The Fractional Energy Balance Equation for Climate projections through 2100

Roman Procyk[1], Shaun Lovejoy[1], and Raphael Hébert[2,3]

[1]Physics Dept., McGill University, 3600 rue University, Montreal, Quebec, H3A 2T8, Canada
[2]Alfred-Wegener Institute Helmholtz Center for Polar and Marine Research, Telegrafenberg A45, 14473 Potsdam, Germany
[3]Institute of Geosciences, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam, Germany

Correspondence: Roman Procyk (roman.procyk@mail.mcgill.ca)

Abstract.

We produce climate projections through the 21^{st} century using the fractional energy balance equation (FEBE)~~which is~~: a generalization of the standard ~~EBE~~energy balance equation (EBE). The FEBE can be derived ~~either~~ from Budyko-Sellers models~~or phenomenologically by applying the~~, or phenomenologically through the application of scaling symmetry to energy storage processes~~. It is~~, easily implemented by changing the integer order of the storage (derivative) term in the EBE to a fractional value~~near 1/2.~~.

The FEBE ~~has two shape~~ is defined by three parameters: a ~~scaling exponent H and~~ fundamental shape parameter, a timescale and an amplitude, corresponding to, respectively, the scaling exponent h, the relaxation time τ~~; its amplitude parameter is~~, and the equilibrium climate sensitivity (ECS). Two additional parameters were needed for the forcing: an aerosol re-calibration factor α to account for the large aerosol uncertainty, and a volcanic intermittency correction exponent ν. A Bayesian framework based on historical temperatures and natural and anthropogenic forcing series was used for parameter estimation. Significantly, the error model was not ad hoc, ~~but was~~ rather predicted by the model itself: the internal variability response to white noise internal forcing.

The 90% Confidence Interval (CI) of the ~~shape parameters were $H = [0.33, 0.44]$~~ exponent and relaxation time were $h = [0.33, 0.44]$ (median=0.38), and $\tau = [2.4, 7.0]$ (median=4.7) years compared to the usual EBE ~~$H = 1$~~$h = 1$, and literature values of τ typically in the range ~~2-8 years. We found that aerosols~~ 2–8 years. Aerosol forcings were too strong, requiring a decrease by an average factor $\alpha = [0.2, 1.0]$ (median=0.6)~~and~~; the volcanic intermittency correction exponent was $\nu = [0.15, 0.41]$ (median=0.28) compared to standard values $\alpha = \nu = 1$. The overpowered aerosols support a revision of the global modern (2005) aerosol forcing 90% CI to a narrower range [-1.0,-0.2]Wm^{-2}. ~~The~~ For the IPCC forcings, the key parameter ECS = [1.6, 2.4]K (median = $2.0K$) compared with the IPCC AR5 range [1.5, 4.5]K (median = $3.2K$). Similarly, we found the transient climate sensitivity (TCR) = [1.2, 1.8]K (median = $1.5K$) compared to the AR5 range TCR = [1.0, 2.5]K (median =$1.8K$). As ~~commonly~~ often seen in other observational-based studies, the FEBE values for climate sensitivities are therefore somewhat lower but still consistent with those in IPCC AR5.

Using these parameters we made projections to 2100 using both the Representative Carbon Pathways (RCP) and Shared Socioeconomic Pathways (SSP) scenarios~~and shown alongside the~~, and compared them to the corresponding CMIP5 ~~/6 MME~~and

CMIP6 multi-model ensembles (MMEs). The FEBE ~~hindprojections (1880-2019~~historical reconstructions (1880–2020) closely follow observations~~(~~, notably during the 1998–2014 slowdown ("hiatus"~~, 1998-2015).~~). We also reproduce the internal variability with the FEBE and statistically validate this against centennial scale temperature observations. Overall the FEBE projections were 10-15% lower but due to their smaller uncertainties, their 90% CIs lie completely within the GCM 90% CIs. ~~The FEBE thus complements and supports the GCMs~~This agreement means that the FEBE validates the MME and vice versa.

1 Introduction

The Earth is a complex, heterogenous system with turbulent atmospheric and oceanic processes operating over scales ranging from millimeters up to planetary scales. When considered by time scale, there are three main regimes: the weather, macroweather and climate ~~((Lovejoy and Schertzer, 2013), (Lovejoy, 2013))~~(Lovejoy and Schertzer, 2013; Lovejoy, 2013). From dissipation times ~~up until scales of about~~ ranging within the scale of ten days - the lifetime of planetary structures - fluctuations in the temperature and other atmospheric quantities increase with time scale: this is the weather regime. Beyond this - in macroweather - fluctuations generally decrease with scale: longer ~~and longer averages tend to converge~~averaged tending more towards convergence. Eventually, ~~in the industrial epoch at a scale of ≈ 20 years,~~ this is reversed and fluctuations again tend to increase~~. This marks~~ , marking the beginning of the climate regime~~,~~. In the industrial epoch this occurs at a scale of ≈ 20 years, while in the pre-industrial epoch the transition is at centuries or millennia and the regime continues up to Milankovitch scales ~~((Lovejoy, 2015b), (Lovejoy, 2019b))~~(Lovejoy, 2015b, 2019b).

A major challenge is determining the Earth's decadal and centennial response to anthropogenic and natural perturbations. At the moment, projection uncertainties - famously exemplified in the range 1.5–$4.5K$ for a CO_2 doubling – are quite large so that for many purposes, including the development of mitigation policies, the development of complementary approaches ~~is desirable. In~~ are needed. When considering alternatives, ~~perhaps the most important point is that~~ although perturbations to the Earth system can be quite varied, when compared to the mean solar radiation, over the past and future decades, those of interest are of the order of only a few percent. This allows diverse forcings to be conveniently approximated by their equivalent radiative forcings. It also explains why ~~–~~ - in spite of their highly nonlinear weather dynamics - that to a good approximation, ~~GCM~~ General Circulation Model (GCM) macroweather and climate responses to external perturbations are typically linear (as quantified for CMIP5 models in ~~(Hébert and Lovejoy, 2018)~~ Hébert and Lovejoy (2018)) but with stochastic internal variability.

In order to construct macroweather and climate models, beyond linearity and stochasticity, we require additional model constraints, the classical one being energy balance. Starting with the first Energy Balance Models (EBMs) proposed by ~~(Budyko, 1969) and (Sellers, 1969)~~Budyko (1969) and Sellers (1969), EBMs have been extensively used for understanding the climate ~~((North, 1975), (North et al., 1981, Reviews))~~(North, 1975; North et al., 1981, Reviews; Trenberth et al., 2014; North and Kim, 20

. In this paper, we will only consider EBMs for the globally averaged temperature~~that can be obtained by latitudinally averaging Budyko-Sellers models~~. The resulting “zero dimensional” energy balance equation (EBE) is a first order linear differential equation, it can alternatively be obtained by considering the Earth to be a uniform slab of material (“box”) radiatively exchanging heat with outer space. Such box models usually involve at least two boxes and they assume Newton’s law of cooling as well as ad hoc assumptions relating surface temperature gradients to the rate of heat exchange.

Energy conservation is an important symmetry principle, yet when implemented in box type models, it violates another symmetry: scale invariance. This is because box models are integer ordered differential equations whose response functions (Green’s functions) are exponentials. In order to respect the scaling, these “climate response functions” (CRFs) have therefore been postulated to be scaling (~~power laws~~power-law). However, the use of pure power law CRFs ~~(e.g. (Rypdal, 2012), (Myrvoll-Nilsen et al., 2020)~~ (e.g. Rypdal, 2012; Myrvoll-Nilsen et al., 2020) leads to divergences: the "runaway Green’s function effect" (Hébert and Lovejoy, 2015) so that if the Earth is perturbed by even an infinitesimal step function forcing, its temperature monotonically increases without ever attaining thermodynamic equilibrium: its ~~equilibrium climate sensitivity (ECS)~~ ECS is infinite. Whereas the classical EBMs conserve energy but violate scaling, the pure scaling CRF models are scaling but violate energy conservation. Such models can only make projections by using forcings that start and then return to zero.

~~(Hébert et al., 2021)~~ Hébert et al. (2021) proposed taming the divergences by cutting off the power law CRFs at small scales. The resulting model was scaling at long times and when forced by step functions, ~~it~~ reaches thermodynamic equilibrium. With this truncated power law CRF and using Bayesian techniques ~~(Hébert et al., 2021)~~ Hébert et al. (2021) were able to make climate projections through 2100 with the IPCC RCP scenario forcings that were coherent with the MME 90% confidence interval. Furthermore, using the historical part of each GCM simulation, ~~they were able to accurately reproduce the~~ corresponding GCM climate ~~projection: for the Earth’~~projections were accurately reproduced, meaning (in regards to the Earth’s globally averaged temperature~~,~~) both models were ~~thus~~ effectively equivalent. The ~~main drawback~~ caveat was that the CRF model truncation was somewhat ad hoc~~. In addition, due to the truncation, it was~~ , and therefore only useful at decadal or longer scales.

To make more realistic models, the key issue is energy storage. Storage is a consequence of imbalances in incoming short wave and outgoing long wave radiation and it must be accounted for in applications of the energy balance principle (Trenberth et al., 2009). As pointed out in ~~((Lovejoy, 2019b), (Lovejoy, 2019a))~~ Lovejoy (2019a, b) and developed in ~~(Lovejoy et al., 2021)~~ Lovejoy et al. (2021) it is sufficient that the scaling principle not be applied to the CRF, but rather to the storage term in the EBE. ~~Rather than~~ In lieu of the energy being stored by uniformly heating a box, energy is instead stored in a hierarchy of structures from small to large, each with time constants that are power laws of their sizes. This conceptual shift can be implemented simply by changing the integer order of the storage (derivative) term in the EBE to a fractional value: the Fractional Energy Balance Equation (FEBE). While ~~(Lovejoy et al., 2021)~~ Lovejoy et al. (2021) derived the FEBE in a phenomenological manner, ~~((Lovejoy, 2021a), (Lovejoy, 2021b), (Lovejoy, 2020))~~ Lovejoy (2021a, b) showed how it could instead be derived from the continuum mechanics heat equation used in the Budyko-Sellers models. Indeed, by extending Budyko-Sellers models from 2D to 3D (i.e. to include the vertical) and imposing the (correct) conductive – radiative surface boundary conditions, one immediately obtains fractional order equations for the surface temperature. In other words, nonclassical fractional equations and long memories turn out to be necessary consequences of the standard Budyko-Sellers approach.

To understand the FEBE's key new features, recall that linear differential equations can be solved with Green's functions; in the classical integer ordered case, these are based on exponentials. However in the general case where one or more terms are of fractional order, they are instead based on "generalized exponentials", themselves based on power laws. In the FEBE, there are two distinct power law regimes with a transition at the relaxation time (estimated to be of the order of a few years, see below). While the low frequency Green's function can be very close to ~~Hébert'~~ Hébert et al. (2021)'s truncated power law CRF, the high frequency regime is able to produce internal variability coherent with the observed scaling and fractional Gaussian noise used ~~by ((Lovejoy et al., 2015), (Del Rio Amador and Lovejoy, 2019), (Del Rio Amador and Lovejoy, 2020))~~ for skillfully forecasting the stochastic (internal) variability at monthly, seasonal, interannual (macroweather) scales (Lovejoy et al., 2015; Del Rio A . In short, there are theoretical arguments as well as empirical evidence that the FEBE accurately models the Earth's temperature response to both internal and external forcing over macroweather and climate time scales.

~~This paper has 3 sections: methods and materials, results, and conclusions. In methods and materials section, we~~ The following text will introduce the standard EBE and generalize it to the FEBE, describe the radiative forcing, temperature and GCM simulations that will be used, and ~~finally we~~ introduce Bayesian inference for determining the model and forcing parameters. ~~In the results section,~~ Using these we present the probability distribution functions for the parameters, estimate Equilibrium Climate Sensitivity (ECS) and Transient Climate Response (TCR), and finally we produce global projections to 2100 using the Representative Carbon Pathways (RCPs) and Shared Socioeconomic Pathways (SSPs) which ~~are a little cooler than the~~ we compare to corresponding CMIP5 ~~GCMs Multi-Model Ensemble (MME)and the currently available~~ and CMIP6 ~~GCMs with which they are compared~~GCM Multi-Model Ensembles (MMEs).

2 Methods and Material

2.1 The FEBE

The zero-dimensional FEBE may be written:

$$\tau^{\cancel{H}h}{}_{-\infty}D^{\cancel{H}h}_{t}T+T=\cancel{\lambda F}s\mathcal{F};\quad \mathcal{F}(t)=F(t)+f(t),\quad 0\le \cancel{H}h\le 1 \tag{1}$$

~~((Lovejoy, 2019a)), (Lovejoy et al., 2021))~~ (Lovejoy, 2019a; Lovejoy et al., 2021). Where $T(t)$ is the Earth temperature anomaly with respect to a reference temperature ~~$(T(-\infty)=0)$~~ $(\lim_{t\to-\infty}T(t)=0)$, τ is the relaxation time, ~~λ~~ s is the climate sensitivity, ~~$F(t)$~~ $\mathcal{F}(t)$ is the anomalous external radiative forcing ~~, and H~~ which is the sum of the stochastic $f(t)$ and deterministic $F(t)$ components, and h is the order of the ~~("Weyl") fractional derivative :~~

Caputo fractional derivative (see Podlubny, 1999):

$${}_{-\infty}D^{\cancel{H}h}_{t}T=\cancel{\frac{1}{\Gamma(1-H)}}\frac{1}{\Gamma(1-h)}\int_{-\infty}^{t}(\cancel{t'-s}t-u)^{\cancel{H}-h}T'(\cancel{t'})\cancel{ds}dt',\quad T'(\cancel{s}u)=\cancel{\frac{dT}{ds}}\frac{dT}{du} \tag{2}$$

~~(~~Γ is the Gamma function~~)~~. If this derivative is integrated by parts and the limit $\cancel{H \to 1}$ $h \to 1$ is taken, using $\cancel{(T(-\infty)=0)}$, $\cancel{{}_{-\infty}D_t^H T = \frac{dT}{dt}}$ $\lim_{t\to+\infty} T(t) = 0$, ${}_{-\infty}D_t^h T = \frac{dT}{dt}$ so that we recover the standard box EBE ~~. The FEBE thus generalizes the EBE and the power law FEBE relaxation time~~ $\cancel{\tau}$ ~~generalizes the exponential EBE relaxation time. Physically~~ $\cancel{T/\lambda}$ ~~corresponds to the linearized energy flux (rate per area) of long wave black body surface emission and since~~ $\cancel{F}$ ~~is the flux of shortwave forcing, we see that the imbalance~~ $\cancel{F - T/\lambda}$ ~~represents the flux of energy conducted into storage, it is proportional to the derivative term. At any instant, the storage is thus proportional to the power law weighted integral of past imbalances. The exception is the box model with~~ $\cancel{H=1}$ ~~where the relationship is instantaneous.~~ (Lovejoy et al., 2021).

If we solve the FEBE using Green's functions, we obtain:

$$T(t) = \cancel{\lambda} s \int_{-\infty}^{t} G_{\cancel{\delta,H}\,0,h}(\cancel{t-s}\,t-u)\,\cancel{F}\,\mathcal{F}(\cancel{s}\,u)\,\cancel{ds}\,du \tag{3}$$

Where $\cancel{G_{\delta,H}}$ $G_{0,h}$ is the impulse (Dirac) response Green's function, for the FEBE it is given by:

$$G_{\cancel{\delta,H}\,0,h}(t) = \begin{cases} \tau^{-1}\left(\frac{t}{\tau}\right)^{h-1} E_{h,h}\left(-\left(\frac{t}{\tau}\right)^{h}\right); & t \geq 0 \\ 0; & t < 0 \end{cases} \tag{4}$$

Where:

$$E_{\alpha,\beta}(z) = \sum_{k}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)} \tag{5}$$

is the "α,β order Mittag-Leffler function" (these and most of the following results are in the notation of ~~(Podlubny, 1999))~~ Podlubny (1999). The condition $\cancel{G=0}$ $G_{h,x}(t) = 0$ for $t < 0$ $(x \in \mathbb{N})$ is needed to respect causality, in what follows we will implicitly assume that for all Green's functions. The Mittag-Leffler functions are often called "generalized exponentials", the classical $\cancel{H=1}$ $h = 1$ box model is the (exceptional) ordinary exponential: $E_{1,1}(z) = e^z$.

Mathematically when $\cancel{0<H<1}$ $0 < h < 1$, the FEBE is a "fractional relaxation equation" ~~,~~ where τ quantifies the slow, power law approach to a new thermodynamic equilibrium. Rather than express solutions in terms of the impulse response $\cancel{G_{\delta,H}}$ $G_{0,h}$, it is ~~more physical and~~ often more convenient to use the step response $\cancel{G_{\Theta,H}:}$

$G_{1,h}$:

$$G_{\cancel{\Theta,H}\,1,h}(t) = \int_{0}^{t} G_{\cancel{\delta,H}\,0,h}(\cancel{s}\,u)\,\cancel{ds}\,du = \left(\frac{t}{\tau}\right)^{\cancel{H}\,h} E_{\cancel{H,H+1}\,h,h+1}\left(-\left(\frac{t}{\tau}\right)^{\cancel{H}\,h}\right). \tag{6}$$

~~(~~$\cancel{\Theta(t)}$ ~~is the step or Heaviside function, the integral of the Dirac function), and:~~

Such that the temperature response can be written as:

$$T(t) = \cancel{\lambda} s \int_{-\infty}^{t} G_{\cancel{\Theta,H}\,1,h}(\cancel{t-s}\,t-u)\,F'(\cancel{s}\,u)\,\cancel{ds}\,du, \quad F'(\cancel{s}\,u) = \cancel{\frac{dF}{ds}}\,\frac{dF}{du}. \tag{7}$$

$G_{1,h}$ has the advantage of being dimensionless, and also has a simple interpretation as being the response to a step forcing such as that found in numerical CO_2 doubling experiments. At high frequencies ($t \ll \tau$), important for modelling and predicting the internal variability we have:

$$G_{\delta,H,high\,0,h,high}(t) = \frac{1}{\tau\Gamma(H)}\frac{1}{\tau\Gamma(h)}\left(\frac{t}{\tau}\right)^{H-1\,h-1}; \quad G_{\Theta,H,high\,1,h,high}(t) = \frac{1}{\Gamma(H+1)}\frac{1}{\Gamma(h+1)}\left(\frac{t}{\tau}\right)^{H\,h}; \quad t \ll \tau \tag{8}$$

These correspond to taking the first terms in the series expansions for the Mittag-Leffler functions in eqns. 4, 6. If we consider the response to Gaussian white noise forcing, ~~then $G_{\delta,H} \propto t^{H-1}$~~ $\gamma(t)$, then $G_{0,h} \propto t^{h-1}$ implies that $T(t)$ is approximately a fractional Gaussian noise (fGn) with statistical scaling exponent ~~$H_I = H - 1/2$~~ h (when forced by a Gaussian white noise, the FEBE response is exactly a fractional Relaxation noise ~~, see (Lovejoy, 2020)). By applying global scale Haar fluctuation analyses (Del Rio Amador and Lovejoy, 2019) found $H_I \approx -0.1$ corresponding to $H \approx 0.4$.~~ (see Lovejoy, 2019a).

To see if this is compatible with the value estimated from the low frequency response to external forcings consider the low frequency behaviour ($t \gg \tau$), important for modelling and projecting the multidecadal responses to external forcing:

$$G_{\delta,H,low\,0,h,low}(t) = \frac{-1}{\tau\Gamma(-H)}\frac{-1}{\tau\Gamma(-h)}\left(\frac{t}{\tau}\right)^{-1-H\,-1-h}; \quad G_{\Theta,H,low\,1,h,low}(t) = 1 - \frac{1}{\Gamma(1-H)}\frac{1}{\Gamma(1-h)}\left(\frac{t}{\tau}\right)^{-H\,-h}; \quad t \gg \tau \tag{9}$$

(note ~~$\Gamma(-H) < 0$ for $0 < H < 1$~~ $\Gamma(-h) < 0$ for $0 < h < 1$). In the box, ~~$H = 1$~~ $h = 1$, case we have exactly ~~$G_{\Theta,1}(t) = 1 - e^{-t/\tau}$ so that when $H < 1$~~ $G_{1,1}(t) = 1 - e^{-t/\tau}$ whereas when $h < 1$, the exponential approach to thermodynamic equilibrium is replaced by a power law. ~~(Hébert et al., 2021) used $G_\Theta = 1 - \left(1 + \frac{t}{\tau}\right)^{H_F}$ with $H_F = -0.5$ corresponding to $H = -H_F = 0.5$~~ Hébert et al. (2021) used $G_1 = 1 - \left(1 + \frac{t}{\tau}\right)^{H_F}$ with $H_F \approx -0.5^{+0.4}_{-0.5}$ corresponding for $t \gg \tau$ to $h = -H_F \approx 0.5$ which is thus the same h value as that corresponding to the internal forcing. It is thus plausible that the FEBE models both high and low frequency regimes with the ~~exponent $H \approx 0.4$. We discuss this further below~~unique exponent $h \approx 0.4$. Indeed it was this empirical finding that pre-dated and motivated the discovery of the FEBE.

2.2 Data

2.2.1 Radiative Forcing Data

We consider natural and anthropogenic sources of external forcing: solar and volcanic, ~~and anthropogenic forcing~~greenhouse gas and aerosol. We use the ~~approximate~~ standard semi-empirical carbon dioxide concentration to forcing relationship (Myhre et al., 1998):

$$F_{CO_2}(\rho) = 3.71 W m^{-2} log_2 \frac{\rho}{\rho_0}. \tag{10}$$

Where F_{CO_2} is the forcing due to carbon dioxide, ρ is the concentration of carbon dioxide and ρ_0 is the preindustrial concentration of carbon dioxide which we take to be 277ppm (Solomon, 2007).

We follow the CMIP5 recommendations for anthropogenic and solar forcing, while volcanic forcing is unprescribed (Taylor et al., 2012). The anthropogenic CMIP6 radiative forcings follow ~~(Smith et al., 2018a)~~Smith et al. (2018a).

2.2.2 Greenhouse Gas Forcing

The global climate is warming and most of the observed changes are due to increases in the concentration of anthropogenic greenhouse gases (GHGs) (IPCC, 2013). Future anthropogenic forcing is prescribed in the Representative Concentration Pathways (RCPs), established by the IPCC for CMIP5 simulations: we considered RCP 2.6, RCP 4.5, and RCP 8.5 (Meinshausen et al., 2011b). RCP 6.0 was omitted in this study since fewer CMIP5 modelling groups performed the associated run. In the CMIP6 simulations the anthropogenic forcings are prescribed in the Shared Socioeconomic Pathways (SSPs) (Meinshausen et al., 2020); we investigate SSP ~~126~~ 1-26 (strong mitigation)~~and SSP 585~~ , SSP 2-45 (middle of the road) and SSP 5-85 (strong emission) scenarios, ~~the~~ designated as high priority for IPCC AR6 and are counterparts to the ~~most distinct previous scenarios RCP 2.6 and 8.5~~previous RCP scenarios above.

The RCP scenarios are derived from estimates of emissions computed by a set of Integrated Assessment Models (IAM), these emissions are converted to concentrations using the Model for the Assessment of Greenhouse-gas Induced Climate Change ~~(MAGICC, (Meinshausen et al., 2011a)),~~ (MAGICC, Meinshausen et al., 2011a), while for the SSP scenarios the emissions are converted to forcings using the Finite Amplitude Impulse Response ~~(FAIR) model ((Smith et al., 2018a))~~model (FAIR, Smith et al., 2018a) . These scenarios will allow us to compare our results from the FEBE with CMIP5/6 simulations.

The wide spread of the scenarios allows for the investigation of the consequences of various future policies, from strong mitigation (RCP 2.6, SSP ~~126) to business as usual~~ 1-26) to no-policy reference (RCP 8.5, SSP ~~585~~5-85) shown in ~~fig.~~ figure 1 (bottom). For ~~RCP2.6 and SSP 126~~RCP 2.6 and SSP 1-26, the strongest mitigation scenarios, the total radiative forcing has a peak at approximately $3Wm^{-2}$ around the year 2050 and declines thereafter due to large scale deployment of negative emission technologies. ~~RCP4.5 is a stabilization scenario~~RCP 4.5 and SSP 2-45 are stabilization scenarios, with the total radiative forcing rising until the year 2070 and with stable concentrations after the year 2070. While ~~RCP8.5 and SSP 585 is a~~ RCP 8.5 and SSP 5-85 are continuously rising radiative forcing ~~pathway, "business as usual",~~ pathways, in which the radiative forcing levels by the end of the ~~21st century at~~ 21^{st} century reaches approximately $8.5Wm^{-2}$~~, and most closely follows current emissions~~ . Current emissions fall somewhere between the $8.5Wm^{-2}$ and $4.5Wm^{-2}$ scenarios.

In this paper we use the forcing due to carbon dioxide equivalent, $F_{CO2_{Eq}}$, as the measure of our anthropogenic forcing, F_{Ant}, given in the RCP and SSP scenarios. The anthropogenic forcing corresponds to the effective radiative forcing produced by ~~Long Lived Greenhouse Gases (GHGs)~~ long lived GHGs F_{GHG}: carbon dioxide, methane, nitrous oxide and fluorinated gases, controlled under the Kyoto protocol, and ozone depleting substances, controlled under the Montreal Protocol. We show the anthropogenic forcings for each RCP and SSP scenario in figure 1.

2.2.3 Aerosol Forcing

Aerosols are a strong component of radiative forcing associated with anthropogenic emissions, resulting from a combination of direct and indirect aerosol effects. There exists high uncertainty of the aerosol forcing, arising from a poor understanding of how clouds respond to aerosol perturbations (Penner et al., 2001; Ramaswamy et al., 2001), compared to the fairly well constrained GHG forcing~~, thus following ((Padilla et al., 2011), (Hébert et al., 2021)) we~~ . We therefore follow (Forest et al., 2002; Harvey and Kaufmann, 2002; Forest et al., 2006; Padilla et al., 2011; Hébert et al., 2021) and introduce the aerosol linear scaling factor α to account for our poor knowledge of aerosol forcing.

We obtained the CMIP5 aerosol forcing ~~by subtracting~~ from the total $CO_{2_{EQ}}$ forcing by subtracting the combined effective radiative forcing ~~from~~ of the gases controlled by the Kyoto protocol, F_{Kyt}, and from those controlled under the Montreal protocol, F_{Mtl}. F_{Mtl} is given in CFC-12 equivalent concentration and we use the relation from ~~(Ramaswamy et al., 2001)~~ Ramaswamy et al. (2001) to convert this to Wm^{-2}.

The total amount of aerosol forcing in 2005 given at the 90% ~~confidence interval (CI)~~ CI in the IPCC AR5 is $[-1.9, -0.1]Wm^{-2}$, but since then attempts have been made to better constrain this value; ~~(Stevens, 2015) demonstrates that an aerosol forcing~~ Stevens (2015) argues that extreme aerosol forcings (more negative than $-1Wm^{-2}$ ~~is~~) are implausible. Using results from ~~(Murphy et al., 2009), Stevens~~ Murphy et al. (2009), Stevens (2015) supports tightening the upper and lower bounds of the aerosol forcing, revising it to ~~be~~ $[-1.0, -0.3]Wm^{-2}$ although the wider range from the IPCC's AR5 is still supported by the more comprehensive study by Bellouin et al. (2020).

The prescribed CMIP6 SSP aerosol forcing, $F_{Aer_{SSP}}$, ~~contain~~ contains contributions from aerosol-radiation interactions and from cloud interactions: F_{ari} and F_{aci} (Smith et al., 2018a). F_{ari} includes the direct radiative effect of aerosols, in addition to rapid adjustments due to changes in the atmospheric temperature, humidity and cloud profile (formerly the semi-direct effect), and is calculated using multi-model results from Aerocom (Myhre et al., 2013). F_{aci} describes how aerosols affect clouds in the radiation budget and is calculated from the aerosol model of ~~(Stevens, 2015)~~ Stevens (2015), which includes a logarithmic dependence of F_{aci} on sulphates, black carbon and organic carbon emissions ~~.~~ - the source of the difference in aerosol forcing shapes between $F_{Aer_{RCP}}$ and $F_{Aer_{SSP}}$ shown in figure 1 (bottom).

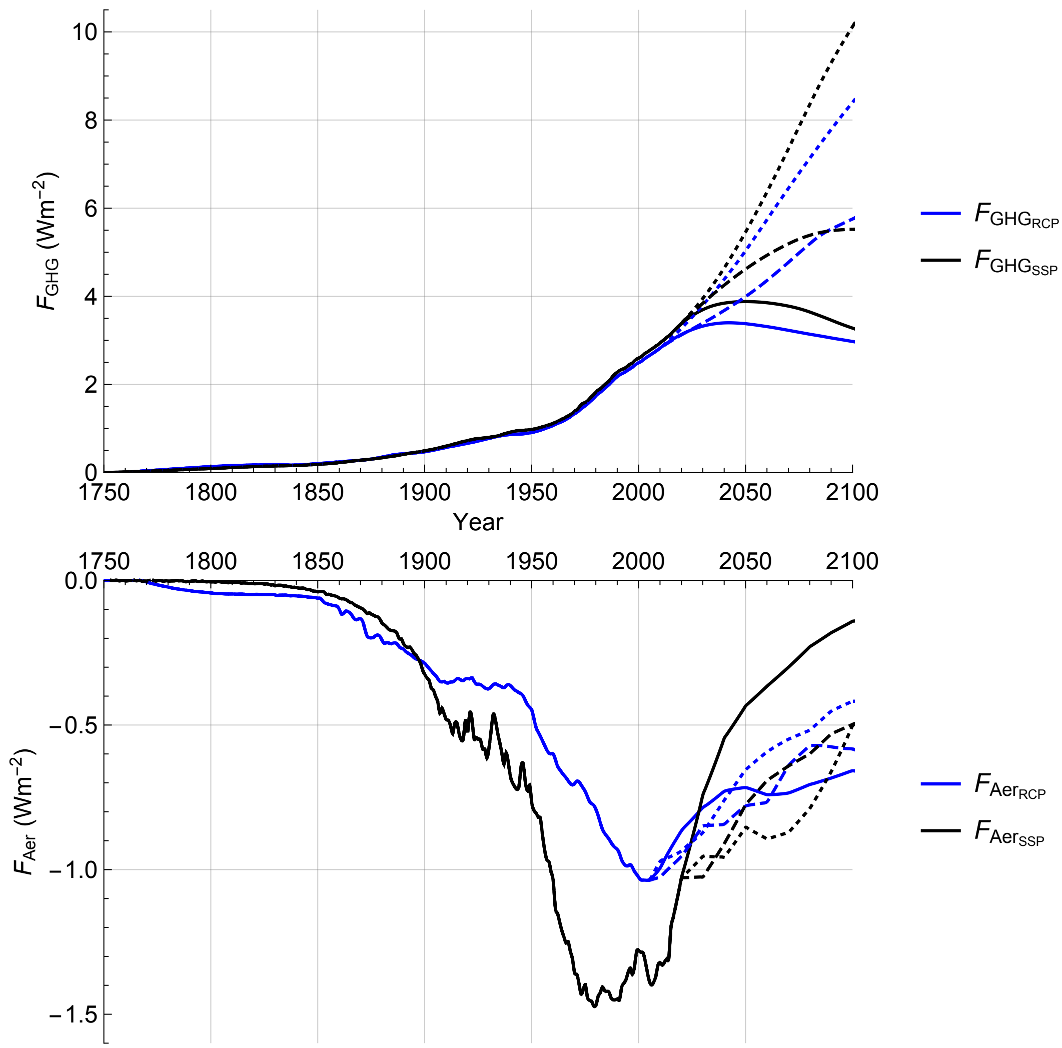

Figure 1. (top) The ~~total~~ anthropogenic forcing series, the sum of the greenhouse gas forcing F_{GHG} and respective aerosol forcing series $F_{Aer_{RCP}}$ (black) or $F_{Aer_{SSP}}$ (blue) are shown over the historical period and projection period until 2100 for RCP 2.6/SSP ~~126~~ 1-26 (solid), RCP 4.5/SSP 2-45 (dashed), and RCP 8.5/SSP ~~585~~ 5-85 (dotted).
(bottom) The anthropogenic aerosol forcing series used, $F_{Aer_{RCP}}$ (blue) and $F_{Aer_{SSP}}$ (black) following the same scheme as above. Updated from ~~(Hébert et al., 2021)~~ Hébert et al. (2021).

2.2.4 Solar Forcing

The other external forcings considered are solar and volcanic; there exist other natural forcings such as mineral dust and sea salt, but they are small and will be implicitly included with the internal variability. We use the CMIP5 recommendation for solar forcing, F_{Sol}, a reconstruction obtained by regressing sunspot and faculae time series with total solar irradiance (TSI) (Wang et al., 2005), shown in ~~fig.~~ figure 2. Following ~~(Meinshausen et al., 2011b),~~ Meinshausen et al. (2011b), the solar forcing anomaly is calculated as the change in solar constant over the average value of the two ~~year~~ 11-year solar cycles from 1882 to 1904 divided by 4 (~~average insolation~~the effective fraction of the surface of the Earth which is exposed to the sun) and multiplied by 0.7 (representing planetary co-albedo). To extend solar forcing to the future we follow CMIP5 and reproduce solar cycle 23 (the last one prior to 2008) as ~~a proxy for~~ the assumed future solar forcing.

2.2.5 Volcanic Forcing

The volcanic forcings series, F_{Vol}, used in this study was generated from the volcanic optical depths, τ_V. Over the 1850 to 2012 period we use the approximate relation: $F_{Vol} \approx -27Wm^{-2}\tau_V$, obtained from the Goddard Institute for Space Science (GISS) website (Sato, 2012). We follow ~~(Hébert et al., 2021)~~Hébert et al. (2021), extending the series to 1765 using the optical depth reconstruction ~~(Crowley et al., 2008)~~of Crowley et al. (2008), and setting volcanic forcing to zero for the future.

It ~~has been shown by (Lewis and Curry, 2015)~~ is well established that volcanic forcing must be scaled down by 40-50% in order to produce a comparable effect on surface temperature, and thus most EBMs linearly scale volcanic forcing (Tomassini et al., 2007; Ring . However the amplitude of the volcanic forcing is not the only issue~~,~~ ; volcanic forcings are highly intermittent (spiky). The intermittency can be quantified in a ~~multi-fractal framework ((Lovejoy and Schertzer, 2013), and (Lovejoy and Varotsos, 2016))~~ multifractal framework (Lovejoy and Schertzer, 2013; Lovejoy and Varotsos, 2016) by the intermittency parameter C_1 which corresponds to the fractal codimension (i.e. 1-D, where D, is the fractal dimension of the part of series that gives the dominant contribution to the mean of the series) characterizing the sparseness of volcanic "spikes" of mean amplitude. Since linear response models do not alter the intermittency, the volcanic series must first be non-linearly transformed before being introduced into a linear response framework. With the effective volcanic forcing F_{Vol_ν}, the volcanic intermittency correction exponent ν and the mean of the whole volcanic series $\langle F_{Vol} \rangle$, we follow ~~(Hébert et al., 2021)~~ Hébert et al. (2021) using a non-linear relation to change the intermittency so that the transformed signal can be linearly related to the temperature:

$$\frac{F_{Vol_\nu}}{\langle F_{Vol} \rangle} = \frac{F_{Vol}^{\nu}}{\langle F_{Vol}^{\nu} \rangle} \quad (11)$$

The normalization is such that the mean is unchanged: $\langle F_{Vol_\nu} \rangle = \langle F_{Vol} \rangle$ (this is slightly different than the normalization used in Hébert et al. (2021)). The volcanic intermittency correction exponent, ν, required to reduce the intermittency parameter of the volcanic forcing, C_{1,F_V}, to equal the corresponding parameter of the temperature response, C_{1,T_V}, can be calculated theoretically using:

$$C_{1,F_V}\nu^{\alpha_{MF}} = C_{1,T_V} \tag{12}$$

where α_{MF} is the multifractality index of the volcanic forcing~~(Lovejoy and Schertzer, 2013)~~, C_1 is the codimension of the mean ~~. For the volcanic forcing intermittency , $C_{1,F_V} \approx 0.16$, the temperature response intermittency, $C_{1,T_V} \approx 0.03$, and~~ (see ch. 4, Lovejoy and Schertzer, 2013).

The volcanic response appears to be non-linear as the intermittency ("spikiness", sparseness of the spikes) parameter C_1 changes from about $C_{1,F_V} \approx 0.16$ for the input volcanic forcing to $C_{1,T} \approx 0.03$ for the temperature response: the latter is therefore much less intermittent than the former although it is possible that the estimated C_1 changes slightly due to finite size effects and internal variability. Assuming $\alpha_{MF} \approx 1.5$, we find an approximate but plausible theoretical estimate of the volcanic intermittency correction exponent $\nu \approx 0.3$ ~~((Lovejoy and Schertzer, 2013) table 11.8, (Lovejoy and Varotsos, 2016))~~(Lovejoy and Schertzer, 2013; Lovejoy and Varotsos, 2016).

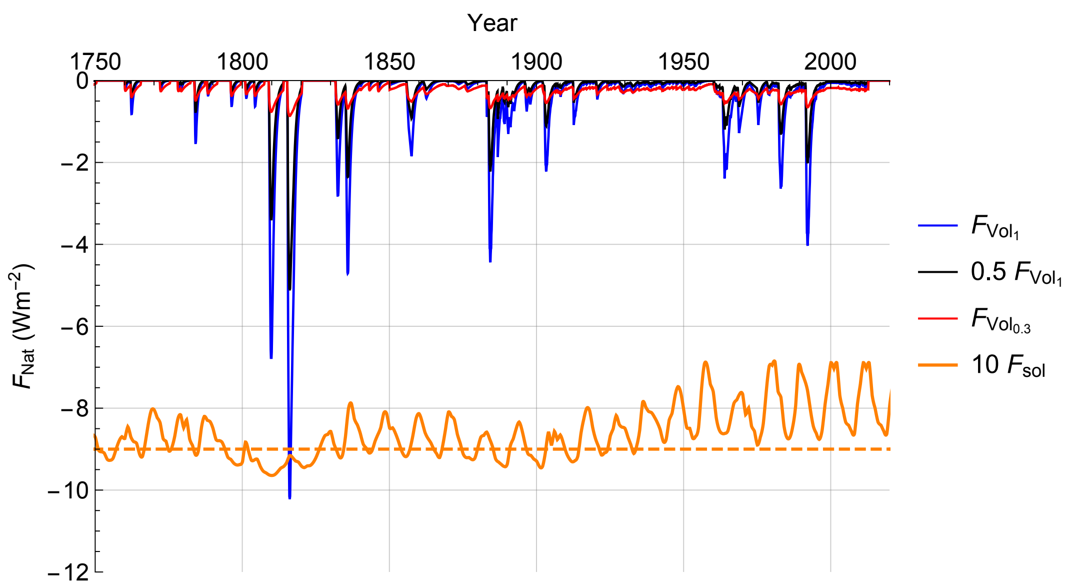

Figure 2. Volcanic forcing F_{Vol_1} (blue) is shown alongside two ~~non-linearly~~ transformed versions~~. Linearly~~: linearly damped by a constant 0.5 coefficient (black), and non-linearly transformed using equation with $\nu = 0.3$ (red). The solar forcing F_{Sol} (orange) has been shifted down by -9 and amplified by a factor of 10 for clarity. Adapted from ~~(Hébert et al., 2021)~~Hébert et al. (2021)

2.2.6 Internal Stochastic Forcing

We consider the standard assumption about internal variability that it is forced by a Gaussian “delta correlated” white noise:

$$f(t) = \sigma\gamma(t); \quad \langle\gamma(t)\rangle = 0; \quad \langle\gamma(t)\gamma(u)\rangle = \delta(t-u), \tag{13}$$

where $f(t)$ is the noise at infinite resolution, $\gamma(t)$ is a "unit" white noise and σ is its amplitude. When averaged to resolution $\tau_r = 1$ month, the average forcing has amplitude $\langle f_{\tau_r}^2\rangle^{1/2} = \sigma_\tau = \frac{\sigma}{\sqrt{\tau_r}}$. In comparison, the internal variability of the mean observational temperature series is equal to the observed series with the forced temperature response removed. We take the global annually averaged monthly temperature anomaly to be $\sigma_{T,\tau_r} \approx \pm 0.14$°C, where τ_r is the resolution (taken to be monthly in this case).

Using Lovejoy et al. (2021) and σ_{T,τ_r}, we can relate σ_{T,τ_r} and σ_{f,τ_r}:

$$\sigma_{f,\tau_r} = \frac{\sigma_{T,\tau_r} K_h}{s}\left(\frac{\tau}{\tau_r}\right)^h, \tag{14}$$

$$K_h = \sqrt{\frac{\pi}{2cos(\pi\left(h-\frac{1}{2}\right))\Gamma(-1-2h)}}. \tag{15}$$

Where K_h is a standard normalization constant, τ is the relaxation time, and s is the climate sensitivity parameter; eq. 14 is an approximation valid at short time scales $\tau_r \ll \tau$. If we introduce a white noise forcing, with the standard deviation calculated using eq. 14, the FEBE response will correspond to an internal variability term with realistic amplitude and autocorrelation structure.

Working in a linear framework we write the forcing series, $\mathcal{F}$, as the sum of ~~anthropogenic and natural forcings:~~ the deterministic forcings, F,(GHG, aerosol, solar and volcanic) and the white noise forcing:

$$\cancel{F}\mathcal{F}(\alpha,\nu;t) = F_{GHG}(t) + \alpha F_{Aer}(t) + F_{Sol}(t) + F_{Vol_\nu}(t) + \sigma_{f,\tau_r}\gamma_{\tau_r}(t); \quad F(t) = \langle\mathcal{F}(t)\rangle = F_{GHG}(t) + \alpha F_{Aer}(t) + F_{Sol}(t) + F_{Vol_\nu}(t), \tag{16}$$

where $\gamma_{\tau_r}(t)$ is a unit white noise at resolution τ_r and $\langle\cdot\rangle$" is the mean ensemble (statistical) average.

2.2.7 Surface Air Temperature Data and CMIP5/6 Simulations

We used five historical records of surface air temperature for our analysis each spanning the period ~~1880-2019~~1880-2020, with median monthly temperature anomalies in relation to the reference period of 1880-1910: HadCRUT4 (Morice et al., 2012), the Cowtan & Way reconstruction version 2.0 ~~(C&W) (Cowtan and Way, 2014b)~~(C&W, Cowtan and Way, 2014b, a; Cowtan et al., 2015), GISS Surface Temperature Analysis ~~(GISTEMP) (Hansen et al., 2010)~~(GISTEMP, Lenssen et al., 2019), NOAA Merged Land Ocean Global Surface Temperature Analysis Dataset ~~(NOAAGlobalTemp) ((Smith et al., 2008), (Zhang et al., 2019))~~ (NOAAGlobalTemp, and Berkley Earth Surface Temperature ~~(BEST) (Rohde et al., 2013)~~(BEST, Rohde and Hausfather, 2020).

The HadCRUT4 dataset is a combination of the sea-surface temperature records: HadSST3 compiled by the Hadley Centre of the UK Met Office along with land surface station records: CRUTEM4 from the Climate Research Unit in East Anglia; the

Cowtan and Way dataset uses HadCRUT4 as raw data, but ~~address~~ interpolates missing data that would lead to bias especially at high latitudes by infilling missing data using an optimal interpolation algorithm (kriging); we use the dataset with land air temperature anomalies interpolated over sea-ice. The GISTEMP dataset combines the Global Historical Climate Network version 3 (GHCNv3) land surface air temperature records with the Extended Reconstructed Sea Surface Temperature version 4 (ERSST) along with the temperature dataset from the Scientific Community on Antarctic Research (SCAR) and is compiled by the Goddard Institute for Space Studies; the NOAA National Climate Data Center uses GHCNv3 and ERSST but applies different quality controls and bias adjustments. The final data, BEST, makes use of its own land surface air temperature product along with a modified version of HadSST.

The CMIP5 models selected have monthly historical simulation outputs available over the 1860 to 2005 period along with outputs of scenario runs from 2005 to 2100 for RCP 2.6, RCP 4.5, and RCP 8.5, summarized in table A1. The ~~available~~ CMIP6 model outputs ~~for the historical and SSP scenarios are provided in (Forster et al., 2020), and~~ have monthly historical simulations from 1860 to 2014 and future projections based on the SSP scenarios 1-26, 2-45 and 5-85 (Forster et al., 2020), climate sensitivity of models ~~is taken from (Flynn and Mauritsen, 2020),~~ are summarized in table A2 (Flynn and Mauritsen, 2020)
.

2.3 Bayesian Parameter Estimation

In this section we establish a procedure to estimate the probability distribution associated with the climate sensitivity: $\cancel{\lambda}$ s, model parameters: τ, ~~H and the~~ h and forcing parameters: α, ν. To estimate them, we relate the forcing to surface air temperature data using the FEBE ~~in~~ with a multi-parameter Bayesian technique. To apply Bayesian inference we require temperature observations, a statistical model that relates forcing data to temperature, and prior information about the model parameters (priors). Bayesian inference is chosen due to its ability to better constrain model parameters by using information from different sources including data and models.

~~In this framework ,~~ Through this framework each parameter combination ($\cancel{H}$ h, τ for $\cancel{G_{\delta,H}}$ $G_{0,h}$ and α, ν for F as well as $\cancel{\lambda}$ s) produces a time-dependent forced response which is associated with a likelihood that depends on how well the corresponding model output matches the observational temperature records over the historic period. To see how this works, recall that the FEBE describes the temperature response to the ~~total~~ sum of the external deterministic forcing $F(t)$ and internal stochastic forcing ~~$\gamma(t)$:~~

$\sigma\gamma(t)$:

$$T(t) = T_{ext}(t) + T_{int}(t); \quad \begin{matrix} T_{ext}(t) = sG_{0,h}(t) * F(t) \\ T_{int}(t) = sG_{0,h}(t) * \sigma\gamma(t) \end{matrix} \cancel{.}, \tag{17}$$

~~Where~~ where T_{ext}, T_{int} are the responses. Any given set of parameters (~~H~~ h, τ for ~~$G_{\delta,H}$~~ $G_{0,h}$ and α,ν for F as well as ~~λ) implies a series~~ s) defines a forced temperature response $T_{ext}(t)$; and ~~hence when compared to~~ when removed from the observation temperature series, ~~it implies~~ they define a series of residuals:

$$T_{res}(t) = T(t) - T_{ext}(t) = T_{int}(t) = \cancel{\lambda} s G_{\cancel{\delta,H}\,0,h}(t) * \sigma\gamma(t). \tag{18}$$

The residuals are thus equal to the internal temperature variability, i.e. the response to the internal forcing ~~$\gamma(t)$~~ $\sigma\gamma(t)$. Here we make the usual assumption that $\gamma(t)$ is a Gaussian white noise so that $T_{res}(t) = T_{int}(t)$ is a fractional Relaxation noise process (fRn, ~~(Lovejoy, 2019a)~~ Lovejoy (2019a)). However, for scales shorter than the relaxation time τ (of the order of years), the fRn process is very close to a fractional Gaussian noise (fGn) process (due to the approximation ~~$G_{\delta,H} \approx G_{\delta,high,H}(t)$~~ $G_{0,h} \approx G_{0,high,h}(t)$, eq. 8). Thus, rather than making an ad hoc assumption about the statistics of the residuals, in our approach the statistics are given by the model itself (a key improvement from Hébert et al. (2021)). The fGn approximation takes into account the strong power law correlations induced by the fractional derivative term in the FEBE and it is generally valid except at the low frequencies that only weakly influence the likelihood function. ~~It is already a much more accurate residual model than the standard Auto Regressive order one (~~ An fGn model for the residuals is more realistic with respect to the autocorrelation function of temperature data (Lovejoy et al., 2015) and thus produces more conservative confidence interval in comparison to other exponential decorrelation models such as an AR(1) ~~) assumption which models temporal correlation as a (short range) exponential function. Rather than making an ad hoc assumption about the statistics of the residuals, in our approach the statistics are given by the model itself. It should be noted that if the residuals are fGn, then the uncertainties are larger than for AR(1) (or other exponential decorrelation models) in spite of the fact that our residuals are more realistic.~~ since the latter underestimate the decorrelation time, and thus overestimate the effective sample size.

~~We therefore estimate the likelihood of any parameter set from the maximum likelihood function that the residuals are an fGn process:~~

$$\cancel{\mathcal{L}(\lambda,H,\tau,\alpha,\nu|T(t)) = Pr(T(t)|\lambda,H,\tau,\alpha,\nu).}$$

To calibrate the FEBE, we take the time-dependent forced response calculated for each parameter combination and remove it from the temperature series to obtain a series of residuals which represent an estimator of the historical internal variability. The likelihood function ($\mathcal{L}$) corresponds to the probability ("Pr") of observing the series $T(t)$ conditioned on the parameters: ~~$\lambda,H,\tau,\alpha,\nu$~~ s,h,τ,α,ν (right hand side), assuming the residuals are a fGn process with parameter ~~H~~ h, and zero mean~~.~~:

$$\mathcal{L}(s,h,\tau,\alpha,\nu|T(t)) = Pr(T(t)|s,h,\tau,\alpha,\nu). \tag{19}$$

~~The fGn likelihood function is a posterior probability; using~~ Using Bayes' rule, we can obtain the ~~a priori~~ posterior probability distribution function (PDF) for our parameters ~~:~~

~~$$Pr(\lambda, H, \tau, \alpha, \nu | T(t)) = \frac{Pr(T(t)|\lambda, H, \tau, \alpha, \nu)\pi(\lambda, H, \tau, \alpha, \nu)}{Pr(T(t))}$$~~

~~where $\pi(\lambda, H, \tau, \alpha, \nu)$ is the~~ using the likelihood function (an a priori probability) and the prior distribution for the parameters~~.~~, $\pi(s, h, \tau, \alpha, \nu)$:

$$Pr(s, h, \tau, \alpha, \nu | T(t)) = \frac{Pr(T(t)|s, h, \tau, \alpha, \nu)\pi(s, h, \tau, \alpha, \nu)}{Pr(T(t))} \tag{20}$$

We use the following Mathematica 12.2 (Wolfram Research, Inc., 2020) functions: LogLikelihood[proc, data], FractionalGaussianNoiseProcess[μ, σ', h'], and EstimatedProcess[data, proc] to calculate the maximum likelihood of those residuals to be a fGn corresponding to our error model. Note that the Hurst exponent h' used within Mathematica 12.2 describes the scaling behaviour of the associated fractional Brownian motion obtained by integrating the fGn. The notation $h = h' - 1/2$ corresponds to the associated parameter in Lovejoy et al. (2015) which directly describes the scaling associated with the fluctuations of the fGn itself.

The priors chosen here are intended to reflect knowledge about the historical climate system. Following ~~(Del Rio Amador and Lovejoy, 20~~ ~~who estimated H~~ Del Rio Amador and Lovejoy (2019) who estimated h from the statistics of the response of the internal forcing, the prior distribution for the scaling parameter is taken to be a normal distribution centered around 0.4 with a standard deviation of 0.1 (twice that of ~~the original work),~~ Del Rio Amador and Lovejoy (2019), i.e N(0.4,0.1)). For the relaxation time~~,~~ τ, we use the normal distribution of the fast time response of the "two-box" exponential model that corresponds to ~~$H = 1$~~ $h = 1$, found by ~~(Geoffroy et al., 2013)~~ Geoffroy et al. (2013) for a suite of 12 CMIP5 GCMs: N(4yrs,2yrs), with the standard deviation doubled of the original work so as to be a weakly informative prior. When considering the aerosol scaling parameter, α, we take the prior distribution to be a normal distribution, N(1.00,0.55) which has a 90% CI and mean coherent with the IPCC AR5 best range for the modern value of aerosol forcing, $F_{Aer} \approx -1.0 Wm^{-2}$, in the series we used. For the remaining two parameters, ~~λ~~ s and ν, we assume non-informative uniform priors over the range of parameters; ~~$\lambda \in [1.0, 4.0]$~~ $s \in [1.0, 4.0]$ and $\nu \in [0.0, 1.0]$. All prior distributions are independent.

~~We approximate our five dimensional~~ Using Bayes, eq. 20, we then fit a multivariate Gaussian distribution to our five-dimensional parameter space, posterior distribution $Pr(s, h, \tau, \alpha, \nu | T(t))$, which will be used to draw sets of parameters to generate future forced temperature projections. The multivariate Gaussian approximation is built by using the means and variances of all parameters through integrating the joint probability to obtain five marginal probabilities, and calculating the covariance between all pairwise parameters using their "joint" marginal distributions as to take into account potentially large correlations. The five dimensional posterior parameter space, ~~$(\lambda, H, \tau, \alpha, \nu)$ by a multinormal distribution:~~

$(s, h, \tau, \alpha, \nu)$ is thus defined by a multivariate normal distribution:

$$P(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{5}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-(\boldsymbol{x}-\boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})/2}, \tag{21}$$

where ~~$\boldsymbol{x} = \{\lambda, \tau, H, \alpha, \nu\}$~~ $\boldsymbol{x} = \{s, \tau, h, \alpha, \nu\}$, the vector of the means is $\boldsymbol{\mu}$ and the covariance matrix ~~that will be used for future projections~~ $\boldsymbol{\Sigma}$~~; it takes into account potentially large correlations between the parameters~~.

3 Results

~~In this section, using~~ Using Bayes' theorem as described above, we derive probability density functions (PDFs) for the model and forcing parameters of the FEBE from the mean likelihood functions of the five observational datasets. ~~We treat the~~ The different observational datasets are treated as dependent due to the use of overlapping raw data, with the differences between series coming partly from the different processing of the raw data by different teams. This corresponds to putting the datasets into a Bayesian framework where each has equal a priori probability: HadCRUTv4, C&W, GISTEMP, NOAAGlobalTemp and BEST ~~.~~

$(n=5)$.

$$Pr(\cancel{\lambda}s,\cancel{H}h,\tau,\alpha,\nu|T(t)) = \frac{1}{n}\sum_{i=1}^{n} Pr(\cancel{\lambda}s,\cancel{H}h,\tau,\alpha,\nu|T_i(t)) \tag{22}$$

Following IPCC methodologies, we report the "very likely" confidence interval at the 90% confidence level throughout this work along with median estimates for the all ensemble spreads. The complete suite of model and forcing parameters and climate sensitivities are summarized in tables 1 and 2. In addition we include a comparison of the same parameters for the Half-order EBE (HEBE) ~~$(H=1/2)$~~ $(h=1/2)$ that is a consequence of the continuum heat equation ~~((Lovejoy, 2021a), (Lovejoy, 2021b)),~~ (Lovejoy, 2021a, b), as well as with the precursor Scaling Climate Response Function (SCRF) model (~~(Hébert et al., 2021)~~ Hébert et al. (2021)) which differs primarily in the treatment of high frequencies ~~, shown~~ in table 3.

3.1 The Model: Green's Function Parameters: ~~H,τ~~ h,τ

3.1.1 The Scaling Exponent ~~H~~ h

The model is characterized by ~~H~~ h and τ, where ~~H~~ the exponent h of the FEBE is the most fundamental. For ~~H~~ h, we found a 90% CI of $[0.33, 0.44]$, with a median value of 0.38 when using $F_{Aer_{RCP}}$, and while using $F_{Aer_{SSP}}$ we found a similar median of 0.38 with 90% CI of [0.32,0.44]. We can already note that it is close to the HEBE value ~~$H=\frac{1}{2}$~~ $h=\frac{1}{2}$ and other empirical estimates for power law impulse Green's functions $(G(t)\approx t^{-H_F-1})$ with ~~$H=-H_F=0.5$ ((Hébert et al., 2021), (Lovejoy et al., 2017)~~ $h=-H_F\approx 0.5^{-0.4}_{+0.5}$ (Lovejoy et al., 2017; Hébert et al., 2021). The NOAA dataset differs the most from all others, the exact cause of the difference is not clear although it arises from the MLOST dataset's use of a complex frequency algorithm with low-frequency tuning (Smith et al., 2008). This low-frequency tuning along with the spatio-temporal smoothing applied in the MLOST dataset is likely the cause of a slightly higher h (i.e. a smoother temperature response).

3.1.2 The Relaxation Time τ

The second model parameter is the relaxation time ~~scale~~ τ that characterizes the approach to equilibrium. τ is a difficult parameter to determine since from the point of view of parameter estimation, it is inversely correlated with ~~λ~~ s: a large τ can be somewhat compensated by a smaller ~~λ~~ s and vice versa. ~~Since the one box (EBE) model is the special $H=1$ case of the FEBE, we can use box model estimates of τ for our prior distribution. For this purpose, we used the (Geoffroy et al., 2013)~~

~~For each observational dataset and their average, PDFs are shown for the model parameters: the scaling parameter H (left), and the transition time τ (right). Shown are the PDFs for parameter estimation based on both $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{SSP}}$ (dashed). The average PDFs of the five observation datasets using $F_{Aer_{RCP}}$ is shown as the main result with shading, with darker 5% tails.~~

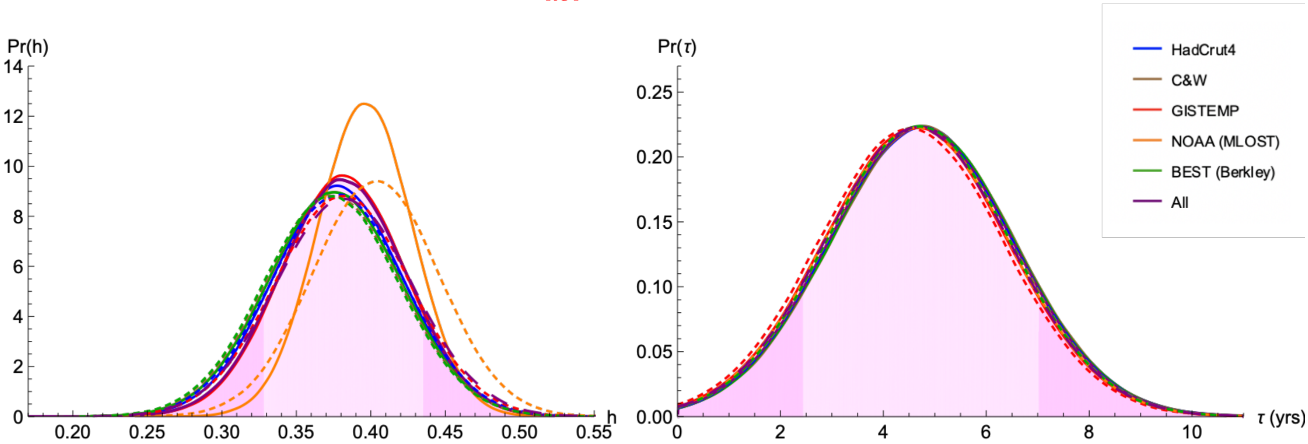

Figure 3. For each observational dataset and their average, PDFs are shown for the model parameters: the scaling parameter h (left), and the transition time τ (right). Shown are the PDFs for parameter estimation based on both $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{SSP}}$ (dashed). The average PDFs of the five observation datasets using $F_{Aer_{RCP}}$ is shown as the main result with shading, with darker 5% tails.

~~two-box relaxation times found by fitting a dozen GCM outputs: 4.1 ± 1.1 years (the fast box τ). This prior is close to other box model fast relaxation times: $\tau = 8.5 \pm 2.5$ years(Schwartz, 2008), $\tau \approx 4$ years (Held et al., 2010), $\tau = 4.3 \pm 0.6$ years (Rypdal and Rypdal, 2014). So as not to be very constraining, we doubled the (Geoffroy et al., 2013) parameter spreads, choosing an N4,2prior distribution of τ. With this prior, we obtained the a posteriori median value of 4.7 years and~~ We obtained the a posteriori median value of 4.7 years and 90% CI of [~~2.4,~~ 2.4,7.0] ~~years when using~~ years when using $F_{Aer_{RCP}}$, and nearly identical results using $F_{Aer_{SSP}}$.

Presented in figure 4 (top) are the step-response Green's function, $G_1(h,\tau;t)$, of the FEBE with the parameters h and τ along with its 90% CI, shown alongside the IPCC two-box model Green's function (IPCC, 2013; Held et al., 2010; Geoffroy et al., 2013) . Considering $G_1(t)$ (blue), at scales below a few years where the box models or the Hébert et al. (2021) truncated scaling model are smooth, the FEBE has a singular response. This enables it to accurately reproduce the statistics of the internal variability as well as to be more sensitive to volcanic forcings. Even up to scales of 25 years, the $G_1(t)$ (blue) responds much faster than the IPCC (black), yet the approach to the asymptotic value 1 corresponding to energy balance is substantially slower. This can also be seen in the ramp-response Green's functions, $G_2(t)$ (bottom). For comparison, each was normalized by the value at 70 years - the standard ramp time for TCR Collins et al. (2013). At multi-year resolution (ignoring the high frequency variability), over the scale of the anthropocene there is little difference between the FEBE and IPCC, with FEBE having a more gradual response. This contributes to the somewhat cooler FEBE centennial scale projections when compared with those from the two-box model.

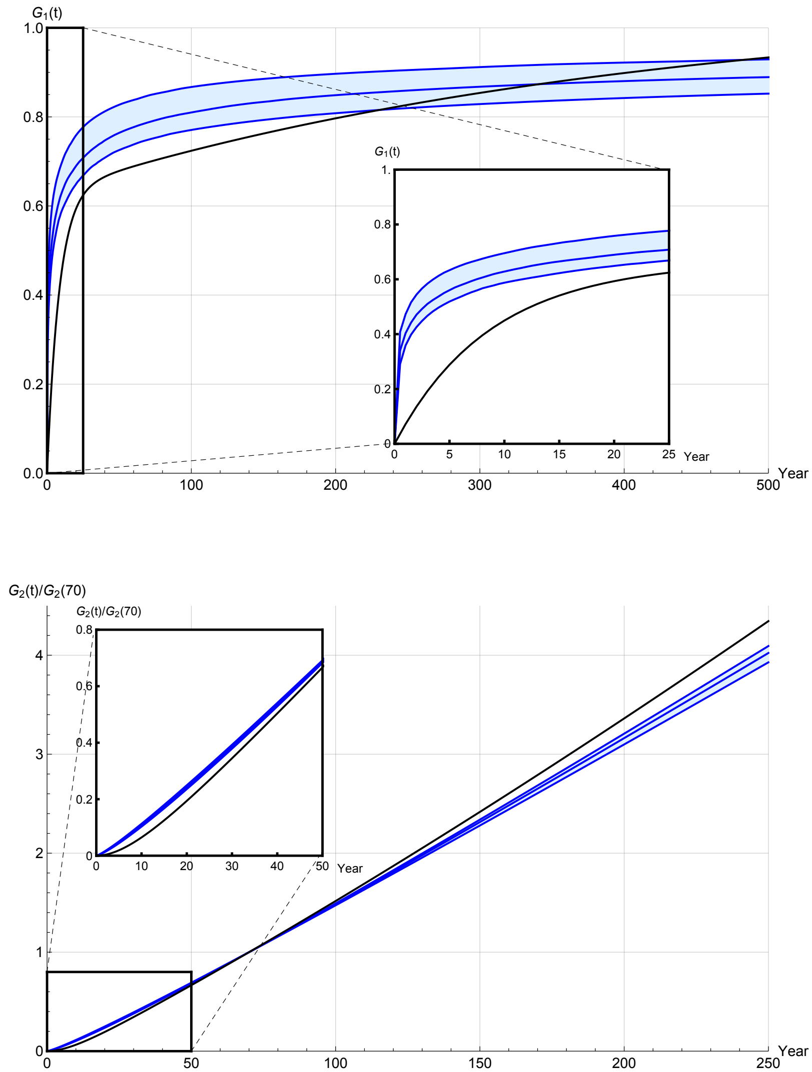

Figure 4. (top) The median and 90% CI of the FEBE step-response Green's function, $G_1(h, \tau; t)$, compared to the IPCC two-box model Green's function (black). (bottom) The median and 90% CI of the FEBE normalized ramp-response Green's function, $G_2(h, \tau; t)$, compared to the IPCC two-box model Green's function (black).

3.2 Characterizing the Forcing

3.2.1 Aerosol Linear Scaling Factor α

The aerosol linear scaling factor α that effectively re-calibrates the aerosol forcing (~~fig.~~ figure 5 left, solid line) was ~~given a prior normal distribution, N(1.00,0.55) which has a 90% CI coherent with the modern (2005) value for $F_{Aer} \approx -1.0Wm^{-2}$, and the a posteriori distribution was~~ found to have a median value of 0.6 with a 90% CI of [0.2,1.0] ~~.~~ for the CMIP5 $F_{Aer_{RCP}}$ series. However when using the CMIP6 sulphate emissions based aerosol forcing series, $F_{Aer_{SSP}}$, we find support for a weaker and better constrained aerosol forcing, recalibration α with a median of 0.33 and 90% CI of [0.05, 0.61] (~~fig.~~ figure 5 left, dashed line). In both cases an aerosol recalibration factor of 1 corresponds to the modern (2005) aerosol forcing value of about $-1.0Wm^{-2}$, but we find both cases that $\alpha < 1$. The result from two independent aerosol forcing series again shows that the forcing associated with aerosols is still widely uncertain and overpowered, supporting post-AR5 studies that found aerosol forcings simulated by GCMs were unrealistic ~~((Zhou and Penner, 2017), (Sato et al., 2018), (Bellouin et al., 2020)),~~ (Zhou and Penner, 2017; Sato et al., 2018; Bellouin et al., 2020), and that aerosol forcing was weaker when climate feedbacks were allowed (Nazarenko et al., 2017).

3.2.2 Volcanic Intermittency Correction Exponent ν

The volcanic intermittency correction exponent ν ~~, given an uninformative uniform prior, $\nu \in [0,1]$ (where $\nu = 0$ implies a constant mean forcing and the original series if $\nu = 1$),~~ was found to have a posterior median value of 0.28 with 90% CI of [0.15,0.41] when using $F_{Aer_{RCP}}$ and similar median value 0.28 with 90% CI of [0.16,0.40] when using $F_{Aer_{SSP}}$ (recall $\nu = 0$ implies a constant mean forcing and the original series is recovered with $\nu = 1$). Both contain the theoretically calculated ν within their 90% CI ($\nu = 0.32$). This result confirms that volcanic forcing is generally overpowered since $\nu = 1$ has nearly null probability as seen in figure 5. Thus, the original volcanic series described without the intermittency correction does not reproduce well, within the FEBE model presented, the cooling events observed in instrumental records following eruptions: ~~the~~ volcanic cooling would be overestimated. As noted in the case for the exponent, h, the NOAA dataset noticeably differs from the others, the spatio-temporal smoothing applied in the MLOST dataset is likely the cause of a lower ν (i.e. a smoother volcanic forcing).

In ~~fig. 6 (left)~~ figure 6 we compare the total forcing series, $F_{Tot}(t)$ (black), IPCC AR5, eq. (16) where $\alpha = \nu = 1$, with the adjusted forcing series, $F_{Tot}(\alpha, \nu; t)$ (blue). During the historical period, the intermittency and strength of the strong volcanic events is greatly reduced and in the recent past the median adjusted forcing series is higher than the unadjusted forcing due to the reduced aerosol forcing strength. This adjusted forcing series consequently contributes to a lower climate sensitivity, presented in the following section, due to the historic negative forcings of volcanoes and aerosols being adjusted to closer match historical observations, eliminating the need for a high climate sensitivity to compensate.

~~Presented in fig. 4 (top) are the step-response Green's function, $G_1(H, \tau; t)$, of the FEBE with the parameters H and τ along with its'~~

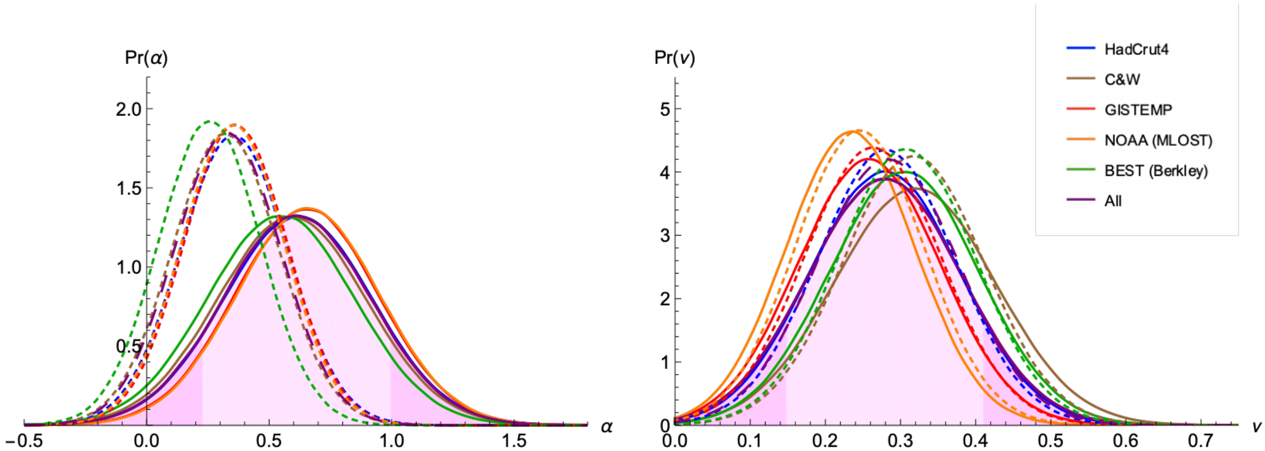

Figure 5. For each observational dataset and their average, PDFs are shown for the forcing parameters: the aerosol scaling factor α (left), and the volcanic intermittency correction exponent ν (right). Again, shown are the PDFs for parameter estimation based on both $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{SSP}}$ (dashed). The average PDFs of the five observation datasets using $F_{Aer_{RCP}}$ is shown as the main result with shading, with darker 5% tails

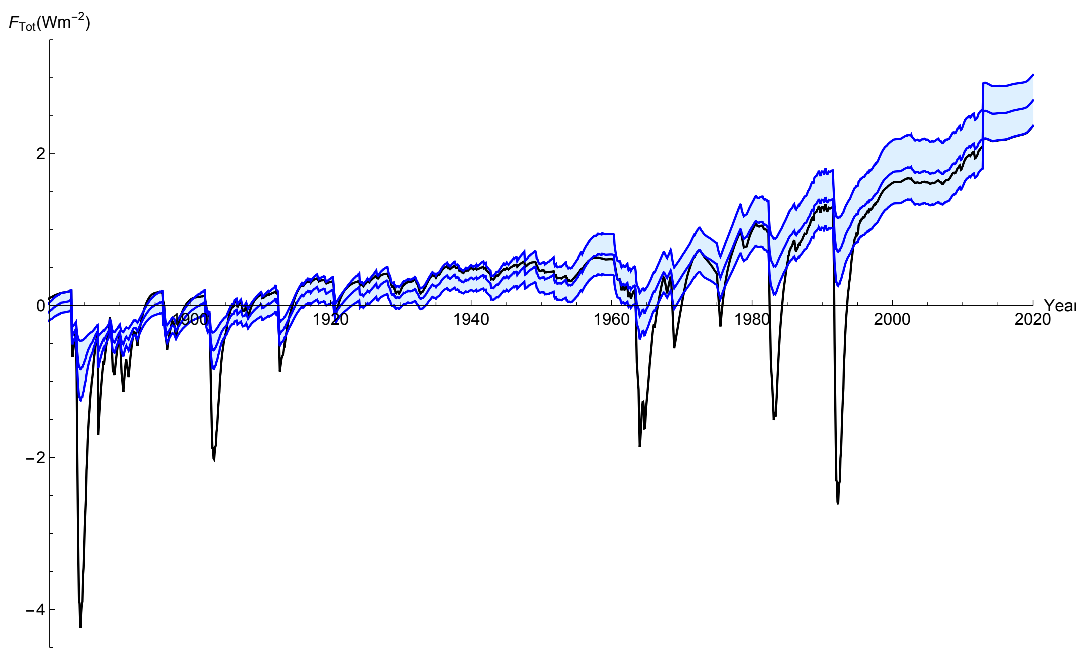

Figure 6. The total historic (1880–2020) forcing series prescribed by the IPCC using, $F_{Aer_{RCP}}$ (black) compared to the adjusted forcing, $F_{Tot}(\alpha, \nu; t)$ (blue) which takes into account aerosol and volcanic corrections, shown ~~over the historical period (1880-2019)~~ with 90% CI.

3.3 Climate Sensitivity

3.3.1 Climate Sensitvity Parameter, s

The climate sensitivity parameter s, refers to the equilibrium change in the annual GMST following a unit change in radiative forcing. Its inverse is the climate feedback parameter, the increase in radiation to space per unit of global warming. We find s to have a median value of 0.56 $K(Wm^{-2})^{-1}$ with 90% CI [0.45,~~shown alongside the IPCC two-box model Green's function (IPCC, 2013). Considering $G_{1,H}$ (blue), at scales below a few years (τ) whereas the box models or the (Hébert et al., 2021) truncated scaling model are smooth, the FEBE has a singular response. This enables it to accurately reproduce the statistics of the internal variability as well as to be more sensitive to volcanic forcings. Even up to scales of 25 years, the $G_{1,H}$ (blue) responds much faster than the IPCC (black), yet the approach to the asymptotic value 1 corresponding to thermodynamic equilibrium is substantially slower. This can also be seen in the ramp-response Green's functions (bottom). For comparison, each was normalized by the value at 70 years - the standard ramp time for TCR. At multi-year resolution (ignoring the volcanic and internal variability), over the scale of the anthropocene there is little difference between the FEBE and IPCC, with FEBE having a more gradual response. This contributes to the somewhat cooler FEBE centennial scale projections when compared with those from the two-box model.~~

~~(top) The median and~~ 0.67] $K(Wm^{-2})^{-1}$ using $F_{Aer_{RCP}}$, and when using $F_{Aer_{SSP}}$ we find median 0.52$K(Wm^{-2})^{-1}$ with 90% CI ~~of the FEBE step-response Green's function, $G_1(H,\tau;t)$, compared to the IPCC two-box model Green's function (black)~~[0.43,0.61] $K(Wm^{-2})^{-1}$ (figure 7). ~~(bottom) The~~ Both on the lower end of the CMIP5 MME climate sensitivity parameter of median 1$K(Wm^{-2})^{-1}$ and 90% CI ~~of~~ [0.5, 1.5]$K(Wm^{-2})^{-1}$ but within the 90% CI. Although both estimates are below the CMIP6 MME 90% CI [0.63, 1.50]$K(Wm^{-2})^{-1}$, with a median of 0.92$K(Wm^{-2})^{-1}$ which has been criticized as being too high (Zelinka et al., 2020; Tokarska et al., 2020; Flynn and Mauritsen, 2020).

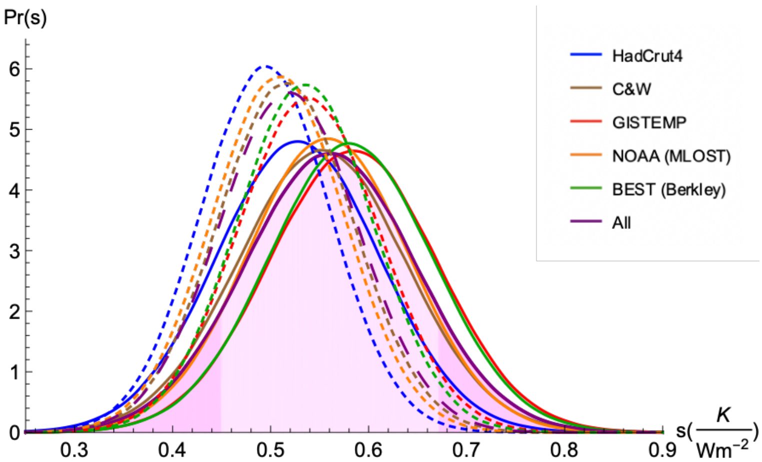

Figure 7. For each observational dataset and their average, PDFs are shown for the ~~FEBE normalized ramp-response Green's function~~climate sensitivity parameter s (the ECS, ~~$G_2(H,\tau;t)$~~here in units of $K(Wm^{-2})^{-1}$), ~~compared to the IPCC two-box model Green's function~~ $F_{Aer_{RCP}}$ (~~black~~solid) and $F_{Aer_{SSP}}$ (dashed).

3.4 ~~Climate Sensitivity~~

3.3.1 Equilibrium Climate Sensitivity

~~Climate sensitivity is a key measure used in climate modelling.~~ Two standard types of climate sensitivity ~~are~~ used for inter-model comparisons: Equilibrium Climate Sensitivity (ECS) and Transient Climate Response (TCR) ~~. Climate sensitivity describes the amount of warming in the atmosphere associated with increases in atmospheric carbon dioxide (CO_2).~~ our results are summarized in table 2.

If atmospheric CO_2 ~~were~~ was increased to double pre-industrial concentrations and then held there, the planet would only slowly reach a new thermodynamic equilibrium. This delay is largely because the world's oceans take a long time to heat up in response to the enhanced greenhouse effect. The Equilibrium Climate Sensitivity (ECS) is the amount of warming achieved when the entire climate system reaches 'equilibrium' or the ~~stable~~ steady-state temperature response to a doubling of CO_2~~, it is equal to λ when measured in units of $K/(CO_2$ doubling)~~. By the definition of the temperature response to external forcings in eq. 7, the climate sensitivity parameter is the equilibrium climate sensitivity. The two are equivalent to within a constant factor: the number of Wm^{-2} per CO_2 doubling, the standard value being $3.71Wm^{-2}/(CO_2$ doubling) (IPCC, 2013).

The PDF for ECS shown in figure 8 (left), for both aerosols series was found to have a 90% CI of $[1.6, 2.4]K$ and a median value of $2.0K$ when using $F_{Aer_{RCP}}$, and median of $1.8K$ and 90% CI [1.5,2.2] using $F_{Aer_{SSP}}$ (see table 2). These results are lower than those found in the CMIP5 MME which had a best value of $3.2K$, but our 90% CI bounds are more narrow, laying within the CMIP5 MME range of $[1.9, 4.5]K$. Although when we consider the expanded ECS 90% CI of $[1.5, 4.5]K$ considered in ~~(IPCC, 2013)~~IPCC (2013), which takes into account both the CMIP5 MME and historical estimates, we see that the FEBE estimates are wholly within this range and much less uncertain. For the CMIP6 MME which has a 90% CI of $[2.0, 5.5]K$ and mean estimate $3.7K$, our best estimate using the corresponding $F_{Aer_{SSP}}$ is slightly below the lower confidence due to the upward shift of ECS estimates seen in CMIP6 models (Zelinka et al., 2020), but again has a more narrow CI.

~~The climate sensitivity parameter, λ, refers to the equilibrium change in the annual GMST following a unit change in radiative forcing. By the definition of the temperature response to external forcings in eq. 7, the climate sensitivity parameter is the equilibrium climate sensitivity. The two are equivalent to within a constant factor: the number of Wm^{-2} per CO_2 doubling, the standard value being $3.71Wm^{-2}/(CO_2$ doubling) IPCC (2013). Its inverse is the climate feedback parameter, the increase in radiation to space per unit of global warming. We find λ to have a median value of 0.56 $K(Wm^{-2})^{-1}$ with 90% CI 0.45,0.67$K(Wm^{-2})^{-1}$ using $F_{Aer_{RCP}}$ (fig. 7). When using $F_{Aer_{SSP}}$ we find median 0.52$K(Wm^{-2})^{-1}$ and 90% CI 0.43,0.61$K(Wm^{-2})^{-1}$. Both on the lower end of the CMIP5 MME climate sensitivity parameter of $1 \pm 0.5K(Wm^{-2})^{-1}$ but within the 90% CI although both estimates are below the CMIP6 MME 90% CI 0.63, 1.50$K(Wm^{-2})^{-1}$ which has been criticized as being too high ((Zelinka et al., 2020), (Tokarska et al., 2020), (Flynn and Mauritsen, 2020)).~~

~~For each observational dataset and their average, PDFs are shown for the climate sensitivity parameter λ (the ECS, here in units of $K(Wm^{-2})^{-1}$), $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{SSP}}$ (dashed)~~

3.3.2 Transient Climate Response

Conventionally, TCR quantifies the ~~changes~~ temperature change that would occur if CO_2 levels increase by 1% (compounded) per year until they double ($\approx$ 70 years). Since the CO_2 forcing is logarithmically dependent on CO_2 concentration, the TCR is then simply the global temperature increase that has occurred at the point in time that a linearly increasing forcing reaches double pre-industrial levels.

The derived ~~PDF for TCR is shown in fig.~~ PDFs for TCR are shown in figure 8 (middle) and summarized in table 2. Our TCR was found to have a 90% CI of $[1.2, 1.8]K$ with a median of $1.5K$ when using $F_{Aer_{RCP}}$, while when using $F_{Aer_{SSP}}$ we find a median 1.4K and 90% CI of [1.1,1.6]K. Both estimates are lower and more constrained, but within the 90% CI given by the CMIP5 MME: a 90% CI of $[1.2, 2.4]K$ and a best value of $1.8K$, and by the CMIP6 MME: 90% CI of $[1.2, 2.8]K$ with best value of $2.0K$.

The ECS and TCR estimates using the SSP scenarios with the FEBE are lower than those using RCP due to the overly strong aerosols over the historical period in the SSPs which require a lower aerosol linear factor along with lower ECS to best match the historical temperature record. The difference between the shape of the RCP and SSP aerosol forcing can also account for this.

The TCR-to-ECS ratio is a non-dimensional measure of the fraction of committed warming already realised after a steady increase in radiative forcing, in this case a doubling of CO_2, this quantity is generally referred to as realised warming fraction (RWF) ~~((Stouffer, 2004), (Solomon et al., 2009), (Millar et al., 2015))~~(Stouffer, 2004; Solomon et al., 2009; Millar et al., 2015) . A model with a low RWF will indicate that global warming may continue for centuries after emissions have stopped. We present the TCR-to-ECS ratio in ~~fig.~~ figure 8 (right), having a 90% CI [0.70, 0.78] and median 0.73 using ~~$F_{Aer_{RCP}}$~~F_{RCP} parameters. Similar results are found using ~~$F_{Aer_{SSP}}$~~F_{SSP} parameters, a median of 0.72 and 90% CI [0.71, 0.79]. ~~The~~ From figure 8 and table 2, we see that the TCR-to-ECS ratio is higher than both generations of MME 90% CI, a consequence of lower ECS and TCR values, and similar uncertainty.

In the next section we show that with a lower and more constrained climate sensitivity parameter (figs. 7 and 8), the adjusted ~~forcing~~ forcings (figure 6) and long memory process of the FEBE ~~(fig. 6) produces~~ produce future projections that tend to be cooler than the CMIP5/6 projections, yet remain within their 90% CI.

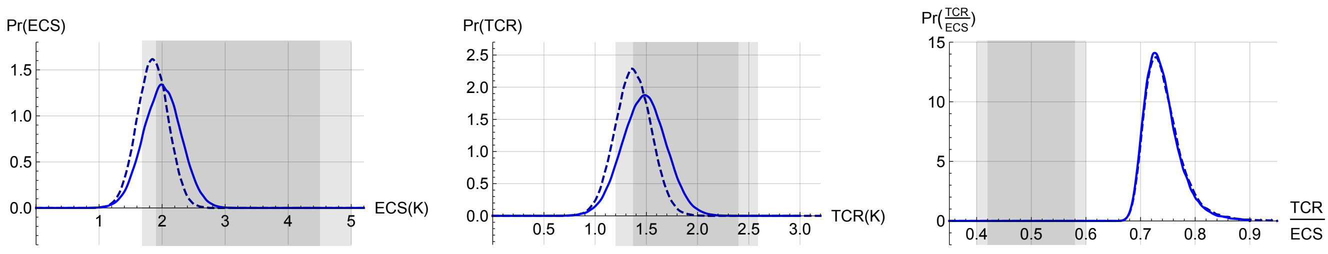

Figure 8. The PDFs for ECS (left), TCR (middle) and the TCR:ECS ratio (right) are derived using $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{SSP}}$ (dashed). The associated 90% CI (bars under the axis), the CMIP5 MME 90% CI (dark gray shading), and the CMIP6 MME 90% CI (light gray shading).

Table 1. Model and Forcing parameter medians for FEBE calibrated over the historical period (~~1880-2019~~1880–2020) using $F_{Aer_{RCP}}$ and $F_{Aer_{SSP}}$, along with their corresponding 90% confidence intervals.

	Median ~~H~~h	~~H~~ h 90% CI Range	Median τ [years]	τ 90% CI Range [years]	Median α	α 90% CI Range	Median ν	ν 90% CI Range	Median ~~λ~~ s $[\frac{K}{Wm^{-2}}]$	~~λ~~ s 90% CI Range $[\frac{K}{Wm^{-2}}]$
$F_{Aer_{RCP}}$	0.38	[0.33, 0.44]	4.7	[2.4, 7.0]	0.6	[0.2, 1.0]	0.28	[0.15, 0.41]	0.56	[0.45, 0.67]
$F_{Aer_{SSP}}$	0.38	[0.32, 0.44]	4.7	[2.4, 7.0]	0.33	[0.05, 0.61]	0.28	[0.16, 0.40]	0.52	[0.43, 0.61]

Table 2. The calculated ECS and TCR medians using both parameters corresponding to $F_{Aer_{RCP}}$ and $F_{Aer_{SSP}}$, along with their corresponding 90% confidence intervals.

	Median TCR [K]	TCR 90% CI Range [K]	Median ECS [K]	ECS 90% CI Range [K]	Median TCR/ECS Ratio	TCR/ECS Ratio 90% CI Range
$F_{Aer_{RCP}}$	1.5	[1.2, 1.8]	2.0	[1.6, 2.4]	0.73	[0.70, 0.78]
$F_{Aer_{SSP}}$	1.4	[1.1, 1.6]	1.8	[1.5, 2.2]	0.74	[0.71, 0.79]

Table 3. Model and Forcing parameter medians using $F_{Aer_{RCP}}$ for FEBE, the phenomological HEBE ~~$(H = \frac{1}{2})$~~ $(h = \frac{1}{2})$ and the ~~Sealing CRF~~ SCRF model (Hébert et al., 2021) calibrated over the historical period, along with their corresponding 90% confidence intervals.

	Median ~~H~~h	~~H~~h 90% CI Range	Median τ (years)	τ 90% CI Range (years)	Median α	α 90% CI Range	Median ν	ν 90% CI Range	Median ECS ~~λ~~$[K]$	ECS ~~λ~~90% CI Range $[K]$
FEBE	0.38	[0.33, 0.44]	4.7	[2.4, 7.0]	0.6	[0.2, 1.0]	0.28	[0.15, 0.41]	2.0	[1.6, 2.4]
HEBE	1/2	-	4.7	[2.4, 7.0]	0.48	[0.10, 0.86]	0.33	[0.16, 0.51]	1.8	[1.4, 2.3]
SCRF	0.5	[0.3, 0.7]	2.0	-	0.8	[0.1, 1.3]	0.55	[0.25, 0.85]	2.3	[1.8, 3.7]

4 Projections

4.1 ~~Hindprojections: 1880-2019~~

~~With the full suite~~ With the collection of model and forcing parameters, ~~we now use the FEBE~~ the FEBE was used to reconstruct the ~~forced temperature response~~ temperature over the historical period~~and make projections~~, as well as make projections of the forced temperature response for the coming century using forcings prescribed by the RCP and SSP scenarios. ~~In this section we present comparisons of the FEBE observation-based projections with those from the GCMs in the CMIP5 MME, and all the currently available CMIP6 GCMs analyzed by (Forster et al., 2020).~~

The CI provided for the MME corresponds to the spread between the different GCMs~~.~~, "structural uncertainty", while for the FEBE it is parameter uncertainty (Bretherton, 2012). In both cases, the projections are deterministic but with uncertainty limits due to their respective model uncertainties. Both yield an estimate of the forced response but with qualitatively different uncertainty bounds.

~~From the multidimensional parameter space we draw from the multinormal distribution (~~For the FEBE, the spread of the forced projections is purely from the uncertainty in the parameters: the contribution to uncertainty from internal variability has been averaged out (it is effectively the average over an infinite ensemble of realizations of internal variability). In order to make projections we therefore draw samples of parameters from the (correlated) multidimensional parameter space (approximated by the multivariate normal distribution in eq. 21), by using a Monte Carlo method~~to generate~~. Once a random set of parameters has been chosen, realizations of the forced temperature response ~~(~~are generated using eq. 3 ~~) using~~ and a numerical convolution. It should be noted this Monte Carlo sampling is simply a convenient numerical technique for performing high dimensional probability space integrals, it does not imply any stochasticity in the projections which although are parametrically uncertain, nevertheless have purely deterministic forcing. However, the Monte Carlo methods do introduce standard Monte Carlo numerical uncertainty, but this was made quite small by using a large number (500) of Monte Carlo realizations. Once we have our ensemble of projections, we remove the pre-industrial baseline (~~1880-1910~~such that the temperature anomaly over

1880–1910 is zero) and calculate the desired confidence intervals of the forced response~~, the historical reconstructed period~~. We consider the historical period coinciding with the range of observation temperature records (1880-2020) and make all comparisons to this period, acknowledging that the CMIP5 GCMs historical reconstruction ended in 2005 and for the CMIP6 GCMs in 2014.

4.1 Reliability and Historical Reconstructions: 1880–2020

In this section we present the full historical reconstruction using the FEBE observation-based projections with those from the GCMs in the CMIP5/6 MME. In order to make a proper comparison with data we must include both the forced deterministic temperature response, with its purely parametric uncertainty, as well as the internal variability of the mean observational temperature series, estimated to be $\approx \pm 0.14$°C (monthly resolution). The two uncertainties were combined assuming the statistical independence of the internal forcing and the parametric uncertainty: the errors therefore add in quadrature.

An important characteristic of probabilistic forecasts is their reliability that quantifies the difference between the forecast and actual probability distributions. For example, consider a set of predictions derived from ensemble forecasts, in some realizations it is predicted that the chance of above-average seasonal-mean temperature for the coming season will be 70%. If the probabilistic forecast system is reliable, then one can expect that in 70% of these predictions the actual seasonal-mean temperature will be above average (Annan and Hargreaves, 2010; Weisheimer and Palmer, 2014). In figure 9, we can verify the reliability of the FEBE. We see that as expected, the temperature observations fall closely within the 90% CI of the FEBE historical reconstruction (i.e. the ensemble average of the response to both internal and external forcing). More precisely, at the monthly resolution in figure 9, the historical mean temperature (red) is within the 90% CI of the FEBE forced response (with internal variability added) 89.9% of the months using the RCP scenario (left) or 90.2% of the months using the SSP scenario (right). The accuracy of this uncertainty verifies both the underlying model and Bayesian parameter estimation method.

This is expected for a reliable model and is an analogous validation of probabilistic aspects of the projection as unlike weather forecasts where we have many past test cases, climate change projections cannot be calibrated in the same manner (Stainforth et al., 2007; Tebaldi and Knutti, 2007; Knutti et al., 2010). In both reconstructions it is possible that the end of war (1945) temperature spike which lies out of the FEBE 90% CI may be explained due to biases associated with bucket and engine room intake measurements (Chan and Huybers, 2021).

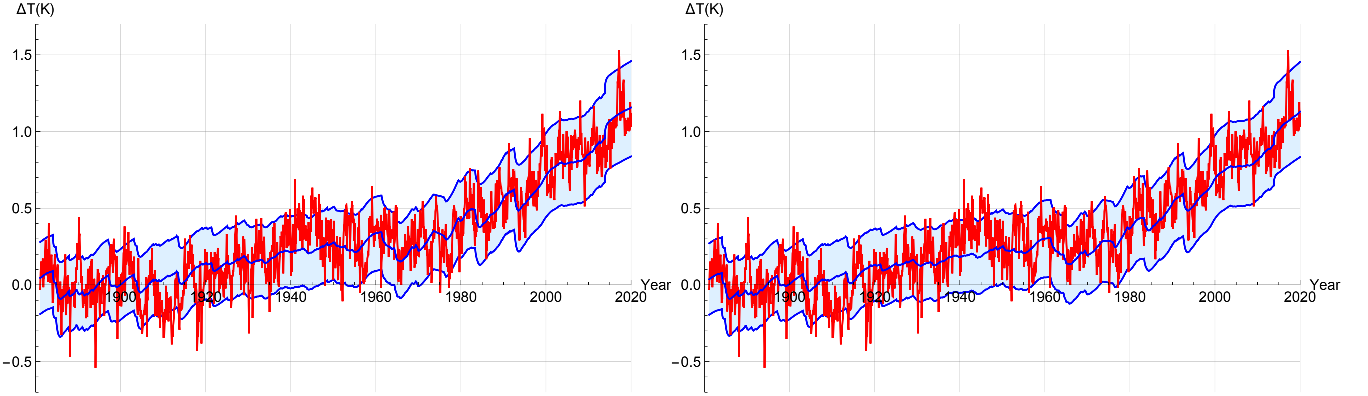

Figure 9. (left) The historical reconstruction (forced temperature response and internal variability) of the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue) alongside mean of 5 observational temperature series (red) at monthly resolution; 90% CI (due to parametric uncertainty and internal variability) are indicated (shaded). (right) Same as left except using $F_{Aer_{SSP}}$ parameters and forcing.

4.2 The Amplitude of the Internal Forcing

The small scale limit of the validity of the FEBE is not known, although it is likely to be ∼1 month (roughly the weather-macroweather transition time scale). Justification comes from the success of the high frequency FEBE limit that successfully forecasts monthly and seasonal temperatures (Del Rio Amador and Lovejoy, 2019, 2021a, b). As discussed earlier (eq. 14) the FEBE predicts the (stochastic) response to the internal forcing. The standard deviation of $f(t)$ is the amplitude of the internal forcing assumed to be a Gaussian white noise, which can be estimated using eqs. 14, and 15, and $\sigma_{T,\tau_r} \approx \pm 0.14$°C. Using our F_{RCP} (and F_{SSP}) parameter estimates, we find a mean estimate of the forcing standard deviation, σ_{f,τ_r}, to be 3.2 Wm^{-2} (~~hindprojections) presented in fig. 9~~ 3.3 Wm^{-2}) and 90% CI of [2, 1, 4.2] Wm^{-2} ([2.3, 4.3] Wm^{-2}) (at a monthly resolution). If we introduce a white noise forcing with σ_{f,τ_r} amplitude the FEBE recreates the amplitude of the internal temperature variability response.

This estimate of the internal variability forcing can be compared with that of Harries and Belotti (2010) who examine the net energy flux balance at the top of atmosphere (TOA) measured using observations from polar-orbiting spacecraft (at monthly scale). The early observations, using the Nimbus experiments, show an internal variability of the $4.1 \pm 4.0 Wm^{-2}$, while more modern measurements (CERES) in the 2000s show variability of between ± 2 and ~~projections to 2100 in fig. 10~~± 4 Wm^{-2} generally laying a few Wm^{-2} of zero. Thus our estimate of the internal forcing variability is within estimates of the TOA net energy flux balance.

4.3 Statistical Evaluation of the FEBE

It was shown in Lovejoy et al. (2021) that the FEBE roughly predicts both high- and low-frequency scaling regimes, using a simple "ramp" model that included the deterministic external and stochastic internal variabilities. In figure 10 (left) we show one realization of the full FEBE, including the deterministic and stochastic forcings, with median parameters calibrated earlier using $F_{Aer_{RCP}}$ (blue) and $F_{Aer_{SSP}}$ (light blue); the five observation temperature series are shown alongside (gray -

shifted up). We compare the model statistics with the 5 globally averaged temperature series using their root-mean-square Haar fluctuations, shown in figure 10 (right). The Haar fluctuation for a series $T(t)$, $\Delta T(\Delta t)$ is the difference between the average of the first and second halves of the interval Δt. This is a convenient way to characterize variability as a function of time scale in real space, valid for increasing or decreasing average fluctuations. By applying global scale Haar fluctuation analyses Del Rio Amador and Lovejoy (2019) found $H \approx -0.1$ corresponding to $h = H + 1/2 \approx 0.4$.

~~Throughout 1880-2005 (~~Below Milankovitch time scales, there are three main scaling regimes observed in the atmosphere: the weather, macroweather and climate Lovejoy (2013). In the macroweather regime, longer than the lifetime of planetary structures (~ 10 days), temperature fluctuations decrease with scale until a transition probably occurs to the climate regime where fluctuations begin to increase. In the industrial epoch this scale is ~ 20 years, while in the ~~overlapping observational and~~ pre-industrial epoch this scales transition occurs at centuries or millennia Lovejoy (2015b). Over the scale of 1 year to about 10 years (the macroweather regime), the FEBE and the observational temperature series have an approximate slope (indicated by the straight reference line in figure 10) of $h \approx 0.4$. We see a transition in both the FEBE and observations at $\Delta t \gtrsim 10$ years: the transition to the climate regime where fluctuations begin to increase with scale. The fact that the FEBE's fluctuations at the climate regime track the observational data well gives confidence in using the FEBE for multidecadal projections.

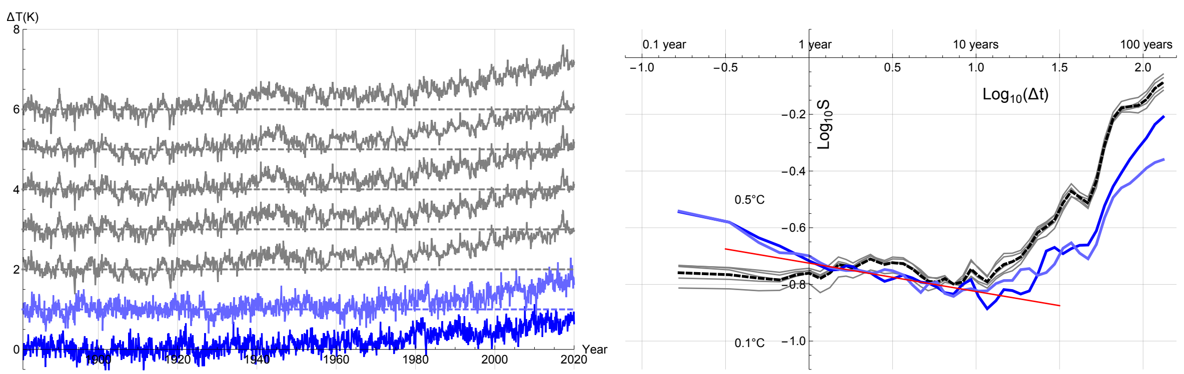

Figure 10. (left) The historical reconstruction (forced and internal temperature response) of the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue) and $F_{Aer_{SSP}}$ (light blue) alongside the 5 observational temperature series (gray - shifted up) at monthly resolution. (right) The Root Mean Square Haar fluctuation structure function $S(\Delta t) = \langle \Delta T(\Delta t)^2 \rangle^{\frac{1}{2}}$ for FEBE reconstruction using $F_{Aer_{RCP}}$ (blue) and $F_{Aer_{SSP}}$ (light blue), and the five globally averaged monthly-resolution temperature time series (gray; mean is shown in dashed black). The reference (red) line has the slope of the approximate median estimate of the scaling exponent $h \approx 0.4$ ($H = h - \frac{1}{2} \approx -\frac{1}{10}$).

4.4 Evaluating the FEBE using Hindprojections Including the Slowdown

We have shown that the FEBE hindcasts are reliable (sec 4.1), that they have realistic internal forcings (sec. 4.2) and realistic statistical variability (sec. 4.3). Here we evaluate their deterministic responses using hindprojections.

In figure 11, we compare the 90% CI of the historical temperature observations with the median forced response of both the FEBE using the RCP (left) and SSP (right) historical forcing compared to both the CMIP5 ~~MME period) , the reconstructed~~

~~forced temperature response from the FEBE, using $F_{Aer_{RCP}}$, and the mean of the~~ (left) and CMIP6 (right) MMEs. In the inset of figure 11, we show the slowdown ("hiatus') period (1998-2014).

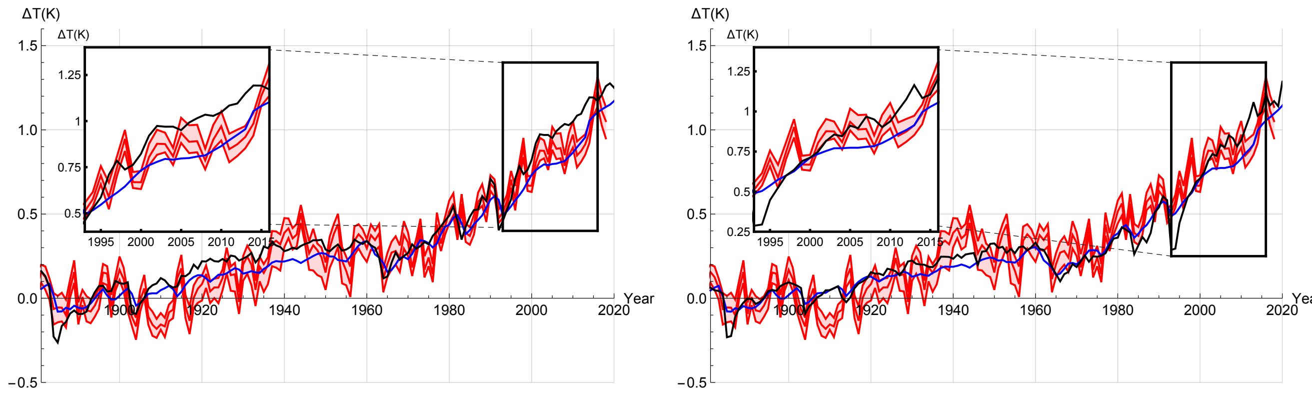

Figure 11. (left) The median historical forced component of the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue), and the median of the CMIP5 MME (black) alongside mean of 5 observational temperature series (red) with their 90% CI indicated (shaded). (right) The median historical forced component of the FEBE, with parameters calibrated using $F_{Aer_{SSP}}$ (blue), and the median of the CMIP6 MME (black) alongside mean of 5 observational temperature series (red) with the 90% CI indicated (shaded).

Throughout the historical period, the hindprojection of the FEBE and the median of the CMIP5~~MME are close, seen in fig. 9). The spread between the forced temperature response and~~ /6 MME are close. Between 1915–1960 the ~~five observational datasets decreases as we move closer to the present. Between 1915-1960 periodthe~~ CMIP5/6 MME is consistently warmer than the ~~historical reconstruction~~ FEBE hindprojection and historical temperature records, although generally by less than 0.05K. The slowdown in global warming during the first decade of the 21st century, termed as ~~'the pause' or 'the hiatus' ((Kaufmann et al., 2011), (Meehl et al., 2011), (Medhaug et al., 2017))~~ the slowdown ("hiatus") (Kaufmann et al., 2011; Meehl et al., 2011 , is tracked closely by the FEBE ~~reconstruction~~ hindprojection while the CMIP5/6 MME overshoots (by $0.1K$ to $0.2K$), a well studied divergence between GCMs and observations~~(fig. 9 lower plot). This supports (Lovejoy, 2015b) who~~ , shown in figure 11 (insets).

Following the monthly resolution reliability confirmation in section 4.1 we can now perform a quantitative comparison between the amount of time the FEBE and CMIP5/6 MME median response is within the bounds of the observational temperature series 90% CI performed with annual resolution data. The median FEBE hindprojection using $F_{Aer_{RCP}}$ is within the 90% CI of the observational temperature series over the whole historic period 47% of the years and over the slowdown is within 70% in comparison to the CMIP5 MME median which is within the whole historic period only 39% and over the slowdown 17%. While the median FEBE hindprojection using $F_{Aer_{SSP}}$ similar results are found, over the whole period: 45% and over the slowdown 35%, in comparison to the CMIP6 MME median which is within the whole historic period 39% and over the slowdown is 30%. In can be seen in both cases that the CMIP MME is generally warmer than the FEBE forced component notably over the period of the slowdown. We see that indeed, the FEBE median forced component in both cases captures the slowdown rather accurately. This supports (Lovejoy, 2015a, b) which found that the ~~hiatus~~ slowdown ("hiatus") could be well predicted by a stochastic fGn model (comparable with the present hindprojection) and concluded ~~that the problem was~~ the issue to be GCM overprojection.

~~The reconstruction of the forced temperature response over the historical period (1880-2019) using the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue), compared with the CMIP5 MME projection (black) and mean of 5 observational temperature series (red); 90% CI are indicated (shaded). The whole historical period shown (top), along with the 'hiatus/pause' where the CMIP5 MME median is consistently above the temperature records while the FEBE median tracks the observations.~~

4.5 Projections through to 2100

~~The~~ We now consider the FEBE projections to 2100. At first, the temperature increase in each case is nearly identical; the future pathways only ~~begin to diverge~~ diverging into their respective scenarios roughly two decades after their beginning (RCPs begin in ~~2000~~2005, SSPs begin in ~~2010), so at first the temperature increase in each case is nearly identical~~2014). Further into the future, the warming rate begins to depend more on the specified scenario, the highest being in RCP 8.5/SSP ~~585 (fig. 10 bottom~~5-85 (figure 12 c, f) while significantly lower in RCP 2.6/SSP ~~126 (fig. 10 top~~1-26 (figure 12 a, d; tables 4, 5), particularly after about 2050 when the global surface temperature response stabilizes (and declines thereafter). Of particular interest are the low emissions scenarios, RCP 2.6/SSP ~~126~~1-26, demonstrating the potential of strong mitigation policies and speculative negative emission technologies where anthropogenic forcing starts decreasing around the mid-2040s. In the CMIP5 MME, the temperature stays below $2K$ throughout the 21st century, whereas the corresponding median FEBE temperature projection never exceeds $1.5K$. ~~In comparison,~~ Comparing projected warming at 2100 for the RCP 2.6/SSP 1-26 scenario, the FEBE projection reaches a median warming of $1.2K$ with 90% CI of $[1.1, 1.4]K$ while the CMIP5 MME has a 90% CI of $[0.9, 2.4]K$ and median warming of $1.7K$. When considering the CMIP6 projections for SSP ~~126 (figure ??, top~~1-26 (figure 12, d) the median temperature exceeds $2K$ beginning near 2050, whereas the corresponding FEBE projection is consistently lower, only crossing the $1.5K$ threshold briefly. At 2100, the CMIP6 projected temperature reaches $2.2K$ with 90% CI of [1.5, 2.8]K while the FEBE projects a median temperature of $1.5K$ and a narrower spread of [1.3, 1.8]K.

While the forcing of the (perhaps most realistic) middle scenario, RCP 4.5/SSP 2-45, stabilizes in the mid 2060s, the temperature projections continue rising throughout the 21st century for both FEBE and the CMIP5~~MME (fig. 10 middle~~/6 MME (figure 12b, e). At 2100 the FEBE and CMIP5 MME ~~projects a~~ project the temperature reaching $1.9K$ [1.6,2.2]K, and $2.6K$ [1.8, 3.2]K respectively shown in figure 12b. A key point to note is that the FEBE RCP 4.5 projection remains below ~~2~~2.5K of warming by 2100, while the CMIP5 MME is well beyond this threshold. ~~At the time of writing the corresponding SSP simulations were not available~~Looking at the CMIP6 projections for SSP 2-45 (figure 12, e) the median temperature exceeds $2K$ beginning near 2050, whereas the corresponding FEBE projection is consistently lower, and begins to diverge after 2050. At 2100, the CMIP6 projected temperature reaches $3K$ with 90% CI of [12.1, 4.2]K while the FEBE projects a median temperature of $2.3K$ and a narrower spread of [1.8, 2.8]K.

The projections of both the FEBE and the CMIP5 MME for the ~~business as usual~~ strong emission scenario, RCP 8.5, show alarming warming rates of $3.5K$ with 90% CI $[2.9,4.1]K$, and $4.8K$ with 90% CI $[3.5, 6.0]K$ in 2100 shown in figure ~~10 (bottom)~~12c, f. The same quickly increasing trend is seen in the CMIP6 SSP ~~585~~ 5-85 scenario with temperatures in 2100 reaching a staggering $6.2K$ with 90% CI $[4.5, 7.0]K$, while the FEBE projection although lower at $3.8K$ and having a tighter bound of $[3.5, 4.5]K$ shows the dire consequences of no mitigation. All results shown in figure 12 are summarized in tables 4 and 5.

The FEBE projections are slightly different depending on the forcing series used, either the RCPs or the SSPs, almost solely because of different aerosol forcing while the other ~~forcing~~ forcings are practically unchanged. By applying the aerosol recalibration factor we constrain both aerosol series over the past to produce the same relative forcing thus there is very little difference over the historical ~~hindprojections of FEBE~~ FEBE hindprojections. Moving into the future, in the SSP scenarios the aerosol forcing is quickly reduced to zero while in the RCP scenarios the aerosols ~~in the future~~ are reduced but not fully eliminated causing the FEBE projections with SSP scenarios to be warmer than their RCP counterparts.

The discrepancy in aerosol forcing strength (~~fig.~~ figure 1) at 2100 between $F_{Aer_{RCP}}$ and $F_{Aer_{SSP}}$ in the strong mitigation scenario (RCP 2.6/SSP ~~126) is $0.6Wm^{-2}$~~1-26) is $\approx 0.6Wm^{-2}$. With the total GHG forcings being nearly identical, the ~~the SSP 126~~ SSP 1-26 scenario is closer to a RCP ~~3.3~~ 3.2 scenario than the RCP 2.6 scenario. Therefore for RCP 2.6, we can roughly attribute 30% of the increased future warming shown in CMIP6 in comparison with CMIP5 (~~figs. 10, ??~~figure 12) to the strong reduction of future aerosol emissions. For the MME, in RCP 2.6, this corresponds to about $0.5K$ in 2100. The corresponding values for RCP 8.5 is 7% and $0.6K$ in 2100.

Although the FEBE projections are consistently about 15% cooler than the CMIP5 MME, ~~the~~ due to the its smaller uncertainty the FEBE 90% CI lies entirely within the corresponding CMIP5 CI. ~~Being in complete agreement, the two projections methods mutually~~ Both projection methods support each other and are thus complementary. When compared to CMIP6 projections although most of 90% CIs overlap, the median CMIP6 temperatures are ~~by~~ nearly 65% warmer than the corresponding FEBE median~~. The main reason being~~ , mainly caused by their overpowered aerosols ~~((Zelinka et al., 2020) , (Flynn and Mauritsen, 2020))~~ (Zelinka et al., 2020; Flynn and Mauritsen, 2020) and previously mentioned ~~removal of future aerosol forcing~~discrepancy in the future aerosol removal as compared to the RCPs.

~~The forced temperature response projected using the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue) compared with the CMIP5 MME projection (black); 90% CI are indicated (shaded). The projections until 2100, for RCP 2.6 (top), RCP 4.5 (middle) and RCP 8.5 (bottom), are shown.~~

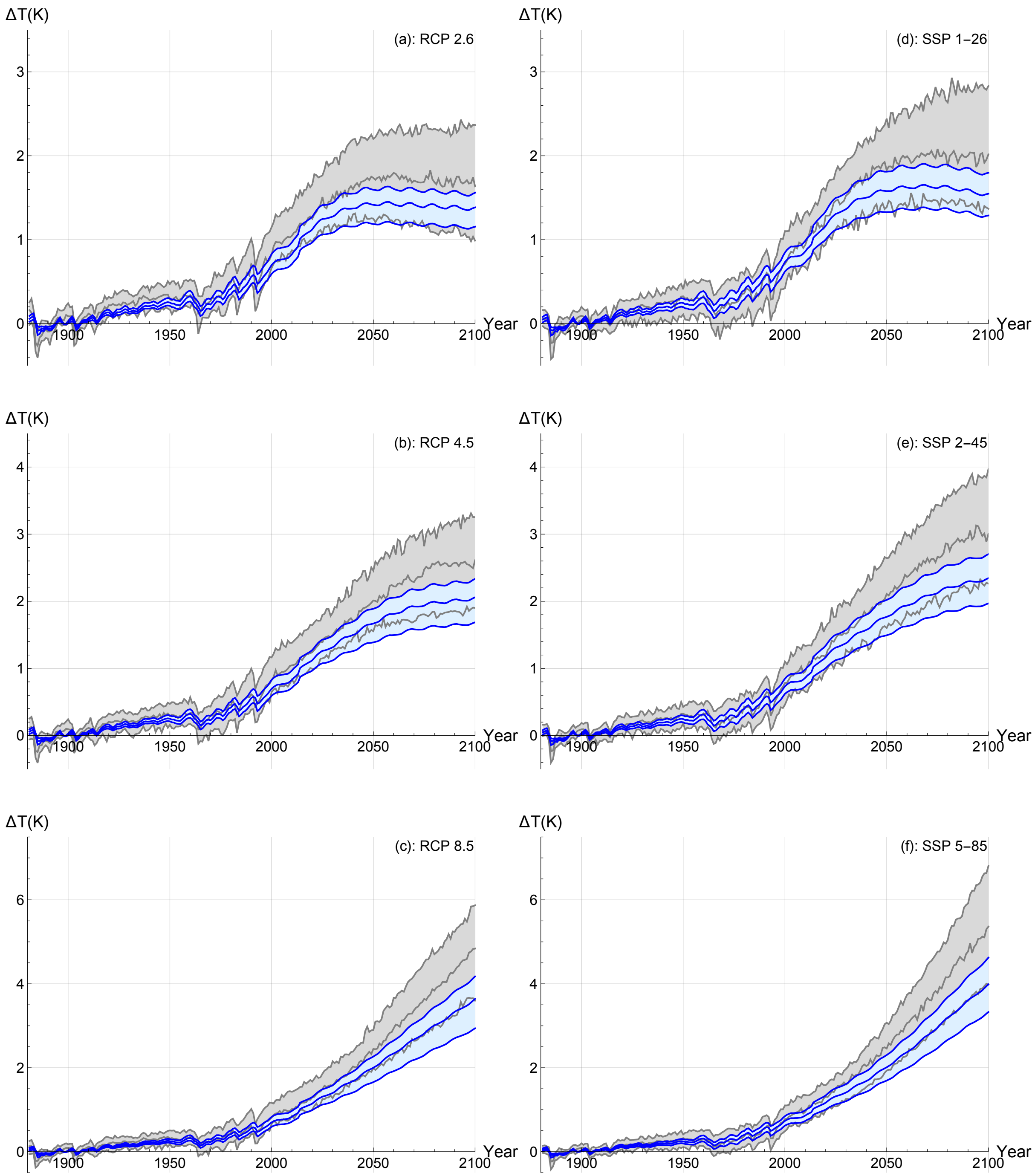

Figure 12. The deterministic forced temperature response projected using the FEBE (blue), with parameters calibrated using ~~$F_{Aer_{SSP}}$~~ $F_{Aer_{RCP}}$ (~~blue~~a, b, c) and $F_{Aer_{SSP}}$ (d, e, f) compared with the ~~CMIP6~~ CMIP5/6 MME projection (black); 90% CI from the parametric uncertainty are indicated (shaded). The projections until 2100, for RCP 2.6/SSP ~~126~~ 1-26 (top), RCP 4.5/SSP 2-45 (middle) and RCP 8.5/SSP ~~585~~ 5-85 (bottom), are shown.

Table 4. The 90% CI of projected warming relative to pre-industrial reference period (1880 –1910) for the RCP scenarios analysed in this study based on the FEBE and the CMIP5 MME. Summary of figure 12 (a, b, c).

	RCP 2.6		RCP 4.5		RCP 8.5	
	FEBE	**CMIP5**	**FEBE**	**CMIP5**	**FEBE**	**CMIP5**
2020-2040	[1.1, 1.5]K	[1.2, 1.9]K	[1.1, 1.5]K	[1.3, 1.9]K	[1.2, 1.6]K	[1.4, 2.0]K
2040-2060	[1.2, 1.6]K	[1.3, 2.2]K	[1.4, 1.9]K	[1.6, 2.6]K	[1.7, 2.3]K	[2.0, 3.0]K
2060-2080	[1.2, 1.6]K	[1.2, 2.3]K	[1.6, 2.2]K	[1.8, 3.0]K	[2.2, 3.1]K	[2.6, 4.3]K
2080-2100	[1.2, 1.6]K	[1.1, 2.4]K	[1.6, 2.3]K	[1.8, 3.2]K	[2.7, 3.8]K	[3.3, 5.3]K

Table 5. The 90% CI of projected warming relative to pre-industrial reference period (1880 –1910) for the SSP scenarios analysed in this study based on the FEBE and the CMIP6 MME. Summary of figure 12 (d, e, f).

	SSP 1-26		SSP 2-45		SSP 5-85	
	FEBE	**CMIP6**	**FEBE**	**CMIP6**	**FEBE**	**CMIP6**
2020-2040	[1.2, 1.6]K	[1.2, 1.9]K	[1.2, 1.6]K	[1.2, 2.0]K	[1.2, 1.7]K	[1.2, 2.0]K
2040-2060	[1.3, 1.8]K	[1.4, 2.3]K	[1.5, 2.0]K	[1.6, 2.7]K	[1.7, 2.3]K	[1.9, 3.0]K
2060-2080	[1.4, 1.9]K	[1.5, 2.6]K	[1.8, 2.4]K	[1.9, 3.2]K	[2.3, 3.2]K	[2.8, 4.4]K
2080-2100	[1.3, 1.8]K	[1.4, 2.8]K	[1.9, 2.6]K	[2.2, 3.8]K	[3.0, 4.2]K	[3.6, 6.0]K

~~((Schurer et al., 2017), (Smith et al., 2018b), (Iseri et al., 2018)) have given evidence that there are important tipping points~~

4.6 Probabilities of Exceeding Critical Warming Thresholds

We can also use the FEBE to estimate the probability of exceeding various warming thresholds. Important tipping points have been established which could lead to irreversible changes in major ecosystems and the planetary climate if certain threshold in warming are exceeded ~~. Therefore~~ (Schurer et al., 2017; Smith et al., 2018b; Iseri et al., 2018) . Using the FEBE and CMIP5/6 MME we calculate the probability of temperature exceeding $1.5K$ and $2.0K$ ~~for both the FEBE and the CMIP5/6 MME shown in figs. 11 and 12.~~ (figure 13).

According to the FEBE for the low emission scenario RCP 2.6, it is unlikely to exceed the $1.5K$ threshold in 2100 ($< 10\%$) while it is much more likely to exceed this threshold according to CMIP5 MME (~~probability of~~ 67%). The FEBE has a negligible probability of exceeding $2K$ while the CMIP5 MME has a 26% ~~chance~~probability. While in the SSP ~~126~~ 1-26 scenario, the FEBE peaks at below 50% of exceeding $1.5K$ and has a negligible probability of exceeding $2K$ as before; it is nearly inevitable to cross $1.5K$ threshold according to the CMIP6 MME, while the $2K$ threshold being exceeded hovers around 60% even under strong mitigation.

In the RCP 4.5 scenario, the probability of the FEBE exceeding the $1.5K$ threshold is extremely likely ($> 95\%$) ~~and~~ although it occurs much later than that projected by the CMIP5 MME: occurring 20 years after the ~~GCMs~~MMEs: 2070~~.~~; for the SSP

2-45 scenario we see the FEBE trail the CMIP6 MME until around 2035, when exceeding $1.5K$ becomes very likely for both near 2045. Similarly, the $2K$ overshoot, as projected by the FEBE will be avoided with a probability of $< 40\%$ but will most likely not be avoided according to the CMIP5 MME with an 89% probability of exceeding $2K$. Again we see the FEBE lag behind the CMIP6 MME, the probability of exceeding $2K$ is lower at the same time as compared to the CMIP6 MME, before they begin to converge around 2080, approaching a very likely probability to exceed the threshold.

For the final high emission, business as usual, RCP 8.5 and SSP ~~585~~ 5-85 scenarios; both the FEBE and CMIP5/6 MME project that exceeding the $1.5K$ threshold is virtually inevitable by 2100. Although in the FEBE projection it is extremely likely that this threshold is exceeded nearly 15 years after the CMIP5/6 MME projections of 2040. The same is found for the $2K$ threshold, with both the FEBE and CMIP5/6 MME exceeding the threshold about 15 years after the $1.5K$ threshold. These results are all summarized in table 6.

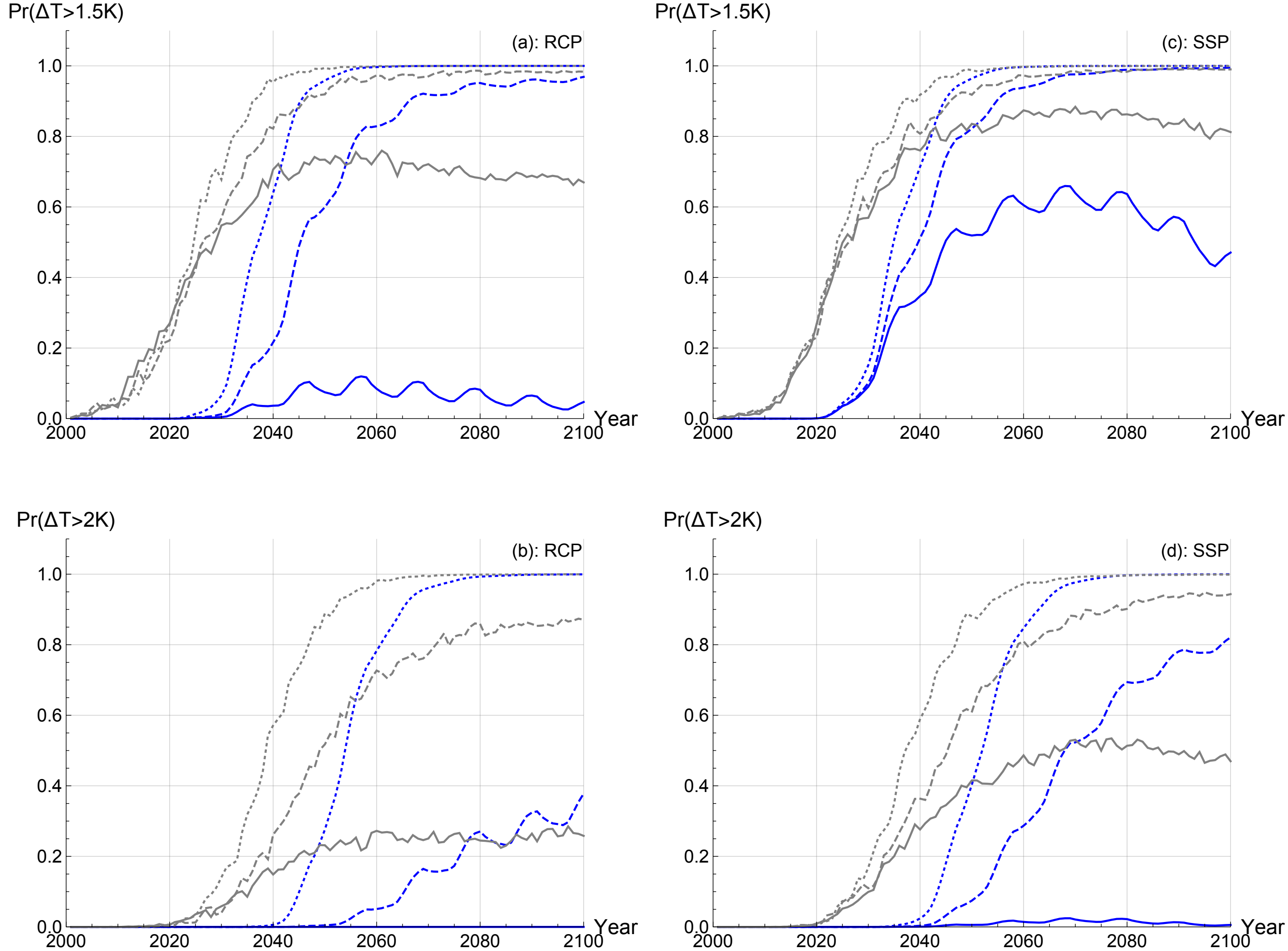

Figure 13. The probability for the global mean surface temperature of exceeding a $1.5K$ threshold (top: a,c), and a $2K$ (bottom - b,c) are given as a function of years for the FEBE (blue), using $F_{Aer_{RCP}}$ (~~blue~~a,b) or $F_{Aer_{SSP}}$ (c, d) and for the CMIP5/6 MME (black). The three ~~RCP~~ scenarios are considered for each case: RCP 2.6/SSP 1-26 (solid), RCP 4.5/SSP 2-45 (dashed), and RCP 8.5/SSP 5-85 (circles).

5 Discussion

In the following section we summarize the key results presented earlier in the paper: model and forcing parameters (see table 1), the climate sensitivities (see table 2), the projected warming at 2100 (see tables 4 and 5) and the probabilities of exceeding warming thresholds of $1.5K$ ~~threshold (top),~~ and ~~a 2~~$2.5K$ ~~(bottom) are given as a function of years~~ (see table 6).

The two parameters that characterize the model, h and τ, were estimated. The fundamental scaling exponent h, was found to have a median value of 0.38 and 90% CI of $[0.33, 0.44]$ using $F_{Aer_{RCP}}$, and similar median value 0.38 and 90%

Table 6. ~~The probability for~~ List of RCP and SSP scenarios analysed in this study and the ~~global mean surface temperature~~ probabilities of exceeding ~~a~~ 1.5°C or 2°C based on the FEBE and the CMIP5/6 MME. Summary of figure 13.

	Probability of Exceeding 1.5°C at 2100			Probability of Exceeding 2°C at 2100		
	FEBE	**CMIP5**	**CMIP6**	**FEBE**	**CMIP5**	**CMIP6**
RCP 2.6	0.1%	67.0%	-	0.0%	25.8%	-
RCP 4.5	96.9%	98.4%	-	37.8%	87.1%	-
RCP 8.5	100%	100%	-	100%	100%	-
SSP 1-26	47.1%	-	81.2%	0.0%	-	47.0%
SSP 2-45	99.5%	-	99.0%	82.1%	-	94.3%
SSP 5-85	100%	-	100%	100%	-	99.9%

CI of $[0.32, 0.44]$ for ~~the FEBE~~ $F_{Aer_{SSP}}$. Both estimates are near h estimated for the Scaling Climate Response Function (Hébert et al., 2021) and the phenomenological HEBE (Lovejoy, 2021a, b). The relaxation time scale τ, characterizing the approach to equilibrium was found to have median value of 4.7 years and 90% CI of [2.4,7.0] years when using $F_{Aer_{RCP}}$, and nearly identical results using $F_{Aer_{SSP}}$. The estimated relaxation time is comparable to other box model fast relaxation times (Schwartz, 2008; Held et al., 2010; Geoffroy et al., 2013; Rypdal and Rypdal, 2014) as the one box (EBE) model is a special case of the FEBE where $h = 1$.

The FEBE model also adjusts the deterministic forcings, notably the aerosol and volcanic forcing series which must be scaled (the former linearly and the latter non-linearly) as for the temperature response to best match historical temperature records. From our analysis we find a more constrained aerosol forcing. For the $F_{Aer_{RCP}}$ we found a median recalibration factor α of 0.6 with 90% CI [0.2,1.0]. Following Stevens (2015) this supports a revision of the global modern (2005) aerosol forcing 90% CI to a narrower range [-1.0,-0.2]Wm^{-2}. Using the CMIP6 aerosols, $F_{Aer_{SSP}}$, we found a median α value of 0.33 and 90% [0.05, 0.61] implying a weaker and more tightly constrained modern (2005) aerosol forcing of [-0.9,-0.1]Wm^{-2}. For volcanism, the non-linear intermittency exponent, ν, was found have median value of 0.28 with 90% CI of [0.15, 0.41] using $F_{Aer_{RCP}}$ and a median 0.28 with similar 90% CI of [0.16, 0.40] using $F_{Aer_{SSP}}$.

In comparison to IPCC AR5 and to the CMIP6 MME, we find lower likely ranges for the climate sensitivity parameter, ECS and TCR when using the FEBE with $F_{Aer_{RCP}}$ (or $F_{Aer_{SSP}}$). For projections, perhaps the most important parameter is s, climate sensitivity parameter which determines the temperature response following an increase in forcing. We find s to have a median value of 0.56 $K(Wm^{-2})^{-1}$ with 90% CI [0.45,0.67] $K(Wm^{-2})^{-1}$ using $F_{Aer_{RCP}}$, and when using $F_{Aer_{SSP}}$ we find median 0.52$K(Wm^{-2})^{-1}$ with 90% CI [0.43,0.61] $K(Wm^{-2})^{-1}$. Again see a lower median for the ECS in comparison to the IPCC AR5 (and CMIP6 MME) estimates for their corresponding forcings, the 90% CI range is reduced from [1.5, 4.5]K (~~blue)~~[2.0, 5.5]K) to [1.6, 2.4]K ([1.5, 2.2]K) and the median value is lowered from 3.0K (3.7K) to 2.0K (1.8K). Several recent observation-based studies (Otto et al., 2013; Skeie et al., 2014; Johansson et al., 2015; Lewis and Curry, 2015, 2018) have also reported lower ECS upper bounds. We also estimated the derived quantity, the Transient Climate Response (TCR), the temperature increase following a linear doubling in forcing over 70 years. For the TCR, where the 90% CI range shrinks from $[1.0, 2.5]K$ ($[1.2, 2.8]K$) to $[1.2, 1.8]K$ ($[1.1, 1.6]K$) and the median estimate decreases from $1.8K$ ($2.0K$) to $1.5K$ ($1.4K$).

With all necessary parameters of the FEBE calibrated on observational temperature series we show that the FEBE is a reliable model able to reconstruct the historical temperatures (see section 4.1), can reproduce the response amplitude to a internal white noise forcing (see section 4.2), and produces realistic temperature fluctuations over a wide range of scales (see section 4.3). Having shown that the FEBE may reproduce historical temperatures and their statistics, we then produce deterministic temperature projections to 2100 using the RCPs: 2.6, 4.5, 8.5 and SSPs: 1-26, 2-45, 5-85 comparing them to their respective CMIP5 and ~~for the~~ CMIP6 ~~MME~~ MMEs relative to the pre-industrial baseline of 1880–1910. In the low emission scenario RCP 2.6 (SSP 1-26) the FEBE projects the 90% CI of the temperature at 2100 to be $[1.2, 1.6]$K ($[1.3, 1.8]$K) as compared to the CMIP5 (~~black). The two~~ CMIP6) MME of $[1.1, 2.4]$K ($[1.4, 2.8]$K). The middle scenario, RCP 4.5 (SSP 2-45) the FEBE projects warming reaching $[1.6, 2.3]$K ($[1.9, 2.6]$K), narrower than the CMIP5 (CMIP6) MME warming of $[1.8, 3.2]$K ($[2.2, 3.8]$K). While in the high emission scenario, RCP 8.5 (SSP 5-85) both the FEBE and CMIP5 (CMIP6) MME project extreme temperature increases of $[2.7, 3.6]$K ($[3.0, 4.2]$K) and $[3.3, 5.3]$K ($[3.6, 6.0]$K), highlighting the need for strong emission mitigation.

During the Paris Conference in 2015, nations of the world strengthened the United Nations Framework Convention on Climate Change by agreeing to holding the increase in the global average temperature to well below $2°C$ above pre-industrial levels and pursuing efforts to limit the temperature increase to $1.5°C$. According to our projections, crossing either of these thresholds is delayed with respect to the CMIP5/6 MME projections but will eventually happen if strong mitigation is not implemented. To avert a $1.5K$ warming, drastic cuts would have to be made to global greenhouse emissions, similar to that in RCP 2.6 (and SSP 1-26), for which we found <10% (<50%) probability of exceeding $1.5K$ in comparison to the CMIP5 (CMIP6) MME which projects a 67% (>80%). Both the FEBE and CMIP5/6 projections have temperatures surpassing $1.5K$ in scenarios with weak or no mitigation: RCP 4.5/SSP ~~scenarios are considered for each case:~~ 2-45, and RCP 8.5/SSP ~~126 (solid),~~ 5-85, albeit the FEBE projects this occurring nearly two decades later than the GCMs. The $2K$ threshold is projected to be avoided by both the FEBE and CMIP5/6 MME if we follow low emission scenarios of RCP 2.6 and SSP 1-26. The opposite is true for any other emission scenarios; the exceeding of the $2K$ threshold in this coming century would surely occur. Thus our model reinforces that only strong mitigation scenarios, such as RCP 2.6 and SSP 1-26, will avoid excursions over these $1.5K$ and ~~SSP 585 (circles).~~ $2K$ thresholds, although it remains to be seen whether negative emission technologies are feasible and whether the appropriate policies are implemented.

6 Conclusions

~~Multidecadal projections have~~ Ever since the first climate models at the end of the 1970's, multidecadal projections have had large uncertainties with the wide ECS confidence limits of $1.5\text{-}4.5K$ essentially unchanged~~for decades~~. For policy makers, the most deleterious consequence of large uncertainties is that projections emanating from quite diverse future scenarios have significant overlap. For example, up until 2050, the RCP 2.6 and 8.5 scenarios can both claim to respect the $2K$ threshold - albeit with rather different probabilities (figure 13). Large overlaps imply a disconnection between policies (mitigation scenarios) and outcomes (temperatures). Now that governments have committed themselves to keeping industrial epoch temperature increases to below $2K$ (and aim at $1.5K$), we face an uncertainty crisis (Lovejoy, 2019b).

One way of reducing this uncertainty is by developing complementary types of models. In this paper we directly constructed such a model in the macroweather regime (roughly one month and up) based on the physically principles of energy conservation and scaling: the Fractional Energy Balance Equation (FEBE). Although originally derived phenomenologically, it was recently discovered ~~((Lovejoy, 2021a), (Lovejoy, 2021b))~~ (Lovejoy, 2021a, b) that the FEBE could be derived as a consequence of classical ~~(Budyko, 1969) and (Sellers, 1969)~~(Budyko, 1969; Sellers, 1969), Energy Balance Models (EBMs) that have been regularly used to determine the Earth's latitudinal temperature variations, its stability to perturbations and to study past and future climate states. The key was to introduce a vertical coordinate that allows for the application of the correct conductive-radiative surface boundary condition, needed for correctly determining the energy storage. A surprising consequence is that even the classical (integer ordered) continuum mechanics heat equation used by Budyko and Sellers implies that the surface temperature obeys a fractional ordered energy balance equation. The FEBE's fractional storage terms imply that the system has a long memory so that when calibrated by observational data, its responses to past forcings are constrained to respect the historical climate.

The FEBE is a parsimonious model with only two shape parameters: an exponent $\cancel{H}h$, and relaxation time scale τ; the classical EBE ~~is the~~ $\cancel{H=1}$ (box model) is the $h=1$ special case. In order to make FEBE projections, we use a Bayesian parameter estimation approach ~~. (Hébert et al., 2021)~~ similar to that used in Hébert et al. (2021). The latter used a Climate Response Function that at long times ($>\tau$) was close to the corresponding FEBE Green's function, but that was a different power law at shorter time scales. While the ~~(Hébert et al., 2021)~~ Hébert et al. (2021) CRF was justified on the basis of scale invariance, the FEBE has a stronger physical basis since it respects both scaling as well as energy conservation, and the long memory is explicitly situated in energy storage mechanisms. From the practical (projection) point of view, the main advantage is that the FEBE directly handles the short time scales (down to a month or less). This allows the FEBE to directly take into account the internal variability: a stochastic white noise forcing and the FEBE response. The ability to model the forced response to both external and internal forcing improves FEBE parameter estimates and contributes to lowering the corresponding projection uncertainties.

Bayesian inference allows for a robust probabilistic parameter characterization. The basic external forcings were those prescribed for the historical part of the CMIP5/6 GCMs and these were constrained by five monthly, global resolution empirical temperature series (since 1880). The internal forcing was assumed to be a Gaussian white noise and, since to a good approximation, the FEBE white noise response is a fractional Gaussian noise (fGn), the latter was taken as the Bayesian inference error model.

In order to estimate the parameters, the forcing series required two adjustments. The most important was the aerosol recalibration parameter α which linearly scales the aerosol forcing to take into account the increasing evidence that the CMIP5 and CMIP6 aerosol cooling was too strong ~~((Padilla et al., 2011), (Hébert et al., 2021), (Zelinka et al., 2020), (Tokarska et al., 2020)~~

~~, (Flynn and Mauritsen, 2020))~~(Padilla et al., 2011; Hébert et al., 2021; Zelinka et al., 2020; Tokarska et al., 2020; Flynn and Mauritsen, 20 . The former aerosol series ($F_{Aer_{RCP}}$) was based both on uncertain data but also on uncertain modelling assumptions, especially about the direct and indirect effects of aerosols. Whereas the latter ($F_{Aer_{SSP}}$) is based on global sulphate production and derived from an alternative model than that in CMIP5.

~~From our analysis we find a more constrained aerosol forcing. For the $F_{Aer_{RCP}}$ we found a median recalibration factor α of 0.6 with 90% CI 0.2,1.0. Like (Stevens, 2015) this supports a revision of the global modern (2005) aerosol forcing 90% CI to a narrower range -1.0,-0.2Wm^{-2}. Using the CMIP6 aerosols, $F_{Aer_{SSP}}$, we found a median α value of 0.33~~ The forcings and ~~90% 0.05, 0.61implying a weaker and more tightly constrained modern (2005) aerosol forcing of -0.9,-0.1Wm^{-2}. For volcanism, the non-linear intermittency exponent, ν, was found have median value of 0.28 with 90% CI of 0.15, 0.41using $F_{Aer_{RCP}}$ and a median 0.28 with similar 90% CI of 0.16, 0.40using $F_{Aer_{SSP}}$.~~

~~With these adjustments to the deterministic part of the forcing, the two parameters that characterize the model, H and τ, were estimated. The most fundamental, H, the scaling parameter was found to have a median value of 0.38 and 90% CI of $[0.33, 0.44]$ using $F_{Aer_{RCP}}$, and similar median value 0.38 and 90% CI of $[0.32, 0.44]$ for $F_{Aer_{SSP}}$. Both estimates are near to the scaling parameter estimated for the SCRF (Hébert et al., 2021) and the phenomenological HEBE ((Lovejoy, 2021a), (Lovejoy, 2021b)). The relaxation time scale τ, characterizing the approach to equilibrium was found to have median value of 4.7 years and 90% CI of 2.4,7.0years when using $F_{Aer_{RCP}}$, and nearly identical results using $F_{Aer_{SSP}}$. The estimated relaxation time is comparable to other box model fast relaxation times ((Geoffroy et al., 2013), (Schwartz, 2008), (Held et al., 2010), (Rypdal and Rypdal, 2014)) as the one box (EBE) model is a special case of the FEBE where $H = 1$.~~

~~In comparison to IPCC AR5 (and to the CMIP6 MME), we find lower likely ranges for ECS and TCR when using the FEBE with $F_{Aer_{RCP}}$ (or $F_{Aer_{SSP}}$). For projections, perhaps the most important parameter is λ, the equilibrium climate sensitivity (ECS) which determines the temperature following an increase in forcing. We see a lower median for the ECS in comparison to the IPCC AR5 (and CMIP6 MME) estimates for their corresponding forcings , the 90% CI range is reduced from 1.5, 4.5K (2.0, 5.5K) to 1.6, 2.4K (1.5, 2.2K) and the median value is lowered from 3.0K (3.7K) to 2.0K (1.8K). We also estimated the derived quantity, the Transient Climate Response (TCR), the temperature increase following a linear doubling in forcing over 70 years. For the TCR, where the 90% CI range shrinks from 1.0,2.5K (1.2, 2.8K) to 1.2,1.8K (1.1, 1.6K) and the median estimate decreases from 1.8K (2.0K) to 1.5K (1.4K). Several recent observation-based studies ((Otto et al., 2013), (Skeie et al., 2014), (Johansson et al., 2015), (Lewis and Curry, 2015), (Lewis and Curry, 2018)) have also reported lower ECS upper bounds.~~

~~The forcings and~~ parameters combined with the RCP and SSP scenarios allow us to make projections through to 2100, we did this for RCP 2.6 (SSP ~~126~~1-26), 4.5 (SSP 2-45), and 8.5 (SSP ~~585~~5-85). Overall, the observational based FEBE projections had uncertainties that are smaller by more than a factor of two in comparison to the CMIP5/6 MME uncertainties. However, the two modelling approaches have quite different sources of uncertainty. Whereas the CMIP5/6 uncertainty is purely due to differences in the climates of the GCMs ("structural uncertainties"), the FEBE uncertainty is "parametric" and it depends largely on the uncertainty of the historical forcings and temperatures, in particular, those associated with aerosols. In fact, a byproduct of the model and Bayesian framework is that we are able to more tightly constrain aerosol forcing, supporting recent literature findings of weaker historical aerosol cooling. As a consequence, the FEBE projections are consistently a little cooler than those of the CMIP5/6 MME but ~~still within the uncertainty bounds of the latter, effectively complementing the GCMs.~~ with uncertainties about half of those of the MME, it still lies within the MME uncertainty bounds. The qualitatively different FEBE thus effectively complements the GCMs.

~~During the Paris Conference in 2015, nations of the world strengthened the United Nations Framework Convention on Climate Change by agreeing to holding "the increase in the global average temperature to well below $2°C$ above pre-industrial levels and pursuing efforts to limit the temperature increase to $1.5°C$". According to our projections, crossing either of these thresholds is delayed with respect to the CMIP5/6 MMEprojections but will eventually be passed if strong mitigation is not implemented. To avert a $1.5K$ warming, drastic cuts would have to be made to global greenhouse emissions, similar to that in RCP 2.6 (and SSP 126), for which we found <10% (<50%) probability of exceeding $1.5K$ in comparison to the CMIP5 (CMIP6) MME which projects a 67% (>80%). Both the FEBE and CMIP5/6 projections have temperatures surpassing $1.5K$ in scenarios with weak or no mitigation: RCP 4.5, RCP 8.5 and SSP 585, albeit the FEBE projects this occurring nearly two decades later than the GCMs. The $2K$ threshold is projected to be avoided by both the FEBE and CMIP5/6 MME if we follow low emission scenarios of RCP 2.6 and SSP 126. The opposite is true for any other emission scenarios; the exceeding of the $2K$ threshold in this coming century would surely occur. Thus our model reinforces that only strong mitigation scenarios, such as RCP 2.6 and SSP 126, will avoid excursions over these $1.5K$ and $2K$ thresholds, although it remains to be seen whether negative emission technologies are feasible and whether the appropriate policies are implemented.~~

In this paper we have presented the FEBE model and projections to 2100 which is an observational-based model physically based on energy and scale symmetries, complementary to GCMs. Future work will explore the full (regional, 2D) FEBE model (Lovejoy, 2021b), which hopefully will constrain and improve future projections. ~~Once the full suite of CMIP6 projections are released, the~~ In (Lovejoy et al., 2021; Lovejoy, 2021a), the FEBE is shown to plausibly reproduce the annual cycle at monthly resolution, in particular to explain the lag between the temperature maximum and the maximum in the radiative forcing. The FEBE could be also used to help understand the generational differences between CMIP models. We can also calibrate the FEBE on the historical runs of the CMIP models in order to perform a feedback analysis to investigate the differences between how models treat their volcanic and aerosol forcings through the parameters ν and α. Updating our parameter estimates from calibrations on GCMs allows for GCM-FEBE hybrid projections. Extensions to precipitation may also be possible at global and regional scales since the FEBE model is consistent with space-time scaling processes in historical precipitation data ~~(de Lima and Lovejoy, 2015).~~ de Lima and Lovejoy (2015). With tighter constraints on ECS and TCR from the FEBE we can better estimate future warming when bringing together multiple lines of evidence such as that done in Sherwood et al. (2020). The FEBE once expanded spatially provides a flexible framework which can be calibrated directly on observations, providing a direct representation of forcing to response relationships.

Data availability. RCP concentrations can be found at https://tntcat.iiasa.ac.at/RcpDb/dsd?Action=htmlpage&page=welcome. SSP radiative forcings are provided at https://doi.org/10.5281/zenodo.3515339. CMIP5/6 model outputs are available at https://esgf-node.llnl.gov.

Appendix A

Table A1. List of CMIP6 Models and model climate parameters.~~* Denotes models used in fig. ??~~

Model	ECS (K)	TCR (K)	TCR-to-ECS ratio
~~MIROC6*~~ MIROC6	2.60	1.58	0.61
~~IPSL-CM6A-LR*~~ IPSL-CM6A-LR	4.50	2.39	0.53
~~CNRM-CM6-1*~~ CNRM-CM6-1	4.82	2.23	0.46
~~BCC-CSM2-MR*~~ BCC-CSM2-MR	3.07	1.60	0.52
~~MRI-ESM2*~~ MRI-ESM2	3.11	1.67	0.54
~~CanESM5*~~ CanESM5	5.58	2.75	0.49
~~CESM2*~~ CESM2	5.15	1.99	0.39
GISS-E2-1-H	2.99	1.81	0.61
GISS-E2-1-G	2.60	1.66	0.64
SAM0-UNICON	3.30	2.08	0.63
E3SM-1-0	5.09	2.91	0.57
~~UKESM1-0-LL*~~ UKESM1-0-LL	5.31	2.79	0.53
~~CNRM-ESM2-1*~~ CNRM-ESM2-1	4.75	1.82	0.38
BCC-ESM1	3.29	1.77	0.54
~~CESM2-WACCM*~~ CESM2-WACCM	4.65	1.92	0.41
MIROC-ES2L	2.66	1.51	0.57
~~EC-EARTH3-VEG*~~ EC-EARTH3-VEG	3.93	2.76	0.70
HADGEM3-GC31-LL	5.46	2.47	0.45
NORCPM-1	2.78	1.55	0.56
~~GFDL-CM4*~~ GFDL-CM4	3.79	-	-
GFDL-ESM4	2.56	-	-
NESM3	4.50	-	-
NORESM2-LM	2.49	1.48	0.59
NORESM2-LM	2.49	1.48	0.59
MPI-ESM1-2-HR	2.84	1.57	0.55
INM-CM4-8	1.81	1.30	0.72
Ensemble Mean $\pm$ Std:	3.74$\pm$1.11	1.98$\pm$0.48	0.55 $\pm$ 0.09

Table A2. List of CMIP5 Models and climate sensitivity parameters.

Model	ECS (K)	TCR (K)	TCR-to-ECS ratio
MPI-ESM-LR	3.48	1.94	0.56
MPI-ESM-MR	3.31	1.93	0.58
MPI-ESM-P	3.31	1.96	0.59
MIROC5	2.70	1.49	0.55
MIROC-ESM	4.68	2.15	0.46
IPSL-CM5B-LR	2.58	1.44	0.56
IPSL-CM5A-MR	4.03	1.96	0.49
IPSL-CM5A-LR	3.97	1.94	0.49
ISM-CM4	2.01	1.22	0.61
CSIRO-Mk3.6.0	4.05	1.76	0.43
CNRM-CM5	3.21	2.04	0.64
CNRM-CM5-2	3.40	1.63	0.48
BNU	3.98	2.58	0.65
BCC-CSM1.1	2.81	1.74	0.62
BCC-CSM1.1(m)	2.77	2.00	0.72
BCC-GCCM3	2.65	1.58	0.60
NORESM1-M	2.75	1.34	0.49
ACCESS1.0	3.76	1.72	0.46
CanESM2	3.71	2.37	0.64
GFDL-ESM2M	2.33	1.23	0.53
GFDL-CM3	3.85	1.85	0.48
CCSM4	2.90	1.64	0.57
FGOALS-g2	3.39	1.42	0.41
GISS-E2-H	2.33	1.69	0.73
GISS-E2-R	2.06	1.41	0.68
HADGEM2-ES	3.96	2.38	0.60
Esemble Mean $\pm$ Std:	3.20$\pm$0.70	1.75$\pm$0.38	0.56 $\pm$ 0.09

Author contributions. SL: conceptualization of the study. RH: design of methods for model calibration. RP: development the model code and prepared the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. S. Lovejoy acknowledges some support from the National Science and Engineering research Council (Canada). R. Hébert has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 716092 and grant agreement no. 772852. We thank D. ~~Clark~~ Clarke and L. Del Rio Amador for helpful discussions, and M. Willard-Stepan for help in editing the manuscript. The work profited from discussions at the CVAS working group of the Past Global Changes (PAGES) programme.

References

Annan, J. D. and Hargreaves, J. C.: Reliability of the CMIP3 ensemble, Geophysical Research Letters, 37, https://doi.org/https://doi.org/10.1029/2009GL041994, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GL041994, 2010.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A.-L., Dufresne, J.-L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, Reviews of Geophysics, 58, e2019RG000 660, https://doi.org/10.1029/2019RG000660, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019RG000660, e2019RG000660 10.1029/2019RG000660, 2020.

Bretherton, S.: A National Strategy for Advancing Climate Modeling, The National Academies Press, Washington, DC, https://doi.org/10.17226/13430, https://www.nap.edu/catalog/13430/a-national-strategy-for-advancing-climate-modeling, 2012.

Budyko, M. I.: The effect of solar radiation variations on the climate of the Earth, tellus, 21, 611–619, 1969.

Chan, D. and Huybers, P.: Correcting Observational Biases in Sea-Surface Temperature Observations Removes Anomalous Warmth during World War II, Journal of Climate, pp. 1 – 44, https://doi.org/10.1175/JCLI-D-20-0907.1, https://journals.ametsoc.org/view/journals/clim/aop/JCLI-D-20-0907.1/JCLI-D-20-0907.1.xml, 2021.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A. J., and Wehner, M.: Long-term climate change: Projections, commitments and irreversibility, pp. 1029–1136, Cambridge University Press, Cambridge, UK, https://doi.org/10.1017/CBO9781107415324.024, 2013.

Cowtan, K. and Way, R.: Update to 'Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends'. Temperature reconstruction by domain: version 2.0 temperature series., https://doi.org/10.13140/RG.2.1.4728.0727, 2014a.

Cowtan, K. and Way, R. G.: Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends, Quarterly Journal of the Royal Meteorological Society, 140, 1935–1944, https://doi.org/10.1002/qj.2297, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2297, 2014b.

Cowtan, K., Hausfather, Z., Hawkins, E., Jacobs, P., Mann, M. E., Miller, S. K., Steinman, B. A., Stolpe, M. B., and Way, R. G.: Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures, Geophysical Research Letters, 42, 6526–6534, https://doi.org/https://doi.org/10.1002/2015GL064888, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL064888, 2015.

Crowley, T. J., Zielinski, G., Vinther, B., Udisti, R., Kreutz, K., Cole-Dai, J., and Castellano, E.: Volcanism and the little ice age, PAGES news, 16, 22–23, 2008.

de Lima, M. I. P. and Lovejoy, S.: Macroweather precipitation variability up to global and centennial scales, Water Resources Research, 51, 9490–9513, https://doi.org/10.1002/2015WR017455, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017455, 2015.

Del Rio Amador, L. and Lovejoy, S.: Predicting the global temperature with the Stochastic Seasonal to Interannual Prediction System (StocSIPS), Climate Dynamics, 53, 4373–4411, https://doi.org/10.1007/s00382-019-04791-4, https://doi.org/10.1007/s00382-019-04791-4, 2019.

Del Rio Amador, L. and Lovejoy, S.: Using scaling for seasonal global surface temperature forecasts: StocSIPS, Climate Dynamics, submitted, 2020.

Del Rio Amador, L. and Lovejoy, S.: Long-Range Forecasting as a Past Value Problem: Untangling Correlations and Causality With Scaling, Geophysical Research Letters, 48, e2020GL092 147, https://doi.org/https://doi.org/10.1029/2020GL092147, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL092147, e2020GL092147 2020GL092147, 2021a.

Del Rio Amador, L. and Lovejoy, S.: Using regional scaling for temperature forecasts with the Stochastic Seasonal to Interannual Prediction System (StocSIPS), Climate Dynamics, https://doi.org/10.1007/s00382-021-05737-5, 2021b.

Flynn, C. M. and Mauritsen, T.: On the Climate Sensitivity and Historical Warming Evolution in Recent Coupled Model Ensembles, Atmospheric Chemistry and Physics Discussions, 2020, 1–26, https://doi.org/10.5194/acp-2019-1175, https://www.atmos-chem-phys-discuss.net/acp-2019-1175/, 2020.

Forest, C. E., Stone, P. H., Sokolov, A. P., Allen, M. R., and Webster, M. D.: Quantifying Uncertainties in Climate System Properties with the Use of Recent Climate Observations, Science, 295, 113–117, https://doi.org/10.1126/science.1064419, https://science.sciencemag.org/content/295/5552/113, 2002.

Forest, C. E., Stone, P. H., and Sokolov, A. P.: Estimated PDFs of climate system properties including natural and anthropogenic forcings, Geophysical Research Letters, 33, https://doi.org/https://doi.org/10.1029/2005GL023977, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GL023977, 2006.

Forster, P. M., Maycock, A. C., McKenna, C. M., and Smith, C. J.: Latest climate models confirm need for urgent mitigation, Nature Climate Change, 10, 7–10, https://doi.org/10.1038/s41558-019-0660-0, https://doi.org/10.1038/s41558-019-0660-0, 2020.

Geoffroy, O., Saint-Martin, D., Olivié, D. J., Voldoire, A., Bellon, G., and Tytéca, S.: Transient climate response in a two-layer energy-balance model. Part I: Analytical solution and parameter calibration using CMIP5 AOGCM experiments, Journal of Climate, 26, 1841–1857, 2013.

Gregory, J. M. and Andrews, T.: Variation in climate sensitivity and feedback parameters during the historical period, Geophysical Research Letters, 43, 3911–3920, https://doi.org/https://doi.org/10.1002/2016GL068406, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL068406, 2016.

Hansen, J., Ruedy, R., Sato, M., and Lo, K.: Global surface temperature change, Rev. Geophys., 48, RG4004, https://doi.org/10.1029/2010RG000345, 2010.

Harries, J. E. and Belotti, C.: On the Variability of the Global Net Radiative Energy Balance of the Nonequilibrium Earth, Journal of Climate, 23, 1277 – 1290, https://doi.org/10.1175/2009JCLI2797.1, https://journals.ametsoc.org/view/journals/clim/23/6/2009jcli2797.1.xml, 2010.

Harvey, L. D. and Kaufmann, R. K.: Simultaneously constraining climate sensitivity and aerosol radiative forcing, Journal of Climate, 15, 2837–2861, 2002.

Hébert, R. and Lovejoy, S.: Interactive comment on "Global warming projections derived from an observation-based minimal model" by K. Rypdal, Earth System Dynamics, 7, 51–70, https://doi.org/10.5194/esd-7-51-2016, https://www.earth-syst-dynam.net/7/51/2016/, 2015.

Hébert, R. and Lovejoy, S.: Regional Climate Sensitivity- and Historical-Based Projections to 2100, Geophysical Research Letters, 45, 4248–4254, https://doi.org/10.1002/2017GL076649, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL076649, 2018.

Hébert, R., Lovejoy, S., and Tremblay, B.: An Observation-based Scaling Model for Climate Sensitivity Estimates and Global Projections to 2100, Climate Dynamics, 2021.

Held, I., Winton, M., Takahashi, K., Delworth, T., Zeng, F., and Vallis, G. K.: Probing the Fast and Slow Components of Global Warming by Returning Abruptly to Preindustrial Forcing, Journal of Climate, 23, 2418–2427, https://doi.org/10.1175/2009JCLI3466.1, https://doi.org/10.1175/2009JCLI3466.1, 2010.

Huang, B., Menne, M. J., Boyer, T., Freeman, E., Gleason, B. E., Lawrimore, J. H., Liu, C., Rennie, J. J., Schreck, C. J., Sun, F., Vose, R., Williams, C. N., Yin, X., and Zhang, H.-M.: Uncertainty Estimates for Sea Surface Temperature and Land Surface Air Temperature in NOAAGlobalTemp Version 5, Journal of Climate, 33, 1351 – 1379, https://doi.org/10.1175/JCLI-D-19-0395.1, https://journals.ametsoc.org/view/journals/clim/33/4/jcli-d-19-0395.1.xml, 2020.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/CBO9781107415324, www.climatechange2013.org, 2013.

Iseri, Y., Yoshikawa, S., Kiguchi, M., Tawatari, R., Kanae, S., and Oki, T.: Towards the incorporation of tipping elements in global climate risk management: probability and potential impacts of passing a threshold, Sustainability Science, 13, 315–328, https://doi.org/10.1007/s11625-018-0536-7, https://doi.org/10.1007/s11625-018-0536-7, 2018.

Johansson, D. J., O'Neill, B. C., Tebaldi, C., and Häggström, O.: Equilibrium climate sensitivity in light of observations over the warming hiatus, Nature Climate Change, 5, 449–453, 2015.

Kaufmann, R. K., Kauppi, H., Mann, M. L., and Stock, J. H.: Reconciling anthropogenic climate change with observed temperature 1998-2008, Proceedings of the National Academy of Sciences of the United States of America, 108, 11 790–11 793, https://doi.org/10.1073/pnas.1102467108, https://pubmed.ncbi.nlm.nih.gov/21730180, 2011.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., and Meehl, G. A.: Challenges in Combining Projections from Multiple Climate Models, Journal of Climate, 23, 2739 – 2758, https://doi.org/10.1175/2009JCLI3361.1, https://journals.ametsoc.org/view/journals/clim/23/10/2009jcli3361.1.xml, 2010.

Lenssen, N., Schmidt, G., Hansen, J., Menne, M., Persin, A., Ruedy, R., and Zyss, D.: Improvements in the GISTEMP uncertainty model, J. Geophys. Res. Atmos., 124, 6307–6326, https://doi.org/10.1029/2018JD029522, 2019.

Lewis, N. and Curry, J.: The Impact of Recent Forcing and Ocean Heat Uptake Data on Estimates of Climate Sensitivity, Journal of Climate, 31, 6051–6071, https://doi.org/10.1175/JCLI-D-17-0667.1, https://doi.org/10.1175/JCLI-D-17-0667.1, 2018.

Lewis, N. and Curry, J. A.: The implications for climate sensitivity of AR5 forcing and heat uptake estimates, Climate Dynamics, 45, 1009–1023, https://doi.org/10.1007/s00382-014-2342-y, https://doi.org/10.1007/s00382-014-2342-y, 2015.

Lovejoy, S.: What Is Climate?, Eos, Transactions American Geophysical Union, 94, 1–2, https://doi.org/10.1002/2013EO010001, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013EO010001, 2013.

Lovejoy, S.: Using scaling for macroweather forecasting including the pause, Geophysical Research Letters, 42, 7148–7155, https://doi.org/https://doi.org/10.1002/2015GL065665, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL065665, 2015a.

Lovejoy, S.: A voyage through scales, a missing quadrillion and why the climate is not what you expect, Climate Dynamics, 44, 3187–3210, https://doi.org/10.1007/s00382-014-2324-0, https://doi.org/10.1007/s00382-014-2324-0, 2015b.

Lovejoy, S.: Fractional relaxation noises, motions and the fractional energy balance equation, Nonlinear Processes in Geophysics Discussions, 2019, 1–52, https://doi.org/10.5194/npg-2019-39, https://www.nonlin-processes-geophys-discuss.net/npg-2019-39/, 2019a.

Lovejoy, S.: Weather, Macroweather and Climate: our random yet predictable atmosphere, Oxford U. Press, 2019b.

Lovejoy, S.: The fractional heat equation, Nonlinear Processes in Geophysics, submitted, 2020.

Lovejoy, S.: The Half-order Energy Balance Equation, Part 1: The homogeneous HEBE and long memories, Earth System Dynamics Discussions, pp. 1–36, https://doi.org/10.5194/esd-2020-12, https://www.earth-syst-dynam-discuss.net/esd-2020-12/, 2021a.

Lovejoy, S.: The Half-order Energy Balance Equation, Part 2:The inhomogeneous HEBE and 2D energy balance models, Earth System Dynamics Discussions, pp. 1–44, https://doi.org/10.5194/esd-2020-13, https://www.earth-syst-dynam-discuss.net/esd-2020-13/, 2021b.

Lovejoy, S. and Schertzer, D.: The Weather and Climate: Emergent Laws and Multifractal Cascades, Cambridge University Press, https://doi.org/10.1017/CBO9781139093811, 2013.

Lovejoy, S. and Varotsos, C.: Scaling regimes and linear/nonlinear responses of last millennium climate to volcanic and solar forcings, Earth Syst. Dynam, 7, 133–150, 2016.

Lovejoy, S., Del Rio Amador, L., and Hebert, R.: The ScaLIng Macroweather Model (SLIMM): using scaling to forecast global-scale macroweather from months to decades, Earth System Dynamics, 6, 637, 2015.

Lovejoy, S., Del Rio Amador, L., and Hébert, R.: Harnessing Butterflies: Theory and Practice of the Stochastic Seasonal to Interannual Prediction System (StocSIPS), pp. 305–355, https://doi.org/10.1007/978-3-319-58895-7_17, 2017.

Lovejoy, S., Procyk, R., Hébert, R., and Del Rio Amador, L.: The fractional energy balance equation, Quarterly Journal of the Royal Meteorological Society, n/a, https://doi.org/https://doi.org/10.1002/qj.4005, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4005, 2021.

Medhaug, I., Stolpe, M., Fischer, E., and Knutti, R.: Reconciling controversies about the 'global warming hiatus', Nature, 545, 41–47, https://doi.org/10.1038/nature22315, 2017.

Meehl, G., Arblaster, J., Fasullo, J., Hu, A., and Trenberth, K.: Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods, Nature Climate Change, 1, 360–364, https://doi.org/10.1038/nclimate1229, 2011.

Meinshausen, M., Raper, S. C. B., and Wigley, T. M. L.: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 – Part 1: Model description and calibration, Atmospheric Chemistry and Physics, 11, 1417–1456, https://doi.org/10.5194/acp-11-1417-2011, https://www.atmos-chem-phys.net/11/1417/2011/, 2011a.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M., Lamarque, J.-F., Matsumoto, K., Montzka, S., Raper, S., Riahi, K., et al.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic change, 109, 213, 2011b.

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., Smith, S. J., van den Berg, M., Velders, G. J. M., Vollmer, M. K., and Wang, R. H. J.: The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500, Geoscientific Model Development, 13, 3571–3605, https://doi.org/10.5194/gmd-13-3571-2020, https://gmd.copernicus.org/articles/13/3571/2020/, 2020.

Millar, R. J., Otto, A., Forster, P. M., Lowe, J. A., Ingram, W. J., and Allen, M. R.: Model structure in observational constraints on transient climate response, Climatic Change, 131, 199–211, 2015.

Morice, C. P., Kennedy, J. J., Rayner, N. A., and Jones, P. D.: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2011JD017187, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD017187, 2012.

Murphy, D., Solomon, S., Portmann, R., Rosenlof, K., Forster, P., and Wong, T.: An observationally based energy balance for the Earth since 1950, Journal of Geophysical Research: Atmospheres, 114, 2009.

Myhre, G., Highwood, E. J., Shine, K. P., and Stordal, F.: New estimates of radiative forcing due to well mixed greenhouse gases, Geophysical research letters, 25, 2715–2718, 1998.

Myhre, G., Samset, B. H., Schulz, M., Balkanski, Y., Bauer, S., Berntsen, T. K., Bian, H., Bellouin, N., Chin, M., Diehl, T., Easter, R. C., Feichter, J., Ghan, S. J., Hauglustaine, D., Iversen, T., Kinne, S., Kirkevåg, A., Lamarque, J.-F., Lin, G., Liu, X., Lund, M. T., Luo, G., Ma, X., van Noije, T., Penner, J. E., Rasch, P. J., Ruiz, A., Seland, Ø., Skeie, R. B., Stier, P., Takemura, T., Tsigaridis, K., Wang, P., Wang, Z., Xu, L., Yu, H., Yu, F., Yoon, J.-H., Zhang, K., Zhang, H., and Zhou, C.: Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations, Atmospheric Chemistry and Physics, 13, 1853–1877, https://doi.org/10.5194/acp-13-1853-2013, https://www.atmos-chem-phys.net/13/1853/2013/, 2013.

Myrvoll-Nilsen, E., Sørbye, S. H., Fredriksen, H.-B., Rue, H., and Rypdal, M.: Statistical estimation of global surface temperature response to forcing under the assumption of temporal scaling, Earth System Dynamics, 11, 329–345, https://doi.org/10.5194/esd-11-329-2020, https://www.earth-syst-dynam.net/11/329/2020/, 2020.

Nazarenko, L., Rind, D., Tsigaridis, K., Del Genio, A. D., Kelley, M., and Tausnev, N.: Interactive nature of climate change and aerosol forcing, Journal of Geophysical Research: Atmospheres, 122, 3457–3480, https://doi.org/10.1002/2016JD025809, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD025809, 2017.

North, G. R.: Theory of Energy-Balance Climate Models, Journal of the Atmospheric Sciences, 32, 2033–2043, https://doi.org/10.1175/1520-0469(1975)032<2033:TOEBCM>2.0.CO;2, https://doi.org/10.1175/1520-0469(1975)032<2033:TOEBCM>2.0.CO;2, 1975.

North, G. R. and Kim, K.-Y.: Energy Balance Climate Models, John Wiley & Sons, 2017.

North, G. R., Cahalan, R. F., and Coakley Jr, J. A.: Energy balance climate models, Reviews of Geophysics, 19, 91–121, 1981, Reviews.

Otto, A., Otto, F. E., Boucher, O., Church, J., Hegerl, G., Forster, P. M., Gillett, N. P., Gregory, J., Johnson, G. C., Knutti, R., et al.: Energy budget constraints on climate response, Nature Geoscience, 6, 415–416, 2013.

Padilla, L. E., Vallis, G. K., and Rowley, C. W.: Probabilistic Estimates of Transient Climate Sensitivity Subject to Uncertainty in Forcing and Natural Variability, Journal of Climate, 24, 5521–5537, https://doi.org/10.1175/2011JCLI3989.1, https://doi.org/10.1175/2011JCLI3989.1, 2011.

Penner, J., Andreae, M., Annegarn, H., Barrie, L., Feichter, J., Hegg, D., Achuthan, J., Leaitch, R., Murphy, D., Nganga, J., and Pitari, G.: Aerosols, their Direct and Indirect Effects, Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, 289-348 (2001), 2001.

Podlubny, I.: Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications, Elsevier, 1999.

Proistosescu, C., Donohoe, A., Armour, K. C., Roe, G. H., Stuecker, M. F., and Bitz, C. M.: Radiative feedbacks from stochastic variability in surface temperature and radiative imbalance, Geophysical Research Letters, 45, 5082–5094, 2018.

Ramaswamy, V., Boucher, O., Haigh, J., Hauglustaine, D., Haywood, J., Myhre, G., Nakajima, T., Shi, G., and Solomon, S.: Radiative Forcing of Climate Change, pp. 349–416, Cambridge University Press, 2001.

Ring, M. J., Lindner, D., Cross, E. F., and Schlesinger, M. E.: Causes of the Global Warming Observed since the 19th Century, Atmospheric and Climate Sciences, 02, 401–415, https://doi.org/10.4236/acs.2012.24035, 2012.

Rohde, R., Muller, R., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., Wurtele, J., Groom, D., and Wickham, C.: A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011, Geoinfor Geostat: An Overview 1: 1, of, 7, 2, 2013.

Rohde, R. A. and Hausfather, Z.: The Berkeley Earth Land/Ocean Temperature Record, Earth System Science Data, 12, 3469–3479, https://doi.org/10.5194/essd-12-3469-2020, https://essd.copernicus.org/articles/12/3469/2020/, 2020.

Rypdal, K.: Global temperature response to radiative forcing: Solar cycle versus volcanic eruptions, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2011JD017283, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD017283, 2012.

Rypdal, M. and Rypdal, K.: Long-Memory Effects in Linear Response Models of Earth's Temperature and Implications for Future Global Warming, Journal of Climate, 27, 5240–5258, https://doi.org/10.1175/JCLI-D-13-00296.1, https://doi.org/10.1175/JCLI-D-13-00296.1, 2014.

Sato, M.: Forcings in GISS Climate Model: Stratospheric Aerosol Optical Thickness., https://data.giss.nasa.gov/modelforce/strataer/, 2012.

Sato, Y., Goto, D., Michibata, T., Suzuki, K., Takemura, T., Tomita, H., and Nakajima, T.: Aerosol effects on cloud water amounts were successfully simulated by a global cloud-system resolving model, Nature Communications, 9, 985, https://doi.org/10.1038/s41467-018-03379-6, https://doi.org/10.1038/s41467-018-03379-6, 2018.

Schurer, A. P., Mann, M. E., Hawkins, E., Tett, S. F. B., and Hegerl, G. C.: Importance of the pre-industrial baseline for likelihood of exceeding Paris goals, Nature Climate Change, 7, 563–567, https://doi.org/10.1038/nclimate3345, https://doi.org/10.1038/nclimate3345, 2017.

Schwartz, S. E.: Uncertainty in climate sensitivity: causes, consequences, challenges, Energy & environmental science, 1, 430–453, 2008.

Sellers, W. D.: A global climatic model based on the energy balance of the earth-atmosphere system, Journal of Applied Meteorology, 8, 392–400, 1969.

Sherwood, S., Webb, M. J., Annan, J. D., Armour, K., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., et al.: An assessment of Earth's climate sensitivity using multiple lines of evidence, Reviews of Geophysics, 58, e2019RG000 678, 2020.

Skeie, R. B., Berntsen, T., Aldrin, M., Holden, M., and Myhre, G.: A lower and more constrained estimate of climate sensitivity using updated observations and detailed radiative forcing time series, Earth System Dynamics, 5, 139–175, https://doi.org/10.5194/esd-5-139-2014, https://www.earth-syst-dynam.net/5/139/2014/, 2014.

Smith, C. J., Forster, P. M., Allen, M., Leach, N., Millar, R. J., Passerello, G. A., and Regayre, L. A.: FAIR v1.3: a simple emissions-based impulse response and carbon cycle model, Geoscientific Model Development, 11, 2273–2297, https://doi.org/10.5194/gmd-11-2273-2018, https://www.geosci-model-dev.net/11/2273/2018/, 2018a.

Smith, D. M., Scaife, A. A., Hawkins, E., Bilbao, R., Boer, G. J., Caian, M., Caron, L.-P., Danabasoglu, G., Delworth, T., Doblas-Reyes, F. J., Doescher, R., Dunstone, N. J., Eade, R., Hermanson, L., Ishii, M., Kharin, V., Kimoto, M., Koenigk, T., Kushnir, Y., Matei, D., Meehl, G. A., Menegoz, M., Merryfield, W. J., Mochizuki, T., Müller, W. A., Pohlmann, H., Power, S., Rixen, M., Sospedra-Alfonso, R., Tuma, M., Wyser, K., Yang, X., and Yeager, S.: Predicted Chance That Global Warming Will Temporarily Exceed 1.5 °C, Geophysical Research Letters, 45, 11,895–11,903, https://doi.org/10.1029/2018GL079362, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL079362, 2018b.

Smith, T. M., Reynolds, R. W., Peterson, T. C., and Lawrimore, J.: Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006), Journal of Climate, 21, 2283–2296, 2008.

Solomon, S.: Climate Change 2007 the physical science basis: contribution of Working Group I to the Fourth Assessment Report of the IPCC, Cambridge University Press, 2007.

Solomon, S., Plattner, G.-K., and Friedlingstein, P.: Irreversible climate change due to carbon dioxide emissions, Proceedings of the National Academy of Sciences, USA, 2009.

Stainforth, D., Allen, M., Tredger, E., and Smith, L.: Confidence, uncertainty and decision-support relevance in climate predictions, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365, 2145–2161, https://doi.org/10.1098/rsta.2007.2074, https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2007.2074, 2007.

Stevens, B.: Rethinking the Lower Bound on Aerosol Radiative Forcing, Journal of Climate, 28, 4794–4819, https://doi.org/10.1175/JCLI-D-14-00656.1, https://doi.org/10.1175/JCLI-D-14-00656.1, 2015.

Stouffer, R. J.: Time scales of climate response, Journal of Climate, 2004.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498, 2012.

Tebaldi, C. and Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365, 2053–2075, https://doi.org/10.1098/rsta.2007.2076, https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2007.2076, 2007.

Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., and Knutti, R.: Past warming trend constrains future warming in CMIP6 models, Science Advances, 6, https://doi.org/10.1126/sciadv.aaz9549, https://advances.sciencemag.org/content/6/12/eaaz9549, 2020.

Tomassini, L., Reichert, P., Knutti, R., Stocker, T. F., and Borsuk, M. E.: Robust Bayesian Uncertainty Analysis of Climate System Properties Using Markov Chain Monte Carlo Methods, Journal of Climate, 20, 1239 – 1254, https://doi.org/10.1175/JCLI4064.1, https://journals.ametsoc.org/view/journals/clim/20/7/jcli4064.1.xml, 2007.

Trenberth, K. E., Fasullo, J. T., and Kiehl, J.: Earth's Global Energy Budget, Bulletin of the American Meteorological Society, 90, 311 – 324, https://doi.org/10.1175/2008BAMS2634.1, https://journals.ametsoc.org/view/journals/bams/90/3/2008bams2634_1.xml, 2009.

Trenberth, K. E., Fasullo, J. T., and Balmaseda, M. A.: Earth's energy imbalance, Journal of Climate, 27, 3129–3144, 2014.

Wang, Y.-M., Lean, J., and Sheeley Jr, N.: Modeling the Sun's magnetic field and irradiance since 1713, The Astrophysical Journal, 625, 522, 2005.

Weisheimer, A. and Palmer, T. N.: On the reliability of seasonal climate forecasts, Journal of The Royal Society Interface, 11, 20131 162, https://doi.org/10.1098/rsif.2013.1162, https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2013.1162, 2014.

Wolfram Research, Inc.: Mathematica, Version 12.2, https://www.wolfram.com/mathematica, champaign, IL, 2020.

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of Higher Climate Sensitivity in CMIP6 Models, Geophysical Research Letters, 47, e2019GL085 782, https://doi.org/10.1029/2019GL085782, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL085782, e2019GL085782 10.1029/2019GL085782, 2020.

Zhang, H., Huang, B., Lawrimore, J., Menne, M., and Smith, T. M.: Global Surface Temperature Dataset (NOAAGlobalTemp), Version 4.0,[NOAA Global Surface Temperature Data], https://doi.org/doi:10.7289/V5FN144H [01/03/2020], 2019.

Zhou, C. and Penner, J. E.: Why do general circulation models overestimate the aerosol cloud lifetime effect? A case study comparing CAM5 and a CRM, Atmospheric Chemistry and Physics, 17, 21–29, https://doi.org/10.5194/acp-17-21-2017, https://www.atmos-chem-phys.net/17/21/2017/, 2017.

Ziegler, E. and Rehfeld, K.: TransEBM v. 1.0: Description, tuning, and validation of a transient model of the Earth's energy balance in two dimensions, Geoscientific Model Development Discussions, 2020, 1–36, https://doi.org/10.5194/gmd-2020-237, https://gmd.copernicus.org/preprints/gmd-2020-237/, 2020.