Estimating radiative forcing with a nonconstant feedback parameter and linear response

Hege-Beate Fredriksen[1,1], Maria Rugenstein[2,2], and Rune Graversen[1,1]

[1]Department of Physics and Technology, UiT The Arctic University of Norway
[2]Colorado State University, Fort Collins, USA

November 30, 2022

Abstract

A new algorithm is proposed for estimating time-evolving global forcing in climate models. The method is a further development of the work of Forster et al. (2013), taking into account the non-constancy of the global feedbacks. We assume that the non-constancy of this global feedback can be explained as a time-scale dependence, associated with linear temperature responses to the forcing on different time scales. With this method we obtain stronger forcing estimates than previously assumed for the representative concentration pathway experiments in the Coupled Model Intercomparison Project Phase 5 (CMIP5). The reason for the higher future forcing is that the global feedback parameter is more negative at shorter time scales than at longer time scales, consistent with the equilibrium climate sensitivity increasing with equilibration time. Our definition of forcing provides a clean separation of forcing and response, and we find that linear temperature response functions estimated from experiments with abrupt quadrupling of CO$_2$ can be used to predict responses also for future scenarios. In particular, we demonstrate that for most models, the response to our new forcing estimate applied on the 21st century scenarios provides a global surface temperature up to year 2100 consistent with the output of coupled model versions of the respective model.

Estimating radiative forcing with a nonconstant feedback parameter and linear response

Hege-Beate Fredriksen[1], Maria Rugenstein[2] and Rune Graversen[1,3]

[1]Department of Physics and Technology, UiT the Arctic University of Norway, Tromsø, Norway
[2]Colorado State University, Fort Collins, USA
[3]The Norwegian Meteorological Institute, Norway

Key Points:

- We present a new method for estimating radiative forcing and apply it to abrupt4xCO_2, 1%CO_2, historical, and future scenario experiments
- Including a time-scale dependent feedback parameter results in stronger forcing estimates for the 21st century
- The temperature responses to the new forcing are well described by a linear response

Corresponding author: Hege-Beate Fredriksen, `hege-beate.fredriksen@uit.no`

Abstract

A new algorithm is proposed for estimating time-evolving global forcing in climate models. The method is a further development of the work of Forster et al. (2013), taking into account the non-constancy of the global feedbacks. We assume that the non-constancy of this global feedback can be explained as a time-scale dependence, associated with linear temperature responses to the forcing on different time scales. With this method we obtain stronger forcing estimates than previously assumed for the representative concentration pathway experiments in the Coupled Model Intercomparison Project Phase 5 (CMIP5). The reason for the higher future forcing is that the global feedback parameter is more negative at shorter time scales than at longer time scales, consistent with the equilibrium climate sensitivity increasing with equilibration time. Our definition of forcing provides a clean separation of forcing and response, and we find that linear temperature response functions estimated from experiments with abrupt quadrupling of CO_2 can be used to predict responses also for future scenarios. In particular, we demonstrate that for most models, the response to our new forcing estimate applied on the 21st century scenarios provides a global surface temperature up to year 2100 consistent with the output of coupled model versions of the respective model.

1 Introduction

Diagnosing the magnitude of a climate forcing is necessary to determine the climate responses to this forcing. However, defining a clear separation between forcing and response is challenging, and no clear distinction exists (Sherwood et al., 2015). In this paper we attempt to apply a separation within a linear temperature response framework, incorporating also the possibility of globally nonconstant atmospheric feedbacks. We test this method on models participating in the Coupled Model Intercomparison Project Phase 5 (CMIP5).

In the most common forcing-feedback framework, the radiative imbalance at the top of the atmosphere (N) is described as

$$N = \lambda T + F, \tag{1}$$

where T is the temperature response, λ is the feedback parameter, and F is the radiative forcing, all evaluated as global means. According to this equation, forcing is the initial radiative imbalance, before the global mean surface temperature starts to respond. However, as discussed by Hansen et al. (2005); Richardson et al. (2019), there are many ways of defining the forcing, allowing various rapid adjustments before diagnosing the radiative imbalance. Forcing estimates are therefore method and model dependent. Some studies even consider multi-annual adjustments associated with ocean inertia (Williams et al., 2008; M. Rugenstein, Gregory, et al., 2016; Menzel & Merlis, 2019). A motivation for this study is therefore to find an estimation method aiming for a clean separation between forcing and response. By design, our method aims at finding the forcing estimates that are the most predictable for the surface temperature responses.

The uncertainties associated with forcing estimates are large, not only due to the different rapid adjustments between models (Smith et al., 2018), but also due to differences in the parameterizations of the radiative transfer (Soden et al., 2018). The instantaneous forcing spread contributes to about half of the total intermodel spread in forcing (Chung & Soden, 2015), and the remaining spread is largely due to fast cloud adjustments (Zelinka et al., 2013). These uncertainties have led to an effort aiming at better characterizing the forcing used for the new CMIP6 model versions (Forster et al., 2016; Pincus et al., 2016). These studies recommend using fixed-SST forcing, largely due to the reduced level of noise by this method as compared to regression-based estimates. Fixed-SST forcing estimates are made by diagnosing the top of atmosphere radiative imbalance after fixing the sea-surface temperatures and letting the atmosphere adjust. These effective forcings include rapid adjustments, e.g. atmospheric temperature and cloud adjustments, and are found to be better predictors of global surface temperature responses than instantaneous forcing estimates (Richardson et al., 2019). There is, however, substantial land warming in these simulations. Our main motivation is to improve forcing estimates based on already existing simulations, which can be used for models where fixed-SST forcing is unavailable, and to circumvent the problem of land warming in fixed-SST simulations.

In experiments with a time-varying forcing, forcing estimates may be even more uncertain than in idealized experiments with constant forcing. Forster et al. (2013), hereafter F13, computes forcing time series $F(t)$ by rearranging Eq. (1). Their method consists of first determining λ following the regression method of Gregory et al. (2004) using idealized step-forcing simulations, and then using time series of $N(t)$ and $T(t)$ from any experiment to compute what they call adjusted forcing:

$$F(t) = N(t) - \lambda T(t) \tag{2}$$

We note that adjusted forcing in F13 does not mean the same as adjusted forcing in Hansen et al. (2005), where the latter allows only fast stratospheric adjustments to take place before the forcing is estimated from the top of the atmosphere imbalance in an idealized step-forcing experiment. Forcing estimates based on regressions in a Gregory plot, such as in Andrews et al. (2012) and F13 are what Forster et al. (2016) refers to as regression-based methods, assuming a constant feedback parameter.

However, several recent studies have pointed out that λ is not a constant (Armour et al., 2013; Geoffroy, Saint-Martin, Bellon, et al., 2013; Andrews et al., 2015; Gregory & Andrews, 2016; Proistosescu & Huybers, 2017; M. Rugenstein et al., 2020). Armour et al. (2013) demonstrate that locally constant feedbacks can result in a globally time-dependent feedback parameter because the pace of sea surface temperatures (SST) equilibration depends on the location, weighting the local feedbacks differently with time. Other studies demonstrated that also locally, feedbacks change magnitude with equilibration time (e.g. Andrews et al., 2015; Andrews & Webb, 2018; M. Rugenstein, Caldeira, & Knutti, 2016; Proistosescu & Huybers, 2017; Dong et al., 2019, 2020) and also throughout the historical time period (Paynter & Frölicher, 2015; Gregory & Andrews, 2016; Armour, 2017; Marvel et al., 2018; Dessler, 2020). The tropical Pacific, the relative warming of midlatitude or global oceans to the West Pacific warm pool, the North Atlantic, and the mid- and high latitudes have all been suggested to influence global feedbacks (e.g. Winton et al., 2010; Trossman et al., 2016; Andrews & Webb, 2018; Dong et al., 2020; Zelinka et al., 2020). The mechanism most often invoked is the dependence of lower tropospheric stability on the ratio of local and remote SSTs. Regions warming faster than the West Pacific warm pool ? which sets the temperature of the free troposphere through deep convection ? show a reduced lower tropospheric stability, a decrease in low-cloud coverage, and thus, a strong cloud and net radiative effect at the top of the atmosphere (e.g. Zhou et al., 2016; Ceppi & Gregory, 2017). In the CMIP6 models, the shortwave cloud feedbacks in the extratropics appear to be more important for the nonconstancy of λ than clouds in the tropics (Zelinka et al., 2020; Bacmeister et al., 2020), but the relatively short record of global cloud observations makes it difficult to assess cloud modeling against the observations (Loeb et al., 2020). Other studies highlight the dependence of feedbacks on temperature and radiative forcing (Meraner et al., 2013; Rohrschneider et al., 2019; Bloch-Johnson et al., 2021).

The nonconstancy of λ implies that the forcing definition in Eq. (2) is ambiguous. This is particularly apparent for strong temperature responses, when λT more strongly affects the determination of the value of F. Here the magnitude and time-dependence of λ are particularly important. Larson and Portmann (2016) demonstrated for instance that λ obtained from regressions in the first 20 yr time period of abrupt4xCO_2 gives higher forcing estimates compared to regressions in 150 yr time period. This is one of several reasons why Forster et al. (2016) recommends fixed-SST methods instead of regression methods to determine the forcing.

We explore how an alternative definition of effective forcing with a time-scale dependent λ differs from estimates by F13. To compute these alternative estimates, we decompose the temperature response assuming it responds linearly to the forcing, and we demonstrate that the linear temperature response to the new forcing is close to the modelled temperature response in future scenarios for most CMIP5 models. By a linear response, we mean the temperature response determined from a linear non-homogeneous system of differential equations, whose solution can be expressed as a convolution between a Green's function and the forcing. Our results suggest that this forcing estimate appears more appropriate for estimating temperature responses using linear response models than previous estimates.

Our method is an iterative routine, starting with the F13 estimate of forcing, then computing the linear response to this forcing, which is further used to compute a new forcing estimate, etc., until convergence to a final forcing estimate is obtained. Theory and methods are described in Section 2, and the results are shown in Section 3. In Section 4, we discuss the assumptions made in our method, and how it compares to other forcing estimates, before we conclude in Section 5.

2 Theory and methods

The time-scale dependence of λ is analysed by making use of the same decomposition as in Proistosescu and Huybers (2017), hereafter PH17. While PH17 use the method to better understand estimates of climate sensitivity, we are interested in the intersect of the fit with the vertical axis, the initial radiative imbalance. We also estimate parameters using a different approach, mainly because our method simplifies the comparison to methods based on single regression estimates in Gregory plots. The equations that will be presented in this section provide interpretations of the different λ's that may appear in a Gregory plot, as well as interpretations of "forcing estimates" based on regressions on decadal to centennial time scales. The method is based on the assumption that the temperature response can be decomposed into a sum of K components $T = \sum_{n=1}^{K} T_n$, where each component is the exponential temperature response to the forcing on the time scale τ_n [yrs],

$$T_n(t) = c_n \exp(-t/\tau_n) * F(t). \tag{3}$$

The $*$ denotes a convolution, and the factors c_n $[\frac{Km^2}{W}]$ are the amplitudes of the temperature responses per unit forcing. As further explained in the next subsection, this temperature decomposition can be interpreted as either approximating different global-scale processes (such as mixed-layer versus deep ocean responses to forcing) or as regions responding with different pace to the forcing (such as the tropics in general versus regions of upwelling or deep ocean convection). c_n therefore depends on both the feedbacks and thermal inertia associated with different regions, and the fraction of the global area involved in the response at time scale τ_n.

Furthermore, the method assumes that constant feedback parameters λ_n exist, with $n = 1, ..., K$ associated with each time scale, such that the terms in Eq. (1) can be decomposed into the following sums:

$$N(t) = \sum_{n=1}^{K} N_n(t) = F(t) + \sum_{n=1}^{K} \lambda_n T_n(t) = F(t) + \lambda(t)T(t) \tag{4}$$

By rewriting Eq. (4), PH17 noted that the time-variation of $\lambda(t)$ can be explained as a weighted average of the feedbacks associated with different components $T_n(t)$ of the global temperature:

$$\lambda(t) = \frac{\sum_{n=1}^{K} \lambda_n T_n(t)}{\sum_{n=1}^{K} T_n(t)} \tag{5}$$

We note that in a 4xCO_2 experiment, we define the forcing to be a constant, and the slope $\lambda(t)$ must be interpreted as the slope of a line drawn between the fixed forcing F and a point $(T(t), N(t))$. This slope may differ from a linearization around a point $(T(t), N(t))$ by regressing a range of points (see discussion on feedback definitions in M. A. A. Rugenstein and Armour (2021)).

Armour et al. (2013) suggested a similar decomposition, but interpreted the components as locally constant feedbacks multiplied by local temperatures with different time evolution. However, recent studies suggest that non-local feedbacks are also important (Andrews et al., 2015; Zhou et al., 2016; Dong et al., 2019; Bloch-Johnson et al., 2020), meaning that temperature changes in one region, and in particular the West Pacific, can influence feedbacks globally.

2.1 Linear model and response

A simple model of temperature changes in the climate system can be constructed by considering different boxes or components that store and exchange energy. If assuming that all anomalous heat fluxes are linearly related to temperature anomalies in the system, the heat uptake in all boxes can be written into a linear non-homogeneous system

$$\mathbf{C}\frac{d\mathbf{T}(t)}{dt} = \mathbf{K}\mathbf{T}(t) + \mathbf{F}(t) \tag{6}$$

By choosing the vector of temperature change components $\mathbf{T}$ to be K-dimensional, the system describes K components that will respond on K different time scales, and the vector $\mathbf{F}$ the atmospheric forcing acting directly on each component. The vector $\mathbf{F}$ could in principle contain different forcings in different regions. The heat capacities $[\frac{Wyr}{m^2K}]$ associated with each component are along the diagonal of the diagonal $K \times K$ matrix $\mathbf{C}$, and coefficients for heat exchange between components and heat loss to the atmosphere $[\frac{W}{m^2K}]$ constitute the matrix $\mathbf{K}$. The left-hand side of this equation describes the heat uptake of each component, and the sum of all heat uptakes must equal the net radiative imbalance N. In this sum of all components, all fluxes between components cancel out, and the sum reduces to Eq. (4).

Linear systems like this have been widely studied, often using one, two or three boxes (e.g. Geoffroy, Saint-Martin, Olivié, et al., 2013; Fredriksen & Rypdal, 2017). Symmetric matrices $\mathbf{K}$ will describe diffusive heat fluxes depending on the temperature difference between two boxes, and feedback parameters will appear on its diagonal. Non-symmetric parts may be due to the dependence of temperature anomalies in one box only. For instance change in sinking processes due to temperature anomalies in the North Atlantic regarded as one box, may by mass continuity induce horizontal mass and hence energy fluxes from adjacent ocean basins regarded as other boxes, independent of the temperature change in these boxes. $\mathbf{K}$ may also incorporate heat fluxes to the deep ocean if assuming they can be modelled as linear functions of temperature components (e.g. Held et al., 2010; Geoffroy, Saint-Martin, Olivié, et al., 2013).

By applying the method variation of parameters, it can be shown that the solution to Eq. (6) is (see the supporting information):

$$\mathbf{T}(t) = \int_{-\infty}^{t} e^{(t-s)\mathbf{C}^{-1}\mathbf{K}}\mathbf{C}^{-1}\mathbf{F}(s)\, ds, \tag{7}$$

showing that the temperature at time t is a response to the forcing experienced at all previous times s. If the matrix $\mathbf{C}^{-1}\mathbf{K}$ has only negative eigenvalues, $-1/\tau_n$, the solution for each temperature component $T_k(t)$ will be a weighted sum of K exponential responses to the global average forcing F with time scales τ_n (where the weights β_n are determined by eigenvalues, eigenvectors, and heat capacities),

$$T_k(t) = \int_{-\infty}^{t} \sum_{n=1}^{K} \beta_n e^{(s-t)/\tau_n} F(s)\, ds \tag{8}$$

Furthermore, the global surface temperature is a weighted average of the components $T_k(t)$:

$$\overline{T(t)} = \sum_{n=1}^{K} c_n \int_{-\infty}^{t} e^{(s-t)/\tau_n} F(s)\, ds \tag{9}$$

where we define the new weights c_n to be an area-weighted average of the weights β_n. If the forcing is not the same in all regions, Eq. (9) is still valid if the regional forcings are scaled versions of the global average forcing. We recognize Eq. (9) as a convolution between a Green's function $G(t)$ and a forcing $F(t)$, consistent with Eq. (3): $T(t) = G(t) * F(t) = \int_{-\infty}^{t} G(t-s)F(s)ds$, with $G(t) = \sum_{n=1}^{K} G_n(t) = \sum_{n=1}^{K} c_n \exp(-t/\tau_n)$, assuming negative eigenvalues.

2.2 Estimating linear response in abrupt $4xCO_2$ experiments

To simplify the estimation of parameters of these responses (time scales τ_n and amplitudes c_n), we start by fixing the time scales, such that T and N depend linearly on the remaining parameters c_n. We find that the exact choice of time scales is not important, as long as we choose them well separated, and within the range of expected time scales. Annual time scales are important over land and shallow ocean areas, while decadal and centennial time scales are particularly important in ocean regions with mixing to the deeper oceans, and hence higher thermal inertia. Following PH17, we use three different time scales. They find three time scales to be the smallest number that well describes the temperature responses. In addition as explained later, we will assume the existence of a fourth time scale explaining slower temperature responses than can be observed in the records studied in this paper.

We analyse data from 21 CMIP5 models, available at `https://esgf-node.llnl.gov/projects/cmip5/`. The variables used are global annual averages of surface air temperatures (tas), and net top-of-atmosphere radiation, computed as the difference between incoming shortwave radiation and outgoing longwave and shortwave radiation (rsdt - rlut - rsut). To minimize the effect of possible model drifts, the temperature $T(t)$ and the variables used to compute the net top of atmosphere radiation $N(t)$ time series are defined as deviations from linear trends in the corresponding time period of the control run (trend values for the abrupt4xCO_2 period are given in Table S1, and are very small). With this definition we also avoid non-zero means of $N(t)$ in equilibrium, which is the case for many models (Forster et al., 2013).

The shortest time scale τ_1 is chosen to be a random number between 1 and 6 years, the second time scale τ_2 is a random factor between 5 and 10 multiplied by τ_1, and the third is a randomly chosen time scale between 80 and 1000 years. The random choice is done 1000 times for each model, and finally, for each model, we keep the set of τ_n with the best (least squares) fit to the modelled temperature evolution for 150 years after an abrupt quadrupling of CO_2. The resulting parameters are dependent on the length of the time series used. If using longer time series the longest time-scale responses may change the most, but these are also the least important for our 21st century analyses.

The temperature response for these step-forcing experiments can be found by computing the integrals in Eq. (9) with a constant forcing F_{4xCO_2} for $t > 0$. This integral results in

$$T_{4\text{xCO}_2}(t) = \sum_{n=1}^{K} a_n(1 - e^{-t/\tau_n}) \tag{10}$$

where $a_n = c_n \tau_n F_{4\text{xCO}_2}$ is the equilibrium temperature of each component, and the equilibrium climate sensitivity (ECS) is defined as $\frac{1}{2}\sum_{n=1}^{K} a_n$ (equilibrium response to a doubling of CO_2).

The expression for N is derived as:

$$\begin{aligned}
N_{4\text{xCO}_2}(t) &= F_{4\text{xCO}_2} + \sum_{n=1}^{K} (\lambda_n T_n(t)) \\
&= F_{4\text{xCO}_2} + \sum_{n=1}^{K} \left(\lambda_n a_n (1 - e^{-t/\tau_n})\right) \\
&= F_{4\text{xCO}_2} + \sum_{n=1}^{K} \lambda_n a_n - \sum_{n=1}^{K} \lambda_n a_n e^{-t/\tau_n} \\
&= -\sum_{n=1}^{K} \lambda_n a_n e^{-t/\tau_n}
\end{aligned}$$

where we in the last step set that $F_{4\text{xCO}_2} + \sum_{n=1}^{K} \lambda_n a_n = 0$, due to the constraint that $N \to 0$ when $t \to \infty$. Introducing the notation that $b_n = -a_n \lambda_n$ gives us $N_{4\text{xCO}_2}(t) = \sum_{n=1}^{K} N_n(t) = \sum_{n=1}^{K} b_n e^{-t/\tau_n}$, and $F_{4\text{xCO}_2} = -\sum_{n=1}^{K} \lambda_n a_n = \sum_{n=1}^{K} b_n$.

The parameters a_n, b_n could be found using linear regression, but that does sometimes violate the physical assumption that these should have the same sign as the forcing. Therefore we have used the non-negative least squares algorithm to ensure positive parameters. This is used only for finding a_n, and the resulting temperature responses are shown in Figure 1 b). This method could in principle also have been used to find b_n, but this does not seem to provide a sufficiently good fit on the short scales. Instead, λ_n are determined in a Gregory plot, and then used to compute $b_n = -\lambda_n a_n$.

2.3 Algorithm for estimating $\boldsymbol{\lambda_n}$

The λ_n, $n = 1, \ldots, K$ are all determined from linear fits in a Gregory plot, as shown in Figure 1 a). We start with estimating λ_3 corresponding to time scale τ_3, then we estimate λ_2, and finally λ_1. We assume that the sum $\sum_{n=1}^{3} a_n$ underestimates the equilibrium response, since the sum excludes the response on the multi-millennial scale τ_4. However, we assume τ_4 is so large that we can make the following approximations for $t \leq 150$ years:

$$T_4(t) = a_4 \left(1 - e^{-t/\tau_4}\right) \approx 0 \tag{11}$$

$$N_4(t) = b_4 e^{-t/\tau_4} \approx b_4 \tag{12}$$

Hence $T(t) \approx \sum_{n=1}^{3} T_n(t)$ and $N(t) \approx b_4 + \sum_{n=1}^{3} N_n(t)$, where b_4 could be interpreted as a constant heat flux going into the deeper oceans, hereby not leading to surface warming on short time scales. We made the somewhat arbitrary choice of setting $\tau_4 = 5000$ years, and assume $\lambda_4 = \lambda_3$. The results are not sensitive to the choice of τ_4 as long as the approximations in Eqs. (11) and (12) hold. In the 150 year long runs considered in this paper, we have no information about λ_4, but longer runs show that the feedback parameter changes little on the longer time scales (M. Rugenstein et al., 2020).

Determining λ_3**:** We consider only temperatures larger than the equilibrium temperature of the first two components, such that $T_1(t) + T_2(t) \approx a_1 + a_2$, and we have: $N(t) \approx -\lambda_3(a_3 - T_3(t)) + b_4$. The total temperature is therefore approximated by $T(t) \approx$

$a_1+a_2+T_3(t)$, resulting in $N(t) \approx -\lambda_3(a_1+a_2+a_3-T(t))+b_4$. This shows that N is approximately a linear function of T with slope λ_3 for $T > a_1 + a_2$. Therefore, λ_3 is computed by linear regression of these points, and the equilibrium temperature found by following this line until $N = 0$. This equilibrium estimate should be higher than $\sum_{n=1}^{3} a_n$, and the difference is a_4. Whenever the unphysical result $a_4 < 0$ is obtained, we exclude the chosen time scales from our analysis.

Determining λ_2: First we subtract our estimates of $T_3(t)$, $T_4(t)$ and $N_3(t)$, $N_4(t)$ from the time series $T(t)$ and $N(t)$, respectively. We then obtain estimates of $T_1(t)+T_2(t)$ and $N_1(t)+N_2(t)$, and these points are the dark gray dots in Figure 1a). For $a_1 < T_1(t) + T_2(t) < a_1 + a_2$, $T_1(t) + T_2(t)$ is approximately $a_1 + T_2(t)$, and should equal the equilibrium value $a_1 + a_2$ when $N_1(t) + N_2(t) = 0$. In this range, $N_1(t) + N_2(t) \approx -\lambda_2(a_2 - T_2(t))$, approximately linearly related to $T_1(t)+T_2(t)$. Therefore, λ_2 is estimated using a least-squares algorithm forcing the linear fit to go through the point $(a_1+a_2, 0)$.

Determining λ_1: We subtract estimates of $(T_2(t), N_2(t))$ from the dark gray dots to obtain estimates of $T_1(t)$ and $N_1(t)$ (light gray dots in Figure 1). We have now $N_1(t) \approx -\lambda_1(a_1-T_1(t))$, and we can, as previously, use least squares to compute λ_1, forcing the linear fit to pass the point $(a_1, 0)$.

In the least squares fits, we also include an upper time limit to the set of points to be included in the calculation. This limit is set to the first time step after reaching 99% of the equilibrium temperature of the component of interest. In this way, our slope is associated with the response on the particular time scale τ_n, and little influenced by the fluctuations around the equilibrium values. Changing this limit to e.g. 90% or 95% has only minor effects on the results. Feedback parameters associated with fluctuations around the base state, or more precisely, radiative restoring coefficients are studied in several papers (Colman & Power, 2010; Colman & Hanson, 2013; Lutsko & Takahashi, 2018; Bloch-Johnson et al., 2020). Depending on the model, they can be similar or different from those associated with the final fluctuation after a quadrupling of CO_2 (M. Rugenstein et al., 2020), and they may also differ from feedbacks associated with forced responses (e.g. Zhou et al., 2015; Dessler & Forster, 2018).

When all a_n, λ_n are estimated, we compute $b_n = -\lambda_n a_n$ and we finally have our estimate of $F_{4\text{xCO}_2} = \sum_{n=1}^{4} b_n$. That is, the sum of the initial radiative imbalance of all 4 components.

2.4 New estimates of effective forcing time series

Using our parameter estimates from the previous subsections, we can for any experiment use the global mean evolutions of $T(t)$ and $N(t)$ to compute a new estimate of the effective forcing as follows:

1. Compute $F(t)$ using F13's method (a single estimate of λ), and take this as the initial estimate of the effective forcing.
2. Use this forcing estimate and amplitudes $c_n = \frac{a_n}{\tau_n F_{4\text{xCO}_2}}$ estimated from 4xCO_2 experiments to compute the components $T_n(t)$ from Eq. (3) by performing convolution integrals.
3. A new estimate of $F(t)$ can then be computed as:
$$F(t) = N(t) - \sum_n \lambda_n T_n(t) \tag{13}$$
4. Repeat steps 2-3 until convergence of $F(t)$. We have used 20 iterations.

We demonstrate how the method can be applied to study the forcing for 1% CO_2 experiments, the historical period and the four representative concentration pathways (RCPs) RCP2.6, RCP4.5, RCP6.0 and RCP8.5.

Table 1. Estimated parameters, where we define F_{2x} and T_{2x} to be half the forcing and equilibrium temperature estimated for a quadrupling of CO_2. The parameters in parentheses $(-\lambda)$, (F_{2x}) and (T_{2x}) are estimated from a single linear regression over years 1-150 in a Gregory plot. The results differ slightly from the numbers reported from the Gregory method by Andrews et al. (2012), possibly because of minor differences in the way global annual average values are constructed. For one model (GFDL-ESM2G) the best fit consists of two exponential responses, where we estimate $a_2 = 0$ and report $\lambda_2 = b_2/a_2$ as 'NaN'.

	τ_1	τ_2	τ_3	$-\lambda_1$	$-\lambda_2$	$-\lambda_3$	$(-\lambda)$	F_{2x}	(F_{2x})	T_{2x}	(T_{2x})
ACCESS1-0	2.43	12.79	231.10	1.30	1.12	0.56	0.78	3.72	2.97	4.33	3.83
ACCESS1-3	1.13	5.80	150.10	1.46	1.30	0.56	0.82	3.60	2.89	4.12	3.53
CanESM2	2.86	26.39	279.11	1.30	1.01	0.91	1.04	4.24	3.83	3.83	3.69
CCSM4	1.04	5.52	197.28	1.32	1.77	0.90	1.18	4.02	3.47	3.19	2.94
CNRM-CM5	1.45	10.71	392.15	1.38	1.09	1.22	1.14	3.87	3.71	3.20	3.25
CSIRO-Mk3-6-0	1.62	11.29	308.98	1.86	1.12	0.41	0.63	3.94	2.58	4.94	4.08
GFDL-CM3	3.28	32.58	98.81	1.21	0.80	0.63	0.75	3.61	2.99	4.24	3.97
GFDL-ESM2G	2.98	17.50	291.97	1.76	NaN	0.90	1.29	3.65	3.09	2.67	2.39
GFDL-ESM2M	1.03	5.77	240.02	1.52	1.58	1.22	1.38	3.58	3.36	2.52	2.44
GISS-E2-H	1.56	10.43	186.27	2.02	1.83	1.40	1.65	4.21	3.81	2.39	2.31
GISS-E2-R	1.51	10.61	232.40	2.98	1.02	1.42	1.79	5.09	3.78	2.25	2.11
HadGEM2-ES	1.01	8.39	367.62	1.96	0.89	0.35	0.63	4.02	2.90	5.91	4.61
inmcm4	1.02	5.65	597.43	1.90	1.48	1.28	1.43	3.18	2.98	2.14	2.08
IPSL-CM5A-LR	1.72	16.54	163.83	1.03	0.84	0.58	0.75	3.43	3.10	4.55	4.13
IPSL-CM5B-LR	1.21	8.01	80.30	2.39	1.11	0.91	1.02	3.64	2.64	2.68	2.60
MIROC-ESM	1.78	11.32	266.35	1.96	0.92	0.68	0.91	5.37	4.26	5.21	4.67
MIROC5	2.77	15.17	89.28	1.72	1.43	1.36	1.52	4.38	4.13	2.80	2.72
MPI-ESM-LR	1.81	9.20	202.56	1.30	1.50	0.86	1.13	4.53	4.09	3.91	3.63
MPI-ESM-MR	1.02	6.23	158.54	2.27	1.45	0.94	1.18	5.15	4.07	3.67	3.46
MRI-CGCM3	1.42	11.61	233.73	2.22	1.34	0.96	1.25	4.05	3.24	2.76	2.60
NorESM1-M	1.75	9.34	273.12	1.87	1.52	0.78	1.11	3.88	3.10	3.17	2.80
Model mean	1.73	11.73	231.26	1.73	1.28	0.90	1.12	4.04	3.38	3.50	3.20
Standard deviation	0.69	6.58	115.35	0.45	0.30	0.31	0.31	0.56	0.50	1.02	0.78

3 Results

The results of the linear response fit for $T(t)$ and $N(t)$ following an abrupt quadrupling of CO_2 are given for the model NorESM1-M in Figure 1, and the estimated parameters are listed in Table 1. We note from Figure 1a) that both the forcing and equilibrium temperature estimates are higher than when obtained from a straight line fit. The narrow spread of the light blue lines also indicate that the choice of time scales is of little importance, and hence not affecting the overall conclusions. Similar plots are shown for the other models listed in Table 1 in the Supporting information. The uncertainty in both the forcing estimate and ECS estimate vary substantially from model to model. Models with a rapid initial warming, such as GISS-E2-R, have fewer points constraining the regression estimate for the shortest time scale, implying larger uncertainty of the forcing.

An overview of all our estimates of the 4xCO_2 forcing is presented in Figure 2. In addition, we compare our forcing estimates to regression estimates done for years 1-20 and years 1-150. In all except one model, the 1-20 year regression gives a higher estimate

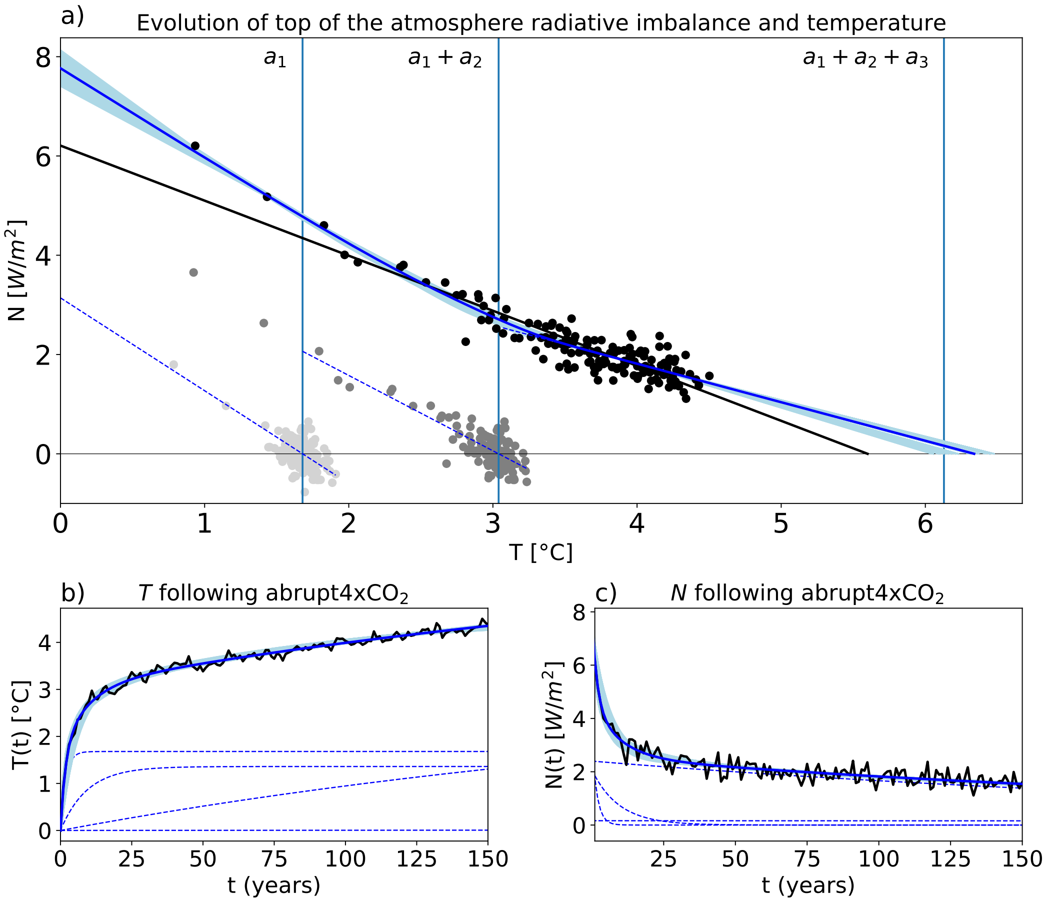

Figure 1. Results for NorESM1-M: a) The black dots and line is a conventional Gregory plot, the light blue lines (recognized as light-blue shading) are our fits to the black points with 1000 different choices of time scales, and the dark blue fit is when using the best (least squares) fits for the temperature in b). Vertical blue lines are the sums of equilibrium temperatures $\sum_{n=1}^{m} a_n, \quad m = 1, 2, 3$. The dark (light) gray dots are N vs. T after subtracting components associated with the third (and second) time scales, and the dashed blue lines are fits to these dots. b) The black curve is the climate model temperature output, and the light blue curves are best fits to the modelled temperature using 1000 different choices of time scales. The dark blue curve is the best fit, and the dashed blue curves are the individual components due to the four time scales which are summed to obtain this fit. c) As panel b), but for the change in net top of the atmosphere radiation.

than the 1-150 year regression. And in all but two models, our best forcing estimate is even higher than estimates obtained from regression of years 1-20. The fixed-SST 4xCO_2 forcing estimates reported by Andrews et al. (2012) are higher than regression-based estimates over 150 years for most of the models where this is available, but smaller than our new forcing estimates.

Using global annual means of $N(t)$ and $T(t)$ from the coupled models, we continue by testing the algorithm described in Section 2.4 for 1% CO_2 experiments. In these experiments we expect a linearly increasing forcing, because to first order, for small increases in CO_2 the forcing depends logarithmically on the CO_2 concentration (Myhre et al., 1998)

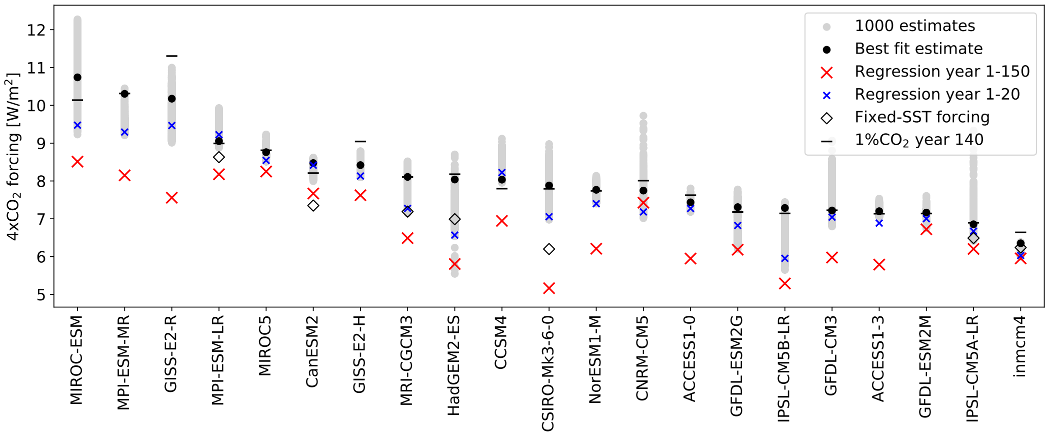

Figure 2. A summary of the $4xCO_2$ forcing estimates made in this paper, to provide an overview of their uncertainties and how they compare to regression estimates. The $1\%CO_2$ estimates are the linear fits to the estimated 1% CO_2 forcing time series evaluated in year 140, the time of quadrupling (except for the models GFDL-ESM2G and GFDL-ESM2M, where the estimates are instead twice the doubling estimates in year 70). Fixed-SST estimates are taken from Andrews et al. (2012) for the models where these are available.

(see limitations of this discussed in Byrne and Goldblatt (2014); Etminan et al. (2016); Gregory et al. (2015); Bloch-Johnson et al. (2021)). A linear increase is indeed what we observe for NorESM1-M in Figure 3, for both the initial and the new forcing estimate. For the new estimate we note a high consistency between the climate model temperature output and the linear response to the forcing. This result suggests that our method can successfully construct forcing estimates that well predicts the surface temperature responses. Results for other models are similar, and are shown in the supporting information. After 140 years of 1% increase the CO_2 concentration is quadrupled, and the linear fit to the 1% CO_2 forcing time series evaluated in year 140 is yet another estimate of F_{4xCO_2}, which we include in Figure 2. For most models this estimate is close to our best estimates determined from abrupt4xCO_2 experiments.

Next we apply the algorithm to the historical and RCP experiments to compute forcing estimates for the time period 1850 - 2100. Our new forcing estimate for the historical and RCP8.5 experiment for NorESM1-M diverges from the forcing estimate using a single feedback parameter when approaching the end of the 21st century (Figure 4a). The difference is about 2 W/m^2 in 2100, and smaller differences are seen during the historical period. As a result, the sum of the linear temperature responses we compute by convolving with the two forcing estimates according to Eq. (3) also diverge (dashed curves in Figure 4b), reaching a difference of almost 1 degree in year 2100. We note that the linear response to our new forcing (dashed blue curve) is remarkably close to the climate model temperature output, indicating that our alternative forcing definition and linear response assumption is the better approximation for this model. This result holds also for the other RCP scenarios (see Figures S109 - S111 in the supporting information).

By computing the time-varying feedback parameter $\lambda(t)$ using Eq. (5), we find a generally higher magnitude than the single estimate of λ. During the historical period the global temperature response is often close to 0, causing high fluctuations in the estimated $\lambda(t)$. The estimate becomes more stable for the future scenarios, where we find a slowly decreasing magnitude of $\lambda(t)$, consistent with a higher weighting of the slow responses. For all years in the experiment, the magnitude of $\lambda(t)$ is still considerably higher

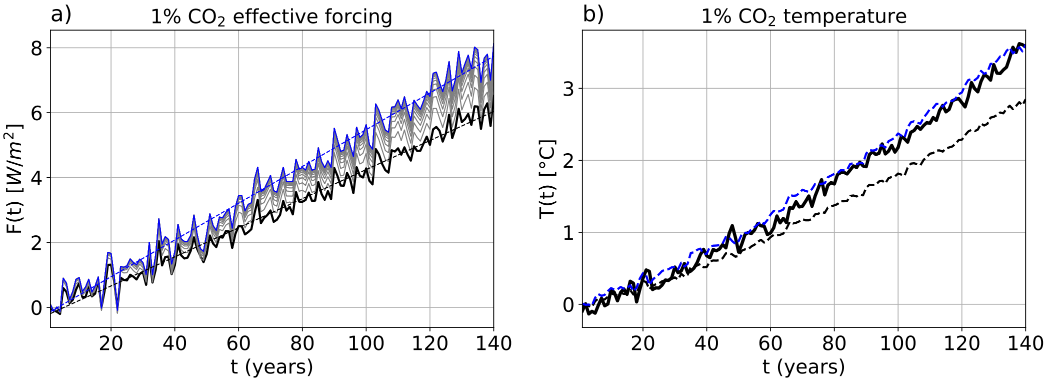

Figure 3. Results for NorESM1-M: a) The black curve is the forcing computed as in F13, using a single and constant value of λ. The gray curves are the iterations of the algorithm described in section 2.4, using three different λ's, and the blue curve the new forcing obtained by convergence after 20 iterations. The dashed lines are linear fits to the initial and final forcing estimates. b) The thick black curve is the modelled temperature change, and the black and blue dashed curves the linear responses to the black and blue curves in a), applying the same response function as estimated in Figure 1 b).

than the single regression estimate, hence the term $-\lambda(t)T(t)$ gives a higher contribution to the forcing estimate. This effect on the forcing is however only visible when the temperature response is strong.

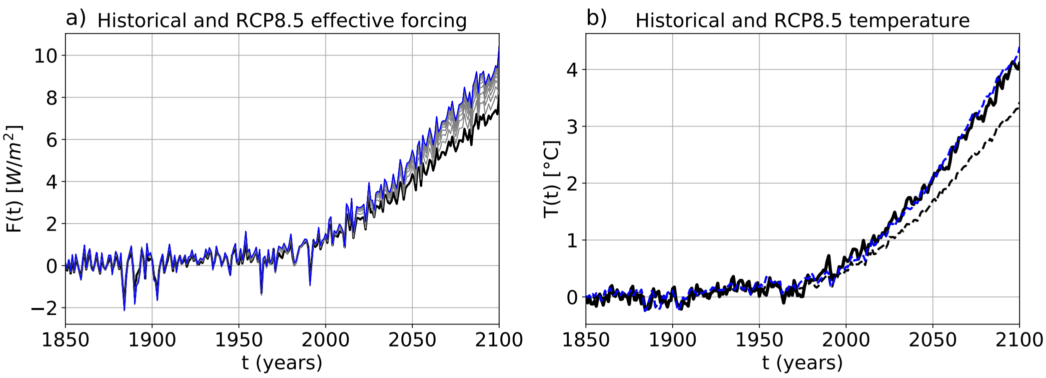

Figure 4. Similar to Figure 3, but for NorESM1-M historical and RCP8.5 experiment.

Repeating the analysis in Figure 4 for all models and RCP scenarios shows that the method presented here works well for many models, but not all (Figures in supporting information). A summary of these results are given in Figure 5, where panel a) compares the mean estimated forcing over years 2091-2100 using the two different methods. The names of the scenarios are constructed to reflect the intended forcing in the end of the 21st century (van Vuuren et al., 2011), and these forcing levels are also shown for comparison. We find that model estimates using F13's method are centered at lower values, while our new forcing estimates are centered close to or slightly above the intended levels. However, the intended forcing is difficult to prescribe as it depends on model-specific fast adjustments, so we can only expect these to be approximate values. The GISS-E2-

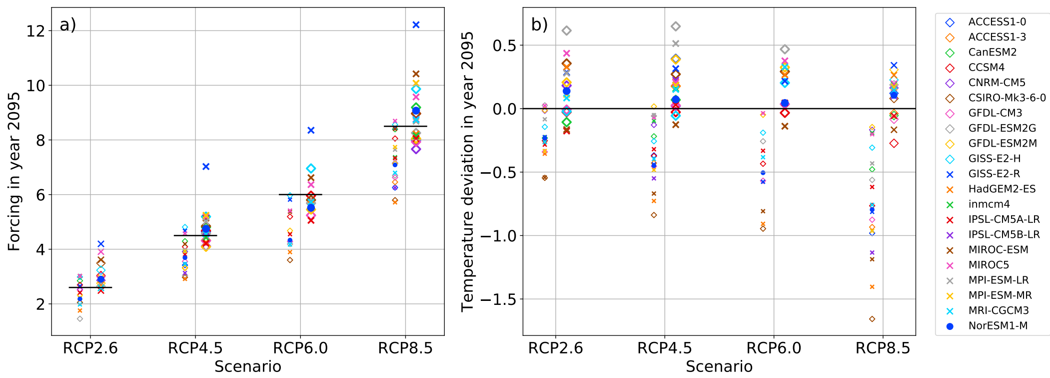

Figure 5. Estimated year 2095 forcing (a) and temperature difference between the result of the linear response and the climate model output (b). For each scenario, the left points show results using F13's method, and the right points show results using our method. Values in year 2095 are computed by averaging over the ten years 2091-2100. The forcing levels 2.6, 4.5, 6.0 and 8.5 W/m^2 are also shown for reference in a) as horizontal black lines.

R model might be considered as an outlier, and its response to abrupt4xCO_2 is also visually different from the other models.

Consistent with the increase in forcing level, we observe an increase in the estimated linear temperature responses in panel b). The linear responses to F13 forcing are mostly lower than the climate model temperature output, and the responses to our new forcing are scattered around, with a center slightly above. Some deviation from the climate model temperature is expected due to internal variability, and to assess this expected uncertainty, we refer to the model spread of the Community Earth System Model Large Ensemble (CESM-LE) (Kay et al., 2015). Here 40 model simulations for the historical + RCP8.5 scenarios from the same model show a model spread of around 0.4 K, which is attributed to internal variability.

Using F13 forcing, the linear response is within these uncertainties for only a few models. For the new forcing, more models are within this uncertainty range than outside. There are also other uncertainties to consider, e.g. associated with our parameter estimation method, probably making the expected uncertainty interval larger than 0.4 K. The uncertainty due to internal variability is also model-dependent (Olonscheck et al., 2020), hence it is difficult to identify models where our linear response hypothesis and forcing estimation method fail.

We note also that the uncertainty of the future scenario forcing estimates is strongly related to the uncertainty of the 4xCO_2 forcing, since both are highly influenced by λ_1 (the inter-model correlation between our 4xCO_2 and RCP8.5 forcing is 0.82). This is particularly apparent for the GISS-E2-R model, where the response of the first few years is so abrupt that forcing estimates, and hence linear responses, are uncertain with both our and F13's estimation method.

In the two models CNRM-CM5 and MIROC5 the two forcing estimates are very similar, because the feedback is close to constant for all years. For these models we find also that the forcing estimate based on a single feedback parameter gives a slightly better estimate of the linear response. So if the global feedback in fact is constant for all years considered here, using all years in the regression should give a more certain estimate of the feedback parameter, and therefore also more certain forcing estimates.

For the three models GFDL-ESM2G, GFDL-ESM2M, and inmcm4 we find that our method is performing less well (see Figures in the Supporting information). The reason is probably linked to the almost constant $4xCO_2$ temperature responses over years $\sim 20-70$, $\sim 20-60$ and $\sim 20-120$, respectively. Our linear response with exponentially relaxing temperatures always predicts continuously increasing temperatures, which therefore poorly approximates these $4xCO_2$ global temperatures. The flattening of the response could possibly be linked to changes in the ocean circulation, e.g. a slowdown of the Atlantic meridional overturning circulation. In that case, linear systems with complex eigenvalues giving oscillatory responses could be an alternative solution. Hence, we will not disregard linear response in these results, but leave further testing of including oscillations in the responses to future studies.

4 Discussion

For most abrupt$4xCO_2$ experiments the Gregory plot follows a convex curve, hence our forcing estimates are mostly higher than those found from simple regression analyses over 150 years (Andrews et al., 2012), or using only the first 20 years (Andrews et al., 2015; Larson & Portmann, 2016). As suggested by PH17, this convexity could be explained by considering different feedback parameters associated with the different time scales of the responses. The time-scale dependence of the feedback parameter could be due to feedbacks varying in strength at different time scales, or it could be regionally different feedbacks weighted differently with time in the global average when the pattern of surface warming evolves. Since it is likely a combination of these circumstances, an interpretation of our parameters could be summarized into: λ_1: Average of annual-scale feedbacks in regions with strong annual-scale responses, λ_2: Average of decadal-scale feedbacks in regions with strong decadal-scale responses, λ_3: Average of centennial-scale feedbacks in regions with strong centennial-scale responses. Or as we come back to later, this description could also be considered an approximation of feedbacks changing with climate state.

The fixed-SST estimation method does not include time-variation and uncertainties in the feedback parameter. Instead, extra model simulations are made with SSTs fixed to climatological values, and the top of atmosphere radiative imbalance is diagnosed. A drawback of this method is that atmospheric and land surface temperatures are allowed to change. Hence the global temperature anomaly is not 0 when the radiative imbalance is diagnosed, and the forcing estimate is therefore contaminated with fast feedback processes associated with land warming. The fixed-SST estimates should be more comparable to our radiative imbalance after some months of adjustments of $T(t)$ and $N(t)$, and Figure 2 shows that they are indeed lower than our estimates for the models where they are available.

Ideally the fixed-SST method should be extended to fix the land surface temperatures also, in order to provide a consistent framework where forcing and feedbacks are well separated. Due to technical difficulties, this has only been done for one complex global climate model so far (Andrews et al., 2021). As discussed by Andrews et al. (2021), several methods have been suggested to correct fixed-SST estimates to account for effects of land temperature changes. One could for instance extrapolate the estimate to $T = 0$ using Eq. (2) given that we know the feedback parameter, or use radiative kernels (Richardson et al., 2019). Richardson et al. (2019) call these estimates Adjusted Effective radiative forcing, and find also these to be the best predictors for global surface temperatures because they have the efficacies closest to 1.

Efficacy factors are introduced to correct for differences in how strong the climate response is to different forcing agents, due to e.g differences in rapid adjustments, or effects of a forcing being concentrated in certain regions. Forcing in experiments considered in this study are dominated by CO_2, a well-mixed greenhouse gas. Other forcings present during the historical period and future scenarios could be more spatially inhomogeneous, e.g. aerosols, and contribute to different spatial patterns of the response. We neglect this effect when applying the parameters estimated for abrupt4xCO_2 experiments to other experiments, and assume the regional patterns to evolve similarly for different experiments. During the historical period, a changing feedback parameter will only result in weak changes in our forcing estimate since the temperature responses are still relatively weak. But if applying our method to strong forcings other than CO_2, the possible effect of efficacies should be investigated first.

When estimating a time-varying forcing, an alternative to fixing the SSTs to climatological values (as employed in RFMIP) is to prescribe the SSTs to e.g. the simulated historical values from the coupled model (as employed in AerChemMIP). These methods produce relatively similar results (Forster et al., 2016), and will both have a lower uncertainty than regression-based estimates. Regression-based estimates are influenced by changes in $T(t)$ arising due to internal variability, e.g. El Niño events, which could drive changes in $N(t)$. In prescribed-SST methods the temperature-driven changes in $N(t)$ is subtracted, resulting in a reduced noise level in the forcing estimate (Forster et al., 2013).

The theory described in this paper does not include an explicit temperature-dependence of the feedback parameter (Rohrschneider et al., 2019; Bloch-Johnson et al., 2021), since it is assumed that Eq. (6) is linear and $\mathbf{K}$ is independent of temperature. However, our estimation algorithm does not clearly distinguish between a time-scale dependence and a temperature-dependence of the feedbacks, since these dependencies are intrinsically linked. In particular, the strong temperature responses to 4xCO_2 is invoked on the long time scales, where the responses to the shorter time scales have already been realised, hereby affecting the feedback parameters if they have temperature dependence. If the 4xCO_2 responses have temperature-dependent feedbacks, the model needed to explicitly explain them becomes nonlinear, and our linear approach may perform less well in providing responses to other scenarios with weaker or stronger temperature responses than that of 4xCO_2. We believe this only causes smaller errors in the temperature responses studied here, but it is a potential explanation for our forcing and responses for the future scenarios being slightly overestimated.

Linear response theory is widely used to describe responses of climate variables. If a forcing record is known, linear response is a computationally cheap tool to estimate e.g. temperature responses compared to running a fully coupled climate model. Many studies assume a Green's function taking a certain form, with unknown parameters that need to be estimated. For box models taking the form of Eq. (6) the Green's function is a sum of exponential functions, but power-laws with fewer parameters have also been used with success (Rypdal & Rypdal, 2014; Fredriksen & Rypdal, 2017). Linear responses to RCP forcing are often studied using a non-parametric approach developed by Good et al. (2011). In the supporting information we show how this method relates to our linear model. This method was used in Good et al. (2013) to find the response to RCP scenarios using the forcing computed by F13. They use this to simulate only differences between RCP scenarios, while we attempt to simulate the full temperature evolution since the historical runs started until year 2100. Another difference to our approach is that we obtain a smoother estimate of the expected response to forcing, with fluctuations only coming from the forcing, while the responses of Good et al. (2013) are themselves influenced by internal variability.

Larson and Portmann (2016) note that the non-parametric model written in matrix form: $\mathbf{Y} = \mathbf{X}\Delta\mathbf{F}/F_0$ can be inverted to estimate the forcing increments $\Delta\mathbf{F}$, which can further be summed up to find the forcing time series. In this equation $\mathbf{Y}$ is a vector of the time evolution of a climate variable, and $\mathbf{X}$ is a matrix containing the same variable in the abrupt4xCO_2 experiment. Their resulting forcing estimate depends also on the forcing estimate F_0 from the abrupt4xCO_2 experiment, which introduces a potential source of bias in the estimate. Internal variability from **X** and **Y** can lead to a very noisy estimate, but some of this is removed when they replace the original abrupt4xCO_2 time series with a fitted exponential response. With our method we also greatly reduce the influence of internal variability from the experiment where the forcing is to be estimated by smoothing it with our linear response to the estimated forcing. So we can say that there is a trade-off between a noisy estimate and having more parameters to be estimated. The method by Larson and Portmann (2016) is treated as an alternative to the F13 method, but here we show how the F13 method and the linear response can be put into one framework. While Larson and Portmann (2016) can demonstrate that their method is not directly dependent of a changing feedback parameter, our method also has the power to explain why this can be the case.

5 Conclusions

The method presented here cleanly separates between forcing and responses to forcing, where the estimated parameters from abrupt4xCO_2 experiments are used to determine forcing and surface temperature responses for other experiments. The resulting RCP forcing estimates at the end of the 21st century is closer to the target levels than previous estimates by F13. Our high forcing estimates are strongly influenced by the high magnitude of the feedback parameter λ_1 at annual time scales. Unfortunately this value is uncertain, as it depends crucially on the first few years of adjustment. Using more ensemble members of abrupt4xCO_2 experiments may help constrain the estimate of λ_1 (M. Rugenstein, Gregory, et al., 2016). More members would also constrain regression estimates of forcing in general (Forster et al., 2016).

Forcing based on fixed-SST methods is often higher than the regression estimate over 150 years (Andrews et al., 2012; Tang et al., 2019), has a smaller uncertainty and is more computationally efficient (Forster et al., 2016). However, these forcing estimates are only available for a few models and scenarios in CMIP5. They will be available for more models and scenarios in CMIP6 (Smith et al., 2020), but far from all. The forcing estimation method presented here could therefore be a valuable supplement in the cases where fixed-SST forcing is unknown, particularly for models where a linear relation between N and T is a poor approximation. Improved forcing estimates could help to quantify the dependency of forcing value on CO_2 concentration in studies comparing e.g. 0.5x, 2x, 4x, 8x CO_2, and temperature dependence of feedbacks (Bloch-Johnson et al., 2021).

Putting forcing, linear responses, and nonconstancy of the global feedback parameter into a unified framework provides also an important insight into why the traditional regression-based forcing estimates may be too low. Furthermore, it suggests how these methods can be improved to provide better forcing estimates, resolving the problems caused by assuming a constant feedback parameter in regression-based methods (Forster et al., 2016).

Acknowledgments
The CMIP5 data are available at `https://esgf-node.llnl.gov/projects/cmip5/`. The forcing estimates from this paper will be stored in `https://dataverse.no/`, and can be accessed through `https://doi.org/10.18710/IHUVTB`. Our python code will be permanently stored in zenodo (link will be inserted when paper is accepted. The code is currently available at `https://github.com/Hegebf/CMIP5-forcing`). We thank Timothy Andrews and an anonymous reviewer for helpful suggestions improving our manuscript.

References

Andrews, T., Gregory, J. M., & Webb, M. J. (2015). The Dependence of Radiative Forcing and Feedback on Evolving Patterns of Surface Temperature Change in Climate Models. *Journal of Climate*, *28*(4), 1630–1648. https://doi.org/10.1175/JCLI-D-14-00545.1

Andrews, T., Gregory, J. M., Webb, M. J., & Taylor, K. E. (2012). Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophysical Research Letters*, *39*, L09712. https://doi.org/10.1029/2012GL051607

Andrews, T., Smith, C. J., Myhre, G., Forster, P. M., Chadwick, R., & Ackerley, D. (2021). Effective Radiative Forcing in a GCM With Fixed Surface Temperatures. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD033880. https://doi.org/10.1029/2020JD033880

Andrews, T., & Webb, M. J. (2018). The Dependence of Global Cloud and Lapse Rate Feedbacks on the Spatial Structure of Tropical Pacific Warming. *Journal of Climate*, *31*(2), 641-654. https://doi.org/10.1175/JCLI-D-17-0087.1

Armour, K. C. (2017). Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks. *Nature Climate Change*, *7*, 331 – 335. https://doi.org/10.1038/nclimate3278

Armour, K. C., Bitz, C. M., & Roe, G. H. (2013). Time-Varying Climate Sensitivity from Regional Feedbacks. *Journal of Climate*, *26*, 4518–4534. https://doi.org/10.1175/JCLI-D-12-00544.1

Bacmeister, J. T., Hannay, C., Medeiros, B., Gettelman, A., Neale, R., Fredriksen, H. B., … Otto-Bliesner, B. (2020). CO2 increase experiments using the Community Earth System Model (CESM): Relationship to climate sensitivity and comparison of CESM1 to CESM2. *Journal of Advances in Modeling Earth Systems*, *12*, e2020MS002120. https://doi.org/10.1029/2020MS002120

Bloch-Johnson, J., Rugenstein, M., & Abbot, D. S. (2020). Spatial radiative feedbacks from internal variability using multiple regression. *Journal of Climate*, *33*, 4121–4140. https://doi.org/10.1175/JCLI-D-19-0396.1

Bloch-Johnson, J., Rugenstein, M., Stolpe, M. B., Rohrschneider, T., Zheng, Y., & Gregory, J. M. (2021). Climate Sensitivity Increases Under Higher CO_2 Levels Due to Feedback Temperature Dependence. *Geophysical Research Letters*, *48*, e2020GL089074. https://doi.org/10.1029/2020GL089074

Byrne, B., & Goldblatt, C. (2014). Radiative forcing at high concentrations of well-mixed greenhouse gases. *Geophysical Research Letters*, *41*, 152-160. https://doi.org/10.1002/2013GL058456

Ceppi, P., & Gregory, J. M. (2017). Relationship of tropospheric stability to climate sensitivity and Earth's observed radiation budget. *Proceedings of the National Academy of Sciences*, *114*(50), 13126–13131. https://doi.org/10.1073/pnas.1714308114

Chung, E.-S., & Soden, B. J. (2015). An assessment of methods for computing radiative forcing in climate models. *Environmental Research Letters*, *10*(7), 074004. https://doi.org/10.1088/1748-9326/10/7/074004

Colman, R. A., & Hanson, L. I. (2013). On atmospheric radiative feedbacks associated with climate variability and change. *Climate Dynamics*, *40*(1), 475–492. https://doi.org/10.1007/s00382-012-1391-3

Colman, R. A., & Power, S. B. (2010). Atmospheric radiative feedbacks associated with transient climate change and climate variability. *Climate Dynamics*, *34*, 919-933. https://doi.org/10.1007/s00382-009-0541-8

Dessler, A. E. (2020). Potential Problems Measuring Climate Sensitivity from the Historical Record. *Journal of Climate*, *33*(6), 2237-2248. https://doi.org/10.1175/JCLI-D-19-0476.1

Dessler, A. E., & Forster, P. M. (2018). An Estimate of Equilibrium Climate Sensitivity From Interannual Variability. *Journal of Geophysical Research: Atmospheres*, *123*(16), 8634-8645. https://doi.org/10.1029/2018JD028481

Dong, Y., Armour, K. C., Zelinka, M. D., Proistosescu, C., Battisti, D. S., Zhou, C., & Andrews, T. (2020). Intermodel Spread in the Pattern Effect and Its Contribution to Climate Sensitivity in CMIP5 and CMIP6 Models. *Journal of Climate*, *33*(18), 7755-7775. https://doi.org/10.1175/JCLI-D-19-1011.1

Dong, Y., Proistosescu, C., Armour, K. C., & Battisti, D. S. (2019). Attributing Historical and Future Evolution of Radiative Feedbacks to Regional Warming Patterns using a Green's Function Approach: The Preeminence of the Western Pacific. *Journal of Climate*, *32*(17), 5471-5491. https://doi.org/10.1175/JCLI-D-18-0843.1

Edwards, C., & Penney, D. (2007). *Differential equations and boundary value problems: Computing and modelling (Fourth edition)*. Pearson.

Etminan, M., Myhre, G., Highwood, E. J., & Shine, K. P. (2016). Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing. *Geophysical Research Letters*, *43*(24), 12,614–12,623. https://doi.org/10.1002/2016GL071930

Forster, P. M., Andrews, T., Good, P., Gregory, J. M., Jackson, L. S., & Zelinka, M. (2013). Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *Journal of Geophysical Research*, *118*, 1139–1150. https://doi.org/10.1002/jgrd.50174

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., … Schulz, M. (2016). Recommendations for diagnosing effective radiative forcing from climate models for CMIP6. *Journal of Geophysical Research: Atmospheres*, *121*(20), 12,460–12,475. https://doi.org/10.1002/2016JD025320

Fredriksen, H.-B., & Rypdal, M. (2017). Long-range persistence in global surface temperatures explained by linear multibox energy balance models. *Journal of Climate*, *30*, 7157–7168. https://doi.org/10.1175/JCLI-D-16-0877.1

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D., & Tytéca, S. (2013). Transient Climate Response in a Two-Layer Energy-Balance Model. Part II: Representation of the Efficacy of Deep-Ocean Heat Uptake and Validation for CMIP5 AOGCMs. *Journal of Climate*, *26*(6), 1859–1876. https://doi.org/10.1175/JCLI-D-12-00196.1

Geoffroy, O., Saint-Martin, D., Olivié, D. J. L., 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. *Journal of Climate*, *26*, 1841–1857. https://doi.org/10.1175/JCLI-D-12-00195.1

Good, P., Gregory, J. M., & Lowe, J. A. (2011). A step-response simple climate model to reconstruct and interpret AOGCM projections. *Geophysical Research Letters*, *38*, L01703. https://doi.org/10.1029/2010GL045208

Good, P., Gregory, J. M., Lowe, J. A., & Andrews, T. (2013). Abrupt CO2 experiments as tools for predicting and understanding CMIP5 representative concentration pathway projections. *Climate Dynamics*, *40*(3), 1041–1053. https://doi.org/10.1007/s00382-012-1410-4

Gregory, J. M., & Andrews, T. (2016). Variation in climate sensitivity and feedback parameters during the historical period. *Geophysical Research Letters*, *43*(8), 3911–3920. https://doi.org/10.1002/2016GL068406

Gregory, J. M., Andrews, T., & Good, P. (2015). The inconstancy of the transient climate response parameter under increasing CO_2. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *373*, 20140417. https://doi.org/10.1098/rsta.2014.0417

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., … Williams, K. D. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, *31*, L03205. https://doi.org/10.1029/2003GL018747

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., … Zhang, S. (2005). Efficacy of climate forcings. *Journal of Geophysical Re-*

search: Atmospheres, *110*(D18). https://doi.org/10.1029/2005JD005776

Hasselmann, K., Sausen, R., Maier-Reimer, E., & Voss, R. (1993). On the cold start problem in transient simulations with coupled atmosphere-ocean models. *Climate Dynamics*, *9*(6), 53–61. https://doi.org/10.1007/BF00210008

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

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., ... Vertenstein, M. (2015). The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bulletin of the American Meteorological Society*, *96*(8), 1333-1349. https://doi.org/10.1175/BAMS-D-13-00255.1

Larson, E. J. L., & Portmann, R. W. (2016). A Temporal Kernel Method to Compute Effective Radiative Forcing in CMIP5 Transient Simulations. *Journal of Climate*, *29*(4), 1497-1509. https://doi.org/10.1175/JCLI-D-15-0577.1

Loeb, N. G., Wang, H., Allan, R. P., Andrews, T., Armour, K., Cole, J. N. S., ... Wyser, K. (2020). New generation of climate models track recent unprecedented changes in earth's radiation budget observed by ceres. *Geophysical Research Letters*, *47*(5), e2019GL086705. https://doi.org/10.1029/2019GL086705

Lutsko, N. J., & Takahashi, K. (2018). What Can the Internal Variability of CMIP5 Models Tell Us about Their Climate Sensitivity? *Journal of Climate*, *31*(13), 5051-5069. https://doi.org/10.1175/JCLI-D-17-0736.1

Marvel, K., Pincus, R., Schmidt, G. A., & Miller, R. L. (2018). Internal Variability and Disequilibrium Confound Estimates of Climate Sensitivity From Observations. *Geophysical Research Letters*, *45*(3), 1595-1601. https://doi.org/10.1002/2017GL076468

Menzel, M. E., & Merlis, T. M. (2019). Connecting Direct Effects of CO2 Radiative Forcing to Ocean Heat Uptake and Circulation. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2163-2176. https://doi.org/10.1029/2018MS001544

Meraner, K., Mauritsen, T., & Voigt, A. (2013). Robust increase in equilibrium climate sensitivity under global warming. *Geophysical Research Letters*, *40*(22), 5944–5948. https://doi.org/10.1002/2013GL058118

Myhre, G., Highwood, E. J., Shine, K. P., & Stordal, F. (1998). New estimates of radiative forcing due to well mixed greenhouse gases. *Geophysical Research Letters*, *25*(14), 2715–2718. https://doi.org/10.1029/98GL01908

Olonscheck, D., Rugenstein, M., & Marotzke, J. (2020). Broad consistency between observed and simulated trends in sea surface temperature patterns. *Geophysical Research Letters*, *47*(10), e2019GL086773. https://doi.org/10.1029/2019GL086773

Paynter, D., & Frölicher, T. L. (2015). Sensitivity of radiative forcing, ocean heat uptake, and climate feedback to changes in anthropogenic greenhouse gases and aerosols. *Journal of Geophysical Research: Atmospheres*, *120*(19), 9837–9854. https://doi.org/10.1002/2015JD023364

Pincus, R., Forster, P. M., & Stevens, B. (2016). The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6. *Geoscientific Model Development*, *9*(9), 3447–3460. https://doi.org/10.5194/gmd-9-3447-2016

Proistosescu, C., & Huybers, P. J. (2017). Slow climate mode reconciles historical and model-based estimates of climate sensitivity. *Sciences Advances*, *3*, e1602821. https://doi.org/10.1126/sciadv.1602821

Richardson, T. B., Forster, P. M., Smith, C. J., Maycock, A. C., Wood, T., Andrews, T., ... Watson-Parris, D. (2019). Efficacy of Climate Forcings in PDRMIP Models. *Journal of Geophysical Research: Atmospheres*, *124*(23),

12824-12844. https://doi.org/10.1029/2019JD030581
Rohrschneider, T., Stevens, B., & Mauritsen, T. (2019). On simple representations of the climate response to external radiative forcing. *Climate Dynamics*, *53*(5), 3131–3145. https://doi.org/10.1007/s00382-019-04686-4
Rugenstein, M., Bloch-Johnson, J., Gregory, J., Andrews, T., Mauritsen, T., Li, C., … Knutti, R. (2020). Equilibrium Climate Sensitivity Estimated by Equilibrating Climate Models. *Geophysical Research Letters*, *47*(4), e2019GL083898. https://doi.org/10.1029/2019GL083898
Rugenstein, M., Caldeira, K., & Knutti, R. (2016). Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes. *Geophysical Research Letters*, *43*(18), 9877-9885. https://doi.org/10.1002/2016GL070907
Rugenstein, M., Gregory, J. M., Schaller, N., Sedláček, J., & Knutti, R. (2016). Multiannual Ocean–Atmosphere Adjustments to Radiative Forcing. *Journal of Climate*, *29*(15), 5643-5659. https://doi.org/10.1175/JCLI-D-16-0312.1
Rugenstein, M. A. A., & Armour, K. C. (2021). Three Flavors of Radiative Feedbacks and Their Implications for Estimating Equilibrium Climate Sensitivity. *Geophysical Research Letters*, *48*(15), e2021GL092983. https://doi.org/10.1029/2021GL092983
Rypdal, M., & Rypdal, K. (2014). Long-Memory Effects in Linear Response Models of Earth's Temperature and Implications for Future Global Warming. *Journal of Climate*, *27*, 5240–5258. https://doi.org/10.1175/JCLI-D-13-00296.1
Sherwood, S. C., Bony, S., Boucher, O., Bretherton, C., Forster, P. M., Gregory, J. M., & Stevens, B. (2015). Adjustments in the Forcing-Feedback Framework for Understanding Climate Change. *Bulletin of the American Meteorological Society*, *96*(2), 217-228. https://doi.org/10.1175/BAMS-D-13-00167.1
Smith, C. J., Kramer, R. J., Myhre, G., Alterskjær, K., Collins, W., Sima, A., … Forster, P. M. (2020). Effective radiative forcing and adjustments in cmip6 models. *Atmospheric Chemistry and Physics*, *20*(16), 9591–9618. https://doi.org/10.5194/acp-20-9591-2020
Smith, C. J., Kramer, R. J., Myhre, G., Forster, P. M., Soden, B. J., Andrews, T., … Watson-Parris, D. (2018). Understanding Rapid Adjustments to Diverse Forcing Agents. *Geophysical Research Letters*, *45*(21), 12,023-12,031. https://doi.org/10.1029/2018GL079826
Soden, B. J., Collins, W. D., & Feldman, D. R. (2018). Reducing uncertainties in climate models. *Science*, *361*(6400), 326–327. https://doi.org/10.1126/science.aau1864
Tang, T., Shindell, D., Faluvegi, G., Myhre, G., Olivié, D., Voulgarakis, A., … Smith, C. (2019). Comparison of Effective Radiative Forcing Calculations Using Multiple Methods, Drivers, and Models. *Journal of Geophysical Research: Atmospheres*, *124*(8), 4382-4394. https://doi.org/10.1029/2018JD030188
Trossman, D. S., Palter, J. B., Merlis, T. M., Huang, Y., & Xia, Y. (2016). Large-scale ocean circulation-cloud interactions reduce the pace of transient climate change. *Geophysical Research Letters*, *43*(8), 3935–3943. https://doi.org/10.1002/2016GL067931
van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., … Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change*, *109*, 5–31. https://doi.org/10.1007/s10584-011-0148-z
Williams, K. D., Ingram, W. J., & Gregory, J. M. (2008). Time Variation of Effective Climate Sensitivity in GCMs. *Journal of Climate*, *21*(19), 5076-5090. https://doi.org/10.1175/2008JCLI2371.1
Winton, M., Takahashi, K., & Held, I. M. (2010). Importance of Ocean Heat Uptake Efficacy to Transient Climate Change. *Journal of Climate*, *23*(9), 2333-2344. https://doi.org/10.1175/2009JCLI3139.1
Zelinka, M. D., Klein, S. A., Taylor, K. E., Andrews, T., Webb, M. J., Gregory,

J. M., & Forster, P. M. (2013). Contributions of Different Cloud Types to Feedbacks and Rapid Adjustments in CMIP5. *Journal of Climate*, *26*(14), 5007–5027. https://doi.org/10.1175/JCLI-D-12-00555.1

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., … Taylor, K. E. (2020). Causes of Higher Climate Sensitivity in CMIP6 Models. *Geophysical Research Letters*, *47*(1), e2019GL085782. https://doi.org/10.1029/2019GL085782

Zhou, C., Zelinka, M. D., Dessler, A. E., & Klein, S. A. (2015). The relationship between interannual and long-term cloud feedbacks. *Geophysical Research Letters*, *42*(23), 10,463-10,469. https://doi.org/10.1002/2015GL066698

Zhou, C., Zelinka, M. D., & Klein, S. A. (2016). Impact of decadal cloud variations on the Earth's energy budget. *Nature Geoscience*, *9*(12), 871–874. https://doi.org/10.1038/ngeo2828

Supporting Information for "Estimating radiative forcing with a nonconstant feedback parameter and linear response"

Hege-Beate Fredriksen[1], Maria Rugenstein[2]and Rune Graversen[1,3]

[1]Department of Physics and Technology, UiT the Arctic University of Norway, Tromsø, Norway

[2]Colorado State University, Fort Collins, USA

[3]The Norwegian Meteorological Institute, Norway

Contents of this file

1. Text S1 to S2
2. Figures S1 to S111
3. Table S1

Introduction This document repeats Figures 1, 3 and 4 for all models and available RCP scenarios. For a description of the figures, see the NorESM1-M figures in the main manuscript.

Text S1 and S2 elaborates some of the mathematics needed to 1) derive the temperature response, and 2) understand the relationship between our linear model and the non-parametric linear models considered in other papers.

Table S1 in the end lists the piControl trend values used when subtracting linear trends from the variables of the abrupt4xCO_2 experiment.

Text S1: Deriving temperature response

To show that Eq. (7) is the solution of Eq. (6), we start by rewriting to:

$$\frac{d\mathbf{T}(t)}{dt} = \mathbf{C}^{-1}\mathbf{K}\mathbf{T}(t) + \mathbf{C}^{-1}\mathbf{F}(t)$$

We consider first the homogeneous problem

$$\frac{d\mathbf{T}(t)}{dt} = \mathbf{A}\mathbf{T}(t)$$

where $\mathbf{A} = \mathbf{C}^{-1}\mathbf{K}$. The matrix of possible solutions $\mathbf{x}_i(t)$ to this problem is the fundamental matrix

$$\mathbf{\Phi}(t) = [\mathbf{x}_1(t)\,|\,\mathbf{x}_2(t)\,|\,\ldots\,|\,\mathbf{x}_n(t)].$$

$e^{\mathbf{A}t}$ is a fundamental matrix when $\mathbf{A}$ consists of constant coefficients, since

$$\frac{d\mathbf{\Phi}(t)}{dt} = \frac{de^{\mathbf{A}t}}{dt} = \mathbf{A}e^{\mathbf{A}t} = \mathbf{A}\mathbf{\Phi}(t).$$

According to the variation of parameters formula for first-order linear systems $\frac{d\mathbf{x}}{dt} = \mathbf{P}(t)\mathbf{x} + \mathbf{f}(t)$, a particular solution is given by

$$\mathbf{x}_p(t) = \mathbf{\Phi}(t)\int \mathbf{\Phi}(t)^{-1}\mathbf{f}(t)dt$$

(see e.g. Edwards and Penney (2007)). For our problem, this means that the particular solution is

$$\mathbf{x}_p(t) = e^{\mathbf{A}t}\int e^{-\mathbf{A}t}\mathbf{C}^{-1}\mathbf{F}(t)dt.$$

Given an initial value $\mathbf{T}(0) = \mathbf{T}_0$, the full solution can be written as

$$\mathbf{T}(t) = e^{\mathbf{C}^{-1}\mathbf{K}t}\mathbf{T}_0 + \int_0^t e^{(t-s)\mathbf{C}^{-1}\mathbf{K}}\mathbf{C}^{-1}\mathbf{F}(s)\,ds,$$

or alternatively, if we know the full history of the system,

$$\mathbf{T}(t) = \int_{-\infty}^t e^{(t-s)\mathbf{C}^{-1}\mathbf{K}}\mathbf{C}^{-1}\mathbf{F}(s)\,ds.$$

Text S2: Relation to non-parametric impulse-response models

To find the relation between linear model considered here and linear models considered in e.g. Larson and Portmann (2016); Good, Gregory, and Lowe (2011); Good, Gregory, Lowe, and Andrews (2013), we start with the general equation from the end of Section 2.1: $T(t) = \int_{-\infty}^{t} G(t-s)F(s)ds$. As noted by Hasselmann, Sausen, Maier-Reimer, and Voss (1993), such a convolution integral can also describe a general climate state variable $\Phi(t)$ that responds linearly to a forcing:

$$\Phi(t) = \int_0^t G(t-s)F(s)ds \tag{1}$$

assuming $F(t) = 0$ for $t \leq 0$. If the forcing takes the form of a unit step-function, which is 0 for $t \leq 0$ and 1 for $t > 0$, the climate response is:

$$R(t) = \int_0^t G(t-s)ds$$

and $\frac{dR}{dt} = G(t)$. By performing an integration by parts, we note that Eq. (1) can be rewritten to

$$\Phi(t) = \int_0^t \frac{dF}{ds} R(t-s)ds \tag{2}$$

where the additional term $R(0)F(t) - R(t)F(0) = 0$ because $R(0) = 0$ and $F(0) = 0$. Discretizing this integral using time steps of years results in the same type of sum used by Larson and Portmann (2016); Good et al. (2011, 2013):

$$\Phi_i = \sum_{j=0}^{i} \Delta F_j R_{i-j} \tag{3}$$

If using a response to a step-forcing ΔF_s instead of the unit response, the response needs to be normalized by ΔF_s.

References

Edwards, C., & Penney, D. (2007). *Differential equations and boundary value problems: Computing and modelling (Fourth edition).* Pearson.

Good, P., Gregory, J. M., & Lowe, J. A. (2011). A step-response simple climate model to reconstruct and interpret AOGCM projections. *Geophysical Research Letters*, *38*, L01703. https://doi.org/10.1029/2010GL045208

Good, P., Gregory, J. M., Lowe, J. A., & Andrews, T. (2013). Abrupt CO2 experiments as tools for predicting and understanding CMIP5 representative concentration pathway projections. *Climate Dynamics*, *40*(3), 1041–1053. https://doi.org/10.1007/s00382-012-1410-4

Hasselmann, K., Sausen, R., Maier-Reimer, E., & Voss, R. (1993). On the cold start problem in transient simulations with coupled atmosphere-ocean models. *Climate Dynamics*, *9*(6), 53–61. https://doi.org/10.1007/BF00210008

Larson, E. J. L., & Portmann, R. W. (2016). A Temporal Kernel Method to Compute Effective Radiative Forcing in CMIP5 Transient Simulations. *Journal of Climate*, *29*(4), 1497-1509. https://doi.org/10.1175/JCLI-D-15-0577.1

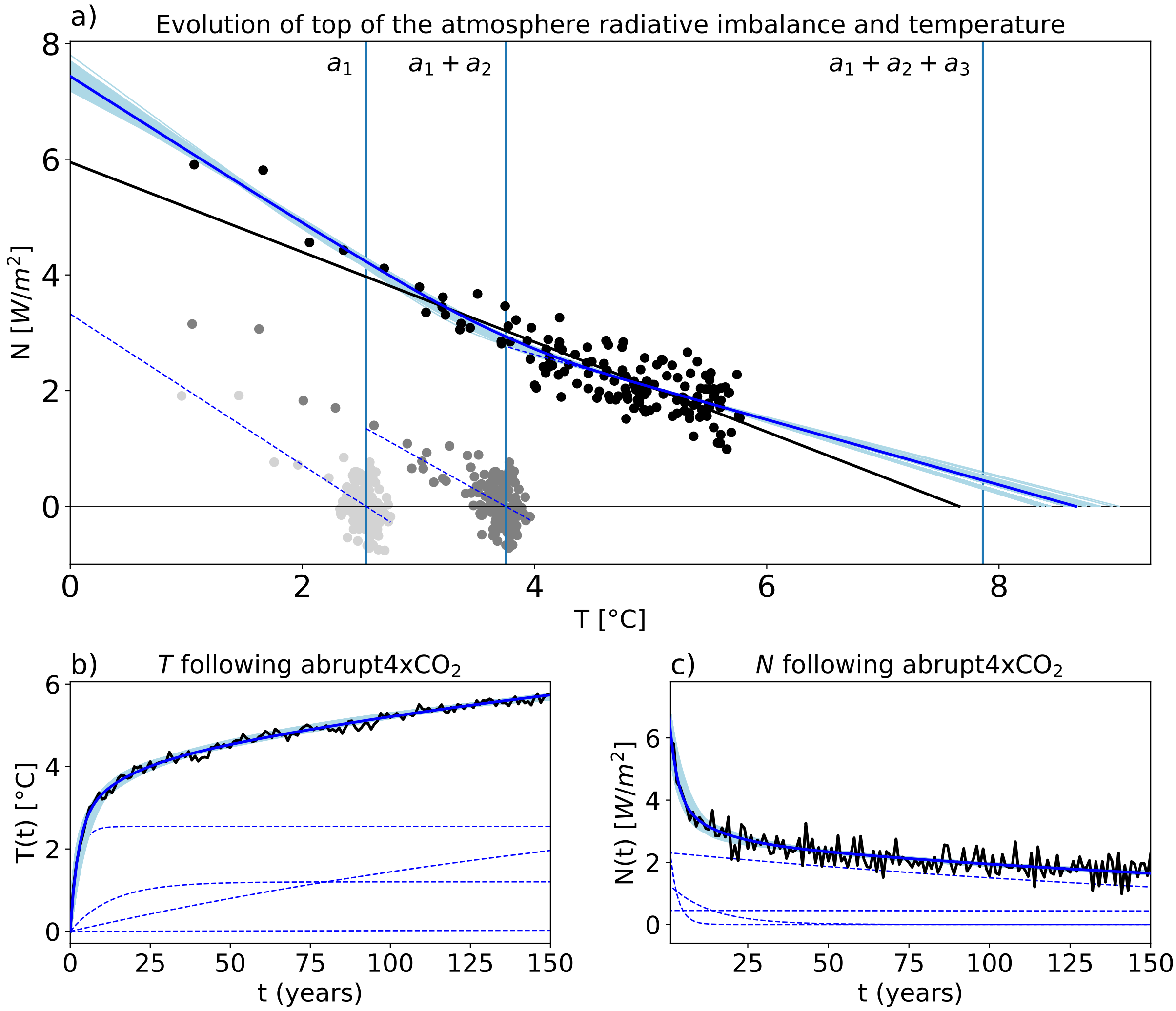

Figure S1. As Figure 1, but for the model ACCESS1-0.

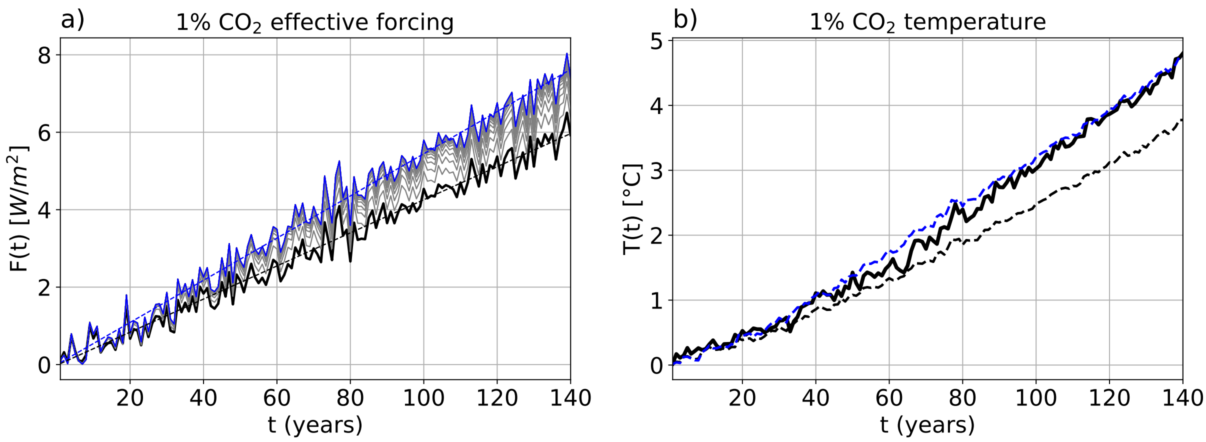

Figure S2. As Figure 3, but for the model ACCESS1-0.

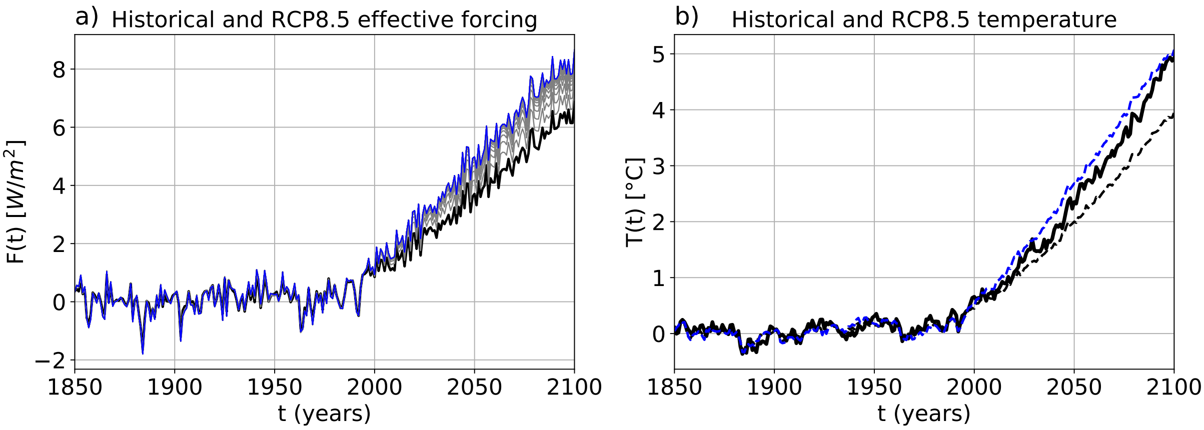

Figure S3. As Figure 4, but for the model ACCESS1-0.

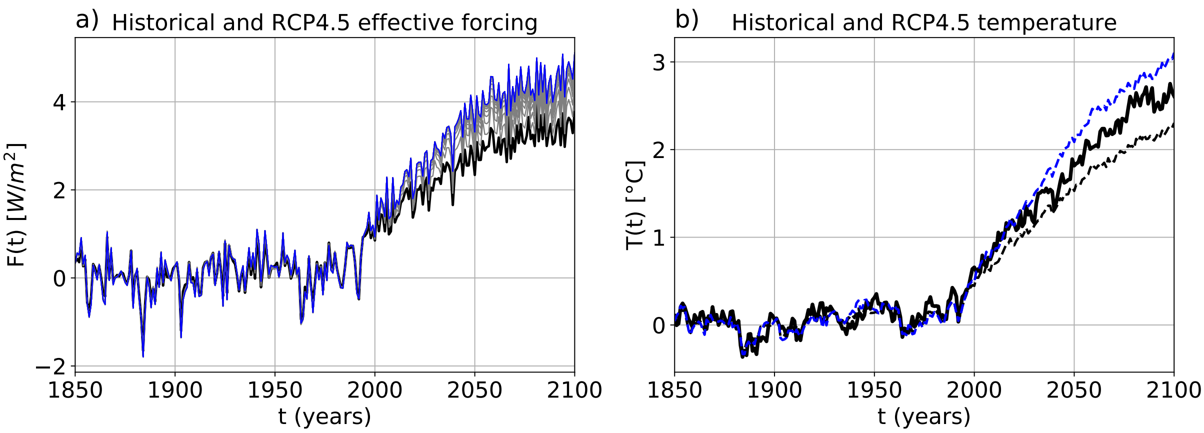

Figure S4. As Figure 4, but for the model ACCESS1-0 and experiment RCP4.5.

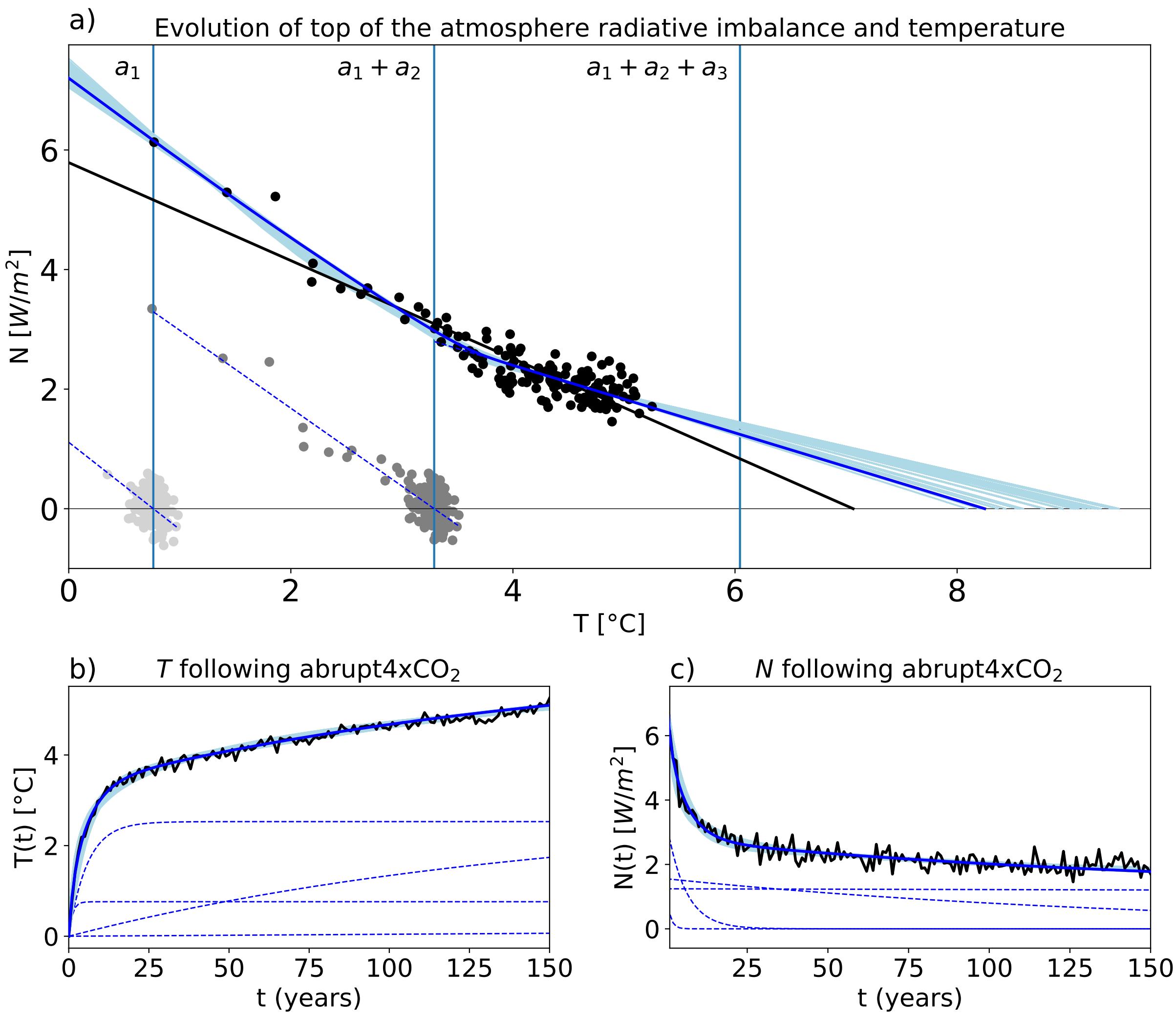

Figure S5. As Figure 1, but for the model ACCESS1-3.

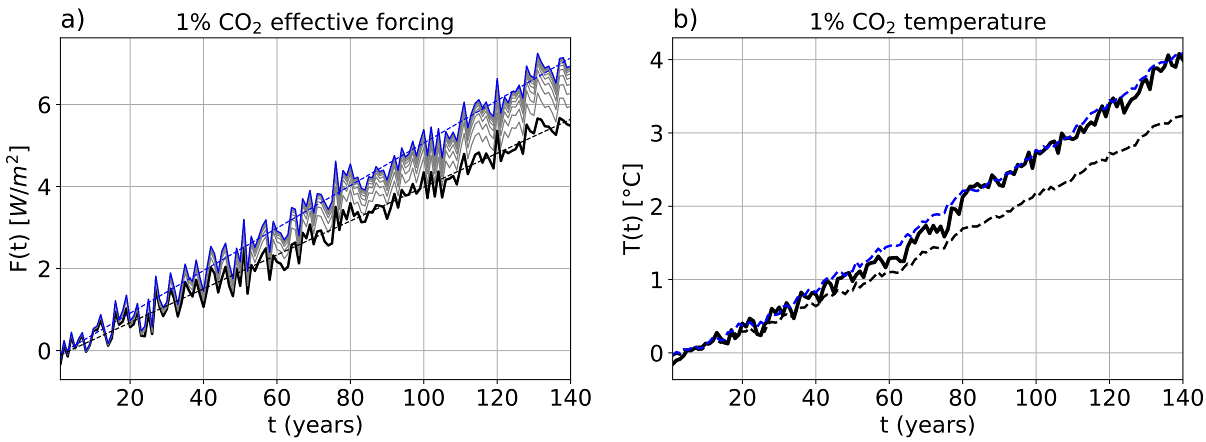

Figure S6. As Figure 3, but for the model ACCESS1-3.

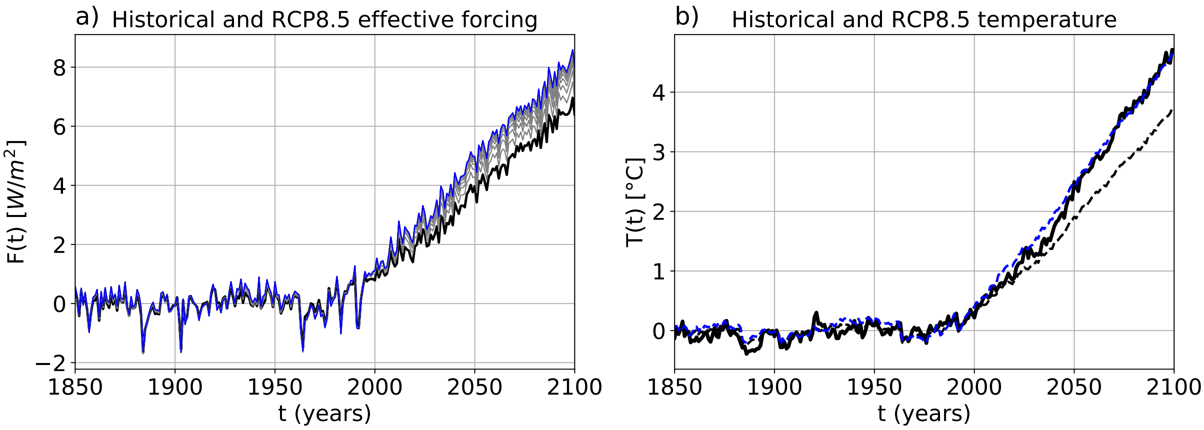

Figure S7. As Figure 4, but for the model ACCESS1-3.

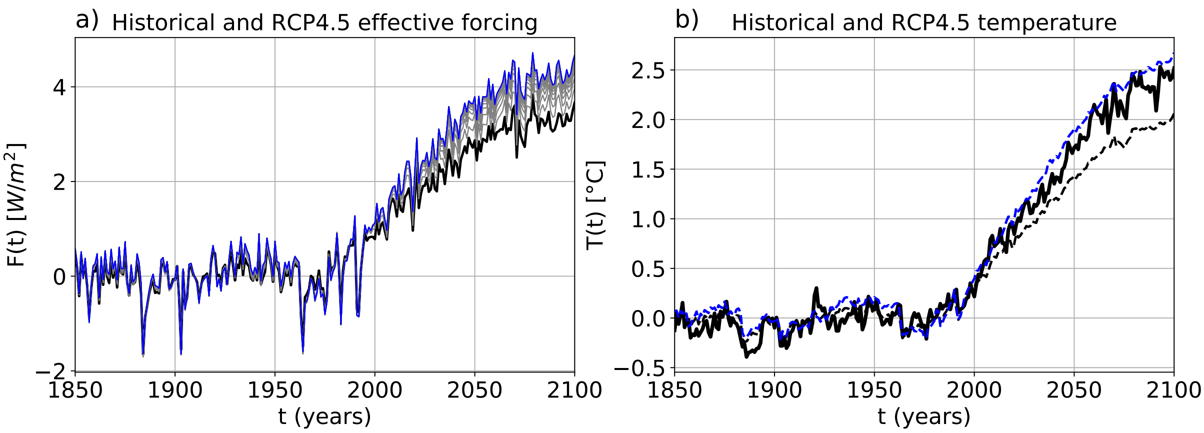

Figure S8. As Figure 4, but for the model ACCESS1-3 and experiment RCP4.5.

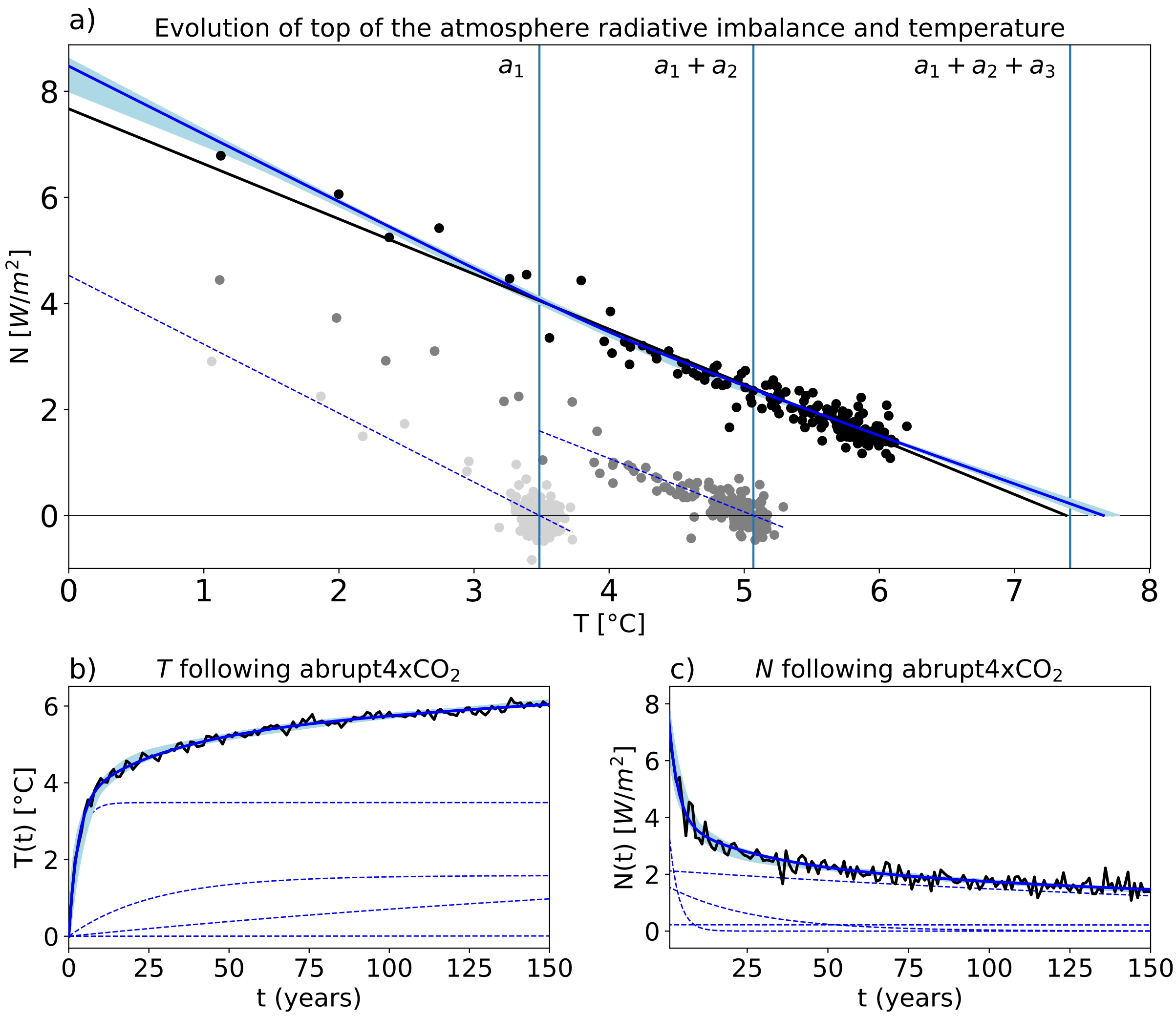

Figure S9. As Figure 1, but for the model CanESM2.

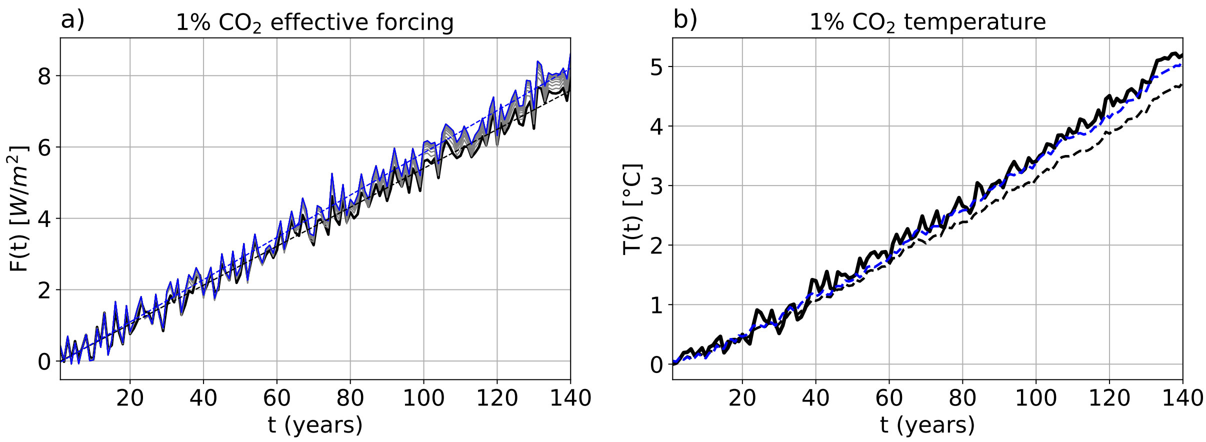

Figure S10. As Figure 3, but for the model CanESM2.

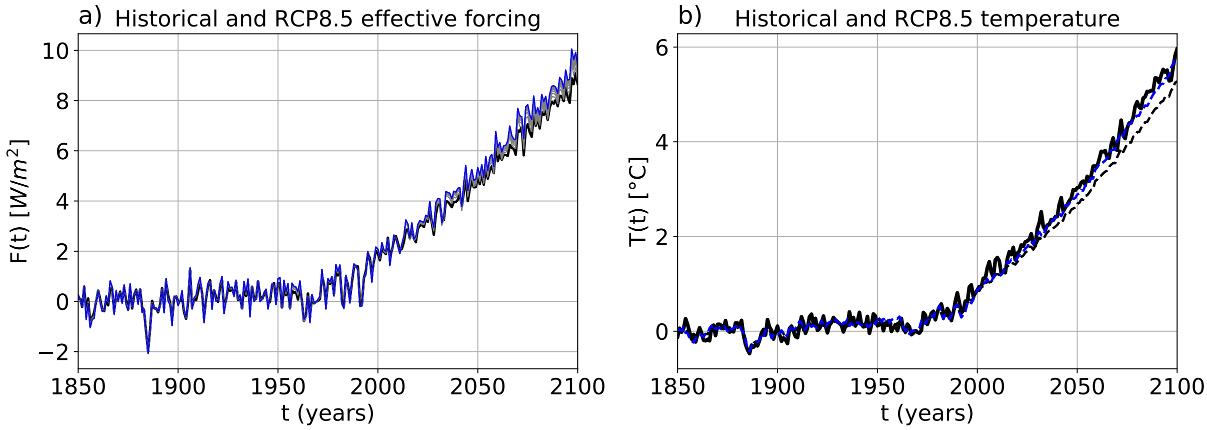

Figure S11. As Figure 4, but for the model CanESM2.

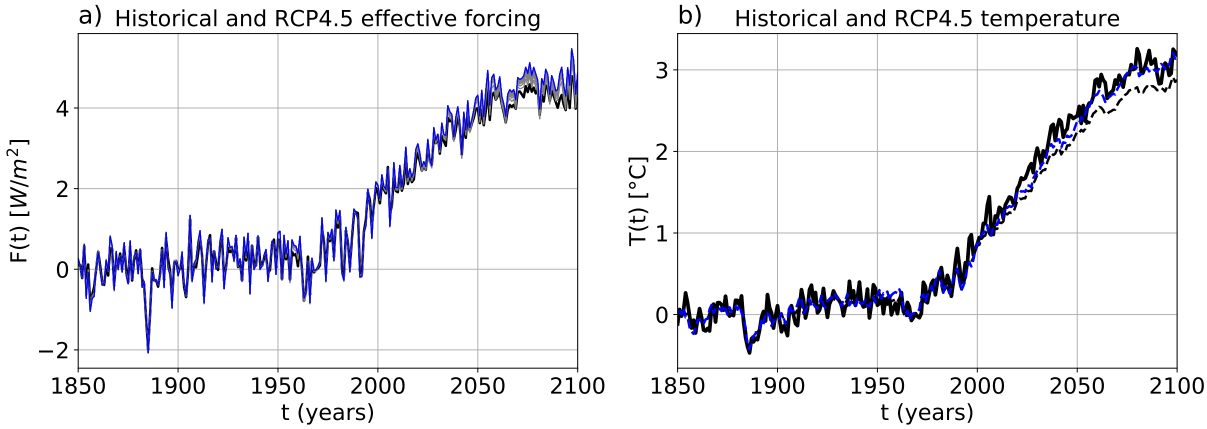

Figure S12. As Figure 4, but for the model CanESM2 and experiment RCP4.5.

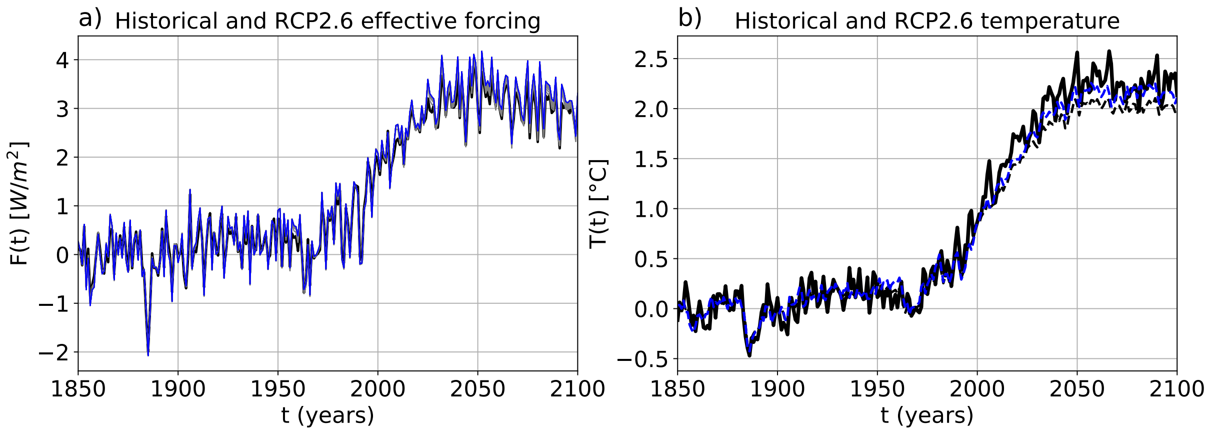

Figure S13. As Figure 4, but for the model CanESM2 and experiment RCP2.6.

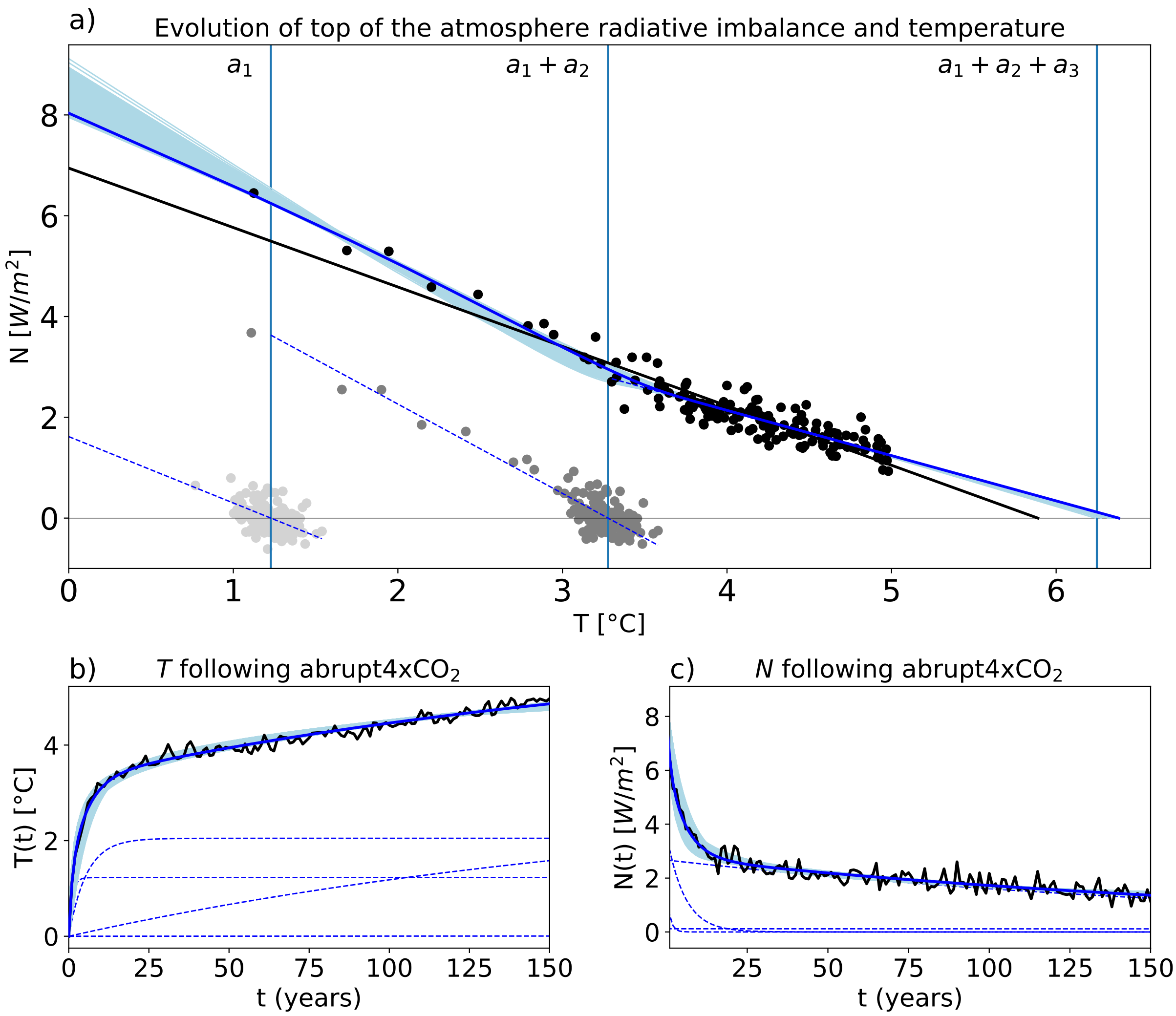

Figure S14. As Figure 1, but for the model CCSM4.

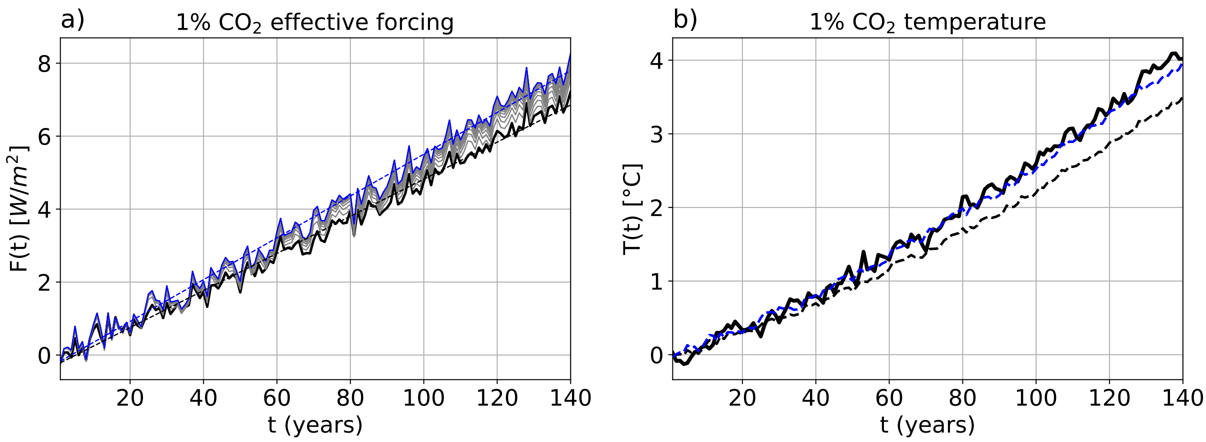

Figure S15. As Figure 3, but for the model CCSM4.

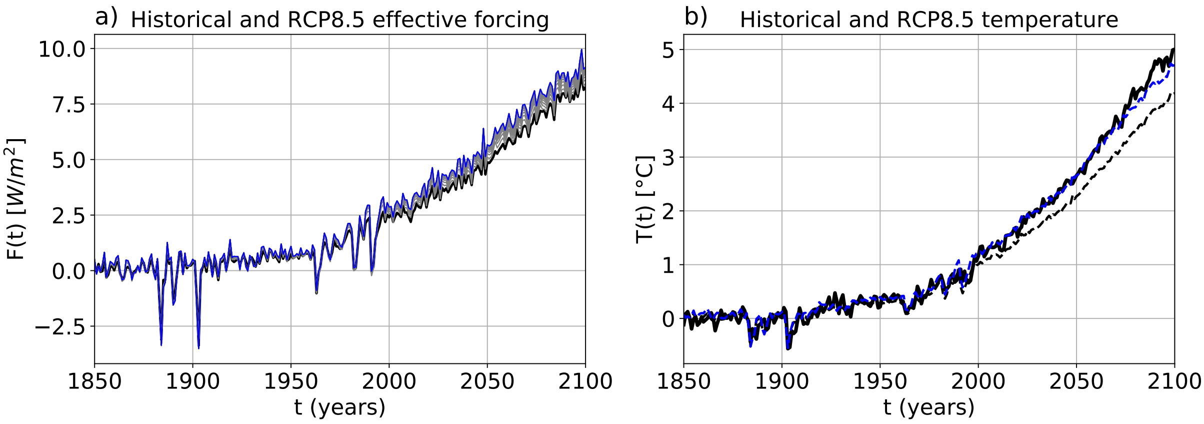

Figure S16. As Figure 4, but for the model CCSM4.

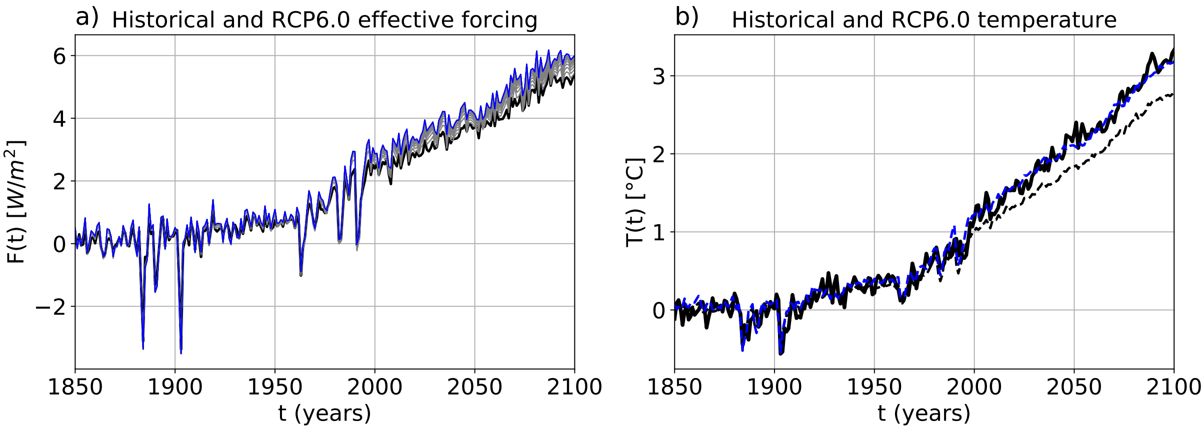

Figure S17. As Figure 4, but for the model CCSM4 and experiment RCP6.0.

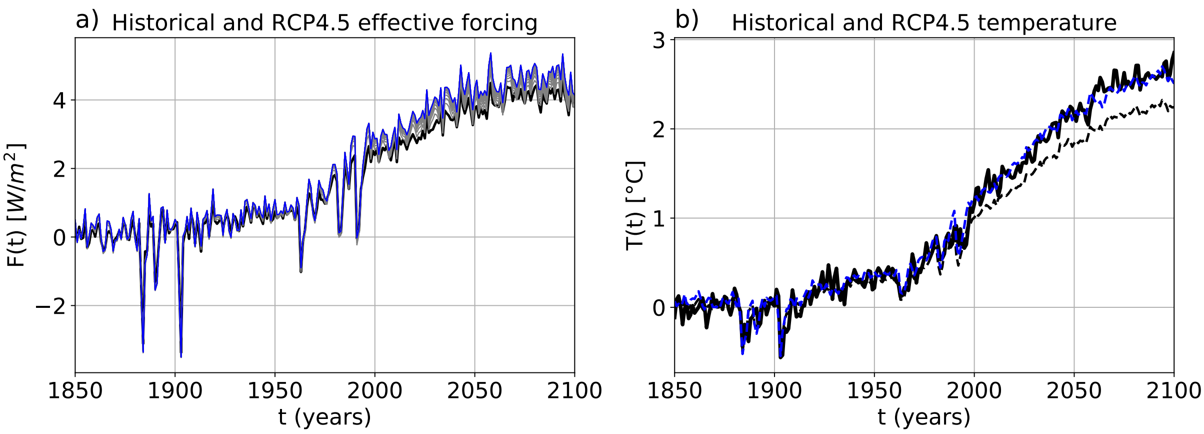

Figure S18. As Figure 4, but for the model CCSM4 and experiment RCP4.5.

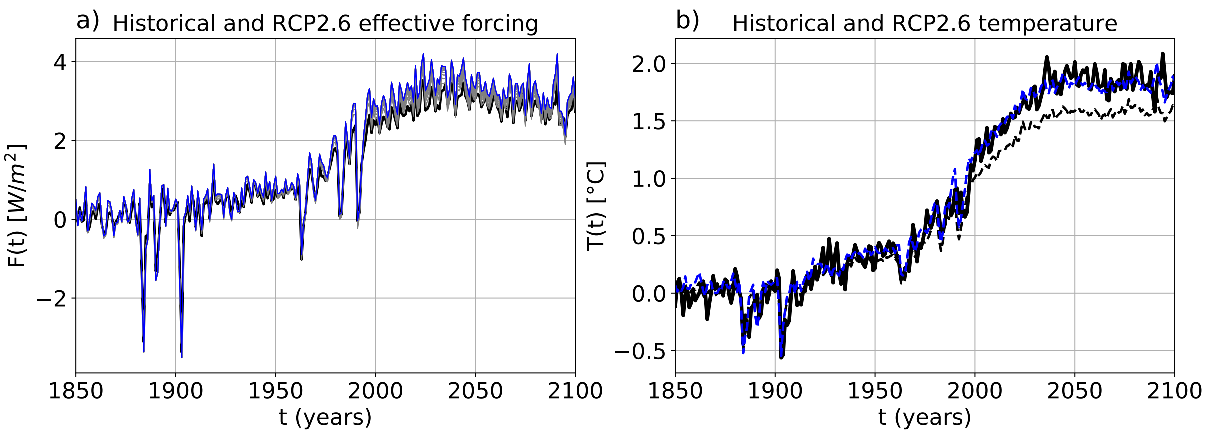

Figure S19. As Figure 4, but for the model CCSM4 and experiment RCP2.6.

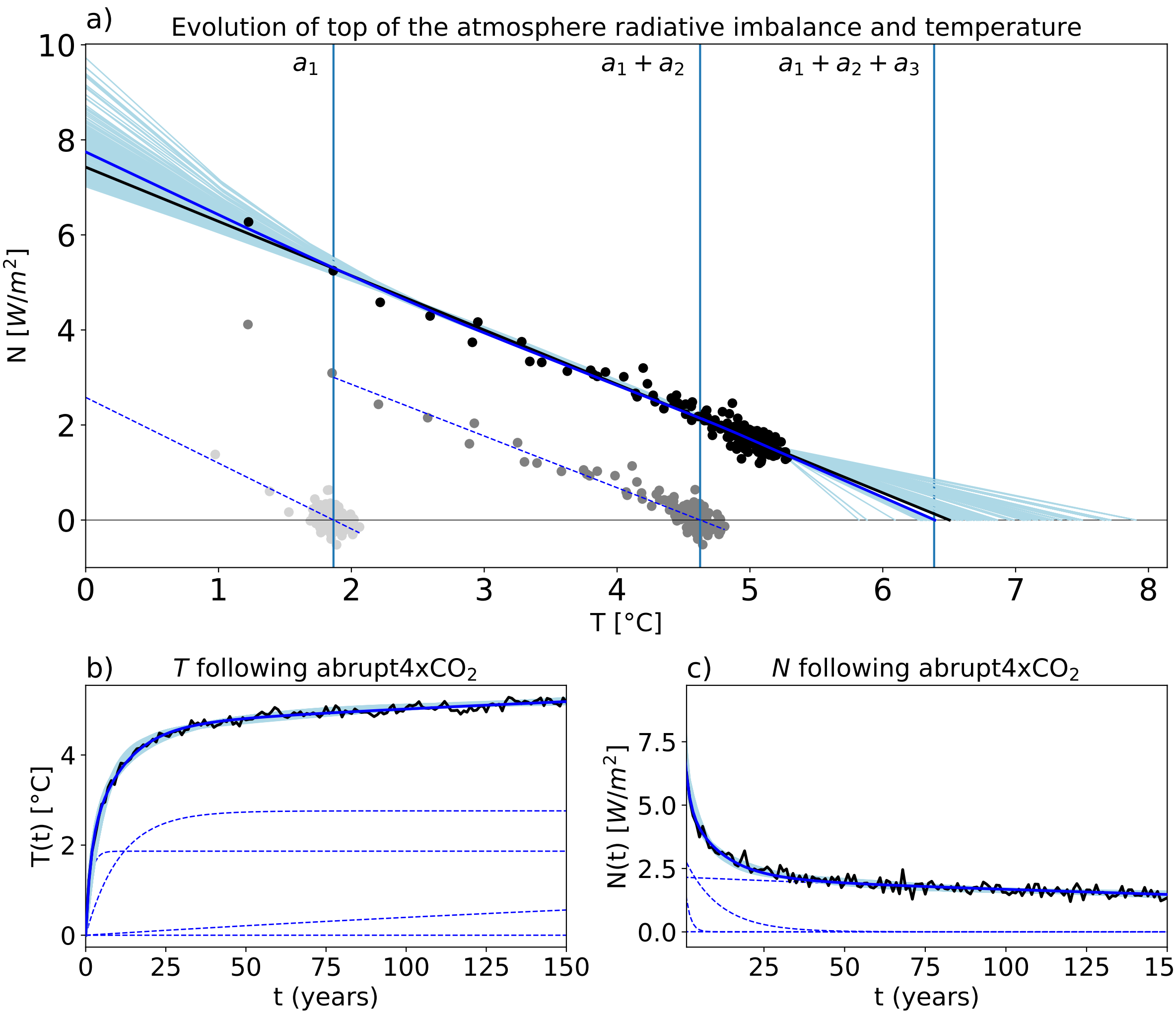

Figure S20. As Figure 1, but for the model CNRM-CM5.

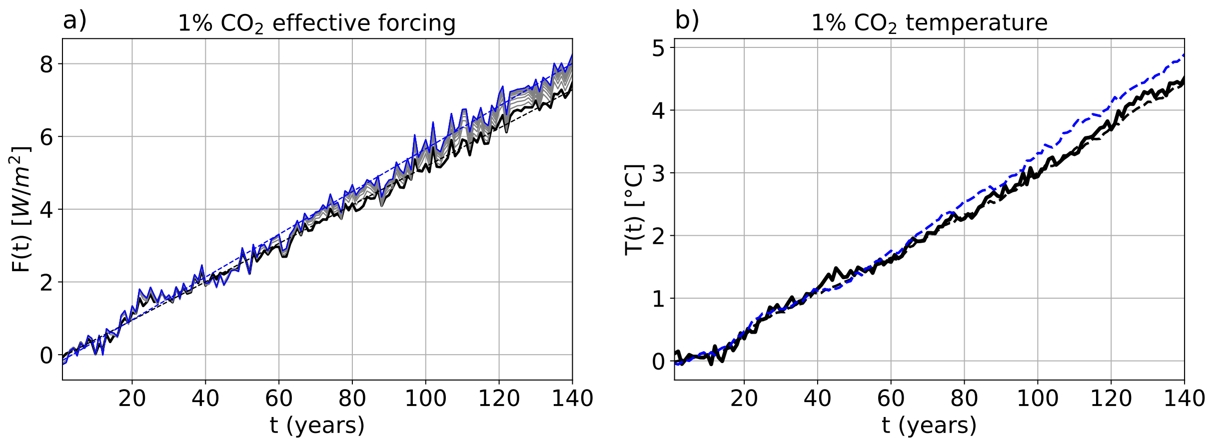

Figure S21. As Figure 3, but for the model CNRM-CM5.

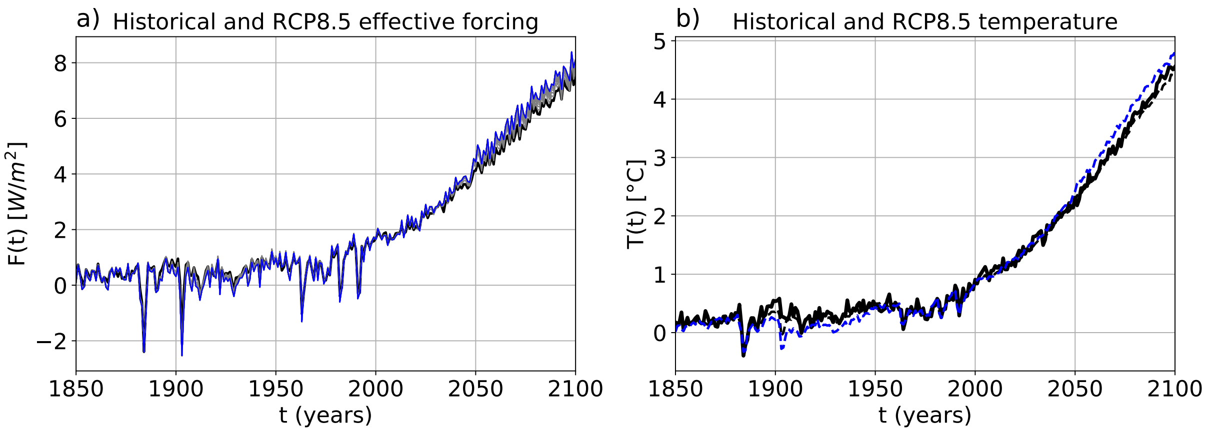

Figure S22. As Figure 4, but for the model CNRM-CM5.

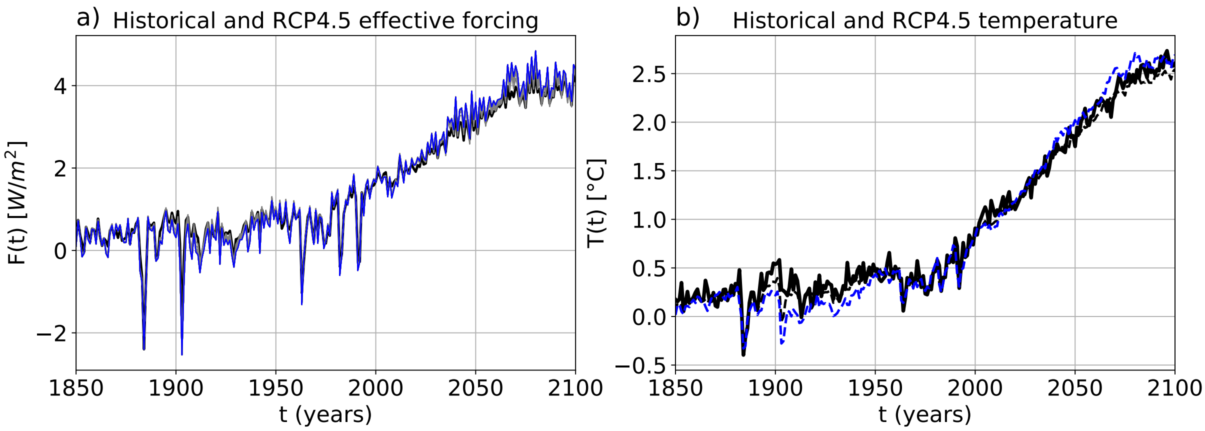

Figure S23. As Figure 4, but for the model CNRM-CM5 and experiment RCP4.5.

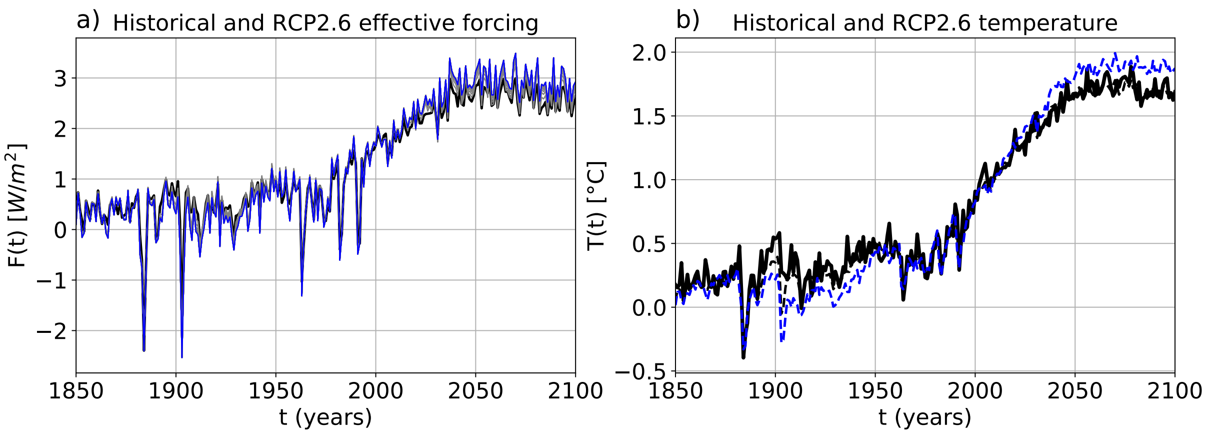

Figure S24. As Figure 4, but for the model CNRM-CM5 and experiment RCP2.6.

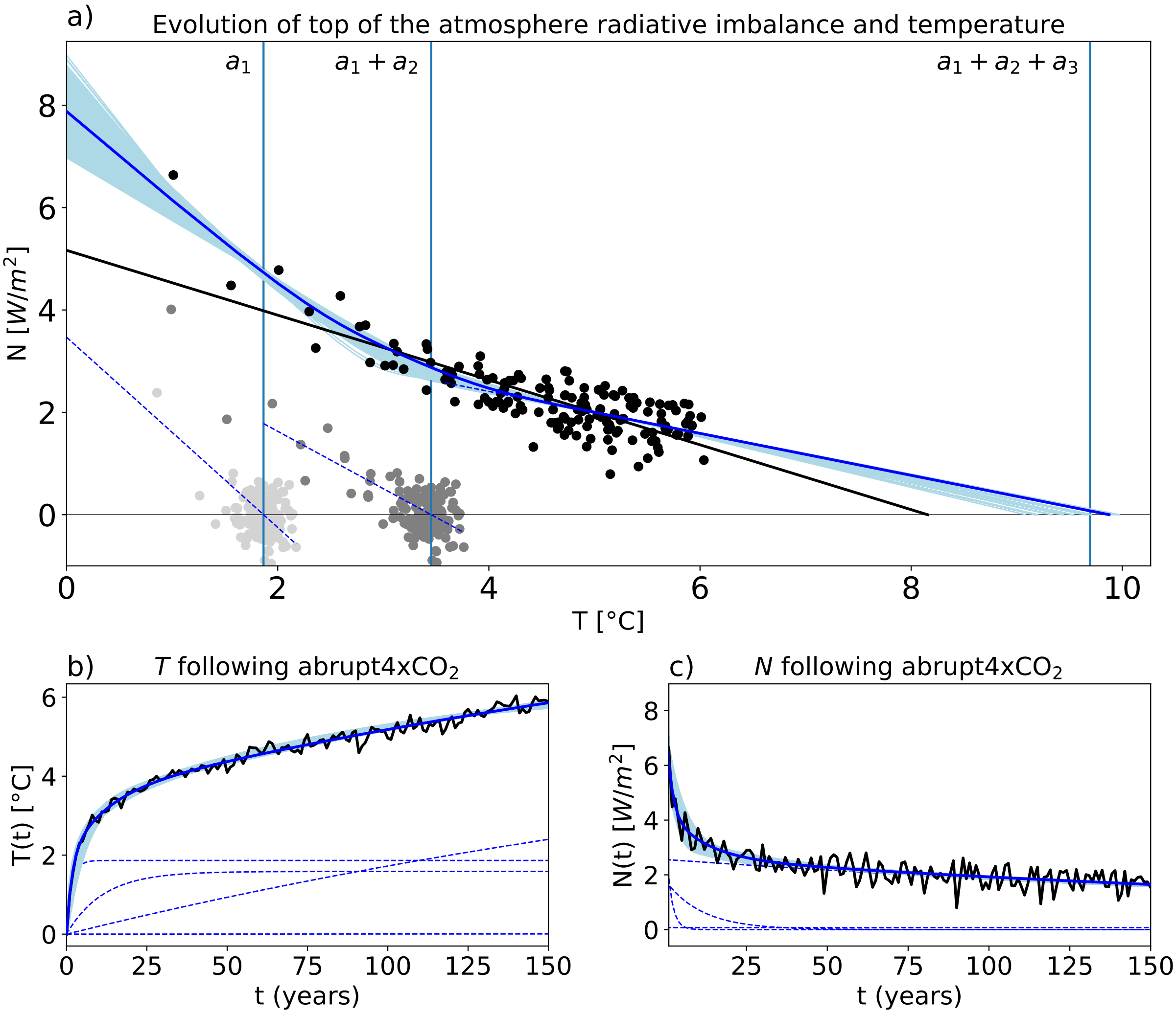

Figure S25. As Figure 1, but for the model CSIRO-Mk3-6-0.

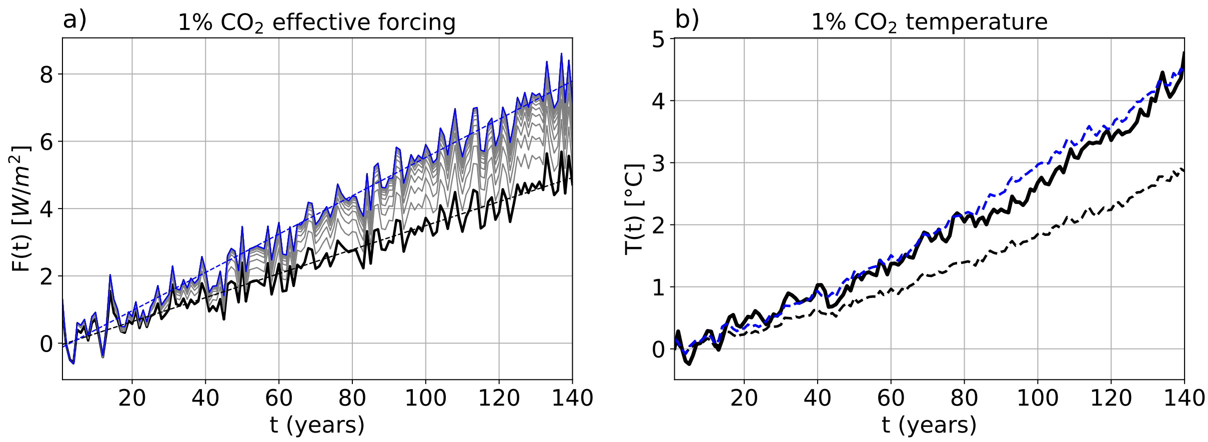

Figure S26. As Figure 3, but for the model CSIRO-Mk3-6-0.

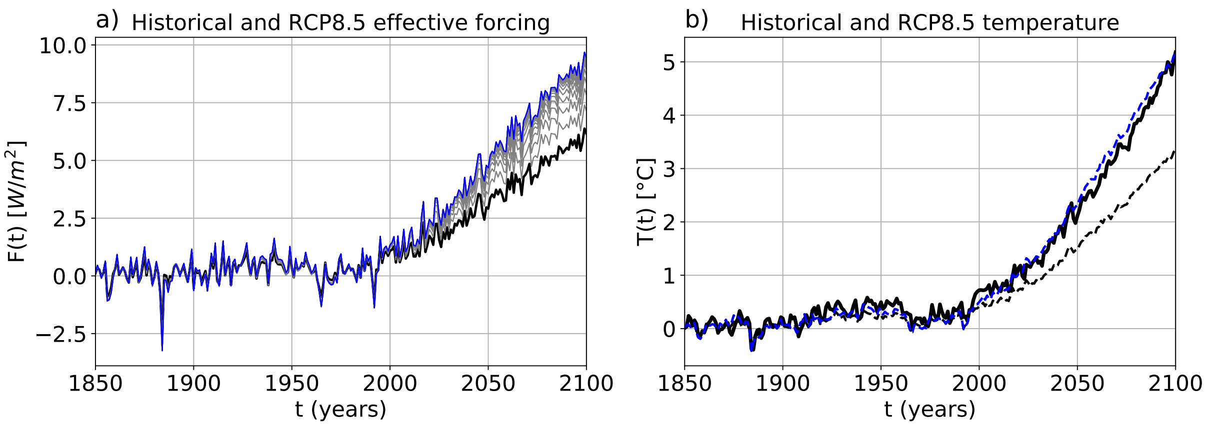

Figure S27. As Figure 4, but for the model CSIRO-Mk3-6-0.

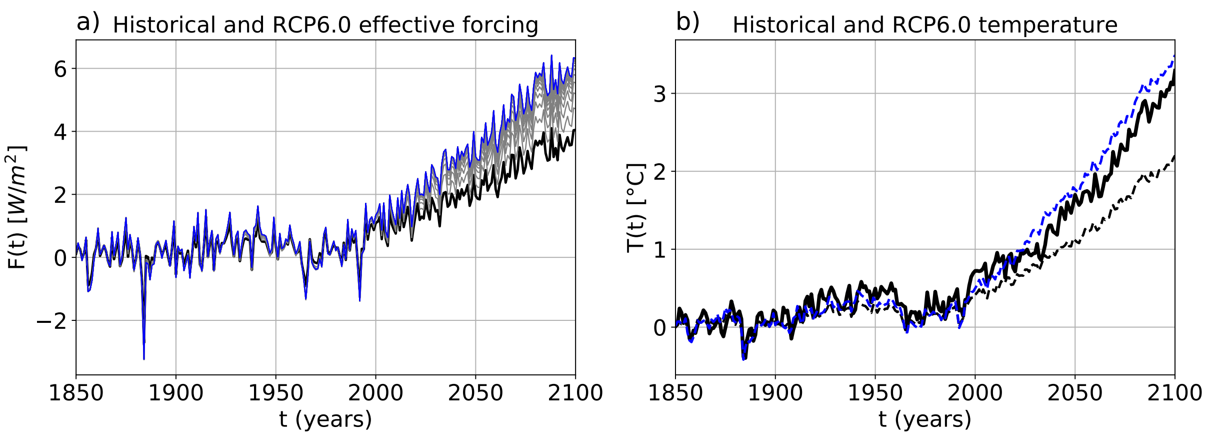

Figure S28. As Figure 4, but for the model CSIRO-Mk3-6-0 and experiment RCP6.0.

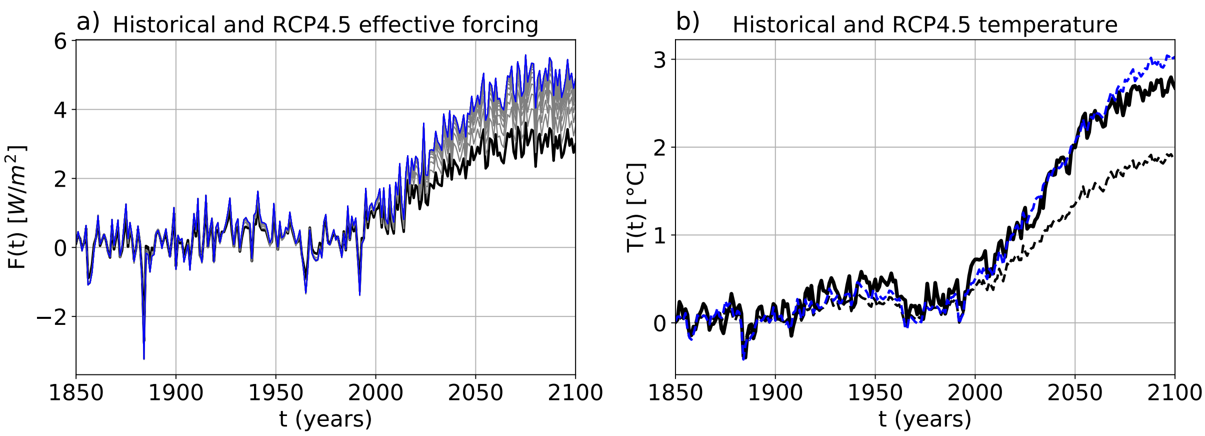

Figure S29. As Figure 4, but for the model CSIRO-Mk3-6-0 and experiment RCP4.5.

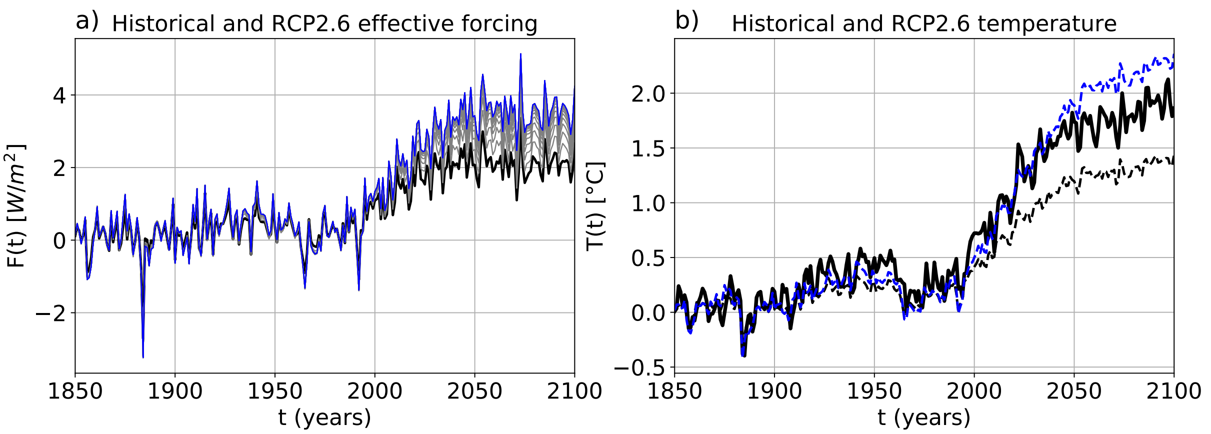

Figure S30. As Figure 4, but for the model CSIRO-Mk3-6-0 and experiment RCP2.6.

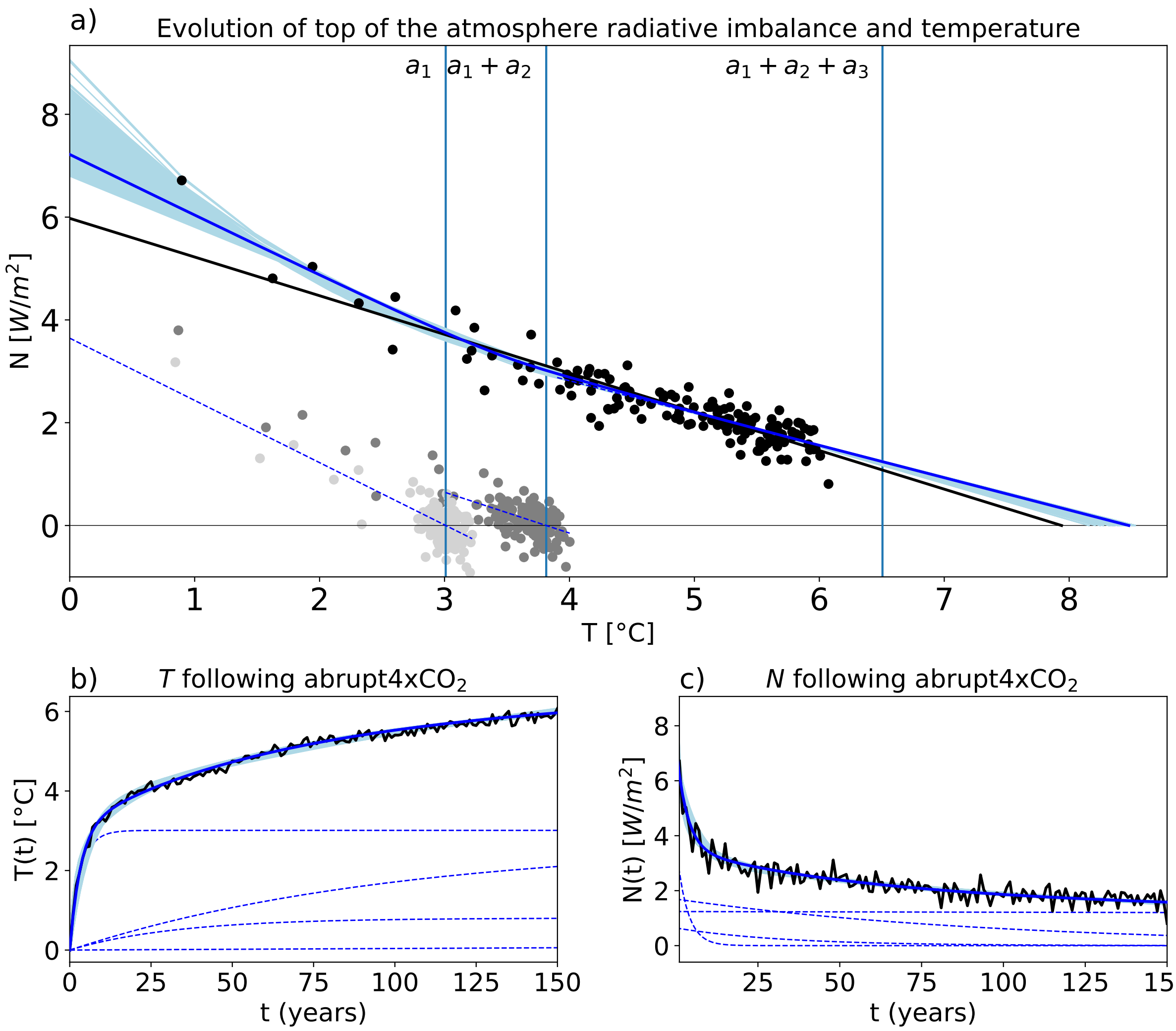

Figure S31. As Figure 1, but for the model GFDL-CM3.

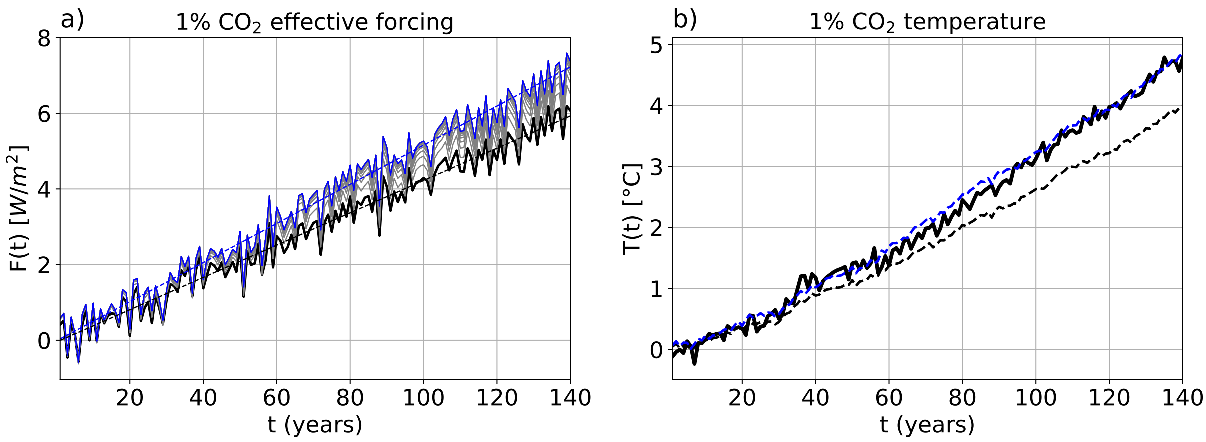

Figure S32. As Figure 3, but for the model GFDL-CM3.

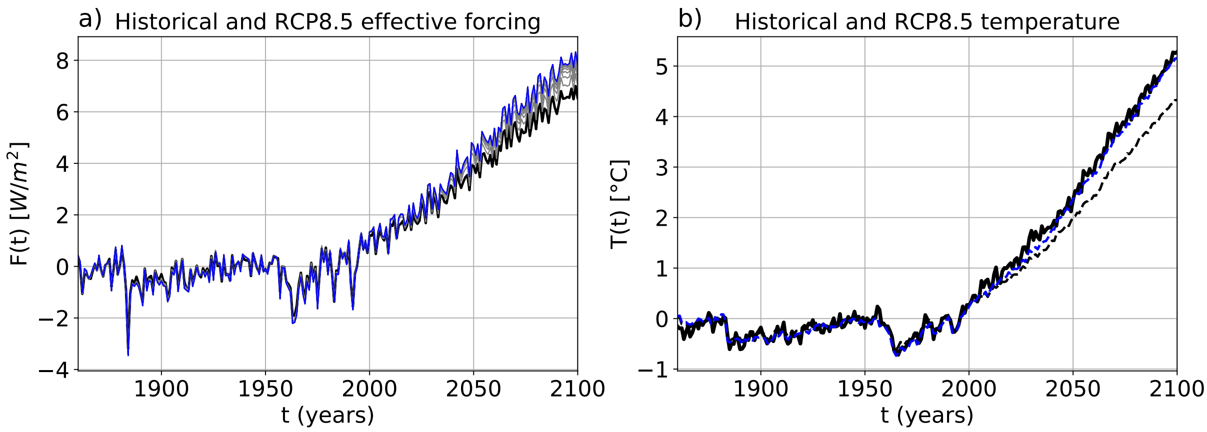

Figure S33. As Figure 4, but for the model GFDL-CM3.

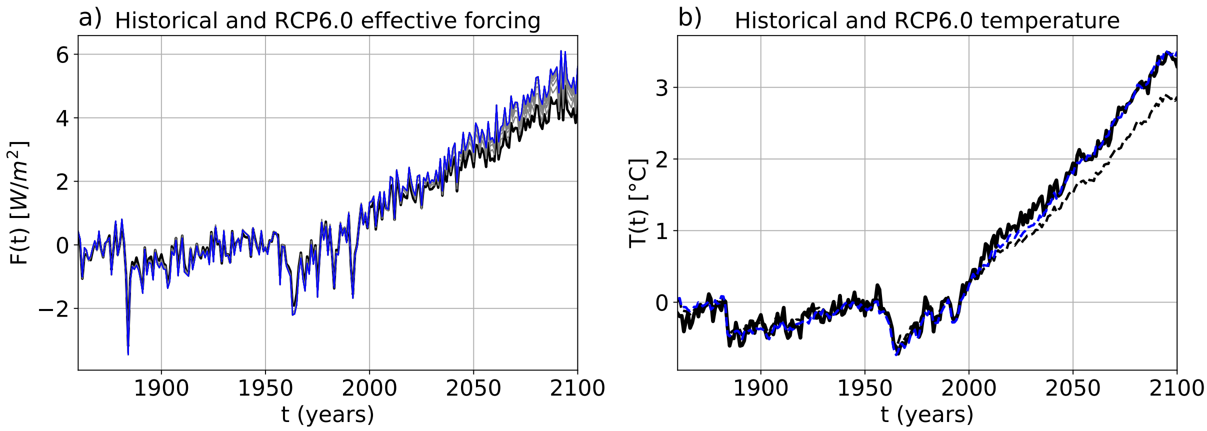

Figure S34. As Figure 4, but for the model GFDL-CM3 and experiment RCP6.0.

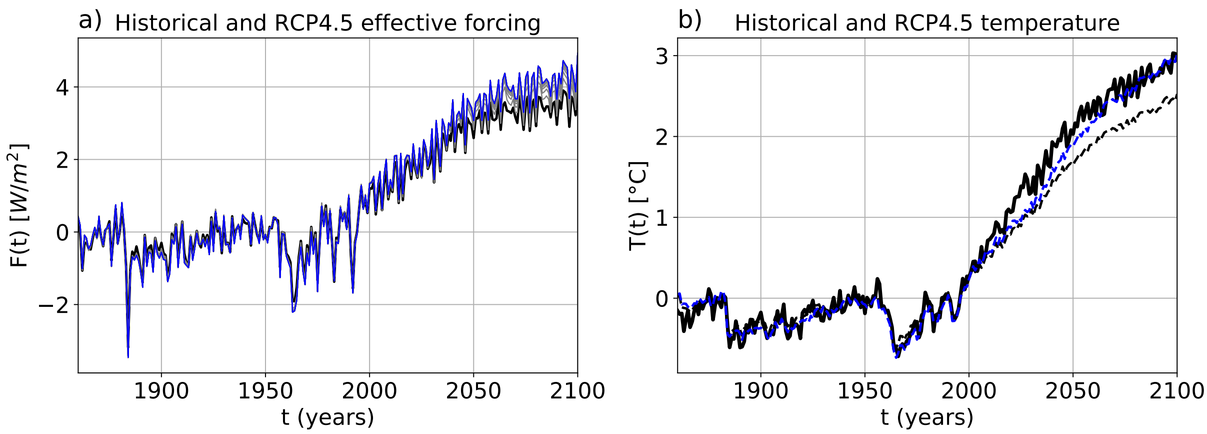

Figure S35. As Figure 4, but for the model GFDL-CM3 and experiment RCP4.5.

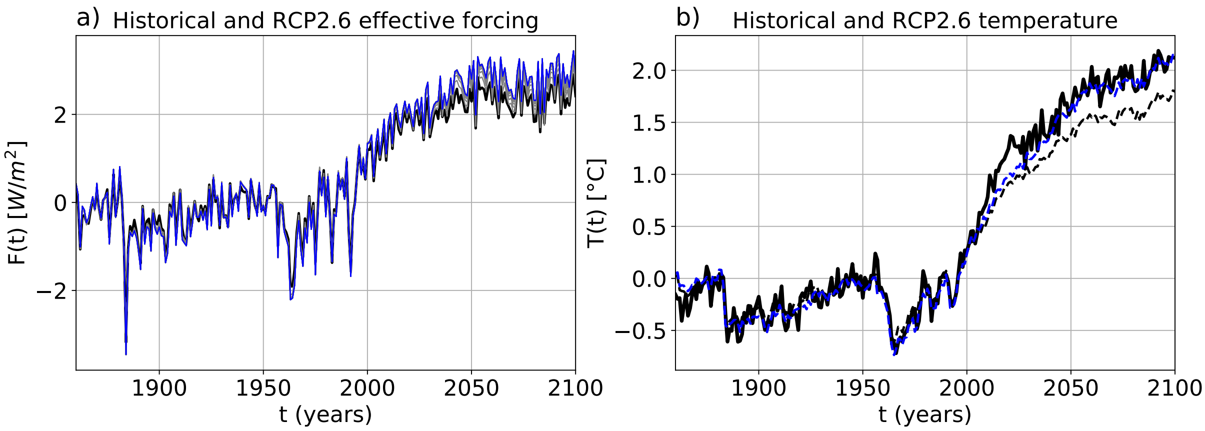

Figure S36. As Figure 4, but for the model GFDL-CM3 and experiment RCP2.6.

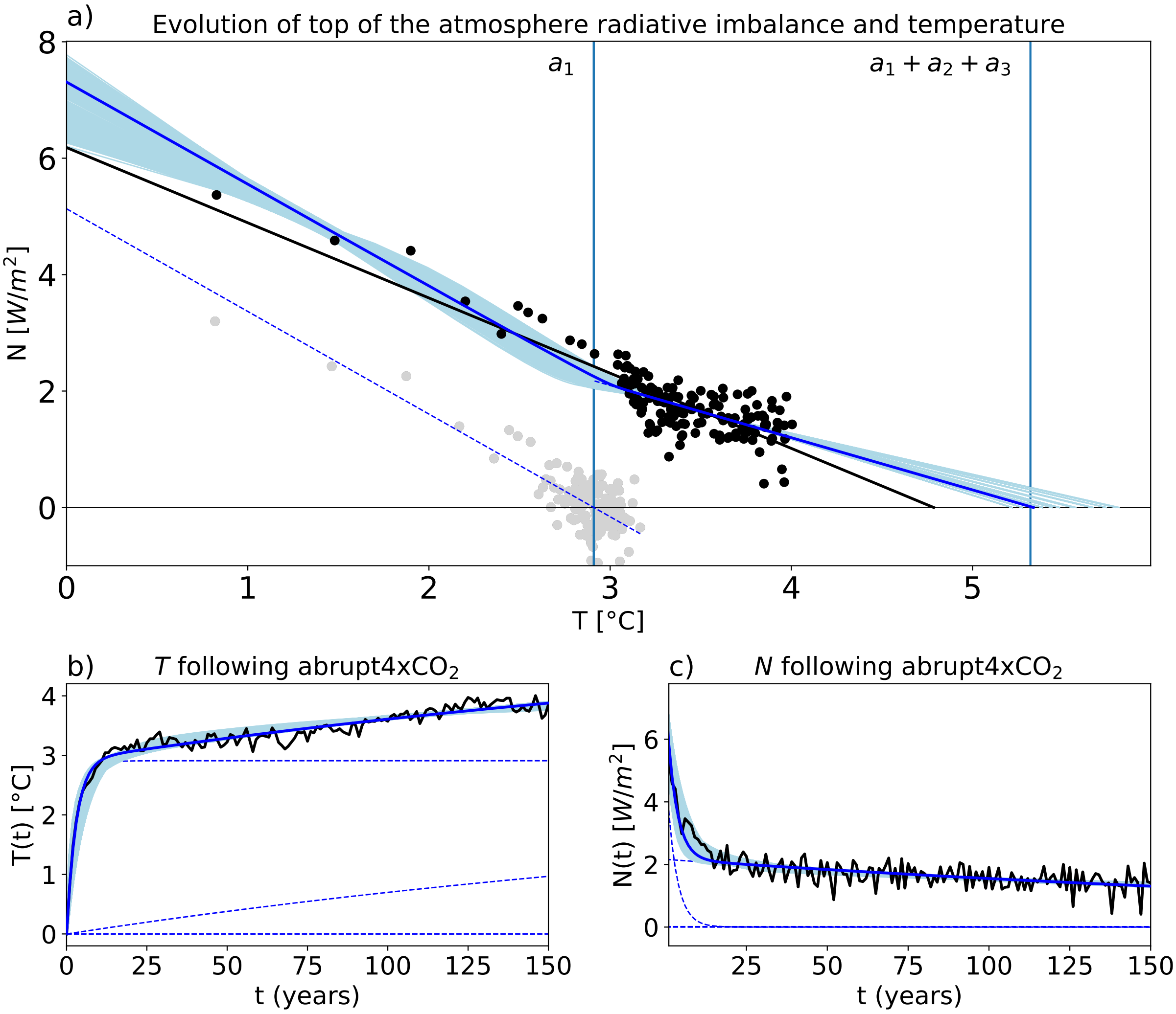

Figure S37. As Figure 1, but for the model GFDL-ESM2G.

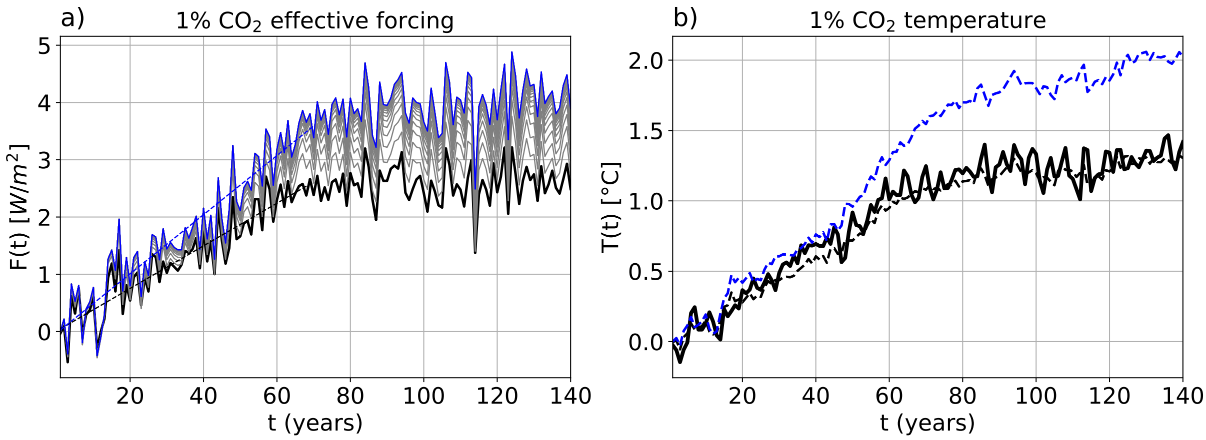

Figure S38. As Figure 3, but for the model GFDL-ESM2G. Note that in this model CO_2 increases only until we reach a doubling after 70 years.

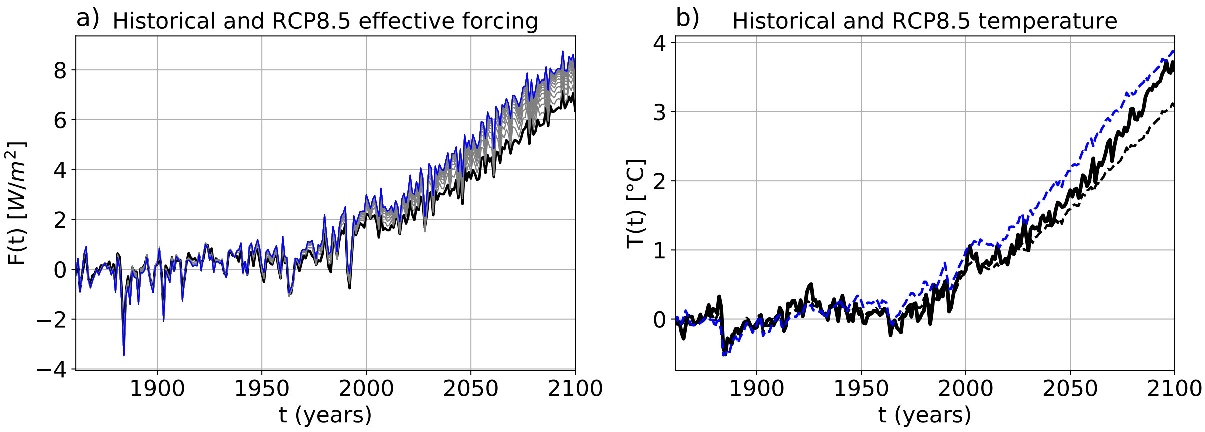

Figure S39. As Figure 4, but for the model GFDL-ESM2G.

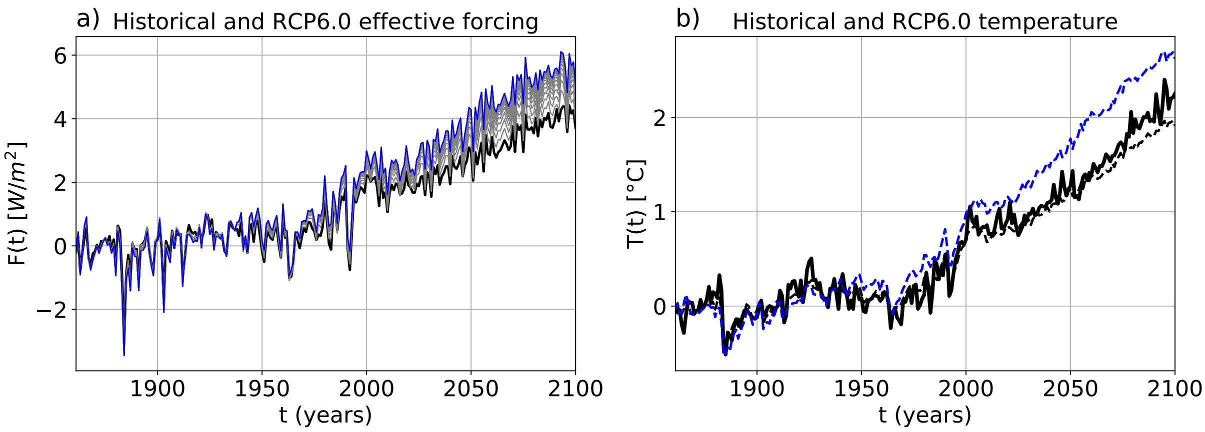

Figure S40. As Figure 4, but for the model GFDL-ESM2G and experiment RCP6.0.

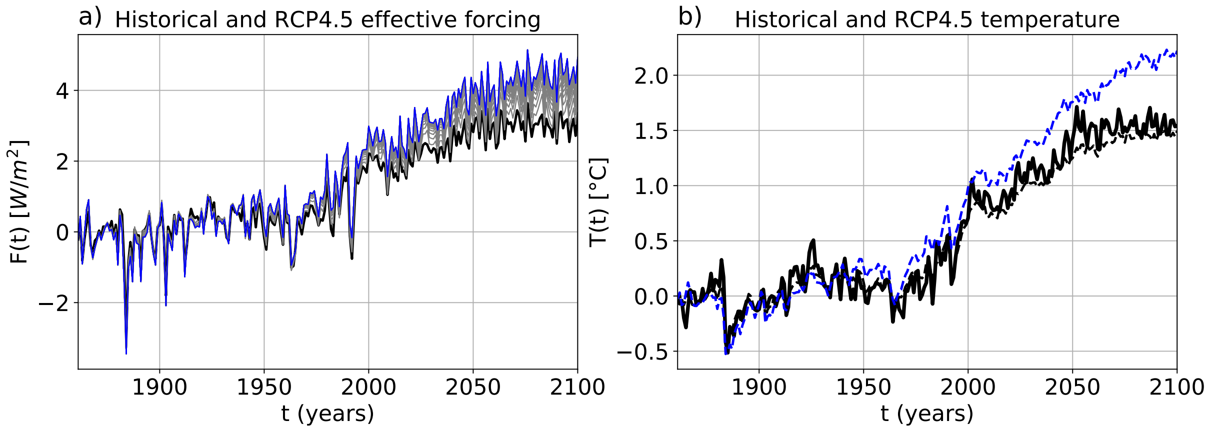

Figure S41. As Figure 4, but for the model GFDL-ESM2G and experiment RCP4.5.

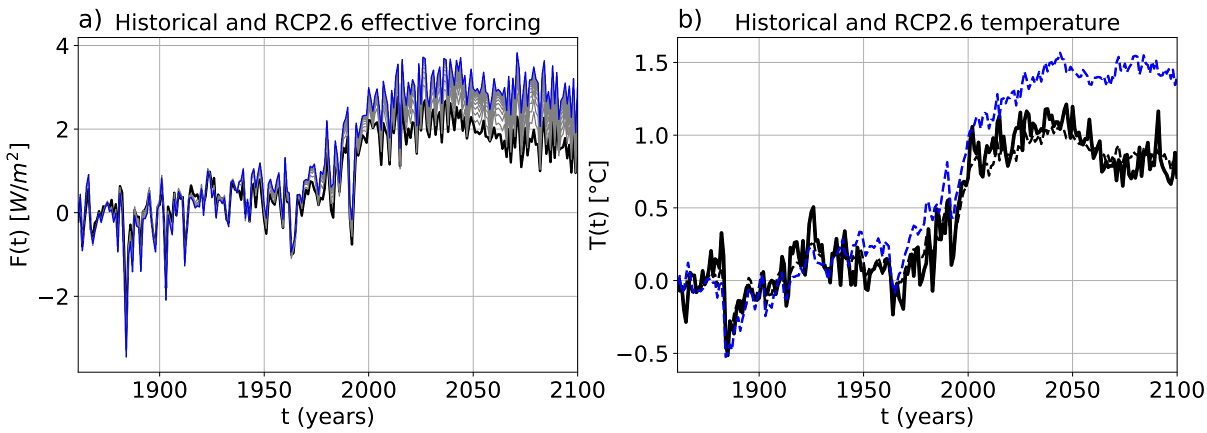

Figure S42. As Figure 4, but for the model GFDL-ESM2G and experiment RCP2.6.

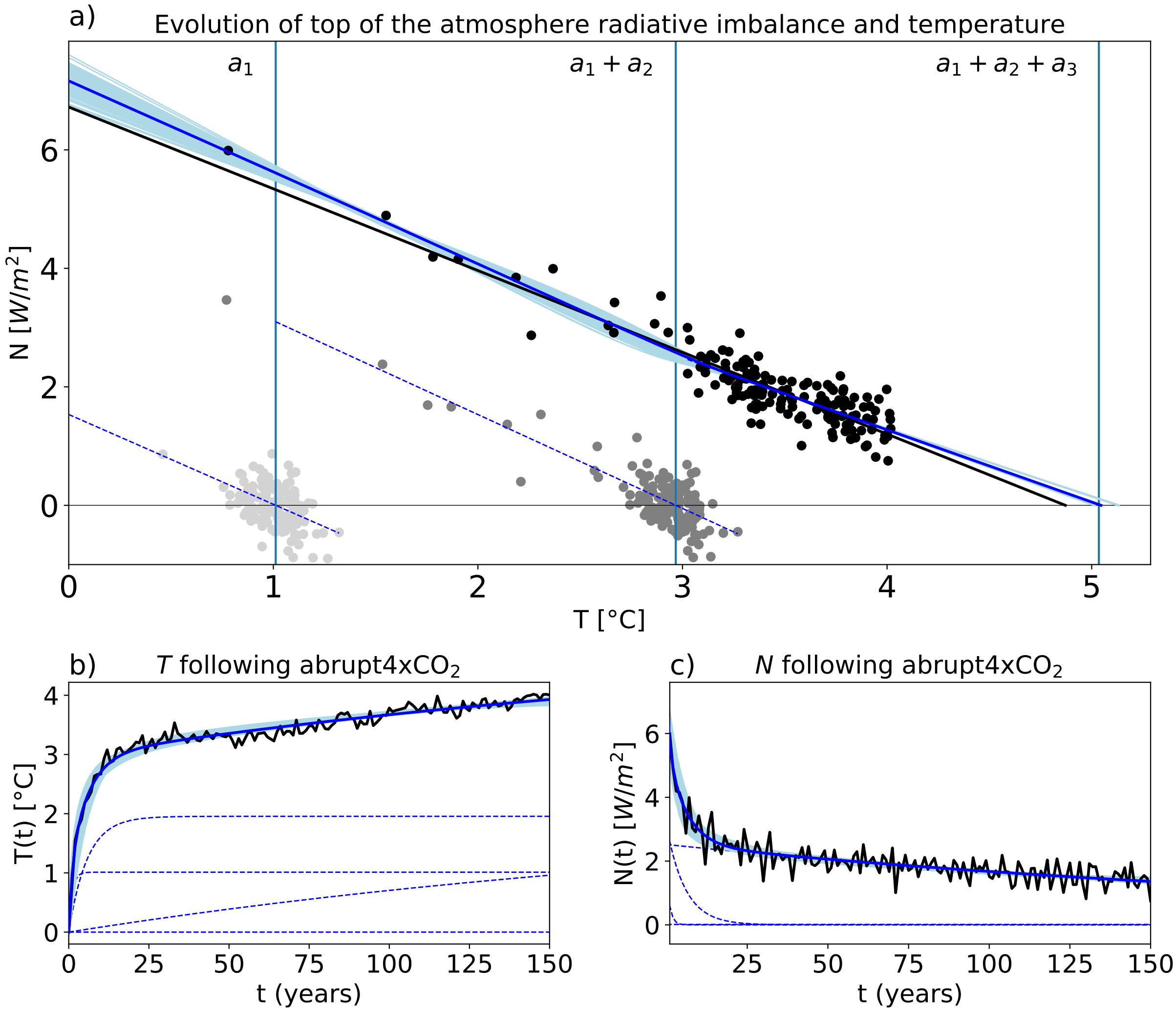

Figure S43. As Figure 1, but for the model GFDL-ESM2M.

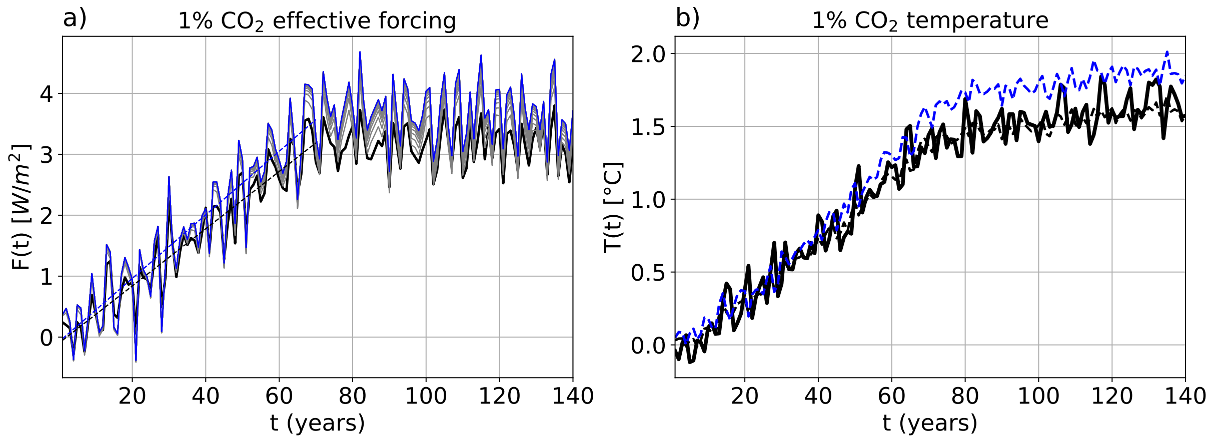

Figure S44. As Figure 3, but for the model GFDL-ESM2M. Note that in this model CO_2 increases only until we reach a doubling after 70 years.

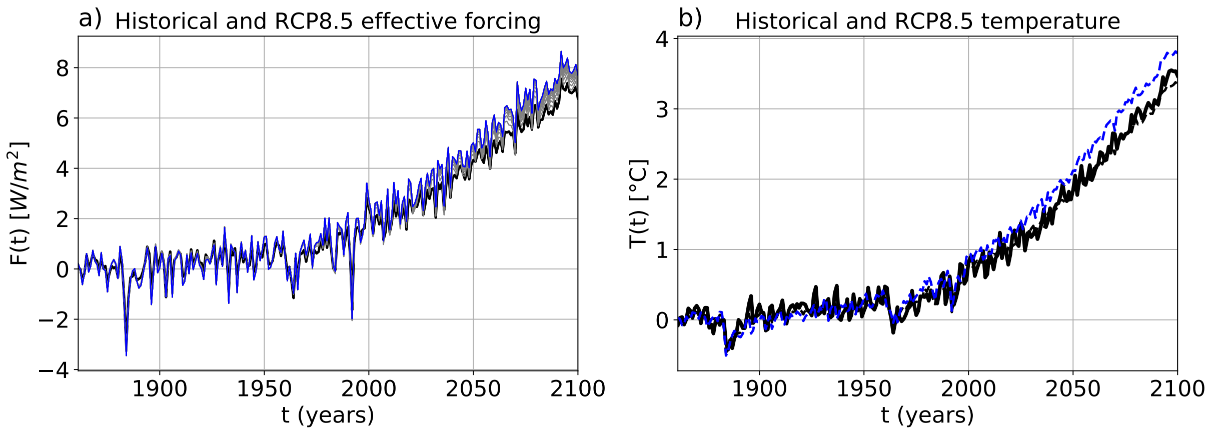

Figure S45. As Figure 4, but for the model GFDL-ESM2M.

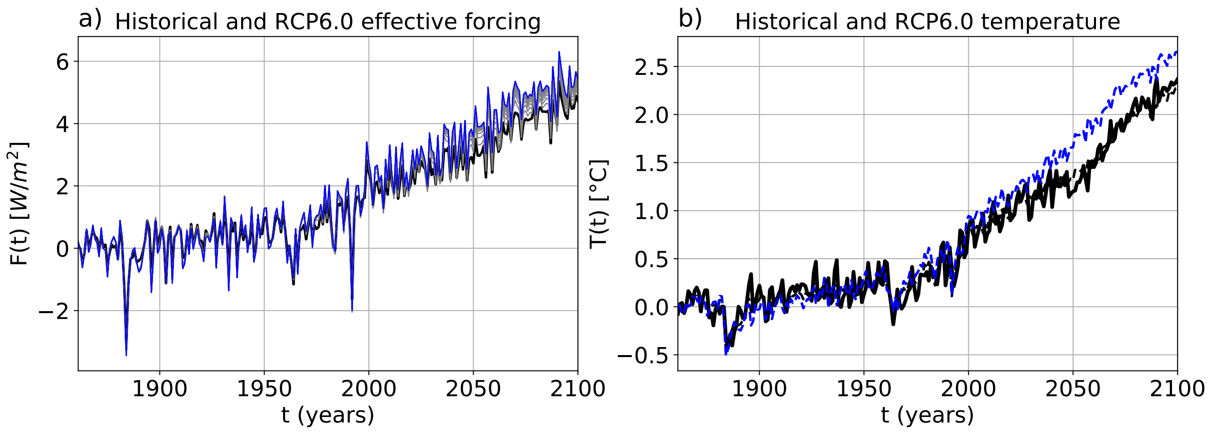

Figure S46. As Figure 4, but for the model GFDL-ESM2M and experiment RCP6.0.

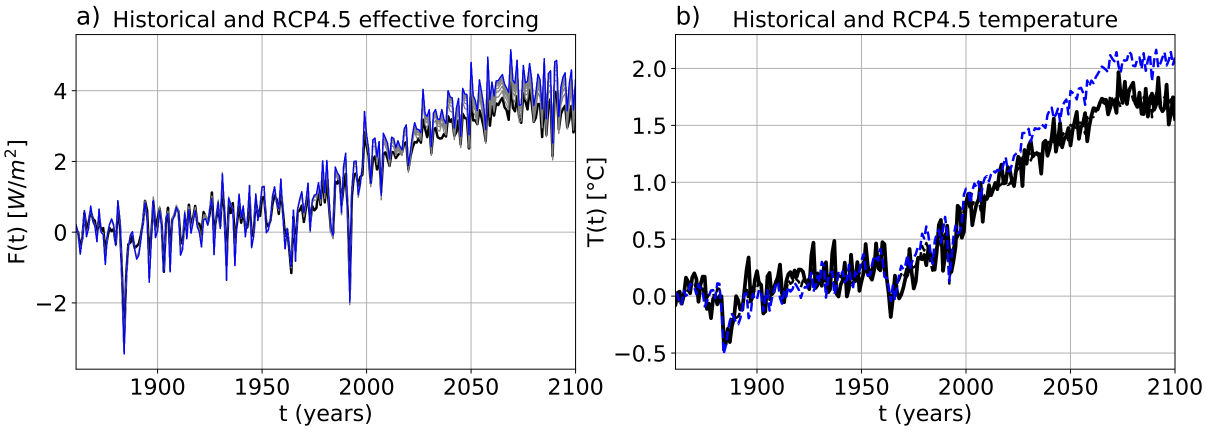

Figure S47. As Figure 4, but for the model GFDL-ESM2M and experiment RCP4.5.

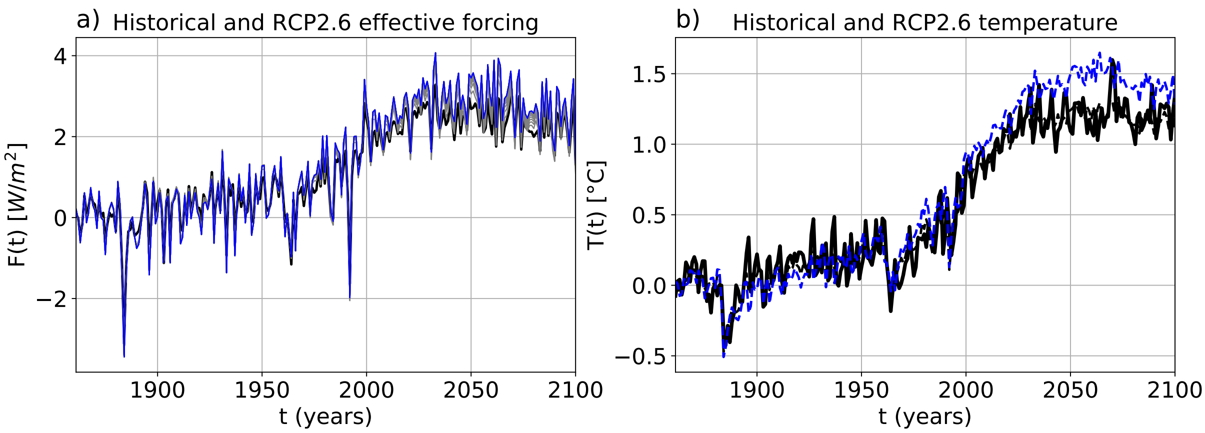

Figure S48. As Figure 4, but for the model GFDL-ESM2M and experiment RCP2.6.

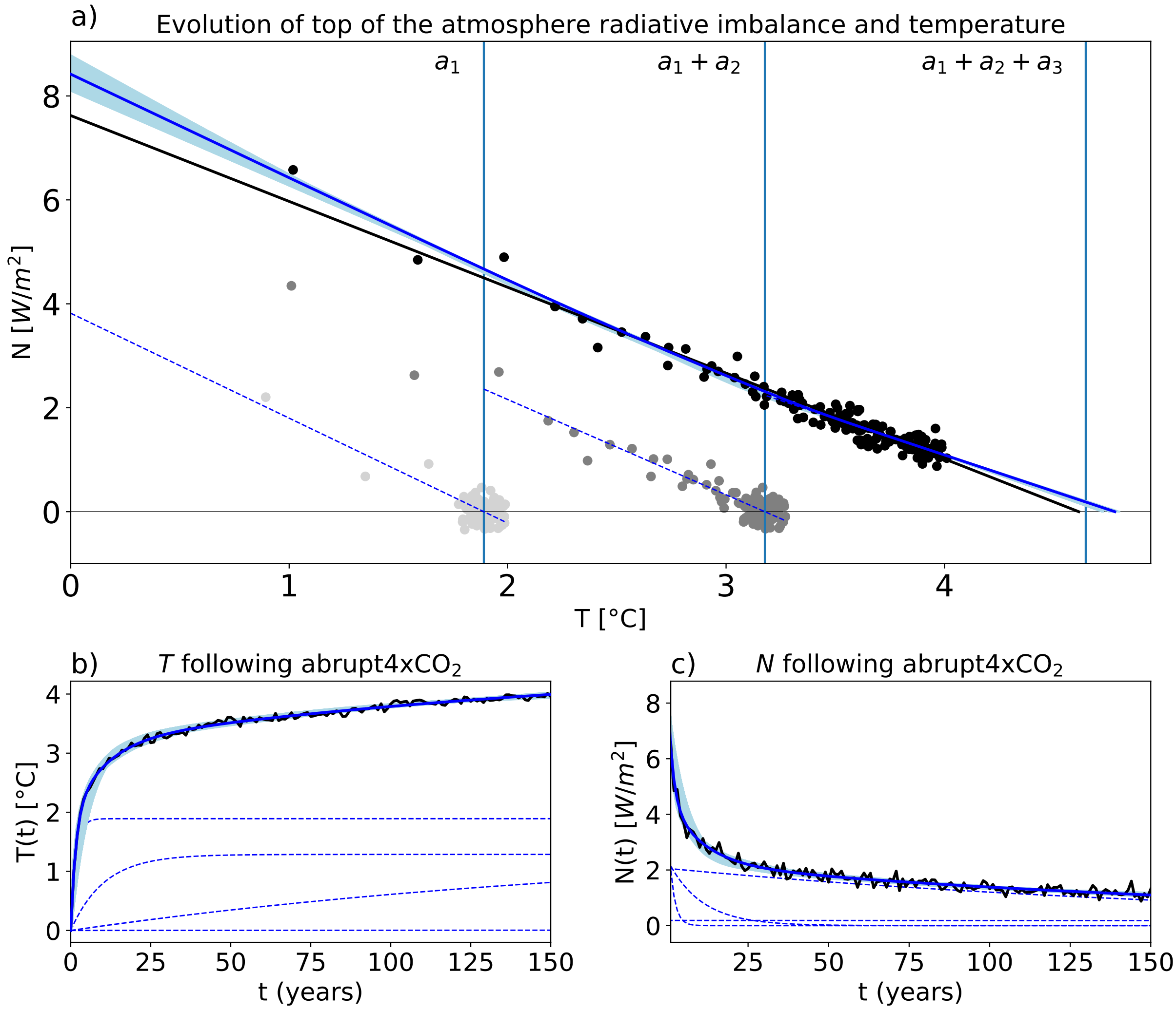

Figure S49. As Figure 1, but for the model GISS-E2-H.

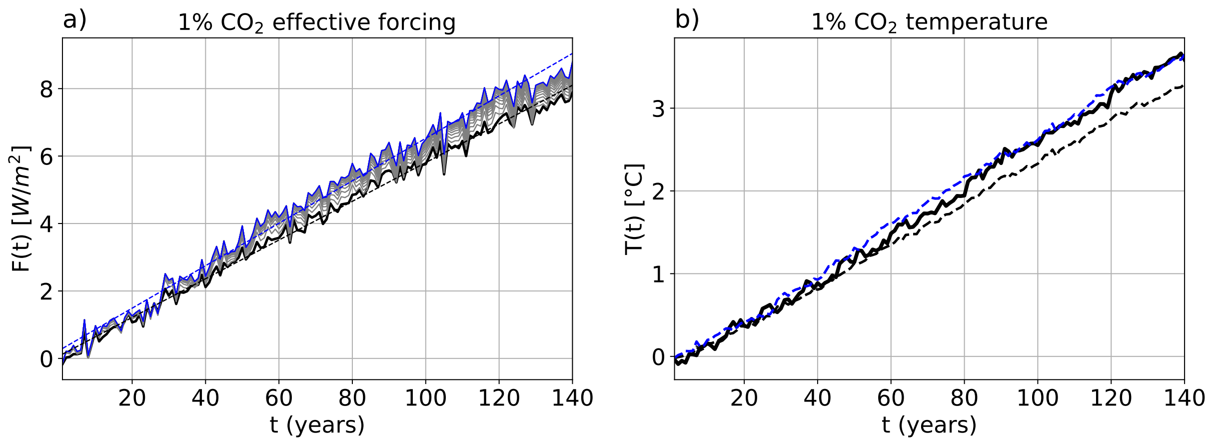

Figure S50. As Figure 3, but for the model GISS-E2-H.

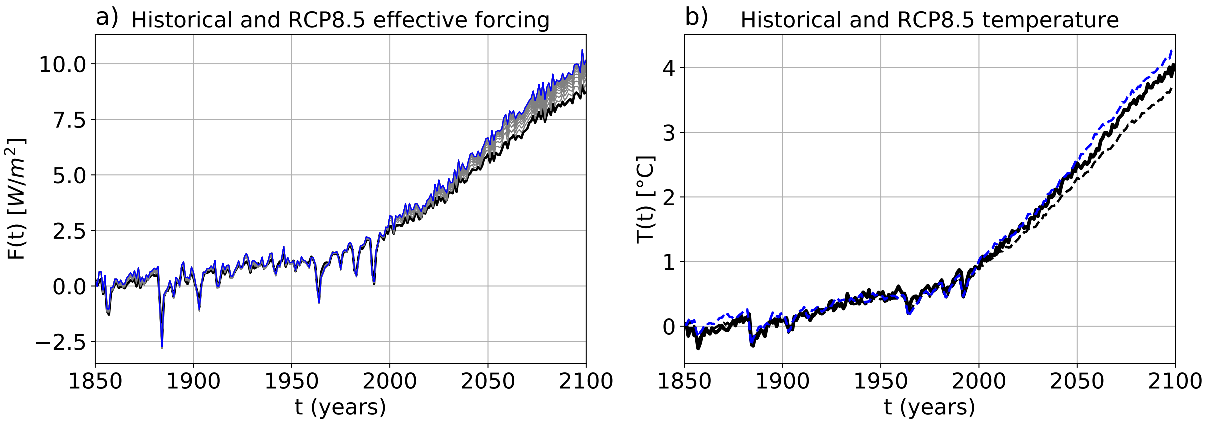

Figure S51. As Figure 4, but for the model GISS-E2-H.

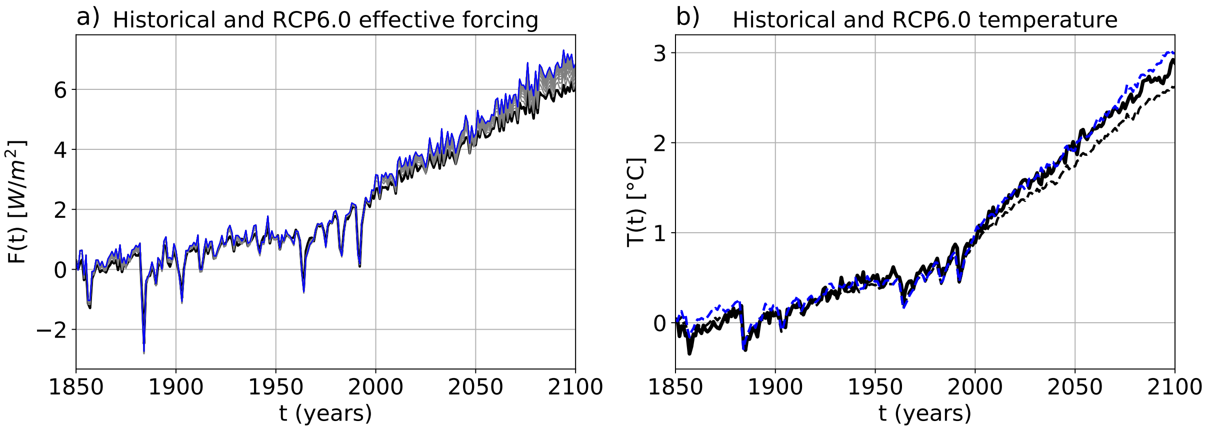

Figure S52. As Figure 4, but for the model GISS-E2-H and experiment RCP6.0.

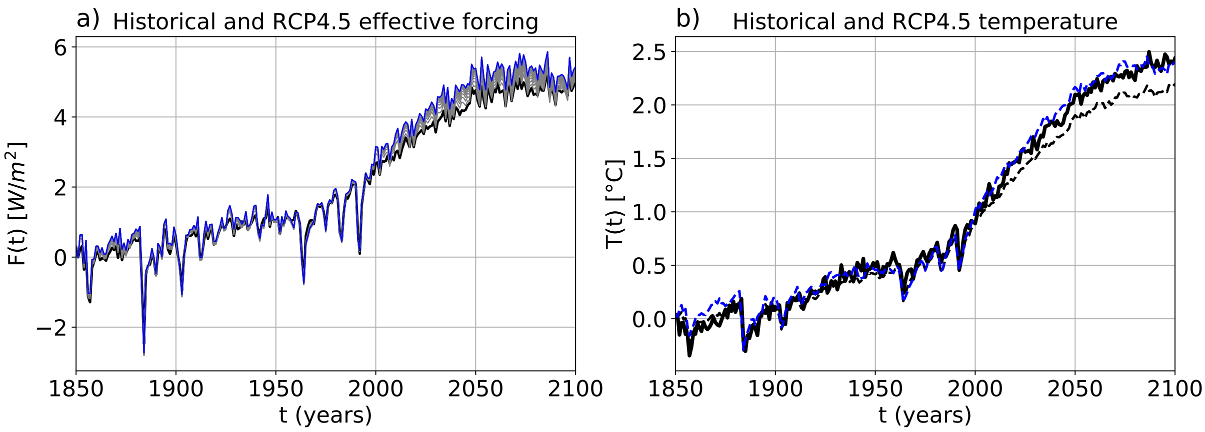

Figure S53. As Figure 4, but for the model GISS-E2-H and experiment RCP4.5.

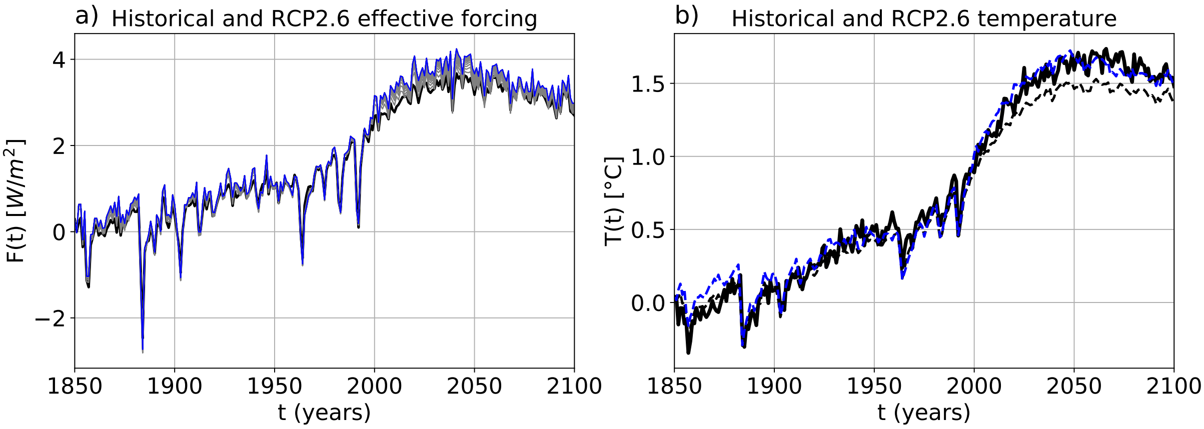

Figure S54. As Figure 4, but for the model GISS-E2-H and experiment RCP2.6.

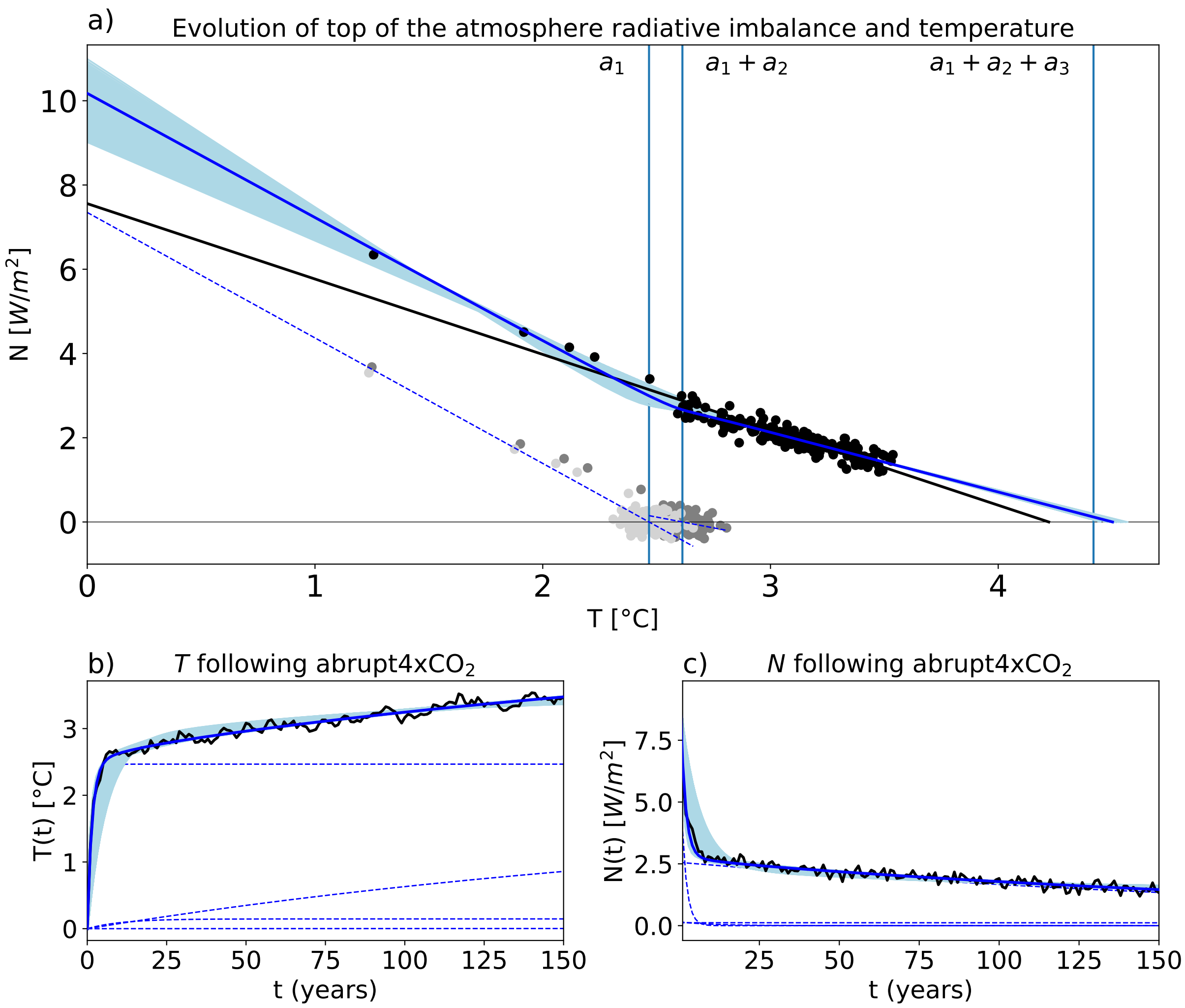

Figure S55. As Figure 1, but for the model GISS-E2-R.

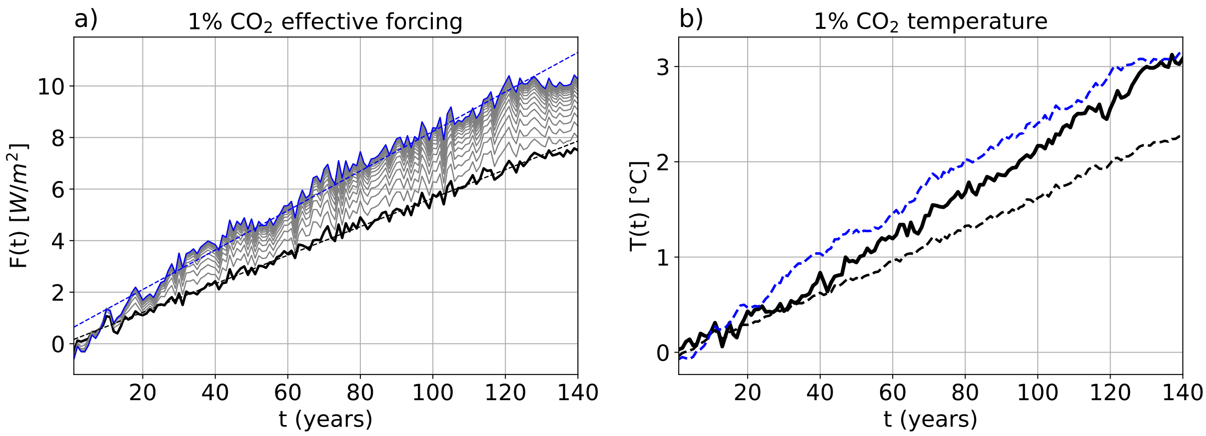

Figure S56. As Figure 3, but for the model GISS-E2-R.

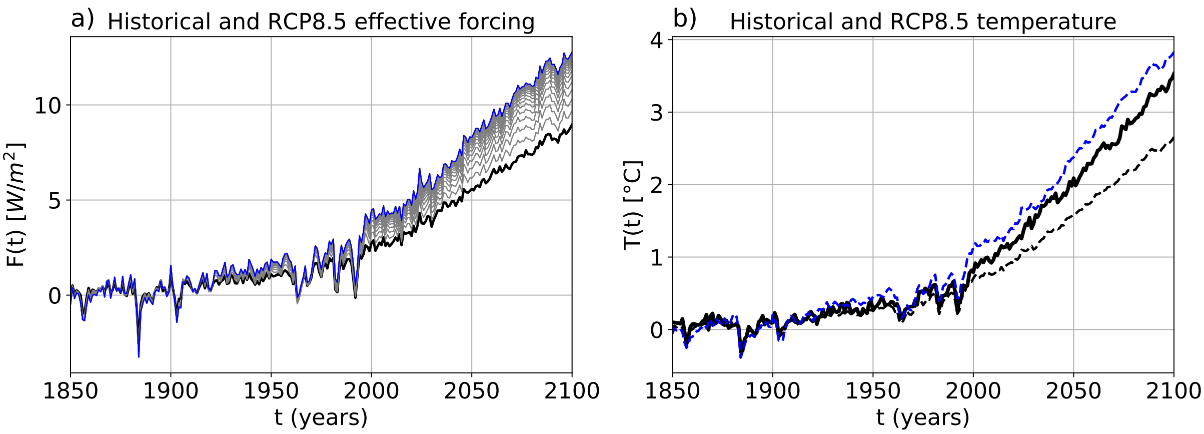

Figure S57. As Figure 4, but for the model GISS-E2-R.

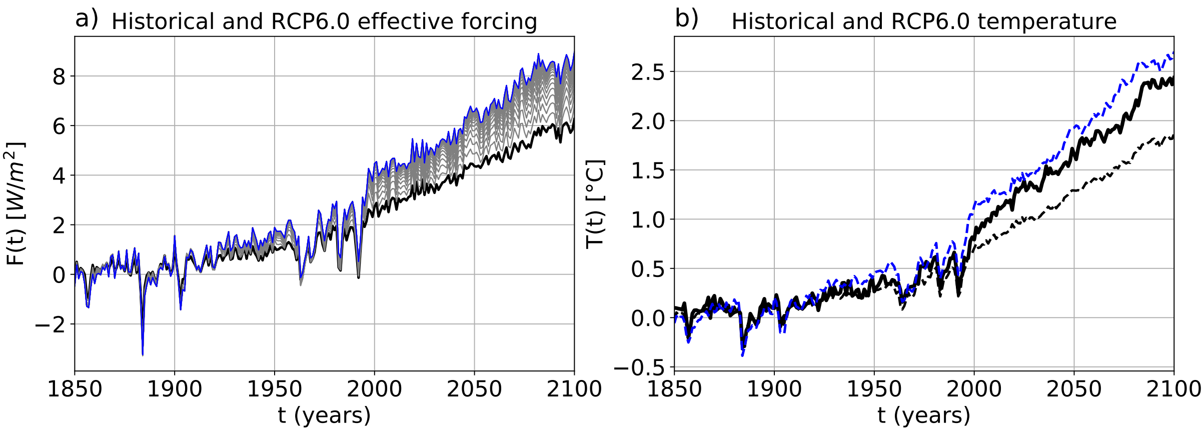

Figure S58. As Figure 4, but for the model GISS-E2-R and experiment RCP6.0.

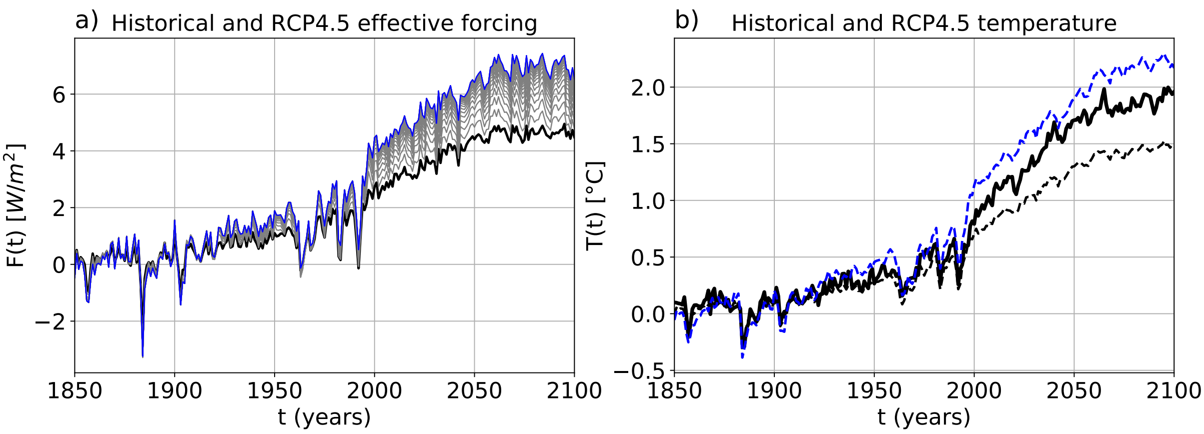

Figure S59. As Figure 4, but for the model GISS-E2-R and experiment RCP4.5.

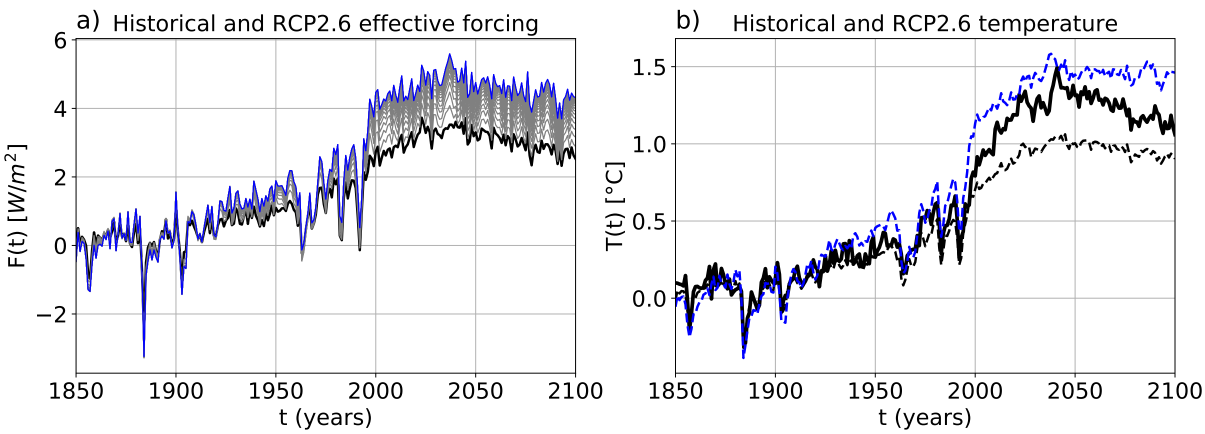

Figure S60. As Figure 4, but for the model GISS-E2-R and experiment RCP2.6.

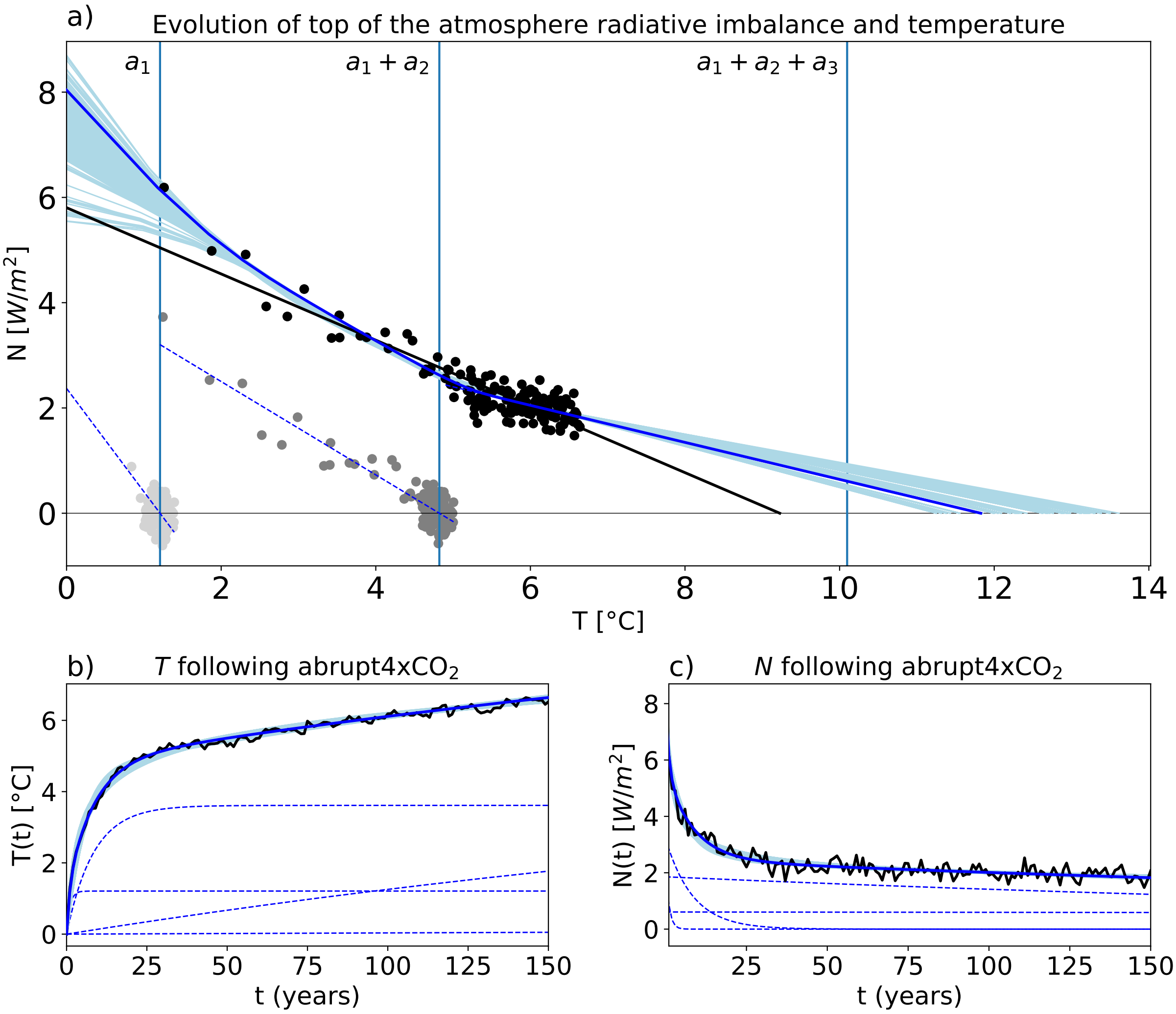

Figure S61. As Figure 1, but for the model HadGEM2-ES.

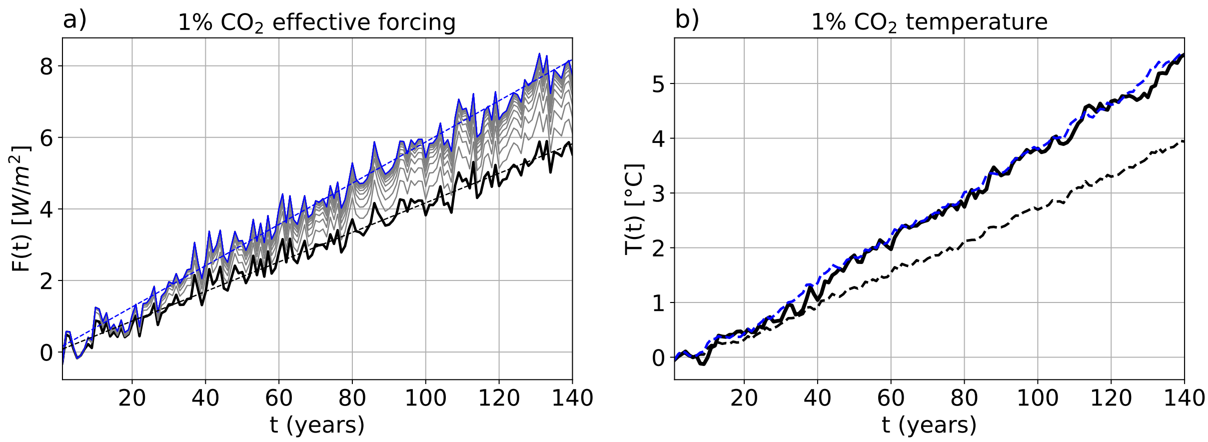

Figure S62. As Figure 3, but for the model HadGEM2-ES.

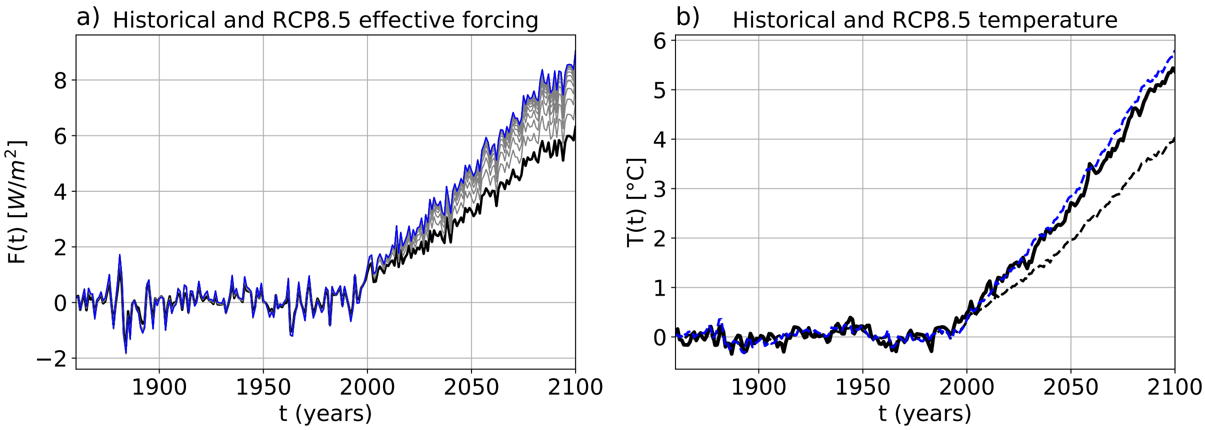

Figure S63. As Figure 4, but for the model HadGEM2-ES.

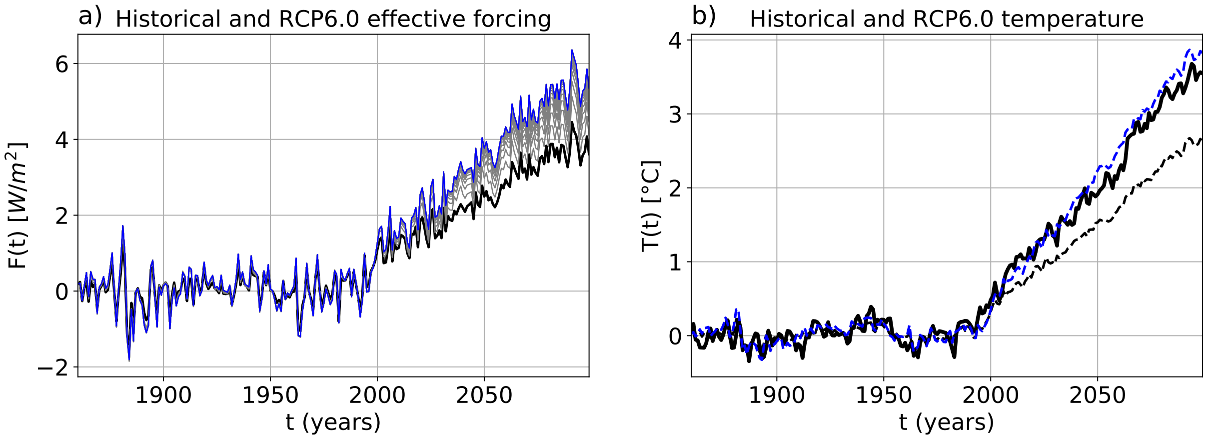

Figure S64. As Figure 4, but for the model HadGEM2-ES and experiment RCP6.0.

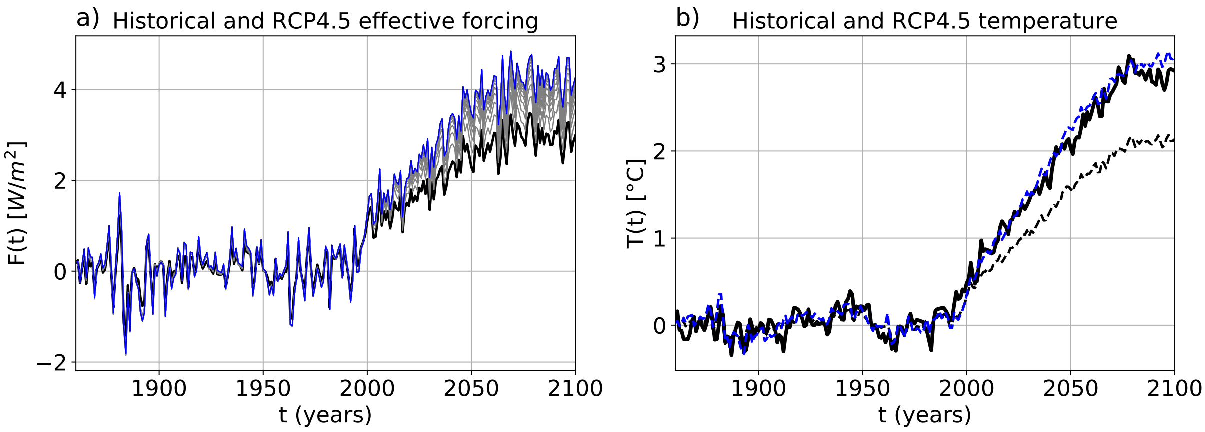

Figure S65. As Figure 4, but for the model HadGEM2-ES and experiment RCP4.5.

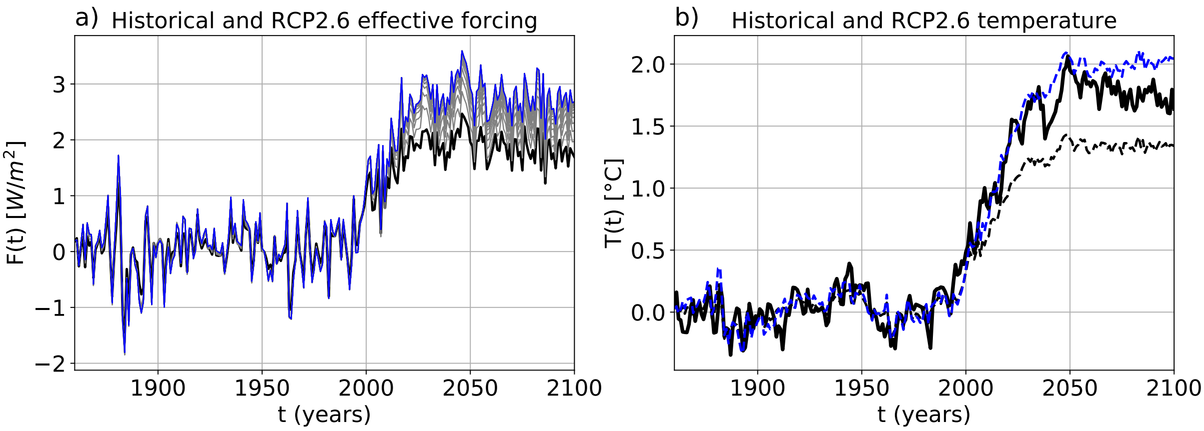

Figure S66. As Figure 4, but for the model HadGEM2-ES and experiment RCP2.6.

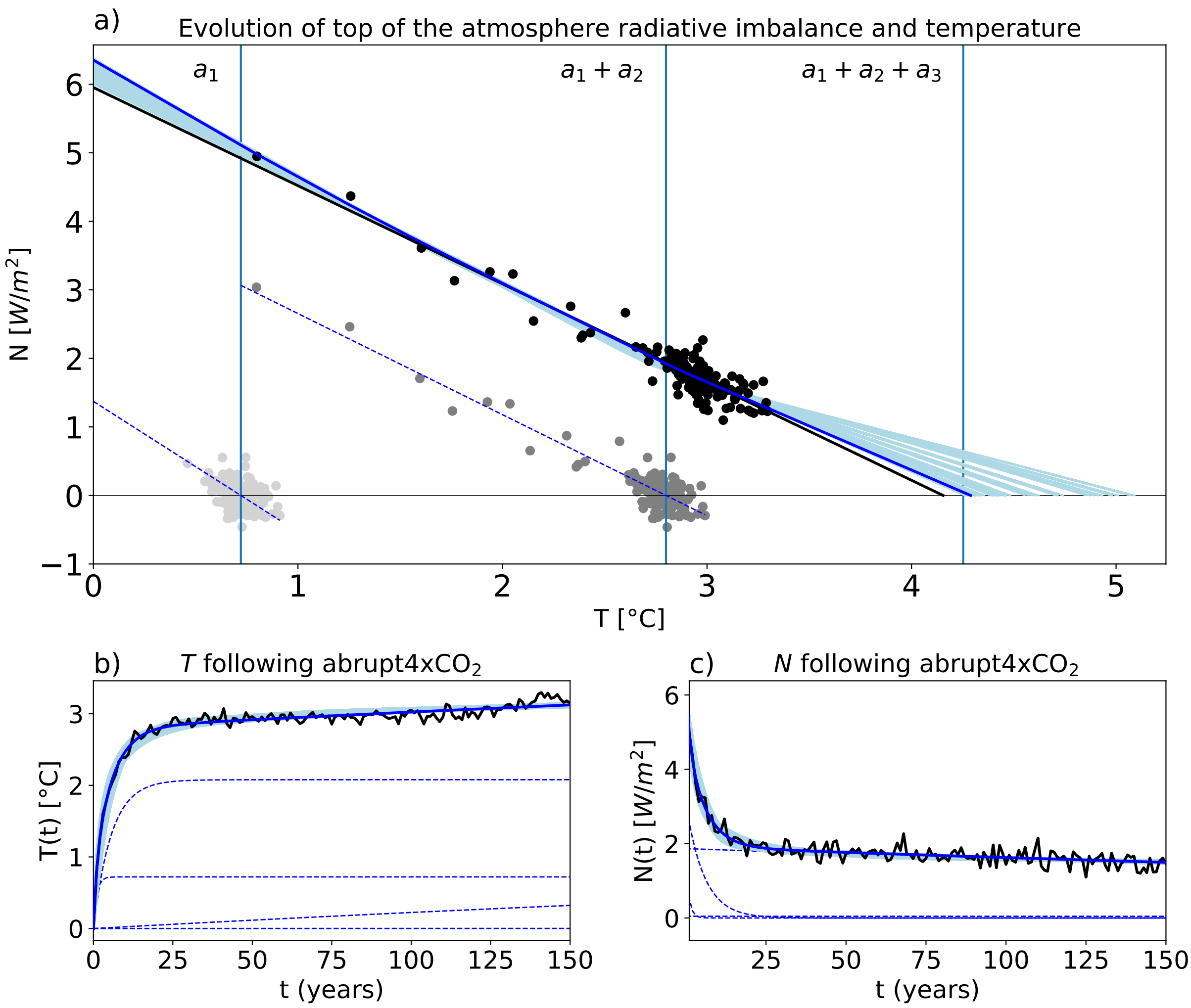

Figure S67. As Figure 1, but for the model inmcm4.

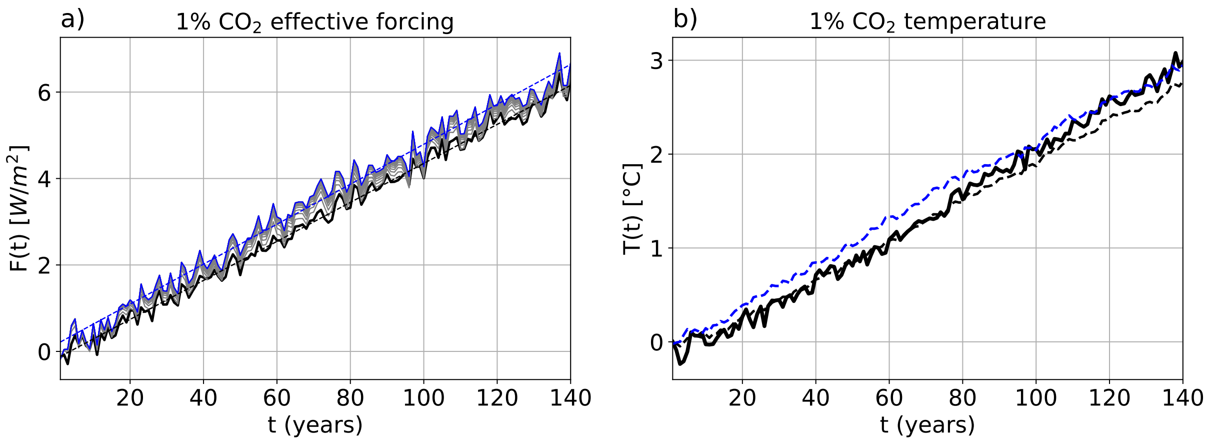

Figure S68. As Figure 3, but for the model inmcm4.

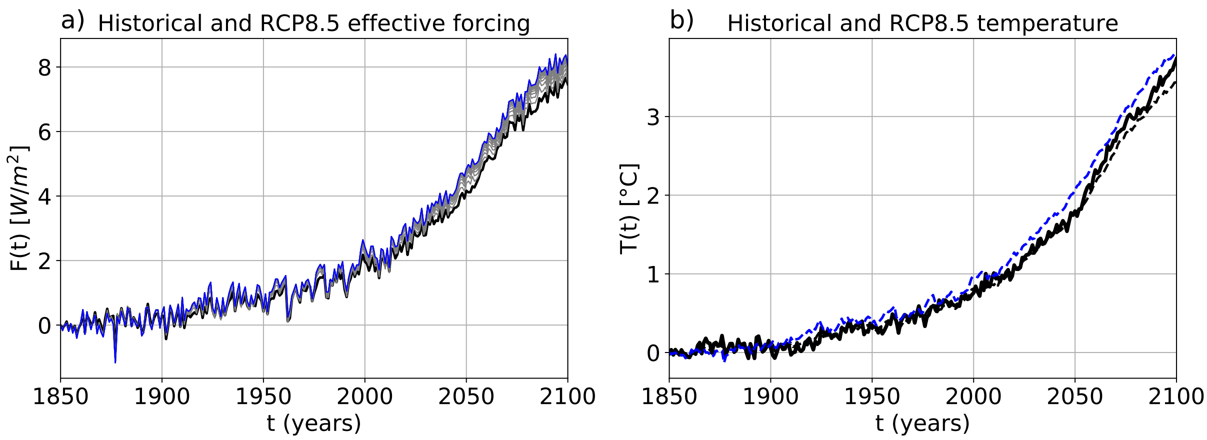

Figure S69. As Figure 4, but for the model inmcm4.

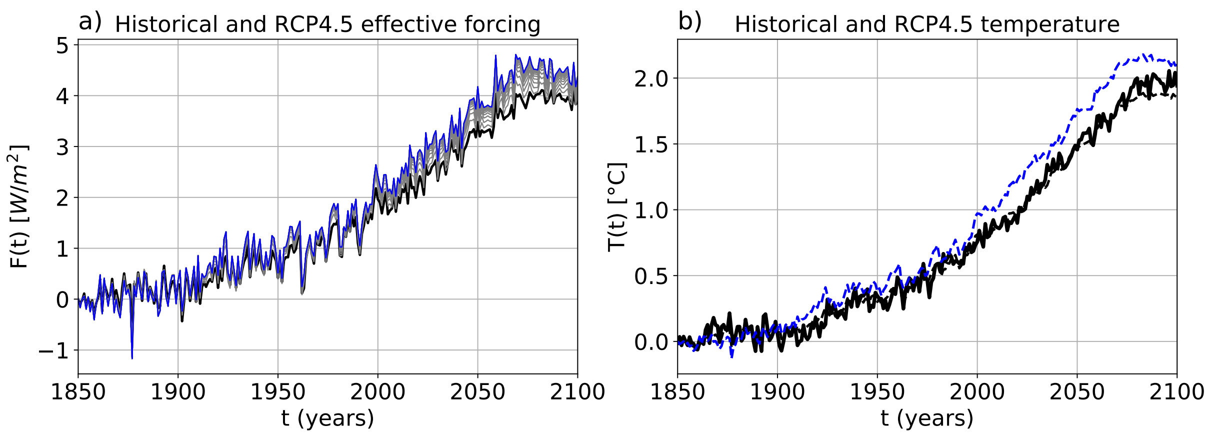

Figure S70. As Figure 4, but for the model inmcm4 and experiment RCP4.5.

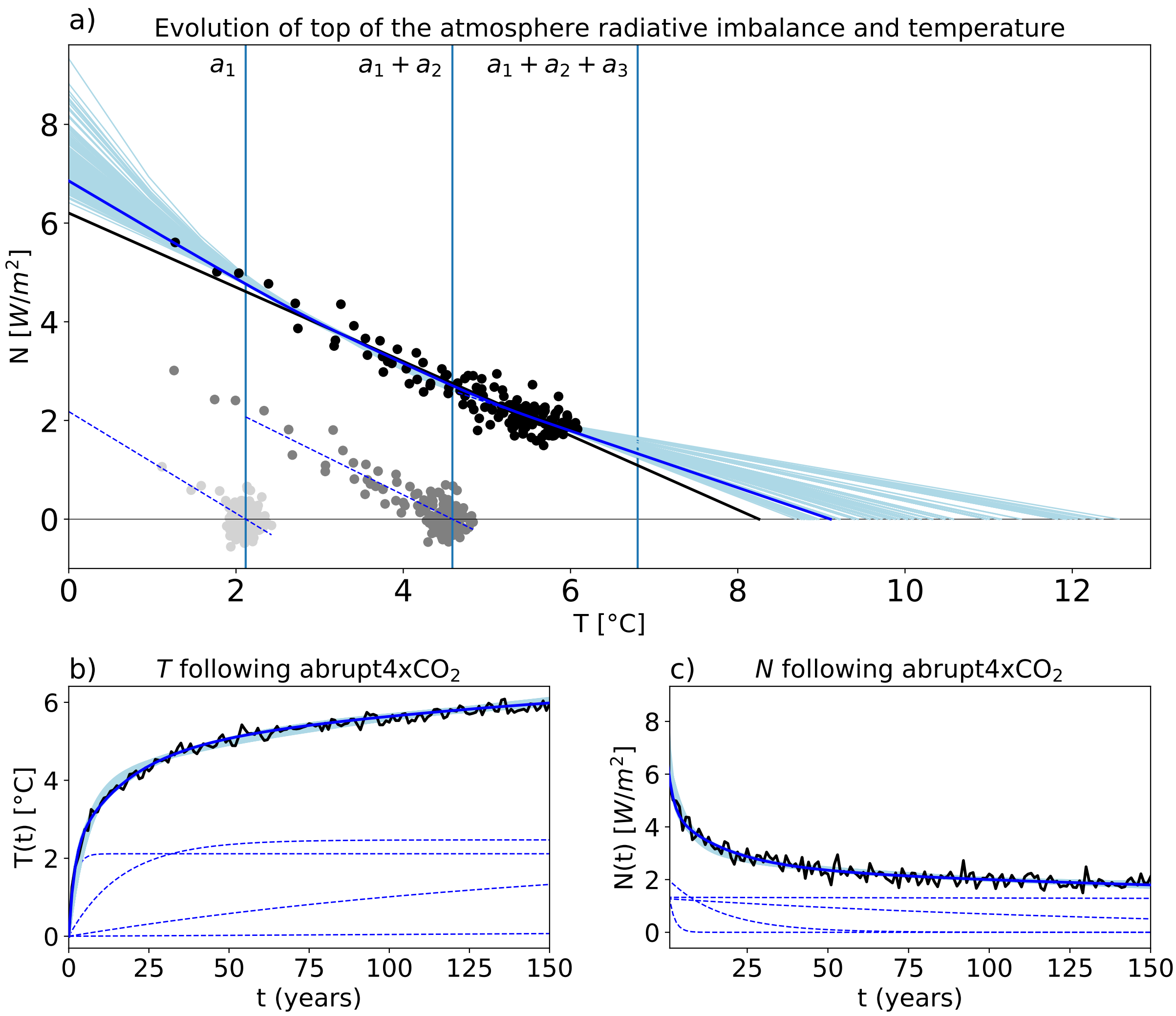

Figure S71. As Figure 1, but for the model IPSL-CM5A-LR.

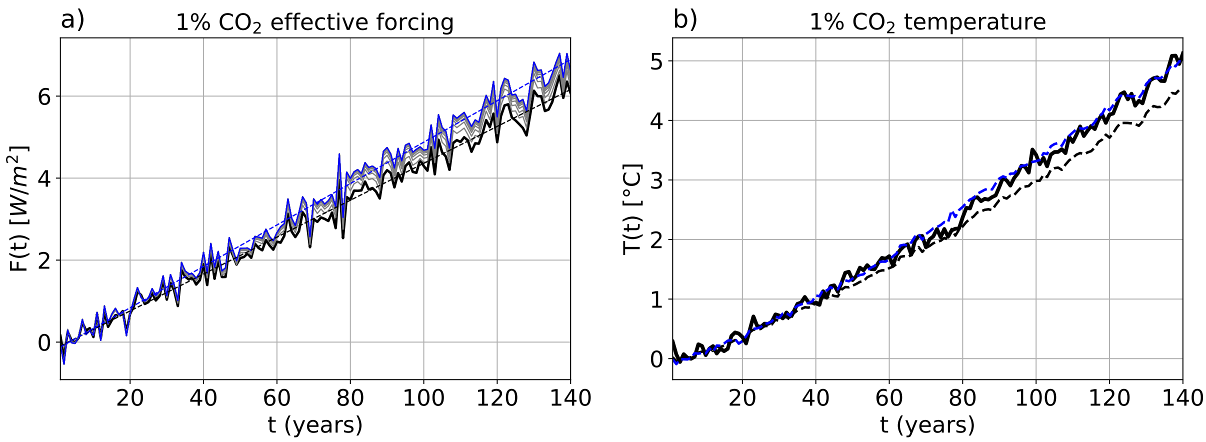

Figure S72. As Figure 3, but for the model IPSL-CM5A-LR.

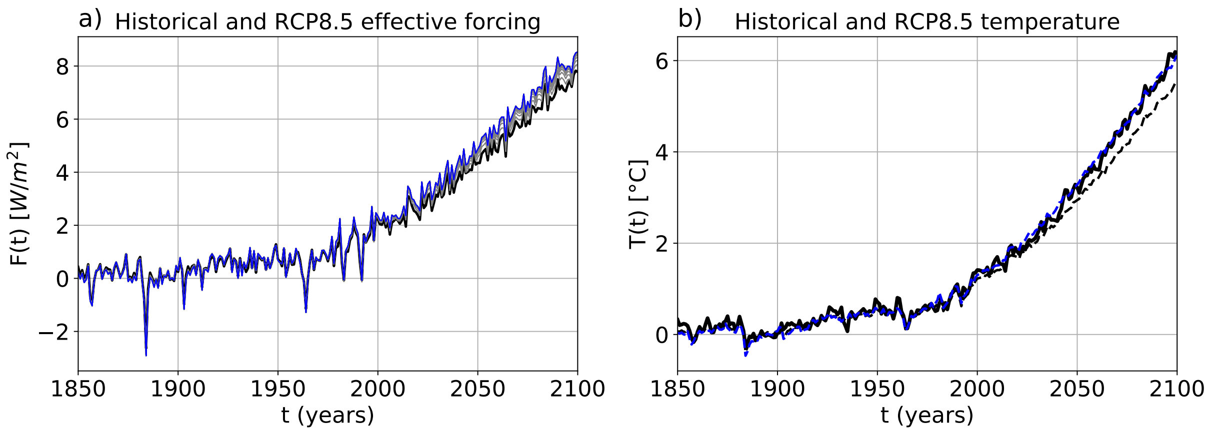

Figure S73. As Figure 4, but for the model IPSL-CM5A-LR.

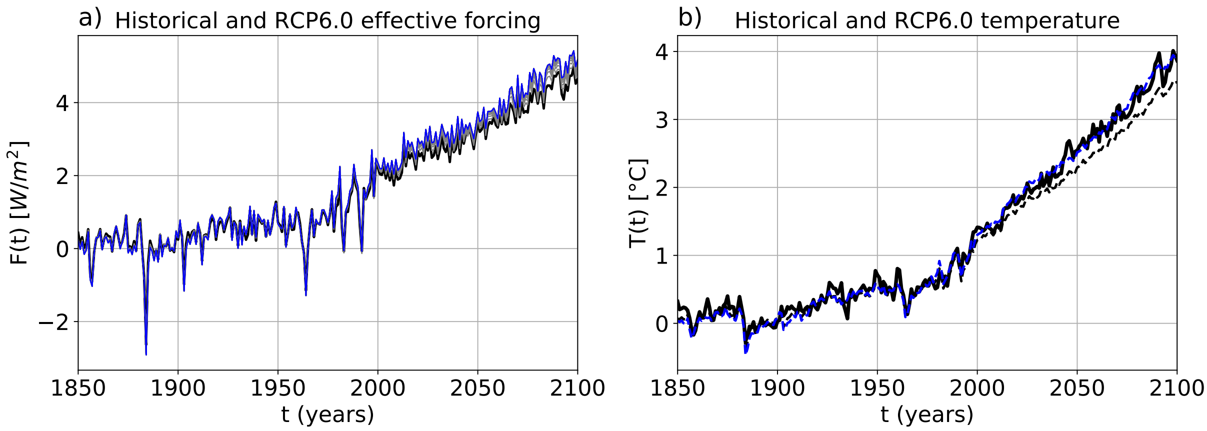

Figure S74. As Figure 4, but for the model IPSL-CM5A-LR and experiment RCP6.0.

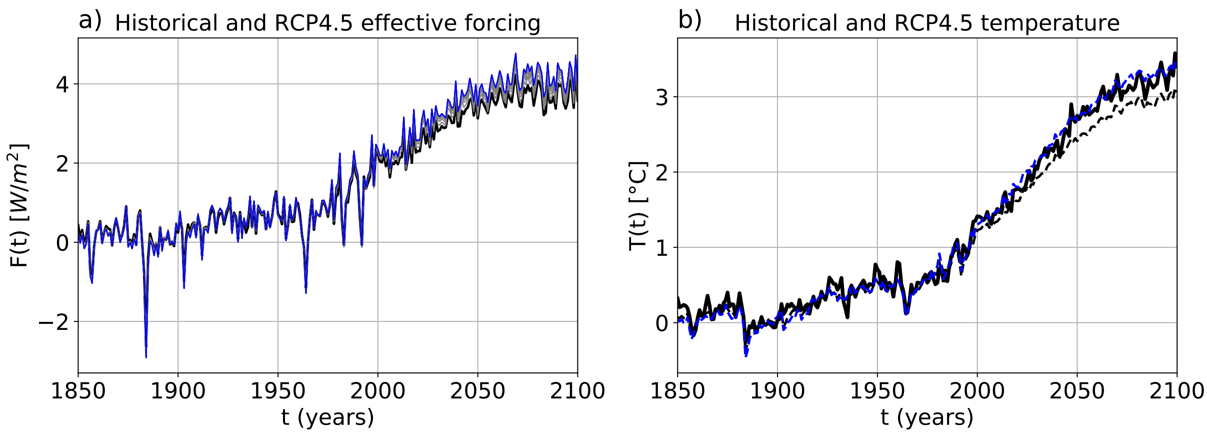

Figure S75. As Figure 4, but for the model IPSL-CM5A-LR and experiment RCP4.5.

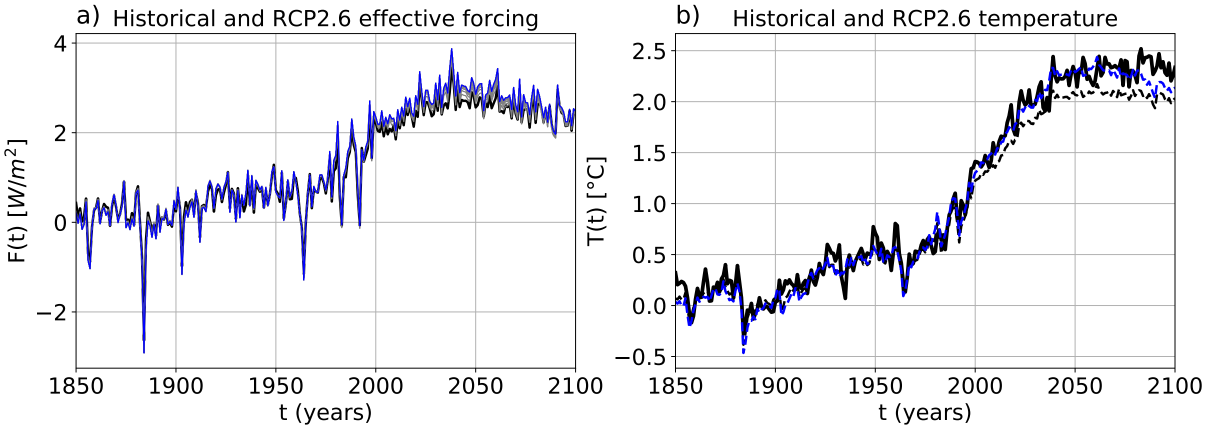

Figure S76. As Figure 4, but for the model IPSL-CM5A-LR and experiment RCP2.6.

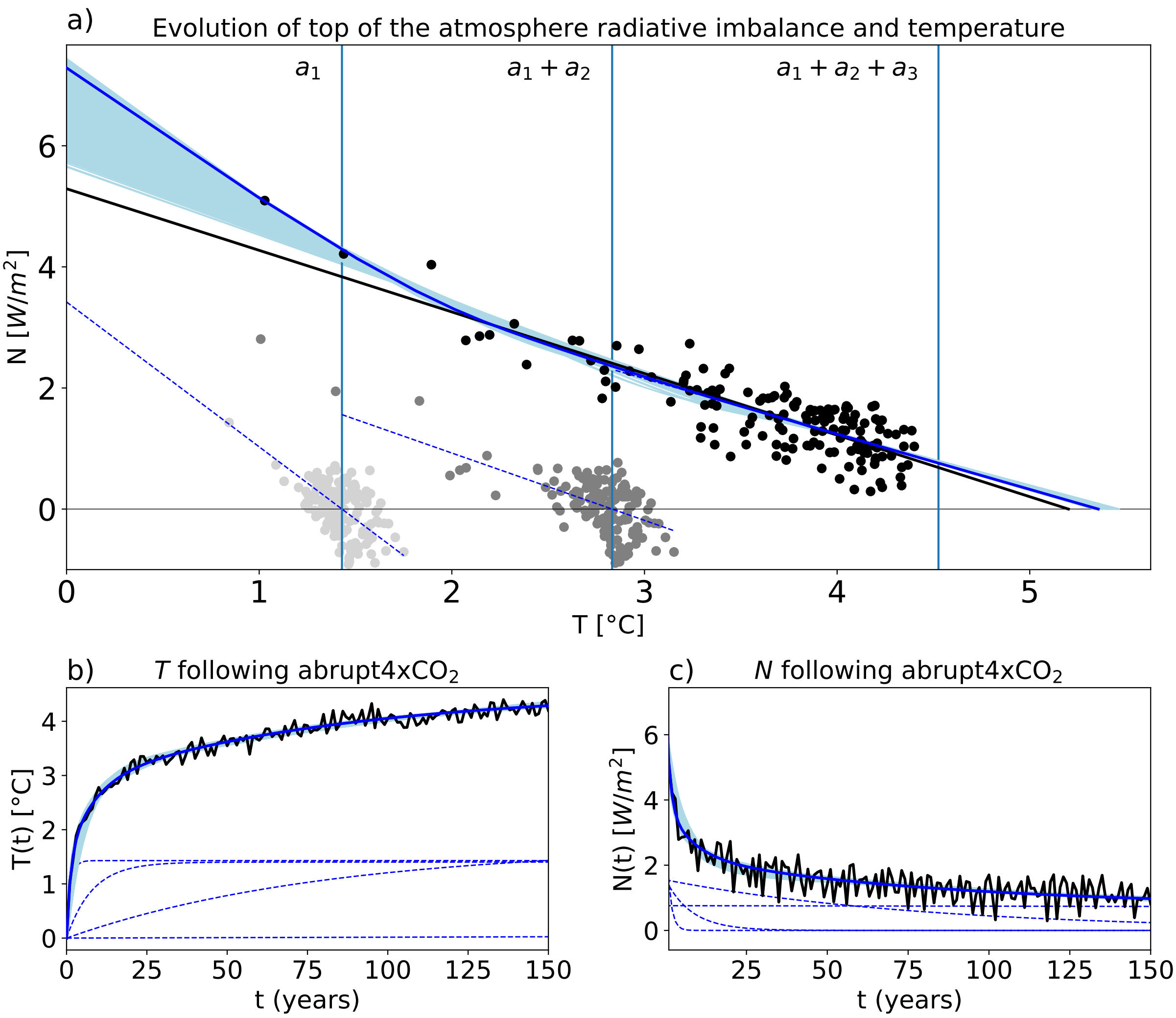

Figure S77. As Figure 1, but for the model IPSL-CM5B-LR.

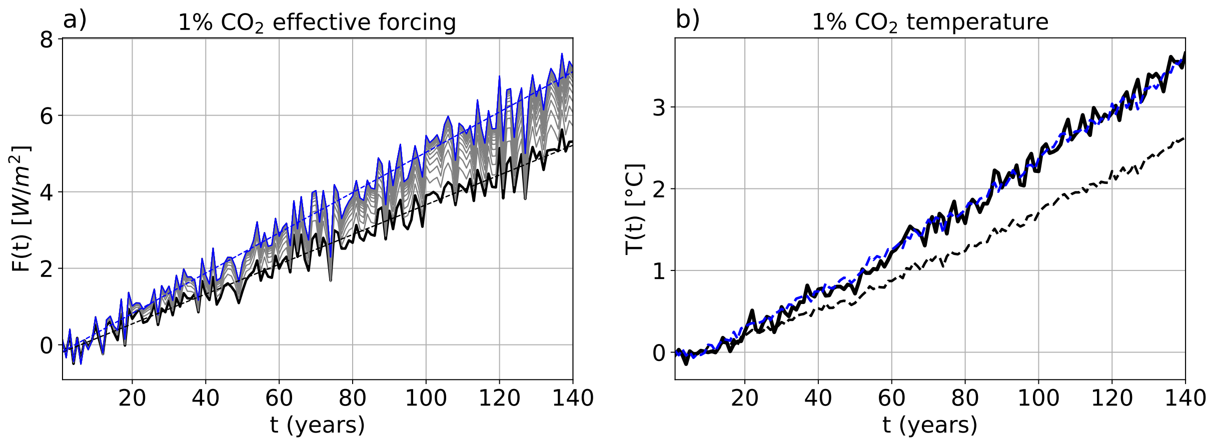

Figure S78. As Figure 3, but for the model IPSL-CM5B-LR.

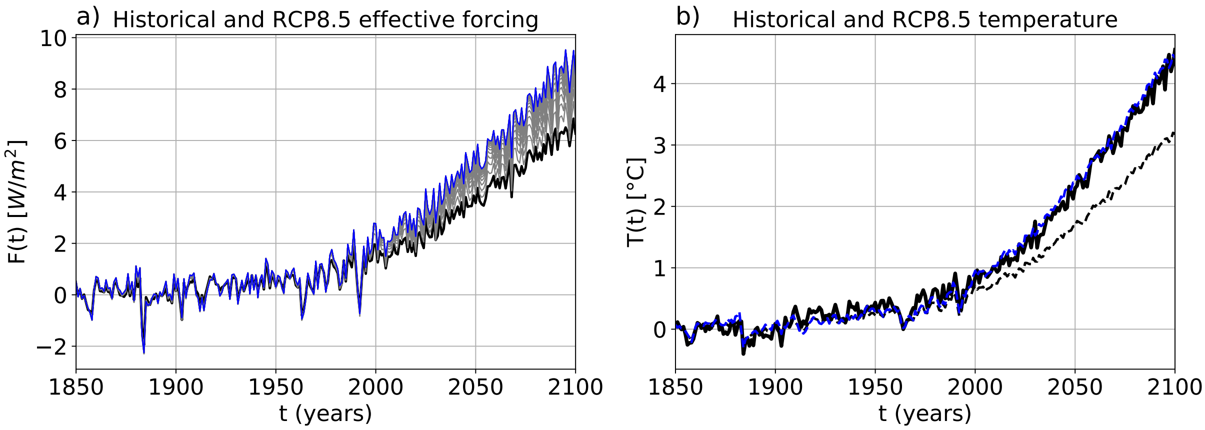

Figure S79. As Figure 4, but for the model IPSL-CM5B-LR.

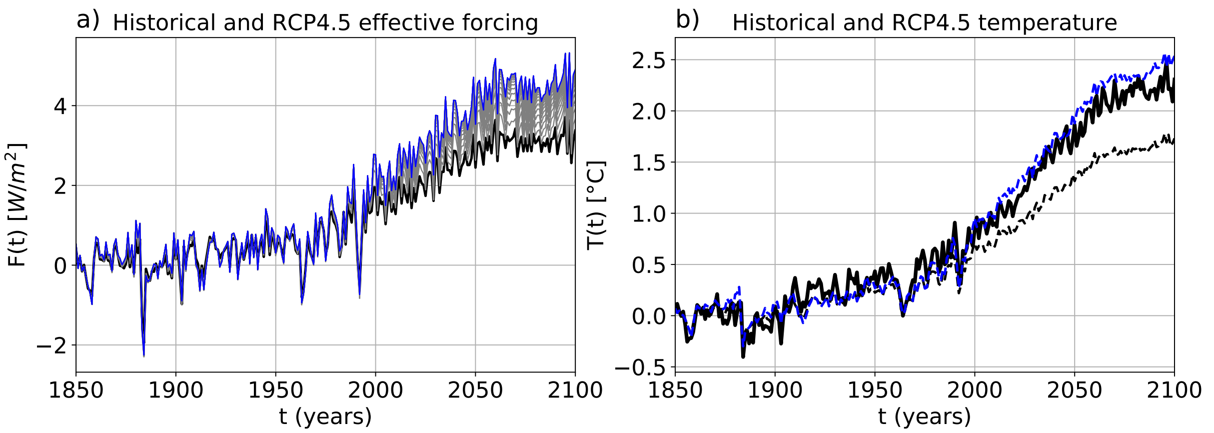

Figure S80. As Figure 4, but for the model IPSL-CM5B-LR and experiment RCP4.5.

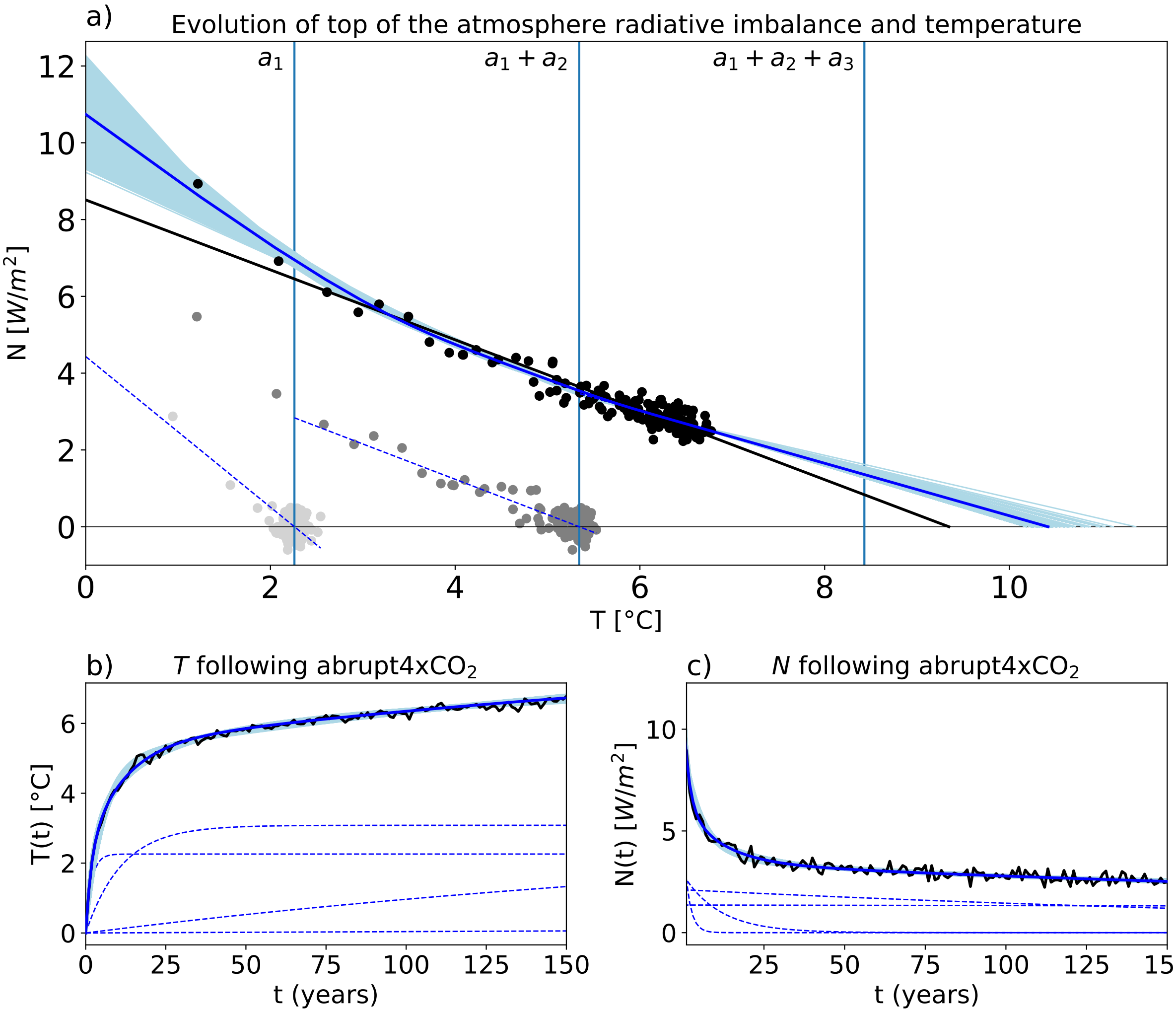

Figure S81. As Figure 1, but for the model MIROC-ESM.

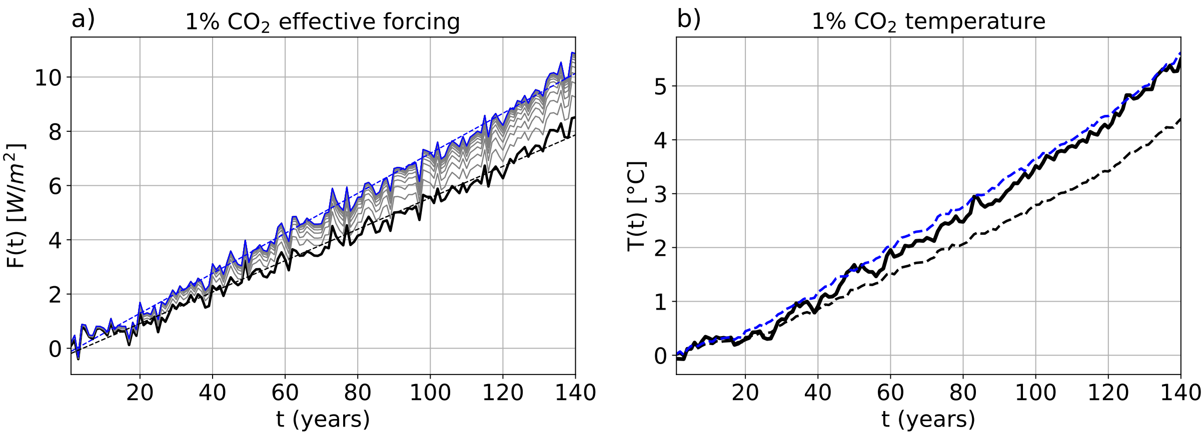

Figure S82. As Figure 3, but for the model MIROC-ESM.

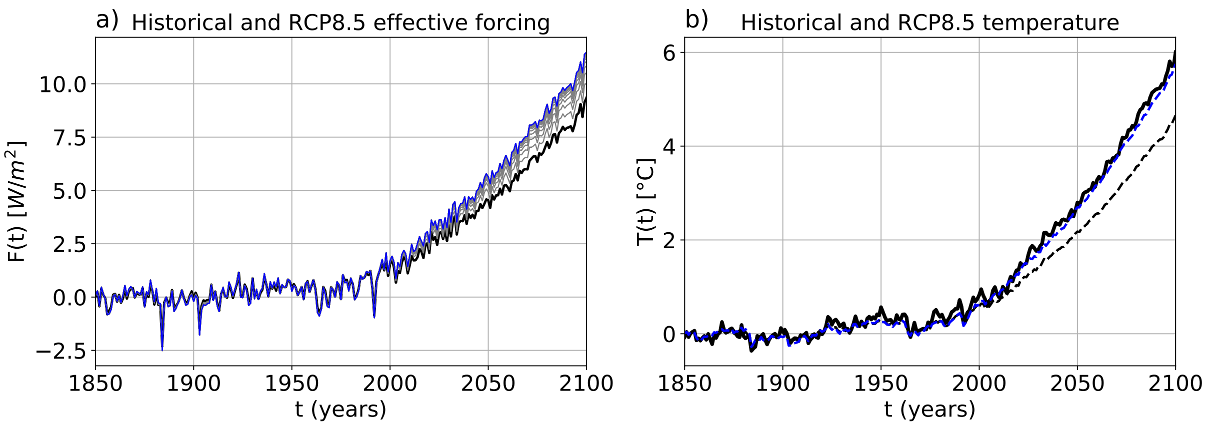

Figure S83. As Figure 4, but for the model MIROC-ESM.

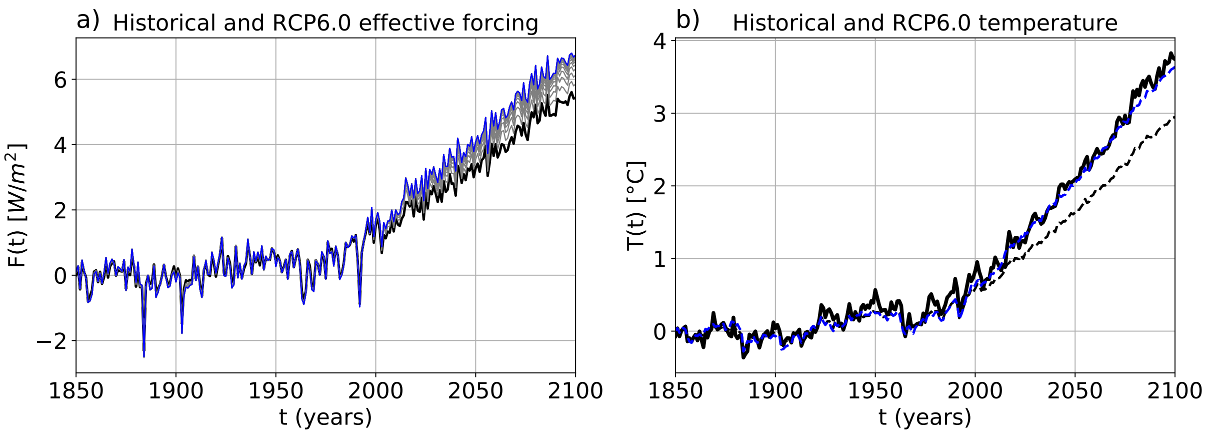

Figure S84. As Figure 4, but for the model MIROC-ESM and experiment RCP6.0.

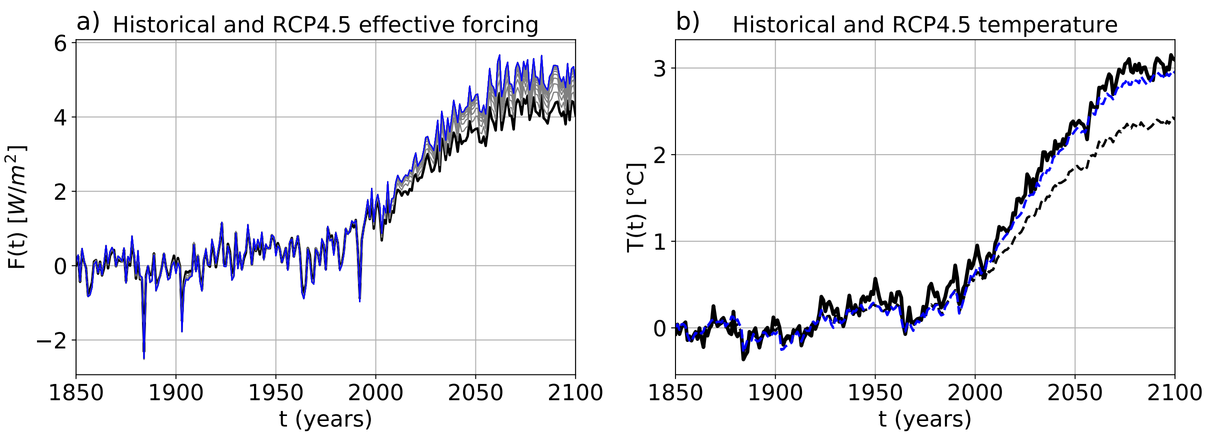

Figure S85. As Figure 4, but for the model MIROC-ESM and experiment RCP4.5.

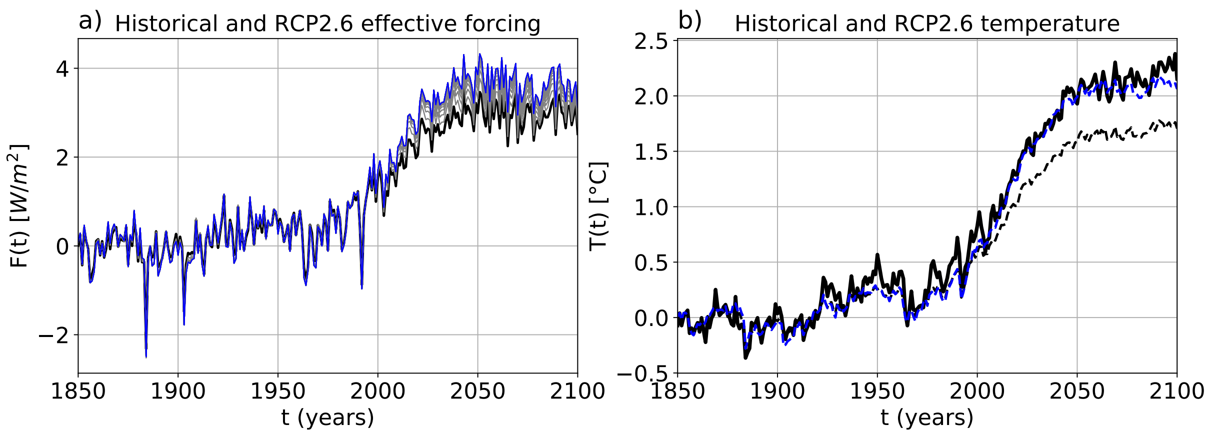

Figure S86. As Figure 4, but for the model MIROC-ESM and experiment RCP2.6.

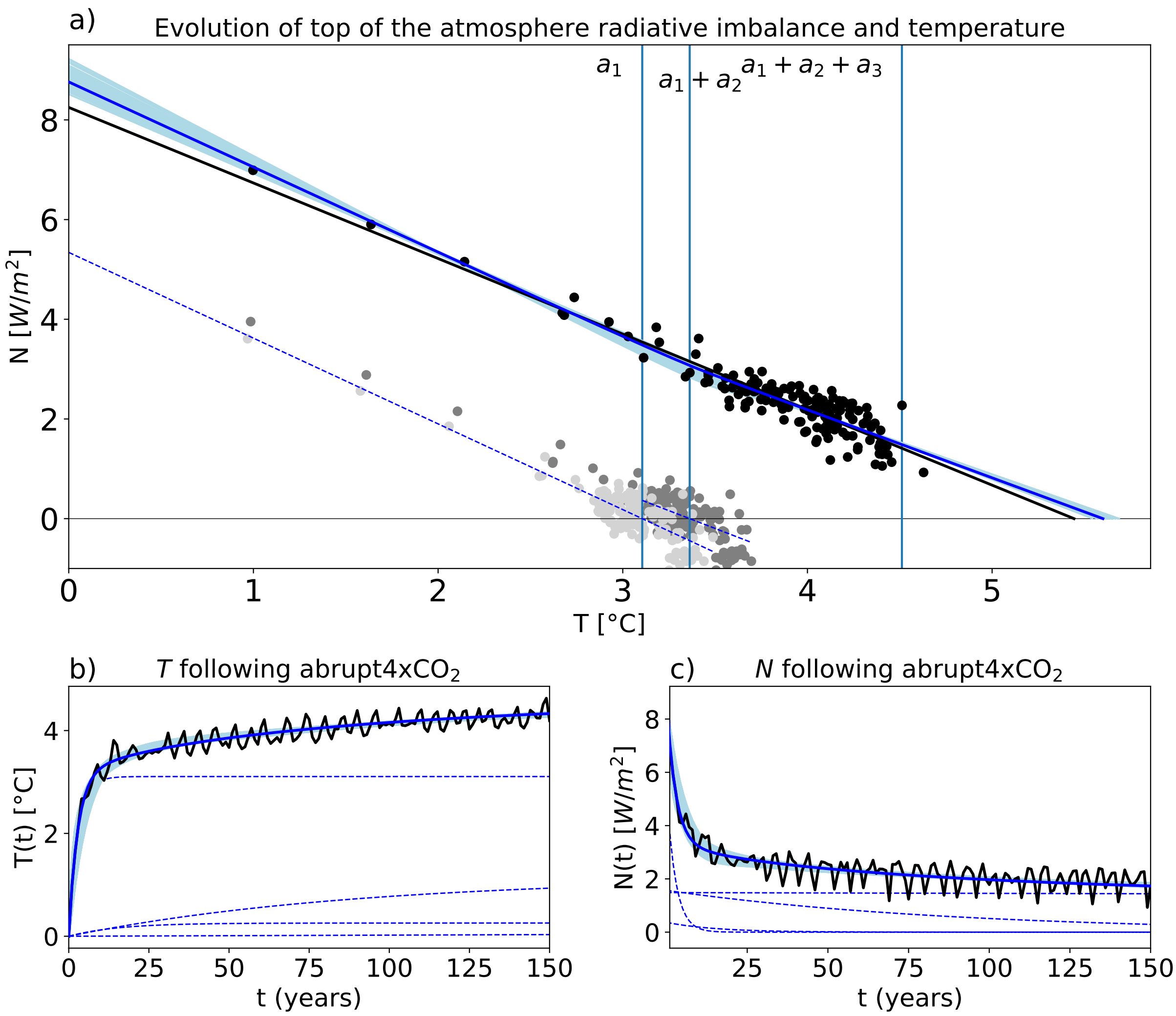

Figure S87. As Figure 1, but for the model MIROC5.

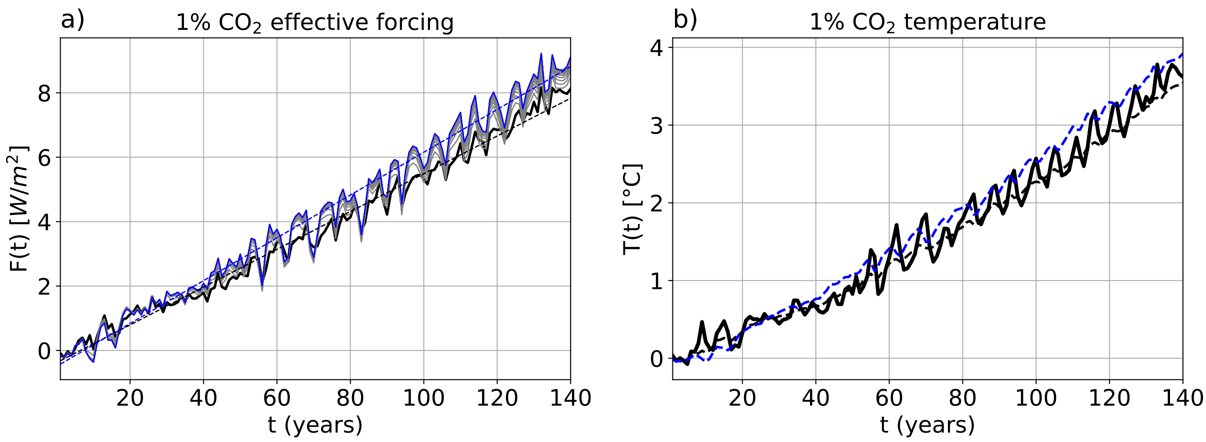

Figure S88. As Figure 3, but for the model MIROC5.

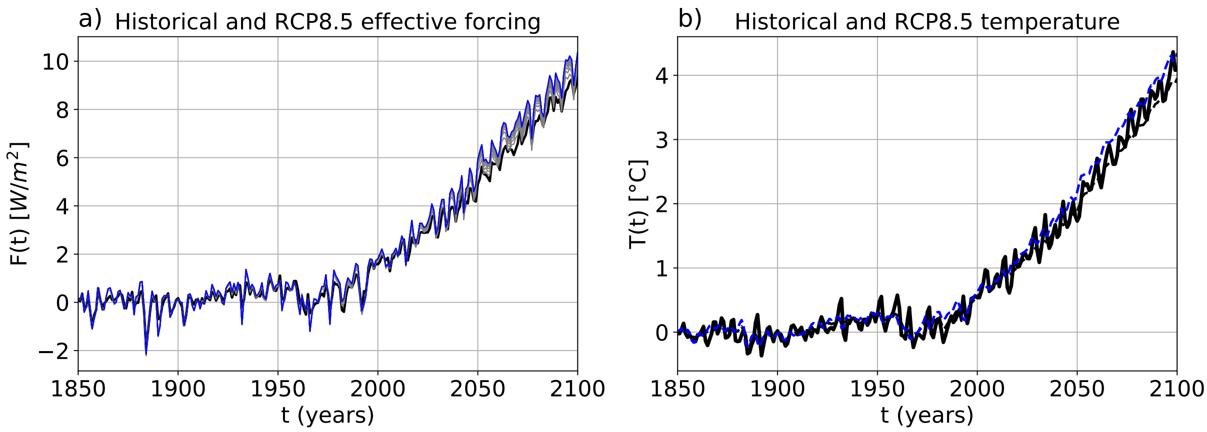

Figure S89. As Figure 4, but for the model MIROC5.

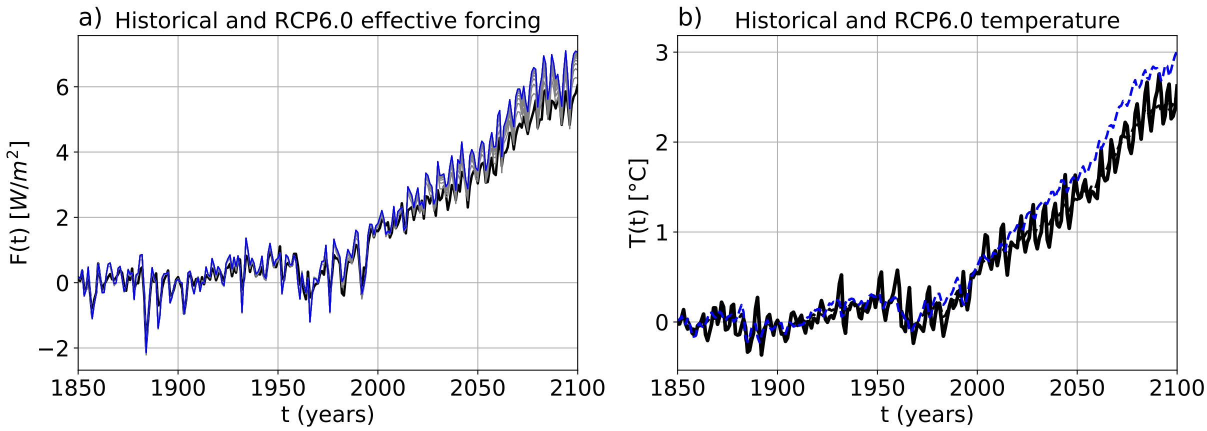

Figure S90. As Figure 4, but for the model MIROC5 and experiment RCP6.0.

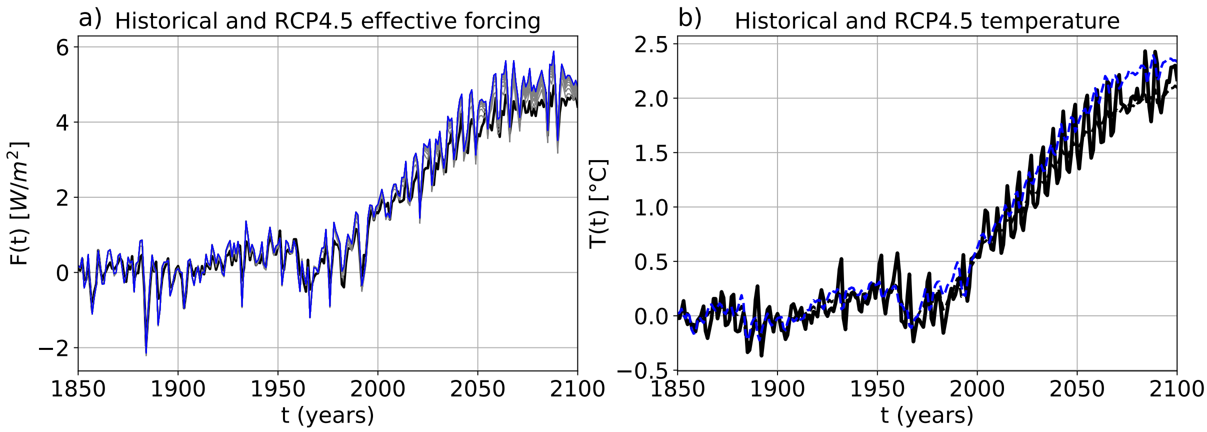

Figure S91. As Figure 4, but for the model MIROC5 and experiment RCP4.5.

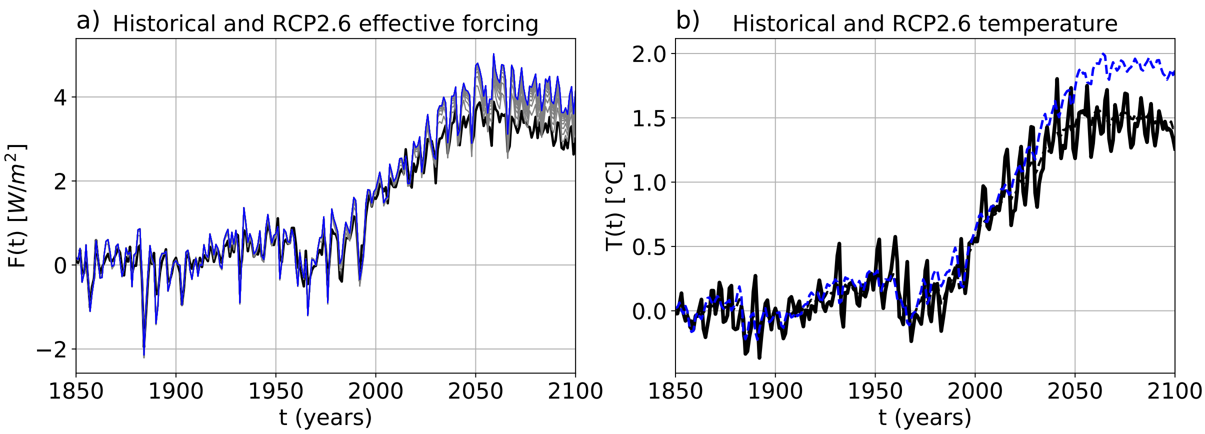

Figure S92. As Figure 4, but for the model MIROC5 and experiment RCP2.6.

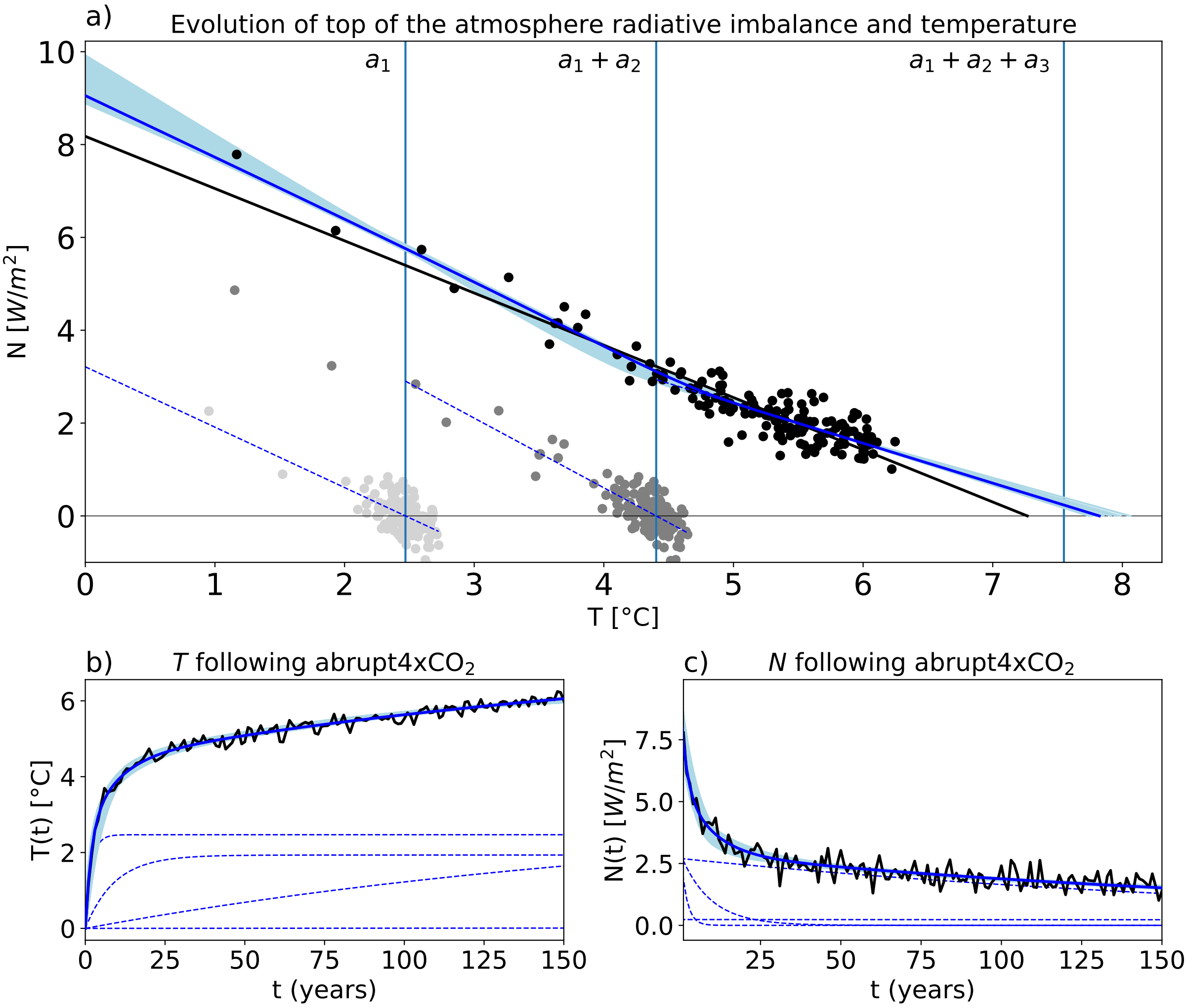

Figure S93. As Figure 1, but for the model MPI-ESM-LR.

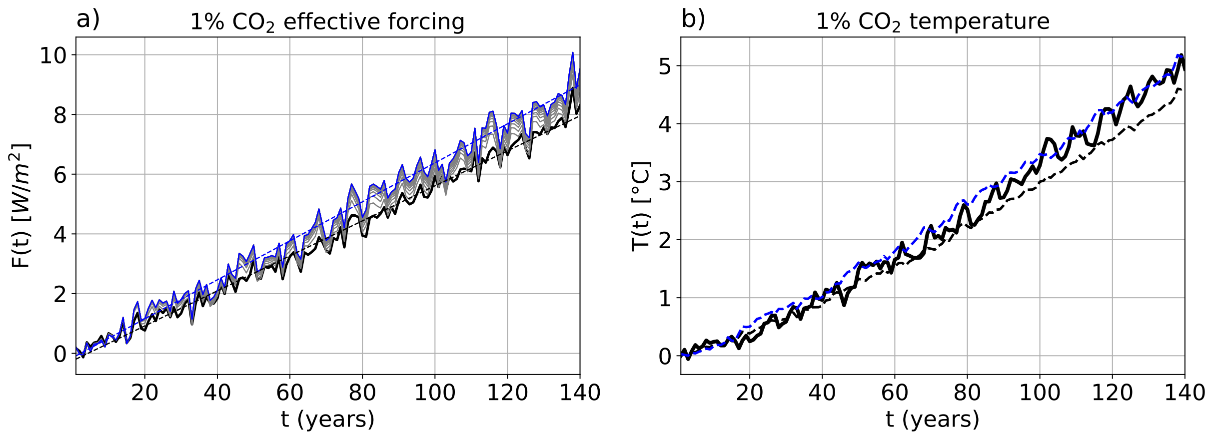

Figure S94. As Figure 3, but for the model MPI-ESM-LR.

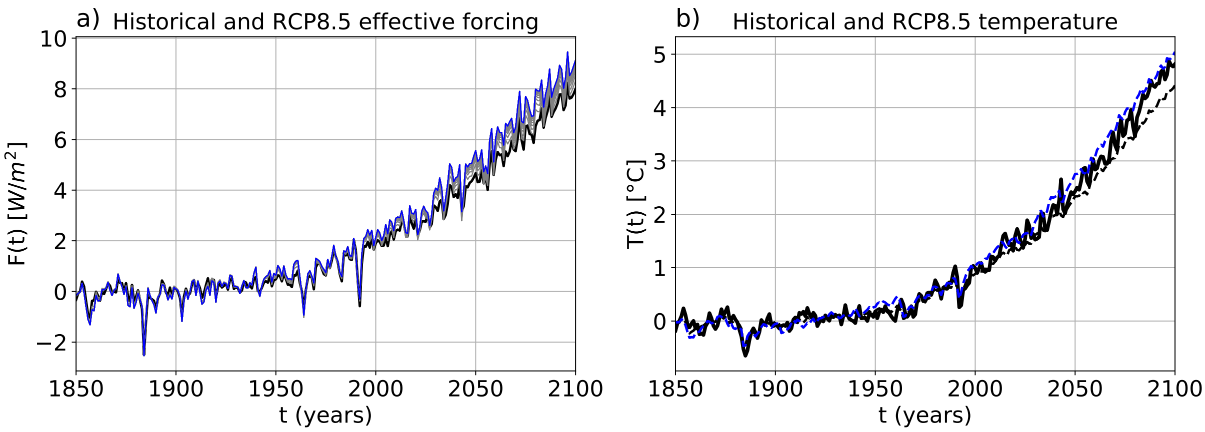

Figure S95. As Figure 4, but for the model MPI-ESM-LR.

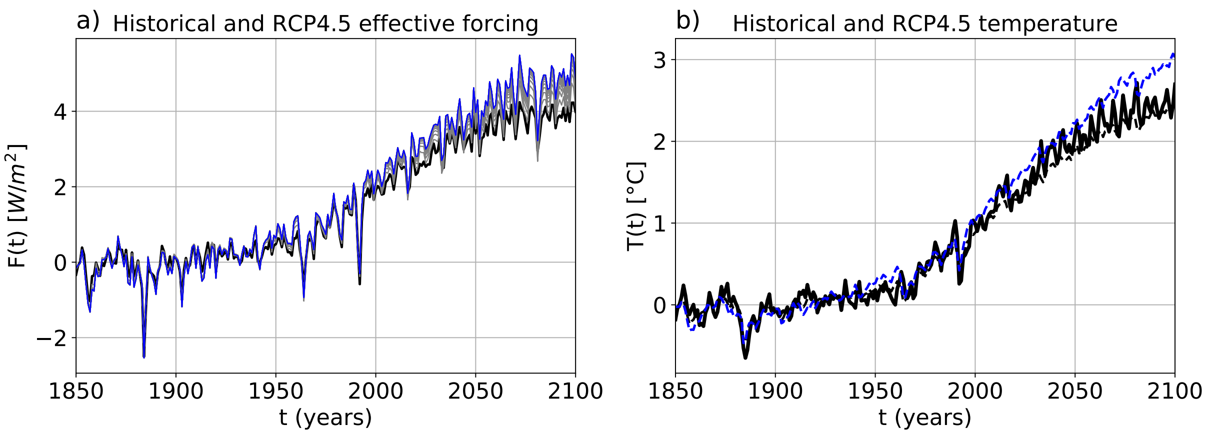

Figure S96. As Figure 4, but for the model MPI-ESM-LR and experiment RCP4.5.

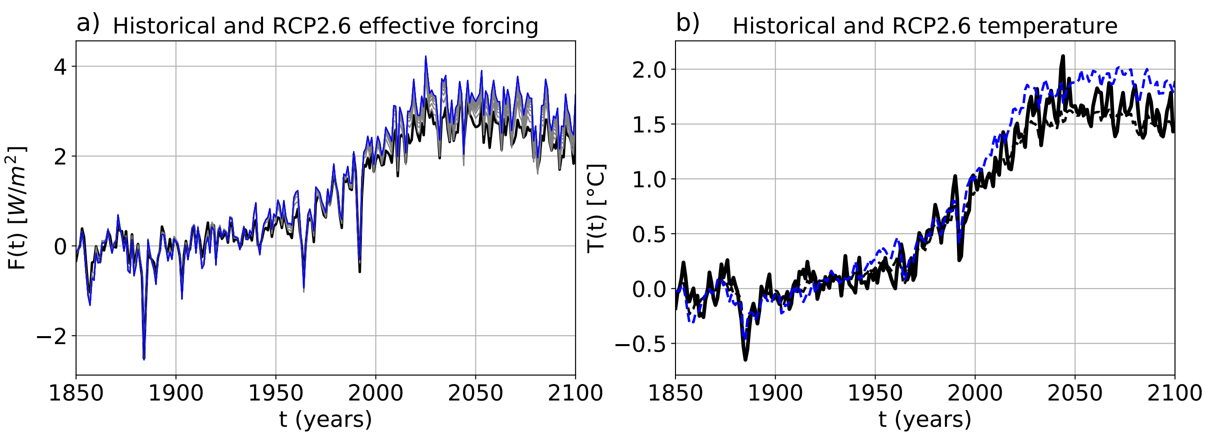

Figure S97. As Figure 4, but for the model MPI-ESM-LR and experiment RCP2.6.

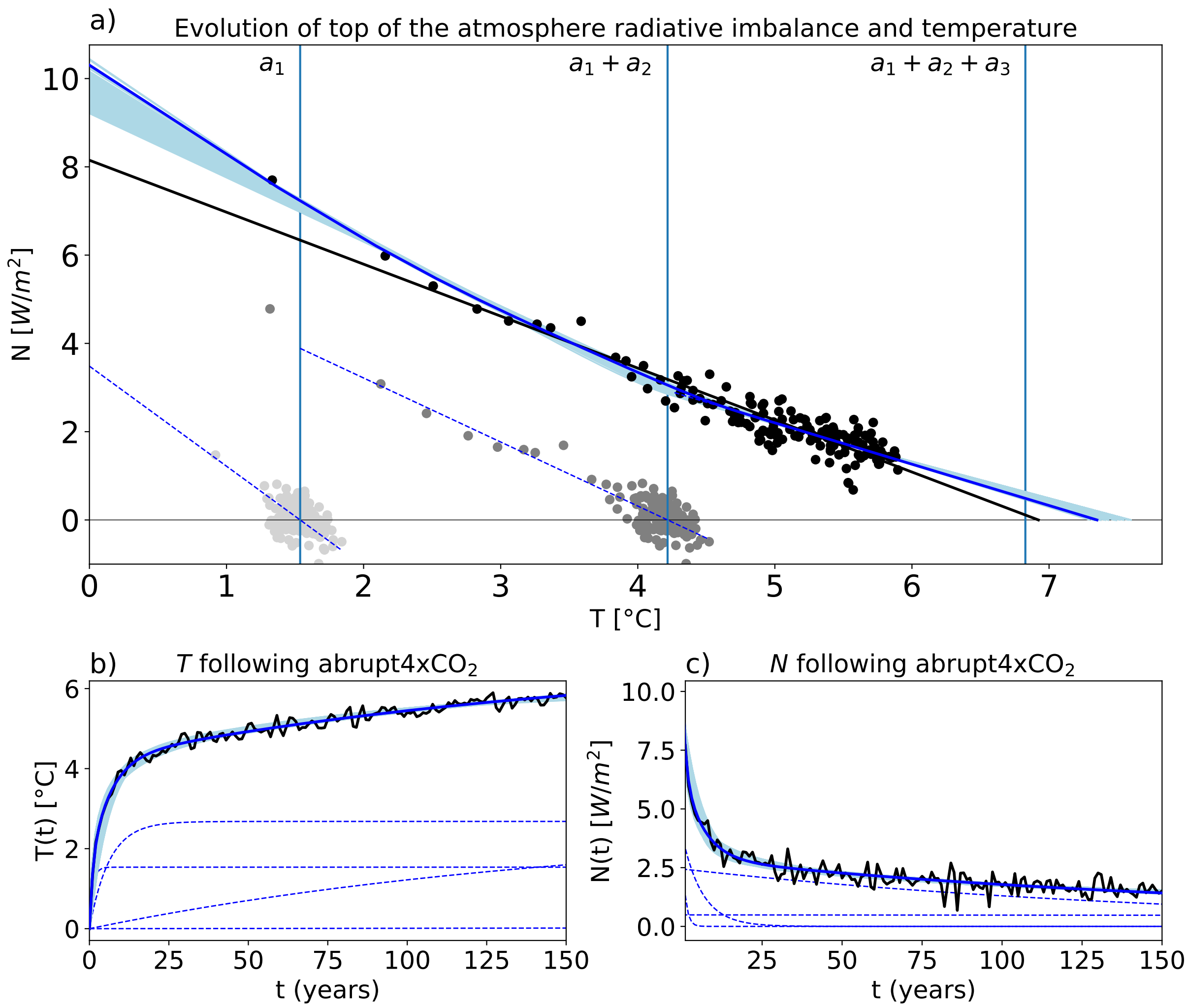

Figure S98. As Figure 1, but for the model MPI-ESM-MR.

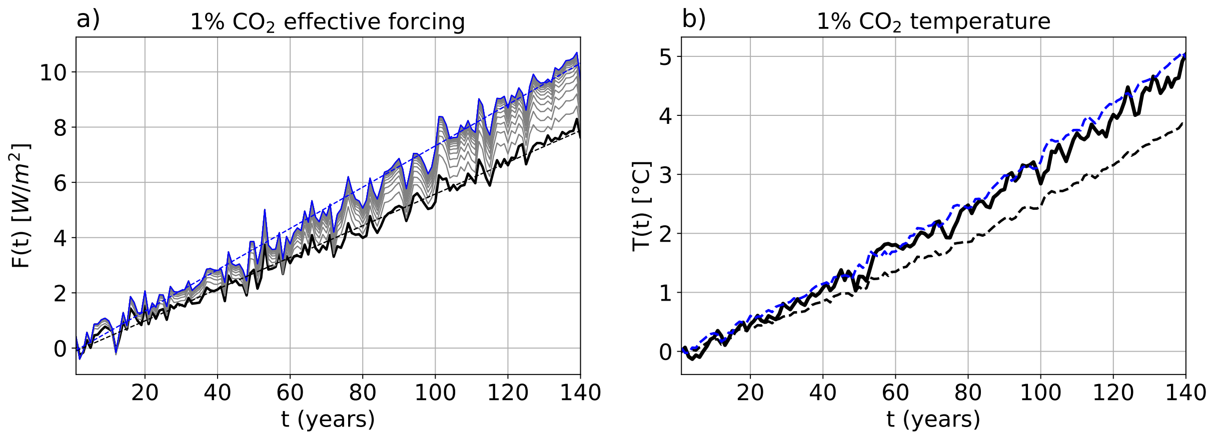

Figure S99. As Figure 3, but for the model MPI-ESM-MR.

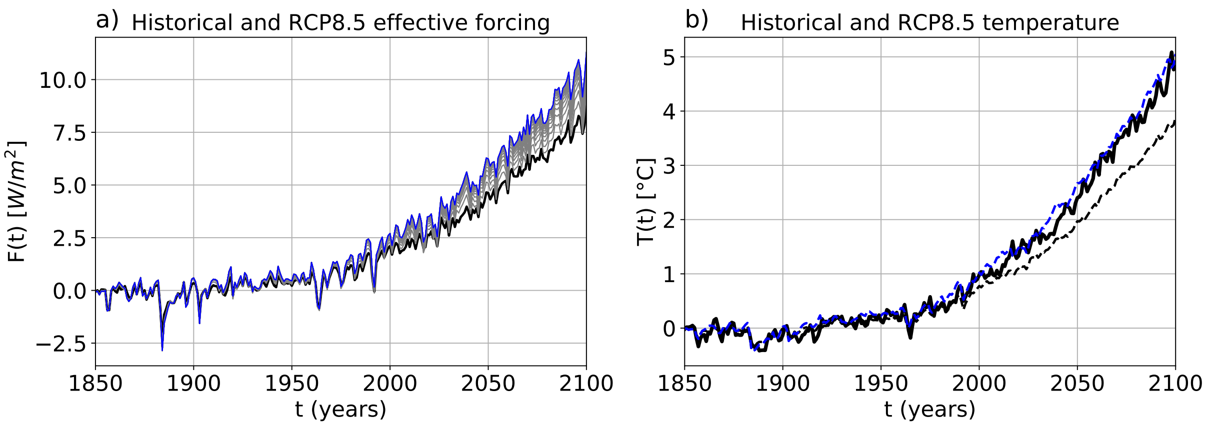

Figure S100. As Figure 4, but for the model MPI-ESM-MR.

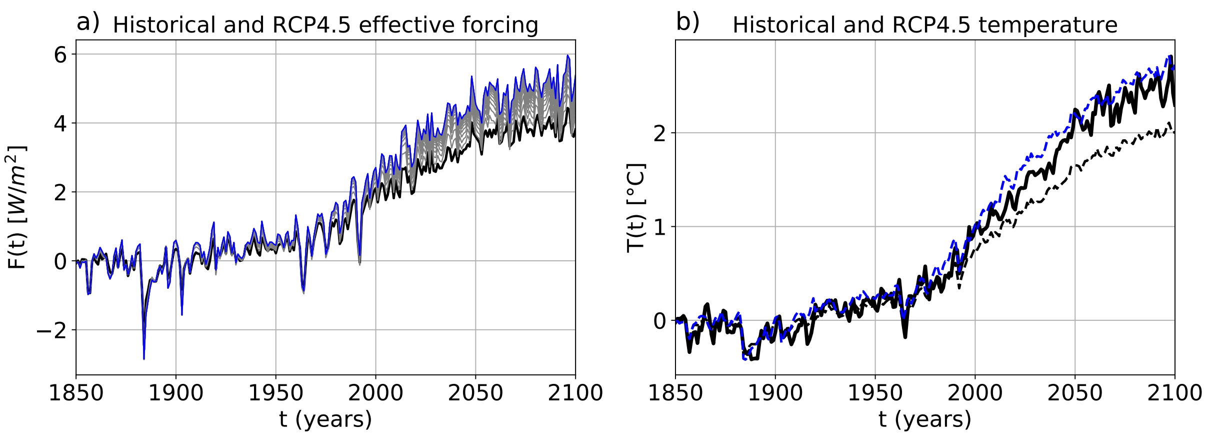

Figure S101. As Figure 4, but for the model MPI-ESM-MR and experiment RCP4.5.

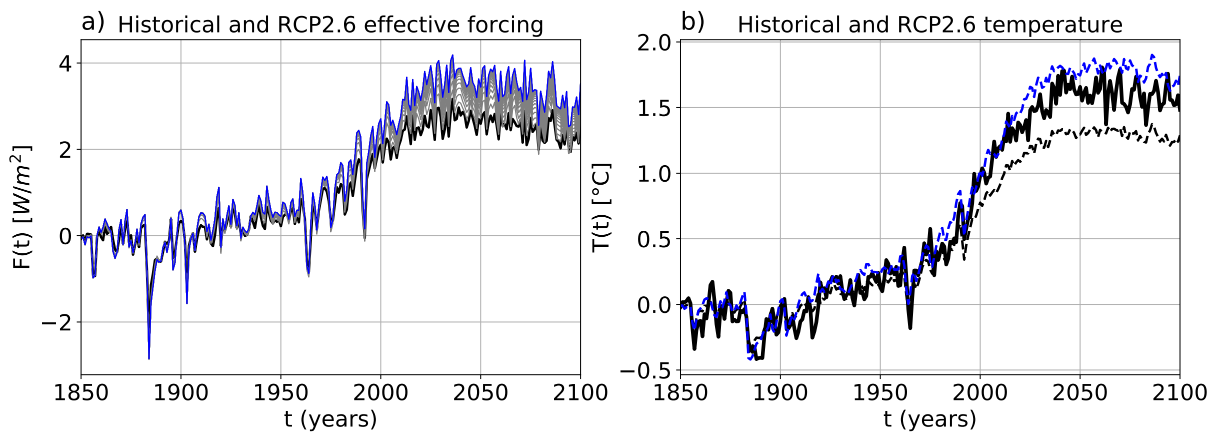

Figure S102. As Figure 4, but for the model MPI-ESM-MR and experiment RCP2.6.

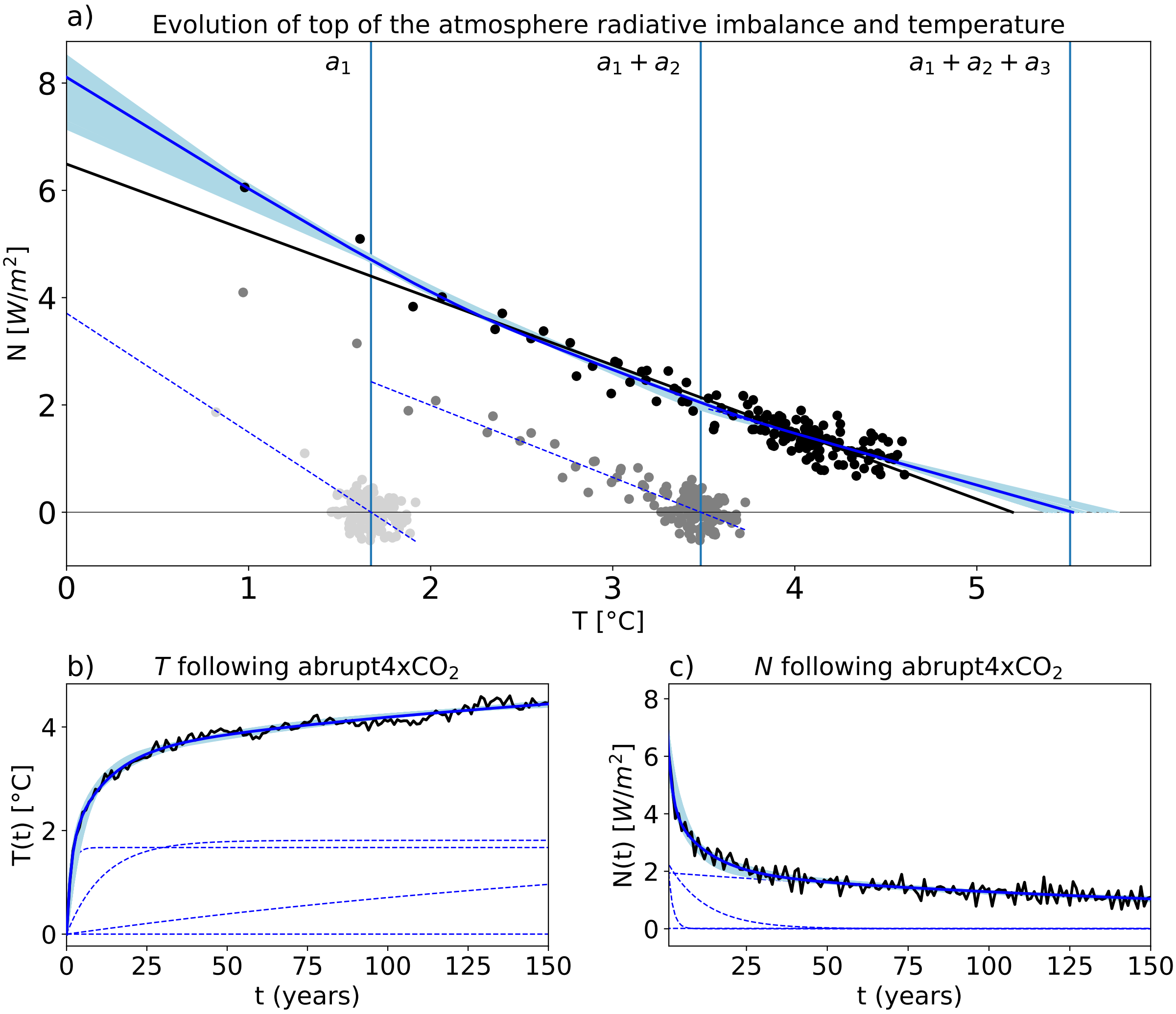

Figure S103. As Figure 1, but for the model MRI-CGCM3.

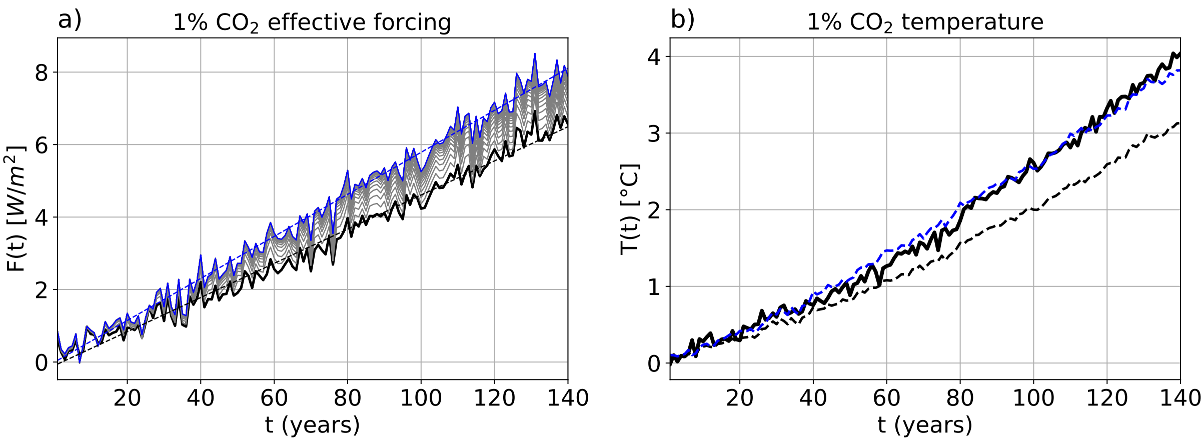

Figure S104. As Figure 3, but for the model MRI-CGCM3.

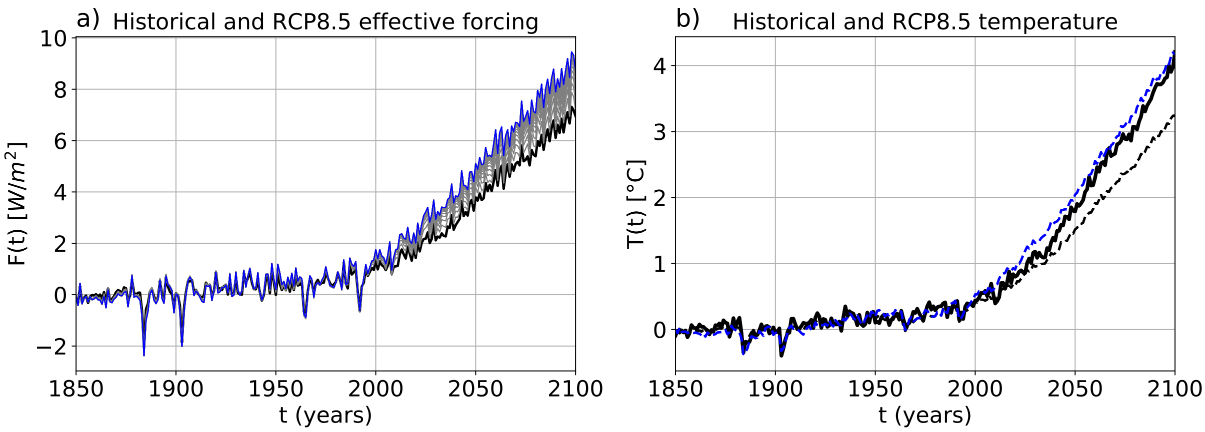

Figure S105. As Figure 4, but for the model MRI-CGCM3.

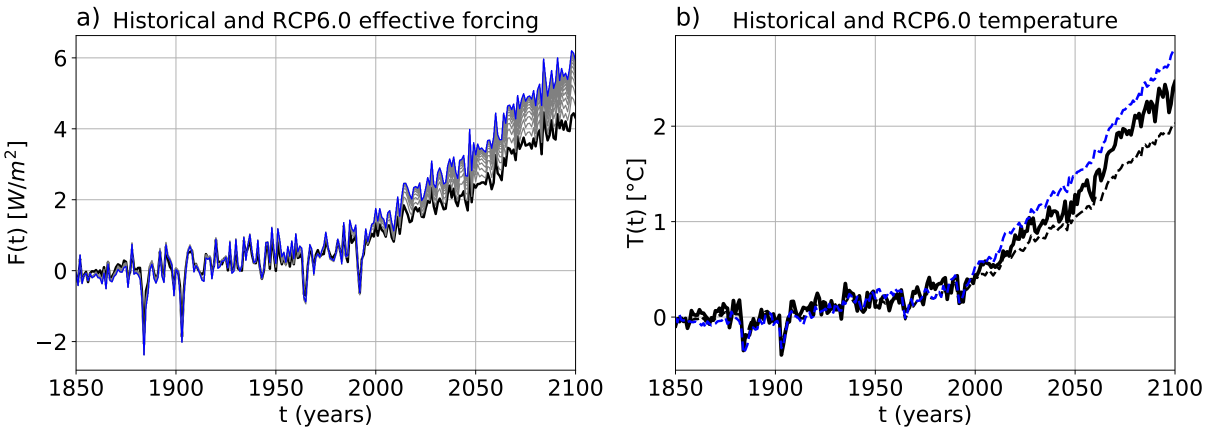

Figure S106. As Figure 4, but for the model MRI-CGCM3 and experiment RCP6.0.

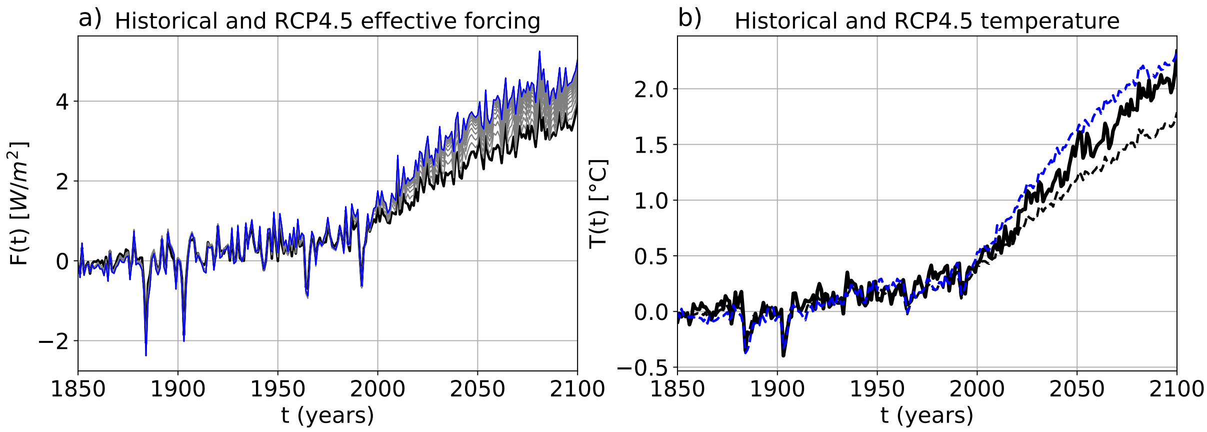

Figure S107. As Figure 4, but for the model MRI-CGCM3 and experiment RCP4.5.

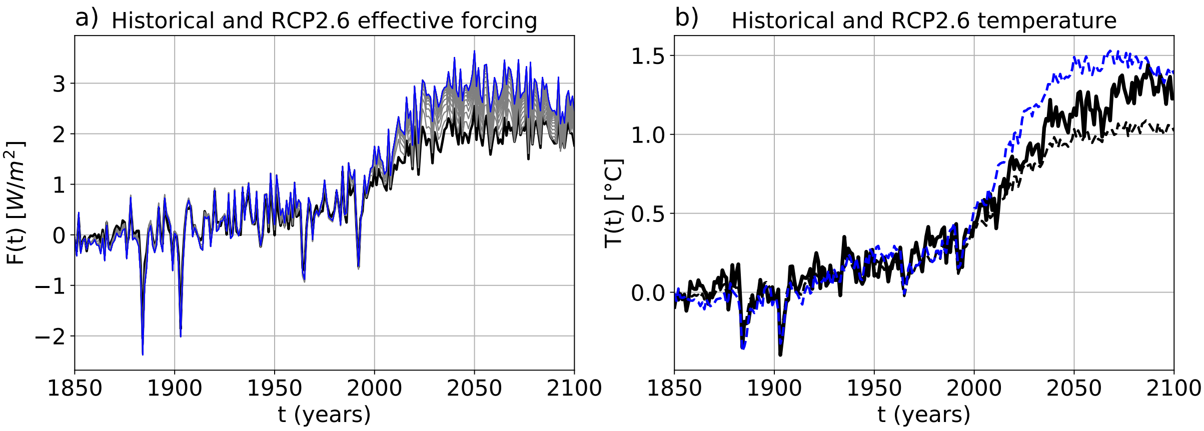

Figure S108. As Figure 4, but for the model MRI-CGCM3 and experiment RCP2.6.

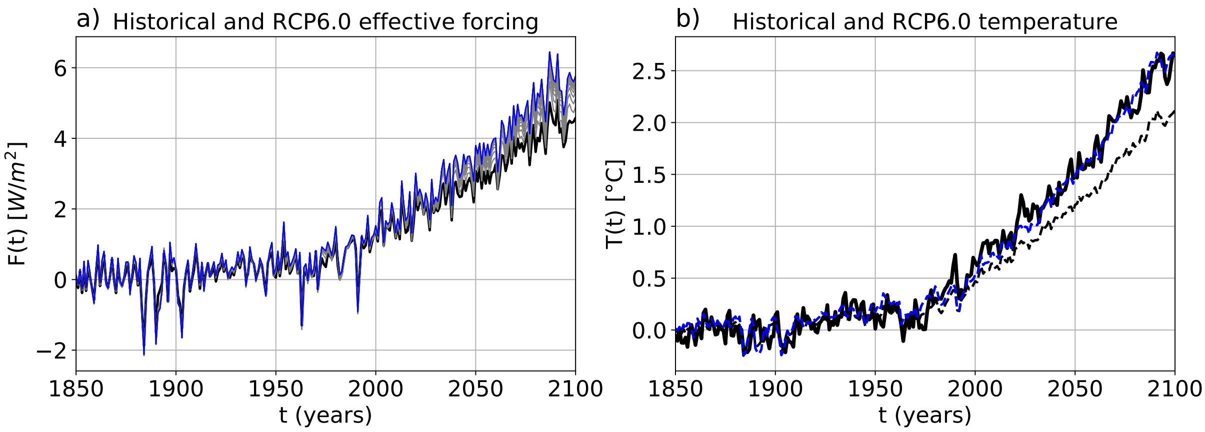

Figure S109. As Figure 4 with the model NorESM1-M, but for the experiment RCP6.0.

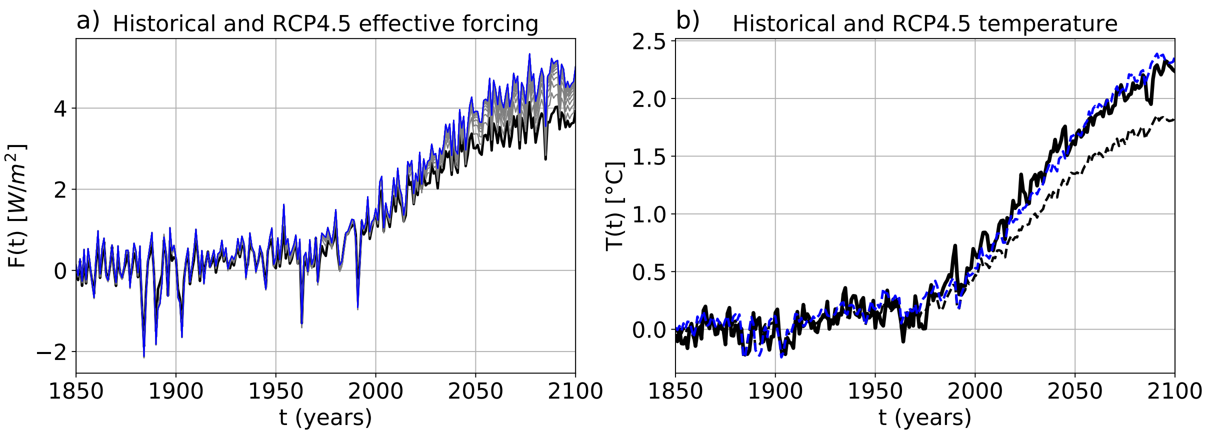

Figure S110. As Figure 4 with the model NorESM1-M, but for the experiment RCP4.5.

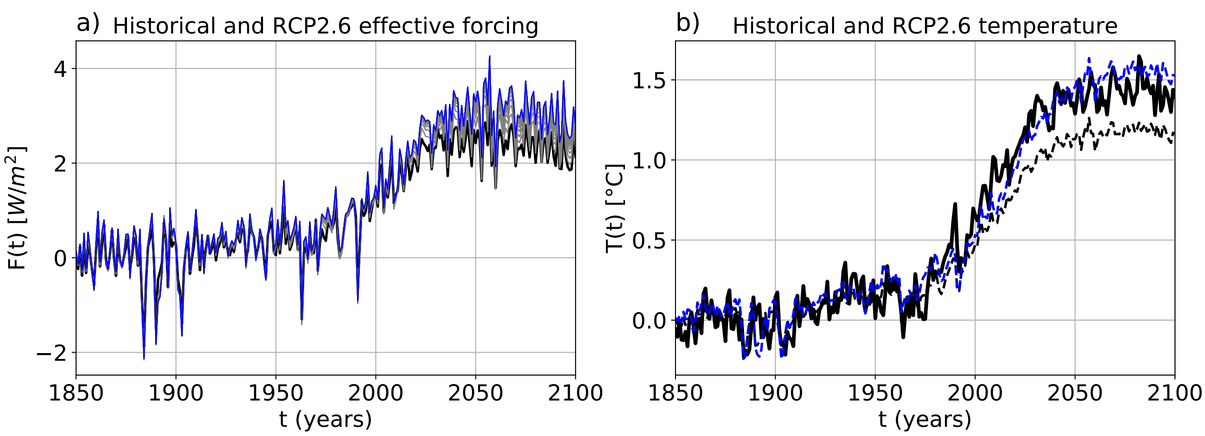

Figure S111. As Figure 4 with the model NorESM1-M, but for the experiment RCP2.6.

Table S1. The piControl trends computed over the 150 year period after branching of the abrupt4xCO_2 experiment. Temperature trends have units °C per year, and the top of atmosphere radiation components have units W/m^2 per year.

	ΔT/year	Δrlut/year	Δrsdt/year	Δrsut/year
ACCESS1-0	1.05e-03	1.52e-03	1.66e-06	-1.31e-03
ACCESS1-3	4.76e-04	1.14e-03	1.79e-07	-6.73e-04
CanESM2	1.80e-04	2.02e-04	9.70e-17	-2.18e-04
CCSM4	-5.03e-04	-4.86e-04	-1.28e-15	7.58e-04
CNRM-CM5	1.16e-03	1.61e-03	-1.09e-06	-9.32e-04
CSIRO-Mk3-6-0	6.71e-04	8.82e-04	-3.76e-09	-1.09e-03
GFDL-CM3	8.09e-04	1.51e-03	-1.29e-15	-9.08e-04
GFDL-ESM2G	-1.04e-03	-1.86e-03	0.00e+00	1.93e-03
GFDL-ESM2M	-1.11e-04	-3.93e-04	0.00e+00	-3.21e-04
GISS-E2-H	8.76e-04	9.77e-04	-1.02e-15	-6.38e-04
GISS-E2-R	5.33e-04	7.95e-04	-3.75e-16	-6.00e-04
HadGEM2-ES	-2.84e-04	1.06e-04	0.00e+00	1.67e-04
inmcm4	-7.63e-04	-1.18e-03	1.48e-06	1.03e-03
IPSL-CM5A-LR	-3.86e-04	-3.81e-04	1.02e-10	1.01e-03
IPSL-CM5B-LR	1.36e-03	2.84e-03	-2.39e-10	-1.62e-03
MIROC-ESM	6.67e-04	7.70e-04	-2.85e-06	-9.82e-05
MIROC5	-3.60e-04	-2.79e-04	4.66e-10	6.02e-04
MPI-ESM-LR	-5.64e-05	2.76e-04	-8.26e-07	-1.28e-04
MPI-ESM-MR	2.89e-05	8.32e-05	-1.40e-07	-1.91e-05
MRI-CGCM3	2.93e-04	1.07e-03	2.75e-06	-4.39e-04
NorESM1-M	-4.72e-04	-9.77e-04	1.14e-10	3.36e-04
min	-1.04e-03	-1.86e-03	-2.85e-06	-1.62e-03
max	1.36e-03	2.84e-03	2.75e-06	1.93e-03