Reviewer 2:

Comments/suggestions by line number:

Fig. S1 caption: state the value of the blue & red line contours shown below 30 hPa.

We have added "The ±2.5m/s isotach is included in blue and red below 30hPa"

136-137: If this refers to the positive (red) regression coefficients, is CNRM the exception (not IPSL)?

We are referring to the blue (negative) coefficients, as these track the QBO winds themselves and their downward horseshoe to the troposphere. This is now clarified.

Fig. 4 caption: from the panel d5 title it seems to be the same region as for the models (220-240E, 55-70N).

Thanks for catching this! corrected

157: insert " (any of December, January or February)" after "available data" to help clarify these are the available months for this test (if I've understood correctly).

Clarified to "months in **DJF from the available data"**

188-189: Uncertainty in the observed response is so large (confidence intervals in Fig. 4b5) that it may not be possible to say that the models' seasonality differs from that observed.

For HadGEM3 we can conclude there is a discrepancy. The huge amount of HadGEM3 output available allows us to conclude that this discrepancy is robust, and not just sampling variability, as the error bars don't overlap. We have added this to the text.

205: typo: "Scandinavia"

corrected

222: Weaker westerlies? Based on the negative values in the Pacific boxes in Fig. 6 or Fig. S11. I guess some plots show stronger westerlies on the poleward side of this feature, but it looks to me like the negative blob is more consistent across all plots. (Line 231 also mentions weaker westerlies.)

Thanks for catching this. We indeed meant weaker westerlies. Corrected.

269-270: Not sure what features are being highlighted as robust in South Asia or East Asia. For HadGEM and MIROC at least there seems enough agreement across experiments that these patterns represent a real response, but the two models don't look very similar to me in South and East Asia.

We have modified the projection for this figure to focus more tightly on the tropics and subtropics. There is enhanced precipitation in South and East Asia in the models, and this is the point we are trying to make. However the details of where the precipitation increases is model dependent.

284: Fig S10 shows ERA5. Fig S12 uses 1970-onward but for tas. Was a figure mistakenly left out of the Supplemental?

We intended to refer to Figure S11 and not S10. This has been corrected. Thanks for catching this.

286: warm anomaly --> cold anomaly

Corrected. Thanks for catching this.

287: Meant to refer to Fig. S12 here? (in addition to Fig 9)

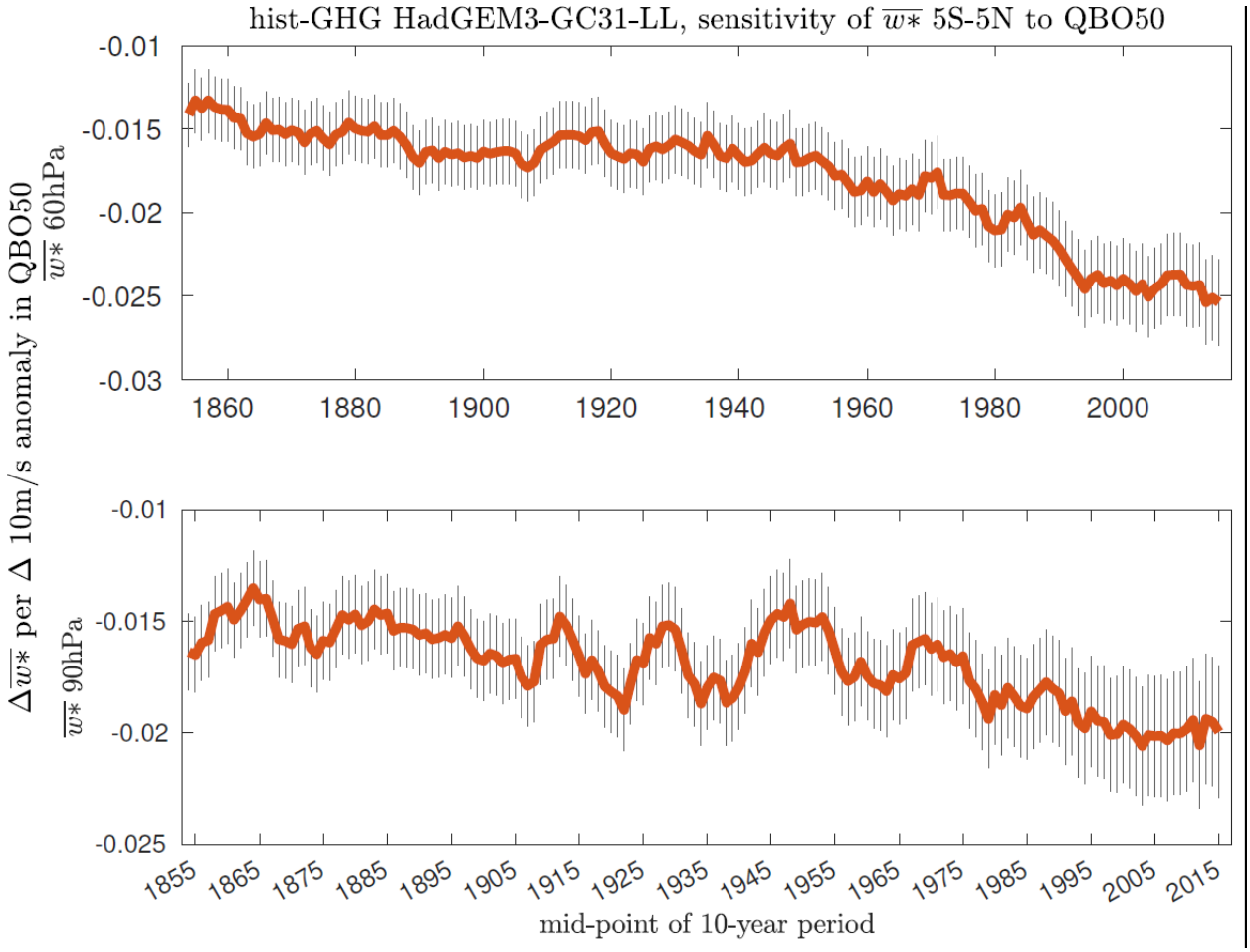
"and S12" has been added.

301: This argument seems to assume the vertical shear remains the same, but if the QBO weakens then the shear weakens. Is that decrease large enough to affect this conclusion? e.g., in eq. (2) could N and the T anomaly both decrease, leaving w* about the same? Are there any runs of the same models with w* available that could be used to check? (I don't imagine a large ensemble would be needed for this.)

Our proposed mechanism involves the weakening of background N, not the N associated with the QBO. The reduction in the background decadal mean N is due to changes in GHGs directly, and is a well understood thermodynamic/radiative process (Vallis 2015) in a warmed climate. This has been clarified.

The consequence of this is the w* per m/s of QBO winds should increase. w* is available for just one model - HadGEM3 -and for this model we now explicitly

provide evidence for this effect. Below is the regression coefficient of equatorial w*bar at 60hPa and at 90hPa per 10m/s change in QBO winds at 50hPa for rolling 10-year chunks of data in the hist-GHG experiment from HadGEM3. Note that the regression coefficient is always negative (i.e., eQBO leads to upwelling), but that the relationship intensifies in a warmed climate with higher GHGs as discussed in the main text. The 95% uncertainty bounds on the regression coefficient as given by a Student-t test is indicated with a vertical line. Units are mm/sec per 10m/s. We have included this figure in the Supplementary Material.



312: "Prescribed ozone changes invigorate the QBO in both models": the QBO amplitude in wind and/or temperature increases? Is a reference to Butchart et al. 2023 appropriate? (https://doi.org/10.1029/2023GL104401)

Garfinkel et al 2025 showed the invigoration in QBO winds, and this is now clarified

We added a citation to Garfinkel et al and to Butchart et al 2023 here.

332: "leads to" --> "is consistent with"

Changed as suggested

375: Line 280 said that this increase is not statistically significant, so if this finding is mentioned in the conclusions I think at least that caveat should be mentioned.

“though not significantly” has been added

380-384: It seems to me this hinges on what exactly is meant by the term "teleconnection". If it refers to a response magnitude, e.g. size of a polar-vortex composite difference, then the models do underestimate these responses. If it refers to the mechanism(s) by which a QBO influence leads to a response, the regression results argue that the models do represent the mechanisms but underestimated QBO amplitude results in a weak response. If so, perhaps that point could be made more explicit here.

We agree this could be worded more precisely. We have changed “biases in teleconnections themselves” to “biases in the processes that connect QBO winds with remote regions”