Can ozone depletion and global warming interact to produce rapid climate change?

Dennis L. Hartmann*, John M. Wallace, Varavut Limpasuvan, David W. J. Thompson, and James R. Holton

Department of Atmospheric Sciences, University of Washington, Seattle, WA 98195-1640

Edited by Susan Solomon, National Oceanic and Atmospheric Administration, Boulder, CO, and approved December 16, 1999 (received for review October 29, 1999)

The atmosphere displays modes of variability whose structures exhibit a strong longitudinally symmetric (annular) component that extends from the surface to the stratosphere in middle and high latitudes of both hemispheres. In the past 30 years, these modes have exhibited trends that seem larger than their natural background variability, and may be related to human influences on stratospheric ozone and/or atmospheric greenhouse gas concentrations. The pattern of climate trends during the past few decades is marked by rapid cooling and ozone depletion in the polar lower stratosphere of both hemispheres, coupled with an increasing strength of the wintertime westerly polar vortex and a poleward shift of the westerly wind belt at the earth's surface. Annular modes of variability are fundamentally a result of internal dynamical feedbacks within the climate system, and as such can show a large response to rather modest external forcing. The dynamics and thermodynamics of these modes are such that strong synergistic interactions between stratospheric ozone depletion and greenhouse warming are possible. These interactions may be responsible for the pronounced changes in tropospheric and stratospheric climate observed during the past few decades. If these trends continue, they could have important implications for the climate of the 21st century.

A clearly recognizable but unexpected pattern of trends over the past 30 years has recently emerged in global climate records. What is causing it and what it portends for the climate of the 21st century remain uncertain. The first indications of this new pattern can be traced back to the 1990 Intergovernmental Panel of Climate Change (IPCC) assessment of the status of global warming, which reported that winter and springtime temperatures over northern Europe and Siberia were remarkably warm during the 1980s relative to the reference period 1951–1980 (1). Initially, these temperature increases were considered to be an integral part of the greenhouse warming “fingerprint” predicted by climate models. However, it has since become apparent that the warming of these regions is linked to the observed trend toward the “high index” polarity of the Northern Hemisphere annular mode (NAM) during this period (2, 3).

The NAM is a recurrent pattern of wintertime climate anomalies (departures from normal) that is closely related to the North Atlantic Oscillation (4–6) and has also been called the Arctic Oscillation (7, 8). Its temporal variability can be represented by the leading principal component of the sea-level pressure (SLP) field, a weighted average of grid-point values of SLP throughout the hemisphere (7, 9, 10) (Fig. 1). SLP at various stations rises and falls in association with this pattern, much as a drumhead moves up and down in association with its preferred mode (i.e., spatial pattern) of oscillation. However, unlike the drumhead in which the fluctuations tend to be periodic, NAM-related climatic fluctuations are irregular and occur on a wide range of time scales. An analogous structure, referred to here as the Southern Hemisphere annular mode (SAM), has been identified as the leading mode of variability in the Southern Hemisphere (11–14).

The relationship of surface to stratospheric expressions of the NAM can be seen by comparing the NAM SLP pattern with its regression onto the 50-hPa height field during winter (Fig. 1). The regression with the surface pattern bears a strong resemblance to the leading mode of variability of the stratospheric polar vortex (15, 16). The high index state of the NAM connotes an anomalously strong polar vortex and vice versa (16–18). When winter months are classified on the basis of the strength of the westerly jet that encircles the polar cap region at lower stratospheric (∼20 km) levels, the ones in which this “polar vortex” is abnormally strong have been shown to correspond to high index months of the NAM when the tropospheric westerlies along 55°N are anomalously strong. This relationship has been simulated in experiments with numerical models (15, 19).

The NAM has exhibited an apparent systematic trend toward its high-index polarity starting around 1970 (3, 20, 21). Pressures have fallen over the Arctic (22) and have risen in lower latitudes whereas the westerlies have strengthened at subpolar latitudes and have weakened at lower latitudes. These changes are evident at all levels and during all seasons, but they are most pronounced at lower stratospheric levels during winter (3). The SAM also appears to show a trend toward its high index state in recent years (23, 24).

The basic structures of the annular modes of variability result from the internal dynamics of the atmosphere. The observed structure, amplitude, and time scale of both the NAM and SAM in the troposphere can be simulated in an atmospheric climate model with the climatological sea surface temperature distribution (19, 25–27). This means that the annular modes are free, internal modes of tropospheric variability that will occur in the absence of any external forcing. Such annular variability results from strong interactions between the time-mean flow and the eddies that are superimposed on it. In particular, poleward eddy momentum fluxes interact strongly with the zonal flow to sustain north–south displacements of the midlatitude westerlies. In the Southern Hemisphere, the interaction is primarily between the zonal flow and transient baroclinic eddies with periods between 2 and 7 days, and annular mode variability occurs in all seasons (14). In the Northern Hemisphere, the transient eddies are also important, but the momentum fluxes by quasistationary waves in the Atlantic Ocean sector seem to dominate the zonal momentum budget (25, 26, 28).

We argue in this paper that the dynamical linkages between annular variability in the troposphere and stratosphere are very strong, and that changes above or below the tropopause influence the behavior of the annular modes at all levels. We show evidence that the tropospheric annular mode is associated with strong modulation of the flow of planetary wave amplitude into the polar stratosphere. When it is in its high index state, the tropospheric jet is displaced poleward of its climatological
Fig. 1. Monthly mean sea-level pressure [in units of 1,000-hPa height (Left) and 50-hPa height (Right)] regressed onto November–April monthly mean values of the NAM pattern index. Units are meters per standard deviation of pattern amplitude.

Position, vertical shear in the high latitude troposphere and lower stratosphere is increased, and planetary wave activity is ducted away from the polar region. The surface wind changes associated with the tropospheric annular mode also alter the forcing of planetary waves. In particular, the source latitudes for planetary waves move poleward with the tropospheric winds, and the distribution of upward planetary wave flux shifts from zonal wavenumber 1 to wavenumber 2, indicating differences in the shape of the vortex at all levels. Both changes in planetary wave sources and propagation favor a colder and stronger stratospheric polar vortex when the zonal winds are displaced poleward. The trends in the stratosphere and troposphere since 1970 are thus consistent with the idea that this mechanism links them together. The causes of these trends could be stratospheric ozone depletion, greenhouse gas (GHG) increases, natural climate variability, or some combination of all three.

Trends in the Stratosphere. The temperature of the polar winter stratosphere is controlled by a balance between dynamical heating induced by adiabatic compression in the high latitude, sinking branch of a global-scale wave-driven meridional circulation, and radiative cooling. The radiative cooling rate depends on the concentration of GHG, and in sunlit regions on the concentration of ozone, which absorbs solar energy in the ultraviolet. The strength of the meridional circulation also depends on the upward flux of planetary wave activity into the stratosphere (29), which is stronger in the Northern Hemisphere (NH) because of the stronger orographic and land-sea thermal contrasts there. Thus, temperatures in the Antarctic stratosphere are colder and closer to radiative equilibrium throughout the winter, when they are cold enough to support the formation of polar stratospheric clouds whose droplets are the sites of chemical reactions that convert chlorine- and bromine-containing substances to photochemically active forms. The existence of a strong vortex isolates the cold, chemically modified polar air (30). With the return of sunlight in early spring, chlorine- and bromine-catalyzed photochemical reactions destroy most of the stratospheric ozone in the Antarctic polar cap region. The springtime “ozone hole” developed in the late 1970s in response to the positive trend in atmospheric chlorine caused by human activities (31–33).

The differences between the stratosphere in the Southern Hemisphere (SH) and NH are indicative of the important interactions among dynamics, radiation and chemistry. Because the stratosphere is very nearly in geostrophic and hydrostatic balance, the strength of the wintertime westerly vortex that encircles the polar cap region is proportional to the temperature contrast between the polar cap region and lower latitudes. Consistent with the lower temperatures in the polar cap region, the wintertime SH polar vortex is much stronger and longer lasting than its NH counterpart, as indicated by the darker blue in Fig. 2. The SH wintertime stratospheric polar vortex forms about a month earlier in autumn than its NH counterpart, and it persists about 2 months later into the spring (34).

The wintertime westerly vortex interacts strongly with the flux of planetary wave activity into the stratosphere from below (29, 35). If the vortex is properly conditioned or the planetary waves are sufficiently strong, planetary waves propagating up from the troposphere can give rise to abrupt midwinter warmings. Planetary wave forcing in the SH is much weaker and vortex variability is much less in winter than in the NH (Fig. 2, inner ring of numbers). The weaker wave-driven meridional circulation during the SH winter is reflected in the relative warmth of the tropical tropopause during that season (36).

With the trend toward the high index polarity of the NAM in recent decades, the NH stratospheric wintertime vortex has been getting colder and stronger (3, 37–40) and persisting longer in the springtime (41–43), reflecting the decreased incidence of midwinter warmings (44). The cooling trends are largest in the active periods leading up to the collapse of the westerly vortex in springtime, as indicated by the outer ring of numbers in Fig. 2. Consistent with the colder ambient conditions, indications of an incipient ozone hole have appeared in the NH (45–47). Some of the ozone decline during the 1980s (48) and 1990s (49) in the NH is a result of reduced ozone transport associated with decreased planetary wave activity. Changing distributions of...
other trace gases also indicate a slowing of the meridional overturning of the stratosphere in recent years (50, 51). The SH polar vortex has exhibited little change during midwinter, when it may already be as strong and cold as it can get. The vortex with its ozone hole has been persisting even later in the spring, however.

Can the Troposphere Drive Stratospheric Annular Variability? The tropospheric wind variability associated with annular modes is characterized by a meridional displacement in the midlatitude westerlies with opposite changes in zonal wind strength around 35 and 55 degrees of latitude in the Northern Hemisphere and 40 and 60 degrees of latitude in the Southern Hemisphere (3, 13, 26). The maximum flux of wave energy from the tropopause to the stratosphere occurs near 60 degrees of latitude in both hemispheres (52). The poleward lobe of the NAM zonal wind anomaly dipole has large amplitude and vertical shear near the tropopause at the latitude of maximum upward planetary wave flux near 60 degrees of latitude (Fig. 3). Linear wave theory suggests that the upward propagation of planetary waves from the troposphere to the stratosphere is mediated by the static stability and vertical gradients of zonal wind near the tropopause (53). Limpasuvan and Hartmann (26) showed that the NAM significantly modulates the index of refraction for planetary waves near the subpolar tropopause in the region that Chen and Robinson (53) showed is critical for controlling the passage of planetary wave activity from the troposphere to the stratosphere. Other studies have shown important variations in planetary wave properties associated with annular variability (54, 55).

Fig. 3 shows zonal wind EP flux cross sections for those times during the Northern Hemisphere winter (December–February) when the NAM index based on sea-level pressure data is more than 1.5 standard deviations above (high index, poleward jet displacement) and below (equatorward jet displacement) its mean. Data used herein are from the 1958–1998 National Centers for Environmental Prediction/National Center for Atmospheric Research reanalysis (56). The EP flux vectors approximate the direction of wave activity flux, and the divergence indicates the magnitude of zonal flow acceleration produced by the waves (57). EP cross sections are shown for the sum of zonal wavenumbers 1 through 3. The longer waves 1 and 2 account for almost all of the flux above 200 hPa. The EP flux vectors in the upper troposphere and lower stratosphere bend more equatorward during high index periods, indicating that planetary waves are bent away from the polar vortex when the winds are shifted poleward and the vortex is strong. When the vortex is expanded and weaker in the stratosphere, waves are more likely to bend poleward. Because an equatorward-pointing EP flux vector corresponds to poleward meridional eddy momentum flux, the eddy westerly momentum flux anomalies are toward the pole during high index states and equatorward when the vortex is expanded. These momentum fluxes help to sustain a strong circumpolar jet in the high phase and to spread and weaken the polar vortex in the low phase. The contours of positive EP flux divergence in the high minus low difference plot indicate that the eddy flux anomalies drive westerly wind anomalies in the polar lower stratosphere when the winds there are stronger.

The upward flux of planetary wave activity is proportional to the poleward eddy heat flux. Table 1 shows the eddy heat and momentum fluxes by zonal waves 1 and 2 at 100 hPa during high and low NAM periods. The total heat flux is about the same during high and low NAM periods, but the flux shifts from wave 1 to wave 2 when the vortex is stronger. Wave 2 propagates upward and poleward less effectively than wave 1 (58), so that the shift in upward planetary wave flux from wave 1 to wave 2 may be important in explaining the stronger vortex associated with the high NAM polarity. The poleward momentum fluxes by both

Table 1. Poleward eddy heat and momentum fluxes at 100 hPa, area-averaged over the latitude belts from 45°N to the pole during the winter season for zonal wavenumbers 1 and 2

<table>
<thead>
<tr>
<th></th>
<th>DJF</th>
<th>High</th>
<th>Low</th>
<th>Hi-Lo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heat 1</td>
<td>115</td>
<td>94</td>
<td>145</td>
<td>-50</td>
</tr>
<tr>
<td>Heat 2</td>
<td>69</td>
<td>109</td>
<td>48</td>
<td>61</td>
</tr>
<tr>
<td>Momentum 1</td>
<td>-47</td>
<td>-29</td>
<td>-113</td>
<td>84</td>
</tr>
<tr>
<td>Momentum 2</td>
<td>-45</td>
<td>101</td>
<td>-75</td>
<td>176</td>
</tr>
</tbody>
</table>

First column is the seasonal mean for all data (DJF), followed by the 1.5 standard deviation high composite (High), low composite (Low), and the difference between them (Hi-Lo). Units are °K m s^-1 for heat flux and m^2 s^-2 for momentum flux.

Fig. 3. Composites for periods of high and low NAM index and their difference (Left, Center, and Right) during the December–March period. The zonal wind composites are on the top (units are m s^-1). The EP flux cross sections for the sum of zonal wavenumbers 1 to 3 are on the bottom. Positive contours are red.
wave 1 and wave 2 are much greater in the high NAM composite, and these fluxes act to strengthen the polar vortex.

Mid-Winter Stratospheric Warnings and the NAM. Because the NAM is associated with zonal wind variations at all levels during the winter season, one might expect an association between midwinter stratospheric warnings and a low NAM index. We define a midwinter warming as an occasion when the zonal wind averaged over the region poleward of 60°N and at or above the 70-hPa surface is easterly. We limit consideration to the December–February period to exclude spring vortex breakdowns from the sample. Fig. 4 is a histogram showing the number of days in which a midwinter warming is in progress versus the NAM SLP index for those days. Major warming days are more numerous when the NAM index is negative (227) than when it is positive (92) during the period 1958–1998. For days when the NAM index is 1.5 standard deviations below its mean value, there are 63 major warming days, versus only 4 when the daily NAM index is 1.5 standard deviations above its mean value. A highly significant relationship exists between the occurrence of midwinter stratospheric warmings and low values of NAM SLP index. This relationship is consistent with the recent trend toward high NAM index and infrequent sudden stratospheric warmings (44).

Can Trends in the Stratosphere Influence Surface Climate? From the above discussion, it is clear that the longitudinally symmetric variations in the troposphere and stratosphere are closely related. Much theoretical and observational evidence exists to support the notion that variability in the troposphere can drive variability in the stratosphere. The view that the stratosphere can have an important influence on events in the troposphere is less commonly held. Scientists studying the atmospheric response to volcanic eruptions were among the first to provide evidence that the strengthening of the stratospheric polar vortex could impact the climate at the earth’s surface. The vortex is observed to be abnormally cold and strong during the winter season after major volcanic eruptions (59–62). The stronger vortex may result from meridional gradients in heating associated with volcanic aerosols. Volcanic aerosols absorb solar radiation and heat the lower stratosphere, but this heating does not occur in the polar regions during the winter months when sunlight does not reach the pole. A uniformly distributed stratospheric aerosol would thus drive an enhanced equator-to-pole temperature decrease, leading to stronger zonal winds in the subpolar stratosphere during winter (63, 64). These enhanced westerly winds may duct planetary waves into the tropics, leading to a further dynamical cooling of the stratosphere. Major eruptions also tend to be followed, the next winter, by anomalously high values of the NAM index at the earth’s surface (65). Volcanic aerosols also affect the ozone concentration (66), which may leave a long-lasting effect in the lower stratosphere, where photochemical and transport time scales can be long compared with a season. Because volcanic aerosols exert only a small direct influence on the global troposphere, it has been argued that the apparent NAM response to the eruptions is induced by processes operating at stratospheric levels.

Further evidence that conditions in the stratosphere can influence the troposphere has been provided by Baldwin and Dunkerton (44), who used a 40-year National Centers for Environmental Prediction/National Center for Atmospheric Research reanalysis data set to show that annular anomalies at 10 hPa (~30 km in altitude) are best correlated with annular mode variations in the troposphere about 3 weeks later. This apparent time lag between events in the stratosphere and events in the troposphere has also been noted in other studies (67, 68). Downward propagation of stratospheric warming events into the troposphere also occurs for some modeled warming events (69). One modeling study has shown that the observed annular trends in the troposphere can only be produced in a model with good resolution of the stratosphere (19), while another has produced this trend in a model without a detailed representation of the stratosphere (70).

Several dynamical mechanisms by which the stratosphere can influence the troposphere can be postulated. Vertical propagation of planetary waves from the troposphere to the stratosphere is very sensitive to vertical wind shear near the tropopause (53), and this shear varies in synchrony with observed variations of the annular modes (25). The planetary wave, zonal flow interaction that leads to downward propagation of wind anomalies in the stratosphere (35, 71) could continue, perhaps in modified form, into the troposphere. A change in the zonally averaged potential vorticity distribution in the lower stratosphere can result from altered radiative forcing or wave, mean-flow interaction in the stratosphere. Hartley et al. (72) use potential vorticity inversion to argue that changes in zonal flow in the lower stratosphere should induce significant changes in the tropopause height and tropospheric flow. Baldwin and Dunkerton (44) argue that the redistribution of mass in the stratosphere may have enough influence on surface pressure to affect tropospheric flow. The annular modes of variability in the troposphere are free internal modes of variability driven by positive feedbacks between mean wind structure and eddy momentum fluxes (14, 26, 73–76). Because these modes exhibit large amplitude variability at low frequencies in the absence of applied forcing, it is possible that their mean state is sensitive to very modest forcing. Such forcing could come from the stratosphere. To summarize then, wave propagation, potential vorticity induction, and mass redistribution in the stratosphere can all influence the troposphere dynamically. Moreover, because the annular modes of the troposphere are free modes of variability, they may respond to modest stratospheric forcing with a significant change in tropospheric structure that is supported by strong dynamical feedbacks within the troposphere.

Recent stratospheric ozone reductions and GHG increases both cause significant radiative cooling in the stratosphere (77, 78). Annual mean responses to the changes of carbon dioxide, ozone, and stratospheric water vapor since 1979 each show about 2°C cooling in the polar lower stratosphere in a recent climate model simulation (78). It is possible, in principle, to induce a change in the NAM-index toward the high index state through radiative forcing provided by recent trends in O3, CO2, and H2O in the stratosphere. All of these composition trends appear to enhance the meridional temperature gradient across subpolar latitudes in the lower stratosphere when introduced into a climate model (78). These thermal gradients lead to increased vertical shear of the zonal wind and strengthening the polar vortex. These wind changes can, in turn, reduce the poleward

Fig. 4. Histogram of the number of days with major midwinter stratospheric warming conditions during December–March, 1958–1998, as a function of the daily value of the NAM index.
penetration of planetary waves and weaken the meridional circulation. Stratospheric ozone decline has produced a strong NAM-like wintertime response at the earth’s surface in one model (79, 80), but has been found to be relatively unimportant in other models (15, 19, 27). Ozone decline and GHG increases could also be contributing to the delay in the breakdown of the polar vortex in springtime.

It is observed that poleward ozone transport in the NH has declined during the last couple decades (48, 49), which is consistent with the stronger and more stable polar vortex observed in recent years. From arguments presented here it appears that decreased ozone in the polar lower stratosphere caused by reduced transport can lead to a stronger vortex and further reductions in planetary wave ozone transport. This positive feedback could sustain long-lasting anomalies that would be similar to the trends observed in the NH. Such a transport-feedback could also act to amplify any external forcing of the system, such as chlorine-induced ozone depletion or GHG effects on temperature.

That the temperature trends in the lower stratosphere resemble the response expected from the local radiative effects of ozone decline and GHG increases does not preclude the possibility that they could be attributable to other forcing mechanisms. Increasing concentrations of GHG induce surface warming as well as radiative cooling in the lower stratosphere. If tropical sea surface temperatures increase at the same time, tropical upper tropospheric warming will be amplified by the changes in the moist adiabatic lapse rate (81–83). The combined effect will be an enhanced meridional temperature gradient in the layer from 10–15 km above the surface, which is in the troposphere in the tropics and in the stratosphere in high latitudes. This temperature gradient will give rise to vertical shear of the zonal wind in the extratropical lower stratosphere and lead to the same positive wave propagation feedbacks described above for ozone depletion. The enhanced lower stratospheric cooling over the polar cap region renders the stratospheric ozone layer more susceptible to ozone-destroying chemicals (84, 85). Because of these dynamical, chemical, and radiative feedbacks, the ozone hole and greenhouse warming issues are much more strongly coupled than has widely been assumed.

Conclusion and Implications for the Climate of the 21st Century. Natural annular modes of internal variability in the troposphere and stratosphere are strongly linked by dynamical processes. These dominant modes of natural variability project strongly onto trends associated with both stratospheric ozone depletion and GHG forcing. Stratospheric ozone depletion and GHG warming may both be producing increased meridional temperature gradients in the extratropical lower stratosphere and upper troposphere, and thus acting synergistically to produce surprisingly large trends in both surface and stratospheric climate. The NAM-related wintertime climate trends and ozone losses have not been in evidence long enough to entirely rule out the possibility that they are a reflection of coupled atmosphere/ocean variability. However, it seems quite likely that they are at least in part human-induced.

If ozone-depleting chemicals are primarily responsible not only for the thinning of the stratospheric ozone layer but also for the trends in the NAM and SAM indices, then the production curbs agreed to in the Montreal Protocol should slow these trends within the next decade or two. On the other hand, if the trends in the indices are primarily due to the buildup of GHG, as suggested by experiments with some climate models (19, 70), then the long-term prognosis is less clear. In the presence of a continuing (GHG-induced) trend toward a colder polar stratosphere, it might be necessary to impose stricter curbs on the use of ozone destroying chemicals to keep ozone losses within specified bounds.

Predicting the future course of NAM-related climate change is further complicated by the possibility of a transition of the Arctic Ocean to an ice free state during the 21st century. The tendency toward a more westerly (counterclockwise, or cyclonic) circulation within and around the polar cap region may be contributing to the recent retreat of the Arctic pack ice during summertime as well as the thinning of the perennial pack ice (86–89). A more cyclonic wind stress, as occurs in the high index polarity of the NAM, favors divergence of ice and surface water out of the Arctic, opening up leads, and thinning the layer of cold, fresh water that insulates the pack ice from the warmer, saltier waters underneath (22). Another possible complicating factor is the oceanic thermohaline circulation, whose sinking branch lies along the edge of the pack ice in the far reaches of the North Atlantic. In recent years, the conditions that favor bottom water formation (i.e., strong outflows of cold, dry air across the ice edge) have been observed less frequently over the Greenland Sea and more frequently over the Labrador Sea, suggesting that bottom water formation has been occurring farther westward (90, 91). Whether such a shift will, in time, serve to change the intensity or basic character of the thermohaline circulation has yet to be determined. There are indications that the latitude of the north wall of the Gulf Stream has shifted northward slightly in recent decades in response to the trend in the NAM (92). A continued northward shift would favor additional warming over Eurasia and the Arctic, above and beyond that associated with the trend in the NAM itself. Coupled interactions such as these could conceivably give rise to NAM-related variability on time scales much longer than the historical record.

This work was supported by National Science Foundation grants ATM-9873691, ATM-9322480, and ATM-9707069 and by the National Oceanic and Atmospheric Administration through Cooperative Agreement NA67RJ0155, Joint Institute for the Study of the Atmosphere and Ocean (JISAO) Task 2, and through its support of the Hayes Center (JISAO contribution no. 738).
