Creation of CAPS-IBI Software and Lake Nutrient Modeling: Components of the Massachusetts Comprehensive Wetlands Assessment and Monitoring Program

Final Report

Wetlands Program Development Grant Cooperative Agreement # 96168701

Scott Jackson, Kevin McGarigal, Ethan Plunkett,
Elizabeth Homa and Brad Compton

Department of Environmental Conservation, University of Massachusetts Amherst

And

Robert English

Daystar Computing and Research Services

July 28, 2015

Introduction

The University of Massachusetts Amherst (UMass) has been working since 2006 with the Massachusetts Department of Environmental Protection (MassDEP) and the MA Office of Coastal Zone Management (MCZM) to develop and implement a comprehensive Wetlands Assessment and Monitoring Program for Massachusetts. Our assessment tools, based on EPA's *Application of Elements of a State Water Monitoring and Assessment Program for Wetlands* (April 2006), are summarized as follows.

Landscape Level 1 Assessment: The Conservation Assessment and Prioritization System (CAPS) is a landscape level model that predicts ecological integrity based on GIS-derived metrics representing stressors in the landscape.[1] The CAPS output has been rigorously verified by UMass based on a substantial amount of taxa abundance data collected to date using statistical calibration based on fitted regression models and maximum likelihood methods to predict the value of the stressor metrics.

Level 2/3 Site Level Assessment Method (SLAM): SLAM development began in 2006 and first focused on forested wetlands because they are the wetland type with the most alteration in MA, and on Salt Marshes because of MCZM experience during sampling over the past decade. To date a total of 256 forested wetland sites have been sampled in the Chicopee, Millers, Concord and Taunton River Watersheds and 175 salt marsh sites have been sampled along the entire MA Coast. Through a process of testing and verification field data have been used to determine whether CAPS IEI and the individual CAPS stressor metrics (e.g. habitat loss, connectedness, etc.) are related to ecological condition and to quantify those relationships through the development of IBIs. Data previously collected by MassDEP from 490 sites were used to develop IBIs for wadable streams. Using data from intensive sampling methods we expect to create simplified sampling methods (RAMs) of known effectiveness. For example, results from multi-taxa assessments of forested wetlands are being used to create robust IBIs that depend only on collection of plant data.

Continuous Aquatic Life Use (CALU): Our assessment approach is based on the relationship between IEI (i.e. constraints on biological condition from the surrounding landscape) and IBI (i.e. actual condition of a site based on field assessments). We describe this relationship as the Continuous Aquatic Life Use (CALU) approach because both the IEI and the IBIs yield scores that are continuous throughout their range and on the same scale so it is not necessary to create tiers like the traditional Tiered Aquatic Life Use (TALU) model.

[1] For information on the role of CAPS as part of the Massachusetts Wetlands Monitoring and Assessment Program go to: http://www.umasscaps.org/applications/wetlands-assessment.html. Related reports are available at: http://www.umasscaps.org/docs_reports/index.html.

Integrating the landscape-based assessment (Level 1) and the site level assessments (Level 2/3) is at the core of the Massachusetts Monitoring and Assessment Program. The landscape-based assessment produces an Index of Ecological Integrity (IEI) as a means of scoring a wetland's context on a scale that is comparable to the Generalized Stressor Gradient. Site-based assessments produce Indices of Biological Integrity (IBI) on the same scale as IEI and are comparable to the Biological Condition Gradient. Work funded by this Grant Cooperative Agreement furthered this approach through 1) development of additional IBIs for salt marshes and forested wetlands and making them easier to implement, and 2) modeling nutrient loading for lakes and their associated wetlands as a step toward developing a comprehensive nutrient loading metric for use in CAPS.

Further development and testing of IBIs for salt marshes, forested wetlands, and wadable streams, and creation of IBI software

REVISE IBIs BASED ON DATA FROM ADDITIONAL FIELD SAMPLING

Since our original IBI analyses in 2013 additional field data collection increased the number of salt marsh and forested wetland sites available for analysis (Table 1). For salt marshes the additional sites came from the same geographic area as the original analyses. However, the additional forested wetland sites came from an expansion in geographic scope to include sites in the Taunton River watershed (the other watersheds previously included are the Miller's, Chicopee and Concord River watersheds).

Table 1. Number of sites included in the 2013 and 2015 IBI analyses.

Wetland Type	2013 Analysis	2015 Analysis
Forested Wetland	214	250
Salt Marsh	130	164
Wadable Stream	490	490

The analyses presented in this report were similar to those conducted in 2013. However, the application of IBIs developed in one portion of the state (Miller's, Chicopee and Concord watersheds) to a different part of the state (Taunton watershed) revealed some weaknesses in the previous approach and led to changes in the analyses used this time around.

IBI Methodology[2]

We developed separate IBIs for each major taxonomic group (e.g., vascular plants, macroinvertebrates) and stressor metric in each ecological system. The development of separate IBIs for each taxonomic group reflects a practical concern over the comparative costs and benefits of collecting and identifying different taxa. Having separate IBIs for different taxonomic groups and stressor metrics also affords us great flexibility in using the observed biotic condition to indicate the nature of the stressor(s) affecting the system. Table 2 lists the various taxonomic groups and sampling technique used in the IBI analyses along with the number of sites and taxa sampled.

[2] This is an abbreviated summary of our IBI methodology highlighting changes made to the original approach. For more detail on the IBI analyses see the 2013 report "E**Error! Main Document Only.**mpirically Derived Indices of Biotic Integrity for Forested Wetlands, Coastal Salt Marshes and Wadable Freshwater Streams in Massachusetts" available at http://umasscaps.org/pdf/CAPS%20IBI%20Report%20Sept%2015%202013%20Final.pdf.

Table 2. Number of sites (N) and number of taxa sampled by ecological system and taxonomic group from 1984 through 2012 for the purpose of developing Indices of Biotic Integrity (IBIs). The number of taxa include the number of separate taxa across taxonomic levels from Species to Phylum that were considered in the development of the IBI and only includes taxa that occurred at 10 or more sites. Similarly the number of sites includes only sites that were used to make the IBIs; some sites are not tallied here if they were sampled for some but not all of the taxonomic groups used by the IBIs.

	Ecological System					
	Forested wetland		Salt marsh		Wadable streams	
Taxonomic group	N	Taxa	N	Taxa	N	Taxa
vascular plants	250	401	167	45	--	--
bryophytes	250	123	--	--	--	--
macroinvertebrates						
quadrats	--	--	167	37	--	--
D-net sweeps	--	--	167	48	--	--
auger	--	--	167	37	--	--
kick nets	--	--	--	--	490	294

Step 1. Taxonomic data summary

The first step involved summarizing the species abundance data at each site. For each site, we created counts of each taxon's abundance at each taxonomic level, including Species, Genus, Family, Order, Class and Phylum. This means that an individual in a sample identified to Species was counted again at the Genus level and, depending on the taxonomic group, the Family, Order, Class and Division/Phylum levels as well. If an individual was only identified to Order, then it was only counted at the Order or higher level. We treated the abundance of each taxon at each taxonomic level as a separate dependent variable in the regression models below, and treated abundance as a Binomial response with a trial size equal to the total specimen count and/or as an unbounded Poisson response (with an offset to account for sampling effort), as appropriate. As one of several measures to safeguard against model overfitting, given the generally large number of taxa relative to the number of sites, we dropped all taxa that were observed at fewer than 10 sites.

Step 2. Regression

The second step was to fit individual responses for each taxon. Specifically, we modeled the relationship between each taxon (dependent variable) and each stressor metric (independent variable) with two functional forms and eight error models. The three-parameter logistic function (Equation 1) allowed for threshold responses of taxa to the gradient (note, the third parameter allows the upper asymptote to exceed one) while the constrained quadratic exponential (Equation 2) allowed for Gaussian and exponential responses to the gradient.

$$y_i = \frac{a}{1+be^{-cx_i}} + error_i \tag{1}$$

$$y_i = e^{(a+bx_i+cx_i^2)} + error_i \tag{2}$$

where y_i = the abundance of a taxon at the ith site, x_i = the value of the stressor metric at the ith site, $error_i$ = the error associated with the prediction at the ith site, and *a*, *b*, and *c* are parameters to be estimated. Note, in Equation 2 we constrained *c* to always be negative to prevent U-shaped distributions (i.e., where abundance peaks at low and high levels of the metric and is lowest in the middle) which we deemed ecologically implausible. Depending on the values of the parameters *a*, *b*, and *c*, these two functional forms can take on a wide variety of shapes, including monotonically increasing or decreasing, unimodal and sigmoidal curves, that represent plausible alternatives for how species' might respond to anthropogenic stressor gradients.

With four to eight suitable error models and two functional forms, we had 8-16 alternative models for each taxon. However, we dropped any model from further consideration if any of the following conditions were met: 1) the model failed to fit; 2) the delta AIC of the model was greater than 10; or 3) the fit predicted negative abundance (unrealistic) or abundance that was more than twice the maximum observed in the training data (these were often fits that behaved strangely at extreme values of the independent variable). For all retained models, we used AIC model weights to estimate the relative quality of each of the models based on how many parameters they had and how well they fit the data (Burnham and Anderson 2002). Note, we did not average these models at this step, but left that for the next step associated with statistical calibration, as described below.

Step 3: Statistical calibration

The third step involved the procedure known as statistical calibration (Jongman et al. 1995). Calibration involves using the estimated parameters (*a*, *b*, and *c*) from the regression in step 2 and the observed value of the dependent variable (y_i), and estimating the value of the independent variable (x_i) – essentially, regression in reverse. Specifically, we used the fitted models from step 2 to predict the log-likelihood of different values of the stressor metric at each site based on the abundance of taxa. The result is a log-likelihood curve that indicates the relative probability of the stressor metric being any particular value given the observed abundance of the taxon at a particular site. We generated log-likelihood curves for each site from the 8-16 different statistical models and then averaged them based on the AIC weights to make a single log-likelihood curve for each site and taxon.

We performed steps 2 and 3 on 20 cross-validation groups; in each group a different 5% of the sites was omitted and thus withheld from the model fitting process in step 2. In step 3, the stressor metric value of each site was then predicted for each taxon based on the models from which the site was omitted. In this manner, no site was simultaneously used for both model fitting (in step 2) and model prediction (in step 3). This cross validation is used to select taxa and was not fully external to the model process; we are therefore calling it an inner cross validation. This inner cross validation helps identify taxa (step 4) for which the model fits are more robust to new or different data but we no longer believe that it gives an accurate assessment of model performance.

Step 4: Taxa selection

The fourth step involved selecting the group of taxa that produce the most accurate predictions. Specifically, we added together the log-likelihood curves of individual taxa from step 3 to make a prediction for the site based on multiple taxa; the value of the stressor metric with the maximum log-likelihood was the predicted metric value for the site. We compared the performance of two different procedures for selecting taxa before selecting a preferred method.

<u>Method 1</u>.--In this method, we used a *stepwise* procedure to select the taxa, starting with the taxon that, by itself, produced the most accurate stressor metric (cross-validated) prediction based on the coefficient of concordance (Lin 1989, 2000) and then incrementally added the taxon that increased the concordance correlation coefficient of the (cross-validated) prediction the most; i.e. the conditional improvement in concordance. The concordance coefficient measures the agreement between the observed value of the stressor metric and our predicted value; a perfect concordance correlation of one occurs when the points fall on a perfect diagonal line with an intercept of zero and slope of one. Note, while the final IBI for field application was constructed using the full dataset (i.e., without cross-validation) for model fitting and calibration, the taxa were always selected based on the cross-validation procedure to avoid the erroneous selection of taxa that were overfit to the dataset. Unless otherwise noted, we report the cross-validated coefficients of concordance.

One of the challenges we faced was determining when to stop in the forward stepwise taxon selection process. As a third hedge against model overfitting and as a means of determining how many taxa to retain in the final IBI, we tested the significance of each taxon's fit against pseudotaxa, as follows. We created 1,000 pseudotaxa by permuting the data from the original taxa. For each pseudotaxa, we performed the same model fitting (step 2) and calibration (step 3) as the real taxa. Then during taxon selection (step 4), we compared each selected taxon's improvement in fit (i.e., concordance correlation) to the improvement in fit garnered by each of the 1,000 pseudotaxa to estimate the significance of the improvement in fit of each taxon. We used this significance test to decide how many taxa to include in the final prediction set; we included all taxa in the stepwise process up until the first taxon that didn't produce a significant increase in prediction accuracy, where significance was evaluated at the 0.05, 0.1 and 0.2 alpha levels. Lastly, for comparative purposes, we also continued the stepwise selection process until the maximum concordance was realized.

<u>Method 2</u>.--In this method, we used the *marginal* significance of each taxon based on the comparison to the 1,000 pseudotaxa, as described above. Specifically, for each taxon we computed the (cross-

validated) coefficient of concordance and compared it to the distribution of concordances of the pseudotaxa (i.e., the distribution of expected concordances by chance alone). We computed a *p*-value for each taxon by determining the proportion of the pseudotaxa distribution of concordances greater than or equal to the observed concordance for each taxon. We included all (marginally) significant taxa in the IBI, where significance was evaluated at the 0.05, 0.1 and 0.2 alpha levels. In this method we simply included all taxa with significant marginal concordances; whereas in the previous method we included taxa in a stepwise process based on their conditional improvement in concordance.

A major challenge faced with either taxa selection method is determining which taxa to include in the pool available for selection. Because we fit models to many different taxa (e.g., vascular plants, bryophytes and macroinvertebrates) depending on the ecological system, and at multiple taxonomic levels, we had many options. While our approach is amenable to the selection of any available taxa at any taxonomic level, for practical reasons we opted to create a limited set of IBIs that balanced performance and ease of data collection. In the prior round of IBI development we focused on making sure that we had separate IBIs for each taxonomic group as well as for all the possible groups combined. In this round we focused on IBIs that combined several of the most effective groups. In forested wetlands we chose to build IBIs out of the combined vascular plant (the most effective single group in the prior analysis) and bryophytes (a group that was moderately effective but relatively easy to collect and analyze). For saltmarshes we made IBIs that used all the taxa (plants and macroinvertebrates). In salt marshes we also tried treating invertebrate larvae as separate taxa from adults but ultimately dropped this approach as it reduced performance. In streams we continued to build IBIs using macroinvertebrates.

Our original methodology used a conditional 0.1 threshold for taxa inclusion. We observed poor performance of these IBIs when applied to the Taunton data and reconsidered the approach for taxa selection and performance evaluation. In the original methodology we evaluated the performance of the IBIs using the same cross validation data that were used to select the taxa. It appears that this biased the performance estimate upwards. In this round we use an external cross validation to evaluate model performance which we believe is unbiased. Further, using an unbiased estimate of performance led us to choose the marginal 0.2 criterion for selecting taxa as it outperformed the conditional 0.1 which we had previously used.

Step 5: Outer cross validation

Previously our randomized testing in which we conducted the entire process on pseudospecies revealed that low concordances were achieved with psuedospecies alone. This showed that the inner cross validation that was used to select species was not a true independent estimate of performance. In this round we decided to directly estimate performance by conducting the whole IBI fitting process in steps 1:4 within a four-fold “outer” cross validation. In this outer cross validation we divided the data in four groups and repeatedly conducted all of the prior steps on each combination of 3 groups and then predicted the fourth group. Thus the holdout data in this cross validation were used only to evaluate the entire IBI process. One side effect of this change is that our estimates of performance have decreased from prior estimates that used the inner cross validation. When we apply the IBIs to new data we used models fit to the entire dataset.

IBI Results

Tables 3, 4 and 5 present the results of our IBI analyses. We used an R-squared value of 0.5 as the cut-off for choosing IBIs that we considered credible enough to use; these are the IBIs included in the IBI software. R-squared values below 0.5 may be statistically significant but the relationships are weak enough that we believe they are of limited usefulness. In this run 16 IBIs were constructed for wadable stream of which ten were credible ($R^2 > 0.5$). Five of 15 IBIs constructed for forested wetlands had R^2 values above 0.5. None of the 22 IBIs constructed for salt marshes were considered credible. The best performing salt marsh metric (connectedness) had an R^2 value of 0.32; IEI (the second best metric) had an R^2 of 0.27. As a result none of the salt marsh IBIs were incorporated into the IBI software.

Table 3. IBIs constructed for salt marshes (IBIs in bold are considered credible, R2 > 0.5)

Metric	Datasets	R-Squared
Connectedness	Macroinvertebrates and plants	0.32
IEI	Macroinvertebrates and plants	0.27
Salt marsh ditching	Macroinvertebrates and plants	0.26
Connectedness	Macroinvertebrates	0.25
IEI	Macroinvertebrates	0.21
Habitat loss	Macroinvertebrates and plants	0.16
Salt marsh ditching	Macroinvertebrates	0.14
Tidal restrictions	Macroinvertebrates and plants	0.12
Similarity	Macroinvertebrates	0.11
Edge predators	Macroinvertebrates and plants	0.09
Feral predators	Macroinvertebrates	0.09
Wetland buffer insults	Macroinvertebrates and plants	0.08
Habitat loss	Macroinvertebrates	0.07
Similarity	Macroinvertebrates and plants	0.07
Wetland buffer insults	Macroinvertebrates	0.06
Feral predators	Macroinvertebrates and plants	0.05
Road traffic intensity	Macroinvertebrates and plants	0.04
Edge predators	Macroinvertebrates	0.03
Tidal restrictions	Macroinvertebrates	0.03
Road traffic intensity	Macroinvertebrates	0.02
Mowing & plowing intensity	Macroinvertebrates and plants	-0.04
Mowing & plowing intensity	Macroinvertebrates	-0.04

Table 4. IBIs constructed for wadable streams (IBIs in bold are considered credible, R2 > 0.5)

Metric	Datasets	R-Squared
Imperviousness	**Macroinvertebrates**	**0.73**
Watershed habitat loss	**Macroinvertebrates**	**0.72**
Sediment loading	**Macroinvertebrates**	**0.71**
IEI.s*	**Macroinvertebrates**	**0.69**
Nutrient loading	**Macroinvertebrates**	**0.68**
Change in phosphorus	**Macroinvertebrates**	**0.65**
Road traffic intensity	**Macroinvertebrates**	**0.64**
Hydrologic alteration	**Macroinvertebrates**	**0.61**
IEI	**Macroinvertebrates**	**0.60**
Habitat loss	**Macroinvertebrates**	**0.58**
Connectedness	Macroinvertebrates	0.49
Edge predators	Macroinvertebrates	0.44
Aquatic connectedness	Macroinvertebrates	0.41
Change in nitrogen	Macroinvertebrates	0.38
Mowing & plowing intensity	Macroinvertebrates	0.19
Dam intensity	Macroinvertebrates	0.004

*IEI.s is IEI calculated for all streams considered as a single category

Table 5. IBIs constructed for forested wetlands (IBIs in bold are considered credible, R2 > 0.5)

Metric	Datasets	R-Squared
IEI	**Bryophytes & vascular plants**	**0.57**
Connectedness	**Bryophytes & vascular plants**	**0.56**
Habitat loss	**Bryophytes & vascular plants**	**0.52**
Edge predators	**Bryophytes & vascular plants**	**0.51**
Invasive earthworm	**Bryophytes & vascular plants**	**0.50**
Invasive plants	Bryophytes & vascular plants	0.50
Similarity	Bryophytes & vascular plants	0.49
Watershed habitat loss	Bryophytes & vascular plants	0.46
Nutrient loading	Bryophytes & vascular plants	0.44
Sediment loading	Bryophytes & vascular plants	0.39
Road traffic intensity	Bryophytes & vascular plants	0.34
Edge effect (microclimate alterations)	Bryophytes & vascular plants	0.33
Aquatic connectedness	Bryophytes & vascular plants	0.32
Road salt	Bryophytes & vascular plants	0.31
Wetland buffer insults	Bryophytes & vascular plants	0.25

Continuous Aquatic Life Use (CALU)

One of the uses of the CAPS derived IBIs is to evaluate wetlands to see if their biological condition is commensurate with the landscape in which they are located. Many IBIs are developed using reference sites and test (impacted) sites but not the full disturbance gradient. Tiers (as used in Tier Aquatic Life Use or TALU standards) are essentially a means for dealing with uncertainty when IBIs are not developed as dose-dependent relationships between biological condition and stressors. CAPS IEI scores are a continuous rather than binary approach for defining reference conditions used in the development of IBIs. When IBIs are developed to correspond to a continuous stressor gradient (consistent with the Biological Condition Gradient concept) then it is no longer necessary to have tiered criteria tied to specific Classes or Qualifiers.

CAPS provides an approach to the establishment of numeric criteria for aquatic life use that is consistent with TALU but eliminates the need to develop tiers. We call this new approach CALU for Continuous Aquatic Life Use standards. Because both IEI and IBIs yield scores that are continuous throughout their range it is not necessary to create Tiers or Classes for wetlands and water bodies in order to have meaningful criteria for aquatic life use.

The CALU approach is based on the relationship between IEI (representing the constraints on biological condition due to the nature of the surrounding landscape) and IBI, which represents the actual condition of a site based on assessments conducted in the field. By defining an acceptable range of variability around this relationship it is possible to create numeric criteria for biological condition (a range of acceptable IBI scores) based on each site's particular landscape context (IEI score). Figure 1 is a CALU plot for forested wetland IEI.

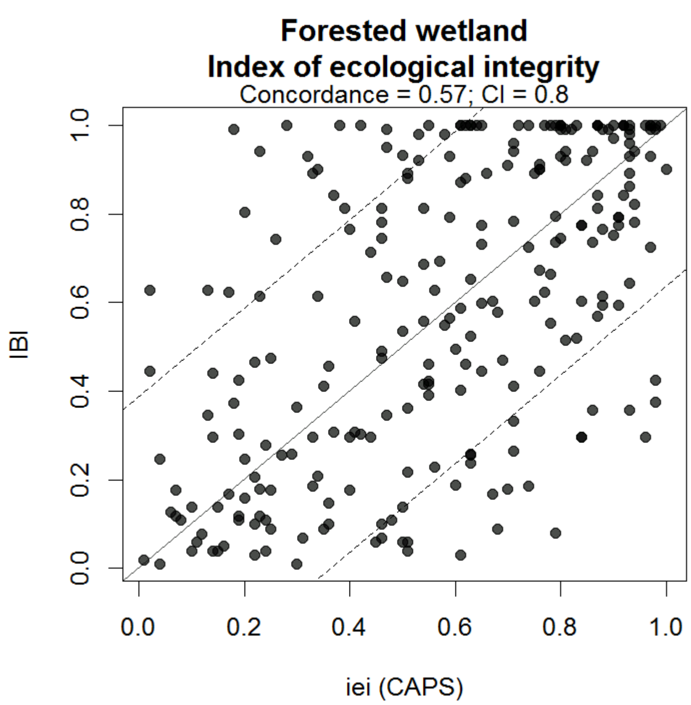


Figure 1. Continuous Aquatic Life Use (CALU) for IEI in forested wetlands

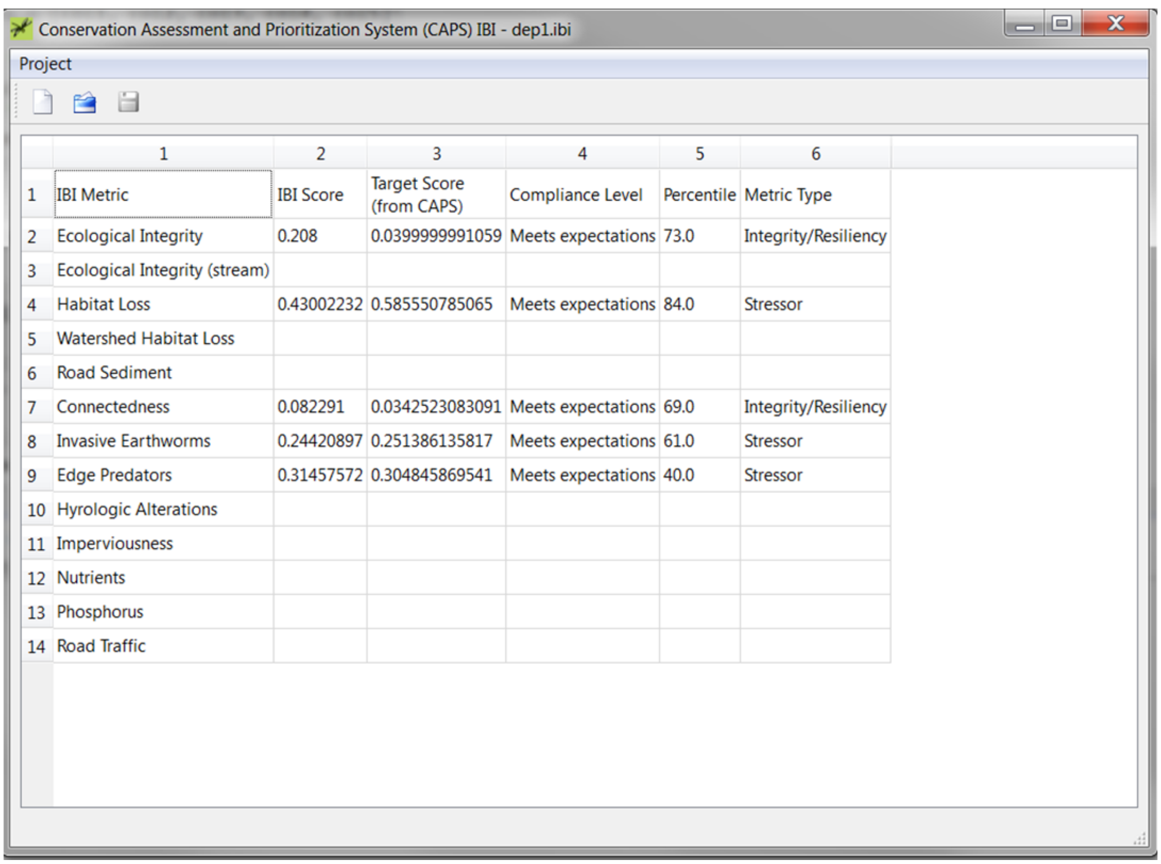
We have created CALU plots for all credible IBIs and the CALU concept is built into the IBI software. When users get results from the IBI software these results will include the percentile for each IBI value. This percentile is the basis for determining whether the site “exceeds expectations” (>90th percentile), “meets expectations” (10th to 90th percentile) or “fails to meet expectations” (<10th percentile).

CREATE SOFTWARE AND USER’S GUIDE FOR IMPLEMENTING IBIs

A new software program, CAPS-IBI was created to allow users to easily compute IBI scores on their own field data for forested wetlands and wadable streams. Because we were not able to construct credible IBIs from the salt marsh data this wetland type was not included in the software.

CAPS-IBI currently is intended to run primarily on Microsoft Windows operating systems (XP, Vista and Windows 7 and Windows 8), although files are available that make it possible for very experienced computer users to run the software on nearly any operating system. Hardware requirements include a minimum of 512MB of RAM and about 300MB of free disk space. CAPS-IBI is written in the Python programming language and makes use of the Quantum GIS API and the Inno Setup 5 installer program.

The software can be downloaded from this link: https://sourceforge.net/projects/capsibi/files/. The User’s Manual is included at Appendix A to this report. Figure 2 is a screen shot of the output from the CAPS-IBI software.



	1	2	3	4	5	6
1	IBI Metric	IBI Score	Target Score (from CAPS)	Compliance Level	Percentile	Metric Type
2	Ecological Integrity	0.208	0.0399999991059	Meets expectations	73.0	Integrity/Resiliency
3	Ecological Integrity (stream)					
4	Habitat Loss	0.43002232	0.585550785065	Meets expectations	84.0	Stressor
5	Watershed Habitat Loss					
6	Road Sediment					
7	Connectedness	0.082291	0.0342523083091	Meets expectations	69.0	Integrity/Resiliency
8	Invasive Earthworms	0.24420897	0.251386135817	Meets expectations	61.0	Stressor
9	Edge Predators	0.31457572	0.304845869541	Meets expectations	40.0	Stressor
10	Hyrologic Alterations					
11	Imperviousness					
12	Nutrients					
13	Phosphorus					
14	Road Traffic					

Figure 2. A screenshot from CAPS-IBI showing output from the analysis

USE OF CAPS-IBI SOFTWARE TO ANALYZE DATA FROM MassDEP ASSESSMENTS

The new forested wetlands IBIs have been applied to five mitigation sites that were permitted by MassDEP as variance projects. Their position on the CALU plot for IEI is shown in Figure 2. Full details of this assessment will be included in a wetland mitigation report currently in preparation. UMass personnel are working with MassDEP staff to use the CAPS-IBI software and the CALU approach to evaluate 40 forested wetland sites in the Chicopee River watershed.

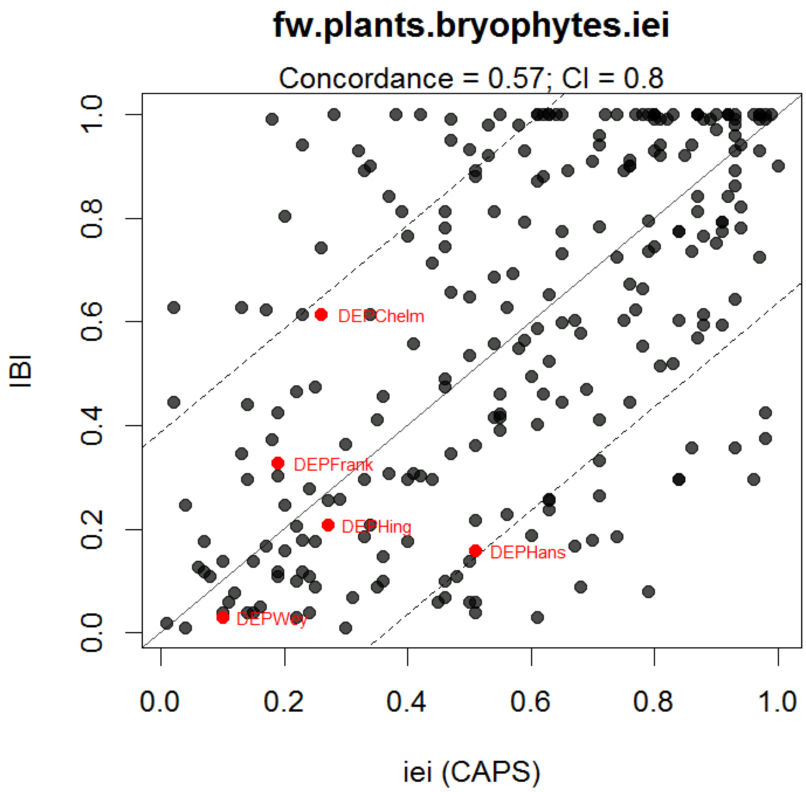


Figure 3. CALU plot for IEI in forested wetlands with five mitigation sites permitted by MassDEP as variance projects

Estimating alteration of lake phosphorus concentrations from anthropogenic basin alterations

SUMMARY

The goal of this study was to estimate for lakes in Massachusetts the degree of alteration of phosphorus concentrations associated with anthropogenic basin modifications. We developed multiple linear regression models, using principal component analysis to guide independent variable selection, to estimate current, altered phosphorus concentrations from a range of both natural and anthropogenic basin characteristics. Natural nutrient concentrations were then estimated by simulating basins with no alterations, allowing the difference between the natural and altered nutrient concentrations to be investigated. The hope is that we will be able to model not only nutrient concentrations in lakes but be able to relate those concentrations to the natural trophic status of those ecosystems. This would then serve as the basis for a stressor metric for inclusion in CAPS.

The significant variables in the final model for all lakes were cranberry bogs, septic systems, transportation, unpaved roads and waste disposal sites in watersheds, all correlated with increases in lake phosphorus concentrations. When only mainland (non-CAPE) watersheds are considered, unpaved roads and waste disposal sites were the anthropogenic basin modifications that were identified as significant variables in predicting lake phosphorus.

DATA

This study was conducted on lakes whose basins are completely within the boundaries of the Commonwealth of Massachusetts. Data for the lake total phosphorus (TP) concentrations, lake depth and lake surface area came from the National Water Quality Portal with samples from 1997 – 2010, and the Massachusetts Department of Environmental Protection (Mass DEP) Division of Watershed Management (DWM) WPP final water quality data for the 2005-2010 monitoring years. MassDEP data were received from Tom Dallaire at Mass DEP. The NWQp data were downloaded from http://www.waterqualitydata.us/ using the query shown below in Figure 4.

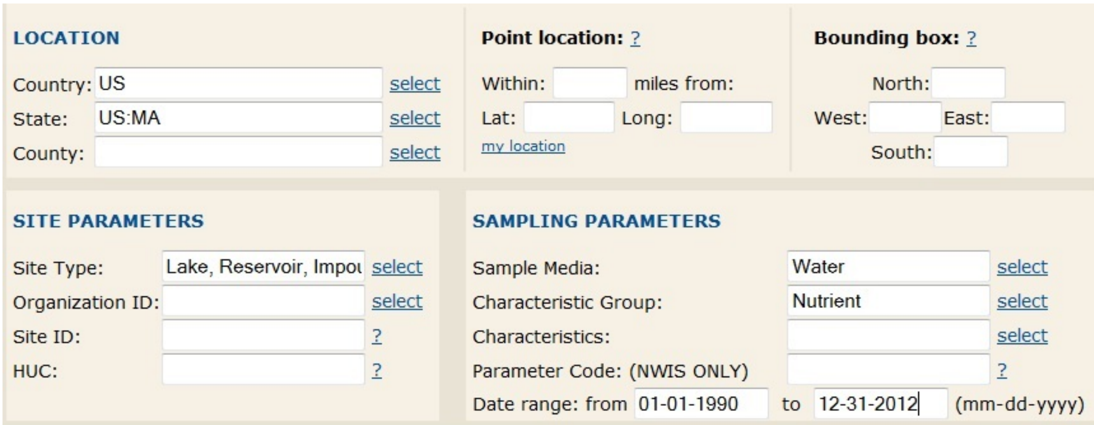


Figure 4 National Water Quality Portal query used to download nutrient data for MA lakes

Of the 81 lakes with phosphorus MassDEP sampling data from 2005-2010, the 53 shown in purple in **Figure** (below) also had surface area and max depth data. Of the 161 NWQp lakes with phosphorus sampling data from 1997-2010, 85 had surface area and max depth data. These 85 are also shown in green in **Figure** .

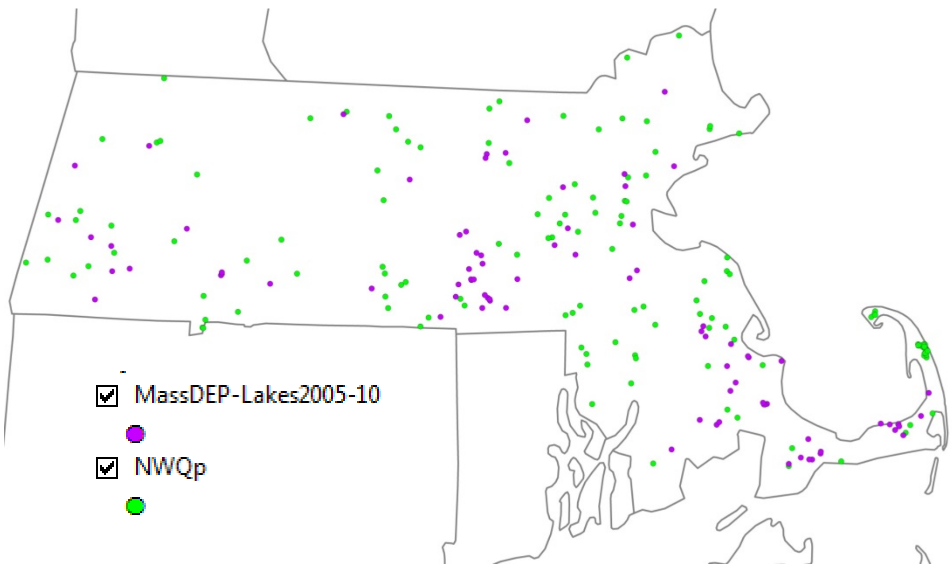


Figure 5 Locations of lakes used in this study

Maximum depth data for over 600 lakes in Massachusetts was provided by Marie Francoise Hatte at the Water Resources Research Center (WRRC) at UMass (http://wrrc.umass.edu/research/acid-rain-monitoring-project). The WRRC depth data included the PALSARIS, a unique numerical code for Massachusetts water bodies: 5 digits for lakes, 7 digits for streams with the first two digits signifying the watershed. This identifier is the same as PALIS used by MassDEP, and this ID was used to link the data sets. (Visual verification was also done by mapping all the lakes and depth data and confirming that locations matched.)

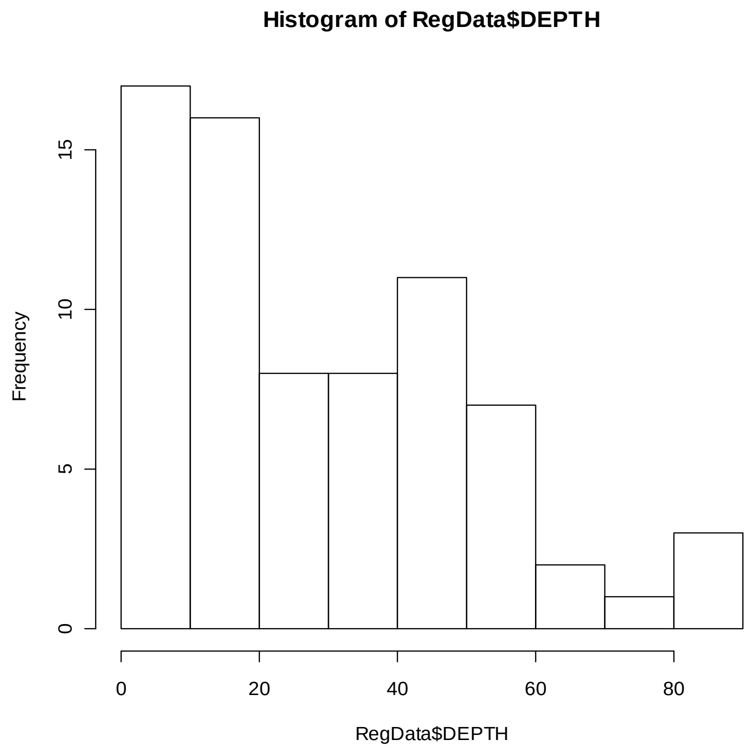


Figure 6. A histogram of the depths (feet) for the MA lakes selected for this study

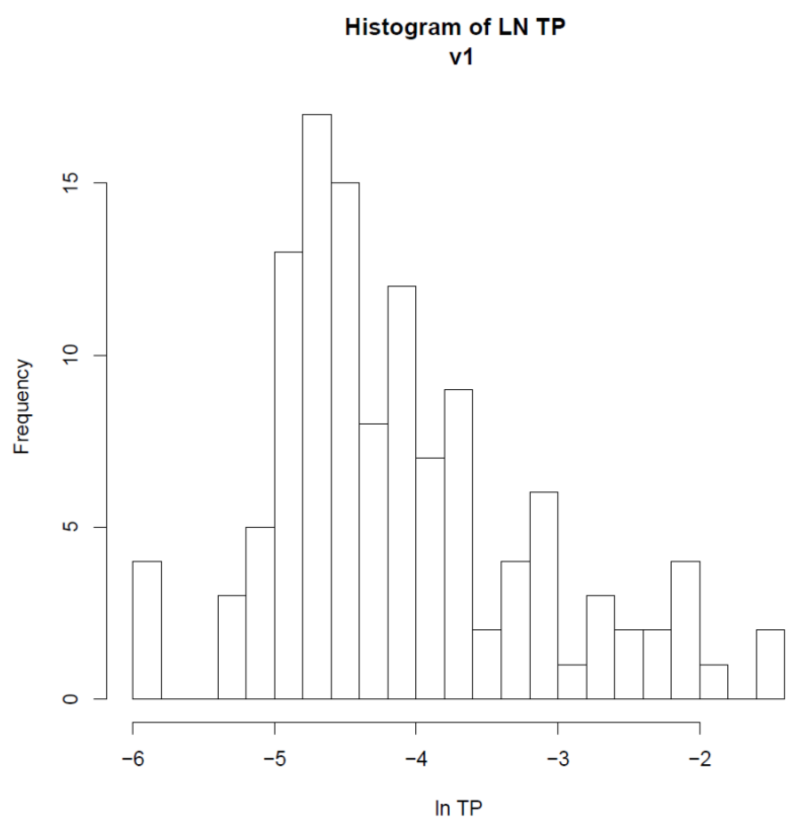


Figure 7. Histogram of ln(average TP) for the lakes in this study

Selected lakes had 1 – 18 measurements with a median of 2 samples per lake (Figure 8). After removing basins smaller than 0.5 km^2, and removing any lakes with data errors, there were a total of 127 lakes available for the analysis. The average phosphorus concentration for the lakes selected ranged from non-detect (at MDL of 0.005 mg/L) and a max of 0.233 mg/L. The mean of all the selected lakes was 0.030 mg/L.

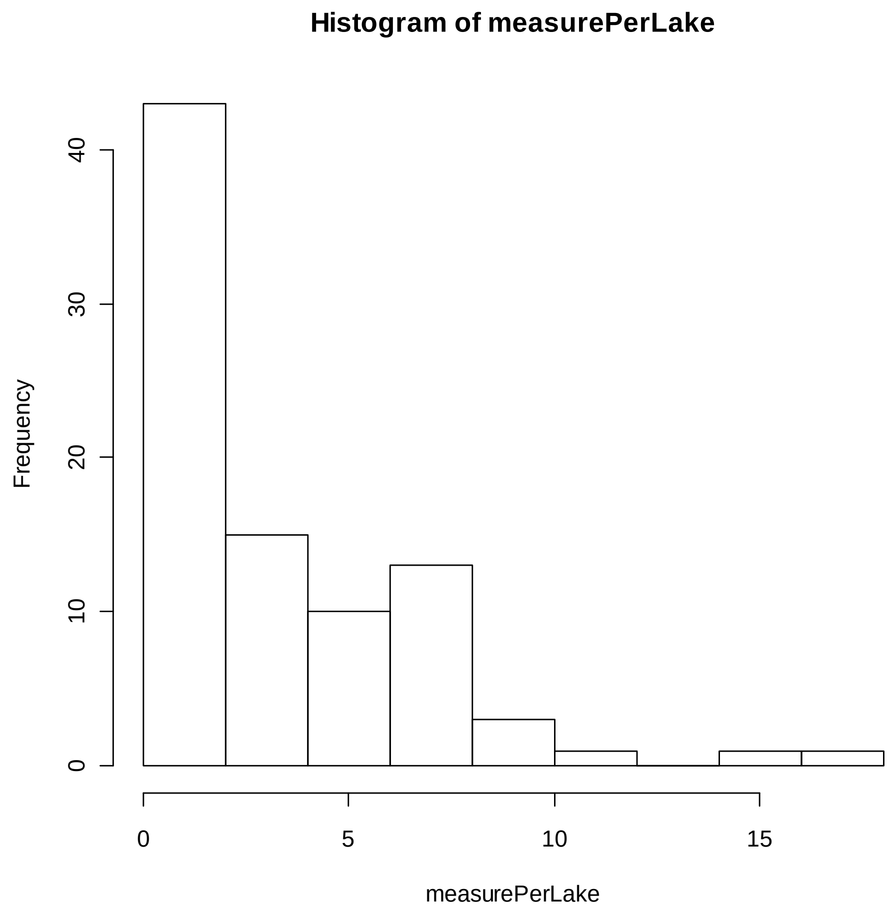
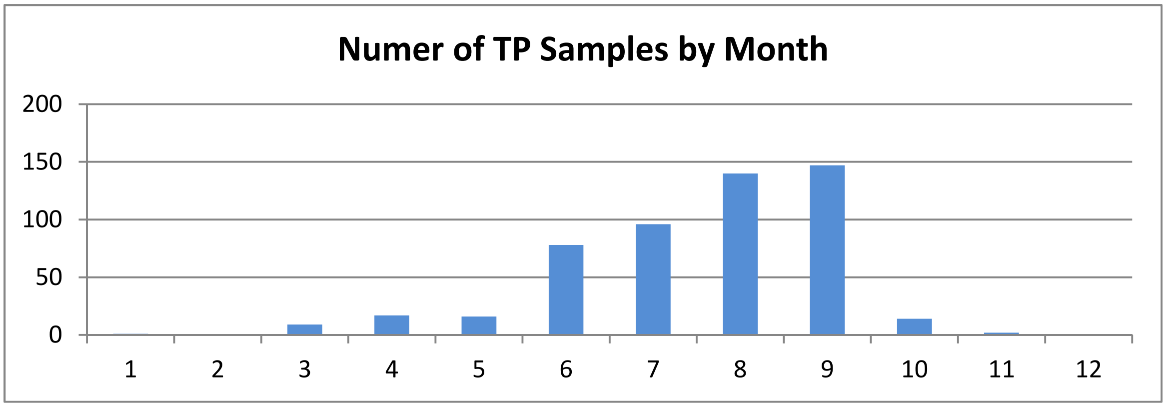


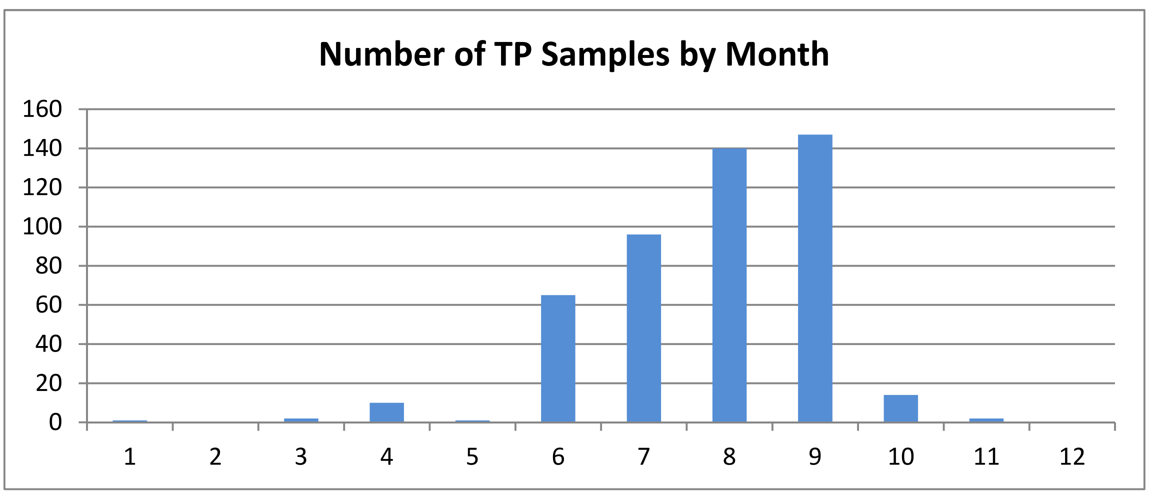
Figure 8. Histogram of the number of samples for each lake in the study



Month	1	2	3	4	5	6	7	8	9	10	11	12
# Samples	1	0	9	17	16	78	96	140	147	14	2	0

Figure 9. The number of phosphorus samples by month

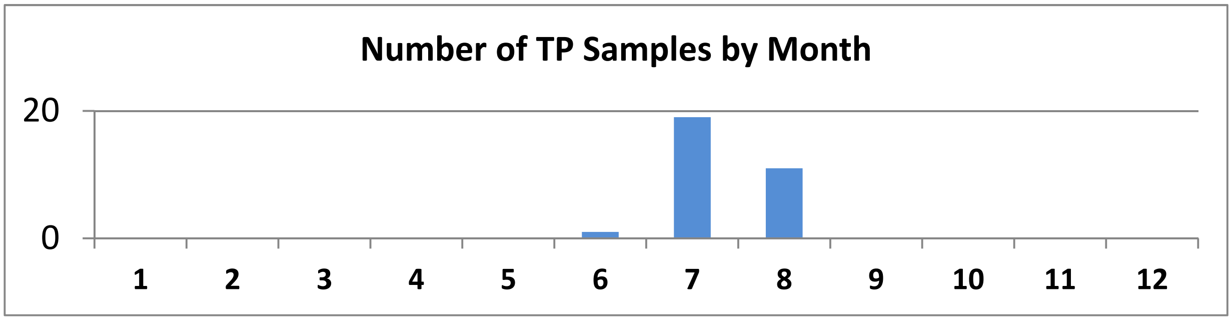
There were a total of 520 TP samples for the lakes selected, including 42 for Walden Pond. Most of the Walden samples were taken in the spring for a USGS lead study in 1997 – 2000. Because of the spring timing of all the samples, and the fact that Walden Pond has no surface inflows or outflows, it was removed from the set of lakes selected for this study.



Month	1	2	3	4	5	6	7	8	9	10	11	12
# Samples	1	0	2	10	1	65	96	140	147	14	2	0

Figure 10. The number of phosphorus samples by month, excluding samples at Walden Pond

All TP samples excluding Walden Pond comes to a total of 478. Most of the samples are taken in the months of June through September (Figure 10). Lakes with only one sample were sampled in June, July or August (**Figure**).



Month	1	2	3	4	5	6	7	8	9	10	11	12
# Samples	0	0	0	0	0	1	19	11	0	0	0	0

Figure 11. Months in which samples were taken for lakes with only one sample

The coefficient of variation (standard deviation / mean) was evaluated for the set of samples for each lake (see **Figure** 2 and **Figure 1**3.). Given the increased variability in samples for one lake when the hypolimnion samples were included, the decision was made to only include the phosphorus concentration in the epilimnion of lakes. All samples labelled as “Near bottom” (Mass DEP data) were removed. For samples without this label (NWQp data), a ratio was calculated of the sample depth to the maximum lake depth. A sample was labelled as “Near bottom” if the sample depth was > 60% of the maximum lake depth. The phosphorus concentration was then calculated as the average value of all measurements not “Near bottom”.

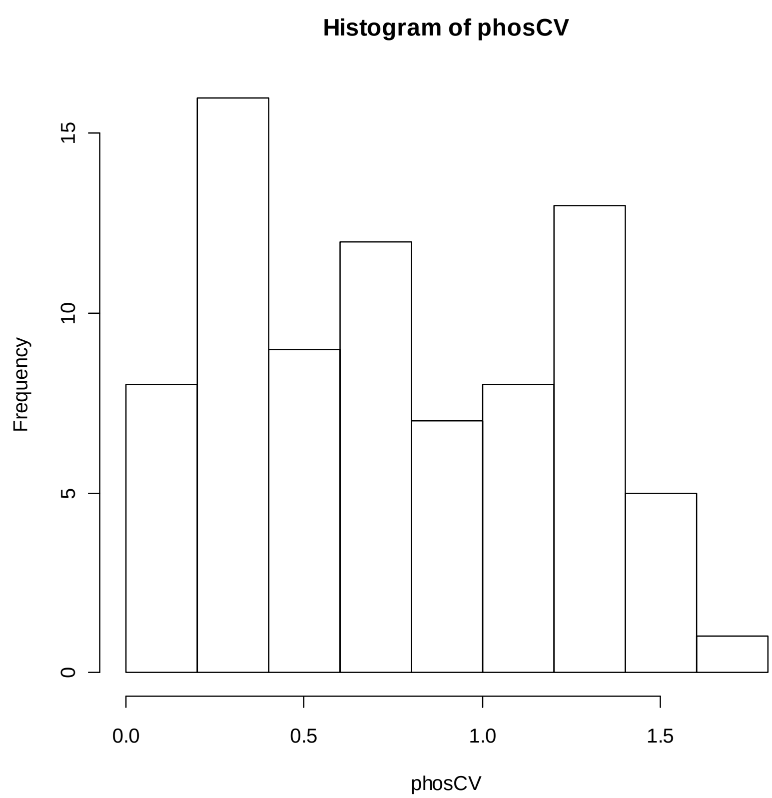


Figure 12. Histogram of the coefficient of variation of the phosphorus samples at all depths for each lake

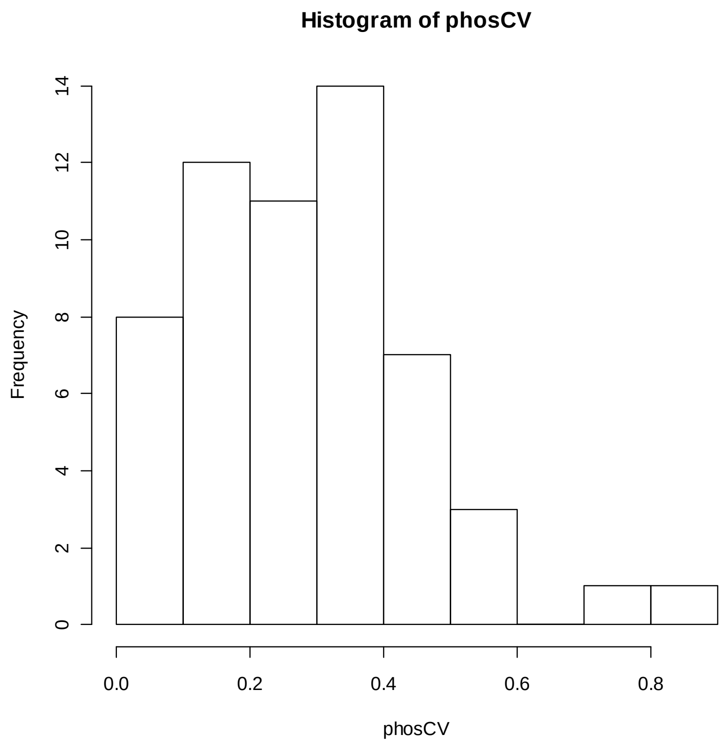


Figure 1 Histogram of the coefficient of variation of the phosphorus samples for each lake excluding those taken in the hypolimnion (or "Near bottom")

Basin Data

Tables 6, 7 and 8 list the basin characteristics that were used as potential independent variables in the regression model.

Table 6. Anthropogenic basin characteristics considered as potential independent variables in the regression model

Variable	Source	Range
Population	2010 census data	0.005 – 8.6
Discharge	2000-2005 in Mass DEP / SYE wateruse db	0 - 15027
Septic	1990 census data · % households on septic	1 - 1257
Imperv	Mass GIS 2006	0.8 – 36.7%
Land Use %†	CAPS Land Use	various

Table 7. Natural basin characteristics considered as potential independent variables in the regression model

Variable	Source
Basin Area (5 – 213 km^2)	CAPS delineation
Climate*	PRISM
Bedrock Lithology**	Mass GIS 2004, Group A
Land Use %†	CAPS Land Uses

†Land Uses: http://www.umasscaps.org/index.html

*Climate: Max Temp, Min Temp, Mean Precipitation from the PRISM 800m climate data for 1981-2010: http://www.prism.oregonstate.edu/products/matrix.phtml?vartype=tmin&view=maps

**Bedrock Lithology: http://www.mass.gov/anf/research-and-tech/it-serv-and-support/application-serv/office-of-geographic-information-massgis/datalayers/bedrock-lithology-.html. See Group A layer.

Table 8. Natural lake characteristics considered as potential independent variables in the regression model

Variable	Source
Lake surface area	Mass DEP, NWQp
Lake depth (max)	UMass WRRC

METHODS

A multiple linear regression model was used to estimate lake phosphorus concentration from the basin and lake characteristics.

$$\ln(c) = \beta_0 + \beta_1 ln(X_1 + 1) + \beta_2 ln(X_2 + 1) + \cdots \beta_{\mathrm{n}} ln(X_{\mathrm{n}} + 1)$$

where:

c	average nutrient concentration
X_i	basin characteristic i
β_i	model coefficient for basin characteristic X_i
n	the number of basin characteristics used as independent variables
i	1 to n

We employed a natural log transformation of both the independent and dependent variables. Natural log transformation is consistent with previous regression models relating basin characteristics to lake or stream nutrient concentrations (Sorrano 2008, Dodds and Oakes 2004). Alternative transformations were evaluated both numerically, by considering univariate correlations between variables, as well as visually, by examining plots of independent against dependent variables. No improvement was found when compared to the log-log model results.

Adding one to the value of the basin or lake characteristics before taking the natural log allows a correct mapping of boundary conditions between the value of the term ($\ln(X+1)$) in the linear regression model and the value of X, so that when $X = 0$, $\ln(X+1) = 0$. This approach allows for the correct representation of the 'removal' of anthropogenic modifications from the regression equation by setting the value of the corresponding terms to zero.

Because the natural log is only defined for values above zero, any basin characteristic with negative values was shifted to be nonnegative by adding the minimum value plus a small increment to all values. Only one required shifting (min temp) as the remaining natural variables were all non-negative.

We used principal component analysis (PCA) to select a subset of the 67 highly intercorrelated basin characteristics considered as potential independent variables in each regression. Variables with the highest eigenvector loadings, determined using a scree test, within intercorrelated sets of variables in the first set of components were maintained in the set of candidate variables. The variable reduction process involved determination of which characteristics had the highest univariate correlation coefficients with the dependent variable, and maintenance of variance inflation factors (VIFs) below 10, along with the analysis of the PCA loadings.

The 28 variables selected from the PCA were then input to stepwise regression algorithm, written in the statistical language R (R Development Core Team, 2006), which minimized the Akaike information criterion (AIC) to estimate the regression parameters and develop the final equations. One outlier was removed based on a high Cook's D (Kutner, 2005).

Weightings for the basin characteristics were calculated based on an 'aquatic distance' of each 30m square cell to the target point which in this case is the location of the sample taken for the nutrient concentration measurement. The 'aquatic distance' is calculated based on the land use of each cell traversed to reach the target point, whether the cell contains a stream channel, and the slope of the land surface of the cell. The approach is based on a method presented by Randhir et al., 2001.

Models with weighted and non-weighted basin characteristics were compared. Variables weighted by aquatic distance had significantly higher univariate correlations with the dependent variables and the models using the weighted basin variables resulted in higher adjusted R^2 values, so only weighted variables were included in the final models.

RESULTS

There are two sets of results shown here. Both include only summer, eplimnion samples for the study lakes. The first set of results includes all lakes and the second excludes lakes in the Southeastern part of Massachusetts and all of Cape Cod due to the domination of groundwater flow in those areas.

The first regression model including samples of all lakes had 11 variables and an adjusted R^2value of 0.64 (Tables 9 and 10).

Table 9. The number of parameters (p), adjusted R2, mean square error (MSE) and standard error (SE) for the model including all lakes

p	11
ADJR2	0.638
MSE	0.312
SE	0.650

Table 10. The model coefficients for the regression including all study lakes

Intercept	-0.775
septic	0.170
Transportation	0.760
Waste_disposal	2.206
Spectator_recreation	-0.707
Unpaved_road	0.150
Cranberry_bog	0.285
Forest	-0.583
Shrub_swamp	0.223
Water_lentic	-0.188
Water_lotic	0.267
DEPTH_IN_FEET	-0.566

Note in Table 10 that the sign is negative for the spectator_recreation variable, indicating that the model estimates an increase in spectator_recreation, weighted by aquatic distance to the lake, will decrease the lake phosphorus concentration. This is potentially an unexpected result, though our previous stream nutrient model showed the same behavior. The negative relationship between Forest and Water_lentic (lake water) is expected, as is the increasing lake depth being associated with a decrease in phosphorus.

The anthropogenic basin characteristics which were predicted to increase phosphorus are septic systems, transportation, waste disposal, unpaved roads and cranberry bogs. As we can see in **Figure** below, septic systems for the study lakes are estimated to have the most consistent impact on changing phosphorus concentrations. All of the study lakes are shown to have septic systems. Where the existence of the other alterations such as transportation or cranberry bogs are not in many basins, but some have a higher impact than the septic systems when they are there (Figure 15).

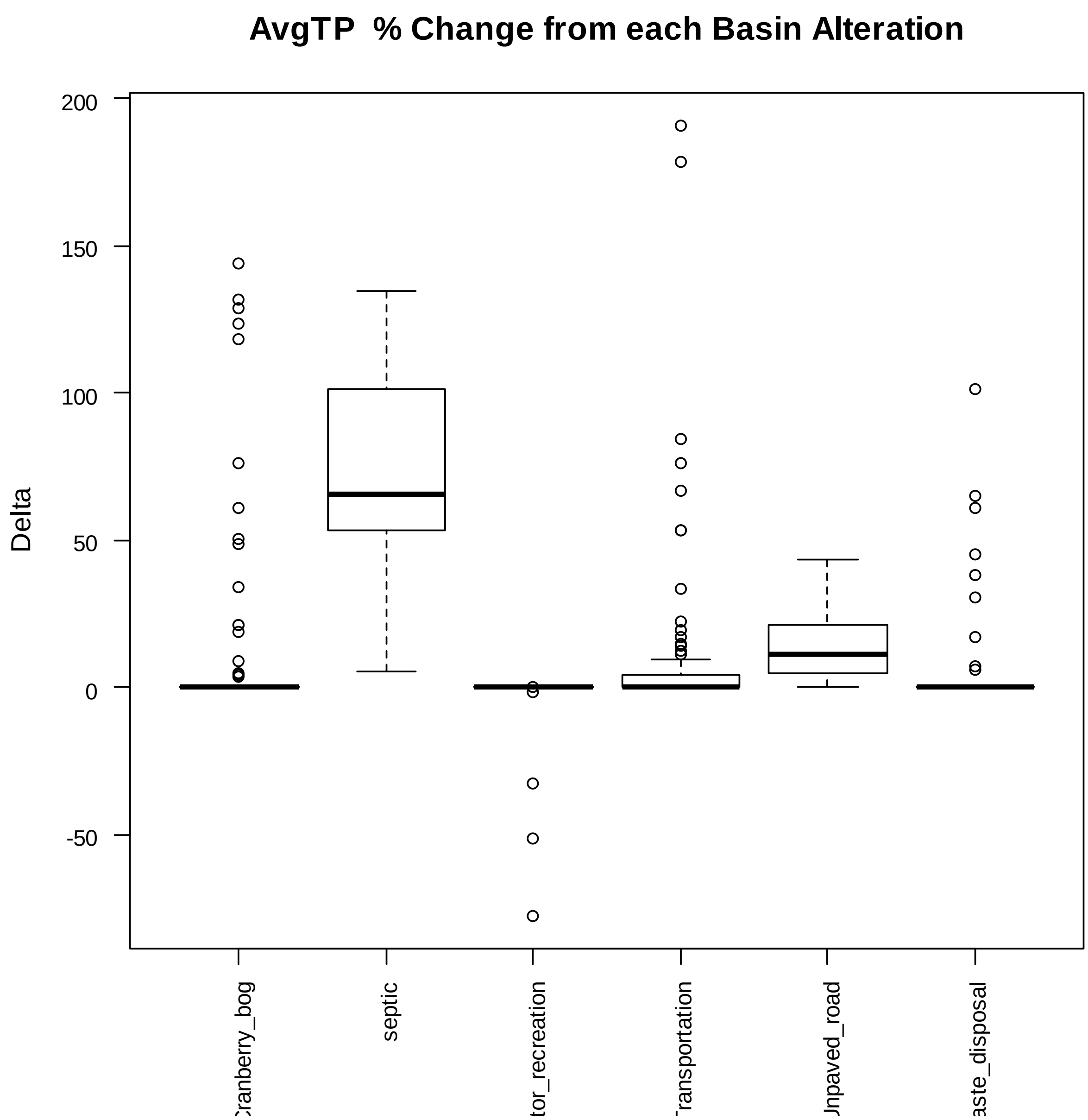


Figure 14. Percent change in TP due to each anthropogenic basin modification in the final model including all study lakes. Every study lake value of the variable is included here, even values of 0. Variables shown (from left to right) are cranberry bog, septic, spectator recreation, transportation, unpaved road, and waste disposal.

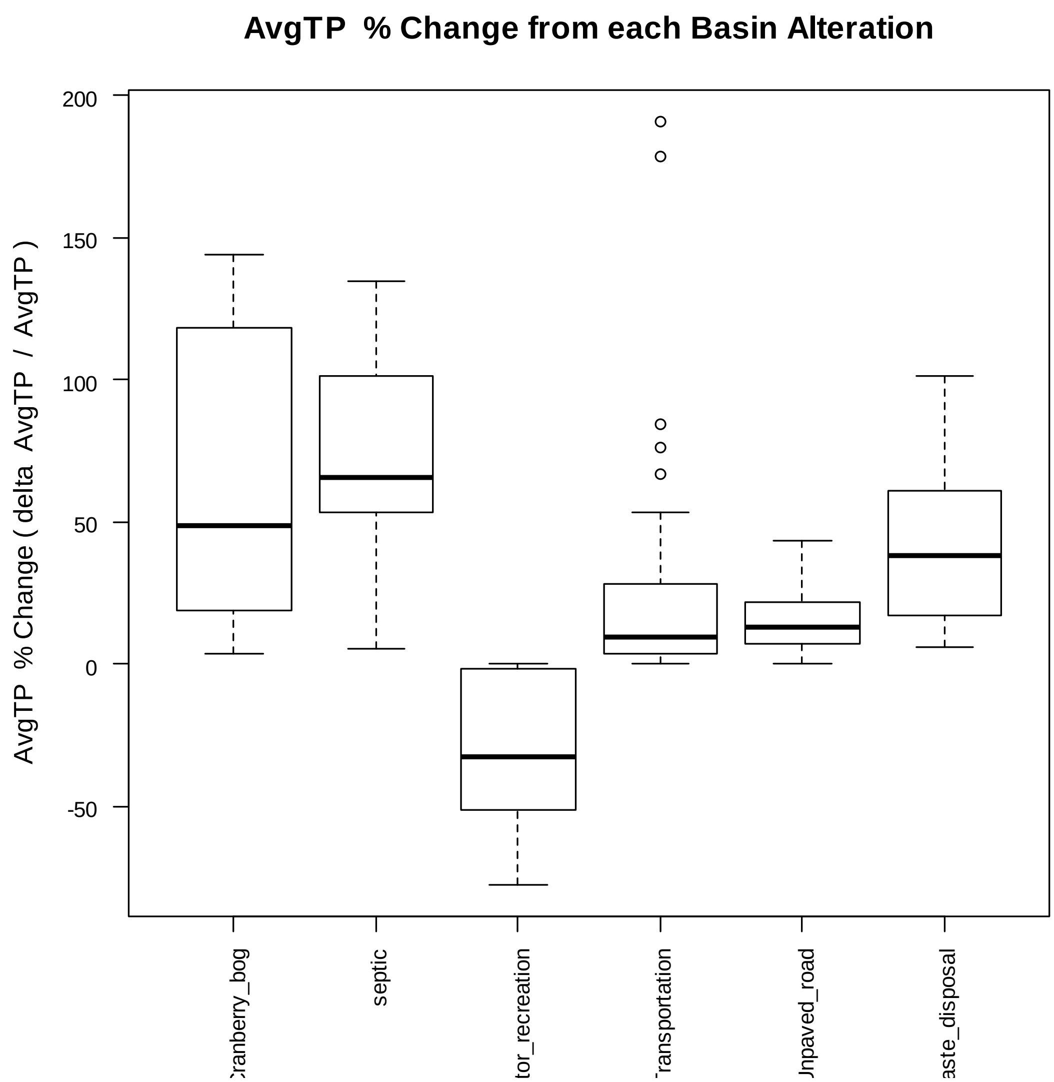


Figure 2 Percent change in TP due to each anthropogenic basin modification in the final model including all study lakes. Only non-zero values of the variables are included here. The ranges, therefore, show the impact on phosphorus for the basins that have those particular modifications. Variables shown (from left to right) are cranberry bog, septic, spectator recreation, transportation, unpaved road, and waste disposal.

The second regression model, for summer, eplimnion samples of lakes not including those in SE Mass and on Cape Cod had 9 final variables and an adjusted R^2 value of 0.70 (Tables 11 and 12).

Table 11. The number of parameters (p), adjusted R2, mean square error (MSE) and standard error (SE) for the model excluding lakes in SE and on Cape Cod

p	9
ADJR2	0.696
MSE	0.257
SE	0.579

The anthropogenic basin characteristics that were predicted to increase phosphorus are waste disposal, and unpaved roads.

Table 12. The model coefficients for the regression the model excluding lakes in SE and on Cape Cod

Intercept	-6.697
Waste_disposal	4.799
Spectator_recreation	-0.465
Unpaved_road	0.374
maxtemp	2.395
Forest	-0.605
Shrub_swamp	0.266
Water_lentic	-0.177
Water_lotic	0.384
DEPTH_IN_FEET	-0.638

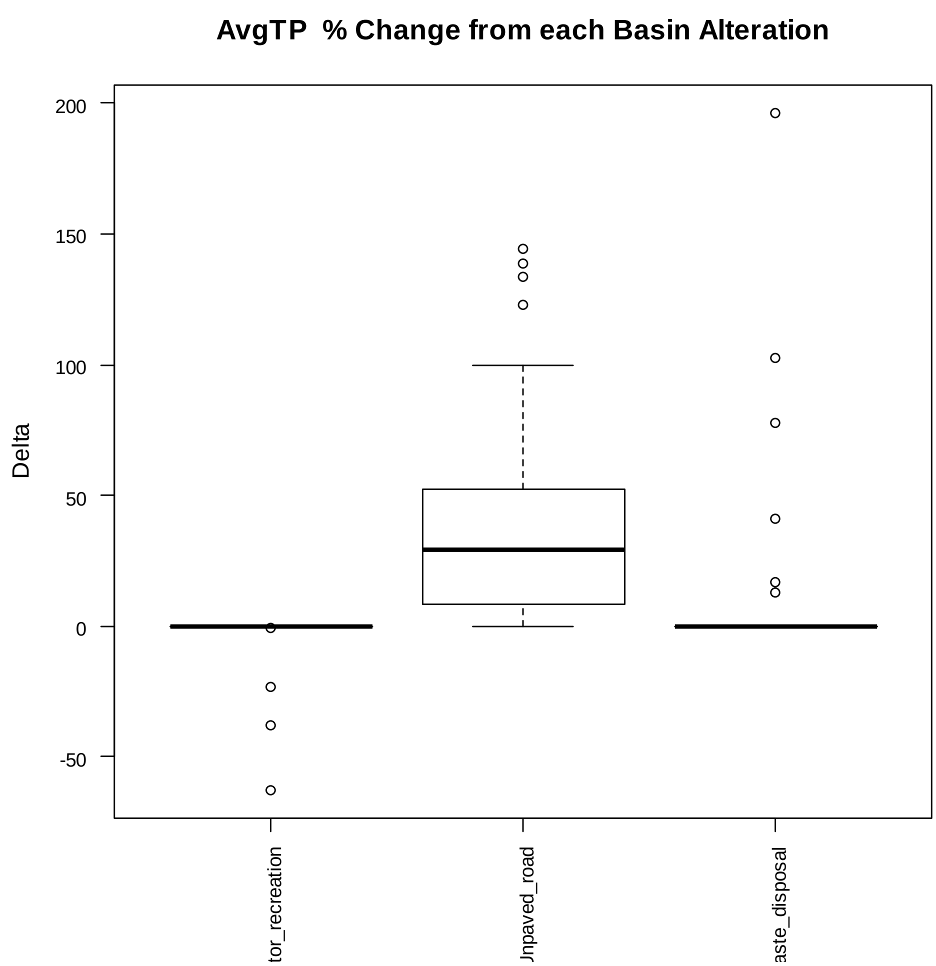


Figure 16.Percent change in TP due to each anthropogenic basin modification in the model excluding lakes in SE and on Cape Cod. Every study lake value of the variable is included here, even values of 0. Variables shown (from left to right) are spectator recreation, unpaved road, and waste disposal.

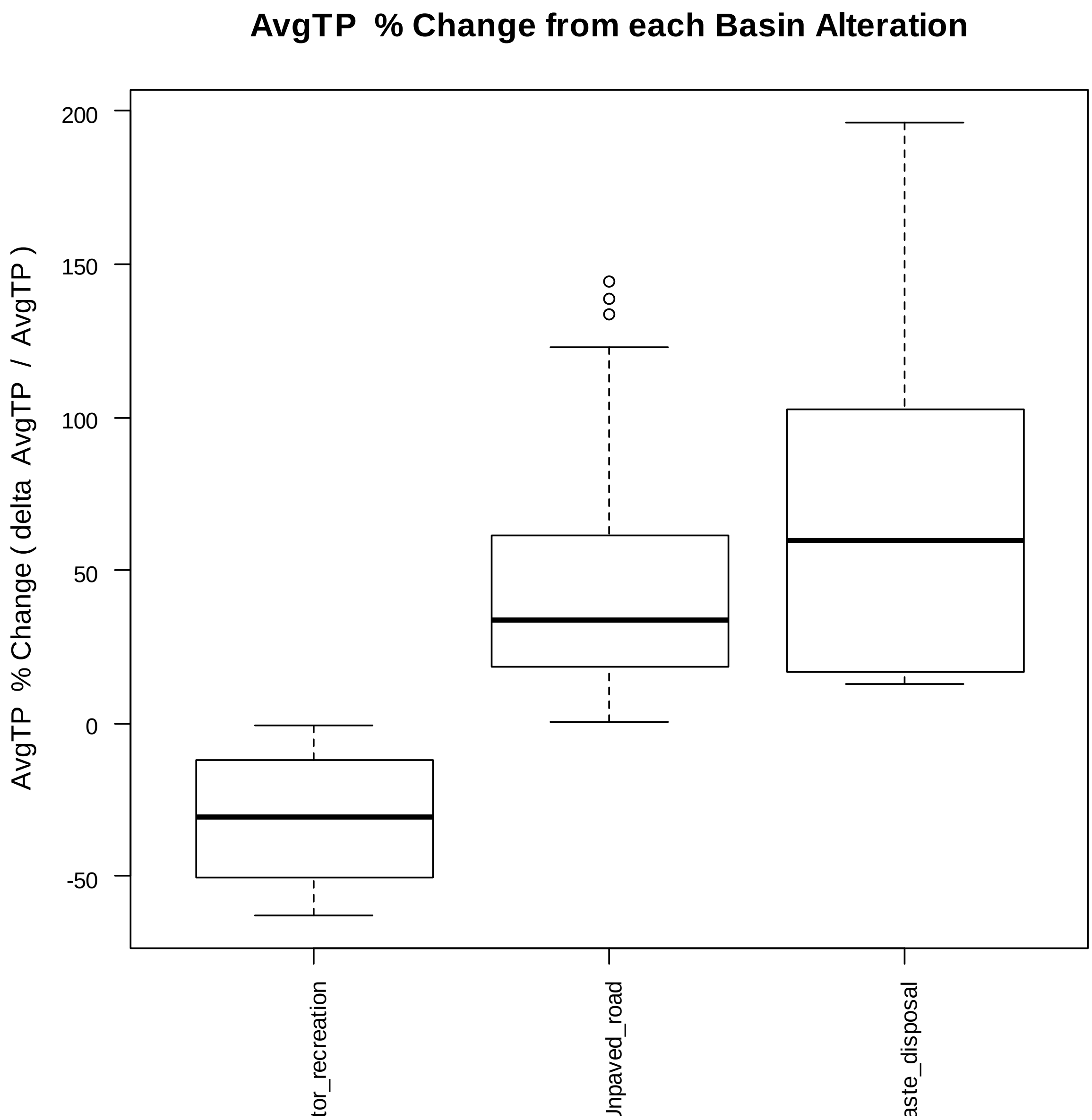
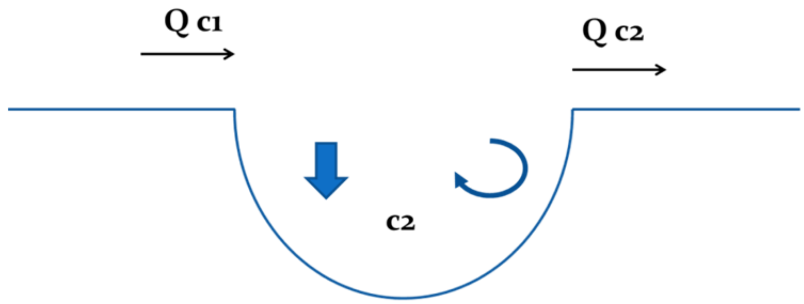


Figure 17. Percent change in TP due to each anthropogenic basin modification in the model excluding lakes in SE and on Cape Cod. Only non-zero values of the variables are included here. The ranges, therefore, show the impact on phosphorus for the basins that have those particular modifications. Variables shown (from left to right) are spectator recreation, unpaved road, and waste disposal.

AN ALTERNATIVE APPROACH TO CONSIDER



An alternative structure for the regression model could be considered which adds mechanistic knowledge based on a mass balance model of phosphorus concentration. The change in phosphorus concentration over time determined by the mass of phosphorus entering the lake (W), minus the mass flowing out of the lake (Qc) minus the mass of phosphorus settling out of the water column ($vA_s c$).

$$V\frac{dc}{dt} = W - Qc - vA_s c$$

In the steady state conditions, the change in phosphorus is 0 and inflows are equal to the outflows. So,

$$W = Qc + vA_s c$$

$$W = c(Q + vA_s)$$

Thus an alternative regression to our first approach could be to use the following structure:

$$c(Q + vA_s) \sim \beta_0 + \beta_{n1} ln(X_{n1} + 1) + \beta_{n2} ln(X_{n2} + 1) + \cdots$$

where:

c	average nutrient concentration
X_{nj}	basin characteristic j
β	model coefficients
Q	flow out of the lake
A_s	lake surface area
v	settling rate
W	nutrient loading rate

Note that this is different from previous regressions done on nutrient loadings estimated by inflow and stream concentration. The proposed model above instead uses outflow and in-lake concentration. Without flow data for all of the lake outflows the flow would need to be estimated.

Another opportunity to improve the model would be to collect data for dissolved oxygen for all the study lakes and note which lakes have hypolimnions that become anoxic (all the oxygen is depleted). These conditions can lead to the release of TP from the lake sediments and the sediments become a possible source instead of a sink for phosphorus.

More robust statistical methods could also be applied to learn about the relationships between the basins and phosphorus concentrations. Regression trees or neural networks should be investigated. In any case a cross validation should be run.

Additional research on alternate models for the aquatic distance calculation could be considered. It would be enlightening to compare the current model with a fixed riparian corridor, a flow path with two simple transport parameters such as proposed by Van Sickle (2008), or a more complex model of nutrient attenuation across the particular land use of each 30m cell.

Knowledge regarding ground water flows could be included in basin delineations if possible.

Additional lake nutrient data could be collected and possibly a model for estimating nitrogen from basin characteristics could also be developed.

NEXT STEPS

We have developing a model for evaluating the nutrient status of lakes. Our models for lake phosphorus had good R-squared values (0.64 – 0.7). In order to use these models to create a CAPS metric for lake phosphorus we will need comprehensive data on lake depths. Given the advances in remote sensing technologies these data might become available in the near future. There are also alternative approaches that could be explored for modeling lake phosphorus.

Our goal is to eventually be able to create a CAPS metric that doesn't just model phosphorus loading but evaluates the degree to which phosphorus concentrations have increase over natural conditions due to anthropogenic land use and waste disposal. The approach that we used for stream hydrologic alterations and stream nutrient status appears to be workable for evaluating phosphorus stress in lakes. Once we have more comprehensive data on lake depth we should be able to include phosphorus as a stressor metric for evaluating the ecological integrity of lake ecosystems in CAPS.

REFERENCES:

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications, 8*(3), 559-568.

Cohn, T. A., L. L. Delong, E. J. Gilroy, R. M. Hirsch, and D. K. Wells (1989), Estimating Constituent Loads, *Water Resour. Res., 25*, 937-942.

Dodds, W., & Oakes, R. (2004). A technique for establishing reference nutrient concentrations across watersheds affected by humans. *Limnology and Oceanography-Methods, 2*, 333-341.

Friesz, PJ and Colman, JA, (2001), Geohydrology and Limnology of Walden Pond, Concord, Massachusetts, Water-Resources Investigations Report 01-4137

Jones, K., A. Neale, M. Nash, R. Van Remortel, J. Wickham, K. Riitters, and R. O'Neill (2001), Predicting nutrient and sediment loadings to streams from landscape metrics: A multiple watershed study from the United States Mid-Atlantic Region*, Landscape Ecol., 16*, 301-312.

Keller, S., T. Q. Zhang, S. Webb, R. Brugam, K. Johnson, and Z. -. Lin (2008), Effects of suburban land use on phosphorus fractions and speciation in the Upper Peruque Creek, Eastern Missouri, Water Environ. Res., 80, 316-323.

Kutner, M. H. (2005)*, Applied Linear Statistical Models,* 1396 pp., McGraw-Hill Irwin, Boston.

Lottig, N. R., E. H. Stanley, P. C. Hanson, and T. K. Kratz (2011), Comparison of regional stream and lake chemistry: Differences, similarities, and potential drivers, Limnol. Oceanogr., 56, 1551-1562.

MADEP-Division of Watershed Management, QUALITY ASSURANCE PROGRAM PLAN Surface Water Monitoring & Assessment, 2005-2009

Moore, R. B., C. M. Johnston, K. W. Robinson, and J. R. Deacon. 2004. Estimation of total nitrogen and phosphorus in New England streams using spatially referenced regression models, U.S. Geological Survey Scientific Investigations Report 2004-5012. U.S. Geological Survey, Pembroke, New Hampshire, USA.

Randhir, T., R. O'Connor, P. Penner, and D. Goodwin (2001), A watershed-based land prioritization model for water supply protection*, For. Ecol. Manage., 143*, 47-56.

Ruddy, B.C., Lorenz, D.L., and Mueller, D.K., 2006, County-Level Estimates of Nutrient Inputs to the Land Surface of the Conterminous United States, 1982–2001: U.S. Geological Survey Scientific Investigations Report 2006-5012, 17 p.

Soranno, P. A., S. L. Hubler, S. R. Carpenter, and R. C. Lathrop (1996), Phosphorus loads to surface waters: A simple model to account for spatial pattern of land use, Ecol. Appl., 6, 865-878.

Soranno, P. A., Cheruvelil, K. S., Stevenson, R. J., Rollins, S. L., Holden, S. W., Heaton, S., et al. (2008). A framework for developing ecosystem-specific nutrient criteria: Integrating biological thresholds with predictive modeling. *Limnology and Oceanography, 53*(2), 773-787.

Van Sickle, J. and C. B. Johnson (2008), Parametric distance weighting of landscape influence on streams*, Landscape Ecol., 23*, 427-438.

Vogel, R. M., B. E. Rudolph, and R. P. Hooper (2005), Probabilistic behavior of water-quality loads*, Journal of Environmental Engineering-Asce, 131*, 1081-1089.

Appendix A: CAPS-IBI Software User’s Manual

APPENDIX B: Land Use Land Cover Variables

ALTERATIONS:

Commercial
Industrial
Urban open
Urban public
Transportation
Mining
Waste disposal
Junkyard
Multi-family residential
High-density residential
Medium-density residential
Low-density residential
Spectator recreation
Participatory recreation
Golf
Water based recreation
Marina
Cemetery
TOTAL URBAN

Cropland
Cranberry_bog
Nursery
Orchard
Pasture
TOTAL PLANTED

Powerline shrubland
Open land
TOTAL BARREN

NATURAL:

Forest
Forested_wetland
TOTAL FOREST
Shrub_swamp
Bog
Shallow_marsh
Deep_marsh
Vernal_pool
TOTAL LOWLAND

Water_lentic
Water_lotic
TOTAL OPEN WATER