Details of the stochastic simulation

A schematic of how residuals between the change point representation and the time series are calculated is included in Figure 1.

In order to eliminate the need to specify an initial fit to estimate the autocorrelation matrix, we perform two subsequent optimizations. During the initialization, we use very rough estimates of the autocorrelation matrix of the residuals and the modeling uncertainty: identity and the standard deviation of the observations, respectively. An automated first guess is proposed by randomly generating the n x_i values and interpolating the y_i values from the observations. This quickly orients the optimization toward the correct region. The optimization is then allowed to run for an initialization period of 25,000 iterations, with 50 individual walkers.

The best fit encountered in the initialization is used to start the true optimization (250,000 iterations with 50 walkers): modeling uncertainty is re-estimated as the mean of the absolute residuals on the best fit; and the autocorrelation matrix is reestimated by taking the autocorrelation of the residuals of the best fit. A histogram of the proposals accepted by the second optimization is taken as the final result.

To conserve computational time, the individual temperature series are run with 10,000 initialization iterations and 100,000 optimization iterations, and the Savitsky-Golay filtered series with 1000 initialization iterations and 50,000 optimization iterations. Tests show that runs with 10,000 optimization iterations achieve reasonable convergence.

Parallel MCMC methodology

To propose updates to the walkers, we apply what Goodman and Weare (2010) refer to as a "stretch move". Consider the ensemble of walkers, in this case representing potential piecewise linear fits $\mathbf{X}$ and an individual walker $\boldsymbol{X}_k^j$ in the ensemble at proposal step j. We select another walker $\boldsymbol{X}_h^j$ from the complementary ensemble $\mathbf{X}_{[k]}^j$, composed of all of the other walkers. Then, a proposal is made to update $\boldsymbol{X}_k$ to W:

$$\boldsymbol{X}_k^j \to W = \boldsymbol{X}_h^j + Z\left(\boldsymbol{X}_k^j - \boldsymbol{X}_h^j\right) \tag{1}$$

GW define the following probability distribution to generate stochastic variable Z:

$$g(Z) \propto \left\{ \frac{1}{Z}\ if\ Z \in \left[\frac{1}{A_Z}, A_Z\right];\ 0\ otherwise \right. \tag{2}$$

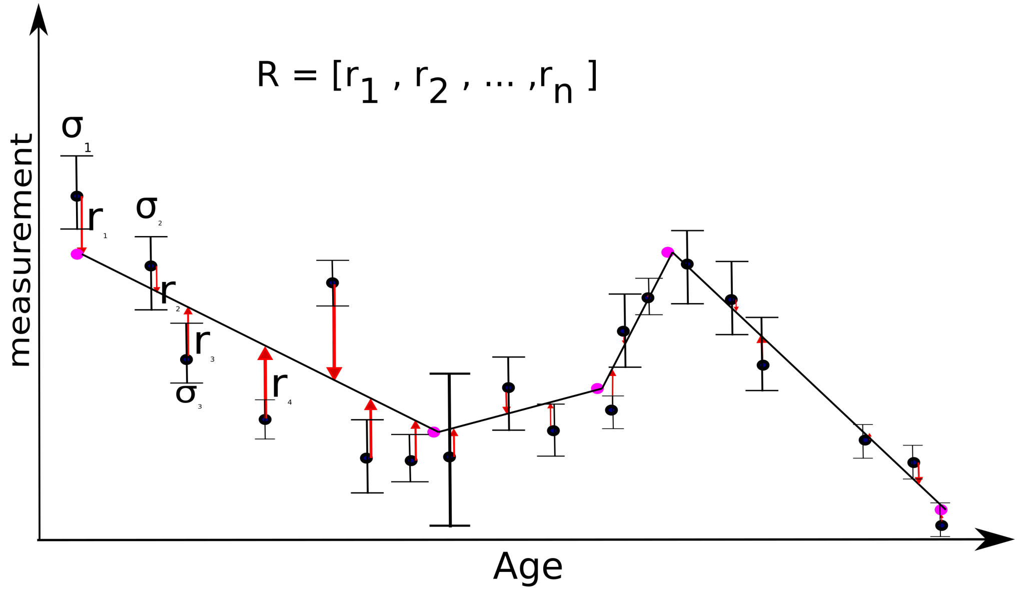

Figure 1. Schematic of the calculation of residuals between a data series and its change point representation. Here, the data series is shown as black points, with 1σ error bars. Change points are shown in pink, and the interpolations between them as black lines. The differences between the data points and the interpolations are shown as red arrows; the residual vector is composed of this distance divided by the uncertainty σ at each point.

where A_Z is a user-defined constant. In this study, the constant A_Z is set to 2, following Foreman-Mackey et al. (2013). Proposals are accepted or rejected with acceptance probability

$$P_{\boldsymbol{X}_k^j \rightarrow \boldsymbol{X}_k^{j+1} = W} = \min\left\{1,\ Z^{j-1}\frac{exp(-J(W)))}{exp(-J(X_k^j)))}\right\}. \tag{3}$$

BIC and change point sensitivity tests

We calculate the BIC for both series individually, and an additive, normalized BIC to pick the number of change points used. The additive BIC gives 7 as the best number of points to coherently fit both series. Using the individual criteria, 6 points would have been chosen for temperature, and 9 for CO_2, but comparing different numbers of points makes it less clear if the changes represented are actually coherent in timescale. These are shown in Figure 2.

We run tests to test the sensitivity of our results to the use of 5-9 change points rather than 8 (Figure 3). It is clear from these figures that 5-point fits cannot accurately represent either series; at 6 points, the temperature series begins to be well-represented, and at 9 points the CO_2 series is appropriately represented. The 7, 8 and 9 point fits do not show significant differences, except at the ACR onset.

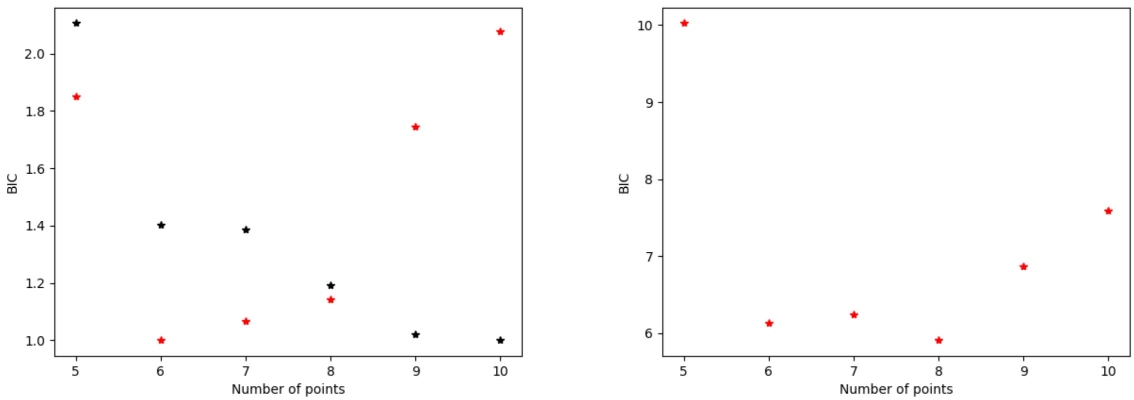

Figure 2. Individual (left, CO_2 in black and ATS3 in red) and cumulative BIC (right) for 5,6,7,8,9 and 10 point (x-axis) runs.

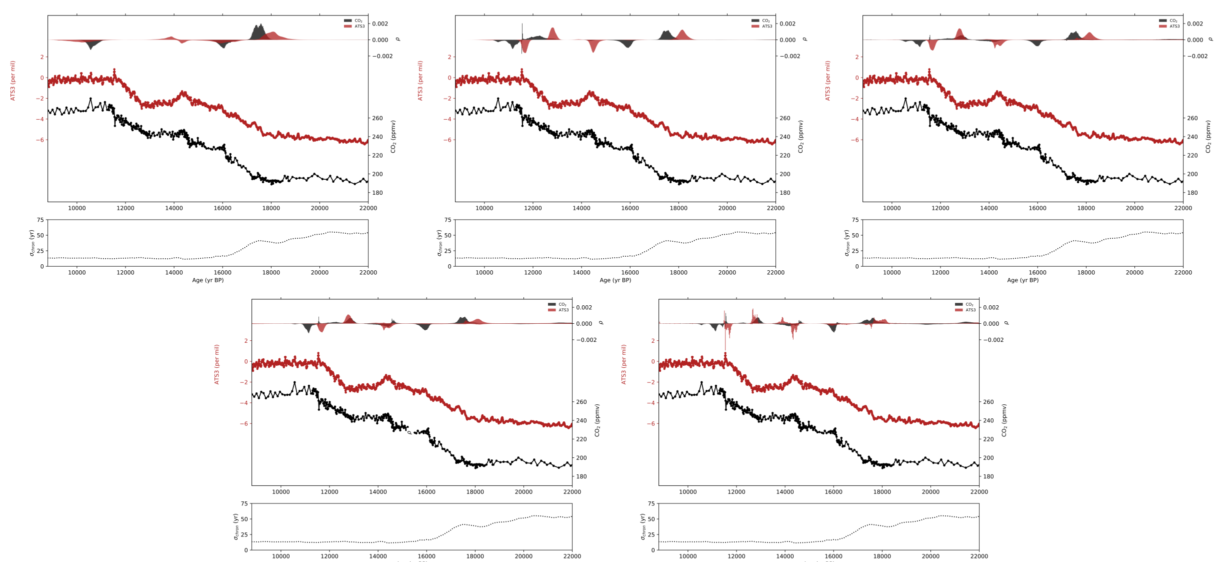

Figure 3. Atmospheric CO_2 (black) and ATS3 histograms of probable change points from the 5,6,7, 8 and 9 (from top left to bottom right) runs of LinearFit.

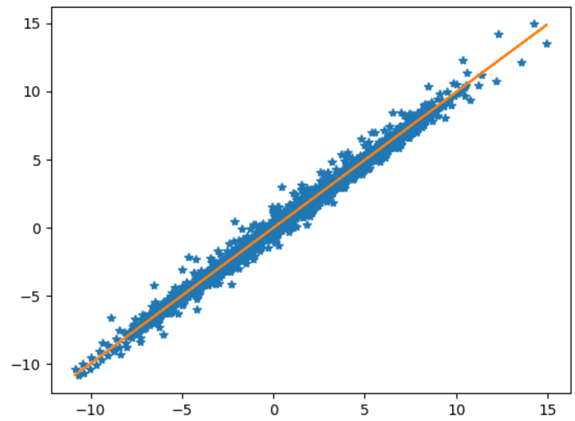

Figure 4. Plot of residuals r_i (x axis) against the predictions made using our AR(1) model $r_i = r_{i-1} \cdot a^{t_i - t_{i-1}}$ (y axis) shown as blue dots, for the CO_2 series after the initialization MCMC procedure. The orange line represents a perfect model fit.

Test of AR(1) residuals

We test our AR(1) model by plotting adjacent residuals of the CO_2 series against their predictions, using the robust AR1 model adapted to unevenly spaced time series described in the methods section. The results are shown in figure 4.

A note on Gaussian Estimates

We avoid providing estimates of change point timings in Gaussian form, i.e. as a mean ± a standard deviation. This is because the change point histograms we calculate are often skewed, and sometimes even multimodal, with multiple, separate probability peaks in a given time period where change is likely. These peaks can be due to sub-millenial scale variations at these points in the two series.

A test with known lags and covariance

.

We use artificially generated series to test the capacity of our method to fit change points in series with four known change points (plus two end points). For each series, a covariance matrix is used to generate noise with a red component, with a correlation coefficient of 0.8 at 50 years and 0.64 at 100 years away from the central point, and a uniform white component. The noise is scaled to 10% of the standard deviation of the change point values. The results of this test, along with the original change points, are shown in Figure 5.

In Figure 5, note that all of the change points are correctly identified in the histograms, with uncertainties on the order of 100 years. However, a small, but incorrect probability peak is generated around 13 ka. We conclude that gradual changes, such

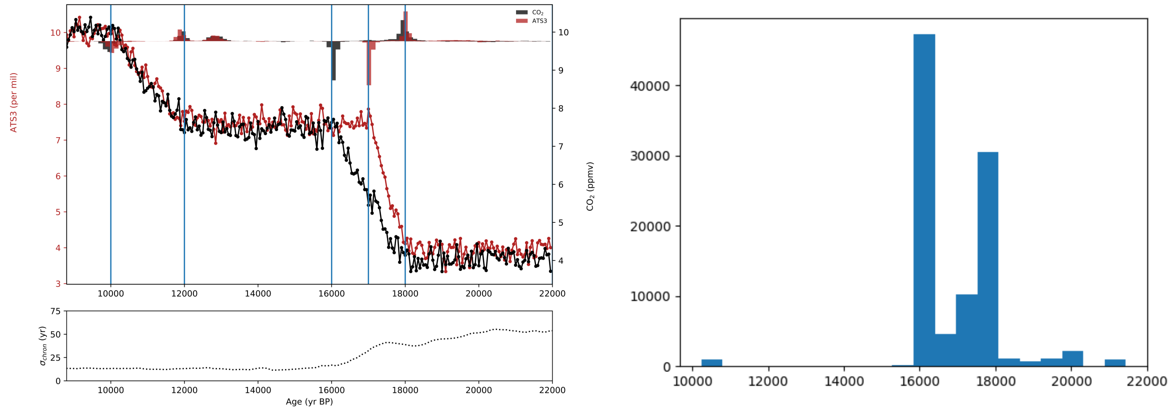

Figure 5. Artificially generated series and generated change point distributions.

as that around 12 ka, are slightly more difficult to extract from correlated noise even when this noise is modeled, adding some methodological uncertainty which is reflected in the histograms.

References

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J.: Emcee: The MCMC Hammer, Publications of the Astronomical Society of the Pacific, 125, 306, doi:10.1086/670067, 2013.

Goodman, J. and Weare, J.: Ensemble samplers with affine invariance, Comm app math comp sci, 5, 65–80, 2010.