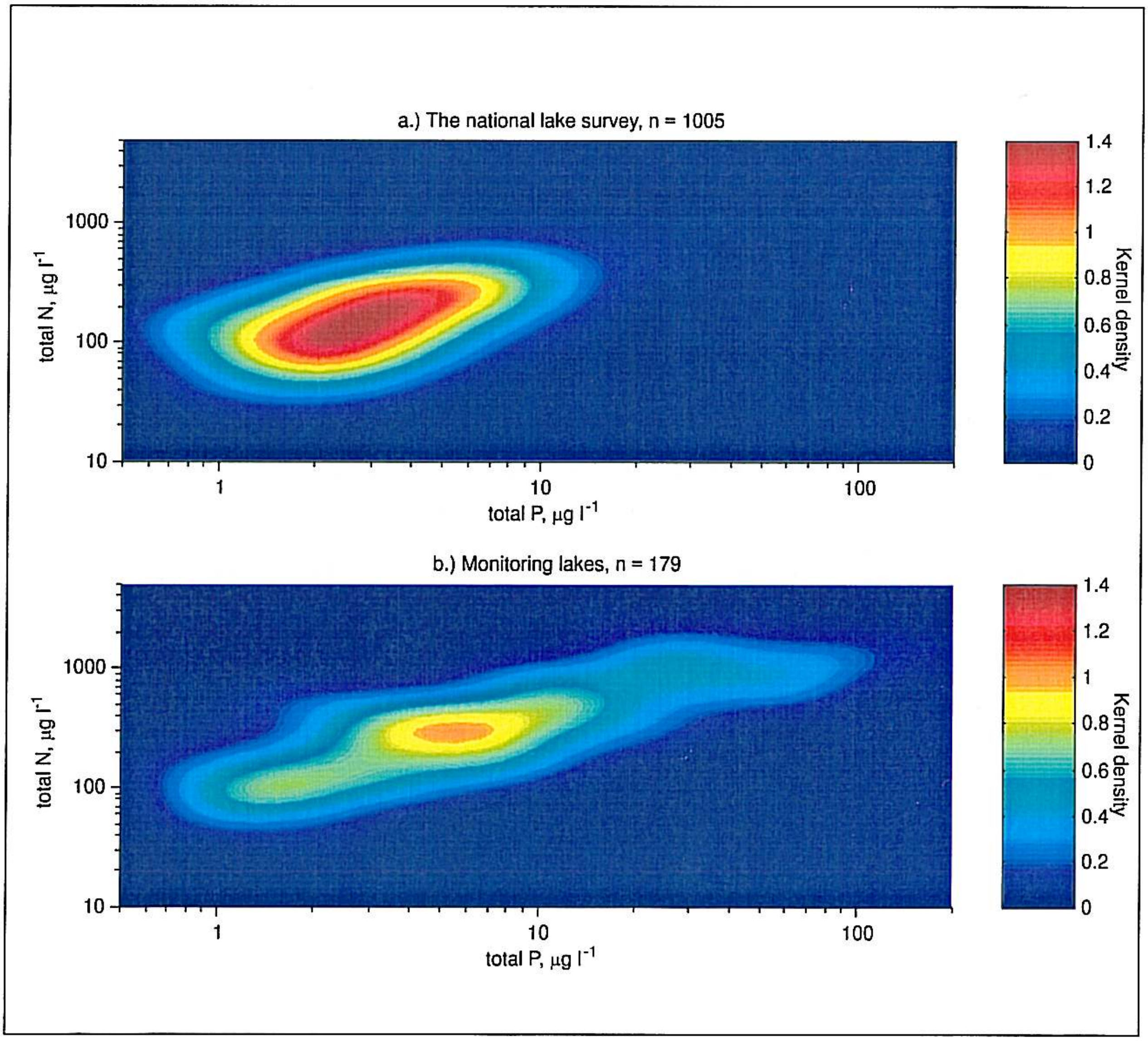
REPORT SNO 3997-99

How can Norwegian monitoring programs for lakes and rivers be used to fulfill the requirements of EUROWATERNET?



Norwegian Institute for Water Research

REPORT

Main Office	Regional Office, Sørlandet	Regional Office, Østlandet	Regional Office, Vestlandet	Akvaplan-NIVA A/S
P.O. Box 173, Kjelsås	Televeien 1	Sandvikaveien 41	Nordnesboder 5	
N-0411 Oslo	N-4890 Grimstad	N-2312 Ottestad	N-5008 Bergen	N-9005 Tromsø
Norway	Norway	Norway	Norway	Norway
Phone (47) 22 18 51 00	Phone (47) 37 29 50 55	Phone (47) 62 57 64 00	Phone (47) 55 30 22 50	Phone (47) 77 68 52 80
Telefax (47) 22 18 52 00	Telefax (47) 37 04 45 13	Telefax (47) 62 57 66 53	Telefax (47) 55 30 22 51	Telefax (47) 77 68 05 09
Internet: www.niva.no				

Title	Serial No.	Date
How can Norwegian monitoring programmes for lakes and rivers be used to fulfil the requirements of EUROWATERNET?	3997-99	January 1999
	Report No. Sub-No. O-95028	Pages Price 14
Author(s) Eirik Fjeld Arne Henriksen Tor S. Traaen	Topic group Eutrophy, surface water	Distribution
	Geographical area Norway	Printed NIVA

Client(s)	Client ref.
Norwegian Pollution Control Authority (SFT)	

Abstract

The European Topic Centre on Inland Waters (ETC/IW) has designed and tested an information and monitoring network, called EUROWATERNET. Technical Guidelines for Implementation of EUROWATERNET has been worked out, which gives the definition and concept of the network, as well as guidelines for selecting river stations, lakes and groundwater. The Guidelines have been tested for selection of lakes and rivers based on existing Norwegian monitoring programmes. The national lake survey of 1995 (n = 1005) gives the water chemical characteristics representative for the whole lake population in Norway. A method for adjusting for the non-random sampling was applied in order to provide representative water chemistry statistics (total nitrogen and phosphate) from a biased sample of monitoring lakes. These monitoring lakes (n = 179) were selected in order to give a good coverage also of mesotrophic and eutrophic lakes. Bivariate Gaussian kernel density estimates were used in order to obtain smoothed frequency distributions of N and P for the two datasets. These frequency distributions were then used to estimate different weights for the lakes in the monitoring lake data set. By applying these weights in the statistical calculations, water chemistry statistics representative for the whole population of Norwegian lakes can be calculated – based on data from a biased sample of monitoring lakes. To design a river monitoring program similar to the lake monitoring programme would require knowledge of the properties of an unbiased sample of the river population in Norway.

4 keywords, Norwegian	4 keywords, English
1. Overvåking	1. Monitoring
2. Innsjøer	2. Lakes
3. Elver	3. Rivers
4. Næringssalter	4. Nutrients

Merete Johannessen Ulstein
Project manager

ISBN 82-577-3594-9

Dag Berge
Head of research department

How can Norwegian monitoring programmes for lakes and rivers be used to fulfil the requirements of EUROWATERNET?

Eirik Fjeld
Arne Henriksen
Tor S. Traaen

Preface

The European Environmental Agency (EEA) has a political mandate from the EU Council of Ministers to produce objective, reliable and comparable information to allow the Commission, Member States and the general public to judge the effectiveness of policy and the needs for policy development. The European Topic Centre on Inland Waters (ETC/IW) has designed and tested an information and monitoring network, called EUROWATERNET. This network will provide the EEA with information that it needs to meet the requirements of its customers including the European Commission, other policy makers, national regulatory bodies and the general public. Information is required on the following topics:

- the status of Europe's inland water resources, quality and quantity (status and trend assessments)
- how that status relates and responds to pressures on the environment (cause-effects relationships)

Technical Guidelines for Implementation of EUROWATERNET has been worked out, which gives the definition and concept of the network, as well as guidelines for selecting rivers, lakes and groundwater stations. The proposed Technical Guidelines have been tested for selection of lakes and rivers based on existing Norwegian monitoring programmes.

Oslo, January 1999

Merete Johannessen Ulstein

Contents

Summary

The European Topic Centre on Inland Waters (ETC/IW) has designed and tested an information and monitoring network, called EUROWATERNET. Technical Guidelines for Implementation of EUROWATERNET has been worked out, which gives the definition and concept of the network, as well as guidelines for selecting river stations, lakes and groundwater. The proposed Technical Guidelines have been tested for selection of lakes and rivers based on existing Norwegian monitoring programmes. The guidelines propose a lake density of 1 lake per 1 750 km^2. For Norway this equates 185 lakes, representing 1% of the lake population larger than 0.1 km^2. A statistically random sample of such a size would likely not include any reasonable number of eutrophic lakes. Such a sampling regime is therefore not particularly well suited to monitor changes in the “tails” of the statistical distributions of interest. In order to achieve a more optimal monitoring with regard to trends in eutrophication, we suggest that the Norwegian monitoring programme should be based on a sample of lakes with bias towards eutrophic lakes. The national lake survey of 1995 (n = 1005) gives the water chemistry characteristics representative for the whole lake population in Norway. Because the current statistical properties of important lake characteristics are known, we may base the further monitoring on a smaller dataset with bias towards eutrophic lakes, and adjust for the non-random sampling to calculate statistics representative for the whole lake population. In this report we describe a method for adjusting for the non-random sampling in order to provide representative water chemistry statistics (total nitrogen and phosphorus). The monitoring lakes (n = 179) were selected in order to give a good coverage also of mesotrophic and eutrophic lakes. This dataset is well suited for monitoring changes in both eutrophic and oligotrophic lakes, whereas eutrophic lakes only will be scarcely represented in a randomly chosen sample of equal size. In order to calculate representative water chemistry statistics (such as means, percentiles, frequency distributions etc.) from a biased dataset, over-represented types of lakes must be given less weight than under-represented lakes. We used bivariate Gaussian kernel density estimates in order to obtain smooth frequency distributions of N and P for the two datasets. Kernel density estimates (KDE) can be regarded as smoothed histograms, and are well suited for comparing bivariate (or multivariate) data. These frequency distributions were then used to estimate different weights for the lakes in the monitoring lake data set. By employing these weights in the statistical calculations, we calculate water chemistry statistics representative for the whole population of Norwegian lakes – based on data from a biased sample of monitoring lakes. For the time being, it is not possible to provide a selection of Norwegian monitoring sites for rivers that can be claimed to be representative for the Norwegian river population. To design a river monitoring program similar to the lake monitoring program we would need to know the properties of a an unbiased sample of the river population in Norway.

1. Introduction

The Technical Guidelines for Implementation of EUROWATERNET (Nixon et al. 1998) gives the definition and concept of this information and monitoring network, as well as guidelines for selecting river stations, lakes and groundwater.

It is stated that "*Monitoring is expensive and is unlikely to be undertaken purely for the "European need". It is essential therefore that the EUROWATERNET is firmly based on national programmes as the 'zero' or 'low cost' option. By and large national networks are likely to be more than adequate (in terms of numbers of stations, frequency of monitoring and determinands monitored) to meet the EEA needs.*" It is further stated that "*The immediate aim is for Member Countries to establish basic networks for rivers and lakes, and to test the proposal for the groundwater network*".

We have tested the proposed Technical Guidelines given for selection of lakes and rivers based on existing national monitoring programmes.

2. The lake surveys

National lake surveys were conducted in Sweden, Finland and Norway in the autumn 1995 on the initiative of the environmental authorities. The project was subsequently expanded to include Denmark, Russian Kola, Russian Karelia, Scotland and Wales (Henriksen et al. 1998).

5 690 lakes of a total of approximately 155 000 lakes (>0.04 km^2) in the study area were sampled. In most countries the lakes were selected statistically from national lake registers. This enabled the calculation of distribution curves for all measured parameters, representing the whole lake population in each country. This selection should be in good agreement with the EEA recommendations. In the following we refer to the Norwegian lakes of this dataset as "the national lake survey" (n = 1005), and we regard it – for all practical means – as unbiased with respect to eutropication parameters.

For Norway there are 38 845 lakes with a surface area larger than 0.04 km^2 and of these 17 627 lakes have surface area larger than 0.1 km^2, the lower size limit given in the Guidelines. Of these only 7 have surface areas larger than 100 km^2. The guidelines propose a lake density of 1 lake per 1 750 km^2. For Norway this equals 185 lakes, representing 1% of the lake population larger than 0.1 km^2.

Norway has two impact monitoring networks for lakes, one for monitoring acidification and one for monitoring eutrophication. It is quite obvious that a pure statistical selection of monitoring sites would give a very inefficient monitoring network to reveal trends and cause/effect relationships both for eutrophication and for acidification. The routine monitoring of acidification of lakes in Norway consist of 200 lakes that are sampled one time per year during/soon after fall overturn. The lakes were selected to be sensitive and are distributed in a gradient of deposition impact, with more lakes in high impact areas than reference areas. The network also ensures that all main regions are represented in the selection.

For monitoring of eutrophication, 400 lakes were selected, mostly on the basis of population density in the catchment area. The yearly monitoring (4–6 samples during the growing season) involves a selection of approximately 30–50 of these lakes. Hence, the total number of monitored lakes in Norway exceeds the recommended density. The network is, however, not a "general network", but consists of two separate impact networks. Sampling frequency and monitoring parameters also reflect the needs of the specific impact assessment. This makes the use of one common network less suitable.

If we consider the lowest total phosphorus value used in the classification criteria in the Dobris report (EEA, 1995), 25 µg P l^{-1}, less than 2% of the Norwegian lakes exceeds this value. For Norway this implies that one hundred randomly selected lakes are likely to give 2 eutrophic lakes (and it might give zero!). This does not, however, mean that eutrophication is a negligible problem in Norway. There are probably 500 to 1 000 eutrophic lakes in Norway, approximately equal to the total number of lakes in Denmark! These lakes are frequently situated in areas having high population density and are easily accessed for recreation (boating, fishing, bathing etc.) and are even used for drinking water supply. Hence, the water quality of these lakes has a high attention by the public, implying a high public pressure for monitoring and remedial action. One might add that the median value for Norwegian lakes is 2.5 µg P l^{-1}, so whatever improvement one make in some hundred eutrophic lakes it will not be detected in a reasonable number of randomly selected sites. The median value of P for Norwegian lakes is not likely to change in the next decades or even century, whatever measures are taken.

How can we then use the existing lake monitoring programs to evaluate trends in the total lake population in Norway? For this purpose we selected 79 lakes from the acidification program with yearly data back to 1986, and 100 lakes from the eutrophication program evenly distributed in five classes according to total phosphorus concentrations (µg P l^{-1}): <7, 7–11, 11–20, 20–50, >50. The latter selection was also stratified with respect to geographic distribution and lake size.

These lakes ("the monitoring lakes", n = 179) were selected in order to give a good coverage also of mesotrophic and eutrophic lakes. This biased dataset are well suited for monitoring changes in both eutrophic and oligotrophic lakes, whereas eutrophic lakes only will be sparsely represented in a randomly chosen sample of equal size. However, in order to calculate representative water chemistry statistics (such as means, percentiles, frequency distributions etc.) from such a biased dataset, over-represented types of lakes must be given less weight than under-represented lakes. In the following we will describe a method for adjusting for non-random sampling in order to provide unbiased water chemistry statistics of monitored lakes.

3. N and P in the lakes

The concentration of total phosphorous and nitrogen in lakes from the large national survey of 1995 varied between 0.5–181 µg P l^{-1} and 18–5420 µg N l^{-1}, with a median concentrations of 2.5 µg P l^{-1} and 155 µg N l^{-1} (Figure 1 a). Corresponding figures for the monitoring lakes were 0.5–655 µg P l^{-1} and 35–3207 µg N l^{-1}, with a median concentrations of 6.0 µg P l^{-1} and 327 µg N l^{-1} (Figure 1 b).

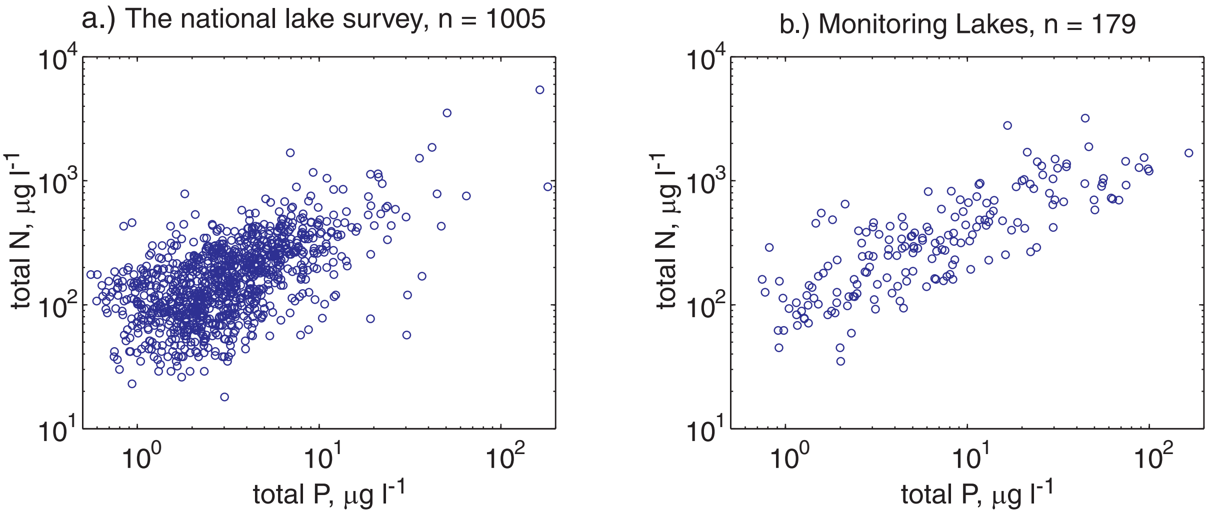


Figure 1. Scatterplots of total phosphorus and nitrogen in the two datasets. The national lake survey is representative for the N and P distribution of Norwegian lakes (area > 0.04 km^2). The Monitoring lakes represent a biased sample with respect to N and P.

4. Kernel density estimation

We used bivariate Gaussian kernel density estimates in order to obtain smooth frequency distributions for the two datasets. Kernel density estimates (KDE) can be regarded as smoothed histograms, and are well suited for comparing bivariate (or multivariate) data. In the univariate case, a kernel density estimate is defined by

$$\hat{f}(x,h) = \frac{1}{nh}\sum_{1}^{n}\phi\left(\frac{x - X_i}{h}\right) \qquad \text{(EQ 1)}$$

where the kernel density function ϕ usually is the standard normal density

$$\phi = \left(1/\sqrt{2\pi}\right)\exp\left(-x^2/2\right) \qquad \text{(EQ 2)}$$

The parameter h is called the window size and determines the amount of smoothing that is applied to the data. Larger values of h produce a smoother density estimate. We can think of Eq. 1 as a situation where we add up the areas of n little Gaussian density curves centred at X_i, each having a standard deviation of h. The result will be a smooth curve which gives us the kernel density estimate of the data.

In a bivariate situation a kernel density estimate is defined as

$$\hat{f}(x,y,h) = \frac{1}{nh_1h_2}\sum_{1}^{n}\phi\left(\frac{x - X_i}{h_1}, \frac{y - Y_i}{h_2}\right) \qquad \text{(EQ 3)}$$

where h_1 and h_2 are the windows sizes in the X and Y directions, respectively.
The criteria for selecting the optimal windows sizes is not well developed for the bivariate case. A to large window size will over-generalise or over-smooth, while a to small size will under-smooth the data. The window size must therefore be chosen carefully.

For our estimates we have used the Kernel Density Estimation Toolbox by Beardah (Nottingham Trent Univ., UK) implemented in MATLAB (Mathworks. Inc.), and the above description of KDE is partly based on Baerdah and Baxter (1995).

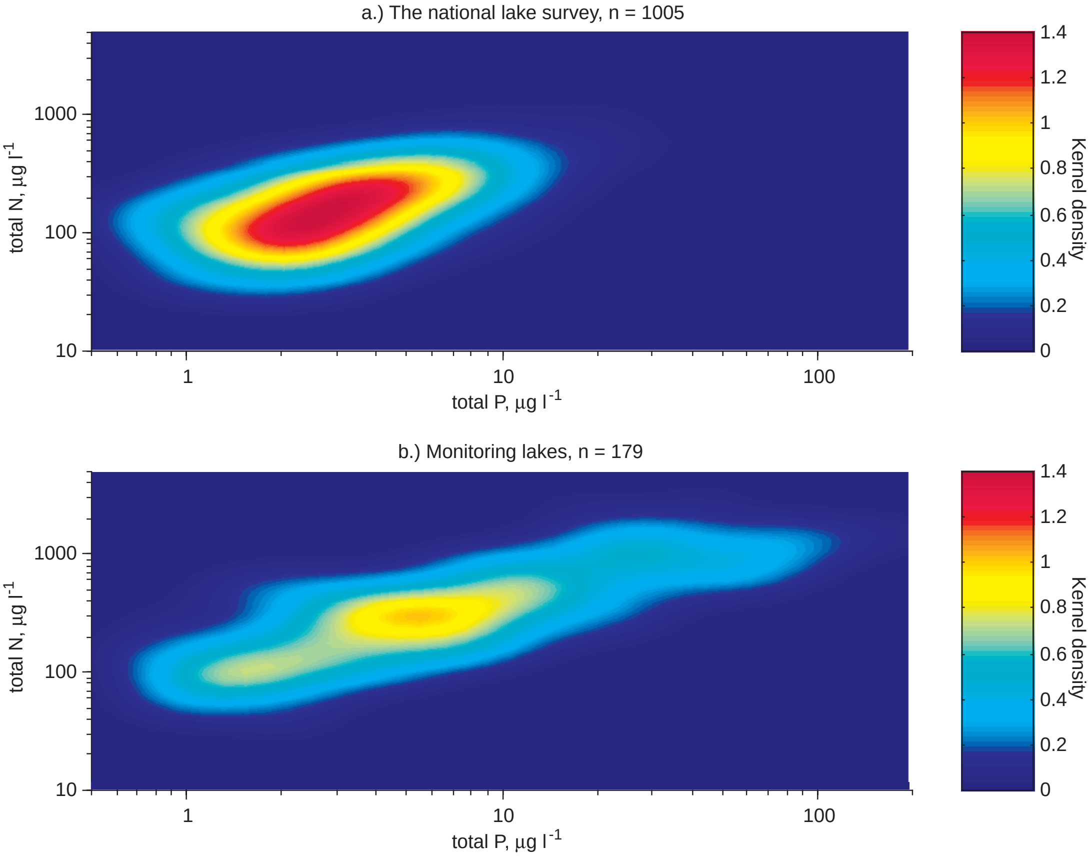


Figure 2. Kernel density estimates of the distribution of total phosphorus and nitrogen in the lakes from the national lake survey (a.), and from monitoring lakes dataset (b.).

The bivariate kernel density estimates of the two datasets are given in figure 2. The kernel density estimates can be thought of as relative frequency distributions, and these figures clearly demonstrate the very different distributions of total N and P in the two datasets. An alternative view (three-dimensional contourplot) of the two distributions is given in Figure 3 and 4.

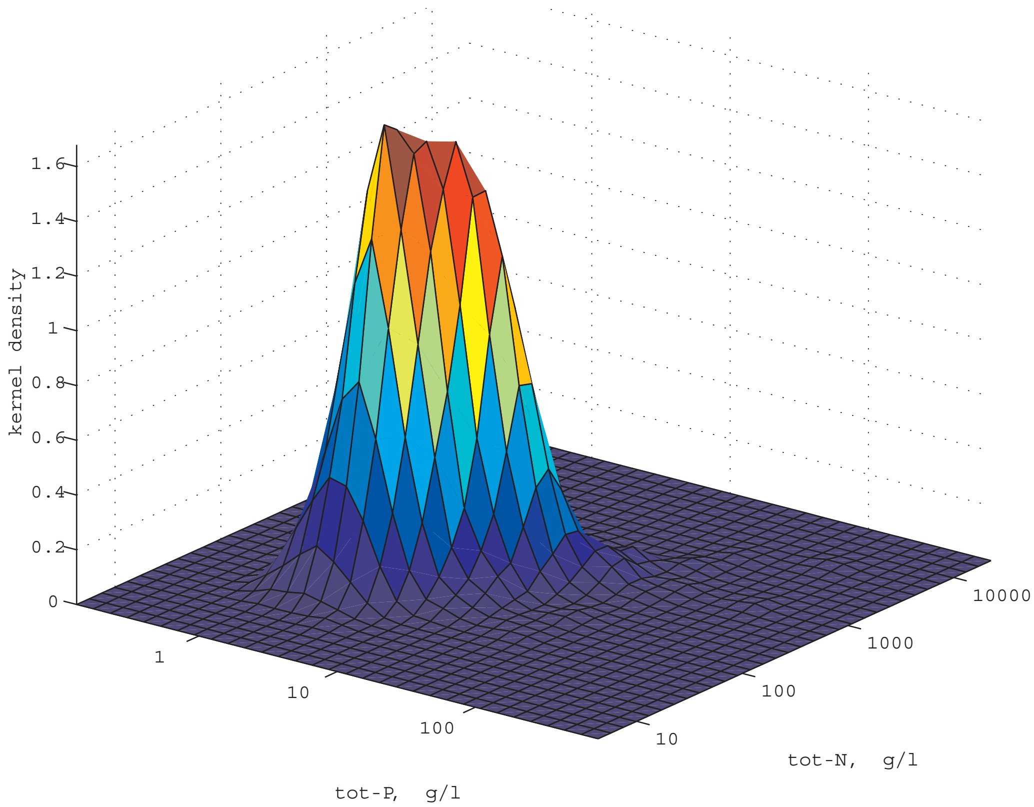


Figure 3. Kernel density estimates of the national lake survey dataset (unbiased sampling, n = 1005).

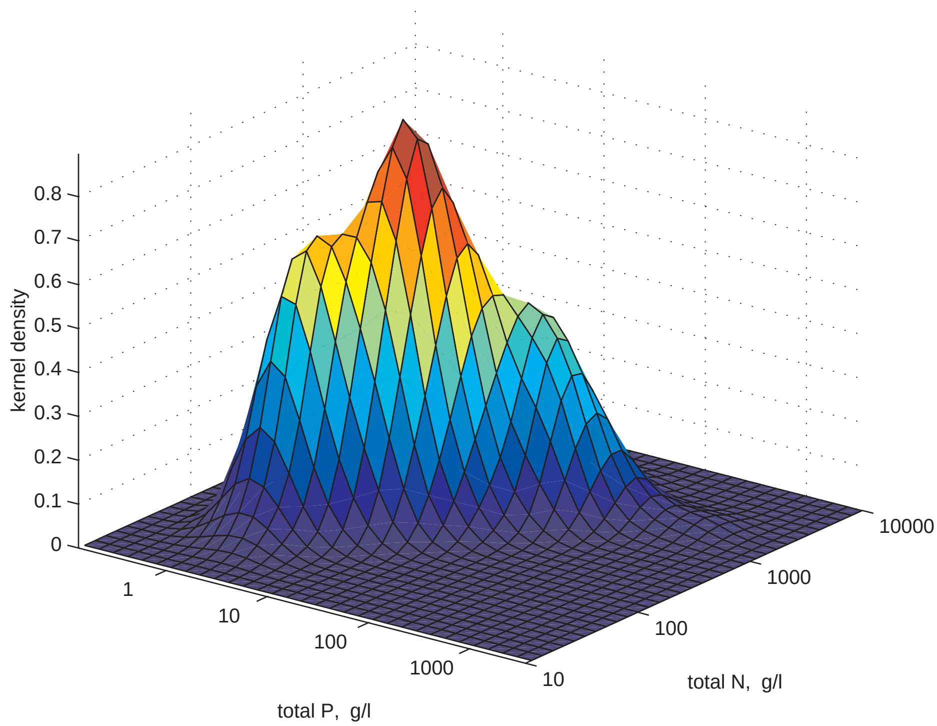


Figure 4. Kernel density estimates of the monitoring lakes dataset (biased sampling, n = 179)

5. Adjusting for biased sampling

In order to use the lakes from the biased dataset to provide descriptive statistics representative for the whole population of Norwegian lakes, the influence of over-represented and under-represented types of lakes must be adjusted in such a way that its frequency distribution mirrors that of a random chosen dataset. Such adjustments can be obtained by giving the lakes from the biased dataset different weights in the statistical calculations.

We can use the kernel density distributions from the two datasets to obtain such weights. The weight for a lake with a given concentration of total N and P (x_i,y_j) is calculated as

$$w(x_i, y_j) = \frac{\hat{f}(x_i, y_j)^*}{\hat{f}(x_i, y_j)} \qquad \text{(EQ 4)}$$

where the nominator and denominator are the kernel density estimates based on the distributions of from the unbiased (the reference survey) and biased survey, respectively. A lake with a kernel density estimate which is twice as high as compared to the reference survey will therefore have a weight of 0.5.

When used in statistical calculations, it is often convenient to scale the weights so that they sum to the number of lakes included in the dataset, which in this case are 179. These weights can be used to adjust for the biased sampling in the stratified survey in order to produce a frequency distribution and other basic statistics that are representative for the whole population of Norwegian lakes (see the figure 5), or in statistical hypothesis testing with regard to future monitoring.

This approach illustrates how existing monitoring network can be used to detect future changes in the total population of lakes. This approach, however, can only be used if the distribution of the whole lake population is known.

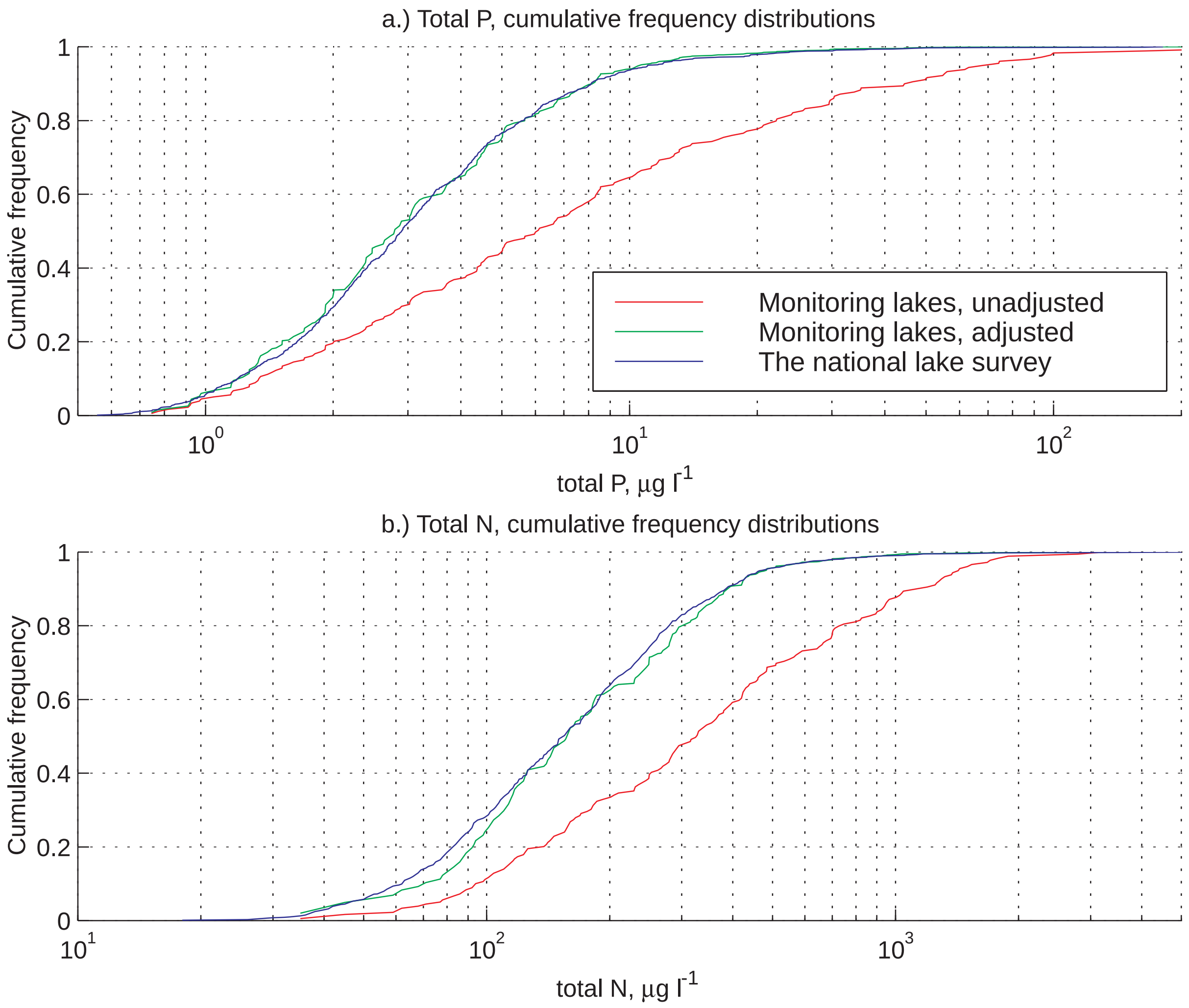


Figure 5. Cumulative frequency distributions for total phosphorus and nitrogen for the biased (the monitoring lakes) and unbiased (the Thousand lakes Survey) datasets. The adjusted distribution is based on weighted values in order to adjust for the effect of the biased sampling scheme. The unadjusted percentiles are based on non-weighted data.

6. Rivers

Also for rivers "the basis of EUROWATERNET is the information derived from existing national and/or regional monitoring networks within each Member Country". Member Countries are asked to select rivers and river stations according to the criteria described in the Guidelines.

For the time being, it is not possible to provide a selection of Norwegian monitoring sites for rivers that can be claimed to be representative for the Norwegian river population. To design a river monitoring program similar to the lake monitoring program we would need to know the properties of an unbiased sample of the river population in Norway.

7. REFERENCES

Nixon, S.C., Grath, J. and Bøgestrand, J 1998. EUROWATERNET The European Environment Agency's Monitoring and Information Network for Inland Water Resources. Technical Guidelines for Implementation. ETC/IW Ref: P031/98/1, June 1998. WRc, Medmenham, UK.

European Environment Agency (EEA) 1995. Europe’s Environment: The Dobris Assessment. Luxemborg: Office for Official Publications of the European Communities. 1995 - XXVI, 712 pp.

Henriksen, A., Skjelkvåle, B.L., Mannio, J., Wilander, A., Harriman, R., Curtis, C., Jensen, J.P., Fjeld, E. and Moiseenko, T. 1998. Northern European Lake Survey – 1995. Finland, Norway, Sweden, Denmark, Russian Kola, Russian Karelia, Scotland and Wales. Ambio 27, 80-91.

Baerdah, C.C. and Baxter, M.J. 1995. MATLAB routines for kernel density estimation and the graphical representation of archeological data. Nottingham Trent University Department of Mathematics Research report 2/95.