Limnol. Oceanogr., 59(6), 2014, 2231–2239

doi:10.4319/lo.2014.59.6.2231

Response to comments: Quantification of the extent of cultural eutrophication of natural lakes in the United States

The purpose of our original study evaluating the extent of cultural eutrophication of U.S. lakes (Bachmann et al. 2013) was to test two basic assumptions used in the nutrient guidance manual of the U.S. Environmental Protection Agency (U.S. EPA 2000*a*) to set reference conditions for lakes in the various nutrient ecoregions of the country. We found that the extent of cultural eutrophication in both reference and non-reference lakes was not as great as assumed in the U.S. EPA recommended methods and that these findings should be taken into consideration when setting numeric nutrient criteria. Seemingly, we were not clear in placing our study in the context of the U.S. EPA methodology, so Smith et al. (2014) misinterpreted the conclusions of our study. Specifically, we did not state that cultural eutrophication was not real and that it could not be found in the United States. We also did not challenge the concept of nutrient criteria for lakes.

Smith et al. (2014) and McDonald et al. (2014) pointed out some problems with bias in our sample of lakes with good sediment cores and also had some good suggestions for improving our study, which we used to reanalyze our data. We found the new paleolimnological results apply directly to 60% of the natural lakes in the United States as defined by the probabilistic sampling model of the National Lakes Assessment (NLA) of 2007 (U.S. EPA 2010). We show that for this group of lakes on average there was no net increase in the diatom-inferred concentrations of total phosphorus (TP) or total nitrogen (TN) in going from the bottoms of the sediment cores, representing about 1850, and the tops of the cores, representing 2007. We also confirmed that for the reference lakes used in the various ecoregions of the NLA study there was no net increase in the diatom-inferred concentrations of TP or TN from the bottoms to the tops of the cores, thus calling into question the U.S. EPA's use of the 75th percentile of reference lakes for setting nutrient criteria or for evaluating the nutrient condition for lakes in the NLA study. Furthermore, we examined different ways of quantifying cultural eutrophication and how they are related to our examination of the methods used by the U.S. EPA to set reference conditions for TP and TN in the various ecoregions of the United States. Our conclusion remains that the methodology used to set the recommended reference conditions for the various ecoregions of the United States (U.S. EPA 2000*b*) likely overestimates the number of lakes that are in need of nutrient reductions based on pre-settlement conditions.

Recalculations

McDonald et al. (2014) correctly pointed out that we did not use the weights assigned to each lake as a part of the probabilistic design of the NLA study and also identified some problems in our interpretation of the codes used in the U.S. EPA spreadsheets to identify the lakes classified as probability and reference lakes. We reanalyzed our data with the corrected lists of lakes as indicated by McDonald et al. (2014) and also used the sample weights to find weighted averages for changes in TP, TN, specific conductance, and pH.

The sum of the weights for all 437 of the natural probability lakes in the NLA is 29,308, representing the population of natural lakes discussed in the NLA. The 233 natural probability lakes with good sediment cores represent 17,469 lakes and account for 60% of the total population of U.S. natural lakes in the NLA study rather than the entire population, because we do not have good cores for the other 40% of the lakes (Fig. 1). There is little difference between corrected results (Table 1) and those previously presented (Bachmann et al. 2013). For the lakes with good cores, there is still a statistically significant negative value for the average change in TP and positive changes in specific conductance and pH. There was no significant change in TN. The reference lakes also showed a statistically significant negative change in TP and a positive change for specific conductance. There were no statistically significant changes for TN and pH in the reference lakes. We also found for this sample, representing 60% of U.S. natural lakes, that the percentages of lakes classified into the four trophic states based on TP at the tops and bottoms of the cores are about the same in 2007 as they were in 1850 (Table 2). Likewise, we also found that for TP and TN the percentages of lakes rated as Good using the NLA methodology were about the same in 2007 as in 1850 and the percentages rated as Poor were somewhat lower in 2007 than in 1850 (Table 3). The frequency distribution of TP in 2007 is essentially the same as it was in 1850, with a slight decrease in the average concentrations for this group of lakes (Fig. 2).

We also followed the model of Herlihy et al. (2013), who used the reference lake approach with measured concentrations of TP and TN in the NLA lakes to determine the criteria for these nutrients in each of the ecoregions and calculated the number for both man-made and natural lakes combined that exceeded these values. We used the same method for all the natural lakes alone to find the percentages in each ecoregion that exceeded both the 75th and the 95th percentiles of their respective reference lake distributions. We also calculated exceedances for the 233 lakes with good sediment cores and the 204 lakes without good cores (Table 4).

Smith et al. (2014) and McDonald et al. (2014) both criticized our paleolimnological sample as being biased because it did not have a good representation of lakes with high sedimentation rates and because the lakes with good cores had a lower percentage of watershed disturbances and

Corresponding author: rbach@ufl.edu

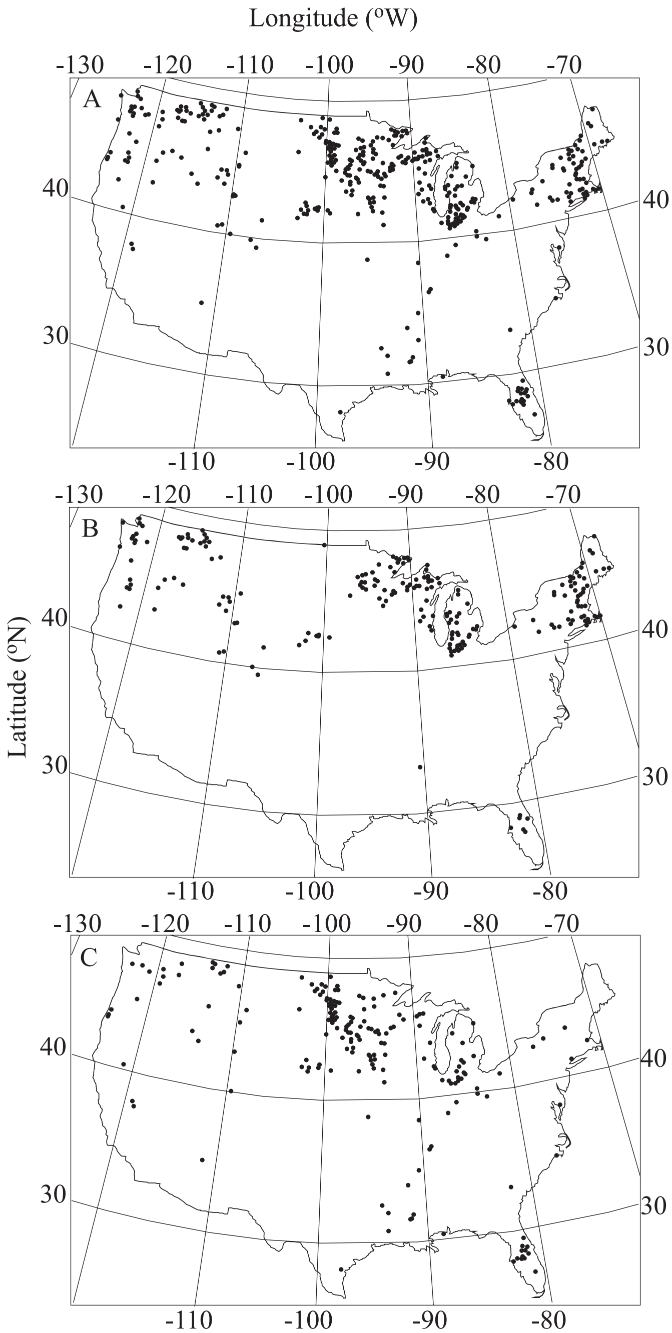

Fig. 1. (A) Locations of all natural lakes in the NLA study. (B) NLA lakes with good sediment cores. (C) Lakes without good sediment cores.

tended to be less productive than the lakes without good cores. We note that the geographic distribution of natural lakes in the United States is not even, but the lakes tend to be concentrated in the glacially disturbed regions (Fig. 1A) and in the western mountains, where we had most of our lakes with good cores (Fig. 1B). We found that 87% of the 29,308 natural lakes represented in the NLA are located north of latitude 40 degrees north. Large areas of the country do not have natural lakes; thus, many states are not represented in this survey. The major geographic area without good cores is Ecoregion 6, the glaciated plains, just south of the well-sampled glacial areas to the north (Fig. 1C). The alluvial lakes are not represented because they are unsuited for sediment core analyses; however, they are small in number compared to the glacial lakes. It is true that the average percentage watershed disturbance is lower for the lakes with good cores than for lakes without good cores (Smith et al. 2014); however, because the lakes with good cores do have a broad range of watershed disturbances extending to 75% (Fig. 3), they represent more than pristine, undisturbed lakes. Likewise, the lakes without good cores also include many lakes with minimal watershed disturbance (Fig. 3C).

Table 1. Weighted-average changes in the diatom-inferred concentrations of $\log_{10}$ total phosphorus (TP), $\log_{10}$ total nitrogen (TN), $\log_{10}$ specific conductance (SC), and pH during the period from pre-settlement times to 2007 for the 233 probability lakes with good cores and unweighted averages for the 50 reference lakes. The group of probability lakes represents 60% of the natural lakes in the United States. For TP, TN, and SC, the changes in $\log_{10}$ units are converted to percentages and are shown in parentheses.

Variable	Probability lakes (%)	Reference lakes (%)
TP	−0.084(−17.6)*	−0.131(−26.0)*
TN	−0.020(−4.5)	−0.026(−5.9)
SC	0.129(34.6)*	0.096(24.7)*
pH	0.082*	0.031

* The change is significantly different from 0 with $p = 0.05$.

Even though our paleolimnological sample covers only 60% of U.S. lakes, we will show that, for the purposes of our study, it does not matter that we do not have data on all lakes. We also show that, in using the reference lake method with measured surface water samples, the group of lakes without good cores has a higher percentage of lakes above the 75th and 95th percentiles than do the lakes with good cores; however, for the entire population of U.S. lakes the percentages exceeding the 75th and 95th percentiles are not as high as assumed in the U.S. EPA methodology (Table 4).

U.S. EPA reference conditions

To clarify the hypotheses that we were testing in the original paper (Bachman et al. 2013), we need to outline the recommended procedures in the nutrient criteria manual that led to our study. According to the manual (U.S. EPA 2000*a*), individual states and Indian tribes in the United States are required to set numerical nutrient criteria for TP and TN for the lakes and reservoirs within their jurisdictions.

Table 2. Percentages of lakes in four trophic states in 1850 and 2007 based on weighted averages of 233 lakes with good cores and representing 60% of the natural lakes in the United States.

Trophic state	1850	2007
Oligotrophic	48	49
Mesotrophic	23	16
Eutrophic	27	32
Hypereutrophic	2	3

Table 3. Percentages of lakes in the categories Good, Fair, and Poor based on diatom-inferred concentrations of TP and TN in the tops and bottoms of short sediment cores from the National Lakes Assessment in 2007 (U.S. EPA 2010). Placement of the lakes in the categories was based on the distributions of the concentrations of TP and TN in the reference lakes for the nutrient region in which a lake was located. The distributions are based on weighted averages for 233 lakes representing 60% of the natural lakes in the United States.

	Total phosphorus		Total nitrogen	
	1850	2007	1850	2007
Good	59	59	67	65
Fair	22	31	19	27
Poor	19	10	14	9

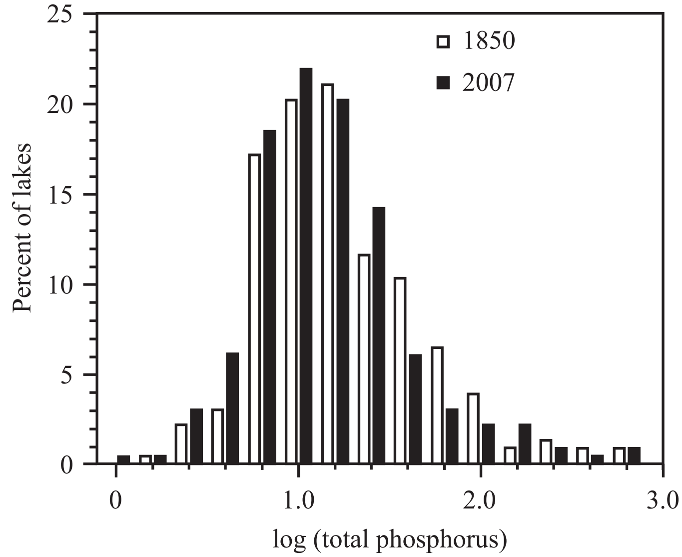

Fig. 2. Distribution of diatom-inferred total phosphorus concentrations in 233 lakes representing 60% of U.S. natural lakes for 1850 and 2007.

Lakes or reservoirs with concentrations of TP and/or TN that exceed the criteria are considered to be impaired and are subject to regulatory activities that may include activities to reduce the concentrations of TP and/or TN to levels below the criteria concentrations. To provide guidance to the states in setting numeric nutrient criteria, the U.S. EPA (2000*a*) published a guidance manual with methods that could be used to set these criteria. They also used those methods to set what they called reference conditions for each of the U.S. ecoregions.

Our Fig. 4 is a redrawing of a figure from the nutrient guidance manual (U.S. EPA 2000*a*) that illustrates the basis for their reference condition recommendations. The left-hand curve represents the frequency curve for current TP concentrations in reference lakes in an ecoregion and is the basis for their preferred method of setting the reference condition. It is assumed that those lakes with TP concentrations higher than the concentration at the 75th percentile are the result of anthropogenic additions of phosphorus, so the recommended reference condition would be placed at the 75th percentile of reference lakes. This percentile was chosen because all lakes were considered to have been affected by human development to some degree by landscape disturbance and atmospheric pollution. Further, it was stated that the upper 25th percentile represented an appropriate margin of safety and excluded the effects of outliers (U.S. EPA 2000*a*). The NLA also used the same reference lake procedure to identify lakes rated as Fair or Poor from those rated as Good for TP and TN. Presumably, the frequency distribution prior to anthropogenic disturbance would have been to the left of the current distribution because none of the lakes would have been artificially eutrophied prior to that time (Fig. 5A,B). Because this percentile is applied to all natural lakes in the United States, it implies that 25% of natural lakes of reference quality in the United States have had sufficient anthropogenic nutrient loading to increase their nutrient concentration above the normal pre-settlement conditions for their respective nutrient ecoregions. This is the assumption we wanted to test in our original paper (Bachmann et al. 2013).

The right-hand curve (Fig. 4) represents the current frequency distribution of TP for a representative sample of all lakes in an ecoregion and is the basis for an alternative method to be used when reference lake data are not available. This method assumes that all or most of the lakes have been subjected to a degree of anthropogenic nutrient loading that justifies shifting the curve to the right, such that the location of the lower 25th percentile now has a concentration about the same as the 75th percentile of the reference lakes. This procedure was considered by the U.S. EPA to be a surrogate for the reference lake method, and, at the time the manual was published, available data indicated that the lower 25th percentile of data from a sample representative of the entire population of lakes in an ecoregion roughly approximated the upper 25th percentile of the reference data. It was noted that in regions of intense cultural enrichment a lower percentile ($< 25\%$) of lakes must be selected to avoid establishing criteria based on degraded conditions (U.S. EPA 2000*a*).

Whereas the reference lake method was preferred, in practice the U.S. EPA used the 25th percentile method with a 10 yr series of data on lakes from government databases to determine reference conditions for each of the nutrient ecoregions and published them in a series of ambient water quality criteria recommendations (for example, U.S. EPA 2000*b*) to support the development of nutrient criteria for states and Indian tribes. The U.S. EPA encouraged states and tribes to use the recommended reference conditions in developing their nutrient criteria for lakes. The reference conditions themselves were not criteria, and the states were encouraged to collect more data to refine them to their conditions in order to come up with scientifically defensible alternatives. However, if a state did not choose to set their own criteria, the U.S. EPA could use the reference conditions as the nutrient criteria for that state. Because this is the method that was used to set the reference condition for the lakes in all ecoregions of the United States, it implies that 75% of the natural lakes in the United States have had sufficient anthropogenic nutrient loading to increase their nutrient concentration above the recom-

Table 4. Percentages of natural lakes exceeding the 75th percentile and the 95th percentile of reference lakes for TP and TN for each nutrient ecoregion. The percentages for each nutrient ecoregion are weighted averages based on surface-water samples and the weights for each NLA lake sampled. There were no natural lakes sampled in the Southern Appalachian ecoregion. Weighted averages are presented for all natural lakes, natural lakes with good cores, natural lakes without good cores, and for man-made lakes.

Nutrient ecoregions	No. of NLA probability lakes	Percent exceeding 75th percentile for TP	Percent exceeding 95th percentile for TP	Percent exceeding 75th percentile for TN	Percent exceeding 95th percentile for TN
Coastal Plains	1662	59	30	55	1
Western Mountains	2770	31	17	40	27
Xeric West	71	40	35	54	20
Grass Plains	126	8	0	20	20
Cultivated Great Plains	1925	77	68	91	91
Temperate Plains	4328	68	29	73	44
Southern Glaciated	7624	32	1	45	3
Northern Glaciated	10,066	34	13	33	10
Southeast Plains and Piedmont	736	72	1	17	1
Weighted averages					
All natural lakes	29,308	43	17	47	19
Natural lakes with good cores	17,469	30	9	38	9
Natural lakes without good cores	11,839	61	28	61	35
All man-made lakes	20,238	41	20	45	19

mended reference condition. We specifically tested that hypothesis in our original paper to assess the extent of cultural eutrophication in U.S. lakes as viewed by the U.S. EPA, the primary regulatory agency (Bachmann et al. 2013).

According to the guidance manual (U.S. EPA 2000*a*), the two recommended methods were not based on quantitative scientific studies but were qualitative in nature. The manual stated: "The quarterly increments were chosen as a reasonable division of the data sets recognizable by the public, and the upper 25th percentile and lower 25th percentile as reasonable and traditional fractions of the range and frequency of distribution." This is understandable because, at that time, data were not available from probability-based samples of U.S. lakes. These data became available for the first time with the NLA study and gave us an opportunity to test the two main assumptions.

Quantification of cultural eutrophication for a population of lakes

Both the NLA, in rating the nutrient condition of U.S. lakes, and the U.S. EPA, in setting nutrient criteria, recognize that there is a range of TP and TN concentrations in the lakes of an ecoregion when they are in an undisturbed state. They set an upper threshold concentration that can be used as a criterion to identify lakes at the current time that had increases in TP or TN concentrations that placed them outside the normal range for undisturbed lakes in that ecoregion. Our study was concerned with the percentages of lakes that exceed the threshold values. These lakes would be rated as Fair or Poor in the NLA or as in need of nutrient reduction for cultural eutrophication by the U.S. EPA.

Another measure of cultural eutrophication would be to measure the percentage of lakes that currently show a statistically significant increase in their nutrient concentrations since a pre-settlement time. This is different from the percentage of lakes exceeding a threshold value, because they are measuring different things. For example, in the left-hand curve in Fig. 4 a lake whose TP concentration increased from 10 μg L^{-1} to 25 μg L^{-1} would not count in the percentage exceeding the threshold; however, an increase from 25 μg L^{-1} to 40 μg L^{-1} would count as an impaired lake because the TP concentration is raised above the normal range for the ecoregion. Likewise, in the population of lakes in an ecoregion as a whole, if 75% of the lakes showed some increase over time it would not necessarily mean that 75% exceed the threshold that would place them outside of the pristine range. In general, the percentage of lakes showing changes is going to be greater than the percentage of those that exceed a nutrient criterion. It is noted that protection from nutrient increases for lakes that currently have nutrient concentrations less than the upper bound is provided through antidegradation rules that prevent increases even though concentrations do not reach the nutrient criterion (U.S. EPA 2000*a*).

A third measure would be the average change in TP for a population of lakes over a period of time. This was the approach that we used for the lakes representing 60% of the U.S. natural lakes because we did not have an error term for the diatom-inferred concentrations of TP and TN that would have allowed us to identify which lakes had shown a statistically significant increase in TP or TN. In this case, an average change that did not exceed 0 indicated there was no change in the nutrient concentrations for the sample lakes as a whole.

Use of the 75th percentile of reference lakes

Reference lakes are used to estimate the normal range of nutrient concentrations in an ecoregion when they are in an

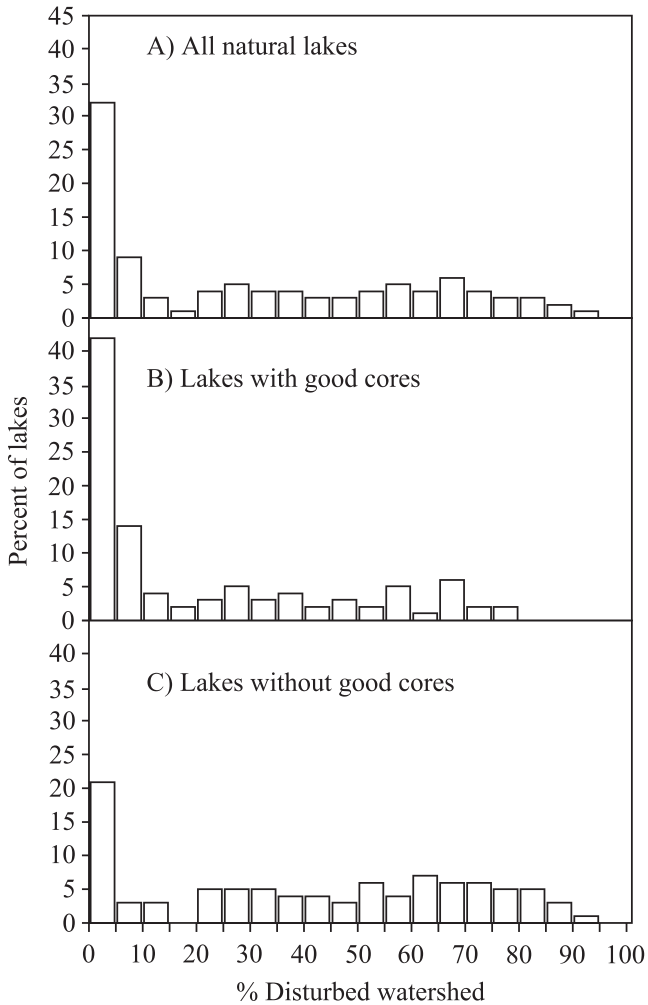

Fig. 3. (A) The frequency distributions of percentages of watersheds disturbed for all natural lakes in the NLA. (B) The same for lakes with good cores. (C) The same for lakes without good cores. Following Smith et al. (2014) the percentage disturbed is the sum of percentages of the watersheds classified as cropland, pastures, and developed lands.

undisturbed state. If we knew the predisturbance distributions of TP and TN, we could follow the common practice in statistical testing and use a range that encompasses 95% of the values in the population (Kilgour et al. 1998). The 95th percentiles would be logical thresholds to represent the upper bounds of the distributions (Fig. 5A) because we are only concerned with lakes with increases in concentrations of TP or TN. Modern lakes with concentrations above the 95th percentile would have a good probability of being outside the normal range and, thus, subject to regulation for nutrient reductions.

In using reference lakes to set nutrient criteria, the U.S. EPA (2000*a*) assumed that the pre-settlement distribution of TP or TN (Fig. 5A) was shifted to the right sufficiently by the current time (Fig. 5B) such that lakes with concentrations of TP or TN above the 75th percentile would represent lakes that needed nutrient reductions to restore them to a pre-settlement condition. Similarly, the NLA used the 75th

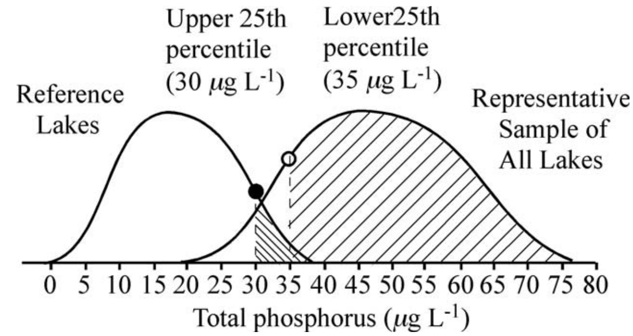

Fig. 4. Conceptual illustration used in the U.S. EPA nutrient guidance manual (U.S. EPA 2000*a*) to demonstrate the assumptions used to set nutrient reference conditions using either reference lakes or a representative sample of all lakes in an ecoregion. Redrawn from their fig. 6.2.

percentile as the dividing point to separate Good lakes from Fair and Poor lakes for nutrient condition.

As a test of this assumption, both we (Bachmann et al. 2013) and Herlihy et al. (2013) examined the changes in TP and TN going from the bottoms to the tops of the 50 short cores for reference lakes in the NLA study and came to the same result. Herlihy et al. (2013) found that the median change for TP was -1.17 μg L^{-1} and the median change in TN was -25.3 μg L^{-1} and concluded that "the vast majority of the 50 reference lakes where we were able to sample sediment diatoms had no significant increases in TP or TN concentrations over time." Although they did not test their results for statistical significance, their findings were in agreement with our results: using the average of log-transformed values, TP showed a decrease and TN did not change in going from the bottoms to the tops of the reference lake cores (Table 1).

As a consequence of this finding, there is no evidence that for the reference lakes the distributions of TP or TN in 2007 are any higher than their distributions in 1850 (Fig. 5A,C). The result is that the use of the 75th percentile of 2007 values overestimates the number of nutrient-impaired lakes; and, if this is used as a criterion, we will always have at least 25% of the lakes in our population deemed impaired for regulatory purposes or in Fair or Poor conditions per the NLA protocols. This would be true if there were no change at all from 1850 to the present. For example, when we applied the 75th percentile standard to the 1850 values at the bottoms of the cores for the reference lakes, we found that 24% of the TP values and 22% of the TN values in the reference lakes would have exceeded the nutrient criteria in 1850 and, under the NLA methodology, would be classified as Fair or Poor. Because this was a time prior to extensive development, they should all have been classified as Good, or only 5% would be classified as Fair or Poor if one used the 95th percentile criterion.

These findings cast doubt on the assumption that the 75th percentile of the distribution of reference lakes can be used to determine which lakes had increased enough to place them out of the expected range for pristine lakes in that ecoregion. The end result is that the use of the 75th percentile of reference lakes for a nutrient criterion will

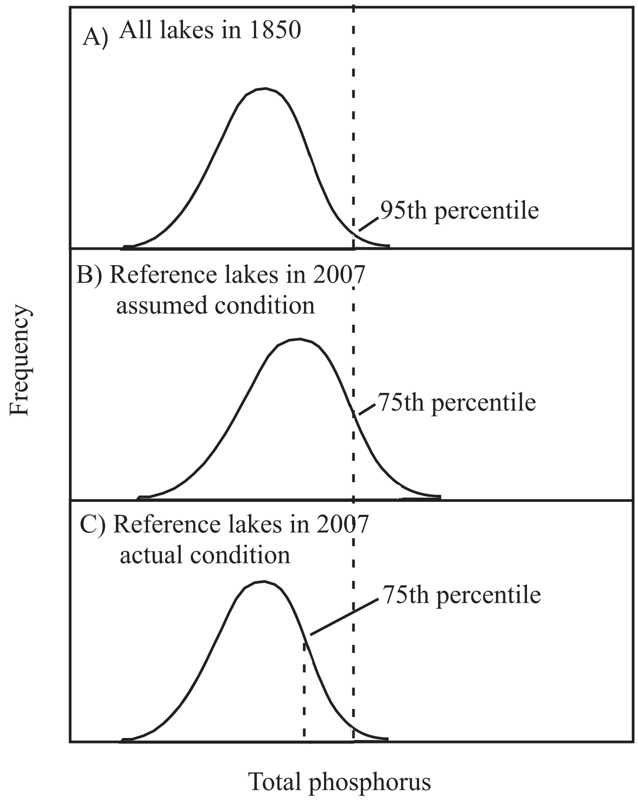

Fig. 5. (A) Frequency distribution of total phosphorus concentrations for lakes in a nutrient ecoregion at some time prior to anthropogenic development. The vertical dashed line at the 95th percentile represents a potential nutrient criterion for that ecoregion. (B) Frequency curve for reference lakes at the current time following development, under the assumption that the curve has shifted to the right due to some cultural eutrophication such that the potential criterion lies at the 75th percentile. (C) Frequency curve for reference lakes based on finding that reference lakes have shown no significant increase in average TP since 1850. The use of the 75th percentile of this curve would overestimate the number of nutrient-impaired lakes.

overestimate the number of affected lakes by about 25% (Fig. 5C). These findings are the basis for our statement that the extent of cultural eutrophication is overstated and that this should be taken into account in setting nutrient criteria (Bachmann et al. 2013).

Use of the 25th percentile of all lakes in an ecoregion

The use of the 25th percentile of the concentrations of TP and TN from a representative group of lakes from an ecoregion to set nutrient criteria is based on the assumption that it will yield about the same result as the 75th percentile of a group of reference lakes from the same ecoregion (U.S. EPA 2000*a*). Herlihy et al. (2013) tested this assumption with nutrient concentrations from surface-water samples, using the 75th percentile of reference lakes in each ecoregion in the NLA study to set nutrient criteria for all natural and man-made lakes, and compared the results with criteria established using the 25th percentile of all lakes. They found that nutrient criteria based on the population 25th percentiles were lower than criteria obtained with other approaches, often by a factor of 2–6, in almost all ecoregions. They concluded that population 25th percentiles cannot be used as a surrogate for the 75th percentile of reference lakes or paleolimnological approaches.

The corollary to the 25th percentile method is that it assumes that 75% of a representative sample of all lakes in an ecoregion now have concentrations of TP and TN that exceed the upper end of the distributions of their concentrations as they existed in 1850. Recall that the U.S. EPA (2000*b*) used this method to set reference conditions for lakes in all ecoregions in the United States. There are several ways to test this assumption. We used the Herlihy et al. (2013) study as an example; with measured nutrient concentrations from surface waters in 2007, we used the 75th percentile and 95th percentile of reference lakes as a criterion for all the natural lakes in the NLA study and calculated the percentages of lakes that would exceed these nutrient criteria in each ecoregion. We found, using the 75th percentile, that only 43% and 47% of all the lakes had concentrations exceeding the criteria for TP and TN, respectively (Table 4). If we take into account the finding that the reference lakes did not show an increase in TP and TN from 1850 to 2007, and if we use the 95th percentiles of reference lake concentrations of TP and TN as criteria, then 17% of the natural lakes would be declared as nutrient impaired for TP and 19% for TN (Table 4). In either case, there is no evidence that 75% of U.S. lakes have levels of TN and TP that exceed the upper bounds of pre-settlement distributions.

Paleolimnological approach

Another approach is to use the results of our analysis of the sediment core samples available for the 233 lakes that represent 60% of the natural lakes in the United States (Table 1). The average changes in the concentrations of TP and TN for this group were not greater than 0, and the frequency distribution of TP in these lakes in 2007 is very close to the distribution in 1850 (Fig. 2), with a slight decline in the average concentrations (Table 1). Also, for this group of lakes, the proportions of lakes in the four trophic state categories did not change from 1850 to the present (Table 2), nor did the proportions of lakes in the NLA categories of Good, Fair, and Poor change from 1850 to 2007 (Table 3). The positive and negative values for individual lakes can be attributed in part to random experimental error that would come from finding the difference between two measures that both are subject to sampling error. Any other changes would represent a balance between the lakes that had increases in nutrient concentrations due to anthropogenic activities, climatic cycles, or other factors vs. the lakes that had decreases in nutrient concentrations due to anthropogenic activities, climatic cycles, or other factors. We do not have evidence that the distributions of TP and TN in this group of lakes have been changed by cultural eutrophication.

Even though we could not obtain good core data from all lakes and there may be a basis for supposing that the 40% of the lakes without good cores might be more likely to show increases in TP and TN over time (McDonald et al.

2014; Smith et al. 2014), mathematically it would be impossible for 75% of all U.S. lakes to exceed the nutrient criteria. This finding is another reason for our statement that the extent of cultural eutrophication is overstated and that this should be taken into account in setting nutrient criteria (Bachmann et al. 2013).

Reservoirs and small lakes

Our original study was limited to the lakes sampled in the NLA with paleolimnological data. This excluded man-made lakes and lakes with surface areas less than 0.04 km^2. Herlihy et al. (2013) used the reference lake methodology with 2007 surface-water samples to find the numbers of natural and man-made lakes combined that would exceed the 75th percentile of the reference lakes in each ecoregion. We used the same methodology to find the percentage of man-made lakes alone that would exceed both the 75th and the 95th percentiles of the same reference lakes for TP and TN. For the 75th percentile, 41% of the man-made lakes had exceedance for TP and 45% had exceedances for TN. If we can assume that the reference lake methodology for lakes applies equally well for natural and man-made lakes, the 95th percentile might yield a better estimate, as explained previously. In this case, the percentage exceedances for TP and TN in man-made lakes are 20% and 19%, respectively.

Support from the literature

Both Smith et al. (2014) and McDonald et al. (2014) use the literature to argue that our conclusions about the extent of cultural eutrophication in U.S. lakes must be wrong. They reviewed studies of cultural eutrophication, including discussions of nutrient loadings from agriculture and atmospheric deposition, models of nutrient concentrations in U.S. rivers, and case studies of eutrophication of lakes in Connecticut and in several European countries. These studies suggest that there might be a large percentage of U.S. lakes that have been culturally eutrophied; however, with the exception of Herlihy et al. (2013), none of these studies has used a probabilistic sample of U.S. lakes to estimate the percentages of natural lakes in the United States that have shown increases in TP or TN that exceed the normal range for the lakes in that ecoregion. We have already shown how the Herlihy et al. (2013) study supports our findings. There are no other national studies, but there are two regional studies in the United States using a paleolimnological approach that also support our conclusions.

Northeastern lakes study—For the NLA, the U.S. EPA (2010) did not date the bottoms of the short cores to determine whether they reached a pre-1850 level but relied on their previous experience in conducting short core studies on lakes in the northeastern United States to determine whether they were confident they had reached that depth. As a part of the U.S. EPA Environmental Monitoring and Assessment Program (Larsen et al. 1991), a probabilistic study of lakes and reservoirs in the northeastern United States (Maine, New York, New Hampshire, Massachusetts, Vermont, Connecticut, New Jersey, and Rhode Island) was carried out in the mid-1990s (Dixit and Smol 1994; Dixit et al. 1999). Short sediment cores were used to find diatom-inferred values for TP at the tops and bottoms of the cores. In that study, the bottoms of the cores were examined using pollen and ^{210}Pb dating, and only 3 of the 159 natural lakes had sediment cores that did not reach a pre-1850 level.

As a check on the dating of the NLA cores, we used a t-test to compare the mean diatom-inferred concentration of TP at the bottom of the NLA cores from lakes in the same northeastern states (8.3 μg L^{-1}) with the mean values for TP at the bottom of cores dated as pre-1850 in the northeastern lakes study (8.2 μg L^{-1}). We found no statistically significant difference (p = 0.91) between the means for the two studies. This test would support the contention that, at least for the northeastern lakes, the NLA core samples with confidence ratings of "Yes" (to the question of whether sediment at the bottom of the core represented a time prior to extensive anthropogenic development) were equivalent to those of cores dated with pollen and ^{210}Pb in the earlier study.

The northeastern lakes study also had several lakes with replicated cores that could be used to estimate the random error associated with the difference between the diatom-inferred TP concentrations at the tops and bottoms of the same core. There were 10 lakes with two cores, 11 lakes with three cores, 6 lakes with four cores, and 2 lakes with five cores. We used these data to calculate a pooled variance of 0.0196 and a standard deviation of 0.14 log units for the difference between the top and bottom TP estimates. We used this standard deviation and a t-value of 2 to calculate the 95% confidence limits on the difference between the diatom-inferred concentrations of TP at the tops and bottoms of these cores, plus or minus 0.28 log units. For the 156 northeastern lakes and 6 reservoirs with cores extending back to pre-1850, 20 lakes had increases in diatom-inferred TP that exceeded the 95% confidence limit of 0.28 log units and 6 showed decreases (Fig. 6). This would imply that about 12% of the northeastern U.S. lakes might have had their TP increased by cultural eutrophication, 4% decreased, and about 84% showed no significant change. We do not have reference lake data to determine whether any of the lakes with increases exceeded the normal range for their ecoregion. The data of Dixit et al. (1999) showed no important change in trophic states for the lakes in the northeastern states. For the pre-1850 time periods, the percentages of oligotrophic, mesotrophic, and eutrophic lakes were 47%, 49%, and 3%, respectively. In the 1990s, those percentages were 50%, 46%, and 4%. Whittier et al. (2002) also analyzed the data from that study and concluded that the population of natural lakes in the northeastern United States had changed very little since before 1850.

Minnesota lakes study—The results of the paleolimnological study of Minnesota lakes (Ramstack et al. 2004) and cited by Smith et al. (2014) also are in agreement with our findings that the percentage of U.S. lakes suffering from cultural eutrophication is less than 75%. They reported on

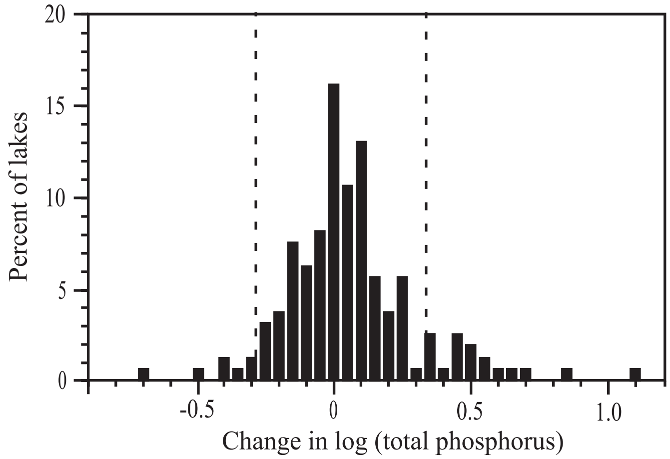

Fig. 6. Frequency distribution of changes in total phosphorus in lakes of the northeastern U.S. between pre-1850 and the early 1990s. The dashed lines indicate the 95% confidence limits on a mean of 0 based on replicated samples. Changes between the dashed lines are not statistically significant. Data are from study of Dixit et al. (1999).

diatom-inferred concentrations of TP from long sediment cores taken in 52 lakes. The lakes are located in three major ecoregions of the state: Northern Lakes and Forests (NLF), North Central Hardwood Forests (NCHF), and the Western Corn Belt Plains (WCBP). A subset of the NCHF region included urban lakes in the Twin Cities Metropolitan Area (METRO). Whereas the lakes were not randomly selected, Smith et al. (2014) and the authors of that study (Ramstack et al. 2004) considered them to be qualitatively representative of the larger population of lakes in each of the ecoregions. The cores were dated with ^{210}Pb, and the reported changes in TP represent a time period from 1800 to the mid-1990s. The numbers of lakes in each ecoregion were found in Heiskary et al. (1987), and the number of METRO lakes (883) was provided by the Minnesota Department of Natural Resources.

The percentage of Minnesota lakes that Ramstack et al. (2004) considered to have statistically significant increases in TP ranged from 0% to 33%, depending on the ecoregion (Table 5). The weighted average for the 10,900 Minnesota lakes was 17% for lakes with statistically significant increases in TP, which is similar to the 12% for the northeastern lakes. Again, we do not know how many of them have had increases in TP that place them above the normal ranges for their ecoregions.

Regional differences in lake density and cultural eutrophication

Both comments (McDonald et al. 2014; Smith et al. 2014) noted differences in the numbers of lakes in the different ecoregions. Herlihy et al. (2013) demonstrated this fact for all of the natural and man-made lakes combined, as we did for the natural lakes of the NLA (Table 4). Most of the natural lakes are located in just four of the ecoregions (Northern Glaciated, Southern Glaciated, Temperate Plains, and Western Mountains). Together, they account for 84% of the natural lakes in the NLA target population, so it is natural that they have a dominant effect on national averages, as noted by the comments to our paper (McDonald et al. 2014; Smith et al. 2014).

Table 5. Results of a paleolimnological study of Minnesota lakes (Ramstack et al. 2004) in four regions with numbers of lakes showing statistically significant increases in TP since 1800.

Region	No. of lakes in region	No. of lakes examined	Lakes with significant increases	
			No.	%
NLF	5558	20	1	5
NCHF	3882	9	3	33
WCBP*	577	5	0	0
METRO	883	18	6	33

Weighted average for 10,600 Minnesota lakes = 17% with increases in TP.
* Does not include the 883 lakes in the METRO portion of this ecoregion.

When either the 75th percentile or the 95th percentile of reference lakes approach is used to find the percentage of lakes in which the concentrations of TP and TN exceed those percentiles, there is a wide range in exceedances among the different ecoregions. This is true for the combined natural and man-made lakes (Herlihy et al. 2013), for the natural lakes (Table 4) in the NLA, and also for the Minnesota lakes (Table 5). This reinforces the idea that we cannot apply a single percentage of eutrophied lakes (e.g., 75%) to set the nutrient criteria in a given ecoregion. We recommend that in any geographic region the process of setting criteria should start out with knowledge of the natural distributions of TP and TN in the lakes in question. We did this for lakes in the state of Florida (Bachmann et al. 2012*a*,*b*), where there are 47 subregions of the 3 major ecoregions (Griffith et al. 1997), thus representing a diversity of expected nutrient concentrations. We agree with Herlihy et al. (2013) that, for natural U.S. lakes, the paleolimnological approach is to be favored where possible because it is a direct measure of nutrient conditions before extensive human development.

Proposed research

Among our study (Bachmann et al. 2013) and those reviewed here and in the literature, there is no evidence to support the hypothesis that 75% of U.S. lakes are culturally eutrophied to the extent that their current concentrations of TP and TN exceed the normal range for their ecoregions prior to European settlement. We have presented various analyses that suggest a much lower percentage. Because it is important to have a scientific basis for setting numeric nutrient criteria in lakes, we suggest that improved paleolimnological surveys be conducted in the various ecoregions of the United States to provide a valid picture of the natural trophic states of lakes prior to European settlement. Lakes should be selected on a probabilistic basis, and the bottoms of the cores should be dated to ensure that they extend into a pre-settlement time period. If that level is not reached on the first sampling of a lake, additional longer cores in that lake should be taken to obtain a pre-settlement sample of sediment. Additional lakes should be sampled to replace lakes in the sample design that cannot yield a good core for one reason or another. In addition, replicate cores should be analyzed from several lakes in order to derive error terms that can be used to determine the statistical confidence limits on changes in the concentrations of TP and TN from the bottoms to the tops of the cores. This would allow the determination of which individual lakes had statistically significant changes in TP and/or TN over time. It would also be valuable to reexamine the diatom inference methodology to see whether improvements in precision and accuracy could be possible.

We agree with Bennion et al. (2011) on the use of the paleolimnological method to establish reference conditions for lake restoration. We suggest that the distributions of diatom-inferred concentrations of TP and TN from the pre-settlement period in each ecoregion be used to establish the reference conditions for nutrient criteria. One could use the 95th percentile as a reference condition or other measures of the normal range as outlined by Kilgour et al. (1998). This would eliminate the potential bias in picking reference sites, because the lakes would have been selected on a probabilistic basis. It would also eliminate assumptions about changes in nutrient concentrations in the reference lakes following European settlement. Also, in agreement with Bennion et al. (2011), it must be recognized that it may be impossible to return every body of water to pre-1850 conditions and that alternative restoration targets may have to be selected.

Roger W. Bachmann, Mark V. Hoyer, and Daniel E. Canfield, Jr.

Fisheries and Aquatic Sciences
School of Forest Resources and Conservation
University of Florida, Gainesville, Florida

Acknowledgments

Steve Paulsen, U.S. Environmental Protection Agency (U.S. EPA), provided us with the paleolimnological data from the Environmental Monitoring and Assessment Program Northeastern Lakes study funded by the U.S. EPA, and Alan Herlihy provided information on the codes used to distinguish reference and probability lakes in the National Lakes Assessment spreadsheets. Dana Bigham produced maps of lake distributions. Four reviewers and Ronnie Glud provided constructive comments on our first draft.

References

Bachmann, R. W., D. L. Bigham, M. V. Hoyer, and D. E. Canfield, Jr. 2012*a*. Factors determining the distributions of total phosphorus, total nitrogen and chlorophyll *a* in Florida lakes. Lake Reserv. Manage. **28:** 10–26, doi:10.1080/07438141.2011.646458

———, ———, ———, and ———. 2012*b*. A strategy for establishing numeric nutrient criteria for Florida lakes. Lake Reserv. Manage. **28:** 84–91, doi:10.1080/07438141.2012.667053

———, M. V. Hoyer, and D. E. Canfield, Jr. 2013. The extent that natural lakes in the United States of America have been changed by cultural eutrophication. Limnol. Oceanogr. **58:** 945–950.

Bennion, H., R. W. Battarbee, C. D. Sayer, G. L. Simpson, and T. A. Davidson. 2011. Defining reference conditions and restoration targets for lake ecosystems using paleolimnology: A synthesis. J. Paleolimnol. **45:** 533–544, doi:10.1007/s10933-010-9419-3

Dixit, S. S., and J. P. Smol. 1994. Diatoms as indicators in the environmental monitoring and assessment program-surface waters (EMAP-SW). Environ. Monit. Assess. **31:** 275–307.

———, ———, D. F. Charles, R. M. Hughes, S. G. Paulsen, and G. B. Collins. 1999. Assessing water quality changes in the lakes of the northeastern United States using sediment diatoms. Can. J. Fish. Aquat. Sci. **56:** 131–152, doi:10.1139/f98-148

Griffith, G. E., D. E. Canfield, Jr, C. A. Horsburgh, and J. M. Omernik. 1997. Lake regions of Florida. EPA/R-97/127. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory.

Heiskary, S. A., C. B. Wilson, and D. P. Larsen. 1987. Analysis of regional patterns in lake water quality: Using ecoregions for lake management in Minnesota. Lake Reserv. Manage. **3:** 337–344, doi:10.1080/07438148709354789

Herlihy, A. T., N. C. Kamman, J. C. Sifneos, D. Charles, M. D. Enache, and R. J. Stevenson. 2013. Using multiple approaches to develop nutrient criteria for lakes in the conterminous USA. Freshw. Sci. **32:** 367–384, doi:10.1899/11-097.1

Kilgour, B. W., K. N. Somers, and D. E. Matthews. 1998. Using the normal range as a criterion for ecological significance in environmental monitoring and assessment. Ecoscience **5:** 542–550.

Larsen, D. P., D. L. Stevens, A. R. Selle, and S. G. Paulsen. 1991. Environmental monitoring and assessment program—surface waters: A northeast lake pilot. Lake Reserv. Manage. **7:** 1–11, doi:10.1080/07438149109354249

McDonald, C. D., and others. 2014. The importance of using representative samples for large-scale paleolimnological inference: Comment on Bachmann et al. (2013). Limnol. Oceanogr. **59:** 2226–2230, doi:10.4319/lo.2014.59.6.2226

Ramstack, J. M., S. C. Fritz, and D. R. Engstrom. 2004. Twentieth century water quality trends in Minnesota lakes compared with presettlement variability. Can. J. Fish. Aquat. Sci. **61:** 561–576, doi:10.1139/f04-015

Smith, V. H., W. K. Dodds, K. E. Havens, D. R. Engstrom, H. W. Paerl, B. Moss, and G. E. Likens. 2014. Comment: Cultural eutrophication of natural lakes in the United States is real and widespread. Limnol. Oceanogr. **59:** 2217–2225, doi:10.4319/lo.2014.59.6.2217

U.S. Environmental Protection Agency. 2000*a*. Nutrient criteria technical guidance manual: Lakes and reservoirs. EPA-822-B00-001. U.S. Environmental Protection Agency, Office of Water.

U.S. Environmental Protection Agency. 2000*b*. Ambient water quality criteria recommendations. Information supporting the development of state and tribal nutrient criteria for lakes and reservoirs in nutrient ecoregion XIII. EPA 822-B-00-014. U.S. Environmental Protection Agency, Office of Water.

U.S. Environmental Protection Agency. 2010. National lakes assessment: A collaborative survey of the nation's lakes. EPA 841-R-09-001. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, D.C.

Whittier, T. R., S. G. Paulsen, D. P. Larsen, S. A. Peterson, A. T. Herlihy, and P. R. Kaufmann. 2002. Indicators of ecological stress and their extent in the population of northeastern lakes: A regional-scale assessment. BioScience **52:** 235–247, doi:10.1641/0006-3568(2002)052[0235:IOESAT]2.0.CO;2

Associate editor: Ronnie Nøhr Glud

Received: 05 March 2014
Accepted: 01 August 2014
Amended: 04 August 2014