

Identifying cardinal dates in phytoplankton time series to enable the analysis of long-term trends

Susanne Rolinski · Heidemarie Horn ·
Thomas Petzoldt · Lothar Paul

Received: 15 February 2006 / Accepted: 24 May 2007 / Published online: 12 July 2007
© Springer-Verlag 2007

Abstract Phenology and seasonal succession in aquatic ecosystems are strongly dependent on physical factors. In order to promote investigations into this coupling, methods of characterising annual time series of phytoplankton were derived and applied to a 31-year data set from Saldenbach Reservoir (Saxony, Germany). Field data are often scarce and irregularly sampled, particularly in the transition period from winter to spring, so reliable methods of determining cardinal dates in the time series are necessary. The proposed methods were used to determine the beginning, maximum and end of the spring mass development of phytoplankton by estimating the inflexion points (A), fitting a Weibull-type function (B) and fitting linear segments to the logarithmic

values (C). For the data set from Saldenbach Reservoir, all three methods proved to be relevant to the analysis of long-term trends. Differences between the maxima determined by the different methods seemed small, but there were deviations when the maximum was related to physical factors such as ice-out. The Weibull-type fit gave the most reliable and comprehensible results and is recommended for trend analyses. For all methods, long-term analysis of the duration of the spring mass development and the duration of the spring full circulation revealed a period of consistently low values (1975–1990) followed by a period of higher values (1990–2005). These periods were also identified for the date of ice-out, although in this case there was a period of high values followed by a period of low values. A sensitivity analysis that compared results from subsampled time series with increasing time intervals indicated that a minimum of one sample every three weeks is needed to obtain reliable results.

Communicated by Ulrich Sommer.

Priority program of the German Research Foundation—contribution 16.

Electronic supplementary material The online version of this article (doi:10.1007/s00442-007-0783-2) contains supplementary material, which is available to authorized users.

S. Rolinski (✉)
Institute of Hydrobiology, Technische Universität Dresden,
01062 Dresden, Germany
e-mail: Susanne.Rolinski@tu-dresden.de

H. Horn
Limnology of Reservoirs, Neunzehnhain Ecological Station,
Saxon Academy of Sciences, Neunzehnhainer Str. 14,
09514 Lengefeld, Germany

T. Petzoldt
Institute of Hydrobiology, Technische Universität Dresden,
01062 Dresden, Germany

L. Paul
Neunzehnhain Ecological Station, Technische Universität
Dresden, Neunzehnhainer Str. 14, 09514 Lengefeld, Germany

Keywords Spring mass development · Climate change · Saldenbach Reservoir · Limnology · Physical–biological coupling · Peak detection

Introduction

One of the most pronounced characteristics of plankton communities is their strong dependence on physical factors. In earlier investigations on long-term trends of plankton, characteristic dates were determined in a way that was adapted to the study sites (Adrian et al. 1999; Weyhenmeyer et al. 1999; Gerten and Adrian 2000; Blenckner et al. 2002; Straile et al. 2003b; Wiltshire and Manly 2004). The data-handling and aggregation performed to identify events in the seasonal cycle of the phytoplankton time series differed

among these earlier studies. In most cases, the day of year with the maximum biomass was considered to define the spring mass development or, when fluctuations perturbed the time series, a weighted average was determined (Wiltshire and Manly 2004; Edwards and Richardson 2004). These cardinal dates were then related to physical variables or proxies of them, such as the NAO index (Straile et al. 2003b), temperature thresholds or mean values (Wiltshire and Manly 2004) or ice-out (Adrian et al. 1999).

Meanwhile, the scope of the studies was widened in order to perform intercomparisons between different lakes (Straile and Adrian 2000; Weyhenmeyer 2001; Anneville et al. 2004; Mooij et al. 2005) and to obtain a mechanistic interpretation of cardinal dates in relation to the physical environment (and e.g. changing climatic conditions). However, a relationship between variables of different nature which stem from different measuring methods and sampling intervals is only meaningful when the definitions of characteristic points are objective and comparable between different studies. Thus, reliable and reproducible methods of determining cardinal dates are a prerequisite.

Analysis of climatological trends in biological systems should be undertaken in seasons when either the most drastic changes in the physical environment occur or when the coupling of biological to physical processes is strong (Horn et al. 1994; Adrian et al. 1999; Benndorf et al. 2001; Walther et al. 2002). In mid-latitudes, the closest link of the physical environment to the biological system is expected to occur in spring. Temperature, light and mixing conditions change from an unfavourable winter situation to more beneficial values for most species. During this period, nonphysical factors are not yet limiting. Later on in the year, trophic interactions and nutrient availability become more prominent and often mask the physical influence (Sommer et al. 1986). As a consequence, we focussed in this paper on the spring mass development (SMD) of phytoplankton.

Our aim was to provide methods of determining the cardinal dates corresponding to the start, maximum and end of phytoplankton spring mass development. Several methods were tested on long-term series of diatoms from the Saldenbach Reservoir (Saxony, Germany) (Horn and Horn 1995; Horn 2003a, 2003b), which have undergone climatic as well as trophic changes in the last 31 years. During this 31-year period, the yield and duration of the phytoplankton spring blooms have varied considerably, so these data may provide a basis for comparing methods (Horn et al. 1994; Horn and Horn 1995).

Objectifying the determination of cardinal dates enhances the possibility of performing year and study site comparisons for other ecological variables too, and thus enhances our ability to investigate and compare long-term trends reliably.

Study site

Saldenbach Reservoir (Table 1) is localised in the Erzgebirge (Saxony, Germany) and serves mainly as a store of drinking water. Its catchment is densely settled (86 ind km^{-2}) and used primarily for agricultural purposes (73%). It has been investigated continuously every week or every other week since 1975 (e.g. Horn 2003a, 2003b) in terms of biological (phytoplankton and zooplankton to species level), chemical (e.g. nutrients, oxygen, pH) and physical variables (temperature, turbidity among others). For a detailed description of the reservoir and the sampling techniques used, see Horn et al. (1994); Horn and Uhlmann (1995) and Paul (2003).

The reservoir is relatively deep (maximum depth 48 m, see Table 1), dimictic, and the winterly ice coverage lasts for two months on average. Stagnation periods occur during winter under ice and during summer between May and October/November. The durations of the mixing periods in autumn and spring depend on ice-on and ice-out. The day of ice-out was taken to be the first day in spring with less than 50% ice coverage of the reservoir surface and no subsequent freeze-up. Potential mixing occurs in spring when the reservoir is ice-free and the temperature at the bottom is below or equal to $4 \text{ }^{\circ}\text{C}$. Thus, the period between ice-out and the first date with a bottom temperature of above $4 \text{ }^{\circ}\text{C}$ was called the mixing period. Stable summer stratification was assumed when the surface temperature exceeded $10 \text{ }^{\circ}\text{C}$.

Methods

Characterisation of spring mass development

The methods we describe and compare here were used to characterise the SMD of phytoplankton. Typical SMDs are expected to yield high plankton abundances, beginning in spring and ending in early summer, often in a clear water phase (Sommer et al. 1986). In dimictic lakes they are usually initiated by the sudden improvement in the underwater light conditions after ice-out along with the alternating

Table 1 Characteristics of the Saldenbach Reservoir

Location	$13^{\circ}14'E$, $50^{\circ}44'N$ 25 km south of Chemnitz (Saxony, Germany)
Volume ^a	$22.4 \times 10^6 \text{ m}^3$
Surface ^a	146.4 ha
Maximum depth ^a	47.8 m
Average depth ^a	15.3 m
Mean retention time	0.7 a

^a Reservoir entirely filled, surface level at altitude 438.8 m

turbulent events and weakly stratified conditions that occur during the warming of the water column (Horn and Paul 1984; Huisman et al. 1999; Backhaus et al. 2003). The peak ends rather abruptly due to either high sedimentation losses (low turbulence) and/or nutrient depletion and/or zooplankton grazing after onset of thermal stratification (see e.g. Sommer et al. 1986; Horn and Horn 2000).

Even in water bodies that usually produce an SMD, such as Saldenbach Reservoir, its formation can vary considerably between years (Horn and Horn 1993). As an example, in 1978 the dominant group, the diatoms, clearly peaked after ice-out and broke down towards the middle of the year, so that the beginning and end of the development were easily identified (see Fig. 1, top), whereas in 2004 the development began before ice-out and included several peaks (see Fig. 1, bottom).

In order to develop procedures that can detect cardinal dates in all time series, some general rules or definitions are given.

Beginning of the spring phytoplankton mass development (BS)

We defined the beginning of the spring phytoplankton mass development as the beginning of a phase with exponential growth. For this determination, it was not the amount of biomass or a fixed threshold value that was crucial but instead its increase per time interval.

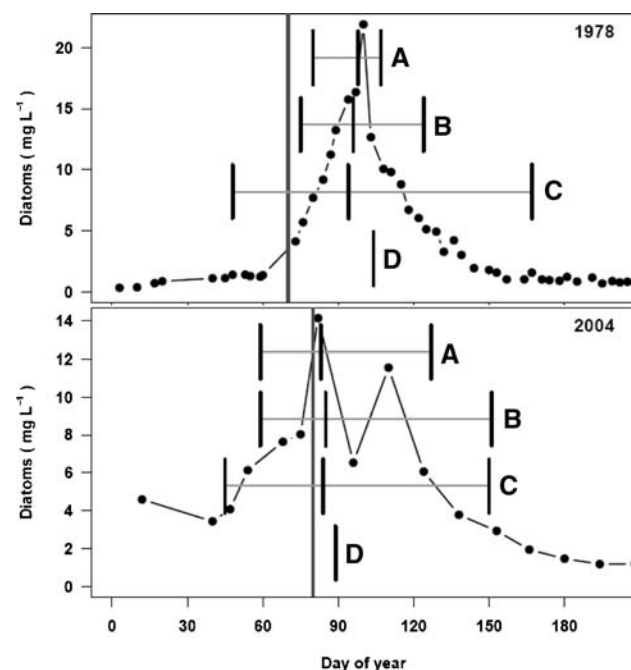


Fig. 1 Sum of diatoms in 1978 (top) and 2004 (bottom) and cardinal dates (BS, MS, ES, connected by grey horizontal line) for methods A–D. Day of ice-out is given as a vertical line

Maximum of the spring bloom (MS)

The maximum of the spring bloom coincides with the equilibrium of growth and mortality. This is not necessarily identical to the date of highest biomass. Phytoplankton biomass may change significantly day by day so that even weekly sampling may not ensure that sampling occurs on the actual day of the bloom maximum. Additionally, counting uncertainties might lead to the overrating of the date of an outlier. Nevertheless, a good estimate for the date of the maximum can be achieved by smoothing the rising and falling sides of the maximum.

End of the spring development (ES)

The end of the spring development is reached when the decreasing phase after MS passes into the summer development, which is a phase of lower values that fluctuate without a particular trend and, under high crustacean grazing pressure, is usually considered to mark the beginning of the clear water phase (Sommer et al. 1986). This transition may occur at any biomass level, so that its definition cannot involve a simple biomass threshold value.

Deviations from this pattern do, however, exist, especially in monomictic deep (e.g. Lake Constance, Straile et al. 2003a) or shallow lakes (e.g. Müggelsee, Adrian et al. 1999) with large surfaces (wind fetches) and, even in strong winters, only brief ice coverage (also observed in the Netherlands, as reviewed by Mooij et al. 2005). In such waters, spring algal biomass frequently does not exceed winter values, or several biomass peaks evolve in quick succession, driven by nutrient availability and weather conditions (light, mixing). Additionally, high values in summer may not permit us to postulate a clear water phase. In those cases, a characteristic spring bloom cannot be identified.

For this study, the sum of diatoms is considered to represent the phytoplankton community, since they are the dominant species in spring in Saldenbach Reservoir and they form almost the entire SMD. In spring, 80–90% of the phytoplankton biomass belongs to diatoms and in summer at least 28% (Horn 2003a). The taxa *Asterionella formosa*, *Fragilaria crotonensis*, *Diatoma elongatum*, *Aulacoseira italica*, *Synedra* spp. and centric diatoms shared most of the diatom biomass and alternated in terms of their proportions of the total biomass in different years.

Determination of relevant time interval

In order to identify the cardinal dates of the SMD, it is necessary to understand what we mean by “spring” (the

time window of interest) and “mass development” (the required biomass).

In order to exclude summer and autumn developments, only the first half of the year before midyear was taken into account, here taken to be the 19th of July or day 200. Years in which the maximum biomass occurred before March or under ice were considered to be winter developments and were omitted too. To be recognised as an SMD, the biomass should exceed 1 mg l^{-1} for the mesotrophic Saldenbach reservoir, where summer mean values (April–September) of between 1 and 5 mg l^{-1} are expected (Vollenweider 1976). In the application shown here, a spring bloom of diatoms (sum of all species) occurred in each of the 31 years available. This may not be the case when the SMD of just one species is studied; e.g. *Asterionella formosa* did not exhibit a SMD in six of these 31 years.

In a second step, relevant peaks were detected. Sampling dates that exhibited higher biomasses than before and after were initially considered as peaks, and the number of peaks was successively reduced by identifying subpeaks. Subpeaks are characterised by either a biomass of less than 10% of the maximum value or by a sufficiently good connection to the neighbouring peak. To evaluate the connection, neighbouring peaks were compared with the smallest value in between. When the smaller of the two peaks did not exceed a threshold multiple (here heuristically determined as nine-fourths) of the pit in between the neighbouring peaks, it was tagged as subpeak. This procedure was repeated until only relevant peaks remained. When more than one peak remained, the time series was limited to the sequence between the minima before and after the first peak after ice-out. The relevant peak was determined solely to identify the time window of interest, whereas the following methods were applied to the original time series.

We used four methods to describe the cardinal dates BS, MS and ES. Details of the algorithms are given in Appendix A.

Method A: points of inflexion

MS_A was taken to be the first peak after ice-out of the smoothed curve (Fig. 2). To determine the beginning and the end of the SMD, data were smoothed and logarithmically transformed. Differences were evaluated with a lag of one week. Peaks (pits) of the differences denote maximum increases (decreases). Therefore, the date of the first relevant peak difference before MS_A was taken to be BS_A , and the pit after MS_A to be ES_A . The term “relevant” denoted a preference for dates after ice-out for BS_A and before midyear for ES_A .

Method B: Weibull function

A function of Weibull-type is characterised by an exponential transition between zero and one. It can be extended to the product

$$w(x) = \left(d + \exp \left(- \left(\frac{x}{e} \right)^f \right) \right) \times \left(1 - a \exp \left(- \left(\frac{x}{b} \right)^c \right) \right) \quad (1)$$

where x is an independent variable (in days of the year) and a, \dots, f are free parameters to be fitted. The function allows for different slopes in the increasing and decreasing sections as well as nonzero baselines before and after.

The first step was to determine the parameters a, \dots, f by applying a nonlinear regression of the function w to the values after they had been scaled to the interval [0,1]. The date of the maximum of w was taken as MS_B . Second, the area under the curve was calculated as the integral of the daily values before and after MS_B separately (see Fig. 2). BS_B denoted the date of the 10% quantile of the area before MS_B and ES_B denoted the date of the 90% quantile of the area after MS_B .

Method C: linear segments

Linear segments were defined that follow the logarithmic values. The time series was divided into four segments including a constant beginning, increasing and decreasing sections, and a constant end. The dates at which the segments intersected were varied between the beginning of the year and midyear, respectively. BS_c , MS_c and ES_c ensured the smallest deviation from the measured values.

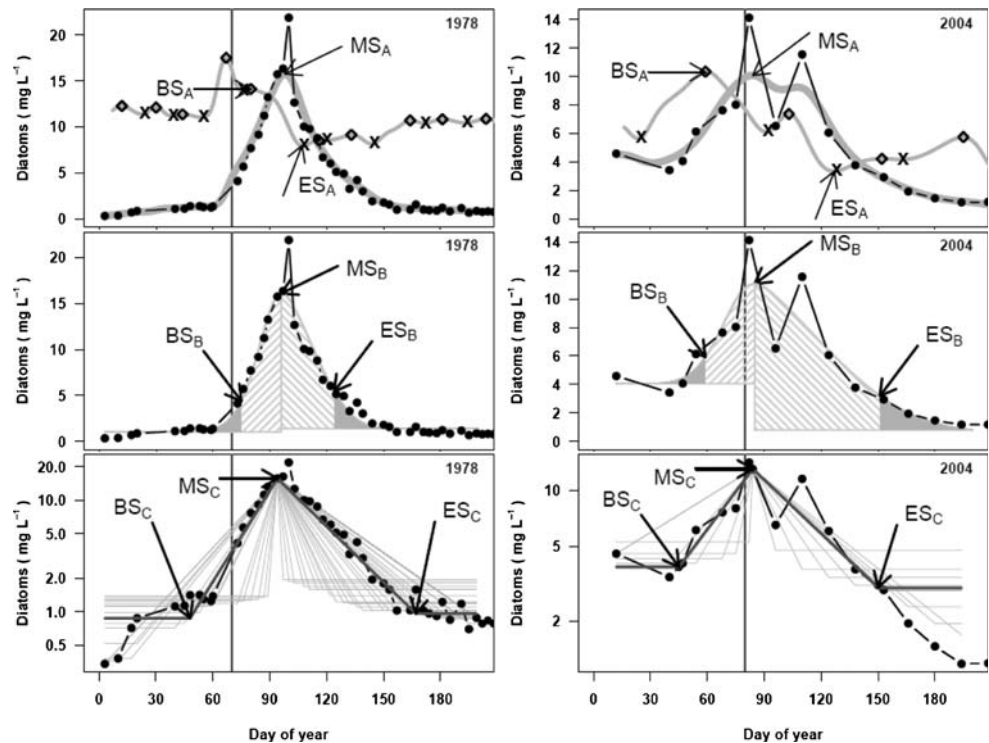
Method D: mean diatom day

Method D was included to provide a comparison with earlier studies. The biomass weighted average day, the mean diatom day (after Wiltshire and Manly 2004) was defined as $MS_D = \sum y_i d_i / \sum y_i$ where the biomass was y_i on day d_i .

Sensitivity analysis

All methods were applied to data series with reduced sampling dates in order to analyse their sensitivity to the time lag between two consecutive samples. Since most sampling programs result in equally spaced dates, subsets with regular intervals of 14, 21 and 28 days were chosen. The results from applying methods A–D to these subsets were checked for deviations from the results obtained when the original time series was analysed, and the methods

Fig. 2 Schematic derivations of methods A (top), B (middle) and C (bottom, note logarithmic scale) for the years 1978 (left) and 2004 (right). Method A: strongly smoothed curve (thick grey) for the determination of MS; derivative (thin grey) with peaks (diamonds) and pits (crosses) for the selection of BS and ES. The derivative (dv) is transformed for better visibility to $\max(\text{data}) \cdot (0.5 + 0.3\text{dv}/\max(\text{ldvl}))$. Method B: fitted Weibull function (grey) and area below the curve before (shading upwards) and after MS (shading downwards). BS and ES occur at the dates of 10% quantiles (filled in grey). Method C: possible fits with linear segments (thin grey lines) and fit with smallest deviation from values on logarithmic scale (dark grey line) which determines BS and ES. Day of ice-out is given as a vertical line



were compared with respect to the mean, median and quantile values.

Trend analysis and breakpoint determination

Long-term developments were analysed with respect to trends and breakpoints, i.e. structural changes in either the mean value or in the slope. Breakpoints that indicate a change in the mean value (i.e. a step function) can be visualised by the normalising transformations (also known as z-scores) $z_i = \frac{(x_i - \bar{x})}{s}$, where \bar{x} is the sample mean and s is the standard deviation. Summing these z-scores cumulatively permits changes (and the significance of each change) to be identified via an OLS-CUSUM test (Zeileis et al. 2003).

To determine breaks in the trends in the time series, the “breakpoints” method of Zeileis et al. (2003) was applied, which is available as a function in the package *strucchange* (Zeileis et al. 2002) in the data analysis system R (R Development Core Team 2005). In this method, the optimal number of breakpoints is determined using a model selection approach and by applying a series of piecewise linear regression models between phytoplankton biovolume and time. The minimum adequate model (i.e. the optimal number of breakpoints) is the model with the lowest Bayesian information criterion (BIC), which is a measure of the compromise between maximising the goodness of fit and minimising the number of parameters.

Trends were derived using the partial Mann–Kendall test of Hamed and Rao (1998) and Yue and Wang (2004). The trend itself was determined as the median value of all differences between the values and their successors (see formula and determination of P value in Appendix B). Trends were considered to be significant when $P < 0.05$. Trends were determined either for the time series from 1975 to 2005 or for subsets, depending on the estimated breakpoints.

Correlation analysis

Relations between time series were determined by Spearman correlation coefficients (estimated as Spearman’s rank coefficient due to possible ties).

Results

Each method was applied to each year of the 31-year data set of the Saldenbach Reservoir. Two years with typical but not extreme formations of an SMD were chosen, 1978 and 2004, to illustrate the methods. The development of diatoms in 1978 (Fig. 1, top) followed the peak formation as described above. Exponential growth after ice-out (day 70) was followed by a rapid decline after the beginning of summer stratification (day 108). The situation was different in 2004, when the highest diatom biomasses were again formed after ice-out (Fig. 1, bottom) but the time series

was characterised by a double maximum, considerable diatom growth under ice (initiated in a short period in February without ice coverage), and a smaller number of sampling dates. The latter aspect means that each value determined and its counting error becomes more important.

For well-formed peaks like the one for 1978 (Fig. 1, top), the determined date of the maximum (MS) varied only slightly between day 94 and 104 depending on the method. However, even in this case, differences between methods in terms of the beginning and end of the SMD could be detected. Although method A used slightly smoothed curves and including a time lag in the temporal gradient determination, it was the most sensitive of the methods to differences in biomass increase (Fig. 2, top left). It yielded later beginnings and earlier ends than the other methods. In contrast, method C (Fig. 2, bottom left) gave earlier beginnings and later ends than the other methods, as the linear segments used only roughly approximated the shape of the curve. Method B (Fig. 2, middle left) therefore seemed to be the most appropriate. The Weibull function considered the onset before and after the peak as well as differences in biomass increase and decline. This method determined the date when a certain proportion of the biomass under the curve was reached reasonably well, regardless of whether small irregularities were present in the time series.

In years with fewer sampling dates and a double maximum, such as 2004 (Fig. 1, bottom), the determined maximum (MS) tended to coincide with the peak with the highest biomass, even though the peak contained only one measurement. Here, even strong smoothing in method A (Fig. 2, top right) did not cause the two peaks to merge. Also, the Weibull function fitted (Fig. 2, middle right) was tilted to the left towards the maximum of the curve. This was also the case for the linear segments (Fig. 2, bottom right). While the rapid increase that occurred just before ice-out led to similar estimates for the beginning from all of the methods, the slower decline of the peak resulted in significant deviations between the estimates the different methods provided for the end. The earliest end, ES, was given by method A (Fig. 2, top right), following the

sharpest decline after the second peak. For the Weibull function (Fig. 2, middle right), the relatively small offset and the large area under the second half of the peak caused method B to give a later ES than that given by method A, while method C (Fig. 2, bottom right) gave a similar ES value to method B.

This drift in the cardinal dates became obvious when displayed in box-and-whisker plots for all years (Fig. 3). Results for the maximum were rather similar for all methods. MS ranged from day of the year 58 to 173 (median between day 111 and 115) (Fig. 3, left), and from day 3 to 103 (median between 38 and 43) with respect to ice-out (Fig. 3, right). The plots showed that there was a tendency for method C to give the earliest BS and the latest ES determination no matter whether day of year was considered or day after ice-out (median values of 71, BS, and 174, ES, in comparison to 86 and 149 for method A).

An analysis of the sensitivity of the methods to a decrease in the number of sampling dates was performed. Data points were reduced so that the intervals between sampling dates were at least 14, 21 or 28 days. Thus, the sampling interval which ranged originally from 4 to 33 days (mean value 11 days) was extended so that in the three subsampled series mean values of 17, 23 and 29 days were realized. The sensitivity to the sampling interval (in terms of the variation in the determined cardinal dates) differed among the methods (Fig. 4). Generally, the deviations from the cardinal dates determined for the original time series increased as the sampling interval increased, as shown by the box-and-whisker difference plots. The median values for MS only deviated from zero (by 1 or 2 days) for monthly sampling, and the quantiles increased to eight days at maximum (method C). The beginning of the SMD, BS, occurred later as the sampling interval increased [median values between -3 (B) and -5 days (C); quantiles between -9 and -13 days (C)]. The opposite was true for the end of the SMD, which occurred earlier as the sampling interval increased [median values between 1 (C) and 6 days (A); quantiles between 7 (B) and 19 days (C)]. With a few exceptions (e.g. ES_A), the variance increased with the sampling interval. Taking the difference between the 25%

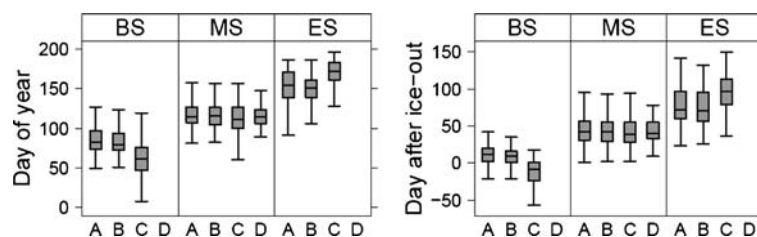


Fig. 3 Boxplots of the cardinal dates for diatom development in terms of day of year (*left*) or day after ice-out (*right*) for methods A–D for all years studied. Each *box* covers the range between quartiles,

with the median given as a *line* and the standard deviation as a *whisker*

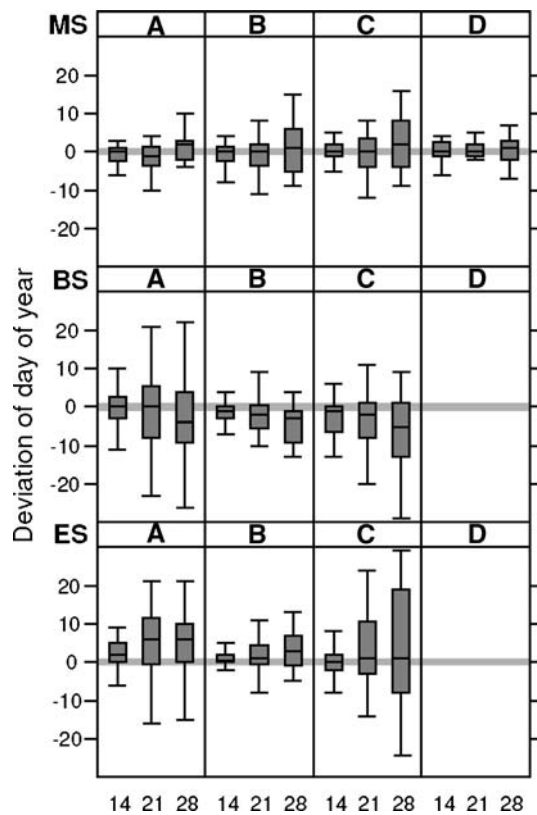


Fig. 4 Boxplots of the deviations of the cardinal dates determined using a reduced time series (sampling interval of 14, 21 and 28 days) from those determined using the original time series for methods A to D. Each *box* covers the range between quartiles, with the median given as a *line* and the standard deviation as a *whisker*

and 75% quantiles as a measure of the deviation from the original cardinal date, the deviations were smaller for MS (A:5, B:11, C:12, D:5 days) than for BS (A:13, B:8, C:14 days) and ES (A:10, B:8, C:27 days).

Year-to-year variations from 1975 to 2005 in the timing of the diatom SMD (Fig. 5) differed considerably between methods. Up to 1989, ice-out ranged between day 65 and 102; it dropped in 1990 to day 19 and then exhibited a decreasing trend with large interannual variations between 1991 and 2002 (significant trends between 1975 and 1987 of 1.92 days per year, $P = 0.024$, and between 1984 and 2002 of -2.2 days per year, $P = 0.040$). In Saldenbach Reservoir, the length of the mixing period was inversely correlated to the day of ice-out ($r^2 = 0.68$, $P < 0.001$), whereas the beginning of the summer stratification was independent of the day of ice-out ($r^2 = 0.05$, $P = 0.23$) and the duration ($r^2 = 0.04$, $P = 0.30$) or end of the spring mixing period ($r^2 = 0.01$, $P = 0.55$). Thus, the later the ice melted, the shorter the period of strong mixing. The beginning of the SMD, BS, showed a clear relation to the day of ice-out and varied accordingly. Correlations to the day of ice-out were highly significant for all methods for

BS (Table 2). Values obtained using method A (Fig. 5, upper left) were later than those given by method B (upper right), and much later than those given by method C (lower left). The latter often determined that the beginning occurred before ice-breakup; this only occurred in exceptional years (e.g. 1986) when the other methods were used. The maximum of the SMD, MS, showed a trend towards earlier occurrence after 1992 (significant for method B, with slope -4.527 , $P = 0.002$, and method C, slope -4.846 , $P = 0.001$) and a higher year-to-year variability in the period before. For this criterion, correlations with day of ice-out were significant for all methods (Table 2), and differences between methods were small, although method D exhibited fewer variations throughout the investigation period. The end, ES, exhibited similar trends and variations to MS, but no correlations with ice-out or the duration of spring mixing were observed. The determined date of ES increased upon moving from method A to B to C, as discussed for the years 1978 and 2004. The tendency for the duration of the SMD (ES–BS) to increase when moving from method A to B to C was reflected in the time series. The duration was significantly correlated to ice-out and the duration of the spring mixing period (Table 2 for DS). The local climate, as described by the day of ice-out, was connected weakly but significantly ($r = 0.51$, $P = 0.004$) to large-scale climatic signals as the Winter North Atlantic Oscillation Index¹, but no correlations of the NAO with the determined cardinal dates could be found.

Despite systematic deviations between the methods concerning the timing of SMD, they all showed prolonged SMD durations since 1990 (Fig. 5). This impression was confirmed by the normalising transformations of the duration of the SMD (duration = ES – BS) (Fig. 6). All methods detected durations that were mostly shorter than the average (indicated by negative anomalies) between 1975 and 1990 and longer than the average afterwards. Cumulative sums of the anomalies that were >5 SD (compare the lines in Fig. 6) indicated consistent periods of low and high values and a reversal point in 1991. This step-like structural change in the time series was significant for all methods with respect to the mean value (P values from OLS-CUSUM-test for method A: $P < 0.001$, B: $P = 0.0011$, and C: $P = 0.01$), and the Mann–Kendall test gave a significant increasing trend over the whole period ($P < 0.05$ for all methods).

In order to determine trends in subsets of the cardinal date time series, breakpoints were calculated using stepwise linear regression models. Trend analysis with the Mann–Kendall test was performed for these sections (Fig. 7). This permitted the identification of multiple breakpoints. For MS, breakpoints were detected for methods A, B and C in

¹ source: <http://www.cgd.ucar.edu/cas/jhurrell/nao.stat.winter.html>

Fig. 5 Long-term trends in cardinal dates from 1975 to 2005 for the sum of diatoms for methods A–D and day of ice-out (grey line). The shaded area is enclosed by BS and ES; the black line denotes MS

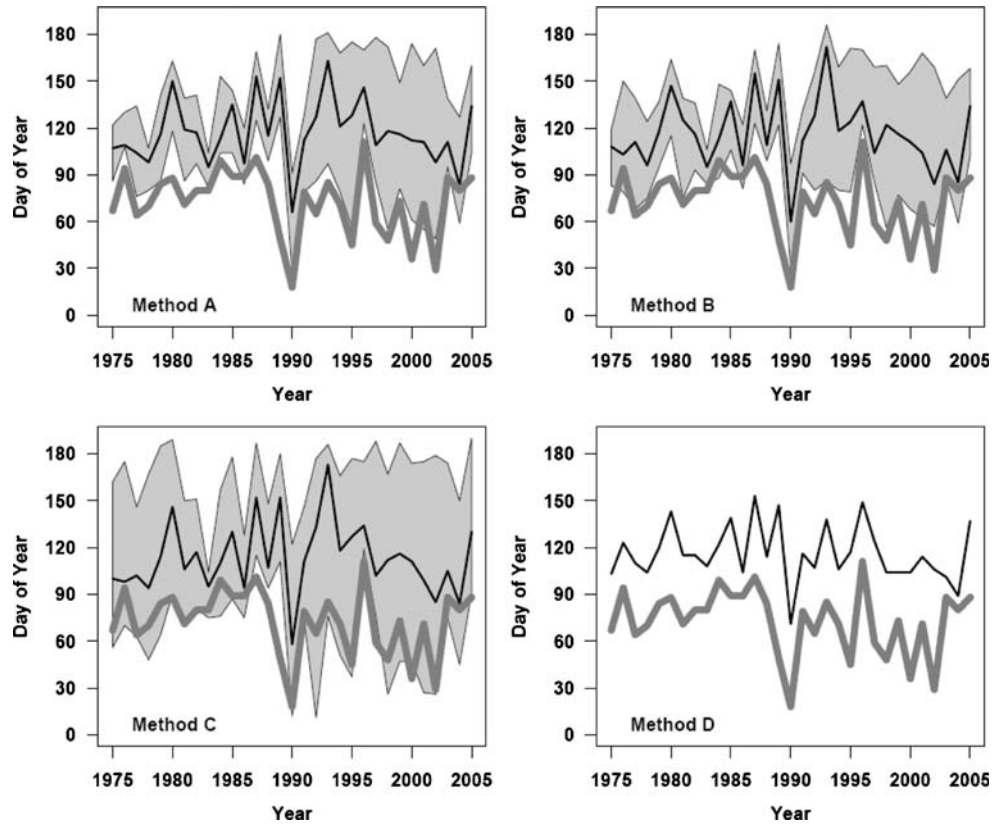


Table 2 Statistical analysis of correlations between MS, BS, ES and the duration of the SMD (DS) determined by methods A–D versus the duration of the mixing period in spring (mixis duration) and the day of ice-out

Method	Mixis duration		Ice-out	
	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>
MS				
A	-0.30	0.11	0.40	0.024*
B	-0.33	0.071*	0.41	0.021*
C	-0.28	0.13	0.38	0.032*
D	-0.47	0.008**	0.54	0.002**
BS				
A	-0.70	<0.001***	0.74	<0.001***
B	-0.60	<0.001***	0.66	<0.001***
C	-0.69	<0.001***	0.68	<0.001***
ES				
A	0.16	0.38	-0.05	0.79
B	-0.60	0.97	0.11	0.55
C	-0.01	0.94	0.11	0.56
DS				
A	0.69	<0.001***	-0.63	<0.001***
B	0.52	0.003**	-0.49	0.006**
C	0.58	<0.001***	-0.51	0.003

Significance levels are given according to *** < 0.001, ** < 0.01, * < 0.05, • < 0.1

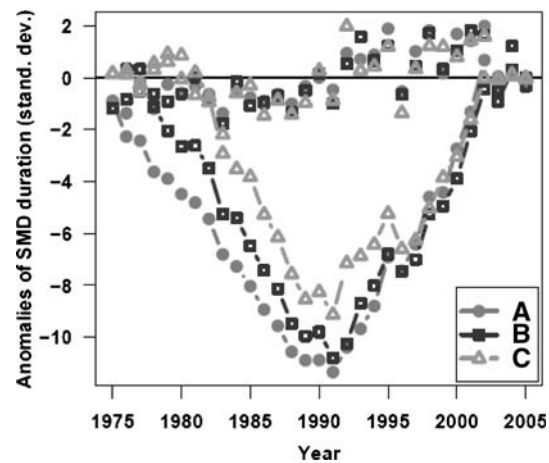


Fig. 6 Individual deviations from the long-term mean of the duration of SMD for methods A–C (symbols only) and cumulative sums (lines and symbols)

1989 and 1993, with significant negative trends in the last section. The beginning and end of the SMD produced heterogeneous results. For BS, method A gave a negative trend for the entire time series and method C gave positive trends in all sections. For ES, significant breakpoints for method A were surrounded by a positive and a negative trend, whereas method B gave an overall positive trend. A significant breakpoint was detected for the duration, DS

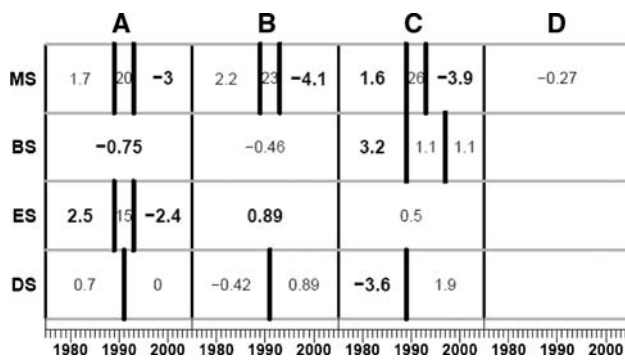


Fig. 7 Breakpoints and trend analysis for segments of MS, BS, ES and the duration DS of the SMD for methods A–D according to the partial Mann–Kendall test. Slopes are given for the entire time series or subsets between breakpoints (denoted by vertical lines). Significant slopes ($P < 0.05$) are given in **bold**

(methods A and B in 1991, C in 1989), but for method C the time series began with a significant negative trend. The variations in breakpoints and trends for BS and ES can be related to the method used and are discussed in the next section.

Discussion

Interannual variations in phytoplankton development can be detected when their seasonal occurrence is described systematically. A reduction of the data from a scattered time series to characteristic points simplifies comparisons between different years. In recent studies on the influence of the physical environment, the date of peak phytoplankton biomass has usually been considered to be a characteristic point (Adrian et al. 1999; Weyhenmeyer et al. 1999; Gerten and Adrian 2000; Winder and Schindler 2004; Wiltshire and Manly 2004). Others, focussing on zooplankton development and the beginning of the clear water phase, selected the date of the maximum (Scheffer et al. 2001) or the above-threshold values (Straile 2000) of the Secchi depth or the minimum chlorophyll *a* concentrations after SMD (Scheffer et al. 2001). This choice restricts the characteristic point to the set of sampling dates available, which are sometimes scarce and randomly selected. A continuous function not only smoothes and reduces counting errors, it also considers neighbouring values and accounts for autocorrelations in the time series. Furthermore, the selection of the maximum value to use in an investigation of climatic effects appears to be an unfortunate one, since it represents the point when the balance of growth and mortality of the plankton population changes towards the dominance of loss terms. Thus, MS is influenced by many more processes than just those directly related to climatic conditions. The beginning of an SMD,

on the other hand, corresponds to the date when the conditions for a spring bloom are favourable, and so it should be primarily controlled by physical factors. The derivation of this date via a clearly defined procedure is even more important in this case, since the counting errors are especially high when phytoplankton abundances are low.

Comparing the results from the chosen methods, each method appears to exhibit distinct properties. The determination of the maximum, MS, with method A gave reliable estimates when sampling dates were rare and they deviated within a reasonable range (2–3 days). An appropriate bandwidth had to be chosen in relation to the sampling rate in order to smooth the time series effectively. When applying method A, the determination of BS and ES was a rather sophisticated process of eliminating subpeaks, which involved choosing the appropriate time window in which the peaks and pits of the derivative of the time series were ordered according to relevance. Its results were rather sensitive to the shape of the phytoplankton development distribution due to the need to identify inflexion points. This sensitivity led to strong systematic deviations (later values for BS and earlier for ES) when the sampling interval was over two weeks. Thus, method A detected points of abrupt change in the seasonal phytoplankton development.

Method B was very successful when time series were restricted to an appropriate period. When there were rather low values before and after the peak, the function adjusted to almost any kind of peak very flexibly. It required the application of nonlinear regression, but estimates of the cardinal dates were plausible and independent of small fluctuations at any period in the time series. Even when the sampling interval was three weeks or more, deviations were small and there were no systematic shifts in the median value. Thus, method B reproduced the overall properties of the peak in terms of location, scale and shape.

Method C comprised a rather robust but also rough approach. The endpoints BS and ES were close to the anticipated values when there was clear peak formation but could deviate substantially with scattered winter or summer values. Consequently, in well-defined cases method C was not better than the other methods, whereas in more problematic years the most conservative estimates were derived. This is also reflected in the rather large deviations it yielded for sampling intervals of more than two weeks. Thus, method C gave the largest estimate for the duration of significant biomass occurrence.

Method D was only suited to determining the maximum but was easy to apply. The overall variations were smaller for this method, and the estimated dates for MS tended towards the center of the time window considered. Without the selection of a relevant peak beforehand, the values of subsequent peaks were also included in the calculation, and

the MS was then calculated to occur later in the year. This characteristic made the method rather insensitive to the sampling interval, so deviations due to fewer data points were small. Thus, method D was a reliable but less sensitive indicator of the cardinal date MS.

Concerning the duration of the SMD, all methods yielded values below the mean in the period before 1991 and above the mean thereafter (Fig. 6). Consequently, the cumulative sums of the anomalies decreased in the first and increased in the second period. Both periods were characterised by different means, medians and variances, but a trend in the first section was only detected by method C. Lower values before 1991 and higher afterwards were observed for the duration of spring full circulation (corresponding to the duration of the SMD) and vice versa (i.e. higher values initially then lower values) for the day of ice-out (Horn 2003a). Since the diatom growth was restricted to the period of mixing in Saldenbach Reservoir (Horn 2003a) and the beginning of the summer stratification was relatively fixed (standard deviation of nine days in comparison to 22 days for ice-out), the duration of the SMD was significantly correlated with the mixing period and ice-out (Table 2). The later the ice melted, the shorter the period of strong mixing and, consequently, the shorter the SMD. Reverse correlations are valid for the beginning but none for the end of the SMD (Table 2). Thus, variations in ice-out caused the variations in the mixing period and determined the beginning and the duration of the SMD. Those relations were postulated by Horn (2003a) for the Saldenbach Reservoir (independent of the actual nutrient supply), but they can now be substantiated by statistical analysis using our methodology. Our approach thus allowed a more comprehensive and unbiased assessment of such relations (e.g. Adrian et al. 1999; Weyhenmeyer et al. 1999; Gerten and Adrian 2000, 2001; Blenckner et al. 2002). Preliminary results obtained by applying the algorithms to time series for five lakes and mesocosms including monomictic and shallow lakes were promising and could be achieved without further modifications.

The correlations of the cardinal dates considered here (BS, MS, ES) with ice-out decreased the later the dates appeared in the year. The beginning (BS) was strongly correlated, the maximum (MS) was significantly but weakly correlated, and the end (ES) was found to be independent of the date at which the ice thawed (Table 2). This agreed with other findings about the persistence of climatic signals from winter to summer (Straile et al. 2003b; Edwards and Richardson 2004). In deep and dimictic lakes, the ice coverage hinders heat exchange and kinetic forcing between the atmosphere and the water during winter, so that the strongest impact on the entire water column occurs in spring before the onset of stratification (Straile et al. 2003b). Thus, the effect of cold or

warm winter conditions on the epilimnion lasts only until the end of the mixing period. For the Saldenbach Reservoir and similar water bodies, it seems appropriate to relate the beginning of the SMD to physical factors. However, in the case of proceeding winter warming and an absence of ice coverage, the mixing regime may become monomictic and the phytoplankton composition and development may change completely, as suggested earlier (Blenckner et al. 2002; Walther et al. 2002; Winder and Schindler 2004; Durant et al. 2005; Köhler et al. 2005).

When considering the changes and trends in the times series for determined cardinal dates (Fig. 7), two break-points were detected for the maximum MS identically by methods A, B and C; this was not the case for method D, however, suggesting an artefact or predominant random effects. For the beginning BS, only method C determined breakpoints. However, the second breakpoint for method C in 1997 resulted from the beginning being found by method C to occur before ice-out in 1992 and 1995 (compare Fig. 5), which overrated small increases in biomass, like in 1978 (compare Fig. 1). For the end, ES, overall positive trends were determined by methods B and C (significant for B). Here, method A gives similar results to those it gave for MS, presumably because, among all of the methods, it gives the ES closest to the MS (compare Fig. 1). For the duration DS, all of the methods gave one breakpoint. For method C again the very early values for BS in 1992 and 1995 produced a slight deviation in the breakpoint (1989 instead of 1991, as given by methods A and B) and a negative trend in this period. Thus, method B provided results that give reliable and individually well-adapted estimates not only for the maximum but also the more sensitive cardinal dates (BS, ES, DS).

In most investigations, the day of the year with maximum biomass was taken as MS (Adrian et al. 1999; Gerten and Adrian 2000; Weyhenmeyer 2001; Winder and Schindler 2004), and sampling intervals ranged between weekly and monthly. In winter, samples were often taken once a month (Winder and Schindler 2004) or fortnightly (Gerten and Adrian 2000), so that the transition to the beginning of an SMD was covered by only few data points. The sensitivity of any correlation approach to investigating the effects of the climate on plankton dynamics to the underlying data was obvious; the inclusion of more years of data in the analysis of long-term data from Müggelsee led to different conclusions concerning the influence of the NAO on the *Daphnia* spring peak (compare Straile and Adrian 2000; Gerten and Adrian 2000). Similarly, slight differences in cardinal dates related to plankton dynamics (when estimated by different methods) may influence the analysis of climate-driven changes. This calls for the use of clearly defined methods that are as objective as possible when identifying such cardinal dates. The evaluation of the

correlation between cardinal dates and climatic factors in the Saldenbach Reservoir, which was only touched upon here, clearly demonstrates the potential of the proposed methods.

Conclusions

Among the proposed approaches, method B, which used a fit to a Weibull-type function, was the most promising for the detection of climate signals in long-term plankton series. It required standard programming techniques and a sufficient data set but it gave reliable and comprehensible results. Method A was preferred when there were rapid changes in the series, and method C evaluated the maximum duration of the SMD. The results from Method B lay somewhat between those from A and C when 10% of the area under the curve was chosen as threshold, but this can be adjusted to narrower or wider selections with other values for the quantiles. The properties of the methods were consistent and allowed us to perform trend analyses. Values for the maximum (MS) were consistently determined with at least triweekly data, whereas the beginning and end of the SMD were much more sensitive, so bi-weekly data sampling is recommended, especially in spring. The proposed methods proved suitable for systematizing the determination of cardinal dates, which is a precondition for quantifying their relationships to control factors such as physical (ice-out, water temperature or mixing conditions) and chemical variables (nutrient concentrations).

In this work, the proposed methods were tested for just one water body, since our aim was to provide a detailed description of their application and to derive conclusions concerning their limits and capabilities. Preliminary analyses show that these methods can be applied to data sets from different water bodies to give reliable and comprehensible results, just as they did for Saldenbach Reservoir. Despite a certain degree of complexity, they can easily be adapted into standard routines (in our case implemented in R) in further applications. These methods used to characterise cardinal dates in plankton time series may provide valuable tools for the integration of various data sets and for generating a more comprehensive view of the variables that drive ecological phenomena and processes.

Acknowledgments This study is supported by the German Research Foundation (DFG) under grant (PA 1202/1) of priority programme 1162 AQUASHIFT. The authors are grateful to Karsten Rinke and Stephan Hülsmann for fruitful discussions, Otto Richter for bringing up the Weibull function, and two anonymous referees for their detailed, persevering and constructive suggestions and their encouragement of the sensitivity analysis.

References

- Adrian R, Walz N, Hintze T, Hoeg S, Rusche R (1999) Effects of ice duration on plankton succession during spring in a shallow polymictic lake. *Freshw Biol* 41(3):621–634
- Anneville O, Souissi S, Gammeter S, Straile D (2004) Seasonal and inter-annual scales of variability in phytoplankton assemblages: comparison of phytoplankton dynamics in three perialpine lakes over a period of 28 years. *Freshw Biol* 49:98–115
- Backhaus JO, Hegseth EN, Wehde H, Irigoien X, Hatten K, Logemann K (2003) Convection and primary production in winter. *Mar Ecol Prog Ser* 251:1–14
- Benndorf J, Kranich J, Mehner T, Wagner A (2001) Temperature impact on the midsummer decline of *Daphnia galeata*: an analysis of long-term data from the biomanipulated Bautzen Reservoir (Germany). *Freshw Biol* 46:199–211
- Blenckner T, Omstedt A, Rummukainen M (2002) A Swedish case study of contemporary and possible future consequences of climate change on lake function. *Aquat Sci* 64:171–184
- Durant JM, Hjermmann DO, Anker-Nilssen T, Beaugrand G, Mysterud A, Pettorelli N, Stenseth NC (2005) Timing and abundance as key mechanisms affecting trophic interactions in variable environments. *Ecol Lett* 8(9):952–958
- Edwards M, Richardson AJ (2004) Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430:881–884
- Gerten D, Adrian R (2000) Climate-driven changes in spring plankton dynamics and the sensitivity of shallow polymictic lakes to the North Atlantic Oscillation. *Limnol Oceanogr* 45(5):1058–1066
- Gerten D, Adrian R (2001) Differences in the persistency of the North Atlantic Oscillation signal among lakes. *Limnol Oceanogr* 46(2):448–455
- Hamed KH, Rao AR (1998) A modified Mann–Kendall trend test for autocorrelated data. *J Hydrol* 204(1–4):182–196
- Horn H (2003a) The relative importance of climate and nutrients in controlling phytoplankton growth in Saldenbach reservoir. *Hydrobiologia* 504:159–166
- Horn W (2003b) Long-term development of the crustacean plankton in the Saldenbach reservoir (Germany)—changes, causes, consequences. *Hydrobiologia* 504:185–192
- Horn H, Horn W (1993) Sedimentary losses in the reservoir Saldenbach: flux and sinking velocities of dominant phytoplankton species. *Int Revue ges Hydrobiol* 78(1):39–57
- Horn W, Horn H (1995) Interrelationships between crustacean zooplankton and phytoplankton. *Hydrobiologia* 307:231–238
- Horn H, Horn W (2000) Sedimentation—the main loss factor in waters dominated by diatoms. Results of long-term investigations. *Int Rev Hydrobiol* 85(2–3):191–208
- Horn H, Paul L (1984) Interactions between light situation, depth of mixing and phytoplankton growth during spring period of full circulation. *Int Revue ges Hydrobiol* 69(4):507–519
- Horn H, Uhlmann D (1995) Competitive growth of blue–greens and diatoms (*Fragilaria*) in the Saldenbach Reservoir, Saxony. *Wat Sci Tech* 32(4):77–88
- Horn W, Horn H, Paul L (1994) Long-term trends in the nutrient input and in-lake concentrations of a drinking water reservoir in a dense populated catchment area (Erzgebirge, Germany). *Int Revue ges Hydrobiol* 79(2):213–227
- Huisman J, Van Oostveen P, Weissing FJ (1999) Critical depth and critical turbulence: two different mechanisms for the development of phytoplankton blooms. *Limnol Oceanogr* 44(7):1781–1787
- Köhler J, Hilt S, Adrian R, Nicklisch A, Kozerski HP, Walz N (2005) Long-term response of a shallow, moderately flushed lake to

- reduced external phosphorus and nitrogen loading. *Freshw Biol* 50:1639–1650
- Mooij WM, Hülsmann S, Domis LND, Nolet BA, Bodelier PLE, Boers PCM, Pires LMD, Gons HJ, Ibelings BW, Noordhuis R, Portielje R, Wolfstein K, Lammens EHRR (2005) The impact of climate change on lakes in The Netherlands: a review. *Aquat Ecol* 39(4):381–400
- Paul L (2003) Nutrient elimination in pre-dams: results of long-term studies. *Hydrobiologia* 504:298–295
- R Development Core Team (2005) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, <http://www.r-project.org>. Cited 8th June 2007
- Scheffer M, Straile D, van Nes EH, Hopper H (2001) Climatic warming causes regime shifts in lake foodwebs. *Limnol Oceanogr* 46:1780–1783
- Sommer U, Gliwicz ZM, Lampert W, Duncan A (1986) The PEG-model of seasonal succession of planktonic events in fresh waters. *Arch Hydrobiol* 106(4):433–471
- Straile D (2000) Meteorological forcing of plankton dynamics in a large and deep continental European lake. *Oecologia* 122(1):44–50
- Straile D, Adrian R (2000) The North Atlantic Oscillation and plankton dynamics in two European lakes—two variations on a general theme. *Glob Change Biol* 6(6):663–670
- Straile D, Jöhnk K, Rosseknecht H (2003a) Complex effects of winter warming on the physicochemical characteristics of a deep lake. *Limnol Oceanogr* 48(4):1432–1438
- Straile D, Livingstone D, Weyhenmeyer GA, George DG (2003b) The response of freshwater ecosystems to climate variability associated with the North Atlantic Oscillation. In: Hurrell JW, Kushnir Y, Ottersen G, Visbeck M (eds) *The North Atlantic Oscillation—climatic significance and environmental impact* (AGU Geophysical Monograph Series). American Geophysical Union, Washington, DC, 134:263–279
- Venables QN, Ripley BD (1999) *Modern applied statistics with S-plus*, 3rd edn. Springer, New York
- Vollenweider RA (1976) Advances in defining critical loading levels for phosphorus in lake eutrophication. *Mem Ist Ital Idrobiol* 33:53–83
- Walther GR, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, Fromentin JM, Hoegh-Guldberg O, Bairlein F (2002) Ecological responses to recent climate change. *Nature* 416:398–395
- Weyhenmeyer GA (2001) Warmer winters: are planktonic algal populations in Sweden's largest lakes affected? *Ambio* 30(8):565–571
- Weyhenmeyer GA, Blenckner T, Pettersson K (1999) Changes of the plankton spring outburst related to the North Atlantic Oscillation. *Limnol Oceanogr* 44(7):1788–1792
- Wiltshire KH, Manly BFJ (2004) The warming trend at Helgoland Roads, North Sea: phytoplankton response. *Helgoland Mar Res* 58(4):269–273
- Winder M, Schindler DE (2004) Climate change uncouples trophic interactions in an aquatic ecosystem. *Ecology* 85(8):2100–2106
- Yue S, Wang C (2004) The Mann–Kendall test modified by effective sample size to detect trend in serially correlated hydrological series. *Water Resour Manage* 18(3):201–218
- Zeileis A, Leisch F, Hornik K, Kleiber C (2002) strucchange: An R package for testing for structural change in linear regression models. *J Stat Softw* 7(2):1–38
- Zeileis A, Kleiber C, Krämer W, Hornik K (2003) Testing and dating of structural changes in practise. *Comput Stat Data Anal* 44(1–2):109–123