We would like to thank Prof. Hitchman and the anonymous reviewer for their helpful and constructive feedback and the editor for overseeing the review process. The comments have helped us improve the manuscript by expanding the discussion on gravity wave drag and including supporting evidence from previous studies. We have revised the manuscript and figures accordingly. All changes are indicated using tracked changes, with additions shown inline and deletions marked with strikethrough. Our responses are provided below, with references to the corresponding line numbers in the tracked manuscript.

Response to Reviewer #1 (Prof. Hitchman)

This paper describes the effects of zonal wind nudging of the QBO on the amplitude and phase of the SAO (QBOi phase 2) in ~10 numerical models. It shows that models with better gravity wave drag parameterizations, higher model tops, and perhaps less diffusion near the model top can be more readily improved by QBO nudging, with regard to fidelity in representing the SAO. In the abstract it is stated that most models show an improvement with a reduced easterly SAO bias in response to the corrected QBO bias and that changed gravity wave drag in response to the improved QBO is attributed as the forcing term having the largest overall impact, which is parameterized unresolved gravity wave drag. This paper functions well as a medium of exchange of information for the community and it is useful and clear. It shows that gravity wave drag is of primary importance in explaining their results, but does not elaborate on the mechanism. I recommend that the authors include a description of what they think is happening in the models with regard to changes in gravity wave drag and recommend minor revision.

1. The primary conclusion in the abstract, restated on lines 109-110 and elsewhere, is robust, but the paper does not say much more about any physical process. I encourage the authors to provide a description of how the gravity wave filtering works, with regard to stating what happens to westward travelling waves and to eastward travelling waves for changes in a given wind layer imposed by the nudging. What are the gravity wave parameterizations used? Which models try to represent deep convective gravity waves from first principles? What are the ways that parameterized gravity waves can be absorbed in these models, going upward into the mesosphere, where both shear and reduced density can cause them to break?

   Thank you for this helpful comment. Based on this suggestion and related comments below, we have made modifications throughout the text to include more discussion on gravity waves. In the Introduction, we have included the discussion of wave filtering and saturation (lines 64-70), and mentioned how QBO winds influence GW propagation to SAO altitudes (lines 111-115). The Data and Methods section now provides further details on the GW parameterization schemes and their sources (lines 207-218). In the Results section, we have modified the discussion of how QBO wind nudging affects GWD (lines 548-553),

and we have expanded this further in the Summary and Discussion section (lines 675-682).

2. l2-22: It would be helpful to briefly characterize vertical resolutions in the QBO and SAO layers. Does vertical resolution matter for parameterized gravity wave drag?

   Thank you for this suggestion. We have added a paragraph in the Data and Methods section discussing the vertical resolution of the models (lines 220-225). Vertical resolution can influence wave propagation and critical-level filtering, thereby affecting the vertical distribution and magnitude of wave forcing.

3. l148-149 and elsewhere: This is an important issue. I suggest that you describe a bit more about the upper layer in the models. Which models use Rayleigh friction? In your model standard, HadGEM3, what is the damping/numerical diffusion profile into the mesosphere?

   Thank you so much for the comment. We have now added more description on the damping based on the responses we received from different modelling groups (lines 165-171). For HadGEM3, the only explicit diffusion mechanisms in the model are semi-implicit off-centering and damping of the vertical velocity near the model top.

4. l164 – I agree that nudging zonal wind is the right way to do it, in order to have the right relationship between temperature anomaly and vertical motion for a wave-driven MMC (e.g., downward / warm). If some models nudged temperature, then in the conclusions you could suggest more strongly that this might be a major reason why their results weren't as hoped. Vertical motion would be upward in an imposed warm anomaly in QBO W shear in a model with a radiative code, which is opposite to our understanding.

   Thank you for the comment and apologies if the text was unclear. None of the models are introducing temperature nudging. Most of the models are nudged towards zonal-mean zonal wind, while a few are nudged towards zonally varying zonal wind.

5. l180-190: You could introduce the idea here that another significant metric for evaluating model processes could be the difference in amplitude between the first and second phases of the SAO in each year. The reason that I mention this is that your definition of SAO blends both the first and second together. You could have a mean and a difference "PC" to intercompare models.

   Thank you very much for this insightful suggestion. We agree that examining the amplitude differences between the first and second phases of the SAO would provide a useful additional analysis for evaluating model performance. Although our current analysis considers the two phases together, Figures 2 and 3 provide some insight into their seasonal asymmetry. A more systematic analysis would be a valuable extension of this work and an interesting direction for future research.

6. l215-220 - It may also be the case that vertical wind structure near 10 hPa is more important than near 50 hPa for filtering gravity waves, which might affect the more immediately-adjacent SAO layer more readily. But maybe that is not true. What do you think?

   Thank you for the comment. Since gravity waves filtered at 50 hPa encounter additional QBO wind layers as they propagate upward, whereas waves filtered at 10 hPa are closer to the SAO region, it is possible that 10hPa vertical structure might have a more direct impact. However, reference level selection is somewhat unaffected by this. For e.g., for HadGEM3, comparing current work with Jaison et al. (2024) shows that using either 50 hPa or 10 hPa yields broadly similar results.

7. l245: You effectively treat HadGEM3 as the reference model in section 3.1, which seems fine. Do you think that the property of a higher model top is the primary reason? Suggest describe the damping properties of HadGEM3 in the upper layers a bit more.

   Thank you for the comment. We think that a higher model top, along with relatively high number of vertical levels and lower vertical grid spacing, and negligible damping at SAO altitudes, all makes HadGEM3 desirable. Some models from Distinct group also have model top at 0.01 hPa. However, they have either a smaller number of vertical levels, or strong damping at SAO altitudes.

8. l296-301 – In light of this discussion, what does this imply about the vertical profile of gravity wave drag? Perhaps add the thread of how it is that the models err here, which can be elaborated on in the conclusions?

   Thank you so much for the suggestion. In this section, we were focusing on how the SAO characteristics have changes between Control and Nudged, and haven't started looking at how the forcing terms are affected. We have added a line here (lines 346-347) and pointed readers to the Results section where more details are added now (lines 548-553). And this is elaborated in the Summary and Discussions (lines 675-682).

9. Are there any models that represent the seasonal variation in deep convective gravity wave driving? If the greatest tropical deep convection occurs in DJFM, then a model which can represent that might be expected do better at the 1st v 2nd phase SAO amplitude asymmetry.

   Thank you for the interesting suggestion. As now added in the main text (lines 210-215), CAS-ESM, E3SM2 and CESM2 (CAME6) follows Beres et al. (2004) for deep convective sources. But not all of these models have good SAO amplitude asymmetry. HadGEM3 has a GW source which is proportional to square root of precipitation and captures the spatial and temporal variability of GW generation. And HadGEM3 does have a reasonable SAO amplitude asymmetry. These comparisons suggest that deep-convective gravity-wave sources may help improve the simulated SAO in some models, but it is not sufficient on its own to ensure a realistic amplitude asymmetry.

10. l357-358: The time mean vertical profile and how it is maintained is important. You could comment on quantification of GWD among models for the annual budget. It would give further useful information about GW behavior in the models.

    Thank you for the comment. While section 3.1 and figure 4 only focus on time mean vertical profile of wind, section 3.2, figure S2 and figure 7 gives useful information on how GWD contributes to the momentum budget and how it varies among models. Section 3.2 is further edited to explain how different forcing terms differs among models (lines 514-517). We apologise if we have misunderstood the reviewer's comment and addressed a different issue.

11. l495-497: How are the gravity wave momentum fluxes for westerly and easterly gravity waves affected in each layer due to improving the QBO wind structure?

    Thank you for this comment. We have expanded the discussion to explain how changes in the QBO wind structure may affect gravity-wave propagation to SAO altitudes (lines 548-553, 568-570). Although we have not explicitly shown the gravity-wave stress profiles, we discuss the implications for GW propagation based on the GWD profiles shown in Fig. 7.

12. l505: Figure 8d shows the influence of the SAO on the QBO down to 50 hPa, due to the MMCs associated with the SAO. I think this is similar to what was shown by Coy (circa 1990).

    Thank you for this comment. We were unable to identify the specific paper by Coy (circa 1990) referred to here. If the reviewer could provide the full reference, we would be happy to consult it.

13. l511-512: Please describe for the reader what this statement implies about what is missing in the models re GWD.

    Thank you for the comment. We have added the implication.

I notice that in this paper tropical GWD is deemed important for the westerly SAO phase, but I think it is also important for the easterly phase. The cross-equatorial flow is primarily a gravity-wave driven flow, a response to the large winter and summer gravity wave drag patterns. As air travels meridionally through the tropics it is exposed to a lot of deep convective gravity wave drag. In addition to GWD driving the pole to pole circulation during the SAO easterly phase, gravity wave drag integrated along the pathway from one hemisphere to the other also helps to explain why a maximum in easterly wind does not occur right at the equator, which would occur in inviscid flow starting with zero zonal velocity at a pole. Tropical westerly GWD, in addition to treatment of inertial instability parameterization in models on the winter subtropical side, can help to limit easterly advection.

Thank you for these final remarks. From previous observational studies, GWD is identified important for SAO easterlies during the rare instances when QBO filtering of westward GWs is less effective. We have now added this point to the Introduction. We also agree that westerly GWD may limit the advection of easterly momentum. Furthermore, although the Brewer–Dobson circulation associated with SAO easterlies is strongly associated with extratropical planetary-scale Rossby waves, we agree that small-scale gravity waves also contribute strongly to this and could play a role in maintaining the SAO easterlies.

Response to Reviewer #2

The paper "Quasi-Biennial Oscillation Modulation of the Semi-Annual Oscillation in QBOi Models" by Jaison et al. is an important contribution that compares the stratopause semiannual oscillation (SAO) simulated in a number of model simulations performed for the QBOi initiative. Model runs nudged to the QBO in the ERA5 reanalysis are compared with free model runs (Control) and improvements of the SAO are investigated. The SAO in the MERRA2 reanalysis is taken as a reference because MERRA2 assimilates MLS satellite observations and the SAO is therefore at least partly constrained by observations.

Two groups of QBOi models are identified. The SAO in the first group is improved by nudging to the ERA5 QBO. By removing a westerly bias of the QBO, the easterly bias of the SAO is ameliorated. The second group of models shows a mixed response to QBO nudging. The authors attribute the improvements of the SAO to a more realistic filtering of gravity waves by the QBO in the stratosphere below SAO altitudes.

Overall, this is a well-written and important paper that is recommended for publication in WCD after minor revisions.

A list of minor and technical comments is given below.

MINOR COMMENTS

(1) l.53: Please provide reference for these numbers, e.g. Garcia et al., JGR, 1997

Thank you for the comment. We have added the reference (line 53).

(2) l.59: You should mention that from satellite observations it seems that significant driving of SAO easterlies by gravity waves seems to be linked to rare cases when the

QBO wind filtering of easterly gravity waves by the QBO is less effective (Ern et al., Ann. Geophys., 2015).

Ern, M., Preusse, P., and Riese, M.: Driving of the SAO by gravity waves as observed from satellite, Ann. Geophys., 33, 483-504, https://doi.org/10.5194/angeo-33-483-2015, 2015.

Thank you for pointing this. We have added the description (lines 59-61).

(3) l.67: You should mention the reference Burrage et al., JGR, 1996 also here, as this paper introduces the original HRDI observations.

Thank you for highlighting this. We have added the reference (line76).

(4) l.84: you should mention that driving of the SAO westerly phase by gravity waves was also shown directly from satellite observation of gravity waves (Ern et al., ACP, 2021).

Thank you for the helpful comment. We have added the information (lines 95-96).

(5) l.89: You should include the reference Li et al., J. Clim., 2025 that investigates the performance of CMIP6 models for simulating the SAO. Further, you should explain what is common and what different between the QBOi Control run and the CMIP6 simulations described in Li et al., 2025.

Li, X., Han, Y., Xia, Y., Tan, X., Li, S., Wen, K., Wang, X., Li, Y., and Xie, F. (2025), Evaluation of CMIP6 Model Performance in Simulating the Stratospheric Semiannual Oscillation, J. Climate, 38, 4281-4297, doi:10.1175/JCLI-D-25-0045.1

Thank you for highlighting this paper. We have now included a reference to Li et al. (2025) in the Introduction and briefly summarised their key findings based on coupled model simulations from the CMIP6 historical experiments (lines 99-101).

(6) l.157: You should emphasize that nudging to ERA5 makes sense because on zonal average the ERA5 QBO should be relatively close to observations (e.g., Ern et al., ACP, 2023).

Ern, M., Diallo, M. A., Khordakova, D., Krisch, I., Preusse, P., Reitebuch, O., Ungermann, J., and Riese, M.: The quasi-biennial oscillation (QBO) and global-scale tropical waves in Aeolus wind observations, radiosonde data, and reanalyses, Atmos. Chem. Phys., 23, 9549-9583, https://doi.org/10.5194/acp-23-9549-2023, 2023.

Thank you for the comment. We have added the justification (lines 180-182).

(7) l.171: You should mention that there is a feedback between ozone and the QBO, and that interactive ozone will likely lead to changes in the simulated QBO (e.g., Ming et al., JGR, 2025).

Ming, A., Hitchcock, P., Orbe, C., and Dube, K. (2025). Phase and amplitude relationships between ozone, temperature, and circulation in the quasibiennial oscillation. Journal of Geophysical Research: Atmospheres, 130, e2024JD042469. https://doi.org/10.1029/2024JD042469.

Thank you for highlighting this. We have added the information (lines 198-199).

(8) l.185-190: Did you check how sharp is the spectral peak of the SAO?

Is no significant spectral power lost by interannual variations of the SAO that would broaden the spectral peak?

Thank you very much for this comment. We have not explicitly examined the SAO spectral peak in this study. However, previous studies, such as Pascoe et al. (2005), Jaison et al. (2024) and Li et al. (2025), have shown that, unlike the QBO, the SAO exhibits a distinct spectral peak at a periodicity of 6 months. As our aim is to evaluate whether the models can adequately capture the amplitude and phase of variability at this periodicity, which is a characteristic feature of the SAO, we have used the 6-month periodicity as the cutoff in our analysis, following a similar approach to Smith et al. (2019).

(9) l.268/269, about Fig.2: My impression is that, except for HadGEM, at 1hPa near the zero wind line several models show downward propagation of the easterly SAO phases which seems to be less the case for MERRA2 and observations (e.g. Garcia et al, 1997). This is particularly seen for MIROC and LMDz and should be mentioned.

Thank you so much for the comment. We agree that some downward propagation of the easterly phase is evident, particularly in MIROC6.1 (p1) nudged and LMDz6. In the other models, however, the zero-wind line (solid line), marking the onset of the easterly phase, occurs at approximately the same time over a range of altitudes. Although the downward tilt of the easterly-phase peak is evident (dashed line). We are highlighting MIROC 6.1(p1) and LMDz6 in the main text (lines 312-315), we are happy to reconsider this interpretation if the reviewer feels that the other models should also be highlighted.

(10) Caption of Fig.2: mention that the figure refers to the "HadGEM3-similar group"

Thank you for the comment. We have added ‘Similar group models’

(11) Caption of Fig.3: mention that the figure refers to the "HadGEM3-distinct group"

Thank you for the comment. We have added ‘Distinct group models’

(12) Fig.6: Optional comment: Would it make sense to include a distribution plot for MERRA2 (for a time period when MLS is assimilated!)?

Thank you for this optional suggestion. While including a MERRA-2 distribution plot could provide additional context, since figures 4 and 5 already provide reasonable

comparisons with MERRA-2, we have chosen not to include this additional analysis in the present study.

(13) Fig.07: Optional comment: Would it make sense to include also MERRA2 TEM diagnostics in the figure?

Thank you for this comment. We have not included MERRA-2 in Fig. 6 because the figure shows Control-minus-Nudged differences in the forcing terms, and including the absolute forcing profiles from MERRA-2 could make the comparison confusing. The primary aim of this analysis is to understand what drives the changes in the SAO winds in response to nudging. In contrast, Jaison et al. (2024) discuss the MERRA-2 TEM diagnostics and compare them with those from HadGEM3. However, if the reviewer suggests a comparison with MERRA-2 necessary, we would be happy to reconsider this.

(14) Fig.9: to better separate upper and lower row of panels, please introduce more white space

Thank you for the comment. We have modified the figure now.

(15) l.479: You should mention that the finding of peak GWD a bit below the maximum change in the wind is as expected. It has been shown that eastward forcing of the stratopause SAO happens mainly in the zonal wind shear zones (e.g., Ern et al., 2021).

Thank you so much for the suggestion and for highlighting that the maximum eastward forcing happens in the zonal wind shear zones. However, since we are showing the Control minus Nudged profile of wind and forcing terms in the figure, rather than their absolute profiles or the corresponding wind shear, we are cautious about directly inferring this relationship from the figure.

(16) l.619-621: You should include the information that already some evidence exists that high-frequency gravity waves are important for driving the SAO. As has been argued by Ern et al., Ann. Geophys., 2015 wave saturation that is not directly linked to critical levels seems to play also an important role

Thank you so much for the comment. We have now highlighted this and made edits in introduction (line 70) and discussion (lines 689-690).

TECHNICAL COMMENTS

Thank you for the technical comments. We have made the following changes.

(1) General comment: please capitalize first letter of "figure", or "table" if followed by a figure, or table number. So far this is quite mixed throughout the paper.

(2) l.65/66: Barking Sands -> Barking Sands, Hawaii

(3) l.168: analyed -> analysed

(4) l.179: have a parametrized non-orographic GW->use a non-orographic GW parameterization

(5) l.202: transition -> transitions

(6) l.249 is used -> are used

(7) l.322: remove blank after "models"

(8) Fig.4: Please check whether in the caption red and blue are interchanged for Fig.4d-f!

(9) In Table 1, please remove vertical bar in front of "-3.1" in the Control run wind column for MRI.

(10) l.561 pattern -> patterns

(11) l.575 they -> they are

We haven't made this change as 'they' seems more appropriate than 'they are'.