Supplementary Online Materials for "Dating Hiatuses: A Statistical Model of the Recent Slowdown in Global Warming – and the Next One"

J. Isaac Miller[1] and Kyungsik Nam[1]

[1]University of Missouri

Correspondence: J. Isaac Miller (millerjisaac@missouri.edu)

S.1 Updating the Forcing Series

We update the forcing series of Hansen *et al.* (2017) to 2016 as follows. We regress the first three series, CO_2, CH_4, N_2O, (in W/m^2) onto the natural log of the series given by NOAA[1] in ppm for carbon dioxide and in ppb for methane and nitrous oxide. We then predict the 2016 forcings using the natural log of the 2016 NOAA data.

In contrast, forcing from CFCs is changing very slowly, so we set 2016 to be the average of 2004-2015. We take the same moving average approach to estimate the 2016 forcings for ozone, tropospheric aerosols & surface albedo, and volcanoes as for CFCs. Volcanic activity over this period was not trivial, as noted by Vernier *et al.* (2011) and Neely *et al.* (2013), but neither are there any major eruptions on the order of Mount Pinatubo in 1991. Finally, solar data in 2016 is updated by imposing the 2015-16 percent change from NASA.[2]

S.2 Estimating the Oceanic Multidecadal Oscillation

Ocean cycles in mean temperature data – and the AMO in particular – have been estimated a number of ways in the literature. A key problem in estimating the cycle is removing the long-run trend due to global climate change. A common method for this purpose is linear detrending of GMT (Enfield *et al.*, 2001; Wyatt and Curry, 2014), although linear detrending has been criticized for this purpose by the IPCC (Bindoff *et al.*, 2013). Approaches using stochastic trends include those of Trenberth and Shea (2006), who use temperatures in other oceans to detrend the Atlantic, and Lenton *et al.* (2017), who use global mean temperature to detrend regions in the Atlantic and Pacific.

Another problem that we must avoid is over-fitting the statistical model in equation (2) in the paper. As an example of over-fitting, consider 2016, which was an unusually warm El Niño year. A variable constructed by simply detrending sea surface temperatures would have a particularly high value for 2016. Regressing GMT onto detrended GMT or sea surface temperatures

[1]Downloaded from www.columbia.edu/~mhs119/GHGs on May 15, 2017.

[2]Downloaded from solarscience.msfc.nasa.gov on May 15, 2017.

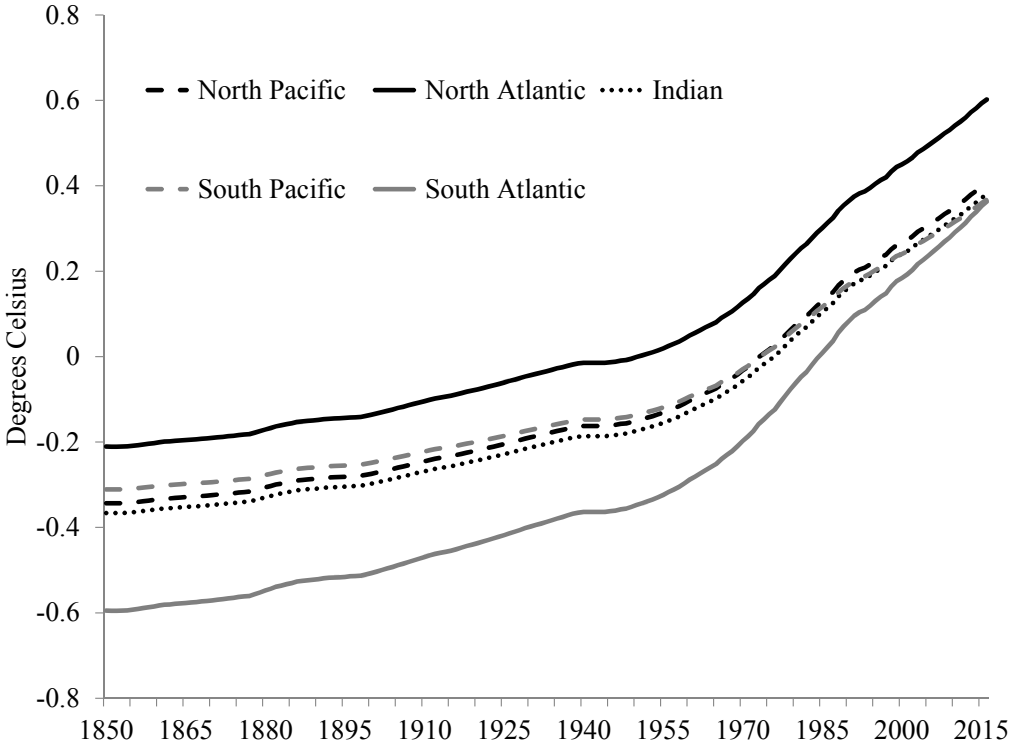

Figure S.1. Heterogeneous Oceanic Temperature Trends. Fitted anomalies from regressing average temperature anomalies of five ocean regions onto a constant and WMGHGs.

would show a superficially good fit, in the sense that the model could not distinguish between secular cyclical variability and idiosyncratic noise.

Our approach handles these two problems by filtering out both long-run and short-run information from the time series of temperature anomaly distributions. We decompose the temperature anomaly T_a (see Section C) into a long-run trend component T_t, a stationary multidecadally oscillating component T_s, and a noise component T_n, so that $T_a = T_t + T_s + T_n$. Similar to the AMO, we refer to T_s as the Oceanic Multidecadal Oscillation (OMO).

In order to estimate T_s, we first divide up the HadSST3 data into oceans: North Atlantic, South Atlantic, North Pacific, South Pacific, Indian, defined according to NOAA.[3] We then calculate the mean for each ocean at each year and then detrend those means. Rather than using a linear time trend or using the trends from other oceans, we regress these means into a constant and WMGHGs, reflecting models in the literature, like that in our equation (2), which assume a long-run relationship between temperature and WMGHGs. Doing so for each ocean separately allows for heterogeneous local climate sensitivity – i.e., for the sea surface temperatures in different oceans to be influenced by WMGHG differently over the long term.

[3] See www.nodc.noaa.gov/woce/woce_v3/wocedata_1/woce-uot/summary/bound.htm.

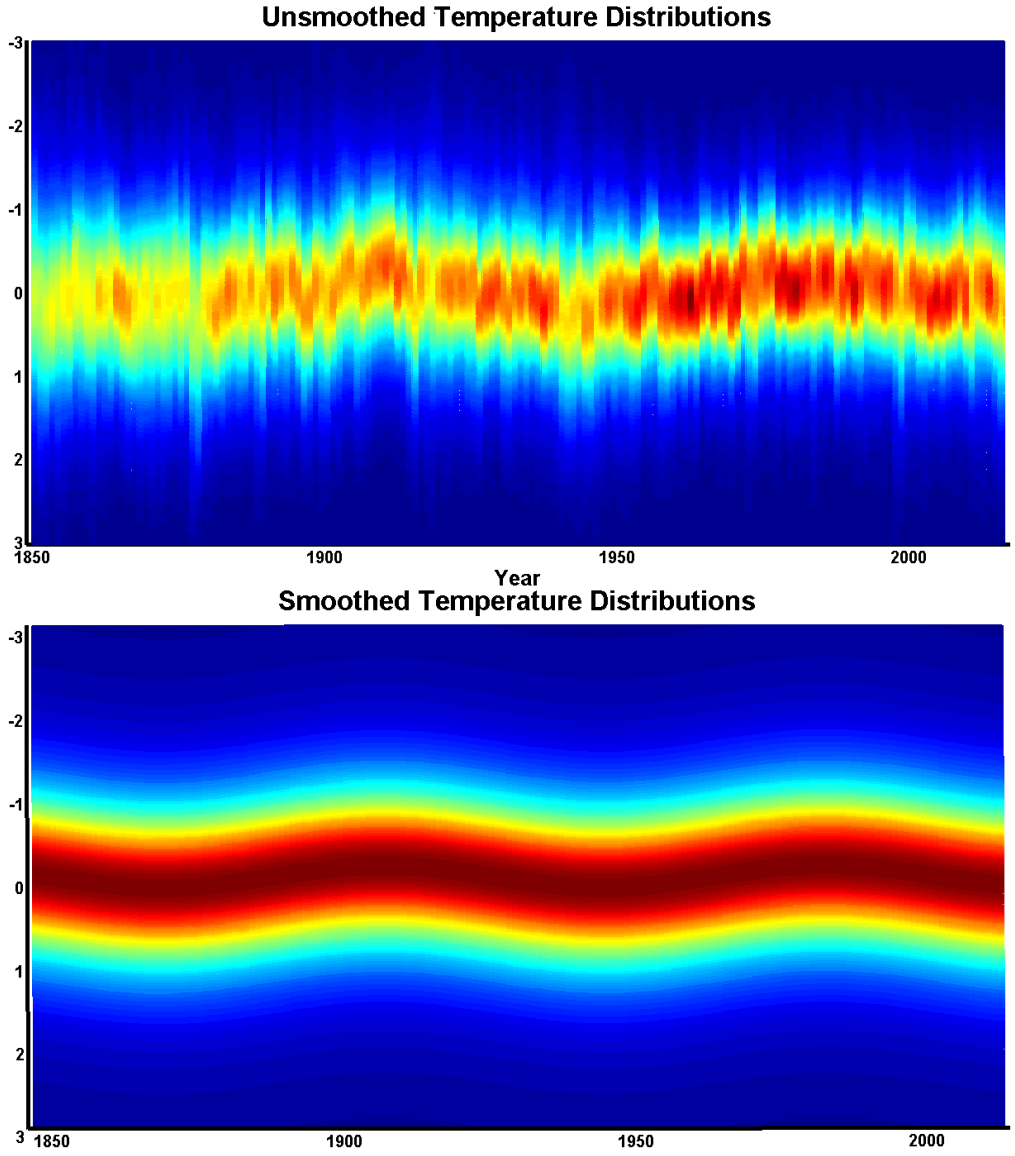

Figure S.2. Unsmoothed and Smoothed Anomaly Distributions Over Time. Top: Unsmoothed anomaly distribution from heterogeneously detrended sea surface temperature anomalies. Bottom: Smoothed anomaly distribution from fitting the unsmoothed distribution to a sine function, used to represent the OMO.

Figure S.1 shows the predicted temperature trends from these regressions. Note that the Pacific and Indian Oceans share a common warming trend, while those of the Atlantic are quite a bit different. The Atlantic is warming faster than the other oceans, with the South Atlantic catching up to the Indian and Pacific from a colder starting point and the North Atlantic becoming increasingly warmer.

	Linear/Homog		Trenberth-Shea		WMGHG/Heterog	
	est.	s.e.	est.	s.e.	est.	s.e.
θ_1	0.13	0.11	0.12	0.13	0.13	0.11
θ_2	14.66	2.71	13.53	4.99	13.75	2.95
θ_3	-0.11	1.61	0.28	3.47	-0.21	1.67
θ_4	-0.03	0.08	-0.03	0.10	-0.00	0.08

Table S.1. Periodic Function Estimation Results. Results from fitting the WMGHG/heterogeneously detrended OMO, linearly/homogeneously detrended OMO, and Trenberth-Shea (2006) AMO to the nonlinear regression in (S.1).

In order to detrend the distribution of sea surface temperatures, we simply subtract the estimated trend for each ocean from the series of temperatures in each of the 5° by 5° boxes in that ocean. A standard nonparametric density estimation technique (Gaussian kernel with Silverman bandwidth) is used to estimate the density $f_t^{s+n}(r)$ of heterogeneously detrended sea surface temperature anomalies. Detrending removes T_t, so that $f_t^{s+n}(r)$ reflects the density of $T_s + T_n$. We omit 0.5% of the outliers in each tail, which we believe is an adequate threshold to ameliorate well-known boundary problems from kernel density estimation without substantively altering the moments of the distribution. Figure S.2 (top panel) shows the density $f_t^{s+n}(r)$ of the stationary temperature distribution for each year.

Next, we smooth $f_t^{s+n}(r)$ by removing short-run noise T_n. To do so, we first calculate the spatial mean $\int_-^+ r f_t^{s+n}(r)dr$, which could be referred to as the heterogeneously detrended oceanic mean temperature. We fit the result to a single sine function, estimating

$$\int_-^+ r f_t^{s+n}(r)dr = \theta_1 \sin(\theta_2(t/T) + \theta_3) + \theta_4 + e_t, \tag{S.1}$$

using nonlinear least squares. We abbreviate the endpoints of the integral simply by $-$ for r^- and $+$ for r^+ throughout the SOM.

Nonlinear least squares estimates a periodic function with an amplitude of 0.130°C, a vertical shift of -0.000°C (nearly zero), a period given by $2\pi/13.75 \times T \simeq 76$ years, and a phase shift given by $0.21/13.75 \times T \simeq 3$ years (Table S.1). This period is roughly consistent with the Wyatt-Curry "stadium wave" with a half-period cooling regime of 31-38 years from about 1940 to about 1975. The years over the sample in which the OMO has a neutral effect – neither cooling nor warming – are (approximately) 1852, 1928, and 2004.

As a comparison, Table S.1 and Figure S.3 compare the OMO and periodic function estimated in this manner with an OMO and periodic function estimated using linear detrending and with the AMO signal of Trenberth and Shea (2006)[4] and similarly estimated periodic function. The linear detrending method estimates a shorter period of 72 years, while the Trenberth-Shea AMO signal has a longer period of 78 years. Although the AMO signal has a longer period, the phase shift is negative, so that

[4]Downloaded from www.cgd.ucar.edu/cas/catalog/climind/AMO.html on July 17, 2017.

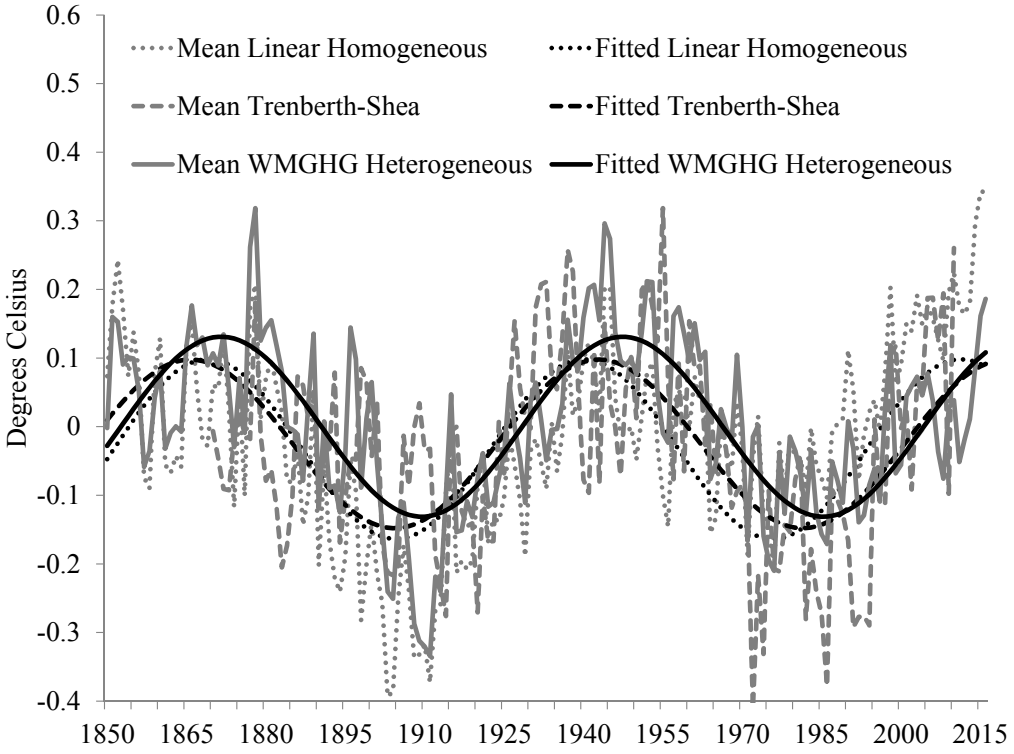

Figure S.3. Mean and Fitted Oscillations. Mean and fitted oscillations using a single linear trend, using heterogeneous WMGHG detrending, and using data from Trenberth and Shea (2006).

the next peak occurs just after the end of the sample. In contrast, the linear detrending method shows a peak in about 2011 – in stark contrast to the recent high temperatures in 2015 and 2016.

Now, in order to estimate the distribution of T_s from that of $T_s + T_n$, we construct a distribution that is changing only in mean over time. To that end, we first create a measure of the average distribution $f^n(r)$ of "de-cycled" anomalies T_n with trend and periodic function removed. The density $f_t^{s+n}(r)$ is already detrended, but in order to remove the multidecadal cycle, we change the support by subtracting the heterogeneously detrended oceanic mean temperature $T_s = \hat{\theta}_1 \sin(\hat{\theta}_2(t/T) + \hat{\theta}_3) + \hat{\theta}_4$ estimated from equation (S.1) from each temperature anomaly in each year. For example, the density function estimated for a temperature anomaly of $1°$C in year t becomes the "de-cycled" density function estimate at $(1 - T_s)°$C and, most importantly, the density at $T_s°$C becomes $0°$C. We then average the densities at each temperature anomaly r across the sample, 1850-2016, obtaining a "de-cycled" density $f^n(r)$.

Finally, we create a series of estimated densities $f_t(r)$, smoothed versions of $f_t^{s+n}(r)$. In order to do so, we reverse the procedure described above by adding the heterogeneously detrended oceanic mean temperature T_s to $f^n(r)$ in order change the support back to r. We then remove outlying anomalies outside the original support. In this way, we estimate the smoothed density $f_t(r)$ of the OMO, displayed in the bottom panel of Figure S.2. The methodology clearly extracts a density that (a)

appears to be stationary, devoid of a long-run stochastic or deterministic trend from warming, (b) appears to be smooth, devoid of idiosyncratic noise, and (c) appears to capture the multidecadal cycle.

S.3 Energy Balance Model

We define location $\ell \in G = L \cup O$ where the set G is all locations on the globe and L and O are sets of land and ocean locations. Location ℓ may be given as a latitude-longitude pair, in which case the integrals over ℓ below become double integrals over latitude and longitude. Letting the index $j = G, L, O$, $n_j = \int_j d\mu_\ell$ with counting measure μ denotes the number of locations in each set.

Adapting the energy balance model (EBM) of North (1975) and North and Cahalan (1981) to accommodate external forcing, we may write

$$C(\ell)dT_t(\ell) = QS(\ell)a(\ell) - (A + BT_t(\ell)) + D(\ell, T_t(\ell)) + h_t(\ell) + \varepsilon_t(\ell), \tag{S.2}$$

where C is heat capacity, T_t is temperature at time t, Q is the solar constant, S is solar irradiance, a is co-albedo, $A + BT_t$ is emitted energy, D is a linear-in-temperature approximation to the heat diffusion term in their model, h_t is radiative forcing, and ε_t is stochastic forcing. The stochastic forcing term is assumed by North *et al.* (1981) to be idiosyncratic, but we do not require this assumption.[5,6]

Historical temperature data sets typically express temperature in terms of anomalies from a base period in order to ameliorate well-known measurement errors. Accordingly, we decompose temperature as temperature during a base period b plus the temperature anomaly, $T_t = T_b + T_t^a$. The two forcing components may also be decomposed into base plus anomaly, expressed as $h_t = h_b + h_t^a$ and $\varepsilon_t = \varepsilon_b + \varepsilon_t^a$. Adapting the EBM in (S.2) to accommodate temperature anomalies allows

$$(C + B)T_t^a - CT_{t-1}^a = D(\ell, T_t^a) + h_t^a + \varepsilon_t^a, \tag{S.3}$$

by subtracting $C(\ell)dT_b(\ell) = 0$ from both sides, discretizing the derivative to a unit increment, and suppressing the location argument for now.

Because $C, B > 0$, $\pi = C/(C + B) < 1$, so that the autoregressive component may be inverted. Doing so yields

$$T_t^a = (C + B)^{-1} \sum_{i=0}^{\infty} \pi^i [D(\ell, T_{t-i}^a) + h_{t-i}^a + \varepsilon_{t-i}^a] \simeq \frac{1}{B}[D(\ell, T_t^a) + h_t^a + \varepsilon_t^a],$$

where the approximation results from a Beveridge-Nelson-type decomposition (see Phillips and Solo, 1992). The approximation is more valid when the data are cointegrated or cotrending, in which case the neglected terms have a lower asymptotic order.

[5] North and Cahalan (1981) assume co-albedo to be a function of temperature as well as latitude, but more recent studies show that co-albedo is effectively constant in temperature at a given latitude (Stephens *et al.*, 2015; Stevens and Schwartz, 2012).

[6] Alexeev *et al.* (2005) note the effectiveness of modeling the diffusion coefficient D as a function of temperature in order to capture polar amplification, which is not our aim.

What does the stochastic forcing term ε_t^a represent? Aside from noisy measurement of the data, it also accounts for otherwise missing components of temperature changes. Most notably missing are natural variability, such as changes in the ocean heat uptake, and other natural cycles, such as ENSO. As proxies, we employ the OMO, given by $T_t^s(\ell)$, and the SOI often used as a proxy for ENSO, given by S_t. Timmermann *et al.* (1999) note the possibility that external forcings may correlate with more frequent and/or severe ENSO cycles. By including the SOI in the model, we are implicitly assuming that the correlation is reflected in the SOI.

An alternative way to capture natural variability might be to allow for separate meridional ocean transport, along the lines of Rose and Marshall (2009), which could likely be accomplished along the lines of Pretis (2015) using a model that cointegrates surface temperature with deep ocean heat content. However, keeping in mind that our aim is to model hiatus periods that may be sparsely distributed over the historical record, the short time span over which ocean heat content is measured precludes this approach.

In such a model, surface temperatures, deep ocean heat content, and forcings share a single stochastic trend, so that the marginal value of the deep ocean heat content relative to forcings is natural variability. Hence, omitting deep ocean heat content does not cause a spurious regression, but rather relegates this stationary variability to the error term. Our proxies, the OMO and SOI, allow us to explicitly model the primary multidecadal and interannual sources of this variability.

Because the EBM does not explicitly include these indicators, it is natural to model them nonparametrically. To this end, we specify the model as

$$T_a = \alpha_0 + \alpha_1 h_t(\ell) + \alpha_1 D(\ell, T_a(\ell)) + b(T_t^s(\ell)) + c(S_t) + \eta_t(\ell), \tag{S.4}$$

where

$$b(T_t^s(\ell)) = \begin{cases} \sum_{i=1}^{m_T} \gamma_i^T b_i(T_t^s(\ell)) \text{ for } \ell \in O \\ 0 \text{ for } \ell \in L \end{cases}$$

$$c(S_t) = \sum_{i=1}^{m_S} \gamma_i^S c_i(S_t)$$

are two generic series expansions intended to capture possibly nonlinear effects of these indicators. The OMO captures variability over the ocean, so this component is set to zero over land. The SOI is a single indicator. The last term $\eta_t(\ell)$ contains the original stochastic forcings and their lags and an allowance for finite-order approximation error of the two expansions. We may think of this term representing residual forcings, and it almost certainly exhibits temporal correlation.

Now we aggregate across locations to obtain a global model. To this end, let $O(T_s) = \{\ell \in O : T_s(\ell) = T_s\}$ be a subset of O over which the $T_s(\ell)$ has the same numerical value, let $n_{O(T_s)} = \int_{O(T_s)} d\mu_\ell$ denote the number of locations in $O(T_s)$ that have the value T_s, and let f be the probability density function of $T_s(\ell)$ in O with support $[r^-, r^+]$. Note that

$$\int_{-}^{+} b(r) f(r) ds = \int_{-}^{+} \left[n_{O(T_s)}^{-1} \int_{O(T_s)} b(T_s(\ell)) d\mu_\ell \right] f(T_s) dT_s = n_O^{-1} \int_{O} b(T_s(\ell)) d\mu_\ell,$$

which means we can aggregate all of the functions b across ocean locations or we can aggregate all of the functions b with the same observed argument and then aggregate them again with weights given by the frequency of each argument. Defining

$T_a = n_G^{-1} \int_G T_a(\ell) d\mu_\ell$, $h = n_G^{-1} \int_G h(\ell) d\mu_\ell$, $\eta = n_G^{-1} \int_G \eta(\ell) d\mu_\ell$, integrating across locations, and noting that the diffusion term $D(\ell, T_a(\ell))$ is constrained to integrate to zero by the first law yields the EBM in equation (1) in the paper.

S.4 Estimating the Energy Balance Model

We approximate the functions b and c nonparametrically using a series of polynomial and trigonometric functions known as the flexible Fourier functional form, which Park *et al.* (2010) analyze using a semiparametric cointegrating regression much like ours. This form may be written as

$$\begin{aligned} b_j^v(v) &= v^j \text{ for } j = 1, ..., p_1 \\ &= \cos 2\pi k v \text{ for } j = p_1 + 2k - 1 \text{ and } k = 1, ..., q_1 \\ &= \sin 2\pi k v \text{ for } j = p_1 + 2k \text{ and } k = 1, ..., q_1 \end{aligned}$$

and analogously for $c_j^v(v)$, for $v \in [0,1]$. Using this notation, $m_T = p_1 + 2q_1$ and $m_S = p_2 + 2q_2$. It is important that these functions are defined over the unit interval, so let $b_j(r) = (r^+ - r^-) b_j^v((r - r^-)/(r^+ - r^-))$ and $c_j(S) = (S^+ - S^-) c_j^v((S - S^-)/(S^+ - S^-))$, where S^+ and S^- are the maximum and minimum observed SOI. Thus,

$$\int b_j(r) f_t(r) dr = \int (r^+ - r^-) b_j^v((r - r^-)/(r^+ - r^-)) f_t(r) dr$$

holds, making estimation convenient by simply multiplying b_j^v by the range of values in the domain of the OMO.

The optimal orders $(p_1, q_1, p_2, q_2) = (2, 0, 1, 0)$ – i.e., $m_T = 2$ and $m_S = 1$ – are jointly determined by Schwarz-Bayesian and Hannan-Quinn information criteria evaluated using least squares with p_1, p_2 up to 3 and q_1, q_2 up to 2. With $m_S = 1$, SOI enters linearly and the regressor is thus simply $(S_t - S^-)$.[7] The results are given in Table S.2.

S.5 Estimation of Uncertainties

S.5.1 Uncertainties from Estimating the OMO

Estimating the OMO relies on a statistical approximation, and we employ a parametric bootstrap strategy similar to that of Poppick *et al.* (2017) to account for uncertainty in estimation. Specifically, after fitting the periodic function in (S.1), we fit the residuals to an AR(1) and redraw from the residuals of the fitted AR(1). We re-create the regressand using a re-created AR(1) error with fitted autoregressive parameter (0.49). We then re-estimate all the parameters of the nonlinear regression. We conduct 999 bootstrap replications in this manner, and the sample paths plotted in Figure 2 in the paper reflect the periodic functions with 0.05 and 0.95 quantiles of θ_2 in (S.1), which determines the period.

[7] All models are expected to be cointegrated, because we simply add stationary series to the model with $(p_1, q_1, p_2, q_2) = 0$, which we found to be stationary (see footnote 1 in the paper). More formally, we run residual-based augmented Dickey-Fuller tests which strongly rejected no cointegration up to four lags. In addition, variable addition tests proposed by Park *et al.* (2010) with added variables t^2 and t^3 fail to reject the null of cointegration.

	OLS	CCR		OLS	CCR	
	est.	est.	s.e.	est.	est.	s.e.
α_0	-0.315	-0.313	0.026	-14.505	-13.968	13.848
α_1	0.412	0.430	0.029	0.431	0.439	0.016
α_2	0.079	0.081	0.038	0.054	0.055	0.022
γ_1				9.463	9.084	10.079
γ_2				-8.878	-8.502	10.236
δ_1						
α_0	-0.212	-0.214	0.055	-19.897	-19.446	12.474
α_1	0.407	0.423	0.029	0.425	0.431	0.015
α_2	0.092	0.094	0.038	0.069	0.069	0.020
γ_1				13.440	13.119	9.080
γ_2				-12.899	-12.569	9.221
δ_1	-0.055	-0.053	0.026	-0.062	-0.061	0.013

Table S.2. EBM Estimation Results. Results from estimating the model in (2) in the paper using least squares (OLS) and asymptotically normal canonical cointegrating regression estimates (CCR) (Park *et al.*, 2010) with $p_1 = 0, 2$, $p_2 = 0, 1$, and $q_1, q_2 = 0$ in (S.4).

Our bootstrap differs from a typical bootstrap in that we do not re-estimate only the period and restrict the other parameters to their fitted values. Re-estimating all of the parameters allows uncertainty about the period to be correlated with uncertainty about the phase shift in particular, because uncertainty in both parameters affects our dating of the function's optima.

S.5.2 Approximations to Uncertainties in Forcings

Myhre *et al.* (2013) estimate forcings (in W/m^2) with 90% confidence for 2010 as follows: from WMGHGs: 2.83 (2.54 to 3.12); from ozone: 0.07 (0.02 to 0.12); from tropospheric aerosols & surface albedo: -0.35 (-0.85 to 0.15); and from solar irradiance: 0.05 (0.00 to 0.10). Volcanic forcings are estimated to be -0.06 (-0.08 to -0.04) over 1999-2002 and -0.11 (-0.15 to -0.08) over 2008-2011. Myhre *et al.* (2013) state that uncertainty about forcing from WMGHGs is about $\pm 10\%$, and indeed this roughly corresponds to the numerical values given for 2010. The numerical values given for volcanic forcings over the two periods correspond to $\pm 33\%$ within rounding error. If we make a similar assumption regarding the other forcings, then uncertainties about forcings from ozone, tropospheric aerosols & surface albedo, and solar are $\pm 70\%$, $\pm 143\%$, and $\pm 100\%$ respectively in 2010.

We assume that these percentages are roughly the same each year, as they are for WMGHGs and appear to be for volcanoes. In this way, we generate intervals for each forcing in each year to be consistent with the given data, in the spirit of Poppick *et al.* (2017) but without a bootstrap. This procedure inherently but realistically allows heteroskedasticity, because the uncertainty grows as the forcing's value grows. However, it does not allow for heteroskedasticity due to more precise measurements

over time, e.g., as noted by Myhre *et al.* (2013) for solar. We further assume that the uncertainties in the forcings are purely idiosyncratic in the sense that they are neither time-dependent nor mutually dependent.

Approximating the distributions of these uncertainties by a Gaussian distribution centered at the reported value, we can estimate the variance by dividing the difference in the quantiles by 2×1.645 and squaring the result, which reverses the formula $\pm 1.645\sqrt{\mathbf{var}(v_t)}$ to calculate 90% intervals from the variance of a mean-zero Gaussian random variable v_t. Doing so generates a 5×5 variance/covariance matrix for each time period, with the estimated variance of each of the five forcings along the main diagonal and zeros elsewhere.

The average of the diagonals, reflecting the variances of the uncertainty for each forcing (WMGHGs, ozone, aerosols, solar, and volcanic respectively), is estimated to be $(0.007, 0.001, 0.284, 0.002, 0.012)'$ over 1850-2016 and $(0.029, 0.003, 0.969, 0.004, 0.000)'$ over 1999-2013, the recent fifteen-year hiatus period. As expected, forcings from aerosols are estimated to be the most uncertain. The uncertainties over the hiatus period are generally larger than those over the whole sample, reflecting the larger magnitudes of the forcings near the end of the sample. In contrast, volcanic forcings are nearly zero during the hiatus period, reflecting the absence of a volcanic eruption with a major impact on global climate.

S.5.3 Contribution of Regressor Uncertainty to Coefficient Estimators

Uncertainty in the regressors may be treated as "classical measurement error" in the parlance of the econometrics literature, which is known to cause bias in the coefficient estimates. Because we observe forcings with error, we may denote our observation of forcings by $h_t = h_t^0 + v_t$, where $h_t^0 = (h_{1t}^0, h_{2t}^0)'$, with h_{1t}^0 and h_{2t}^0 denoting respectively the sum of non-volcanic forcings and volcanic forcing if the forcings could be observed without uncertainty. Similarly, $v_t = (v_{1t}, v_{2t})'$ such that v_{1t} and v_{2t} are respectively the sum of uncertainties about non-volcanic forcings, estimated as described above, and uncertainty about volcanic forcing. v_t has a mean of zero and its components have variances given by $\sigma_{v1,t}^2$ and $\sigma_{v2,t}^2$. The former is the sum of the variances of the non-volcanic forcings, as the covariances are assumed to be zero.

In a cointegrating model like ours, bias in the long-run relationship is not hard to fix. In fact, although it was not designed to do so, the feasible CCR methodology of Park *et al.* (2010) already takes into account this bias. Using a closely related model, Miller (2010, Theorem 2) shows the CCR estimator to be consistent, asymptotically normal, and asymptotically unbiased, with a variance that takes into account the measurement uncertainty.[8]

S.5.4 Contribution of Uncertainty to Forecasts

Explaining the contributions of the uncertainty to the missing heat of the 1998-2013 episode requires a measure of in-sample fit of T_t^a for some arbitrary time period $t = 0$, given by $\hat{T}_0^a = h_0^{*\prime}\hat{\alpha} + x_0'\hat{\gamma} + w_0\delta$. For simplicity, denote the right-hand side by $z_0'\hat{\pi}$ with $z_t = (h_t^{*\prime}, x_t', w_t)' = (1, h_{1t}, h_{2t}, x_t', w_t)'$ and $\pi = (\alpha', \gamma', \delta)' = (\alpha_0, \alpha_1, \alpha_2, \gamma', \delta)'$ and let $z_t^0 = (1, h_{1t}^0, h_{2t}^0, x_t', w_t)'$

[8] We do not model uncertainty in measuring temperatures, which is expected to be smaller than uncertainties in the forcings. Also, we do not model the effect of uncertainty in the volcanic forcings on the coefficient estimate, because the uncertainty is much smaller than the uncertainty for the other forcings – nearly zero – over the 1998-2013 period.

and $\varpi_t = (0, v_{1t}, v_{2t}, 0, 0)$, such that $z_t = z_t^0 + \varpi_t$. The variance of the uncertainty in $\hat{T}_0^a$ is given by

$$\mathbf{var}(\hat{T}_0^a | z_0^0) = z_0^{0\prime}\mathbf{var}(\hat{\pi}|z_0^0)z_0^0 + \mathbf{var}(\varpi_0'\hat{\pi}|z_0^0) + 2z_0^{0\prime}\mathbf{cov}(\hat{\pi}, \varpi_0'\hat{\pi}|z_0^0)$$

using this notation.

If we could observe z_0^0, a 90% uncertainty interval for $\hat{T}_0^a$ would be

$$z_0^{0\prime}\hat{\pi} \pm 1.645\sqrt{z_0^{0\prime}\mathbf{var}(\hat{\pi}|z_0^0)z_0^0}$$

from the first term. Instead, rewrite the variance as

$$\begin{aligned}\mathbf{var}(\hat{T}_0^a | z_0^0) = {} & z_0^{0\prime}\mathbf{var}(\hat{\pi}|z_0^0)z_0^0 + \mathbf{var}(\hat{\alpha}_1|z_0^0)\bar{\sigma}_{1v}^2 + \mathbf{var}(\hat{\alpha}_2|z_0^0)\bar{\sigma}_{2v}^2 \\ & + \left[2z_0^{0\prime}\mathbf{cov}(\hat{\pi}, \varpi_0'\hat{\pi}|z_0^0)\right] + \left[\mathbf{var}(\varpi_0'\hat{\pi}|z_0^0) - (\mathbf{var}(\hat{\alpha}_1|z_0^0)\bar{\sigma}_{1v}^2 + \mathbf{var}(\hat{\alpha}_2|z_0^0)\bar{\sigma}_{2v}^2)\right]\end{aligned}$$

where $\bar{\sigma}_{1v}^2$ and $\bar{\sigma}_{2v}^2$ are temporal averages that estimate the variances $\sigma_{v1,t}^2$ and $\sigma_{v2,t}^2$ at $t = 0$. Specifically, we use the averages over 1999-2013 given above, so that $\bar{\sigma}_{2v}^2$ is effectively zero.

Ignoring the two terms in brackets and setting $\bar{\sigma}_{2v}^2 = 0$, a 90% uncertainty interval for $\hat{T}_0^a$ given by

$$z_0'\hat{\pi} \pm 1.645\sqrt{z_0^{0\prime}\mathbf{var}(\hat{\pi}|z_0^0)z_0^0 + \mathbf{var}(\hat{\alpha}_1|z_0^0)\bar{\sigma}_{1v}^2}$$

takes into account uncertainty in the non-volcanic forcings. Because the uncertainty in the regressors is correlated with the uncertainty in the estimator, the bracketed terms are not zero, but we expect that they will be small.

As with predictions from any linear model, the variance of the out-of-sample conditional forecasts is augmented by the estimated variance of the error term η_t. In that case, we drop $\mathbf{var}(\hat{\alpha}_1|z_0^0)\bar{\sigma}_{1v}^2$, because we are conditioning on specific data and there are no measurement errors, and we use least squares rather than CCR, in order to minimize mean squared forecast error.

References

Alexeev, V.A., Langen, P.L., and Bates, J.R.: Polar amplification of surface warming on an aquaplanet in "ghost forcing" experiments without sea ice feedbacks, Clim Dynam, 24, 655-666, https://doi.org/10.1007/s00382-005-0018-3, 2005.

Bindoff, N.L., Stott, P.A., AchutaRao, K.M., Allen, M.R., Gillett, N., Gutzler, D., Hansingo, K., Hegerl, G., Hu, Y., Jain, S., Mokhov, I.I., Overland, J., Perlwitz, J., Sebbari, R., and Zhang, X.: Detection and attribution of climate change: From global to regional, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.M., Cambridge University Press, Cambridge, 867-952, 2013.

Enfield, D.B., Mestas-Nunez, A.M., and Trimble, P.J.: The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S., Geophys Res Lett, 28, 2077-2080, https://doi.org/10.1029/2000GL012745, 2001.

Hansen, J., Sato, M., Kharecha, P., von Schuckmann, K., Beerling, D.J., Cao, J., Marcott, S., Masson-Delmotte, V., Prather, M.J., Rohling, E.J., Shakun, J., Smith, P., Lacis, A., Russell, G., and Ruedy, R.: Young people's burden: requirement of negative CO2 emissions, Earth Syst Dynam, 8, 577-616, https://doi.org/10.5194/esd-8-577-2017, 2017.

Lenton, T.M., Dakos, V., Bathiany, S., and Scheffer, M.: Observed trends in the magnitude and persistence of monthly temperature variability, Sci Rep-UK, 7, 5940, https://doi.org/10.1038/s41598-017-06382-x, 2017.

Miller, J.I.: Cointegrating regressions with messy regressors and an application to mixed-frequency series, J Time Ser Anal, 31, 255-277, https://doi.org/10.1111/j.1467-9892.2010.00660.x, 2010.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and natural radiative forcing., in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.M., Cambridge University Press, Cambridge, 659-740, 2013.

Neely III, R.R., Toon, O.B., Solomon, S., Vernier, J.-P., Alvarez, C., English, J.M., Rosenlof, K.H., Mills, M.J., Bardeen, C.G., Daniel, J.S., and Thayer, J.P.: Recent anthropogenic increases in SO2 from Asia have minimal impact on stratospheric aerosol, Geophys Res Lett, 40, 999-1004, https://doi.org/10.1002/grl.50263, 2013.

North, G.R.: Theory of energy-balance climate models, J Atmos Sci, 32, 2033-2043, https://doi.org/10.1175/1520-0469(1975)032<2033:TOEBCM>2.0.CO;2, 1975.

North, G.R. and Cahalan, R.F.: Predictability in a solvable stochastic climate model, J Atmos Sci, 38, 504-513, https://doi.org/10.1175/1520-0469(1981)038<0504:PIASSC>2.0.CO;2, 1981.

North, G.R., Cahalan, R.F., and Coakley, Jr., J.A.: Energy balance climate models, Rev Geophys and Space Ge, 19, 91-121, https://doi.org/10.1029/RG019i001p00091, 1981.

Park, J.Y., Shin, K., and Whang, Y.J.: A semiparametric cointegrating regression: Investigating the effects of age distributions on consumption and saving, J Econometrics, 157, 165-178, https://doi.org/10.1016/j.jeconom.2009.10.032, 2010.

Phillips, P.C.B., and Solo, V.: Asymptotics for linear processes, Ann Stat, 20, 971-1001, https://doi.org/10.1214/aos/1176348666, 1992.

Poppick, A., Moyer, E.J., and Stein, M.L.: Estimating trends in the global mean temperature record, Advances in Statistical Climatology, Meteorology, and Oceanography, 3, 33-53, https://doi.org/10.5194/ascmo-3-33-2017, 2017.

Pretis, F.: Econometric Models of Climate Systems: The Equivalence of Two-Component Energy Balance Models and Cointegrated VARs, University of Oxford, Department of Economics Discussion Paper, Number 750, 2015.

Rose, B.E.J. and Marshall, J.: Ocean heat transport, sea ice, and multiple climate states: Insights from energy balance models, J Atmos Sci, 66, 2828-2843, https://doi.org/10.1175/2009JAS3039.1, 2009.

Stephens, G.L., O'Brien, D., Webster, P.J., Pilewski, P., Kato, S. and Li, J.: The albedo of Earth, Rev Geophys, 53, https://doi.org/10.1002/2014RG000449, 2015.

Stevens, B. and Schwartz, S.E.: Observing and modeling Earth's energy flows, Surv Geophys, 33, 779-816, https://doi.org/10.1007/s10712-012-9184-0, 2012.

Timmermann, A., Oberhuber, J., Bacher, A., Esch, M., Latif, M., and Roeckner, E.: Increased El Niño frequency in a climate model forced by future greenhouse warming, Nature, 398, 694-697, https://doi.org/10.1038/19505, 1999.

Trenberth, K.E. and Shea, D.J.: Atlantic hurricanes and natural variability in 2005, Geophys Res Lett, 33, L12704, https://doi.org/10.1029/2006GL026894, 2006.

Vernier, J.-P., Thomason, L.W., Pommereau, J.-P., Bourassa, A., Pelon, J., Garnier, A., Hauchecorne, A., Blanot, L., Trepte, C., Degenstein, D., and Vargas, F.: Major influence of tropical volcanic eruptions on the stratospheric aerosol layer during the last decade, Geophys Res Lett, 38, L12807, https://doi.org/10.1029/2011GL047563, 2011.

Wyatt, M.A. and Curry, J.A.: Role for Eurasian Arctic shelf sea ice in a secularly varying hemispheric climate signal during the 20th century, Clim Dynam, 42, 2763-2782, https://doi.org/10.1007/s00382-013-1950-2, 2014.