Response to Reviewer 1

This paper introduces an interesting concept of how to account for the long-term effects (such as changing climate feedback parameter) in the carbon budgets framework. However, I found it challenging to understand this study in the context of applications to carbon budgets, and how these findings should be interpreted. I would recommend revising the framing of this paper to make it more relevant and easier to follow for readers familiar with the carbon budgets literature.
Many thanks to the reviewer for the extensive review - and I recognise (from both reviews) that the initial submission required better context in the existing carbon budgets literature. I have endeavoured to incorporate the suggestions and better frame the paper in the revised version.

The paper could also benefit from clarifications, and consistency with the most recent literature on carbon budgets and TCRE. In particular, the role of non-CO2 forcing on hysteresis of the effective TCRE curves discussed in the paper should be clearly separated from making claims on hysteresis in TCRE alone (which applies to CO2-induced warming only). I included several suggestions that potentially could help to clarify the points of confusion. Also, I would suggest putting the findings of this paper in the context of the overall uncertainties in carbon budgets (see IPCC Special Report, Table 2.2, Chapter 2, for a summary of different uncertainties), which I suppose are much larger than the uncertainties in carbon budgets due to changing climate feedbacks.

Point well taken - I have worked to make clear in the revised version that there is potential for hysteresis from a number of factors, of which this study only deeply considers one - with better reference to the SR15 conclusions.

Furthermore, I would suggest discussing the uncertainties in the observational datasets explicitly in the main text, and how they affect the results of constraining the simple model.
Agreed, I have structured the methods to be inline with the document - including the observational uncertainties.

Major points

1. Model description

It is unclear what model is used in this study -is it a version of the FaIR model with additional components that would account for changing long-term feedbacks, or is it a simpler impulse-response model that contains less processes than FaIR? (the references are a bit vague).

Sorry about this - I've made the current version more clear. It's a Green's function solution of the core carbon-climate equations in FAIR - not the whole model (i.e. there's no complex chemistry - just a single bulk forcing term for aerosols.)

I would suggest to include basic description of the model in the main text (e.g. on lines 70-75) in the context of recent climate model emulators (or how it differs component-wise from FaIR) for the readers to have a brief idea of how the climate response is determined without the need of referring to the appendix.

The methods are now inline with the document, with an expanded description of the model

Lines 180-185: It is unclear how the emulator used here differs from the FaIR emulator? Is this an extension of FaIR that accounts for the possibility of changing feedbacks, or is it a simplified version of it.

This should be clearer now - it's a green's function implementation of the core dynamical assumptions which is fast enough to run MCMC calculations (I started out using FAIR itself, but the solver was too slow to run the number of iterations needed to calculate posterior distributions).

2. Observational constraints from the historical period

It is unclear how the observational uncertainty both in the observed warming and in the estimates of cumulative CO2 emissions from the global carbon project affects the results.

Thanks for this point. I've made efforts to expand the discussion of observational uncertainty - and have updated the analysis to better consider these aspects.

The approach of Millar and Friedlingstein 2018 (MF18) is not possible here - given the MCMC optimization is computationally demanding so repeating the probabilistic assessment for each member of Cowtan and Way (CW hereon) observational ensemble would not be practical, nor would the results be particularly meaningful (an ensemble of posterior distributions).

Given this - I've attempted to ensure that observational uncertainties are appropriately considered in the model's parameter space. For the case of climate sensitivity parameters, showing that the range of 20th C warming is consistent with the CW spread. For emissions uncertainty - I've introduced a new parameter which introduces uncertainty into the emissions in a given year to account for land use emissions uncertainty. A prior on this parameter is now chosen such that distribution of cumulative emissions in 2016 is consistent with MF18.

I would suggest either illustrating it on Figure 1 or at least discussing the following points in the main text:

Figure 1: 'observed' cumulative CO2 emissions – Please include references to the observational datasets in the perhaps in the figure caption, and specify if they include the total CO2 emissions (from fossil fuels and land use change?)

Done - now total land use and fossil emissions from GCP 2019.

If so, the uncertainty on estimated CO2 land use change emissions in the historical period is quite large (even up to +/-50 percent for the annual E luc emissions), and it should be indicated on the figure or at least mentioned in the text and the figure caption. (e.g. see Table 5 from the recent Global Carbon Project 2019).

Thanks for this point. The revised manuscript allows for uncertainty in historical cumulative emissions by introducing a new parameter in the model, for which the prior is manually adjusted to replicate (very well) the distribution of historical emissions in MF18, see new Figure S3.

Figure 1: observed warming from HadCRUT4 – is it adjusted for the blending-masking effects? If not, it is not like-for-like comparison with the global (and complete coverage) climate models' output. In such case, at least a caveat in the figure caption and a short mention of this point would be useful. (e.g. see Cowtan et al. 2015; Richardson et al. 2016, 2018).

I've shifted to using the Cowtan/Way 2015 ensemble median (for calibration target) and ensemble to assess the sigma T parameter which conveys the degree to which we trust that data in the MCMC calibration

Uncertainties in the other observation-based quantities (heat content, paleo and RWF) should be discussed, as some of those inputs/constraints have narrower uncertainties, while others are a lot larger.

I've made clear for other constraints (heat, Paleo-ECS and RWF) that they are idealized - they illustrate what the effect would be on our confidence in the event that we knew that data.

Lines 70-75: Please discuss the uncertainty in the observational parameters that are used to constrain the model output. Also, perhaps include a figure showing the observation-based priors used.

The confidence in temperatures is now covered in the context of the discussion of the selection of sigma_T, and informed by the range of observed warming seen in the Cowtan-Way ensemble.

Do historical emissions include emissions from land use change? If so, the uncertainty on cumulative emissions is much larger than the uncertainty resulting from observed temperature.

Confidence in cumulative emissions is replicated from Millar/Friedlingstein 2018 - itself informed by uncertainty estimates in GCP2016. This is represented in the model with a scaling parameter on emissions, which is calibrated to represent this uncertainty.

Also, the discussion regarding constraints from the historical record could use the following reference and a short discussion:
Millar, R. J. Friedlingstein, P. The utility of the historical record for assessing the transient climate response to cumulative emissions.Phil. Trans. R. Soc. A 376, 20160449 (2018).

Well noted -this is now discussed in the introduction

References:

Cowtan, K. et al. Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. Geophysical Research Letters 42, 6526–6534 (2015).
Richardson, M., Cowtan, K., Hawkins, E. Stolpe, M. B. Reconciled climate response estimates from climate models and the energy budget of Earth. Nature Climate Change 6, 931 (2016).
Richardson, M., Cowtan, K. Millar, R. J. Global temperature definition affects achievement of long-term climate goals. Environ. Res. Lett. 13, 054004 (2018).

3. TCRE definition, non-CO2 forcing and the effective TCRE hysteresis

Please note that the definition of TCRE should be applied to CO2-induced warming alone. If calculating carbon budgets directly from RCP scenarios that are subject to CO2 and non-CO2 forcing, please refer to the Effective TCRE (Matthews et al. 2016). The current version of the manuscript confuses these two concepts, referring to TCRE even if non-CO2 forcing is present, making the arguments difficult to follow, since the effective TCRE, per definition, is not necessarily linear, due to the non-linearities arising from non-CO2 forcing. This should be clarified throughout the text.

Point well taken, and apologies for this confusion. I have made efforts to clarify the definitions throughout.
Based on earlier studies (e.g. MacDougall et al. 2015; Tokarska et al. 2019), I would expect that the apparent hysteresis behaviour depends on non-CO2 forcing scenario, and I am not convinced that observational constraints address this non-linearity.
I have now noted these papers - but my paper is exploring uncertainties which are not present in these studies. MacDougal 2015 considers only a single feedback timescale (i.e. constant sensitivity parameter) - therefore is omitting the major development considered here. Tokarska 2019 uses only a single model configuration in which thermal effects and carbon cycle nonlinearities cancel to produce a near constant TCRE - but the results are not generalised to all possible configurations of the model, Neither study is addressing the key issue here - whether historical temperatures can constrain the free parameters of model which allows for feedbacks on multiple timescales.

Furthermore, if considering TCRE to CO2-emissions alone (with no non-CO2 influence), TCRE would likely be fully reversible (no hysteresis)- e.g. see Figure 2a in MacDougall et al. 2015.

Figure S3 shows the response to CO2 emissions alone for each posterior parameter distribution. The effective TCRE is clearly different to the TCRE, as would be expected, but the hysteresis behaviour is not strongly influenced by the non-CO2 forcing - arising primarily from unresolved uncertainty on fraction of warming to date which is explained by slow timescale and fast timescale feedbacks.

Thus, regarding Figure 1, I would suggest discussing the effect on CO2-only response separately, as I suppose most of these non-linearities arises due to the specific nonCO2 emission scenarios, and is not necessarily an inherent property of TCRE alone. One way to address this issue would be to repeat the analysis using CO2-only simulations (according to RCP 2.6 scenarios), to illustrate if such hysteresis also arises in the absence of non-CO2 forcing.

Thanks for this suggestion - I have conducted the sensitivity study as suggested, which illustrates that although non-CO2 forcing assumptions do scale inferred TCRE, they do not play a strong role in hysteresis on a multi-century timescale (illustrated in supplemental figure S4). Scenarios with non-CO2 forcers set to zero show different apparent TCRE in RCP8.5 (as expected), but the hysteresis behaviour in RCP2.6 remains primarily a function of the choice of prior assumptions on the model behavior.

References:

MacDougall, A. H., Zickfeld, K., Knutti, R. Matthews, H. D. Sensitivity of carbon budgets to permafrost carbon feedbacks and non-CO2 forcings. Environ. Res. Lett. 10, 125003 (2015).

Tokarska, K. B., Zickfeld, K. Rogelj, J. Path independence of carbon budgets when meeting a stringent global mean temperature target after an overshoot. Earth's Future (2019).

Specific comments:

Lines 15-25: Please note that TCRE refers to CO2-only induced warming (originally defined in simulations where atmospheric CO2 concentrations increase at a rate of 1

Corrected as suggested

Lines 20-25: I found this sentence confusing and inaccurate: 'the range of TCRE values observed in Earth System Models (ESMs) can be used to infer model-based carbon budgets which are compatible with 1.5 and 2 degree Celsius targets of the Paris Agreement. . ." Is this referring to model-based TCRE that is used then in conjunction with other quantities (such as estimates of observed warming and future warming from non-CO2 forcing, as in Rogelj et al. 2019a framework) to infer remaining carbon budgets? Or is this sentence referring to carbon budgets at 1.5 C and 2.0C directly inferred from ESM output, as in AR5, for example? (in that case, those budgets already account for CO2 and non-CO2 warming in RCP scenarios, for example), but those budgets are not calculated directly from TCRE.

Paragraph completely rewritten in light of the reviewer's comment.

Lines 115-125: Since CO2 emissions follow different trajectories, but non-CO2 forcing follows the RCP 2.6 trajectory for each section, it is unclear how is the role of non-CO2 forcing distinct from CO2-induced changes? TCRE is pathway independent for CO2 emissions (both positive and negative), but if non-CO2 emissions are evolving in time (according to the RCP 2.6

scenario), the effective TCRE (to CO2 and non-CO2 forcing) is scenario-dependent, and heavily depends on the chosen non-CO2 scenario.
e.g. Reference: Mengis, N., Partanen, A.-I., Jalbert, J. Matthews, H. D. 1.5 ◦ C carbon budget dependent on carbon cycle uncertainty and future non-CO2 forcing. Sci Rep 8, 5831 (2018).

I've added a paragraph at the end of section 2.1 to discuss non-CO2 forcers. But the CO2-only experiments in Figure S3 show fairly convincingly that this is not the major factor in explaining the possibility for hysteresis without the RWF prior.

4. Robustness of TCRE under negative emissions

Please note that there are several recent studies using climate models of different complexity, including comprehensive ESMs and EMICs, that should be cited on lines 30-40. Currently, the paper gives an impression that this topic has not been studied in depth, while quite the opposite is true. Some discussion of these more recent studies would also be helpful on lines 25-40.

Regarding Earth system response to negative emissions in ESMs and EMICs:

Ehlert, D. Zickfeld, K. Irreversible ocean thermal expansion under carbon dioxide removal. Earth System Dynamics 9, 197–210 (2018).
Tokarska, K. B. Zickfeld, K. The effectiveness of net negative carbon dioxide emissions in reversing anthropogenic climate change. Environ. Res. Lett. 10, 094013 (2015).
Thanks - these are added in a new dedicated paragraph in the introduction on EMICs

Jones, C. D. et al. Simulating the Earth system response to negative emissions. Environmental Research Letters 11, 095012 (2016).
Mentioned in a dedicated paragraph on ESMs

Regarding TCRE behaviour under negative emissions:

Zickfeld, K., MacDougall, A. H. Matthews, H. D. On the proportionality between global temperature change and cumulative CO2 emissions during periods of net negative CO 2 emissions. Environ. Res. Lett. 11, 055006 (2016).

Noted in the EMIC discussion
Tokarska, K. B., Zickfeld, K. Rogelj, J. Path independence of carbon budgets when meeting a stringent global mean temperature target after an overshoot. Earth's Future (2019).
Noted in the EMIC discussion

MacDougall, A. H., Zickfeld, K., Knutti, R. Matthews, H. D. Sensitivity of carbon budgets to permafrost carbon feedbacks and non-CO2 forcings. Environ. Res. Lett. 10, 125003 (2015).
Noted in the EMIC discussion

Lines 30-35: '[TCRE] robustness in complex models under large negative emissions is relatively unexplored' – There are at least several recent studies that look at ESM model responses under different amounts of negative emission scenarios, and reversibility of TCRE after an overshoot (see several examples above).

Noted in the ESM discussion

Lines 80-85 claim that the TCRE relationship is not robust under negative emissions. However, it is unclear what fraction of this hysteresis behaviour is due to non-CO2 forcing. In intermediate-complexity model (UVic ESM) TCRE is reversible under negative CO2 emissions. At least a discussion of this claim in the context of these two following studies would be helpful here.

This paragraph has been highly restructured, with a more extensive literature review. I now note in the EMIC paragraph in the introduction that the net negative emission cumulative emission behavior is well tested in the U.Vic model. It is also noteworthy that the vast majority of the literature on the matter is conditional on the structural assumptions in a single EMIC - with very few studies formally sampling uncertain parameters of the model in tests of reversibility.

Lines 105-110: Figure 1b is not discussed in the previous section. I find it unconvincing why the TCRE framework would not hold under negative emissions even if model output is constrained by temperature and cumulative CO2 emissions (see major point above regarding reversibility). Please explain more your claim, possibly process-wise). Also, this hysteresis in the effective TCRE shown here may arise due to time-dependent non-CO2 forcing. Please see my comments below and in the above section 3 regarding the separation of CO2 and non-CO2 effects on the reversibility of TCRE.

As noted in the previous section, in this setup, the non-CO2 forcers are not a significant factor in the potential for hysteresis (the dominant factor being whether a prior is assumed for RWF or long term ECS). The model with the RWF prior does show more hysteresis in the all-forcing RCP2.6 compared with the CO2-only RCP2.6 (Figure S4(c)), but the choice of prior is by far the dominant constraint on hysteresis-like behavior in the model (compare Figure S4(a) with S4(c))

Line 80: This paragraph suggests that TCRE relationship is not robust under negative emissions. However, TCRE (due to CO2 emissions alone, as originally defined) has shown to be reversible in overshoot scenarios with negative emissions, including RCP scenarios (see major point 3 above).

I would argue it has not been shown to be true in a general sense - it has been demonstrated, for the most part, to hold in the UVic model, with limited evidence on shorter timescales in ESMs, and for simple climate models making strong structural assumptions on feedback timescales.

The non-linearity probably arises due to time-varying non-CO2 forcing. This should be clarified here, As for the point above - supplemental plot S3 shows that the potential for hysteresis arises from the timescale dynamics of the thermal response, not from non-CO2 forcing assumptions.

and please refer to effective TCRE, if non-CO2 forcing is included.

corrected

Also, this behaviour depends on non-CO2 forcing scenario, and I am not convinced that observational constraints address this

See response to major point 3 and figure S3

Non-linearity.

Lines 85 to 90: I would suggest discussing the effect on CO2-only response separately, as I suppose most of these non-linearities arises due to the specific non-CO2 emission scenarios, and is not necessarily an inherent property of TCRE alone.

See response to major point 3 and figure S3

Line 107: 'cumulative emissions framework is not guaranteed to hold under negative emissions' – This is a strong statement, which I am not convinced about. I would expect non-CO2 forcing in RCP 2.6 to be responsible for this hysteresis, and if considering RCP 2.6 CO2-only simulation, this hysteresis effect would be a lot smaller, if at all Present?

See response to major point 3 and figure S3

5. Committed warming (after emissions reach net-zero):

The paper refers to peak warming occurring after the emissions reach net-zero. However, there is no discussion with the literature on the committed warming occurring after emissions are stopped, which is directly relevant to the carbon budgets framework. For example, a short mention in the introduction (e.g. lines 25-40) and a discussion of how this paper fits within earlier studies would be valuable.

Ehlert, D. Zickfeld, K. What determines the warming commitment after cessation of CO 2 emissions? Environ. Res. Lett. 12, 015002 (2017).

MacDougall, A. H. et al. Z. M. MacDougall, A.H, Frölicher, T.L., Jones, C.D., Rogelj,J., Matthews, H.D., Zickfeld K., Arora, V.K., Barrett, N.J., Brovkin, V., Burger, F.A., Eby, M., Eliseev, A.V., Mokhov, I.I., Hajima, T., Holden, P.B., Jeltsch-Thömmes,A., Séférian, R., Michou, M., Shaffer, G., Sokolov, A., Wiltshire, A., Ziehnand, T., Menviel, L. How much warming remains in the pipeline? A multi-model analysis of the CO2 zero emission commitment. (discussion paper/ in review). https://www.biogeosciencesdiscuss.net/bg-2019-492/

Thanks for these suggestions - the papers are now discussed in the introduction.

Also, regarding the parameter choices and how they influence TCRE – perhaps it would be valuable to discuss the results of this study in the context of earlier studies, for example:

MacDougall, A. H., Swart, N. C. Knutti, R. The Uncertainty in the Transient Climate Response to Cumulative CO2 Emissions Arising from the Uncertainty in Physical Climate Parameters. J. Climate 30, 813–827 (2016).
Thanks - paper noted in introduction.

Lines 170-175: I would suggest also discussing the zero-emission commitment (see examples above), which suggests, that on average, ZEC is close to zero for CO2 emission pathways (in ESMs).
Noted - for decadal timescales - but the focus of this study is century timescales. The MacDougal paper itself shows a large diversity of response on century timescales (Figure 3a, https://www.biogeosciences-discuss.net/bg-2019-492/) , with only a subset of models having performed integrations long enough to assess the long term response.
Also, part of the difference between the threshold exceedance and avoidance budgets may be non-CO2 forcing, which is not part of the TCRE relationship (it would be part of the effective TCRE, which, however, is not expected to be linear due to non-CO2 influence). Please see above major comments regarding framing.
I've now noted this point in the discussion.

Line 60: Please note that recent studies suggest that the peak warming after emissions are zeroed is likely to be close to zero (see examples above)
Note that the focus of this study is on century timescales - where the MacDougal 2019 review shows a large diversity of sign and magnitude of post-cessation warming.

6. Long-term uncertainties in the context of overall carbon budget uncertainties

Lines 215-220: It would be good to put it in the context of other uncertainties on carbon budgets (see IPCC SR Ch2, Table 2.2). I would expect that other uncertainties such as permafrost carbon cycle feedbacks, model response to non-CO2 forcings, and nonCO2 forcing scenario uncertainty are still the dominant sources of uncertainness in the remaining carbon budgets.

Thanks for this suggestion. I've quantified the impact of the prior assumption uncertainty on 2100 budgets in the results section, added a new figure S12 to illustrate budgets as a function of time in each experiment. I've also added a discussion paragraph to consider these uncertainties in the context of other factors, as you suggest, tabulated in SR1.5 Table 2.2

Minor suggestions:

Title: I would suggest for the title to reflect more that the scope of this paper is also focusing on carbon budgets and TCRE framework
Fair point.

Title is now: The role of prior assumptions in carbon budget calculations

Abstract: The abstract gives the impression that the main source of uncertainties for near-term policy decisions is future negative emissions capacity and the long-term response to climate forcers as the main sources of uncertainty in the near-term policy decisions. However, carbon budgets and related net-zero emissions targets are subject to much larger transient uncertainty from the future non-CO2 forcing (at the time of 1.5C or target warming level), and climate models' response to non-CO2 forcing, which varies largely among models, contributing to a large spread in the remaining carbon budgets. (See IPCC Special Report, Chapter 2, Table 2.2. therein for quantification of different sources of uncertainties). While non-CO2 forcing is not the main scope of this paper, I would suggest revising this framing to avoid the misconception about the key sources of uncertainties in the remaining carbon budgets and near-term emission Targets.

Thanks for this - and I agree, non-CO2 forcing uncertainty should be discussed in the abstract.

Abstract: 'definite cumulative emissions budget' -I would suggest following terminology from Rogelj et al. 2019a framework, for consistency with other studies. Please specify if that is referring to the total or remaining budget?

Done - abstract rewritten to support the Rogelj model (a paper which came out during the late stages of writing)

Lines 90-100: While it is an interesting discussion, it is unclear how it relates to the transient timescales shown on Figure 1. (For example, it would be interesting to see the emulator behaviour until year 3000, for example, to assess the effect of non-linearities discussed in this study).

Added Supplemental Figure S16 to show evolution up to year 3000 for all simulations.

Figure 2 c. I found this plot confusing, and it is unclear what the baselines are. Following the SR1.5 and Rogelj et al 2019b recommendations, I would suggest plotting only the warming since 2006-2016, and offset it (as in SR 1.5 Table 2.2), so that the 1.5C and 2.0C target levels are clearly readable,

Done. x- and y-axis now anomalies from 2010.

Are cumulative emissions since 2010 or since 2020? (the figure caption and x-axis labels are inconsistent or confusing).

Now 2010 throughout.

Similarly, I suggest using the present-day warming baseline (as in SR 1.5), for consistency, in the whisker plots.

Done.

Perhaps, to clarify the point of this figure, it would be also useful to show whisker plots for the remaining carbon budgets at the time when 1.5C and 2.0C target is reached for the first time (before the overshoot), which would help to illustrate the difference in the transient and long-term budgets.

Thanks for this suggestion. Done.

However, they are not expected to be the same due to the ongoing non-CO2 forcing contribution. This point would need to be clarified as well.

As noted in response to major point 3 - this is a factor, but a secondary one to the choice of prior on thermal response parameters.

Lines 130-135: I found this paragraph unclear and confusing. It seems to be comparing carbon budgets calculated from scenarios that non-CO2 forcing is constantly evolving over time, with carbon budget estimates directly inferred from TCRE, but how that latter estimate accounts for the future contribution from non-CO2 forcing?

I've removed the TCRE estimate comparison from the text and Figure 2d, now just citing the SR1.5.

Lines 135-155 and Figure 3: It is unclear how the budgets can be compared for the different time periods, given that they entail different levels of non-CO2 forcing that is evolving in time in the simulations considered in this study. (i.e. since those budgets depend on the future non-CO2 forcing levels that differ, how can they be compared in a like-for-like manner?)

The non-CO2 emissions are fixed here at RCP2.6 emission levels, but the forcing is allowed to vary as an uncertain parameter in the model configuration - and that uncertainty is represented in the vertical spread of the distribution of points in Figure 3. I've made efforts to make this clearer in the text.

Lines 200-205: Perhaps a brief discussion in the context of more recent literature would be interesting (e.g. see Rogelj et al. 2019)

Thanks - restructured such that the end of the discussion explicitly supports the Rogelj framing.

Lines 205-215: Please note that carbon budgets should be calculated from anthropogenic warming estimate (Rogelj et al. 2019b. Haustein et al. 2017), which is not subject to internal variability. Reference: Haustein, K. et al. A real-time Global Warming Index. Scientific Reports 7, 15417 (2017).

I disagree that the Haustein estimate is not subject to internal variability, for reasons I laid out in this realclimate piece:

http://www.realclimate.org/index.php/archives/2017/10/1-5oc-geophysically-impossible-or-not/

The thesis of which was that the anthropogenic warming estimate is itself particularly sensitive to temperatures in the last few years of the timeseries, and repeating the approach in a large ensemble produces a distribution of forced warming estimates varying by over 0.2K. As such, I'm reluctant to recommend a regression approach as the preferred means of assessing forced warming trends for carbon budgeting.

That said - in the context of the present paper, I agree that the regression reconstruction would be more skillful by mid-century - and have cited these references in this context.

Response to reviewer 2

Summary: This study explores the long-term warming of climate for heavily mitigated scenarios, discussing the reasons behind changes in the sensitivity of warming to cumulative carbon emitted over time. An efficient model ensemble is generated and integrated with both prescribed concentration pathways (e.g. Figure 1), and an adaptive pathway algorithm to generate warming scenarios that restore towards the desired warming targets (e.g. Figure 2).

The study discusses how the required mitigation efforts are affected by changes in the sensitivity of surface warming to cumulative carbon emitted over time (the TCRE), and how a constant TCRE framework may not be able to account for such effects.

I found the study, as currently written, difficult to place in the context of existing literature. In particular, this study is missing comparisons to existing literature using observationally-constrained ensembles to explore the future carbon budget, adaptive mitigation pathways, and the time evolutions of effective climate sensitivity and transient climate response to emission.

Before such comparisons are made, it is difficult to say what in this study is new, and how it fits within existing knowledge.

Many thanks to the reviewer,

I have made efforts to clarify the framing in the revised version. The central focus of the study, and novel aspects are:

1 - the adequacy of TCRE-based carbon budgeting for temperature stabilization targets and the potential for hysteresis in the cumulative carbon-temperature relationship.
2 - how including different types of common prior assumptions (which vary across the literature) in the Bayesian model can alter the appropriateness of meta-frameworks for policy such as carbon budgeting.

The use of adaptive pathways and the simple model itself were never intended to by the novel aspects of this study (the adaptive mitigation pathway concept follows previous studies - Sanderson et al 2016, Sanderson et al 2017). That said, I apologise for missing important comparisons in the methodological aspects of the study. I've endeavoured to place the revised version in the context of the studies raised - which are certainly relevant.

In particular, the Goodwin studies are very relevant and make good efforts to quantify the effects of prior assumptions - but they do not focus on the question of reversibility and hysteresis. However, there are a number of studies and approaches in common usage which make stronger structural assumptions (fixed lambda or RWF) which are used to justify the cumulative emissions budgeting framework. The point of this study is to examine those assumptions, and how they influence model dynamics if imposed.

Major points: Significant areas of existing literature missing from discussion.

1. Observation-constrained ensembles exploring the future carbon budget

The study as currently written is missing comparisons to existing literature on observation-constrained calculations of future carbon budget for this century (Goodwin et al., 2018a) and out to year 2300 (Goodwin et al., 2018b) generated using the WASP model.

Many apologies for these omissions. I now discuss the papers in both results and introduction.

More details are required for the method used here for generating an observationally constrained ensemble. A full methodology needs to be presented containing prior assumptions, observational constraints and how the observational constraints are applied.
The method has been expanded, and now incorporated into the main text body.

This method used here should then be compared to the Monte Carlo plus history matching method presented in Goodwin et al. (2018a).

This is a good point. I now discuss the Goodwin paper at the start of the optimization chapter - because it represents an important design choice. In FAIR, at least - the Python code requires seconds to run - which made MCMC quite impractical to optimize the model (though a brief inspection suggests that in WASP, as compiled C++ code, might be fast enough).

The solution I proposed was to recode the core elements of the FAIR as a pulse-response model, fast enough that full MCMC is feasible to estimate the posterior (i.e. thousands of simulations per second) - which avoids the design issues of history matching (difficulty in automation of the selection of cutoff values and sampling uniformly from the posterior).

But - I accept that my approach has downsides, the model is an accurate representation of the FAIR core dynamics - but every aspect of the model and forcing must be represented in the pulse/response framework, which might limit further developments (e.g. a more detailed breakdown of non-CO_2 forcers

Note that both Goodwin et al. (2018a) and (2018b) studies adopt an efficacy on the ocean heat uptake, which is equivalent to allowing the effective climate sensitivity to change over time (where the ocean heat uptake efficacy is greater than 1 the effective climate sensitivity in the

present day is less than the equilibrium climate sensitivity on multi-century timescales, and where the efficacy is less than 1 the effective climate sensitivity for the present day is greater than the equilibrium value). As such, both studies allow the effective climate sensitivity to vary implicitly over time, and do not assume a relationship between TCR and ECS like the FaIR model studies (an approach which this study critiques).
I've now explicitly noted in the discussion that WASP does not contain this prior assumption.

The fact that no relationship is assumed between TCR and ECS in this study is currently discussed, but to assess whether the method is novel is needs to be compared to the WASP model methodology, for which this is also true.
The novelty in this study is not the omission of the TCR/ECS relationship - but the point that including it or not has large implications on the robustness of the TCRE framework.

2. Comparison to adaptive mitigation pathway algorithms in the literature

The Goodwin et al (2018b) study in Earth's Future presents and uses an adaptive mitigation pathway approach to restore a large ensemble of observation-constrained efficient model simulations to 1.5 and 2.0 ◦ C targets – in a very similar manor to the results presented in this study in Figure 2. Given the similarity of the method, the results in this study should be compared to this existing Goodwin et al (2018b) study in the literature. The 'Adaptive scenario design' used here (Figure 2, Appendix A1.2) should be compared to the 'Adaptive Mitigation Pathway' algorithm presented in Goodwin et al (2018b) and used in Brown et al. (2018) and Nicholls et al (2018). The resulting model output in this study, for compatible carbon emission pathways, should then be compared to the similar output generated in these previous studies in the literature.

I have now cited Goodwin 2018b as also using an adaptive scenario design. However, the approach considered here follows our earlier works, which predate the Goodwin studies (Sanderson et al 2016, used in Sanderson et al 2017) - which detail the methodology for adaptive mitigation pathways used here. As such, I didn't consider this aspect of my present study to be novel. I am simply applying an established approach to produce idealized pathways. Apologies if this was unclear in the previous version.
There are, however, some interesting distinctions between the two approaches. Sanderson 2016 is 'forward looking' - i.e. scenario parameters are iteratively adjusted until targets are met, while Goodwin 2018b represents to some degree the state of knowledge of decision makers based on observed TCRE. I've noted this at the start of Section 2.2.

3. Comparisons to existing literature on the reasons behind continued warming after emissions cease and the non-constancy of the TCRE

This study is missing comparisons to existing literature:

on the reasons behind continued surface warming after emissions cease (e.g. Frölicher et al., 2014; Williams et al., 2017a)

These are now cited in the introductory paragraphs on ZEC.

on the reasons behind near-constancy (or otherwise) of the TCRE (e.g. Goodwin et al., 2015; Williams et al., 2017b).

I now cite both papers in the introduction in the paragraph introducing TCRE.

One of the potential reasons that TCRE is non-constant is a change in effective climate sensitivity. However, other possible reasons are discussed in the studies mentioned here. The reasons behind nonconstancy of the TCRE in the model simulations presented here should be quantified in a way that relates to previous studies in the literature such as these.

This is an interesting point. The pulse-response framing here and the gradient framing of Goodwin 2015 potentially give two different perspectives on the role of the ocean which are perhaps complementary. In the Goodwin framework - the TCRE at a given time is decomposed into the dependence of surface warming on radiative forcing, the fractional dependence of radiative forcing from atmospheric and the dependence of radiative forcing from atmospheric CO 2 on carbon emissions, each of which evolve in time and whose interaction explains why TCRE remains constant (or not).

The Greens function/pulse response method used here could provide a different framing, with discrete responses to emissions which emerge on different exponentially decaying timescales. The model is already defined in terms of the set of exponential decays which describe p_CO2 as a function of emissions, and a second set which define warming as a function of p_CO2. The response parameters do not evolve in time because the time evolution is coded into the pulse response. As such - one can clearly see the effect of a prior on timescales of model response (e.g. imposing the RWF prior suppresses the long timescale sensitivity parameter).

I see these two frameworks as being complementary, the advantage of this approach being that it provides fixed parameters which can then be estimated for the real world (rather than time-evolving diagnostics) - and have added a paragraph to this effect in the discussion. A comprehensive use of the Greens function description as a framework for decomonsing system response to be compared the Goodwin framework is beyond scope in an already long paper, given it’s not the focus - but it would make for an interesting followup which I’d be keen to work on.

4. Constraints on the time-evolution of effective climate sensitivity in the literature

One effect leading to a change in TCRE over different response timescales is that the effective climate sensitivity also evolves over different response timescales. Again, here the discussion is missing sections of the literature.

For example, Goodwin (2018) generates a large ensemble of model simulations with explicitly time-evolving effective climate sensitivity, and then uses historic observations to constrain how effective climate sensitivity evolves on different response timescales. This reveals an increase in effective climate sensitivity over time (Goodwin, 2018) that will, crucially for this study, affect future mitigation requirements for a given warming target. Rohling et al. (2018) presents a similar account of time-evolving effective climate sensitivity from a paleo-perspective. These studies should be discussed and the results of this study compared to these previous findings.

I now discuss this study in some length in the discussion as a proposal for future work. Broadly - the non-stationarity of EffCS and TCRE mean that attempts to quantify these parameters from observations should be qualified - and that an alternative is to explicitly calibrate carbon and thermal feedbacks on different timescales. To do this properly requires future work - specifically considering additional information (other than global mean evolution) which might be used to constrain system response at different timescales.

It should also be noted that unlike the FaIR model methodology that is currently discussed for comparison (e.g. section 2, Lines 170-185), the WASP model methodology in Goodwin et al. (2018a; 2018b) and Goodwin (2018) does not assume a prior relationship between TCR and ECS or a near-constant TCRE. Therefore, comparisons to the WASP methodology deserve a separate discussion in section 2.

Now noted - I broadly cite the Goodwin 2018 approach as being an example of using a geological prior on ECS, in contrast to FAIR which uses a prior on RWF.

References:

Sanderson BM, O'Neill BC, Tebaldi C. What would it take to achieve the Paris temperature targets?. Geophysical Research Letters. 2016 Jul 16;43(13):7133-42.

Sanderson BM, Xu Y, Tebaldi C, Wehner M, O'Neill BC, Jahn A, Pendergrass AG, Lehner F, Strand WG, Lin L, Knutti R. Community climate simulations to assess avoided impacts in 1.5 and 2 C futures. Earth System Dynamics. 2017 Sep 19;8(3):827-47.

Relevant Changes to Manuscript

- Methods now inline with document
- Title altered following suggestion of reviewer 1
- Clarified framing of paper in response to reviewer 2
- Significantly extended literature review, notably Goodwin papers.
- Added discussion of Monte-Carlo vs. history matching approaches
- Discussed prior adaptive mitigation scenarios - including Sanderson 2016 (followed in this study) and Goodwin 2018
- Clarification of results in context of wider literature and other uncertainties, using IPCC SR1.5 framing
- Clarification of model design, and relationship to FAIR
- Added discussion of observational uncertainties and introduced parameter to represent uncertainty in historical land use emissions
- Changed observational carbon estimates to GCP2019
- Changed observational temperature estimate to Cowtan and Way
- Added new figure S3 to illustrate uncertainty in land use emissions
- Added new Supplemental Figure S4 to show the impact of non-CO2 forcers on hysteresis, and associated discussion in main text
- Expanded literature discussion of negative emission commitments and committed warming
- Added supplemental FIgure S16 to show millennial carbon budget implications

~~Constraints on long term warming~~ The role of prior assumptions in ~~a climate mitigation scenario~~carbon budget calculations

Benjamin Sanderson[1,2]

[1]CERFACS, Toulouse, France

[2]National Center for Atmospheric Research, Boulder CO, USA

Correspondence: Benjamin Sanderson (sanderson@cerfacs.fr)

Abstract. Cumulative emissions budgets and net-zero emission target dates are often used to frame climate negotiations (Frame et al., 2014; Millar et al., 2016; Van Vuuren et al., 2016; Rogelj et al., 2015b; Matthews et al., 2012). However, their utility for near-term policy decisions is confounded by an uncertainties in future negative emissions capacity (Fuss et al., 2014; Smith et al., 2016; Larkin et al., 2018; Anderson and Peters, 2016), the role of non-CO2 forcers (MacDougall et al., 2015) and in long term Earth System response to ~~climate forcers (Rugenstein et al., 2019; Knutti et al., 2017; Armour, 2017)which~~ forcing (Rugenstein et al., 2019; Knutti et al., 2017; Armour, 2017). Such uncertainties may impact the utility of an ~~indefinite~~ absolute carbon budget if peak temperatures occur significantly after net zero emissions are achieved, the likelihood of which ~~in a simple model is~~ is shown here to be conditional on prior assumptions about the long term dynamics of the Earth System. ~~Here we illustrate that the risks associated with near term positive emissions can be framed using a definite cumulative emissions budget with a 2040 time horizon, allowing~~ In the context of these uncertainties, we show that the necessity and scope for negative emissions deployment later in the century ~~to be better informed by observed warming over coming decades~~can be conditioned on near term emissions, providing support for a scenario framework which focuses on emissions reductions rather than absolute budgets(Rogelj et al., 2019b).

Introduction

The climate policy discussion has adopted some convenient frameworks which act as proxies for the drivers and consequences of climate change. For example, it is broadly assumed that climate risks scale with global mean temperature (O'Neill et al., 2017). International climate agreements have thus been framed in this context (United Nations, 2015), necessitating Earth system parameters which relate future emissions trajectories to temperatures. This relationship is often framed through the Transient Climate Response to cumulative carbon Emissions (TCRE - the ratio of the globally averaged transient CO_2 induced surface temperature change per unit carbon dioxide emitted, ~~Rogelj et al. (2019a); Allen et al. (2009); Millar et al. (2016); Matthews et al. (20~~ ~~). This convenient relationship allows the direct translation of temperature targets into available carbon budgets and derives from the apparent~~ (Rogelj et al., 2019a; Allen et al., 2009; Millar et al., 2016; Matthews et al., 2009; Gillett et al., 2013)).

This near linear relationship between cumulative emissions and surface temperatures is seen in many climate simulations ~~(England et al., 2009). Under the TCRE framework, the emissions budget constraint is dependent on a single parameter and the~~

~~range of TCRE values observed in Earth System Models (ESMs)can be used to infer model-based carbon budgets which are compatible with the 1.5 and 2 degree Celsius targets of the Paris agreement (Gillett et al., 2013)~~on decadal to century timescales provides a basis for cumulative carbon budgets corresponding to temperature targets (England et al., 2009; Gillett et al., 2013) , though its application to real-world carbon budgets is complicated by the effect of non-CO2 forcers. The "effective TCRE" (Matthews et al., 2017a) is thus the warming rate per unit carbon dioxide emitted in a scenario where forcers other than CO2 are acting on the system (such as aerosols and other greenhouse gases), which adds some uncertainty to the estimation of carbon budgets (Mengis et al., 2018; Rogelj et al., 2015a).

~~Meanwhile, understanding~~ Understanding of how the Earth System reaches equilibrium in response to climate forcing has advanced ~~; recent~~ in recent years; a number of studies have highlighted that existing 150 year simulations are insufficiently short to assess the Equilibrium Climate Sensitivity (ECS, the equilibrium response of surface temperatures to a doubling of carbon dioxide concentrations) of General Circulation Models, and assuming a single feedback parameter associated with Effective Climate Sensitivity (Gregory et al., 2004) can lead to a significant underestimation of long term response (Gregory and Andrews, 2016; Geoffroy et al., 2013; Senior and Mitchell, 2000; Winton et al., 2010; Armour et al., 2013; Li et al., 2013; Rose et al., 2014; Andrews et al., 2018).

What is less clear at present is whether these findings have any relevance for the use of ~~(effective)~~ TCRE in emissions policy decisions. ~~Though the cumulative-emissions temperature relationship is relatively~~ The TCRE framework is robust in transient scenarios in which emissions remain mostly positive ~~(Zickfeld et al., 2012; Krasting et al., 2014; Herrington and Zickfeld, 2014) , its robustness in complex models~~ (Zickfeld et al., 2012; Krasting et al., 2014; Herrington and Zickfeld, 2014; Goodwin et al., 2015) , and its value can be to some degree constrained by emissions and observed temperatures to date - even in the context of observational uncertainties (Millar and Friedlingstein, 2018). This path-independence has been explained by the fact that both heat and carbon are absorbed into the ocean on similar timescales, the former acting to realize warming in response to forcing while the latter reduces the forcing itself (Williams et al., 2016).

However, the robustness of temperature-cumulative emissions scaling in Earth System Models under large negative emissions ~~is relatively unexplored~~ on longer timescales is less well understood (Boucher et al., 2012; Vichi et al., 2013; Cao and Caldeira, 2010). Although an experimental design to test the long term robustness of TCRE under zero or negative emissions (Jones et al., 2019) ~~or the dynamics of equilibrium response to forcing (Rugenstein et al., 2019)~~ have been proposed and would be highly valuable, only a small selection of Earth System Models have ~~not generally~~ performed this type of experiment to date~~. Though such experiments have been performed in simple climate (Ricke and Caldeira, 2014; Millar et al., 2017c) and some intermediate complexity models (Zickfeld and Herrington, 2015) where~~ , finding large uncertainties in land and ocean carbon sinks (Jones et al., 2016) and in the long-term dynamics of equilibrium response to forcing (Rugenstein et al., 2019).

Earth systems model of intermediate complexity (EMICs) allow a more compuationally tractable integration of long timescale changes and in these cases, cumulative emissions-temperature proportionality has been ~~observed, it rests to thoroughly test whether these findings arise due to oversimplified modelstructure or prior assumptions on model parameters~~found to be relatively insensitive to emissions pathway (Zickfeld and Herrington, 2015; Tokarska and Zickfeld, 2015; Tokarska et al., 2019a; Zickfeld . However, many of these results are conditional on the structural assumptions of a single EMIC: the U.Vic Model (Weaver et al., 2001)

. Within this structure, parametric sensitivities for TCRE itself have been comprehensively tested (MacDougall et al., 2017) and reversibility in the U.Vic model has been tested to a degree (Ehlert and Zickfeld, 2018), but uncertainties remain in these results due to structural assumptions and parametric choices in the U.Vic model.

Simple climate models allow for very fast simulations which are capable of wide-scale parameter searches, but in many cases results are still subject to structural assumptions. For example, a fixed climate feedback parameter (Ricke and Caldeira, 2014; MacDougall a or a prior constraint on fraction of equilibrium warming which has already been realized to date (Millar et al., 2017c). These assumptions have been called into question by recent advances in understanding on Earth System response timescales (Rugenstein et al., 20 . Other models are less structurally constrained, but assume prior information on the equilibrium climate sensitivity of the real world (Goodwin et al., 2018b). The effect of this set of assumptions on the TCRE framework has not been assessed.

A number of studies have considered the "Zero Emission Warming Commitment" (ZEC), or the warming expected after emissions cease. This quantity can potentially be positive or negative in different models (MacDougall et al., 2020; Ehlert and Zickfeld, 201 and modifications to the cumulative emissions/ carbon budgeting framework have been proposed ~~(Rogelj et al., 2019a) to allow for additional corrections for non-CO_2 forcings (Rogelj et al., 2015a),~~ (Rogelj et al., 2019a; Froelicher and Paynter, 2015) to allow continued post-zero emissions temperature evolution ~~(Jones et al., 2019)~~ and unforeseen earth-system feedbacks or ’tipping-points’ which change biosphere or climate feedbacks (Brook et al., 2013). An complementary framework proposes a policy framework focused on net zero emissions and associated peak warming (Rogelj et al., 2019b). However, these frameworks are most useful if the zero emissions commitment is a small and finite correction to the net carbon budget, which is only true if peak warming occurs within a small number of decades of net-zero emissions.

Aside from physical modeling uncertainties in the long term stability of the TCRE assumption, indefinite carbon budgeting in policy making requires the combination of the effects of near term emissions reductions (Knutti et al., 2016; Rogelj et al., 2016a; Eom et al., 2015) and long term carbon removal technology which is subject to large socioeconomic, technological and physical uncertainties (Fuss et al., 2014; Smith et al., 2016; Larkin et al., 2018).

Similarly, the framing of climate policy in terms of a net zero emissions target also combines decarbonization of infrastructure (of which some sectors are highly difficult (Bataille et al., 2018)) and mid-century negative emissions capacity. These two components are conceptually different~~-~~; the former is at least partly a function of structural choices which are currently available, while the latter is conditional on deeply uncertain biophysical (Smith et al., 2016), technological (Lomax et al., 2015) and social (Anderson and Peters, 2016) factors.

Here, we consider long term emissions scenarios in a simple model informed by recent advances in understanding in the thermal response of the Earth system to climate forcing on a range of timescales (Armour et al., 2013; Geoffroy et al., 2013; Winton et al., 2010; Held et al., 2010; Proistosescu and Huybers, 2017; Rugenstein et al., 2016), and how prior assumptions on model parameters have an impact on the long term robustness of a cumulative carbon emissions budget and the possible commitment to long term negative emissions to maintain a stable climate. We discuss the plausibility of ~~the Earth exhibiting hysteresis behavior of~~ hysteresis in global mean temperature as a function of cumulative emissions and ~~that peak warming may occur~~ of peak warming occurring significantly after net zero emissions have been achieved.

Finally, we propose that a policy approach which relies primarily on indefinite carbon budgets is not useful in the light of large geophysical and socioeconomic uncertainties, and that more robust decisions can be made if near term mitigation priorities are decided independently from absolute commitments on long term negative emissions capacity, which can be revised later (Rogelj et al., 2019b). Furthermore, we show that global temperature evolution on the timescale of the mid 21st century would enable a better constraint on future negative emissions requirements for temperature stabilization.

1 ~~Results~~Methods

1.1 ~~Can transient observations constrain model response?~~Model Description

We first consider to what degree historical observations can constrain the long term coupled carbon-climate evolution of the Earth System. ~~To address this, we consider a~~ In order to produce a posterior parameter distribution conditioned on observations (and thus uncertainties in system response), there are various strategies (Emerick et al., 2011).

Our approach here is employ Bayesian calibration, a Markov-Chain Monte Carlo (MCMC) optimization (Goodman and Weare, 2010) in which a posterior parameter distribution is iteratively calculated by such that the sample density is representative of an underlying likelihood function. This approach is generally considered as an accurate approach but the number of model iterations required is often too computationally demanding to be practical (Oliver and Chen, 2011).

Computationally efficient alternatives include "History Matching" approaches which rule out members of a random sample which are not consistent with observations (Goodwin et al., 2018b; Williamson et al., 2013), an approach which can approximate the posterior in a computationally efficient manner subject to careful treatment of stochastic errors and prior assumptions (Liu et al., 2003). However, in the present study, the use of MCMC is made feasible through the use of a fast two timescale thermal response model ~~, with timescales of response representing the deep ocean and shallow ocean response (as in Proistosescu and Huybers ( ). This~~ (comparable to those used in Proistosescu and Huybers (2017); Geoffroy et al. (2013); Smith et al. (2018); Millar et al. (2017c)).

The thermal model in FAIR represents temperatures as a combination of two components with fast and slow inherent timescales:

$$\frac{dT_n}{dt} = \frac{q_n F - T_n}{d_n}; T = \sum_n T_n; n = 1, 2, \tag{1}$$

where T_n is global mean temperature and for each timescale n. T_n is the component of warming associated with that timescale, q_n is the feedback parameter and d_n is the response timescale. We consider the heat flux into the shallow and deep ocean to be functions of the same timescale:

$$R_n = r_n(F - T_n/q_n); R = \sum_n R_n; \sum_n r_n = 1; n = 1, 2 \tag{2}$$

where r_n is an efficacy factor for heat absorbed by the deep ($n = 1$) or shallow ($n = 2$) ocean, which sum to unity given the boundary condition that $R(0) = F(0) = F_{4xCO2}$ at $t = 0$ (allowing just one degree of freedom r_1 - the fraction of heat which is allocated to deep ocean storage).

The thermal model is made sufficiently fast for MCMC calibration using the particular solution to the step-change in forcing, which can be convoluted with a generic forcing timeseries to provide a general solution (Ruelle, 1998; Ragone et al., 2016; Lucarini et al., 2 . The particular solutions for temperature and radiation response to a step change in forcing F_{4xCO2} at time $t = 0$ can be expressed as a sum of exponential decay functions:

$$T_p(t) = F_{4xCO2} \sum_{n=1}^{2} q_n (1 - exp(-t/d_n)) \tag{3}$$

$$R_p(t) = F_{4xCO2} \sum_{n=1}^{2} r_n (exp(-t/d_n)), \tag{4}$$

where $T_P(t)$ is the annual global mean temperature and $R_p(t)$ is the net top-of atmosphere radiative imbalance at time t, and F_{4xCO2} is the instantaneous global mean radiative forcing associated with a quadrupling of CO_2, taken here to be $3.7 Wm^{-2}$(Myhre et al., 2013).

The thermal model is coupled to a ~~simple~~ emissions driven pulse model (~~as in Myhre et al. (2013); Millar et al. (2017c); Smith et al. (201 , see additional material)~~ in which each unit of emitted carbon dioxide is allocated to one of four pools with its own representative decay time~~. We then ask whether~~). The carbon scheme has four atmospheric carbon pools R_i (where $i = 0..3$, following Myhre et al. (2013)) with dissipation timescales τ_i as detailed in Table 1. Each unit pulse of emissions is allocated to each of the four pools with a fraction a_i:

$$\frac{dR_i}{dt} = a_i E(t) - \frac{R_i}{\tau_i}, \tag{5}$$

for which the solution for a unit emissions pulse $\delta(t)$ can be written:

$$R_i(t) = a_i (1 - e^{-t/\tau_i}). \tag{6}$$

A generic emissions time-series $E(t)$ can then be expressed as a sum of discrete pulses, allowing the corresponding carbon pools $C_i(t)$ to be expressed as a sum of pulse-responses $R_i(t)$

$$C_i(t) = \int_0^t \frac{dE(t')}{dt} R_i(t - t') dt'. \tag{7}$$

Atmospheric CO_2 concentrations C are calculated as the sum of the four pools $C(t) = C_0 + \sum_i C_i$(t), and are converted into a radiative forcing estimate assuming the standard logarithmic relationship:

$$F(t) = \frac{F_{4xCO2}}{ln(4)} ln\left(\frac{C(t)}{C_0}\right) + f_r F_{aer} + F_{other}, \quad (8)$$

where f_r is a free parameter to allow scaling of aerosol forcing (conceptually allowing for forcing uncertainty in the historical timeseries), and $F_{otherAnt}$ is all other anthropogenic and and natural forcers (summed from (Meinshausen et al., 2011b)). The thermal response is calculated by expressing the numerical time derivative of the forcing timeseries $F(t)$ where the change in forcing in a given time-step in a given year $\Delta F(t')$ is $[F(t') - F(t'-1)]$. The forcing timeseries can thus be expressed a series of step functions, and T_p from equation A1 can be used to calculate the integrated thermal response.

$$T(t) = \sum_{t'=0}^{t} \Delta F(t') \sum_{n=1}^{2} q_n \left(1 - exp\left(\frac{-(t-t')}{d_n}\right)\right), \quad (9)$$

Heat fluxes into the deep $(D(t))$ and shallow $(H(t))$ ocean components are represented by numerical integration of the slow (n=1) and fast (n=2) pulse response components of $R_p(t)$ in Equation 4:

$$D(t) = r_1 \sum_{t'=0}^{t} \Delta F(t') exp\left(\frac{-(t-t')}{d_1}\right), \quad (10)$$

$$H(t) = (1-r_1) \sum_{t'=0}^{t} \Delta F(t') exp\left(\frac{-(t-t')}{d_2}\right), \quad (11)$$

This is again performed in a computationally efficient manner using MATLAB's 'filter' function.

1.1.1 Model Optimization

We then assess the degree to which the physical parameters of this simple model (detailed in Table 1) can be constrained by historical transient information. The Earth system configuration of the pulse model has time-series inputs emissions of CO_2, along with radiative estimates from Meinshausen et al. (2011b) of non-CO2 forcing agents. We optimize the thermal model parameters for 2 timescales, the carbon dissipation parameters for 4 pools and the non-CO2 forcing factor f_r.

Optimization is conducted with the Goodman and Weare (2010) MCMC implementation, using flat initial parameter distributions as shown in Table 1, 200 walkers and 50,000 iterations for each optimization. Cost functions are computed for global mean temperature (T), global CO_2 concentrations (C), Shallow Ocean Heat Content (H) and Deep Ocean Heat Content (D):

$$E_T = \sum_t \left(\frac{(T(t) - T_{GCM}(t))}{\sqrt{2}\sigma_T}\right)^2$$

$$E_C = \sum_t \left(\frac{(C(t) - C_{GCM}(t))}{\sqrt{2}\sigma_C}\right)^2, E_H = \sum_t \left(\frac{(H(t) - H_{GCM}(t))}{\sqrt{2}\sigma_H}\right)^2, E_D = \sum_t \left(\frac{(D(t) - D_{GCM}(t))}{\sqrt{2}\sigma_D}\right)^2, \quad (12)$$

Long name	Symbol	Default	Min	Max
Geological re-absorption fraction	a_0	0.26	0.1	.3
Deep ocean invasion/equilibration fraction	a_1	0.14	0.1	.3
Biospheric uptake/ocean thermocline invasion fraction	a_2	0.22	0.1	.3
Rapid Biospheric uptake/ocean thermocline invasion fraction*	a_3	n/a	n/a	n/a
Geological re-absorption timescale $(years)$**	τ_0	10^6	10^6	10^6
Deep ocean invasion/equilibration timescale $(years)$	τ_1	200	200	1000
Biospheric uptake/ocean thermocline invasion timescale $(years)$	τ_2	40	40	100
Rapid biospheric uptake/ocean mixed-layer invasion timescale $(years)$	τ_3	1	1	10
Thermal equilibration of deep ocean Sensitivity (KWm^{-2})	q_1	0	0	10*
Thermal adjustment of upper ocean Sensitivity (KWm^{-2})	q_2	0	0	10
Thermal equilibration of deep ocean timescale $(years)$	d_1	239	80	3000
Thermal adjustment of upper ocean timescale $(years)$	d_2	30	1	40
Fraction of forcing in deep ocean response	r_1	0	0.33	0.5
Fraction of forcing in upper ocean response	r_2	0	0.33	0.5
Non-CO2 Forcing ratio	f_r	0.7	1	1.3
Emissions scaling ratio	s_e	0.8	1	1.2

Table 1. A table showing default model parameter values and minimum and maximum values used in model optimization. *deep ocean thermal response is limited to zero for 2 timescale model. *a_3 is calculated as the $1-\sum_{i=1:3}(a_i)$

. **following Millar et al. (2017c), deep ocean carbon uptake timescale is not included in the optimization (the timescale is effectively infinite: sufficiently longer than the scenarios considered here for the a_3 pool to not absorb significant carbon).

where σ_T represents the confidence in observed temperature values. To estimate this value, we use 2000-2019 annual global mean temperature anomalies from 1850-1900 in the HadCRUT-CW 100 member observational ensemble, where σ_T is the standard deviation of 2000 point (20 years, and 100 ensemble members), which represents uncertainty due to both natural variability and observational processing uncertainties (Cowtan and Way, 2013; Cowtan et al., 2015).

For σ_C, we lack an unforced standard deviation estimate - so a normalization constant of $\sigma_C = 0.3ppm$ was chosen empirically to produce a ± 1 ppmv range in 2016 observed concentrations in the posterior distribution (though uncertainties in emissions are much larger, and represented with the emissions scaling parameter s_e.

Shallow and Deep Ocean heat uptake (in cases where they are used) is taken as the 0-300m and 300m+ heat content respectively in Zanna et al. (2019), with σ_H and σ_D taken as 1850-1950 standard deviations from the same dataset. Confidence estimates in these timeseries is not available, so σ_H and σ_D nominally represents uncertainty due to natural variability - so "C,T, Heat" results should be considered to be an idealized estimate of how ocean heat information could constrain models if we were confident in that information.

In the 'C, T constraint' case, optimization is conducted using $-E_T$ and $-E_C$ as log likelihoods in the MCMC optimizer, with parameter boundaries as listed in Table 1. The 'C, T, Heat constraint' case uses the sum of $-E_T$, $-E_C$, E_D and $-E_H$ cost functions. The 'C,T, paleo' case is implemented using the likely value and upper bound on Earth System Sensitivity from Goodman and Weare (2010) fit the median and 90th percentile of a gamma distribution for equilibrium. The 'C,T, RWF' constraint is implemented using a log-normal prior on Transient Climate Response with 5–95 percentiles of 1.0–2.5 K as in Millar et al. (2017c), and a Gaussian prior on RWF (the ratio between LTE and TCR) with mean 0.6, and 5th and 9th percentiles of 0.45 and 0.75. The emissions scaling parameter is subject to Gaussian prior which was adjusted such that uncertainty in 5-95% cumulative CO_2 emissions in 2016 reflects observational uncertainties. It was found empirically that a Gaussian prior with a mean scaling parameter of 1, and standard deviations of 0.1 well represented published uncertainties, largely attributable to uncertain land use emissions (Le Quéré et al., 2017; Millar and Friedlingstein, 2018) (see Figure S3).

2 Results

2.1 The impact of prior assumptions on carbon dynamics

We consider a number of different constraint assumptions on model parameters and how they influence the range of future projections under different scenarios (Figure 1). If the model parameters are conditioned only on historical emissions and temperature (Figure 1(a,b)), transient warming under continued positive emissions is well constrained, such that temperatures follow the ~~effective~~ TCRE relationship under a high emission scenario (RCP8.5, Riahi et al. (2011)) emissions. However, the relationship is not robust under long term negative emissions in a decarbonization scenario (RCP2.6, Van Vuuren et al. (2011)) where some model variants in the posterior parameter distribution allow hysteresis in which temperatures continue to rise over the following centuries under a regime of net negative emissions.

Adding information on historical deep and shallow ocean heat content (Zanna et al., 2019) does not significantly constrain the system (Figure 1(a,c)). However, assuming addition information about long term equilibrium climate sensitivity is known from paleo-climate data ~~(Royer et al., 2011)~~(Royer et al., 2011; Goodwin et al., 2018b), does provide constraint on the degree of possible hysteresis (Figure 1(d)) as does the assumption of a known Realized Warming Fraction (RWF, the fraction of present day warming relative to equilibrium warming associated with current forcing) which is a very strong constraint on ~~cumulative-emissions temperature proportionality~~TCRE-like behavior. This prior, used in Millar et al. (2017b) produces a model configuration in which a proportional relationship between cumulative emissions-temperature is robust during both positive and negative phases of the emissions scenario (Figure (Figure 1(e)).

This raises the question of the degree to which we are confident in our knowledge of the values of ECS and RWF. In Millar et al. (2017b), the RWF prior is derived from the observation that the Transient Climate Response (TCR, the warming at the time of CO_2 doubling in a transient simulation where CO_2 increases by 1 percent per year) and Effective Climate Sensitivity (EffCS) are correlated in the CMIP5 ensemble (Millar et al., 2015) (where EffCS is the estimation of equilibrium response through the linear extrapolation of temperature change as a function of net top of atmosphere radiative imbalance in an instantaneous CO_2 quadrupling experiment (Gregory et al., 2004)).

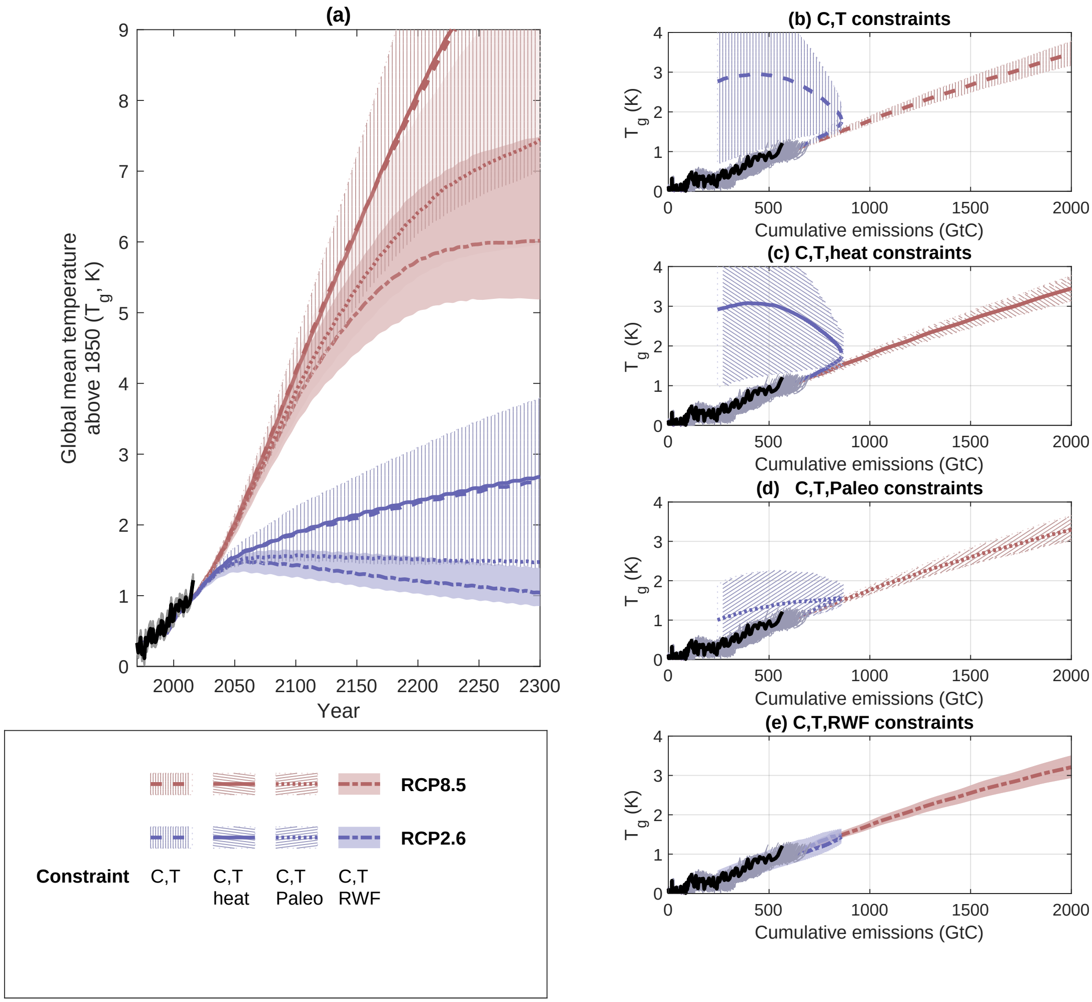


Figure 1. Posterior distributions of future global mean temperature projections constrained by 1850-2016 historical temperatures in a range of scenarios, priors and structural choices as a function of (a) time and (b-e) cumulative emissions of carbon (with 1000 years of climate evolution plotted from 1851-2850). Colored lines represent RCP8.5 (red) and RCP2.6 (blue). (b) and dashed lines in (a) show 2-timescale model posterior constrained using emissions (C) and temperature (T) only, (c) and solid lines in (a) are constrained using C,T and ocean heat content (H), (d) and dot-dash lines in (a) use C,T and RWF. (e) and dotted lines C,T and a paleoclimate prior on ECS. Shaded regions indicate the 10-90th percentile range. Solid black lines show observed ~~HadCRUT values~~ global mean temperature median estimate (Cowtan and Way, 2013) and ~~historical~~ most likely estimates of combined land use and fossil fuel emissions (Le Quéré et al., 2017). Grey lines show uncertainties in observed temperature-cumulative emissions following Millar and Friedlingstein (2018).

However, the Equilibrium Climate Sensitivity (ECS), realized over a multi-century to millennial timescale, is often significantly greater than the Effective Climate Sensitivity (Rugenstein et al., 2016; Knutti et al., 2017) and its value may not be well constrained by observed warming (Proistosescu and Huybers, 2017; Andrews et al., 2018). As such, and it is not apparent that the long-term ECS in a model like Myhre et al. (2013) can be constrained by TCR (with large implications for millennial-scale temperature evolution, as seen in Supplemental Figure S16).

These prior assumptions strongly impact the range of possible behavior under strong negative emissions in RCP2.6. However, under RCP8.5, the ensembles constrained by historical temperatures show a near-linear relationship between cumulative emissions and temperature, irrespective of prior assumptions and constraints used (Figure 1(b-e), red lines), this can be broadly understood by considering that in RCP8.5, radiative forcing continues to increase at current rates and thus long term warming is broadly a function of TCR, which is itself constrained by historical temperature evolution.

The scenarios considered here are multi-gas, with both CO2 emissions and non-CO2 forcers. As expected (Mengis et al., 2018) , non-CO2 forcing assumptions can alter the effective TCRE seen in transient RCP8.5 simulations and RCP2.6 projections on shorter timescales of less than a century (see Supplemental Figure S4), however the potential for hysteresis on longer timescales is similar in multi-gas and CO2 only experiments.

2.2 Implications for meeting Paris temperature targets

If we consider a 'high risk' world where ECS (and its relationship to TCR) is not independently constrained, corresponding to subplot (b) in Figure 1, the cumulative emissions framework is not guaranteed to hold under negative emissions, and the concept of an indefinite cumulative carbon budget associated with a temperature target may not be helpful for near-term carbon mitigation planning (results for other prior assumptions are shown in the additional material).

We illustrate this in some idealized cases~~where~~ , using adaptive scenarios in which emissions are adjusted in order to achieve 1.5 and 2 degree C climates are achieved post ~~2100.~~ 2100 (similar to those considered in Sanderson et al. (2016b, 2017); Goodwin et al. (20). The Sanderson et al. (2016b) approach allows iteration of scenarios such that targets can be met in almost all cases, but the optimization is "forward looking" (in contrast to Goodwin et al. (2018a), which simulates decisions made in response to observed warming to date without perfect knowledge of the future). Here, we follow a similar strategy to Sanderson et al. (2016b) , where scenarios are designed using a small number of parameters which are then optimized to meet a stabilization target post-2100.

Scenarios are conducted in 3 phases: before 2020 is the 'historical' period, where emissions follow RCP2.6 (which is broadly consistent with observations before 2020). Between 2020 and 2040, the 'uninformed' period, CO_2 emissions follow one of a range of linear mitigation pathways such that 2040 CO_2 emissions are chosen at random for each scenario, ranging from $0GtC/yr$ to $10GtC/yr$ (our focus here is on low emission futures, and we do not consider here futures where emissions increase post-2020).

Each ensemble member uses a single parameter set draw from the posterior distribution of models calculated during the MCMC constraint of model parameter space in Section 1.1.1. Emissions follow RCP2.6 from 1850 until 2020, after which CO_2 emissions are by a 'pchip' spline which is fixed at a number of points, the first of which are 2010 and 2020 RCP2.6

emissions - ensuring a smooth transition from the RCP time-series to the post-2020 timeseries. An 'uninformed' emissions trajectory takes place from 2020 to 2040, where emissions evolve from RCP2.6 2020 levels (10.26GtC/yr) to a 2040 emissions level drawn randomly from a uniform distribution with bounds at $0GtC/yr$ and $10GtC/yr$.

Post 2040, in the 'adaptive' period, an emission scenario is calculated iteratively to achieve temperature stabilization at a defined target post-2100, allowing for a temperature overshoot before 2100 with a large but finite lower limit on net negative emissions capacity in line with the largest negative emissions values seen in the integrated assessment literature for 1.5 degree temperature stabilization targets ($-20GtC/yr$, First (2018)). Non-CO_2 gas emissions follow RCP2.6 throughout the simulation in all cases (~~as such these scenarios cannot~~ clearly, these scenarios should not be treated as socioeconomically plausible scenarios, rather as idealized illustrations of Earth System Response to a range of forcing pathways).

Parametric control of the adaptive phase is achieved by specifying 3 time points (the first, tp_1 in the range 2060-2100, the second (tp_2) in the range 2101-2300 and the third tp_3 fixed at the end of the simulation in 2764. Each time point is associated with an emissions rate $Ep_{1,2,3}$ which are each weakly constrained to lie in the range -40 to +10 GtC/yr. Optimization uses MATLAB's fmincon algorithm to find optimal values of $tp_{1,2}$ and $Ep_{1,2,3}$, where the model is run iteratively for a given physical parameter set to find a solution which minimizes the RMSE from the desired annual mean global mean temperature timeseries target (1.5 or 2.0C, in this case) over the date range 2100-2500.

The temperature trajectories are illustrated in Figure 2(a). Each member of the posterior distribution of possible simple climate models in Figure 1(a,b) is then paired with a random 2020-2050 emissions reduction pathway and then a post-2050 emissions pathway is calculated to optimize for stabilization at 1.5 or 2 degrees post-2100. This framework allows us to consider what would be required for long term stabilization in a model configuration where the cumulative emissions-temperature relationship does not necessarily hold.

The resulting scenarios are idealized, some requiring a very rapid switch to large net-negative values after 2040 in order to stabilize temperatures at 1.5C (Figure 2(b)), and such rapid decarbonization may not be achievable in reality ~~(Sanderson et al., 2016a)~~ (Sanderson et al., 2016b), but we can learn some useful properties of the system response by studying the relationships between near term and long term emissions commitments. Non-CO2 emissions remain at RCP2.6 levels in all cases (though the non-CO2 forcing varies as a function of the f_r parameter).

The range of long-term emission trajectories for temperature stabilization is diverse (Figure 2(c)), ~~allowing for large positive or negative fluxes over the following centuries~~ in some cases ~~as global mean temperatures remain stable (by construction). This implies that although in nearly all cases, temperatures have stabilized by~~ requiring large negative emissions in the latter half of the 21st century to achieve temperature stabilization after 2100 (Figure 2(a))~~, the~~. The cumulative carbon budget plume allows for a 1.5C(2.0C) ~~budget of -250 to 200GtC (75 to 650GtC~~ post-2010 budget of -300 to 400GtC (0 to 900GtC) by 2100, a budget which continues to grow more uncertain over the centuries which follow (Figure 2(c,d)). ~~This is in contrast to the indefinite cumulative carbon budget for~~ Most of the 1.5~~or 2 degrees calculated from assumed effective TCRE - which is relatively tightly constrained as 160-200GtC (300-380GtC) for~~ C simulations overshoot the target in the latter half of the 21st century (Figure 2(a)), and post-2010 budget for initial exceedance of 1.5~~C (2.0C) stabilization after corrections for present day warming due to non-CO2 gases (Figure 2(d)~~C is more tightly constrained at 250-400GtC (most 2C simulations do not significantly overshoot).

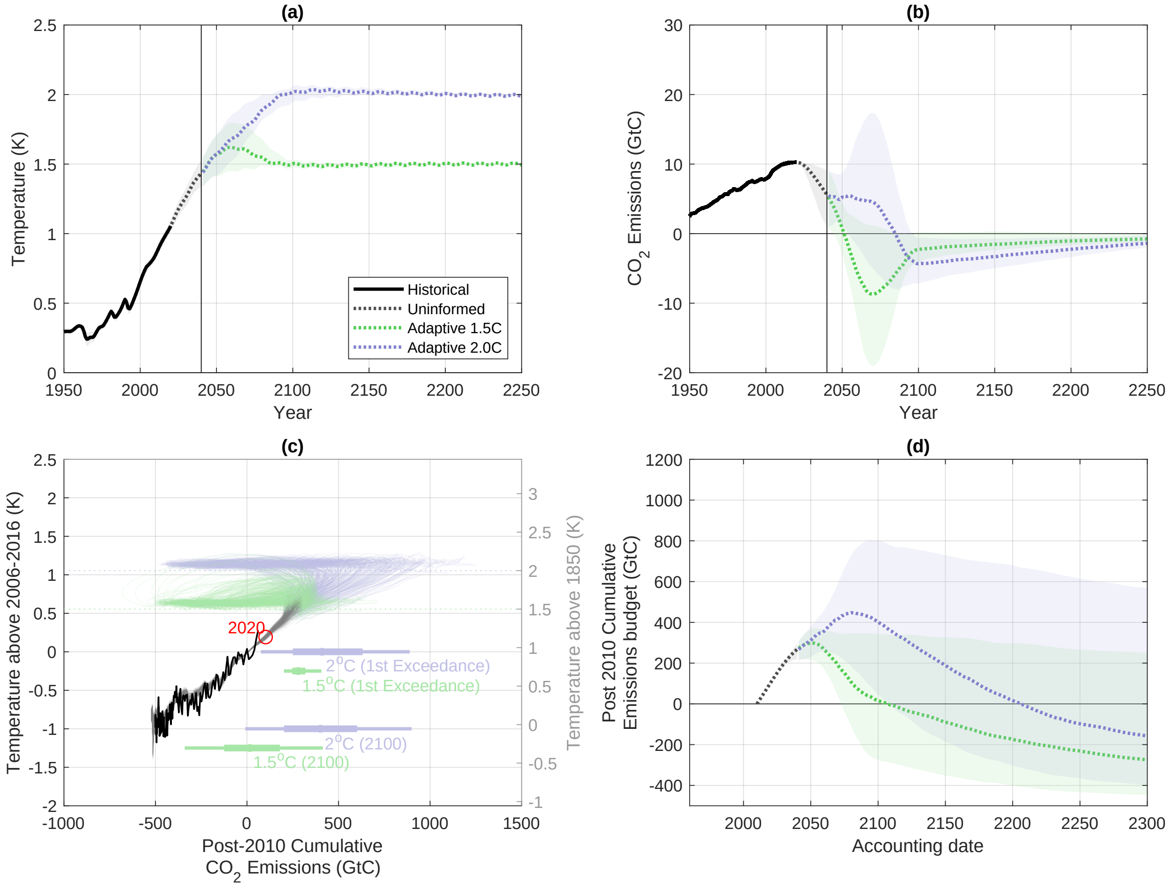


Figure 2. Plots showing idealized pathways to 1.5 or 2.0C temperature stabilization for an ensemble of coupled carbon-climate model configurations. (a) shows the global mean temperature as a function of time for 1.5 and 2.0C stabilization ensembles (b) shows emissions in the historical, uniformed and adaptive stages of the simulation (c) shows the global mean temperatures above ~~pre-industrial/2006-2015~~ 2006-2016 (left/right axis) levels as a function of post-2010 cumulative CO_2 emissions while (d) shows the cumulative carbon emissions total for ensemble members as a function of time. Shaded regions in (a,b,d) indicate 10th-90th percentile range of the ensemble distribution, while dotted lines shown the 50th percentile. Gray/blue/black areas refer to uninformed/adaptive for 2.0C/adaptive for 1.5C respectively. Box/whisker plots in (c) show the long term cumulative carbon budget assessed in 2100 for 1.5 and 2.0C stabilization from 1850-2500. Box/whisker plots in (d) show the ~~effective~~ TCRE estimate of carbon budget with (median shown by '+') and without (median shown by 'x') non-CO2 gas correction. Red circle shows ensemble mean warming and post-2010 cumulative emissions in 2020.

Figure 3. Plots showing (a) the relationship between mid-century cumulative carbon budgets and (b) mid-century warming and associated likelihoods of long term carbon removal requirements for temperature stabilization. (a) shows the ensemble relationship between the net carbon emitted between 2020 and 2040 (uninformed period in Figure 1) and the associated range of possible carbon removal required later in the century in the adaptive phase for 1.5C (green) and 2.0C (blue) stabilization. Filled circles represent an individual ensemble member, while shaded blue/green areas represent a moving estimate of the 10-90th percentile range of the 2.0C/1.5C distribution (solid blue/green lines are 2.0/1.5C median. (b) shows 2050-2100 allowable carbon budget as a function of 2050 warming above pre-industrial levels. Dots and shading show ensemble distribution as in (a). Horizontal box/whisker plots show 10th,25th,50th,75th and 90th percentiles of 2050 warming consistent with labeled 2020-2040 carbon budgets and the associated percentage reduction in 2040 emissions relative to 2020. Gray bar shows the range of reference 2100 net carbon budgets considered for end of century 1.5 degree overshoot scenarios in the IPCC spacial report on 1.5 degrees (First, 2018)).

This large uncertainty in the face of long term stabilization scenarios draws into question the utility of an indefinite carbon budget ~~, hence we~~ (in the case where we have no prior information on equilibrium response). We can consider to what degree we can constrain future response using a definite budget with a 2020-2050 timeframe (Figure 3). Firstly, even in the face of ~~response timescale uncertainty~~possible hysteresis of temperature as a function of cumulative carbon emissions, there is a linear relationship between 2020-2040 budgets and associated late century carbon removal rates required for stabilization (Figure 3(a)).

For example, if a late century net carbon emission of -2.9 GtC/yr is assumed for late century (corresponding to the central estimate of 1.5 degree, low overshoot stabilization from the IPCC Special Report on 1.5C warming (First, 2018), a 50 percent chance of 1.5 degrees requires a 2020-2040 budget of 150GtC, which would require a 60 percent cut in emissions from present day levels by 2040. A 75 percent chance of meeting the target would require a 2020-2040 budget of 100GtC - requiring just over 100 percent cut in carbon emissions by 2040.

Here again, the choice of prior constraint on model parameters has an important effect. If the Paleoclimate ~~(or RWF) constraint on ECS is usedas in Figure 1(c, or d)~~or RWF is used, a 75 percent chance of 1.5 degrees given ~~the~~ an assumed -2.9GtC/yr late century removal rate would allow a 160GtC(or 220GtC) budget from 2020-2040 (see Additional Material Figure S14(c,~~or d).~~ d)). Similarly, estimated carbon budgets become more consistent with TCRE derived estimates if an RWF prior is used, with 2100 budgets of 120-430 GtC (500-900GtC) for 1.5C (2.0C). This can be compared with the IPCC SR1.5 assessment of 115-230GtC (320-550GtC) respectively, which includes uncertainties in non-CO2 emissions and forcings and long timescale carbon cycle feedbacks.

~~However, in all cases, by~~ These findings support the framing of emissions policy in terms of near term emissions reductions rather than indefinite carbon budgets (Rogelj et al., 2019b). By mid 21st century, observed warming will provide a good indication of the degree of negative emissions required for stabilization - as the average realized warming in 2040-2060 provides quite a strong constraint on budgets for the latter half of the century (Figure 3(b)). The degree of possible mid-century warming can be reduced by minimizing the 2020-2040 carbon budget, but there still exists uncertainty due to the degree of thermal inertia in the system as greenhouse gas concentrations stabilize.

The strong relationship between mid-century warming and late century carbon removal requirements for 1.5 or 2.0C stabilization occurs because 2040-2060 warming can be potentially decreased either by fortuity (with a small value of real-world equilibrium climate sensitivity) or by action (by minimizing near-term emissions), both of which reduce late century net carbon removal requirements. Conversely, high climate sensitivity or slow decarbonization would both result in greater mid-century warming and greater necessity for negative emissions deployment.

3 Discussion

Recent climate policy discussions have been framed in the context of a carbon budget, an allowable net total of cumulative emissions which are consistent with a desired limit on planetary warming (Allen et al., 2009; Millar et al., 2016). Nuances in the estimation of this budget have been noted relating to bias correction of existing models (Millar et al., 2017a), the compensation for the effects of non-CO2 anthropogenic emissions ~~(Rogelj et al., 2015a)~~ (Rogelj et al., 2015a; MacDougall et al., 2015; Mengis et al., 2018) and the need for additional carbon fluxes for temperature stabilization after net-zero emissions have been achieved ~~(Rogelj et al., 2016b; Jon~~ ~~.~~

~~But in the current framework, these~~ (Rogelj et al., 2016b; Jones et al., in review; Mengis et al., 2018). These factors are deemed to be corrections to the ~~effective~~ TCRE-computed carbon budgets (Rogelj et al., 2019a), and value of ~~effective~~ TCRE informed by a combination of model response historical records of global surface temperatures (Gillett et al., 2013; Steinacher and Joos, 2016) form the basis for published model estimates on carbon budgets for temperature stabilization (Matthews et al., 2017a, a).

~~Limitations in the applicability of the cumulative-emissions/temperature relationship due to the response timescales of the Earth System have been noted before (Rogelj et al., 2019a) in terms of~~ It has been noted before that at any given time, the TCRE can be expressed as a product of 3 components: the the ~~discrepancy between~~ dependence of surface warming on radiative forcing, the fractional dependence of radiative forcing on atmospheric CO_2 and the dependence of atmospheric CO 2 on carbon emissions (Goodwin et al., 2015) - but each of these elements can potentially evolve in time as feedbacks are realized on different timescales (Rogelj et al., 2019a; Goodwin et al., 2018a). This has been addressed by introducing "Threshold Avoidance Budgets" and "Threshold Exceedance Budgets" (Rogelj et al., 2016b) which differ due to the lag of peak temperatures after net-zero emissions have been achieved as slower timescale components of the system equilibrate or due the effects of non-CO2 forcers. But, the scale of these effects is generally assumed to be small - on the order of 1-2 decades (Ricke and Caldeira, 2014; Zickfeld and Herrington, 2015). Idealized experiments to assess zero-emission warming commitment(MacDougall et al., 2020) in both EMICs and ESMs suggest the ZEC is small on a 50 year timescale but uncertain on a century timescale, with a large diversity of magnitude, sign and rate of warming post-cessation of emissions.

~~However, as we have seen, models can be~~ It has also been demonstrated that effective climate sensitivity likely evolves in time (Goodwin, 2018; Rohling et al., 2018), which will influence TCRE (Goodwin et al., 2015) and thus carbon budgets for a given temperature target (Goodwin et al., 2018b); thus attempts to quantify fixed real world estimates for TCRE or effective climate sensitivity must be qualified for long timescales (Rugenstein et al., 2019) or extended net negative emissions(Ehlert and Zickfeld, 2018)

. In this study, the pulse-response formulation allows for the idealized separation of process response both in the evolution of atmospheric CO2 in response to emissions and in the thermal response of the system to forcing, allowing the an illustration of how prior assumptions impact feedbacks on different timescales. Future work should consider further how these fixed parameters of the carbon-climate system can be further independently constrained and integrated with existing understanding of time-evolving net climate feedbacks.

We find that the pulse-response model is not constrained to follow ~~cumulative-emissions/temperature proportionality under negative emissions~~TCRE-like behaviour without prior knowledge of equilibrium climate sensitivity. Considering other simple models, such priors are often used (either explicitly or implicitly). The parameters of the FAIR (Millar et al., 2017c; Smith et al., 2018) simple climate model, for example, are constrained using a prior on RWF ~~. This constraint arises due to~~ (whereas projected uncertainty ranges using other models such as Goodwin et al. (2018b) use no such prior). The constraint in FAIR is justified with an observed relationship between *Effective* Climate Sensitivity and TCR in CMIP models, and is thus likely overly constraining on possible model behavior consistent with state of art GCMs (see Additional Material section S1).

Other models do not explicitly constrain RWF, but do constrain equilibrium climate sensitivity - the WASP model (Goodwin et al., 2015; considers a multiple timescales of response and a geological prior on equilibrium warming response to emissions, which acts to preclude the

Similarly, in the MAGICC model (Meinshausen et al., 2011a), non-stationary feedbacks are represented in two ways - using an allowance for an oceanic surface and and land surface feedback strengths, as well as having forcing dependent feedback strengths. However, ECS values calculated using MAGICC when calibrated as an emulator of CMIP GCM simulations remain very close to the Effective Climate Sensitivities of the target model (Meinshausen et al., 2011a) - even though in some cases we know that the true ECS realized in millennial time-frames is significantly greater than the EffCS value (Rugenstein et al., 2019). This requires further research, but is possibly explained by the consensus that multiple feedback timescales arise from warming patterns associated with shallow and deep ocean warming (Li et al., 2013; Geoffroy et al., 2013). Representing feedbacks as a function of the warming of the ocean *surface* warming is therefore a strong structural assumption which may not capture this effect.

~~Indeed, recent~~ Recent work has made clear that the long timescale response of the Earth system is not well constrained by past observations (Proistosescu and Huybers, 2017; Andrews et al., 2018), drawing into question whether recent transient warming is able to constrain Equilibrium Climate Sensitivity (Otto et al., 2013) or the Realized Warming Fraction (Millar et al., 2015). In the absence of these constraints, we cannot rule out without additional data that the slow timescale response of the Earth System associated with deep ocean warming may lead to a world which exhibits a (relatively) low TCR but a high ECS realized over centuries or millennia (Rugenstein et al., 2019) which, as we show here, may complicate the use of an indefinite carbon budget for temperature targets.

Here, we find that these factors result in large uncertainties on remaining carbon budgets until 2100, with the possibility of unless prior information is assumed on the value of ECS or RWF (Supplemental Figure S10). Using an RWF prior, carbon budgets for 1.5C and 2C are broadly consistent with TCRE-derived estimates in Rogelj et al. (2018), but removing this prior reduces the lower bound of the budget from positive 120GtC with a RWF prior (as assessed in 2100 for 1.5C stabilization) to negative 300GtC if the prior is removed. These factors are in addition to existing uncertainties arising from non-CO2 forcing and scenario assumptions (approximately ±200GtC in long term budgets) and uncertainties in pre-industrial temperatures (approximately ±100GtC in long term budgets) (Rogelj et al., 2018).

Other sources of information which may yet resolve the uncertainty. Independent information to constrain ECS from paleoclimate (Royer et al., 2011) or process understanding (Sherwood et al., 2014; Zhai et al., 2015; Tian, 2015; Tan et al., 2016; Cox et al., 2018) may help constrain the potential for temperature hysteresis. But many constraints to date have considered only *effective* climate sensitivity (Gregory et al., 2004) - whereas it is increasingly clear that both the timescale and amplitude of climate feedbacks need to be constrained in order to understand Earth System response to future forcing pathways (Armour et al., 2013). Such avenues could and should be explored further.

The pulse response model of the type used here is also a simplification of global response, albeit a commonly used one (Joos et al., 2013) - which resolves the degrees of freedom in the range of responses exhibited in physical Earth System Models. The anthropogenically forced warming in 2040-2060 would be subject to internal variability of order 0.1C (Dai et al., 2015; Rogelj et al., 20 which could potentially be improved with detection approaches (Haustein et al., 2017). As such, observed mid-century warming would be of some value in constraining negative emissions requirements later in the century which spans nearly 0.6C over the ensemble range (Figure 3(b)).

Clearly, the models used here are idealizations. Emission rates and rates of change are not constrained by technological or societal limitations, and only CO2 pathways are modified from the RCP2.6 scenario - and so results are only illustrative of how the Earth System might respond to different hypothetical pathways. Finding pathways for technology and policy which can actually achieve these pathways is a question for Integrated Assessment Models. However, the present standard approach of producing scenarios through forward-looking solvers (O'Neill et al., 2016) is unable to capture the risk highlighted here associated with actors who act today with imperfect knowledge about future technology (Fuss et al., 2014; Anderson and Peters, 2016) and Earth System response.

~~The pulse response model of the type used here is also a simplification of global response, albeit a commonly used one (Joos et al., 2013) - which resolves the degrees of freedom in the range of responses exhibited in physical Earth System Models. However, the model only resolves the central tendency of the long term equilibration of the Earth System to a forcing change, without any estimate of climate variability. The real-world climatological temperature in 2040-2060 would be subject to internal variability (Kay et al., 2015), but such variation in annual mean temperatures is of the order 0.1-0.2C (Rogelj et al., 2017), and decadal average deviations from climatology of global mean temperature due to internal variability are of order 0.1C (Dai et al., 2015), which implies that by 2060, observed mid-century warming will be of some value in constraining negative emissions requirements later in the century which spans nearly 0.7C over the ensemble range (Figure 3(b))~~This has led to a call to frame of policy in terms of near-term emissions which are compatible with projected peak levels of warming (Rogelj et al., 2019b).

~~In summary, even~~ The results of this study support this logic. Even in the presence of large uncertainty on long term response to emissions, near-term climate policy can be well posed through the use of a time-limited net carbon budget, or equivalently, a near-term commitment for a percentage reduction in emissions by a certain date (Sachs et al., 2016; Kaya et al., 2019). ~~Such a framework allows near-term emissions reduction requirements to broadly be considered separately from the negative emission fluxes required for temperature stabilization, the feasibility of which remains deeply uncertain (Fuss et al., 2014; Anderson and Peters, 2016), and does not require waiting for peak warming to occur (Rogelj et al., 2019b) in order to inform the required scale of negative emissions capacity (especially in the theoretical case where peak warming occurs significantly after net zero emissions are reached).~~

Observed warming over the coming decades will provide additional information on our commitments to implement negative emissions infrastructure for temperature stabilization - commitments which may or may not prove feasible to realize. But a near-term budget would provide decision-makers with the tools to assess the risk of failure to meet temperature targets as a function of clearly defined targets for near-term decarbonization.

Data availability. CMIP5 and CMIP6 data are available through a distributed data archive developed and operated by the Earth System Grid Federation (ESGF).

Code and data availability. Code for this study is available on Github at https://github.com/benmsanderson/matlab_pulse

Appendix: ~~Methods~~

A1 ~~Simple Climate Model Implementation~~

~~The temperature portion of the code allows for 2 representative temperatures, each with an equilibration timescale d_j (for 2 timescales, j following Myhre et al. (2013); Millar et al. (2017c)), producing a simple model for temperature and radiation response to a step change in forcing:~~

$$P(t) = F_{4xCO2} \sum_{n=1}^{3} q_n (1 - exp(-t/d_n))$$

$$R(t) = F_{4xCO2} \sum_{n=1}^{3} r_n (exp(-t/d_n)),$$

~~where $P(t)$ is the annual global mean temperature and $R(t)$ is the net top-of atmosphere radiative imbalance (used only for the calculation of Effective Climate Sensitivity), and F_{4xCO2} is the instantaneous global mean radiative forcing associated with a quadrupling of CO_2, taken here to be $3.7 W m^{-2}$(Myhre et al., 2013).~~

~~Constraining thermal parameters from historical temperatures and emissions requires a consideration of both the carbon cycle as well as other relevant climate forcers. MCMC optimization of even a simple model of this form requires 10^7 or more calculations, so a very rapid model is required for computational tractability. This study employs a fast pulse-response model~~

~~to represent the response of surface global mean surface temperatures to emissions. The model is implemented as a digital filter in MATLAB (see attached code) - allowing efficient computation and enabling Markov-Chain Monte Carlo parameter estimation for the physical parameters.~~

~~The carbon scheme is a simple pulse dissipation model, with four atmospheric carbon pools R_i (where $i = 0..3$, following Myhre et al. (2013)) with dissipation timescales τ_i as detailed in Table 1. Each unit pulse of emissions is allocated to each of the four pools with a fraction a_i:~~

$$\frac{dR_i}{dt} = a_i E(t) - \frac{R_i}{\tau_i},$$

~~for which the solution for a unit emissions pulse $\delta(t)$ can be written:~~

$$R_i(t) = a_i(1 - e^{-t/\tau_i}).$$

~~A generic emissions time-series $E(t)$ can then be expressed as a sum of discrete pulses, allowing the corresponding carbon pools $C_i(t)$ to be expressed as a sum of pulse-responses $R_i(t)$~~

$$C_i(t) = \int_0^t \frac{dE(t')}{dt} R_i(t - t')dt'.$$

~~Atmospheric CO_2 concentrations C are calculated as the sum of the four pools $C(t) = C_0 + \sum_i C_i$(t), and are converted into a radiative forcing estimate assuming the standard logarithmic relationship:~~

$$F(t) = 5.4 ln\left(\frac{C(t)}{C_0}\right) + f_r F_{ext}(t)$$

~~following Myhre et al. (2013), and all other forcings (aerosols, and non-CO2 greenhouse gases) are combined into a single term $F_{ext}(t)$ using global mean RCP values from Meinshausen et al. (2011b). Uncertainty in the amplitude of non-CO_2 forcings is simply represented simply by an uncertainty factor f_r, which is also optimized in the course of the MCMC calibration (Table 1). The thermal response is calculated by expressing the derivative of the forcing timeseries $F(t)$ as a series of step functions and using the CO2 quadrupling response T_p from equation A1 to calculate the integrated thermal response.~~

$$T(t) = \int_0^t \frac{\frac{dF}{dt}(t')}{F_{4xCO2}} T_p(t - t')dt'.$$

~~This is again performed in a computationally efficient manner using MATLAB's 'filter' function.~~

A0.1 ~~Model Optimization~~

The earth system configuration of the pulse model has time-series inputs emissions of CO_2, along with radiative estimates from Meinshausen et al. (2011b) of non-CO2 forcing agents. We optimize the thermal model parameters for 2 timescales $[\mathbf{q}, \mathbf{d}, \mathbf{r}]$, the carbon dissipation parameters $[\mathbf{a}, \tau]$ for 4 pools and the non-CO2 forcing factor f_r. Optimization, as for the 4xCO2 case is conducted with the Goodman and Weare (2010) MCMC implementation, using flat initial parameter distributions as shown in Table 1, 200 walkers and 50,000 iterations for each optimization. Cost functions are computed for global mean temperature (T), global CO_2 concentrations (C), Shallow Ocean Heat Content (H) and Deep Ocean Heat Content (D):

$$E_T = \sum_t \left(\frac{(T(t) - T_{GCM}(t))}{\sqrt{2}\sigma_T} \right)^2$$

$$E_C = \sum_t \left(\frac{(C(t) - C_{GCM}(t))}{\sqrt{2}\sigma_C} \right)^2, E_H = \sum_t \left(\frac{(H(t) - H_{GCM}(t))}{\sqrt{2}\sigma_H} \right)^2, E_D = \sum_t \left(\frac{(D(t) - D_{GCM}(t))}{\sqrt{2}\sigma_D} \right)^2,$$

where σ_T is defined as for the abrupt-CO2 case as the standard deviation of HadCRUT 1850-1950 values. For σ_C, we lack an unforced standard deviation estimate - so a normalization constant of $\sigma_C = 0.3ppm$ was chosen empirically to produce a ± 1 ppmv range in 2016 observed concentrations in the posterior distribution. Shallow and Deep Ocean heat is taken as the 0-300m and 300m+ heat content respectively in Zanna et al. (2019), with σ_H and σ_D taken as 1850-1950 standard deviations from the same dataset.

In the 'C, T constraint' case, optimization is conducted using $-E_T$ and $-E_C$ as log likelihoods in the MCMC optimizer, with parameter boundaries as listed in Table 1. The 'C, T, Heat constraint' case uses the sum of $-E_T$, $-E_C$, E_D and $-E_H$ cost functions. The 'C,T, paleo' case is implemented using the likely value and upper bound on Earth System Sensitivity from Goodman and Weare (2010) fit the median and 90th percentile of a gamma distribution for equilibrium. The 'C,T, RWF' constraint is implemented using a log-normal prior on Transient Climate Response with 5–95 percentiles of 1.0–2.5 K as in Millar et al. (2017c), and a Gaussian prior on RWF (the ratio between LTE and TCR) with mean 0.6, and 5th and 9th percentiles of 0.45 and 0.75.

A0.1 Adaptive scenario design

We propose an ensemble of simulations which achieve post-2100 stabilization at the 1.5 and 2.0C levels referred to in the Paris Agreement (United Nations, 2015). Each ensemble member uses a single parameter set draw from the posterior distribution of models calculated during the MCMC constraint of model parameter space in Section 1.1.1. Emissions follow RCP2.6 from 1850 until 2020, after which CO_2 emissions are by a 'pchip' spline which is fixed at a number of points, the first of which are 2010 and 2020 RCP2.6 emissions - ensuring a smooth transition from the RCP time-series to the post-2020 timeseries. An 'uninformed' emissions trajectory takes place from 2020 to 2040, where emissions evolve from RCP2.6 2020 levels (10.26GtC/yr) to a 2040 emissions level drawn randomly from a uniform distribution with bounds at $0GtC/yr$ and $10GtC/yr$.

~~Post 2050, the emissions are defined by an 'adaptive' phase - with 3 time points (the first, tp_1 in the range 2060-2100, the second (tp_2) in the range 2101-2300 and the third tp_3 fixed at the end of the simulation in 2764. Each time point is associated with an emissions rate $Ep_{1,2,3}$ which are each weakly constrained to lie in the range -40 to +10 GtC/yr. Optimization uses MATLAB's fmincon algorithm to find optimal values of $tp_{1,2}$ and $Ep_{1,2,3}$, where the model is run iteratively for a given physical parameter set to find a solution which minimizes the RMSE from the desired annual mean global mean temperature timeseries target (1.5 or 2.0C, in this case) over the date range 2100-2500.~~

~~Long name Symbol Default Min Max Geological re-absorption fraction a_0 0.26 0.1 .3 Deep ocean invasion/equilibration fraction a_1 0.14 0.1 .3 Biospheric uptake/ocean thermocline invasion fraction a_2 0.22 0.1 .3 Rapid Biospheric uptake/ocean thermocline invasion fraction* a_3 n/a n/a n/a Geological re-absorption timescale $(years)$** τ_0 10^6 10^6 10^6 Deep ocean invasion/equilibration timescale $(years)$ τ_1 200 200 1000 Biospheric uptake/ocean thermocline invasion timescale $(years)\tau_2$ 40 40 100 Rapid biospheric uptake/ocean mixed-layer invasion timescale $(years)$ τ_3 1 1 10 Thermal equilibration of deep ocean Sensitivity (KWm^{-2}) q_1 0 0 10* Thermal adjustment of upper ocean Sensitivity (KWm^{-2}) q_2 0 0 10 Thermal equilibration of deep ocean timescale $(years)$ d_1 239 80 3000 Thermal adjustment of upper ocean timescale $(years)$ d_2 30 1 40 Fraction of forcing in deep ocean response r_1 0 0.33 0.5 Fraction of forcing in upper ocean response r_2 0 0.33 0.5 Non-CO2 Forcing ratio f_r 0.7 1 1.3 A table showing default model parameter values and minimum and maximum values used in model optimization. *deep ocean thermal response is limited to zero for 2 timescale model. *a_3 is calculated as the $1-\sum_{i=1:3}(a_i)$. **following Millar et al. (2017c), deep ocean carbon uptake timescale is not included in the optimization (the timescale is effectively infinite: sufficiently longer than the scenarios considered here for the a_3 pool to not absorb significant carbon).~~

Author contributions. The author performed all analysis and writing for this project

Competing interests. The author declares no competing interests

Acknowledgements. This work is funded by the French National Research Agency, project number ANR-17-MPGA-0016. Benjamin Sanderson is an affiliate scientist with the National Center for Atmospheric Research, sponsored by the National Science Foundation.

References

Allen, M. R., Frame, D. J., Huntingford, C., Jones, C. D., Lowe, J. A., Meinshausen, M., and Meinshausen, N.: Warming caused by cumulative carbon emissions towards the trillionth tonne, Nature, 458, 1163, 2009.

Anderson, K. and Peters, G.: The trouble with negative emissions, Science, 354, 182–183, 2016.

Andrews, T., Gregory, J. M., Paynter, D., Silvers, L. G., Zhou, C., Mauritsen, T., Webb, M. J., Armour, K. C., Forster, P. M., and Titchner, H.: Accounting for changing temperature patterns increases historical estimates of climate sensitivity, Geophysical Research Letters, 45, 8490–8499, 2018.

Armour, K. C.: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks, Nature Climate Change, 7, 331, 2017.

Armour, K. C., Bitz, C. M., and Roe, G. H.: Time-varying climate sensitivity from regional feedbacks, Journal of Climate, 26, 4518–4534, 2013.

Bataille, C., Åhman, M., Neuhoff, K., Nilsson, L. J., Fischedick, M., Lechtenböhmer, S., Solano-Rodriquez, B., Denis-Ryan, A., Stiebert, S., Waisman, H., et al.: A review of technology and policy deep decarbonization pathway options for making energy-intensive industry production consistent with the Paris Agreement, Journal of Cleaner Production, 187, 960–973, 2018.

Boucher, O., Halloran, P. R., Burke, E. J., Doutriaux-Boucher, M., Jones, C. D., Lowe, J., Ringer, M. A., Robertson, E., and Wu, P.: Reversibility in an Earth System model in response to CO2 concentration changes, Environ. Res. Lett., 7, 024 013, https://doi.org/10.1088/1748-9326/7/2/024013, 2012.

Brook, B. W., Ellis, E. C., Perring, M. P., Mackay, A. W., and Blomqvist, L.: Does the terrestrial biosphere have planetary tipping points?, Trends in Ecology & Evolution, 28, 396–401, 2013.

Cao, L. and Caldeira, K.: Atmospheric carbon dioxide removal: long-term consequences and commitment, Environ. Res. Lett., 5, 024 011, https://doi.org/10.1088/1748-9326/5/2/024011, 2010.

Cowtan, K. and Way, R.: Coverage bias in the HadCRUT4 temperature record, QJR Meteorol. Soc, 2013.

Cowtan, K., Hausfather, Z., Hawkins, E., Jacobs, P., Mann, M. E., Miller, S. K., Steinman, B. A., Stolpe, M. B., and Way, R. G.: Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures, Geophysical Research Letters, 42, 6526–6534, 2015.

Cox, P. M., Huntingford, C., and Williamson, M. S.: Emergent constraint on equilibrium climate sensitivity from global temperature variability, Nature, 553, 319, 2018.

Dai, A., Fyfe, J. C., Xie, S.-P., and Dai, X.: Decadal modulation of global surface temperature by internal climate variability, Nat. Clim. Change, 5, 555–559, https://doi.org/10.1038/nclimate2605, 2015.

Ehlert, D. and Zickfeld, K.: What determines the warming commitment after cessation of CO2 emissions?, Environmental Research Letters, 12, 015 002, 2017.

Ehlert, D. and Zickfeld, K.: Irreversible ocean thermal expansion under carbon dioxide removal, Earth System Dynamics, 9, 197–210, 2018.

Emerick, A. A., Reynolds, A. C., et al.: Combining the ensemble Kalman filter with Markov chain Monte Carlo for improved history matching and uncertainty characterization, in: SPE Reservoir Simulation Symposium, Society of Petroleum Engineers, 2011.

England, M. H., Gupta, A. S., and Pitman, A. J.: Constraining future greenhouse gas emissions by a cumulative target, Proceedings of the National Academy of Sciences, 106, 16 539–16 540, 2009.

Eom, J., Edmonds, J., Krey, V., Johnson, N., Longden, T., Luderer, G., Riahi, K., and Van Vuuren, D. P.: The impact of near-term climate policy choices on technology and emission transition pathways, Technological Forecasting and Social Change, 90, 73–88, 2015.

First, P. J.: Global warming of 1.5 C An IPCC Special Report on the impacts of global warming of 1.5 C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, IPCC, 2018.

Frame, D. J., Macey, A. H., and Allen, M. R.: Cumulative emissions and climate policy, Nature Geoscience, 7, 692, 2014.

Froelicher, T. L. and Paynter, D. J.: Extending the relationship between global warming and cumulative carbon emissions to multi-millennial timescales, Environmental Research Letters, 10, 075 002, 2015.

Fuss, S., Canadell, J. G., Peters, G. P., Tavoni, M., Andrew, R. M., Ciais, P., Jackson, R. B., Jones, C. D., Kraxner, F., Nakicenovic, N., et al.: Betting on negative emissions, Nature climate change, 4, 850, 2014.

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D., and Tytéca, S.: Transient climate response in a two-layer energy-balance model. Part II: Representation of the efficacy of deep-ocean heat uptake and validation for CMIP5 AOGCMs, Journal of Climate, 26, 1859–1876, 2013.

Gillett, N. P., Arora, V. K., Matthews, D., and Allen, M. R.: Constraining the ratio of global warming to cumulative CO2 emissions using CMIP5 simulations, Journal of Climate, 26, 6844–6858, 2013.

Goodman, J. and Weare, J.: Ensemble samplers with affine invariance, Communications in applied mathematics and computational science, 5, 65–80, 2010.

Goodwin, P.: How historic simulation–observation discrepancy affects future warming projections in a very large model ensemble, Climate Dynamics, 47, 2219–2233, 2016.

Goodwin, P.: On the time evolution of climate sensitivity and future warming, Earth's Future, 6, 1336–1348, 2018.

Goodwin, P., Williams, R. G., and Ridgwell, A.: Sensitivity of climate to cumulative carbon emissions due to compensation of ocean heat and carbon uptake, Nature Geoscience, 8, 29–34, 2015.

Goodwin, P., Brown, S., Haigh, I. D., Nicholls, R. J., and Matter, J. M.: Adjusting mitigation pathways to stabilize climate at 1.5 C and 2.0 C rise in global temperatures to year 2300, Earth's Future, 6, 601–615, 2018a.

Goodwin, P., Katavouta, A., Roussenov, V. M., Foster, G. L., Rohling, E. J., and Williams, R. G.: Pathways to 1.5 C and 2 C warming based on observational and geological constraints, Nature Geoscience, 11, 102–107, 2018b.

Gregory, J., Ingram, W., Palmer, M., Jones, G., Stott, P., Thorpe, R., Lowe, J., Johns, T., and Williams, K.: A new method for diagnosing radiative forcing and climate sensitivity, Geophysical Research Letters, 31, 2004.

Gregory, J. M. and Andrews, T.: Variation in climate sensitivity and feedback parameters during the historical period, Geophysical Research Letters, 43, 3911–3920, 2016.

Haustein, K., Allen, M., Forster, P., Otto, F., Mitchell, D., Matthews, H., and Frame, D.: A real-time global warming index, Scientific reports, 7, 1–6, 2017.

Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F., and Vallis, G. K.: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing, Journal of Climate, 23, 2418–2427, 2010.

Herrington, T. and Zickfeld, K.: Path independence of climate and carbon cycle response over a broad range of cumulative carbon emissions, Earth Syst. Dyn., 5, 409–422, https://doi.org/10.5194/esd-5-409-2014, 2014.

Jones, C., Frölicher, T., Koven, C., MacDougall, A., Matthews, D., Zickfeld, K., Rogelj, J., and Tokarska, K.: ZEC-MIP: Quantifying the Zero Emissions Commitment, GMD, https://doi.org/https://doi.org/10.5194/gmd-2019-153, in review.

Jones, C. D., Ciais, P., Davis, S. J., Friedlingstein, P., Gasser, T., Peters, G. P., Rogelj, J., van Vuuren, D. P., Canadell, J. G., Cowie, A., et al.: Simulating the Earth system response to negative emissions, Environmental Research Letters, 11, 095 012, 2016.

Jones, C. D., Frölicher, T. L., Koven, C., MacDougall, A. H., Matthews, H. D., Zickfeld, K., Rogelj, J., Tokarska, K. B., Gillett, N., Ilyina, T., Meinshausen, M., Mengis, N., Seferian, R., and Eby, M.: The Zero Emission Commitment Model Intercomparison Project (ZECMIP) contribution to CMIP6: Quantifying committed climate changes following zero carbon emissions, Geosci. Model Dev. Discuss., pp. 1–18, https://doi.org/10.5194/gmd-2019-153, 2019.

Joos, F., Roth, R., Fuglestvedt, J., Peters, G., Enting, I., Bloh, W. v., Brovkin, V., Burke, E., Eby, M., Edwards, N., et al.: Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis, Atmospheric Chemistry and Physics, 13, 2793–2825, 2013.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S., Danabasoglu, G., Edwards, J., et al.: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability, Bulletin of the American Meteorological Society, 96, 1333–1349, 2015.

Kaya, Y., Yamaguchi, M., and Geden, O.: Towards net zero CO 2 emissions without relying on massive carbon dioxide removal, Sustainability Science, pp. 1–5, 2019.

Knutti, R., Rogelj, J., Sedláček, J., and Fischer, E. M.: A scientific critique of the two-degree climate change target, Nature Geoscience, 9, 13, 2016.

Knutti, R., Rugenstein, M. A., and Hegerl, G. C.: Beyond equilibrium climate sensitivity, Nature Geoscience, 10, 727, 2017.

Krasting, J. P., Dunne, J. P., Shevliakova, E., and Stouffer, R. J.: Trajectory sensitivity of the transient climate response to cumulative carbon emissions, Geophys. Res. Lett., 41, 2520–2527, https://doi.org/10.1002/2013GL059141, 2014.

Larkin, A., Kuriakose, J., Sharmina, M., and Anderson, K.: What if negative emission technologies fail at scale? Implications of the Paris Agreement for big emitting nations, Climate policy, 18, 690–714, 2018.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., et al.: Global carbon budget 2017, Earth System Science Data Discussions, pp. 1–79, 2017.

Li, C., von Storch, J.-S., and Marotzke, J.: Deep-ocean heat uptake and equilibrium climate response, Climate Dynamics, 40, 1071–1086, 2013.

Liu, N., Oliver, D. S., et al.: Evaluation of Monte Carlo methods for assessing uncertainty, SPE Journal, 8, 188–195, 2003.

Lomax, G., Lenton, T. M., Adeosun, A., and Workman, M.: Investing in negative emissions, Nature Climate Change, 5, 498, 2015.

Lucarini, V., Ragone, F., and Lunkeit, F.: Predicting climate change using response theory: Global averages and spatial patterns, Journal of Statistical Physics, 166, 1036–1064, 2017.

MacDougall, A. H. and Friedlingstein, P.: The origin and limits of the near proportionality between climate warming and cumulative CO2 emissions, Journal of Climate, 28, 4217–4230, 2015.

MacDougall, A. H., Zickfeld, K., Knutti, R., and Matthews, H. D.: Sensitivity of carbon budgets to permafrost carbon feedbacks and non-CO2 forcings, Environmental Research Letters, 10, 125 003, 2015.

MacDougall, A. H., Swart, N. C., and Knutti, R.: The uncertainty in the transient climate response to cumulative CO2 emissions arising from the uncertainty in physical climate parameters, Journal of Climate, 30, 813–827, 2017.

MacDougall, A. H., Frölicher, T. L., Jones, C. D., Rogelj, J., Matthews, H. D., Zickfeld, K., Arora, V. K., Barrett, N. J., Brovkin, V., Burger, F. A., Eby, M., Eliseev, A. V., Hajima, T., Holden, P. B., Jeltsch-Thömmes, A., Koven, C., Menviel, L., Michou, M., Mokhov, I. I., Oka, A., Schwinger, J., Séférian, R., Shaffer, G., Sokolov, A., Tachiiri, K., Tjiputra, J., Wiltshire, A., and Ziehn, T.: Is there warming in the pipeline?

A multi-model analysis of the zero emission commitment from CO_2, Biogeosciences Discussions, 2020, 1–45, https://doi.org/10.5194/bg-2019-492, https://www.biogeosciences-discuss.net/bg-2019-492/, 2020.

Matthews, H. D., Gillett, N. P., Stott, P. A., and Zickfeld, K.: The proportionality of global warming to cumulative carbon emissions, Nature, 459, 829, 2009.

Matthews, H. D., Solomon, S., and Pierrehumbert, R.: Cumulative carbon as a policy framework for achieving climate stabilization, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 370, 4365–4379, 2012.

Matthews, H. D., Landry, J.-S., Partanen, A.-I., Allen, M., Eby, M., Forster, P. M., Friedlingstein, P., and Zickfeld, K.: Estimating carbon budgets for ambitious climate targets, Current Climate Change Reports, 3, 69–77, 2017a.

Meinshausen, M., Raper, S. C., and Wigley, T. M.: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6–Part 1: Model description and calibration, Atmospheric Chemistry and Physics, 11, 1417–1456, 2011a.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M., Lamarque, J.-F., Matsumoto, K., Montzka, S., Raper, S., Riahi, K., et al.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic change, 109, 213, 2011b.

Mengis, N., Partanen, A.-I., Jalbert, J., and Matthews, H. D.: 1.5 C carbon budget dependent on carbon cycle uncertainty and future non-CO 2 forcing, Scientific reports, 8, 1–7, 2018.

Millar, R., Allen, M., Rogelj, J., and Friedlingstein, P.: The cumulative carbon budget and its implications, Oxford Review of Economic Policy, 32, 323–342, 2016.

Millar, R. J. and Friedlingstein, P.: The utility of the historical record for assessing the transient climate response to cumulative emissions, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376, 20160 449, 2018.

Millar, R. J., Otto, A., Forster, P. M., Lowe, J. A., Ingram, W. J., and Allen, M. R.: Model structure in observational constraints on transient climate response, Climatic Change, 131, 199–211, 2015.

Millar, R. J., Fuglestvedt, J. S., Friedlingstein, P., Rogelj, J., Grubb, M. J., Matthews, H. D., Skeie, R. B., Forster, P. M., Frame, D. J., and Allen, M. R.: Emission budgets and pathways consistent with limiting warming to 1.5C, Nat. Geosci., 10, 741–747, https://doi.org/10.1038/ngeo3031, 2017a.

Millar, R. J., Fuglestvedt, J. S., Friedlingstein, P., Rogelj, J., Grubb, M. J., Matthews, H. D., Skeie, R. B., Forster, P. M., Frame, D. J., and Allen, M. R.: Emission budgets and pathways consistent with limiting warming to 1.5 C, Nature Geoscience, 10, 741, 2017b.

Millar, R. J., Nicholls, Z. R., Friedlingstein, P., and Allen, M. R.: 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, 7213–7228, 2017c.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and natural radiative forcing, pp. 659–740, Cambridge University Press, Cambridge, UK, https://doi.org/10.1017/CBO9781107415324.018, 2013.

Oliver, D. S. and Chen, Y.: Recent progress on reservoir history matching: a review, Computational Geosciences, 15, 185–221, 2011.

O'Neill, B. C., Tebaldi, C., Vuuren, D. P. v., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., et al.: The scenario model intercomparison project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461–3482, 2016.

O'Neill, B. C., Oppenheimer, M., Warren, R., Hallegatte, S., Kopp, R. E., Pörtner, H. O., Scholes, R., Birkmann, J., Foden, W., Licker, R., et al.: IPCC reasons for concern regarding climate change risks, Nature Climate Change, 7, 28, 2017.

Otto, A., Otto, F. E., Boucher, O., Church, J., Hegerl, G., Forster, P. M., Gillett, N. P., Gregory, J., Johnson, G. C., Knutti, R., et al.: Energy budget constraints on climate response, Nature Geoscience, 6, 415, 2013.

Proistosescu, C. and Huybers, P. J.: Slow climate mode reconciles historical and model-based estimates of climate sensitivity, Science Advances, 3, e1602 821, 2017.

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

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P.: RCP 8.5—A scenario of comparatively high greenhouse gas emissions, Climatic Change, 109, 33, 2011.

Ricke, K. L. and Caldeira, K.: Maximum warming occurs about one decade after a carbon dioxide emission, Environ. Res. Lett., 9, 124 002, https://doi.org/10.1088/1748-9326/9/12/124002, 2014.

Rogelj, J., Meinshausen, M., Schaeffer, M., Knutti, R., and Riahi, K.: Impact of short-lived non-CO2 mitigation on carbon budgets for stabilizing global warming, Environmental Research Letters, 10, 075 001, 2015a.

Rogelj, J., Schaeffer, M., Meinshausen, M., Knutti, R., Alcamo, J., Riahi, K., and Hare, W.: Zero emission targets as long-term global goals for climate protection, Environmental Research Letters, 10, 105 007, 2015b.

Rogelj, J., Den Elzen, M., Höhne, N., Fransen, T., Fekete, H., Winkler, H., Schaeffer, R., Sha, F., Riahi, K., and Meinshausen, M.: Paris Agreement climate proposals need a boost to keep warming well below 2 C, Nature, 534, 631, 2016a.

Rogelj, J., Schaeffer, M., Friedlingstein, P., Gillett, N. P., van Vuuren, D. P., Riahi, K., Allen, M., and Knutti, R.: Differences between carbon budget estimates unravelled, Nat. Clim. Change, 6, 245–252, https://doi.org/10.1038/nclimate2868, 2016b.

Rogelj, J., Schleussner, C.-F., and Hare, W.: Getting It Right Matters: Temperature Goal Interpretations in Geoscience Research, Geophys. Res. Lett., 44, 10,662–10,665, https://doi.org/10.1002/2017GL075612, 2017.

Rogelj, J., Shindell, D., Jiang, K., Fifita, S., Forster, P., Ginzburg, V., Handa, C., Kheshgi, H., Kobayashi, S., Kriegler, E., et al.: Mitigation pathways compatible with 1.5 C in the context of sustainable development, 2018.

Rogelj, J., Forster, P. M., Kriegler, E., Smith, C. J., and Séférian, R.: Estimating and tracking the remaining carbon budget for stringent climate targets, Nature, 571, 335, 2019a.

Rogelj, J., Huppmann, D., Krey, V., Riahi, K., Clarke, L., Gidden, M., Nicholls, Z., and Meinshausen, M.: A new scenario logic for the Paris Agreement long-term temperature goal, Nature, 573, 357–363, 2019b.

Rohling, E. J., Marino, G., Foster, G. L., Goodwin, P. A., Anna, S., and Köhler, P.: Comparing climate sensitivity, past and present, Annual Review of Marine Science, 2018.

Rose, B. E., Armour, K. C., Battisti, D. S., Feldl, N., and Koll, D. D.: The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake, Geophysical Research Letters, 41, 1071–1078, 2014.

Royer, D., Pagani, M., and Beerling, D.: Geologic constraints on earth system sensitivity to CO 2 during the Cretaceous and early Paleogene, Earth System Dynamics Discussions, 2, 211–240, 2011.

Ruelle, D.: General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium, Physics Letters A, 245, 220–224, 1998.

Rugenstein, M., Bloch-Johnson, J., Abe-Ouchi, A., Andrews, T., Beyerle, U., Cao, L., Chadha, T., Danabasoglu, G., Dufresne, J.-L., Duan, L., et al.: LongRunMIP–motivation and design for a large collection of millennial-length AO-GCM simulations, Bulletin of the American Meteorological Society, https://doi.org/https://doi.org/10.1175/BAMS-D-19-0068.1, 2019.

Rugenstein, M. A., Caldeira, K., and Knutti, R.: Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes, Geophysical Research Letters, 43, 9877–9885, 2016.

Sachs, J. D., Schmidt-Traub, G., and Williams, J.: Pathways to zero emissions, Nature Geoscience, 9, 799, 2016.

Sanderson, B. M., O'Neill, B. C., and Tebaldi, C.: What would it take to achieve the Paris temperature targets?, Geophys. Res. Lett., 43, 7133–7142, https://doi.org/10.1002/2016GL069563, 2016a.

Sanderson, B. M., O'Neill, B. C., and Tebaldi, C.: What would it take to achieve the Paris temperature targets?, Geophysical Research Letters, 43, 7133–7142, 2016b.

Sanderson, B. M., Xu, Y., Tebaldi, C., Wehner, M., O'Neill, B. C., Jahn, A., Pendergrass, A. G., Lehner, F., Strand, W. G., Lin, L., et al.: Community climate simulations to assess avoided impacts in 1.5 and 2 C futures, Earth System Dynamics, 8, 827–847, 2017.

Senior, C. A. and Mitchell, J. F.: The time-dependence of climate sensitivity, Geophysical Research Letters, 27, 2685–2688, 2000.

Sherwood, S. C., Bony, S., and Dufresne, J.-L.: Spread in model climate sensitivity traced to atmospheric convective mixing, Nature, 505, 37, 2014.

Smith, C. J., Forster, P. M., Allen, M., Leach, N., Millar, R. J., Passerello, G. A., and Regayre, L. A.: FAIR v1. 3: A simple emissions-based impulse response and carbon cycle model, Geoscientific Model Development, 11, 2273–2297, 2018.

Smith, P., Davis, S. J., Creutzig, F., Fuss, S., Minx, J., Gabrielle, B., Kato, E., Jackson, R. B., Cowie, A., Kriegler, E., et al.: Biophysical and economic limits to negative CO 2 emissions, Nature climate change, 6, 42, 2016.

Steinacher, M. and Joos, F.: Transient Earth system responses to cumulative carbon dioxide emissions: linearities, uncertainties, and probabilities in an observation-constrained model ensemble, Biogeosciences, 13, 1071–1103, 2016.

Tan, I., Storelvmo, T., and Zelinka, M. D.: Observational constraints on mixed-phase clouds imply higher climate sensitivity, Science, 352, 224–227, 2016.

Tian, B.: Spread of model climate sensitivity linked to double-Intertropical Convergence Zone bias, Geophysical Research Letters, 42, 4133–4141, 2015.

Tokarska, K. B. and Zickfeld, K.: The effectiveness of net negative carbon dioxide emissions in reversing anthropogenic climate change, Environmental Research Letters, 10, 094 013, 2015.

Tokarska, K. B., Schleussner, C.-F., Rogelj, J., Stolpe, M. B., Matthews, H. D., Pfleiderer, P., and Gillett, N. P.: Recommended temperature metrics for carbon budget estimates, model evaluation and climate policy, Nature Geoscience, 12, 964–971, 2019a.

Tokarska, K. B., Zickfeld, K., and Rogelj, J.: Path independence of carbon budgets when meeting a stringent global mean temperature target after an overshoot, Earth's Future, 2019b.

United Nations: Paris Agreement, https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-7-d&chapter=27&clang=_en, 2015.

Van Vuuren, D. P., Stehfest, E., den Elzen, M. G., Kram, T., van Vliet, J., Deetman, S., Isaac, M., Goldewijk, K. K., Hof, A., Beltran, A. M., et al.: RCP2. 6: exploring the possibility to keep global mean temperature increase below 2 C, Climatic Change, 109, 95, 2011.

Van Vuuren, D. P., Van Soest, H., Riahi, K., Clarke, L., Krey, V., Kriegler, E., Rogelj, J., Schaeffer, M., and Tavoni, M.: Carbon budgets and energy transition pathways, Environmental Research Letters, 11, 075 002, 2016.

Vichi, M., Navarra, A., and Fogli, P. G.: Adjustment of the natural ocean carbon cycle to negative emission rates, Clim. Change, 118, 105–118, https://doi.org/10.1007/s10584-012-0677-0, 2013.

Weaver, A. J., Eby, M., Wiebe, E. C., Bitz, C. M., Duffy, P. B., Ewen, T. L., Fanning, A. F., Holland, M. M., MacFadyen, A., Matthews, H. D., et al.: The UVic Earth System Climate Model: Model description, climatology, and applications to past, present and future climates, Atmosphere-Ocean, 39, 361–428, 2001.

Williams, R. G., Goodwin, P., Roussenov, V. M., and Bopp, L.: A framework to understand the transient climate response to emissions, Environmental Research Letters, 11, 015 003, 2016.

Williams, R. G., Roussenov, V., Frölicher, T. L., and Goodwin, P.: Drivers of continued surface warming after cessation of carbon emissions, Geophysical Research Letters, 44, 10–633, 2017.

Williamson, D., Goldstein, M., Allison, L., Blaker, A., Challenor, P., Jackson, L., and Yamazaki, K.: History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble, Climate dynamics, 41, 1703–1729, 2013.

Winton, M., Takahashi, K., and Held, I. M.: Importance of ocean heat uptake efficacy to transient climate change, Journal of Climate, 23, 2333–2344, 2010.

Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J., and Heimbach, P.: Global reconstruction of historical ocean heat storage and transport, Proceedings of the National Academy of Sciences, 116, 1126–1131, 2019.

Zhai, C., Jiang, J. H., and Su, H.: Long-term cloud change imprinted in seasonal cloud variation: More evidence of high climate sensitivity, Geophysical Research Letters, 42, 8729–8737, 2015.

Zickfeld, K. and Herrington, T.: The time lag between a carbon dioxide emission and maximum warming increases with the size of the emission, Environ. Res. Lett., 10, 031 001, https://doi.org/10.1088/1748-9326/10/3/031001, 2015.

Zickfeld, K., Arora, V., and Gillett, N.: Is the climate response to CO2 emissions path dependent?, Geophysical Research Letters, 39, 2012.

Zickfeld, K., MacDougall, A. H., and Matthews, H. D.: On the proportionality between global temperature change and cumulative CO2 emissions during periods of net negative CO2 emissions, Environmental Research Letters, 11, 055 006, 2016.

Appendix : Supplementary Material

S1 Prior assumptions on Realized Warming Fraction

The RWF constraint refers to the observation in Millar et al. (2015) that in CMIP5, at least, the TCR and ECS values are well correlated, such that the ratio of the two values $TCR : ECS$ lies between 0.45 and 0.75 throughout the ensemble. This is apparent by considering Figure S2(a), which shows the joint CMIP5 distribution for the two quantities. The consistency of this relationship informed the parameter distribution choice for a pulse-response model ensemble used in Millar et al. (2017c), where a relationship between ECS and TCR was represented in the distribution of plausible models.

However, as is shown in Proistosescu and Huybers (2017) and in Figure S1, for most models in the CMIP5 and CMIP6 ensembles, the constant feedback ECS is likely an underestimate of true equilibrium response to forcing. Fitting a 2-timescale model directly to model output time-series allows the computation of the possible values of equilibrium climate sensitivity consistent with the first 150 years of simulation (Figure S1). In many cases the extrapolated equilibrium temperature is not strongly constrained by the 150 year simulations. However, some features are generally discernible: in most models the uncertainty distribution contains values which are generally greater than the Effective Climate Sensitivity estimate (Gregory et al., 2004).

Figure S2 shows EffCS and ECS as a function of TCR for each model in CMIP5 (and some available models in CMIP6). ECS is estimated as in Gregory et al. (2004), while EffCS is calculated by performing an MCMC fit (as in section 1.1.1 using the 150 year global mean surface temperature timeseries of each CMIP5 and CMIP6 model's abrupt4xCO2 simulation to the 2 timescale pulse response model (itself forced by a step function forcing timeseries corresponding to 7.2Wm^{-2} after the first timestep). The resulting posterior distribution of the combined value of q_1 and q_2 then informs the range of plausible values of ECS which are consistent with the CMIP5 or CMIP6 simulation.

It is notable that there is no discernible relationship between possible values of ECS (fitted here with a flat prior allowing values between 0 and 40) and TCR in CMIP5 because the values of ECS are not strongly constrained. As such, there is little basis to assume that the equilibrium sensitivity of the system is well constrained by the TCR in the form of an RWF prior.

S2 Joint distributions of parameters in MCMC optimization

Figures S5, S6, S7 and S8 and show pairwise posterior distributions of the parameters optimized using historical emissions and ~~HadCRUT~~ HadCRUT-CW temperature evolution from 1850-2016 in the 'C.T', 'C,T.H', 'C,T, RWF' and 'C,T,Paleo' constraints respectively.

In all cases, it is apparent that there are a range of solutions allowing for different timescales of response in the pulse-response model which can describe climate evolution to date within the provided constraints. For example, in Figure S5, there is a trade-off between q_2 and d_2 parameters, which represent equilibrium climate response on fast timescales and timescale of the shallow ocean thermal response of the system. There is also a trade-off between the q_2 and q_1 (the component of ECS associated with slow feedbacks associated with warming of the deep ocean component). Introducing a constraint on ocean heat

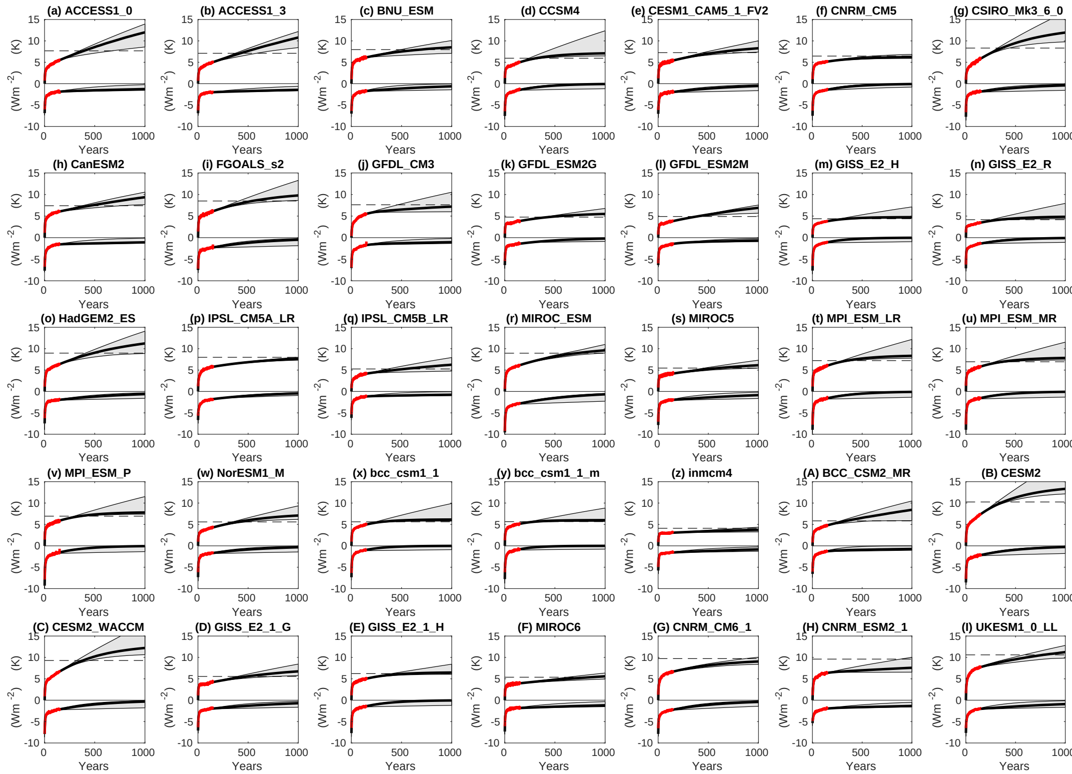
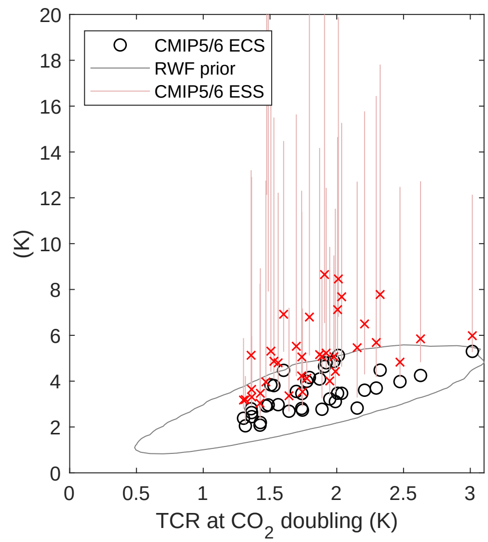


Figure S1. A Figure showing the temperature response to a quadrupling of carbon dioxide. Red points show annual mean temperature from the model's 150 year simulation, the red central line shows the best fitting 2 timescale pulse response model, while the pink area shows the 10th and 90th percentiles of the MCMC distribution of possible pulse response models. The dashed black line shows the Effective climate sensitivity(Gregory et al., 2004).

content, at least in this configuration, constrains r_1, the fraction of heat absorbed by the deep ocean - but has only a minor constraining effect on ECS S6.

The 'RWF' cases (Figure S7), however, has a significant effect on ECS. The deep ocean/slow component of ECS (q_1) is constrained to be small by the RWF constraint (which states that 40-60 percent of equilibrium warming associated with current greenhouse gas concentrations has already been realized). S8) constrain the sum of the 2 equilibrium responses q_n, which is apparent in the parameter distributions by the truncated distribution for q_1 in Figure S8 and increased values for q_2, implying that a greater fraction of present day warming is explained by warming on decadal timescales. The paleo constraint has a similar, but less dramatic effect - constraining the upper bound of the distribution for q_1 relative to the 'C,T' case.

Figure S2. Black circles show Effective Climate Sensitivity (calculated as a constant feedback extrapolation following Gregory et al. (2004)) as a function of Transient Climate Response (warming at time of CO2 concentration doubling in each model's 1 percent CO2 ramping experiment. Each point shows one model in the CMIP5/6 ensemble, and the circle shows the 5th percentile of the prior joint distribution for ECS and TCR used in Millar et al. (2017c). Red whisker plots show the relationship between TCR and ECS calculated using the 2-timescale pulse-response model fits to the abrupt4xCO2 simulation of the corresponding CMIP5 model.



S3 Sensitivity of scenario pathways to different historical constraints

The main paper considers scenarios in the case where only historical emissions and concentrations are known, but we can also consider the impact of different historical constraints on the pathways for 1.5C and 2.0C climate stabilization pathways. The temperature pathways are achievable, irrespective of the historical ensemble constraints used due to the large allowed negative emissions fluxes in the late 21st century (Figure S9), but the range of possible future emissions is dependent upon the prior assumptions (Figure S11) - the use of either the RWF or Paleo constraints tends to reduce the post-2050 negative emissions burden.

It is apparent that, as for the RCP2.6 case in Figure 1(b-e), there are large differences in the cumulative-emissions/temperature behavior for the different ensembles (Figure S12). Solutions with substantial hysteresis are possible with 'C,T', 'C,T,Heat' and 'C,T,Paleo' constraints - but not in the case of the 'C,T,RWF' constraint. Associated net 2100 cumulative carbon budgets for 1.5 and 2C stabilization also vary by prior - with significantly larger allowances when the RWF or Paleo constraints are employed.

The relationship between mid-century temperatures and late century carbon removal requirements, however, remains relatively robust irrespective of ensemble constraint (Figure S13), though expected mid-century warming for a given amount of 2020-2050 net emissions is reduced by 0.1-0.2K if the RWF or Paleo constraint is used (Figure S15). For example, a 2020-2050 budget of 100GtC produces a likely 2050 warming of 1.6K if 'C,T' constraints are used, and 1.45K if 'C,T,RWF is used'.

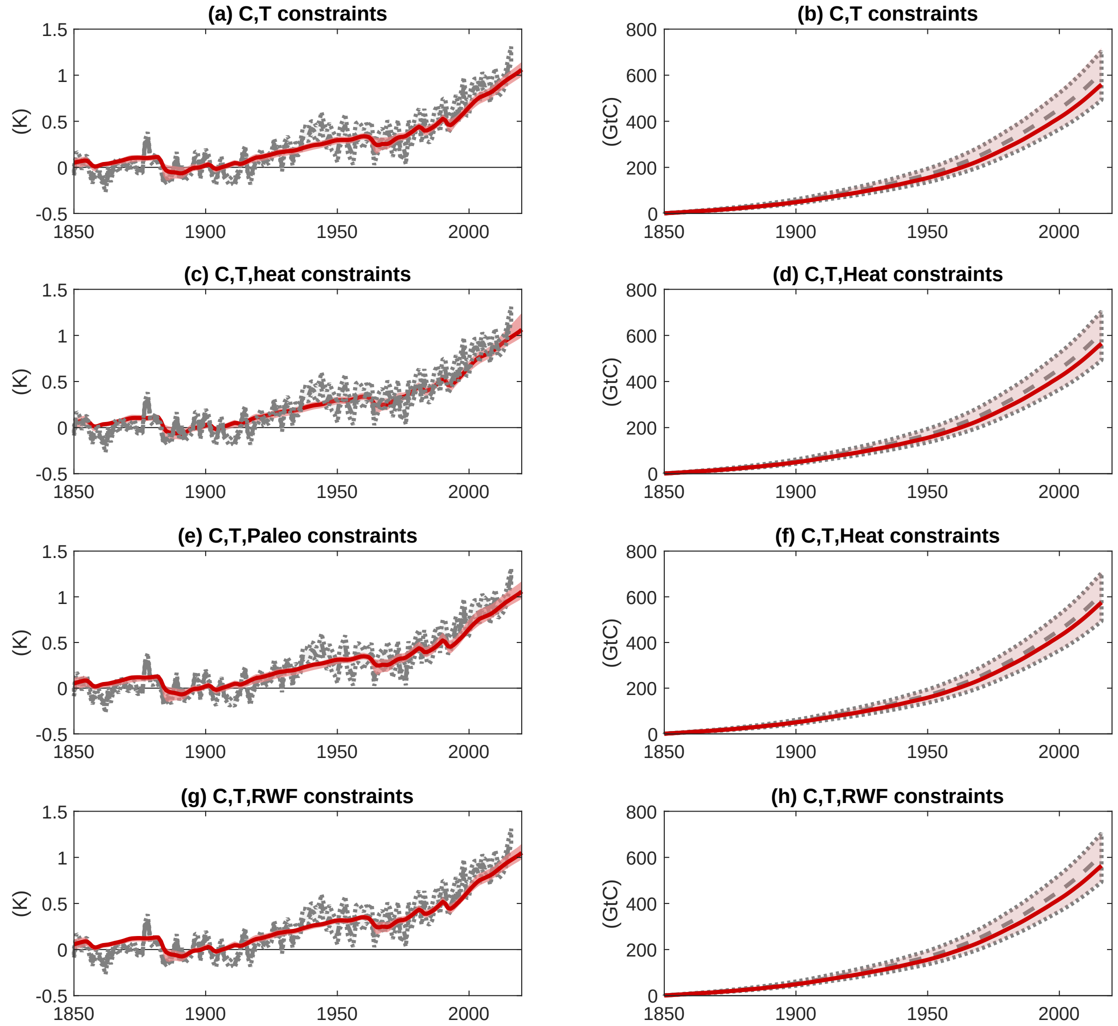


Figure S3. Illustration of ensemble spread (given different prior assumptions) in global mean temperature (a,c,e,g) and cumulative carbon emissions (b,d,f,h) for the observed period. Observational median is grey line, while 5-95% observational uncertain ranges are shown in grey shade (Cowtan and Way (2013) for temperature and Le Quéré et al. (2017); Millar and Friedlingstein (2018) for cumulative emissions). Ensemble median in the historical period is shown in red, with 5-95% ensemble range in red shade.

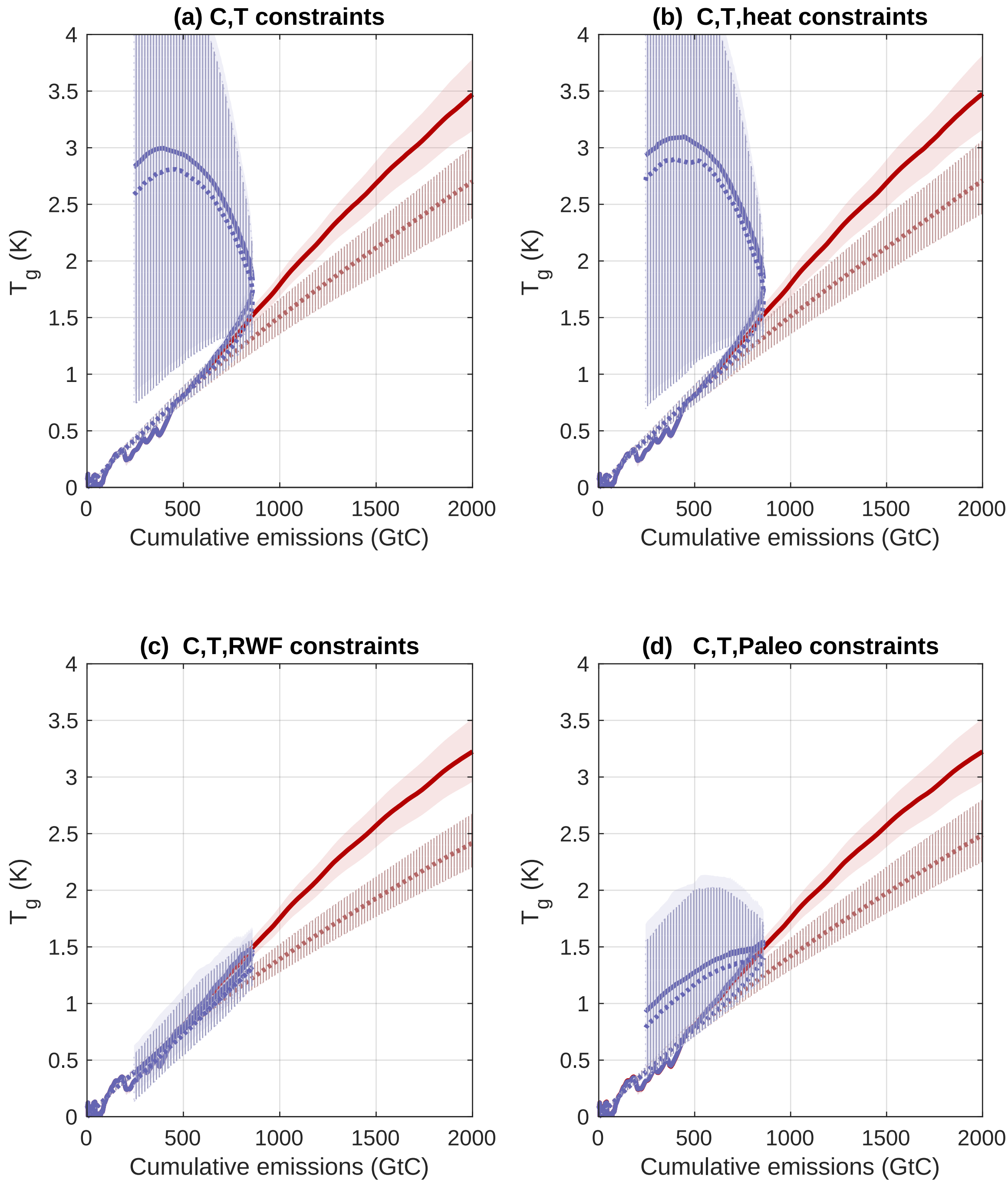


Figure S4. Illustration of the impact of non-CO2 forcers on model behaviour on temperature-cumulative emissions relationships, using different prior assumptions . Each subplot shows the 10-90th percentile ranges of projected response to cumulative emissions corresponding to Figure 1 in the main study. Solid/solid shade regions show the all forcing simulation, while hatched/dotted lines show the response of the same ensemble to a CO2-emissions only simulation, with all other forcings set to zero. RCP8.5/RCP2.6 are shown in red/blue respectively.

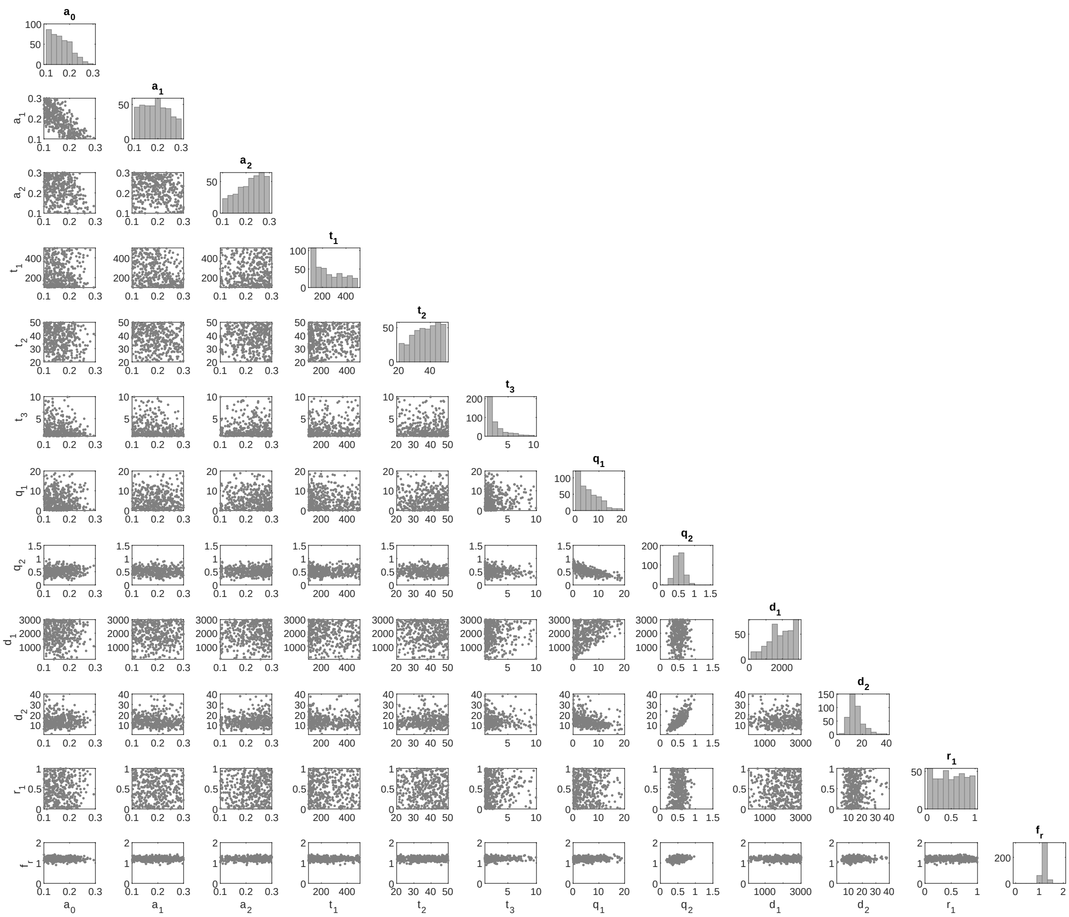


Figure S5. A 'corner-plot' showing pairwise posterior parameter distributions for models constrained using HadCRUT temperature anomalies from 1850-2016. Plots on the diagonal show parameter distributions for each of the parameters in Table 1 considered in the MCMC optimization only historical emissions and temperatures. Off-diagonal plots illustrate 2 dimensional distributions.

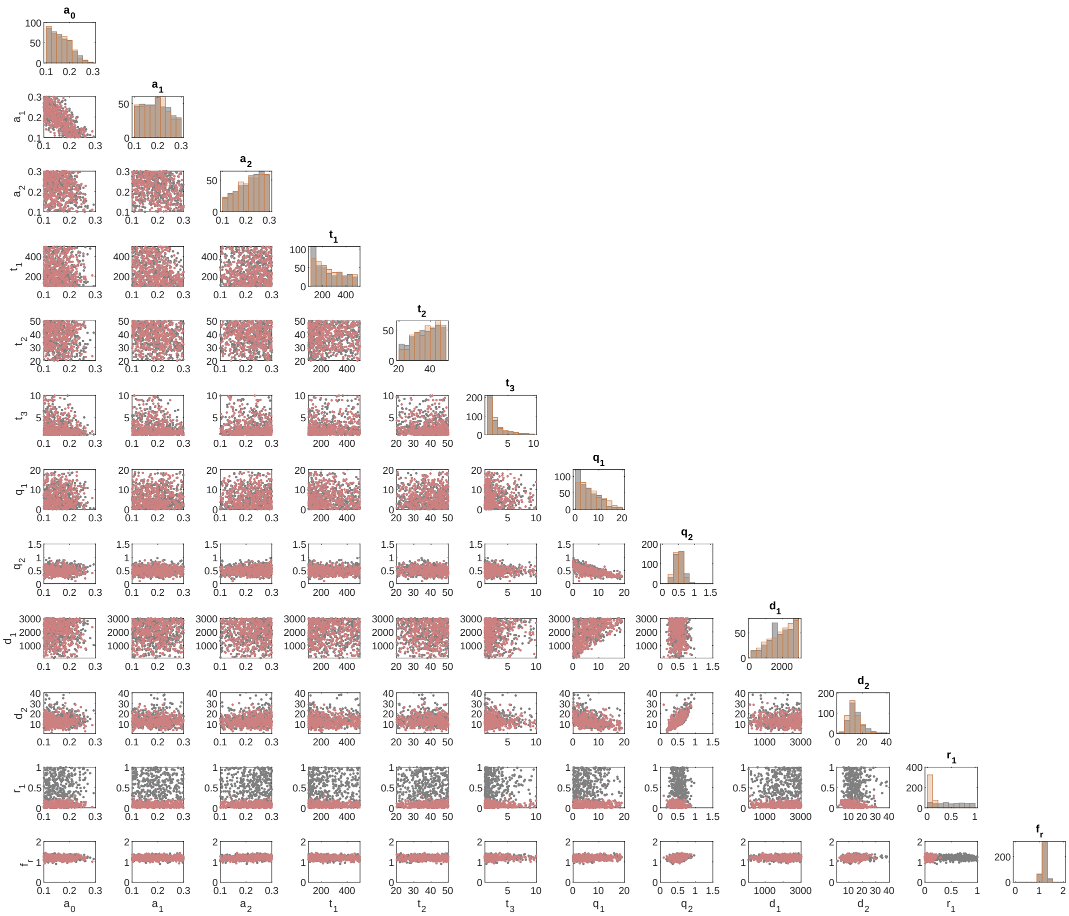


Figure S6. As for Figure S5, but with 'C,T' only constraints in gray and 'C,T and Heat' constraints i red.

.

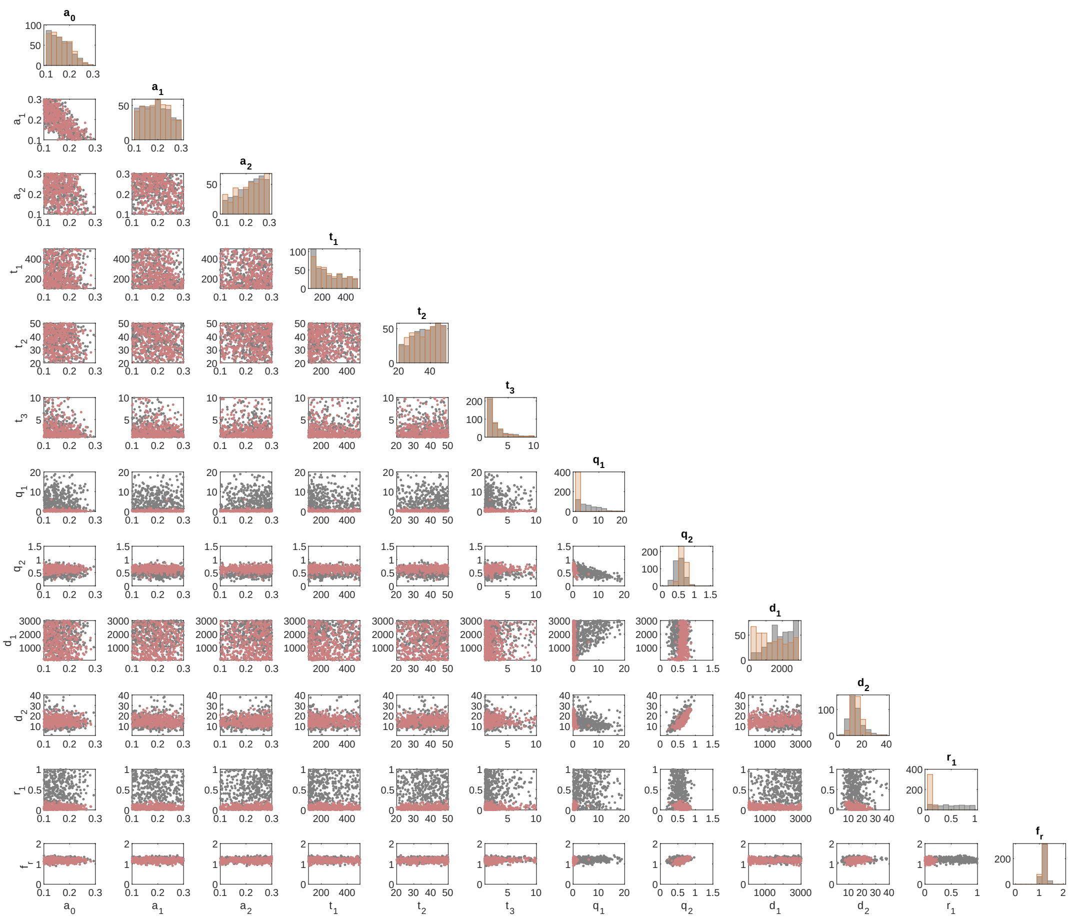


Figure S7. As for Figure S5, but with 'C,T' only constraints in gray and using 'C,T,RWF' constraints in red.

.

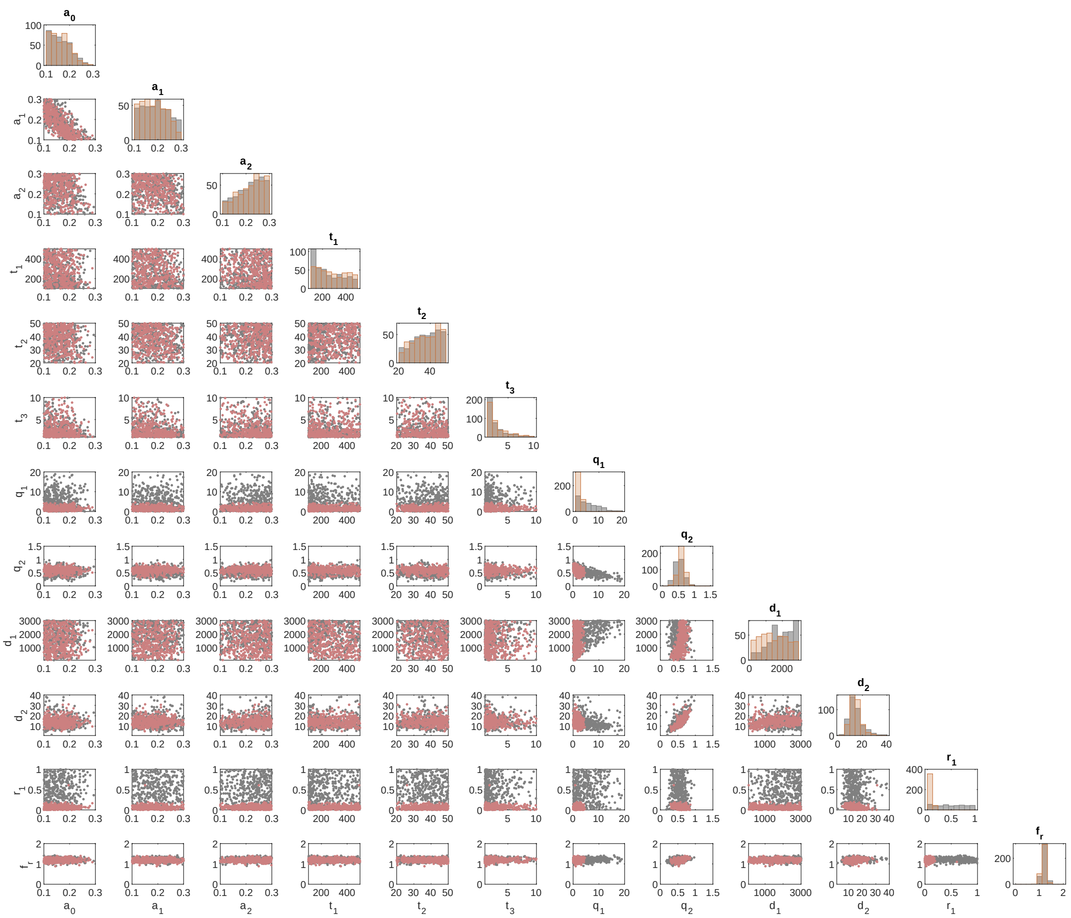


Figure S8. As for Figure S5, but with 'C,T' only constraints in gray and 'C,T,Paleo' constraints in red.

.

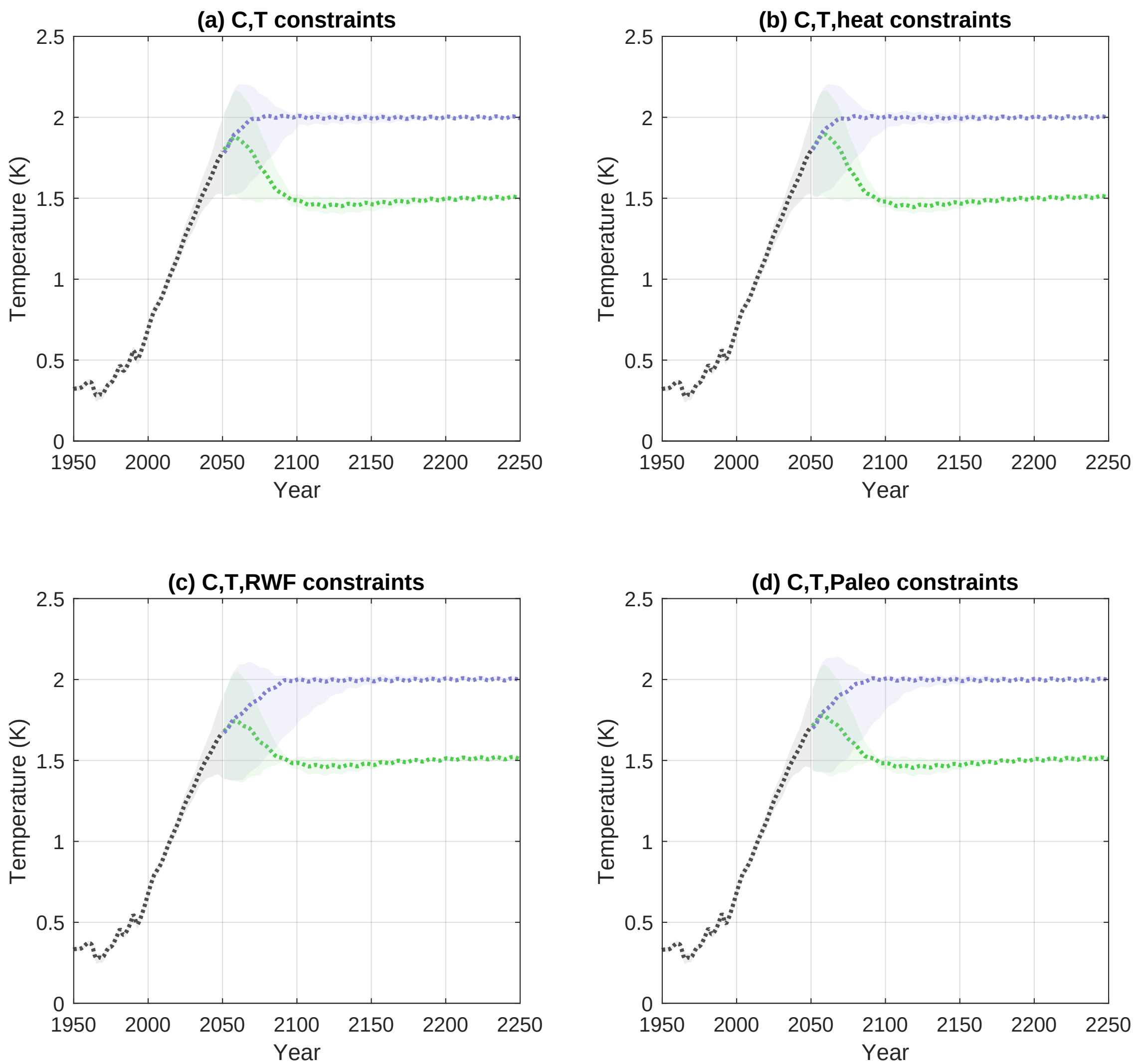


Figure S9. As for Figure 2(a), for each of the constraints considered in Figure 1(b-e).

.

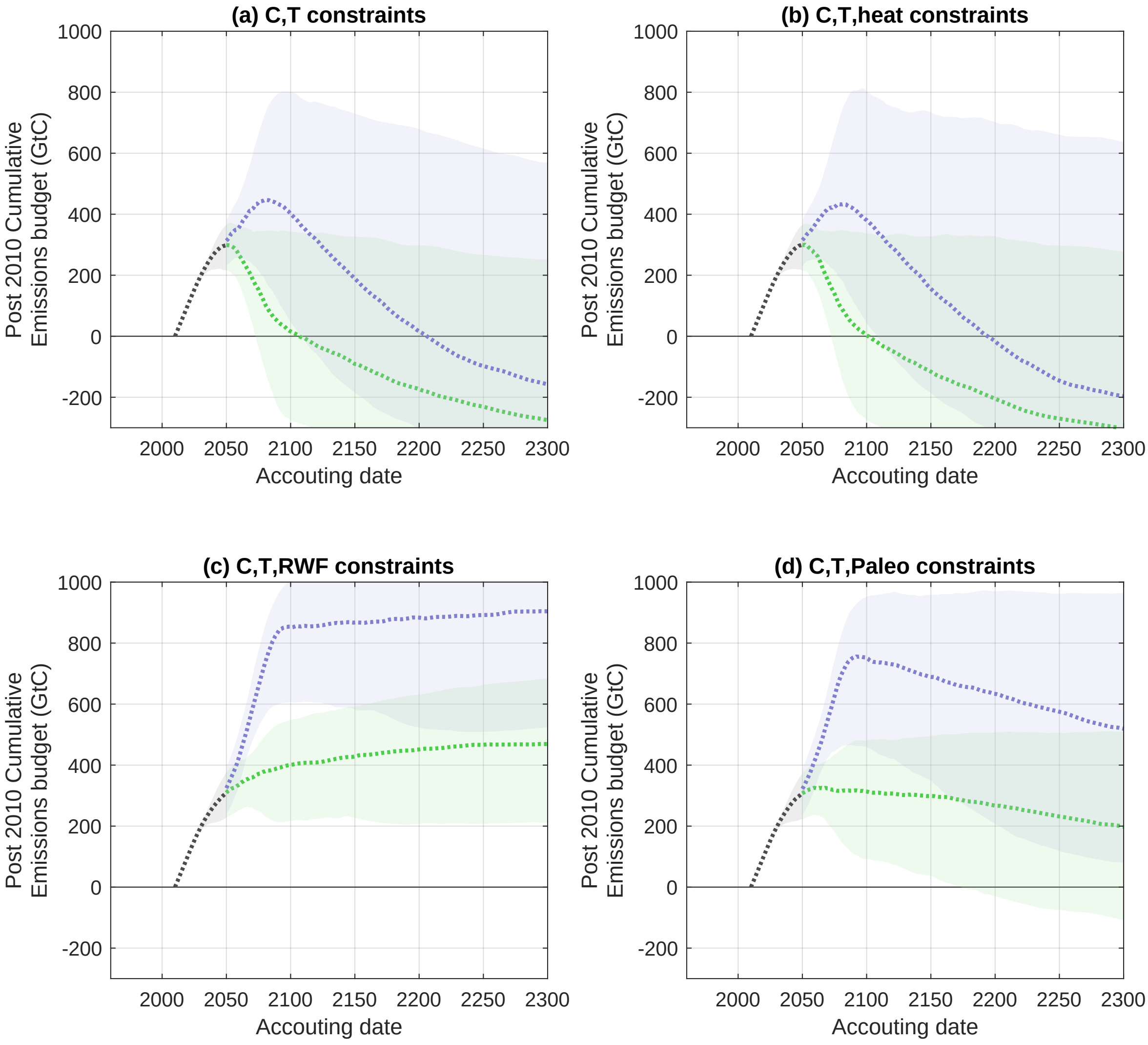


Figure S10. As for Figure 2(d), for each of the constraints considered in Figure 1(b-e).

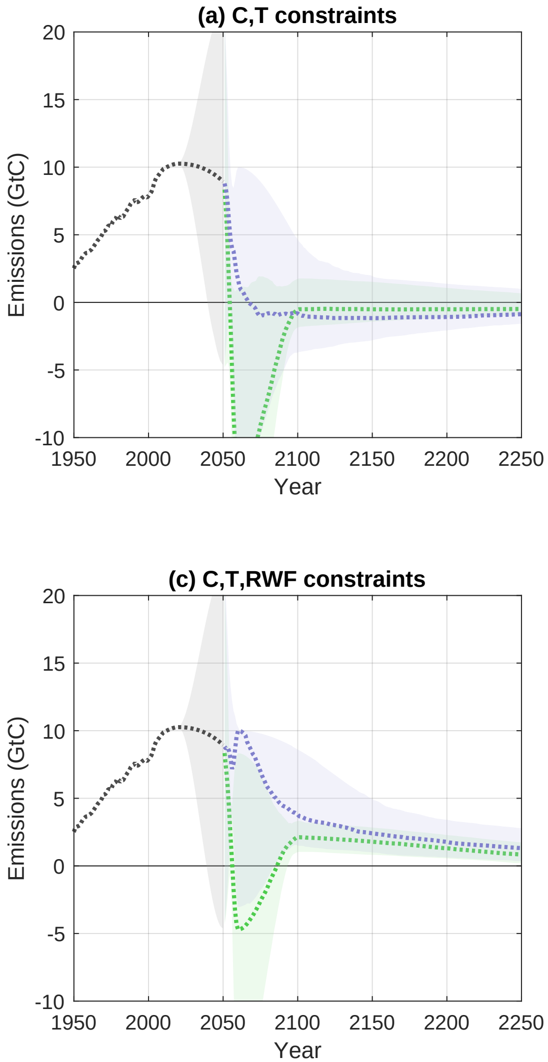


Figure S11. As for Figure 2(b), for each of the constraints considered in Figure 1(b-e).

.

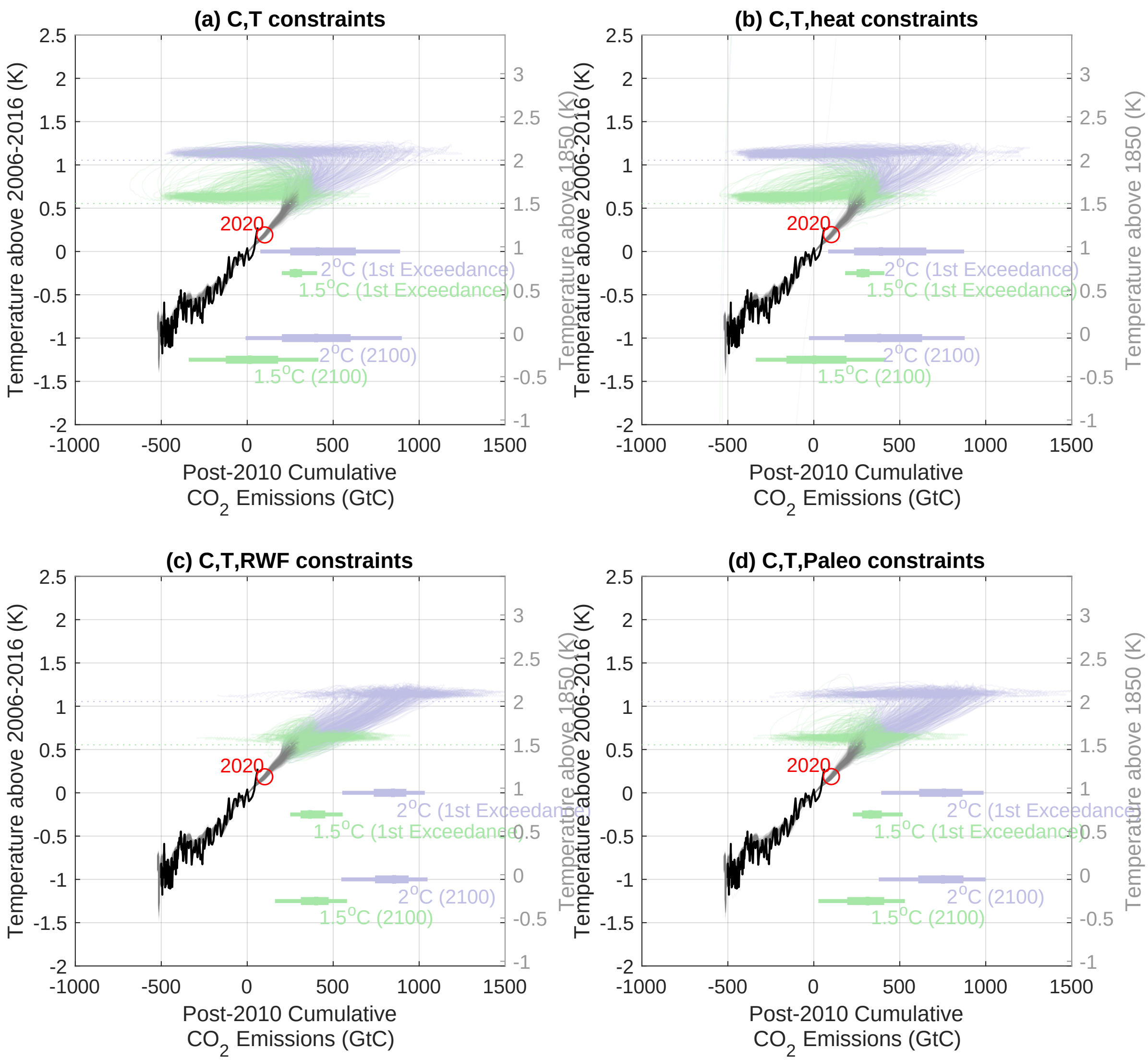


Figure S12. As for Figure 2(c), for each of the constraints considered in Figure 1(b-e).

.

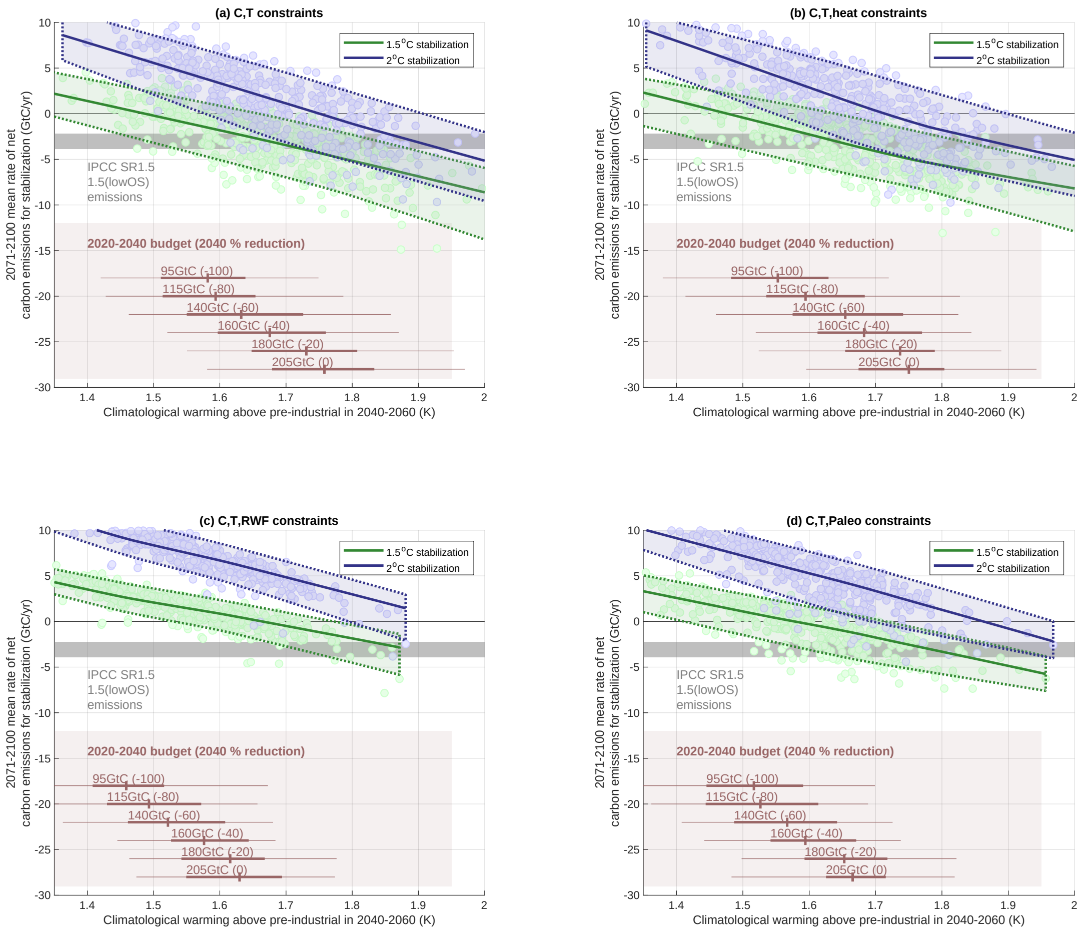


Figure S13. As for Figure 3(b), for each of the constraints considered in Figure 1(b-e).

.

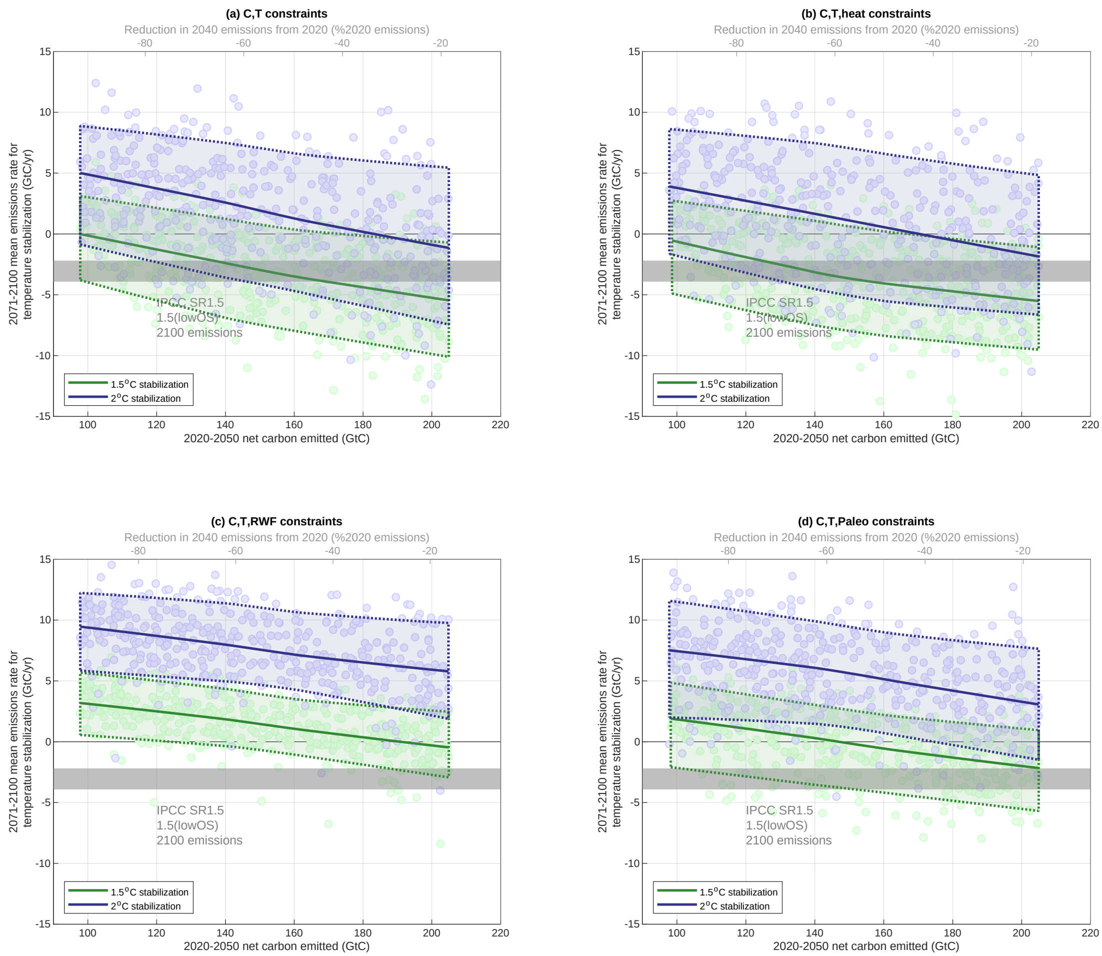


Figure S14. As for Figure 3(a), for each of the constraints considered in Figure 1(b-e).

.

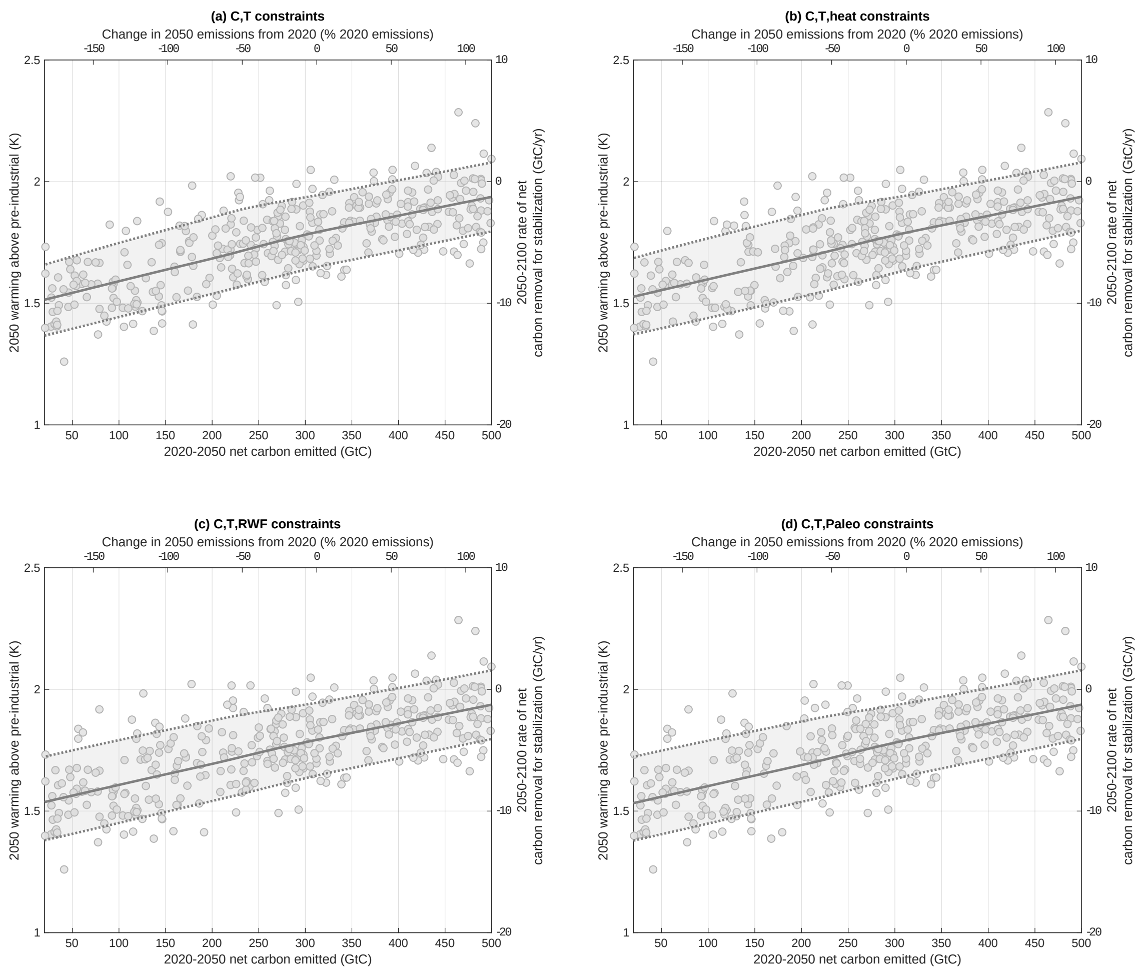


Figure S15. As for Figure 3(b), for each of the constraints considered in Figure 1(b-e).

.

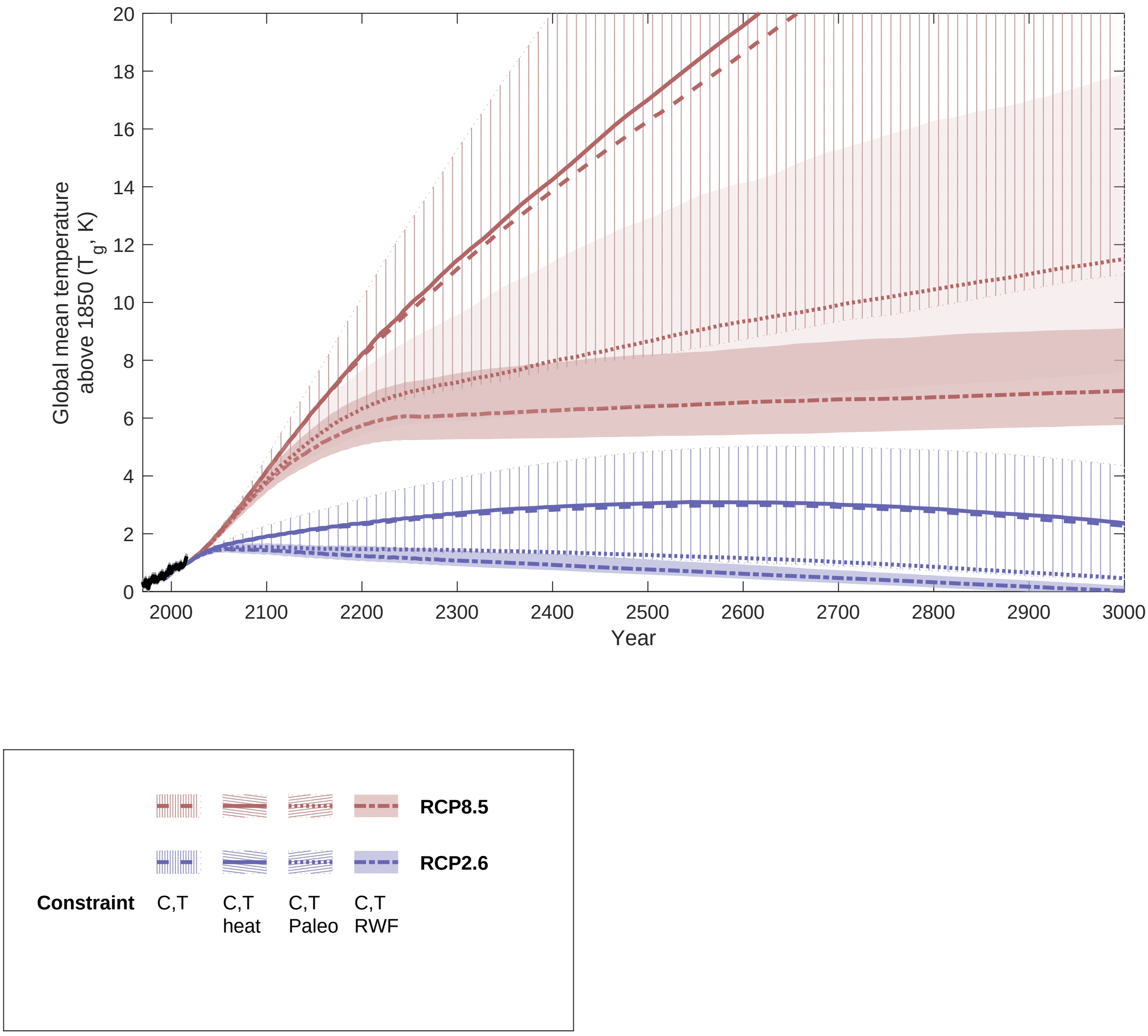


Figure S16. As for Figure 1(a), with the time axis showing millennial-scale evolution, with 2300 emissions remaining constant until the year 3000 in all cases.

However, the 2050-2100 negative emission requirements are broadly similar for a given level of observed 2050 warming, irrespective of the constraint. For example, if observed warming in 2050 is 1.5C - this corresponds to allowable 2050-2100 emissions of -150 to +200GtC if C,T constraints are used, and 0 to +200GtC if C,T and RWF are used.