Age Models in Large Scale Synthesis

Well constructed chronologies are critical to reliable paleoecological inference. Analysis that relies strongly on chronologies must take into account the methods' inherent assumptions and the parameters used in the development of the chronologies. A first step in large-scale paleodata syntheses is to review and, as needed, revise chronologies. The selection of a single age-modeling software helps reduce variability among chronologies due to differing assumptions among models, but the choice of age model parameters, and assumptions about synthesis data are equally important. This paper uses pollen records and associated chronological controls from the Neotoma Paleoecological Database (http://neotomadb.org) to examine the biases and assumptions made in chronology construction on an aggregate dataset. We examine the challenges of missing uncertainty estimates in 210Pb dates, and the use of zero-length hiatuses to account for changing accumulation rates with depth. We also use this exercise to provide prescriptive guidance for chronology construction in the large-scale re-analysis of records.

A third of pollen age models in Neotoma have chronologies expressed in radiocarbon years only. The analysis of newly constructed age models indicates broader patterns that may change our interpretations of chronologies. Direct calibration of interpolated ages results in differences of 31 -- 470 cal yrs BP when compared to newly constructed age models. 210Pb dates show a linear relationship between 210Pb ages and Bacon calibrated ages, but Bacon calculated uncertainties are XXX higher than the reported 210Pb uncertainty, increasing exponentially with age. Accumulation rates show a pattern similar to other studies, with an inflection point at approximately 200 years before present. Establishing biostratigraphic markers for the EuroAmerican settlement era, and using a hiatus at the point of inflection for sediment accumulation at this time period results in age estimates that are, on average 44 years younger in the new Bacon models.

Parameters for the best-fit Bacon model show some general trends. Best-fit model section thickness increases with absolute model age, and a framework in which each core is reconstructed multiple times with a grid of parameters for section thickness, memory and accumulation provides a basis for broad-scale reconstruction that accounts for regional and local-scale variability in accumulation patterns for sediment over time.

This paper uses these reconstructions as a basis to solve several issues with chronology construction, and to point to outstanding issues in age-model reconstruction that can be solved with new methodological approaches. In particular, choices about model parameters, assumptions about sediment accumulation processes and product representation must be addressed for us to move forward as a community.

Table of Contents

# Introduction

Advances in computing have improved our ability to store and process large complex data sets. These advances have motivated scientists to address hypotheses that rely on access to paleoecological data sets. Paleoecological databases streamline this process by providing the means to store, curate, and query structured data across specified spatial and temporal bounds. These databases also provide the opportunity to more easily undertake synthetic analysis and extend our current understanding of ecological and climatic patterns beyond modern observational data (Brewer et al., 2012; Uhen et al., 2013). Most records stored in paleoecological databases are associated with the chronology or ages of the constituent samples; these chronologies are usually the result of the application of an age-depth model. Given this relationship between sample age and the underlying chronological model, the chronology assigned to a sedimentary archive can strongly affect the interpretation of a record, and indeed, geochronology is central to much of our understanding of the evolution of the earth system [harrison2016geochronology]. Paleoecologists have been aware of the varying methodologies and limitations in age modeling capabilities for some time, and the issues caused by these limitations are well noted (Grimm et al., 2009; Liu et al., 2012).

Recent efforts towards methodological consistency in synthesis analysis have focused on standardizing age models for records within a database (Giesecke et al., 2014) and establishing regional benchmarks (Blois et al., 2011; Flantua et al., 2016). Age model standardization in Europe relied on the Clam software (Blaauw, 2010), as did efforts with the North American Pollen Database (Blois et al., 2011). The paleoecological community has also begun to recognize value in the explicit accounting of chronological uncertainty in processes (**???**). Recent developments in age modelling include several Bayesian approaches that allow for the quantification of age uncertainty in sediment records (*e.g.*, Blaauw and Christen, 2011; Parnell and Parnell, 2016; Ramsey and Lee, 2013). The need for databases to reliably store uncertainty data and the generating models arises from the new imperative to propogate uncertainty in paleoecological analyses (**???**). The desire to store these models presents new challenges to database managers and developers by increasing data structure complexity and storage requirements.

Despite recent methodological advances in age modeling, many paleoecological records preceding these developments include extant age models. As new and extant methods diverge, the limitations of age models developed using extant methods become more problematic. Blaauw (2012) noted that subjectivity and flexibility in age-depth modeling may introduce additional "researcher degrees of freedom" that can lead to greater rates of false-positive relationships in paleoecological research. However, in practice it is difficult to identify and quantify these effects on existing age-models, especially when only a limited description of the methodology is available. The reliability of extant age models may be questionable for several reasons: 1) a lack robust of uncertainty quantification, 2) the construction of many records was carried out in radiocarbon years, or using outdated radiocarbon calibration curves, and 3) underlying inconsistency in reconstruction methodologies resulting in potential biases. However, the value of extant age models cannot be ignored; it is the inherent value of these records that has motivated standardization efforts.

Reliable sediment age estimates require accurate dating of high quality stratigraphic control points, well supported conversion from radiometric (often radiocarbon) to calendar years, robust and reliable models that mehcanistically or empirically represent the underlying process of sediment deposition and accumulation, and clear quantification of uncertainty that is able to take into account measurement and process uncertainty.

Chronology construction is dependant on the availability of dated stratigraphic control points from a sediment core. These control points may be geochronological (dated material), geostratigraphic (*e.g.*, the "modern" core top), or biostratigraphic (changes in pollen assemblages associated with dated changes on the landscape). Geochronological ages are often uncertain due to analytical errors in laboratory radiocarbon dating processes (Ward and Wilson, 1978), the calibration of certain age types (*e.g.*, radiocarbon to calendar year calibration; Reimer et al., 2013), and potential differences between the ages of dated material and the age of deposition of the sediment matrix (Blois et al., 2011). Geostratigrahic markers may have fewer sources of uncertainty, the core top age is assumed to be the year of sampling and tephras are often widespread which provides broader support for dated ages, however, sedimentary mixing may introduce uncertainty. Biostratigraphic control points are determined by the examination (usually visual) of changes in biological proxy assemblages throughout a core. However, time series of proxy data are often noisy, and in practice identifying changes in composition is both difficult and subjective (Dawson et al., 2016). There remains uncertainty in the identification of compositional shifts related to the chronological/stratigraphic sampling density (Liu et al., 2012), and, in the case of landscape-scale phenomena such as the mid-Holocene Hemlock decline (Bennett and Fuller, 2002; Davis, 1981), the need to assign temporal bounds to the landscape-scale phenomenon that caused the compositional shift.

Chronologies are approximations of physical processes that lead to variablility in sediment accumulation rates within depositional basins. While methods such as Bacon (Blaauw and Christen, 2011) include an autocorrelation term intended to reflect physical processes, they do not explicitly attempt to model the physical process of sedimentation. Theoretical work extends our understanding of sedimentation within lake basins (Bennett and Buck, 2016), describing broad relationships between basin shape and long term patterns of accumulation that can improve the mechanistic characterization of sedimentation processes, but broad-scale empirical support would require links between paleoecological resources and bathymetric data that is yet unavailable. Varved lake records can act as controls to validate model precision and accuracy (Trachsel and Telford, 2017). Recent work compiling published varved records (Ojala et al., 2012) provides further insight into chronology accuracy. However varved records are often spatially clustered, and are the result of a range of physical, biological and climatic conditions that lead to the possibility that these systems may not reflect broader controls on lacustrine systems without rhythmites, varves or other laminated sediments.

The development of the first radiocarbon calibration curve (IntCal98; Hughen et al., 1998) allowed for the association of radiocarbon years with calendar years, based on material with known ages and associated 14C dates. Continued improvement of the calibration curve has resulted in improved chronology development, but chronologies are rarely updated as new curves are generated (Grimm et al., 2014). Offset and uncertainty are influenced by the method used to re-calibrating legacy models using only radiocarbon years. Comparisons between re-calibrated age models against directly re-calibrated dates highlight two issues: the possibility of systematic offsets resulting in biases in age estimates, and the need for direct intervention and systematic rules-based chronology construction in synthesis work.

Differences between age models and direct recalibration.

**Figure 1**. *Diagram of two methods for generating calibrated dates from radiocarbon age-models. (a) Each age (in radiocarbon years) is calibrated directly (red arrows) from an existing model (black dashed line) or (b) the age models is reconstructed using linear interpolatation (red dashed line) from calibrated radiocarbon ages (red line). Negative values indicate that recalibration of interpolated 14C dates followed by linear interpolation of calibrated ages provides systematically older ages than direct recalibration of interpolated ages. (d) Using the chronological controls of the default age model from Neotoma records, calibrate chronological controls reported in radiocarbon years, and then use linear interpolation. Any records with age-reversals were rejected to simplify this illustrative example.*

Histograms of age model types against year of publication

**Figure 2**. *Number and type of default chronologies for North American pollen records in the Neotoma Paleoecological Database based on the original year of publication for the dataset. Within Neotoma the default dataset chronology may be updated by subsequent researchers when this information is provided to a data steward.*

We take the position that developing new age models with recalibrated stratigraphic control points that can account for temporal uncertainty is the ideal approach. However, large scale data analysis across heterogeneous data can be complex, thus the use of harmonized age models, that integrate ideas about the underlying mechanics of the processes linking depth and time become important (Blaauw and Christen, 2011). The Neotoma database (Grimm, 2008) contains 2961 global pollen records that can be used for paleoecological analysis. These records have been obtained from publications that span a time period from 1908 to 2016, with more than half the records coming from before 1983.

This paper builds on work from Dawson *et al*. (Dawson et al., 2016), which rebuilt existing age models using the Bacon age-depth model, by explicitly detailing the key challenges posed by generating a large number of Bayesian age models and quantifying the effects of decisions on aggregate outcomes. We outline the decision making process around age model parameter selection, summarize differences between the original chronologies and the new chronologies, highlight differences between chronologies generated using two age model types and highlight best practices for chronologies within large-scale synthesis. We also identify limitations of current methods, that could serve to improve chronology construction in the future.

# Key Issues in Generating Age Models from Neotoma

## Use of posthoc calibration of individual ages in radiocarbon chronologies

Across the Neotoma database (here we refer to North America only), many records within the database still record chronologies using only radiocarbon years (Figure 2). The transition from age models using only radiocarbon years to those with calibrated radiocarbon years within Neotoma is dramatic. The final radiocarbon model appears to be from 1998, following this we see no more radiocarbon models. Along with this transition, we are seeing a second transition from simple linear models to more complex models using flexible Bayesian methods. A critical question then becomes, when faced with records generated using only uncalibrated dates, should we calibrate radiocarbon dates, generate age models *de novo*, or ignore the records altogether?

While not the preferred method, direct recalibration of interpolated ages does occur within the Neotoma ecosystem. For example, the temporal search function within the Neotoma Explorer (<http://apps.neotomadb.org/explorer>), Tilia (<http://tiliait.org>) and the Neotoma API (<http://api.neotomadb.org>) all use a lookup table that directly recalibrates ages in radiocarbon years. However, this process results in systematic biases in both synthetic data and in the actual Neotoma data (Figure 1).

In the paper we take the position that develping new age models based on recalibrated stratigraphic control points that account for uncertainty is the ideal approach. However, large scale data analysis across heterogeneous data can be complex, thus the use of harmonized age models, that integrate ideas about the underlying mechanics of the processes linking depth and time become important (Blaauw and Christen, 2011). Developments since Blois *et al*. (2011), and workshops such as the PAGES-sponsored Age Models, Chronologies, and Databases Workshop (Grimm et al., 2014) have begun the process of outlining the methods required to undertake large-scale efforts to re-build age models.

## Choice of Age Model Type and Software

Estimating age-depth relationships and uncertainties requires that decisions be made about the type of age modeling approach, choice of software implementation, suitable chronological controls and model parameters (Blois et al., 2011; Giesecke et al., 2014; Grimm et al., 2014). We here pursue Bayesian approaches for their ability to flexibily estimate the ages of pollen samples with robust uncertainty. Several Bayesian age modelling software packages exist, including Bacon, Bcal, Bchron, and Oxcal. We use Bacon because it 1) implements a Bayesian framework that estimates sample age posteriors, 2) is widely used by paloecologists, and 3) accounts for some of the key components in the sediment deposition and accumulation process. Using Bacon to estimate age-depth relationships requires the specification of: 1) chronological controls, which include radiometric dates, biostratigraphic markers, and their uncertainty, 2) priors on the accumulation rate and memory, 3) values for the resolution and structure of divisions within the sediment core (section thicknesses and hiatuses). To ensure a consistent methodology among sediment core in the UMW domain, we developed a set of standard decisions that are described below.

## Choice of Age Controls and Types

Neotoma contains 88 different chronology control types, including various pollen stratigraphic controls, radiocarbon and other isotopic dates and marine isotope stages. Since this analysis was focused primarily on sedimentary pollen records, most of the chronological controls used reflect more recent control types, including 210Pb, 137Cs, and 14C dates, as well as biostratigraphic dates associated with events of known or inferred dates, such as historical fires and the *Tsuga* decline. Core tops are often used, and have highly certain dates, however, several older records use interpolated dates for the core top when sediment was lost in the coring process. These cores should be treated with caution, and this is often noted in the core description.

## Chronological Uncertainty

Methods of understanding an managing radiocarbon uncertainty are well established, however the use of 210Pb and biostratigraphic dates are less well understood in the context of sedimentary archives, particularly in large scale synthesis work. Biostratigraphic events often require indentification by the researcher and may be time-transitive while uncertainties associated with 210Pb may appear to narrow to properly model ages using the Bacon software.

### Dealing with zero-value 210Pb errors

Bacon requires that all chronological markers be associated with defined errors. Historically, some 210Pb data entered into Neotoma were entered without error reporting. Of the 398 210Pb age controls in Neotoma, 148 have no error reported (Figure 3). Binford (1990) reports *"Ninety-five per cent confidence intervals range from about 1 -- 2 years at 10 years of age, 10 -- 20 at 100 years, and 80 -- 90 at 150 years old."* Using this assessment we fit a smooth linear function to assign 95% confidence intervals for all 210Pb dates with missing uncertainty data. These confidence intervals were then divided by 2 (and rounded up to the nearest integer) to obtain standard deviations to be used in the Bacon model (Figure 3).



**Figure 3.** *Reported uncertainty for 210Pb dates within Neotoma, by age, with the total count of unassigned samples by age bin (lower). Estimated uncertainty, using a model relating reported 210Pb uncertainty to age, is then assigned to all 210Pb dates with unreported uncertainty. Present is defined as radiocarbon present, or 1950 CE.*

### Expert elicitation exercise to identify biostratigraphic events

Biostratigraphic events have been used extensively as chronological controls within age models, but are often poorly constrained with respect to the identification of the "rise" event. The *Ambrosia* rise is a well known phenomenon in the eastern United States, and elsewhere, that appears to be contemporaneous with Euro-American settlement, land clearance and the initiation of intensive agriculture in the region (McAndrews, 1968, 1988). In the Upper Midwest, significant increases in *Ambrosia*, *Rumex*, and/or Poaceae are typically coincident with the settlement horizon. Because of this feature it is possible to establish a biostratigraphic chronological control. When the rise can be identified visually it can be used as a stratigraphic marker with an age defined as either immediately preceeding (when the "pre-settlement" sample is used), or immedately after. The pre-settlement sample has been identified for many of the fossil pollen cores in the Neotoma database. To reduce expert bias in the selection of the 'pre-settlement' sample we asked a team of experts to identify pre-settlement samples based on pollen diagrams depicting proportional changes as a function of depth for key indicator species and the ten most abundant arboreal taxa, with no temporal scale. More details of this procedure are available in Dawson et al. (2016) and in Kujawa et al. (Kujawa et al., in review).

The elicitation exercise provides an independently assesed estimate of biostratigraphic change in the pollen records that can be used as a biostratigraphic marker. Gridded Public Land Survey (PLS) datasets in this region provide a historical record of the minimum date of presence for legacy forests (Goring et al., 2016). Goring et al. (2016) use historical records of forest cover from the 1820s - 1903 to provide estimates of forest cover prior to EuroAmerican settlement. The gridded datasets associated with the paper also provides the PLS sampling year; we subtract 50 from the value for grid cells in which sites fall to assign a "pre-settlement" period for each core, with an age uncertainty (stanadard deviation) of 50 years. Ideally, methods would include asymmetric uncertainty, since, in the case of the settlement horizon, we know the minimum age of, but not the most recent age for potential settlement.

Five cores were located within grid cells without digitized PLS data. In these cases, we look to the surrounding grid cells to estimate the maximum sampling year. One site in the Upper Peninsula of Michigan was assigned a maximum sampling year of 1860. The four remaining cores were in the Lower Peninsula of Michigan, and were assigned a maximum year of sampling of 1840.

## Bacon Settings

### Accumulation Rate Priors

Bacon provides a default accumulation rate for all depths based on Goring *et al*. (2012)'s survey of Holocene accumulation rates in eastern North America. Empirical age-depth curves suggest that the age-depth relationship is non-linear, as a result of lower sediment compaction in upper sediments (Goring et al., 2012), basin shape (Bennett and Buck, 2016) and patterns of deposition and sediment transport operating on longer time scales (Goring et al., 2012; Webb and Webb, 1988) and acceleration of erosion rates during the Anthropocene [REF]. Narrowing the geographic and temporal range for the target reconstructions made it neccessary to re-asses mean accumulation rates, focusing on a narrower temporal and spatial window. Regionally, there is evidence that after the time of settlement, sediment accumulation is much faster (fewer years per cm of sediment accumulation) than before settlement (more years per cm of sediment accumulation), although bulk density may decrease.

Many sites show an inflection point in sedimentation rates at or around the Anthropocene horizon in the Upper Midwest. This is likely a combination of both decreased compation in upper sediments and changes in land use and erosion. To generate accumulation rate priors for the Upper Midwest we estimated accumulation rates by pooling the mean age and change in depth for adjacent chronological markers. These accumulation rates were pooled into "modern" and "pre-settlement" groups. For each group a mean equal to the empirical mean of accumulation rates for that time period, and a variance equal to double the empirical variance from the grouped accumulation rates was calculated. Doubling the empirical variance represents a conservative approach to modeling uncertainty, in that it accounts for additional variability in accumulation rates that was not observed in the data.

The modern accumulation rate prior, from more recent sedimentation rates was assigned to the post-settlement portion of cores with "settlement" horizons assigned through the Expert Elicitation. The 2K prior was assigned for the pre-settlement portion of the cores, and for the entirety of cores without identified "pre-settlement" horizons. Cores with two priors were assigned an "instantaneous" (10yr) hiatus at the pre-settlement sample depth.

### Determining section thickness

Bacon works by diving a core into sections whose lengths are determined by the user (Blaauw and Christen, 2011). Sections are the atomic unit of a Bacon model, and various parameters are fit within each section. In particular, the memory parameter defines the flexibility of the accumulation rates between adjacent sections, thus a model with large section thicknesses will be, by virtue of having fewer overall sections, less flexible. However, narrow section thicknesses result in very large run-times, and may show other, unanticipated problems associated with finding fits in multi-parameter space.

Dividing a core into sections is an approximation of time discretization [REF - Andria?], abstracted through the process of deposition. Consistent section widths across all cores (and thus consistent discretization) should be a goal. However, the internal unmeasurable variability in sedimentation rates, and uncertainty in the age-depth models themselves, makes it difficult to implementat consistent section thicknesses. To provide prescriptive widths, but still allow flexibility as needed in the Bacon modelling, each core was run with widths of 5, 10, 15 and 20cm. Model fit was assessed visually, and the best fit model was subsequently chosen to be the default calibrated age model for that core.

# Results

## Age Controls

### 210Pb Errors



**Figure 4**. *Reported and estimated 210Pb ages from the Bacon models show strong accordance, for the most part. Uncertainties within the constructed Bacon models at the depths of the 210Pb samples show higher uncertainty from Bacon models than from the original uncertainty models. The autoregressive nature of the Bacon model, the influence of the memory parameter, and the nature of the Bayesian model itself mean that uncertainty is propagated through the core, and as such low uncertainty in the individual date (x-axis) does not result in low Bacon uncertainty.*

Bacon ages associated with 210Pb ages show a linear relationship, with the Bacon ages older than the reposted 210Pb ages (slope = 1.1241429, p < 0.01; Figure 4). Uncertainty estimates for the Bacon chronologies are also consistently larger than the 210Pb error estimates for the point samples and the difference increases with depth (slope = 1.9762292, p < 0.01).

Only three lakes in the region had original age models that included 210Pb dates and reported estimates of uncertainty for the chronology. Bacon estimates for these records show higher model uncertainty than the original age model, but the difference in modeled uncertainty is site-specific (Figure 5). For Crooked Lake (Brugam et al., 1997) the uncertainty difference is high and consistent with depth/age (triangles, Figure 5), while for Fish Lake (Umbanhowar Jr, 2004) the difference is small (squares; Figure 5). Brown's Bay shows the highest uncertainty, but only at the deepest depth (circles; Figure 5).

Comparison of 210Pb dates measured and reconstructed error at those depths.

**Figure 5**. *For the three cores within the region with both reported 210Pb and reconstructed Bacon uncertainties, the latter are consistently higher, although the magnitude of the difference varies by record. Fish Lake shows the smallest difference, while reconstructed ages for Brown's Bay shows the largest difference.*

### Expert Elicitation

On average, the original modeled ages were older than the biostratigraphic ages assigned based on the dates for PLS in the vicinity of the focal lakes. Bacon models showed better affinity to the assigned biostratigraphic dates based on expert identification of the *Ambrosia* rise in the region. This age differential means that Bacon models are, on average, 44 years younger at the settlement horizon.

Of the 163 records with assigned settlement ages from the elicitation exercise, 55 of the original age models have a modeled ages at the assigned settlement horizon that are younger than 1800 CE, while only 17 of the Bacon models have a modeled age for the settlement horizon younger than 1800 CE.



**Figure 6**. *The relationship between assigned 'settlement' horizons, based on Public Land Survey records and the expert elicitation exercise, and the modeled age of the horizon within the final age-depth model for each individual record. The settlement horizon plotted against the modeled age within the original age model (left); and against the newer Bacon model (right). A best fit linear regression is plotted in blue, with a 1SE envelope in gray. The negative relationship in the original models may reflect an over estimation of sedimentation rates in upper sediments as a result of linear age-model fits. Higher variability in the orginal relationship may represent greater inter-researcher variability in settlement horizon estimation.*

The linear models fit to the settlement horizons in show a negative slope for the relationship between the assigned settlement age and the modeled age (Figure 6). This may be because many of the original linear models over-estimate sedimentation rates in the upper sediments since they assign an inflection point at the horizon. Higher variability is likely a result of both variability in detecting and assigning the settlement horizon among researchers, but also in the process for assigning the age associated with settlement regionally. Researchers may have used historical documents, oral histories or other records. The use of the standardized PLS dates in te new models reduced this source of variability.

## Bacon Settings

### Accumulation Rate Priors

**Table 1**. *Empirical estimates of accumulation rates in the Upper Midwestern United States from Neotoma records.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Group | Mean | Variance | Rate | Shape |
| Modern | 3.02 | 5.70 | 0.5 | 3 1.6 |
| 2K | 15.00 | 250 | 0.0 | 6 0.9 |

Accumulation rates (*yr/cm* of sediment accumulation) for records in the region show clear differentiation between accumulation rates within the last 200 years and accumulations prior to the last 200 years. Means and variance are presented in **Table 1** along with the Gamma rate and shape parameters. Accumulation rates (the mean *yr cm-1* deposition) change dramatically, dropping by five-fold in the modern period.



**Figure 7**. *Sediment accumulation rates for cores in the study region with samples in the last 2000 years (in calibrated radiocarbon years, with 0 at 1950 CE). The upper panel shows the accumulation rate in years per cm against the interval midpoint age in calibrated radiocarbon years. The dashed vertical line indicates 1850 CE, the approximate date used to represent major land-use change as a result of EuroAmerican settlement. Note the strong change in accumulation rates between the periods before and after EuroAmerican settlement. The lower panel shows the distribution of accumulation rates for the pre- and post settlement eras, highlighting the difference in central tendency.*

Deposition rates increase with time across the pre-/post-settlement interval (Figure 7a), however Bacon does not support the ability to allow time-dependent priors, except through the use of sequential hiatuses, with varying accumulation rates. Unfortunately, each interval bounded by a hiatus must have more than one section contained within it, and thus, the implementation of sequential hiatuses becomes problematic and highly site-specific since it depends on sampling interval and the length of the pre-/post-settlement sequence within the core.

### Section Thickness

Bacon models were fit with section thicknesses for values of 5, 10, 15 and 20cm for each core (Figure 8). Optimum section thickness was assigned from the best fit model. A generalized linear model (GLM) using a gamma family shows a significant relationship between total core length and best-fit section thickness ( = 82.7, p > 0.001). This relationship indicates that the principle of wider section thicknesses for Bacon models on longer cores holds generally, but, the distribution of best-fit thicknesses (Figure 8) also appears to indicate that many shorter cores are well served by wide section thicknesses. The GLM accounts for only 26% of total deviance, and should be treated as indicative, not prescriptive.



**Figure 8**. *Bacon section thickness as a function of total core length for best-fit Bacon models on each of the 282 sediment cores with reconstructed age models from the upper Midwestern United States used in this study. A curve (blue line) fit with a generalized linear model with gamma family indicates the predicted relationship between maximum core depth and the best-fit core thickness (gray shading represents 1 standard error).*

# Discussion

Large scale synthesis of paleoecological records relies on our ability to distinguish synchronous, time transient and asynchronous events at a range of spatial scales. For this reason chronology has become central to understanding the past operation of earth system processes (Harrison et al., 2016). Standardized age models, across data sets within a synthesis project reduces a significant source of variability, and, using Bayesian approaches, can improve the ability of researcher to take temporal uncertainty into account when assessing the influence of climatic and biotic factors in driving vegetation change on the landscape. Our approach in this paper, and the associated analysis in Dawson et al. (Dawson et al., 2016) and elsewhere has shown what is possible for large-scale sysntehsis projects, and also, we hope, can serve as a roadmap for future endevours.

Databases such as Neotoma provide a resource that can leverage aggregate data to improve individual results. The finding of and prescription for accumulation rate inflections associated with the pre-/post-settlement horizon provide one example. The expert elicitation exercise has resulted in age models with settlement horizons of more recent age that originally assigned by investigators. This might imply more rapid vegetation change in the post-settlement era, and may change interpretations of vegetation-climate relationships in this era as temporal uncertainty has also increased, as evidenced by our assessment of 210Pb dates.

A challenge for paleoecologists is that site level effects can often play a significant role in sediment accumulation, pollen accumulation and the interpretation of model results. When working with aggregate data it is possible to lose sight of these site level effects, and the potential benefits of doing batch runs with the latest software might make a researcher overlook the peculiarities associated with individual records. Knowing the data well and taking time to examine results in the context of the original publications is critical to evaluating the quality of a record. In this exercise a number of records were rejected, a number of records had modifications made and some peculiar records were retained based on documetation within the primary literature. There is no alternative to knowing your data well.

## Modelling Recommendations

This work has highlighted a number of future opportunities for age-depth model development. The last decade has seen rapid improvements in modeling software and our understanding of the limitations of chronologies (Telford et al., 2004; Trachsel and Telford, 2017), but work remains. This work can be broadly categorized into parameters, processes and products:

### Parameters

Most modern age-modeling software provides for the use of only a single uncertainty model, generally normally distributed error. This is at odds with certain age constraints, such as the "modern" sample, which has an absolutely known age, or the pre-settlement horizon (based on the PLS) which provides a certain "older" date, but an uncertain younger boundary, and as such reflects a truncated distribution. Providing a broader range of uncertainty distributions could improve modelling of certain features used as chronological controls in age models.

Our work and the work of others has indicated the variability of certain key parameters through time. Along with accumulation rates, memory parameters may also change through time, as sediment source changes or as a result of long term changes in precipitation variability evident in regional climate reconstructions [REF]. Accomodating these changes currently requires the use of a zero-length hiatus in Bacon, but allowing time/depth varying parameters using priors drawn from (for example) Neotoma or the Europrean Pollen Database [REF] would provide significantly more control for the models.

Given the large number of parameters required to fit models it is only natural that, at times, changes in parameter value can have unintuitive changes in model performance or fit. Changing memory parameters should increase or decrease the flexibility of models, but section thickness can also influence flexibility, and these may have interacting effects. By allowing users to pass vectors of values for each parameter it may be possible to produce a "landscape" of fits with a single call, from which the best fit model may be selected.

## Process

Sedimentation is a physical process that is driven by climatic factors, changes in sediment source, changes in basin size or shape, and autotrophic effects in the case of palustrine environments. In addition, apparent sedimentation may be affected by the process of coring (as a result of compaction) and in generating composite cores when multiple drives are required to collect a complete sedimentary sequence. When composite cores are used the information relating to how the cores were spliced together is rarely included, but can introduce depth uncertainty in the composite core. Work on assessing fit is being undertaken with Corelyzer and indep within the ODP program, but this information is not yet being introduced into age-depth modelling software.

External influcences on sedimentation rate may appear in secondary proxies, for example, grain size or taxonomic composition in peat sequences. Currently there is no ability to use this information to constrain age model accumulation or flexibility. Providing these secondary controls would add additional complexity to age models, but could improve overall model fit, particularly as XRF and other associated measurements become more commonplace.

### Product

The age model, as a component of paleoecological studies, is increasing critical, as the use of paleoecological records expands. As such, the decisions that go into age-model construction must be preserved in a codified manner. The development of widely accepted provenance standards (e.g., WC3's PROV), and their adoption and promotion within large institutions [REF] mean that this information can be preserved, but adoption of these standards remains a challenge. Community support for standards, and the development and continued participation within research coordination networks such as C4P or EarthRates are then critical for developing these tools as community standards. Codifying these choices and examining them in aggregate can also help reduce the requirements for oversight as the volume of information can provide us with a method by which we can quantify the "art" of age-depth modeling. For example, as the IntCal curve changes, Bacon models that have been preserved in Neotoma with their full suite of parameters may simply be re-run, and clearly attributed as more recent versions. This was a goal of the INQUA Age Modelling meeting (Grimm et al., 2014), but its implementation remains challenging.

|  |  |
| --- | --- |
| Challenge | Solution |
| **Parameter** |  |
| Age uncertainties vary in the shape of their distributions. | Allow uncertainty models to include other distributions, including truncated distributions. |
| Accumulation rates change through time but priors are fixed. | The current solution is to use 'zero-length' hiatuses, which serve to reset the accumulation rate, but allowing models of change based on empirical studies would improve fit. |
| Parameter values can have significant, but unintuitive effects on model fit | Allowing models to accept vectors of possible values may improve fit and provide a better understanding of the interacting relationships between model parameters. |
| **Process** |  |
| Uncertainty in depth is not represented | Estimates of overlap within composite cores, based on correlation from geochemical or visual cues may be included |
| Models don't use secondary information. | Secondary parameters may be used to indicate rapid changes in memory, or accumulation. |
| **Product** |  |
| Age models require significant oversight | Codify certain decisions, provide testing suites for fit parameters. |
| Ensuring that model outputs are clear and the decisions made about parameters are justified. | Improved uptake of reproducible workflows and use of provenance tools. |

## Future development

The tools and techniques for generating age models continues to increase. BChron (Parnell and Parnell, 2016), Bacon STAN (https://github.com/andrewdolman/baconr/), CSciBox (https://www.cs.colorado.edu/~lizb/cscience.html), OxCal (Ramsey and Lee, 2013) all exist as alternative tools, and geochronR (https://github.com/nickmckay/GeoChronR) exists as a resource for managing age models in the R programming environment. It is unlikely the community will select a single resource for modeling chronologies, but the adoption of certain standards across the paleogeosciences can help improve interoperability of data, and ultimately, lead to improvements in our overall understanding of many of the exciting research questions that drive us as a community and as individuals.

# Conclusions

Large-scale synthesis using Quaternary paleoecological records may require the development of purpose-built chronologies. Careful attention to chronology development can result in substantial changes to site-level chronologies, and systematic offsets at a regional scale. The development of new chronologies must consider choices at the scale of individual chronological controls, at the scale of the site, and at regional scales. Using a matrix of parameters with empirical support, and choosing among multiple realizations is likely the most effective method for developing high quality chronologies for use in synthesis research. Once models have been constructed, a secondary goal is the preservation and encoding of the choices that went into model development. High quality paleoecological research increasingly depends on high quality chronological support, and as such, researchers must pass on the information they have gained, in developing site level and regional chronologies, in a well supported and effectively communicated manner.

# Acknowledgements

This work is a contribution of the PalEON Project (<http://paleon-project.org>). SJG would like to thank support from NSG, CEG and AMG. SJG is supported through NSF Grants NSF-1541002, NSF-1550855, NSF-1241868, and NSF-1550707.

# References

Bennett, K., Fuller, J., 2002. Determining the age of the mid-Holocene *Tsuga canadensis* (hemlock) decline, eastern North America. The Holocene 12, 421–429.

Bennett, K.D., Buck, C.E., 2016. Interpretation of lake sediment accumulation rates. The Holocene 26, 1092–1102.

Binford, M.W., 1990. Calculation and uncertainty analysis of 210Pb dates for PIRLA project lake sediment cores. Journal of Paleolimnology 3, 253–267.

Blaauw, M., 2012. Out of tune: The dangers of aligning proxy archives. Quaternary Science Reviews 36, 38–49.

Blaauw, M., 2010. Methods and code for ’classical’ age-modelling of radiocarbon sequences. Quaternary Geochronology 5, 512–518.

Blaauw, M., Christen, J.A., 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process. Bayesian Analysis 6, 457–474.

Blois, J.L., Williams, J.W.J., Grimm, E.C., Jackson, S.T., 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.

Brewer, S., Jackson, S.T., Williams, J.W., 2012. Paleoecoinformatics: Applying geohistorical data to ecological questions. Trends in Ecology & Evolution 27, 104–112.

Brugam, R.B., Giorgi, M., Sesvold, C., Johnson, S.M., Almos, R., 1997. Holocene vegetation history in the Sylvania Wilderness Area of the western Upper Peninsula of Michigan. American Midland Naturalist 62–71.

Davis, M.B., 1981. Outbreaks of forest pathogens in Quaternary history. Proceedings of the fourth International Palynological Conference.

Dawson, A., Paciorek, C.J., McLachlan, J.S., Goring, S., Williams, J.W., Jackson, S.T., 2016. Quantifying pollen-vegetation relationships to reconstruct ancient forests using 19th-century forest composition and pollen data. Quaternary Science Reviews 137, 156–175.

Flantua, S., Blaauw, M., Hooghiemstra, H., others, 2016. Geochronological database and classification system for age uncertainties in Neotropical pollen records. Climate of the Past 12, 387–414.

Giesecke, T., Davis, B., Brewer, S., Finsinger, W., Wolters, S., Blaauw, M., Beaulieu, J.-L. de, Binney, H., Fyfe, R.M., Gaillard, M.-J., Gil-Romera, G., Knaap, P.W.O. van der, Kunes, P., Kühl, N., Leeuwen, J.F.N. van, Leydet, M., Lotter, A.F., Ortu, E., Semmler, M.S.S., Bradshaw, R.H.W., 2014. Towards mapping the late Quaternary vegetation change of Europe. Vegetation History and Archaeobotany 23, 75–86.

Goring, S., Williams, J., Blois, J., Jackson, S., Paciorek, C., Booth, R., Marlon, J., Blaauw, M., Christen, J., 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.H., McLachlan, J.S., others, 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, e0151935.

Grimm, E., 2008. Neotoma: An ecosystem database for the Pliocene, Pleistocene, and Holocene. Illinois State Museum Scientific Papers E Series 1.

Grimm, E.C., Blaauw, M., Buck, C., JW, W., 2014. Age models, chronologies, and databases workshop: Complete report and recommendations. PAGES Magazine 22, 104.

Grimm, E.C., Maher, L.J., Nelson, D.M., 2009. The magnitude of error in conventional bulk-sediment radiocarbon dates from central North America. Quaternary Research 72, 301–308.

Harrison, M., Baldwin, S., Caffee, M., Gehrels, G., Schoene, B., Shuster, D., Singer, B., 2016. Geochronology: It’s about time. Eos 97, 12–13.

Hughen, B.K., McCormac, G., Plicht, J. van der, Spurk, M., 1998. INTCAL98 radiocarbon age calibration, 24,000-0 cal BP. Radiocarbon 40, 1041–1083.

Kujawa, E.R., Goring, S.J., Dawson, A., Calcote, R., Grimm, E.C., Hotchkiss, S.C., Jackson, S.T., Lynch, E.A., McLachlan, J., St-Jacques, J.-M., Umbanhowar, C., Jr, Williams, J.W., in review. The effect of anthropogenic land cover change on pollen-vegetation relationships in the American Midwest. The Anthropocene.

Liu, Y., Brewer, S., Booth, R.K., Minckley, T.A., Jackson, S.T., 2012. Temporal density of pollen sampling affects age determination of the mid-Holocene hemlock (*Tsuga*) decline. Quaternary Science Reviews 45, 54–59.

McAndrews, J., 1968. Pollen evidence for the protohistoric development of the “Big Woods” in Minnesota. Review of Palaeobotany and Palynology 7, 201–211.

McAndrews, J.H., 1988. Human disturbance of North American forests and grasslands: The fossil pollen record, in: Vegetation History. Springer, pp. 673–697.

Ojala, A., Francus, P., Zolitschka, B., Besonen, M., Lamoureux, S., 2012. Characteristics of sedimentary varve chronologies–a review. Quaternary Science Reviews 43, 45–60.

Parnell, A., Parnell, M.A., 2016. Package “bchron”.

Ramsey, C.B., Lee, S., 2013. Recent and planned developments of the program oxcal. Radiocarbon 55, 720–730.

Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Bronk Ramsey, C., Brown, D.M., Buck, C.E., Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Haflidason, H., Hajdas, I., Hatté, C., Heaton, T.J., Hogg, A.G., Hughen, K.A., Kaiser, K.F., Kromer, B., Manning, S.W., Reimer, R.W., Richards, D.A., Scott, E.M., Southon, J.R., Turney, C.S.M., Plicht, J. van der, 2013. Selection and treatment of data for radiocarbon calibration: An update to the International Calibration (IntCal) criteria. Radiocarbon 55, 1923–1945.

Telford, R.J., Heegaard, E., Birks, H.J.B., 2004. All age–depth models are wrong: But how badly? Quaternary science reviews 23, 1–5.

Trachsel, M., Telford, R.J., 2017. All age–depth models are wrong, but are getting better. The Holocene 27, 860–869.

Uhen, M.D., Barnosky, A.D., Bills, B., Blois, J., Carrano, M.T., Carrasco, M.A., Erickson, G.M., Eronen, J.T., Fortelius, M., Graham, R.W., others, 2013. From card catalogs to computers: Databases in vertebrate paleontology. Journal of Vertebrate Paleontology 33, 13–28.

Umbanhowar Jr, C.E., 2004. Interaction of fire, climate and vegetation change at a large landscape scale in the Big Woods of Minnesota, USA. The Holocene 14, 661–676.

Ward, G.K., Wilson, S., 1978. Procedures for comparing and combining radiocarbon age determinations: A critique. Archaeometry 20, 19–31.

Webb, R.S., Webb, T., III, 1988. Rates of sediment accumulation in pollen cores from small lakes and mires of eastern North America. Quaternary Research 30, 284–297.