Reviewer #1

Minor points:

The last sentence in the abstract “implying that for strong ENSO events a stratospheric pathway may not be necessary for the ENSO-North Atlantic teleconnection” : recommend change “necessary” to other words such as “dominant” since all the experiments still have a stratospheric pathway although maybe not as important as the tropospheric one.
We agree, changed

Line 130: reorder the Supplementary Figures to let Figure 1 appear first in the text.
The supplementary figures have been reordered

Line 144: “La Niño” -> “La Niña” - fixed

Line 145: “ad” -> “and” - fixed

Line 187: “La Nina” -> “La Niña” - fixed

Figure 3: How about in OBS? Or in other words, does the author believe that the weak role played by the stratosphere is a character of the current models or also true in OBS? Better mention this somewhere in the text.
We have added ERA5 to Figure 3 (also to Figure 5), and discuss it in the text. We think our model runs show evidence that for strong ENSO events the tropospheric teleconnection is dominant, which we point out in the discussion, i.e. *"This result appears consistent with that by Bell et al. (2009) who also found that for strong ENSO events the tropospheric teleconnections dominate."*, although we don't think the observational evidence we present in this paper necessarily confirms that result.

Line 193: suggest cite the corresponding figure before “because of ….” - agree, added "Figure 4 c"

Supplementary Figure 5: suggest unify the subplots’ Y coordinates - agree, merged into a single colorbar

Figure 5 and Supplementary Figure 5: It looks that the HF in FullBC and TropBC (Figure 5(k) and 5(l)) is more sensitive to the AL change than those in CTRL and StratBC (Figure 5(i) and 5(j)), which is different from the results shown in Supplementary Figure 5. I notice that the season used for calculation is also different, which may cause the inconsistency? - Agree, we have used NDJFM for both plots for consistency (Supp Fig 5 has moved to Fig 6)

Reviewer #2

The authors have carried out a set of bias-correction experiments in the ECHAM 6 model investigating the role of basic-state model biases on the teleconnection between ENSO and the NAO. By imposing a climatologically varying set of fixed tendencies to the temperature and divergence fields in the troposphere, stratosphere, or both, they can partially correct model biases in different regions of the model. They then impose tropical pacific SST anomalies modeled after El Nino or La Nina states and analyse the impact on the NAO, considering established pathways for the teleconnection through the stratospheric polar vortex.

They find that while the bias correction does affect the response of the polar vortex to ENSO, the NAO response to ENSO itself is largely unaffected. They conclude from this that, in the model, the teleconnection between ENSO and the NAO does not depend on the stratospheric pathway, and that it is not sensitive to the model biases corrected for by the bias correction methodology employed. Another interesting hypothesis presented is that the response of near-tropopause heat fluxes (indicative of vertically propagating waves) to the Aleutian low may saturate at large amplitudes.

I have a few questions and concerns outlined below; my main criticism is that the authors should discuss their results more quantitatively in the context of the observed teleconnection. The text could also have used another read through to catch typos.

If these concerns are meaningfully addressed the paper should meet the standard for publication.

Major comments

Structure of Biases

I have the same criticism of this revised paper as I did of the discussion paper: the authors discuss the effects of model bias on the ENSO/NAO teleconnection without clearly showing or discussing the detailed structure of the biases. The supplementary figures (1 through 3) are far more meaningful than the profiles of RMSE shown in Figure 1. They should be included in the main text in addition (or in lieu of) the current panels a-f of Figure 1 and the discussion in 3.1 should be rewritten in terms of the structure of the temperature and wind biases that are corrected.

In particular, there are strong cold biases (I assume what is shown is model - reanalyis, though this is not specified) at high latitudes near and above the tropopause evident in supplemental figure 2. This is included in the TropBC run, but the bias may be better thought of as a lower-stratospheric bias. This could be indicative of an extratropical tropopause that is too high in the model. The TropBC run reduces the high latitude temprature anomalies at the expense of introducing cold anomalies in the tropical upper troposphere. The wind biases that are corrected are in thermal wind balance with these anomalies (as expected), but this makes sense of the significant reduction in RMSE found in the mid-latitude stratosphere in TropBC and FullBC.

Thank you for the constructive criticism regarding this section, and apologies for not correcting in the first round of reviews. We have replaced the RMSE profiles in Figure 1 with zonal mean plots from the supplementary figures, as they are much more descriptive. We have re-written section 3.1, outlining the specific biases, particularly those relevant for ENSO teleconnections, and details of where the biases have been reduced in the different experiments.

Structure of ENSO response

The differences between the modeled and observed teleconnections are very substantial, particularly with regards to the stratospheric pathway of the teleconnection. This is material to the central conclusions of the manuscript with regards to the relative importance of the tropospheric pathway in this model: the model is not properly capturing the stratospheric pathway in its basic configuration or in the bias correction runs. This is particularly evident with regards to the response of the heat flux metric (panel 2c), which is highly nonlinear in the observations, with a positive response to either signed tropical Pacific SST anomlies. This leads to a highly nonlinear stratospheric response (panels d and e), and an NAO response with a seasonal structure that more closely resembles the stratospheric seasonal cycle than the seasonal cycle of the tropical divergence response.

Statements made in lines 163 and 207-8 are far too speculative and the hypothesis that this non-linearity is consistent with sampling variability should be quantitatively tested. This can be done by subsampling periods from the model runs as long as the observational record to determine if something like the observational signal could have occurred plausibly by chance within the context of the models variability. Given the large structural differences seen in Fig. 2 this comparison needs to be done more quantitatively.

These differences also need to be explicitly discussed both in section 3 and in the conclusions.

We have approached this in two ways. Firstly, since the manuscript was first submitted more ERA5 data has become available, with a back extension to 1950. We have included this extra data to increase the number of ENSO events we can compare to. We find the HF response to be less non-linear (perhaps a detailed study of the ERA5 Back Extension ENSO response would be of interest, but beyod the scope of this paper). Secondly, as suggested, we have subsampled our model runs. We show the subsampling results using the JFM value of HF, since this is the period with the greatest difference between model and reanalysis. We find the observed values are within the upper adjacent values, indicating the observational values are within the sampling uncertainty of all model runs. Therefore, while the reviewer is right pointing to the differences between observed and simulated teleconnections, our tests indicate that there is no evidence that the models do not capture the observed stratospheric pathway of ENSO teleconnections. Observational records are simply too short to constraint the magnitude of ENSO teleconnection even with respect to the sign of the response.

Specific comments:

l 34: This isn't consistent with Figure 2; the NAO response is negative, not positive. Similar non-linearities are seen in HF and stratospheric responses.
That was worded badly, corrected

l 158: 'almost opposite' this is ambiguous and confusing - the text could be read to mean that the La Nina response is also an increase in HF but with CTRL and StratBC runs showing half as strong a response as Full BC and TropBC.
Agree, the sentence was re-written.

l 162-3: what is the 'canonical' response here? the observations?
This was poorly worded and has been changed

l 273: I find supplemental figure 5 quite helpful in supporting this argument - I think it's worth showing in the main text.
Agree, it has been moved to the main text

Methodology:

Are the temperature and divergence bias correction tendencies computed from a nudging run in which only temperature and divergence are relaxed, or ar they taken from a run in which all four fields are nudged?
All four fields - divergence, vorticity, temperature, log surface pressure - were used in the nudging stage, then just divergence and temperature for the bias correction. We have added some clarification to Section 2. This setup was found experimentally, to have the greatest reduction in biases without introducing additional errors. Because the key goal of the procedure is to reduce biases we believe our approach is well justified.

Fig. 4: Since we are comparing magnitudes of responses in different runs, the shading should show a confidence interval (which is relevant for comparison), not a measure of variability (which is not). Changed

Typos

l57: affect - fixed

l81: nudged - fixed

l145 symmetric and actually - fixed

l163: response. - fixed

l185-187: Please clarify - text has been rewritten

l190: which effect? - The effect of a weaker or stronger vortex on the NAO. Text added

l222: its - fixed

l269: closer to observations. - fixed