Point of No Return and Optimal Transitions in CMIP5

Matthias Aengenheyster[1]†

[1]Institute for Marine and Atmospheric Research Utrecht, Department of Physics and Astronomy, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands

The evolution of the climate in terms of global mean surface temperature (GMST) under anthropogenic forcing scenarios is explored with respect to the warming targets of 1.5 K and 2 K as set in the Paris agreement. For the first time we successfully use Linear Response Theory to combine a large number of complex climate models from the CMIP5 ensemble, allowing us to determine the GMST response to any CO_2 forcing scenario. We construct a simple stochastic model, forced by CO_2 emissions, that well recovers ensemble mean and variance of the CMIP5 GMST. The Point of No Return, the point in time when it is too late to reach a target, is derived, and we show its dependence on cumulative emissions, climate uncertainty, risk tolerance and stringency of efforts to combat climate change. This yields an easy-to-communicate result to inform debate on climate change action. Based on our assessment the 1.5 K target is not deemed reachable anymore, and only a limited time remains to start ambitious efforts to reach the 2 K goal. In addition, using simple economic assumptions we find optimal pathways for energy transitions to a carbon-free era and show how they depend on the warming target.

1. Introduction

The Earth System is currently in a state of rapid change, in particular rapid warming, that is unprecedented in the historical temperature and geological greenhouse gas concentrations records (Pachauri *et al.* 2014). This change is primarily driven by the rapid increase in atmospheric concentrations of greenhouse gases (GHG) due to anthropogenic emissions since the industrial revolution (Pachauri *et al.* 2014; Myhre *et al.* 2013a). Changes in natural physical and biological systems are already being observed (Rosenzweig *et al.* 2008), and efforts are made to determine the "anthropogenic impact" on particular (extreme weather) events (Haustein *et al.* 2016). Nowadays, the question is not so much *if* the climate will change as a result of human interference (which has led to the phrasing of the term 'anthropocene' (Crutzen 2002)) but by *how much* it will change, and whether this will to *dangerous* anthropogenic interference with the climate. The exact dynamics of the climate system in a high-GHG world (such as brought about by unmitigated business-as-usual emission pathways) are not well known and models differ in their predictions for which there may be fundamental reasons (Roe and Baker 2007). In addition, researchers have attempted to identify points in the climate system, labeled 'tipping points' (Lenton *et al.* 2008; Cai *et al.* 2016) or 'critical transitions' (Lucarini and Bódai 2017) that could be triggered by warming above some threshold and lead to large-scale, irreversible change with dramatic consequences for ecosystems, human societies and economies. Lists of such elements commonly include the Atlantic Meridional Overturning Circulation (AMOC), the land-based ice sheets in Greenland and Antarctica, the Indian summer monsoon, and the El Niño Southern

† Email address for correspondence: m.aengenheyster@students.uu.nl

Oscillation (ENSO) (Lenton *et al.* 2008; Cai *et al.* 2016). How exactly the tipping would proceed and what the respective thresholds are is often not well known (Lenton *et al.* (2008), for example, use 'expert elicitation').

As much of the dynamics of the climate behavior is unknown and the system itself is characterized by feedbacks, chaos and stochastic dynamics (Dijkstra 2013), the debate on action on climate change is focused on the question of *risk* and how the *probability* of dangerous climate change can be reduced. In the scientific and political discussion targets on 'allowable' warming (in terms of change in Global Mean Surface Temperature (GMST) relative to pre-industrial conditions) have turned out to be very salient, with the 2 K warming threshold commonly seen as the highest justifiable threshold to avoid the worst effects (Pachauri *et al.* 2014). Indeed, in the Paris Conference it was agreed to attempt to limit warming below 1.5 °C (United Nations 2015). When we speak of "targets" in this context it is to be understood as ensuring non-exceedance of the thresholds.

In this work we want to contribute to the debate on if, and how, these targets can be reached and what they imply on the transition the global economy would have to undergo. In particular, taking into account the chaotic and stochastic nature of the climate and the risk tolerance of decision makers, we want to do so in a probabilistic fashion. Fundamentally, we ask:

If we want to limit warming to some threshold in the year 2100, while accepting a certain risk tolerance of actually exceeding it, what is the latest point in time we have to start to ambitiously reduce fossil fuel emissions?

The point in time when it is 'too late' to act in order to reach the warming target is called the Point of No Return (PONR). Clearly, the value of PONR will depend on a number of things, such as the climate system dynamics and the tools at our disposal to reduce emissions. Crucially, we require a model of the global climate that is a) accurate enough that it gives a realistic picture of the behavior of GMST under climate change scenarios, b) is forced by fossil fuel emissions, c) is simple enough to be evaluated for a very large number of scenarios and d) provides information about risk, i.e. it cannot be purely deterministic.

Fulfilling these requirements sets the outline of this thesis. We start by constructing a simple yet accurate model of GMST change. As the most comprehensive dataset available we use model data from the CMIP5 experiments and apply Linear Response Theory to inform the ensemble behavior (section 2). To our knowledge this is the first attempt to tackle CMIP5 with this methodology. We then continue in section 3 to build a simple State Space Model of seven variables that well represents CMIP5 GMST mean and variance, giving us a simple, easy-to-use, fast-to-integrate tool that is physically motivated, constrained by complex model data and enables us to study a variety of interesting problems. This is what we proceed to do in section 4 where we – with no or very few assumptions on the global economy – determine two metrics that give an intuitive grasp on the behavior of the climate system under anthropogenic forcing: First, the Safe Carbon Budget (SCB), the maximum permissible amount of cumulative fossil CO_2 emissions that still allows us to reach a given target on global warming, and second the Point of No Return (PONR). We show how both metrics depend on the stochastic nature of the climate, the risk we are willing to accept and, for the PONR, the tools at our disposal to combat emissions.

After finding in section 4 simple, easy-to-communicate results that still have a very strong physical backing, we tackle in section 5 the question of determining the *optimal* transition to the carbon-free era, assuming we want to meet a given target. For this we introduce some basic economic theory and find welfare-optimizing transition pathways under different formulations of climate damages. In section 6 we discuss our results and conclude with an outlook on future applications of our methodology.

2. Linear Response Theory on the CMIP5 Ensemble

As part of the CMIP5 experiments† many large-scale coupled climate models were subjected to the same forcings for a number of scenarios (Taylor *et al.* 2012). As all models were designed to represent the same (physical) processes but used different formulations, parameterizations, resolutions and implementations, the underlying idea is that the ensemble of results from different models offers a glimpse into the (statistical) properties of the real climate and that the drawbacks of each model compensate, with the ensemble mean giving a better image of the "truth" than any individual model.

We extend this view and perceive each of these model runs as one possible realization of the real climate, treating them as statistically equivalent. With this view we can apply ideas and methods from statistical physics, namely Linear Response Theory (LRT), that allow us to examine the behavior of the climate system via ensemble statistics.

In this section we briefly present the mathematical framework and the data used, before we apply the theory to the data.

2.1. *Linear Response Theory*

Ruelle (Ruelle 1998a,b) have shown how to obtain the response of general non-equilibrium systems to weak forcings (Linear Response Theory, LRT, in this context thereby often called Ruelle Response Theory (RRT)). This approach was successfully applied to a conceptual climate model (Lucarini and Sarno 2011) and to simplified GCMs (Ragone *et al.* 2016; van Zalinge *et al.* 2017). In these studies LRT has performed well and was able to predict the impact of arbitrary CO_2 forcing scenarios on GMST. At the core of LRT stands the expansion of statistical properties of a chosen observable ϕ (such as mean and variance of GMST) under the forcing $f(t)$ as a perturbation series

$$\phi_f(t) = \phi_0 + \sum_{n=1}^{\infty} \phi_f^{(n)}(t) \tag{2.1}$$

with unperturbed observable ϕ_0 and perturbative terms of order n given by $\phi_f^n(t)$. Linear response theory cuts the series after the first term

$$\phi_f^{(1)}(t) = \int_{-\infty}^{+\infty} G_\phi^{(1)}(t') f(t-t') dt' \tag{2.2}$$

† Coupled Model Intercomparison Project 5

with $G_\phi^{(1)}$ the first order Green's function. With this function known the (linear) response to any forcing can be calculated. Taking the Fourier transform of Eq. 2.2 reveals that

$$\tilde{\phi}_f^{(1)}(\omega) = \chi_\phi^{(1)}(\omega)\tilde{f}(\omega) \tag{2.3}$$

$$\Rightarrow G_\phi(t) = \mathcal{F}^{-1}\left(\frac{\tilde{\phi}_f^{(1)}(\omega)}{\tilde{f}(\omega)}\right) \tag{2.4}$$

with inverse Fourier transform $\mathcal{F}^{-1}$ and susceptibility $\chi_\phi^{(1)}(\omega) = \mathcal{F}\left(G_\phi^{(1)}(t)\right)$. Hence in principle, once the time evolution of observable ϕ under forcing $f(t)$ is known one can apply Eq. 2.4 to determine the Green's function. However, this approach involves the practical complication that it requires taking numerical forward and inverse Fourier transforms of timeseries, many of which may not be 'nice' for that purpose (e.g. not periodic in the window considered) so the result of Eq. 2.4 may not be well-behaved. Therefore, in practice we use the special case of a step-wise forcing $f(t) = A\theta(t)$ with constant A and Heaviside function $\theta(t)$. Plugging this in Eq. 2.2 the expression simplifies to

$$G_\phi^{(1)}(t) = \frac{1}{A}\frac{d}{dt}\left[\phi_f^{(1)}(t)\right] \tag{2.5}$$

which is much more well-behaved and will be used from now on.

2.2. *Data*

The backbone of our study is the CMIP5 data obtained from the German Climate Computing Center (DKRZ)†. We use the pre-industrial control runs (labeled piControl), the forcing scenarios with abrupt CO_2 quadrupling (abrupt4xCO2) and smooth 1% CO_2 increase per year (1pctCO2). Besides those idealized experiments we use the RCP (Representative Concentration Pathways) scenarios labeled 2.6, 4.5, 6.0 and 8.5 that represent very moderate to very high forcing scenarios deemed to be plausible (the numbers give the radiative forcing in $\mathrm{W\,m^{-2}}$ in 2100) (Taylor *et al.* 2012). Their forcings (concentrations (Meinshausen *et al.* 2011) and emissions (van Vuuren *et al.* 2007; Clarke *et al.* 2007; Fujino *et al.* 2006; Riahi *et al.* 2007)) are obtained from the RCP Database‡.

We focus on the annual-mean area-weighted Global Mean Surface Temperature (GMST). For the RCP scenarios GMST is directly available from KNMI's Climate Explorer¶. We decide to use only those ensemble members for which the control run and at least one perturbation run are available, leading to 34 members for the abrupt and 39 for the smooth forcing experiment. Considering those members from the RCP runs that we also have in the abrupt forcing run, we have 25 members for RCP2.6, 30 for RCP4.5, 19 for RCP6.0 and 29 for RCP8.5.

For illustration we show the properties of the RCP4.5 ensemble in Fig. 1, together with the distribution of deviations from the mean response. We consider the gaussianity satisfactory for now and proceed to compute the linear response from the ensemble mean of the simulations.

† World Data Center for Climate at `https://cera-www.dkrz.de/` and the ESGF Node at DKRZ at `https://esgf-data.dkrz.de`

‡ available at `http://tntcat.iiasa.ac.at/RcpDb` and `http://www.pik-potsdam.de/~mmalte/rcps/`

¶ `http://climexp.knmi.nl`

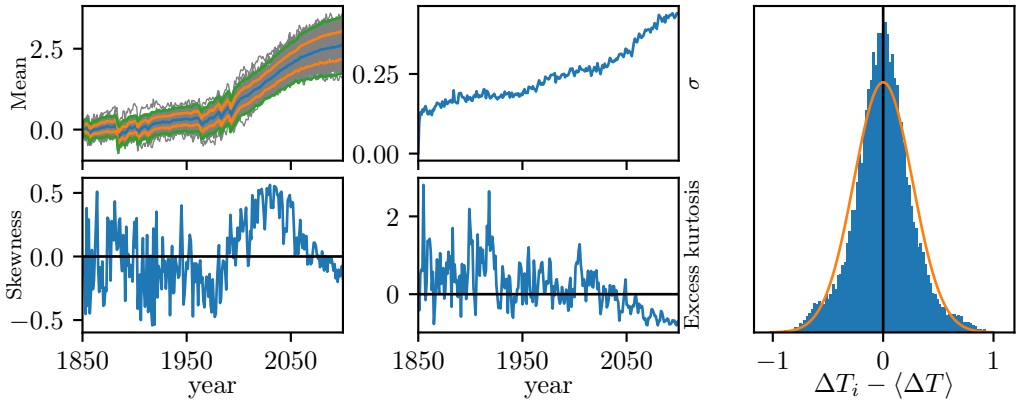

Figure 1: Statistics of RCP4.5 ensemble. Left: Time evolution of ensemble mean (top left, mean in blue, $\pm 1\sigma$ in orange, $\pm 2\sigma$ in green), standard deviation (top right), skewness (bottom left) and excess kurtosis (bottom right). Right: Probability density distribution of deviations from the mean response over 1850-2100, with fitted normal distribution in orange.

2.3. *Application*

The CO_2 concentration as a function of time for the abrupt quadrupling and smooth CO_2 increase is defined as

$$C_{CO2,abrupt}(t) = C_0(3\theta(t) + 1) \tag{2.6}$$

$$C_{CO2,smooth}(t) = \begin{cases} C_0 & , \quad t \leqslant 0 \\ C_0 1.01^t & , \quad t > 0 \end{cases} \tag{2.7}$$

with time in years from the start of the forcing, pre-industrial CO_2 concentration C_0 and Heaviside function $\theta(t)$. We obtain the radiative forcing due to CO_2 relative to pre-industrial conditions from

$$\Delta F = \alpha_{CO2} \ln\left(\frac{C_{CO2}(t)}{C_0}\right) \tag{2.8}$$

with $\alpha_{CO2} = 5.35\,\mathrm{W\,m^{-2}}$ (Myhre *et al.* 2013a). The Green's function is computed as outlined in section 2.1 from the abrupt forcing case as the time derivative of the response

$$G_T(t) = \frac{1}{\Delta F_{abrupt}} \frac{d}{dt} \Delta T_{abrupt} \tag{2.9}$$

where $\Delta F_{abrupt}(t) = const. = \ln(4C_0/C_0) = \ln(4)$. As discussed, we obtain the temperature evolution (i.e. the deviation from the pre-industrial state) for any forcing via the convolution of the Green's function and an arbitrary forcing function ΔF_{any}

$$\Delta T_{any}(t) = \int_0^t G_T(t') \Delta F_{any}(t - t')\, dt' \tag{2.10}$$

In order to judge the quality of our approach we use equation 2.10 to compute the response for our two scenarios. As equation 2.9 is exact we expect with $\Delta F_{any} = \Delta F_{abrupt}$ to exactly reproduce the abrupt CMIP5 response. In addition, for LRT to be useful we have to be able to reasonably reproduce the smooth $1\,\%\,\mathrm{yr^{-1}}$ response with

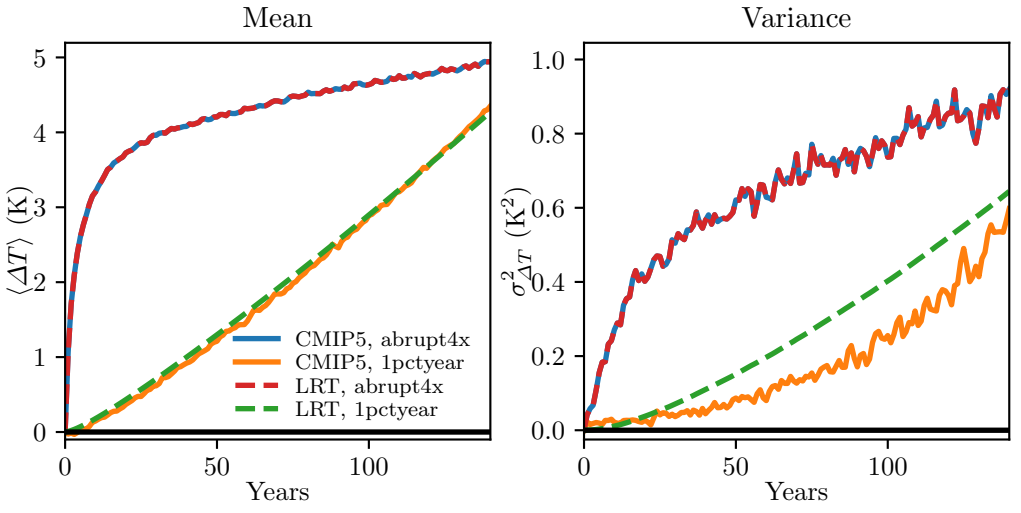

Figure 2: Ensemble mean (left) and variance (right) of temperature response from CMIP5 (solid) and LRT reproduction (dashed). Year 0 gives the start of the perturbation.

$\Delta F_{any} = \Delta F_{smooth}$. We show the results in Fig. 2. We find that LRT applied to the abrupt perturbation recovers perfectly – as required – the abrupt response and is well able to recover the response to a smooth forcing. The fit is very good for the mean reponse and also captures the variance quite well. This is not a trivial finding, given the drastic difference in time evolution of the forcing, and the – until now – unanswered question whether LRT can be applied to the CMIP5 ensemble at all.

After computing the Green's function from the abrupt forcing scenarios and testing the performance on the smooth forcing run we turn to the more realistic RCP scenarios. We compute the radiative forcing according to Eq. 2.8 and temperature perturbations via Eq. 2.10. However, in the RCP simulations CO_2 does not tell the whole story. Unlike the idealized simulations they include both non-fossil CO_2 emissions and non-CO_2 GHG emissions, most notably CH_4 and N_2O. We include them by scaling up the radiative forcing obtained from CO_2 emissions by a factor $A = 1.3$ to best reproduce the RCP temperature evolution. We show the resulting reconstruction of RCP temperatures from CO_2 concentrations, overlaid with CMIP5 data, in Fig. 3, and see the good agreement. An alternative approach would be to use concentration timeseries for CO_2, CH_4 and N_2O and compute their respective radiative forcings, then summing to obtain the total. We briefly show this in appendix A but no substantially improved agreement with the CMIP5 ensemble is found.

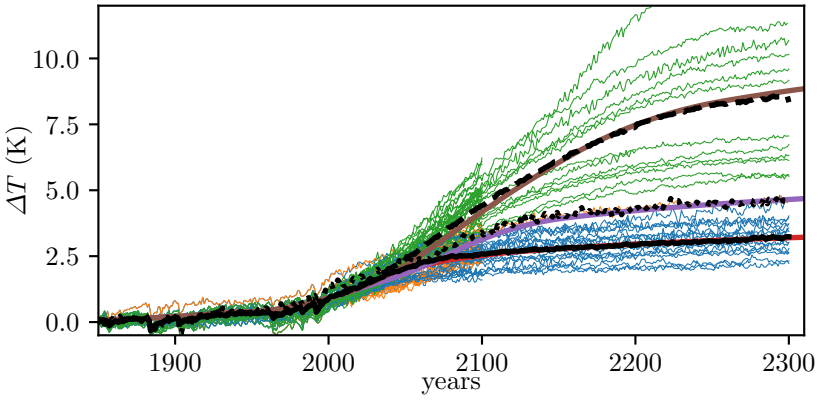

Figure 3: Reconstruction of RCP temperature evolution from concentration pathways using CO_2 only. Blue, orange and green lines gives CMIP5 data for RCP4.5, RCP6.0 and RCP8.5, respectively, with the ensemble mean given in solid (RCP4.5), dotted (RCP6.0) and dashed (RCP8.5) black. Reconstruction using CO_2 radiative forcing in red (RCP4.5), purple (RCP6.0) and brown (RCP8.5)

3. Building a State Space Model

We have shown in the preceding section that it is possible to apply Linear Response Theory to the CMIP5 ensemble to obtain a Green's function that captures the mean response of the ensemble to changes in atmospheric CO_2 concentration (and also to CH_4 and N_2O, if desired). However, we want to go a step further. We are interested in *steering* the climate in some sense, and this necessarily involves influencing the carbon *fluxes* to the atmosphere, be they (positive) fossil fuel emissions or negative emissions, e.g. via bioenergy with carbon capture and storage (BECCS). If and how BECCS could be important for tackling climate change is part of an ongoing discussion (see for example Fuss *et al.* (2014)), but not the topic of this thesis where we focus on fossil emissions. We use this section to build a model that relates (positive) fossil CO_2 emissions to GMST changes. We start by describing the Carbon Model that relates emissions to concentrations, go on to present the full model in response function form, and then translate it to state space where we model the evolution of GMST probability density functions (PDFs) by including stochasticity. We define suitable initial conditions and close with the presentation of the reconstruction of CMIP5 RCP results by our simple model.

3.1. *Carbon Model*

The global carbon cycle and its formulation in terms of response functions have been studied by Joos *et al.* (2013) and included in the Fifth Assessment Report (Myhre *et al.* 2013a), the results of which we will use here. Joos *et al.* (2013) conduct a multi-model study of many carbon models of varying complexity under different background states and forcing scenarios. Their main result is the fit of a three-timescale exponential with constant offset (Eq. 3.1) to the ensemble mean of responses to a 100 GtC emission pulse to a present-day climate.

$$G_{CO2}(t) = a_0 + \sum_{i=1}^{3} a_i e^{-\frac{t}{\tau_i}} \tag{3.1}$$

Coefficients $a_i, i = 0 \dots 3$ and timescales $\tau_i, i = 1 \dots 3$ are determined using least-square fits on the multi-model mean. CO_2 concentrations are then determined equivalently to the temperature model

$$C_{CO2}(t) = \int_0^t G_{CO2}(t - t') \; E_{CO2}(t') \; dt' \tag{3.2}$$

with instantaneous CO_2 emissions E_{CO2}. Joos *et al.* (2013) also give results for pulses of different sizes (100 and 5000 GtC) administered either to a pre-industrial or present-day world. The shapes of these response function vary substantially. As one would expect, a very large pulse leads to saturation of natural carbon sinks and a larger fraction remaining in the atmosphere. A similar behavior occurs when administering the pulse to a present-day (high CO_2) rather than a pre-industrial (low CO_2) world. In this study we use Joos' 'default' case as we are interested in a present-day world. Note however that for high-emission scenarios cumulative emissions are substantially higher than 100 GtC (relative to 2015, $>$ 700 GtC for RCP4.5 and beyond 1800 GtCfor RCP8.5 until 2100). There the 5000 GtC case could be a better, though more pessimistic fit, as it implies slower CO_2 decay in the atmosphere.

In the previous section we noted the possibility to compute radiative forcing based on gases other than CO_2. Then also for these gases one has to construct a model relating emissions to concentrations. This can be done but does not substantially outperform using scaled CO_2 radiative forcing only. Due to the added complexity of initial conditions and emission scenarios for these gases we decided against this option, but briefly present it in appendix B.

3.2. *Full Response Function Model*

Collecting the information from sections 2 and 3.1 we now assemble our model in response function form.

$$C_{CO2}(t) = C_{CO2,0} + \int_0^t G_{CO2}(t') \; E_{CO2}(t - t') \; dt' \tag{3.3a}$$

$$\Delta F_{CO2}(t) = A \; \alpha_{CO2} \ln(C_{CO2}(t)/C_0) \tag{3.3b}$$

$$\Delta T(t) = \Delta T_0 + \int_0^t G_T(t') \Delta F_{CO2}(t - t') dt' \tag{3.3c}$$

It relates fossil CO_2 emissions E_{CO2} to GMST perturbation ΔT relative to pre-industrial levels, with initial conditions for CO_2 $C_{CO2,0}$ and GMST perturbation ΔT_0. Note that this is quite a simple model with few "knobs to turn". The only really free parameter is the constant A that scales up CO_2-radiative forcing to take into account non-fossil CO_2 and non-CO_2 GHG emissions.

In the full model we recalibrate the constant A (previously set to $A = 1.3$ (section 2.3)) to $A = 1.48$† in order to optimize the agreement with CMIP5. It is found empirically but if we let – as a bullpark estimate – 75% of CO_2 emissions be fossil and

† it becomes $A = 1.2$ when including CH_4 and N_2O

10% of radiative forcing come from non-CO_2 GHG, then fossil CO_2 represents 67.5% of radiative forcing so its contribution needs to be scaled up by $A = 1/0.675 \approx 1.48$, providing a sort of a-posteriori rational for the value of A. Internally emissions need to be converted from $\mathrm{GtC\,yr^{-1}}$ to $\mathrm{ppm\,yr^{-1}}$ which is done using the respective molar masses and the mass of the Earth's atmosphere:

$$E(\mathrm{ppm\,yr^{-1}}) = \frac{1}{M_C}\frac{M_{air}}{m_{atm}}\ 10^{12}\,\mathrm{kg\,GtC^{-1}}\ 10^{6}\,\mathrm{ppm}\ \ E(\mathrm{GtC\,yr^{-1}}) \tag{3.4}$$

with $M_{air} = 28.97\,\mathrm{kg\,kmol^{-1}}$ the mean molecular weight of air, $M_C = 12.011\,\mathrm{kg\,kmol^{-1}}$ the mean molecular weight of carbon and $m_{atm} = 5.1352 \times 10^{18}\,\mathrm{kg}$ the total mass of the atmosphere, yielding $E(\mathrm{ppm\,yr^{-1}}) = 0.46969\ E(\mathrm{GtC\,yr^{-1}})$. In Tab. 1 we summarize the model's ten parameters.

C_0 (ppm)	a_0	a_1	a_2	a_3
278	0.2173	0.2240	0.2824	0.2763
A	α ($\mathrm{W\,m^{-2}}$)	τ_1	τ_2	τ_3
1.48	5.35	394.4	36.54	4.304

Table 1: Response Function Model Parameters. All timescales τ_i are in years and the carbon model amplitudes a_i are dimensionless for E in $\mathrm{ppm\,yr^{-1}}$.

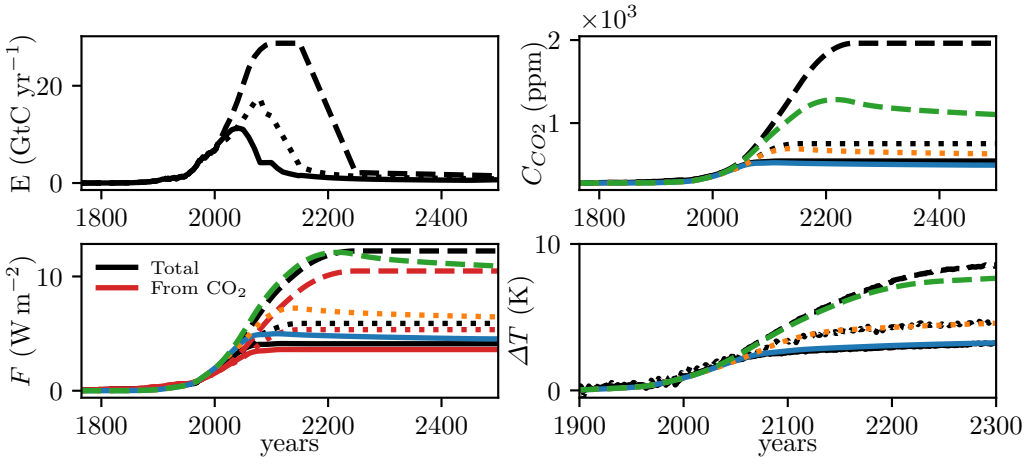

Figure 4: Reconstruction of RCP results using the Reponse Function Model. In all panels, solid lines refer to RCP4.5, dotted to RCP6.0 and dashed lines to RCP8.5. Black lines (and red lines in the bottom left panel) show RCP data while colors (blue: RCP4.5, orange: RCP6.0, green: RCP8.5) give our reconstruction.
Top left: Fossil CO_2 emissions. **Top right:** CO_2 concentrations from RCP and reconstructed using G_{CO2}. **Bottom left:** Total anthropogenic radiative forcing (black) and radiative forcing from CO_2 only (red) (both from RCP) and reconstructed forcing using the relations above. **Bottom right:** Temperature perturbation from CMIP5 RCP (ensemble mean) and the our reconstruction.

In Fig. 4 we show the results obtained for RCP emissions. For very high emission scenarios it is not possible to recover the CO_2 concentrations correctly. This is because the response is not pulse-size independent. In reality for such high emissions natural sinks saturate. However, the up-scaling of radiative forcing is quite successful, yielding a good temperature reconstruction.

3.3. *Stochastic State Space Model*

We now have successfully built a response function model that relates fossil CO_2 emissions to GMST perturbations. However, the model still contains a data-based temperature response function that is somewhat difficult to handle, and it informs only about the *mean* CMIP5 response. But our main motivation is to obtain new insights on the possible transitions to a 'safe', fossil-free, state. These transitions necessarily depend strongly on the variance of the climate and on the risk we are willing to take while making the transition. This variance is quite substantial, as evident from Fig. 3. We have shown (Fig. 2) to be able to build a Green function not only for the temperature mean (left) but also the variance (right). However, we have nothing comparable for the carbon model (though we could build one), and more importantly, no information on how the variance is transmitted from the carbon model through the radiative forcing and the temperature model to the GMST (i.e., we do not have equations of motion for variance). Therefore we now translate our response function model to state space and incorporate the variance via suitable stochastic terms.

The first important step is to rewrite the response functions in a way that can be implemented in state space which is easiest by writing them as sums of exponentials. For the carbon model Joos *et al.* (2013) already do this. For the temperature model it turns out we can well represent G_T with a two-timescale exponential with a small constant offset:

$$G_T(t) = \sum_{i=0}^{2} b_i e^{-t/\tau_{bi}} \tag{3.5}$$

We let the small constant offset b_0 decay over a long time scale $\tau_{b0} = 400\,\mathrm{yr}$ that cannot really be fit from the 140 yr abrupt forcing runs but is required to prevent a perpetual temperature increase to constant radiative forcing with $C > C_0$ and allow temperatures to stabilize at some level. We obtain three 'temperature reservoirs' analogous to the carbon reservoirs in the carbon model, and arrive at the seven-dimensional dynamical system shown in Tab. 2 with parameters in Tab. 3. In the deterministic version ($\sigma_{C2} = \sigma_{T0} = \sigma_{T2} = 0$) the model reproduces the results from the Green's function approach. However, the major benefit is that now we can include stochasticity and model (the spread of) uncertainty throughout the model. We introduce additive noise to the carbon model such that the standard deviation of the model response to an emission pulse as reported by Joos *et al.* (2013) is recovered. For the temperature model we introduce (small) additive noise to recover the (small) CMIP5 control run standard deviation. We observe that in the CMIP5 RCP runs the ensemble variance increases with rising ensemble mean. This calls for the introduction of (substantial) multiplicative noise, which we introduce in ΔT_2, letting these random shocks decay again over an 8-year timescale. The main purpose of this noise is to strongly increase temperature variance with mean temperature. It affects temperature trajectories by introducing substantial shocks that are (especially at high temperatures) almost certainly unphysical when looking at individual timeseries. However, the focus here is on ensemble statistics.

$dC_P = a_0 E dt$	$\Delta F = A\ \alpha \ln(C/C_0)$
$dC_1 = (a_1 E - \frac{1}{\tau_1} C_1)dt$	$d\Delta T_0 = (b_0 \Delta F - \frac{1}{\tau_{b0}} \Delta T_0)dt + \sigma_{T0}\ dW_t$
$dC_2 = (a_2 E - \frac{1}{\tau_2} C_2)dt + \sigma_{C2}\ dW_t$	$d\Delta T_1 = (b_1 \Delta F - \frac{1}{\tau_{b1}} \Delta T_1)dt$
$dC_3 = (a_3 E - \frac{1}{\tau_3} C_3)dt$	$d\Delta T_2 = (b_2 \Delta F - \frac{1}{\tau_{b2}} \Delta T_2)dt \quad + \sigma_{T2} \Delta T_2\ dW_t$
$C = C_P + \sum_{i=1}^{3} C_i$	$\Delta T = \sum_{i=0}^{2} \Delta T_i$

Table 2: Stochastic State Space Model. Carbon model on left, temperature model on the right.

a_0	a_1	a_2	a_3	τ_1	τ_2	τ_3
0.2173	0.2240	0.2824	0.2763	394.4	36.54	4.304
C_0 (ppm)	b_0	b_1	b_2	τ_{b0}	τ_{b1}	
278	0.00115176	0.10967972	0.03361102	400	1.42706247	
A	α ($\mathrm{W\,m^{-2}}$)	σ_{C2} ($\mathrm{ppm/yr^{1/2}}$)	σ_{T0} ($\mathrm{K/yr^{1/2}}$)	σ_{T2} ($\mathrm{yr^{-1/2}}$)	τ_{b2}	
1.48	5.35	0.65	0.015	0.13	8.02118539	

Table 3: Stochastic State Space Model Parameters. All timescales are in years, the carbon model amplitudes a_i are dimensionless for E in $\mathrm{ppm\,yr^{-1}}$, the temperature model amplitudes b_i are in $\mathrm{K\,W^{-1}\,m^2\,yr^{-1}}$.

For the numerical integration let

$$X_t = (C_P, C_1, C_2, C_3, \Delta T_0, \Delta T_1, \Delta T_2)^T \tag{3.6}$$

$$\mathcal{A}(X_t, t) = \begin{pmatrix} a_0 E \\ a_1 E - \frac{1}{\tau_1} C_1 \\ a_2 E - \frac{1}{\tau_2} C_2 \\ a_3 E - \frac{1}{\tau_3} C_3 \\ b_0 (A\alpha \ln\left(\frac{C_P + C_1 + C_2 + C_3}{C_0}\right)) - \frac{1}{\tau_{b0}} \Delta T_0 \\ b_1 (A\alpha \ln\left(\frac{C_P + C_1 + C_2 + C_3}{C_0}\right)) - \frac{1}{\tau_{b1}} \Delta T_1 \\ b_2 (A\alpha \ln\left(\frac{C_P + C_1 + C_2 + C_3}{C_0}\right)) - \frac{1}{\tau_{b2}} \Delta T_2 \end{pmatrix}, \ \mathcal{B}(X_t) = \begin{pmatrix} 0 \\ 0 \\ \sigma_{C2} \\ 0 \\ \sigma_{T0} \\ 0 \\ \sigma_{T2} \Delta T_2 \end{pmatrix} \tag{3.7}$$

$$dX_t = \mathcal{A}(X_t, t)dt + \mathcal{B}(X_t)\ dW_t \tag{3.8}$$

We employ the strong order-1 Milstein scheme

$$X_{n+1} = X_n + \mathcal{A}(X_n, t_n)\Delta t + \mathcal{B}(X_n)\Delta W_n + \frac{1}{2}\mathcal{B}(X_n)\mathcal{B}'(X_n)\left((\Delta W_n)^2 - \Delta t\right) \tag{3.9}$$

$$\mathcal{B}'(X_t) = (0, 0, 0, 0, 0, 0, \sigma_{T2})^T \tag{3.10}$$

$$\mathcal{B}(X_t)\mathcal{B}'(X_t) = (0, 0, 0, 0, 0, 0, \sigma_{T2}^2)^T X_t \tag{3.11}$$

3.4. *Initial Conditions*

Initial conditions are straightforward for pre-industrial conditions where we simply set

$$C_{int} = [278.05158, 0, 0, 0]^T \tag{3.12a}$$

$$\Delta T_{int} = [0, 0, 0]^T \tag{3.12b}$$

that is, we assign the entire pre-industrial carbon budget to the permanent reservoir and set the remaining reservoirs to zero. This is justified as we assume that the pre-industrial conditions are in steady-state with $E = 0, C = C_0, \Delta F = 0$ on average and hence all non-permanent reservoirs eventually relax to zero. We may also want to start from a later date, such as the present. Then setting initial conditions is more complicated as available data on carbon and temperature history naturally only give total $C, \Delta T$ but not our intermediate budgets. We solve this problem by starting our model at pre-industrial conditions and run it without noise until a specific year with historical emissions†. They agree very closely with the RCP emission scenarios up to 2005 (from when on the scenarios differ). As we can see in Fig. 5 the temperature evolution of our model forced with historical emissions closely matches the observed temperatures‡, excluding internal variability. The reconstruction of CO_2 concentrations¶ is less good, mainly for three reasons. First, our model is forced by fossil emissions while for early times total CO_2 emissions are dominated by emissions due to land use change (Le Quéré *et al.* 2016) so we observe a substantial deviation. When approaching the present fossil fuel emissions begin to dominate the signal. Secondly, as elaborated above, our model is forced with CO_2 only, incorporating other GHG such as methane and nitrous oxide only via a constant factor. Third, the model is pulsesize-independent and tuned to the present-day. We hence find that our relation between CO_2 emissions and temperature evolution is apparently quite good but the intermediate CO_2 concentration – which we are not directly interested in – is fitted less well. We take the initial conditions for starting dates other than 1765 from the reference run (Tab. 4).

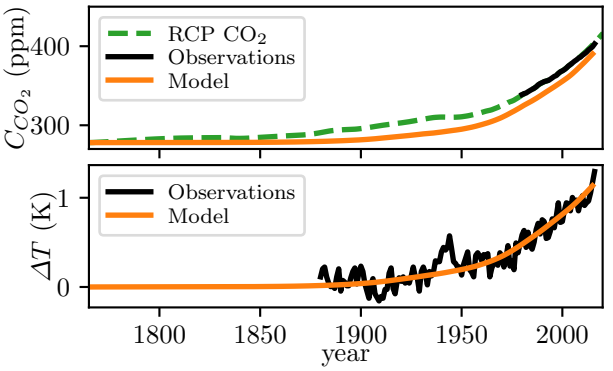

Figure 5: Reconstruction based on historical fossil CO_2 emissions

† historical emissions as fossil fuel and cement production emissions from (Le Quéré *et al.* 2016), accessed 28^{th} March, 2017

‡ historical GMST from `https://data.giss.nasa.gov/gistemp/graphs_v3/`, accessed 28^{th} March, 2017

¶ historical CO_2 concentrations from `https://www.esrl.noaa.gov/gmd/ccgg/trends/global.html`, accessed 28^{th} March, 2017

Year	C_P	C_1	C_2	C_3	ΔT_0	ΔT_1	ΔT_2
2000	306.251	26.833	18.914	3.579	0.074	0.296	0.444
2005	309.941	30.282	21.013	3.974	0.085	0.328	0.496
2015	319.281	39.048	26.719	5.216	0.110	0.408	0.617

Table 4: Initial conditions for several starting years

3.5. *Application to RCP Scenarios*

In order to test the quality of our stochastic model we start it in 1765 at pre-industrial conditions and run it forward, comparing with CMIP5 RCP results. We show PDFs and 2σ envelopes in Fig. 6. We find that the model is well able to reproduce the CMIP5 RCP behavior under the different forcing scenarios. As these are very different in terms of rate of change and total cumulative emissions this is not a trivial finding. In addition RCP2.6 contains substantial negative emissions, responsible for the downward trend that the model correctly reproduces. The mean response for RCP8.5 is underestimated, likely because the carbon model is pulsesize-independent and hence carbon sinks are too strong at high concentrations. Also note that the uncertainty in the carbon cycle plays a rather minor role compared to the temperature model (Fig. 6, compare third and fourth row). The temperature perturbation ΔT is very closely lognormally distributed while for weak forcings (RCP2.6, RCP4.5 to some degree) the distribution is approximately normal. The CO_2 concentration is normally distributed in all cases.

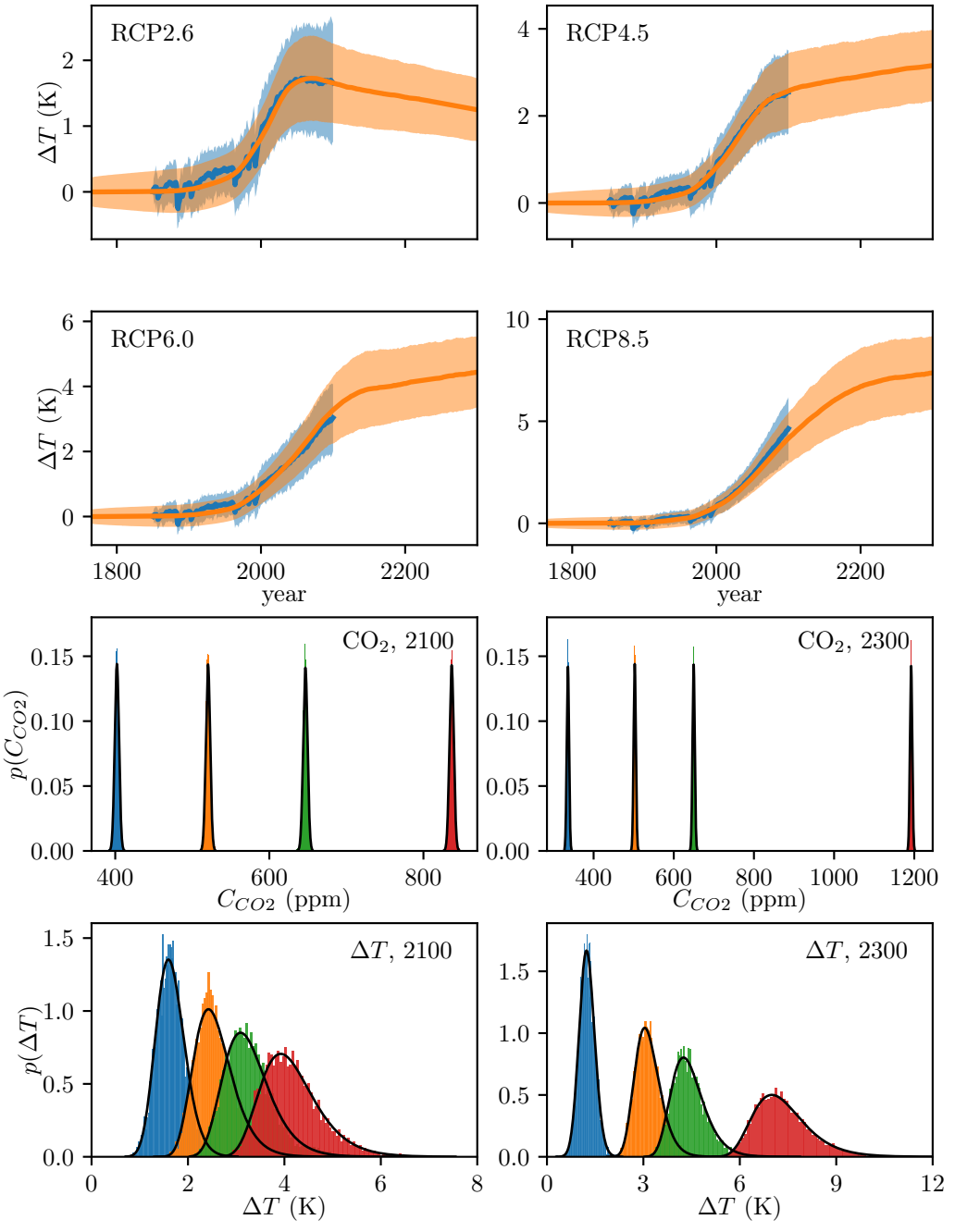

Figure 6: Stochastic State Space Model applied to RCP scenarios.
Upper two rows: Ensemble mean and 2σ envelopes of CMIP5 RCPs (blue) and stochastic model (orange).
Lower two rows: Probability density functions for CO_2 and ΔT in the years 2100 and 2300 based on 5000 ensemble members, and driven by forcing from RCP2.6 (blue), RCP4.5 (orange), RCP6.0 (green) and RCP8.5 (red). In black we fit normal distributions to CO_2 and lognormal distributions to ΔT.

4. Physical Transition Constraints

With the simple climate model now available we address the question of transitioning from the present-day (2015) to a carbon-free era in such a way as to avoid catastrophic climate change.

While many criteria can be formulated (e.g. on carbon concentration, sea level rise etc.) and in reality the impacts (extreme event distributions, flooding due to sea level rise, desertification, shifts in precipitation patters etc.) will likely strongly vary geographically, the most common global criteria that are easiest to communicate are in terms of GMST thresholds, such as done at the Paris Climate Conference in 2015, in particular the 2 K target, understood as necessary to avoid large and irreversible damages, including possible tipping points, and the 1.5 K target, formulated in the Paris agreement as the more desirable target. Questions have been raised on whether the commitments are sufficient to reach either target (Rogelj *et al.* 2016).

We have noted the issue of model errors and climate uncertainty due to inherent chaos and randomness. Therefore the question is not only of the *target* threshold to reach but also the *risk* one is willing to take to exceed it. In this section we consider two questions:

(i) What is the maximum amount of cumulative CO_2 emissions that allows us to stay below the 1.5, and 2 K target, as a function of the risk we are willing to take?
(ii) What is the latest point in time where we have to start taking mitigating action in order to stay below either target with a chosen risk tolerance, as a function of the aggressiveness of the action we deem feasible?

The answer to the first question gives us an emission target that must not be exceeded. We label this the Safe Carbon Budget (SCB), and it is a function of temperature target and risk tolerance. The second question yields a year in which it is too late to take action. This we call the Point of No Return (PONR), and it is a function of temperature target, risk tolerance and the aggressiveness of the (emission reduction) strategies available.

We call the temperature target T_{max} as the maximum warming allowable and describe the risk tolerance by the parameter β, the probability of staying below a given target. For example, $T_{max} = 2\,\mathrm{K}, \beta = 0.9$ corresponds to a 90% probability of staying below 2 K warming, i.e. 90 out of 100 realizations of the stochastic model, started in 2015 and integrated until 2100, do not exceed $\Delta T = 2\,\mathrm{K}$ in the year 2100.

For the PONR we need to define emission reduction strategies, so we require a model on how CO_2 emissions E are going to evolve in the future. Let the world economy be described by

$$Y = Y_0 e^{gt} \tag{4.1a}$$

$$En = \gamma_0 e^{-r_\gamma t} Y \tag{4.1b}$$

$$E = (1-a)(1-m)En \tag{4.1c}$$

with Gross World Product (GWP) Y with present-day value Y_0 and GWP growth rate g, global energy use En with initial energy efficiency γ_0 and rate of change r_γ and crucially, abatement and mitigation rates a and m. Parameter values are given in Tab. 5. The mitigation rate describes the share of non-GHG producing forms of energy production (wind, solar, water etc.), while the abatement rate gives the share of fossil energies the

emissions of which are not released, but captured and stored away (Carbon Capture and Storage (CCS)).

Y_0 (T\$)	g (yr^{-1})	γ_0 ($\mathrm{tC}\,\$^{-1}$)	r_γ (yr^{-1})
73	2%	1.4×10^{-4}	0

Table 5: Economic Parameters for Emissions Scenarios

4.1. *Safe Carbon Budget*

Using the full stochastic model we integrate an 8000-member ensemble for each of 6000 emission scenarios† starting in 2015. We compute the temperature distribution in 2100 ΔT_{2100} as a function of cumulative CO_2 emissions E_Σ. In Fig. 7 we show the 50^{th}, 67^{th}, 90^{th} and 95^{th} percentiles of ΔT_{2100} as function of E_Σ. In black the same

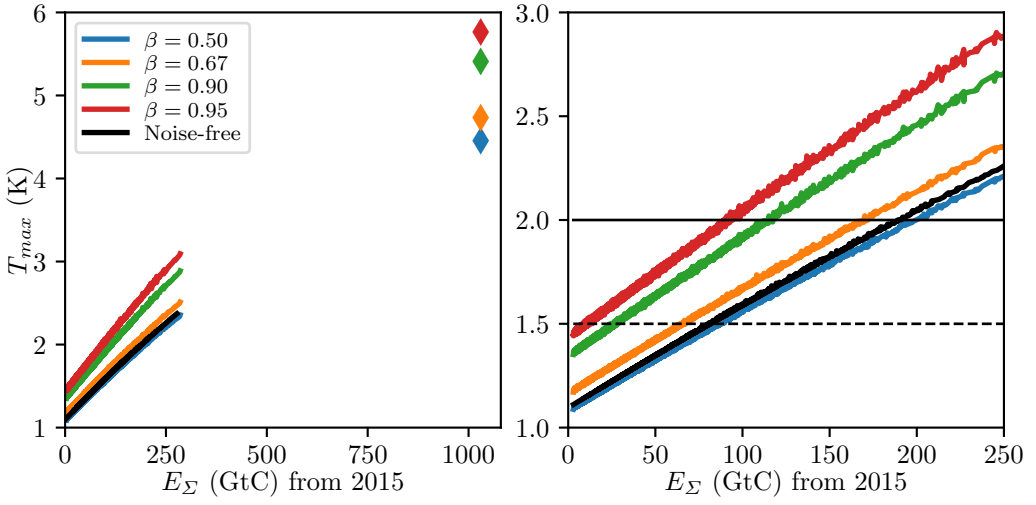

Figure 7: T_{max} in 2100 such that $p(\Delta T_{2100} \leqslant T_{max}) = \beta$ as a function of cumulative emissions for different β. Black curve gives deterministic results with noise terms set to zero. Zoom on the right with the 2 K (solid) and 1.5 K targets (dashed) as horizontal lines. Diamonds give 'business as usual' with $a(t) = 0, m(t) = 0.14$. Note in that case there is still substantial warming 'in the pipeline' as $E(t = 2100) > 0$.

calculation for the deterministic case without stochasticity is shown. It gives the certain relationship between cumulative emissions and warming in 2100 in a purely deterministic climate and approximately corresponds to the ensemble mean warming. Take a moment to realize how vastly higher 'business as usual' emissions‡ and warmings are, giving a taste for the scale of the problem.

† These scenarios are generated from Eq. 4.1*c* by letting $a(t) = 0$ and $m(t) = \min(1, m_0 + m_1 t)$, with $0 \leqslant m_0 \leqslant 0.7$ and m_1 drawn from a beta distribution ($\alpha = 1.2, \beta = 3$), with the [0,1] interval scaled such that $m = 1$ latest in 2080.

‡ We here define 'business as usual' by keeping abatement and mitigation at current values $a = 0, m = 0.14$ and using Eq. 4.1 to compute emissions.

The relations $\Delta T_{2100} = f(E_\Sigma)$ can be very well described by expressions of the type

$$\Delta T(E_\Sigma) = a \ln\left(\frac{E_\Sigma}{b} + 1\right) + c \tag{4.2}$$

with suitably fit coefficients a, b, c. Indeed, for this range of emissions a linear fit would be reasonable, as done by Allen *et al.* (2009). However our expression also works for cumulative emissions in the range of business as usual (when fitting parameters on suitable emission trajectories, not shown). It is straightforward to find the allowable emissions (the Safe Carbon Budget) for any combination of T_{max}, β from Fig. 7, as shown in Tab. 6. Note the drastic reduction in allowable emissions when enforcing the

β	0.5	0.67	0.9	0.95	Noise-free
$T_{max} = 1.5\,\mathrm{K}$	86.67	65.56	26.66	10.49	81.21
$T_{max} = 2.0\,\mathrm{K}$	199.03	169.64	115.13	92.29	189.53

Table 6: Safe Carbon Budget (in GtC since 2015) as function of threshold and safety probability.

target with a higher probability (following the horizontal lines from right to left in Fig. 7), or when moving from the deterministic, noise-free calculation to the stochastic situation with high β. This effect is particularly drastic for the 1.5 K target where the allowable emissions go essentially to zero for $\beta = 0.95$, given that emissions in 2015 where in the order of 9.9 GtC (Le Quéré *et al.* 2016). This raises grave doubts about the practical – or even theoretical – feasibility of the 1.5 K target.

We address the sensitivity of the Safe Carbon Budget to the relevant model parameters by varying each one by $\pm 10\%$ and running the calculation to see how the obtained value (for each scenario) changes (results shown in appendix C). The biggest effects on the SCB are found for the initial condition of the large carbon reservoirs and the radiative forcing parameters A, α, C_0 that are essentially fixed constants. The parameters of the carbon model (a_i, τ_i) do not have big impacts on the found SCB, on the order of $0 - 8$ GtC, with the larger numbers found for larger absolute values of SCB, limiting the relative effect. Also, possible variations in these parameters are likely not independent, potentially canceling each other.

Varying the temperature-model parameters can have quite noticeable effects, exceeding 20% for large and exceeding 50% for small values of SCB. The model is particularly sensitive to changes in the intermediate timescale (b_2, τ_{b2}). Fortunately the sensitivity to the long timescale, which is less well constrained, is much smaller. Similar to the carbon model, variations here are likely to offset each other to some degree.

The sensitivity of SCB to the noise amplitudes is small. It is largest for the multiplicative noise amplitude that is responsible for much of the spread of the temperature distribution (so increasing σ_{T2} decreases the SCB) but smaller than the climate parameter sensitivities.

4.2. *Point of No Return*

The Point of No Return (PONR) gives the the first year where it is no longer possible to reach a defined 'Safe State'. For our purposes, and similar to studies like van Zalinge *et al.* (2017) we define this state in terms of the probability distribution of GMST $p(\Delta T)$ in 2100:

Definition 1 (Point of No Return). *The Point of No Return (PONR) is the time t_P from which on no allowed $[a(t), m(t)]$ such that $0 \leqslant a(t), m(t) \leqslant 1, t_P \leqslant t \leqslant t_f$ can be chosen to fulfill*

$$p(\Delta T(t_f) \leqslant T_{max}) \geqslant \beta \tag{4.3}$$

Clearly there is a close connection between the Point of No Return and the Safe Carbon Budget. Indeed, one could define a Point of No Return also in terms of the ability to reach the Safe Carbon Budget. As there is a one-to-one relation between cumulative emissions and warming the Point of No Return in "Carbon space" has been defined in section 4.1. Its location in time however depends crucially on how fast a transition to a carbon-neutral economy is feasible, i.e. the meaning of the word "allowed" in the definition.

We resort to three crude but illustrative choices to model the abatement and mitigation rates:

(i) Extreme Mitigation (EM): At time $t_a > t_0$, we set $m = 1$, resulting in $E = 0$ from then onward. This corresponds to the most extreme physically possible scenario and therefore serves as an absolute upper bound – essentially we drop the restriction "allowed" in the PONR definition.

(ii) Fast Mitigation (FM): From time t_a onwards, both a, m increase by 0.05 per year. This results in zero emissions in less than 20 years.

(iii) Ambitious Mitigation (AM): As FM, but the increase is 0.02 per year, which is less extreme but still substantially faster than currently realized.

In the meantime, we are being pessimistic and assume that for $t_0 < t < t_a$ we remain with current values $m = 0.14$ and $a = 0$. In Fig. 8 we show probabilities for staying below the 1.5 and 2.0 K thresholds in 2100 as a function of t_a for EM, FM and AM, respectively. We can see that the remaining 'window of action' is small, but exists. For example, we can still reach the 2 K target with a probability of 67% when starting Fast Mitigation in the year 2035. However, reaching the 1.5 K target appears unlikely and is even impossible when requiring a high safety probability. The Point of No Return for the different targets and probabilities is given in Tab. 7, showing how it is close (or even already passed) for low thresholds and/or high safety requirements. This presents a major result of this work. We see several things here. First, the PONR is *close*. It is latest in the 2040s, and that is only when allowing for setting all CO_2 emissions to zero instantaneously. The more relevant numbers lie in the 2030s or earlier, and considering the speed of large-scale political and economic transformations, urgent action is warranted. Second, the PONR is *passed* for the 1.5 K target. We would have five years to start the extremely ambitious FM policy to avoid 1.5 K – and that only with 67% probability. Third, note the shifts for changing T_{max}, β and the policy. Switching from 1.5 to 2 K buys an additional 15 years. Allowing a one-third, instead of one-tenth exceedance risk, buys an additional seven years. Allowing for the more aggressive FM policy instead of AM buys an additional ten years. Note that these timeframes are not

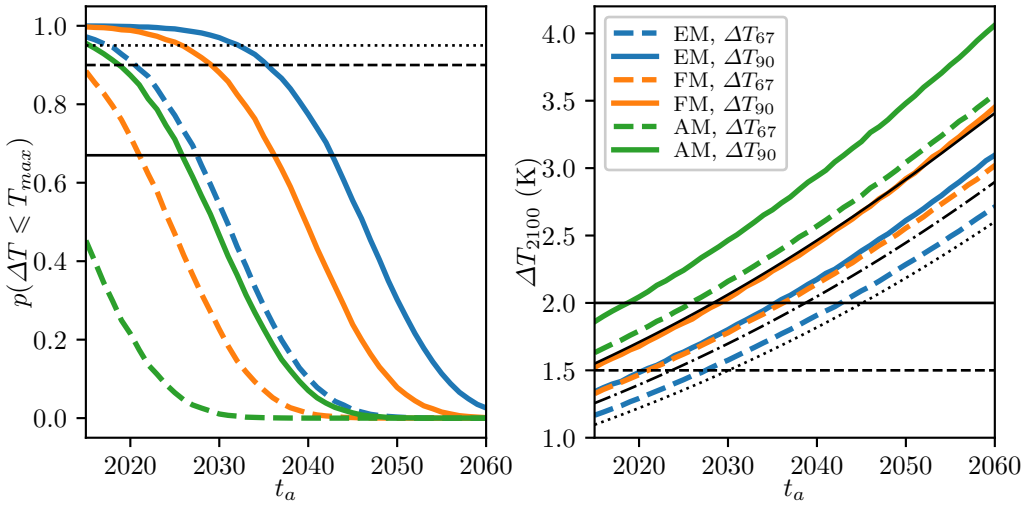

Figure 8: **Left:** Probability of staying below 'safe' temperature thresholds in 2100 when starting measures in year t_a for EM (blue), FM (orange), AM (green) with targets $T_{max} =$ 2 (solid) and $T_{max} = 1.5$ (dashed). Horizontal lines give $p = 0.67$ (solid), 0.90 (dashed) and 0.95 (dotted).
Right: 67^{th} and 90^{th} temperature percentiles in 2100 when starting measures in year t_a. Horizontal lines give $\Delta T_{2100} = 2.0\,\mathrm{K}$ (solid) and 1.5 K (dashed). In thin black lines we show the temperature in 2100 in the deterministic case of EM (dotted), FM (dash-dot) and AM (solid).

β	0.5	0.67	0.9	0.95	noise-free
EM					
$T_{max} = 1.5\,\mathrm{K}$	2031	2028	2021	2018	2031
$T_{max} = 2.0\,\mathrm{K}$	2047	2043	2036	2033	2046
FM					
$T_{max} = 1.5\,\mathrm{K}$	2025	2022	–	–	2024
$T_{max} = 2.0\,\mathrm{K}$	2040	2037	2030	2026	2039
AM					
$T_{max} = 1.5\,\mathrm{K}$	–	–	–	–	–
$T_{max} = 2.0\,\mathrm{K}$	2030	2026	2019	2016	2029

Table 7: Point of No Return as function of threshold and safety probability.

very large compared to political timescales. A two-term government delaying action could easily increase the exceedance risk from 10% to 30%.

As in case of the SCB, we address the sensitivity of the PONR values to the model parameters by varying each one by ±10% (results are shown in appendix D). The sensitivities are generally small and in no way change our message qualitatively. The effect of initial conditions and carbon model parameters in small, often even unnoticeable (with the exception of the permanent carbon reservoir, due to its large size). As for the case of the SCB we find large sensitivities to the fixed radiative forcing parameters

A, α, C_0. We find the most relevant effects of changes in the temperature model parameters. For example, a 10% error in τ_{b2} can move the PONR by 3-4 years. An interesting effect is the case of r_γ, the energy-saving progress (reduction in energy-intensity of a unit of GWP) which is zero by default. Increasing it to 1% or 2% has little or no effect on *close* PONR (e.g. 2020) but is capable of delaying *late* PONR by up to 11 years, and the effect is more substantial for the less ambitious scenarios. This is an interesting finding in itself, showing that in the long run increasing energy efficiency can play an important role in avoiding the PONR.

4.3. *Effects of Stochasticity in GWP Growth Rate g*

For the PONR computation we have used the very simple economic assumption of constant output growth of $2\,\%\,\mathrm{yr}^{-1}$. It is worth asking whether this is not an oversimplification and if uncertainty in future economic growth, i.e. stochasticity in the economic model, might change our assessment. We test this by using, instead of Eq. 4.1,

$$dg = -(g - g_0)dt + \sigma_g dW_t \tag{4.4a}$$

$$dY = gY dt \tag{4.4b}$$

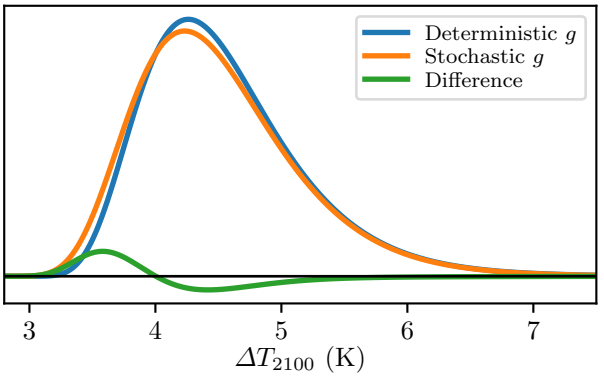

Figure 9: Distribution of ΔT in 2100 assuming constant g (blue) and stochastic g according to equations (4.4)

with $g_0 = 2\,\%\,\mathrm{yr}^{-1}, \sigma_g = 1.5\,\%\,\mathrm{yr}^{-1}$ and initial values g_0, Y_0. This introduces substantial uncertainty in the state of the global economy, and thereby also in the fossil fuel emissions. Taking a 'business-as-usual' scenario with $a = 0, m = 0.14$ we find for global output Y in the year 2100 a mean value of $\mu_Y = 397\,\mathrm{T\$}$ and standard deviation $\sigma_Y = 54\,\mathrm{T\$}$. We show the distribution of ΔT in 2100 in Fig. 9. Clearly, the GMST distribution turns out to be very insensitive to the noise in g. This is of course a very encouraging finding that increases confidence in our results.

4.4. *Using statistics over the entire time horizon*

Another point worth mentioning is the fact that here we have focused exclusively on the warming state *in 2100*. Our values for SCB and PONR do not ensure that the climate is in a 'safe' state until then (overshoots) or afterwards (long-term adjustment). As our emission scenarios generally include very rapid emission reduction we do find overshoots which may occur before 2100. Also the long timescales in our model lead to adjustments on timescales of several centuries (we do not put a lot of confidence in those long-term results as the model parameters are not tuned to such cases and hence badly constrained).

It is easily possible to change the definitions above and enforce non-exceedance of temperature thresholds for *all* timesteps. This leads to values as shown in Tab. 8 and reduces the Safe Carbon Budgets by $15 - 20\,\mathrm{GtC}$. It does not substantially change the

picture drawn so far. The temperature exceedances induced by using Tab. 6 instead of Tab. 8 are generally small, just the probabilities are quite sensitive, especially at small emissions/warmings where the PDFs of ΔT are relatively narrow.

β	0.5	0.67	0.9	0.95	Noise-free
$T_{max} = 1.5\,\mathrm{K}$	70.51	49.79	10.84	–	64.57
$T_{max} = 2.0\,\mathrm{K}$	178.31	151.81	97.44	72.71	173.87

Table 8: Safe Carbon Budget (in GtC since 2015) as function of threshold and safety probability when ensuring no-exceedance between 2015 and 2300.

5. Optimal Transitions

In the previous section we have used our physical climate model, and some very basic economic assumptions. We now want to go a step further and determine what would be the *optimal* (i.e. welfare-maximizing) pathways to reduce emissions such that warming is constrained below dangerous levels, which we phrase in terms of the SCB.

In the following we give an introduction in the economic theory we are going to use, and proceed to determine and discuss optimal transition pathways for two different formulations of climate damages.

5.1. *Economic Background*

We model the world economy in the style of a very simplified Integrated Assessment model (IAM, such as DICE (Nordhaus 1991), see also Dietz and Stern (2015); Rezai and van der Ploeg (2016); Rezai and van der Ploeg (2017)). Global economic output Y, energy production En and emissions E follow Eq. 4.1, that is we keep the growth rate g of global output constant as this simplifies the analysis and no large effect of a stochastic g on ΔT was found (section 4.3). Global consumption as fraction c of GWP is given as

$$c = \frac{C}{Y} = (1 - D)(1 - \Lambda) \tag{5.1}$$

that is, consumption equals output if both climate damages D and energy-related costs Λ vanish, and both act to reduce consumption. The cost Λ associated with the different types of energy generation, here differentiated between 'normal' fossil energy production, fossil energy with abatement, and renewable energies, is expressed as

$$\Lambda = \Big[\left(G_0 e^{-r_E t} + \frac{1}{\theta_a} A_1 e^{-r_A t} a^{\theta_a} \right) (1 - m) + H_0 m + \frac{1}{\theta_m} m^{\theta_m} H_1 e^{-r_R t} \Big] \gamma_0 e^{-r_\gamma t} \tag{5.2}$$

with the term in round brackets giving, for the non-mitigated energy fraction, the costs of fossil fuel extraction and abatement while the remaining summands give the cost of renewable energy usage. We see the main feature that marginal costs increase with abatement and mitigation but decrease over time (for positive r_A, r_R), assuming that technological progress will continue to reduce the cost of renewable technologies (as is already being observed (Farmer and Lafond 2016)).

Progressing climate change is associated with damages such as those brought about by heat waves, droughts and sea level rise. There exist several expressions in the literature

to parameterize these damages, such as letting them depend on the carbon stock (D_{carbon}), having the benefit of simplicity, or using a more or less convex function of global warming (D_{temp}) (Dietz and Stern 2015). For the temperature damage function we use the expression by Ackerman and Stanton (2012) that introduces substantial convexity to represent the high damages associated with very large warmings, and show both possible formulations in Eq. 5.3.

$$D_{carbon} = d(C - C_0) \qquad (I) \qquad (5.3a)$$

$$D_{temp} = (1 + \zeta_1 \Delta T^{\zeta_2} + \zeta_3 \Delta T^{\zeta_4})^{-1} \qquad (II) \qquad (5.3b)$$

Our aim is then to maximize global discounted welfare W as a function of consumption

$$\max_{a(t),m(t)} W = \int_0^\infty u(c(t)) Y e^{-SDRt} dt \qquad (5.4)$$

$$u(c(t)) = \frac{c(t)^{1-IIA}}{1 - IIA} \qquad (5.5)$$

with utility function of global consumption $u(c)$, intergenerational inequality aversion IIA, and social discount rate $SDR = \rho + (IIA - 1)g$ with rate of time patience ρ †. The benefit of the simple carbon-based damage function is that much of the optimization problem can be solved analytically, yielding direct expressions for abatement and mitigation fractions

$$a(t) = \left(\frac{P(t) e^{r_A t}}{A_1} \right)^{\epsilon_a} \quad , \; 0 \leqslant a(t) \leqslant 1 \qquad (5.6a)$$

$$m(t) = \left(\frac{G_0 e^{-r_E t} + \frac{1}{\theta_a} A_1 e^{-r_A t} a(t)^{\theta_a} + (1 - a(t)) P(t) - H_0}{H_1 e^{-r_R t}} \right)^{\epsilon_m} \quad , \; 0 \leqslant m(t) \leqslant 1 \qquad (5.6b)$$

where the carbon price P (the price to be paid for emitting carbon) is a constant ratio of GWP

$$P(t) = \tau^{U,C} Y(t) \qquad (5.7)$$

with τ a free parameter that needs to be determined based on the optimization target. We differentiate "unconstrained" and "constrained" optimization. In the former case τ is simply chosen such that welfare according to Eq. 5.4 is maximized, irrespective of the warming this entails. We label this as τ^U. In the latter case welfare is maximized *under the constraint that* warming is kept below some warming with some probability, that is, a chosen SCB (section 4.1) is not exceeded. In this case we label the found value τ^C. It is easy to see that we have $\tau^U \leqslant \tau^C$ as the climate-combating efforts get stricter with increasing P and the constrained policy will be *at least as strict* as the unconstrained one. The equality applies when the unconstrained optimization leads to cumulative emissions that undercut the SCB.

† This formulation for $u(c)$ is called a CRRA (constant relative risk aversion) social utility function, where we have the *index of relative risk aversion* RRA equal to IIA.

Therefore, for each time step t_i in the carbon-temperature model we have

$$D(C) = 1 - d(C - C_0) \tag{5.8a}$$

$$Y = Y_0 e^{g t_i} \tag{5.8b}$$

$$En = \gamma_0 e^{-r_\gamma t_i} Y \tag{5.8c}$$

$$P = \tau^{U,C} Y \tag{5.8d}$$

$$a = \left(\frac{P e^{r_A t_i}}{A_1} \right)^{\epsilon_a} \tag{5.8e}$$

$$m = \left(\frac{G_0 e^{-r_E t_i} + \frac{1}{\theta_a} A_1 e^{-r_A t_i} a^{\theta_a} + (1-a)P - H_0}{H_1 e^{-r_R t_i}} \right)^{\epsilon_m} \tag{5.8f}$$

$$E = (1-a)(1-m)En \tag{5.8g}$$

E is then fed into the climate model defined in section 3.3. The entire model with climate and economic components has $\tau^{U,C}$ as its only free control parameter.

For a start in 2015 we let $Y_0 = 74\,\mathrm{T\$}$, $E(2015) = 9.9\,\mathrm{GtC}$ (see (Le Quéré *et al.* 2016)) and an assumed $a = 0, m = 0.1401$ (World Energy Council 2016) which leads to

$$\gamma_0 = \frac{1}{(1-a)(1-m)} \frac{E}{Y} = 1.556 \times 10^{-4}\,\mathrm{tC\,\$^{-1}} \tag{5.9}$$

We show the remaining economic parameters in Tab. 9.

g (yr^{-1})	r_γ (yr^{-1})	r_A (yr^{-1})	r_E (yr^{-1})	r_R (yr^{-1})	A_1 ($\$\,\mathrm{tC}^{-1}$)
0.02	0	0.0125	-0.001	0.0125	2936
θ_a	ϵ_a	θ_m	ϵ_m	H_0 ($\$\,\mathrm{tC}^{-1}$)	H_1 ($\$\,\mathrm{tC}^{-1}$)
2	$1/(\theta_a - 1) = 1$	2.8	$1 - /(\theta_m - 1) = 0.56$	515	1150
d (tC^{-1})	ζ_1	ζ_2	ζ_3	ζ_4	G_0 ($\$\,\mathrm{tC}^{-1}$)
1.9×10^{-14}	0.00245	2	5.021×10^{-6}	6.76	515

Table 9: Economic model parameters

5.2. *Optimization with Carbon Damage Function*

Starting with the carbon damage function, we solve for $\tau^{U,C}$ and show the solutions for the unconstrained and constrained optimization in Fig. 10.

5.2.1. *Unconstrained Optimization*

We find τ^U by maximizing welfare without any constraints on the allowed cumulative emissions. We obtain

$$\tau^U = 1.5556\,\$\,\mathrm{T\$}^{-1} \tag{5.10}$$

The mitigation fraction reaches $m = 1$ in 2091, from when on emissions are zero. CO_2 concentrations peak in 2080 at 501 ppm and global warming rises to 2.46 K (90^{th} percentile of 2.76 K), beyond levels considered 'safe', as caused by cumulative emissions of 269.8 GtC. The initial carbon price is 115.12 $ and reaches 630.15 $ by 2100.

5.2.2. *Constrained Optimization*

Clearly, simply maximizing welfare does not limit warming to 2 K or even 1.5 K. Therefore, we employ the additional constraint of limiting warming in 2100 to some maximum value T_{max} with prescribed probability β, that is we constrain cumulative emissions E_Σ to the SCB. This being a purely deterministic quantity simplifies our investigation. We find the value of τ^C that maximizes welfare under the constraint $E_\Sigma \leqslant E_{safe}$ (Tab. 10). As an example we may wish to stay below $T_{max} = 2\,\mathrm{K}$ warming with a probability of $\beta = 90\,\%$. This gives $SCB = 115.13\,\mathrm{GtC}$ (Tab. 6), and $\tau^C = 3.34\,\$\,\mathrm{T}\$^{-1}$ (Tab. 10). We start with a carbon price of 287 \$ and the fossil fuel era ends in 2066. Our economic formulation assumes that economic actors will instantly adapt to any carbon price. This is why the initial mitigation ratio is already 45% as the carbon price starts so high that many actors immediately switch to renewable energies. It may be questioned whether this is a realistic assumption, and indeed how fast a transition is feasible. We will address this in a following section.

β	0.5	0.67	0.9	0.95	Noise-free
$T_{max} = 1.5\,\mathrm{K}$	4.10	4.91	7.78	10.84	4.28
$T_{max} = 2.0\,\mathrm{K}$	2.10	2.43	3.34	3.93	2.20

Table 10: Optimal value for τ^C (in $\$\,\mathrm{T}\$^{-1}$) as function of threshold and safety probability.

We examine the results for the constrained case, noise-free and with $\beta = 0.9$, for $T_{max} = 2.0\,\mathrm{K}$ and $T_{max} = 1.5\,\mathrm{K}$, together with the unconstrained case, in Tab. 11 and the left column of Fig. 10. We see large differences between the scenarios. In particular the 'unconstrained' optimal transition is too slow to prevent high warming, and the inclusion of stochasticity in the climate tightens the budget and calls for a faster, more ambitious transition.

	max CO_2 (ppm)	E_Σ (GtC)	t_f (yr)
unconstrained	501.29	269.84	2091
constrained, $T_{max} = 2\,\mathrm{K}$, $\sigma_i = 0$	462.68	189.53	2082
constrained, $T_{max} = 2\,\mathrm{K}$,$\sigma_i \neq 0$	427.71	115.13	2070
constrained, $T_{max} = 1.5\,\mathrm{K}$,$\sigma_i = 0$	411.43	81.21	2063
constrained, $T_{max} = 1.5\,\mathrm{K}$,$\sigma_i \neq 0$	390.27	26.66	2045
	$P(t = 2015)$ (\$)	$a(t = 2015)$	$m(t = 2015)$
unconstrained	115.12	0.039	0.275
constrained, $T_{max} = 2\,\mathrm{K}$, $\sigma_i = 0$	162.93	0.055	0.332
constrained, $T_{max} = 2\,\mathrm{K}$, $\sigma_i \neq 0$	247.41	0.084	0.416
constrained, $T_{max} = 1.5\,\mathrm{K}$, $\sigma_i = 0$	317.06	0.108	0.474
constrained, $T_{max} = 1.5\,\mathrm{K}$, $\sigma_i \neq 0$	575.96	0.196	0.643

Table 11: Key numbers for different constraints, comparing unconstrained and constrained optimization, including stochasticity ($\sigma_i = (\sigma_C, \sigma_{T0}, \sigma_{T2}) \neq 0$) and without it ($\sigma_i = 0$). t_f gives the end of the fossil fuel era where $E = 0$.

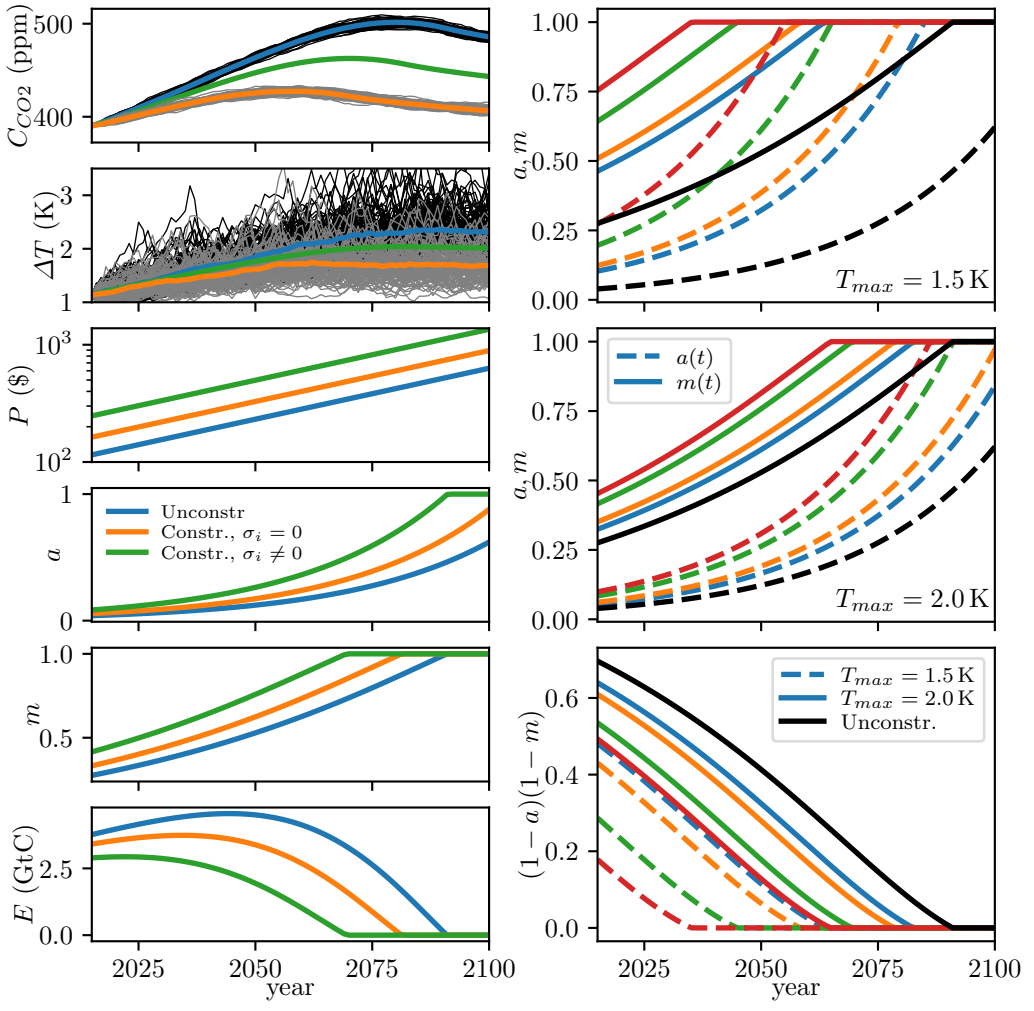

Figure 10: Optimal Transitions with carbon damage function
Left column: Trajectories for optimization in the unconstrained (blue) and constrained case without (orange) and with (green) stochasticity, for $T_{max} = 2\,\mathrm{K}, \beta = 0.9$.
Right column: Transition pathways m (solid) and a (dashed) with $\beta = 0.5$ (blue), 0.67 (orange), 0.9 (green) and 0.95 (red) for $T_{max} = 1.5$ (top) and $2\,\mathrm{K}$ (middle), and the unconstrained case in black. The bottom panel shows the instantaneous emission reduction $(1-a)(1-m)$ for $T_{max} = 1.5$ (dashed) and $2\,\mathrm{K}$ (solid)

5.3. *Optimization with Temperature Damage Function*

After showing the results for the carbon damage function we now move to the more realistic temperature damage function. The problem now is that we do not anymore have analytical expressions for $a(t), m(t)$ and instead have to solve the optimization problem numerically. In principle this enormously increases the number of optimization parameters from one (τ^U, τ^C) to $2N$ where N is the number of timesteps. We treat this

problem by using a basinhopping algorithm encapsulating the numerical optimizer†. The enormous parameter space makes finding the solution very slow and stops us from using very long time horizons to ensure full convergence of the integral in Eq. 5.4. The result as shown in the left column of Fig. 11 does appear reasonable, compared with results shown for the carbon optimization before (higher starting values and faster transition for stricter targets) but the solution is not perfect and does not look fully converged. Most likely the basinhopping algorithm has not been able to find the global optimum in the enormous parameter space.

Due to the problems of finding the global optimum in a parameter space of dimension well beyond 200, we introduce some simplifications. Prescribing some 'reasonable' shapes for the transition pathways drastically reduces the parameter space so we can be much more certain of having found the true optimum in the reduced space.

$$I : P(t) = P_0 e^{g_P t} \tag{5.11a}$$

$$II : a(t) = a_0 e^{g_a t} \quad , \quad m(t) = m_0 e^{g_m t} \tag{5.11b}$$

$$III : a(t) = a_0 + a_1 t \quad , \quad m(t) = m_0 + m_1 t \tag{5.11c}$$

$$IV : a(t) = a_0 e^{g_a t} + a_1 t \quad , \quad m(t) = m_0 e^{g_m t} + m_1 t \tag{5.11d}$$

For case I we assume that the change to the temperature damage function does not affect the relationship between P, and a, m (Eq. 5.6a, 5.6b), which appears reasonable for small warmings where damages according to D_{carbon} and D_{temp} are very similar, while in the cases II, III, IV we assume a, m to be independent. We can then diagnostically compute P by inverting Eq. 5.6a:

$$P(t) = A_1 e^{-r_A t} a(t)^{\theta_a - 1} \tag{5.12}$$

which obviously does not describe $P(t)$ after the time where $a = 1$.

We compare the welfare returned by paths $I - IV$ with the optimization in the large parameter space. It turns out that case IV yields the best welfare of these five cases. This appears surprising as the space is so much more restricted. However this likely confirms our suspicion that the global optimum had not yet been found in the large parameter space, and that the six-parameter parameterization was able to closely approximate the true optimum. We show the optimal result from case IV in the right column of Fig. 11.

Again tighter constraints mean earlier transitions. We notice that for the 2 K target the lines lie very close together. With the exception of $\beta = 0.95$ it is actually beneficial not to use up the entire budget. That is, the unconstrained optimization suggests such an ambitious pathway that it fulfills the $\beta = 0.9$ constraint. We see the big difference when comparing with the optimization using the carbon damage function in Fig. 10: Using the temperature damage function leads to a substantially more ambitious transition as the high-emission, high-warming trajectories are avoided due to more convex damages.

† We use the python library `scipy.optimize`

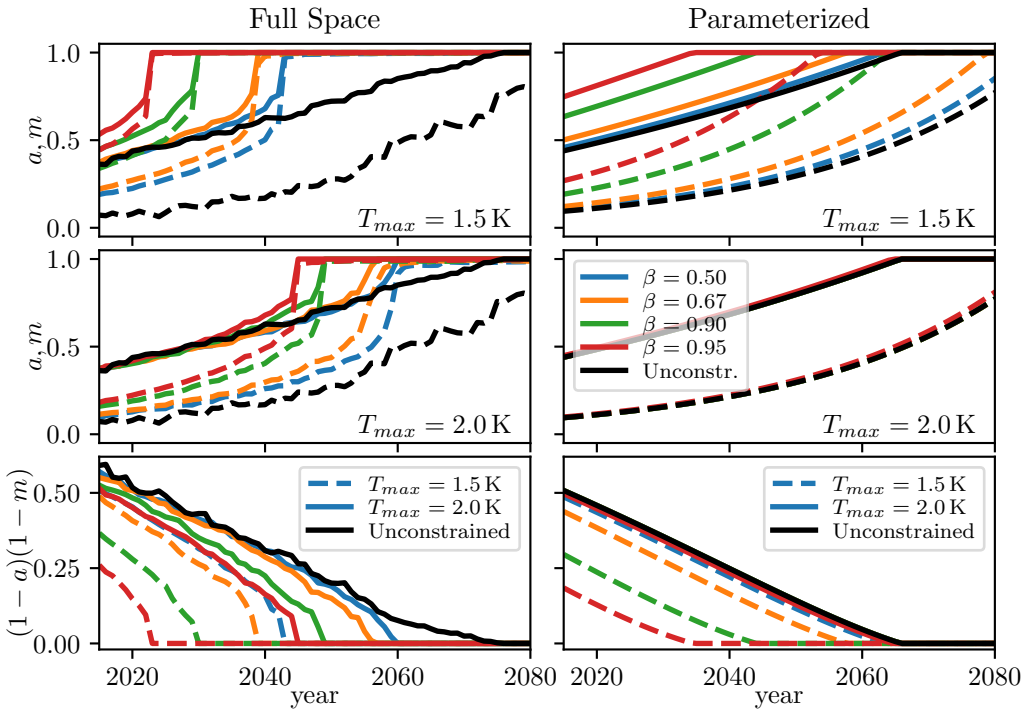

Figure 11: Optimal transition pathways a (dashed) and m (solid) when controlling $0 \leqslant a, m \leqslant 1$. **Left:** Optimization in large parameter space, **Right:** Optimizing $a_0, m_0, a_1, m_1, g_a, g_m$ according to equation (5.11*d*). Temperature thresholds 1.5 K (top) and 2.0 K (middle) and safety probabilities as given by the respective SCBs. The bottom row shows the instantaneous emission reduction due to abatement and mitigation $(1-a)(1-m)$.

5.4. *Controlling Rates of Transition*

In the previous section we determined optimal transition pathways that are welfare-maximizing while being constrained to the SCB. However, there remains a serious point of critique. We have let the *starting point* of the transition pathway free, i.e. we have chosen initial a, m purely based on the most optimal value in terms of welfare but not based on the question whether it is feasible. Under such assumptions *of course* there is always a possible transition pathway that fulfills a given (positive) SCB as we could simply start with $m = 1$. However, this scenario is not particularly realistic. For this reason, we briefly consider the situation where we do not control a, m directly but their *rate of change*.

Let us introduce two new state space variables for the abatement and mitigation rate

$$da = F_a \, dt \tag{5.13a}$$
$$dm = F_m \, dt \tag{5.13b}$$

Their only dynamic is being forced by an external control. By setting the initial values of a, m to present-day values and choosing the optimal F_a, F_m we get a better handle

on *feasible* transitions. We start by attempting to find the global optimum in the large parameter space with the same technique as before and by limiting $0 \leqslant F_a, F_m \leqslant 0.1$. This is of course a very high value but it is so different from the rates of change observed before that we should see a clear effect.

As before we also compute the transition in a reduced parameter space. Parameterizing the transition pathways is now a lot more difficult than before due to the sharp variations in the rate of change. We find optimal parameter values for F_a, F_m as described by

$$F_a(t) = A_a + B_a e^{-C_a t} + D_a e^{E_a t} \tag{5.14a}$$

$$F_m(t) = A_m + B_m e^{-C_m t} + D_m e^{E_m t} \tag{5.14b}$$

We constrain all these parameters according to Tab. 12.

$0 \leqslant A_a \leqslant 0.1$	$0 \leqslant B_a \leqslant 0.1$	$0 \leqslant C_a \leqslant 0.5$	$0 \leqslant D_a \leqslant 0.1$	$0 \leqslant E_a \leqslant 0.08$
$0 \leqslant A_m \leqslant 0.1$	$0 \leqslant B_m \leqslant 0.2$	$0 \leqslant C_m \leqslant 0.5$	$0 \leqslant D_m \leqslant 0.1$	$0 \leqslant E_m \leqslant 0.08$

Table 12: Constraints for F_a, F_m parameterization

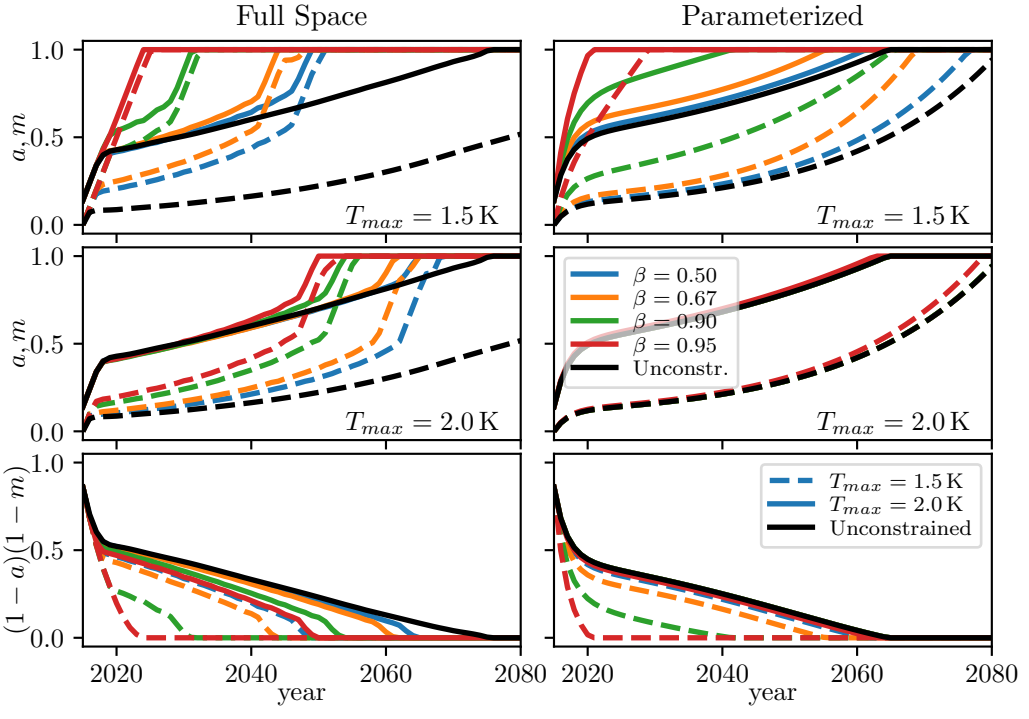

Figure 12: Optimal transition pathways a (dashed) and m (solid) when controlling $0 \leqslant F_a, F_m \leqslant 0.1$. **Left:** Optimization in large parameter space, **Right:** Optimizing F_a, F_m according to Eq. 5.14. Temperature thresholds 1.5 K (top) and 2.0 K (middle) and safety probabilities as given by the respective SCBs. The bottom row shows the instantaneous emission reduction due to abatement and mitigation $(1 - a)(1 - m)$.

In Fig. 12 we show the solutions for the full parameter space and parameterized transition. Comparing Fig. 12 with Fig. 11 shows that the transitions are qualitatively very similar. The pathways start from present-day values and converge as quickly as allowed to scenarios similar to the ones shown in Fig. 11. Computing the welfare of the two optimization results reveals that finding the global optimum again is easier in the parameterized case, as it performs better. The exception is the case of $T_{max} = 1.5\,\mathrm{K}, \beta = 0.95$, where the optimal path appears to be an increase at the fastest possible rate, which is higher in the full optimization.

The result shown in Fig. 12 serves as a good illustration of how the pathways in terms of F_a, F_m quickly approach those in terms of a, m with free the initial rates (Fig. 11). However, this very fast initial adjustment is not very feasible either, as the maximal rates of increase obtain the unrealistic value of 0.1. In Fig. 13 we therefore show a more realistic solution with stricter bounds on F_a, F_m. This has the consequence that for $T_{max} = 1.5\,\mathrm{K}, \beta = 0.9, 0.95$ it is no longer possible to meet the SCB, as the 'minimal possible' emissions are 50 GtC. Note how again for $T_{max} = 2\,\mathrm{K}$ the solutions for $\beta = 0.5, 0.67, 0.9$ and the unconstrained case coincide and yield $E_\Sigma = 107\,\mathrm{GtC}$. The overall shape is very similar to Fig. 12, excluding the fast initial adjustment. Fig. 13 represents our best guess for a feasible optimal transition. It reaffirms our conclusion from section 4.2 that reaching the 1.5 K target is extremely unlikely.

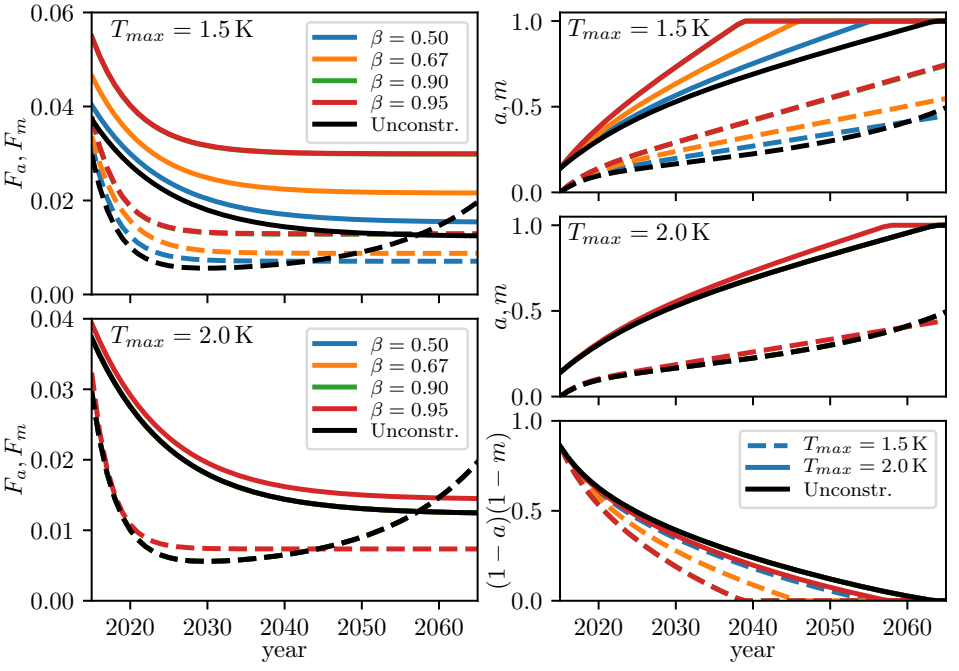

Figure 13: Optimal transition pathways when controlling F_a, F_m with more realistic bounds. All colors as in top left: $\beta = 0.5$ (blue), 0.67 (orange), 0.9 (green) and 0.95 (red) and unconstrained solution in black. **Left:** F_a (dashed) and F_m (solid) for $T_{max} = 1.5\,\mathrm{K}$ (top) and 2 K (bottom).
Right: a (dashed) and m (solid) for $T_{max} = 1.5\,\mathrm{K}$ (top) and 2.0 K (middle). The bottom right shows $(1-a)(1-m)$, with pathways for 1.5 K (dashed) and for 2 K (solid).

6. Conclusion and Outlook

Concluding the thesis, we take a step back and put our findings in a larger context. Primarily, we have done three things:

(i) We have shown that it is possible to apply Linear Response Theory to the CMIP5 ensemble, and built a simple model that gives a good representation of its statistical properties.

(ii) We have computed two simple, easy-to-communicate metrics with a solid physical underpinning and minimal assumptions, the Safe Carbon Budget and the Point of No Return, that capture warming targets and risk tolerance,

(iii) Under additional assumptions on economic dynamics, we have found possible optimal transition pathways to a carbon-free era in order to reach global warming targets.

In this context we have in particular pointed out the strong constraints that are put on future emissions by restricting GMST increase to below 2 K, and the near-impossibility of reaching the 1.5 K target with remotely realistic emission scenarios. An energy transition *substantially* more ambitious than assumed in RCP2.6 would be required to stay below 1.5 K with some robust probability, and whether that is feasible is highly doubtful. With all other RCP scenarios exceeding 2 K is virtually certain in the near future, as we can show by the instantaneous probability of exceeding the respective thresholds in Fig. 14. This is a sobering finding in light of the bold ambition in the Paris agreement, and adds to the sense of urgency to act quickly in the soon-to-be-closing window before the thresholds have been crossed.

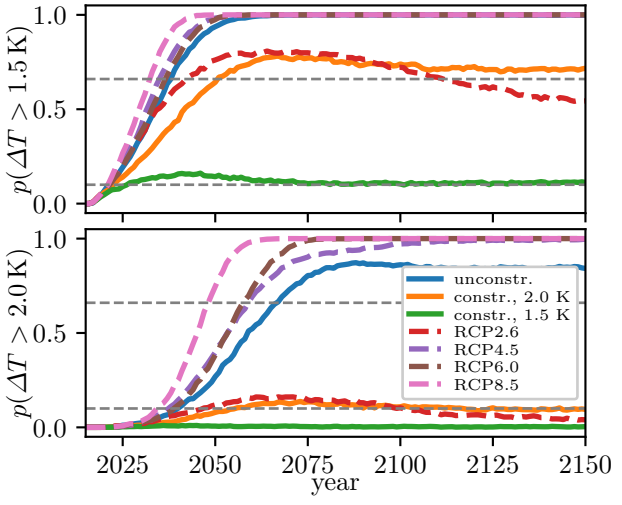

Figure 14: Instantaneous probability to exceed 1.5 K (top) and 2.0 K (bottom). RCP emission scenarios are shown as dashed lines while solid lines show the unconstrained (blue) and constrained cases (2 K (orange) and 1.5 K (green)) from section 5.2. Dashed horizontal lines give $p = 0.1$ and 0.67, respectively.

Our analysis heavily relies on the application of LRT to the CMIP5 ensemble and this has turned out to be very successful. Starting from an (not particularly large) ensemble of quite diverse models we were able to very well recover the ensemble mean and variance of the global warming signal. This is encouraging and calls for further research into the applicability of LRT, also to other variables available from CMIP5, such as sea level rise, sea ice cover, or measures of the atmospheric and oceanic circulation. It is likely though that – especially considering the rather small ensembles available – the results will be worse for less 'smooth' and well-behaved quantities. For example, Lucarini *et al.* (2016) applied LRT to spatial patterns of a 200-member ensemble of PLASIM. This appears out of reach for CMIP5 unless much larger ensembles should become available.

There are of course limitations to the model we have constructed. Our focus has been to accurately represent the basic statistical properties (mean and variance) of the CMIP5 RCP ensemble, in order to study overall warming probabilities as function of emissions. One could argue that we implicitly assume Gaussian PDFs as we only use the variance to judge the 'spread' (Fig. 6) we tune our stochastic model to. We justify this procedure by its good prediction of variance in the idealized experiments, the approximate gaussianity of ΔT for small warmings and until 2100†, and the fact that the small number of model runs make it difficult to justify fitting any particular non-Gaussian distribution. Of course, in reality the CMIP5 ensemble is most likely not normally distributed for high emissions scenarios (see e.g. Fig. 3). Indeed the models differ so much in their predictions for RCP8.5 beyond 2100 that one may classify them in different groups of models, with our methods applied separately, to get a 'high' and a 'low' response. In addition, we have focused much less on the physical realism of the individual realizations of the stochastic process as compared to the ensemble. As mentioned, the large multiplicative noise factor leads – especially at high mean warmings – to an immensely volatile trajectory that in all likelihood is not physical. It might be a worthy endeavor to investigate how this could be improved, for example to also recover higher-order statistics or timeseries properties that we did not cover here. It may be possible to combine various noise terms to obtain better individual timeseries. One might also involve further methods for timeseries analysis that we did not discuss here, such as delay differential equations or recurrent neural networks.

By our choice of model we limited ourselves to response functions represented as sums of exponentials. Several times we noted the issue already realized in Joos *et al.* (2013) that the real climate's carbon model is not pulse-independent and hence using a single constant response function has inherent problems, in particular when running very different emission scenarios, because the efficiency of the natural carbon sinks to the ocean and land reservoirs is a function both of temperature and the reservoir sizes. Therefore, expectedly, our model has problems reproducing CO_2 concentration pathways (Fig. 4), a price we accept to pay as we focus on the temperature reproduction. In a new work Millar *et al.* (2017) draw a different conclusion from studying a similar problem and introduce in their new FAIR model response functions that dynamically adjust parameters based on warming to represent sink saturation. Consequently, their model gives much better results in terms of CO_2 concentrations. It would be an interesting lead for future research to conduct our analysis here (in terms of SCB, PONR and optimal transitions) with other simple models (such as FAIR or MAGICC) to discover similarities and differences.

In our work we have used a simple economic framework mainly to

(i) describe how emissions progress from now into the future in a 'business-as-usual' scenario in order to derive the PONR, and to

(ii) formulate simple expressions for climate damages, costs associated with combating climate change, and global welfare to optimize for.

While the first item includes relatively few assumptions (mainly that economic growth and energy efficiency are approximately constant), the second item represents a simple view of the world with many obvious drawbacks. We have ourselves addressed the issue

† This is our main interest, and in fact, the 140-year idealized experiments hardly justify much longer projection horizons.

that it assumes the possibility of instantaneous adjustment to any carbon price, and thereby any level of abatement and mitigation. Also our model is formulated purely in terms of global output and does not consider for example capital stock, labor (Dietz and Stern 2015) and fixed investments in fossil fuel technology (such as coal power plants). We also do not address the issue of *how* the carbon price or the abatement/mitigation ratios could be set in practice. Common options would be carbon taxes and subsidies on renewable energies (Rezai and van der Ploeg 2016; Golosov *et al.* 2014) and have many inherent problems themselves (Green paradox, free-riding, heterogeneous actors etc.). In addition, we have also assumed that abatement and/or mitigation can indeed reach $a, m = 1$, leading to a completely emission-free world. Combined with only representing CO_2 in our model this is probably (over-)optimistic, as some emission sources (land use change, agriculture) cannot be reduced to zero. The relevant actors deciding on emissions, such as companies, do not perform the welfare maximization we have done here as climate damages are currently an unpriced externality. Thereby the economic problem becomes more a political problem of which countries (resp. their governments) have the political will to act and internalize these damages. Due to these drawbacks, the economic part of this thesis is more an illustrative exercise than a realistic projection of likely economic developments. We do stress, however, that many of these drawbacks do not apply to our physical-based metrics SCB and PONR.

Putting the message of this thesis in one sentence, we argue that if no fast action is taken the Earth is very likely to warm by more than two degrees, while 1.5 degrees of warming are already essentially unavoidable. If this is a vision to be avoided at all costs, it may be necessary to think about additional possibilities for action. The prolonged negative emissions present in the RCP2.6 scenario successfully decrease the GMST again. This hints at the potential applicability of negative emission (e.g. BECCS) or geoengineering (e.g. sulfate injections) technologies if mitigation efforts are judged to be insufficient. It would be interesting to think about how to include either into our framework and how to treat them economically. For example, one might ask the question by how much a particular geoengineering approach would delay the PONR.

Acknowledgements

There are many people that deserve a heartfelt thank you. First and foremost I want to thank Henk for introducing the topic to me and for being an endless source of enthusiasm, encouragement and support. Many thanks to Rick for his economic expertise, his creativity in coming up with new ideas, and a delicious coffee in a beautiful Oxford book shop. Qingyi always had an open door and mind to discuss this and other research. Many thanks to Femke for many physics- and non-physics-related coffee breaks. Over the years, many people have helped me do what I like to do, so I also want to thank my inspiring former academic and professional colleagues and supervisors, in particular Jürgen, Martin, Joachim, Robert, Brian, Jonathan, Jörg, Björn and Yuri. A great many more people have opened new perspectives, brought up new ideas or simply been supportive by being who they are. Thank you to you all. Finally, I thank my parents, Margret and Ulrich, for everything.

REFERENCES

Ackerman, F., and E. A. Stanton (2012), Climate Risks and Carbon Prices: Revising the Social Cost of Carbon, *Economics: The Open-Access, Open-Assessment E-Journal*, *6*(2012-10), 1, doi:10.5018/economics-ejournal.ja.2012-10.

Allen, M. R., D. J. Frame, C. Huntingford, C. D. Jones, J. A. Lowe, M. Meinshausen, and N. Meinshausen (2009), Warming caused by cumulative carbon emissions towards the trillionth tonne, *Nature*, *458*(7242), 1163–1166, doi:10.1038/nature08019.

Cai, Y., T. M. Lenton, and T. S. Lontzek (2016), Risk of multiple interacting tipping points should encourage rapid CO2 emission reduction, *Nature Climate Change*, *6*(5), 520–525, doi:10.1038/nclimate2964.

Clarke, L. E., J. A. Edmonds, H. D. Jacoby, H. M. Pitcher, J. M. Reily, and R. G. Richels (2007), Scenarios of Greenhouse Gas Emissions and Atmospheric Concentrations Synthesis, *Tech. rep.*, Department of Energy, Office of Biological & Environmental Research, Washington, D.C.

Crutzen, P. J. (2002), Geology of mankind, *Nature*, *415*(6867), 23–23, doi:10.1038/415023a.

Dietz, S., and N. Stern (2015), Endogenous Growth, Convexity of Damage and Climate Risk: How Nordhaus' Framework Supports Deep Cuts in Carbon Emissions, *The Economic Journal*, *125*(583), 574–620, doi:10.1111/ecoj.12188.

Dijkstra, H. A. (2013), *Nonlinear Climate Dynamics*, 529 pp., Cambridge University Press, Cambridge, doi:10.1017/CBO9781139034135.

Farmer, J. D., and F. Lafond (2016), How predictable is technological progress?, *Research Policy*, *45*(3), 647–665, doi:10.1016/j.respol.2015.11.001.

Fujino, J., R. Nair, M. Kainuma, T. Masui, and Y. Matsuoka (2006), Multi-gas Mitigation Analysis on Stabilization Scenarios Using Aim Global Model, *The Energy Journal*, *SI 2006*(01), 343–354, doi:10.5547/ISSN0195-6574-EJ-VolSI2006-NoSI3-17.

Fuss, S., *et al.* (2014), Betting on negative emissions, *Nature Climate Change*, *4*(10), 850–853, doi:10.1038/nclimate2392.

Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014), Optimal Taxes on Fossil Fuel in General Equilibrium, *Econometrica*, *82*(1), 41–88, doi:10.3982/ECTA10217.

Haustein, K., F. E. L. Otto, P. Uhe, N. Schaller, M. R. Allen, L. Hermanson, N. Christidis, P. McLean, and H. Cullen (2016), Real-time extreme weather event attribution with forecast seasonal SSTs, *Environmental Research Letters*, *11*(6), 064,006, doi:10.1088/1748-9326/11/6/064006.

Joos, F., *et al.* (2013), Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis, *Atmospheric Chemistry and Physics*, *13*(5), 2793–2825, doi:10.5194/acp-13-2793-2013.

Le Quéré, C., *et al.* (2016), Global Carbon Budget 2016, *Earth System Science Data*, *8*(2), 605–649, doi:10.5194/essd-8-605-2016.

Lenton, T. M., H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber (2008), Tipping elements in the Earth's climate system, *Proceedings of the National Academy of Sciences*, *105*(6), 1786–1793, doi:10.1073/pnas.0705414105.

Lucarini, V., and T. Bódai (2017), Edge states in the climate system: exploring global instabilities and critical transitions, *Nonlinearity*, *30*(7), R32–R66, doi:10.1088/1361-6544/aa6b11.

Lucarini, V., and S. Sarno (2011), A statistical mechanical approach for the computation of the climatic response to general forcings, *Nonlinear Processes in Geophysics*, *18*(1), 7–28, doi:10.5194/npg-18-7-2011.

Lucarini, V., F. Ragone, and F. Lunkeit (2016), Predicting Climate Change Using Response Theory: Global Averages and Spatial Patterns, *Journal of Statistical Physics*, doi:10.1007/s10955-016-1506-z.

Meinshausen, M., *et al.* (2011), The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, *Climatic Change*, *109*(1-2), 213–241, doi:10.1007/s10584-011-0156-z.

Millar, R. J., Z. R. Nicholls, P. Friedlingstein, and M. R. Allen (2017), A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions, *Atmospheric Chemistry and Physics*, *17*(11), 7213–7228, doi:10.5194/acp-17-7213-2017.

Myhre, G., *et al.* (2013a), Anthropogenic and Natural Radiative Forcing, in *Climate Change*

2013 - The Physical Science Basis, edited by Intergovernmental Panel on Climate Change, chap. 8, pp. 659–740, Cambridge University Press, Cambridge, doi:10.1017/CBO9781107415324.018.

Myhre, G., *et al.* (2013b), Anthropogenic and Natural Radiative Forcing Supplementary Material, in *Climate Change 2013 - The Physical Science Basis*, edited by Intergovernmental Panel on Climate Change, chap. 8, Cambridge University Press, Cambridge.

Nordhaus, W. D. (1991), To Slow or Not to Slow: The Economics of The Greenhouse Effect, *The Economic Journal*, *101*(407), 920, doi:10.2307/2233864.

Pachauri, R. K., *et al.* (2014), *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*, 151 pp., IPCC, Geneva, Switzerland.

Ragone, F., V. Lucarini, and F. Lunkeit (2016), A new framework for climate sensitivity and prediction: a modelling perspective, *Climate Dynamics*, *46*(5-6), 1459–1471, doi:10.1007/s00382-015-2657-3.

Rezai, A., and F. van der Ploeg (2016), Intergenerational Inequality Aversion, Growth, and the Role of Damages: Occam's Rule for the Global Carbon Tax, *Journal of the Association of Environmental and Resource Economists*, *3*(2), 493–522, doi:10.1086/686294.

Rezai, A., and F. van der Ploeg (2017), Second-Best Renewable Subsidies to De-carbonize the Economy: Commitment and the Green Paradox, *Environmental and Resource Economics*, *66*(3), 409–434, doi:10.1007/s10640-016-0086-3.

Riahi, K., A. Grubler, and N. Nakicenovic (2007), Scenarios of long-term socio-economic and environmental development under climate stabilization, *Technological Forecasting and Social Change*, *74*(7), 887–935, doi:10.1016/j.techfore.2006.05.026.

Roe, G. H., and M. B. Baker (2007), Why Is Climate Sensitivity So Unpredictable?, *Science*, *318*(5850), 629–632, doi:10.1126/science.1144735.

Rogelj, J., *et al.* (2016), Paris Agreement climate proposals need a boost to keep warming well below 2°C, *Nature*, *534*(7609), 631–639, doi:10.1038/nature18307.

Rosenzweig, C., *et al.* (2008), Attributing physical and biological impacts to anthropogenic climate change, *Nature*, *453*(7193), 353–357, doi:10.1038/nature06937.

Ruelle, D. (1998a), General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium, *Physics Letters A*, *245*(3-4), 220–224, doi:10.1016/S0375-9601(98)00419-8.

Ruelle, D. (1998b), Nonequilibrium statistical mechanics near equilibrium: computing higher-order terms, *Nonlinearity*, *11*(1), 5–18, doi:10.1088/0951-7715/11/1/002.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An Overview of CMIP5 and the Experiment Design, *Bulletin of the American Meteorological Society*, *93*(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.

United Nations (2015), *Adoption of the Paris Agreement*, Framework Convention on Climate Change, 21st Conference of the Parties, Paris.

van Vuuren, D. P., M. G. J. den Elzen, P. L. Lucas, B. Eickhout, B. J. Strengers, B. van Ruijven, S. Wonink, and R. van Houdt (2007), Stabilizing greenhouse gas concentrations at low levels: an assessment of reduction strategies and costs, *Climatic Change*, *81*(2), 119–159, doi:10.1007/s10584-006-9172-9.

van Zalinge, B. C., Q. Y. Feng, M. Aengenheyster, and H. A. Dijkstra (2017), On determining the Point of no Return in Climate Change, *Earth System Dynamics Discussions*, pp. 1–17, doi:10.5194/esd-2016-40.

World Energy Council (2016), World Energy Resources 2016, *Tech. rep.*, World Energy Council, London.

Appendix A. Radiative Forcing for non-CO_2 gases

In section 2.3 we described how we compute the radiative forcing using CO_2 concentration only, and here describe the alternative approach to include CH_4 and N_2O. Following (Myhre *et al.* 2013b, Table 8.SM.1) we give the radiative forcing in $\mathrm{W\,m^{-2}}$ for CO_2, CH_4 and N_2O as

$$\Delta F_{CO2} = \alpha_{CO2} \ln(C_{CO2}/C_0) \tag{A 1}$$

$$\Delta F_{CH4} = \alpha_{CH4}(\sqrt{C_{CH4}} - \sqrt{M_0}) - (f(C_{CH4}, N_{00}) - f(M_0, N_{00})) \tag{A 2}$$

$$\Delta F_{N2O} = \alpha_{N2O}(\sqrt{C_{N2O}} - \sqrt{N_0}) - (f(M_{00}, C_{N2O}) - f(M_{00}, N_0)) \tag{A 3}$$

$$f(M, N) = 0.47 \ln\left(1 + 2.01 \times 10^{-5}(M \cdot N)^{0.75} + 5.31 \times 10^{-15} M (M \cdot N)^{1.52}\right) \tag{A 4}$$

where C_0 is the pre-industrial (1750) CO_2 concentration, M_{00}, N_{00} are the present-day (2011), and M_0, N_0 the pre-industrial (1750) CH_4 and N_2O concentrations. The total radiative forcing for all considered contributers is then simply the sum

$$\Delta F_{tot} = A \sum_i \Delta F_i \tag{A 5}$$

with scaling constant $A = 1.1$ to take into account still disregards contributions (land-use emissions, CFCs etc.).

We show the result, overlaid with RCP simulation results, in Fig. 15.

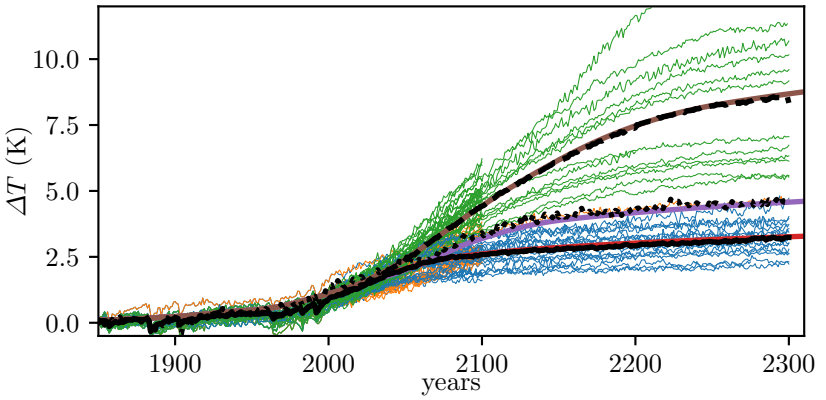

Figure 15: Reconstruction of RCP temperature evolution from concentration pathways using CO_2, CH_4, N_2O. Blue, orange and green lines gives CMIP5 data for RCP4.5, RCP6.0 and RCP8.5, respectively, with the ensemble mean given in solid (RCP4.5), dotted (RCP6.0) and dashed (RCP8.5) black. Reconstruction using CO_2 radiative forcing and CMIP5 Green's function in red (RCP4.5), purple (RCP6.0) and brown (RCP8.5).

The agreement when using all three gases is slightly better (in particular for the high emissions scenario RCP 8.5) than when only using CO_2.

Appendix B. Emission Model for non-CO_2 gases

We mentioned in section 3.1 the possibility to model CH_4 and N_2O in a similar fashion to CO_2 using response functions. For some applications and research questions this may indeed be warranted, as on short timescales their contribution may reach or exceed the role of CO_2 (Myhre *et al.* 2013a, Figure 8.32).

Following Myhre *et al.* (2013a) the response functions for CH_4 and N_2O are simpler than for CO_2 and can just be case as exponentials, so

$$G_{CO2}(t) = a_0 + \sum_{i=1}^{3} a_i e^{t/\tau_i} \tag{B 1}$$

$$G_{CH4}(t) = e^{t/\tau_{CH4}} \tag{B 2}$$

$$G_{N2O}(t) = e^{t/\tau_{N2O}} \tag{B 3}$$

and the time evolutions are computed as

$$C_{CO2}(t) = \int G_{CO2}(t-t')\ E_{CO2}(t')\ dt' \tag{B 4}$$

$$C_{CH4}(t) = \int G_{CH4}(t-t')\ E_{CH4}(t')\ dt' \tag{B 5}$$

$$C_{N2O}(t) = \int G_{N2O}(t-t')\ E_{N2O}(t')\ dt' \tag{B 6}$$

with the instantaneous emissions and response function for each gas. From these instantaneous concentrations we compute their radiative forcing (appendix A) and use it in the temperature response function.

Appendix C. Safe Carbon Budget Parameter Sensitivity

We show the results for the parameter sensitivity study of the SCB calculation described in section 4.1. Sensitivities were determined for all discussed values of T_{max}, β. Here we show sample values for a small ($T_{max} = 1.5\,\mathrm{K}, \beta = 0.95$), intermediate ($T_{max} = 1.5\,\mathrm{K}, \beta = 0.5$), and large ($T_{max} = 2.0\,\mathrm{K}, \beta = 0.5$) SCB.

T_{max}, β	1.5 K, 0.95	1.5 K, 0.5	2.0 K, 0.5
undisturbed	10.49	86.67	199.03
Initial Conditions			
C_P	–, 75.26	–, 77.41	–, 80.11
C_1	7.14, -7.05	6.96, -6.94	6.87, -6.45
C_2	0.85, -0.91	0.92, -0.96	0.78, -0.91
C_3	-0.06, 0.01	0.0, -0.05	0.19, -0.11
ΔT_0	1.4, -1.46	1.92, -1.92	2.2, -1.84
ΔT_1	-0.01, -0.05	-0.01, -0.03	0.17, 0.01
ΔT_2	-0.04, -0.11	-0.0, 0.03	-0.0, -0.12
Climate Parameters			
a_0	0.55, -0.49	4.17, -3.81	9.14, -8.14
a_1	0.41, -0.44	3.61, -3.39	7.99, -7.13
a_2	0.17, -0.11	1.07, -1.01	3.0, -2.55
a_3	-0.04, 0.0	0.07, -0.09	0.31, 0.08
τ_1	1.59, -1.28	2.03, -1.68	2.64, -2.01
τ_2	1.75, -1.73	3.0, -2.9	5.18, -4.68
τ_3	-0.08, -0.07	0.01, -0.06	-0.14, 0.13
A	23.26, –	33.77, -27.0	50.61, -39.92
α	23.32, –	33.76, -27.04	50.48, -39.84
C_0	–, 78.33	-86.44, 83.71	-94.33, 91.34
b_0	3.24, -3.17	5.79, -5.61	8.38, -7.72
b_1	5.13, -4.9	9.49, -8.8	13.99, -13.25
b_2	13.36, –	16.16, -14.38	24.6, -21.48
τ_{b0}	0.92, -0.74	1.11, -0.97	1.56, -1.17
τ_{b1}	5.23, -4.97	9.53, -8.88	14.63, -13.12
τ_{b2}	15.66, –	15.83, -14.01	24.11, -21.02
Stochastic Parameters			
σ_{C2}	0.1, -0.15	-0.02, -0.0	-0.02, 0.01
σ_{T0}	1.68, -2.01	0.29, -0.29	0.06, -0.42
σ_{T2}	4.08, -4.14	-1.26, 1.44	-1.96, 2.22

Table 13: Sensitivity of Safe Carbon Budget to parameter variations. Values in difference in GtC from the undisturbed value (first row). First and second numbers give 10% parameter decrease and increase, respectively. No value being shown implies no positive SCB could be calculated. The fields corresponding to the radiative forcing parameters A, α, C_0 are colored in cyan, while the most sensitive climate model parameters b_2, τ_{b2} are given in orange.

Appendix D. Point of No Return Parameter Sensitivity

We show the results of the parameter sensitivity study of the PONR calculation discussed in section 4.2. We performed the study for all scenarios listed in Tab. 7 and show them here for sample scenarios with PONR values that are close (EM,$T_{max} = 1.5\,\mathrm{K}, \beta = 0.95$), far away (EM,$T_{max} = 2.0\,\mathrm{K}, \beta = 0.5$), and intermediate (EM,$T_{max} = 1.5\,\mathrm{K}, \beta = 0.5$).

T_{max}, β	1.5 K, 0.95	1.5 K, 0.5	2.0 K, 0.5
undisturbed	2031	2018	2047
Initial Conditions			
C_P	10, -12	12, –	7, -9
C_1	1, -1	1, -2	0, -1
C_2	1, 0	0, 0	0, 0
C_3	0, 0	0, 0	0, -1
ΔT_0	1, 0	0, 0	0, -1
ΔT_1	0, 0	0, 0	0, 0
ΔT_2	0, 0	0, 0	0, 0
Climate Parameters			
a_0	1, 0	0, 0	1, -1
a_1	1, 0	0, 0	1, -1
a_2	1, 0	0, 0	0, -1
a_3	0, 0	0, 0	0, 0
τ_1	1, 0	0, 0	0, -1
τ_2	1, 0	0, -1	0, -1
τ_3	0, 1	0, 0	0, 0
A	6, -4	5, –	5, -5
α	5, -4	5, –	5, -5
C_0	–, 12	–, 14	-13, 9
b_0	1, -1	1, -1	1, -1
b_1	2, -1	1, -1	1, -2
b_2	3, -2	3, –	2, -3
τ_{b0}	1, 0	0, 0	0, -1
τ_{b1}	2, -1	1, -1	1, -2
τ_{b2}	3, -2	4, –	3, -3
Economic Parameters			
Y_0	2, -1	0, 0	2, -2
g	1, 0	0, 0	0, -1
γ_0	2, -1	0, 0	2, -3
r_γ	2, 3	0, 0	4, 11
Stochastic Parameters			
σ_{C2}	1, 0	0, 0	0, 0
σ_{T0}	0, 0	0, 0	0, 0
σ_{T2}	0, 0	1, -1	0, 0

Table 14: Sensitivity of Point of No Return to parameter variations. Values in difference in years from the undisturbed case (first row). First and second numbers give 10% parameter decrease and increase, respectively. Exception is r_γ (in orange) which is zero by default. First and second numbers here give $r_\gamma = 0.01$ and $r_\gamma = 0.02$, respectively. No value being shown implies no PONR was found, i.e. it is already passed. All three columns refer to the EM scenario. The fields corresponding to the radiative forcing parameters A, α, C_0 are colored in cyan, while the most sensitive climate model parameters b_2, τ_{b2} are given in orange.