Geophysical Research Letters

RESEARCH LETTER
10.1029/2026GL124347

More Stabilizing Radiative Feedbacks in CO_2 Removal Scenarios

Ivan Mitevski[1] and **Gabriel A. Vecchi**[1,2]

[1]Department of Geosciences, Princeton University, Princeton, NJ, USA, [2]High Meadows Environmental Institute, Princeton University, Princeton, NJ, USA

Key Points:

- Radiative feedbacks are nearly identical during CO_2 increase and removal for low-to-moderate forcing ($\leq 3\times CO_2$)
- At higher forcings ($\geq 4\times CO_2$), feedbacks become more stabilizing (more negative) during the removal phase, driven by cloud feedbacks
- North Atlantic low-level cloud loss from an Atlantic Meridional Overturning Circulation overshoot reduces shortwave reflection, keeping the removal phase warmer

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

I. Mitevski,
mitevski@princeton.edu

Citation:

Mitevski, I., & Vecchi, G. A. (2026). More stabilizing radiative feedbacks in CO_2 removal scenarios. *Geophysical Research Letters*, *53*, e2026GL124347. https://doi.org/10.1029/2026GL124347

Received 13 MAY 2026
Accepted 30 AUG 2026

Abstract Recent studies show that global mean surface temperature exhibits a transient hysteresis during CO_2 removal, where temperatures remain elevated after CO_2 concentrations return to pre-industrial levels. This is largely attributed to ocean thermal inertia rather than an irreversible climate response; however, the role of radiative feedbacks remains unclear. Here, using a fully coupled Earth System Model, we find that radiative feedbacks are nearly identical for CO_2 increase and subsequent removal up to 3 times pre-industrial CO_2 concentrations. However, at $4\times CO_2$ and higher forcings, feedbacks become more stabilizing during CO_2 removal. Through a hierarchy of experiments we show that this is driven by low-level clouds in the North Atlantic, linked to warmer sea-surface temperatures from an overshoot of Atlantic Meridional Overturning Circulation (AMOC) strength during CO_2 removal. These results suggest that AMOC-mediated feedbacks contribute to the transient temperature hysteresis in future CO_2 removal scenarios by keeping the removal phase warmer.

Plain Language Summary Global temperatures do not immediately return to pre-industrial levels after CO_2 is removed, rather, the return takes many years—a temporary (transient) delay often called "hysteresis." While this is often attributed to the ocean's high heat capacity and its ability to resist rapid temperature changes, the role of radiative feedbacks has been unclear. We show that for smaller changes (up to three times the pre-industrial levels), feedbacks remain consistent during CO_2 increase and removal. However, at higher CO_2 levels, the removal process triggers more stabilizing radiative feedbacks that act to keep temperatures elevated, adding to this transient delay. We find this is primarily caused by a low-level cloud loss in the North Atlantic, resulting from an overshoot in the Atlantic Meridional Overturning Circulation. Our study emphasizes an additional mechanism by which the transient global mean surface temperature hysteresis can be modulated.

1. Introduction

The atmospheric CO_2 concentration has increased by approximately 50% since pre-industrial times, driving the majority of the historical global mean surface temperature rise (ΔT) since the mid-19th century (IPCC, 2021). To mitigate further warming and limit ΔT to below 2°C, some proposed socio-economic pathways (e.g., SSP1-1.9 and SSP1-2.6) include net-negative emissions, where CO_2 is artificially removed from the atmosphere at a greater rate than it is emitted (Fuss et al., 2014; Keller et al., 2018; Rogelj et al., 2018). However, studies on CO_2 removal scenarios (Boucher et al., 2012; Held et al., 2010) have shown that even after CO_2 is reduced, ΔT does not fully return to pre-industrial levels but remains elevated, indicating a transient hysteresis—a temporary offset that relaxes on longer timescales rather than a permanent, irreversible change (Eliseev et al., 2014). Beyond temperature, other components of the climate system, including precipitation (Cao et al., 2011; Wu et al., 2010), low-level clouds (Boucher et al., 2012), and ocean stratification (Boucher et al., 2012), have also been found to exhibit transient hysteresis, suggesting a broader, slowly reversible climate response during CO_2 removal.

The slow progress in emissions reductions has intensified interest in CO_2 removal, leading to initiatives like the Carbon Dioxide Removal Model Intercomparison Project (CDR-MIP, Keller et al. (2018)) to investigate the implications of various CO_2 removal strategies. CDR-MIP experiments, typically employing transient CO_2 removal scenarios (1% per year) from initial CO_2 concentrations of 2× or 4× pre-industrial levels, have revealed transient hysteresis-like responses in key climate variables. This behavior, indicating a lack of full reversibility upon returning to pre-industrial conditions, has been identified in ΔT (Qu & Huang, 2023; Song et al., 2023; S.-K. Kim et al., 2022), precipitation (Jeltsch-Thömmes et al., 2020; Song et al., 2023), the Atlantic Meridional

Overturning Circulation (AMOC) (Schwinger et al., 2022), and other large-scale dynamics metrics (S.-Y. Kim et al., 2023; Hwang et al., 2024). At leading order, this ΔT hysteresis reflects the ocean's thermal inertia: heat stored during the CO_2 increase is released only slowly during removal, so ΔT stays elevated after CO_2 returns to pre-industrial levels, without implying a true equilibrium irreversibility (Held et al., 2010). Its magnitude has been attributed to the rate of CO_2 decrease, the model's climate sensitivity, and the forcing history (Jeltsch-Thömmes et al., 2020; Qu & Huang, 2023). It has also been attributed to the forcing asymmetry between the two phases, since the increase pushes the system away from the pre-industrial state while the removal pushes it back toward pre-industrial CO_2 from a warmer, non-equilibrated state. These behaviors are consistent with the IPCC assessment that some CO_2-induced changes are only slowly reversible (IPCC, 2021), though the specific role of radiative feedbacks in this transient hysteresis has yet to be investigated.

Understanding whether radiative feedbacks are consistent during CO_2 increase and removal is crucial for evaluating the effectiveness of CO_2 removal in reversing global warming. The role of radiative feedbacks in this transient ΔT hysteresis has not yet been thoroughly investigated, primarily due to limitations in previous experimental designs, including the use of transient rather than abrupt CO_2 experiments and the lack of accurate effective radiative forcing estimates (Forster et al., 2016). Given that radiative feedbacks are nonlinear and state-dependent when CO_2 concentrations are increased and decreased from the pre-industrial state (Bloch-Johnson et al., 2021; Eisenman & Armour, 2024; Mitevski, Orbe et al., 2021; Mitevski et al., 2022, 2023), it is essential to: (a) determine whether radiative feedbacks are consistent during CO_2 removal from a future warmer state, and (b) whether this consistency depends on the magnitude of warming in that state.

2. Methods

Fully coupled Earth System Model runs. We conduct experiments using the large ensemble version of the Community Earth System Model 1 (CESM1-LE) (Kay et al., 2015). The CESM1-LE includes the Community Atmosphere Model version 5 (CAM5, with 30 vertical levels) and the Parallel Ocean Program version 2 (POP2, with 60 vertical levels), both having an approximate horizontal resolution of 1° in all model components. We perform fully coupled 150-year experiments, following the standard CMIP protocol (Eyring et al., 2016) to diagnose radiative feedbacks, where we abruptly increase CO_2 concentrations from pre-industrial levels to 2×, 3×, 4×, 5×, 6×, 7×, and 8×CO_2. At year 151, immediately following the 150-year increase simulation, we abruptly reduce CO_2 to pre-industrial levels and extend the simulations for an additional 150 years. The models used here do not include dynamic glaciers or dynamic vegetation, both of which are known to cause hysteresis behavior. Additionally, we do not consider the carbon cycle but focus on the climate system's response to changes in radiative forcing. Throughout, ΔT denotes the global mean surface temperature anomaly relative to the long-term mean of the 1×CO_2 pre-industrial control, computed over the full length of the control simulation, for both the CO_2 increase and removal experiments. Because this reference is a time-invariant climatology rather than a year-varying control time series, subtracting it is a constant offset and leaves all regression slopes reported here unchanged.

Slab ocean model runs. To isolate the ocean's contribution to any potential hysteresis behavior in CO_2 removal scenarios, we also run the slab ocean version of the CESM1-LE model, using prescribed ocean heat flux (QFLUX) from a pre-industrial control run. We perform abrupt CO_2 increase and removal experiments at 2×, 3×, 4×, 5×, and 6×CO_2 (a subset of the fully coupled forcing range), each consisting of a 60-year increase simulation followed by a 60-year removal simulation; the shorter 60-year simulations are sufficient because the slab ocean equilibrates within a few decades. We use both the fully coupled and slab ocean versions of CESM1-LE to analyze the roles of radiative feedbacks and ocean circulation changes in CO_2 removal scenarios.

Radiative feedbacks. We calculate the net feedback parameter, λ_{net}, by regressing the annual global mean net top-of-atmosphere radiation anomaly ΔN (downward positive) onto the annual global mean surface temperature anomaly ΔT, both taken with respect to the long-term mean of the 1×CO_2 pre-industrial control, over each CO_2 increase and removal simulation, for both the fully coupled (150-year) and the slab ocean (60-year) experiments. For the removal experiments the regression spans years 1–150 of the removal branch. The individual feedbacks are diagnosed using the radiative kernel method (Shell et al., 2008; Soden et al., 2008), in which each feedback is quantified from the top-of-atmosphere radiative response to a unit change in atmospheric temperature, surface temperature, water vapor (WV), or surface albedo. We compute these feedbacks using the CESM–CAM5 radiative kernels of Pendergrass et al. (2018) as implemented in the ClimKern package (Janoski, Mitevski, Kramer, Previdi, & Polvani, 2025), with the water-vapor feedback evaluated using the Zelinka method (Zelinka et al., 2020). The cloud feedback is computed as a residual: the all-sky top-of-atmosphere radiative response minus the all-sky radiative forcing and the summed non-cloud (Planck, lapse-rate, water-vapor, and surface-albedo) all-sky kernel feedbacks. Because it is computed as a residual, this term includes both the cloud feedback and any error in the kernel decomposition. Regional feedback contributions (Figure 2d) are computed as the area-weighted sum of the local feedback over the region divided by the *global* surface area, so that contributions from complementary regions sum to the global-mean feedback; the North Atlantic is defined as box 0–60°N, 80°W–10°E.

Estimated inversion strength. We compute Estimated Inversion Strength (EIS) following the method outlined by Wood and Bretherton (2006). For the CO_2 increase run, EIS is computed as the difference between the 150 years of the 1×CO_2 run and years 140–150 of the 4×CO_2 run. For the CO_2 removal scenarios, EIS is computed as the difference between years 140–150 of the 4×CO_2 increase run and years 140–150 of the 4×CO_2 removal run. We plot the EIS difference between the CO_2 increase and removal runs in Figures S1m–S1o in Supporting Information S1. The end-of-period difference maps of low cloud (Figure 3c and Figures S1g–S1i in Supporting Information S1) are referenced in the same way: the CO_2 increase fields relative to the pre-industrial control, and the CO_2 removal fields relative to the end of the corresponding CO_2 increase run.

3. Results

We first examine the global mean surface temperature and net top-of-atmosphere (TOA) radiation responses in the fully coupled abrupt CO_2 increase and removal experiments (Figure 1). The global mean surface temperature, ΔT, increases with rising CO_2 concentrations (Figure 1a). At the same time, the net TOA radiation (Figure 1c) starts with a large imbalance at the beginning of the experiment and then decreases to approximately 0.5 Wm^{-2} for the 2×CO_2 run and 2.5 Wm^{-2} for the 8×CO_2 run, remaining positive and far from equilibrium even after 150 years. During the subsequent removal phase (Figures 1b and 1d), ΔT decreases but by year 300 does not return to zero, exhibiting a similar transient ΔT hysteresis as found in the 1% yr^{-1} CO_2 increase and removal runs in the CDRMIP studies (Keller et al., 2018; S.-K. Kim et al., 2022). The magnitude of the net TOA radiation response to CO_2 removal (Figure 1d) is similar to the CO_2 increase, with a large imbalance at the beginning of the CO_2 removal (around year 151) that decreases toward the end of the run (around year 300) while remaining nonzero. The net TOA radiation imbalance at the end of the CO_2 removal runs (Figure 1d) is smaller in magnitude than at the end of the CO_2 increase runs, though in both phases the runs remain far from equilibrium after 150 years.

We next focus on the net radiative feedback parameter λ_{net} (Figure 2a), computed as the 150-year regression of net TOA radiation imbalance and global mean surface temperature (see Methods). We find that λ_{net} exhibits a non-monotonic behavior across the different CO_2 experiments during CO_2 increase (red), reaching a minimum at 4×CO_2. Previously, Mitevski et al. (2023) documented the non-monotonic behavior of λ_{net} at 4×CO_2 and attributed it to the collapse of the AMOC. This collapse induced cooling in the North Atlantic, which generated more low-level clouds. These clouds reflected shortwave radiation back to space, producing a more stabilizing (negative) feedback λ_{net}. Here we focus on the CO_2 removal experiments (blue) and find that for 2×CO_2 and 3×CO_2, the λ_{net} remains similar between CO_2 increase and removal. However, at 4×CO_2 and higher forcing, λ becomes more negative (i.e., more stabilizing) in the CO_2 removal runs compared to the CO_2 increase runs. A more negative λ_{net} means that the atmosphere is more efficient in equilibrating net TOA imbalance per degree global warming (λ_{net} is in Wm^{-2} K^{-1}). As we show below, this arises from a loss of reflective low cloud rather than from enhanced radiative cooling, so it acts to keep temperatures elevated during the removal phase.

To identify which individual feedback is contributing to the more stabilizing (negative) λ_{net} in the CO_2 removal runs, we decompose the net feedback parameter λ from Figure 2a into cloud feedbacks (Figure 2b) and non-cloud feedbacks (Figure 2c), the latter including lapse rate, Planck, WV, and albedo feedbacks. We find that the non-cloud feedbacks remain relatively similar between the CO_2 increase (red) and removal (blue) across the different CO_2 forcings, and the majority of the difference in λ between the CO_2 increase and removal runs from 4×CO_2 to 8×CO_2 can be attributed to cloud feedbacks. Specifically, from 4× to 8×CO_2, the difference in cloud feedbacks between CO_2 increase and removal accounts for the difference in λ_{net}. The global-mean feedback difference is dominated by the cloud feedback not only relative to the summed non-cloud term but also relative to each non-cloud feedback considered individually (lapse-rate, Planck, WV, and surface albedo).

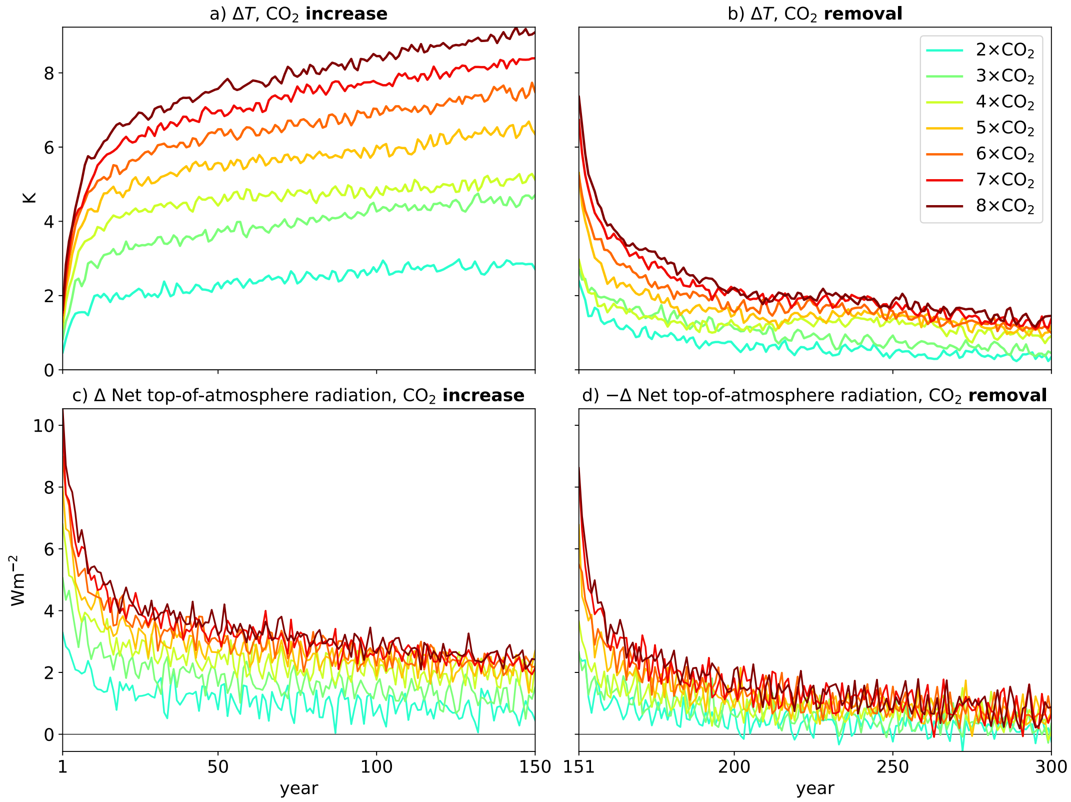

Figure 1. Global mean surface temperature response ΔT in (a) 150-year abrupt CO_2 increase experiments, where atmospheric CO_2 levels are raised from $1\times CO_2$ to 2×, 3×, 4×, 5×, 6×, 7×, and $8\times CO_2$, and (b) subsequent 150-year abrupt CO_2 removal experiments, in which CO_2 levels are reduced back to $1\times CO_2$. Also shown is the net top-of-atmosphere radiation imbalance for both CO_2 increase and removal, corresponding to panels (c) and (d), respectively. During CO_2 removal the net top-of-atmosphere radiation imbalance is negative; panel (d) therefore plots its negative ($-\Delta N$) so that it can be compared directly with the CO_2 increase imbalance in panel (c).

Focusing on the $4\times CO_2$ runs, we analyze the global mean net feedback parameter λ_{net} difference between CO_2 increase and removal (Figure 2d, left, black bar) which represents the difference between the red and blue point at $4\times CO_2$ in Figure 2a. Decomposing this difference into cloud feedbacks (blue) and non-cloud feedbacks (gray), we find that cloud feedbacks make λ_{net} more negative, confirming that cloud feedbacks primarily drive the stronger stabilizing feedbacks in the CO_2 removal runs, while non-cloud feedbacks act to reduce the overall difference. Examining regional contributions, we find that approximately 80% of this difference originates from the North Atlantic (black bars, Figure 2d, right), with cloud feedbacks (blue) again being the dominant factor. Having established that cloud feedbacks primarily drive the more stabilizing λ_{net} in CO_2 removal runs, contributing to the increased ΔT hysteresis, we next investigate the mechanisms behind the cloud feedback.

To investigate the mechanisms behind the more stabilizing feedbacks in CO_2 removal scenarios over the North Atlantic, we analyze maps showing the difference between the $4\times CO_2$ removal and the $4\times CO_2$ increase run. The difference in the net feedback parameter λ_{net} (Figure 3a) confirms that most of the stabilizing (negative) feedbacks are concentrated in the subtropical and northeastern Atlantic (blue regions). The total cloud feedback maps (Figure 3b) further reveal that the enhanced negative net feedback in the North Atlantic is primarily driven by cloud feedbacks. A more detailed decomposition into shortwave and longwave cloud components indicates that the shortwave cloud feedback is responsible for this enhanced negative feedback in the North Atlantic (not shown).

This more stabilizing cloud feedback originates in the North Atlantic surface temperature response, which is in turn driven by changes in the AMOC. During CO_2 increase, the AMOC weakens and eventually collapses at $4\times CO_2$ and higher concentrations (Figure 4a), a phenomenon previously documented in Mitevski et al. (2023), cooling the North Atlantic (Figure S1j in Supporting Information S1). In contrast, during CO_2 removal, while

10.1029/2026GL124347

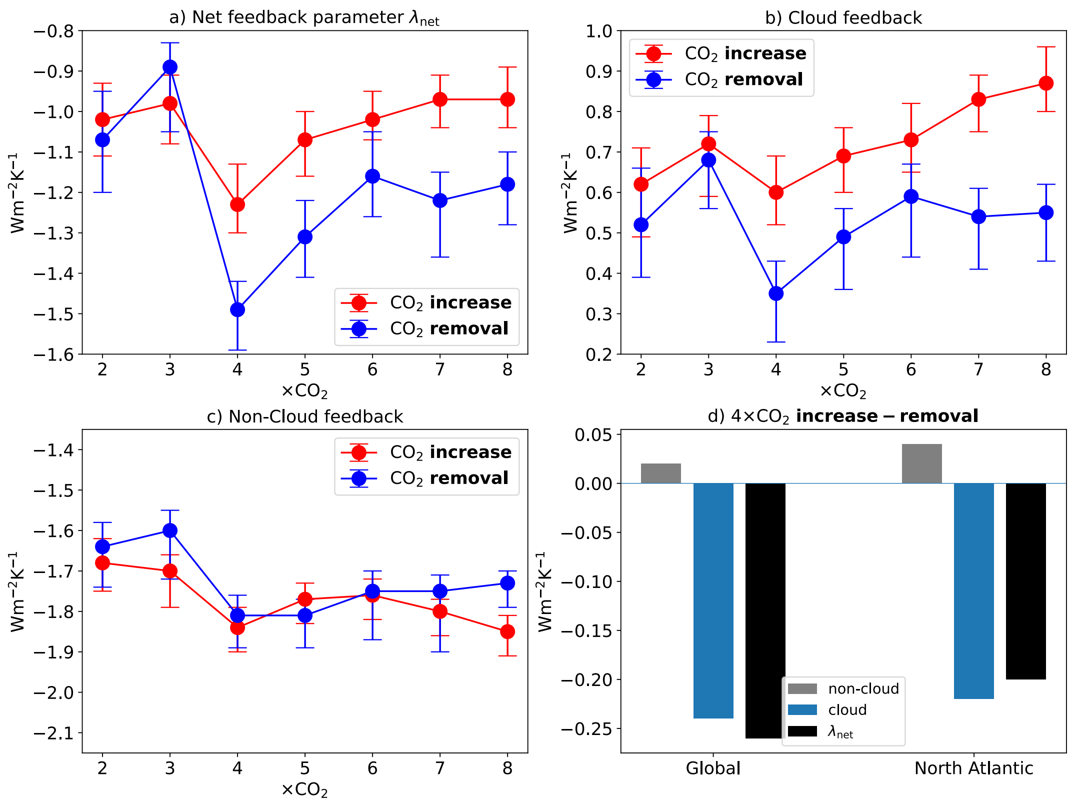

Figure 2. (a) Net radiative feedback parameter calculated from a 150-year Gregory regression (see methods) for fully coupled abrupt CO_2 increase runs (red) and CO_2 removal runs (blue). (b) Cloud feedback calculated as a residual from kernel calculations. (c) The remaining feedbacks—Planck (PL), lapse rate, water vapor, and albedo (AL)—calculated using kernels. (d) Decomposition of the difference in net feedback at 4×CO_2 between blue and red in the Global and North Atlantic regions. Regional contributions are area-weighted sums of the local feedback divided by the *global* surface area (see Methods), so the North Atlantic and non–North Atlantic contributions sum to the global value; the North Atlantic is defined as 0–60°N, 80°W–10°E. Error bars denote the 95% confidence interval of the Gregory-regression slope, estimated from a 10,000-resample bootstrap. Note that the y-range for panels (a) and (b) is the same (0.8 Wm^{-2}) to facilitate comparison.

global surface temperatures cool, the AMOC recovers and overshoots, strengthening beyond its 1×CO_2 baseline (Figure 4b); this overshoot has been attributed to salinity buildup in the subtropical North Atlantic during the CO_2 increase phase when the AMOC weakens (An et al., 2021; Jackson et al., 2014; Wu et al., 2011). Because the overshoot during removal exceeds the preceding collapse during increase at 4×CO_2 and higher, the North Atlantic warms anomalously during removal even as the global atmosphere cools, and this anomalous warming exceeds in magnitude the cooling that occurred during CO_2 increase (Figures S1j and S1k in Supporting Information S1).

These AMOC-driven surface temperature changes control the North Atlantic low-level clouds through their effect on lower-tropospheric stability, which we quantify with the estimated inversion strength (EIS; see Methods); the low-level clouds in turn influence the radiative budget by reflecting incoming shortwave radiation back to space (Figure 3c). During CO_2 increase, the North Atlantic cooling increases the EIS, enhancing low-level cloud cover and shortwave reflection and stabilizing the climate, producing the minimum in the net feedback parameter λ_{net} at 4×CO_2 (Figure 2a). During CO_2 removal, the anomalous North Atlantic warming reduces the EIS (Figures S1m–S1o in Supporting Information S1), so low-level cloud cover decreases more than it previously increased (red regions in Figure 3c), reducing shortwave reflection. Because the EIS reduction during removal exceeds the EIS increase during CO_2 increase, the low-level cloud loss likewise exceeds the earlier gain; this asymmetry amplifies the stabilizing (negative) cloud feedback, explaining why λ_{net} is more stabilizing in the CO_2 removal scenario. Across all seven forcing levels, the asymmetry in the North Atlantic cloud feedback between removal and increase scales with the asymmetry in the AMOC change (Figure S2 in Supporting Information S1), tying the feedback difference specifically to AMOC behavior rather than to ocean dynamics generally.

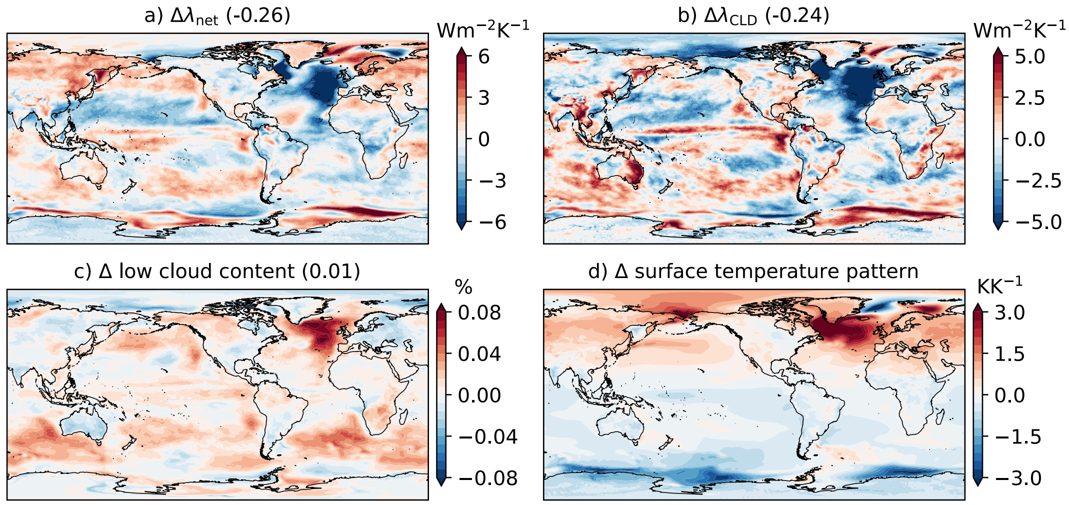

Figure 3. Maps showing the difference between CO_2 removal and CO_2 increase at 4×CO_2 for (a) the net cloud feedback parameter, (b) cloud feedbacks, (c) low cloud amount, and (d) the surface temperature pattern. Panels (a, b), and (d) are computed by regression: (a) and (b) from 150-year regressions of local top-of-atmosphere radiation against the global mean surface temperature response (Wm^{-2} K^{-1}), and (d) from the regression slope of local surface temperature against global mean surface temperature (local warming per degree of global warming, K K^{-1}); each of these panels shows the CO_2 removal pattern minus the CO_2 increase pattern. The low cloud amount (c) is an end-of-period difference over the final decade (years 140–150) of each 150-year simulation: for the CO_2 increase scenario, the response relative to the pre-industrial control; for the CO_2 removal scenario, the response relative to the end of the CO_2 increase run. Maps for the individual 4×CO_2 increase and removal runs are shown in Figure S1 in Supporting Information S1.

To further confirm that the more stable feedbacks arise from the surface temperature response in the North Atlantic, we conduct experiments using a "slab ocean" model with ocean heat transport prescribed from a pre-industrial control simulation (Figure 5). By imposing a fixed ocean heat flux, these experiments lack AMOC dynamics and therefore prevent North Atlantic cooling during CO_2 increase (when AMOC collapses) and warming during CO_2 removal (when AMOC overshoots). Hence, the slab ocean model experiments result in a simple warming and cooling response with fast equilibration timescales. The global surface temperature increases with CO_2 increase (Figure 5a) and equilibrates within the first 30 years. Similarly, the net TOA radiation imbalance (Figure 5c) approaches zero on the same timescale. At year 61, CO_2 is returned to 1×CO_2 and the surface temperature cools to a near pre-industrial state (Figure 5b) as the net TOA radiation imbalance equilibrates (Figure 5d). These experiments with the slab ocean model show negligible hysteresis in the surface temperature response (Figure 5b), while the net TOA radiation budget reaches full equilibration (Figure 5d). Furthermore, the radiative feedbacks (Figure 5e) remain nearly identical between the CO_2 increase (red) and

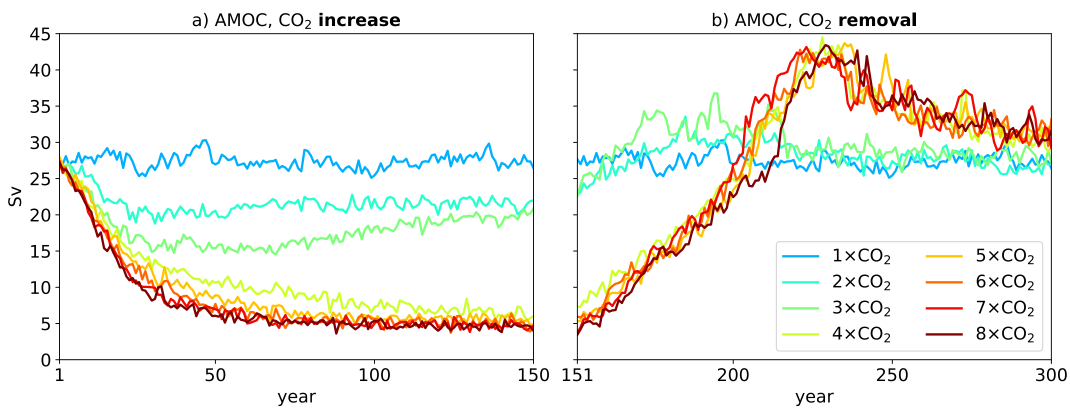

Figure 4. The Atlantic Meridional Overturning Circulation in abrupt (a) CO_2 increase and (b) CO_2 removal experiments with CESM1-LE.

10.1029/2026GL124347

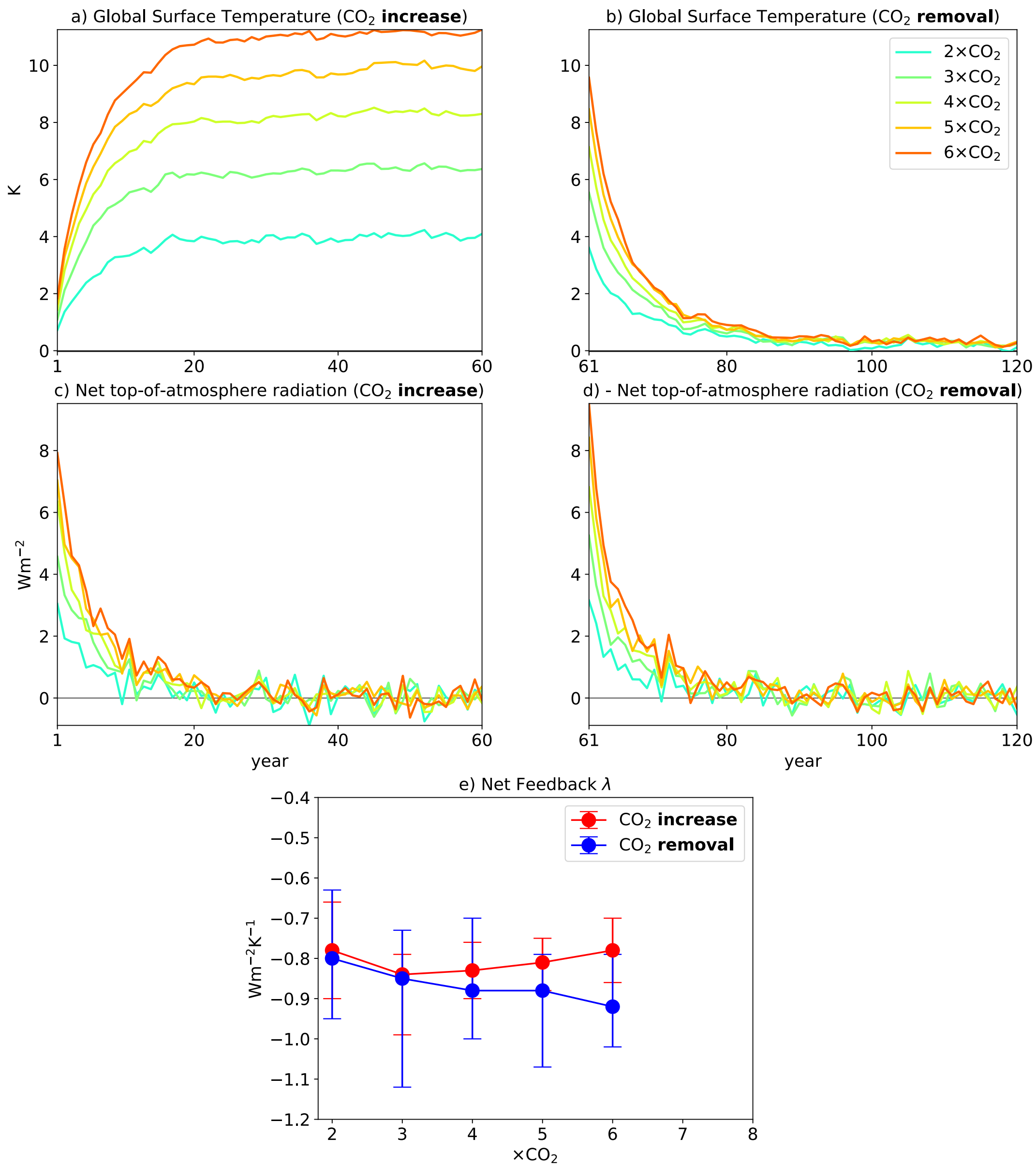

Figure 5. Results from the slab ocean model with prescribed ocean heat flux (see Methods). Global mean surface air temperature response (ΔT) for (a) CO_2 increase and (b) CO_2 removal. Net top-of-atmosphere radiation imbalance for (c) CO_2 increase and (d) CO_2 removal. (e) Net radiative feedback parameter in CO_2 increase (red) and removal (blue); error bars denote the 95% confidence interval of the Gregory-regression slope, estimated from a 10,000-resample bootstrap. During CO_2 removal the net top-of-atmosphere radiation imbalance is negative; panel (d) plots its negative ($-N$) so that it can be compared directly with the CO_2 increase imbalance in panel (c).

removal (blue) simulations; the small increase–removal difference at the highest forcing (6×CO_2) lies within the regression uncertainty. This serves as an additional confirmation that the more stabilizing (negative) feedbacks during CO_2 removal in the fully coupled ocean configuration are a direct consequence of the distinct SST responses in the North Atlantic.

4. Discussion and Outlook

We have investigated radiative feedbacks in CO_2 removal scenarios starting from 2×CO_2 up to 8×CO_2 from a future warmer state with a fully coupled Earth System Model and have found that the feedbacks are mostly the same during CO_2 increase and removal up to 3×CO_2. At 4×CO_2 and higher forcing, the feedbacks become more stabilizing (negative) when CO_2 is removed. We link the more stabilizing feedbacks to the low level cloud feedbacks in the North Atlantic associated with an AMOC overshoot during CO_2 removal. Although this transient temperature hysteresis is set mainly by ocean thermal inertia, these AMOC-mediated feedbacks provide an additional modulation that emerges only at higher forcing.

For low-to-moderate forcing (≤3×CO_2), the AMOC changes little and the North Atlantic low-cloud feedback is similar under CO_2 increase and removal, so the feedbacks are nearly identical and the transient temperature hysteresis arises mainly from ocean thermal inertia, as is widely understood. At 4×CO_2 and higher forcing, however, the AMOC collapses during increase and overshoots during removal, and because the overshoot exceeds the preceding collapse the North Atlantic loses more low-level cloud during removal than it gained during increase. Less shortwave is then reflected, which opposes the cooling and keeps the removal phase warmer than it otherwise would be; our 150-year regression diagnoses this as a more stabilizing (more negative) feedback. This AMOC-mediated cloud feedback is therefore an additional contribution to the transient hysteresis that emerges only when the AMOC declines substantially, on top of the inertial effect that dominates at lower forcing. Our slab ocean experiments, which exclude interactive AMOC dynamics and thus preclude the overshoot, show neither this cloud loss nor the feedback asymmetry, confirming that the AMOC-mediated SST response drives it. This AMOC-mediated control on North Atlantic low clouds and the associated shortwave feedback has been documented across a range of models and experimental designs (Hahn et al., 2025; He et al., 2017; Lin et al., 2019; Trossman et al., 2016; Zhang et al., 2010).

The more stabilizing feedbacks during CO_2 removal at the 4×CO_2 threshold that we found imply that the Earth system becomes more "efficient" at radiating energy back to space per degree of global mean temperature change. It is important to keep in mind that the hysteresis behavior does not imply irreversibility of the climate system on equilibration timescales as we have shown that our slab ocean simulations with ocean heat transport prescribed from a pre-industrial control simulation show no hysteresis (Figure 5), but simply a different pathway to getting back to pre-industrial CO_2 concentrations. The more stabilizing feedback therefore adds to, rather than reduces, the residual warmth that persists after CO_2 returns to pre-industrial levels. Quantifying how much it adds is left to future work, since isolating it from the dominant ocean-inertia contribution requires dedicated experiments and its magnitude should depend on the rate of CO_2 change (abrupt vs. gradual).

Future work should assess the robustness of this AMOC-cloud feedback mechanism across the CMIP6 multi-model ensemble, particularly in models with varying AMOC sensitivities and cloud parameterizations. Also, given that the magnitude of the AMOC overshoot may depend on the timescale of the forcing (Hankel, 2025), further research is needed to determine how different rates of CO_2 removal influence the more stabilizing feedbacks. As CO_2 removal strategies become a more prominent component of climate mitigation discussions, understanding the changing feedbacks during CO_2 removal is an important component for better projecting the requirement for reversing global warming.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

The radiative kernels and associated data used in this study are archived in the ClimKern kernel and data repository (Janoski, Mitevski, & Kramer, 2025), described by Janoski, Mitevski, Kramer, Previdi, and Polvani (2025). The CESM1-LE model output analyzed here was previously described in Mitevski, Orbe, et al. (2021), Mitevski et al. (2022), and Mitevski et al. (2023), and is available at Mitevski, Polvani, and Orbe (2021).

Acknowledgments
Without implying their endorsement, we thank Kyle Armour, Nadir Jeevanjee, Maria Rugenstein, Lily Hahn, Jiang Zhu, and Haozhe He for helpful discussions at various points during the course of this work. IM was supported by the HESS Postdoctoral Fellowship in the Department of Geosciences at Princeton University, GV was supported by the Simons Foundation (MPS-SRM-00006584).

References

An, S.-I., Shin, J., Yeh, S.-W., Son, S.-W., Kug, J.-S., Min, S.-K., & Kim, H.-J. (2021). Global cooling hiatus driven by an amoc overshoot in a carbon dioxide removal scenario. *Earth's Future*, *9*(7), e2021EF002165. https://doi.org/10.1029/2021EF002165

Bloch-Johnson, J., Rugenstein, M., Stolpe, M. B., Rohrschneider, T., Zheng, Y., & Gregory, J. M. (2021). Climate sensitivity increases under higher co2 levels due to feedback temperature dependence. *Geophysical Research Letters*, *48*(4), e2020GL089074. https://doi.org/10.1029/2020GL089074

Boucher, O., Halloran, P. R., Burke, E. J., Doutriaux-Boucher, M., Jones, C. D., Lowe, J., et al. (2012). Reversibility in an Earth system model in response to co2 concentration changes. *Environmental Research Letters*, *7*(2), 024013. https://doi.org/10.1088/1748-9326/7/2/024013

Cao, L., Bala, G., & Caldeira, K. (2011). Why is there a short-term increase in global precipitation in response to diminished co2 forcing? *Geophysical Research Letters*, *38*(6). https://doi.org/10.1029/2011gl046713

Eisenman, I., & Armour, K. C. (2024). The radiative feedback continuum from snowball Earth to an ice-free hothouse. *Nature Communications*, *15*(1), 6582. https://doi.org/10.1038/s41467-024-50406-w

Eliseev, A. V., Demchenko, P. F., Arzhanov, M. M., & Mokhov, I. I. (2014). Transient hysteresis of near-surface permafrost response to external forcing. *Climate Dynamics*, *42*(5–6), 1203–1215. https://doi.org/10.1007/s00382-013-1672-5

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., et al. (2016). Recommendations for diagnosing effective radiative forcing from climate models for cmip6. *Journal of Geophysical Research: Atmospheres*, *121*(20), 12–460. https://doi.org/10.1002/2016jd025320

Fuss, S., Canadell, J. G., Peters, G. P., Tavoni, M., Andrew, R. M., Ciais, P., et al. (2014). Betting on negative emissions. *Nature Climate Change*, *4*(10), 850–853. https://doi.org/10.1038/nclimate2392

Hahn, L. C., Lutsko, N. J., Eisenman, I., Luongo, M. T., & Xie, S.-P. (2025). Contribution of amoc decline to uncertainty in global warming via ocean heat uptake and climate feedbacks. *Journal of Climate*, *38*(21), 6245–6259. https://doi.org/10.1175/JCLI-D-24-0752.1

Hankel, C. (2025). The effect of co_2 ramping rate on the transient weakening of the Atlantic meridional overturning circulation. *Proceedings of the National Academy of Sciences*, *122*(1), e2411357121. https://doi.org/10.1073/pnas.2411357121

He, J., Winton, M., Vecchi, G., Jia, L., & Rugenstein, M. (2017). Transient climate sensitivity depends on base climate ocean circulation. *Journal of Climate*, *30*(4), 1493–1504. https://doi.org/10.1175/JCLI-D-16-0581.1

Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F., & Vallis, G. K. (2010). Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *Journal of Climate*, *23*(9), 2418–2427. https://doi.org/10.1175/2009JCLI3466.1

Hwang, J., Son, S.-W., Garfinkel, C. I., Woollings, T., Yoon, H., An, S.-I., et al. (2024). Asymmetric hysteresis response of mid-latitude storm tracks to co2 removal. *Nature Climate Change*, *14*(5), 496–503. https://doi.org/10.1038/s41558-024-01971-x

Ipcc, A. W. (2021). *Ar6 climate change 2021: The physical science basis*. Intergovernmental Panel on Climate Change.

Jackson, L. C., Schaller, N., Smith, R. S., Palmer, M. D., & Vellinga, M. (2014). Response of the Atlantic meridional overturning circulation to a reversal of greenhouse gas increases. *Climate Dynamics*, *42*(11), 3323–3336. https://doi.org/10.1007/s00382-013-1842-5

Janoski, T. P., Mitevski, I., & Kramer, R. J. (2025). ClimKern kernel and data repository (version 1.2.0) [dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.14743752

Janoski, T. P., Mitevski, I., Kramer, R. J., Previdi, M., & Polvani, L. M. (2025). Climkern v1.2: A new python package and kernel repository for calculating radiative feedbacks. *Geoscientific Model Development*, *18*(10), 3065–3079. https://doi.org/10.5194/gmd-18-3065-2025

Jeltsch-Thömmes, A., Stocker, T. F., & Joos, F. (2020). Hysteresis of the Earth system under positive and negative co2 emissions. *Environmental Research Letters*, *15*(12), 124026. https://doi.org/10.1088/1748-9326/abc4af

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., et al. (2015). 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*(8), 1333–1349. https://doi.org/10.1175/BAMS-D-13-00255.1

Keller, D. P., Lenton, A., Scott, V., Vaughan, N. E., Bauer, N., Ji, D., et al. (2018). The carbon dioxide removal model intercomparison project (cdrmip): Rationale and experimental protocol for cmip6. *Geoscientific Model Development*, *11*(3), 1133–1160. https://doi.org/10.5194/gmd-11-1133-2018

Kim, S.-K., Shin, J., An, S.-I., Kim, H.-J., Im, N., Xie, S.-P., et al. (2022). Widespread irreversible changes in surface temperature and precipitation in response to co2 forcing. *Nature Climate Change*, *12*(9), 834–840. https://doi.org/10.1038/s41558-022-01452-z

Kim, S.-Y., Choi, Y.-J., Son, S.-W., Grise, K. M., Staten, P. W., An, S.-I., et al. (2023). Hemispherically asymmetric hadley cell response to co_2 removal. *Science Advances*, *9*(30), eadg1801. https://doi.org/10.1126/sciadv.adg1801

Lin, Y.-J., Hwang, Y.-T., Ceppi, P., & Gregory, J. M. (2019). Uncertainty in the evolution of climate feedback traced to the strength of the Atlantic meridional overturning circulation. *Geophysical Research Letters*, *46*(21), 12331–12339. https://doi.org/10.1029/2019GL083084

Mitevski, I., Dong, Y., Polvani, L. M., Rugenstein, M., & Orbe, C. (2023). Non-monotonic feedback dependence under abrupt co2 forcing due to a north Atlantic pattern effect. *Geophysical Research Letters*, *50*(14), e2023GL103617. https://doi.org/10.1029/2023GL103617

Mitevski, I., Orbe, C., Chemke, R., Nazarenko, L., & Polvani, L. M. (2021). Non-monotonic response of the climate system to abrupt co2 forcing. *Geophysical Research Letters*, *48*(6), e2020GL090861. https://doi.org/10.1029/2020GL090861

Mitevski, I., Polvani, L. M., & Orbe, C. (2021). Abrupt and transient CO2 experiments with CESM-LE [dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.5725084

Mitevski, I., Polvani, L. M., & Orbe, C. (2022). Asymmetric warming/cooling response to co2 increase/decrease mainly due to non-logarithmic forcing, not feedbacks. *Geophysical Research Letters*, *49*(5), e2021GL097133. https://doi.org/10.1029/2021GL097133

Pendergrass, A. G., Conley, A., & Vitt, F. M. (2018). Surface and top-of-atmosphere radiative feedback kernels for CESM-CAM5. *Earth System Science Data*, *10*(1), 317–324. https://doi.org/10.5194/essd-10-317-2018

Qu, X., & Huang, G. (2023). The primary factors influencing the cooling effect of carbon dioxide removal. *npj Climate and Atmospheric Science*, *6*(1), 215. https://doi.org/10.1038/s41612-023-00547-4

Rogelj, J., Popp, A., Calvin, K. V., Luderer, G., Emmerling, J., Gernaat, D., et al. (2018). Scenarios towards limiting global mean temperature increase below 1.5°c. *Nature Climate Change*, *8*(4), 325–332. https://doi.org/10.1038/s41558-018-0091-3

Schwinger, J., Asaadi, A., Goris, N., & Lee, H. (2022). Possibility for strong northern hemisphere high-latitude cooling under negative emissions. *Nature Communications*, *13*(1), 1095. https://doi.org/10.1038/s41467-022-28573-5

Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Using the radiative kernel technique to calculate climate feedbacks in NCAR's community atmospheric model. *Journal of Climate*, *21*(10), 2269–2282. https://doi.org/10.1175/2007JCLI2044.1

Soden, B. J., Held, I. M., Colman, R., Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, *21*(14), 3504–3520. https://doi.org/10.1175/2007JCLI2110.1

Song, S.-Y., Yeh, S.-W., Allan, R. P., Xie, S.-P., An, S.-I., & Park, H.-S. (2023). Climate sensitivity controls global precipitation hysteresis in a changing co2 pathway. *npj Climate and Atmospheric Science*, *6*(1), 156. https://doi.org/10.1038/s41612-023-00484-2

Trossman, D. S., Palter, J. B., Merlis, T. M., Huang, Y., & Xia, Y. (2016). Large-scale ocean circulation-cloud interactions reduce the pace of transient climate change. *Geophysical Research Letters*, *43*(8), 3935–3943. https://doi.org/10.1002/2016GL067931

Wood, R., & Bretherton, C. S. (2006). On the relationship between stratiform low cloud cover and lower-tropospheric stability. *Journal of Climate*, *19*(24), 6425–6432. https://doi.org/10.1175/JCLI3988.1

Wu, P., Jackson, L., Pardaens, A., & Schaller, N. (2011). Extended warming of the northern high latitudes due to an overshoot of the Atlantic meridional overturning circulation. *Geophysical Research Letters*, *38*(24). https://doi.org/10.1029/2011GL049998

Wu, P., Wood, R., Ridley, J., & Lowe, J. (2010). Temporary acceleration of the hydrological cycle in response to a co2 rampdown. *Geophysical Research Letters*, *37*(12). https://doi.org/10.1029/2010gl043730

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*(1), e2019GL085782. https://doi.org/10.1029/2019GL085782

Zhang, R., Kang, S. M., & Held, I. M. (2010). Sensitivity of climate change induced by the weakening of the Atlantic meridional overturning circulation to cloud feedback. *Journal of Climate*, *23*(2), 378–389. https://doi.org/10.1175/2009JCLI3118.1