VATTEN 64:259–272. Lund 2008

# VARIABILITIES AND UNCERTAINTIES IN IMPORTANT COASTAL WATER VARIABLES

*by ANDREAS C. BRYHN, LARS HÅKANSON and JENNY M. EKLUND*
*Dept. of Earth Sciences, Uppsala Univ., Villav. 16, 752 36 Uppsala, Sweden*
*corresponding author, e-mail: andreas.bryhn@geo.uu.se*



## Abstract

The focus of this work is on general patterns in uncertainty as well as temporal and spatial variability in key water variables in coastal science and management. These patterns are essential since they regulate how many samples must be taken to get reliable mean or median values characterising coastal water quality and which variables are most suitable for monitoring and predictive modelling. We present results concerning coefficients of variation, correlations, regressions, variations in data from different time periods, and confidence intervals for empirical mean values. We use data from Ringkøbing Fjord (Denmark, N. Europe), Chesapeake Bay (Eastern U.S.) and other coastal marine sites to illustrate the basic principles related to patterns in variability. We have shown that total and particulate N and P generally have much lower coefficients of variability (CV) than dissolved inorganic nutrient fractions. The latter are, hence, of limited use in predictive models for coastal management. Total nitrogen (N) and phosphorus (P) were, on the other hand, found to be useful predictors of two standard bioindicators, the Secchi depth (a measure of water clarity) and chlorophyll-a concentrations (a measure of phytoplankton biomass or production).

*Key words* – Variability, uncertainty, water variables, CV, predictive power, coastal areas, management

## 1. Introduction and aim

The variability and representativity of variables is a major concern in understanding and interpreting changes in ecosystems and in developing practically useful predictive models. Even if it would be possible to refine empirical analyses to the extent that measurement uncertainty would be completely eliminated, the natural variability due to physical, chemical and biological processes would still severely limit our ability to explain and predict environmental changes. Thus, data scarcity, variability and uncertainty are important constraints in coastal management (de Jonge, 2000; Irvine, 2004). The basic aim of this work is to analyse changes in a set of standard water variables using simple, operational statistical methods.

Variables may vary both within and among systems and the magnitude of these variations influences the predictive power of both empirical and mechanistically-based dynamic models. During periods with changing wind conditions, most bioindicators (such as Secchi depth and chlorophyll-a concentrations) are likely to vary much both horizontally and vertically (see, e.g., Bloesch, 1995; Weyhenmeyer, 1996; Moldan and Billharz, 1997; Bortone, 2005). This means that there will be large uncertainties in single measurements. The same argument is valid not just for bioindicators in coastal systems but for all water variables in natural aquatic systems (Comin et al., 2004).

This work focuses on coefficients of variation (CV = SD/MV, SD = standard deviation, MV = mean value), variations in data from different time periods, correlations and regressions between coastal water variables, and confidence intervals for empirical mean values. Data from several coastal areas in Europe and North America are compared, although much of the work is based on data from Chesapeake Bay, United States (Cooper and Brush, 1993; Boesch et al., 2000), and Ringkøbing Fjord, Denmark (Nielsen et al., 2004; Petersen et al., 2008; Håkanson et al., 2007a) because for these areas there are extensive data sets available.

For empirical data to be meaningful in management contexts, they need to convey statistically significant information that can also be used to in, e.g., model comparisons (Peters, 1991). A basic aim of this study is to investigate the extent to which water variables fluctuate at different sites and at different times of the year, and to test which variables are best suited for predicting the Secchi depth and chlorophyll concentration. Predicting such easily measured and interpretable operational effect variables accurately is essential in estimating the probable outcome of various coastal management strategies (Peters, 1991).

## 2. Brief background on the study sites and sampling programs

The two main sites of this study are Ringkøbing Fjord and Chesapeake Bay. Mean values and standard deviations, which will be discussed, describe conditions in entire ecosystems (= defined coastal areas rather than at individual sampling sites), since there is a well documented interest in analyzing conditions on an ecosystem scale (Arhonditsis et al., 2000; Boesch et al., 2000; Petersen et al., 2008).

Ringkøbing Fjord, Denmark (56°0' N, 8°15' E), is a lagoon which in many ways may be regarded as a brackish lake since it is only connected to the sea through a narrow sluice. The area of the lagoon is 300 $km^2$, its mean depth is 2 m, and its maximum depth is 5 m. It has undergone several dramatic ecological regime shifts during recent decades (Håkanson et al., 2007a; Petersen et al., 2008). The latest regime shift occurred in 1996 and was demonstrated by marked changes in median annual values for total phosphorus (TP) concentrations, suspended particulate matter (SPM), Secchi depth, chlorophyll-a concentrations, salinity, macrophyte cover and biomass of soft clams. Water chemistry data were registered as averages between surface and bottom waters during 1980–1986. During 1987–1997, integrated vertical sampling was performed, while samples from 1998–2004 were taken from 1 m below the water surface. The water column in shallow and exposed Ringkøbing Fjord is subject to frequent mixing by wind and wave action and there are only small concentration differences between monitoring stations.

Chesapeake Bay (37–40° N, 75–77° W), the largest estuary in the U.S.A., is one of the most intensively studied coastal areas of the world since it has been greatly affected by nutrient inputs, intensive fishing and other human activities (Cooper and Brush, 1993; Boesch et al., 2000). It has an area of 11500 $km^2$, a mean depth of 6.5 m and a maximum depth of 36 m. 200 monitoring stations with the most data available were selected from an extensive online data set (Chesapeake Bay Program, 2007). Data were available from several depths, and at 28 of the stations, a clear summer thermocline (a vertical temperature difference of 4°C or more) was detected. Therefore, data were divided into two datasets covering surface waters (200 stations) and bottom waters (28 stations).

## 3. Statistical methods

The methods described in this section have previously been found to be useful in aquatic science and management, partly due to their simplicity. The number of data (n; sampled from the same statistical population) of a certain variable required for acquiring a mean value with high certainty can be calculated from the sampling formula, which is derived from the common t-test (Håkanson, 1984):

$$n = (1.96 \cdot CV/L)^2 + 1 \qquad (1)$$

where L is the error in the mean value at the 95 % confidence level. Alternatively, eq. 1 may instead be used to calculate L, given that n is known from an available monitoring program.

The predictive power is a fundamental concept in aquatic science and management. It expresses the accuracy of model predictions and may be seen as a quantitative indicator of how well we understand scientific relationships (Peters, 1991). A common and simple way to measure the predictive power is to use the coefficient of determination ($r^2$; r = the correlation coefficient) between modelled and empirical data (Håkanson and Peters, 1995; Håkanson, 1999). Yves Prairie (1996) has produced some very useful results illustrating the practical utility of models for predictions of individual y-values. If the uncertainty bands in a regression are wide apart when modelled values are compared to empirical data, then the model can produce useless predictions for individual y-values. The usefulness of the predictions and the $r^2$ value display a non-linear relationship (Prairie, 1996), which conveys that the predictive power of $r^2$ values up to 0.65–0.75 is not very much different from the predictive power at $r^2$=0. However, predictive power increases very rapidly for $r^2$ values higher than 0.75, which underlines the importance of searching for high $r^2$ values.

One way to estimate the highest possible $r^2$ of a predictive model is to compare two empirical samples of the same variable since one can generally not expect models to predict better than what one can measure. The $r^2$ value generated from a regression between such similar samples is abbreviated as $r_e^2$ (Håkanson, 1999). A second method for estimating the highest possible predictive power is to use the reference $r^2$, $r_r^2$. From a statistical point of view, the following equation (see Håkanson, 1999 for a more detailed motivation) gives $r_r^2$ as a function of (1) the number of samples ($n_i$) for each $y_i$-value in the regression, (2) the number of data points in the regression (N), (3) the standard deviations related to all individual data points in the regression and (4) the range of the y-variable:

$$r_r^2 = 1 - 0.66 \cdot CV_y^2 \quad (2)$$

where $CV_y$ (or CV for simplicity) is the characteristic within-site variability for the y-variable. It is valid for actual (non-transformed) y-values.

Regression analyses could be performed for many reasons, e.g., to compare model predictions (on the x-axis) with empirical data (on the y-axis), to test hypotheses about relationships, and to develop statistical/empirical models. Many textbooks examine regression analyses (see, e.g., Draper and Smith, 1966; Cooley and Lohnes 1971; Mosteller and Tukey, 1977; Pfaffenberger and Patterson, 1987; Newman, 1993). One requirement in simple regression analyses is that data series are normally distributed. If not, normality may be achieved by subjecting the data to log-transformations or to other mathematical operations (see Håkanson and Peters, 1995). A multivariate regression includes one y-variable and several x-variables and the criteria used in this study for the x-variables were:

- They had to be widely used in monitoring programs and commonly used as input variables in predictive modelling. This excluded chlorophyll (Chl) and the Secchi depth (Sec), since they are usually treated as y-variables in modelling (Håkanson and Peters, 1995).
- They had to correlate significantly with y-variables in bivariate regressions as well as within the multiple regression.
- These correlations had to be of the same sign; i.e., an x-variable that was positively correlated with a y-variable was excluded if its contribution in the multiple regression was negative, and vice versa.

The last two criteria are necessary in order not to let bivariate and multivariate regressions convey contradictory information.

# 4. Results and discussion

## 4.1. Coefficients of variation and predictive power

As stressed in section 2, there are no clear concentration gradients in Ringkøbing Fjord due to intensive mixing. Table 1 presents randomly selected data from Ringkøbing Fjord on total phosphorus (TP) concentrations to illustrate a fundamental problem related not just to TP-variations, but to changes in most water variables, not just for this period from Ringkøbing Fjord, but generally, for most coastal areas. The table gives mean values (MV), medians (M50), standard deviations (SD) and coefficients of variations (CV) on a monthly basis. The mean value for TP for April is 102 and for May 60 µg/l. Given a CV-value of 0.94 (for April in table 1), eq. 1 conveys that one would need 340 samples to determine the true mean value with an error less than 10 % of the mean (L = 0.1). However, for TP there are, on average, only 6 samples available for each month in this period (1982) from this coastal area where a very comprehen-

*Table 1. Example of how the available data from the three monitoring stations in Ringkøbing Fjord were analysed regarding TP-concentrations (in µg/l) in 1982. MV = mean value, M50 = median, n = number of samples, SD = standard deviation, CV = coefficient of variation (SD/MV), L = error in percent of the mean value (with a 95 % probability) and values for the 95 % confidence interval for the mean values (plus and minus). Horizontal bars connect rows which contain data from the same month.*

| Date | St. 1 | St. 2 | St. 3 | MV | M50 | SD | CV | n (L=0.1) | L (%) | plus 95 % CI | minus 95 % CI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 82-02-10 | 210 | 100 | 90 | 133 | 100 | 67 | 0.50 | 97 | 69 | 266 | 0.2 |
| 82-04-01 | 30 | 230 | 220 | 102 | 45 | 96 | 0.94 | 340 | 83 | 293 | –90 |
| 82-04-20 | 50 | 40 | 40 | | | | | | | | |
| 82-05-12 | 80 | 50 | 50 | 60 | 50 | 17 | 0.29 | 33 | 40 | 95 | 25 |
| 82-06-02 | 50 | 60 | 70 | 53 | 60 | 23 | 0.42 | 69 | 37 | 98 | 8 |
| 82-06-29 | 10 | 70 | 60 | | | | | | | | |
| 82-07-22 | 110 | 10 | 10 | 43 | 10 | 58 | 1.33 | 681 | 185 | 159 | –72 |
| 82-08-04 | 40 | 110 | 70 | 88 | 75 | 56 | 0.64 | 158 | 56 | 201 | –24 |
| 82-08-24 | 190 | 80 | 40 | | | | | | | | |
| 82-09-09 | 30 | 170 | 190 | 130 | 170 | 87 | 0.67 | 173 | 93 | 304 | –44 |
| 82-10-25 | 170 | 90 | 200 | 153 | 170 | 57 | 0.37 | 54 | 51 | 267 | 40 |
| 82-12-09 | 120 | 170 | 160 | 150 | 160 | 26 | 0.18 | 13 | 24 | 203 | 97 |

Table 2. *Compilation of CV-values (their mean values and standard deviations) for different water variables and different time periods from Ringkøbing Fjord for the period 1980 to 2004.*

| | Monthly (n ≈ 6) | Spring (n ≈ 18) | Winter (n ≈ 18) | Summer (n ≈ 18) | Fall (n ≈ 18) | 3-months (n ≈ 18) | Year (n ≈ 72) | All (n ≈ 1500) |
|---|---|---|---|---|---|---|---|---|
| *Sal* | | | | | | | | |
| MV | 0.10 | 0.19 | 0.10 | 0.15 | 0.16 | 0.15 | 0.23 | 0.35 |
| SD | 0.07 | 0.10 | 0.07 | 0.07 | 0.08 | 0.09 | 0.08 | – |
| *TP* | | | | | | | | |
| MV | 0.40 | 0.46 | 0.37 | 0.34 | 0.47 | 0.40 | 0.48 | 0.67 |
| SD | 0.26 | 0.19 | 0.18 | 0.22 | 0.16 | 0.20 | 0.13 | – |
| *Chl* | | | | | | | | |
| MV | 0.37 | 0.53 | 0.28 | 0.36 | 0.56 | 0.42 | 0.52 | 0.87 |
| SD | 0.21 | 0.19 | 0.15 | 0.20 | 0.38 | 0.26 | 0.14 | – |
| *SPM* | | | | | | | | |
| MV | 0.40 | 0.75 | 0.46 | 0.54 | 0.64 | 0.59 | 0.71 | 0.78 |
| SD | 0.18 | 0.37 | 0.23 | 0.24 | 0.24 | 0.30 | 0.28 | – |
| *TN* | | | | | | | | |
| MV | 0.16 | 0.33 | 0.14 | 0.25 | 0.18 | 0.23 | 0.40 | 0.51 |
| SD | 0.10 | 0.10 | 0.07 | 0.14 | 0.11 | 0.13 | 0.14 | – |
| *Sec* | | | | | | | | |
| MV | 0.22 | 0.38 | 0.21 | 0.29 | 0.40 | 0.31 | 0.38 | 0.71 |
| SD | 0.13 | 0.15 | 0.10 | 0.20 | 0.15 | 0.17 | 0.14 | – |
| *Temp* | | | | | | | | |
| MV | 0.15 | 0.52 | 0.12 | 0.44 | 0.56 | 0.39 | 0.56 | 0.56 |
| SD | 0.15 | 0.14 | 0.04 | 0.15 | 0.27 | 0.23 | 0.09 | – |

sive monitoring program is carried out. With 6 available monthly data, one has to accept that the error for the calculated mean value in April is 83 % and that the 95 % confidence interval for the individual data from this month varies between –90 and +293 µg/l. The data scarcity in combination with the variability imply that one would need many more samples than available to be able to draw any statistically significant conclusions on whether there is a change in TP-concentrations between months 4 and 5 in this example.

Table 2 compiles results regarding salinity, TP, chlorophyll-a, suspended particulate matter (SPM), total nitrogen (TN), Secchi depth and water temperature, and how these variables varied from 1980–2004. CV-values are different for different time periods and CVs have therefore been calculated for monthly data (there are about 5–9 samples available for most months), spring (months 3, 4 and 5), winter (months, 12, 1 and 2), summer (months 6, 7 and 8) and fall (9, 10 and 11). We also calculated CVs based on monthly values for all 3-month periods, as well as CVs based on values for each year and CVs for all the data. This makes it possible to identify some interesting patterns in variability; i. e., to distinguish whether there are periods when the variability is smaller so that more representative, less uncertain, mean or median values can be determined. Some of those results are compiled in fig. 1 for TP, water temperature, SPM and salinity. From table 2 and fig. 1, one can note:

- The CVs based on monthly data are very high for SPM and TP (0.40), high for chlorophyll (0.37), lower for Secchi depth (0.22), and low for TN (0.16), temperature (0.15) and salinity (0.10).
- The CVs based on annual data are even higher: 0.71 for SPM (a very high value), between 0.5 and 0.7 for chlorophyll, TP and temperature, lower for TN, Secchi depth and lowest for salinity (0.23).

If there are only 6 samples (2 from each of the 3 monitoring stations) available for a month and hence about 12·6 = 72 samples for annual data, one can see that the mean or median annual values may be determined with less statistical uncertainty than the mean/median monthly values, in spite of the fact that the CV-values are higher for annual data. For example, for monthly data on TP, CV = 0.4, n = 6 gives that the mean value can be estimated with an error (L) less than 36 % of the mean (with a 95 % certainty). For annual data, CV = 0.48, n = 6·12 (24 samples from each of the 3 monitoring stations) gives that the error (L) around the mean value is 11 % of the mean. So, given the marked variability in most of these water variables, and the relatively

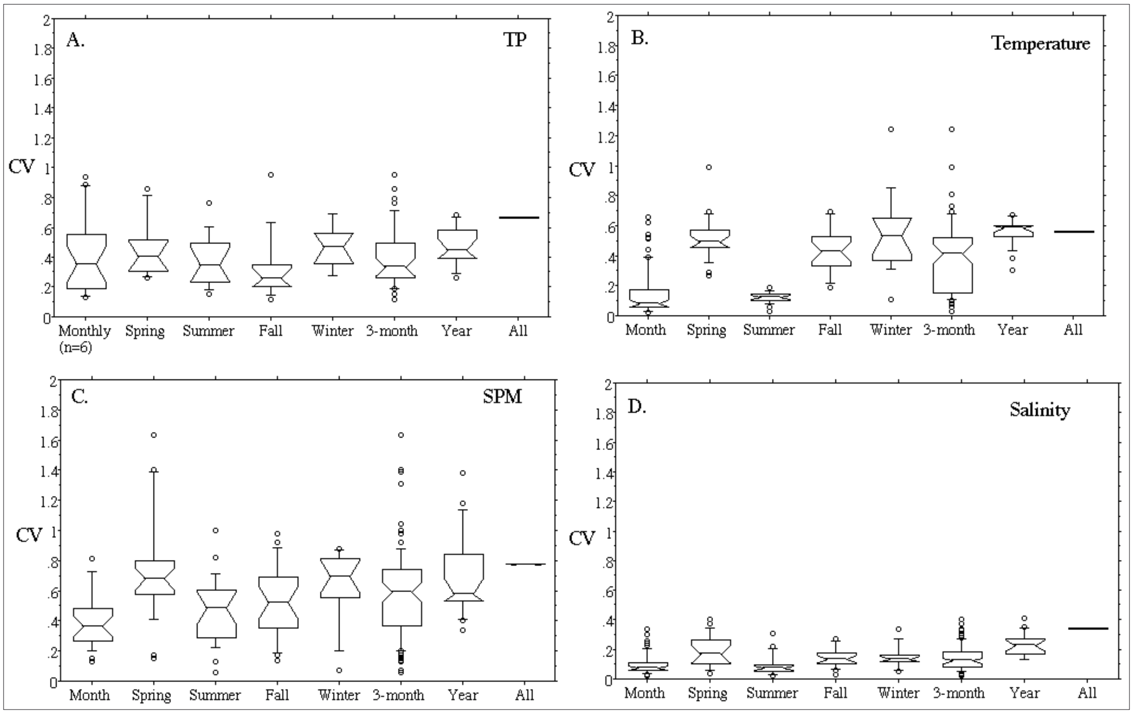


Fig. 1. *Box-and-whisker-plots (showing medians, quartiles, percentiles and outliers) for coefficients of variation (CV) for (A) total phosphorus, (B) water temperature, (C) suspended particulate matter (SPM) and (D) salinity from Ringkøbing Fjord based on monthly data (there are on average 6 samples available for each month), seasonal data from spring (months 3, 4 and 5), summer (months 6, 7, 8), fall (months 9, 10 and 11) and winter (months 12, 1 and 2), based on all 3-month periods, based on annual samples (from 14 years) and based on all samples from the period 1980 to 2004.*

few monthly data available, it becomes important to focus on long-term changes on a seasonal and/or an annual basis. Fig. 2 gives a compilation of such annually-based CVs for salinity (smallest CV), Secchi depth, TN, TP, chlorophyll, temperature and SPM (largest CV). These CVs will also be used in the following to provide confidence intervals for the empirical data.

Is the variability in water variables smaller or greater in Ringkøbing Fjord than in other coastal areas? Table 3 gives a comparison between average CV-values for open water areas (in the Baltic Sea) and coastal areas (in the Baltic Sea). Generally, the CVs for SPM, chlorophyll, TN and TP are about a factor of 2 higher in Ringkøbing Fjord than in Baltic coastal areas, and about the same as at individual sites in the open Baltic Sea. The main reason for this very high variability in Ringkøbing Fjord may seem like a paradox: The lagoon is comparatively large and shallow and dominated by resuspension events. This means a considerable mixing, but also great temporal and spatial variability especially before, during and after storms, in particular for variables such as SPM, which, by definition are related to many cycles of sedimentation and resuspension (Håkanson, 2006).

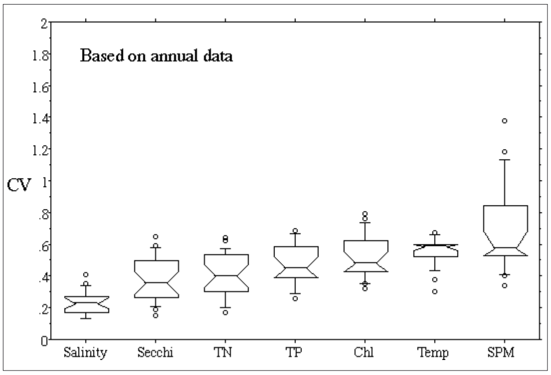


Fig. 2. *Box-and-whisker-plots (showing medians, quartiles, percentiles and outliers) for coefficients of variation based on annual data for salinity, Secchi depth, total nitrogen concentrations (TN), total phosphorus (TP), chlorophyll-a (Chl), water temperature and suspended particulate matter concentrations (SPM) based on data from Ringkøbing Fjord 1980 to 2004. Data from all three monitoring stations have been included.*

Table 3. *Characteristic summer or growing season coefficients of within-site or within-system variation (CV), highest reference $r^2$-values ($r_r^2$; see eq. 2), and the required number of samples (n) if the mean value error (L) should be 10 % or less. Variables from (A) sites in the open brackish Baltic Sea (months 6–10; based on Håkanson and Eckhell, 2005), (B) brackish coastal areas in the Baltic (months 6–8; based on Wallin et al., 1992 and Nordvarg, 2001), and (C) Ringkøbing Fjord (months 6–8). The context of the table is explained in the running text.*

| | CV | $r_r^2$ | n (L=0.1) |
|---|---|---|---|
| *A. Baltic open water sites* | | | |
| Sal | 0.07 | 0.997 | 3 |
| Temp | 0.40 | 0.89 | 62 |
| SPM | 0.67 | 0.70 | 173 |
| *B. Baltic coastal areas* | | | |
| TN | 0.13 | 0.99 | 7 |
| TP | 0.16 | 0.98 | 11 |
| Sec | 0.19 | 0.98 | 15 |
| Chl | 0.25 | 0.96 | 25 |
| O2 | 0.26 | 0.96 | 27 |
| DIP | 0.28 | 0.95 | 31 |
| DIN | 0.31 | 0.94 | 38 |
| SPM | 0.34 | 0.92 | 45 |
| *C. Ringkøbing Fjord* | | | |
| Sal | 0.15 | 0.99 | 10 |
| TN | 0.25 | 0.96 | 25 |
| Sec | 0.29 | 0.94 | 33 |
| TP | 0.34 | 0.92 | 45 |
| Chl | 0.36 | 0.91 | 51 |
| Temp | 0.44 | 0.87 | 75 |
| SPM | 0.54 | 0.81 | 113 |
| NOx | 0.61 | 0.75 | 144 |
| DIN | 0.71 | 0.67 | 195 |
| Ort-P | 0.76 | 0.62 | 223 |
| NHx | 0.78 | 0.60 | 235 |

Table 4 extends this comparison to surface and bottom waters of Chesapeake Bay, assuming that changes in the whole water body could merit system-wide management action. Data included dissolved inorganic nitrogen (DIN), dissolved nitrogen (DN), dissolved organic nitrogen (DON), dissolved inorganic phosphorus (DOP), dissolved phosphorus (DP), orthophosphate (OrtP), particulate nitrogen (PN) and particulate phosphorus (PP). Organic nutrients had higher CVs in bottom waters than in surface waters, while this relationship was the opposite for inorganic nutrients. In general, CVs were much more similar in Chesapeake Bay to Ringkøbing Fjord than to the Baltic Sea sites. To compare and complement the data given in tables 3 and 4, one can note that Weston et al. (2004) have presented the following CVs for chlorophyll-a for marine open water areas (from the North Sea), yearly CV = 0.68; $r_r^2$ = 0.69 (from eq. 2); and median monthly CV = 0.34; $r_r^2$ = 0.93.

Table 4. *Characteristic coefficients of variation (CV) at one typical (with conditions compared to representative conditions in the whole water body) monitoring station highest reference $r^2$-values ($r_r^2$; see eq. 2), and the required number of samples (n) if the mean value error (L) should be 10 % or less. Data from Chesapeake Bay, months 6–8. Variables from (A) surface waters and (B) deep waters. The context of the table is given in the text.*

| | CV | $r_r^2$ | n (L=0.1) |
|---|---|---|---|
| *A. Chesapeake Bay, surface waters* | | | |
| Temp | 0.08 | 0.995 | 4 |
| Sal | 0.18 | 0.98 | 14 |
| DON | 0.23 | 0.97 | 21 |
| TN | 0.24 | 0.96 | 24 |
| Sec | 0.26 | 0.95 | 28 |
| DN | 0.28 | 0.95 | 30 |
| TP | 0.35 | 0.92 | 49 |
| PN | 0.37 | 0.91 | 53 |
| PP | 0.41 | 0.89 | 66 |
| SPM | 0.55 | 0.80 | 117 |
| DP | 0.56 | 0.79 | 122 |
| Chl | 0.59 | 0.77 | 134 |
| DOP | 0.61 | 0.76 | 142 |
| OrtP | 0.72 | 0.66 | 200 |
| DIN | 0.97 | 0.38 | 360 |
| *B. Deep waters* | | | |
| Sal | 0.12 | 0.99 | 7 |
| Temp | 0.13 | 0.99 | 8 |
| DN | 0.23 | 0.96 | 22 |
| TN | 0.23 | 0.96 | 22 |
| DON | 0.24 | 0.96 | 23 |
| TP | 0.41 | 0.89 | 66 |
| PN | 0.46 | 0.86 | 83 |
| PP | 0.49 | 0.84 | 92 |
| DIN | 0.53 | 0.81 | 111 |
| DP | 0.56 | 0.79 | 123 |
| OrtP | 0.62 | 0.75 | 147 |
| Chl | 0.70 | 0.67 | 192 |
| SPM | 0.73 | 0.64 | 208 |
| DOP | 0.83 | 0.54 | 267 |

Arhonditsis et al. (2000) reported similar data variability from the semi-enclosed Gulf of Gera in the Mediterranean for chlorophyll (taken between June 1996 and October 1997), CV = 0.60; $r_r^2$ = 0.76. Apparently, CVs may be higher in enclosed or semi-enclosed water bodies than in more open waters.

In addition, tables 3 and 4 revealed a clear pattern regarding CVs of several marine water variables, which enabled us to group them into four categories. Sal was always among the variables with the lowest CVs. DN, DON and TN had intermediate CV levels, while PN, PP and TP had high CVs. Chl, DIN, DIP, DOP, DP, SPM, and fractions of DIN and DIP had very high CVs. The large difference in CVs means (1) that variables with low CVs may be predicted with much higher preci-

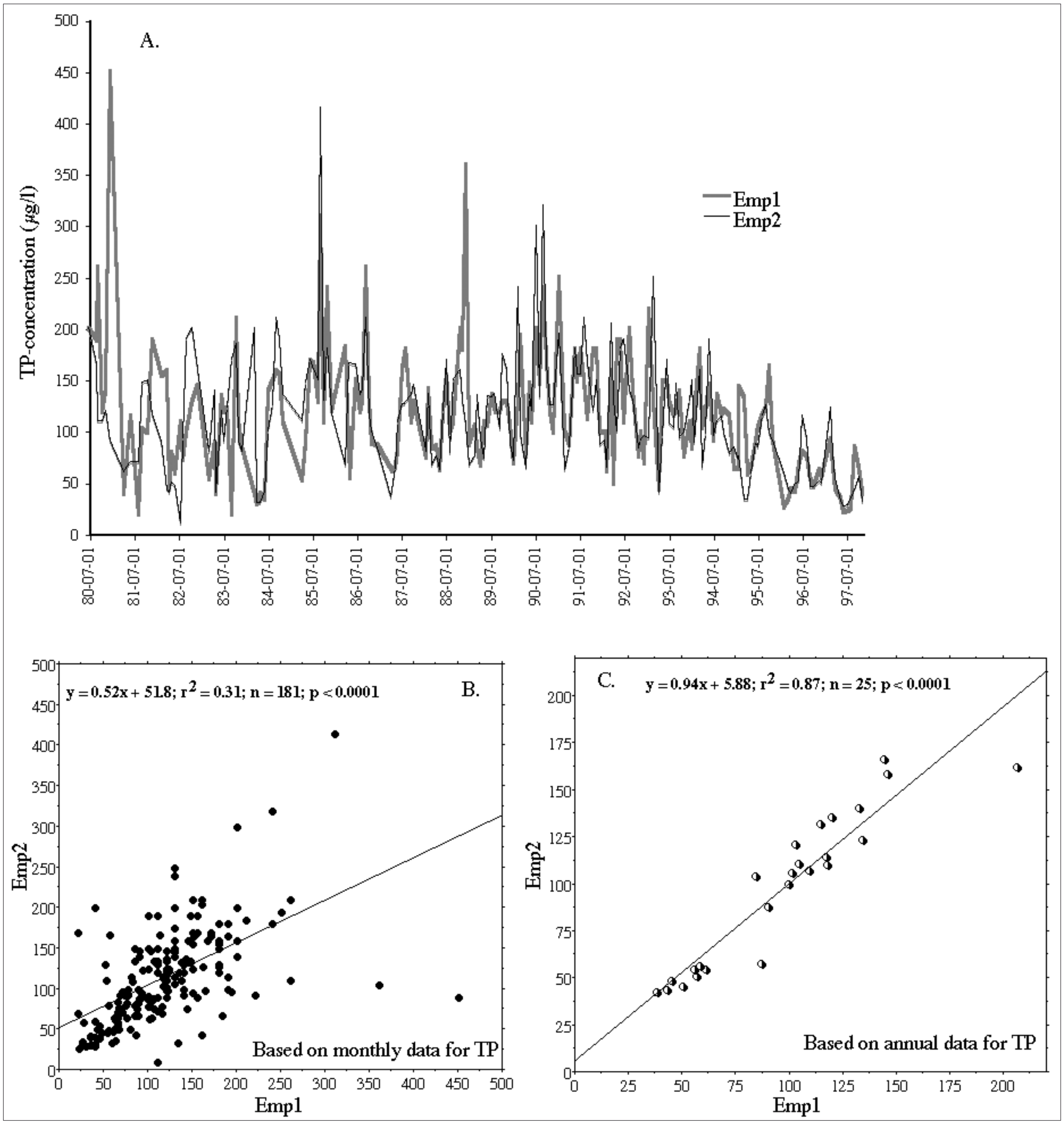


Fig. 3. *A. The correspondence between two sets of randomly selected empirical mean values (Emp1 and Emp2) from monthly samples taken in Ringkøbing Fjord during the period 1980 to 1997. B. The same data as in fig. A but compared by means of a regression giving the regression line, the slope (0.52), the $r^2$-value (0.31), the number of data (n = 181) and the statistical certainty (p < 0.0001). C. The same as in fig. B but based on annual data. This gives the slope = 0.94 and the $r^2$-value = 0.87 (the number of data = 25). The context of the table is explained in the running text.*

sion than those with high CVs given a similar number of samples and (2) that 100–1700 % more samples must be taken from the "very high CV" group of variables compared to the "intermediate CV" group in order to establish a mean value for the growing season with an error of 10 % or less.

To test and illustrate the first method for determining the highest expected predictive power ($r_e^2$, see Section 3), TP-data from Ringkøbing Fjord were used. Fig. 3A shows two parallel series of data on TP-concentrations where 3 data have been randomly selected in the series called Emp1 and 3 other samples used for Emp2; these two samples represent monthly mean values in this well-mixed water body and should represent the same thing. Each data point thus consisted of two mean values of three randomly chosen data each from the six available monthly samples. Then, the $r_e^2$ value (see the methods section) calculated from 181 monthly samples (from the period 1980 to 2004) is 0.31, which is rather low but highly significant ($p < 0.0001$) since the number of data is so large (fig. 3B). The slope of the regression line is 0.52 which is quite different from 1 and underlines the poor correspondence between different monthly mean values. This variability includes analytical and methodological uncertainties but mainly depends on the significant monthly variability for TP (and most other water variables) in this lagoon (and other coastal bays), which is also reflected in the very high characteristic CV for TP in Ringkøbing Fjord. The higher the CV-value, the more difficult it will be to establish representative and reliable empirical mean values of the given variable. From fig. 3C, one can note that very significant improvements in predictive power can be expected for annual values, as compared to monthly values ($r_e^2 = 0.87$ versus $r_e^2 = 0.31$) and the slope of the regression line is much closer to 1 for yearly data (0.94 compared to 0.52). Evidently, the $r_e^2$ value depends on, e.g., the total number of samples in the regression and the number of

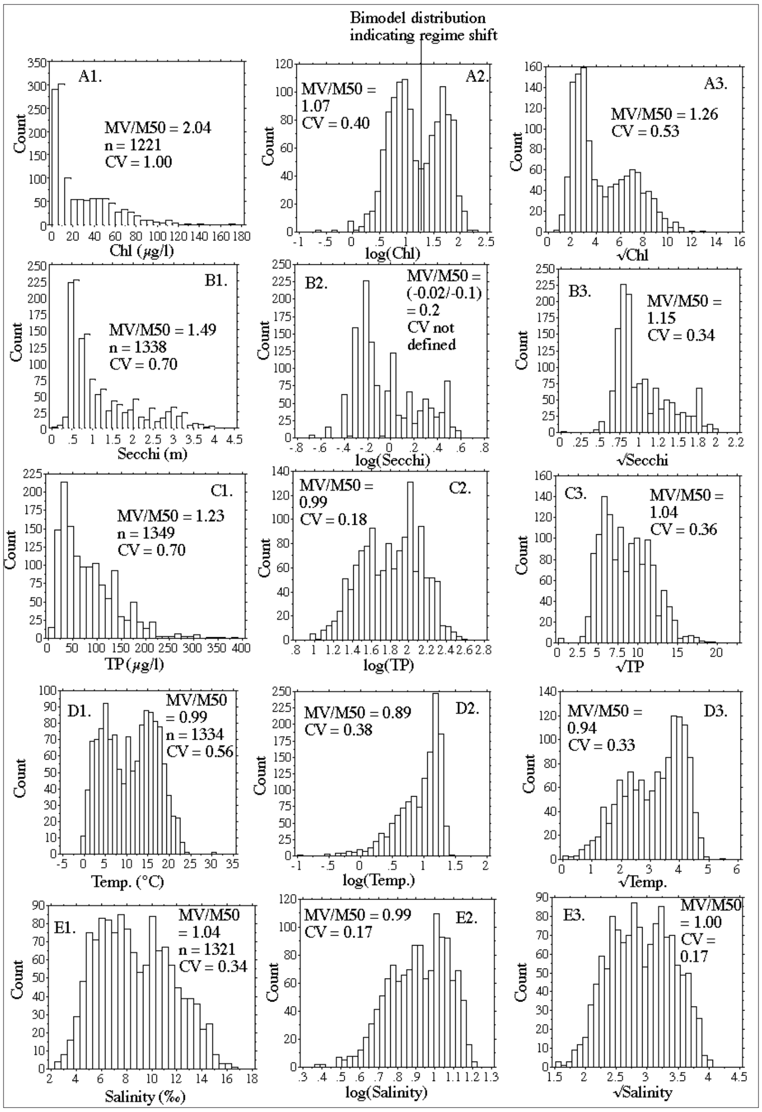


Fig. 4. *Frequency distributions based on all available data from Ringkøbing Fjord for the period 1980 to 2004 for: A. chlorophyll-a concentrations (chl), B. Secchi depths (Sec), C. total-P concentrations (TP), D. temperatures (Temp) and E. salinity. The figure gives information on frequency distributions based on the actual data, for logarithmic transformations of the actual and for exponential transformations (of the type √X), statistics (minimum, maximum, mean and median values) as well on the ratio between mean (MV) and median (M50) value as a measure of the normality of the frequency distribution.*

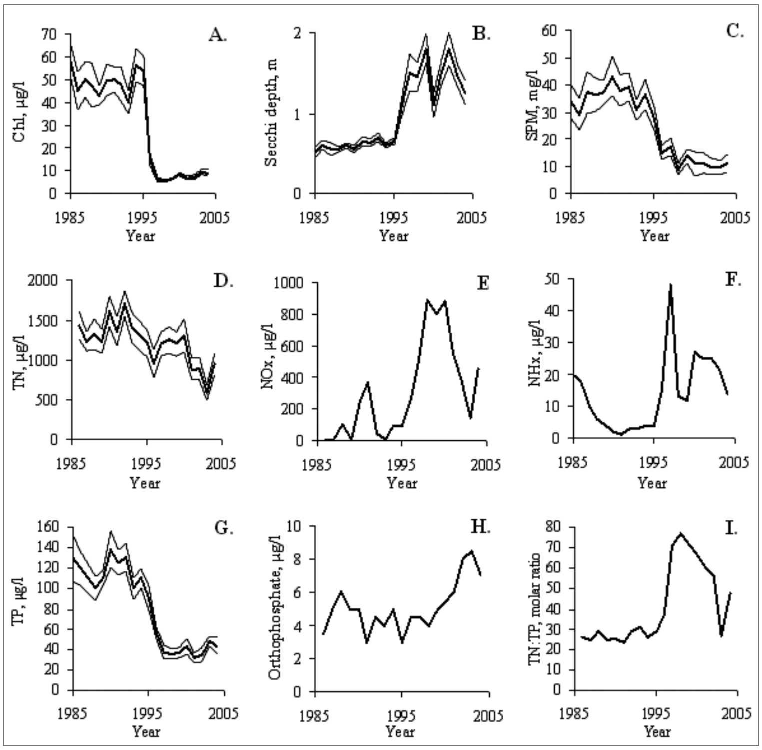


Fig 5. *Nine curves illustrating changes in median annual values (thick lines) in selected variables during the period 1981 to 2004 when there has been a regime shift in Ringkøbing Fjord. Five variables are given with uncertainty bands (thin lines) based on the error in the annual mean value at the 95 % confidence level.*

samples for each individual value. According to table 3, one should not expect modelled TP-values for the growing season to yield higher $r^2$-values than 0.92 when regressed against empirical data, a value which lies slightly above the yearly $r_e^2$ value (fig. 3). TN, Sec and Chl could be predicted with similar or higher precision since their $r_r^2$-values were between 0.91 and 0.96, while SPM and dissolved inorganic nutrients could be predicted with lower precision ($r_r^2$ at 0.60-0.81 according to table 3).

## 4.2. Regressions and variability

The aim of this section is to utilize data from Ringkøbing Fjord and Chesapeake Bay to illustrate inherent and general problems related to regressions and empirical models based on uncertain water variables. Fig. 4 gives frequency distributions for selected water variables (chlorophyll, TP, SPM, Secchi depth and TN) using the entire data set from Ringkøbing Fjord for the period 1980 to 2004. One can note that all frequency distributions are positively skewed; the ratio between the mean value and the median indicates the degree of normality, and all these variables have MV/M50-ratios larger than 1. This also means that a higher degree of normality may be obtained after log-transformation (see Håkanson and Peters, 1995).

Fig. 5 illustrates changes in important water variables (median annual values) in Ringkøbing Fjord, and in this section, we will use regressions to try to statistically explain variations in the two target bioindicators (chlorophyll-a concentration and Secchi depth). The basic question here is: how much of the variation in median annual values can be statistically explained by the given x-variables TN/TP-ratio, TP, OrtP, TN, nitrate+nitrite ($NO_x$), ammonium+ammonia ($NH_x$), and SPM? Fig. 6 shows the regressions between logarithmic values for median chlorophyll concentrations (Chl) during the summer period versus the corresponding values for TP and TN. There is a much closer agreement ($r^2$ = 0.94) between TP and Chl than between TN and Chl ($r^2$ = 0.75) in this lagoon. Table 5 gives results for many regressions between chlorophyll (as y-variable) and potential x-variables and table 6 gives similar results for Secchi depth as target y-variable. From these results, one may conclude:

- The results depended very much on the season of the year; the best results are generally obtained for data from the summer period.
- Better correlations were obtained for log-median values than for log-mean values (data not displayed) because most frequency distributions for most variables are not normal but log-normal (see fig. 4).

Table 5. *Relationships between different nutrients, nutrient ratios, or suspended particulate matter and chlorophyll-a in Ringkøbing Fjord. Six different averaging methods have been used on median log values: annual values of all data, all data adjusted to give equal weight to each of the four seasons, spring values, summer values, autumn values and winter values. All correlations are based on log-values. Data are from the period 1986–2004, although series that only include TP, $NH_x$, Chl, Secchi and SPM also cover 1985 (except for winter values). From Håkanson et al. (2007b).*

| | All data | Season adjusted | Spring | Summer | Autumn | Winter |
|---|---|---|---|---|---|---|
| $r^2$ (TP vs Chl) | 0.96<br>p | 0.94<br>p | 0.86<br>p | 0.94<br>p | 0.93<br>p | 0.56<br>p |
| $r^2$ (OrtP vs Chl) | 0.23<br>n | 0.32<br>n | 0.03<br>n | 0.35<br>p | 0.57<br>m | 0.30<br>m |
| $r^2$ (TN vs Chl) | 0.33<br>p | 0.24<br>p | 0.06<br>n | 0.75<br>p | 0.58<br>p | 0.13<br>n |
| $r^2$ ($NO_X$ vs Chl) | 0.50<br>m | 0.06<br>n | 0.00<br>n | 0.19<br>m | 0.27<br>m | 0.00<br>n |
| $r^2$ ($NH_X$ vs Chl) | 0.42<br>m | 0.59<br>m | 0.55<br>m | 0.24<br>m | 0.52<br>m | 0.35<br>m |
| $r^2$ (DIN vs Chl) | 0.56<br>m | 0.09<br>n | 0.00<br>n | 0.22<br>n | 0.33<br>m | 0.02<br>n |
| $r^2$ (TN:TP vs Chl) | 0.81<br>m | 0.77<br>m | 0.73<br>m | 0.68<br>m | 0.63<br>m | 0.46<br>m |
| $r^2$ (DIN:OrtP vs Chl) | 0.45<br>m | 0.00<br>n | 0.01<br>n | 0.51<br>m | 0.15<br>n | 0.33<br>n |
| $r^2$ (SPM vs Chl) | 0.89<br>p | 0.81<br>n | 0.70<br>p | 0.72<br>p | 0.78<br>p | 0.24<br>p |

Significance at the $p<0.05$ level; p = positive, m = negative, n = not significant, $r^2$ = coefficient of determination, TP = total phosphorus, TN = total nitrogen, OrtP = orthophosphate, Chl = chlorophyll, $NO_X$ = nitrate + nitrite, $NH_X$ = ammonium + ammonia, DIN = dissolved inorganic nitrogen = $NO_X$ + $NH_X$, SPM = suspended particulate matter.

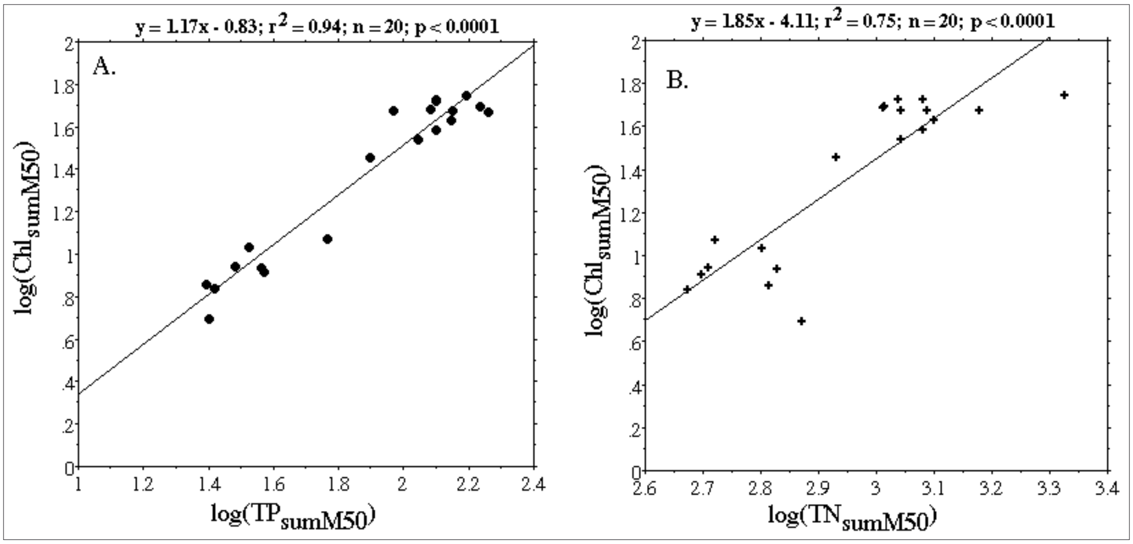


Fig. 6. *Two regressions based on median values from the summer period (months 6, 7 and 8; using data from 1980 to 2004) between chlorophyll and total phosphorus concentrations (A) and between chlorophyll and total nitrogen concentrations (logarithmic data for Ringkøbing Fjord; if we use data from 1986 to 2004, when the data are more reliable but fewer, the $r^2$-value is 0.76), B. The figure also gives the regression lines, $r^2$-values, number of data (n) and the statistical certainties (p) of the regressions.*

Table 6. *Relationships between nutrients, Chl-a and the Secchi depth in Ringkøbing Fjord. Median log-values. See table 5 for explanations and abbreviations. Data are from the period 1986–2004, although series that only include TP, $NH_x$, Chl, Sec and SPM also cover 1985 (except for winter values).*

| | All data | Season adjusted | Spring | Summer | Autumn | Winter |
|---|---|---|---|---|---|---|
| $r^2$ (TP vs Sec) | 0.94 m | 0.95 m | 0.92 m | 0.93 m | 0.83 m | 0.88 m |
| $r^2$ (OrtP vs Sec) | 0.22 n | 0.31 p | 0.01 n | 0.39 m | 0.49 p | 0.01 n |
| $r^2$ (TN vs Sec) | 0.42 m | 0.33 m | 0.19 n | 0.76 m | 0.56 m | 0.42 m |
| $r^2$ ($NH_X$ vs Sec) | 0.41 p | 0.52 p | 0.45 p | 0.19 n | 0.29 p | 0.22 n |
| $r^2$ ($NO_X$ vs Sec) | 0.48 p | 0.01 n | 0.03 n | 0.52 p | 0.14 n | 0.28 m |
| $r^2$ (DIN vs Sec) | 0.52 p | 0.03 n | 0.03 n | 0.56 p | 0.18 n | 0.13 n |
| $r^2$ (TN:TP vs Sec) | 0.71 p | 0.69 p | 0.61 p | 0.77 p | 0.48 p | 0.61 p |
| $r^2$ (DIN:OrtP vs Sec) | 0.42 p | 0.03 n | 0.02 n | 0.62 p | 0.06 n | 0.06 n |
| $r^2$ (Chl vs Sec) | 0.95 m | 0.83 m | 0.78 m | 0.94 m | 0.78 m | 0.35 m |
| $r^2$ (SPM vs Sec) | 0.86 m | 0.87 m | 0.79 m | 0.73 m | 0.84 m | 0.79 m |

Significance at the $p<0.05$ level; p = positive, m = negative, n = not significant

- There were major differences among the x-variables in how they correlate to chlorophyll and Secchi depth. TP was by far the best predictor for both chlorophyll and Secchi depth in Ringkøbing Fjord. Table 7a shows results from stepwise multiple regressions on yearly median values between five potential x-variables (TP, TN, $NO_x$, $NH_x$ and salinity) and three different y-variables (Secchi depth, Chl and SPM). Out of the various x-variables, only TP was included in all three cases. When the stepwise multiple regression was repeated for median summer values (see table 7b), the pattern was similar for Secchi and Chl, but not for SPM, where TN and $NO_x$ were included as x-variables. TN also showed strong individual correlation with summer SPM ($r^2$ = 0.86 compared to 0.77 between TP and SPM). SPM was always rejected as an additional explanatory x-variable for Secchi and Chl.
- Nitrogen or ratios based on nitrogen or different forms of nitrogen generally co-vary with the two target bioindicators less well than TP in this lagoon.

Table 7. *Stepwise multiple regression models based on data from Ringkøbing Fjord.*

| y-variable | n (years) | Step | F | $r^2$ | x-variable | Model |
|---|---|---|---|---|---|---|
| *A. Based on yearly median values.* | | | | | | |
| log (Chl) | 20 | 1 | 421 | 0.96 | $x_1$=log (TP) | $y=-4.239+1.7107 \cdot x_1$ |
| log (Sec) | 20 | 1 | 262 | 0.94 | $x_1$=log (TP) | $y=3.361-0.8170 \cdot x_1$ |
| log (SPM) | 20 | 1 | 155 | 0.89 | $x_1$=log (TP) | $y=-1.239+1.007 \cdot x_1$ |
| *B. Based on summer median values.* | | | | | | |
| log (Chl) | 20 | 1 | 280 | 0.94 | $x_1$=log (TP) | $y=-0.833+1.175 \cdot x_1$ |
| log (Sec) | 20 | 1 | 254 | 0.93 | $x_1$=log (TP) | $y=1.494-0.8129 \cdot x_1$ |
| log (SPM) | 19 | 1 | 104 | 0.85 | $x_1$=log (TN) | $y=-6.914+3,395 \cdot x_1$ |
| | | 2 | 67 | 0.89 | $x_2$=log ($NO_x$) | $y=-5.589+3.028 \cdot x_1-0.2028 \cdot x_2$ |

Table 8. *Bivariate correlations ($r^2$ values) from surface water variables in Chesapeake Bay, months 6–8.*

| | DIN | DN | DON | DOP | DP | OrtP | PN | PP | TN | TP | Chl | Sec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Chl | 0.01<br>n | 0.06<br>p | 0.27<br>p | 0.26<br>p | 0.03<br>p | 0.00<br>n | 0.69<br>p | 0.52<br>p | 0.32<br>p | 0.22<br>p | | |
| Sec | 0.07<br>m | 0.31<br>m | 0.34<br>m | 0.23<br>m | 0.12<br>m | 0.14<br>m | 0.36<br>m | 0.51<br>m | 0.50<br>m | 0.38<br>m | 0.38<br>m | |
| SPM | 0.02<br>p | 0.07<br>p | 0.07<br>p | 0.04<br>p | 0.16<br>p | 0.20<br>p | 0.12<br>p | 0.55<br>p | 0.17<br>p | 0.45<br>p | 0.17<br>p | 0.56<br>m |

Significance at the $p<0.05$ level; p = positive, m = negative, n = not significant.

Table 8 shows which variables can be used to predict Chl, Sec and SPM in Chesapeake Bay with high certainty. Here, phosphorus as a predictor was not always superior to nitrogen as in Ringkøbing Fjord. Those nutrients or nutrient fractions which were the best predictors were generally total or particulate nitrogen or phosphorus, followed by dissolved organic nutrients and dissolved inorganic nutrients. An exception was the OrtP vs. SPM relationship which was relatively strong. According to table 9, it seems that various fractions of both nitrogen and phosphorus are needed to explain areal variations in Chl and Sec, while TP and OrtP were enough to explain variations in SPM. A comparison with the $r_r^2$ value for Chl in table 4 (0.77) gives at hand that Chl cannot be predicted with much higher certainty than in table 9 ($r^2$=0.79). However, $r_r^2$ values for both Sec and SPM are much higher than their respective $r^2$ values in table 9 (0.95 compared to 0.76 and 0.80 compared to 0.65).

To put these regressions between nutrients and bioindicators of eutrophication (tables 5–9) into a wider comparative context, it is necessary to compare a much greater number of aquatic ecosystems. In such a comparative study, Guildford and Hecky (2000) found a much stronger correlation ($r^2$=0.60 compared to 0.08) between TP and Chl than between TN and Chl at several ocean sites. Rather similarly, Håkanson et al. (2007b) found that TP and Sal in combination correlated slightly more strongly with Chl ($r^2$=0.71) than TN and Sal did ($r^2$=0.68) in a wide range of aquatic systems. Conversely, Smith (2006) found TN to be a better predictor of Chl than TP ($r^2$=0.84 compared to 0.60) and that TN and TP are strongly mutually correlated ($r^2$=0.55).

It may be argued that TN and TP are unsuitable for predicting Chl, since the phytoplankton which produce Chl contain large concentrations of nitrogen and phosphorus, increasing the risk of autocorrelation. However, TN and TP may alternatively be seen as proxies of Chl which have common causes (riverine nutrient inputs). Tables 5–9 clearly demonstrate that concentrations of dissolved inorganic nutrients are of very limited use in predictive coastal science and management. Even though batch experiments in laboratories often show that dissolved inorganic nutrients is what phytoplankton consume, concentrations of dissolved inorganic nutrients poorly reflect their availability since they are very rapidly regenerated (Dodds, 2003). Instead, Chl can be predicted with much higher certainty from TN, TP, PN or PP, as this study has shown.

These results do not specifically address the very interesting and much debated issue on limiting or regulating nutrient. However, Schindler (1977, 1978) showed that bioassays and inorganic nutrient concentrations may be irrelevant for the management of aquatic systems since phosphorus governs the long-run concentra-

Table 9. *Stepwise multiple regression models based on surface water data from Chesapeake Bay, months 6–8.*

| y-variable | n (stations) | Step | F | $r^2$ | x-variable | Model |
|---|---|---|---|---|---|---|
| log (Chl) | 191 | 1 | 422 | 0.69 | $x_1$=log (PN) | $y=1.548+1.022\cdot x_1$ |
| | | 2 | 346 | 0.79 | $x_2$=log(PP) | $y=1.530+0.7614\cdot x_1+0.4445\cdot x_2$ |
| log (Sec) | 186 | 1 | 244 | 0.57 | $x_1$=log (SPM) | $y=1.596-0.6341\cdot x_1$ |
| | | 2 | 275 | 0.75 | $x_2$=log (TN) | $y=3.410-0.4692\cdot x_1+0.6927\cdot x_2$ |
| | | 3 | 194 | 0.76 | $x_3$=log (DOP) | $y=3.113-0.4765\cdot x_1-0.5169\cdot x_2-0.1929\cdot x_3$ |
| log (SPM) | 117 | 1 | 155 | 0.63 | $x_1$=log (PP) | $y=-0.5251+1.090\cdot x_1$ |
| | | 2 | 106 | 0.65 | $x_2$=log (OrtP) | $y=-0.5050+1.007\cdot x_1+0.1006\cdot x_2$ |

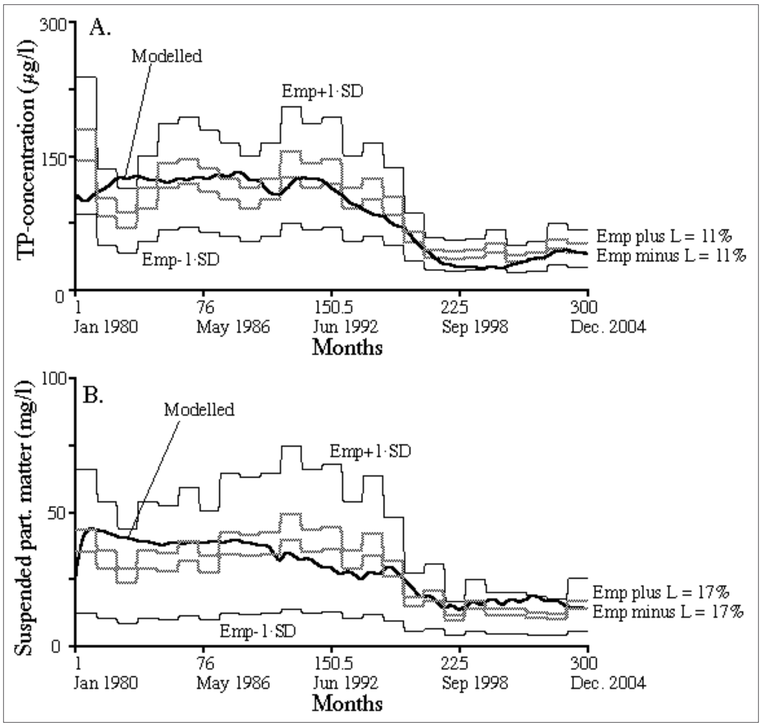


Fig. 7. *A comparison between empirical data and dynamically modelled values for TP-concentrations (fig. A) and SPM-concentrations (fig. B) in Ringkøbing Fjord and 70 % (1· annual SD) confidence intervals related to the uncertainty in the individual empirical data (CV = 0.48 for TP and 0.71 for SPM) and the error bands for the mean annual values (L = 11 % for TP and L = 17 % for SPM).*

tions Chl in lakes, while any short-term nitrogen deficits are counteracted over the year through nitrogen fixation. This mechanism has also been demonstrated for the brackish Baltic Proper (Savchuk and Wulff, 1999) and the global ocean (Tyrrell, 1999).

## 5. Concluding comments

We have presented CV-values for many standard water variables for coastal areas. High CV-values imply that uncertainty bands are relatively wide if few samples are taken. Fig. 7 presents two types of uncertainty bands defined for TP and SPM, first the wide 95 % confidence intervals for individual data, then the error bands in the mean values (11 % for TP and 17 % for SPM; as calculated from eq. 1) related to the characteristic annual CV-values (0.48 for TP and 0.71 for SPM, see table 2) and the available number of data for each year (n = 72). These uncertainty bands are very useful in modelling and fig. 7 illustrates how they can be applied to dynamically modelled values (from Håkanson et al., 2007a) in comparison with empirical median annual data.

Furthermore, we have suggested and tested methods for motivating spatial and temporal scales in coastal modelling and management, and discussed methods for determining the predictive power and overall usefulness of various potential driving variables in regression models for important coastal bioindicators.

Finally, we have found that TN, TP, PN and PP show much lower variability and are more strongly correlated to Chl and Sec than DIN and dissolved inorganic phosphorus (DIP) or fractions thereof. The former are therefore much more relevant than the latter for predictive coastal management. Unfortunately, many coastal monitoring programs do not include TN, TP, PN or PP but only DIN and DIP, and many management related studies (e. g., Tyrrell, 1999; Arhonditsis et al., 2000; Newton et al., 2003) are solely based on DIN and DIP pools when it comes to nutrient fluxes. Thus, the results from this paper may improve the future design of coastal monitoring programs.

### Acknowledgements

This work has been carried out within the framework of the Thresholds-project, and integrated EU project (no., 003933-2), and we would like to acknowledge the financial support from EU and the constructive cooperation within the project. Jens W Hansen and his colleagues in the marine group of the County of Ringkjøbing are greatly acknowledged for collecting and delivering data and for valuable comments to the manuscript. We are equally grateful to the Chesapeake Bay Program for sharing monitoring data online.

## Literature references


Arhonditsis, G., Tsirtsis, G., Angelidis, M. O., and M. Karydis, 2000. Quantification of the effects of nonpoint nutrient sources to coastal marine eutrophication: applications to a semi-enclosed gulf in the Mediterranean Sea. Ecological Modelling, 129: 209–227.

Bloesch J., 1995. Mechanisms, measurement and importance of sediment resuspension in lakes. Marine and Freshwater Research, 46: 295–304.

Boesch, D. F., Brinsfield, R. B. and Magnien, R. E., 2000. Chesapeake Bay Eutrophication: Scientific Understanding, Ecosystem Restoration, and Challenges for Agriculture. Journal of Environmental Quality, 30: 303–320.

Bortone, S.A. (ed.), 2005. Estuarine indicators. CRC Press, Boca Raton, 531 p.

Chesapeake Bay Program, 2007. Water quality database, http://www.chesapeakebay.net/data/index.cfm

Comin, F. A., Menéndez, M., Herrera, J. A., 2004. Spatial and temporal scales for monitoring coastal aquatic ecosystems. Aquatic Conservation: Marine and Freshwater Ecosystems, 14 (Suppl. 1): S5–S17.

Cooley, W.W. and Lohnes, P.R., 1971. Multivariate data analysis. John Wiley & Sons, Inc., New York, 364 p.

Cooper, S. R. and Brush, G. S., 1993. A 2,500-Year History of Anoxia and Eutrophication in Chesapeake Bay. Estuaries, 16: 617–626.

Dodds, W. K., 2003. Misuse of inorganic N and soluble reactive P concentrations to indicate nutrient status of surface waters. Journal of the North American Benthological Society, 22:171–181.

Draper, N. R. and Smith, H., 1966. Applied Regression Analysis. Wiley, New York, 352 p.

Guildford, S.J. and Hecky, R.E., 2000. Total nitrogen, total phosphorus, and nutrient limitation in lakes and oceans: is there a common relationship? Limnology and Oceanography, 45: 1213–1223.

Håkanson, L., 1984. Sediment sampling in different aquatic envi ronments: statistical aspects. Water Resources Research, 20: 41–46.

Håkanson, L., 1999. Water pollution – methods and criteria to rank, model and remediate chemical threats to aquatic ecosystems. Backhuys Publishers, Leiden, 299 p.

Håkanson, L., 2006. Suspended particulate matter in lakes, rivers and marine systems. –The Blackburn Press, New Jersey, 331 pp.

Håkanson, L. and Eckhéll, J., 2005. Suspended particulate matter (SPM) in the Baltic – new empirical data and models. Ecological Modelling, 189: 130–150.

Håkanson, L. and Peters, R.H., 1995. Predictive limnology. Methods for predictive modelling. SPB Academic Publishing, Amsterdam, 464 p.

Håkanson, L., Bryhn, A. C., and Eklund, J. M., 2007a. Modelling phosphorus and suspended particulate matter in Ringkøbing Fjord to understand regime shifts. Journal of Marine Systems (in press).

Håkanson, L., Bryhn, A. C., and Hytteborn, J. K., 2007b. On the issue of limiting nutrient and predictions of cyanobacteria in aquatic systems. Science of the Total Environment, 379: 89–108.

Irvine, K. 2004, Classifying ecological status under the European Water Framework Directive: the need for monitoring to account for natural variability. Aquatic Conservation: Marine and Freshwater Ecosystems, 14: 107–112.

de Jonge, V.N., 2000. Importance of temporal and spatial scales in applying biological and physical process knowledge in coastal management, an example for the Ems estuary. Continental Shelf Research, 20: 1655–1686.

Moldan, B. and Billharz, S., (eds), 1997. Sustainibility indicators. Wiley, see http://www.icsu-scope.org/downloadpubs/scope58/.

Mosteller, F. and Tukey, J. W., 1977. Data Analysis and Regression: a second cours in statistics. Addison-Wesley Publ. Reading, Mass. 588 p.

Newman, M.C., 1993. Regression analysis of log-transformed data: statistical bias and its correction. Environmental Toxicology and Chemistry, 12: 1129–1133.

Newton, A., Icely, J. D., Falcao, M., et al., 2003. Evaluation of eutrophication in the Ria Formosa coastal lagoon, Portugal. Continental shelf research, 23:1945–1961.

Nielsen, M.H., Rasmussen, B. and Gertz, F., 2004. A simple model for water level and stratification in Ringkøbing Fjord, a shallow, artificial estuary. Estuarine, Coastal and Shelf Science, 65: 235–248.

Nordvarg, L., 2001. Predictive models and eutrophication effects of fish farms . Dr. Thesis, Uppsala University, Sweden.

Peters, R. H., 1991. A critique for ecology. Cambridge University Press, Cambridge, UK, 366 pp.

Petersen, J.K., Hansen, J.W., Laursen M.B., Clausen, P., Carstensen, J. and Conley, D., 2008. Regime shift in a marine coastal ecosystem. Ecological Applications, 18: 497–510.

Pfaffenberger, R.C. and Patterson, J.H., 1987. Statistical methods. Irwin, Illinois, 1246 p.

Prairie, Y., 1996. Evaluating the predictive power of regression models. Canadian Journal of Fisheries and Aquatic Sciences, 53: 490–492.

Savchuk, O. and Wulff, F., 1999. Modelling regional and large-scale response of Baltic Sea ecosystems to nutrient reductions. Hydrobiologia, 393: 35–43.

Schindler, D.W., 1977. Evolution of phosphorus limitation in lakes. Science, 195: 260–262.

Schindler, D.W., 1978. Factors regulating phytoplankton production and standing crop in the world's freshwaters. Limnology and Oceanography, 23: 478–486.

Tyrrell, T., 1999. The relative influences of nitrogen and phosphorus on oceanic primary production. Nature, 688 400: 525–531.

Wallin, M., Håkanson, L. and Persson, J., 1992. Load models for nutrients in coastal areas, especially from fish farms (in Swedish with English summary). Nordiska ministerrådet, 1992:502, Copenhagen, 207 p.

Weston , K., Jickells, T. D., Fernand, L., and E. R. Parker, 2004. Nitrogen cycling in the southern North Sea: consequences for total nitrogen transport. Estuarine, Coastal and Shelf Science, 59: 559–573.

Weyhenmeyer, G.A., 1996. The significance of lake resuspension in lakes. Uppsala dissertations from the faculty of science and technology 225. Sweden. ISBN 91-554-3820-2.