Solar wind dynamic pressure effect on planetary wave propagation and synoptic-scale Rossby wave breaking

Hua Lu,1 Christian Franzke,1 Olivia Martius,2 Martin J. Jarvis,1 and Tony Phillips1

Received 19 November 2012; revised 27 March 2013; accepted 27 March 2013; published 30 May 2013.

We provide statistical evidence of the effect of the solar wind dynamic pressure ($P_{sw}$) on the northern winter and spring circulations. We find that the vertical structure of the Northern Annular Mode (NAM), the zonal mean circulation, and Eliassen-Palm (EP)-flux anomalies show a dynamically consistent pattern of downward propagation over a period of ~45 days in response to positive $P_{sw}$ anomalies. When the solar irradiance is high, the signature of $P_{sw}$ is marked by a positive NAM anomaly descending from the stratosphere to the surface during winter. When the solar irradiance is low, the $P_{sw}$ signal has the opposite sign, occurs in spring, and is confined to the stratosphere. The negative $P_{sw}$ signal in the NAM under low solar irradiance conditions is primarily governed by enhanced vertical EP-flux divergence and a warmer polar region. The winter $P_{sw}$ signal under high solar irradiance conditions is associated with positive anomalies of the horizontal EP-flux divergence at 55°N–75°N and negative anomalies at 25°N–45°N, which corresponds to the positive NAM anomaly. The EP-flux divergence anomalies occur ~15 days ahead of the mean-flow changes. A significant equatorward shift of synoptic-scale Rossby wave breaking (RWB) near the tropopause is detected during January–March, corresponding to increased anticyclonic RWB and a decrease in cyclonic RWB. We suggest that the barotropic instability associated with asymmetric ozone in the upper stratosphere and the baroclinic instability associated with the polar vortex in the middle and lower stratosphere play a critical role for the winter signal and its downward propagation.


1. Introduction

Solar activity affects the ozone-sensitive stratospheric circulation by means of photochemistry or catalytic reactions. The increase of solar ultraviolet (UV) irradiance at the solar maximum enhances the ozone concentration, making the equatorial upper stratosphere warmer compared to the solar minimum [Hood et al., 1993; Haigh, 1994; Gray et al., 2010]. This thermal perturbation increases the equator to winter-pole temperature gradient and thereby alters the upward propagation of planetary waves. Via a dynamic feedback between the planetary waves and the mean flow, it leads to a strengthened polar vortex. It has been proposed that the solar UV forcing, which originates in the upper equatorial stratosphere, may interact with the planetary waves and affect the polar vortex dynamically; such solar-induced effects propagate poleward and downward as polar vortex anomalies during winter [Kodera and Kuroda, 2002]. General circulation models (GCMs), however, tend to produce a much smaller atmospheric response to the UV forcing than the signals of the 11 year solar cycle found in reanalysis data sets [Matthes et al., 2006; Gray et al., 2010]. In particular, direct radiative effects due to total solar or UV irradiances appear to be too small to cause the temperature responses at high latitudes [Labitzke et al., 2006; Lu et al., 2009], suggesting that either certain amplification processes are not adequately represented in GCMs or that solar irradiance may not be the only solar forcing affecting the radiative balance of the stratosphere [Gray et al., 2010].

Ionized energetic particle precipitation (EPP) generated during solar wind disturbances produces odd nitrogen ($\text{NO}_x$) and odd hydrogen ($\text{HO}_x$) species in the Earth’s mesosphere and lower thermosphere (MLT) [Brasseur and Solomon, 2005; Randall et al., 2005]. During polar winter, EPP-$\text{NO}_x$ can be transported into the stratosphere in polar darkness where it has a lifetime of several months or longer, allowing it to engage in catalytic ozone destruction [Solomon et al., 1982; Randall et al., 2005, 2009; Reddmann et al., 2010]. Several studies have shown that the EPP-$\text{NO}_x$ produced routinely in the MLT by low to medium energetic auroral electron precipitation events may lead to circulation changes in the stratosphere and the troposphere [Seppälä et al., 2009; Reddmann et al., 2010; Baumgaertner et al., 2011; Semeniuk et al., 2011].
[4] The dynamic response to the intrusion of EPP-NO\textsubscript{x} produced by low-energy electrons at the MLT region have been studied using chemistry climate models [Callis et al., 1996; Langematz et al., 2005; Rozanov et al., 2005; Baumgaertner et al., 2011; Semeniuk et al., 2011]. Most of the models showed a significant ozone loss in the stratosphere and mesosphere but with varying magnitude and spatial extent. Some models showed that direct downward descent of EPP-NO\textsubscript{x} into the stratosphere rarely occurred and suggested that the effects on stratospheric O\textsubscript{3} and the temperature in the polar middle and upper stratosphere are indirect or have a dynamic origin [Marsh et al., 2007]. Others showed that EPP-NO\textsubscript{x} may contribute to 5% of stratospheric NO\textsubscript{x} which may cause a reduction of high-latitude stratospheric ozone for months after the intrusion [Reddmann et al., 2010]. Earlier models estimated a large-scale temperature reduction in the stratosphere with significant ozone loss at low and midlatitudes [Langematz et al., 2005; Rozanov et al., 2005]. Most recent modeling studies however showed that the temperature response in Northern Hemisphere (NH) winter is marked by an alternating vertical structure at high latitudes, with a warmer signal in the upper stratosphere and the mesosphere and a cooling signal in the lower stratosphere [Baumgaertner et al., 2011; Semeniuk et al., 2011]. In winter months, the anomalous circulation pattern in the lower stratosphere and troposphere in response to EPP-NO\textsubscript{x} appears to resemble a positive Northern Annular Mode (NAM) signature, similar to that which has been obtained from reanalysis data in response to the geomagnetic Ap index [Lu et al., 2008a; Seppälä et al., 2012].

[5] Studies have shown that solar wind related parameters such as the geomagnetic Ap index are good proxies for EPP-NO\textsubscript{x} [Siskind, 2000; Seppälä et al., 2007; Lu et al., 2008b]. Statistically significant correlations between stratospheric and tropospheric circulation anomalies and solar wind parameters have been reported [e.g., Lu et al., 2008a, 2008b; Woollings et al., 2010; Seppälä et al., 2009]. Boberg and Lundstedt [2002] found that the variation of the winter North Atlantic Oscillation (NAO) is correlated with the electric field strength of the solar wind. Changes of surface temperature have also been linked to geomagnetic perturbations [Seppälä et al., 2009]. The signature of the open solar flux, which is a measure of annual to decadal variation of solar wind strength and geomagnetic activity, was found over the North Atlantic/Eurasian sector, where there were significant changes in the occurrence of blocking and the winter mean surface temperature. The open solar flux signal was found to be statistically stronger than that of the F10.7 cm solar flux, a proxy for solar irradiance; it implies that solar wind perturbations may have a stronger influence than solar irradiance at high latitudes during winter [Woollings et al., 2010]. Lu et al. [2008b] found that the springtime polar stratospheric temperature response was opposite to what would be expected from catalytic reactions from in situ EPP-NO\textsubscript{x} and stratospheric ozone. They consequently suggested that the stratospheric response is more likely due to an indirect dynamic effect. A recent modeling study of Kvissel et al. [2012] shows that EPP-NO\textsubscript{x} enhancement induces anomalously stronger planetary waves in the stratosphere in association with a weaker and warmer polar vortex in spring, in good agreement with the observational study by Lu et al. [2008b]. Kvissel et al. [2012] suggest a dynamic pathway that involves a poleward transfer of existing stratospheric nitric acid which interacts with zonally asymmetric O\textsubscript{3} and short wave heating. Nevertheless, it remains unclear as to how the upper atmospheric perturbations are transferred dynamically downward to influence the lower stratosphere and troposphere circulations.

[6] Lu et al. [2008a] showed in a statistical study that northern winter stratospheric wind and temperature variations are related to the variation of solar wind dynamic pressure, which is a key parameter that describes the strength of the Sun’s coronal magnetic field and determines the geomagnetic Ap index and open solar flux [Lockwood et al., 2009]. Lu et al. [2008a] found that, during January–March, the solar wind dynamic pressure signal in the zonal mean zonal wind is marked by an equivalent barotropic structure, in which the wind anomalies extend from the surface into the stratosphere and have no phase shift or tilt with height from the surface to 30 hPa. It is associated with a strengthening of the prevailing westerly winds at 50°N–70°N and weakening of the winds at 30°N–40°N, resembling the positive phase of the Northern Annular Mode (NAM) that is closely related to the NAO at the surface [Feldstein and Franzke, 2006]. The winter signal is only robust at the solar maximum or when the F10.7 cm solar flux is high. Lu et al. [2008a] suggested that solar wind perturbations induce equatorward planetary wave refractions that cause the wind and temperature signals seen in the reanalysis data. Most recently, Seppälä et al. [2012] tested the wave refraction hypothesis proposed by Lu et al. [2008a] using the monthly mean Eliassen-Palm (EP) flux along with the zonal mean zonal wind and temperature from ERA-40 and ERA-Interim reanalyses. Seppälä et al. found that a significant enhancement of equatorward refraction of planetary waves in the upper stratosphere is associated with a high geomagnetic Ap index in the period January–March. The geomagnetic Ap signature in the EP flux is marked by anomalous convergence (divergence) in the upper (lower) stratosphere, corresponding to a warmer (colder) upper (lower) stratosphere. The corresponding zonal mean zonal wind response resembles the positive phase of the NAM, in good agreement with Lu et al. [2008a] and Baumgaertner et al. [2011].

[7] The NAM has an equivalent barotropic vertical structure and extends from the Earth’s surface into the stratosphere [Thompson and Wallace, 2000]. During the NH winter, the NAM is the leading mode of variability in the midlatitudes and high latitudes, describing the north-south shifting of the jet stream. While the NAM is directly linked to the strength of the polar vortex in the stratosphere [Baldwin et al., 2003], in the troposphere it is related to the position of the eddy-driven jet and is due to baroclinic wave breaking [Gerber and Vallis, 2007]. Baroclinic instability arises from horizontal temperature gradients, converts potential into kinetic energy, and is, for example, responsible for the formation of cyclones and is often associated with changes of the vertical wind shear. Wave-mean flow interaction, such as baroclinic instability, plays a key role in determining the variation of the NAM and its coupling from the stratosphere to the troposphere [Thompson et al., 2002; Simpson et al., 2009]. To understand the positive correlation between the NAM and the solar wind dynamic pressure or solar wind induced geomagnetic activity previously reported [Lu et al., 2008a; Baumgaertner et al., 2011; Semeniuk et al., 2011; Seppälä et al., 2012], we need...
to examine to what extent it affects the planetary wave propagation and breaking in the stratosphere and how they are linked to the strength of the polar vortex.

[5] Previous studies have shown that the stratospheric anomalies associated with the polar vortex may descend into the troposphere through an alteration of the vertical wind shear, therefore affecting synoptic-scale Rossby wave breaking (RWB) near the tropopause, which is a potential mechanism linking the tropospheric and stratospheric circulations [Kunz et al., 2009a, 2009b]. When they break, synoptic-scale Rossby waves affect the position and strength of the midlatitude jet stream through meridional redistribution of momentum. Thus, RWB is the major driving mechanism of the tropospheric NAM/NAO [Franzke et al., 2004; Feldstein and Franzke, 2006]. There are two main types of RWB: anticyclonic (LC1) and cyclonic (LC2) [Thornicroft et al., 1993]. An increase of LC1 (decrease of LC2) type of RWB in the Atlantic basin is consistently linked to the positive (negative) NAM/NAO [Franzke et al., 2004; Martius et al., 2007; Kunz et al., 2009b]. In addition, LC1 (LC2) RWB is associated with a poleward (equatorward) shift in the jet stream over the Atlantic. The stratospheric-driven, low-frequency NAM anomalies may arise in the troposphere because of more or less frequent RWB events for an extended period [Martius et al., 2007; Kunz et al., 2009a]. In order to understand how solar forcing affects the tropospheric NAM, it is crucial to examine whether or not there is any systematic change of RWB near the tropopause on a monthly to intraseasonal time scale.

[6] In this paper, we use the ERA-40 and ERA-Interim data sets to examine the responses of planetary waves as well as anomalous synoptic-scale RWBs to the changes in the solar wind dynamic pressure. We focus on the temporal progression of the NAM, zonal mean zonal wind and temperature, and EP-flux divergence in relation to changes in the solar wind dynamic pressure (Psw). Psw is a heliospheric parameter and has been observed to cause large-scale auroral brightening with associated geomagnetic storms, which produces EPP-NOx in the MLT [Lyons et al., 2008; Yang et al., 2011]. Studies show that an increase of Psw under southward interplanetary magnetic field (IMF) conditions triggers a global enhancement of auroral intensity at all local times instantaneously [Boudouridis et al., 2003].

[10] Psw is one of a few key parameters that describe the coupling between the solar wind and the Earth’s magnetosphere [Svalgaard, 1977; Finch and Lockwood, 2007]. The parameters that went into estimating Psw were measured well away from the Earth’s atmosphere including using spacecraft orbiting around the L1 Lagrangian point between the Earth and the Sun, which is ~1.5 million km from the Earth. This is well outside of Earth’s magnetopause in the IMF and hence is completely unaffected by the Earth’s atmosphere. Richardson et al. [1998] have shown that, for 6-hourly averages, correlations for each of solar wind speed, density and flux at the L1 point, and the same parameter in the near-Earth plasma environment are all around 0.6. By contrast, the geomagnetic indices, such as the Ap index, are derived through indirect measurement of ionospheric currents that reach the maximum between 110 and 140 km above the Earth’s surface. These are not only influenced by the solar wind but also by the magnitude and the direction of the IMF and the secular changes in the Earth’s magnetic fields [Finch and Lockwood, 2007]. Ap is also affected by the observer’s ability to identify solar quiet day variation (Sq), i.e., diurnal variation of the geomagnetic field, which is affected by tidal dynamics [Mayaud, 1980]. It is well documented that the atmospheric tides in the lower thermosphere can vary dramatically from day to day due to complex dynamical interactions of the neutral atmosphere, including the nonlinear interaction of nonmigrating tides with solar-induced tides [Hagan and Forbes, 2002]. Consequently, the Ap index can be affected by the atmospheric parameters that we are using as tracers for the solar wind influence on the Earth’s atmosphere and is therefore not a truly independent variable for analyses based on correlations. In addition, the Ap signal appears to be contaminated by the occurrence of major sudden stratospheric warming (SSWs) [Lu et al., 2008b; Seppälä et al., 2012], while the Psw signal is insensitive to the SSWs [Lu et al., 2008a; this study]. For these reasons, we use Psw to quantify the solar wind. Nevertheless, it can be shown that similar but weaker results to those presented here can be obtained by using the geomagnetic Ap index.

[11] In this study, we examine the relationship between the solar wind dynamic pressure, EP-flux divergence, and mean-flow response. We focus on variations that occur at time scales of from a few weeks to ~2 months, which are known to be associated with the stratosphere-driven time scales in terms of stratosphere-troposphere dynamic coupling [Baldwin et al., 2003]. Unlike the earlier studies that used monthly data [e.g., Lu et al., 2008a], the results reported here are based on daily data. This provides more detailed temporal information including the life cycle of the signals and the downward propagation of the circulation anomalies. Analogous to Lu et al. [2008a], the analysis is carried out according to high and low solar irradiances (referred to as HS and LS, respectively) to examine the potential solar UV conditioning of the stratospheric circulation. To investigate how the solar wind dynamic pressure affects the NAM in the troposphere, we also examine one of its key dynamical drivers: Rossby wave breaking (RWB) near the tropopause [Martius et al., 2007].

2. Data and Methods

[12] We use monthly mean geopotential height and daily temperature and zonal and meridional winds from the European Centre for Medium-range Weather Forecasts (ECMWF) ERA-Interim data [Dee et al., 2011] for the period 1979–2009, backward extended with ECMWF ERA-40 reanalysis [Uppala et al., 2005] for 1958–1978 to study the stratospheric and tropospheric response. For the majority of our analysis except for synoptic-scale waves, we use 2.5° horizontal resolution on 23 pressure levels from 1000 hPa to 1 hPa. This horizontal resolution and vertical extent are sufficient for an examination of the large-scale planetary waves and mean flow interaction throughout the stratosphere and the troposphere. Although it is not ideal to merge two datasets derived from different data assimilation models, we have done this in order to maximize the length of the data set for a more robust statistical analysis. It also means that data from the presatellite era have been included, which are known to be less reliable in the upper stratosphere. Nevertheless, qualitatively similar results can be obtained either from the full-length ERA-40 (January 1958 to September 2001) or from ERA-Interim (January 1979 to December 2009). The
results reported here make use of all available data blended from ERA-40 and ERA-Interim.

The daily northern annular mode (NAM) index is derived as the leading empirical orthogonal function (EOF) of daily zonal mean zonal wind over 20°N–90°N, from the blended ERA-40 and ERA-Interim reanalyses for the same period. Details about this method can be found in Baldwin and Thompson [2009]. To focus on intraseasonal to seasonal variations, a 180 day high-pass filter was applied to all time series of atmospheric variables such as zonal mean zonal wind \( u \), zonal mean temperature \( T \), the NAM, and the EP fluxes. Thus, the effect of climate variability on longer time scales are likely eliminated.

The NAM variability is linked to wave-induced momentum transport. The extratropical quasi-geostrophic zonal mean zonal momentum equation is governed by the transformed Eulerian-mean equations [Andrews et al., 1987]

\[
\frac{\partial \bar{u}}{\partial t} = f v + \nabla \cdot \left( \frac{\nabla F}{p_0 a \cos \phi} \right)
\]

(1)

where

\[
\nabla F = (a \cos \phi)^{-1} \frac{\partial}{\partial \phi} (F^\phi \cos \phi) + \frac{\partial F^{\phi \cos \phi}}{\partial z}
\]

(2)

and

\[
F^\phi = -p_0 a \cos \phi u
\]

\[
F^{\phi \cos \phi} = p_0 a \cos \phi \left[ f - (a \cos \phi)^{-1} (\bar{u} \cos \phi) \right] \sqrt{\theta / \theta_0}
\]

(3)

where \( u, v, \) and \( \theta \) denote the zonal wind, meridional wind, and potential temperature; theobar over the zonal mean and primes denote the departure from the zonal mean, \( f \) the Coriolis parameter, \( \phi \) the latitude, \( a \) the radius of the Earth, and \( \theta \) the log-pressure vertical coordinate. Equation (1) can be understood such that the acceleration of the zonal mean wind is equal to the Coriolis force acting on the residual mean meridional circulation of the flow \( (f \bar{v}) \) and wave divergence \( (\nabla \cdot \nabla F) \) as the friction in the upper troposphere and stratosphere is small and can be neglected. \( \nabla \cdot \nabla F \) is essentially the wave property required for driving departures from the radiatively determined state [Dickinson, 1968]. The divergence of the Eliassen-Palm (EP) flux \( \nabla \cdot F \) consists of the meridional divergence of the zonal momentum flux \( (\bar{F}^\phi) \) and vertical divergence of the heat flux \( (F^{\phi \cos \phi}) \). We show in Figure 1 that the extratropical zonal mean zonal wind can be accelerated when \( \nabla \cdot F > 0 \) through the meridional convergence of the eddy momentum flux or the vertical convergence of the heat flux or decelerated when \( \nabla \cdot F < 0 \) due to an enhanced wave breaking. The dynamics that give rise to \( \nabla \cdot F \) affect the zonal mean zonal flow, and thus the NAM should be reflected in the temporal variability of the zonal mean zonal wind \( \bar{u} \) and temperature \( T \). If the atmospheric circulation response to the solar wind dynamic pressure originates from changes in dynamics, the associated anomalies of \( \bar{u} \) and \( \nabla \cdot F \) should obey equation (1), and the temperature anomalies should behave in the same way as what is depicted in Figure 1.

The solar wind dynamic pressure \( P_{sw} \) measured in Geocentric Earth Magnetic (GEM) coordinates was extracted from National Aeronautics and Space Administration Goddard Space Flight Center (NASA/GSFC) OMNIWeb site (http://omniweb.gsfc.nasa.gov/). This data set is produced from solar wind and interplanetary magnetic field measurements from 15 geocentric satellites and three spacecraft in orbit around the L1 Sun-Earth Lagrange point and has been carefully compiled through cross-calibration. \( P_{sw} \) has been calculated by NASA-OMNI as \( P_{sw} = 1.6726 \times 10^{-8} N_{sw} V_{sw}^2 \), where \( N_{sw} \) is the flow density in number of particles per cubic centimeter and \( V_{sw} \) is the solar wind speed in kilometers per second.
In physical terms, $P_{\text{sw}}$ represents the momentum flux of the solar wind and has a unit of nanopascals (nPa). Daily averages of $P_{\text{sw}}$ from January 1963 to December 2009, covering ~4.5 solar cycles, are used in this study. The time series of monthly mean $P_{\text{sw}}$ for the period of 1963–2009 are shown in Figure 2. As we seek the dynamic response to the accumulated effect of EPP-NOX from the MLT region over a period of a few months, high-frequency, short-term changes in $P_{\text{sw}}$ will be largely averaged out and are therefore not accounted for by this study.

As can be seen in Figure 2, the correlation between monthly $P_{\text{sw}}$ and the well-known 11 year solar cycle (i.e., 10.7 cm solar flux shown as the black line) is low and not statistically significant at the $p=0.05$ level, largely due to the high-frequency variation of $P_{\text{sw}}$. Nevertheless, the low-frequency $P_{\text{sw}}$ does have a quasi-decadal variation, with a minimum occurring near the solar maxima. Before 1990, the decadal-scale minima of $P_{\text{sw}}$ were followed by a rapid increase of $P_{\text{sw}}$ during the declining phase of the 11 year solar cycle then a slow decrease until the next solar maximum. The noticeably smaller increase in $P_{\text{sw}}$ since 1990, compared to previous cycles, contributes to the low correlation between $P_{\text{sw}}$ and the 10.7 cm solar flux [Richardson and Kasper, 2008]. This weak statistical dependence allows the signals associated with the 11 year solar irradiance and the solar wind dynamic pressure to be separated. However, as we will demonstrate in section 3, their mechanisms may not be mutually exclusive, and they could be acting together to either enhance or cancel the overall atmospheric responses.

<table>
<thead>
<tr>
<th>Year Group</th>
<th>Year List</th>
</tr>
</thead>
</table>

The subgroups for the HS (high solar irradiance) and LS (low solar irradiance) conditions are given in the bottom four rows. The years affected by major volcanic eruptions (1983–84, 1992–1993) and major ENSO events (1972–73, 1982–83, 1997–98) are listed as bold and italic.

**Table 1. The Years With Low and High Solar Wind Dynamic Pressure (LP and HP) for January–March Mean (Top Two Rows)**

**Figure 3.** Daily running regression between solar wind dynamic pressure ($P_{\text{sw}}$) and the NAM in time (x axis) and altitude (y axis) cross-section. In the left-hand panels, the daily NAM is used. In the right-hand panels, a 31 day running window is applied to the NAM. Contours are at $\pm 0.5$, $\pm 1$, $\pm 1.5$, ... nPa$^{-1}$. The area covered by white dots indicates significance at the $p=0.05$ level. (a) When all the data are included; (b) when solar irradiance is high (HS); (c) when solar irradiance is low (LS). The data affected by major volcano eruptions are excluded. Only low-frequency $P_{\text{sw}}$ (>365 days) is used (see Figure 2). Similar results are obtainable with a low-pass filter with a cutoff period of $t \geq 90$ days.
At the latitude bands that showed the most significant changes, which are reported here. On the basis of those results, we identify the latitude or latitudinal band of the tropospheric midlatitude jet stream, it is necessary to study the anomalous values of $\Psi^{(0)}$, $\Psi^{(2)}$, and $\Psi$ separately. For a given latitude, daily values of $\Psi^{(0)}$, $\Psi^{(2)}$, and $\Psi$ were calculated based on daily ERA-40 and ERA-Interim data and using equations (1) to (6). The running regression was performed by including all data from 1965 to 2009 as well as by subgrouping the data according to the level of the solar irradiance. This allows us to assess the modulating effect of the solar UV in the stratosphere on the $P_{sw}$ signature. Daily data of 10.7 cm solar radio flux

\[
\Psi^{(0)} = (\rho_0 a \cos \phi)^{-1} \nabla F^{(0)} = \rho_0 a \cos \phi^{-1} \frac{\partial F^{(0)}}{\partial z} \cos \phi \tag{4}
\]

\[
\Psi^{(2)} = (\rho_0 a \cos \phi)^{-1} \nabla F^{(2)} = (\rho_0 a \cos \phi)^{-1} \frac{\partial F^{(2)}}{\partial z} \cos \phi \tag{5}
\]

\[
\Psi = \Psi^{(0)} + \Psi^{(2)} = (\rho_0 a \cos \phi)^{-1} \nabla F \tag{6}
\]

to diagnose the eddy momentum flux, heat flux, and total EP-flux contributions to the changes detected in the mean flow. Analyzing $\Psi^{(0)}$, $\Psi^{(2)}$, and $\Psi$ separately allows us to evaluate whether the horizontally propagating waves (indicated by $\Psi^{(0)}$) or whether the vertically propagating waves (largely determined by $\Psi^{(2)}$) are most responsible for the signal in the circulation anomalies, i.e., anomalies detected in $\tilde{u}$ and $\tilde{T}$. In addition, as the meridional redistribution of momentum flux plays a unique role in affecting the position and strength of the tropospheric midlatitude jet stream, it is necessary to study the anomalies of $\Psi^{(0)}$, $\Psi^{(2)}$, and $\Psi$ separately. For a given latitude, daily values of $\Psi^{(0)}$, $\Psi^{(2)}$, and $\Psi$ were calculated based on daily ERA-40 and ERA-Interim data and using equations (1) to (6). The running regression was performed by including all data from 1965 to 2009 as well as by subgrouping the data according to the level of the solar irradiance. This allows us to assess the modulating effect of the solar UV in the stratosphere on the $P_{sw}$ signature. Daily data of 10.7 cm solar radio flux

\[
\frac{d}{dz} \frac{F}{a} \cos \phi \tag{1}
\]

\[
\frac{d}{dt} \frac{F}{a} \cos \phi \tag{2}
\]

\[
\frac{d}{dz} \frac{F}{a} \cos \phi \tag{3}
\]
As an additional diagnostic, a composite analysis of geopotential height ($Z$), in which the data were separated into high and low $P_{sw}$ (HP and LP hereafter) for $Alldata$, $HS$ and $LS$ conditions, was also carried out, where HP and LP are defined following the same approach as for $F_{10.7}$. As an example, the years for the January–March mean in the HP and LP cases as well as HSHP, HLSL, LSHP and LSLP cases are listed in Table 1. Nevertheless, the year groups can change slightly if different months are used to take the mean value of $P_{sw}$, and this is the case in performing our running regression from November to May. To examine the vertical connection of the $P_{sw}$ signals, the composite analysis was done at six different pressure levels, of which three were in the stratosphere (5, 10, and 50 hPa) and the other three were in the troposphere (200, 500, and 925 hPa). The full meridional and zonal structure of the difference in geopotential height between HP and LP during the extended NH winter period (October to March) was studied. The composite differences were taken for a sequence of overlapping 2 month intervals. This allows us to investigate the evolution of the responses, especially the temporal consistency with the $P_{sw}$ signal in the NAM.

Figure 5. Same as the right-hand panels of Figure 3, but the NAM is replaced by zonal mean temperature averaged for 10°N (left) and 80°N (right). Contours are at ±0.5, ±1, ±1.5, ... K nPa$^{-1}$ for the left-hand panels and ±2, ±4, ±6, ... K nPa$^{-1}$ for the right-hand panels.

($F_{10.7}$, [10$^{-22}$ W m$^{-2}$ Hz$^{-1}$]) from the National Geophysical Data Center (NGDC, http://ngdc.noaa.gov) were used to estimate the 11 year solar irradiance cycle. We define $F_{10.7} < -0.25$ and $F_{10.7} > 0.25$ to be low and high solar irradiance or conventionally called high and low solar activity, where $F_{10.7}$ stands for the median-normalized values of $F_{10.7}$; transition periods where $|F_{10.7}| \leq 0.25$ were excluded from the analysis. Qualitatively similar results can be obtained if other definitions, such as mean instead of median, are used. Hereinafter, we call the regression in which no subgrouping was applied the $Alldata$ case, while the regression applied to only low or high $F_{10.7}$ years are abbreviated as $HS$ or $LS$ case, respectively. The running regressions between various atmospheric variables and $P_{sw}$ over the period of 1965–2009 are performed using a 31 day running window with a daily time step. The results are displayed in the form of time-pressure height cross-sections for winter and spring (1 November to 31 May, $x$ axis), from 1000 hPa to 1 hPa ($y$ axis), and for $Alldata$, $HS$, and $LS$ conditions (three separate panels from top to bottom).

As an additional diagnostic, a composite analysis of the vertical connected $P_{sw}$ signals, the composite analysis was done at six different pressure levels, of which three were in the stratosphere (5, 10, and 50 hPa) and the other three were in the troposphere (200, 500, and 925 hPa). The full meridional and zonal structure of the difference in geopotential height between HP and LP during the extended NH winter period (October to March) was studied. The composite differences were taken for a sequence of overlapping 2 month intervals. This allows us to investigate the evolution of the responses, especially the temporal consistency with the $P_{sw}$ signal in the NAM.

[22] The intrinsic timescale of NAM variability in the stratosphere is of the order of several weeks, reflecting the relatively slow interactions between the zonal flow and planetary-scale waves. In the troposphere, the variability of the NAM is of the order of about 2 weeks and is largely driven by interactions between the zonal flow and rapidly evolving baroclinic waves. As Rossby wave breaking and their induced momentum fluxes are the major drivers of the tropospheric NAM, additional diagnostics are needed to examine how the solar wind dynamic pressure might affect them. In this study, the RWB events are identified using the 1° × 1° horizontal resolution ERA-40 and ERA-Interim data sets on the 310–360 K isentropic level with 5 K intervals. Our identification follows the method of Wernli and Sprenger [2007] that identifies the presence or absence of a RWB event for every 6 h time interval on various tropopause intersecting isentropes. The number of RWB events is then aggregated vertically and then stratified into anticyclonic (LC1) and cyclonic (LC2) types [Martius et al., 2007]. The difference between the two types of RWB, LC1 and LC2 (or anticyclonic and cyclonic) is that LC1 is associated with an equatorward momentum flux, whereas LC2 is associated with a poleward momentum flux [Thorncroft et al., 1993]. Composite analysis (between HP
and LP conditions) is then performed to study if there is any significant difference in RWB occurrence frequency over the winter months.

3. Results

3.1. The \( P_{sw} \) Signal in the NAM, \( \bar{u} \), and \( T \)

[25] Figure 3 shows the running regression between the daily NAM and \( P_{sw} \) (left-hand panels), and the 31 day averaged NAM and \( P_{sw} \) (right-hand panels). In both daily and 31 day averaged NAM cases, very similar NAM anomalies are obtained except that the daily case shows a larger magnitude of response and variation. Under Alldata conditions, a positive \( P_{sw} \) signal in the NAM is detected in the troposphere from late February to March and a negative signal in the stratosphere from mid-March to May, while no significant NAM anomalies are found during winter months in the stratosphere. Under HS conditions, however, a positive \( P_{sw} \) signal in the NAM is detected. It propagates from the stratosphere to the troposphere in December–March, followed by a weaker negative signal in February–April. In March–May, a significant negative \( P_{sw} \) signal in the NAM is detected under LS conditions, where the signal propagates from the upper to the lower stratosphere. LS years appear to dominate the overall \( P_{sw} \) signal in the NAM. As a whole, the \( P_{sw} \) signal in the NAM suggests that a top-down mechanism is most likely at play both under HS and LS conditions. The main difference between the signal under HS and LS conditions appears to be the timing and the strength of the responses.

[24] Figure 4 shows the running regression between \( P_{sw} \) and \( \bar{u} \) at 35°N (left) and 55°N (right). A similar response pattern can be obtained at the latitude range of 25°N–45°N and 50°N–70°N, respectively. A \( P_{sw} \) signal in \( \bar{u}_{35N} \) occurs in January–March only under HS conditions, and is marked by a downward propagation of easterly anomalies from 1 hPa to the surface. The descent of the easterly anomalies appears to occur nearly instantaneously at 1–5 hPa, and then the descent becomes slower in the lower stratosphere and troposphere. Under LS conditions, alternating westerly and easterly anomalies are shown, which are significant at the 0.1 level. But they are short lived, and there is no clear descent from the upper stratosphere into the lower atmosphere. There is also a 45–60 day time delay when they are compared to the HS case. In contrast, downward propagation of the \( P_{sw} \) signal in \( \bar{u}_{55N} \) occurs under both HS and LS conditions. Those high-latitude signals are characterized by oscillating westerly anomalies followed by easterly anomalies. The signal descends from the upper to the lower stratosphere over a period of ~45 days. While the magnitude of the response under HS and LS conditions is similar, the main difference between the signal under HS and LS conditions is that the response occurs earlier (starting in December) under HS conditions and later (starting in February) under LS conditions. The delayed response under LS conditions results in no significant response in winter under Alldata conditions. In spring, easterly \( \bar{u}_{35N} \) anomalies are detected for Alldata, HS, and LS cases, implying a weaker springtime polar vortex is in general associated with high \( P_{sw} \), consistent with previous studies [Lu et al., 2008b; Kvissel et al., 2012; Seppälä et al., 2012]. As a whole, these \( P_{sw} \) signals in \( \bar{u} \) are also dynamically consistent with the NAM responses (see Figure 3).

[25] If \( P_{sw} \) induced circulation anomalies do obey the dynamic reasoning shown in Figure 1, then opposite temperature anomalies must be detected at the low and high-latitude upper stratosphere. For this purpose, Figure 5 shows the running regression between \( P_{sw} \) and \( T \) at 10°N (left) and 80°N (right). Note that very similar response patterns can be obtained for the latitude belts of 5°N–15°N and 70°N–90°N, respectively. Under HS conditions, the temperature response in the low-latitude stratosphere is characterized by a warming signal in November–December followed by a cooling signal in the upper stratosphere in January–February. The signal in the polar latitudes is marked by an opposite response with a cooling signal followed by a warming signal descending from the upper to the lower stratosphere over a period of a month or so. Under LS conditions, the temperature response in the low-latitude stratosphere is mostly confined to the upper stratosphere, oscillating from cooling to warming anomalies in winter months. This is dynamically consistent with the zonal wind responses at 25°N–45°N (see Figure 4). Weaker but similar responses can also be found in spring and the middle stratosphere. At high latitudes, nearly the same oscillation pattern to that under HS conditions is found except for a time delay of ~45 days. Similar to the signal in the NAM, the temperature response is only statistically significant above the 95% confidence level in spring under Alldata conditions and is dominated by LS years. These results suggest that \( P_{sw} \) may have a significant effect on the magnitude as well as the timing of stratospheric final warmings.

Figure 6. Same as the right-hand panels of Figure 3, but the NAM is replaced by zonal mean temperature averaged at 40°N. Contours are at ±0.5, ±1, ±1.5, . . . K nPa⁻¹.
Apart from the temperature at low and high latitudes, we also find that the temperature response at middle latitude behaves quite differently under HS and LS conditions. Figure 6 shows the running regression between $P_{sw}$ and $T$ at 40°N for Alldata, HS, and LS conditions. Under HS conditions, a warming signal in January–March descends from 10 hPa to the surface with a cooling signal being present in the upper stratosphere (1–5 hPa) in January–February. The warming signal in the stratosphere is much weaker and insignificant under Alldata and LS conditions. In March, the tropospheric signal is also of opposite sign under HS and LS conditions. They cancel each other and consequently result

Figure 7. Composite differences (high minus low $P_{sw}$, i.e., $HP - LP$) of December–January mean geopotential height (in meters) at 10 and 50 hPa (top to bottom) for the period of 1965–2009 under Alldata, HS, and LS conditions (left to right). The area shaded by white back-slash-dotted lines indicates significance at the $p = 0.05$ level. The latitude extent is 10°N–90°N.

Figure 8. Same as Figure 7, except for January–March mean geopotential height at 200 hPa (top) and 500 hPa (bottom).
rather than a near annular pattern, with negative latitude response is marked by a wave number 1 pattern. These levels were selected as they represent the strongest (middle), and ZHP-LP number 1 pattern with a negative response pattern is also found under the Pacific and more zonally varying when they move from the upper stratosphere to the lower troposphere. It is worth noting that the January–March mean at 50 hPa, an annular mode pattern signature can be observed at 200 hPa and 500 hPa only under HS conditions. At both pressure levels, the regions with statistical significance of \( p = 0.05 \) are confined to the Atlantic sector. The \( P_{sw} \) signals in tropospheric geopotential height \( Z \) are also consistent with those in the NAM (see Figure 1), both in sign and timing. Under HS conditions, the response is predominantly annular with positive \( Z_{HP-LP} \) at high latitudes and negative \( Z_{HP-LP} \) at midlatitudes, consistent with a positive NAM and a stronger, less disturbed polar vortex. Under LS conditions, no significant \( Z_{HP-LP} \) can be found. Those signals in the stratospheric geopotential height are consistent with that in the NAM (see Figure 1), both in sign and timing of the signals.

### 3.2. The \( P_{sw} \) Signal in Geopotential Height \( Z \)

[27] Figure 7 shows the composite difference in the December–January mean NH geopotential height between HP and LP conditions (\( Z_{HP-LP} \), for Alldata (left), HS (middle), and LS (right), at 10 hPa (top) and 50 hPa (bottom). These levels were selected as they represent the strongest differences in \( Z \). At 10 hPa, the height response \( Z_{HP-LP} \) is most robust under Alldata conditions, where the signal is characterized by positive height anomalies at middle latitudes and negative anomalies at polar latitudes. The strongest significant signals are the positive \( Z_{HP-LP} \) over northwest Europe and the northwest Pacific. A similar response pattern is also found under HS conditions except that the difference is statistically significant only over the northwest Pacific, and the negative difference in the high latitudes are more zonal and widespread than under Alldata conditions. Under LS conditions, however, the high-latitude response is marked by a wave number 1 pattern rather than a near annular pattern, with negative \( Z_{HP-LP} \) over the Pacific and northern America and positive \( Z_{HP-LP} \) over the Pacific and northern Europe.

[28] At 50 hPa and under Alldata conditions, the lower stratospheric height response \( Z_{HP-LP} \) resembles a near wave number 1 pattern with a negative \( Z_{HP-LP} \) anomaly over northern Canada and a positive \( Z_{HP-LP} \) anomaly over Europe.

![Figure 9](image-url)

**Figure 9.** Same as the right-hand panels of Figure 3 but with the NAM replaced by horizontal EP-flux divergence \( \Psi^{(0)} \) at 35°N–45°N (left) and at 55°N–75°N (right). Contours are at \( \pm 0.5, \pm 1, \pm 2, \pm 3 \), \( \ldots \times 10^{-5} \) m s\(^{-2}\) Pa\(^{-1}\).
to descend into the troposphere. Such a signal is dynamically consistent with a flow deceleration at ~35°N and acceleration at ~55°N (see Figure 4). The \( P_{sw} \) signal in \( \Psi^{(0)} \) under LS conditions is confined to the upper stratosphere with negative anomalies in March and positive anomalies from late April to May. It is clear no downward descent of \( \Psi^{(0)} \) beyond 10 hPa can be observed under LS conditions.

[31] Figure 10 shows the running regression of the vertical EP-flux convergence term \( \Psi^{(z)} \), averaged over the latitude band 35°N–75°N. From November to May, the response in \( \Psi^{(z)} \) from 35–75°N to \( P_{sw} \) is marked by oscillating positive, negative, and then positive anomalies which descend from the upper stratosphere to the lower stratosphere over a period of ~10–15 days. Compared with the temperature response (see Figure 5), positive \( \Psi^{(z)}_{35–75°N} \) anomalies (i.e., planetary wave divergence) are associated with anomalous cooling at high latitudes and warming at low latitudes, while the opposite holds for negative \( \Psi^{(z)}_{35–75°N} \). In late winter and spring, the stronger wave driving (i.e., the negative \( \Psi^{(z)}_{35–75°N} \) anomalies) is consistent with a weaker polar vortex and adiabatic warming over the pole (see Figures 4 and 5). Thus, Figure 10 suggests that the response of vertically propagating waves at 35°N–75°N is dynamically consistent with the temperature response in the upper to middle stratosphere. As with the temperature responses, the main difference of the \( \Psi^{(z)}_{35–75°N} \) responses under HS and LS conditions is the timing. That is, the descent of the responses occurs 30–45 days earlier under HS conditions than under LS conditions.

[32] We found that the response patterns of \( \Psi^{(z)} \) shown in Figure 10 are very robust for Alldata and LS conditions; they hold true for any latitude band from 35°N to 75°N though the signal is the strongest around 55°N, where the zonal wind is the strongest. For HS conditions, the descent of \( \Psi^{(z)} \) signal occurs primarily at 55°N poleward. Figure 11 shows the running regression of \( \Psi^{(z)} \), averaged over 35°N–45°N under HS conditions. In December–March, the signal is marked by weak but statistically significant positive \( \Psi^{(z)} \) anomalies at 5–100 hPa and large negative \( \Psi^{(z)} \) anomalies above. Unlike the high-latitude response that tends to descend from 1 hPa to 30 hPa (see Figure 10), no clear descent is observed for the \( \Psi^{(z)}_{35–45°N} \) anomalies. A close examination of a series of 10° latitude bands suggests that as the positive anomalies in January–March get weaker with an increase of latitude, the negative anomalies get stronger and descend. These signals suggest that, when solar UV is high, the solar wind dynamic

---

**Figure 10.** Same as the right-hand panels of Figure 3 but with the NAM replaced by the vertical EP-flux divergence term \( \Psi^{(z)} \) at 35°N–75°N. The contours are ±0.5, ±1, ±2, ±4, ±6, \( \cdots \times 10^{-7} \) m s\(^{-2} \) Pa\(^{-1} \). Note the vertical extent (y axis) is 200–1 hPa because the regression is too noisy below 200 hPa.

**Figure 11.** Same as the middle panel of Figure 10 but for \( \Psi^{(z)} \) at 35°N–45°N under HS conditions.

---

3.3. The \( P_{sw} \) Signal in EP flux

[30] The downward propagation of zonal mean zonal wind anomalies in the stratosphere is known to be closely associated with anomalies in stratospheric wave drag caused by the vertical extent of the propagation of planetary-scale waves [Andrews et al., 1987]. To examine how the \( P_{sw} \) signals in the NAM, \( \bar{u} \), \( T \), and \( Z \) are linked to wave-induced momentum transfer, we show in Figure 9 the running regression of the horizontal EP-flux convergence term \( \Psi^{(0)} \), averaged over the latitude bands of 20°N–40°N (left) and 50°N–70°N (right). The \( P_{sw} \) signal in \( \Psi^{(0)}_{20–40°N} \) is generally marked by negative \( \Psi^{(0)} \) anomalies in winter and positive \( \Psi^{(0)} \) anomalies in spring, as seen in the Alldata case. A nearly opposite response is found in \( \Psi^{(0)}_{50–70°N} \). The signals are most significant in the upper stratosphere. This general response pattern is strongly modulated by the 11 year solar irradiance cycle. Under HS conditions, the \( P_{sw} \) signal in \( \Psi^{(0)} \) occurs earlier and appears...
The pressure effect on the eddy heat flux at middle latitudes is marked by wave divergence in the midstratosphere and wave convergence in the upper stratosphere. This is dynamically consistent with the temperature response shown in Figure 5e, where high-latitude cooling in the midstratosphere is followed by a warming signal above. Another feature worth noting is the negative response at 100–200 hPa in November–January, which intensifies and extends into later months with an increase of latitude. It indicates that solar wind dynamic pressure may have a detectable effect on synoptic-scale Rossby wave breaking, and the effect may actually start in early winter. A poleward movement of the negative Ψ anomalies in middle to later winter implies that synoptic-scale Rossby wave breaking intensifies, and the effect moves poleward as the winter progresses. It is noticeable that the $P_{sw}$ signal in $Ψ_{35–45N}$ leads that in temperature by ~15 days in the upper stratosphere, suggesting that the wind and temperature response shown in Figures 4b, 4e, 5b, and 5e is of a dynamic origin and it is generated in the upper stratosphere or above.

Figure 12 shows the $P_{sw}$ signals in total EP-flux term $Ψ = (ρ d a \cos Φ)^{-1} \nabla \cdot F$ averaged at 35°N–45°N and 55°N–75°N. At 35°N–45°N, the signal is characterized by anomalous wave convergence at 1–10 hPa from mid-December to March. Such $Ψ$ anomalies are noticeably stronger and extend lower into the middle and lower stratosphere under $HS$ conditions than under $LS$ conditions. At 55°N–75°N, the $P_{sw}$ signal in $Ψ$ is marked by wave divergence at 1–20 hPa from December to February and followed by wave convergence at 1–100 hPa in March–April. The high-latitude wave response can be seen

![Figure 12](image12.png)

Figure 13. Same as the right-hand panels of Figure 3 but with the NAM replaced by the meridional circulation term $f \tilde{v}$ at 60°N. Contours are at ±0.5, ±1, ±2, ±4, ±6, ... × 10^{-5} m s^{-2} nPa^{-1}.
more clearly under LS conditions than under HS conditions. Overall, those $\Psi$ anomalies suggest that more planetary waves are refracted equatorward when $P_{sw}$ is higher than its average during NH winter. Under HS conditions, the wave refraction is also accompanied by more planetary wave breaking in the upper stratosphere at both high and low latitudes. Compared with Figures 9–11, it is clear that the vertical EP flux determines the $P_{sw}$ signal in $\Psi$ under Alldata and LS conditions, while the horizontal EP flux plays an additional role in determining the $P_{sw}$ signal under HS conditions.

Figure 13 shows the $P_{sw}$ signals in the meridional residual circulation term $f\nu^*/C_2$ at 60°N for Alldata, HS, and LS conditions (top to bottom). This is to investigate the effects on the Brewer-Dobson (BD) circulation. In the Alldata case, the meridional circulation is anomalously strengthened in March followed by weakened meridional circulation in April–May. Under HS conditions, the meridional circulation weakens at 30–50 hPa in December–January and strengthens above 10 hPa and below 70 hPa in December–March, followed by a weakening of the meridional circulation in April–May. Under LS conditions, there is a weakening of meridional circulation during January–February, followed by a strengthening of the circulation in March–April. Also, Figure 13 suggests that the $P_{sw}$ effect on the BD circulation is weak during winter in Alldata case and under LS conditions, and if there is any, the effect is mainly on its upper branch rather than on its lower branch. Under HS conditions, the $P_{sw}$ effect on the BD circulation during NH winter is observed both in the upper stratosphere and in the lower stratosphere, implying that changes of large-scale planetary waves as well as synoptic-scale Rossby waves are involved.

3.4. The $P_{sw}$ Effect on Synoptic-Scale Rossby Wave Breaking

[35] In order to investigate how the solar wind dynamic pressure may ultimately affect the tropopause-level synoptic structure, we examined the change of frequency in synoptic-scale Rossby wave breaking near the tropopause. The upper panel of Figure 14 shows that when $P_{sw}$ is high (HP), anticyclonic wave breaking (LC1) occurs more frequently (up to 10%) at the latitude band of 20°N–40°N, while cyclonic wave breaking (LC2) occurs less frequently (up to 6%) at 40°N–60°N during January–March. These RWB anomalies are not evenly distributed zonally; the most significant signatures are primarily confined to the Atlantic region. For the same region, very similar RWB anomaly patterns can be found for HS conditions (the lower panel of Figure 14). The enhancement of LC1 type of RWB becomes more zonal and extends deeper into Eurasia under HS conditions. Under LS conditions, there are no significant differences in RWB between HP and LP conditions (not shown). These responses are in good agreement with the $\Psi(\theta)$ anomalies shown in Figure 9 and are consistent with the
known link between momentum flux and the tropospheric NAM [Limpasuwann and Hartmann, 2000; Robinson, 2006]. It implies that a higher $P_{sw}$ is associated with an increase of LC1-type RWB and a decrease of LC2-type RWB, which leads to a positive NAM/NAO and a poleward shifted jet stream over the North Atlantic region [Franzke et al., 2004; Feldstein and Franzke, 2006]. The RWB anomalies in Figure 14 are also consistent with Figure 4. An increase of the vertical wind shear in the lowermost stratosphere through a strengthening of the polar vortex can change the wave-breaking characteristics and consequently alter the momentum fluxes in such a way as to drive the eddy-driven midlatitude jet poleward to high latitudes [Robinson, 2006; Chen and Held, 2007]. In midwinter, the momentum fluxes associated with the breaking waves lead to a northward shift of the jet. The ensuing adjustment processes lead to a stronger meridional temperature gradient in the area of the jet stream and a decrease of the temperature gradient in the latitudinal band where the jet stream used to be.

[36] It is known that a positive feedback exists between this type of synoptic-scale Rossby wave breaking (LC1 and LC2) and the associated momentum fluxes and the latitudinal location of the eddy-driven jet [Franzke et al., 2004]. This feedback mechanism could serve as an amplification mechanism for the initially small $P_{sw}$ effect in the stratosphere and as a mechanism for the dynamical transfer into the troposphere.

[37] It is possible that the $P_{sw}$ signals discussed above may have been contaminated by other factors such as the major volcanic eruptions and El Niño Southern Oscillation (ENSO). The same analysis was carried out by excluding the years affected by major volcanic eruptions and/or by major ENSO events (see Table 1). We found that excluding the years affected by major volcanic eruptions enhances the signal in terms of statistical significances. Excluding the years affected by major ENSO does not alter the response in magnitude and the spatial patterns but results in slightly lower statistical significances. Thus, it is unlikely that the $P_{sw}$ signals we detected in the ERA-reanalysis data have been caused by either the major volcanic eruptions or the ENSO. This is consistent with previous findings of Lu and Jarvis [2011] and Seppälä et al. [2012]. Furthermore, the wave-breaking characteristics we found are also distinct from those associated with the major teleconnection patterns, such as the NAO in terms of the magnitude of frequency distributions and PNA (Pacific/North American Pattern) in terms of the location of the wave-breaking anomalies [Martius et al., 2007].

4. Discussion

[38] To some extent, our results are comparable to a recent modeling study by Baumgaertner et al. [2011], who showed that the zonal mean temperature response to geomagnetic activity induced by EPP-N0X in NH winter is marked by anomalous warming in the upper stratosphere and mesosphere and is accompanied by anomalous cooling in the middle to lower stratosphere. Here we also find that the cooling effect extends into the troposphere during late winter under HS conditions. However, it compares less well with the modeling of Semeniuk et al. [2011], which detects significant temperature responses at midlatitudes. In our analysis, the signal descends from the stratosphere to the troposphere with alternating positive and negative signals with a period of ~45 days. The timing of the signal appears to be modulated by the 11 year solar irradiance cycle. Without grouping the data according to the level of solar UV, it can be expected that a range of response patterns can be obtained if the analysis is carried out for each calendar month. The oscillating nature of the response can lead to signal cancelation if the regression or composite analysis is carried out for the entire winter and spring. Here we show that interannual variations of the wintertime NAM in response to changes of solar wind dynamic pressure are closely linked to a strengthened mean polar vortex, highlighting the important role of stratospheric dynamics. In the troposphere, we find that the zonal mean zonal wind response to the solar wind dynamic pressure is similar to the zonal wind anomalies of Fig. 4d of Haigh et al. [2005], who performed a sensitivity study using a simple mechanistic model with an increased equator-to-pole temperature gradient imposed to the model in the stratosphere.

[39] Here, we propose two possible mechanisms to explain the observed stronger solar wind dynamic pressure signature under HS conditions. The first possible mechanism is a change of the meridional structure of the flow. Higher solar UV input induces anomalous warming in the low-latitude stratosphere. This leads to a greater temperature gradient from equator to midlatitudes and therefore stronger westerly winds in the subtropical upper stratosphere. A modeling study by Gerber [2012] shows that a stronger vortex creates a waveguide deeper into the stratosphere, allowing waves to propagate to higher altitudes. This raises the breaking level of planetary waves and deepens the Brewer-Dobson circulation vertically. The net effect on the extratropical circulation is an anomalous cooling in the lower stratosphere and warming above. The strengthening of the polar vortex combined with a moderate increase in planetary waves induced by the enhanced $P_{sw}$ results in anomalously more planetary waves being refracted equatorward; therefore, more momentum flux is transferred poleward [Dickinson, 1968; Kodera and Kuroda, 2002]. Such feedback processes between the planetary waves and polar vortex further induce a stronger temperature gradient from the middle latitude to the pole and a positive NAM in the stratosphere. Nevertheless, further investigations are needed to understand how and why the enhancement of $P_{sw}$ induces more planetary wave breaking in the upper stratosphere.

[40] The second possible mechanism is the flow asymmetry induced by solar UV and stratospheric flow interaction at low latitudes. Under HS conditions, solar UV and ozone interaction in the low-latitude stratosphere induces an additional asymmetry of the flow [Albers and Nathan, 2012]. Using a general circulation and chemistry model, Gabriel et al. [2011] found that, for northern winter, the 11 year solar cycle signal in middle atmospheric temperature, ozone, and water vapor is characterized by a pronounced wave number 1 pattern and a significant increase in amplitude and a horizontal phase shift in terms of the wave-1 pattern from solar minimum to maximum. As such, those authors suggested that the solar UV variability induces these changes by modulating zonally asymmetric radiative heating and subsequently modulating planetary wave propagation. Hirota [1967] showed that longitudinally varying flows may be barotropically unstable, even
when their zonal mean is stable. A small asymmetry to the flow would tend to increase the growth rate and the period of the unstable disturbances [Hartmann, 1983]. We speculate that an enhancement of solar UV leads to a zonally asymmetric radiative heating induced by asymmetric ozone in the low-latitude stratosphere. This leads to a localized negative potential vorticity gradient in northern winter, which consequently enhances barotropic instability of the polar vortex. This leads to a planetary-scale wave growing on the equatorward side of the polar vortex.

[41] A widely held perception of the upper stratosphere is that it is the region whose large-scale dynamics are dominated by slowly varying planetary waves propagating upward from below. Dissipating planetary waves are responsible for large negative values of EP-flux divergence $\nabla \cdot F$. However, a decrease of wave activity in the lower stratosphere with an increase of wave activity in the upper stratosphere may play an important role for barotropic energy exchange between the mean zonal flow and planetary waves. It was demonstrated by Hartmann [1983] that the unstable mode of the planetary waves with a period of a few weeks may grow on the equatorward side of the polar vortex near the region with localized negative potential vorticity gradients near the stratosphere. These unstable waves are able to derive momentum from a region with reversed potential vorticity gradient and then feed energy to the higher latitude flows, in a similar way as is shown in our Figure 9. The energy generation of the unstable planetary waves arises from an equatorward momentum flux near the edge of the region of a reversed potential vorticity gradient. The periods of those unstable modes are of the order of 1–3 weeks which is very close to the time scales of the variations of the wave forcing from the troposphere and the horizontal structure of the unstable modes is also similar to that of upward propagating waves. Cnossen and Lu [2011] showed that the 11 year solar UV signature in the early winter stratosphere is mainly confined to the upper stratosphere, and the atmospheric response is characterized by a poleward advection of relatively warmer low-latitude air, which would induce a localized reversal of potential vorticity which in turn may enhance the barotropic instability seen in the HS case (Figure 9). In late winter or spring, a more upright polar vortex interacts with the strengthened planetary waves propagating upward from the troposphere, and barotropic instability is less likely to occur. Thus, the increase of wave drag in the stratosphere leads to an EP-flux convergence and a weaker and warmer than normal polar vortex. It is possible that those unstable modes would hardly rise above the noise level, and it is therefore difficult to isolate the effect caused by the region of a reversed potential vorticity gradient for individual winters. Nevertheless, its effect may be statistically detectable over an extended time period when internal variability and noise are likely to be averaged out. On the basis of our analysis and the aforementioned studies, we thus speculate that the interaction between waves propagating upward from the troposphere and the instability of the stratospheric polar vortex is the mechanism through which we observe the $P_{sw}$ signal.

[42] Under HS conditions at $35^\circ$–$45^\circ$ N, the $P_{sw}$ signal in the vertical EP-flux divergence $\Psi^{(z)}$ is marked by a negative $\Psi^{(z)}$ anomaly extending downward from the upper stratosphere to the tropopause region (see Figure 10). This vertical structure with the temperature response at 40$^\circ$N (cooling in the upper stratosphere and warming at 10–100 hPa; see Figure 6) indicates that instability of planetary waves might be involved in the $P_{sw}$ modulation of $\Psi^{(z)}$. The upward propagation of planetary waves from the troposphere to the stratosphere is known to be sensitive to the vertical shear of the zonal wind near the tropopause region. Planetary wave propagation into the stratosphere becomes weaker than average with an increased vertical wind shear and vice versa [Chen and Robinson, 1992]. In the upper stratosphere, however, wave breaking increases with a colder vortex, because potential vorticity gradients along the edge of the vortex create a waveguide higher into the stratosphere [Gerber, 2012]. This effect becomes particularly noticeable in the presence of topography and with an increased meridional temperature gradient. Results shown in Figure 10 are consistent with this theoretical explanation. In essence, a colder vortex creates a waveguide higher in the stratosphere, raising the breaking level of planetary waves, causing more wave breaking in the upper stratosphere and less wave breaking in the lower stratosphere.

[43] Figure 11 also indicates that a reduction of baroclinic instability at ~10 hPa and an increase of baroclinic instability near the tropopause are associated with a higher $P_{sw}$ The positive NAM might be linked to the baroclinic instability associated with the polar vortex [Tanaka and Tokinaga, 2002]. The baroclinic instability induces a strengthening of westerly anomalies at high latitudes by extracting momentum from lower latitudes. This results in an increase of eastward shear with height and a reduction of eddy drag toward the pole that is accompanied by an increased horizontal temperature gradient in the region. The increase of baroclinic instability near the tropopause may therefore play a key role in bringing stratospheric changes into the troposphere. Further research is needed to examine whether or not $P_{sw}$ modulates the unstable planetary wave modes and which process is involved to increase the baroclinic instability near the tropopause. Research of this kind would help to understand what causes the significant weakening of the zonal wind at $35^\circ$N and the positive temperature anomalies at $40^\circ$N under HS conditions.

[44] In the troposphere, the zonal mean-flux change in relation to $P_{sw}$ variations mainly occurs in late winter (January–March) and under HS conditions. As the winter progresses, the stratospheric anomalies propagate downward and induce a stronger vertical wind shear above the tropopause [Limpasuvan and Hartmann, 1999; 2000; Kunz et al., 2009a]. Through a positive feedback between zonal wind anomalies and synoptic-scale Rossby waves, the associated baroclinic instability draws the energy from the available potential energy associated with horizontal temperature gradients by extracting momentum from lower latitudes and feeding it into higher latitudes. As a result, the stratospheric changes are brought further downward into the troposphere [Simpson et al., 2009]. The strengthening of the lower stratospheric wind increases the eastward phase speed of the Rossby waves, leading to increased subtropical wave breaking and poleward shift in the eddy momentum flux convergence [Chen and Held, 2007]. This process also results in more anticyclonic synoptic-scale RWB at lower latitude and less cyclonic RWB at high latitude. This is consistent with a poleward momentum flux convergence and therefore a poleward shift of the eddy-driven jet in the troposphere [Franzke et al., 2004; Martius et al., 2007; Kunz et al.,...
2009a]. The strengthening of the lower stratospheric winds reinforces the barotropic anomaly in the zonal wind structure and the associated anticyclonic shear, which further amplifies the equatorward refraction of eddy activity [Simpson et al., 2009; Barnes and Hartmann, 2010].

5. Conclusions

[45] We studied the possible influence of the solar wind dynamic pressure ($P_{sw}$) on the stratospheric and tropospheric circulations in NH winter and spring in the blended ERA-40 and ERA-Interim data. Our main results are the following:

[46] 1. The vertical structure of the zonal mean circulation and EP-flux anomalies shows a dynamically consistent pattern of downward propagation, implying that stratospheric anomalies affect the tropospheric circulation; thus, a top-down mechanism plays a key role. As the winter progresses, these anomalies move downward steadily over a period of ~45 days. This applies to both HS and LS conditions, except that there is a time delay under LS conditions. This delay is likely due to the reduced solar UV in the upper stratosphere leading to a reduced equator-to-pole temperature gradient. As a result, the winter signal is stronger under HS conditions, while the spring signal is stronger under LS conditions. The overall signal is dominated by that associated with LS conditions.

[47] 2. In good agreement with previous studies [Lu et al., 2008a, 2008b; Seppälä et al., 2012], the response of the zonal mean zonal wind is characterized by a strengthening of the westerly winds at ~55°N and weakening of the winds at ~35°N. The response of the zonal mean temperature is marked by cooling in the high-latitude lower stratosphere and warming near the stratosphereopause. The weakening of the winds at ~35°N is most significant under HS conditions.

[48] 3. At 10 hPa, the geopotential height response to $P_{sw}$ is characterized by positive anomalies at middle latitudes and negative anomalies at polar latitudes. The strongest significant signals are over northwest Europe and the northwest Pacific. Under HS conditions, the response of the winter geopotential height shows a nearly equivalent barotropic structure that extends from the stratosphere to the troposphere. In the stratosphere, the response is marked by positive heights at 20°–40°N and negative heights at high latitudes, consistent with a positive NAM. In the troposphere, a similar pattern is revealed, but the midlatitude anomalies are mostly confined to the two ocean basins where the jet streams dominate. The dipole pattern in the geopotential height suggests that baroclinic instability associated with a strong vortex in the middle to lower stratosphere is likely to be the responsible mechanism for the downward propagation of the winter $P_{sw}$ signals. No noticeable stratosphere-troposphere coupling is detected for LS conditions under which the dominant wave forcing is mainly controlled by an enhancement of the upward eddy heating flux in the stratosphere in spring.

[49] 4. These mean-flow responses are dynamically consistent with the anomalous EP-flux divergence in the lower stratosphere and convergence in the upper stratosphere, suggesting that planetary waves play a key role in the solar wind dynamic pressure signature that we observe in the blended ERA-40 and ERA-Interim data.

[50] 5. It appears that the $P_{sw}$ signal depends also on the strength of the polar vortex or the meridional temperature gradient in the early winter stratosphere as suggested by the signal under HS conditions. A strong polar vortex helps to increase equatorward wave refraction. This is consistent with the recent findings of Seppälä et al. [2012].

[51] 6. There is a significant equatorward shift of synoptic-scale Rossby wave breaking near the tropopause during January–March which corresponds to an increased frequency of anticyclonic RWB at 20°–40°N and a decrease in cyclonic RWB at 40°–65°N. The wave-breaking anomalies are mainly found in winters when the solar irradiance flux is higher than average, and the significant effect is confined to the Atlantic region where the eddy-driven jet exists. The enhancement of anticyclonic RWB anomalies extends downstream into Eurasia under HS conditions, agreeing with the earlier findings of the tropospheric circulation anomalies and a reduction of blocking associated with the open solar flux [Woollings et al., 2010]. The pattern of the wave breaking suggests that a persistent poleward shift of the jet in January–March is associated with a positive NAM.

Acknowledgments. H.L. and M.J.J. were supported by grant NE/I010173/1 of the UK Natural Environment Research Council (NERC). We acknowledge use of ECMWF reanalysis data sets and NASA/GSFC’s Space Physics Data Facility’s OMNIWeb service and OMNI data. We are grateful to three anonymous reviewers for their constructive comments.

References


