BOREAL ENVIRONMENT RESEARCH 2: 303–315 ISSN 1239-6095
Helsinki 11 November 1997 

# Statistical modelling of phosphate variations in the Baltic proper

Charlotta Pers, Åsa Danielsson and Lars Rahm

*Department of Water and Environmental Studies, Linköping University, S-581 83 Linköping, Sweden*

Pers, C., Danielsson, Å. & Rahm, L. 1997. Statistical modelling of phosphate variations in the Baltic proper. *Boreal Env. Res.* 2: 303–315. ISSN 1239-6095

A statistical model of the variation of phosphate concentration in the upper layers of the Baltic proper was formed, with the aim of studying the magnitude of this variation and to be able to adjust corrupted time series. The large variation observed is only partially explained by seasonal variations and geographical structures. Despite inclusion of long-term trends and annual variations, the residuals showed a substantial correlation within and between the time series.

## Introduction

The Baltic Sea is affected by changing natural and anthropogenic conditions making it vulnerable (Rosenberg *et al*. 1990). It has narrow and shallow connections between its major basins and the North Sea, which do not allow sufficient water and mass exchange. Also the freshwater supply is small in comparison to the volume of the basin. Together these conditions create a system with a long response time. In addition, a permanent halocline, at a level below the bottom of the entrance to the Baltic proper, hampers the exchange processes between the salinity stratified deep water and the well-mixed surface layer. This makes the Baltic proper an efficient pollution trap (Wulff *et al*. 1990) heavily loaded with nutrients from natural and anthropogenic sources.

The awareness of the sensitivity of the Baltic Sea has lead to the formation of national and international monitoring programs. These programs have generated many time series which all have substantial variations in their observational variables. A considerable proportion of these variations are assumed to be explained by seasonality and spatial gradients (Kahma and Voipio 1989, Sandén and Danielsson 1995). Bergström and Carlsson (1994), on the other hand, focused on the large interannual variations in freshwater supply. Obviously long-term changes represent only a minor part of the observed variations.

The sampling frequency of the various monitoring programs is too low to properly describe the seasonal changes and/or annual variations. Storms and ice cover during winter prohibit sampling leading to gaps in the time series. Small changes in the start of the spring bloom due to ice cover, storms, insolation, stratification etc., affect the nutrient concentrations detected by the monitoring programs (Nõmmann 1990) leading to variable results for single monitoring cruises. Further, time series with equidistant sampling are usually necessary inputs to nutrient mod-

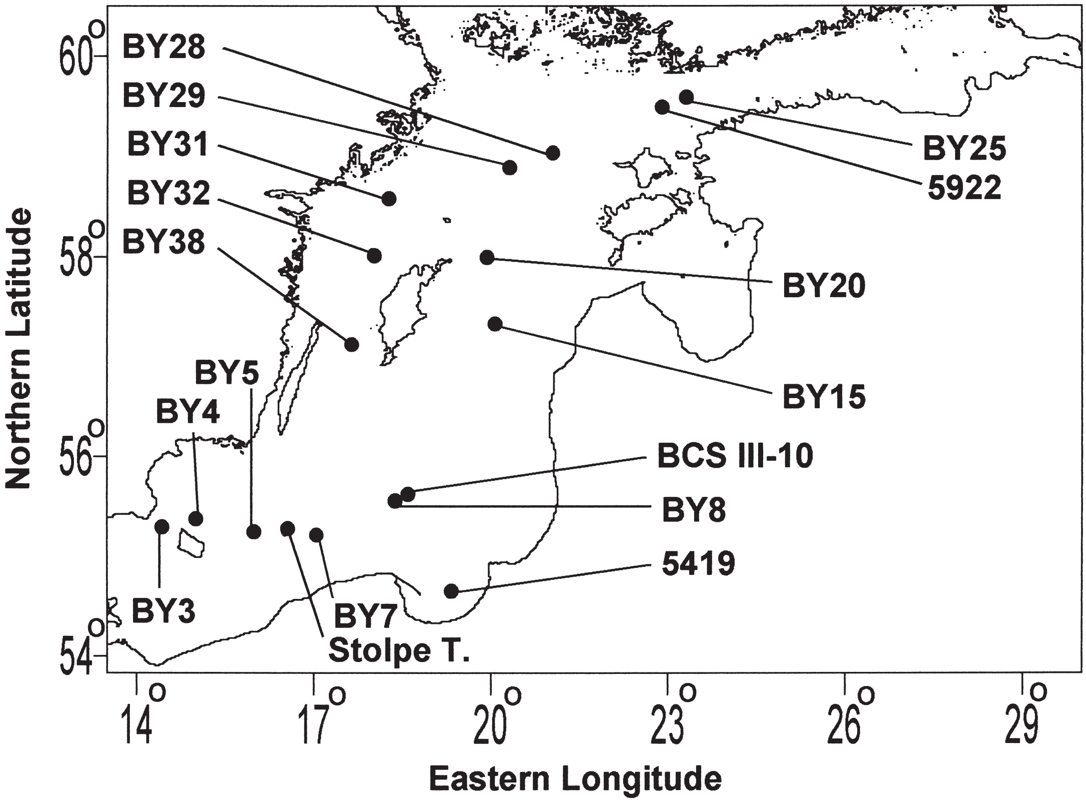


**Fig. 1**. The sampling stations in the Baltic proper.

els. Corrupted time series are often filled using linear interpolation, without regard to existing knowledge about spatio-temporal variations in nutrient concentrations.

Nutrient concentrations have large seasonal variations. This is not surprising as they are tightly coupled to the primary production and the subsequent regeneration of organic matter. A rapid decrease in the inorganic concentrations during the spring bloom is followed by a slow recovery during the summer and autumn. The formation of a strong but shallow thermocline in the upper layers during summer inhibits the exchange of water and dissolved substances with deeper layers. A second, minor, plankton bloom may occur during autumn. A more pronounced increase in nutrient concentrations occurs during late autumn and winter. This is caused by increased mixing and erosion of the halocline, due to storms and seasonal turnovers. The intensity of the exchange processes is probably similar for the entire Baltic proper, since the forcing functions act on a scale similar to or larger than this region.

Phosphate is chosen as a model substance in the present work. This is motivated by its central role in the nutrient dynamics, although it is not a limiting nutrient of primary production in the Baltic proper. While most limnetic systems and the slightly brackish Bothnian Bay are phosphorus limited, the Baltic proper is nitrogen limited (Granéli *et al*. 1990). The interannual variation of phosphate concentration can be substantial, e.g. due to climatological reasons. The mean annual supply of phosphorus to the Baltic proper, including import from adjacent basins, is roughly 10% of its total amount in the water mass (Wulff and Stigebrandt 1989). This means that the Baltic proper is rather sensitive to changes in the total annual supply of phosphorous from the rivers and the seasonal distribution of phosphorous.

The objective of this work is to describe the variations in phosphate concentration within the Baltic proper surface layer, using statistical models. These models are used to "fill the gaps" in present and historic time series. The interpolation procedure includes an estimation of the uncertainty of the created time series.

## Data material

The data used in this study was acquired from several off-shore monitoring stations in the Baltic proper (Fig. 1). To get sufficient amount of data and minimize the variance due to different ships and laboratories, data from Baltic Environmental Database (BED) at the Department of Systems Ecology, Stockholm University, was used. A permanent quality control is carried out within BED, where questionable or erroneous data are expelled. Observations made before 1970 have low reliability due to analytical problems, and consequently only those after 1970 were included.

The data had been irregularly sampled in time. In this study there is a demand of equal spacing between observations. Therefore, the data was divided into intervals of one month. Lack of observations prohibited higher resolution. Every time series is characterised by its position and observation depth. When there were more than one observation during one month at the same position and depth, the median value was used. During the exploration of data, some outliers were omitted (0.1% of the original data). Although only the most frequently visited stations were included in this study, the chosen time series lack 40–80% of their observations (Table 1).

A pronounced variation within each year as well as between years is clearly visible for all time series. For example at station BY15, during winter (December to April) the phosphate concentration was high and varied considerably between years, while during the summer (June and July) it was rather low and stable (Fig. 2). The phosphate concentration is obviously linked to the biological activity.

Sandén and Danielsson (1995) found no significant differences in phosphate concentrations in the latitudinal direction in the Baltic proper. Due to lack of data, longitudinal variations could not be analysed. Hence, the Baltic proper is in the forthcoming modelling assumed to have a homogeneous seasonal variation. Three stations (BY5, BY15 and BY31) located in the major subbasins of the Baltic proper were chosen for a minor study of the homogeneity of seasonal variation. They are assumed to be representative for their respective subbasins (though only the surface layers are considered). These three stations are also the most frequently visited stations in the Baltic. In several studies BY15 is assumed to represent the entire Baltic proper (Rahm 1985, Wulff

**Table 1**. Estimated mean phosphate concentration level, number of observations (*n*), and estimated trend slope for all time series. Non-significant trend slopes are denoted by *n.s.*, *** denotes significance level less than 1%, ** denotes significance level less than 5% and * denotes significance level less than 10%.

| Station | Depth (m) | Mean phosphate concentration (mmol m$^{-3}$) | *n* | Slope (mmol m$^{-3}$ yr$^{-1}$) |
|---|---|---|---|---|
| BY3 | 0 | 0.33 | 73 | *n.s.* |
| | 10 | 0.34 | 74 | *n.s.* |
| | 20 | 0.37 | 75 | *n.s.* |
| BY4 | 0 | 0.30 | 61 | *n.s.* |
| | 10 | 0.29 | 60 | *n.s.* |
| | 20 | 0.32 | 63 | *n.s.* |
| BY5 | 0 | 0.32 | 148 | *n.s.* |
| | 5 | 0.29 | 126 | *n.s.* |
| | 10 | 0.33 | 161 | *n.s.* |
| | 15 | 0.31 | 122 | *n.s.* |
| | 20 | 0.34 | 162 | *n.s.* |
| BY7 | 0 | 0.28 | 72 | *n.s.* |
| | 10 | 0.29 | 72 | *n.s.* |
| | 20 | 0.30 | 72 | *n.s.* |
| BY8 | 0 | 0.27 | 72 | *n.s.* |
| | 10 | 0.25 | 55 | *n.s.* |
| | 20 | 0.26 | 75 | 0.0060* |
| BY15 | 0 | 0.28 | 145 | *n.s.* |
| | 1 | 0.31 | 50 | *n.s.* |
| | 5 | 0.26 | 100 | *n.s.* |
| | 10 | 0.28 | 164 | *n.s.* |
| | 15 | 0.26 | 100 | *n.s.* |
| | 20 | 0.30 | 152 | *n.s.* |
| BY20 | 0 | 0.27 | 95 | *n.s.* |
| | 10 | 0.26 | 94 | *n.s.* |
| | 20 | 0.30 | 95 | *n.s.* |
| BY25 | 0 | 0.33 | 66 | *n.s.* |
| | 10 | 0.34 | 68 | *n.s.* |
| | 20 | 0.38 | 66 | *n.s.* |
| BY28 | 0 | 0.32 | 129 | 0.0060** |
| | 5 | 0.27 | 82 | 0.0025* |
| | 10 | 0.31 | 127 | 0.0061** |
| | 15 | 0.28 | 82 | *n.s.* |
| | 20 | 0.34 | 128 | 0.0064*** |
| BY29 | 0 | 0.24 | 81 | – 0.0033** |
| | 10 | 0.28 | 84 | – 0.0039*** |
| | 20 | 0.30 | 85 | *n.s.* |
| BY31 | 0 | 0.28 | 105 | *n.s.* |
| | 5 | 0.27 | 87 | *n.s.* |
| | 10 | 0.30 | 133 | *n.s.* |
| | 15 | 0.30 | 83 | *n.s.* |
| | 20 | 0.32 | 111 | *n.s.* |
| BY32 | 0 | 0.25 | 67 | *n.s.* |
| | 5 | 0.24 | 49 | *n.s.* |
| | 10 | 0.28 | 79 | *n.s.* |
| | 15 | 0.24 | 49 | *n.s.* |
| | 20 | 0.28 | 72 | *n.s.* |
| BY38 | 0 | 0.36 | 106 | *n.s.* |
| | 5 | 0.27 | 78 | *n.s.* |
| | 10 | 0.35 | 115 | *n.s.* |
| | 15 | 0.29 | 73 | *n.s.* |
| | 20 | 0.39 | 119 | 0.0050** |
| BCS III-10 | 10 | 0.32 | 78 | *n.s.* |
| | 20 | 0.35 | 82 | 0.0060* |
| STOLPE T. | 0 | 0.30 | 66 | *n.s.* |
| | 10 | 0.32 | 60 | *n.s.* |
| | 20 | 0.33 | 64 | *n.s.* |
| 5922 | 0 | 0.33 | 78 | *n.s.* |
| | 10 | 0.35 | 74 | *n.s.* |
| | 20 | 0.36 | 76 | *n.s.* |
| 5419 | 0 | 0.21 | 56 | *n.s.* |
| | 10 | 0.26 | 55 | *n.s.* |
| | 20 | 0.28 | 54 | *n.s.* |

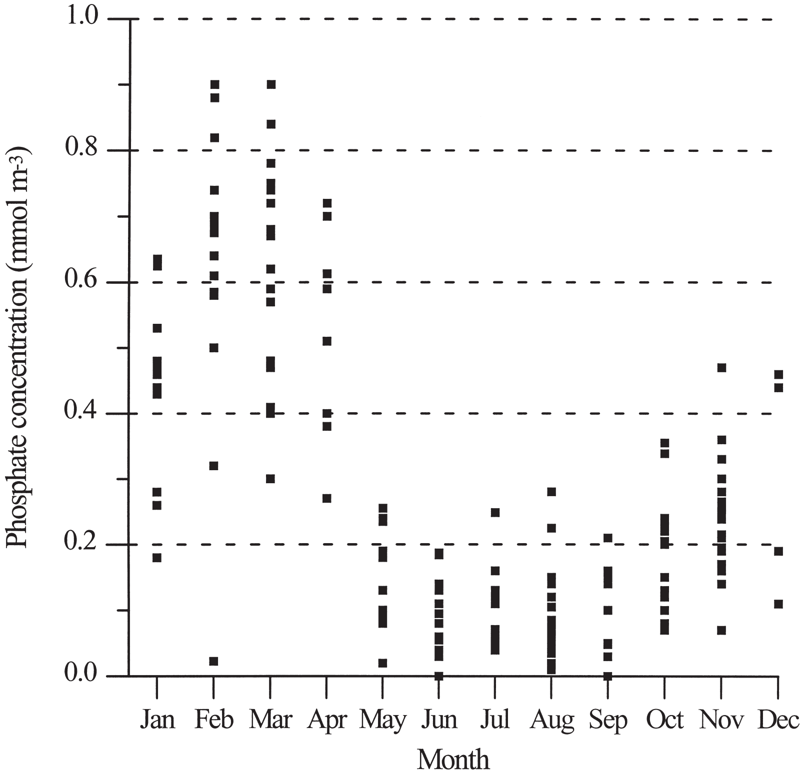


**Fig. 2**. Phosphate concentration against month. Data from station BY15, 10 m, 1970–93.

and Rahm 1989). It is therefore interesting to compare the results from these stations to the results from all stations in the Baltic proper.

## Models and methods

To describe the variations in phosphate concentration, a statistical modelling approach was used. The behaviour of the concentration is explained by different components of variation. Two models were taken under consideration. The first model (M1) takes into consideration the mean level, a seasonal component and a random variation, while the second model (M2) also includes a linear trend and annual variations. Both models are additive time series models. They both assume the random variation to be independent of phosphate concentration.

### Model 1

M1 assumes the nutrient concentration to be composed of three parts: a spatially dependent mean value ($\mu$) around which the concentration varies, a seasonal variation ($S$) representing the seasonal deviation from the mean level, and a random variation ($\varepsilon$). The model of phosphate concentration ($X$) is described by

$$X(p,d,y,m) = \mu(p,d) + S(m) + \varepsilon(p,d,y,m) \quad (1)$$

where: $p$ = spatial location (station), $d$ = sampling depth, $y$ = year and $m$ = month.

The mean nutrient concentration and the seasonal variation are regarded as unknown deterministic functions, while the residual is assumed to be a stochastic variable.

The mean value of nutrient concentration ($\mu$) is specific for each time series, and it is dependent on the sampling station and depth. It is estimated by a weighted average of observed nutrient concentrations:

$$\hat{\mu}(p,d) = \frac{1}{12}\sum_{m}\left[\frac{1}{n_y}\sum_{\forall y} x(p,d,y,m)\right] \quad (2)$$

where: $x$ = the observed nutrient concentration and $n_y$ = the number of observations at a given station and depth during a specific month throughout the years.

The seasonal variation is regarded as spatially homogeneous, i.e. independent of sampling position and depth, and consequently it is equal for all time series. It is estimated as an average of all stations, depths and years at month $m$, minus the average of all estimated mean levels:

$$\hat{S}(m) = \frac{1}{n_{pdy}} \sum_{\forall p,d,y} x(p,d,y,m) - \frac{1}{n_{pd}} \sum_{\forall p,d} \hat{\mu}(p,d) \quad (3)$$

The residuals are assumed independent and identically distributed, with unknown distribution and expected value zero. They are estimated by the remains, i.e. by the observed nutrient concentration minus the estimated mean level and seasonal variation:

$$\hat{\varepsilon}(p,d,y,m) = x(p,d,y,m) - \hat{\mu}(p,d) - \hat{S}(m) \quad (4)$$

## Model 2

M2 includes also a linear trend and an annual variation in comparison to M1,

$$\begin{aligned} X(p,d,y,m) &= \mu(p,d) + S(m) \\ &+ \underbrace{T(p,d,y,m) + A(d,y) + \eta(p,d,y,m)}_{\varepsilon(p,d,y,m)} \end{aligned} \quad (5)$$

where: $T$ = linear trend, $A$ = annual variation and $\eta$ = residual.

The trend is characteristic for individual time series, while the annual variation is assumed independent of sampling station and month. The average nutrient concentration, seasonal variation, linear trend and annual variation are regarded as unknown deterministic functions. The residuals are assumed to be stochastic, independent and identically distributed (i.i.d.), with unknown distribution and expected value zero. In M2, $\varepsilon$ is not expected to fulfil the assumptions that are stated in M1 since it incorporates annual variations and long-term trend as well as random variations. The seasonal variation and mean level were estimated as previously.

The trend ($T$) was estimated using the method proposed by Hirsch and Slack (1984) and Hirsch *et al.* (1982). It is robust against non-Gaussian distribution and extreme values. It handles seasonality, but assumes identically distributed variables. When the trend test gave a $p$-value less than 0.1, the estimated trend slope of the method was used in the model, otherwise the trend was assigned the value zero.

$$\hat{T}(p,d,y,m) = \begin{cases} \mathrm{slope}(p,d)\left[y + \frac{m}{12} - \left(1982 + \frac{0.5}{12}\right)\right] & p-\text{value} < 0.1 \\ 0 & p-\text{value} \geq 0.1 \end{cases} \quad (6)$$

The linear trend was pivoted around year 1982 and the correction 0.5 was used to put the date to the middle of each month. As seen in equation (7), the annual variation was estimated by the average of $\varepsilon$ calculated first over months and then over stations.

$$\hat{A}(d,y) = \frac{1}{n_p} \sum_{\forall p} \left[ \frac{1}{n_m} \sum_{\forall m} \varepsilon(p,d,m,y) \right] \quad (7)$$

This indirect weighting procedure was applied so that stations with many observations during a particular year would not have greater impact than stations with few observations during the same year.

Finally the residuals ($\eta$) are estimated:

$$\hat{\eta}(p,d,y,m) = \hat{\varepsilon}(p,d,y,m) - \hat{T}(p,d,y,m) - \hat{A}(d,y) \quad (8)$$

## Methods for prediction

It is often necessary to estimate values where there is a lack of data in time series, e.g. for use in future budget calculations. In this study, the procedure for predicting missing values in time series was composed of two phases. First, the observed time series was decomposed into its components, according to above. Second, a new time series ($\hat{x}$) was created by summing all deterministic components:

$$\hat{x}(p,d,y,m) = \hat{\mu}(p,d) + \hat{S}(m) + \hat{\varepsilon}(p,d,y,m) \quad (\mathrm{M1})$$

(9)

$$\begin{aligned} \hat{x}(p,d,y,m) &= \hat{\mu}(p,d) + \hat{S}(m) \\ &+ \hat{T}(p,d,y,m) + \hat{A}(d,y) + \hat{\eta}(p,d,y,m) \end{aligned} \quad (\mathrm{M2})$$

The random residual components are predicted by their expected value, i.e. zero. The gaps in the original series were then replaced by values from the new series resulting in a 'filled' series. This

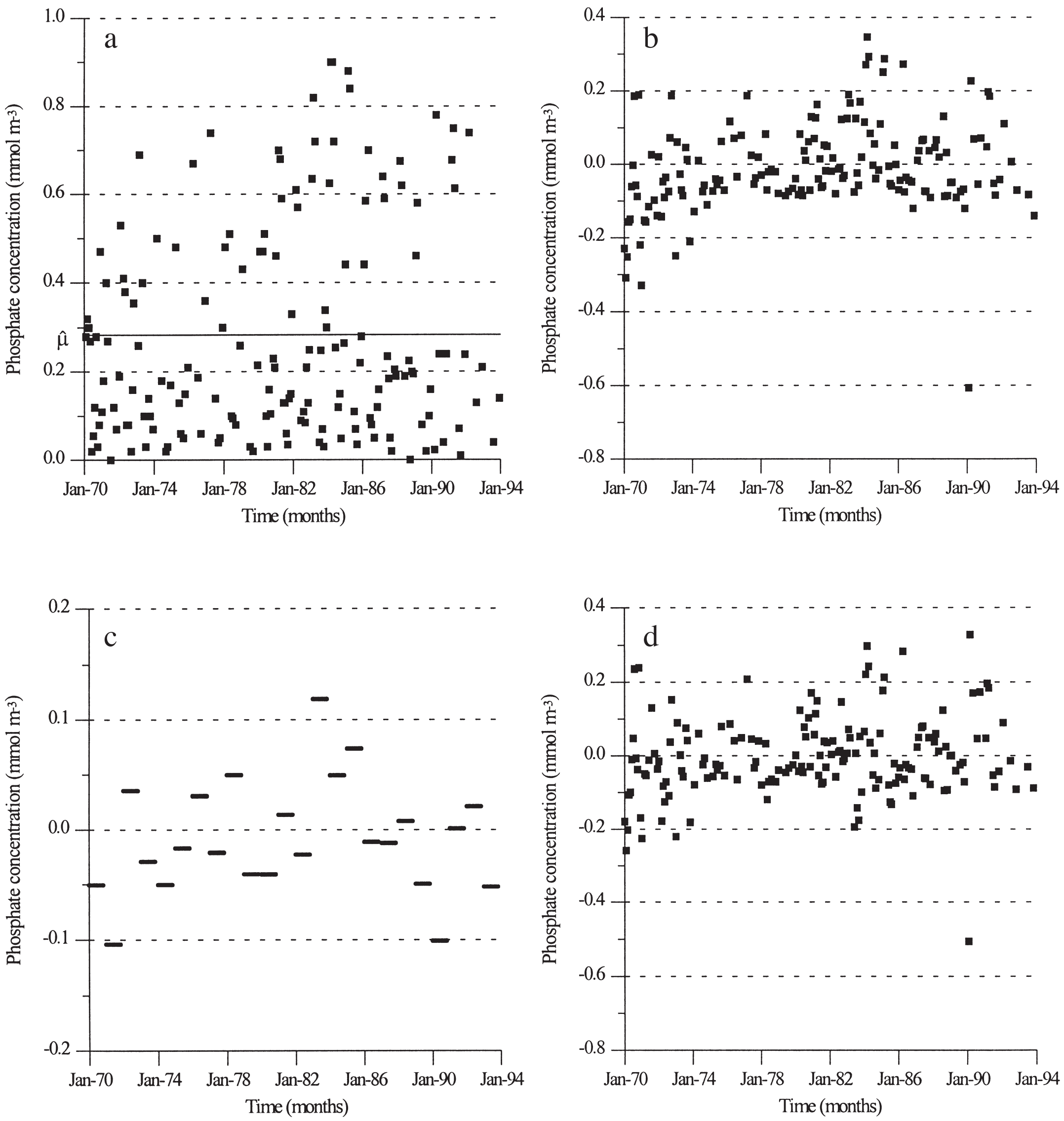


**Fig. 3**. Estimated components of time series of phosphate concentration from BY15 at 10 m depth from 1970–93. — a: Mean level ($\hat{\mu}$) and observed phosphate concentrations. — b: Residuals of phosphate concentration of M1. — c: Annual variation of phosphate concentration. — d: Residuals of phosphate concentration of M2.

method can be used to predict phosphate concentration in the immediate future. Then the unknown annual variation was estimated to be zero, i.e. its mean value.

## Methods for model test

To evaluate the models, their components were estimated for time series of a slightly shorter period, 1970–91. Thereafter the components obtained from the models were used to predict the following two years. The models were investigated by comparing the mean sum of squares of the difference between predicted concentration and observed concentration ($MSS_{pred}$):

$$MSS_{pred} = \frac{1}{n} \sum_{y=92,93} \sum_{\forall p,d,m} \left[ x_{pred}(p,d,y,m) - x(p,d,y,m) \right]^2 \quad (10)$$

The mean sum of squares of the difference between the observed values and those of the estimated

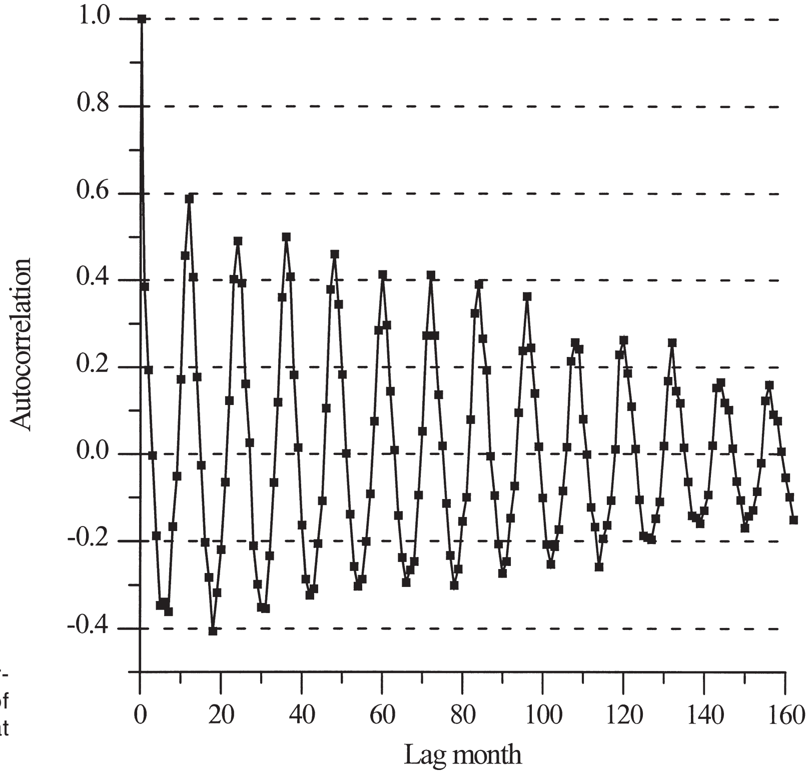


**Fig. 4**. Estimated autocorrelation for time series of phosphate concentration at station BY15 at 10 m depth.

time series from 1970–93 ($MSS_{interp}$) becomes:

$$MSS_{interp} = \frac{1}{n} \sum_{y=92,93} \sum_{\forall p,d,m} \left[ \hat{x}(p,d,y,m) - x(p,d,y,m) \right]^2 \quad (11)$$

## Results

The model components were estimated for all stations in Fig. 1 (Fig. 3) and for three selected stations (BY5, BY15 and BY31) separately. The results from the two models are presented together to allow comparison.

The autocorrelation was estimated for the observed time series, ignoring seasonality or possible long term trend. The estimated autocorrelation shows a fairly long time dependence (Fig. 4). An intersection at about 14 years is noted. In addition, the autocorrelation clearly reflects the seasonal variation.

### Mean level and seasonal variation

The mean level of phosphate concentrations of all time series vary by a factor two over the area, but generally increase with depth (Table 1).

The seasonal variation of phosphate are of the same magnitude as the mean level (Fig. 5). During the summer, the concentrations approach zero. Highest concentrations are obtained in February, right before the start of primary production with about twice the assumed mean levels.

For the entire basin and the three selected stations the seasonal variation is rather similar. Only the northernmost station (BY31) shows a marked difference during late winter and early spring. Station BY15 has its maximum concentration later than the average of the entire basin and a later, but faster decrease during spring. The seasonal variation of BY5 is almost identical with the one of the entire basin.

### Trends and annual variations

The examination of M2 gave only a few statistically significant trends ($p$-value < 0.1). The stations with significant trends (BY8, BY28, BY29, BY38 and BCS III-10) are located in the middle

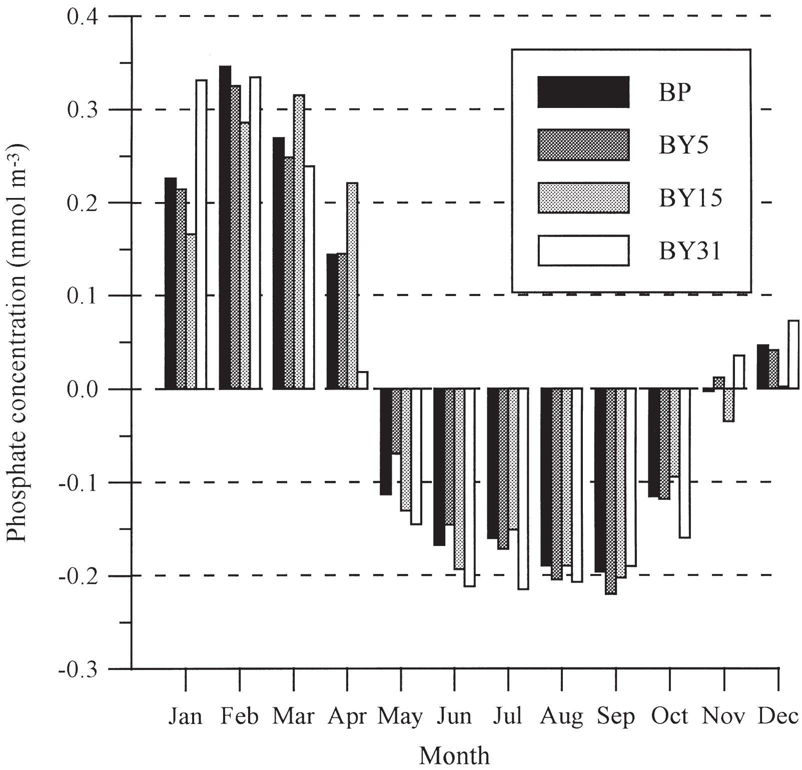


**Fig. 5**. Estimated seasonal variation of phosphate concentration for the entire Baltic proper (BP), and for three selected stations.

of the Baltic proper (Table 1 and Fig. 1). Both positive and negative trends were detected, with a maximum slope of 2% $yr^{-1}$.

The annual variation in phosphate concentration at 10 m depth is large, although smaller than the mean concentration (Fig. 3a and 3c). It is clear that the seasonal variation is larger than either the annual variation or the trend. Obviously, the two latter components explain very little of the variation of $\varepsilon$, as it is less than the variation remaining after their removal (Fig. 3d). Still, the introduction of trend and annual variation improve the estimates, since the residuals are smaller, although marginally, and seem to have a less systematic behaviour (Fig. 3b and d).

## Residuals

A successful modelling would render residuals that are insignificant in magnitude. In our case a few residuals is comparable with the mean level. The majority of residuals is smaller though. This means that the random variation of concentration is large and the model will only describe the variations partially.

The autocorrelation was estimated for the residual time series of M1 and M2. For the first thirty months it varied between – 0.2 and 0.3 (compare with the autocorrelation based on the original times series at BY15; Fig. 4). Seasonal variations may be found in the residuals, see Fig. 6.

The dependence between residual time series at different stations and depths was investigated by estimating the correlation coefficient. For residuals of M1, less stations, only 25% (for M2, 23%) had a high dependency (correlation coefficient above 0.5), compared to 80% of the original time series. The correlation between depths was investigated separately. These correlations were generally high. Almost all of them had correlation coefficients above 0.5.

## Model test

M1 gives better predictability (smallest $MSS_{pred}$) of the two models, while M2 gives a better description of the data (smallest $MSS_{interp}$). The differences in mean sum of squares were small for

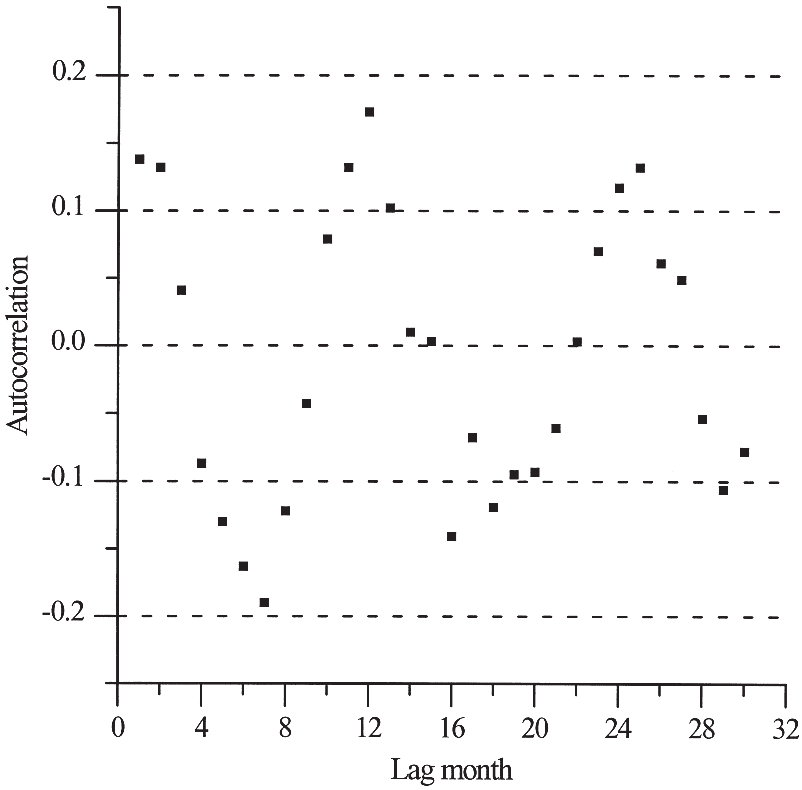


**Fig. 6**. Estimated autocorrelation for residuals of phosphate concentration of M2 at station BY15 at 10 m depth.

all four cases (Table 2). For M2 the sum of squares differ by ~ 28% in favour of the interpolation ($MSS_{interp}$ , i.e. the residuals), while for M1 the difference is only 6%.

### Estimating missing values

According to the model test, M2 is the better model for interpolation. It is therefore used to replace missing values in the time series. The estimated time series for phosphate concentration of BY15 series includes some negative values, which of course is physically impossible (Fig. 7a). Gaps in the original time series were replaced with estimated values if they were larger than zero and otherwise by zero (Fig. 7b).

Since the residuals came from an unknown distribution, percentiles were used as a description of the uncertainty in the estimation of missing values. There are only minor differences in percentiles between the models, but the 5th and 95th percentiles indicate that the residuals of M2 are less variable than those of M1 (Fig. 8a and b). The effect of the large uncertainty on the estimated time series is illustrated in Fig. 7c. The percentiles suggest that the distribution of the residuals $\varepsilon$ and $\eta$ is similar for most months.

## Discussion

Two statistical models were applied to time series of phosphate concentration of the Baltic proper. One model was based on mean value and seasonal variation, while the second also included long-term trends and annual variations. The estimated autocorrelation for the observed time se-

**Table 2**. Mean sum of squares for predicted phosphate concentration ($MSS_{pred}$) and mean sum of squares for interpolated phosphate concentration ($MSS_{interp}$) for models 1 (M1) and 2 (M2).

| | M1 | M2 |
|---|---|---|
| $MSS_{pred}$ | 0.0148 | 0.0157 |
| $MSS_{interp}$ | 0.0139 | 0.0123 |

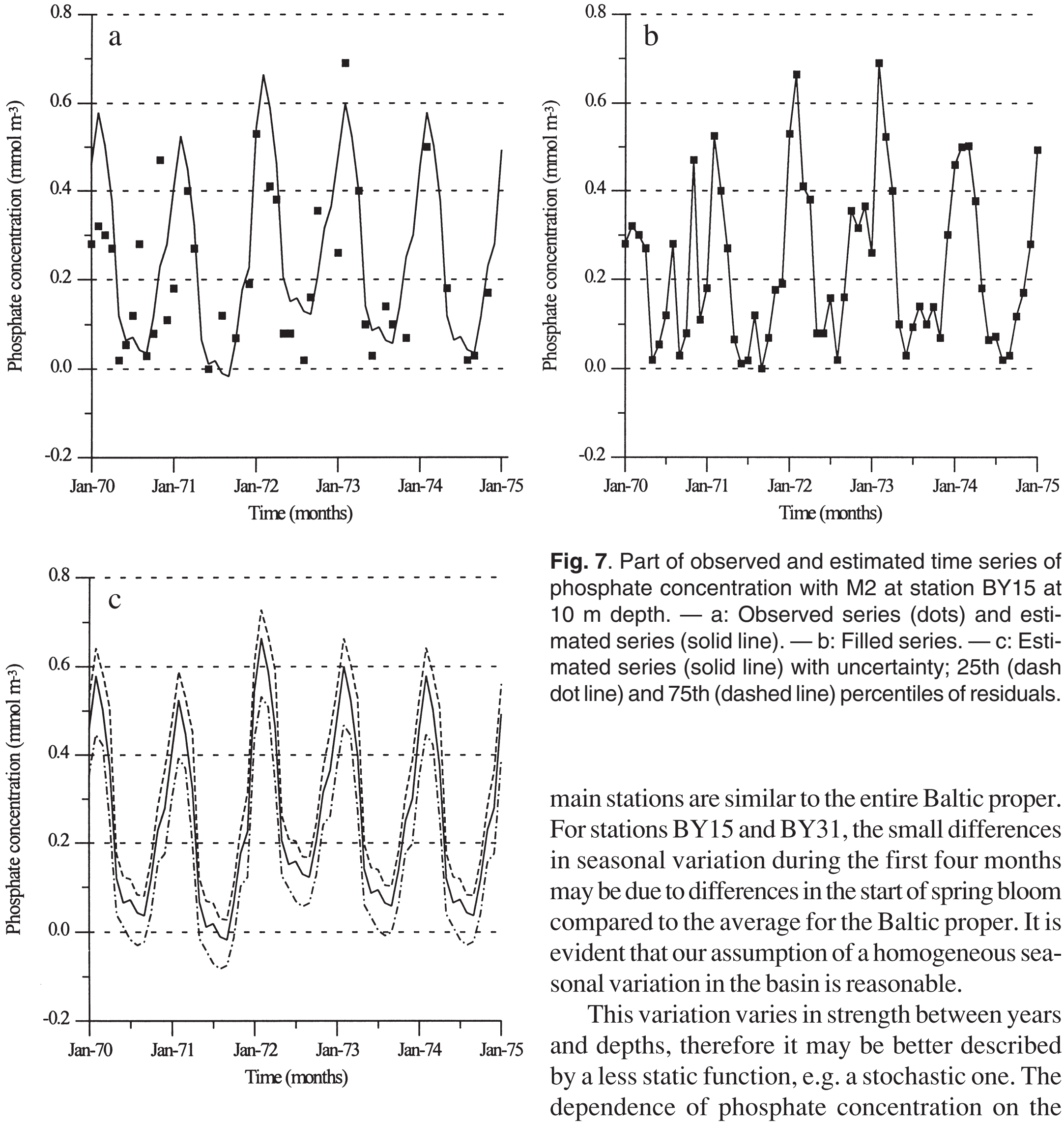


**Fig. 7**. Part of observed and estimated time series of phosphate concentration with M2 at station BY15 at 10 m depth. — a: Observed series (dots) and estimated series (solid line). — b: Filled series. — c: Estimated series (solid line) with uncertainty; 25th (dash dot line) and 75th (dashed line) percentiles of residuals.

ries showed a time dependence in accordance with the characteristic time scales for nutrients estimated by Wulff and Stigebrandt (1989).

Seasonal variations and spatial gradients explain only a part of the variation in the observations. Despite explicit modelling of these factors, large variations were still found in the residuals. There may be several reasons for this. Some of the variations are completely random while others may be due to measurement error. One possibility is the coarse time resolution. We analysed on a monthly basis, excluding important processes acting on shorter time scales, e.g. diurnal variations.

Estimates of the seasonal variation for the three main stations are similar to the entire Baltic proper. For stations BY15 and BY31, the small differences in seasonal variation during the first four months may be due to differences in the start of spring bloom compared to the average for the Baltic proper. It is evident that our assumption of a homogeneous seasonal variation in the basin is reasonable.

This variation varies in strength between years and depths, therefore it may be better described by a less static function, e.g. a stochastic one. The dependence of phosphate concentration on the amplitude of seasonal variations was assumed to be minor. Therefore, an additive model was applied instead of a multiplicative model. According to the percentiles of the estimated residuals this may not have been an entirely correct assumption (Fig. 8). There is a larger scatter in data and higher phosphate concentrations during the winter months than during the summer months.

From a statistical point of view there is the question of how well the components have been estimated. As mentioned earlier, the seasonal Kendall test for trends (Hirsch and Slack 1984) assumes the variables to be identically distributed. The percentiles of residuals in Fig. 8 show that this is not likely

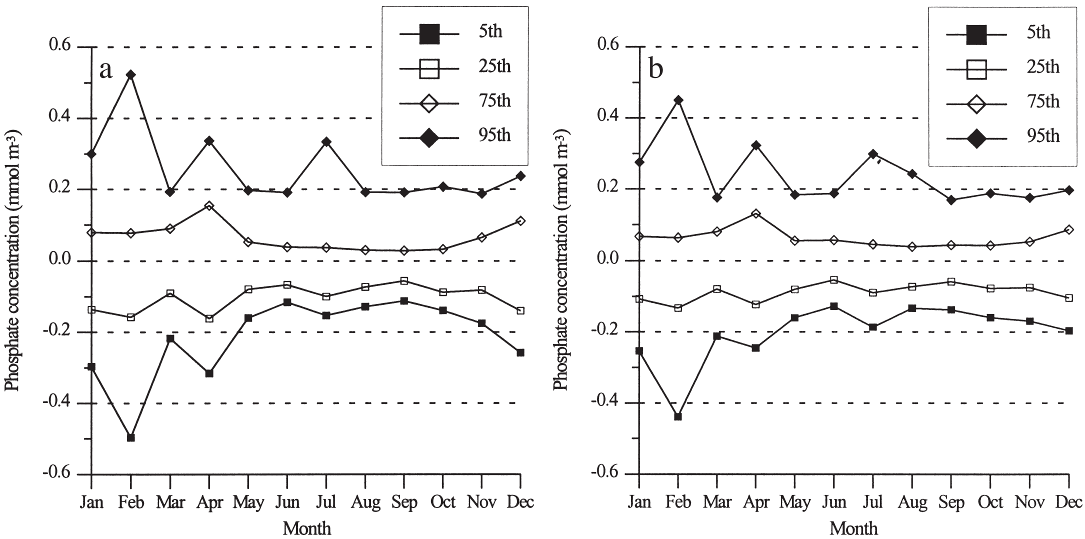


**Fig. 8**. Percentiles of residuals of phosphate concentration for the entire Baltic proper. Samples from 1970–93. — a: Residuals of M1. — b: Residuals of M2.

the case. In addition, there may be bias towards summer concentrations in the component estimations. Although the time resolution was set on a monthly basis, because of available observations, there were months completely lacking observations. This was especially true for the winter season, when sampling was sparse. This imbalance was reduced by estimating the mean levels with weighted averages, but it could still have hampered the estimates of the annual variation.

Long term trends are probably also present in the time series. Sandén *et al.* (1991) and Sandén and Rahm (1993) found significant monotonous nutrient trends using data from 1968–90. In our study it is evident that a monotonous, linear trend may not be the best description (Fig. 9). Unfortunately, possible non-linear trends could not be taken into consideration, due to the high amount of missing data for some of the time series. Instead a linear trend with the above mentioned annual variation was used in M2.

The annual variation of phosphate concentration is large, although smaller than the mean phosphate concentration. Fig. 9 clearly indicates that the seasonal variation is greater than the annual variation. Climatic factors are probably responsible for these variations. As an example the annual phosphorus load from rivers is relatively large compared to the phosphorous pool in the water mass (about one-tenth, Wulff *et al.* 1990), making riverine load important.

The investigation of residuals does not fully corroborate our prior assumptions of independence, but the residuals show much less time dependence than the original phosphate concentrations. However, these estimates may be unreliable, as the correlation estimates assume that the residual variance is constant, an assumption that may be unreasonable in this case (Fig. 8a and b). The residual time series seems to be only weakly correlated in time, but the correlations between the different time series are large. This contradicts our assumption of independence and is true irrespective of model.

Seasonal variations may be found in the residuals, see Fig. 6. This is evident when comparing the seasonal variation estimates of the three specific stations in Fig. 5. Since these estimates are not identical, seasonal components must remain in the residuals.

For the model of stations BY5, BY15, and BY31 the residuals $\hat{\eta}$ are equal or less than the residuals for the model of entire basin. Since the residuals of the general model do not differ substantially from those of the reduced models, the conclusion must be that the model applied to the entire basin is acceptable in comparison to our specialised models.

The result of the model test is that the two models are complementary to each other. M2 is better for interpolation, but M1 performs best in prediction.

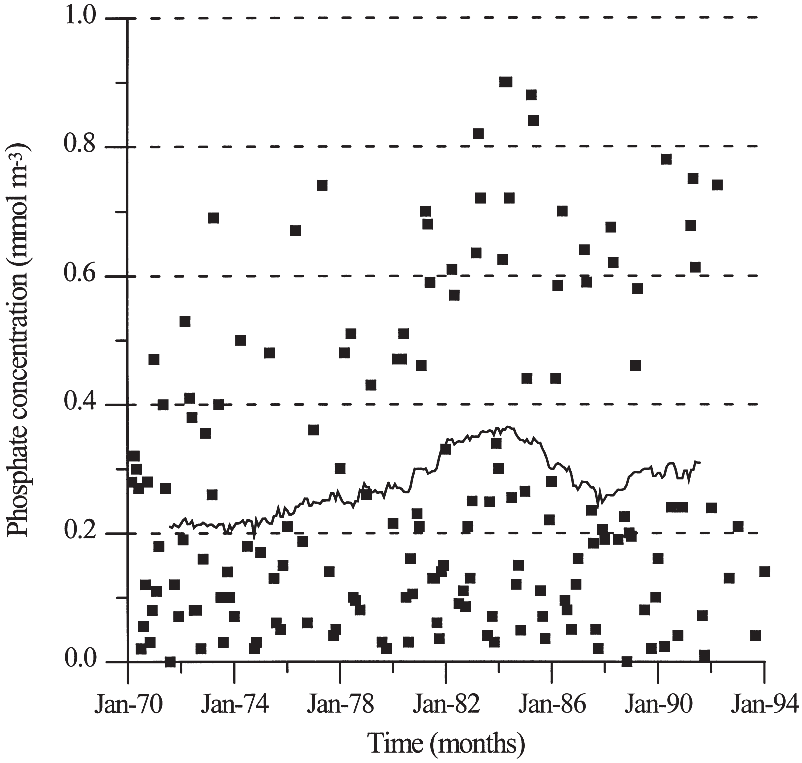


**Fig. 9**. Observed phosphate concentration (dots), with three year moving average (solid line) at station BY15 at 10 m depth.

## Conclusions

Four conclusions can be drawn from this study:

— the seasonal variation and the spatial dependence do not fully explain the large variation in phosphate concentration in the photic zone
— even with an inclusion of annual variation and linear trend the variation is not fully explained
— the seasonal variation is rather homogeneous in the surface layer
— the method is useful for time series completion

Acknowledgement: This study was financially supported by the Swedish Environmental Protection Agency (SEPA).

## References

Bergström S., & Carlsson B. 1994. River runoff to the Baltic Sea: 1950–1990. *Ambio*. 23: 280–287.

Granéli E., Wallström K., Larsson U., Granéli W. & Elmgren R. 1990. Nutrient limitation of primary production in the Baltic Sea area. *Ambio*. 19: 142–151.

Hirsch R. M. & Slack J. R. 1984. A nonparametric trend test for seasonal data with serial dependence. *Water Resour. Res*. 20: 727–732.

Hirsch R. M., Slack J. R. & Smith R. A. 1982. Techniques of trend analysis for monthly water quality data. *Water Resour. Res*. 18: 107–121.

Kahma K. & Voipio A. 1989. Seasonal variation of some nutrients in the Baltic Sea and the interpretation of monitoring results. ICES *C.M. 1989/C:31*.

Nõmmann S. 1990. *Physical control of phytoplankton growth in the Baltic Sea: A multitude of spatio-temporal scales*, Ph.D.-thesis, Stockholm University.

Rahm L. 1985. On the diffusive salt flux of the Baltic proper. *Tellus*, 37: 87–96.

Rosenberg R., Elmgren R., Fleischer S., Jonsson P., Persson G. & Dahlin H. 1990. Marine eutrophication case studies in Sweden. *Ambio*. 19: 102–108.

Sandén P. & Danielsson Å. 1995. Spatial properties of nutrient concentrations in the Baltic Sea. *Environmental Monitoring and Assessment*. 34: 289–307.

Sandén P. & Rahm L. 1993. Nutrient trends in the Baltic Sea. *Environmetrics*. 4: 75–103.

Sandén P., Rahm L. & Wulff F. 1991. Non-parametric trend test of Baltic Sea data. *Environmetrics*. 2: 263–278.

Wulff F. & Rahm L. 1989. Optimizing the Baltic sampling

programme: the effect of using different stations in calculation of total amount of nutrients. *Beitr. Meeresk.* 60: 61–66.
Wulff F. & Stigebrandt A. 1989. A time-dependent budget model for nutrients in the Baltic Sea. *Global Biogeochemical Cycles*. 3: 63–78.
Wulff F., Stigebrandt A. & Rahm L. 1990. Nutrient Dynamics of the Baltic Sea. *Ambio*. 19: 126–133.

Received 20 August 1996, accepted 31 August 1997