https://doi.org/10.5194/egusphere-2025-1179
Preprint. Discussion started: 24 March 2025

Comment on "Can uncertainty in climate sensitivity be narrowed further?" by Sherwood and Forest (2024)

Nicholas Lewis[1]

[1] Bath, UK; independent of any institution

Correspondence to: Nicholas Lewis (nhlewis@btinternet.com)

Abstract. This comment addresses assertions made by Sherwood and Forest (2024) [SF24] regarding the narrowing of the range of equilibrium climate sensitivity (ECS), particularly at the low end. SF24 challenged a previous study by Lewis (2022) [L22] that found a narrower and substantially lower ECS level. This comment clarifies that, contrary to SF24's claims, L22 did not rule out a high ECS level based on historical evidence, and did identify and correct errors in Sherwood et al. (2020). Those errors included use of an invalid likelihood estimation method that, ironically, substantially underestimated likelihood at high ECS levels for their historical evidence. This comment also discusses the role of priors in Bayesian ECS estimation and explains why the subjective Bayesian approach favoured by SF24 risks producing unreliable inference for uncertain parameters such as ECS. Finally, the importance of considering structural uncertainties in climate models, particularly concerning tropical warming patterns, is extended beyond the points raised by SF24. Such uncertainties could affect ECS estimation not only from historical period evidence but also from climate process understanding, paleoclimate data and emergent constraints, but seem more likely to suggest existing ECS estimates are too high than too low.

1 Introduction

In the Sherwood and Forest (2024) Opinion article "Can uncertainty in climate sensitivity be narrowed further?" published in Atmospheric Chemistry and Physics (hereafter SF24), the authors express doubts that the uncertainty range for equilibrium climate sensitivity (ECS) has been further narrowed since the publication of Sherwood et al. (2020) (hereafter S20). They note that the observationally driven ECS range in S20 was approximately adopted in the relevant chapter (Forster et al., 2021) of the IPCC Sixth Assessment Report (AR6).

SF24's authors state that the new study claiming the largest revision in the range for ECS is Lewis (2022) (hereafter L22), which "asserts a narrower and substantially lower ECS level using the basic S20 methodology with various updates". This comment addresses the erroneous claims that SF24 made about L22, addresses problems with the subjective Bayesian approach described by SF24, and discusses the challenges in narrowing the range of climate sensitivity posed by uncertainties in the realism of global climate model (GCM) simulated long-term tropical warming patterns.

https://doi.org/10.5194/egusphere-2025-1179
Preprint. Discussion started: 24 March 2025

2 Critique of SF24's Claims Regarding L22

SF24 states concerning L22:

"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."

There were indeed differences in opinion on Bayesian priors between L22 and S20. And S20's decision to not adjust for CO_2 forcing increasing slightly faster than logarithmically with concentration, unlike L22, is arguably a defensible difference of opinion. The minor effects on ECS estimation of these two differences of opinion almost cancelled out.

However, L22 also found actual errors, inconsistencies and indefensible methodological choices in S20. L22 Sections 5.1 and Supporting Information S2 showed that S20's method of likelihood estimation was invalid, and resulted in major underestimation of the historical evidence likelihoods at high climate sensitivities. Moreover, L22 Sections 4.1 and Supporting Information S1 showed that, based on an ensemble of 26 GCMs, S20's ECS estimates from both Process and Historical evidence were biased approximately 16% high, due to S20 inappropriately using a fixed sea surface temperature (SST) simulation based rather than a regression based estimate of doubled CO_2 forcing ($F_{2\times CO2}$) to derive ECS from estimated climate feedback. As L22 showed, in general GCM $F_{2\times CO2}$ estimates based on the land-warming corrected fixed SST simulation method (which it found agreed very closely, at their median, with the AR6 best estimate of $F_{2\times CO2}$) are significantly higher than 150-year regression based $F_{2\times CO2}$ estimates. The latter are based on the *y*-intercept from a linear regression of planetary radiative imbalance (N) against the increase in global mean surface temperature (ΔGMST) as simulated in a GCM over 150 years following an abrupt quadrupling in CO_2 concentration (abrupt4xCO2), after scaling to a doubling of CO_2, with the *x*-intercept taken as the GCM's estimated ECS. In almost all GCMs, climate feedback weakens over the first few decades of the 150 year simulation period, reflecting evolving tropical warming patterns (a forced pattern effect). That being so, N is not linearly related to ΔGMST, and such regression based $F_{2\times CO2}$ values are bound to underestimate true $F_{2\times CO2}$.[1] Nevertheless, to be arithmetically correct, ECS estimates intended (as they are in both S20 and L22) to be on the same basis as 150-year regression based GCM ECS estimates, and therefore calculated by dividing observationally derived estimates of the slope in such a regression into an estimate of $F_{2\times CO2}$, require that the $F_{2\times CO2}$ estimate used is likewise regression based.

SF24 criticizes L22's results by claiming:

"Instead, the reduction and narrowing of the ECS probability density function (PDF) resulted from a selective use of evidence - most importantly, a decision to reject the possibility of a large "pattern effect" on historical sea surface temperature (SST), even though this continues to be strongly supported by new studies, 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 a high ECS level."

[1] Reflecting the cause of this underestimation, the data in L22 Table S1 show a 0.98 correlation across the 26 GCMs between the measure of the forced pattern effect and the degree to which regression based $F_{2\times CO2}$ underestimates fixed SST based $F_{2\times CO2}$.

https://doi.org/10.5194/egusphere-2025-1179
Preprint. Discussion started: 24 March 2025

This claim is totally false. On the contrary, L22 finds (Table 8) that the historical record does not rule out a high ECS level. The standard 90% (5% - 95%) uncertainty range that L22 arrived at using only data from the historical record was 1.2 °C to 7.6 °C. Where, in addition, the common prior assumption that ECS is positive and does not exceed 20 °C was made, the range became 1.15 °C to 6.1 °C. Neither of these ranges rule out a high ECS level. Both 7.6 °C and 6.1 °C substantially exceed S20's main combined-evidence ECS 95% bound of 4.7 °C.

S20's combined all-evidence median (50% probability) baseline ECS estimate was 3.1 °C, with 66% ('likely') and 90% ranges of respectively 2.6 - 3.9 °C and 2.3 - 4.7 °C. The corresponding values in L22 were 2.16 °C, 1.75 - 2.7 °C, and 1.55 - 3.2 °C. Without the two revisions criticized by SF24, the L22 all-evidence ECS 5% bound and median estimates would have been unchanged and the 95% bound would have increased by only 0.1 °C. Excluding historical evidence entirely would similarly have left the L22 baseline ECS 5% bound and median estimates unchanged, with the 95% bound increasing by only 0.2 °C. Thus, the reduction and narrowing of the ECS PDF in L22 had almost nothing to do with the revisions it made to assumptions about the pattern effect, or aerosol cooling, over the historical period.

It is not surprising that the historical record does not rule out a high ECS level, since, in both S20 and L22, aerosol forcing is the largest source of uncertainty. Aerosol forcing reduces, to an uncertain extent, the magnitude of the denominator of the energy budget formula for estimating ECS. That results in there being a non-negligible probability, indeed a substantial one on S20's assumptions, that the denominator is small or negative, implying a very high or unbounded ECS, but a negligible probability that the denominator is very large, thus ruling out ECS being very low. It is surprising that the SF24 authors failed to recognize that it is process and, particularly for S20, paleoclimate evidence that most constrained the ECS upper bound both in L22 (Table 8) and in S20 (Table 10).

Both S20 and L22 also estimated an effective climate sensitivity over the historical period (S_{hist}), which is not adjusted for any pattern effect, and is somewhat less affected than ECS by assumptions about aerosol cooling. Employing a sampling method, S20 derived a 90% range for S_{hist} of 1.9 to 14.4 °C, with a median of 3.1 °C, using an aerosol forcing distribution that assigned a 16% probability to it being even stronger (more negative) than −2.0 Wm^{-2}. L22 down weighted the probability of very strong aerosol forcing from that assumed in S20, but without assuming the most likely level of aerosol forcing was weaker than in S20, and derived a 90% range for S_{hist} of 1.3 to 4.3 °C (median 2.1 °C). As noted in S20, without criticism, Tokarska et al. (2020) likewise effectively down weighted very strong aerosol forcing as less consistent with observations. That resulted in their 90% range for S_{hist} being 1.3 to 3.1 °C, with a median of 2.1 °C – identical results to L22 save that the 95% bound in L22 was considerably higher.

The pattern effect – the dependence of outgoing radiation to space on the geographical pattern of SST warming – is thought to result in ECS being higher than S_{hist}. Of the three studies that SF24 cites as supporting a large historical pattern effect, neither Heede and Fedorov, 2021, nor Chao et al., 2022, consider the full historical period (from the second half of the 19th century on), which is what S20 and L22 both use. They only consider respectively post-1980 and post-2000 periods, and thus do not in fact support SF24's claims. Andrews et al. (2022) did estimate the pattern effect over the full historical period

https://doi.org/10.5194/egusphere-2025-1179
Preprint. Discussion started: 24 March 2025

(1871 - 2010), as 0.48 $Wm^{-2}°C^{-1}$, in line with the S20 estimate of 0.5 $Wm^{-2}°C^{-1}$, when regressing annual mean historical simulation data from atmosphere-only GCMs driven by SST evolution in the non-spliced HadiSST1 dataset rather than in the outlier AMIPII SST dataset. The Andrews et al. (2022) estimate reduces to 0.41 $Wm^{-2}°C^{-1}$ when regressing pentadal mean data, to suppress bias from responses to interannual fluctuations (Lewis and Mauritsen 2021). If the resulting individual model estimates are weighted equally by modelling centre rather than by model, recognising that models from the same centre have structural similarities, that 0.41 $Wm^{-2}°C^{-1}$ estimate reduces further to 0.36 $Wm^{-2}°C^{-1}$, in line with the L22 estimate of 0.35 $Wm^{-2}°C^{-1}$.

3 Subjective vs. Objective Bayesian Approaches

Regarding statistical issues, as noted by SF24 priors on ECS and other climate system parameters have been a contentious issue since the first Bayesian ECS studies. SF24 points out that most climate scientists have had little formal training in probability or statistics – which perhaps accounts for none of the S20 authors or peer reviewers realising that the S20 likelihood estimation method was invalid. SF24 argues that probability is most useful as a quantification of what someone expects, rather than a quasi-objective calculation based on the chosen physical models and data used. That is, they favour a Subjective Bayesian rather than an Objective Bayesian or a frequentist approach. However, as explained in L22, it is essential for scientific inference that the statistical methods used are calibrated, in the sense that the uncertainty ranges they generate closely approximate confidence intervals. Objective Bayesian methods involve use of mathematical, noninformative priors that are generally intended to produce uncertainty ranges that are (at least approximately) true confidence intervals; frequentist methods share that intention. Subjective Bayesian methods are not designed to do so, and their uncertainty ranges may be very ill-calibrated when (as in ECS estimation) the data are insufficiently strong to dominate the influence of the prior. S20 used a Subjective Bayesian approach, while L22 employed an Objective Bayesian approach.
There are also other problems with adopting a Subjective Bayesian approach. First, the prior and data-likelihood may well both be based on the same evidence, at least to some extent. Scientists' expectations as to the value of ECS can only rationally be based on some set of observational evidence and/or on climate model behaviour. It would be reasonable to use the well-established fact that the climate system is not extremely unstable to require a prior to rule out ECS being negative or exceedingly high (over 20 °C, say). However, a considerable part of the observational data-based evidence available to scientists will already be reflected in the likelihood function(s) produced by the physical models and data used for estimating ECS, so it is duplication to also use it in formulating a prior. Moreover, since observational evidence is used in the construction of climate models, and because related aspects of climate model behaviour will have been used to quantify data-variable distributions used in estimating the ECS likelihood, climate model behaviour should not shape the prior, particularly given the concern that ECS in such models may be unrealistic. Secondly, even if the chosen prior does produce well-calibrated uncertainty ranges when used to infer ECS from a data-likelihood reflecting one set of evidence, the standard

https://doi.org/10.5194/egusphere-2025-1179
Preprint. Discussion started: 24 March 2025

Bayesian method of simply updating that initial posterior PDF by data-likelihoods reflecting other sets of evidence may well not produce well-calibrated uncertainty ranges reflecting the combined evidence (Lewis 2018; Lewis and Grunwald 2018). For historical and, particularly, process evidence, the uniform (flat) prior for climate feedback that S20 selected was in fact close to the mathematical, noninformative prior used in L22, although for paleoclimate evidence it differed substantially (L22 Figure 3). However, the uniform prior for ECS selected for use with observational evidence in the IPCC Fourth Assessment Report was very far from noninformative, and resulted in huge overestimation of the chance of ECS being very high; only the upper bound placed on that prior (20 °C for most studies) prevented almost 100% probability being assigned to near-infinite ECS values. The same applies to S20's ECS estimate from historical evidence when using a 0 to 20 °C uniform prior for ECS (Table 6); the median estimate was 8.5 °C, with a 95% bound (existing purely due to the prior upper limit of 20 °C) of 18.6 °C. By contrast, when using a noninformative prior with S20's historical evidence inputs, the L22 (Table 4) median ECS estimate was only 4.2 °C, with a 95% bound of 13.7 °C, despite being after correcting S20's underestimation of its historical evidence likelihood at high ECS levels.

4 Structural Uncertainties in Climate Models

SF24 raises the issue of structural uncertainty in GCMs and other forward models. SF24 notes that the inability of GCMs to simulate the pattern of warming over recent decades could have explanations that have very different ramifications for ECS from those if the pattern of warming is a transient phenomenon or due to missing aerosol forcing mechanisms. Specifically, if the reduction in the Pacific east-west temperature gradient that occurs after a decade or two in almost all GCM CO_2-forced simulations, which underlies the pattern effect, is unrealistic, as some recent studies suggest it may be (Lee et al. 2022, Lee et al. 2024), the weakening of net climate feedback in GCMs over 150 years after an abrupt quadrupling of CO_2 concentration would not occur. That would imply that most current ECS estimates, including those based on process understanding and on emergent constraints as well as on historical warming – all of which generally reflect the aforesaid weakening of climate feedback in GCMs – may be significantly biased upwards. Paleoclimate ECS estimates adjusted on the assumption that ECS is lower in cooler climates, as in the S20 and L22 LGM estimates, might also be biased upwards. Hua et al. (2024) found that, contrary to GCM simulations, the Pacific east-west temperature gradient was weaker at the LGM than preindustrially, consistent with SST in the West Pacific warm pool being more sensitive to greenhouse gas forcing than SST in the eastern equatorial Pacific (Seager et al. 2019), and suggesting ECS is higher, not lower, in cooler climates. The possibility that long term tropical warming patterns simulated by GCMs are significantly wrong could be one of the most important of the omitted structural uncertainties in ECS estimation about which SF24 expresses concern. However, that uncertainty points primarily to the possibility that ECS estimates are too high, not too low. Even if the Pacific east-west temperature gradient, and hence net climate feedback, does eventually weaken to the extent simulated by GCMs, a multidecadal to centennial delay in that weakening occurring could imply a significantly lower warming response this century, as a fraction of ECS, than would otherwise be the case.

https://doi.org/10.5194/egusphere-2025-1179
Preprint. Discussion started: 24 March 2025

SF24 claims that probability distributions for ECS in S20 and AR6 remain approximately valid but that subsequent studies, including L22, omit important structural uncertainties. However, L22 did not omit any structural uncertainties that were included in S20. The equations used in L22 for estimating ECS were identical to those in S20, save for the inclusion of an additional structural uncertainty concerning the ratio of regression based to fixed SST based estimates of $F_{2\times CO2}$. Moreover, the uncertainty estimates for each input variable in L22 were almost all the same or greater than those used in S20, except for the historical pattern effect and aerosol forcing – the changed input distributions of which had a negligible effect on the L22 combined evidence ECS estimation.

5 Conclusions

S20 substantially narrowed the uncertainty in ECS, primarily by rejecting lower values of ECS that had been included in climate assessments since 1979 and through the 2013 IPCC Fifth Assessment Report, and its observationally-driven ECS approach and range was approximately adopted in IPCC AR6. By pointing out deficiencies in the S20 analysis and adopting input data estimates based on newer evidence and in a few cases alternative appraisals of existing evidence, L22 provided an analysis that re-emphasized the lower values of ECS. While S20 estimated the probability of ECS being below 2.3°C as only 5%, L22 estimated it to be over 50%. S24 provides no evidence nor valid argument against L22's estimate.

The reluctance of climate scientists to move away from mathematically unsound Subjective Bayesian approaches for estimating ECS or other uncertain climate system parameters is deeply concerning. Obtaining reliable results requires either adopting an Objective Bayesian approach, employing a noninformative prior, or using frequentist statistical methods.

While structural uncertainty in forward models, in particular GCMs, may well be underestimated, SF24's claim that probability distributions for ECS in S20 remain approximately valid but that subsequent studies omit important structural uncertainties cannot be justified in respect to L22.

Code and data availability

Simulation data used to re-estimate HadiSST1 driven pattern effect results in Andrews et al. (2022) using pentadal mean ordinary least squares regression was downloaded from https://zenodo.org/records/6799004 on 15 February 2025. The regression code used was the lm function in R.

Competing interests

The contact author has declared that he does not have any competing interests.

Acknowledgements

The author acknowledges Timothy Andrews for publicly archiving global annual mean simulation data used in Andrews et al. (2022).

Financial support

No financial support was obtained for this work.

References

Andrews, T., Bodas-Salcedo, A., Gregory, J. M., Dong, Y., Armour, K., Paynter, D., Lin, P., Modak, A., Mauritsen, T., Cole, J. N., and Medeiros, B.: On the effect of historical SST patterns on radiative feedback, J. Geophys. Res.-Atmos., 127, e2022JD036675, https://doi.org/10.1029/2022JD036675, 2022.

Bradshaw, C. D., Lunt, D. J., Flecker, R., and Davies-Barnard, T.: Disentangling the roles of late Miocene palaeogeography and vegetation – Implications for climate sensitivity, Palaeogeogr. Palaeoclimatol. Palaeoecol., 417, 17-34, https://doi.org/10.1016/j.palaeo.2014.10.003, 2015.

Forster, P., Storelvmo, T., Armour, K. C., Collins, W., Dufresne, J-L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The Earth's energy budget, climate feedbacks, and climate sensitivity. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V. et al. (eds.)], Cambridge University Press, 2021.

Hou, A., Jonkers, L., Ford, H. L., and Ho, S. L.: El Nino-like tropical pacific ocean cooling pattern during the last glacial maximum, Commun. Earth Environ., 5, 587, https://doi.org/10.1038/s43247-024-01740-w, 2024.

Lee, S., L'Heureux, M., Wittenberg, A.T., Seager, R., O'Gorman, P. A., Johnso, N. C.: On the future zonal contrasts of equatorial Pacific climate: Perspectives from Observations, Simulations, and Theories. npj Clim Atmos Sci 5, 82 (2022). https://doi.org/10.1038/s41612-022-00301-2

Lee, S., Bannon, P.R., Park, M., Clark, J.P. Zonal Contrasts of the Tropical Pacific Climate Predicted by a Global Constraint. Asia-Pac J Atmos Sci 60, 669–678 (2024). https://doi.org/10.1007/s13143-024-00373-5

Lewis, N.: Combining independent Bayesian posteriors into a confidence distribution, with application to estimating climate sensitivity, J. Stat. Plan. Infer., 195, 80-92, https://doi.org/10.1016/j.jspi.2017.09.013, 2018.

Lewis, N.: Objectively combining climate sensitivity evidence, Clim. Dynam., 60, 3139-3165, https://doi.org/10.1007/s00382-022-06468-x, 2022.

Lewis, N., and Grunwald, P.: Objectively combining AR5 instrumental period and paleoclimate climate sensitivity evidence, Clim. Dynam., 50, 2199-2216, https://doi.org/10.1007/s00382-017-3744-4, 2018.

https://doi.org/10.5194/egusphere-2025-1179
Preprint. Discussion started: 24 March 2025

Lewis, N., and Mauritsen, T.: Negligible unforced historical pattern effect on climate feedback strength found in HadISST-based AMIP simulations, J. Climate, 34, 39-55, https://doi.org/10.1175/JCLI-D-19-0941.1, 2021.

Rugenstein, M., Bloch-Johnson, J., Gregory, J., Andrews, T., Mauritsen, T., Li, C. et al: Equilibrium climate sensitivity estimated by equilibrating climate models, Geophys. Res. Lett., 47, e2019GL083898, https://doi.org/10.1029/2019GL083898, 2020.

Seager, R., Cane, M., Henderson, N., Lee, D. E., Abernathey, R., and Zhang, H.: Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases, Nat. Clim. Change, 9, 517-522, https://doi.org /10.1038/s41558-019-0505-x, 2019.

Sherwood, S. C., and Forest, C. E.: Opinion: Can uncertainty in climate sensitivity be narrowed further?, Atmos. Chem. Phys., 24, 2679-2686, https://doi.org/10.5194/acp-24-2679-2024, 2024.

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.

Tokarska, K.B., Hegerl, G.C., Schurer, A.P., Forster, P.M. and Marvel, K.: Observational constraints on the effective climate sensitivity from the historical period. Environmental Research Letters, 15(3), p.034043, 2020.