

## Response to Reviewer comments

We thank the reviewers for their useful feedback and encouraging comments which have improved the manuscript. This process has been very helpful to sharpen up the manuscript and prompt further analysis.

We have tried to address all the comments, although some of the analysis we deem to be out of the scope of this specific paper. We emphasise that there is plenty more research that could and should be done on the climate responses to the regional MCB. This paper hopes to help direct some of that research for the wider research community.

We have made edits to the manuscript following your comments. We present the reviewers' comments in black and our responses in blue. Changes made to form the revised manuscript are referred to by line number of the new manuscript (e.g. L50 refers to line 50 of the new document). The edits were made using tracked changes to help the reviewers see what changes have been implemented.

**RC1:** Anonymous Referee #1, 04 Jan 2026

### **General Comments**

This manuscript presents a systematic and carefully constructed framework for designing multi-region MCB deployments based on a large ensemble of regional patch simulations using UKESM1.0. The optimization strategy—sampling and filtering over more than one million possible combinations using multiple climate response targets—is both novel and well motivated, and the comparison with a midlatitude-only deployment provides a useful benchmark within the existing MCB literature.

With revisions that more fully address the methodological limitations and clarify their implications, I believe this work would make a strong and valuable contribution to the literature on MCB deployment design.

### **Major Comments**

1. The paper's central conclusions depend strongly on assumptions of linearity, additivity, and ERF-based scaling of coupled climate responses, which are not sufficiently stress-tested in the current analysis. While the optimized deployments are compelling as a proof-of-concept, the manuscript itself documents substantial nonlinearity and diminishing returns in ERF with

increasing injection rate. These features raise important questions about the robustness of the scaling framework used to construct the optimized solutions. Equation (2) estimates the response of a MCB combination by scaling the 50 Tg yr<sup>-1</sup> coupled responses using ERF ratios derived from fixed-SST simulations. While this approach is pragmatic, the manuscript does not justify why ERF(50 Tg yr<sup>-1</sup>) is used as the sole reference point, despite the availability of ERF estimates at 5 and 10 Tg yr<sup>-1</sup>. Although the authors note substantial variability in the coupled responses at 5 and 10 Tg yr<sup>-1</sup>, there is also appreciable variability at 50 Tg yr<sup>-1</sup>. Moreover, since most optimized combinations allocate only small fractions (5–10 Tg yr<sup>-1</sup>) to many regions, ERF(5 Tg yr<sup>-1</sup>) or ERF(10 Tg yr<sup>-1</sup>) may be more appropriate scaling references. Please consider a sensitivity test in which Eq. (2) is evaluated using ERF(5 Tg yr<sup>-1</sup>) or ERF(10 Tg yr<sup>-1</sup>) as the reference and discuss how the optimization results might change if the scaling reference were injection-rate dependent.

Equation 2 does incorporate the ERF(5/10 and 50 Tg yr<sup>-1</sup>) data within the  $ERF_i(x_i)$  variable. For example, if  $x_{NP} = 5 \text{ Tg yr}^{-1}$  then the ERF scaled contribution to response R from the NP region is calculated as  $\frac{ERF_{NP}(5 \text{ Tg yr}^{-1})}{ERF_{NP}(50 \text{ Tg yr}^{-1})} \times r_{NP}(50 \text{ Tg yr}^{-1})$ . The  $r_i(50 \text{ Tg yr}^{-1})$  was used as the reference coupled response in this manuscript, somewhat pragmatically because the signal in assessing responses is largest.

To explore the sensitivity of the optimised solution to the choice of coupled response, we repeated the analysis using the climate responses to the 10 Tg yr<sup>-1</sup> deployments as the reference (i.e., adapted Equation 2 below). The ERF (5/10 and 50 Tg yr<sup>-1</sup>) data are still used within the  $ERF_i(x_i)$  variable but  $r_i(10 \text{ Tg yr}^{-1})$  is used as the reference climate response.

$$R = \sum_{i=1}^{N=14} \frac{ERF_i(x_i)}{ERF_i(10 \text{ Tg yr}^{-1})} \times r_i(10 \text{ Tg yr}^{-1})$$

Filtering Step	Response	Target	Target range	Number of Combinations remaining after filtering step
1	$\Delta T$ (°C)	-0.973	$\pm 0.213$	60,290
2	$\Delta P$ (mm day <sup>-1</sup> )	-0.0632	$\pm 0.0139$	475
3	$\Delta SSI$ (10 <sup>6</sup> km <sup>2</sup> )	3.36	$\pm 0.92$	219
4	$\Delta SOI$ (Pa)	-1.86	$\pm 30.95$	38
5	NH $\Delta T$ (°C)	-1.301	$\pm 0.262$	18
6	SH $\Delta T$ (°C)	-0.644	$\pm 0.163$	1

Using the modified equation 2 certainly highlights the sensitivity of the optimised solution to the choice of using the 5/10/50 Tg coupled responses as the reference in our methodology. While we acknowledge the appreciable variability of the 50 Tg yr<sup>-1</sup> responses, we believe it is still preferable to use these

rather than the 5 or 10 Tg yr<sup>-1</sup> coupled responses because the MCB response signal to the larger perturbation is better characterized relative to the internal variability noise.

Quantification of errors across the simulations:

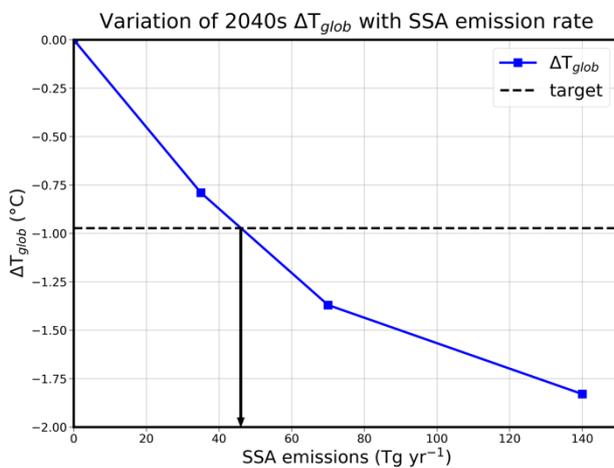
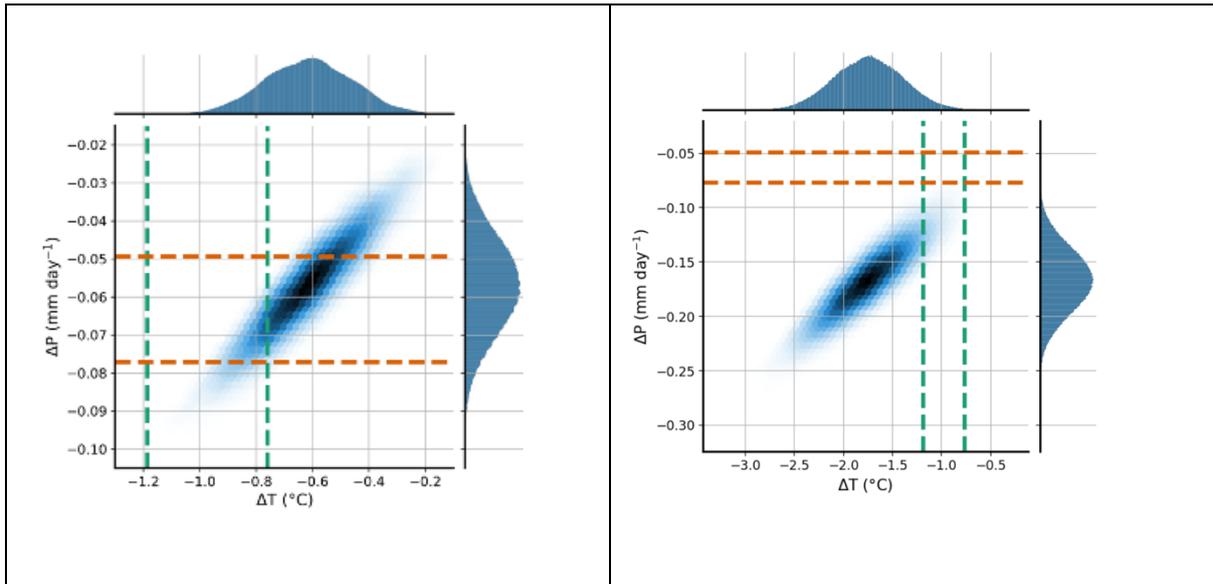
The mean and median (across the 14 MCB regions) of the standard deviation of the annual mean responses divided by the 2040s mean response.

	5 Tg	10 Tg	50 Tg	5 Tg	10 Tg	50 Tg
Variable	Mean	Mean	Mean	Median	Median	Median
Temperature	3.03	1.44	0.56	1.79	0.78	0.53
Precipitation	6.43	1.78	0.68	2.38	1.38	0.66
Arctic Sea Ice	6.45	7.02	3.7	1.73	1.48	1.06
SOI	29.42	6.82	11.16	6.85	4.07	1.93
ERF	2.23	1.12	0.31	1.91	0.78	0.27

The table shows that the 50 Tg yr<sup>-1</sup> simulations have the highest signal to noise ratio. The lower error as a fraction of the absolute response for the ERF in this table also illustrates how ERF is better constrained in our simulations than the climate responses which motivates our methodology focusing on scaling using ERF.

When we compare the distribution of temperature and precipitation responses when using the 10Tg/yr simulation ERF and response (modified Eq 2), we see that the estimated global mean cooling and drying distributions for 50Tg/yr combinations are much larger than the original estimates using the 50Tg/yr simulation ERF and responses (Fig. 8):

Original: 50 Tg yr <sup>-1</sup> reference	Modified: 10 Tg yr <sup>-1</sup> reference
$R = \sum_{i=1}^{N=14} \frac{ERF_i(x_i)}{ERF_i(50 Tg yr^{-1})} \times r_i(50 Tg yr^{-1})$	$R = \sum_{i=1}^{N=14} \frac{ERF_i(x_i)}{ERF_i(10 Tg yr^{-1})} \times r_i(10 Tg yr^{-1})$
Most efficient cooling ~ -1K for 50Tg	Most efficient cooling ~ -2.5K for 50Tg



When we compute the  $\Delta T$  from simulations with equal injections in all 14 regions at once, this plot suggests that a  $50 \text{ Tg yr}^{-1}$  deployment would produce about a 1K cooling. Thus, using the  $10 \text{ Tg yr}^{-1}$  coupled responses tends to overestimate the cooling and drying in a  $50 \text{ Tg/yr}$  simulation.

To address the comment, we have added the following text to make it clearer that the 5 and  $10 \text{ Tg yr}^{-1}$  ERF data is used and the reasoning behind this methodology (L404-410):

'Therefore, this directly uses the ERF data from the 5, 10 and  $50 \text{ Tg yr}^{-1}$  prescribed SST simulations, and interpolates the ERF for values between 10 and  $50 \text{ Tg yr}^{-1}$ . The  $50 \text{ Tg yr}^{-1}$  coupled responses themselves have high variability. However, owing to the larger rate of SSA emission they have higher signal-to-noise ratios, and the variability is less significant compared to the absolute mean coupled response than for the 5 and  $10 \text{ Tg yr}^{-1}$  simulations. Thus, we focus on using  $50 \text{ Tg yr}^{-1}$  as a pragmatic decision made in the absence of many ensemble members and longer simulations that might otherwise reduce the uncertainty in the 5 and  $10 \text{ Tg yr}^{-1}$  coupled responses.'

2. The manuscript states that an MCB combination strategy should avoid seeding only the Northern Hemisphere in order to minimize southward ITCZ shifts and associated impacts on vulnerable regions such as South Asia and the Sahel.

However, the optimization framework does not include an explicit ITCZ metric, nor does it constrain NH and SH precipitation ( $\Delta P$ ) separately as target variables. Is there a specific reason these metrics were not included? Given the strong allocation of MCB patches in the NH in the optimized deployments, it seems important to more directly incorporate hemispheric or ITCZ-related precipitation responses into the optimization targets.

We agree that an ITCZ metric would have been an interesting target to incorporate into this methodology, although it was not prioritised in our analysis.

To explore how the inclusion of ITCZ constraints into our methodology would have impacted the optimised solution, we have incorporated a NH/SH dP target and an ITCZ metric (Adam et al. (2016) - Adam, O., Bischoff, T., and Schneider, T.: Seasonal and interannual variations of the energy flux equator and ITCZ. Part I: Zonally averaged ITCZ position, *Journal of Climate*, 29, 3219–3230, <https://doi.org/10.1175/JCLI-D-15-0512.1>, 2016). This metric uses the precipitation median between 20N/20S as a definition of ITCZ position.

These do slightly redistribute the SSA in optim-14 and optim-6 as they further remove some of the 172 combinations left (to 13 and 11 respectively). The Adam et al., (2016) ITCZ constraint interestingly removes the SSA from NP but this is compensated by an increase in WNP (note these areas do overlap).

Region	Optim-14 (original)	Optim-14 (with NH/SH dP constraint)	Optim-14 (with Adam 2016 ITCZ constraint)
NO	15.93	18.08	17.27
NP	5.55	3.08	0.0
NEP	0.49	0.77	0.0
SEP	0.00	0.00	0.0
SP	5.64	6.92	9.09
WNP	3.78	3.46	8.64
NWP	0.52	0.77	0.0
SWP	0.96	0.38	0.45
NA	1.02	0.38	0.45
NWA	0.55	0.77	0.45
SEA	0.81	0.77	0.45
SA	9.91	12.31	11.36
NI	1.69	0.00	0.45
SEI	3.14	2.31	1.36

Region	Optim-6 (original)	Optim-6 (NH/SH dP constraint)	Optim-5 (Adam 2016 ITCZ constraint)
NO	18.123	19.58	18.10
NP	6.314	3.33	
NEP			

SEP			
SP	6.416	7.5	9.52
WNP	4.3	3.75	9.05
NWP			
SWP			
NA			
NWA			
SEA			
SA	11.274	13.33	11.90
NI			
SEI	3.572	2.5	1.43

Due to a change in HPC at the Met Office in the UK, running the UKESM1.0 model in the same configuration is not possible and we believe due to the small impact on optim-14/6, it is best to stick with the optimised simulations we have analysed in part 5 of our paper.

We do agree a consideration of the ITCZ is worth further stressing in our paper. We have analysed the outcomes of NH/SH dP and the ITCZ metric (Adam et al., 2016) for optim-14, optim-6 and Midlat (none of which incorporated these as targets during optimisation) and have added this to our paper.

To address the comment, we have added the following text (L596-602):

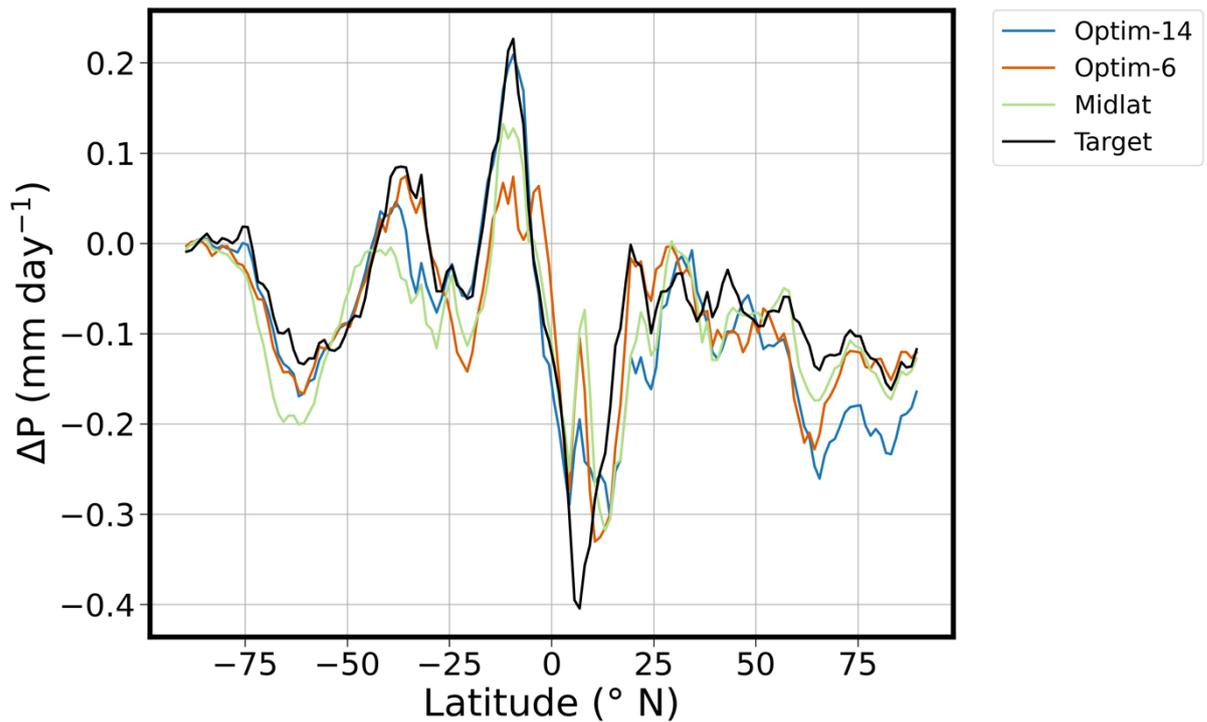
'The choice of metrics influences the optimisation; therefore, this methodology could be extended to incorporate other climate response targets that would further restrict the 172 combinations. For example, an optimised deployment could consider impacts on the position of the ITCZ, along with other climate response considerations. Although an ITCZ target was not incorporated in the optimisation in this study, impacts of the optimised MCB deployments on the ITCZ are considered in Section 5.'

And additional results and discussion of ITCZ:

Table 5 (L674)

Response	Target	Target range	Optim-14	Optim-6	Midlat
NH ΔP (mm/day)	-0.124	±0.026	-0.156 ± 0.055	-0.128 ± 0.065	-0.127 ± 0.050
SH ΔP (mm/day)	-0.002	±0.006	-0.015 ± 0.060	-0.030 ± 0.060	-0.046 ± 0.040
ITCZ (°)	0.109	±0.052	0.115 ± 0.235	0.260 ± 0.284	0.292 ± 0.185

Figure 15 (L693)



We have added the following discussion of the new table and figure at the end of section 5 (L658-672):

'While an ITCZ target was not explicitly incorporated in this optimisation, ensuring balanced  $\Delta T_{NH}$  and  $\Delta T_{SH}$  responses prevents shifts in tropical precipitation latitude to first order. We verify this in Table 5 which summarises 3 potential targets that could be used following the methodology in this paper; Northern Hemispheric mean precipitation anomaly ( $DP_{NH}$ ), Southern Hemispheric mean precipitation anomaly ( $DP_{SH}$ ), and ITCZ position, which is taken to be the latitude position of the precipitation median for precipitation between latitudes  $20^\circ$  N and  $20^\circ$  S (Adam et al., 2016). Optim-14, Optim-6 and Midlat were not optimised for these targets. For  $DP_{NH}$ , Optim-6 and Midlat match well with the target and all three simulations counteracting the increase in NH precipitation under SSP2-4.5 for this model. However, all three cause too much SH drying with respect to the  $DP_{SH}$  target. Optim-14 has the best match with the target ITCZ position, which agrees with the zonal mean precipitation shown in Figure 15. Fig. 15 shows that all three of Optim-14, Optim-6 and Midlat result in southward shifts of the ITCZ. With respect to the target, all three optimised deployments reproduce similar features, with Optim-14 producing the best match to the target peak at  $\sim 10^\circ$  S. Optim-14 and Optim-6 both result in too much drying north of  $50^\circ$  N compared to Midlat and the target, which is the result of MCB in the NO region in these two optimised strategies. Incorporating an ITCZ target in future analyses could help further constrain the MCB response on precipitation patterns,. However, including larger ensembles and longer runs

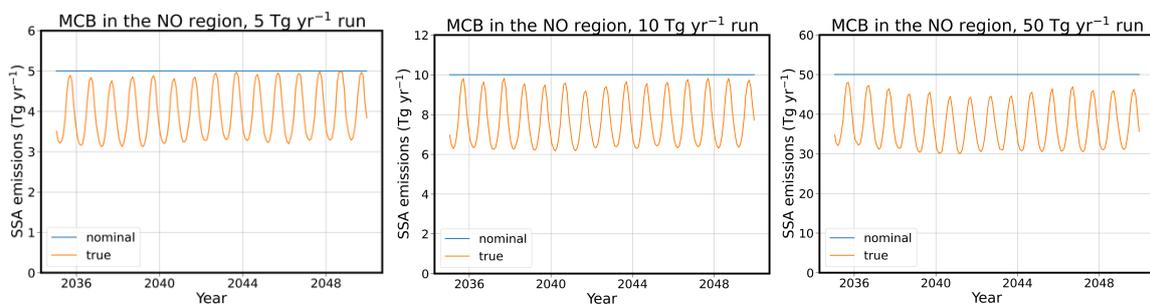
may be necessary enable a more robust analysis and optimization, particularly on regional precipitation anomalies.'

### Minor comments

1. Line 119: According to Henry et al. (2025), sea-salt emissions in September in the NO are approximately twice as large as in other months, due to the seasonal maximum in open-ocean area. Given that the NO region plays a prominent role in the optimized deployments, it would be helpful to describe the seasonal variability more clearly in the main text, rather than only mentioning it briefly.

We have done some more analysis on the emission rates for the NO region MCB simulations:

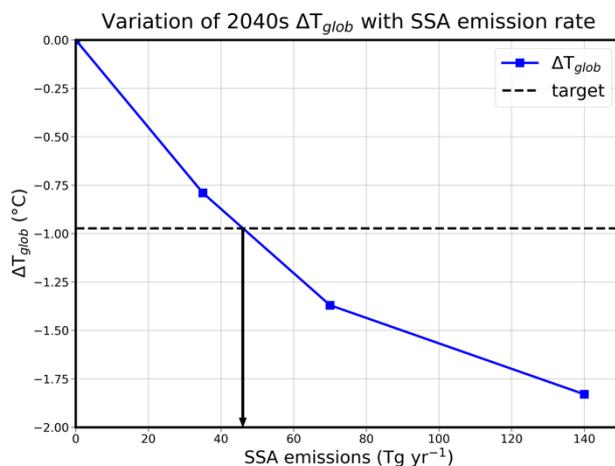
Nominal NO region MCB SSA emission rate	Average maximum monthly emission rate as % of nominal (September)	Average minimum monthly emission rate as % of nominal (March)	September : March SSA emission rate ratio (2 s.f.)
5 Tg yr <sup>-1</sup>	98%	64%	1.5
10 Tg yr <sup>-1</sup>	96%	63%	1.5
50 Tg yr <sup>-1</sup>	92%	62%	1.5



We have edited the discussion of MCB in NO region to the following (L137-142):

'In the NO region, the 5, 10 and 50 Tg yr<sup>-1</sup> SSA emissions rates are nominal. The true emission rate is scaled by the sea ice free fraction of a given grid cell, which results in a seasonal variation of SSA emission rates, along with slight annual variation due to the seasonal and inter-annual variability in sea ice coverage, as discussed in Henry et al. (2025). The SSA emission reach their annual maximum when the sea ice area is at its minimum in September, when the emission rate is approximately 1.5 times the minimum emission rate (typically in March) for a given injection rate.'

- Line 227: Including a simple figure would improve transparency and help readers better understand the rationale for selecting 50 Tg yr<sup>-1</sup> as the reference deployment magnitude.



- Added Figure 8 showing the linear interpolation between the 35 and 70 Tg/yr runs where the SSA emission is shared equally across all 14 regions. This interpolation indicates a 46 Tg/yr would be sufficient for the desired cooling. (L381)

## Technical corrections

- Line 190: It is unclear whether  $\Delta T$  and  $\Delta P$  refer to global-mean values. If so, please state this explicitly and ensure that the same convention is applied consistently throughout the manuscript.

This referred to global mean values. The use of  $\Delta T$  was not consistent throughout the manuscript so we have changed the convention for the revised manuscript as follows:

$\Delta T$  = surface air temperature anomaly.

$\Delta T_{glob}$  = global mean surface air temperature anomaly.

$\Delta T_{NH}$  = Northern Hemisphere surface air temperature anomaly.

$\Delta T_{SH}$  = Southern Hemisphere surface air temperature anomaly.

The same convention is applied to precipitation.

- Line 211: "With comparison to the ERF responses in Figure 3...". This appears to refer to Figure 2, which presents the ERF results. Please check and correct the figure numbering if necessary.

Corrected (L319)

3. Figure 8: For clarity, it would be helpful to note in the caption that the black arrows are arbitrary or conceptual, rather than representing specific quantitative values, as they could otherwise be misleading.

Text in the figure caption changed to (L364) (now figure 7):

“The shading shows the target range and the black arrows provide the magnitude for a conceptual target restoration.”

4. Table 4: The reported value of  $\Delta T$  for Optim-14 appears to be missing a minus sign. Please check and correct if necessary.

Corrected

5. Line 387: For clarity, please consider using a more explicit expression than “NO MCB” such as “MCB in the NO region” to improve readability for non-specialist readers.

Amended throughout manuscript

**RC2:** Anonymous Referee #2, 15 Jan 2026

Summary

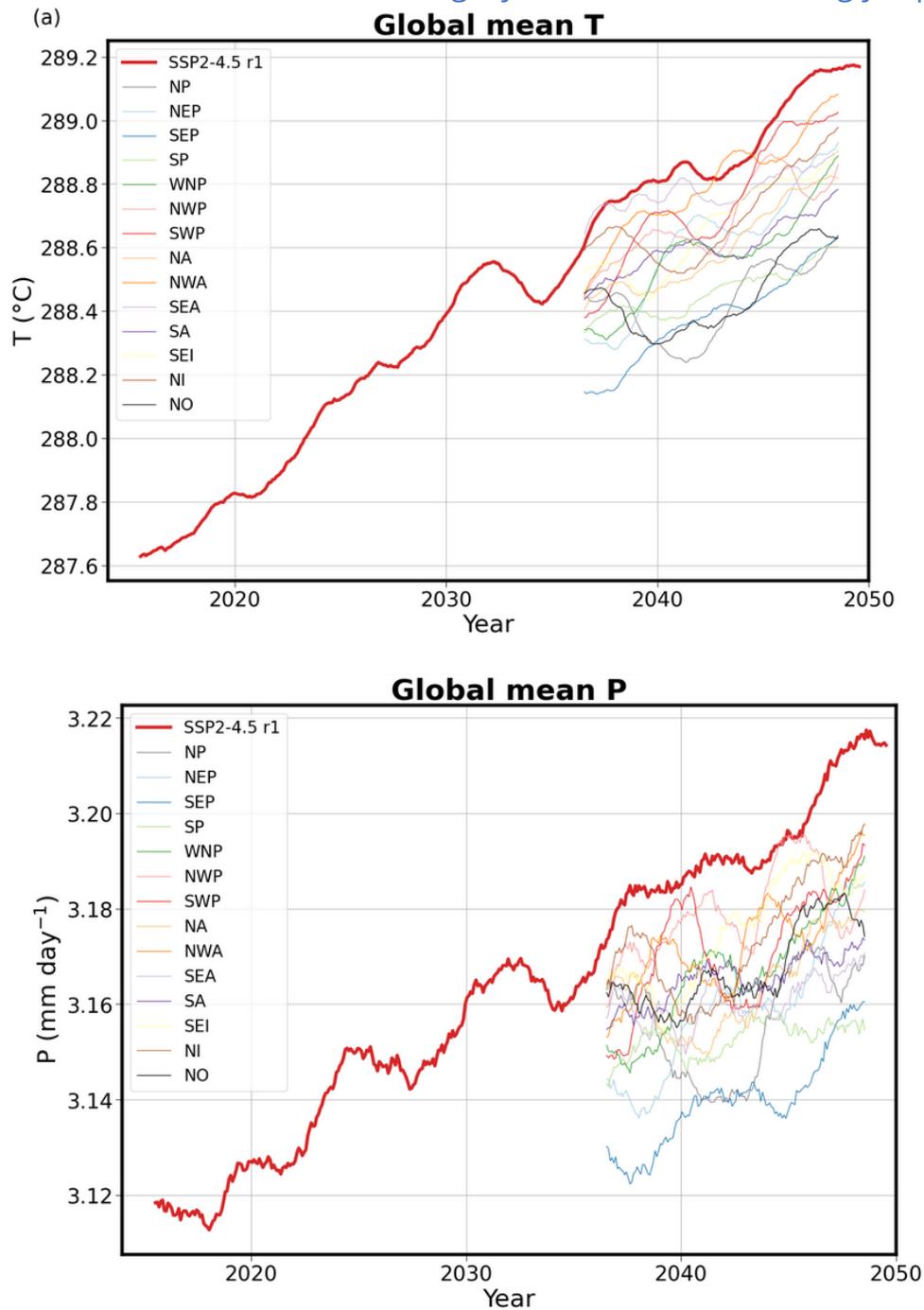
Mason et al. present a methodology for designing patchwork MCB strategies to optimize global climate targets in UKESM1.0. I believe this manuscript will be a valuable contribution to the field by systematically exploring the efficacy of different MCB regions that have been otherwise been somewhat arbitrarily chosen in the literature. I do have some questions regarding their methodology and analysis that should be addressed before publication.

Major comments

1. Length of simulations and ensemble size. Is 5 years long enough to spin up a fully coupled model and 10 years long enough to get a statistically significant signal? Time series of the global T and P response might help address whether the model is adequately spun up after 5 years. Additionally, in L211, do you mean the ERF responses in Figure 2? If so, the error bars on Fig.2 say standard error while Fig. 6 says standard deviation. If what is plotted indeed matches the captions, then it makes sense that the error in Fig. 6 would be smaller than Fig. 2. The authors should make the error calculation consistent between the plots to draw a fair comparison, as this is a key assumption that must be satisfied to warrant the methodology in the next set of simulations. Otherwise, I think that a few more ensemble members would be necessary to constrain the variability and robustness of the methodology. Given that there is high internal variability with concentrated

regional scale perturbations, I'm not convinced that 1 ensemble member of a 10-year average is sufficient to truly develop a robust optimal strategy.

Global T and P time series (rolling 3 year means for the 50 Tg yr<sup>-1</sup> patches):



Linear regression through the annual means in the 2040s, for the data plotted above, retrieves the following gradients. The standard error of the linear regression is used to produce a 95% confidence interval for the SSP245 r1 2040s gradient. Asterisks mark gradients that fall outside this confidence interval range

for the SSP245 r1 gradient, indicating the climate response is not in quasi-equilibrium with respect to the MCB perturbation.

Simulation	2040s global mean temperature gradient (K yr <sup>-1</sup> )	2040s global mean precipitation gradient (mm day <sup>-1</sup> yr <sup>-1</sup> )
SSP245 r1	0.0436 +/- 0.0224 (95% confidence interval) 0.021 -> 0.066	0.003335 +/- 0.00227 (95% confidence interval) 0.0011->0.0056
NP	0.054	0.0049
NEP	0.035	0.0033
SEP	0.040	0.0024
SP	0.027	0.00052*
WNP	0.038	0.0031
NWP	0.036	0.00056*
SWP	0.056	0.0035
NA	0.041	0.0033
NWA	0.039	0.0030
SEA	0.014*	0.0013
SA	0.029	0.0017
SEI	0.048	0.0030
NI	0.065	0.0058*
NOx	0.040	0.0020

Only the SEA falls outside the confidence interval range for temperature, while for precipitation SP, NWP and NI fall outside the confidence interval. Thus, from the 50 Tg yr<sup>-1</sup> simulation T and P timeseries we argue the model is adequately spun up for this study. We agree that longer simulations, for example 30 years, allowing more than 5 years of spin-up would reduce the risk of model spin up influencing the analysis.

Quantification of errors across the simulations (as justification for using ERF scaling). Calculated as the mean or median fractional error across the 14 MCB simulations (standard deviation of annual mean responses divided by magnitude of the 2040s mean response):

	5 Tg	10 Tg	50 Tg	5 Tg	10 Tg	50 Tg
Variable	Mean	Mean	Mean	Median	Median	Median
Temperature	3.03	1.44	0.56	1.79	0.78	0.53
Precipitation	6.43	1.78	0.68	2.38	1.38	0.66
Arctic Sea Ice	6.45	7.02	3.7	1.73	1.48	1.06

SOI	29.42	6.82	11.16	6.85	4.07	1.93
ERF	2.23	1.12	0.31	1.91	0.78	0.27

This table illustrates how ERF is better constrained, with lower fractional error and thus higher signal-to-noise ratio, in our simulations than the climate responses and motivated our methodology.

Ideally, we would have liked to do this analysis with several ensemble members. Given the large array of simulations used in this study (pairs of coupled and prescribed SST for 14 regions at 3 injection rates) simply increasing the ensemble size by 1 would require too many additional ESM simulations. We agree it's important to highlight the limitations of the set of simulations we use.

We have adjusted the error bars and figure caption in Fig. 2 (L207) to be consistent with those of Fig. 6. The standard deviation for the ERF in general remains significantly smaller when compared to the coupled responses.

We have corrected the reference to Figure 2 you noticed. (L319)

We have added the following discussion of the robustness of our methodology to the conclusion (L728-734)

'Moreover, this study relies on a single 15-year coupled simulations for each region and given injection rate, which is a significant limitation of this analysis. Longer simulations and/or larger ensembles would help to improve this analysis by allowing better uncertainty quantification and enabling the development of more robust optimised strategies.'

2. Linear additivity. While I can squint and convince myself that Fig. 3o and 3p look vaguely similar latitudinally, I find it much less convincing that there is agreement in the spatial patterns in Fig. 4o and 4p. While I do see traces of equatorial drying in the Atlantic, Indian, and West Pacific in both, the East Pacific drying in Fig. 4p is missing from 4o and almost all the moistening in the sum of regions is absent from the simultaneous simulation. The claim in L164-166 needs further justification given that the core methodology of this paper requires the linear additivity assumption to be satisfied. I agree with the statement in L236-237 that the climate responses are variable for the lower emission rates, but I am confused why scaling the response from the 50 Tg/y deployment would be a more robust proxy for smaller regional perturbations than 5 and 10 Tg/y simulations you've already run. Does this method "overcome the variability" in a physically realistic way? Would we expect the spatial pattern of responses to small perturbations to necessarily be scaled down versions of the response to larger perturbations (especially for hydrological variables)?

Whether this methodology overcomes the variability in a physically realistic way is very interesting. With the exception of SOI, our methodology targets bulk properties that span very large areas. While the spatial patterns of the responses may vary (e.g. more spatially extensive drying over the Atlantic in response to increasing SSA injection in SEA) this methodology assumes that their contribution to these bulk properties can be estimated by this ERF scaling.

Figure 12 (revised manuscript) shows some improvements to the spatial correlation scores with respect to Midlat, so our methodology does improve the estimation of more optimised MCB strategies. However, our methodology prioritises targeting large scale climate metrics. Alternative optimisation methodologies that target the minimisation of regional differences from a target period, which has been considered for stratospheric aerosol injection (Brody et al., 2025), would be a worthwhile pursuit of future research.

We have added the following text to the discussion section (L724-7):

'It is likely that the optimal emission distribution will depend on the climate targets. Incorporating spatial correlation into the optimization methodology by emphasising the minimisation of regional scale differences from a target period, rather than just targeting large scale climate metrics, could reduce some of the residual regional warming or precipitation changes that remain in Optim-14 and Optim-6 (Brody et al., 2025).'

And added Brody, E., Zhang, Y., MacMartin, D. G., Visoni, D., Kravitz, B., and Bednarz, E. M.: Using optimization tools to explore stratospheric aerosol injection strategies, *Earth System Dynamics*, 16, 1325–1341, <https://doi.org/10.5194/esd-16-1325-2025>, 2025. as a reference, which applies this alternative methodology to Stratospheric Aerosol Injection optimization.

3. SOI climate response target. Targeting a mean value for T, P and SSI makes sense to me because there are clear increasing (T and P) and decreasing (SSI) trends for each of those variables (L269-270). It's less clear that this target makes sense for SOI which is an oscillation. Is there a robust trend in SOI between 2015-2050 under global warming? Would a metric that captures changes magnitude and frequency separately be more appropriate for an oscillatory variable? For example, I could imagine a combination of patches that achieves the baseline mean SOI by decreasing the magnitude and frequency of El Niños in favor of La Niña conditions, which was described earlier in the paper as an unfavorable outcome. In other words, I wonder if this metric might measure a "successful" outcome for the wrong reasons. In Table 4, the delta SOI value for Optim-14 is wildly different from the Optim-6 and Midlat experiments and the farthest away from the target. Any speculation as to why this occurs? Could this be avoided by carefully selecting particular combinations from the 172 combinations rather than the mean to obtain

the distribution? This is the only target that is missed with substantial error (error>mean) compared to the other targets but is not mentioned in the text. Failure modes of an “optimal” strategy are important to disclose and discuss. They could even give insight as to how to improve the methodology (either the formulation of the metric or selection of patches).

Given the limited length of our simulations we believe attempts to analyse the magnitude and frequency separately are not feasible.

We incorporated our SOI target because recent studies have highlighted that tropical east Pacific forcing can cause substantial shifts in ENSO mean state (as seen in the large impact on SOI from SEP deployment). Thus, we include the metric to try to prevent any optimised combination from excessively affecting the ENSO mean.

Although not shown in Table 4, the SOI for Optim-14 is within the target range highlighted in Table 2 owing to the target range of  $\pm 31$  Pa. Thus, the SOI has a large range relative to the target. As a result we are less concerned about the magnitude of the differences between optim-14/6 and Midlat, as they are small relative to the variability. We do not interpret the anomalies as a missed target as they are all consistent with the target within the uncertainty. We have added a target range column to Table 4. (L646)

Response	Target	Optim-14	Optim-6	Midlat
$\Delta$ SOI (Pa)	$-1.86 \pm 30.95$	$19.7 \pm 121.4$	$-3.5 \pm 192.7$	$-17.1 \pm 152.6$

We have now looked at the Nino 3.4 index as an alternative measure of ENSO mean state for our optimised combinations. We compare the mean SST in the 5N-5S, 170W-120W region of our baseline against the optimised combinations.

(Barnston et al. 1997) Nino 3.4 index

<https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni> - “The Niño 3.4 index typically uses a 5-month running mean, and El Niño or La Niña events are defined when the Niño 3.4 SSTs exceed +/- 0.4C for a period of six months or more. “

Simulation	Mean SST in 5N-5S, 170W-120W region	1 STD of annual means
SSP245 2014-33	300.16	$\pm 0.85$
Optim-14 2040s	300.11	$\pm 0.67$
Optim-6 2040s	300.37	$\pm 0.64$
Midlat 2040s	300.20	$\pm 0.71$

We find this also indicates the optimised combinations do not show significant impacts on the ENSO mean state. However, we agree that the SOI and SST analysis does not fully consider the magnitude and frequency of ENSO events and that it would be advisable to include these into ENSO optimised MCB in future studies.

We have amended the discussion of the choice of SOI target (L283-5):  
'SOI is an oscillatory variable, so targets quantifying impacts on the magnitude and frequency of ENSO would be preferable for MCB optimisation. However, given our coupled simulation length of 15 years this analysis is not feasible. Instead, a mean  $\Delta$ SOI target is included to prevent significant shifts to the SOI mean state and to avoid selecting MCB strategies that lead to large La Nina-like responses (Haywood et al., 2023; Chen et al., 2025)'

And added to the discussion section (L730-3):  
'Longer simulations would allow for more sophisticated analyses of the impacts of MCB on ENSO, and alternative targets that consider the magnitude and frequency of El Niño and La Niña events would be interesting to incorporate into future optimisation analyses, along with ITCZ targets.'

#### Minor comments

1. L44-45: Should add reference to Xing et al. (2025) that explicitly analyzes the ENSO response to subtropical MCB (the authors find suppression of ENSO variability and a mean-state La Niña tendency).

#### References

Xing, C., Stevenson, S., Fasullo, J., Harrison, C., Chen, C., Wan, J., et al. (2025). Subtropical marine cloud brightening suppresses the El Niño–Southern Oscillation. *Earth's Future*, 13, e2025EF006522. <https://doi.org/10.1029/2025EF006522>

Thank you for the reference, we have added this (L53-5):  
'Additionally, Xing et al. (2025) found MCB in the subtropical eastern Pacific results in a mean-state La Niña tendency and suppresses ENSO variability, reducing ENSO amplitude by approximately 61%.'

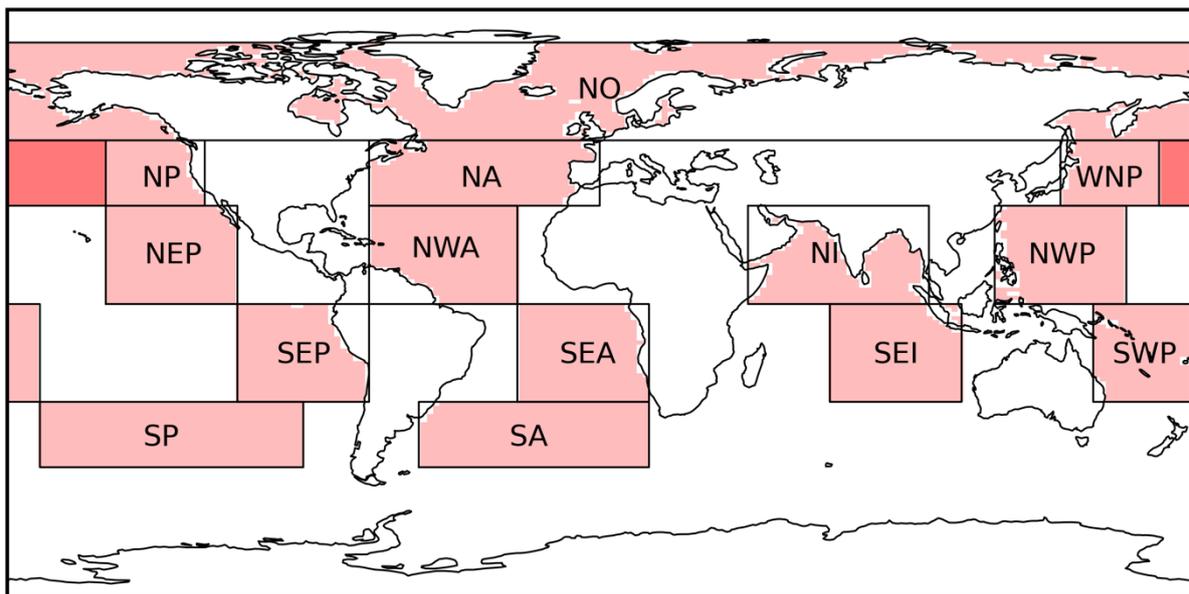
2. L69: Maybe instead of "holistically" you could say "simultaneously"? Holistically is a bit vague since you arbitrarily chose a few large-scale climate metrics to analyze.

Done (L84)

- L106-107: How do you define “open ocean”? Do you use a land/ocean mask? Is there further open ocean vs. coastal grid cell delineation? What is the % ocean area for the boxes (add column to Table 1)? Perhaps you could update Figure 1 to show the regions where SSA is injected (shading?) rather than boxes. I also think there is an extra box drawn between NP and WNP without a label (also shaded in the Optim-14 and Optim-6 experiments but doesn't appear to be one of the 14 regions?).

In this study open ocean is defined as grid cells with a land fraction = 0.0. We use a land fraction mask where only grid cells with this land fraction of 0.0 are used for SSA emission in a given region. Emissions for the NO are scaled by sea ice fraction (see response to reviewer 1's minor comment 1).

We have added a column to table 1 with the open ocean area of a given grid box, and updated figure 1.



The extra box is in fact the overlap of the WNP and NP boxes. To make this clearer, we have added the following text to the figure caption of Fig. 1 (L132): 'The red shading shows the open ocean grid points of a given box, showing where SSA are emitted. Note that the WNP and NP regions overlap, illustrated by the darker shading, as outlined in Table 1.'

- L119-120: Over what period are the 10-year transient AMIP simulations and what are the fixed SSTs?

The AMIP simulations are over the period 1979-1989.

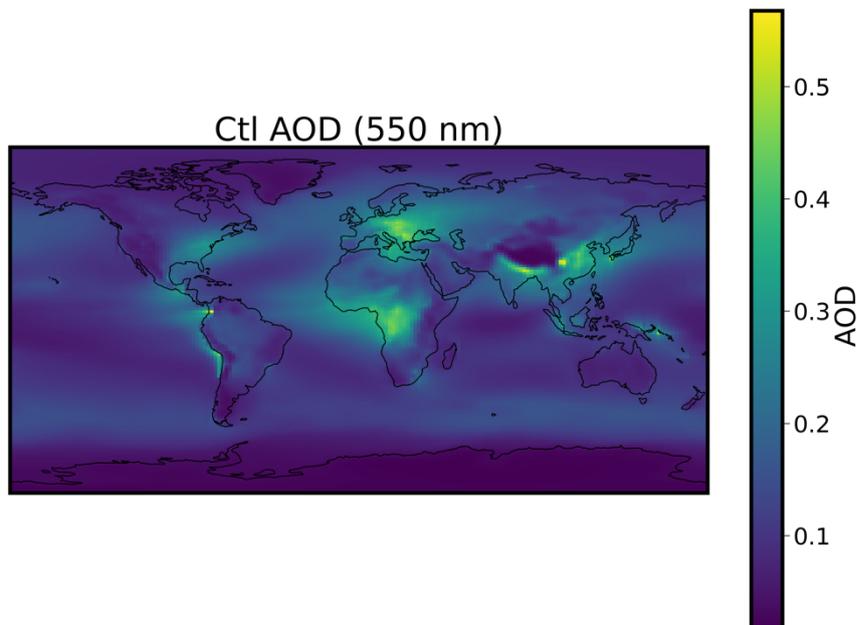
'Fixed SST' here means the sea surface temperatures are prescribed. There is no coupling (between the atmosphere and ocean) in these simulations for ERF diagnosis. The SSTs are for this 1979-1989 period.

We have changed 'fixed SST' to 'prescribed SST' to make this clearer and added 'from 1979 to 1989' to the description of the AMIP simulations (L143).

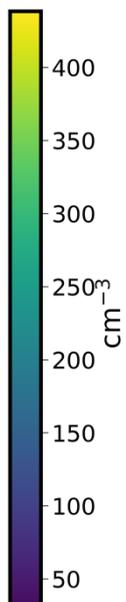
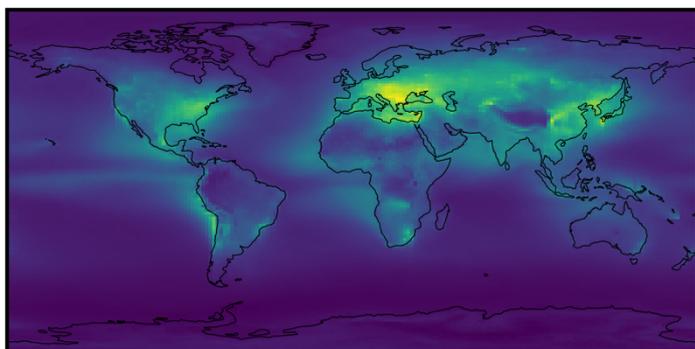
5. L140: What do you mean by pristine stratocumulus? The NEP is often considered a major stratocumulus cloud deck and the third region in early MCB studies, yet it does not have as strong of radiative forcing. It also appears that DRF dominates these stratocumulus regions, which is counter intuitive given the elevated low cloud fraction, while the SP and SA have the largest CRF response. Perhaps some maps of cloud fraction, cloud liquid water path and other related cloud properties would help.

This was an erroneous use of the word 'pristine'. We meant to refer to the relatively low CDNC of these clouds that make them susceptible to MCB (Jones et al., 2009), but have removed the word pristine from the manuscript.

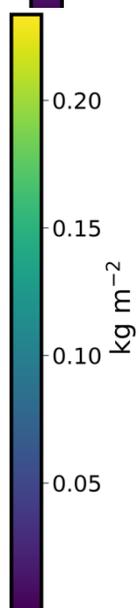
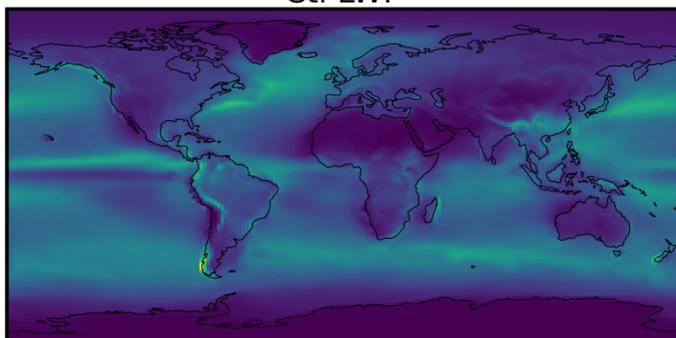
To see the background conditions of the AMIP simulations we have plotted AOD, CDNC, LWP and effective radius for the control simulation:

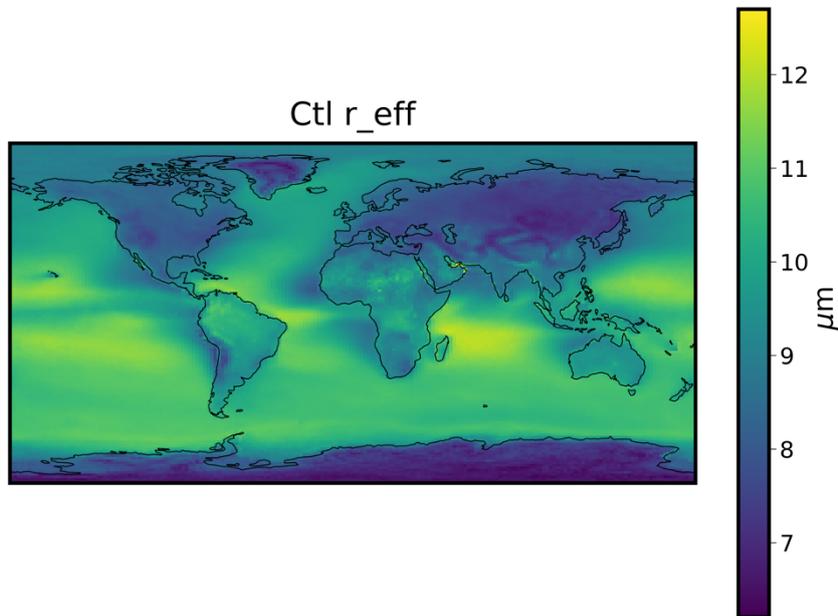


Ctl CDNC



Ctl LWP





Changes in SW cloud forcing, cloud cover and LWP for 25 Tg yr<sup>-1</sup> SSA in each of the NEP, SEP and SEA are shown in Figure 5 of Rasch et al. (2024). (<https://gmd.copernicus.org/articles/17/7963/2024/>). For UKESM1 the increase in cloud cover for SEA and SEP is more significant than for the NEP box. Hirasawa et al., 2026 compared the SW cloud ERF across different regions and found that NEP tends to have less low cloud fraction relative to SEP and SEA, which partially accounts for the weaker MCB forcing in that region.

We have added to the discussion (L180):

'The SEA and SEP give the most negative ERF for the 50 Tg yr<sup>-1</sup> injection rate, whereas the NEP ERF for this injection rate is approximately a third weaker. Previous studies found that the NEP region has both less climatological low cloud cover and smaller increases in cloud cover relative to SEP and SEA in UKESM1, which causes the area-averaged ERF susceptibility (Rasch et al., 2024, Hirasawa et al., 2026).'

6. L152: I would make it clear here that you are transitioning away from the AMIP runs to the coupled simulations for the rest of the section/manuscript.

We have split section 3 into section 3.1 "AMIP simulations" and 3.2 "Coupled simulations" to make this transition clearer.

7. L161: Are the +/- values one standard deviation? How were these errors calculated?

Yes. For the 700 Tg/yr simulation the value is the standard deviation of the ten annual mean differences that comprise the 2040s mean response. For the sum of the 14 x 50 Tg/yr responses the variability is combined in quadrature: The variances of the ten annual mean differences for each of the 14 regions are summed and divided by n;  $SD = \sqrt{\frac{\sum Var(X)}{N}}$

8. L181-189: Why does NH MCB lead to an interhemispheric temperature imbalance but SH MCB does not? Looking at Fig. 5b, I would argue that SH MCB does influence the ITCZ position (otherwise it would be a bunch of flat lines), but there is no clear trend as in the NH MCB cases. In fact, the scale is larger on the y-axis in Fig. 5b than 5a, which suggests that the magnitude of changes (at least for SEP) are large just not coherent.

We have adjusted the wording to make it clearer that SH does influence ITCZ but no clear trend (L269):

'For the southern hemisphere (SH) deployments, we also see changes in ITCZ position but the trend is not as clear as for the NH deployments.'

9. L202-203: It seems important to note that most of the regions have an insignificant SSI response, even the 50 Tg/yr experiments (Fig. 6c). Only SA, WNP, and NO look like they have error bars that don't cross 0, which is especially interesting since SA is one of the farthest seeding regions from the Arctic. How do you explain the mechanism for that result? Is the SA patch triggering a teleconnection in the Arctic?

We thank the reviewer for their comment. We believe examining this teleconnection and the mechanism behind it is beyond the scope of this current study.

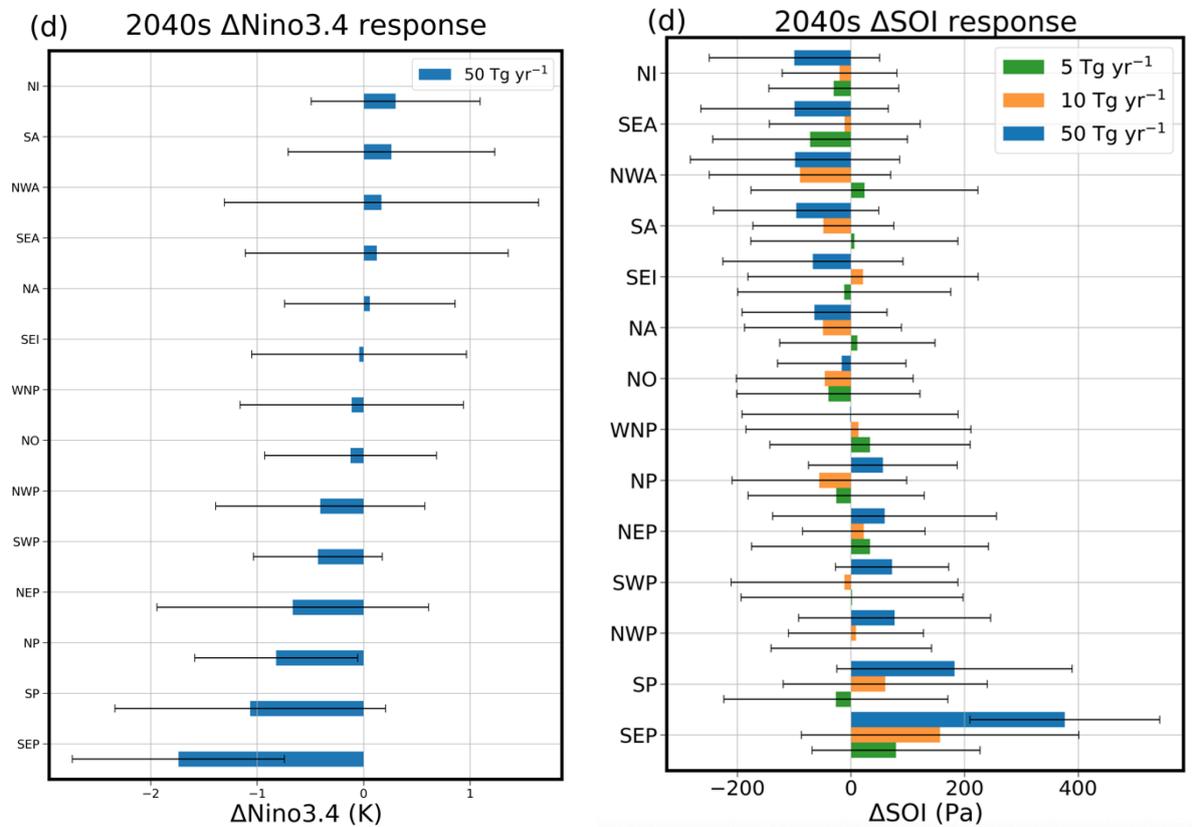
We have added a line to note the insignificant SSI responses of most of the patches (L291):

'Fig. 6(c) shows that most of the regions have an insignificant SSI response'

10. L204-205: Similar story with the SOI response—it appears that not only is the SEP the most sensitive, but it's also the only statistically significant bar in Fig. 6d. Is this result sensitive to your choice in metric to quantify ENSO? For example, perhaps you could decompose the ENSO response into an oceanic (e.g. thermocline slope) and atmospheric (e.g., Walker Cell Index) component to get a clearer picture of the ENSO response?

Given the limited length of our simulations and single ensemble member assessing the mechanisms driving the ENSO response is out of the scope of this study. We have looked at the Nino 3.4 as an alternative metric to quantify ENSO mean state (as discussed in the response to major comment 3). We present

below a figure comparing the SOI metric we use with this alternative ENSO metric. It also shows that the SEP pushes the mean state the most (also with a La Niña like tendency). The NP response is significant for this alternative metric, but there appears to be broad agreement with the findings of using SOI.



We have added lines noting that the SEP is the only statistically significant bar (L291-6):

'MCB in the SEP region exhibited the largest impact on SOI, as shown in Fig. 6(d), with the only statistically significant SOI response of all the patch simulations for SEP at 50 Tg yr<sup>-1</sup> SSA emission rate. This suggests that ENSO is most sensitive to seeding in the SEP region, with a positive SOI response of 377 Pa for a 50 Tg yr<sup>-1</sup> injection, making a La Niña state more likely. We note that studies using other fully coupled models have highlighted the strong La Niña response to aerosol radiative perturbations over the SEP region (Fasullo et al., 2024).'

11. L224-227: It's not clear from this text what delta T you are interpolating to get to 50 Tg/y. Maybe a figure would help visualize the interpolation. Referencing a future section (4.2) seems backwards. You should just describe what you want to say here first.

Added Figure 8 showing the linear interpolation between the 35 and 70 Tg/yr runs where the SSA emission is shared equally across all 14 regions. This interpolation indicates a 46 Tg/yr would be sufficient for the desired cooling. We have also moved the discussion of targets forward so as to avoid referencing a future section. The climate response targets are now introduced in section 4.1. (L332)

12. L229-231: Why 10 x 5 Tg/y shares into some combination of the 14 regions? Is this a computational constraint or other scientific rationale?

This choice balanced speed of processing with the utility of multiples of 5 Tg allowing the base set of simulations to be used directly and reducing interpolation.

Have added text (L388):

'This choice balances the computational constraint of handling a large number of combinations. Using multiples of 5 Tg yr<sup>-1</sup> shares allows the responses from the 5, 10 and 50 Tg yr<sup>-1</sup> base simulations to be used directly, with interpolation only necessary for values between 10 and 50 Tg yr<sup>-1</sup>.'

13. Fig 7: should specify that the delta T and P are global mean over the 2040s?

Figure caption amended (L426) (now Figure 9)

14. L273: Echoing the other reviewer, should NH/SH delta P should be climate response targets to capture changes in ITCZ shifts.

We agree that an ITCZ metric would be an interesting target to incorporate into this methodology, although it was not prioritised in our analysis. We refer the reviewer to the earlier response, to reviewer 1's major comment 2.

15. L335: How are you calculating the correlation score? Are you comparing the 2040s mean MCB value vs 2014-2033 target (2 values per grid cell)? Or are there multiple ensemble members for these simulations?

We use area-weighted spatial Pearson correlation scores. Yes, the correlation scores compare the 2040s mean MCB value vs 2014-2033 target, 2 values per grid cell, weighted by grid cell area to contribute to the overall correlation. There is a single ensemble member for these simulations.

16. L390-395: Beyond other ESMs, more ensemble members and longer runs seem like necessary next steps for this work to interrogate uncertainty within the model itself. It's also possible that more sophisticated optimization algorithms including AI could be useful for sampling different combinations of patches rather than this statistical approach.

We have added (L729):

'Longer simulations and/or larger ensembles would help to improve this analysis by allowing better uncertainty quantification and enabling the development of more robust optimised strategies.'

We recognize there are important opportunities to examine different, more sophisticated optimization algorithms and in the future one of the authors plans to use a neural network to sample optimised combinations from the base set of simulations.