LIMNOLOGY
and
OCEANOGRAPHY: METHODS

*Limnol. Oceanogr.: Methods* 11, 2013, 213 224


# Comparison of linear and cubic spline methods of interpolating lake water column profiles

*Ryan P. North*[1,2*] *and David M. Livingstone*[1]
[1]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Department of Water Resources and Drinking Water, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland
[2]ETH Zurich, Department of Environmental Systems Science, Institute of Biogeochemistry and Pollution Dynamics, Universitätsstrasse 16, CH-8092 Zürich, Switzerland

## Abstract

Two commonly used methods of interpolating lake water column profiles—two-point linear interpolation and cubic spline interpolation—were compared, and their relative performance assessed using "leave-k-out" cross-validation. Artificial "pseudo-gaps" of various sizes were created in measured water column profiles of four representative variables (water temperature, oxygen concentration, total phosphorus concentration, and chloride concentration) from the Lake of Zurich by removing measured data from the profiles. The pseudo-gaps were then filled using each of the two interpolation methods. The performance of each interpolation method was assessed based on the root mean square error, mean bias error, and maximum absolute bias error of the interpolated values in relation to the original measured values. The performance of the interpolation methods varied with depth, season, and profile shape. When the profiles were homogeneous both methods performed well, but when the profiles were heterogeneous, linear interpolation generally performed better than cubic spline interpolation. Although the data generated by cubic spline interpolation were less biased than those generated by linear interpolation, there were more instances of extreme errors. The results of this study suggest that linear interpolation is generally preferable to cubic spline interpolation for filling data gaps in measured lake water column profiles.

Studies of long-term changes in lake ecosystems often require the analysis of historical water column data that were originally measured at inconsistent sampling depths. To facilitate the comparison of such data, some form of interpolation is usually employed to yield estimates of the data at standard depths (e.g., Livingstone 2003; Coats et al. 2006; Rempfer et al. 2010) before going on to produce time-series based on these data. Because of their simplicity and ease of use, two-point linear interpolation (between the two measured values on either side of a gap) and cubic spline interpolation (which incorporates information from several measurements on each side of a gap) are both commonly used for this purpose. However, despite the importance of interpolation accuracy when standardizing profile sampling depths, very little published information is available on this topic. In the interest of consistency between studies, it would be advantageous to formalize the selection of an interpolation method for lake water column profiles.

In other contexts, comparisons have been conducted of various interpolation methods. Most of these comparisons have focused on temporal interpolation (Amritkar and Kumar 1995; Baltazar and Claridge 2002; Claridge and Chen 2006) or spatial interpolation in two dimensions (Holdaway 1996; Eischeid et al. 2000; Skaugen and Andersen 2010). These studies applied a multitude of different interpolation methods, and found that each one performed differently depending on the data set being considered. There is therefore no "one-size-fits-all" interpolation method that is best for all data sets. However, the simplest form of interpolation often proves to be the optimal choice (e.g., Chen and Claridge 2000; Baltazar and Claridge 2002; Claridge and Chen 2006). With this in mind,

*Corresponding author: E-mail: Ryan.North@eawag.ch

***Acknowledgments***

We would like to thank all who contributed to the collection of the lake data used in this study. The data were kindly provided by Wasserversorgung Zürich, Switzerland; Amt für Abfall, Wasser, Energie und Luft, Canton of Zurich, Switzerland; Amt für Umweltschutz, Canton of Zug, Switzerland; Istituto Scienze della Terra, Scuola Universitaria della Svizzera Italiana, Switzerland; and Fondazione Mach – Istituto Agrario di S. Michele all'Adige, Italy. We also thank the entire Environmental Isotopes Group at Eawag for discussions that improved the manuscript greatly. This research forms part of project "HYPOX," funded under the European Commission's Seventh Framework Programme (grant no. 226213).

DOI 10.4319/lom.2013.11.213

this study is confined to the two comparatively simple interpolation methods mentioned above—two-point linear interpolation and cubic spline interpolation—for regularizing spatial sampling intervals and filling data gaps in lake profiles. The study aims to obtain greater assurance of compatibility, quality, and reliability when analyzing lake profiles with inconsistent sampling intervals or missing data.

## Materials and procedures

### Data

The study compared the two interpolation methods by creating artificial data gaps, referred to henceforth as "pseudo-gaps," in measured water column profiles. The pseudo-gaps were then filled using each of the two interpolation methods. The relative ability of each interpolation method to reconstruct missing profile data was assessed based on the accuracy with which it was able to fill the pseudo-gaps.

This procedure requires many lake profiles that have been measured at standard sampling depths with no data gaps. Multi-annual data sets from five lakes in Switzerland (Lake of Zurich, Greifensee, Lake of Lugano, Aegerisee, Lake of Walenstadt) and one in northern Italy (Lake of Garda) were examined to find the longest, most complete data set in which the profiles were sampled consistently at the same standard depths. The most suitable data set was found to be that from the Lake of Zurich from 1976 to 2010. During this period, the lake was sampled at approximately monthly intervals a total of 420 times at its deepest point. Sampling depths were consistent at 0.3, 1, 2.5, 5, 7.5, 10, 12.5, 15, 20, 30, 40, 60, 80, 90, 100, 110, 120, 130, and 135 m. All profiles with data missing from any of these 19 standard depths were excluded from the study from the outset. However, there were only few such profiles. For total phosphorus concentration, for example, 19 profiles had to be excluded based on this criterion; for other variables used in this study, even fewer profiles needed to be excluded. Further details on the Lake of Zurich data set are given by Zimmermann et al. (1991), Livingstone (2003), and Jankowski et al. (2006).

To assess the representativeness of the Lake of Zurich profiles, we qualitatively compared the profiles of physical (e.g., temperature), chemical (e.g., oxygen), and biological (e.g., chlorophyll *a*) variables from the Lake of Zurich (1976-2010) with corresponding profiles from the other five lakes listed above. The profiles from all six lakes were found to be similar in shape and seasonal variation, and we therefore concluded that the Lake of Zurich profiles were broadly representative of lakes in the European Alpine region. However, as this comparison only considered lakes from this particular region, caution should be exercised when applying the results of this article beyond temperate climate lakes.

To provide a range of profile shapes, four representative lake variables were included in the analysis: water temperature, oxygen concentration ([$O_2$]), total phosphorus concentration (TP), and chloride concentration ([$Cl^-$]). Measured profiles of all four variables in the Lake of Zurich during a typical year (1993) are shown in Fig. 1. These particular lake variables were selected as together they cover the majority of lake profile types: e.g., orthograde (temperature), positive clinograde

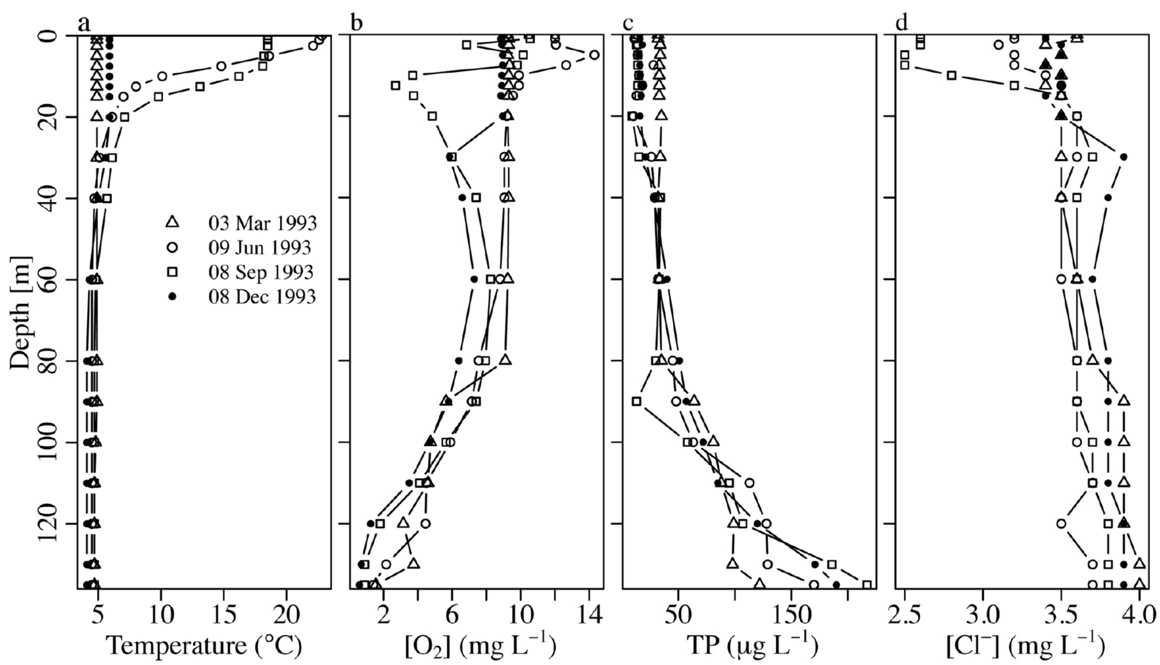


**Fig. 1.** Water column profiles of (a) temperature, (b) oxygen concentration ([$O_2$]), (c) total phosphorus concentration (TP), and (d) chloride concentration ([$Cl^-$]), measured in the Lake of Zurich in 1993.

(temperature, $[O_2]$), negative clinograde (TP, $[Cl^-]$), positive heterograde ($[O_2]$), negative heterograde ($[Cl^-]$), and positive-negative heterograde ($[O_2]$). It should be stressed here that this analysis is not concerned with explaining the behavior of these variables themselves, but with the methodology necessary to cope with interpolating the range of profile shapes represented by these variables at different times of the year.

Based on their shapes, the profiles can be roughly divided into three depth zones. Within the first depth zone, comprising the epilimnion and metalimnion (0–20 m), profile shapes vary the most, both spatially and temporally. In this depth zone in summer and fall, temperatures are constant within the mixed surface layer but decrease uniformly through the stratified thermocline. $[O_2]$ and $[Cl^-]$ show a great degree of variability with depth, in contrast to TP, which remains approximately constant. In the first depth zone in winter and spring, profiles of all four variables vary little with depth. Temperature, $[O_2]$, and $[Cl^-]$ show a marked seasonal variability in profile shape, but seasonal variability in TP is much less pronounced. In the second depth zone, the upper hypolimnion (20–80 m), all variables have approximately uniform profiles regardless of season. In the third depth zone, the lower hypolimnion (80–136 m), temperature and $[Cl^-]$ profiles are approximately constant with depth, whereas $[O_2]$ decreases with depth and TP increases with depth.

**Procedure**

In the following, the term "spatial sampling interval" refers to the vertical distance between adjacent measurements in the same sampling profile. The term "spatial data gap" refers to one or more consecutively missing measurements from the standard set of depths, and the "spatial data gap size" is the number of missing measurements. Note that because spatial sampling intervals are not uniform, the length in meters of a spatial data gap of a given size is also not uniform. The core functions of the R language and environment were used for all aspects of the analysis (R Development Core Team 2011).

All available temperature data from all six lakes mentioned above were analyzed to determine typical spatial data gap sizes. We found that less than 3% of spatial data gaps contained more than three consecutively missed measurements (Fig. 2). Interpolation is thus most often needed to fill spatial gaps of between one and three measurements. This range was therefore chosen as the range of spatial pseudo-gap sizes to be used in the comparison.

The comparison of spatial interpolation methods used a leave-k-out cross-validation, where $k = 1,2,3$ is the pseudo-gap size (method derived from Baltazar and Claridge 2002; see Arlot and Celisse 2010 for a full review of cross-validation). Given a complete profile $y_i$ ($i = 1 \ldots n$) of $n$ observations (where $n = 19$ in the specific case of the Lake of Zurich data set), the procedure for creating and filling 1-point pseudo-gaps was as follows. First, observation $y_2$ was removed to create a new profile with one 1-point pseudo-gap between observations $y_1$ and $y_3$. This gap was then filled by interpolating across it (either by two-point linear interpolation or by cubic spline interpolation), creating a new value, $x_2$. The interpolation error for this pseudo-gap was calculated as $e_2 = x_2 - y_2$. Starting from the original profile each time, and going down the profile from $j = 2$ to $j = n - 1$, the process was conducted $n - 2$ times, yielding $n - 2$ new profiles, each with one interpolated pseudo-gap at a unique depth. The first and last points of the measured profile, $y_1$ and $y_n$, were always retained to avoid endpoint effects. For a pseudo-gap size of $k = 2$, the same procedure was used, except that observations $y_j$ and $y_{j+1}$ were removed each time for $j = 2 \ldots n - 2$. For a pseudo-gap size of $k = 3$, observations $y_j$, $y_{j+1}$, and $y_{j+2}$ were removed each time for $j = 2 \ldots n - 3$. An interpolation error was calculated for every point within each interpolated pseudo-gap; e.g., for every pseudo-gap of size 2, interpolation errors $e_j = x_j - y_j$ and $e_{j+1} = x_{j+1} - y_{j+1}$ were calculated.

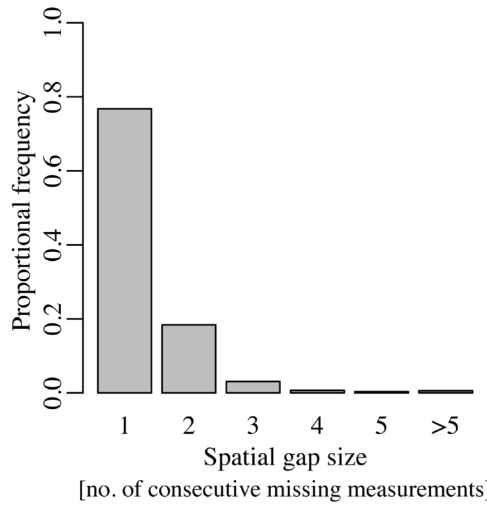


**Fig. 2.** Proportional frequency of occurrence of spatial data gaps of a given size (expressed as the number of consecutive missing measurements) within temperature profiles normally measured at a set of standard depths from which at least one measurement was missing. The plot is based on profiles from the Lake of Zurich (1972–2010; 470 profiles), Greifensee (1942–2002; 609 profiles), Lake of Lugano (1987–2008; 416 profiles), Aegerisee (1950–1994; 278 profiles), Lake of Walenstadt (1972–2010; 391 profiles), and Lake of Garda (1991–2008; 219 profiles). Of the 2383 profiles, 354 contained at least one missing measurement, and there were a total of 810 spatial data gaps of varying size.

Linear interpolation across a pseudo-gap was conducted solely based on the two measured values bounding the pseudo-gap—i.e., on the two values $y_{j-1}$ and $y_{j+1}$ adjacent to the removed value $y_j$—and is therefore extremely local. Because of the nature of the spline function, cubic spline interpolation yields a value that is influenced not only by the measured values directly bounding the pseudo-gap, but also by several measured values on either side. However, because the effect of a constraint on a spline diminishes rapidly with increasing distance (Emery 2001), spline-interpolated values are still influenced predominantly by values measured in the vicinity of the pseudo-gap.

Before analysis, the calculated interpolation errors for each variable were sorted in three different ways: (i) by pseudo-gap size only; (ii) by pseudo-gap size and then by the month in which the measurement was made; and (iii) by pseudo-gap size, by month, and then by measurement depth. The three resulting groups of interpolation errors provided a varied assessment of the performance of the interpolation methods as well as insight into how sampling practices (which affect, for example, the density of measurements as a function of depth) influence interpolation accuracy.

### Assessment

The statistical measures chosen to assess quantitatively the accuracy of each interpolation method were the root mean square error (RMSE), the mean bias error (MBE), and the maximum absolute bias error (MABE). The RMSE, defined as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum e_j^2}{n_e}}, \tag{1}$$

where $n_e$ is the total number of interpolated data points within a given range of depths and profiles, gives an overall measure of the accuracy of the interpolation within this range. The MBE is defined as:

$$\mathrm{MBE} = \frac{\sum e_j}{n_e}. \tag{2}$$

A negative MBE represents a consistent underestimate of the observed value by the interpolated value and a positive MBE represents a consistent overestimate. Both RMSE and MBE are well-accepted and commonly used statistical measures that have been employed for the same purpose in similar studies (e.g., Amritkar and Kumar 1995; Baltazar and Claridge 2002; Neilsen et al. 2010). The MABE, defined as:

$$\mathrm{MABE} = \max\left(\left|e_j\right|\right), \tag{3}$$

was included additionally to provide insight into the magnitude of error each interpolation method could potentially create. Since both RMSE and MBE are averages, they do not provide an indication of the range of error associated with each of the interpolation methods.

To assess the significance of the results, the RMSE, MBE, and MABE were compared with the measurement uncertainty (MU), on the assumption that any error less than the MU is not significant. Not all MUs were known, nor were they necessarily consistent throughout all measurement years. Therefore, as a conservative approach, the largest known MU for a given variable was chosen for comparison purposes (Table 1). The relative performance of the two interpolation methods was assessed in terms of the difference of the RMSE (or MBE, or MABE) associated with two-point linear interpolation and the RMSE (or MBE, or MABE) associated with cubic spline interpolation. This difference will be referred to henceforth as the "RMSE difference" (or "MBE difference," or "MABE difference").

**Pseudo-gap size**

Interpolation errors were grouped by variable and by pseudo-gap size to calculate RMSE values (Table 1). For temperature, $[O_2]$, and TP, the RMSE exceeded the corresponding MU for all pseudo-gap sizes, but for $[Cl^-]$ the RMSE exceeded the MU only for the cubic spline interpolation of a 3-point pseudo-gap. For all variables and pseudo-gap sizes, the RMSE associated with two-point linear interpolation was always less than or

**Table 1.** The root mean square errors (RMSEs) and maximum absolute bias errors (MABEs) associated with the linear interpolation and cubic spline interpolation of lake profiles over pseudo-gaps of various sizes, and the corresponding differences (cubic spline minus linear). The variables are temperature (T), oxygen concentration ($[O_2]$), total phosphorus concentration (TP), and chloride concentration ($[Cl^-]$). Pseudo-gap sizes vary from 1 to 3, where the size of a pseudo-gap is the number of missing data points within the gap. For comparison purposes, the measurement uncertainty (MU) for each variable and pseudo-gap size is also listed. The profiles employed were measured in the Lake of Zurich during 1976-2010.

| | | Linear | | Cubic spline | | Difference | |
|---|---|---|---|---|---|---|---|
| **Variable [units] (± MU)** | **Pseudo-gap size** | **RMSE** | **MABE** | **RMSE** | **MABE** | **RMSE** | **MABE** |
| T [°C] | 1 | 0.5 | 5.3 | 0.5 | 7.5 | 0.0 | 2.2 |
| (± 0.1°C) | 2 | 0.8 | 7.0 | 0.8 | 11.8 | 0.1 | 4.8 |
| | 3 | 1.1 | 8.7 | 1.4 | 27.8 | 0.3 | 19.2 |
| $[O_2]$ [mg L$^{-1}$] | 1 | 0.8 | 7.1 | 1.0 | 9.9 | 0.2 | 2.8 |
| (± 0.3 mg L$^{-1}$) | 2 | 1.0 | 8.2 | 1.5 | 14.4 | 0.5 | 6.2 |
| | 3 | 1.2 | 8.7 | 2.2 | 32.8 | 1.0 | 24.1 |
| TP [µg L$^{-1}$] | 1 | 8.3 | 131.0 | 9.6 | 123.9 | 1.3 | –7.1 |
| (± 1.8 µg L$^{-1}$) | 2 | 9.7 | 152.4 | 11.8 | 149.6 | 2.1 | –2.8 |
| | 3 | 11.6 | 165.1 | 15.4 | 238.9 | 3.9 | 73.7 |
| $[Cl^-]$ [mg L$^{-1}$] | 1 | 0.1 | 2.3 | 0.2 | 2.3 | 0.0 | 0.1 |
| (± 0.27 mg L$^{-1}$) | 2 | 0.2 | 2.2 | 0.2 | 2.6 | 0.1 | 0.4 |
| | 3 | 0.2 | 2.2 | 0.3 | 4.2 | 0.2 | 2.0 |

equal to that associated with cubic spline interpolation. Using a pseudo-gap size of 3 as an example, temperature RMSEs were 1.1°C for linear interpolation and 1.4°C for cubic spline interpolation; $[O_2]$ RMSEs were 1.2 mg $L^{-1}$ for linear interpolation and 2.2 mg $L^{-1}$ for cubic spline interpolation; TP RMSEs were 11.6 µg $L^{-1}$ for linear interpolation and 15.4 µg $L^{-1}$ for cubic spline interpolation; and $[Cl^-]$ RMSEs were 0.2 mg $L^{-1}$ for linear interpolation and 0.3 mg $L^{-1}$ for cubic spline interpolation.

For all four variables, the difference of the RMSE associated with linear interpolation and the RMSE associated with cubic spline interpolation increased rapidly with pseudo-gap size (Table 1). For a pseudo-gap size of 1, there was little difference between the two interpolation methods, but as the pseudo-gap size grew larger, cubic spline interpolation became increasingly less accurate than linear interpolation. For $[Cl^-]$, the RMSE differences were however very small at all pseudo-gap sizes, implying that for this variable the choice of interpolation method is not critical.

The MABEs associated with linear interpolation were consistently less than those associated with cubic spline interpolation, except for TP (pseudo-gap size of 1 and 2). The MABEs tended to increase with pseudo-gap size, as did the MABE differences, which were generally much larger for a pseudo-gap size of 3 than for a pseudo-gap size of 1 or 2. This means that the increase in MABE with increasing pseudo-gap size was less in the case of linear interpolation than in the case of cubic spline interpolation. For all variables and pseudo-gap sizes, the MABEs exceeded the corresponding MUs by at least an order of magnitude.

**Season and month**

When grouped by pseudo-gap size and month, the RMSE values for both methods showed a strong seasonal cycle in all cases, with values being lowest in winter and spring, increasing during summer, and peaking in fall (Fig. 3). For temperature, the RMSEs associated with both methods exceeded the MU from April through December. For $[O_2]$ and TP the same was true throughout the entire year. For $[Cl^-]$ the RMSEs associated with linear interpolation were below the MU for 1-point and 2-point pseudo-gaps, and only exceeded the MU during August and September for a 3-point pseudo-gap. The RMSEs associated with the cubic spline interpolation of $[Cl^-]$ exceeded the MU from August through November for a 2-point pseudo-gap, and from June through November for a 3-point pseudo-gap.

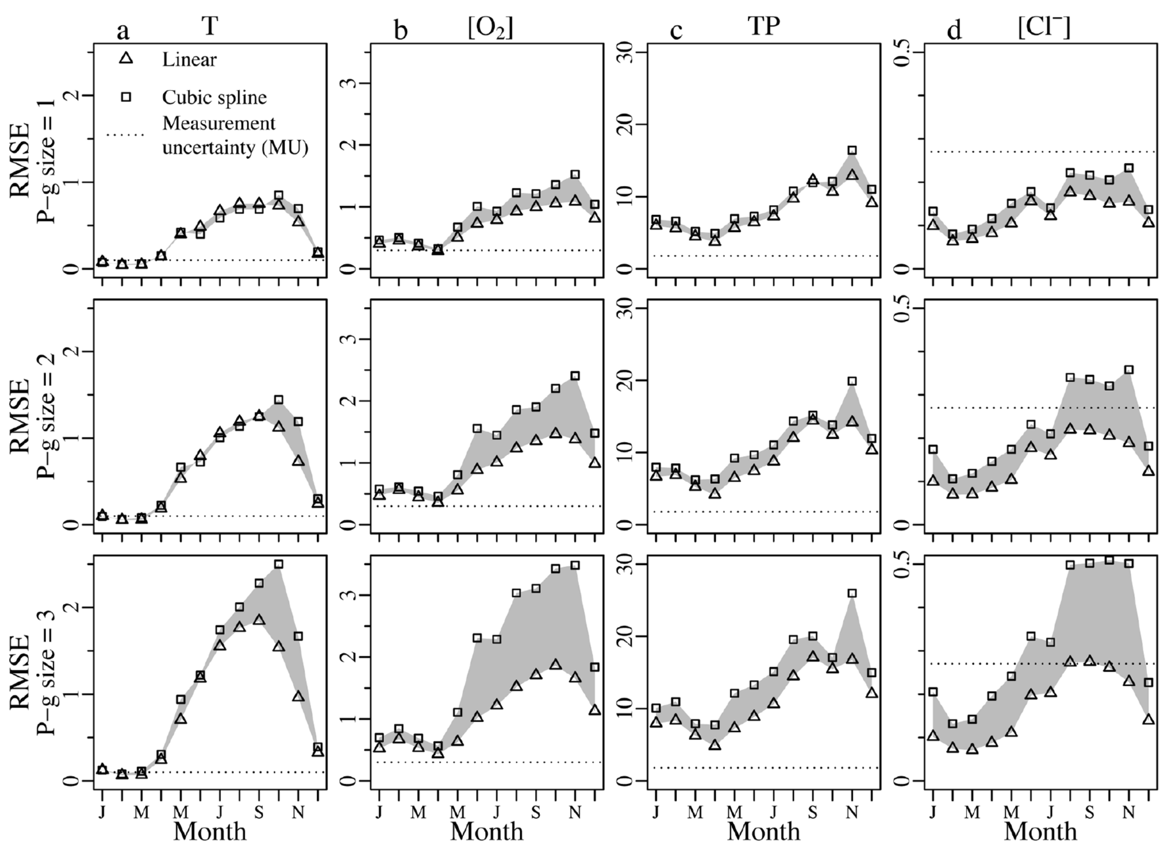


**Fig. 3.** Seasonal variation in the root mean square error (RMSE) associated with linear interpolation and with cubic spline interpolation, for pseudo-gap sizes of 1, 2, and 3. The variables are (a) temperature (T, °C), (b) oxygen concentration ($[O_2]$, mg $L^{-1}$), (c) total phosphorus concentration (TP, µg $L^{-1}$), and (d) chloride concentration ($[Cl^-]$, mg $L^{-1}$). Dotted lines indicate measurement uncertainty (MU). The shaded areas indicate the difference of the RMSEs associated with each of the two methods.

In general, in all months the RMSEs for both methods increased with increasing pseudo-gap size, but as the rate of increase was higher for cubic-spline interpolation than for linear interpolation, the RMSE differences (Fig. 3, shaded areas), and therefore the improvement of linear interpolation over cubic spline interpolation, also increased with pseudo-gap size. For a pseudo-gap size of 1, the RMSE differences for all four variables were small throughout most of the year. As the pseudo-gap size was increased to 2 and 3, the RMSE differences became progressively larger, particularly from approximately July to November. There was a strong seasonality in the RMSE differences (Fig. 3, shaded areas), which were smallest in winter and spring but underwent an increase in summer to reach a maximum in fall.

**Measurement depth**

The influence of measurement depth on interpolation accuracy was investigated by successively grouping the interpolation errors by variable, pseudo-gap size, depth, and month, before calculating the RMSEs. As general patterns were similar for all pseudo-gap sizes, only the results for a pseudo-gap size of 2 are shown (Fig. 4). For all three pseudo-gap sizes, the previously discussed seasonality in the RMSE (i.e., minimum in winter and spring, maximum in fall) and dependence on pseudo-gap size (i.e., RMSE increasing with pseudo-gap size) was once again evident. Additionally, the results showed a concentration of high RMSEs in the metalimnion and upper hypolimnion, centered around a depth of 20 m (Fig. 4). The RMSEs of TP were also high near the lake bottom, while for the other three variables the RMSEs were low below approximately 80 m, remaining near or below the respective MUs for all three pseudo-gap sizes. The RMSEs for [$Cl^-$] exceeded the MU from August to November for depths above 20 m for linear interpolation and above 40 m for cubic spline interpolation. This depth limitation of [$Cl^-$] was present for all three pseudo-gap sizes; however, the frequency of occurrence of cases in which the RMSE exceeded the MU tended to increase with pseudo-gap size.

For the majority of depths and months, the RMSEs associated with linear interpolation were lower than those associated with cubic spline interpolation (Fig. 4). For all three pseudo-gap sizes, the RMSE differences were largest at depths above 80 m for temperature, [$O_2$], and [$Cl^-$], and below 100 m for TP (Fig. 4, shaded areas). Between June and November, the RMSE associated with linear interpolation occasionally exceeded that associated with cubic spline interpolation for temperature in the lower epilimnion and the metalimnion, and for TP at the lake bottom. However, the number of such occurrences was comparatively low.

**Mean bias error**

The effectiveness of the two interpolation methods was also assessed for all three pseudo-gap sizes using the MBE and the same grouping method used for the RMSE (i.e., pseudo-gap size, month, and depth). For the most part, the MBEs remained below or near the respective MUs. Therefore, only the MBEs for a pseudo-gap size of 2, sorted by depth and month, are presented as an example (Fig. 5). In the cases when the MBE did exceed the MU, the magnitudes of the MBE associated with cubic spline interpolation were smaller than those associated with linear interpolation. The MBEs for temperature, [$O_2$], and [$Cl^-$] showed similar seasonal and spatial patterns to those exhibited by the RMSEs; i.e., the MBEs were low in winter and spring and high in summer and fall, with the largest bias occurring in the metalimnion and upper hypolimnion (Fig. 5). The MBE for TP followed the same seasonal pattern but the highest values occurred below 100 m. The MBE differences for temperature, [$O_2$], and [$Cl^-$] generally increased with increasing pseudo-gap size, as the magnitude of the MBEs of both methods increased proportionally. Although the MBE difference for TP also increased with pseudo-gap size, this was mainly accounted for by an increase in the MBE associated with linear interpolation, as the MBE associated with cubic spline interpolation changed very little with pseudo-gap size.

## *Discussion*

Two-point linear interpolation and cubic spline interpolation are both simple to apply, but from the outset, each can be seen to have both advantages and disadvantages. For instance, any minimum or maximum that lies in a gap between two measured values clearly cannot be satisfactorily represented using linear interpolation. Linear interpolation will always overestimate the minima in profiles and underestimate the maxima, resulting in reduced variability within the profile and a bias toward the mean. The values produced by cubic spline interpolation do not suffer from this bias and may yield more realistic values for the extrema within a profile, but this comes at a cost: because spline-interpolated data can lie outside the range of measured data, unrealistically low minima and unrealistically high maxima may result. This tends to happen when adjacent measurements are spatially close but differ substantially in magnitude from one another, which is most likely to result when the sampling interval is small but measurement error is large. The likelihood of this kind of overshoot makes it necessary to check spline-interpolated data to ensure that they are physically reasonable. This can be done either by eye or by adding an additional automatic stage to the interpolation process to flag values that are physically impossible (e.g., negative concentrations) or unlikely (e.g., extremely high concentrations or extremely high concentration gradients within the interpolated profile, or interpolated concentrations that differ excessively from the measured values above and below them). If profile interpolation is to be completely automated, with no checking by eye, then linear interpolation is safer because unpleasant surprises are much less likely. However, the interpolated profile resulting from a spline interpolation is by definition continuous with respect to its first and second derivatives, which is not the case for a linearly interpolated profile. Thus the gradients of linearly interpolated profiles are discontinuous, making flux calculations unreliable.

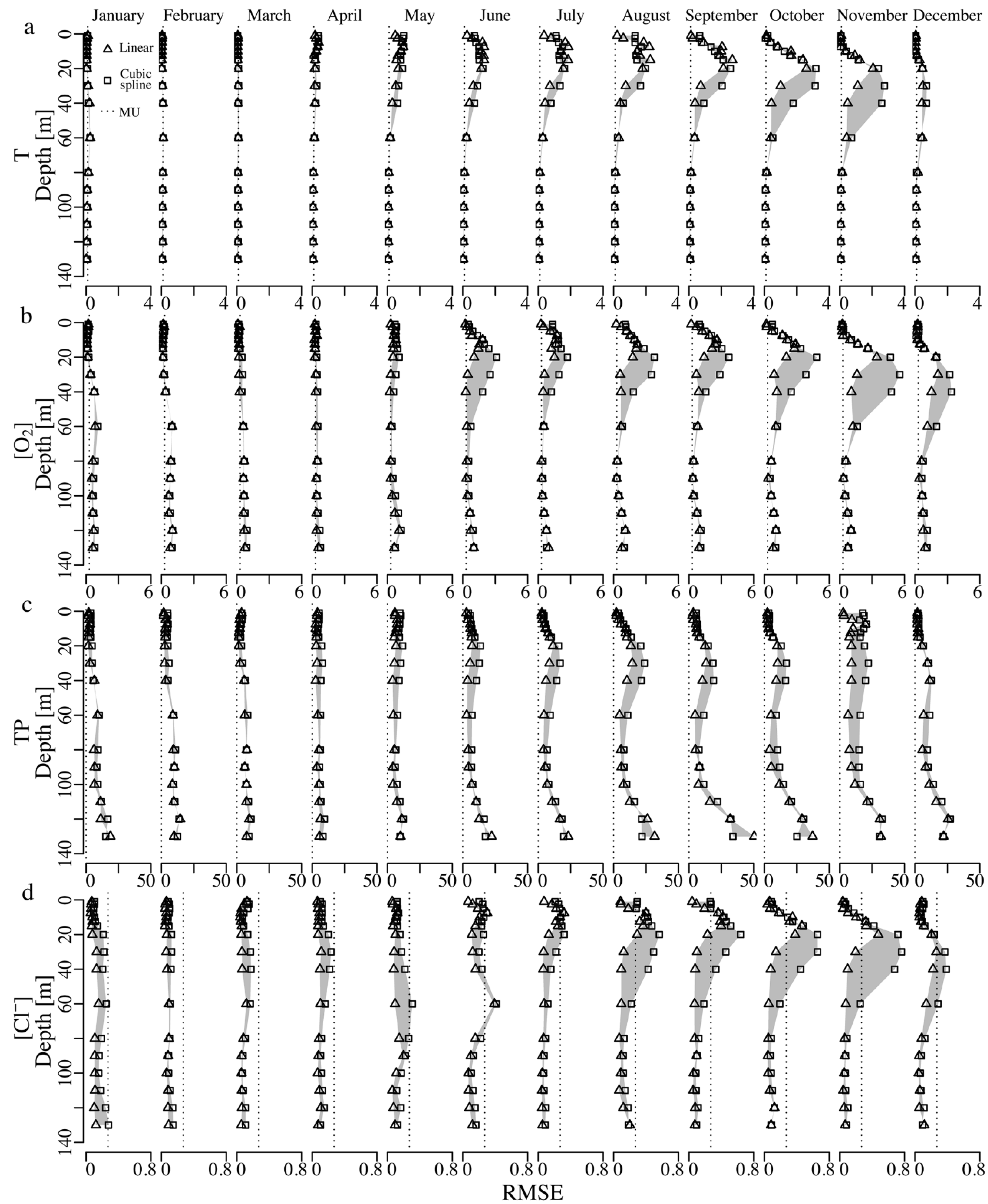


**Fig. 4.** Monthly profiles of the root mean square errors (RMSEs) associated with linear interpolation and cubic spline interpolation. Calculated interpolation errors for a pseudo-gap size of 2 were grouped by measurement depth and month. The variables are (a) temperature (T, °C), (b) oxygen concentration ([$O_2$], mg $L^{-1}$), (c) total phosphorus concentration (TP, µg $L^{-1}$), and (d) chloride concentration ([$Cl^-$], mg $L^{-1}$). Dotted lines indicate measurement uncertainty. The shaded areas indicate the difference of the RMSEs associated with each of the two methods.

This study has shown the errors associated with two-point linear interpolation to be consistently smaller than those associated with cubic spline interpolation. Additionally, the analysis provides useful insight into how interpolation is affected by season, depth, and spatial data gap size. For all four variables investigated, the two methods performed equally well throughout the entire profile in winter and spring, when the RMSEs and MBEs associated with each of the two methods did not generally exceed the MU (Figs. 4, 5). In summer and fall the same was true below approximately 60 m for water tem-

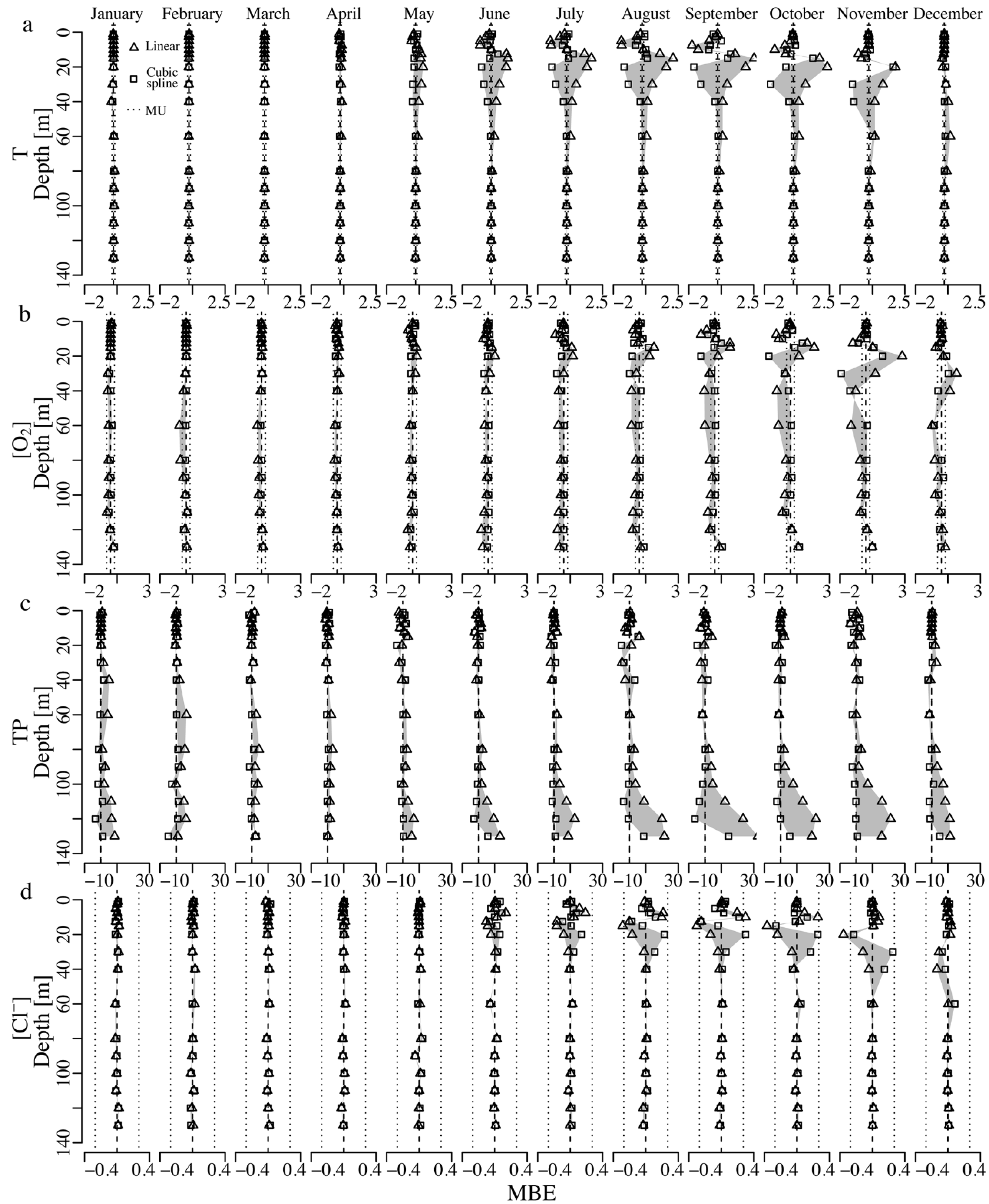


**Fig. 5.** Monthly profiles of the mean bias error (MBE) for linear interpolation and cubic spline interpolation. Calculated interpolation errors for a pseudo-gap size of 2 were grouped by measurement depth and month. The variables are (a) temperature (T, °C), (b) oxygen concentration ($[O_2]$, mg $L^{-1}$), (c) total phosphorus concentration (TP, µg $L^{-1}$), and (d) chloride concentration ($[Cl^-]$, mg $L^{-1}$). Dotted lines indicate measurement uncertainty. The shaded areas indicate the difference of the MBEs associated with each of the two methods.

perature, $[O_2]$, and $[Cl^-]$ (but not for TP). Where there was a measurable difference in performance, linear interpolation produced consistently smaller errors than cubic spline interpolation. Although the RMSE increased with pseudo-gap size for both methods, the errors associated with cubic spline interpolation increased at a faster rate than those associated with linear interpolation, and as a result the RMSE difference also increased with increasing pseudo-gap size. In the rare situations in which the RMSE associated with cubic spline interpolation was smaller than that associated with linear interpolation, the difference between the RMSEs was small and did not increase with pseudo-gap size (e.g., temperature in Fig. 4).

Despite the fact that the RMSE values associated with linear interpolation were consistently lower than those associated with cubic spline interpolation (Fig. 4), the bias associated with the latter was smaller (Fig. 5). Cubic spline interpolation both overestimated and underestimated the true values over a wide range of error magnitude, whereas linear interpolation consistently produced overestimates or underestimates (depending on the profile shape), but within a smaller range of error magnitude. The difference between the methods can be illustrated using density distribution plots of the errors (Fig. 6). The distributions of the errors associated with cubic spline interpolation have a broad, shallow peak near zero, with long tails, whereas those associated with linear interpolation are either slightly negative or slightly positive, with a high, narrow peak, and short tails.

The reason for the difference in the density distributions is made clear by the example shown in Fig. 7. In Fig. 7a, observed water temperatures at 5 m, 7.5 m, and 10 m have been removed to create a 3-point pseudo-gap, which is then filled using both two-point linear interpolation and cubic spline interpolation. The interpolation errors (listed in Fig. 7a) associated with the cubic spline interpolation are slightly smaller than those associated with the linear interpolation (see also Fig. 4). In Fig. 7b, a 3-point pseudo-gap created by removing the measurements at 20 m, 30 m, and 40 m, and interpolated similarly, results in the opposite situation: the interpolation error associated with the two-point linear interpolation is much smaller than that associated with the cubic spline interpolation. There are two key differences between the two scenarios. The first is in the magnitude of the errors: while cubic spline interpolation errors can be small (Fig. 7a) they can also be large (Fig. 7b), whereas linear interpolation errors are consistently small. The second is in physical plausibility: under normal conditions, the spline-interpolated profile of Fig. 7b is not physically realistic (assuming the temperature of maximum density to be 4°C, then between approximately 20 m and 30 m denser water would be overlying lighter water, and between about 30 m and 40 m water temperatures would be < 0°C). The large interpolation errors depicted in Fig. 7b occur only when data points are missing in regions where gradients are changing; i.e., where $d^2C/dz^2 \neq 0$, where C is concentration or temperature and z is depth. In regions where $d^2C/dz^2$ is zero or close to zero, less information is lost when data points are missing, making it easier for either interpolation method to accurately fill the data gap.

The density distribution of errors associated with the linear interpolation of temperature ($T$) at 20 m depth, where $d^2T/dz^2$ can be large during certain times of the year, differs clearly from that associated with the linear interpolation of the other three variables at the same depth (Fig. 6). Instead of a high, narrow peak, the error distribution for temperature has a small peak near zero and a large shoulder up to approximately 2°C. At 20 m depth, the temperature profile is typically either uniform (e.g., during winter) or is transitioning

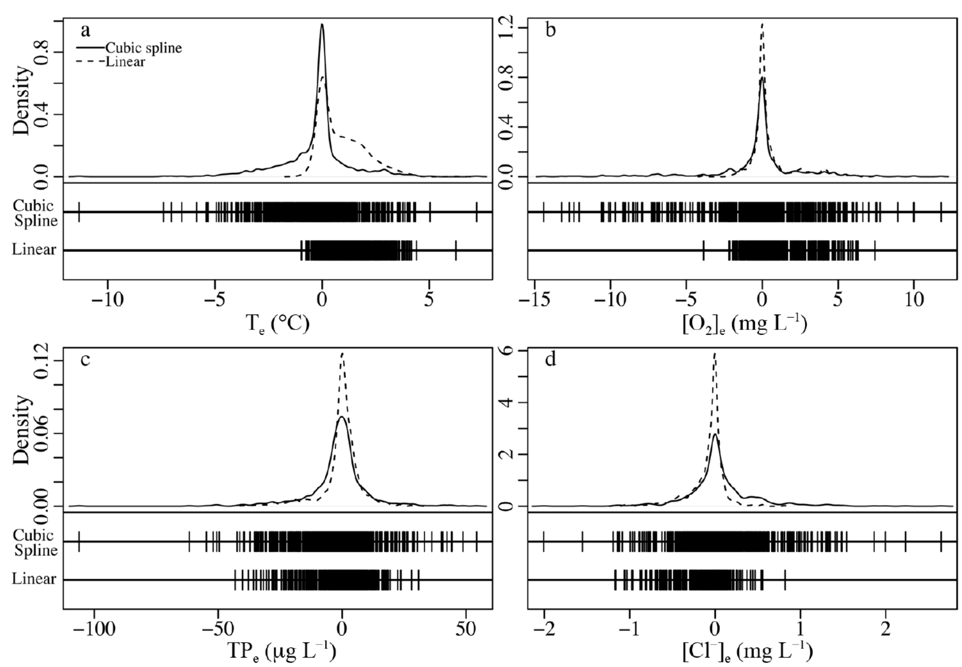


**Fig. 6.** Density distribution plots of interpolation errors (interpolated minus observed) for (a) temperature ($T_e$), (b) oxygen concentration ($[O_2]_e$), (c) total phosphorus concentration ($TP_e$), and (d) chloride concentration ($[Cl^-]_e$) at 20 m depth for all three pseudo-gap sizes (1, 2, and 3) combined. Two types of distribution plot are shown for each variable to highlight the width and height of peaks, as well as tail sizes.

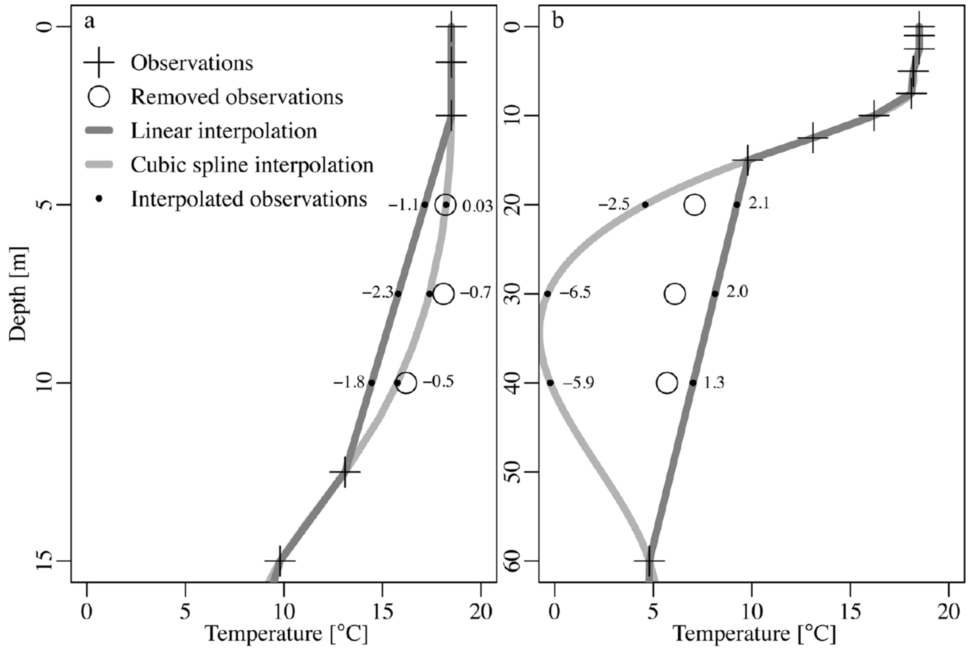


**Fig. 7.** Two examples of the linear interpolation and cubic spline interpolation of part of a temperature profile (measured in the Lake of Zurich on 08 Sep 1993) for a pseudo-gap size of 3, showing (a) a situation in which cubic spline interpolation is the more accurate (interpolation over the pseudo-gap from 2.5 to 12.5 m depth) and (b) a situation in which two-point linear interpolation is the more accurate (interpolation over the pseudo-gap from 15 to 60 m depth). The magnitudes of the interpolation errors for both methods are marked on the figure. The profiles of observed temperature are plotted at the standard measurement depths, whereas the interpolated profiles are shown with a uniform sampling interval of 1 m to highlight differences between the results of the two interpolation methods.

from a low to a high gradient (e.g., during summer stratification, Fig. 7). In the former case, linear interpolation can accurately interpolate any data gaps and the interpolation errors are small, giving rise to the peak in the distribution near zero. In the latter case, linear interpolation will consistently overestimate the observed values, and interpolation errors will be positive (Fig. 7b). As the pseudo-gap size increases, the magnitudes of the interpolation errors also increase, leading to a concentration of interpolation errors $T_e$ within the approximate range $0°C \leq T_e \leq 2°C$ (Fig. 6a). This can be seen clearly in Fig. 7b, where a 3-point pseudo gap produces a linear interpolation error of $T_e = 2.1°C$ at a depth of 20 m. The contrast in density distribution between cubic spline interpolation and linear interpolation at 20 m depth illustrates very well the trade-offs between the advantages and disadvantages of the two interpolation methods discussed above. During the stratification period, $d^2T/dz^2$ is large, and as a result, linear interpolation consistently overestimates the temperature, but without creating unrealistically high or low values. Cubic spline interpolation is able to accurately interpolate the temperature at 20 m (peak at approximately zero), but does so at the risk of creating unrealistically high or low values, which are reflected in the large tails in the density distribution of $T_e$ (Fig. 6a).

Interpolation errors associated with both linear and cubic spline interpolation show a strong seasonal dependence. Errors are low in winter and spring, increase during summer and are high in fall. The same seasonal pattern also holds true for the RMSE difference and for the MBE difference. The homogeneity of the winter and spring profiles makes it easy for both methods to interpolate all three pseudo-gap sizes accurately. During summer and fall, the existence of gradients in the profiles (Fig. 1) makes interpolation difficult, particularly over large gaps (e.g., Fig. 7). Even TP profiles, which do not show much seasonal variation themselves (Fig. 1), show a seasonal pattern in their RMSE values (Fig. 3c).

Figs. 4 and 5 show that most of the seasonal pattern in RMSE and MBE can be attributed to interpolation errors in and near the metalimnion (and in the case of TP, near the lake bottom). For water temperature, $[O_2]$, and $[Cl^-]$, interpolation errors were small and rarely exceeded MU below approximately 60 m. This pattern of error distribution also held true for the RMSE difference between the two methods. As discussed above, changing gradients in a profile result in large errors during summer and fall. As these gradients are located around the metalimnion and upper hypolimnion (and near the lake bottom for TP), this is where the largest interpolation errors occur. Below approximately 60 m, profiles are generally homogeneous (except for TP), and are therefore easier to interpolate accurately. The depth dependence of the magnitude of the interpolation errors can be linked to the physics of an individual lake. A lake that does not mix regularly will develop varying thermal and chemical gradients, giving rise to large interpolation errors. In contrast, a lake subjected to consistently strong mixing events will have weaker thermal or chemical stratifications, minimizing changes in gradients and therefore interpolation errors. Existing information on typical lake profiles and mixing patterns can therefore give a preliminary indication of how successful empirical interpolation is likely to be in filling any data gaps.

This discussion has focused on a comparison of the size of the errors produced by the two interpolation methods. However, the size of the errors should also be considered in the context of the usefulness of the interpolated values. The MABE values in Table 1 show that for all pseudo-gap sizes and variables, both interpolation methods have the potential to produce large errors. Although the magnitude of the MABEs is disconcerting, density distribution plots of the errors (Fig. 6) show that extremely large MABEs are in fact rare, and that most interpolation errors (concentrated in the vicinity of the peaks in Fig. 6) are small. Nevertheless, because both interpolation methods have the potential to produce large errors in a number of situations, further steps (e.g., visual inspection of the interpolated profiles) may be necessary to identify unacceptably large errors.

## Comments and recommendations

The two methods compared in this study, two-point linear interpolation and cubic spline interpolation, are two of the simplest gap-filling methods available. Yet the results of this study show that both linear and cubic spline interpolation can provide reasonably accurate results when interpolating gaps of various sizes in lake profiles. For both methods, interpolation errors were smallest when the pseudo-gap size was small, and increased as the pseudo-gap size grew. However, the rate of increase with pseudo-gap size was lower for linear interpolation than for cubic spline interpolation, and as a result, linear interpolation was substantially more accurate than cubic spline interpolation when interpolating over large pseudo-gaps. Two-point linear interpolation is therefore generally recommended for the interpolation of data gaps in lake profiles, as it is likely to yield more accurate and more consistent results than cubic spline interpolation.

However, the limitations discussed above must be taken into consideration. In addition, when either interpolation method is applied, it is important to be aware of the size and location of data gaps, and to verify the results of the interpolation. Interpolation is in a sense a necessary evil, and Figs. 3-6 show that errors associated with any automatic interpolation can be unacceptably high. Automatic profile interpolation, therefore, needs to be supplemented with an automatic method of detecting physically implausible interpolation errors, and ideally, with visual inspection of the interpolated profiles. Furthermore, the results suggest that neither method can satisfactorily interpolate a 3-point pseudo-gap, so that for gap sizes of 3 or more points, a process-based physical model may be the only feasible method of interpolation.

This study has also highlighted the importance of selecting proper sampling intervals. The data set used for the analysis seemed well-designed, with measurement densities high near the lake surface, decreasing through the metalimnion and upper hypolimnion, and increasing again slightly toward the lake bottom. However, the cross-validation analysis found the highest interpolation errors for all variables in the metalimnion and upper hypolimnion, and near the lake bottom for TP. Clearly, smaller spatial sampling intervals at these depths would not only reduce interpolation errors, but also capture gradients and variations with greater accuracy.

## References

Amritkar, R. E., and P. P. Kumar. 1995. Interpolation of missing data using nonlinear and chaotic system analysis. J. Geophys. Res. 100:3149-3154 [doi:10.1029/94JD01531].

Arlot, S., and A. Celisse. 2010. A survey of cross-validation procedures for model selection. Statist. Surv. 4:40-79 [doi:10.1214/09-SS054].

Baltazar, J. -C., and D. E. Claridge. 2006. Study of cubic splines and Fourier series as interpolation techniques for filling in short periods of missing building energy use and weather data. ASME J. Sol. Energy Eng. 128:226-230 [doi:10.1115/1.2189872].

Chen, H., and D. E. Claridge. 2000. Procedures for filling short gaps in energy use and weather data, p. 314-326. *In* 12th symposium on improving building systems in hot and humid climates, San Antonio, May 15-17, 2000. Energy Systems Laboratory.

Claridge, D. E., and H. Chen. 2006. Missing data estimation for 1–6 h gaps in energy use and weather data using different statistical methods. Int. J. Energy Res. 30:1075-1091 [doi:10.1002/er.1207].

Coats, R., J. Perez-Losada, G. Schladow, R. Richards, and C. Goldman. 2006. The warming of Lake Tahoe. Clim. Change. 76:121-148 [doi:10.1007/s10584-005-9006-1].

Eischeid, J. K., P.A. Pasteris, H. F. Diaz, M. S. Plantico, and N. J. Lott. 2000. Creating a serially complete, national daily time series of temperature and precipitation for the western United States. J. Appl. Meteorol. 39:1580-1591 [doi:10.1175/1520-0450(2000)039<1580:CASCND>2.0.CO;2].

Emery, W. J. 2001. Statistical methods and error handling. *In* W. J. Emery and R. E. Thomson [eds.], Data analysis methods in physical oceanography. Elsevier.

Holdaway, M. 1996. Spatial modeling and interpolation of monthly temperature using kriging. Clim. Res. 6:215-225 [doi:10.3354/cr006215].

Jankowski, T., D. M. Livingstone, H. Bührer, R. Forster, and P. Niederhauser. 2006. Consequences of the 2003 European heat wave for lake temperature profiles, thermal stability, and hypolimnetic oxygen depletion: Implications for a warmer world. Limnol. Oceanogr. 51:815-819 [doi:10.4319/lo.2006.51.2.0815].

Livingstone, D. M. 2003. Impact of secular climate change on the thermal structure of a large temperate central European lake. Clim. Change. 57:205-225 [doi:10.1023/A:1022119503144].

Neilsen, D., G. Duke, B. Taylor, J. Byrne, S. Kienzle, and T. V. der Gulik. 2010. Development and verification of daily gridded climate surfaces in the Okanagan basin of British Columbia. Can. Wat. Resour. J. 35:131-154 [doi:10.4296/cwrj3502131].

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Rempfer, J., D. M. Livingstone, C. Blodau, R. Forster, P. Niederhauser, and R. Kipfer. 2010. The effect of the exceptionally mild European winter of 2006-2007 on temperature and oxygen profiles in lakes in Switzerland: A foretaste of the future? Limnol. Oceanogr. 55:2170-2180 [doi:10.4319/lo.2010.55.5.2170].

Skaugen, T., and J. Andersen. 2010. Simulated precipitation fields with variance-consistent interpolation. Hydrol. Sci. J. 55:676-686 [doi:10.1080/02626667.2010.487976].

Zimmermann, U., R. Forster, and H. Sontheimer. 1991. Long-term change in the water quality of Lake Zurich, Upper Lake Zurich, and the Lake of Walenstadt. Tech. Rep., City of Zurich Water Supply. [In German.]

*Submitted 11 October 2012*
*Revised 11 March 2013*
*Accepted 19 March 2013*