Supplements

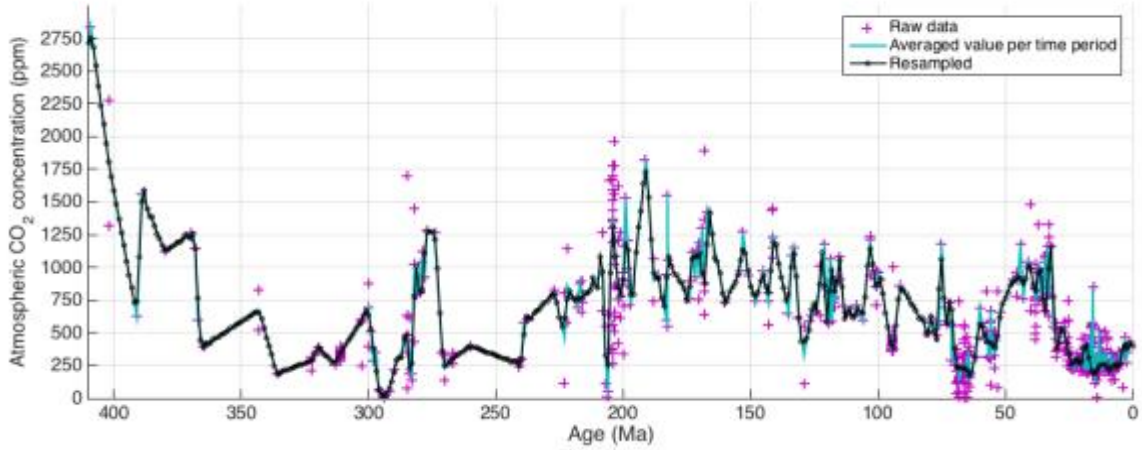

Figure S1: Filtering and resampling of raw proxy-CO_2 data (purple crosses). One median value was taken for each time step that displayed multiple observations. The result (blue) has one atmospheric CO_2 (ppm) value per time step. The proxy-CO_2 data was subsequently resampled using the *resample* function in MATLAB (black circles).

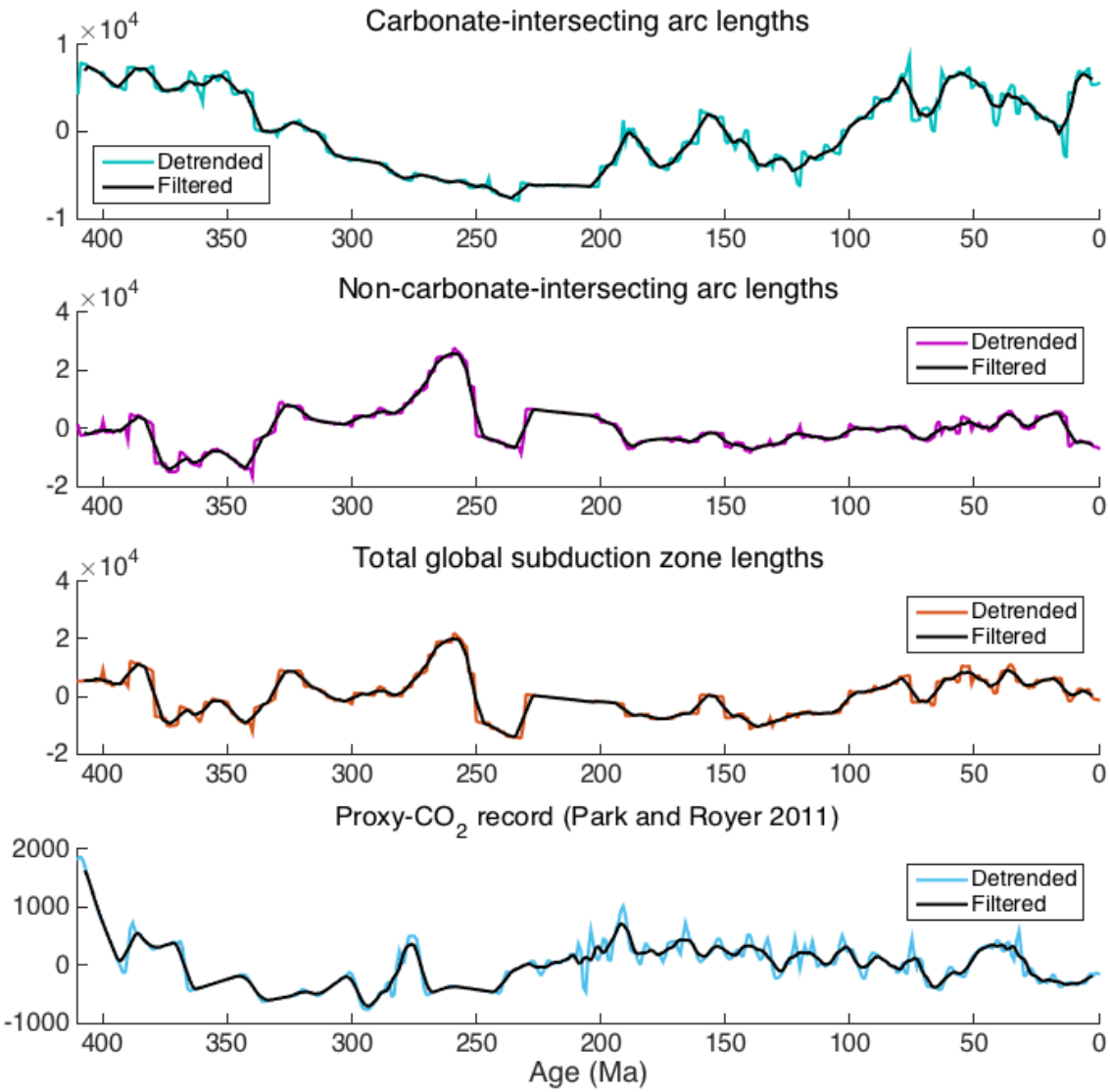

Figure S2: Results from detrending and filtering. A moving average filter was applied with a window size of 7.

S3. Continuous wavelet transform (CWT)

The application of the continuous wavelet transform has been discussed in detail by Kaiser (2010), however the method is briefly discussed below. We used a Morlet mother or analysing function because it provides a higher frequency resolution than other analysing functions and because it is complex, which is useful for discerning information about the amplitude and phase of oscillations in our time series. The shape of the Morlet function means that it

performs better than other mother wavelets in capturing both the positive and negative oscillations in each time series (Mallat, 1999). The Morlet wavelet consists of a complex exponential function modulated by a derivative of the Gaussian distribution function ($e^{-\eta^2/2}$), given by the following formula:

$$\psi_0(t) = \pi^{-1/4}\, e^{i\omega_0 t} e^{-t^2/2}, \quad (1)$$

where ψ_0 is the non-dimensional wavelet value at time t and ω_0 is the wave number set automatically in the MATLAB function, which adjusts the scale resolution.

The continuous wavelet transform (CWT) allows for the identification of periodic components in time series. It is simply the convolution of a time series x_n and a set of wavelets ψ of various widths calculated following:

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \psi^* \left[\frac{(n'-n)\delta t}{s}\right], \quad (2)$$

where (*) denotes the complex conjugate. The parameter s indicates the variation of the wavelet 'widths' or scale, which is shifted along the time series of 410 Ma (n = 410) at 1 Myr intervals (δt = 1) at the time positions n.

The set of scales for wavelet analysis are set so that all frequencies are sampled adequately from the smallest possible frequency, which is twice the time-sampling resolution ($2\delta t$) to the largest, which is half the series length. Hence, we examine scales ranging from 2 to 128 Myr periods. Edge effects in the CWT occur as the wavelet reaches the finite edges of the signal, where the transform procedure creates artefacts at the boundary. To account for edge effects, a decorrelation calculation known as the cone of influence (COI) is computed for each signal following the procedure described in Torrence and Compo (1998). Time series must be padded with zeroes before the wavelet analysis and later removed. However, this introduces discontinuities at the end points of the series which decrease the amplitude of the wavelet spectrum near the edges (Torrence and Compo, 1998). The COIs in Figure 2 and 3 are the white translucent areas where spectral information is less reliable, and only within the cone are edge effects negligible. The statistical significance of peaks in wavelet power for all wavelet analyses is determined at the 5% significance level by comparing wavelet power to the null hypothesis of a

stationary process with a given coloured-noise background power spectrum (Grinsted et al., 2004). Such a statistical test is applicable as both time series have a probability density function that is approximately normally distributed. We used a first order autoregressive process [AR(1)] based on the approach described in Allen and Smith (1995) because it is a basic model of red noise (i.e. power decays with increasing frequency) that is commonly used in geophysical data. If a peak in the wavelet power spectrum is significantly greater than the AR1 noise, then it is significant at the 5% level and 95% confidence intervals are constructed as black contours, such as in Figures 2 and 3. For more detail on statistical significance testing, refer to Torrence and Compo (1998) and Grinsted et al. (2004).

S4. Cross-wavelet transform (XWT) and wavelet coherence (WTC)

The XWT and WTC were computed to quantify interdependence between the arc time series and the CO_2 proxy record.

The cross-wavelet transform (XWT) of the two time series *x* and *y* can be defined following Eq. (3):

$$|W^{xy}| = |W^{x}W^{y*}|, \quad (3)$$

where (*) is the complex conjugation. The cross-wavelet transform was calculated using a complex Morlet wavelet to find areas of covariance between the CO_2 proxy data and the modelled results in the transform domain; that is, where frequency signal components in the series match. The local phase difference between two individual signals *x* and *y* is equivalent to the phase angle of the XWT, visualised by a map of arrows in the XWT plot (Fig 3). The phase map illustrates whether the CO_2 proxy data lags or leads the modelled results. Where the two series share common power, the phase arrows are locked or slowly varying through time and the angle of the phase arrow from vertical is used to quantify the phase relationship. West-pointing arrows suggest the two series are in phase where the peaks are correlated, whereas east-pointing arrows indicate the series are in anti-phase. Arrows pointing to the north indicate that the model is lagging the proxy-CO_2 data and arrows pointing to the south indicate that the model is leading the proxy-CO_2 data. Phase-locked behaviour is consistent with the proxy data and modelled results having combined high power, but are insufficient in significance testing interrelationships (Maraun and Kurths, 2004).

The statistical significance of cross-wavelet power against an AR1 null hypothesis was carried out following Torrence and Compo (1998) and Grinsted et al. (2004).

Wavelet coherence between two time series is a measure of the strength of the phase relationships calculated by the XWT, and can be used to estimate the causality between two signals. To investigate coherence of two series *x* and *y* in time frequency space, the following squared wavelet coherence estimator is as follows:

$$R_n^2(s) = \frac{\left|S\left(s^{-1}W_n^{xy}(s)\right)\right|^2}{S(s^{-1}|W_n^x(s)|^2)\cdot S\left(s^{-1}\left|W_n^y(s)\right|^2\right)}, \quad (4)$$

where *S* represents a localised smoothing operator. For a detailed description of the smoothing operation, see Torrence and Compo (1998) and Grinsted et al. (2004). In sum, The WTC between two time series is the squared cross-spectral power between two time series, normalised by the local power on a scale of 0 to 1 (Torrence and Compo, 1998). The statistical significance level of WTC is estimated against red noise using Monte Carlo methods as outlined in Torrence and Compo (1998). Due to smoothing, the localisation of wavelet power in time is diluted. In addition, normalisation coefficients may result in regions of coherence with a high confidence level may exist even when common power is low. Both these effects necessitate that frequency domain correlations in the WTC be paired with time-localised areas of high joint power in the XWT to find meaningful causal relationships. Some interpretations of XWT and WTC suggest that regions of shared power are only significant when periodicities persist for an interval multiple times larger than the wavelength period (Katsavrias et al., 2012). We have not applied this approach because we hypothesise that the complex feedbacks in the climate system may result in non-continuous and ephemeral causal relationships between atmospheric CO_2 and volcanigenic forcing. The significance of this relationship is not captured fully by the persistence of a particular waveband of a particular periodicity, and we therefore geologically contextualise all regions of co-varying power to reach meaningful conclusions from the data.

Supplemental References

Allen, M. R., and Smith, L. A.: Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise, Journal of Climate, 9, 3373-3404, 1996.

Grinsted, A., Moore, J. C., and Jevrejeva, S.: Application of the cross wavelet transform and wavelet coherence to geophysical time series, Nonlinear processes in geophysics, 11, 561-566, 2004.

Kaiser, G.: A friendly guide to wavelets, Springer Science & Business Media, 2010.

Katsavrias, C., Preka-Papadema, P., and Moussas, X.: Wavelet analysis on solar wind parameters and geomagnetic indices, Solar Physics, 280, 623-640, 2012.

Mallat, S.: A wavelet tour of signal processing, Academic press, 1999.

Maraun, D., and Kurths, J.: Cross wavelet analysis: significance testing and pitfalls, Nonlinear Processes in Geophysics, 11, 505-514, 2004.

Torrence, C., and Compo, G. P.: A practical guide to wavelet analysis, Bulletin of the American Meteorological society, 79, 61-78, 1998.