Variance reduction for trend analysis of hydrochemical data from brackish waters

Claudia Libiseller[1] and Anders Nordgaard[1]

Abstract

We propose one parametric and one non-parametric method for detection of monotone trends in nutrient concentrations in brackish waters. Both methods take into account that temporal variation in the quality of such waters can be strongly influenced by mixing of salt and fresh water, thus salinity is used as a classification variable in the trend analysis. With the non-parametric approach, Mann-Kendall statistics are calculated for each salinity level, and the parametric method involves the use of bootstrap estimates of the trend slope in a time series regression model. In both cases, tests for each salinity level are combined in an overall trend test.

1 Introduction

Cultural eutrophication of coastal and marine ecosystems has become a widespread problem over the past decades. The most severe effects can be found in semi-enclosed waters such as the Baltic Sea (Richardson and Jørgensen, 1996; Wulff *et al.*, 2001) and estuaries, in which there is a limited exchange of water. Considerable efforts have been made to reduce this eutrophication, and both the Helsinki Commission (HELCOM) and the Oslo-Paris Commission (OSPAR) have recommended a 50% reduction of nutrient inputs to the sea. However, the impact of implemented measures can be difficult to verify for two reasons. First, there can be considerable time lags between measures and effects (Grimvall *et al.*, 2000). Second, the mixing of salt and fresh water can introduce natural fluctuations in water quality that can result in prolonged concealment of important anthropogenic trends.

In the present study, we examined whether the human impact on phosphorus concentrations in brackish waters can be elucidated by taking into account variation in the salinity of the collected water samples. In particular, we investigated how information about salinity classes can be incorporated into

[1] Department of Mathematics, Division of Statistics, Linköping University, SE-58183 Linköping, Sweden

parametric and nonparametric test for trends in phosphorus. Splitting collected data into two or more subsets is widely practised in the environmental sciences. For example, Fryer and Nicholson (2002) used the body length of fish to divide data on mercury concentrations in muscle tissue from these animals into two groups, and they subsequently applied smoothers to examine temporal changes in the two time series. Simmonds *et al.* (1997) split tropospheric data into samples representing "background" and "polluted" air masses and thereafter assessed temporal trends by computing moving averages. In the current study, we combined trend tests for different subsets of data into an overall test using one non-parametric and one parametric method. The nonparametric approach we used is based on multivariate Mann-Kendall tests for monotone trends (Hirsch and Slack, 1984; Lettenmaier, 1988; Libiseller and Grimvall, 2002), and the parametric approach involves bootstrap confidence intervals for the trend slope in linear regression models (Efron and Tibshirani, 1993).

2 Natural variation in the Baltic Sea

At the surface of the Baltic Sea, the water is well mixed by the action of the wind, and the concentrations of nutrients exhibit pronounced seasonal variation. There is a strong halocline below the surface layer. Ocean water from the North Sea (Kattegat) can enter the Baltic proper through the Straits, but the salinity of the upper layer of the sea only occasionally reaches a level high enough to enable mixing with the much saltier bottom water. On average, the major inflows of ocean water that are required to allow such mixing occur only about nine times a century, and, between those events, conditions are often hypoxic or anoxic in the stagnant bottom water (Stigebrandt, 2001).

3 Data

The data set we analysed included monthly measurements of nutrient concentrations and other hydrochemical parameters collected at a large number of sampling sites in the Baltic Sea during the period 1989–1998. At each site, information was gathered at different depths, with a resolution of 5–10 m. We chose to use data from a sampling site in the Western Gotland Basin (57°07'N; 17°40'E) to illustrate the statistical procedures.

4 Trend testing of total phosphorus in the Baltic Sea

As indicated in Figure 1, the major inflows of sea water can render conventional univariate trend tests inappropriate. The temporal variation in phosphorus has a long-wave component, and there is a striking coincidence between the medium- and long-term variation in phosphorus and salinity. The close relationship between levels of phosphorus and salinity is further illustrated in Figure 2, which shows that there was a considerable reduction in variance when the phosphorus concentrations were plotted against salinity instead of depth.

Together, the results presented in Figures 1 and 2 suggest that salinity should be regarded as a covariate in the tests for nutrient trends. The work described below demonstrates how this can be achieved by using nonparametric and parametric approaches.

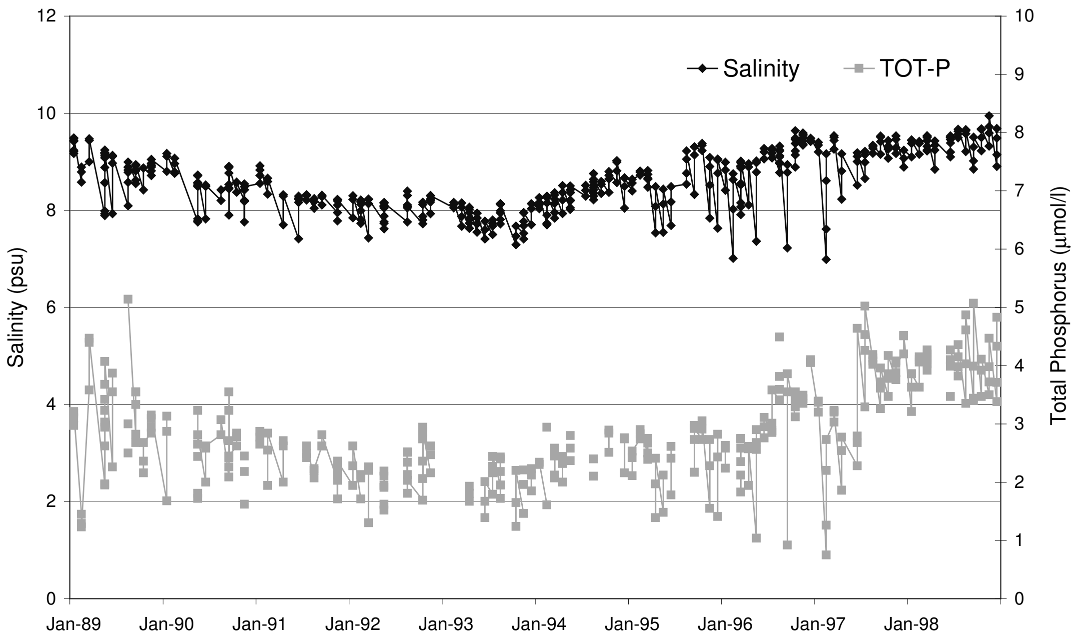


Figure 1. Time series of total phosphorus concentrations and salinity in the Western Gotland Basin. The graph shows data recorded at a depth of more than 80 m.

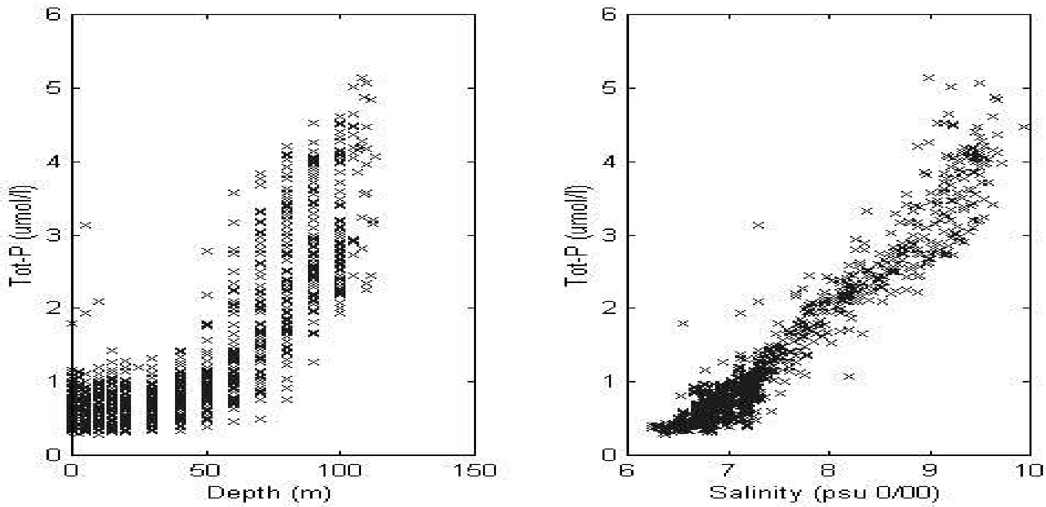


Figure 2. Total phosphorus concentrations in the Western Gotland Basin plotted against depth (left) and salinity (right).

4.1 A non-parametric trend test

The classical univariate Mann-Kendall test (Mann, 1945; Kendall, 1975) is a non-parametric test for randomness against trend, and the test statistic for a time series y_k, $k = 1,...,n$, is computed as follows:

$$T = \sum_{j<i} sign(y_i - y_j) \qquad 1 \le j < i \le n \qquad (4.1.1)$$

Under the null hypothesis of no trend, and if there are no ties or missing values, the test statistic is asymptotically normally distributed, with mean 0 and variance $n(n-1)(2n+5)/18$. Over the past decades, several other versions of Mann-Kendall tests have been developed to accommodate ties, seasonality, and missing values (Hirsch and Slack, 1984). In addition, a simple estimator of the covariance between two Mann-Kendall statistics (Dietz and Killeen, 1981) paved the way for combined trend tests using several Mann-Kendall statistics.

To perform an overall trend test for all seasons and salinity levels, we used what is known as the covariance sum test that was developed by Lettenmaier (1988) to detect trends at a network of stations. For this test, the test statistic can be written

$$Z = \mathbf{1}^T \mathbf{T} \Big/ \sqrt{\mathbf{1}^T \mathbf{\Gamma} \mathbf{1}} \qquad (4.1.2)$$

where **T** is the vector of all univariate Mann-Kendall statistics, **Γ** is the variance-covariance matrix of that vector, and **1** is a vector with all elements equal to 1. Under the null hypothesis of no trend, Z converges weakly to a standard normal distribution as $n \to \infty$.

4.2 A parametric approach

Let us assume that, within each salinity class l $(l=1,\ldots,K)$, the time series of nutrient concentrations has a linear trend and a deterministic seasonal pattern. Then we formulate a time series regression model

$$y_{t,i} = \beta_0 + \beta_1 t + \sum_{j=1}^{S-1} \beta_{sj} x_j + \sum_{l=1}^{K-1} (\beta_{cl}^{(1)} u_l + \beta_{cl}^{(2)} u_l \cdot t) + \varepsilon_{t,i} \, ; t = 1,\ldots,n \, ; i = 1,\ldots,m_t$$

(4.2.1)

where S is the number of seasons, $x_1,\ldots,x_{S-1}$ denote dummy variables for the seasons, $u_1,\ldots,u_{K-1}$ are salinity class dummies (= 1 for salinity class $l+1$, 0 otherwise), and ε_t are random errors that may be correlated. The indicated model may have several observations (m_t) for each time point, and we propose the following overall trend parameter:

$$\beta = w_{(1)} \beta_1 + \sum_{l=1}^{K-1} w_{(l+1)} \cdot (\beta_1 + \beta_{cl}^{(2)}) \tag{4.2.2}$$

where $w_{(1)},\ldots,w_{(K)}$ are weights proportional to the number of observations in the K salinity classes. An unbiased estimator of β is

$$\hat{\beta}_T = w_{(1)} \hat{\beta}_1 + \sum_{l=1}^{K-1} w_{(l+1)} \cdot (\hat{\beta}_1 + \hat{\beta}_{cl}^{(2)}) = \hat{\beta}_1 + \sum_{l=1}^{K-1} w_{(l+1)} \cdot \hat{\beta}_{cl}^{(2)} = \mathbf{w}' \cdot \hat{\boldsymbol{\beta}} \tag{4.2.3}$$

where $\hat{\beta}_1, \hat{\beta}_{cl}^{(2)}$ are the ordinary least squares (OLS) estimates (stacked in the vector $\hat{\boldsymbol{\beta}}$), and $\boldsymbol{w}$ is the row vector $(0,1,0,\ldots,0,w_2,\ldots,w_K)$. A studentised version of (4.2.3) becomes

$$t = \frac{\hat{\beta}_T - \beta}{s\sqrt{\mathbf{w}' \cdot (\boldsymbol{X}'\boldsymbol{X})^{-1} \cdot \mathbf{w}}} \tag{4.2.4}$$

where s^2 is the mean square error, and $(\boldsymbol{X}^{'}\boldsymbol{X})^{-1}$ is the information matrix of the regression model.

The statistic t can be used to construct a confidence interval for β, provided we have some information about the underlying probability distribution of the data. However, to avoid making any assumptions about this distribution, we can compute confidence intervals for β using a bootstrap approach.

The bootstrap in its original form (Efron, 1979) does not allow for any dependencies between the observations. Nevertheless, several investigators have proposed different ways of applying the bootstrap approach to such cases. We elected to use what is designated the ARMA bootstrap method (Kreiss and Franke 1992).

Applying OLS to the regression model (4.2.1) gives a set of residuals $e_1,...,e_N$, where N is the total number of observations. Because several observations are taken for each time point, the ARMA model can not be applied directly to these residuals. Instead, the median residual for each time point is calculated and is denoted e_t. Based on descriptive studies of such residuals, the following auto-regressive model of order 2 is proposed:

$$\tilde{e}_t = \varphi_{i,1}\tilde{e}_{t-1} + \varphi_{i,2}\tilde{e}_{t-2} + a_t \tag{4.2.5}$$

In this model, the φ-coefficients are estimated by OLS, and residuals are subsequently calculated as

$$\hat{a}_t = \tilde{e}_t - \hat{\varphi}_1\tilde{e}_{t-1} + \hat{\varphi}_2\tilde{e}_{t-2}, t \geq 2 \qquad (4.2.6$$

These residuals are standardised to zero mean and then resampled by ordinary bootstrap. From the resampled residuals $\hat{a}_1^*,...,\hat{a}_n^*$, resampled versions $\tilde{e}_{t,i}^*$, $i = 1,\ldots,m_t$, of $\tilde{e}_t$ are calculated using equation (4.2.5) and suitable initial values. Finally, resampled versions of y_t are calculated according to

$$y_{t,i}^* = \hat{\beta}_0 + \hat{\beta}_1 t + \sum_{j=1}^{S-1}\hat{\beta}_{sj}x_j + \sum_{l=1}^{K-1}(\hat{\beta}_{cl}^{(1)}u_l + \hat{\beta}_{cl}^{(2)}u_l \cdot t) + \tilde{e}_{t,i}^*\,; t = 1,\ldots,n; i = 1,\ldots,m_t \tag{4.2.7}$$

and resampled OLS estimates of the β-parameters can be obtained.

It is now possible to calculate a resampling distribution for the statistic t, that is, we compute resampled versions of t

$$t^* = \frac{\hat{\beta}_T^* - \hat{\beta}_T}{s_* \sqrt{\mathbf{w}' \cdot (\boldsymbol{X}'\boldsymbol{X})^{-1} \cdot \mathbf{w}}} \tag{4.2.8}$$

by repeating the bootstrap procedure a large number of times. The quantities $\hat{\beta}_T^*$ and s_*^2 are the counterparts of $\hat{\beta}_T$ and s^2 from the OLS estimation on $\{y_{t,i}^*\}$.

To determine whether trends are similar in different salinity classes, we can simply apply the corresponding bootstrap procedure to assess the estimates of $\{\beta_{cl}^{(2)}\}$.

5 Results

To illustrate the proposed tests, we divided the above-mentioned total phosphorus data for the Western Baltic Basin and into three salinity classes (6.2–7.4, 7.4–8.5, and 8.5–9.6 psu) and four seasons (Dec–Feb, Mar–May, Jun–Aug, and Sep–Nov). We chose to use a small number of classes to ensure that there would be a representative number of observations for each combination of salinity class and season. Two-sided tests and confidence intervals were applied to identify trends.

5.1 Non-parametric trend test

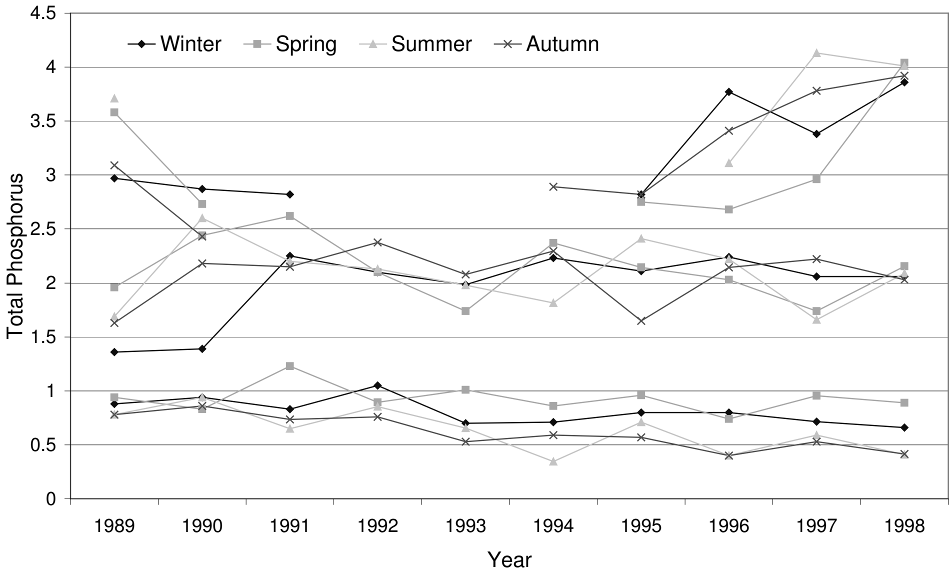
The Mann-Kendall test statistics were summed over the four seasons to obtain such statistics for each salinity level (Table 1). Also, the sum of all trend test statistics is given to represent a test statistic for the whole water column (designated all levels combined in Table 1). This overall test statistic did not suggest a significant trend. However, a significant downward trend in total phosphorus did appear for the group with low salinity levels. Figure 3 shows the analysed series of monthly median phosphorus and salinity levels divided into four seasons and three salinity classes.

Table 1. Trends in total phosphorus determined by Mann-Kendall tests performed using salinity-stratified data and summation over seasons

Salinity level	Z
Low	–2.5*
Medium	–0.17
High	1.52
All levels combined	–1.68

*Test statistic significant at the 5% level

(a)



(b)

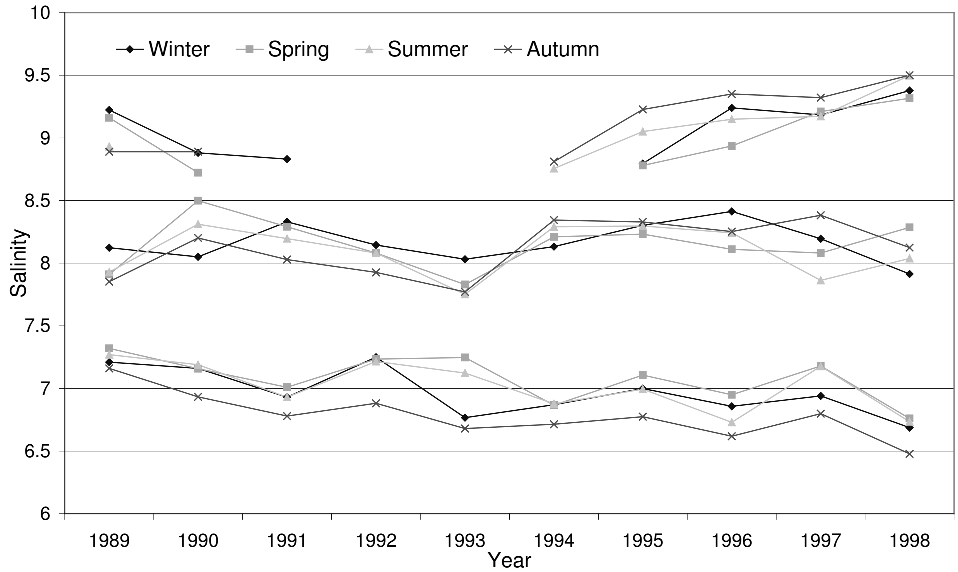


Figure 3. Time series of total phosphorus concentrations (a) and salinity (b) in the Western Gotland Basin divided into three salinity levels and four seasons.

5.2 Parametric trend test

We applied the parametric trend test to the same number of seasons as in the non-parametric approach (i.e., four). Figure 4 shows a histogram of 1,000 values of t^* for trends in total phosphorus in the Western Gotland Basin.

According to the bootstrap principle, this distribution is a good estimate of the true distribution of t. Therefore, we computed the 2.5 and 97.5 percentiles and used these to construct 95% confidence intervals for β_T through equation (4.2.4) and correspondingly for the three $\beta_{cl}^{(2)}$, and the results are given in Table 2.

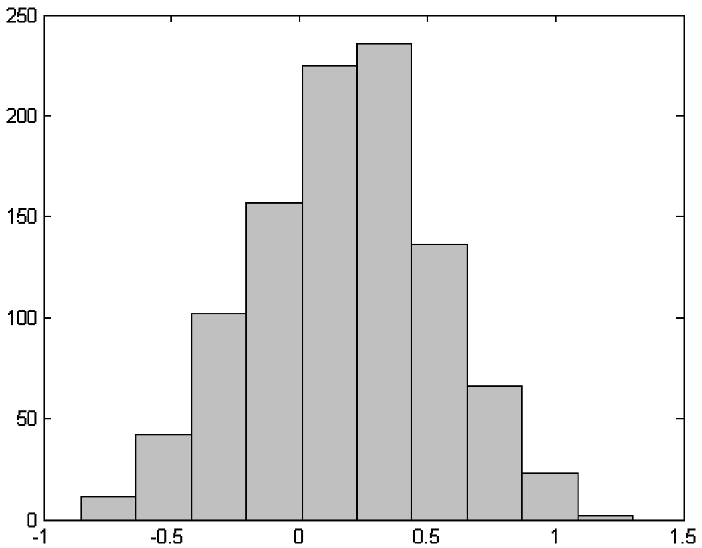


Figure 4. Histogram of 1,000 values of t^* for trends in total phosphorus in the Western Gotland Basin.

Table 2. Confidence intervals for the trend slope at different salinity levels and for the combined slope representing all levels

	Confidence interval	
Salinity level	Lower 95%	Upper 95%
Low	–0.0081	–0.0061
Medium	–0.0022	–0.00062
High	0.0213	0.0227
All levels combined	–0.0023	–0.00081

6 Conclusions and discussion

Variation in total phosphorus concentrations in the Baltic Sea is highly dependent on the mixing of waters of different origin. In particular, time series of such data is influenced by a large undulating fluctuation caused by inflowing salt- and phosphorus-rich water from the North Sea. The objective of our study was to reduce this natural variation by dividing the time series into subgroups according to levels of salinity. We examined trends for each of the series, and then combined the results to obtain an overall statistic for the entire water column.

Fryer and Nicholson (2002) used a similar approach when they divided mercury concentrations in muscle tissue into two categories on the basis of the body length of the fish and subsequently applied smoothers to examine trends. These investigators found it difficult to compute joint inference for both subseries due to the correlation between the two groups, but they proposed a way to circumvent this problem. In contrast, with the approaches we used, it was not difficult to combine the different test statistics into a test statistic for the entire water column, because it was possible to compute the covariances between the test statistics or trend estimates for the different layers, and the combined statistics were corrected for such dependencies.

Simmonds *et al.* (1997) have also split data into two series to examine trends in "background" and "polluted" air masses. By comparison, the series of data used in our study represented samples of Baltic Sea water, which exhibited different salinities due to varying amounts of mixing of waters of divergent origin (primarily fresh water from rivers and ocean water from the North Sea). Accordingly, our results of the individual trend statistics indicate a negative trend in total phosphorus in fresh waters.

The method we have proposed here is comparable to generally accepted strategies that take seasonal variation into account, and that use time series divided into different seasons (months) and compute a trend test statistic for each group. The success of such a procedure depends on the availability of several years of observations. However, the cycle under consideration in our study (i.e., fluctuation in phosphorus concentration caused by the mixing of salt and fresh waters) was much longer, more precisely about nine years, thus the available observations covered only one cycle. Accordingly, it was difficult to assess changes in the time series that were not due to this smooth fluctuation.

Consequently, we also suspect that the dependence of total phosphorus concentration on the level of salinity would persist after the separation into three salinity levels. A possibility to improve the proposed method is to include salinity as a numerical covariate, instead of solely for the purpose of classification. This can be easily achieved with a multivariate non-parametric trend test, such as the partial Mann-Kendall test (Libiseller and Grimvall, 2002), or by including salinity not only as dummy variable in the regression procedure.

The two different approaches we applied yielded comparable results. More precisely, both methods suggest that phosphorus concentrations are decreasing in waters with low salinity levels. The findings for the series with high salinity (representing the bottom layer of the Baltic Sea) indicate an increase in phosphorus, which might be explained by hypoxic conditions prevailing in recent years.

7 Acknowledgements

This work is part of the project *IMPACT* (EC-IST-1999-11313), which is funded by the IST Programme of the European Community.

8 Bibliography

Dietz, E.J. and Killeen, T.J. (1981). A nonparametric multivariate test for monotone trend with pharmaceutical applications. *Journal of the American Statistical Association* 76, 169-174.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics* 7, 1-26.

Efron, B. and Tibshirani, R.J. (1993). *An introduction to the bootstrap.* Monographs on Statistics and Applied Probability 57, Chapman &Hall.

Fryer, R.J. and Nicholson, M.D. (2002). Assessing covariate-dependent contaminant time-series in the marine environment. *ICES Journal of Marine Sciences* 59, 1-14.

Grimvall, A., Stålnacke, P., and Tonderski, A. (2000). Timescales of nutrient losses from land to sea- a European perspective. *Ecological Engineering* 14, 363-371.

Hirsch, R.M. and Slack, J.R. (1984). A nonparametric trend test for seasonal data with serial dependence. *Water Resources Research* 20, 727-732.

Kendall, M.G. (1975). *Rank Correlation Methods*. Charles Griffin, London.

Kreiss, J.P. and Franke, J. (1992). Bootstrapping stationary autoregressive moving-average models. *Journal of Time Series Analysis* 13, 297-317.

Lettenmaier, D. P. (1988). Multivariate nonparametric tests for trend in water quality. *Water Resources Bulletin* 24(3), 505-512.

Libiseller, C. and Grimvall, A. (2002). Performance of partial Mann-Kendall tests for trend detection in the presence of covariates. *Environmetrics* 13, 71-84.

Mann, H.B. (1945). Non-parametric tests against trend. *Econometrica* 13, 245-259.

Richardson, K. and Jørgensen, B.B. (1996). Eutrophication: Definition, History and Effects, in: *Coastal and Estuarine Studies: Eutrophication in Coastal marine Ecosystems.* Jørgensen B. B. and Richardson K. (eds.), American Geophysical Union.

Simmonds, P.G., Seurling, S., Nickless, G., and Derwent, R.G. (1997). Segregation and interpretation of ozone and carbon monoxide measurements by air mass origin at the TOR station Mace Head, Ireland from 1987 to 1995. *Journal of Atmospheric Chemistry* 28, 45-59.

Stigebrandt, A. (2001). Physical oceanography of the Baltic Sea, in F. Wulff *et al.* (Eds.*)*: *A Systems Analysis of the Baltic Sea*, Ecological Studies 148, Springer-Verlag.

Wulff, F., Rahm, L., and Larsson P. (*Eds.)* (2001). *A system analysis of the Baltic Sea*, Ecological Studies 148, Springer-Verlag.