The Fractional Energy Balance Equation for Climate projections through 2100

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

[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

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

Abstract.

We produce climate projections through the 21st century using the fractional energy balance equation (FEBE) which is a generalization of the standard EBE. The FEBE can be derived either from Budyko-Sellers models or phenomenologically by applying the 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 parameters: a scaling exponent H and relaxation time τ; its amplitude parameter is 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 predicted by the model itself: the internal variability response to white noise internal forcing.

The 90% Confidence Interval (CI) of the shape parameters were $H = [0.33, 0.44]$ (median=0.38), $\tau = [2.4, 7.0]$ (median=4.7) years compared to the usual EBE $H = 1$, and literature values τ typically in the range 2-8 years. We found that aerosols were too strong 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 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 seen in other observational-based studies, the FEBE values 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 CMIP5/6 MME. The FEBE hindprojections (1880-2019) closely follow observations (notably during the "hiatus", 1998-2015). Overall the FEBE 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.

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)). From dissipation times up until scales of about 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. Eventually, in the industrial epoch at a scale of ≈ 20 years, this is reversed and fluctuations again tend to increase. This marks the beginning of the climate regime, in the pre-industrial epoch the transition is at centuries or millennia and the regime continues up to Milankovitch scales ((Lovejoy, 2015), (Lovejoy, 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 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 macroweather and climate responses to external perturbations are typically linear (as quantified for CMIP5 models in (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), EBMs have been extensively used for understanding the climate ((North, 1975), (North et al., 1981, Review)). 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). However the use of pure power law CRFs (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) 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.

60   (Hébert et al., 2020) 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., 2020) 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's

65   globally averaged temperature, both models were thus effectively equivalent. The main drawback was that the CRF model truncation was somewhat ad hoc. In addition, due to the truncation, it was 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. As pointed out in ((Lovejoy, 2019b), (Lovejoy, 2019a)) and developed in (Lovejoy et al., 2020) it is sufficient that the scaling principle not

70   be applied to the CRF, but rather to the storage term in the EBE. Rather than 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., 2020) derived the FEBE in a phenomenological manner, ((Lovejoy, 2020b), (Lovejoy, 2020c), (Lovejoy, 2020a)) showed how it could instead be derived

75   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 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

80   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's truncated power law, the high frequency regime is able to produce internal variability coherent with the observed scaling and fractional Gaussian noise used by ((Lovejoy et al.,

85   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. 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 introduce

90   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, 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 CMIP5 GCMs Multi-Model Ensemble (MME) and the currently available CMIP6 GCMs with which they are compared.

2 Methods and Material

2.1 The FEBE

The zero-dimensional FEBE may be written:

$$\tau^H {}_{-\infty}D_t^H T + T = \lambda F; \quad 0 \leq H \leq 1 \tag{1}$$

((Lovejoy, 2019a)), (Lovejoy et al., 2020)) Where $T(t)$ is the Earth temperature anomaly with respect to a reference temperature $(T(-\infty) = 0)$, τ is the relaxation time, λ the climate sensitivity, $F(t)$ the anomalous external radiative forcing, and H the order of the ("Weyl") fractional derivative:

$$_{-\infty}D_t^H T = \frac{1}{\Gamma(1-H)} \int_{-\infty}^{t} (t'-s)^H T'(t')ds, \quad T'(s) = \frac{dT}{ds} \tag{2}$$

(Γ is the Gamma function). If this derivative is integrated by parts and the limit $H \to 1$ is taken, using $(T(-\infty) = 0)$, $_{-\infty}D_t^H T = \frac{dT}{dt}$ so that we recover the standard box EBE. The FEBE thus generalizes the EBE and the power law FEBE relaxation time τ generalizes the exponential EBE relaxation time. Physically T/λ corresponds to the linearized energy flux (rate per area) of long wave black body surface emission and since F is the flux of shortwave forcing, we see that the imbalance $F - T/\lambda$ 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 $H = 1$ where the relationship is instantaneous.

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

$$T(t) = \lambda \int_{-\infty}^{t} G_{\delta,H}(t-s)F(s)ds \tag{3}$$

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

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

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)). The condition $G = 0$ for $t < 0$ is needed to respect causality. The Mittag-Leffler functions are often called "generalized exponentials", the classical $H = 1$ box model is the (exceptional) ordinary exponential: $E_{1,1}(z) = e^z$.

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

$$G_{\Theta,H}(t) = \int_0^t G_{\delta,H}(s)ds = \left(\frac{t}{\tau}\right)^H E_{H,H+1}\left(-\left(\frac{t}{\tau}\right)^H\right) \tag{6}$$

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

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

At high frequencies ($t \ll \tau$), important for modelling and predicting the internal variability we have:

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

If we consider the response to Gaussian white noise forcing, then $G_{\delta,H} \propto t^{H-1}$ implies that $T(t)$ is approximately a fractional Gaussian noise (fGn) with statistical scaling exponent $H_I = H - 1/2$ (when forced by a Gaussian white noise, the FEBE response is exactly a fractional Relaxation noise, see (Lovejoy, 2020a)). By applying global scale Haar fluctuation analyses (Del Rio Amador and Lovejoy, 2019) found $H_I \approx -0.1$ corresponding to $H \approx 0.4$.

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:

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

(note $\Gamma(-H) < 0$ for $0 < H < 1$). In the box, $H = 1$, case we have exactly $G_{\Theta,1}(t) = 1 - e^{-t/\tau}$ so that when $H < 1$, the exponential approach to thermodynamic equilibrium is replaced by a power law. (Hébert et al., 2020) used $G_\Theta = 1 - \left(1 + \frac{t}{\tau}\right)^{H_F}$ with $H_F = -0.5$ corresponding to $H = -H_F = 0.5$ which is thus the same as that corresponding to the internal forcing. It is thus plausible that the FEBE models both regimes with the exponent $H \approx 0.4$. We discuss this further below.

2.2 Data

2.2.1 Radiative Forcing Data

We consider natural and anthropogenic sources of external forcing: solar and volcanic, and anthropogenic forcing. We use the approximate 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.

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).

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) we investigate SSP 126 (strong mitigation) and SSP 585 (strong emission) scenarios, the counterparts to the most distinct previous scenarios RCP 2.6 and 8.5.

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)), while for the SSP scenarios the emissions are converted to forcings using the Finite Amplitude Impulse Response (FAIR) model ((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 (RCP 8.5, SSP 585) shown in fig. 1 (bottom). For RCP2.6 and SSP 126, the strongest mitigation scenarios, the total radiative forcing has a peak at approximately $3 W m^{-2}$ around the year 2050 and declines thereafter due to large scale deployment of negative emission technologies. RCP4.5 is a stabilization scenario, 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 continuously rising radiative forcing pathway, "business as usual", in which the radiative forcing levels by the end of the 21st century at approximately $8.5 W m^{-2}$, and most closely follows current emissions.

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

170   by Long Lived Greenhouse Gases (GHGs) F_{GHG}: carbon dioxide, methane, nitrous oxide and fluorinated gases, controlled

under the Kyoto protocol, 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

175   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, compared to the fairly well constrained GHG forcing, thus following ((Padilla

et al., 2011), (Hébert et al., 2020)) we 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 the combined effective radiative forcing

180   from 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) to convert this to Wm^{-2}.

The total amount of aerosol forcing in 2005 given at the 90% confidence interval (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 more

negative than $-1Wm^{-2}$ is implausible. Using results from (Murphy et al., 2009), Stevens supports tightening the upper and

185   lower bounds of the aerosol forcing, revising it to be $[-1.0, -0.3]Wm^{-2}$.

The prescribed CMIP6 SSP aerosol forcing, $F_{Aer_{SSP}}$, contain 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

190   radiation budget and is calculated from the aerosol model of (Stevens, 2015), which includes a logarithmic dependence of F_{aci}

on sulphates, black carbon and organic carbon emissions.
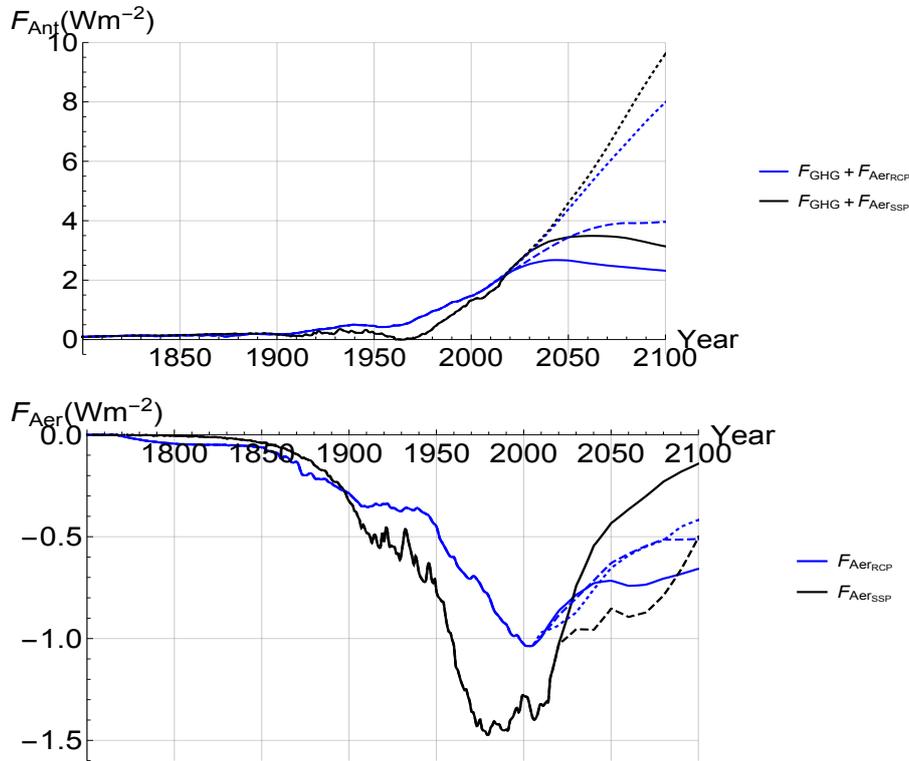
Figure 1. (top) The total anthropogenic forcing series, the sum of the greenhouse gas forcing F_{GHG} and respective aerosol forcing series $F_{Aer_{RCP}}$ (black) or $F_{Aer_{SSP}}$ (blue) are shown over the historical period and projection period until 2100 for RCP 2.6/SSP 126 (solid), RCP 4.5 (dashed), and RCP 8.5/SSP 585 (dotted). (bottom) The anthropogenic aerosol forcing series used, $F_{Aer_{RCP}}$ (blue) and $F_{Aer_{SSP}}$ (black). Updated from (Hébert et al., 2020).

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

195 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. 2. Following (Meinshausen et al., 2011b), 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) and multiplied by 0.7 (representing planetary co-albedo). To extend solar forcing to the future we reproduce solar cycle 23 (the last one prior to 2008) as a proxy for future solar forcing.

200 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: $F_{Vol} \approx -27 W m^{-2} \tau_V$, obtained from the Goddard Institute for Space Science

(GISS) website (Sato, 2012). We follow (Hébert et al., 2020), extending the series to 1765 using the optical depth reconstruction (Crowley et al., 2008), and setting volcanic forcing to zero for the future.

205   It has been shown by (Lewis and Curry, 2015) 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. 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)). Since linear response models do not alter the intermittency, the volcanic series must first be non-linearly transformed before being introduced into a linear

210 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., 2020) using a non-linear relation to change the intermittency so that the transformed signal can be linearly related to the temperature:

$$\frac{F_{Vol_\nu}}{\langle F_{Vol} \rangle} = \frac{F_{Vol}^\nu}{\langle F_{Vol}^\nu \rangle} \tag{11}$$

The normalization is such that the mean is unchanged: $\langle F_{Vol_\nu} \rangle = \langle F_{Vol} \rangle$. The volcanic intermittency correction exponent,

215 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

220 $\alpha_{MF} \approx 1.5$, 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)).

Working in a linear framework we write the forcing series as the sum of anthropogenic and natural forcings:

$$F(\alpha, \nu; t) = F_{GHG}(t) + \alpha F_{Aer}(t) + F_{Sol} + F_{Vol_\nu} \tag{13}$$

2.2.6   Surface Air Temperature Data and CMIP5/6 Simulations

225 We used five historical records of surface air temperature for our analysis each spanning the period 1880-2019, 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, 2014), GISS Surface Temperature Analysis (GISTEMP) (Hansen et al., 2010), NOAA Merged Land Ocean Global Surface Temperature Analysis Dataset (NOAAGlobalTemp) ((Smith et al., 2008), (Zhang et al., 2019)) and Berkley Earth Surface Temperature (BEST) (Rohde et al., 2013).

230   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;
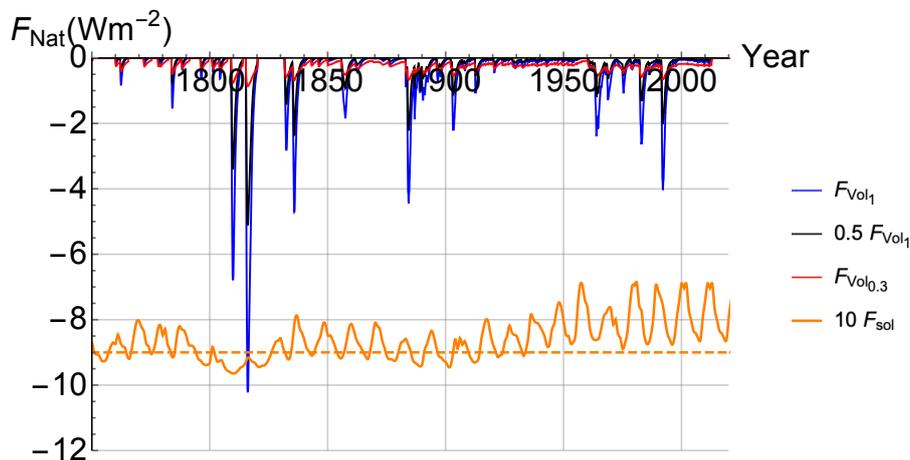
Figure 2. Volcanic forcing F_{Vol_1} (blue) is shown alongside two non-linearly transformed versions. Linearly damped by a constant 0.5 coefficient (black), and non-linearly transformed using equation with $\nu = 0.3$ (red). The solar forcing F_{Sol} (orange) has been shifted down by -9 and amplified by a factor of 10 for clarity. Adapted from (Hébert et al., 2020)

the Cowtan and Way dataset uses HadCRUT4 as raw data, but address 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 climate sensitivity of models is taken from (Flynn and Mauritsen, 2020), summarized in table A2.

2.3 Bayesian Parameter Estimation

In this section we establish a procedure to estimate the probability distribution associated with the climate sensitivity: λ, model parameters: τ, H and the forcing parameters: α, ν. To estimate them, we relate the forcing to surface air temperature data using the FEBE in 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, each parameter combination (H, τ for $G_{\delta,H}$ and α, ν for F as well as λ) 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 external deterministic forcing $F(t)$ and internal stochastic forcing $\gamma(t)$:

255 $$T(t) = T_{ext} + T_{int}(t); \quad \begin{matrix} T_{ext} = \lambda G_{\delta,H}(t) * F(t) \\ T_{int} = \lambda G_{\delta,H}(t) * \gamma(t) \end{matrix} . \tag{14}$$

Where T_{ext}, T_{int} are the responses. Any given set of parameters (H, τ for $G_{\delta,H}$ and α, ν for F as well as λ) implies a series $T_{ext}(t)$; and hence when compared to observation temperature series, it implies a series of residuals:

$$T_{res}(t) = T(t) - T_{ext}(t) = T_{int}(t) = \lambda G_{\delta,H}(t) * \gamma(t). \tag{15}$$

The residuals are thus equal to the internal temperature variability, i.e. the response to the internal forcing $\gamma(t)$. Here we make
260 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)). 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 $G_{\delta,H} \approx G_{\delta,high,H}(t)$, eq. 8). The fGn approximation takes into account the strong power law correlations induced by the fractional derivative term in the FEBE and it is valid except at the low frequencies that only weakly influence the likelihood function. It is already a much more accurate residual model than
265 the standard Auto Regressive order one (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.

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

$$\mathcal{L}(\lambda, H, \tau, \alpha, \nu | T(t)) = Pr(T(t) | \lambda, H, \tau, \alpha, \nu). \tag{16}$$

The likelihood function ($\mathcal{L}$) corresponds to the probability ("Pr") of observing the series $T(t)$ conditioned on the parameters: $\lambda, H, \tau, \alpha, \nu$ (right hand side), assuming the residuals are a fGn process with parameter H, and zero mean.

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

$$Pr(\lambda, H, \tau, \alpha, \nu | T(t)) = \frac{Pr(T(t) | \lambda, H, \tau, \alpha, \nu) \pi(\lambda, H, \tau, \alpha, \nu)}{Pr(T(t))} \tag{17}$$

where $\pi(\lambda, H, \tau, \alpha, \nu)$ is the prior distribution for the parameters.

The priors chosen here are intended to reflect knowledge about the historical climate system. Following (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), 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 $H = 1$, found by (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, $F_{Aer} \approx -1.0 W m^{-2}$, in the series we used. For the remaining two parameters, λ and ν, we assume non-informative uniform priors over the range of parameters; $\lambda \in [1.0, 4.0]$ and $\nu \in [0.0, 1.0]$.

We approximate our five dimensional parameter space, $(\lambda, H, \tau, \alpha, \nu)$ by a multinormal distribution:

$$P(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{5}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-(\boldsymbol{x}-\boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})/2}, \tag{18}$$

where $\boldsymbol{x} = \{\lambda, \tau, H, \alpha, \nu\}$, 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 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 different observational datasets 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.

$$Pr(\lambda, H, \tau, \alpha, \nu | T(t)) = \frac{1}{n} \sum_{i=1}^{n} Pr(\lambda, H, \tau, \alpha, \nu | T_i(t)) \tag{19}$$

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) ($H = 1/2$) that is a consequence of the continuum heat equation ((Lovejoy, 2020b), (Lovejoy, 2020c)), as well as with the precursor Scaling Climate Response Function (SCRF) model ((Hébert et al., 2020)) which differs primarily in the treatment of high frequencies, shown in table 3.
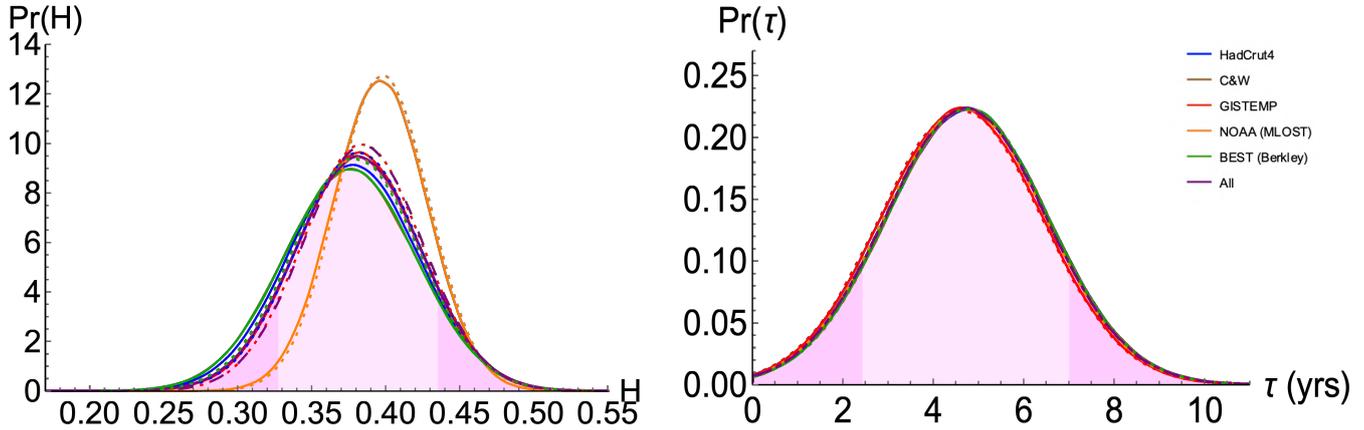
Figure 3. For each observational dataset and their average, PDFs are shown for the model parameters: the scaling parameter H (left), and the transition time τ (right). Shown are the PDFs for parameter estimation based on both $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{SSP}}$ (dashed). The average PDFs of the five observation datasets using $F_{Aer_{RCP}}$ is shown as the main result with shading, with darker 5% tails.

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

3.1.1 The Scaling Exponent H

The model is characterized by H and τ, where H is the most fundamental. For 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 $H = \frac{1}{2}$ 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., 2020), (Lovejoy et al., 2017)).

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: a large τ can be somewhat compensated by a smaller λ 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) 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 N[4,2] prior distribution of τ. With this prior, we obtained the a posteriori 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 $F_{Aer_{SSP}}$.

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. 4 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

325   the a posteriori distribution was found to have a median value of 0.6 with a 90% CI of [0.2,1.0].

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. 4 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

330   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)), 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

335   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 $F_{Aer_{SSP}}$. 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 4. Thus, the original volcanic series described without the intermittency correction does not reproduce well, within the FEBE model presented, the cooling events observed in

340   instrumental records following eruptions: the cooling would be overestimated.

In fig. 5 (left) we compare the total forcing series, $F_{Tot}(t)$ (black), IPCC AR5, eq. (13) 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

345   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.
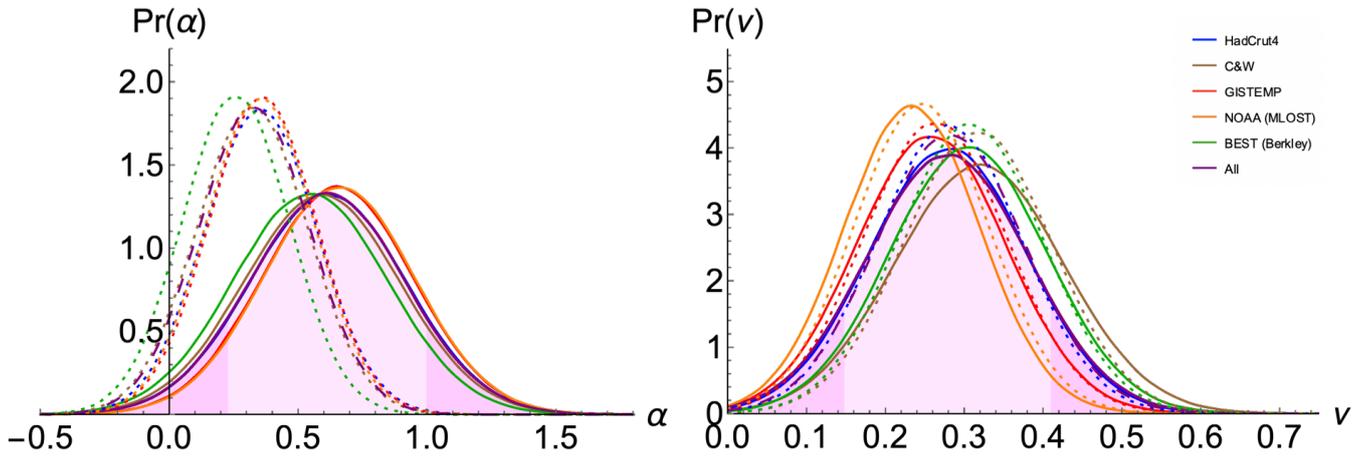
Figure 4. For each observational dataset and their average, PDFs are shown for the forcing parameters: the aerosol scaling factor α (left), and the volcanic intermittency correction exponent ν (right). Again, shown are the PDFs for parameter estimation based on both $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{SSP}}$ (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
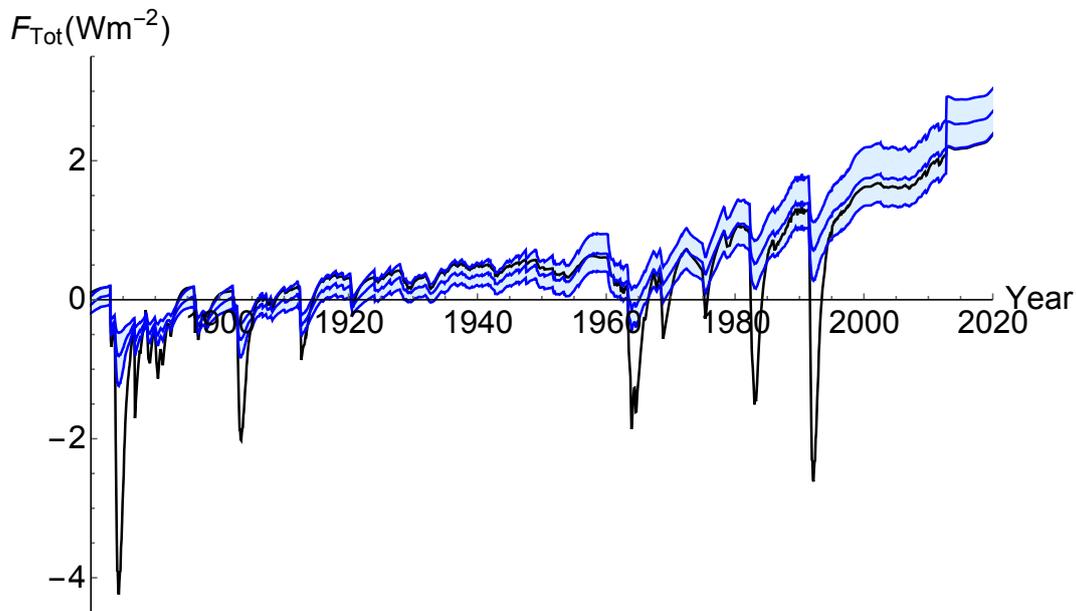
Figure 5. The total forcing series prescribed by the IPCC using, $F_{Aer_{RCP}}$ (black) compared to the adjusted forcing, $F_{Tot}(\alpha, \nu; t)$ (blue) which takes into account aerosol and volcanic corrections, shown over the historical period (1880-2019) with 90% CI.

Presented in fig. 6 (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). Considering $G_{1,H}$ (blue),

at scales below a few years (τ) whereas the box models or the (Hébert et al., 2020) truncated scaling model are smooth, the

350 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

355 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.
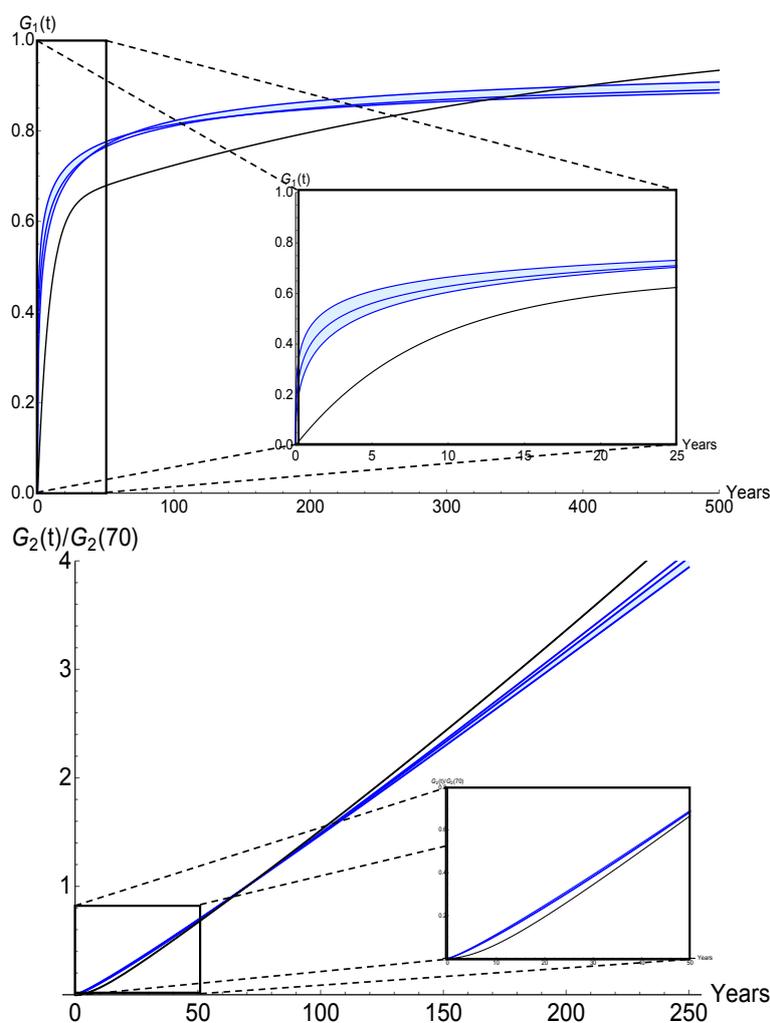
Figure 6. (top) The median and 90% CI of the FEBE step-response Green's function, $G_1(H, \tau; t)$, compared to the IPCC two-box model Green's function (black). (bottom) The median and 90% CI of the FEBE normalized ramp-response Green's function, $G_2(H, \tau; t)$, compared to the IPCC two-box model Green's function (black).

3.3 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-
360 model comparisons: Equilibrium Climate Sensitivity (ECS) and Transient Climate Response (TCR). Climate sensitivity de-
scribes the amount of warming in the atmosphere associated with increases in atmospheric carbon dioxide (CO_2).

If atmospheric CO_2 were 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
365 when the entire climate system reaches 'equilibrium' or the stable temperature response to a doubling of CO_2, it is equal to λ
when measured in units of $K/(CO_2$ doubling).

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 $F_{Aer_{SSP}}$. 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
370 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), 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.
375 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%
380 CI [0.45,0.67] $K(Wm^{-2})^{-1}$ using $F_{Aer_{RCP}}$ (fig. 7). When using $F_{Aer_{SSP}}$ we find median $0.52K(Wm^{-2})^{-1}$ 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)).
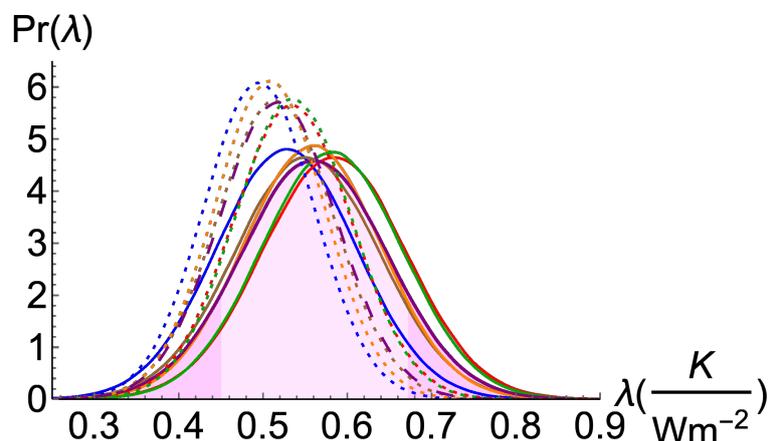
Figure 7. For each observational dataset and their average, PDFs are shown for the climate sensitivity parameter λ (the ECS, here in units of $K(Wm^{-2})^{-1})$, $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{SSP}}$ (dashed)

3.3.2 Transient Climate Response

385 Conventionally, TCR quantifies the changes that would occur if CO_2 levels increase by 1% (compounded) per year until they double (≈ 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. 8 (middle). Our TCR was found to have a 90% CI of $[1.2, 1.8]K$ with a median

390 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 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

395 (RWF) ((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. 8 (right), having a 90% CI [0.70, 0.78] and median 0.73 using $F_{Aer_{RCP}}$. Similar results are found using $F_{Aer_{SSP}}$, a median of 0.72 and 90% CI [0.71, 0.79]. 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.

400 In the next section we show that with a lower and more constrained climate sensitivity parameter (figs. 7 and 8), the adjusted forcing and long memory process of the FEBE (fig. 5) produces future projections that tend to be cooler than the CMIP5/6 projections, yet within their 90% CI.
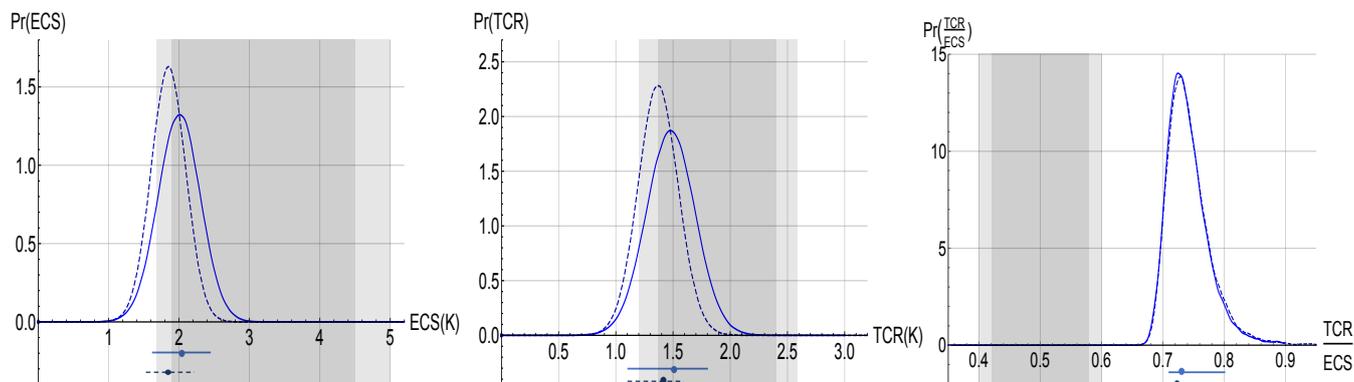
Figure 8. The PDFs for ECS (left), TCR (middle) and the TCR:ECS ratio (right) are derived using $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{SSP}}$ (dashed). The associated 90% CI (bars under the axis), the CMIP5 MME 90% CI (dark gray shading), and the CMIP6 MME 90% CI (light gray shading).

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

	Median 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 $[\frac{K}{Wm^{-2}}]$	λ 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 ($H = \frac{1}{2}$) and the Scaling CRF model (Hébert et al., 2020) calibrated over the historical period, along with their corresponding 90% confidence intervals.

	Median 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

405   With the full suite of model and forcing parameters, we now use the FEBE to reconstruct the forced temperature response over the historical period and make projections 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.

410   From the multidimensional parameter space we draw from the multinormal distribution (eq. 18) using a Monte Carlo method to generate realizations of the forced temperature response (eq. 3) using a numerical convolution. Once we have our ensemble of projections, we remove the pre-industrial baseline (1880-1910) and calculate the desired confidence intervals of the forced response, the historical reconstructed period (hindprojections) presented in fig. 9 and projections to 2100 in fig. 10.

   Throughout 1880-2005 (the overlapping observational and CMIP5 MME period), the reconstructed forced temperature
415 response from the FEBE, using $F_{Aer_{RCP}}$, and the mean of the CMIP5 MME are close, seen in fig. 9). The spread between the forced temperature response and the five observational datasets decreases as we move closer to the present. Between 1915-1960 period the CMIP5 MME is consistently warmer than the historical reconstruction 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)), is tracked closely by the
420 FEBE reconstruction while the CMIP5 MME overshoots (by $0.1K$ to $0.2K$), a well studied divergence between GCMs and observations (fig. 9 lower plot). This supports (Lovejoy, 2015) who found that the hiatus could be well predicted by a stochastic fGn model (comparable with the present hindprojection) and concluded that the problem was GCM overprojection.
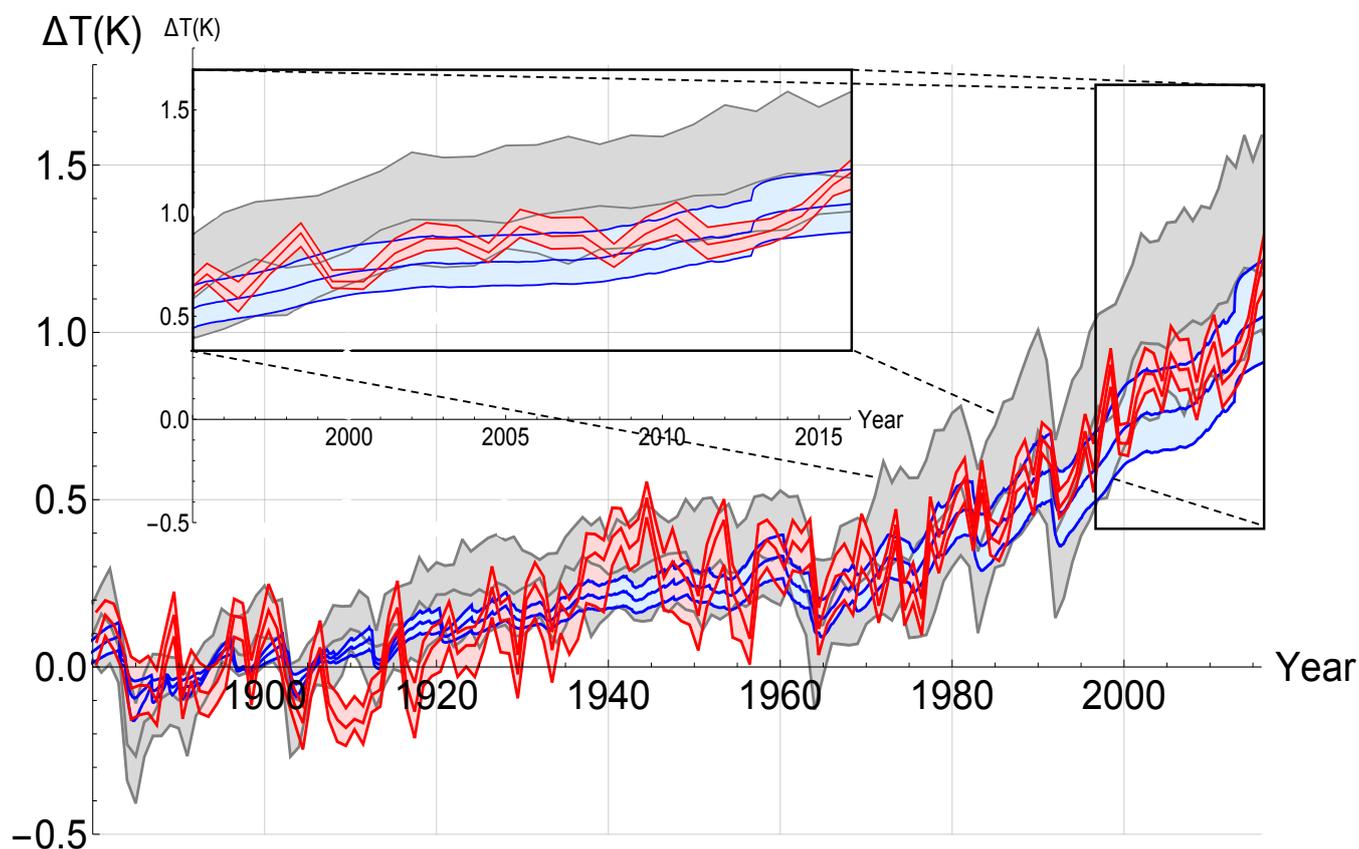
Figure 9. The reconstruction of the forced temperature response over the historical period (1880-2019) using the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue), compared with the CMIP5 MME projection (black) and mean of 5 observational temperature series (red); 90% CI are indicated (shaded). 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.

4.2 Projections through to 2100

The future pathways only begin to diverge into their respective scenarios roughly two decades after their beginning (RCPs
425 begin in 2000, SSPs begin in 2010), so at first the temperature increase in each case is nearly identical. 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) while significantly lower in RCP 2.6/SSP 126 (fig. 10 top), 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, demonstrating the potential of strong mitigation policies and speculative negative emission technologies where anthropogenic forcing starts
430 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, at 2100, the FEBE projection reaches $1.2K$ 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 11, top) 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,
435   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, stabilizes in the mid 2060s, the temperature projections continue rising throughout the 21st century for both FEBE and the CMIP5 MME (fig. 10 middle). At 2100 the FEBE and CMIP5 MME projects a temperature reaching $1.9K$ $[1.6,2.2]K$, and $2.6K$ $[1.8, 3.2]K$ respectively. A key point to
440   note is that the FEBE RCP 4.5 projection remains below $2K$ 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.

The projections of both the FEBE and the CMIP5 MME for the business as usual 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). The same quickly increasing trend is seen in the CMIP6 SSP 585 scenario with temperatures in 2100 reaching a staggering $6.2K$ with
445   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.

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 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
450   the historical hindprojections of FEBE. 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. 1) at 2100 between $F_{Aer_{RCP}}$ and $F_{Aer_{SSP}}$ in the strong mitigation scenario (RCP 2.6/SSP 126) is $0.6Wm^{-2}$. With the total GHG forcings being nearly identical, the the SSP 126 scenario is closer to
455   a RCP 3.3 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, 11) 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 FEBE 90% CI lies entirely within the corresponding CMIP5 CI. Being in complete agreement, the two projections methods mutually support each other
460   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 their overpowered aerosols ((Zelinka et al., 2020), (Flynn and Mauritsen, 2020)) and previously mentioned removal of future aerosol forcing.
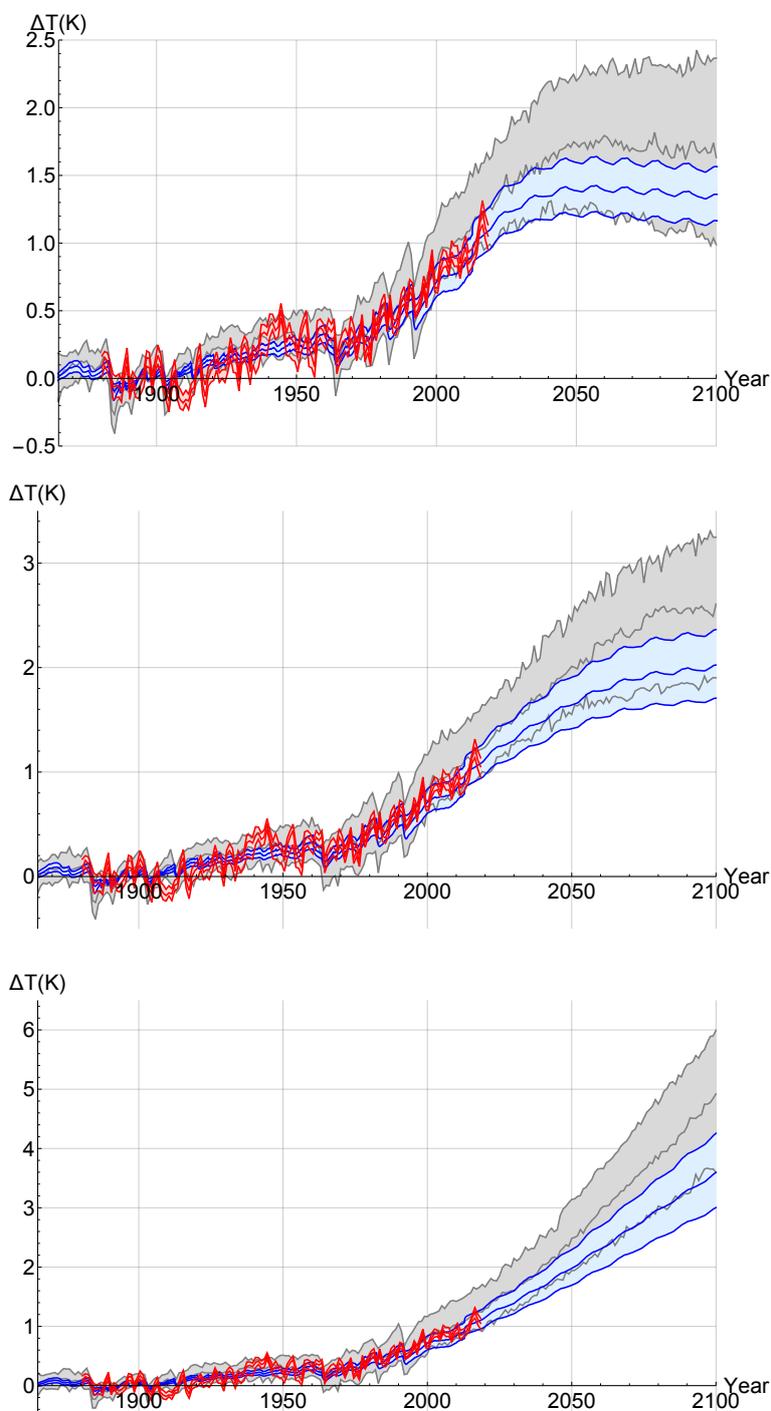
Figure 10. The forced temperature response projected using the FEBE, with parameters calibrated using $F_{Aer_{RCP}}$ (blue) compared with the CMIP5 MME projection (black); 90% CI are indicated (shaded). The projections until 2100, for RCP 2.6 (top), RCP 4.5 (middle) and RCP 8.5 (bottom), are shown.
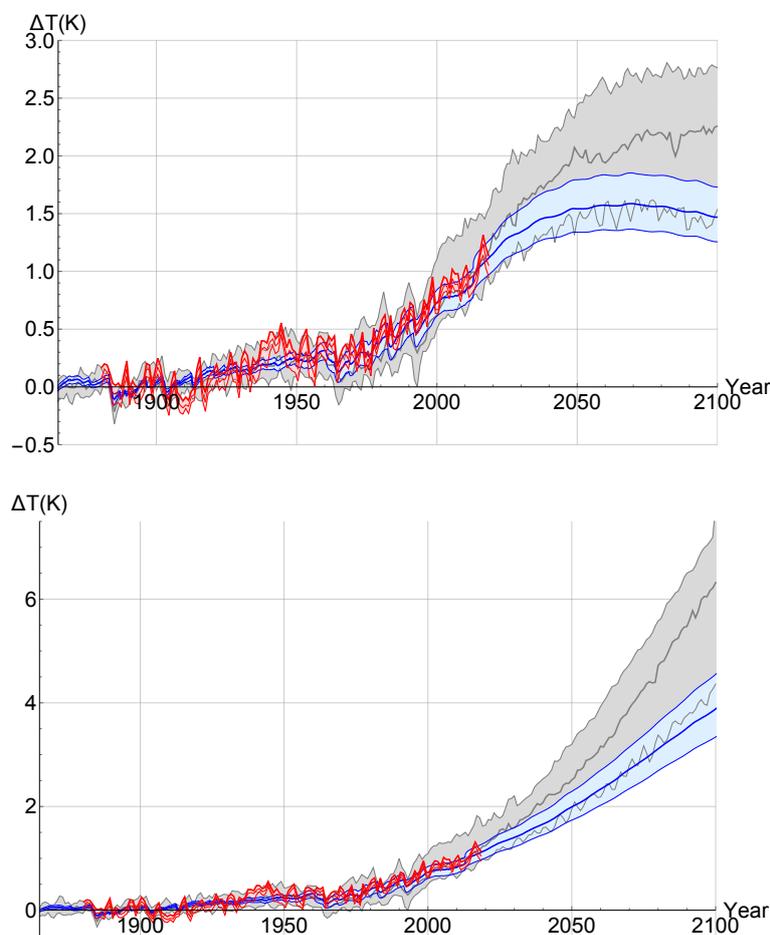
Figure 11. The forced temperature response projected using the FEBE, with parameters calibrated using $F_{Aer_{SSP}}$ (blue), compared with the CMIP6 projection (black); 90% CI are indicated (shaded). The projections until 2100, for SSP 126 (top), and SSP 585 (bottom), are shown.

((Schurer et al., 2017), (Smith et al., 2018b), (Iseri et al., 2018)) have given evidence that there are important tipping points which could lead to irreversible changes in major ecosystems and the planetary climate if certain threshold in warming are exceeded. Therefore we calculate the probability of temperature exceeding $1.5K$ and $2.0K$ for both the FEBE and the CMIP5/6 MME shown in figs. 12 and 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. While in the SSP 126 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 it occurs much later than that projected by the CMIP5 MME: occurring 20 years after the GCMs: 2070. 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 $2K$.

For the final high emission, business as usual, RCP 8.5 and SSP 585 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.
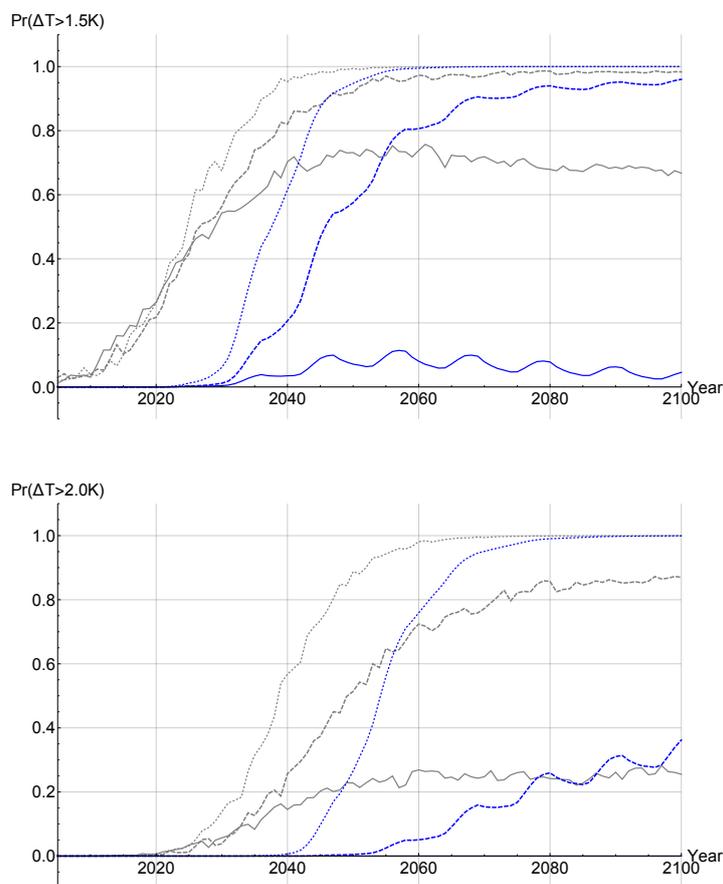
Figure 12. The probability for the global mean surface temperature of exceeding a $1.5K$ threshold (top), and a $2K$ (bottom) are given as a function of years for the FEBE, using $F_{Aer_{RCP}}$ (blue) and for the CMIP5 MME (black). The three RCP scenarios are considered for each case: RCP 2.6 (solid), RCP 4.5 (dashed), and RCP 8.5 (circles).
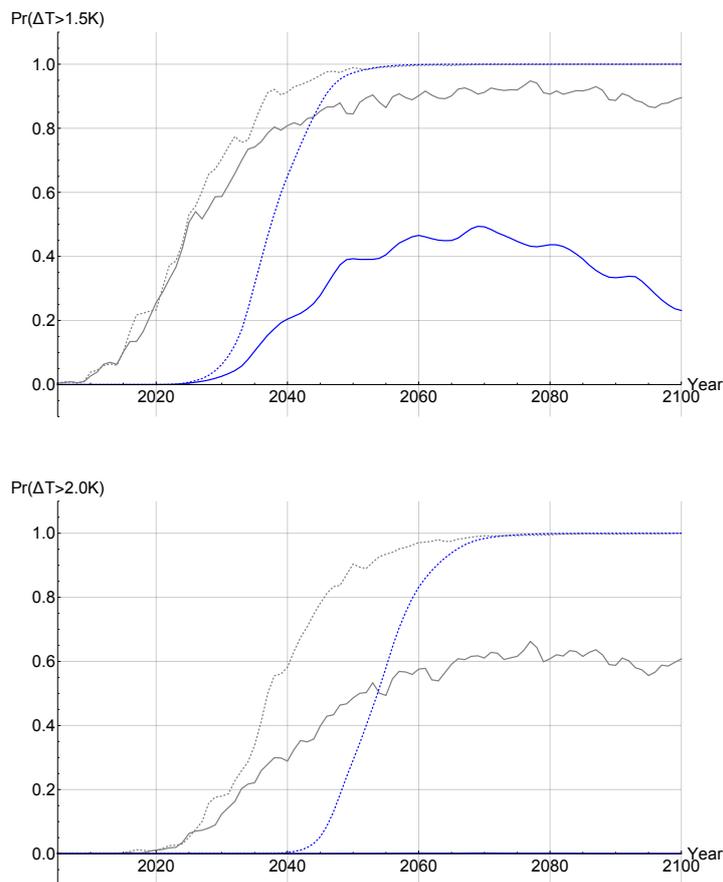
Figure 13. The probability for the global mean surface temperature of exceeding a $1.5K$ threshold (top), and a $2K$ (bottom) are given as a function of years for the FEBE, using $F_{Aer_{SSP}}$ (blue), and for the CMIP6 MME (black). The two SSP scenarios are considered for each case: SSP 126 (solid), and SSP 585 (circles).

5   Conclusions

Multidecadal projections have large uncertainties with ECS confidence limits of 1.5-$4.5K$ essentially unchanged for decades. For policy makers, the most deleterious consequence 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
485 - albeit with rather different probabilities. Large overlaps imply a disconnection between policies (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.

One way of reducing this uncertainty is by developing complementary types of models. In this paper we directly constructed a model in the macroweather regime (roughly one month and up) based on the physically principles of energy conservation

490 and scaling: the Fractional Energy Balance Equation (FEBE). Although originally derived phenomenologically, it was recently discovered ((Lovejoy, 2020b), (Lovejoy, 2020c)) that the FEBE could be derived as a consequence of classical (Budyko, 1969) and (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

495 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 H, and relaxation time scale τ; the classical

500 EBE is the $H = 1$ special case. In order to make FEBE projections, we use a Bayesian parameter estimation approach. (Hébert et al., 2020) 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., 2020) 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

505 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 pre-

510 scribed 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 re-

515 calibration 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., 2020), (Zelinka et al., 2020), (Tokarska et al., 2020), (Flynn and Mauritsen, 2020)). 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.

520 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,-0.2]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

525 median 0.28 with similar 90% CI of [0.16, 0.40] using $F_{Aer_{SSP}}$.

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., 2020) and the phenomenological HEBE ((Lovejoy, 2020b),

530 (Lovejoy, 2020c)). 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 ((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

535 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, 4.5]K ([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). We also estimated the derived quantity, the Transient Climate Response (TCR), the temperature increase following a linear doubling in

540 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). 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

545 did this for RCP 2.6 (SSP 126), 4.5, and 8.5 (SSP 585). 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 differences in the climates of the GCMs ("structural uncertainties"), the FEBE uncertainty 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 frame-

550 work 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.

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

555 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 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 $1.5K$

560   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

565   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, 2020c), which hopefully will constrain and improve future projections. Once the full suite of CMIP6 projections are

released, the FEBE could be also used to help understand the generational differences between CMIP models. Extensions to

570   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).

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. 11

Model	ECS (K)	TCR (K)	TCR-to-ECS ratio
MIROC6*	2.60	1.58	0.61
IPSL-CM6A-LR*	4.50	2.39	0.53
CNRM-CM6-1*	4.82	2.23	0.46
BCC-CSM2-MR*	3.07	1.60	0.52
MRI-ESM2*	3.11	1.67	0.54
CanESM5*	5.58	2.75	0.49
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*	5.31	2.79	0.53
CNRM-ESM2-1*	4.75	1.82	0.38
BCC-ESM1	3.29	1.77	0.54
CESM2-WACCM*	4.65	1.92	0.41
MIROC-ES2L	2.66	1.51	0.57
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*	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 ± Std:	3.74±1.11	1.98±0.48	0.55 ± 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 ± Std:	3.20±0.70	1.75±0.38	0.56 ± 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 and L. Del Rio Amador for helpful discussions.

References

Annan, J. D. and Hargreaves, J. C.: Using multiple observationally-based constraints to estimate climate sensitivity, Geophysical Research Letters, 33, https://doi.org/10.1029/2005GL025259, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GL025259, 2006.

585 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.

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

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, 2014.

Cox, P. M., Huntingford, C., and Williamson, M. S.: Emergent constraint on equilibrium climate sensitivity from global temperature vari-
595   ability, Nature, 553, 319–322, 2018.

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.

600 Del Rio Amador, L. and Lovejoy, S.: Predicting the global temperature with the Stochastic Seasonal to Interannual Prediction System (Stoc-SIPS), 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.

605 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.

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.

610 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.

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

Hébert, R. and Lovejoy, S.: Interactive comment on "Global warming projections derived from an observation-based minimal model" by K.
615   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., Lovejoy, S., and Tremblay, B.: An Observation-based Scaling Model for Climate Sensitivity Estimates and Global Projections to 2100, Climate Dynamics, 2020.

Held, I. M., 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.

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.

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.

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.: 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, 2015.

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, 2020a.

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

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

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., Hebert, R., and Del Rio Amador, L.: The Fractional Energy Balance Equation, Quart. J. of the Roy. Met. Soc., http://www.physics.mcgill.ca/~gang/eprints/eprintLovejoy/esubmissions/FEBE.physics.3.6.20bis.pdf, 2020.

660 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,
665 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.

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
670 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
675 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,
680 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
685 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.

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

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.

695   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.

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.

700   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.

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.

Rypdal, K.: Global temperature response to radiative forcing: Solar cycle versus volcanic eruptions, Journal of Geophysical Research: At-
705   mospheres, 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.

710   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,
715   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.

Skeie, R. B., Berntsen, T., Aldrin, M., Holden, M., and Myhre, G.: A lower and more constrained estimate of climate sensitivity using
720   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.

725   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/
730   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., Plattner, G.-K., and Friedlingstein, P.: Irreversible climate change due to carbon dioxide emissions, Proceedings of the National Academy of Sciences, USA, 2009.

735 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.

740 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.

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.

745 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.

750 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.