Minimal impact of model biases on northern hemisphere ENSO teleconnections.

Nicholas L. Tyrrell[1], Alexey Yu. Karpechko[1]

[1]Finnish Meteorological Institute, Helsinki, 00500, Finland

5 *Correspondence to*: Nicholas L. Tyrrell (nicholas.tyrrell@fmi.fi)

Abstract. Correctly capturing the teleconnection between the El Niño–Southern Oscillation (ENSO) and Europe is of importance for seasonal prediction. Here we investigate how systematic model biases may affect this teleconnection. A two–step bias–correction process is applied to an atmospheric general circulation model to reduce errors in the climatology. The bias–corrections are applied to the troposphere and stratosphere independently and jointly~~separately and together~~ to produce 10 a range of climates. ENSO type sensitivity experiments are then performed to reveal the impact of differing climatologies on the ENSO–Europe teleconnections.

The bias–corrections do not affect the response of the tropical atmosphere nor the Aleutian Low to the strong ENSO anomalies imposed in our experiments. However, in El Niño experiments the anomalous upward wave flux and the response of the northern hemisphere polar vortex differ between the climatologies. We attribute this to a reduced sensitivity of the 15 upward wave fluxes to the Aleutian Low response in the bias-correction experiments, where the reduced biases result~~results~~ in a deepened Aleutian Low in the base state. Despite the differing responses of the polar vortex, the NAO response is similar between the climatologies, implying that for strong ENSO events the~~a~~ stratospheric pathway may not be the dominant pathway~~necessary~~ for the ENSO–North Atlantic teleconnection.

1 Introduction

20 The El Niño–Southern Oscillation (ENSO) has been shown to influence European climate via tropospheric and stratospheric teleconnections. Although ENSO is a key driver of global variability on seasonal to annual timescales, it's effect on Europe is less robust (Brönnimann~~nninman~~ 2007), and exhibits decadal variability (Rodríguez–Fonseca et al., 2016). The large seasonal variability in the mid–latitude Northern hemisphere, and relatively low number of observed ENSO events create some difficulty in measuring the effect in observational data. The ENSO–Europe teleconnection begins with anomalous
25 convection in the tropical Pacific, and during El Niño events this leads to increased divergence in the upper troposphere, creating a Rossby wave source (Hoskins and Karoly, 1981). The anomalous Rossby waves propagate to the Northern Pacific where they strengthen the wintertime Aleutian Low. There are multiple possible connections between the North Pacific anomalies and the North Atlantic (Jiménez–Esteve and Domeisen, 2018), with a tendency for a negative North Atlantic Oscillation during El Niño event. For the stratospheric connection, as reviewed by Domeisen et al., (2019), the deepened
30 Aleutian low can lead to upward propagating waves, particularly of wavenumber 1, which travel into the stratosphere and weaken the wintertime stratospheric polar vortex. For strong vortex weakening events, such as sudden stratospheric warmings, anomalies can propagate down to the troposphere and project onto the Northern Annular Mode, and the North Atlantic Oscillation (NAO) (Butler et al., 2014). The result is also a tendency for a negative NAO during El Niño events. Mezzina ~~The opposite is approximately true for La Niña events, but the anomalous response is weaker (e.g. Jiménez–Esteve~~
35 ~~and Domeisen, 2019). Mezzini~~ et al. (2020) suggest the NAO-like patterns that result from ENSO variability are distinct from the NAO, and results~~result~~ from a Rossby wave train from the tropics to the North Atlantic which does not affect NAO variability. Some aspects of the teleconnection are approximately opposite for La Niña events; there is reduced upper tropospheric divergence and resultant Rossby wave, and a shallower Aleutian Low, but the anomalous response is weaker with a less consistent extratropical response (e.g. Jiménez–Esteve and Domeisen, 2019).

40 Both the tropospheric and stratospheric teleconnection pathway can be simulated with climate models of sufficient resolution (Cagnazzo and Manzini, 2009, Bell et al. 2009). Models also allow for large numbers of ENSO events to be simulated, which has revealed non–linearities in teleconnections (Frauen, et al., 2014, Jiménez–Esteve and Domeisen, 2019, Garfinkel et al., 2019). However, for confidence in modelling results, we need an understanding of the deficiencies of models. A fully coupled model needs to correctly represent both the complex dynamics of the ENSO ocean–atmosphere interactions to
45 generate the convective anomalies that drive the teleconnections, and the mean climatology so the anomalies interact with the base state correctly. For example, the convective response of tropical Pacific is dependent on the mean state of the Walker circulation (Bayr et al. 2018). The location and strength of the convective response is then important in controlling the location of the extratropical pressure response (Bayr et al., 2019), and can lead to non–linearities. In addition to the patterns of climatological SSTs, the state of the tropical (e.g. Quasi-Biennial Oscillation (QBO) phase) and extratropical

atmosphere can influence the response of the NAO and polar vortex (Garfinkel et al., 2007), and biases in the subtropical jet can affect the propagation of Rossby waves from the tropics to the extratropics (Li et al., 2020).

The technique of flux correcting SSTs has been used to study the effect of model biases on ENSO dynamics (Spencer et al., 2007, Manganello~~Magnello~~ and Huang, 2009, Dommenget et al., 2014) or seasonal forecasting (Magnusson et al., 2013a, Magnusson et al., 2013b). Empirical corrections are added to the coupling between the ocean and atmosphere to push the model towards the observed climatology. It is possible to use a similar technique on the prognostic atmospheric variables of a model. This bias–correction technique was used by Kharin and Scinocca (2012), and artificially decreased biases were associated with an increase in predictive skill on seasonal timescales. Simpson et al., (2013a, 2013b) used the technique to study the impact of jet latitude bias on the Southern Annular Mode (SAM), although they did not see improvements in the persistence of the SAM when they reduced biases in the jet. When Chang et al. (2019) used a similar bias–correction technique they found an improvement in the North Pacific jet and North American rainfall climatology, and a modest improvement in seasonal forecast skill. Tyrrell et al., (2020) investigated how climatological biases affect~~effect~~ the relationship between the Eurasian snow extent and the wintertime polar vortex, and found that the .~~The~~ strength of the vortex had ~~was found to have~~ only a small effect on its response to a tropospheric forcing, however, the downward propagation of stratospheric anomalies was sensitive to the tropospheric circulation.

In this study we have used a similar bias correction technique to probe the impact of climatological biases on the communication of ENSO anomalies from the tropical Pacific to the North Atlantic and European sector. The technical details of the model and experiments are described in section 2, the results of the bias corrections and ENSO experiments are described in section 3, and a discussion and conclusions are presented in section 4.

2 Data and Methods

We use the ECHAM6 spectral atmospheric model (Stevens et al., 2013), run with a horizontal truncation of T63 and 95 vertical levels with a model top at 0.02 hPa. The bias correction technique follows Kharin and Scinocca, 2012, and is similar to that described in Tyrrell et al., 2020 (T20). It is a two–step process; first, the bias correction terms are calculated in a nudged training stage. The model's prognostic variables – divergence, vorticity, temperature, and log of surface pressure – are all nudged towards ERA–Interim for 30 years, and the nudging tendencies are recorded every 6 hours. Then the nudging tendencies for the divergence and temperature are used to create a year-long climatology of correction terms. This climatology is then smoothed in time with a Gaussian filter with a 25 day window, and it represents the inherent bias in the model's prognostic variables. Secondly, the divergence and temperature ~~climatology of~~ correction terms are~~is~~ added to the free running model as an additional tendency term at each time step. An important difference between the nudged and bias corrected runs is that the bias correction terms are independent of the current model state, so the model can respond to perturbations, whereas during the nudged run the model is tightly constrained to the reanalysis. The technical details of the bias correction are outlined in T20, with two differences for the current experiments. For the training step the model was

nudged to ERA–Interim data from 1979–2009, whereas in T20 only the years 1979–1989 were used. The ~~This did not impact the results, with the~~ resulting bias correction terms being very similar and this did not impact the results. The second difference to T20 was that the only bias correction terms used for this study were the divergence and temperature, rather than

85   the divergence, vorticity, temperature and log of surface pressure. During the training stage all the model's prognostic variables were nudged~~nudge~~ towards ERA Interim, and~~but~~ it was found that using only two of the temperature, divergence and vorticity of the bias correction terms gave the best results for reducing the biases in the winds and temperature. Through testing different combinations we found that bias correcting only the divergence and temperature lead to the biggest decreases in the climatological biases of the control run.

90   The bias corrections were applied on model levels between approximately 850 hPa and 2.6 hPa, and three types of bias correction runs were performed; the troposphere and stratosphere were corrected in FullBC, the stratosphere only in StratBC, and the troposphere only in TropBC (see Table 1 for details). Then ENSO SST forcing experiments were conducted with each of these bias corrected climatologies. To generate the SST pattern we used a regression of the Niño3.4 timeseries and HadISST SSTs from 1979–2009. Only positive regression values between 30°S and 30°N and east of 150°E in the Pacific

95   Ocean were used for the pattern, and the regression values were multiplied by 1.5 to strengthen the response, which corresponds approximately to an El Niño or La Niña forcing magnitude of 1.5K. Climatological SSTs, using HadISST data from 1979-2009, were used for the control (CTRL)~~CTRL~~ run. The ENSO pattern was added to (El Niño), or subtracted from (La Niña) the SST climatology in the tropical Pacific, with climatological SSTs used everywhere else. The ENSO anomaly pattern was kept constant in time, i.e., the anomaly did not vary seasonally relative to the climatological SST. Each

100   experiment was run for 100 years. It should be noted that using a regression to generate ENSO patterns results in symmetric El Niño/La Niña magnitudes, whereas from observations El Niño anomalies tend to be larger than La Niña and have a slightly different structure. This simplification, along with a constant ENSO forcing and climatological SSTs outside the Pacific Ocean basin, has the advantage of reducing the number of controlling parameters when analysing the results of the bias-corrections, which was the main aim of the research. However, the simplifications should be considered when

105   comparing the results with observations, particularly in relation to the intra-seasonal and early winter ENSO-Atlantic connection (e.g. King et al., 2018) that may be driven by SSTs and rainfall away from the Pacific (Ayarzagüena et al., 2018, Abid et al., 2021).

To calculate the ~~root-mean-square error and~~ biases between the model and reanalysis (Figure~~Fig.~~ 1) we use ERA Interim 1979–2009, since that data was used to train ~~for~~ our bias correction scheme. However, when analysing the response to El

110   Niño and La Niña runs we use ~~,~~ the newer ERA5 data ~~is used~~ (Hersbach et al., 2020), including the ERA5 Back Extension (Bell et al., 2020), from 1950 to 2021. To composite~~since it is a longer time series. To divide~~ the data the NINO3.4 index was used for DJF (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php). El Niño years defined as NINO3.4 above 0.9 K (13~~8~~ years), La Niña years NINO3.4 below -0.9 K (16~~8~~ years) and the years between -0.5 K and 0.5 K were defined as neutral years (19~~13~~ years). The slightly stricter threshold of +/- 0.9 K was used to define the

115 El Niño/La Niña years to include only stronger events. The years included in the ERA5 composite are listed in Supplementary Table 1.

3 Results

3.1 Reduced model biases

The climatological biases of the model's wind and temperature vary with latitude, height and season. Here we focus on the
120 extended winter season (November-March, NDJFM) because this is the season when the ENSO teleconnection to the Northern Hemisphere is the strongest. In Figure 1 we show the mean NDJFM biases in the zonal mean zonal wind, zonal temperature and mean sea level pressure. The biases are calculated as 100 years of the model climatology minus the 1980–2009 ERA Interim climatology. The largest tropospheric biases in UZ in the CTRL (Figure 1 a) are associated with the subtropical jet, which is too poleward, but in general the UZ and temperature biases are small in the troposphere, in
125 particular in the tropical troposphere. Figure 1 a and e show that the stratospheric vortex is too weak and warm in the CTRL, and a cold bias at 200hPa (Figure 1 e) in the high latitudes in indicative of a too high extratropical tropopause. The sea level pressure (SLP) biases in the control run (Figure 1 i) show an annular pattern of low pressure around 60°-80°N, and high pressure to the south of that over Africa and Asia and the north Pacific, which results in a weak Aleutian Low. Supplementary Figure 1 shows that the spatial pattern of biases in the geopotential height at 300hPa is very similar to that in
130 SLP suggesting that the biases are nearly barotropic.

The bias corrections are applied at different model levels, hence, the biases in control run are reduced to different extents in the three bias corrected runs. The bias in the subtropical jet is reduced in the FullBC (Figure 1 b) and TropBC (Figure 1 d) runs but not in StratBC (Figure 1 c). In the stratosphere the bias towards a too weak and too warm stratospheric polar vortex is almost completely removed in FullBC, and reduced in StratBC. In TropBC the bias in stratospheric UZ is worsened
135 slightly (Figure 1 d), however, the stratospheric temperature bias is slightly improved compared to CTRL (compare Figure 1 e and h). The tropopause temperature bias in the high latitudes is reduced in FullBC and TropBC, at the expense of introducing cold anomalies in the tropical upper troposphere (Figure 1 f, h). The Aleutian Low anomaly is reduced in FullBC (Figure 1 j) and TropBC (Figure 1 l), but not in StratBC (Figure 1 k), and the reduction in bias, like the bias itself, is barotropic as shown in Supplementary Figure 1. Note that the strong westerly bias in the equatorial stratosphere is not
140 corrected. The bias is associated with the Quasi-Biennial Oscillation (QBO), internally generated in our model, and the lack of bias correction is partly due to the fact that our approach with annually varying bias-correcting tendencies is not optimized for the QBO, as discussed in Karpechko et al. (2021).

Overall, the bias correction technique is effective at reducing biases throughout the atmosphere, and the different bias correction experiments allow us to isolate biases in various atmospheric features such as the polar vortex, subtropical jet and
145 the Aleutian Low, which are relevant for ENSO teleconnections to the high latitudes.

The climatological biases of the model's wind and temperature vary with latitude and height. In Fig. 1 we show the root mean square error (RMSE) of the zonal mean zonal wind (UZ) for three regions; the tropics (20°S – 20°N), mid–latitudes (20°N – 50°N) and the polar region (50°N – 90°N; the polar region is extended to 50°N to capture the full extent of the stratospheric polar vortex). The RMSE is calculated relative to the 1980–2009 ERA Interim climatology. It should be noted that the control run (green dashed line) performs very well in many regions, particularly the tropics, with UZ RMSE values below 2 m s^{-1}. The errors are small in the lower troposphere, and tend to increase with height. In Fig. 1 panels d–e we show the percentage change in the UZ RMSE, between the CTRL run and the three bias–corrected runs. These figures demonstrate the effectiveness of the different height–dependent bias correction experiments. For example in Fig. 1e, in the mid–latitudes the corrections in the TropBC experiment reduce the UZ RMSE by around 40% in the troposphere. Although the corrections stop at around 100 hPa their effect continues into the stratosphere, but weakens. As the tropospheric corrections weaken, the corrections in StratBC lead to around a 25% decrease in UZ RMSE in the upper stratosphere. In the mid-latitudes FullBC the UZ RMSE is reduced throughout the atmosphere with reductions of over 50% in the upper troposphere–lower stratosphere region. The same is somewhat true for the tropics and polar regions, although with some exceptions such as the tropical stratosphere, where errors are increased in the FullBC and TropBC, which may be related to the QBO. In the polar atmosphere, the improvements in the FullBC are mostly seen between 100-200 hPa and in the upper stratosphere. The TropBC also leads to increased errors in the polar stratosphere, hence, tropospheric bias-corrections may degrade the mid-stratosphere. Additionally, in the StratBC experiment there is no significant improvement below 20hPa (midlat) & 50hPa (polar), even though the bias corrections are applied from 100hPa.

The SLP bias is shown in Fig. 1 g–j. Here we see similarities between the CTRL and StratBC, with a large bias in the Aleutian Low, and the FullBC and TropBC, where that Aleutian low bias is reduced. The reduced bias in surface pressure, and structure of the pressure field is a barotropic feature, as shown in the geopotential height at 300hPa in Supplementary Figure 3. Additional details of the bias corrections are included in Supplementary figure 1, where we show the zonal mean zonal wind, and Supplementary figure 2, the zonal mean temperature. Overall, the bias–correction technique is effective at reducing errors in the climatological zonal winds throughout the atmosphere, with similar results for the temperature. The reductions in biases are larger in the extra–tropics than tropics, and the vertical profile of corrections can be effectively controlled.

3.2 Teleconnection response to ENSO

We trace the path of ENSO anomalies from the tropical Pacific to the northern hemisphere polar regions and the North Atlantic. Following from Fig. 3 in Jiménez–Esteve and Domeisen, 2019, our Figure~~Fig.~~ 2 shows the anomalous response of indices chosen to highlight the ENSO teleconnection to the North Atlantic. The El Niño and La Niña forcing does not vary seasonally in our experiments, thus is not shown. ERA5 values are included for reference (see Supplementary Table 1 for years included in the ERA5 composites), although direct comparison with the model runs is difficult due to the idealized experimental setup. The convective response of the tropical atmosphere to SST anomalies is represented by the meridional divergent wind at 100 hPa defined in the region 0°–20°N, 160°–220°E (Figure 2 a~~Fig. 2a~~). As expected, the positive anomalies for El Niño are greater than the negative La Niña~~o~~ anomalies, however, the ERA5 anomalies are more symmetric ~~ad actually slightly larger~~ for La Niña. We also see there is no significant difference between the experiments. This is not surprising given the small biases in the tropical troposphere~~atmosphere~~, and reasonably small improvements in the tropics

between the control and bias corrected experiments (Figure 1). The anomalous divergence creates a Rossby wave that leads to a deepening (El Niño) or weakening (La Niña) of the Aleutian low. We measure this using an Aleutian Low Index (ALI), defined as the SLP between 35°–60°N, 180°–240°E, indicated by the green box in Figure 1 iFig. 1 g. The response of the ALI is proportional to the tropical divergence, with the anomalous negative El Niño response being greater than the positive La Niña response. Again, there are no clear differences in the anomalous response between the different climatologies. However, in contrast to the tropical regions, the FullBC and TropBC runs have reduced Aleutian Low biases compared to the CTRL and StratBC runs (Figure 1 i–lFig. 1 g–j), implying that the response of the Aleutian Low to an ENSO signal is not dependant on model biases. The modelled ALI anomalies have a greater magnitude thanare very similar to ERA5.

The next step in the teleconnection is the response of heat flux at 100 hPa, 45°–75°N (HF). The anomalous HF for both El Niño and La Niña shows differences between the bias corrected runs. For an El Niño forcing, the CTRL and StratBC runs show an increase in HF with significant values (indicated by black crosses) for the DJF and JFM three–month means, whereas the FullBC and TropBC anomalies are about half as strong and have no significant values. For the La Niña forcing all the models show a negative anomalous HF, with the absolute value of the anomaly being weaker than the El Niño response. The positive HF response to El Niño in the FullBC and TropBC was about half as strong as the CTRL and StratBC response, whereas for the La Niña response the CTRL and TropBC anomalous negative HF was around half the value of the StratBC and FullBC runs. There were no significant HF values at the 5% level for any of the La Niña experiments. The lack of significance could be partly due to the weaker Rossby wave source associated with a La Niña, and the high variability of the HF. The response of HF in ERA5 exhibits more variability than that in the model. The ERA5 HF El Niño anomalies peak in JFM, where they are larger than in the all model runs. La Niña HF anomalies in ERA5 are negative in early winter and change to positive in late winter, with the late winter positive response being of opposite sign to what is seen in our model. To test if this was due to sampling uncertainty in observations we subsampled our model for periods of 16 years, to match La Nina events in ERA5, and the subsampling was performed 500 times, as shown in Supplementary Figure 2. The HF was used since this was a controlling factor in the stratospheric response, and the HF value for Jan-Feb-Mar is shown, since this differed the most between ERA5 and the model. It was found that the observed values of the La Niña HF are within the sampling uncertainty of all model runs, indicating that there is no evidence that the models do not capture the observed stratospheric pathway of ENSO teleconnections.

The next step in the teleconnection is the response of heat flux at 100 hPa, 45°–75°N (HF). The anomalous HF for both El Niño and La Niña shows differences between the climatologies. For an El Niño forcing the CTRL and StratBC runs show an increase in HF with significant values (indicated by black dots) for the DJF and JFM three–month means, whereas the FullBC and TropBC anomalies are about half as strong and have no significant values. The results are almost opposite for the La Niña forcing, except the CTRL and TropBC anomalous HF was around half the value of the StratBC and FullBC runs, and there were no significant HF values for any of the La Niña experiments. The lack of significance could be partly due to the weaker Rossby wave source associated with a La Niña, and the high variability of the HF. All El Niño model runs

We measure the response of the stratospheric polar vortex with the zonal mean zonal wind at 60N and 10 hPa (UZ60). UZ60 is well predicted by the HF values. Namely, for an El Niño forcing the CTRL and StratBC experiments show a significant weakening, whereas the response of FullBC and TropBC only show a slight weakening of the vortex (Figure 2 dFig. 2d). Likewise with the La Niña forcing the FullBC and StratBC show a greater strengthening than the CTRL and TropBC. These results are also true for the lower stratosphere as measuredmeasure with the zonal mean zonal wind at 60N and 100hPa (Figure 2 e). The Fig. 2e). We also show the seasonal mean of the full zonal mean wind response is shown in Supplementary figure 34. As with the ALI response, the response of the polar vortex does not appear to be affected by biases in the strength of the vortex, and is instead fully explained by the heat flux response, which is discussed further in the next section. Both the El Niño and La Niña Comparing the late-winter response in ERA5 changes sign throughout the extended winter season at both 10hPa and 100hPa, which is not seen in the models. However, the observed values are not inconsistent with zero in any season, suggesting that the observational records may be too short to constraint the sign of the stratospheric pathway of the ENSO teleconnectionto the models, the CTRL and StratBC anomalies are similar to ERA5 at 10hPa, but the FullBC and TropBC anomalies are similar at 100hPa.

The differences in the magnitude of the stratospheric responsesstratospheric response are not mirrored in the response of the NAO, despite the well-known connection between the vortex and the NAO and the importance of a realistic stratosphere for the ENSO–North Atlantic teleconnection (for example, as shown by Cagnazzo and Manzini, 2009, with an older version of the model used in with a similar model to our study). There is a weaker FullBC response in early winter UZ60 to La Nina, and correspondingly weaker NAO response in early winter. However, for the other runs the strength of the polar vortex anomaly has no clear connection with the response of the NAO. This is evident when comparing the El Niño response of the CTRL and StratBC, to FullBC and TropBC, where the latter two have a weak UZ60 response but a strong NAO response. The NAO response to El Niño in ERA5 is also negative but weaker than the models, and the NAO response shows statistical significance which is not seen in the stratospheric indices for ERA5. ERA5 also shows a weak insignificant positive NAO during mid and late winter.

To determine the reason for the weak connection between the stratospheric anomalies and the NAO in the models, we investigate scatter plots of uz 60N 10hPa and the NAO index, as shown in Figure 3. For this figure we chose to show the variability within each experiment (i.e. each of 100 years of DJF means for El Niño, Neutral and La Niña, and for ERA5, 13 El Niño years, 16 La Niña years and 19 neutral years) to better understand the time-mean sensitivity of the ENSO teleconnection, as well as the sensitivity of the stratosphere-troposphere couplings within different bias corrected runs. The positive regression slopes in Figure 3 show that there is the expected relationship between UZ60 and the NAO for each year within each experiment – that is, a weaker vortex (lower values of UZ60) is associated with a more negative NAO. For each model the large crosses in Figure 3 show the mean value of El Niño years (red crosses), La Niña years (blue crosses) and neutral years (green crosses). There is an indication that the positive correlation between vortex strength and NAO is also

apparent between the mean values of the El Niño, and Neutral/La Niña years, with a much smaller signal between the Neutral and La Niña experiments. However, the large variability within each experiment means that the ENSO-forced difference is relatively small. Although causality is not explained, the figure demonstrates that the stratosphere-troposphere coupling does not play a dominant role in the ENSO-EU teleconnection in our experiments. The effect of a weaker or stronger vortex on the NAO is relatively small compared to variability, and hence, the different polar vortex responses between the bias correction experiments do not translate neatly into different magnitudes of the NAO response. One could also hypothesize that a weaker polar vortex response to El Niño in FullBC and TropBC may cause an NAO response of similar magnitude to that in CTRL and StratBC (Figure 2 f) because of an increased sensitivity of NAO to the stratospheric variability. However, the similarity of the UZ60-NAO correlation coefficients between the experiments does not support this hypothesis. Instead, it reveals that the sensitivity of the NAO to the stratospheric variability remains unchanged by the bias corrections. Figure 3 a shows the NAO and UZ60 for ERA5, divided into ENSO phases. The behaviour is broadly similar to the model, although the regressions for both El Niño and La Niña are not statistically significant. Overall, for our experiments with a strong ENSO forcing, the stratosphere plays only a minor role in the NAO response.

To determine the reason for the weak connection between the stratospheric anomalies and the NAO, we investigate scatter plots of uz 60N 10hPa and the NAO index, as shown in Figure 3. For this figure we chose to show the variability within each experiment (i.e. 100 years of DJF means for El Niño, Neutral and La Niña) to better understand the time-mean sensitivity of ENSO teleconnection as well as the sensitivity of the stratosphere-troposphere couplings to different climatologies. The figure shows that there is the expected relationship between UZ60 and the NAO within each experiment – that is, a weaker vortex is associated with a more negative NAO. There is an indication that this relationship is also apparent between the El Niño, and Neutral/La Niña years, with a much smaller signal between the Neutral and La Nina experiments. However, the large variability within each experiment means that the forced difference is relatively small. Although causality is not explained, the figure demonstrates that the stratosphere-troposphere coupling does not play an important role in the ENSO-EU teleconnection in our experiments. The effect is relatively small compared to variability, and hence, the different polar vortex responses between the bias correction experiments do not translate neatly into different magnitudes of the NAO response. One could also hypothesize that a weakened polar vortex response to El Niño in FullBC and TropBC can cause a NAO response of similar magnitude to that in CTRL and StratBC because of an increased sensitivity of NAO to the stratospheric variability. However, the similarity of the UZ60-NAO correlation coefficients between the experiments does not support this hypothesis. Instead, it reveals that the sensitivity of NAO to the stratospheric variability remained unchanged. Overall, for our experiments the stratosphere plays only a minor role in the NAO response.

3.3 A mechanism for differences in the simulatedDifference in polar vortex El Niño response

To investigate the cause of the different HF response between the experiments it is necessary to consider the effect of the bias corrections. To do that we must consider the absolute values and anomalous response together. In Figure 4 we show the DJF ALI, HF and UZ60, for La Niña, neutralNeutral and El Niño conditions. ERA5 is included to demonstrate the biases. Figure 4a shows the reduced bias in the ALI in the FullBC and TropBC experiments compared to the CTRL and StratBC experiments. The figure also shows the deepening and weakening of the ALI for the ENSO forcing is fairly constant between the bias correction runsclimatologies, hence, the larger biases in the CTRL and StratBC experiments do not impact the response of the AL. In Figure 4b we again see the reduced bias in the FullBC and TropBC experiments for the HF. However,

290 there is a smaller change between neutral and El Niño conditions for the FullBC and TropBC compared to the CTRL and StratBC. In other words, in the model the HF is less sensitive to the deepening of the AL in the bias corrected runs ~~climatologies~~ where the AL is already deeper due to the reduced bias. In ERA5 the AL is deeper, yet responds more like the biased runs~~The same is not true in ERA5~~, although we caution against using this to dismiss the model results, due to the possibility of sampling error (e.g. Supplementary figure 2~~low number of events (around 8–9 per ENSO phase~~).

As shown in Figure~~Fig.~~ 2d, the response of the polar vortex is controlled by the HF response in the different climatologies.
295 Figure 4c reiterates this, and also demonstrates that the biases in the polar vortex do not impact its response to anomalous wave forcing. In neutral conditions the UZ60 biases in the StratBC and FullBC are around 5 m s^{-1} less than the CTRL and TropBC biases. However, for El Niño conditions the FullBC (less bias, stronger vortex) and TropBC (more biased, weaker vortex) have a weak response, and the StratBC (less bias, stronger vortex) and CTRL (more biased, weaker vortex) have a strong response.

300 We have shown that the stratospheric response to an El Niño forcing is partially dependant on model biases, and seems to be related to the sensitivity of the HF to the AL. In Figure~~Fig.~~ 5 we use regressions of HF onto SLP to show how the effectiveness of wave forcing by the Aleutian low changes between neutral and El Niño conditions. Similar to Figure 3, we again consider the variability within each bias-correction and ENSO experiment, to understand the time-mean response. Figure 5 a–d shows the regression of monthly HF onto monthly SLP for the extended winter months (November–March) in
305 neutral ENSO conditions. The areas of SLP that are most strongly associated with HF are the Aleutian low region and Siberia, with weaker connections over Greenland and North America. In neutral years these features are very similar between the control and the bias corrected runs, which means the deeper Aleutian low has not affected its~~it's~~ association with wave driving. Figure 5 e–h is the same regression in years with an El Niño forcing. Rather than testing the difference between neutral and El Niño years, this is now a measure of variability during El Niño years. The connection between SLP
310 in the Aleutian low and HF is lower in the CTRL and StratBC El Niño runs, but is now absent in the FullBC run and very weak in the TropBC runs. Therefore, for an equally large AL anomaly, there would be less wave forcing in the FullBC and TropBC. There appears to be a threshold for the depth of the Aleutian Low, below which any additional anomalies do not result in additional wave forcing. During La Niña years the regression values over the AL region are slightly stronger in the FullBC and TropBC, but this did not lead to differences in the response of the HF (i.e. Figure 4b, La Niña to neutral
315 changes). ~~4 b). The HF was plotted against the AL SLP anomalies to test for a non-linear saturation in the wave forcing by the Aleutian Low (Supplementary Figure 5). For variability within the experiments there is an indication that the relationship between HF and AL breaks down with low AL values, and for the FullBC and TropBC El Niño experiments the relationship actually reverses slightly. Although it does not conclusively explain the differences between the bias correction experiments, it shows that the HF-AL relationship does change in tandem with the absolute value of the AL, and the behaviour is fairly~~
320 ~~consistent within each set of bias-correction experiments for different ENSO forcings.~~

In Figure 6 the HF was plotted against the AL SLP anomalies to look for a non-linear saturation in the wave forcing by the Aleutian Low. For variability within the experiments there is an indication that the relationship between HF and AL breaks

down with low AL values, as indicated by the steeper regression line for La Nina years (blue) and flatter line for El Niño years (red). For the FullBC and TropBC El Niño experiments the relationship actually reverses slightly, shown by the positive regression in Figure 6 c and e (red line). ERA5 does not show changes to the HF-AL between the ENSO composites, however, the correlations are weak so it's difficult to make conclusions based on that data. Although Figure 6 does not conclusively explain the differences between the bias correction experiments, it shows for our model that the HF-AL relationship does change in tandem with the absolute value of the AL, and the behaviour is fairly consistent within each set of bias-correction experiments for different ENSO forcings.

4 Discussion and Conclusions

By applying bias correction terms~~a bias-correction technique~~ to the divergence and temperature tendencies of an atmospheric model we have reduced biases in the tropospheric and stratospheric mean states to create various climatologies. With the different climatologies we have performed idealized ENSO forcing experiments to test the role of biases in ENSO teleconnections. There were only small reductions in the bias in the tropics, and there was no difference in the convective response to ENSO between the bias correction experiments. Likewise, the anomalous response of the Aleutian low to El Niño and La Niña forcing was similar between the experiments, despite ~~the~~ reduced biases in the Northern hemisphere extratropical sea level pressure. Li et al (2020) showed an equatorward jet bias can dampen the response of the Aleutian Low to a tropical Rossby wave source. The jet in the CTRL run has a slight poleward bias (Figure 2 a), which is improved ~~troposphere. Reduced biases in the Aleutian low SLP did, however, lead to differences in the anomalous upward wave flux associated with a deepened low due to an El Niño forcing. A stronger polar vortex~~ in the FullBC and TropBC runs, but this ~~didn't~~~~experiments with stratospheric bias correction did not~~ affect the ~~anomalous~~ response of the Aleutian low~~vortex~~. Hence, reductions in the subtropical jet biases for our model is likely not important for the ENSO teleconnection. We ~~we~~ find that reducing certain climatological biases in the surface pressure and wind speed does not significantly affect the response of the Aleutian low to Rossby wave forcing. Reduced biases in the Aleutian low SLP did, however, lead to differences in the anomalous upward wave flux associated with a deepened low due to an El Niño forcing, so the model's ability to generate a planetary wave flux may be dependent on biases in surface pressure. The response of the polar vortex was ~~also~~ shown to be dependent on the upward planetary wave forcing, and not affected by local biases in the strength of the vortex. A stronger polar vortex in the experiments with stratospheric bias corrections did not affect the anomalous response of the vortex to wave forcing. The ~~However, the model's ability to generate a planetary wave flux may be dependent on biases in surface pressure. Because the~~ NAO response was shown to not be ~~is not~~ sensitive to the stratospheric representation nor the stratospheric ~~or~~ response, we conclude that in our experiments - with an ~~with~~ SST forcing that corresponding to large ENSO events - it is dominated by tropospheric teleconnections. This result appears consistent with ~~that by~~ Bell et al. (2009) who also found that for strong ENSO events the tropospheric teleconnections dominate.

To validate our model results we compare them with ERA5 reanalysis from 1950 to 2021. A threshold of 0.9K in the NINO3.4 region was used to composite large ENSO events in the reanalysis, resulting in 13 El Niño and 16 La Niña events (all years listed in Supplementary Table 1). When comparing the model results to ERA5 the main differences occurred for the HF and stratospheric ENSO response. ERA5 exhibits more variability throughout winter which is not present in the model. However, our tests suggested that sampling errors in the reanalysis (see Supplementary figure 2) may be the reason for the difference. Note that in order to simplify comparison between model experiments we applied a simple constant ENSO SST forcing in all model runs, which is not an optimal approach for comparison with observations. A more realistic experimental setup would be better suited to examine the differences with observed teleconnections. Also, we note that despite the variability of the stratospheric response in ERA5, the observed NAO response was more similar to the model, which again points to dominant role of tropospheric teleconnections for strong El Niño events.

Although one motivation behind artificially bias correcting the model was to investigate how the response to various forcings might improve if the biases were reduced, it should be noted that the ECHAM atmospheric model has already been shown to have a realistic response to an ENSO forcing (Manzini et al. 2006, Cagnazzo and Manzini, 2009). The ENSO teleconnection, the Siberian snow–polar vortex connection investigated in Tyrrell et al. (2020), and the QBO teleconnection investigated in Karpechko et al. (2021) are all relevant to seasonal forecasting, but the bias correction technique is unlikely to be used for operational forecasting. Hence, for these experiments the bias corrections are a tool that is used not necessarily to improve the response relative to observations, but rather to explore the sensitivity of the response to climatological biases.

It was interesting to find that response of the Aleutian low and the stratospheric polar vortex was not affected by the climatological biases that we reduced. These two features are important in the ENSO–to–Europe teleconnection and had large reductions in bias due to the corrections. These features are also both forced by planetary waves; horizontally propagating waves from anomalous convection in the tropical Pacific, or vertically propagating waves from the northern hemisphere troposphere to the stratosphere. Hence, model biases in the depth of the Aleutian low, or the magnitude of the polar vortex winds, do not appear to strongly affect their response to wave forcing.

The control and bias corrected runs differed in the magnitude of wave forcing caused by the deepening Aleutian low due to the El Niño forcing. We theorize that this was due to the relationship between the depth of the Aleutian low and its effectiveness at wave forcing. The two experiments with bias corrections in the troposphere both had a deeper Aleutian low, which was closer to observations. Although the magnitude of ALI anomaly was the same, the runs with a deeper Aleutian low had reduced wave forcing for El Niño conditions. Regressions between SLP and HF showed that lower Aleutian low SLP was associated with a decreasing correlation between Aleutian low SLP and HF. Therefore, we speculate that the reduced wave forcing when the troposphere was bias corrected in the FullBC and TropBC, was due to the lower climatological SLP values in the Aleutian low area. It appears that at some threshold of low values of SLP, further anomalies in the Aleutian low do not result in anomalous upward waves. The opposite was not true for the La Niña conditions, since there appears to be no maximum values where the relationship between HF and Aleutian low SLP changes. Additionally, the

non–linearity in the El Niño/La Niña atmospheric response (e.g. Frauen et al., 2014) means that the La Niña response is smaller, making it more difficult to distinguish robust differences between the climatologies. By artificially bias correcting an atmospheric model, we have shown that some aspects of ENSO teleconnections are very robust to the specific model biases we corrected, while more subtle interactions of anomalies with the basic state can impact the overall response. A deeper understanding of the influence of inherent model biases on teleconnections can guide future model development, and also aid in the physical understanding of these important teleconnections.

Data availability

The climatological means of all model experiments, for the variables used in this paper are available at: https://figshare.com/articles/dataset/ECHAM6_Bias_Correction_ENSO/13311623. The full timeseries is available upon request to Nicholas Tyrrell. ERA-Interim and ERA5 data Copernicus Climate Change Service Climate Data Store (CDS, https://climate.copernicus.eu/climate-reanalysis). The ECHAM6 model is available to the scientific community under a version of the MPI-M license http://www.mpimet.mpg.de/en/science/models/license/. The HadISST SST and sea ice data are available from the U.K. Met Office https://www.metoffice.gov.uk/hadobs/hadisst/.

Author contributions

NLT conducted the model runs, analysis, and wrote the first draft. AYK contributed to the interpretation of the results, and improving the final manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors would like to acknowledge Sebastian Rast, John Scinocca, Slava Kharin, and Michael Sigmond for invaluable technical and scientific help. N. L. T. and A. Y. K. are funded by the Academy of Finland (Grants 333255, 286298, and 294120). Two anonymous reviewers greatly helped to improve the manuscript.

References

Abid, M.A., Kucharski, F., Molteni, F., Kang, I.S., Tompkins, A.M. and Almazroui, M.: Separating the Indian and Pacific Ocean impacts on the Euro-Atlantic response to ENSO and its transition from early to late winter, J. Clim., 34(4), pp.1531-1548, doi: 10.1175/JCLI-D-20-0075.1, 2021

Ayarzagüena, B., Ineson, S., Dunstone, N.J., Baldwin, M.P. and Scaife, A.A.: Intraseasonal effects of el niño–southern oscillation on North Atlantic climate, J. Clim., 31(21), pp.8861-8873, doi: 10.1175/JCLI-D-18-0097.1, 2018.

Baldwin, M.P., Stephenson, D.B., Thompson, D.W., Dunkerton, T.J., Charlton, A.J. and O'Neill, A.: Stratospheric memory and skill of extended-range weather forecasts, Sci., 301(5633), pp.636-640, doi:10.1126/science.1087143, 2003.

Bayr, T., Latif, M., Dommenget, D., Wengel, C., Harlaß, J., and Park, W.: Mean-state dependence of ENSO atmospheric feedbacks in climate models, Clim. Dyn., 50, 3171-3194, doi:10.1007/s00382-017-3799-2, 2018.

Bayr, T., Domeisen, D.I.V. and Wengel, C.: The effect of the equatorial Pacific cold SST bias on simulated ENSO teleconnections to the North Pacific and California, Clim. Dyn. 53, 3771–3789, doi:10.1007/s00382-019-04746-9, 2019.

Bell, C. J., Gray, L. J., Charlton–Perez, A. J., Joshi, M. M., and Scaife, A. A.: Stratospheric communication of El Niño teleconnections to European winter, J. Clim., 22, 4083– 4096, doi:10.1175/2009JCLI2717.1, 2009.

Bell, B., Hersbach, H., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz Sabater, J., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., Thépaut, J-N. (2020): ERA5 hourly data on pressure levels from 1950 to 1978 (preliminary version). Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://cds.climate.copernicus-climate.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels-preliminary-back-extension?tab=overview

Brönnimann, S.: Impact of El Niño–southern oscillation on European climate, Reviews of Geophysics, 45, doi:10.1029/2006RG000199, 2007.

Butler, A. H., Polvani, L. M., and Deser, C.: Separating the stratospheric and tropospheric pathways of El Niño–Southern Oscillation teleconnections. Environmental Research Letters, 9(2), 9, doi:10.1088/1748-9326/9/2/024014, 2014.

Cagnazzo, C., and Manzini, E.: Impact of the stratosphere on the winter tropospheric teleconnections between ENSO and the North Atlantic and European region, J. Clim., 22, 1223–1238, doi:10.1175/2008JCLI2549.1, 2009.

445 Chang, Y., Schubert, S. D., Koster, R. D., Molod, A. M., and Wang, H.: Tendency Bias Correction in Coupled and Uncoupled Global Climate Models with a Focus on Impacts over North America, J. Clim., 32 (2), 639–661, doi:10.1175/JCLI-D-18-0598.1, 2019.

Domeisen, D. I., Garfinkel, C. I., and Butler, A. H.: The teleconnection of El Niño Southern Oscillation to the
450 stratosphere, Reviews of Geophysics, 57(1), 5-47, doi:10.1029/2018RG000596, 2019.

Dommenget, D., Haase, S., Bayr, T., and Frauen, C.: Analysis of the Slab Ocean El Niño atmospheric feedbacks in observed and simulated ENSO dynamics, Clim. Dyn., 42(11–12), 3187-3205, doi:10.1007/s00382-014-2057-0, 2014.

455 Frauen, C., Dommenget, D., Tyrrell, N. L., Rezny, M. and Wales, S.: Analysis of the Nonlinearity of El Niño–Southern Oscillation Teleconnections, J. Clim., doi:10.1175/JCLI-D-13-00757.1, 2019.

Garfinkel, C. I., and Hartmann, D. L.: Effects of the El Niño–Southern Oscillation and the quasi--biennial oscillation on polar temperatures in the stratosphere, J. Geophys. Res.-Atmos., 112(D19), doi:10.1029/2007JD008481, 2007.
460
Garfinkel, C. I., Weinberger, I., White, I. P., Oman, L. D., Aquila, V., and Lim, Y. K.: The salience of nonlinearities in the boreal winter response to ENSO: North Pacific and North America, Clim. Dyn., 52, 4429–4446, doi:10.1007/s00382-018-4386-x, 2019.

465 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz–Sabater, J., ... and Simmons, A.: The ERA5 global reanalysis, Q. J. Roy. Met. Soc., 146(730), 1999-2049, doi:10.1002/qj.3803, 2020.

Hoskins, B. J., and Karoly, D. J.: The steady linear response of a spherical atmosphere to thermal and orographic forcing. J. Atmos. Sci., 38(6), 1179–1196, doi:10.1175/1520-0469-1981-038, 1981.
470
Jiménez–Esteve, B., and Domeisen, D. I. V.: The tropospheric pathway of the ENSO–North Atlantic teleconnection, J. Clim., 31(11), 4563–4584, doi:10.1175/JCLI-D-17-0716.1, 2018.

Jiménez–Esteve, B., and Domeisen, D. I. V.: Nonlinearity in the North Pacific atmospheric response to a linear ENSO
475 forcing, Geophys. Res. Lett., 46(4), 2271-2281, doi:10.1029/2018GL081226, 2019.

Karpechko, A. Yu., N. L. Tyrrell and S. Rast (2021), Sensitivity of QBO teleconnection to model circulation biases, Quart. J. Roy. Meteor. Soc., https://doi.org/10.1002/qj.4014.

480 Kharin, V. V., and Scinocca, J. F.: The impact of model fidelity on seasonal predictive skill, Geophys. Res. Lett., 39(18), doi:10.1029/2012GL052815, 2012.

King, M.P., Herceg-Bulić, I., Bladé, I., García-Serrano, J., Keenlyside, N., Kucharski, F., Li, C. and Sobolowski, S.: Importance of late fall ENSO teleconnection in the Euro-Atlantic sector, Bull. Am. Met. Soc., 99(7), pp.1337-1343, doi:
485 10.1175/BAMS-D-17-0020.1, 2018.

Li, R. K., Woollings, T., O'Reilly, C., & Scaife, A. A. Effect of the North Pacific tropospheric waveguide on the fidelity of model El Niño teleconnections. J. Clim., 33(12), 5223-5237, 2020.

490 Magnusson, L., Alonso-Balmaseda, M., Corti, S., Molteni, F. and Stockdale, T.: Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors, Clim. Dyn. 41, 2393–2409, doi:10.1007/s00382-012-1599-2, 2013a.

Magnusson, L., Alonso–Balmaseda, M. and Molteni, F.: On the dependence of ENSO simulation on the coupled model
495 mean state, Clim. Dyn., 41, 1509–1525, doi:10.1007/s00382-012-1574-y, 2013b.

Manganello, J. V. and Huang, B.: The influence of systematic errors in the Southeast Pacific on ENSO variability and prediction in a coupled GCM. Clim. Dyn., 32, 1015–1034, doi:10.1007/s00382-008-0407-5, 2009.

500 Manzini, E., Giorgetta, M. A., Esch, M., Kornblueh, L., and Roeckner, E.: The Influence of Sea Surface Temperatures on the Northern Winter Stratosphere: Ensemble Simulations with the MAECHAM5 Model, J. Clim., 19 (16): 3863–3881, doi:10.1175/JCLI3826.1, 2006.

Mezzina, B., García-Serrano, J., Bladé, I., Palmeiro, F.M., Batté, L., Ardilouze, C., Benassi, M. and Gualdi, S.: Multi-model
505 assessment of the late-winter extra-tropical response to El Niño and La Niña, Clim. Dyn., 1-22. doi: 10.1007/s00382-020-05415-y, 2020.

Rodríguez–Fonseca, B., Polo, I., García–Serrano, J., Losada, T., Mohino, E., Mechoso, C. R., and Kucharski, F.:. Are

Atlantic Niños enhancing Pacific ENSO events in recent decades? Geophy. Res. Lett., 36(20), L20705,
510   doi:10.1029/2009GL040048, 2009.

Simpson, I. R., Hitchcock, P., Shepherd, T. G, and Scinocca, J. F.: Southern annular mode dynamics in observations and models. Part I: The influence of climatological zonal wind biases in a comprehensive GCM, J. Clim., 26, 3953–3967, doi:10.1175/JCLI-D-12-00348.1, 2013a.
515

Simpson, I. R., Hitchcock, P., Shepherd, T. G, and Scinocca, J. F.: Southern annular mode dynamics in observations and models. Part II: Eddy feedbacks, J. Clim., 26, 5220–5241, doi:10.1175/JCLI-D-12-00495.1, 2013b.

Spencer, H., Sutton, R., and Slingo, J. M.: El Niño in a Coupled Climate Model: Sensitivity to Changes in Mean State
520   Induced by Heat Flux and Wind Stress Corrections. J. Clim., 20, 2273–2298, doi:10.1175/JCLI4111.1, 2007.

Stevens, B., and Co–authors: Atmospheric component of the MPIM Earth System Model: ECHAM6. J. Adv. Mod. Earth Sy., 5 (2), 146–172, doi:10.1002/JAME.20015@10.1002/(ISSN)1942-2466.MPIESM1, 2013.

525   Tyrrell, N. L., Karpechko, A. Y., and Rast, S.: Siberian snow forcing in a dynamically bias–corrected model, J. Clim. 33(24), 10455-10467, doi:10.1175/JCLI-D-19-0966.1, 2020.

Table 1. Experiment details and run names.

Bias corrections	ENSO Neutral	El Niño	La Niña
None	CTRL	CTRL_EN	CTRL_LN
850 hPa to 2.6 hPa	FullBC	FullBC_EN	FullBC_LN
100 hPa to 2.6 hPa	StratBC	StratBC_EN	StratBC_LN
850 hPa to 100 hPa	TropBC	TropBC_EN	TropBC_LN

530

535

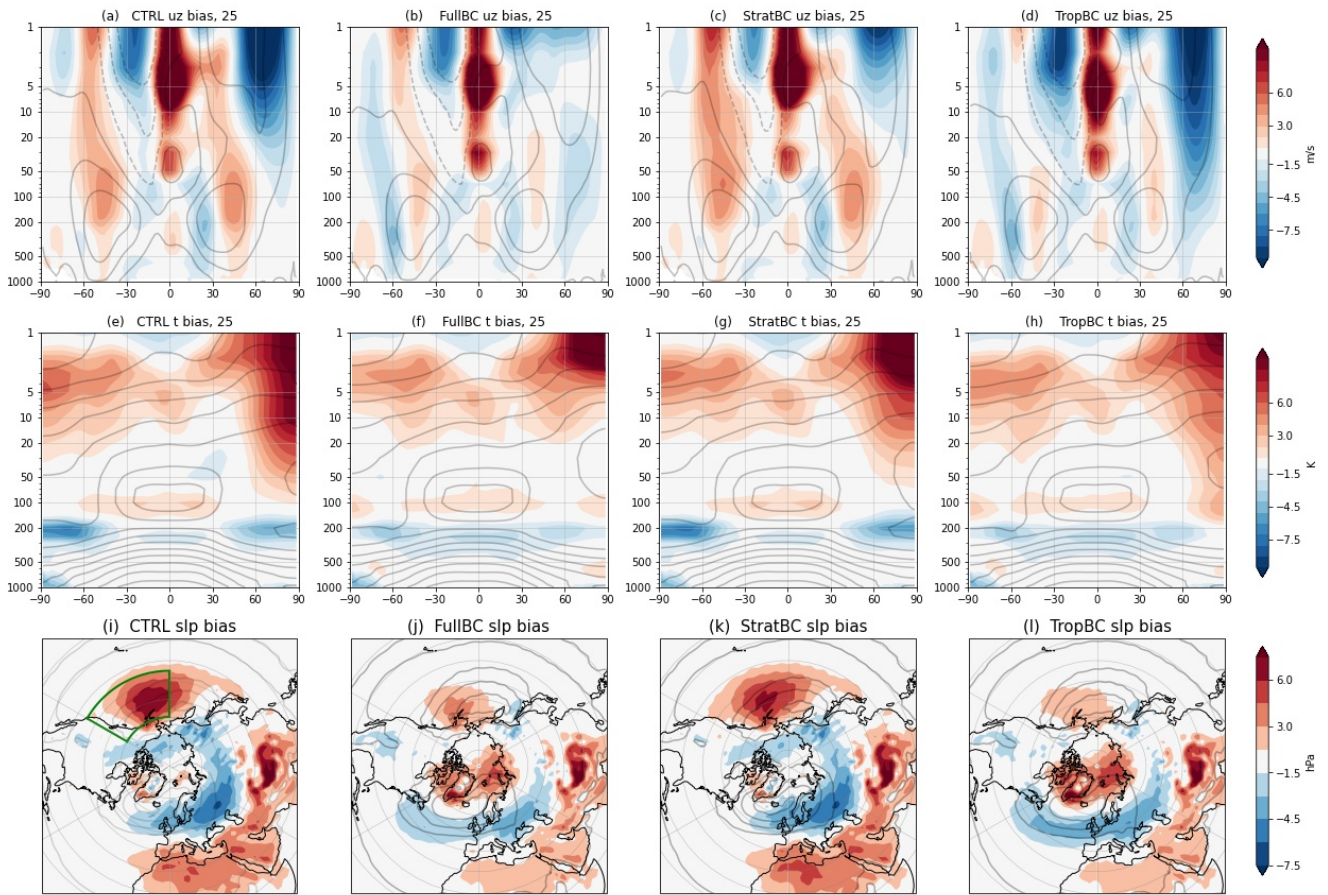



Figure 1: Nov-Mar bias of zonal mean zonal wind (top row), zonal mean temperature (middle row) and MSLP (bottom row) in the four experiments; CTRL (a, e, i), FullBC (b, f, j), StratBC (c, g, k), and TropBC (d, h, l). The bias is calculated as the difference between model and ERA Interim climatology (1979-2009). Grey contours show model climatology. Negative winds in the top rows are marked with dashed lines. Contours are drawn at intervals of 10 m/s for zonal winds (-30, -20, ..., 30 m/s), at 10K for temperatures (200, 210, ..., 290K), and at 5 hPa for SLP (990, 995, ..., 1030hPa). The green box in (i) shows the area of the Aleutian Low Index used in Figure 2, 4 and 6.

Figure 2: Panels a–c shows RMSE of UZ in the four experiment, in the tropics (20°S to 20°N), the mid–latitudes (20°N to 50°N) and the pole (50°N to 90°N). RMSE calculated using difference between model and ERA Interim climatology. Panels d–f are the percentage change of UZ RMSE between the control run and the three bias–corrected runs, with negative values showing improvement in the climatology of the bias corrected runs. Panels g–j show the bias in the control and bias corrected experiments, calculated as the experiment minus ERA interim. The green box in (g) shows area of the Aleutian Low Index.

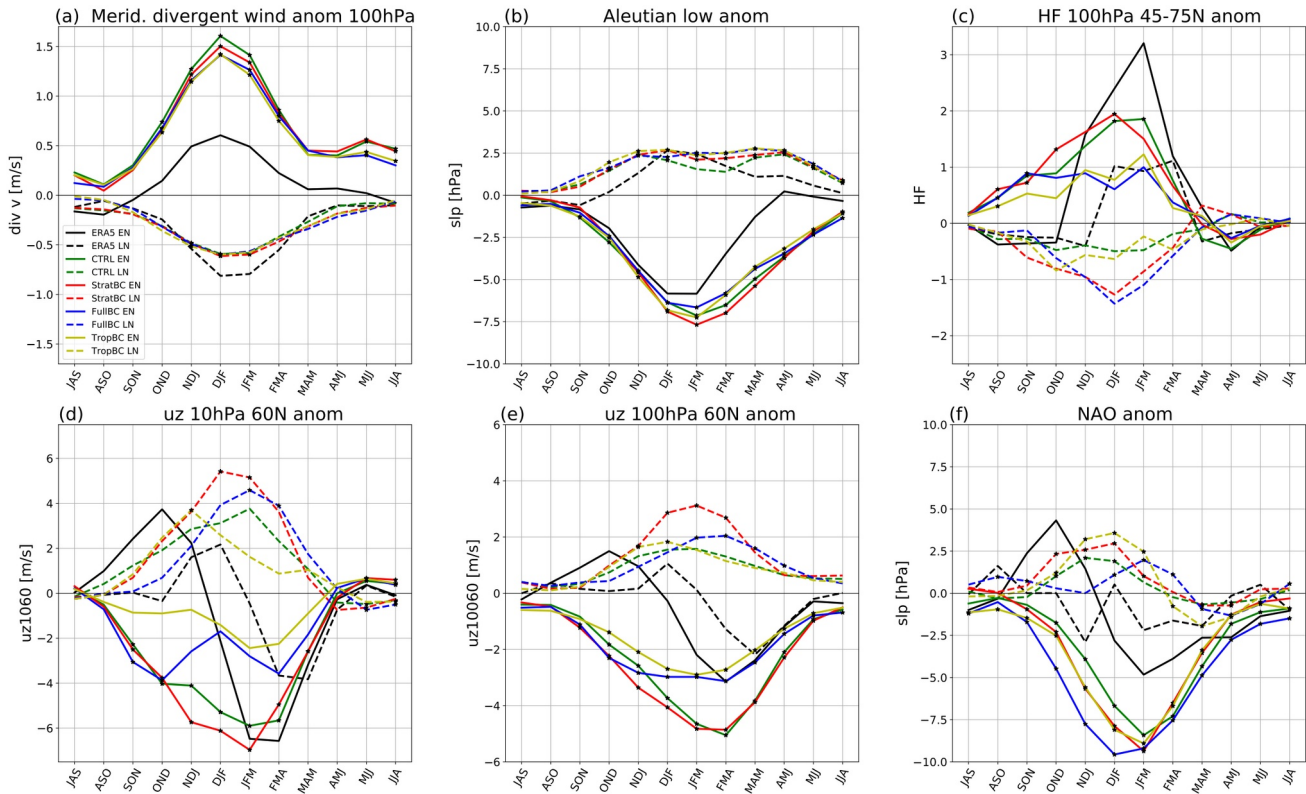



Figure 2. Progression of anomalies from the ENSO region to the stratospheric polar vortex and NAO, for the model and ERA5. Timeseries uses three–month means, and black ~~crosses~~dots indicate significance at p < 0.05 in the model, red crosses are used for ERA5. For model runs solid lines show 100 years of the El Niño run minus 100 years neutral run, dashed lines show 100 years of the La Niña run minus 100 years neutral run. ERA5 data is from 1950-2021 and shows ~~anomalies show~~ the difference between a composite of 138 El Niño (black solid lines) or 16 ~~or~~ La Niña years (black dashed lines), and 1912 neutral years.

~~Figure 3. Scatterplot with regression line of mean DJF values of uz at 60N 10hPa and the NAO index, for (a) CTRL, (b) FullBC, (c) StratBC and (d) TropBC. Neutral years are green, El Niño years are red, and La Niña years are blue. The large crosses indicate the mean value for each El Nino/La Nina/Neutral experiment.~~

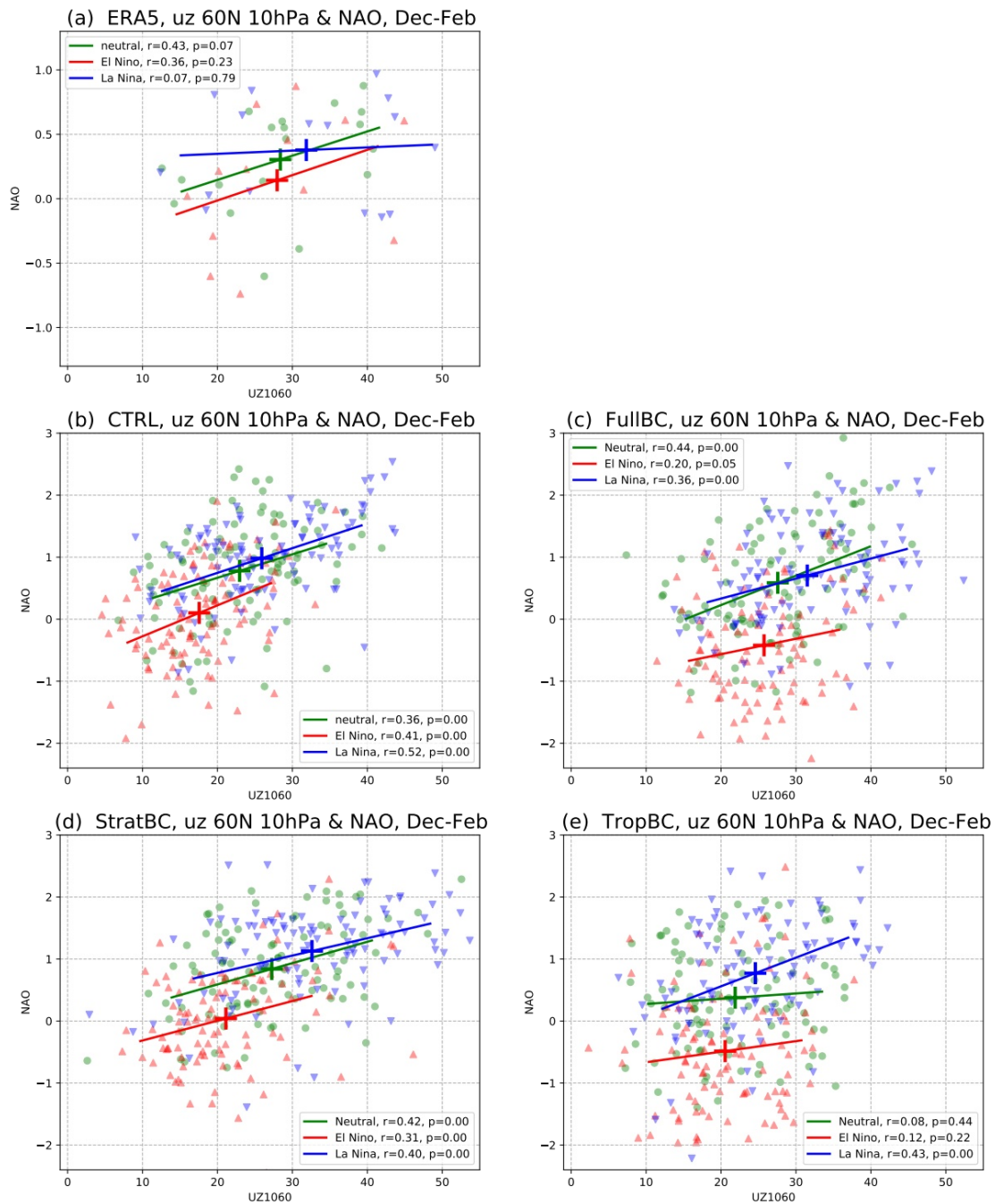



Figure 3. Scatterplot with regression line of mean DJF values of uz at 60N 10hPa and the NAO index, for (a) ERA5 (1950-2021), (b) CTRL, (c) FullBC, (d) StratBC and (e) TropBC. Neutral years are green, El Niño years are red, and La Niña years are blue. The large crosses indicate the mean value for each El Niño/La Nina/Neutral experiment. Correlation and p-values shown in the legend.

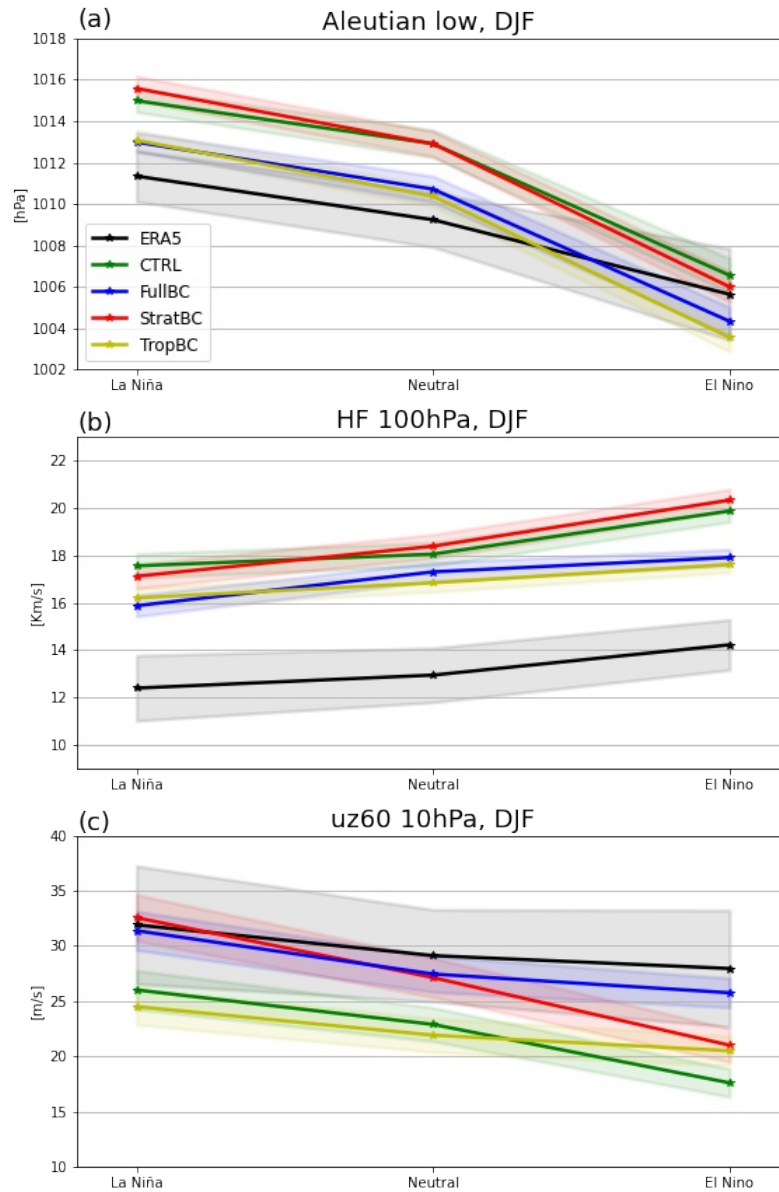
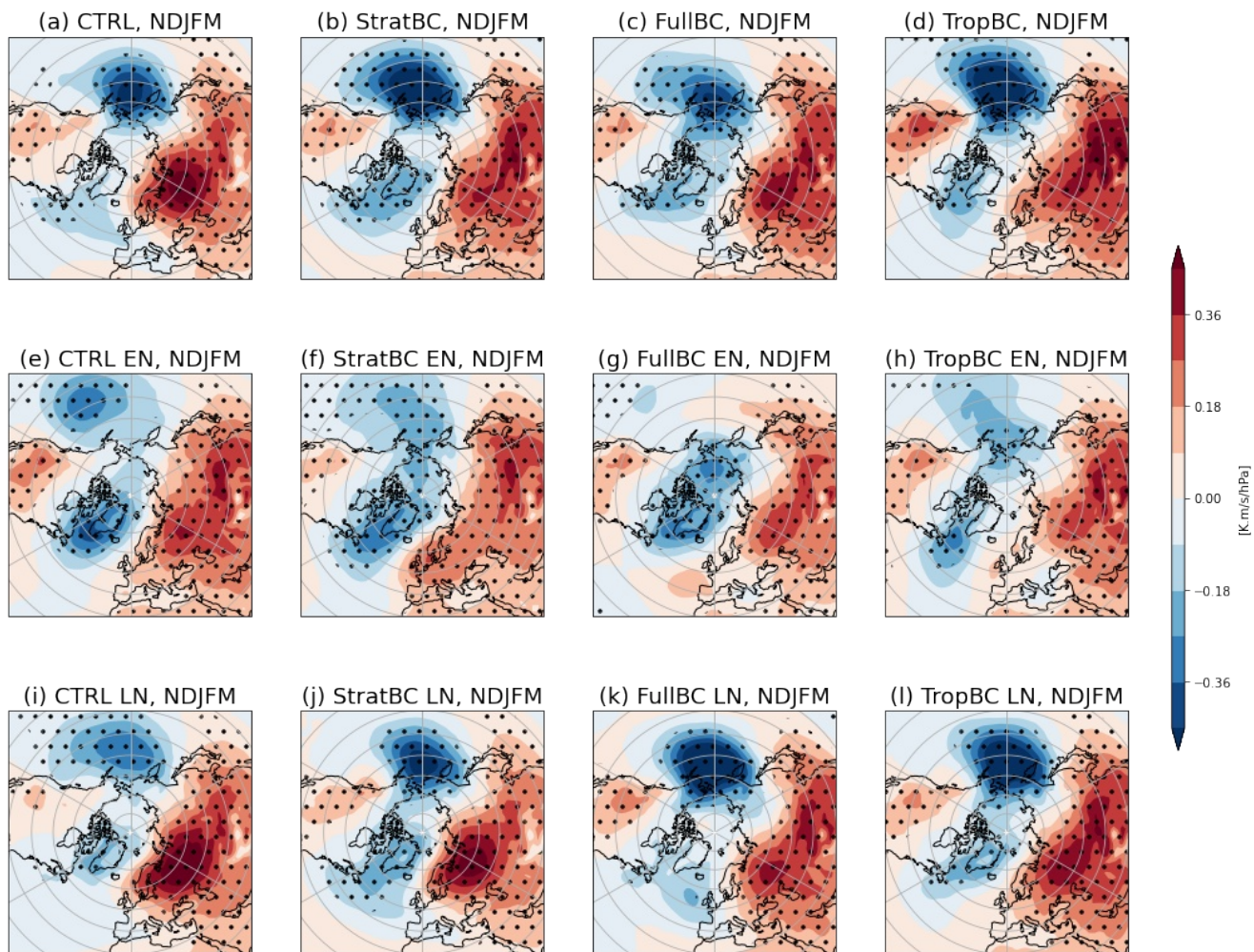



Figure 4. DJF values in the model and ERA5 of (a) Aleutian Low Index, (b) heat flux between 45N–75N at 100 hPa and (c) UZ60 for La Niña, neutral and El Niño years. Shading shows 95% confidence interval.

570

575 | **Figure 5. Regression of monthly HF and monthly sea level pressure for extended winter months (Nov–Mar). Top row is neutral years, middle row is El Niño years, bottom row is La Niña years. Stippling indicates significance at p < 0.05.**

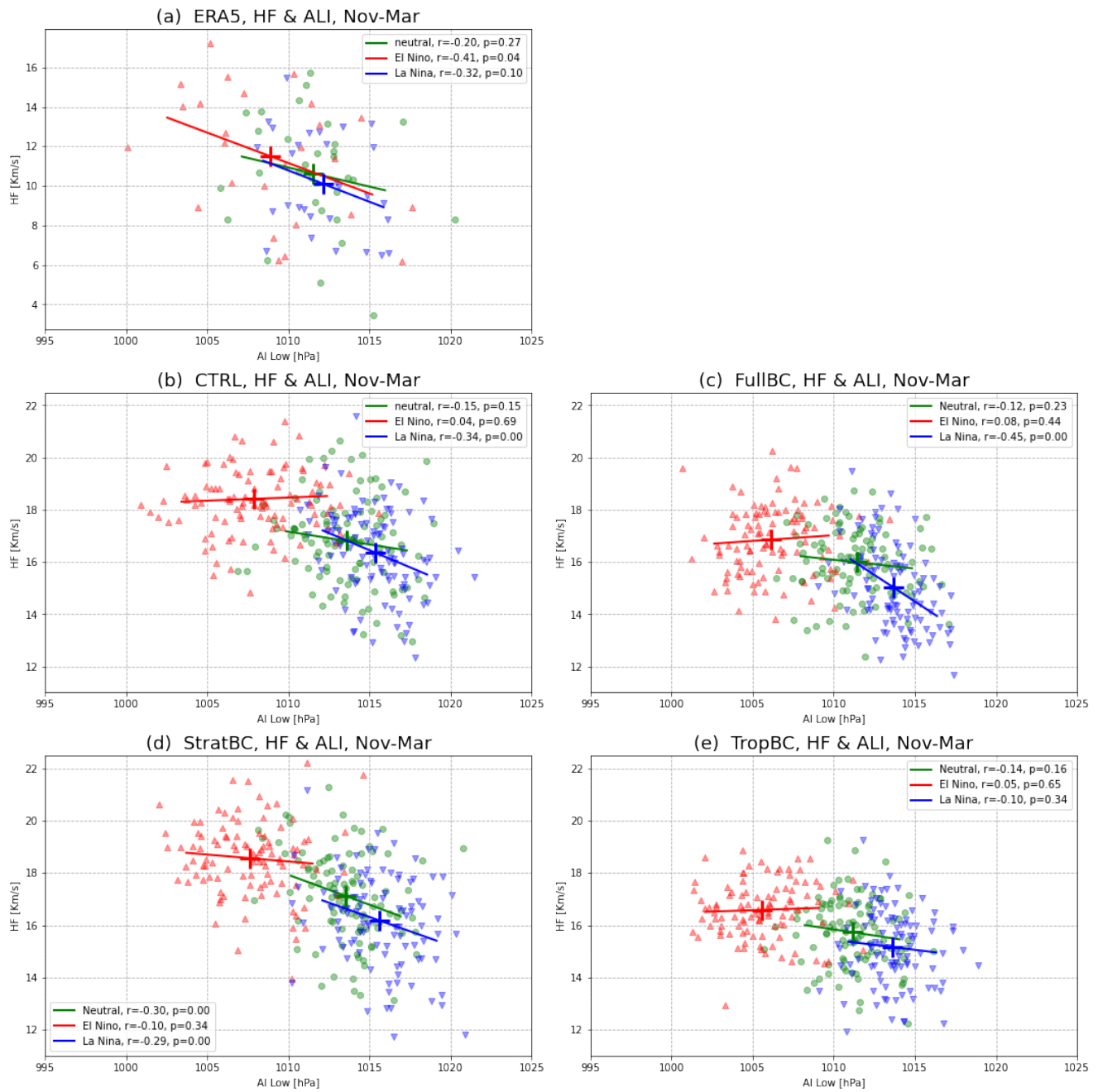



Figure 6. Scatterplot with regression line of mean NDJFM values of the Aleutian Low Index and the heat flux between 45N–75N at 100 hPa, for (a) ERA5 (1950-2020), (b) CTRL, (c) FullBC, (d) StratBC and (e) TropBC. Neutral years are green, El Niño years are red, and La Niña years are blue. Correlation and p-values shown in the legend.

580