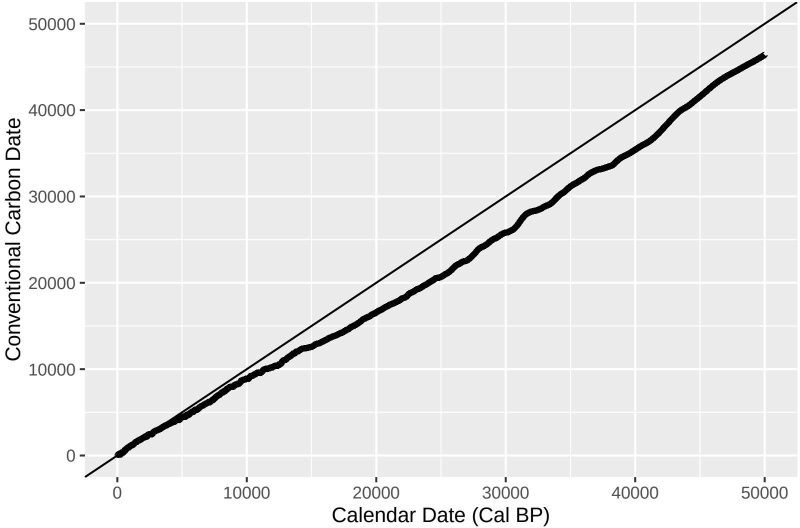
Analyzing Radiocarbon Data with Temporal Order Constraints

Reilly Villanueva

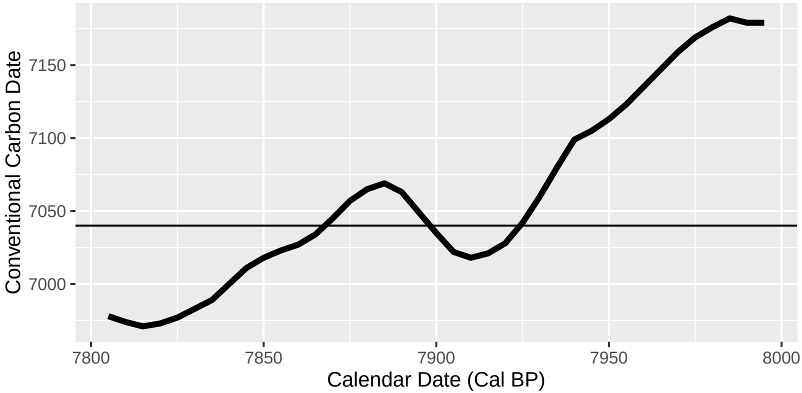
Department of Mathematics

May 6th, 2016

Calibration Curve



Calibration Curve



Bayes Theorem

B is an event.

$A_1, \ldots, A_n$ is a set of disjoint outcomes.

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

Bayes for Probability Densities

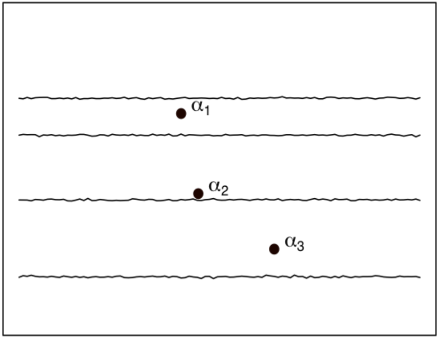
X for conventional carbon dates.

θ for ages of samples.

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta}$$

Temporal Order Constraints

For samples $\alpha_1, \alpha_2, \alpha_3$ as shown



We assume ordered ages: $0 < \theta_1 < \theta_2 < \theta_3 < 50000$

Numerical Integration

Bayes:

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta}$$

Approximating the Integral:

$$\int_a^b f(x|\theta)\pi(\theta)d\theta \approx \sum_{i=a}^{b} f(x|\theta_i)\pi(\theta_i)\Delta\theta_i$$

Markov Chains

Let Θ_i be a random variable, and let θ_i be a single draw from the random variable Θ_i.

A **Markov chain** is a sequence $\Theta_1, \Theta_2, \ldots$ of random variables where the conditional distribution of Θ_{n+1} given $\theta_1, \ldots, \theta_n$ depends only on θ_n.

A sequence of random variables $\Theta_1, \Theta_2, \ldots$ is **stationary** if for every positive integer k, the distribution of the k-tuple $(\theta_{n+1}, \ldots, \theta_{n+k})$ does not depend on n.

Gibbs Sampler

Begin with the Markov chain at state t, written as $x^t = (x_1^t, \ldots, x_p^t)$.

1. Generate $X_1^{t+1} \sim f_1(x_1 | x_2^t, \ldots, x_p^t)$.
2. Generate $X_2^{t+1} \sim f_2(x_2 | x_1^{t+1}, x_3^t, \ldots, x_p^t)$.

   $\vdots$

p. Generate $X_p^{t+1} \sim f_p(x_p | x_1^{t+1}, \ldots, x_{p-1}^{t+1})$.

Metropolis-Hastings

Let h be the (not necessarily normalized) probability density of the stationary distribution we are constructing.

1. While at position x_i, propose a move to position y having probability density conditioned on x_i, written as $q(y|x_i)$.
2. Let $h(x)$ be the density at the sample when h is parameterized by x. Compute the Hastings ratio:

$$r(x_i, y) = \frac{h(y)q(y|x_i)}{h(x_i)q(x_i|y)}.$$

3. Accept the proposed move with probability $a(x_i, y) = \min(1, r(x_i, y))$.

Highest Posterior Density Region

Proposal Density for Samples with Carbon Dates

θ_j^i is the i-th estimate for the j-th sample's age and x_j is the radiocarbon date of the j-th sample

$$f(x_j|\theta_{j-1}^{i-1}, \theta_{j+1}^{i-1}) = \begin{cases} N(\theta_j^{i-1}, \sigma^2), & \text{if } \frac{\theta_j^{i-1} - \theta_{j-1}^{i-1}}{2} < \theta_j^i < \frac{\theta_{j+1}^{i-1} - \theta_j^{i-1}}{2} \\ 0, & \text{otherwise} \end{cases}$$

Proposal Density for Samples without Carbon Dates

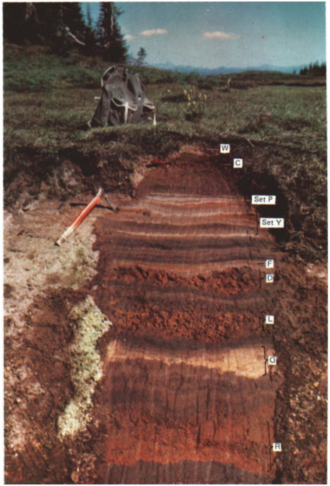
z_k is the sample without a carbon date

$$f(z_k|\theta^i_{k-1}, \theta^i_{k+1}) = \begin{cases} c, & \text{if } \theta^i_{k-1} < \theta^i_k < \theta^i_{k+1} \\ 0, & \text{otherwise} \end{cases}$$

where c is some constant

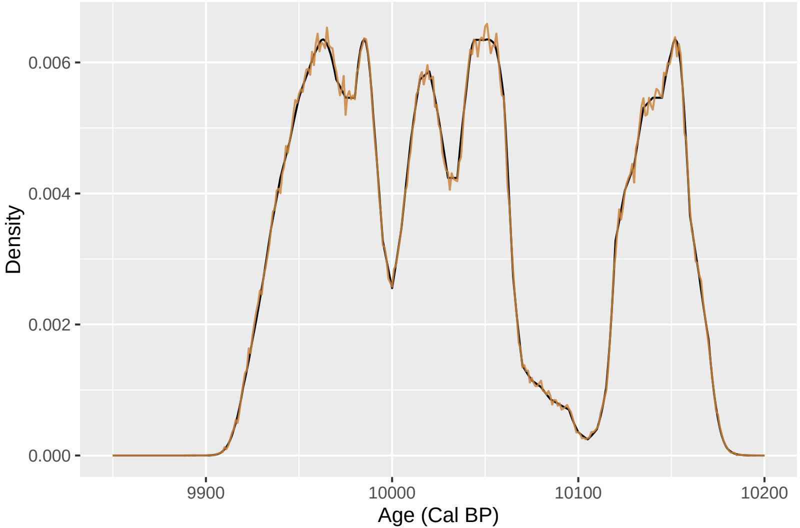
R Tephra Data

Description	Carbon-14 age (years BP)	Location relative to R tephra
Coniferous needle	8905±20	20.5cm above
Twig	8920±60	directly above
Charcoal	8890±40	directly under
Peat	8760±80	50cm below
Small twig	8990±60	100cm+ below

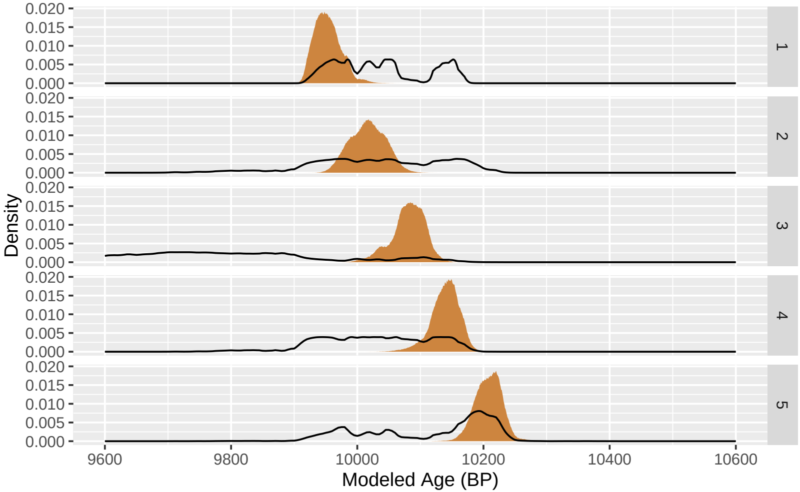


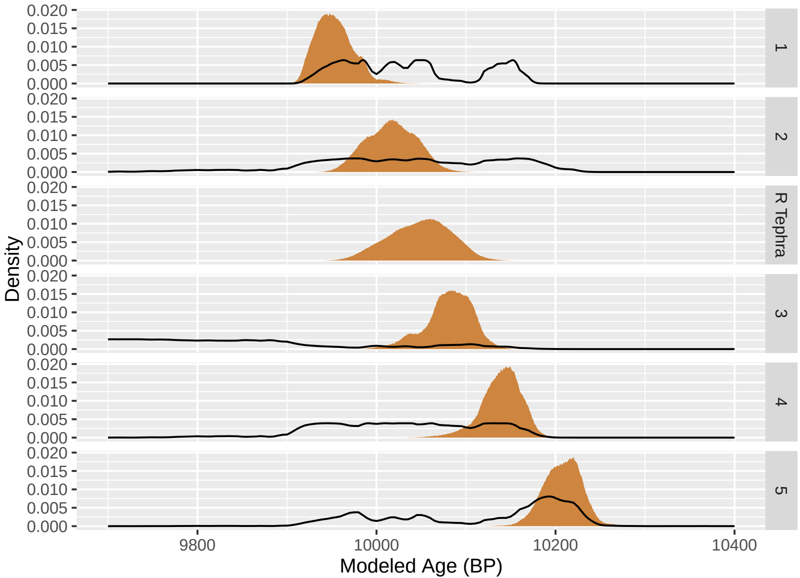
Credit: U.S. Geological Survey Department of the Interior/USGS

Comparing Numerical Integration and MCMC



Posterior Densities





Density
0.020
0.015
0.010
0.005
0.000
1
2
R Tephra
3
4
5
9800
10000
10200
10400
Modeled Age (BP)