Sudden stratospheric warmings during El Niño and La Niña: sensitivity to atmospheric model biases.

Nicholas L. Tyrrell[1], Juho M. Koskentausta[1,2], Alexey Yu. Karpechko[1].

1. Meteorological Research Unit, Finnish Meteorological Institute, Helsinki, 00500, Finland

2. Faculty of Science, University of Helsinki, Helsinki, Finland

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

Abstract.

The number of sudden stratospheric warmings (SSWs) per year is affected by the phase of the El Niño–Southern Oscillation (ENSO), yet there are discrepancies between the observed and modeled relationship. We investigate how systematic model biases in atmospheric winds and temperatures may affect the ENSO-SSW connection. A two-step bias-correction process is applied to the troposphere, stratosphere or full atmosphere of an atmospheric general circulation model. ENSO type sensitivity experiments are then performed by adding El Niño and La Niña sea surface temperature (SST) anomalies to the model's prescribed SSTs, to reveal the impact of differing climatologies on the ENSO–SSW teleconnection.

The number of SSWs per year is overestimated in the control run, and this statistic is improved when stratospheric biases are reduced in both the stratosphere and troposphere. The seasonal cycle of SSWs is also improved by the bias corrections. The composite SSW responses in the stratospheric zonal wind, geopotential height and surface response are well represented in both the control and bias corrected runs. The model response of SSWs to ENSO phase is more linear than in observations, in line with previous modeling studies, and this is not changed by the reduced biases. However, the ratio of wave-1 events to wave 2 events as well as the tendency to have more wave trend of more wave-1 events during El Niño years than La Niña years is improved in the bias corrected runs.

1 Introduction

The El Niño-Southern Oscillation (ENSO) can impact the northern hemisphere wintertime stratospheric variability, and the prevalence of sudden stratospheric warmings (SSWs). Understanding the ENSO-SSW link can help interpreting seasonal model predictions and improve seasonal forecasts. The increased convection in the tropical east Pacific during an El Niño event triggers a Rossby wave train that strengthens and deepens the Aleutian low (Bell et al. 2009; Cagnazzo and Manzini 2009). This leads to constructive linear interference of the planetary waves and an increased wave flux into the stratosphere, and hence, a weakened stratospheric polar vortex. During El Niño years the polar vortex is, on average, weaker than in neutral years, and El Niño is also associated with an increase in the number of SSWs (Domeisen et al., 2019).

Although La Niña is the opposite phase to El Niño, the negative SST anomalies tend to be weaker, more westward, and have a different time evolution (Hoerling et al., 1997, Larkin & Harrison, 2002, Frauen et al., 2014). The decrease in convection in the topical east Pacific associated with La Niña still leads to a shallower Aleutian low, decreased wave flux and a stronger polar vortex (Iza et al., 2016, Jiménez-Esteve and Domeisen, ~~, B., & Domeisen, D. I. V.,~~ 2019, Domeisen et al., 2019). The anomalous La Niña response is weaker than El Niño due in part to the weaker response of the tropical convection and Rossby wave forcing (Trascasa-Castro et al., 2019). The changes to the vertical wave activity flux seem a valid dynamical argument as to why El Niño might lead to more SSWs and La Niña lead to less SSWs, however, the observational record is not so clear. There is a higher chance of an SSW during El Niño years, but there is also an increase in SSW frequency associated with La Niña years (Butler et al., 2014). However, there may be sampling errors due to the relatively short observational record (Domeisen et al., 2019), and the La Niña-SSW relationship is sensitive to the SSW definition (Song & Son 2018). Modeling studies show the increased likelihood of an SSW during an El Niño, and show a decreased~~decrease~~ likelihood of SSWs during La Niña years (Polvani et al., 2017, Song & Son, 2018). It is unclear if the discrepancy between models and observations is due to the low number of observed ENSO and SSW events in observations, or non-linearities in the ENSO teleconnections which the models are unable to simulate (Domeisen et al., 2019).

The role of mean state model biases has been investigated for some aspects of the ENSO-SSW teleconnection. Biases in the tropical Pacific SSTs can lead to different ENSO dynamics (Bayr et al., 2018), and affect the position of the North Pacific sea level pressure response (Bayr et al., 2019). Mean state biases in the extratropical circulation can affect the propagation of Rossby waves (Li et al., 2020), and their impact on North Pacific SSTs (Dawson, et al., 2011). The impact of climatological biases on the mean ENSO-to-northern hemisphere teleconnection was discussed in Tyrrell and Karpechko (2021), using output from the same modeling experiments as in this paper (see Section 2). It was found that mean state of the Aleutian low changed the response of the polar vortex to an El Niño forcing by modulating the upward wave flux to the stratosphere. Biases in the strength of the polar vortex did not impact its anomalous response to ENSO, and the NAO response was not impacted by biases.

In this paper we investigate how the climatological biases in atmospheric winds and temperatures affect the relationship between ENSO and northern hemisphere SSWs. We use a bias correction technique to reduce atmospheric biases at specific levels to create different climates, within which we can run ENSO-like SST perturbation experiments. The bias correction technique and data are described in Section 2, in Section 3 we present the bias reductions and mean ENSO response (3.1), the statistics of SSWs (3.2), downward propagation and the surface response (3.3), and the heat flux response (3.4). A discussion and conclusions are presented in Section 4.

2 Data and Methods

2.1 Bias corrections

We used the ECHAM6 atmospheric model (Stevens et al., 2010), with a horizontal truncation of T63 and 95 levels in vertical with a model top at 0.02 hPa. It was run in bias-corrected and biased modes and with SST perturbation experiments. The bias correction process follows Kharin and Scinocca (2012), and has been used to study the effects of model biases on the Eurasian snow extent-polar vortex connection (Tyrrell et al., 2020), Quasi-Biennial Oscillation teleconnections (Karpechko et al., 2021) and the ENSO-northern hemisphere winter teleconnections (Tyrrell and Karpechko, 2021) and involves two steps: first, the dynamic variables of the model (divergence, vorticity, temperature, and log of surface pressure) are nudged towards ERA-Interim reanalysis. During this step the nudging tendencies are recorded every 6 hours. Forty years of nudging tendencies are then composited and smoothed to create an annual climatology of the nudging tendencies. This climatology represents the inherent biases in the model. In the second step, the nudging tendency climatology is added to the model as an additional tendency at each timestep, in order to correct the biases in the model's climatology. For the second step it was experimentally found that the biggest reduction in biases occurred when only the divergence and temperature were corrected. The dynamic variables of ECHAM6 are solved using a spectral decomposition of the globe, which allows for nudging and bias correcting on specific wavenumbers. Wavenumbers below n = 21 were nudged and corrected, which means that features below about 1000 km were not corrected. The bias corrections can also be applied at different height levels, and three experiments were performed with bias corrections in the troposphere only, TropBC, stratosphere only, StratBC, and full atmosphere, FullBC (details in Table 1). The critical difference between the nudged and bias corrected runs is that when the model is nudged it is very tightly constrained towards observations, whereas when the bias correction tendencies are applied the model can still respond realistically to perturbations. Additional details of the bias correction scheme are available in Tyrrell et al., (2020) and Tyrrell and Karpechko, (2021).

2.2 El Niño and La Niña experiments

Simplified ENSO SST sensitivity experiments were performed using the bias corrected climatologies as described in Tyrrell and Karpechko, (2021). For the ENSO SST pattern we used a regression of the Niño3.4 time series and HadISST SSTs from 1979–2009. Only the positive regression values between 30°S and 30°N and east of 150°E in the Pacific Ocean were used, and the regression values were multiplied by 1.5 to strengthen the response, corresponding to an El Niño or La Niña forcing magnitude of 1.5K. There are also ENSO related SST anomalies outside the tropical Pacific which were excluded from the perturbed SSTs forcing. Although they can be important for some ENSO teleconnections, they are primarily a response to the tropical Pacific forcing and occur at time lag (Tyrrell et al., 2014), so they were excluded to reduce the complexity of the forced ENSO signal. Climatological SSTs using HadISST data from 1979-2009 were used outside the tropical Pacific, and for the control run (CTRL). The ENSO anomaly was kept constant in time, i.e., the anomaly did not vary seasonally, and each experiment was run for 100 years.

The ERA5 Reanalysis data from 1979-2019 (Hersbach et al., 2020) was used as a reference to compare to the model results. El Niño and La Niña years were defined by the DJF value of the Oceanic Niño of ERSST.v5 SST anomalies in the Niño 3.4 region (5N—5S, 120—170W), from the NOAA CPC website (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php, last access: 16 June 2021), and using a threshold of +/- 0.5K, see Supplementary Table 1.. This results in 13 El Niño years with 9 SSWs, 13 La Niña years with 11 SSWs, and 13 neutral years with 6 SSWs. The relatively low number of El Niño/La Niña years and SSWs means that few of the reanalysis ENSO results have statistical significance, and they may be dependent on the temperature threshold for defining ENSO events. As such, the reanalysis is included as a reference, but a more in depth analysis focusing on ERA5 – and other observational data sets – would be required to fully verify and explain those results.

The SSW central date is defined using the Charlton-Polvani criteria (Charlton and Polvani, 2007), defined as the first day when zonal mean zonal wind at 60N and 10hPa (uz_{60}) is easterly (i.e. $uz_{60} < 0$ m/s). The reversal has to occur during 1st November - 31st March. After an SSW has been detected, winds must return to westerlies for 20 consecutive days before another SSW is detected (as in Butler et al. 2017) to avoid multiple detection of the same event, and uz_{60} must return to westerlies for at least 10 consecutive days before 30 April to exclude final warming.

3 Results

3.1 Reduced model biases and mean ENSO response

The bias corrections are applied globally at different pressure levels. The reductions in biases have a three-dimensional structure which has relevance to the ENSO teleconnection to the stratospheric vortex and the Northern hemisphere, and this was explored in Tyrrell and Karpechko (2021). As this paper focuses on SSWs, the reduced model biases in the wintertime polar vortex are of particular interest. In Figure 1a we show the seasonal progression of uz_{60} using the mean daily values for

the 100 year model runs, and 41 years of ERA5 data. The standard deviation for ERA5 and CTRL is also shown as shading. The CTRL run (blue) has a too weak vortex compared to ERA5 from October to January, and this bias is reduced by

115 approximately half in the FullBC and StratBC runs. The bias corrections in TropBC actually increase the bias in the polar vortex in November-December. All model runs effectively capture the polar vortex strength during February and March. As shown in Figure 1 b the interannual variability of the vortex strength is relatively well simulated in CTRL, and the bias corrections do not significantly change the variance. The largest difference between the reanalysis and the model is in January when ERA5 exhibits increased variance, which is not simulated by any of the model runs. The mean difference in

120 daily uz_{60} between El Niño and neutral years, and La Niña and neutral years is shown in Figure 1 c and d for the model (i.e. daily mean of 100 El Niño or La Niña years minus 100 neutral years), and ERA5 (15 El Niño years, 13 La Niña years, minus 13 neutral years). The CTRL, FullBC and StratBC runs have the strongest mean El Niño response throughout winter, although only the StratBC has a statistically significant response in January (as shown at the 5% level by bold lines). The CTRL and FullBC show a weaker response in January. The TropBC has only a weak El Niño response throughout winter.

125 The CTRL has the weakest La Niña response and the StratBC the strongest with a persistent response from December to March. The mean daily ERA5 response to both El Niño and La Niña shows large variability with little significance in the response. For certain months the ERA5 response is opposite to that seen in the models, e.g. the Febuary-March La Niña response, and at times it shows a similar magnitude and sign, e.g. the La Niña response in January or the El Niño response in March. The mean ENSO response was studied in more detail in Tyrrell and Karpechko (2021), where seasonal mean values

130 indicated that in early winter the models and ERA5 disagreed on sign of the El Niño response and agreed on the La Niña, then in late winter they agreed on the El Niño response and disagreed on La Niña.

Before analyzing SSW responses we assess the ability of the model to capture the timescales of variability. This is explored in Figure 2, following Fig. 1 from Baldwin et al., 2003. Using the geopotential height averaged over the polar cap (60N –

135 90N) (Zcap), and at pressure levels from 1000hPa to 1hPa, the figure shows the time in days when the autocorrelation function drops to 1/e. The day-to-day variability is smoothed with a Gaussian filter (σ = 26 days). The CTRL run captures the timescales of the variability in the winter stratosphere reasonably well, but it is slightly too weak in early winter. The timescales are shorter in the FullBC, and again in TropBC runs, and are slightly longer in StratBC. In all experiments, the timescales are shorter in El Niño experiments and longer in La Niña experiments than in the corresponding neutral

140 experiments. The relationship between the strength of the polar vortex and the timescales of variability was tested in Figure 3, which plots the DJF UZ 60N 10hPa against the DJF timescales of variability averaged from 150hPa to 50hPa. Figure 3a shows each ENSO phase for each model separately, so a weaker or stronger vortex strength may be due to the ENSO phase or the bias corrections. A stronger vortex corresponds to longer timescales of variability, and a weaker vortex corresponds to shorter timescales, with a correlation coefficient of $r = 0.62$. We examine this more closely in Figure 3b by averaging each

145 ENSO phase (i.e. the mean of CTRL_EN, FullBC_EN, TropBC_EN, and StratBC_EN, etc. El Niño is up-triangles, La Niña is down triangles and neutral is the circles), and each bias corrected run (i.e. the mean of FullBC, FullBC_EN, FullBC_LN.

Colored crosses). We see that as the vortex strengthens and weakens by ENSO phase, the timescales of variability change accordingly. However, changes to the vortex strength due to the bias corrections do not correspond neatly to changes to timescales of variability. A stronger vortex has weaker dynamical variability and is driven by slow radiative processes, which may explain the vortex-variability timescales relationship between ENSO phases. On the other hand, at least in the case of CTRL-FullBC, the relationship does not hold, because FullBC has both stronger vortex and shorter variability; therefore, application of the bias correction technique may have affected the timescales of the variability.

3.2 SSW statistics

The statistics of SSWs are detailed in Table 1 and Figure 4. The years used for compositing ERA5 are shown in Supplementary Table 1. The number of SSWs per year is overestimated in CTRL (1.12 SSWs/year) in comparison to ERA5 (0.63 SSWs/year). This statistic is made more realistic in FullBC (0.71 SSWs/year), but there is only a small improvement in the StratBC (1.04 SSWs/year) and TropBC (1.10 SSWs/year) runs. The significance of the change in the number of SSWs due to the bias corrections was tested using a Monte Carlo simulation (Figure 4). Specifically, in each experiment, 100 years were randomly chosen with replacement and the number of SSWs in the sample was calculated. The procedure was repeated 10,000 times to obtain a distribution. The 5-95th percentiles of the distributions are shown in Figure 4 with solid error bars. We see that the FullBC is significantly different from CTRL for neutral (green markers) and El Niño years (red markers), while TropBC differs only for El Niño years. The StratBC runs do not differ significantly from the CTRL for any ENSO phase, although in all cases its distributions lie between those of CTRL and FullBC.

The difference between the models and ERA5 was also tested with Monte Carlo simulations. This was done by choosing the same number of years from the model experiments as there were in the corresponding ERA5 samples. For example, 13 years were taken from the El Niño experiments to match the number of El Niño years in ERA5. The 5-95th percentiles from those simulations are also shown in Figure 4 as the dashed vertical lines that extend beyond the solid error bars (since there is greater uncertainty with fewer chosen years). The ERA5 statistics are shown as dashed horizontal lines. Two lines are shown to compare the experiments without an ENSO forcing with ERA5. The green dashed line is the number of SSWs per year for the full ERA5 period (41 years) and the black dot-dashed line is for the 15 neutral years. Although the full period includes El Niño and La Niña years it is not dependant on one phase, and includes a large number of years, whereas the neutral-only years have a small sample size. The model distribution is again estimated by choosing either 41 years (green dashed vertical line which corresponds to the green dashed horizontal line) or 15 years (black dot-dashed vertical line which corresponds to the black dot-dashed horizontal line). The results show that the FullBC is the only run that consistently captures the SSW statistics of ERA5 in all experiments. Note all La Niña experiments are consistent with ERA5 as well as with each other.

~~The statistics of SSWs are detailed in Table 1. The number of SSWs per year is overestimated in CTRL (1.11 SSWs/year) in comparison to ERA5 (0.65 SSWs/year). This statistic is made more realistic in FullBC (0.70 SSWs/year), but there is only a small improvement in the StratBC (1.03 SSWs/year) and TropBC (1.09 SSWs/year) runs.~~ Consistent with previous modeling studies (e.g. Polvani et al. 2017), SSW frequency is increased during El Niño years and decreased during La Niña years in all

model experiments. For all model experiments, except TropBC, the number of SSW during El Niños is nearly twice as large as that during La Niñas. In TropBC, the exceedance is 40%. For ERA5 we find that SSW frequency is increased during both El Niños (0.69 SSWs/year) and La Niñas (0.85 SSWs/year), consistent with previous studies. The years used for compositing ERA5 are shown in Supplementary Table 1.

The seasonal evolution of SSWs frequency is shown in Figure 5. To explore the differences in seasonal evolution more clearly, the number of SSWs in each month is divided by the total number of SSWs for each experiment, similarly for ERA5. This gives the percentage of the annual total SSWs in each month. Compared to ERA5 there is not enough seasonal variation in CTRL, with too many SSWs in November , December, and March, and too few in January and February Feb. The seasonal variation is improved slightly in FullBC, although the seasonal cycle is still underestimated. In StratBC and TropBC the SSW seasonal statistics are not improved as much as in FullBC. In particular, TropBC almost has an inverse of the seasonal relationship of SSWs compared to ERA5, with the most SSWs in November. There are no consistent changes to the seasonality of SSWs with El Niños or La Niñas in the model or ERA5. Note that the model does not have a seasonally evolving ENSO pattern which may affect the seasonality of SSWs in the ENSO experiments but not in experiments with neutral ENSO conditions which have seasonally evolving SSTs. Yet the neutral ENSO experiments similarly lack seasonal SSW variations.

3.3 SSW downward propagation and surface response

Figure 6 shows the SSW composite of normalized Zcap, and then the anomalous values in El Niño an La Niña years relative to neutral years. For ERA5 the ENSO phases are normalized using the standard deviation for all ERA5 years. The CTRL run simulates the downward propagation of stratospheric anomalies after an SSW reasonably well, although it underestimates the tropospheric response in comparison to ERA5. For neutral years (Figure 6 d, g, j, m) the CTRL and StratBC runs have run has the weakest tropospheric response, and the FullBC run has the strongest response which , and is most similar to that ERA5. All runs show a weaker stratospheric response during El Niño years in comparison to neutral years, both before and after SSWs (i.e. Figure 6 e, h, k, n), however this doesn't always correspond and this corresponds to a slightly weaker tropospheric response. This is more clearly shown in Figure 7, which shows a 30 and 90 day time-mean of Figure 6. We see that in the CTRL and FullBC runs the weaker stratospheric response in El Niño years corresponds to a weaker Conversely, the models show a stronger stratospheric response during La Niña years (Figure f, i, l, o). In FullBC and StratBC in particular, this corresponds with a strong tropospheric response (Figure 7 b, c). In the TropBC and StratBC runs (Figure 7 d, e) the change in the stratospheric response due to . The ENSO response in ERA5 varies from the models. During El Niño is not as pronounced (i.e. the SSW response is only slightly weaker than in years there is a stronger response (relative to neutral years). Correspondingly, there is only a small difference in the tropospheric SSW response in TropBC, while in StratBC the response is actually stronger than in neutral years (for the 30 day mean) or very similar (for the 90 day mean) before SSW events, with a slightly stronger stratospheric response and weaker tropospheric response after SSW events. Whereas during La Niña years the normalized Zcap response is weaker before, and stronger after SSW events.

The models show a stronger stratospheric response during La Niña years (Figure 6 f, i, l, o). In FullBC and StratBC in particular, this corresponds with a strong tropospheric response. The ENSO response in ERA5 differs from the models. During El Niño years there is a stronger response (relative to neutral years) before SSW events, with a slightly stronger stratospheric response and weaker tropospheric response after SSW events. Whereas during La Niña years the normalized Zcap response is weaker before, and stronger after SSW events.

The sea level pressure response to SSWs is well represented in all model experiments and is similar across different climatologies, i.e. bias correction does not greatly affect the surface response. Figure 6 shows the composites of absolute sea level pressure anomalies averaged over 30 days after the central dates of SSWs, and the differences between this quantity in El Niño minus neutral years (middle column), and La Niña minus neutral years (right column). A negative Arctic Oscillation (AO) pattern following SSWs is seen in all runs. The negative AO pattern is stronger in La Niña experiments for the FullBC and StratBC runs. These runs both have a stronger La Niña , which relates to the stronger stratospheric Zcap response (in Figure 6 i, o, and Figure 7 h, j), however, there is not a linear relationship between the stratospheric ENSO response and the surface response. TropBC has a smaller La Niña stratospheric response and surface pressure response (Figure 6 l), but CTRL has a large stratospheric La Niña response (Figure 6 f and Figure 7 g) without a surface pressure response (Figure 6 f). Similarly, the and o. The weaker Zcap response in Figure 6 during El Niño years can also be seen in the weaker negative AO response in CTRL and TropBC, but not FullBC or StratBC (Figure 8 e, h, k, n). The 2-meter temperature response was expected to be quite weak in the model runs, since the same climatological SSTs were used for all runs (except SST anomalies prescribed in tropical Pacific in El Niño and La Niña experiments) which dampens the near surface temperature anomalies, however, there was a La Niña – El Niño difference of 0.4K across Eurasia in the monthly averaged 2-meter temperature (not shown).

3.4 Heat fluxHeatflux and wave 1/wave 2 response

We now look at the wave forcing that causes SSWs. Figure 9 shows the SSW composite anomalies for the heat flux at 100hPa, 45N-75N. The black lines show all wave numbers, the red lines are wave 1, and the blue lines are wave 2. Solid lines indicate dates when the composite anomalies are significantly different from zerosignificance at the 90% confidence level. The ratios of wave 1 to wave 2 SSW events are also listed in Table 1, where each event is defined based on the average heat flux for the 10 days preceding an SSW. The CTRL run has too small wave1/wave2 flux ratio of 62/38 compared to the ERA5 ratio of 77/23, i.e. there are too many wave 2 events in CTRL; ERA5 has 0.15 wave 2 SSWs per year, and CTRL has 0.41 wave 2 SSWs per year. This ratio is improved in the bias correction experiments, with the FullBC (76/24) being most similar to ERA5, and a smaller improvement in StratBC (67/33) and TropBC (68/32). As expected, in ERA5 the wave 2 flux is weaker in El Niño years and stronger in La Niña years, hence, La Niña events have a smaller wave 1/wave 2 ratio than El Niño events (e.g. Garfinkel & Hartmann, 2008). This is simulated reasonably well in the experiments, but the relationship is weaker. For all climatologies the El Niño years have a larger wave 1/wave 2 ratio than La Niña years, however, in CTRL the La Niña experiment has a larger wave 1/wave 2 ratio than the neutral experiment. The total heat flux

anomaly before an SSW is smallest in El Niño and largest in La Niña in all climatologies. Since the anomalies are calculated with respect to each experiment's own background flux, which is largest in El Niño and smallest in La Niña experiments, the result explains the larger frequency of SSWs in El Niño and small frequency in La Niña. It happens because during El Niños, an SSW can be induced by a weaker wave activity pulse which happen more frequently; however, a larger wave activity pulse that occurs more rarely is required to induce an SSW during La Niñas. Note that in all experiments as well as in ERA5 the larger flux during La Niñas is due to increased wave 2 contribution; however, only in StratBC the wave 2 increase is larger than that of wave 1, which is also seen in ERA 5.

4 Discussion and Conclusions

The ECHAM6 atmospheric model was run with bias-correcting tendencies added to the temperature and divergence at each timestep. The bias corrections were added at different levels – the stratosphere (StratBC), troposphere (TropBC), or the full atmosphere (FullBC) – to create a range of climates with reduced biases. SST forcing experiments were conducted within these climates by applying a positive or negative ENSO pattern in the tropical Pacific. The seasonal mean response is explored in Tyrrell and Karpechko (2021). In and in this paper we have focused on the relationship between the ENSO forcing and SSWs.

For the years without an ENSO forcing the number of SSWs is overestimated in our control run in comparison with ERA5. This is largely due to the polar vortex being too weak in the CTRL run. When the strength of the vortex is improved in the FullBC the SSW statistics also improve. There is a insignificantsmaller improvement in the StratBC runs, despite the improvement in the strength of the vortex being similar to FullBC. The polar vortex strength was not corrected remains weak in the TropBC run, and there is no significant improvement in the number of SSWs. The lack of stratospheric bias correction in TropBC indicates that the stratospheric biases do not originate in the tropospheric circulation biases but are more likely resulting from orographic and non-orographic gravity wave drag parameterizations (e.g. Eichinger et al., 2020)parameterizations. The seasonal variation of SSWs is too small in the CTRL run compared to ERA5, with too many SSWs in November and March. This is slightly improved in FullBC, but not in StratBC or TropBC. The duration of an SSW is well simulated, i.e. the number of days that UZ < 0 m/s after an SSW is not significantly different from ERA5 in any of the model runs suggesting that it is controlled by basic processes such as radiative relaxation, well represented in the model. Likewise, the downward propagation and surface response is similar between ERA5 and the control run, and not affected by the bias corrections. The ratio of wave 1 to wave 2 events is too small in CTRL, and this is improved in FullBC, and to a lesser extent in StratBC and TropBC.

The ERA5 reanalysis data suggests that there is an increase in SSWs in both La Niña and El Niño years, when compared to neutral years (Table 1). This is based on a fairly low number of events; depending on the threshold used to define ENSO there are around 10-15 El Niño or La Niña years, with around 0.6 – 0.9 SSWs per year. This makes it difficult to statistically confirm the observed ENSO-SSW relationship. Our model results differ from observations and are in line with other

modelling studies, which show a more linear relationship between ENSO and SSWs. The increase in wave 1 events in El Niño years and wave 2 events in La Niña years is captured by the bias corrected runs, but not by the control run. The timescales of variability in the stratosphere were tested with the autocorrelation of Zcap, and it was found that the weakening (El Niño) and strengthening (La Niña) of the polar vortex due to ENSO phases explains~~explained~~ changes to the timescales of the variability. However, similar strength changes to the vortex by the bias corrections did not relate directly into similar changes to the timescales of the variability, suggesting that other factors associated with bias correction procedure affect the timescales.

The impact of ENSO phase on Zcap response to SSWs was fairly consistent amongst the models, with a weaker lower stratospheric Zcap SSW relative response (i.e. less warming/less weakened vortex) during El Niño years, and stronger Zcap relative response (more warming, weaker vortex) during La Niña years. This relationship between ENSO phase and the stratospheric SSW response did not consistently lead to a similar relationship between the MSLP response averaged over 30 days after SSWs amongst the models. All the model runs showed a negative AO response, with FullBC and StratBC having a stronger response in La Niña years and CTRL and TropBC having a weaker response in El Niño years. The composite eddy heat flux ~~HF~~ showed that a larger anomalous wave forcing is required for an SSW to occur during La Niña years, compared to neutral and El Niño years, and the El Niño anomalous wave forcing required to trigger an SSW was slightly smaller than in neutral years. This relationship is associated with the fact that the anomalous forcing is calculated with respect to each experiment's~~experiments~~ own climatology, and there is a larger background wave forcing in the El Niño experiments and a smaller wave forcing in the La Niña experiments. Consequently, a relatively small anomalous forcing is required to induce an SSW in the El Niño experiments, and~~amd~~ a large anomalous forcing is required in La Niña experiments. The additional wave forcing during La Niña years primarily came from an increase in wave 2 events, and likewise the reduced wave forcing during El Niño years was associated with less wave 2 forcing. This result is similar in the ERA5 data, although with more extreme differences in wave 1 and 2 between the ENSO phases. Indeed, in ERA5 La Niña years the magnitude of wave 2 forcing is greater than that of wave 1, which was only reproduced~~only occurred~~ in the StratBC model runs.

Overall, we show that reductions in both stratospheric and tropospherc biases ~~improvements to the strength of the polar vortex~~ can improve the SSW statistics of a model in relation to the number of SSWs per year, and the ratio of wave 1 and wave 2 events. Whether the improvements lead ~~leads~~ to a more realistic ENSO-SSW relationship is unclear given the large uncertainty in the observed statistics.

Data availability

Data availability. The climatological means of all model experiments for the variables used in this paper are available at https://doi.org/10.6084/m9.figshare.13311623.v2 (Tyrrell and Karpechko, 2020). The full time series is available upon request to Nicholas Tyrrell. ERA-Interim and ERA5 data can be found at Copernicus Climate Change Service Climate Data

Store (CDS, https://doi.org/10.24381/cds.bd0915c6, Hersbach et al., 2018). 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/ (Max-Planck-Institut für Meteorologie, 2020). The HadISST SST and sea ice data are available from the U.K. Met Office https://www.metoffice.gov.uk/hadobs/hadisst/ (Met Office Hadley Centre, 2020; Rayner et al., 2003).

Author Contributions

JK conducted analysis and contributed to the manuscript, NLT conducted the model runs and analysis and wrote the first draft. AYK contributed to the interpretation of the results and improving the final manuscript. The reviews from Ronald Kwan Kit Li and an anonymous referee greatly helped to improve the paper.

Competing interests.

The authors declare that they have no conflict of interest.

Acknowledgments

The authors would like to acknowledge Sebastian Rast, John Scinocca, Slava Kharin, and Michael Sigmond for invaluable technical and scientific help.

Financial support.

This research has been supported by the Academy of Finland (grant nos. 286298, 333255, and 294120).

References

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 tele-connections to European winter, J. Climate, 22, 4083–4096, https://doi.org/10.1175/2009JCLI2717.1, 2009.

335 Butler, A. H., Sjoberg, J. P., Seidel, D. J., & Rosenlof, K. H.: A sudden stratospheric warming compendium. Earth System Science Data, 9(1), 63-76, 2017.

Butler, A. H., L. M. Polvani, and C. Deser, 2014: Separating the stratospheric and tropospheric pathways of El Niño–Southern Oscillation tele- connections. Environ. Res. Lett., 9, 024014, https://doi.org/ 10.1088/1748-9326/9/2/024014.

340 Butler, A. H., Polvani, L. M., and Deser, C.: Separating the stratospheric and tropospheric pathways of El Niño–Southern Oscillation teleconnections, Environ. Res. Lett., 9, 024015, https://doi.org/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. Climate, 22, 1223–1238, https://doi.org/10.1175/2008JCLI2549.1, 2009.
345

Charlton, A. J., and Polvani, L. M.,: A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. J. Climate, 20, 449–469, 2007.

Dawson, A., Matthews, A. J., and Stevens, D. P.: Rossby wave dynamics of the North Pacific extra-tropical response to El
350 Niño: Importance of the basic state in coupled GCMs. Climate Dynamics, 37(1–2), 391–405, 2011.

Domeisen, D. I., Garfinkel, C. I., and Butler, A. H.: The teleconnection of El Niño Southern Oscillation to the stratosphere, Rev. Geophys., 57, 5–47, https://doi.org/10.1029/2018RG000596, 2019.

355 Eichinger, R., Garny, H., Šácha, P., Danker, J., Dietmüller, S., and Oberländer-Hayn, S.: Effects of missing gravity waves on stratospheric dynamics; part 1: climatology. Climate Dynamics, 54(5), 3165-3183, https://doi.org/10.1007/s00382-020-05166-w, 2020.

Garfinkel, C. I., and Hartmann, D. L.: Different ENSO teleconnections and their effects on the stratospheric polar vortex, J.
360 Geophys. Res., 113, D18114, doi:10.1029/2008JD009920, 2008.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J. N.: ERA5 hourly data on pressure levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set],
365 https://doi.org/10.24381/cds.bd0915c6, 2018.

Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Dahlgren,

P., De Chiara, G., Dee, D. P., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hogan, R. J., Hólm, E. V., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hoerling, M. P., Kumar, A., and Zhong, M.: El Niño, La Niña, and the nonlinearity of their teleconnections. J. Climate, 10, 1769–1786, doi:10.1175/1520-0442(1997)010, 1997.

Iza, M., Calvo, N., & Manzini, E.: The stratospheric pathway of La Niña. Journal ofClimate, 29(24), 8899–8914, 2016.

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

Karpechko, A. Yu., Tyrrell, N. L., and Rast, S.: Sensitivity of QBO teleconnection to model circulation biases, Q. J. Roy. Meteor. Soc., 147, 2147–2159, https://doi.org/10.1002/qj.4014, 2021.

Kharin, V. V. and Scinocca, J. F.: The impact of model fidelity on seasonal predictive skill, Geophys. Res. Lett., 39, L18803, https://doi.org/10.1029/2012GL052815, 2012.

Larkin, N. K., and Harrison, D. E.: ENSO warm (El Niño) and cold (La Niña) event life cycles: Ocean surface anomaly patterns, their symmetries, asymmetries, and implications. J. Climate, 15, 1118–1140, doi:10.1175/1520-0442(2002)015, 2002.

Li, R. K., Woollings, T., O'Reilly, C., and Scaife, A. A.: Effect of the North Pacific tropospheric waveguide on the fidelity of model El Niño teleconnections, J. Climate, 33, 5223–5237, 2020.

Max-Planck-Institut für Meteorologie: Availability & Licenses, available at: https://mpimet.mpg.de/en/science/models/availability-licenses, last access: 13 January 2020.

Met Office Hadley Centre: Hadley Centre Sea Ice and Sea Surface Temperature data set (HadISST), available at: https://www.metoffice.gov.uk/hadobs/hadisst/, last access: 13 January 2020.

Polvani, L. M., Sun, L., Butler, A. H., Richter, J. H., and Deser, C.: Distinguishing stratospheric sudden warmings from ENSO as key drivers of wintertime climate variability over the North Atlantic and Eurasia. J. Climate, 30(6), 1959–1969. https://doi.org10.1175/JCLI-D-16-0277.1, 2017.

Song, K., and Son, S-W.: Revisiting the ENSO–SSW relationship. J. Climate 31(6) 2133-2143, 2018.

405

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., and Brokopf, R.: Atmospheric component of the MPI-M Earth system model: ECHAM6, J. Adv. Model. Earth Sy., 5, 146–172, https://doi.org/10.1002/jame.20015, 2013.

410   Trascasa-Castro, P., Maycock, A. C., Yiu, Y. Y. S., and Fletcher, J. K.: On the linearity of the stratospheric and Euro-Atlantic sector response to ENSO. Journal of Climate 32(19) 6607-6626, 2019.

Tyrrell, N. L., Dommenget, D., Frauen, C., Wales, S., Rezny, M.: The influence of global sea surface temperature variability on the large-scale land surface temperature, Clim. Dyn., 44, 2159–2176 doi:10.1007/s00382-014-2332-0, 2015.

415

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.

Tyrrell, N. L., and Karpechko, A. Yu.: Minimal impact of model biases on northern hemisphere ENSO teleconnections,
420   Weather Clim. Dynam., 3, 1–13, https://doi.org/10.5194/wcd-3-1-2021, 2021

Table 1. Experiment names and details for bias correction and ENSO experiments. Number of SSWs per year calculated from 100 years in the model experiments, and 41 years of ERA5 data. For ERA5~~ERA-5~~, the SSW frequency in the third column is shown for all years and only for years with a neutral ENSO in brackets. The wave1:wave2 ratio is based on the heat flux~~heatflux~~ at 100hPa, 45N-75N.

425

Bias corrections	Experiment	Neutral		El Niño		La Niña	
		SSW/yr	wave1:wave2	SSW/yr	wave1:wave2	SSW/yr	wave1:wave2
-	ERA5	*0.6~~3~~7 (0.4~~0~~6)*	*77:23*	*0.69~~0~~*	*100:0*	*0.85*	*45:55*
None	CTRL	*1.12*	*62:38*	*1.66*	*70:30*	*0.81*	*70:30*
850-2.6 hPa	FullBC	*0.71*	*76:24*	*1.07*	*71:29*	*0.58*	*62:38*
100-2.6 hPa	StratBC	*1.04*	*67:33*	*1.38*	*74:26*	*0.72*	*51:49*
850-100 hPa	TropBC	*1.10*	*68:32*	*1.23*	*75:25*	*0.86*	*60:40*
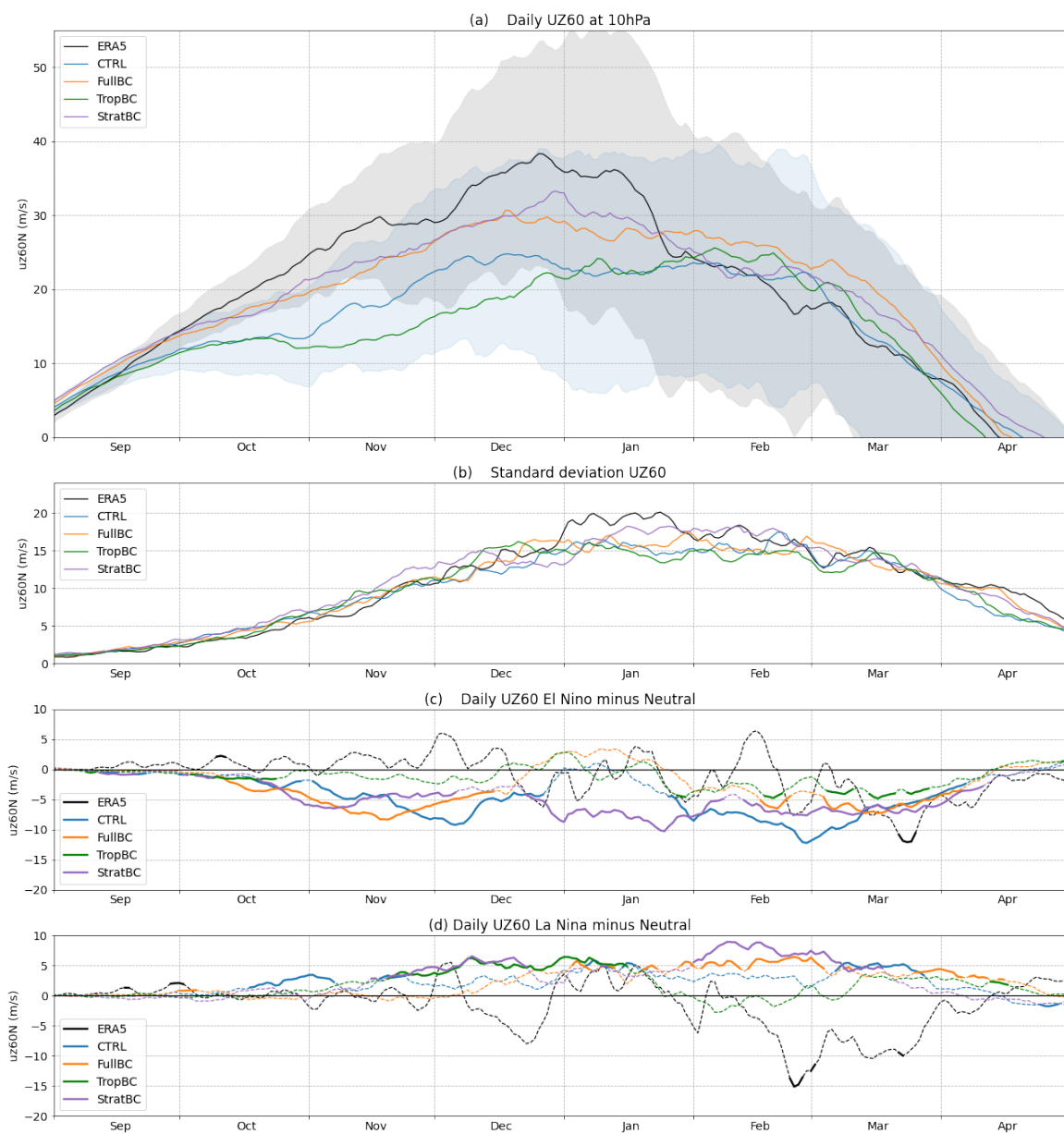
Figure 1: a) Mean daily zonal mean zonal winds at 60°N and 10hPa for ERA5 (1979-2019) and control and bias-correction experiments (100 years each). Shading shows one standard deviation for ERA5 (grey) and CTRL (blue). b) Mean daily standard deviation for ERA5 and the experiments. c) Daily uz_{60} El Niño response calculated as El Niño years minus neutral years, e.g. CTRL_EN – CTRL. d) Daily uz_{60} La Niña response calculated as La Niña years minus neutral years. For c) and d) dashed line shows the mean El Niño/La Niña response, and a solid line indicates responses significant at the 5% level.
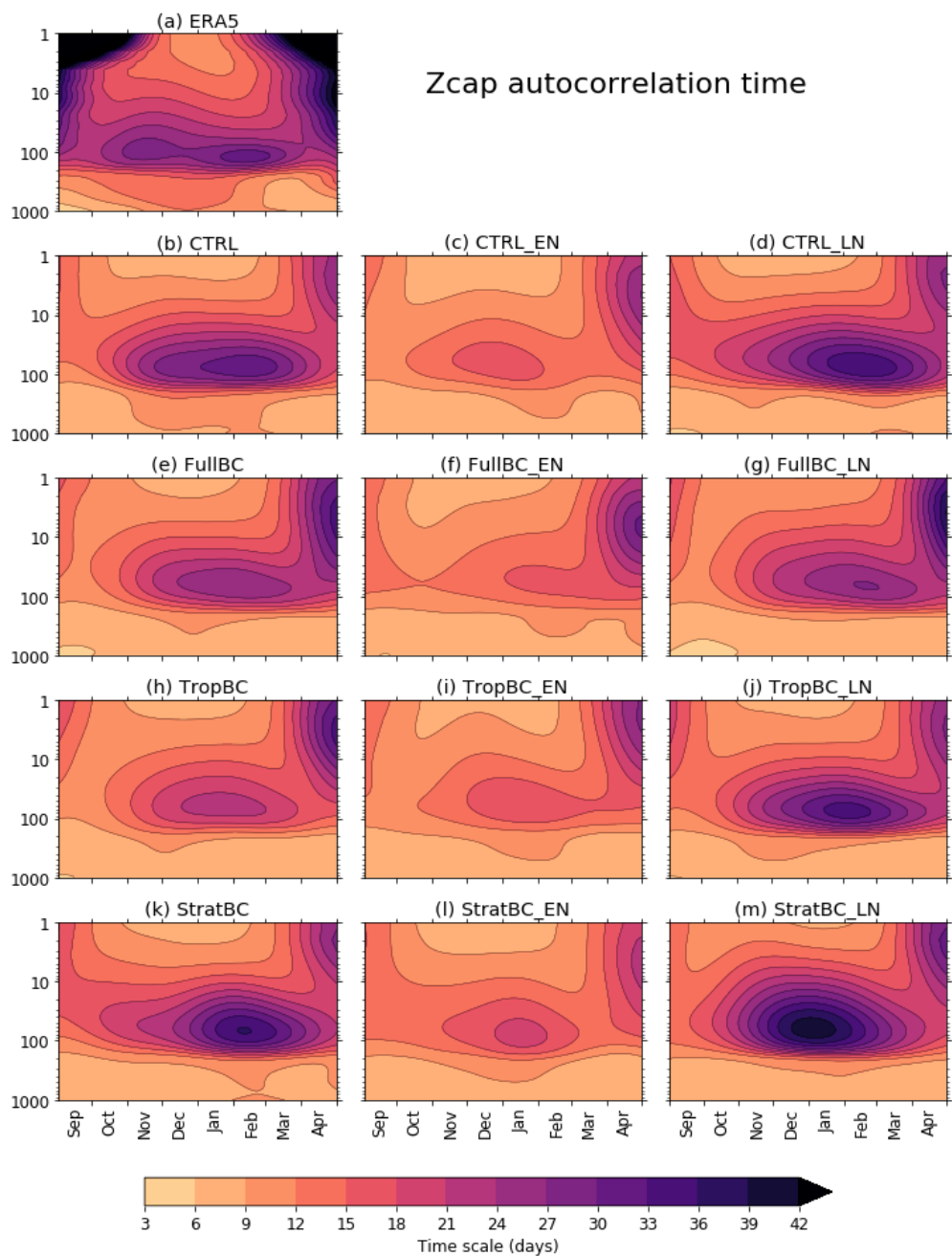
Figure 2: Time-scales of polar cap (60°N-90°N area average with cosine weighting) geopotential height (Zcap) variability. The timescales are defined as days when the autocorrelation function drop to 1/e. The timeseries are smoothed in time with a Gaussian filter (σ = 26 days), following Fig. 1 from Baldwin et al. 2003. ERA5 data from 1979-2019, 100 years for each model run (neutral, El Niňo, and La Niňa).
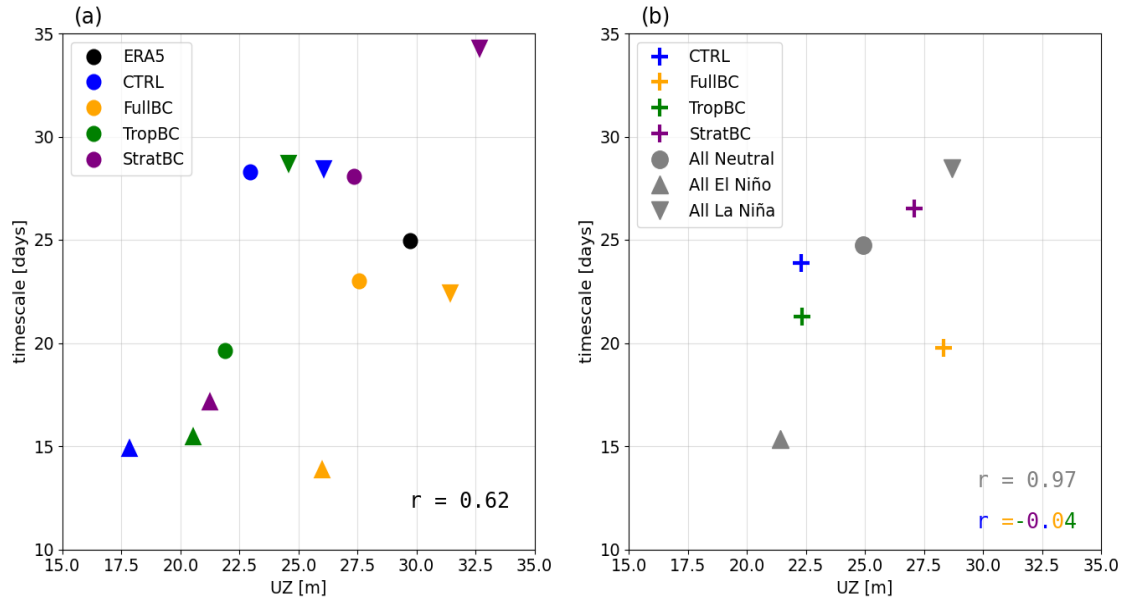
Figure 3: Mean DJF zonal mean zonal wind at 60°N 10hPa plotted against weighted average from 50-150hPa of time-scales of polar cap geopotential height (Zcap) from Figure 2. (a) Circles are neutral ENSO conditions, upward-triangles are El Niño experiments**, downward-triangles are La Niña** experiments**. Correlation coefficient for all models and ENSO phases:** *r = 0.62* **(b) Colored crosses show the mean of El Niño, La Niña and neutral conditions for each experiment (corr. coef.:** *r = -0.04*). **Grey circle and triangles show the multi-model mean for each ENSO phase (corr. coef.:** *r = 0.97*). **ERA5 data from 1979-2019, 100 years for each model run (neutral, El Niño, and La Niña).**

445
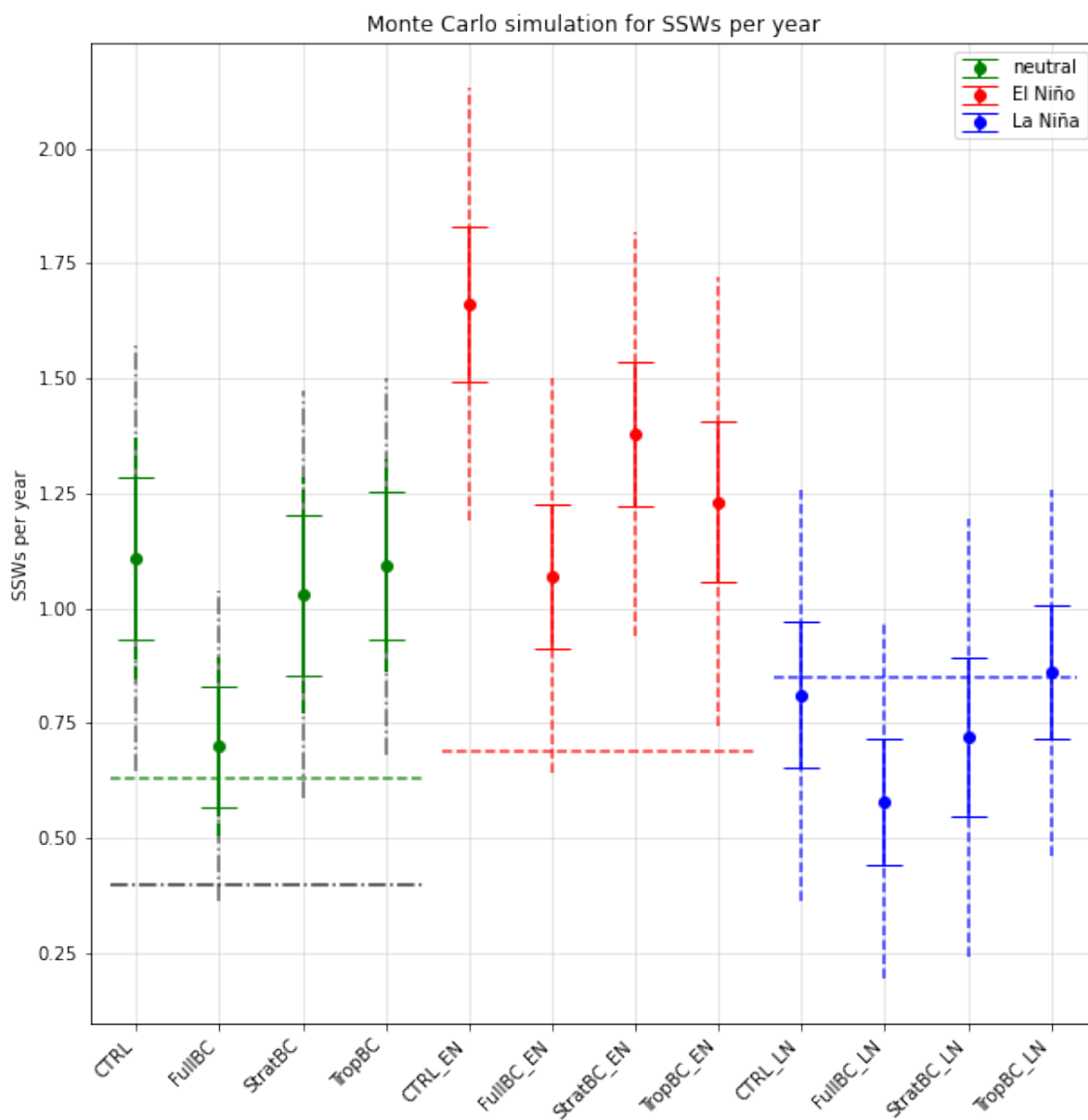
Figure 4: Estimates of variability for number of SSWs per year using Monte Carlo simulations (N=10000). The horizontal dashed lines show the ERA5 SSWs per year (green is all 41 years, black dot-dashed is 15 neutral years, red is 13 El Niño years, blue is 13 La Niña years). Solid error bars show the 5th-95th percentile for Monte Carlo simulation choosing 100 random years from each 100 year simulation. Dashed error bars show the 5th-95th percentile for Monte Carlo simulations where the number of years choosen matchs the number of years in ERA5 (green: all 41 years, black dot-dashed: 15 neutral years, red: 13 El Niño years, blue: 13 La Niña years).
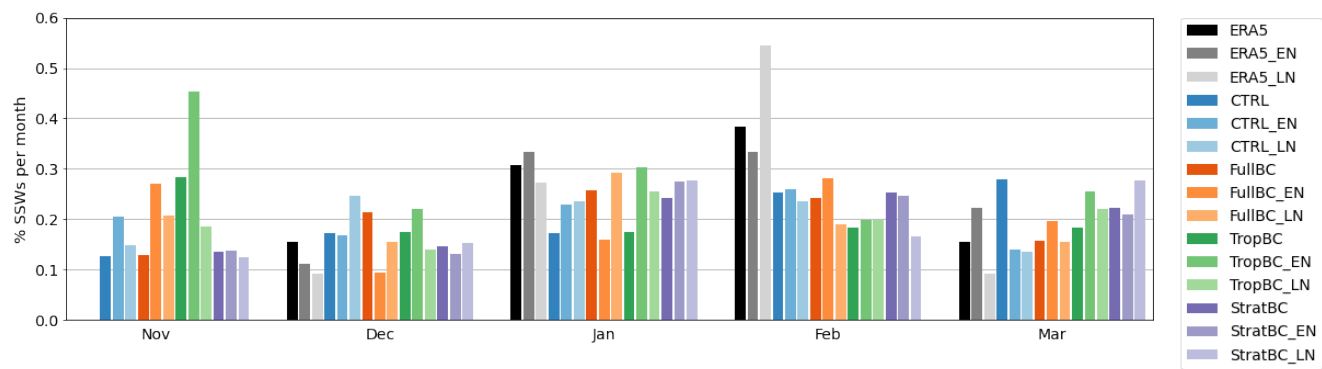
Figure 5: Monthly distribution of SSW frequencies, for neutral, El Niño and La Niña years. The bars are normalized by dividing the number of SSWs in each month by the total number of SSWs for each experiment. ERA5 data from 1979-2019, 100 years for each model run (neutral, El Niño, and La Niña).

460
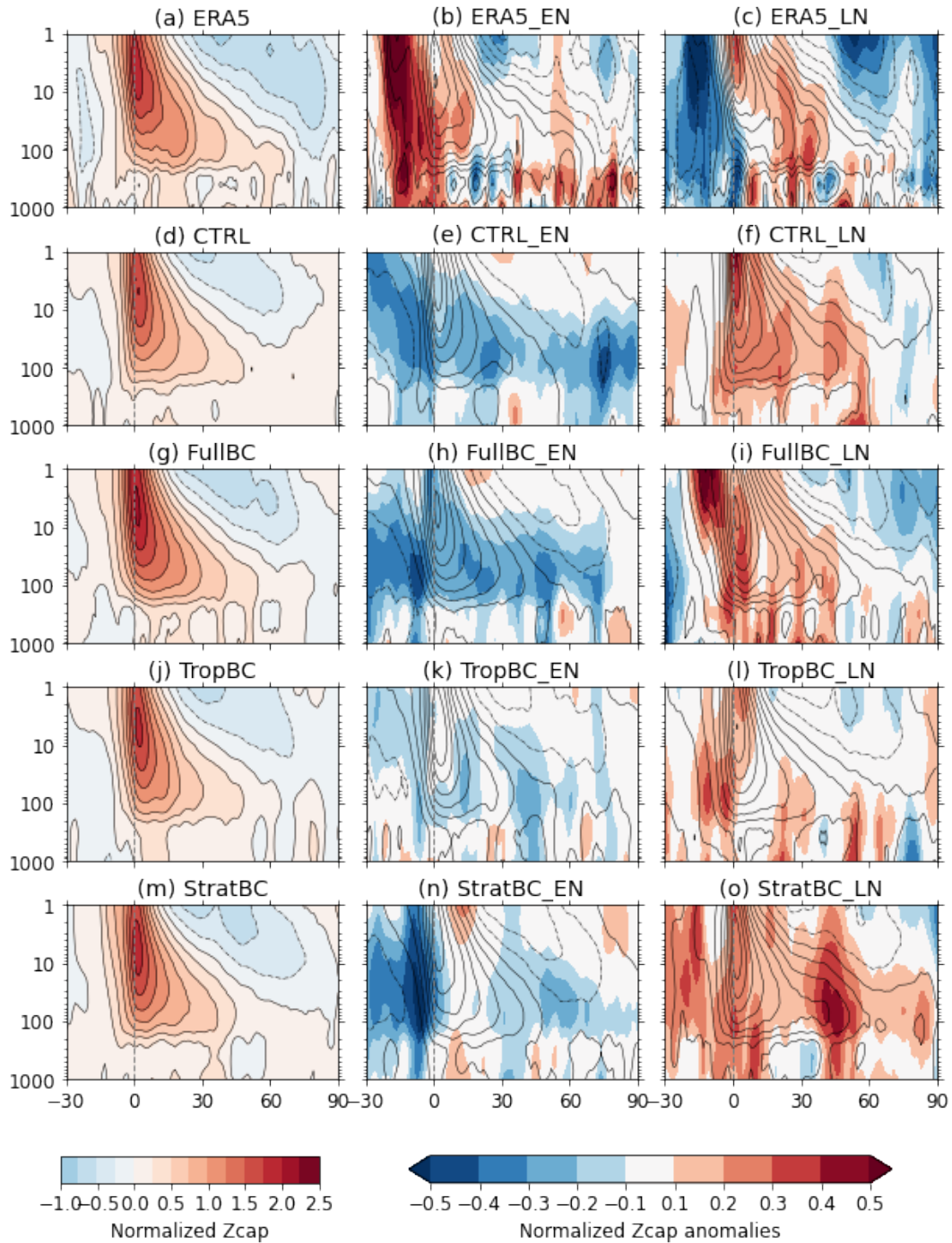
Figure 6: SSW response of normalized polar cap (60°N-90°N) geopotential height, composited around SSW day zero (shown as dashed vertical grey line). Left column shows neutral years for the experiments and all years for ERA5. Middle column shows SSW response in El Niño years with contours (negative values dashed) and the El Niño anomalous SSW response in colours (normalized polar cap geopotential height in El Niño years minus neutral years). Right column is the same for La Niña. ERA5 data from 1979-2019, 100 years for each model run (neutral, El Niño, and La Niña).
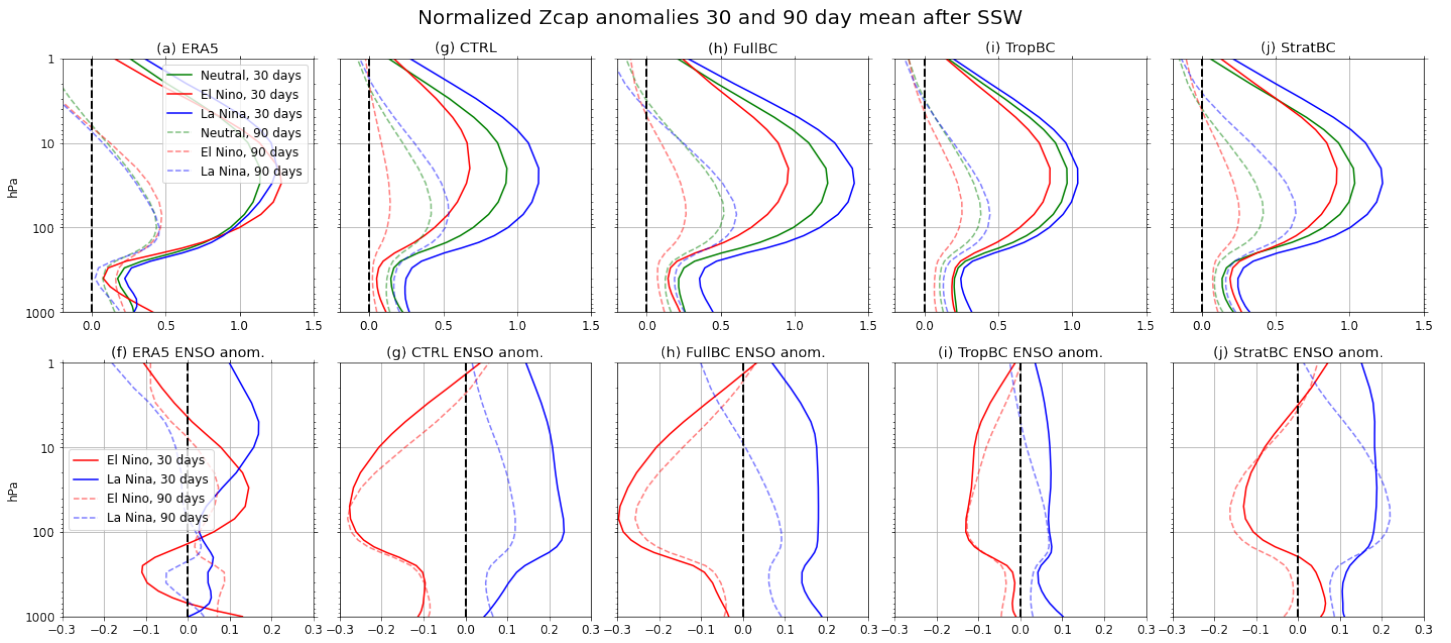
Figure 7: Top row; Normalized polar cap (60°N-90°N) geopotential height, composited around day zero of and time-averaged for 30 and 90 days after SSWs, for neutral years (green), El Niño years (red) and La Niña years (blue). Bottom row shows the change in the mean Zcap SSW difference from neutral years due to El Niño (red) and La Niña (blue)

470
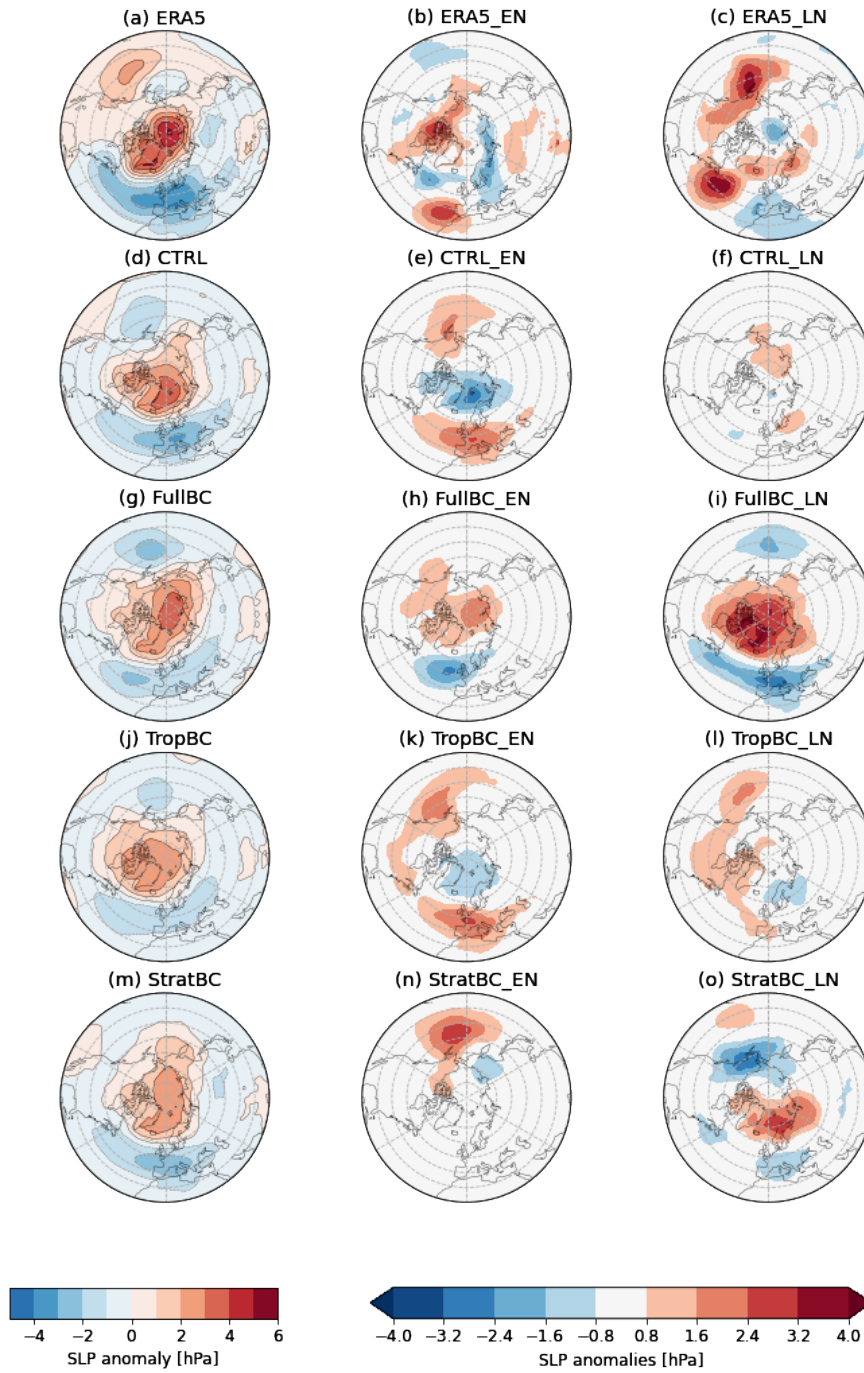
Figure 8: Left column: Composites of absolute SLP anomalies averaged over 30 days after the central dates of SSWs, for neutral ENSO conditions for the models and all years for ERA5. Middle and right column show difference between SLP anomalies during El Niño (middle) and La Niña (right), minus neutral condition anomalies. ERA5 data from 1979-2019, 100 years for each model run (neutral, El Niño, and La Niña).
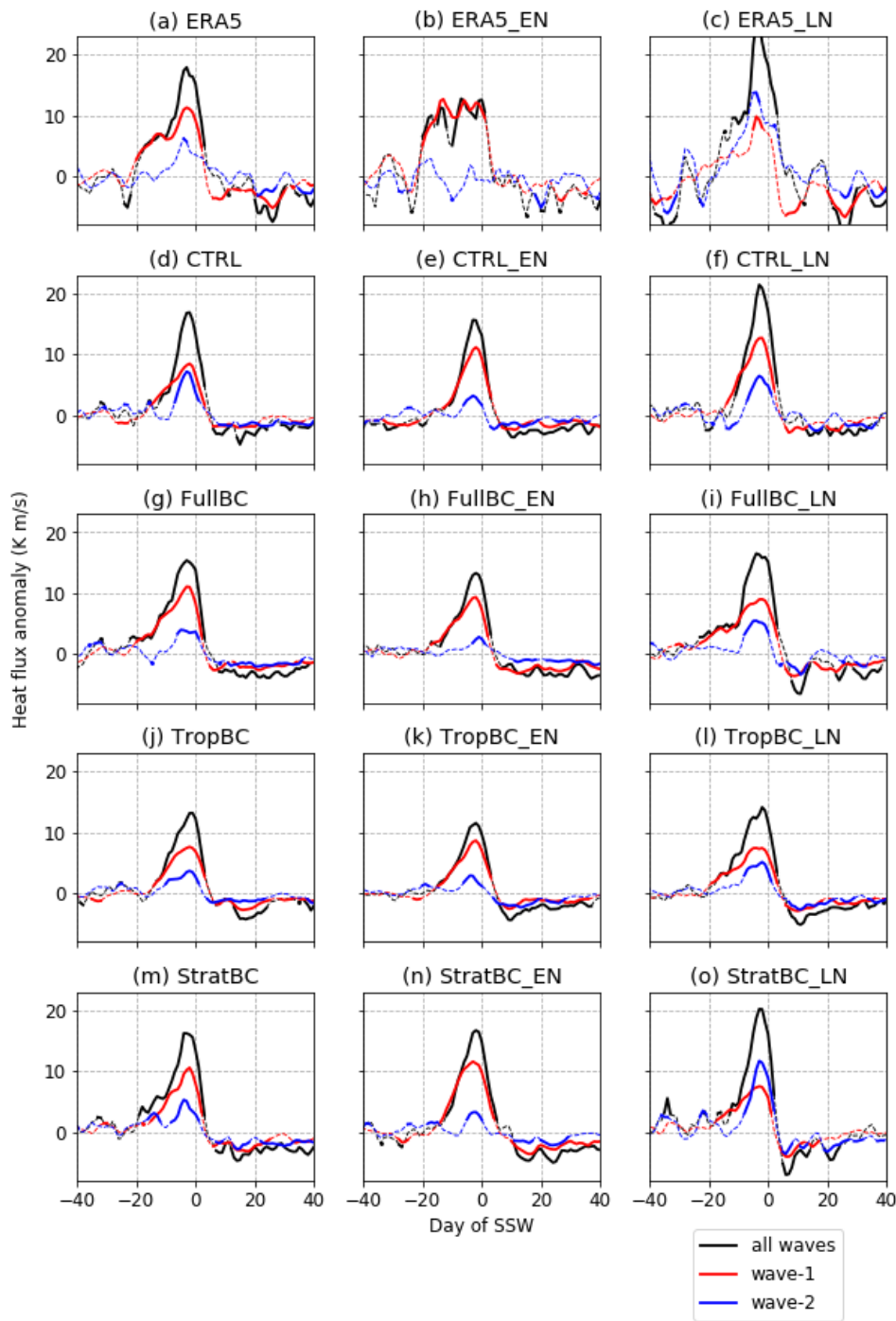
Figure 9: Heat flux anomaly at 100hPa, 45°N-75°N for SSW events composited on day zero of SSWs, defined for wave 1 (red), wave 2 (blue) and all waves (black). In model experiments, the anomalies are calculated with respect to that experiment's climatology. In ~~ERA5~~ERA-5, the anomalies for all years as well as for ENSO years are calculated with respect to the ~~ERA5~~ERA-5 climatology. ERA5 data from 1979-2019, 100 years for each model run (neutral, El Niño, and La Niña).

480