Matters arising

In response to *Delayed emergence of a global temperature response after emission mitigation*, Samset et al., 2020.[1]

Uncertainty in near-term temperature evolution must not obscure assessments of climate mitigation benefits

Alexandrine Lanson[1], Peter Pfleiderer[1,2], Flavio Lehner[3,4], Carl-Friedrich Schleussner[1,2]

[1] Climate Analytics, Berlin, Germany

[2] IRITHESys, Humboldt University, Berlin, Germany

[3] Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland

[4] Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, USA

In a recent study, Samset et al.[1] reported that due to the imprint of natural variability, the effects of emission mitigation will only be perceived through global temperature with a multi-decadal delay. Their analysis, also including a decomposition into the effects of mitigating individual climate forcers, is highly relevant and timely, but does not fully substantiate all conclusions made by the authors. Here, we provide essential context around the findings by Samset et al.[1] of multi-decadal delays of mitigation benefits. We also express concerns with their conceptual approach towards assessing a discernible warming response under different greenhouse gas concentration pathways. A broader debate on how to best assess and communicate emerging effects of climate mitigation in the light of natural variability seems warranted.

To start simple, increased atmospheric greenhouse gas concentrations lead to increased radiative forcing and thus warming[2]. Important differences exist between long- and short-lived climate forcers that need to be taken into account when assessing warming under different emissions trajectories. But for CO2 as the dominant greenhouse gas, reduced emissions directly translate into reduced warming. The IPCC's Fifth Assessment Report has assessed a transient climate response to emissions (TCRE) of 0.2°C to 0.7°C of global mean temperature (GMT) increase per 1000 Gt CO2[2]. The Representative Concentration Pathways (RCPs) deployed by Samset et al.[1], are distinct greenhouse gas concentration pathways and consequently differences in their mean warming response as modelled by the reduced-complexity climate model MAGICC[3] are evident in Figure 1 of Samset et al.[1] within years and even before 2020.

So why is it then that Samset et al.[1] do conclude that it may take decades for the effects of mitigation to emerge? The explanation lies in the imprint of natural variability on the near-term temperature trajectory and their assessment of its effects. Natural variability, including multi-decadal modes linked to ocean dynamics, dominates the uncertainty of global mean temperature evolution on decadal timescales[4]. Thus, the effects of mitigation might not be immediately perceived when assessing a single GMT trajectory on short time scales.

The results and interpretations of Samset et al.[1] rely on their conception of emergence following an approach by Tebaldi and Friedlingstein[5]. They define the year of emergence of a significant signal as "*the first year when at least 66% of the baseline-scenario pairs are statistically significantly different*"[1] (baseline-scenarios pairs being RCP4.5-mitigation scenario pairs), using a Student's t-test ($p<0.05$).

The methodological choice needs to be critically reviewed in the light of the question it tries to address. Rather than testing for the effects of mitigation on a given warming trajectory, this test assesses when any possible GMT trajectory under a mitigation scenario would be discernible from any possible GMT evolution under a reference scenario (or 66% of those randomly combined samples). This is very different from assessing the actual effects of mitigation on an individual trajectory, or the ensemble response. Naturally, robust differences in such a test will only emerge after the mitigation signal dominates over natural variability. Samset et al.[1] find that it requires about 0.2°C of anthropogenic warming difference for this test to yield robust results of emergence. Discernible differences in climate impacts such as extreme temperature or long-term sea level rise can already be detected for similar GMT differences[6].

The authors argue that their approach is the appropriate way to assess the question of emergence as *"[the] emergence of a climate mitigation signal beyond natural variability can never be proven, as we would be comparing to an unknown, counterfactual world."*[1]

That assertion is at least debatable. A range of well-established approaches exist to assess the anthropogenic warming contribution in the presence of natural variability[4,7,8,9]. At any given point in time, we will thus be able to assess the effects of mitigation on the anthropogenic warming trend (with some uncertainty around it, which will however be much smaller than the irreducible uncertainty portrayed in Samset et al.[1]). It is also worth recalling that the Paris Agreement refers to anthropogenic climate change only, excluding natural variability[10]. So assessments of *"the progress made towards the ambitions of the Paris Agreement"*[1] would not need to rely on approaches like the one proposed by Samset et al.[1]

Furthermore, counterfactual worlds resembling observed modes of natural variability can be assessed to reconcile observed and modelled warming trends. The scientific community has done so extensively when assessing the so-called warming hiatus period in the early 21st century, during which GMT increase had slowed, and the role of natural variability in explaining it[1112]. There is no reason to believe that this will not be possible going forward. Indeed, different strands of detection and attribution research, such as on extreme weather events, are commonly dealing with even stronger presence of natural variability and are perfectly able to quantify partial contributions of anthropogenic climate change[13].

A key methodological challenge of the Samset et al.[1] approach is also apparent as it does not allow for a clear distinction between two factors influencing the test's significance: the magnitude of the forced warming response and the sample size (time series length). In their analysis, Samset et al.[1] assess GMT trajectories from 2021 onward. Thus, the Student's t-test is performed with very small sample sizes in the near future. We have illustrated the effects of warming difference and sample size in Figure 1 for constant warming differences. We find that a minimum of about 10 years (or after 2030) is required to robustly detect a constant 0.15°C temperature difference. For an emerging warming difference between scenarios over time (see Figure 1), robust detection will only be possible considerably later. The core findings of Samset et al.[1] of a delayed emergence of robust differences between the RCPs therefore depend at least in part on statistical effects resulting from their methodological choices rather than climate system uncertainty.

This serves as an illustration of how much methodological choices affect the outcome and require very careful communication and explanation. Policy makers or the general public may not understand the implications of different approaches. Rather than trying to "*manage expectations*"[1], scientists might thus be better served by engaging in a dialogue with policy makers and society on what may constitute relevant mitigation benefits. To avoid communication challenges Samset et al.[1] rightly foresee around the issue of future GMT development, transparency about the methodological approaches and the scope of the analysis is key. Imprecisions can be a disservice to the science-policy interface. Irreducible uncertainty in near-term climate projections must not obscure the messaging around our understanding of the response of the climate system to reducing greenhouse gas emissions.

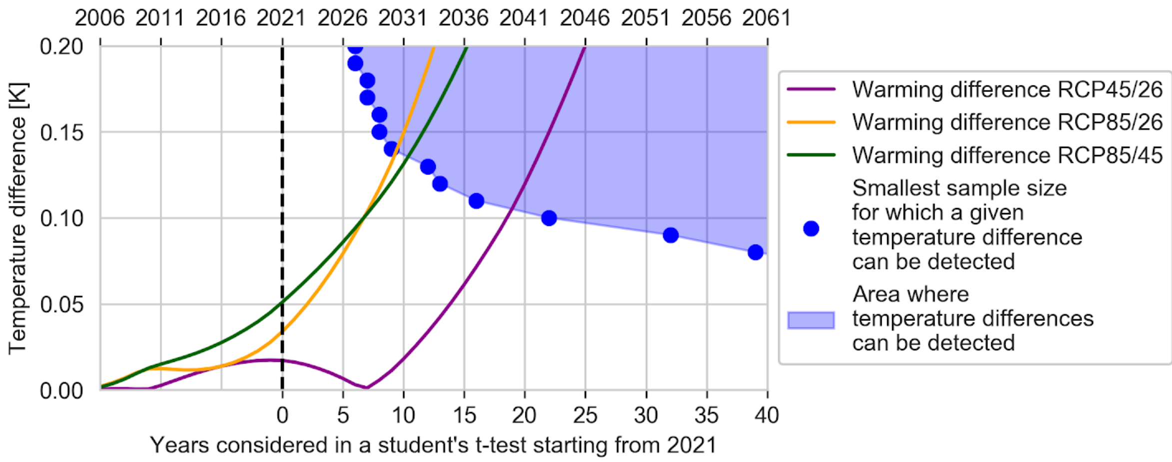

Figure 1: **Detectability of temperature differences using the Samset method:** For different constant temperature differences and different sample sizes, a Student's t-test is applied to pairs of annual global mean temperature variability time series from CESM1 LENS[1415]. Blue dots indicate the minimal sample size for which a given temperature difference is detected by 66% of the pairs at a 0.05 significance level, the significance threshold chosen by Samset et al. The area right of the blue dots is the area for which temperature differences can be detected. Solid lines show emerging warming differences in the forced GMT response between RCP scenarios (from MAGICC6).

Competing interests: The authors declare no competing interests.

Author contribution: AL and CFS conceptualised the response and wrote the manuscript. PP led the analysis on Fig. 1. All authors contributed to the writing process.

[1] Samset, B. H., Fuglestvedt, J. S. & Lund, M. T. Delayed emergence of a global temperature response after emission mitigation. *Nat. Commun.* **11**, 1–10 (2020).

[2] Intergovernmental Panel on Climate Change. Summary for Policymakers. in Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change 1–30 (Cambridge University Press, 2014).

[3] Meinshausen, M., Raper, S. C. B. & Wigley, T. M. L. Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 – Part 1: Model description and calibration. *Atmos. Chem. Phys.* **11**, 1417–1456 (2011).

[4] Lehner, F. *et al.* Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth System Dynamics* **11**, 491–508 (2020).

[5] Tebaldi, C. & Friedlingstein, P. Delayed detection of climate mitigation benefits due to climate inertia and variability. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 17229–17234 (2013).

[6] Pfleiderer, P., Schleussner, C.-F., Mengel, M. & Rogelj, J. Global mean temperature indicators linked to warming levels avoiding climate risks. *Environ. Res. Lett.* **13**, 064015 (2018).

[7] Haustein, K. et al. A real-time Global Warming Index. *Sci. Rep.* **7**, 1–6 (2017).

[8] Lehner, F., Deser, C. & Terray, L. Toward a New Estimate of 'Time of Emergence' of Anthropogenic Warming: Insights from Dynamical Adjustment and a Large Initial-Condition Model Ensemble. *J. Clim.* **30**, 7739–7756 (2017).

[9] Sippel, S. *et al.* Uncovering the Forced Climate Response from a Single Ensemble Member Using Statistical Learning. *J. Clim.* **32**, 5677–5699 (2019).

[10] Rogelj, J., Schleussner, C.-F. & Hare, W. Getting It Right Matters: Temperature Goal Interpretations in Geoscience Research. *Geophysical Research Letters* vol. 44 10,662–10,665 (2017).

[11] Medhaug, I., Stolpe, M. B., Fischer, E. M. & Knutti, R. Reconciling controversies about the 'global warming hiatus'. *Nature* **545**, 41–47 (2017).

[12] Marotzke, J. & Forster, P. M. Forcing, feedback and internal variability in global temperature trends. *Nature* **517**, 565–570 (2015).

[13] Stott, P. A. *et al.* Attribution of extreme weather and climate-related events. *Wiley Interdiscip. Rev. Clim. Change* **7**, 23–41 (2016).

[14] Kay, J. E. *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. *Bull. Am. Meteorol. Soc.* **96**, 1333–1349 (2015).

[15] Sanderson, B. M., Oleson, K. W., Strand, W. G., Lehner, F. & O'Neill, B. C. A new ensemble of GCM simulations to assess avoided impacts in a climate mitigation scenario. *Clim. Change* **146**, 303–318 (2018).