

# Calibrating biogeochemical and physical climate proxies from non-varved lake sediments with meteorological data: methods and case studies

Lucien von Gunten · Martin Grosjean ·  
Christian Kamenik · Marian Fújak ·  
Roberto Urrutia

Received: 11 July 2011 / Accepted: 12 February 2012 / Published online: 28 February 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** Lake sediment records are underrepresented in comprehensive, quantitative, high-resolution (sub-decadal), multi-proxy climate reconstructions for the past millennium. This is largely a consequence of the difficulty of calibrating biogeochemical lake sediment proxies to meteorological time series (calibration-in-time). Thanks to recent methodological advances, it is now possible. This paper outlines a step-by-step, specifically tailored methodology, with practical suggestions for calibrating and validating biogeochemical proxies from lake sediments to meteorological data. This approach includes: (1) regional climate data; (2) site selection; (3) coring and core selection; (4) core chronology; (5) data acquisition;

and (6) data analysis and statistical methods. We present three case studies that used non-varved lake sediments from remote areas in the Central Chilean Andes, where little a priori information was available on the local climate and lakes, or their responses to climate variability. These case studies illustrate the potential value and application of a calibration-in-time approach to non-varved lake sediments for developing quantitative, high-resolution climate reconstructions.

**Keywords** Sedimentology · Paleolimnology · Climate change · Paleoclimate · Numerical methods · Calibration in time

**Electronic supplementary material** The online version of this article (doi:10.1007/s10933-012-9582-9) contains supplementary material, which is available to authorized users.

L. von Gunten · M. Grosjean · C. Kamenik  
Oeschger Centre for Climate Change Research,  
University of Bern, Bern, Switzerland  
e-mail: martin.grosjean@oeschger.unibe.ch

C. Kamenik  
e-mail: christian.kamenik@giub.unibe.ch

L. von Gunten (✉) · M. Grosjean · C. Kamenik  
Institute of Geography, University of Bern, Erlachstrasse  
9a T3, 3012 Bern, Switzerland  
e-mail: lucien.vongunten@giub.unibe.ch

L. von Gunten  
Climate System Research Center, University  
of Massachusetts, Amherst, MA, USA

L. von Gunten  
Geological Institute, Swiss Federal Institute  
of Technology, Zürich, Switzerland

M. Fújak  
Surface Waters Research, EAWAG, Dübendorf,  
Switzerland  
e-mail: marian.fujak@eawag.ch

R. Urrutia  
Centro de Ciencias Ambientales EULA-Chile,  
Universidad de Concepción, Concepción, Chile  
e-mail: rurrutia@udec.cl

## Introduction

High-resolution (annual-decadal), quantitative, multi-proxy regional and hemispheric climate reconstructions have become cornerstones in the debate about climate variability, anthropogenic perturbations to the climate system and questions regarding the extent to which paleoclimate information may help constrain projections of future climate at regional scales (Luterbacher et al. 2004; Hegerl et al. 2006; Neukom et al. 2010, 2011). Lake sediment data, however, are very rarely used for such purposes (Mann et al. 2008; Neukom et al. 2011). This is likely a consequence of the methodological difficulties of producing time series from lake sediment proxy data that meet the requirements for comprehensive multi-proxy climate reconstructions: (1) the proxy time series must be highly (ideally annually) resolved, continuous and have regular time intervals; (2) the chronology must be accurate; and (3) the lake-sediment climate proxies must be calibrated to meteorological data and expressed in quantitative terms of a climate variable (e.g. surface air temperature or precipitation) or an index (e.g. Palmer Drought Severity Index, Grosjean et al. 2009). These criteria are generally very difficult to meet using lake sediments, particularly non-varved sediments from remote areas, with low sedimentation rates, where local meteorological data and ecological observations are sparse or missing. Such lakes, however, are common in practice.

Two approaches for proxy-climate calibrations exist (Birks 1998). One is ‘Calibration-in-space,’ which involves developing a modern training dataset that models biological (e.g. diatoms, foraminifera and chironomids) responses to an environmental variable (e.g. temperature), followed by development of a transfer function that can be applied to a sediment core to reconstruct the environmental variable of interest (Imbrie and Kipp 1971; Sachs et al. 1977; Walker et al. 1997; Bigler and Hall 2002). This approach cannot be used for physical and biogeochemical lake sediment proxies because every lake has a unique geomorphological and hydrological setting. The other approach, however, ‘Calibration-in-time’ (CIT), can be used, whereby a time series of proxy variables within a sediment core is calibrated against a time series of environmental (meteorological) observations during a period of common observation (calibration period; Cook et al. 1994).

The CIT approach has at least three fundamental requirements: (1) a highly accurate and precise chronology of the lake sediments during the calibration period; (2) newly established calibration models for every lake; and (3) very long time series of climate data to develop a good calibration model. Data acquisition of biogeochemical proxies is often less time consuming and less expensive compared to biological proxies. Non-destructive scanning techniques provide measurements at high resolution (<2 mm) and acquisition of large data sets is relatively rapid (Zolitschka et al. 2001; Rein and Sirocko 2002; Rothwell 2006). This is an enormous advantage for developing millennia-long, highly resolved climate reconstructions.

The CIT approach has been applied successfully to varved sediments because highly accurate chronologies can be developed and sediment sampling and analyses can be performed at annual resolution (Francus et al. 2002; Kalugin et al. 2007; Trachsel et al. 2010a; Kaufman et al. 2011; Elbert et al. 2012). Unfortunately, the vast majority of lakes do not produce varves. This lack of annual laminations makes the CIT approach more challenging because dating uncertainties during the calibration period, which are typically derived from  $^{210}\text{Pb}$  profiles, are larger. There are, however, some successful examples from the Austrian Alps (Koinig et al. 2002), south-central Alaska (McKay et al. 2008), Central Chile (von Gunten et al. 2009a) and Tasmania (Saunders et al. pers. commun.).

Over the past few years, we have systematically explored new ways and refined methods to calibrate high-resolution biogeochemical and physical proxies in lake sediments to observed climate data. This paper outlines the methodological framework required for calibrating biogeochemical or physical proxies in non-varved sediments to meteorological time series. The method is specifically designed and optimized for producing quantitative, high-resolution (sub-decadal) climate reconstructions. We use three case studies from lakes with organic and lithoclastic sediments to demonstrate the application of this approach. All of the lakes are not ideal for CIT in the sense that they are non-varved and located in remote areas of the Central Chilean Andes and the Coastal Range where local meteorological data are not available. Two lakes have relatively low sedimentation rates (8–12 cm/100 years). Surprisingly, our results show that the

tailor-made methodology worked successfully in all three lakes and that standard biogeochemical proxies could be calibrated to meteorological data. We also include recommendations for evaluating the quality of climate data from reanalysis data sets, establishing highly accurate chronologies, sediment sub-sampling, optimizing standard analytical methods, optimizing statistical methods for calibration and verification, and testing the robustness of the calibration model. This methodology has the potential to be applied in many other lakes.

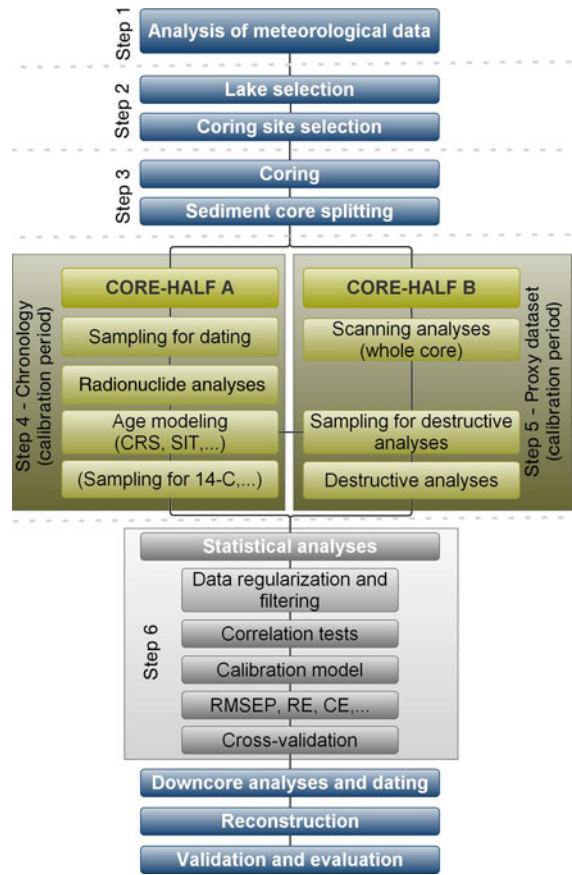
**Materials and methods**

The overall methodology

The overall methodology is designed to establish a calibration-in-time between a time series of biogeochemical lake sediment proxies and a time series of climate data in non-varved lakes, with the objective of using this calibration for quantitative reconstruction of past climate (Fig. 1). The starting point is a lake in a remote area, where little or no a priori knowledge exists with respect to the ability of any organic or inorganic sediment variable to predict a climate variable such as temperature or precipitation, and where there is no long instrumental record of local climate observations.

*Step 1: Regional climate data*

Typically, the length of the climate time series determines the length of the calibration and verification periods. Ideally, the calibration period should be >100 years long to: (1) obtain a sufficient number of independent observations and Degrees of Freedom (DF) for statistical analysis; (2) capture multi-decadal variability; and (3) obtain a large range of observations for broad calibration. In the absence of local climate data, reanalysis data (gridded monthly surface air temperature and precipitation values from 1850/1901 onward; Mitchell and Jones 2005; Brohan et al. 2006) can be considered. In such cases, the quality of the reanalysis data needs to be assessed with all available (short and discontinuous) local and regional instrumental meteorological data. For short calibration periods (i.e. <50 years), the likelihood of obtaining a



**Fig. 1** Flow chart showing the main steps to establish a calibration-in-time between a time series of biogeochemical lake sediment proxies and a time series of climate data in a non-varved lake

statistically significant ( $p < 0.05$ ) correlation is small if the data need to be smoothed.

Subsequently, the dominant frequency bands of variability shown in the meteorological data should be assessed, preferably by statistical analyses. It is necessary to know how many data points are needed to describe the variability in the meteorological time series. Ideally, there is pronounced (multi-) decadal variability, with large amplitudes, i.e. strong variability in a frequency band that is larger than the sampling resolution of the proxy data, the expected dating uncertainty and the potential autocorrelation of the proxy data produced, for instance, by bioturbation. In the case of weak decadal, but strong interannual climate variability, such as an ENSO-dominated region, calibration is difficult because even a small dating error, e.g. offset by one year, may obscure

correlations with meteorological data or lead to spurious ones.

Generally speaking, the prominent frequency domains of climate variability determine the minimum resolution of the sampling interval needed. As a rule of thumb, the time span of the sampling is about three times greater than the temporal resolution of the calibration (and climate reconstruction) envisaged (Koinig et al. 2002). Ideally, the amplitude of variability during the calibration period is large because reconstructions should not be made for data points outside the calibration range.

### *Step 2: Catchment, lake and coring site*

Every lake has a unique geomorphological and hydrological setting and is influenced in different ways by environmental and/or climate factors (Last and Smol 2001; Wetzel 2001). Sound knowledge of the lake system and factors controlling sediment formation is important. Every statistical relationship between a lake sediment climate proxy variable and a climate variable (i.e. the calibration) requires a causal, process-based explanation. In the absence of long observational monitoring series, e.g. data from sediment traps, conceptual model analogues from existing literature may help. Therefore, at least basic physical and chemical water column measurements (e.g. temperature, conductivity, dissolved oxygen, pH) should be collected in the field and a detailed geomorphological map of surface processes and land cover in the catchment should be made (Last and Smol 2001). An accurate bathymetric and seismic survey helps to select the best coring site (Scholz 2001).

### *Step 3: Coring and core selection*

Perfect preservation of the sediment–water interface and the top part of the sediments is imperative for accurate  $^{210}\text{Pb}$  dating and high-resolution (mm-scale) sampling. Sediment mixing should be avoided. We use a 60-mm-diameter UWITEC gravity corer with a transparent PVC liner that allows immediate observation of core stratigraphy.

In the laboratory, cores should be split longitudinally into two halves (A and B), photographed and described using standard protocols (Schnurrenberger

et al. 2003; Ohlendorf et al. 2011). Subsequent analyses of samples for chronology, as well as climate proxy data sets for the calibration and the reconstruction are made on the same core using both halves. This is done to achieve an accurate stratigraphic match between the chronology and the multi-proxy dataset. Stratigraphic correlation of different cores at the precision of the sampling interval required (mm-scale) is not usually possible, except when lake sediments are finely laminated.

Sediment strata should be as straight and flat as possible, with no bending towards the edges of the PVC liner. Ideally strata should be perpendicular to the liner to allow mm-scale sampling of syndepositional sediments. Here, syndepositional means 1–2 years. If sediment bending at the edges is substantial, a sediment slab of known geometry may be cut longitudinally from the center of the half core. This typically requires freezing the sediment with liquid nitrogen and cutting with a band saw.

### *Step 4: Establishing the chronology (core-half A)*

Core-half A is used for analyses that enable preparation of the age–depth model for the calibration period. High accuracy and high precision of the chronology are equally important.

Sampling for  $^{210}\text{Pb}$  measurements should be done at the highest possible resolution that still allows for high-precision measurements and therefore, small uncertainties. Irregular sampling according to microstratigraphic features yields better results and allows one to precisely exclude event layers from the chronology (Arnaud et al. 2002; von Gunten et al. 2009b). The minimum mass of sediment needed for precise activity measurement depends on the method and apparatus used, activity of the radionuclides in the sediment and the counting time (Appleby 2001). The minimum sample mass for a given lake sediment can be experimentally determined from surface sediments or a parallel core. In our experience, a sampling interval of  $0.5 \pm 0.1$  cm, equal to  $\sim 3\text{--}4$  g dry mass from half a 60-mm-diameter core, yields good results for detection of gamma ray.

Procedures to develop highly accurate and precise chronologies for young sediments (<150 years) have been optimized in various ways (Appleby 2001, 2008; Arnaud et al. 2002; von Gunten et al. 2009b). Significant improvements to the accuracy of the

chronology can be made by: (1) systematic comparison of different numerical models to obtain  $^{210}\text{Pb}$ -based chronologies including Constant Flux Constant Sedimentation (CFCS), Constant Initial Concentration (CIC), Constant Rate of Supply (CRS) (Appleby 2001), and Sediment Isotope Tomography (SIT) (Carroll et al. 1995); (2) constraining the numerical age-depth models with independent time markers, typically  $^{137}\text{Cs}$  or polychlorinated biphenyls; and (3) cross-validating the chronology with independent diagnostic stratigraphic markers of known age, e.g. volcanic ash or flood layers, seismites, exotic pollen, or spheroidal carbonaceous particle (SCP) profiles. For a review and examples, see von Gunten et al. (2009b).

The other important feature related to CIT is the probabilistic nature of chronologies. We can conceive of an ensemble of chronologies, with a likelihood determined by the dating uncertainty at each data point (Goslar et al. 2009). This would lead to an ensemble of calibrations that are a function of the probabilistic chronology. This problem is very demanding of computer resources. Koinig et al. (2002) proposed a practical alternative, smoothing the proxy data time series, such that the length of the Gaussian or triangular (low pass) filter is chosen according to the maximum (or average) dating uncertainty in the period considered. This has fundamental consequences: (1) the dating uncertainty has a direct influence on the temporal resolution of the final reconstruction, i.e. on the degree to which the high-frequency signal of climate variability may be retained in the reconstruction; and (2) smoothing has direct effects on the statistical properties of the calibration model (variance and, hence, autocorrelation, DF and significance levels). Thus, minimizing dating uncertainties during the calibration period is a fundamental step. Here, the SIT model has a unique advantage, that uncertainty can be numerically reduced by constraining the model with stratigraphic markers that are dated accurately and precisely. This is particularly important for the early twentieth century, which often is the beginning of the calibration period, when CRS-derived  $^{210}\text{Pb}$  dates typically have large uncertainties and would require substantial smoothing. Finally, samples for  $^{14}\text{C}$  analysis or other dating techniques are collected from lower parts of core-half A to establish the chronology for the reconstruction period.

### *Step 5: Establishing the proxy data set (core-half B)*

Because it is not known a priori which of the numerous measured sediment proxies are climate-sensitive and will finally result in a proxy-climate calibration model, a standard set of organic and inorganic variables can be used. The former variables target lake productivity as a function of climate, whereas the latter target transport processes of allochthonous lithogenic material and variable precipitation of carbonates as a function of climate. Details on specific proxies are outlined in the case studies. The basic concept is to: (1) measure a multi-proxy data set for the calibration period, which includes a minimum of  $\sim 50$ – $80$  samples; (2) assess the sensitivity of each sediment proxy, individually or combined using Principal Components Analyses, to climate variables, with correlation analysis corrected for autocorrelation; and (3) measure only those sediment proxies that passed the calibration tests and perform well as down-core climate predictors.

The choice of proxies also depends on the number of samples that need to be processed for the calibration and reconstruction, as well as the sample mass required for analysis. Rapid, high-resolution, cost-effective and non-destructive methods such as scanning visible (370–730 nm) reflectance spectroscopy (VIS-RS, Rein and Sirocko 2002) or scanning X-ray fluorescence (XRF, Zolitschka et al. 2001) are favored because they provide large multivariate datasets with a series of organic and inorganic proxies in a relatively short amount of time. These methods should be applied to un-oxidized sediments of core-half B before any sediment sub-sampling has taken place. It is generally recommended that work be done at the highest possible resolution over the calibration period and that data be combined or smoothed later, if necessary.

Destructive analytical methods should be applied only after the chronology has been established. This is important for statistical reasons because a sufficient effective sample size is required for calibration. In our experience, an effective sample size of  $\sim 20$  (roughly 80–100 samples, but corrected for autocorrelation) was the lower limit. Higher sampling resolution, however, does not necessarily result in a higher number of effective samples. For example, bioturbation may blur any high-frequency variability. From a practical point of view, sampling at 2-mm intervals is

often the limit for soft sediments. Cores may be frozen with liquid nitrogen to allow for finer sampling. This is useful for cores from lakes with low sedimentation rates,  $<1 \text{ mm year}^{-1}$ .

This sampling strategy usually leads to a very small mass in each sample (200–500 mg) for analyzing multiple proxies. Laboratory methods, e.g. sequential extraction and analysis, should be optimized to account for such small sample masses.

Furthermore, we recommend that all proxies measured in ‘concentration’ units (i.e. %,  $\text{mg g}^{-1}$ ) be converted into ratio or ‘flux’ values ( $\text{mg cm}^{-2} \text{ year}^{-1}$ ) using the age-depth model. This makes proxies independent from changes in Mass Accumulation Rate (MAR) and matrix dilution effects.

#### *Step 6: Data analysis and statistical methods*

Because of the absence of varves, measurement data are for irregular time durations. The irregular proxy data time series’ are regularized (interpolated) to annual values to fit the calendar scale of the meteorological data. Usually, mean annual and averaged seasonal meteorological series are most appropriate, but any significant time span can be considered.

For a number of reasons, including dating uncertainty, sampling error, regularization, autocorrelation of the proxy data and variance stabilization, the annualized proxy time series must be low-pass filtered with a time-invariant, typically Gaussian or triangular filter. The length of the filter should be chosen so that various metrics of the calibration statistics, such as the correlation coefficient,  $p$  value, Reduction of Error (RE), Coefficient of Efficiency (CE) and Root Mean Squared Error of Prediction (RMSEP) are optimized, and the high-frequency variability in the time series is retained.

Next, the homogenized and stepwise-filtered (unfiltered raw annual, 3-, 5- and 7-yr triangular filtered) multiple proxy data are individually, or combined as Principal Components, correlated with the meteorological data using Pearson or Spearman correlation tests. Correlation tests accounting for leads and lags may also be used. It is important that: (1) the meteorological data are filtered in the same way as the proxy data; and (2) significance levels ( $p$  values) are calculated from the number of DF that has been corrected for autocorrelation and trends ( $p_{\text{corr}}$ , according to Dawdy and Matalas 1964; Trenberth 1984). Strong autocorrelation in the sediments, for example

due to bioturbation or a limited number of data points in the calibration period, decreases the effective sample size and DF and thus reduces the likelihood of a good calibration.

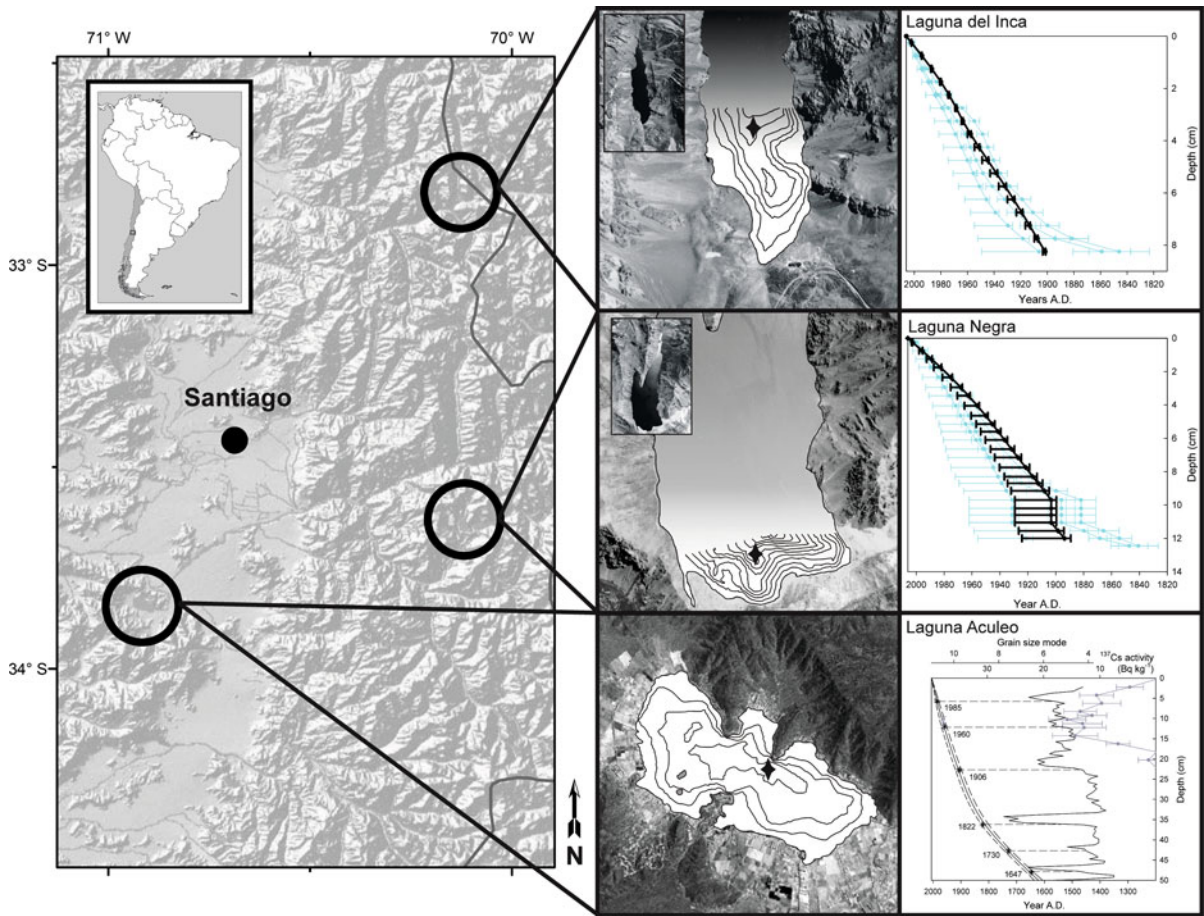
The proxy-climate correlation matrix is then screened for statistically significant ( $p < 0.05$ ) relationships. At this point it is important to eliminate significant correlations that have no causal relationship (‘*False Positives*’) by retaining only those significant correlations that have a process-based causal explanation and/or for which independent proxies reflecting similar environmental variables or processes show comparable results.

Finally, the calibration model is established using an appropriate, usually regression-based model and measurements for the quality of the calibration statistics are calculated (RE, CE, RMSEP; Cook et al. 1994). In theory, the data should be normally distributed (Shapiro–Wilk test; Shapiro and Wilk 1965). Typically, Ordinary Least Squares (OLS) or Major Axis Regression methods are used. Each has different effects on the calibration statistics (Trachsel et al. 2010b). Usually RE, CE and RMSEP are used as quality criteria; however, for climate reconstructions the match with the amplitude of the instrumental target or the ratio between the amplitude and the RMSEP might also be relevant (Legendre and Legendre 1998; Trachsel et al. 2010b).

In tree-ring research, RE, CE and RMSEP statistics are typically calculated from a “split-period approach” (calibration and validation periods; Cook et al. 1994). With non-varved sediments, the number of independent observations is usually much smaller. Therefore, the entire observation period can be used for calibration and the quality of the calibration (verification) is assessed by (k-fold) cross-validation (Legendre and Legendre 1998). Leave-one-out and k-fold cross-validation, however, underestimates RMSEP in autocorrelated time series (Telford and Birks 2009). Using the entire observational period for calibration has the advantage that, in most cases, a broader range of variability is encompassed in the calibration model. To produce the downcore reconstruction, the calibration model is used to transform the biogeochemical data into a climate state variable.

#### Case study area, lakes and sediments

The three remote case study lakes lie in Central Chile (32–34°S), in the western Andes (Fig. 2). The climate



**Fig. 2** Locations of the study lakes, coring sites and core chronologies. **Bold lines** in the age-depth graphs represent the  $^{210}\text{Pb}$  models and uncertainties chosen for Laguna del Inca and

Laguna Negra. Isolines of the bathymetric maps are 10 m (Laguna del Inca, Negra) and 2 m (Laguna Aculeo) contours

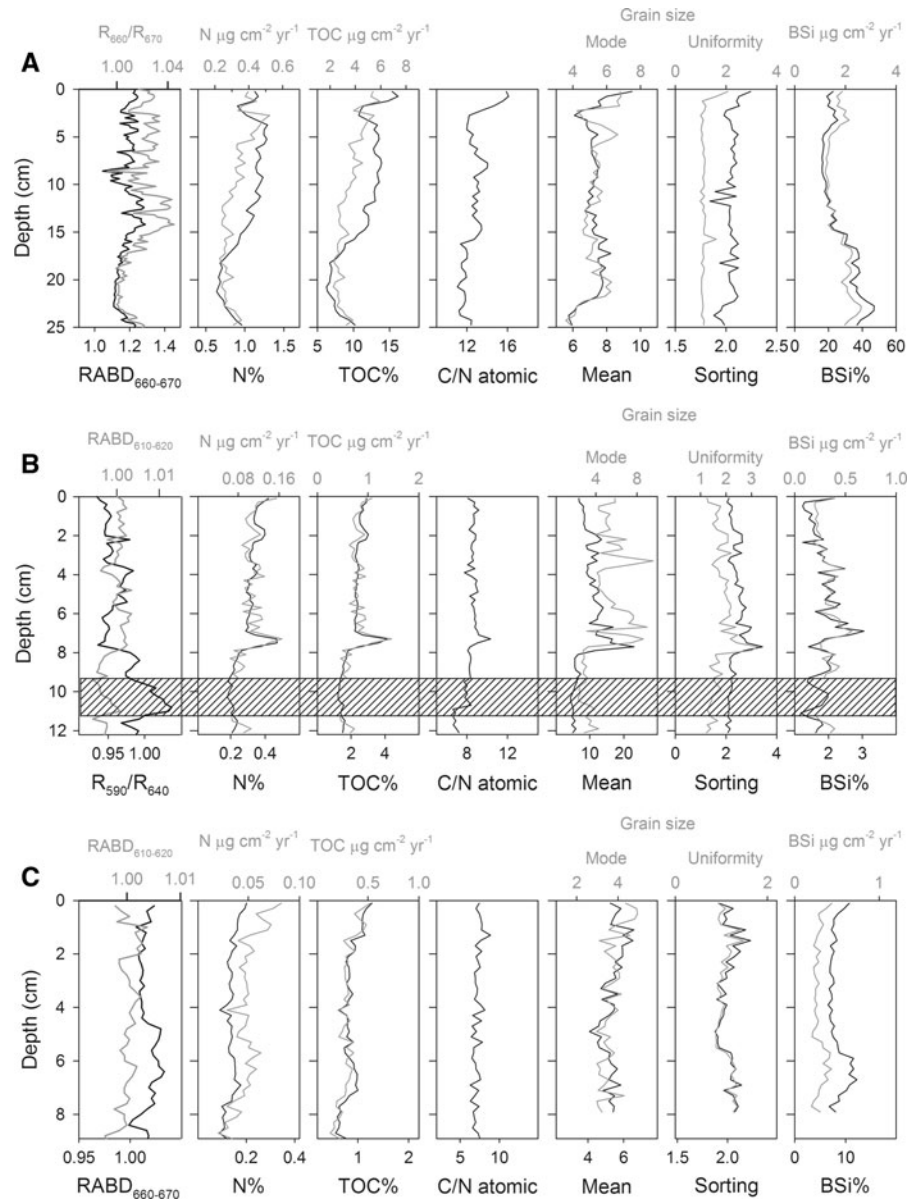
is warm-temperate, with hot (18–20°C), dry summers, cool (7–10°C), humid winters and very high radiation. In summer, the frontal systems of the Southern Hemisphere westerly winds are blocked by a very stable anticyclone over the south-east Pacific (SPA), while in austral winter, northward displacement and weakening of the SPA allows northerly progression of the westerly winds (Aceituno 1988). Interannual variability in precipitation is related to ENSO: warm (cold) ENSO phases result in enhanced (reduced) precipitation (Garreaud et al. 2009).

Laguna Aculeo (33°50'S/70°54'W, 355 m a.s.l.) is a warm, neutral (pH = 7.2), polymictic, hypereutrophic lake of tectonic origin located 50 km south-east of Santiago on the eastern slope of the Coastal Range. It covers an area of 11 km<sup>2</sup> and has a maximum depth of 7 m. The catchment ranges up to

2,280 m a.s.l. and is mostly forested. The lower slopes are used for agriculture and residential housing. The lake sediments are classified as anoxic, dark grey (2.5Y 4/1 to 10YR 4/1, terminology after Munsell Color 1994), massive diatomaceous ooze. A detailed description of the sediments is given in von Gunten et al. (2009a). The chronology and multi-proxy data set discussed here (Fig. 3a) were generated from the same core.

Laguna Negra (33°38'S/70°08'W, 2,680 m a.s.l.) is an oligotrophic lake with a neutral (mean pH = 7.0) and well-oxygenated water column. It is a large (5.7 km<sup>2</sup>) and deep (>120 m) dimictic, glacio-tectonic lake. The catchment extends up to 4,600 m a.s.l., consists of basaltic to dacitic lavas, granodioritic intrusiva and pyroclastic rocks, and is partly glaciated in the north-east (Sernageomin 2003). Vegetation

**Fig. 3** Raw data of a subset of the standard biogeochemical and physical variables measured in **a** Laguna Aculeo, **b** Laguna Negra and **c** Laguna del Inca. The 1906 seismite in Laguna Negra is marked by the shaded area



cover is very scarce (<10%) and is mainly composed of grasses and small shrubs. The lake is fed by sub-surface runoff from the permafrost areas, surface runoff from the glaciers (Glaciar Echaurren) and rainfall. There is little evidence of eutrophication during the past few decades (Fig. 3b). The sediment core consists of massive dark grayish-brown (4/2 2.5Y–4/2 10YR) silt with an aquatic moss (*Fissidens*) layer between 7.2 and 7.6 cm depth. The lithoclastic fraction consists mainly of chlorite, quartz, plagioclase (albite) and mica. The chronology of Laguna

Negra was developed using a constrained and validated <sup>210</sup>Pb SIT model, which has been specifically optimized for CIT. The method is described in detail in von Gunten et al. (2009b).

Laguna del Inca (32°50'S/70°08'W, 2,840 m a.s.l.) is a dimictic oligotrophic lake with neutral pH (mean = 6.9). It is deep (>120 m) and has a surface area of 1.6 km<sup>2</sup>. The catchment is large (~46 km<sup>2</sup>), steep and extends above 4,600 m. The lake freezes in winter. The geology consists of basaltic to dacitic lavas and pyroclastic rocks (Sernageomin 2003). In



contrast to Laguna Negra, the catchment of Laguna del Inca is not glaciated, but has large areas with extended rock glaciers and permafrost in the north. Vegetation cover is very scarce (<10%) and is mainly composed of grasses and small shrubs near the lake. Sediments are classified as massive (partly laminated) dark brown (10 YR 3/3) to dark reddish grey (5YR 4/2) silty-clay. It remains inconclusive whether the laminae represent annual cycles or events. Weak eutrophication is recorded since ca. 1980 (von Gunten et al. 2009c). The chronology was established with a  $^{210}\text{Pb}$  SIT model, constrained and validated with the 1964  $^{137}\text{Cs}$  peak, a SCP profile and the seismite related to the AD 1906 earthquake (von Gunten et al. 2009c).

#### Methods and data for the case study lakes

##### *Climate data*

In the absence of adequate local instrumental data, gridded reanalysis data were extracted from the HadCRUT3 ( $5^\circ \times 5^\circ$ , AD 1850–2006; Brohan et al. 2006) and CRU TS 2.1 ( $0.5^\circ \times 0.5^\circ$ , AD 1901–2000; Mitchell and Jones 2005) temperature and precipitation reanalysis datasets, and from the  $2.5^\circ \times 2.5^\circ$  observed land surface precipitation dataset (AD 1901–1997; Dai et al. 1997). Large-scale circulation indices such as the Southern Oscillation Index (SOI) back to AD 1866 (Allan et al. 1991), El Niño-3 Index (AD 1408–1978; D'Arrigo et al. 2005) and a Sub-tropical Anticyclone Belt position reconstruction (AD 1500–1975; Bradley 1992) were also used.

Quality of the reanalysis datasets was tested by comparing them to raw data from the nearest instrumental meteorological station. All datasets showed identical patterns and trends, and were highly significantly correlated to each other (Supplementary Online Material Fig. S1).

##### *Field and laboratory methods*

Short sediment cores were taken with an UWITEC gravity corer equipped with an action hammer (Laguna Negra and Laguna del Inca) and a modified Livingstone piston corer (Laguna Aculeo). The choice of coring sites was based on bathymetric surveys. Sediment cores were sealed and stored under cold ( $4^\circ\text{C}$ ) and dark conditions.

After being opened, sediment cores were photographed and their stratigraphy was described. Thin sections for micro-facies analysis were prepared for Laguna Negra and Laguna de Inca. The uppermost 25 cm of each core-half A were then stratigraphically sampled at spacing of 0.4–0.7 cm, according to facies, for  $^{226}\text{Ra}$ ,  $^{210}\text{Pb}$  and  $^{137}\text{Cs}$  measurements, i.e. gamma ray counting. The same sub-samples were analyzed for additional chronostratigraphic markers (i.e. SCP profiles, pollution markers, earthquake layers; von Gunten et al. 2009a, b, c) to constrain the  $^{210}\text{Pb}$ -based CRS and SIT models, and to verify the chronology for the calibration period (1880–2006).

After the depth-age models were established and the resolution determined that would provide 50–100 data points for the calibration period, the corresponding core-halves (B) were scanned with VIS-RS at 2-mm intervals, corresponding to <1–2 years, with a Gretag-Spectrolino (GretagMcBeth, Switzerland). Information and technical details about this method can be found in Rein and Sirocko (2002), Rein (2003) and Trachsel et al. (2010a).

Using the algorithms defined by Rein and Sirocko (2002), three variables were extracted from the spectra: (1) the relative absorption band depth at 660–670 nm ( $\text{RABD}_{660-670}$ ), which is a proxy for total sedimentary chlorin; (2) the ratio of reflectance at 660–670 nm ( $\text{R}_{660}/\text{R}_{670}$ ), which is a proxy for the degree of photopigment diagenesis; and (3) the ratio of reflectance at 590–640 nm ( $\text{R}_{590}/\text{R}_{640}$ ), which is a proxy for the lithogenic content, specifically illite, chlorite and mica (Trachsel et al. 2010a). In Laguna Negra and Laguna Inca, the relative absorption band depth at 610–620 nm ( $\text{RABD}_{610-620}$ ) was also used, which is indicative of albite, as it absorbs light around 620 nm (USGS Spectral Library). The algorithm used to calculate  $\text{RABD}_{610-620}$  is:

$$\text{RABD}_{610-620} = [(\text{R}_{590} + \text{R}_{640})/2]/\text{R}_{\min(610;620)} \quad (1)$$

Albite was determined by X-ray diffraction (XRD) in the sediments of Laguna Negra and is diagnostic for sediment sources from the granodioritic intrusive, with 40–50% feldspars forming the bedrock underneath the Echaurren Glacier (Sernageomin 2003). In the sediment core,  $\text{RABD}_{610-620}$  is highly negatively correlated with  $\text{R}_{590}/\text{R}_{640}$  ( $r = -0.76$ ,  $p < 0.001$ ) and highly positively correlated with the grain size median ( $r = 0.56$ ,  $p < 0.001$ ), indicating the lithogenic sensitivity of the

proxy RABD<sub>610–620</sub>. XRD analysis confirmed low chlorite/albite Peak Intensity Ratios in samples with high RABD<sub>610–620</sub>.

The cores from Laguna Negra and Laguna del Inca were sampled at 2-mm intervals and the core from Laguna Aculeo was sampled at 5-mm intervals. Sediment sub-samples (100–200 mg) were freeze-dried and analyzed for total organic carbon (TOC) and total nitrogen (N) content using a Vario Macro Elemental Analyzer. Inorganic carbon was removed with HCl (10%) prior to analysis. Biogenic silica (BSi) was leached according to Ohlendorf and Sturm (2008) from organic-free (30% H<sub>2</sub>O<sub>2</sub>) samples and measured with inductively coupled plasma atomic emission spectroscopy (ICP-OES). BSi concentrations were corrected for inorganic silicon derived from silicate minerals, using the Al-concentration in the leachate (Al:Si ratio of 2:1). After BSi extraction, the remaining mineroclastic sediment samples were treated with HCl (10%) and used for laser grain-size measurements on a Mastersizer 2000. From the grain size distributions, we determined the following variables: mean, median, sorting, mode and uniformity. Finally, BSi, TOC and N fluxes were calculated using the MARs derived from the chronology.

### Statistical methods

To synchronize the proxy data to the time intervals of the meteorological data, all proxy data were first regularized by linear interpolation to annual intervals. The regularized annual series were then smoothed with 3-, 5-, 7- and 9-year triangular filters to account for resampling errors and dating uncertainties (Koinig et al. 2002). For the calibration, meteorological data were treated with the same filter that was applied to the proxy data. Pearson's product-moment correlation coefficient was used and P values were corrected for autocorrelation and trends (Trenberth 1984). In this study we use the terms “highly significant correlation” for  $p < 0.01$ , “significant correlation” for  $p < 0.05$  and “non-significant correlation” for  $p > 0.05$ . For simplicity, the examples here only use OLS with scaling (Cook et al. 1994; Esper et al. 2005). To cover the full range of climate variability in the calibration period and to keep the number of DF high, the entire reanalysis period was used for the calibration models. The quality of the reconstruction was assessed with split-period cross-validation, for which the twentieth

century was split into a calibration and verification period. RMSEP, RE and CE were calculated after Cook et al. (1994). Finally, tenfold cross-validated RMSEP was calculated over the entire calibration period to estimate the prediction error.

The most recent sediments of Laguna Aculeo (AD 1997–2005) were excluded from calibration because the disturbed and waterlogged sediments at the sediment–water interface were difficult to sample and the C/N ratios of these sediments have anomalously high values (molar C/N > 15), which suggests the presence of non-aquatic sources of sediment organic matter. In Laguna del Inca, sediments younger than 1980 were excluded because of eutrophication (von Gunten et al. 2009c). In Laguna Negra the *Fissidens* plant layer and the data of the 1906 seismite were excluded from the calibration. R software was used for statistical analysis (R Development Core Team 2011). The scripts are available at: <http://tinyurl.com/Rscripts>.

## Results

Figure 3a–c display parts of the organic and inorganic proxy data sets for the three case study lakes. Data series that were not correlated with climate are not shown.

The organic sediments of Laguna Aculeo (Fig. 3a) showed considerable variability in all biochemical proxies (R<sub>660</sub>/R<sub>670</sub>, N, TOC, C/N and BSi), while the physical proxies (grain-size mode, mean, uniformity and sorting) remained relatively stable. The correlation analysis, in which all proxies were correlated with seasonal and annual temperature and precipitation, revealed highly significant correlations between the proxies for primary production in the lake and temperature (summer and annual), while no significant correlations were found between the physical proxies and climate (SOM Table S1). No proxy showed a significant correlation with precipitation or a climate index (e.g. El Niño-3). The highest correlations were found between RABD<sub>660;670</sub> and R<sub>660</sub>/R<sub>670</sub>, and austral summer DJF temperatures (CRU TS2.1,  $r = 0.77$  [0.84],  $p < 0.001$  [ $<0.01$ ] for 3 [5]-year filtered data). TOC and N were also significantly correlated with unfiltered and 3-year filtered DJF temperatures.

It appeared from the statistical analysis that, in Laguna Aculeo, productivity proxies (photopigments, TOC and N) were sensitive to summer temperature, but not to precipitation, whereas none of the physical proxies seemed to be sensitive to climate (data not shown). von Gunten et al. (2009a) concluded that in hypereutrophic Laguna Aculeo, primary production is not limited by nutrient availability, but rather by light availability, which in turn is related to temperature. Also, photo-oxidation, which influences the degree of photopigment diagenesis, is dependent on light availability and thus covaries with temperature. This interpretation was also supported by the observation that no eutrophication effect was found in the lake sediments of the 20<sup>th</sup> century, despite large land use changes and anthropogenic impact.

The minerogenic sediments of Laguna Negra showed variability in all measured proxies, suggesting considerable changes in water-column productivity in the lake, and transport of clastic material from the catchment into the lake during the past 100 years. The high values of TOC, N and grain size between 7.2 and 7.6 cm are related to the *Fissidens* layer and were excluded from the correlation (Fig. 3b).

For Laguna Negra, highly significant correlations were found between proxies for organic content (N%, N flux, TOC%, TOC flux) and the lithogenic RABD<sub>610–620</sub> when compared to annual and summer temperatures (Electronic Supplementary Material ESM, Table S1b). Best correlations resulted for the relationship between RABD<sub>610–620</sub> and austral summer and annual temperatures (HadCRUT3,  $r = 0.78$  [0.81],  $p < 0.001$  [ $<0.01$ ], 3 [5]-year filtered). Comparison with precipitation and other climate indices did not yield any significant results.

In Laguna Negra, both biological and physical correlations with temperature were plausible. It is well established that biological activity in high-altitude lakes can be strongly related to temperature (Meyers 1997; Battarbee et al. 2002; Koinig et al. 2002). RABD<sub>610–620</sub>, on the other hand, appeared to reflect higher relative proportions of albite in the lithoclastic component. In the catchment of Laguna Negra, albite is diagnostic for the granodioritic bedrock underneath the glaciated part in the NW (Glaciar Echaurren) and would be mobilized preferentially during warm summers when glacier melt and sediment transport from this area is enhanced. Conceptually, enhanced melt-water flux would also lead to stronger undercurrents in

the lake, when relatively coarse-grained suspended sediments (with albite) would be conveyed to the distal parts of the lake and the coring site. However, this process has not been observed or measured in the field.

The sediments of Laguna del Inca are almost exclusively lithogenic (average N: 0.1%, TOC: 0.7%). All organic proxies (TOC, N, RABD<sub>660–670</sub>) except BSi showed only little variability over most of the calibration period. In contrast, the concentration of BSi and grain size parameters showed much larger changes (Fig. 3c). Significant negative correlations were found between organic proxies (BSi, N, TOC concentrations), physical proxies (grain size uniformity and sorting) and temperature (SOM Table S1c). The highest negative correlation was found between BSi concentration and summer temperature (CRU TS2.1,  $r = -0.75$  [ $-0.83$ ],  $p < 0.01$  [0.02], 3 [5]-year filtered). No significant correlation was found for precipitation. Although the concentrations of BSi, N and TOC were highly negatively correlated with warm season temperatures, the fluxes of the same proxies correlated only weakly with temperature (data not shown). This suggests that the negative correlation between organic matter concentration in the sediments and temperature is largely controlled by matrix effects (dilution effects), as a function of MAR. Conceptually, warm temperatures release larger water masses from the vast permafrost areas and rock glaciers in the north of the catchment. The main hydrologic tributaries then transport particles, MAR is enhanced, and consequently, the concentration of organic matter in the sediments (inferred from TOC, N and BSi) decreases. It appeared that variability in MAR, as well as matrix effects that are a consequence of variable sediment transport, obscured possible positive relations between primary production and temperature as discussed in Laguna Negra. This is not surprising, because primary production in Laguna del Inca is much smaller than in Laguna Negra. The lake-to-catchment ratio of Laguna del Inca (0.035) is only about one-fourth that of Laguna Negra (0.13), suggesting that catchment processes such as sediment transport play a more dominant role in Laguna del Inca.

Figure 4a–c and Table 1 show the calibration curves for the best climate proxy in each of the three lakes for the annual, 3- and 5-year filtered data and the respective calibration statistics. In all three lakes, the selected proxy data reproduced the decadal temperature trends. Even sub-decadal to interannual variability was partly

resolved in Laguna Aculeo and Laguna Negra, whereas in Laguna del Inca, sub-decadal temperature variability was not resolved. Laguna del Inca has the lowest sedimentation rates ( $\sim 0.8 \text{ mm yr}^{-1}$ ), which limits the maximum possible sampling resolution for BSi measurements. In general, even the 5-year filtered proxy data showed smaller variability than the 5-year filtered climate data, suggesting that the lake sediment proxy series contains additional ‘smoothers,’ enhancing auto-correlation. Possible causes are the lake system itself (e.g. bioturbation) or the propagation of sampling errors. The point, however, is that the technical resolution (near-annual data) does not necessarily mean that the full climate variability at this temporal resolution is actually recorded in the sediments. The calibration statistics were very good for Laguna Negra and Laguna Aculeo, with high RE and CE values.

For Laguna del Inca, the calibration quality (RE and CE values) was lower. It appears that sub-decadal to decadal-scale variability of the BSi proxy during the calibration period (AD 1945–1978) is minimal and did not follow the observed temperature cooling from 1960 AD onwards (Fig. 4c). This offset might be an effect of early eutrophication, but in this context it illustrates that pronounced decadal variability is very important to obtain good calibration statistics (see Step 1 Overall Methodology). The calibration statistics in Laguna del Inca were much better if the calibration and verification periods are switched (calibration: AD 1911–1944, validation: AD 1945–1997; RE = 0.8 and CE = 0.1). The RMSEPs were very low in all three lakes (0.24–0.29°C), attesting to the good quality of all calibrations and low error of prediction. Also, the RMSEP (uncertainty) was low compared with the amplitude of decadal-scale variability.

## Discussion

In this study, we focus discussion on the methodological aspects of the calibration-in-time CIT approach rather than on the results from the individual case studies per se. Although our calibration test showed that it was possible to apply the CIT approach successfully in all three study lakes, some of the methodological steps are critical. In the following section, we discuss two of those important steps: (1) optimization of the calibration and (2) plausibility of the calibration and implications for climate reconstructions.

## Optimizing the calibration

Here we discuss three issues: (1) quality and choice of the climate data, (2) optimum sampling resolution and (3) optimum resolution for the calibration and reconstruction. Many lakes are located in remote areas where sufficiently long local meteorological data are missing. Our case study examples from the central Andes show that reanalysis data sets may be suitable substitutes for local climate information. In this case, the quality of reanalysis data needs to be assessed using available shorter, and perhaps discontinuous observational data. For some applications and paleoclimate interpretations, the use of reanalysis data as a target for the proxy-climate regression is recommended. We can express meteorological data from a local weather station adjacent to a given lake as a combination of the regional climate signal (plus noise) and the local climate signal (plus noise), for instance:

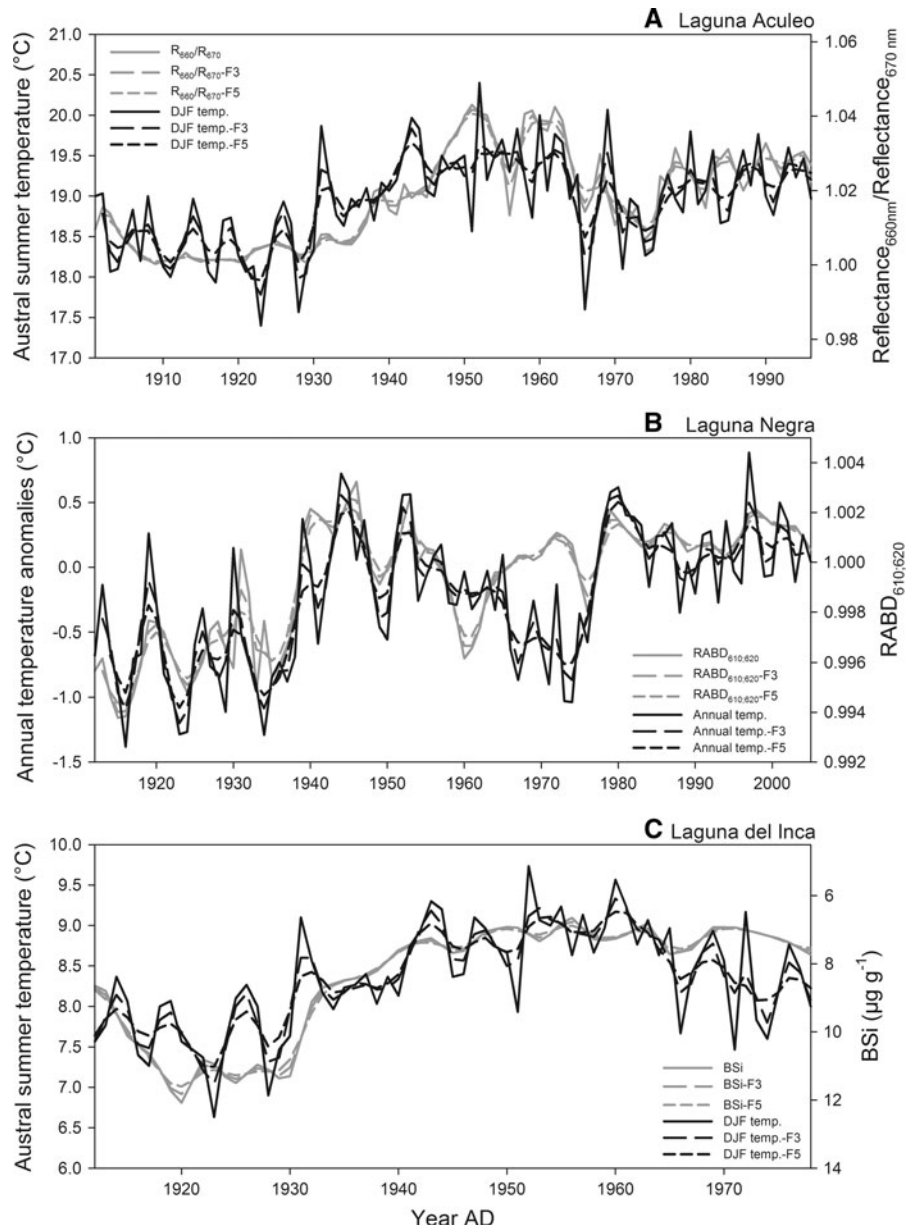
$$T = (T_{\text{reg}} + \varepsilon_{\text{reg}}) + (T_{\text{loc}} + \varepsilon_{\text{loc}}) \quad (2)$$

where  $T$  is the temperature and  $\varepsilon$  is the noise. We can then regress the lake sediment proxies against the local meteorological data. In the case where the local component of the climate signal is very strong, e.g. in complex mountain topography, the regression statistics might be very good, but the ability of the site to predict regional climate might be poor. If proxy data from a lake are regressed against a larger area, as represented in reanalysis data ( $0.5^\circ \times 0.5^\circ$  or  $5^\circ \times 5^\circ$ ), the regression statistics might be poorer, but local effects are eliminated, i.e. appear as uncertainty in the calibration model. The size of the climate target influences the correlation with the local proxy, which partly explains why correlations with the CRU TS 2.1 and the HadCRUTS3 data sets yield different results (Tables S1–3). The value of this approach is revealed in that the created proxy dataset compares better to other proxy records or model runs from the region. Alternatively, the regional significance of local information should be tested with spatial correlation analysis.

Our examples illustrate the importance of strong decadal-scale variability with large amplitudes in the climate data and in the proxy data time series for successful CIT. These conditions are not fulfilled in Laguna del Inca in the calibration period and explain the poor RE and CE statistics.

The minimum number of sub-samples  $N_{\text{min}}$  (data points in the calibration period) and the sampling

**Fig. 4** Selected sediment proxies plotted against the reanalysis temperatures in the three lakes for annual, 3-year (F3) and 5-year (F5) triangular filtered data. **a** Laguna Aculeo: CRU TS 2.1 austral summer temperature and VIS-RS  $R_{660nm}/R_{670nm}$ , **b** Laguna Negra: HadCRUT3 mean annual temperature and VIS-RS  $RABD_{610,620}$ , **c** Laguna del Inca: CRU TS 2.1 austral summer temperature and BSi concentration



intervals for a successful CIT depends on several factors: (1)  $N_{min}$  should be large enough that the major features of the meteorological time series (variability structure) can be described; (2) samples should be taken at the highest possible temporal resolution, which depends on the minimum sample mass needed to acquire the multi-proxy data set and on the core cutting device; in this respect, the minimum sampling interval is a function of the core diameter and the sediment density; (3) the sampling interval depends on the sedimentation rate, and the temporal resolution of

the original raw data set (sampling) should be about 3–4 times greater than the final resolution envisaged for the calibration and the reconstructions, to avoid errors as a consequence of resampling and filtering procedures; and (4) the sampling resolution is a function of the degree of autocorrelation in the sediment stratigraphy. For instance, sediment mixing (e.g. bioturbation, pools and reservoirs in the cycling of bio-geochemical species within the lake, or imperfect sampling along irregular sediment surfaces) enhances autocorrelation of the proxy data time series,

**Table 1** Calibration statistics for the best calibration for each lake (all 5-year filtered)

Laguna	Linear regression		Pearson r	RE	CE	RMSEP (°C)
Aculeo	Cal.	1949–1996	0.80	0.79	0.56	0.24
	Verif.	1903–1948	0.81			
Negra	Cal.	1960–2003	0.45	0.89	0.83	0.26
	Verif.	1914–1959	0.93			
Inca	Cal.	1945–1978	0.32	0.11	−1.77	0.29
	Verif.	1912–1944	0.86			

*Cal.* calibration period, *Verif.* verification period, *RE* reduction of error, *CE* coefficient of efficiency, *RMSEP* root mean squared error of prediction (tenfold cross-validated)

which in turn reduces the effective sample size (and DF, Dawdy and Matalas 1964) and determines the variability that is recorded in the proxy series. Sampling at much higher resolution does not improve the proxy series.

Generally, many independent observations (low lag-1 autocorrelation, e.g. low bioturbation) yield a more robust correlation (low  $p$  value). Both the proxy time series and the climate time series are typically autocorrelated and one must adjust for the DF in the statistical tests (Trenberth 1984). At the same time, progressive data smoothing typically increases the correlation coefficients and removes errors due to the probabilistic chronology of the proxy series (Koinig et al. 2002), and also reduces the RMSE. Therefore, there is a trade-off between positive effects of data filtering (higher correlation coefficients R, RE and CE, reduced RMSE) and negative effects (loss of DF due to higher autocorrelation, loss of significance, and loss of high-frequency variability in the calibration and reconstruction). Thus, the question of optimized temporal resolution (degree of filtering) of the calibration and reconstruction arises. For example, Fig. 5 shows results for R, RE, CE, RMSEP, DF,  $p$ -values and the respective 95 and 99% confidence levels, with stepwise filtering from 1- to 9-year filters for Laguna Aculeo. It appears from this empirical sensitivity test that, in Laguna Aculeo, the optimum is obtained with a 5-year triangular filter. It also appears (Fig. 5b) that this is precisely the span of the filter when the lag-1 autocorrelations in both time series (climate data and proxy data) converge, i.e. when the sediment mixing bias (bioturbation) disappears. This degree of smoothing represents the upper limit of temporal resolution that is meaningful for a paleoclimate reconstruction.

Using similar tests (data not shown) we found the optimized configuration for Laguna del Inca with a

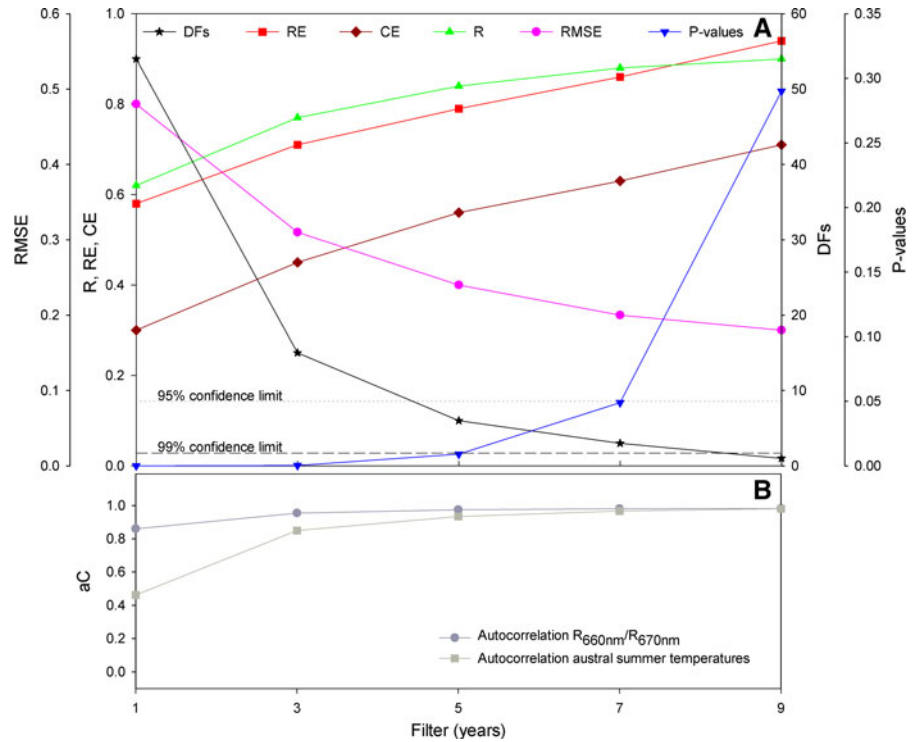
filter of 5–7 years, and for Laguna Negra with a filter of 3–5 years. This is also visible in the calibration curves shown in Fig. 4a–c. It is, however, important to recall the importance of the accuracy and precision of the chronology. The lower the quality of the chronology (accuracy and precision), the larger the span of the filter required to achieve a decent correlation coefficient, but then the likelihood for statistical significance ( $p$  value) decreases rapidly.

Because many of the factors used for optimization are not known at the time when sediment samples are taken, we recommend, as a rule of thumb, a minimum of 50, but preferably 80–100 data points for a 100-year long calibration/verification period. This explains again why the chronology needs to be established before the sediments are cut (Steps 4 and 5 Overall Methodology). Our examples show that even lakes with relatively low sedimentation rates ( $\sim 10$  cm  $100$  year $^{-1}$ ) are suitable.

#### Plausibility of the calibrations and reconstructions

We used a set of organic and inorganic sediment variables and applied a systematic comparison with seasonal temperature and precipitation data (correlation matrix). We found that all three case study lakes show high sensitivity to summer and/or annual temperatures. Depending, however, on the lake type, catchment configuration and sediment formation, the lakes respond to different processes and, consequently different proxies display response to the same climate variable. This common response is surprisingly consistent and supports the causal nature of the statistical relationship. The probability is very small that the relationship between sediment proxies and warm season temperature appears by coincidence in three lakes, using different proxies. It is interesting that no

**Fig. 5** Calibration statistics for Laguna Aculeo for 1-9-year triangular filtered data **(a)** *DFs* degrees of freedom, *RE* reduction of error, *CE* coefficient of efficiency, *R*: Pearson correlation coefficient, *RMSEP*: root mean squared error of prediction and **b** Autocorrelation (*aC*) as a function of filtering for austral summer temperature and  $R_{660}/R_{670}$



proxy in any of the lakes showed sensitivity to precipitation or cold season temperature. In terms of climate reconstruction, information about temperature is valuable because all the other known paleoclimate proxies for this area are sensitive to precipitation (Villalba et al. 2009).

Presentation of climate reconstructions using downcore measurements is beyond the scope of this article. The next logical step, however, is to measure those proxies that yielded good calibration results. von Gunten et al. (2009a) showed in Laguna Aculeo that highly resolved (sub-decadal) reconstructions for the past ca. 1,000 years can be made, and Neukom et al. (2011) found that the von Gunten et al. (2009a) reconstruction from Laguna Aculeo is one of the best predictors for southern South American temperatures and is highly consistent with regional tree ring, coral and ice core data.

When applying CIT for climate reconstructions it is important to keep in mind that a good calibration does not necessarily imply that a good down-core reconstruction can be obtained. The main caveat of the CIT method is the assumption of stationarity, i.e. that the relationship between the proxy and climate variable as determined for the calibration period has remained

constant. Also, very long time series ( $>10^3$  years) of sediment proxies may be affected by non-climatic trends related to factors such as lake ontogeny. On shorter time scales, perturbations such as damming, landslides, changes in vegetation and hydrology, glacier fluctuations (Last and Smol 2001) may be problematic. Finally, the amplitude of variability during the reconstruction should not exceed the range of calibration, that is, the regression should not be extrapolated beyond experienced conditions.

### Conclusions

In this study we used a methodological approach to calibrate biogeochemical and sedimentological proxy data from non-varved lake sediments to meteorological data. We used data from three lakes in central Chile and showed that the CIT approach works even under difficult conditions, i.e. in remote areas where there is little information available and lakes display relatively slow sedimentation rates. We used standard organic and inorganic biogeochemical and physical analyses and applied statistical methods to produce

and validate a regression-based calibration with meteorological data.

From the examples in this study, we conclude that a successful CIT has several prerequisites:

1. A qualitatively good and long (ideally >100 years) meteorological time series for the study area. In the absence of local meteorological data, reanalysis data can be considered. The climate data should present strong decadal variability because interannual variability is very hard to capture with non-varved sediment data as a consequence of dating uncertainties and bioturbation.
2. Basic knowledge of catchment and lake processes. Suitable lakes should have sedimentation rates >10 cm 100 year<sup>-1</sup>, show little bioturbation or sediment disturbance (turbidites) and must not show evidence of major anthropogenic disturbances, e.g. eutrophication or hydrologic modification of the catchment.
3. A highly accurate and precise chronology. In non-varved lakes, we recommend routine calculation of all three standard <sup>210</sup>Pb models (CIC, CRS and SIT) and validation with independent stratigraphic markers. Independent stratigraphic markers are worth the effort and are important to reduce the error of the chronology at the beginning of the calibration period (~1900).
4. Continuous analysis of sediment cores at very high resolution over the calibration period, with a minimum of 50, but ideally 80–100 data points. A large number of independent observations are important for the calibration statistics. The advantage of high-resolution, non-destructive scanning methods is obvious. For destructive methods, we recommend establishing the chronology first, followed by sampling of the core.
5. Rigorous testing of the calibration and strong conceptual understanding of the processes underlying the statistical proxy-climate relationship.

The overall methodology and protocols presented here need further refinements, testing and improvements. We hope, however, that our proposed methodology and practical suggestions facilitate and stimulate more work on calibration and quantification of climate reconstructions from non-varved lake sediments.

**Acknowledgments** We thank D. Fischer, A. Müller, B. Rein and A. Zwyssig for technical and laboratory assistance, H. Alonso, I. Alvial, A. Araneda, X. Boës, M. Espinoza and C. Salvetti for help during fieldwork, S. Köchli for making the ICP-OES measurements and K. Saunders for her comments on an earlier draft of the manuscript. Aguas Andinas and DIFROL are acknowledged for permission to conduct research. This project was funded by the Swiss National Science Foundation (NF-200021-107598, NF 200020\_121869 and PBBEP2-126056) and the Chilean Swiss Joint Research Programme No. CJRP-1001.

## References

- Aceituno P (1988) On the Functioning of the Southern Oscillation in the South-American Sector 1 Surface Climate. *Mon Weath Rev* 116:505–524
- Allan RJ, Nicholls N, Jones PD, Butterworth IJ (1991) A Further Extension of the Tahiti-Darwin SOI, Early ENSO Events and Darwin Pressure. *J Clim* 4:743–749
- Appleby PG (2001) Chronostratigraphic techniques in recent sediments. In: Last WM, Smol JP (eds) *Tracking environmental change using lake sediments: basin analysis coring and chronological techniques*, vol 1. Kluwer Academic Publishers, Dordrecht, pp 171–201
- Appleby PG (2008) Three decades of dating recent sediments by fallout radionuclides: a review. *Holocene* 18:83–93
- Arnaud F, Lignier V, Revel M, Desmet M, Beck C, Pourchet M, Charlet F, Trentesaux A, Tribouvillard N (2002) Flood and earthquake disturbance of Pb-210 geochronology (Lake Anterne, NW Alps). *Terra Nova* 14:225–232
- Battarbee RW, Thompson R, Catalan J, Grytnes J-A, Birks HJB (2002) Climate variability and ecosystem dynamics of remote alpine and arctic lakes: the MOLAR project. *J Paleolimnol* 28:1–6
- Bigler C, Hall RI (2002) Diatoms as indicators of climatic and limnological change in Swedish Lapland: a 100-lake calibration set and its validation for paleoecological reconstructions. *J Paleolimnol* 27:97–115
- Birks HJB (1998) Numerical tools in palaeolimnology - progress, potentialities, and problems. *J Paleolimnol* 20:307–332
- Bradley RS (1992) Climate since AD 1500 database. IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series # 92-015. NOAA/NGDC Paleoclimatology Program, Boulder, USA
- Brohan P, Kennedy JJ, Harris I, Tett SFB, Jones PD (2006) Uncertainty estimates in regional and global observed temperature changes: a new data set from 1850. *J Geophys Res Atmos* 111, doi:10.1029/2005JD006548
- Carroll J, Lerche I, Abraham JD, Cisar DJ (1995) Model determined sediment ages from Pb-210 profiles in unmixed sediments. *Nucl Geophys* 9:553–565
- Cook ER, Briffa KR, Jones PD (1994) Spatial regression methods in dendroclimatology—a review and comparison of two techniques. *Int J Clim* 14:379–402
- D'Arrigo R, Cook ER, Wilson RJ, Allan R, Mann ME (2005) On the variability of ENSO over the past six centuries. *Geophys Res Lett* 32, doi:10.1029/2004GL022055



- Dai A, Fung IY, DelGenio AD (1997) Surface observed global land precipitation variations during 1900–88. *J Clim* 10:2943–2962
- Dawdy DR, Matalas NC (1964) Statistical and probability analysis of hydrologic data part, part III: analysis of variance, covariance and time series. In: Chow VT (ed) *Handbook of applied hydrology a compendium of water-resources technology*. McGraw-Hill, New York, pp 68–90
- Elbert J, Grosjean M, von Gunten L, Urrutia R, Fischer D, Wartenburger R, Ariztegui D, Fujak M, Hamann Y (2012) Quantitative high-resolution winter (JJA) precipitation reconstruction from varved sediments of Lago Plomo 47°S, Patagonian Andes, AD 1530–2001. *Holocene*. doi: [10.1177/0959683611425547](https://doi.org/10.1177/0959683611425547)
- Esper J, Frank DC, Wilson RJS, Briffa KR (2005) Effect of scaling and regression on reconstructed temperature amplitude for the past millennium. *Geophys Res Lett* 32, doi:[10.1029/2004GL021236](https://doi.org/10.1029/2004GL021236)
- Francus P, Bradley RS, Abbott MB, Patridge W, Keimig F (2002) Paleoclimate studies of minerogenic sediments using annually resolved textural parameters. *Geophys Res Lett* 29(20), doi:[10.1029/2002GL015082](https://doi.org/10.1029/2002GL015082)
- Garreaud RD, Vuille M, Compagnucci R, Marengo J (2009) Present-day South American climate. *Palaeogeogr Palaeoclimatol Palaeoecol* 281(3–4):180–195
- Goslar T, Van Der Knaap WO, Kamenik C, Van Leeuwen JFN (2009) Free-shape <sup>14</sup>C age-depth modelling of an intensively dated modern peat profile. *J Quat Sci* 24:481–499
- Grosjean M, von Gunten L, Trachsel M, Kamenik C (2009) Calibration-in-time: Transforming biogeochemical lake sediment proxies into quantitative climate variables. *PAGES News* 17(3):108–110
- Hegerl GC, Crowley TJ, Hyde WT, Frame DJ (2006) Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature* 440:1029–1032
- Imbrie J, Kipp NG (1971) A new micropaleontological method for quantitative paleoclimatology. Application to a late Pleistocene Caribbean core. In: Turekian KK (ed) *The late Cenozoic glacial ages*. Yale Univ Press, New Haven, pp 71–131
- Kalugin I, Daryin A, Smolyaninova L, Andreev A, Diekmann B, Khlystov O (2007) 800-yr-long records of annual air temperature and precipitation over southern Siberia inferred from Teletskoye Lake sediments. *Quat Res* 67:400–410
- Kaufman CA, Lamoureux SF, Kaufman DS (2011) Long-term river discharge and multidecadal climate variability inferred from varved sediments, southwest Alaska. *Quat Res* 76:1–9
- Koinig KA, Kamenik C, Schmidt R, Agusti-Panareda A, Appleby PG, Lami A, Prazakova M, Rose N, Schnell OA, Tessadri R, Thompson R, Psenner R (2002) Environmental changes in an alpine lake (Gossenköllesee, Austria) over the last two centuries - the influence of air temperature on biological parameters. *J Paleolimnol* 28:147–160
- Last WM, Smol JP (2001) *Tracking environmental change using lake sediments: basin analysis, coring, and chronological techniques*, vol 1. Kluwer Academic Publishers, Dordrecht
- Legendre P, Legendre L (1998) *Numerical ecology*. Elsevier, Amsterdam
- Luterbacher J, Dietrich D, Xoplaki E, Grosjean M, Wanner H (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* 303:1499–1503
- Mann ME, Zhang Z, Hughes MK, Bradley RS, Miller SK, Rutherford S, Ni F (2008) Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *PNAS* 105(36):13252–13257
- McKay NP, Kaufman DS, Michelutti N (2008) Biogenic silica concentration as a high-resolution, quantitative temperature proxy at Hallet Lake, south-central Alaska. *Geophys Res Lett* 35, doi:[10.1029/2007GL032876](https://doi.org/10.1029/2007GL032876)
- Meyers PA (1997) Organic geochemical proxies of paleoceanographic, paleolimnologic and paleoclimatic processes. *Org Geochem* 27:213–250
- Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int J Clim* 25:693–712
- Neukom R, Luterbacher J, Villalba R, Küttel M, Frank D, Jones PD, Grosjean M, Esper J, Lopez L, Wanner H (2010) Multi-centennial summer and winter precipitation variability in southern South America. *Geophys Res Lett* 37, doi:[10.1029/2010GL043680](https://doi.org/10.1029/2010GL043680)
- Neukom R, Luterbacher J, Villalba R, Küttel M, Frank D, Jones PD, Grosjean M, Wanner H, Aravena J-C, Black DE, Christie DA, D'Arrigo R, Lara A, Morales M, Soliz-Gamboa C, Srur A, Urrutia R, von Gunten L (2011) Multi-proxy summer and winter surface air temperature field reconstructions for southern South America covering the past centuries. *Clim Dyn* 37:35–51
- Ohlendorf C, Sturm M (2008) A modified method for biogenic silica determination. *J Paleolimnol* 39:137–142
- Ohlendorf C, Gebhardt C, Hahn A, Kliem P, Zolitschka B (2011) The PASADO core processing strategy - A proposed new protocol for sediment core treatment in multidisciplinary lake drilling projects. *Sediment Geol* 239:104–115
- R Development Core Team (2011) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna
- Rein B (2003) *In situ Reflektionspektroskopie und digitale Bildanalyse - Gewinnung hochauflösender Paläoumweltdaten mit fernerkundlichen Methoden*. Habilitationsschrift, Universität Mainz, Mainz
- Rein B, Sirocko F (2002) In situ reflectance spectroscopy – analysing techniques for high-resolution pigment logging in sediment cores. *Int J Earth Sci* 91:950–954
- Rothwell RG (2006) *New techniques in sediment core analysis*. Geological Society, special publication 267, London
- Sachs HM, Webb T III, Clark DR (1977) Paleocological transfer functions. *Annu Rev Earth Planet Sci* 5:159–178
- Schnurrenberger D, Russell J, Kelts K (2003) Classification of lacustrine sediments based on sedimentary components. *J Paleolimnol* 29:141–154
- Scholz CA (2001) Application of seismic sequence stratigraphy in lacustrine basins. In: Last WM, Smol JP (eds) *Tracking environmental change using lake sediments: vol 1, Basin analysis, coring and chronological techniques*. Kluwer Academic Publishers, Dordrecht, pp 7–22
- Sernageomin (2003) *Mapa Geológico de Chile: versión digital*. Servicio Nacional de Geología y Minería, Publicación Geológica Digital, Santiago, Chile

- Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52:591–611
- Telford R, Birks HJB (2009) Evaluation of transfer functions in spatially structured environments. *Quat Sci Rev* 28: 1309–1316
- Trachsel M, Grosjean M, Schnyder D, Kamenik C, Rein B (2010a) Scanning reflectance spectroscopy (380–730 nm): a novel method for quantitative high-resolution climate reconstructions from minerogenic lake sediments. *J Paleolimnol* 44:979–994
- Trachsel M, Grosjean M, Larocque-Tobler I, Schwikowski M, Blass A, Sturm M (2010b) Quantitative summer temperature reconstruction derived from a combined biogenic Si and chironomid record from varved sediments of Lake Silvaplana (south-eastern Swiss Alps) back to AD 1177. *Quat Sci Rev* 29:2719–2730
- Trenberth KE (1984) Some effects of finite sample size and persistence on meteorological statistics. *Mon Weath Rev* 112:2359–2368
- Villalba R, Grosjean M, Kiefer T (2009) Long-term multi-proxy climate reconstructions and dynamics in South America (LOTRED-SA): state of the art and perspectives. *Palaeogeogr Palaeoclimatol Palaeoecol* 281:175–179
- von Gunten L, Grosjean M, Rein B, Urrutia R, Appleby PG (2009a) A quantitative high-resolution summer temperature reconstruction based on sedimentary pigments from Laguna Aculeo, Central Chile, back to AD 850. *Holocene* 19(6):873–881
- von Gunten L, Grosjean M, Beer J, Grob P, Morales A, Urrutia R (2009b) Age modeling of young non-varved lake sediments: methods and limits. Examples from two lakes in Central Chile. *J Paleolimnol* 42:401–412
- von Gunten L, Grosjean M, Eggenberger U, Grob P, Urrutia R, Morales A (2009c) Pollution and eutrophication history AD 1800–2005 as recorded in sediments from five lakes in Central Chile. *Glob Planet Change* 68:198–208
- Walker IR, Levesque AJ, Cwynar LC, Lotter AF (1997) An expanded surface-water palaeotemperature inference model for use with fossil midges from eastern Canada. *J Paleolimnol* 18:165–178
- Wetzel RG (2001) *Limnology*. Elsevier Academic Press, London
- Zolitschka B, Mingram J, van der Gaast S, Jansen JHF, Naumann R (2001) Sediment logging techniques. In: Last WM, Smol JP (eds) *Tracking environmental change using lake sediments: Basin analysis, coring and chronological techniques*, vol 1. Kluwer Academic Publishers, Dordrecht, pp 137–154