Role of wave–mean flow interaction in sun–climate connections: Historical overview and some new interpretations and results

Terrence R. Nathan \textsuperscript{a,}\textsuperscript{*}, John R. Albers \textsuperscript{a}, Eugene C. Cordero \textsuperscript{b}

\textsuperscript{a} Atmospheric Science Program, Department of Land, Air and Water Resources, University of California, Davis, CA 95616, USA

\textsuperscript{b} Department of Meteorology and Climate Science, San Jose State University, San Jose, CA 95192, USA

**Abstract**

Quasi-decadal variations in solar irradiance – termed the 11-year solar cycle (SC) – have been linked to variations in a variety of atmospheric circulation features, including the polar vortex, the Brewer–Dobson circulation, and the quasi-biennial oscillation. These features share an underlying commonality: they are all rooted in wave–mean flow interaction. The purpose of this paper is to provide a historical overview of the connection between the SC and wave–mean flow interaction and to propose a more complete theoretical framework for solar modulated wave–mean flow interaction that includes both zonal-mean and zonally asymmetric ozone as intermediaries for communicating variations in solar spectral irradiance to the climate system. We solve a quasi-geostrophic model using the WKB formalism to highlight the physics connecting the SC to planetary wave-drag. Numerical results show the importance of the zonally asymmetric ozone field in mediating the effects of solar variability to the wave-driven circulation in the middle atmosphere.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Since the 11-year solar sunspot cycle was identified by Heinrich Schwabe (1843) more than one hundred and fifty years ago, interest in the 11-year solar cycle (SC) as a purveyor of atmospheric variability has seemed to wax and wane with the SC itself. Early studies relating the SC to weather and climate variability were often fraught with difficulties. Among the difficulties was the sparse data record, which was often too short to yield reliable statistics. In some cases the correlations were significant only at specific locations (Barnston and Livezey, 1989), while in other cases what appeared to be significant correlations were either mitigated or vanished as additional data became available or different statistical analyses were used (Laut, 2003). In those studies where the statistical analyses appeared sound, the results were often met with skepticism owing to an absence of a plausible physical mechanism to explain the results.

A resurgence in sun–climate research associated with the SC essentially began about 20-years ago. At that time, Labitzke and van Loon (1988) showed that when the winter northern polar stratospheric temperatures are stratified according to the phase of the equatorial quasi-biennial oscillation (QBO), a quasi-decadal variation is observed that is consistent with the SC. Because the QBO exerts its influence far beyond its tropical seat of origin (Baldwin et al., 2001), affecting such atmospheric circulation features as the Northern Annular mode (Coughlin and Tung, 2001), Northern Hemisphere polar vortex (Lu et al., 2008), and the timing of stratospheric sudden warmings (Gray et al., 2004), it was believed that the SC modulation of the QBO might serve as a pathway for communicating solar variations to other atmospheric circulation features. Although initially viewed with caution, Labitzke and van Loon’s SC–QBO connection continues to be reinforced by other observational and modeling studies (e.g., Soukharev, 1999; Labitzke and van Loon, 2000; Gray et al., 2001; Labitzke, 2001; McCormack, 2003; Labitzke, 2004; Cordero and Nathan, 2005; Salby and Callaghan, 2006; Soukharev and Hood, 2006; Camp and Tung, 2007; McCormack et al., 2007; Smith and Matthes, 2008).

Since Labitzke and van Loon’s (1988) study, important advances have been made in sun–climate research. For example, monitoring of solar irradiance (total and spectral) from ground and satellite-based platforms have better constrained the radiative forcing in chemistry–climate models (Fröhlich and Lean, 2004). Observations of the middle atmosphere have affirmed a connection between quasi-decadal variations in ozone (and other...
chemical constituents) and quasi-decadal variations in solar spectral irradiance (Soukharev and Hood, 2006). And atmospheric general circulation models have improved to the point that mechanisms for communicating solar variability to the climate system can be evaluated with greater confidence (Haigh and Blackburn, 2006).

The advances in sun–climate research notwithstanding, questions remain that are important to policy-makers and scientists alike. The questions relate largely to the attribution of solar radiative forcing to global climate change over the past century and particularly over the past few decades. The central issue is quantifying the relative contributions of solar and anthropogenic forcing to global warming. Although some statistical studies have suggested that trends in solar forcing can account for a large fraction of the observed trends in global temperature, upon closer scrutiny the studies have proven specious (this issue is lucidly suggested that trends in solar forcing can account for a large fraction of the observed trends in global temperature, upon closer scrutiny the studies have proven specious (this issue is lucidly suggested that trends in solar forcing can account for a large fraction of the observed trends in global temperature, upon closer scrutiny the studies have proven specious (this issue is lucidly discussed in Benestad and Schmidt, 2009). Recent studies estimate that the solar contribution to 20th century warming ranges from 7% to 10% (Benestad and Schmidt, 2009; Lean, 2010). Specifically, based on a suite of global climate simulations, Benestad and Schmidt conclude: “the most likely contribution from solar forcing a global warming is 7 ± 1% for the 20th century and is negligible for warming since 1980.” Employing an empirical model, Lean (2010) concludes: “trends in solar irradiance in the past century contribute global warming of 10% or less.”

For the remainder of this paper, we focus primarily on the 11-year SC modulation of wave–mean flow interaction in the middle atmosphere (the region between about 10 and 100 km). In the following section (Section 2), we present a historical overview of the connection between variations in solar irradiance, stratospheric ozone, and wave–mean flow interaction. In Section 3, we propose a theoretical framework for solar modulated wave–mean flow interaction that is more complete than previously espoused. Our framework is built around two ozone pathways that together serve as intermediaries for communicating changes in solar irradiance to the wave-driven circulation. One pathway hinges on zonal-mean ozone and the other on zonally asymmetric ozone. The role of zonally asymmetric ozone (ZAO) in sun–climate connections has not been explored, save for the mechanistic study of the solar-modulated QBO by Cordero and Nathan (2005). There is, however, an ample body of theoretical, observational, and modeling work showing the importance of ZAO to both the extratropical and tropical circulations (e.g., Garcia and Hartmann, 1980; Nathan, 1989; Nathan and Li, 1991; Nathan et al., 1994; Echols and Nathan, 1996; Cordero et al., 1998; Cordero and Nathan, 2000; Gabriel et al., 2007; Nathan and Cordero, 2007; Crook et al., 2008; Waugh et al., 2009). In view of this body of work, we hypothesize that ZAO may also play an important role in the sun–climate problem. We support this hypothesis using a mechanistic chemistry–climate model, first analyzed analytically in Section 3.3, and then numerically in Section 3.4. We close with a brief discussion of the remaining issues surrounding solar-modulated wave–mean flow interaction and how our proposed theoretical framework might assist in the interpretation of results obtained from chemistry–climate models.

2. Historical overview

Exploring the connection between solar variability and climate has engaged and challenged scientists for decades. In this section, we provide an overview of the developments and current understanding of the physics that connect variations in solar irradiance (total and spectral) to wave–mean flow interaction. We begin with an overview of the connection between solar irradiance, ozone, and temperature, and follow with a simple theoretical framework that serves as a conceptual guide for interpreting the literature on solar-modulated wave–mean flow interaction in the middle atmosphere.

2.1. Solar irradiance, ozone, and temperature

By compositing proxy, ground-based, and satellite data, reconstructions of solar variability have been extended back several centuries (see Fig. 1). Variations in solar irradiance have been recorded at ground-based stations for more than a century and from satellites since 1979 (Lean 1997). Satellite measurements of total solar irradiance (TSI) show variations associated with the 11-year solar cycle of only < 0.1% (Fröhlich and Lean, 2004). Given such small variation in TSI, the challenge has been, and continues to be, to account for what appears to be a disproportionately large response in the climate system. Thus, much effort has been devoted to identifying the mechanism(s) that can amplify and communicate variations in solar activity to Earth’s climate system.

Sun–climate mechanisms can be broadly categorized as direct and indirect (Gray et al., 2005), which may combine to produce a response in the climate system that is greater than their sum (Meehl et al., 2009). Direct mechanisms are associated with variations in TSI, whereas indirect mechanisms are typically associated with variations in solar spectral irradiance (SSI), primarily in the UV. In contrast to total TSI, which varies by < 0.1% over the SC, variations in SSI over the SC show much larger variations in the UV: 20% at 140–155 nm; 10% at 170–190 nm; and 8% at 200 nm (Lean, 1997). These relatively large UV variations affect the photochemical production and destruction of stratospheric ozone, and thus influence the temporal and spatial variability of ozone (Soukharev and Hood, 2006).

Because stratospheric ozone absorbs UV radiation, it also plays a central role in shaping the thermal structure of the middle atmosphere. Ground-based and satellite observations have been used to measure the variations in stratospheric ozone associated with SC changes in UV radiation (e.g. Wang et al., 1996; Randel and Wu, 2007; Soukharev and Hood, 2006). One of the principle

Fig. 1. Reconstruction of total solar irradiance from different analyses (top) and from sunspot number (bottom) from 1610 through 2004. The upper envelope of the shaded region denotes irradiance variations associated with the 11-year solar cycle, while the lower envelope denotes the total irradiance from the Lean (2000) reconstruction (Wang et al., 2005 and Tapping et al., 2007). The figure is adapted from Lean (2010).
challenges in quantifying these solar-induced ozone changes is that only two complete SCs have been recorded since satellite observations of ozone have been available. In addition, changes in satellite instruments and signal contamination by volcanic aerosols must be accounted for in making accurate assessments of the SC influence on ozone. Despite these challenges, an improving collection of observations has established a consensus regarding the relationship between the SC and ozone: observations of global and tropical total ozone vary by 2–3% over the SC (Calisesi and Matthes, 2006). Total ozone anomalies vary considerably with latitude, and the magnitude of these changes is partially dependent on the observed dataset. In Fig. 2, the annual solar regression coefficients from SAGE II observations show maxima in the upper stratosphere in the tropics, and in the middle stratosphere at 40°N and 30°S. A seasonal analysis of these changes also shows very sharp gradients in ozone change, especially in the winter high latitudes (Soukharev and Hood, 2006). Thus, the large-scale structure of SC related ozone anomalies is characterized by positive ozone anomalies (during solar maximum) in the upper stratosphere (40–50 km) and lower stratosphere (below 25 km), with a weak signal present in the middle stratosphere (∼32 km). While lower stratospheric ozone changes are persistent across different observing platforms, this feature is not well captured by models (WMO, 2007). Various analyses have suggested that this lower stratospheric ozone response may be dynamical in origin (Hood and Soukharev, 2003; Hood, 2007).

Studies of observed and reanalysis datasets have attempted to quantify the stratospheric temperature response to the 11-year SC modulation of ozone (Labitzke, 2001; Labitzke et al., 2002; Crooks and Gray, 2005). In the upper stratosphere, it appears that during solar maximum, temperatures are generally warmer in the lower and middle latitudes than during solar minimum (e.g., Claud et al., 2008). This agrees with Labitzke and van Loon’s (1988) analysis of radiosonde data. Analyses by Hood (2004) showed more structure in the temperature response in low latitudes: positive temperatures at 48 km; negative at 32 km; and positive again between 16 and 20 km. Additional studies using NCEP and ERA-40 reanalysis data again found a positive response in the lower equatorial stratosphere (Haigh et al., 2005; Crooks and Gray, 2005). Significant uncertainties exist, however, regarding the potential relationship between these signals in lower stratospheric tropical ozone, the QBO, and volcanic eruptions (Crooks and Gray, 2005). The uncertainties exist in part from the inability of a linear multiple regression analysis to separate potential non-linear interactions between the SC and the QBO.

At high latitudes, the polar vortex is stronger, less disturbed, and thus colder during solar maximum compared to solar minimum (Gray et al., 2004; Camp and Tung, 2007). Claud et al. (2008) show a colder polar vortex during the entire winter in the Southern Hemisphere and during December–January in the Northern Hemisphere winter. However, during February–March, the vortex is actually warmer. A plausible mechanism to explain the response of the polar vortex to solar forcing is that normally during solar maximum, ozone heating in the upper tropical stratosphere increases the meridional temperature gradient and thus promotes a stronger (and colder) vortex. Concurrently, the Brewer–Dobson circulation (BDC) is weaker (warmer equatorial region in the lower stratosphere). This changes during February–March, however, when the BDC is enhanced, which produces local cooling in the equatorial stratosphere and local warming in the high latitude stratosphere. An Eliassen–Palm (EP) flux analysis suggests that the enhanced BDC may be due to enhanced tropospheric wave driving (Claud et al., 2008). Studies have also found a relationship between the SC, high latitude circulation, and the QBO (Gray et al., 2004). Identifying precisely how these features are connected, however, may require a longer period of observations.

2.2. Wave–mean flow interaction

Many prominent atmospheric circulation features that are closely associated with wave–mean flow interaction have been shown to be modulated by solar activity. These SC-modulated features include the polar vortex (Kodera and Kuroda, 2002), planetary wave activity/refraction (Shindell et al., 1999; Rind et al., 2002), stratospheric sudden warmings (Gray et al., 2004; Gray et al., 2006), the Brewer–Dobson circulation (Kodera and Kuroda, 2002), annular modes (Baldwin and Dunkerton, 2005), and the equatorial quasi-biennial oscillation (Gray et al., 2001; McCormack, 2003; Cordero and Nathan, 2005). Yet, despite these studies, the precise pathways for communicating solar variability from the middle to lower atmosphere have yet to be fully understood, particularly with regard to the role of ozone. In the following two subsections, we provide an overview of the developments connecting the SC, stratospheric ozone, and wave–mean flow interaction.

2.2.1. Theory and prior work

We begin our overview by introducing a theoretical framework that will ease interpretation of the physics connecting solar variability with wave–mean flow interaction. To do so, we appeal to the divergence of Eliassen–Palm (EP) flux, which measures the wave-driving of the zonal-mean circulation (Andrews and McIntyre, 1976). To show in conceptual terms the connection between solar-modulated wave propagation and attenuation and the divergence of the EP flux, we use the quasi-geostrophic framework. We assume the zonally uniform background flow is slowly varying, which permits a steady-state, phase-integral (WKB) solution for the disturbance streamfunction field of the form (Nathan and Hodyss, 2010),

$$\phi(x,y,z) = a \exp \left[ \left( kx + \int idy + \int mdz \right) \right] + c.c.$$  \hspace{1cm} (1)

where $K = ki + lj + mk$ is the wavevector; $k$ is the (constant) zonal wavenumber, and $l(y,z)$ and $m(y,z)$ are local, slowly varying (complex) wavenumbers in the meridional and vertical directions, respectively; $a(y,z), K(y,z),$ $\nabla \phi(y,z)$ is the slowly varying disturbance amplitude, which is a function of $\phi(y,z)$, the wavevector, and $\phi(y,z)$, a function comprising the background flow variables and
solar-modulated ozone heating; and c.c. denotes the complex conjugate of the preceding term.

Based on Eq. (1), the divergence of the EP flux (planetary wave drag) can be written to lowest order as

$$\nabla \cdot F \propto -a^2 (k l_i + m m_i) \exp \left( -\int l_i \, dy - \int m_n \, dz \right) , \tag{2}\n$$

where constants have been omitted for simplicity. The real and imaginary parts of the wavenumbers measure, respectively, wave propagation and wave attenuation. Together $l_i$ and $m_n$ determine the wave refraction in the latitude-height plane. In the absence of wave damping, for which $l_i = 0$ and $m_n = 0$, $\nabla \cdot F$ vanishes, in accordance with the Charney and Drazin (1961) non-acceleration theorem.

In the context of the SC-modulation of the zonal-mean flow, changes in ozone heating due to changes in SSI affect both propagation and attenuation. As Eq. (2) shows, propagation and attenuation are multiplicative, thus making clear the nonlinear connection between changes in SSI, planetary wave activity, and the zonal-mean flow. This point will be addressed in greater detail in Section 3, where we explicitly make the connection between SSI, ozone chemistry and transport, and planetary wave drag (PWD). For now, Eq. (2) will serve as a conceptual guide for interpreting previous work connecting variations in SSI with the zonal-mean circulation.

Perhaps the most cited pathway for connecting variations in solar activity to the climate system, one which is intimately connected to wave-mean flow interaction, is what is commonly referred to as the solar/UV/planetary wave mechanism (Geller, 2006). This mechanism operates as follows. The variations in solar spectral irradiance at primarily ultraviolet (UV) wavelengths produce variations in the photochemical production of ozone in the stratosphere. In turn, these ozone changes produce changes in the radiative heating and meridional temperature gradient, which, via thermal wind balance, perturb the spatial distribution and strength of the stratospheric zonal-mean winds. These solar-perturbed changes in zonal-mean wind produce changes in the index of refraction of the planetary waves, resulting in changes in PWD.

The conceptual framework for the solar/UV/wave mechanism can be traced back to Hines (1974). Based on Charney and Drazin’s (1961) seminal work on the vertical propagation of planetary wave activity, which showed that planetary waves generated in the troposphere could propagate to great heights during Northern Hemisphere winter, Hines hypothesized that variations in solar activity could perturb the winds in the upper atmosphere, thus causing changes in planetary wave reflection that could affect the interference between upward and downward propagating waves. The result would be a change in the tropospheric wave pattern.

Geller and Alpert (1980) placed Hines’ (1974) hypothesis on firmer dynamical footing by employing a quasi-geostrophic planetary wave model, where the stratospheric winds were perturbed in a manner to mimic the changes that were assumed to occur from variations in solar activity. They concluded that “planetary wave coupling between the troposphere and the upper stratosphere appears to be a plausible mechanism to give a tropospheric response to solar activity.” Bates (1980, 1981) also employed a quasi-geostrophic framework to examine how changes in solar irradiance might affect radiative damping and stratospheric wind, thus altering planetary wave structure in the troposphere. Like Geller and Alpert, Bates concluded that “variations in solar UV radiation can lead to changes in the mean temperature and wind distributions in the stratosphere...[which] can influence the stratospheric propagation characteristics of planetary waves, leading to changes in the steady-state interference pattern of these waves at all levels.” But there was another assertion made by Bates, one that is not only omitted from the solar/UV/wave mechanism as traditionally cited, but one that has been largely ignored when connecting variations in solar activity to changes in planetary wave activity. Bates asserted that planetary waves could be directly affected by solar activity by modifying the ozone photochemistry to affect the wave damping (attenuation). But as we show in Section 3.3 below, even the notion of solar-modulated wave damping is incomplete; variations in SSI also will affect the heating caused by ozone transport, which, depending on the ozone distribution and the ratio of dynamical to radiative–photochemical time scales, may actually reduce the damping of the planetary waves.

Hines (1974), Geller and Alpert (1980), and Bates (1980, 1981) all inferred the effects of solar variability through perturbations to the zonal-mean wind. These studies were followed by others that relied on more direct connections between solar variability, ozone, temperature, and wind (Callis et al., 1981; Haigh, 1996; Shindell et al., 1999; Balachandran et al., 1999; Shindell et al., 2001; Kodera and Kuroda, 2002; Rind et al., 2002; Haigh and Blackburn, 2006). Haigh (1996), for example, used a two-dimensional radiative–chemical–transport model to show a highly nonlinear relationship between changes in SSI and stratospheric ozone with a “consequent change in latitudinal temperature gradient...and potentially planetary wave activity.” Shindell et al. (1999) employed a stratospheric general circulation model with realistic irradiance and ozone changes and showed that in Northern Hemisphere middle latitudes, solar heating differences consistent with the SC produced changes in the latitudinal temperature gradient and horizontal shear of the zonal-mean wind. These changes increased the quasi-geostrophic potential vorticity thereby altering planetary wave refraction. Rozanov et al. (2008) showed using a coupled-chemistry climate model that “enhancement of solar spectral irradiance [over the 11-year solar cycle] leads to an acceleration of the polar night jets and suppression of the Brewer–Dobson circulation...” Fig. 3 shows the changes in zonal-mean wind over the solar cycle obtained by Rozanov et al. At ~55 km altitude near 60°N, the zonal-mean wind increases by ~5 m/s from solar minimum to solar maximum. Qualitatively similar results have been obtained by many others (e.g., Balachandran et al., 1999; Shindell et al., 2001; Rind et al., 2002). But it is Balachandran et al. (1999) who state most succinctly what has become the mantra for the solar/UV/wave mechanism: “UV changes associated with solar variability bring about changes in the conditions of propagation and dissipation of planetary waves.”
waves." These changes alter the wave driving of the zonal-mean circulation [see Eq. (2)], which are evident not only in the zonal-mean winds but in the residual circulation as well. Solar forcing of the residual circulation may then yield, via the “downward control” principle (Haynes et al., 1991), changes in the lower atmosphere.

The solar/UV/wave mechanism hinges on solar induced ozone changes directly affecting the zonal-mean temperature and thus the zonal-mean wind. The solar/UV/wave mechanism does not form a complete dynamical picture, however. In reality, the zonal-mean wind can also be indirectly affected by solar induced changes in the planetary waves themselves. As shown by Nathan (1989), Nathan and Li (1991), Nathan et al. (1994), and Nathan and Cordero (2007), zonally asymmetric ozone can produce changes in the spatial structure of forced extratropical Rossby waves. These changes alter wave transience and wave dissipation, which produce zonally averaged wave fluxes. In the transformed Eulerian-mean formalism (Andrews and McIntyre, 1976), the divergence of these wave fluxes drives both the zonal-mean flow and the residual mean meridional (Brewer–Dobson) circulation. The solar/UV/wave mechanism should, therefore, be supplemented with the zonal-mean changes that arise from the zonally asymmetric ozone field.

In the following section, we hypothesize a solar/UV/wave mechanism that accounts for both zonal-mean and zonally asymmetric ozone. Preliminary results are then presented that support our hypothesis.


3.1. Pathways

Fig. 4 illustrates our theoretical framework for communicating and amplifying variations in solar activity to the wave-driven circulation. The framework hinges on two key pathways. Pathway I contains the traditional solar/UV/wave mechanism as well as the effects of eddy ozone flux convergences, where the latter arises from the zonally asymmetric ozone field (ZAO). Pathway II involves solar modulation of the planetary wave drag (PWD) by the ZAO. The indirect coupling between the solar forcing and ZAO can be explained via reasoning given by Nathan and Cordero (2007). Briefly, as planetary waves propagate from the troposphere into the stratosphere, wave-like (zonally asymmetric) perturbations in the wind and temperature fields produce wave-like perturbations in the ozone field. All three wave fields are coupled to each other as well as to the zonal-mean fields of wind, temperature, and ozone. Consequently, the fluxes produced by the wave fields, which are embodied in the PWD, also depend on the solar-modulated, zonal-mean background fields. With this in mind, consider the two pathways shown in Fig. 4.

Along pathway I, the zonal-mean ozone field is modulated by variations in SSI—an externally forced top-down effect—and by eddy-ozone flux convergences produced by planetary wave activity forced from below—an internally forced bottom-up effect. The former effect is the traditional solar/UV/wave mechanism, whereas the latter effect is due to the ZAO field. These effects combine to produce latitudinal variations in zonal-mean ozone heating, which produce latitudinal changes in zonal-mean temperature. By thermal wind balance, there are corresponding changes in the zonal-mean wind, leading to refraction, attenuation, and perhaps partial reflection of the planetary waves. The planetary waves can be further modulated via the Holton and Tan (1980) mechanism, whereby solar-induced variations in the QBO affect the subtropical zero wind line, causing a change in the planetary wave guide.

Along pathway II, ZAO directly modulates the PWD, which alters both the zonal-mean wind and the Brewer–Dobson circulation. The PWD exerted by ZAO may manifest in three ways:
through local changes in wave–mean flow interaction, which may result in a downward propagating signal (Plumb and Semeniuk, 2003); through “downward control,” whereby the ozone-induced PWD causes a mean meridional circulation and a simultaneous mass adjustment in the surface pressure (Haynes et al., 1991); and through the refraction and downward reflection of vertically propagating planetary waves (Perlwitz and Harnik, 2003). The precise way in which solar variability modulates downward propagating signals produced by these mechanisms remains unclear.

3.2. Coupled-chemistry mechanistic model

To examine the connection between variations in solar spectral irradiance, the zonal-mean and zonally asymmetric ozone fields, and the dynamical circulation, we employ the radiative–photochemical–dynamical model of Nathan and Cordero (2007) in combination with the quasigeostrophic form of the transformed Eulerian-mean equations. Briefly, the model atmosphere is forced from below and confined to a mid-latitude channel of width L centered at 60° N. The quasi-geostrophic flow is linearized about a steady, zonally averaged basic state that is in radiative–photochemical equilibrium. As in Nathan and Cordero, the basic state zonal wind \( \beta \) is assumed to vary only with height, an assumption that facilitates identifying the connection between variations in SSI, planetary wave induced ozone heating, and the zonal-mean circulation. The linear response of the perturbation fields to wave-induced ozone heating (OH) and Newtonian cooling (NC) is described by coupled equations for quasi-geostrophic potential vorticity, \( q \), and ozone volume mixing ratio, \( \gamma \), which in log-pressure coordinates take the form

\[
\frac{\partial}{\partial t} + \mathbf{U} \cdot \nabla q = \frac{1}{\rho} \frac{\partial}{\partial z} \left( \frac{\rho \beta \phi}{\sigma} q \right) + \nabla \cdot \left( \mathbf{B} \right),
\]

(3)

where the perturbation potential vorticity, \( q \), and basic state potential vorticity gradient, \( \beta_e \), are, respectively,

\[
q = \nabla^2 \phi + \frac{1}{\rho} \frac{\partial}{\partial z} \left( \frac{\rho \beta \phi}{\sigma} \right),
\]

(5)

\[
\beta_e = \beta - \frac{1}{\rho} \frac{\partial}{\partial z} \left( \frac{\rho \beta \phi}{\sigma} \right).
\]

(6)

The diabatic heating rate per unit mass, \( Q \), and the ozone production and destruction term, \( S \), are, respectively,

\[
Q = \Gamma_1 \gamma + \Gamma_2 \int_0^z \frac{\rho(x)/\rho_0}{H_0} \gamma dz - f_0 \frac{H_0}{K} \Gamma_1 \frac{\partial \phi}{\partial z},
\]

(8)

\[
S = -\xi_0 \gamma + \xi_2 \int_0^z \frac{\rho(x)/\rho_0}{H_0} \gamma dz - f_0 \frac{H_0}{K} \xi \frac{\partial \phi}{\partial z}.
\]

(9)

The first two terms on the right-hand side (rhs) of the diabatic heating rate, \( Q \), constitute the zonally asymmetric variations in OH produced by the planetary waves. The first term on the rhs of Eq. (8) is the local OH rate and the second term, called the shielding term, is the non-local OH rate that arises from variations in ozone above a given level. The third term on the rhs of Eq. (8) represents the effects of long wave radiative cooling, which we model as Newtonian cooling (NC; Dickinson, 1973). Like the diabatic heating rate, the first two terms on the rhs of the ozone production and destruction term, \( S \), are due to local and non-local (shielding) ozone perturbations, respectively. The third term on the rhs of Eq. (9) constitutes the ozone production/destruction that arises from temperature perturbations. The oxygen-only radiative–photochemical model is based on Nathan and Li (1991), which we have updated using reactions rates based on Sander et al. (JPL) (2006); an accounting of ozone catalytic loss cycles involving hydrogen, nitrogen, and chlorine based on the approximation of Haigh and Pyle (1982); temperature-dependent ozone absorption cross sections of Molina and Molina (1986); and enhancement of solar radiation due to multiple scattering, surface reflection, and clouds and aerosols using the approximation of Nicolet et al. (1982). The details of the radiative–photochemical model are described in Appendix A.

The boundary conditions at the channel side-walls require that the meridional velocity vanish, i.e., \( \partial \phi / \partial x = 0 \) at \( y = 0, L \). For the analytical analysis presented in Section 3.3, a radiation condition is imposed, which requires that the upward energy flux \( \rho_0 \mathbf{w} \rightarrow 0 \) as the height \( z \rightarrow \infty \) (the overbar denotes a zonal average). At the lower boundary, a planetary wave is imposed that forces the model circulation (see Section 3.4 for details).

To examine solar modulated wave–mean flow interaction, Eqs. (3)–(9) must be supplemented with the transformed Eulerian mean equations, which can be written as (Holton, 2004)

\[
\frac{\partial}{\partial t} - f_0 \mathbf{v} = \rho_0^{-1} \nabla \mathbf{F} + \mathbf{X} + \mathbf{F}^\prime.
\]

(10)

\[
\frac{\partial}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{F} + \mathbf{H} = \mathbf{U} + \mathbf{F}^\prime.
\]

(11)

\[
\frac{\partial}{\partial y} + \mathbf{U} \cdot \nabla \mathbf{F} + \mathbf{H} = \mathbf{U} + \mathbf{F}^\prime.
\]

(12)

\[
\frac{\partial}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{F} + \mathbf{H} = \mathbf{U} + \mathbf{F}^\prime.
\]

(13)

\[
\frac{\partial}{\partial t} - f_0 \mathbf{v} + RH^{-1} \frac{\partial}{\partial y} \mathbf{F} = \mathbf{X} + \mathbf{F}^\prime.
\]

(14)

In the above equations, the overbar denotes zonally averaged basic state quantities; \( \mathbf{X} \) is the zonal-mean zonal wind; \( \mathbf{v} \) and \( \mathbf{w} \) are the residual meridional and vertical velocities, respectively; \( \nabla \mathbf{F} \) is the divergence of Eliassen–Palm flux, which is a proxy for the planetary wave drag associated with the large-scale waves; \( \mathbf{X} \) is the wave drag associated with the small-scale (i.e., gravity) waves; \( \mathbf{F} \) is the zonal-mean friction; \( \mathbf{U} \) is the net zonal-mean diabatic heating rate, which, like that for the perturbation field, is comprised of Newtonian cooling and local and non-local ozone heating; \( \mathbf{S} \) is the zonal-mean production and destruction of ozone; and \( \mathbf{G} \) is the net zonal-mean wave–ozone fluxes. A listing of the model parameters and variables is presented in Table 1.

For the remainder of this study, we focus our attention on the changes in the zonal-mean flow arising from the solar-modulated wave drag exerted by the resolved planetary waves, measured by \( \nabla \mathbf{F} \). We will not consider contributions to the zonal-mean flow tendency arising from the wave drag exerted by the small scale waves \( \mathbf{X} \) and friction \( \mathbf{F}^\prime \).

The theoretical basis that underlies the connection between solar variability, planetary wave drag, zonally asymmetric ozone heating, and the zonal-mean circulation is most easily expressed by considering the steady-state form of Eqs. (3)–(14), for which Eqs. (10) and (12) can be combined and the radiation boundary condition applied to yield the steady-state form of the
residual vertical velocity
\[ w^* = -\rho_0 \frac{\partial}{\partial z} \left[ \int_0^z \nabla \cdot F dz \right]. \] (15)

This equation makes clear the notion of “downward control.” In particular, an upper-level perturbation in planetary wave drag \( \nabla \cdot F \), solar induced or otherwise, is such that \( w^* \) is zero above that level and non-zero below.

### 3.3. Analytical results

To obtain an expression for the solar-modulated PWD, which operates along pathway II in Fig. 4, we neglect the shielding terms\(^1\) in Eqs. (8) and (9) and assume that the background distributions for wind, temperature, and ozone are slowly varying. As in Nathan and Cordero (2007), we formalize the slowly varying assumption by introducing the new vertical coordinate \( z^* = z \) such that \( \partial / \partial z \rightarrow \partial / \partial z + \partial / \partial z^* \), where \( \partial / \partial z \) is non-dimensional. This new vertical coordinate and corresponding derivative transformation, together with the assumption that the coefficients in Eqs. (3)–(9) vary only with height, allows steady-state solutions for the streamfunction and ozone fields to be chosen WKB in form (Bender and Orszag, 1978)

\[ \phi(z^*) = \phi_0 \exp \left( i \int_0^z \frac{(\partial \phi_0(z^*))}{\partial z} + m_1(\zeta^*) d \zeta^* \right), \] (17)

and

\[ \gamma(z^*) = b(z^*) \phi(z^*) = |b(z^*)| \exp(i\gamma(z^*)) \phi(z^*), \] (18)

where \( \phi_0 \) is the constant determined from the lower boundary condition. In Eq. (17) \( m_0(\zeta^*) \) is the (complex) ozone-modified refractive index (OMRI), which describes the leading order approximation to the propagation and attenuation of the planetary wave field; \( m_1(\zeta^*) \) is the higher order correction to the OMRI. In Eq. (18) \( |b(z^*)| \) is the amplitude (modulus) of the ozone field and \( \gamma(z^*) \) is the phase angle between the ozone and streamfunction fields.

To make further progress in analytically interpreting the effects of solar-modulated ozone on planetary wave drag, we assume that the ozone heating and Newtonian cooling are small, i.e., \( O(\beta) \), for which the WKB solution Eq. (17) is approximated by

\[ \phi(z^*) = \frac{\phi_0}{\sqrt{M_0}} \exp \left( i \int_0^z \frac{(\tilde{M}\phi(\zeta^*))}{\partial \zeta^*} + M_1(\zeta^*) d \zeta^* \right), \] (19)

where

\[ M_0 = \left[ \sigma \left( \frac{\partial k}{\partial \zeta} (k^2 + \beta^2) \right) - \frac{1}{4} \right]^{1/2}, \] (20)

is the local refractive index (or local vertical wavenumber) originally derived by Charney and Drazin (1961), while

\[ M_1 = M_0 + i M_{1r}, \] (21)

where the expressions controlling wave propagation \( M_{1r} \) and wave attenuation due to damping \( M_{1i} \) are, respectively,

\[ M_{1r} = \frac{\Gamma_{1r} M_0}{1 + \phi_0} \left[ \frac{c_M}{\frac{\partial k}{\partial \zeta}} \frac{\partial k}{\partial \zeta} \right] \left[ \frac{c_2}{M_1^{1/2}} \frac{\partial}{\partial \zeta} \right] \left[ \frac{1}{M_1^{1/2}} \left( \frac{\partial k}{\partial \zeta} \right)^2 \right]^2 \] (22)

and

\[ M_{1i} = \frac{\Gamma_{1i} M_0}{1 + \phi_0} \left[ \frac{c_M}{\frac{\partial k}{\partial \zeta}} \frac{\partial k}{\partial \zeta} \right] \left[ \frac{c_2}{M_1^{1/2}} \frac{\partial}{\partial \zeta} \right] \left[ \frac{1}{M_1^{1/2}} \left( \frac{\partial k}{\partial \zeta} \right)^2 \right]^2 \] (23)

where as Nathan and Cordero (2007) derived simplified expressions for the ozone-modified propagation and attenuation that were valid in either the dynamically controlled lower stratosphere.

---

\(^1\)The shielding terms only have a modifying effect on the local ozone heating and ozone production/destruction terms in Eq. (8) and (9) (see Nathan and Li (1991)). Thus, neglecting the shielding terms will not affect the qualitative conclusions reached for the analytical presentation in this section. The shielding terms are retained in the numerical results section (see §3.4).
or photochemically controlled upper stratosphere, Eqs. (22) and (23) are valid throughout the stratosphere, except near zero wind lines, where $\mathbf{v} = 0$, or near reflecting surfaces, where $\mathbf{M}_0 = 0$. For wintertime climatological wind profiles, $\mathbf{v}$ generally does not vanish. Reflecting surfaces may exist for such profiles, however, which would require additional technical analysis.

In particular, near reflecting surfaces, where $M_0 = 0$, the WKB solution Eq. (16) and consequently Eqs. (22) and (23) become invalid. As described in Nathan and Cordero (2007), to obtain a uniformly valid WKB expansion would require obtaining WKB solutions for the regions below, within, and above the reflecting region, and subsequently matching the solutions. Above the reflection region the wave field would consist solely of an evanescent wave owing to imposition of the radiation condition. Below the reflecting region the solution would consist of both upward and downward propagating waves, with the reflection coefficient determined by the ratio of the upward to downward propagating wave amplitudes. Although the calculation of a uniformly valid WKB approximation and corresponding reflection coefficient is in principle straightforward, in reality the calculations are technically difficult. The difficulty is heightened if the diabatic processes arising from Newtonian cooling and solar-modulated ozone physics are not assumed small, as we have assumed, but rather are permitted to enter the WKB solutions at $O(1)$. In this case, no real WKB turning points would exist in the presence of the diabatic processes. As Boyd (1998) lucidly describes, to determine the reflection coefficient in this case would require examining the turning points in the complex plane, subsequently matching the WKB solutions across the propagation, breakdown, and evanescent regions. Carrying out this procedure, however, is beyond the scope of the present study. It will suffice here to simply recognize that $M_0 = 0$ corresponds to a first approximation, to a reflecting surface in the presence of weak $O(\varepsilon)$ ozone heating effects. In Eqs. (22) and (23) the $\zeta_j = 1 - \varepsilon^j$ are positive constants, which are defined in Table 1; $D_0 = 1 - 1/4M_0^2\beta^2$ and $D_0 = 1 - 1/2M_0\beta$ are both positive except near reflecting surfaces where $M_0 = 0$ or near zero wind lines where $\mathbf{v} = 0$. The ratio of the dynamical to radiative and photochemical time scales are defined as $\tau_D(\varepsilon) = \tau_D / \tau_k$ and $\tau_P(\varepsilon) = \tau_P / \tau_k$, respectively.

The expressions for wave propagation ($M_0$) and wave attenuation ($M_1$) contain three terms that are due solely to the ZAO field. These terms originate from (i) vertical advection of zonal-mean ozone by the wave field; (ii) meridional advection of zonal-mean ozone by the wave field; and (iii) photochemically accelerated cooling. Because each term is a function of the zonal-mean wind, temperature, and ozone fields, which are functions of SSL (see Appendix A), it follows that each ZAO term also is a (nonlinear) function of SSL. Thus, along pathway II in Fig. 4, the ZAO ozone field imparts three solar-modulated physical processes to the PWD which are absent in the traditional solar/UV/wave mechanism. The relative importance of each term depends on the ratio of dynamical to photochemical time scales, which are strong functions of altitude (see, for example, Fig. 3 in Nathan and Cordero, 2007). In the lower stratosphere, meridional ozone advection and vertical ozone advection may augment or oppose each other depending on the spatial distributions of zonal-mean ozone and zonal-mean wind. In the upper stratosphere, photochemically accelerated cooling dominates and always augments the Newtonian cooling. In the mid-stratosphere, ozone transport and ozone chemistry are of comparable importance. What distinguishes the zonal-mean ozone and zonally asymmetric ozone fields as intermediaries for communicating variations in SSL to the zonal-mean circulation is that the latter depends on both eddy ozone transports and eddy ozone photochemistry. Although Bates (1980) hypothesized the solar-induced changes in ozone photochemistry might be important to sun–climate connections, he was apparently unaware of the full nature of ZAO in communicating solar variability to the zonal-mean circulation, while others have only inferred that variations in SSL could affect both planetary wave propagation and damping (e.g., Balachandran et al., 1999). Eqs. (22) and (23) explicitly show the physics that connects variations in SSL and ZAO to wave propagation and wave damping, which together modify the planetary wave drag [see Eq. (2)]. In the following Section, we present some numerical results for the middle atmosphere that reinforce the analytical results presented above.

### 3.4. Some numerical results

To assess how solar-modulated zonal asymmetries in ozone affect the planetary wave drag, we now present some numerical results based on the quasi-geostrophic model described above. The dynamical portion of the model is based on the one-dimensional (in height) model of Holton and Mass (1976), which accounts for the interaction between a single planetary wave and the zonal-mean flow, while the radiative–photochemical portion of the model is based on Nathan and Li (1991). Our mechanistic chemistry–dynamical model has predictive equations for the zonal-mean and wave portions of the wind, temperature, and ozone fields. The model accounts for local ozone heating as well as the heating that arises from ozone perturbations above a given level (called the “shielding effect”). The 11-year solar cycle has been incorporated into our model by altering the solar spectral irradiance between the wavelengths of 175 and 400 nm based on tabulations in Lean (1997) and Lean et al. (1997). Specifically, we have chosen the variations in solar spectral irradiance (SSI) between solar minimum and solar maximum to vary from 10% at 170–190 nm, 3–4% at 250 nm, and less than 0.5% above of 300 nm. These variations in SSL alter the model by modifying the ozone heating and production/destruction rates in the radiative–photochemical terms appearing in both the zonal-mean and wave equations.

At the upper boundary, we assume, as in Holton and Mass (1976), that the wave energy flux vanishes at the upper boundary, thus preventing any spurious reflections that may contaminate solutions in the region of interest. At the lower boundary ($z_B = 10$ km), a planetary wave was imposed that grows monotonically from zero to an asymptotic steady-state, such that $\phi(x,y,z_0) = g_B h_0 f_0 \left[ 1 - \exp(-\lambda t/\tau) \right] \exp(\iota \omega t) \sin \theta; n = k a_0$, $\cos \theta_0$ is the quantized zonal wavenumber, $g_B$ is the acceleration of gravity, $a_0$ is the Earth’s radius, $\theta_0 = 60^\circ$, and $l = \pi / a_0$. Model simulations have been carried out for planetary wave $n = 1$, $\tau = 2.5 \times 10^5$ s, and $h_0 = 35$ geopotential meters, a bottom forcing value that causes the zonal-mean flow to remain westerly as it evolves to a steady-state. The model was initialized with climatological profiles of zonal-mean wind, temperature, and ozone consistent with late winter (February) in the Northern Hemisphere.

We carried out four numerical experiments (see Table 2). These experiments are: (1) zonal-mean ozone (ZMO) only (zonal asymmetric ozone (ZAO) is suppressed), with reference values of SSL (experiment ZMO/Ref-SOL); (2) ZMO only (ZAO is suppressed), but with enhanced SSL consistent with the solar cycle (SC) (experiment ZMO/En-SOL); (3) combined ZMO and ZAO with reference SSL (experiment ZAO/Ref-SOL); and (4) combined ZMO and ZAO, but with enhanced SSL (experiment ZAO/En-SOL).
Table 2
Experiments. The experiments outlined in the text can be separated into two categories: (i) experiments using reference SSI versus experiments using enhanced SSI consistent with the 11-year solar cycle; and (ii) experiments using zonal-mean ozone only versus experiments with combined zonal-mean ozone and zonally asymmetric ozone.

<table>
<thead>
<tr>
<th>Zonal-mean ozone only (ZMO)</th>
<th>Reference SSI</th>
<th>Enhanced SSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZMO/Ref-SOL</td>
<td>ZMO/En-SOL</td>
<td></td>
</tr>
<tr>
<td>ZA0/Ref-SOL</td>
<td>ZA0/En-SOL</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5. Zonal-mean wind (m/s) for: (1) zonal-mean ozone (ZMO) only (zonally asymmetric ozone (ZAO) is suppressed) and reference values of solar spectral irradiance (SSI) (ZMO/Ref-SOL; blue solid line); (2) ZMO only (ZAO is suppressed), but with enhanced SSI due to the 11-year solar cycle (ZMO/En-SOL; blue dashed line); (3) ZMO and ZAO combined with reference SSI (ZAO/Ref-SOL; red solid line); and (4) ZMO and ZAO combined with enhanced SSI (ZAO/En-SOL; red dashed line). (For interpretation of the reference to color in this figure legend the reader is referred to the web version of this article.)

Fig. 5 shows equilibrated zonal-mean wind profiles for the four experiments. We consider first the effects of enhanced SSI in which ZAO is artificially suppressed (only ZMO operates in the model). Comparison of ZMO/Ref-SOL and ZMO/En-SOL shows that in the upper stratosphere and lower mesosphere, the enhanced SSI increases the zonal-mean wind by 1–4 m/s between 50 and 65 km in height. This result is consistent with the work of Rozanov et al. (2008) (see Fig. 3) who found a statistically significant SC-induced increase in the zonal-mean wind of 3–5 m/s between 50 and 65 km in height at 60°N latitude. In the lower stratosphere, between about 20 and 35 km, Rozanov et al. find that the zonal-mean wind decreases slightly, though the result was not deemed statistically significant. Nonetheless, we also find a slight decrease in the zonal-mean wind at the same latitude and height. This lower stratospheric response is intriguing and requires further investigation to see if it is due to local forcing by the solar-modulated ozone field or perhaps due to a solar-induced signal originating from above and migrating downward.

We now examine the effects of enhanced SSI when both ZMO and ZAO are operating in the model. Thus, we return to Fig. 5 and compare the equilibrated zonal-mean wind profiles for ZAO/Ref-SOL and ZAO/En-SOL. In the upper stratosphere and lower mesosphere, the enhanced SSI causes an increase in the zonal-mean wind, similar to the case above when zonal-mean ozone alone was considered. There are two important differences, however, between the experiments with ZMO alone and the experiments with both ZMO and ZAO. First, the enhanced SSI case has a slightly larger effect on the equilibrated zonal-mean wind when ZAO is included (−1–4 m/s increase in the ZMO case versus −2–5 m/s in the ZMO plus ZAO case). Second, comparison between ZMO/En-SOL and ZAO/En-SOL shows that the ZAO increases the altitude of maximum zonal-mean wind by ~5.0 km. In contrast, for the ZMO alone case, an enhanced SSI has no effect on the altitude of maximum zonal wind (compare ZMO/Ref-SOL with ZMO/En-SOL), whereas for the combined ZMO and ZAO case, an enhanced SSI increases the altitude of maximum zonal-mean wind by ~2.5 km.

We have carried out some additional experiments to assess the contributions of pathways I and II in Fig. 4 to the changes in zonal-mean wind caused by enhanced SSI. We also have assessed the relative contributions of wave–ozone transport and wave–ozone photochemistry to V F in pathway II (see analytical analysis in Section 3.3). We find that pathways I and II contribute about equally to the solar-induced change in the zonal-mean wind. Pathway II, however, is mostly responsible for the upward shift in the wind maximum. Within pathway II, we find that the wave–ozone advection and wave–ozone photochemistry are of comparable importance, but it is the photochemistry, i.e., the photochemically accelerated cooling, that is primarily responsible for the altitude increase in the maximum zonal-mean wind.

4. Concluding remarks

For more than 35-years, the solar/UV/wave mechanism has served as the theoretical bedrock for connecting variations in solar spectral irradiance (SSI) to latitudinal changes in zonal-mean ozone, with concomitant changes in the zonal-wind and planetary wave activity. We have shown that this oft-quoted solar/UV/wave mechanism is incomplete; it does not account for the role of SC-modulated changes in planetary wave drag due to zonally asymmetric ozone. We have presented a more complete theoretical framework, one which includes both zonal-mean and zonally asymmetric ozone as intermediaries for communicating variations in SSI to those circulation features driven by wave-mean flow interaction. This theoretical framework hinges on two pathways. Along pathway I, the zonal-mean ozone field is modulated by variations in SSI — an externally forced top-down effect — and by eddy-ozone flux divergences produced by planetary wave activity forced from below—an internally forced bottom-up effect. The former effect is the traditional solar/UV/wave mechanism, whereas the latter effect is due to the zonally asymmetric ozone field. Along pathway II, variations in SSI affect the zonally asymmetric ozone to modulate the planetary wave drag (PWD), which alters both the zonal-mean wind and the Brewer–Dobson circulation.

To place our theoretical framework on firmer footing, we have employed a quasi-geostrophic model of the extratropical circulation that couples radiative transfer, ozone photochemistry, and the dynamical circulation. Based on a WKB analysis, we identify the physics that connect the 11-year solar cycle (SC) to the planetary wave–drag, a body force that embodies the effects of wave propagation and wave attenuation due to damping. Together these SC-modulated wave properties drive the zonal-mean wind and residual circulation.

Using our quasi-geostrophic model, we have presented some numerical results that underscore the importance of the zonally asymmetric ozone (ZAO) field in communicating the effects of solar variability to the wave-driven circulation in the middle atmosphere. By varying the solar spectral irradiance consistent with changes over the 11-year solar cycle, we have shown that ZAO can have a significant impact on the equilibrated (steady-state) zonal-mean wind. This result sparks the following question: how can the solar modulated changes in the zonal-mean wind that we have obtained, which are confined to the middle atmosphere, be communicated downward to affect surface climate? Before providing possible answers to this question, we first emphasize that our results are for an equilibrated circulation,
wherein the wave fields and zonal-mean fields have asymptotically achieved a steady-state. Based on recent a recent study by Hardiman and Haynes (2008), who showed that in a model similar to ours but without ozone or solar variability, transient wave forcing can produce a “spectacular” change in the circulation, a change in which the downward propagating signal is much stronger than for steady-wave forcing. Thus, Hardiman and Haynes (2008) provide a possible answer to the question we posed above. That is, under transient wave forcing, the effects of solar-modulated ZAO may penetrate far deeper into the atmosphere to affect the circulation of the troposphere.

The importance of ZAO to mediating the effects of solar variability in the stratosphere may also extend its influence to the troposphere via the northern and southern annular modes (AMs), which are the dominant patterns of extratropical climate variability on time scales of weeks to months (e.g., Thompson and Wallace, 1998, 2000). Because AMs exist from Earth’s surface to the stratosphere, they provide a possible pathway for communicating solar-induced stratospheric perturbations downward to affect surface climate (e.g., Baldwin and Dunkerton, 1999, 2001; Thompson et al., 2004). Indeed, several modeling studies have suggested that changes in solar activity may modulate AM structures by modifying planetary wave activity and the Brewer–Dobson circulation (Kuroda and Shibata, 2006; Kodera and Kuroda, 2005; Baldwin and Dunkerton, 2005; Tourpali et al., 2005). These studies share a common feature: they all hinge, either directly or indirectly, on planetary waves to communicate and amplify the solar signal to the AMs. In fact, several authors have suggested that wave dynamics may play a critical role in enhancing perturbations local to the stratosphere to AM structures and tropospheric climate (e.g., Kushnir and Polvani, 2004; Song and Robinson 2004). While these research studies suggest that the response of AMs to changes in the stratosphere depend crucially on wave activity, very few studies have fully investigated how perturbations to the stratospheric ozone field—particularly the zonally asymmetric part of the ozone field—may serve as a mediator for communicating the solar signal to the northern and southern AMs. The few studies that have investigated the role of zonally asymmetric ozone (ZAO) in communicating the solar signal to the circulation have confirmed the importance of ZAO in modulating the QBO (Cordero and Nathan, 2005) and the extratropical stratosphere (Nathan and Cordero, 2007). However, the role of ZAO in communicating variations in solar activity to AMs to affect surface climate remains an important, yet unexplored topic.

In addition to examining the effects of SSI on transient wave forcing and the annular modes, further work is also needed to assess how the variations in SSI affect planetary wave structure, particularly with regard to downward reflection. In particular, how sensitive is the planetary wave reflecting layer to SC variations in SSI? Moreover, to sharpen our model as an interpretive tool, we will need to include background fields that vary in both latitude and height. This will allow, at the least, for a more accurate assessment of the effects of solar-modulated ZAO on wave refraction and wave attenuation due to damping and will provide a more complete theoretical basis for understanding how the solar modulation of ZAO affects the Brewer–Dobson circulation.

The quasi-geostrophic, mechanistic model employed here can provide guidance on how to begin to interpret the results obtained from simple atmospheric general circulation models and perhaps even chemistry–climate models. For example, our analytical analysis (Eqs. (22) and (23)) shows the importance of zonally asymmetric ozone to the planetary wave drag, particularly the role played by wave–ozone transports, which depend explicitly on the meridional and vertical gradients of zonal-mean ozone. In light of 21st century projections showing that stratospheric ozone will undergo changes due to reductions in ozone-depleting substances and increases in well-mixed greenhouse gases (WMO, 2007), it remains unclear how these changes in ozone distribution might mediate future cyclical variability and secular trends in solar forcing to the climate system. This will be among the challenges for future sun–climate research.

**Acknowledgments**

Support for this work was provided in part by the NSF Grant ATM-0733698 (T.N.); NASA Living with a Star (LWS) Program Grant NNG05GM55G (T.N. and E.C.), and NASA/NRL Grant NNH08AI67I (T.N.); NSF Faculty Early Career Development ( CAREER ) Program Grant ATM-0449996 (E.C.).

**Appendix A**

### A.1. Radiative/photochemical parameterizations

The ozone heating coefficients can be written as (Lindzen and Goody, 1965; Nathan and Li, 1991)

\[
\Gamma_1(z,t) = \frac{q_1(z,t)N_0}{m_0},
\]

and

\[
\Gamma_2(z,t) = \frac{L_1(z,t)N_0\rho}{m^2_0 \cos \mu},
\]

where

\[
q_1(z,t) = \int_0^\infty F(v) \sigma_1(v) \exp[-\sigma_1(v)\tilde{x}_1(z,t) - \sigma_2(v)\tilde{x}_2(z)] dv,
\]

\[
L_1(z,t) = \int_0^\infty \frac{F(v) \sigma_1^2(v) \exp[-\sigma_1(v)\tilde{x}_1(z,t) - \sigma_2(v)\tilde{x}_2(z)]}{hv} dv,
\]

and

\[
\tilde{x}_1(z,t) = \frac{1}{\cos \mu} \int_0^{\tilde{\mu}_0} \frac{\tilde{n}(\tilde{\mu},t) d\tilde{\mu}}{\rho(\tilde{\mu},t)\gamma(\tilde{\mu},t)}
\]

In the above equations, \( F(v) \) and \( \sigma(v) \) are the solar flux and the absorption cross section at frequency \( v \); respective; \( \tilde{x}_1, \tilde{x}_2, \) and \( \gamma \) are the slant path column density, number density, and mass mixing ratio of molecular oxygen \( (i=2) \) and ozone \( (i=3) \). The solar zenith angle is the \( \mu \) and \( N_0 \) is the Avogadro’s number.

The coefficients appearing in the ozone production/destruction parameterizations are based on the following photochemical model:

<table>
<thead>
<tr>
<th>Reactions</th>
<th>Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>( O_3 + h\nu \rightarrow 2O_2 )</td>
<td>( J_{21}n_2 )</td>
</tr>
<tr>
<td>( O_3 + h\nu \rightarrow O_2 + O_2 )</td>
<td>( J_{31} )</td>
</tr>
<tr>
<td>( O_3 + h\nu \rightarrow O + O_2 )</td>
<td>( J_{31} )</td>
</tr>
<tr>
<td>( O + O_2 + M \rightarrow O_3 + M )</td>
<td>( k_{12}n_3n_2n_m )</td>
</tr>
<tr>
<td>( O + O_3 \rightarrow 2O_2 )</td>
<td>( k_{13}n_3 )</td>
</tr>
</tbody>
</table>

Here, \( k_{12} \) and \( k_{13} \) are the temperature dependent reaction rate coefficients and

\[
J_1(z,t) = \int_0^\infty [F(v)\sigma(v)\exp[-\sigma(v)\tilde{x}_1(z,t) - \sigma(v)\tilde{x}_2(z)]/hv] dv,
\]

is the photodissociation rate for ozone \( (i=3) \) and oxygen \( (i=2) \).
The ozone production/destruction coefficients take the form
\[ \zeta_1(z, t) = -\frac{4C_f J_f(z, t)}{n_2} \exp[-K/T(z, t)], \] (A11)
\[ \zeta_2(z, t) = \frac{N_0 \rho_0}{m_c \cos \phi} \left[ -2\gamma_2 J_2(z, t) + \frac{\zeta_1(z, t)}{2} \right] L_2(z, t), \] (A12)
\[ \zeta_3(z, t) = \frac{\zeta_1(z, t)}{2T(z, t)} \left[ 1 + \frac{K}{T(z, t)} \right], \] (A13)

where
\[ L_2(z, t) = \int_0^\infty [F(v)\sigma_1(v)\exp(-\sigma_2 v)\rho_b(v)]/v^2dv \] (A14)
for ozone \((i=3)\) and oxygen \((i=2)\), and \(C_1 = 3.18 \times 10^{22}\) molecules \(^{-1}\). The temperature dependent reaction rate, \(K\), is defined following the approximation of Haigh and Pyle (1982):
\[ K = 2 + f_1 + f_2 + f_3, \] (A15)
where \(f_1, f_2\), and \(f_3\) are the ratios of the rate of destruction of oxygen due to catalytic cycles involving \(H_2O, NO,\) and \(Cl\), respectively, due to the rate of destruction due to ozone only reactions; \(a = 2570 K, b = 630 K, c = 1400 K,\) and \(d = 260 K\) (Sander et al., 2006 [IPL]).


