Statistical approaches and tools for IntCal20

Christopher Bronk Ramsey, Tim Heaton, Maarten Blaauw, Paul Blackwell, Paula Reimer, Ron Reimer, and Marian Scott

LEVERHULME TRUST

Some key challenges

- High frequency solar Miyake-type events
- Records with uncertainty in calendar age
- Floating tree-ring sequences (eg late glacial)
- Reservoir and dead-carbon effects

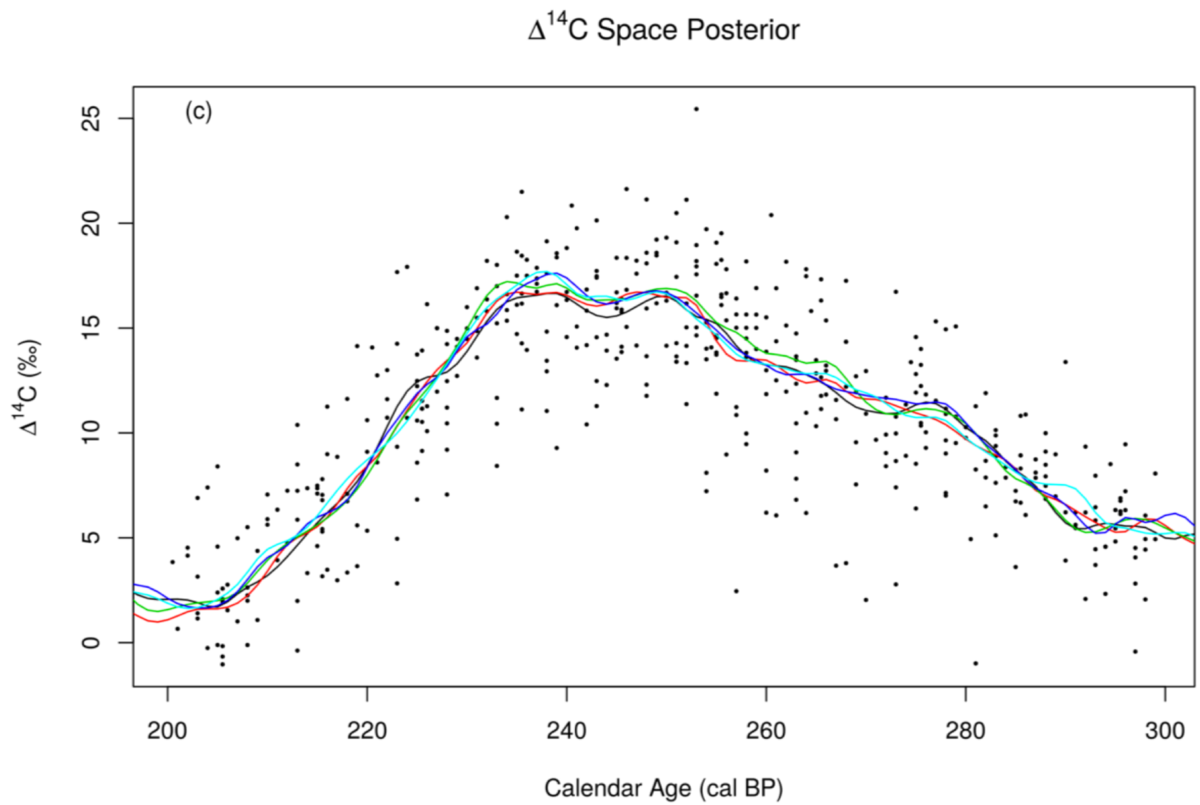
New approach

- Bayesian splines
- Ability to deal with Geophysical constraints
- Posterior information generated
- More rapid code to run

Updated tools

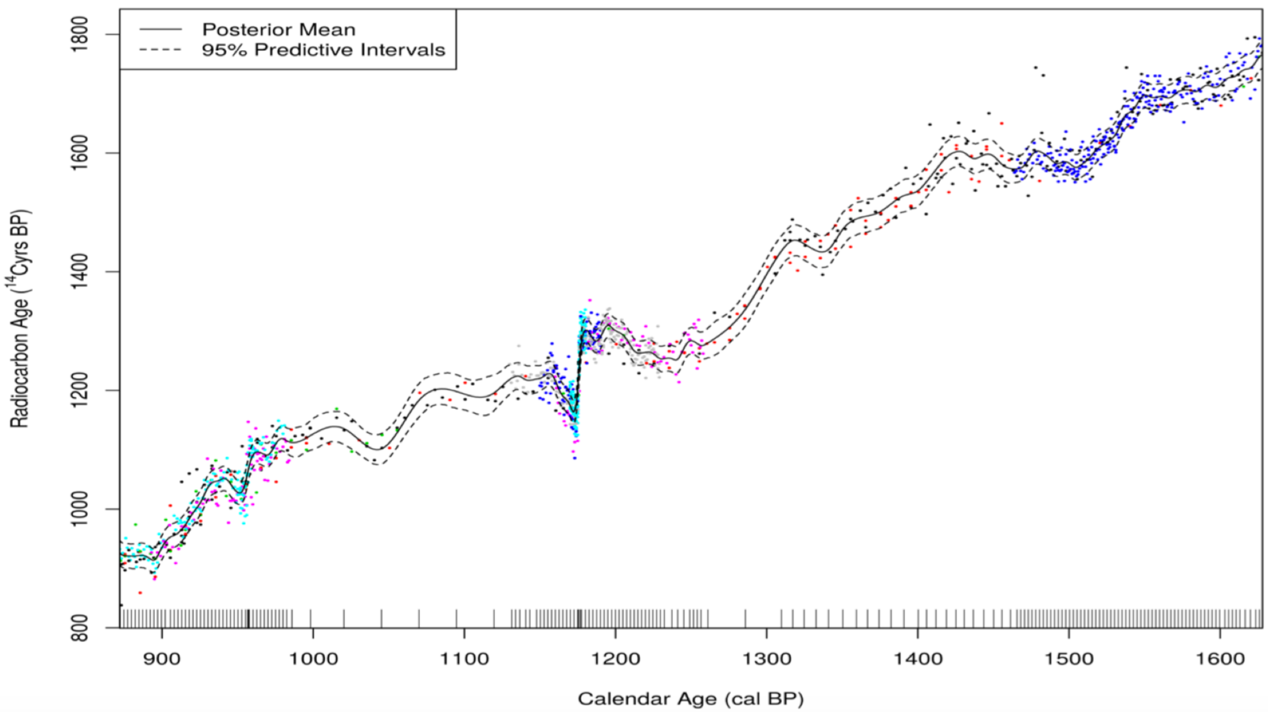
- Calib
- OxCal
- Bacon
- IntChron (INTIMATE)

Generates multiple curve realisations

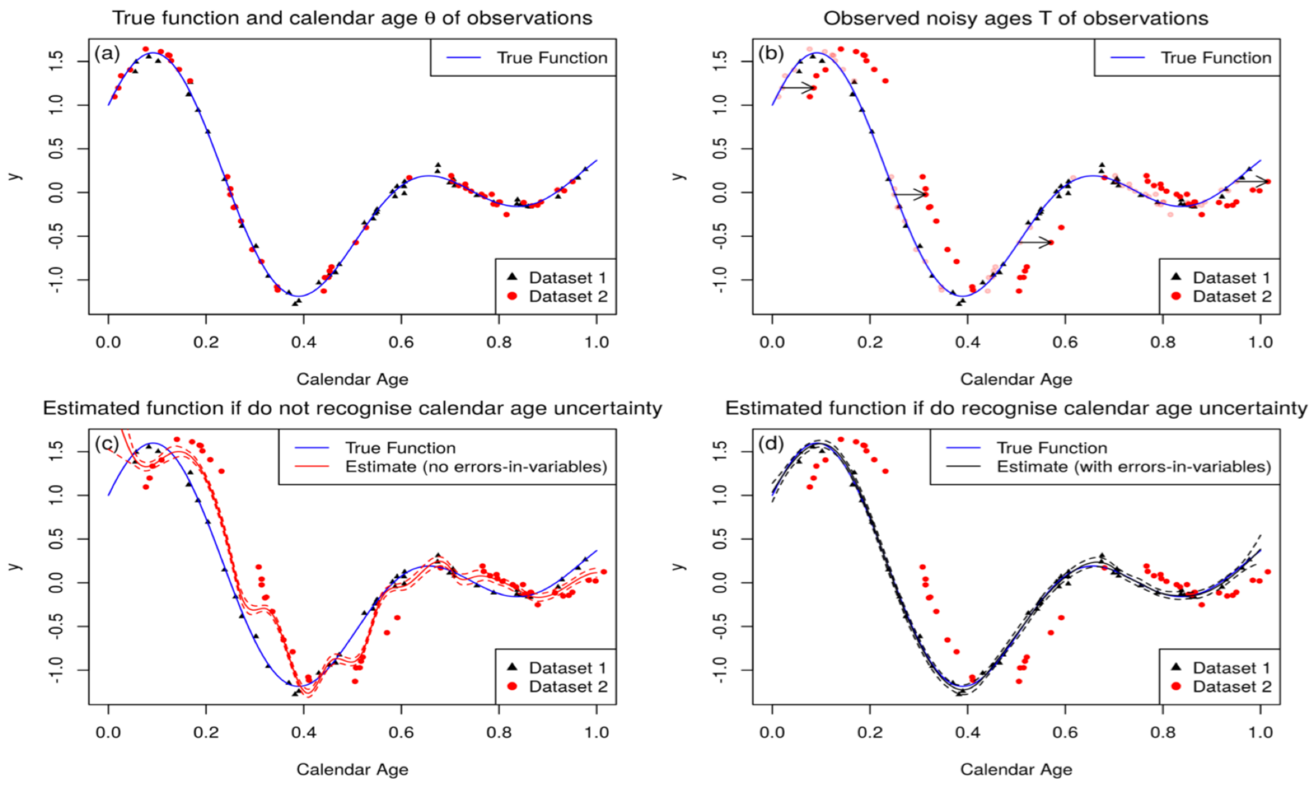


Miyake events

Can choose more knots where need more detail e.g. Miyake events



Merging of timescales



Marine and speleothem data

- Marine and speleothem determinations do not directly measure atmospheric ^{14}C
- Reservoir ages and DCF:

$$X_i = \mu(\theta_i) + R_j(\theta_i) + \epsilon_i$$

where $R(\theta)$ is term specific to set j.

- Marine Reservoir Ages — estimated via a OGCM with coastal shift
- Dead carbon fractions — varying around an unknown mean
- Incorporated similarly to errors-in-variables (but Up-Down ^{14}C shift as opposed to L-R θ shift)

Additional information

Statistical approaches and tools for IntCal20

Christopher Bronk Ramsey, Tim Heaton, Maarten Blaauw, Paul Blackwell, Paula Reimer, Ron Reimer and Marian Scott

IntCal Statistics Group

christopher.ramsey@arch.ox.ac.uk

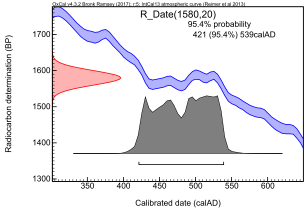
LEVERHULME TRUST

Poster Overview

- Summary of IntCal20: Aims, Data and Challenges;
- New statistical method of Bayesian spline regression;
- Incorporating unique features in the data:
  - Tree rings (ca. 0 – 14000 cal BP) — blocking, keeping detail, Miyake events;
  - Further back in time (ca. 14000 – 55000 cal BP) — uncertain calendar ages; reservoir/dcf effect; heavy tails.
- Updates to Bacon, Calib and OxCal.

Idea of radiocarbon calibration: IntCal

- Proportion of atmospheric ^{14}C fluctuated significantly over time
- Need to adjust "radiocarbon dates" via *calibration curve*
- IntCal20 curve provides historic estimate of ^{14}C from 0 – 55,000 cal BP
- Find all calendar ages θ consistent with observed radiocarbon age X



This is an inverse problem so Bayesian statistics is natural

$$\pi(\theta|X) \propto f(X|\theta)\pi(\theta)$$

where $f(X|\theta)$ is *likelihood* of observing ^{14}C determination X if it came from calendar year θ, given by the calibration curve.

IntCal20 Component Datasets

IntCal20 based on > 12,900 ^{14}C determinations where calendar age θ is known (either exactly or estimated). Data split into two categories:

Back to 14,190 cal BP:

- Dendrochronologically dated trees - many annual measurements

Further back (up to 55,000 cal BP):

- Speleothems e.g. Hulu Cave (Cheng et al., 2018)
- Corals e.g. Barbados, Tahiti (Bard et al., 1990)
- Macrofossils e.g. Lake Suigetsu (Bronk Ramsey et al., 2012)
- Forams e.g. Cariaco Basin (Hughen et al., 2004)
- Floating ^{14}C tree-ring sequences — Bølling-Allerød (Adolphi et al., 2017) and SH kauri (Turney et al., 2010)

Need to combine all these datasets together

Key data challenges

- Tree-rings a mix of annual and multi-year ^{14}C measurements
- Variable data density and sharp Miyake-type events
- Uncertainty on calendar ages of some ^{14}C determinations
- Floating tree-ring sequences (with relative but no absolute calendar ages)
- Indirect measurements of atmospheric ^{14}C — speleothems and marine samples
- Potential over-dispersion in observed ^{14}C samples (additional sources of variation)

We'll discuss those in red.

A new statistical methodology

IntCal20 uses new methodology:

- Quicker to update curves (permit more investigation);
- Still rigorous, captures uniqueness of the data;
- Ideally Bayesian (consistent with calibration).

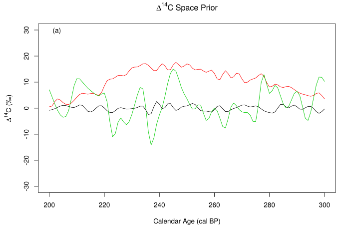
We selected:

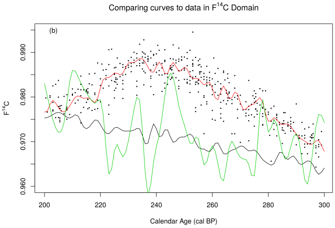
- Bayesian splines;
- Flexible to incorporate unique features of the data;
- Provides posterior information of independent interest.

Bayesian Splines

Bayesian Methods:

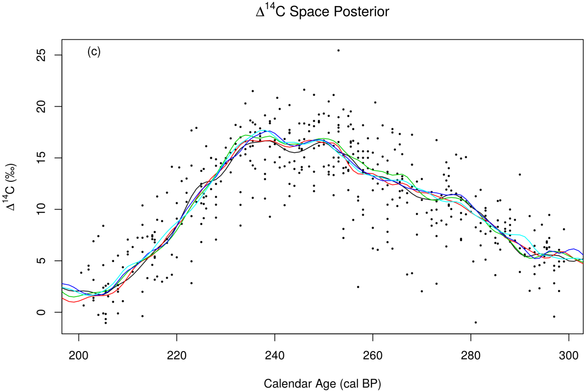
- Prior — attempts to capture beliefs, penalizes roughness in Δ^{14}C
- Observed data — provides a likelihood to combine with prior based on closeness to ^{14}C samples
- Combine into posterior — updated beliefs in light of observed data





Bayesian Splines: Posterior

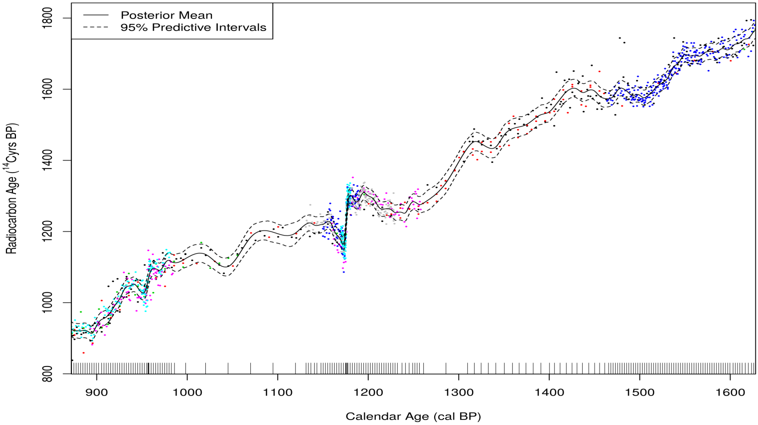
Use MCMC — outputs lots of plausible curves we summarise



Unique aspects of IntCal data

Challenges: Variable Data Density and Miyake-type events

Can choose more knots where need more detail e.g. Miyake events



Challenges: Merging timescales I

Errors-in-variables

- Not all calendar ages of older determinations are known exactly e.g. varve counting, wiggle-matching, palaeoclimate tie-pointing, floating tree-ring sequences;
- We only observe (X_i, T_i), where T_i noisy observations of true calendar age θ_i:

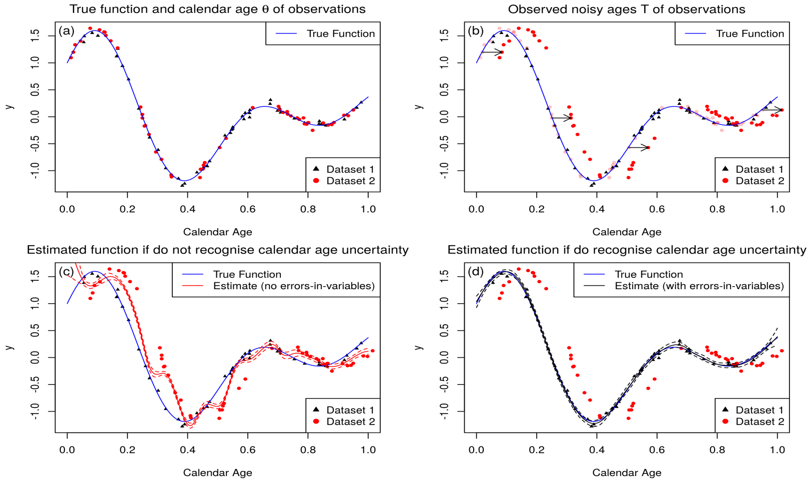
$$\begin{aligned} X_i &= \mu(\theta_i) + \epsilon_i \\ T_i &= \theta_i + \eta_i \end{aligned} \qquad \text{can't just be simplified to} \qquad \begin{aligned} X_i &= \mu(\theta_i) + \epsilon_i \\ T_i &= \theta_i \end{aligned}$$

- Some timescales need registering/merging;

Hope that:

- If multiple records show same features then keep;
- Features seen only in one record are likely noise so smoothed.

Challenges: Merging timescales II



Challenges: Indirect Measurements

- Marine and speleothem determinations do not directly measure atmospheric ^{14}C
- Reservoir ages and DCF:

$$X_i = \mu(\theta_i) + R_j(\theta_i) + \epsilon_i$$

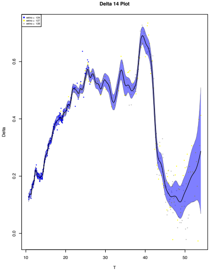
where $R(\theta)$ is term specific to set j.

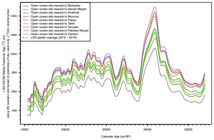
- Marine Reservoir Ages — estimated via a OGCM with coastal shift
- Dead carbon fractions — varying around an unknown mean
- Incorporated similarly to errors-in-variables (but Up-Down ^{14}C shift as opposed to L-R θ shift)

Challenges: Estimating MRAs

- Create preliminary ^{14}C curve based upon Hulu-cave only
- Use as input to 3D LSG OGCM (Butzin et al., 2020)
- Provide first-order approximation to MRAs for each dataset
- Apply constant coastal shifts and add variation to make consistent with overlap with atmospheric trees

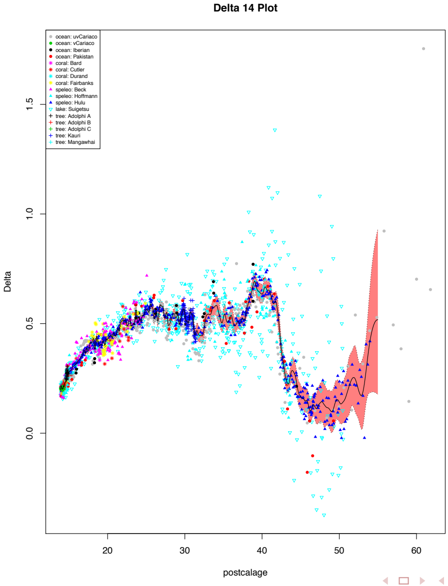
Cariaco is a unique case dealt with differently



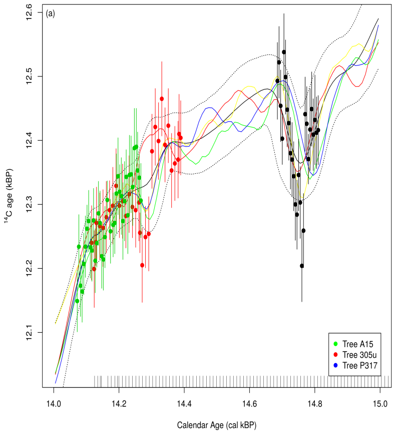


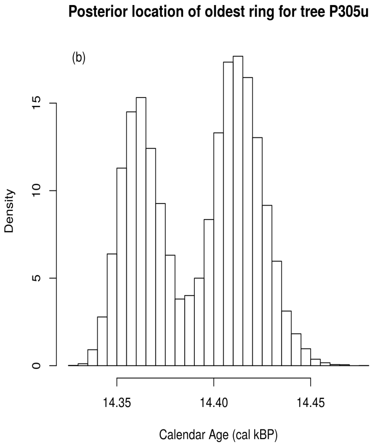
Outputs and Implications

IntCal20: Estimated $\Delta^{14}C$



IntCal20: Realisations and Internal Calibration





IntCal20: User Differences and Benefits

- Higher annual detail — will increase multimodality in calibrated age estimates
- Potential use of realisations to include more information on curve currently lost in summarisation
- See our other talk

Conclusion

- IntCal20 has new methodology based on Bayesian splines;
- Runs much more quickly;
- More flexible and can investigate modelling choices;
- Provides output of potential further interest.