

JGR Atmospheres



RESEARCH ARTICLE

10.1029/2025JD044403

Key Points:

- The stratospheric polar vortex response to Arctic sea-ice loss is weak and not statistically significant in most models
- Diverse stratospheric responses across models arise partly due to differences in the models' simulated basic states
- HadGEM3-GC31-MM, with a realistic basic state, simulates a robust weakening and shift of the polar vortex, with more frequent sudden warmings

Correspondence to:

R. Mudhar,
rm811@exeter.ac.uk

Citation:

Mudhar, R., Seviour, W. J. M., Screen, J. A., Thomson, S. I., & Turrell, C. (2025). Diverse and weak simulated stratospheric responses to future Arctic sea-ice loss. *Journal of Geophysical Research: Atmospheres*, 130, e2025JD044403. <https://doi.org/10.1029/2025JD044403>

Received 21 MAY 2025

Accepted 9 OCT 2025

Diverse and Weak Simulated Stratospheric Responses to Future Arctic Sea-Ice Loss

Regan Mudhar¹ , William J. M. Seviour^{1,2}, James A. Screen¹ , Stephen I. Thomson¹ , and Charles Turrell¹ 

¹Department of Mathematics and Statistics, University of Exeter, Exeter, UK, ²Global Systems Institute, University of Exeter, Exeter, UK

Abstract Climate models project that the Arctic Ocean could see an ice-free summer by the middle of this century. Through coordinated simulations, the Polar Amplification Model Intercomparison Project (PAMIP) aims to elucidate the causes and consequences of future Arctic sea-ice loss. There is particular interest in understanding the mechanisms by which midlatitude weather and climate may be impacted, including via the “stratospheric pathway”; Arctic sea-ice loss and associated warming is proposed to induce a wave-driven weakening of the wintertime stratospheric polar vortex, which could subsequently impact tropospheric circulation. However, this is not well understood: studies do not so far find a robust stratospheric response to Arctic sea-ice loss in either strength or sign. Here, we conduct novel analysis of the stratospheric response in thirteen PAMIP-contributing models, looking beyond the typical time- and zonal-mean diagnostics. Although our results overall confirm the lack of robust response, one model, HadGEM3-GC31-MM, has a statistically significant equatorward shift in vortex latitude, deceleration of vortex winds, and increase in sudden stratospheric warmings. Its response is found to be highly state-dependent, significant only in the easterly phase of the quasi-biennial oscillation (QBO). Though we cannot comprehensively conclude why models simulate this range of responses, our analysis does highlight areas for consideration in future work to better constrain the stratospheric response to Arctic sea-ice loss. We explore the role of ensemble size, resolution and basic state, including zonal-mean winds in the polar and midlatitude stratosphere and upper troposphere, as well as the QBO.

Plain Language Summary Since the 1980s, the Arctic has warmed much faster than the rest of the world and summertime sea-ice has halved. This dramatic change is thought to have far-reaching impacts. It may, for example, affect the jet stream and the weather systems it carries over northern America, Asia and Europe. One suggested pathway for impacts goes via the stratosphere: Arctic change could trigger atmospheric disturbances that disrupt the typically strong and stable band of winds circling the north pole in the winter, known as the “stratospheric polar vortex.” We use climate models to study changes in the characteristics and behavior of the vortex in response to possible future Arctic sea-ice loss. We find no consistent change in the average state nor variability of the vortex across models. However, one model does show a significant weakening and shift of the vortex. Though it is tricky to unpick why certain models differ from the others, our findings highlight that future studies into the response to Arctic sea-ice loss may benefit from careful consideration of model resolution, ensemble size, and representation of the stratosphere's basic state, including the polar vortex, tropical lower stratosphere, and midlatitude upper troposphere/lower stratosphere winds.

1. Introduction

In the Arctic Ocean, significant loss of sea-ice is considered a major contributor to strong near-surface atmospheric warming (Screen & Simmonds, 2010), in conjunction with other atmospheric and oceanic processes (see Jenkins et al., 2024). With ongoing human emissions, climate models project a sea-ice-free Arctic summer by the middle of this century (Notz & SIMIP Community, 2020; Notz & Stroeve, 2016). Whilst the consequences of sea-ice loss likely extend well beyond the Arctic, the precise nature of the changes in atmospheric circulation are uncertain. As a result, the impacts on midlatitude weather and climate, and on the communities that live across the northern midlatitudes, are also uncertain (Cohen et al., 2020).

One of the proposed mechanisms linking Arctic sea-ice loss and associated warming to midlatitude circulation is the so-called “stratospheric pathway.” The hypothesis is that Arctic change modulates upward-propagating planetary waves which, if sufficiently strong, can disturb the stratospheric polar vortex (Kim et al., 2014;

© 2025. The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Kretschmer et al., 2018). In particular, strong wave-driving can increase the chances of sudden stratospheric warmings (SSWs): dramatic breakdowns of the wintertime vortex whose own anomalies can influence the surface for up to weeks at a time (Baldwin et al., 2021; Domeisen & Butler, 2020; Kolstad et al., 2022). However, the stratospheric pathway mechanism is not certain—including the influence of Arctic sea-ice loss on the stratospheric polar vortex itself. Some studies also find that, though the stratosphere may respond to Arctic sea-ice loss, stratosphere–troposphere coupling anomalies may not necessarily always extend to the surface (Cai et al., 2012; Xu et al., 2023).

In general, studies find the simulated stratospheric response to greenhouse gas forcings are model-dependent (Ayarzagüena et al., 2020; Karpechko et al., 2022, 2024). Arctic sea-ice loss is one component of greenhouse gas-forced climate change that is often studied in isolation; stratospheric responses to sea-ice loss also appear to be model-dependent. There are a variety of systematic model biases which are understood to contribute to the range of such responses (Smith et al., 2017; Sun et al., 2015). One way in which this can happen is through modification of a model's “basic state,” with some studies noting the dependency of the stratospheric response on both oceanic and atmospheric state. In particular, there appears to be a role for the climatological stratosphere itself. Idealized model experiments find that the vortex response to polar warming is sensitive to both the climatological polar vortex and “neck” winds (Mudhar et al., 2024). This neck region between the stratospheric polar vortex and tropospheric jet has similarly been identified as a key region in studies with more complex models. This includes the response to general greenhouse gas forcings (Karpechko et al., 2024), as well as Arctic-specific forcings. For example, Sigmond and Sun (2024) propose that the climatological neck winds control the amount of anomalous sea-ice loss-induced planetary wave activity that can propagate into the stratosphere. Other studies have found a dependency on the tropical lower-stratosphere winds, or quasi-biennial oscillation (QBO) phase, which may modify the climatological vortex, and hence its response to forcing, as well as the propagation of anomalous wave activity into the vortex directly (Labe et al., 2019; Xu et al., 2024).

There has not yet been a systematic and/or detailed multi-model comparison of the stratospheric response to Arctic sea-ice loss, with the majority of relevant studies discussed above using a single model. Here we make use of simulations from the Polar Amplification Model Intercomparison Project (PAMIP), part of the Sixth phase of the Coupled Model Intercomparison Project (CMIP6), to do so. PAMIP involves numerous international modeling groups performing coordinated large ensemble sea-ice loss experiments (Smith et al., 2018). Previous analysis of the PAMIP simulations of future Arctic change suggest a robust negative north Atlantic oscillation (NAO) response across atmosphere-only models, with an equatorward shift of the tropospheric jet (Smith et al., 2022). Although the sign of the jet shift appears consistent, the magnitude varies widely across the models, implying large uncertainty in any downstream changes in midlatitude weather and climate. The stratospheric response is even more uncertain, with no consistent sign or strength of change in winter-mean zonal-mean stratospheric polar vortex winds across models (Smith et al., 2022; Zheng et al., 2023). This stratospheric uncertainty may also contribute to the range of magnitudes of the simulated tropospheric jet shift, though is likely not critical for the direction (Mudhar et al., 2024; Smith et al., 2022).

In this study, we investigate the response of the stratosphere to future Arctic sea-ice loss using PAMIP simulations. As well as predominantly using individual models, the research to-date has typically focused on the change in the zonal-mean strength of time-mean stratospheric westerlies within the wintertime vortex. We build on this by examining the change in vortex geometry and variability, too. We aim to address two questions:

1. How robust is the stratospheric response to Arctic sea-ice loss in the models, both in terms of mean state and variability?
2. If responses differ significantly between models, what drives those differences?

2. Methods

2.1. Model Experiments

We utilize coordinated PAMIP experiments run with 13 atmospheric global climate models (Smith et al., 2018). Experiment 1.1 *pdSST*–*pdSIC* simulates the “present day” climate by prescribing the surface with present day estimates of sea surface temperatures (SSTs) and sea-ice concentration (SIC). Experiment 1.6 *pdSST*–*futArcSIC* is the same as 1.1 but the Arctic SIC is based on that expected to result from a global warming of 2°C compared to pre-industrial conditions, with future SSTs used only where sea-ice is lost compared to in 1.1.

Table 1

Models Used, Resolution (Degrees Latitude $\times$ Degrees Longitude $\times$ Number of Vertical Levels $\times$ Model Top in hPa), and Ensemble Size (Monthly/Daily)

Institute	Model	Resolution	Members
Alfred Wegener Institute Bremerhaven; Max Planck Institute for Meteorology, Hamburg	AWI-CM-1-1-MR	$0.94 \times 0.94 \times 95 \times 0.01$	100/100
Canadian Centre for Climate Modelling and Analysis	CanESM5	$2.8 \times 2.8 \times 49 \times 1$	300/100
Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique	CNRM-CM6-1	$1.0 \times 1.0 \times 91 \times 0.01$	300/100
Institute of Atmospheric Physics, Beijing	FGOALS-f3-L	$1.0 \times 1.0 \times 32 \times 2.19$	100/100
University of Exeter	HadGEM3-GC31-LL	$1.25 \times 1.875 \times 85 \times 0.005$	200/191
Met Office UK	HadGEM3-GC31-MM	$0.55 \times 0.83 \times 85 \times 0.005$	300/300
Institute Pierre Simon Laplace	IPSL-CM6A-LR	$1.26 \times 1.25 \times 79 \times 0.01$	200/200
University of Tokyo/National Institute for Environmental Studies/Japan Agency for Marine-Earth Science and Technology	MIROC6	$1.4 \times 1.4 \times 81 \times 0.004$	100/100
National Center of Atmospheric Research	CESM2	$1.25 \times 1.0 \times 32 \times 2.25$	200 ^a /–
US Department of Energy/University of California Irvine	E3SMv1	$1.0 \times 1.0 \times 72 \times 0.12$	200 ^b /–
Barcelona Super Computing Centre	EC-EARTH3	$1.0 \times 1.0 \times 91 \times 0.01$	150 ^a /–
Norwegian Meteorological Institute	NorESM2-LM	$1.9 \times 2.5 \times 32 \times 3.6$	100/–
Research Center for Environmental Changes, Academia Sinica (Taiwan)	TaiESM1	$0.9 \times 1.25 \times 30 \times 2$	199 ^c /–

Note. The top eight models are those with daily data available. ^aOnly 100 members are available for temperature and geopotential height. ^bOnly 298 members are available for geopotential height. ^cOnly 91 are members available for pdSST-futArcSIC .

The difference between 1.6 and 1.1 can thus be considered the simulated response to future Arctic sea-ice loss and associated local changes in SSTs (Smith et al., 2018, 2022). Each simulation starts on 1 April and runs for 14 months, though we focus on the northern hemisphere extended winter period from October to April. In PAMIP, all of the models are required to have at least 100 ensemble members (Table 1) and are forced with the same SSTs and SIC.

We analyze monthly-mean zonal wind, temperature and geopotential height from all models listed, and daily-mean zonal wind data from a subset of these. Our analysis of vortex variability requires daily data and we find that the appropriate variables at this frequency are only available for eight of the PAMIP-contributing models; in Sections 3 and 4 we make sure to highlight where we focus only on these models. This subset of models samples a range of horizontal resolutions, vertical levels and model tops, as well as stratospheric polar vortex climatologies and responses, making it a reasonably representative sample of the wider PAMIP multi-model ensemble. We also compare zonal wind and geopotential height-derived diagnostics to that of ERA5 (Hersbach et al., 2020), using daily data for the extended winter periods of 1979/80 to 2021/22 at a horizontal resolution of 1° and with 37 vertical levels up to 1 hPa.

2.2. Diagnostics

We look at the change in the stratospheric polar vortex mean state and its variability between experiments using the 10 hPa, 60°N zonal wind, or that derived from each models' closest grid point. We refer to the former as $U_{10,60}$. For variability, we use standard deviation of these winds ($\sigma_{10,60}$), SSW frequency and final warming (FW) date, all calculated from daily data. We also look at zonal wind in the neck region at 70 hPa and averaged over $45\text{--}55^\circ\text{N}$ (U_{neck} and σ_{neck}). Unless otherwise stated, these zonal wind diagnostics are zonal, ensemble, and northern winter (December–February, DJF) means.

SSWs are identified following the Charlton and Polvani (2007) approach for “major” warmings. Here, we look for a reversal in the sign of daily $U_{10,60}$ over extended winter, requiring that (a) winds must return to westerly for at least 20 consecutive days between reversals and (b) that they return to westerly for at least 10 consecutive days before 30 April. Given that the PAMIP experiments are 14-month timeslices, with the first 2 disregarded for spin-up we treat each ensemble member as a “year.” FW dates, the last date in spring when $U_{10,60}$ reverses, are also identified using daily data extending to the end of May (Ayarzagüena et al., 2020; Butler & Gerber, 2018). Applying these criteria to the ERA5 data gives a mean SSW frequency of 0.61 year^{-1} and a median FW date of

~13 April. As they require daily data, SSW frequency and FW date are only analyzed using the eight models that have such data available.

For some of our analysis, we separate ensemble members by QBO phase. We use the zonal-mean early winter (October–November) 30 hPa, 5°S–5°N-mean zonal wind as our QBO index (U_{QBO} , Labe et al., 2019; Richter et al., 2020). We then classify ensemble members into westerly (QBO-W) or easterly (QBO-E) phases as follows,

$$U_{\text{QBO}} \begin{cases} \geq 5 \text{ m s}^{-1} & : \text{QBO-W} \\ \leq -5 \text{ m s}^{-1} & : \text{QBO-E,} \end{cases} \quad (1)$$

such that members with $-5 \text{ m s}^{-1} < U_{\text{QBO}} < 5 \text{ m s}^{-1}$ are classified as “neutral.”

We also look at the geometry of the stratospheric polar vortex by applying moment diagnostic methods (described in detail in Seviour et al., 2013) to December–March monthly-mean 10 hPa geopotential height. Such “vortex moments” provide a 2D geometrical description of the vortex by essentially fitting an equivalent ellipse to a contour that defines the vortex edge from which we can infer the location of its center (as described by the centroid latitude, ϕ , and longitude, λ) and how “stretched” it is (described by the ratio of its major to minor axis, or aspect ratio, a). ERA5 data reveals a mean northern hemisphere winter vortex that is slightly elliptical, with $a \sim 1.6$, and centered just north of 75°N.

Finally, we calculate wave-1 refractive index using monthly-mean zonal wind and temperature ($n_{k=1}^2$; Matsuno, 1970; Mudhar et al., 2024). Refractive index generally describes regions in which waves are able to propagate: waves decay in regions of negative n^2 , such that wave propagation is easiest in regions of positive n^2 , indicating a kind of “waveguide.” By using wavenumber $k = 1$, we specifically focus on the waveguide for planetary-scale waves. Thus, analyzing $n_{k=1}^2$ enables us to gain some insight into potential changes in upward wave propagation that could contribute to the stratospheric response.

Given the inconsistencies between some of the models' daily versus monthly ensemble sizes, depending on the data available (Table 1), we only compare diagnostics calculated using data that has consistent time resolution and/or ensemble members. In other words, we do not draw relationships between monthly and daily data, since the latter's members may be a subset of the former. This is particularly key when comparing mean strength of winds to their variability (e.g., standard deviation or SSW frequency). Where appropriate, we also identify where mean responses are significant using a bootstrapping approach. This is done by randomly resampling data from the two experiments N times with replacement, where N is the total number of ensemble members, then taking the difference. This is done 1,000 times (quicker than but comparable with 10,000), with the resultant distribution of differences used to identify whether the sign of change is significant to the 97.5% confidence level; that is, if <2.5% of responses are of a different sign, then we say the response is significant.

3. Stratospheric Response to Arctic Sea-Ice Loss in PAMIP Simulations

Across PAMIP-contributing models, the average response to future Arctic sea-ice loss is a weakening of the winter stratospheric winds—but this is not robust. We see this across all of the models analyzed here (Figure 1), including the smaller subset of models with daily data, corroborating previous work by Smith et al. (2022). The individual models clearly disagree on the sign of the response, though changes are fairly small in most (Figures 1a–1m). CNRM-CM6-1, E3SMv1, HadGEM3-GC31-LL, MIROC6, and NorESM2 all show a strengthening of $U_{10,60}$, though none with any significance. The rest of the models show a weakening, but only CanESM5 and HadGEM3-GC31-MM have a significant response in polar latitudes as high as 10 hPa; CanESM5 does not show a significant $U_{10,60}$ response specifically, though TaiESM1 does.

Similarly, the daily evolution of $U_{10,60}$ over the extended winter shows fairly small changes between climatology (pdSST–pdSIC) and the future experiment (pdSST–futArcSIC, Figure 2) across our eight-model subset. Within DJF, CanESM5 and HadGEM3-GC31-MM have the largest decelerations, though only the latter has a consistently significant response (from around January onwards). Interestingly, CNRM-CM6-1 seems to experience a late, and significant, strengthening in March.

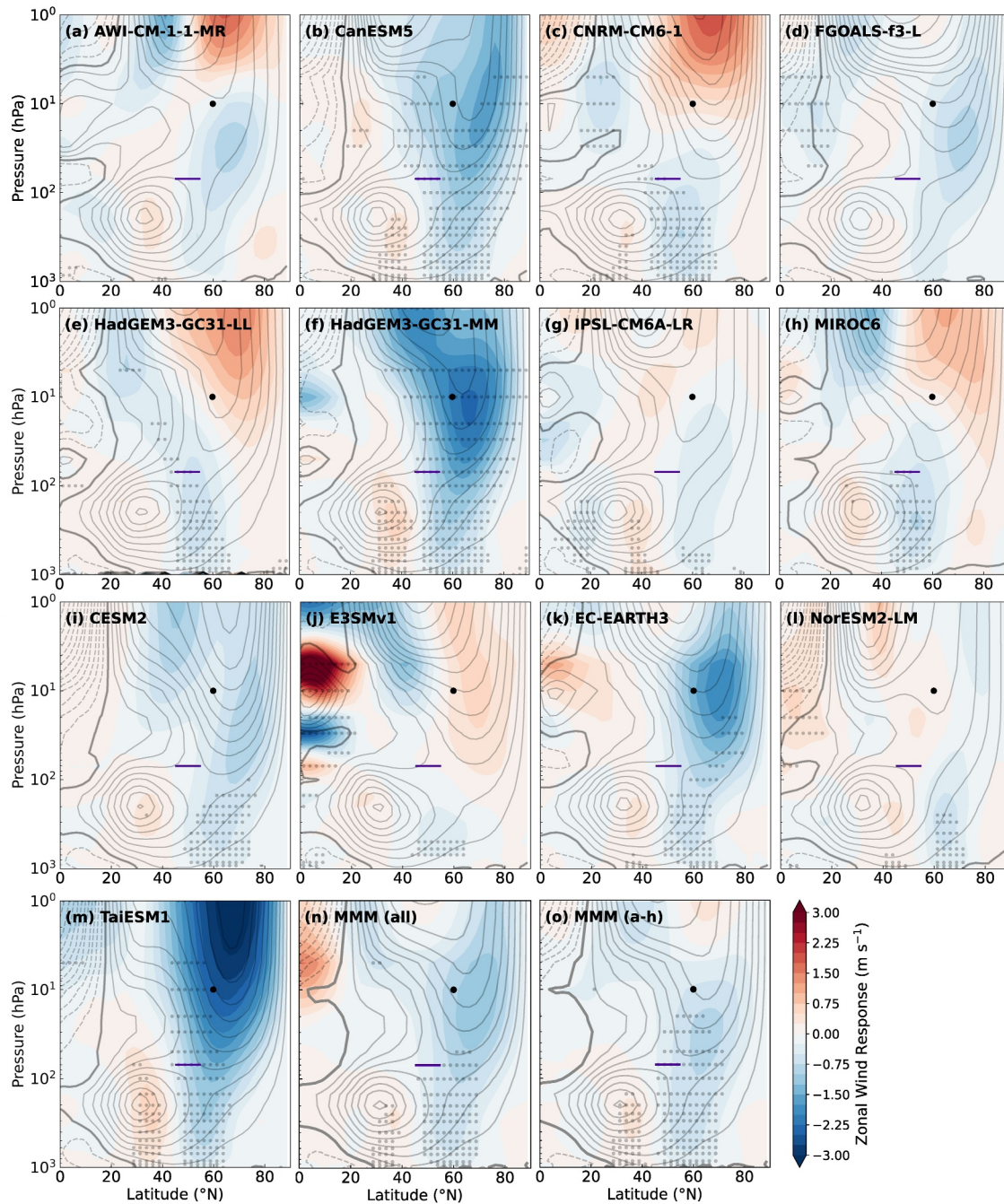


Figure 1. Zonal, DJF and ensemble-mean zonal wind response (colors) to Arctic sea-ice loss per model, overlaid with climatological (pdSST–pdSIC) contours, including the zero wind line (thick contour). The stippling indicates where the response is statistically significant, the purple line the neck region at 70 hPa, 45–55°N, and the black dot 10 hPa, 60°N. Panels (n) and (o) show the multi-model mean (MMM) for all models and the subset with daily data, respectively, with stippling where the response is significant *and* ≥ 10 of the 13, or ≥ 6 of the 8, models agree on the sign. All the models' data has been latitudinally regridded to 2.5° to support comparison.

Relatedly, there is no consistent sign of change in mean SSW frequency across this subset (Figure 3a), nor in SSW timings throughout extended winter (not shown). HadGEM3-GC31-MM is the only model with a significant response: an increase in SSWs in response to future Arctic change. The average timing of FWs also does not significantly change in any of the models, with differences between the two experiments' median (and mean) FW dates a matter of days (Figure 3b). Despite small changes in the average FW date, some models' overall FW date distributions do clearly change. For example, the range of FW dates contracts for CNRM-CM6-1 and HadGEM3-

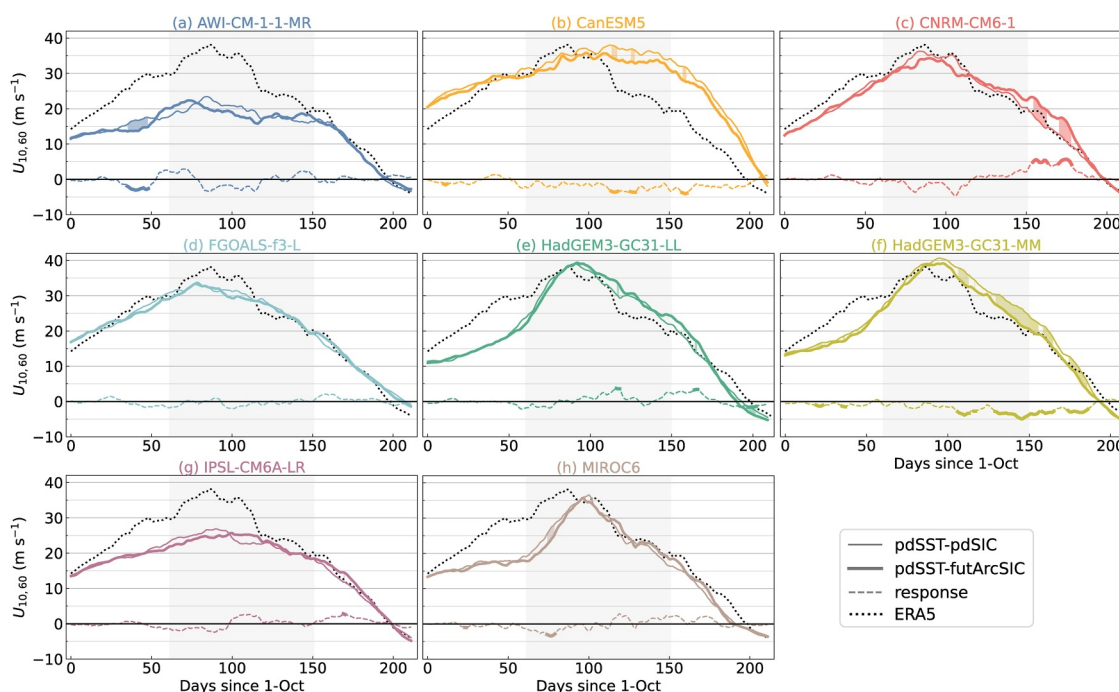


Figure 2. Daily evolution of $U_{10,60}$ from October to April (DJF shaded in gray) in the two experiments, as well as the response (dashed line), for each of the eight models with daily data. Fill between pdSST-pdSIC and pdSST-futArcSIC lines and a bold response line indicates periods with a statistically significant change. The evolution of ERA5 $U_{10,60}$ is marked on each by the dotted line.

GC31-LL, becoming slightly more evenly spread in the former. Whereas the range of FW dates extends for HadGEM3-GC31-MM and IPSL-CM6A-LR, with indication of dates as early as February in the latter. Though CMIP6 model simulations of abrupt4 $\times$ CO₂ showed a robust change in FW dates versus picontrol (Ayarzagüena et al., 2020), the PAMIP protocol uses a much weaker, and ultimately different type of, forcing. Nevertheless, given that Arctic sea-ice loss is a key component in climate change generally, we believe our findings support previous conclusions that sea-ice loss contributes to the simulated spread in CO₂-driven projections (Kretschmer et al., 2020).

Finally, most models' polar vortices do not significantly shift (or stretch) in response to Arctic sea-ice loss, contrary to the robust shift seen in models under greenhouse gas forcings (Karpechko et al., 2022). Here, the most notable changes in vortex location seem to be the statistically significant shifts equatorward in CanESM5 and HadGEM3-GC31-MM's mean centroid latitudes (Figure 4b). Although CMIP6 experiments, ranging from 4 $\times$ CO₂ to the mildest SSP scenarios, show a consistently eastward shift in average vortex longitude (Karpechko et al., 2022), most of the PAMIP-contributing models show a small westward shift (8 out of the 13 models, Figure 4c). As for FWs, this suggests that Arctic sea-ice loss is unlikely to be the main driver of the responses to general greenhouse gas forcings but could contribute to the spread. Meanwhile an almost even split of models' mean aspect ratios become marginally smaller/larger, with little qualitative change in distributions across ensemble members (Figure 4a). It is worth noting that there are large mean state biases here, with all the models having a vortex that is on average too poleward and too circular compared to ERA5 (dotted black lines in Figure 4), similar to results found for historical CMIP6 simulations (Hall et al., 2021).

4. Exploring Potential Causes of Model Uncertainty

The above results paint a picture of an overall weak yet diverse response to future Arctic sea-ice loss across the PAMIP-contributing models. Now we explore what might drive differences between them, including why HadGEM3-GC31-MM might be the only one with a consistently statistically significant stratospheric response for a number of measures, by looking at factors including ensemble size, resolution, and simulated basic state.

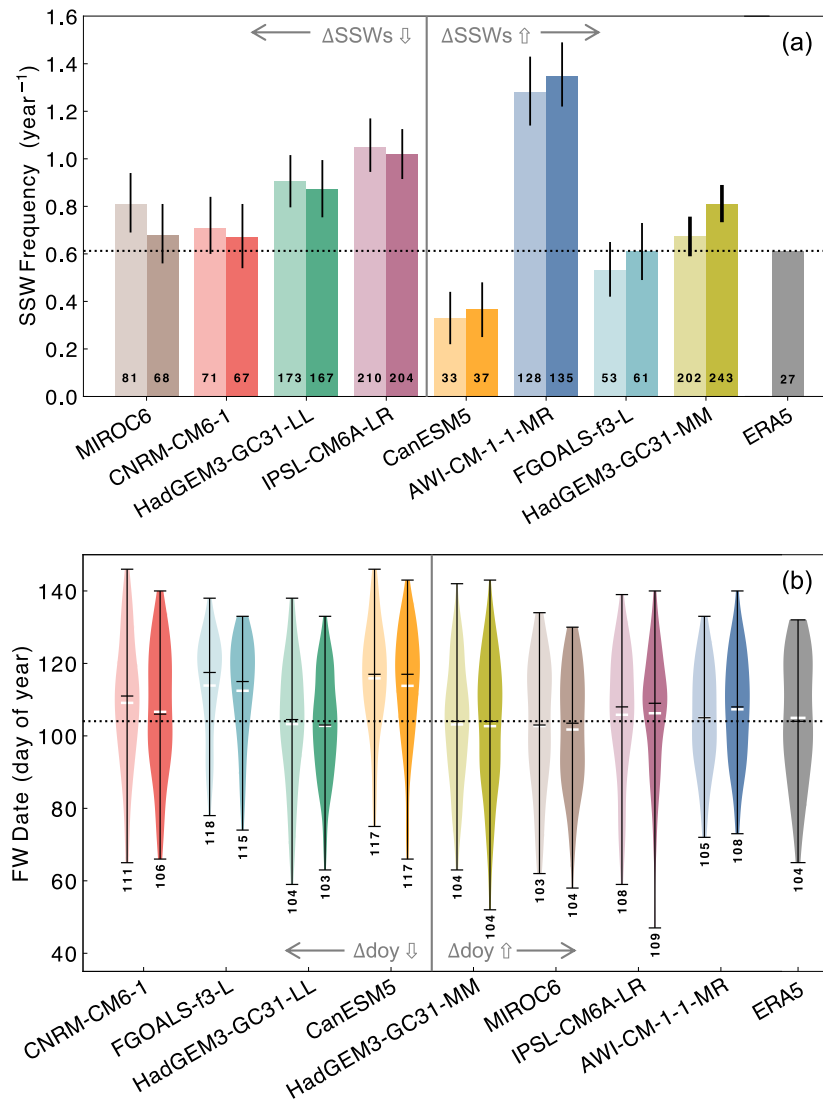


Figure 3. For the eight models with daily data, (a) mean SSW frequency in pdSST-pdSIC (light) and pdSST-futArcSIC (dark) per model. The absolute count is at the bottom of each bar and the black lines mark the 97.5% confidence levels (thick lines for model(s) whose means are significantly different). Bars are ordered L-R by the size of the difference between the two experiments' mean SSW frequency, that is, the biggest decrease to increase. (b) FW timing per experiment (colors as in a) per model, with the near-central black bar marking the median, upper/lower the extrema, and white the mean. Median FW day of the year is below each violin, where 120 corresponds to ~30 April. Violins are ordered L-R by the size of difference between the two experiments' median FW dates, that is, largest shift earlier to later. The ERA5-derived SSW frequency and median FW date are marked with the horizontal dotted line.

4.1. Role of Ensemble Size

The models analyzed here have a range of ensemble sizes, from 100 to 300 members (Table 1). To assess the robustness of models' stratospheric responses, and the dependency of that response on ensemble size, we perform a bootstrapping test similar to that described in Section 2.2. In this case, we find stratospheric wind responses ($\Delta U_{10,60}$) and conduct the bootstrapping process for ensemble sizes $n = 1$ up to N , the total number of ensemble members in each model. Given the sign of the ensemble-mean $U_{10,60}$ response (i.e., strengthening or weakening), for each of the sub-ensembles, p^n is the probability across the 1,000 bootstrapping resamples that the mean change with n members is different to that with N :

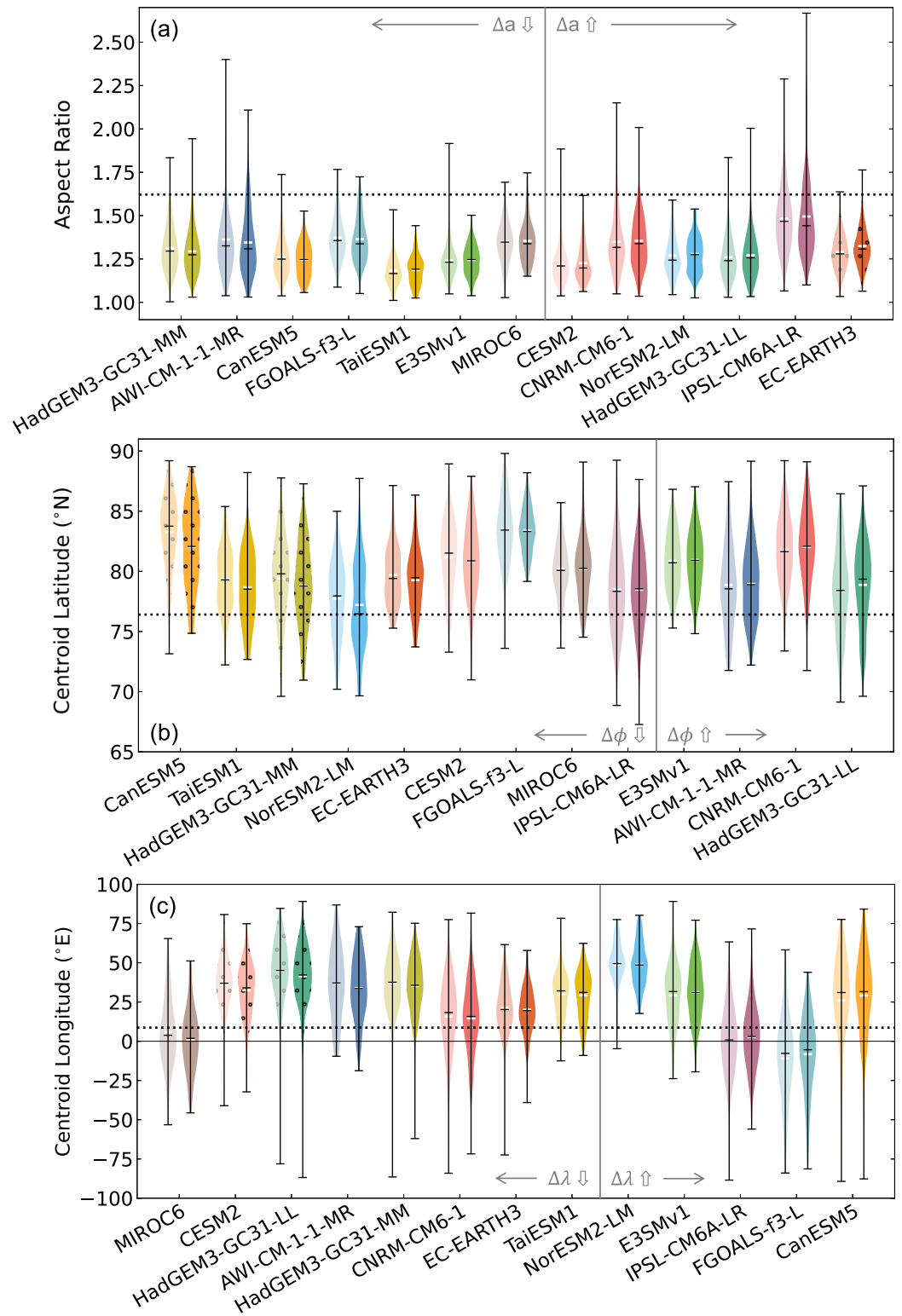


Figure 4. As in Figure 3b but for vortex (a) aspect ratio, a , (b) centroid latitude, ϕ , and (c) centroid longitude, λ , and for all models (using monthly data), with the ERA5-derived values marked by the dotted lines. Violins are ordered L-R by change in median values; that is, models whose vortices become more to less circular, more equator- to poleward, and more west- to eastward, respectively. The spotty pattern indicates models whose change in the mean between experiments is significant.

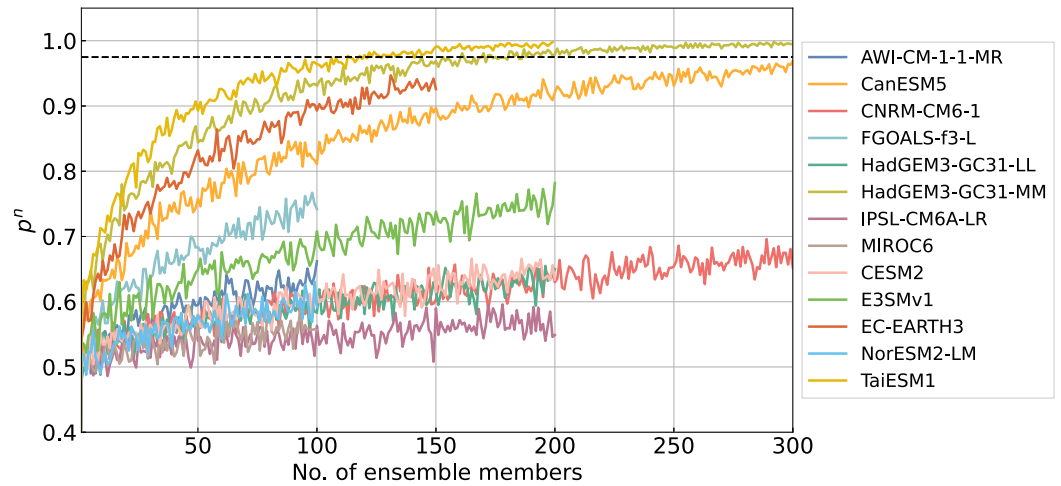


Figure 5. p^n (Equation 2) calculated using $U_{10,60}$ for resampled sub-ensembles up to each model's total ensemble size. The black dashed line marks the 97.5% level.

$$p^n = \begin{cases} p(\Delta U_{10,60}^n > 0), & \text{if } \Delta U_{10,60}^N > 0 \\ p(\Delta U_{10,60}^n < 0), & \text{if } \Delta U_{10,60}^N < 0. \end{cases} \quad (2)$$

This gives an indication of the size of the ensemble that could be sufficient to achieve a statistically significant response of any sign in each model's $U_{10,60}$.

Some studies propose that the lack of robust response may, along with model uncertainty, be due to internal variability obscuring both near-surface and stratospheric responses. These studies find that even the minimum of 100 ensemble members in PAMIP is insufficient to capture the atmospheric response, and in particular robustly identify the stratospheric pathway in the NAO response (Liang et al., 2024; Peings et al., 2021; Sigmond & Sun, 2024). We reach a similar conclusion for the stratospheric response itself: for most of the models analyzed here, even 300 members may not be enough. Figure 5 shows that only HadGEM3-GC31-MM and TaiESM1 achieve a statistically significant $U_{10,60}$ response, emerging around $n = 200$ and 150 , respectively. We do however note that, as in Table 1, TaiESM1 has less than half the number of ensemble members available for pdSST-futArcSIC as in pdSST-pdSIC. The emergence of significance does vary when using the “mismatched” ensembles, as above, compared to when using only the same 91 members in both—in the latter case, significance emerges even before $n = 100$. Nevertheless, we take this result to mean that there is something else intrinsic to these models' stratospheric response to sea-ice forcing relative to internal variability—and that HadGEM3-GC31-MM's significant response cannot solely be due to its large ensemble.

4.2. Role of Resolution

Although horizontal resolution has been suggested to affect responses to Arctic sea-ice loss (Steffing et al., 2021), the evidence for its influence, especially on the stratospheric response, is inconclusive. For example, Figure 2 of Smith et al. (2022) shows that the higher horizontal resolution version of the OpenIFS model (OpenIFS-511) has a larger deceleration relative to the lower resolution version (OpenIFS-159) in response to Arctic sea-ice loss (Steffing et al., 2021). The version with more vertical levels (OpenIFS-1279, 137 levels; 60 in the stratosphere where $p < 100$ hPa) also shows a vortex strengthening compared to weakening in that with the same horizontal resolution but $\sim 30\%$ fewer levels (OpenIFS-511, 91 levels; 38 in the stratosphere). We do not analyze the OpenIFS simulations here, due to data availability, but we can look across the models we do have, which span a range of resolutions (Table 1).

Across the models, we do not find any clear relationship between vortex response and horizontal resolution (Figure 6a). The three models with the largest $U_{10,60}$ weakening have resolutions of 1° or finer, but CanESM5,

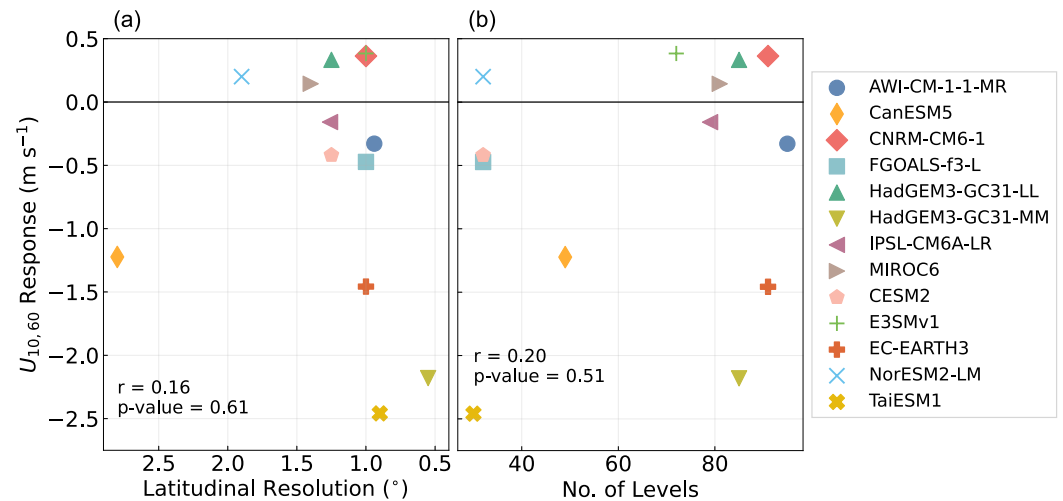


Figure 6. $U_{10,60}$ response, calculated from monthly data, versus (a) horizontal and (b) vertical model resolutions, with Pearson r and associated p -value shown for relationships across all models.

with the next largest weakening, has the coarsest resolution of all the models analyzed, making it a notable outlier. Interestingly, our model ensemble includes two versions of the HadGEM3 model whose predominant difference is their horizontal resolution (Andrews et al., 2020). Comparing their $pdSST$ – $pdSIC$ zonal winds indicates that the higher resolution version has a statistically significantly stronger vortex, with a similarly significant difference between their stratospheric responses (Figure 7). But unlike with OpenIFS, the stratospheric response is of opposite sign between the two versions of HadGEM3 and, unfortunately, without data from more models run at multiple different resolutions, it is difficult to conclude whether higher horizontal resolution is synonymous with a stronger vortex and thus results in (greater) deceleration in response to Arctic sea-ice loss. Furthermore, we might expect vertical resolution to be particularly important for resolving stratospheric processes. Yet we find that models with a high number of levels and model top are not necessarily more similar to ERA5 (see e.g., HadGEM3-GC31-MM vs. AWI-CM-1-1-MR in Figure 8 and Hall et al., 2021), nor are those with a low number of levels and model top any less likely to simulate vortex weakening (e.g., CanESM5 and TaiESM1 in Figure 6b).

Previous work, including as part of PAMIP, had hinted at a role for resolution in the diversity of simulated stratospheric responses, but only two PAMIP-contributing models were specifically run with different resolutions. Though both those models display significant differences in their simulated responses for different resolutions, the overall relationship across the multi-model ensemble is not straightforward: we find no obvious resolution dependency across our 13 models (Figure 6). There are too many additional model differences to be able to clearly identify the role of model resolution in the stratospheric response; to do so would require targeted simulations with more models that are beyond the scope of this study. Nevertheless, it is an important question to try to address, especially given the often-discussed computational cost savings that could be made by running with lower resolution in favor of more ensemble members.

4.3. Role of the Basic State

Previous work (discussed in Section 1) highlights the importance of models' simulated basic state in driving differences in models' forced responses. Figure 7 shows that although only a few of the models analyzed here have significantly different $U_{10,60}$ responses, most of them have significantly different climatologies. Furthermore, across the eight models with daily data, a weaker polar vortex is typically a more variable one, in terms of SSW frequency (Figure 9a). For a model such as AWI-CM-1-1-MR, very weak climatological $U_{10,60}$ may indicate a simulated vortex state that is simply more susceptible to anomalous wave activity-driven wind reversals (Figures 2a and 3a). Whereas for a model such as CanESM5, the very strong $U_{10,60}$ may instead limit the ability of waves to reverse vortex westerlies, with much fewer SSWs (Figures 2b and 3a).

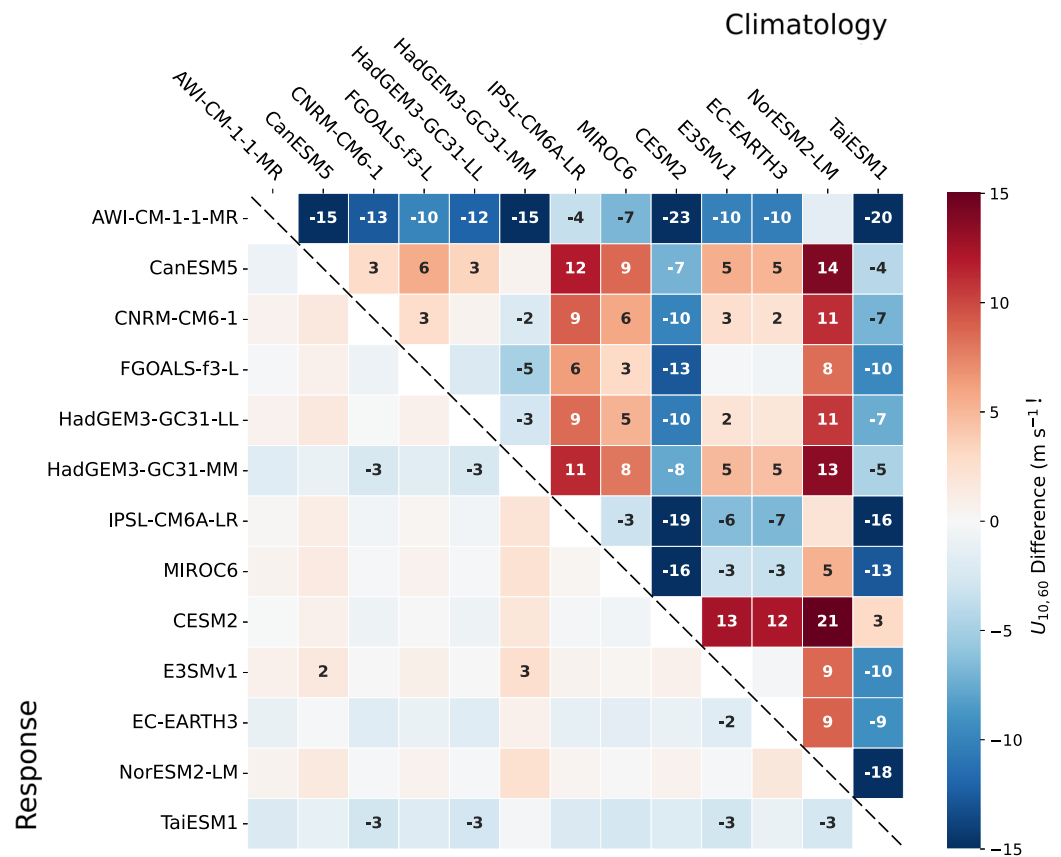


Figure 7. The difference between each models' pdSST-pdSIC $U_{10,60}$ (climatology, above diagonal line) and response (below line). The colors indicate the size of difference and numbers show that difference only where it is statistically significant (i.e., the p -value from a two-tail t -test is ≤ 0.025). The differences are calculated as row minus column; for example, the top right box indicates that AWI-CM-1-1-MR has a 20 m s⁻¹ weaker $U_{10,60}$ than TaiESM1.

With a focus on the models with daily data, it then seems that, in response to Arctic change, the stronger the vortex in pdSST-pdSIC, the larger its deceleration and associated increase in SSW frequency (Figure 9c), corroborating previous experiments with an idealized model (Mudhar et al., 2024). Indeed, in our eight-model subset, the change in SSW frequency is well-correlated with the response in both the mean strength and variability of the vortex, which themselves are strongly negatively correlated (Figure 9b). The change in SSW frequency tracking vortex strength is unsurprising, but there is also some clear dependency of the mean vortex response on its own basic state.

Factors that influence the stratospheric state are typically ones that modify wave propagation. For example, phenomena such as ENSO or QBO (see Section 4.4), or factors that result from model differences. One such example is differences in gravity wave parametrizations that impact simulated vortex strength and/or neck winds (Sigmond & Sun, 2024). We use the wave-1 refractive index squared ($n_{k=1}^2$) to infer a waveguide for planetary waves (following Mudhar et al., 2024). Within the waveguide, this neck region between the subtropical jet and stratospheric polar vortex has previously been highlighted as being particularly key: it is the region through which tropospheric wave activity can travel to ultimately be able to interact with the stratospheric vortex. As the $n_{k=1}^2$ calculation partly depends on zonal winds, U_{neck} can be used as a proxy for the waveguide there; stronger U_{neck} typically means more positive $n_{k=1}^2$, implying a better waveguide in the neck. The optimal waveguide enables upward-propagating waves to reach and break in the mid-stratosphere, resulting in a weaker, more variable vortex, with relatively high incidence of SSWs. This requires sufficient stratosphere-troposphere coupling, in turn characterized by sufficiently strong neck and polar vortex winds: this is a state in which the stratosphere is not isolated from the troposphere due to winds being too weak (no tropospheric interaction at all) nor too strong

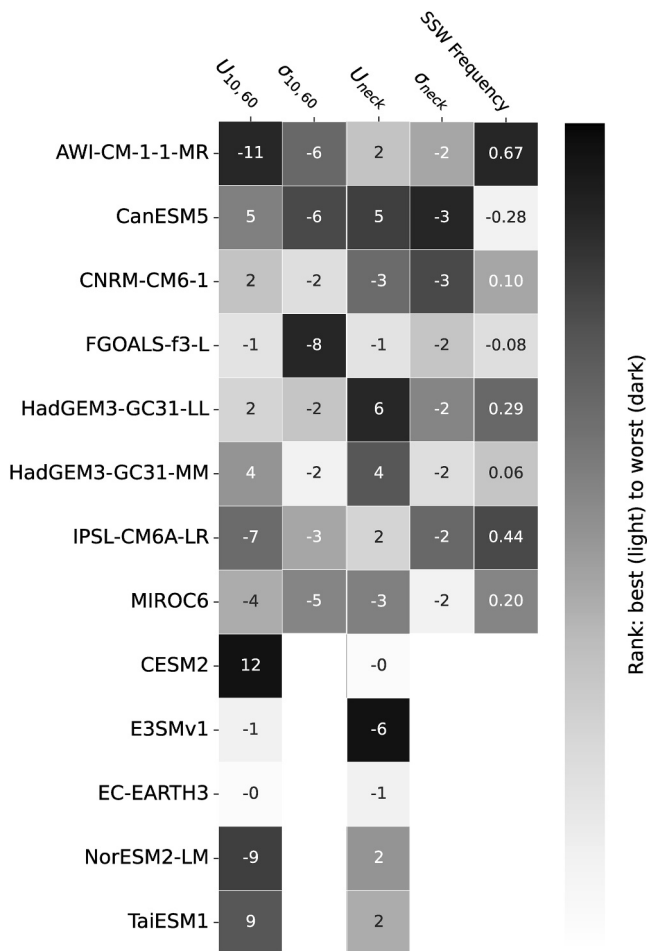


Figure 8. Difference between models' and ERA5-derived values for different diagnostics: $U_{10,60}$ and U_{neck} are calculated from monthly data and the rest from daily, where available. The numbers are the size of the difference (U and σ in m s^{-1} , SSW frequency in year^{-1}) and the colors indicate the ranking of models from best (light gray) to worst (black) based on absolute error; for example, the top left box indicates that AWI-CM-1-1-MR's $U_{10,60}$ is 11 m s^{-1} weaker than in ERA5, the second biggest mean bias of all the models for $U_{10,60}$.

(reflective of tropospheric wave activity), and is something that can be highly model-dependent. Here, climatological $n_{k=1}^2$ varies across the models, but is overall positive from around the midlatitude mid-troposphere into the polar stratosphere (not shown), thus theoretically conducive to upward wave propagation. Across the models with daily data, there is a positive correlation for climatological U_{neck} versus $U_{10,60}$ and $\sigma_{10,60}$ (Figure 9a), but it is weak. There is essentially no relationship between climatological U_{neck} versus $U_{10,60}$ across the full ensemble. This suggests that there are differences beyond just neck winds that cause some of the models to have a better-connected stratosphere and troposphere than others.

Despite differences in climatology, there is some agreement on the response of refractive index across the PAMIP-contributing models. Previous studies have highlighted U_{neck} as a control on the ability of waves to reach the stratosphere, particularly in response to Arctic sea-ice loss (Albers & Birner, 2014; Mudhar et al., 2024; Sigmond & Sun, 2024). We find that the neck region $n_{k=1}^2$ predominantly becomes more positive in response to Arctic sea-ice loss (Figure 10). The notable exceptions are E3SMv1, HadGEM3-GC31-LL, and IPSL-CM6A-LR, which for them suggests a restriction of waves being able to propagate into and weaken the vortex; the first two do show a clear mean vortex strengthening but the latter does not (Figure 1). Most of the other models show a degradation of the waveguide poleward of the tropospheric jet (browns around $40\text{--}60^\circ\text{N}$, $100\text{--}200 \text{ hPa}$ in Figure 10), just below an enhancement of $n_{k=1}^2$ that extends well into the stratosphere — though changes are small. Consistent with their zonal wind responses (Figure 1), CanESM5 and HadGEM3-GC31-MM show statistically significant enhancements of refractive index from the neck region into the mid-stratosphere; the strong enhancement in TaiESM1 is more on the equatorward side of the vortex. Interestingly, models with $U_{10,60}$ strengthening in response to the sea-ice forcing tend to have weak neck winds in pdsST–pdsIC, with the exception of HadGEM3-GC31-LL. Although for models with $U_{10,60}$ weakening, those with the strongest neck winds in pdsST–pdsIC also show both the largest vortex and neck wind deceleration, such as CanESM5, HadGEM3-GC31-MM, and TaiESM1. This implies that these models have an improved waveguide that allows anomalous wave activity, including that additionally generated by the Arctic sea-ice loss, to propagate up into, then break within, the vortex, manifested as a higher SSW frequency and weaker vortex.

Overall, we find a negative correlation between climatological $U_{10,60}$ and its response (Figure 9c for models with daily data, correlation is slightly weaker at $r = -0.48$ across all 13): the stronger the vortex climatologically, the greater its deceleration in response to Arctic sea-ice loss. We also find that the models with daily data whose polar vortices decelerate most tend to have the largest increase in SSW frequency (Figure 9b). The change in SSW frequency also corresponds well to the change in neck wind variability (σ_{neck}), which itself is highly correlated to the change in vortex variability ($\sigma_{10,60}$, Figure 9b). This suggests that models whose winds in the lower- to mid-stratosphere become more variable in response to Arctic sea-ice loss similarly see an increase in SSW frequency, and vice versa for a decrease, consistent with changes in the waveguide.

4.4. Role of the QBO

The quasi-biennial oscillation (QBO) in the tropical lower stratosphere is another factor that can directly influence the stratospheric state. The easterly phase (QBO-E) is typically associated with a weaker, more disturbed, polar vortex than westerly (QBO-W, Holton & Tan, 1980). Labe et al. (2019) have further proposed that under the influence of Arctic warming, the weaker vortex typical of QBO-E may weaken more than that of QBO-W. Walsh et al. (2025) discuss the potential mechanisms behind the contrasting vortex responses to sea-ice loss in the QBO-

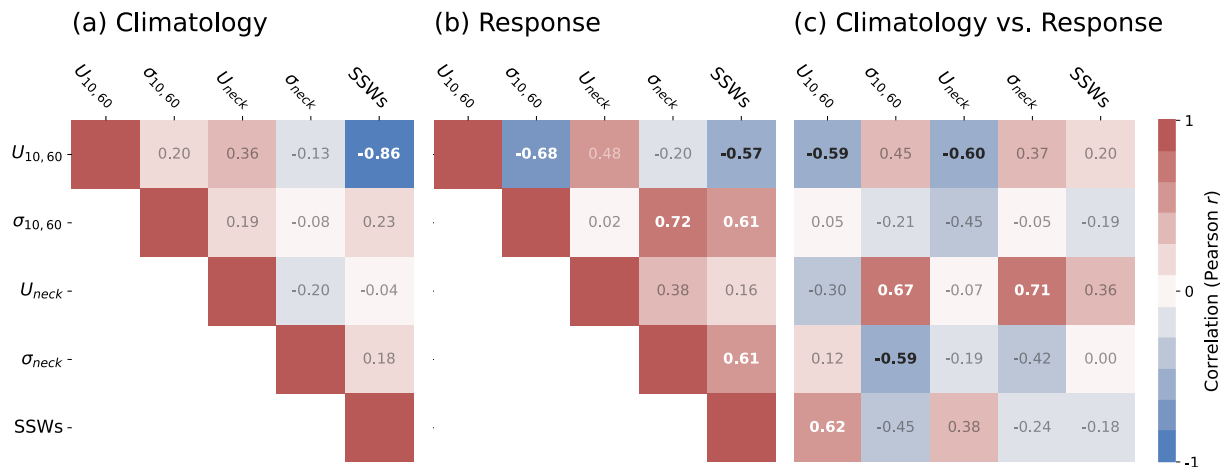


Figure 9. Pearson correlations between different daily zonal wind-derived diagnostics' (a) climatological (pdSST–pdSIC) values and (b) response, and (c) the relationship between diagnostics' response (rows) with climatology (columns), all for the eight models with daily data only. The colors and numbers indicate the sign and strength of the correlation, with $|r| \geq 0.5$ bolded. In (c), the top right value of +0.2 indicates a weak positive correlation between $U_{10,60}$ response and climatological SSW frequency, whereas the bottom left value of +0.62 indicates a strong positive correlation between SSW frequency response and climatological $U_{10,60}$.

E and W simulations with HadGEM3-GC31-MM. They identify constructive interference with climatological wave-1 geopotential height in the upper troposphere/lower stratosphere in QBO-E, and destructive interference in QBO-W. They also highlight increased upward wave activity corresponding to regions of enhanced wave-1 refractive index in QBO-E, resulting in notable wave convergence in the polar stratosphere that coincides with the significant deceleration there. This mechanistic description, for just one of the PAMIP-contributing models, corroborates previous work by Labe et al. (2019) with another model. To build on this, we have generalized their investigations to a larger sample of the PAMIP multi-model ensemble.

Among the 13 models analyzed here, CanESM5, CESM2, FGOALS-f3-L, NorESM2-LM, and TaiESM1 have previously been identified as models that are unable to simulate the QBO (Lee et al., 2020; Richter et al., 2020). Given the persistence of the initial conditions during these 14-month (approximately half the QBO period) time-slice simulations, we view even the models that do not have an internally generated QBO as sampling different QBO phases. Within PAMIP, the two HadGEM3 models were specifically initialized to have an approximately equal number of ensemble members in both phases in the winter, however IPSL-CM6A-LR also has members in all states (Figure 11). In the October–November period used to classify the QBO phase (Equation 1), we find that eight of the models are predominantly in QBO-E (Figure 12a) and only three have >20% of members in QBO-W (HadGEM3-GC31-LL/MM and IPSL-CM6A-LR, Figure 12a). Across the models with daily data, there are a comparable proportion of members in QBO-W as in a neutral state, but the most overall are in QBO-E (true using both monthly and daily data sets).

Averaging across all models' members (MMM) reveals marginal differences in the climatological $U_{10,60}$ between QBO phases (Figure 12c). $U_{10,60}$ appears slightly stronger in QBO-W than QBO-E (and neutral), though the inter-model spread is large. The $U_{10,60}$ response is then especially interesting. It seems that in spite of the inter-model spread in climatological $U_{10,60}$, there is clear agreement between models, and thus in the MMM, in terms of the sign of the $U_{10,60}$ response when splitting by QBO phase: $U_{10,60}$ weakens under QBO-E and strengthens under QBO-W (Figure 12d). This split of response by phase is particularly seen in HadGEM3-GC31-LL, HadGEM3-GC31-MM and IPSL-CM6A-LR, which have members in both phases. Though this MMM picture may seem contrary to the results not split by phase previously, in which stronger climatological polar vortices decelerated more, it is consistent with the expectation that a weaker polar vortex under QBO-E should weaken further with Arctic forcing (Labe et al., 2019). Furthermore, mean SSW frequency increases in response to sea-ice loss under QBO-E, but decreases under QBO-W (Figure 12b) for those models with daily data. Though neither change is statistically significant, the sign of change is consistent with our previous results, whereby a sea-ice loss-driven weakening of $U_{10,60}$ is accompanied by an increase in SSW frequency (i.e., QBO-E members) and vice versa for a strengthening of $U_{10,60}$ (QBO-W members).

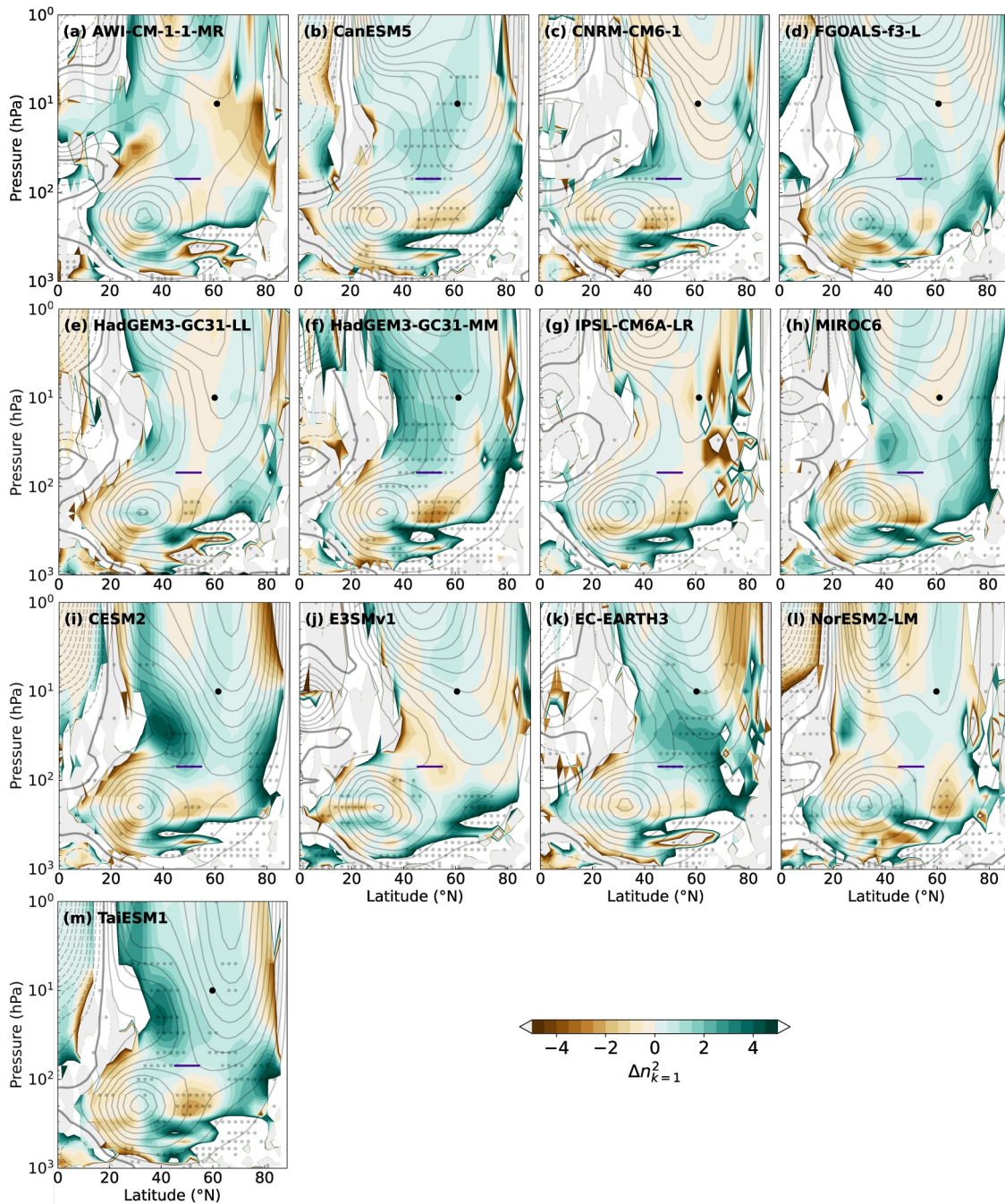


Figure 10. As in Figure 1 but for wave-1 refractive index squared, $\Delta n_{k=1}^2$ and without the multi-model mean.

Compared to previous studies such as Labe et al. (2019) or Walsh et al. (2025), identifying a clear role for QBO phase in the stratospheric response to Arctic sea-ice loss is complicated by the different modeling approaches across the PAMIP-contributing models. We suggest that the QBO modulation of the vortex response is not solely about the mean polar vortex state (the MMM values across phases are not overly dissimilar in Figure 12c), but likely modulates the response through its relation to other changes in the basic state, similar to Walsh et al. (2025). Crucially, this model ensemble is overall biased toward the QBO-E phase (Figure 12a), which suggests a biasing of the MMM response toward a vortex weakening (as in Figure 1n) and SSW frequency increase for the models with daily data. To that end, the importance of sampling the different phases within models, and not just across the model spread, is clear: QBO phase can affect the vortex response, possibly even so far as the sign. We also suggest

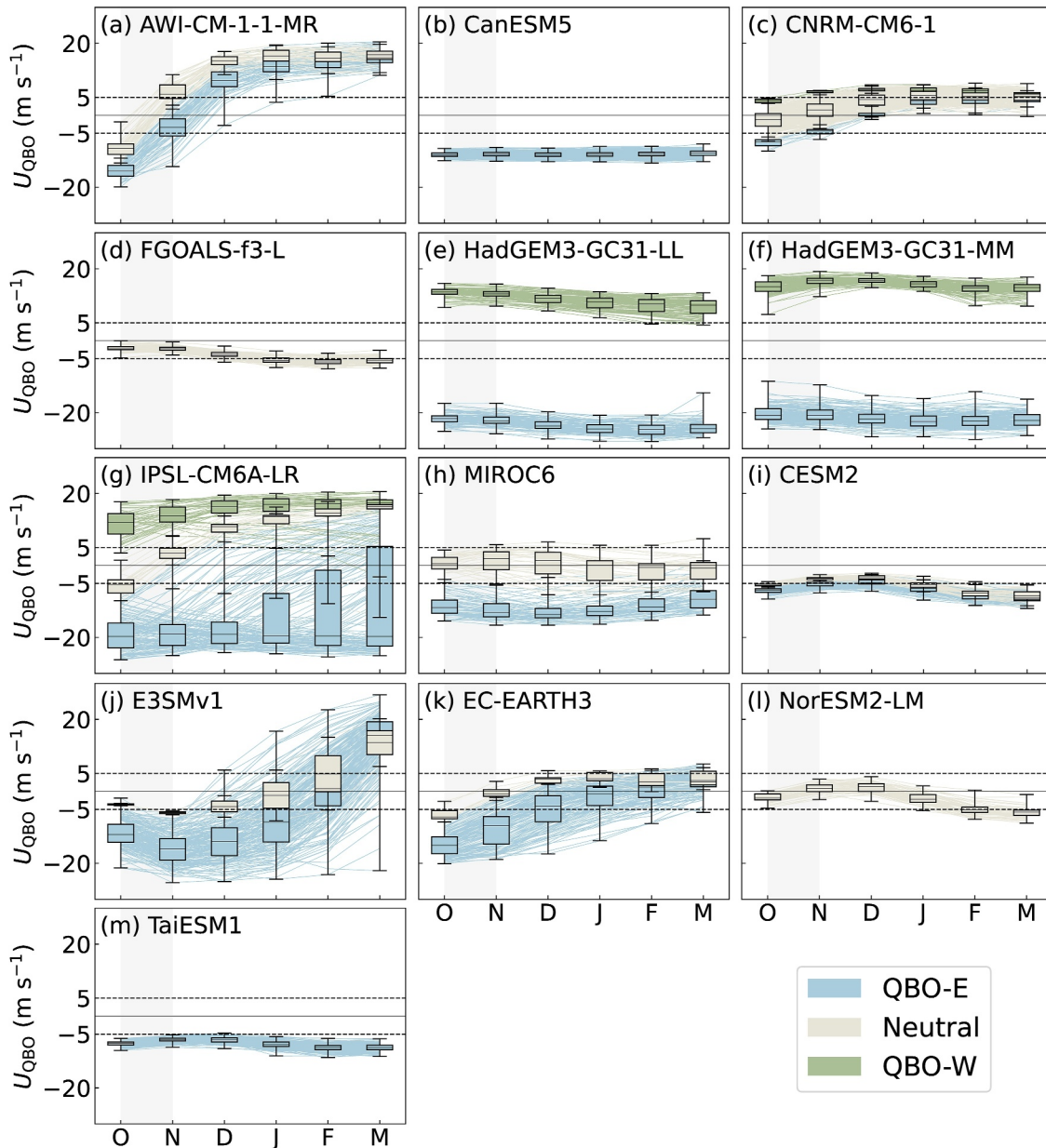


Figure 11. Monthly evolution of the pdSST–pdSIC simulation of U_{QBO} per model, with members split into QBO phases depending on winds' sign and strength (black dashed horizontal lines mark $\pm 5 \text{ m s}^{-1}$) in October–November (gray-shaded region, diagnostic description in Section 2.2). The colored lines show evolution of individual members in each phase, with the horizontal lines on the box-and-whiskers marking the median and extrema of the distribution.

taking particular care when initializing the models. For example, not choosing initial conditions from a point in a run where the QBO is too weak or perhaps in a transitional phase. Furthermore, evolution of AWI-CM-1-1-MR and IPSL-CM61-LR U_{QBO} (Figures 11a and 11g) suggest it may even be worth using nudging to maintain conditions throughout a simulation (e.g., Butchart et al., 2018). Finally, we note that HadGEM3-GC31-MM under QBO-E has by far the largest magnitude of $U_{10,60}$ change (disregarding the few CNRM-CM6-1 members), whereas the response of its members under QBO-W are comparable to that of other models, suggesting that the robust response of HadGEM3-GC31-MM when not split by QBO phase appears to come only from its QBO-E members.

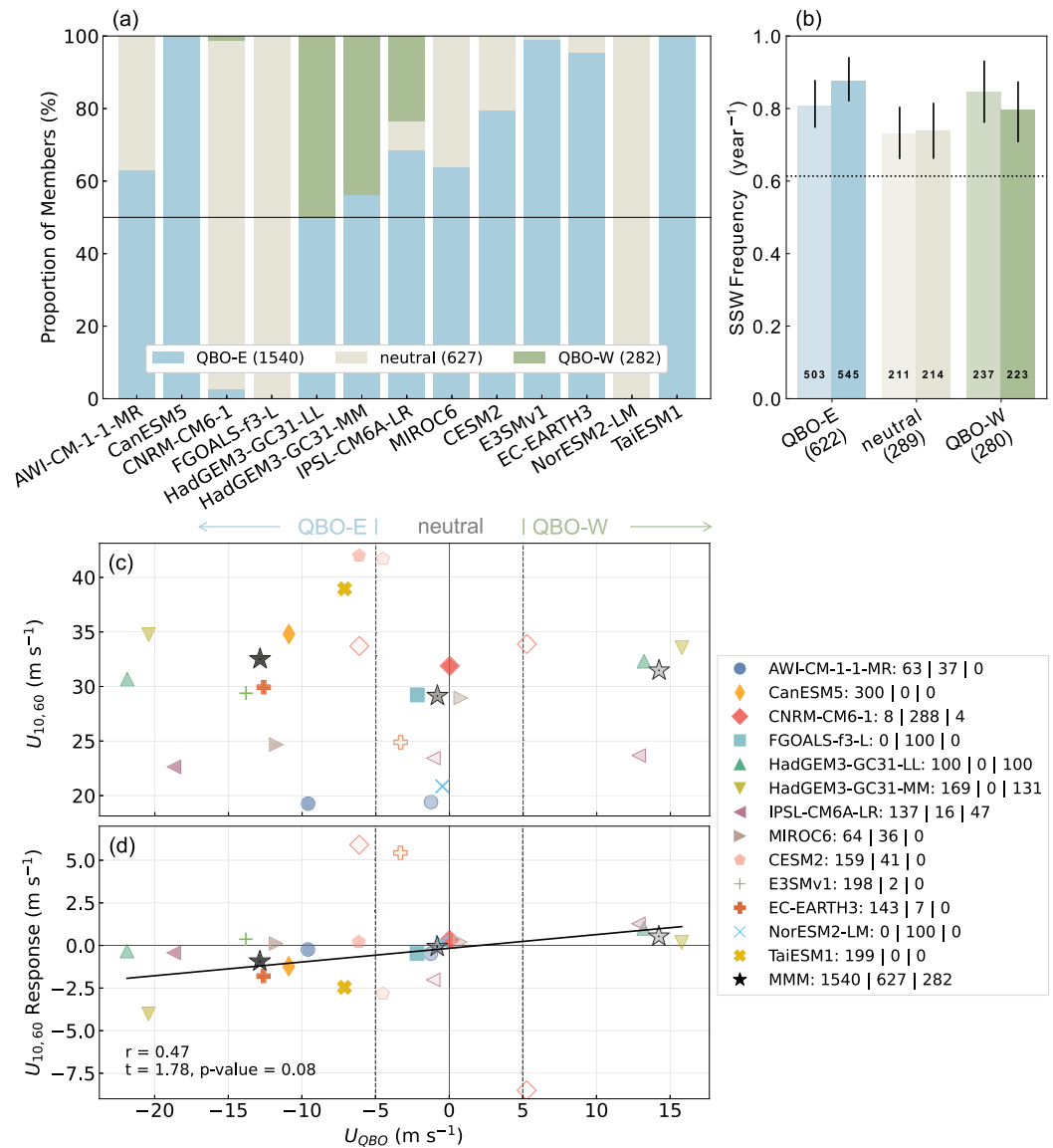


Figure 12. (a) The proportion of ensemble members in each QBO phase per model using monthly data; proportions for the first eight models using daily data are comparable. The black line marks 50% and the key shows the total number of members in each phase in brackets. (b) SSW frequency as in Figure 3a (where the lighter and darker colors are $pdSST-pdSIC$ and $pdSST-futArcSIC$, respectively), but calculated from a composite of the eight models' ensemble members in each QBO phase (total number of members in brackets). Then U_{QBO} versus (c) $pdSST-pdSIC$ $U_{10,60}$ and (d) its response using monthly-mean data split by QBO phase, including a weighted line of best fit and r across all models, and two-tail t -test with associated p -value to indicate the significance of the difference between the multi-model mean (MMM, star) vortex responses of QBO-E and W members. The more opaque the symbol the greater the proportion of members in a given phase: the key shows numbers of members in QBO-E | neutral | QBO-W per model.

5. Discussion and Conclusions

The response to a single forcing, such as Arctic sea-ice loss, cannot be directly observed. Nevertheless, it is useful to study responses to individual forcings in order to understand how they may contribute to the overall trend of global climate change. For the response to future Arctic sea-ice loss, previous studies have utilized emergent constraints from model simulations to estimate the real world response (e.g., Karpechko et al., 2024; Smith et al., 2022). Our analysis does not unfortunately present a compelling argument for such a constraint for the stratosphere. To gain an insight into a possible “true” response, then, one could choose to jointly assess how well a

model simulates the present and how robustly it simulates the future. In that case, we would be particularly interested in models that both (most) accurately represent the world as we know it, and have a statistically significant response to the additional forcing. For the former, most of the models simulate a wintertime vortex (and neck winds), that are biased too strong compared to ERA5 (Figure 8), in turn affecting timings and magnitude of the stratospheric evolution over extended winter (Figure 2); around half of the models with daily data have average FW dates that are too late and most of those eight models overestimate the frequency of SSWs (Figure 8). Altogether, HadGEM3-GC31-MM is one of the least-biased models: its DJF-mean $\text{pdSST}-\text{pdSIC } U_{\text{neck}}$ and $U_{10,60}$ are slightly stronger than in ERA5 (Figure 8), but its extended winter evolution of the latter, median FW date, and mean SSW frequency are very similar (Figures 2 and 3). However, we cannot know which aspects of the present climate are important for the response, so a “better” simulation of the aforementioned quantities may not necessarily lead to a more realistic response.

When attempting to understand that “realistic” response, it is worth noting that this study comes up against a number of limitations related to model resolution, parametrizations, and the use of atmosphere-only versus coupled models. First, we use atmosphere-only rather than coupled atmosphere-ocean models; yet, it has been suggested that atmosphere-only models underestimate the response to sea-ice loss compared to that in coupled models (Deser et al., 2015; Labe et al., 2020). By using fixed SSTs in these atmospheric models, simulated variability is likely lower than, and thus not directly comparable to, that of the real world. As such, we cannot necessarily expect a direct relation between models' responses and model bias: we certainly see that models with mean SSW frequency or FW date close to ERA5 are overall not consistent in the sign of their responses (Figure 3). It also means that we do not consider the proposed dependency of stratosphere–troposphere coupling on the background oceanic state such as the El Niño–Southern Oscillation (ENSO; Shen et al., 2024, 2025; Walsh et al., 2025), Atlantic multidecadal variability (Liang et al., 2024), interdecadal Pacific variability (Hu & Guan, 2018; Simon et al., 2022), or warming SSTs generally (Hu et al., 2018; Yu et al., 2024). Although isolating the impact of sea-ice loss in coupled simulations is technically challenging (see e.g., England et al., 2022; Lewis et al., 2024), future work could compare the stratospheric responses between atmosphere-only and coupled models to quantify the modulating effects of oceanic feedback (e.g., SST anomalies) on the Arctic sea-ice–stratospheric polar vortex relationship, thereby providing a more targeted research outlook. Next, our preliminary analysis in Section 4.2, corroborating similar previous work, suggests that resolution could affect the strength of response but a lack of data means we are unable to really explore its relative role. And finally, previous work suggests that small or no stratospheric polar vortex change could result from canceling effects of sea-ice loss in the Atlantic versus Pacific sectors (McKenna et al., 2018; Sun et al., 2015; Xu et al., 2023). One study utilizing PAMIP's regional sea-ice loss experiments, run with just a few models, found a particularly interesting result for HadGEM3-GC31-MM: there was a large vortex weakening in response to Barents–Kara Sea sea-ice loss but no such compensating strengthening with Sea of Okhotsk sea-ice loss, as seen in other models (Xu et al., 2024). This could explain the HadGEM3-GC31-MM's comparatively large stratospheric weakening in our own study. However, Xu et al. (2024) did not specifically examine the additive effects of these separate regional experiments on the winter vortex, which are additionally not expected to be linear (Screen, 2017). Although future work that analyses the PAMIP regional sea-ice loss experiments would aid in the understanding, the experiments we have analyzed here all have the same boundary conditions, enforcing simultaneous sea-ice loss over the entire Arctic region. This means that the model differences apparent in our results cannot be explained by differences in the regional patterns of sea-ice loss.

Furthermore, in this study we find that the thirteen PAMIP-contributing models analyzed display a diversity of responses to future Arctic sea-ice loss. For the stratospheric polar vortex that is in terms of both its mean state and variability, with responses varying in both sign and strength. With a subset of the models, the changes in SSW frequency closely follow changes in the vortex strength and variability (Figure 9b). The models with the greatest vortex weakening show an enhanced neck region waveguide (Figure 10), allowing greater upward wave flux, accompanied by strong deceleration of the vortex (Figure 1) and an increase in SSW frequency (Figure 3a), corroborating previous studies that highlight the key role of the neck (Mudhar et al., 2024; Sigmond & Sun, 2024). But the lack of significantly different simulated responses unfortunately limits our ability to say whether the models are actually responding differently, and if so, why. Regardless, we find that stratospheric responses are well-correlated with climatological state (Figure 9a) and that most models' simulated stratospheres are significantly different from each other (Figure 7). We also note our focus on stratospheric diagnostics defined at the

10 hPa level; consideration of lower stratospheric levels (≥ 70 hPa) could mean more models display a significant response, but do not believe this would notably alter our results.

Along with its relatively small climatological biases, HadGEM3-GC31-MM also shows a large and robust response to future Arctic sea-ice loss across its 300-member ensemble, with a strong stratospheric signal in the mean that seems to emerge even at smaller ensemble sizes (Figure 5) and for a subset of members under QBO-E (Figure 12d). Although we find this model's response to be particularly compelling, these characteristics do not necessarily mean that HadGEM3-GC31-MM gets closest to the “true” response. Nevertheless, besides a large ensemble (Section 4.1), it has the added benefits of having one of the finest resolutions (Section 4.2) and near-even sampling of the two QBO phases (Section 4.4)—features that other models with significant vortex weakening, CanESM5 and TaiESM1, are not able to compete with. Unfortunately, we are overall unable to identify one single factor that may explain its large and significant response, though the strong response under QBO-E is especially interesting, and may perhaps explain TaiESM1's large vortex weakening. Although future modeling studies into stratospheric responses may benefit from running larger ensembles to find significant responses (similar to Peings et al., 2021; Liang et al., 2024; Sigmond & Sun, 2024), we emphasize that the simulated basic state seems to play a key role in models' responses. Our analysis especially highlights the QBO phase as being highly important in experiments such as these, as it influences both climatological vortex state and response (Figure 12). With this subset of PAMIP-contributing models biased toward QBO-E, we suggest that it is possible that we are seeing a multi-model ensemble polar vortex response that is biased toward a weakening. We believe that future studies would thus benefit from ensuring models have simulations in all phases.

In greenhouse gas-forced climate change projections, the response of the stratosphere is uncertain. The response to sea-ice loss represents one source of uncertainty in polar vortex projections, but few studies have so far investigated the extent to which model dependency in the stratospheric response to climate change is driven by model dependency in the response to Arctic sea-ice loss specifically. This is the first study to look across an ensemble of Arctic sea-ice loss-forced complex models, with a focus on the stratosphere. Though the causes of model differences in the stratospheric response to sea-ice loss remain unclear, simulated basic state is an apparent control on vortex response—including, but not limited to, QBO phase. And it is not only the polar vortex mean state that changes in response to sea-ice loss, but also vortex variability, including SSW frequency, and downward coupling to the troposphere. Overcoming the limitations discussed above, including considerations for future inter-model comparisons, are not just crucial for understanding the stratospheric response to Arctic climate change, but for the surface too. Further work is required to better understand the role of the stratosphere in future climate change, but we believe that inter-model comparison protocols such as PAMIP can provide invaluable insight.

Conflict of Interest

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

Data Availability Statement

PAMIP data sets are available from the CMIP data archive (CMIP6 Data Archive [Dataset], 2025), except for that of HadGEM3-LL which is available from Eade (2021). Refractive index was calculated using code based on the `aostools` python package (Jucker, 2021). ERA5 reanalysis data are freely available from the European Centre for Medium-Range Weather Forecasts (Hersbach et al., 2020).

References

- Albers, J. R., & Birner, T. (2014). Vortex preconditioning due to planetary and gravity waves prior to sudden stratospheric warmings. *Journal of the Atmospheric Sciences*, 71(11), 4028–4054. <https://doi.org/10.1175/JAS-D-14-0026.1>
- Andrews, M. B., Ridley, J. K., Wood, R. A., Andrews, T., Blockley, E. W., Booth, B., et al. (2020). Historical simulations with HadGEM3-GC3.1 for CMIP6. *Journal of Advances in Modeling Earth Systems*, 12(6), e2019MS001995. <https://doi.org/10.1029/2019MS001995>
- Ayarzagüena, B., Charlton-Perez, A. J., Butler, A. H., Hitchcock, P., Simpson, I. R., Polvani, L. M., et al. (2020). Uncertainty in the response of sudden stratospheric warmings and stratosphere–troposphere coupling to quadrupled CO₂ concentrations in CMIP6 models. *Journal of Geophysical Research: Atmospheres*, 125(6), e2019JD032345. <https://doi.org/10.1029/2019JD032345>
- Baldwin, M. P., Ayarzagüena, B., Birner, T., Butchart, N., Butler, A. H., Charlton-Perez, A. J., et al. (2021). Sudden stratospheric warmings. *Reviews of Geophysics*, 59(1), e2020RG000708. <https://doi.org/10.1029/2020RG000708>
- Butchart, N., Anstey, J. A., Hamilton, K., Osprey, S., McLandress, C., Bushell, A. C., et al. (2018). Overview of experiment design and comparison of models participating in phase 1 of the SPARC Quasi-Biennial Oscillation initiative (QBOi). *Geoscientific Model Development*, 11(3), 1009–1032. <https://doi.org/10.5194/gmd-11-1009-2018>

Acknowledgments

RM and CT are funded by a NERC GW4+ Doctoral Training Partnership studentship from the Natural Environment Research Council (NE/S007504/1). WJMS, JAS and SIT are supported by Natural Environment Research Council Grant ArctiCONNECT (NE/V005855/1). This work used JASMIN, the UK's collaborative data analysis environment (Lawrence et al., 2013, <https://www.jasmin.ac.uk>).

- Butler, A. H., & Gerber, E. P. (2018). Optimizing the definition of a sudden stratospheric warming. *Journal of Climate*, 6(31), 2337. <https://doi.org/10.1175/JCLI>
- Cai, D., Dameris, M., Garny, H., & Runde, T. (2012). Implications of all season Arctic sea-ice anomalies on the stratosphere. *Atmospheric Chemistry and Physics*, 12(24), 11819–11831. <https://doi.org/10.5194/acp-12-11819-2012>
- Charlton, A. J., & Polvani, L. M. (2007). A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *Journal of Climate*, 20(3), 449–469. <https://doi.org/10.1175/JCLI3996.1>
- CMIP6 Data Archive. (2025). [Dataset]. Retrieved from <https://aims2.llnl.gov/search>
- Cohen, J., Zhang, X., Francis, J., Jung, T., Kwok, R., Overland, J., et al. (2020). Divergent consensus on Arctic amplification influence on midlatitude severe winter weather. *Nature Climate Change*, 10(1), 20–29. <https://doi.org/10.1038/s41558-019-0662-y>
- Deser, C., Tomas, R. A., & Sun, L. (2015). The role of ocean-atmosphere coupling in the zonal-mean atmospheric response to Arctic sea ice loss. *Journal of Climate*, 28(6), 2168–2186. <https://doi.org/10.1175/JCLI-D-14-00325.1>
- Domeisen, D. I. V., & Butler, A. H. (2020). Stratospheric drivers of extreme events at the Earth's surface. *Communications Earth & Environment*, 1(1), 59. <https://doi.org/10.1038/s43247-020-00060-z>
- Eade, R. (2021). PAMIP HadGEM3-LL CMIP6 [Dataset]. <https://doi.org/10.5281/zenodo.5127891>
- England, M. R., Eisenman, I., Till, & Wagner, J. W. (2022). Spurious climate impacts in coupled sea ice loss simulations. *Journal of Climate*, 22(35), 7401–7411. <https://doi.org/10.1175/JCLI-D-21>
- Hall, R. J., Mitchell, D. M., Seviour, W. J., & Wright, C. J. (2021). Persistent model biases in the CMIP6 representation of stratospheric polar vortex variability. *Journal of Geophysical Research: Atmospheres*, 126(12), e2021JD034759. <https://doi.org/10.1029/2021JD034759>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis [Dataset]. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Holton, J. R., & Tan, H.-C. (1980). The influence of the equatorial quasi-biennial oscillation on the global circulation at 50 mb. *Journal of the Atmospheric Sciences*, 10(37), 2200–2208. [https://doi.org/10.1175/1520-0469\(1980\)037<2200:TIOTEQ>2.0.CO;2](https://doi.org/10.1175/1520-0469(1980)037<2200:TIOTEQ>2.0.CO;2)
- Hu, D., & Guan, Z. (2018). Decadal relationship between the stratospheric arctic vortex and Pacific decadal oscillation. *Journal of Climate*, 31(9), 3371–3386. <https://doi.org/10.1175/JCLI-D-17-0266.1>
- Hu, D., Guan, Z., Tian, W., & Ren, R. (2018). Recent strengthening of the stratospheric Arctic vortex response to warming in the central North Pacific. *Nature Communications*, 9(1), 1697. <https://doi.org/10.1038/s41467-018-04138-3>
- Jenkins, M. T., Dai, A., & Deser, C. (2024). Arctic climate feedback response to local sea-ice concentration and remote sea surface temperature changes in PAMIP simulations. *Climate Dynamics*, 62(12), 10599–10620. <https://doi.org/10.1007/s00382-024-07465-y>
- Jucker, M. (2021). Scaling of Eliassen-Palm flux vectors [Software]. *Atmospheric Science Letters*, 22(4), e1020. <https://doi.org/10.1002/asl.1020>
- Karpechko, A. Y., Afargan-Gerstman, H., Butler, A. H., Domeisen, D. I., Kretschmer, M., Lawrence, Z., et al. (2022). Northern Hemisphere stratosphere-troposphere circulation change in CMIP6 models: 1. Inter-model spread and scenario sensitivity. *Journal of Geophysical Research: Atmospheres*, 127(18), e2022JD036992. <https://doi.org/10.1029/2022JD036992>
- Karpechko, A. Y., Wu, Z., Simpson, I. R., Kretschmer, M., Afargan-Gerstman, H., Butler, A. H., et al. (2024). Northern Hemisphere stratosphere-troposphere circulation change in CMIP6 models: 2. Mechanisms and sources of the spread. *Journal of Geophysical Research: Atmospheres*, 129(13), e2024JD040823. <https://doi.org/10.1029/2024JD040823>
- Kim, B., Son, S. W., Min, S. K., Jeong, J. H., Kim, S. J., Zhang, X., et al. (2014). Weakening of the stratospheric polar vortex by Arctic sea-ice loss. *Nature Communications*, 5(1), 4646. <https://doi.org/10.1038/ncomms5646>
- Kolstad, E. W., Lee, S. H., Butler, A. H., Domeisen, D. I., & Wulff, C. O. (2022). Diverse surface signatures of stratospheric polar vortex anomalies. *Journal of Geophysical Research: Atmospheres*, 127(20), e2022JD037422. <https://doi.org/10.1029/2022JD037422>
- Kretschmer, M., Coumou, D., Agel, L., Barlow, M., Tziperman, E., & Cohen, J. D. (2018). More-persistent weak stratospheric polar vortex states linked to cold extremes. *Bulletin of the American Meteorological Society*, 99(1), 49–60. <https://doi.org/10.1175/BAMS-D-16-0259.1>
- Kretschmer, M., Zappa, G., & Shepherd, T. G. (2020). The role of Barents-Kara sea ice loss in projected polar vortex changes. *Weather and Climate Dynamics*, 1(2), 715–730. <https://doi.org/10.5194/wcd-1-715-2020>
- Labe, Z., Peings, Y., & Magnusdottir, G. (2019). The effect of QBO phase on the atmospheric response to projected Arctic sea ice loss in early winter. *Geophysical Research Letters*, 46(13), 7663–7671. <https://doi.org/10.1029/2019GL083095>
- Labe, Z., Peings, Y., & Magnusdottir, G. (2020). Warm Arctic, cold Siberia pattern: Role of full Arctic amplification versus sea ice loss alone. *Geophysical Research Letters*, 47(17), e2020GL088583. <https://doi.org/10.1029/2020GL088583>
- Lawrence, B. N., Bennett, V. L., Churchill, J., Jukes, M., Kershaw, P., Pascoe, S., et al. (2013). Storing and manipulating environmental big data with JASMIN [Software]. *IEEE Big Data*, 68–75. <https://doi.org/10.1109/BigData.2013.6691556>
- Lee, W. L., Wang, Y. C., Shiu, C. J., Tsai, I. C., Tu, C. Y., Lan, Y. Y., et al. (2020). Taiwan Earth System Model Version 1: Description and evaluation of mean state. *Geoscientific Model Development*, 13(9), 3887–3904. <https://doi.org/10.5194/gmd-13-3887-2020>
- Lewis, N. T., England, M. R., Screen, J. A., Geen, R., Mudhar, R., Seviour, W. J. M., & Thomson, S. I. (2024). Assessing the spurious impacts of ice-constraining methods on the climate response to sea-ice loss using an idealised aquaplanet GCM. *Journal of Climate*, 37(24), 6729–6750. <https://doi.org/10.1175/jcli-d-24-0153.1>
- Liang, Y.-C., Kwon, Y.-O., Frankignoul, C., Gastineau, G., Smith, K. L., Polvani, L. M., et al. (2024). The weakening of the stratospheric polar vortex and the subsequent surface impacts as consequences to Arctic sea ice loss. *Journal of Climate*, 1(37), 309–333. <https://doi.org/10.1175/jcli-d-23-0128.1>
- Matsuno, T. (1970). Vertical propagation of stationary planetary waves in the winter Northern Hemisphere. *Journal of the Atmospheric Sciences*, 27(6), 871–883. [https://doi.org/10.1175/1520-0469\(1970\)027<0871:VPOSPW>2.0.CO;2](https://doi.org/10.1175/1520-0469(1970)027<0871:VPOSPW>2.0.CO;2)
- McKenna, C. M., Bracegirdle, T. J., Shuckburgh, E. F., Haynes, P. H., & Joshi, M. M. (2018). Arctic sea ice loss in different regions leads to contrasting Northern Hemisphere impacts. *Geophysical Research Letters*, 45(2), 945–954. <https://doi.org/10.1002/2017GL076433>
- Mudhar, R., Seviour, W. J., Screen, J. A., Geen, R., Lewis, N. T., & Thomson, S. I. (2024). Exploring mechanisms for model-dependency of the stratospheric response to arctic warming. *Journal of Geophysical Research: Atmospheres*, 129(10), e2023JD040416. <https://doi.org/10.1029/2023JD040416>
- Notz, D., & SIMIP Community. (2020). Arctic sea ice in CMIP6. *Geophysical Research Letters*, 47(10), e2019GL086749. <https://doi.org/10.1029/2019GL086749>
- Notz, D., & Stroeve, J. (2016). Observed Arctic sea-ice loss directly follows anthropogenic CO₂ emission. *Science*, 354(6313), 747–750. <https://doi.org/10.1126/science.aag2345>
- Peings, Y., Labe, Z. M., & Magnusdottir, G. (2021). Are 100 ensemble members enough to capture the remote atmospheric response to +2°C Arctic sea ice loss? *Journal of Climate*, 34(10), 3751–3769. <https://doi.org/10.1175/JCLI-D-20-doi:10.1175/JCLI-D-20>

- Richter, J. H., Anstey, J. A., Butchart, N., Kawatani, Y., Meehl, G. A., Osprey, S., & Simpson, I. R. (2020). Progress in simulating the Quasi-Biennial Oscillation in CMIP models. *Journal of Geophysical Research: Atmospheres*, 125(8), e2019JD032362. <https://doi.org/10.1029/2019JD032362>
- Screen, J. A. (2017). Simulated atmospheric response to regional and pan-arctic sea ice loss. *Journal of Climate*, 30(11), 3945–3962. <https://doi.org/10.1175/JCLI-D-16-0197.1>
- Screen, J. A., & Simmonds, I. (2010). The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature*, 464(7293), 1334–1337. <https://doi.org/10.1038/nature09051>
- Seviour, W. J., Mitchell, D. M., & Gray, L. J. (2013). A practical method to identify displaced and split stratospheric polar vortex events. *Geophysical Research Letters*, 40(19), 5268–5273. <https://doi.org/10.1002/grl.50927>
- Shen, X., Kretschmer, M., & Shepherd, T. G. (2024). A forensic investigation of climate model biases in teleconnections: The case of the relationship between ENSO and the Northern Stratospheric Polar Vortex. *Journal of Geophysical Research: Atmospheres*, 129(19), e2024JD041252. <https://doi.org/10.1029/2024JD041252>
- Shen, X., Kretschmer, M., & Shepherd, T. G. (2025). Quantifying the state-dependent causal effect of Barents–Kara Sea ice loss on the stratospheric Polar Vortex in a large ensemble simulation. *Climate Dynamics*, 305(63), 305. <https://doi.org/10.1007/s00382-025-07802-9>
- Sigmond, M., & Sun, L. (2024). The role of the basic state in the climate response to future Arctic sea ice loss. *Environmental Research: Climate*, 3(3), 031002. <https://doi.org/10.1088/2752-5295/ad44ca>
- Simon, A., Gastineau, G., Frankignoul, C., Lapin, V., & Ortega, P. (2022). Pacific Decadal Oscillation modulates the Arctic sea-ice loss influence on the midlatitude atmospheric circulation in winter. *Weather and Climate Dynamics*, 3(3), 845–861. <https://doi.org/10.5194/wcd-3-845-2022>
- Smith, D. M., Dunstone, N. J., Scaife, A. A., Fiedler, E. K., Copsey, D., & Hardiman, S. C. (2017). Atmospheric response to Arctic and Antarctic sea ice: The importance of ocean-atmosphere coupling and the background state. *Journal of Climate*, 30(12), 4547–4565. <https://doi.org/10.1175/JCLI-D-16-0564.1>
- Smith, D. M., Eade, R., Andrews, M. B., Ayres, H., Clark, A., Chripko, S., et al. (2022). Robust but weak winter atmospheric circulation response to future Arctic sea ice loss. *Nature Communications*, 13(1), 727. <https://doi.org/10.1038/s41467-022-28283-y>
- Smith, D. M., Screen, J. A., Deser, C., Cohen, J., Fyfe, J., García-Serrano, J., et al. (2018). The Polar Amplification Model Intercomparison Project (PAMIP) contribution to CMIP6: Investigating the causes and consequences of polar amplification. *Geoscientific Model Development*, 12(3), 1–42. <https://doi.org/10.5194/gmd-2018-82>
- Steffing, J., Semmler, T., Zampieri, L., & Jung, T. (2021). Response of northern hemisphere weather and climate to arctic sea ice decline: Resolution independence in polar amplification model intercomparison project (PAMIP) simulations. *Journal of Climate*, 34(20), 8445–8457. <https://doi.org/10.1175/JCLI-D-19-1005.1>
- Sun, L., Deser, C., & Tomas, R. A. (2015). Mechanisms of stratospheric and tropospheric circulation response to projected Arctic sea ice loss. *Journal of Climate*, 28(19), 7824–7845. <https://doi.org/10.1175/JCLI-D-15>
- Walsh, A., Screen, J. A., Scaife, A. A., Smith, D. M., & Eade, R. (2025). Interdependent extratropical atmospheric responses to Arctic sea-ice loss, QBO and ENSO. *Journal of the Atmospheric Sciences*.
- Xu, M., Screen, J. A., Tian, W., Zhang, J., Zhang, C., & Yu, H. (2024). Influence of regional sea ice loss on the Arctic Stratospheric Polar Vortex. *Journal of Geophysical Research: Atmospheres*, 129(15), e2023JD040571. <https://doi.org/10.1029/2023JD040571>
- Xu, M., Tian, W., Zhang, J., Screen, J. A., Zhang, C., & Wang, Z. (2023). Important role of stratosphere–troposphere coupling in the Arctic mid-to-upper tropospheric warming in response to sea-ice loss. *npj Climate and Atmospheric Science*, 6(1), 9. <https://doi.org/10.1038/s41612-023-00333-2>
- Yu, H., Screen, J. A., Xu, M., Hay, S., & Catto, J. L. (2024). Comparing the atmospheric responses to reduced Arctic sea ice, a warmer ocean, and increased CO₂ and their contributions to projected change at 28C global warming. *Journal of Climate*, 37(23), 6367–6380. <https://doi.org/10.1175/JCLI-D-24-0104.1>
- Zheng, C., Wu, Y., Ting, M., Screen, J. A., & Zhang, P. (2023). Diverse Eurasian temperature responses to Arctic sea ice loss in models due to varying balance between dynamical cooling and thermodynamical warming. *Journal of Climate*, 36(24), 8347–8364. <https://doi.org/10.1175/jcli-d-22-0937.1>