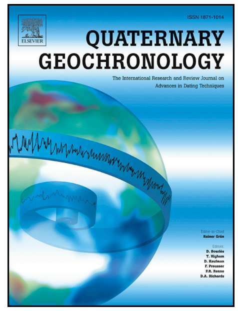
Reconstructing the accumulation history of a saltmarsh sediment core: which age-depth model is best?

Accepted Manuscript



Reconstructing the accumulation history of a saltmarsh sediment core: Which age-depth model is best?

Alexander J. Wright, Robin J. Edwards, Orson van de Plassche, Maarten Blaauw, Andrew C. Parnell, Klaas van der Borg, Arie F.M. de Jong, Helen M. Roe, Katherine Selby, Stuart Black

PII: S1871-1014(16)30090-5

DOI: 10.1016/j.quageo.2017.02.004

Reference: QUAGEO 825

To appear in: *Quaternary Geochronology*

Received Date: 22 July 2016

Revised Date: 8 February 2017

Accepted Date: 8 February 2017

Please cite this article as: Wright, A.J., Edwards, R.J., van de Plassche, O., Blaauw, M., Parnell, A.C., van der Borg, K., de Jong, A.F.M., Roe, H.M., Selby, K., Black, S., Reconstructing the accumulation history of a saltmarsh sediment core: Which age-depth model is best?, *Quaternary Geochronology* (2017), doi: 10.1016/j.quageo.2017.02.004.

Reconstructing the accumulation history of a saltmarsh sediment core: Which age-depth model is best?

Alexander J. Wright[a], Robin J. Edwards[b]*, Orson van de Plassche[a†], Maarten Blaauw[c], Andrew C. Parnell[d], Klaas van der Borg[e], Arie F. M. de Jong[e], Helen M. Roe[c], Katherine Selby[f], Stuart Black[g]

[a] Department of Marine Biogeology, Faculty of Earth & Life Sciences, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, the Netherlands.

[b] School of Natural Sciences, Trinity College Dublin, Dublin 2, Republic of Ireland.

[c] School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, BT7 1NN, UK.

[d] School of Mathematics and Statistics, Insight Centre for Data Analytics, University College Dublin, Belfield, Dublin 4, Republic of Ireland.

[e] Faculty of Science, Department of Physics, Universiteit Utrecht, 3508 TA Utrecht, the Netherlands.

[f] Environment Department, University of York, Heslington, York, YO10 5DD, UK.

[g] School of Archaeology, Geography and Environmental Science, University of Reading, Reading RG6 6AB, UK.

[†]Deceased

*Corresponding Author. Email: robin.edwards@tcd.ie

Keywords: relative sea-level reconstruction; saltmarsh; age-depth model; Bayesian statistics

ABSTRACT

Saltmarsh-based reconstructions of relative sea-level (RSL) change play a central role in current efforts seeking to quantify the relationship between climate and sea-level rise. The development of an accurate chronology is pivotal, since errors in age-depth relationships will propagate to the final record as alterations in both the timing and magnitude of reconstructed change. A range of age-depth modelling packages are available but differences in their theoretical basis and practical operation mean contrasting accumulation histories can be produced from the same dataset.

We compare the performance of five age-depth modelling programs (Bacon, Bchron, Bpeat, Clam and OxCal) when applied to the kinds of data used in high resolution, saltmarsh-based RSL reconstructions. We investigate their relative performance by comparing modelled accumulation curves against known age-depth relationships generated from simulated stratigraphic sequences. Bpeat is particularly sensitive to non-linearities which, whilst maximising the detection of small rate changes, has the potential to generate spurious variations, particularly in the last 400 years. Bacon generally replicates the pattern and magnitude of change but with notable offsets in timing. Bchron and OxCal successfully constrain the known accumulation history within their error envelopes although the best-fit solutions tend to underestimate the magnitude of change. The best-fit solutions of Clam generally replicate the timing and magnitude of changes well, but are sensitive to the underlying shape of the calibration curve, performing poorly where plateaus in atmospheric ^{14}C concentration exist.

We employ an ensemble of age-depth models to reconstruct a 1500 year accumulation history for a saltmarsh core recovered from Connecticut, USA based on a composite chronology comprising 26 AMS radiocarbon dates, ^{210}Pb, ^{137}Cs radionuclides and an historical pollen chronohorizon. The resulting record reveals non-linear accumulation during the late Holocene with a marked increase in rate around AD1800. With the exception of the interval between AD1500 and AD1800, all models produce accumulation curves that agree to within ~10 cm at the century-scale. The accumulation rate increase around AD1800 is associated with the transition from a radiocarbon-based to a ^{210}Pb-dominated chronology. Whilst repeat analysis excluding the ^{210}Pb data alters the precise timing and magnitude of this acceleration, a shift to faster accumulation compared to the long-term rate is a robust feature of the record and not simply an artefact of the switch in dating methods. Simulation indicates that a rise of similar magnitude to the post-AD1800 increase (detrended increase of ~16 cm)

is theoretically constrained and detectable within the radiocarbon-dated portion of the record. The absence of such a signal suggests that the recent rate of accumulation is unprecedented in the last 1500 years. Our results indicate that reliable (sub)century-scale age-depth models can be developed from saltmarsh sequences, and that the vertical uncertainties associated with them translate to RSL reconstruction errors that are typically smaller than those associated with the most precise microfossil-based estimates of palaeomarsh-surface elevation.

1. Introduction

Constructing an accurate accumulation history is a vital but non-trivial component of most sediment-based palaeoenvironmental reconstructions (Telford et al., 2004; Blaauw and Heegaard, 2012). This is exemplified by the current generation of 'high resolution' relative sea-level (RSL) studies seeking to employ saltmarsh sediments as late Holocene 'tide gauges' (see Barlow et al., 2013). In this approach the age and altitude of palaeomarsh-surfaces (PMS) (Figure 1a) are combined with estimations of the height above sea level at which they formed (Figure 1b) in order to reconstruct the RSL change experienced at a study site (Figure 1c). Microfossils such as foraminifera are used to infer PMS height whilst age control is provided by AMS radiocarbon dating of saltmarsh plant remains. Whilst some microfossil samples are directly dated, the age of others must be inferred by interpolation between dated horizons. Although this situation is not unique to RSL reconstruction, establishing an accurate age-depth relationship is particularly important for saltmarsh-based studies since it directly impacts the magnitude of the reconstructed change as well as determining its timing (see Figure 1c and 1d). As core collection typically targets high marsh environments, the resulting RSL reconstruction is primarily controlled by the sediment accumulation history (Edwards, 2007).

In recent years, several software tools have been developed to assist in the process of chronology construction. Whilst some packages employ classical statistical methods to develop age-depth models (e.g. Clam: Blaauw, 2010), the use of Bayesian statistics has become increasingly common (Parnell et al., 2011; Parnell and Gehrels, 2015). Variations in underlying theory and its practical application mean that each model handles data differently and, in this way, a single dataset can produce a diversity of accumulation histories. In fact, Blaauw and Heegaard (2012) note that model choice is the greatest source of uncertainty in age-depth modelling. Previous work highlights that each modelling approach has particular strengths and weaknesses, with no single model out-performing all others in every situation (Parnell et al., 2011). Consequently, comparative assessment of model performance using simulated and real data is an important step to ensure that informed choices are made during chronology construction (e.g. Telford et al., 2004; Blockley et al., 2007). Furthermore, since inaccurate accumulation histories can give rise to spurious RSL signals, it is important to ensure that any inferred rate changes are not simply artefacts of the calibration process or switches between dating method (Gehrels et al., 2005; Barlow et al., 2013).

In this paper we present a new, well-dated saltmarsh sediment core from Connecticut, USA, covering the last 1500 years which is typical of sequences targeted in ‘high resolution’ RSL studies (e.g. Kemp et al., 2011, 2013). We use a suite of simulations to evaluate the performance of five age-depth modelling packages (Bacon, Bchron, Bpeat, Clam and OxCal) in order to address the following questions: 1) Do age-depth models introduce spurious accumulation rate changes?; 2) Can we tell if recent accumulation rates are without precedent given down-core changes in dating approach and resolution?

2. Saltmarsh core and age data

A 1.82 m-thick sequence of high saltmarsh peat was recovered from Pattagansett River marsh in Connecticut, USA (Figure 2). Twenty-six samples for AMS radiocarbon dating were collected at 6 cm intervals below 29 cm depth to produce a 1500 year-long record with an average of one radiocarbon date every 60 calendar years (Figure 3, Table B.1). This radiocarbon-based chronology was supplemented by pollen and short-lived radionuclide data from the upper 64 cm of the sequence (Figure 4, Table 1, Table B.2).

An initial manual wiggle-match of the radiocarbon data to the calibration curve (van de Plassche et al., 2001) confirms the predominantly linear nature of the age-depth profile and the absence of significant hiatuses (Figure 3). This is supported by the lithostratigraphy (Figure 2c) which indicates consistent accumulation within a high marsh environment (abundant *Spartina patens* rhizomes with uniform $\delta^{13}C$ signatures (Table B.1)). The resulting late Holocene accumulation rate of 1.1 mm/yr matches estimates of the underlying rate of glacio-isostatic adjustment (GIA) for the region (1.0 ± 0.2 mm/yr, Donnelly et al., (2004); 1.1 ± 0.1 mm/yr, Engelhart et al., (2009)), implying that the effects of sediment compaction in this shallow core are negligible. Forward extrapolation of this long-term rate fails to intersect with the modern surface by ~13 cm (Figure 3b, 4f), indicating that an increase in accumulation rate must have occurred in the most recent portion of the record. This inference is confirmed by both a simple linear interpolation from the core top to the *Ambrosia* chronohorizon (mean accumulation rate of 1.7 mm/yr since AD1650) or from the ^{210}Pb and ^{137}Cs data (mean accumulation rates of 2.1 mm/yr since AD1850 or 2.6 mm/yr since AD1963). The local rate of RSL rise recorded by the tide gauge at New London is 2.3 mm/yr since AD1938.

Whilst this simple approach of comparing linear trends is sufficient to identify the existence of a recent acceleration in saltmarsh accumulation rate, it cannot reliably quantify it given the range of possible rates (1.6 mm/yr – 2.8 mm/yr), or unequivocally date the timing of its onset. More importantly it is unable to address the question of whether a change of similar magnitude occurred in the earlier, radiocarbon-dated portion of the record, which is masked within the larger age error envelope.

Age-depth modelling has been used to refine the timing and significance of recent changes identified in RSL records and to decrease the magnitude of age error envelopes by considering the stratigraphic ordering of dates within a sediment core (e.g. Kemp et al., 2011). However, given the differences in performance and underlying theory, it is unclear which approach will produce the most precise and accurate accumulation history for a particular sediment core. In the following section, we use simulations to produce a series of known accumulation histories against which we can evaluate the performance of the different age-depth modelling packages. Whilst numerous permutations of synthetic data are possible (e.g. uneven sampling intervals, varying age precision etc), the characteristics of the simulated dataset will influence relative model performance. Consequently, we develop a series of synthetic dates that emulate the sampling resolution and dating precision of the Pattagansett core chronology.

3. Age-depth simulation and modelling

3.1 Developing synthetic sedimentary sequences

We develop hypothetical age-depth scenarios to serve as targets for the chronological modelling programs (Figure 5, Appendix A). We initially consider a linear age-depth profile (Simulation 1) reflecting constant accumulation at a rate of 1.1 mm/yr (the long-term linear rate of the Pattagansett core). We simulate the process of radiocarbon-based chronology construction by 'sampling' a hypothetical core at 6 cm depth intervals and then 'decalibrating' the known calendar age to a radiocarbon date. We follow the method of Michczyński (2007) which uses the calibration curve to convert a calendar age into a radiocarbon age which is then assigned an error term to emulate a radiocarbon date. We use an error term of ± 35 yrs thereby producing a synthetic dataset of comparable resolution and precision to the Pattagansett record (Figure 5a). Finally, we include two age markers (along with the core-top) to simulate the provision of the age constraints provided by pollen and short-lived radionuclide data.

We then explore the reconstruction of variable accumulation rates (Simulations 2-6) by superimposing an oscillating (sinusoidal) term upon the background linear rise (Figure 5b, Figure 5c, Appendix A). We vary the amplitude and the period of this oscillating term whilst ensuring sediment age increases consistently with depth in core. The magnitudes of the detrended oscillations range from 6 – 21 cm (Table A.1); the former being the smallest theoretically detectable signal based on our sampling resolution and the latter being the largest possible oscillation that does not violate the principle of superposition. A sinusoidally oscillating term is selected for operational simplicity and is not intended to imply that 'real' RSL oscillations are necessarily periodic. Instead, we use multiple simulations to gauge the capacity of different models to reliably capture non-linear changes of varying magnitude. We present these data as detrended signals since this is the format commonly used for comparison with models and between regions with differing background rates of RSL rise (e.g. Engelhart et al., 2009; Gehrels, 2010; Kemp et al., 2011; Barlow et al., 2014; Kopp et al., 2016).

3.2 Age-depth models

The synthetic data are processed by five age-depth modelling packages that are freely available and can be run on a desktop computer. Four of these programs (Bacon: Blaauw & Christen, 2011; Bchron: Haslett & Parnell, 2008; Bpeat: Blaauw & Christen, 2005; Clam: Blaauw, 2010) are written for the free, open-source statistical environment R (R Development Core Team, 2010), whilst OxCal (Bronk Ramsey, 1995, 2001, 2009a) is a stand-alone package that can be run on-line or downloaded (c14.arch.ox.ac.uk). Clam (Blaauw, 2010) employs classical age-depth modelling, provides both numerical best-fit and confidence interval interpolations and was developed as a quick and transparent way to produce age-depth models. The remaining programs employ a Bayesian statistical approach which accommodates the introduction of additional 'prior' information to assist in refining the probability distributions of age data (see Parnell et al., 2011 for a review). For example, applying the principle of superposition means that models do not produce accumulation histories with age reversals and confidence intervals become narrower.

Bpeat (Blaauw & Christen, 2005) provides numerical best-fit interpolations, graphical grey-scale summaries of uncertainty, and essentially functions as an advanced form of 'wiggle match dating'. Bacon (Blaauw & Christen, 2011) provides numerical best-fit and confidence interval interpolations, graphical grey-scale summaries of uncertainty, and is superficially similar to Bpeat in terms of its tuneable parameters (see Appendix A). Bchron (Haslett & Parnell, 2008) provides numerical best-fit and confidence interval interpolations and is fully automated so does not require extensive preliminary analysis to determine optimal parameters. Finally, OxCal (Bronk Ramsey, 1995, 2001, 2008, 2009a; Bronk Ramsay and Lee, 2013) provides numerical confidence interval interpolations but no best-fit solution. It also has additional functionality in the manner in which outliers are identified during age-depth modelling (Bronk Ramsey, 2009b).

Further details of the theoretical basis and operation of each of the models are provided in the publications that accompany them and useful comparative reviews of a subset of packages have been made by Blockley et al. (2007) and Parnell et al. (2011). Whilst the number of model development runs (>100) means the details cannot be presented here, we summarise the key outcomes of these analyses, and document the selection of parameters where they deviate from the default values (Appendix A). The nature of the models (e.g. use of Monte Carlo sampling) means that results may vary slightly between runs made with identical settings. Consequently, during model evaluation and development, we considered the output from multiple runs, and present results as the mean of three runs per reconstruction. The final selection of parameters (Table 2) was made to optimise the fit between model output and the suite of simulated curves, whilst ensuring choices were parsimonious and avoided over-fitting (Blaauw & Heegaard, 2012).

We assess the performance of these models by comparing the accuracy and precision of the detrended profiles. We measure accuracy in terms of how closely a best-fit model solution approximates the target accumulation history, and the extent to which this known curve is contained within the error envelope of the reconstruction. The magnitude of the error envelope is used to indicate model precision, and hence increased model precision must be accompanied by better model fit if the reconstruction is still to be deemed accurate. Quantitative measures of overall goodness-of-fit are included in Table A.2.

3.3 Modelling linear accumulation

Figure 6 presents the detrended accumulation histories produced by each of the modelling programs for the linear age-depth scenario. Since accumulation is constant throughout, any deviation from a horizontal line indicates the potential for spurious rate changes to be introduced during the calibration and interpolation process.

In general, we consider all models to have accurately reconstructed the linear accumulation scenario in that the best-fit curves do not deviate substantially from a straight line (misfits < 5 cm), and the real profile is always contained within the confidence intervals (Figure 6a, Figure 6b). This is an important result as it demonstrates that reconstructions produced by any of these programs do not produce spurious oscillations linked to the underlying structure of the radiocarbon calibration curve (see Gehrels et al., 2005; Gehrels & Woodworth, 2013; Barlow et al., 2013), at least not when based on the kind of well-dated sequence considered here.

Small differences in model reconstructions do arise indicating variations in their sensitivity to calibration curve shape. The best-fit curves of Bpeat and Clam are most susceptible to this effect during the last 400 years of the record and the wide Clam confidence intervals indicate reduced precision at certain points, equivalent to age uncertainties of up to ~150 years (Figure 6d).

3.4 Modelling non-linear accumulation

Non-linear scenarios reveal the potential for real rate changes to be distorted or masked within a predominantly radiocarbon-dated sequence. We begin by considering a signal of ~21 cm (Simulation 6, Table A.1) which is of comparable magnitude to the recent (c. 100-200 yrs) detrended increase in RSL rise reported from the Atlantic coast of North America (e.g. Gehrels, 2010; Kemp et al. 2011).

Figure 7 presents the simulated accumulation curve along with the reconstructed curves produced by the various programs. We initially compare model performance by asking three questions: 1) Does the model consistently detect accumulation rate change? 2) Does the model accurately represent the magnitude of change? 3) Does the model reliably reproduce the pattern of change?

All models unambiguously detect the accumulation rate changes and this is clearly reflected in both the best-fit solutions and confidence intervals (Figure 7a, Figure 7b). The magnitude of change is excellently reproduced by the best-fit reconstructions of Bpeat. The best-fit curves for Clam and Bacon reliably capture the magnitude of some oscillations, but are not consistent throughout the sequence, encountering particular difficulties in the last few hundred years of the record. The best-fit solution of Bchron consistently underestimates the peak magnitude of change.

The nature of the Bpeat program means that the oscillating curve is essentially represented by a series of linear segments. Whilst these do an excellent job of approximating the upward limb of each oscillation, the falling limbs appear as isolated or disjointed collections of points, effectively resembling hiatuses that correlate with phases of extremely low or zero accumulation. These falling limbs are associated with significant age misfits (Figure 7e). Whilst the best-fit curve for Clam does a good job of replicating the pattern of change for the earlier oscillations, the narrow confidence intervals associated with its reconstructions do not always circumscribe the actual accumulation curve, and consequently may give the impression of false precision. The difficulties encountered in the last few hundred years, reflecting the underlying structure of the radiocarbon calibration curve, are also evident as larger confidence intervals that still do not always contain the real accumulation history (Figure 7b).

Whilst Clam and Bacon indicate broadly similar magnitudes of change, there is a phase offset in the Bacon reconstruction which results in a tendency for both the best-fit curve and the confidence intervals to lead the real accumulation curve. This produces large misfits (particularly for age) and the appearance of poorer overall performance (Figure 7e), even though the general shape of the confidence intervals are a reasonable approximation of the underlying signal. This temporal offset may be linked to the use of a sinusoidal term (e.g. an aliasing effect), or may reflect our choice of ‘section thickness’ in the Bacon setup (Appendix A). Irrespective of the precise cause, these between-model differences are indicative of the kinds of temporal uncertainty associated with model choice and the reconstruction process, even where all models employ data with the same sampling frequency. In this instance, whilst inter—model differences are typically of the order of c. 50 years, they may rise to a century or more (Figure 7e). Overall, Bchron and Oxcal outperform the other programs in terms of their ability to reliably capture known accumulation variability within their confidence intervals (Figure 7b).

To explore further the issue of signal detectability we repeat the process using a series of simulations with oscillations of differing magnitude (Table A.1, Appendix A). These results indicate that the ability to consistently detect rate changes begins to fail with oscillations ~10 cm in magnitude (i.e. Simulation 3). For example whilst Bpeat identifies the existence of every oscillation, it fails to reliably capture the magnitude of every change (Figure A.10c). Although none of the other best-fit solutions accurately reflect this scale of oscillation, the confidence intervals of Bchron and OxCal continue to perform well by circumscribing the actual accumulation curve and providing indications of its non-linear form (Figure A.13c, Figure A.14c).

Figure 8 shows a simulated curve with oscillations of ~13 cm (Simulation 4) which are comparable in magnitude to the recent increase in accumulation recorded in the Pattaganssett record (Figures 3 & 4). All models recognise the existence of the oscillations, with the best-fit curve for Bpeat most closely approximating their magnitude (Figure 8a). In this instance, the best-fit curve of Clam outperforms that of Bacon which has become somewhat unstable, perhaps linked to the greater significance of phase-shifts in a scenario with shorter period oscillations (Figure 8c). Once again, whilst the best-fit solution for Bchron underestimates the magnitude of change, both its confidence intervals, and those of OxCal, do a good job of delimiting the target accumulation curve (Figure 8b).

Collectively, these results demonstrate an accumulation signal of ~21 cm (Simulation 6), comparable to the increases in RSL rise reported from other sites along the Atlantic coast of USA, will be detectable within the radiocarbon-dated portion of the record irrespective of the age-depth modelling program employed (Figure 7). Conversely, signals with a magnitude of less than ~10 cm (Simulation 3) will likely be circumscribed by the confidence intervals (Figure A.3c) but may not be accurately resolved by a best-fit solution (Figure A.2c) given the quality of the data, vertical sampling interval and the underlying background accumulation rate.

Whilst the choice of modelling program influences the detail of the final best-fit accumulation curve, differences between models only translate to centimetre-scale vertical discrepancies in their reconstructions (Figure A.7). These offsets are generally small when compared to the size of the confidence intervals associated with each model. As the lower limits of signal detection are approached, inter-model differences tend to become more pronounced with different models ‘failing’ in contrasting ways. An important exception to this general pattern is the relatively poor performance of all models in the last 400 years of the record reflecting the underlying shape of the radiocarbon calibration curve. Whilst vertical offsets may be subtle, misfits in the reconstructed timing of changes can be of the order of a century or more.

4. Developing an age-depth model for the saltmarsh core

The simulations presented in Section 3 are tailored to exploring model performance when applied to a dataset with a radiocarbon-dating precision (±35 yrs) and effective sampling resolution (1 date every c. 60 yrs) comparable to our Connecticut saltmarsh core (Section 2). These provide information on the magnitude of the detrended signal that may be reliably detected within the radiocarbon-dated portion of our record (~13 cm or more). Oscillations smaller than this may be constrained within the confidence intervals but will not be accurately discernible in envelope shape or associated best-fit curves. Subtle changes of ~5 cm are equivalent to the misfits associated with modelling linear accumulation and so can effectively be regarded as indistinguishable from 'noise'.

In light of the differences in performance outlined in Section 3, we employ an ensemble of age-depth models to utilise the relative strengths of the different approaches and infer additional information from the discrepancies between reconstructions. We exclude Bacon from this analysis due to the 'phase-shift' effect noted in simulation (Section 3.4).

Applying Occam's razor (and in the absence of evidence to the contrary) the assumption of a linear accumulation rate is a reasonable starting place for chronological model development. More complicated accumulation histories only need be invoked when this linear assumption fails to adequately describe the data. The sensitivity of Bpeat to non-linearity (Section 3.3) makes it an excellent first-assessment tool. If Bpeat suggests limited divergence from a linear profile, we can be confident that we are not missing any significant rate changes. Where Bpeat does identify potential rate changes, we can use the best-fit solution to provide an indication of their likely location, and to get an approximate magnitude of the detrended signal involved. The cost of this sensitivity is that Bpeat has the greatest potential to produce spurious 'jumps' where none exist, notably around the c. AD1700 'threshold' in the calibration curve (e.g. Figure 6a).

Once this initial framework is in place, Bchron or OxCal can be used to provide confidence intervals on the basis that they consistently circumscribe the simulated accumulation curve (Section 3.4). Whilst the extremes of these confidence intervals will tend to overestimate the magnitude of an actual oscillation (Figure 8b), the best-fit solution of Bchron has a tendency to smooth or dampen the oscillation (Figure 8a), with this becoming more pronounced as dating precision reduces. Therefore as a final step, it may be instructive to consult the best-fit solution of Clam since this tends to provide a middle-ground reconstruction against which the extremes of Bpeat and Bchron/OxCal can be evaluated, particularly in the earlier (pre-AD1600) portion of the record (Figure 8e).

4.1 Evaluating the model ensemble

The initial screening run using Bpeat provides strong evidence for non-linear accumulation within the record (Figure 9a). Changes in the early portion of the sequence are small (~5 cm) and therefore below the limit of reliable detection inferred from simulation. More marked variation is apparent after AD1500 with a reduction in rate, followed by a short interval of quasi-uniform accumulation before the most recent acceleration commenced around AD1800. Whilst this pronounced oscillation (detrended rise of 26 cm) is much larger than anything experienced during the preceding millennium, simulations indicate that Bpeat 'failure' may overestimate the magnitude of change during this time interval (Figure 8a, Figure 8c).

Adding the Bchron / OxCal confidence intervals and best-fit solution refines the initial accumulation history outlined by Bpeat (Figure 9b), constraining the maximum size of any pre-AD1500 detrended change to ~13 cm or less and placing the c. AD1800 rise between ~9 and 18 cm. Both the confidence intervals and the best fit solution (Bchron) indicate pre-AD1500 oscillations that are larger than any artefacts noted in the linear simulation (Figure 6), suggesting they are real features of the record. The post-AD1500 rate reduction is essentially absent from the Bchron / Oxcal reconstructions and so the subsequent detrended rise is correspondingly smaller. This more muted picture of change is consistent with the tendency for the Bchron best-fit curve to smooth variability evident in the simulations (Figure 8a).

Finally, the best-fit curve of Clam reconstructs oscillations in the pre-AD1500 portion of the record which equate to a detrended signal of ~12 cm and are generally contained within the Bchron / Oxcal confidence intervals (Figure 9c). The only departure from this pattern is following the post-AD1500 deceleration when the curve plots just below the confidence intervals between AD1600 and AD1800, giving a detrended recent rise of ~21 cm.

4.2 Model sensitivity to age data selection

To investigate the effect of a switch in dating method, we repeat the age-depth model runs for our saltmarsh core with the ^{210}Pb data removed (Figure 10b). The impact of this change on the best-fit reconstructions is minimal for Bchron and Clam, whilst its effect on Bpeat is to shift the major inflection in accumulation rate from AD1800 to AD1700. In contrast a marked post-AD1700 impact is seen in the confidence intervals of OxCal and Bchron, the latter of which in particular expands significantly until constrained by the ^{137}Cs marker.

The difference in behaviour between Bpeat, Bchron and Clam can be attributed to the manner in which they incorporate the pollen chronohorizon data and use it to constrain which side of the

AD1650 horizon contemporaneous radiocarbon dates are placed (Figure 3b). To illustrate this effect, we repeat our analysis with the pollen chronohorizon also removed (Figure 10c). The best-fit solutions of Bchron and Clam are not significantly affected, and there is no substantial further expansion of the Oxcal and Bchron confidence intervals. In contrast, the best-fit solution of Bpeat alters dramatically, effectively smoothing out the large post-AD 1500 rate reduction and producing a reconstruction that approximates that of Bchron. It is interesting to note that removal of this age constraint produces a less 'rigid' reconstruction in the earlier portion of the record, with Bpeat now closely tracking the Bchron best-fit solution and adding further support for non-linear change prior to AD1500.

As a final illustration of sensitivity, we remove the radiocarbon date at 65 cm depth (adjacent to the pollen chronohorizon) which plots as a potential outlier in the original linear 'wiggle-match' (Figure 3a). Whilst the best-fit curve of Bchron is not significantly impacted, the Clam and Bpeat reconstructions more closely align and the best-fit curves plot close to that of Bchron for the period AD1500-1600 (Figure 10d). Collectively, these model runs indicate that Bchron and Oxcal produce the most 'stable' reconstructions and that as data are removed the best-fit solutions of Bpeat and Clam tend to converge toward that of Bchron.

4.3 Towards a 'consensus' accumulation curve

We combine these reconstructions to develop an informal 'consensus' accumulation curve (Figure 10e). With the exception of the period between AD1500 and AD1800, all models show excellent agreement (within ~5 cm of each other). Our consensus curve is constrained within the Bchron and Oxcal confidence intervals, respects all points where the individual age-depth profiles overlap, and remains within ~10cm of all best-fit solutions. For the interval centred on AD800, our curve approximates the best-fit solution of Bchron on the basis that Bpeat does not register a large oscillation at this point. Between AD1000 and AD1300 our curve closely tracks the best-fit solution of Clam on the basis that a rate reduction is evident in all models whilst simulation results suggest the best-fit solution of Bchron is likely to smooth this signal. Between AD1300 and AD1400, the best-fit solutions of all models are essentially indistinguishable and show an accelerated rate of rise which is also mirrored in the confidence interval trends. Whilst the small magnitude of this signal (~ 5cm) is below the reliable limits of detection indicated by simulation, the agreement between models suggests that an accelerated rate of rise sometime during the 13th and 14th centuries is likely, although its magnitude cannot be accurately determined.

After AD1400, the best-fit solutions begin to diverge and our consensus curve initially tracks that of Clam and Bpeat on the basis of the smoothing-tendency associated with Bchron. The consensus curve then diverges from both that of Bpeat and Clam and instead tracks the lower limit of the Bchron and Oxcal confidence intervals. This solution is selected on the basis that simulations indicate Bpeat and Clam are prone to producing spurious signals in this time interval, whilst the combined confidence intervals of Bchron and Oxcal consistently circumscribe the target curves during simulation. In effect, it produces a best-fit solution that lies midway between the extremes of Bchron and Bpeat. From AD1800 onward the best fit solutions converge as they enter the more tightly constrained portion of the chronology, and are essentially indistinguishable during the 19th and 20th centuries. An inflection centred around AD1800 is clear in all chronologies, as is the stepped nature of the final portion of the curve with a brief slowdown centred on AD1900 interrupting the accelerated rate of the last 200 years.

4.4 Are recent accumulation rates unprecedented?

It is clear that the upper portion of our core from Pattagansett, which post-dates AD1800, accumulated faster than the background rate experienced over the last 1500 years. The detrended magnitude of this recent rise is between ~9 – 26 cm (equivalent to accumulation rates of 1.6 – 2.4 mm/yr) although the results of simulation suggest that these extremes are likely under- and over-estimates of the real signal. Instead, the consensus 'best-fit' curve places the rise at ~16 cm which, whilst equivalent to a century-scale accumulation rate of ~1.9 mm/yr, includes an interval of reduced rate centred around AD1900. This accords well with the accumulation rates inferred by simple linear interpolation of the pollen and short-lived radionuclide data (Table 1).

The simulation results indicate that a signal of 16 cm would be accurately resolved in the radiocarbon-dated portion of the record. Whilst it is possible that an oscillation of up to ~13 cm could be accommodated within the confidence intervals of the accumulation curve prior to AD1800, simulations indicate that these intervals tend to overestimate the magnitude of change. This fact, coupled with the limited response of Bpeat which simulations show to be sensitive to non-linearities, suggests that a pre-AD 1800 signal of the order of ~10 cm or less is the most plausible interpretation of the data. On this basis, we conclude that accumulation during the last two centuries occurred at a century-scale rate that is without precedent in the previous 1300 years of the record.

Similar accelerations in accumulation rate (translated into increases in the rate of RSL rise) have been documented in a number of saltmarshes around the globe (Kemp et al. 2009, 2011; Gehrels & Woodworth, 2013). Whilst simulations like those presented here would be needed to determine if the noted increases are larger than any signal that could be masked within the age-depth uncertainties particular to each record, our results provide support for the contention that recent rates of RSL rise along parts of the Atlantic coast of N. America are without precedent for much of the Common Era (e.g. Kemp et al., 2013, 2015; Kopp et al., 2016). In their synthesis sea-level reconstructions, Kopp et al. (2016) conclude that global sea level variability over the pre-20th century Common Era was smaller than the ±25 cm estimated in the IPCC fifth assessment report (Mason-Delmotte et al., 2013) and instead was very likely to be between ~±7 cm to ~±11 cm. Our simulations indicate that even the smaller of these signals (ie a 14 cm 'oscillation') would be detectable if expressed as an accumulation rate change in a well-dated saltmarsh core with similar properties to our material from Pattagansett.

4.5 Implications for the use of saltmarshes as 'geological tide gauges'

Geological data are required to extend the duration of instrumental records in order to address topical questions relating to the timing, magnitude, spatial pattern and significance of sea-level change (Gehrels 2010; Mason-Delmotte et al., 2013; Miller et al., 2013). Saltmarsh sediments have attracted particular interest due to the fact that they can furnish near-continuous, (sub)centennial- and decimetre-scale records that overlap with tide gauge data and extend back many centuries into the past. Proxy records that are precise enough to permit meaningful comparison with tide gauges are at the limits of resolution, both of the methodologies employed to develop them, and of the sedimentary archives from which they are extracted (Edwards, 2007). Consequently, whilst the use of saltmarshes as geological tide gauges is now an established technique, its application requires detailed knowledge of the sediments and the proxies employed, and careful consideration of the uncertainties associated with reconstructions of age and altitude (Gehrels & Shennan, 2015; Shennan, 2015).

Barlow et al. (2013) highlight the need to evaluate age models and suggest that particular caution is required when interpreting RSL changes that may reflect the underlying structure of the radiocarbon calibration curve, or which coincide with the junction between chonological methods. The results of our simulations and the comparative application of multiple age-depth modelling approaches permit some more detailed comments to be made on these subjects with the important caveat that they apply to well-dated sequences such as our Pattagansett core which is devoid of any significant hiatuses.

Firstly, whilst simple interpolation of radiocarbon data does have the potential to introduce spurious rate changes that mirror the calibration curve (Gehrels et al., 2005), our linear simulations demonstrate that when dealing with a well-dated sequence, all of the age-depth modelling approaches we consider are not significantly influenced by this phenomenon.

Secondly, by necessity, all chronologies that cover the intersection between instrumental and geological data will be derived from a composite of chronological methods. The fact that the junction between ^{210}Pb and ^{14}C records is coincident with the timing of a potentially significant rate change means that simply extrapolating and comparing two linear trends is prone to error. However, since the age-depth models take into consideration age uncertainties, there is no *a priori* reason that a switch in dating approach will result in a marked rate change in best-fit solutions. Instead, the shift in resolution and precision will be expressed as a change in the width of confidence intervals as is clearly illustrated by the reconstructions from Pattagansett (Figure 10). Hence, whilst the most significant rate change of our 1500 year record occurs close to the boundary between dating approaches, it is not an artefact of this switch in chronometers.

Whilst the presence of an acceleration is a robust feature of our record, the exact magnitude and timing of the change, and the precision with which it can be established, are influenced by the ^{210}Pb data, the supporting chronological information provided by the pollen chronohorizon and the choice of modelling program employed. In our example, the post-AD1800 detrended accumulation rate ranged from 1.6 – 2.4 mm/yr depending on which age-depth model was selected, and this uncertainty exists before accounting for additional error terms that ultimately influence a RSL reconstruction (e.g. underlying GIA rate, PMS height reconstruction etc). Similarly, age-misfits varied between models when applied to simulated data with a resolution / precision comparable to our saltmarsh core (Figure 7e, Figure A.4, Figure A.5). Encouragingly errors were typically less than ~50 years for much of the record, but could rise to a century or more at certain points, with no modelling program being completely immune to this effect which reflects the underlying shape of the calibration curve. This is noteworthy since there is particular interest in trying to pin-point the timing of any recent acceleration in the rate of RSL rise with a view to better understanding the drivers and mechanisms responsible (e.g. Gehrels & Woodworth, 2013; Long et al., 2014; Kopp et al. 2016).

Gehrels & Woodworth (2013) attempt to distil this kind of detailed information from seven saltmarsh records but choose to exclude all data points that are not directly dated on the basis that age-depth modelling can introduce spurious signals. This conservative approach was justified given that only two of the sites possessed sequences with sufficiently well-constrained chronologies to produce the kinds of records described above. This limitation exists despite the records being a carefully selected sub-set of the available data, chosen on the basis of their comparatively high quality. This reinforces the fact that the chronological requirements for the use of saltmarsh sequences as geological tide gauges are extremely exacting and have rarely been met for practical reasons such as cost of analysis and access to suitable sedimentary sequences. For example, irregularly spaced dates, changes in the type of dated material and sequences with varied lithology, all present additional challenges when age-depth modelling. Simulations such as those performed here, using synthetic data designed to emulate the characteristics of the sedimentary sequences of interest, are useful exploratory tools for assessing model performance and gauging record resolution.

Whilst a comprehensive assessment of all these variables is beyond the scope of this paper, we briefly examine the influence of dating precision by repeating our simulations using synthetic radiocarbon dates with ^{14}C age errors of ± 70 years, comparable to radiocarbon dates reported in some of the older saltmarsh literature (e.g. Nydick et al., 1995) and ± 10 years, similar to the pooled high precision AMS dates of some more recent work (e.g. Kemp et al., 2009). The results are illustrated in Figure 11 for an oscillation of ~13 cm (Simulation 4). The best-fit solutions based on lower precision dates fail to reliably resolve the oscillation (Figure 11c) and the confidence intervals for all models are expanded yet do not always circumscribe the simulated curve (Figure 11f). In contrast, the high precision dates reduce confidence interval width (increased precision) whilst still generally constraining the simulated accumulation curve (retained accuracy). However, the depth and age misfits of the best-fit solutions are not significantly altered by the use of high-precision dates since they remain ultimately tied to the shape of the calibration curve. Instead, the use of complementary forms of chronological information, such as stable lead isotope or other dated pollution markers, will be required to further refine these chronologies (e.g. Gehrels et al., 2006, 2008; Kemp et al., 2012; Marshall, 2015).

Finally, it is important to acknowledge that record resolution is not simply a product of down-core sampling frequency and age precision, but is instead conditioned by the accumulation characteristics of the individual sediment core. For example, in regions of rapid RSL rise (e.g. high GIA-related subsidence), the creation of accommodation space permits rapid sediment accumulation, resulting in a higher temporal sampling resolution for a given down-core sampling interval. When considering an oscillating RSL term, the background accumulation rate also determines the maximum size of oscillation that can be accommodated before sediment over-printing occurs. Hence, in locations with low background accumulation rates, the magnitude of the resolvable signal is reduced. Consequently, the comparison of RSL records from regions of contrasting GIA, even following detrending, is not always straightforward. Simulations using synthetic data tailored to the particular characteristics of each record may prove useful tools for evaluating the significance of apparent inter-record differences.

5. Summary and conclusions

The use of saltmarshes as geological 'tide gauges' requires the development of precise and accurate accumulation histories for the sediment cores used to furnish the proxy data. Advances in age-depth modelling coupled with detailed dating of sedimentary sequences using a combination of AMS radiocarbon, short-lived radionuclide and historical chronohorizon techniques, mean robust (sub)century-scale reconstructions are possible. Next generation RSL reconstruction methods will combine age-depth relationships and PMS estimates within a single numerical framework (e.g. Cahill et al., 2016), but the resulting reconstructions are still governed by the age-depth model choice. The importance of evaluating the performance of each module in the assembled hierarchical model increases with the complexity of data manipulation, as the direct connection between raw data and resulting reconstruction is obfuscated incrementally.

We compare the performance of five age-depth modelling programs through the use of simulation and subsequent application to a real saltmarsh sediment core. On the basis of our results we conclude:

- Simulations constructed to emulate the sampling resolution and data quality of a real sedimentary record provide valuable insights into the relative performance of age-depth models, whilst indicating the smallest change that can theoretically be resolved;
- No single modelling package out-performs all others, but an ensemble approach can exploit different model strengths to produce a 'consensus' estimate of accumulation history;

- In a well-dated sequence, inter-model differences in reconstruction are generally smaller than the error terms associated with them, and translate to vertical errors that are typically less than the uncertainties associated with microfossil-based PMS reconstruction;
- Age-depth modelling does not generate spurious oscillations related to the underlying structure of the radiocarbon calibration curve when applied to well-dated sequences such as our example core from Pattagansett River marsh, Connecticut, USA;
- Whilst the interval between AD1500 and AD1800 is particularly challenging for age-depth models based on radiocarbon dating, an increase in accumulation relative to the background rate is noted at Pattagansett and this is not an artefact generated by a switch between dating methods;
- Precisely delimiting the timing of the recent increase in accumulation rate is reliant on the provision of complementary (i.e. non-radiocarbon) age data, but the balance of evidence suggests marsh surface rose more during the last 200 years than at any other comparable period in this 1500 year-long record.

ACKNOWLEDGEMENTS

This work was supported by: the Dutch National Research Programme on Global Air Pollution and Climate Change (Ocean-climate variability and sea level in the North Atlantic region since AD 0); the Vrije Universiteit Amsterdam (Coastal Records); the European Union (Simulations, observations & palaeoclimatic data: climate variability over the last 500 years) to OvdP. AW and RE wrote the manuscript. The Pattagansett core chronological data were provided by AdJ and KvdB (radiocarbon), HR and KS (pollen), and SB (^{210}Pb). AP and MB provided assistance with age model runs (Bpeat, Clam, Bacon, Bchron). This paper is a contribution to IGCP Project 639, “Sea-level change from minutes to millennia”.

References

Appleby, P.G., 2001. Chronostratigraphic techniques in recent sediments. In: Last, W.M., Smol, J.P. (Eds.), Tracking Environmental Change Using Lake Sediments. Volume 1: Basin Analysis, Coring and Chronological Techniques Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 171-203.

Appleby, P.G., 2008. Three decades of dating recent sediments by fallout radionuclides: a review 1. The Holocene 18, 83-93.

Barlow, N.L.M., Shennan, I., Long, A.J., Gehrels, W.R., Saher, M.H., Woodroffe, S. A., Hillier, C., 2013. Salt marshes as late Holocene tide gauges. Glob. Planet. Change 106, 90–110. doi:10.1016/j.gloplacha.2013.03.003

Barlow, N.L.M., Long, A.J., Saher, M.H., Gehrels, W.R., Garnett, M.H., Scaife, R.G., 2014. Salt-marsh reconstructions of relative sea-level change in the North Atlantic during the last 2000 years. Quat. Sci. Rev. 99, 1-16. doi: 10.1016/j.quascirev.2014.06.008

Blaauw, M., 2010. Methods and code for “classical” age-modelling of radiocarbon sequences. Quat. Geochronol. 5, 512–518. doi:10.1016/j.quageo.2010.01.002

Blaauw, M., Christen, J.A., 2005. Radiocarbon peat chronologies and environmental change Journal of the Royal Statistical Society Series C-Applied Statistics 54, 805-816.

Blaauw, M., Christen, J.A., 2011. Flexible Paleoclimate Age-Depth Models Using an Autoregressive Gamma Process. Bayesian Analysis 6, 457-474.

Blaauw, M., Heegaard, E., 2012. Estimation of Age-Depth Relationships. In H.J.B. Birks et al. (eds.), Tracking Environmental Change Using Lake Sediments. Developments in Paleoenvironmental Research 5. Springer. DOI 10.1007/978-94-007-2745-8 12

Blockley, S.P.E., Blaauw, M., Bronk Ramsey, C., van der Plicht, J., Blockley, S., 2007. Building and testing age models for radiocarbon dates in Lateglacial and Early Holocene sediments. Quat. Sci. Rev. 26, 1915-1926.

Bronk Ramsey, C., 1995. Radiocarbon calibration and analysis of stratigraphy: The OxCal program. Radiocarbon, 37(2), 425-430.

Bronk Ramsey, C., 2001. Development of the radiocarbon calibration program OxCal. Radiocarbon, 43(2A), 355-363.

Bronk Ramsey, C., 2008. Deposition models for chronological records. Quaternary Science Reviews, 27(1-2), 42-60.

Bronk Ramsey, C., 2009a. Bayesian analysis of radiocarbon dates. Radiocarbon, 51(1), 337-360.

Bronk Ramsey, C., 2009b. Dealing with outliers and offsets in radiocarbon dating. Radiocarbon, 51(3), 1023-1045.

Bronk Ramsey, C., Lee, S., 2013. Recent and Planned Developments of the Program OxCal. Radiocarbon, 55(2-3), 720-730.

Brugham, R.B., 1978. Pollen indicators of land-use change in Southern Connecticut. Quaternary Research, 9, 349-362.

Cahill, N., Kemp, A.C., Horton, B.P., Parnell, A.C., 2016. A Bayesian hierarchical model for reconstructing relative sea level: from raw data to rates of change. Clim. Past Discuss. 12.2, 525-542. doi:10.5194/cpd-11-4851-2015

Clark, J.S., Overpeck, J.T., Webb, T III, Patterson, W.A. III., 1986. Pollen stratigraphic correlation and dating of barrier-beach peat sections. Review of Palaeobotany and Palynology, 47, 145-168

Donnelly, J.P., 2004. Coupling instrumental and geological records of sea-level change: Evidence from southern New England of an increase in the rate of sea-level rise in the late 19th century. Geophys. Res. Lett. 31, 2–5. doi:10.1029/2003GL018933

Edwards, R., 2007. Sea levels: resolution and uncertainty. Prog. Phys. Geogr. 31, 621–632. doi:10.1177/0309133307087086

Engelhart, S.E., Horton, B.P., Douglas, B.C., Peltier, W.R., Tornqvist, T.E., 2009. Spatial variability of late Holocene and 20th century sea-level rise along the Atlantic coast of the United States. Geology 37, 1115–1118. doi:10.1130/G30360A.1

Gehrels, R., 2010. Sea-level changes since the Last Glacial Maximum: an appraisal of the IPCC Fourth Assessment Report. J. Quat. Sci. 25, 26–38. doi:10.1002/jqs.1273

Gehrels, W., Kirby, J., Prokoph, a, Newnham, R., Achterberg, E., Evans, H., Black, S., Scott, D.,

2005. Onset of recent rapid sea-level rise in the western Atlantic Ocean. Quat. Sci. Rev. 24, 2083–2100. doi:10.1016/j.quascirev.2004.11.016

Gehrels, W.R., Hayward, B.W., Newnham, R.M., Southall, K.E., 2008. A 20th century acceleration of sea-level rise in New Zealand. Geophys. Res. Lett. 35, 1–5. doi:10.1029/2007GL032632

Gehrels, W.R., Marshall, W.A., Gehrels, M.J., Larsen, G., Kirby, J.R., Eiriksson, J., Heinemeier, J., Shimmield, T., 2006. Rapid sea-level rise in the North Atlantic Ocean since the first half of the nineteenth century. The Holocene 16, 949–965. doi:10.1177/0959683606hl986rp

Gehrels, W.R., Shennan, I., 2015. Sea level in time and space: revolutions and inconvenient truths. J. Quat. Sci. 30, 131–143. doi:10.1002/jqs.2771

Gehrels, W.R., Woodworth, P.L., 2013. When did modern rates of sea-level rise start? Glob. Planet. Change 100, 263–277. doi:10.1016/j.gloplacha.2012.10.020

Haslett, J., Parnell, A., 2008. A simple monotone process with application to radiocarbon-dated depth chronologies. Journal of the Royal Statistical Society Series C-Applied Statistics 57, 399-418

Kemp, A.C., Hawkes, A.D., Donnelly, J.P., Vane, C.H., Horton, B.P., Hill, T.D., Anisfeld, S.C., Parnell, A.C., Cahill, N., 2015. Relative sea-level change in Connecticut (USA) during the last 2200 yrs. Earth Planet. Sci. Lett. 428, 217–229. doi:10.1016/j.epsl.2015.07.034

Kemp, A.C., Horton, B.P., Culver, S.J., Corbett, D.R., van de Plassche, O., Gehrels, W.R., Douglas, B.C., Parnell, A. C., 2009. Timing and magnitude of recent accelerated sea-level rise (North Carolina, United States). Geology 37, 1035–1038. doi:10.1130/G30352A.1

Kemp, A.C., Horton, B.P., Donnelly, J.P., Mann, M.E., Vermeer, M., Rahmstorf, S., 2011. Climate related sea-level variations over the past two millennia -Supporting Information. Proc. Natl. Acad. Sci. U. S. A. 108, 11017–22. doi:10.1073/pnas.1015619108

Kemp, A.C., Horton, B.P., Vane, C.H., Bernhardt, C.E., Corbett, D.R., Engelhart, S.E., Anisfeld, S.C., Parnell, A.C., Cahill, N., 2013. Sea-level change during the last 2500 years in New Jersey, USA. Quat. Sci. Rev. 81, 90–104. doi:10.1016/j.quascirev.2013.09.024

Kemp, A.C., Sommerfield, C.K., Vane, C.H., Horton, B.P., Chenery, S., Anisfeld, S., Nikitina, D., 2012. Use of lead isotopes for developing chronologies in recent salt-marsh sediments. Quat. Geochronol. 12, 40–49. doi:10.1016/j.quageo.2012.05.004

Kopp, R.E., Kemp, A.C., Bittermann, K., Horton, B.P., Donnelly, J.P., Gehrels, W.R., Hay, C.C., Mitrovica, J.X., Morrow, E.D., Rahmstorf, S., 2016. Temperature-driven global sea-level variability in the Common Era. Proc. Natl. Acad. Sci. 1–8. doi:10.1073/pnas.1517056113

Long, A.J., Barlow, N.L.M., Gehrels, W.R., Saher, M.H., Woodworth, P.L., Scaife, R.G., Brain, M.J., Cahill, N., 2014. Contrasting records of sea-level change in the eastern and western North Atlantic during the last 300 years. Earth Planet. Sci. Lett. 388, 110-122. DOI: 10.1016/j.epsl.2013.11.012

Marshall, W., 2015. Chronohorizons: indirect and unique event dating methods for sea-level reconstructions. In Shennan, I., Long, A.J., Horton, B.P., (eds.) Handbook of Sea-Level Research. John Wiley & Sons. 373-385.

Michczynski, A., 2007. Is it possible to find a good point estimate of a calibrated radiocarbon date? Radiocarbon 49, 393-401.

Masson-Delmotte, V., M. Schulz, A. Abe-Ouchi, J. Beer, A. Ganopolski, J.F. González Rouco, E. Jansen, K. Lambeck, J. Luterbacher, T. Naish, T. Osborn, B. Otto-Bliesner, T. Quinn, R. Ramesh, M. Rojas, X. Shao and A. Timmermann, 2013. Information from Paleoclimate Archives. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Miller, K.G., Kopp, R.E., Horton, B.P., Browning, J. V, Kemp, A.C., 2013. A geological perspective on sea-level rise and its impacts along the U.S. mid-Atlantic coast. Earth's Futur. 1, 3–18. doi:10.1002/2013EF000135

Nydick, K.R., Bidwell, A.B., Thomas, E., Varekamp, J.C., 1995. A sea-level rise curve from Guilford, Connecticut, USA. Mar. Geol. 124, 137–159. doi:10.1016/0025-3227(95)00037-Y

Parnell, A.C., Buck, C.E., Doan, T.K., 2011. A review of statistical chronology models for high-resolution, proxy-based Holocene palaeoenvironmental reconstruction. Quat. Sci. Rev. 30, 2948–2960. doi:10.1016/j.quascirev.2011.07.024

Parnell, A.C., Gehrels, W.R., 2015. Using chronological models in late Holocene sea-level reconstructions from saltmarsh sediments. In Shennan, I., Long, A.J., Horton, B.P., (eds.) Handbook of Sea-Level Research. John Wiley & Sons. 500-513.

Shennan, I., 2015. Handbook of sea-level research: framing research questions. In Shennan, I., Long, A.J., Horton, B.P., (eds.) Handbook of Sea-Level Research. John Wiley & Sons. 3-25.

Telford, R., Heegaard, E., Birks, H., 2004. All age–depth models are wrong: but how badly? Quat. Sci. Rev. 23, 1–5. doi:10.1016/j.quascirev.2003.11.003

van de Plassche, O., Edwards, R.J., van der Borg, K., de Jong, A.F.M., 2001. C-14 wiggle-match dating in high-resolution sea-level research. Radiocarbon 43, 391-402.

Table 1 Summary of chronological data

Data Type	Depth (cm)	Age (yrs AD)	Comment
Core top / surface	1 ± 0.5	2001 ± 1	Date of core retrieval
^{137}Cs	10 ± 1	1963 ± 1	63 samples, 29 depths with activity: AD1963 peak in thermonuclear fallout correlate with peak activity in ^{137}Cs. Linear rate = 2.6 ± 0.2 mm/yr
^{210}Pb	1 – 42	1998 - 1799	63 samples, 48 depths with activity: age model constrained by AD1963 marker using piecewise CRS approach (Constant Rate of Supply, Appleby in Last and Smol, 2001; Appleby, 2008). Linear rate ~ 2.1 mm/yr
Pollen	61 ± 3	1650 ± 50	Ragweed (*Ambrosia*) rise at 58 cm (after AD1640) correlated with historical timing of early European settlement in the region (Brugham, 1978; Clark et al., 1986): assigned a conservative ± 50 age uncertainty term. Linear rate = 1.6 – 1.9 mm/yr
New London tide gauge	-	1938 – 2006	2.3 mm/yr
^{14}C dates (PMS depths, calibrated ages)	26±3 - 176±3	1953 - 431	26 AMS dated samples
^{14}C wiggle match rate	26 - 176	1888 - 511	1.1 mm/yr (also equivalent to rate of GIA): under-predicts position of present day marsh surface by 13.4 cm

Table 2 Summary of model specifications used in the simulations. See Appendix A for further details.

Model	Parameters
Bacon	Mean accumulation rate (α) = 1.0mm/yr; Section thickness = variable
Bchron	Automated procedure; Includes depth uncertainty of ± 3 cm for dated samples
Bpeat	Mean accumulation rate (α) = 1.0mm/yr; No. of sections = 15; HiatusA= 0.5
Clam	Run length = 100,000 iterations (exclude age reversals); Span = 0.3; smoothed spline
Oxcal	P_Sequence; k=2; General outlier model

Figure Captions

Figure 1. Illustration of how palaeomarsh-surface (PMS) accumulation dominates the reconstructed relative sea-level (RSL) record. (a) Radiocarbon-dated plant macrofossils fix PMS position at particular points in time, producing an age-depth plot. (b) PMS elevation above mean sea level is reconstructed from sample foraminiferal content, producing a depth-elevation plot. (c) Age-depth modelling assigns a date to each foraminiferal sample to produce a reconstruction of PMS elevation change over time. The modelled accumulation curve influences the timing and shape of the reconstructed RSL change. (d) The resulting RSL reconstructions, which are typically presented following removal of the long-term (linear) trend, are strongly influenced by the choice of age-depth model.

Figure 2. Core site location and summary lithostratigraphy for Pattagansett River marsh, Connecticut, USA. NL = New London tide gauge.

Figure 3. (a) Linear 'wiggle match' of AMS radiocarbon dates from Pattagansett River marsh (Core PY) showing the global fit on the IntCal09 calibration curve. (b) Calibrated radiocarbon dates (2σ) plotted alongside chronohorizons provided by an historical pollen marker (green) and the peak in ^{137}Cs (red). Forward projection of the long-term linear trend (1.1 mm/yr) underestimates the marsh surface by ~13cm.

Figure 4. Composite chronological dataset spanning the post-AD1600 period. (a) *Ambrosia* pollen abundance levels increasing above 2% indicate land clearance and provide a chronohorizon dating to AD1650 ± 50 years. (b-e) Gamma spectrometry results including excess lead (total ^{210}Pb – ^{226}Ra), ^{137}Cs and ^{241}Am. The peak in atmospheric thermonuclear weapons testing and subsequent partial nuclear test ban treaty (AD1963 ± 2 years) is correlated with the ^{137}Cs maximum and subsequent rapid fall, and the lower peak in ^{241}Am. (f) The composite chronology derived from excess ^{210}Pb results (piecewise constant rate of supply model) is shown as horizontal black bars, alongside the calibrated radiocarbon dates (2σ) shown as grey crosses, and the pollen (green) and ^{137}Cs (red) chronohorizons.

Figure 5. Simulated accumulation curves emulating the sampling resolution and precision of the Pattagansett River saltmarsh core for: (a) linear; and (b-c) non-linear modelling scenarios (see Table

B.1 for details). Upper graphs show simulated age-depth curves (solid black lines) and synthetic radiocarbon sampling points (black boxes). The ‘decalibrated’ radiocarbon dates derived from these points of known age are plotted as grey crosses. Additional chronohorizons are shown as green (pollen) and red (^{137}Cs) squares. Lower graphs show the simulated curves following detrending for a long-term (linear) accumulation rate of 1.1 mm / yr.

Figure 6. Graphs of best-fit (a, c) and ±95% confidence interval (b, d) generated by the various age modelling programs for Simulation 1 (linear). Data are plotted as misfits in depth (a, b) and age (c, d) between the simulated accumulation curve and the reconstructed curves produced by the age-depth models. Line colours and envelope shading refer to the particular modelling programs indicated on the figure.

Figure 7. Graphs of best-fit (a, c, e) and ±95% confidence interval (b, d, f) generated by the various age modelling programs for Simulation 6 (~21 cm oscillation). The detrended simulated (target) accumulation curve is plotted alongside the reconstructed curves produced by the age-depth models (a, b). Data are also plotted as misfits in depth (c, d) and age (e, f) between the simulated and reconstructed accumulation curves. Line colours and envelope shading refer to the particular modelling programs indicated on the figure.

Figure 8. Graphs of best-fit (a, c, e) and ±95% confidence interval (b, d, f) generated by the various age modelling programmes for Simulation 4 (~13 cm oscillation). The detrended simulated (target) accumulation curve is plotted alongside the reconstructed curves produced by the age-depth models (a, b). Data are also plotted as misfits in depth (c, d) and age (e, f) between the simulated and reconstructed accumulation curves. Line colours and envelope shading refer to the particular modelling programs indicated on the figure.

Figure 9. Detrended accumulation curves for the Pattagansett River marsh core produce by: (a) Bpeat best-fit; (b) Bchron best-fit with Bchron and Oxcal confidence intervals; (c) Clam best-fit. Symbols indicate location and type of age data used in age-depth modelling. Line colours and envelope shading refer to the particular modelling programs indicated on the figure.

Figure 10. A comparison of detrended accumulation curves for the Pattagansett River marsh core illustrating the influence of dataset composition on age-depth modelling. Reconstructions are the best-

fit curves (Bpeat, Bchron, Clam) and confidence intervals (Bchron, Oxcal) developed: (a) from all chronological data; (b) following exclusion of the ^{210}Pb chronohorizon; (c) following exclusion of the both ^{210}Pb and pollen chronohorizons; (d) following exclusion of both chronohorizons and possible ^{14}C outlier. An informal ‘consensus’ accumulation curve based on the complete dataset is shown in (e). See text for discussion.

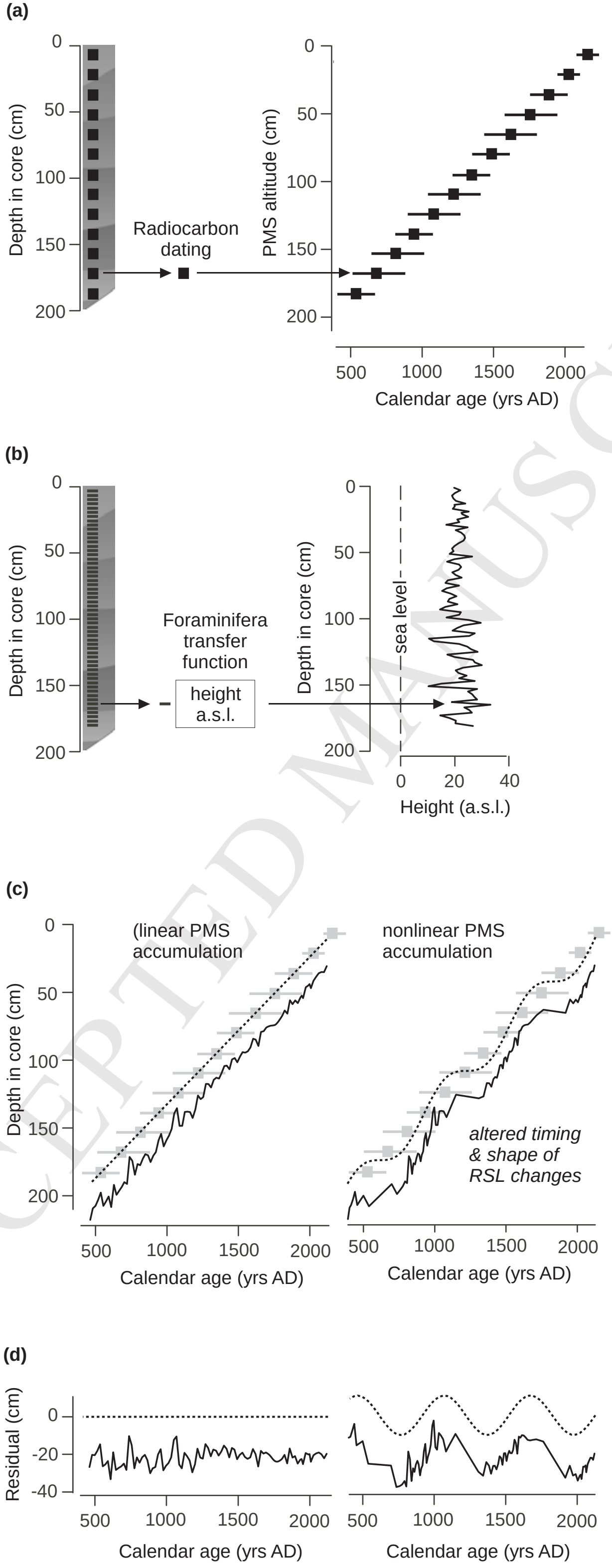
Figure 11. An illustration of the influence that radiocarbon-date precision has on the capacity of age-depth modelling programs to accurately resolve non-linear accumulation based on Simulation 4 (~13 cm oscillation). Reconstructions are developed from synthetic data with a precision of ± 10 ^{14}C yr (a, d), ± 35 ^{14}C yr (b, e) and ± 70 ^{14}C yr (c, f). Graphs of best-fit (a, b, b) and ±95% confidence interval (d, e, f) generated by the various modelling programmes are plotted alongside the simulated (target) accumulation curve.

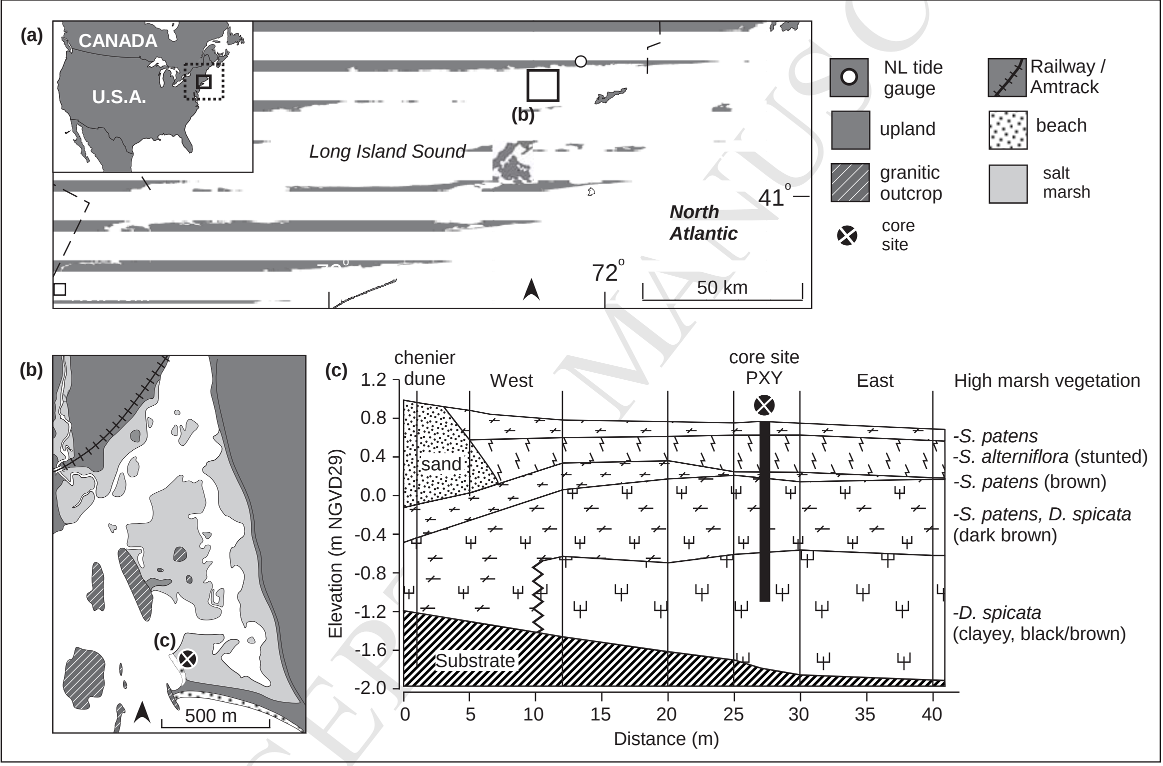
Appendices

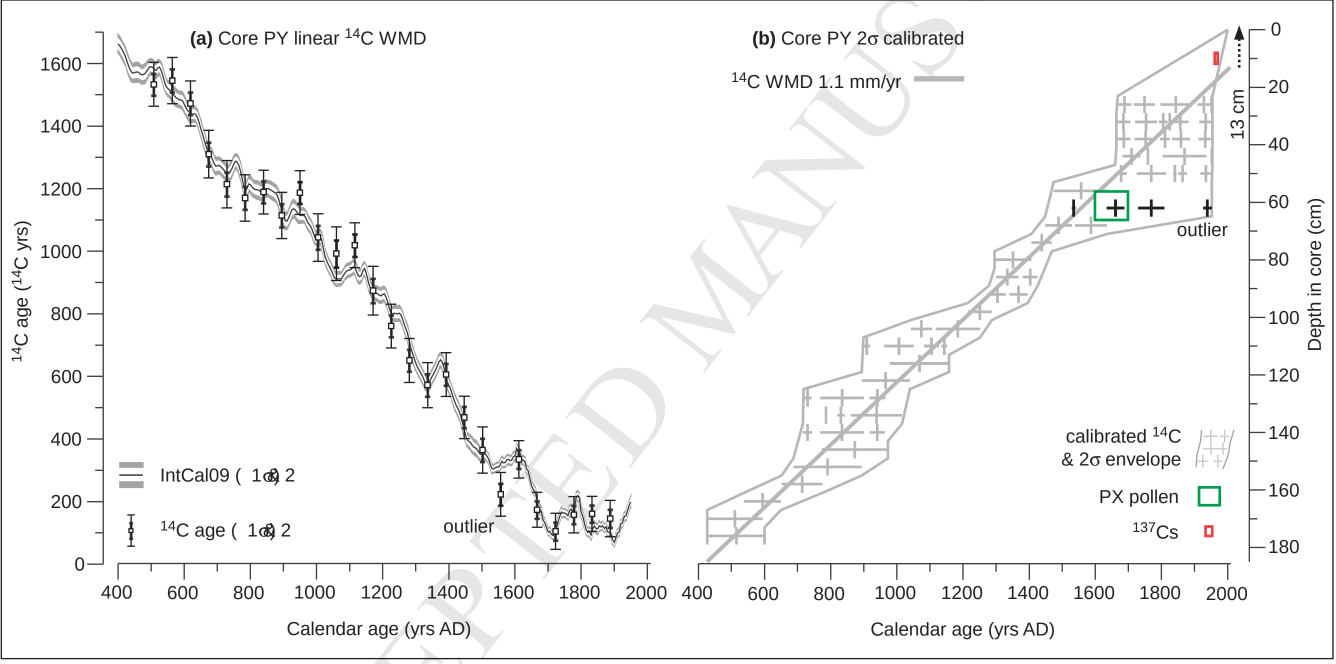
Appendix A: Supplementary information summarising age-depth modelling packages, model scenarios and model run outputs

Appendix B: Details of age data for Pattagansett River saltmarsh core

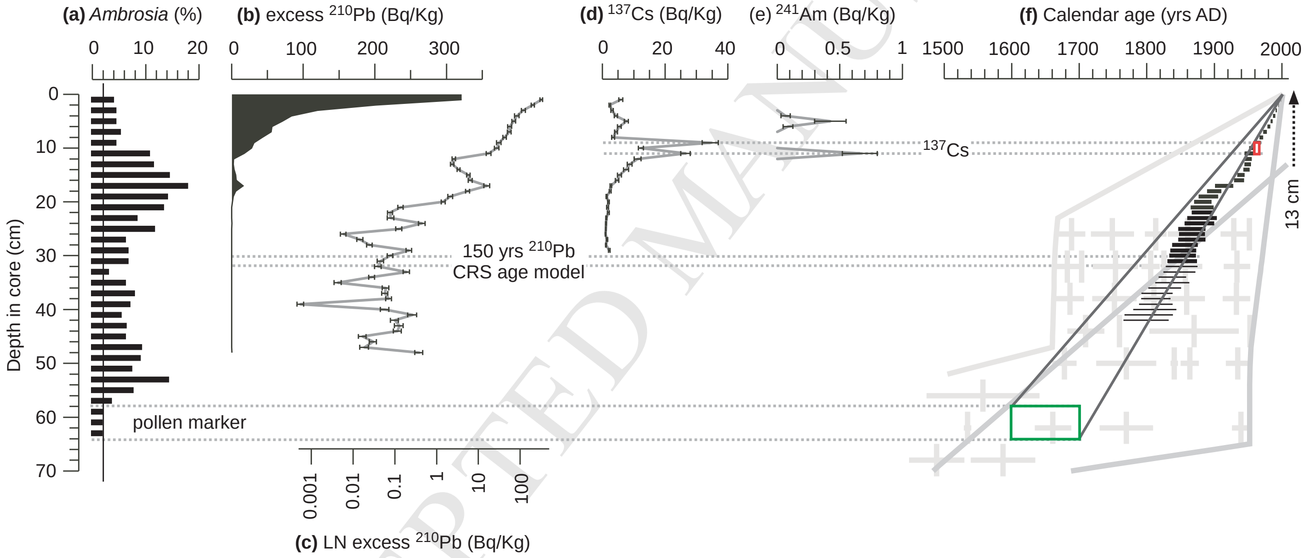
ACCEPTED MANUSCRIPT


(a)
Depth in core (cm)
Radiocarbon dating
PMS altitude (cm)
Calendar age (yrs AD)
(b)
Depth in core (cm)
Foraminifera transfer function
height a.s.l.
Depth in core (cm)
sea level
Height (a.s.l.)
(c)
Depth in core (cm)
(linear PMS accumulation
nonlinear PMS accumulation
altered timing & shape of RSL changes
Calendar age (yrs AD)
Calendar age (yrs AD)
(d)
Residual (cm)
Calendar age (yrs AD)
Calendar age (yrs AD)

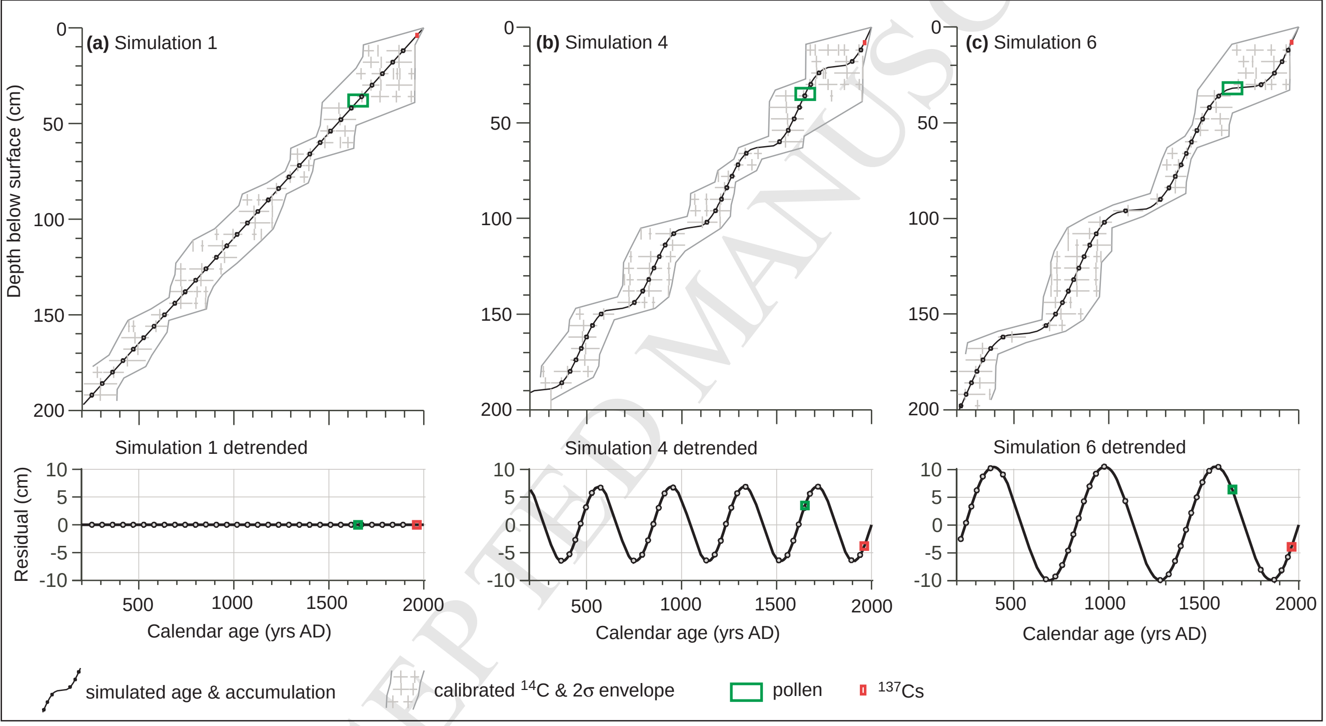

(a)
CANADA
U.S.A.
NL tide gauge
upland
granitic outcrop
core site
Railway / Amtrack
beach
salt marsh
(b)
Long Island Sound
North Atlantic
41°
72°
50 km
(b)
500 m
(c)
chenier dune
West
core site PXY
East
High marsh vegetation
sand
Substrate
-S. patens
-S. alterniflora (stunted)
-S. patens (brown)
-S. patens, D. spicata (dark brown)
-D. spicata (clayey, black/brown)
Elevation (m NGVD29)
1.2
0.8
0.4
0.0
-0.4
-0.8
-1.2
-1.6
-2.0
Distance (m)
0
5
10
15
20
25
30
35
40



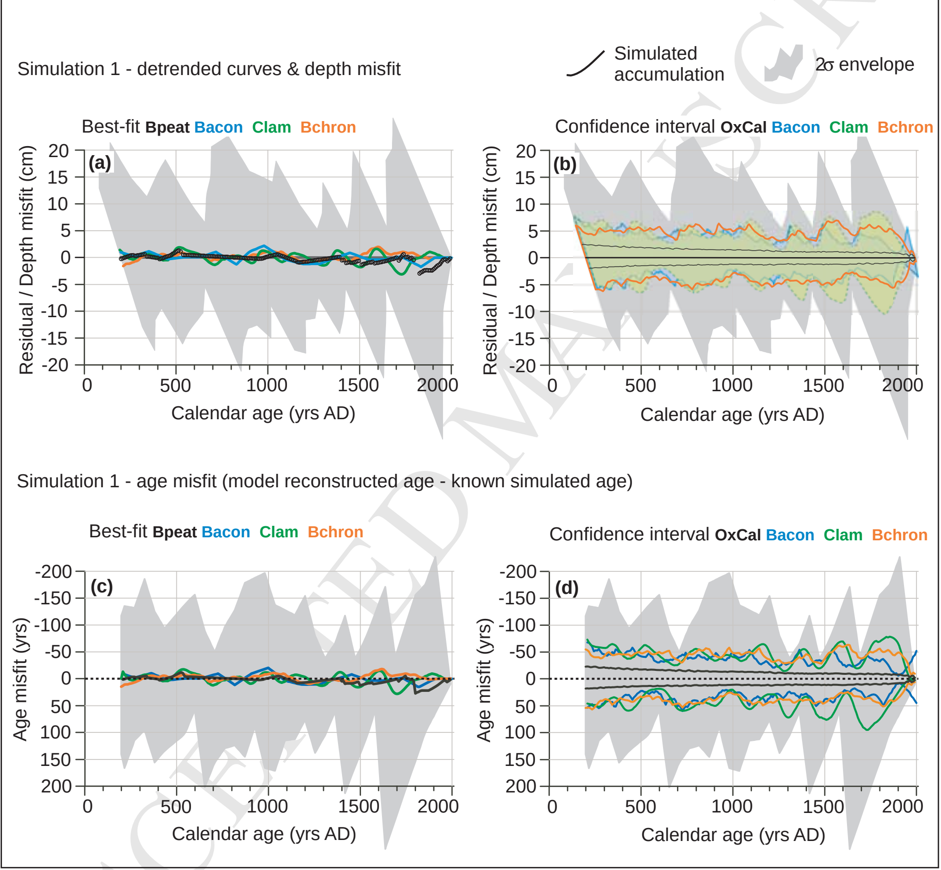
(a) Core PY linear 14C WMD
(b) Core PY 2σ calibrated
14C age (14C yrs)
IntCal09 (1σ & 2
14C age (1σ & 2
outlier
Calendar age (yrs AD)
14C WMD 1.1 mm/yr
13 cm
outlier
calibrated 14C
& 2σ envelope
PX pollen
137Cs
Depth in core (cm)

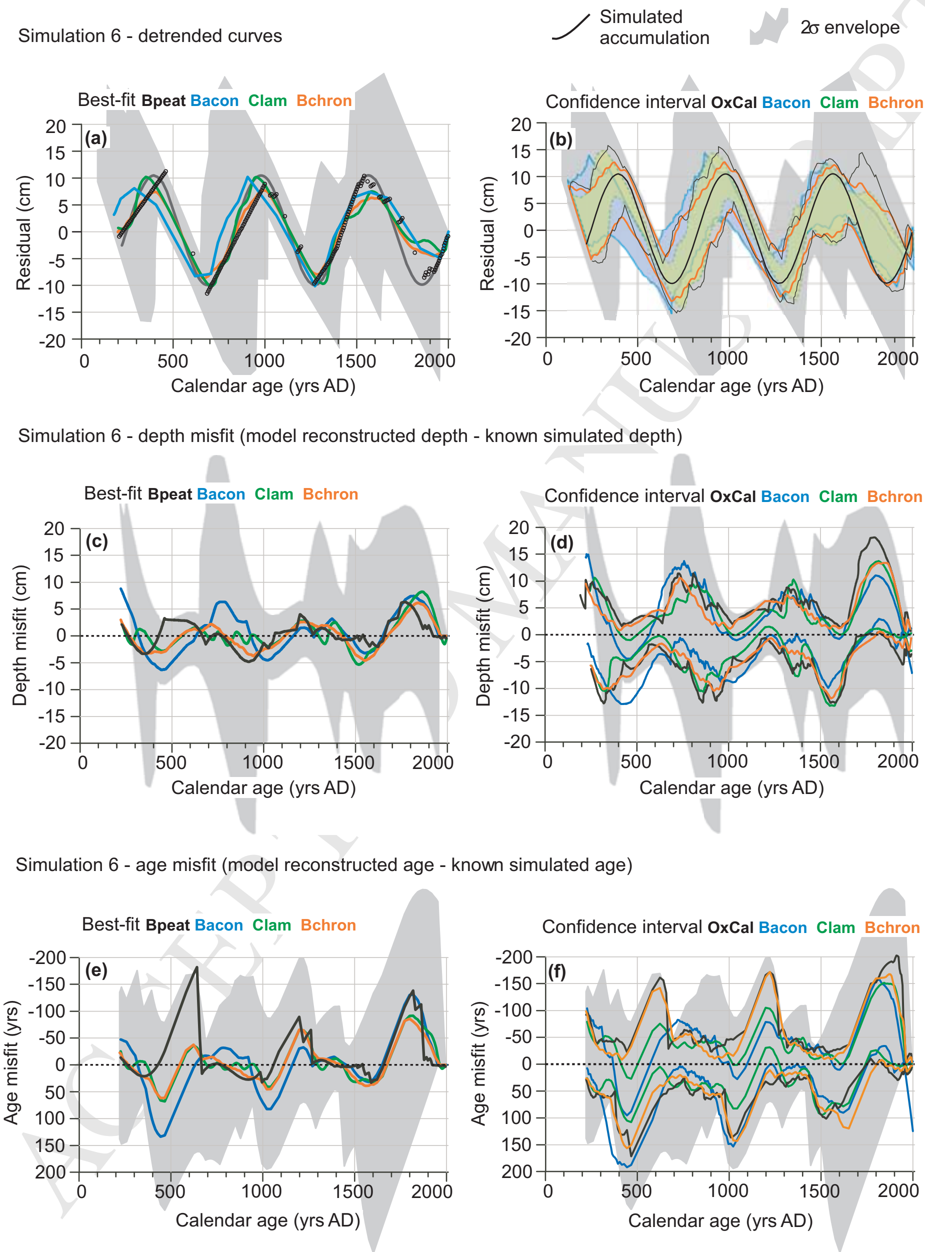


(a) Ambrosia (%)
0 10 20
(b) excess ^{210}Pb (Bq/Kg)
0 100 200 300
(d) ^{137}Cs (Bq/Kg)
0 20 40
(e) ^{241}Am (Bq/Kg)
0 0.5 1
(f) Calendar age (yrs AD)
1500 1600 1700 1800 1900 2000
Depth in core (cm)
0 10 20 30 40 50 60 70
150 yrs ^{210}Pb
CRS age model
^{137}Cs
13 cm
pollen marker
0.001 0.01 0.1 1 10 100
(c) LN excess ^{210}Pb (Bq/Kg)

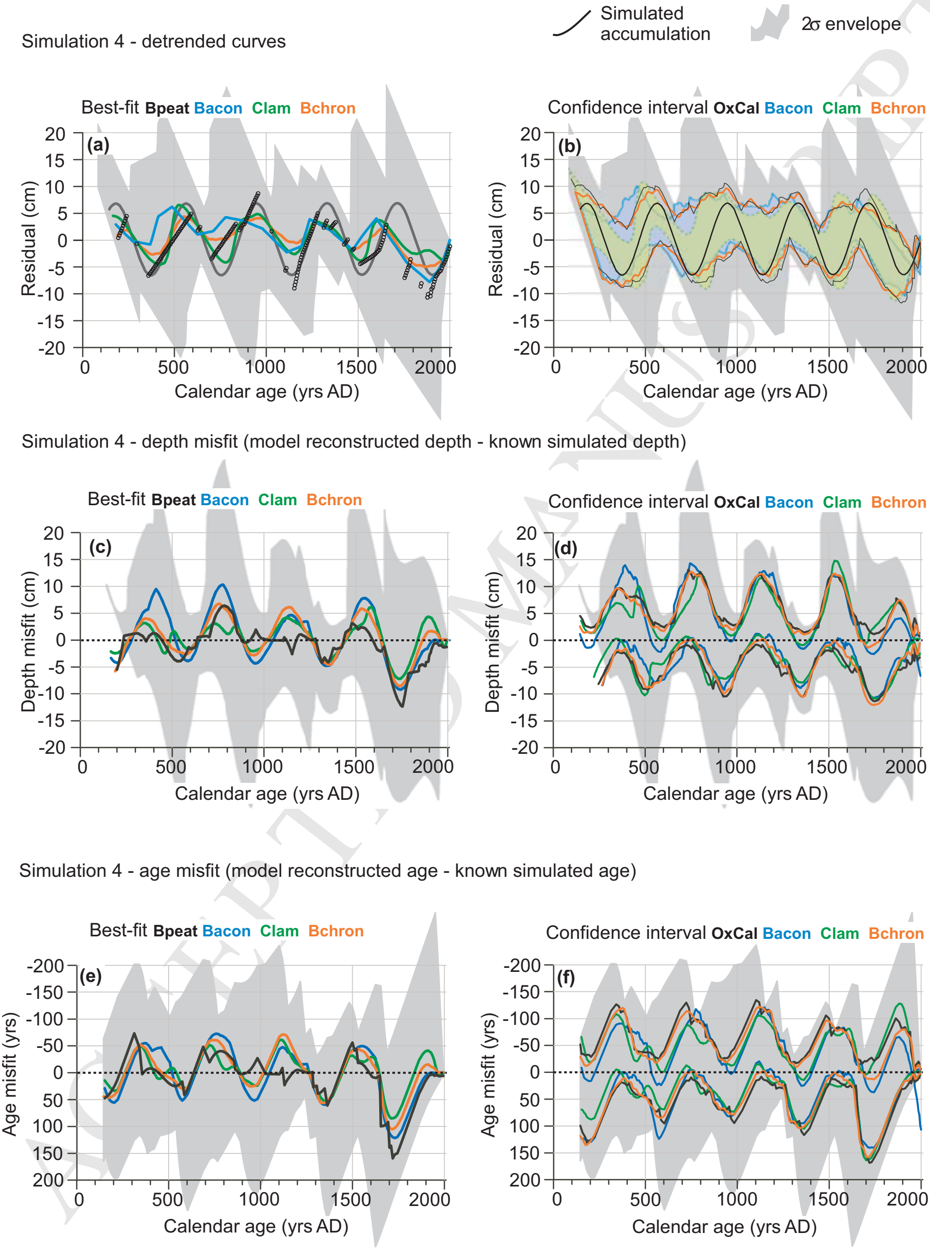


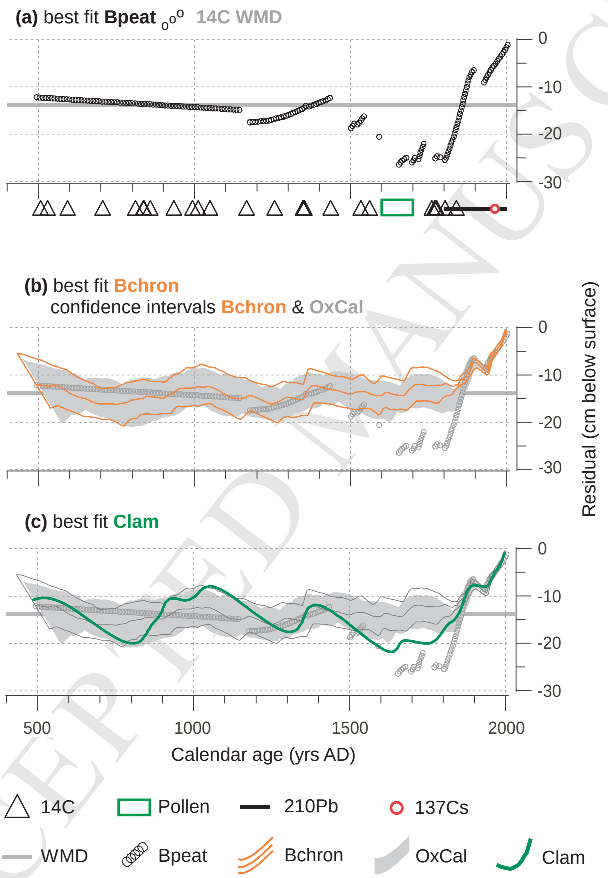
(a) Simulation 1
(b) Simulation 4
(c) Simulation 6
Depth below surface (cm)
0
50
100
150
200
Simulation 1 detrended
Simulation 4 detrended
Simulation 6 detrended
Residual (cm)
10
5
0
-5
-10
500
1000
1500
2000
Calendar age (yrs AD)
simulated age & accumulation
calibrated ^{14}C & 2σ envelope
pollen
^{137}Cs

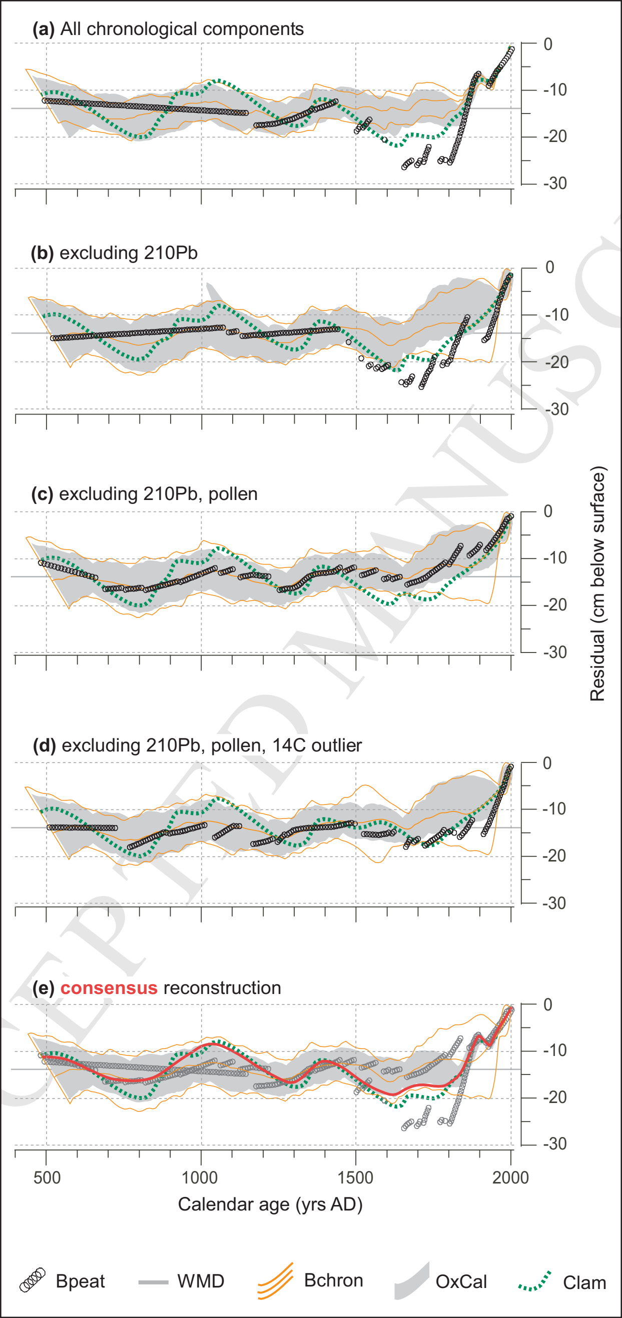

Simulation 1 - detrended curves & depth misfit
Simulated accumulation
2σ envelope
Best-fit Bpeat Bacon Clam Bchron
Confidence interval OxCal Bacon Clam Bchron
(a)
(b)
Residual / Depth misfit (cm)
20
15
10
5
0
-5
-10
-15
-20
0
500
1000
1500
2000
Calendar age (yrs AD)
Simulation 1 - age misfit (model reconstructed age - known simulated age)
Best-fit Bpeat Bacon Clam Bchron
Confidence interval OxCal Bacon Clam Bchron
(c)
(d)
Age misfit (yrs)
-200
-150
-100
-50
0
50
100
150
200
Calendar age (yrs AD)

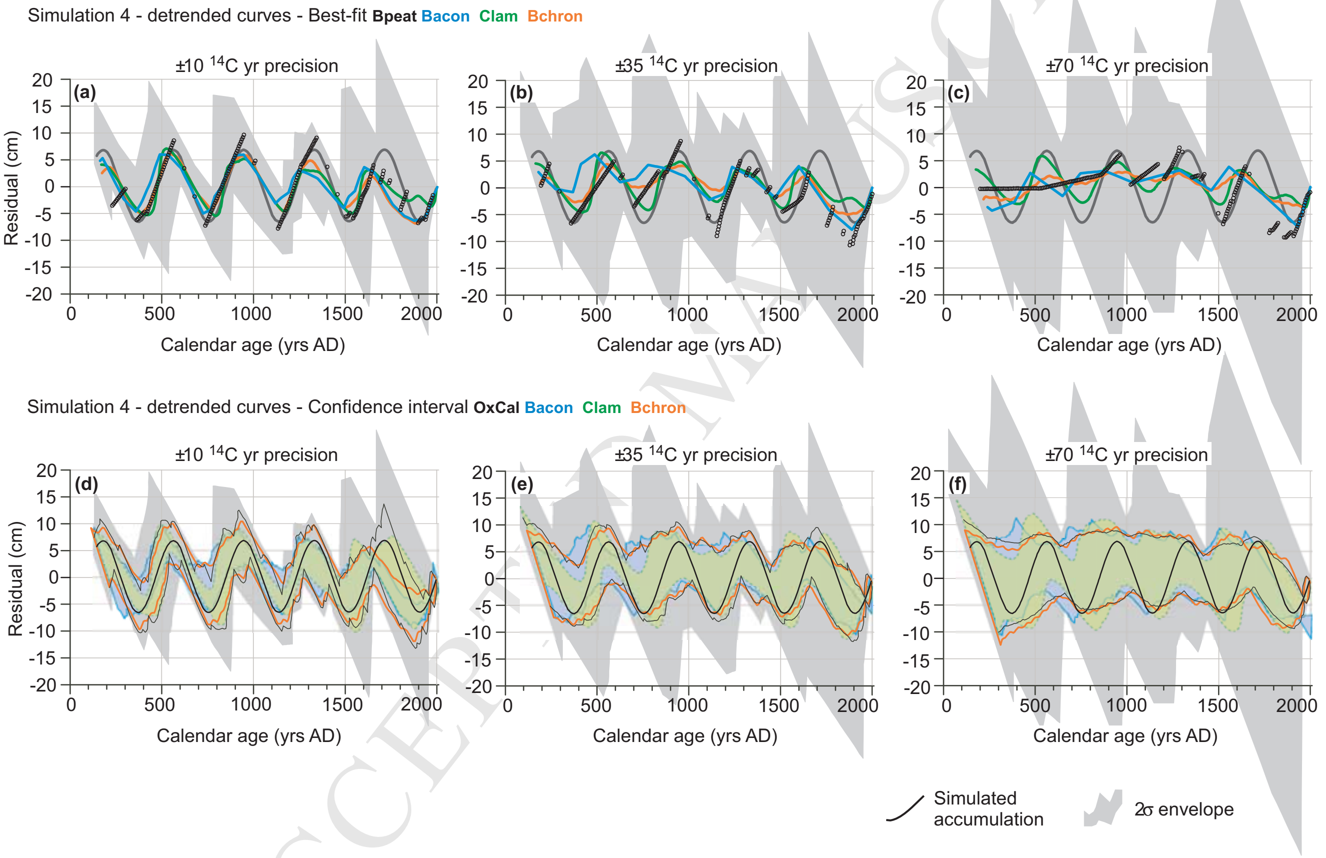


Simulation 6 - detrended curves
Simulated accumulation
2σ envelope
Best-fit Bpeat Bacon Clam Bchron
(a)
Residual (cm)
Calendar age (yrs AD)
Confidence interval OxCal Bacon Clam Bchron
(b)
Residual (cm)
Calendar age (yrs AD)
Simulation 6 - depth misfit (model reconstructed depth - known simulated depth)
Best-fit Bpeat Bacon Clam Bchron
(c)
Depth misfit (cm)
Calendar age (yrs AD)
Confidence interval OxCal Bacon Clam Bchron
(d)
Depth misfit (cm)
Calendar age (yrs AD)
Simulation 6 - age misfit (model reconstructed age - known simulated age)
Best-fit Bpeat Bacon Clam Bchron
(e)
Age misfit (yrs)
Calendar age (yrs AD)
Confidence interval OxCal Bacon Clam Bchron
(f)
Age misfit (yrs)
Calendar age (yrs AD)


Simulated accumulation
2σ envelope
Simulation 4 - detrended curves
Best-fit Bpeat Bacon Clam Bchron
(a)
Residual (cm)
Calendar age (yrs AD)
Confidence interval OxCal Bacon Clam Bchron
(b)
Residual (cm)
Calendar age (yrs AD)
Simulation 4 - depth misfit (model reconstructed depth - known simulated depth)
Best-fit Bpeat Bacon Clam Bchron
(c)
Depth misfit (cm)
Calendar age (yrs AD)
Confidence interval OxCal Bacon Clam Bchron
(d)
Depth misfit (cm)
Calendar age (yrs AD)
Simulation 4 - age misfit (model reconstructed age - known simulated age)
Best-fit Bpeat Bacon Clam Bchron
(e)
Age misfit (yrs)
Calendar age (yrs AD)
Confidence interval OxCal Bacon Clam Bchron
(f)
Age misfit (yrs)
Calendar age (yrs AD)

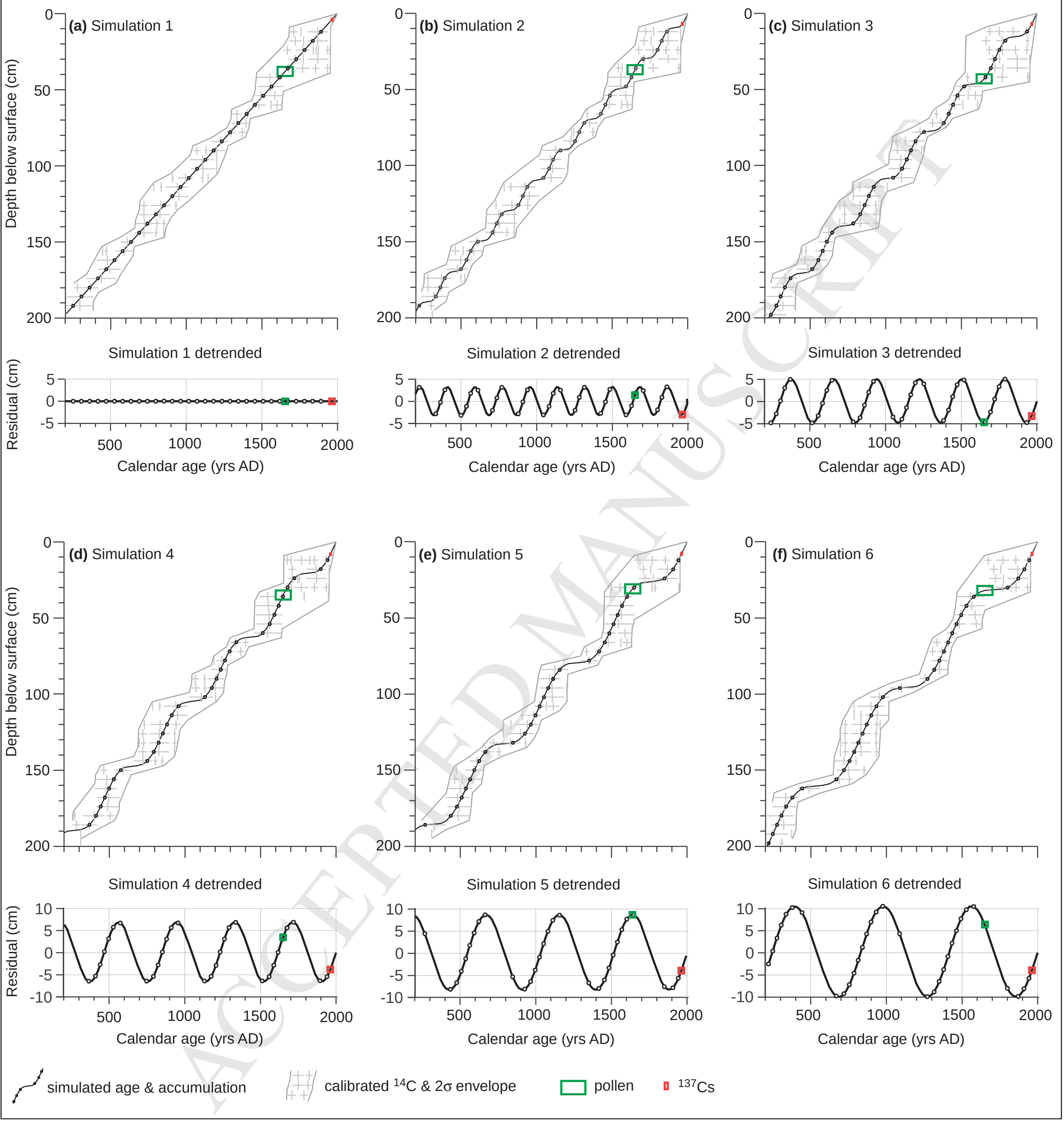

(a) best fit Bpeat 14C WMD
0
-10
-20
-30
(b) best fit Bchron
confidence intervals Bchron & OxCal
0
-10
-20
-30
Residual (cm below surface)
(c) best fit Clam
0
-10
-20
-30
500
1000
1500
2000
Calendar age (yrs AD)
14C
Pollen
210Pb
137Cs
WMD
Bpeat
Bchron
OxCal
Clam


(a) All chronological components
(b) excluding 210Pb
(c) excluding 210Pb, pollen
(d) excluding 210Pb, pollen, 14C outlier
(e) consensus reconstruction
0
-10
-20
-30
Residual (cm below surface)
500
1000
1500
2000
Calendar age (yrs AD)
Bpeat
WMD
Bchron
OxCal
Clam



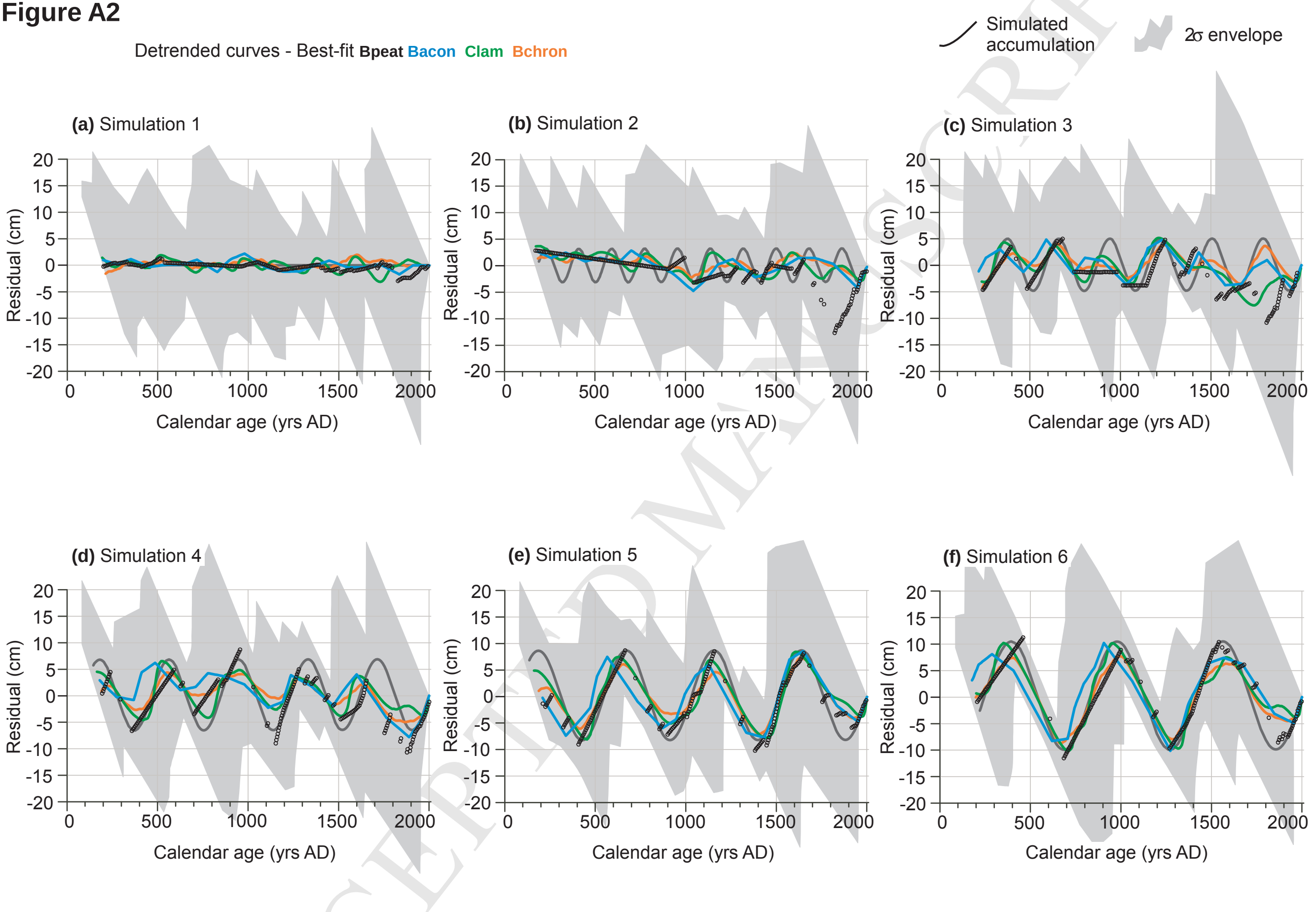
Simulation 4 - detrended curves - Best-fit Bpeat Bacon Clam Bchron
±10 14C yr precision
±35 14C yr precision
±70 14C yr precision
(a)
(b)
(c)
Residual (cm)
Calendar age (yrs AD)
Simulation 4 - detrended curves - Confidence interval OxCal Bacon Clam Bchron
(d)
(e)
(f)
Simulated accumulation
2σ envelope

Figure A1



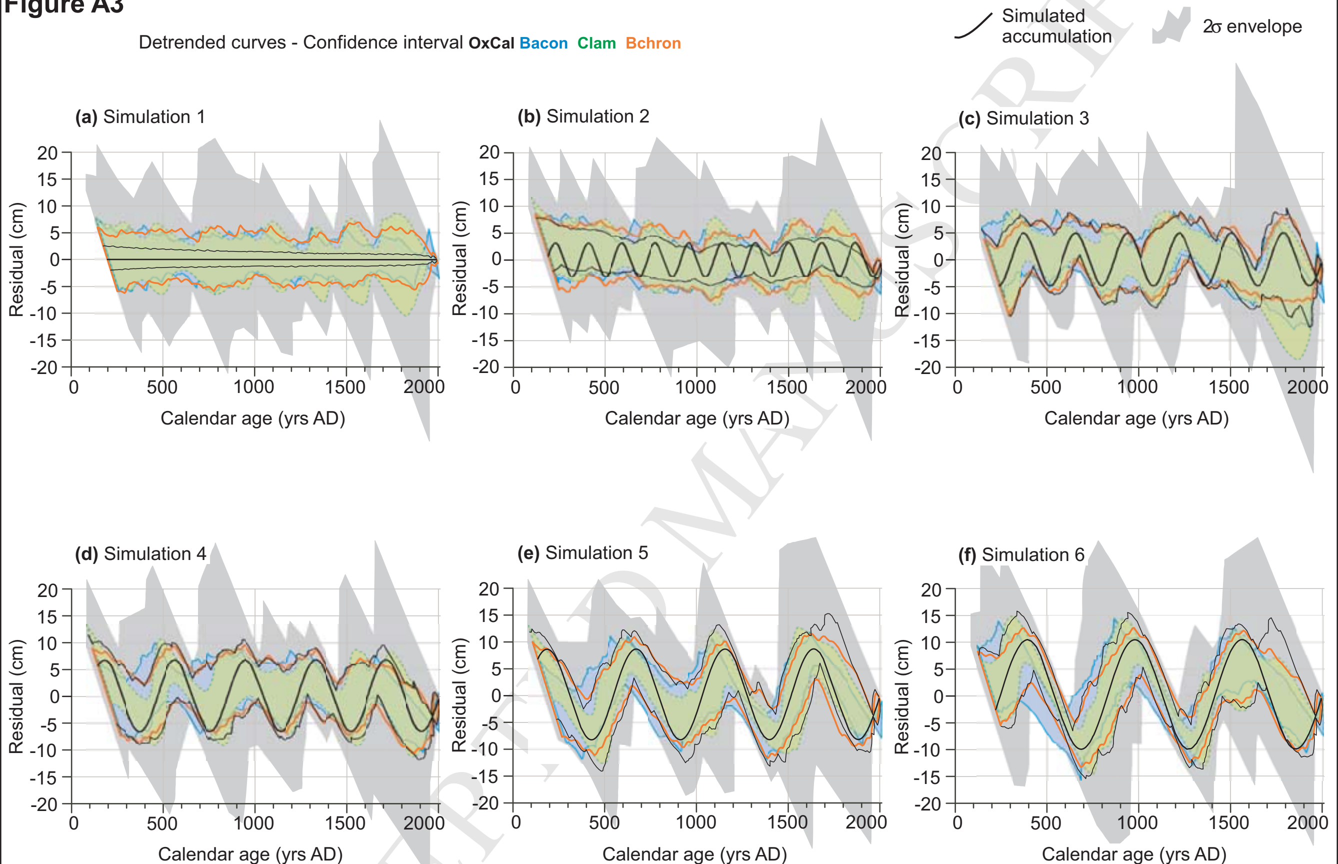
(a-f) 2σ calibrated and detrended 14C palaeomarsh surface accumulation simulations 1 to 6 and associated calibrated 14C age-depth envelope limited to the period 200-2000 yrs AD in this illustration for (a) linear and (b-f) nonlinear sinusoid variability tailored to cores PX and PY: GIA subsidence (0.11 cm/yr), down-core sampling (6 cm), age markers (pollen, 137Cs, surface), –35 ^{14}C yrs (1σ) average 14C measurement precision. Magnitude of trough-to-peak variability is close to the maximum allowed by the available accommodation space which is a combination of GIA subsidence (0.11 cm/yr) and peak-to-peak time interval for each simulation. (d) Simulation 4 nonlinear acceleration is equivalent to cores PXY modern acceleration

Figure A2



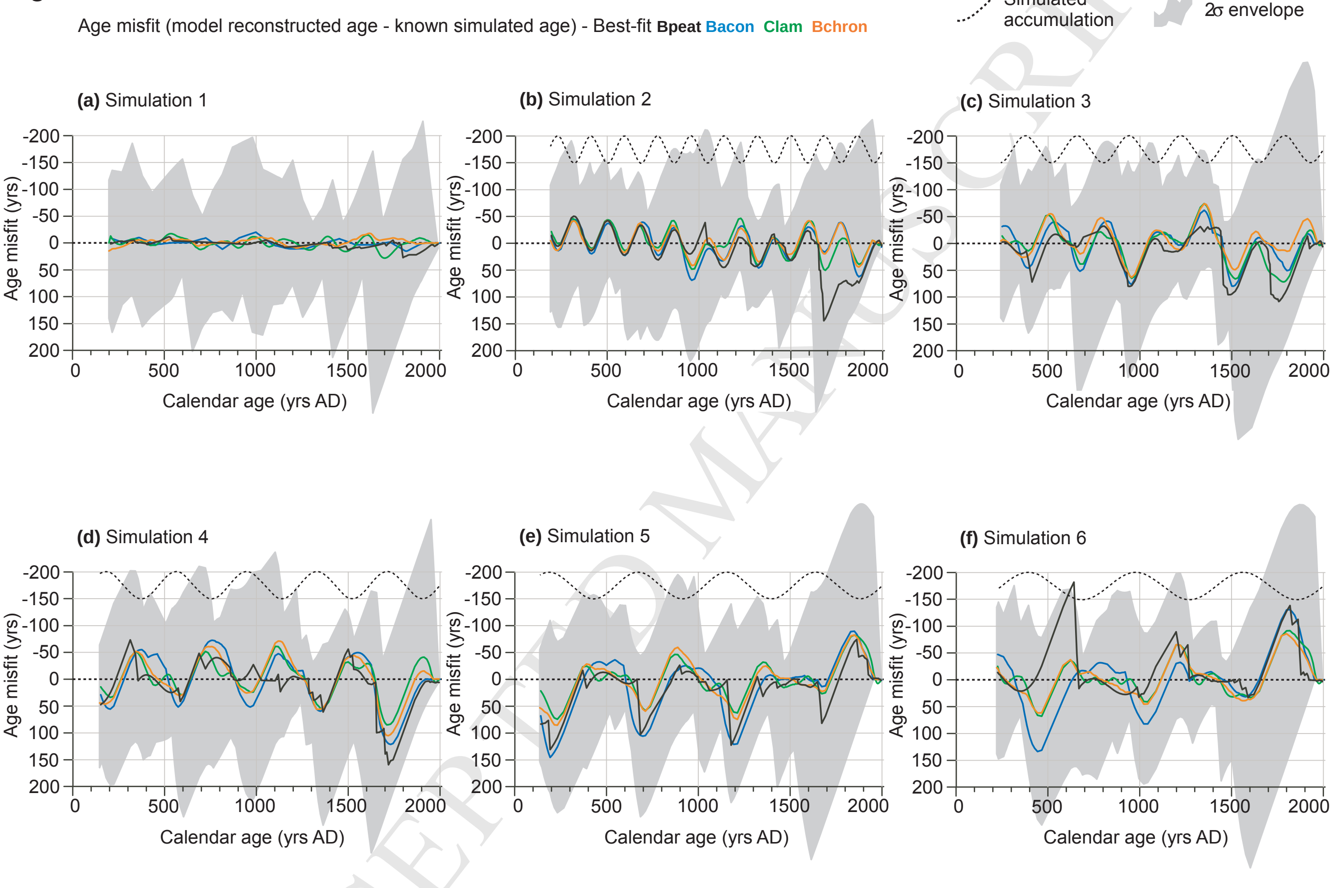
(a-f) Detrended curves (–35 ^{14}C yr precision) best fit model results grouped to compare the influence of calibration/model related artifacts (a Simulation 1) and success at predicting nonlinear palaeomarsh surface (PMS) accumulation (b-f Simulation 2 to 6). Black line represents known accumulation; age-depth envelope (grey shade, Y-axis not scaled to fit these due to excessive space requirements) encompasses individually calibrated 14C, Bpeat (black circles, mean of 3 runs using 15 sections), Bacon (blue line, mean of 3 runs), Clam (green line, 100,000 iterations using spline width 0.3), Bchron (orange line, mean of 3 standard runs). Bpeat results are represented by individual maximum a posteriori (MAP), Bacon the average MAP with step size 10 cm for 14C precision 35 yrs (–1σ), Clam smoothing spline individual run weighted-mean, Bchron mean average of the mode (50%).

Figure A3



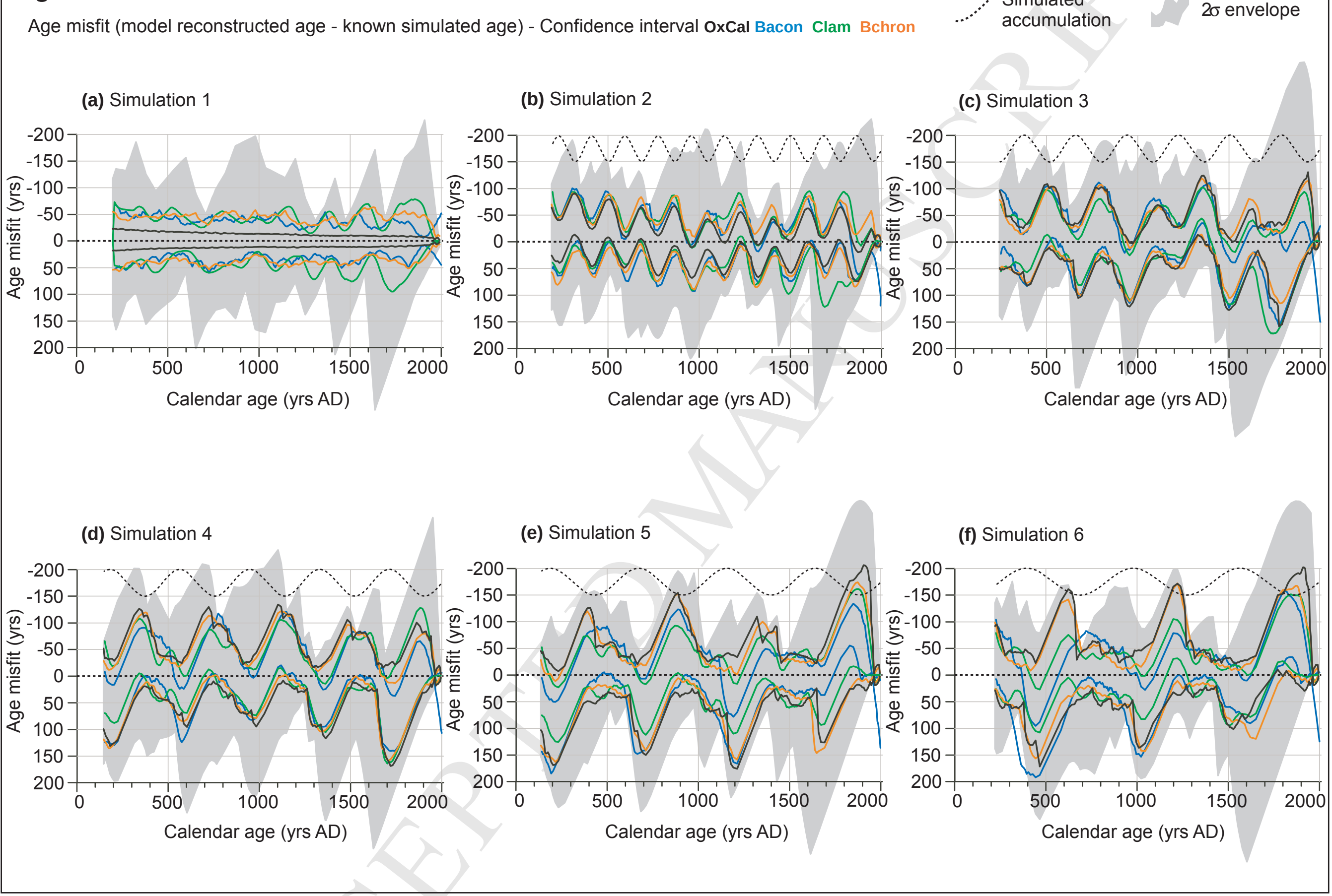
(a-f) Detrended curves (±35 ^{14}C yr precision) 95% confidence interval (CI) model results grouped to compare model success at constraining linear (a Simulation 1) and nonlinear (b-f Simulation 2 to 6) palaeomarsh surface (PMS) accumulation. Black line represents known accumulation; age-depth envelope (grey shade, Y-axis not scaled to fit these due to excessive space requirements) encompasses individually calibrated 14C only, Bacon (blue envelope, mean of 3 runs), Clam (green envelope, 100,000 iterations using spline width 0.3), Bchron (orange lines, mean of 3 standard runs), OxCal (thin black lines, mean of 3 runs, P_Sequence K=2 auto, General outlier model. Bacon results are represented by the 95% probability intervals (PI) with step size 10 cm for 14C precision of 35 yrs (±1σ), Clam by the 95% confidence intervals (CI), Bchron by the 95% highest posterior density region (HDR defined between 2.5% and 97.5%), OxCal by the 95% highest probability density range (HPD defined between from and to 95.4%).

Figure A4



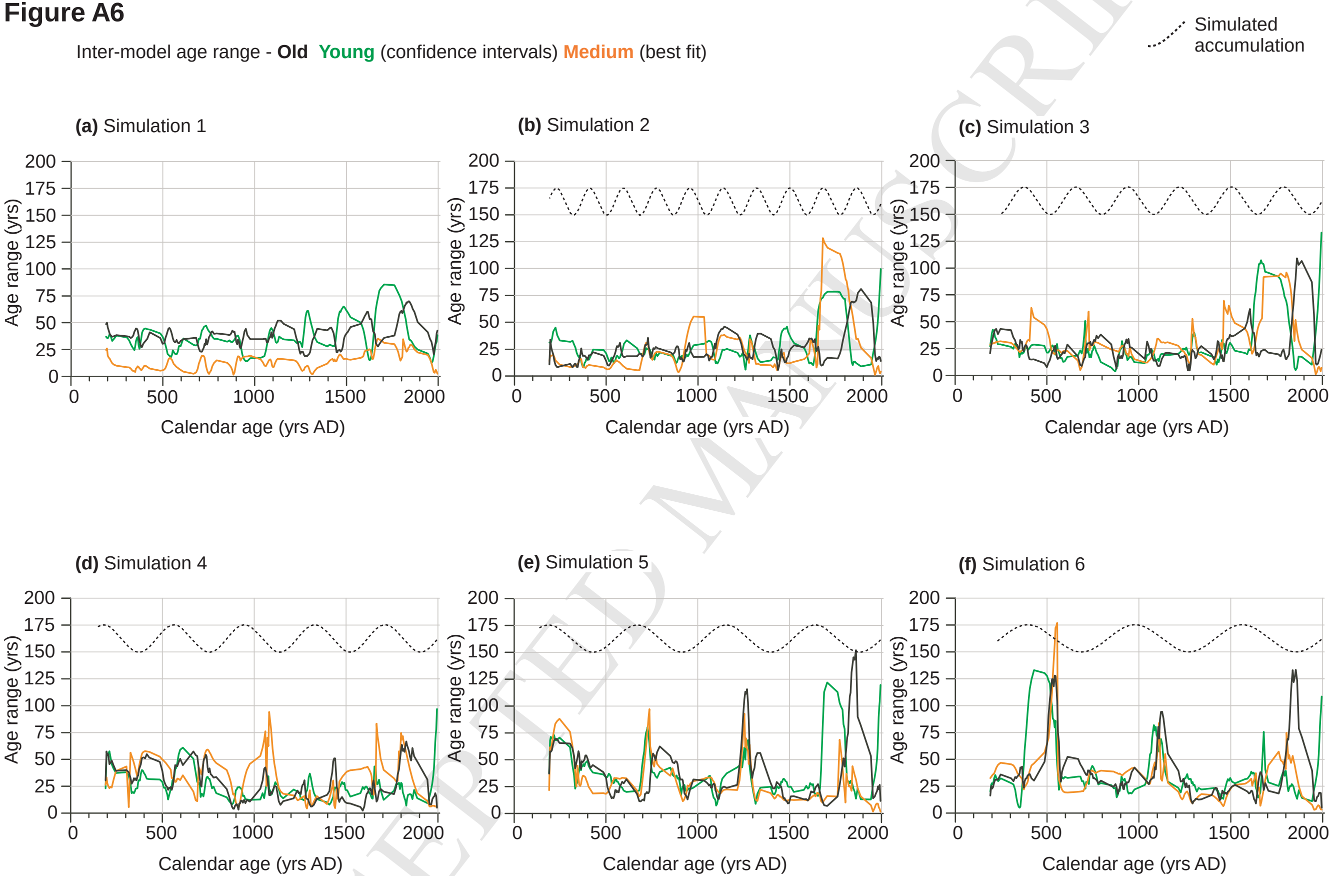
(a-f) Age misfit (model reconstructed age - known simulated age, –35 ^{14}C yr precision) for best-fit model results grouped to compare the influence of calibration/model related artifacts (a Simulation 1) and success at predicting nonlinear palaeomarsh surface (PMS) accumulation (b-f Simulation 2 to 6). Black dashed line represents known accumulation; age-depth envelope (grey shade, Y-axis not scaled to fit these due to excessive space requirements) encompasses individually calibrated 14C, Bpeat (black line, mean of 3 runs using 15 sections), Bacon (blue line, mean of 3 runs), Clam (green line, 100,000 iterations using spline width 0.3), Bchron (orange line, mean of 3 standard runs). Bpeat results are represented by individual maximum a posteriori (MAP), Bacon the average MAP with step size 10 cm for 14C precision 35 yrs (–1σ), Clam smoothing spline individual run weighted-mean, Bchron mean average of the mode (50%).

Figure A5



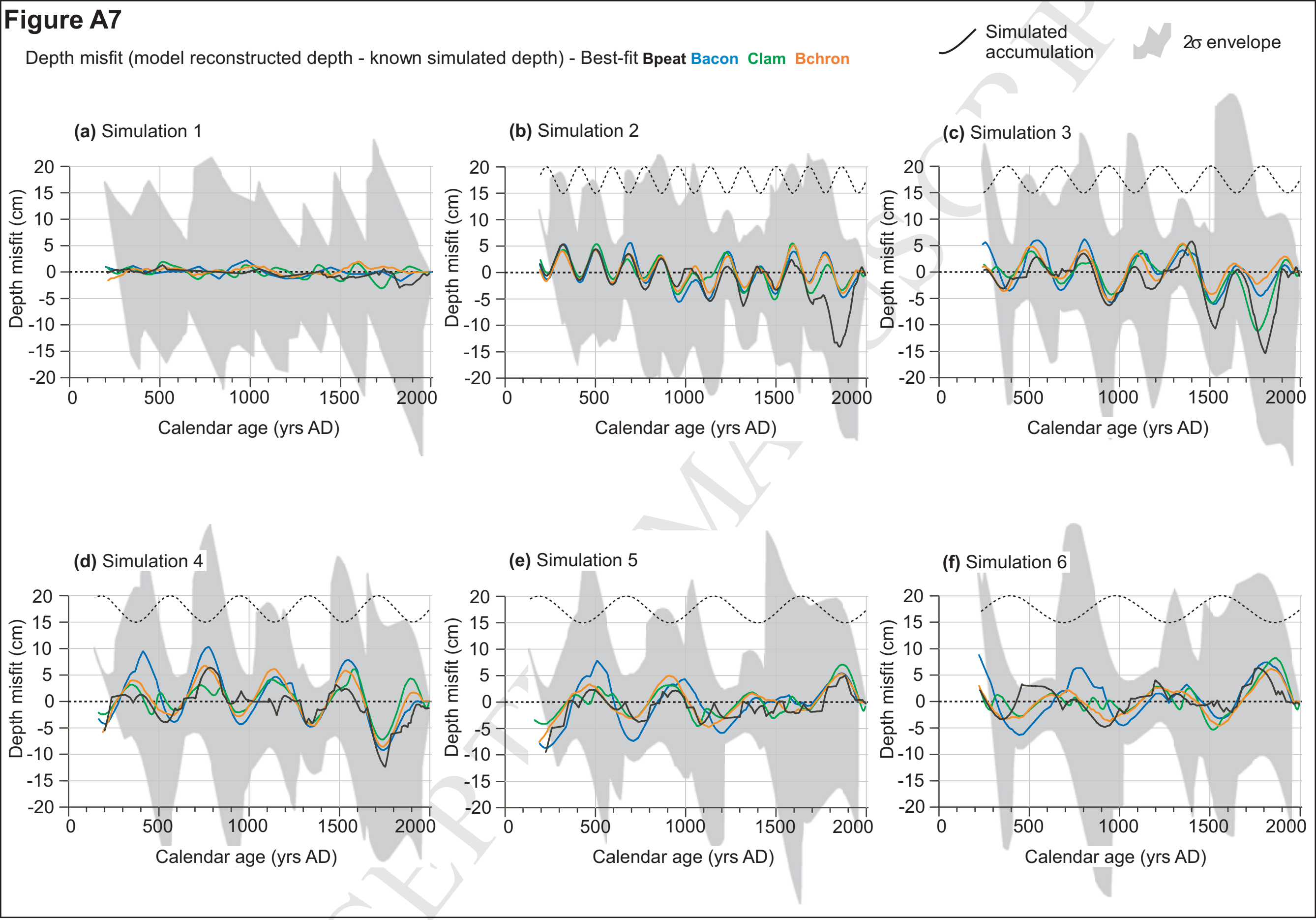
(a-f) Age misfit (model reconstructed age - known simulated age, –35 ^{14}C yr precision) –95% confidence interval (CI) model results grouped to compare model success at constraining linear (a Simulation 1) and nonlinear (b-f Simulation 2 to 6) palaeomarsh surface (PMS) accumulation. NOTE - when any CI envelope crosses the zero line (black dashed) it has no longer successfully constrained the simulated age-depth sequence. Black line dashed line represents known accumulation; age-depth envelope (grey shade, Y-axis not scaled to fit these due to excessive space requirements) encompasses individually calibrated 14C only, Bacon (blue lines, mean of 3 runs), Clam (green lines, 100,000 iterations using spline width 0.3), Bchron (orange lines, mean of 3 standard runs), OxCal (black lines, mean of 3 runs, P_Sequence K=2 auto, General outlier model. Bacon results are represented by the 95% probability intervals (PI) with step size of 10 cm for 14C precision of 35 yrs (–1σ), Clam by the 95% confidence intervals (CI), Bchron by the 95% highest posterior density region (HDR defined between 2.5% and 97.5%), OxCal by the 95% highest probability density range (HPD defined between from and to 95.4%).

Figure A6

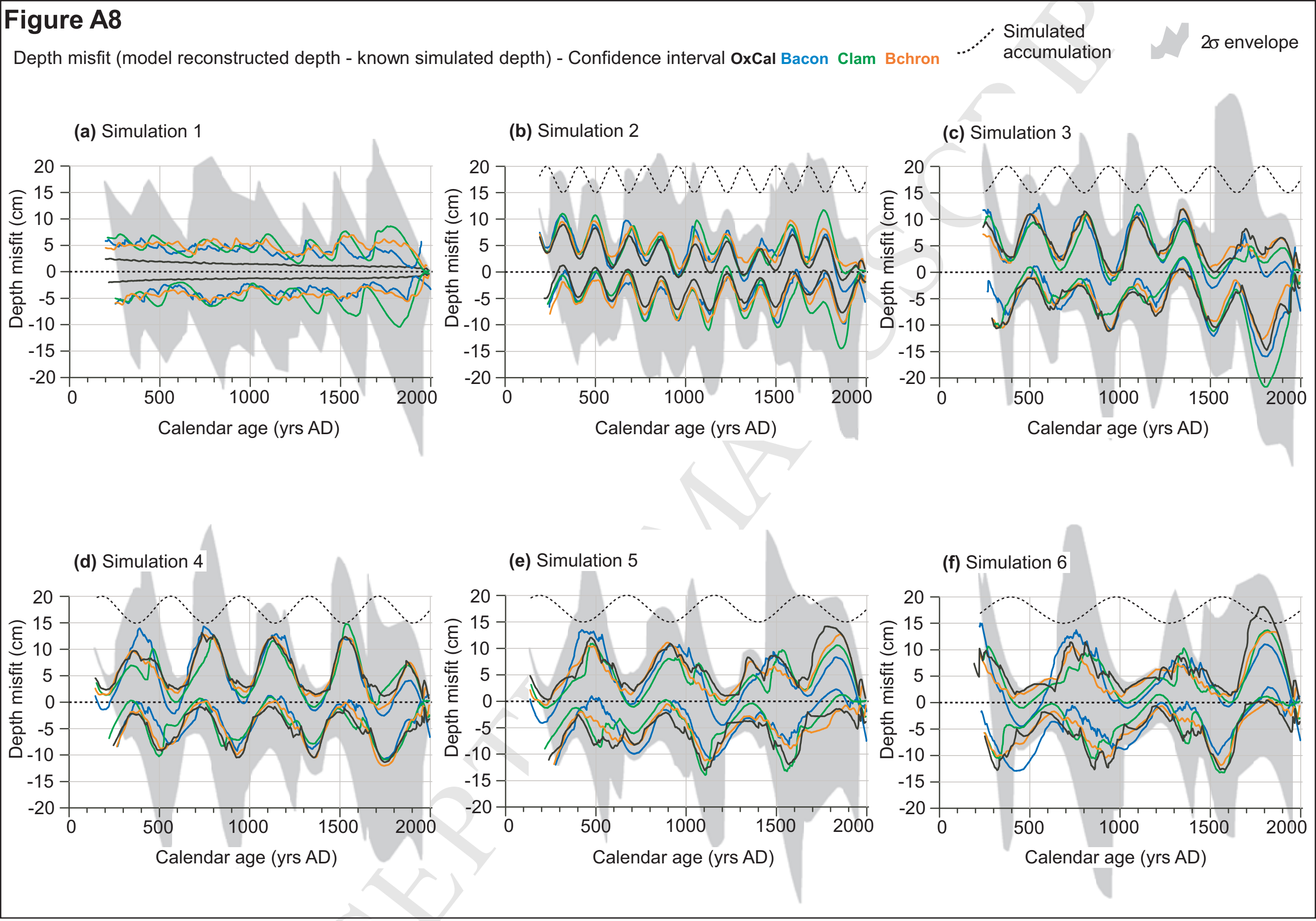


(a-f) Inter-model age range –35 ^{14}C yr precision (youngest - oldest, all models to capture maximum range) for Bpeat (mean of 3 runs using 15 sections), Bacon (mean of 3 runs), Clam (100,000 iterations using spline width 0.3), Bchron (mean of 3 standard runs). Bpeat results are represented by individual maximum a posteriori (MAP), Bacon the average MAP with step size 10 cm for 14C precision 35 yrs (–1σ), Clam smoothing spline individual run weighted-mean, Bchron mean average of the mode (50%).

Figure A7

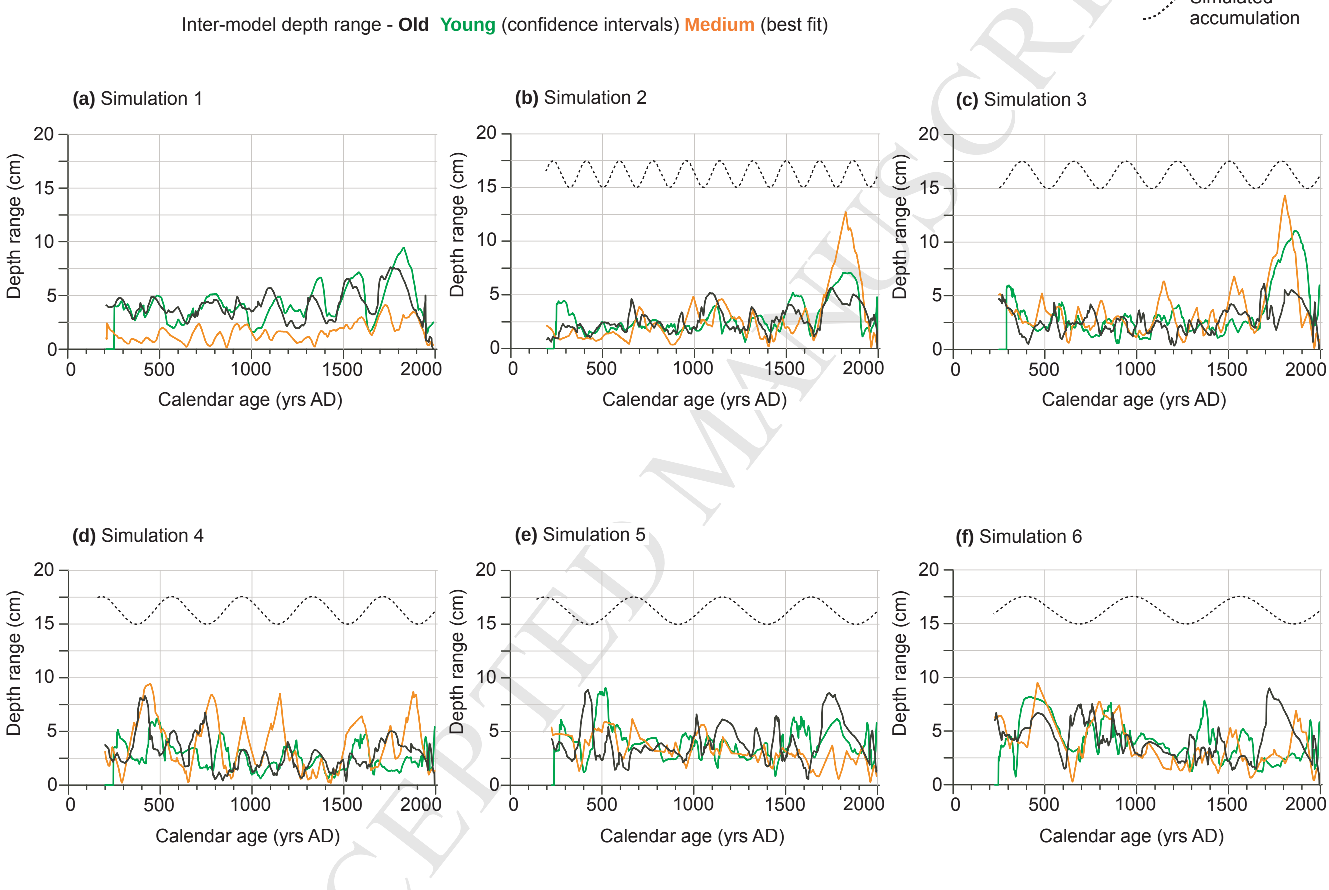


(a-f) Depth misfit (model reconstructed depth - known simulated depth, ±35 ^{14}C yr precision) for 'best-fit model results grouped to compare the influence of calibration/model related artifacts (a Simulation 1) and success at predicting nonlinear palaeomarsh surface (PMS) accumulation (b-f Simulation 2 to 6). Black dashed line represents known accumulation; age-depth envelope (grey shade, Y-axis not scaled to fit these due to excessive space requirements) encompasses individually calibrated 14C, Bpeat (black line, mean of 3 runs using 15 sections), Bacon (blue line, mean of 3 runs), Clam (green line, 100,000 iterations using spline width 0.3), Bchron (orange line, mean of 3 standard runs). Bpeat results are represented by individual maximum a posteriori (MAP), Bacon the average MAP with step size 10 cm for 14C precision 35 yrs (±1σ), Clam smoothing spline individual run weighted-mean, Bchron mean average of the mode (50%).



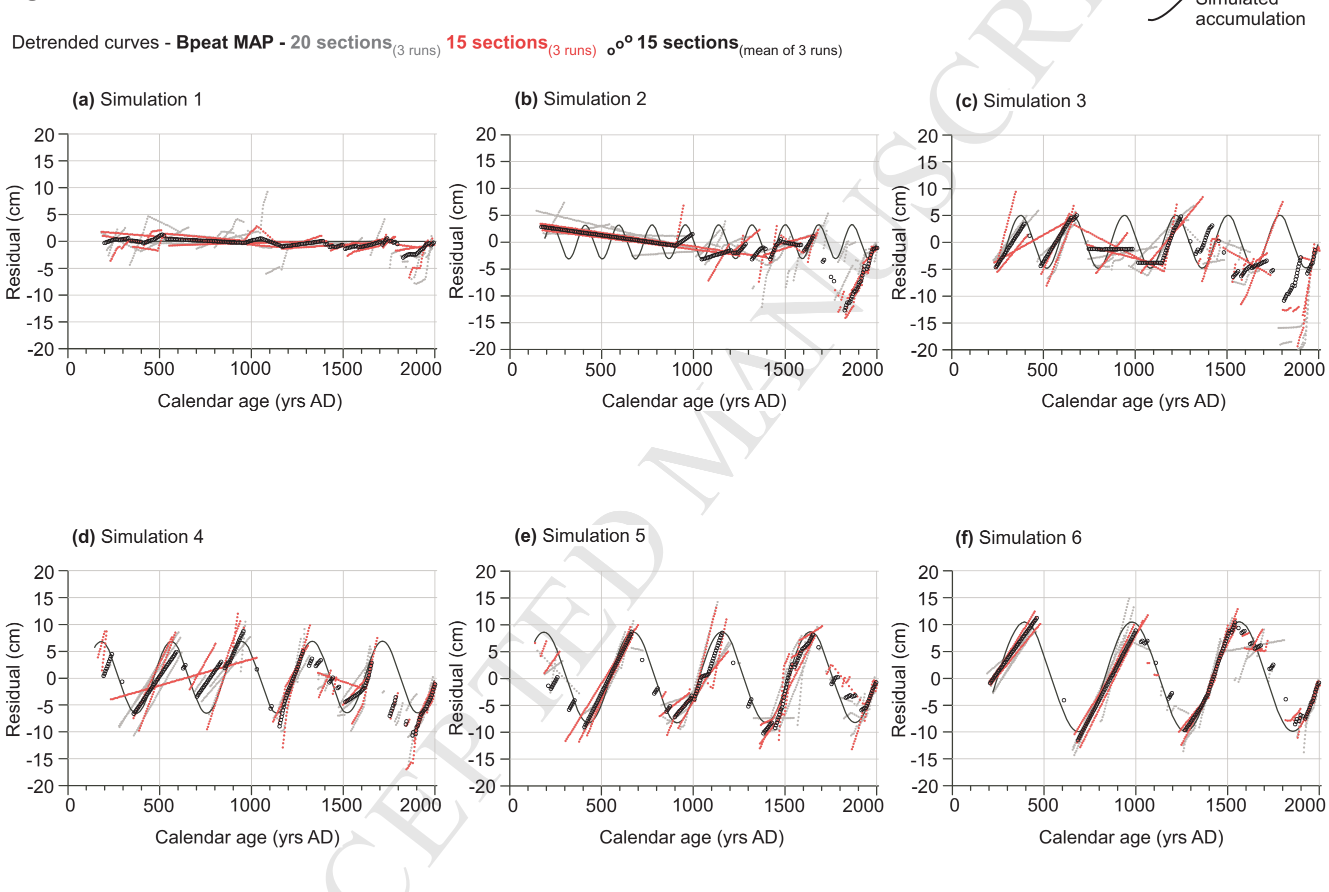
(a-f) Depth misfit (model reconstructed depth - known simulated depth, ±35 ^{14}C yr precision) for ±95% confidence interval (CI) model results grouped to compare model success at constraining linear (a Simulation 1) and nonlinear (b-f Simulation 2 to 6) palaeomarsh surface (PMS) accumulation. NOTE - when any CI envelope crosses the zero line (black dashed) it has no longer successfully constrained the simulated age-depth sequence. Black line dashed line represents known accumulation; age-depth envelope (grey shade, Y-axis not scaled to fit these due to excessive space requirements) encompasses individually calibrated 14C only, Bacon (blue lines, mean of 3 runs), Clam (green lines, 100,000 iterations using spline width 0.3), Bchron (orange lines, mean of 3 standard runs), OxCal (black lines, mean of 3 runs, P_Sequence K=2 auto, General outlier model. Bacon results are represented by the 95% probability intervals (PI) with step size of 10 cm for 14C precision of 35 yrs (±1σ), Clam by the 95% confidence intervals (CI), Bchron by the 95% highest posterior density region (HDR defined between 2.5% and 97.5%), OxCal by the 95% highest probability density range (HPD defined between from and to 95.4%).

Figure A9



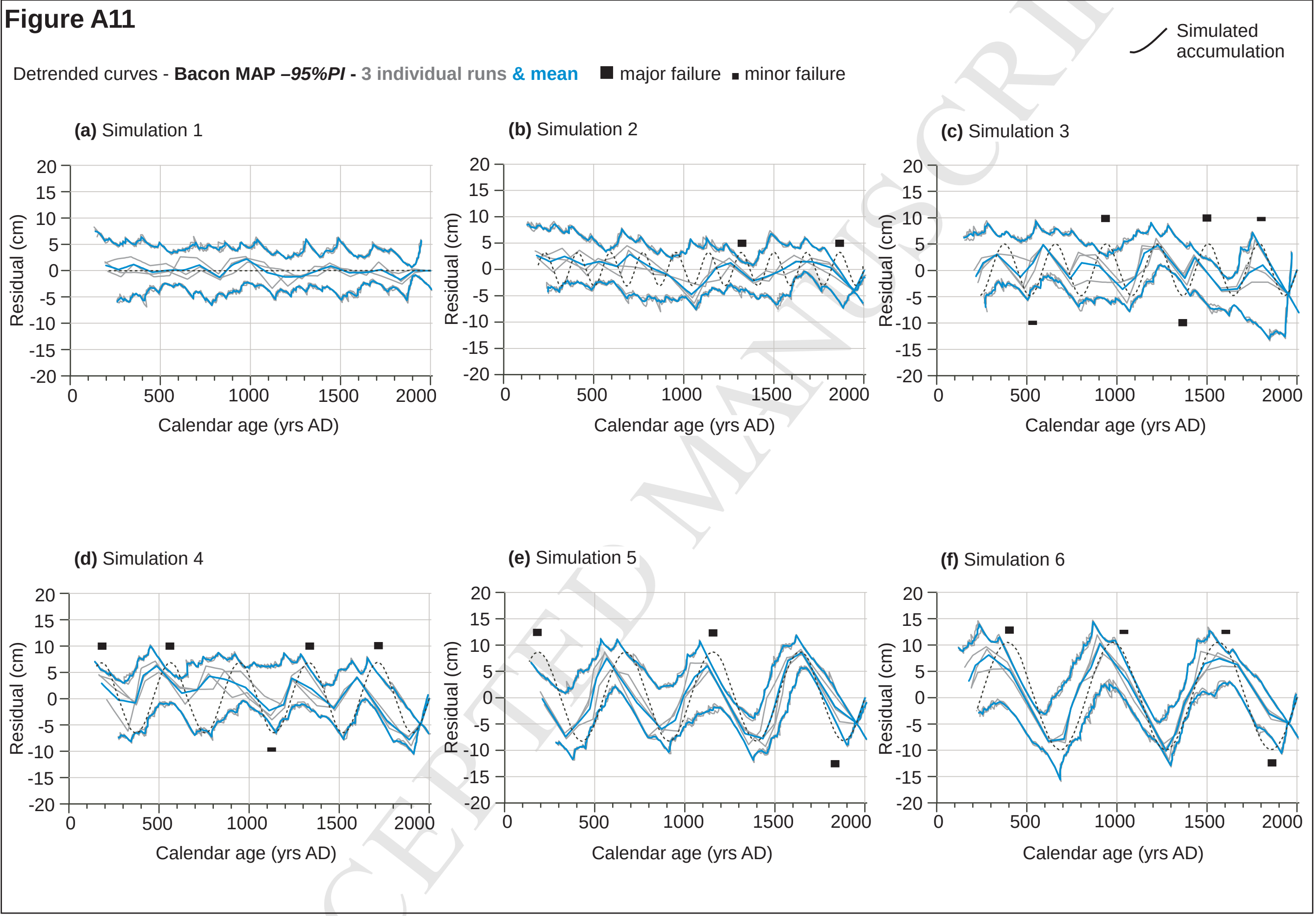
(a-f) Inter-model depth range –35 ^{14}C yr precision (smallest - largest, all models to capture maximum range) for Bpeat (mean of 3 runs using 15 sections), Bacon (mean of 3 runs), Clam (100,000 iterations using spline width 0.3), Bchron (mean of 3 standard runs). Bpeat results are represented by individual maximum a posteriori (MAP), Bacon the average MAP with step size 10 cm for 14C precision 35 yrs (–1σ), Clam smoothing spline individual run weighted-mean, Bchron mean average of the mode (50%).

Figure A10



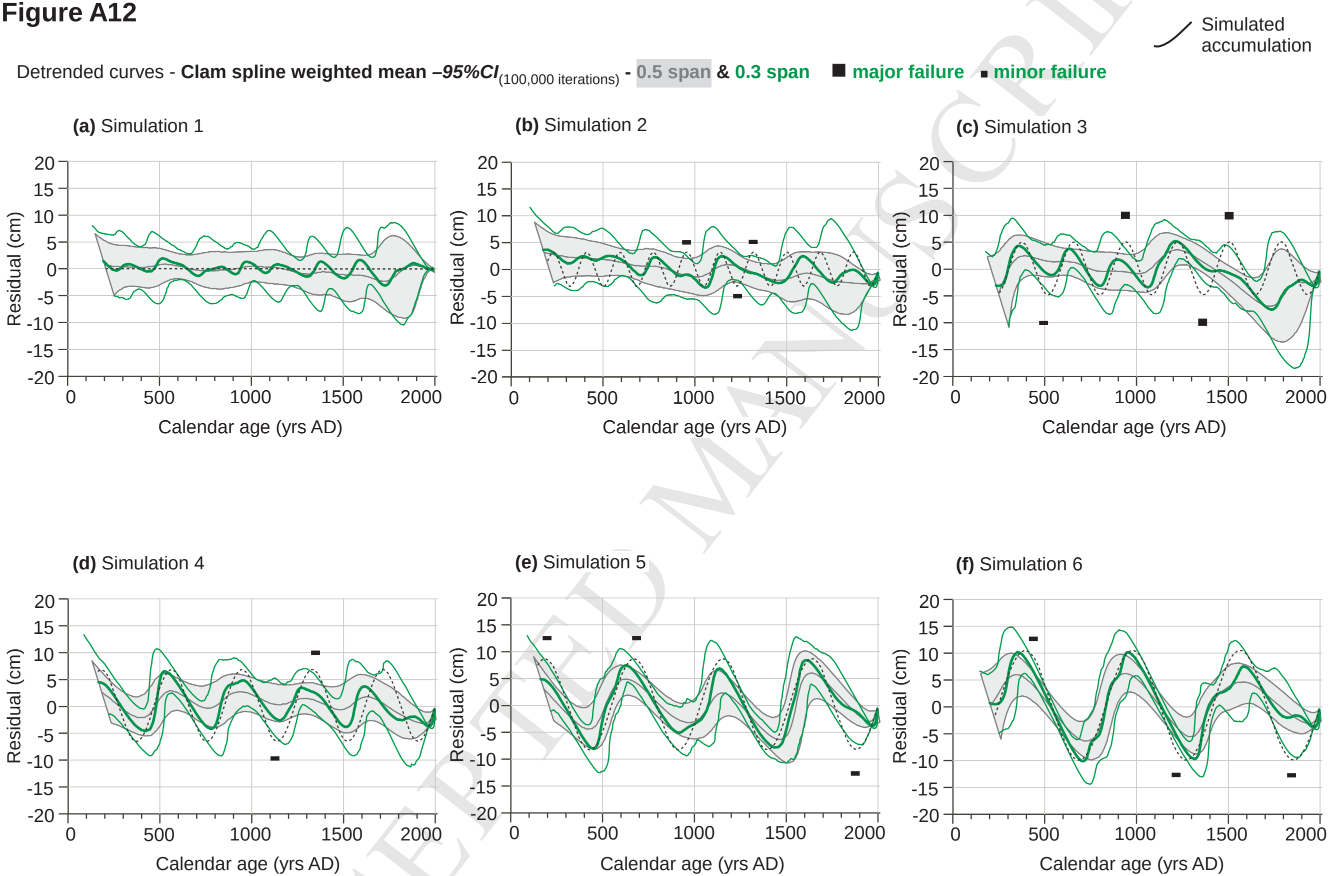
(a-f) Bpeat detrended curves (±35 ^{14}C yr precision) best fit maximum a posteriori (MAP) results for 3 runs of 15 and 20 sections, illustrate the sensitivity for incorporating calibration artefacts (linear) and allow qualitative judgement of the success with which nonlinear (sinusoidal) palaeomarsh surface accumulation has been reconstructed.

Figure A11



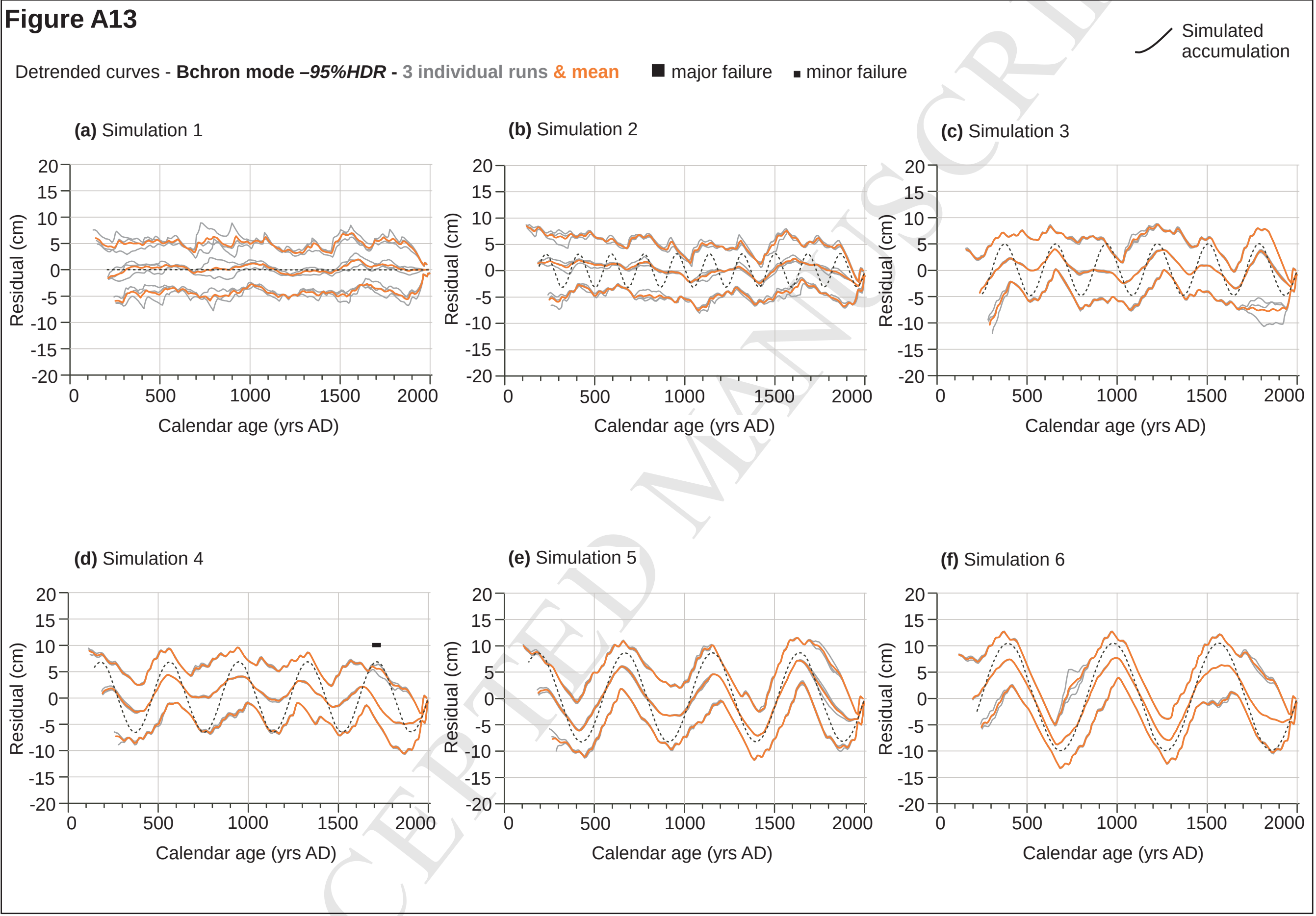
(a-f) Bacon detrended curves (–35 ^{14}C yr precision) best fit maximum a posteriori (MAP) results with 95% probability intervals (PI) and mean summaries, illustrate the sensitivity for incorporating calibration artefacts (linear) and allow qualitative judgement of the success with which the MAP has reconstructed nonlinear (sinusoidal) palaeomarsh surface accumulation and whether probability intervals have fully contained it (black cube - clear excursion, black line - minor excursion).

Figure A12



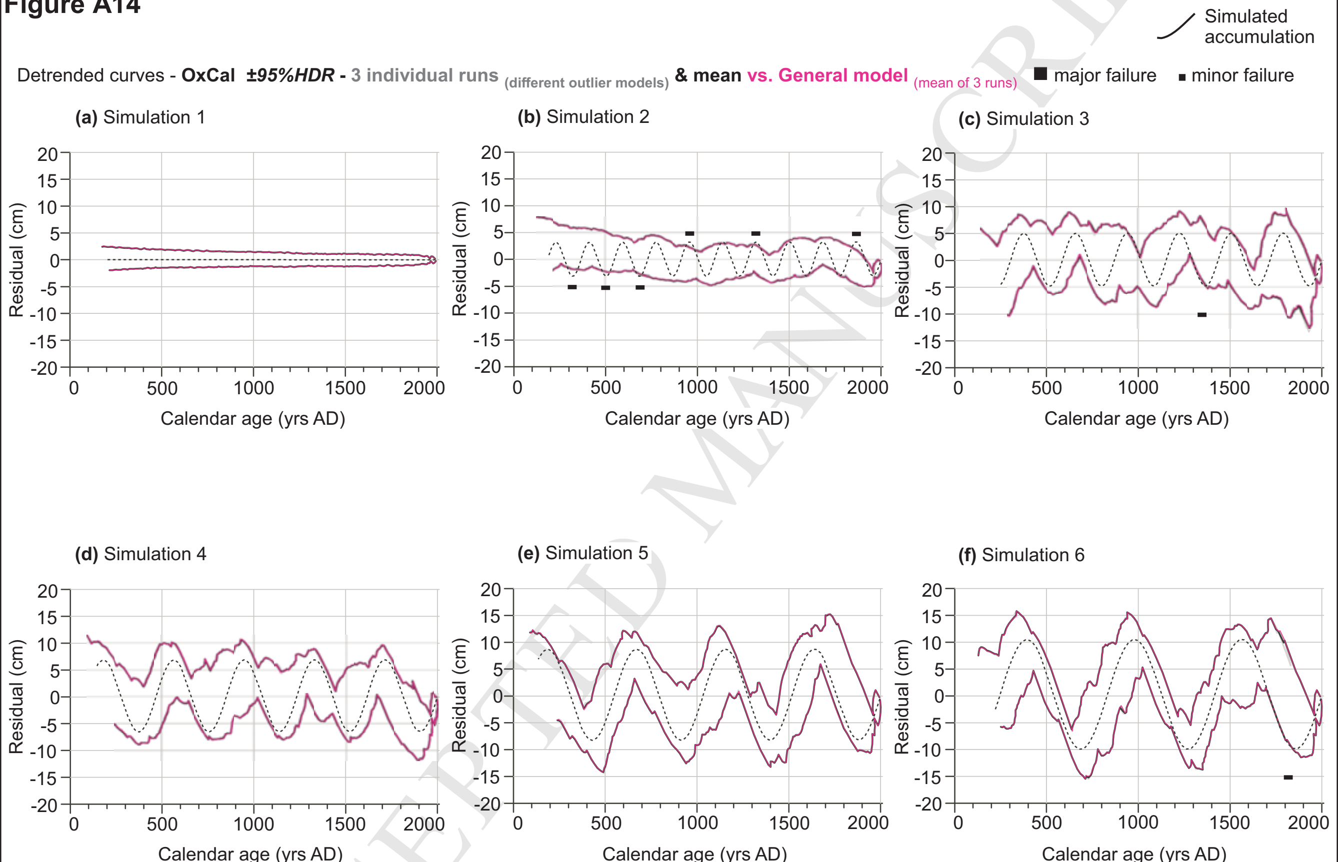
(a-f) Clam detrended curves (–35 ^{14}C yr precision) smooth spline 0.3 and 0.5 span best fit weighted mean results with 95% confidence intervals (CI) and mean summaries, illustrate the sensitivity for incorporating calibration artefacts (linear) and allow qualitative judgement of the success with the 0.3 weighted mean has reconstructed nonlinear (sinusoidal) palaeomarsh surface accumulation and whether confidence intervals have fully contained it (black cube - clear excursion, black line - minor excursion). Span of 0.3 is clearly more sensitive than 0.5, both vastly lower than the programme default 0.75 (not illustrated).

Figure A13



(a-f) Bchron detrended curves (–35 ^{14}C yr precision) best fit mode results with 95% highest posterior density regions (HDR) and mean summaries, illustrate the sensitivity for incorporating calibration artefacts (linear) and allow qualitative judgement of the success with the mode has reconstructed nonlinear (sinusoidal) palaeomarsh surface accumulation and whether HDR have fully contained it (black cube - clear excursion, black line - minor excursion).

Figure A14



(a-f) OxCal detrended curves (±35 ^{14}C yr precision) 95% highest posterior density region (HDR defined between 2.5% and 97.5%) using P_Sequence K=2 auto, Ssimple, Rscaled & General outlier models (grey lines), mean summary (black) and mean summary of having run with the General outlier model only (mean 3 runs), illustrate the sensitivity for incorporating calibration artefacts (linear) and allow qualitative judgement of the success with the HDR have fully contained the nonlinear (sinusoidal) palaeomarsh surface accumulation (black cube - clear excursion, black line - minor excursion).

Wright et al. - Reconstructing the accumulation history of a saltmarsh sediment core: Which age-depth model is best?

Appendix A: Supplementary information summarising age-depth modelling packages, model scenarios and model run outputs

Summary of model operation and setup parameters

Age-depth modelling was performed using Bacon (Blaauw & Christen, 2011), Bchron (Haslett & Parnell, 2008), Bpeat (Blaauw & Christen, 2005) and Clam (Blaauw, 2010) in the free, open-source statistical environment R (R Development Core Team, 2010). OxCal (Bronk Ramsey, 1995, 2001, 2009a) was executed via the online interface.

Bpeat

Bpeat provides numerical best-fit interpolations and grey-scale summaries. The former comprises the single iteration which best fits the model (*Maximum a Posteriori* - MAP), whilst the latter illustrates the full range of iterations for any given model run, but is not amenable to detrending or further analysis. We present 'best-fit' solutions based on the mean MAP results from three runs.

The user can specify the number of rate changes and the program then identifies the depth(s) at which these rate changes occur (so called change-point linear regression). The program can also detect hiatuses by accommodating age gaps between the end of one linear segment and the beginning of another. The user can adjust how the program deals with hiatuses and the extent to which accumulation rate may change between individual segments of the core, as well as setting a prior probability threshold for the identification of outliers.

Bpeat was run using a mean accumulation rate (α value) of 1.0 mm/yr (to match our simulated sequences). The number of user-defined sections was varied between 5 and 20, with 15 proving to be optimal. Fewer sections resulted in insensitivity to non-linearities, whilst more numerous sections commonly resulting in failure to produce a coherent age-depth profile. Following preliminary analysis of a range of values (0.005 – 2.0) a 'HiatusA' parameter of 0.5 was selected on the basis of good fit with simulated curves, and reflecting the low probability and duration of hiatuses associated with the Connecticut core.

<u>Prior parameter settings – altered within the R interface</u>

name=.dat file “name” within similarly named folder

nsecs=number of sections (2) (2, 5, 10, 15)

mindepth=minimum core depth cm (0)

maxdepth=maximum core depth cm (200)

RemoveExtremes=remove 14C probabilities falling outside calibration curve (FALSE)

OUT=outlier analysis 1=yes, 0=no (1)

OUTLPPROB= outlier probability 0 to 1.0 (0.05)

<u>Prior parameter settings - altered within the “constants_template.R” file</u>

ALPHAM=*G_PDF: mean core accumulation rate yrs/cm (10) (10)

ALPHASTD=*G_PDF: standard deviation accumulation rate yrs/cm (5) (5)

EPSILON=*G_PDF: larger values = greater section dependency (5) (5)

HIATUSA=*G_PDF: ‘shape’ higher values = more ‘peaked’ PDF (0.005) (0.5)

HIATUSB=*G_PDF: ‘rate’ duration 1/2=short, 1/2000=long (1/200) (1/200)

Bacon

Bacon provides numerical best-fit and confidence interval interpolations, grey scale summaries and is superficially similar to Bpeat in terms of its tuneable parameters, with section ‘thickness’ operating in a similar manner to number of sections. As before, the mean accumulation rate is set at 1.0 mm/yr and the influence of section thickness was explored in multiple runs. Whilst the selection of small section thicknesses tended to produce smoothed reconstructions, larger thicknesses had the effect of shifting accumulation rates out of phase with known variability. The precision of the radiocarbon dates also influenced the effect of section thickness with the result that different optimal values were determined for the different precisions applied here. Bacon automatically handles outliers based on student-t distributions with wider tails than a normal distribution.

<u>Prior parameter settings – altered within the R interface</u>

core=.dat file “name” within similarly named folder

res=section thickness cm (5) [nsecs] (20 to 2.5 in steps of 2.5)

d.min=minimum core depth cm (0)

d.max=maximum core depth cm (200)

default.acc	default accumulation rate shape (2) & mean (10) [ALPHA]
acc.shape	*G_PDF: higher values result in more ‘peaked’ distributions (4)
acc.mean	*G_PDF: controls the mean rate yrs/cm (10)
default.mem	section dependency strength (4) & mean (0.7) [EPSILON]
mem.strength	*G_PDF: larger values = more ‘peaked’ distributions (4)
mem.mean	*G_PDF: controls the dependency PDF mean (0.7)
default.hiatus	default known/unknown hiatus shape (1) & mean (100) [HIATUS]
hiatus.depths	location of any known hiatus depths cm
hiatus.shape	*G_PDF: larger values = more ‘peaked’ distributions (1)
hiatus.mean	*G_PDF: controls the hiatus PDF mean (100)

Bchron

Bchron (v. 3.1.4) provides numerical best-fit and confidence interval interpolations which are performed between pairs of dated levels assuming ‘piecewise linear’ sediment accumulation in a manner referred to as ‘stochastic linear interpolation’ (Parnell et al., 2008 p. 1875). Whilst the program proved time consuming to install and run, it has the great advantage of being fully automated and therefore does not require extensive preliminary analysis to determine optimal parameters. Bchron is the only program that allows for depth ranges to be included for a given sample, thereby accounting for the palaeomarsh-surface range applied to radiocarbon-dated plant macrofossils. Inclusion of this depth uncertainty (i.e. ±3 cm) has the effect of increasing the width of confidence intervals which subsequently do a better job of constraining known accumulation variability.

Clam

Clam (v. 2.0) employs classical age-depth modelling, provides both numerical best-fit and confidence interval interpolations and was developed as a quick and transparent way to produce age-depth models. It is a useful 'first-step' tool for exploring how choices made during the modelling process (e.g. interpolation method, inferred presence of hiatuses etc.) may influence the resulting chronology. Whilst less sophisticated than its Bayesian counterparts, Clam employs Monte Carlo algorithms to sample from, and thus reflect, the multi-modal probability distributions associated with calibrated radiocarbon dates. It will endeavour to fit all dated levels (i.e. there is no automatic outlier detection) and can produce models with age reversals, although there is an option to exclude these once generated. Clam will then interpolate between dated points either by applying a (global) linear solution or some form of curve (e.g. a smoothed polynomial or locally weighted spline). We used model runs employing 100,000 iterations and excluded all iterations with age-reversals. Preliminary runs using the default span (0.75) proved unsatisfactory as substantial smoothing of oscillations occurred. Further analysis revealed that a span of 0.3 coupled with a smoothed spline produced the optimal 'best-fit' solution, capturing the amplitude of simulated change whilst generating confidence intervals that circumscribed most of the known variability.

OxCal

Oxcal (online v. 4.2) provides numerical confidence interval interpolations and includes several different types of age-depth model. We used P_Sequence which is the most appropriate for the kind of depositional context considered here (Bronk Ramsey, 2008). Similar to Bchron it employs an incremental sedimentation model but in this instance the size of the sedimentation 'event' is a tuneable parameter (k) which determines how many increments are required to complete the entire sequence. Varying k impacts rigidity of the entire age-depth model and we ran a series of model evaluations (k values ranging from 0.1 to 1000) before employing a nominal k value of 2, whilst allowing the model to adjust this within a specified range. Oxcal has additional functionality in the manner in which outliers are identified during age-depth modelling. We compared the S_simple, R_scaled and General outlier models before opting for the latter.

Table A.1 Attributes of nonlinear simulated accumulation

Parameter	SIM 2	SIM 3	SIM 4	SIM 5	SIM 6
Period (yrs) peak-to-peak	200 yrs	300 yrs	400 yrs	500 yrs	600 yrs
Resolution (no.) peak-to-peak samples	3.7	5.5	7.3	9.2	11.0
Linear GIA (cm) peak-to-peak contribution	22.0 cm	33.0 cm	44.0 cm	55.0 cm	66.0 cm
Amplitude (± cm) applied & [max. possible]	±3.2 cm [±3.5 cm]	±5.0 cm [±5.3 cm]	±6.7 cm [±7.1 cm]	±8.5 cm [±8.8 cm]	±10.3 cm [±10.6 cm]
Total acceleration (cm yrs) trough-to-peak	17.4 cm in 100 yrs	26.5 cm in 150 yrs	35.4 cm in 200 yrs	44.5 cm in 250 yrs	53.6 cm in 300 yrs
Linear GIA (cm) trough-to-peak contribution	11.0 cm	16.5 cm	22.0 cm	27.5 cm	33.0 cm
Detrended acceleration (cm yrs) trough-to-peak	6.4 cm in 100 yrs	10.0 cm in 100 yrs	13.4 cm in 200 yrs	17.0 cm in 250 yrs	20.6 cm in 300 yrs

Summary of nonlinear sinusoidal simulation (SIM) attributes tailored to the Pattagansett PXY cores. Linear glacial isostatic adjustment (GIA) applied in all instances is equivalent to 0.11 cm/yr (i.e. SIM 1).

Table A.2 Summary goodness-of-fit for each non-linear simulation and modelling approach. Figures indicate the percentage of predicted values outside the 95% confidence interval for age and depth (not available for Bpeat). Values greater than 5% indicate the extent to which confidence intervals were too narrow (over-estimate of precision). Further details of model misfits are represented graphically in Figures A2 – A14.

Age Misfit	SIM 2	SIM 3	SIM 4	SIM 5	SIM 6
Oxcal	17.7%	2.5%	0.0%	0.0%	1.5%
Bacon	17.7%	18.2%	26.8%	30.3%	18.2%
Bchron	0.0%	3.0%	8.6%	1.5%	1.5%
Clam	9.6%	12.2%	9.6%	16.8%	12.7%
Depth Misfit	**SIM 2**	**SIM 3**	**SIM 4**	**SIM 5**	**SIM 6**
Oxcal	19.1%	5.0%	0.0%	0.0%	4.4%
Bacon	17.3%	23.2%	29.8%	30.8%	30.1%
Bchron	0.0%	5.4%	9.2%	0.0%	2.5%
Clam	10.5%	19.0%	15.2%	20.7%	22.3%

Wright et al. - Reconstructing the accumulation history of a saltmarsh sediment core: Which age-depth model is best?

Appendix B: Details of age data for Pattagansett River salt-marsh core

Table B.1 Accelerator mass spectrometry ^{14}C results

Lab no. (UtC-)	Depth (cm)	PMS (cm)	$\delta^{13}C$ (p.mil)	^{14}C age ±1σ
12834	29-30	26±3	-13.4	145±29
12835	35-36	32±3	-13.0	160±28
12836	41-42	38±3	-12.9	157±29
12837	47-48	44±3	-12.9	104±29
12838	53-54	50±3	-13.0	173±28
12839	59-60	56±3	-13.0	334±30
12840	**65-66**	**62±3**	**-13.4**	**222±35**
12841	71-72	68±3	-13.9	364±37
12842	77-78	74±3	-13.5	468±34
12843	83-84	80±3	-13.4	605±35
12844	89-90	86±3	-13.4	571±36
12845	95-96	92±3	-13.5	650±35
12846	101-102	98±3	-13.6	760±35
12847	107-108	104±3	-13.8	873±39
12848	113-114	110±3	-13.8	1018±36
12849	119-120	116±3	-14.3	991±43
12850	125-126	122±3	-13.8	1043±38
12851	131-132	128±3	-13.5	1186±35
12852	137-138	134±3	-13.9	1113±37
12853	143-144	140±3	-14.3	1188±35
12854	149-150	146±3	-14.0	1169±37
12855	155-156	152±3	-13.8	1213±38
12856	161-162	158±3	-14.0	1309±38
12857	167-168	164±3	-13.9	1471±36
12858	173-174	170±3	-14.3	1544±37
12859	179-180	176±3	-14.7	1532±35

All dated material consists of *Spartina patens* rhizomes. (Depth) sample depth in core; (PMS) estimated depth of palaeo-marsh surface; ($\delta^{13}C$) abundance of ^{13}C relative to ^{12}C with respect to PDB reference; (^{14}C age ±1σ) ^{14}C age in years before present (BP) with associated 1σ error and normalised to $\delta^{13}C$ = -25‰. Possible outlier based on linear wiggle-match shown in **bold**.

Table B.2 Gamma spectrometry results

Depth (cm)	DM (g)	CDD (g/cm^3)	xs ^{210}Pb (Bq/kg)	± (%)	^{137}Cs (Bq/kg)	± (%)	^{241}Am (Bq/kg)	± (%)	pwCRS (yrs)	± (yrs)
1	12.085	0.19	321.23	6.88	5.86	10.42	-	-	2.47	0.17
2	13.243	0.40	201.54	8.88	2.34	11.31	-	-	6.04	0.54
3	10.508	0.56	119.68	10.75	3.02	13.32	-	-	9.37	1.02
4	9.997	0.72	83.86	12.86	4.32	12.21	0.07	54.27	12.69	1.65
5	9.119	0.86	70.86	10.09	7.65	8.37	0.42	29.64	16.44	1.67
6	11.639	1.04	56.50	10.86	5.43	10.56	0.09	44.42	20.54	2.25
7	12.085	1.23	55.09	10.68	4.32	10.64	-	-	26.01	2.81
8	8.697	1.37	42.58	8.88	3.42	13.42	-	-	31.59	2.84
9	12.085	1.55	31.25	12.20	*34.42	7.53	-	-	37.13	4.59
10	12.764	1.75	27.81	13.05	12.31	6.53	-	-	43.86	5.81
11	13.352	1.96	17.60	13.07	*26.52	5.78	0.66	21.31	49.65	6.59
12	11.315	2.14	2.60	9.76	11.21	9.75	-	-	50.69	5.03
13	12.085	2.33	2.38	9.52	8.65	8.49	-	-	51.76	5.01
14	35.102	2.88	3.37	8.56	7.54	10.52	-	-	53.72	4.68
15	12.085	3.07	5.77	9.35	5.43	11.15	-	-	61.64	5.49
16	10.346	3.23	6.42	11.42	4.67	12.31	-	-	64.40	7.34
17	12.259	3.42	16.03	15.76	2.65	10.53	-	-	86.68	13.62
18	12.413	3.61	5.55	10.66	2.43	12.35	-	-	101.33	10.76
19	12.085	3.80	2.14	13.33	1.31	12.61	-	-	109.93	14.59
20	21.075	4.13	1.44	10.88	1.86	13.67	-	-	118.07	12.77
21	10.56	4.30	0.14	14.42	1.62	14.57	-	-	119.01	17.06
22	10.034	4.45	0.08	13.24	1.88	14.67	-	-	118.85	15.74
23	12.273	4.64	0.08	18.34	1.25	15.15	-	-	119.45	21.91
24	9.233	4.79	0.45	17.87	1.10	13.63	-	-	123.16	22.01
25	8.601	4.92	0.13	16.21	1.07	10.68	-	-	134.32	20.15
26	9.197	5.07	0.01	15.41	0.97	11.78	-	-	134.37	19.16
27	10.017	5.22	0.01	16.28	1.44	12.47	-	-	134.52	20.27
28	13.763	5.44	0.02	15.17	1.11	10.68	-	-	144.78	18.93
29	12.352	5.63	0.22	15.06	2.17	12.31	-	-	147.24	19.16
30	11.035	5.80	0.08	15.31	-	-	-	-	148.19	19.63
31	31.165	6.29	0.05	17.00	-	-	-	-	148.81	21.90
32	31.036	6.78	0.04	18.16	-	-	-	-	149.41	23.51
33	31.165	7.26	0.19	17.85	-	-	-	-	152.67	23.68
34	30.807	7.74	0.03	15.31	-	-	-	-	163.21	20.40
35	13.724	7.96	0.00	19.05	-	-	-	-	163.30	25.40
36	20.628	8.28	0.06	17.93	-	-	-	-	174.59	24.13
37	13.492	8.49	0.06	16.94	-	-	-	-	185.90	23.02

38	20.352	8.81	0.07	15.91	-	-	-	-	187.67	21.90
39	18.845	9.10	0.00	18.03	-	-	-	-	187.68	24.82
40	14.387	9.33	0.06	22.96	-	-	-	-	189.28	31.98
41	14.498	9.55	0.27	24.24	-	-	-	-	198.14	35.91
42	8.633	9.69	0.10	22.04	-	-	-	-	202.25	33.56
43	8.369	9.82	0.13	23.79	-	-	-	-	208.54	67.73
44	7.618	9.94	0.12	21.99	-	-	-	-	215.66	76.44
45	6.156	10.04	0.02	20.10	-	-	-	-	216.85	83.54
46	8.092	10.16	0.03	19.89	-	-	-	-	219.13	93.65
47	7.945	10.29	0.02	23.43	-	-	-	-	220.65	99.98
48	7.881	10.41	0.38	21.40	-	-	-	-	-	-

Results consist of (DM) sample dry mass, (CDD) cumulative dry density, (xs ^{210}Pb) excess ^{210}Pb provided by total ^{210}Pb minus ^{226}Ra, (pwCRS) ‘piecewise’ constant rate of supply age-depth model using a core top age of AD2002 and AD1963 ^{137}Cs spike at 9 cm core depth.

Wright et al. - Reconstructing the accumulation history of a saltmarsh sediment core: Which age-depth model is best?

Highlights

- The performance of five age-depth modelling programs is evaluated using synthetic and real data
- Reconstruction accuracy and precision varies but no single model is best
- Simulation reveals the smallest resolvable accumulation change in a core
- No models produce spurious oscillations that will distort sea-level reconstructions
- Increased accumulation rate in our core since AD1800 is not an artefact of data type