ticktack: A Python package for carbon box modelling

Utkarsh Sharma 1*, Qingyuan Zhang 1*, Jordan Dennis 1, and Benjamin J. S. Pope 1,2 ¶

1 School of Mathematics and Physics, The University of Queensland, St Lucia, QLD 4072, Australia
2 Centre for Astrophysics, University of Southern Queensland, West Street, Toowoomba, QLD 4350, Australia ¶ Corresponding author * These authors contributed equally.

DOI: 10.21105/joss.05084

Summary

Radiocarbon measurements from tree rings allow us to recover measurements of cosmic radiation from the distant past, and exquisitely calibrate carbon dating of archaeological sites. But in order to infer cosmic production rates from raw ΔC¹⁴ data, we need to model the entire global carbon cycle, from the production of radiocarbon in the stratosphere and troposphere to its uptake by the oceans and biosphere. Many such competing models exist, in which the Earth system is partitioned into ‘boxes’ with reservoirs of C¹², C¹⁴, and coefficients of flow between them.

ticktack¹ is the first open-source package for carbon box modelling, allowing you to specify your own model or load a model with the same parameters as several leading closed-source models. Built in Python on Google Jax (Bradbury et al., 2018), it solves the carbon box ordinary differential equations using diffrax (Kidger, 2021) with arbitrary parametric models of production rates. This forwards model is connected via a simple API to Bayesian inference using the MCMC engine emcee (Foreman-Mackey et al., 2013).

Statement of need

Radiocarbon dating is a fundamental tool of modern archaeology, used for scientific dating of organic samples. The radioactive decay of carbon-14 (or ‘radiocarbon’) is like a clock ticking from the moment of death of an organism, so that if you know the initial radiocarbon fraction of a sample you can use this to infer its age. Radiocarbon is produced by cosmic radiation striking the upper atmosphere, and this varies slowly with time, so it is necessary to have a ‘calibration curve’ of the natural variation of the atmospheric radiocarbon fraction with time. Because tree-rings can be assigned single-year dates by the science of dendrochronology, they can be used to accurately determine this calibration curve going back thousands of years (Reimer et al., 2013, 2020; Suess, 1970).

Not only is this useful for archaeology, but also for astrophysics and geophysics: this calibration curve encodes a history of the cosmic ray flux at Earth. And it contains surprises: single-year spikes in radiocarbon production, equivalent to several years’ worth arriving at once, called ‘Miyake events’ after their discovery by Miyake et al. (2012). These occur every thousand years or so, and have been used to date to single-year precision archaeological finds as significant as the first European presence in the Americas (Kuitems et al., 2021). The most widely accepted hypothesis is that these are the result of extreme solar particle events (Usoskin et al., 2013; Usoskin & Kovaltsov, 2021), orders of magnitude bigger than the largest ever observed in the instrumental era (Cliver et al., 2022), but considerable uncertainty remains as to their origin and detailed physics.

¹https://github.com/SharmaLlama/ticktack

When radiocarbon is produced, it filters through the entire Earth system, through the atmosphere, into the oceans, and into the biosphere. To quantitatively model tree-ring radiocarbon time series, both to infer long-term trends in cosmic radiation and the parameters of these Miyake events, it is therefore necessary to model the entire global carbon cycle. This is usually done with carbon box models (CBMs) ([Dorman, 2004]), in which the global carbon distribution is partitioned into discrete reservoirs (e.g. the stratosphere, troposphere, surface and deep oceans, long and short lived biota, etc), and modelled as a system of ordinary differential equations (ODEs) with linear couplings between reservoirs - a vectorised diffusion equation with a time-varying production term. Data are usually represented in terms of $\Delta C^{14}$, or fractional difference in radiocarbon content relative to a standard, and we usually want to infer a parametric or nonparametric reconstruction of the production term with analytic or (preferably) Bayesian methods. A number of implementations of such models exist ([Brehm et al., 2021; Büntgen et al., 2018; Güttler et al., 2015; Miyake et al., 2017]), but not only are these all closed-source codes, but also make different physical and computational assumptions so that results are not straightforwardly reproducible and comparable.

We introduce a new open-source alternative, ticktack, written in Python and using Jax. We employ an object-oriented framework, with classes for

- Box and Flow, a lightweight interface for specifying reservoirs and the flows between them, which can be compiled to a
- CarbonBoxModel, which stores these metadata and whose main method run solves the CBM ODE for an arbitrary production function and time steps using diffrax;
- SingleFitter, which stores a single tree's $\Delta C^{14}$ data and a CarbonBoxModel as attributes, and can sample their log-likelihood using emcee;
- MultiFitter, which takes a list of SingleFitter objects and provides a similar interface to their joint log-likelihood.

We include preset configurations of the CarbonBoxModel that implement reservoirs and coefficients for the 11-box Güttler et al. (2015), 4-box Miyake et al. (2017), and 22-box Büntgen et al. (2018) and Brehm et al. (2021) models.

We have applied this in our team’s accompanying science paper ([Zhang et al., 2022]) to systematically analyse all extant public data on the six known Miyake events, finding no relationship to the solar cycle, and hints of a nonzero duration in several of the events. Moreover we hope that our toolkit will be adopted more widely in the radiocarbon community, and encourage other developers to contribute to the ongoing development of this open-source code base.

**Documentation & Case Studies**

In the accompanying documentation, we have several worked examples of applications of ticktack to real and simulated data:

- fitting a single dataset with emcee;
- fitting multiple datasets with a MultiFitter;
- nonparametric direct inversion of the ODE using an analytic solution;
- flat production, illustrating optional features of the ODE solver;
- nonparametric inference using a Gaussian process, using a GP to interpolate a more robust nonparametric inversion of the data.

Figures produced by the single-dataset tutorial are shown in Figure 1.
Figure 1: Left: Cornerplot of posterior samples (Hinton, 2016). Right: Predictive posterior draws for a super-Gaussian spike with sinusoidal 11-year solar cycle, overlaid on the original 774 CE discovery data from Miyake et al. (2012).

Acknowledgements

We acknowledge contributions from Mike Dee, Ulf Büntgen, Andrea Scifo, and Margot Kuitems during the genesis of this project. We are grateful for the financial support of the inaugural Big Questions Institute Fellowship at the University of Queensland.

References


