Ice Core Dating using Probabilistic Programming

UNIVERSITY OF CAMBRIDGE

British Antarctic Survey

Aditya Ravuri [1] Tom R. Andersson [*2] Ieva Kazlauskaite [12]

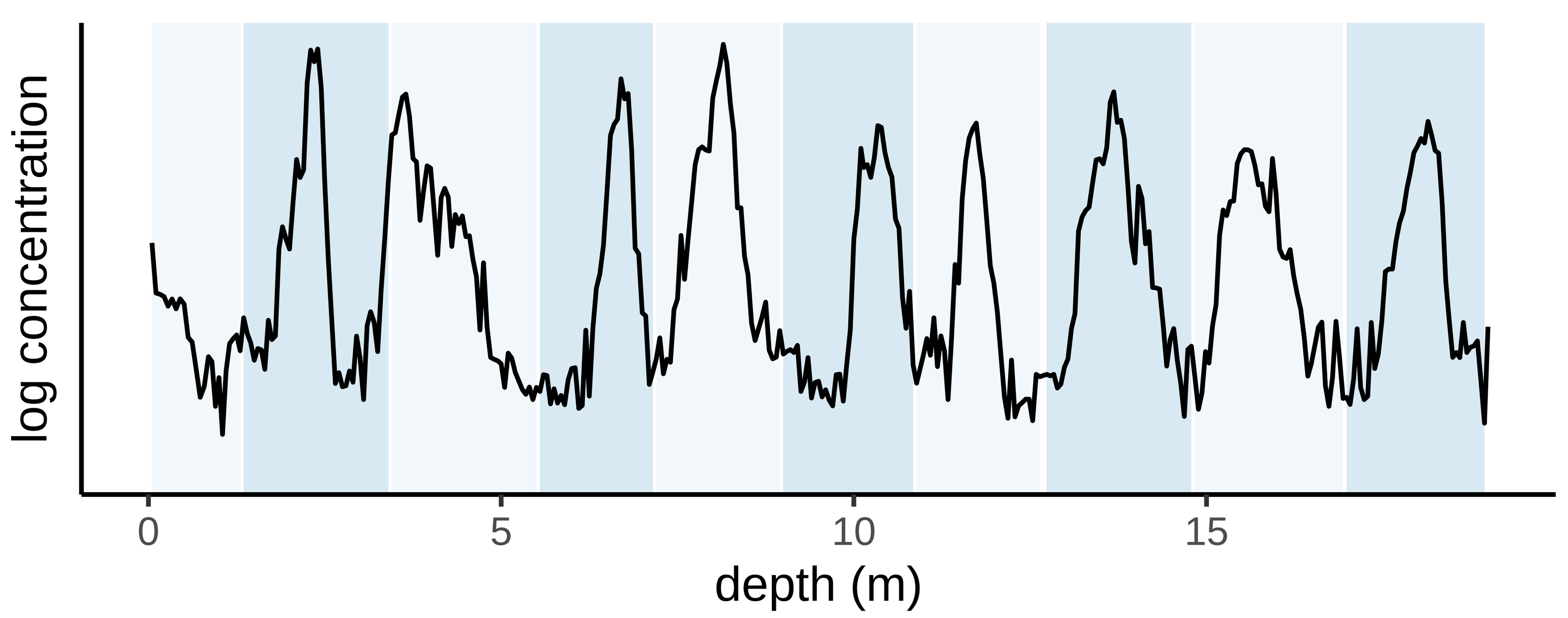
Will Tebbutt [1] Richard E. Turner [1] Mark Girolami [1] Max Fryer [1] J. Scott Hosking [23] Neil D. Lawrence [1] Markus Kaiser [12]

[1]University of Cambridge [2]British Antarctic Survey [3]The Alan Turing Institute

The Ice Core Dating Problem

1. Climate → atmospheric chemical *proxies*
2. Precipitation on ice sheet → annual layers in proxies
3. Seasonal proxy depth-series → chronology (what we infer)
4. Chronology → past climate conditions (what scientists want)

Manual layer counting is arduous [1] → automate with uncertainty!

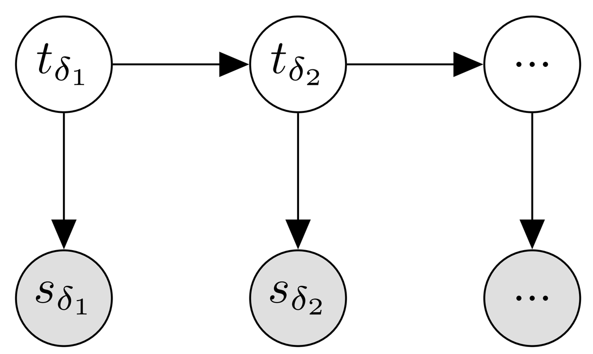


Our contribution

- A series of models for ice core dating
- Implemented using **probabilistic programming languages (PPLs)** for fully automatic inference

Graphical model

Hidden Markov model (HMMs) with latent time values, t_{δ_i}, indexed by discrete depth values δ_i



Discrete domain and index: Classical HMMs

$$\text{Transitions: } t_{\delta_i} | t_{\delta_{i-1}} \sim \mathit{Categorical}(P_{t_{\delta_{i-1}}}) \tag{1}$$

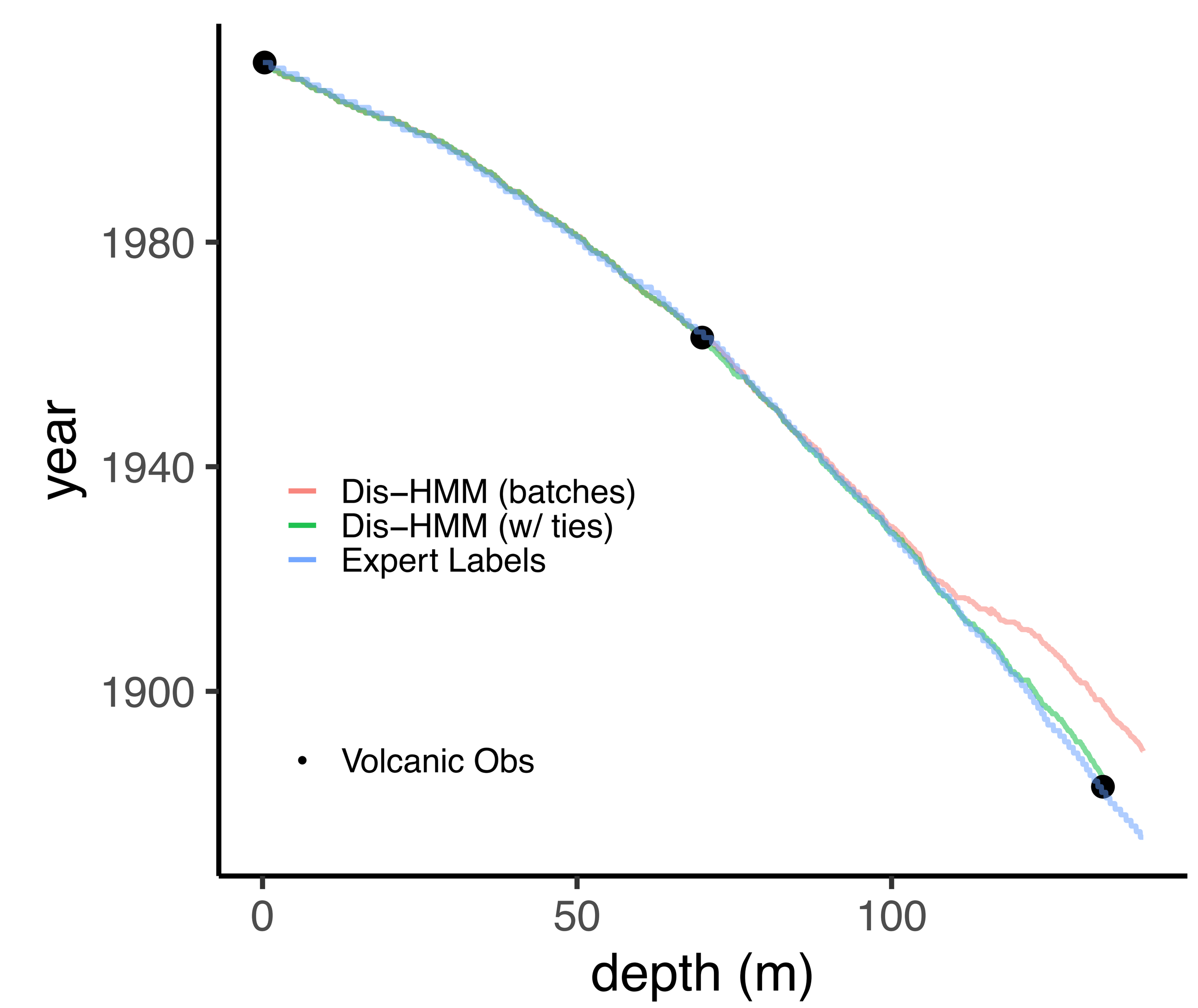
$$\text{Observations: } s_i | t_{\delta_i} \sim \mathcal{N}(a \cos(2\pi t_{\delta_i}) + b, \sigma^2) \tag{2}$$

$$\text{Transition matrix } P = \begin{bmatrix} \times & \times & 0 & \dots \\ 0 & \times & \times & \dots \\ 0 & 0 & \times & \ddots \end{bmatrix}, \tag{3}$$

with $t_{\delta_i} \in \{1/12, 2/12, \dots\}$, ensuring t_{δ_i} is monotonically increasing.

Hierarchical observation model

- Allow for parameters a, b in (2) to change with each data point with a hierarchical prior placed over a_i, b_i.
- Latent parameters (a_i, b_i) must be marginalized. MCMC is expensive, so we use variational inference (VI).
- Dated volcanic events can be incorporated, constraining the depth-time mapping.

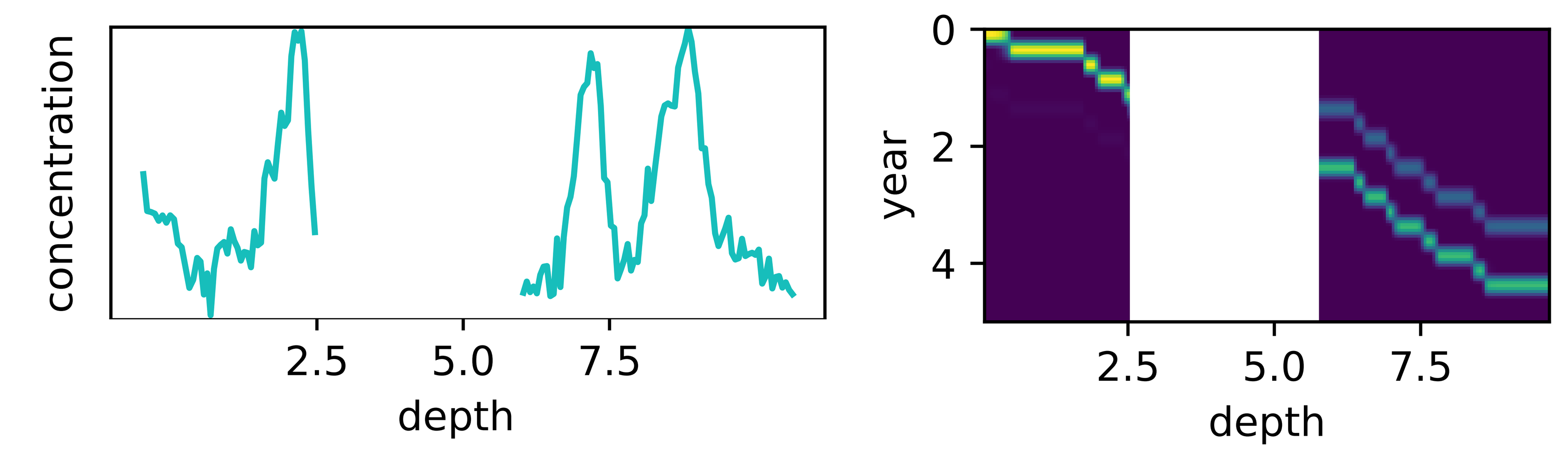


Discrete domain and continuous index: Cts-HMMs

Continuous depth index δ_i → continuous index Markov process:

$$\mathbb{P}(t_{\delta_i} | t_{\delta_{i-1}}) = \exp\left((\delta_i - \delta_{i-1})\mathbf{Q}\right).$$

With transition rate matrix $\mathbf{Q}$, can capture uncertainty over a missing data section with *bimodal* latent posterior:



Future work: An extension to SDEs

Depth & time both continuous → stochastic differential equations. For example, with monotonic sample paths:

$$\begin{bmatrix} dz_\delta \\ dt_\delta \end{bmatrix} = \begin{bmatrix} -\theta z_\delta \\ -\mu^+(z_\delta, t_\delta, \delta) \end{bmatrix} d\delta + \begin{bmatrix} \sqrt{2\theta} \\ 0 \end{bmatrix} d\mathbf{W}_\delta.$$

The Promise of Probabilistic Programming

Probabilistic programming languages promise to enable automatic inference and fast model prototyping, while ensuring maintainability.

In practice, inference is only possible in limited model classes, and some forms of inference are not scalable.

[1] Mai Winstrup.
A hidden markov model approach to infer timescales for high-resolution climate archives.
In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, AAAI'16, 2016.