Atmospheric
Chemistry
and Physics

Surface temperature effects of recent reductions in shipping SO$_2$ emissions are within internal variability

Duncan Watson-Parris[1], **Laura J. Wilcox**[2], **Camilla W. Stjern**[3], **Robert J. Allen**[4], **Geeta Persad**[5], **Massimo A. Bollasina**[6], **Annica M. L. Ekman**[7], **Carley E. Iles**[3], **Manoj Joshi**[8], **Marianne T. Lund**[3], **Daniel McCoy**[9], **Daniel M. Westervelt**[10], **Andrew I. L. Williams**[11], **and Bjørn H. Samset**[3]

[1]Scripps Institution of Oceanography and Halıcıoğlu Data Science Institute,
University of California San Diego, La Jolla, CA, USA
[2]National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK
[3]CICERO Center for International Climate Research, Oslo, Norway
[4]Department of Earth and Planetary Sciences, University of California, Riverside, CA, USA
[5]Jackson School of Geosciences, University of Texas at Austin, Austin, TX, USA
[6]School of GeoSciences, University of Edinburgh, Edinburgh, UK
[7]Department of Meteorology and Bolin Centre for Climate Research,
Stockholm University, Stockholm, Sweden
[8]Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, UK
[9]Department of Atmospheric Science, University of Wyoming, Laramie, WY, USA
[10]Lamont-Doherty Earth Observatory, Columbia Climate School, New York, NY, USA
[11]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA

Correspondence: Duncan Watson-Parris (dwatsonparris@ucsd.edu)

Abstract. In 2020, the International Maritime Organization (IMO) implemented strict new regulations on the emissions of sulfate aerosol from the world's shipping fleet. This can be expected to lead to a reduction in aerosol-driven cooling, unmasking a portion of greenhouse gas warming. The magnitude of the effect is uncertain, however, due to the large remaining uncertainties in the climate response to aerosols. Here, we investigate this question using an 18-member ensemble of fully coupled climate simulations evenly sampling key modes of climate variability with the NCAR model, the Community Earth System Model version 2 (CESM2). We show that, while there is a clear physical response of the climate system to the IMO regulations, including a surface temperature increase, we do not find global mean temperature influence that is significantly different from zero. The 20-year average global mean warming for 2020–2040 is $+0.03$ °C, with a 5 %–95 % confidence range of $[-0.09, 0.19]$, reflecting the weakness of the perturbation relative to internal variability. We do, however, find a robust, non-zero regional temperature response in part of the North Atlantic. We also find that the maximum annual mean and ensemble mean warming occurs around 1 decade after the perturbation in 2029, which means that the IMO regulations have likely had very limited influence on observed global warming to date. We further discuss our results in light of other, recent publications that have reached different conclusions. Overall, while the IMO regulations may contribute up to 0.16 °C $[-0.17, 0.52]$ to the global mean surface temperature in individual years during this decade, consistent with some early studies, such a response is unlikely to have been discernible above internal variability by the end of 2023 and is in fact consistent with zero throughout the 2020–2040 period.

1 Introduction

Anthropogenic aerosols play a complex and dual role in Earth's climate system. On the one hand, they contribute to atmospheric pollution, adversely affecting air quality and public health. On the other hand, they mostly exert a net cooling effect on the climate by increasing the albedo, or reflectivity, of the atmosphere, thereby reducing the amount of solar radiation that reaches the Earth's surface (Bellouin et al., 2020). Sulfate aerosols, for instance, scatter sunlight directly and enhance cloud brightness by increasing the number of water droplets in clouds, further reflecting sunlight away from the Earth (e.g. Albrecht, 1989; Twomey, 1974). Anthropogenic aerosol emissions thereby currently induce a global average cooling which (partially) offsets greenhouse-gas-driven warming by around 0.5 [0.22–0.96] °C compared to pre-industrial times (Forster et al., 2021). The magnitude of aerosol cooling is a key uncertainty in climate science and hinders our ability to accurately predict the magnitude and timing of future warming (Watson-Parris and Smith, 2022).

Environmental and health concerns associated with anthropogenic aerosols have led to international efforts to reduce their emission, with potentially significant consequences for the climate (Wall et al., 2022). In January 2020, the International Maritime Organization (IMO) took a significant step in this direction by implementing stringent regulations to curb sulfur dioxide (SO$_2$; a precursor for sulfate aerosol) emissions from the global shipping fleet (IMO 2019), which at the time contributed around 14 % of all anthropogenic sources of these pollutants (Christensen et al., 2022). Given the substantial share of global trade transported by sea, and the corresponding volume of emissions from shipping, the impact of these regulations on the global climate system was anticipated to be notable and to provide a useful experiment to quantify broader aerosol impacts (Christensen et al., 2022). Similar regulatory efforts in other sectors, and national efforts in major industrial nations such as China (Li et al., 2017; Samset et al., 2019; van der A et al., 2017), also aim to improve air quality by reducing emissions of SO$_2$ and other aerosol precursors or species, thereby posing a challenge to climate scientists: to quantify and predict how these reductions will influence Earth's energy balance and the ongoing rate and pattern of warming.

While focused studies have found evidence of the effect of the change in shipping emissions on cloud brightness (Diamond, 2023; Watson-Parris et al., 2022), it is challenging to discern any signal in large-scale cloud properties because the observed covariability does not always flow causally from the observed microphysical properties (Glassmeier et al., 2021; Stevens and Feingold, 2009). The instantaneous radiative forcing due to aerosol–cloud interactions from the 2020 change in shipping emissions is estimated to be 0.5 W m^{-2} in the annual mean within shipping corridors (Diamond, 2023). Best estimates of the global mean effective radiative forcing (ERF) resulting from an 80 % drop in shipping emissions range from 0.035 to 0.15 W m^{-2} across multiple models and methodologies (Gettelman et al., 2024; Skeie et al., 2024; Yoshioka et al., 2024). Yuan et al. (2024) report a forcing of 0.2 W m^{-2} averaged over the global ocean only, which is consistent with these global estimates.

A possible discernible climate influence of the IMO regulations became a topic for discussion in 2023, when observed global mean surface temperatures (GMSTs) set a record. The 2023 GMST anomaly exceeded predictions based on long-term climate change trends and internal variability, including the El Niño–Southern Oscillation (ENSO), by more than 0.2 °C, causing speculation that the reduction in SO$_2$ emissions from shipping could be one of the driving factors (Schmidt, 2024). Based on the estimates of ERF given above, however, calculations with simple energy balance models (EBMs) suggest that the warming associated with shipping changes since 2020 is unlikely to exceed 0.05 °C by 2023, with a long-term response of 0.07 °C (Gettelman et al., 2024). A weakness of the EBMs in this case, however, is that they generally assume a spatially homogeneous forcing and thus cannot account for the spatial heterogeneity in ocean feedbacks or climate responses to aerosol forcing, which is known to be substantial (Persad and Caldeira, 2018; Shindell et al., 2009; Westervelt et al., 2020).

To provide a quantitative estimate of the magnitude and pattern of the climate response to the IMO shipping regulations, and the role it may have played in recent record surface temperatures, it is therefore crucial to also have estimates using fully coupled Earth system models (ESMs) that take into account a broad range of aerosol–climate interactions and their spatial heterogeneity, along with internal variability and its potential feedback on transient climate forcing. The latter means that it is necessary to use an ensemble of model simulations that is sufficiently large to discern a statistically significant temperature response to a weak perturbation. Even with ESMs, structural uncertainty and ensemble size create disagreement, as evidenced by recent studies of surface warming estimates due to shipping emission reductions. For instance, Yoshioka et al. (2024) found a global mean temperature increase of 0.04 °C, averaged over 2020 to 2049, in response to a 0.13 W m^{-2} ERF in HadGEM3-GC3.1, while Quaglia and Visioni (2024) found a global temperature increase of 0.2 °C by 2030 in response to an approximately 0.2 W m^{-2} radiative perturbation in the Community Earth System Model version 2 (CESM2; for a 90 % reduction in shipping emissions).

Here, we present estimates of the transient surface temperature implications of the recent IMO regulations, using a large (18-member) ensemble of fully coupled transient simulations with CESM2. We show the ensemble mean response over time and discuss the implications of the sample size for the ability to quantify any forced warming. Also, as it is conceivable that the current specific phases of ENSO and the Atlantic Multi-decadal Oscillation (AMO) could be particularly (in)sensitive to radiative perturbations in the shipping corri-

dors (e.g. Wang et al., 2022), we have designed the ensemble to sample different modes of climate variability. Finally, as there have already been a range of studies quantifying the temperature response to the IMO regulations leading to differing, if not opposite, conclusions, we provide a broader discussion on the challenges and limitations of disentangling the effect of shipping emissions using currently available climate modelling methodologies.

2   Methods

CESM2 (Danabasoglu et al., 2020) consists of the Community Atmosphere Model version 6 (CAM6; Bogenschutz et al., 2018), the Parallel Ocean Program version 2 (POP2; Danabasoglu et al., 2012), the Community Land Model version 5 (CLM5; Lawrence et al., 2019), and the Community Ice Code version 5 (CICE; Hunke et al., 2013). Aerosols in CAM6 are represented by the four-mode version of the Modal Aerosol Module (MAM4; Liu et al., 2016). We note that 2.5 % of SO$_2$ emissions from the international shipping sector are emitted as sulfate aerosol at the surface level and into the accumulation mode (Emmons et al., 2020). The cloud microphysics scheme is version 2 of the Morrison–Gettelman scheme (Gettelman and Morrison, 2015). Simulations are performed at approximately 1° horizontal resolution.

Our baseline experiments come from archived simulations performed as part of the CESM2 Large Ensemble (CESM2-LE) project (Rodgers et al., 2021). From 2015 onwards, CESM2-LE uses aerosol/precursor gas emissions (including SO$_2$/SO$_4$ from international shipping), land use changes, and greenhouse gas concentrations from the Shared Socioeconomic Pathway 3-7.0 (SSP3-7.0; O'Neill et al., 2016). Our perturbation experiments are initialised from the CESM2-LE baseline experiments beginning in 2015 and integrated through 2040. The perturbation simulations are identical to the baseline simulations, except for a uniform 80 % reduction in SO$_2$/SO$_4$ emissions from international shipping starting in 2020 (consistent with the IMO regulations) and extending through 2040. Thus, taking a difference (perturbation minus baseline) isolates the effects of the decrease in SO$_2$ emissions from international shipping.

CESM2 has a relatively strong anthropogenic aerosol forcing when quantified in isolation and a high climate sensitivity compared to other ESMs (see Fig. 1; Schlund et al., 2020; Zelinka et al., 2023). Its oceanic response has also been shown to be particularly sensitive to aerosol emissions (Fasullo et al., 2023; Hassan et al., 2021). For this particular perturbation, however, CESM2 shows a similar radiative forcing to other ESMs (Skeie et al., 2024). Our model setup is the same as the one used for CESM2 in Gettelman et al., 2024, who report an ERF of 0.11 W m^{-2} for a 100 % ± 25 % reduction in shipping SO$_2$ emissions. This places CESM2 near the
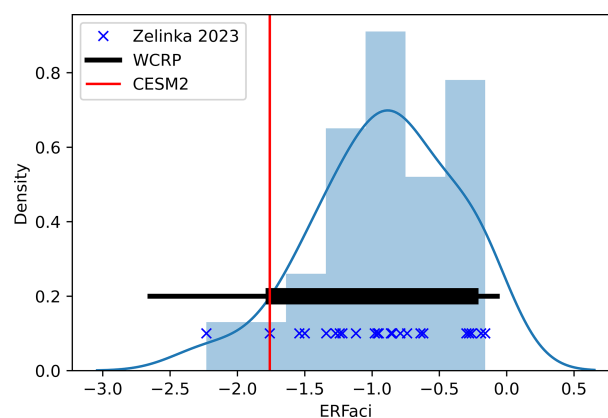

Figure 1. Distribution of the effective radiative forcing due to aerosol–cloud interactions (ERFaci) in CMIP6 models as assessed by Zelinka et al. (2023) in blue and Bellouin et al. (2020) in black, with CESM2 highlighted in red.

mean of the ERF estimates recently provided (Skeie et al., 2024; Gettelman et al., 2024).

To help understand the possible importance of the imposed shipping emission perturbation relative to internal climate variability, we perform 18 ensemble simulations. These ensemble members all use CESM2-LE realisations that feature 11-year running mean smoothed CMIP6 biomass burning emissions, including members 1011-001, 1031-002, 1051-003, 1091-005, 1111-006, 1131-007, 1151-008, 1171-009, 1191-010, 1231-011, 1231-012, 1231-016, 1231-018, 1251-012, 1281-017, 1281-020, 1301-015, and 1301-017. All comparisons and differences are calculated with respect to the 18 corresponding unperturbed ensemble members. Recent satellite data introduce more interannual variability into the biomass burning dataset than data sources used before 1997 and after 2014, so smoothing reduces the variability in biomass burning fluxes over 1990–2020 (Fasullo et al., 2022). Furthermore, to help understand the possible importance of dominant modes of climate variability (e.g. ENSO and AMV) to the climate impacts associated with the shipping perturbation, 8 of the above ensemble members feature ENSO-neutral conditions, 5 feature ENSO-positive conditions, and 5 feature ENSO-negative conditions. The AMV index is also evenly sampled in the ensemble, for each ENSO state, and spans −0.15 to 0.15.

The simulated responses are compared to observed SST changes, diagnosed from the HadCRUT v5.0.2 analysis (Morice et al., 2021). We fit a locally weighted regression (LOESS) model over time to each 5° × 5° grid cell to model the long-term SST trends (2020–2040) and then subtract this to discern the anomaly in 2023.

In the following, unless otherwise specified, estimates for surface temperature change are for the period 2020–2040. All significance tests are performed using a two-sided Student's t-test for the null hypothesis that two independent

samples (drawn from two distributions with equal variance) have identical average (expected) values.

3   Results

Figure 2a shows the relative change in near-surface SO$_2$ concentration in the reduction scenario compared to baseline and demonstrates that the majority of the changes in near-surface SO$_2$ occur over the Northern Hemisphere oceans and in particular over the North Atlantic and northeastern Pacific – clearly aligned with the main international shipping routes. Figure 2b shows the change in near-surface SO$_2$ concentrations over the simulation period both globally and over the North Atlantic. The gradual reduction in the change over the period is due to the underlying shipping emission reductions in the baseline SSP3-7.0 scenario.

Despite these widespread and regionally significant changes in SO$_2$ concentrations (and corresponding SO$_4$ concentrations, not shown), the temperature response in these simulations is negligible. Based on our sample of 18 simulations, we calculate a global 20-year mean surface temperature change from the IMO regulations of +0.03 °C, with a 5 %–95 % confidence range of [−0.09, 0.19]. Figure 3a shows the temperature change with respect to the baseline simulations as a global mean and over the North Atlantic, neither of which show significant spatial mean warming over the 20-year study period. In fact, over the first 5 years, there is no statistically robust change in temperature found anywhere on the globe (see Fig. 3b), demonstrating the low strength of the perturbation with respect to the model's simulated internal variability in the Earth system. Averaging over 20 years, a robust localised signal starts to appear during 2020–2040 in a small region of the North Atlantic (see Fig. 3c) where there is a statistically significant local warming of around 0.2 °C. Figure 3d shows that the anomalous warming observed in the region in 2023 (see Methods) does broadly correspond with the pattern of simulated warming in response to the IMO regulations but only as discerned after 20 years, visualised in Fig. 3e (which is a close-up of Fig. 3c).

Despite some degree of similarity between the recent observed anomalous warming and the simulated temperatures shown in Fig. 3e, the timescales do not match. The observed anomalies occurred 3 years after the emissions changes, while the average model response over the first 5 years (2020–2025) shows no significant warming in the region (Fig. 3b). As noted above, it is possible that the particular phase of ENSO or AMO could have made the North Atlantic particularly susceptible to such a perturbation in this short period of time in observations, but, by sub-sampling our ensemble based on these characteristic modes, we still find no evidence that the shipping emission changes could have contributed significantly to the observed global temperature changes (see Fig. 4).

As mentioned above, detecting the climate impacts of a relatively small externally forced perturbation, such as the estimated global top-of-atmosphere radiative forcing associated with the IMO shipping regulations (e.g. +0.12 ± 0.03 W m^{-2}; Gettelman et al., 2024), is complicated by the influence of internal climate variability (Deser et al., 2012, 2020). Figure 5 shows the actual evolution of our 18 ensemble members, with and without the IMO regulations, compared to the HadCRUT5 global surface temperature anomaly data series. The observed evolution is clearly within the range sampled by CESM2, in both emission scenarios. This difficulty is compounded over smaller (e.g. regional) spatial scales and shorter (e.g. decadal) timescales. The ability to robustly separate and quantify a forced signal in the climate system is the goal and motivation of large ensembles (e.g. Kay et al., 2015; Kirchmeier-Young et al., 2016; Maher et al., 2019; Rodgers et al., 2021; Simpson et al., 2023), whereby dozens or more independent climate model ensemble members are generated using identical external forcing but different initial climate states. Since ensemble members will in general feature different timing of internal climate variability, which essentially represents noise (i.e. the component of the signal that is not externally forced), averaging over a larger number of ensemble members reduces such noise, allowing a more robust quantification of the externally forced signal.

Figure 6 shows the important influence of internal climate variability on the global mean temperature response (dT) in our CESM2 shipping perturbation experiments and the importance of having a sufficient ensemble size to robustly detect a forced response. Figure 6a shows the impact of randomly selecting N of our 18 ensemble members with replacement (i.e. a bootstrapping analysis). For each combination, we calculate the corresponding ensemble global mean temperature response. The figure shows the results for $N = 1000$ samples. For small N, the spread in the ensemble mean ΔT is quite large (approximately −0.09 to 0.125 °C for $N = 5$). For $N = 10$, the spread is reduced, but it still exceeds ±0.05 °C. As N continues to increase, however, the spread converges to our 18-ensemble-member ΔT.

To further illustrate the importance of ensemble size when dealing with perturbations that have weak responses relative to internal variability, Fig. 6b–d show example combinations of 10 unique ensemble members (as used, for example, by Quaglia and Visioni, 2024). Statistical testing and hatching is done as for our Fig. 3 (see Methods). Figure 6b shows a 10-member combination consistent with our 18-member mean response ($\Delta T = 0.03$ °C). Figure 6c and d show 10-member combinations with ΔT at the edges of, but still within, the 9 %–95 % confidence interval ($\Delta T = \pm 0.05$ °C).

While panels b–d of Fig. 6 all represent deliberate picking of ensemble members, the analysis illustrates how, even with a decently sized ensemble of 10 members, one could conclude that the IMO shipping regulations may lead to substantial global mean warming or global mean cooling. This illustrates the importance of internal climate variability in de-
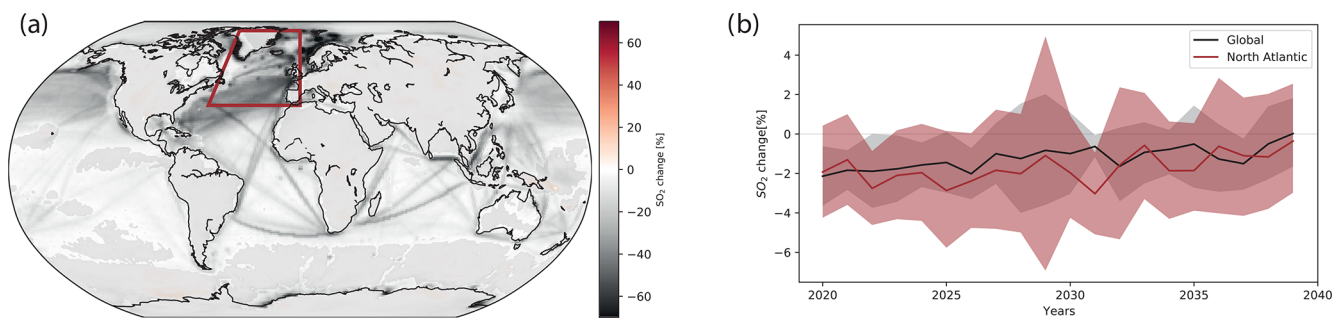
Figure 2. (a) Relative change in SO₂ near the surface due to the shipping emission perturbation, averaged over the whole 2020–2040 period, where stippling represents locations where the null hypothesis of "no change" cannot be rejected at $P < 0.05$. **(b)** Annual mean evolution of SO₂, averaged over the entire globe (black line; shading shows ± 1 standard deviation in ensemble member spread) and over the North Atlantic region (red line), as indicated on the map. All changes are relative to the baseline Shared Socioeconomic Pathway 3-7.0 (SSP3-7.0).
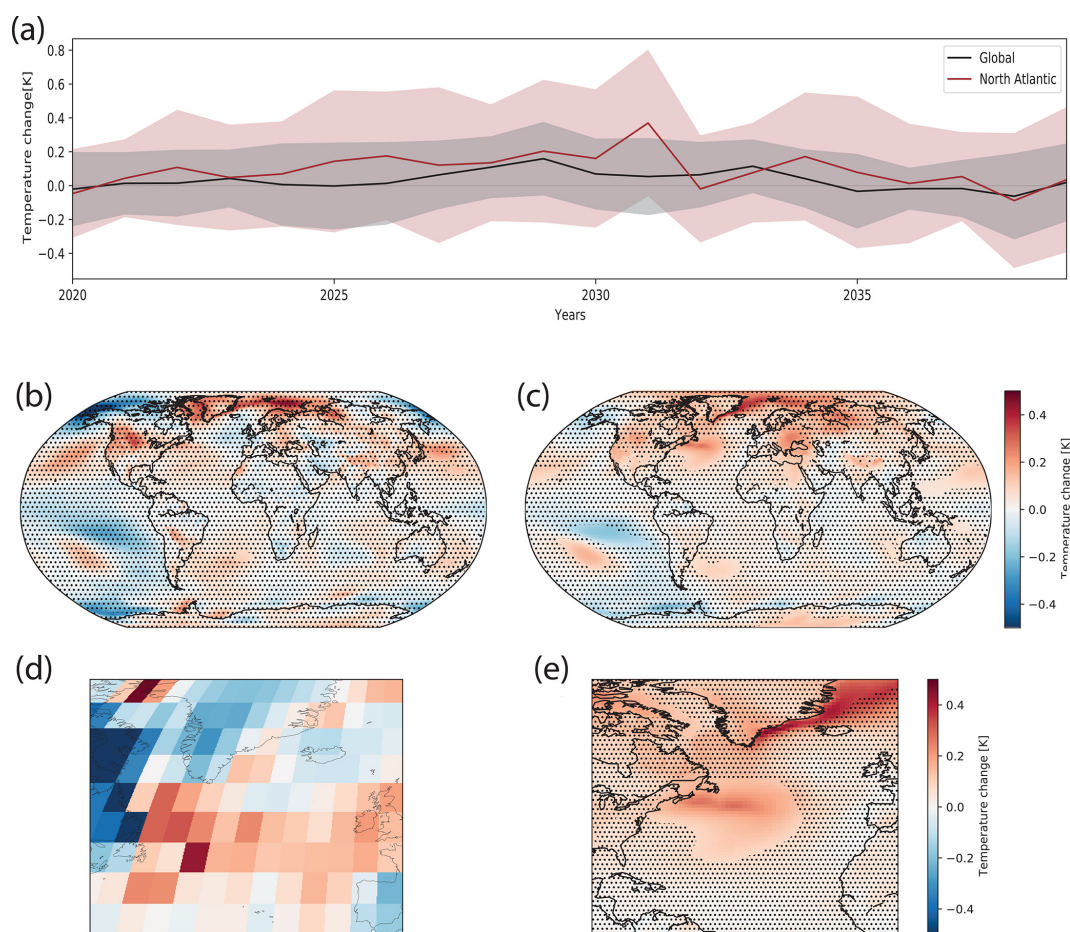
Figure 3. (a) Global and North Atlantic annual mean evolution in temperature change (shading shows ± 1 standard deviation in ensemble member spread). **(b)** Annual mean temperature change for 2020–2025. **(c)** Annual mean temperature change for 2020–2040. **(d)** Observed anomalous warming in North Atlantic sea-surface temperature in 2023. **(e)** A close-up of panel **(c)** showing the corresponding region of the North Atlantic. Stippling in panels **(b)**, **(c)**, and **(e)** represents locations where the null hypothesis of "no change" cannot be rejected at $P < 0.05$. All changes are relative to the corresponding baseline CESM2 LENS simulations of SSP3-7.0.
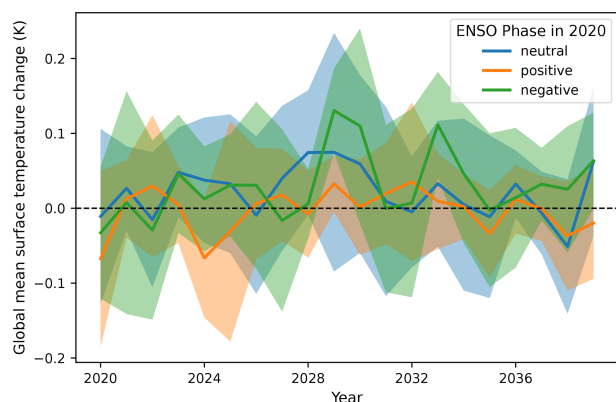
Figure 4. Global mean change in surface temperature due to shipping emission changes with respect to the baseline SSP3-7.0 simulations, subdivided into three ENSO phases in 2020: neutral (blue; 8 members), positive (orange; 5 members), and negative (green; 5 members). Shading shows the inter-member spread, represented by ±1 standard deviation.
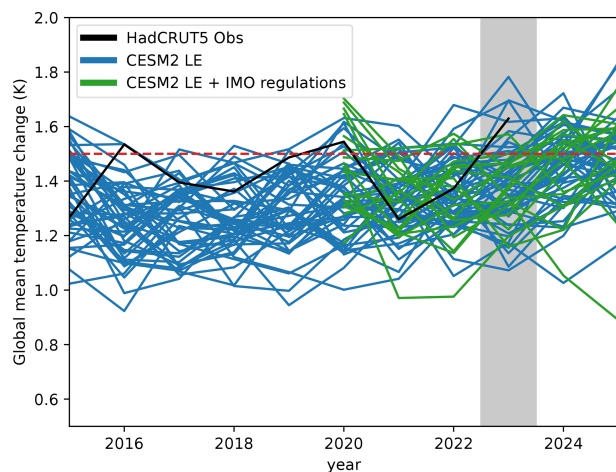

Figure 5. Global mean surface temperature change with respect to piControl for the baseline historical + SSP3-7.0 CESM2 LENS ensemble (blue) and the perturbed shipping emission ensemble performed in this study (green), overlaid on the HadCRUT SST observational estimate (with respect to the 1850–1900 average; black). The year 2023 is highlighted. The dashed red line represents the 1.5 °C Paris Agreement warming threshold.

tecting a global mean temperature response associated with the IMO shipping regulations and the importance of a sufficiently large ensemble size to reduce the risk of spurious conclusions. We note that our choice of 18 ensemble members, as with most experimental designs, also represents a trade-off between additional information gained versus increased computational expense. Clearly, however, a moderately large ensemble size of 10, which has been shown to be sufficient in some contexts (e.g. Monerie et al., 2022), is not sufficient to make robust claims regarding the impact of shipping SO₂

reductions on global mean temperature, in particular during its early transient evolution.

4 Discussion and conclusions

The strict new fuel regulations introduced by the IMO provided a valuable experiment to better understand the role of changing anthropogenic aerosol in the climate system, particularly as an analogue to other, current and future, efforts to improve air quality globally. Using an 18-member ensemble of simulations from CESM2, we find that the global temperature response to the IMO regulations that came into force in 2020 is +0.03 °C, with a 5 %–95 % confidence range of [−0.09, 0.19] for the period 2020–2040. This result, which is consistent with a null hypothesis of no discernible global mean temperature response, is at the low end of other estimates of this warming effect, from simple energy balance models (Gettelman et al., 2024; Yuan et al., 2024) and global climate models (Quaglia and Visioni, 2024; Yoshioka et al., 2024), which suggested global warming between 0.04 °C (Yoshioka et al., 2024) and 0.2 °C (Quaglia and Visioni, 2024) associated with the shipping regulations over decadal timescales. To understand the differences between these conclusions and to explain why our analysis provides an important bound on the role of shipping emission changes in recently observed temperature extremes, it is necessary to discuss in some detail the methodologies and framings behind the other results.

Some recent studies of the climate impact of the IMO regulations used EBMs to calculate the global temperature response from the effective radiative forcing (Gettelman et al., 2024; Yuan et al., 2024). These simple models are useful in many situations, notably for comparing the climate implications of known emissions from industrial sectors, regions, or scenarios. However, for absolute climate impacts, they require substantial assumptions to be made about the sensitivity and timescale of the responses, and critically do not account for regional heterogeneity of responses or climate feedbacks, such as influences on ocean circulation or sea ice, or of internal variability. Ocean feedbacks may be particularly important in the case of shipping emission changes which are focused over the Northern Hemisphere oceans, where it is conceivable that the ocean mixed layers may warm more efficiently than, for example, the Southern Ocean (e.g. Ma et al., 2020). Region-specific cloud responses and teleconnections to aerosol changes over Northern Hemisphere oceans may also differ from those to, for example, sulfur emission changes in Asia (Burney et al., 2022; Persad and Caldeira, 2018), which tend to dominate the total global response to recent sulfur emission changes used to calibrate simple EBMs. Coupled climate models provide estimates of the temperature response to the IMO regulations that take such feedbacks and pattern dependencies into account.
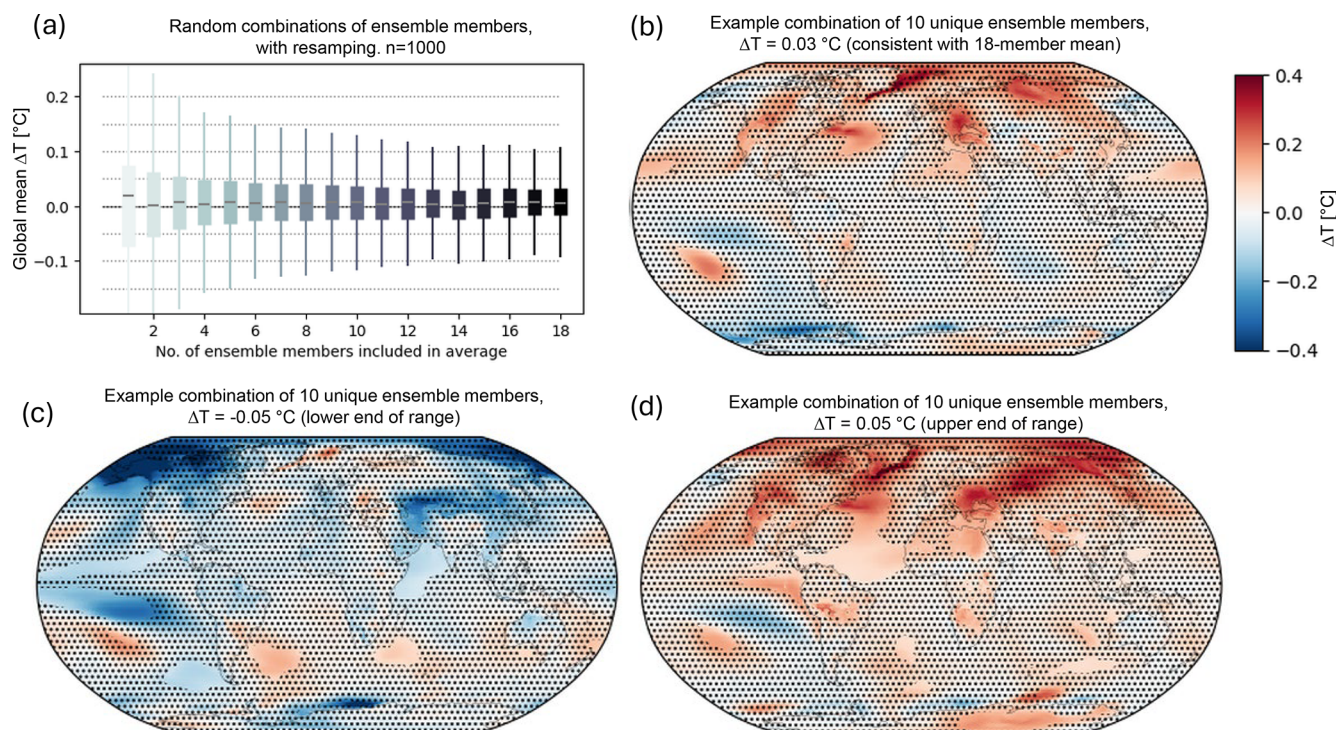
Figure 6. Ensemble size is crucial for quantifying a forced signal for weak perturbations. **(a)** Spread in the 2020–2040 global annual mean surface temperature change (ΔT) when randomly sampling and averaging N (given on the x axis) ensemble members out of a total of 18 available members 1000 times, with resampling. Each spread includes 1000 combinations, illustrating an increasing robustness of ΔT with the number of ensemble members. Boxes show the median and 5 %–95 % range, and whiskers show the maximum and minimum values. **(b)** Example combination of 10 ensemble members, where the global annual mean ΔT is consistent with the 18-member mean ($\Delta T = 0.03$ °C). Hatching shows grid points not significant at the 95 % confidence level, as for Fig. 2. **(c)** Same as panel **(b)** but for an example combination consistent with ΔT at the lower end of the 10-member mean range. ($\Delta T = -0.05$ °C). **(d)** Same as panel **(b)** but for an example combination consistent with ΔT at the upper end of the 10-member mean range. ($\Delta T = 0.05$ °C).

Despite the differences in the approach, our estimate of the global warming due to the IMO regulations is similar to the estimate by Gettelman et al. (2024) of 0.07 °C by 2030, which is based on the FaIR EBM. The EBM approach taken by Yuan et al. (2024), however, which diagnoses a global temperature anomaly of 0.17 °C at equilibrium, overstates the response to their forcing by a factor of 1.4. This is because Yuan et al. (2024) calculate a global temperature response using a global climate feedback parameter but also using an ocean area mean ERF. A global feedback parameter should be used with a global forcing in this context, which would reduce their forcing and temperature estimates by a factor of 0.7, the fraction of global area that is ocean, from 0.2 to 0.14 W m^{-2} and 0.17 to 0.12 °C in equilibrium, respectively. Uncertainty in the climate feedback parameter should also be considered in this estimate. Using a feedback parameter equivalent to that of CO$_2$, the AR6 likely range for ECS of 2.5–4.0 °C (which may both be overly simplistic assumptions) and 2 × CO$_2$ forcing of 3.9 ± 0.5 W m^{-2} (90 % range) would produce an equilibrium warming due to a forcing of 0.14 W m^{-2} of 0.09–0.14 °C (90 % range). Altogether, 7 years, the timescale used by Yuan et al. (2024) in their cal-

culation, is approximately the time taken for the upper ocean to reach equilibrium, so, based on Gregory et al. (2024), one would expect to see around two-thirds of the equilibrium response (i.e. 0.06–0.10 °C) in that time. With these factors taken into account, the estimates presented by Yuan et al. (2024) would have been consistent with those of Gettelman et al. (2024) and those presented in this work.

ESM estimates of the temperature response to the IMO regulations initially appear to show similar diversity to those based on EBMs. However, this can be seen to largely be due to a difference in the framing of the experiments and reporting of the results, rather than a substantive difference in the main conclusions. Using HadGEM3-GC3.1-LL, Yoshioka et al. (2024) estimated a global mean warming of 0.04 °C (averaged over 2020–2049), which is in agreement with our estimate of 0.03 [−0.09, 0.19] °C (averaged over 2020–2040). Their model has a similar ERF of 0.13 W m^{-2}, and they use coupled transient simulations with 12 ensemble members per experiment. Their global mean warming estimate is similar to that presented in this study, but the pattern of warming is markedly different, with most of the significant warming over SE Asia and the eastern Pacific. Conversely, Quaglia

and Visioni (2024) estimate a temperature increase of 0.2 °C by 2030 due to the introduction of the IMO regulations. They use the same model that we have used in this study and a similar experiment design, with transient simulations initialised from the CESM2 Large Ensemble. Although their temperature response is 1 order of magnitude larger than our stated response, it is not inconsistent with our results. We show in Fig. 3a that the global temperature response to the reduction in shipping emissions peaks in 2029 at 0.16 $[-0.17, 0.52]$ °C. Our best estimate of the warming by 2030 is 0.07 $[-0.14, 0.28]$ °C, where the range is ± 1 standard deviation and encompasses a warming of 0.2 °C. However, our 18-member ensemble shows that the global ensemble mean warming over 2020–2040 is not statistically significantly different from zero ($p = 0.18$), even in 2030 ($p = 0.054$).

The apparently large discrepancy between our CESM2 numbers and those from Quaglia and Visioni (2024) also demonstrates the importance of framing. We report a 2020–2040 mean value, and they report the value for 2030. Both our Fig. 3a and Fig. 6 in Yoshioka et al. (2024) show that the maximum global temperature response to the 2020 emission change occurs around 2030, and the estimate of the temperature response by 2030 is consistent across all three studies. Yoshioka et al. (2024) also find a long-term mean response of 0.04 °C, averaged over 2020–2049, which is more consistent with our estimate of 0.03 °C. However, while we show in Fig. 3a that the global warming in response to the IMO regulations is not significant, Yoshioka et al. (2024) present significant decadal mean warming in their Fig. 6. However, this is based on ± 1 standard error (SE), while we show ± 1 standard deviation. Yoshioka et al. (2024) use 12 members, so their standard deviation of $\sqrt{12} \cdot$ SE would also indicate that these decadal mean values were always within 1 standard deviation of 0, consistent with our results.

Using large ensembles of simulations differing only by the initial conditions is an important technique for distinguishing forced responses from internal variability, particularly when looking for regional responses, or considering small forcings. Here, our large ensemble confirms that the global response to IMO regulations in CESM2 cannot be distinguished from internal variability, at least in terms of global mean surface temperature. Our best estimate is not significantly different from zero, and subsampling our 18-member ensemble to produce estimates of the global temperature response based on different ensemble sizes always returns a mean estimate that is not significantly different from 0 at the 5 % level. In fact, given the ensemble variance and the number of simulations available, the minimum detectable effect size over the full 20 years at a significance level of 0.05 is approximately 0.048 °C. The warming due to the ship emission changes in CESM2 is therefore very likely less than 0.05 °C. A 10-member ensemble, as used by Quaglia and Visioni (2024), can only significantly detect an effect larger than 0.07 °C. To demonstrate this, 10 members subsampled from within our 18-member ensemble can give a global warming estimate of

0.05 °C or a global cooling of 0.05 °C, averaged over 2020–2040.

Large ensembles also make it possible to characterise the spatial pattern of the response and identify physically robust responses that are consistent across ensemble members. North Atlantic; Greenland, Iceland, and Norway (GIN); South Atlantic; and eastern Pacific warming are common features of all our subsampled ensembles. However, only the North Atlantic and GIN sea warming are physically robust in addition to being significant at the 5 % level. The difference between the pattern of warming shown in this work and in Yoshioka et al. (2024) likely has a large component of internal variability, in addition to the effects of structural differences between the models used.

Our results highlight the challenge in rapidly attributing observed extreme events to evolving or temporary anthropogenic changes, particularly given the large internal climate variability on annual to decadal timescales. By crowdsourcing computing resources, we were able to rapidly generate an ensemble of fully coupled Earth system model simulations of sufficient size to quantify the forced response and its confidence interval. However, this highlights a deficiency in current medium-term climate attribution tasks. As climate change is broadly acknowledged to be increasingly contributing to the extreme events experienced by millions of people around the world, climate scientists are increasingly tasked with understanding and accurately attributing them, but the resources to conduct this at scale are limited and ad hoc. An operational climate body that was specifically tasked with running decadal-scale attribution studies (Stevens, 2024), or the development of trusted methods for separating forced signals from variability (Samset et al., 2022), would provide a valuable resource as the demand for such information accelerates. The IMO shipping regulations do lead to a relatively small forced global mean surface warming in our results, consistent with its moderate aerosol ERF and also with a zero response when internal variability is taken into account. However, we note that other, future and ongoing, broader aerosol emission reductions, although uncertain (Persad et al., 2023), are expected to lead to a much larger aerosol ERF (e.g. Wilcox et al., 2023) and may thus still drive significant global warming and regional climate change impacts (e.g. Allen et al., 2020, 2021).

Author contributions. DWP, LJW, RJA, AW, and BHS designed the experiments, and DWP, CWS, RJA, and GP carried them out. DWP, LJW, CWS, and BHS performed the analysis and prepared the article with contributions from all co-authors.

References

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1230, https://doi.org/10.1126/science.245.4923.1227, 1989.

Allen, R. J., Turnock, S., Nabat, P., Neubauer, D., Lohmann, U., Olivié, D., Oshima, N., Michou, M., Wu, T., Zhang, J., Takemura, T., Schulz, M., Tsigaridis, K., Bauer, S. E., Emmons, L., Horowitz, L., Naik, V., Noije, T. van, Bergman, T., Lamarque, J.-F., Zanis, P., Tegen, I., Westervelt, D. M., Sager, P. L., Good, P., Shim, S., O'Connor, F., Akritidis, D., Georgoulias, A. K., Deushi, M., Sentman, L. T., John, J. G., Fujimori, S., and Collins, W. J.: Climate and air quality impacts due to mitigation of non-methane near-term climate forcers, Atmos. Chem. Phys., 20, 9641–9663, https://doi.org/10.5194/acp-20-9641-2020, 2020.

Allen, R. J., Horowitz, L. W., Naik, V., Oshima, N., O'Connor, F. M., Turnock, S., Shim, S., Sager, P. L., van Noije, T., Tsigaridis, K., Bauer, S. E., Sentman, L. T., John, J. G., Broderick, C., Deushi, M., Folberth, G. A., Fujimori, S., and Collins, W. J.: Significant climate benefits from near-term climate forcer mitigation in spite of aerosol reductions, Environ. Res. Lett., 16, 034010, https://doi.org/10.1088/1748-9326/abe06b, 2021.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A.-L., Dufresne, J.-L., Feingold, G., Fiedler, S., Forster, P., Gettelman, A., Haywood, J. M., Lohmann, U., Malavelle, F., Mauritsen, T., McCoy, D. T., Myhre, G., Mülmenstädt, J., Neubauer, D., Possner, A., Rugenstein, M., Sato, Y., Schulz, M., Schwartz, S. E., Sourdeval, O., Storelvmo, T., Toll, V., Winker, D., and Stevens, B.: Bounding Global Aerosol Radiative Forcing of Climate Change, Rev. Geophys., 58, e2019RG000660, https://doi.org/10.1029/2019rg000660, 2020.

Bogenschutz, P. A., Gettelman, A., Hannay, C., Larson, V. E., Neale, R. B., Craig, C., and Chen, C.-C.: The path to CAM6: coupled simulations with CAM5.4 and CAM5.5, Geosci. Model Dev., 11, 235–255, https://doi.org/10.5194/gmd-11-235-2018, 2018.

Burney, J., Persad, G., Proctor, J., Bendavid, E., Burke, M., and Heft-Neal, S.: Geographically resolved social cost of anthropogenic emissions accounting for both direct and climate-mediated effects, Sci. Adv., 8, eabn7307, https://doi.org/10.1126/sciadv.abn7307, 2022.

Christensen, M. W., Gettelman, A., Cermak, J., Dagan, G., Diamond, M., Douglas, A., Feingold, G., Glassmeier, F., Goren, T., Grosvenor, D. P., Gryspeerdt, E., Kahn, R., Li, Z., Ma, P.-L., Malavelle, F., McCoy, I. L., McCoy, D. T., McFarquhar, G., Mülmenstädt, J., Pal, S., Possner, A., Povey, A., Quaas, J., Rosenfeld, D., Schmidt, A., Schrödner, R., Sorooshian, A., Stier, P., Toll, V., Watson-Parris, D., Wood, R., Yang, M., and Yuan, T.: Opportunistic experiments to constrain aerosol effective radiative forcing, Atmos. Chem. Phys., 22, 641–674, https://doi.org/10.5194/acp-22-641-2022, 2022.

Danabasoglu, G., Bates, S. C., Briegleb, B. P., Jayne, S. R., Jochum, M., Large, W. G., Peacock, S., and Yeager, S. G.: The CCSM4 Ocean Component, J. Climate, 25, 1361–1389, https://doi.org/10.1175/jcli-d-11-00091.1, 2012.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kin-

nison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), J. Adv. Model. Earth Sy., 12, e2019MS001916, https://doi.org/10.1029/2019ms001916, 2020.

Deser, C., Knutti, R., Solomon, S., and Phillips, A. S.: Communication of the role of natural variability in future North American climate, Nat. Clim. Change, 2, 775–779, https://doi.org/10.1038/nclimate1562, 2012.

Deser, C., Phillips, A. S., Simpson, I. R., Rosenbloom, N., Coleman, D., Lehner, F., Pendergrass, A. G., DiNezio, P., and Stevenson, S.: Isolating the Evolving Contributions of Anthropogenic Aerosols and Greenhouse Gases: A New CESM1 Large Ensemble Community Resource, J. Climate, 33, 7835–7858, https://doi.org/10.1175/jcli-d-20-0123.1, 2020.

Diamond, M. S.: Detection of large-scale cloud microphysical changes within a major shipping corridor after implementation of the International Maritime Organization 2020 fuel sulfur regulations, Atmos. Chem. Phys., 23, 8259–8269, https://doi.org/10.5194/acp-23-8259-2023, 2023.

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G., Kinnison, D., Lamarque, J., Marsh, D., Mills, M. J., Tilmes, S., Bardeen, C., Buchholz, R. R., Conley, A., Gettelman, A., Garcia, R., Simpson, I., Blake, D. R., Meinardi, S., and Pétron, G.: The Chemistry Mechanism in the Community Earth System Model Version 2 (CESM2), J. Adv. Model. Earth Sy., 12, e2019MS001882, https://doi.org/10.1029/2019MS001882, 2020.

Fasullo, J. T., Lamarque, J., Hannay, C., Rosenbloom, N., Tilmes, S., DeRepentigny, P., Jahn, A., and Deser, C.: Spurious Late Historical-Era Warming in CESM2 Driven by Prescribed Biomass Burning Emissions, Geophys. Res. Lett., 49, e2021GL097420, https://doi.org/10.1029/2021gl097420, 2022.

Fasullo, J. T., Rosenbloom, N., and Buchholz, R.: A multiyear tropical Pacific cooling response to recent Australian wildfires in CESM2, Sci. Adv., 9, eadg1213, https://doi.org/10.1126/sciadv.adg1213, 2023.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J. L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity, Cambridge University Press, https://doi.org/10.1017/9781009157896.009, 2021.

Gettelman, A. and Morrison, H.: Advanced Two-Moment Bulk Microphysics for Global Models. Part I: Off-Line Tests and Comparison with Other Schemes, J. Climate, 28, 1268–1287, https://doi.org/10.1175/jcli-d-14-00102.1, 2015.

Gettelman, A., Christensen, M. W., Diamond, M. S., Gryspeerdt, E., Manshausen, P., Stier, P., Watson-Parris, D., Yang, M., Yoshioka, M., and Yuan, T.: Has Reducing Ship Emissions Brought Forward Global Warming?, Geophys. Res. Lett., 51, e2024GL109077, https://doi.org/10.1029/2024gl109077, 2024.

Glassmeier, F., Hoffmann, F., Johnson, J. S., Yamaguchi, T., Carslaw, K. S., and Feingold, G.: Aerosol-cloud-climate cooling overestimated by ship-track data, Science, 371, 485–489, https://doi.org/10.1126/science.abd3980, 2021.

Gregory, J. M., Bloch-Johnson, J., Couldrey, M. P., Exarchou, E., Griffies, S. M., Kuhlbrodt, T., Newsom, E., Saenko, O. A., Suzuki, T., Wu, Q., Urakawa, S., and Zanna, L.: A new conceptual model of global ocean heat uptake, Clim. Dynam., 62, 1669–1713, https://doi.org/10.1007/s00382-023-06989-z, 2024.

Hassan, T., Allen, R. J., Liu, W., and Randles, C. A.: Anthropogenic aerosol forcing of the Atlantic meridional overturning circulation and the associated mechanisms in CMIP6 models, Atmos. Chem. Phys., 21, 5821–5846, https://doi.org/10.5194/acp-21-5821-2021, 2021.

Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., and Elliott, S.: CICE: the Los Alamos Sea Ice Model Documentation and Software User's Manual Version 5. Los Alamos National Laboratory Technical Report, No. LA-CC-06-012, 2013.

IMO: IMO 2020: Consistent Implementation of MARPOL Annex VI, International Maritime Organization, ISBN 978-92-801-17189, 2019.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J.-F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M.: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability, B. Am. Meteorol. Soc., 96, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1, 2015.

Kirchmeier-Young, M. C., Zwiers, F. W., and Gillett, N. P.: Attribution of Extreme Events in Arctic Sea Ice Extent, J. Climate, 30, 553–571, https://doi.org/10.1175/JCLI-D-16-0412.1, 2016.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., Kampenhout, L. van, Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., Broeke, M. van den, Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Martin, M. V., and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, J. Adv. Model. Earth Sy., 11, 4245–4287, https://doi.org/10.1029/2018ms001583, 2019.

Li, C., McLinden, C., Fioletov, V., Krotkov, N., Carn, S., Joiner, J., Streets, D., He, H., Ren, X., Li, Z., and Dickerson, R. R.: India Is Overtaking China as the World's Largest Emitter of Anthropogenic Sulfur Dioxide, Sci. Rep.-UK, 7, 14304, https://doi.org/10.1038/s41598-017-14639-8, 2017.

Liu, X., Ma, P.-L., Wang, H., Tilmes, S., Singh, B., Easter, R. C., Ghan, S. J., and Rasch, P. J.: Description and evaluation of a new four-mode version of the Modal Aerosol Module (MAM4) within version 5.3 of the Community Atmosphere Model, Geosci. Model Dev., 9, 505–522, https://doi.org/10.5194/gmd-9-505-2016, 2016.

Ma, X., Liu, W., Allen, R. J., Huang, G., and Li, X.: Dependence of regional ocean heat uptake on anthropogenic warming scenarios, Sci. Adv., 6, eabc0303, https://doi.org/10.1126/sciadv.abc0303, 2020.

Maher, N., Milinski, S., Suarez-Gutierrez, L., Botzet, M., Dobrynin, M., Kornblueh, L., Kröger, J., Takano, Y., Ghosh, R.,

Hedemann, C., Li, C., Li, H., Manzini, E., Notz, D., Putrasahan, D., Boysen, L., Claussen, M., Ilyina, T., Olonscheck, D., Raddatz, T., Stevens, B., and Marotzke, J.: The Max Planck Institute Grand Ensemble: Enabling the Exploration of Climate System Variability, J. Adv. Model. Earth Sy., 11, 2050–2069, https://doi.org/10.1029/2019ms001639, 2019.

Monerie, P.-A., Wilcox, L. J., and Turner, A. G.: Effects of Anthropogenic Aerosol and Greenhouse Gas Emissions on Northern Hemisphere Monsoon Precipitation: Mechanisms and Uncertainty, J. Climate, 35, 2305–2326, https://doi.org/10.1175/JCLI-D-21-0412.1, 2022.

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set, J. Geophys. Res.-Atmos., 126, e2019JD032361, https://doi.org/10.1029/2019JD032361, 2021.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geosci. Model Dev., 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

Persad, G. G. and Caldeira, K.: Divergent global-scale temperature effects from identical aerosols emitted in different regions, Nat. Commun., 9, 3289, https://doi.org/10.1038/s41467-018-05838-6, 2018.

Persad, G., Samset, B. H., Wilcox, L. J., Allen, R. J., Bollasina, M. A., Booth, B. B. B., Bonfils, C., Crocker, T., Joshi, M., Lund, M. T., Marvel, K., Merikanto, J., Nordling, K., Undorf, S., van Vuuren, D. P., Westervelt, D. M., and Zhao, A.: Rapidly evolving aerosol emissions are a dangerous omission from nearterm climate risk assessments, Environ. Res.: Clim., 2, 032001, https://doi.org/10.1088/2752-5295/acd6af, 2023.

Quaglia, I. and Visioni, D.: Modeling 2020 regulatory changes in international shipping emissions helps explain anomalous 2023 warming, Earth Syst. Dynam., 15, 1527–1541, https://doi.org/10.5194/esd-15-1527-2024, 2024.

Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., Edwards, J., Kim, J.-E., Simpson, I. R., Stein, K., Stuecker, M. F., Yamaguchi, R., Bódai, T., Chung, E.-S., Huang, L., Kim, W. M., Lamarque, J.-F., Lombardozzi, D. L., Wieder, W. R., and Yeager, S. G.: Ubiquity of human-induced changes in climate variability, Earth Syst. Dynam., 12, 1393–1411, https://doi.org/10.5194/esd-12-1393-2021, 2021.

Samset, B. H., Lund, M. T., Bollasina, M., Myhre, G., and Wilcox, L.: Emerging Asian aerosol patterns, Nat. Geosci., 12, 582–584, https://doi.org/10.1038/s41561-019-0424-5, 2019.

Samset, B. H., Zhou, C., Fuglestvedt, J. S., Lund, M. T., Marotzke, J., and Zelinka, M. D.: Earlier emergence of a temperature response to mitigation by filtering annual variability, Nat. Commun., 13, 1578, https://doi.org/10.1038/s41467-022-29247-y, 2022.

Schlund, M., Lauer, A., Gentine, P., Sherwood, S. C., and Eyring, V.: Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6?, Earth Syst. Dynam., 11, 1233–1258, https://doi.org/10.5194/esd-11-1233-2020, 2020.

Schmidt, G.: Climate models can't explain 2023's huge heat anomaly – we could be in uncharted territory, Nature, 627, 467, https://doi.org/10.1038/d41586-024-00816-z, 2024.

Shindell, D. T., Faluvegi, G., Koch, D. M., Schmidt, G. A., Unger, N., and Bauer, S. E.: Improved Attribution of Climate Forcing to Emissions, Science, 326, 716–718, https://doi.org/10.1126/science.1174760, 2009.

Simpson, I. R., Rosenbloom, N., Danabasoglu, G., Deser, C., Yeager, S. G., McCluskey, C. S., Yamaguchi, R., Lamarque, J.-F., Tilmes, S., Mills, M. J., and Rodgers, K. B.: The CESM2 Single-Forcing Large Ensemble and Comparison to CESM1: Implications for Experimental Design, J. Climate, 36, 5687–5711, https://doi.org/10.1175/JCLI-D-22-0666.1, 2023.

Skeie, R. B., Byrom, R., Hodnebrog, Ø., Jouan, C., and Myhre, G.: Multi-model effective radiative forcing of the 2020 sulfur cap for shipping, Atmos. Chem. Phys., 24, 13361–13370, https://doi.org/10.5194/acp-24-13361-2024, 2024.

Stevens, B.: A Perspective on the Future of CMIP, AGU Adv., 5, e2023AV001086, https://doi.org/10.1029/2023AV001086, 2024.

Stevens, B. and Feingold, G.: Untangling aerosol effects on clouds and precipitation in a buffered system, Nature, 461, 607, https://doi.org/10.1038/nature08281, 2009.

Twomey, S.: Pollution and the planetary albedo, Atmos. Environ, 1967, 1251–1256, https://doi.org/10.1016/0004-6981(74)90004-3, 1974.

van der A, R. J., Mijling, B., Ding, J., Koukouli, M. E., Liu, F., Li, Q., Mao, H., and Theys, N.: Cleaning up the air: effectiveness of air quality policy for SO$_2$ and NO$_x$ emissions in China, Atmos. Chem. Phys., 17, 1775–1789, https://doi.org/10.5194/acp-17-1775-2017, 2017.

Wall, C. J., Norris, J. R., Possner, A., McCoy, D. T., McCoy, I. L., and Lutsko, N. J.: Assessing effective radiative forcing from aerosol–cloud interactions over the global ocean, P. Natl Acad. Sci. USA, 119, e2210481119, https://doi.org/10.1073/pnas.2210481119, 2022.

Wang, G., Cai, W., Santoso, A., Wu, L., Fyfe, J. C., Yeh, S.-W., Ng, B., Yang, K., and McPhaden, M. J.: Future Southern Ocean warming linked to projected ENSO variability, Nat. Clim. Change, 12, 649–654, https://doi.org/10.1038/s41558-022-01398-2, 2022.

Watson-Parris, D.: Surface temperature changes due to shipping emissions changes simulated with CESM2, Zenodo [data set], https://doi.org/10.5281/zenodo.15185389, 2025.

Watson-Parris, D. and Smith, C. J.: Large uncertainty in future warming due to aerosol forcing, Nat. Clim. Change, 12, 1111–1113, https://doi.org/10.1038/s41558-022-01516-0, 2022.

Watson-Parris, D., Christensen, M. W., Laurenson, A., Clewley, D., Gryspeerdt, E., and Stier, P.: Shipping regulations lead to large reduction in cloud perturbations, P. Natl. Acad. Sci. USA, 119, e2206885119, https://doi.org/10.1073/pnas.2206885119, 2022.

Westervelt, D. M., Mascioli, N. R., Fiore, A. M., Conley, A. J., Lamarque, J.-F., Shindell, D. T., Faluvegi, G., Previdi, M., Correa, G., and Horowitz, L. W.: Local and remote mean and extreme temperature response to regional aerosol emissions reductions, Atmos. Chem. Phys., 20, 3009–3027, https://doi.org/10.5194/acp-20-3009-2020, 2020.

Wilcox, L. J., Allen, R. J., Samset, B. H., Bollasina, M. A., Griffiths, P. T., Keeble, J., Lund, M. T., Makkonen, R., Merikanto, J., O'Donnell, D., Paynter, D. J., Persad, G. G., Rumbold,

S. T., Takemura, T., Tsigaridis, K., Undorf, S., and Westervelt, D. M.: The Regional Aerosol Model Intercomparison Project (RAMIP), Geosci. Model Dev., 16, 4451–4479, https://doi.org/10.5194/gmd-16-4451-2023, 2023.

Yoshioka, M., Grosvenor, D. P., Booth, B. B. B., Morice, C. P., and Carslaw, K. S.: Warming effects of reduced sulfur emissions from shipping, Atmos. Chem. Phys., 24, 13681–13692, https://doi.org/10.5194/acp-24-13681-2024, 2024.

Yuan, T., Song, H., Oreopoulos, L., Wood, R., Bian, H., Breen, K., Chin, M., Yu, H., Barahona, D., Meyer, K., and Platnick, S.: Abrupt reduction in shipping emission as an inadvertent geoengineering termination shock produces substantial radiative warming, Commun. Earth Environ., 5, 281, https://doi.org/10.1038/s43247-024-01442-3, 2024.

Zelinka, M. D., Smith, C. J., Qin, Y., and Taylor, K. E.: Comparison of methods to estimate aerosol effective radiative forcings in climate models, Atmos. Chem. Phys., 23, 8879–8898, https://doi.org/10.5194/acp-23-8879-2023, 2023.