Climate Dynamics
https://doi.org/10.1007/s00382-020-05521-x

An observation-based scaling model for climate sensitivity estimates and global projections to 2100

Raphaël Hébert[1] · **Shaun Lovejoy**[2] · **Bruno Tremblay**[3]

Received: 30 May 2017 / Accepted: 3 November 2020
© The Author(s) 2020

Abstract
We directly exploit the stochasticity of the internal variability, and the linearity of the forced response to make global temperature projections based on historical data and a Green's function, or Climate Response Function (CRF). To make the problem tractable, we take advantage of the temporal scaling symmetry to define a scaling CRF characterized by the scaling exponent H, which controls the long-range memory of the climate, i.e. how fast the system tends toward a steady-state, and an inner scale $\tau \approx 2$ years below which the higher-frequency response is smoothed out. An aerosol scaling factor and a non-linear volcanic damping exponent were introduced to account for the large uncertainty in these forcings. We estimate the model and forcing parameters by Bayesian inference which allows us to analytically calculate the transient climate response and the equilibrium climate sensitivity as: $1.7^{+0.3}_{-0.2}$ K and $2.4^{+1.3}_{-0.6}$ K respectively (likely range). Projections to 2100 according to the RCP 2.6, 4.5 and 8.5 scenarios yield warmings with respect to 1880–1910 of: $1.5^{+0.4}_{-0.2}K$, $2.3^{+0.7}_{-0.5}$ K and $4.2^{+1.3}_{-0.9}$ K. These projection estimates are lower than the ones based on a Coupled Model Intercomparison Project phase 5 multi-model ensemble; more importantly, their uncertainties are smaller and only depend on historical temperature and forcing series. The key uncertainty is due to aerosol forcings; we find a modern (2005) forcing value of $[-1.0, -0.3]$ Wm^{-2} (90 % confidence interval) with median at -0.7 Wm^{-2}. Projecting to 2100, we find that to keep the warming below 1.5 K, future emissions must undergo cuts similar to RCP 2.6 for which the probability to remain under 1.5 K is 48 %. RCP 4.5 and RCP 8.5-like futures overshoot with very high probability.

Keywords Global mean temperature · Projections · Climate sensitivity · RCP scenarios · Global warming · Scaling

1 Introduction

The atmosphere is a complex system involving turbulent processes operating over a wide range of scales starting from millimeters at the Kolmogorov dissipation scale up to the size of the Earth, spanning over 10 orders of magnitudes in space. The dynamics are sensitive to initial conditions and there are deterministic predictability limits that are roughly equal to the eddy turn-over time (lifetime) of structures. For planetary scale structures in the atmosphere, the overall deterministic prediction limit of about 10 days corresponds to the scaling transition timescale τ_w from the *weather regime* to the *macroweather regime* (Lovejoy and Schertzer 2013a).

The atmospheric components of GCMs exhibit the same weather-macroweather scaling transition as the atmosphere and similar predictability limits. Beyond this horizon, the internal variability has to be interpreted stochastically so that a single GCM run is only one realization of the random

✉ Raphaël Hébert
raphael.hebert@awi.de

Shaun Lovejoy
lovejoy@physics.mcgill.ca

Bruno Tremblay
bruno.tremblay@mcgill.ca

[1] Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Telegrafenberg A45, Potsdam 14473, Germany

[2] Department of Physics, McGill University, 3600 rue University, Montréal, Québec H3A 2T8, Canada

[3] Department of Atmospheric and Oceanic Sciences, McGill University, 845 rue Sherbrooke Ouest, Montréal, Québec H3A 0G4, Canada

Published online: 18 December 2020

process; at these timescales, weather models effectively become stochastic macroweather generators. For projections over multi-decadal timescales and beyond, multi-model ensembles (MME) that include several models are used. The mean of the MME is taken to obtain the deterministic forced component of temperature variability and average out the internal variability (Collins et al. 2013).

Emergent properties of the Earth's climate, i.e. properties which are not specified a priori, are then inferred from GCM simulations. The *equilibrium climate sensitivity* (ECS) is such a property; it refers to the expected temperature change after an infinitely long time following a doubling in carbon dioxide (CO_2) atmospheric concentration. Another is the *transient climate response* (TCR), which is defined as the change in temperature after a gradual doubling of CO_2 atmospheric concentration over 70 years at a rate of 1% per year. However, it is not clear whether such emergent properties from computational models can be taken as genuine features of the natural world. The difficulty is that each GCM has its own climate ("structural uncertainty") and this leads to very large discrepancies in ECS and TCR between GCMs; this underscores the need for qualitatively different approaches which can narrow down the properties of the real climate directly from observations.

The ecological consequences of global warming could be dire; therefore, better constraining climate sensitivity is of utmost importance in order to meet the urgency of adjusting economical and environmental policies. Since the United States National Academy of Sciences report (Charney et al. 1979), the likely range for the ECS has not changed and remains $[1.5, 4.5]\,K$. A likely range corresponds to a 66% confidence interval (CI) and a very likely range corresponds to a 90% CI (Mastrandrea et al. 2010). In the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC), the lower limit was revised upward by 0.5 K, in accordance with the very likely range for ECS of GCMs from the Coupled Model Intercomparison Project phase 3 (CMIP3) ([2.1, 4.4] K). In the Fifth Assessment Report (AR5), the lower limit returned to that of AR3 and earlier assessment reports because of new observation-based results, while the very likely range of CMIP phase 5 (CMIP5) GCMs ([1.9, 4.5] K) remained very close to the CMIP3 one.

In this paper, we extend the approach of Hébert and Lovejoy (2018) to make climate projections through 2100. The approach is based on historical data and a simple model of the system memory based on scaling symmetries. The output of our model is then evaluated against the instrumental record using Bayes' rules in order to obtain a probabilistic estimate of its parameters.

The paper is structured in 3 sections : methods and material, results and conclusion. The methods and material section is divided in three sub-sections. First, we introduce the linear response framework and then describe the scaling climate response function considered. Secondly, the radiative forcing, temperature and GCM simulations used will be described, and thirdly, we explain the method used for the estimation of the model parameters. The results section is also divided into three parts. The first sub-section presents the probability distribution functions for the parameters and applies them to decompose the anthropogenic and natural forced signals, and the internal variability. The second subsection estimates the ECS and TCR with the parameters found, and the third uses the same parameters to produce global projections to 2100 which are better constrained than a 32 CMIP5 GCMs MME with which they are compared.

2 Methods and material

2.1 The linear response framework

The approach used in this study builds on the work of Hasselmann and other authors who worked within the linear response framework applied to the climate (Budyko, Sellers, Schwarz, Li and Jarvis, Held et al., Von Hateren, Rypdal and Rypdal, Dijkstra, Geoffroy et al., Marshall et al.). Below we first provide a review of this work for context. The reader solely interested in the current approach can jump to Sect. 2.2 without loss of continuity.

The internal components of the Earth system are often far from thermodynamic equilibrium, yet, taken as a whole, the Earth is not so far from an energy balance with outer space, and, at any moment, the difference between the incoming and outgoing energy fluxes is stored in the soil, ocean and atmosphere. The deviations from energy balance are typically small—at the level of a few percent—and this justifies the linear response framework. In this paper, we consider "zero-dimensional" energy balance models in which the Earth's globally averaged temperature $T(t)$ and forcing $F(t)$ are anomaly time series, i.e. they represent deviations from their reference values.

The earliest linearized temperature response models were the Budyko-Sellers energy balance models based on the heat equation (Budyko 1969; Sellers 1969). These were originally one dimensional (zonally averaged) models, which, when globally averaged, are equivalent to the single "box" model (Hasselmann et al. 1993). Global energy balance box models are models of the temperature based on a homogeneous "box". The box has a spatially uniform temperature that stores energy according to its heat capacity, density and size. If there is a single box, and one asssumes Newton's law of cooling and that the heat crossing the surface is proportional to the first derivative of the order differential relationship with

temperature, then, when perturbed, the Earth's temperature will relax in an exponential way to its new steady-state temperature. When extra boxes are added, they mutually exchange heat, leading to a total response that is the sum of exponentials.

Hasselmann et al. (1993, 1997) already noted that it was desirable to use the more general linearized framework of response functions. This, they argued, was because empirical box models with a small number of degrees of freedom "lose the detailed information on the climate state and therefore cannot be readily constrained to conform to the detailed linearized dynamics of a more realistic CGCM climate model." In this context, the response functions are called "climate response functions" (CRFs) and, following Hasselmann et al. (1993), we have a choice between the equivalent impulse CRFs $G_\delta(t)$ and step CRFs $G_\Theta(t)$. The subscript "δ" indicates Dirac delta function and "Θ" is for its integral, the step (Heaviside) function ($\Theta = 1$ for $t > 0$ and $\Theta = 0$ for $t \leq 0$).

Hasselmann et al. (1993) and especially Hasselmann et al. (1997) already pointed out the advantages of the step CRF that relates the forcing and temperature via:

$$T(t) = s \int_{-\infty}^{t} G_\Theta(t - t')F'(t')dt', \tag{1}$$

where the $F'(t)$ is the time derivative of the forcing $F(t)$ and s is the equilibrium climate sensitivity (ECS) with units of K per doubling of CO_2 (see Eq. 15 below). While Hasselmann et al. (1993) incorporated the sensitivity into the definition of $G_\Theta(t)$, writing the response with the separate factor s has the advantage that $G_\Theta(t)$ is dimensionless and s and $F(t)$ have their usual dimensions. For generality, we have also extended the range of integration to cover the entire past. In advocating Eq. 1, Hasselmann et al. (1997) pointed out that "The formulation of the climate response in terms of a response integral (i.e. step response) rather than in the traditional form of a differential equation for a box model has further advantages: it is not limited to simple low-order differential equations...". Another advantage, later emphasized by Marshall et al. (2014, 2017),) is that, physically, basing the theory on $G_\Theta(t)$ is equivalent to studying the temperature response in classical CO_2 doubling experiments; other advantages are discussed below. The equivalence between the impulse and step CRFs arises because the step function $\Theta(t)$ is the integral of $\delta(t)$ so that:

$$G_\delta(t) = G'_\Theta(t) \tag{2}$$

The temperature response in terms of $G_{\delta(t)}$ rather than $G_{\Theta(t)}$ can thus easily be obtained by integrating Eq. 1 by parts to yield:

$$T(t) = s \int_{-\infty}^{t} G_\delta(t - t')F(t')dt' \tag{3}$$

In the integration by parts, we used Eq. 2 and the boundary conditions $G_\delta(0) = 0$ and $F(-\infty) = 0$. Since causality requires $G_\Theta(t) = G_\delta(t) = 0$ when $t \leq 0$, the former condition $G_\delta(0) = 0$ is satisfied by physical systems. Similarly, the relation $F(-\infty) = 0$ is not a restriction as it can be regarded simply as the definition of a convenient reference level of the forcing.

Unfortunately, without more assumptions or information, the linear framework of Eq. 1 (or Eq. 3) is unmanageably general. In order to make progress, Hasselmann et al. (1997) proposed a response function consisting of a sum of N exponentials - effectively an N box model (although without using differential equations: the boxes were only implicit). Nevertheless, they ultimately chose $N = 3$ out of practical necessity—so as to fit GCM outputs. Following the more usual procedure of deriving the impulse responses from linear differential equations (where impulse CRFs are called "Green's functions"), Li and Jarvis (2009) used Laplace transforms to explicitly show that polynomial forcings of nth ordered differential equations (with constant coefficients and with n an integer), can quite generally be reduced to sums of exponentials. However, in the application part of their paper, they nevertheless used the value N = 3. The N exponential model was later advocated by van Hateren (2013), by the IPCC AR5 (2013, section 8.SM.11.2), and more recently by Frederiksen and Rypdal (2017). However, each exponential has its own amplitude and time constant so that even if we exclude the climate sensitivity parameter, the N exponential models have 2N parameters. This rapidly becomes an unmanageably large number so that in practice $N = 2$ or $N = 3$ are often chosen.

An interesting exception is van Hateren (2013) who used $N = 6$, but avoided fitting the implicit $2N = 12$ independent parameters by linking the amplitudes and time constants by a power law, calling the resulting four parameter model a "fractal climate response" model. His resulting step CRF model specifies four parameters: low and high frequency truncations, a logarithmic oscillation frequency and an overall scaling exponent. In the model we develop below, we eliminate the unnecessary oscillations and low frequency cutoff, thus reducing the step CRF to only two parameters. Using van Hateren's empirically fitted parameters yields an impulse response that over the range of about 6 months to 1000 years is within a factor of ≈ 2 of our two-parameter model described below. For comparison, the key exponent H (denoted by q in his notation) was estimated as -0.15 compared to our value of $-0.5^{+0.4}_{-0.5}$ (below).

An approach related to the scaling CRF we develop below is that of Procyk et al. (2020) which is based on the Earth energy balance. A fractional generalization of conventional box models was considered and solved using Mittag–Leffler functions, often called generalized

exponentials, also characterized by a scaling exponent H which was estimated as $H \in [0.33, 0.44]$ and is equivalent to the negative of the SCRF's H (below).

Although these authors proposed exponentials largely on mathematical grounds, the majority of linear response theory applications attempt to give physical interpretations of their parameters, especially their time constants, and these have not been very satisfactory. If each exponential can be modelled by a box that effectively stores heat, then it is not clear what the box should represent physically. If one chooses the atmosphere (e.g. Dijkstra 2013), then one obtains a short relaxation time τ of the order of days, whereas if one chooses the ocean, then a wide range of time scales can be obtained depending on the thickness of the relevant ocean layer.

Several estimates of the fast τ, which would correspond to the rapidly equilibriating mixed-layer of the ocean, find values below 10 years: $\tau = 8.5 \pm 2.5$ years (Schwartz 2008), $\tau \approx 4$ years (Held et al. 2010), $\tau \in [1, 6]$ years (Geoffroy et al. 2013), $\tau = 4.3 \pm 0.6$ years (Rypdal and Rypdal 2014). The estimates of the slow component, which would correspond to the deep ocean, are widely divergent and Geoffroy et al. (2013) concludes it should be between 60 and 700 years, an interval spanning an order of magnitude well above what can be directly probed from historical observations. The IPCC AR5 suggests a double exponential model with $\tau = 8.4$ years and $\tau = 409.5$ years.

But the box models are overly simple: in reality, the earth is highly heterogeneous with dynamical processes redistributing energy over a huge number of degrees of freedom, and a multitude of storage mechanisms, covering a wide ranges of scales. What we really need is a phenomenological model that is approximately valid for the globally averaged temperature, an equation that is valid at time scales longer than the weather scales (about 10 days). Once we accept that our equation is at best valid for averages (globally and over weather scales), it is no longer necessary to constrain the model to a single—or even a small number of degrees of freedom—slabs or boxes.

When $F(t)$ is a step function, Eqs. 1 and 3 describe a linear system that is perturbed by a forcing and that subsequently relaxes to a new steady-state temperature. The problem is that the range of time scales over which the system reacts to a forcing is huge; indeed, due to scaling symmetries, as recognized by Rypdal (2012) and van Hateren (2013), the response is closer to a power law. We therefore seek to go beyond exponentials, while still restricting our attention to CRFs that correspond to processes which can relax to a stable state of energy balance. To see what constraints such "generalized" relaxation imposes, the step CRF is particularly convenient.

For example, for a physical system, a finite step forcing:

$$F(t) = F_0 \Theta(t); \quad \Theta(t) = \begin{cases} 1 \text{ when } t \geq 0 \\ 0 \text{ when } t < 0 \end{cases} \tag{4}$$

must give a finite response. Since in Eq. 1 we included the extra sensitivity factor s, without loss of generality we can consider only normalized step CRF such that:

$$lim_{t \to \infty} G_\Theta(t) = 1 \tag{5}$$

Since sF_0 is the new steady-state temperature, s is the usual ECS. In addition, G_Θ should be constrained to functions such that a steady state is established monotonically; we should exclude step responses that oscillate or that overshoot the steady state before returning to it:

$$G'_\Theta(t) = G_\delta(t) > 0 \tag{6}$$

In addition, the response to a positive forcing should be positive so that combining these constraints, and using Eqs. 5, 6 and causality ($G_\Theta(0) = 0$) we obtain :

$$0 \leq G_\Theta(t) \leq 1 \tag{7}$$

Systems whose step CRFs respect Eqs. 6 and 7 thus define physically plausible generalized relaxation processes. As an example, the classical relaxation box/exponential model yields:

$$\begin{aligned} G_{\Theta,box}(t) = 1 - e^{\frac{-t}{\tau}}; \quad G_{\delta,box}(t) = \tau^{-1} e^{\frac{-t}{\tau}}; \quad t > 0 \\ G_{\Theta,box}(t) = G_{\delta,box}(t) = 0; \quad t \leq 0 \end{aligned} \tag{8}$$

where τ is the "relaxation time", the characteristic time associated with the return to a state of energy balance. We see that the approach to the steady-state is exponentially fast. Although the box model is usually specified via a differential equation, from the above, we see that it could equivalently be specified by $G_{\Theta,box}(t)$; indeed, it is easy to verify that $G_{\Theta,box}(t)$ satisfies the standard box-model relaxation equation:

$$\tau \frac{dG_{\Theta,box}}{dt} + G_{\Theta,box} = \Theta(t) \tag{9}$$

Taking derivatives, we also confirm that $G_{\delta,box}(t) = G'_{\Theta,box}(t)$ is indeed the impulse response for the operator.

Other simple CRF's have been proposed, notably the Dirac function itself with a lag t_0:

$$G_\Theta(t) = \Theta(t - t_0); \quad G_\delta(t) = \delta(t - t_0) \tag{10}$$

Which implies:

$$T(t) = sF(t - t_0) \tag{11}$$

This has been used by Lean and Rind (2008) (with $t_0 = 10$ years as part of multiple regression), and directly by Lovejoy (2014) who varied t_0 in the range 0–20 years. In both cases,

the forcing was at annual or longer temporal resolution so the response was not really instantaneous as Eq. 10 implies.

2.2 Scaling relaxation processes

The Dirac CRF has essentially no memory, and the box model has an exponentially decaying one, neither is satisfactory, unless they are combined into a model with multiple response times. To be more realistic, the CRF should satisfy scaling symmetries (respected for example by the GCMs in control runs); we should therefore choose functions that correspond to scaling (power law) relaxation processes. The simplest scaling CRF (SCRF) that satisfies Eqs. 5 and 6 is:

$$G_\Theta(t) = \begin{cases} 1 - \left(1 + \frac{t}{\tau}\right)^H & \text{when} \quad t \geq 0 \\ 0 & \text{when} \quad t < 0 \end{cases} \tag{12}$$

where the requirement $H < 0$ is needed so that $\lim_{t\to\infty} G_\Theta(t) = 1$ and the truncation at a small timescale τ is necessary so that $G_\Theta(0) = 0$. This step SCRF describes a power law relaxation process, thus converging more slowly, and realistically, to a steady-state than typical exponential models (Fig. 1), with scaling exponent H ($H < 0$) and a corresponding impulse SCRF:

$$G_\delta(t) = \begin{cases} -\frac{H}{\tau}\left(1 + \frac{t}{\tau}\right)^{H-1} & \text{when} \quad t \geq 0 \\ 0 & \text{when} \quad t < 0 \end{cases} \tag{13}$$

so that the impulse CRF is a also a truncated power law.

Rypdal (2012) already proposed a similar CRF with $H > 0$, which has the advantage of not needing the truncation τ at small time scales. This allows the modelling of the high-frequency with the same scaling by the simple addition of a random white noise forcing to the deterministic forcing. This came at the expense of divergence at large time scales, the *runaway Green's function effect* (Hébert and Lovejoy 2015), since any finite increase in forcing would lead to an ever increasing temperature response, i.e. an infinite ECS. An interesting alternative to model high- and low-frequency regimes with scaling regimes can be obtained by the fractional generalization of the energy-balance equation, leading to a CRF with a similar (convergent) low-frequency behaviour as the SCRF (Procyk et al. 2020).

It is straightforward to analytically calculate the expected temperature increase using the truncated SCRF in Eq. 13 in response to specific forcing scenario such as to recover TCR and ECS (See Appendix A for details). With the same forcing scenario as ECS, we also define ECS_{500} as the expected temperature 500 years after the CO_2 doubling rather than at infinity. This is a more relevant ECS measure from a human perspective and helps to illustrate the contribution to the ECS of the very long-memory beyond 500 years. The ratio of TCR to ECS (Eq. A10) changes from unity for $H \to -\infty$ to zero for $H \geq 0$ where ECS diverges while TCR remains finite; conversely, the fraction of warming left between 500 years and infinity ($(ECS - ECS_{500})ECS^{-1}$) goes from zero to unity when H goes from negative infinity to zero and greater (Fig. 2).

2.3 Data

2.3.1 Radiative forcing data

In this paper, we consider three sources of external forcing: solar and volcanic which are natural, and anthropogenic forcing which involves several forcing agents produced

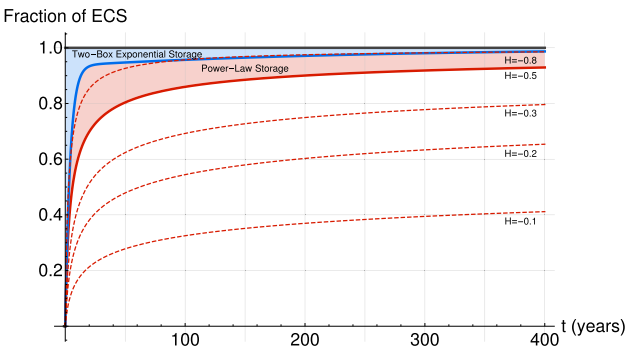

Fig. 1 A nondimensional comparison of the step SCRF (Eq. 12, red) with the classical two-box exponential step CRF (Eq. 8, blue). The nondimensional (step) forcing (black) is also shown, and the difference between this forcing and the response is the rate of energy storage. The parameters for the two-box exponential are the best estimates from Geoffroy et al. (2013): $\tau_{fast} = 4.1\,years$, $\tau_{slow} = 249\,years$, $C = 7.3\,year\,Wm^{-2}K^{-1}$ and $C_0 = 106\,year\,Wm^{-2}K^{-1}$ while for the SCRF we use $\tau = 2\,years$ and show the curves for different H values indicated on the graph

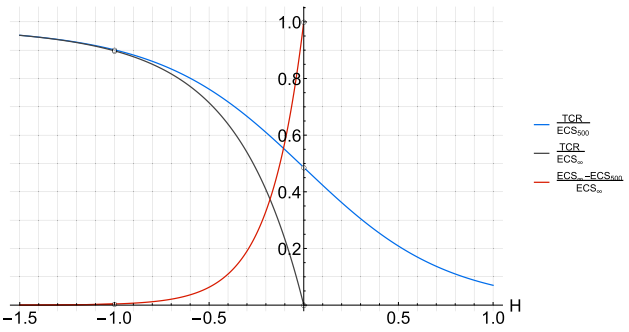

Fig. 2 The analytical ratio between TCR and ECS is shown as a function of the scaling exponent H for a high frequency cutoff $\tau = 2\,years$. The ratio of TCR to ECS_{500} is also shown for equilibrium defined at 500 *years* (black) along with the true equilibrium at infinity (blue) and the leftover warming fraction (red) between 500 years and infinity

Table 1 Summary of the effective radiative forcing increase from 1750 to 2005 for different anthropogenic sources as reported in the IPCC AR5

Forcing sources considered	Lower 5%	Median	Upper 5%
CO_2	0.40	0.45	0.50
All GHG	0.64	0.71	0.78
Aerosol direct effects	−0.24	−0.13	−0.02
Aerosol indirect effects	−0.49	−0.19	−0.08
Total anthropogenic forcing	0.21	0.43	0.55
Forcing sources neglected	**Lower 5%**	**Median**	**Upper 5%**
Tropospheric ozone	0.06	0.11	0.16
Stratospheric ozone	−0.04	−0.01	0.02
Stratospheric water vapour from NH_4	0.01	0.02	0.03
Land use changes	−0.07	−0.04	−0.01
Surface albedo from black carbon	0.00	0.03	0.06
Contrails	0.002	0.003	0.11

Given are the median, the upper 5% bound and the lower 5% bound in units of $\Delta F_{2\times\,CO_2}$

by humans. The forcing is usually expressed in W m^{-2}; however the climate sensitivity is commonly measured in K per doubling of CO_2. Therefore, it is convenient to also define forcing as a fraction of the forcing imparted by a doubling of CO_2 concentration: $\Delta F_{2\times\,CO_2}$. The generally accepted (approximate) carbon dioxide concentration to forcing relationship is:

$$F_{CO_2}(\rho) = 3.71\ \text{W m}^{-2} \log_2 \frac{\rho}{\rho_0} \tag{14}$$

where F_{CO_2} is the forcing due to carbon dioxide, ρ is the carbon dioxide concentration and ρ_0 is its pre-industrial value which we take to be 277 *ppm*. Therefore,

$$\Delta F_{2\times\,CO_2} = F_{CO_2}(2\rho_0) = 3.71\ \text{W m}^{-2} \tag{15}$$

a) Greenhouse Gas Forcing

Anthropogenic influences on the climate have been recognized as the main driver of the global warming characteristic of the last century, and the related forcing is mostly due to historical changes in atmospheric composition. Future anthropogenic forcing is prescribed in four scenarios, the Representative Concentration Pathways (RCPs), established by the IPCC for CMIP5 simulations : RCP 2.6, RCP 4.5, RCP 6.0 and RCP 8.5 (Meinshausen et al. 2011), shown in Fig. 3. They are named according to the total radiative forcing in Wm^{-2} expected in the year 2100 and are motivated by complex economic projections, expected technological developments, and political decisions. The scenarios allow us to verify and compare results from our observations-based SCRF model with CMIP5 simulations. Generally, RCP 6.0 was left out of the analysis since fewer CMIP5 modeling groups performed the associated runs.

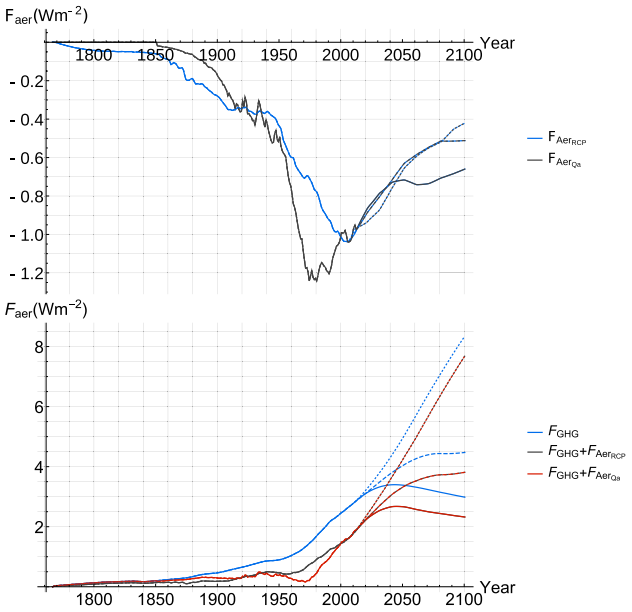

Fig. 3 (top) The anthropogenic aerosol forcing series used, $F_{Aer_{RCP}}$ (blue) and $F_{Aer_{Q_a}}$ (black), are shown over the historical period and over the projection period until 2100 for RCP 2.6 (solid), RCP 4.5 (dashed), and RCP 8.5 (dotted); $F_{Aer_{Q_a}}$ was extended with the $F_{Aer_{RCP}}$ series. (bottom) The greenhouse gas forcing series F_{GHG} (blue) and the total anthropogenic forcing series, adding $F_{Aer_{RCP}}$ (black) or $F_{Aer_{Q_a}}$ (red) to F_{GHG}, are shown over the historical period and projection period for the 3 RCP scenarios considered, as above

The measure of anthropogenic forcing F_{Ant} used in this paper is the carbon dioxide equivalent F_{CO_2EQ} series given in the RCP scenarios. It corresponds to the combined effective radiative forcing produced by Long Lived Greenhouse Gases (GHG) F_{GHG}: carbon dioxide, methane, nitrous oxide and fluorinated gases, controlled under the

Kyoto protocol, ozone depleting substances, controlled under the Montreal Protocol, and aerosols.

Chapter 8 of the IPCC AR5 reports the increase from 1750 to 2005 of effective radiative forcings with a very likely (90%) confidence interval (CI); we summarized them in Table 1 to evaluate the relative uncertainty of the different anthropogenic forcing agents. Note that we will report likely and very likely (symmetrical) CI at the 66% and 90% confidence level, respectively, throughout this work (i.e. ± 1 and ± 1.645 standard deviations respectively), in accordance with the IPCC. The largest forcing increase stems from GHG, in particular carbon dioxide, and it has a relatively small uncertainty.

b) Aerosol forcing There are also negative contributions to anthropogenic forcing from aerosols' direct effect and indirect cloud albedo effects, both with very high relative uncertainties. The total anthropogenic change in effective radiative forcing is certainly positive, due to the strong GHG forcing, but the large uncertainty on aerosol forcing strongly dominates the total uncertainty. We therefore introduce the aerosol linear scaling factor α as an extra parameter to scale aerosol forcing (see Eq. 21 below).

The aerosol forcing in the RCP files $F_{Aer_{RCP}}$ is given implicitly; it can be obtained by subtracting the combined effective radiative forcing from gases controlled under the Kyoto protocol, F_{Kyt}, and from those controlled under the Montreal protocol, F_{Mtl} (Fig. 3) from the CO_{2EQ} forcing. F_{Mtl} is given in CFC-12 equivalent concentration and we use the relation from Ramaswamy et al. (2001) to convert back to W m^{-2}.

The very likely CI given for the modern value, defined in 2005, of total aerosol forcing in the IPCC AR5 is $[-1.9, -0.1]$ W m^{-2}, but Stevens (2015) (S15) demonstrates that a forcing more negative than -1 W m^{-2} is implausible and suggests, combined with results from Murphy et al. (2009) tightening the upper bound to -0.3 W m^{-2}, that the interval be revised to $[-1.0, -0.3]$ W m^{-2}.

S15 proposes a three parameter model to derive the aerosol direct and indirect forcing directly from anthropogenic sulfur dioxide emissions Q_a :

$$F_{Aer_{Q_a}} = \gamma Q_a + \beta \log \frac{Q_a}{\bar{Q}_n} \tag{16}$$

where Q_a is the annual anthropogenic sulfur dioxide emissions, γ is the direct effect coefficient, β is the indirect effect coefficient and $\bar{Q}_n$ is the mean natural sulfur dioxide atmospheric source strength. For a given set of parameters, we can obtain a forcing series which is highly correlated (r = 0.95) with $F_{Aer_{RCP}}$, and also with Q_a (r = -0.996) itself since the log term is approximately linear near $\bar{Q}_n$. Therefore, we also include in the analysis an alternate aerosol forcing series for comparison which takes Q_a as a linear surrogate for the total aerosol forcing such that :

$$F_{Aer_{Q_a}} \approx \gamma^* Qa \tag{17}$$

where the coefficient γ^* is taken to be equal to -9.3×10^{-6} W m$^{-2} Tg^{-1} yr$ so that the modern (2005) value of $F_{Aer_{Q_a}}$ and $F_{Aer_{RCP}}$ are both about -1.0 W m^{-2}. This method allows us to derive an aerosol forcing series (Fig. 3) directly from sulfur dioxide emission data that does not rely on GCMs. We decided to present as the main result estimates made using F_{AerRCP} given it is based on a more accepted aerosol forcing series closer to that used in the MME we use for comparison, and results based on F_{AerQa} are presented alongside to show the imporant downward impact on all warming metrics that would result if it were shown to be a more reliable indication of aerosol forcing.

c) Neglected anthropogenic forcing

Left out from our analysis are other sources of anthropogenic forcing for which only the direct radiative forcing increase is reported in AR5: tropospheric ozone, stratospheric ozone, statospheric water vapour from CH_4, surface albedo from land use changes, surface albedo from black carbon aerosol on snow and ice and contrails. They were neglected given the uncertainty on the shape of the response and their small combined impact. Their total should be positive and therefore, the median estimates of sensitivity presented in this paper will possibly be biased high by a small amount.

d) Solar forcing

The two main natural forcings are solar and volcanic, but others include natural aerosols such as mineral dust and sea salt which will not be considered as they are not externally forced and depend on the internal variability of the system.

We use the recommended solar forcing F_S for CMIP5 experiments shown in Fig. 4 (along with the volcanic

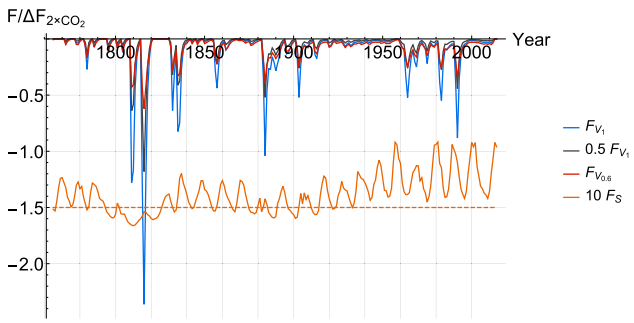

Fig. 4 Volcanic forcing F_{V_1} (blue) is shown alongside two damped versions. The black one is linearly damped by a constant 0.5 coefficient while the red one, $F_{V_{0.6}}$, is damped using Eq. 18 with $\nu = 0.6$. The solar forcing F_S (orange) has been shifted down by -1.5 and amplified by a factor of 10 for clarity

forcing). It is reconstructed by regressing sunspot and faculae time series with total solar irradiance (TSI) (Wang et al. 2005). To obtain the solar perturbation to radiative forcing, the TSI is divided by 4 due to the spherical geometry of the Earth, multiplied by the average co-albedo of the Earth (about 0.7) and the average value of the two 11-year solar cycles from 1882 to 1904 is removed to obtain an anomaly. For the future, we take the solar cycle 23 (the last one before 2008) and reproduce it to extend the series as was recommended for the CMIP5 experiments.

e) Volcanic forcing

Contrary to other forcings, there was no standard volcanic forcing series prescribed for CMIP5 experiments. The volcanic forcing F_V used here was derived from the optical depth τ_V using the approximate relation $F_V \approx -27$ W m$^{-2}\tau_V$ for instantaneous forcing. The series for τ_V and the relation to forcing were obtained from the Goddard Institute for Space Science (GISS) website (Sato 2012). The volcanic forcing covers the period from 1850 to 2012 and it was kept null for its extension into the future as was prescribed for CMIP5 experiments. We extend it back to 1765 using the optical depth reconstruction of Crowley et al. (2008). To obtain the radiative forcing, the series was multiplied by a factor of -24 W m^{-2} so that the total forcing following the Pinatubo eruption from 1991 to 1996 is equal to the Sato (2012) dataset over the same period.

The response to volcanic forcing is crucial in improving the estimation of parameters, especially the scaling exponent H, as it provides insight into the climate system on fast timescales (inter-annual); within this scaling model, there is no characteristic scale and it has implications for longer timescales. Volcanic forcing is peculiar as it is strong and highly intermittent. The intermittency can be quantified by the parameter C_1 which corresponds to the fractal codimension (i.e. 1 minus the fractal dimension) characterizing the sparseness of volcanic "spikes" of mean amplitude (see Lovejoy and Schertzer 2013b; Lovejoy and Varotsos 2016), with large eruptions producing negative forcing spikes of magnitude greater than $\Delta F_{2\times \mathrm{CO}_2}$; on the other hand, the instantaneous temperature response is weaker than expected using linear response theory.

It was found that even though volcanic forcing dwindles away quickly, it has noticeable effects on the climate at decadal timescales and longer by sharply reducing the ocean heat intake (Church et al. 2005; Stenchikov et al. 2009). This, in fact, corresponds to the physical mechanism behind the long-range memory to forcing in the linear response framework, and it acts to reduce the instantaneous impact of the volcanic forcing. Gregory et al. (2016) found that this reduction in ocean heat intake explains most of the discrepancy between forcing and response, but also added that "the magnitude of the volcanic forcing [...] may be smaller in AOGCMs (by 30 % for the HadCM3 AOGCM) than in off-line calculations that do not account for rapid cloud adjustment".

The volcanic response appears to be non-linear as the intermittency ("spikiness", sparseness of the spikes) parameter C_1 changes from about $C_{1F_V} \approx 0.2$ for the input volcanic forcing to $C_{1T} \lesssim 0.1$ for the temperature response :the latter is therefore much less intermittent than the former. Theoretically, a linear response model with a power-law Green's function cannot alter the intermittency parameter, although these estimates are sensitive to finite size effects and internal variability (Lovejoy and Varotsos 2016).

Since the volcanic forcing is too strong and too intermittent, using it within the SCRF framework requires the use of an effective volcanic forcing series if we are to use it in the linear response framework. The following theoretically motivated non-linear relation damps the amplitude of the volcanic forcing while also reducing its intermittency parameter :

$$\frac{F_{V_\nu}}{< F_V >} = \left[\frac{F_V}{< F_V >}\right]^\nu \tag{18}$$

where F_{V_ν} is the damped effective volcanic forcing, ν is the damping exponent and $< F_V >$ is the mean F_V. The damping exponent required to reduce the intermittency parameter of the volcanic temperature response, C_{1T_V}, can be theoretically calculated using the relation :

$$C_{1F_V}\nu^{\alpha_{MF}} = C_{1T_V} \tag{19}$$

where α_{MF} is the multifractality index of the volcanic forcing (e.g. Lovejoy and Schertzer 2013b). For $\alpha_{MF} \approx 1.6$, we find $\nu \lesssim 0.65$. This calculation might underestimate ν, because the temperature variability also includes the response to the less intermittent forcing (anthropogenic and solar) as well as the internal variability, which is quasi-Gaussian with $C_1 \approx 0$. Given the difficulty in estimating C_1 and α_{MF} to calculate ν, we introduce ν as a free parameter and estimate it directly.

Another avenue would be to simply introduce a linear scaling factor to reduce the amplitude of the volcanic forcing, but this would not change the nondimensional spikiness. The two methods are approximately equal for medium size eruptions, but strong peaks get reduced further in the non-linear damping case (see Fig. 4). From the point of view of maximizing the variance explained by the forced response in the temperature record, the linear damping method would produce very similar result, but we choose the non-linear damping method simply because it is better justified theoretically and decreases the intermittency parameter C_1.

2.3.2 Surface air temperature data and cmip5 simulations

Our analysis was performed on 5 observational records of surface air temperature each spanning at least the period

1880–2014: HadCRUT4 (Morice et al. 2011), the Cowtan and Way reconstruction version 2.0 (C and W) (Cowtan and Way 2014), GISS Surface Temperature Analysis (GISTEMP) (Hansen et al. 2010; GISTEMP Team 2015), NOAA Merged Land Ocean Global Surface Temperature Analysis Dataset (NOAAGlobalTemp, formerly known as MLOST) (Smith et al. 2008) and Berkeley Earth Surface Temperature (BEST) (Rohde et al. 2013).

The HadCRUT4 dataset is a combination of the sea-surface temperature records compiled by the Hadley Centre of the UK Met Office with the land surface air temperature records compiled by the Climate Research Unit in East Anglia; the C and W dataset uses HadCRUTv4 as raw data, but aims to address coverage bias by infilling missing data by kriging; the dataset with land air temperature anomalies interpolated over sea-ice was used. GISTEMP is produced by the Goddard Institute for Space Studies by combining the Global Historical Climate Network version 3 (GHCNv3) land surface air temperature records with the Extended Reconstructed Sea Surface Temperature version 4 (ERSST) and the temperature dataset from the Scientific Community on Antarctic Research (SCAR). NOAA National Climatic Data Center also uses GHCNv3 and ERSST, but with different quality controls and bias adjustements. BEST uses its own land surface air temperature product combined with a modified version of HadSST.

The CMIP5 models used are presented in Table 2. The 32 selected GCMs have historical simulation outputs available for the period from 1860 to 2005 and outputs of scenario runs over 2005–2100 for RCP 2.6, RCP 4.5 and RCP 8.5.

2.4 Parameter estimation

We have now presented the SCRF and we need to establish a procedure to estimate its parameters s, τ, and H as well as the forcing parameters α and ν. To achieve this, we relate the temperature and forcing data introduced above through the SCRF in a multi-parameter Bayesian estimation scheme.

The radiative forcing data is used to produce a theoretical forced temperature response by means of convolution with the SCRF. The convolution was implemented numerically as a discrete sum :

$$\Delta T_{Forced}(t) = -s\frac{H}{\tau}\sum_{i=1}^{N}\left(\frac{t+\tau-t'_i}{\tau}\right)^{H-1}F(t'_i)\Delta t' \tag{20}$$

where ΔT_F is the forced temperature response, H is the scaling exponent, s is the sensitivity to integrated radiative forcing, τ is the high-frequency cutoff and $F(t_i)$ is the annual forcing series $F(t)$ linearly interpolated to a resolution such that $\Delta t' = t'_{i+1} - t'_i = 0.05$ year; this resolution is much smaller than $\tau \approx 2$ years (see below) and it was found to be sufficient to produce a temperature response with negligible error compared to the analytic response for different polynomial forcing scenarios while remaining computationally efficient.

Due to the assumption of linearity, the forcing series used can be written as the sum of the constituent forcings :

$$F(t) = F_{GHG}(t) + \alpha F_{Aer}(t) + F_S(t) + F_{V_\nu} \tag{21}$$

This includes the two extra parameters discussed earlier: the aerosol linear scaling factor α and the volcanic non-linear damping exponent ν. This allows us to take into account the uncertainty on the forcing when estimating model parameters; the uncertainties on F_{GHG} and F_S are thus neglected, because they are overwhelmed by the uncertainty on F_{Aer}. The uncertainties add in quadrature and therefore, if we leave out the uncertainty on $F_{GHG}$when adding to the uncertainties of the forcing from the aerosol direct and indirect effects, we only underestimate the total uncertainty (at the 90% confidence level) by about $0.01\Delta F_{2\times\,\mathrm{CO}_2}$ (using the errors given in Table 1 for the total change over the 1750–2005 period). The amplitude of F_V is sensitive to ν and its parameter spread is informative of the related uncertainty in the volcanic forcing.

All the series used for parameter estimation were adjusted to the same anomaly level so that their mean values were zero over the reference period of 1880–1910. There are thus five parameters to determine. A time-dependent forced response $\Delta T_{Forced}(s, H, \tau, \alpha, \nu; t)$ is calculated for each parameter combination and removed from the temperature series to obtain a series of residuals which represent an estimator $\widehat{\Delta T}_{Internal}$ of the historical internal variability.

$$\widehat{\Delta T}_{Internal}(s, H, \tau, \alpha, \nu; t|\Delta T_{Obs}) = \Delta T_{Obs}(t) - \Delta T_{Forced}(s, H, \tau, \alpha, \nu; t) \tag{22}$$

where ΔT_{Obs} is an observational temperature dataset. This allows us to calculate the likelihood function $\mathcal{L}$ to be maximized which corresponds to the probability Pr of the internal variability to follow our assumed error model :

$$\mathcal{L}(s, H, \tau, \alpha, \nu|\Delta T_{Obs}) = Pr(\Delta T_{Obs}|s, H, \tau, \alpha, \nu) \tag{23}$$

The error model we use is a fractional Gaussian noise (fGn) with zero mean (see Lovejoy et al. 2015, 2017), and therefore the residuals are not independently distributed. This model approximates well the scale dependence of the internal variability, and so even if it is misspecified, it will be superior to non-parametric approaches (Poppick et al. 2017; Lovejoy et al. 2016).

Using Bayes' rule, we can derive a probability distribution function (PDF) for the parameters of interest:

$$Pr(s, H, \tau, \alpha, \nu|\Delta T_{Obs}) = \frac{Pr(\Delta T_{Obs}|s, H, \tau, \alpha, \nu)\pi(s, H, \tau, \alpha, \nu)}{Pr(\Delta T_{Obs})} \tag{24}$$

Table 2 CMIP5 models used are presented here along with the modeling centres which produced them

Modeling Center (or Group)	Institute ID	Model Name
Beijing Climate Center, China Meteorological Administration	BCC	BCC-CSM1.1 (1)
		BCC-CSM1.1(m) (2)
College of Global Change and Earth System Science, Beijing Normal University	GCESS	BNU-ESM (3)
Canadian Centre for Climate Modelling and Analysis	CCCMA	CanESM2 (4)
National Center for Atmospheric Research	NCAR	CCSM4 (5)
Community Earth System Model Contributors	NSF-DOE-NCAR	CESM1(CAM5) (6)
Centre National de Recherches Météorologiques / Centre Europen de Recherche et Formation Avance en Calcul Scientifique	CNRM-CERFACS	CNRM-CM5 (7)
Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	CSIRO-QCCCE	CSIRO-Mk3.6.0 (8)
EC-EARTH consortium	EC-EARTH	EC-EARTH (9)
LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences and CESS, Tsinghua University	LASG-CESS	FGOALS-g2 (10)
The First Institute of Oceanography, SOA, China	FIO	FIO-ESM (11)
NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL	GFDL-CM3 (12)
		GFDL-ESM2G (13)
		GFDL-ESM2M (14)
NASA Goddard Institute for Space Studies	NASA GISS	GISS-E2-H (p1(15), p2(16), p3(17))
		GISS-E2-R (p1(18), p2(19), p3(20))
National Institute of Meteorological Research/Korea Meteorological Administration	NIMR/KMA	HadGEM2-AO (21)
Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)	MOHC (and INPE)	HadGEM2-ES (22)
Institut Pierre-Simon Laplace	IPSL	IPSL-CM5A-LR (23)
		IPSL-CM5A-MR (24)
Japan Agency for Marine-Earth Science and Technology,	MIROC	MIROC-ESM (25)
Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies		MIROC-ESM-CHEM (26)
Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC	MIROC5 (27)
Max-Planck-Institut für Meteorologie	MPI-M	MPI-ESM-MR (28)
(Max Planck Institute for Meteorology)		MPI-ESM-LR (29)
Meteorological Research Institute	MRI	MRI-CGCM3 (30)
Norwegian Climate Centre	NCC	NorESM1-M (31)
		NorESM1-ME (32)

The superscript number in paranthesis identifies the GCMs for later references

where $\pi(s, H, \tau, \alpha, \nu)$ corresponds to the prior distribution for the parameters.

The prior distribution for the high-frequency cutoff τ is taken to be a Dirac-delta at 2 years. From the point of view of the parameter estimation, this choice has a marginal effect on the results since it is the ratio of H to τ which is most determinant for sensitivity estimates at equilibrium, and H is allowed to vary. This weak assumption makes the estimation procedure more manageable by reducing the number of free parameters to four : H, s, α and ν. See Appendix B for a detailed discussion concerning this choice of τ, and the sensitivity of the result to this choice.

Given no previous knowledge of s and ν, we simply assume a non-informative uniform prior over the range of parameters considered. We know that the scaling exponent H should be negative since the ECS would diverge if $H \geq 0$, and therefore, we need a prior in H which discards such non-physical values. For each value of H, there is a unique TCR to ECS ratio which can be calculated (shown on Fig. 2). We make use of this fact to derive a prior in H based on the TCR to ECS ratios from an ensemble of 23 CMIP5 GCMs (Yoshimori et al. 2016). In order to minimize the impact of this choice, we take a wide uniform distribution covering the mean ratio plus or minus 4 times the standard deviation,

i.e. between 0.22 and 0.91, which yields a corresponding uniform distribution in H between −1.1 and −0.1. Regarding α, we take as the prior distribution a normal distribution $N(1.00, 0.55)$ which has a 90% CI coherent with the modern value for F_{Aer} from the IPPC AR5, $[-1.9, -0.1]$ W m^{-2}, since the modern value of $F_{Aer} \approx -1.0$ W m^{-2} in the series we used. The efficacy of the parameter estimation method is tested on synthetic temperature series in Appendix C.

3 Results

3.1 Scaling climate response function

In this section, using Bayes' theorem as described above, we derived a PDF for the parameters of the SCRF from the mean likelihood of the five observational datasets available: HadCRUTv4, C&W, GISTEMP, NOAAGlobalTemp and BEST. Two different series were used for the aerosol forcing in equation 21: $F_{Aer_{RCP}}$ and $F_{Aer_{Q_a}}$, and the results are compared. Estimates and CIs are rounded to the resolution used for the likelihood function of the variable.

The PDF for the aerosol linear scaling factor α, using $F_{Aer_{RCP}}$ (Fig. 5: bottom left, solid line), yields a 90% CI of [0.1, 1.3] with median at 0.8; for comparison, a uniform prior in α, and a prior consistent with S15, would yield the median values of 0.6 and 0.7 with CIs of [0.2, 1.2], and [0.4, 0.9] respectively; the posterior result is therefore not sensitive to the prior assumption. Given that in 2005 $F_{Aer_{RCP}} \approx -1$ W m^{-2}, the negative value of α translates into confidence intervals for the modern aerosol forcing. It is interesting to note that we recovered a 90% CI close to what had been argued by S15, namely $[-1.0, -0.3]$ W m^{-2}, thus reinforcing his claim to decrease the uncertainty on aerosol forcing using a completely independent approach, although more comprehensive estimates still support the wider range from the IPCC's AR5 (Bellouin et al. 2020). On the other hand, the result using $F_{Aer_{Q_a}}$ supports a significantly weaker (better constrained) aerosol forcing with an α median value at 0.2 and a 90% of $[-0.2, 0.6]$ (Fig. 5: bottom left, dashed). The sulfur dioxide emission series Q_a, and therefore $F_{Aer_{Q_a}}$, shows a sharp increase between 1950 and 1980, sharper than F_{AerRCP}. Meanwhile, the global mean temperature only decreases marginally and therefore, α for $F_{Aer_{Q_a}}$ needs to be smaller for a good agreement.

The volcanic damping exponent ν was found (when using $F_{Aer_{RCP}}$) to have a 90% CI of [0.25, 0.85] with median value at 0.55 (Fig. 5: bottom right, solid line) and using $F_{Aer_{Q_a}}$ yielded a slightly higher median ν of 0.60 with a 90% CI of [0.30, 0.90] (Fig. 5: bottom right, dashed line). The datasets which tend towards a lower ν (i.e. smoother volcanic forcing), namely NOAAGlobalTemp and GISTEMP, are also those with stronger statio-temporal filtering, and therefore, a smoother volcanic response. These results confirm that volcanic forcing is generally overpowered since $\nu = 1$ has practically null probability in both cases. This means that using the original volcanic forcing series described above without the non-linear damping does not reproduce well, within the SCRF model presented, the cooling events observed in the instrumental records following eruptions: the cooling would be too strong. It also shows that the volcanic contribution is essential since ν values near zero,

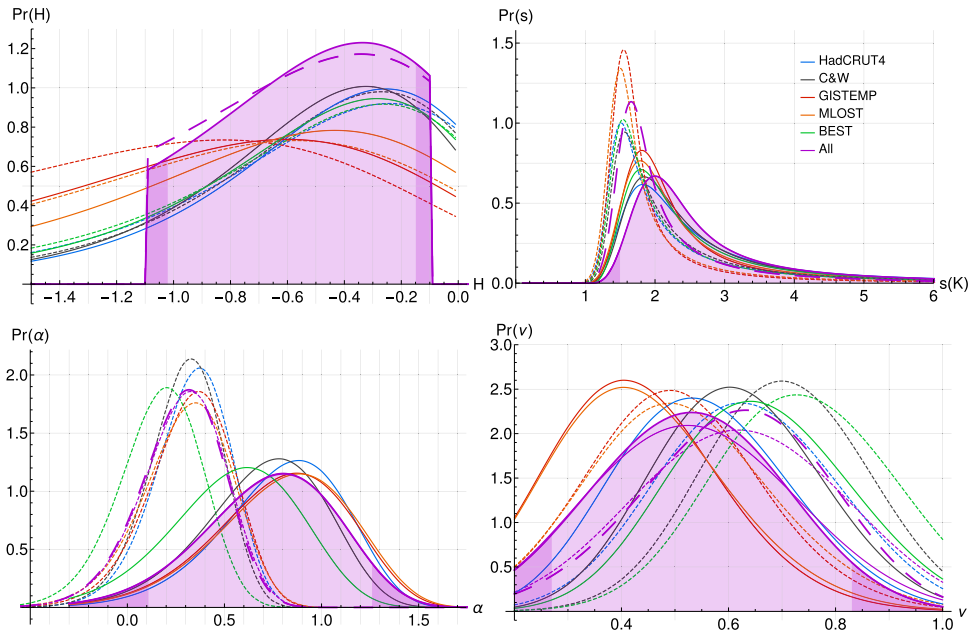

Fig. 5 For each observational dataset and also their average, PDFs are shown for the scaling exponent H (top left), the equilibrium climate sensitivity s in K per doubling of CO_2 (top right), the volcanic damping exponent ν (bottom right), and the aerosol linear scaling factor α (bottom left); for each case the probabilities over the remaining parameters were integrated out. Shown alongside are the corresponding PDFs for the parameter estimation based on both F_{AerRCP} (solid) and $F_{Aer_{Q_a}}$ (dashed). The average PDFs (purple) from the five observational datasets are shown after the prior distribution has been applied; the one using F_{AerRCP} is shown as the main result with shading and darker 5% tails

which effectively corresponds to a constant (weak) mean volcanic forcing, are also ruled out. We also obtain from Eq. 19, taking $\alpha_{MF} \approx 1.6$ and $C_{1F_V} \approx 0.2$, that the intermittency parameter C_1 of the effective volcanic forcing, and also the linear volcanic temperature response, is $C_{1T_V} = 0.07^{+0.08}_{-0.05}$ at the 90% confidence level for the $F_{Aer_{RCP}}$ result, and $C_{1T_V} = 0.09^{+0.07}_{-0.06}$ for the $F_{Aer_{Q_a}}$ result.

The most crucial parameter in our model is its scaling exponent H which is the main determinant for ECS estimates made later. We found, when using $F_{Aer_{RCP}}$, a 90% CI of $[-1.0, -0.1]$, with median value at -0.5 (Fig. 5: top left, solid line). Using $F_{Aer_{Q_a}}$ did not significantly change the median estimate of H and the 90% CI (Fig. 5: top left, dashed line). The purpose of H is somewhat similar to that of an ocean diffusivity parameter in a one-dimensional SCM and we see that in fact, the datasets using HadSST: HadCRUTv4, C&W and BEST, yielded PDFs closer to each others than to those using ERSST: NOAAGlobalTemp and GISTEMP.

We can therefore identify and separate the anthropogenic and natural components of the forced variability, ΔT_{Anthro} and $\Delta T_{natural}$ respectively, from the observational temperature series ΔT_{Obs} to obtain the internal variability $\Delta T_{Internal}$, see Fig. 6.

$$\Delta T_{Internal}(t) = \Delta T_{Obs}(t) - \Delta T_{Anthro}(\lambda, H, \tau, \alpha, t) - \Delta T_{Natural}(H, \lambda, \tau, \nu, t) \quad (25)$$

ΔT_{Anthro} and $\Delta T_{Natural}$ are obtained by the convolution of the associated forcing with the SCRF; solar and volcanic forcing in the case of $\Delta T_{Natural}$, and greenhouse gases and aerosol forcing in the case of ΔT_{Anthro}. Our result confirms that much of the 20th century warming is indeed human-induced, while much of the decadal scale variability can be attributed to natural forcing, and internal processes. In the projection period, after 2015, $\Delta T_{Natural}$ brings a positive contribution to the temperature anomaly since the prescribed solar forcing is a repetition of the anomalously high solar cycle 23. In reality, the natural forcing for the future will be quite different than those prescribed here for conformity with CMIP5 experiments.

3.2 Climate sensitivity

The paleoclimate record has been used to estimate a likely range for ECS (equivalent to our parameter s) by comparing the temperature and forcing changes over past periods such as the last millennium (Hegerl et al. 2006: [1.9, 4.3] K) and the last glacial maximum (LGM) (Schmittner et al. 2011:[1.4, 3.5] K; von der Heydt et al. 2014: [2.0, 2.6] K) There is evidence that the climate sensitivity depends on the mean climate state and therefore the modern ECS does not necessarily correspond to past ECS (Crucifix 2006; Yoshimori et al. 2011). In fact, there are many methodological challenges in order to fairly compare those different

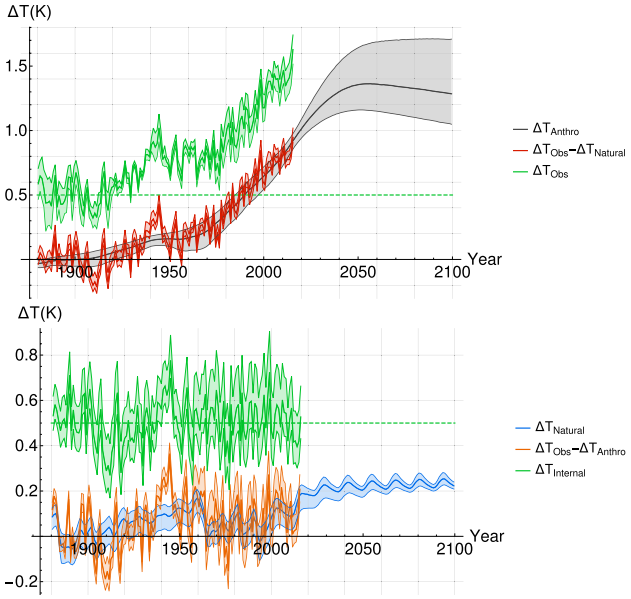

Fig. 6 (top) The mean observational temperature series ΔT_{obs} (green and shifted up by 0.5 K) has $\Delta T_{Natural}$ removed and the residual $\Delta T_{Obs} - \Delta T_{Natural}$ (red) is compared with ΔT_{Anthro} (black) calculated using $F_{Aer_{RCP}}$; they are highly correlated (r=0.94). (bottom) ΔT_{Anthro} is removed from ΔT_{Obs} and the residual $\Delta T_{Obs} - \Delta T_{Anthro}$ (orange) is compared with $\Delta T_{Natural}$ (blue); the are correlated with r = 0.51. Once $\Delta T_{Natural}$ is also removed, we obtain $\Delta T_{Internal}$ (green and shifted up by 0.5 K), which is thus uncorrelated with $\Delta T_{Natural}$. The error (shaded) given for each curve correponds to the 90% CI derived from the estimated parameters, except in the case of the observational data where it correponds to the spread between the datasets (1.645 standard deviation, by analogy with a 90% CI)

estimates, and in their extensive review on the subject, PALEOSENS Project Members (2012) considered 21 different paleoclimate studies in order to derive a [2.2, 4.8] K likely range for ECS.

An alternative approach developed to obtain ECS from the instrumental period are one-dimensional simple climate models (SCMs) with few parameters, among which the ECS is specified rather than being measured as an emergent property. The outputs of SCMs are then evaluated against the observational record, usually within a Bayesian framework, to find the best parameter combination and derive a range for the ECS : [2.1, 7.1] K (Meinshausen et al. 2009), [1.2, 3.5] K (Aldrin et al. 2012), [1.5, 5.2] K (Bodman et al. 2013), [0.9, 3.2] K (Skeie et al. 2014) and [1.9, 3.3] K (Johansson et al. 2015). See Bodman and Jones (2016) for a thorough review of this approach.

Recently, Sherwood et al. (2020) have conducted a comprehensive studies of all those lines of evidence and concluded that an ECS below 2 K was hard to reconcile with feedback-, historial- and paleoclimate-based estimates, while estimates paleoclimate-based estimates provide the strongest evidence against an ECS above 4.5 K. Their Bayesian

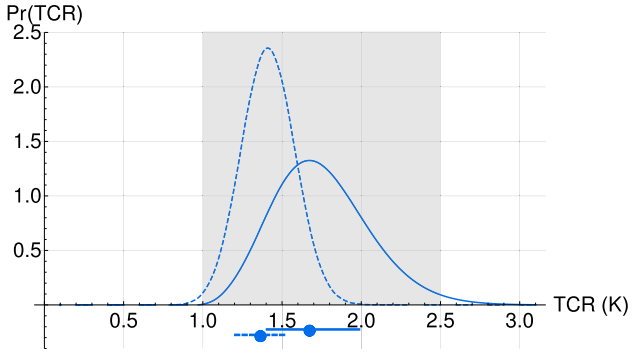

Fig. 7 The PDF for TCR is calculated using $F_{Aer_{RCP}}$ (solid) and $F_{Aer_{Q_a}}$ (dashed). The associated likely intervals (66%) (bars under the axis) are tighter than the IPCC likely range (gray shading) with lower median

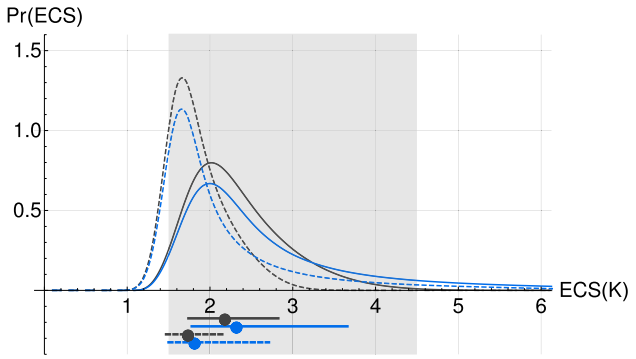

Fig. 8 Same as Fig. 7, but for ECS (blue) and ECS_{500} (black)

analysis yielded a likely range for ECS of [2.6, 3.8] K, with 3.1 K at median, and a very like range [2.3, 4.7] K.

Similarly to SCMs, with probabilistic estimates of our model parameters it is straightforward to calculate the associated temperature response to any forcing scenario; this allows us to derive PDFs for common measures of climate sensitvity: TCR and ECS. In addition, we define ECS_{500} as the temperature change 500 years after a step-function doubling in carbon dioxide concentration instead of at infinity.

Using Eq. A9 for each parameter combination with the associated probability assigned, we derived the PDFs for TCR shown in Fig. 7 using a uniform prior in TCR. Our result is slightly lower and better constrained than the one given in the IPCC AR5: a [1.0, 2.5] K likely range with best value at around 1.8 K. Using $F_{Aer_{RCP}}$, we found a median TCR of 1.7 K with a likely range of [1.4, 2.0] K, and when using $F_{Aer_{Q_a}}$ the median is revised downward to 1.3 K with an even slimmer likely range of [1.1,1.5] K.

The PDFs for ECS (Fig. 8) are calculated with a uniform prior distribution in ECS and a likely range subset of the corresponding IPCC AR5 likely range of [1.5,4.5] K was found. Using F_{AerRCP}, the ECS PDF found is skewed towards higher ECS with a likely range of [1.8,3.7] K with its median at 2.3 K. The result for F_{AerQ_a} yields a likely range of [1.5,2.7] K with its median at 1.8 K. Both results are coherent with the IPCC AR5 likely range, albeit on its lower side. A large fraction of the expected warming for the models with very high ECS only occurs at long timescale as a consequence of the slow convergence time when $H \rightarrow 0^{-}$. This is obvious when considering ECS_{500}, calculated with Eq. A5, for we obtain more symmetrical distributions than for ECS: a likely range of [1.7, 2.8] K with its median at 2.2 K for F_{AerRCP}, and a likely range of [1.5, 2.1] K with its median at 1.7 K. The median ECS we found when using $F_{Aer_{RCP}}$, which is derived from GCM experiments, is 0.6 K higher than when using $F_{Aer_{Q_a}}$ which is entirely based on historical observations. The latter is close to other observation-based estimates with low-ECS (Ring et al. 2012; Skeie et al. 2014) while the former is closer to high-ECS observation-based estimates (Meinshausen et al. 2009; Bodman et al. 2013; Otto et al. 2013; Johansson et al. 2015) as well as that of Sherwood et al. (2020), but both are lower than the 3 K best value reported in AR5 which is very close to CMIP3 and CMIP5 GCMs based estimates. All the ECS results are summarized in Table 3

3.3 Projections to 2100

Using Eq. 20, we are now able to use the SCRF to reconstruct the forced temperature variability over the historical period and make projections for the coming century according to the RCP scenarios. We compare these purely observations-based projections with those from the CMIP5 MME considered (32 GCMs). The CI given for the MME correspond to the spread between the different GCMs, i.e., the structural uncertainty, since a full probabilistic characterization of uncertainty is generally not possible with GCMs given the hundreds of degrees of freedom involved. In addition, we also show the SCRF projections after recalibrating the sensitivity coefficient s on the historical part of the MME, thus confirming that its forced response is also linear and that the memory in the MME projections is close to that of the SCRF. Furthermore, the ability of the SCRF method was shown for the majority of individual CMIP5 models by calibrating parameters over the low-emission scenario RCP 2.6 and projecting the medium- and high-emission scenarios, RCP 4.5 and RCP 8.5 (Hébert 2017).

Over the 1860–2000 period, the reconstructed forced temperature variability produced by the SCRF, whether using F_{AerRCP} or F_{AerQ_a}, and the mean of the CMIP5 MME track each other closely (Fig. 9, top left). There is only a small gap between the two over the 1915–1960 period when the CMIP5 MME is consistently warmer, but generally by less

Table 3 Summary of ECS from other observational studies

Source	Mean	Median	Likely (17–83%)
IPCC			
First Assessment Report (1990)		2.5	1.5–4.5
Second Assessment Report (1995)		2.5	1.5–4.5
Third Assessment Report (2001)			1.5–4.5
Fourth Assessment Report (2007)		3	2–4.5
Fifth Assessment Report (2013)			1.5–4.5
Feedback framework			
Gregory et al. (2002)		6.1	1.6-∞
Lindzen and Choi (2011)	0.7		0.6–1.0
Otto et al. (2013)	2.0		1.2–3.9
SCMs			
Meinshausen et al. (2009)			2.1–7.1
Ring et al. (2012)	1.5–2.		
Bodman et al. (2013)	3.2	3.1	1.5–5.2
Skeie et al. (2014)	1.8	1.7	0.9–3.2
Johansson et al. (2015)	2.6		1.9–3.3
Comprehensive			
Sherwood et al. (2020)		3.1	2.6–3.8
Current study when using $F_{Aer_{RCP}}$			
ECS_{500}	2.3	2.2	1.7–2.8
ECS(Uniform Prior)	2.8	2.3	1.8–3.7
Current study when using $F_{Aer_{Q_a}}$			
ECS_{500}	1.9	1.7	1.5–2.2
ECS(Uniform Prior)	2.2	1.8	1.5–2.7
GCMs			Very likely (5–95%)
CMIP3 (AR4 Table 8.2)	3.2	3.2	2.1–4.4
CMIP5 (AR5 Table 9.5)	3.22	2.89	1.9–4.5

Some results were taken directly from the summaries in Bodman et al. (2016) and Lewis and Crok (2014)

than 0.05 K. The variations of the instrumental data around the mean forced components, i.e. the internal variability, are greater in the first half of the 20th century than in the second. Between 1905 and 1935, the internal variability shows a prolonged negative fluctuation under the mean forced variability, and then a large positive fluctuation until 1945. Subsequently and until 2015, the departures from the mean are less persistent.

After 2000, the SCRF reconstruction accurately follows the so-called hiatus while the CMIP5 MME overshoots. This was also shown by Lovejoy (2015) with a simple Dirac CRF and an effective sensitivity to CO_2. In Schmidt et al. (2014), the overshoot of the CMIP5 MME is attributed to a combination of conspiring factors, mainly errors in volcanic and solar input, in representation of aerosols and in the evolution of El-Niño. In fact, an impulse-response model, similar to what we are using here, is used by Schmidt et al. (2014) to accurately retro-adjust the CMIP5 projection ensemble. We did not investigate the effect of those adjustments on the CMIP5 MME for future projections and simply considered the original GCM results.

The low-emission scenario, RCP 2.6, is of particular interest since the dominant anthropogenic forcing starts decreasing around the mid-2040s and allows us to observe the "warming in the pipeline" (Hansen et al. 2011) due to the large thermal inertia of oceans, which is modeled here by the long-range memory to radiative forcing of the SCRF. This means that the upper bound of the projection (higher H) keeps increasing up to 2100 even though the forcing has been decreasing since 2045, while the lower bound (lower H) has less memory and begins to decrease as early as 2055. The median SCRF surface temperature projection for $F_{Aer_{RCP}}$ (or $F_{Aer_{Q_a}}$; hereafter, the SCRF result using $F_{Aer_{Q_a}}$ is given in brackets after the result using F_{AerRCP}) under RCP 2.6 exhibits a behaviour close to a stabilization over the 2043–2100 period, reaching, in 2100, $1.5^{+0.4}_{-0.2}$ K (or $1.3^{+0.3}_{-0.2}K$) at the 90% confidence level compared to $1.7 \pm 0.8K$ for the MME (Fig. 9, top right).

The behaviour of the SCRF projection is in fact similar to the MME result over 2043–2100 in that regard, which has an average projection which also shows stabilization, and thus we can project the MME rather well after recalibrating the sensitivity coefficient s. Out of the 32 CMIP5 models analyzed and numbered for reference in Table 2, 12 showed a significant cooling trend between 2043 and 2100 (2, 5, 10, 11, 13, 15, 16, 18, 22, 24, 28, 29) while 9 showed a significant warming trend over the same period (4, 6, 7, 8, 12, 21, 25, 26, 30); the remaining 11 models did not show a significant trend in either direction.

The forcing of the middle scenario, RCP 4.5, stabilizes in the mid 2060s, but the temperature projections, in Fig. 9 (bottom left), continue rising until 2100, both with the SCRF and the CMIP5 MME, and reaching $2.3^{+0.7}_{-0.5}K$ (or $1.9^{+0.4}_{-0.3}K$), at the 90 % confidence level, and $2.6 \pm 0.8K$ respectively. Both projections for the business as usual scenario, RCP 8.5, show warming at a staggering rate up to $4.1^{+1.3}_{-0.9}K$ (or $3.4^{+0.7}_{-0.5}K$) and $4.8 \pm 1.3K$, respectively, in 2100 as shown on Fig. 9 (bottom right). For the RCP 4.5 and RCP 8.5 scenarios, the SCRF projection has 25% (44%) and 15% (46%) less uncertainty than the MME spread and the median is colder by 0.3 K (or 0.7 K) and 0.7 K (or 1.4 K) respectively. In fact, the SCRF projection approximately corresponds to the lower half (or the lower quarter) of the CMIP5 MME projection range.

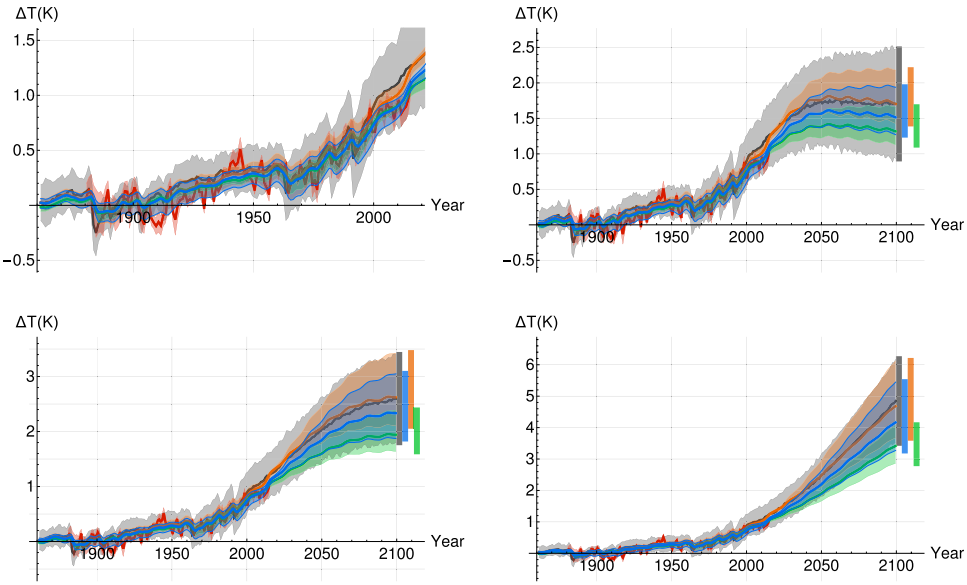

Fig. 9 The median forced temperature variability is projected using the SCRF, with the parameters calculated using $F_{Aer_{RCP}}$ (blue) or $F_{Aer_{Q_q}}$ (green), and compared with the CMIP5 MME projection (black) and a SCRF projection of the MME (orange) as well as the historical temperature observations (red); 90% CI are indicated along the projections (shaded) and for the year 2100 (thick vertical lines). The historical period (top left) and the projections until 2100, for RCP 2.6 (top right), RCP 4.5 (bottom left) and RCP 8.5 (bottom right), are shown

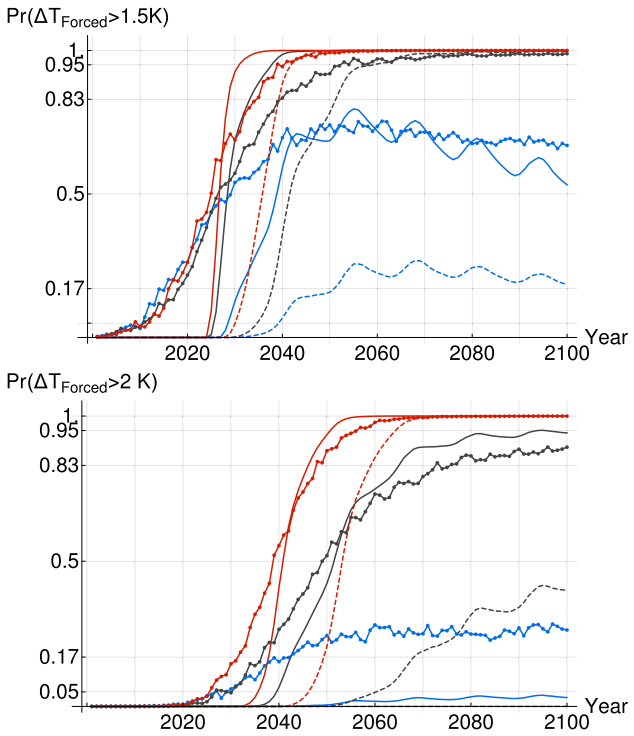

Fig. 10 The probability for the global mean surface temperature of exceeding a 1.5 K threshold (top), and a 2 K threshold (bottom) are given as a function of years for the SCRF, using F_{AerRCP} (solid) and using $F_{Aer_{Q_q}}$ (dashed), and for the CMIP5 MME (circles). The three RCP scenarios are considered for each case: RCP 2.6 (blue), RCP 4.5 (black) and RCP 8.5 (red)

International negotiations often invoke 1.5 K and 2 K threshold not to be crossed in order to avoid a major climatic upheaval. The probability when those temperature are reached under each scenario can be calculated for the SCRF and the CMIP5 MME (Fig. 10). For the latter, we assumed the global mean temperature of the MME every year to be normally distributed in order to obtain the probability of having crossed the given threshold at that time. Generally, the increase in probability as a function of years is sharper for the SCRF than for the CMIP5 MME given the smaller uncertainty on the projections.

For the low-emission scenario RCP 2.6, it is likely that the 1.5 K threshold would be exceeded in 2100 according to the SCRF projections with 54% probability (or 19%), while it was slightly more likely for the CMIP5 MME with 67% probability. It is very likely (>97%) that the 2 K threshold would not be crossed in 2100 according to the SCRF projection while the CMIP5 MME still shows 26% probability of exceeding.

For the medium-emission scenario RCP 4.5, the SCRF asserts it is extremely likely (>95%) that the global temperature will be above 1.5 K as early as 2038 (or 2059); for the CMIP5 MME, it becomes extremely likely after 2050. The 2 K threshold on the other hand will not as certainly be crossed in 2100 according to the SCRF projections as it has 94% (or 40%) of overshooting, similarly to the CMIP5 MME which shows a 89% probability of overshoot in 2100.

For the high-emission scenario RCP 8.5, all methods show a high probability of a warming exceeding 2 K before 2100. According to the SCRF, the risk of overshooting 1.5 K is negligible before 2024 (or 2028), but extremely likely after 2036 (or 2047), similarly to the CMIP5 MME which reaches the 95% probability of overshooting 1.5 K in 2038. The 2 K threshold is also extremely likely to be crossed about 15 years later in 2055 for both the SCRF and CMIP5 MME (or 2068).

4 Conclusion

Multidecadal climate projections rely almost exclusively on deterministic global climate models (GCMs) in spite of the fact that there are still very large structural uncertainties between Coupled Model Intercomparison Project phase 5 (CMIP5) GCMs, i.e. each has its own climate, rather than the real world climate. Climate skeptics have argued that IPCC projections are untrustworthy precisely because they are entirely GCM based. While this conclusion is unwarranted, it underscores the need for independent and qualitatively different approaches. It is therefore significant that the alternative GCM-free approach we present here yields comparable results albeit with smaller uncertainty.

This motivated us to elaborate a model, based on the scaling of climate processes, for the response of the global mean air surface temperature of the Earth to external forcing: the scaling climate response function (SCRF). The forced component of temperature variability is reconstructed from external forcing within the linear response framework with a power-law scaling Green's function truncated at high-frequency. The stochastic component ultimately due to the internal turbulent dynamics of the atmospheric system is approximated by a fractional Gaussian noise process, as was proposed in the ScaLing Macroweather Model (SLIMM) (Lovejoy et al. 2015). Similarly, GCMs yield stochastic internal variability with an approximately linear mean forced response (Meehl et al. 2004), and we showed that in fact the SCRF model can project their forced response rather accurately.

Our model is robust and by restricting the parameter space, it allows for a full probabilistic characterization of uncertainty by Bayesian inference. A by-product of our analysis is a better constrained aerosol forcing since we found the aerosol linear scaling factor α to be within a 90 % CI of [0.1, 1.3] for the RCP aerosol forcing $F_{Aer_{RCP}}$. This supports a revision of the global modern aerosol forcing 90 % confidence interval to a narrower $[-1.3, -0.1]$ W m^{-2}, similar to Stevens (2015). On the other hand, we obtain a very weak aerosol forcing if instead we use $F_{Aer_{Q_a}}$, which was reconstructed directly from sulfur dioxide emissions using a linearized version of Stevens' proposed model (Eq. 16). While the difference between the aerosol series might arise from a misreprensation of aerosol effects in GCMs which were used to produce F_{AerRCP}, errors could also arise because of deviations from linearity with respect to SO_2 emissions due to other aerosol species not explicitly taken into account to produce $F_{Aer_{Q_a}}$.

Following others (Church et al. 2005; Stenchikov et al. 2009; Lovejoy and Varotsos 2016), we also found that the volcanic forcing was generally over-powered and overly intermittent, or too "spikey", to produce results, within the SCRF framework, consistent with instrumental data. An effective volcanic forcing with lower-intermittency was obtained with a non-linear damping by the exponent ν. It was found to be within [0.25,0.85] at the 90% confidence level using F_{AerRCP} (or [0.30,0.90] using $F_{Aer_{Q_a}}$), yielding a corresponding CI of [0.02,0.15] (or [0.03,0.16] using $F_{Aer_{Q_a}}$) for the intermittency parameter C_1 of the effective volcanic forcing, which is compatible with the lower intermittency of the temperature.

Our analysis supports better constrained TCR and ECS likely range than the IPCC AR5. When using $F_{Aer_{RCP}}$ (or $F_{Aer_{Q_a}}$), the range shrinks from [1.0, 2.5] K to [1.4, 2.0] K for the TCR (or [1.2, 1.5] K) and from $[1.5, 4.5]\,K$ to $[1.8, 3.7]\,K$ for the ECS (or [1.5, 2.7] K); the median estimates also decrease from 1.8 K to 1.7 K (or 1.4 K) for the TCR and from 3.0 K to 2.4 K (or 1.8 K) for the ECS. This agrees with other recent observation-based studies (Otto et al. 2013; Skeie et al. 2014, and Johansson et al. 2015) which also support a downward revision of the ECS upper 17% bound by at least half a degree. In addition, the ECS_{500} was found to be significantly smaller, $2.2^{+0.6}_{-0.5}K$ (or $1.7^{+0.4}_{-0.2}K$), than the ECS. This implies that if the ECS is on the higher end of the CI, then a large fraction of the warming would be experienced hundreds of years after a potential stabilization of anthropogenic forcing. An important and rather conservative claim supported by this evidence is therefore that the upper 5% ECS bound and median of AR5 can be safely revised downward to 4.0 K and 2.5 K. The lower 5% bound of 1.5 K, on the other hand, remains reliable.

Our ECS likely range is therefore better constrained than paleoclimate estimates such as Hegerl et al. (2006) and PALEOSENS Project Members (2012) who found $[1.9, 4.3]\,K$ and $[2.2, 4.8]\,K$ respectively. Our method is strictly based on modern instrumental data and the low uncertainty could be an "epoch bias", i.e. if climate sensitivity depends on the climate mean state, then an estimate based on only one period would not necessarily be representative.

Our historical approach also decreases the uncertainty on projections by more than a factor of two, even with the large uncertainty associated with aerosol forcing, compared to the structural uncertainty of a multi-model ensemble (MME) of CMIP5 global climate models for RCP scenarios. However, the uncertainties compared are qualitatively different since it is not possible to charaterize probabilistically the entire set of parameters involved in GCMs. The structural uncertainty in a MME is rather based on the dispersion between the GCM runs, each produced with a definite set of parameters. Our approach on the other hand does not yield any structural uncertainty

since we assumed a single CRF which is able to produce a wide range of climate responses.

The SCRF projections to 2100 are entirely independent of the GCMs. Still, they are within the uncertainty bounds of the latter, effectively providing an independent confirmation of the GCM projections. This eliminates one of the key climate skeptic arguments: projections are not reliable since they are solely GCM-based. This conclusion is therefore important not only for scientists but also for policy makers, stressing the need for mitigation and adaptation measures. The SCRF is a the first of a new family of approaches that directly model the temperature at monthly scales and longer, see for example Procyk et al. (2020) that, in addition to scaling, directly exploits the principles of energy-balance (the fractional energy-balance equation) and produces projections and climate sensitivity estimates in a similar range, although with a better estimation of the scaling exponent H by also utilizing the macroweather high-frequency response.

According to our projections made to 2100, to avert a 1.5 K warming, future emissions will be required to undergo drastic cuts similar to RCP 2.6, for which we found a 46% probability to remain under the said limit; it is virtually certain that RCP 4.5 and RCP 8.5-like futures would overshoot. Even a 2.0 K warming limit would surely be surpassed by 2100 under RCP 8.5 and probably also under RCP 4.5, with only a 6% chance of remaining under the limit. The safest option remains RCP 2.6 which we project to remain under 2.0 K with very high confidence. The question remains whether it is at all realistic given that it relies strongly on the massive deployment of speculative negative emission technologies.

On the other hand, our model has obvious limitations since it assumes a linear stationary relationship between forcing and temperature, neglecting nonlinear interactions which could arise as the system evolves, as it currently warms. In particular, so-called tipping points could be reached in the coming century which would lead to a breakdown of the linear model proposed. Such potential behaviours are of critical value for improving future projections, but they have not yet been observed with high confidence even in GCMs. This underlines the need to exploit paleoclimate archives to achieve a better understanding of low-frequency natural variability, namely the transition scale from the macroweather regime to the climate regime. In this study, we have assumed the increased variability in the climate regime to be strictly a result of forcing, but internal modes of variability could also have a significant contribution for longer timescales.

More relevant to human activities and adaptation policies are regional projections which are also almost entirely produced by GCMs. In addition to the discrepancy between their global mean response, CMIP5 GCMs show widely varying spatial patterns of warming over the last century between themselves, and with significant differences from those observed in gridded instrumental temperature datasets (Hébert and Lovejoy 2018). Future work should explore the possibility of data driven models at the regional scale for climate projection. Already, it was found by Lovejoy and de Lima (2015) that in the macroweather regime, statistical space-time factorization holds for temperature and precipitation, both in instrumental datasets and in GCMs. This implies the possibility of developing linear response models for regional projections, although the main obstacle foreseen will be to identify the forced signal in the stronger regional internal variability.

Acknowledgements Raphaël Hébert's graduate stipend was funded by Lovejoy's Natural Sciences and Engineering Research Council of Canada (NSERC)-Discovery grant: "Multifractal Geophysics". RH has also received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement #716092 and #772852). This project is a contribution to the international training program ArcTrain Canada and also to the research program of Québec-Océan. Bruno Tremblay acknowledges NSERC-Discovery, the Canadian Sea Ice and Snow Evolution (CanSISE) network funded by the NSERC Climate Change and Atmospheric Research (CCAR) program, and also the Marine Environmental Observation Prediction and Response (MEOPAR) network. This work is a contribution to the NSF Office of Polar Program Grant #1928126. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and I thank the climate modeling groups (listed in Table 2 of this thesis) for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix: Response of the SCRF to theoretical forcing scenarios

In this section, we calculate the expected temperature response for different forcing scenario under the SCRF of Eq. 13. The expected forced temperature response for an N^{th}-order polynomial forcing function $F_{poly} = \sum_{i=0}^{N} a_i t^i \Theta(t)$ can be written as:

$$\Delta T_{poly}(t) = \int_{-\infty}^{t} \sum_{i=0}^{N} a_i {t'}^i s \frac{(-H)}{\tau} \left(\frac{t - t' + \tau}{\tau} \right)^{H-1} \Theta(t') dt' \tag{A1}$$

$$\Delta T_{poly}(t) = -s \frac{H}{\tau} \sum_{i=0}^{N} a_i \sum_{k=0}^{i} \binom{i}{k} \frac{(t+\tau)^{i-k}}{H+k} \frac{(-1)^k}{\tau^{H-1}} ((t+\tau)^{H+k} - \tau^{H+k}) \tag{A2}$$

where a_i is the i^{th}-order coefficient of the forcing polynomial and the Heaviside function $\Theta(t)$ makes the forcing null for negative times (the forcing is assumed to be zero for $t < 0$). This is convenient not only since the ECS and TCR experiments are specific cases, but also because it also allows to calculate the theoretical response for any forcing which can be approximated by a polynomial. Of course, any orthogonal basis of functions could be used. Note that $\forall H \in \{0, -1, -2, ..., -N\}$, where N is the order of the polynomial, there is a singularity, but the limit at such points is the same from both sides so that we can redefine it the value of the limit without further complications.

To measure ECS, the forcing scenario used has $a_0 = \Delta F$ and $a_i = 0$ for all $i > 0$, which corresponds to the following step-function :

$$F_{step}(t) = \Delta F \Theta(t) \tag{A3}$$

where ΔF corresponds to a doubling in CO_2 and $\Theta(t)$ is again the Heaviside step function. We can calculate the temperature change for this step-increase in forcing at any time following the increase:

$$\Delta T_{step}(t) = -s\Delta F \left[\frac{(\tau + t)^H - \tau^H}{\tau^H} \right] \tag{A4}$$

which can be rewritten as:

$$\Delta T_{step}(t) = s\Delta F \left[1 - \left(\frac{t}{\tau} + 1 \right)^H \right] \tag{A5}$$

To obtain the ECS, we only need to take $t \to \infty$:

$$ECS = \Delta T_{eq} = \Delta T_{step}(t \to \infty) = \begin{cases} s\Delta F & H < 0 \\ \infty & 0 \geq H \end{cases} \tag{A6}$$

ECS is thus the temperature change at infinity when a steady-state is reached; in this model this requires $H < 0$. If $H \geq 0$, we effectively have a runaway greenhouse model as the temperature keeps increasing monotonically. As the scaling exponent approaches 0 from the left, $H \to 0^-$, the convergence time increases and then diverges for a positive exponent. We also define ECS for a more concrete horizon of 500 years :

$$ECS_{500} = \Delta T_{step}(500\, years) \tag{A7}$$

This allows us to calculate the warming fraction left for a given scaling exponent H between half a millennium and infinity.

To calculate TCR, a different forcing scenario is used where a doubling of carbon dioxide takes place over 70 years, increasing at a rate of 1% per year, and then held constant. The forcing produced by carbon dioxide is proportional to the logarithm of its concentration and therefore, this exponentially increasing carbon dioxide scenario corresponds to a linear increase in forcing with $a_1 = \Delta F \tau_{tr}^{-1}$ and $a_i = 0 \; \forall i \neq 1$:

$$F_{tr}(t) = \begin{cases} 0 & t < 0 \\ \Delta F \frac{t}{\tau_{tr}} & 0 < t < \tau_{tr} \\ \Delta F & \tau_{tr} < t \end{cases} \tag{A8}$$

where τ_{tr} is the time over which the doubling takes place, that is 70 years, and ΔF is the final change in forcing. We can again calculate the temperature response analytically with the SCRF :

$$\Delta T_{tr}(t) = \begin{cases} \frac{-s\Delta F}{\tau^H \tau_{tr}(H+1)} \left[(\tau + t)^{H+1} - \tau^H (\tau + t(H+1)) \right] & 0 \leq t \leq \tau_{tr} \\ \frac{-s\Delta F}{\tau^H} \left[\frac{(\tau+t)^{H+1} - (\tau+t-\tau_{tr})^H (\tau+t+\tau_{tr}H))}{\tau_{tr}(H+1)} + (\tau + t - \tau_{tr})^H - \tau^H) \right] & \tau_{tr} \leq t \end{cases} \tag{A9}$$

The temperature change at $\Delta T_{tr}(t = \tau_{tr})$, that is when the forcing reaches its final value, is then defined as the TCR. This measure is not representative of the ECS, although they become equal for sufficiently large negative values of the scaling exponent H, but it is even more relevant for near-term projections as the TCR will be the main contribution to global warming as long as the exponential rise in carbon dioxide concentration continues. If we take $t \to \infty$, we notice that the left term in the bracket vanishes and we are left with Eq. A6. We thus see that the steady-state response does not depend on the forcing scenario as $\Delta T_{tr}(t \to \infty) = \Delta T_{step}(t \to \infty)$, but only on the change in forcing. It is convenient to analytically calculate the ratio of TCR to ECS as a function of H and τ :

$$\frac{TCR}{ECS} = \frac{T_{tr}(\tau_{tr})}{T_{step}(t \to \infty)} \tag{A10}$$

This ratio is shown on Fig. 2 for $\tau = 2\,years$ with equilibrium defined at infinity and at 500 years. It decreases monotonically to zero for the former as the scaling exponent approaches zero from the left and the ECS diverges. The gap between TCR and ECS therefore deepens for higher values of H, which explains the sensitivity of ECS to H since the observed warming, similar to the TCR, is a smaller fraction of the ECS. On the other hand, for increasingly negative values of H, the TCR to ECS goes to one as the memory of the model becomes infinitely small, effectively becoming a Dirac δ function. Also shown, the warming fraction left after 500 years of constant forcing informs us about the longest component of the memory of the model, and that therefore, for values of $H \approx -0.5$ and greater, a sizeable fraction of the warming will occur beyond that horizon.

Appendix: inner scale of the linear response

In this section, we present an argument in favour of the inner scale of the linear response at approximately 2 years. We mentioned that the physical mechanism which we approximate by integrating the radiative forcing history with the SCRF is the accumulation of heat, mostly, in the oceans which in turn drives the forced temperature response in the atmosphere and at the sea-surface. Therefore, we expect a strong correlation between the two at large time scales where the forced variability overwhelms the internal one, and thus a similar large scale scaling behaviour which should be dominated by the anthropogenic effects. It may seem surprising that up until now, the fundamental issue of land-ocean coupling has, to our knowledge, not been investigated empirically. The likely reason is that it requires the use of Haar, or other appropriately defined, fluctuations. The Haar fluctuation over an interval Δt is simply the difference between the mean over the first and second half of the interval. It is a convenient way to characterize variability as a function of time scale in real space and it is valid whether fluctuations increase or decrease with scale, which is necessary here in order to avoid spurious statistical issues at low-frequencies (Lovejoy and Schertzer 2012)

Figure 11 compares the mean Haar fluctuations as a function of timescale of the global sea-surface temperature (SST), the land-surface air temperature (LST) and the global combined land-ocean surface air temperature (SAT). The forced component of variability dominates for large timescales and we see that SST, LST and SAT have similar forced variability, higher for LST and weaker for SST as expected, and a scaling behaviour akin to the forcing for time scales above 20 years. For short timescales, the internal variability dominates and this leads to diverging behaviour between the SST, LST and the SAT since

Table 4 The ECS and ECS_{500} are calculated for different values of the inner scale τ

τ	Median ECS (K)	ECS likely range (K)	Median ECS_{500} (K)	Likely range ECS_{500} (K)
1 year	3.4	[2.2,7.6]	2.3	[1.9,2.8]
2 years	2.3	[1.9,3.4]	2.1	[1.8,2.6]
4 years	2.0	[1.7,2.4]	1.9	[1.7,2.3]
Combined	2.4	[1.8,4.4]	2.1	[1.8,2.6]

A combination of the three results weigthed by their likelihood is also shown

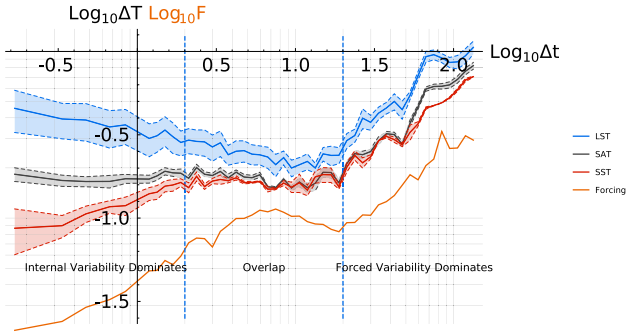

Fig. 11 The mean Haar fluctuations (in units of K) are shown as a function of timescale Δt (in years), adjusted to annual resolution from monthly resolved series, for LST (blue), SAT (black) and SST (red), and the shaded areas correspond to the standard deviation between all the datasets considered. Vertical dashed lines indicate the domains over which internal or forced variability dominates, or where they are of similar magnitude. At high-frequency, the scaling behaviours differ due to diverging turbulent dynamics of the internal variability. The same analysis performed for the forcing series (orange; in number of CO_2 doublings) shows a similar scaling behaviour at low-frequency, where the forced variability dominates, between the forcing and the temperature series. The series used are described in the data sub-section

the turbulent dynamics of the ocean and the atmosphere are different. The driving energy rate density, the turbulent ϵ that determines the weather-macroweather transition scale τ_w, is about 10^{-3} W Kg^{-1} in the atmosphere, but only about 10^{-8} W Kg^{-1} in the ocean surface layer (Lovejoy and Schertzer 2010). The transition time scale is proportional to $\epsilon^{-\frac{1}{3}}$ so that the weather-macroweather transition in the SST is visible around 2 years rather than at a sub-monthly scale in the atmosphere, about 10 days.

The SAT should scale similarly to LST in the macroweather regime since they are both atmospheric fields with similar turbulent dynamics, but we observe a discrepancy on Fig. 11 : as expected the monthly LST series do not show the sub-monthly transition, but the scaling of SAT below 2 years shows a flattening. It is possible this is just a bias resulting from the inclusion of SST as

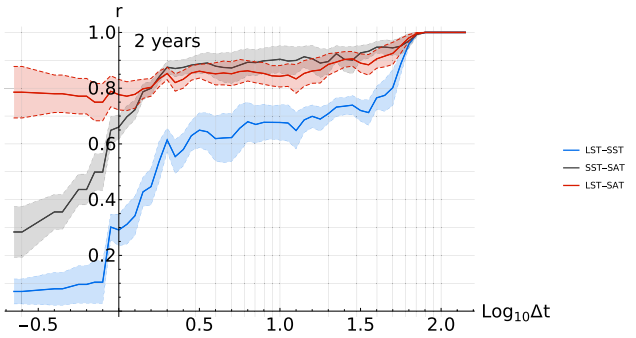

Fig. 12 The average correlation coefficient r between the first-order Haar fluctuations of LST-SAT (red), SST-SAT (black), and LST-SST (blue) are shown as a function of timescale Δt in *years*. It was calculated for an ensemble of all possible pairs between the LST, SST and SAT datasets considered, and the shaded areas correspond to the one standard deviation. There is a rapid increasing trend in correlation between land and ocean as a function of timescale under 2 years and a slowly increasing trend above 2 years. To improve the statistics of the fluctuations at large timescale, the fluctuations from reversed series are also considered; this means there are only 2 fluctuations for timescales above half of the series length and only 1 or -1 correlation values are possible; thus the sharp increase above 40 years is not quantitatively significant

air-surface temperature in global SAT, which could artificially decrease the high-frequency variability of SAT, but more investigation needs to be done to confirm the claim.

The model presented is not expected to reproduce the internal variability, which is treated as an independant stochastic process, and the high-frequency cutoff τ was introduced as the smallest scale over which the linear response approximation is expected to hold. As mentioned above, this should translate in correlation between the ocean and the atmosphere forced responses. Figure 12 shows the mean pairwise correlation between the first-order Haar fluctuations of SST, LST and SAT at different timescales. We observe a steep increase to high correlation between all pairs up to 2 years; the SAT series are composed from the LST and SST series and as such it is not surprising to find high correlation, but the completely independent LST and SST series confirm the same transition. This supports our interpretation that as we consider larger time scales, the internal variability averages out and the air surface temperature becomes proportional to the ocean temperature as a consequence of the low-frequency dominance of the forced response from integrated radiative forcing on both. This motivates us to fix the model parameter $\tau = 2$ years for simplicity.

Result for ECS and ECS_{500} are calculated for $\tau \in \{1, 2, 3\}$ *years* in order to evaluate the sensitivity of the result to this choice (Table 4). We also give an average of the three results weigthed by their relative likelihood (which yields weights of 0.38,0.33 and 0.29 respectively), and the result is in fact very close to the $\tau = 2$ years value used. The likelihood function therefore varies slowly along the direction of τ, with the 3 values almost equally probable, and does not allow to constrain the parameter well by itself. The parameter estimation with a smaller τ yields an H closer to zero, and therefore the high ECS tail thickens as ECS increases as $H \to 0^-$, and as ultimately diverges for $H \geq 0$. However, the convergence time also increases as $H \to 0^-$ and so the larger part of the response will happen at multi-centennial, multi-millenial and even much longer timescales. Therefore, the choice of τ can have a sizable impact for the theoretical ECS at infinity, but the effect will be marginal for any centennial projections and quasi-equilibrium climate sensitivity such as ECS_{500}.

Appendix: Numerical estimation of the SCRF parameters

This appendix provides extra information concerning the SCRF model parameter estimation technique and tests the efficacy of the method on synthetic time series.

The parameters for the fractional Gaussian noise (fGn) (which is used as a model for the internal variability)

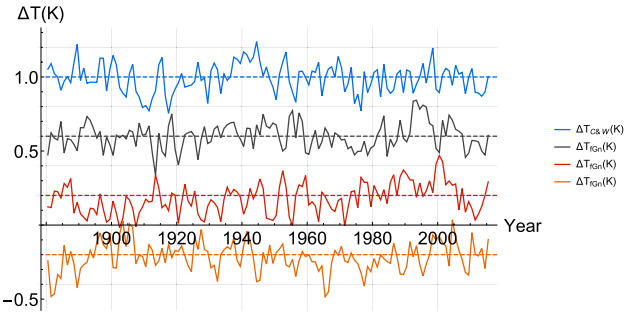

Fig. 13 The forced temperature response is removed from the C&W instrumental temperature series to obtain the associated internal variability series (blue). The best volatility σ and Hurst exponent H' parameters to describe the series as an fGn process were estimated as 0.10 and −0.26 respectively. Three single realizations of an fGn process with the same parameters are shown (black, red and orange) are shown as examples. Each series has null mean and fluctuates about its associated reference line (dashed)

Table 5 Parameters for the fractional Gaussian noise process used for the error model when calculating the likelihood function for each dataset

Dataset	σ	H'
HadCRUT4	0.10	−0.24
C&W	0.10	−0.26
GISTEMP	0.12	−0.17
NOAAGlobalTemp	0.12	−0.15
BEST	0.11	−0.23

The mean μ was set to zero, so only the volatility σ and exponent H' are given

are estimated with an iterative procedure. A first estimate is made from the internal variability obtained from a likely combination of parameters; it is then used as the error model to find the parameter combination with maximum likelihood whose associated internal variability serves to obtain the second estimate of the fGn parameters; the maximization process is repeated until the parameters converged, which occurred after only 2 steps. Figure 13 shows the internal variability from the Cowtan & Way (C&W) instrumental temperature series and three other realizations of the same estimated fGn process (Table 5). For CMIP5 GCMs, it would be simpler since the fGn parameters could be estimated directly from unforced control runs which correspond to realizations of the internal variability in each GCM. The following *Mathematica 10.4* (Wolfram Research Inc 2016) functions were used to perform the above calculations : LogLikelihood[proc,data], FractionalGaussianNoiseProcess[μ,σ,h] and EstimatedProcess[data,proc]. Note that the Hurst exponent h used within *Mathematica 10.4* describes the scaling behaviour of the associated fractional Brownian motion obtained by integrating the fGn. The notation $H' = h - 1$ corresponds to the associated parameter in Lovejoy et al. (2015) which directly describes the scaling associated with the fluctuations of the fGn itself. H' should not be confused with the scaling exponent H

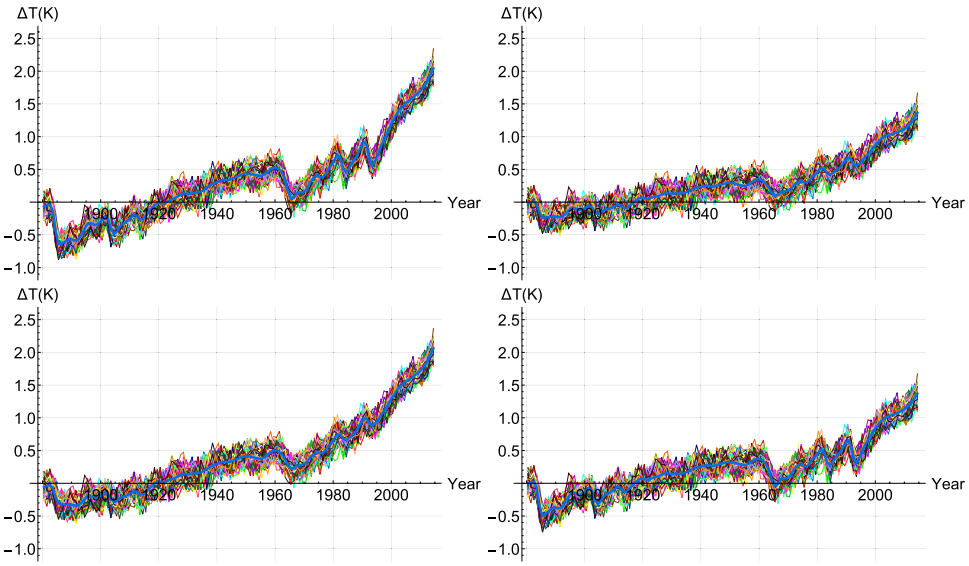

Fig. 14 Each graph shows the simulated forced temperature response (thick, blue) with the following (H, ν) parameter combination : $(-0.2, 0.8)$ (top left), $(-0.6, 0.5)$ (top right), $(-0.2, 0.5)$ (bottom left)and $(-0.6, 0.8)$(bottom right). To each forced response, 25 realizations of a fGn were added (thin, colors) to produce synthetic series of known parameters

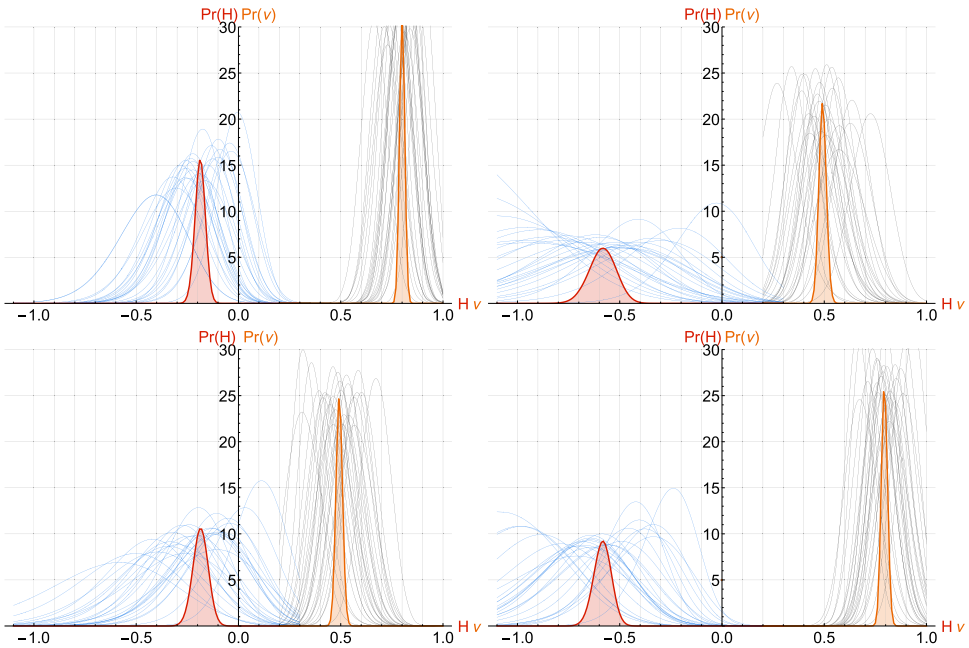

Fig. 15 The resulting PDFs for each of the 25 realizations of fGn considered for the parameters H (blue) and ν (black) are shown for simulations produced with the following (H, ν) parameter combination : $(-0.2, 0.8)$ (top left), $(-0.6, 0.5)$ (top right), $(-0.2, 0.5)$ (bottom left) and $(-0.6, 0.8)$(bottom right). In addition, the re-normalized product of the individual PDFs, after they were re-interpolated at 0.01 increments, are shown for H (red) and ν (orange). The PDFs for the individual realizations have been multiplied by a factor of 5 for clarity

Table 6 Summary of results for test simulations of four (H, ν) parameter combinations. The median H and ν given along with the associated confidence intervals are obtained from the combination of the PDFs of the 25 individual realizations of fGn

H	ν	Median H	90% CI	% correct	Median ν	90% CI	% correct	δ_s
−0.2	0.8	−0.18	[−0.22, −0.14]	96 %	0.80	[0.78,0.82]	92 %	0.23 ± 0.01
−0.6	0.5	−0.57	[−0.68, −0.47]	96 %	0.49	[0.47,0.52]	88 %	0.123 ± 0.006
−0.2	0.5	−0.19	[−0.25, −0.13]	96 %	0.49	[0.47,0.52]	88 %	0.25 ± 0.01
−0.6	0.8	−0.58	[−0.65, −0.51]	96 %	0.80	[0.77,0.82]	88 %	0.113 ± 0.006

The % correct corresponds to the number of realizations for which the true parameter lies in the 90% CI. The average relative error δ_s for all realizations is given with its one standard deviation

used to produce the forced temperature response using the SCRF.

To show the efficacy of our method, we test it on simulations generated by adding realizations of fGn, with $\sigma = 0.1$ and $H' = -0.25$, onto forced temperature response generated with Eq. 20. Given that it is computationally heavy to calculate the likelihood of a fGn, we restrict the test simulations to $\alpha = 1$, $s = 1$, $H \in \{-0.2, -0.6\}$ and $\nu \in \{0.5, 0.8\}$. This yields 4 parameter combinations to which are added 25 realizations of fGn with known parameters $(\mu, \sigma, h) = (0, 0.1, 0.75)$, shown on Fig. 14.

Figure 15 shows the resulting PDF for H and ν with α known and the best s found. The likelihood function was calculated $\forall H \in [-1.1, 0.3]$ and $\forall \nu \in [0.2, 1.0]$ both in increments of 0.05. For each combination of H and ν, the best s is found; first an estimate s_0 is made by least-square regression and then a maximum likelihood estimate is made $\forall s \in [0.85 s_0, 1.15 s_0]$ in increments of $0.05 s_0$. The likelihood function is then interpolated to find the best s and its associated PDF using Bayes' rule with a uniform prior distribution; the relative error δ_s for each s, defined as $\delta_s = \sigma_s s^{-1}$ where σ_s is the one standard deviation defined from the PDF, was found to be very small and consistent between the different parameter combinations (H, ν). Results are summarized for each set of simulations considered in Table 6.

Therefore, we can recover the parameters accurately with this Bayesian approach with a fGn error model; the 90% confidence intervals are effective, but they remain large and widely varying for single realizations because of the stochastic internal variability.

Decreasing the resolution in H to 0.1 does not alter the result significantly and for the parameter estimation from observational data it will be sufficient to calculate the likelihood function $\forall s \in [0.085 s_0, 1.15 s_0], \forall H \in [-1.5, 0.2]$ in increments of 0.1, $\forall \nu \in [0.2, 1.0]$ in increments of 0.05 and $\forall \alpha \in [-0.3, 2.]$ in increments of 0.1. Since the relative error on s is small and consistent, we neglect the uncertainty on this parameter for simplicity and only consider the best value for each H, α and ν parameter combination. The resulting 3-dimensional likelihood surface was then interpolated to higher resolution with a second-order interpolation and used to calculate the probabilities of the climate sensitivities and projections corresponding to each parameter combination.

References

Aldrin M, Holden M, Guttorp Skeie RB, Myhred G, Berntsen TK (2012) Bayesian estimation of climate sensitivity based on a simple climate model fitted to observations of hemispheric temperatures and global ocean heat content. Environmetrics 23:253–271. https://doi.org/10.1002/env.2140

Bellouin N, Quaas J, Gryspeerdt E, Kinne S, Stier P, Watson-Parris D, Boucher O, Carslaw KS, Christensen M, Daniau A-L, Dufresne J-L, Feingold G, Fiedler S, Forster P, Gettelman A, Haywood JM, Lohmann U, Malavelle F, Mauritsen T, McCoy DT, Myhre G, Mülmenstädt J., Neubauer D, Possner A, Rugenstein M, Sato Y, Schulz M, Schwartz SE, Sourdeval O, Storelvmo T, Toll V, Winker D, Stevens B (2020) Bounding Global Aerosol Radiative Forcing of Climate Change. Rev Geophys 58, e2019RG000660. https://doi.org/10.1029/2019RG000660

Bodman RW, Jones RN (2016) Bayesian estimation of climate sensitivity using observationally constrained simple climate models. WIREs Clim Change 7:461–473. https://doi.org/10.1002/wcc.397

Bodman RW, Karoly DJ, Rayner PJ (2013) Uncertainty in temperature projections reduced using carbon cycle and climate observations. Nat Clim Change 3:725–729. https://doi.org/10.1038/NCLIMATE1903

Budyko MI (1969) The effect of solar radiation variations on the climate of the Earth. Tellus 21:611–619. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

Charney JG, Arakaw A, Baker DJ et al (1979) Carbon dioxide and climate: a scientific assessment. National Academy of Sciences Press, Washington, DC

Church JA, White NJ, Arblaster JM (2005) Significant decadal-scale impact of volcanic eruptions on sea level and ocean heat content. Nature 438(7064):74–7. https://doi.org/10.1038/nature04237

Clement AC, Seager R, Cane MA, Zebiak SE (1996) An ocean dynamical thermostat. J Clim 9:2190–2196. https://doi.org/10.1175/1520-0442(1996)009<2190:AODT>2.0.CO;2

Collins M, Knutti R, Arblaster J, Dufresne JL, Fichefet Y, Friedlingstein P, Gao X, Gutowski WJ, Johns T, Krinner G, Shongwe M, Tebaldi C, Weaver AJ, Wehner M (2013) Long-term Climate Change: Projections, Commitments and Irreversibility. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V,

Midgley PM (eds.)]. Cambridge University. Press, Cambridge, United Kingdom and New York, NY, USA
Cowtan K, Way RG (2014) Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends, Q.J.R. Meteorol. Soc. 140:1935–1944. https://doi.org/10.1002/qj.2297
Crowley TJ, Zielinski G, Vinther B, Udisti R, Kreutzs K, Cole-Dai J, Castellano E (2008) Volcanism and the little ice age. PAGES Newslett 16:22–23. https://doi.org/10.1029/2002GL0166335
Crucifix M (2006) Does the last glacial maximum constrain climate sensitivity? Geophys Res Lett 33:L18701. https://doi.org/10.1029/2006GL027137
Dijkstra H (2013) Nonlinear climate dynamics. Cambridge University Press, Cambridge, pp 357
Frederiksen HB, Rypdal M (2017) Long-range persistence in global surface temperatures explained by linear multibox energy balance models. J Clim 30:7157–7168. https://doi.org/10.1175/JCLI-D-16-0877.1
GISTEMP Team (2015) GISS surface temperature analysis (GISTEMP). NASA Goddard Institute for Space Studies. https://doi.org/10.1175/1520-0442(2002)015<3117:AOBEOT>2.0.CO;2Accessed 1 May 2017
Gallavotti G, Cohen EGD (1995) Dynamical ensembles in stationary states. J Stat Phys 80:931–970. https://doi.org/10.1002/env.2140
Geoffroy O, Saint-Martin D, Olivié DJ, Voldoire A, Bellon G, Tytéca S (2013) Transient climate response in a two-layer energy-balance model. part I: analytical solution and parameter calibration using CMIP5 AOGCM experiments. J Clim 26:1841–1857. https://doi.org/10.1002/env.2140
Goldblatt C, Watson AJ (2012) The runaway greenhouse: implications for future climate change, geoengineering and planetary atmospheres. Phil Trans R Soc A 370:4197–4216. https://doi.org/10.1002/env.2140
Gregory JM, Forster PM (2008) Transient climate response estimated from radiative forcing and observed temperature change. J Geophys Res 113:D23105. https://doi.org/10.1029/2008JD010405
Gregory JM, Stouffer RJ, Raper SCB, Stott PA, Rayner NA (2002) An observationally based estimate of the climate sensitivity. J Clim 15:3117–3121. https://doi.org/10.1175/1520-0442(2002)015%3c3117:AOBEOT%3e2.0.CO;2
Gregory JM, Andrews T, Good P et al (2016) Small global-mean cooling due to volcanic radiative forcing. Clim Dyn. https://doi.org/10.1007/s00382-016-3055-1
Haar A (1910) Zur Theorie der orthogonalen Funktionensysteme. Math Ann 69:331. https://doi.org/10.1002/env.2140
Hansen J (2004) Defusing the global warming time bomb. Sci Am 290(3):68–77
Hansen J, Ruedy R, Sato M, Lo K (2010) Global surface temperature change. Rev Geophys 48:RG4004. https://doi.org/10.1002/env.2140
Hansen J, Sato M, Kharecha P, von Scuckmann K (2011) Earth's energy imbalance and implications. Atmos Chem Phys 11:13 421–449. https://doi.org/10.5194/acp-11-13421-2011
Hasselmann K, Hasselmann S, Giering R, Ocana V, Storch HV (1997) Sensitivity study of optimal CO_2 emission paths using a simplified structural integrated assessment model (SIAM). Clim Change 37:345–386. https://doi.org/10.1002/env.2140
Hasselmann K, Sausen R, Maier-Reimer E, Voss R (1993) On the cold start problem in transient simulations with coupled atmosphere-ocean models. Clim Dyn 9:53–61
van Hateren JH (2013) A fractal climate response function can simulate global average temperature trends of the modern era and the past millennium. Clim Dyn 40:2651. https://doi.org/10.1029/2002GL0166335
Hegerl GC, Crowley TJ, Crowley Hyde WT, Frame DJ (2006) Climate sensitivity constrained by temperature reconstructions over the past seven centuries. Nature 440:1029–1032. https://doi.org/10.1002/env.2140
Held IM, Winton M, TakahashiK Delworth T, Zeng F, Vallis GK (2010) Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. J Clim 23:2418–2427
von der Heydt AS, Köhler P, van de Wal RSW, Dijkstra HA (2014) On the state dependency of fast feedback processes in (paleo) climate sensitivity. Geophys Res Lett 41:6484–6492. https://doi.org/10.1029/2002GL0166335
Huybers P, Curry W (2006) Links between annual, Milankovitch and continuum temperature variability. Nature 441:329–332. https://doi.org/10.1038/NCLIMATE1903
Hébert R, Lovejoy S (2015) Interactive comment on "Global warming projections derived from an observation-based minimal model" by K. Rypdal. Earth Syst Dynam Discuss 6:C944–C953
Hébert R, Lovejoy S (2018) Regional climate sensitivity and historical based projections to 2100. Geophys Res Lett. https://doi.org/10.1002/env.2140
Hébert R (2017) A Scaling Model for the Forced Climate Variability over the Anthropocene. Master Thesis, McGill University, Québec, Canada
Jansen E, Overpeck J, Briffa KR, Duplessy J-C, Joos F, Masson-Delmotte V, Olago D, Otto-Bliesner B, Peltier WR, Rahmstorf S, Ramesh R, Raynaud D, Rind D, Solomina O, Villalba R, Zhang D (2007) Palaeoclimate. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Climate change 2007: the physical science basis. Contribution of Working Group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge
Johansson DJA, O'Neill BC, Tebaldi C, Häggström O (2015) Equilibrium climate sensitivity in light of observations over the warming hiatus. Nat Clim Change 5:449–453. https://doi.org/10.1038/NCLIMATE1903
Kirtman B, Power SB, Adedoyin JA et al. (2013) Near-term Climate Change: Projections and Predictability. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
Klimont Z, Smith SJ, Cofala J (2013) The last decade of global anthropogenic sulphur dioxide: 2000–2011 emissions. Environ. Res. Lett. 8:014003. https://doi.org/10.1038/NCLIMATE1903
Knutti R, Rugenstein MAA (2015) Feedbacks, climate sensitivity and the limits of linear models. Phil Trans R Soc A 373:20150146. https://doi.org/10.1038/NCLIMATE1903
Lean JL, Rind DH (2008) How will Earth's surface temperature change in future decades? Geophys Res Lett 36:L15708. https://doi.org/10.1038/NCLIMATE1903
Levitus S, Antonov J, Boyer T (2005) Warming of the world ocean, 1955–2003. Geophys Res Lett 32:L02604. https://doi.org/10.1038/NCLIMATE1903
Lewis N, Crok M (2013) A sensitive matter, a report from the global warming policy foundation. pp 67
Li S, Jarvis A (2009) Long run surface temperature dynamics of an A-OGCM: The HadCM3 4xCO_2 forcing experiment revisited. Clim Dyn 33:817–825
Lindzen RS, Choi Y (2011) On the observational determination of climate sensitivity and its implications. Asia-Pacific J Atmos Sci 47:377. https://doi.org/10.1038/NCLIMATE1903
Lovejoy S (2013) What is climate? Eos Trans AGU 94:1–2. https://doi.org/10.1038/NCLIMATE1903
Lovejoy S (2014a) Return periods of global climate fluctuations and the pause. Geophys Res Lett 41:4704–4710. https://doi.org/10.1038/NCLIMATE1903

Lovejoy S (2014b) Scaling fluctuation analysis and statistical hypothesis testing of anthropogenic warming. Clim Dyn 42:2339–2351. https://doi.org/10.1038/NCLIMATE1903

Lovejoy S (2017) How accurately do we know the temperature of the surface of the earth? Clim Dyn 49:4089. https://doi.org/10.1002/wcc.397

Lovejoy S, de Lima MIP (2015) The joint space-time statistics of macroweather precipitation, space-time statistical factorization and macroweather models. Chaos 25:075410. https://doi.org/10.1002/wcc.397

Lovejoy S, del Rio Amador L, Hébert R (2015) The Scaling LInear Macroweather model (SLIM): using scaling to forecast global scale macroweather from months to decades. Earth Syst Dynam Discuss 6:489–545. https://doi.org/10.1002/wcc.397

Lovejoy S, Schertzer D (2010) Towards a new synthesis for atmospheric dynamics: space-time cascades. Atmos Res 96:1–52. https://doi.org/10.1016/j.atmosres.2010.01004

Lovejoy S, Schertzer D (2012) Haar wavelets, fluctuations and structure functions: convenient choices for geophysics. Nonlinear Processes Geophys 19:513–527. https://doi.org/10.5194/npg-19-1-2012

Lovejoy S, Schertzer D (2013) The weather and climate: emergent laws and multifractal cascades. Cambridge University Press, Cambridge, p 480

Lovejoy S, Schertzer D, Varon D (2013) Do GCM's predict the climate. or macroweather? Earth Syst Dyn 4:1–16. https://doi.org/10.1002/wcc.397

Lovejoy S, Varotsos C (2016) Scaling regimes and linear/nonlinear responses of last millennium climate to volcanic and solar forcings. Earth Syst Dyn 7:133–150. https://doi.org/10.1002/wcc.397

Lovejoy S, Schertzer D (2013a) Low-Frequency Weather and the Emergence of the Climate. In: Sharma AS, Bunde A, Dimri VP, Baker DN (eds) Extreme Events and Natural Hazards: The Complexity Perspective. American Geophysical Union (AGU), pp 231–254. https://doi.org/10.1029/2011GM001087

Lovejoy S, Del Rio Amador L, Hébert R (2017) Harnessing butterflies: theory and practice of the stochastic seasonal to interannual prediction system (StocSIPS). In: Tsonis A (ed) Advances in nonlinear geosciencs. Springer International Publishing. https://doi.org/10.1007/978-3-319-58895-7

Lovejoy S, Del Rio Amador L, Hébert R, De Lima I (2016) Giant natural fluctuation models and anthropogenic warming: Giant Natural Fluctuations. Geophys Res Lett. https://doi.org/10.1002/2016GL070428

Marshall J, Armour KC, Scott JR, Kostov Y, Hausmann U, Ferreira D, Shepherd TG, Bitz CM (2014) The ocean's role in polar climate change: asymmetric Arctic and Antarctic responses to greenhouse gas and ozone forcing. Philos T R Soc A 372:20130040. https://doi.org/10.1002/wcc.397

Marshall J, Scott J, Proshutinsky A (2017) "Climate response functions" for the Arctic Ocean: a proposed coordinated modelling experiment. Geosci Model Dev 10:2833–2848. https://doi.org/10.1002/wcc.397

Mastrandrea MD, Field CB, Stocker TF, Edenhofer O, Ebi KL, Frame DJ, Held H, Kriegler E, Mach KJ, Matschoss PR, Plattner GK, Yohe GW, and Zwiers FW (2010) Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties. Intergovernmental Panel on Climate Change (IPCC)

Matsumoto K (2007) Radiocarbon-based circulation age of the world oceans. J Geophys Res 112:C09004. https://doi.org/10.1029/2007JC004095

Meehl GA, Washington WM, Ammann CM, Arblaster JN, Wigley TML, Tebaldi C (2004) Combinations of Natural and Anthropogenic Forcings in Twentieth-Century Climate. J Clim 17:3721–3727. https://doi.org/10.1175/1520-0442(2004)017%3c3721:CONAAF%3e2.0.CO;2

Meinshausen M, Meinshausen N, Hare W, Raper SCB, Frieler K, Knutti R, Frame DJ, Allen MR (2009) Greenhouse-gas emission targets for limiting global warming to 2^oC. Nature 458:1158–1162. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

Meinshausen M, Smith SJ, Calvin K, Daniel SJ, Kainuma MLT, Lamarque JF, Matsumoto JF, Montzka SA, Raper SCB, Riahi K, Thomson A, Velders GJM, van Vuuren DPP (2011) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Clim Change 109:213–241. https://doi.org/10.1007/s10584-011-0156-z

Morice CP, Kennedy JJ, Rayner NA, Jones PD (2011) Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set. J Geophys Res 117:D08101. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

Myhre G, Shindell D, Brèon FM, Collins W, Fuglestvedt J, Huang J, Koch D, Lamarque JF, Lee D, Mendoza B, Nakajima T, Robock A, Stephens G, Takemura T, Zhang H (2013) Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA

Otto A, Otto FEL, Boucher O, Church J, Hegerl G, Forster PM, Gillett NP, Gregory J, Johnson GC, Knutti R, Lewis N, Lohmann U, Marotzke J, Myhre G, Shindell D, Stevens B, Allen MR (2013) Energy budget constraints on climate response. Nat Geosci 6:415–416. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

PALEOSENS Project Members (2012) Making sense of palaeoclimate sensitivity. Nature 491:683–691. https://doi.org/10.1038/nature11574

Poppick A, Moyer EJ, Stein ML (2017) Estimating trends in the global mean temperature record. Adv Stat Climatol Meteorol Oceanograp 3(1):33–53

Procyk R, Lovejoy S, Hébert R (2020) The Fractional Energy Balance Equation for Climate projections through 2100. Earth Syst Dyn Discuss. https://doi.org/10.5194/esd-2020-48

Ragone F, Lucarini V, Lunkeit F (2014) A new framework for climate sensitivity and prediction. Clim Dyn 46:1459. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

Ramaswamy V et al. (2001) Radiative forcing of climate change. Climate Change 2001: The Scientific Basis. Houghton JT et al. (Eds.) Cambridge University Press, pp 349–416

Rhein M, Rintoul SR, Aoki S, Campos E, Chambers D, Feely RA, Gulev S, Johnson GC, Josey SA, Kostianoy A, Mauritzen C, Roemmich D, Talley LD, Wang F (2013) Observations: Ocean. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM. Eds. Cambridge University Press

Ring MJ, Lindner D, Cross EF, Schlesinger ME (2012) Causes of the Global Warming Observed since the 19^{th} Century. Atmos Clim Sci 2(4):401–415. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

Rohde R, Muller RA, Jacobsen R, Muller E, Perlmutter S, Rosenfeld A, Wurtele J, Groom D, Wickham C (2013) A new estimate of the average earth surface land temperature spanning 1753 to 2011. Geoinfor Geostat Overv 1:1. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

Ruelle D (2009) A review of linear response theory for general differentiable dynamical systems. Nonlinearity 22:855–870

Rypdal K (2012) Global temperature response to radiative forcing: solar cycle versus volcanic eruptions. J Geophys Res 117:D06115. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

Rypdal K (2015) Global warming projections derived from an observation-based minimal model. Earth Syst Dynam Discuss 6:1789–1813. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

Rypdal M, Rypdal K (2014) Long-memory effects in linear-response models of Earth's temperature and implications for future global warming. J Clim 27:5240–5258. https://doi.org/10.1111/j.2153-3490.1969.tb00466.x

Rypdal K, Rypdal M (2016) Review of revised paper by Lovejoy and Varotsos. Earth Syst Dynam Discuss 7:133–150. https://doi.org/10.5194/esd-7-133-2016

Sato M (2012) Forcings in GISS climate model : stratospheric aerosol optical thickness. https://doi.org/10.1038/nature04237. Accessed Jan 2017

Schmidt GA, Shindell DT, Tsigaridis K (2014) Reconciling warming trends. Nature Geosci 7(3):158–160. https://doi.org/10.1038/nature04237

Schmittner A, Urban N, Shakun J, Mahowald N, Clark P, Bartlein P, Mix A, Rosell-Melé A (2011) Climate sensitivity estimated from temperature reconstructions of the last glacial maximum. Science 334(6061):1385–1388. https://doi.org/10.1038/nature04237

Schwartz SE (2007) Heat capacity, time constant, and sensitivity of Earth's climate system. J Geophys Res 112:D24S05. https://doi.org/10.1029/2007JD008746

Schwartz SE (2008) Reply to comments by G. Foster et al., R. Knutti et al., and N. Scafetta on "Heat capacity, time constant, and sensitivity of Earth's climate system". J Geophys Res. 113:D15105. https://doi.org/10.1029/2008JD009872

Sellers WD (1969) A Global Climatic Model Based on the Energy Balance of the Earth-Atmosphere System. J Appl Meteor 8:392–400. https://doi.org/10.1175/1520-0450(1969)008<0392:AGCMBO>2.0.CO;2

Sherwood S, Webb MJ, Annan JD, Armour KC, Forster PM, Hargreaves JC, Hegerl G, Klein SA, Marvel KD, Rohling EJ, Watanabe M, Andrews T, Braconnot P, Bretherton CS, Foster GL, Hausfather Z, Heydt AS von der, Knutti R, Mauritsen T, Norris JR, Proistosescu C, Rugenstein M, Schmidt GA, Tokarska KB, Zelinka MD (2020) An assessment of Earth's climate sensitivity using multiple lines of evidence. Rev Geophys. https://doi.org/10.1029/2019RG000678

Skeie RB, Berntsen T, Aldrin M, Holden M, Myhre G (2014) A lower and more constrained estimate of climate sensitivity using updated observations and detailed radiative forcing time series. Earth Syst Dyn 5:139–175. https://doi.org/10.1038/nature04237

Smith SJ, van Aardenne J, Klimont Z, Andres RJ, Volke A, Delgado Arias S (2011) Anthropogenic sulfur dioxide emissions: 1850–2005. Atmos Chem Phys 11:1101–1116. https://doi.org/10.1038/nature04237

Smith TM, Reynolds RW, Peterson TC, Lawrimore J (2008) Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006). J Clim 21:2283–2296. https://doi.org/10.1038/nature04237

Stenchikov G, Delworth TL, Ramaswamy V, Stouffer RJ, Wittenberg A, Zeng F (2009) Volcanic signals in oceans. J Geophys Res 114:D16104. https://doi.org/10.1038/nature04237

Stevens B (2015) Rethinking the lower bound on aerosol radiative forcing. J Clim 28(12):4794–4819. https://doi.org/10.1175/JCLI-D-14-00656.1

Wang YM, Lean JL, Sheeley NR Jr (2005) Modeling the Sun's magnetic field and irradiance since 1713. Astrophys J. 625:522–538. https://doi.org/10.1029/2002GL0166335

Wolfram Research Inc., Mathematica, Version 10.4, Champaign, IL (2016)

Yoshimori M, Hargreaves JC, Annan JD, Yokohata T, Abe-Ouchi A (2011) Dependency of feedbacks on forcing and climate state in physics parameter ensembles. J Clim 24:6440–6455. https://doi.org/10.1029/2002GL0166335

Yoshimori M, Watanabe M, Shiogama H, Oka A, Abe-Ouchi A, Ohgaito R, Kamae Y (2016) A review of progress towards understanding the transient global mean surface temperature response to radiative perturbation. Prog in Earth and Planet Sci 3:21. https://doi.org/10.1029/2002GL0166335

Zeng X, Geil K (2016) Global warming projection in the 21st century based on an observational data-driven model. Geophys Res Lett 43:10947–10954. https://doi.org/10.1029/2002GL0166335

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.