ACDIS

Research Report

Global Heat Balance Model Parameter Calibration

Chenghao Ding
Department of Nuclear, Plasma, and Radiological Engineering
University of Illinois at Urbana-Champaign

Clifford E. Singer
Program in Arms Control & Domestic and International Security
University of Illinois at Urbana-Champaign

Research of the Program in Arms Control
& Domestic and International Security

University of Illinois at Urbana–Champaign

May 2022

The research for this publication is supported by the University of Illinois. It is produced by the Program in Arms Control & Domestic and International Security at the University of Illinois at Urbana-Champaign.

ACDIS Publication Series: The ACDIS *Occasional Paper* series is designed for circulating the scholarly analytical results of faculty, students, and visiting researchers associated with ACDIS. The ACDIS *Research Reports* series publishes technical reports and findings from security related research. The ACDIS *Commentary* series serves to inform U.S. and international policy decisions. The ACDIS *International Security Policy Brief* series contains previous works with a purpose similar to those in the *Commentary* series. ACDIS *Swords and Ploughshares* contains archives of a periodic journal of collected articles aimed at a general audience. For additional information and to download publications, visit the ACDIS home page at: http://acdis.illinois.edu/

Published 2022 by ACDIS//ACDIS DIN:1.2022

University of Illinois at Urbana–Champaign 359 Armory Building, 505 E. Armory Ave.
Champaign, IL 61820-6237

Series Editor: Jazmin Tejeda

Global Heat Balance Model Parameter Calibration

CHENGHAO DING
Department of Nuclear, Plasma, and Radiological Engineering
University of Illinois at Urbana Champaign

CLIFFORD E. SINGER
csinger@illinois.edu
Program in Arms Control & Domestic and International Security
University of Illinois at Urbana Champaign

Maximizing the joint probability of obtaining series of global average temperatures and rates of global thermal energy accumulation yields estimates of parameters in a global heat balance equation. One such parameter, denoted β with units $^\circ\mathrm{C}/(\mathrm{W/m^2})$, is the equilibrium ratio of global average temperature to a constant radiative forcing in the absence of large volcanic eruptions. With β, a parameter c_{th} determines the rate of relaxation $\mu = 1/(\beta c_{th})$, in inverse years, towards an equilibrium after radiative forcing is made constant. Twentieth and 21st century global average temperatures and 1991–2018 measurements of accumulated global thermal energy associated with radiative imbalance are used to estimate β and μ (and thus c_{th}), the effective efficacy of nominal radiative shielding from tropospheric aerosols, and the global average temperature difference between a reference pre-industrial temperature and an average over two more recent decades.

1. Background

This is the fourth in a series of reports describing components of a revision of an earlier form [1] of the Climate Action Gaming Experiment (CAGE). The previous reports are referred to here as CAGER1, CAGER2, and CAGER3. Their respective titles are Climate Action Game Experiment Motivation and Role of Radiative Forcing, Calibration and Extrapolation of a Simple Global Carbon Balance Model, and Non-anthropogenic Influences on Global Average Temperature. Those three reports describe information needed to calibrate parameters in the global heat balance equation

$$c_{th}\tau' = F - \tau/\beta \tag{1.1}$$

using historical estimates of global average temperature and rates of global thermal energy accumulation. In this equation, τ is operationally defined as the difference between global average temperature (herein just called temperature) and a value that would be in equilibrium with zero radiative forcing. τ' is the annual rate of change of τ. Dividing both sides of this equation by c_{th} illustrates why that parameter can be referred to as the thermal inertia. The larger c_{th} is, the smaller is the prompt effect on τ of increasing radiative forcing.

2. Radiative Forcing

By construction here, β is the ratio τ/F eventually obtained if radiative forcing F approaches a constant. However, with persistent recurrence of large volcanic eruptions, such an equilibrium is never achieved. For this reason, CAGER3 describes a method for removing from estimates of annually averaged temperature the transient effects on τ of large volcanic eruptions. The cooling effects from samples of a statistical model of future volcanic eruptions [2] could be overlain on extrapolations based on the work described here, if that were desired. CAGER3 also describes methods for removing transient effects of Schwabe c. 11 year solar irradiance cycle variations, and of variations of global average temperature correlated with the El Niño Southern Oscillation (ENSO). The results obtained here are thus meant only to describe averages over Schwabe and ENSO variations. Pursuant to the results described in CAGER3, adjustable multipliers of these nominal transient temperature effects are included here to account both for a range of uncertainty

reported for those transient effects [3] and for the different context here than that in which estimates for those effects were derived.

Radiative shielding from tropospheric aerosols has the largest recently estimated uncertainty of all the contributions to radiative forcing used here. The Working Group 1 contribution to the International Panel on Climate Change Sixth Assessment Report (here called AR6) combines such uncertainties as the square root of sums of squares [4]. Following that procedure and setting aside stratospheric aerosols from volcanic eruptions, uncertainty from the sum of tropospheric aerosol radiation interactions and aerosol cloud interactions together is 1.62 times as large as the uncertainty in radiative forcing for all of the other contributions described in CAGER1, CAGER2, and CAGER3. For this reason, total radiative forcing is written as

$$F = F_{na} + c_a F_a \tag{2.1}$$

Here F_a is the (negative) radiative forcing from tropospheric aerosols. F_{na} sums all other radiative forcing except for Schwabe cycle variations and stratospheric aerosols from volcanic eruptions. Figure 2a shows how other radiative forcing F_{na} (i.e. F with $c_a = 0$) is reduced by adding to it negative radiative forcing $c_a F_a$ for $c_a = 1$ and for the value $c_a = 0.37$ discussed below.

3. Global Average Temperature

In this report, temperature data used in estimating the parameters β, μ, and c_a is limited to annually and globally averaged thermometer measurements. How much these estimates differ from a temperature in equilibrium with zero radiative forcing is subject to an amount of uncertainty that turns out to have consequences for the present exercise. Hawkins et al. [5] identified the period 1720–1800 as a longer time to average over that is expected to have about the same radiative forcing as a shorter-term average around 1750. This idea is consistent with an only 0.004°C difference between the 1745–1755 and 1720–1800 averages of a composite of multiple non-thermometer proxies for thermometer temperatures [6].

Hawkins et al. provide two types of estimates used here for differences between 1986–2005 and eighteenth century values. For reasons described below in Appendix A, one such estimate of approximately 0.82°C is used as a reference starting point for the work here. An alternate estimate of approximately 0.70°C, based on some 1720–1800 Dutch and Central England thermometer measurements, is used as an example of how fixing a lower value without allowing for its uncertainty leads the procedure adopted here to produce results that are less compatible with GISTEMP values τ_G from the Goddard Institute for Space Studies [7].

4. Global Heating and Energy Balances

With radiative forcing F approximated as constant in each calendar year, Appendix A notes that the exact analytic solution for the annual average temperature in the kth year after 1749 for $k > 1$ is

$$\tau_k = \left(F_k(1-\kappa)/\mu + \sum_{j=2}^{i-j} \kappa^2 e^{(k-j)\mu} F_j \right) / c_{th} \tag{4.1}$$

Here $\kappa = (1 - e^{-\mu})/\mu$, and $\mu = 1/(\beta c_{cth})$ as noted above. Expanding in μ as a small parameter, κ is approximately $1 - \mu/2$. Thus, in any given year the change in τ averaged over that year from emissions F_k in that year is $(F_k/c_{th})(1-\kappa)/\mu \approx (1/2)F_k/c_{th}$. The contribution to τ averaged over that year due to emissions F_j in one of the previous years depends on the number $(k-j)$ of intervening years approximately as $e^{-(k-j)\mu}F_j/c_{th}$. While these approximations illustrate the structure of the solution, it is the exact result in equation 4.1 that is used here.

Estimates of the accumulated earth energy imbalance (EEI) between incoming and outgoing radiative energy [8] provide a useful supplement to the global average surface temperature data [9].

The rate of change of that energy imbalance is equal to the annual rate of change E'_{th} of global stored energy E_{th}. Appendix A gives the value of a coefficient c_E needed to convert annual changes E'_{th} derived from reported values of E_{th} in ZJ (i.e. in units of 10^{21} Joule) to W/m^2 averaged over the earth's area. Setting that imbalance equal to the left-hand side of equation 1.1 gives

$$c_{th}\tau' = c_E E'_{th} \tag{4.2}$$

Using the EEI data is particularly important if radiative forcing grows exponentially with time, as it approximately does in the middle curve in Figure 2a. That is because for such exponential growth it can be shown that temperature data alone cannot distinguish between a combination of a smaller values of β and c_{th} and a combination of larger values of β and c_{th}.

5. Results

Using the procedure described in Appendix A, maximizing the probability of the temperatures τ_G being obtained as functions of the parameters listed in Table 1 gives the results described in this section. Results are shown in Figure 1a for two assumptions about the difference between the initial condition temperature and the GISTEMP temperature estimates. Including an overbar to indicate an average over the 1986–2005 years used for the above-mentioned Hawkins et al. estimates, the difference between a GISTEMP input entries and a temperature in equilibrium with forcing $F = 0$ is written as $\tau_G - \overline{\tau}_G + \tau_H + \overline{\Delta\tau_{vs}} + \tau_0$. The τ_G entries are differences from a 1951–1980 average, and their 1986–2005 average is larger by $\overline{\tau}_G = 0.42°$C. $\tau_H = 0.82°$C is the above-mentioned estimate of how much the 1986–2005 average temperature exceeds a temperature in equilibrium with pre-industrial radiative forcing. An additional correction $\overline{\Delta\tau_{vs}}$ is needed because Hawkins et al. correct for ENSO-related temperature variations but do not remove transient effects of large volcanic eruptions and Schwabe solar cycle variations as done here. For the cases shown in Figure 1a and 1b respectively, the values of those corrections, rounded to the nearest 0.01°C listed are respectively $\overline{\Delta\tau_{vs}} = 0.02°$C and 0.01°C, as discussed in Appendix A. Then τ_0 is the remaining correction for uncertainty in τ_G that then remains to be estimated, along with the other parameters, by maximizing the probability of obtaining the data given the parameters.

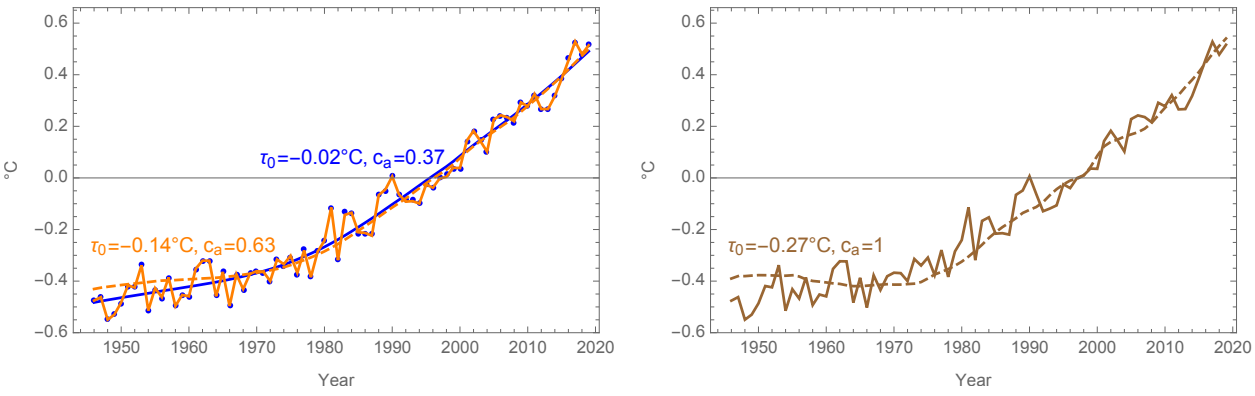

Figure 1. (a) Difference between solution τ_k in equation (4.1) of equation (1.1) and $\tau_H = 0.82°$ with probability maximizing value $\tau_0 = -0.02°$ (solid curve) and fixed value $\tau_0 = -0.14°$ (dashed curve), compared to transient-corrected data (dots and dashed curves) less τ_H. (b) As in Figure 1a but with tropospheric aerosol forcing efficacy $c_a = 1$ and resultant probability maximizing $\tau_0 = -0.27°$C.

The solid curve in Figure 1a adjusts both τ_0 and the multiplier c_a of (negative) tropospheric aerosol forcing to find a probability-maximizing fit. The dashed curve does the same except with parameter τ_0 fixed to be $-0.14°$, as if the limited number of 1720–1800 thermometer measurements referenced by Hawkins et al. provided an accurate estimate of τ_0. With this smaller value of τ_0, the numerator

τ in the loss term τ/β in the heat balance equation is smaller. That is compensated for by making the tropospheric aerosol multiplier c_a larger at 0.63. That compensates by including more negative radiative shielding in the driving function $F_{na} + c_a F_a$ and thus making that driving term also smaller. However, that approach pushes the dashed curve below the best fit curve c. 1980, which in turn is compensated with estimates for β and c_{th} that push the dashed curve above the smooth curve in and before the 1950s.

Table 1. Probability Maximizing Parameters

$\mathrm{Log}_{10}(P/\hat{P})$	0	-3.1	-11.2	Log_{10} of probability ratio
β (°C/(W/m^2))	0.518	0.507	0.581	equilibrium sensitivity
μ (1/yr)	0.068	0.105	0.415	thermal equilibration rate
c_a	0.37	0.63	***1***	tropospheric aerosols multiplier
τ_0 (°C)	-0.02	***-0.14***	-0.27	temperature offset estimate
$\overline{\Delta\tau_{vs}}$ (°C)	0.02	0.02	0.01	transients correction
c_s	0.44	0.44	0.44	solar transient multiplier
c_v	1.05	1.04	0.86	volcanic transient multiplier
c_e	1.08	1.14	1.10	ENSO transient multiplier
c_{th} ((W/m^2)(yr/°C))	28.32	18.84	4.17	thermal inertia
$1/\mu = \beta c_{th}$ (yr)	14.66	9.55	2.41	thermal equilibration timescale

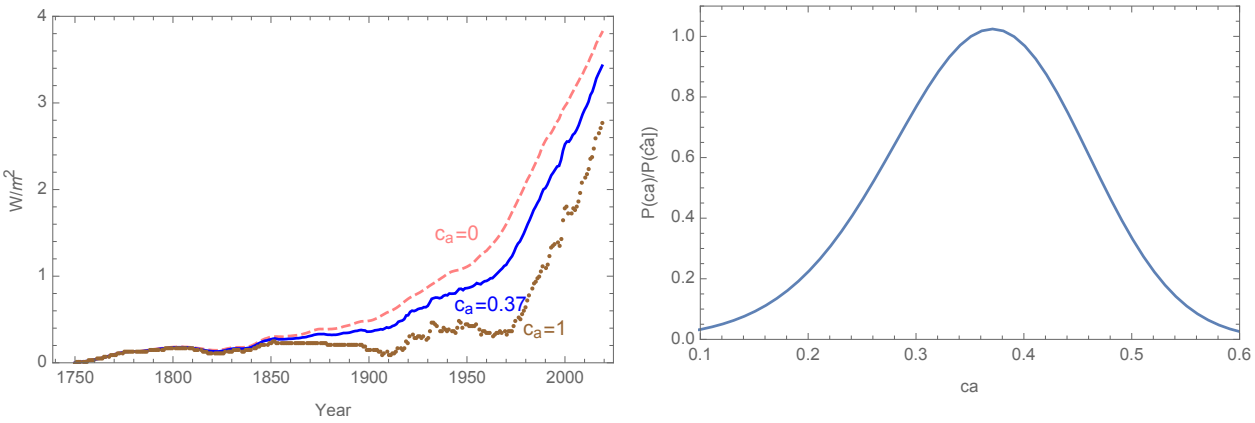

Figure 2. (a) Radiative forcing with different contributions from tropospheric aerosol shielding including none ($c_a = 0$ dashed curve), probability maximizing estimate ($c_a = 0.37$ solid curve), and AR6 nominal values ($c_a = 1$ dots). (b) Ratio of maximized probability to the maximized probability with $c_a = 0.37$, as a function of c_a.

The first two columns of numbers in Table 1 list the fitting parameters for the cases plotted in Figure 1a. (The italic bold type entries in Table 1 are fixed parameters; the others are probability maximizing parameters.) The numbers in Table 1 and most of the numbers on figures and in the text here have at most two digits after the decimal point for visual simplicity. The calculations use unrounded numbers, and some of the same results with more digits are listed in subsequent report CAGER6. The first row in Table 1 also lists base ten logarithms of the ratio of probability of the data given the fitting parameters to that for the $\tau_0 = -0.02$°C and $c_a = 0.37$ case. That probability

measure for the $\tau_0 = -0.14°C$ and $c_a = 0.65$ case is 0.0008 times that for the $\tau_0 = -0.02°C$ and $c_a = 0.37$ case.

If the tropospheric aerosol forcing is fixed at $c_a = 1$, as in Figure 1b, then the probability-maximized fitting parameters call for a solution with $\tau_0 = -0.27°C$. That fit has a very much lower probability of the data given the model parameters, as listed in the first row of numbers in Table 1.

The present work uses a simple energy balance model tuned to the observational record, as suggested by Hawkins et al., albeit one somewhat simpler than the two examples of energy balance models that those authors cited. Given the methodology differences, the small size of a $\tau_0 = -0.02°C$ addition to the Hawkins et al. estimate referred to above may be somewhat fortuitous; but it does point out that the approach used here is not necessarily at odds with that derived using a more complicated analysis.

Figures 2a and 2b are presented to clarify the points made about tropospheric aerosol shielding that are illustrated by the results in Figures 1a and 1b. Figure 2a shows total radiative forcing without tropospheric aerosol shielding ($c_a = 0$), with the nominal AR6 tropospheric aerorsol shielding $c_a = 1$, and with the maximum probability result $c_a = 0.37$. The $c_a = 1$ case has total radiative forcing dropping after 1955 and not recovering to its 1955 value until 1975. Figure 2b plots the ratio of the maximized probability for various fixed values of c_a to that for the probability maximizing value of $c_a = 0.37$. For $c_a > 0.6$, that ratio is very small.

6. Implications of Results

The CAGE project is designed to be accessible both operationally and conceptually to a wide range of policy negotiation simulation designers and participants. This goal places a premium on simplicity of the physical balance equations used. The resulting difference in methodology from studies using one or more complex global climate models (GCMs) limits comparison with results from such complex models largely to qualitative comments. For example, results presented above in Figures 1b and 2b and in Table 1 suggest that using an unmodified AR6 tropospheric aerosol shielding with the equations and solution methods used here is inappropriate. This should not, however, be taken to point out anything more than is already known about the challenges of dealing with tropospheric aerosols in GCMs. In part due to geographic and seasonal differences between aerosols and well-mixed greenhouse gases, the efficacy of a suggested level of global radiative forcing from aerosols is specific to the setup and execution of each GCM. Indeed, the overall concept of global radiative shielding from tropospheric aerosols is more useful for communicating ideas about the importance of its temporal evolution than it is of operational utility in a GCM supplied with tropospheric aerosol source terms and enough atmospheric chemistry and radiative transport detail to track their influences on climate change. Operational constraints on the CAGE design nevertheless require inclusion of a specific formula for radiative forcing of tropospheric aerosols when seeking a solution of global heat balance equation (1.1) that is consistent with the historical observations used.

Also, the formulas used here are based on a widely used notion that radiative forcing can usefully be expressed as a difference from a time period when global average temperature was approximately constant. It is then necessary to specify how that initial temperature differs from ones used in equation (1.1. The dependence on τ_0 of estimates μ and c_a in the first two columns of numbers in Table 1 suggests that it is indeed important to account for uncertainty in the value of that parameter.

One promising feature of the plots in Figure 1a is that the dots and intersections of the jagged lines are visually identical. This is even though the plotted values of transient-adjusted global average temperature data depend on values of the scaling parameters c_s, c_v, and c_e that differ from each other in the calculations leading to the dots and the intersections of the jagged lines. The jagged lines in Figures 1a and 1b are also very similar, with an exception of the value of the transient-corrected temperature data for 1992. That is the year of the peak cooling effect following

the mid-1991 Pinatubo eruption. For Figure 1b, a value of the volcanic scaling parameter with $c_v < 1$ allows the data to dip below the fitting curves in that year. This observation helps account for the difference in 1992 for the jagged lines in Figures 1a and 1b. However, as long as the parameters surveyed are close to the maximum probability parameters, fixing the transient correction scalings to their probability maximizing values can be useful for subsequent work investigating joint probability distributions for the other parameters, as described in subsequent report CAGER6.

As described in Appendix A, a number of statistical tests of assumptions of statistical independence and normality of distributions of residuals between data and equation solutions revealed one outlier that is dealt with as a unique incidence of heteroskedasticity. However, the very small probability ratios listed in the first row of numbers in Table 1 do not account for the possibility of systematic rather than random differences between data and equation solutions. The subsequent study described in report CAGER6 that examines joint probability distributions for up to four of the model parameters based on such a statistical model can thus only put a lower limit on the overall uncertainty associated with historically calibrated extrapolations. Nevertheless such an exercise could at least allow for a systematic approach to opening the question of how negotiation simulation participants respond to information about uncertainties in the physical consequences of implementing different emissions modification policies.

Appendix A. Probability Maximization Methods

The probability density for the data as a function of the model parameters is taken to be $P = P_\tau P_E$ where

$$P_\tau = (2\pi\sigma_\tau^2)^{-n_\tau/2} e^{-((1/2)\boldsymbol{\epsilon}_\tau \cdot \boldsymbol{\epsilon}_\tau/\sigma_\tau^2)} \tag{A.1}$$

and

$$P_E = (2\pi\sigma_{E*}^2)^{-n_{E*}/2} e^{-((1/2)\boldsymbol{\epsilon}_{E*} \cdot \boldsymbol{\epsilon}_{E*}/\sigma_{E*}^2)} (2\pi\sigma_{out}^2)^{-1/2} e^{-((1/2)\epsilon_{out}^2/\sigma_{out}^2)} \tag{A.2}$$

Here $\boldsymbol{\epsilon}_\tau$ is a vector of $n_\tau = 74$ values of $\tau_G + (\tau_H - \overline{\tau}_G + a_0) - \tau$ with the elements of τ given by the analytic solution in equation 4.1. The vector $\boldsymbol{\epsilon}_E$ is a list of $n_{E*} = n_E - 1 = 27$ values of $c_E E'_{th} - c_{th}\tau'$, and ϵ_{out} is the "outlier" value for $c_E E'_{th} - c_{th}\tau'$ computed from differences from the years 2002 and 2001. The constant C_E is equal to $10^{21}/(A_E s_y)$ with $A_E = 5.1 \times 10^{14}\text{m}^2$ the area of the earth, $s_y = 31557600$ the average number of seconds in a year after 1901, and values of E'_{th} in ZJ/yr. The annual growth rates of global energy from the radiative forcing less radiative loss imbalance are approximated as $E_{k+1} - E_k$ with E_k the energy changes since 1960 for 1990–2018 from von Schuckmann et al. [8].

The overall solution to the heat balance equation with F equal to a different constant F_k for each kth year after 1749 is a sum of the solutions τ_k with a zero source in every year except the kth one. Dividing equation 1.1 by c_{th} and setting $g_k = F_k/c_{th}$, $\mu = 1/(c_{th}\beta)$, and $s_k = t - t_k$ gives, in differential equation notation, $d\tau_k/ds_k = g_k u_k - \mu\tau_k$. Here $u_k = 1$ during year t_k and 0 otherwise. Rewriting the result as $d(e^{\mu s_k}\tau_k)/ds_k = g_k u_k e^{\mu s_k}$ and using the initial condition $\tau_k = 0$ when $s_k = 0$ gives $\tau_k = g_k\mu(1 - e^{-\mu s_k})$ for $0 \le s_k \le 1$. Integrating forward with 0 source from a resulting value of $\tau_k = \kappa g_k(1 - e^{-\mu})/\mu = \kappa g_k$ when $s_k = 1$ gives $\tau_k = \kappa g_k e^{-\mu s_k}$ for $s_k \ge 1$. Averaging this result over each year and summing the resulting averages gives the formula above in equation 4.1.

The formula for $\overline{\Delta\tau_{vs}}$ correction to the input temperature data τ_G that is mentioned in the first paragraph in Section 5 is

$$\overline{\Delta\tau_{vs}} = c_v\overline{\tau}_v + c_s\overline{\tau}_s \tag{A.3}$$

Here $\overline{\tau}_v = 0.024°\text{C}$ is the amount by which the average temperature for 1986–2005 after removing transient volcanic cooling exceeds that average without removing that cooling effect if $c_v = 1$. The corresponding correction for the Schwabe cycle effect with $c_s = 1$ is $-0.015°\text{C}$. With the probability maximizing values of c_v and c_s listed in Table 1, the resulting values of $\overline{\Delta\tau_{vs}}$ are 0.018°C, rounded

to 0.02°C in the first colum in Table 1, and 0.014°C rounded to 0.01°C in the third column in Table 1.

Appropriate prior probability distributions for σ_τ and σ_E are taken to be proportional to their inverses. Evaluating $\int_0^\infty d\sigma_\tau \int_0^\infty d\sigma_{E*} \int_0^\infty d\sigma_{out} P/(\sigma_\tau \sigma_{E*} \sigma_{out})$ [10] gives a result that is a constant times

$$(\boldsymbol{\epsilon}_\tau \cdot \boldsymbol{\epsilon}_\tau)^{-n_\tau/2} (\boldsymbol{\epsilon}_{E*} \cdot \boldsymbol{\epsilon}_{E*})^{-n_{E*}/2} (e_{out}^2)^{-1/2} \tag{A.4}$$

The value of the constant multiplying this expression is not needed, since it will be maximized with respect to the model parameters by minimizing

$$L = -(n_\tau/2) \ln(\boldsymbol{\epsilon}_\tau \cdot \boldsymbol{\epsilon}_\tau) - (n_E/2) \ln(\boldsymbol{\epsilon}_{E*} \cdot \boldsymbol{\epsilon}_{E*}) - (1/2) \ln(e_{out}^2) \tag{A.5}$$

There are 74 years from 1946–2019 and 28 averages from 1991.5–2017.5 of years used to compute values of E'_{th}. Starting with temperature data from 1946 avoids a World War II measurement anomaly from 1941–1945 [11]. This approach also avoids a period of lower global geographic coverage for temperature measurements during and before WWII, as illustrated by the upper curve in Figure 3a. Once per decade reported GISTEMP global coverage fractions from 1910 through 1940 are nearly constant with an average of 0.7, similar to that shown in Figure 3a. For GISTEMP, that number then jumps to 1 in 1950 because the GISTEMP averaging method allows for some weight in geographical locations within 1200 km of measurement locations. A recent analysis that includes geographic coverage and additional sources of uncertainty suggests about twice as much GISTEMP uncertainly from 1905 to the start of WWII than after WWII [12]. Figure 3a also compares temperature estimates from the Hadley Centre Climate Research Unit (HADCRUT) [13, 14] to those from the Goddard Institute for Space Studies estimates used here. There are some larger differences in the decade before 1906 than thereafter. No use is made here of temperature data before 1906, not even for comparison with results using temperature data from 1906–2019 to results using temperature data from 1946–2019 as in Table 2. The data range chosen for E'_{th} avoids a period of less complete measurements prior to the 1990s [8]. Not surprisingly, the part before 1990 of the earth energy imbalance estimates plotted in Figure 3b shows larger scatter than in the later estimates.

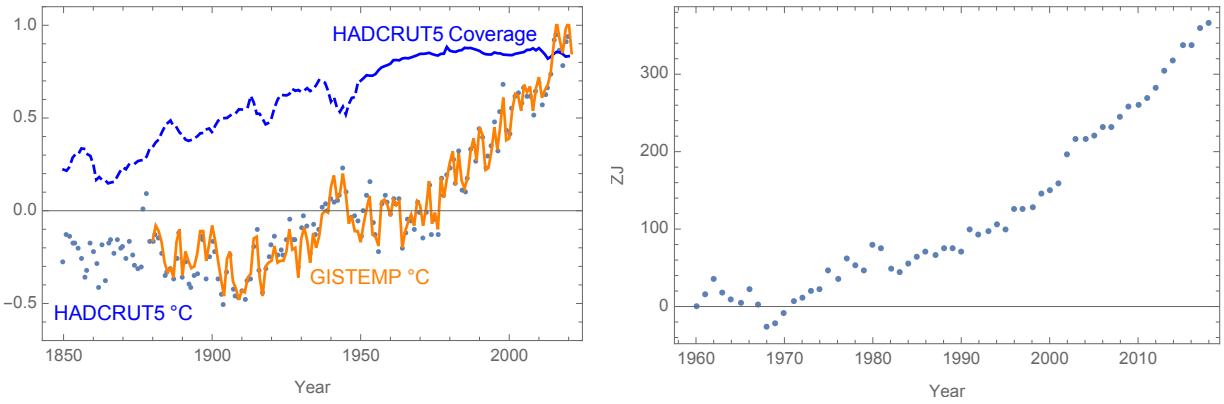

Figure 3. (a) Un-modified annual average temperature estimate differences from a 1951–1980 average (lower jagged curve for GISTEMP and dots for HADCRUT5); and fractional global temperature measurement coverage (upper curve, solid for 1946–2019 GISTEMP data and dashed for earlier data). (b) Earth energy imbalance estimates less 1960 value.

The choice of data ranges used here also avoids years that lead the data sets to include indications of temporal autocorrelation in the residuals between the data and the maximum probability fits. Including temperature data before 1939 leads to an estimated temporal autocorrelation regression

of type AR1. Including data prior to 1991 for the earth energy imbalance leads to an estimated temporal autocorrelation of moving average type MA1. If these correlations were properties of the entire physical system, the AR1 type would indicate a bias in the parameter estimates with AR1 included in the analysis. That should not be a problem here for extrapolating observable future temperatures, since it is the actual temperatures and not the underlying parameter values that are of primary interest for CAGE studies. If the temporal autocorrelation is a property of the earlier data measurements alone, then including the earlier data would only unnecessarily complicate the analysis. Accounting for MA1 temporal autocorrelation would not be expected to bias the parameter estimates. However, it would both change a subsequent joint probability distribution and likely tractably but substantially complicate the computational analysis thereof.

Other questions about statistical assumptions used are whether each of the two data streams used are heteroskedastic, skewed, kurtotic, or interdependent. Testing the maximum deviation ratio of absolute value to standard deviation residuals [15, 16] identifies one earth energy imbalance outlier point at a confidence level of 95.3%. That is for the 2001.5 EEI residual. That result motivated allowing that residual to have a different σ than the other EEI residuals. (For the τ residuals, the corresponding confidence level for the residual with the largest absolute value was 69.9% and not taken to require a heteroskedastic treatment.) The resulting kurtosis of the residuals at 2.50 for temperature and 3.32 for EEI, as compared to 3 for a normal distribution. The skewness of temperature residuals is small at 0.38. With homoskedasticity assumed for all of the EEI residuals, the skewness of the EEI residuals is 0.75. With the 2001.5 EEI residual allowed to have a different σ, the skewness of the rest of the EEI residuals is 0.31. After applying this one heteroskedasticity correction, the departures from normality were not taken to be severe enough to introduce the complication of dealing with non-normal alternatives to equations (A.1) and (A.2).

Table 2. Alternative Statistical Results

Type	β	μ	c_a	τ_0	c_{th}
Reference	0.518	0.068	0.37	-0.02	28.3
Homoskedastic	0.526	0.062	0.36	-0.02	30.5
E'_{th} outlier dropped	0.526	0.062	0.36	-0.02	30.5
Temperature from 1906	0.505	0.066	0.18	0.06	30.1

Table 2 compares results for fits to β, μ, c_a, and τ_0 as in Table 1 to results using three alternative analyses. Except with some more digits to emphasize differences, the numbers in the first row are the same as Table 1, i.e. with the 2001.5 EEI residual allowed to have a different σ. For the next row in Table 2, all of the EEI residuals are forced to have the same σ. In that case, the large outlier forces a larger estimate for $c_{th} = 1/(\beta\mu)$.

Neither of the results from the two different statistical models shown in Table 2 suggests changing the conclusion that an efficacy of $c_a = 1$ for the nominal AR6 tropospheric aerosol shielding is incompatible with the model and data used here. The heteroskedastic approach changes the weighting of portions of the data stream used but does not change the estimates for β appreciably. Dropping an $c_E E'_{th} - c_{th}\tau'$ outlier with a 2.6 times the root mean square has little effect on the probability maximizing values of β and c_a.

There is some interdependence of the underlying global average temperature and earth energy imbalance data. In particular, a small portion of that energy imbalance is stored in the atmosphere, land and ice (about 12%) [8] and in a top ocean layer that is well mixed on a time short compared to the 26-year span of time of the EEI data used. The effective depth of the well mixed top ocean layer has been estimated to be about 75 meters [17]. The value of c_{th} converted to $(\mathrm{J/m^2})/^\circ\mathrm{C}$ and

divided by the ocean fraction 0.709 of the earth's area and by the volumetric heat capacity of sea water gives a total effective average depth of 307 meters for the heated part of the ocean in the global heat balance model, in the approximation that all of the energy imbalance is stored in the ocean. Due at least in part to this overlap in data sources for τ and E_{th}, a subsequent CAGER6 report on uncertainty analysis under the assumption of independence of measurements of global average temperature and earth energy imbalance will just be aimed at estimating a lower limit on the overall uncertainty.

Also shown in Table 2 are parameters estimated with the same analysis used for the reference case but with the temperature data range extended back to 1906. That was done primarily to see whether the mismatches with data shown for alternative cases in Figures 1a and 1b result from using only part of the available temperature data. A temperature difference 0.21°C is subtracted from the 1941–1946 GISTEMP values to account for a World War II sea surface temperature anomaly [11] times the ocean fraction of the earth's surface area. Using the 1906–2019 temperature data thus corrected, the $c_a = 1$ case shows the same type of mismatches with the data as in Figure 1b. The case with the longer temperature data used yields a probability maximization value of $c_a = 0.18$.

Since the data extended back to 1906 have AR1 temporal correlation, the estimates in the last line of Table 2 would be biased to different values if the AR1 autocorrelation were accounted for in the analysis. The parameters listed in Table 2 could nevertheless be used as an alternative to extrapolating future temperatures with AR1 accounted for, sampling how observations chosen from an AR1 distribution differ from the extrapolation, and fitting a curve through the result. However, given that the variation of a WWII measurement anomaly during years 1941–1945 [11] is not very clear, and in light of the lower global coverage back to 1906 shown in Figure 3a, that more complicated approach is not adopted here. Instead, the point of including the $c_a = 0.18$ result in the last line of numbers in Table 2 is to note that using a larger data range does not appear to challenge the conclusion that an efficacy factor of $c_a = 1$ multiplying the nominal AR6 tropospheric aerosol forcing is incompatible with the database and analysis method used here.

References

[1] Singer, C. and L. Matchett, 2015. Climate action gaming experiment: Methods and example results, *Challenges* **6**, 202–228, doi:10.3390/challe6020202.

[2] Ammann, C. M., and P. Naveau. 2010. A statistical volcanic forcing scenario generator for climate simulations, *Journal of Geophysical Research* **115**, Issue D5, March 16, https://doi.org/10.1029/2009JD012550.

[3] Foster, G., and S. Rahmstorf. 2011. Global temperature evolution 1979–2020, *Environmental Research Letters* **6**, 044022, http://dx.doi.org/10.1088/1748-9326/6/4/044022.

[4] IPCC, 2021. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Pèan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.). Cambridge University Press. In Press, https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/.

[5] Hawkins, E., P. Ortega, E. Suckling, A. Schurer, G. Heferl, P. Jones, M. Joshi, T. J. Osborn, V. Masson-Delmotte, J. Mignot, P. Thorne, and G. J. van Oldenborgh. 2017. Estimating changes in global temperature since the preindustrial period, *Bulletin of the American Meteorological Society*, **98**, 1841–1856, https://doi.org/10.1175/BAMS-D-16-0007.1.

[6] Mann, E. M., and P. D. Jones. 2003. Global surface temperatures over the past two millennia, *Geophysical Research Letters* **30**, No. 15, 1820, August 2003, https://doi.org/10.1029/2003GL017814.

[7] GISTEMP Team. 2021. *GISS Surface Temperature Analysis (GISTEMP), version 4.* NASA Goddard Institute for Space Studies, https://data.giss.nasa.gov/gistemp/; https://data.giss.nasa.gov/gistemp/tabledata_v4/GLB.Ts+dSST.csv, accessed December 22, 2021.

[8] von Schuckmann, K., L. Cheng, M. D. Palmer, J. Hansen, C. Tassone, V. Aich, S. Adusumilli, H. Beltrami, T. Boyer, F. J. Cuesta-Valero, D. Desbruyères, C. Domingues, A. García-García, P. Gentine, J. Gilson, M. Gorfer, L. Haimberger, M. Ishii, G. C. Johnson, R. Killick, B. A. King, G. Kirchengast, N. Kolodziejczyk, J. Lyman, B. Marzeion, M. Mayer, M. Monier, D. P. Monselesan, S. Purkey, D. Roemmich, A. Schweiger, S. L. Seneviratne, A. Shepherd, D. A. Slater, A. K. Steiner, F. Straneo, M.-L. Timmermans, and S. E. Wijffels. 2020.

Heat stored in the Earth system: where does the energy go? *Earth System Science Data* **12**, 2013–2041, https://doi.org/10.5194/essd-12-2013-2020.

[9] Aldrin, M., M. Holden, P. Guttorp, R. Skeie, G. Myhre, and T. Bernsten. 2012. Bayesian estimation of climate sensitivity based on a simple climate model fitted to observations of hemispheric temperatures and global ocean heat content, *Environmetrics* **23**, 253–271, doi:101002/env.2140.

[10] Box, G. E. P., and G. C. Tiao. 1992. *Bayesian Inference in Statistical Analysis*, New York: Wiley.

[11] Chan, D., and P. Huybers. 2021. Correcting observational biases in sea surface temperature observations removes anomalous warmth during World War II, *Journal of Climate Change* **34**, 4585-4602, https://doi.org/10.1175/JCLI-D-20-0907.1.

[12] Lenssen, N. J. L., G. A. Schmidt, J. E. Hansen, M. J. Menne, A. Persin, R. Ruedy, and D. Zyss. 2019. Improvements in the GISTEMP uncertainty model, *Journal of Geophysical Research: Atmospheres* **124**, 6307–6326. https://doi.org/10.1029/2018JD029522.

[13] Osborn, T. J., P. D. Jones, D. H. Lister, C. P. Morice, I. R. Simpson, J. Winn, E. Hogan, and I. C. Harris. 2021. Land surface air temperature variations across the globe updated to 2019: the CRUTEM5 dataset. *Journal of Geophysical Research: Atmospheres* **126**, January 27, https://doi.org/10.1029/2019JD032352.

[14] Climate Research Unit. 2021. Temperature, https://crudata.uea.ac.uk/cru/data/temperature/#filfor, accessed September 22, 2021.

[15] National Institute of Standards and Technology. 2012. NIST/SEMATECH e-Handbook of Statistical Methods, Section 1.3.5.17.1, Grubb's Test, https://www.itl.nist.gov/div898/handbook/eda/section3/eda35h1.htm, accessed January 2, 2022.

[16] Zaiontz, C. 2022. Grubb's Test, https://www.real-statistics.com/students-t-distribution/identifying-outliers-using-t-distribution/grubbs-test, accessed January 4, 2022.

[17] Marten, A. L. 2011. Transient temperature response modeling in IAMs: The effects of over simplification on the SSC. *Economics E-journal* **5**, 2011-18, http://dx.doi.org/10.5018/economics-ejournal.ja.2011-18.