Supplementary Information

Idealised emission scenarios:

We generate carbon dioxide emission scenarios by assuming the fractional rate of change in emissions f_E initially continues close to the rate observed since 1980 (1.7% increase/year in the black and red lines in figure 1a, between 1.3 and 3% increase/year in the orange lines: see figure S3a) [31]. After a certain date t_1, f_E is reduced at a constant rate, passing zero (stable emissions) after t_2 years and continuing to fall until it reaches a maximum rate of decline $f_E = -t_3^{-1}$.

Varying t_1 from 2010 to 2050 (5 values), t_2 from 5 to 25 years (5 values) and t_3 from 10 to 190 years (10 values), gives 250 scenarios: the black lines in Fig. 1a (a representative subset plotted for clarity). Varying t_1 to t_3 over a finer grid, varying the initial growth rate and selecting scenarios with cumulative emissions within 1% of 1TtC gives the orange lines. We include emission reduction rates of up to 10% per year, as our focus is on the range of behaviours of the climate system, but it should be noted that such high rates of reduction are generally considered prohibitively expensive.

For each setting of model carbon cycle parameters, historical anthropogenic (fossil and land-use-change) emissions are scaled to give the observed rate of increase in atmospheric CO_2 over 1961-2000 [31] and the same scaling factor applied to future emissions. Hence in computing warming commitments, we assume that the fractional error in past emissions persists into the future, or equivalently that all values of CWC are referenced to our best-estimate of the cumulative CO_2 emissions over 1750-2000 (0.44 TtC).

Constraining climate and carbon cycle response:

We take two approaches to quantifying the range of uncertainty in Cumulative Warming Commitment (CWC) consistent with currently-available information: first, we assess the range of CWCs that are consistent with climate observations and the constraints of a very simple climate model; and second, we assess the range of CWCs we expect from currently-available Earth System Models (ESMs), using the HadSCCCM1 model to emulate the behaviour of these more complex models. ESMs are typically tuned so that some aspects of their behaviour replicate the historic or present day observed climate, but unlike our very simple climate model ensemble, the ensemble of more complex models is not explicitly designed to span the full range of behaviour consistent with observations.

To assess the range of CWCs consistent with climate observations, we require a representation of the coupled climate-carbon-cycle system that captures the essential behaviour of more complex models (and hence, it is hoped, the Earth system itself), but

also sufficiently simple for key parameters to be constrained directly. For the global temperature anomaly, $T(t)$, response to carbon dioxide concentrations $C(t)$, we use a conventional energy balance model coupled to a diffusive ocean [32]: despite its simplicity, such a model allows multiple time-scales in the response, which is essential if we are to interpret the constraints provided by recent trends.

$$a_1 \frac{dT}{dt} = a_3 \ln\left(\frac{C}{C_0}\right) - a_0 T - a_2 \int_0^t \frac{dT(t')}{dt'} \frac{dt'}{\sqrt{t-t'}} \quad (1)$$

The constants a_1 and a_3 are fixed to values corresponding to a 75m ocean mixed layer, 70% ocean coverage and forcing on doubling carbon dioxide of 3.74Wm^{-2}, while a_0 (the feedback parameter which determines ECS) and a_2 (which controls downward diffusion of surface temperature anomalies) are allowed to vary. The introduction of additional complexity, such as upwelling or a variable land-sea temperature difference, results in over-parameterization due to the low number of reasonably unambiguous observational constraints available for a model that only simulates global quantities.

The climate response of the model is constrained by the linear temperature trend attributable to the increase in greenhouse gases over the 20th Century based on "fingerprint" attribution results (the blue curve in figure S1a) [21] and the effective heat capacity of the atmosphere-land-ocean system (figure S1b) which is implied by the combination of observed surface warming [33] and total ocean heat uptake over the period 1955-98 [34]. To convert total warming induced by long-lived greenhouse gases into CO_2-induced warming, we use CO_2 and non-CO_2 forcing timeseries which give 2000 values of 1.47 and 0.85 Wm^{-2} respectively, consistent with refs. [21,30] and unit efficacies, following the majority of studies constraining ECS and transient climate response (TCR) cited in ref. [1]. Revised estimates of forcing and efficacies [35] might reduce our estimate of the fraction of past greenhouse-gas-attributable warming that is due to CO_2, reducing CWC by a corresponding factor. This would, however, also require a revision of observationally-constrained estimates of ECS and TCR. Since these are not the focus of this study, we prefer to use figures consistent with the studies cited in ref. [1] which may give a conservatively high estimate of CWC, noting that other neglected uncertainties, such as the possibility of a_0 changing over time, could introduce a bias in the opposite sense.

We use the best-guess value for ocean heat uptake from [34]: more recent estimates of upper ocean heat content increase suggest somewhat higher values [36], but these are still well within our confidence interval. To facilitate reproduction of our results and tracing the origins of uncertainties, we use parameterized likelihood functions which closely approximate the errors provided by the sources. In all cases, if X is the observable quantity of interest, we assume fractional errors in X, or errors in $\log(X)$, are normally distributed with mean zero and standard error $\sigma_{\log(X)}$: this assigns the same likelihood to equal fractional errors in both directions. The dots on figures S1a and S1b show the relative likelihoods generated by random variations of a_0 and a_2 in equation (1) when forced by observed greenhouse gas increases and likelihoods are assigned to computed surface temperature trends over the full 20th Century and effective heat capacities over 1958-98. The overall likelihood of a given parameter combination is given by the product of the likelihoods of the two constraints, as the observational errors are assumed to be

independent, but the density of dots under the likelihood profile is arbitrary. See ref. [19] for an explanation of the principle of likelihood profiling.

We can plot these likelihoods against any output of the model: figure S1c shows the implied likelihood profile for ECS, or the equilibrium warming response to stabilizing atmospheric CO_2 at double its pre-industrial concentration. Upper and lower horizontal dotted lines show likelihood thresholds corresponding to the 17-83% and 5-95% asymptotic confidence intervals. In the language of the IPCC, it is "likely" and "very likely" that the range over which the likelihood profiles exceed these thresholds covers the correct (but unknown) values of the quantities plotted on the abscissa in figures S1 and S2: note the exact correspondence between this range and the nominal 5-95% confidence interval from ref. [21], shown by red and green diamonds in fig. S1a. The key advantage of the likelihood profile is that it makes clear the information provided by the data and is not sensitive to the manner in which parameters are sampled in the model (the density of dots in figures S1 and S2), in contrast to a Bayesian posterior: provided the ensemble is large enough to delineate the likelihood profile, this derived profile is insensitive to the density of points underneath it. If a forecast quantity is simply proportional to one of the observed quantities used to constrain the model and independent of all the others, then the confidence interval for that forecast quantity will be a correspondingly scaled version of the confidence interval on that observed quantity. Past greenhouse-gas attributable warming, for example, explains much of the variation in CWC, so there is a direct link between these confidence intervals (which, as noted in the main text, would be even more direct if we could assume a negligible ZEC). Likelihood profiling is not new to climate research: it is one way of expressing "optimal fingerprinting" as used by the detection and attribution community and exploring its implications for climate forecasts [21].

The range of values of ECS that lie above the lower threshold gives a 5-95% confidence interval of 2.0-4.8^{o}C, consistent with refs. [1,2]. The upper bound on any estimated interval for ECS is particularly sensitive to details of the analysis because of the non-linear relationship between observable quantities and ECS which creates the "fat tail" on the likelihood profile for the latter. A key physical assumption in this model is that the feedback parameter a_0 is constant: introducing the possibility of climate-state-dependent feedbacks further increases uncertainty in ECS. It would also increase uncertainty in CWC, but to a lesser extent because of the shorter timescales involved in CWC.

Representing the carbon cycle is more challenging because there are more sources of complexity and weaker observational constraints. We use a simple box-diffusion model, assuming, first, that the long-term (millennial timescale) equilibrium CO_2 level C_3 increases with cumulative emissions [4]:

$$\frac{dC_3}{dt} = b_3 E \qquad (2)$$

where E is the net (anthropogenic and natural) emission rate in ppm-equivalent-per-year and b_3 corresponds to the Revelle Buffer Factor. This is held constant at 0.1 to reproduce the 1000-year response of current ESMs [29]: estimates of CWC are insensitive to b_3, whereas the response to a "stabilisation" emission scenario depends on it strongly. CO_2

levels equilibrated between the atmosphere, land-biosphere and upper-ocean exceed this long-term value by an amount C_2 which relaxes towards zero through advection into the deep ocean:

$$\frac{dC_2}{dt} = b_1 E - b_0 C_2 \quad (3)$$

where $b_1+b_3=0.35$ is the level to which the atmospheric CO_2 anomaly would fall after a unit pulse injection in the absence of deep-ocean advection and b_0^{-1} represents an adjustable time-constant of order 200 years. Finally, penetration of atmospheric CO_2 anomalies into the land-biosphere and ocean thermocline is represented by diffusion:

$$\frac{dC_1}{dt} = b_4 E - b_2 \int_0^t \frac{dC_1(t')}{dt'} \frac{dt'}{\sqrt{t-t'}} \quad (4)$$

where C_1 is the excess atmospheric CO_2 over the short-term equilibrium level C_2+C_3, $b_1+b_3+b_4=0.85$ is the level to which the atmospheric CO_2 anomaly would fall within a year of a unit pulse injection and b_2 represents an adjustable diffusivity. Atmospheric CO_2 levels are given by the sum $C=C_1+C_2+C_3$. With appropriate choices of parameters, this model can reproduce the response of more complex ESMs to a pulse injection of CO_2 with significantly fewer adjustable parameters than conventional impulse-response models [37] (impulse-response models replace equation (4) with a set of equations identical to equation (3)). Including an explicit diffusive term means that CO_2 levels do not fall as rapidly following a cessation of emissions as they do in an impulse-response model with the same number of free parameters [38], reflecting the behaviour of some more complex models more faithfully [18]. We represent emissions E by the sum of anthropogenic emissions and a linear temperature feedback:

$$E = E_a + b_5 T' \quad (5)$$

where T' is the temperature anomaly above an exponentially-weighted running mean of the preceding century, representing the timescale over which anomalous sources and sinks re-equilibrate. We use historical temperatures until 2000 and computed temperatures thereafter for T', adjusting the baseline to ensure continuity. More complex representations of the impact of rising temperatures on carbon uptake or release by soils, vegetation and the upper ocean give an overall feedback that is close to linear over small temperature changes [16], but this linearization is unlikely to be valid for temperature anomalies greater than 3-4^{o}C.

For any given values of b_0 and b_5, b_2 is constrained by the observed increase in atmospheric CO_2 from 1961-2000 [31] divided by total anthropogenic emissions over the same period [40,41] (the "net airborne fraction" shown in figure S2a), with natural emissions b_5T' driven by observed temperature anomalies about the 1900-1920 average. The range is dominated by current uncertainty in land-use-change emissions: we use land-use emissions from ref. [41] and assume the fractional uncertainty quoted in ref. [42] interpreted as a 5-95% interval. Note that [42] report this range as a "likely" (one standard error) interval, but this gives a 5-95% interval that includes negative land-use-change emissions, which we believe overstates the uncertainty.

Lacking direct observational constraints on temperature feedbacks in the carbon cycle, we constrain b_5 by driving the model with historical emissions followed by the CO_2

component of the SRES A2 emissions scenario from 1750-2100 both with, $C_{A2}(t)$, and without, $C'_{A2}(t)$, the temperature feedback term. The contribution of temperature feedback to the net airborne fraction, $(C_{A2}(2100)-C'_{A2}(2100))/C'_{A2}(2100)$, is compared to corresponding values from the C^4MIP experiment (coloured symbols in Fig. S2b: colours correspond to the models shown in Fig. 2). The estimated likelihood profile (blue curve in Fig. S2b) is obtained by taking the mean of the C^4MIP ensemble as the best-guess and assuming the largest fractional departure from that mean represents two standard errors. For this comparison we use best-fit values of a_0 and a_2: adjusting these to match those of the C^4MIP models has little impact on estimates of b_5. Finally, we impose a likelihood profile for b_0 consistent with the range of reported values of the long-term response to pulse-injection experiments: see fig. S2c [29].

The response of this simple climate model to the range of emission scenarios shown in figure 1a of the main text, with best-fit values of parameters, is shown in fig. S3: panel a shows the fractional rates of change used to generate the emissions shown in fig. 1a of the main text, panel b shows atmospheric CO_2 concentrations and panel c shows the temperature response, with the same colour coding throughout. Red and orange lines correspond to scenarios with a total cumulative CO_2 emission of 1 TtC: despite the large range in peak and 2050 emissions across these scenarios, they give a very small range of CO_2 concentrations in the late 21st century (the red and all the oranges curves are almost indistinguishable in fig S3b) and an even smaller range of temperatures (fig S3c): hence the timing as well as the peak warming is insensitive to the timing of emissions (for a given emissions total).

The dotted black lines in figs S3 panels b & c show the best guess (and the shaded bar and symbols at 0.44 TtC in Fig. 2 of the main text show the range of uncertainty) in response to a complete cessation of emissions in 2000, the Zero Emissions Commitment or ZEC. It is noticeable that CO_2 concentrations and hence temperatures take of order a century to decline significantly: while our best-fit ZEC is relatively small, consistent with refs. [4,25], we obtain much larger overshoots because we allow a larger range of uncertainty in ECS which in turn feeds back onto the carbon cycle. This behaviour arises because we have described the short-timescale carbon sinks as a diffusive process, driven by gradients in CO_2 partial pressure rather than absolute CO_2 anomalies. Hence, as soon as CO_2 levels stop rising, these short-timescale sinks rapidly equilibrate, leaving only slow advection into the deep ocean, (competing with the positive carbon-cycle feedback), as the sole draw-down mechanism. This may give a relatively pessimistic view of the ZEC, but it is consistent with our physical understanding of how some carbon sinks work, and hence cannot be ruled out. Systematic studies of the response to realistic peaking and declining emissions with more complex models should clearly be a priority.

Since fractional errors in (or errors in the logarithm of) the quantities used to constrain the model are all assumed to be normal, we can compare the impact of each of these constraints on our estimate of SWC simply by reducing their estimated fractional uncertainty (standard error in $\log(X)$) either by a constant amount (we use 0.05) or by a constant fraction (we use 50%). Results are shown in the Table. Both relative and absolute reductions in fractional uncertainty in 20th century warming attributable to

greenhouse gases have a substantial impact on uncertainty in CWC. A halving of uncertainty in greenhouse gas attributable warming has been predicted by 2020, simply through the strengthening signal [43]. Halving fractional uncertainty in the temperature feedback on the carbon cycle also reduces uncertainty in CWC by a substantial (but smaller) amount primarily because uncertainty in this feedback is currently very large. Since this is currently an almost entirely model-based quantity, the experience of research into ECS suggests that such a reduction in uncertainty may take decades of research. Current fractional uncertainty in net airborne fraction is only slightly over 0.05, so subtracting 0.05 (i.e. largely eliminating uncertainty in this quantity) reduces uncertainty in CWC, but halving fractional uncertainty in this quantity has no detectable impact. Reducing uncertainty in the other constraints, either in relative and absolute terms, also has no significant impact on uncertainty in CWC.

The HadSCCCM1 model [10] consists of climate and carbon cycle components, each tuned to the response of more complex ESMs of intermediate complexity and GCM-carbon-cycle models used in the C^4MIP study [15,16]. The tuning was carried out using results from an experiment that begins in the pre-industrial period, follows historic carbon emission estimates up to 2000 then SRES A2 emissions to 2100. No other greenhouse gas forcing changes were included. Whilst the C^4MIP ensemble illustrates the range of behaviour of current ESMs, it is not a comprehensive assessment of model uncertainty: unlike our first simple climate model, parameters in the C^4MIP experiment have not been varied systematically to span the range of behaviour consistent with recent observations. The climate component of HadSCCCM1 also closely follows equation (1). Free parameters controlling ECS and ocean heat uptake are found where possible through a least-squares fit to 1850-2100 global temperatures simulated by the C^4MIP models. In eight cases, however, the optimum fit was not well defined, in which case literature values for ECS for that climate model were adopted, and the ocean heat uptake parameter found by fitting [15]. Note that since the C^4MIP simulations are driven by the SRES A2 scenario to 2100, the fit is dominated by model behaviour over the 21st Century.

The ocean carbon cycle component of HadSCCCM1 consists of two parts. The first is a diffusive flux of carbon through the ocean surface which depends on the difference in concentration of carbon dioxide between the atmosphere and the ocean mixed layer. The second step is the carbon removal from the mixed layer into the deep ocean and is estimated using a linear impulse response function with multiple time modes [37]. The only parameter we alter in the ocean carbon cycle (again through comparison against the C^4MIP simulations) is the depth of the mixed layer as seen by the carbon cycle.

The terrestrial carbon cycle model has both vegetation and soil components stores. The vegetation carbon content is a balance between global average net primary productivity (parameterized as a function of atmospheric carbon dioxide, which asymptotes to a maximum value multiplied by a quadratic function of temperature rise in order to represent the effect of climate change) and vegetation carbon turnover. The vegetation carbon turnover time, which governs the rate at which vegetation carbon is lost to the soil in the form of litter, is a function of vegetation carbon. The soil carbon is then a balance between the litter supply and respiration loss, the latter a function of soil temperature and

carbon content; the soil temperature response is frequently parameterised as a "Q_{10}" dependence (e.g. [39]). As in the temperature and ocean model, these carbon fluxes are calibrated against the corresponding diagnostics from the more complex models in the C^4MIP simulations, in this case the net primary productivity, soil respiration and vegetation and carbon stores. A more comprehensive description of the HadSCCCM1 model set up and validation of its performance against the tuning dataset are available in refs. [10,15].

In a final piece of analysis, we diagnose average (or, equivalently, time-integrated) warming over the 2000-2500 period: this is expected to be more important for certain impacts, such as ice sheet melting, than peak warming. We present these values in Figure S4, plotted against cumulative carbon emissions. Figure S4 has very strong similarities to figure 2 in the main text: we conclude that not only peak temperature but also mean long-term warming is determined by cumulative emissions, and not the particular shape of the emissions profile. The small difference between peak warming and long-term average warming highlights again that, once cumulative emissions have committed us to a certain level of warming, recovery may take many centuries.

Additional References:

32. Hansen, J. *et al*: Climate response times: dependence on climate sensitivity and ocean mixing, *Science*, **229**: 857–859 (1985).
33. Brohan, P. *et al*, Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, *J. Geophys. Res.*, **111**, D12106, doi:10.1029/2005JD006548, (2006).
34. Levitus S., Antonov J. and Boyer T.: Warming of the world ocean, 1955–2003. *Geophys. Res. Lett.* **32**: L02604, DOI: 10.1029/2004GL021592 (2005).
35. Gregory, J. M. and Forster, P. M., transient climate response estimated from radiative forcing and observed temperature change, *J. Geophys. Res.*, **113**, D23105, (2008).
36. Domingues, C. M. *et al*, Improved ocean warming estimates: implications for climate models and sea level rise, *Nature*, **453**, 1090-1093, (2008).
37. Joos, F., Bruno, M., Fink, R., Siegenthaler, U. and Stocker, T.F.: An efficient and accurate representation of complex oceanic and biospheric models of anthropogenic carbon uptake, *Tellus*, **48B**, 397-417, (1996).
38. Socolow, R. H. and Lam, S. H., Good enough tools for global warming policy making, *Phil. Trans. R. Soc. A* **365**, 897–934, (2007).
39. Raich, J & Schlesinger, W.: The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate, *Tellus*, **44B**, 81-99, (1992)
40. Marland, G., T.A. Boden, and R.J. Andres. 2008. Global, Regional, and National CO2 Emissions. In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.
41. Houghton, R.A. 2008. Carbon Flux to the Atmosphere from Land-Use Changes: 1850-2005. In TRENDS: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

42. Denman, K. L. *et al* Couplings Between Changes in the Climate System and Biogeochemistry, in Solomon *et al* (eds.) *Climate Change 2007: the Physical Science Basis.* Cambridge University Press, 2007
43. Stott, P. A. and J. Kettleborough, Origins and estimates of uncertainty in 21st century temperature rise, *Nature*, **416**:723-726

Climate system property, *X*	Most likely value of *X*	5-95% confidence interval	Reduction in fractional uncertainty in CWC due to reducing fractional uncertainty $\sigma_{\log(X)}$ by	
			0.05	50%
20th century warming trend attributable to GHGs	0.97 °C/century	0.73-1.27 °C/century	18%	29%
Effective heat capacity 1955-98	0.70 GJ/°C	0.38-1.30 GJ/°C	1%	3%
Net airborne fraction (AF) 1960-2000	0.43	0.39-0.47	5%	0%
Contribution of temp. feedback to net AF 1766-2100	0.17	0.07-0.39	5%	13%
Rate constant for advection of CO_2 into deep ocean	200 years	133-302 years	0%	0%

Table: Summary of constraints on simple climate model. Columns 4 & 5 show impact of reducing fractional uncertainties in constraints on the fractional uncertainty in CWC, expressed as the percentage change in the ratio between the upper and lower bounds on the 5-95% confidence interval.

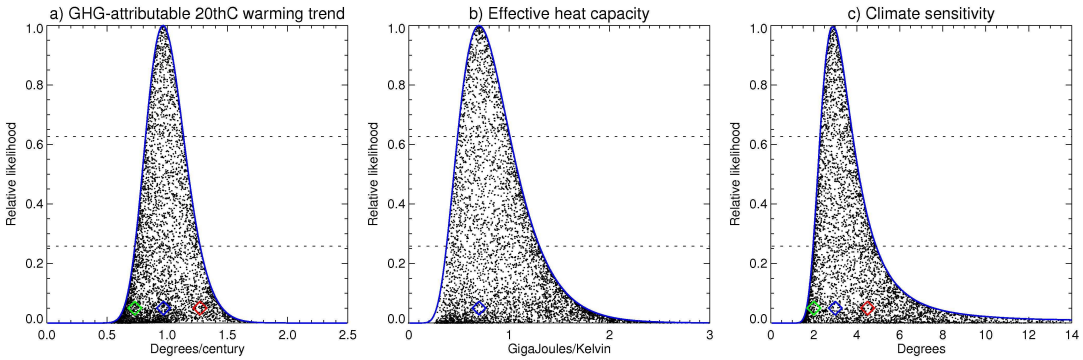

Figure S1: Likelihood profiles for the temperature component of the simple climate model. Dots show relative likelihood of random combinations of parameters a_0 and a_2. Blue curves in panel (a) show likelihood of the 20th century warming trend attributable to the observed greenhouse gas increase, blue, green and red diamonds show best-fit and 5-95% range from ref. [21]. Blue curve in panel (b) shows likelihood of effective heat capacity of the atmosphere/land/ocean system over 1955-1998. Best estimate from refs. [33] and [35] and likelihood profile obtained by summing (in quadrature) fractional errors in pentadal surface warming and ocean heat uptake from refs. [33] and [35] respectively. Blue curve in panel (c) shows the likelihood profile for climate sensitivity implied by these constraints: for comparison, blue, green and red diamonds show best-guess and "likely" (17-83%) uncertainty range in sensitivity from ref. [1].

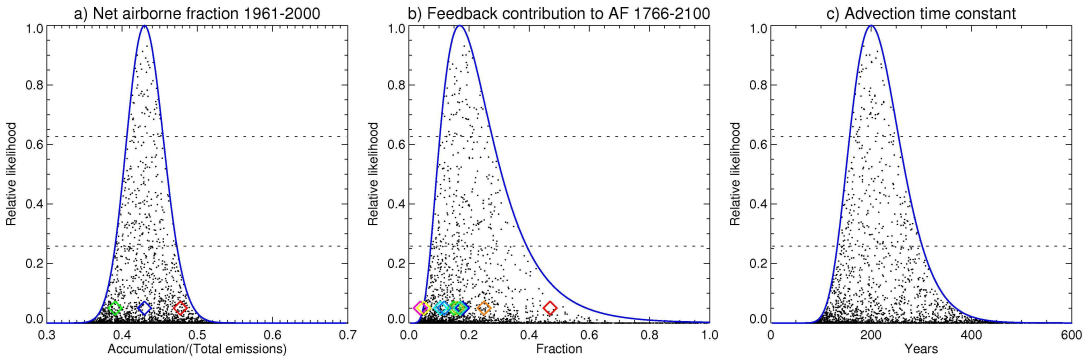

Figure S2: Likelihood profiles for carbon cycle component of simple climate model. Dots show relative likelihood of random combinations of parameters b_0, b_2 and b_5. Blue curve in panel (a) shows estimated likelihood profile for net Airborne Fraction (AF) over the period 1961-2000 based on atmospheric CO_2 measurements at Mauna Loa [31] and anthropogenic CO_2 emissions from fossil fuel burning [40] and land-use change [41] allowing for fractional uncertainty in land-use-change emissions [42]. Blue symbol shows most likely net AF while green and red diamonds show range based on "likely" land-use-change emissions from ref. [42]. Blue curve in panel (b) shows estimated likelihood profile for the contribution of temperature feedbacks to the net AF 1750-2100 under the A2 emissions scenario, with coloured symbols showing the corresponding quantity from the C^{4}MIP experiments. Panel (c) shows the estimated likelihood profile for b_0^{-1}, the time-constant for advection of carbon into the deep ocean.

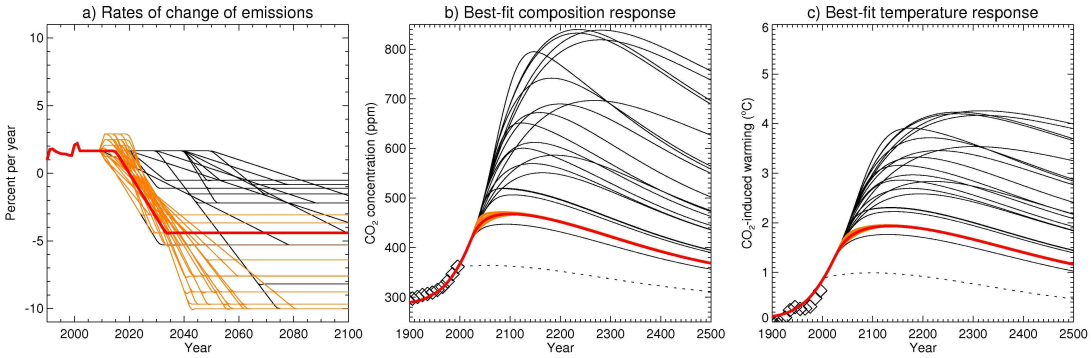

Figure S3: Rates of change of emissions used to generate idealized scenarios and responses with best-fit parameters. Panel a: fractional rates of change of emissions used to generate the idealized emission scenarios shown in figure 1a, main text. Panels b & c: Response of simple climate model with best-fit values of model parameters. Note the small range of CO_2 concentrations and even smaller range of temperature responses to the orange emission profiles shown in panel a, all of which have the same cumulative CO_2 emission. Dotted line shows best-fit response assuming zero emissions after 2000: both CO_2 and temperatures take over a century to decline.

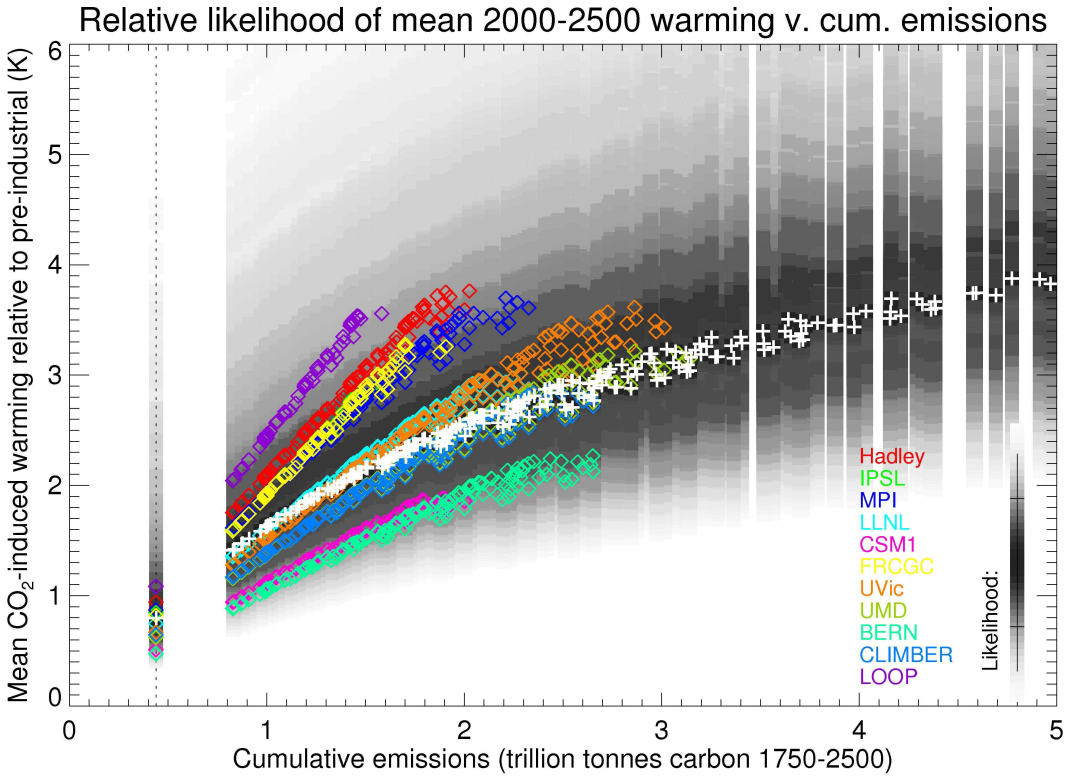

Figure S4: As figure 2 in main text, but plotting mean warming over the period 2000-2500 against cumulative CO_2 emissions.