Earth System
Dynamics

On determining the point of no return in climate change

Brenda C. van Zalinge, Qing Yi Feng, Matthias Aengenheyster, and Henk A. Dijkstra

Institute for Marine and Atmospheric Research Utrecht, Department of Physics, Utrecht University,
Utrecht, the Netherlands

Correspondence to: Qing Yi Feng (q.feng@uu.nl)

Abstract. Earth's global mean surface temperature has increased by about 1.0 °C over the period 1880–2015. One of the main causes is thought to be the increase in atmospheric greenhouse gases. If greenhouse gas emissions are not substantially decreased, several studies indicate that there will be a dangerous anthropogenic interference with climate by the end of this century. However, there is no good quantitative measure to determine when it is "too late" to start reducing greenhouse gas emissions in order to avoid such dangerous interference. In this study, we develop a method for determining a so-called "point of no return" for several greenhouse gas emission scenarios. The method is based on a combination of aspects of stochastic viability theory and linear response theory; the latter is used to estimate the probability density function of the global mean surface temperature. The innovative element in this approach is the applicability to high-dimensional climate models as demonstrated by the results obtained with the PlaSim model.

1 Introduction

In the year 2100, which is as far (or as close) in the future as 1932 is in the past, mankind will be living on an Earth with a different climate than today. At that time, we will know the 2100 global mean surface temperature (GMST) value and its increase, ΔT, above the pre-industrial GMST value. From the then available GMST records, it will also be known whether this change in GMST has been gradual or whether it was rather "bumpy". If the observational effort continues as of today, there will also be an adequate observational record to determine whether the probability of extreme events (e.g. flooding, heat waves) has increased.

The outcomes of these future observations to be made by future generations will strongly depend on the socio-economic and technological developments and political decisions made now and over the next decades. Fortunately, there is a set of tools available to inform decision makers: Earth system models. These models come in different flavours, from global climate models (GCMs) providing details on the development of the ocean–atmosphere–ice–land system to integrated assessment models (IAMs), which also aim to describe the development of the broader socio-economic sys-

tem. During the preparation for the fifth assessment report (AR5) of the Intergovernmental Panel on Climate Change (IPCC), GCM studies have focussed on the climate system response to GHG changes as derived by IAMs from different socio-economic scenarios; the data from these simulations are gathered in the so-called CMIP5 archive (http://cmip-pcmdi.llnl.gov/cmip5/).

Depending on the representation of fast climate feedbacks in GCMs that determine their climate sensitivity, the CMIP5 models project a GMST increase ΔT of 2.5–4.5 °C over the period 2000–2100 (Pachauri et al., 2014). This does not mean that the actual measured value of ΔT in 2100 will be in this interval. For example, the GMST may be well outside this range because of current model errors which misrepresent the strength of a specific feedback. As a consequence, a transition might have occurred in the real climate system which did not occur in any of the CMIP5 model simulations. Another possibility is that the GHG development was eventually far outside of the scenarios considered in CMIP5.

A crucial issue in 2100 will be whether a climate state has been reached in which a dangerous anthropogenic interference (DAI) can be identified (Mann, 2009). In this case, present-day islands will have been swallowed by the ocean

and extreme events will have increased in frequency and magnitude (Smith and Schneider, 2009). These effects are very inhomogeneously distributed over the Earth and lead to enormous socio-economic consequences. If this is the case in 2100, then there is a point in time at which we must have crossed the conditions for DAI. This time, marking the boundary of a "safe" and "unsafe" climate state, obviously depends on the metrics used to quantify the state of the complex climate system.

In very simplified views, this boundary is interpreted as a threshold CO_2 concentration (Hansen et al., 2008) or GMST. The latter, in particular the $\Delta T_c = 2\,°C$ threshold, has become an easy to communicate (and maybe therefore leading) idea to set mitigation targets for greenhouse gas reduction. Emission scenarios have been calculated (Rogelj et al., 2011) such that ΔT will remain below ΔT_c. Although thresholds for GMST have been criticized for being very inadequate regarding impacts (Victor and Kennel, 2014), such a threshold forms the basis for policymaking as set forward in the Paris 2015 (COP21) agreement.

Supposing that measures are being taken to keep $\Delta T < \Delta T_c$, does this mean that we are "safe"? The answer is a simple no, as DAI may still have occurred regionally, such as the disappearance of island chains due to sea level rise (Victor and Kennel, 2014). Hence, attempts have been made to define what "safe" means in a more general way, such as the tolerable windows approach (TWA; Petschel-Held et al., 1999) and viability theory (VT; Aubin, 2009). These approaches also deal with general control strategies to steer a system towards "safety" when needed.

On a more abstract level, both TWA and VT start by defining a desirable (or "safe") subspace V of a state vector x in a general state space X. This subspace is characterized by constraints, such as thresholds for properties of x. For example, when x is a high-dimensional state vector of a GCM, such a threshold could be $\Delta T < \Delta T_c$ for GMST. When the time development (or trajectory) of x is such that it moves outside the subspace V, a control is sought to steer the trajectory back into V. Note that this is an abstract formulation of the mitigation problem when the amplitude of the greenhouse gas emission is taken as a control. Recently, (Heitzig et al., 2016) have added more detail to regions in the space X which differ in their "safety" properties and the amount of flexibility in the control to steer to "safety".

Given a certain desirable subspace of the climate system state vector (e.g. to avoid DIA) and a suite of control options (e.g. CO_2 emission reduction), it is important to know when it is too late to steer the system to "safe" conditions, for example in the year 2100. In other words, when is the point of no return (PNR)? The TWA and VT approaches and the theory in (Heitzig et al., 2016) suffer from the "curse of dimensionality" and cannot be used within CMIP5 climate models. For example, the optimization problems in VT and TWA lead to dynamic programming schemes which have up to now only been solved for model systems with low-dimensional

state vectors. The approach in (Heitzig et al., 2016) requires the computation of regional boundaries in state space, which also becomes tedious in more than two dimensions. Hence, with these approaches it will be impossible to determine a PNR using reasonably detailed models of the climate system.

In this paper, we present an approach similar to TWA and VT, but one which can be applied to high-dimensional models of the climate system. The key to the approach is the estimation of the probability density function of the properties of the state vector x which determine the "safe" subspace V. The PNR problem is coupled to limitations in the control options (e.g. of emissions) and can be defined precisely using these options and stochastic viability theory. The methodology is presented in Sect. 2; to illustrate the concepts, we apply the approach in Sect. 3 to an idealized energy balance model with and without tipping behaviour. In Sect. 4, the application to a high-dimensional climate model is presented using data from the Planet Simulator (PlaSim; Fraedrich et al., 2005). A summary and discussion in Sect. 5 concludes the paper.

2　Methodology

Here we briefly describe the concepts we need from stochastic viability theory and then define the PNR problem, specifically in the climate change context.

2.1　Viable states

Viability theory studies the control of the evolution of dynamical systems to stay within certain constraints on the system state vector (Aubin, 2009). Here we consider finite dimensional deterministic systems with state vector $x \in R^d$ and vector field $f : R^d \to R^d$ given by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x, t). \tag{1}$$

In the general formulation of viability theory, a time-dependent input is also considered on the right-hand side of Eq. (1), which can be used to control the path of the trajectory $x(t)$ in state space.

For our purposes, we only need the concept of a viable state, which is related to constraints on the state vector defining a viable region V in state space, also called the viability constraint set. In the model (Eq. 1) such a set can, for example, be defined by a threshold condition $||x|| < ||x_*||$. An initial condition $x_0 = x(t = 0)$ is called viable if $x(t) \in V$ for all $0 \le t \le t^*$, where t^* is a certain end time. The set of all these initial conditions forms the viability kernel associated with V.

Stochastic extensions of viability theory consider finite dynamical systems defined by stochastic differential equations:

$$d\mathbf{X}_t = \boldsymbol{f}(\mathbf{X}_t, t)dt + \boldsymbol{g}(\mathbf{X}_t, t)d\mathbf{W}_t, \qquad (2)$$

where $\mathbf{X}_t \in R^d$ is a multidimensional stochastic process, $\mathbf{W}_t \in R^n$ is a vector of n-independent standard Wiener processes, and the matrix $\boldsymbol{g} \in R^{d \times n}$ describes the dependence of the noise on the state vector. The normalized probability density function (PDF) $p(\boldsymbol{x}, t)$ can be formally determined from the Fokker–Planck equation associated with Eq. (2).

A stochastic viability kernel V_β consists of initial conditions $\mathbf{X}_0$ for which the system has, for $0 \le t \le t^*$, a probability larger than a value β to stay in the viable region V (Doyen and De Lara, 2010). For example, in a one-dimensional version of Eq. (2) with state vector $X_t \in R$ and a viable region V given by $x \le x^*$, a state X_t is called viable with tolerance probability β_T if

$$\int_{-\infty}^{x^*} p(x, t)dx \ge \beta_T; \qquad (3)$$

otherwise, X_t is called non-viable.

2.2 Linear response theory

In relatively idealized low-dimensional models (such as the energy balance model in Sect. 3), the probability density functions can be easily computed by solving for the Fokker–Planck equation (see Sect. 3.2). However, in order to find the temporal evolution of the PDF of the global mean surface temperature (GMST) under any CO_2-equivalent forcing in high-dimensional climate models, such as PlaSim in Sect. 4, we will use linear response theory (LRT). With this theory, the effect of any small forcing perturbation on the system state can be calculated by running the climate model for only one forcing scenario (Ragone et al., 2016).

In LRT, the expectation value of an observable Φ, when forcing the system with a time-dependent function $f(t)$, can be calculated by computing the convolution of a Green's function $G_{\langle\Phi\rangle}$ and the forcing $f(t)$ according to

$$\langle\Phi\rangle_f(t) = \int_{-\infty}^{+\infty} G_{\langle\Phi\rangle}(\tau) f(t-\tau)d\tau. \qquad (4)$$

To construct this Green's function, the property that the convolution in the time domain is the same as pointwise multiplication in the frequency domain is used. The Fourier transform of Eq. (4) is given by

$$\langle\widetilde{\Phi}\rangle_f(\omega) = \chi_{\langle\Phi\rangle}(\omega)\widetilde{f}(\omega), \qquad (5)$$

with $\chi_{\langle\Phi\rangle}(\omega)$, $\langle\widetilde{\Phi}\rangle_f(\omega)$, and $\widetilde{f}(\omega)$ as the Fourier transforms of $G_{\langle\Phi\rangle}(t)$, $\langle\Phi\rangle_f(t)$, and $f(t)$, respectively. Therefore, once

the time evolution of the expectation value of an observable under a certain forcing is known, the Green's function of this observable can be constructed with Eq. (5), and consequently the linear response of the observable to any forcing can be calculated.

2.3 The point of no return problem

In the climate change context, scenarios of GHG increase and the associated radiative forcing have been formulated as representative concentration pathways (RCPs). In (Pachauri et al., 2014), there are four RCP scenarios (Fig. 1a) ranging from an increase in radiative forcing of $2.6\,\mathrm{Wm}^{-2}$ (RCP2.6) in 2100 (with respect to 2000) to a forcing increase of $8.5\,\mathrm{Wm}^{-2}$ (RCP8.5).

To define the PNR for each of these RCPs, a collection of mitigation scenarios on greenhouse gas emissions has to be considered. These mitigation scenarios will lead to changes in GHG concentrations described by functions $F_\lambda(t)$, where λ is a parameter. For instance, the collection F_λ could result from mitigation measures that lead to an exponential decay to different stabilization levels (measured in CO_2 equivalent, or CO_2 eq.) within a certain time interval. An example of such a collection F_λ is shown by the dashed and dotted red lines in Fig. 1b. The most extreme member of F_λ is defined as the mitigation scenario (represented by a certain value of parameter λ) which has the steepest initial decrease at a certain time t (dashed curve in Fig. 1b).

Along the curve of a certain RCP scenario, there will be a point in time at which action will be taken to reduce emissions of GHG; this is indicated by a time of action t_b. Consider, for example (Fig. 1b), that t_b is chosen as the first year in which the state vector $\mathbf{X}_t$ is no longer viable. A reduction in emissions is, however, not immediately followed by a decrease in CO_2 eq. due to the long residence time of atmospheric CO_2. In addition, there is a delay to take action in emission reduction due to technological, social, economic, and institutional challenges. Hence, emission reduction will only start Δt_1 years after t_b. The CO_2 eq. will, even after emissions have been reduced, also still increase over a time Δt_2. The time at which the CO_2 eq. starts to reduce according to F_λ is indicated by $t_c = t_b + \Delta t$, where $\Delta t = \Delta t_1 + \Delta t_2$. Eventually, $\mathbf{X}_t$ may become viable again and this point in time is indicated by t_e (Fig. 1b).

For a given RCP scenario, tolerance probability β_T, viable region V, and collection F_λ, we define the PNR (π_t) as the first year t_c in which, even when at that moment the most extreme CO_2-equivalent reduction scenario F_λ applies,

a. either $\mathbf{X}_t$ will be non-viable for more than τ_T years, where τ_T is a set tolerance time, or

b. $\mathbf{X}_t$ will be non-viable in the year 2100.

The first PNR, which we will indicate below by π_t^{tol}, is based on limiting the number of years that $\mathbf{X}_t$ is non-viable since
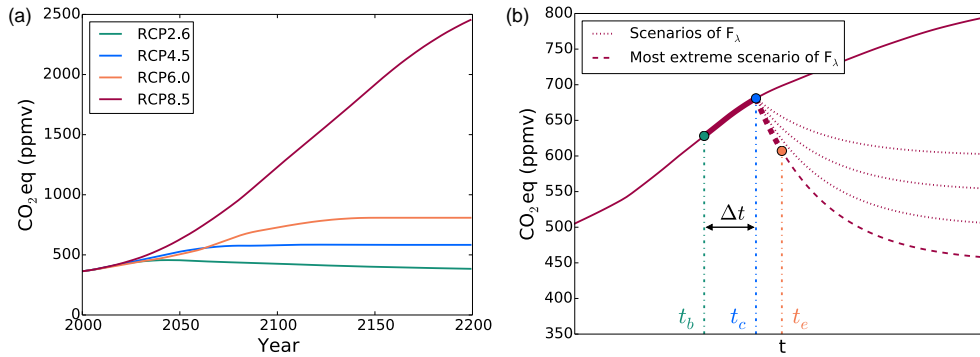
Figure 1. (a) The CO_2-equivalent trajectories of the RCP scenarios used by the IPCC in CMIP5. **(b)** The solid red curve represents a typical RCP scenario. At the time t_b, the climate state becomes non-viable, while at $t = t_c$ a CO_2-equivalent reduction F_λ applies; at time t_e, the climate state is viable again.

(during these years) society is exposed to risks from, for example, extreme weather events. The second PNR, which we will indicate below by π_t^{2100}, imposes no restrictions on how long $\mathbf{X}_t$ is non-viable, but it is only based on $\mathbf{X}_t$ being non-viable at the end of this century. Hence, under the given set of mitigation options, it is guaranteed that the state will have left the viable region by the year 2100. We will use both PNR concepts in the results below.

3 Energy balance model

In this section, we illustrate the concepts and the computation of the PNR for an idealized energy balance model of the Budyko–Sellers type (Budyko, 1969; Sellers, 1969). We will also assume that the CO_2 eq. can be directly controlled, and hence no carbon cycle model is needed to determine CO_2 eq. from an emission reduction scenario.

3.1 Formulation

We use the stochastic extension of the model formulation as in Hogg (2008). The equation for the atmospheric temperature T_t (in K) is given by

$$dT_t = \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6)$$
$$\frac{1}{c_T}\left\{Q_0(1-\alpha(T_t)) + G + A\ln\frac{C(t)}{C_{\text{ref}}} - \sigma\epsilon T_t^4\right\}dt + \sigma_s dW_t.$$

The values and meanings of the parameters in Eq. (6) are given in Table 1. The first term on the right-hand side of Eq. (6) represents the short-wave radiation received by the surface and $\alpha(T)$ is the albedo function given by

$$\alpha(T) = \alpha_0 H(T_0 - T) + \alpha_1 H(T - T_1) \qquad\qquad (7)$$
$$+ \left(\alpha_0 + (\alpha_1 - \alpha_0)\frac{T - T_0}{T_1 - T_0}\right)H(T - T_0)H(T_1 - T).$$

This equation contains the effect of land ice on the albedo, and $H(x) = 1/2(1 + \tanh(x/\epsilon_H))$ is a continuous approximation of the Heaviside function. When the temperature

$T < T_0$, the albedo will be α_0, and when $T > T_1$ it will be α_1 and the albedo is linear in T for $T \in [T_0, T_1]$.

The second term on the right-hand side of Eq. (6) represents the effect of greenhouse gases on the temperature. It consists of a constant part (G) and a part ($A\ln\frac{C(t)}{C_0}$) depending on the mean CO_2-equivalent concentration in the atmosphere (indicated by $C(t)$). The third term on the right-hand side of Eq. (6) expresses the effect of long-wave radiation on the temperature, and the last term represents noise with a constant standard deviation σ_s. The standard value of σ_s, chosen as 3 % of the value of G/c_T, is hence about 0.3 K year^{-1}. The variance in CO_2 concentration originates mostly from seasonal variations, and 3 % is on the high side. Nevertheless, we still use this value because if we take values smaller than 3 % the PDF of the GMST will almost be a delta function, and concepts cannot be illustrated clearly.

3.2 Results: stochastic viability kernels

When using the global mean CO_2-equivalent concentration C in Eq. (6) as a time-independent control parameter, a bifurcation diagram can be easily (numerically) calculated for the deterministic case ($\sigma_s = 0$). In Fig. 2, such diagrams are plotted for C versus the equilibrium temperature T for two values of α_1. To obtain realistic values for the temperature, the temperature equilibria are shifted upwards by 30 K. This is done by substituting T with $T - 30$ and adapting the right-hand side of Eq. (6) such that the new temperature is a steady state. This is obviously a bit artificial here, but we justify it by our aim to illustrate the methodology; the results from more realistic models will follow in Sect. 4. The diagram corresponding to $\alpha_1 = 0.2$ (Fig. 2a) has two saddle-node bifurcations which are absent for $\alpha_1 = 0.45$ (Fig. 2b). From now on, the energy balance model with $\alpha_1 = 0.45$ and $\alpha_1 = 0.2$ will be called the monostable and bistable case, respectively.
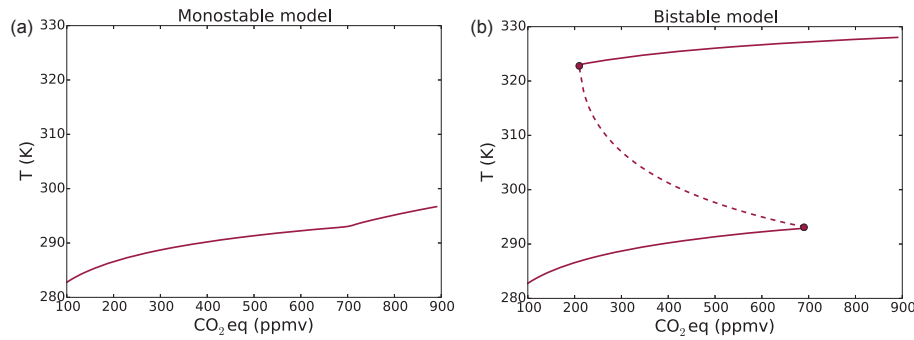
Figure 2. Bifurcation diagram of the deterministic energy balance model for $\alpha_1 = 0.45$ (**a**; monostable model) and $\alpha_1 = 0.2$ (**b**; bistable model). The solid curve represents a stable equilibrium, while the dashed curve represents an unstable equilibrium.

For $\sigma_s \neq 0$, we explicitly determine the normalized PDF $p(x, t)$ by rewriting Eq. (6) as

$$dT_t = f(T_t, t)dt + \sigma_s dW_t, \qquad (8)$$

with $f(T, t) = c_T^{-1}(Q_0(1 - \alpha(T)) + G + A \ln \frac{C(t)}{C_{\text{ref}}} - \sigma \epsilon T^4)$, the Fokker–Planck equation of Eq. (8), which is given by

$$\frac{\partial p}{\partial t} + \frac{\partial (fp)}{\partial x} - \frac{\sigma_s^2}{2} \frac{\partial^2 p}{\partial x^2} = 0. \qquad (9)$$

This differential equation is solved numerically for $p(x, t)$ under any prescribed function $C(t)$ with boundary conditions $p(x_u, t) = p(x_l, t) = 0$, where $x_l = 270$ K, $x_u = 335$ K, and an initial condition $p(x, 0)$ (specified below) satisfies $\int_{x_l}^{x_u} p(x, 0)dx = 1$.

We first show stochastic viability kernels for each initial condition T_0 and C_0, where C_0 is an initial CO_2-equivalent concentration and T_0 is the expectation value of the initial PDF of T_t. As a starting time, we take the year 2030 and suppose that the climate system will be forced by a certain RCP scenario from 2030 to 2200. For every C_0, the original RCP scenario from Fig. 1a is adjusted such that its time development remains the same, but it has C_0 as the CO_2-equivalent concentration in 2030. The PDF of the GMST $p(x, t = 0)$ ($t = 0$ refers to the year 2030) has a prescribed variance (defined by σ_s^2) and expectation value T_0.

In Fig. 3, the stochastic viability kernels are plotted for the energy balance model forced by the RCP4.5 scenario and a viable region V defined by $T \leq 293$ K. The results for the monostable and bistable cases are plotted in Fig. 3a and b, respectively. The colours indicate, for each combination of T_0 and C_0, in which stochastic viability kernel the initial state (C_0, T_0) is located. For example, consider the bistable case and an initial condition of $T_0 = 288$ K and $C_0 = 400$ ppmv; this initial condition is in the kernel V_β with $\beta \geq 0.9$. This means that, with a probability larger than 0.9, a trajectory of the model starting at (C_0, T_0) will remain viable up to the year 2200, where C follows the RCP4.5 scenario. The white areas contain initial conditions that are in a stochastic viability kernel V_β with $\beta < 0.5$.
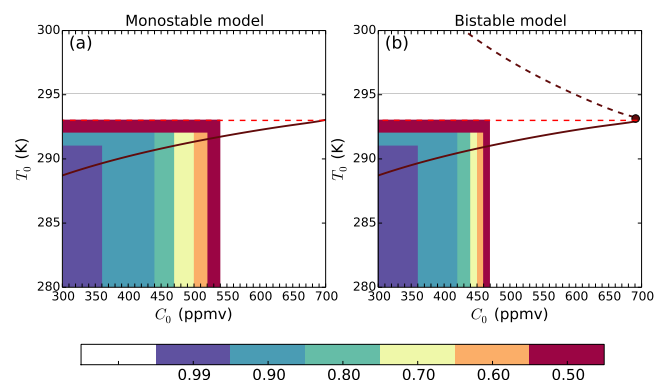

Figure 3. The stochastic viability kernels for the monostable and bistable cases forced by the RCP4.5 scenario. The viable region is defined as $T \leq 293$ K and is indicated by the red dashed line. These plots show, for each combination of T_0 and C_0, in which stochastic viability kernel these initial values are located. The numbers in the colour bar represent the β in V_β. For convenience, the bifurcation diagram of the deterministic model is also shown.

The sensitivity of the stochastic viability kernels with respect to the RCP scenario, the threshold defining the viable region V, and the amplitude of the noise σ_s was also investigated (results not shown). The behaviour is as expected in that the area of the kernels becomes smaller (larger) when noise is larger (smaller), when the threshold temperature is smaller (larger), and when the radiative forcing associated with the RCP scenario is more (less) severe. For example for the RCP6.0 scenario, each combination of T_0 and C_0 (same range as in Fig. 3) is in a V_β with $\beta < 0.5$ for both monostable and bistable cases.

3.3 Results: point of no return

Again for illustration purposes, we assume that a reduction in the emissions will have an immediate effect on the CO_2 eq. such that effectively the CO_2 eq. is controlled. We choose the collection F_λ to consist of mitigation scenarios that exponentially decay to the pre-industrial CO_2-equivalent concentra-

Table 1. Value and meaning of the parameters in the energy balance model given by Eq. (6).

c_T	$5.0 \times 10^8 \, \mathrm{Jm^{-2} \, K^{-1}}$	Thermal inertia	ϵ	1.0	Emissivity
Q_0	$342 \, \mathrm{Wm^{-2}}$	Solar constant/4	α_0	0.7	Albedo parameter
G	$1.5 \times 10^2 \, \mathrm{Wm^{-2}}$	Constant	α_1	0.2 or 0.45	Albedo parameter
A	$2.05 \times 10^1 \, \mathrm{Wm^{-2}}$	Constant	T_0	263 K	Albedo parameter
C_{ref}	280 ppmv	Reference CO_2 concentration	T_1	293 K	Albedo parameter
σ	$5.67 \times 10^{-8} \, \mathrm{Wm^{-2} \, K^{-4}}$	Stefan–Boltzmann constant	ϵ_H	0.273 K	Albedo parameter

tion, which is 280 ppmv. For this exponential decay, we consider different e-folding times τ_d. The most extreme scenario has an exponential decay within 50 years, which corresponds to an e-folding time of $\tau_d = 9$ years. Hence, the collection F_λ is given by (for $\tau_d \geq 9$)

$$F_\lambda(t) = (C_{t_c} - 280) \exp\left(-\frac{t - t_c}{\tau_d}\right) + 280. \qquad (10)$$

In this equation, t_c is the time at which the scenario is applied, and C_{t_c} is the associated CO_2-equivalent concentration at that moment.

Next, we determine PNR values π_t^{tol} for the energy balance model when it is forced by the four different RCP scenarios using a tolerance probability of $\beta_T = 0.9$ and a tolerance time of $\tau_T = 20$ years. The π_t^{tol} values for a system forced with the RCP4.5, RCP6.0, and RCP8.5 scenarios are shown in Fig. 4 for both the monostable and bistable cases. As expected, the more extreme the RCP scenario, the earlier the PNR. This can be easily explained by the fact that when the CO_2-equivalent concentration rises faster, the temperature will become non-viable earlier. Consequently, the PNR will be earlier, since the GMST is only allowed to be nonviable for at most τ_T years. When the model is forced with RCP2.6, there is no PNR for either model. The reason for this is that the CO_2-equivalent concentration will remain low throughout the whole period, and consequently the temperature will stay viable. The value of π_t^{tol} for the bistable case is earlier than the value for the monostable case in each scenario. This can be clarified by the fact that the PDF of the temperature in the bistable case will leave the viable region at a lower CO_2-equivalent concentration because of the existence of nearby equilibria.

The sensitivity of π_t^{tol} versus the tolerance time τ_T and the tolerance probability β_T was also investigated, and the results are as expected (and therefore not shown). A longer tolerance time will shift π_t^{tol} to later times; for example, for the RCP4.5 scenario $\pi_t^{tol} = 2071$, 2088, and 2116 for $\tau_T = 0, 20$ and 50 years for the bistable case (for fixed $\beta_T = 0.9$). With a fixed $\tau_T = 20$ years, the value of π_t^{tol} shifts to smaller values when the tolerance probability is increased. For example, for $\beta_T = 0.80$ and 0.99, the values of π_t^{tol} are 2127 and 2058, respectively, for the bistable case (for $\beta_T = 0.9$, $\pi_t^{tol} = 2088$; see Fig. 4).
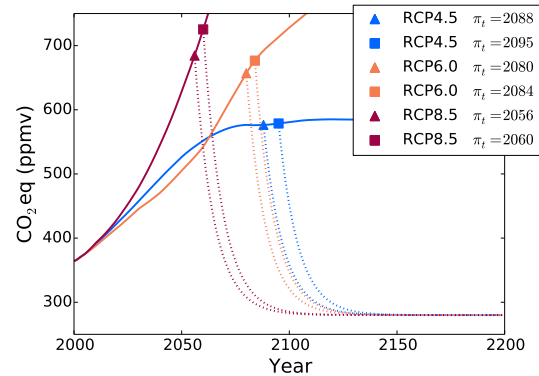

Figure 4. The PNR π_t^{tol} for a system forced with different RCP scenarios, tolerance probability $\beta_T = 0.9$, and tolerance time $\tau_T = t_e - t_b = 20$ years. The triangles indicate the point of no return for the bistable case and the squares for the monostable case. The dotted line is the most extreme scenario of F_λ with an exponential decay to 280 ppmv and an e-folding time of 9 years. Note that for both cases there is no PNR when the model is forced with the RCP2.6 scenario.

4 PlaSim

The results in the previous section have illustrated that a PNR can be calculated when an estimate of the probability density function is available and a collection of mitigation scenarios is defined. We will now apply these concepts to the more detailed, high-dimensional climate model PlaSim, a general circulation model developed by the University of Hamburg (see https://www.mi.uni-hamburg.de/en/arbeitsgruppen/theoretische-meteorologie/modelle/plasim.html).

Atmospheric dynamical processes are modelled using the primitive equations formulated for vorticity, divergence, temperature, and the logarithm of surface pressure. Moisture is included by transport of water vapour (specific humidity). The governing equations are solved using the spectral transform method on a T21 grid (resulting in a horizontal resolution of about 5 to 6° in the midlatitudes). In the vertical, five non-equally spaced sigma (pressure divided by surface pressure) levels are used. Considering that the model has about 100 prognostic variables, this yields an order of magnitude for the model dimensionality of about 10^5. The model is forced by diurnal and annual cycles. PlaSim is freely available, including a graphical user interface facilitating its us-

age, and has been applied to a variety of problems in climate response theory (Ragone et al., 2016).

A main problem here is to determine a relation between the CO_2 eq. (and associated radiative forcing) and the GMST, i.e. a response function. Previous approaches have used a fit of a specific response function (e.g. a power law function) to available observations (Rypdal, 2016). This is more complicated for an approach using stochastic viability theory (applying it did not produce useful results), and hence we proceed by using linear response theory as described in Sect. 2.3.

We use the same data as in Ragone et al. (2016) provided by F. Lunkeit and V. Lucarini (University of Hamburg, Germany). The difference from those in Ragone et al. (2016) is that the seasonal forcing is present, which results in a long-term increase in the GMST of 5 °C (instead of 8 °C in Ragone et al., 2016) under a scenario in which the CO_2 concentration doubles. The reason for this difference is not fully clear but probably results from seasonal rectification effects of nonlinear feedbacks. GMST data from two ensembles were used, each with 200 simulations made with two different CO_2-forcing profiles (all other GHGs are kept constant). For both forcing profiles, the starting CO_2 concentration is set to a value of 360 ppmv, which is representative of the CO_2 concentration in 2000. During the first set of experiments, the CO_2 concentration is instantaneously doubled to 720 ppmv and kept constant afterwards. During the second set of experiments, the CO_2 concentration increases each year by 1 % until a concentration of 720 ppmv is reached. This will take approximately 70 years, and afterwards the concentration is fixed. The total length of the simulations is 200 years. Furthermore, the forcing $f(t)$ in Eq. (4) is taken as the logarithm of the CO_2 concentration, since the radiative forcing scales approximately logarithmically with the CO_2 concentration.

To determine the PDF of GMST under any CO_2-equivalent forcing, we make the assumption that at each point in time the PDF of the GMST is normally distributed. As we have 200 data points for the GMST at each time interval, a χ^2 test was used to analyse the PDFs. For each time, the value of $\chi^2 > 0.05$ and therefore the assumption that the PDF of the GMST is normally distributed appears justified.

The Green's functions for the expectation value and variance of GMST have been calculated with the instantaneously doubling CO_2 profile and the associated ensemble. From the ensemble, at each point in time the expectation value and variance are calculated to obtain the temporal evolution of these two variables. Subsequently, we have found the Green's functions using Eq. (5). To check whether these Green's functions perform well, we compared the temporal evolution of the expectation value and variance of the GMST under the 1 % forcing (calculated with Eq. 4) with those directly generated by PlaSim (Fig. 5).

The expectation value determined with LRT is close to the one directly generated by PlaSim. However, the variance of the ensemble generated by PlaSim is much noisier than the one calculated with LRT. Although the Green's function of

the variance provides only a rough approximation, it has the right order of magnitude and we will use it to calculate the variance of the GMST for other forcing scenarios.

4.1 Results: point of no return under CO_2-equivalent control

We first consider the case without a carbon cycle model, again assuming that the CO_2-equivalent concentration can be controlled directly. The scenarios F_λ chosen for use in PlaSim exponentially decay to different stabilization levels (varying between 400 and 550 ppmv; see Edenhofer et al., 2010). This stabilization level is taken as the parameter λ. We assume that stabilization happens within 100 years, which corresponds to an e-folding time τ_d of about 25 years; the mitigation scenarios F_λ are then given by

$$F_\lambda(t) = \left(C_{t_c} - \lambda\right) \exp\left(-\frac{t - t_c}{\tau_d}\right) + \lambda, \tag{11}$$

where t_c is again the time at which the mitigation scenario is applied, and C_{t_c} is the associated CO_2-equivalent concentration. The most extreme mitigation scenario in F_λ in terms of CO_2-equivalent decrease is the one that stabilizes at a CO_2-equivalent concentration of 400 ppmv.

We next determine the PNR π_t^{2100} by requiring the GMST to be viable in 2100 using a tolerance probability of $\beta_T = 0.90$. The viable region is set at $T \leq 16.15$ °C, which corresponds to temperatures less than 2 °C above the pre-industrial GMST.

The values of π_t^{2100} for all the RCP scenarios are plotted in Fig. 6a. Solid curves show the RCP scenarios, while dashed curves present the most extreme scenario F_λ. For RCP8.5, π_t^{2100} is 10 years earlier than for RCP6.0, since the CO_2-equivalent concentration increases much faster for the RCP8.5 scenario. The mitigation scenario after the point of no return, represented by the dashed line, is the same for all RCP scenarios. This is related to our definition of π_t^{2100}, for which the GMST is required to be viable in 2100. The mitigation scenario that is plotted is the ultimate scenario that guarantees this. It indicates that for each CO_2 scenario the associated π_t^{2100} is given by the intersection of that CO_2-equivalent scenario and the mitigation scenario. This is because an exponential decay to 400 ppmv within 100 years is considered always possible, no matter the CO_2-equivalent concentration at t_c. However, when this concentration becomes too high, this mitigation scenario is no longer very realistic.

The influence of the tolerance probability on π_t^{2100} for the RCP4.5 scenario is plotted in Fig. 6b, for which we only consider a tolerance probability of 0.8, 0.9, and 0.99. When the tolerance probability is higher, it takes longer before the GMST will be viable again and thus the PNR π_t^{2100} will be earlier. However, the differences are very small since the mitigation scenarios that guarantee viability in 2100 for the different tolerance probabilities are very close.
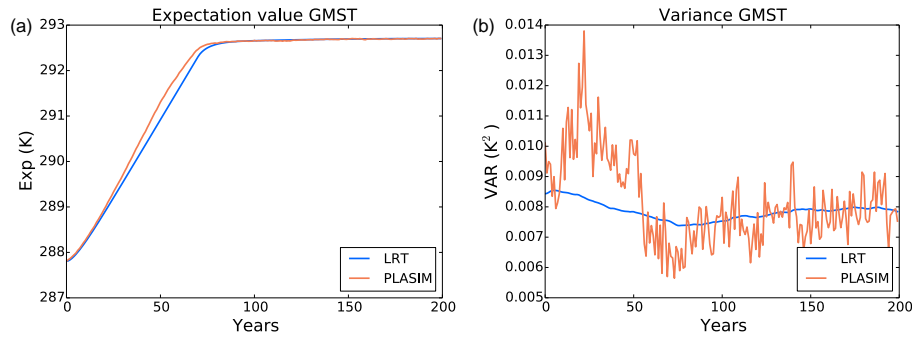
Figure 5. (a) The expectation value and (b) variance of GMST generated by PlaSim (orange) and determined through LRT (blue) for the 1 % CO_2 concentration increase.
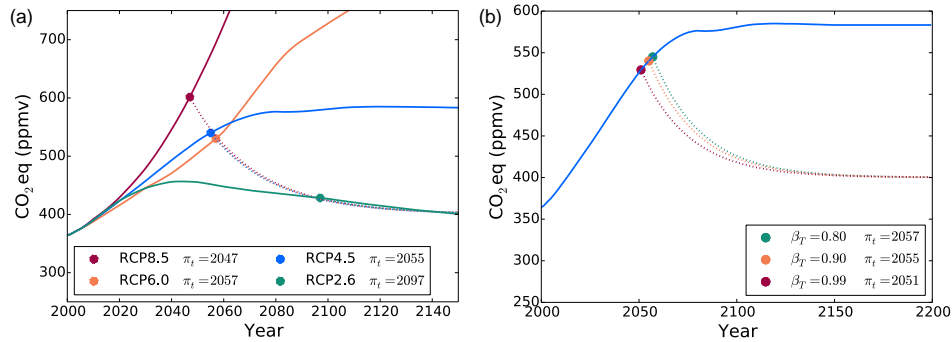

Figure 6. (a) The PNR π_t^{2100} for the RCP2.6, RCP4.5, RCP6.0, and RCP8.5 scenarios for a tolerance probability of $\beta_T = 0.9$ and $\Delta t = 0$. The solid lines represent the RCP scenarios and the dashed line the most extreme scenario from F_λ. Note that these dashed lines coincide. (b) The point of no return for RCP4.5 for different tolerance probabilities.

4.2 Results: point of no return under emission control

Finally, we consider a more realistic case in which emissions are controlled and a carbon model converts emissions to CO_2 eq. A simple carbon model relating emissions E to concentrations C is given by

$$C_{CO_2}(t) = C_{CO_2,0} + \int_0^t G_{CO_2}(\tau)\, E_{CO_2}(t-\tau)\mathrm{d}\tau, \qquad (12)$$

where $C_{CO_2,0}$ is the initial concentration. The Green's function for CO_2 is taken directly from (Joos et al., 2013):

$$G_{CO_2}(t) = a_0 + \sum_{i=1}^{3} a_i e^{t/\tau_i}, \qquad (13)$$

for which the parameters are shown in Table 2. The quantity E_{CO_2} is the CO_2 emission in ppm yr^{-1} that has been converted from ppm yr^{-1} using the carbon molecular weight as $E_{CO_2}[\mathrm{ppm\,yr}^{-1}] = \gamma\, E_{CO_2}[\mathrm{GtC\,yr}^{-1}]$ with $\gamma = 0.46969$ ppm GtC^{-1}. The emissions for the RCP scenarios are taken from (Meinshausen et al., 2011) [1]. The carbon model underestimates CO_2 levels for very high emission scenarios as it does not include saturation of natural CO_2 sinks.

[1] See the database at http://www.pik-potsdam.de/~mmalte/rcps/

Following Table 8.SM.1 of (Myhre et al., 2013), we obtain the changes in radiative forcing compared to pre-industrial (in Wm^{-2}) due to changes in CO_2 as

$$\Delta F_{CO_2} = \alpha_{CO_2} \ln \frac{C_{CO_2}}{C_0}, \qquad (14)$$

where C_0 is the pre-industrial (1750) CO_2 concentration.

We use the same PlaSim ensemble of instantaneous CO_2 doubling runs again to determine a Green's function that relates radiative forcing changes to temperature changes as

$$\Delta T(t) = \int_0^t G_T(\tau)\Delta F(t-\tau)\mathrm{d}\tau, \qquad (15)$$

where G_T is the data-based function determined from LRT. The total radiative forcing is taken as $\Delta F = A\Delta F_{CO_2}$, where we introduce a scaling constant A to correct for the high climate sensitivity of the PlaSim model compared to typical CMIP5 models. Based on trial runs attempting to reconstruct mean CMIP5 RCP temperature trajectories with RCP CO_2 emissions, we choose $A = 0.6$.

For PlaSim, the Green's function G_T, as determined through LRT, is well approximated by a one-timescale ex-

Table 2. Model parameters. No units are given for dimensionless parameters.

$C_{CO_2,0}$ (ppm)	278	α_{CO_2}	5.35	C_0 (ppm)	278
a_0	0.2173	A	0.6	γ (ppm GtC^{-1})	0.46969
a_1	0.2240	a_2	0.2824	a_3	0.2763
τ_1 (yr)	394.4	τ_2 (yr)	36.54	τ_3 (yr)	4.304
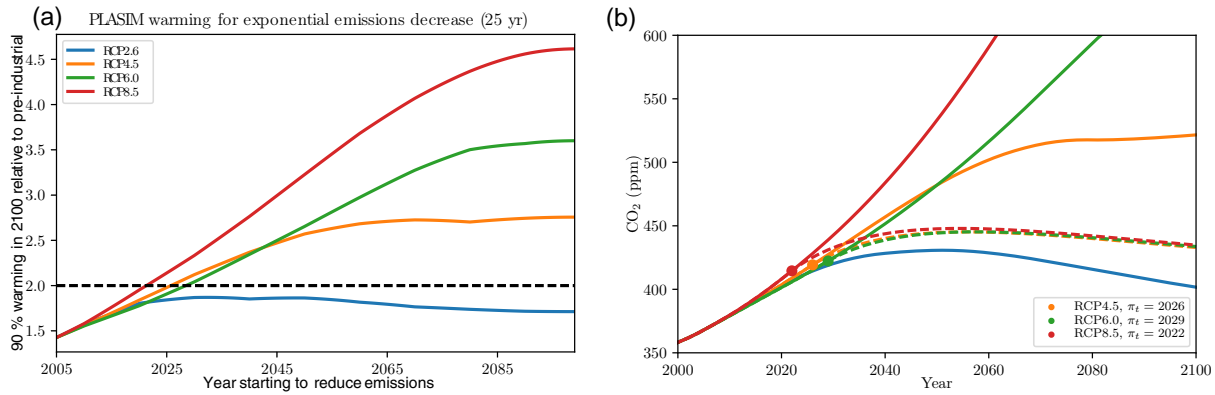
Figure 7. (a) Warming in 2100 when starting exponential CO_2 emission reduction in a given year. **(b)** CO_2 concentration for the four RCP scenarios as computed by the model (solid) and following exponential mitigation starting from the point of no return (dashed).

ponential:

$$G_T(t) \approx b_1 e^{-t/\tau_{b1}}, \tag{16}$$

with $b_1 = 0.25\,\text{KW}^{-1}\,\text{m}^2\,\text{yr}^{-1}$ and $\tau_{b1} = 4.69$ years. We determine a Green's function for the temperature variance in the same way.

To compute the point of no return π_t^{2100} in the carbon–climate model, we start from pre-industrial CO_2 concentrations and take the corresponding initial temperature perturbation as $\Delta T = 0$. We then prescribe the RCP emission scenarios for RCP2.6, RCP4.5, RCP6, and RCP8.5 (that are identical up to the year 2005). At a year $t_b > 2005$ we start the reduction in emissions at an exponential rate; i.e. for $t > t_b$ the emissions follow

$$E_{CO_2}(t) = E_{CO_2}(t_b)\exp\left(-\frac{t - t_b}{\tau_e}\right), \tag{17}$$

where $\tau_e = 25$ years is the e-folding timescale of the emission reduction that we keep constant. Using the carbon model, we compute the instantaneous CO_2 concentrations for each such scenario and use the Green's functions for GMST mean and variance to determine the PDF in the year 2100 for each starting year t_b. Assuming Gaussian distributions (as mentioned, this is well satisfied for the original PlaSim ensemble), we can then easily determine the temperature threshold below which 90 % of the values fall. The first year for which this threshold is above 2 K gives π_t^{2100}. Note that the value of t_c (in Fig. 1b), at which the CO_2 starts to decrease, is determined by the coupled carbon–climate model.

The warming in 2100 predicted by our simple climate model when starting exponential CO_2 emission reduction in a given year is shown in Fig. 7a. The intersections between the RCP curves (solid colour) and the dashed line (representing 2 K warming) provide values of π_t^{2100}. Values do not differ much for the different RCP (4.5, 6.0, and 8.5) scenarios and are before 2030. RCP2.6 does not have a point of no return as its emission scenario is sufficient to keep the warming safely below 2 K. The counter-intuitive lowering of the curve for RCP2.6 (also slightly for RCP4.5) is due to very fast emission reductions in these RCP scenarios. Starting emission reduction at later times may therefore lead to lower total emissions (and hence temperatures). The CO_2 concentration for the four RCP scenarios as computed by the model (solid) and following the exponential mitigation starting from the point of no return (dashed) is shown in Fig. 7b. Note how emissions "still in the pipeline" lead to CO_2 increases even after the reduction is initiated.

Note that this approach does not factor in the uncertainty in the carbon model as we do not have a Green's function propagating the carbon uncertainty through the temperature model. Including this would very likely increase the variance in the PDF and move the point of no return to an earlier year. On the other hand, the PlaSim variance is quite small, so the 90 % threshold is not vastly different from the mean.

5 Discussion

Pachauri et al. (2014) stated with high confidence that "without additional mitigation efforts beyond those in place today, and even with adaptation, warming by the end of the 21st century will lead to high to very high risk of severe, widespread and irreversible impacts globally". If no measures are taken to reduce GHG emissions during this century and if there are no new technological developments that can reduce GHGs in the atmosphere, it is likely that the GMST will be 4 °C higher than the pre-industrial GMST at the end of the 21st century (Pachauri et al., 2014). Consequently, it is important that anthropogenic emissions are regulated and significantly reduced before widespread and irreversible impacts occur. It would help motivate mitigation to know when it is "too late".

In this study we have defined the concept of the point of no return (PNR) in climate change more precisely using stochastic viability theory and a collection of mitigation scenarios. For an energy balance model, as in Sect. 3, the probability density function could be explicitly computed, and hence stochastic viability kernels could be determined. The additional advantage of this model is that a bistable regime can easily be constructed to investigate the effects of tipping behaviour on the PNR. We used this model (with the assumption that CO_2 could be controlled directly instead of through emissions) to illustrate the concept of PNR based on a tolerance time for which the climate state is non-viable. For the RCP scenarios considered, the PNR is smaller in the bistable than in the monostable regime of this model. The occurrence of possible transitions to warm states in this model indeed cause the PNR to be "too late" earlier.

The determination of the PNR in the high-dimensional PlaSim climate model, however, shows the key innovation in our approach, i.e. the use of linear response theory (LRT) to estimate the probability density function of the GMST. PlaSim was used to compute another variant of a PNR based only on the requirement that the climate state is viable in the year 2100. Hence, the PNR here is the time at which no allowed mitigation scenario can be chosen to keep GMST below a certain threshold in the year 2100 with a specified probability. In the PlaSim results, we used a viability region defined as GMSTs lower than 2 °C above the pre-industrial value, but with our methodology, the PNR can be easily determined for any threshold defining the viable region. The more academic case in which we assume that GHG levels can be controlled directly provides PNR (for RCP4.5, RCP6.0, and RCP8.5) values around 2050 (Sect. 4.2). However, the more realistic case in which the emissions are controlled (Sect. 4.3) and a carbon model is used reduces the PNR for these three RCP scenarios by about 30 years. The reason is that there is a delay between the decrease in GHG gas emissions and concentrations.

Although our approach provides new insights into the PNR in climate change, we recognize that there is potential for substantial further improvement. First of all, the PlaSim model has a too-high climate sensitivity compared to CMIP5 models. Although in the most realistic case (Sect. 4.3) we somehow compensate for this effect, it would be much better to apply the LRT approach to CMIP5 simulations. Second, in the LRT approach, we assume the GMST distributions to be Gaussian. This is well justified in PlaSim, as can be verified from the PlaSim simulations, but it may not be the case for a typical CMIP5 model. Third, for the more realistic case in Sect. 4.3, we do not capture the uncertainties in the carbon model and hence in the radiative forcing.

A large ensemble such as that available for PlaSim is not available (yet) for any CMIP5 model. However, we have recently applied the same methodology to two CMIP5 model ensembles, i.e. a 34-member ensemble of abrupt CO_2 quadrupling and a 35-member ensemble of smooth 1 % CO_2 increase per year. The CO_2-quadrupling ensemble was used to derive the Green's function, and then the 1 % CO_2 increase ensemble was used as a check on the resulting response. The probability density function of GMST increase is close to Gaussian for the 1 % CO_2 increase ensemble but clearly deviates from a Gaussian distribution for the 4x CO_2-forcing ensemble, particularly at later times. Although the ensemble is relatively small and the models within the ensemble are different (but many are related), the results for the LRT-determined GMST response (Aengenheyster, 2017) are surprisingly good. This indicates that the methodology has a high potential to be successfully applied to the results of CMIP5 model simulations (and in the future, CMIP6). The applicability of LRT to other observables than GMST can in principle be performed, but the results may be less useful (e.g. due to non-Gaussian distributions).

Because PlaSim is highly idealized compared to a typical CMIP5 model, one cannot attribute much importance to the precise PNR values obtained for the PlaSim model as in Fig. 7. However, we think that our approach is general enough to handle many different political and socioeconomic scenarios combined with state-of-the-art climate models when adequate response functions of CMIP5 models have been determined (e.g. using LRT). Hence, it will be possible to make better estimates of the PNR for the real climate system. We therefore hope that these ideas on the PNR in climate change will eventually become part of the decision-making process during future discussions about climate change.

References

Aengenheyster, M.: Point of No Return and Optimal Transitions in CMIP5, Faculty of Science Theses (Master thesis), Utrecht University archive, available at: https://dspace.library.uu.nl/handle/1874/351607, 2017.

Aubin, J.-P.: Viability theory, Springer Science & Business Media, New York, 2009.

Budyko, M.: Effect of solar radiation variation on climate of Earth, Tellus, 21, 611–619, 1969.

Doyen, L. and De Lara, M.: Stochastic viability and dynamic programming, Syst. Control Lett., 59, 629–634, 2010.

Edenhofer, O., Knopf, B., Barker, T., Baumstark, L., Bellevrat, E., Chateau, B., Criqui, P., Isaac, M., Kitous, A., Kypreos, S., and Leimbach, M.: The economics of low stabilization: model comparison of mitigation strategies and costs, Energ. J., 31, 11–48, 2010.

Fraedrich, K., Jansen, H., Kirk, E., Luksch, U., and Lunkeit, F.: The Planet Simulator: Towards a user friendly model, Meteorol. Z., 14, 299–304, 2005.

Hansen, J., Sato, M., Kharecha, P., Beerling, D., Berner, R., Masson-Delmotte, V., Pagani, M., Raymo, M., Royer, D. L., and Zachos, J. C.: Target atmospheric CO_2: Where should humanity aim?, The Open Atmospheric Science Journal, 2, 217–231, https://doi.org/10.2174/1874282300802010217, 2008.

Heitzig, J., Kittel, T., Donges, J. F., and Molkenthin, N.: Topology of sustainable management of dynamical systems with desirable states: from defining planetary boundaries to safe operating spaces in the Earth system, Earth Syst. Dynam., 7, 21–50, https://doi.org/10.5194/esd-7-21-2016, 2016.

Hogg, A. M.: Glacial cycles and carbon dioxide: A conceptual model, Geophys. Res. Lett., 35, L01701, https://doi.org/10.1029/2007GL032071, 2008.

Joos, F., Roth, R., Fuglestvedt, J. S., Peters, G. P., Enting, I. G., von Bloh, W., Brovkin, V., Burke, E. J., Eby, M., Edwards, N. R., Friedrich, T., Frölicher, T. L., Halloran, P. R., Holden, P. B., Jones, C., Kleinen, T., Mackenzie, F. T., Matsumoto, K., Meinshausen, M., Plattner, G.-K., Reisinger, A., Segschneider, J., Shaffer, G., Steinacher, M., Strassmann, K., Tanaka, K., Timmermann, A., and Weaver, A. J.: Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis, Atmos. Chem. Phys., 13, 2793–2825, https://doi.org/10.5194/acp-13-2793-2013, 2013.

Mann, M. E.: Defining dangerous anthropogenic interference, P. Natl. Acad. Sci. USA, 106, 4065–4066, 2009.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic Change, 109, 213–241, https://doi.org/10.1007/s10584-011-0156-z, 2011.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, T., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing Supplementary Material, in: Climate Change 2013 – The Physical Science Basis, edited by: Jacob, D., Ravishankara, A. R., and Shine, K., chap. 8, Cambridge University Press, Cambridge, 2013.

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q., Dasgupta, P., and Dubash, N. K.: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 2014.

Petschel-Held, G., Schellnhuber, H. J., Bruckner, T., Tóth, F. L., and Hasselmann, K.: The Tolerable Windows Approach: Theoretical and Methodological Foundations, Climatic Change, 41, 303–331, 1999.

Ragone, F., Lucarini, V., and Lunkeit, F.: A new framework for climate sensitivity and prediction: a modelling perspective, Clim. Dynam., 46, 1459–1471, 2016.

Ragone, F., Lucarini, V., and Lunkeit, F.: A new framework for climate sensitivity and prediction: a modelling perspective, Clim. Dynam., 46, 1459–1471, https://doi.org/10.1007/s00382-015-2657-3, 2016.

Rogelj, J., Hare, W., Lowe, J., van Vuuren, D. P., Riahi, K., Matthews, B., Hanaoka, T., Jiang, K., and Meinshausen, M.: Emission pathways consistent with a 2 °C global temperature limit, Nature Publishing Group, 1, 413–418, 2011.

Rypdal, K.: Global warming projections derived from an observation-based minimal model, Earth Syst. Dynam., 7, 51–70, https://doi.org/10.5194/esd-7-51-2016, 2016.

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

Smith, J. B., Schneider, S. H., Oppenheimer, M., Yohe, G. W., Hare, W., Mastrandrea, M. D., Patwardhan, A., Burton, I., Corfee-Morlot, J., Magadza, C. H., and Füssel, H. M.: Assessing dangerous climate change through an update of the Intergovernmental Panel on Climate Change (IPCC) "reasons for concern", P. Natl. Acad. Sci. USA, 106, 4133–4137, 2009.

Van Zalinge, B., Feng, Q. Y., Aengenheyster, M., and Dijkstra, H. A.: PNRESD: Data for "On determining the point of no return in climate change", https://doi.org/10.5281/zenodo.838675, 3 August 2017.

Victor, D. G. and Kennel, C. F.: Climate policy: Ditch the 2 C warming goal, Nature, 514, 30–31, 2014.