A Review of Statistics in Palaeoenvironmental Research

Maarten BLAAUW[✉], J. Andrés CHRISTEN[✉], and
Marco Antonio AQUINO-LÓPEZ[✉]

Palaeoecologists use sequences of fossils within deposits from continents and oceans all over the world in order to produce time-series of past environmental dynamics over decades to millennia or longer. Such information can place current and future environmental change into context, for example by showing how climate, environments, ecosystems and humans interacted during past events, and by enabling verification of climate models through ‘hind-casting’ of such events. Through a meta-analysis and focused literature review of currently used statistical approaches in palaeoecological research, we highlight potential pitfalls and suggest ways forward to a fuller statistical understanding of the possibilities and limitations of palaeoecological studies. Statisticians or at least statistical reasoning should be involved in order to quantify uncertainties across the full analytical pipeline of obtaining, analysing and interpreting fossil time-series and could help optimizing the analytical decisions taken at all these steps.

Supplementary materials accompanying this paper appear online.

Key Words: Chronology; Climate change; Palaeoecology; Palaeoclimatology.

1. INTRODUCTION

Climate, environments and ecosystems (CEEs) are almost never constant but fluctuate over a wide range of time-scales. Current CEEs can be measured and/or manipulated in order to investigate the nature of their dynamics and interactions, and similarly their dynamics can be investigated using measurements of CEEs over time, which can range back to some seasons, years, decades or centuries. However, directly measured time-series cover only some areas of our planet, and no direct observations are available on time-scales from

This article is an invited review.

Maarten Blaauw (✉), Archaeology and Palaeoecology, School of Natural and Built Environment, Queen’s University Belfast, Belfast BT7 1NN, United Kingdom (E-mail: m.blaauw@qub.ac.uk). J. Andrés Christen, Centro de Investigación en Matemáticas, A.P. 402, 36000 Guanajuato, Gto, Mexico. Marco Antonio Aquino-López, Archaeology and Palaeoecology, School of Natural and Built Environment, Queen’s University Belfast, Belfast BT7 1NN, United Kingdom, and Arts and Humanities Institute, Maynooth University, Maynooth, Co. Kildare, Ireland.

© 2019 The Author(s)

Journal of Agricultural, Biological, and Environmental Statistics, Volume 25, Number 1, Pages 17–31
<https://doi.org/10.1007/s13253-019-00374-2>

multiple centuries to millennia or even millions of years. For such time-series, we thus need to rely on physical, chemical or biological *proxies* of CEEs (Jackson 2007). Such proxies can be found as partly fossilized remains (sub-fossils) within sites that have accumulated over long time-scales, and have been used to identify past extinctions, climate perturbations, human impacts and other major environmental events (e.g. Seddon et al. 2014).

The types of research questions approached using such fossil proxy studies vary, with some studies aiming to reconstruct past ecosystems (palaeoecology, e.g. forest composition from pollen in lakes or bogs; Prentice et al. 1987), whereas others aim to decipher climate dynamics (palaeoclimatology, e.g. regional temperature time-series from tree-ring widths; Mann et al. 1998), mainly manifested through its intermediate of the palaeo-environment. Some deposits can be seen as quite simple systems that store information about CEEs as they accumulate over millennia. Ice-sheets from polar or montane regions for example can contain acid peaks which can be interpreted as traces of volcanic eruptions, dust which indicates wind sources, CO₂ in air bubbles which indicate past atmospheric concentrations, and stable isotopes of oxygen and hydrogen within the ice itself which indicate sea-ice cover, precipitation sources and air temperatures (e.g. Steffensen et al. 2008). Other types of deposits form more indirect CEE proxy records since they form part of more intricate environments or ecosystems. For example, trees in temperate regions form annual rings and their width, density, colour and composition can serve as proxies for local climate and immediate environmental conditions such as fires or earthquakes (Swetnam and Betancourt 1990). Lakes and oceans contain sediment that accumulates over millennia and within which sub-fossilized pollen, remains of organisms (e.g. diatoms, foraminifera, and insects), isotopes and other proxy information of local environmental conditions can be found (e.g. indicators of lake levels which can be related to precipitation and temperature; Verschuren et al. 2009). Bogs accumulate over millennia and store pollen, macrofossils such as seeds and leaves, small animals, and isotopes within the resulting peat (e.g. Mauquoy and van Geel 2007). Additional types of fossil deposits containing proxy records include corals, speleothems, and other cave deposits. Since these sites accumulate over time, depth can be used as a proxy for time. By analysing the environmental proxies of a sequence of *d*'s along a core, time series of fossil proxy information are obtained.

A range of statistical issues has to be taken into account when analysing and interpreting these types of proxy records. Here we provide a short overview of the processes involved, assisted by a review of papers published in the Q1-ranked disciplinary journal *Quaternary Science Reviews*. We searched through all papers published in this journal through Volumes 183-181 and selected all that presented proxy records within cores sampled from sedimentary deposits ($n = 50$; other types of studies such as those based purely on modelling or geomorphology were not taken into account). For each paper (Table 1), we analysed any statistical techniques used in the four subsequent steps of (i) selecting which types of archives, sites, cores, and core depths to analyse for a number of proxies, (ii) interpreting the proxies in terms of CEEs, (iii) building chronologies, and (iv) putting the results into a wider spatiotemporal context. After this overview, we look into more detail at the latter two of these steps since the authors are most familiar with research in these areas. While for chronology-building, statistical methods have been widely adopted throughout the fossil proxy communities, this is much less the case for the step of placing studies into context.

Table 1. Meta-analysis of statistical approaches involved during the main steps in palaeoenvironmental research, based on 50 sediment-based studies published in volumes 183-181 (2018) of the journal *Quaternary Science Reviews*.

	Interpretation of proxy information	Chronology building	Placement into context
Approaches	Stable isotopes (12), pollen (11), loss-on-ignition (7), X-ray fluorescence (5), magnetism (5), grain size (5), stratigraphy (3), chronology (2)	Absolute dating using ^{14}C (32), optically stimulated luminescence (4), ^{210}Pb (4), or U/Th (3), Ar/Ar (1) Relative dating using tuning (12), palaeomagnetism (4), annual layer counting (3)	Literature review, comparing timings of events between studies
Statistical techniques	Cluster analysis (CONISS; 6), ordination (PCA/DCA/DNMS; 4), regression/correlation (4), periodicity/wavelet analysis (4), generalized additive modelling (2), transfer function (1), end-member modelling (1), change-point analysis (1), broken-stick modelling (1) None (26)	Classical age-depth models using linear interpolation (14), smooth spline (5), linear/polynomial regression or piece-wise linear regression (3), linear regression with added random noise (1) Bayesian age-depth models using Bacon (12) or Oxcal's P_sequence (7) Layer-counting (1) Not mentioned (7)	Visualisation of timings events among studies (35), conceptual model (8), change-point analysis (1)
Consideration of uncertainties	Uncertainties visualized (17) None (33)	Dates and age-model (19), dates only (7), model only (2) None (22)	None (in-text literature review only; 11) Uncertainty visualisation of none (23), some (13) or all (3) time-series in context

Numbers within brackets indicate amount of studies in the meta-analysis which used that approach. Since none of the studies mentioned any usage of statistics during the stage of sampling for sites, proxies, cores or core depths, no column was added for this step. For more details, see Supplementary Information/Appendix 1.

It is hoped that this overview will encourage a deeper focus on the development and use of statistics during investigations based on fossil proxy records, and thus develop a fuller understanding of the uncertainties involved at all steps.

2. OVERVIEW

As outlined above, here based on a meta-analysis and additional literature we will outline the steps taken to obtain, analyse and interpret fossil time series. The first step is that of selecting study sites and which proxies, cores and core depths to analyse.

2.1. SAMPLING STRATEGIES

Many research teams throughout the world are involved in producing fossil proxy time series. At first sight, the fossil proxy sites available within one of the major (though non-exhaustive) palaeo-databases, the US National Oceanic and Atmospheric Administration's National Climatic Data Center (NOAA-NCDC), appear to cover large parts of the Earth (Fig. 1). However, coverage is patchy, with large areas of for example the Pacific Ocean, Africa, the Amazon and Australia being poorly represented, and tree-ring records covering some temperate and montane regions only. Some of this poor representation is due to absence of sites (e.g. regions that have unsuitable climates to support ice sheets, trees, lakes or bogs), but at smaller scales there could be sampling biases related to distance from laboratories, roads or other infrastructure. Whereas some studies attempt to select sites that cover specific environmental gradients of interest (e.g. Charman et al. 2013), and all studies will have made decisions on how best to spend their limited funds and research time in order to obtain a proxy record that would answer their research questions, very few studies (none of the 50 in Table 1) mention the use of statistics to guide their choice of archive, site, core length and density of proxy analysis. The Quaternary research community could benefit from statistical analyses of sites sampled so far, for example to strengthen our understanding of which regions (e.g. in climate or ecosystem space) are currently under-sampled and would benefit from more research effort, or to identify fossil time-series that show high correlation with other regions or studies and would thus form valuable targets for further research efforts. Once a site is selected, statistical approaches to sampling design could be useful in deciding which depths to analyse (e.g. Christen and Buck 1998; Blaauw et al. 2018).

2.2. INTERPRETING PROXY RECORDS

The spectra of sub-fossils analysed within sediment cores cannot inform us directly about past CEEs. Instead, some kind of modelling (calibration) is required before they can be interpreted in terms of past CEE dynamics. These fossil calibrations are based on the assumption that the same natural laws and processes that operate now have always operated in the past (Uniformitarianism; Hutton 1788). In order to illustrate this, we can consider these laws as functions ($G(\cdot)$) which depend on the state of the environment (μ_e) at time t ($G(\mu_e(t))$). For example, from observing where current species of beetles are found, we can

Figure 1. Global map of palaeoenvironmental sites available within the databases of the National Oceanic and Atmospheric Administration's National Climatic Data Center. Red bars indicate records from trees, green circles lakes, blue stars ice and blue tildes oceans.

infer their preferences for distinct CEEs. If we then encounter the same species of beetles fossilised within a deposit, we can infer that when that deposit was formed, conditions must have been preferable for those beetles. We can thus calibrate our fossils into 'proxy' time-series of CEEs by substituting space (CEEs where species are currently found) for time (periods when fossil specimens of the same species are found within deposits). Even so, since current distributions of biota do not necessarily cover all possible climates and ecosystems, and human activity will have impacted current vegetation patterns, some ancient ecosystems will have no current analogue (Goring et al. 2016). This is equivalent to observing a function in sections with missing parts (Fig. 2).

Many conceptual or statistical methods are used to translate fossil proxies into proxy variables of interest. In the late 1800s and early 1900s, Axel Blytt and Rutger Sernander used the spatial distribution of vegetation types within regions of Scandinavia, in order to interpret peat layers found within Danish bogs as indicating distinct climatic periods in time (von Post 1946). Since the 1980s, spatial distributions of assemblages of living beetles have been used to reconstruct past temperatures through identifying 'mutual climatic ranges' (Bray et al. 2006). More modern quantitative space-to-time transformations include ordination (Hill and Gauch 1980; Birks et al. 2012), weighted-averaging partial-least-squares regression (WA-PLS; ter Braak and Juggins 1993) and transfer functions (Imbrie and Kipp 1971; Juggins et al. 2014); 10% of the papers in Table 1 use such methods. These methods have to be used with care, because many of the techniques were originally aimed at exploratory analysis rather than hypothesis testing, relationships can change over space and time (including through recent human impacts), multiple causal factors can interact, and ecosystems can show nonlinear responses which depend on initial conditions (Juggins 2013). The most commonly used technique (CONISS, 12%), a clustering algorithm based

Figure 2. Scheme of how forward and reverse models are used to link, from top to bottom, time, past environments, proxies, and measurements, through the function $G(\mu_e(t))$.

on squared Euclidian dissimilarity, produces zones to visualise the main trends in (mostly) pollen diagrams (Grimm 1987). The majority of papers within our meta-analysis (52%) use no statistical techniques to interpret their proxies, and only 34% use error bars or other visualisations of proxy uncertainties (even though techniques to do so have been available for decades; Maher 1972).

To summarize, even though a wide range of statistical techniques to interpret proxy time series has been available for decades, our meta-analysis suggests they are under-utilized within the palaeo community. Greater statistical rigour could advance interpretations of fossil time-series. For example, not all biota fossilize equally, causing a degree of bias in the final observed record, and there is also some evidence that fossils can move vertically within cores (Payne and Gehrels 2009). Perhaps statistical tools could be developed to model this. Null-models (Blaauw et al. 2010b; Barr 2017) could be useful in separating signal from noise in fossil time-series (e.g. Correa-Metrio et al. 2014; Turner et al. 2016).

2.3. BUILDING CHRONOLOGIES

On its own, each fossil time series provides only limited information about the dynamics of the environment at that site. Even if say the pollen record of an undated lake core could tell us how the nearby environment has evolved into its current state, only by putting the records onto a common time-scale can multiple records be compared to each other. Through such comparisons, inferences of spatiotemporal patterns can be made, which themselves are essential to test proposed cause-effect mechanisms. For example, a compilation of radio-carbon dates across Ireland indicates that a major societal collapse occurred at around the time of a climate event (Armit et al. 2014) and could thus be argued to show how vulnerable past societies were to climate change. However, since detailed analysis of the available data showed that the societal collapse happened some time *before* said climate event, in this case no such causal link could be established.

Our literature meta-analysis shows that a wide range of absolute and relative dating is used, and that most models that relate sediment core depths to ages are either based on classical methods such as linear interpolation (30%), smooth splines (10%) or regression (6%). Bayesian methods (see next section) include *Bacon* (24%, Blaauw and Christen 2011) and the *P_Sequence* within the Bayesian age-modelling software *Oxcal* (14%, Bronk Ramsey 2008). A minority (14%) do not mention which approaches were used, and most do not show uncertainties (44%). Therefore, even though the step of chronology-building relies more on statistical techniques than do the previous two steps, many studies ignore the associated uncertainties. Chronology-building, including potential areas of future statistical interest, will be discussed in more detail in Sect. 3.

2.4. PLACEMENT INTO CONTEXT

Many studies of past CEE dynamics place their study in context by comparison to other regional records (e.g. cores from nearby lakes), to potential forcing factors (e.g. solar irradiance), and/or to other types of time-series coeval with the study's record (e.g. archaeological records), in order to identify trends or events which then can be interpreted in terms of underlying forcing factors. Indeed, most papers in our analysis (70%, Table 1) plot their CEE record(s) together with other relevant time-series on the same calendar scale (often aided by shading or lines to highlight common dynamics). However, most of these figures do not show estimates of chronological or proxy uncertainties (59%), and 22% do not show figures and rely on in-text discussions of other studies only. A minority of studies present conceptual models to interpret their records (16%), and 1 study used a formal change-point analysis. Thus, much of this critical stage of placing records into context is done without considering the inherent statistical uncertainties. This will be discussed more in Sect. 4, outlining potential research lines including statistics.

3. BUILDING CHRONOLOGIES

Some records form annual layers, and by counting them secure chronologies can be obtained that sometimes span many millennia. Many trees in temperate areas form such layers (rings), several ice sheets in Greenland and elsewhere do, and even some lakes and ocean basins preserve annual bands (varves; De Geer 1912; Zolitschka et al. 2015) during sedimentation, e.g. through seasonal sediment inwash into glacial lakes or seasonal biological activity such as in Lake Suigetsu in Japan (Bronk Ramsey et al. 2012) or the Cariaco Basin offshore of Venezuela (Haug et al. 2001). Distinct patterns of thicker and thinner rings or varves can be used to infer environmental change but also to link chronologies within sites (e.g. within a single forest), across sites within a region (e.g. forests across Ireland or Northwestern Europe), or even between continents (De Geer 1921). Tree time-series can be extended backward in time by overlaying the rings of living trees with those of timbers from historical buildings or fossilised trunks in river beds or bogs.

Many records do not record such annual layering, and for these records other chronological methods have to be used. These are often based on measuring radioactive isotopes of

for example lead, carbon, or uranium/thorium. Radioactive isotopes are unstable and decay at a constant rate (half-life)—based on measuring the remaining concentrations of these isotopes within samples, absolute ages can be calculated with a degree of error. The age calculations depend on the processes involved. For example, ^{210}Pb has a half-life of 22.3 years and can be used to date sediments back to c. 100–200 years in time. Some of the ^{210}Pb in a sediment core will have rained in from the atmosphere, while the rest will have come from *in-situ* decay of parent isotopes. These contributions are unique to each site and must thus be estimated (Aquino-López et al. 2018). Most of the radiocarbon or ^{14}C in organic material will have come from atmospheric $^{14}\text{CO}_2$, which itself is derived from oxidation of ^{14}C particles produced through collisions of cosmic rays with particles high up in the atmosphere. Measurements of ^{14}C in tree-rings of independently obtained known calendar age have shown that atmospheric ^{14}C concentrations have varied over time. Therefore, a calibration curve is required in order to put ^{14}C dates on a calendar scale (Reimer et al. 2013).

Besides such *absolute* dates, there are also *relative* dates, or so-called isochronous events across sites. For example, a volcanic eruption could deposit an ash layer (tephra) across continents, and if it possesses a unique chemical fingerprint, then this layer can be used to align sites that contain the layer in question. If some of these sites carry chronological information about the age of this tephra (e.g. through layer counting in ice cores, or radiocarbon dating in lakes), then this age information can be transferred (or tie-pointed) to other sites that recorded the same tephra. Sometimes the same approach is used to produce site chronologies by tie-pointing assumed synchronous climate events – however often these layers cannot be identified securely, and subsequent climate interpretations can suffer from circular reasoning (Blaauw 2012).

Absolute dates are expensive and time-consuming to obtain, some core depths will not contain enough isotopes for reliable dating, and often cores will contain only few identifiable tephra layers. Therefore, often age-depth modelling has to be used to estimate the ages t of any depth d of a core, $t = F(d)$. Traditionally this is done by drawing some sort of linear interpolation or regression curve through the dated depths (Bennett 1994; Blaauw 2010; see Table 1). However, this is not straightforward, because some measurements have asymmetric errors (e.g. calibrated ^{14}C dates), often some dates appear to be outlying (e.g. through contamination, re-deposition of older sediment, or transport of material between depths in a core), and sediments are unlikely to have accumulated linearly over long stretches of depth or time. Moreover, sometimes unexpected scatter is found among multiple ^{14}C dates (Scott 2013; Christen and Perez 2009). As a result of the above problems, basic age-models likely underestimate chronological uncertainties (Blaauw et al. 2018).

Bayesian age-depth models that use MCMC simulations to simulate a more variable accumulation of sediment, such as *Bpeat* (Blaauw and Christen 2005), *Bchron* (Haslett and Parnell 2008), *Oxcal's P_Sequence* (Bronk Ramsey 2008), and *Bacon* (Blaauw and Christen 2011), have become quite popular among the Quaternary research community. *Bpeat* (Blaauw and Christen 2005) and its successor *Bacon* (Blaauw and Christen 2011) use piece-wise linear models, together with constraints on sedimentation times and how they vary between sections. The prior for sedimentation time was informed by analysing >200 lakes (Goring et al. 2012). The posterior information can be used to estimate the age

Figure 3. Scheme of translating depths to ages. Left panel shows the relative abundance across a core of *Sphagnum tenellum*, a macrofossil indicating very wet conditions in a bog. Middle panel shows the core's *Bacon* (Blaauw and Christen 2011) age-depth model (grey-scale indicates likely ages for depths; blue distributions are calibrated 14C dates). This age-depth model can be used to plot the fossil time-series on the calendar scale (right panel). Data come from Blaauw and Christen (2011).

distribution of any core depth, or even display the possible ages of proxy values (Fig. 3). *Bchron* models sediment accumulation through simulating jumps in both depth and time using Gamma distributions, while *OxCal*'s *P_Sequence* uses a Poisson process to simulate sediment accumulation. The Bayesian age-models listed here deal well with outlying dates, either through simulating the process of dates becoming outliers (Bronk Ramsey 2009) or by using student-t distributions with wide tails (Christen and Perez 2009; Blaauw et al. 2018). Because sediment can be assumed to have accumulated over time without reversals, the above Bayesian approaches enforce age-models to be monotonous. Such prior information on necessarily positive accumulation rates can result in much enhanced chronological precision, especially in cores with dates that overlap in time. For reviews on age-depth modelling, see Blaauw and Heegaard (2012), Parnell et al. (2011) and Trachsel and Telford (2016).

As outlined above, chronology-building has received considerable statistical attention over the past decade or so, and could be thus considered as a relatively well-developed component of numerical palaeoecology. In the future, process-based, forward models could be of use in producing more geologically realistic age-depth models. Such models could start to make use of real-world information such as how the shapes of lake basins shape sedimentation rates (Bennett and Buck 2016). Using physical climate models, sedimentation could be simulated mechanistically, e.g. through modelling how different climate-dependent vegetation types and other landscape factors affect erosion (Merritt et al. 2003) and thus sediment input into lakes. More thorough approaches to estimating time-series from annually banded records would also be welcome (see Comboul et al. 2014; Boers et al. 2017).

4. PLACING FOSSIL TIME-SERIES INTO CONTEXT

As discussed in Sect. 2, by far most studies place their results into context through plotting their environmental inferences, or raw proxy data, against calendar age, together with a number of nearby or otherwise relevant time-series (Table 1), after which any shared

events or other dynamics are interpreted. This raises questions regarding (i) spatial scales, (ii) impacts of the chronological quality of individual time-series, and (iii) potential mechanistic links between the time-series.

Most of the cores containing the fossil time-series outlined above are generally less than 10 cm wide, and thus the reconstructions are essentially based on one-dimensional time-series. These records are then extrapolated to spatial scales at which the records are thought to respond. To start at the smallest scales, even within a core there is some degree of horizontal variability (Fig. 2 of Wörmer et al. 2014). Within bogs, twigs, seeds and other macrofossils can be assumed to have derived from very close to where a core was sampled and are thus used to reconstruct highly local hydrological changes (e.g. Blaauw and Mauquoy 2012). Diatom records within lake sediment indicate ecological dynamics within a lake (or local conditions in specific areas of a lake; Rimet et al. 2016). On slightly larger spatial scales, pollen within a bog or lake is often assumed to have originated both from within or just around the site as well as from the surrounding ecosystems, and can thus be used to reconstruct the past dynamics of forests, agriculture and other vegetation from a wider area around the site. The source area of pollen is a topic of active modelling research (e.g. Sugita 2007; Theuerkauf and Couwenberg 2016)—for example, an increase in the amount of pine pollen within a core could have been caused by a growing pine population in the wider region, but also by a few nearby pine trees starting to produce pollen as they reached maturity. Similarly, a narrow ring found within one tree could indicate a purely local event such as disease, but if replicated within a site, between sites and/or across continents would imply a widely felt event and thus more likely caused by a large-scale factor such as climate. On the largest spatial scales, events in many fossil records are compared and/or aligned to those found in distant Greenland ice cores (e.g. Itambi et al. 2009) or global stacks of ocean sediment records (e.g. Lisiecki and Raymo 2005).

In the quest for identifying large-scale events from multiple fossil time-series across regions or continents (e.g. to test how well such events can be reproduced by climate models), a series of important decisions has to be made. Some fossil records are much better dated than others. Blois et al. (2011) compiled well-dated and poorly dated pollen records within a region, and then aligned pollen events shared between records in order to enhance the chronologies of the poorly dated records. Similarly, layer-counted Greenland ice cores form precisely dated records of climate events and are used as reference frames to which more poorly dated archives on other continents are linked (e.g. Haesaerts et al. 2009). However, even closely spaced ice cores show differences in event timings, erroneous tie-points could be chosen, and circular reasoning could mask real timing differences of events within or across regions (Blaauw 2012). Chronological uncertainties of aligned records have been modelled using Gaussian processes (Heaton et al. 2013), but more work remains to be done on quantifying our confidence in chosen tie-points or in modelling accumulation between tie-points.

Most of the commonly used approaches to interpret proxies are ‘reverse’ models since they start with the outcomes (fossils) in order to infer their causes (past CEEs); this is equivalent to observing $G(\mu_e)$ and inferring μ_e from it (Fig. 2). This introduces several problems, such as poor quantifications of uncertainties, dangers of correlations with non-causal variables, no obvious ways to combine multiple records, and difficulties in estimating

the smoothness in changes over time and in combining multiple records. Some of the above problems could be avoided by building simple or more complicated process-based or forward models, where changes in climate parameters steer changes in environments and ecosystems, which are then registered as changes in fossil spectra—in other words, defining the function $G(\mu_e)$. Blaauw et al. (2010b) developed a basic conceptual model where m fossil types have individual preferences and tolerances for n environmental factors. As these factors vary over time, conditions become more or less favourable for each of the fossil types, causing changes in their relative abundances. Simple conceptual models (Jackson 2012) could also enhance our qualitative understanding of the possibilities and limitations of proxy data. Building on Haslett et al. (2006), Parnell et al. (2016) produced realistic yet computationally demanding forward Bayesian models, creating iterative ‘climate histories’ through linking modular components with proxies (climate, modern vegetation distribution, and dated sediment cores with pollen counts) and assuming some smoothing over time. Other Bayesian compilation approaches include Holmström et al. (2005) and Ilvonen et al. (2016).

It is important to note that most Quaternary research questions do not involve absolute proxy values, but rather their changes over time and the speeds of these changes. This introduces a dependence between neighbouring proxy measurements, as well as on the age-depth models. Some work has been done on identifying changes, including change-point analysis (Blois et al. 2011), probabilistic identification of proxy events (Blaauw et al. 2010a), and directional analysis (Holmström 2010). Some past-climate studies search for periodicities using wavelet analysis (Debret et al. 2007; see also Table 1). Such techniques are sensitive to uneven sample spacing, and to our knowledge, the impacts of chronological uncertainties on wavelet analysis have not yet been investigated.

5. DISCUSSION

As outlined above, a wide range of statistical techniques has been developed and used in order to better reconstruct past climates, environments and ecosystems from fossil records. However, many of the statistical approaches to proxy interpretation rely on exploratory analysis rather than hypothesis-testing. Some of these techniques were developed by the Quaternary community itself, whereas others were done through collaborations with the statistics community. Recent exciting collaborative developments in past CEE dynamics include tipping points (e.g. Scheffer et al. 2009), forward models (Parnell et al. 2016) and artificial intelligence (Anderson et al. 2014). One of the main remaining problems however is that, even though many approaches to quantify uncertainties are now available, the Quaternary community continues to largely ignore them at all steps (Table 1). Also, not all uncertainties can be quantified (Jackson 2012). We would thus welcome continued collaboration between the palaeo and statistics communities (e.g. <https://www.pastearth.net/>, <http://sites.nd.edu/paleonproject/>, <https://www.earthcube.org>). Most of all, efforts should be strengthened to educate the palaeo community about the best ways to dealing with the large uncertainties inherent in palaeoenvironmental studies. More effort should also be made to communicate these uncertainties to climate modellers and other research communities that make use of fossil-derived CEE reconstructions (Jackson 2012).

ACKNOWLEDGEMENTS

We thank the Associate Editor and reviewer Gavin Simpson for their comments, which helped focus this manuscript. JAC is partially funded by CONACYT CB-2016-01-284451, RDECOMM and ONRG grants. MAL was funded by a CONACYT Doctoral Grant.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

[Received November 2017. Accepted August 2019. Published Online September 2019.]

REFERENCES

- Armit, I., Swindles, G. T., Becker, K., Plunkett, G., and Blaauw, M. (2014), "Rapid climate change did not cause population collapse at the end of the European Bronze Age," *PNAS*, 111, 17045–17049.
- Anderson, K., Bradley, E., Rassbach de Vesine, L., Zreda, M., and Zweck, C. (2014), "Forensic reasoning about paleoclimatology: creating a system that works," *Advances in Cognitive Systems*, 3, 221–240.
- Aquino-López, M., Blaauw, M., Christen, J. A., and Sanderson, N. K. (2018), "Bayesian analysis of ^{210}Pb dating," *Journal of Agricultural, Biological and Environmental Statistics*, <https://doi.org/10.1007/s13253-018-0328-7>.
- Barr, W. A. (2017), "Signal or noise? A null model method for evaluating the significance of turnover pulses," *Paleobiology*, 43, 656–666.
- Bennett, K. D. (1994), "Confidence intervals for age estimates and deposition times in late-Quaternary sediment sequences," *The Holocene*, 4, 337–348.
- Bennett, K. D., and Buck, C. E. (2016), "Interpretation of lake sediment accumulation rates," *The Holocene*, 26, 1092–1102.
- Birks, H. J. B., Juggins, S., Lotter, A., and Smol, J. P. (2012), *Tracking Environmental Change Using Lake Sediments Vol. 5: Data Handling and Statistical Techniques*, Springer, Dordrecht, 673 pp.
- Blaauw, M., Christen, J. A. (2005), "Radiocarbon peat chronologies and environmental change," *Applied Statistics*, 54, 805–816.
- Blaauw, M. (2010), "Methods and code for 'classical' age-modelling of radiocarbon sequences," *Quaternary Geochronology*, 5, 512–518.
- Blaauw, M., Wohlfarth, B., Christen, J. A., Ampel, L., Veres, D., Hughen, K. A., Preusser, F., Svensson, A. (2010a), "Were last glacial climate events simultaneous between Greenland and France? A quantitative comparison using non-tuned chronologies," *Journal of Quaternary Science*, 25, 387–394.
- Blaauw, M., Bennett, K. D., and Christen, J. A. (2010b), "Random walk simulations of fossil proxy data," *The Holocene*, 20, 645–649.
- Blaauw, M., and Christen, J. A. (2011), "Flexible paleoclimate age-depth models using an autoregressive gamma process," *Bayesian Analysis*, 6, 457–474.
- Blaauw, M. (2012), "Out of tune: the dangers of aligning proxy archives," *Quaternary Science Reviews*, 36, 38–49.
- Blaauw, M., and Heegaard, E. (2012), "Estimation of age-depth relationships," in Birks, H. J. B., Juggins, S., Lotter, A., Smol, J. P. (editors), *Tracking Environmental Change Using Lake Sediments Vol. 5: Data Handling and Statistical Techniques*, pp 379–413.
- Blaauw, M., and Mauquoy, D. (2012), "Signal and variability within a Holocene peat bog - chronological uncertainties of pollen, macrofossil and fungal proxies," *Review of Palaeobotany and Palynology*, 186, 5–15.
- Blaauw, M., Christen, J. A., Bennett, K. D., and Reimer, P. J. (2018), "Double the date and go for Bayes - impacts of model choice, dating density and quality on chronologies," *Quaternary Science Reviews*, 188, 58–66.

- Blois, J. L., Williams, J. W., Grimm, E. C., Jackson, S. T., and Graham, R. W. (2011), "A methodological framework for assessing and reducing temporal uncertainty in paleovegetation mapping from late-Quaternary pollen records," *Quaternary Science Reviews*, 30, 1926–1939.
- Boers, N., Goswami, B., and Ghil, M. (2017), "A complete representation of uncertainties in layer-counted paleoclimatic archives," *Climate of the Past*, 13, 1169–1180.
- Bray, P. J., Blockley, S. P. E., Coope, G. R., Dadswell, L. F., Elias, S. A., Lowe, J. J., and Pollard, A. M. (2006), "Refining mutual climatic range (MCR) quantitative estimates of palaeotemperature using ubiquity analysis," *Quaternary Science Reviews*, 25, 1865–1876.
- Bronk Ramsey, C. (2008), "Deposition models for chronological records," *Quaternary Science Reviews*, 27, 42–60.
- (2009), "Dealing with outliers and offsets in radiocarbon dating," *Radiocarbon*, 51, 1023–1045.
- Bronk Ramsey, C., Staff, R. A., Bryant, C. L., Brock, F., Kitagawa, H., and van der Plicht, J. (2012), "A complete terrestrial radiocarbon record for 11.2 to 52.8 kyr B.P.," *Science*, 338, 370–374.
- Charman, D. J., Beilman, D. W., Blaauw, M., Booth, R. K., Brewer, S., Chambers, F. M., Christen, J. A., Gallego-Sala, A., Harrison, S. P., Hughes, P. D. M., Jackson, S. T., Korhola, A., Mauquoy, D., Mitchell, F. J. G., Prentice, I. C., van der Linden, M., De Vleeschouwer, F., Yu, Z. C., Alm, J., Bauer, I. E., Corish, Y. M. C., Garnau, M., Hohl, Huang, Y., V., Karofeld, E., Le Roux, G., Loisel, J., Moschen, R., Nichols, J. E., Nieminen, T. M., MacDonald, G. M., Phadtare, N. R., Rausch, N., Sillasoo, Ü., Swindles, G. T., Tuittila, E. -S., Ukonmaanaho, L., Väliranta, M., van Bellen, S., van Geel, B., Vitt, D. H., Zhao, Y. (2013), "Climate-related changes in peatland carbon accumulation during the last millennium," *Biogeosciences*, 10, 929–944.
- Christen, J. A. and Buck, C. E. (1998), "Sample selection in radiocarbon dating," *Applied Statistics*, 47, 543–557.
- Christen, J. A., and Perez, S. E. (2009), "A new robust statistical model for radiocarbon data," *Radiocarbon*, 51, 1047–1059.
- Comboul, M., Emile-Geay, J., Evans, M. N., Mirnategui, N., Cobb, K. M., and Thompson, D. M. (2014), "A probabilistic model of chronological errors in layer-counted climate proxies: applications to annually banded coral archives," *Climate of the Past*, 10, 825–841.
- Correa-Metrio, A., Meave, J. A., Lozano-García, S., and Bush, M. B. (2014), "Environmental determinism and neutrality in vegetation at millennial time scales," *Journal of Vegetation Science*, 25, 627–635.
- Debret, M., Bout-Roumazeilles, V., Grousset, F., Desmet, M., McManus, J. F., Massei, N., Sebag, D., Petit, J. -R., Copard, Y., and Trentesaux, A. (2007), "The origin of the 1500-year climate cycles in Holocene North-Atlantic records," *Climate of the Past*, 3, 569–575.
- De Geer, G. (1912), "Geochronologie der letzten 12000 Jahre," *Geologische Rundschau*, 3, 457–471.
- (1921), "Correlation of late glacial annual clay-varves in North America with the Swedish time scale," *Geologiska Föreningens i Stockholm Förhandlingar*, 43, 70–73.
- Grimm, E. C. (1987), "CONISS: a FORTRAN 77 program for stratigraphically constrained cluster analysis by the method of incremental sum of squares," *Computers & Geosciences*, 13, 13–35.
- Goring, S., Williams, J. W., Blois, J. L., Jackson, S. T., Paciorek, C. J., Booth, R. K., Marlon, J. R., Blaauw, M., and Christen, J. A. (2012), "Deposition times in the northeastern United States during the Holocene: establishing valid priors for Bayesian age models," *Quaternary Science Reviews*, 48, 54–60.
- Goring, S. J., Mladenoff, D. J., Cogbill, C. V., Record, S., Paciorek, C. J., Jackson, S. T., Dietze, M. C., Dawson, A., Matthes, J. J., McLachlan, J. A., and Williams, J. W. (2016), "Novel and lost forests in the upper midwestern United States, from new estimates of settlement-era composition, stem density, and biomass," *PLoS ONE* 11(12), e0151935.
- Haesaerts, P., Borziac, I., Chekha, V. P., Chirica, V., Damblon, F., Drozdov, N. I., Orlova, L. A., Pirson, S., and van der Plicht, J. (2009), "Climatic signature and radiocarbon chronology of middle and late pleniglacial loess from Eurasia: comparison with the marine and Greenland records," *Radiocarbon*, 51, 301–318.
- Haslett, J. and Parnell, A. C. (2008). A simple monotone process with application to radiocarbon-dated depth chronologies. *Journal of the Royal Statistical Society, Series C*, 57, 399–418.
- Haslett, J., Whitley, M., Bhattacharya, S., Mitchell, F. J. G., Allen, J. R. M., Huntley, B., Wilson, S. P., and Salter-Townshend, M. (2006), "Bayesian palaeoclimate reconstruction," *Journal of the Royal Statistical Society, Series A*, 169, 395–438.

- Haug, G. H., Hughen, K. A., Sigman, D. M., Peterson, L. C., and Rohl, U. (2001), "Southward migration of the Intertropical Convergence Zone through the Holocene," *Science*, 293, 1304–1308.
- Heaton, T. J., Bard, E., and Hughen, K. A. (2013), "Elastic tie-pointing—transferring chronologies between records via a Gaussian process," *Radiocarbon*, 55, 1975–1997.
- Hill, M. O., and Gauch, H. G. (1980), "Detrended Correspondence Analysis: An Improved Ordination Technique," *Vegetatio*, 42, 47–58.
- Holmström, L. (2010), "BSiZer," *Wiley Interdisciplinary Reviews: Computational Statistics*, 2, 526–534.
- Holmström, L., Ilvonen, L., Seppä, H., and Veski, S. (2005) "A Bayesian spatiotemporal model for reconstructing climate from multiple pollen records," *The Annals of Applied Statistics*, 9, 1194–1225.
- Hutton, J. (1788), "Theory of the earth," *Transactions of the Royal Society of Edinburgh*, 1, 209–304.
- Ilvonen, L., Holmström, L., Seppä, H., and Veski, S. (2016), "A Bayesian multinomial regression model for palaeoclimate reconstruction with time uncertainty," *Environmetrics*, 27, 409–422.
- Imbrie, J., Kipp, N. G. (1971), "A new micropaleontological method for quantitative paleoclimatology: application to a late Pleistocene Caribbean core," in Turekian, K.K. (Ed.), *The Late Cenozoic Glacial Ages*, Yale University Press, New Haven, pp. 71–181.
- Itambi, A. C., von Döbenek, T., Mulitza, S., Bickert, T., and Heslop, D. (2009), "Millennial-scale northwest African droughts related to Heinrich events and Dansgaard-Oeschger cycles: evidence in marine sediments from offshore Senegal," *Paleoceanography*, 24, PA1205.
- Jackson, S. T. (2007), "Looking forward from the past: history, ecology, and conservation," *Frontiers in Ecology and the Environment*, 5, 455.
- (2012), "Representation of flora and vegetation in Quaternary fossil assemblages: known and unknown knowns and unknowns," *Quaternary Science Reviews*, 49, 1–15.
- Juggins, S. (2013), "Quantitative reconstructions in palaeolimnology: new paradigm or sick science?," *Quaternary Science Reviews*, 64, 20–32.
- Juggins, S., Simpson, G. L., and Telford, R. J. (2014), "Taxon selection using statistical learning techniques to improve transfer function prediction," *The Holocene*, 25, 130–136.
- Lisiecki, L. E., and Raymo, M. E. (2005), "A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records," *Paleoceanography*, 20, PA1003.
- Maher, L. J., Jr. (1972), "Nomograms for computing 0.95 confidence limits of pollen data," *Review of Palaeobotany and Palynology*, 13, 85–93.
- Mann, M. E., Bradley, R. S., Hughes, M. K., 1998. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, 392, 779–787.
- Mauquoy, D., van Geel, B., 2007. *Plant macrofossil methods and studies: Mire and Peat Macros*. In Elias, S.A. (Ed), *Encyclopedia of Quaternary Science*, Elsevier, pp 2315–2336.
- Merritt, W. S., Letcher, R. A., Jakeman, A.J., 2003. A review of erosion and sediment transport models. *Environmental Modelling & Software* 18, 761–799.
- Parnell, A. C., Buck, C. E., and Doan, T. K. (2011), "A review of statistical chronology models for high-resolution, proxy-based Holocene palaeoenvironmental reconstruction," *Quaternary Science Reviews*, 30, 2948–2960.
- Parnell, A. C., Haslett, J., Sweeney, J., Doan, T. K., Allen, J. R. M., and Huntley, B. (2016), "Joint palaeoclimate reconstruction from pollen data via forward models and climate histories," *Quaternary Science Reviews*, 151, 111–126.
- Payne, R., and Gehrels, M. (2009), "The formation of tephra layers in peatlands: an experimental approach," *Catena*, 81, 12–23.
- Prentice, I. C., Berglund, B. E., Olsson, T., 1987. Quantitative forest-composition sensing characteristics of pollen samples from Swedish lakes. *Boreas*, 16, 43–54.
- Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Bronk Ramsey, C., Grootes, P. M., Guilderson, T. P., Hafflidason, H., Hajdas, I., Hatté, C., Heaton, T. J., Hoffmann, D. L., Hogg, A. G., Hughen, K. A., Kaiser, K. F., Kromer, B., Manning, S. W., Niu, M., Reimer, R. W., Richards, D. A., Scott, E. M., Southon, J. R., Staff,

- R. A., Turney, C. S. M., and van der Plicht, J. (2013), "IntCal13 and Marine13 radiocarbon age calibration curves 0-50,000 years cal BP," *Radiocarbon*, 55, 1869–1887.
- Rimet, F., Bouchez, A., and Tapolczai, K. (2016), "Spatial heterogeneity of littoral benthic diatoms in a large lake: monitoring implications," *Hydrobiologia*, 771, 179–193.
- Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., and Sugihara, G. (2009), "Early-warning signals for critical transitions," *Nature*, 461, 53–59.
- Scott, E. M. (2013), "Radiocarbon dating – Sources of error," In Elias, S.E. (Ed.), *Encyclopedia of Quaternary Science* (2nd edition), p 324–328.
- Seddon, A. W., Mackay, A. W., Baker, A. G., Birks, H. J., Breman, E., Buck, C. E., Ellis, E. C., Froyd, C. A., Gill, J. L., Gillson, L., Johnson, E. A., Jones, V. J., Juggins, S., Macias-Fauria, M., Mills, K., Morris, J. L., Nogués-Bravo, D., Punyasena, S. W., Roland, T. P., Tanentzap, A. J., Willis, K. J., Aberhan, M., Asperen, E. N., Austin, W. E., Battarbee, R. W., Bhagwat, S., Belanger, C. L., Bennett, K. D., Birks, H. H., Bronk Ramsey, C., Brooks, S. J., Bruyn, M., Butler, P. G., Chambers, F. M., Clarke, S. J., Davies, A. L., Dearing, J. A., Ezard, T. H., Feurdean, A., Flower, R. J., Gell, P., Hausmann, S., Hogan, E. J., Hopkins, M. J., Jeffers, E. S., Korhola, A. A., Marchant, R., Kiefer, T., Lamentowicz, M., Larocque-Tobler, I., López-Merino, L., Liow, L. H., McGowan, S., Miller, J. H., Montoya, E., Morton, O., Nogué, S., Onoufriou, C., Boush, L. P., Rodriguez-Sanchez, F., Rose, N. L., Sayer, C. D., Shaw, H. E., Payne, R., Simpson, G., Sohar, K., Whitehouse, N. J., Williams, J. W. and Witkowski, A. (2014), "Looking forward through the past: identification of 50 priority research questions in palaeoecology," *Journal of Ecology*, 102, 256–267.
- Swetnam, T. W., Betancourt, J. L. (1990). Fire-Southern Oscillation relations in the Southwestern United States. *Science*, 249, 1017–1020.
- Steffensen, J. P., Andersen, K. K., Bigler, M., Clausen, H. B., Dahl-Jensen, D., Fischer, H., Goto-Azuma, K., Hansson, M., Johnson, S.J., Jouzel, J., Masson-Delmotte, V., Popp, T., Rasmussen, S. O., Röthlisberger, R., Ruth, U., Stauffer, B., Siggaard-Andersen, M. -L., Sveinbjörnsdóttir, Á. E., Svensson, A., White, J. W. C., 2008. High-resolution Greenland ice core data show abrupt climate change happens in a few years. *Science*, 321, 680–684.
- Sugita, S. (2007), "Theory of quantitative reconstruction of vegetation II: all you need is LOVE," *The Holocene*, 17, 243–257.
- ter Braak, C. J. F., and Juggins, S. (1993), "Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages," *Hydrobiologia*, 269, 485–502.
- Theuerkauf, M., and Couwenberg, J. (2016), "The extended downscaling approach: A new R-tool for pollen-based reconstruction of vegetation patterns," *The Holocene*, 27, 1252–1258.
- Trachsel, M., and Telford, R. J. (2016), "All age-depth models are wrong, but are getting better," *The Holocene*, 27, 860–869.
- Turner, T. E., Swindles, G. T., Charman, D. J., Langdon, P. G., Morris, P. J., Booth, R. K., Parry, L. E., and Nichols, J. E. (2016), "Solar cycles or random processes? Evaluating solar variability in Holocene climate records," *Nature Scientific Reports*, 6, 23961.
- Verschuren, D., Sinninghe Damsté, J.S., Moernout, J., Kristen, I., Blaauw, M., Fagot, M., Haug, G. H., CHAL-LACEA project members, 2009. Half-precessional dynamics of monsoon rainfall near the East African equator. *Nature*, 462, 637–641.
- von Post, L. (1946), "The prospect for pollen analysis in the study of Earth's climatic history," *New Phytologist*, 45, 193–217.
- Wörmer, L., Elvert, M., Fuchser, J., Lipp, J. S., Buttigieg, J. S., Zabel, M., and Hinrichs, K. U. (2014), "Ultra-high-resolution paleoenvironmental records via direct laser-based analysis of lipid biomarkers in sediment core samples," *PNAS*, 111, 15669–15674.
- Zolitschka, B., Francus, P., Ojala, A. E. K., and Schimmelmann, A. (2015), "Varves in lake sediments: a review," *Quaternary Science Reviews*, 117, 1–41.