The Structure and Dynamics of the Stratospheric Northern Annular Mode in CMIP5 Simulations

YUN-YOUNG LEE

Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, and Department of Land, Air and Water Resources, University of California, Davis, Davis, California

ROBERT X. BLACK

Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia

(Manuscript received 20 September 2013, in final form 23 September 2014)

ABSTRACT

The structure and dynamics of stratospheric northern annular mode (SNAM) events in CMIP5 simulations are studied, emphasizing (i) stratosphere–troposphere coupling and (ii) disparities between high-top (HT) and low-top (LT) models. Compared to HT models, LT models generally underrepresent SNAM amplitude in stratosphere, consistent with weaker polar vortex variability, as demonstrated by Charlton-Perez et al. Interestingly, however, this difference does not carry over to the associated zonal-mean SNAM signature in troposphere, which closely resembles observations in both HT and LT models. Nonetheless, a regional analysis illustrates that both HT and LT models exhibit anomalously weak and eastward shifted (compared to observations) storm track and sea level pressure anomaly patterns in association with SNAM events.

Dynamical analyses of stratosphere–troposphere coupling are performed to further examine the distinction between HT and LT models. Variability in stratospheric planetary wave activity is reduced in LT models despite robust concomitant tropospheric variability. A meridional heat flux analysis indicates relatively weak vertical Rossby wave coupling in LT models consistent with the excessive damping events discussed by Shaw et al. Eiassen–Palm flux cross sections reveal that Rossby wave propagation is anomalously weak above the tropopause in LT models, suggesting that weak polar vortex variability in LT models is due, at least in part, to the inability of tropospheric planetary wave activity to enter the stratosphere. Although the results are consistent with anomalously weak vertical dynamical coupling in LT models during SNAM events, there is little impact upon attendant tropospheric variability. The physical reason behind this apparent paradox represents an important topic for future study.

1. Introduction

For several decades in the twentieth century, it was widely accepted that the troposphere influences the stratosphere mainly through the upward propagation of Rossby and gravity waves. Sudden stratospheric warmings (SSWs), a primary form of stratospheric variability, are driven by the breaking of planetary-scale Rossby waves emanating from tropospheric altitudes, leading to a westward torque within the stratosphere. The resulting atmospheric consequences of such events include the increase of polar stratospheric temperatures by as much as 50°C and the reversal of stratospheric circumpolar

flow from westerly to easterly (Andrews et al. 1987; Limpasuvan et al. 2004; Matsuno 1971). Similarly, zonal-wind decelerations associated with anomalous upward propagation of Rossby wave activity also generate stratospheric final warming (SFW) events (Black and McDaniel 2007). Analogous studies of polar vortex intensification events have also been performed (Limpasuvan et al. 2005; McDaniel and Black 2005).

The robust downward stratospheric modulation of the tropospheric circulation has become an active research topic since Thompson and Wallace (1998) found that strong polar vortex variability during the winter season is associated with hemispheric-scale tropospheric circulation anomalies. Their discovery led to the adoption of the concept of “annular modes,” which span the troposphere and lower stratosphere. Over the Northern Hemisphere, the northern annular mode (NAM) connects
variations in the stratospheric polar vortex strength to the meridional displacement of the midlatitude tropospheric jet stream, influencing regional climate and weather (Thompson and Wallace 1998; Wittman et al. 2005). Possible mechanisms for downward coupling include so-called downward control (Black 2002; Haynes 2005; Haynes et al. 1991), indirect effects related to the alteration of planetary Rossby wave propagation within the troposphere (Song and Robinson 2004), and vertical wave reflection within the stratosphere, referred to as “downward wave coupling” (Perlwitz and Harnik 2003; Shaw et al. 2010). The latter suggests a close relationship between negative heat flux events in the lower stratosphere and a large-amplitude high-latitude wave 1 pattern in the troposphere that is linked to robust circulation anomalies over the North Atlantic (Shaw and Perlwitz 2013).

A new suite of model simulations available in association with phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) provides an ideal opportunity to assess the capabilities of state-of-the-science atmosphere–ocean general circulation models (AOGCMs) in representing the structure and temporal evolution of the stratospheric northern annular mode (SNAM) and its impact on regional tropospheric weather conditions during boreal winter. One improvement in CMIP5 is that many more models have their highest vertical levels (“top” or “lid”) above 1 hPa than in phase 3 (CMIP3). Previous studies have illustrated that the simulation of stratospheric climate is improved in stratosphere resolving models (Cordero and Forster 2006; Pawson et al. 2000). Therefore, it is expected that some CMIP5 models with high lids and increased vertical resolution within the stratosphere (so-called high-top models) will replicate the observed stratospheric mean state and variability better than the other models (low-top models). The main purpose of the current study is to isolate the differences in the representation and impact of SNAM between high-top and low-top models in CMIP5 simulations.

Several modeling studies have found that enhanced stratospheric representation can reduce the tropospheric biases and variability and additionally provide a more realistic representation of tropospheric climate variability. Boville and Cheng (1988) demonstrate that a wave reflection off of the model lid, caused by a low model top, can alter both the stratospheric and tropospheric circulation. Shaw and Perlwitz (2010) illustrate how wave reflection off the model lid can impact the heat flux distribution. Sassi et al. (2010) suggest that a poorly resolved stratosphere might strengthen the reflection of planetary waves from the lid of low-top models, resulting in less realistic atmospheric variability in the troposphere and stratosphere. Charlton-Perez et al. (2013) found that in low-top models (i) north annular mode events are less persistent, leading to a shorter-lived tropospheric impact than in observations; and (ii) SSW frequency is underestimated. Hardiman et al. (2012) suggest that, compared to low-top models, high-top models better resolve the surface response to the El Niño–Southern Oscillation (ENSO), quasi-biennial oscillation (QBO), and SSWs. Shaw and Perlwitz (2010) highlight that the underestimation of the extreme eddy heat flux in the low-top version of the Canadian Middle Atmosphere Model (CMAM) is due to the excessive Rayleigh damping, rather than only being due to a low-lid height. Similarly, Sigmond et al. (2008) suggest that differences in the doubled CO2 response between high-top and low-top models is mainly due to different parameterizations of orographic gravity wave drag and not the height of the model top. The recent study of Shaw et al. (2014) demonstrates that biases in stratospheric heat flux variability are linked to biases in stratosphere–troposphere coupling concentrated in the North Atlantic region. They further illustrate that the performance of low-top models can vary in relation to the specific nature of the associated stratospheric heat flux bias.

The impact of model-top height upon representing the tropospheric circulation remains incompletely resolved, motivating detailed dynamical analyses of such behavior in CMIP5 models. The current paper provides direct statistical, structural, and dynamical comparisons between high-top and low-top CMIP5 models, focusing on a detailed diagnostic assessment of SNAM and its associated regional tropospheric behavior, including the subtropical jet streams, storm tracks, and lower-tropospheric circulation. In section 2, we outline the dataset and methodology used. In section 3, we present diagnoses of the statistics, spatial structure, time evolution, and tropospheric impact of SNAM in high-top and low-top models. In section 4, we provide a suite of diagnostic analyses focused on deducing the underlying dynamical discrepancies in stratosphere–troposphere coupling between the two model groups. Lastly, a summary and conclusions are given in section 5.

2. Data and methods

The dataset used in this study is the model output from a set of 20 historical AOGCM simulations derived from the CMIP5 archive (complete model expansions are provided in Table 1) (Taylor et al. 2012). Among the 20 simulations considered, 10 have a well-resolved stratosphere (so-called high-top models) while the other 10 have a relatively poorly resolved stratosphere (low-top models). The separation of models into two groups is based mainly upon the lid height, with a threshold of 1 hPa, following the approach of Charlton-Perez et al. 2013. Model
<table>
<thead>
<tr>
<th>Models</th>
<th>Expansion</th>
<th>Lid height</th>
<th>Event No.</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNR</td>
<td>NCEP–NCAR reanalysis dataset</td>
<td>—</td>
<td>24</td>
<td>10.3</td>
</tr>
<tr>
<td>HT: CMCC-CESM</td>
<td>Centro Euro-Mediterraneo per I Cambiamenti Climatici Carbon Cycle Earth System Model</td>
<td>0.01 hPa (80 km)</td>
<td>25</td>
<td>11.5</td>
</tr>
<tr>
<td>HT: GFDL CM3</td>
<td>Geophysical Fluid Dynamics Laboratory (GFDL) Climate Model, version 3</td>
<td>0.01 hPa</td>
<td>27</td>
<td>6.1</td>
</tr>
<tr>
<td>HT: HadGEM2-CC</td>
<td>Hadley Centre Global Environment Model, version 2—Carbon Cycle Earth System Model</td>
<td>85 km</td>
<td>31</td>
<td>8.1</td>
</tr>
<tr>
<td>HT: IPSL-CM5A-LR</td>
<td>L’Institut Pierre-Simon Laplace (IPSL) Coupled Model, version 5A, low resolution</td>
<td>0.04 hPa</td>
<td>29</td>
<td>8.3</td>
</tr>
<tr>
<td>HT: IPSL-CM5A-MR</td>
<td>IPSL Coupled Model, version 5A, mid resolution</td>
<td>0.04 hPa</td>
<td>36</td>
<td>9.5</td>
</tr>
<tr>
<td>HT: IPSL-CM5B-LR</td>
<td>IPSL Coupled Model, version 5B, low resolution</td>
<td>0.04 hPa</td>
<td>33</td>
<td>6.2</td>
</tr>
<tr>
<td>HT: MIROC-ESM</td>
<td>Model for Interdisciplinary Research on Climate, Earth System Model</td>
<td>0.0036 hPa</td>
<td>35</td>
<td>9.0</td>
</tr>
<tr>
<td>HT: MIROC-ESM-CHEM</td>
<td>Model for Interdisciplinary Research on Climate, Earth System Model, Chemistry Coupled</td>
<td>0.0036 hPa</td>
<td>30</td>
<td>10.0</td>
</tr>
<tr>
<td>HT: MPI-ESM-LR</td>
<td>Max Planck Institute Earth System Model, low resolution</td>
<td>0.01 hPa</td>
<td>24</td>
<td>7.1</td>
</tr>
<tr>
<td>HT: MRI-CGCM3</td>
<td>Meteorological Research Institute Coupled Atmosphere–Ocean General Circulation Model, version 3</td>
<td>0.01 hPa</td>
<td>36</td>
<td>7.2</td>
</tr>
<tr>
<td>LT: BCC_CSM1.1</td>
<td>Beijing Climate Center, Climate System Model, version 1.1</td>
<td>2.917 hPa</td>
<td>21</td>
<td>9.4</td>
</tr>
<tr>
<td>LT: BNU-ESM</td>
<td>Beijing Normal University–Earth System Model</td>
<td>2.917 hPa</td>
<td>35</td>
<td>9.1</td>
</tr>
<tr>
<td>LT: CNRM-CM5</td>
<td>Centre National de Recherches Méterologiques Coupled Global Climate Model, version 5</td>
<td>10 hPa</td>
<td>36</td>
<td>8.7</td>
</tr>
<tr>
<td>LT: CSIRO-Mk3.6.0</td>
<td>Commonwealth Scientific and Industrial Research Organisation Mark 3.6.0</td>
<td>4.5 hPa</td>
<td>35</td>
<td>7.3</td>
</tr>
<tr>
<td>LT: GFDL-ESM2G</td>
<td>GFDL Earth System Model with Generalized Ocean Layer Dynamics component</td>
<td>3 hPa</td>
<td>26</td>
<td>15.5</td>
</tr>
<tr>
<td>LT: GFDL-ESM2M</td>
<td>GFDL Earth System Model with Modular Ocean Model, version 4 component</td>
<td>3 hPa</td>
<td>28</td>
<td>11.3</td>
</tr>
<tr>
<td>LT: HadCM3</td>
<td>Hadley Centre Coupled Model, version 3</td>
<td>10 hPa</td>
<td>23</td>
<td>9.8</td>
</tr>
<tr>
<td>LT: INM-CM4.0</td>
<td>Institute of Numerical Mathematics Coupled Model, version 4.0</td>
<td>10 hPa (0.01 sigma)</td>
<td>34</td>
<td>7.9</td>
</tr>
<tr>
<td>LT: MIROC5</td>
<td>Model for Interdisciplinary Research on Climate, version 5</td>
<td>3 hPa</td>
<td>40</td>
<td>7.0</td>
</tr>
<tr>
<td>LT: NorESM1-M</td>
<td>Norwegian Earth System Model, version 1 (intermediate resolution)</td>
<td>3.54 hPa</td>
<td>31</td>
<td>10.4</td>
</tr>
</tbody>
</table>

HT average: 30.6 36.4 8.3 12.8
LT average: 30.9 44.5 9.6 9.4
HT – LT: (11.1%) (0.3%) (8.2%) (99.8%)
ensembles based on the dynamical performance metrics, such as heat flux, rather than lid heights can be more appropriate for the SNAM analysis because dynamical metrics naturally exclude the models with biased dynamics (over Rayleigh damping and excessive reflection) (Shaw et al. 2014). Although the more common lid-height formalism may inadvertently include some models with such biased behavior, here we choose to retain this framework (i) because of the limited sample size and (ii) for consistency with existing research. The results of the current study are based on analyzing a single ensemble member (r1i1p1) for each AOGCM considered since (i) not all required variables are available for each ensemble member, and (ii) the statistical behavior of SNAM obtained for the 56-yr period considered is insensitive to the ensemble member analyzed. Given the different horizontal model grid configurations, for consistency, the output from each simulation is first interpolated to a common 2.5° latitude by 2.5° longitude grid. The common eight vertical levels from 1000 to 10 hPa are selected prior to analysis. To validate the CMIP5 simulations, we analyze parallel observational results derived from the National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) reanalysis dataset (NNR; Kalnay et al. 1996). The time period considered in this study is from 1950 to 2005 (56 years), encompassing the latter portion of the twentieth century, and the boreal winter season is defined as the 120-day period extending from January through April (JFMA), when the polar vortex exhibits the greatest variability.

To identify the spatial structure of SNAM, we applied principal component (PC) analysis to the daily zonal-mean zonal wind anomalies (UA) for NNR and each CMIP5 simulation. The SNAM loading pattern is taken as the first eigenvector of zonal wind for the spatial domain of 45°–90°N, 100–100 hPa, encompassing the lower stratospheric polar vortex during the boreal winter season (JFMA) [following Black and McDaniel (2009)]. The daily SNAM index is defined as the normalized PC time series derived from projecting the daily zonal-mean zonal wind anomaly field upon the loading pattern. One inherent limitation of the current study is that only three vertical levels (10, 50, and 100 hPa) are available for the analysis, given the relatively coarse vertical resolution of the CMIP5 output. Nonetheless, we believe that these levels are sufficient for the SNAM characterization, since they span the primary vertical levels normally sampled in the mid and lower stratosphere.

The typical structural and dynamical behavior associated with SNAM is diagnosed using both linear regression and composite analyses. We isolate characteristic regional tropospheric behavior by regressing anomalous sea level pressure (SLP) and storm track fields against the SNAM index. A local storm track measure, the envelope function of the meridional wind at the 250 hPa (ENV250) (Nakamura 1992; Nakamura and Wallace 1990), is calculated from sequential high-pass and low-pass filtering of the meridional wind field using a 15-point Lanczos filter with a cutoff period of 8 days (Duchon 1979). We also construct composite analyses of large-amplitude positive (POS) and negative (NEG) SNAM events in order to directly compare the characteristic behavior of robust SNAM events among NNR, HT, and LT models. To select POS (NEG) SNAM events, we seek periods when the SNAM index exceeds +2 (falls below −2) standard deviations. Each event is required to persist for at least 3 days, noting that the typical duration time scale for observed large-amplitude SNAM events is ~12 days (McDaniel and Black 2005). To ensure event independence, the interval between consecutive events is required to be at least 15 days. Finally, to isolate boreal winter episodes, we require that the central date of each event lie within the 120-day period encompassing JFMA.

After applying the case identification criteria discussed above, we then perform lead–lag composite analyses to isolate the typical structural and dynamical time evolution of SNAM events. These include pressure–time cross sections of zonal-mean zonal wind anomalies latitudinally averaged over 52.5°–87.5°N (encompassing the polar vortex core). We also contrast separate phases of the short-term SNAM life cycle, namely the onset (lag −5 to −1), mature (lag 0 to +4), and decay (lag +5 to +14) phases, which encompass the period of strongest stratosphere–troposphere coupling and largest amplitude tropospheric response (e.g., McDaniel and Black 2005). The life cycle phases are selected based upon the characteristics of the evolution of zonal-mean zonal wind anomaly identified in section 3b. The period average structure of zonal-mean zonal winds and wave activity is assessed for each of the three life cycle phases defined above. The period-averaged results are only presented for the onset phase since the latter two phases are readily described in terms of the onset phase structure.

Besides zonal-mean zonal wind, composite analyses of dynamically oriented flow measures are performed to isolate possible dynamical distinctions among SNAM events in NNR, HT, and LT models. These include Eliassen–Palm cross sections to diagnose the propagation of zonal-mean Rossby wave activity in the latitude–pressure plane along with its interaction with the zonal-mean flow (Black and McDaniel 2007). We also compare the behavior of planetary-scale wave activity in the upper troposphere and lower stratosphere by performing a
power spectrum analysis of the meridional wind field at 250 and 50 hPa (noting that upper-tropospheric planetary wave activity is a fundamental source of variability for the stratospheric polar vortex).

The general method used to assess statistical significance is a statistical hypothesis test based upon random sampling (Efron and Tibshirani 1993), as the limited number of model ensembles are subject to sampling uncertainties. In contrasting the HT and LT behavior, we assess (i) differences between each group (HT or LT) and NNR, as well as (ii) differences between the two groups. In the former case, the model group biases are normalized at each grid point by the standard deviation of 10,000 random group samples (normalized values are displayed with green and brown shading in Figs. 1–5).

In the latter case, we assess the significance of the difference between the HT and LT composites for the linear regression and certain lag composite analyses of SNAM events. In this procedure, we 1) conduct the 10-model random sampling twice among a total of 20 models (with repetition), 2) calculate the composite difference between the two random groups, 3) repeat this procedure 10,000 times to create the null distribution for each grid point, and 4) assess the relative location of the difference between the HT and LT original composite to the null distribution. If the difference between the HT and LT model groups falls among the lowest 2.5% or the highest 2.5% of the null distribution, then the composite difference between the two model groups is considered statistically significant at the 5% confidence level.

3. Statistics and structure of SNAM in CMIP5 simulations

a. Spatial structure and tropospheric extent of SNAM

During boreal winter, the atmospheric general circulation exhibits two prominent hemispheric-scale westerly features: the tropospheric subtropical jet and the polar night jet (Wallace and Hobbs 2006). In the current study, the analysis of the polar night jet is limited to only its lower part—namely, the lower stratospheric polar vortex—since CMIP5 model output is generally limited to 10 hPa or below. In CMIP5 model simulations, the general structure of the climatological-mean zonal-mean zonal winds generally resembles observations (Fig. 1). However, the overall intensity of both jet features is anomalously strong (green shading) in both HT (40.1 m s$^{-1}$, 24.7 m s$^{-1}$) and LT (37.8 m s$^{-1}$, 25.5 m s$^{-1}$) CMIP5 models compared to NNR (36.3 m s$^{-1}$, 18.9 m s$^{-1}$). Conversely, the high latitudinal tropospheric zonal winds are suppressed (brown shading) in both CMIP5 model groups. Although the two model groups exhibit qualitatively similar deficiencies, there are also some more subtle differences in the pattern and intensity of the polar vortex between HT and LT models. In the LT models, the latitudinal range (vertical depth) of the polar vortex is greater (shallower) than either NNR or HT with an anomalous southward expansion (vertical retraction) of the polar vortex. The peak intensity of polar vortex is slightly stronger in LT (25.5 m s$^{-1}$) than in HT (24.7 m s$^{-1}$). Conversely, the peak intensity of the tropospheric subtropical jet is
weaker in LT (37.8 m s$^{-1}$) than in HT (40.1 m s$^{-1}$). The weaker lower-stratosphere winds in LT models alter the local refractive index for planetary wave propagation, providing a potential partial barrier to vertical and latitudinal wave propagation (Sigmond and Scinocca 2010). This idea is studied further in the diagnostic analyses of section 4.

The intraensemble spread in representing the tropospheric subtropical jet and the stratospheric polar vortex within the HT and LT model ensembles is also studied in terms of the standard deviation within each ensemble (not shown). For both groups (HT and LT), the intermodel variability in representing the stratospheric polar vortex is greater than that for the tropospheric subtropical jet. For the stratospheric polar vortex, the standard deviation within the LT ensemble is larger than that for the HT ensemble. By contrast, the midlatitude tropospheric zonal winds show slightly greater variability in the HT ensemble than in the LT ensemble. Interestingly, the two model groups exhibit a marked difference in the peak latitudinal position of the standard deviation of zonal wind in the vicinity of the tropopause. At 100 hPa, the HT ensemble has maximum spread near 37.5°N, while the LT ensemble has a local maximum near 30°N.

In addition to the climatological-mean pattern discussed above, the CMIP5 models also exhibit deficiencies in representing the temporal variability in zonal-mean zonal wind, particularly at stratospheric altitudes. Linear regressions of zonal-mean zonal wind anomalies with respect to SNAM indices are displayed in Fig. 2. For HT models, although the regressed pattern strongly resembles NNR in terms of the dipole structure, the westerly anomalies within the 10-hPa high-latitude pole are anomalously strong (∼16.7 m s$^{-1}$) compared to NNR (14.3 m s$^{-1}$). Conversely, for LT models, the northern pole is anomalously weak (∼10.0 m s$^{-1}$) at stratospheric altitudes and the peak latitude at 10 hPa is displaced slightly poleward compared to NNR. Furthermore, the southern pole (easterly anomalies) in the LT models is anomalously weak and displaced downward (with peak amplitudes ∼100 hPa) compared to HT and NNR. Our random sample analyses indicate that the differences between HT and LT SNAM structures are statistically significant at and above 100 hPa (Fig. 2). Interestingly, the tropospheric zonal wind anomaly patterns associated with SNAM are considerably more uniform among NNR, HT, and LT (versus the stratospheric signature), with little difference in either magnitude or structure. The primary difference of note is that, in both HT and LT model groups, the westerly anomaly signature near 65°N is a bit too strong (green shading in Fig. 2). Furthermore, there are no statistically significant differences in the tropospheric SNAM signature between the HT and LT models. This suggests that the tropospheric impact of SNAM events does not differ appreciably between the two model groups, even though the stratospheric SNAM signature (and associated polar vortex variability) is quite different.
Given these distinctions in CMIP5 HT and LT models between the stratospheric and tropospheric zonal-wind anomaly signatures, it is useful to quantify the vertical change in the anomaly structure. The anomaly magnitude decrease between the midstratosphere and troposphere can be quantified by assessing the pressure level at which the regressed zonal-wind anomaly magnitude (averaged from 52.5°–87.5°N) decreases to \(1/e\) of the value observed at 10 hPa. Such an analysis reveals that the corresponding downward influence of SNAM events is more efficient (or effective) in LT models than in NNR or HT models.

In addition to the loading pattern analysis, we analyze the tropospheric response to SNAM in terms of regional storm track and near-surface circulation anomalies. Regional regression analyses of the observed envelope function (Fig. 3) reveal a meridional dipole anomaly structure concentrated over the North Atlantic regions. This dipole pattern results in northward (southward) storm track shifts during POS (NEG) SNAM events, paralleling the north–south jet shifts found by Shaw et al. (2014) over the North Atlantic in association with extreme stratospheric heat flux episodes. For both HT and LT models, the SNAM-related storm track anomaly dipole pattern is anomalously weak, with the northern pole over the eastern North Atlantic displaced slightly eastward versus NNR. Although the biases appear accentuated in the LT models, the differences between HT and LT model groups are significant only within limited areas lying outside of the primary anomaly signatures. The most notable difference between the two groups is the relatively weak storm track signature over the Arctic region in the LT models.

The impact of SNAM events upon the near-surface general circulations is assessed via linear regressions of sea level pressure with respect to the SNAM indices (Fig. 4). In NNR, the regressed anomaly pattern closely resembles the North Atlantic Oscillation (NAO) or Arctic Oscillation (AO), both reflecting the surface signature of the NAM. For the CMIP5 simulations (particularly the LT models), the overall intensity of the midlatitude circulation cell over the North Atlantic is weaker than in NNR. We additionally note an appreciable eastward shift of the positive anomaly center in the LT models. As found for the storm track analysis, regions of significant differences between HT and LT models are mainly confined to small localized areas away from the primary circulation anomaly signatures. The one exception is associated with anomalously weak SLP anomalies over the midlatitude central North Atlantic (brown shading), linked to the eastward shift discussed above.

We have noted that the representation of the zonal-mean SNAM structure in the stratosphere is significantly weaker in the LT models compared to HT models (Fig. 2). Even though the storm track and SLP patterns associated with SNAM exhibit some differences in both magnitude and spatial distribution between the HT and LT models, these differences are small and statistically...
insignificant in contrast to the stratospheric signatures. This is quantitatively supported by the spatial regression coefficient of model variables with respect to observed ones. The regression coefficient difference between HT and LT for the UA (0.40) is much larger than those for the horizontal SLP (0.11) and ENV250 (0.25) patterns, respectively. This is an unexpected result, as one might intuitively expect the tropospheric SNAM signature to vary in direct relation to differences observed at stratospheric altitudes.

b. Polar vortex evolution during SNAM events

Before studying the temporal structural evolution of SNAM in terms of composite analyses, we investigate the frequency and duration statistics of extreme events selected in terms of SNAM indices for individual model (see section 2). Over the 55 winter seasons considered, the number of observed NEG SNAM events (39) is a little more than 1.5 times that of POS SNAM events (24), as shown in Table 1. CMIP5 SNAM events are significantly more frequent than NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three categories (POS-HT, POS-LT, and NEG-LT), all characterized by average-frequency enhancements over NNR in three
differences found between HT and LT model groups (final row of Table 1). This is pursued by first producing the probability density function (PDF) for differences in frequency and duration between the two model groups using a random sampling analysis. The actual differences (next-to-last row in Table 1) are then ranked in relation to the PDF distributions. Our results indicate that the differences in frequency (duration) of NEG events rank in the lowest 0.3% (highest 0.2%) of the distribution, suggesting statistical significance. In contrast, differences in frequency (duration) of POS events rank in the lowest 11.1% (lowest 8.2%) of PDF. Thus, we conclude that only the frequency and duration of NEG SNAM events are significantly different between HT and LT models.

Using the identified events, the structural evolution of POS and NEG SNAM events is first studied in terms of
lead–lag composites of zonal-mean zonal wind anomalies averaged over a latitudinal band (52.5°–87.5°N) encompassing the extent of the polar vortex (Fig. 5). For observed SNAM events (NNR; Figs 5a,d), the primary zonal wind strengthening (or weakening) in the stratosphere occurs over a 10-day period beginning around 8 days prior to onset (lag 0), with maximum amplitudes achieved 2 days after onset. This acceleration (deceleration) signature is vertically coherent and extends downward through the lower stratosphere and troposphere all the way to the surface. The stratospheric anomaly decay is considerably more gradual, as it takes over a month for the stratospheric circulation to return to its normal state (noting that the tropospheric zonal wind anomalies weaken much more quickly). Although this asymmetry in SNAM evolution (rapid development and slow decay in the stratosphere) may partly reflect our focus on case onset, it is also likely due to the nature of the underlying dynamical mechanisms (as discussed below).

As for the regressed SNAM structure (Fig. 2), we again find that the stratospheric zonal wind anomaly signature is anomalously weak in the LT models with large-amplitude (>three standard deviations) normalized anomalies for both POS and NEG events. This leads to a significantly different polar vortex evolution between the HT and LT model groups during SNAM events (region of significant differences enclosed by red
contrasts the composite SNAM behavior during the period of most rapid and vertically coherent zonal wind changes (−5 to +14 days) characterized by strong stratosphere–troposphere coupling. Informed by the time–pressure evolution in Fig. 5, we divide the short-term SNAM evolution into three parts: the onset (day −5 to −1), mature (day 0 to +4), and decay (day +5 to +14) phases.

Contrasting POS and NEG events in the three composite pairs (Fig. 5), we find that the magnitude of the midstratospheric zonal wind anomalies is notably stronger in NEG events. In the lower stratosphere and troposphere, however, the magnitudes of POS and NEG events are quite similar. There are also qualitative differences in the time evolution of POS and NEG SNAM events. For POS events, modest stratospheric westerly anomalies are present several weeks prior to event onset, and the onset itself is a bit more gradual. For NEG events, however, there is a persistent period of weak westerly anomalies until about three weeks prior to onset. This is followed by a brief period (7–10 days) of weak easterly anomalies before an abrupt deceleration in the week prior to onset. Such asymmetries are evident in the composite analyses of SSW and vortex intensification events by Limpasuvan et al. (2005). The preexisting westerly wind anomalies and relatively abrupt formation of easterly anomalies is likely due to the dynamically-driven nature of NEG SNAM events similar to SSWs and SFWs. It is well known that, during SSWs and SFWs, the polar vortex decelerates abruptly over a few days in response to strongly enhanced vertical propagation of planetary wave activity from the troposphere (Andrews et al. 1987; Black and McDaniel 2007; Matsuno 1971). Associated dynamical analyses of this behavior will be presented in the next section. In contrast, the more gradual development of POS SNAMs is likely related to a relative lack of dynamical forcing from the troposphere leading to a response on a longer radiative time scale. More specifically, for POS SNAMs, upward Rossby wave propagation from the troposphere is relatively weak (negative anomalies), allowing the stratospheric temperature structure to radiatively relax over a period of several weeks prior to event onset, leading to a more gradual strengthening of the polar vortex.

In addition to zonal wind anomaly fields, we also diagnosed the parallel composite evolution of the total zonal-mean zonal wind field associated with POS and NEG SNAM (not shown for brevity). Both NNR and HT model composites exhibit notable easterly zonal wind signatures just after day 0 during NEG events. Of course, this is a characteristic of SSW events occurring in the Northern Hemisphere during the winter season (Andrews et al. 1987). During a SSW, the stratospheric polar vortex weakens abruptly, leading, in some cases, to an easterly reversal in the zonal flow (Andrews et al. 1987; Limpasuvan et al. 2004). In comparing the central dates of our observed (NNR) NEG SNAM events to SSWs identified in Charlton and Polvani (2007), we find that 16 (of 39) of our NEG SNAM events roughly coincide with SSW events.

Stratospheric final warming events can be another factor. SFW onset usually occurs between mid-March and late May, with a mean value near 14 April (Black and McDaniel 2007). Since the analysis period of the current study spans March and April, NEG SNAM events may encompass certain SFW events. Considering that the life cycles of both SSWs and SFWs are driven abruptly (Black and McDaniel 2007; Limpasuvan et al. 2004), the relative rapidity of NEG SNAM onset is consistent with this behavior. On the other hand, we find that the LT models are unable to represent the transition to easterly flow. For LT models, we suspect that this is due to the relative weakness in the zonal flow variability associated with SNAM events. This is also reflected by a general lack of SSW events occurring in LT models (Charlton-Perez et al. 2013).

Returning to our consideration of the zonal wind anomaly evolution, we quantified the anomaly decay time scales by calculating e-folding times at four levels near and above the tropopause (250, 100, 50, and 10 hPa)
based on the lead–lag composite fields (Fig. 5). The $e$-folding time is defined as the mean time taken for the anomaly signal to decrease to $1/e$ of original (peak) magnitude and is obtained by statistically determining $A$ and $B$ in the equation $y = Ae^{-Bx}$ via least squares fitting and averaging results obtained separately for POS and NEG events. The $e$-folding time determined for NNR SNAM events is about 15.2 (30.4) days in the middle (lower) stratosphere defined at 10 (100) hPa. For HT models, the respective $e$-folding times (18.1 and 38.4 days) are found to be longer than NNR. In contrast, the LT models exhibit the shortest persistence (12.4 and 22.0 days, respectively). The maximum zonal wind anomaly persistence is observed to occur at 100 hPa in all three composite pairs. This vertical structure of anomaly persistence is consistent with the results of the observational study by Gerber et al. (2010). They find that NAM persistence reaches a peak near 100 hPa during JFMA (noting they identify NAM structures at individual pressure levels and apply the $e$-folding time-scale analysis to NAM indices). The average $e$-folding time determined here is shorter than the $e$-folding times found by Gerber et al. (2010). This is likely because the composite wind fields analyzed here are based on very strong SNAM cases, while Gerber’s calculation encompasses all NAM events.

The SNAM persistence behavior shows distinct differences between POS and NEG events. The average $e$-folding times for the stratospheric levels above 250 hPa for POS (NEG) events are 20.8, 24.0, and 16.6 days (25.0, 29.6, and 15.4 days) for NNR, HT, and LT, respectively. This reveals that, for both NNR and HT, the easterly anomaly signatures associated with NEG SNAM events persist longer than the corresponding westerly anomaly signatures of POS SNAM. As mentioned above, the total zonal wind field reverses to an easterly wind regime for NEG SNAM events in both NNR and HT. In this case, Rossby wave propagation is temporarily precluded, leading to a different dynamical regime than found under westerly background winds, requiring a longer time to restore the flow to its normal (westerly) state. As mentioned earlier, for the LT models the zonal winds remain westerly even during NEG SNAM events. This likely provides a dynamical explanation for why there is little difference in POS and NEG SNAM persistence in LT models.

The temporal change in the zonal wind field as a function of latitude and height is analyzed by applying a centered difference operator to the zonal-mean zonal winds during the three evolution stages. During the onset phase (Fig. 6), the NNR composite exhibits a vertically coherent pattern of zonal wind increases (decreases) over high latitudes during POS (NEG) events. At 10 hPa, the zonal wind change pattern exhibits peak magnitudes near 65°N (75°N) for POS (NEG) events. For both POS and NEG NNR events, the latitude of the largest-amplitude zonal wind change slopes northward with increasing altitude (decreasing pressure). Near the tropopause (250 hPa) the latitudinal location is around 60°N (70°N) for POS (NEG) events. For POS events, both HT and LT CMIP5 simulations well represent the vertical structure described above for NNR, except that the change pattern is displaced slightly northward in the models. For NEG events, the CMIP5 models exhibit qualitatively different behavior with an equivalent barotropic vertical structure in HT models and a slight southward vertical tilt in LT models. This suggests that the tropospheric signature of NEG SNAM onset is concentrated at higher latitudes in the CMIP5 models compared to NNR. Considering the magnitude of changes in zonal-mean zonal winds, we find that the stratospheric zonal wind changes in HT models are comparable to those of the NNR cases, while zonal wind changes in the LT models are notably weaker than in NNR. Nonetheless, we find that the magnitudes of concomitant tropospheric zonal wind changes observed in the various model composite categories are comparable to NNR, but with small northward shifts in phase. Parallel analyses of the mature and decay phases (for brevity not shown) reveal much weaker zonal wind change patterns. The mature phase exhibits continued (albeit weak) anomaly intensification concentrated at stratospheric altitudes while the decay phase is characterized by a partial return toward climatological conditions (i.e., the change patterns are roughly opposite in sign to those displayed in Fig. 6).

4. Dynamical analyses of stratosphere–troposphere coupling in CMIP5 simulations

Rossby waves with zonal wavenumbers 1, 2, and 3 are the main source of intraseasonal variability in the stratospheric polar vortex. Such waves are generated in the troposphere via orography, land–ocean thermal contrasts and wave–wave interaction and can propagate upward into the stratosphere where, upon breaking, they act to decelerate the polar vortex. Anomalously strong episodes of upward propagation of planetary wave activity can dramatically reduce the stratospheric westerlies and even reverse the wind direction to become easterly. Rossby wave activity is found to be generally enhanced over the North Atlantic, Eurasia, and North America during the incipient stage of SSW events (as early as 10 days prior to maximum warming) and consists primarily of zonal wavenumbers 1 and 2 (e.g., Nakagawa and Yamazaki 2006). Therefore, the
relative strength of tropospheric planetary wave activity may be a contributing factor to misrepresenting SNAM structure in CMIP5 simulations, particularly in LT models. If anomalous wave activity is insufficiently vigorous in the troposphere, its likely impact on the stratosphere will be correspondingly weak, leading to weaker polar vortex variability. This hypothesis is consistent with previous studies analyzing the general heat flux characteristic of models and discovering that extreme eddy heat flux events are underrepresented in the LT version of the CMAM because of excessive Rayleigh damping (Shaw and Perlwitz 2010; Shaw et al. 2014). Our study focuses directly on SNAM events and considers the tropospheric contribution to SNAM onset.

To test our hypothesis, we next evaluate the strength of planetary wave activity by applying a spatial power spectrum analysis to the 250-hPa meridional wind field (Fig. 7). The power spectrum fields are averaged over the 50°–70°N latitudinal band, which is chosen based on the observed structure of anomalous EP fluxes and local wave driving for NNR events (Fig. 11; the selected band encompasses the region of strongest vertical EP fluxes in the vicinity of the tropopause). The results of NNR show, as expected, anomalously weakened (enhanced) planetary wave activity during the onset of POS (NEG) SNAM events. These signatures are consistent with the possibility of anomalously downward (upward) wave propagation during POS (NEG) events. We note that these anomaly signatures tend to be concentrated within zonal wavenumber 2. The power spectrum structure obtained from the CMIP5 models is not too different from NNR around the time of onset (lag +0). However, they are some notable discrepancies in intensity, time

![Figure 6](http://journals.ametsoc.org/doi/pdf/10.1175/JCLI-D-13-00570.1)
evolution and apparent wave–wave interaction. In NNR, the negative (positive) anomalies in wavenumber 2 appear to originate from preexisting anomaly signatures in wavenumber 4 (3) during POS (NEG) events. Interestingly, the LT models qualitatively mimic this apparent upscale energy transfer for NEG events. HT models, on the other hand, do not show any evidence of such wave–wave interaction. In terms of anomaly magnitudes, both HT and LT models appear to similarly underestimate the planetary wave power anomalies observed during the period of SNAM onset.

Figure 8 shows a parallel power spectrum analysis for 50 hPa (in the stratosphere). The temporal variability of wave activity is consistent with the tropospheric signatures in Fig. 7. During onset, there is a coherent simultaneous suppression (enhancement) of wave activity for POS (NEG) events throughout the wave spectrum analyzed. Interestingly, the weakening of wavenumber 1 is dominant for POS events, while for NEG events wavenumber 2 exhibits the strongest variation. The temporal variation of smaller scale wavenumbers (4–5) almost disappears, which is expected...
because of the effective vertical filtering of smaller-scale Rossby waves at the tropopause. Unlike near the tropopause (250 hPa), in the lower stratosphere (50 hPa) the power spectrum anomaly magnitudes resolved in HT models are much closer to observations than found in LT models, especially for wavenumber 2. The difference in wavenumber 2 power between two model groups is statistically significant around the time of onset. The HT models also better replicate the timing of peak wavenumber 2 amplitude for NEG events.

The vertical propagation of planetary waves is inferred here by comparing lags in the timing of peak anomaly signatures between 250 and 50 hPa. The stratospheric variation slightly lags the tropospheric signal, particularly during POS events. From the comparison of intensity with Fig. 7, we can study the extent to which the tropospheric spectrum of planetary waves propagates into stratosphere or not. The comparison for LT models shows that the temporal variation of planetary wave activity in the stratosphere is anomalously weak, even though the tropospheric variation is comparable to that of HT models. This suggests that some of the upward propagating planetary wave activity in the LT models is either reflected/refracted within the troposphere or is deposited near the tropopause. Thus, the LT models appear to exhibit a specific deficiency in representing the upward transfer of anomalous tropospheric planetary wave activity into the stratosphere. This appears to be a possible explanation for the relatively weak polar vortex variability found in LT models.

Fig. 8. As in Fig. 7, but for the pressure level of 50 hPa. Contour interval is 10 m$^2$ s$^{-2}$. 
Based on the strength of planetary wave activity shown in Figs. 7 and 8, we further analyze the daily meridional heat flux associated with individual wave-numbers in order to characterize the vertical wave coupling between the stratosphere and troposphere. Figure 9 shows the composite evolution of the (latitudinally averaged) anomalous meridional heat flux associated with wavenumber 1 for NNR, HT, and LT models. Meridional heat flux is averaged over 60°–90°N latitudinal band. (b),(c),(e),(f) The difference between each model group and NNR is displayed in terms of a multiple of the standard deviation of 10000 random sample groups (shading). The area enclosed by thick red solid lines in (b),(c),(e),(f) denotes regions where the difference between the two model groups exceeds the 2.5% highest (falls below the 2.5% lowest) threshold from the null distribution of 10000 random groups (indicating significance at the 95% confidence level).

FIG. 9. Lead–lag composites of anomalous meridional heat flux of wave 1 (contours; units: K m s$^{-1}$) for (a),(d) NNR; (b),(e) HT; and (c),(f) LT models. Meridional heat flux is averaged over 60°–90°N latitudinal band. (b),(c),(e),(f) The difference between each model group and NNR is displayed in terms of a multiple of the standard deviation of 10000 random sample groups (shading). The area enclosed by thick red solid lines in (b),(c),(e),(f) denotes regions where the difference between the two model groups exceeds the 2.5% highest (falls below the 2.5% lowest) threshold from the null distribution of 10000 random groups (indicating significance at the 95% confidence level).

Based on the strength of planetary wave activity shown in Figs. 7 and 8, we further analyze the daily meridional heat flux associated with individual wave-numbers in order to characterize the vertical wave coupling between the stratosphere and troposphere. Figure 9 shows the composite evolution of the (latitudinally averaged) anomalous meridional heat flux associated with wavenumber 1 for NNR, HT, and LT models. Positive (negative) values are linked to anomalous upward (downward) wave coupling. Negative anomalies are observed in the weeks leading up to the onset of SNAM POS events, which reflects either a reduction in the amount of upward propagation in relation to climatology or, if large enough in amplitude, an actual downward reflection of wave activity within the stratosphere (e.g., Shaw and Perlwitz 2013). In either case, the wave-induced deceleration of the polar vortex is reduced from climatological values, leading to an effective westerly acceleration of the polar vortex. Conversely, during NEG events there is anomalously enhanced upward wave propagation, serving to increase the amount of planetary wave activity in the stratosphere leading to a weakened polar vortex. The relationship between strong (weak) polar vortex and
anomalously downward (upward) wave coupling follows the results of previous studies (Limpasuvan et al. 2004; Polvani and Waugh 2004). In NNR, both POS and NEG coupling signatures persist for at least 30 days, although the relationship with SNAM onset varies. More specifically, the anomalous downward coupling for POS events persists longer than the corresponding upward coupling signature for NEG events.

The CMIP5 simulations well represent the sign of the anomalous wavenumber 1 coupling during SNAM onset. However, the composite evolution of this coupling in the stratosphere differs from NNR. First of all, the persistence of the coupling signature is shorter than NNR for both POS and NEG events. Second, the coupling signature begins later in the models than in NNR. Therefore, the peak timing of vertical wave coupling occurs later than in NNR, which is clearly seen in NEG events. Interestingly, the postonset asymmetry between POS and NEG events observed in NNR virtually disappears for the LT models. In addition to the temporal evolution, the CMIP5 simulations also have deficiencies in representing the coupling intensity, which is underestimated (particularly for the LT models). Unlike for the stratosphere, the composite evolution of wave coupling within the troposphere is not appreciably different among the three composite categories. This may help explain why the tropospheric variation of zonal winds in the LT models is comparable to that in HT and NNR (e.g., Fig. 5).

As mentioned earlier, a negative heat flux anomaly signature does not, by itself, necessarily indicate a specific dynamical mechanism, since it can be associated with either weakened upward wave coupling or (if sufficiently large) downward wave coupling (Shaw and Perlwitz 2013). This issue is addressed by additionally diagnosing the total meridional heat flux associated with wavenumber 1, which can be more directly connected to the net vertical wave propagation. This analysis is presented in Fig. 10, which illustrates that, overall, the net wave propagation is primarily upward, even during positive SNAM events. For POS SNAM onset, we note that there is a brief transient signature of weak downward coupling near the tropopause in both NNR and HT. The lack of a strong downward coupling signal is not surprising, since it has already been demonstrated that there is not a one-to-one correspondence between downward wave coupling and extreme vortex events (Shaw and Perlwitz 2013).

Our examination of heat flux behavior was originally motivated by the idea that the realistic tropospheric behavior in LT models during SNAM events might be associated with a downward reflection of planetary wave activity in the LT models. However, we find no evidence of a net downward coupling in the LT model composite. We note that during NEG events in both NNR and HT, a brief period of net downward wave coupling occurs just after the time the stratospheric zonal wind recovers to become westerly (Figs. 10d,e). Again, however, this signature is absent in the LT model composite. This likely indicates a fundamental difference between LT and HT models in the representation of stratospheric dynamics. The LT models appear unable to adequately represent some of the fundamental stratospheric dynamical signatures associated with SNAM events. A recent study analyzed the general heat flux characteristics in 31 CMIP5 models and found that LT models exhibit distinct biases associated with two types of unphysical behavior near the model top: wave reflection or excessive model damping (Shaw et al. 2014).

Figure 11 presents a composite diagnostic analysis of the behavior of anomalous wave activity during SNAM onset. The vectors represent the Eliassen–Palm (EP) flux, while the contours represent the net impact of the anomalous wave activity flux upon the zonal-mean zonal wind field. Unlike the previous analysis, this analysis encompasses the entire spectrum of wave time and space scales. EP flux vectors provide direct information on the propagation of Rossby wave activity in the latitude–pressure plane (Edmon et al. 1980). The local convergence (divergence) of the EP flux is proportional to the local deceleration (acceleration) of zonal-mean zonal winds in association with Rossby wave breaking. Thus, the local EP flux convergence (or divergence) represents the so-called wave driving (WD) of the zonal-mean flow (Black and McDaniel 2007; Palmer 1982). We have analyzed EP flux and WD anomalies for the three SNAM evolution phases, and the results for the onset phase are displayed in Fig. 11.

For NNR, the primary EP flux anomaly signature is a downward and equatorward (upward and poleward) flux in midlatitudes for POS (NEG) cases, qualitatively resembling the signatures obtained by Limpasuvan et al. (2004, 2005) during the growth phase of SSW and vortex intensification events, respectively. The CMIP5 simulations well replicate the overall EP flux anomaly pattern, except for an excessive equatorward flux component within the lower stratosphere for POS events. Also, there is excessive meridional refraction into the subtropics for NEG events in the HT models. In terms of intensity, the HT models represent the observed EP flux...
reasonably well. On the other hand, the LT models underestimate the observed strength of the anomalous EP flux, particularly the vertical component, at higher latitudes above the midtroposphere (purple contours).

The patterns of local WD (contours) generally match the change of zonal-mean zonal winds shown in Fig. 6. For NEG events, the convergence of wave activity flux serves to decelerate the stratospheric polar vortex, and this is indicated by a vertically coherent negative WD pattern. In contrast, for POS events, there is robust anomalous downward EP flux within the stratosphere, suggesting that there is either less wave activity available, or there is a vertical trapping of the existing spectrum of planetary wave activity. This leads to a positive WD signature and a strengthening of the stratospheric polar vortex.

Regarding the vertical structure of the WD patterns, the vertical level of maximum WD amplitude in the stratosphere is displaced downward in the CMIP5 composites compared to NNR for both POS and NEG events. This appears somewhat inconsistent with the pattern of zonal wind change. Contrasting the WD magnitudes, we find the local stratospheric WD in the LT models to be weaker than either NNR or HT models, consistent with the zonal wind change patterns observed in Fig. 6. In contrast to the stratospheric comparison, differences in the tropospheric signatures of EP and WD are relatively weak. All three composites exhibit similar tropospheric structures in the direction and strength of the EP flux and the WD magnitude. The results demonstrate that the previously noted LT SNAM behavior of weakened stratospheric polar vortex variability

**Fig. 10.** As in Fig. 9, but for total meridional heat flux of wave 1. Contour interval is 10 K m s$^{-1}$. 
combined with robust tropospheric signatures is dynamically driven.

Figure 12 presents a composite diagnosis of the total wave activity during SNAM onset. General propagation of EP flux is upward in the lower troposphere and stratosphere, while the intensity becomes weak (strong) for POS (NEG) events. Unlike in Fig. 11, the vertical EP flux and WD fields for NNR, HT, and LT models look quite similar in terms of amplitude and structure. However, the meridional component of the EP flux shows some differences (purple contours). Specifically for POS SNAM, both model groups have a statistically significant weakening in the meridional flux at higher latitudes. For POS SNAM events, the observed equatorward EP flux around 40°N in lower stratosphere is anomalously weak in both HT and LT models. On the other hand, during NEG SNAM events, the HT models exhibit an anomalously strong stratospheric equatorward EP flux compared to both NNR and the LT model group.

5. Discussion and conclusions

We have studied how well CMIP5 simulations replicate the stratospheric northern annual mode (SNAM), including its vertical structure, temporal behavior, and attendant stratosphere–troposphere dynamical coupling. The models are categorized in terms of their representation of the stratosphere (high-top vs low-top models). For NNR and individual CMIP5 simulations, SNAM is identified as the leading mode of a principle component analysis of the lower stratospheric zonal-mean zonal wind over the domain encompassing 45°–90°N, 100–10 hPa during boreal winter (JFMA). To analyze the temporal variability of the stratospheric polar vortex and its associated tropospheric footprint, large-amplitude SNAM events are selected from NNR and each CMIP5 simulation. We then contrast the PC loading patterns, associated tropospheric structure, event statistics, and the composite

Fig. 11. Anomaly composite of EP flux (vectors; units: m² s⁻²) and local wave driving of zonal-mean flow (black thin contours; units: m s⁻¹ day⁻¹; contour interval: 0.5 m s⁻¹ day⁻¹) for the period extending from 5 to 1 days prior to event onset for (a),(d) NNR; (b),(e) HT; and (c),(f) LT models. Upper (lower) panels are for positive (negative) SNAM events. Contour interval is 0.5 m s⁻¹ day⁻¹. The difference of vertical flux (purple thick contours of −5, −3, 3, and 5 levels) between each model group and NNR is displayed in terms of a multiple of the standard deviation of 10 000 random sample groups.
time evolution of the zonal wind field and associated dynamical measures.

The loading pattern analysis reveals that CMIP5 models qualitatively replicate the primary structural features of observed SNAM events. However, the HT models slightly overestimate the intensity of associated polar vortex variability, while LT models strongly underestimate polar vortex variability, consistent with the results of Charlton-Perez et al. (2013). In LT models, we find that negative SNAM events are not typically associated with a wind reversal to stratospheric easterlies, consistent with a reduced frequency of SSWs in LT models (Charlton-Perez et al. 2013). Statistical analyses show that for the CMIP5 models, SNAM events are typically more frequent but less persistent than in NNR. Also, observed asymmetries between positive (POS) and negative (NEG) events statistics are not fully replicated by either HT or LT models. Composite evolution analysis of SNAM events reveals that the strongest zonal wind anomalies are less persistent in LT models. Interestingly, these analyses further suggest that the attendant tropospheric zonal wind signature in LT models closely resembles NNR despite suppressed stratospheric polar vortex variability in the LT models (Figs. 2 and 5).

A regression analysis is used to isolate the regional tropospheric manifestations of SNAM. During POS (NEG) SNAM events the North Atlantic storm track shifts poleward (equatorward) and the near-surface large-scale circulation anomaly pattern resembles the positive (negative) phase of the Arctic Oscillation. Both of these signatures are qualitatively well represented in both HT and LT models, with a slight weakening and eastward shift observed for the LT models. Generally speaking, since annular modes exhibit an equivalent barotropic vertical structure, the intensity of the tropospheric signature associated with stratospheric polar vortex variability is expected to be proportional to the

---

**FIG. 12.** Total composite of EP flux (vectors; units: $m^2 s^{-2}$) and local wave driving of zonal-mean flow (contours; units: $m s^{-1} day^{-1}$) for the period extending from 5 to 1 days prior to event onset for (a),(d) NNR; (b),(e) HT and (c),(f) LT models. Upper (lower) panels are for positive (negative) SNAM events. Contour interval is 2.0 $m s^{-1} day^{-1}$. The difference of meridional EP flux (purple thick contours of $-5$, $-3$, $3$, and $5$ levels) between each model group and NNR is displayed in terms of a multiple of the standard deviation of 10 000 random sample groups.
intensity of SNAM. Thus, it is not too surprising that the LT models exhibit a slightly weaker response, given the general weakness of polar vortex variability in LT. In fact, the tropospheric signature in the LT models is perhaps more robust than one would expect based upon a proportional scaling based upon the stratospheric SNAM amplitude (e.g., NNR versus LT).

Parallel dynamical analyses of troposphere–troposphere wave coupling are pursued via a collective consideration of wave power spectra, meridional heat fluxes, EP fluxes and wave driving patterns. We try to delineate the physical reasons behind the relatively deep vertical structure observed in LT models by a consideration of differences in stratosphere–troposphere coupling among NNR, HT, and LT composites.

A spatial power spectrum analysis applied near the tropopause (250 hPa) reveals little difference between HT and LT models in representing the variation of planetary wave activity. If anything, the LT models exhibit a slightly better representation (than HT) of the wave amplitude and scale interaction (Fig. 7). However, a parallel analysis performed in the lower stratosphere (50 hPa) indicates that the spectrum of planetary wave activity in LT models is much weaker than in HT models (Fig. 8). Thus, the (well represented) upper-tropospheric planetary wave activity in LT models appears unable to effectively propagate into the lower stratosphere, leading to suppressed variability in the polar vortex. Taken together, our results suggest an inadequate representation of basic dynamical processes in the lower stratosphere of LT models, perhaps related to insufficient vertical resolution.

Meridional heat flux analyses illustrate that the upward propagation of wavenumber 1 is reduced (enhanced) during POS (NEG) events. During POS event onset, downward wave coupling occurs briefly (~2 days) in the lower stratosphere in both NNR and HT models. Vertically propagating Rossby waves can be reflected downward under certain conditions; therefore, reflected waves may serve to alter tropospheric variability (Shaw and Perlwitz 2013). We initially speculated that the reflection or refraction of tropospheric planetary waves in LT models may compensate for a presumably weak direct downward influence from SNAM, leading to a relatively robust tropospheric signal in LT models. A recent study focused on heat flux behavior in CMIP5 models appears to support this possibility with a finding that downward wave coupling is overrepresented in the subset of LT models (Shaw et al. 2014). However, in our analyses of SNAM events, we find little evidence of enhanced downward coupling in LT models. In fact, the heat flux anomaly magnitudes are suppressed in our LT model composites (vs NNR) for both POS and NEG SNAM events (see Fig. 9), suggesting that our LT model pool is more consistent with a second subset of LT behavior described by Shaw et al. (2014), linked to excessive damping near the model top. In any case, we suggest that downward coupling is not a leading contributor to the robust tropospheric signature observed in association with stratospheric polar vortex variation (SNAM) in LT models. An important caveat to the above discussion is that the model pool considered consists of individual members with differing biases in representing the general nature of heat flux characteristics. Additional research will be required to more fully ascertain the physical reasons behind the robust tropospheric signature observed in the LT model group.

We have diagnosed composite anomalies in the EP flux and zonal wave driving with a focus on the SNAM onset period. The CMIP5 simulations represent the overall EP flux patterns reasonably well, with minor discrepancies in the meridional component. However, LT models notably underestimate the magnitude of the EP flux anomalies at stratospheric altitudes. Nonetheless, all three composites (NNR, HT, and LT) show dynamical consistency between the patterns of anomalous wave driving and the concomitant zonal wind change, suggesting that the transformed Eulerian-mean framework does correctly characterize the nature of the zonal flow dynamics during SNAM onset.

In the stratosphere, anomalous wave driving in LT models is considerably weaker in magnitude than found in NNR and HT models, consistent with the relative weakness in the LT SNAM amplitude and associated EP flux anomaly pattern. Interestingly, the tropospheric signatures of both EP flux and zonal wave driving are well represented in LT models during SNAM onset. This implies that during SNAM onset, LT models are able to correctly capture the tropospheric dynamical response even though the stratospheric response is seriously deficient. Therefore, we conclude that the impact of SNAM events on the tropospheric circulation variability is (i) dynamically driven and (ii) not strongly dependent upon how realistically the models employed represent stratospheric variability. We note, however, that our study focuses on the proximate short-term tropospheric response occurring during the period of SNAM onset. It is unlikely that our conclusions carry over to the weaker-amplitude long time-scale response that is dependent upon SNAM persistence in the lower stratosphere. Although this topic is also of interest, it lies outside the research scope of the current study.
Acknowledgments. This research was conducted under support by the U.S. Department of Energy, Office of Biological and Environmental Research, Award DE-SC0004942. We thank the anonymous peer reviewers for their helpful comments and feedback in the review process.

REFERENCES


