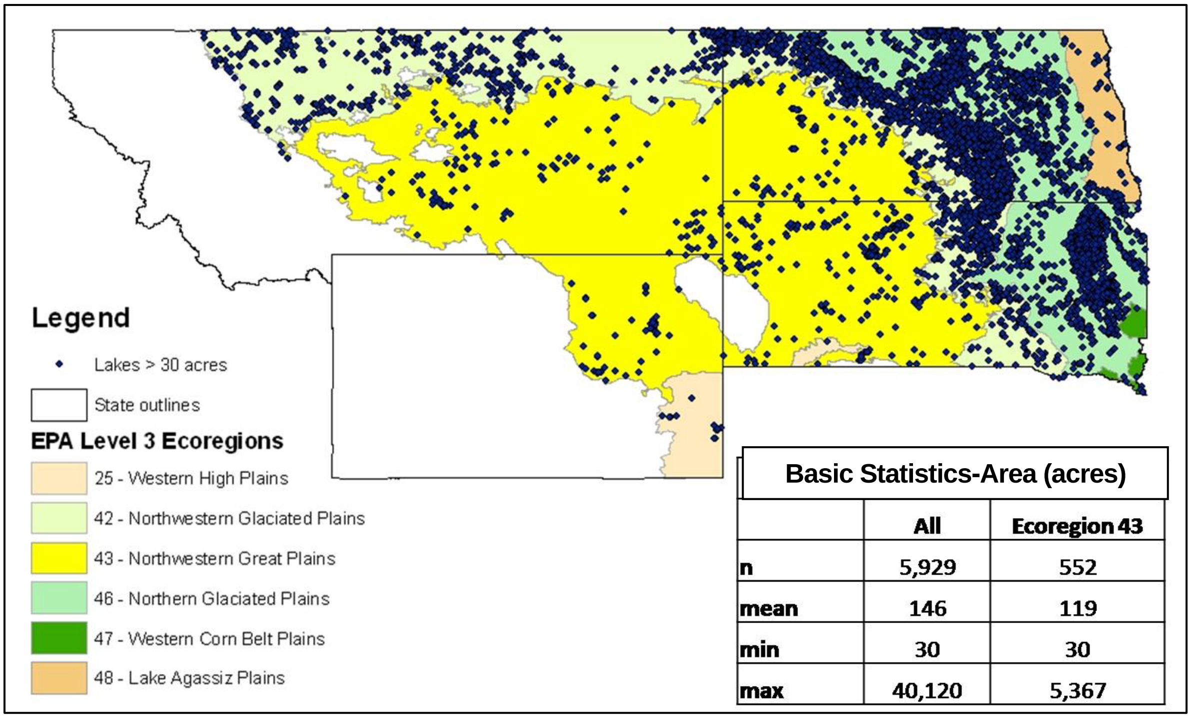
Figure 5a: Master Database Lakes that are <15 acres



Basic Statistics-Area (acres)		
	All	**Ecoregion 43**
n	7,097	1,191
mean	12	12
min	10	10
max	14.9	14.8

Another option may be to set the screening criteria at 30 acres. **Figure 6a** and **Figure 7a** show the results of resetting the minimum surface area screen at 30 acres.

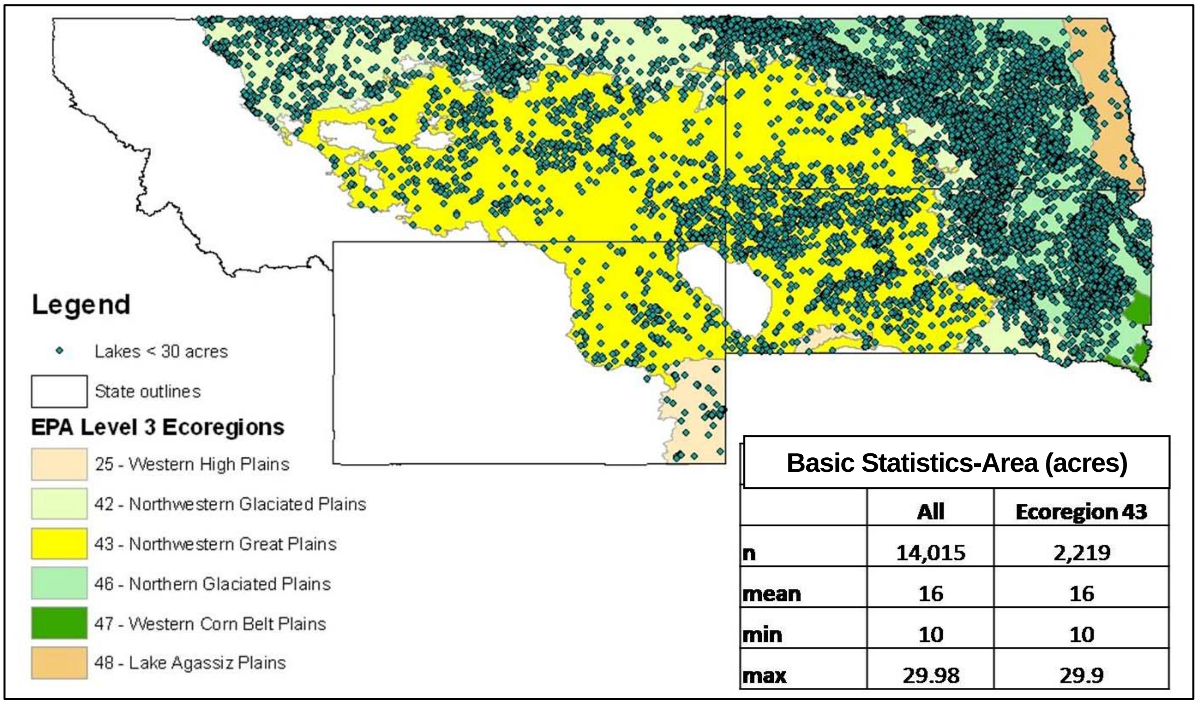
Figure 6a: Contents of the Lakes Master Database using Criteria of Lakes >30 acres



Basic Statistics-Area (acres)

	All	Ecoregion 43
n	5,929	552
mean	146	119
min	30	30
max	40,120	5,367

Figure 7a: Master Database Lakes that are <30 acres



Basic Statistics-Area (acres)

	All	Ecoregion 43
n	14,015	2,219
mean	16	16
min	10	10
max	29.98	29.9

State staff should now use this information to determine of using another minimum surface area screening criteria will help the Lakes Master Database to better reflect the number of lakes that are actually within their state.

Recommendations

No recommendation is explicitly given related to re-defining the minimum lake surface area screening criteria. Since State Agency are more familiar with their state's resources, it is left to these staff to use the information in **Figure 1a** through **Figure 7a** to determine if a different minimum surface area screening criteria would result in a more accurate set of lake data for their states. Depending on the results of this determination, the approach to defining lakes for inclusion in the lake master database could be changed. Although altering the surface area threshold reduces the number of water bodies defined as "lakes" the basic issue of the lack of information about lake volume and drainage area prohibits further analysis. We recommend developing a stratified random sampling approach using the lake surface area and ecoregions to generate the drainage area and volume (or mean depth to compute volume using surface area) data needed to complete the classification. Lakes with actual water quality data could be weighted in the sampling approach.

Section IV: Reservoir Water Quality Data

Section IV: Reservoir Water Quality Data

MEMO

(External Correspondence)

From: Stephanie Johnson, Ph.D., P.E.

Through: Mark R. Deutschman, Ph.D., P.E.

To: Tina Laidlaw

Date: September 20, 2010
(original memo: April 2, 2010)

Cc: File 4965-002
Dennis McIntyre, GLEC

Subject: Revised report on activities related to water quality data under Tasks 2 & 3 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8

This memo addresses the analysis of water quality data as described in Tasks 2 and 3 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8 (i.e. the Nutrient Criteria project). We discuss the data that was provided, the data that was of interest, how Houston Engineering, Inc. (HEI) managed the data, a summary of the available data, and the statistical analyses that were performed. The purpose of this memorandum is to use the available water quality data to assess whether the reservoir classification process is applicable to the Plains States of EPA Region 8 and to determine whether there are sufficient data to complete the water quality modeling in support of establishing nutrient criteria. **This version of the memo is revised based on comments received on the original, dated April 2, 2010.**

Water Quality Data Provided and Analyzed

In earlier phases of the Nutrient Criteria project, HEI solicited the States of North Dakota, South Dakota, Wyoming, and Montana to provide their water quality data for lakes and reservoirs. Of particular interest was data related to developing nutrient criteria for the states, including total phosphorus (TP), total nitrogen (TN), chlorophyll-a (Chl-a), and secchi depth. North Dakota, South Dakota, and Wyoming provided HEI with a spreadsheet or database of the requested water quality data (HEI, 2010). Montana had recently discovered an error in their water quality master database and could not, therefore, provide data of sufficient quality for this study. The databases/spreadsheets provided by ND, SD, and WY contained water quality records for a number of different parameters collected in lakes and reservoirs for dates ranging from 1977 through 2009. For the purposes of this project, HEI focused on those parameters (i.e., stressors and response variables) related to establishing nutrient criteria: TP, TN, Chl-a, and secchi depth. HEI used these data to compute Carlson Trophic State Index (TSI) values based on the TP concentration, the Chl-a concentration, and the secchi depth.

Earlier tasks in this project had HEI assess the morphometry data that is available for lake and reservoirs in the study area. The results of that assessment showed that insufficient data are available for lakes in the area, limiting our ability to eventually model the water quality of those waterbodies (which is the next step in this project). Due to the lack of data, HEI was directed (by EPA) to set the lake data to the side for now and perform the analyses of water quality data only for the area's reservoirs. The remainder of this memo, therefore, pertains only to the water quality data from reservoirs in the reservoir master database, as discussed in the March 5, 2010 project memo (HEI, 2010).

Management of the Data Sources

The first step in the water quality analysis was to combine the states' water quality datasets into a single master dataset and trim out the unnecessary data. Unnecessary data was defined as water quality parameters not related to nutrients and data from reservoirs that are not in the reservoir master database, which was created by HEI during the lake/reservoir classification step (addressed in (HEI, 2010)). A major part of the data management effort was to interpret the water quality data and ensure that it matched between sources. For example, each State had a slightly different way of reporting nitrogen values. For the most part, TN values were not explicitly reported. When TN was not available explicitly in the database, we summed the Total Kjeldahl Nitrogen (TKN) and Nitrite (NO_2) + Nitrate (NO_3) concentrations to estimate the TN. In the absence of NO_2+NO_3 data, TKN values were used to estimate TN (performed for 107 of the 3,787 estimated values). This estimation was deemed "tolerable" since the data show that NO_2+NO_3 are, generally, about 10% of the TKN concentration.

Another common issue in the data was values being reported as being below the detection limit (i.e. 'non-detect' values). Approximately 40% of Chl-a values were 'non-detects'; about 20% of TN values were computed from non-detect values. Less than 10% of TP values had this problem. Numerous procedures exist for dealing with non-detect data, a common problem in analyzing water quality data. For this work, non-detect values were estimated as half of the detection limit. In addition to non-detect data, there were also a number of values reported as "0". For parameters that also had detection limits recorded, zero values were analyzed similar to non-detects; we assumed the value was equal to half of the lowest detection limit reported. When no detection limits were given, however, (mainly for secchi depth and Chl-a data in SD), zero values were disregarded for our calculations.

Data Summary

Water quality data were provided for 178 of the reservoirs contained in the reservoir master database. **Table 1** summarizes the number of reservoirs with water quality data by ecoregion, state, and classification tier. **Table 2** provides a more complete picture, detailing the location, reservoir classification tier, and period of record of each reservoir. It also shows the count and descriptive statistics of the various water quality parameters measured. Note that some reservoirs have only a few samples for each parameter, which may limit their validity for use in the future water quality modeling task of this project. Identifying the reservoirs that are desirable for modeling (i.e., those with a "significant" number of samples and morphometric data) will be done in a later step.

Table 1: Reservoirs with Water Quality Data by Ecoregion, State, and Classification Tier

EPA Level 3 Ecoregion[1]	25	42	43	46	47	48
# of reservoirs with water quality data	3	39	74	56	1	5
State	**North Dakota**		**South Dakota**		**Wyoming**	
# of reservoirs with water quality data	87		88		3	
Reservoir Classification Tier[2]	**I**	**II**	**III**	**IV**		
# of reservoirs with water quality data	116	36	14	12		

[1] EPA Tier 3 Ecoregions: 25 = Western High Plains; 42 = Northwestern Glaciated Plains; 43 = Northwestern Great Plains; 46 = Northern Glaciated Plains; 47 = Western Corn Belt Plains; 48 = Lake Agassiz Plain
[2] The Reservoir Classification Tier is computed as [(Surface area/Drainage area)*Volume], as discussed in the March 5, 2010 memo on the topic (HEI, 2010). The theory behind the classification is that the eutrophication response of the reservoirs will be unique by Tier.

Statistical Analysis

To understand the nature of the nutrient data, descriptive statistics were computed for each of the parameters and for the TSIs. **Table 3** summarizes the results, which show (among other things) that the distributions of TP, Chl-a, and secchi depth are right-skewed and non-normal in distribution (discussed more below).

Table 3: Water Quality Data – Descriptive Statistics

Parameter	N	Mean	Standard Deviation	Min	1st Quartile	Median	3rd Quartile	Max
TP (mg/L)	8,438	0.27	0.62	0.001	0.044	0.16	0.36	48.70
TP TSI	8,438	67.61	23.58	0	54.37	73.26	84.54	155.49
TN (mg/L)	6,999	1.88	5.59	0.025	0.80	1.34	1.84	116.10
Chl-a (ppb)	3,480	27.70	55.84	0.057	3.00	10.30	29.79	676.0
Chl-a TSI	3,480	52.72	14.89	2.465	41.35	53.45	63.87	94.50
Secchi (m)	4,586	1.60	1.43	0.040	0.61	1.07	2.13	11.40
Secchi TSI	4,586	58.89	12.88	24.89	40.07	59.06	67.14	106.4

Data Analysis

<u>Relationships between Stressor and Response Variables</u>

The amount of nutrients in a system drives the ecological condition, including eutrophication. Of the water quality variables analyzed here, TP and TN are the stressor (independent) variables and Chl-a and secchi visibility are the response (dependent) variables (i.e., the amount of TP and/or TN in a system is expected to drive the value of Chl-a and/or secchi depth observed).

To gain insight on the eutrophication response of the region's reservoirs to nutrient levels, the relationships between stressor and response variables were evaluated using simple linear regression methods. The results of this analysis are important for use in the future modeling of in-reservoir water quality, as they will guide the creation of regional stressor-response relationships for use in the models. Six regressions were computed: TP vs. Chl-a; TP vs. secchi depth; TP vs. TN; TN vs. Chl-a; TN vs. secchi depth; and Chl-a vs. secchi depth. Given the non-normal nature of the distributions (**Table 3**), the regressions were performed on the natural logarithms of the data. **Figures 1 through 6** show the results.

Figure 1: Total Phosphorus vs. Chl-a (ln transformed)

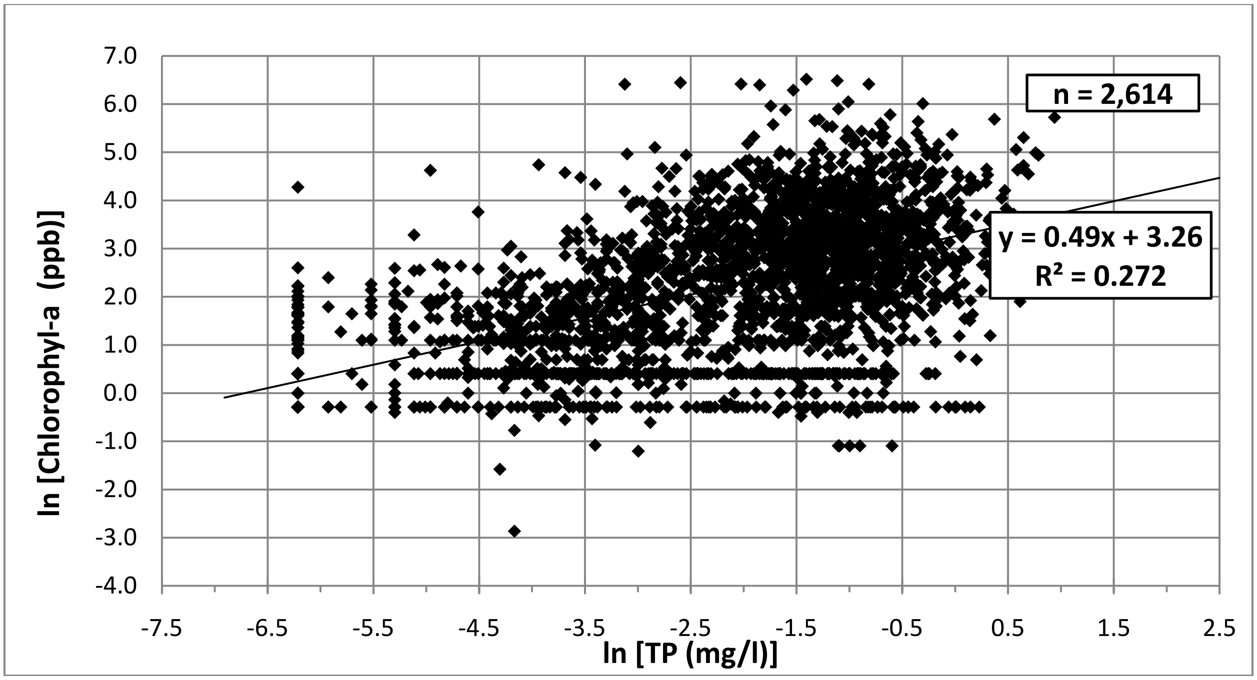


Figure 2: Total Phosphorus vs. Secchi Depth (ln transformed)

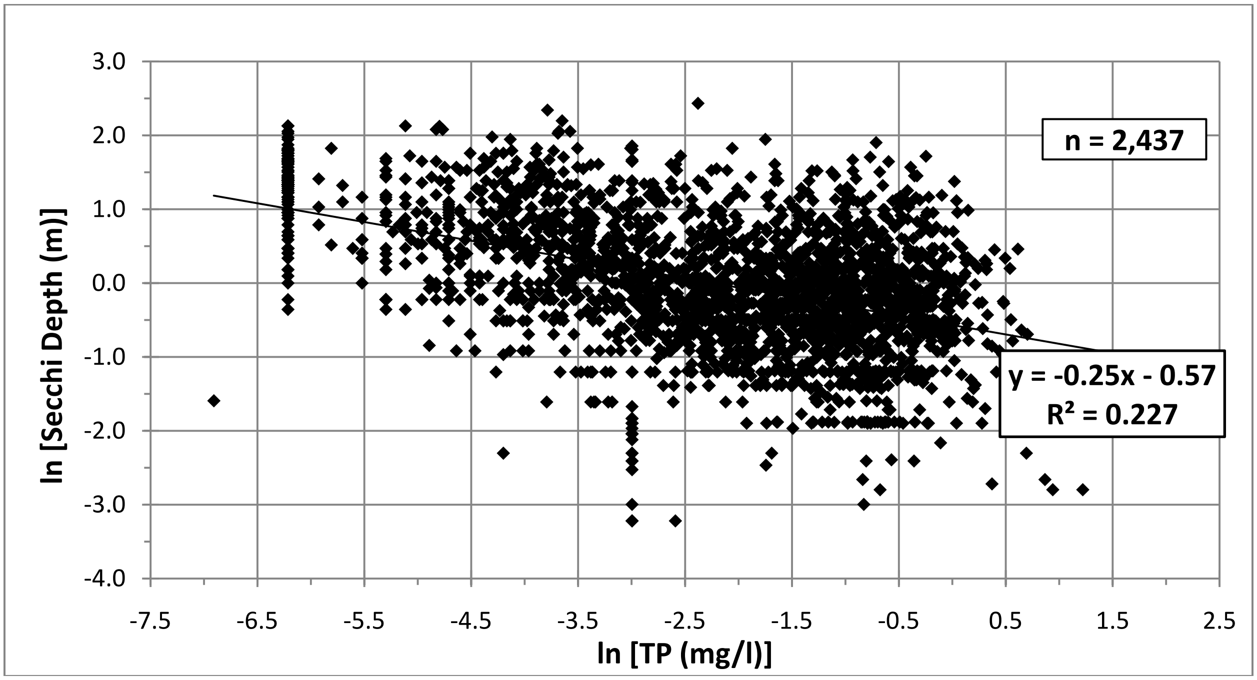


Figure 3: Total Phosphorus vs. Total Nitrogen (ln transformed)

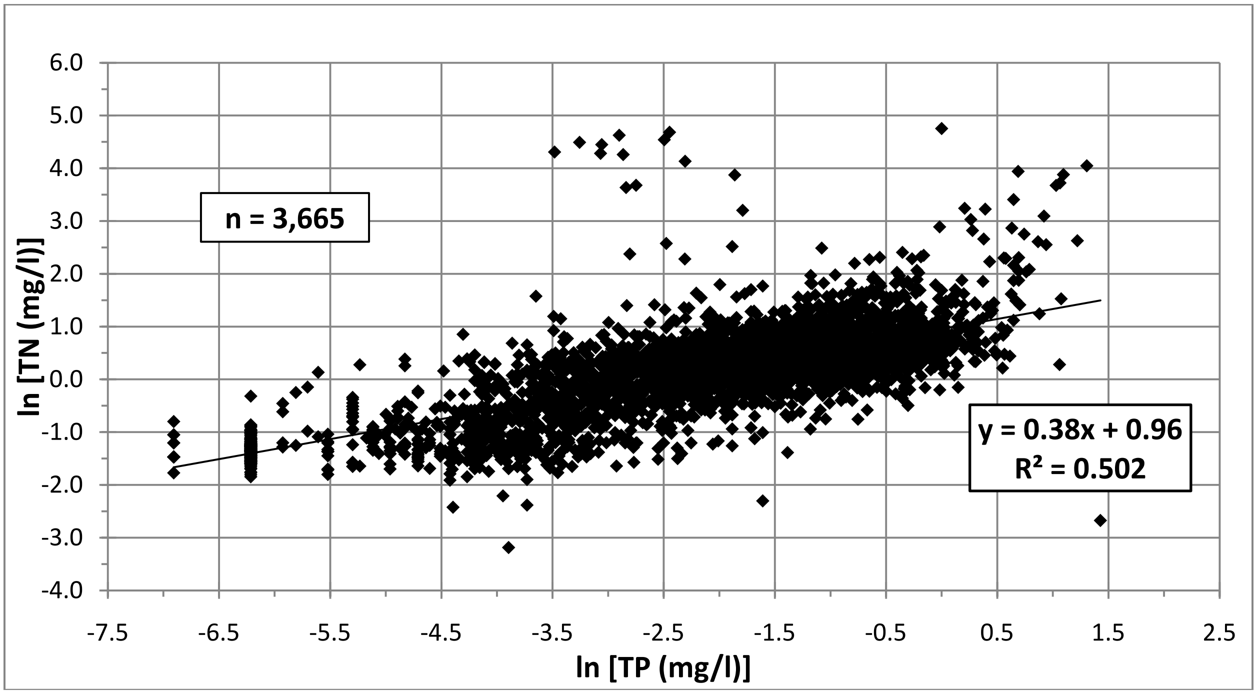


Figure 4: Total Nitrogen vs. Chl-a (ln transformed)

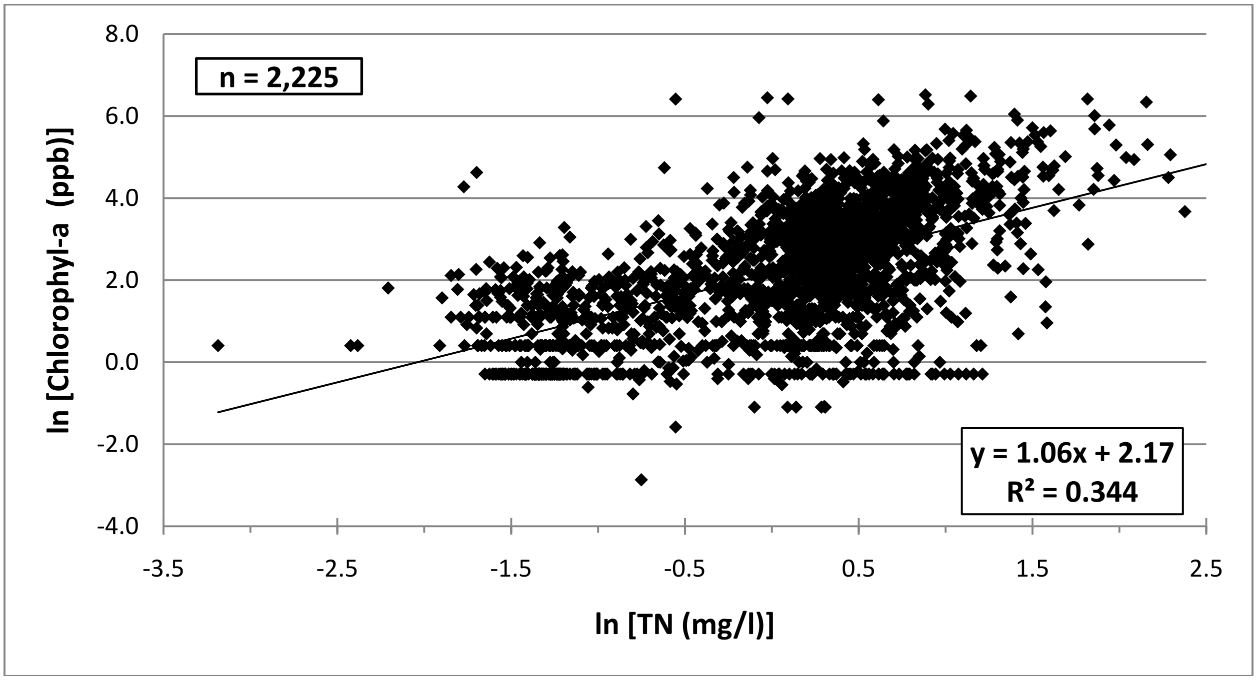


Figure 5: Total Nitrogen vs. Secchi Depth (ln transformed)

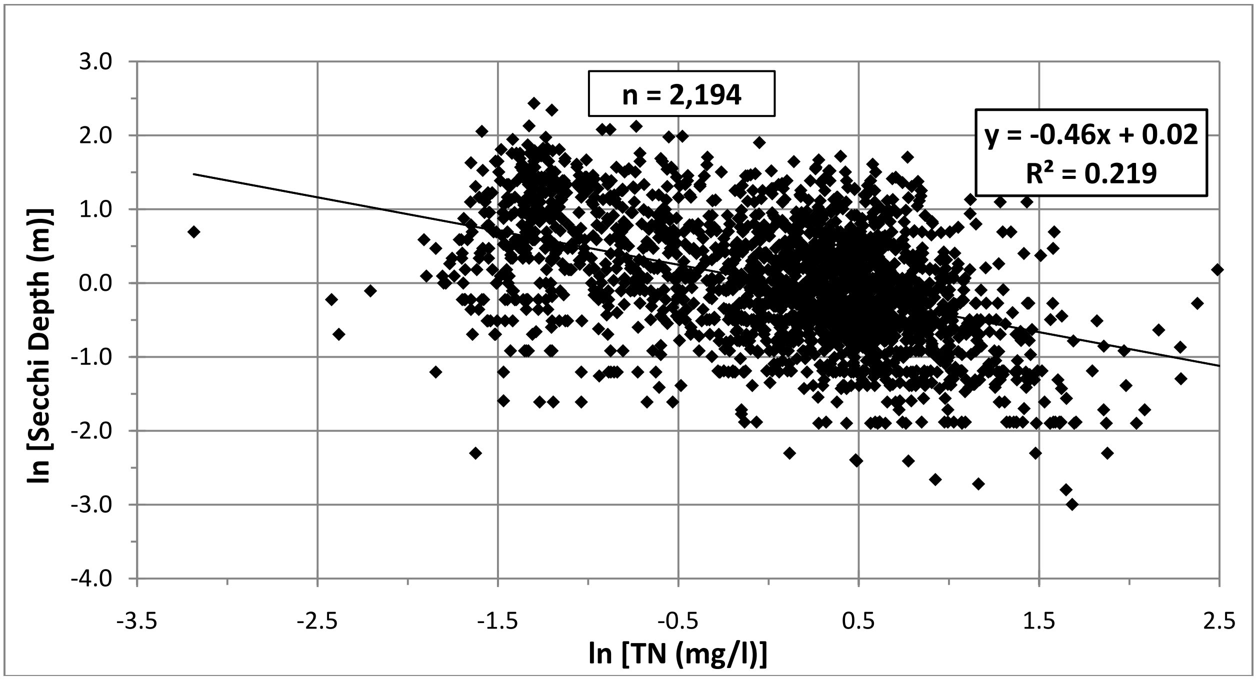
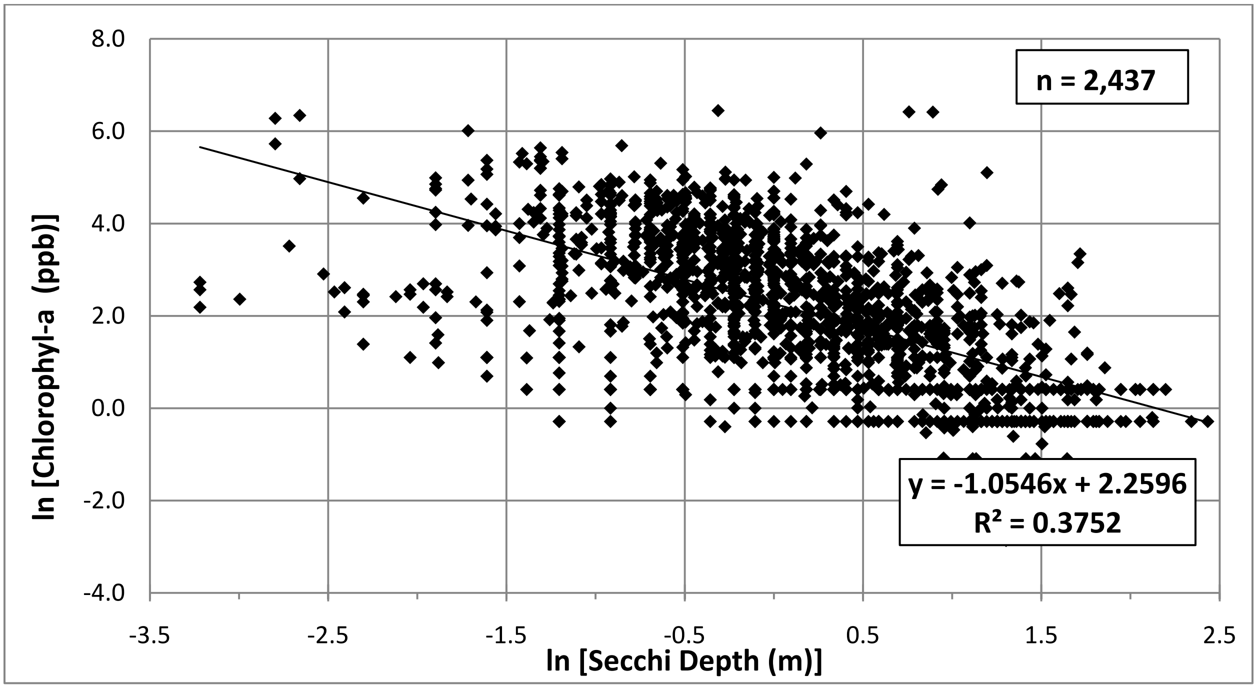


Figure 6: Secchi Depth vs. Chl-a (ln transformed)



The regression plots show a few obvious characteristics of the data sets that are noteworthy. Vertical and horizontal lines of data show the presence of detection limits in the datasets. For example, **Figure 1** shows a vertical line of data points at ln[TP] = -6.21. This line correlates to the 0.004 mg/l detection limit for TP, which we modeled as 0.002 mg/l (half of the detection limit – discussed above). The presence of multiple lines shows that detection limits in the water quality tests varied over time or by laboratory (e.g., a vertical line at ln[TP]=-5.30 correlates to another TP detection limit of 0.01 mg/l). HEI used the water quality data as provided by the states, without performing any additional quality assurance. As a result of our analyses, some potential data errors are highlighted. One example is shown in **Figure 1**, where 21 of the ln[Chl-a] data points lie below the Chl-a detection limit of 1.5 ppb. Given current limitations in laboratory analyses, it is likely that these data points are erroneous. However, including them in the linear regression will not have a significant impact on the results (as they represent 0.8% of the observations). Another potential error in the data is shown in **Table 3**, where the maximum TP value is report as 48.7 mg/l (resulting in a computed maximum TP TSI of 155.49). This value is an order of magnitude greater than the next largest value; including it in our analysis, however, does not significantly affect the results.

Typical of data from natural systems, the regressions in **Figures 1 through 6** show considerable scatter about the best fit line. The general trend of the regressions, however, show the expected result where Chl-a concentrations increase with an increase in TP and secchi depths decrease as nutrient levels increase. **Figure 6** shows the expected trend of decreasing secchi depths as Chl-a concentrations rise. The results in **Figure 3** show that reservoir TN:TP ratios vary widely within the dataset showing an overall average ratio of 2.5. Given the amount of variability, the potential for phosphorus, nitrogen, and phosphorus/nitrogen limitation should be considered in future work.

Comparing Statistical Distributions by Reservoir Classification Tier

One goal of this project is to understand if there is a way to place the reservoirs into groups based on their eutrophication response. If this grouping can be achieved, unique nutrient criteria may be appropriate for each group. In theory, the reservoir classification tier (computed as [(Surface area/Drainage area)*Volume]) was developed as an approach to grouping waterbodies based on eutrophication response. Other options for grouping may be by state or EPA ecoregion.

A qualitative analysis of the impact of grouping can be done through the use of box and whisker plots. These plots show the distribution of the water quality parameter by group. **Figures 7 through 13** show these plots, comparing nutrient and TSI values across reservoir classification tiers. Plots were also made for groups by ecoregion and state; these plots are included as an **Appendix** to this memo.

Figure 7: Distribution of Total Phosphorus Values by Reservoir Classification Tier

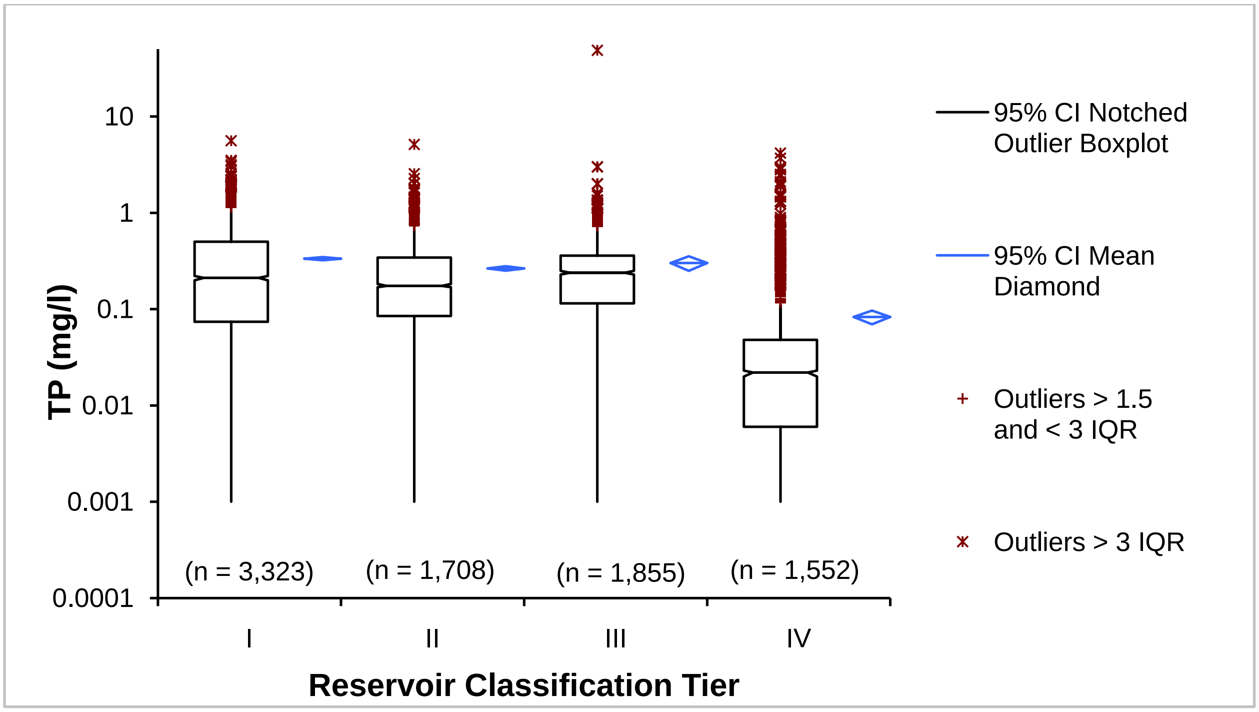


Figure 8: Distribution of Total Phosphorus TSI Values by Reservoir Classification Tier

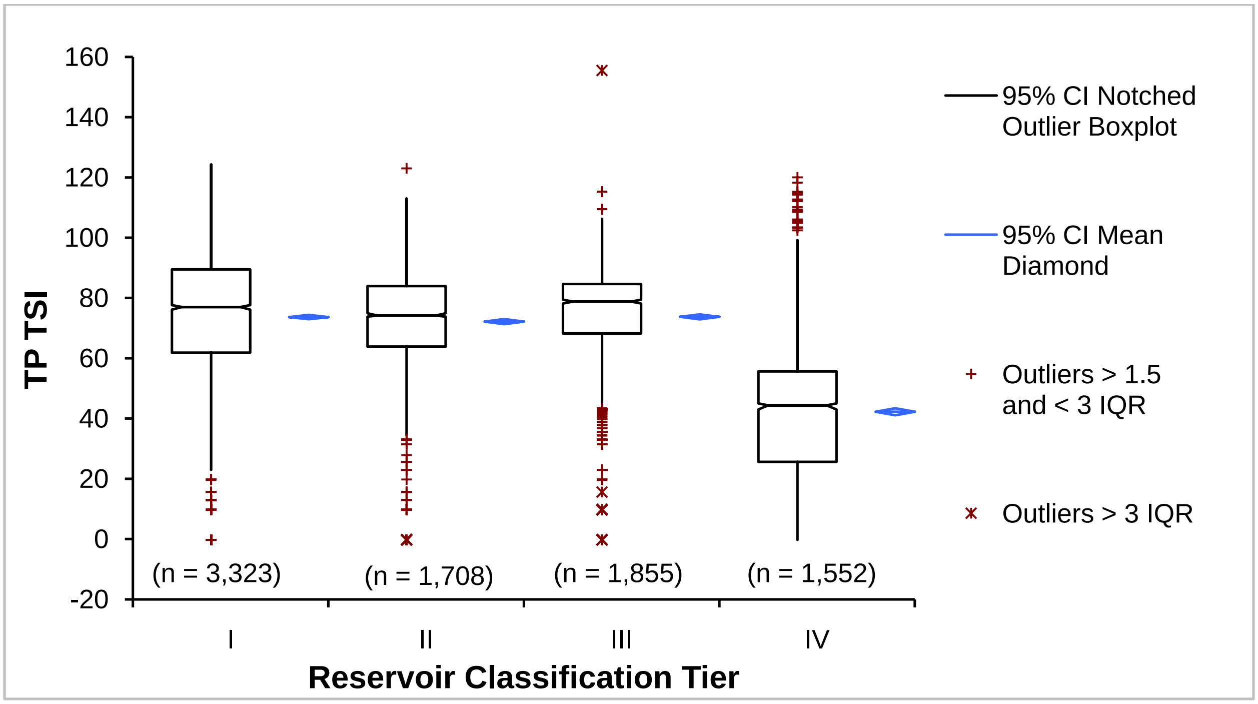


Figure 9: Distribution of Total Nitrogen Values by Reservoir Classification Tier

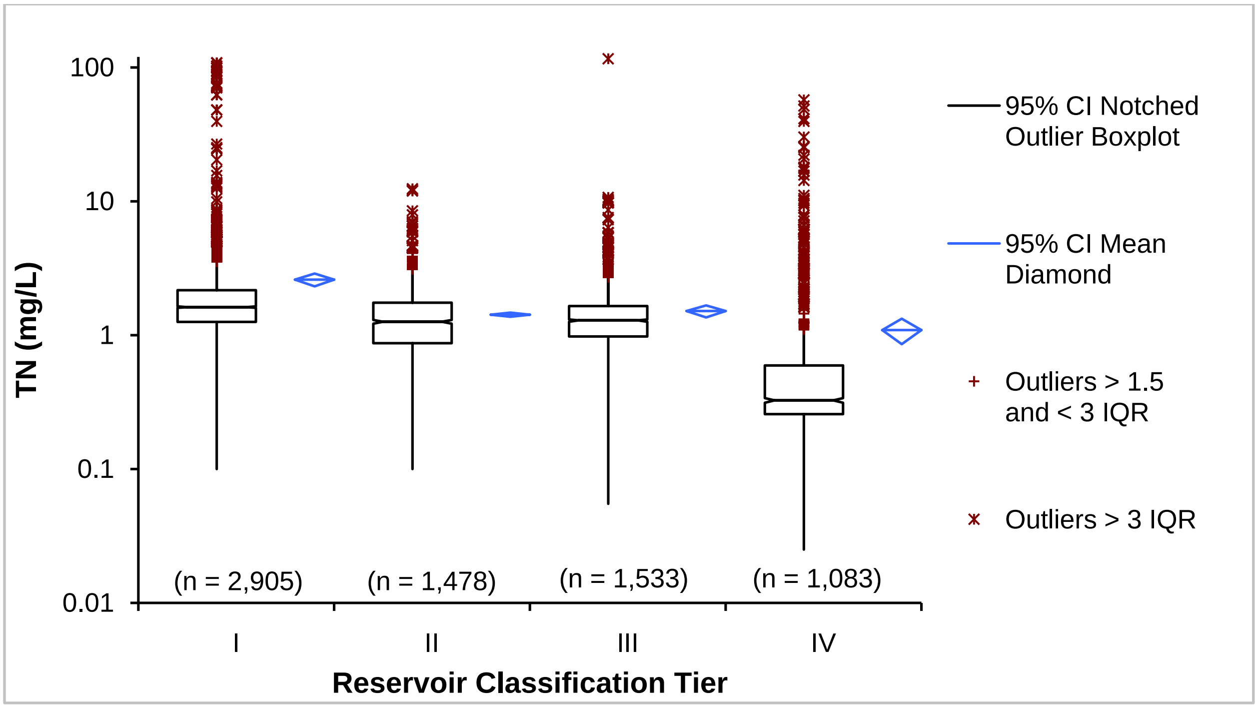


Figure 10: Distribution of Chl-a Values by Reservoir Classification Tier

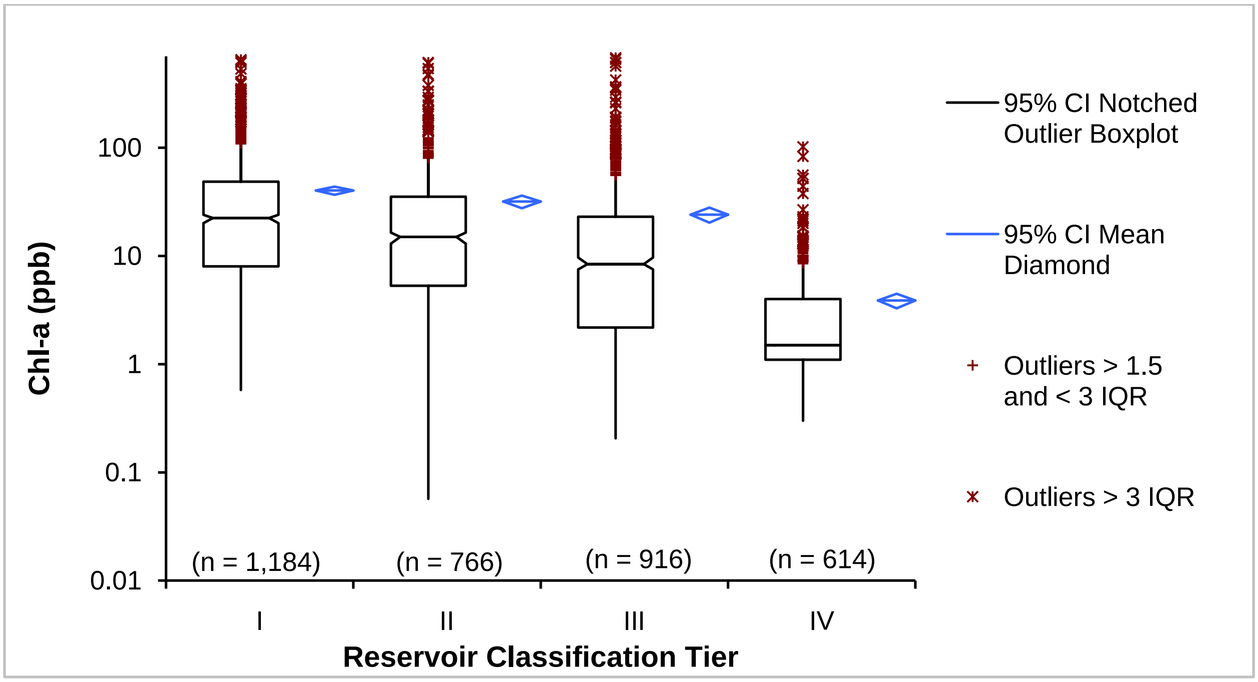


Figure 11: Distribution of Chl-a TSI Values by Reservoir Classification Tier

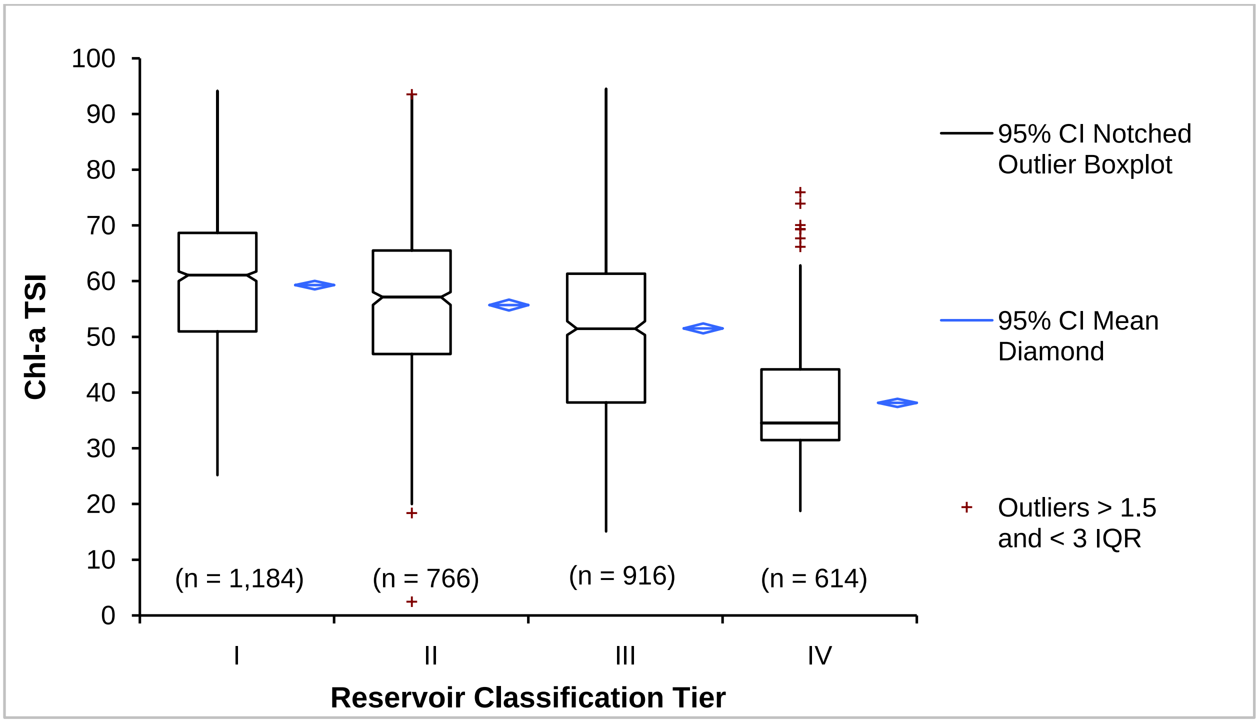


Figure 12: Distribution of Secchi Depth Values by Reservoir Classification Tier

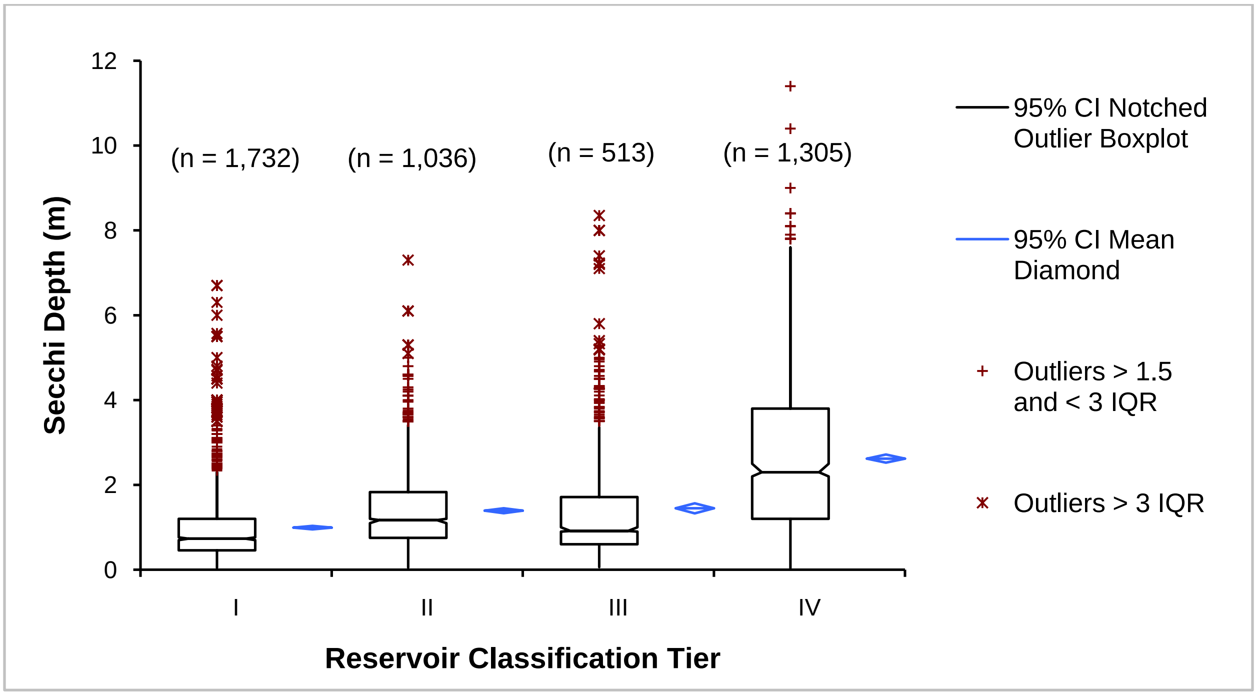
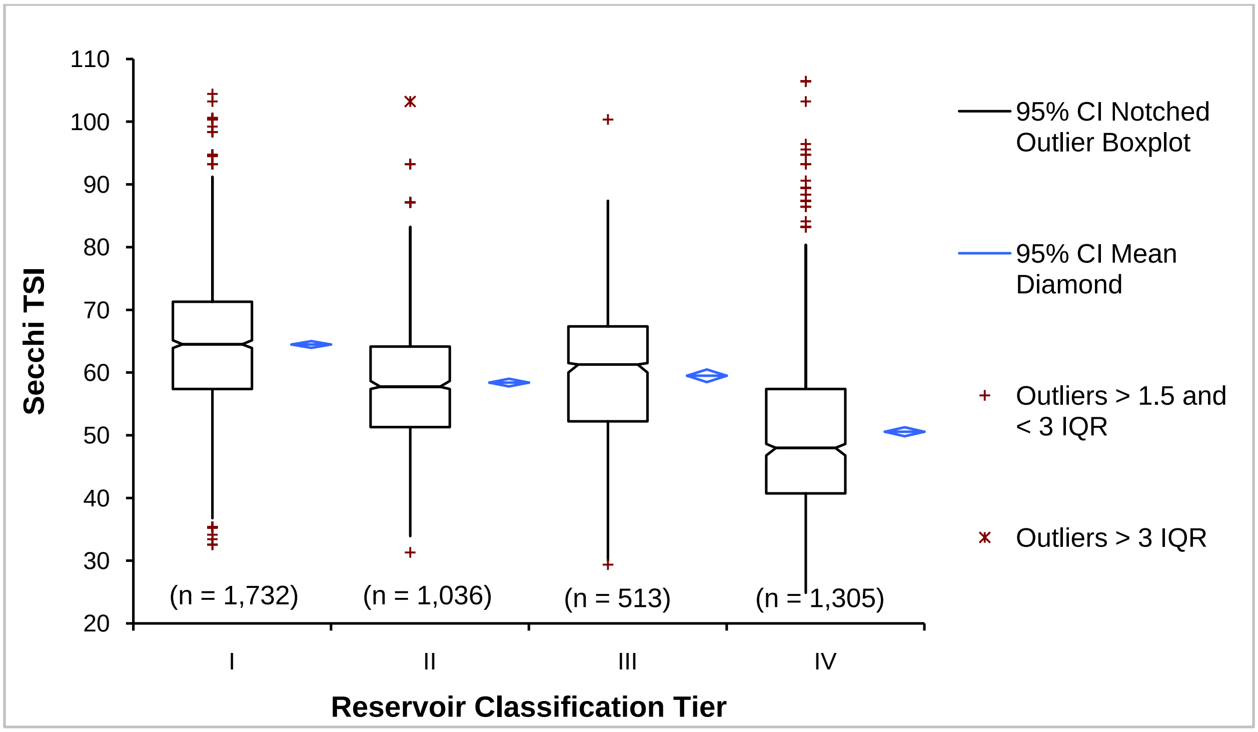


Figure 13: Distribution of Secchi Depth TSI Values by Reservoir Classification Tier



Analysis of Variance – One-Way Testing

Figures 7 through 13 give a qualitative sense of the differences between nutrient and TSI values amongst reservoir classification tiers. To formalize that analysis, we used statistical tests to analyze the variance and determine if the distributions are statistically significantly different. Analysis of variance is a statistical method that considers multiple data sets, categorized by group, and tests whether or not the statistical distributions of the data within each group are statistically significantly equal to one another (often by assuming that the variances are equivalent and testing the mean or median). As applied in this work, the analysis tests whether the distributions of nutrient data amongst reservoir classification tiers are statistically significantly different, as theorized by the approach. To determine which analysis of variance test is most appropriate for this work, the data sets were first tested to see if they could be described using a normal distribution. Normal probability plots and results of the Shapiro-Wilk Test indicate that the distributions of TP, TN, Chl-a, and secchi depth are not normal (as alluded to above). A non-parametric statistical test is, therefore, preferred.

The Kruskal-Wallis Test is a non-parametric statistical test that analyzes different groups of data to see if the median ranks of each group are statistically significantly different from one another. This test was used to compare TP, TN, Chl-a, and secchi depth data between reservoir classification tiers. Results of the analyses are shown in **Tables 4 through 7**.

Table 4: Results of Kruskal-Wallis Test for TP by Classification Tier

TP (mg/L) by Classification Tier	n	Rank sum	Mean rank
1	3,323	16,081,776.5	4,839.54
2	1,708	7,751,798.0	4,538.52
3	1,855	8,963,962.5	4,832.32
4	1,552	2,806,604.0	1,808.38
Kruskal-Wallis' statistic	1,882.65		
X^2 statistic	1,882.65		
DF	3		
p	<0.0001	(chisqr approximation, corrected for ties)	
Bonferroni Contrast	Difference	p	
1 v 2	301.01	<0.0001	
1 v 3	7.21	5.4466	
1 v 4	3,031.16	<0.0001	
2 v 3	-293.80	0.0003	
2 v 4	2,730.14	<0.0001	
3 v 4	3,023.95	<0.0001	

Table 5: Results of Kruskal-Wallis Test for TN by Classification Tier

TN (mg/L) by Classification Tier	n	Rank sum	Mean rank
1	2,905	12,680,489.5	4,365.06
2	1,478	5,072,159.0	3,431.77
3	1,533	5,314,529.5	3,466.75
4	1,083	1,429,322.0	1,319.78
Kruskal-Wallis' statistic	1,795.46		
X^2 statistic	1,795.46		
DF	3		
p	<0.0001	(chisqr approximation, corrected for ties)	

Bonferroni Contrast	Difference	p
1 v 2	933.28	<0.0001
1 v 3	898.31	<0.0001
1 v 4	3,045.28	<0.0001
2 v 3	-34.98	3.4914
2 v 4	2,111.99	<0.0001
3 v 4	2,146.97	<0.0001

Table 6: Results of Kruskal-Wallis Test for Chl-a by Classification Tier

Chl-a (ppb) by Classification Tier	n	Rank sum	Mean rank
1	1,184	2,596,901.5	2,193.33
2	766	1,490,258.5	1,945.51
3	916	1,510,082.0	1,648.56
4	614	459,698.0	748.69
Kruskal-Wallis' statistic	880.02		
X^2 statistic	880.02		
DF	3		
p	<0.0001	(chisqr approximation, corrected for ties)	

Bonferroni Contrast	Difference	p
1 v 2	247.82	<0.0001
1 v 3	544.77	<0.0001
1 v 4	1,444.64	<0.0001
2 v 3	296.95	<0.0001
2 v 4	1,196.81	<0.0001
3 v 4	899.87	<0.0001

Table 7: Results of Kruskal-Wallis Test for Sechhi Depth by Classification Tier

Chl-a (ppb) by Classification Tier	n	Rank sum	Mean rank
1	1,732	2,899,425.0	1,674.03
2	1,036	2,406,833.0	2,323.20
3	513	1,104,191.0	2,152.42
4	1,305	4,107,542.0	3,147.54
Kruskal-Wallis' statistic	928.64		
X^2 statistic	928.64		
DF	3		
p	<0.0001	(chisqr approximation, corrected for ties)	
Bonferroni Contrast	Difference	p	
1 v 2	-649.17	<0.0001	
1 v 3	-478.39	<0.0001	
1 v 4	-1,473.51	<0.0001	
2 v 3	170.78	0.0450	
2 v 4	-824.34	<0.0001	
3 v 4	-995.12	<0.0001	

Results of the Kruskal-Wallis test are summarized in the resultant p-value, which indicates the likelihood that the analyzed data would be observed by chance alone if the null hypothesis (in this case, that the distributions of data amongst the tiers are all equal) were true. In all cases, the tests resulted in a p-value of <0.0001 (i.e., there is less than a 0.1% chance of observing the data that we have if the distributions amongst the tiers are statistically equivalent), indicating that at least one of the distributions of TP, TN, Chl-a, and secchi depth is statistically significantly different from that in the other reservoir classification tiers. Similar results were seen when comparing TP, TN, Chl-a, and secchi amongst ecoregions and states; the distributions of the data were found to be statistically significantly different.

To determine if all of the distributions are different from one another, a series of pairwise Kruskal-Wallis tests were performed using the Bonferroni approach. Results of these analyses are also expressed through a p-value and shown in **Tables 4-7**. The pairwise analysis showed that Tier 1 and 3 TP distributions are not statistically significantly different from one another. It also showed that Tier 2 and 3 TN distributions are not statistically significantly different. All other combinations of Tiers for the parameters were shown to be different.

<u>Analysis of Response Variables by Mean Phosphorus Concentrations</u>

A final analysis was performed to explore how the distributions of Chl-a and secchi depth vary with the mean reservoir TP concentration. Results of this analysis are important as they give insight into options for developing nutrient criteria. For example, earlier analyses showed that the concentration of Chl-a and secchi depth of a reservoir depends on the amount of TP present. If the goal of the nutrient criteria is to control Chl-a levels, one approach to doing so may be to set a nutrient criteria for the maximum mean TP value in a reservoir. In order to set an appropriate mean TP value, however, it would be valuable to understand the Chl-a concentrations that could be expected to result from those TP levels. The following analysis gives insight to these expectations by plotting the distribution of Chl-a and secchi depth as a function of different mean TP categories. The analysis was performed for all of the reservoir data lumped together, as well as for the data separated by reservoir classification tier. Results are presented only for all of the data lumped.

Figure 14 shows a box and whisker plot of Chl-a concentrations grouped by mean TP concentration. Mean TP concentration bin ranges were set based on Carlson's TSI method and the correlating eutrophic state of the waterbody. **Figure 15** shows a similar plot for secchi depths by mean TP.

Figure 14: Distribution of Chl-a by Mean Reservoir TP – All Data

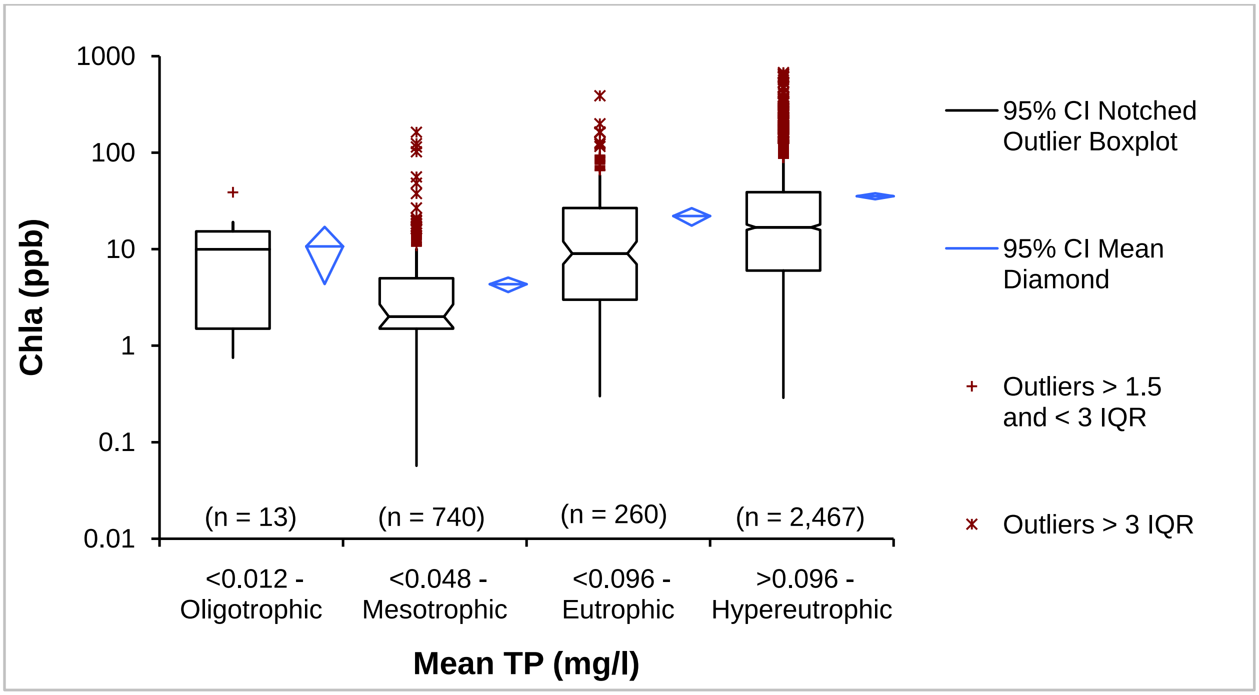
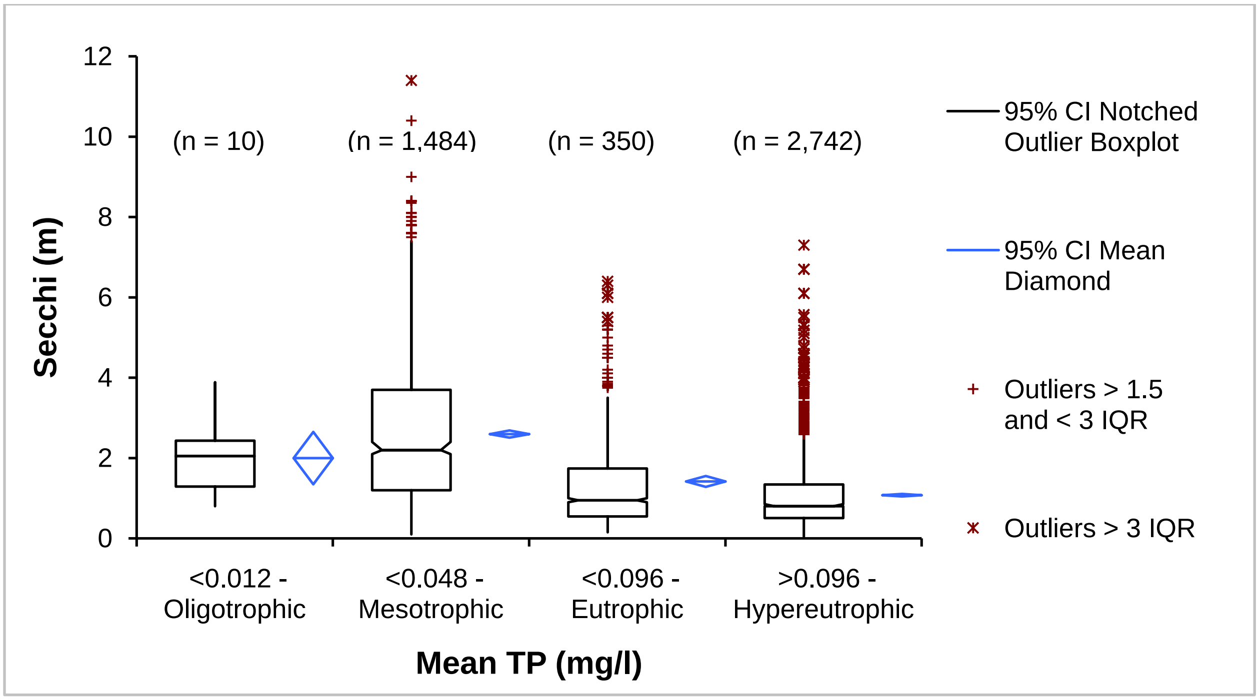


Figure 15: Distribution of Secchi Depth by Mean Reservoir TP – All Data



Results of this analysis generally show the expected pattern of increasing Chl-a concentrations and decreasing secchi depths with increasing mean TP. However, in both cases the expected pattern of low Chl-a and high Secchi depth is not observed under Oligotrophic conditions. This discrepancy is likely due to a lack of data in the Oligotrophic group creating a misleading distribution (there are 13 Chl-a measurements under Oligotrophic conditions in **Figure 14**, while there are 740 for Mesotrophic, 260 for Eutrophic, and 2,467 for Hypereutrophic). Kruskal-Wallis tests confirm that the distributions amongst groups are statistically significantly different. Similar plots and analyses were performed after separating the data by classification tier. Similar results were seen for these smaller groups.

Conclusions

Three of the four states (ND, SD, and WY) provided water quality data to HEI for use in Tasks 2 and 3 of the Region 8 Plains State Nutrient Criteria project. The analyses described within this memorandum focused on those water quality parameters that describe the presence of nutrients in a system and the response to these nutrients. Due to a lack of lake morphometry data in the study area, the water quality analyses were performed only on data from reservoirs that are present in the reservoir master database (HEI, 2010). One hundred and seventy eight reservoirs were included in the analysis, the bulk of which are in ND and SD.

Basic statistics show that the distributions of TP, Chl-a, and secchi depth are all right-skewed, leading them to be lognormally-transformed for the linear regression analysis. Results of the stressor-response variable linear regressions show the expected results with Chl-a increasing and secchi depth declining with an increase in TP and TN. As Chl-a values increase secchi depths decrease. Considerable variability is seen in the relationship of TN to TP, indicating the potential for phosphorus limitation in some reservoirs and nitrogen or co-limitation in others. Outcomes of this analysis will be used in future receiving water quality modeling.

A Kruskal-Wallis test was performed for each parameter, grouping the data by classification tier, by ecoregion, and by state. Results of the tests show that the distributions of TP, TN, Chl-a, and secchi depth are not all statistically significantly equal amongst the tiers. Pairwise analyses, show that some tiers are statistically significantly different, while others are not. A more detailed statistical test (for example, two-way analysis of variance) is recommended to gain further insight to the water quality data and determine if compounding factors may assist in separating water quality responses.

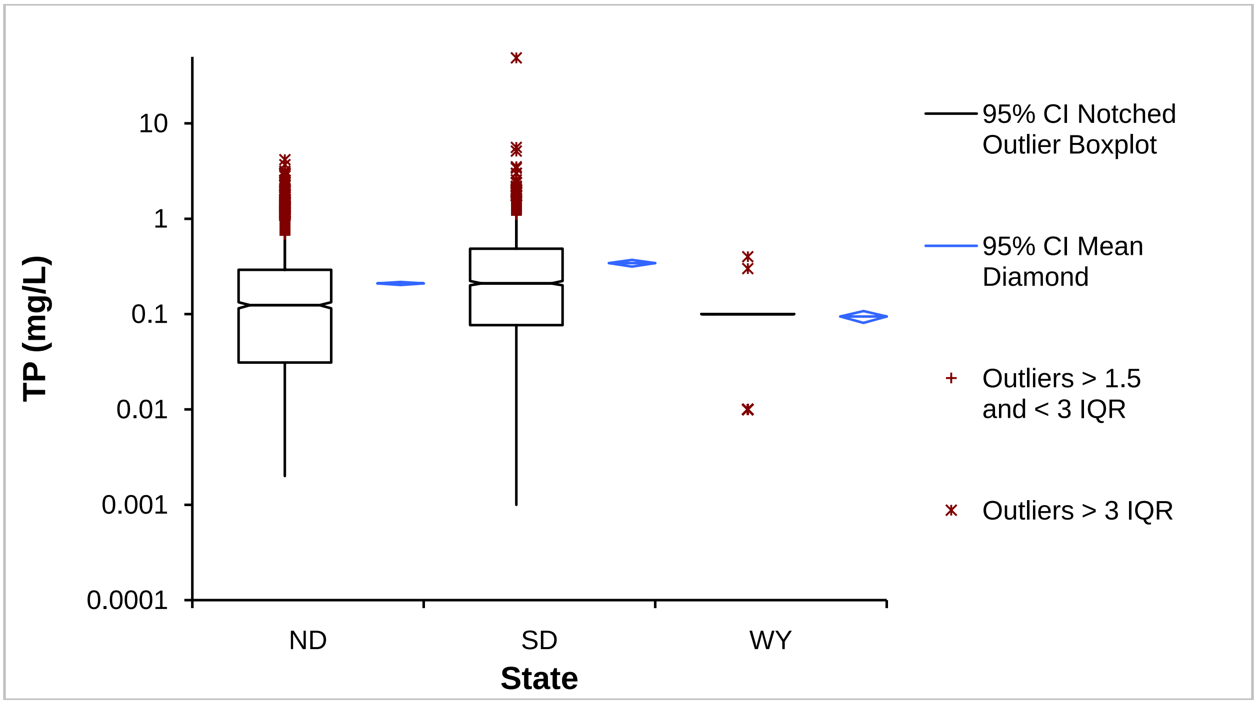
An analysis of Chl-a and secchi depth distributions grouped by mean reservoir TP generally shows the anticipated trend of increasing Chl-a and decreasing secchi depth as mean TP concentrations rise. Kruskal-Wallis tests confirm that the distributions amongst mean TP categories are statistically significantly different. Repeating the analysis while grouping data by classification tier produces a similar result.

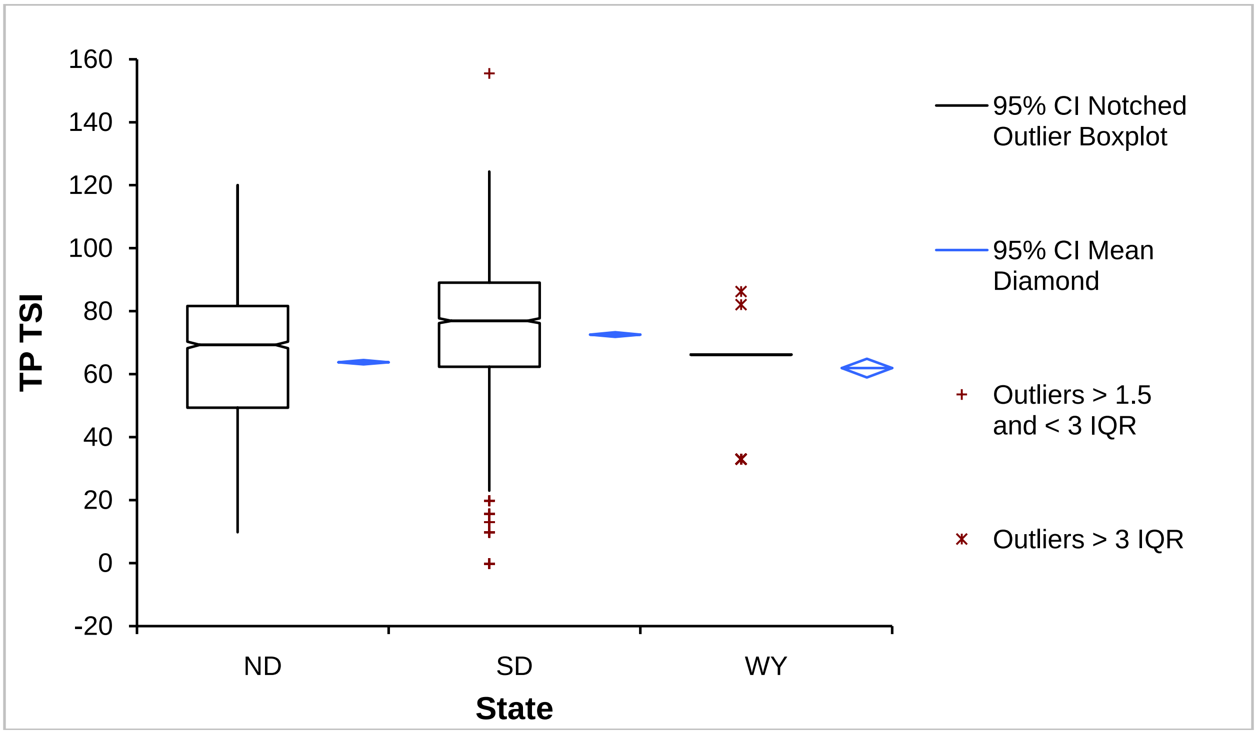
Based on results of the water quality analyses, HEI recommends moving forward into the modeling component of this project. Watershed loading and receiving water models should be setup for a select group of “priority” reservoirs, selected from the list shown in **Table 2**. Priority reservoirs will be chosen based on the amount of water quality data available and the accuracy of the reservoir drainage area and morphometric data. Results of the pairwise Kruskal-Wallis tests do not warrant separate models by tier. Additional statistical analysis may give more insight, but based on the results in this memo, modeling by tier is not recommended. Once the receiving water models are created, results of the “Analysis of Response Variables by Mean Phosphorus Concentrations” section will be used to assist in setting appropriate nutrient criteria.

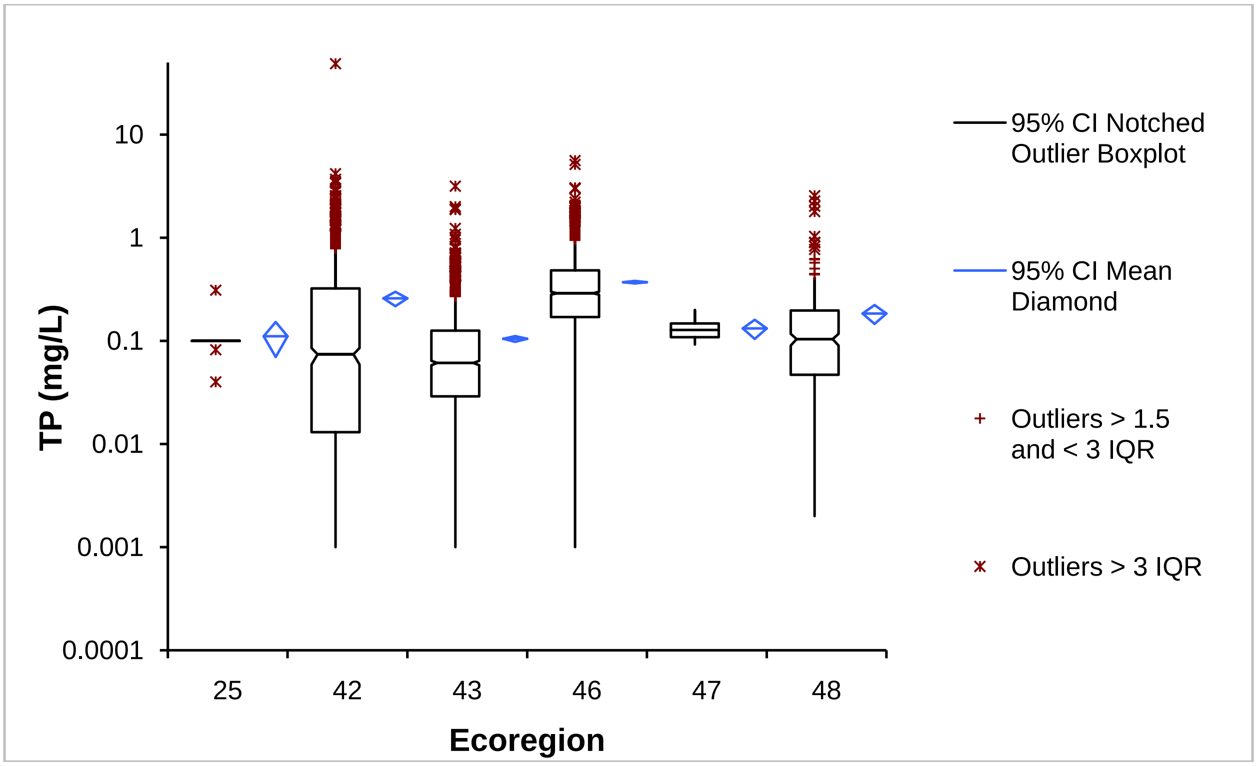
References

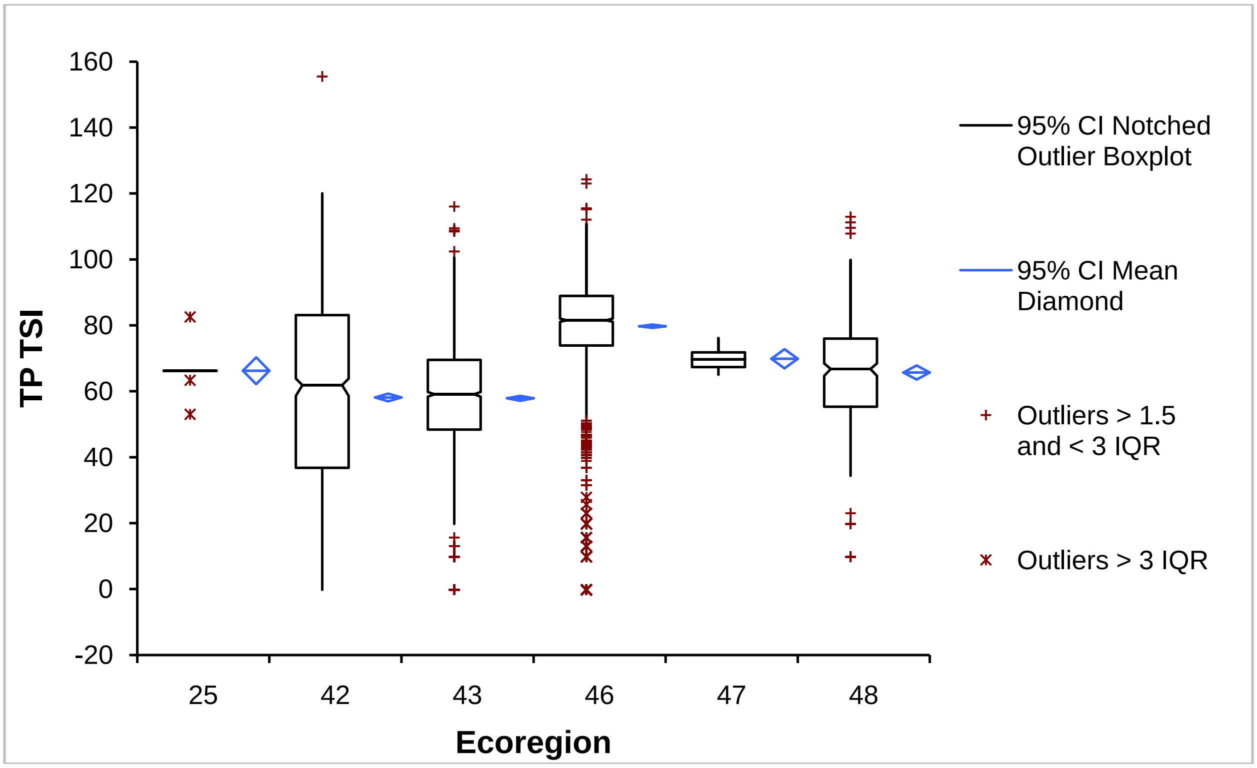
Houston Engineering, Inc. March 5, 2010. “Status report on select activities under Tasks 2 & 3 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8”. Memo to Tina Laidlaw at EPA.

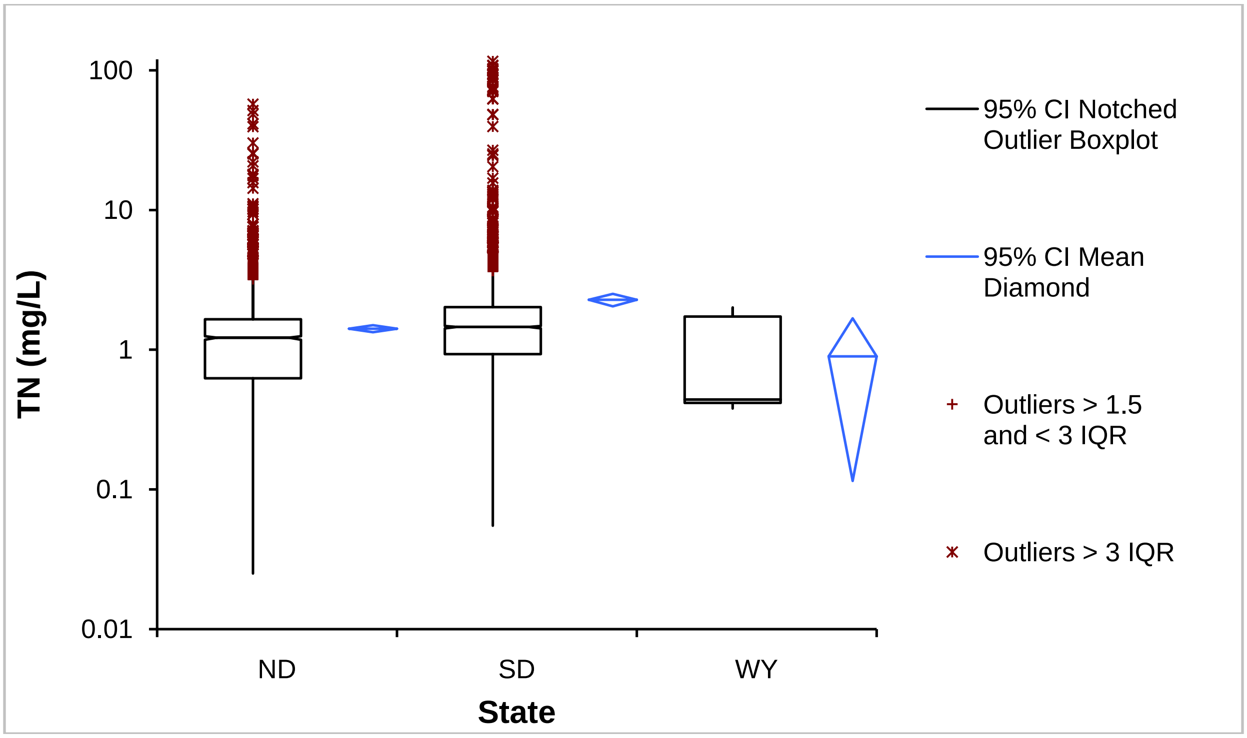
APPENDIX

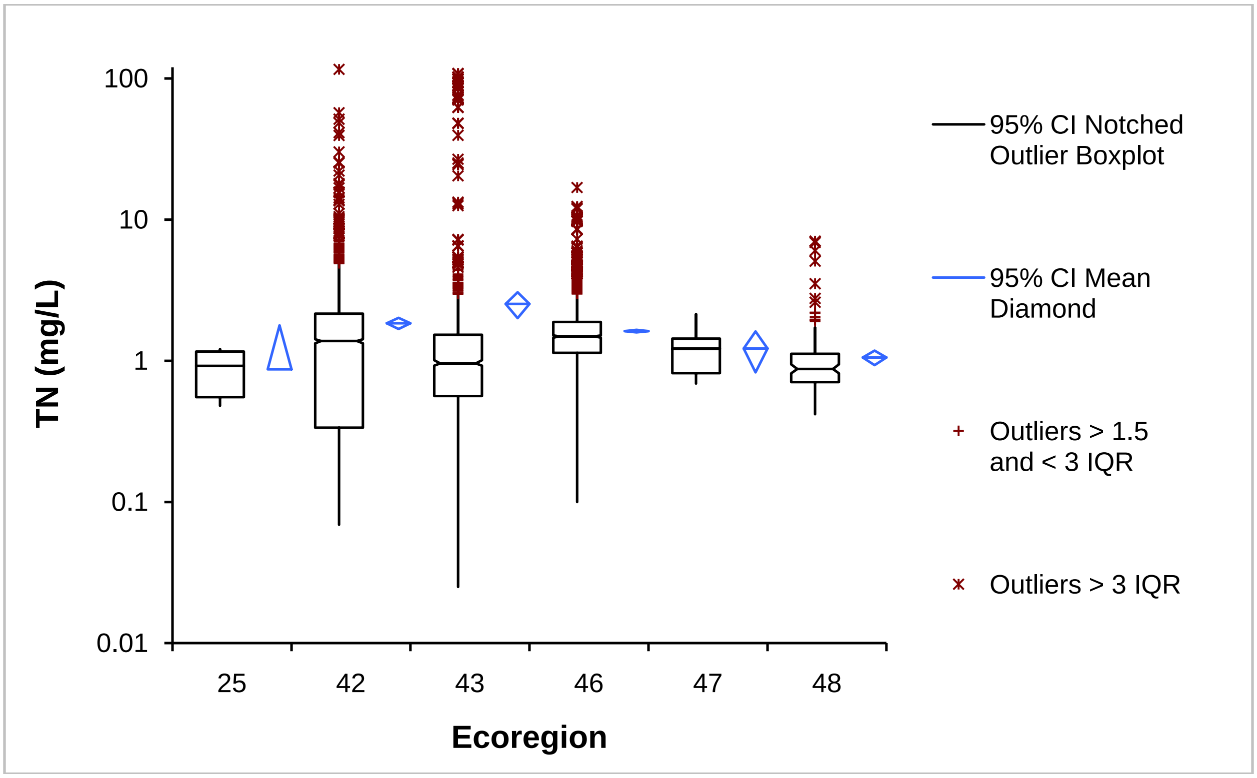

TP (mg/L)
10
1
0.1
0.01
0.001
0.0001
ND
SD
WY
State
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR

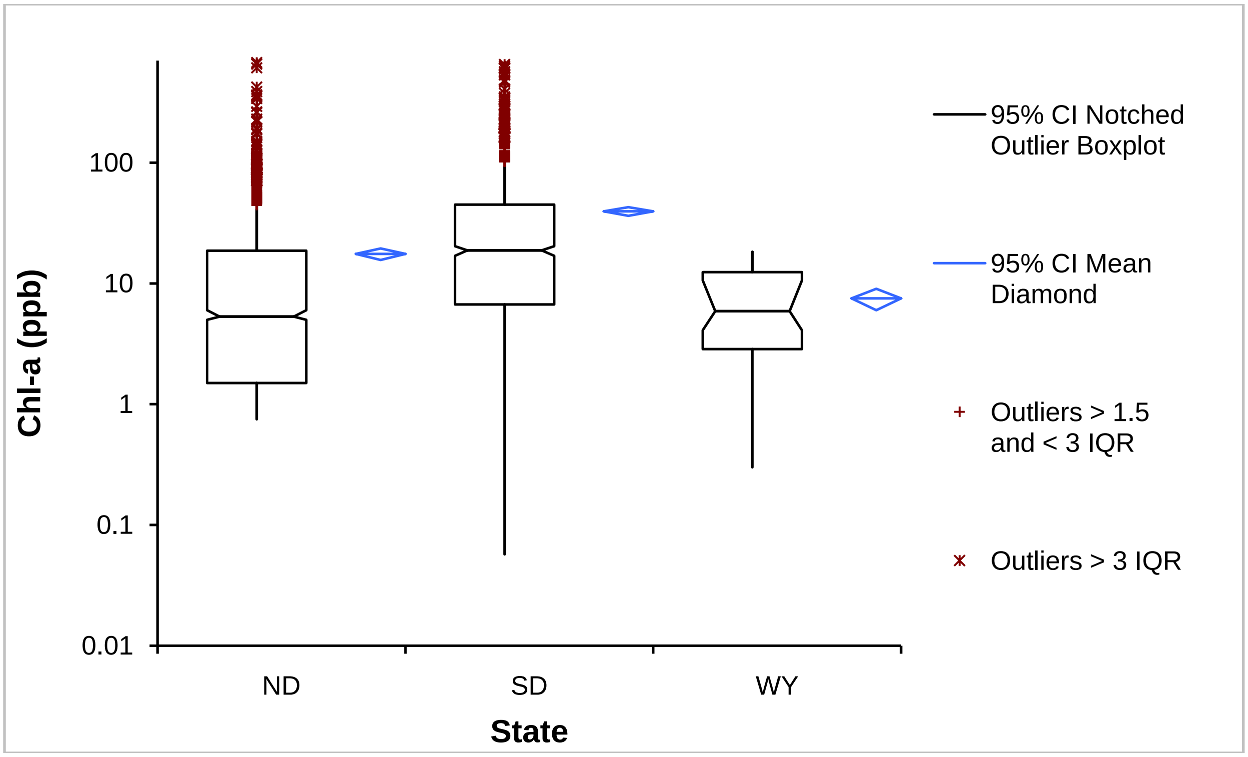

TP TSI
160
140
120
100
80
60
40
20
0
-20
ND
SD
WY
State
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR

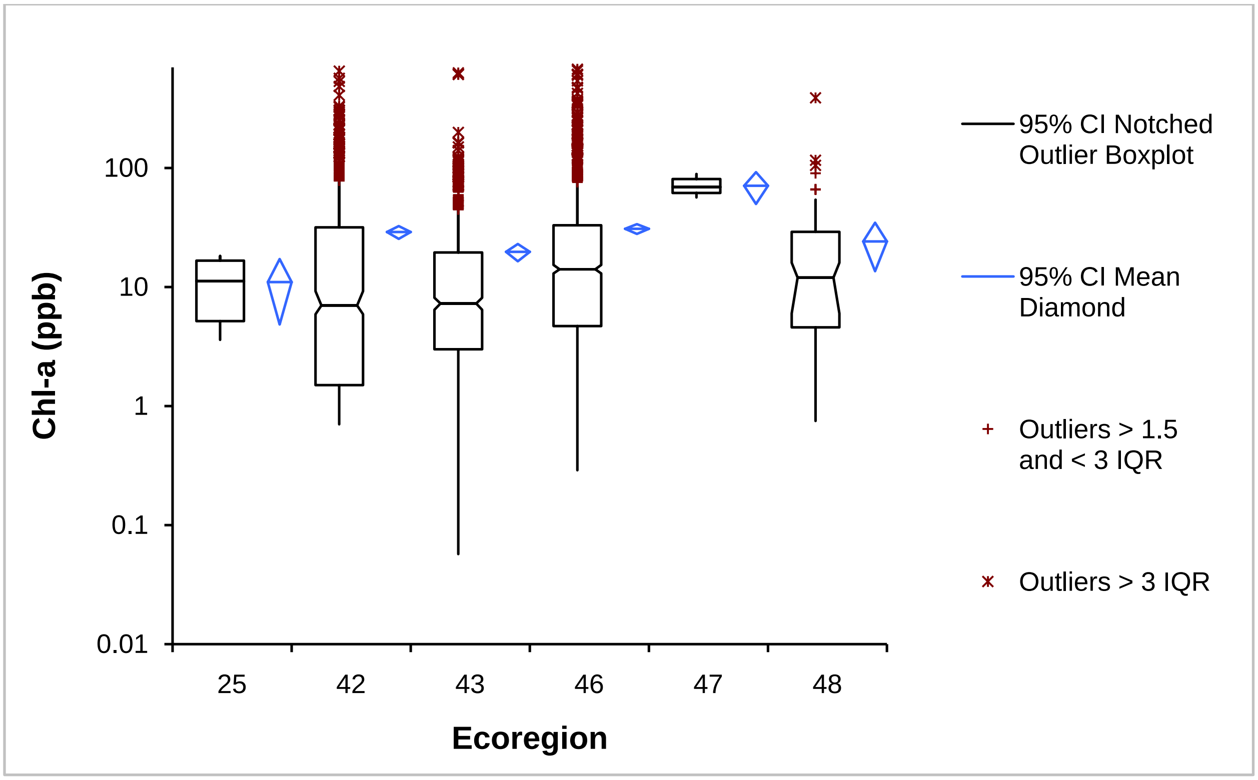

TP (mg/L)
10
1
0.1
0.01
0.001
0.0001
25
42
43
46
47
48
Ecoregion
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR

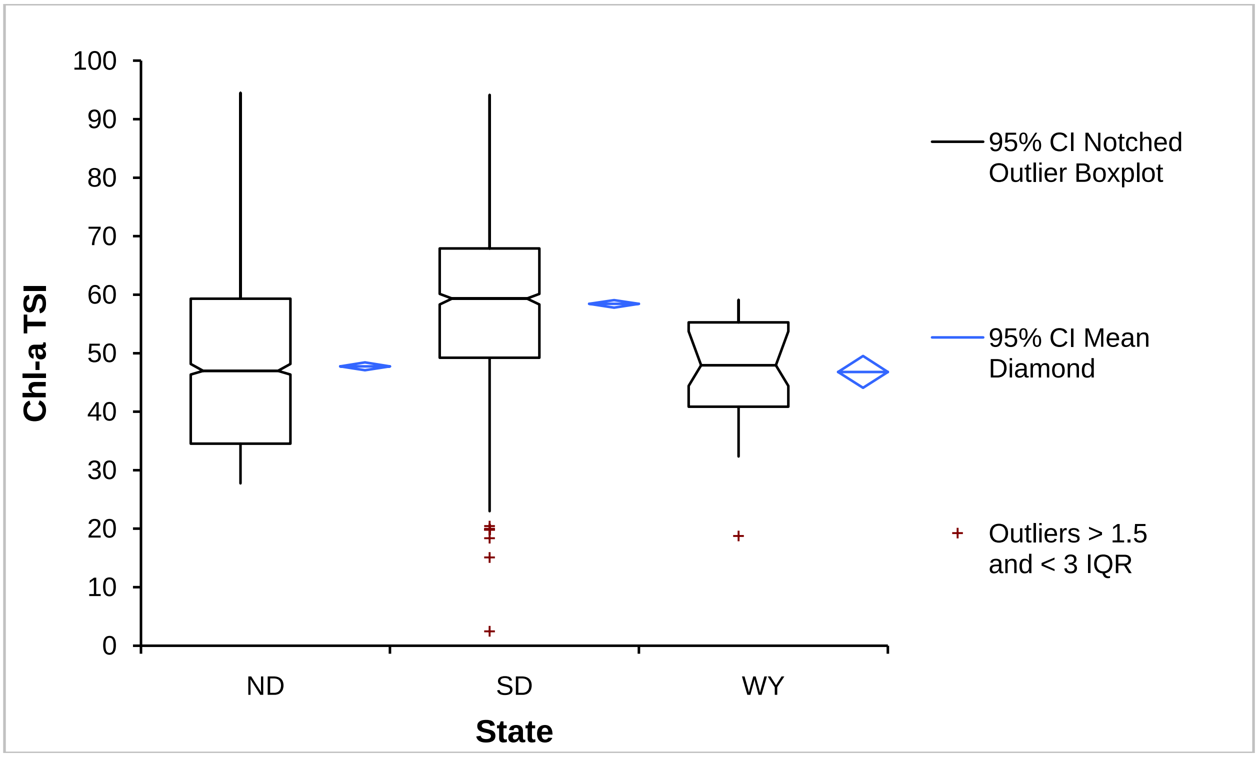

TP TSI
160
140
120
100
80
60
40
20
0
-20
25
42
43
46
47
48
Ecoregion
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR

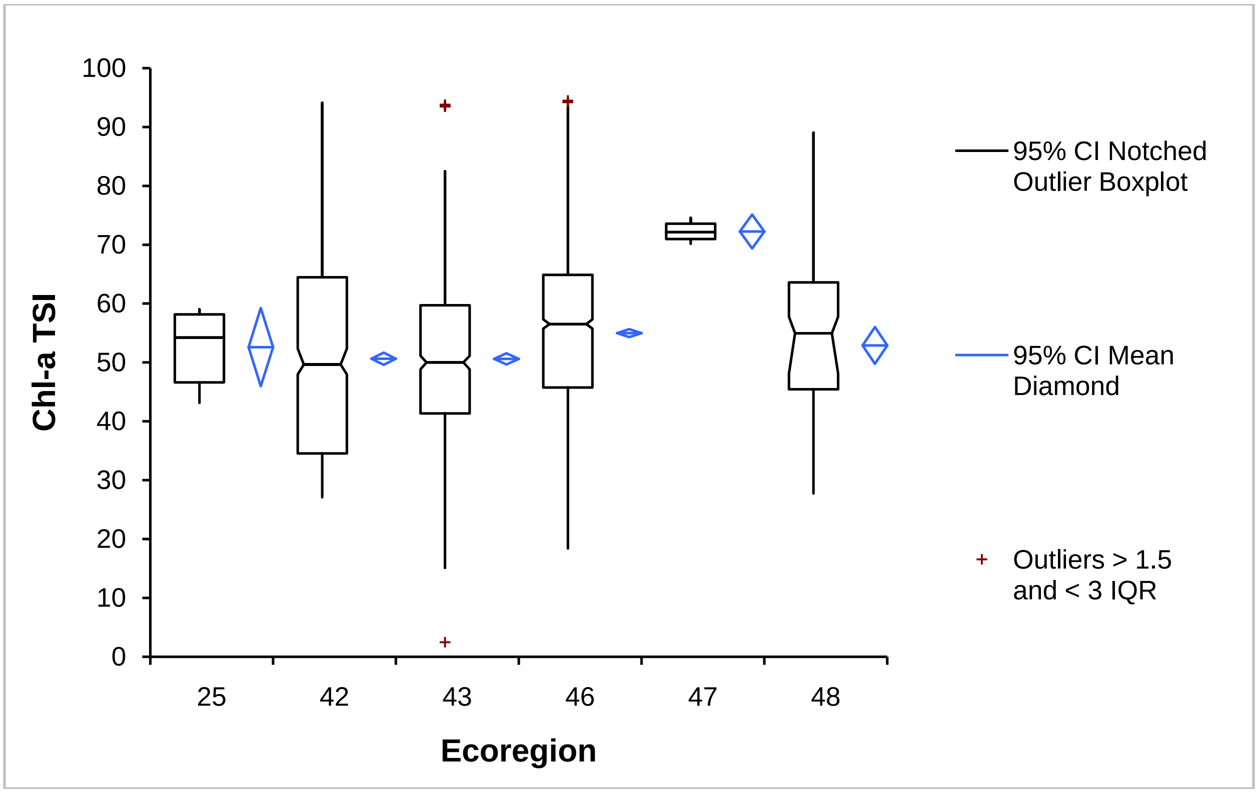

100
10
1
0.1
0.01
TN (mg/L)
ND
SD
WY
State
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR

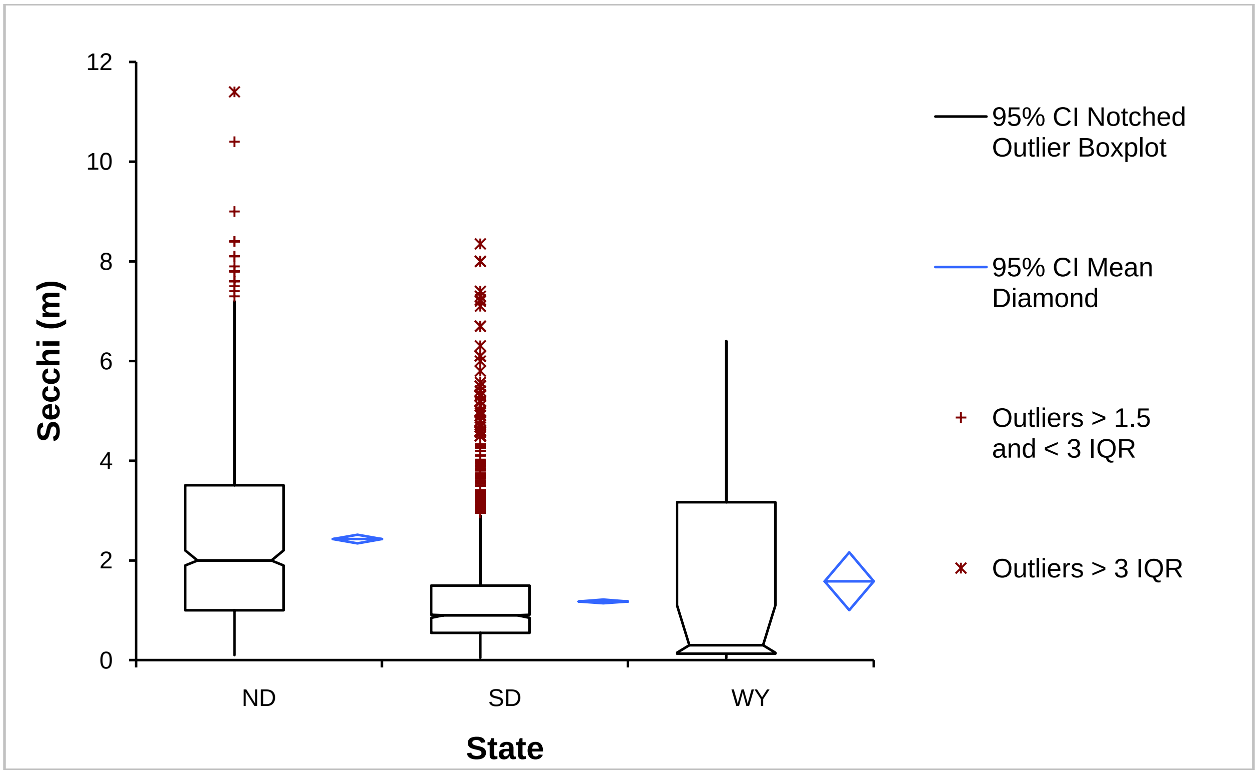

100
10
1
0.1
0.01
TN (mg/L)
25
42
43
46
47
48
Ecoregion
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR

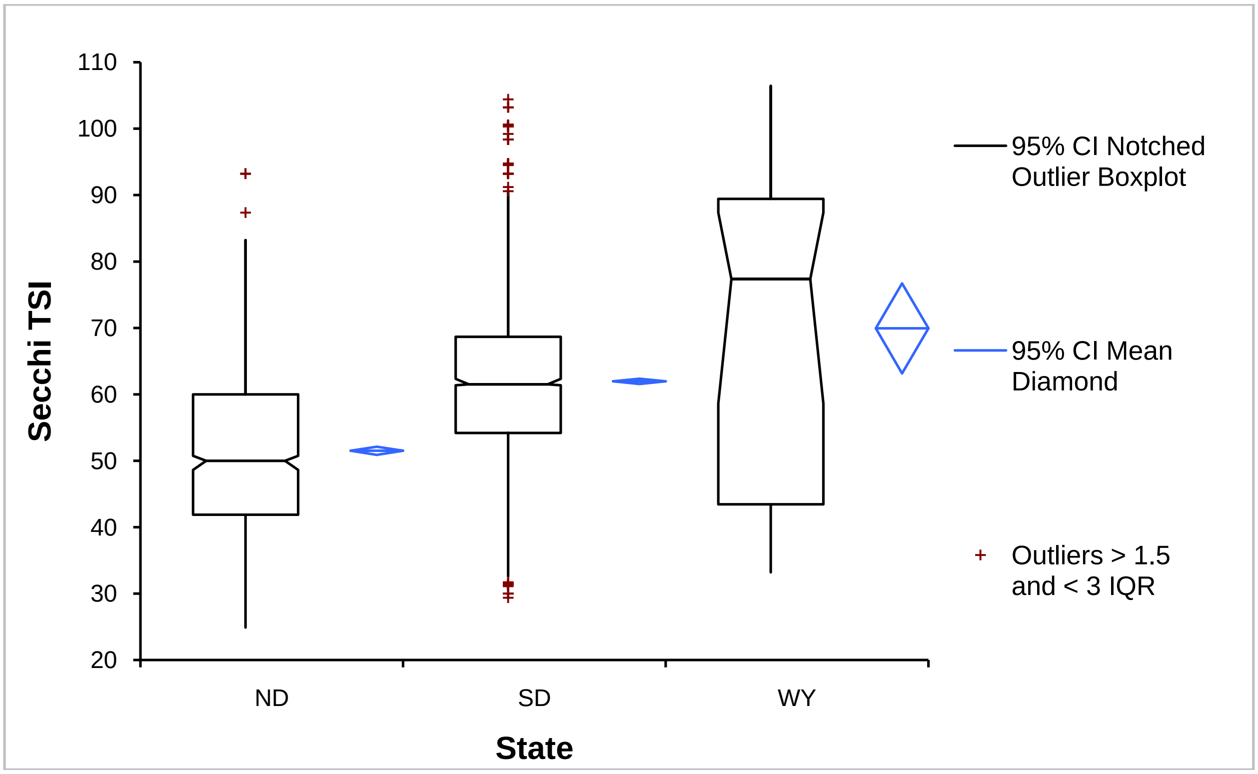

Chl-a (ppb)
100
10
1
0.1
0.01
ND
SD
WY
State
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR

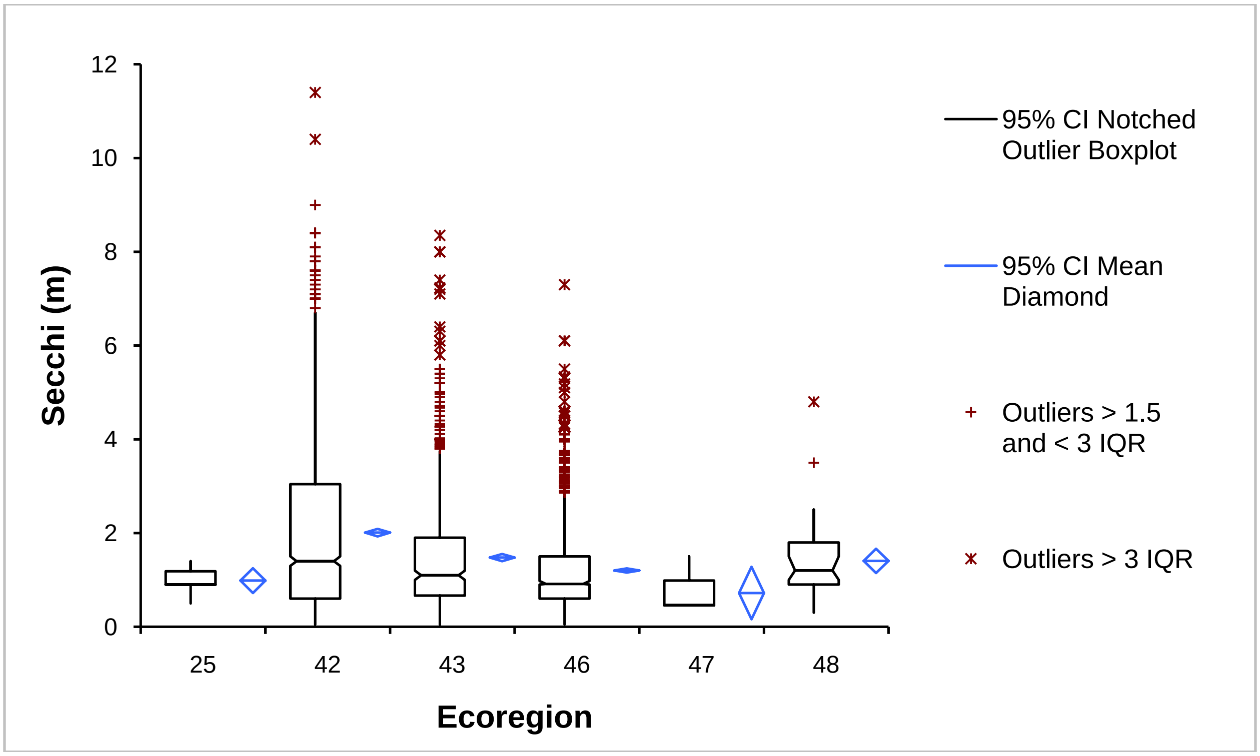

Chl-a (ppb)
100
10
1
0.1
0.01
25
42
43
46
47
48
Ecoregion
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR


Chl-a TSI
100
90
80
70
60
50
40
30
20
10
0
ND
SD
WY
State
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR


Chl-a TSI
100
90
80
70
60
50
40
30
20
10
0
25
42
43
46
47
48
Ecoregion
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR


Secchi (m)
12
10
8
6
4
2
0
ND
SD
WY
State
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR


Secchi TSI
110
100
90
80
70
60
50
40
30
20
ND
SD
WY
State
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR


12
10
8
6
4
2
0
Secchi (m)
25
42
43
46
47
48
Ecoregion
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR
Outliers > 3 IQR

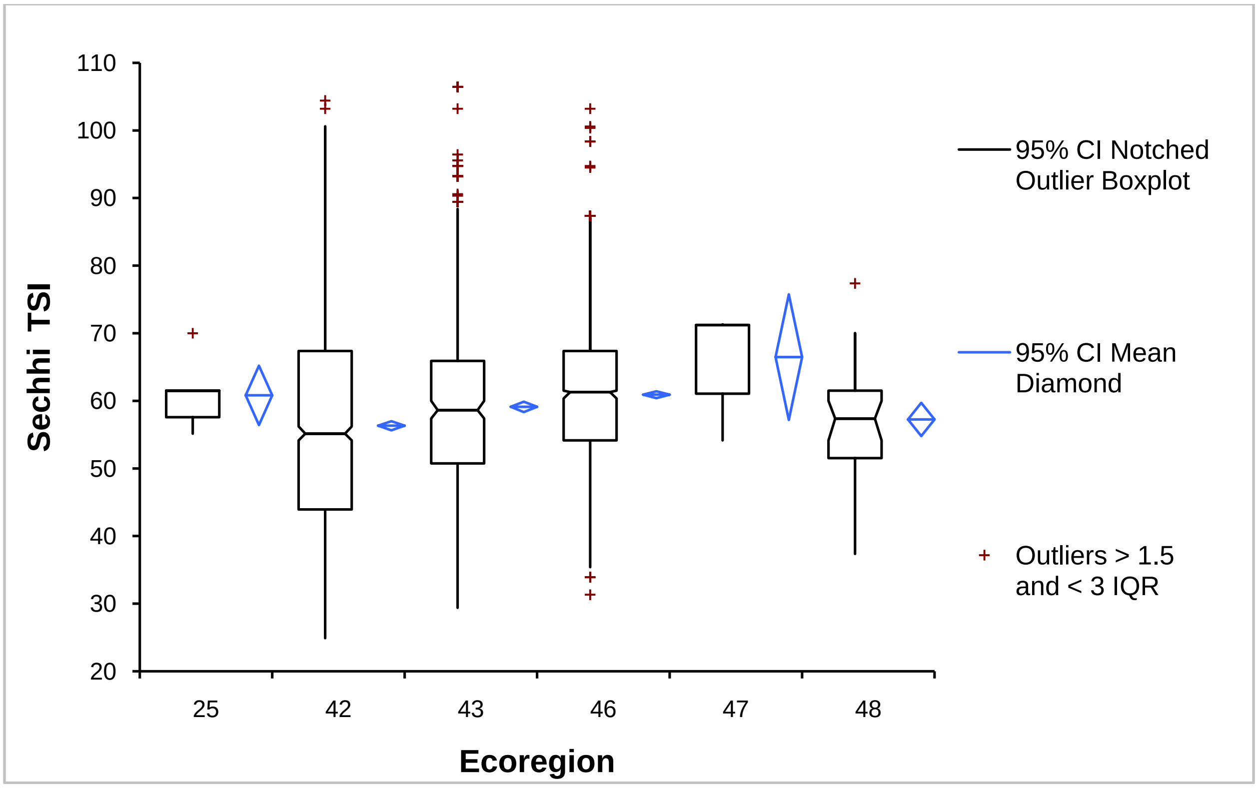

110
100
90
80
70
60
50
40
30
20
Sechhi TSI
25
42
43
46
47
48
Ecoregion
95% CI Notched Outlier Boxplot
95% CI Mean Diamond
Outliers > 1.5 and < 3 IQR

Table 2: Overall Summary of Water Quality Data

				Classification	Sample Date		Secchi Depth (m)				TN (mg/L)				TP (mg/L)				Chl-a (ppb)			
ReservoirName	State	Ecoregion	ComID	Tier	Min	Max	n	Avg	Minimum	Maximum	n	Avg	Minimum	Maximum	n	Avg	Minimum	Maximum	n	Avg	Minimum	Maximum
Academy Lake	SD	42	148207364	1	6/15/1989	8/3/2005	10	0.5	0.05	1.3	19	1.92	1.06	3.2	19	0.822	0.215	1.86	3	79.550	33.49501	102.644
Allen	SD	25	126571649	1	7/11/1996	7/11/1996	1	0.50	0.50	0.50					1	0.310	0.31	0.31				
Alvin	SD	46	130997497	1	6/29/1989	7/24/2007	44	0.81	0.30	1.4	76	2.61	0.1	16.9	76	0.270	0.001	1.55	14	67.196	7.173913	166.16
Amsden Dam	SD	46	145283218	2	7/12/1989	8/12/2008	59	2.3	0.56	7.3	78	1.10	0.16	2.67	66	0.373	0.001	0.901	36	12.731	1.056	146.06
Angostura	SD	43	137355575	3	9/3/1977	7/23/2008	134	2.7	0.25	8.4	248	0.551	0.055	1.58	250	0.026	0.001	0.424	126	9.301	0.20625	162.9375
Armourdale Dam	ND	46	0	1	1/21/1988	9/11/2004	2	0.70	0.60	0.80	53	2.00	1.47	2.56	66	0.339	0.061	1.94	11	25	0.8	66.8
Arnegard Dam	ND	43	143765670	1	7/22/1980	7/16/2009	5	1.2	0.20	3.0	1	0.484	0.484	0.484	9	0.153	0.031	0.25	3	12	3	24.6
Arroda Lake	ND	43	0	1	6/12/1986	1/25/2006	4	1.6	0.70	2.2	4	2.17	1.18	4.84	4	0.037	0.024	0.065	3	3.8	1.5	7.1
Balta Dam	ND	46	143261387	1	7/21/1992	1/24/2006	5	0.38	0.30	0.40	4	2.68	2.3	2.9	11	0.285	0.144	0.507	6	36	14	66.8
Baukol-Noonan Dam	ND	46	144443097	1	7/14/1992	2/23/1993	2	2.0	1.50	2.4					6	0.017	0.002	0.031	2	1.5	1.5	1.5
Bear Butte	SD	43	154903300	2	6/4/1996	5/17/1999	1	1.5	1.50	1.5					1	0.080	0.08	0.08	2	19.26	12.06	26.465
Beaver (STATE) Lake	SD	46	125116353	1	1/29/1990	8/3/2009	57	0.44	0.05	1.7	61	2.24	0.22	5.44	61	0.447	0.001	1.32	23	52	6.6	116.1333
Bisbee-Big Coulee Dam	ND	46	143387941	3	7/30/1991	2/23/1992	2	0.55	0.50	0.60					12	0.821	0.687	0.956	2	33	15	50
Blacktail Dam	ND	42	148778537	2	7/16/1991	10/31/2004	1	1.6	1.60	1.6	75	1.14	0.575	3.74	84	0.129	0.002	1.19	27	14	0.75	66.4
Blickensderfer Dam	ND	43	0	1	2/27/2003	6/17/2003					7	1.69	1.53	1.97	7	0.024	0.015	0.037	1	3	3	3
Blumhardt Lake	ND	42	147905337	1	6/16/2005	11/4/2008	10	0.75	0.30	1.1	13	1.50	1.14	1.96	13	0.077	0.033	0.117	13	27	0.75	81.2
Bowman-Haley Dam	ND	43	131873151	3	8/14/1986	1/31/2001	20	0.95	0.25	2.4	20	1.54	1.09	2.36	83	0.098	0.002	0.432	15	19	2	66
Braddock Dam	ND	42	0	1	7/28/1992	2/25/1993	2	0.85	0.50	1.2					6	0.285	0.163	0.368	2	18	14	22
Brakke Dam	SD	43	144351649	2	6/15/1989	8/11/2009	40	0.74	0.15	1.7	63	0.936	0.21	2.49	63	0.071	0.001	0.217	23	22.554	0.94875	161.733
Brewer Lake	ND	46	143318142	2	9/4/1987	10/19/2005	5	1.1	0.50	1.8	99	1.02	0.653	1.96	117	0.170	0.013	1.24	27	16	0.75	71
Burke Lake	SD	42	148214698	1	2/16/1989	7/22/2008	117	0.50	0.15	2.0	173	2.74	0.11	8.11	173	0.324	0.001	1.56	41	97.159	7.46625	407.715
Bylin Dam	ND	46	149372937	1	7/17/1996	3/11/1997	2	0.80	0.80	0.80					9	0.599	0.496	0.832	2	13	12	13
Byre	SD	43	144351237	1	5/1/2000	8/2/2005	22	0.58	0.12	1.5	43	1.44	0.23	5.46	43	0.104	0.001	0.332	15	18.26	1.2375	81.18
Camels Hump Dam	ND	43	75080387	1	7/22/2005	2/23/2006	2	2.8	2.00	3.6	6	1.02	0.913	1.33	6	0.017	0.01	0.031	1	1.5	1.5	1.5
Campbell (CAMPBELL)	SD	42	145469326	1	7/18/1989	9/17/2008	65	1.5	0.46	4.7	88	2.16	0.12	5.6	24	0.806	0.003	1.34	52	36.655	2.145	303.1462
Carbury Dam	ND	46	0	2	7/21/1992	10/6/2003	2	0.40	0.40	0.40	47	2.26	1.68	3.07	59	0.282	0.086	0.649	15	52	21	106
Carthage	SD	46	145674916	1	6/19/1989	8/29/2006	26	1.3	0.30	5.5	47	1.92	0.16	6.5	47	0.295	0.001	1	33	70.750	9.669001	351.2025
Castle Rock Dam	ND	43	0	1	7/24/2002	3/6/2006	2	1.4	0.90	1.9	9	1.89	1.52	3.01	9	0.147	0.059	0.416	3	5.000	1.5	12
Cedar Lake	ND	43	139875485	1	7/15/1991	3/6/2006	2	0.45	0.40	0.50	5	1.90	1.55	2.51	16	0.133	0.009	0.586	3	15	1.5	23
Clausen Springs	ND	46	143319625	1	7/23/1991	1/24/2006	6	2.2	1.2	4.0	4	1.53	1.26	1.85	13	0.341	0.065	0.991	6	1.275	0.75	2.4
Coal Springs	SD	43	151677934	1	7/25/1989	7/26/2004	12	2.2	0.15	3.1	12	2.18	1.62	3.31	12	0.318	0.085	0.58	6	8.639	2.68	24.12
Corsica	SD	42	128447909	1	6/14/1989	7/24/2006	31	0.36	0.21	0.70	46	2.08	0.1	4.97	46	0.618	0.001	3.51	30	66.01	10.05	281.094
Covell	SD	47	130997150	1	6/29/1989	8/17/2004	5	0.72	0.46	1.5	8	1.22	0.69	2.14	8	0.132	0.092	0.199	4	70.92	56.5125	88.77
Cresbard	SD	46	144267820	1	7/24/1989	7/17/2007	52	0.96	0.18	3.2	92	1.81	0.12	5.41	91	0.769	0.001	2.07	29	30.696	1.794375	115.2113
Crown Butte Dam	ND	43	135580348	1	7/20/1992	9/20/2005	4	2.2	0.70	4.4	64	1.72	0.778	3.81	80	0.210	0.009	3.16	15	32	1.5	150
Curlew	SD	43	126844644	1	12/19/1977	7/28/2005	11	0.66	0.20	1.2	13	0.955	0.3	1.65	14	0.056	0.001	0.12	7	19.87	4.785	52.635
Dante	SD	42	128449065	1	6/14/1989	7/23/2007	33	1.0	0.30	3.6	37	2.28	1.02	5.83	37	0.181	0.03	0.696	21	35.482	0.70125	82.69801
Danzig Dam	ND	43	135579870	1	7/21/1994	1/31/1995	2	1.5	1.1	1.8					6	0.183	0.084	0.29	2	32	24	39
Davis Dam	ND	43	0	1	7/7/1987	2/8/1995	5	2.4	1.0	4.5					9	0.057	0.009	0.112	2	41	15	66
Dead Colt Creek Dam	ND	48	0	1	7/29/1992	7/9/2009	2	1.7	1.2	2.1	41	1.18	0.687	3.52	50	0.167	0.002	1.03	12	13	0.75	40
Derby	SD	43	126553039	1	10/25/1977	8/10/1978					3	0.477	0.29	0.75	2	0.263	0.219	0.307				
Dewbarry	SD	43	151678487	1	7/31/1989	7/31/2003	6	1.1	0.10	1.8	6	3.88	1.21	6.54	6	1.34	0.21	2	3	105.01	94.6275	112.6125
Dickinson Dike	ND	43	70796313	1	8/12/1993	2/23/2006	19	2.7	0.75	5.5	44	0.981	0.62	2.17	53	0.072	0.002	0.714	14	16	0.75	79
East Arroda Lake	ND	43	0	1	6/2/2005	1/25/2006	3	1.5	1.4	1.7	4	1.33	1.26	1.45	4	0.048	0.04	0.053	3	3.9	3.6	4.3
Elm Lake	SD	46	147911290	3	7/19/1989	8/11/2008	81	1.4	0.20	5.2	139	1.21	0.1	2.28	139	0.356	0.004	0.836	57	13.12	0.335	69.345
Epping-Springbrook Dam	ND	42	143752951	2	7/16/1991	1/22/1992	2	1.9	1.2	2.5					9	0.707	0.558	0.884	2	13	1.5	25
Fairfax	SD	42	148215250	1	6/27/2001	8/3/2005	14	1.2	0.37	2.4	10	1.75	0.23	2.8	10	0.716	0.001	0.963	4	34.685	7.218751	71.40376
Fate	SD	43	144351038	1	6/15/1989	7/22/2008	45	0.84	0.15	1.5	63	1.10	0.15	3.53	63	0.078	0.001	0.481	23	7.121	0.5775	35.71615
Faulkton	SD	46	144272084	1	7/24/1989	9/7/2006	62	1.2	0.06	4.0	82	1.91	0.1	3.83	81	0.544	0.0025	1.078	29	69.9	2.68	340.659
Fiddle Creek Dam	SD	43	137355454	1	11/9/1977	8/22/1978					3	0.48	0.28	0.6	3	0.019	0.016	0.022				
Fish Creek Dam	ND	43	135581385	1	7/6/1993	2/7/2006	6	1.6	0.50	3.3	4	1.38	1.14	1.7	15	0.096	0.026	0.426	6	25	0.75	88.1
Flat Creek	SD	43	143227281	2	6/21/1989	8/11/2009	29	0.71	0.20	2.2	20	1.72	1.12	2.86	20	0.137	0.063	0.261	9	36.113	13.84667	69.01125
Fordville Dam	ND	48	149374832	2	7/22/1992	7/21/2009	2	1.4	1.0	1.8	9	1.16	0.676	1.72	17	0.248	0.066	0.502	5	15	3.125	49

				Classification	Sample Date		Secchi Depth (m)				TN (mg/L)				TP (mg/L)				Chl-a (ppb)			
ReservoirName	State	Ecoregion	ComID	Tier	Min	Max	n	Avg	Minimum	Maximum	n	Avg	Minimum	Maximum	n	Avg	Minimum	Maximum	n	Avg	Minimum	Maximum
Fort Meade Blm	SD	43	154903564	**1**	8/2/1999	8/2/1999													1	14.740	14.74	14.74
Freeman	SD	43	128633827	**1**	6/14/1989	8/23/2007	52	1.0	0.25	6.30	51	46	0.15	108.6	52	0.082	0.001	0.384	24	63.484	1.34	199.0725
Frettum Dam	ND	42	145410781	**2**	6/10/2008	8/25/2008	2	1.90	1.20	2.60	2	1.4	1.2	1.65	2	0.023	0.015	0.031	2	3	3	3.7
Froelich Dam	ND	43	145465848	**3**	7/9/1992	2/13/2007	5	1.4	0.70	2.30	4	2.80	1.96	3.58	13	0.316	0.135	0.647	6	33	1.5	118
Gardner (BUFFALO LAKE)	SD	43	131721911	**2**	7/15/1996	7/24/2002	10	0.5	0.20	0.75	3	0.9	0.8	0.99	4	0.060	0.04	0.087	4	6.425	3.09375	11.55
Gascoyne Lake	ND	43	131871786	**1**	5/26/2009	7/16/2009	2	1.0	0.80	1.20									2	8	6	10.7
Geddes	SD	42	148214656	**1**	6/15/1989	7/24/2006	30	0.35	0.15	1.40	45	3	0.1	8.28	45	0.390	0.001	1.12	29	127.632	22.78	323.7772
Glendo Reservoir	WY	43	136457559	**4**	7/7/2004	7/6/2006	21	3.4	0.30	6.40					30	0.1	0.01	0.1	21	3	0.3	6.5
Guernsey Reservoir	WY	25	136458524	**3**	8/5/2004	8/5/2004	6	1.07	0.90	1.40					9	0.1	0.1	0.1	6	11	3.6	18.2
Hanson's	SD	46	125127710	**1**	7/5/1989	6/14/2006	23	0.91	0.61	3.08	42	0.94	0.27	4.26	42	0.116	0.001	0.396	11	20.123	5.73375	55.48125
Harmon Lake	ND	43	0	**1**	5/29/2009	7/14/2009					1	1.87	1.87	1.87	1	0.193	0.193	0.193	3	46	1.5	96.9
Harvey Dam	ND	46	143263044	**1**	6/26/1990	2/10/2003	3	1.80	1.50	2.30	2	1.86	1.84	1.88	11	0.810	0.298	1.07	2	9.250	1.5	17
Hayes	SD	43	128617117	**1**	6/13/1989	7/28/2004	35	1.1	0.30	1.90	60	1.8	0.28	5.19	61	0.196	0.001	0.533	26	28.853	5.6925	119.8312
Heinrich-Martin Dam	ND	46	0	**1**	7/27/1992	11/4/2008	12	1.6	0.40	3.00	13	1.298	1	1.81	22	0.087	0.002	0.219	15	33	0.75	127
Henry (SCOTLAND)	SD	46	125122308	**1**	6/27/1989	8/6/2008	10	1.1	0.30	2.69	20	1.8	0.67	3.23	20	0.280	0.143	0.501	6	15.876	6.1875	51.579
Hiddenwood	SD	42	147897989	**1**	7/18/1989	8/5/2009	85	0.73	0.20	3.81	83	1.5	0.1	5.105	83	0.207	0.002	0.542	7	29.441	7.96125	80.4
Homme Dam	ND	48	0	**2**	6/25/1991	9/9/2006	3	1.2	0.50	1.80	12	2.80	0.844	7.05	18	0.591	0.054	2.55	11	32	2	105
Hurley	SD	42	139474512	**1**	7/7/2004	7/27/2004	8	1.7	0.76	5.57	4	1.75	1.43	2.07	4	0.872	0.771	0.972	4	45.251	28.05	65.54625
Indian Creek Dam	ND	43	0	**1**	7/15/1991	3/6/2006	20	1.6	0.50	5.00	67	1.67	1.26	2.92	76	0.065	0.008	0.274	18	17	1.5	76.4
Isabel	SD	43	143229800	**2**	6/20/1989	8/1/2005	23	1.1	0.60	2.69	29	1.6	0.1	2.94	29	0.218	0.001	0.451	12	84.467	2.68	612.38
Jamestown Reservoir	ND	46	147293683	**3**	5/13/1998	8/13/2009	25	1.0	0.20	2.70	79	1.61	0.722	3.26	84	0.287	0.079	0.654	19	43	5.9	109
Jewett	SD	43	151679493	**1**	6/3/2008	8/12/2008	8	0.88	0.75	1.00	2	1.56	1.28	1.84	2	0.325	0.243	0.406	2	18.652	7.226999	30.07617
Jones (HAND)	SD	46	142197488	**1**	6/12/1989	8/12/2009	56	0.93	0.15	2.90	65	1.7	0.16	3.67	65	0.508	0.001	0.871	41	38.649	4.0425	144.4575
Keyhole Reservoir	WY	43	140523605	**4**	8/6/2002	6/25/2003	24	0.12	0.04	0.19	6	0.9	0.38	2	30	0.1	0.1	0.4	23	11	3	18.3
Kolding Dam	ND	46	0	**1**	7/17/1996	3/12/1997	2	0.90	0.80	1.00					8	0.405	0.051	0.871	2	23	16	29
Kota Ray Dam	ND	43	143754084	**1**	7/13/1992	7/15/2009	4	2.20	1.80	2.80					8	0.064	0.021	0.128	3	5	5	5.7
Kroetche	SD	43	156021192	**2**	7/5/1995	6/26/2000													2	8.341	6.7	9.9825
Kulm-Edgeley Dam	ND	46	147899624	**1**	7/24/1991	1/25/2006	6	1.79	1.10	3.75	4	1.7	1.5	2.14	13	0.962	0.522	3.05	6	14	0.75	36
Kyle	SD	25	126557480	**1**	10/25/1977	8/10/1978					3	0.870	0.48	1.21	2	0.061	0.04	0.082				
Lacreek Refuge Pool #10	SD	43	154733387	**1**	6/3/2008	7/22/2008	8	0.18	0.10	0.30	2	2.190	1.46	2.92	2	0.667	0.542	0.791	2	36.069	18.81	53.328
Lake Ashtabula	ND	46	147442779	**3**	2/4/1987	10/16/2008	3	1.32	1.00	1.75	623	1.485	0.791	10.7	826	0.285	0.062	0.773	605	25	0.75	676
Lake Audubon	ND	43	143777120	**4**	5/31/2005	1/25/2006	3	1.97	1.30	2.40	4	0.538	0.405	0.882	4	0.009	0.006	0.016	3	1.000	0.75	1.5
Lake Darling	ND	46	144451264	**3**	5/20/1997	2/24/1998	9	0.95	0.15	1.80					24	0.157	0.059	0.46	6	33	1.5	128
Lake ILO	ND	43	71591767	**3**	5/21/2009	7/15/2009	2	0.35	0.30	0.40	1	0.764	0.764	0.764	1	0.088	0.088	0.088	1	0.750	0.75	0.75
Lake LaMoure	ND	46	148619903	**2**	7/23/1991	7/27/2009	5	1.3	0.80	2.00	95	1.957	1.16	5.9	155	0.480	0.032	1.77	45	19.489	0.75	80
Lake Oahe	ND	43	139478686	**4**	6/7/1999	9/6/2001	19	1.2	0.60	2.90	84	0.27	0.025	0.593	84	0.027	0.009	0.186	34	4.632	1.5	20
Lake Sakakawea	ND	42	143776425	**4**	6/8/1992	8/4/2009	1076	2.8	0.10	11.40	662	0.30	0.075	0.692	1,051	0.024	0.002	0.896	458	2.942	0.75	102
Lake Tschida	ND	43	70799645	**4**	8/7/1991	10/19/2008	16	1.6	0.20	3.80	88	0.835	0.456	1.94	97	0.057	0.024	0.155	17	11.059	1.5	52
Larimore Dam	ND	48	0	**1**	7/22/1992	10/15/2007	25	1.5	0.30	4.80	96	0.812	0.418	1.91	105	0.094	0.011	0.41	24	42.538	0.75	388
Latham	SD	42	144272629	**1**	6/4/2008	7/29/2008	5	0.66	0.29	0.79	2	3.22	2.18	4.25	2	1.355	1.09	1.62	2	27.935	9.768	46.101
Leland Dam	ND	43	81805697	**1**	7/12/1994	8/1/2007	4	2.08	1.10	3.20	10	1.17	0.806	1.72	19	0.054	0.002	0.162	7	10.379	0.75	22.8
Leola	SD	46	148621069	**1**	5/28/2009	8/5/2009	8	0.40	0.15	0.60	2	2.41	1.85	2.97	2	0.252	0.212	0.292				
Long Lake (Moffit)	ND	42	145419462	**4**	4/30/2003	6/4/2009					137	5.935	0.069	57.3	137	0.616	0.009	4.17				
Louise	SD	46	142196592	**2**	6/12/1989	8/12/2009	50	1.4	0.60	3.40	91	1.469	0.16	2.54	91	0.617	0.001	5.13	24	35.961	4.02	175.54
Loyalton Dam	SD	42	144265448	**1**	7/19/1989	6/6/2007	49	0.52	0.12	1.92	77	2.066	0.23	6.64	78	0.170	0.001	0.896	23	23.750	4.02	53.6
Marindahl Lake	SD	46	123212595	**2**	6/26/1989	8/4/2005	55	1.8	0.70	4.25	47	1.644	0.23	4.675	47	0.137	0.001	0.704	14	24.104	0.34	82.83
Matejcek Dam	ND	46	149374324	**1**	7/30/1991	2/26/1992	2	1.30	1.10	1.50					9	0.5	0.255	1	2	1.500	1.5	1.5
McCloud Reservoir	ND	42	143750907	**1**	5/20/2009	7/15/2009	2	2.05	1.30	2.80									1	1.500	1.5	1.5
McDowel Dam	ND	42	145416586	**1**	8/12/1993	5/14/2009	4	0.95	0.80	1.00	157	1.50	0.668	5.27	170	0.101	0.002	1.09	37	23.127	0.75	110
McGregor Dam	ND	42	143745052	**2**	7/16/1991	10/30/2004	2	0.60	0.50	0.70	71	2.40	1.02	7.93	80	0.226	0.025	1.35	26	32.700	0.75	218
McVille Dam	ND	46	147439530	**1**	7/22/1992	3/10/1993	2	1.25	1.00	1.50					9	0.194	0.125	0.296	2	14	11	17
Menno	SD	46	125126362	**1**	6/19/2000	8/16/2008	2	1.08	1.00	1.16	6	2.3	1.2	4.63	6	0.454	0.149	0.807	4	46.637	9.537	79.497
Mina Lake	SD	46	144261638	**3**	7/19/1989	8/2/2006	52	1.1	0.15	5.33	101	2	0.3	3.17	101	0.795	0.001	1.35	17	59.732	3.58875	612.38
Mirror Lake	ND	43	143225987	**1**	6/14/1990	3/6/2006	20	1.2	0.60	2.30	77	1.51	0.891	3.34	111	0.142	0.009	0.685	31	24	1.5	98.5
Mitchell	SD	46	125127688	**2**	10/20/1986	8/11/2009	357	1.5	0.30	6.10	195	1.23	0.23	12.4	194	0.255	0.001	1.76	254	40.788	0.28875	476.652

				Classification	Sample Date		Secchi Depth (m)				TN (mg/L)				TP (mg/L)				Chl-a (ppb)			
ReservoirName	State	Ecoregion	ComID	Tier	Min	Max	n	Avg	Minimum	Maximum	n	Avg	Minimum	Maximum	n	Avg	Minimum	Maximum	n	Avg	Minimum	Maximum
Moreau #1	SD	43	151678817	1	6/30/2003	8/11/2009	19	1.2	0.55	2.10	8	1.3	0.86	2.07	8	0.3	0.05	0.556	4	34.65	9.9825	70.29
Mott Watershed Dam	ND	43	0	1	2/27/2003	6/18/2003					4	2.93	2.45	3.74	4	0.049	0.025	0.083	2	1.1	0.75	1.5
MT Carmel Dam	ND	46	0	2	6/25/1991	1/31/2006	7	1.9	0.70	3.10	3	2.21	1.64	3.06	13	0.445	0.138	0.828	7	7.9	0.75	21
Murdo	SD	43	128625721	2	6/14/1989	8/23/2007	19	1.5	0.90	4.11	19	1.4	0.84	2.26	19	0.1	0.02	0.471	10	10.5	2.01	26.4825
Nelson Lake	ND	43	132143765	3	7/8/1993	1/26/2006	5	0.54	0.40	0.60	4	0.900	0.649	1.18	13	0.15	0.077	0.284	5	20.7	11.2	28
New Underwood	SD	43	126839431	1	12/19/1977	9/18/1978					4	1.5	1.23	1.7	4	0.11	0.049	0.172				
New Wall	SD	43	126846536	1	12/19/1977	8/1/2006	23	0.74	0.09	1.19	29	1.5	0.94	2.69	29	0.1	0.003	0.707	12	69.1	5.36	629.8
Newell	SD	43	154897883	2	10/17/1978	8/2/2006	61	1.5	0.27	3.96	46	0.4	0.166	0.81	46	0.0	0.001	0.174	14	3.9	0.057	7.6
Newell City Dam	SD	43	154900241	2	6/27/1989	8/8/2007	18	1.6	0.84	3.00	28	1	0.1	1.26	28	0.0	0.001	0.088	13	4.187	1.60875	8.78625
Niagara Dam	ND	46	149087508	1	7/17/1996	3/11/1997	2	1.3	0.80	1.80					7	0.34	0.27	0.546	2	10	5	15
Nieuwsma Dam	ND	42	145467192	1	7/25/1991	2/11/1992	2	1.55	1.10	2.00					8	0.728	0.514	1.78	2	1.5	1.5	1.5
North Lemmon Lake	ND	43	139877901	2	7/12/1990	3/6/2006	5	2.84	2.00	3.80	5	1.18	1.04	1.37	11	0.056	0.014	0.162	2	9.1	6.1	12
Northgate Dam	ND	46	0	1	7/17/1991	10/19/2003	2	2.75	2.10	3.40	51	1.74	1.42	2.23	60	0.47	0.21	1.13	14	25	1.5	88
Oahe	SD	43	0	4	7/18/2007	7/18/2007	1	3.89	3.89	3.89					1	0.0	0.007	0.007				
Odland Dam	ND	43	74884391	1	7/7/1992	7/16/2009	4	1.83	1.00	2.80					9	0.226	0.136	0.329	3	11	3	17.5
Orman Dam	SD	43	154900319	4	10/17/1978	8/12/2009	71	1.4	0.24	3.88	52	0.43	0.11	1.268	54	0.0	0.001	0.119	24	2.87	0.335	13.695
Patterson Lake	ND	43	70796587	2	7/7/1992	3/7/2006	10	0.63	0.10	1.20	21	1.46	0.865	2.45	70	0.18	0.077	0.639	22	15	1.5	68
Pheasant Lake	ND	46	147908966	1	7/24/1991	10/15/2007	3	1.0	0.60	1.40	72	1.60	1.26	2.54	90	0.525	0.238	0.854	18	9.5	0.75	37.1
Pierpont	SD	46	145283002	2	6/27/1989	7/30/2008	25	1.5	0.30	3.05	22	1	0.1	1.97	22	0.1	0.01	0.166	6	27.1	3.50625	93.8
Pigors	SD	46	145283284	1	6/28/2006	8/2/2006	8	1.35	1.00	2.07	4	1.27	1.18	1.43	4	0.45	0.367	0.53	2	9.29	7.02266	11.55
Pipestem Reservoir	ND	46	147307413	2	5/12/1998	10/18/2008	12	0.82	0.40	1.60	50	1.41	0.615	3.84	50	0.247	0.031	1.16	9	23	1.5	64
Platte Lake	SD	42	148201164	2	6/14/1989	7/23/2007	23	0.22	0.05	0.50	30	2	0.1	6.55	30	0.626	0.0025	2.02	11	93.489	23.5125	330.825
Pocasse	SD	42	145468291	3	7/18/1989	8/31/2009	36	0.81	0.06	2.13	49	5.5	0.35	116.1	21	3.0	0.001	48.7	34	51	2.4	566.511
Potts	SD	42	139475843	1	6/7/2004	7/27/2004	9	3.40	1.75	6.70	11	1.1	0.17	1.67	11	0.5	0.001	0.77	5	8.753	5.98125	11.67375
Rahn Dam	SD	42	149716055	1	7/6/1989	7/14/2005	25	0.89	0.30	2.00	45	2	0.1	6.74	45	0.3	0.001	1.912	17	50.05	16.665	99.16
Raleigh Reservoir	ND	43	141672740	1	7/19/2005	5/24/2006	2	2.75	1.50	4.00	6	1.30	0.839	1.44	6	0.075	0.027	0.129	1	17.6	17.6	17.6
Ravine Lake	SD	46	145672739	1	4/12/1988	7/16/2007	42	0.64	0.20	3.96	90	2.0	0.61	8.6	90	0.74	0.117	5.6	10	47.28	14.4375	98.49
Redfield	SD	46	142195353	1	7/12/1989	7/28/2005	84	0.55	0.27	1.52	53	2.08	1.22	5.25	54	0.643	0.244	1.95	8	41.310	6.31125	115.24
Renwick Dam	ND	48	0	2	5/15/1990	10/19/2004	11	1.2	0.60	2.00	36	0.96	0.524	2.77	52	0.2	0.004	0.827	28	12	0.75	54
Richmond	SD	46	148621844	3	6/10/1986	6/12/2007	133	0.77	0.15	2.45	265	2	0.1	10.34	279	0.3	0.001	3	17	28.321	4.0425	85.244
Roosevelt	SD	42	122539585	1	7/6/1989	7/22/2008	44	1.6	0.61	4.00	35	1.7	0.23	3.61	35	0.4	0.001	0.616	20	17.975	5.73375	55.31936
Rose Hill	SD	42	145673345	1	6/12/1989	7/17/2007	31	1.1	0.30	2.60	67	1.6	0.11	5.03	68	0.4	0.001	0.981	39	31.35	2.585	119.5837
Rosette	SD	42	144258738	1	6/25/1991	1/10/2008	35	0.68	0.10	3.00	53	3.0	0.17	6.17	47	0.7	0.001	1.6	31	53.684	1.503333	133.155
Sather Dam	ND	43	81805103	1	7/26/2006	8/1/2007	1	0.80	0.80	0.80	10	1.6	0.957	2.2	10	0.065	0.022	0.135	5	26	0.75	72.6
Schlecht-Thom Dam	ND	46	147898651	1	7/21/1993	1/25/2006	6	1.5	0.25	3.00	4	1.6	1.41	1.8	13	0.554	0.327	0.803	6	8	1	28.5
Schlecht-Weixel Dam	ND	42	0	1	7/21/1993	1/25/2006	6	1.2	0.60	2.20	4	2	1.63	2	13	0.48	0.172	1.5	6	29	1.5	95.1
Shadehill	SD	43	131721781	4	6/21/1989	7/31/2008	68	1.8	0.64	3.72	50	0.90	0.21	3.282	53	0.0	0.001	0.124	30	5.4	0.5775	56
Sharman Dam	SD	43	125221980	1	6/1/1999	6/1/1999													1	38.86	38.86	38.86
Sheep Creek Dam	ND	43	139813368	1	7/8/1992	3/7/2006	4	1.2	0.90	1.60	35	1.49	1.13	2.98	44	0.250	0.058	1.23	9	31	0.75	91
Short Creek Dam	ND	43	0	1	7/17/1991	10/29/2005	2	0.85	0.80	0.90	49	2.60	2.07	3.53	58	0.851	0.117	2.02	22	12	0.75	78
Silver Lake	ND	46	144884079	1	7/29/1992	3/2/1993	2	0.70	0.60	0.80					6	0.280	0.221	0.374	2	19	11	26
Snow	SD	43	146673245	1	8/7/1991	8/7/1991	1	0.71	0.71	0.71	3	1.51	1.27	1.66	3	0.2	0.002	0.447	1	16.08	16.08	16.08
Sorum Strood	SD	43	144077927	1	6/29/1998	6/29/1998													1	19.095	19.095	19.095
Spring Lake	ND	43	148779257	1	4/24/1996	11/4/1997									78	0.057	0.009	0.308	15	3.6	1.5	15
Spring Lake (Bowman)	ND	43	134311313	1	7/5/1994	2/8/1995	2	2.70	2.60	2.80					9	0.112	0.002	0.603	2	8	7	9
Stanley Reservoir	ND	42	143753483	1	5/20/2009	7/15/2009	2	0.85	0.60	1.10	1	2.77	2.77	2.77	1	0.736	0.736	0.736	1	106.000	106	106
Sully (Sully)	SD	42	139477267	1	7/20/1989	8/18/2009	37	0.49	0.06	2.68	29	4.1	0.22	13.85	29	1.092	0.001	3.39	18	191	5.36	650.76
Sully (Tripp)	SD	42	148209430	1	7/6/1989	8/3/2005	7	0.30	0.18	0.35	6	6.53	2.77	15.61	6	0.793	0.65	1.02	2	63.150	35.51625	90.783
Sweetbriar Dam	ND	43	135580310	2	8/30/1991	9/20/2005	6	1.8	0.60	3.10	48	1.17	0.798	1.95	66	0.197	0.045	0.631	14	27	1.5	140
Timber Creek Dam	SD	46	0	4	8/5/2003	8/5/2003																
Tinsdale	SD	43	126841437	1	6/24/1996	6/24/1996	1	0.80	0.80	0.80					1	0.06	0.06	0.06	1	13.735	13.735	13.735
Tioga Dam	ND	42	143749804	2	5/20/2009	7/15/2009	2	2.05	2.00	2.10									1	16.3	16.3	16.3
Tolna Dam	ND	46	147439400	2	7/22/1992	3/3/1993	2	1.60	1.20	2.00					9	0.215	0.055	0.378	2	8	6	10
Tripp	SD	42	128457635	1	6/21/2005	8/3/2005	5	1.92	1.70	2.00	4	2.32	1.62	2.93	4	0.324	0.212	0.483	2	19.092	3.996667	34.188
Upper Des Lacs	ND	46	145064896	4	5/21/1997	2/23/1998	6	0.47	0.20	1.00					11	0.233	0.111	0.385	4	36	1.5	83

ReservoirName	State	Ecoregion	ComID	Classification Tier	Sample Date Min	Sample Date Max	Secchi Depth (m) n	Secchi Depth (m) Avg	Secchi Depth (m) Minimum	Secchi Depth (m) Maximum	TN (mg/L) n	TN (mg/L) Avg	TN (mg/L) Minimum	TN (mg/L) Maximum	TP (mg/L) n	TP (mg/L) Avg	TP (mg/L) Minimum	TP (mg/L) Maximum	Chl-a (ppb) n	Chl-a (ppb) Avg	Chl-a (ppb) Minimum	Chl-a (ppb) Maximum
Vermillion	SD	46	123218588	**2**	7/6/1989	7/24/2007	121	1.3	0.24	5.30	115	1.9	0.12	12.132	115	0.4	0.001	0.812	53	34.882	1.69125	248.8672
Waggoner Lake	SD	43	128629047	**1**	6/13/1989	8/12/2009	43	1.0	0.50	2.44	79	1.2	0.1	4.1	76	0.1	0.001	0.827	35	58.783	4.49625	609.7
Wall (OLD)	SD	42	145672669	**2**	6/16/1998	6/16/1998													1	12.06	12.06	12.06
Wanalain	SD	42	148196841	**1**	7/30/2007	7/30/2007	1	0.58	0.58	0.58					1	1.016	1.016	1.016				
Warsing Dam	ND	46	143262719	**1**	7/22/1992	3/10/1993	2	1.80	1.10	2.50					9	0.157	0.105	0.311	2	18	6	29
Welk Dam	ND	42	145466064	**1**	5/14/1991	1/3/1992	3	1.1	0.50	2.30					10	0.867	0.648	1.08	2	85	71	98
White Clay	SD	43	126572648	**1**	10/25/1977	8/10/1978					3	0.50	0.36	0.63	2	0.045	0.042	0.047				
White Earth Dam	ND	43	0	**1**	7/15/1992	2/23/1993	2	1.15	1.00	1.30					9	0.251	0.216	0.301	2	3	3	3.4
White Lake Dam	SD	46	144884358	**2**	6/28/1989	7/27/2004	44	0.83	0.33	3.25	95	1.5	0.23	4.67	95	0.146	0.001	0.376	23	82.378	6.1875	600.5175
Whitman Dam	ND	46	149374645	**1**	7/29/1991	2/26/1992	2	1.85	1.80	1.90					9	0.479	0.378	0.617	2	1.5	1.5	1.5
Wilmarth	SD	42	125125821	**1**	7/6/1989	7/19/2006	75	1.4	0.60	3.40	82	1.6	0.17	3.93	80	0.697	0.001	1.612	51	23.137	1.27875	95.04
Wilson Dam	ND	46	147903661	**1**	7/21/1993	1/25/2006	6	2.05	1.10	4.50	4	2	1.37	2	13	0.7	0.5	1.2	6	6	1	15.6
Wolf Creek	SD	43	126572646	**1**	10/25/1977	8/10/1978					3	0.67	0.42	0.81	2	0.059	0.035	0.082				
Wylie Pond	SD	46	0	**1**	7/27/1999	2/20/2001	2	0.64	0.64	0.64	22	0.65	0.16	1.35	22	0.024	0.001	0.066	19	4.344	1.77375	18.315
Yankton	SD	46	128450549	**2**	6/27/1989	7/24/2006	22	1.6	0.76	3.00	24	0.51	0.21	1.15	26	0.035	0.003	0.122	10	7.073	3.2175	17.42093

Section V: Two-way Analysis of Variance

Section V: Two-way Analysis of Variance

MEMO
(External Correspondence)

HoustonEngineering Inc.

To:	Tina Laidlaw	**From:**	Stephanie Johnson, Ph.D., P.E.
Date:	September 27, 2010	**Through:**	Mark R. Deutschman, Ph.D., P.E.
Cc:	File 4965-002 Dennis McIntyre, GLEC	**Subject:**	Summary of Results of Two-Way Analysis of Variance: Plains States Nutrient Water Quality Data

Introduction

The following is intended to communicate the results of multiple two-way analysis of variance analyses performed on water quality data associated with EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8 (i.e. the Nutrient Criteria project). These results were previously submitted to the Nutrient Criteria Project Team in a July 15, 2010 e-mail and discussed during a July 28th conference call.

Two-way analysis of variance tests were requested by the Nutrient Criteria Project Team during a conference call on June 7th, 2010. This request came after results of one-way analysis of variance tests (performed on the reservoir classification tier, EPA Tier 3 Ecoregion, and states) showed considerable uncertainty and no clear indication of whether the classification technique "worked" or not (HEI, 2010a). Project Team members believed that the two-way analysis of variance tests may provide more clarity.

Background

The goal of this work was to perform two-way analysis of variance tests on total phosphorus (TP), secchi depth, and chlorophyll-a (Chl-a) by both EPA Tier 3 Ecoregion and reservoir classification tier. The reservoir classification tier is computed as [(Surface area/Drainage area)*Volume], as discussed in the March 5, 2010 memo (HEI, 2010b). The theory behind the classification is that the eutrophication response of the reservoirs will be unique by tier. Since the TP dataset is the most robust, we use it as an example for purposes of explaining the statistical methods used for this work.

The TP dataset contains 8,437 individual TP measurements collected in 168 reservoirs. We desire to know if the TP measurements are statistically significantly different based on the ecoregion, the classification tier, or the ecoregion and classification tier that the data was collected in. Analysis of variance tests can be used to answer the question.

In this analysis, we have six EPA Tier 3 Ecoregions and four reservoir classification tiers. **Table 1** shows how the TP data are distributed amongst these categories. The number of TP measurements is shown first and the number of reservoirs represented by those measurements follows (in parentheses).

Table 1: Quantification of TP Measurements and Number of Reservoirs by Ecoregion and Reservoir Classification Tier (# of Reservoirs)

Ecoregion[1]	Reservoir Classification Tier			
	1	2	3	4
25	3 (2)	0	9 (1)	0
42	1,068 (28)	205 (5)	20 (1)	1,188 (2)
43	986 (43)	357 (11)	360 (5)	353 (8)
46	1,103 (34)	1,059 (14)	1,465 (7)	11 (1)
47	8 (1)	0	0	0
48	155 (2)	87 (3)	0	0

[1] EPA Tier 3 Ecoregions: 25 = Western High Plains; 42 = Northwestern Glaciated Plains; 43 = Northwestern Great Plains; 46 = Northern Glaciated Plains; 47 = Western Corn Belt Plains; 48 = Lake Agassiz Plain

Since the number of observations in each ecoregion-classification tier category is not equal, this dataset is considered “unequal” or “unbalanced”, which limits the statistical tests that can be used for its analysis. Unbalanced datasets must be analyzed using regression models (Helsel and Hirsch, 2002), an example of which is the General Linear Model (GLM).

Statistical Analysis – Methods

GLM analysis allows us to consider more than one factor as being influential on the TP values, by comparing the means of the TP values within each ecoregion, classification tier, and ecoregion-classification tier category. GLM assumes that the residuals of the resultant model are normally distributed and requires that all categories to be analyzed have data within them. The GLM requirement that residuals be normally distributed was not met when modeling TP values. However, log-transforming the TP values resulted in residuals that are more normally distributed. Therefore, the models built for this work used log-transformed (i.e., ln(TP)) values. Statistical analyses were performed using the Minitab software, Version 16.

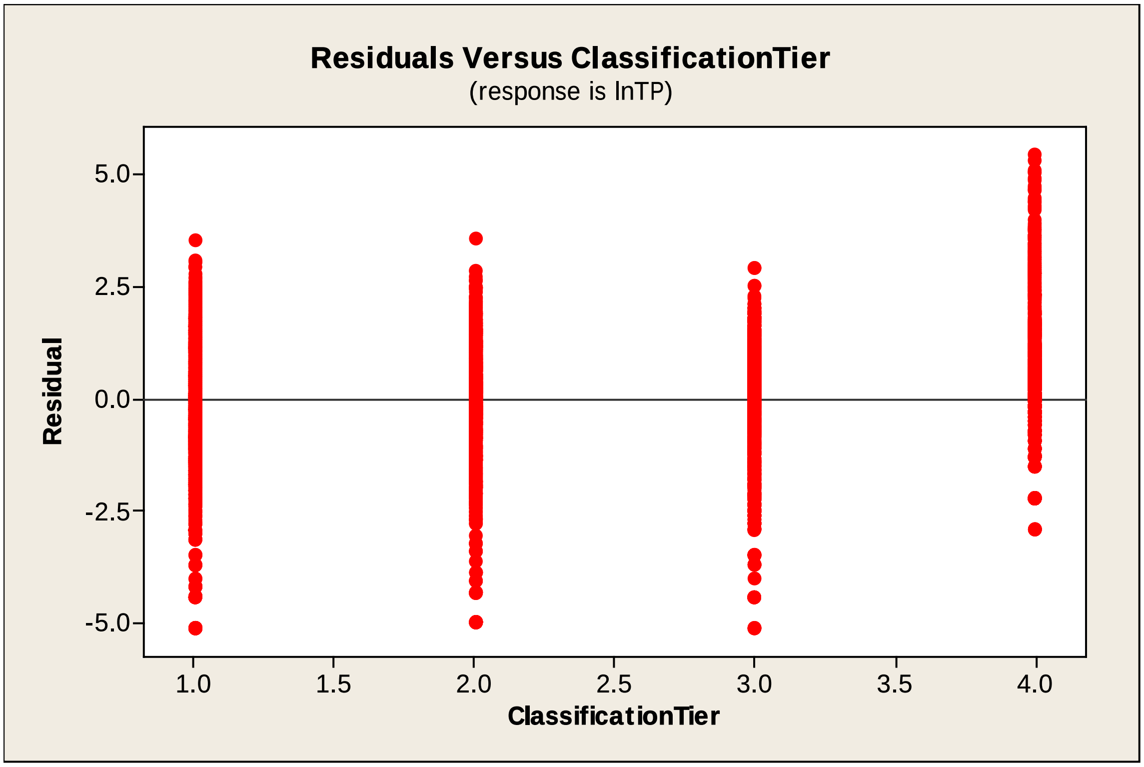
GLM Models for 1-way Analysis of Variance

The first step in this work was to develop a GLM considering only one variable to understand how the inclusion of additional variables would affect the results of the model. The first model created was for ln(TP) vs. Classification Tier; results are shown below.

Results for: Reservoir-TP

General Linear Model: lnTP versus ClassificationTier

```
Factor              Type   Levels  Values
ClassificationTier  fixed       4  1, 2, 3, 4


Analysis of Variance for lnTP, using Adjusted SS for Tests

Source                DF   Seq SS   Adj SS  Adj MS        F      P
ClassificationTier     3   6001.8   6001.8  2000.6  1018.53  0.000
Error               8433  16564.1  16564.1     2.0
Total               8436  22565.9


S = 1.40150   R-Sq = 26.60%   R-Sq(adj) = 26.57%

Grouping Information Using Tukey Method and 95.0% Confidence

ClassificationTier     N  Mean  Grouping
1                   3323  -1.8  A
3                   1854  -1.8  A
2                   1708  -1.9  A
4                   1552  -4.0    B

Means that do not share a letter are significantly different.
```



Results show that the mean ln(TP) values amongst the four classification tiers are not all equal. Pairwise comparison of the values (using the Tukey method with a 95% confidence level), however, shows that Tiers 1, 2, and 3 are not statistically significantly different from one another; Tier 4 is different than the others.

A similar analysis was performed to view the impact of ecoregion on mean ln(TP). Results are shown below.

General Linear Model: lnTP versus Ecoregion

Factor	Type	Levels	Values
Ecoregion	fixed	6	25, 42, 43, 46, 47, 48

Analysis of Variance for lnTP, using Adjusted SS for Tests

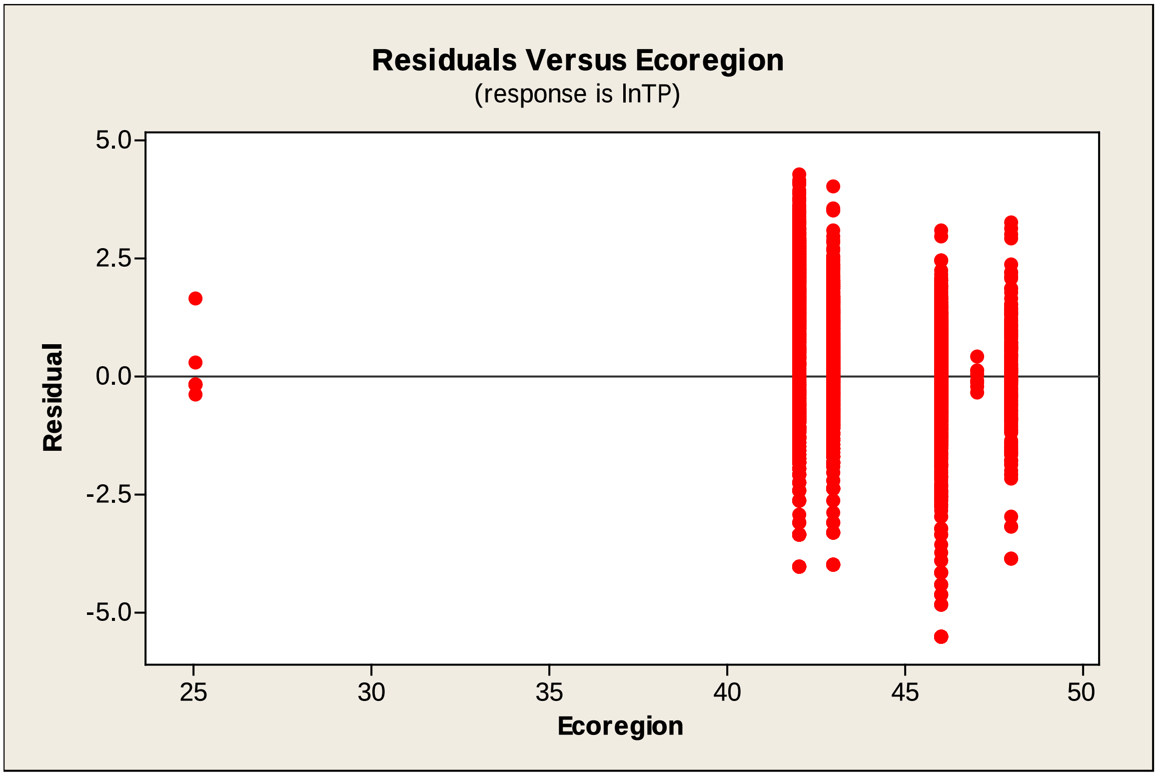
Source	DF	Seq SS	Adj SS	Adj MS	F	P
Ecoregion	5	4646.84	4646.84	929.37	437.27	0.000
Error	8431	17919.04	17919.04	2.13		
Total	8436	22565.88				

S = 1.45787   R-Sq = 20.59%   R-Sq(adj) = 20.55%

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion	N	Mean	Grouping
46	3638	-1.4	A
47	8	-2.0	A B
48	242	-2.3	B
25	12	-2.8	B
42	2481	-2.9	B
43	2056	-2.9	B

Means that do not share a letter are significantly different.



Similar to the analysis based on classification tier, results of this model show that the mean ln(TP) values amongst the six ecoregions are not all equivalent to one another. Pairwise comparison of the values shows that Ecoregions 46 and 47 are not statistically significantly different from one another. Ecoregions 47, 48, 25, 42, and 43 are also not statistically significantly different from one another. Ecoregion 46 and 47, however, are statistically different from Ecoregions 47, 48, 25, 42, and 43.

GLM Model for Two-Way Analysis of Variance

As noted above, the GLM approach requires that all categories used in the model are populated with data. When creating a model to consider ecoregion, classification tier, and ecoregion-classification tier, this is not the case. For example, consider **Table 1** which shows that the 25-II ecoregion-classification tier category doesn't have any data in it. Similarly, the 25-IV, 48-III, 48-IV, 47-II, 47-III, and 47-IV categories are also empty.

To alleviate this problem and enable the GLM approach to be taken, data from ecoregions 25, 47, and 48 were removed from the analysis. The two-way analysis of variance was, therefore, only performed on data from ecoregions 42, 43, and 46. Results of the analysis follow.

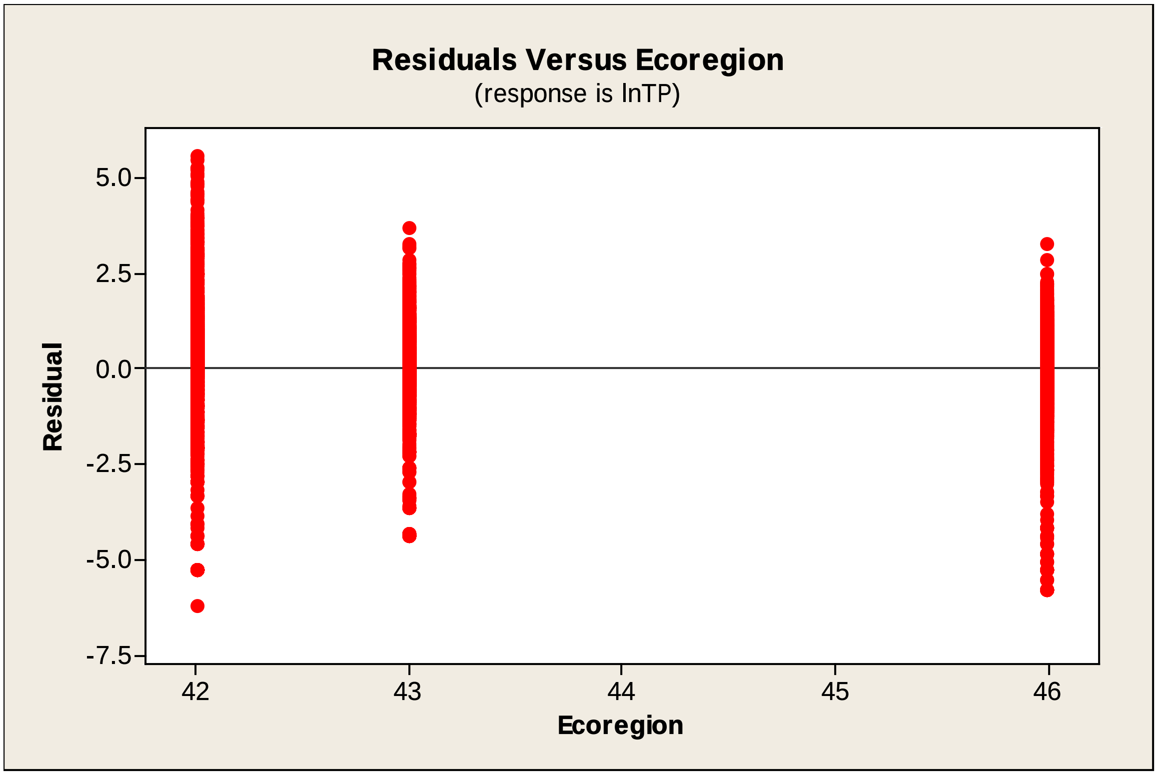
General Linear Model: lnTP versus Ecoregion, ClassificationTier

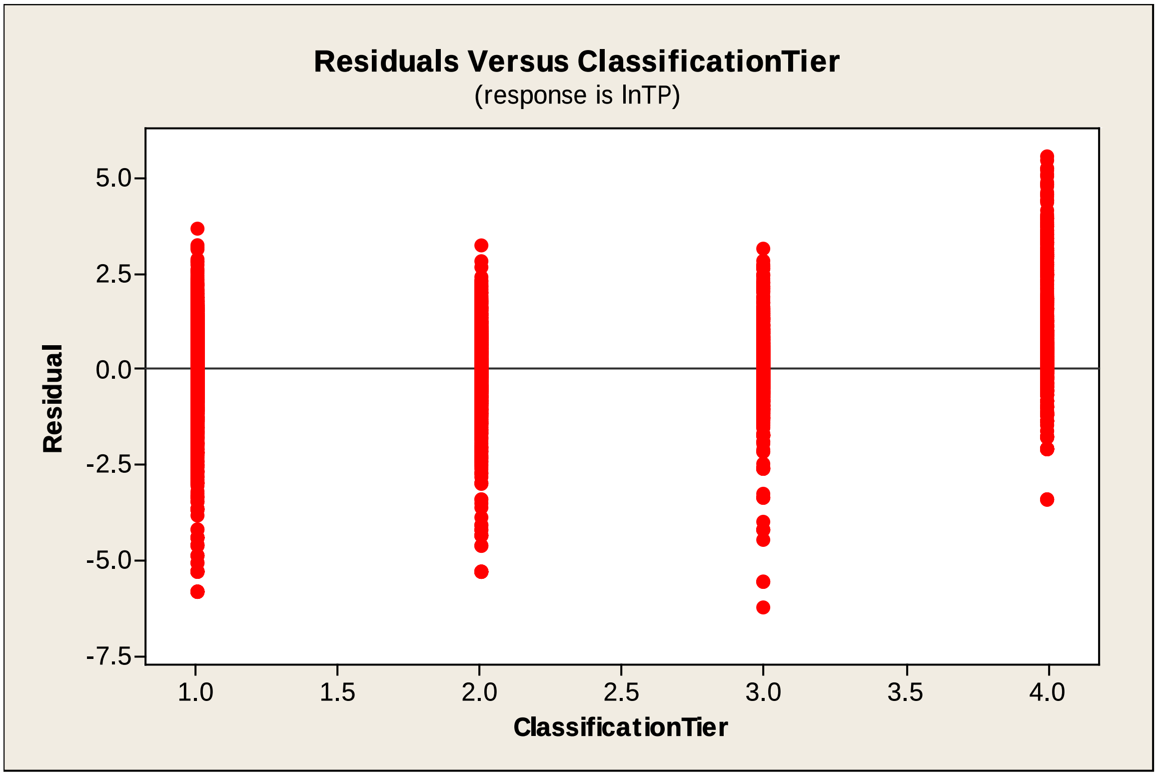
```
Factor              Type   Levels  Values
Ecoregion           fixed       3  42, 43, 46
ClassificationTier  fixed       4  1, 2, 3, 4


Analysis of Variance for lnTP, using Adjusted SS for Tests

Source                         DF    Seq SS    Adj SS  Adj MS       F      P
Ecoregion                       2   4638.15    558.81  279.41  173.18  0.000
ClassificationTier              3   3593.20    233.86   77.95   48.32  0.000
Ecoregion*ClassificationTier    6    823.73    823.73  137.29   85.09  0.000
Error                        8163  13170.18  13170.18    1.61
Total                        8174  22225.27


S = 1.27020   R-Sq = 40.74%   R-Sq(adj) = 40.66%

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion     N  Mean  Grouping
46         3638  -1.4  A
42         2481  -2.1    B
43         2056  -3.1      C

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

ClassificationTier     N  Mean  Grouping
1                   3157  -1.8  A
3                   1845  -1.9  A B
2                   1621  -2.1    B
4                   1552  -3.1      C

Means that do not share a letter are significantly different.
```

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion	ClassificationTier	N	Mean	Grouping
42	3	20	-0.7	A
46	1	1103	-1.1	A
46	3	1465	-1.4	A
46	4	11	-1.6	A B
42	1	1068	-1.6	A B
46	2	1059	-1.6	B
42	2	205	-2.1	B
43	1	986	-2.5	B
43	2	357	-2.6	B
43	4	353	-3.5	C
43	3	360	-3.6	C
42	4	1188	-4.2	D

Means that do not share a letter are significantly different.





Results of the two-way analysis of variance show that the mean ln(TP) values are not all equal amongst ecoregion, classification tier, or ecoregion-classification tier categories. Pairwise comparisons, however, show that while the mean ln(TP) values in this analysis are statistically significantly different amongst ecoregions, they are not all different amongst the classification tier, or ecoregion-classification tier categories.

GLMs for Other Nutrient Data

Similar GLM analyses were performed to consider the effect of ecoregion, classification tier, or ecoregion-classification tier on secchi depth and Chl-a. As in the TP analysis, one-way GLMs were created first, followed by two-way analysis models. Data from Ecoregions 25, 47, and 48 were removed in all cases, due to a lack of data in all ecoregion-classification tier combinations in those areas. Results were similar, with GLM results showing the mean ln(SecchiDepth) and ln(Chl-a) values were not statistically significantly equal amongst all categories in the model. Pairwise comparisons, however, showed that the categories were not all different from one another. Details of these analyses are shown in **Appendices A and B**.

GLMs for Means

The analyses described above and included in **Appendices A and B** were performed on the individual TP, sechhi depth, and Chl-a measurements in each ecoregion, tier, and/or ecoregion-tier combination. These analyses were, therefore, answering the question "are the individual nutrient (and response) values measured in different ecoregions and tiers statistically significantly different?"

Another approach to viewing the difference between the ecoregions, tiers, and/or ecoregion-tier combinations is to ask "are the mean nutrient (and response) values measured in the reservoirs of different ecoregions and tiers statistically significantly different from one another?" To answer that question, the mean TP and Chl-a values were computed for each reservoir and the two-way analysis of variance tests were repeated (this analysis was not performed for secchi depth). Results are shown in **Appendices C and D** and display less variation in mean values than was seen in the raw data (as expected). Outcomes of the pairwise tests show little or no statistically significant differences for the mean Mean(ln(TP)) and mean Mean(ln(Chl-a)) values amongst the ecoregion, classification tier, and ecoregion-classification tier categories.

Conclusion

Results of the two-way analysis of variance tests, using GLMs, did not provide a clear indication that the reservoir classification technique "worked" for the study area reservoirs. Similar to the results of the one-way analysis of variance Kruskal-Wallis testing, results showed that some data categories were different from the others but not that they are each statistically significantly unique. Performing the two-way analysis on the mean reservoir values reduces the uncertainty even more, resulting in less variation and less (or no) difference amongst the groups.

Results of the GLM tests (as presented here) were discussed during a July 28th, 2010 conference call among Project Team members. All participants agreed that the variability in the water quality data was too great to recommend splitting the reservoirs into six or more groups (one for each ecoregion/tier combination) for modeling (which is the next step in the project) and establishing nutrient criteria. Team Members suggested, instead, that two models be built – one for EPA Tier 3 Ecoregion 46 and the other for EPA Tier 3 Ecoregions 42/43. The models will lump the reservoirs of classification tiers 1-3 (in each geographic area) together for simulation and the development of nutrient criteria. Areas outside of Ecoregions 42, 43, and 46 will not be modeled at all, due to their small coverage in the study area and lack of water quality data. Reservoirs in classification tier 4 will also not be included in the modeling because they consistently appear to be statistically significantly different than the other tiers and they represent the largest reservoirs in the study states. These large reservoirs may have site-specific standards developed for them and would, therefore, not be directly subject to findings of this project's modeling effort.

References

Helsel, D.R. and R.M. Hirsch. 2002. Statistical Methods in Water Resources Techniques of Water Resources Investigations, Book 4, chapter A3. U.S. Geological Survey. 522 pp.

Houston Engineering, Inc. March 5, 2010. “Status report on select activities under Tasks 2 & 3 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8”. Memo to Tina Laidlaw at EPA.

Houston Engineering, Inc. September 20, 2010. “Revised report on activities related to water quality data under Tasks 2 & 3 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8”. Memo to Tina Laidlaw at EPA.

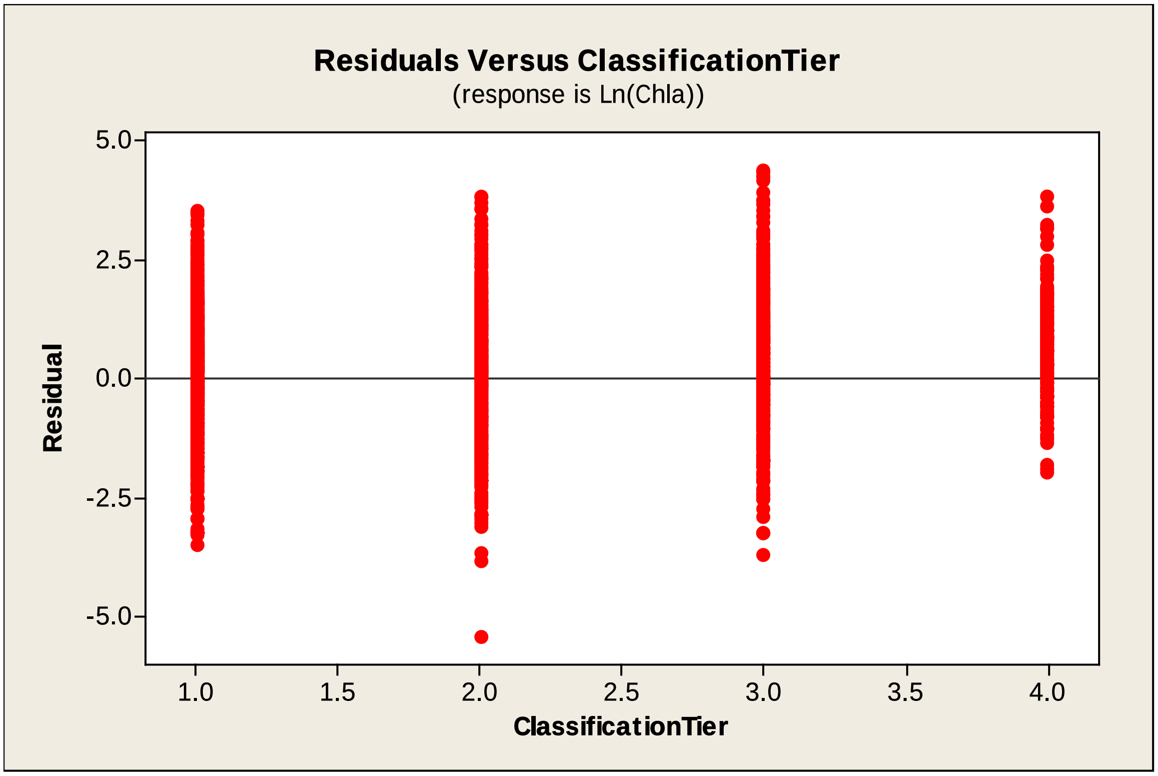
Appendix A: GLM Analysis for Chl-a Data

1-WAY ANALYSIS OF VARIANCE

Run GLM on ln(Chla) data and perform a pairwise analysis using Tukey methods and a 95% level of confidence.

General Linear Model: Ln(Chla) versus ClassificationTier

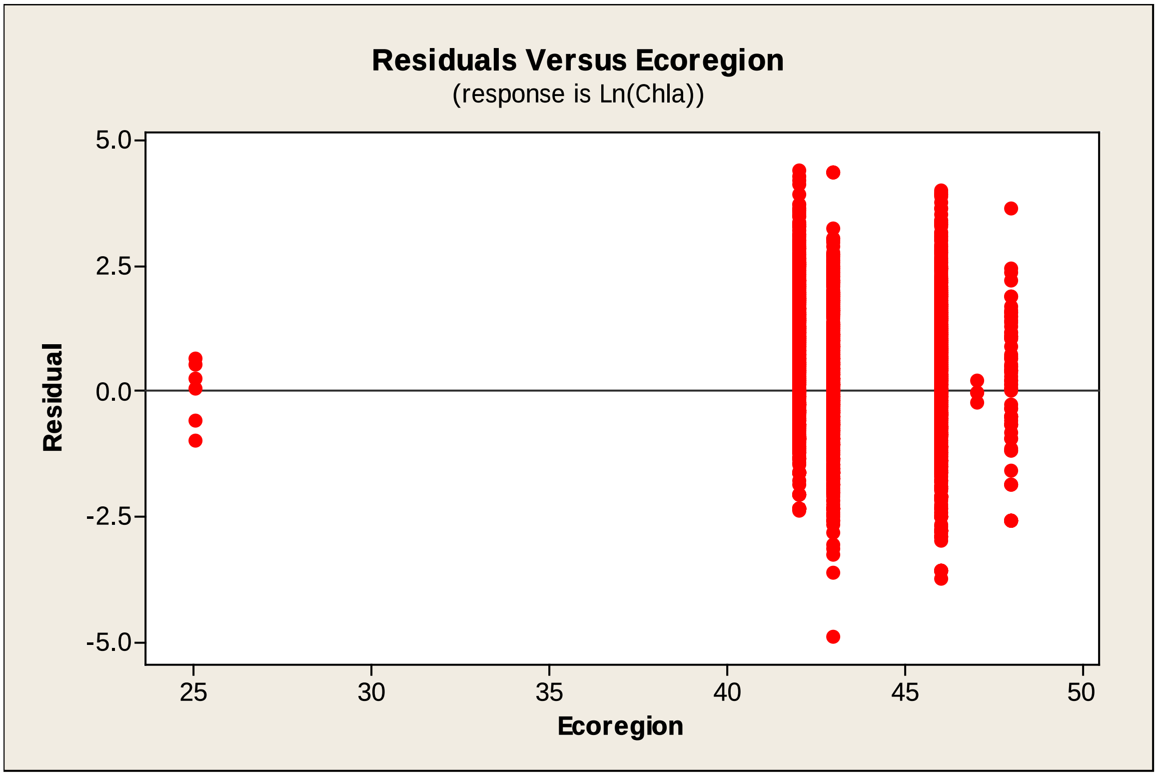
```
Factor              Type   Levels  Values
ClassificationTier  fixed       4  1, 2, 3, 4


Analysis of Variance for Ln(Chla), using Adjusted SS for Tests

Source                DF   Seq SS   Adj SS  Adj MS       F      P
ClassificationTier     3  1969.65  1969.65  656.55  379.00  0.000
Error               3476  6021.57  6021.57    1.73
Total               3479  7991.22


S = 1.31618   R-Sq = 24.65%   R-Sq(adj) = 24.58%


Grouping Information Using Tukey Method and 95.0% Confidence

ClassificationTier     N  Mean  Grouping
1                   1184   2.9  A
2                    766   2.6    B
3                    916   2.1      C
4                    614   0.8        D

Means that do not share a letter are significantly different.
```



General Linear Model: Ln(Chla) versus Ecoregion

```
Factor     Type   Levels  Values
Ecoregion  fixed       6  25, 42, 43, 46, 47, 48


Analysis of Variance for Ln(Chla), using Adjusted SS for Tests

Source       DF    Seq SS    Adj SS  Adj MS      F      P
Ecoregion     5   183.103   183.103  36.621  16.29  0.000
Error      3474  7808.115  7808.115   2.248
Total      3479  7991.217


S = 1.49920   R-Sq = 2.29%   R-Sq(adj) = 2.15%
```

```
Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion     N  Mean  Grouping
47            4   4.2  A
46         1611   2.5  A
48           80   2.3  A B
25            6   2.2  A B
42         1027   2.0    B
43          752   2.0    B

Means that do not share a letter are significantly different.
```



2-WAY ANALYSIS OF VARIANCE

Trim out all reservoirs that are in Ecoregions 25, 47, and 48 (since there's not four tiers in each of these ecoregions) and run a GLM to look at ln(Chla) vs. Ecoregion, Tier, and Ecoregion/Tier

Results for: Worksheet 3

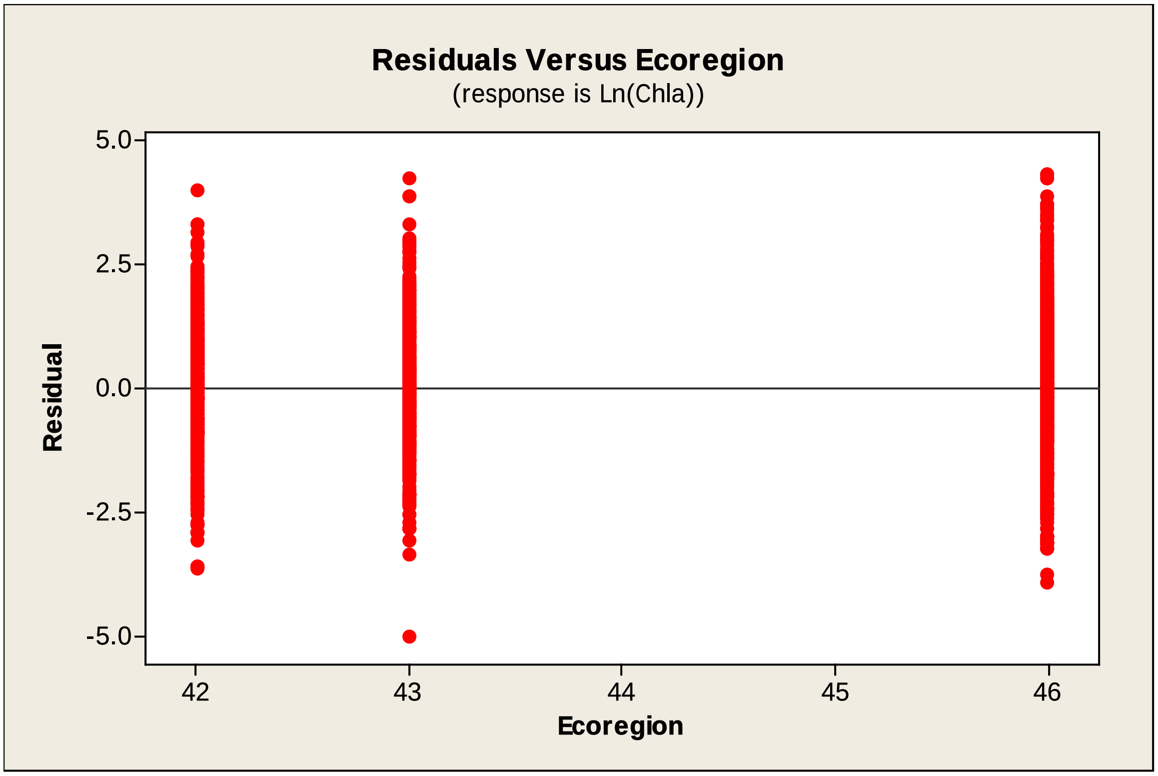
General Linear Model: Ln(Chla) versus Ecoregion, ClassificationTier

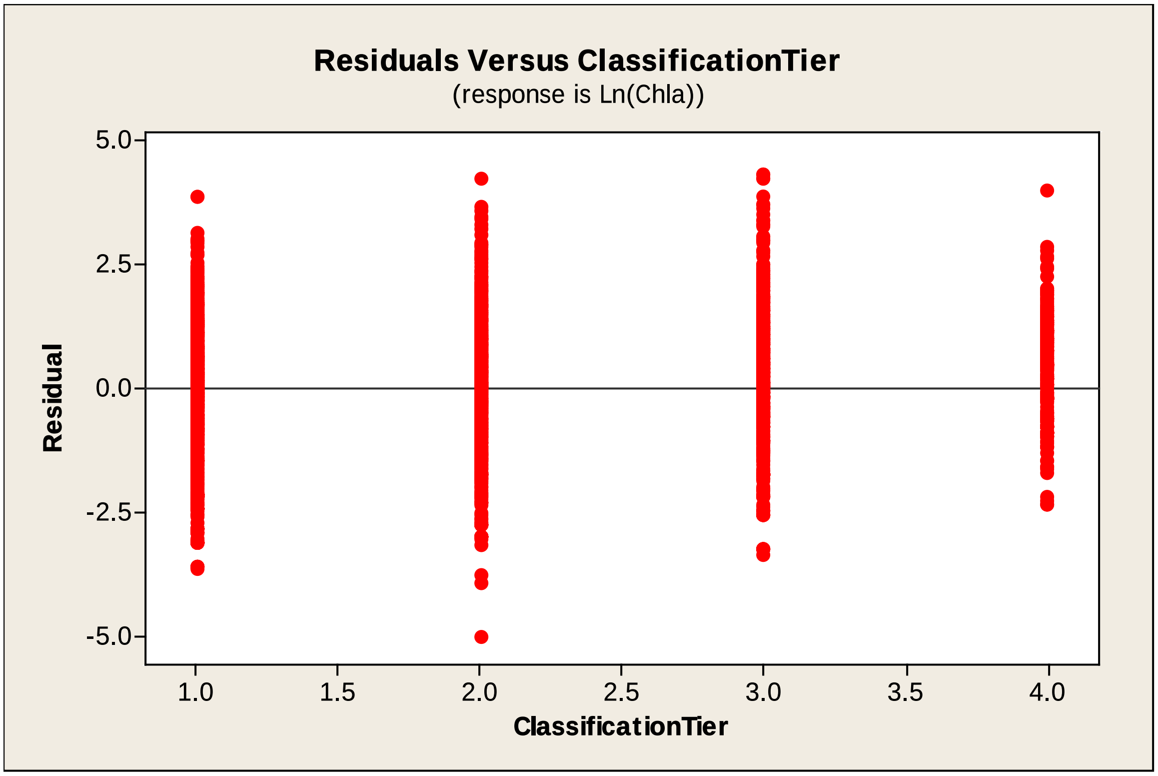
```
Factor              Type   Levels  Values
Ecoregion           fixed       3  42, 43, 46
ClassificationTier  fixed       4  1, 2, 3, 4


Analysis of Variance for Ln(Chla), using Adjusted SS for Tests

Source                          DF   Seq SS   Adj SS  Adj MS      F      P
Ecoregion                        2   167.20    61.94   30.97  18.73  0.000
ClassificationTier               3  1870.17   147.09   49.03  29.64  0.000
Ecoregion*ClassificationTier     6   186.38   186.38   31.06  18.78  0.000
Error                         3378  5587.20  5587.20    1.65
Total                         3389  7810.95


S = 1.28608   R-Sq = 28.47%   R-Sq(adj) = 28.24%

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion     N  Mean  Grouping
46         1611   2.6  A
42         1027   2.4  A
43          752   1.9    B

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

ClassificationTier     N  Mean  Grouping
1                   1144   2.9  A
2                    722   2.5    B
3                    910   2.3    B
4                    614   1.5      C

Means that do not share a letter are significantly different.
```

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion	ClassificationTier	N	Mean	Grouping
42	1	465	3.3	A
42	3	34	3.0	A B
46	1	359	2.8	B
46	4	4	2.7	A B C
46	2	525	2.7	B C
43	1	320	2.5	B C
42	2	70	2.5	B C
46	3	723	2.2	C
43	2	127	2.2	C
43	3	153	1.8	C
43	4	152	1.2	C
42	4	458	0.6	D

Means that do not share a letter are significantly different.




Residuals Versus ClassificationTier
(response is Ln(Chla))
Residual
5.0
2.5
0.0
-2.5
-5.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
ClassificationTier

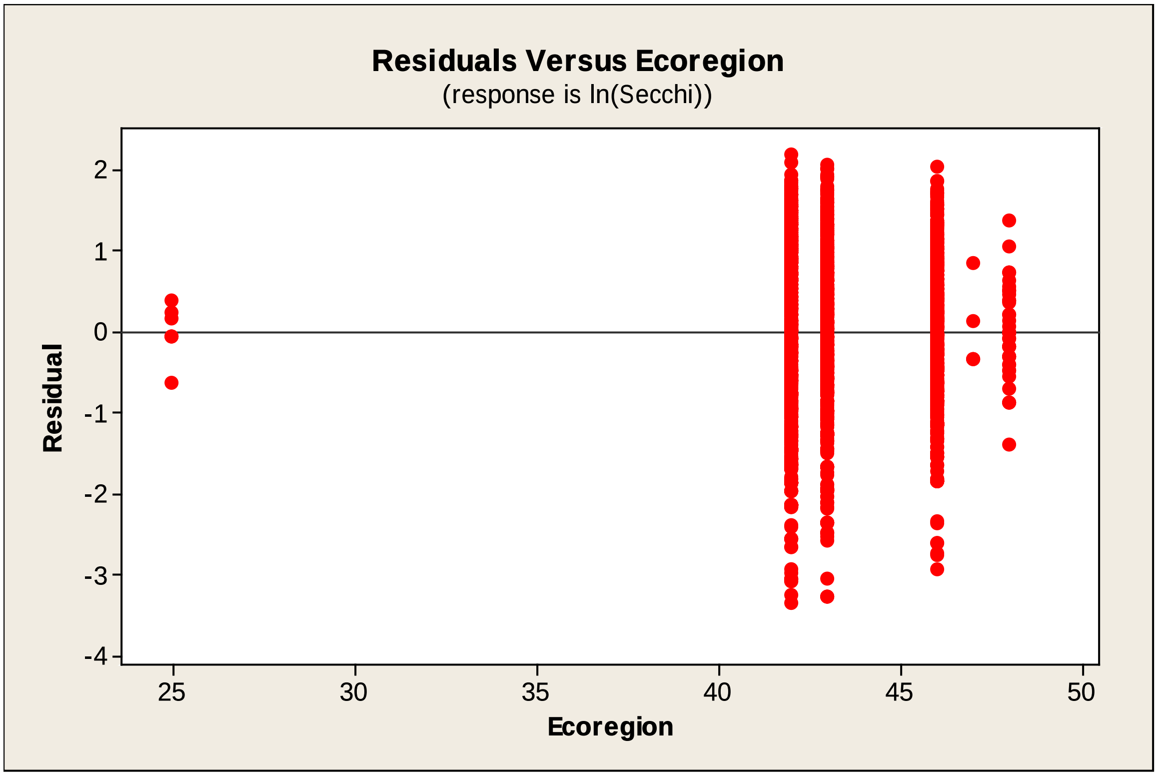
Appendix B: GLM Analysis for Secchi Depth Data

1-WAY ANALYSIS OF VARIANCE

Run GLM on ln(SecchiDepth) data and perform a pairwise analysis using Tukey methods and a 95% level of confidence.

General Linear Model: ln(Secchi) versus Ecoregion

```
Factor     Type   Levels  Values
Ecoregion  fixed       6  25, 42, 43, 46, 47, 48


Analysis of Variance for ln(Secchi), using Adjusted SS for Tests

Source       DF    Seq SS    Adj SS  Adj MS      F      P
Ecoregion     5    92.217    92.217  18.443  23.70  0.000
Error      4580  3564.278  3564.278   0.778
Total      4585  3656.495


S = 0.882172   R-Sq = 2.52%   R-Sq(adj) = 2.42%


Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion     N  Mean  Grouping
42         1891   0.3  A
48           43   0.2  A B
43         1016   0.1    B
25            7  -0.1  A B
46         1624  -0.1    B
47            5  -0.4  A B

Means that do not share a letter are significantly different.
```



General Linear Model: ln(Secchi) versus ClassificationTier

Factor	Type	Levels	Values
ClassificationTier	fixed	4	1, 2, 3, 4

Analysis of Variance for ln(Secchi), using Adjusted SS for Tests

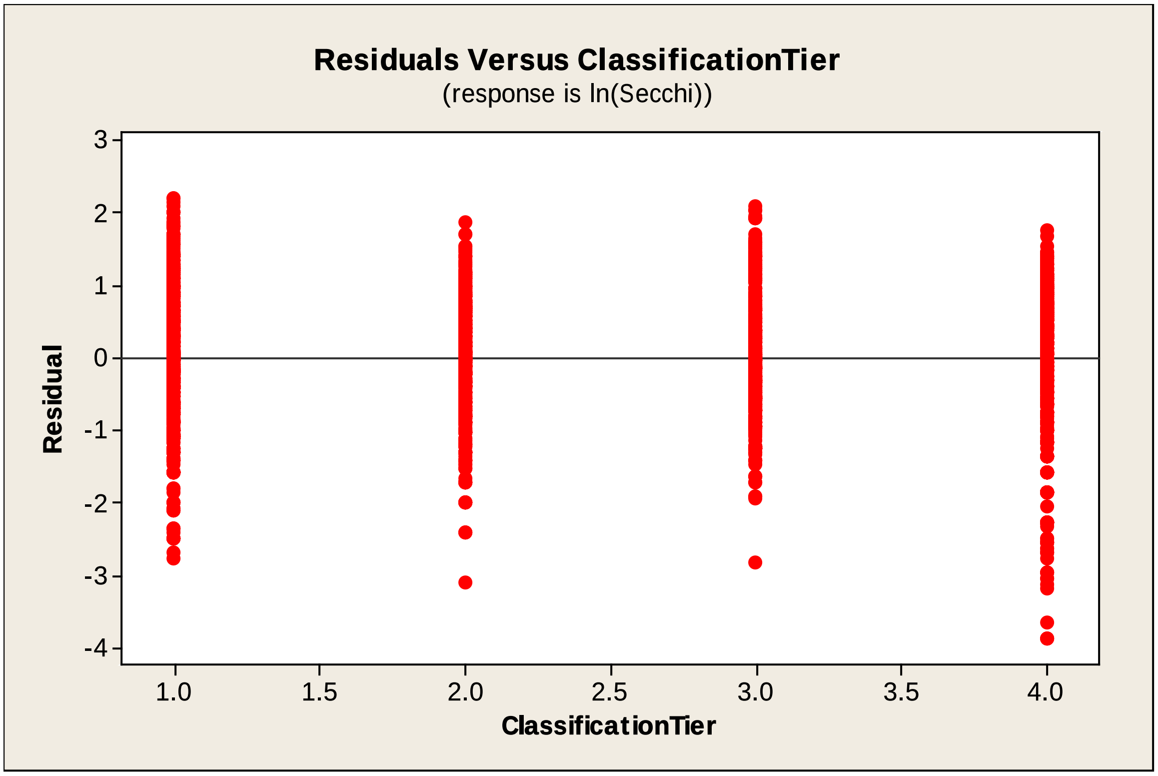
Source	DF	Seq SS	Adj SS	Adj MS	F	P
ClassificationTier	3	695.87	695.87	231.96	358.99	0.000
Error	4582	2960.62	2960.62	0.65		
Total	4585	3656.49				

S = 0.803830   R-Sq = 19.03%   R-Sq(adj) = 18.98%

Grouping Information Using Tukey Method and 95.0% Confidence

ClassificationTier	N	Mean	Grouping
4	1305	0.7	A
2	1036	0.1	B
3	513	0.0	B
1	1732	-0.3	C

Means that do not share a letter are significantly different.


Residuals Versus ClassificationTier
(response is ln(Secchi))
Residual
3
2
1
0
-1
-2
-3
-4
1.0
1.5
2.0
2.5
3.0
3.5
4.0
ClassificationTier

2-WAY ANALYSIS OF VARIANCE

Trim out all reservoirs that are in Ecoregions 25, 47, and 48 (since there's not four tiers in each of these ecoregions) and run a GLM to look at ln(SecchiDepth) vs. Ecoregion, Tier, and Ecoregion/Tier

Results for: Worksheet 3

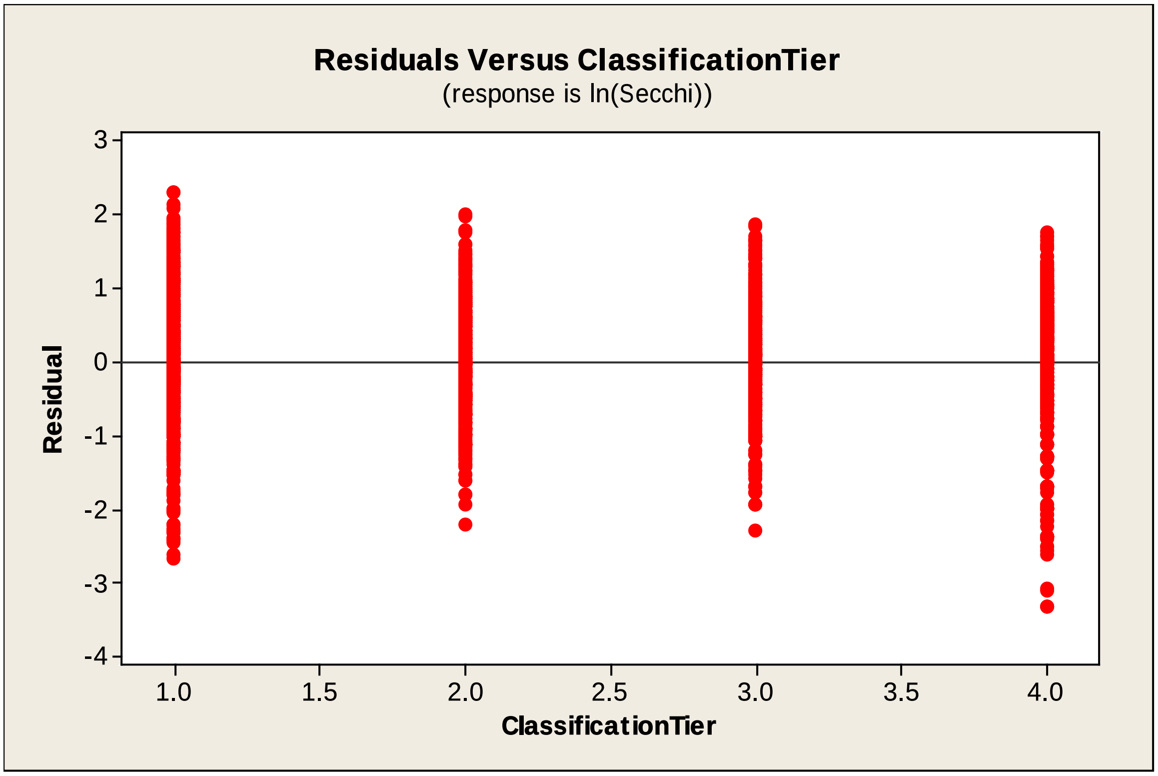
General Linear Model: ln(Secchi) versus ClassificationTier, Ecoregion

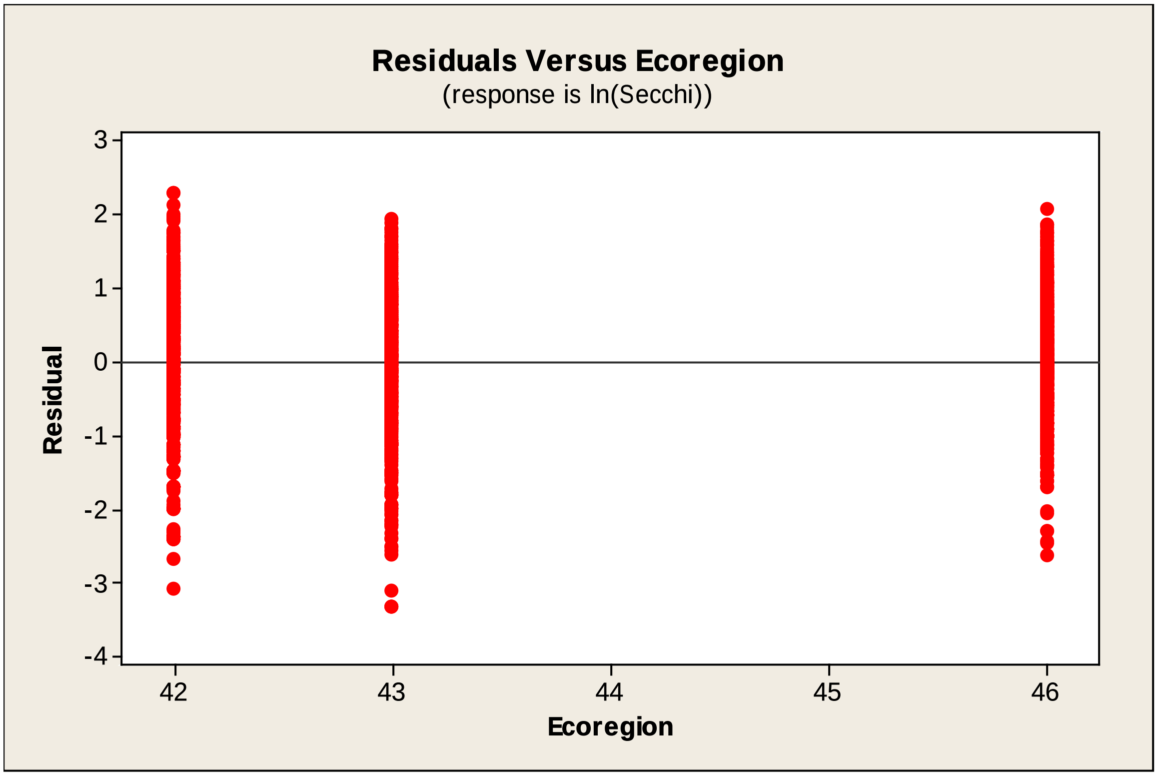
```
Factor              Type   Levels  Values
ClassificationTier  fixed       4  1, 2, 3, 4
Ecoregion           fixed       3  42, 43, 46


Analysis of Variance for ln(Secchi), using Adjusted SS for Tests

Source                          DF    Seq SS    Adj SS  Adj MS      F      P
ClassificationTier               3   701.506    16.255   5.418   9.14  0.000
Ecoregion                        2     1.093    46.815  23.407  39.47  0.000
ClassificationTier*Ecoregion     6   257.574   257.574  42.929  72.40  0.000
Error                         4519  2679.691  2679.691   0.593
Total                         4530  3639.865


S = 0.770054   R-Sq = 26.38%   R-Sq(adj) = 26.20%

Grouping Information Using Tukey Method and 95.0% Confidence

ClassificationTier     N  Mean  Grouping
4                   1305  -0.0  A
3                    507  -0.0  A
1                   1699  -0.3  A
2                   1020  -0.3  A

Means that do not share a letter are significantly different.


Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion     N  Mean  Grouping
43         1016   0.1  A
42         1891  -0.3    B
46         1624  -0.3    B

Means that do not share a letter are significantly different.
```

Grouping Information Using Tukey Method and 95.0% Confidence

ClassificationTier	Ecoregion	N	Mean	Grouping
4	42	1076	0.8	A
3	43	166	0.6	B
2	46	766	0.2	C
4	43	223	0.1	C D
2	43	222	-0.1	D E
1	43	405	-0.1	D E
3	46	305	-0.2	E
1	46	547	-0.4	F
1	42	747	-0.4	F
3	42	36	-0.5	E F G
4	46	6	-1.0	E F G
2	42	32	-1.1	G

Means that do not share a letter are significantly different.





Residuals Versus Ecoregion
(response is ln(Secchi))
Residual
Ecoregion
3
2
1
0
-1
-2
-3
-4
42
43
44
45
46

Appendix C: GLM Analysis for Reservoir Mean TP Values

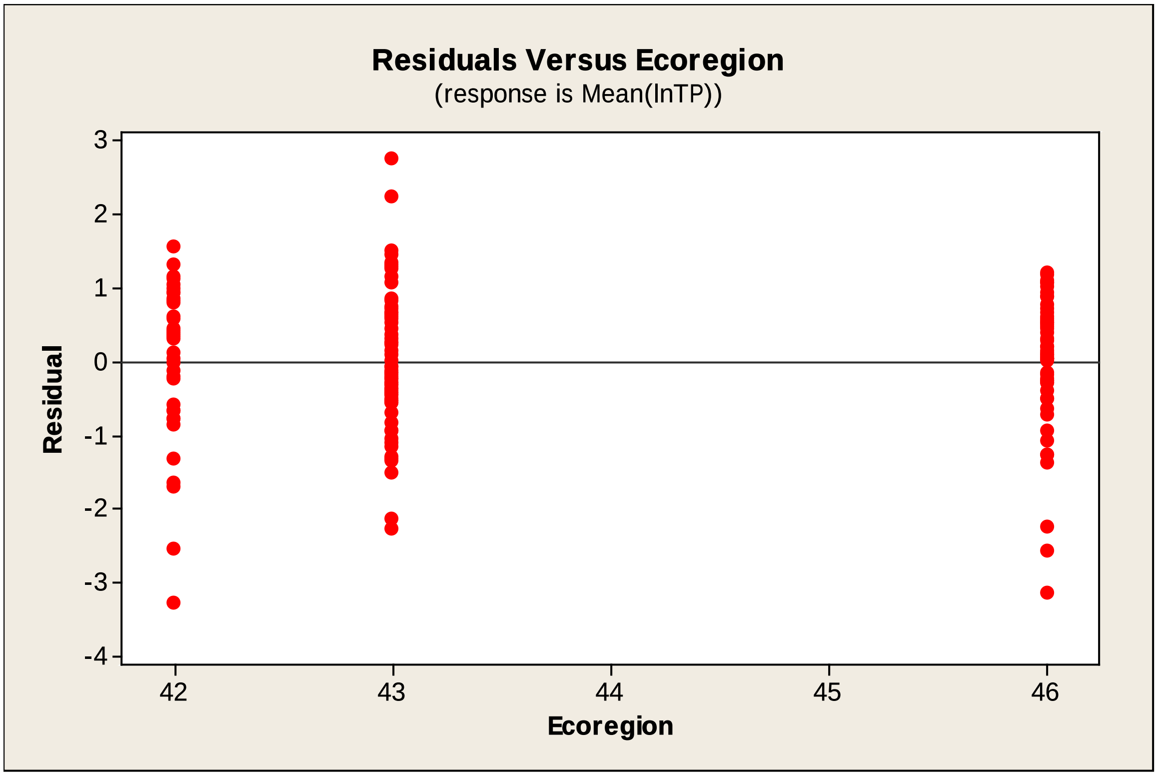
1-WAY ANALYSIS OF VARIANCE

Started with dataset for only Ecoregions 42, 46, and 48 (other Ecoregions trimmed out due to lack of data in all tiers). Run GLM on the Mean TP values (actually use Mean (ln(TP)) to encourage Normality) for each reservoir and compare groups using Tukey method and 95% confidence level.

Results for: Reservoir-TP

General Linear Model: Mean(lnTP) versus Ecoregion

```
Factor     Type   Levels  Values
Ecoregion  fixed       3  42, 43, 46


Analysis of Variance for Mean(lnTP), using Adjusted SS for Tests

Source      DF   Seq SS   Adj SS  Adj MS      F      P
Ecoregion    2   67.958   67.958  33.979  36.20  0.000
Error      156  146.410  146.410   0.939
Total      158  214.368


S = 0.968774   R-Sq = 31.70%   R-Sq(adj) = 30.83%

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion   N  Mean  Grouping
42         36  -1.3  A
46         56  -1.4  A
43         67  -2.7    B

Means that do not share a letter are significantly different.
```



General Linear Model: Mean(lnTP) versus Tier

Factor	Type	Levels	Values
Tier	fixed	4	1, 2, 3, 4

Analysis of Variance for Mean(lnTP), using Adjusted SS for Tests

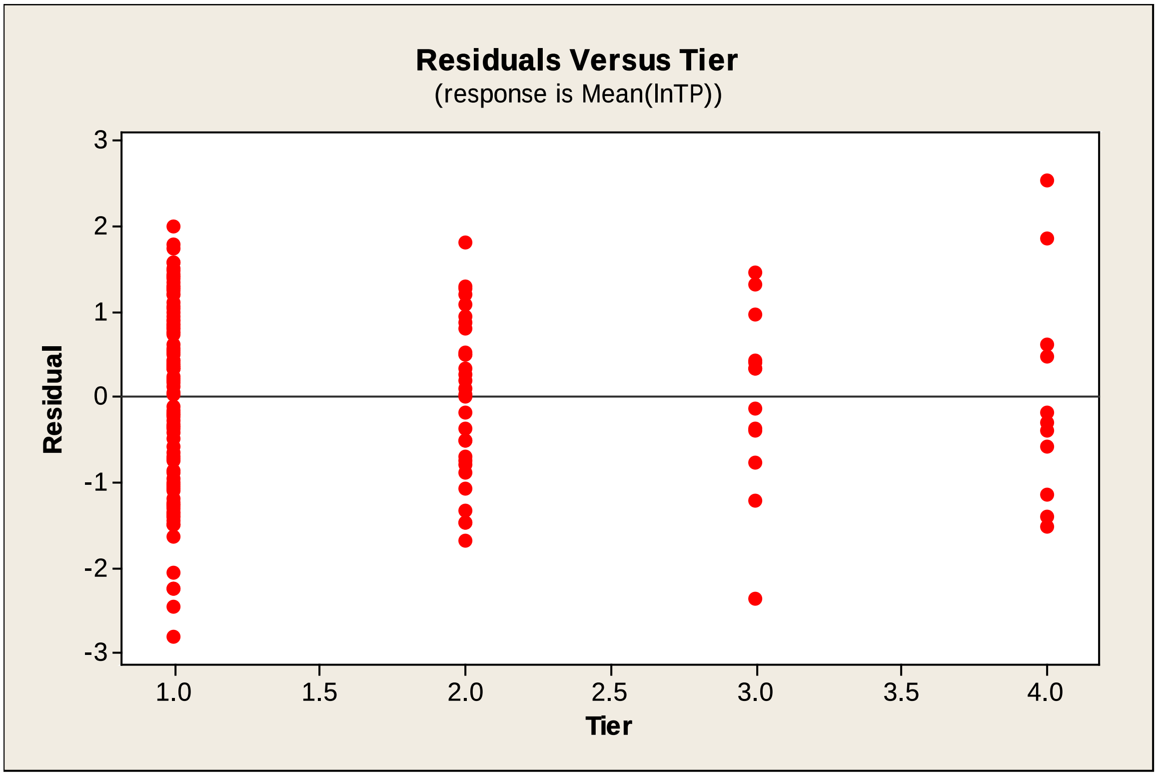
Source	DF	Seq SS	Adj SS	Adj MS	F	P
Tier	3	31.750	31.750	10.583	8.98	0.000
Error	155	182.618	182.618	1.178		
Total	158	214.368				

S = 1.08544   R-Sq = 14.81%   R-Sq(adj) = 13.16%

Grouping Information Using Tukey Method and 95.0% Confidence

Tier	N	Mean	Grouping
3	13	-1.7	A
1	105	-1.7	A
2	30	-2.2	A
4	11	-3.4	B

Means that do not share a letter are significantly different.



2-WAY ANALYSIS OF VARIANCE

General Linear Model: Mean(lnTP) versus Ecoregion, Tier

```
Factor     Type   Levels  Values
Ecoregion  fixed       3  42, 43, 46
Tier       fixed       4  1, 2, 3, 4
```

```
Analysis of Variance for Mean(lnTP), using Adjusted SS for Tests

Source          DF    Seq SS    Adj SS   Adj MS      F      P
Ecoregion        2   67.9581   25.5910  12.7955  15.31  0.000
Tier             3   20.6953   11.4479   3.8160   4.56  0.004
Ecoregion*Tier   6    2.8233    2.8233   0.4706   0.56  0.759
Error          147  122.8910  122.8910   0.8360
Total          158  214.3676


S = 0.914327   R-Sq = 42.67%   R-Sq(adj) = 38.38%
```

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion	N	Mean	Grouping
46	56	-1.5	A
42	36	-1.6	A
43	67	-2.9	B

Means that do not share a letter are significantly different.

Grouping Information Using Tukey Method and 95.0% Confidence

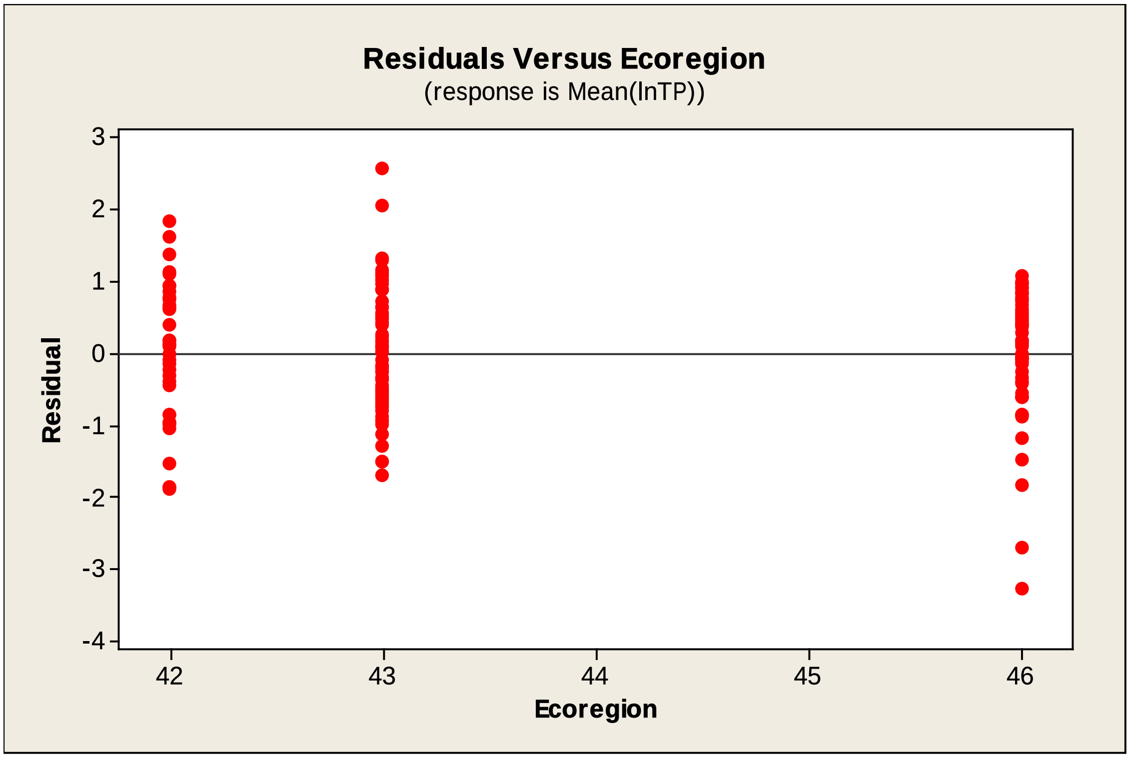
Tier	N	Mean	Grouping
3	13	-1.5	A
1	105	-1.6	A
2	30	-2.2	A
4	11	-2.7	A

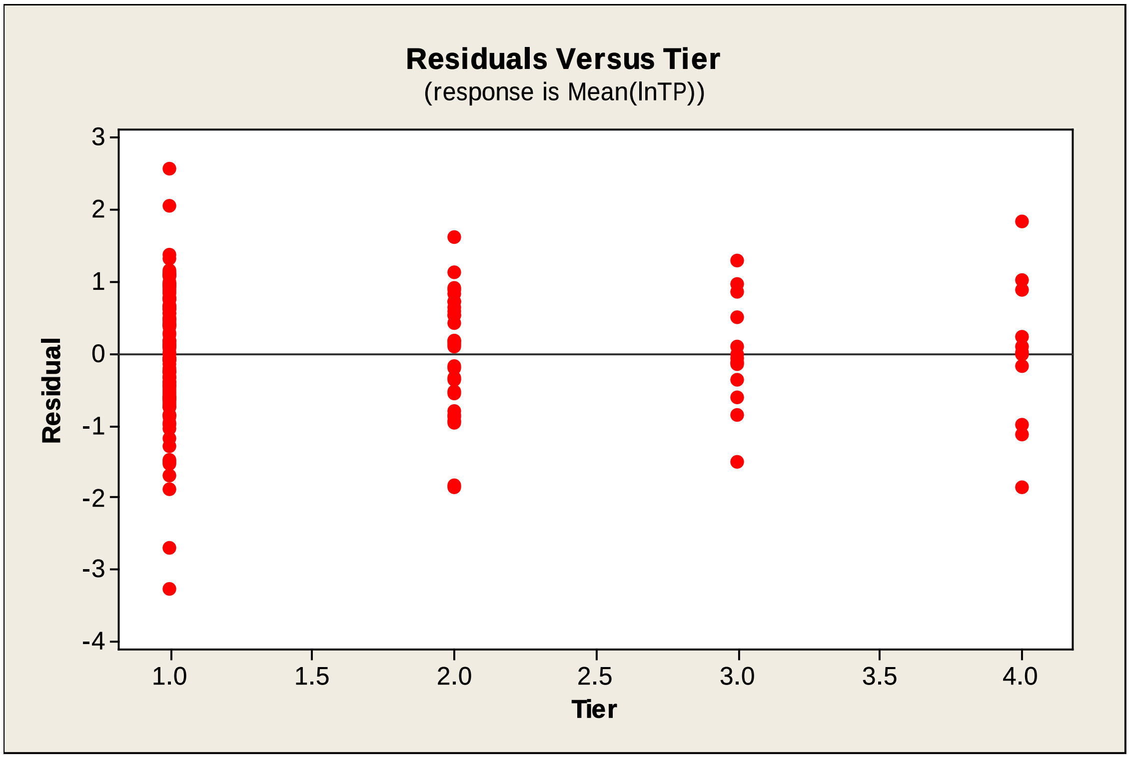
Means that do not share a letter are significantly different.

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion	Tier	N	Mean	Grouping
42	3	1	-0.7	A
42	1	28	-1.1	A
46	3	7	-1.2	A
46	1	34	-1.3	A
46	4	1	-1.6	A
46	2	14	-1.8	A
42	2	5	-2.0	A
43	1	43	-2.5	A
43	3	5	-2.5	A
43	2	11	-2.7	A
42	4	2	-2.7	A
43	4	8	-3.8	A

Means that do not share a letter are significantly different.


Residuals Versus Ecoregion
(response is Mean(lnTP))
Residual
3
2
1
0
-1
-2
-3
-4
42
43
44
45
46
Ecoregion


Residuals Versus Tier
(response is Mean(lnTP))
Residual
3
2
1
0
-1
-2
-3
-4
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Tier

Appendix D: GLM Analysis for Reservoir Mean Chl-a Values

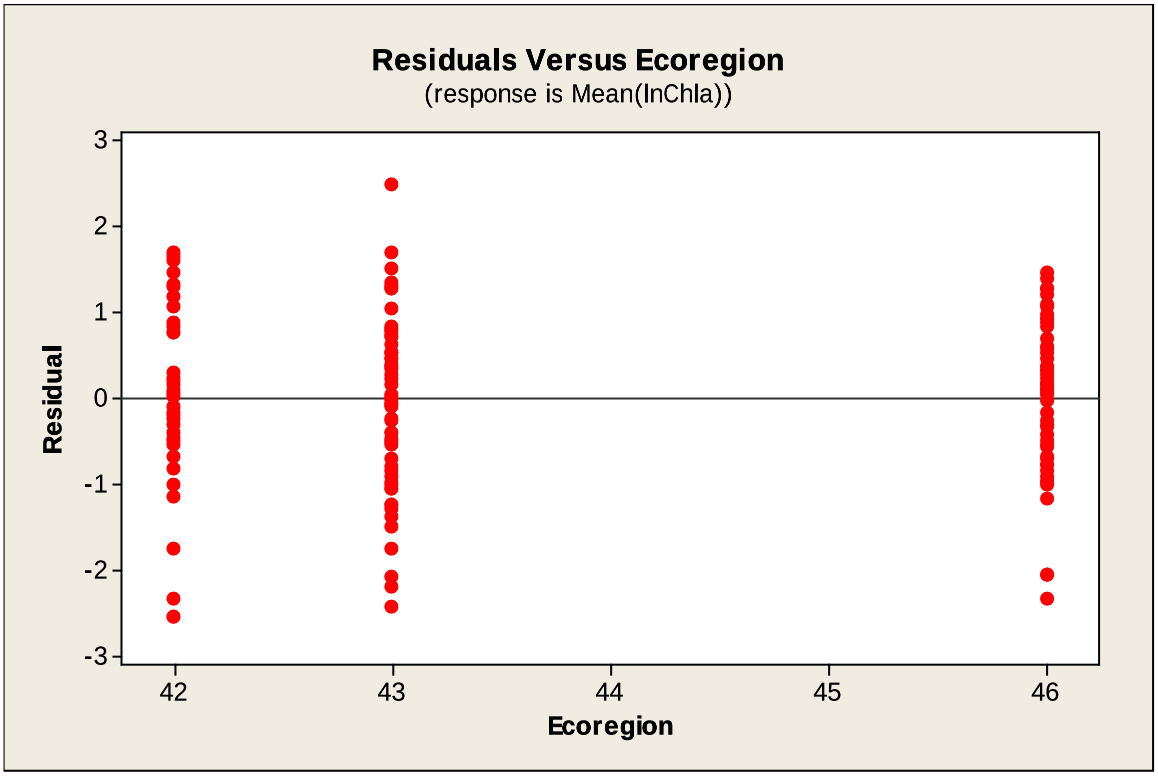
1-WAY ANALYSIS OF VARIANCE

Started with dataset for only Ecoregions 42, 46, and 48 (other Ecoregions trimmed out due to lack of data in all tiers). Run GLM on the Mean Chl-a value (actually use Mean (ln(Chl-a)) to encourage Normality) for each reservoir and compare groups using Tukey methods and 95% confidence level.

Results for: Worksheet 2

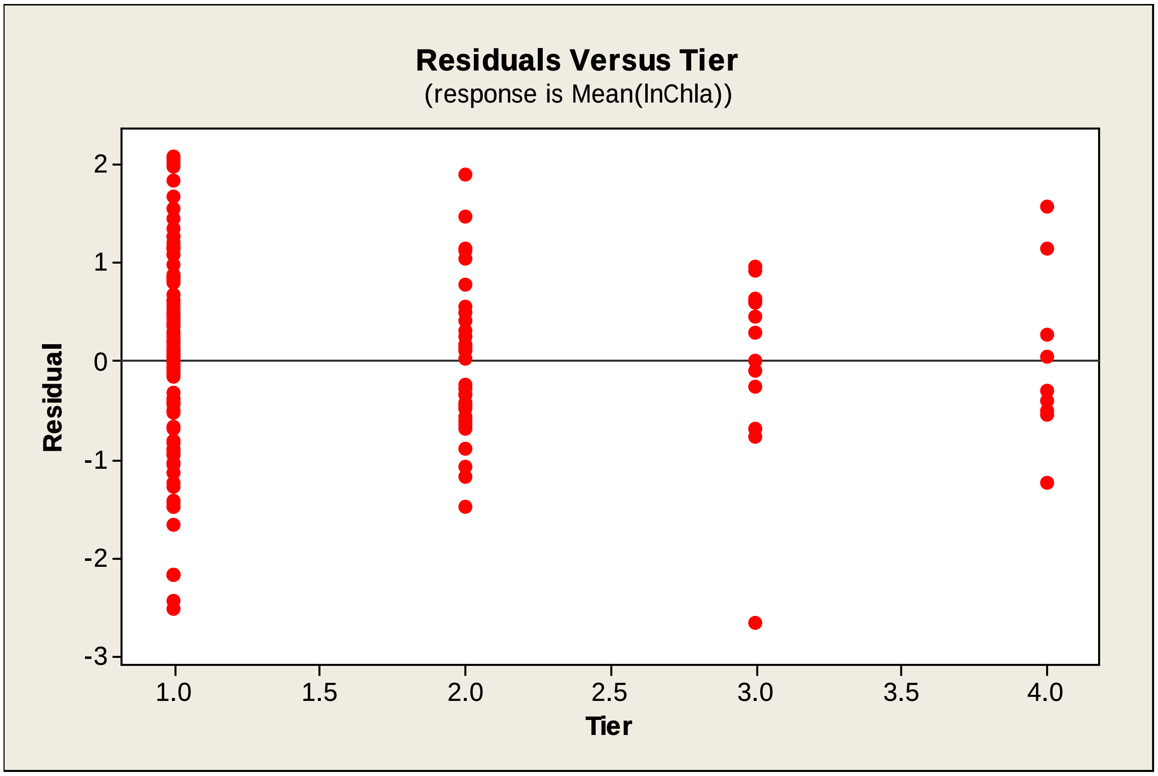
General Linear Model: Mean(lnChla) versus Ecoregion

```
Factor     Type   Levels  Values
Ecoregion  fixed       3  42, 43, 46


Analysis of Variance for Mean(lnChla), using Adjusted SS for Tests

Source      DF   Seq SS   Adj SS  Adj MS     F      P
Ecoregion    2   15.632   15.632   7.816  7.70  0.001
Error      155  157.264  157.264   1.015
Total      157  172.897


S = 1.00728   R-Sq = 9.04%   R-Sq(adj) = 7.87%


Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion   N  Mean  Grouping
42         37   3.0  A
46         55   2.5  A B
43         66   2.1    B

Means that do not share a letter are significantly different.
```



General Linear Model: Mean(lnChla) versus Tier

```
Factor  Type   Levels  Values
Tier    fixed       4  1, 2, 3, 4


Analysis of Variance for Mean(lnChla), using Adjusted SS for Tests

Source   DF   Seq SS   Adj SS  Adj MS     F      P
Tier      3   16.590   16.590   5.530  5.45  0.001
Error   154  156.307  156.307   1.015
Total   157  172.897


S = 1.00746   R-Sq = 9.60%   R-Sq(adj) = 7.83%


Grouping Information Using Tukey Method and 95.0% Confidence

Tier    N  Mean  Grouping
1     103   2.6  A
2      33   2.4  A
3      13   2.4  A
4       9   1.2    B

Means that do not share a letter are significantly different.
```



2-WAY ANALYSIS OF VARIANCE

General Linear Model: Mean(lnChla) versus Ecoregion, Tier

Factor	Type	Levels	Values
Ecoregion	fixed	3	42, 43, 46
Tier	fixed	4	1, 2, 3, 4

Analysis of Variance for Mean(lnChla), using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Ecoregion	2	15.6324	5.0856	2.5428	2.70	0.070
Tier	3	12.6225	6.4783	2.1594	2.30	0.080
Ecoregion*Tier	6	7.3125	7.3125	1.2188	1.30	0.263
Error	146	137.3292	137.3292	0.9406		
Total	157	172.8967				

S = 0.969851   R-Sq = 20.57%   R-Sq(adj) = 14.59%

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion	N	Mean	Grouping
46	55	2.6	A
42	37	2.3	A
43	66	1.9	A

Means that do not share a letter are significantly different.

Grouping Information Using Tukey Method and 95.0% Confidence

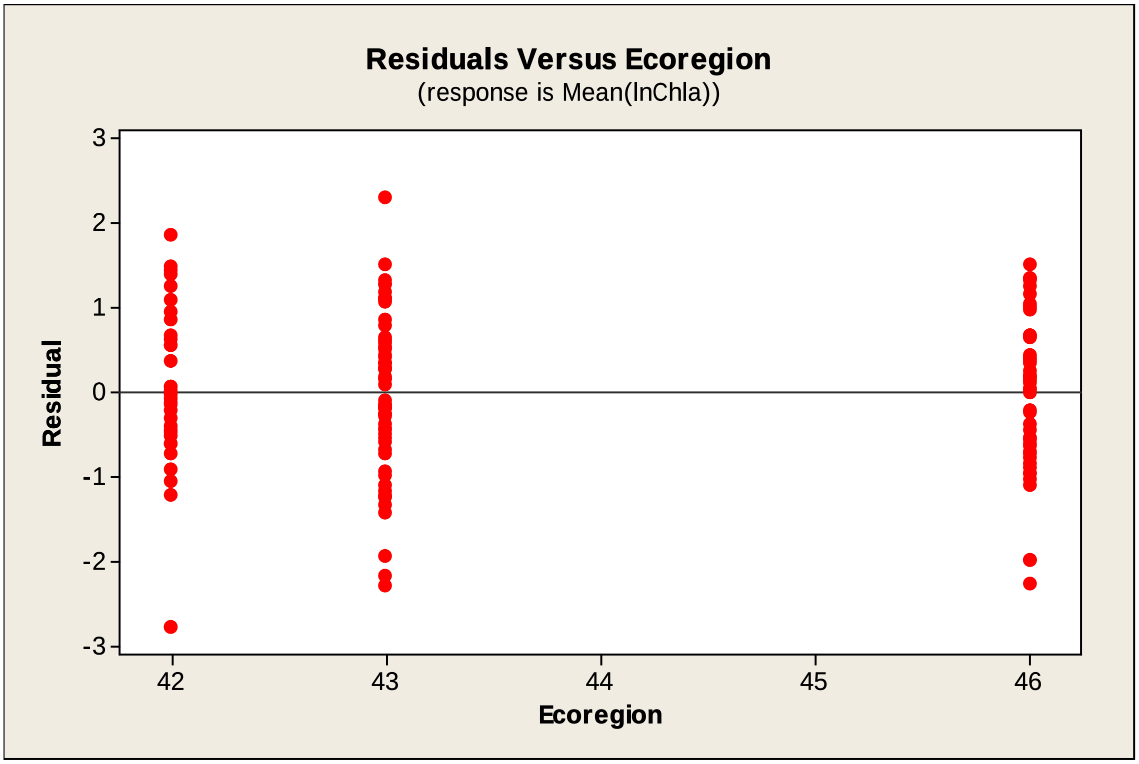
Tier	N	Mean	Grouping
1	103	2.6	A
3	13	2.5	A
2	33	2.4	A
4	9	1.5	A

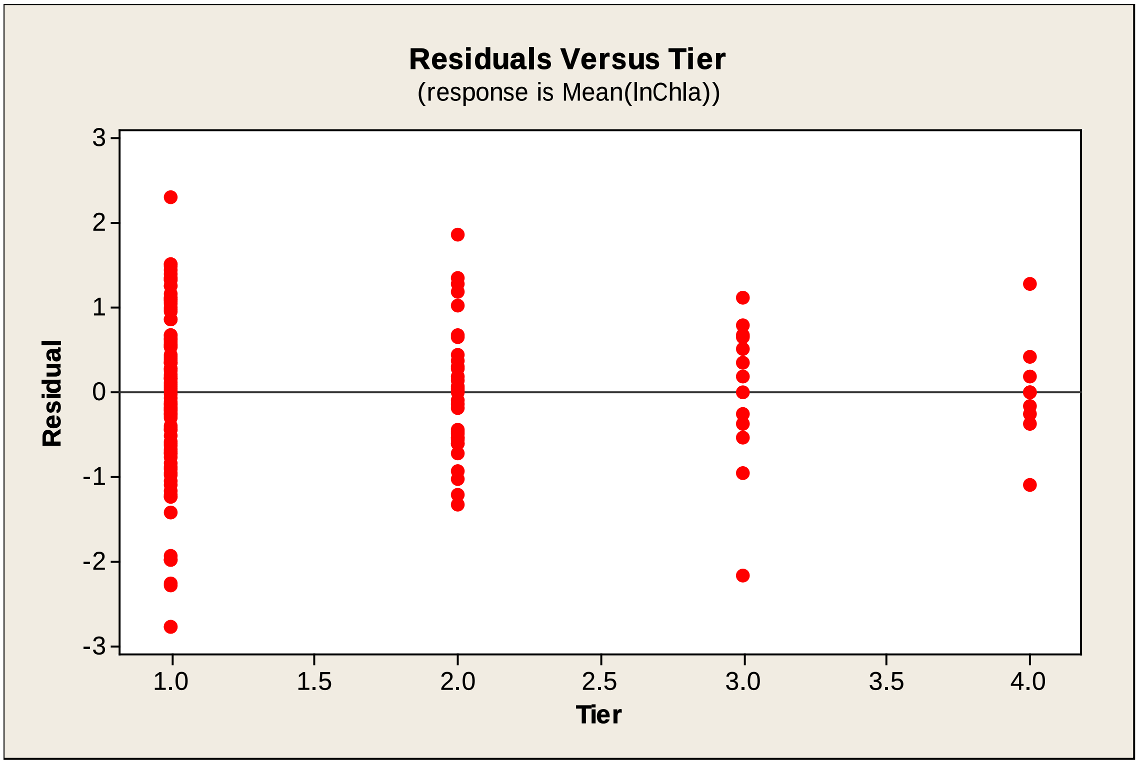
Means that do not share a letter are significantly different.

Grouping Information Using Tukey Method and 95.0% Confidence

Ecoregion	Tier	N	Mean	Grouping
42	1	28	3.2	A
42	3	1	3.0	A B
46	4	1	2.7	A B
46	3	7	2.7	A B
46	2	14	2.5	A B
42	2	7	2.4	A B
46	1	33	2.4	A B
43	1	42	2.3	B
43	2	12	2.2	A B
43	3	5	1.9	A B
43	4	7	1.0	B
42	4	1	0.6	A B

Means that do not share a letter are significantly different.


Residuals Versus Ecoregion
(response is Mean(lnChla))
Residual
3
2
1
0
-1
-2
-3
42
43
44
45
46
Ecoregion


Residuals Versus Tier
(response is Mean(lnChla))
Residual
3
2
1
0
-1
-2
-3
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Tier

Section VI: Model Development & Calibration

Section VI: Model Development & Calibration

MEMO

(External Correspondence)

To: Tina Laidlaw

Date: April 1, 2011

Cc: File 4965-002
Dennis McIntyre, GLEC

From: Stephanie Johnson, Ph.D., P.E.

Through: Mark R. Deutschman, Ph.D., P.E.

Subject: Model Development and Calibration Associated with Task 4 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8

This memorandum summarizes assumptions and methodologies used to develop two regional loading and eutrophication models as a deliverable under EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8 (i.e., the Nutrient Criteria Project). This work builds upon efforts that have been underway since 2009 and summarized in a series of memoranda dated October 8, 2009, March 5, 2010, September 20, 2010 and September 27, 2010. The regional models developed under these efforts and their application are meant to guide the Nutrient Criteria Project Team (consisting of members from the EPA and each impacted Plains State – North Dakota, South Dakota, Wyoming and Montana) as they make their policy decisions about setting nutrient criteria in the Region. Results of the application of the models described herein are discussed in a follow-up memo addressing efforts under Task 5 of the Nutrient Criteria Project.

BACKGROUND

The goal of this portion of the project is to develop and calibrate models that reflect the nutrient loading to and eutrophication response within the reservoirs of the study area. As discussed in the conclusions of the September 27, 2010 memorandum, two loading/ eutrophication models are to be created; i.e., one applicable to reservoirs in EPA Level 3 Ecoregion 46 and a second applicable to reservoirs in EPA Level 3 Ecoregions 42/43. In both cases, the models are applicable to all but the largest of the reservoirs (described in other portions of the project as classification tier 4 reservoirs) since it was assumed that site specific criteria would likely be developed for those waters. Once developed and calibrated, the models will be used to simulate management scenarios in the reservoirs' watersheds and inform the eventual development of nutrient criteria in the area. Therefore, the two models described in this memorandum are created to address nutrient loadings to and eutrophication responses in 934 reservoirs, 157 of which have water quality data available for them. The reservoirs are spatially distributed across the study area as shown in **Figure 1** and **Table 1**.

Figure 1: Modeled Reservoirs in the Study Area

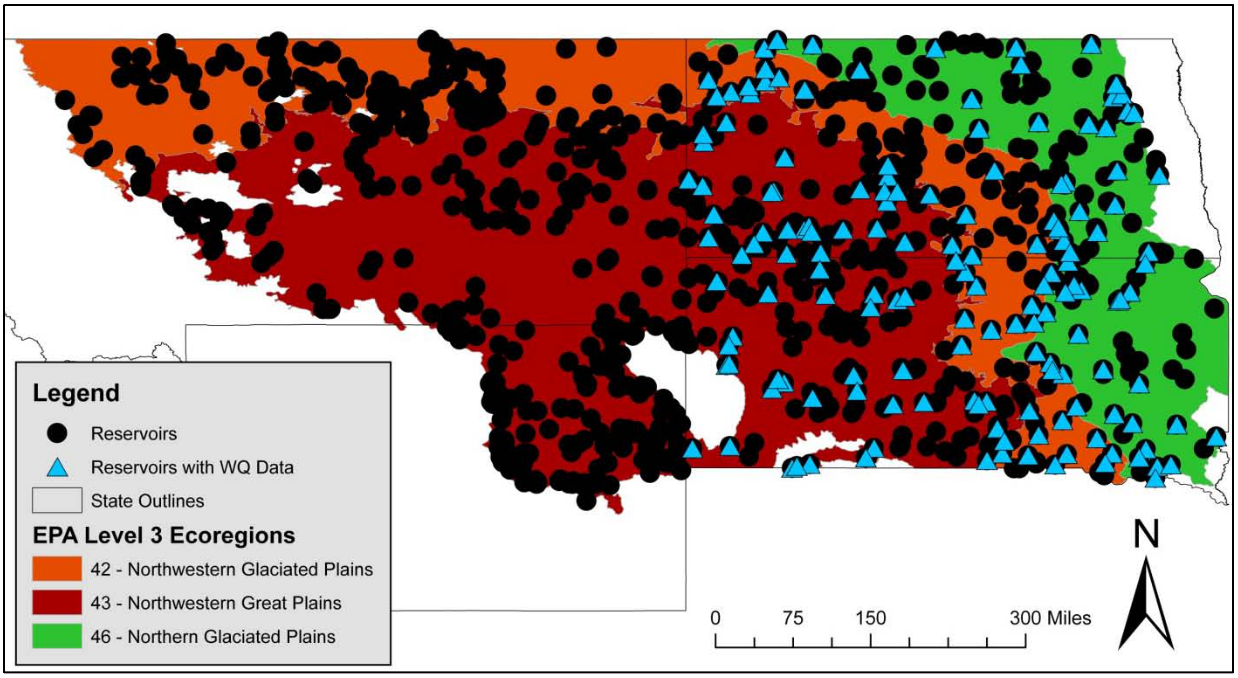


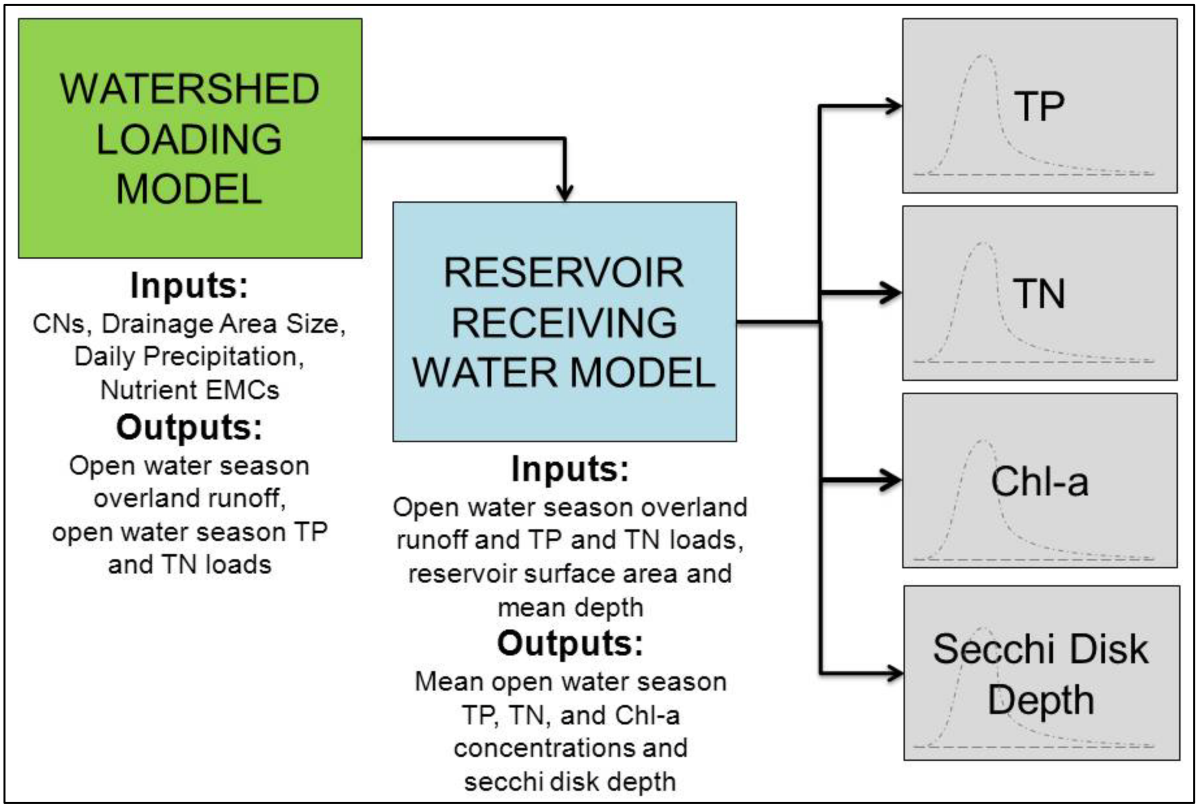
Table 1: Distribution of Reservoirs by State

	North Dakota	South Dakota	Montana	Wyoming
# of Reservoirs	274	218	293	149
# of Reservoirs w/ WQ data	75	82	0	0

METHODOLOGY

The nutrient loading /eutrophication modeling framework is shown in **Figure 2**, where a simple watershed nutrient loading model is used to estimate daily overland surface water runoff volumes and the corresponding total phosphorus (TP) and total nitrogen (TN) growing season loads using a daily time step. The daily runoff volumes and TP and TN loads are summed to compute open water season (March 1 – November 30) values, which then become inputs into a completely mixed steady-state (receiving water) water quality model. The receiving water model used is a variation of the U.S. Army Corps of Engineer's BATHTUB eutrophication model (Walker, 1996); i.e., CNET. The CNET eutrophication model is basically a spreadsheet version of BATHTUB, which uses empirical equations to estimate eutrophication responses in lakes and reservoirs. In this work, CNET uses the nutrient loading and surface water runoff values computed in the watershed loading model to estimate the open water season in-reservoir stressor concentrations of TP, TN, and chlorophyll-a (Chl-a) concentrations, and secchi disk depth (secchi depth) values (i.e., response), based on a series of (user-chosen) equations.

Figure 2: Stochastic Modeling Framework



The CNET (i.e., BATHTUB) model can be developed for any time period provided the assumption of steady-state conditions are valid (Walker, 1996). The BATHTUB model guidance suggests that the turnover ratio of the model (defined as the length of the modeled time period divided by the mass residence time of the waterbody) be greater than 2. In this case, the average TP mass residence times of reservoirs in the study area are approximately 2-3 months. In addition, 98% of Chl-a, 93% of TN, 93% of TP, and 98% of secchi depth data were collected during the open water season. Using this information, the Project Team decided to develop the regional reservoir eutrophication models to simulate conditions during the open water season. Doing so results in a turnover ratio of approximately 3.7, satisfying the guidance set forth in the BATHTUB User's Manual and allowing us to use the water quality analyses performed earlier in this project for model calibration.

The models created for this work do not represent nutrient loadings to or eutrophication responses within one specific study area reservoir, but rather in the population of reservoirs (in general) across each of the modeled regions (i.e., EPA Level 3 Ecoregion 46 and Ecoregions 42/43). For example, the model representing the population of reservoirs in Ecoregion 46 was developed and calibrated to provide insight on how reservoirs across this ecoregion can be expected to respond to nutrient loadings, based on historically observed runoff and in-reservoir water quality data, as described below. Actual (i.e., raw) observations of precipitation, drainage area, land use/cover, and reservoir characteristics were fed into the develop models as summarized in **Table 2**. Outputs of the models were then compared to observed mean annual unit runoff values for hydrology calibration. Annual average water quality was then computed for each reservoir from the empirical data (discussed in the September 20, 2010 memo). Outputs of the modeled in-reservoirs concentrations and associated responses were then compared to these reservoir annual averages across each region.

Table 2: Summary of Model Input and Observed Calibration Data

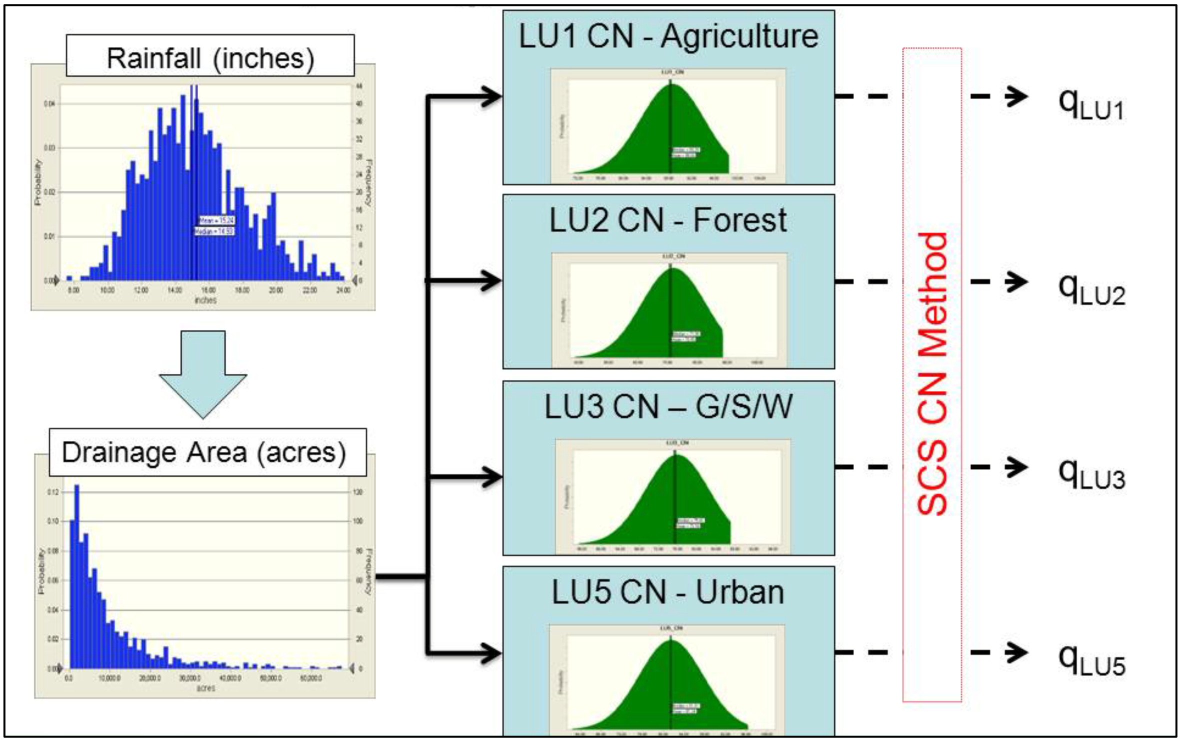
	Purpose	Data Format/Manipulation	Description/Notes
Precipitation	Model Input	Raw daily records	Daily precipitation distributions were developed by (non-winter) season
Drainage Area	Model Input	Raw data	Distribution of areas as recorded in the Reservoir Master Database
Mean Depth	Model Input	Raw data and/or estimated values (as a function surface area)	Distribution of areas as recorded in the Reservoir Master Database
Surface Area	Model Input	Raw data	Distribution of areas as recorded in the Reservoir Master Database
Unit Runoff	Loading Model Calibration	Mean Annual (computed by USGS)	Weighted-area average by modeling region
In-Reservoir Water Quality	Receiving Water Model Calibration	Reservoir annual average	Distribution of reservoir annual means by region

To allow the models to appropriately account for the wide variation in reservoir geometries, amount of runoff, and TP and TN watershed loadings and also to replicate the eutrophication response of the reservoir population, the models use a Monte Carlo approach. The Monte Carlo approach relies on repeated random sampling from a statistical distribution for specific input parameters to compute results, in this case the open-water average TP, TN, and Chl-a concentrations, and the secchi depth. For this work, distributions of the various model inputs (i.e., precipitation, land use curve numbers (CNs) and TP and TN estimated mean concentrations (EMCs), reservoir drainage area sizes, and reservoir surface areas and mean depths) were developed for each of the modeled regions as described later in the memorandum. During each model simulation, the statistical distributions are randomly sampled and the inputs used to drive the eutrophication model equations. This results in an estimated in-reservoir eutrophication response for the stressor (TP and TN) and response (Chl-a and secchi depth) variables based on that particular combination of inputs. Repeatedly performing this random sampling approach and using the model to compute the results, creates a series of eutrophication stressor and response values that reflect the conditions that have historically been observed across each of the modeled regions. This series of response values creates a distribution of estimated in-reservoir Chl-a concentrations and secchi depth values, as shown in **Figure 2**. An example of the computational process is as follows: during each model simulation the model input variable distributions are randomly sampled. During the first simulation, for example, the random sampling may select a precipitation value from a particularly dry year; a reservoir surface area, contributing drainage area size, and reservoir mean depth associated with a large reservoir; CN values associated with a high infiltration land use/land cover within the drainage area; and nutrient estimated mean concentration (EMC) values associated with low nutrient loadings. Entering these values into the modeling equations may result in a low Chl-a concentration and a high secchi depth value. The second simulation, however, may randomly sample values associated with a wet year, a small reservoir, high infiltration, and high nutrient loadings. Therefore, this simulation may result in a high Chl-a concentration and a low secchi depth value. Performing a large number of simulations creates a wide variety of loading and reservoir geometry scenarios, enabling the models to represent the wide variation of conditions and resources that are observed within each of the modeled regions. It also allows the models to account for the uncertainty involved in the model inputs. The result of this modeling approach is then a distribution of expected average open water season in-reservoir water quality in each modeled region (as shown in **Figure 2**).

<u>Watershed Loading Model</u>

The watershed loading model developed for this Task uses the Soil Conservation Service (SCS) CN method to compute surface runoff on a daily basis. The daily runoff value is a function of the population of reservoir drainage areas, period of record daily precipitation, and the curve numbers (CNs) of four general land use types within the population of reservoir drainage area, as shown in **Figure 3**. An estimated daily precipitation value is sampled from the seasonal distributions of precipitation for each modeled region (developed as described in the Modeling Inputs section). A drainage area is then sampled from a representative distribution of these values for the population of reservoirs within the ecoregion and divided into five general land use types (LU1 = agricultural, LU2 = forest, LU3 = grassland/ shrubland/wetlands, LU4 = water, and LU5 = urban) based on a set (average) percent cover for the region (also described in the Modeling Inputs section). Daily runoff from each of the land uses is computed by sampling a CN from each of the four land-based categories (i.e., not water) and combining it with precipitation through the SCS method. The total daily runoff from the drainage area is then computed by combining the unit runoff for each land use type with the area of that land use and summing them up.

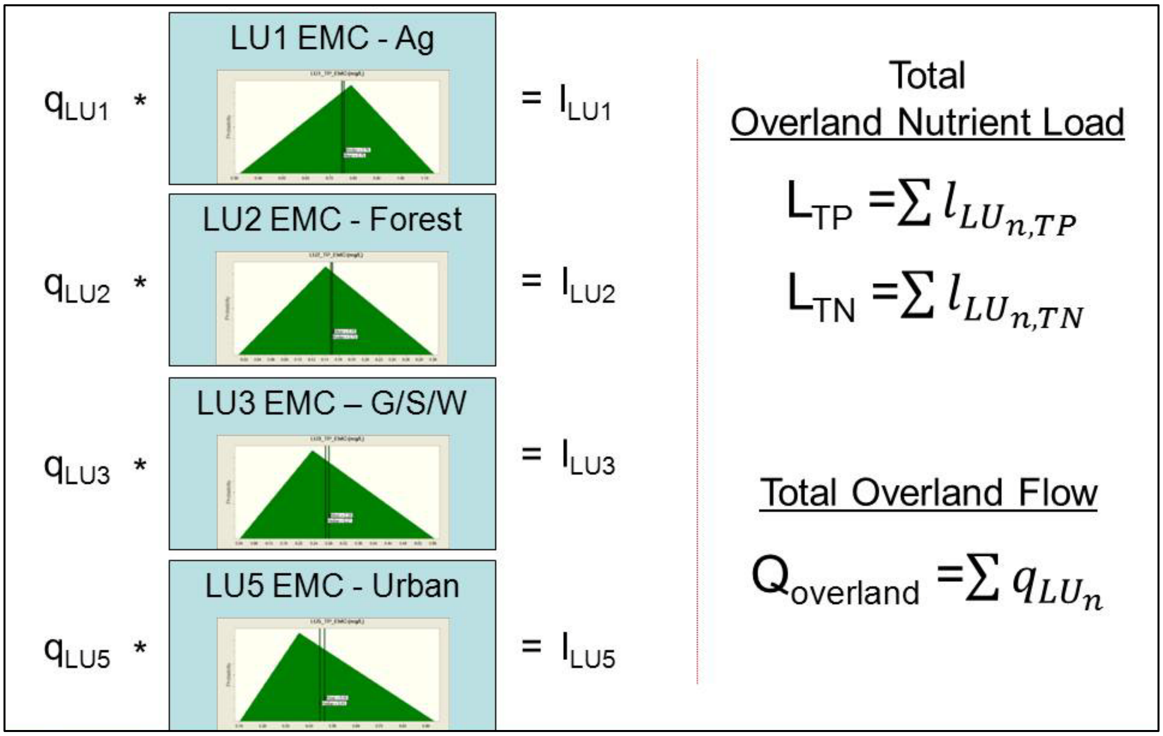
Figure 3: Structure of Watershed Loading Model – Daily Runoff Calculation



G/S/W = grassland/ shrubland/wetlands; q = daily runoff

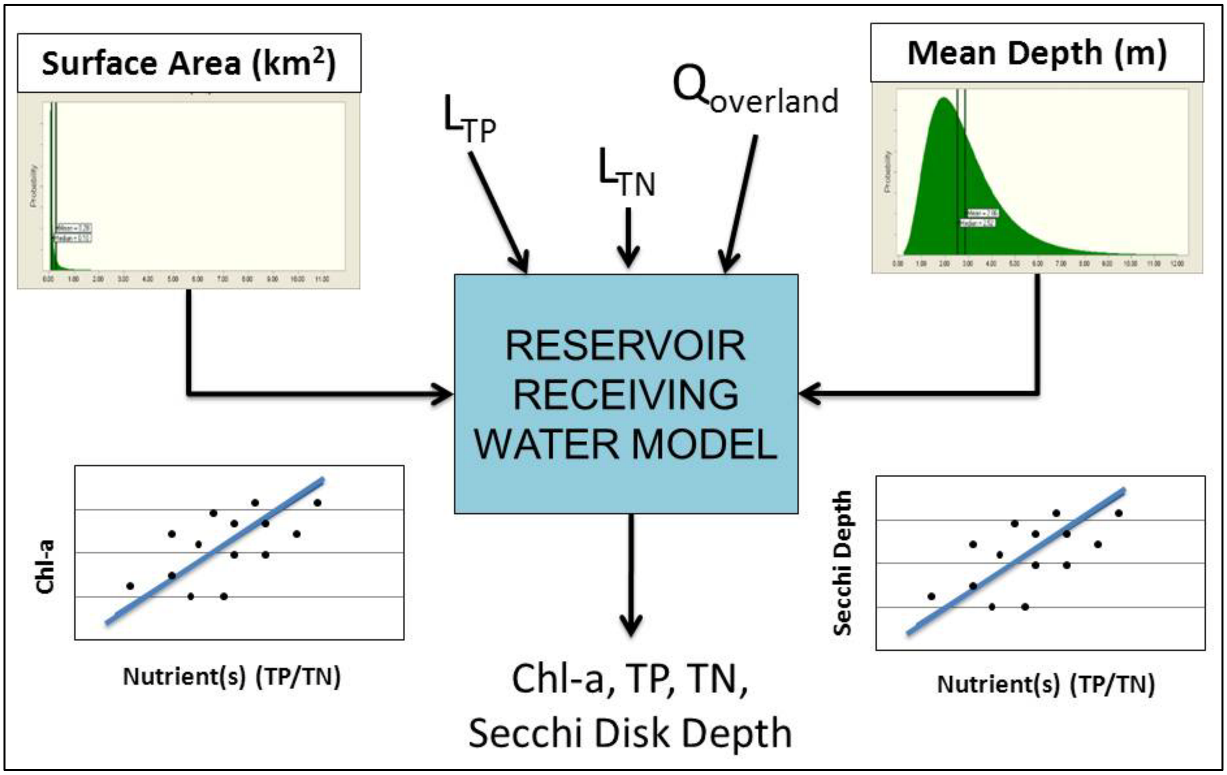
To compute daily pollutant loadings from each land use category, the estimated daily surface water runoff volumes are multiplied by the associated land use TP and TN EMCs (sampled from their input distributions, described in the Modeling Inputs section), as shown in **Figure 4**. The daily land use category nutrient loading values (e.g., daily TP loading from agriculture and forest) are then summed to compute a daily total nutrient loading from each overall drainage area. Daily values (from March 1st to November 30th) are then summed to compute the open water season surface water inflow and pollutant load to the reservoirs. These values are used as inputs to the CNET receiving water model as shown in **Figure 5**.

Figure 4: Structure of Watershed Loading Model – Pollutant Loading and Open Water Season Runoff and Loading Calculations



l = daily pollutant load; L = open water season load; Q = open water season runoff

Figure 5: Structure of Reservoir Receiving Water Model



The CNET receiving water model combines the outputs of the watershed loading model with information on reservoir morphometry through a series of empirical equations. Results of these equations provide estimates of mean open water season TP, TN, and Chl-a concentration, and secchi depth. Additionally, the model estimates the frequency with which nuisance algal blooms will occur within any given year based on the simulated mean Chl-a and a defined intra-annual coefficient of variation (COV) as described in Walker, 1984. Further information on the CNET modeling equations follows.

<u>Receiving Water Model - Empirical Water Quality Equations Used</u>

The CNET (i.e., BATHTUB) receiving water model has a number of different empirical equations for use in simulating nutrient concentrations and the associated eutrophication response of a given waterbody, based on nutrient loading and waterbody (i.e., reservoir) characteristics. Sedimentation models estimate the concentration of nutrients (i.e., TP and TN) that will be present in the water column and available for eutrophication processes as a function of the total load of nutrients entering the water body over the modeling period. Eutrophication response models estimate the amount of Chl-a and the secchi depth based on a combination of nutrient levels, light availability, turbidity, and flushing rate (the independent variables used in the estimates depend on the equation selected). Finally, statistical relationships are used to estimate the likelihood of experiencing nuisance algal blooms in any given year based on the simulated mean Chl-a and a defined intra-annual COV (as described in Walker, 1984).

Sedimentation Models

Two different models were used to simulate mean open water season in-reservoir phosphorus concentrations as a function of phosphorus loading and reservoir characteristics. Model results were then used to determine which model best represents the observed data. The first model used is the “Canfield & Bachmann (1981), Natural Lakes” equation. The second is the “Canfield & Bachmann (1981), Reservoirs + Lakes” equation.

Canfield & Bachmann (1981), Natural Lakes

$P = P_i / [1 + (C_P) * 0.162 (W_p / V)^{0.458} * T]$

Where: C_P = Phosphorus model calibration factor
P = Total phosphorus concentration (mg/m^3)
P_i = Inflow total phosphorus concentration (mg/m^3)
W_p = Total phosphorus loading (kg/yr)
V = Total volume (hm^3)
T = Hydraulic residence time (yrs)

Canfield & Bachmann (1981), Reservoirs + Lakes

$P = P_i / [1 + (CP) * 0.129 (W_p / V)^{0.549} * T]$

One equation, the “Bachmann (1980), Volumetric Load” equation was used to simulate nitrogen sedimentation.

$N = N_i / [1 + (C_n) * 0.159 (W_n / V)^{0.59} * T]$

Where: C_n = Nitrogen model calibration factor
N = Total nitrogen concentration (mg/m^3)
N_i = Inflow total nitrogen concentration (mg/m^3)
W_n = Total nitrogen loading (kg/yr)

Eutrophication Response Models

Mean open water season in-reservoir Chl-a concentrations were simulated using two different empirical equations. The first equation estimates Chl-a concentration as a function of only the TP concentration. The second equation estimates the Chl-a concentration as a function of a combined nutrient, which is a combination of both TP and TN mean open water season concentrations, as shown below.

Chl-a vs. TP

$B = [C_b] * 0.28 * P$

Where: C_b = Chl-a model calibration factor
B = Chl-a concentration (mg/m^3)

Chl-a vs. combined nutrient

$X_{pn} = [P^{-2} + ((N-150)/12)^{-2}]^{-0.5}$

$B_x = X_{pn}^{1.33} / 4.31$

$G = Z_{mix} (0.14 + 0.0039 * F_s)$

$B = C_b * B_x / [(1 + 0.025 * B_x * G) * (1 + G * a)]$

Where: X_{pn} = Composite nutrient concentration (mg/m^3)
B_x = Nutrient-potential Chl-a concentration (mg/m^3)
G = Kinetic Factor
Z_{mix} = Mean depth of mixed layer (m)
F_s = Summer flushing rate (yr^{-1})
a = Nonalgal turbidity (m^{-1}) = 1/S – 0.025 * B
C_b = Chl-a model calibration factor

Water clarity, expressed as the secchi depth, was also modeled using two different empirical equations. The first equation expresses secchi depth as a function of the mean open water season Chl-a concentration and non-algal turbidity; the second equation estimates depth as a function of the mean open water season combined nutrient concentration. Both equations are shown below.

Secchi Depth vs. Chl-a and Non-Algal Turbidity
$S = [C_S] / (a + 0.025 * B)$
Where: C_s = Secchi depth model calibration factor
S = Secchi depth (m)

Secchi Depth vs. Combined Nutrient
$S = [C_S] * 16.2 * X_{pn}^{-0.79}$

MODEL INPUTS

To run the regional reservoir models stochastically, using the Monte Carlo approach, statistical distributions for model input variables were required. These distributions were developed on a modeling region basis, so that the models created would properly reflect watershed loadings and reservoir responses unique to each area; i.e., the population of reservoirs within Ecoregions 46 and 42/43. **Table 1** summarizes the model input distributions and the data they contain. The following sections provide more details on the creation of the distributions and show their results.

Precipitation

To develop distributions representative of the daily rainfall in the reservoir drainage areas of the two modeling regions, data were collected from seventeen National Climatic Data Center (NCDC) sites (shown in **Figure 6**) - five stations in Ecoregion 46 and twelve stations in Ecoregions 42/43. Thirty years of precipitation data (1980 through 2009) were used. To compute values indicative of rainfall across each region, daily rainfall values at each of the monitoring locations were pooled on a seasonal basis ("spring" was defined as March 1 – May 30; "summer" was defined as June 1 – August 30; "fall" was defined as September 1 – November 30). The likelihood of it raining on any given day in each season was then computed. Lastly, distributions of non-zero observed daily precipitation in each modeling region by season were created. The resultant distributions are shown as box and whisker plots in **Figure 7** and **Figure 8**.

Figure 6: NCDC Precipitation Stations used to Compute Model Input Distributions

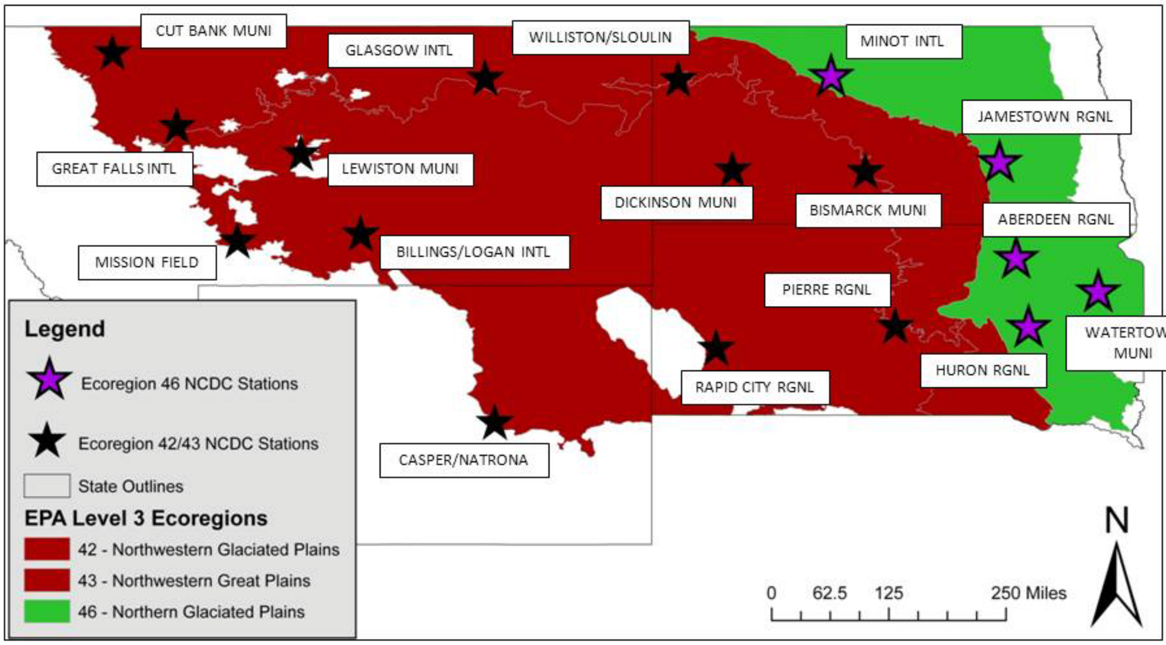


Figure 7: Non-Zero Daily Precipitation in Ecoregion 46

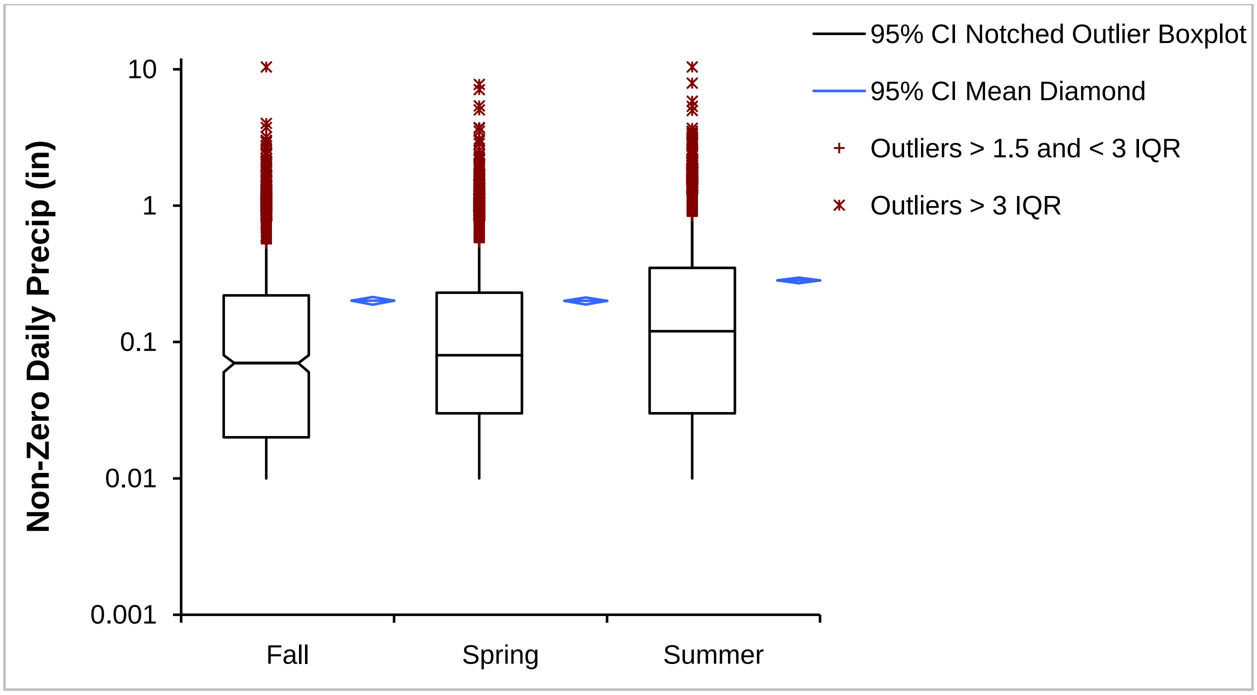
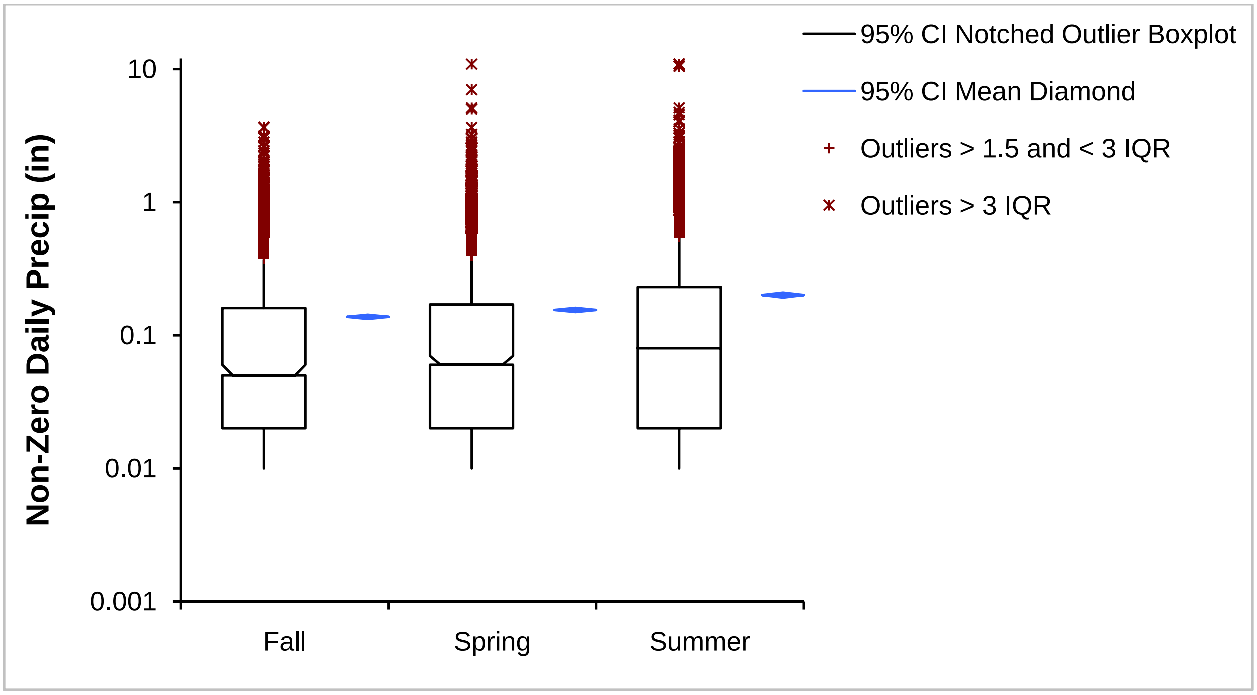


Figure 8: Non-Zero Daily Precipitation in Ecoregions 42/43



To compute runoff (and the associated nutrient load), the watershed loading model simulates rainfall everyday by first determining if it rained or not based on the likelihood of it raining on any given day in the corresponding season (e.g., to simulate precipitation in Ecoregion 46 on April 2, the model would use the fact that, based on 30-years of record at the 5 stations shown in **Figure 6**, spring days are 28% likely to receive rain). If the model determines that it did rain on that day, a rainfall depth is sampled from the corresponding season's (and region's) non-zero daily precipitation distribution; if the model determines that it did not rain, a zero is entered for that day.

Based on the daily data from the five stations in Ecoregion 46, the total open water season precipitation depth in the region ranged from 3.4 to 42.6 inches from 1980 through 2009. The average depth during this time was 18.3 inches. A similar analysis at the twelve stations in Ecoregions 42/43 showed that total open water season precipitation depths were widely variable, ranging from 1.6 to 36.1 inches. The average total open water season depth during this time was 12.7 inches.

<u>Reservoir Morphometry Data</u>

Statistical distributions for reservoir contributing drainage areas, surface areas, and mean depths were developed for each modeled region based on the data in the Master Reservoir Database developed under Task 2 of this project, using the data described in the March 5th, 2010 memo. The resultant distributions of the variables for each modeled region are shown as box and whisker plots in **Figure 9**, **Figure 10**, and **Figure 11**. (Note: the mean depth values shown in **Figure 11** are those that were estimated using the regression equations discussed in the March 5th, 2010 memo.) Statistical analyses show a correlation between reservoir surface area and contributing drainage area and reservoir surface area and mean depth. To properly reflect this statistical interdependence, the correlations between these distributions of reservoir morphometry were included in the modeling. They are also summarized in **Table 3**.

Figure 9: Reservoir Drainage Areas by Modeling Region

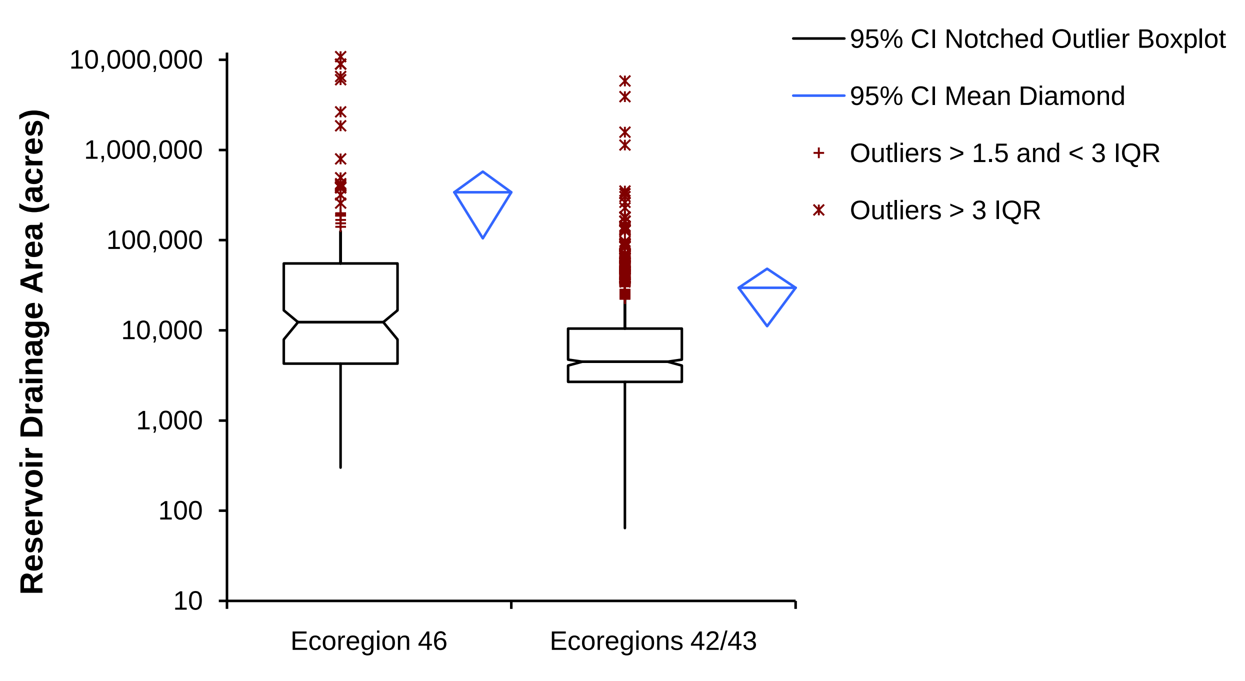


Figure 10: Reservoir Surface Areas by Modeling Region

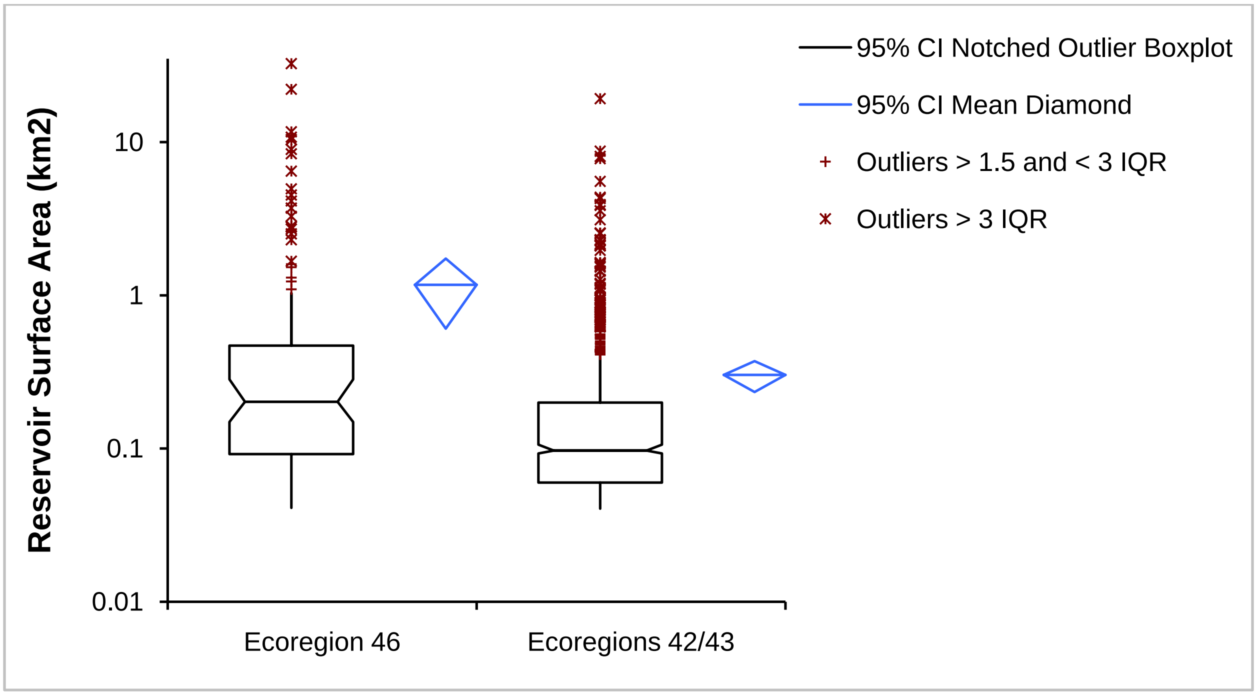


Figure 11: Reservoir (Estimated) Mean Depths by Modeling Region

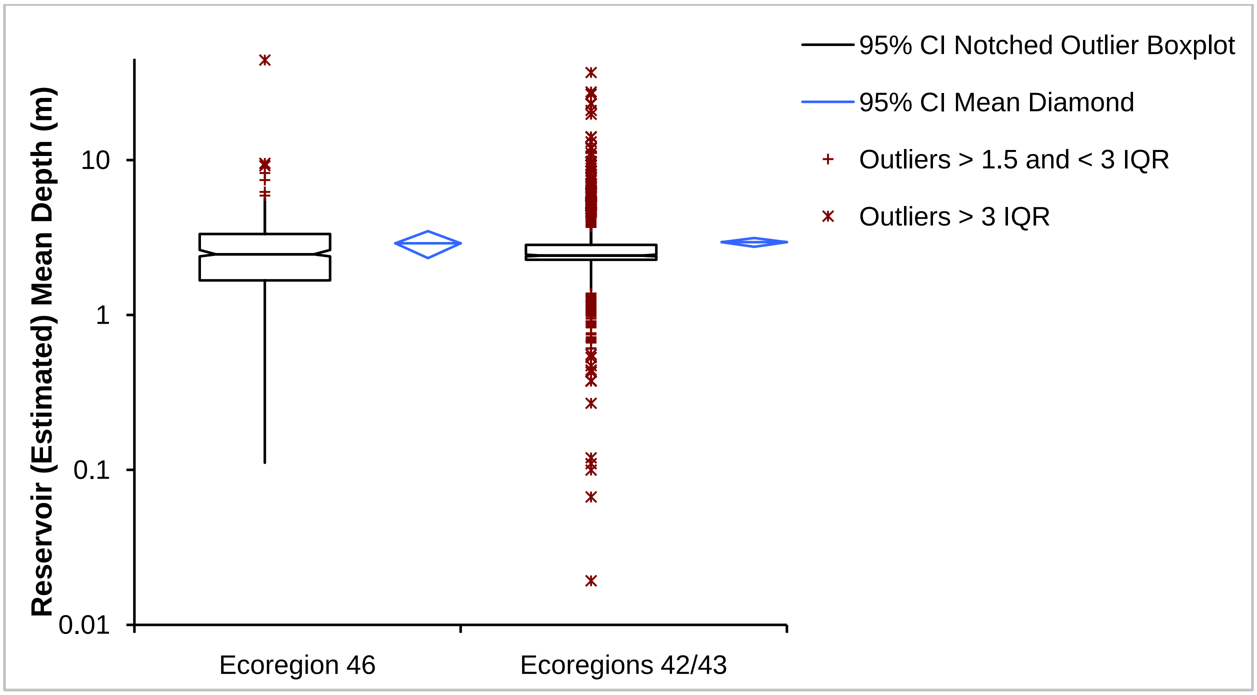


Table 3: Correlation Coefficients between Reservoir Morphometry Variables

	Ecoregion 46	Ecoregion 42/43
Surface Area vs. Mean Depth	0.24	0.29
Surface Area vs. Drainage Area	0.77	0.83

General Land Use Categories

The contributing drainage area data used for this work was mainly retrieved from the National Inventory of Dams, which simply provides a value for the drainage area into each waterbody but has no corresponding geospatial data showing the specific drainage area boundary. Since no other datasource was found to show the actual location/ delineation of the drainage areas contributing to each of the study area's reservoirs, the encompassing 8–digit Hydrologic Units (described by their Hydrologic Unit Codes, or HUCs) were used to estimate these areas. To do this, it was assumed that the land use/cover and soils data within the 8-digit HUC containing the centroid of each reservoir was reflective of the land use/cover and soil type (expressed as hydrologic soil group category) of that reservoir's actual drainage area. Using this assumption, the land use/cover and soils data in the 8-digit HUC could then be used to populate/ compute inputs for the watershed loading model. Given the homogeneity of land use/cover and soils within the study area, this assumption was deemed reasonable.

The 2001 National Land Cover Dataset (NLCD) was used to denote the land use/cover in the study area. For the purpose of analysis, NLCD land use categories were re-classified into five (more general) land use categories and summarized per 8-dight HUC. **Table 4** shows the five general land use categories used in the watershed loading modeling and the average percent of the population of reservoirs contributing drainage areas (represented by its encompassing 8-digit HUC) in each land use by modeling region.

Table 4: Average Percentages of Land Use in Reservoir Drainage Areas by Modeling Region

General Land Use Category	Ecoregion 46	Ecoregions 42/43
Agriculture, row crops (LU1)	71%	25%
Forest, woods (LU2)	2%	2%
Grasslands/shrubs/wetlands, brush (LU3)	20%	70%
Water (LU4)	2%	1%
Urban, impervious area (LU5)	5%	2%

Curve Numbers

Curve numbers (CNs) were computed for the study area by combining GIS data on land use/land cover (2001 NLCD) with soils (SSURGO) to find unique combinations of land use and hydrologic soil group. For each reservoir drainage area (again defined by the encompassing 8-digit HUC), an area-weighted CN was then estimated for each of the major land use types (LU1 – LU5). This analysis resulted in four weighted CNs per reservoir drainage area; LU4 – water – was assumed non-contributing (of pollutant loading) and, therefore, not modeled. Distributions of each general land use category's weighted average CNs where then developed. Results are shown, by modeling region, as box and whisker plots in **Figure 12** and **Figure 13**.

Figure 12: Computed CNs by Land Use Category for Ecoregion 46

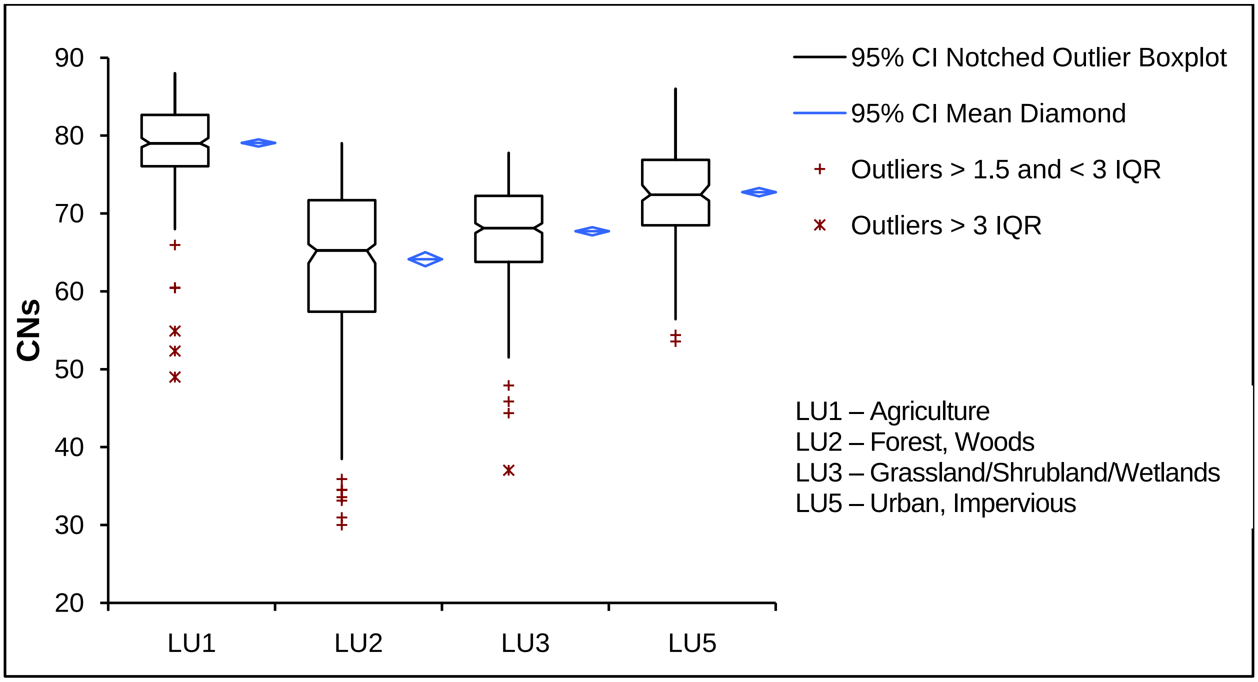
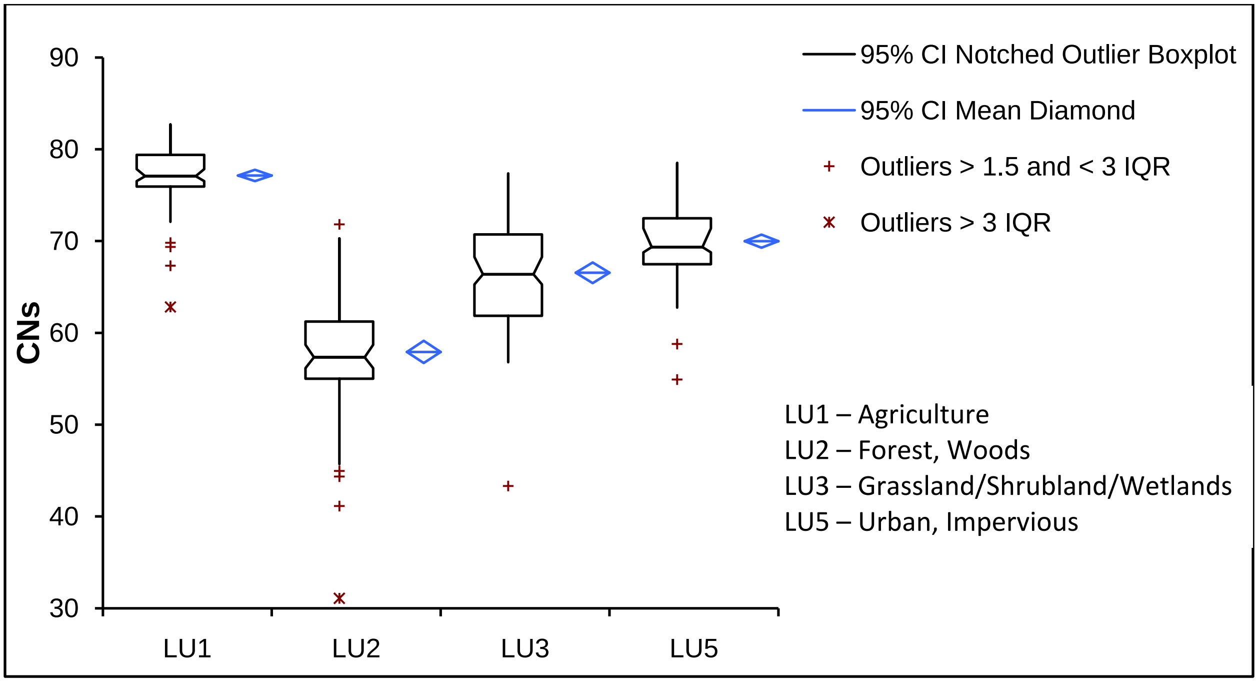


Figure 13: Computed CNs by Land Use Category for Ecoregions 42/43



<u>Nutrient Estimate Mean Concentration Values</u>

Distributions of nutrient EMC values for each general land use category were developed based on literature values and discussions with the Nutrient Criteria Project Team. Given the data available, it was determined that a triangular distribution would be used for the TP EMCs and a lognormal distribution for the TN EMCs. **Table 5** and **Table 6** summarize the nutrient EMC distributions by land use category.

Table 5: Total Phosphorus Estimated Mean Concentrations by Land Use Category

Land Use Category	Distribution Type	Median (mg/L)	Mean (mg/L)	Minimum (mg/L)	Maximum (mg/L)
LU1 – Ag	Triangular	0.76	0.75	0.32	1.14
LU2 – Forest	Triangular	0.15	0.15	0.01	0.30
LU3 – G/S/W	Triangular	0.27	0.28	0.04	0.56
LU5 – Urban	Triangular	0.44	0.46	0.10	0.93

Table 6: Total Nitrogen Estimated Mean Concentrations by Land Use Category

Land Use Category	Distribution Type	Median (mg/L)	Mean (mg/L)	Minimum (mg/L)	Maximum (mg/L)
LU1 – Ag	Lognormal	4.52	6.15	0.80	25.0
LU2 – Forest	Lognormal	0.64	1.92	0.04	25.8
LU3 – G/S/W	Lognormal	3.38	4.62	1.24	22.9
LU5 – Urban	Lognormal	1.72	2.36	0.20	90.1

MODELING ASSUMPTIONS

In addition to creating the model input distributions discussed above, a number of assumptions were made during model development. Some of those assumptions were described in the discussion about the various inputs developed; others were not. The majority of these assumptions were discussed with the Nutrient Criteria Project Team and feedback/approval was provided before work continued. The following summarizes the most notable assumptions made during model development:

- Assume that data from 12 NCDC stations in Ecoregions 42/43 and 5 NCDC stations in Ecoregion 46 are reflective of precipitation across the modeled regions
- Receiving water model steady-state time period for the open water season is defined as March 1 to November 30
- Average annual evaporation = 30 inches/year (assume all evaporation takes place during the modeled open water season – March 1 to November 30)
- Atmospheric TP loading = 30 kg/km^2-yr
- Atmospheric TN loading = 204 kg/km^2-yr
- Average Ecoregion 46 non-algal turbidity = 0.27 m^{-1}
- Average Ecoregions 42/43 non-algal turbidity = 0.02 m^{-1}
- Mixed depth = mean depth (i.e., reservoirs do not stratify)
- Assume no spatial variation in water quality data across reservoirs (i.e., completely mixed)
- Assume model input distributions are reflective of all reservoirs in the modeled regions

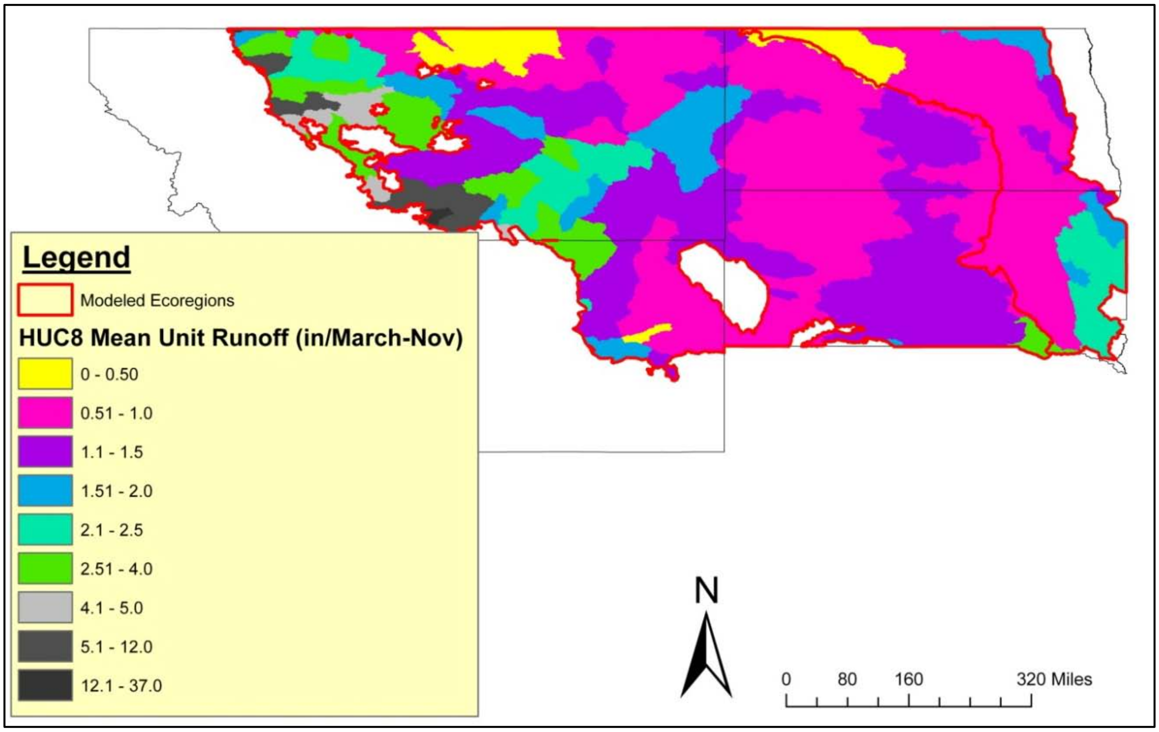
- Assume available water quality data is representative of all reservoirs in the modeled regions
- Assume empirical equations are representative of the eutrophication processes in the reservoirs of our study area
- Assume no internal loading (will be implicitly accounted for in calibration process)
- Assume eutrophication response dependent solely on TP and TN – don't consider Ortho-P and inorganic N in the modeling equations

HYDROLOGY CALIBRATION

The first step in calibrating each regional model was to adjust the watershed loading model hydrology component to match observed and simulated unit runoff values (inches per year per unit area). The watershed loading model pollutant loading component could not be calibrated, since no observed data was available. Observed unit runoff data were provided by the US Geological Survey (USGS), who computed mean unit runoff values for the nation's 8-digit HUCs under their WaterWatch program (http://waterwatch.usgs.gov/index.php/?m=romap3&w=download). These mean unit runoff values were computed from historical flow data from USGS stream gages, using data from 1901 to 2008. Details on the data calculations are discussed here: http://waterwatch.usgs.gov/wwhelps/romap3.html. Ideally, individual flow records from each of the major USGS gaging stations in the study area would have been obtained and unit runoff statistical distributions developed (similar to what was done with the precipitation data) for use in the hydrology calibration. Given the large amount of flow data available and the considerable effort involved in reducing these data, the pre-computed averages for the 8-digit HUC (i.e.,mean basin averages) were deemed sufficient for this study. By using the mean unit runoff values (by the 8-digit HUC),however, some of the variability in the runoff distributions (i.e., the extreme high and low values) has been removed. Therefore, it is expected that simulated runoff distributions should show more variation than the observed. The focus of the hydrology calibration was, therefore, to match the central tendency of the simulated and observed data, noting that the variation in the distributions is expected to differ.

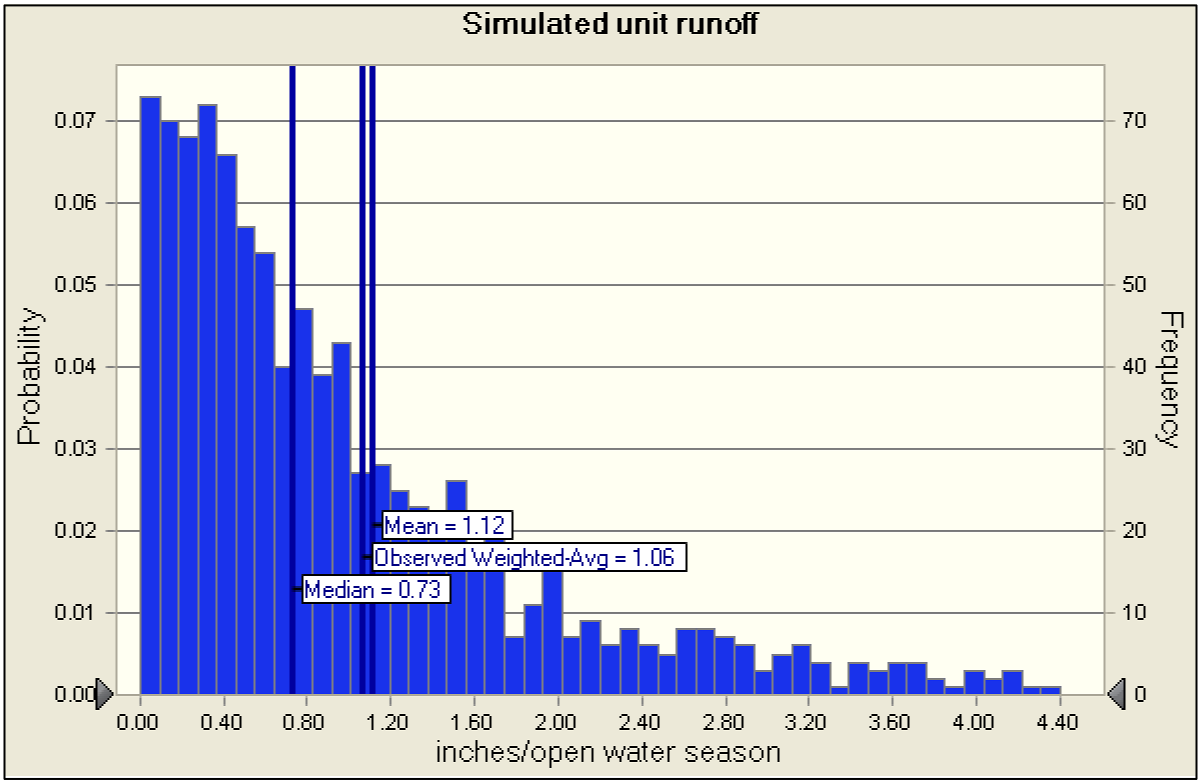
The models were calibrated for the open water season unit runoff. The mean monthly unit runoff values for March through November were summed to compute the mean open water season unit runoff for each 8-digit HUC in the study area. The mean open water season unit runoffs for the study regions were then computed by spatially weighting these averages. **Figure 14** shows the mean open water season unit runoff values for the study area. Based on these data, the weighted-average mean open water season unit runoff value for Ecoregion 46 is 1.06 inches.

Figure 14: Mean Open Water Season Unit Runoff



The hydrology in the model for Ecoregion 46 was calibrated by increasing or decreasing the CNs in the watershed loading model until the estimated means of the modeled and observed unit runoff distributions reasonably matched. All CNs (regardless of LU type) were adjusted by the same percentage during calibration. **Figure 15** shows the result of the Ecoregion 46 hydrology calibration, where simulated CNs were increased by 1% from the original input values, thereby creating an mean simulated open water season unit runoff value of 1.12 inches.

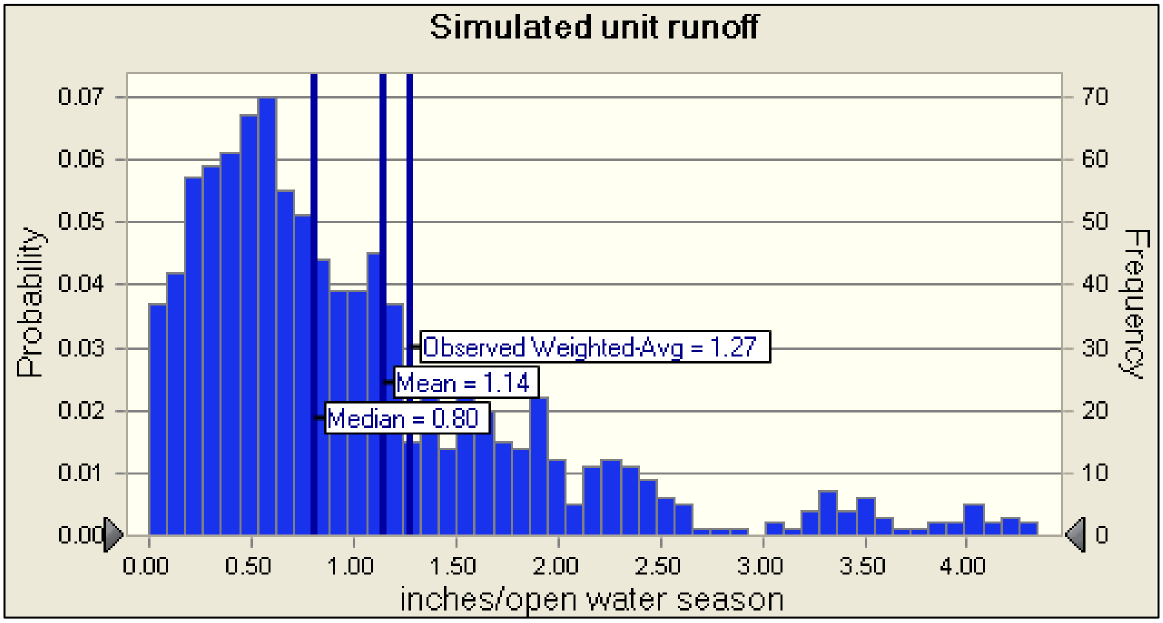
Figure 15: Results of Ecoregion 46 Hydrology Calibration



Using the data shown in **Figure 14**, the observed weighted-average mean open water season unit runoff value for Ecoregions 42/43 is 1.54 inches. However, as shown in **Figure 14**, a small number of HUCs on the western edge of the region have considerably larger mean unit runoff values than the remaining HUCs in the area due to the impact of mountain hydrology in this area (i.e., their unit runoff values are significantly impacted by mountain snow melt). If we disregard those HUCs with mean open water season unit runoff values over 4 inches, since these unit runoff values are not truly reflective of the Plains hydrology, the weighted average mean open water season unit runoff for Ecoregions 42/43 is 1.27 inches.

Trying to balance a realistic distribution of LU1 CNs with matching the central tendency in the unit runoffs, the CN values were increased by 12% from their original input values during the hydrology calibration. This adjustment results in an average simulated open water season unit runoff of 1.14 inches for Ecoregions 42/43, as shown in **Figure 16**.

Figure 16: Results of Ecoregions 42/43 Hydrology Calibration



RESERVOIR RECEIVING WATER MODEL – WATER QUALITY CALIBRATION

The second step in calibrating the regional reservoir models was to compare model results to observed in-reservoir water quality data. In this step, we calibrated the receiving water models to four different water quality variables, including concentrations of TN, TP, and Chl-a, and the associated secchi depth. Water quality data for the model calibration was provided by each state in the study area, as described in the March5th and April 2nd, 2010 Memoranda addressing the data gathering and water quality analysis phases of the project. The raw data were then used to compute the mean annual water quality in each reservoir, which was used to drive calibration. [As noted earlier, the vast majority of water quality data used in this analysis was collected during the open water season; therefore, the mean annual water quality values are assumed equal to mean open water season values for model calibration purposes.] **Figure 17** through **Figure 30** show box and whisker plots of the reservoir-yearly average water quality values computed in the two modeling regions.

The water quality models were calibrated by effectively adjusting the amount of sedimentation in the sedimentation equations and the slope of the stressor-response response relationships in the eutrophication response equations. These adjustments were done by adjusting the equations' calibration coefficients (C_p, C_n, C_b, and C_s). The overall goal in calibrating the water quality models was to adjust the coefficients until the simulated distributions of regional open water season values matched the distributions of the observed data as closely as possible. Emphasis was placed first on matching the central tendency (i.e., median) of the mean open water distributions. The secondary priority was to match the first and third quartiles of the statistical distributions, since these values could have implications in the eventual nutrient criteria development. The water quality model state variables were calibrated in the following order: TP, TN, Chl-a, and secchi depth. When calibrating the Chl-a and secchi depth equations, the recommended/ calibrated TP and TN sedimentation equations were utilized.

Simulated Water Quality – Ecoregion 46

Figure 17 through **Figure 23** show the results of the water quality model calibrations in Ecoregion 46. The individual values in the “Modeled” and “Observed” statistical distributions represent the mean open water season in-reservoir concentrations/values for the population of reservoirs in the region. The calibration coefficients required to attain model calibration are shown on each of the plots and summarized in **Table 7**.

Figure 17: Results of Ecoregion 46 Total Phosphorus Model Calibration (Canfield & Bachman (1981), Natural Lakes)

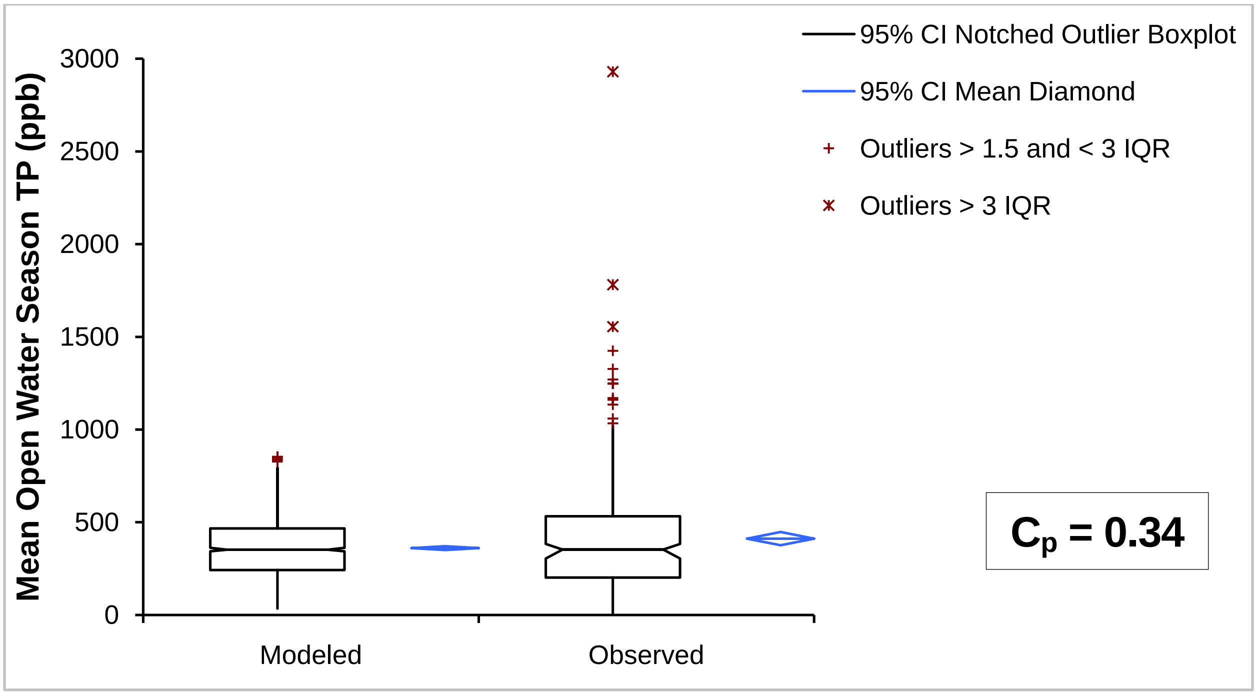


Figure 18: Results of Ecoregion 46 Total Phosphorus Model Calibration (Canfield & Bachman (1981), Lakes + Reservoirs)

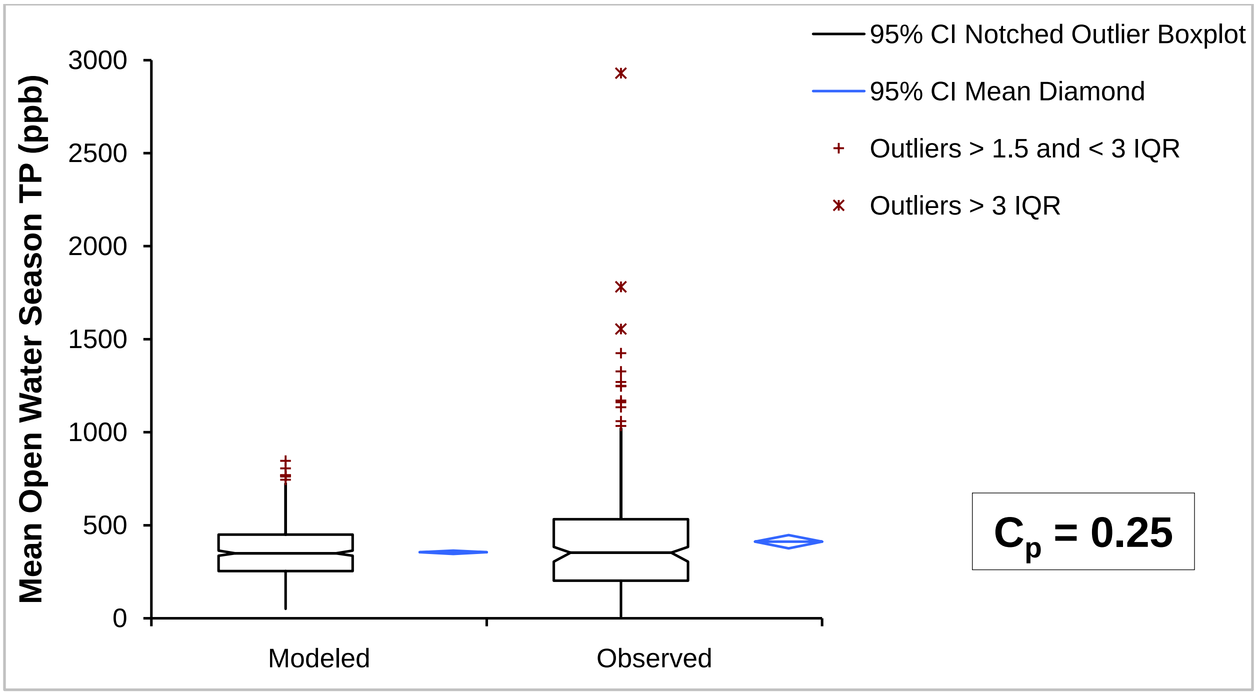


Figure 19: Results of Ecoregion 46 Total Nitrogen Model Calibration (Bachman (1980), Volumetric Load)

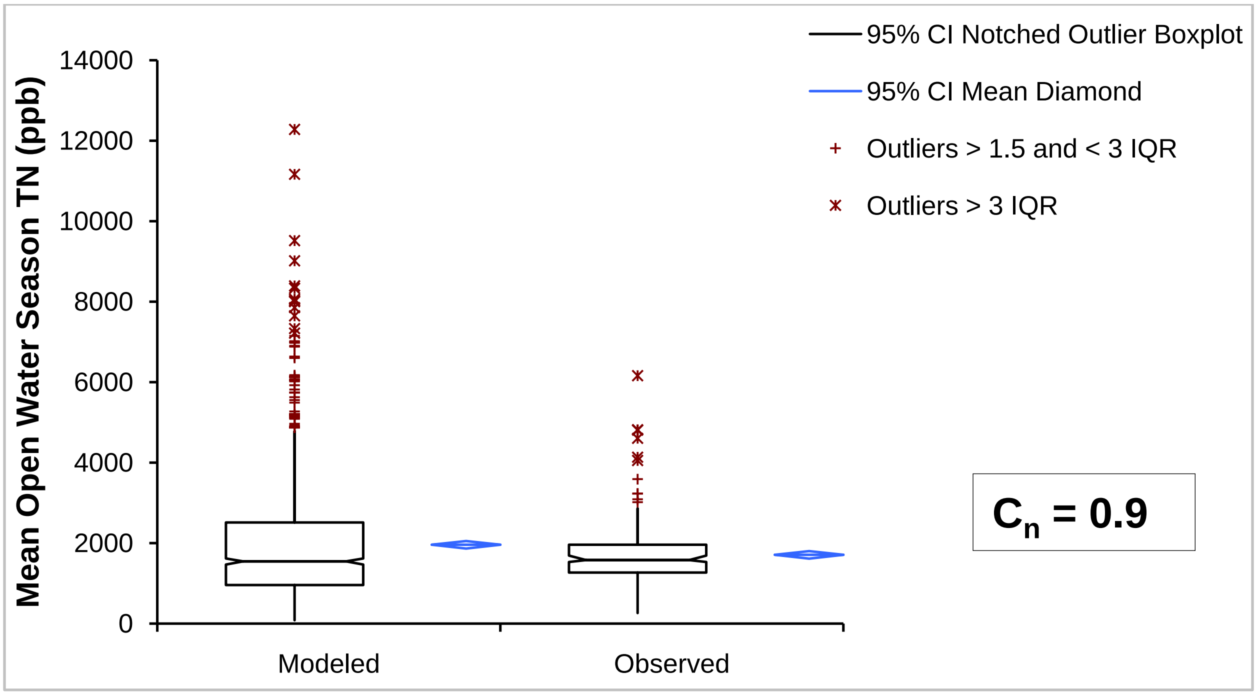


Figure 20: Results of Ecoregion 46 Chlorophyll-a Model Calibration (Chl-a vs. Combined Nutrient)

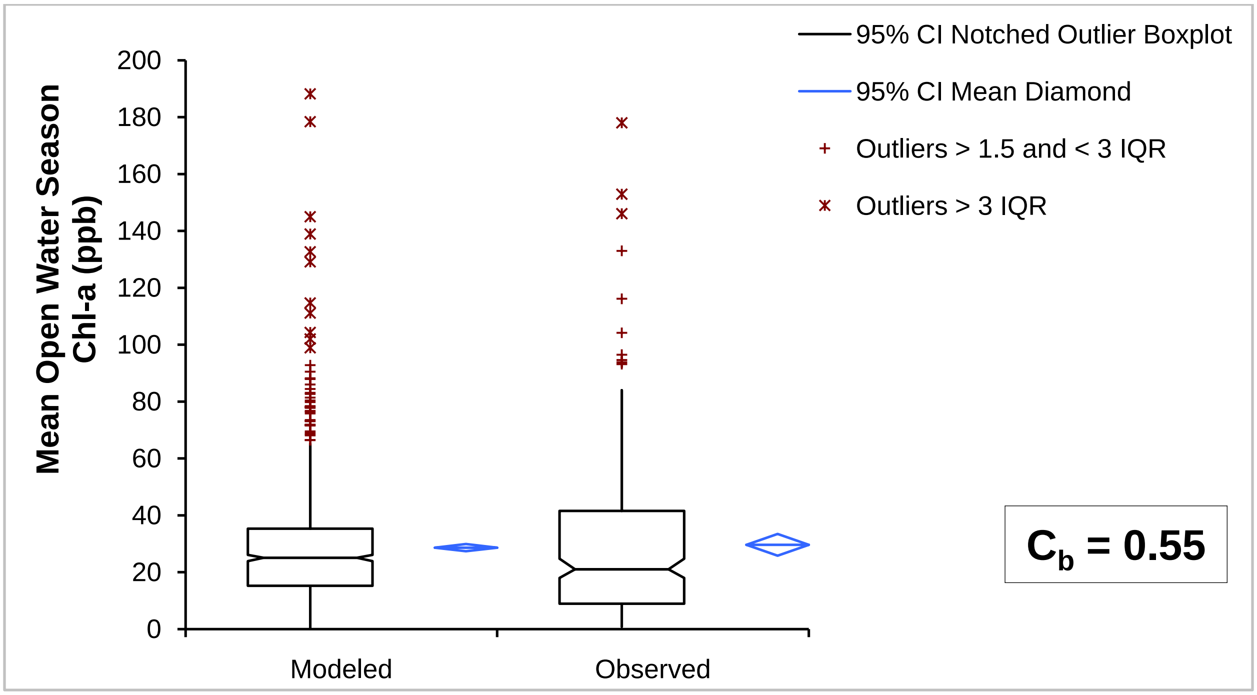


Figure 21: Results of Ecoregion 46 Chlorophyll-a Model Calibration (Chl-a vs. TP)

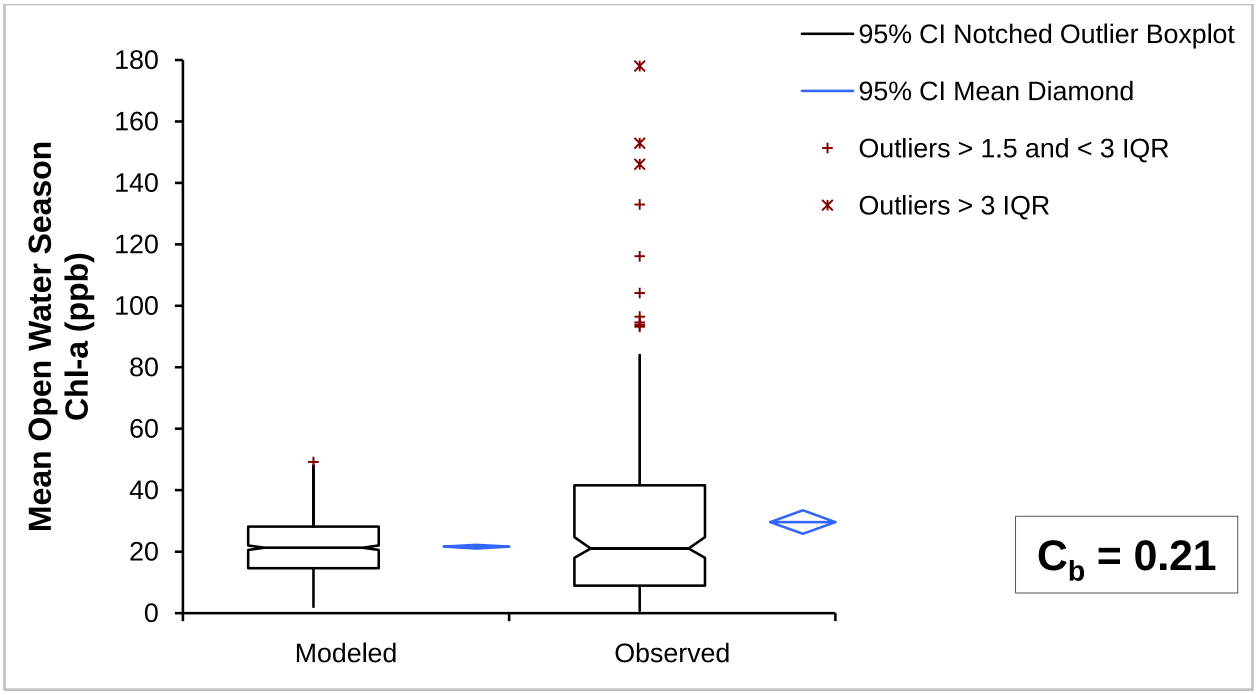


Figure 22: Results of Ecoregion 46 Secchi Depth Model Calibration (Secchi Depth vs. TP and Non-Algal Turbidity)

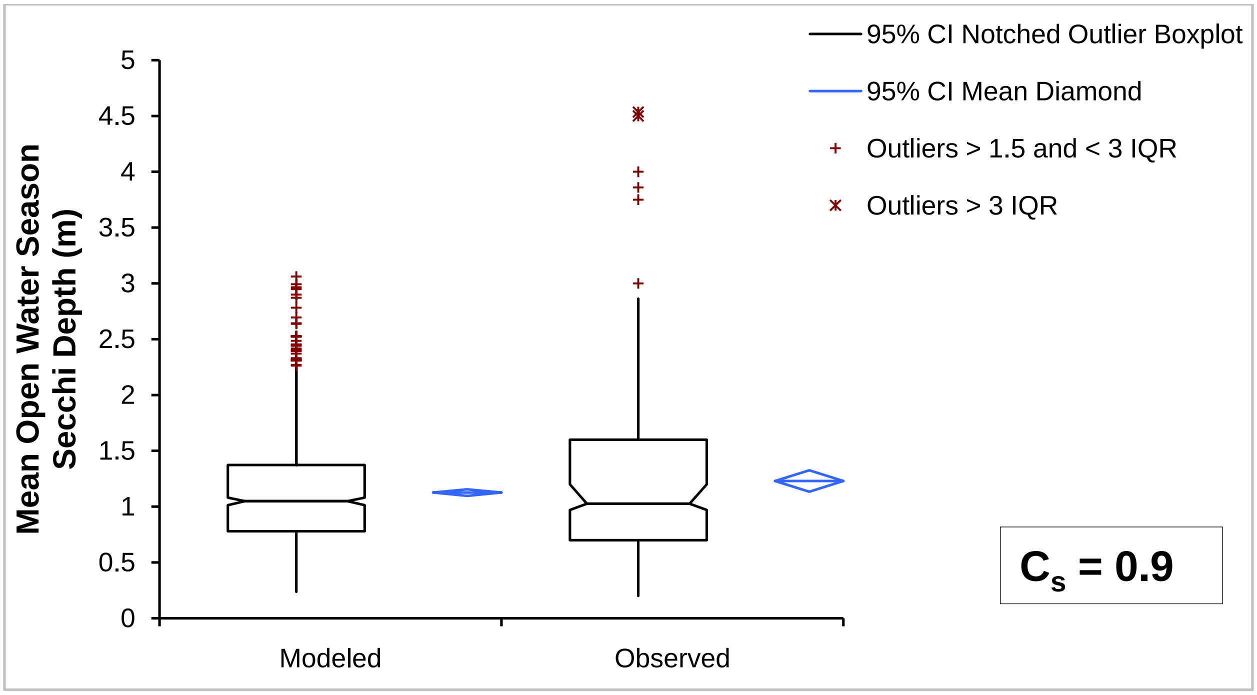


Table 7 summarizes the calibration coefficients used to calibrate the modeling results shown above. In general, the further that the calibration coefficient deviates from 1.0, the poorer the equation reflects the sedimentation or stressor-response relationship of the population of reservoirs. Results of modeling the Ecoregion 46 reservoirs show that the "Canfield & Bachmann (1981), Natural Lakes" model better represents the process of phosphorus sedimentation in these reservoirs than the "Canfield & Bachmann (1981), Reservoirs + Lakes" model. Chlorophyll-a concentrations were better represented by the equation that simulates them as a function of the combined nutrient (rather than the equation that computes them from TP alone). Ecoregion 46 secchi depths were best simulated by the model that estimates them as a function of Chl-a and non-algal turbidity. In fact, the calibration coefficient for the model considering secchi depth as a function of the combined nutrient was so large (greater than 2.5) that the equation was considered inappropriate for modeling secchi depth in Ecoregion 46 and its calibration results are not included in this memo.

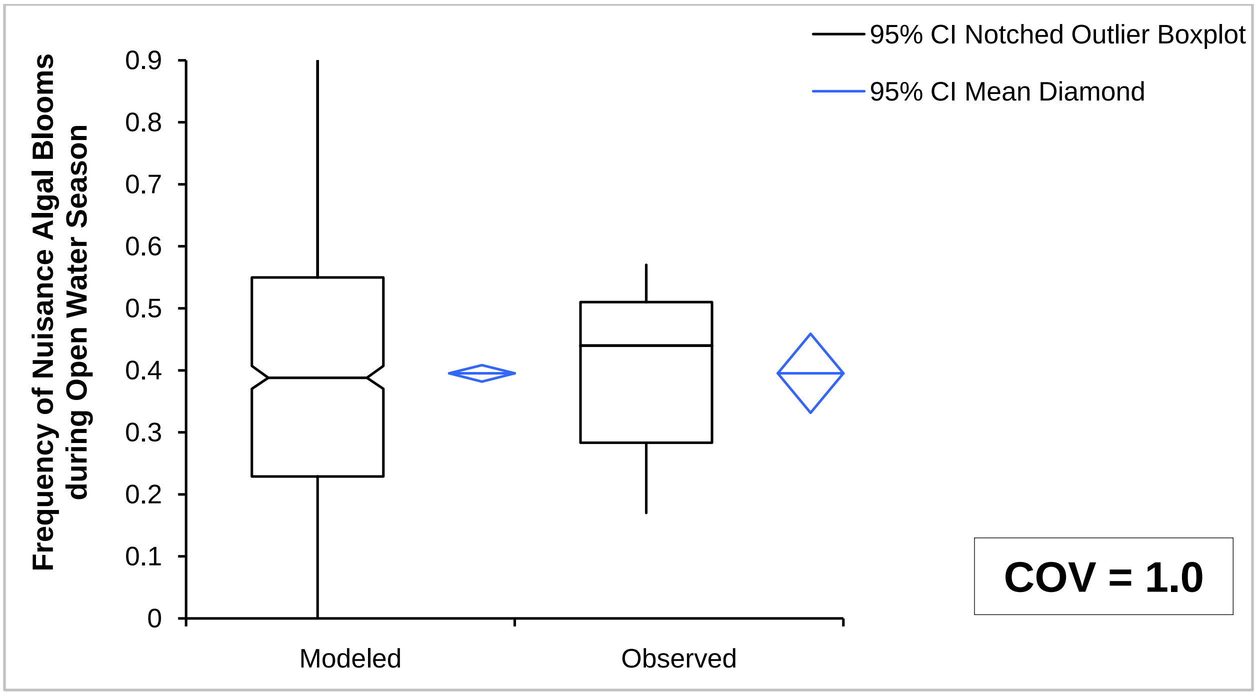
Table 7: Resultant Ecoregion 46 Water Quality Model Calibration Coefficients

Model	Calibration Coefficient	Value after Calibration
Total Phosphorus Sedimentation (Canfield & Bachman (1981), Natural Lakes)	C_p	0.34
Total Phosphorus Sedimentation (Canfield & Bachman (1981), Lakes + Reservoirs)	C_p	0.25
Total Nitrogen Sedimentation (Bachman (1980), Volumetric Load)	C_n	0.9
Chl-a vs. combined nutrient[2]	C_b	0.55
Chl-a vs. TP[2]	C_b	0.21
Secchi depth vs. combined nutrient[1,2,]	C_s	N/A
Secchi depth vs. TP and non-algal turbidity[2]	C_s	0.9

[1] The calibration coefficient had a value of > 2.5, which is considered outside of the acceptable range. Results of using this model for simulating secchi depth in Ecoregion 46 were, therefore, not included in this memo.
[2] The calibrated "Canfield & Bachman (1981), Natural Lakes" and "Bachman (1980), Volumetric Load" equations for TP and TN sedimentation were in use during these simulations.

The last step in calibrating the water quality model for Ecoregion 46 was to adjust the model input representing the intra-annual variation of individual reservoir Chl-a concentrations: the Chl-a distribution COV. As described by Walker (1984), a statistical relationship can be used to estimate the likelihood of experiencing nuisance algal blooms within any given year, based on the mean yearly (or, in this case, open water season) Chl-a concentration and the COV of the statistical distribution of individual Chl-a values within each year. Based on guidance from the Nutrient Criteria Project Team, for the purposes of this modeling, Chl-a concentrations greater than or equal to 20 parts per billion (ppb) were defined to be nuisance conditions. For each year of observed in-reservoir Chl-a concentrations, the number (and percent) of Chl-a values greater than or equal to 20 ppb were counted. A statistical distribution of these values was then created. The simulated likelihood of nuisance algal blooms was then calibrated to this observed value, by adjusting the COV. Calibrating the Ecoregion 46 model to best match the observed frequency of nuisance blooms resulted in a COV value of 1.0. **Figure 23** shows the result.

Figure 23: Results of Ecoregion 46 Frequency of Nuisance Algal Blooms (Chl-a > 20 ppb) Calibration



<u>Simulated Water Quality – Ecoregions 42/43</u>

Similar to Ecoregion 46, the Ecoregions 42/43 receiving water model was also calibrated to simulate water quality in the region's reservoirs. The individual values in the "Modeled" and "Observed" statistical distributions represent the mean open water season in-reservoir concentrations/values for the population of reservoirs in Ecoregion 42/43. **Figure 24** through **Figure 31** show the results of the water quality model calibrations for this region. The calibration coefficients required to attain model calibration are shown on each of the plots and summarized in **Table 8**.

Figure 24: Results of Ecoregions 42/43 Total Phosphorus Model Calibration (Canfield & Bachman (1981), Natural Lakes)

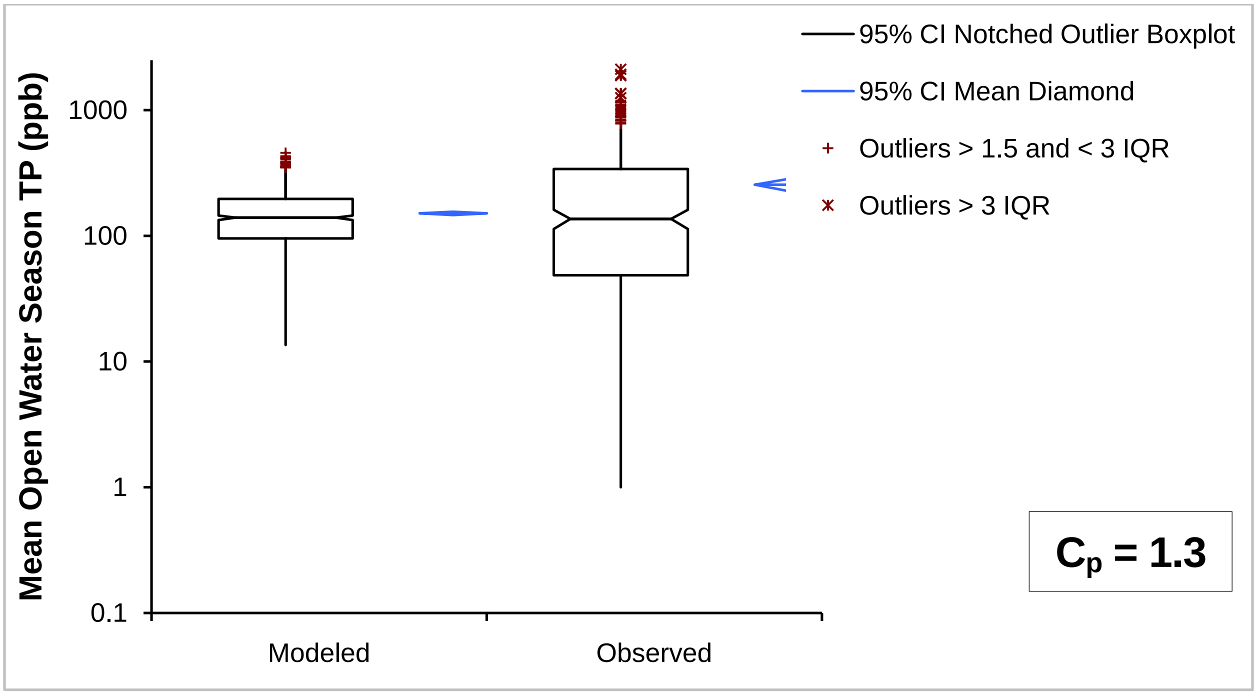


Figure 25: Results of Ecoregions 42/43 Total Phosphorus Model Calibration (Canfield & Bachman (1981), Lakes + Reservoirs)

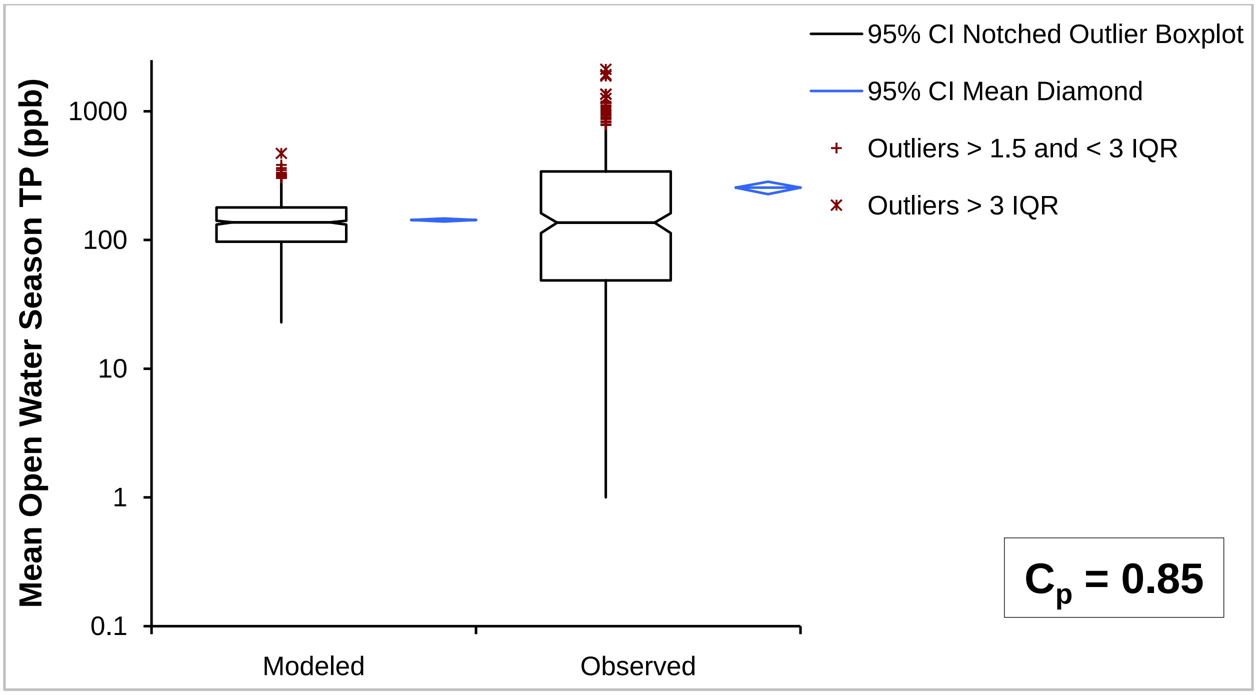


Figure 26: Results of Ecoregions 42/43 Total Nitrogen Model Calibration (Bachman – Volumetric Load)

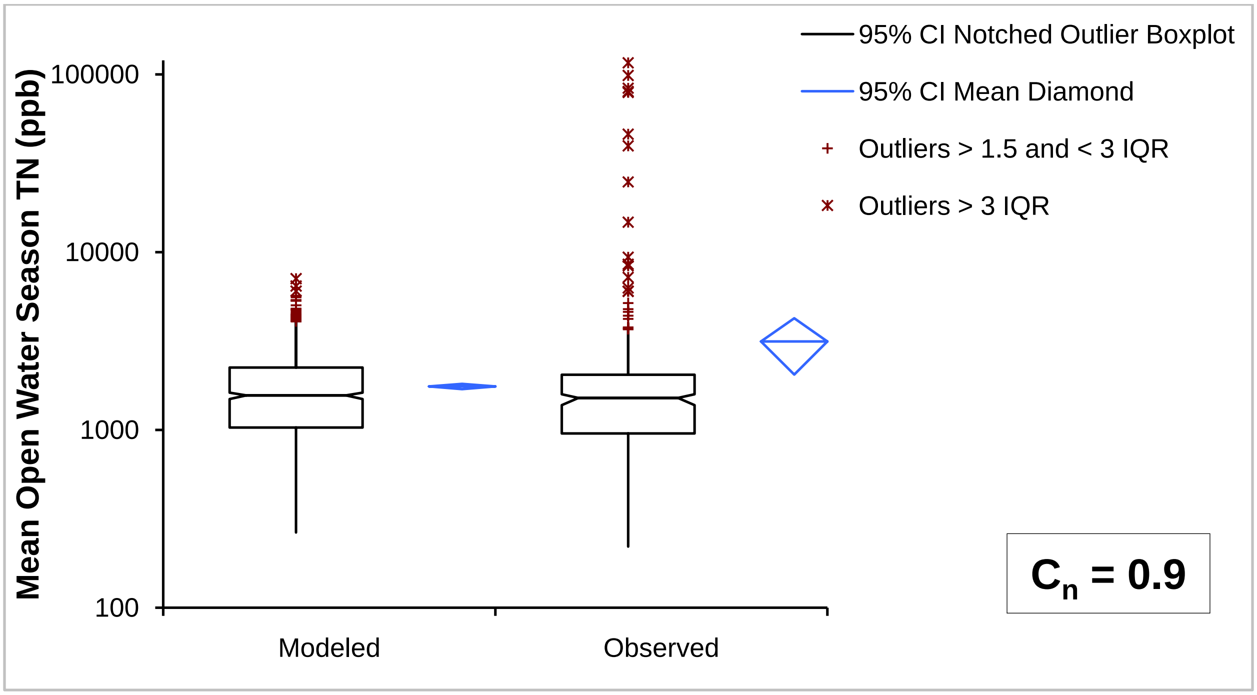


Figure 27: Results of Ecoregions 42/43 Chlorophyll-a Model Calibration (Chl-a vs. Combined Nutrient)

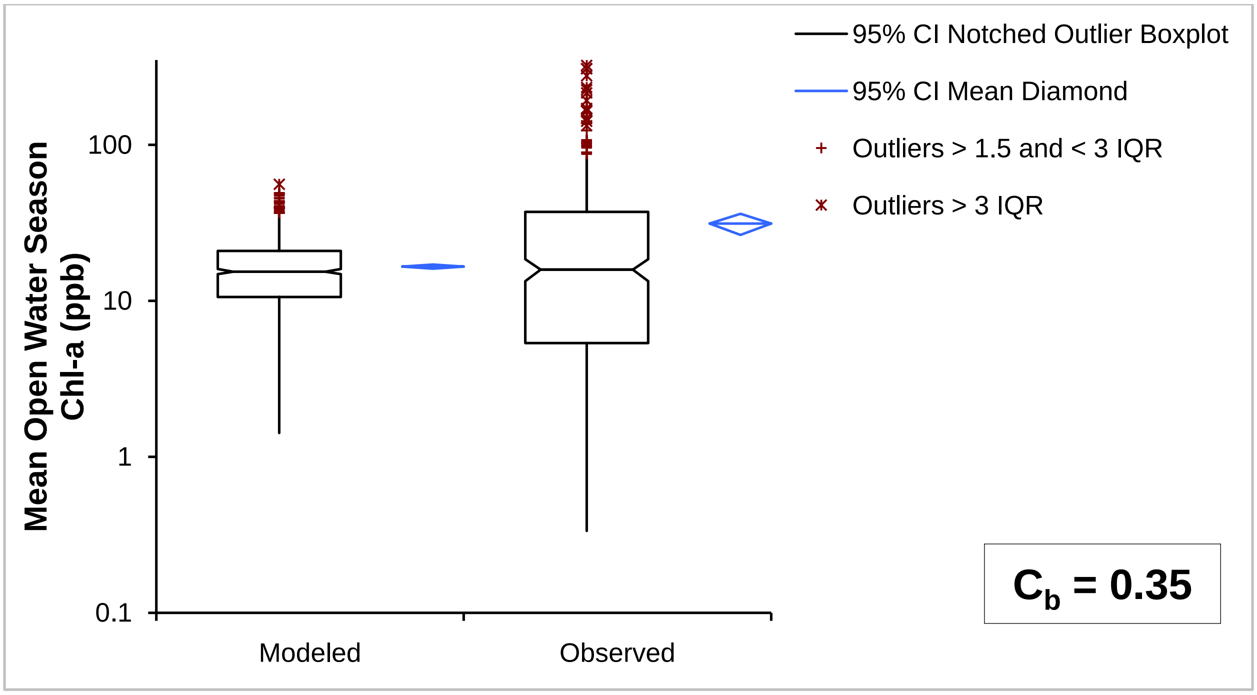


Figure 28: Results of Ecoregions 42/43 Chlorophyll-a Model Calibration (Chl-a vs. TP)

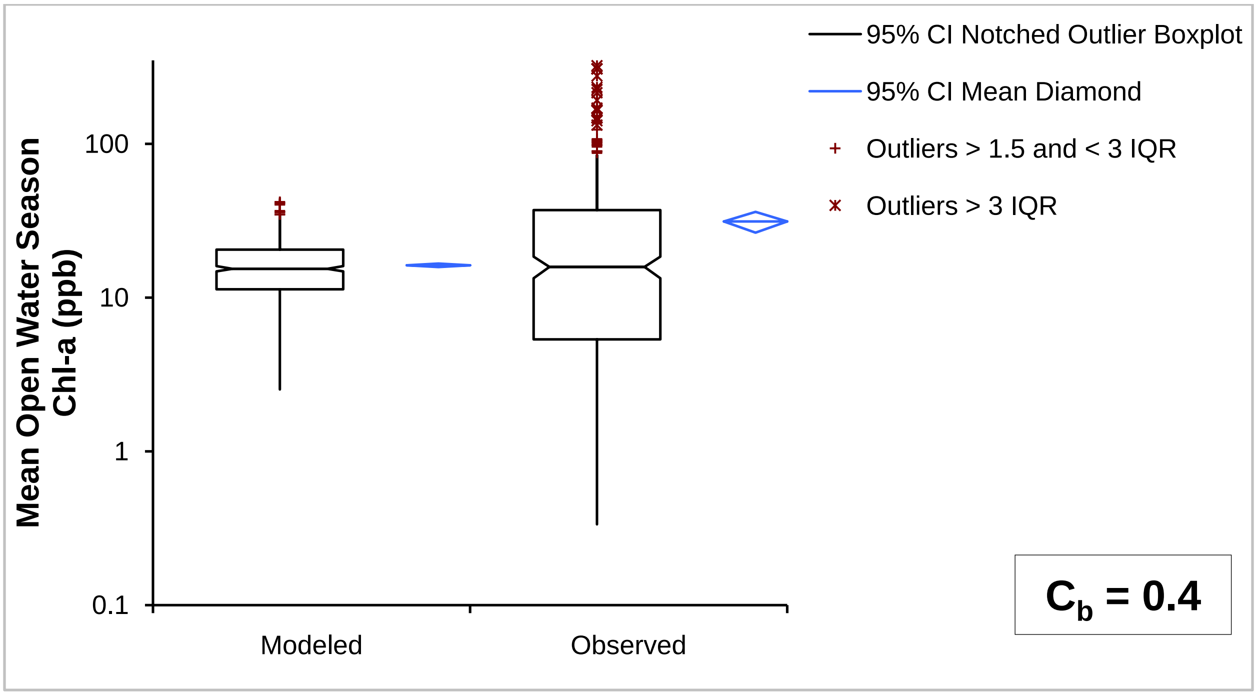


Figure 29: Results of Ecoregions 42/43 Secchi Depth Model Calibration (Secchi Depth vs. Combined Nutrient)

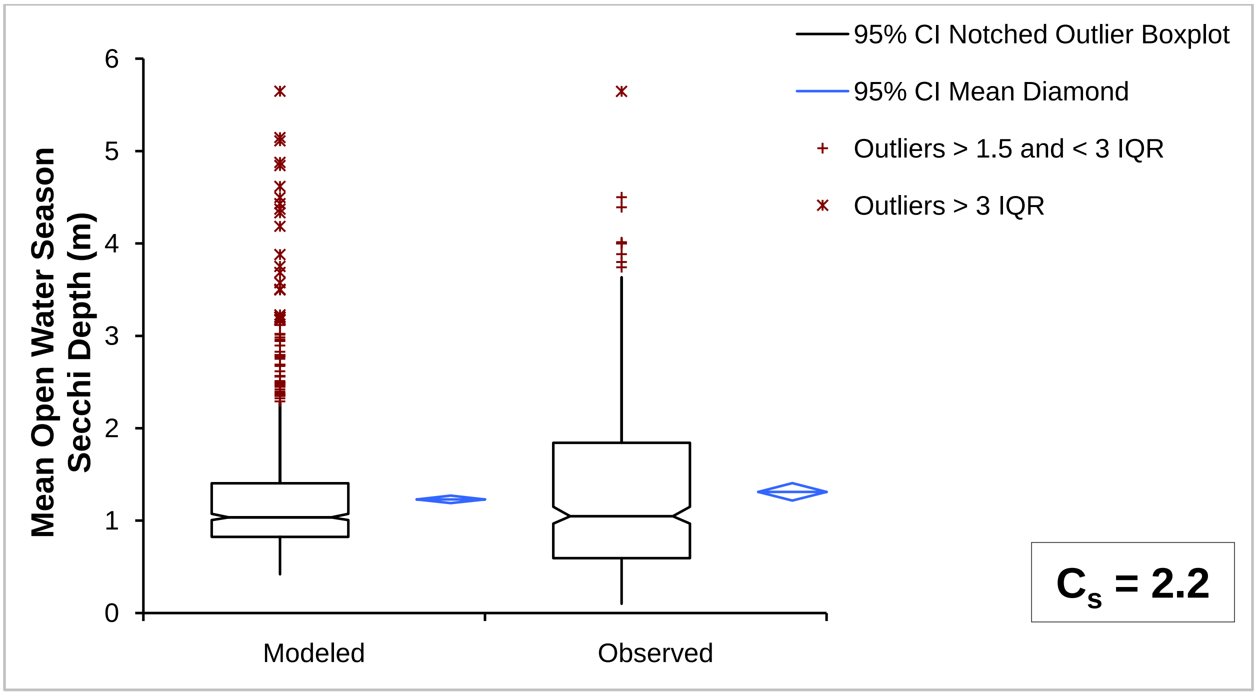


Figure 30: Results of Ecoregions 42/43 Secchi Depth Model Calibration (Secchi Depth vs. TP and Non-Algal Turbidity)

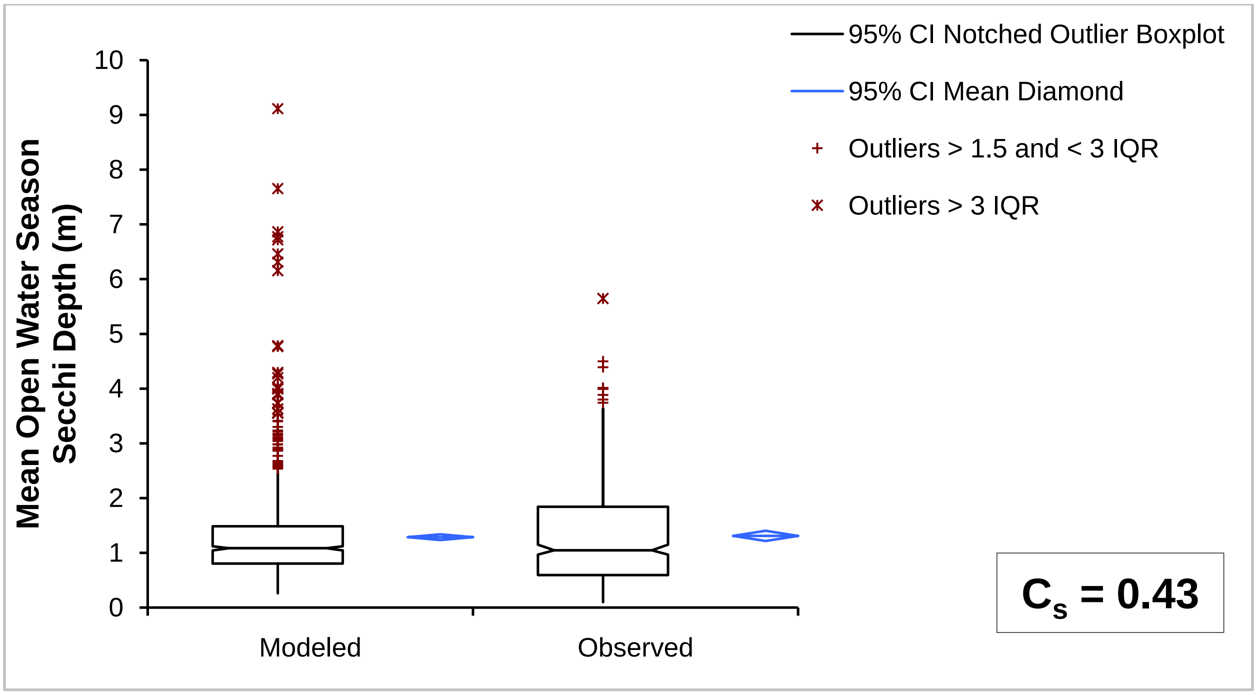


Table 8 summarizes the Ecoregions 42/43 water quality model calibration coefficients. In this region, calibration results show that the "Canfield & Bachmann (1981), Reservoirs + Lakes" model better represents the process of phosphorus sedimentation in Ecoregions 42/43 reservoirs than the "Canfield & Bachmann (1981), Natural Lakes" model. Similar to what was seen in Ecoregion 46, the equation that computes Chl-a as a function of TP does a slightly better job representing the data than that which computes it as a function of the combined nutrient. Again, secchi depths are best reflected by the equation that computes them as a function of Chl-a and non-algal turbidity.

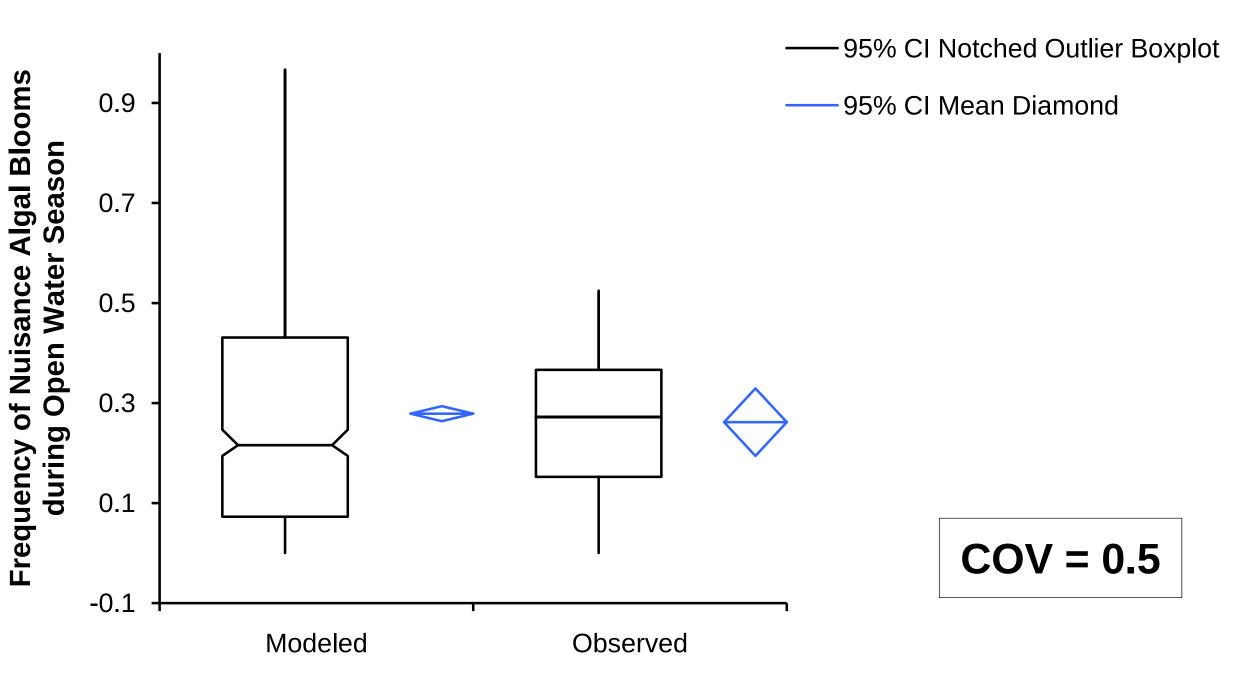
Table 8: Resultant Ecoregions 42/43 Water Quality Model Calibration Coefficients

Model	Calibration Coefficient	Value after Calibration
Total Phosphorus Sedimentation (Canfield & Bachman (1981), Natural Lakes)	C_p	1.3
Total Phosphorus Sedimentation (Canfield & Bachman (1981), Lakes + Reservoirs)	C_p	0.85
Total Nitrogen Sedimentation (Bachman (1980), Volumetric Load)	C_n	0.9
Chl-a vs. combined nutrient[1]	C_b	0.35
Chl-a vs. TP[1]	C_b	0.4
Secchi depth vs. combined nutrient[1]	C_s	2.2
Secchi depth vs. TP and non-algal turbidity[1]	C_s	0.43

[1] The calibrated "Canfield & Bachman (1981), Lakes + Reservoirs" and "Bachman (1980), Volumetric Load" equations for TP and TN sedimentation were in use during these simulations.

Again, the last step in calibrating the Ecoregions 42/43 water quality model was to adjust the Chl-a intra-annual COV to match the simulated and observed frequency of nuisance algal blooms (Chl-a > 20 ppb) in any given year. The calibrated best fit COV for this region is 0.5. **Figure 31** shows the analysis results.

Figure 31: Results of Ecoregions 42/43 Frequency of "Nuisance" Bloom (Chl-a > 20 ppb) Estimation



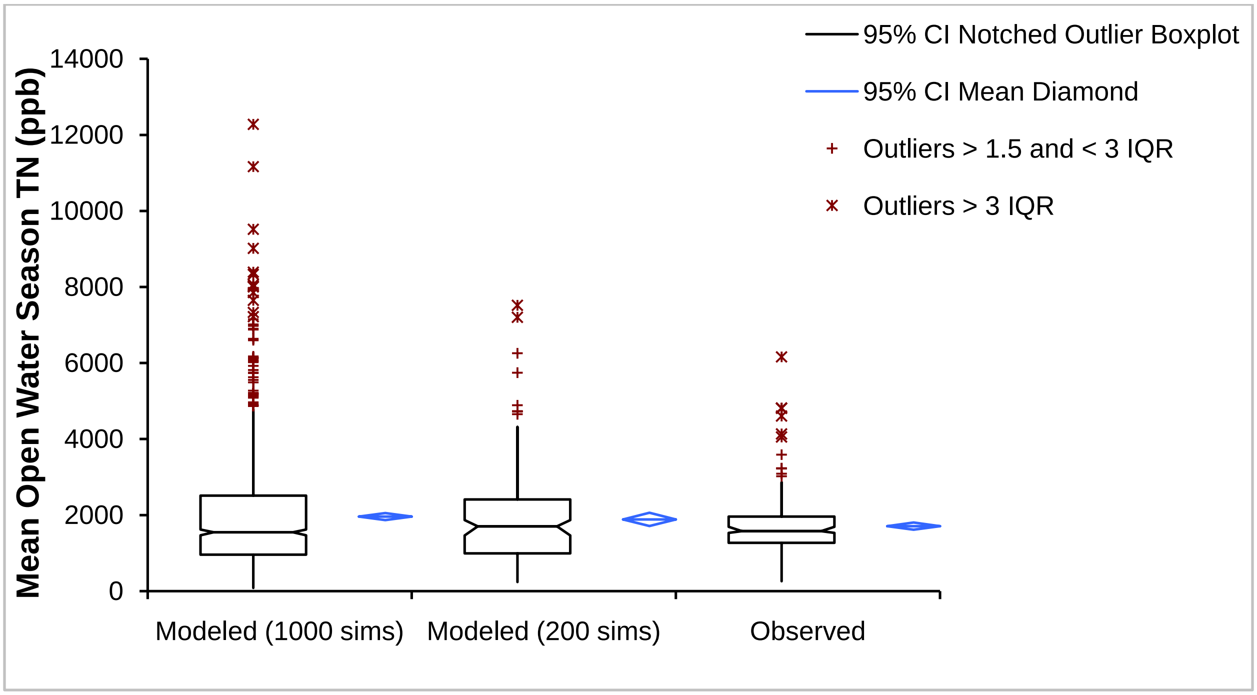
DISCUSSION

The watershed loading/receiving water eutrophication models created in this Task were designed to reflect nutrient loadings to and eutrophication response within the reservoirs of the two modeling regions: EPA Level 3 Ecoregion 46 and EPA Level 3 Ecoregions 42/43. A stochastic modeling approach, using Monte Carlo simulations, was developed to reflect the variation in precipitation, drainage area characteristics, reservoir morphometry, and water quality that's observed across the study area. The calibrated models can now be used to simulate nutrient loads from various management strategies and/or land cover changes within the reservoir drainage areas and their associated impacts on water quality. Results of the model applications should be used to inform the establishment of nutrient criteria in the modeled regions.

Outcomes of the water quality calibration show that, in general, the empirical water quality relationships available through the CNET/BATHTUB model allow for an accurate estimate of the central tendency of the observed reservoir mean annual water quality data. However, the simulation of the distribution of the mean water quality values (e.g., the 25th and 75th percentiles) is poorer. In a few cases, the simulated water quality distributions show more outliers than the observed values. In other cases, the observed values show a large number of outliers. While a number of explanations for this can be given, it is important to note since some methods for setting nutrient criteria rely on these outer percentile and (improper) skewness in the modeled distributions could potentially have an impact on the results.

One cause of the increased number of outliers in the simulated values is likely due to the number of simulations being performed in the stochastic modeling. In this work, the various water quality models were performed for 1,000 simulations, allowing the creation of the estimated value distributions. In comparison, the observed water quality values had between 200 and 300 values for use in creating the observed water quality distributions. If the modeled distributions were created using fewer simulations, the resultant modeled distributions look more similar to the observed distributions. For example, **Figure 32** shows the result of simulating secchi depths in Ecoregion 46 using 50 model simulations in the stochastic modeling instead of the 1,000 used elsewhere. Also shown in the figure is the result that was shown using 1,000 simulations (i.e., the same output shown in **Figure 19** above). This analysis shows that running the model with fewer simulations results in a distribution that's more similar to what's shown with the observed data. This analysis shows that the distribution created by running 1,000 simulations (essentially) represents the distribution that would be expected if 1,000 reservoir-year value combinations were available for analysis and is, therefore, more reflective of the complete population of expected water quality in the modeled region.

Figure 32: Ecoregion 46 TN Estimated with 200 vs. 1,000 Model Simulations



When large numbers of outliers are seen in the observed (mean open water season) data, those outliers were often computed as a function of one or two data points, which increases the likelihood of error. The values created by the CNET model are (in theory) reflective of the water quality that would be seen if a number of representative samples were collected in the regions reservoirs and averaged per waterbody. Therefore (for comparison purposes), the observed mean open water season water quality values would ideally be based on a number of individual samples (five or more, for example) that were collected in each reservoir during each year. Therefore, when the mean values are based on only one or two samples, the potential that the observed value is not representative of mean conditions in the waterbody is increased and the distribution may (misleadingly) be skewed.

Another potential reason behind the simulated water quality distributions not matching the observed distributions is the fact that the empirical equations used in the CNET BATHTUB model were developed for reservoirs with average hydraulic residence times less than 2 years and caution that utilizing the equations for estimating water quality in waterbodies with hydraulic residence times greater than that will introduce some error into the calculations. Average hydraulic residence times for the reservoirs in the study area are estimated at between 1 day and 24 years, with an average hydraulic residence time of 1.1 years. The inclusion of reservoirs with longer residence times, therefore, likely introduces some modeling error.

Given the results of the model calibration and performance, the following empirical equations are recommended for use in modeling the anticipated eutrophication response of the reservoirs in the study area:

- Phosphorus sedimentation in Ecoregion 46: Canfield & Bachman (1981), Natural Lakes
- Phosphorus sedimentation in Ecoregions 42/43: Canfield & Bachman (1981), Lakes + Reservoirs
- Nitrogen Sedimentation: Bachman (1980), Volumetric Load
- Chlorophyll-a Eutrophication Response: Chl-a vs. combined nutrient
- Secchi Depth Eutrophication Response: Secchi depth vs. TP and non-algal turbidity

REFERENCES

Walker, W.W. (1984) Statistical Bases for Mean Chlorophyll-a Criteria. Lake and Reservoir Management: Practical Applications, North American Lake Management Society, Proceedings of Fourth Annual Conference, McAcfee, New Jersey.

Walker, W.W. (1996) Simplified Procedures for Eutrophication Assessment & Prediction: User Manual Instruction Report W-96-2 USAE Waterways Experiment Station, Vicksburg, Mississippi. (Updated September 1999)

Section VII: Modeling Results

Section VII: Modeling Results

MEMO

(External Correspondence)

To: Tina Laidlaw

Date: April 2, 2011

Cc: File 4965-002
Dennis McIntyre, GLEC

From: Stephanie Johnson, Ph.D., P.E.

Through: Mark R. Deutschman, Ph.D., P.E.

Subject: Summary of Modeling Results associated with Task 5 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8

This memorandum addresses the results of applying two regional watershed loading and eutrophication models that were created under Task 4 of this project, as described in an April 1, 2011 memorandum. The models were applied to simulate various land management/cover scenarios in the watersheds contributing to each modeling region's reservoirs. This work was performed under Task 5 of EPA Contract #EP-C-09-001: Development of Nutrient Criteria for Lakes and Reservoirs for North Dakota and other Plains States in Region 8 (i.e. the Nutrient Criteria Project). Results of the modeling may be used to guide the development of nutrient criteria in the study area, providing insight on how area reservoirs are expected to respond to reduced nutrient loadings.

BACKGROUND

As described in the April 1, 2011 memo addressing model development and calibration, two watershed loading/receiving water models were created for the Nutrient Criteria Project study area. One model addresses reservoirs contained in EPA Level 3 Ecoregion 46; the other represents reservoirs in EPA Level 3 Ecoregions 42/43. These models were developed to simulate nutrient loads into and eutrophication responses of reservoirs within the two model regions. The watershed loading models compute the surface water runoff, total phosphorus (TP) load, and total nitrogen (TN) load expected from each region's reservoir drainage areas. The receiving water models then use these data as inputs to estimate the in-reservoir eutrophication responses, including: mean open water season (March 1 – November 30) TP, TN, combined nutrient (a combination of TP and TN), and chlorophyll-a (Chl-a) concentrations, as well as mean open water season secchi disk depths. The likelihood (or frequency) of experiencing nuisance algal blooms within any given year is also estimated through statistical relationships. For the purposes of this work, the Nutrient Criteria Project Team defined a "nuisance" algal bloom as an in-reservoir Chl-a concentration equal to or greater than 20 parts per billion (ppb).

SCENARIO DEVELOPMENT

Each regional model was initially developed to represent the regional reservoirs/drainage areas as they currently exist (i.e., the Base Condition). As described in the April 1, 2011 memorandum, the variables that represent the status of reservoir drainage area conditions in the watershed loading models are: land cover/use curve number (CN), TP estimated mean concentration (EMC), and TN EMC. These variables are also responsible for controlling the amount of surface water runoff, TP load, and TN load that enter the regions' reservoirs. The CN and EMC values are input to the models as a function of general land use categories within the reservoirs' contributing drainage areas. **Table 1** summarizes the mean CN and EMC values used in the models by modeled region and land use category. CN and EMC values were developed without consideration of best management practices (BMPs) or other controls that may be implemented on the landscape to impact the amount of water and/or nutrients contributed from these areas (i.e., values are reflective of what would be seen on un-managed land). As shown in **Table 1**, (unmanaged) agricultural lands generally have higher CN and EMC values than (unmanaged) grasslands (i.e., native cover) resulting in more runoff and higher nutrient loads per modeled area.

Table 1: Average Characteristics of the Modeled General Land Use Categories

General Land Use Category	Average CN (CN, Ecoregion(s))	Mean TP EMC (mg/L)	Mean TN EMC (mg/L)
Agriculture, Row Crops	79 (46); 77 (42/43)	0.75	6.15
Forest, Woods	65 (46); 57 (42/43)	0.15	1.92
Grasslands/Shrubs/Wetlands (G/S/W)	68 (46); 66 (42/43)	0.28	4.62
Urban, Impervious	72 (46); 70 (42/43)	0.46	2.36

In addition to the Base Conditions model, a number of different model scenarios were created. The goal of these modeling scenarios was to understand how different runoff amounts and nutrient loads entering the regions' reservoirs from their contributing drainage areas may impact water quality within the reservoirs' waters. To simulate these reductions in surface water runoff and nutrient loading, each model scenario assumes that some increased fraction of the agricultural land within the reservoir drainage areas would be converted to management and/or cover that is better reflected through the (lower) CN and EMC values associated with native plant covers (i.e., G/S/W) than those of row crop agriculture. Practical (i.e., real world) reasons for this conversion (to lower CNs and EMCs) may be the implementation of agricultural BMPs or other management techniques that serve to reduce the amount of runoff and nutrient load contributed from these areas. Another reason may be the conversion of lands of agricultural row crop production into native cover. For the purposes of this modeling, however, conversion of lands from row crop agriculture to native cover is used simply as a surrogate to generate a range of nutrient loads into the regions' reservoirs for purposes of understanding how the reservoirs may respond.

Table 2 and **Table 3** summarize the conditions included in each model scenario, including the percent of each reservoir drainage area simulated using native cover. **Table 2** summarizes this information for the Ecoregion 46 models; **Table 3** shows it for the Ecoregions 42/43 models. According to the 2001 National Land Cover Dataset (NLCD) data for the study area (the most recently available at the time of model development), the reservoir contributing drainage areas in Ecoregion 46 had, on average, 71% of their land use in agricultural cover in 2001. This value was used in the Base Conditions model. Given this large fraction of agricultural cover under Base Conditions, six additional model scenarios were run for Ecoregion 46 to simulate increasing portions of those lands under native cover. NLCD 2001 data show that Ecoregions 42/43 reservoir drainage areas have a much lower fraction of cover under Base Conditions (25%) so only two additional model scenarios were created for that area. In general, the amount of land in native cover was increased by 10% with each subsequent model run.

Table 2: Ecoregion 46 Modeling Scenarios

Model Scenario	% of Drainage Area Simulated under Native Cover	Median of Simulated Distribution	
		Mean Open Water Season TP Load (kg)	Mean Open Water Season TN Load (kg)
Base Condition	20%	887	5,662
Scenario 1	31%	760	5,097
Scenario 2	41%	665	4,521
Scenario 3	51%	559	4,061
Scenario 4	61%	445	3,558
Scenario 5	71%	345	2,860
Scenario 6	81%	251	2,150

Also shown in **Table 2** and **Table 3** are the medians of the simulated mean open water season TP and TN loads into the regions' reservoirs under each scenario. As expected, as more land is simulated as native cover the associated nutrient loads go down.

Table 3: Ecoregions 42/43 Modeling Scenarios

Model Scenario	Percent of Drainage Area Simulated under Native Cover	Median of Simulated Distribution	
		Open Water Season TP Load (kg)	Open Water Season TN Load (kg)
Base Condition	70%	229	1,765
Scenario A	75%	196	1,539
Scenario B	80%	125	1,120

MODEL RESULTS

The following section presents the resultant in-reservoir eutrophication responses associated with the reduced nutrient loads simulated under each model scenario summarized in **Table 2** and **Table 3**. As nutrient loads decrease, water quality improvements within the regions' reservoirs should be seen. **Figure 1** through **Figure 4** show the results of the various modeling scenarios in Ecoregion 46, concentrating on in-reservoir nutrient and Chl-a concentrations and the associated secchi depths. **Table 4** summarizes all of the modeled reservoir response variables for each simulation.

Figure 1: Ecoregion 46 Simulated Mean Open Water Season TP Concentrations by Model Scenario

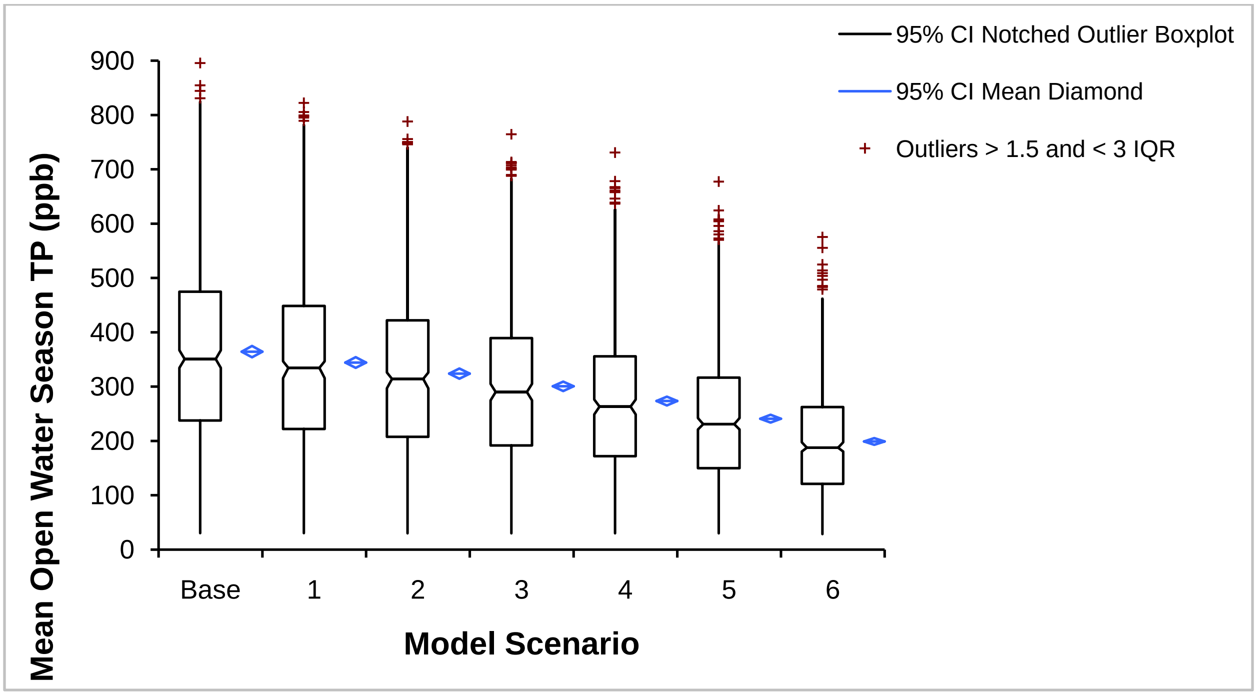


Figure 2: Ecoregion 46 Simulated Mean Open Water Season TN Concentrations by Model Scenario

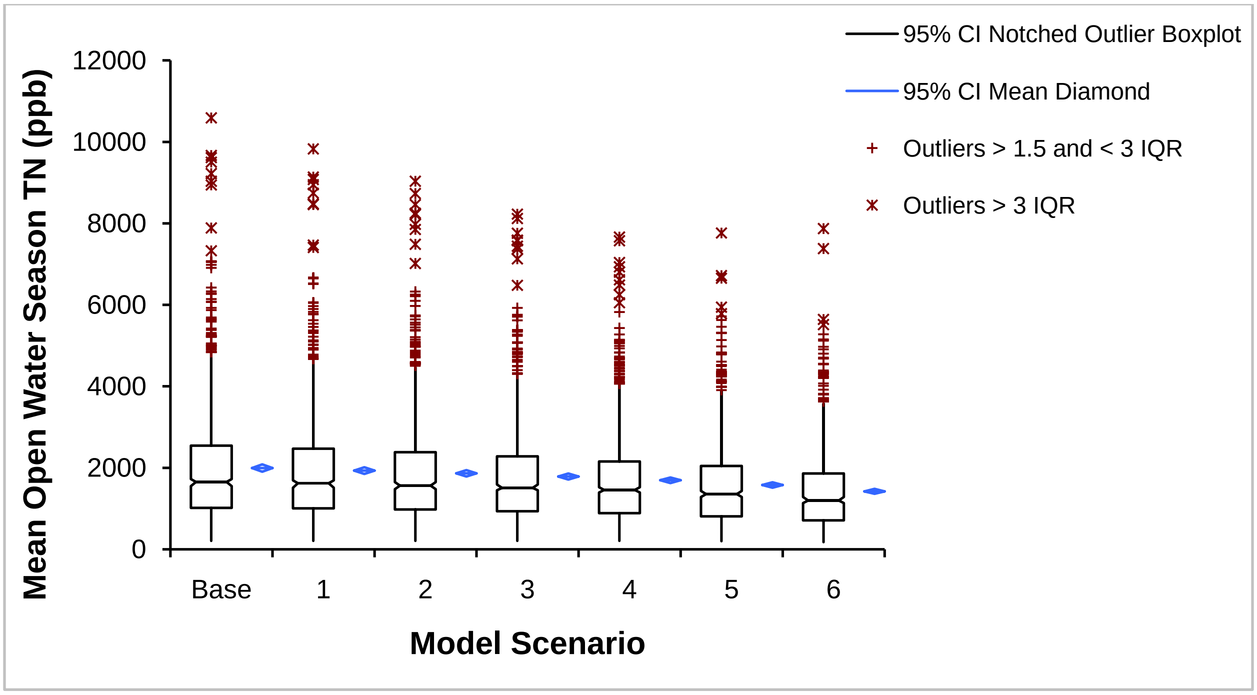


Figure 3: Ecoregion 46 Simulated Mean Open Water Season Chl-a Concentrations by Model Scenario

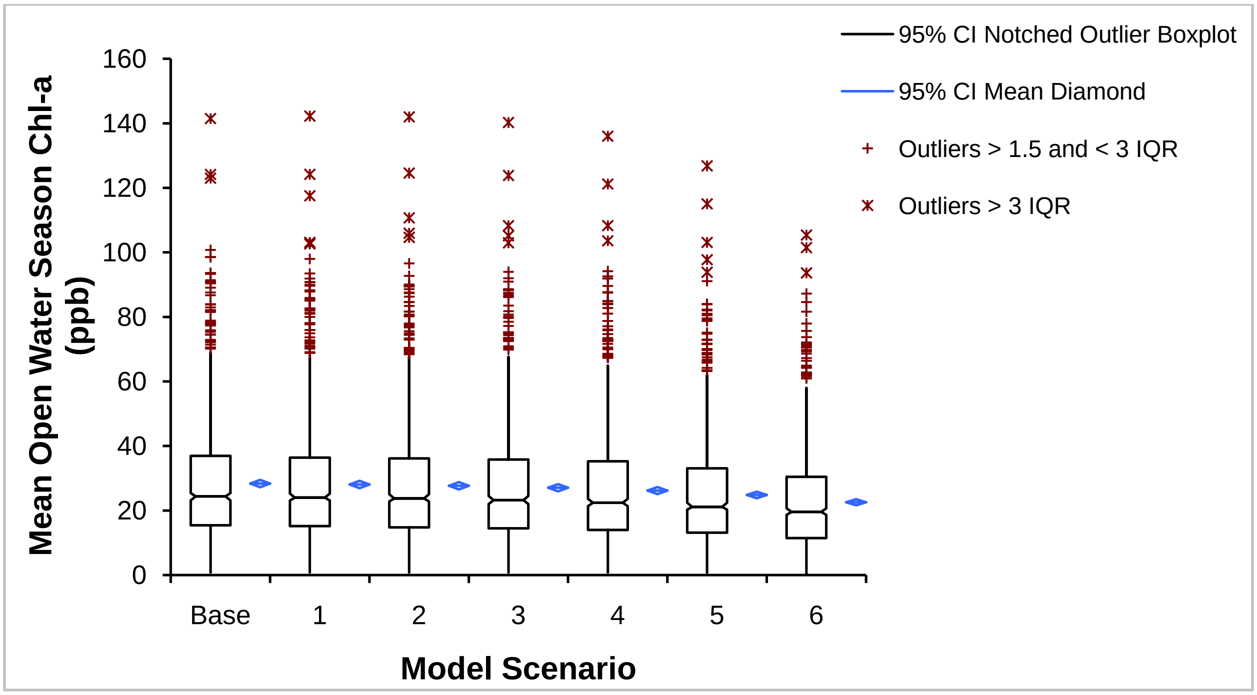


Figure 4: Ecoregion 46 Simulated Mean Open Water Season Secchi Depths by Model Scenario

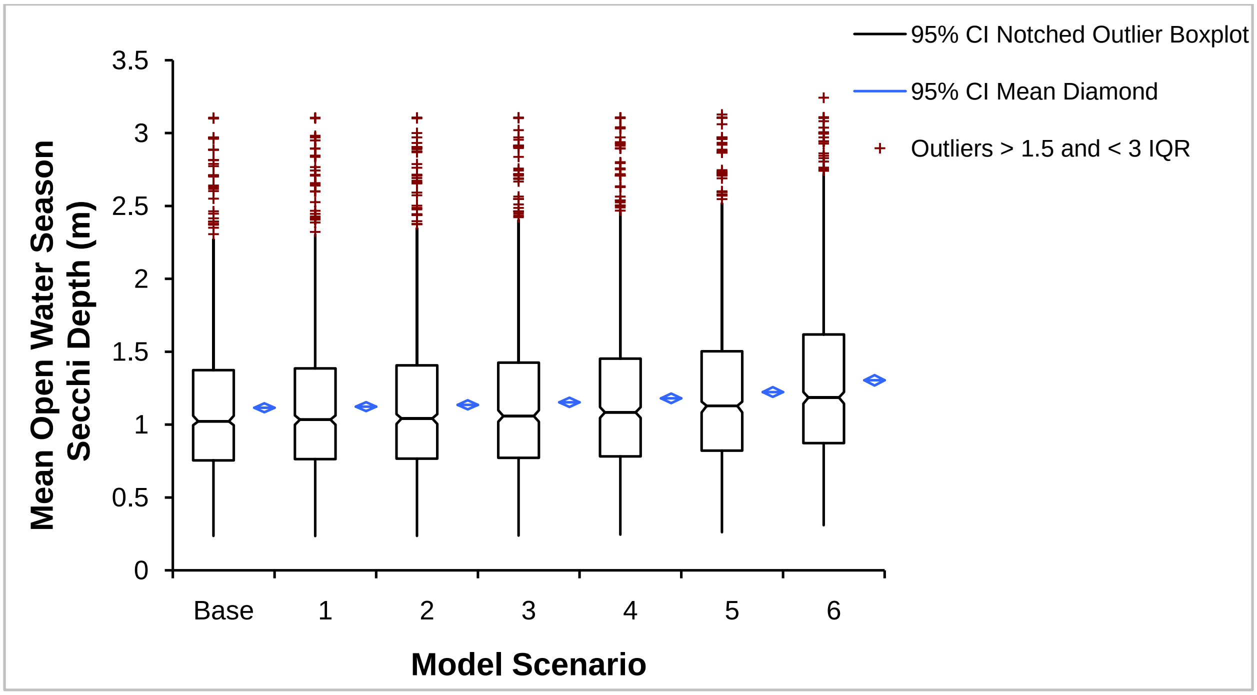


Table 4: Summary of Model Results for Ecoregion 46

Model Scenario	Expected (i.e., Median) Mean Open Water Season Values								Expected Nuisance Algal Bloom Frequency[1]
	TP (ppb)	TN (ppb)	Combined Nutrient (ppb)	Chl-a (ppb)	Secchi Depth (m)	Chl-a TSI	TP TSI	Secchi Depth TSI	
Base Condition	351	1,651	116	24.4	1.02	62.0	88.7	59.7	38.2%
Scenario 1	335	1,620	113	24.0	1.03	61.8	88.0	59.5	37.6%
Scenario 2	314	1,564	109	23.8	1.04	61.7	87.1	59.4	37.2%
Scenario 3	290	1,509	104	23.2	1.06	61.5	86.0	59.2	36.3%
Scenario 4	263	1,457	99	22.4	1.08	61.1	84.6	58.8	35.0%
Scenario 5	231	1,355	90	21.1	1.13	60.5	82.7	58.3	32.8%
Scenario 6	188	1,199	79	19.6	1.19	59.8	79.7	57.5	30.1%

[1] Defined as a Chl-a concentration ≥ 20 ppb.

Figure 5 through **Figure 8** show the modeled in-reservoir eutrophication responses for the three model scenarios run in Ecoregions 42/43. **Table 5** summarizes all of the in-reservoir results.

Figure 5: Ecoregions 42/43 Simulated Mean Open Water Season TP Concentrations by Model Scenario

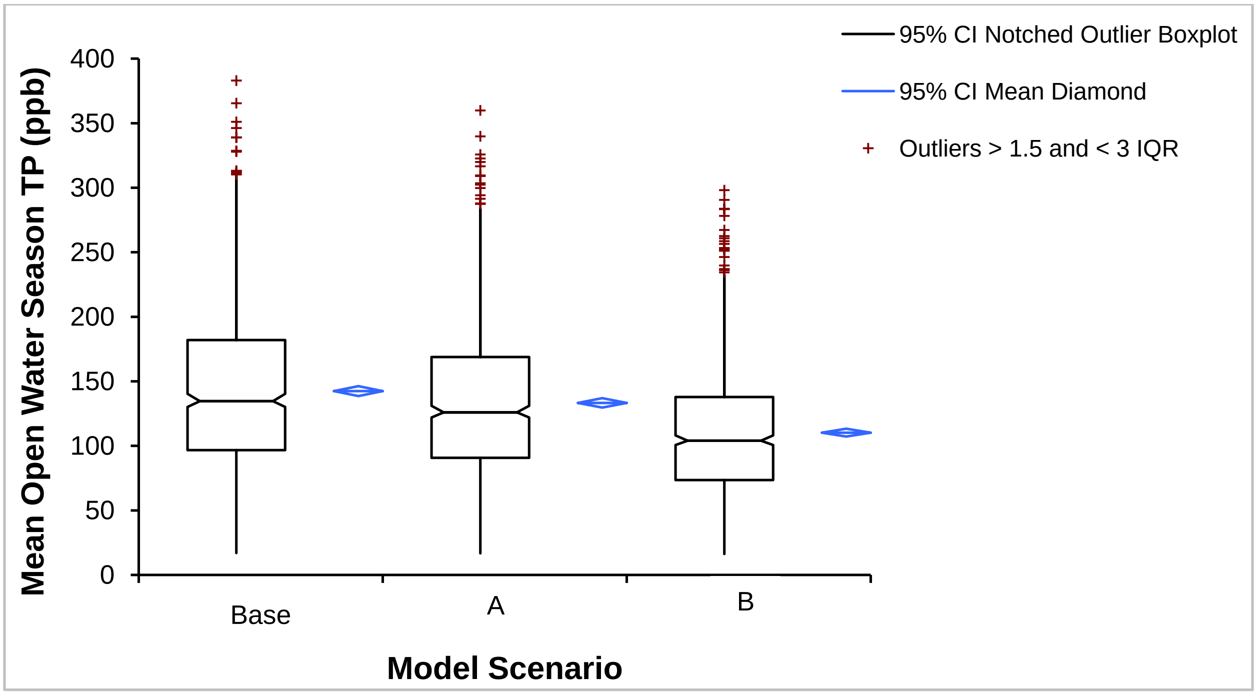


Figure 6: Ecoregions 42/43 Simulated Mean Open Water Season TN Concentrations by Model Scenario

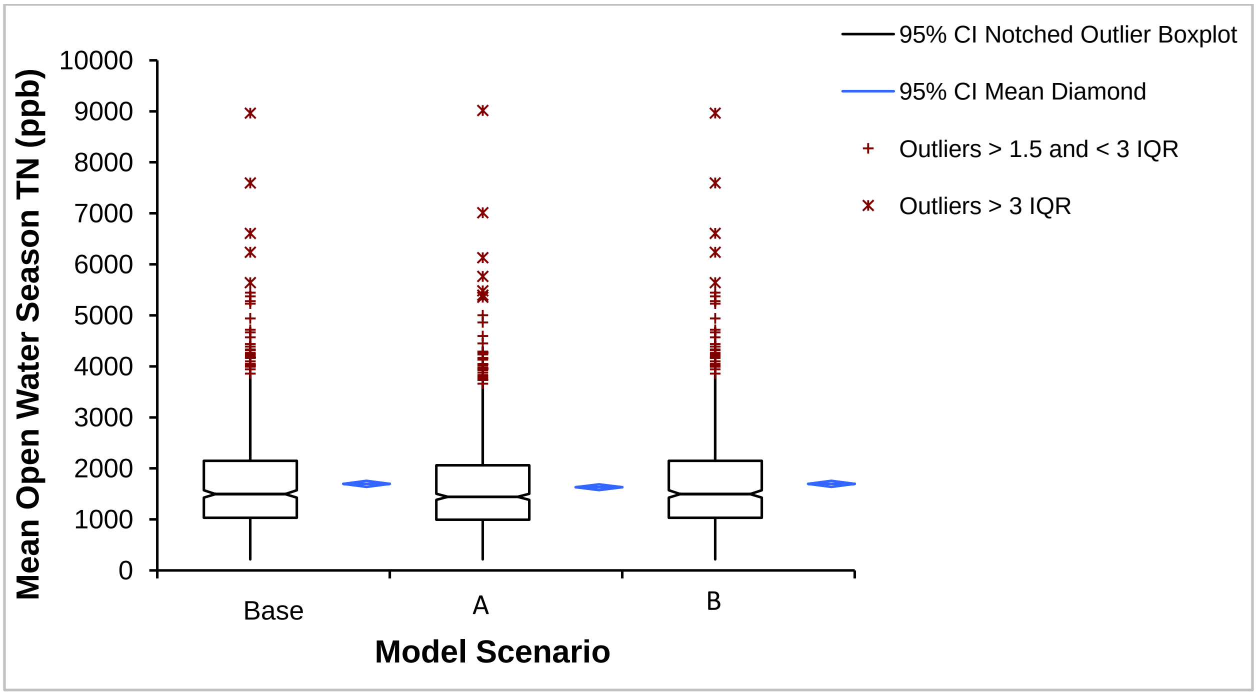


Figure 7: Ecoregions 42/43 Simulated Mean Open Water Season Chl-a Concentrations by Model Scenario

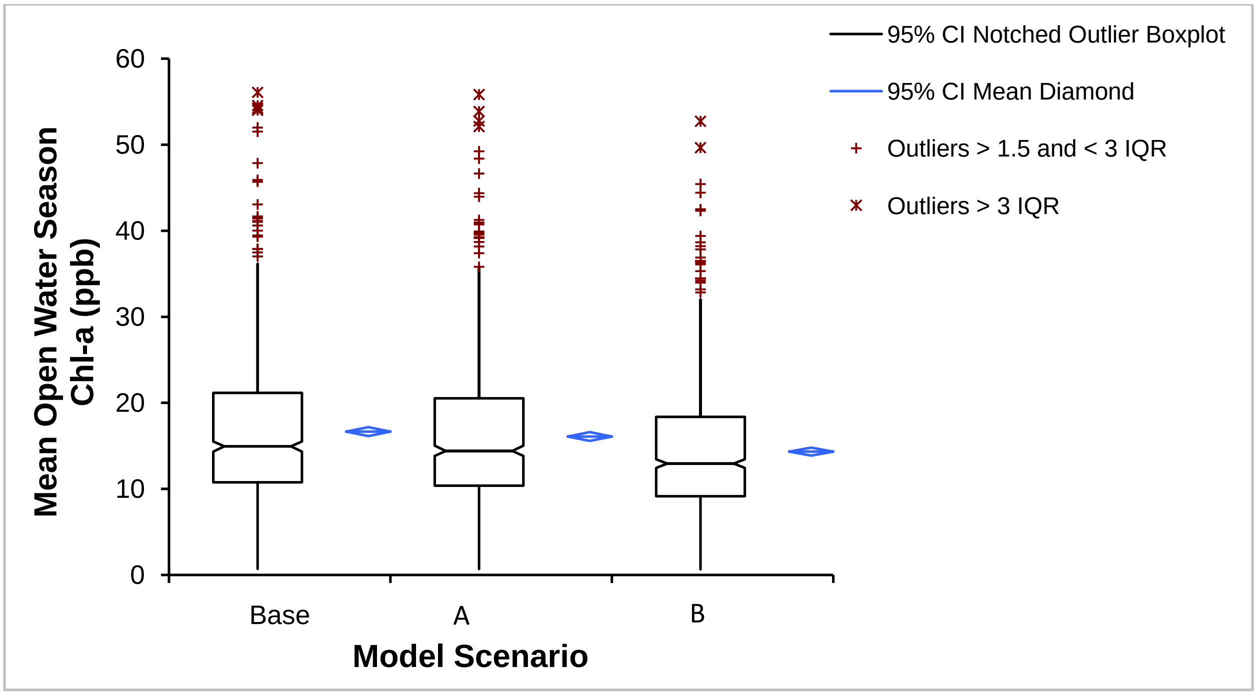


Figure 8: Ecoregions 42/43 Simulated Mean Open Water Season Secchi Depths by Model Scenario

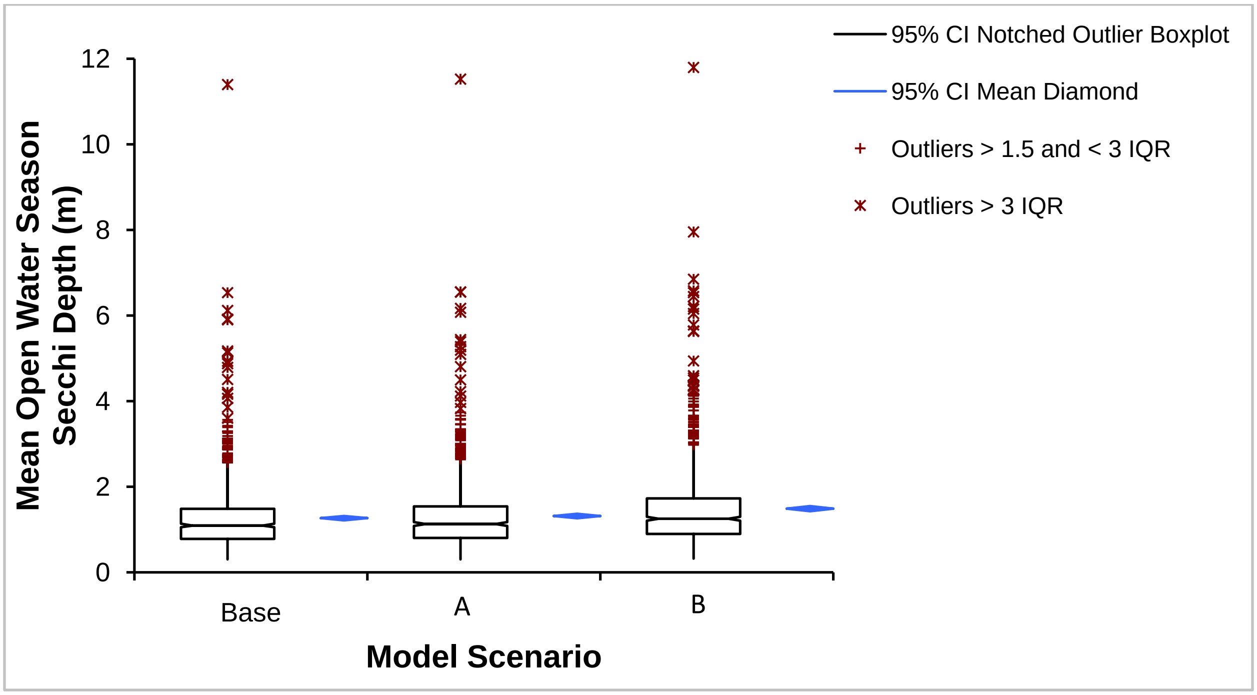


Table 5: Summary of Model Results for Ecoregions 42/43

Model Scenario	Expected (i.e., Median) Mean Open Water Season Values								Expected Nuisance Algal Bloom Frequency[1]
	TP (ppb)	TN (ppb)	Combined Nutrient (ppb)	Chl-a (ppb)	Secchi Depth (m)	Chl-a TSI	TP TSI	Secchi Depth TSI	
Base Condition	135	1,495	82	15.0	1.09	57.2	74.9	58.7	20.3%
Scenario A	126	1,444	78	14.4	1.13	56.8	73.9	58.2	18.3%
Scenario B	104	1,495	68	12.9	1.25	55.7	71.2	56.8	13.1%

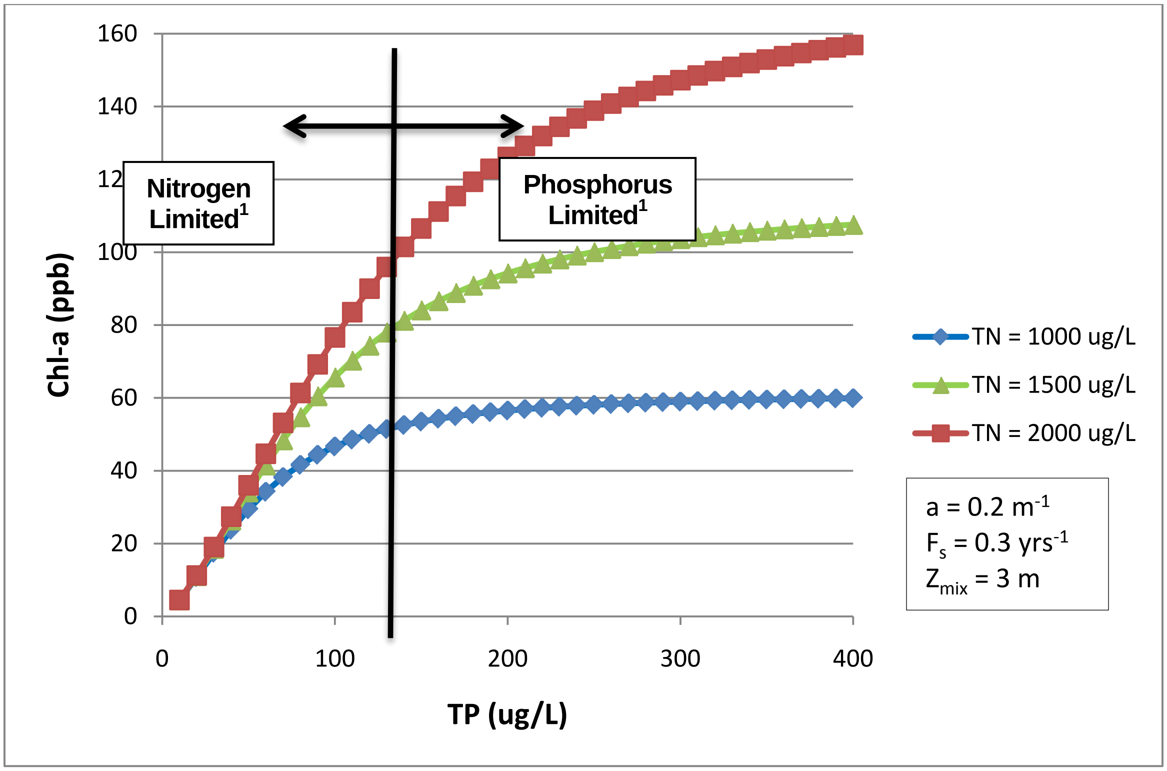
[1] Defined as a Chl-a concentration ≥ 20 ppb.

DISCUSSION

As expected, results of the modeling show that as the nutrient loads into the study area reservoirs reduce, the in-reservoir water quality will improve. The simulated improvements in watershed nutrient loading into the reservoirs, however, are more dramatic than the response seen within the waterbodies themselves. In Ecoregion 46, for example, a 70% reduction in TP and 60% reduction in TN open water season loading (the difference between the Base Conditions and Scenario 6) resulted in a 40% reduction in in-reservoir mean open water season TP concentration, 30% reduction in mean open water season TN concentration, and 20% improvement in mean open water season Chl-a concentration and secchi disk depth. Similar results are seen in the Ecoregions 42/43 model, where 40-45% reductions in nutrient loading lead to 15-20% improvements in in-reservoir mean open water season concentrations and secchi disk depths. In addition (and somewhat related to), results of the Base Conditions modeling show that the reservoirs in the study area are nitrogen limited.

Assuming that reservoirs with a TN:TP ratio of less than 15 are TN limited, under current conditions, reservoirs in the study area are nitrogen limited (the Ecoregions 42/43 model shows a TN:TP ratio of 11 under Base Conditions; the Ecoregion 46 model shows a ratio of approximately 5). Modeled Chl-a concentrations in the study area reservoirs are simulated as a function of a combined nutrient (a combination of TP and TN, as described in the April 1, 2011 memo). The modeled stressor-response relationship is shown in **Figure 9**, where Chl-a is plotted as a function of both TP and TN concentrations (and some assumed constant water body characteristics). As shown in this plot, when waterbodies are nitrogen limited the response in Chl-a concentration is more dependent on reductions in TN than TP. Given that improvements in TP loadings (and associated in-reservoir concentrations) are greater than those in TN values under the model scenarios, in-reservoir water quality improvements are somewhat muted.

Figure 9: Stressor-Response Relationship used in Regional Models



[1] Regions of nitrogen and phosphorus limitation are based on the TN = 2000 ug/L curve.

Similar to the improvement in mean open water season water quality, the frequency of expected nuisance algal blooms (defined by the Project Team as Chl-a concentrations ≥ 20 ppb) is reduced as nutrient loads decrease. Depending on the tolerance of stakeholders/regional citizens for experiencing these types of events, these results may be helpful to the Project Team as they make their decisions on using the model results to inform nutrient criteria standards.

The overall goal of this model development and application was to provide information to the EPA and the Project Team for use in establishing nutrient criteria standards for the Plains States of EPA Region 8. The results presented in **Table 4**, **Table 5**, and the figures of this memorandum provide an appreciation of how different nutrient loads delivered to the area's reservoirs may affect the in-reservoir eutrophication response. The Project Team can now combine these resultants with other considerations (such as citizen perceptions of water quality, state and federal regulations, and other policy goals) to set nutrient criteria protective of the water resources of the area.