Opinion: Can uncertainty in climate sensitivity be narrowed further?

Steven C. Sherwood[1] and Chris E. Forest[2,3,4]

[1]Climate Change Research Centre, UNSW Sydney, Kensington NSW 2052, Australia
[2]Dept. of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, PA 16802, USA
[3]Earth and Environmental Systems Institute, The Pennsylvania State University, University Park, PA, USA
[4]Center for Earth System Models, Analysis, and Data, The Pennsylvania State University, University Park, PA, USA

Correspondence: Steven Sherwood (ssherwood@alum.mit.edu)

Abstract. After many years with little change in community views on Equilibrium Climate Sensitivity (ECS), in 2021 the IPCC concluded that it was much better known than previously. This development underpinned increased confidence in long-term climate changes in that report. Here, we place this development in historical context, briefly assess progress since then, and discuss the challenges and opportunities for further improving our knowledge of this iconic concept. We argue that the probability distributions published in those assessments are still approximately valid; while various subsequent studies have claimed further narrowing, they have omitted important structural uncertainties associated with missing processes, imperfect relationships, or other factors that should be included. The distributions could nonetheless be narrowed in the future, particularly through better understanding of certain climate processes and paleoclimate proxies. Not all touted strategies are truly helpful, however. We also note that ECS does not address risks from the carbon cycle or possible tipping points; and as increasingly strong mitigation (i.e., "net-zero") scenarios are considered, ECS becomes less informative about future climate change compared to other factors such as aerosol radiative forcing and influences on regional change such as ocean dynamics.

1 Introduction

Equilibrium climate sensitivity (ECS)—the eventual rise in Earth's mean surface temperature following a doubling of atmospheric CO_2—has a long history. Early estimates, from the original calculation of Arrhenius (1896) to the seminal National Academy of Sciences (NAS) review (Charney et al., 1979) (aka The Charney Report), were based on first principles, i.e., climate models. The NAS review committee had two early GCMs available, but rather than taking these models at face value, the NAS committee attempted a deep dive into how neglected processes could shift ECS up or down. The authors decided not to change the best-guess ECS from the model average, but allowed for a large uncertainty. This set an admirable precedent for grappling with the uncertainty of ECS given only a partial understanding of the climate system, a situation that remains the case today despite considerable progress.

A key development beginning around the turn of the century was the attempt to infer ECS from observed historical temperature data. This works essentially by using these data to constrain a known relationship between the estimated changes in radiative forcing and warming, where ECS is a parameter that is either adjusted directly (Andronova and Schlesinger, 2001;

Knutti et al., 2002) or indirectly by modifying processes such as cloud feedback within a simple climate model (Forest et al., 2002; Sokolov and Stone, 1998; Libardoni et al., 2019). To use an illustrative representation of the response of the global mean temperature perturbation from equilibrium (T') (e.g., Peixoto and Oort (1992) in Section 2.5), we can use the following equation:

$$C\frac{dT'}{dt} = N(t) = F(t) - \lambda T'(t) \tag{1}$$

where C is an effective heat capacity, N is the net top-of-atmosphere energy imbalance, λ is a feedback factor, and $F(t)$ is the radiative forcing perturbation due to exogenous forcing agents (historical anthropogenic greenhouse gases, aerosols, solar intensity variations, land-use change, and others). While this equation can be made more complicated, it highlights the key features of how the additional energy is stored and how the global mean temperature depends on three terms: C, $F(t)$, and λ. Given the observational records of T' and C (hence N), and F, we can estimate ECS as the ratio of F from a doubling of CO_2 (4 $\mathrm{W\,m^{-2}}$) to λ. In early works, C and λ were thought of as constant, while we now know that the effective heat capacity, C, changes over time as the ocean mixes heat into deeper regions, and feedbacks (i.e., λ), particularly from clouds, can also differ between periods of transient warming and equilibrium (Senior and Mitchell, 2000). These complications have important consequences, discussed below.

Other methods of constraining ECS using observations have exploited the responses to volcanic eruptions, solar cycles, or interannual variations in surface temperature, but have broadly been judged less convincing or useful because the forcing and/or response are qualitatively very different from those of GHG, and because the magnitudes of the signal are small. Meanwhile paleoclimate scientists began drawing inferences about ECS since the 1980's (Edwards et al., 2007) leading to a major assessment by PALAEOSENS (2012); see also von der Heydt et al. (2016). These involve large changes but were not fully accepted by many in the mainstream climate science community, however, due to the perceived potential for errors in proxy temperature records and potentially deep uncertainties around radiative forcings and the role of geologic changes.

IPCC reports have conducted reviews of the ECS literature every 6-7 years but are not permitted to do new research, and so were limited in their ability to apply statistical methods to aggregate research findings. Over time, there emerged an offset between published historical estimates (typically lower) vs. paleo and model estimates (typically higher), although estimates from each of these varied among studies. Faced with a large and growing number of such disparate indicators, successive IPCC reports including the Fifth Assessment (AR5) in 2013 echoed the 1.5-4.5°C likely range originally obtained by Charney et al. (1979) (though IPCC gave it a clearer probabilistic interpretation: 66%-or-greater chance).

A significant revision to the status quo occurred with the publication of the first analysis to incorporate all major sources of evidence using a Bayesian probability framework, by Sherwood et al. (2020), hereafter S20. Rather than seeking consensus among estimates, this framework quantifies probability by asking how likely the aggregate of evidence would be under each possible ECS value. This assessment obtained a "likely" range about half as wide as did AR5, at 2.6-3.9°C for baseline assumptions. This revision was due partly to new evidence and partly to the new approach. The substantial lifting of the low end of the range was due to the combined weight of all evidence against low values, while the reduced high end was because higher values are hard to obtain from the process point of view and suggest more warming than observed since the Last Glacial

Maximum. This latter point represented a step forward by bringing communities together and broadening the acceptance of paleoclimate evidence as a valuable constraint on ECS, at the same time highlighting the unreliability of historical estimates due to differences between past cloud responses and anticipated future ones. The narrower, and observationally driven ECS range was approximately adopted by the IPCC in 2021 (AR6), and the stark difference between this narrower range and the broadening spread in GCMs helped motivate a new "warming levels" approach to projections in AR6 that disconnected them from GCM estimates of ECS, as well as underpinning more confident statements about future climate.

2 Recent Advances?

Quite a few ECS studies have come out since S20, often deploying the same methodology but updating one or more of its ingredients. Some studies have reaffirmed the recent assessments. For example, Zhu et al. (2021) confirm that a GCM with very high ECS predicts excessively large paleoclimate changes contradicted by proxy evidence and then show (Zhu et al., 2022) that good predictions are made by an improved model with an ECS of 4 °C. Researchers are beginning to look at the late Miocene (not used by S20), also inferring higher ECS (Brown et al., 2022) although the reliability of using this time period requires further investigation. Ceppi and Nowack (2021) have used a clever data analysis to find an overall cloud feedback essentially the same as that of S20, although claiming smaller uncertainty. There is new process understanding suggesting positive feedback from mid-level clouds (Stauffer and Wing, 2022), and possibly stronger positive feedback from optical changes in Southern Ocean clouds (Wall et al., 2022b). Myers et al. (2021) find a narrowing and slight reduction of ECS based on a revised observational constraint on low-cloud feedback, but the narrower uncertainty accounts only for observational limitations and relies on optimistic and untested structural assumptions about the independence of cloud types and transferability of observed regional variations to global average cloud behavior.

Among new studies we are aware of, the one claiming the largest revision is by Lewis (2022) who asserts a narrower and substantially lower ECS using the basic S20 methodology with various updates. While this author claims “errors” in S20, looking carefully it appears these are differences in opinion on methodological choices and priors rather than errors, and they moreover were acknowledged to have little effect on the outcome. Instead, the reduction and narrowing of the ECS probability density (PDF) resulted from a selective use of evidence — most importantly, a decision to reject the possibility of a large “pattern effect” on historical SST even though this continues to be strongly supported by new studies (e.g., Heede and Fedorov, 2021; Andrews et al., 2022; Chao et al., 2022), and a downward revision of expected historical aerosol cooling. Together these two departures allowed Lewis to conclude (in contrast to other studies) that the historical record rules out high ECS. Scafetta (2022) and a few other studies have reached a similar conclusion by pointing out that most CMIP GCMs warm too much in recent decades, but these studies ignore the pattern-effect model bias that is contributing to this, as well as the uncertainties in the aerosol forcing trend over such a short period (see e.g., Williams et al., 2022), both of which are likely systematic across the model ensemble. Some of these studies also neglect any impact of internal variability on the global-mean temperature, which has a small effect on trends at the centennial scale but becomes a significant noise source with shorter records. Interestingly another recent study (Wall et al., 2022a) finds a narrower PDF and *stronger* mean of aerosol cooling, and thereby a slightly higher ECS also on the basis of the historical record all other things being equal; however, they only considered sulfate effects on shallow clouds which ignores potentially large effects from less well-understood cloud and aerosol interactions.

It was noted by S20 that an endemic problem in individual studies of ECS up to that time was a failure to account for "structural" uncertainty in the assumptions of forward models used to predict what would be observed given a particular ECS. The quoted uncertainties typically derived only from sampling (in model or data space), conditional on the structural assumptions. Because results can be quite sensitive to such assumptions, past studies sometimes obtained very different PDFs of ECS—sometimes barely overlapping—from essentially the same evidence. In situations where no solid argument is available to rule out particular approaches, one should average results over the various equivalent alternatives (and possibly allow further broadening in case the available approaches share biases or faulty assumptions). Choosing one plausible approach and ignoring others produces an overconfident (too-narrow) PDF. In our judgment, this unfortunately continues to be an issue with every recent study noted above that attempted to quantify probabilities, so we are not convinced that uncertainties have significantly narrowed. Nor do we see a consensus toward raising or lowering the ECS among the credible recent studies, although there is some evidence that a bit more of the positive total cloud feedback may come from types other than low cloud.

3 Opportunities for near-term progress on ECS

So if these studies have not yet significantly "moved the dial" on ECS, is there any hope for this to happen? Following Bayesian reasoning (which we think applies whether or not Bayes mathematical methodology is formally used), outcomes can change either if the priors change or the evidence (or its interpretation) changes.

Priors (on ECS or related variables) have been a contentious issue since the first Bayesian ECS studies. S20 found that switching between two priors for which there was advocacy changed the upper or lower range limits as much as omitting an entire line of evidence—showing that consensus on priors could substantially reduce apparent uncertainty. Disagreement about priors sometimes seems related to disagreement about the meaning of probability itself: does it quantify (a) what someone really expects and would bet big money on, or (b) a quasi-objective calculation conditional on an explicit set of chosen assumptions and parameters? To be most useful for guiding decisions we would argue for (a), which also seems the effective choice of the IPCC, and which motivates the above-stated concerns about structural uncertainty. However, many arguments for priors invoke "objectivity" or "neutrality" (e.g., Jeffreys prior, or "flat" or uniform priors in some arbitrarily chosen variable) which seems to imply (b), as does the neglect of structural uncertainty in so many past studies (presumably because it cannot be treated "objectively" at least in a single study). Most physical scientists are uncomfortable with explicitly subjective judgments in spite of their unavoidable role in probability, and most have had little formal training in probability or statistics. Further cross-disciplinary community discussion of these issues might lead to a stronger consensus and downgrade "priors" as a source of disagreement or apparent uncertainty about ECS.

The more straightforward path to improvement however would be through stronger evidence, or greater confidence in interpreting it. Indeed each of the new studies noted in section 2 does offer new evidence: even if this may have been over-interpreted in individual studies, the cumulative effect will reshape the PDF over time (hopefully by narrowing it), although this will be a slow process because single pieces of new evidence have a small effect on the PDF. Future assessments that again survey all evidence and structural models will be crucial to track this evolution.

All three major lines of evidence (process understanding, the historical warming record, and paleoclimate evidence) could further constrain ECS. Our understanding of cloud processes has advanced markedly over the last decade and continues to do so. In particular, there is huge potential for machine learning to make better use of the enormous satellite datasets for outgoing longwave radiation; some new studies are emerging (Kuma et al., 2023) and we expect major advances to narrow uncertainty in cloud responses to warming and aerosols. High-resolution global models may also help assess existing models, which have more heavily parameterised cloud processes.

There is also clear potential to improve use of historical evidence. Many studies including S20 use only the global-mean, linear temperature and ocean heat content trends to constrain ECS. Forest et al. (2002); Hegerl and Wallace (2002); Libardoni et al. (2019) and some later studies used latitudinal variations or other patterns to help attribute warming to different forcings, which is in principle is more powerful, but doubts are raised about this approach. Specifically, current models struggle to correctly predict even global-scale temperature patterns such as the equator-to-pole gradient in some paleoclimates, Arctic and Antarctic warming in recent decades, and the Pacific east-west temperature gradient that underlies the so-called “pattern effect” noted earlier. Recent work continues to suggest that the latter discrepancy is either a transient phenomenon or perhaps due to missing aerosol forcing mechanisms possibly involving global teleconnections (Heede and Fedorov, 2021; Meehl et al., 2021), but other explanations are also possible that could have very different ramifications for ECS. Resolving these modeling gaps would not only remove a key obstacle to using the global-mean warming record, but would strengthen efforts to use regional temperature patterns to help distinguish the effect of different forcing agents such as aerosols. A promising approach is to use model ensembles to develop consistent explanations for observed patterns of change (e.g., Raghuraman et al., 2023). There also remains significant uncertainty in patterns of ocean surface warming especially prior to the ARGO period (e.g., Gleckler et al., 2016), such that further extension of the (now) high-quality record over time, or retroactive improvement of older estimates via improved statistical approaches or data recovery, could help constrain pattern effects.

Palaeoclimate evidence has strong potential to exploit real-world phenomena that, while not the same as future climate change or the ECS scenario, are closer analogs than any phenomena during the historical period. Models are crucial to bridge the gap between scant proxy records and ECS, where the net palaeo-radiative-forcing is problematic. Advances in paleoclimate modeling (in particular the growing use of state of the art atmosphere-ocean models for paleo simulations) may enable strategies similar to those above for the historical record to be employed for those time periods, and will be made stronger by the further accumulation of proxy evidence and confidence in its interpretation. Further information on past forcings, especially those other than CO_2, would be especially helpful; for example, strong concerns have been raised about the possibility of large, unexpected forcings due to different preindustrial fire and smoke regimes whose plausible range might have been underestimated in past assessments (Mahowald et al., 2023).

4 Approaches that merit caution

It is worth pointing out some advances are sometimes touted as helpful for constraining ECS but would not actually make a material difference, at least until other more urgent advances have been made. For example, more accurate estimates of historical temperature rise or of the current planetary top-of-atmosphere (TOA) energy imbalance would make very little difference, because the associated uncertainties are already much smaller than uncertainties described earlier relating to surface temperature patterns, short-lived forcers, and the preindustrial reference state. The Earth's current TOA budget surplus is surprisingly well known, with reasonable agreement between satellite, ocean and model heat budget estimates (Loeb et al., 2020)—the problem is we don't know how to infer ECS from this information. A key issue is how to assess the net radiative forcing which relies on quantifying the short-term forcing agents and transient non-GHG factors (especially aerosols, natural and anthropogenic), not only in today's atmosphere but also in the preindustrial baseline one for which we have essentially no observations. On the other hand continued precision monitoring of heat content and surface temperatures may, if combined with better modeling, eventually enable better quantification of fast feedbacks via the observed internal variability and trends.

Although the potential of paleoclimate evidence has been noted, another avenue that won't make a big difference is more simulations of paleoclimates using standard approaches and boundary conditions. What is instead needed is better quantification of the forcings and global temperature changes that actually occurred, for as many climate episodes as possible, and higher confidence in our estimated forcing ranges given deep uncertainties. Model simulations can help with this if experiments are properly designed and done with state of the art models. One avenue is to design experiments to better constrain forcing adjustments (changes in cloud cover or other constituents that are brought about by a forcing agent such as an ice sheet via its topography, thus supplementing its direct radiative forcing), and their efficacies (variations in warming per unit radiative forcing). The role of paleo-forcings that are not so important today must always be considered, for example due to different sea level, vegetation burning regimes, or dust sources; indirect effects of these or topography on clouds; etc. Another avenue is to address and possibly resolve apparent inconsistencies among proxy records, which may be possible using models that can simulate the proxies directly. As an example, some paleo-proxy studies have assumed that land and ocean temperature changes are equal, but land changes are actually expected to be greater, which might reduce apparent discrepancies between proxies (Seltzer et al., 2023). It is not surprising that we cannot exactly simulate past temperature patterns with current models, given that we cannot simulate trends correctly in recent decades where observations are very good.

Finally, so-called "emergent constraint" studies continue to appear, which begin with a GCM distribution of ECS and then narrow it by finding observational constraints that correlate with ECS. These can be very useful for clarifying important feedback mechanisms, but cannot really yield reliable PDFs on their own—some prior is required, which is usually taken to be the unconstrained, empirical distribution of ECS among a set of GCMs, although this is not necessary (Renoult et al., 2020). Interpreting a model distribution as a PDF in general is not satisfactory mainly because the models available now share common conceptual errors and structural simplifications, so they may poorly represent the broader set of successful models that in principle could be found (Stainforth et al., 2007). Emergent constraints also need a clear explanatory mechanism to give them a better chance of robustness and to allow them to be combined with (and be known not to duplicate) other known or suspected feedback mechanisms (Hall et al., 2019)—for example if a constraint operates through high-cloud feedback, it may need to be combined with other evidence on middle or low clouds and other feedbacks to get an overall climate sensitivity estimate. Some emergent-constraint studies (e.g., Scafetta, 2022) use global temperature changes as the constraint variable, which might seem to avoid the need to understand processes; but as noted earlier, the danger is that the GCMs in the ensemble share common flaws in the representation of aerosol forcing and pattern effects, which will throw off any constraint that depends on trends, basically hitting the same problem with the historical record noted by S20.

5 The role of ECS during the approach to equilibrium

It is often noted that Earth's climate will never be in equilibrium. This does not mean that ECS is not an important parameter: it does largely determine the severity of warming in higher-emissions scenarios, even while not in equilibrium, and is in fact more accurate in this respect than the Transient Climate Response (TCR) metric originally intended to capture ocean heat uptake uncertainties (Grose et al., 2018). As the world begins to decarbonise, however, lower-emission scenarios are looking more plausible, and for these the ECS (and TCR) are indeed less relevant.

To see why, note that because C in eq. (1) increases substantially with time as anthropogenic heat penetrates deeper into the oceans, climate exhibits a superposition of rapid (years) and slow (centuries) responses to changes in applied forcing. Assuming at some point net-zero is reached (i.e., no net anthropogenic CO_2 emissions), anthropogenic CO_2 will subsequently be drawn down into natural reservoirs on a time scale of decades (though with a significant fraction remaining for millennia), leaving a positive but decreasing radiative forcing. Similarly, other anthropogenic radiative forcing agents (e.g., CH_4 and N_2O) will also have their own stabilization dynamics. To first order, this lingering greenhouse heat input is expected to balance heat transfer from the surface to deep ocean, hence gradually warming the deep ocean without changing the average surface temperature much. The corollary is that CO_2-driven surface warming stops if, and only if, net zero is reached. A concern however is that net-zero scenarios also involve the removal of most anthropogenic aerosols which, depending on their current radiative forcing, could add enough subsequent warming to threaten Paris targets even if net-zero were achieved today (Smith et al., 2019; Sherwood et al., 2022; Dvorak et al., 2022), although removal of other short-lived GHG would swing things the other way if aerosol forcing happens to be weak. This highly uncertain, potential "rogue warming" is conditionally independent of ECS or TCR given the historical forcing and surface temperature at the time net-zero is reached, because we are taking away a (relatively) recently applied forcing rather than starting from equilibrium (Sherwood et al., 2022). Hence, while climate sensitivity affects how much warming occurs between now and attainment of "net zero," it will not matter much thereafter; whereas knowing the aerosol forcing (Kramer et al., 2021, 2019), valued hitherto mainly as a way to help constrain ECS/TCR, becomes very important in its own right.

The smooth approach to net zero described above is based on understanding and models that may not account for all kinds of potential surprises or "tipping points". The carbon cycle affords some obvious ones: the ECS metric by design ignores this by taking the radiative forcing (hence atmospheric composition) as given. Humans, however, are not in total control of the global carbon cycle even though, currently, we are the main player when it comes to CO_2 sources. Many ecosystems including tropical forests are already showing signs of stress and could provide unwanted surges of carbon into the atmosphere even if the direct human contribution subsides (Wunderling et al., 2022). The uptake or outgassing of CO_2 from oceans is also subject to uncertainties and possible ocean circulation changes (e.g. Li et al., 2023). Existing carbon-cycle models typically do not allow for vegetation die-off or other tipping-point behaviours in both land and ocean ecosystems, making this a key uncertainty especially in weaker mitigation scenarios.

The above arguments also consider only the global average, as does ECS itself. In particular, our inability to explain global-scale SST changes in recent decades not only gets in the way of inferring ECS but does not bode well for future projections of regional climate change even if ECS is known. In the coming decades, further adjustments of regional patterns of SST, land surface temperatures, and sea-ice extents could cause substantial regional climate changes even if further increases in global-mean temperature were kept small, and will for example depend on whether unexplained warming patterns persist, or reverse. In general, because we don't live in a static climate and we are uncertain about changes in variability in a future climate, this supports the idea that ECS is not the be-all and end-all metric of future climate change severity. Better explaining these regional changes takes on an urgent role regardless of whether future emissions are high or low, although for different reasons. Any help in reducing *a priori* uncertainty in present-day anthropogenic forcing (i.e., from aerosols) would likewise help both objectives, given the ability of aerosols to drive regional variations in climate and to interfere with our ability to infer ECS.

6 Conclusions

In general, our understanding of cloud feedback and other global climate processes is advancing rapidly. Sixty years ago, the first climate models were being developed and the most pressing question was: How much will Earth warm when we increase CO_2 concentrations? Since then, in every decade, the climate science community has been providing new estimates of plausible bounds on ECS given the models, the theory, and the data over the past five decades. By now this has led to a substantial improvement in our knowledge of climate sensitivity.

However, each new study is not automatically translating into narrower uncertainty in ECS or future change, in spite of what some of them claim. Such convergence will be a slow process that will require continuing development of frameworks to benchmark our understanding and, in particular, to identify and recognise structural uncertainties, missing or new forcings and feedbacks, etc. In statistical inference, it is crucial to test the assumptions of a statistical model (for example that residuals are independent and normally distributed), and if they do not hold, to develop a different model. Inference of ECS from evidence is an analogous exercise, except using more physically motivated models. These models are not perfect; it is important always to assess whether their errors could substantially affect the result—i.e., that they be "fit for purpose." And, to be useful for decision making, our probabilistic estimates of ECS need to include all uncertainties, not just those we can comfortably estimate with standard statistical tools. To be objectively confident about further narrowing will therefore require us to look broadly at the climate system and seek a diversity of perspectives and approaches (Harding, 1995).

Interesting questions for the research community to consider going forward are: Are the climate model components catching up to be able to diagnose feedbacks in a way that is testable against observations? Are the differences in the current climate models converging towards the observable data (and are we sure this isn't due to tuning)? Are models exhibiting the full diversity of behaviour necessary to capture climate-relevant variations (e.g., in surface temperature and cloud cover) to within observational uncertainty? Are we observing the right quantities well enough to test our models (and scientific understanding) in ways needed to constrain predictions?

All relevant lines of evidence show potential to improve our estimates of ECS over time. While use of the historical record has suffered some blows due to previously unappreciated problems in its interpretation, it is likely to make a comeback. For one thing, the historical climate record keeps getting longer. For another, progress is being made in reconciling model warming and cloudiness change patterns with observed data (Zhu et al., 2022; Raghuraman et al., 2023) and this should eventually lead to more confident constraints than have been possible before, using geographically and/or time-resolved approaches. These estimates will clearly benefit from better observations of atmospheric, ocean, and carbon cycle processes. We hope that by the time of AR7 this will have led to further confidence in not only ECS, but also for other ingredients that will determine our future climate.

Author contributions. Both authors wrote this article.

Competing interests. The authors declare no competing interests

Acknowledgements. SCS acknowledges all of his co-authors on the 2020 WCRP assessment for providing so many useful insights into the problem of ECS. CEF acknowledges all of his co-authors for their roles related to improving estimates of ECS.

References

Andrews, T., Bodas-Salcedo, A., Gregory, J. M., Dong, Y., Armour, K. C., Paynter, D., Lin, P., Modak, A., Mauritsen, T., Cole, J. N. S., Medeiros, B., Benedict, J. J., Douville, H., Roehrig, R., Koshiro, T., Kawai, H., Ogura, T., Dufresne, J.-L., Allan, R. P., and Liu, C.: On the Effect of Historical SST Patterns on Radiative Feedback, J. Geophys. Res., 127, https://doi.org/10.1029/2022JD036675, 2022.

Andronova, N. and Schlesinger, M.: Objective estimation of the probability density function for climate sensitivity, J. Geophys. Res., 106, 22 605–22 611, https://doi.org/10.1029/2000JD000259, 2001.

Arrhenius, S.: On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground, Phil. Mag., 41, 237–276, 1896.

Brown, R. M., Chalk, T. B., Crocker, A. J., Wilson, P. A., and Foster, G. L.: Late Miocene cooling coupled to carbon dioxide with Pleistocene-like climate sensitivity, Nat. Geosci., 15, 664+, https://doi.org/10.1038/s41561-022-00982-7, 2022.

Ceppi, P. and Nowack, P.: Observational evidence that cloud feedback amplifies global warming, Proc. Nat. Acad. Sci. US., 118, https://doi.org/10.1073/pnas.2026290118, 2021.

Chao, L.-W., Muller, J. C., and Dessler, A. E.: Impacts of the Unforced Pattern Effect on the Cloud Feedback in CERES Observations and Climate Models, Geophys. Res. Lett., 49, https://doi.org/10.1029/2021GL096299, 2022.

Charney, J., Arakawa, A., Baker, D. J., Bolin, B., Dickenson, R. E., Goody, R. M., Leith, C. E., Stommel, H. M., and Wunsch, C. I.: Carbon Dioxide and Climate:A scientific assessment, Tech. rep., National Academy of Sciences, Woods Hole, MA, 1979.

Dvorak, M. T., Armour, K. C., Frierson, D. M. W., Proistosescu, C., Baker, M. B., and Smith, C. J.: Estimating the timing of geophysical commitment to 1.5 and 2.0 degrees C of global warming, Nat. Clim. Change, 12, 547+, https://doi.org/10.1038/s41558-022-01372-y, 2022.

Edwards, T. L., Crucifix, M., and Harrison, S. P.: Using the past to constrain the future: how the palaeorecord can improve estimates of global warming, Prog. Phys. Geog., 31, 481–500, 2007.

Forest, C. E., Stone, P. H., Sokolov, A. P., Allen, M. R., and Webster, M. D.: Quantifying uncertainties in climate system properties with the use of recent climate observations, Science, 295, 113–117, 2002.

Gleckler, P. J., Durack, P. J., Stouffer, R. J., Johnson, G. C., and Forest, C. E.: Industrial-era global ocean heat uptake doubles in recent decades, Nature Climate Change, 6, 394–398, https://doi.org/10.1038/nclimate2915, 2016.

Grose, M. R., Gregory, J., Colman, R., and Andrews, T.: What Climate Sensitivity Index Is Most Useful for Projections?, Geophys. Res. Lett., 45, 1559–1566, https://doi.org/10.1002/2017GL075742, 2018.

Hall, A., Cox, P., Huntingford, C., and Klein, S.: Progressing emergent constraints on future climate change, Nat. Clim. Change, 9, 269–278, https://doi.org/10.1038/s41558-019-0436-6, 2019.

Harding, S.: "Strong objectivity": a response to the new objectivity question, Synthese, 104, 331–349, 1995.

Heede, U. K. and Fedorov, A. V.: Eastern equatorial Pacific warming delayed by aerosols and thermostat response to CO_2 increase, Nat. Clim. Change, 11, 696+, https://doi.org/10.1038/s41558-021-01101-x, 2021.

Hegerl, G. C. and Wallace, J. M.: Influence of Patterns of Climate Variability on the Difference between Satellite and Surface Temperature Trends, J. Climate, 15, 2412–2428, 2002.

Knutti, R., Stocker, T., Joos, F., and Plattner, G.: Constraints on radiative forcing and future climate change from observations and climate model ensembles, Nature, 416, 719–723, https://doi.org/10.1038/416719a, 2002.

Kramer, R. J., Soden, B. J., and Pendergrass, A. G.: Evaluating Climate Model Simulations of the Radiative Forcing and Radiative Response at Earth's Surface, Journal of Climate, 32, 4089 – 4102, https://doi.org/https://doi.org/10.1175/JCLI-D-18-0137.1, 2019.

Kramer, R. J., He, H., Soden, B. J., Oreopoulos, L., Myhre, G., Forster, P. M., and Smith, C. J.: Observational Evidence of Increasing Global Radiative Forcing, Geophysical Research Letters, 48, e2020GL091 585, https://doi.org/https://doi.org/10.1029/2020GL091585, e2020GL091585 2020GL091585, 2021.

Kuma, P., Bender, F. A. M., Schuddeboom, A., McDonald, A. J., and Seland, O.: Machine learning of cloud types in satellite observations and climate models, Atmos. Chem. Phys., 23, 523–549, https://doi.org/10.5194/acp-23-523-2023, 2023.

Lewis, N.: Objectively combining climate sensitivity evidence, Clim. Dyn., https://doi.org/10.1007/s00382-022-06468-x, 2022.

Li, Q., England, M. H., Hogg, A. M., Rintoul, S. R., and Morrison, A. K.: Abyssal ocean overturning slowdown and warming driven by Antarctic meltwater, Nature, 615, 841–847, 2023.

Libardoni, A. G., Forest, C. E., Sokolov, A. P., and Monier, E.: Underestimating Internal Variability Leads to Narrow Estimates of Climate System Properties, Geophysical Research Letters, 46, 10 000–10 007, https://doi.org/https://doi.org/10.1029/2019GL082442, 2019.

Loeb, N. G., Wang, H., Allan, R. P., Andrews, T., Armour, K., Cole, J. N. S., Dufresne, J.-L., Forster, P., Gettelman, A., Guo, H., Mauritsen, T., Ming, Y., Paynter, D., Proistosescu, C., Stuecker, M. F., Willen, U., and Wyser, K.: New Generation of Climate Models Track Recent Unprecedented Changes in Earth's Radiation Budget Observed by CERES, Geophys. Res. Lett., 47, https://doi.org/10.1029/2019GL086705, 2020.

Mahowald, N. M., Li, L., Albani, S., Hamilton, D. S., and Kok, J.: Opinion: The importance of historical and paleoclimate aerosol radiative effects, Atm. Chem. Phys. Disc., https://doi.org/10.5194/egusphere-2023-1174, 2023, 2023.

Meehl, G. A., Hu, A., Castruccio, F., England, M. H., Bates, S. C., Danabasoglu, G., McGregor, S., Arblaster, J. M., Xie, S.-P., and Rosenbloom, N.: Atlantic and Pacific tropics connected by mutually interactive decadal-timescale processes, Nat. Geosci., 14, 36–43, https://doi.org/10.1038/s41561-020-00669-x, 2021.

Myers, T. A., Scott, R. C., Zelinka, M. D., Klein, S. A., Norris, J. R., and Caldwell, P. M.: Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity, Nature Climate Change, 11, 501–507, 2021.

PALAEOSENS Project Members: Making sense of palaeoclimate sensitivity, Nature, 491, 683–691, 2012.

Peixoto, J. P. and Oort, A. H.: Physics of Climate, American Institute of Physics, New York, 1992.

Raghuraman, S. P., Paynter, D., Menzel, R., and Ramaswamy, V.: Forcing, cloud feedbacks, cloud masking, and internal variability in the cloud radiative effect satellite record, J. Climate, pp. 1–38, https://doi.org/10.1175/JCLI-D-22-0555.1, 2023.

Renoult, M., Annan, J. D., Hargreaves, J. C., Sagoo, N., Flynn, C., Kapsch, M.-L., Mikolajewicz, U., Ohgaito, R., and Mauritsen, T.: A Bayesian framework for emergent constraints: case studies of climate sensitivity with PMIP, Clim. Past, 16, 1715–1735, https://doi.org/https://doi.org/10.5194/cp-16-1715-2020, 2020.

Scafetta, N.: Advanced Testing of Low, Medium, and High ECS CMIP6 GCM Simulations Versus ERA5-T2m, Geophys. Res. Lett., 49, https://doi.org/10.1029/2022GL097716, 2022.

Seltzer, A. M., Blard, P.-H., Sherwood, S. C., and Kageyama, M.: Terrestrial amplification of past, present, and future climate change, Sci. Adv., p. eadf8119, https://doi.org/10.1126/sciadv.adf811, 2023.

Senior, C. A. and Mitchell, J. F. B.: The time-dependence of climate sensitivity, Geophys. Res. Lett., 27, 2685–2688, 2000.

Sherwood, S. C., Sen Gupta, A., and Schwartz, S. E.: Probability of committed warming exceeding 1.5 °C and 2.0 °C Paris targets, Env. Res. Lett., 17, https://doi.org/10.1088/1748-9326/ac6ff6, 2022.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka, M. D.: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence, Rev. Geophys., 58, https://doi.org/10.1029/2019RG000678, 2020.

Smith, C. J., Forster, P. M., Allen, M., Fuglestvedt, J., Millar, R. J., Rogelj, J., and Zickfeld, K.: Current fossil fuel infrastructure does not yet commit us to 1.5 °C warming, Nature Communications, 10, 101, 2019.

Sokolov, A. P. and Stone, P. H.: A flexible climate model for use in integrated assessments, Clim. Dyn., 14, 291–303, 1998.

Stainforth, D. A., Allen, M. R., Tredger, E. R., and Smith, L. A.: Confidence, uncertainty and decision-support relevance in climate predictions, Phil. Trans. Royal Soc. A, 365, 2145–2161, https://doi.org/10.1098/rsta.2007.2074, 2007.

Stauffer, C. L. and Wing, A. A.: Properties, Changes, and Controls of Deep-Convecting Clouds in Radiative-Convective Equilibrium, J. Adv. Model. Earth Sys., 14, https://doi.org/10.1029/2021MS002917, 2022.

von der Heydt, A. S., Dijkstra, H. A., van de Wal, R. S. W., Caballero, R., Crucifix, M., Foster, G. L., Huber, M., Köhler, P., Rohling, E., Valdes, P. J., Ashwin, P., Bathiany, S., Berends, T., van Bree, L. G. J., Ditlevsen, P., Ghil, M., Haywood, A. M., Katzav, J., Lohmann, G., Lohmann, J., Lucarini, V., Marzocchi, A., Pälike, H., Baroni, I. R., Simon, D., Sluijs, A., Stap, L. B., Tantet, A., Viebahn, J., and Ziegler, M.: Lessons on Climate Sensitivity From Past Climate Changes, Curr Clim Change Rep, 2, 148–158, 2016.

Wall, C. J., Norris, J. R., Possner, A., McCoy, D. T., McCoy, I. L., and Lutsko, N. J.: Assessing effective radiative forcing from aerosol-cloud interactions over the global ocean, Proc. Nat. Acad. Sci. US., 119, https://doi.org/10.1073/pnas.2210481119, 2022a.

Wall, C. J., Storelvmo, T., Norris, J. R., and Tan, I.: Observational Constraints on Southern Ocean Cloud-Phase Feedback, J. Climate, 35, 5087–5102, https://doi.org/10.1175/JCLI-D-21-0812.1, 2022b.

Williams, A. I. L., Stier, P., Dagan, G., and Watson-Parris, D.: Strong control of effective radiative forcing by the spatial pattern of absorbing aerosol, Nature Clim. Change, 12, 735+, https://doi.org/10.1038/s41558-022-01415-4, 2022.

Wunderling, N., Staal, A., Sakschewski, B., Hirota, M., Tuinenburg, O. A., Donges, J. F., Barbosa, H. M. J., and Winkelmann, R.: Recurrent droughts increase risk of cascading tipping events by outpacing adaptive capacities in the Amazon rainforest, Proc. Nat. Acad. Sci. US., 119, https://doi.org/10.1073/pnas.2120777119, 2022.

Zhu, J., Otto-Bliesner, B. L., Brady, E. C., Poulsen, C. J., Tierney, J. E., Lofverstrom, M., and DiNezio, P.: Assessment of Equilibrium Climate Sensitivity of the Community Earth System Model Version 2 Through Simulation of the Last Glacial Maximum, Geophys. Res. Lett., 48, https://doi.org/10.1029/2020GL091220, 2021.

Zhu, J., Otto-Bliesner, B. L., Brady, E. C., Gettelman, A., Bacmeister, J. T., Neale, R. B., Poulsen, C. J., Shaw, J. K., McGraw, Z. S., and Kay, J. E.: LGM Paleoclimate Constraints Inform Cloud Parameterizations and Equilibrium Climate Sensitivity in CESM2, J. Adv. Model. Earth Sys., 14, https://doi.org/10.1029/2021MS002776, 2022.