6.1. Introduction

The stratosphere over Antarctica is one of the most inaccessible places on the planet. During the antarctic winter, it extends from about 8 to 55 km above the surface, has temperatures colder than $-90^\circ$C, and winds that are greater than 100 m s$^{-1}$. Yet even this terribly remote and hostile region has felt man’s impact. The antarctic ozone hole is a clear example of how our industrial society can affect the atmosphere even in this remote corner of the earth. The tremendous ozone losses over Antarctica observed each spring have ultimately resulted from man-made chlorine compounds, and these ozone losses have led to increased levels of biologically harmful ultraviolet radiation at the earth’s surface. Understanding the meteorology of the southern stratosphere is the key to our understanding of the antarctic ozone hole.

Knowledge and understanding of the stratosphere in the Southern Hemisphere has increased considerably since appearance of the first monograph on meteorology of the Southern Hemisphere in 1972. The analyses of Labitzke and van Loon (1972) in that publication were original and insightful in many aspects but necessarily hampered by lack of observational data (being based on a single year of monthly mean rawinsonde observations in the lower stratosphere, together with one year of satellite radiance measurements). Over the last 25 years, substantial improvements in the observational database in the SH stratosphere have occurred, due to several factors.

1) The refinement and continued availability of global satellite measurements, leading to daily meteorological analyses of stratospheric circulation.

2) The development of stratospheric data-assimilation systems, based on a blending of observations with global models.

3) Appearance of the ozone hole over Antarctica in the middle 1980s, which prompted observational aircraft campaigns and detailed studies of SH stratospheric dynamics and chemistry.

4) Analyses of stratospheric chemical constituent data from the limb infrared monitor of the stratosphere (LIMS), the total ozone mapping spectrometer (TOMS), the Stratospheric Aerosol and Gas Experiment (SAGE), and the Upper Atmosphere Research Satellite (UARS). In particular, UARS provided a wide range of simultaneous measurements of key constituents tied to stratospheric transport and ozone photochemistry.

Observations of dynamical and constituent behavior from this rich database have fueled parallel leaps in theoretical understanding of stratospheric dynamics, radiation, chemistry, and constituent transport (cf. McIntyre 1992; Holton et al. 1995).

The objective of this chapter is to give an overview of several aspects of the SH stratosphere, based primarily on the long record of observational data now available. Part of our focus is on contrasting behavior in the SH versus that in the (somewhat better-known) NH stratosphere; such contrasts often lead to better conceptual understanding of the SH itself. The general circulation of the SH stratosphere is considered in section 6.3, with focus on the seasonality of the zonal-mean fields and relationships between the mean flow, eddies, and radiative forcing. Daily variability of the stratosphere is discussed in section 6.4, including stratospheric warmings, transient and traveling planetary waves, and inferred smaller-scale features. Interannual variability of the SH stratosphere is analyzed in section 6.5, while section 6.6 focuses on global ozone characteristics and the antarctic ozone hole.

6.2. Data

A majority of the meteorological data presented here are derived from daily stratospheric analyses from NCEP (formerly called the National Meteorological Center). Daily global analyses of stratospheric
geopotential height and temperature have been made since October 1978 for pressure levels covering 1000–
0.4 mb (approximately 0–55 km). Lower-level data
(over 1000–100 mb) are output of the Global Data
Assimilation System used at NCEP. Stratospheric
level data (at 70, 50, 30, 10, 5, 2, 1, and 0.4 mb) are
daily operational analyses produced at the Climate
Prediction Center of NCEP. The stratospheric analyses
are produced using available radiosonde observations
(only in the NH) and satellite-derived temperature
retrievals in a modified successive correction (Cress-
man) analyses (Gelman and Nagatani 1977). Details of
these data may be found in Gelman et al. (1986),
Randel (1992), and Finger et al. (1993). Zonal-mean
and eddy horizontal winds are derived from the geopo-
tential data using balance assumptions (details dis-

We also use stratospheric meteorological analyses
from the United Kingdom Meteorological Office
(UKMO) stratospheric data- assimilation system
(Swinbank and O'Neill 1994). This system uses a
global numerical model of the atmosphere, with fields
continuously adjusted toward available wind and tem-
perature observations as the model is integrated for-
ward in time. Output of these analyses includes tem-
peratures and three-dimensional winds, spanning the
altitude range 1000–0.3 mb (0–57 km). Data are
available from November 1991 to present. Compari-
sions between circulation statistics from NCEP-derived
products and the UKMO assimilation show reasonable
agreement (e.g., Manney et al. 1996); notable differ-
ces include an underestimate of the tropical quasi-
biennial oscillation (QBO) in the NCEP analyses
(compared to rawinsonde observations) and a some-
what colder SH polar vortex in the UKMO data.

6.3. Observed general circulation

A simplified perspective of the stratospheric general
circulation is provided by analysis of the structure and
seasonal variation of the zonally averaged flow. It is
straightforward to separate the governing equations
(for momentum, thermodynamic energy, and mass
conservation) into zonal-mean and eddy components,
yielding the so-called Eulerian-mean equations (see
Andrews et al. 1987, section 3.3). Interpretation of the
resulting equations in terms of physically separating
mean and eddy effects is often confusing, however,
because the Eulerian-mean meridional circulation is
not independent from the Eulerian eddy fluxes (for
example, there is strong cancellation between the eddy
heat-flux convergence and adiabatic cooling associated
with \( \bar{w} \)). In fact, in the idealized case where wave
transience and dissipation is neglected, the eddy and
mean circulation terms exactly cancel, producing zero
net forcing of the zonally averaged flow (Andrews and
McIntyre 1978; Boyd 1976). An alternative and con-
ceptually more physical separation of zonal-mean and
eddy quantities is provided by the so-called trans-
formed Eulerian-mean (TEM) budget equations [written
here based on quasigeostrophic scaling assumptions,
following Andrews et al. (1987, section 7.2)].

\[
\frac{\partial \bar{u}}{\partial t} - 2\Omega \sin \phi \bar{v} = \bar{G} \quad \text{(zonal momentum balance)}
\]

(6.1)

\[
\frac{\partial \bar{T}}{\partial t} + \bar{w} \left( \frac{N^2 H}{R} \right) = \bar{Q} \quad \text{(thermodynamic balance)}
\]

(6.2)

\[
\frac{1}{a \cos \phi} \frac{\partial}{\partial \phi} \left( \bar{w} \frac{v^* \cos \phi}{a} \right) + \frac{1}{\rho_0} \frac{\partial}{\partial z} \left( \rho_0 \bar{w}^* \right)
\]

\[= 0 \quad \text{(mass continuity)}
\]

(6.3)

\[
2\Omega \sin \phi \frac{\partial \bar{u}}{\partial z} + \frac{R}{H} \frac{1}{\partial \phi} \frac{\partial \bar{T}}{\partial \phi} = 0 \quad \text{(geostrophic thermal wind)}
\]

(6.4)

Notation is standard and defined in the Appendix. This
set of TEM equations results from defining trans-
formed mean meridional circulation components \( \bar{v}^*, \bar{w}^* \) as

\[
\bar{v}^* = \bar{v} - \rho_0 \frac{R}{H} \frac{\partial}{\partial z} \left( \frac{\bar{v}^* T^*}{N^2} \right)
\]

\[
\bar{w}^* = \bar{w} + \frac{R}{H} \frac{1}{\partial \phi} \frac{\partial}{\partial \phi} \left( \frac{\bar{v}^* T^*}{N^2} \right)
\]

The terms on the right-hand side describe the mean
circulation components associated with eddy heat-flux
divergence. The components \( \bar{v}^*, \bar{w}^* \) then represent
the difference between the Eulerian means \( \bar{v}, \bar{w} \)
and these eddy-associated terms; this \( \bar{v}^*, \bar{w}^* \)
circulation is hence termed the residual mean circulation. This
residual circulation is more directly linked with diabatic
processes that determine the mean mass flow in the
stratosphere and provides a concise view of transport
in the meridional plane.

From this TEM perspective, changes in the zonal-
mean flow are driven by radiative forcing \( \bar{G} \) of
the thermodynamic field, together with wave driving of
the zonal (or angular) momentum \( \bar{G} \). The ther-
modynamic and zonal momentum fields are in turn
coupled via thermal wind balance, and the residual
circulation \( \bar{v}^*, \bar{w}^* \) acts to redistribute the radiative and wave
forcing effects nonlocally, in such a manner so as to
maintain thermal wind balance. The wave driving term
\( \bar{G} \) is often partitioned into components associated
with large-scale planetary waves (the so-called
Eliassen–Palm (or EP) flux divergence, \( \rho_0 \bar{v} \cdot \nabla \cdot F \) and
those attributable to smaller or unresolved scales ($\tilde{X}$) (e.g., gravity waves):

$$\tilde{G} = \rho_0^{-1} \mathbf{\nabla} \cdot \mathbf{F} + \tilde{X}.$$  

Components of the quasigeostrophic EP flux are given by

$$\mathbf{F} = (F_\phi, F_z)$$

$$= \rho_0 a \cos \phi \left( -u^T \nu', 2\Omega \sin \phi \frac{R}{H} \nu^T \frac{T^z}{N^2} \right),$$  \hspace{1cm} (6.5)  

and estimates of the planetary-wave EP flux and its divergence may be made using stratospheric meteorological analyses. The structure of the stratospheric mean flow, its seasonal evolution, and the fundamental asymmetries between the NH and SH can be understood to first order by analyses of this set of equations [Eqs. (6.1)–(6.4)], using observations and/or models to derive the radiative heating and wave-driving distributions.

As a note, the full primitive equation form of the TEM equations (Andrews et al. 1987, section 3.5) includes an additional component in $F_z$ proportional to the vertical eddy momentum flux $\tilde{w}'u'$. Although it is negligible for large-scale motions, it is the dominant component for gravity-wave forcing of the mean flow. Gravity-wave forcing (which is of considerable importance for the SH stratosphere, as discussed later) may be calculated directly via this covariance term, or included (using a parameterized expression) in the $X$ component of $\tilde{G}$.

\textbf{a. Zonal-mean temperature}

The observed zonal-mean temperature over 0–50 km during January, April, July, and October is shown in Fig. 6.1, illustrating the overall temperature structure of the troposphere and stratosphere, together with its seasonal cycle. These and the following climatological means are based on NCEP data averaged over 1990–95. The heavy dashed line in each panel in Fig. 6.1 denotes the approximate position of the tropopause, as defined by a strong vertical gradient of the static stability parameter $N^2$. The stratosphere is characterized by increasing temperature with altitude above the tropopause, maximizing at the stratopause near 50 km (the upper limit of the data shown in Fig. 6.1). The temperature increase with altitude (i.e., the existence of the stratosphere) is due to the absorption of solar ultraviolet radiation by ozone, the structure of which is strongly latitudinally and seasonally dependent (following the solar declination). Radiative cooling in the stratosphere occurs mainly via infrared emission by carbon dioxide, which is principally a function of stratospheric temperature and less strongly dependent on latitude. The net (total) radiative heating is thus latitudinally and seasonally dependent, resulting in a polar stratosphere that is warm relative to low midlatitudes in SH summer (January) and relatively cold throughout the rest of the year. The cold polar stratosphere is particularly evident in midwinter (July), due to the complete absence of high-latitude heating during polar night.

\textbf{b. Zonal-mean winds}

Figure 6.2 shows the zonal-mean zonal wind for January, April, July, and October. There is a strong easterly jet in the SH summer (January) stratosphere over low latitudes, with winds near $-50$ m s$^{-1}$, which is somewhat stronger than the NH summer easterly jet in July (maximum near $-40$ m s$^{-1}$). Throughout the rest of the year, strong westerly winds cover middle and high latitudes of the SH stratosphere, associated (via thermal wind balance) with the cold polar stratosphere seen in Fig. 6.1. This strong wind maximum is termed the polar night jet, and the associated circulation system is termed the stratospheric polar vortex. Note that the SH winter polar night jet is much stronger than that observed in the NH winter, and the SH winter polar temperatures are much colder than those in the NH; the reason for this asymmetry is discussed below. The seasonal evolution of the SH polar night jet/polar vortex follows the pattern of appearing first in the upper stratosphere in early fall (April), strengthening and descending with time until middle winter, and then disappearing from the top down in spring. These radiative-driven changes occur first in the upper stratosphere because the radiative relaxation time there (of order 5 days) is much faster than that in the lower stratosphere (of order 50 days) (e.g., Gille and Lyjak 1987). This smooth seasonal evolution is interrupted by episodic large-scale disturbances (planetary waves) that originate in the troposphere, propagate vertically, and interact with the mean flow (these are associated with so-called stratospheric warming events). The transition from spring westerlies to summer easterlies in the stratosphere (termed breakdown of the vortex or the final warming) is often accompanied by one of these transient wave events (Newman 1986; Mechoso et al. 1988). The SH vortex breakdown is delayed until very late spring, significantly later than in the NH (as shown further below).

\textbf{c. Wave driving and residual circulation}

A concise diagnostic for the propagation characteristics and mean-flow forcing of large-scale planetary waves is provided by EP flux cross sections (Edmon et al. 1980; Andrews et al. 1987). Figure 6.3 shows climatological average EP flux diagrams for January, April, July, and October. In these diagrams, arrows show components of $F_\phi, F_z$ [Eq. (6.5)], denoting the direction and magnitude of planetary-wave propagation in the meridional plane, and contours show the EP
flux divergence ($\rho C_T \nabla \cdot \mathbf{F}$), which is the planetary-wave component of the wave driving $\tilde{G}$ in Eq. (6.1). The majority of tropospheric eddy activity (predominantly eastward-traveling, medium-scale zonal waves 4–7; Randel and Held 1991) does not propagate into the stratosphere but rather propagates equatorward or is absorbed in the middle upper troposphere (as evidenced by the maxima in wave driving in that region and strong decrease across the tropopause in Fig. 6.3). This lack of propagation of the tropospheric waves into the stratosphere is due to the background wind and thermal structures acting to restrict vertical propagation for all but the largest planetary scales (zonal waves 1–2) during winter (Charney and Drazin 1961; Dickinson 1969). Note that the tropospheric eddies do contribute a component of drag somewhat above the tropopause in Fig. 6.3, and that the residual circulation associated with tropospheric waves extends nonlocally, so that their influence is very important for behavior of the lower stratosphere (Dunkerton 1988).

The amount of wave activity in the stratosphere in Fig. 6.3 exhibits a strong seasonal cycle, with near-zero wave driving by planetary waves in the summer hemisphere (due to exclusion of vertical propagation in summer easterlies—see Fig. 6.2). The wave activity that is observed to propagate into the stratosphere is predominantly of planetary scale, although a rich spectrum of smaller scales are generated in the stratosphere by nonlinear and mixing processes (as described below). There is a strong hemispheric asymmetry in the observed planetary-scale wave driving, as EP fluxes during midwinter in the NH are much stronger than the corresponding time in the SH (compare the January and July patterns in Fig. 6.3). This is ultimately due to the stronger generation of planetary-scale waves in the NH troposphere due to orographic and
Fig. 6.2. Climatological zonal-mean zonal wind in January, April, July, and October. Contour interval is 10 m s\(^{-1}\) and negative values are shaded.

thermal forcing, which is absent in the SH. During spring, stratospheric EP fluxes are larger in the SH than in the NH, so that the seasonal cycle is markedly different between the hemispheres. This hemispheric difference in planetary-wave driving is the primary cause of the differences in wind, temperature, and polar vortex structures between the NH and SH.

The absorption of vertically propagating gravity waves (and the associated wave driving contribution to \(\vec{G}\) in Eq. (6.1)) also exerts a strong influence on the circulation of the middle atmosphere. Gravity waves generated in the troposphere (by topographic forcing, convection, or frontogenesis, for example) propagate vertically and grow exponentially with height, until they reach an altitude where their horizontal phase speed equals the background winds (a critical level) or their amplitudes become so large that they become convectively unstable and "break" (leading to turbulence, small-scale mixing, and dissipation; see Prusa et al. 1996). This breaking typically occurs at mesospheric altitudes and is associated with a convergence of the waves' vertical momentum flux and acceleration of the background mean flow (tending to drag the mean flow toward the phase speed of the waves). The presence of critical lines in the troposphere and stratosphere sharply limits the phase speeds of waves that reach the mesosphere, however; such filtering effects lead to a strong gravity-wave drag in the mesosphere (of order 50–100 m s\(^{-1}\) per day) in both summer and winter (Lindzen 1981; Holton 1983). This momentum deposition is balanced by an intensified mean meridional circulation in the mesosphere (toward the winter pole), and the associated vertical circulations explain the remarkably cold summer and warm winter polar mesopause (~90 km) temperatures, which are both far from radiative equilibrium (see Andrews et al. 1987, section 7.2).
Although gravity waves are primarily absorbed in the mesosphere, their influence extends well down into the stratosphere (via the nonlocal residual circulation, or "downward control"; see Haynes et al. 1991). Garcia and Boville (1994) have studied this in some detail, using a numerical model that includes parameterizations of both planetary- and gravity-wave breaking. Their results demonstrate that mesospheric gravity-wave drag can significantly affect the temperature in the polar winter stratosphere to altitudes as low as 20 km. The effect is more important when planetary-wave driving is relatively weak, as in the SH during midwinter. An important consequence of this result is that an inadequate amount of parameterized gravity-wave drag in numerical models results in a too-cold, intense polar vortex (the so-called "cold pole" problem—e.g., Mahlman and Umscheid 1987 and Boville 1995). Hitchman et al. (1989) discuss similar gravity-wave influences on variations of the stratopause in the SH.

An important component of the global stratospheric flow is the residual mean meridional circulation ($v^*$, $w^*$), whose structure and seasonal evolution is shown in Fig. 6.4. Components of this circulation are calculated from the coupled thermodynamic and continuity equations [Eq. (6.2)–(6.3)], using accurate radiative heating calculations to estimate net radiative forcing (Shine 1987; Rosenfeld et al. 1987; Olaguer et al. 1992; Rosenlof 1995). The important effects of clouds and latent heat transfer preclude such straightforward calculations for the troposphere, and only the inferred stratospheric circulation is shown in Fig. 6.4. The overall stratospheric circulation is upward in the Tropics and downward in winter high latitudes, with strong flow from the summer to winter hemisphere in the upper stratosphere (see Dunkerton 1978; Garcia 1987).
There is a strong seasonal cycle and hemispheric asymmetry observed in the residual circulation, with more intense poleward and downward motion observed during midwinter in the NH than in the SH winter. The upward branch of this circulation in the Tropics also varies seasonally, with upward motion centered in the respective summer subtropics (stronger during NH winter); this circulation is associated with the observed annual cycle in tropical stratospheric temperature and trace constituents (Yulaeva et al. 1994; Rosenlof 1995). Seasonality in the residual circulation is forced by seasonal variations in both radiative forcing ($\tilde{Q}$) and wave driving ($\tilde{G}$), but the strong circulations during winter (and spring in the SH) are attributable primarily to wave forcing (Andrews et al. 1987, section 7.2; Haynes et al. 1991; Holton et al. 1995; see also Garcia and Boville 1994). During these times, the wave driving acts to push the stratosphere away from radiative equilibrium, and the radiative forcing is largely a response to this departure from equilibrium (since $Q$ is primarily a function of temperature). During equinox seasons, the radiative forcing is comparably more important; radiative cooling over the polar cap in autumn (when wave driving is small) leads to a buildup of the meridional pressure gradient and development of the polar jet through the lowering of mass surfaces in polar latitudes relative to lower latitudes.

The large NH–SH differences in residual circulation seen in Fig. 6.4 are due principally to the observed NH–SH differences in wave driving $\tilde{G}$ (see Fig. 6.3), as seen from the steady-state version of Eq. (1), $f \cdot \mathbf{v} = \tilde{G}$. Thus, the weaker midwinter residual circulation in the SH is directly related to the smaller SH wave forcing. This weaker circulation is also consistent with the colder and more intense SH polar vortex (Figs. 6.1–6.2), in that the residual circulation acts to maintain departures from radiative equilibrium (the stron-
ger NH circulation is associated with warmer NH polar temperatures. This latter point is illustrated in the seasonal march of temperatures and residual circulation vertical velocity at the North and South Poles, respectively, shown in Figs. 6.5–6.6. Note that the time axis has been shifted by 6 months between the NH and SH in these plots, to facilitate direct comparisons. These figures clearly show that the much colder SH middle and lower polar stratosphere is associated with a weaker downward circulation in midwinter. A similar comparison of springtime conditions shows stronger downward motion in the SH, in balance with the larger SH wave driving.

The coupled seasonal variations of zonal wind in the middle stratosphere (10 mb), the large-scale wave forcing into the lower stratosphere, and residual mean vertical velocity at 50 mb is illustrated in Fig. 6.7. Here, the 50-mb heat flux is used as a measure of wave activity entering the stratosphere [this is proportional to the vertical component of the EP flux $F_E$—Eq. (6.5)]. [We note that heat flux is a derived quantity whose estimation is sensitive to details of the meteo-

![Temperature North Pole](image1)

![Temperature South Pole](image2)

**Fig. 6.5.** Climatological height-time sections of zonal-mean temperature at the North Pole (top) and South Pole (bottom). Values below 200 K are shaded. Note that the time axes are shifted by six months between the NH and SH.

![Vertical velocity 80 N](image3)

**Fig. 6.6.** Time traces of residual mean vertical velocity ($\bar{w}$) at 80°N (top) and 80°S (bottom), for the 50-, 10-, and 2-mb levels. The left axis refers to the 50- and 10-mb levels, and the right axis is for the 2-mb level. Note the time axes are shifted by six months between the NH and SH.
with downward velocities inside the SH polar vortex inferred from trace constituent observations by Schoeberl et al. 1995; see also Rosenfield et al. 1994). Note also the seasonality in the tropical upwelling in Fig. 6.7, as discussed above. These variations are consistent [in balance, via Eqs. (6.1)–(6.4)] with the colder, more intense SH midwinter polar vortex.

d. Trace constituent transport

Much has been learned about the overall mass flow in the stratosphere by analysis of quasi-conservative chemical tracers. The evolution of such tracers depends on the distribution of sources and sinks, and the relative timescales for transport versus chemical processes. The overall structure of species with long photochemical lifetimes (long compared to transport timescales of weeks to months, such as ozone in the lower stratosphere, nitrous oxide, or methane) is strongly influenced by and bears the signature of transport. Conversely, species with short lifetimes (such as ozone in the upper stratosphere) adjust quickly to the ambient background and are not sensitive to their transport history.

A useful species with a reasonably long lifetime to observe transport effects is methane (CH₄). Methane is generated in the troposphere by biotic activity, is transported into the stratosphere in the Tropics (by the general circulation), and oxidized in the middle and upper stratosphere (yielding a source of stratospheric water vapor). The photochemical lifetime of methane below 45 km is >100 days, so that the stratospheric distribution is determined mainly by the circulation.

Methane measurements from the stratospheric and mesospheric sounder have been used in a number of studies to deduce stratospheric circulation features (Jones and Pyle 1984; Solomon et al. 1986b; Holton and Choi 1988; Stanford et al. 1993). The overall characteristics of methane are similar to other constituents with tropospheric sources and stratospheric photochemical sinks, namely nitrous oxide (N₂O) and the chlorofluorocarbons.

Figure 6.8 shows meridional cross sections of stratospheric methane throughout the seasonal cycle derived from the Halogen Occultation Experiment (HALOE; Russell et al. 1993a) and cryogenic limb array etalon spectrometer (CLAES; Kumer et al. 1994) instruments on UARS. These data were organized according to the background potential vorticity fields (using so-called equivalent latitude mapping): results over most latitudes are derived from HALOE, and CLAES data are used to fill in polar regions (Randel et al. 1998). The methane contours in Fig. 6.8 bulge upward near the Equator, reflecting the mean upward circulation in the Tropics (see Fig. 6.4). The tropical peak moves latitudinally with season, due to seasonal changes in the transport circulation (compare the January and July panels in Figs. 6.4 and 6.8); there is also

Fig. 6.7. Climatological latitude–time sections of (top) zonal-mean zonal wind at 10 mb (contour interval of 10 m s⁻¹); (middle) zonal-average eddy heat flux at 50 mb (contours of ±2, 6, 10, ... K·m s⁻¹); and (bottom) residual mean vertical velocity at 50 mb (contour interval of 0.1 mm s⁻¹).
a double-peaked tropical pattern in April related to the tropical SAO. Enhanced meridional gradients in methane are observed in the subtropics and across the polar vortex regions; these stronger gradients denote regions of minimal horizontal mixing and are sometimes referred to as the subtropical and polar vortex "mixing barriers." The SH winter–spring polar vortex is particularly evident in these data, with remarkably strong methane gradients (near 60°S) resulting from strong downward flow inside the vortex and small horizontal transport across the vortex edge (Hartmann et al. 1989; Schoeberl et al. 1992; Fisher et al. 1993; Manney et al. 1994; Sutton 1994). In contrast, the methane isolines are relatively flat over winter midlatitudes, due to rapid horizontal mixing associated with the midlatitude "surf zone" (McIntyre and Palmer 1983, 1984; Leovy et al. 1985).

Recent simulations of the material flow and transport in the stratosphere have shown behavior very similar to these global observations. Figure 6.9 shows an example of such structure in the SH stratosphere, derived from a seasonal cycle simulation using a three-dimensional transport model (from Sutton 1994; similar results are shown in Fisher et al. 1993; Manney et al. 1994; and Eluszkiewicz et al. 1995). Figure 6.9 shows the locations of particles during the middle of SH winter (30 June), after horizontally stratified initialization three months earlier (1 April) at the locations denoted on the right in Fig. 6.9. Evolution of the particles shows upward motion in the Tropics and poleward–downward motion in the SH extratropics. Particles inside the vortex (poleward of 60°S) remain isolated, due to minimal transport across the vortex edge; conversely, strong mixing is observed in the midlatitude surf-zone region (~20°–50°S). The combination of strong descent and weak mixing in the vortex results in vortex interior air exhibiting characteristics of the mesosphere (as observed by Russell et
Fig. 6.9. Model-generated pattern of air parcels during the middle of SH winter (30 June), three months after horizontally stratified initialization (on 1 April) at the altitudes denoted on the right side (from Sutton 1994). Note the similarity to the SH winter methane patterns shown in Fig. 6.8.

Overall, the modeled structure is remarkably similar in many aspects to that observed (i.e., the July methane patterns shown in Fig. 6.8), and such modeling allows detailed studies of the transport mechanisms that lead to the observed structure.

e. Potential vorticity perspective

In addition to chemical tracers of fluid motion, dynamical tracers are often used to study circulation and transport. Dynamical tracers consist of quasiconservative field variables derived from conventional meteorological analyses, and two important dynamical tracers are potential temperature and potential vorticity. Potential temperature $\theta$ is defined as the temperature a parcel of dry air at temperature $T$ would acquire if expanded or compressed adiabatically to a reference pressure $p_s = 1000$ mb:

$$\theta = T \left( \frac{p_s}{p} \right)^{2/7}.$$  \hspace{1cm} (6.6)

For atmospheric motions that are adiabatic, fluid parcels remain on constant $\theta$ (isentropic) surfaces. Because potential temperature is mainly a function of altitude, $\theta$ may be regarded as a vertical coordinate for fluid parcels; for adiabatic conditions, three-dimensional motion is reduced to two-dimensional flow on isentropic surfaces. In the lower and middle stratosphere, diabatic heating rates are typically less than 1 K day$^{-1}$, making an adiabatic approximation valid for timescales of 1–2 weeks. Thus, tracer distributions and evolution on such timescales are often analyzed on isentropic surfaces.

Potential vorticity is a derived dynamical quantity that is materially conserved for adiabatic frictionless flow. Mathematically, $PV$ is defined as the dot product of the absolute vorticity vector and the potential temperature gradient. For the stratosphere, this can be approximated as

$$PV = -g(\zeta + 2\Omega \sin \phi) \frac{\partial \theta}{\partial p},$$  \hspace{1cm} (6.7)

where $\zeta$ is the (local) vertical component of the relative vorticity derived from the horizontal wind field. The motivation behind PV conservation is discussed at length in McIntyre and Palmer (1983, 1984) and Hoskins et al. (1985) and references therein; PV conservation is the fluid-dynamical generalization of angular momentum conservation in Newtonian mechanics. For adiabatic flow (as noted above, a good approximation for timescales up to several weeks in the lower-middle stratosphere), contours of PV behave as material lines and allow tracing of fluid motions. These ideas are confirmed by the observed good agreement between derived PV and chemical tracer fields for transient events (e.g., Harnett et al. 1989; Leovy et al. 1985; Randel et al. 1993; Manney et al. 1995).

A potential vorticity perspective of the time-mean SH stratospheric structure and seasonal evolution is shown in Fig. 6.10, for the 465 and 840 K isentropic levels. These isentropic levels are near 20 and 30 km, respectively, representing the lower and middle stratosphere. The winter polar vortex is associated with strong horizontal PV gradients; air inside the vortex (at polar latitudes) has PV values that are much higher than those outside, signifying distinctive air masses. The region of strong gradients is associated with the “edge” of the polar vortex and is near the locus of maximum zonal winds. The vortex edge may be objectively defined as the location of the strongest latitudinal gradient of PV (Nash et al. 1996), and the heavy contours in Fig. 6.10 denote the respective vortex edges for each month. Due to PV conservation discussed above, these strong PV gradients present a barrier to north–south mass flux across the vortex edge; the associated restoring mechanism is responsible for oscillatory behavior for large-scale vortex perturbations (planetary Rossby waves). Detailed transport calculations (Bowman 1993a,b; Chen et al. 1994; Chen 1994; Manney et al. 1994) confirm small net mass flux from inside to outside the vortex and vice-versa (discussed later in more detail). The mid-latitude region outside the winter vortex exhibits relatively weaker PV gradients; this is the so-called “surf zone” (McIntyre and Palmer 1984), where north–south transport and mixing is less restricted (and readily observed; see section 6.4). Figure 6.10 also illustrates...
the early winter development of the vortex in the middle stratosphere and the persistence of a strong lower-stratospheric vortex well into SH spring.

For comparison, Fig. 6.11 shows the time-mean PV structure in the NH winter (January) at 465 and 840 K. The NH vortex is less circular and displaced farther off the pole than that in the SH. Such time-mean asymmetries are associated with stationary planetary waves, which originate in the troposphere due to thermal and orographic forcings (and are thus significantly larger in the NH). Stationary waves are comparatively weaker in the SH, the only evidence in the time-average PV fields being in October at 840 K (Fig. 6.10). Comparison of the PV fields in Figs. 6.10–6.11 demonstrate that the SH polar vortex is substantially stronger in midwinter than in the NH (as seen also in the zonal winds in Figs. 6.2 and 6.7). The polar vortex is also much more persistent and robust in the SH, remaining intact well into spring (October); the NH vortex is completely gone by the corresponding time (April—see Fig. 6.7).

f. Climatology of planetary-wave activity

A climatological description of the SH stratosphere should include characterization of the observed planetary-wave variability. This includes zonal scale of the waves and partitioning between stationary and transient components. Climatological characteristics of planetary waves in the SH stratosphere based on satellite observations have been discussed by Harwood (1975), Hartmann (1976), Leovy and Webster (1976), Mechosso and Hartmann (1982), Hartmann et al. (1984), Vienne and Stanford (1982), and Randel (1987, 1988). Details and references on specific traveling waves are discussed in section 6.4; here we show climatological wave spectra based on the 1990–95 NCEP database. There are several possible quantities from which to quantify wave variability: geopotential height, kinetic energy, heat or momentum fluxes, potential vorticity, etc. The planetary-wave signatures in all of these quantities turn out to be qualitatively similar, and we focus here on eddy heat flux at 50 mb (a measure of eddy activity entering the lower stratosphere; see Fig. 6.7).

Figure 6.12 shows latitude–time sections of the 50-mb eddy heat flux (as in Fig. 6.7), separated according to stationary and transient components. These estimates are calculated simply as the time averages of the individual years' stationary (monthly mean) and transient components; transients include both zonally propagating and spatially fixed but time-varying waves. This separation shows that both sta-
Stationary wave heat flux 50mb

Transient wave heat flux 50mb

Fig. 6.12. Climatological 50-mb eddy heat flux, separated according to stationary (top) and transient (bottom) wave components. Contour interval is 2 K m s$^{-1}$, with zero contours omitted.

The stationary component in the SH shows a weak maximum in fall (May), a relative minimum in midwinter, and a large amplitude maximum in spring (September–October). Climatological means clearly show this double-peaked seasonal cycle throughout the stratosphere (e.g., Geller and Wu 1987; Randel 1988), and the midwinter minimum in wave activity has been discussed theoretically by Plumb (1989). The stationary waves in the SH are almost completely associated with zonal wave 1, whereas stationary waves 1–2 are both important in the NH. It is interesting to note that although stationary-wave seasonality and structure in the NH stratosphere is reasonably well understood in terms of forcing from the troposphere and variations in background flow (e.g., Matsuno 1970; Rong-ju and Gambo 1982; Karoly and Hoskins 1982), there is less detailed understanding for the SH stationary waves.

To quantify transient wave behavior, Fig. 6.13 shows zonal wavenumber–frequency co-spectra of transient eddy heat flux at 50 mb, 60°S, for data during three individual years (1991, 1993, and 1995) (calculation details are as in Mechoso and Hartmann).

Fig. 6.13. Space–time co-spectra of transient eddy heat flux at 50 mb, 60°S, for July–October statistics during 1991 (top), 1993 (middle), and 1995 (bottom). Spectra are shown for zonal waves 1–4, separately. Numerical values at right denote the total transient/stationary fluxes (units of K·m s$^{-1}$) and their sums for each year.
1982, or Randel and Held 1991). Statistics are calculated from data covering July–October from each year. Also included are numerical values summing the total transient and stationary components for comparison. There is a substantial amount of variability in details of the spectra from year to year, although the overall character is similar, showing eastward-traveling wave 2 and stationary wave 1 contributing the largest components. The overall heat flux is smaller in 1995, due mainly to relatively weak transient waves. The transient wave 2 spectra show the majority of covariance traveling eastward with periods of 8–25 days (corresponding to zonal phase speeds of 5–10 m s$^{-1}$); a relatively sharp spectral peak with period near 17 days is seen in both 1991 and 1993. This eastward-traveling wave 2 is a commonly observed feature in the SH stratosphere (Harwood 1975; Hartmann 1976; Leovy and Webster 1976; Mecho and Hartmann 1982; Hartmann et al. 1984; Mecho et al. 1985; Randel 1987; Shiotani and Hirota 1985; Shiotani et al. 1990; and Manney et al. 1991b). Its origin has been discussed in terms of vertical propagation from the troposphere (Randel 1987), in situ generation from instability mechanisms in the stratosphere (Mechoso and Hartmann 1982; Hartmann 1983; Manney et al. 1991b), and internal, nonlinear vortex interactions (Lahoz et al. 1996).

6.4. Synoptic variability

In this section we examine the synoptic (day-to-day) variability of the SH stratosphere during 1995, as an example year. We include comparisons with the NH winter 1994–95, to demonstrate the significant differences between hemispheres in day-to-day variability. We analyze time sections of the various quantities discussed in section 6.3, together with synoptic maps illustrating development of a planetary-wave event and the final warming. We also briefly discuss some traveling planetary waves observed in the SH stratosphere.

a. Temperatures, winds, and eddy heat fluxes

The SH stratosphere is characterized by significantly less day-to-day variability than the NH. Figure 6.14 displays the 50-hPa zonal-mean temperatures at 80°N (top) and 80°S (bottom) for the years 1994–95 and 1995, respectively (solid dark line, with the long-term mean shown as the thick white line and the envelope of 1979–95 data indicated by shading). This figure demonstrates minimal variance (on both daily and interannual timescales) in the SH in midwinter. Typically, the 80°S zonal-mean temperatures vary by only a few degrees on a day-to-day timescale during the austral summer, fall, and early winter (January through June). Only in the late winter–spring does this variability begin to increase, prior to polar vortex breakup.

In sharp contrast to the SH, the NH shows a high degree of day-to-day variability in the stratosphere during midwinter, with smaller variability in late spring (top of Fig. 6.14). One of the most striking aspects of these NH temperatures are the rapid warmings during midwinter. These rapid warmings are termed stratospheric sudden warmings and are associated with episodic planetary-wave driving ("bursts") from the troposphere (as shown below). The polar night jet is also decelerated strongly during those polar warming events. Midwinter stratospheric warmings vary in intensity, but are quite common in the NH. On the other hand, the SH winter period is almost completely devoid of significant sudden warmings.

As with the temperatures, the zonal winds in the SH exhibit much less daily variability than the NH. Figure 6.15 displays the 10-mb zonal-mean zonal wind at 60°N (top) and 60°S (bottom) for the years 1994–95 and 1995, respectively, together with long-term means and envelope of daily extremes. Rapid decelerations of the jet are observed during midwinter in the NH, associated with the warming events seen in Fig. 6.14. In contrast, the SH winter variations are weak. The fact that the episodic warming events are tied to bursts of wave activity from the troposphere is demonstrated in Fig. 6.16, which shows the respective eddy heat-flux time series (at 50 mb, 60°N–S) for these years. In the NH, there were five large bursts of wave activity that were associated with reductions of the northern jet.
the final burst in early April reversed the flow to the summer easterly regime. In addition to slowing the zonal wind, this wave activity warms the polar region by increasing the downward circulation in the polar region (see Fig. 6.14). These wave events are the source of the large day-to-day variability observed in the NH.

In contrast to these NH wave events, the SH shows a much weaker wave forcing in midwinter. In early September 1995, there is a burst of wave activity in the SH that decreases the zonal wind at 10 mb by over 15 m s\(^{-1}\). In late September–October 1995, an extended burst of wave activity acts to decrease the vortex strength by a substantial amount. By late November, the 10-hPa zonal wind has reversed to easterlies. These data show that enhanced wave driving during SH spring acts to decelerate the jet, erode the vortex, and reverse the circulation to the austral summer easterly wind regime.

**b. The day-to-day evolution of the polar vortex**

The time series above quantify variations in zonal-mean quantities on a day-to-day basis. Evolution of the three-dimensional vortex may be viewed in terms of vortex area and by inspection of daily maps. Evolution of the area of the vortex during 1995 at 465 and 840 K is shown in Fig. 6.17, together with the 10-mb zonal wind for reference. These diagrams display the area within individual PV contours throughout the seasonal cycle, calculated in terms of "equivalent latitude" (i.e., the latitude of an equivalent PV distribution arranged symmetrically about the pole; see Buchart and Remsberg 1986). The polar vortex is identified by strong meridional PV gradients (see Fig. 6.10); note the earlier formation of the vortex at 840 K and longer persistence of the vortex at 465 K (well into spring). Figure 6.17 also shows a lack of significant high-frequency variability in vortex area throughout winter in the SH, consistent with the zonal-mean view discussed above.

O’Neill and Pope (1990) calculated PV as a function of equivalent latitude and time for the years 1979 to 1988 on the 850-K surface. In their paper, they observed that the surf zone of the Southern Hemisphere begins to develop as the eddy activity begins to increase during late winter and early spring. This is evident in Fig. 6.17 as the area of weak PV gradients at 840 K near 40°S, beginning in August and continuing through spring. A similar strong seasonality in the surf zone is not observed in the lower stratosphere (465 K).

For comparison, Fig. 6.18 shows corresponding plots for the NH winter 1994–95. Large variations in vortex area are observed in midwinter; in particular, there is a significant decrease in the vortex area throughout the stratosphere during the warming events of late January–early February. Note also
clear development of the surf zone (weak midlatitude PV gradients) at 840 K, beginning in January; the NH surf-zone area is significantly wider than that in the SH.

c. A case study of synoptic and planetary-wave activity

The wave-driving event of early September 1995 presents an interesting case study of vortex variability during a springtime wave event in the SH. The heat flux displayed in Fig. 6.16 shows a burst of wave activity between 9 and 12 September. Figure 6.19 displays false-color images of PV on the 460-K isentropic surface (~18 km) during this period. The flow is clockwise, approximately parallel to PV isolines; the jet core lies along the polar-vortex edge where the PV gradient is largest. The panels in Fig. 6.19 all illustrate the basic structure of the polar vortex: 1) strongly negative values of PV in the polar region (i.e., the polar vortex); 2) a strong PV gradient at about 60°S (the vortex edge); and 3) a broad region of weak PV gradients extending to the subtropics (the surf zone).

The first panel of Fig. 6.19 (6 September) shows a relatively circumpolar vortex, which begins to elongate as time progresses, extending well northward of...
60°S. This “lobe” of the vortex is evident on 9 September, reaching the southern tip of South America. The lobe moves slowly eastward between 9 and 13 September and has started to decay by 13 September. On 11 September, a tongue of material extends away from the vortex, up toward South America, and then westward across South America and over the Pacific Ocean. This tongue is an example of wave breaking in the SH (McIntyre and Palmer 1983, 1984). The preferred wave breaking in the South American region is a climatological feature, related to displacement of the vortex toward this region in SH spring (associated with stationary planetary wave 1—see Fig. 6.10). In this particular case, and in most springtime wave-breaking cases, the PV is stripped out of the vortex edge region and does not involve material from deep inside the vortex. This weak transport of material across the edge of the polar vortex is consistent with a number of numerical studies showing that the southern polar vortex is relatively isolated from midlatitudes over the winter period (above 400 K) and that exchange of air across the polar-vortex boundary is relatively weak (Bowman 1993a,b; Bowman and Mangus 1993; Chen 1994; Chen et al. 1994; Manney et al. 1994; McIntyre 1995).

In the middle stratosphere, this September wave event was evidenced in deceleration of the zonal wind at 10 mb by approximately 20 m s⁻¹ (Fig. 15). Figure 6.20 displays false-color images of PV on the 850-K isentropic surface (~30 km) during this period. The behavior of the PV at 850 K is qualitatively similar to that displayed at 460 K. In contrast to 460 K, however, a strong tongue of subtropical air is pulled southward around the edge of the vortex. This tongue is observed as the low PV (white colors), which is pulled out of the subtropics on 10 September over the South Atlantic Ocean, moves across the polar region, and is evident as low PV to the south of Australia on 13 September. A similar event was observed in UARS constituent observations by Randel et al. (1993) and analyzed further by Waugh (1993b).

d. A case study of the final warming

The most spectacular wave events in the SH occur during the vortex breakup in late spring (October–December). The heat flux displayed in Fig. 6.16 shows a number of bursts of wave activity in the October through December period, with diminishing intensity as the vortex weakens. Figure 6.21 displays a sequence of false-color images of PV on the 460-K isentropic
surface during this breakup phase. The vortex temperatures have risen considerably by the November-December time period (see Fig. 6.5). The first panel (6 November) shows a relatively intact circumpolar vortex, which has significantly decreased in strength since the early September period (see Fig. 6.20; note the color-scale change). This vortex continues to decrease in strength during November and early December, with substantial wave-like deformations. By 18 December, the vortex has virtually disappeared. This breakup was atypical in that the vortex persisted into mid-December at this altitude.

In the middle stratosphere, the breakup occurs significantly earlier than at lower levels. Figure 6.22 displays false-color images of PV on the 850-K isentropic surface during the breakup. The images shown in Fig. 6.22 are constructed using trajectory reverse domain filling (RDF), which produces a high-resolution representation of the PV field (see Sutton 1994; Newman and Schoeberl 1995; and Schoeberl and Newman 1995). The RDF field is represented by a dense regular grid of points; these gridded points are then advected backward in time with the trajectory model to an earlier time (in our case, 10 days prior to the date shown), and the PV observation from the analysis at this earlier time is plotted on the dense regular grid for the shown date. The evolution in Fig. 6.22 shows wave-breaking events that pull streamers of polar PV off of the vortex into long tongues of material extending well into the midlatitudes. Such tongues are ubiquitous features of large-amplitude (nonlinear) wave events in the middle stratosphere (e.g., Juckes and McIntyre 1987; Waugh 1993a; Norton 1994; Schoeberl and Newman 1995). These images demonstrate the rapid dispersion of vortex air into midlatitudes during the final warming (see similar results in Manney et al. 1994).

e. Traveling planetary waves

Several distinct modes of traveling planetary waves have been observed in the SH stratosphere. In terms of wave amplitude and importance for the general circulation, the eastward-traveling wave 2 (discussed at the end of section 6.3c) is the largest. Another distinctive wave mode observed in the polar upper stratosphere is the so-called “4-day wave,” discovered by Venne and Stanford (1979). Characteristics of this mode include

1) temperature-variance maxima confined to polar regions (centered near 60°–70°S) in the upper stratosphere and mesosphere (Venne and Stan-
2) eastward-propagating power for zonal waves 1–4; reconstructed synoptic maps reveal one or more warm pools circling the pole with period near 4 days (Prata 1984; Lait and Stanford 1988; Allen et al. 1997). [An example is shown in Fig. 6.23, adapted from Lait and Stanford 1988.]

3) analysis of wave and background mean-flow structure strongly suggests instability of the polar night jet and/or mesospheric subtropical jet as the source of excitation for the 4-day wave (Hartmann 1983; Randel and Lait 1991; Manney 1991; Manney and Randel 1993; Lawrence and Randel 1996; Allen et al. 1997).

Another traveling wave with sharp spectral characteristics is the 2-day wave. This is a westward-propagating, zonal wave 3–4 oscillation observed in the upper stratosphere and mesosphere of the subtropics during summer. It has been documented in satellite observations by Rodgers and Prata (1981), Burks and Leovy (1986), Wu et al. (1993), Randel (1994), Limposuvan and Leovy (1995), and Wu et al. (1996). One of the most interesting features of the 2-day wave is that it is observed in low to middle latitudes of both hemispheres, but only for a period of several weeks following summer solstice. Its origin has been discussed theoretically as being associated with a global normal mode (Salby 1981a,b) or due to dynamic instability of the summer easterly jet (Plumb 1983; Pfister 1985). Randel (1994) showed a combination of normal-mode structure and instability signature that suggests the 2-day wave is a near-resonant mode forced by dynamic instability.

Observational studies of satellite data have also shown evidence of westward-propagating, normal-mode Rossby waves in the upper stratosphere of the SH (Rodgers 1976; Hirota and Hirooka 1984; Venne 1989). These waves correspond to the free or resonant oscillations of the atmosphere and are characterized by 1) planetary-scale horizontal structure, 2) weak vertical phase tilts, and 3) regular westward propagation at (nearly) discrete frequencies, with typical periods of 5–15 days. These modes typically exhibit relatively small amplitudes and are interesting mainly because of their global coherence and similarity to calculated resonant modes (e.g., Salby 1981a).

### 6.5. Interannual variability

Most analyses of interannual variability in the stratosphere have focused on the NH, because of the relatively
Fig. 6.22. Potential vorticity variations at 840 K during breakdown of the SH polar vortex in November 1995.

long record of observations compared to the SH. Labitzke (1982) has analyzed NH stratospheric analyses over 100–10 mb dating from 1957, while Baldwin and Holton (1988) and Dunkerton and Baldwin (1991) have shown results from NCEP data (up to 10 mb) using the weekly analyses available since 1964. The NCEP global data analyzed here span 17 years, providing the best record to date to analyze interannual stratospheric variability in the SH.

Several studies of NH stratospheric variability have focused on observing the response to specified external forcings. These include extratropical effects of the tropical QBO (Holton and Tan 1980, 1982; Labitzke 1982; Dunkerton and Baldwin 1991), the 11-year solar cycle (Labitzke and van Loon 1987, 1995), and ENSO (van Loon and Labitzke 1987; Baldwin and O'Sullivan 1995). Newman and Randel (1988) analyzed interannual variability in the SH stratosphere in relation to

Fig. 6.23. Synoptic maps of polar temperatures in the upper stratosphere (near 45 km) during August 1981; contour interval is 2 K. Labeled features “A” and “EFH” are warm pools rotating about the pole with a period of 4 days. Feature A retains its identity for over seven complete revolutions about the pole (not shown here). [Adapted from Lait and Stanford (1988).]
Here, $\overline{T_i}$ is the monthly average temperature for year $i$, $\overline{T}$ is the ensemble monthly average, $N$ is the number of years available, and the cosine ($\phi =$ latitude) term accounts for geometrical effects (North et al. 1982). As a note, the $\sigma_{int}$ estimates in the Tropics in Fig. 6.24 are probably underestimated (approximately by a factor of two), due to the inability of the NMC data to fully resolve the tropical QBO signal.

Interannual variability in the SH lower stratosphere exhibits a maximum during late winter and spring (August–November) poleward of 50°S—i.e., during the spring warming. Interannual variations are comparatively small during the rest of the year, in particular during midwinter. This variability in the SH is distinctive from the corresponding NH patterns, which exhibit strong polar variability throughout midwinter (December–March). The space–time maxima in interannual variability are similar to those for daily variability (see Fig. 6.14); much of the interannual variance is related to the presence or absence of large stratospheric warming events during a particular month. Interannual variability of zonal wind in the middle stratosphere shows similar seasonality to that in Fig. 6.24 but with the extratropical maxima shifted 10°–15° equatorward (associated with thermal wind balance).

Figure 6.25 shows time series of monthly mean temperatures during November at 100 mb for several Antarctic stations. November data are chosen because of the maximum interannual variability at this time seen in Fig. 6.24. These plots show data from the NCEP reanalysis (Kalnay et al. 1996), together with radiosonde data spanning approximately 1960–97, for a longer-term perspective on interannual variability. The monthly mean radiosonde data here have been constructed from the available daily station observations, using the method of Trenberth and Olson (1989), which minimizes sampling biases. Figure 6.25 shows a significant component of year-to-year see-saw behavior that is (statistically) related to the QBO (as discussed below). The long-term record also suggests that some significant cooling has occurred in the polar lower stratosphere since approximately 1980.

The altitude–time structure of the decadal-scale cooling over Antarctica is quantified in Fig. 6.26, which shows the Antarctic temperature changes between the decades (1986–95 minus 1970–79), derived from radiosonde data (as in Fig. 6.25) (Randel and Wu 1998). These data show strong cooling (of order 6–10 K) in the lower stratosphere (~12–21 km), which maximizes in spring (October–December). Smaller-magnitude cooling (~1 K) persists throughout summer, while no significant changes are observed in winter. These decadal temperature changes exhibit similar space–time characteristics to antarctic ozone depletion since 1980. Furthermore, the observed cooling in Fig. 6.26 is in reasonable agreement with the ozone-hole temperature change calculations of Shine.
(1986), Kiehl et al. (1988), Mahlman et al. (1994), and Ramaswamy et al. (1996). This strongly suggests that the observed polar cooling in spring is primarily a radiative response to the Antarctic ozone hole. Waugh and Randel (1998) show a corresponding delay in the springtime vortex breakdown in the SH lower stratosphere (~2-week delay between the early 1980s and middle 1990s).

The seasonal evolution and interannual standard deviation of zonal-mean wind at 1 mb is shown in Fig. 6.27, illustrating interannual variability in the upper stratosphere (as with Fig. 6.24, the tropical variability may be underestimated in these NCEP-based estimates). Interannual variability in the SH upper stratosphere maximizes in midwinter near 30°S, on the equatorward flank of the climatological jet core. This SH midwinter variability maximum has been discussed by Shiotani et al. (1993) using a subset of the data here, and our discussion follows and extends their results. Figure 6.28 shows the seasonal march of

![Image of graphs showing temperature changes](image)

**Fig. 6.25.** Time series of November average 100-mb temperatures at eight Antarctic radiosonde stations. Solid lines denote radiosonde data, and dashed lines show data from NCEP reanalysis (Kalnay et al. 1996) at the radiosonde locations. The smooth curves in each panel show the decadal-scale variations at each location calculated from the radiosonde data. (From Randel and Wu 1998.)

![Image of Antarctic temperature change](image)

**Fig. 6.26.** Altitude–month profile of temperature differences over Antarctica between the decades (1986–95 minus 1970–79), derived from radiosonde data (see Fig. 6.25). Contour interval is 1 K, and shading denotes a 2-sigma statistically significant difference with respect to natural variability. Data are unavailable during midwinter at the uppermost levels. (From Randel and Wu 1998.)
near 30°–40°S and 60°–70°S, and extend throughout the middle and upper stratosphere. These anomalies correspond to a latitudinal shifting of the midwinter polar night jet; Shiotani et al. (1993) term these states the high-latitude jet and low-latitude jet. Shiotani et al. furthermore show evidence of systematic changes in planetary-wave structure during the respective regimes, such that larger wave amplitudes are observed in midwinter for the high-latitude-jet periods. The cause of this interannual variability is not well understood at present.

**Extratropical influences of the QBO in the SH**

The QBO dominates interannual variability in the tropical stratosphere. While it is not a dominant effect outside of the Tropics, statistical studies (primarily using the longer record of NH data) suggest that some component of middle–high-latitude variability is related to the QBO, in particular the strength of the NH winter polar vortex (e.g., Holton and Tan 1980, 1982; Dunkerton and Baldwin 1991). Evidence of an extra-

monthly mean zonal winds at 1 mb, 30°N and 30°S, for each year of the record here (1979–95); note the time axes are shifted for direct NH–SH seasonal comparisons. This figure illustrates the stronger zonal winds in the SH vs. NH subtropics and highlights the substantial interannual variability in the SH in midwinter (note the hint of similar variability in the December NH data). Time series of the SH anomalies versus year (not shown) reveal a see-saw pattern of positive and negative winds (with few near the long-term mean). These anomalies are not statistically correlated with the tropical QBO.

The spatial structure of the zonal-wind anomalies during July 1989 and 1992 (two opposite extremes) are shown in Fig. 6.29 (these years are noted with the light and heavy dashed lines in Fig. 6.28). The anomalies show a north–south dipole pattern, with centers

![Fig. 6.27. Latitude–time sections at 1 mb of the zonal-mean zonal wind (top) and its interannual standard deviation (Eq. (8)) over 1979–95 (bottom-contour interval 2 m s⁻¹).](image)

![Fig. 6.28. Time traces of monthly mean zonal-mean wind at 1 mb, 30°N (top) and 30°S (bottom). Each line denotes an individual year during 1979–95.](image)
tropical QBO signal is particularly clear in column ozone data and are observed in both the NH and SH (Bowman 1989; Lait et al. 1989; Tung and Yang 1994; Yang and Tung 1994). Randel and Cobb (1994) isolated extratropical QBO patterns in both column ozone and lower-stratosphere temperature using regression analyses tied to the observed QBO tropical winds. An update of their results for lower-stratosphere temperature (shown in Fig. 6.30) demonstrates that temperature variations in the midwinter subtropics and spring polar latitudes are statistically coherent with the tropical QBO winds. The modeling studies of Gray and Dunkerton (1990), O'Sullivan and Young (1993), O'Sullivan and Dunkerton (1994), and Buchart and Austin (1996) provide plausible mechanisms for transmitting the QBO signal to extratropics and are in qualitative agreement with these observations. Although the patterns in Fig. 6.30 are statistically significant, the QBO accounts for (at maximum) only 25%–35% of the interannual variance for seasonal anomalies in the time sample studied here; thus, the QBO is not a dominant mode of interannual variability in extratropics.

6.6. Ozone

Ozone is of prime importance in the stratosphere. Ozone absorbs solar ultraviolet radiation, providing the heat source for the temperature increase with altitude that defines the stratosphere [in the absence of ozone there would be no stratosphere, although the location of the tropopause itself is not very sensitive to the presence of ozone (Thuburn and Craig 1997)]. Ozone also provides the principal screening of this biologically harmful ultraviolet radiation. Thus, the analysis of the observed structure and variability of ozone is a central focus of stratospheric research.
The presence of ozone in the atmosphere was first deduced from observations of a sharp cutoff in the near ultraviolet end of the solar spectrum (Hartley 1880). Chapman (1930) provided a simple theoretical explanation for the formation and vertical structure of the ozone layer. The formation of ozone occurs by the photolysis of molecular oxygen by ultraviolet radiation (200–220 nm), followed by a combination of atomic and molecular oxygen:

\[
O_2 + h\nu \rightarrow O + O, \quad (6.9a)
\]

\[
O + O_2 + M \rightarrow O_3 + M, \quad \text{net: } 3O_2 \rightarrow 2O_3. \quad (6.9b)
\]

Here, \(M\) is a third body (typically a nitrogen or oxygen molecule) needed to conserve energy and momentum. Because the amount of ultraviolet radiation decreases downward from the top of the atmosphere while the amount of molecular oxygen decreases exponentially (with density) away from the earth’s surface, the production of ozone must be a maximum in a layer at some intermediate height centered at low latitudes (where sunlight maximizes). Chapman originally believed that ozone destruction takes place by reaction of ozone with atomic oxygen. It is now known that this mechanism is secondary to ozone destruction by so-called catalytic cycles. Catalytic loss takes place by removing both ozone and atomic oxygen via:

\[
Z + O_3 \rightarrow ZO + O_2, \quad (6.10a)
\]

\[
O + ZO \rightarrow Z + O_2, \quad \text{net: } O + O_3 \rightarrow 2O_2. \quad (6.10b)
\]

where \(Z\) is principally reactive chlorine, bromine, nitrogen, or hydrogen species (additional catalytic cycles are important for winter polar-ozone losses, as discussed in the following sections). Ozone production primarily occurs in the Tropics, but observations show that the ozone concentration is greater in the high-latitude lower stratosphere than in the Tropics, thus showing that transport is a crucial factor in ozone behavior. Because this transport depends in turn on the heating distribution associated with ozone, detailed understanding involves the coupling of chemistry, radiation, and dynamics in the stratosphere.

**a. Climatology of ozone**

The strong wavelength dependence of ultraviolet light absorption by ozone allows deduction of ozone amount by surface and satellite measurements (the latter using ultraviolet radiation backscattered from the surface and clouds). One instrument that provided a long time series of global column-ozone measurements was the TOMS on the **Nimbus 7** satellite (spanning 1978–93; McPeters et al. 1993). Column ozone is the vertically integrated total amount of ozone, often measured in terms of Dobson units (DU), after the pioneering work of G. Dobson. One DU is approximately equal to \(2.6 \times 10^{22}\) molecules of ozone per square meter; typical observed values are of order 300 DU.

Figure 6.31 shows the climatological global structure of column ozone derived from TOMS data. Here,
TOMS data have been longitudinally averaged with respect to the polar-vortex structure in the lower stratosphere using potential vorticity coordinates (Bufton and Remsberg 1986; Schoeberl et al. 1992; Randel and Wu 1995); this allows direct comparisons of the NH and SH column-ozone structure in relation to the vortex (which is also indicated in Fig. 6.31). Column ozone is a minimum over the Equator, increasing toward higher (particularly winter) latitudes. There is a strong seasonal cycle in ozone in extratropics, with maximum amounts over high latitudes in winter and spring of the respective hemispheres. Winter–spring column-ozone amounts are substantially higher in the NH than in the SH, but similarly sized in summer. The observed middle–high-latitude maximum in column ozone, together with the seasonal maximum in winter–spring and the hemispheric differences, are related to global-scale transport from the tropical ozone source region by both mean meridional circulations and eddy processes. Note that the seasonal cycle and interhemispheric differences in residual mean circulation in Fig. 6.4 are consistent with these aspects of observed ozone structure (i.e., flow from Tropics to high latitudes in winter and stronger transport during winter in the NH than in the SH). Inspection of Fig. 6.31 shows that column ozone is a maximum over latitudes near the edge or slightly outside of the vortex in each hemisphere, peaking in spring (March in the NH and October in the SH). This aspect is much more pronounced in the SH, where a “collar,” or maximum of ozone, is seen near 60°S. This maximum away from the pole is due to circulation effects, as the strongest downward velocity in the lower stratosphere occurs near the edge of the polar vortex, particularly in the SH (Schoeberl et al. 1992; Manney et al. 1994; clearly seen in Fig. 6.7c). Column ozone shows a relative minimum inside the vortex in the SH throughout winter–spring, especially pronounced in September–October. The deep spring minimum is mainly due to chemical depletion in the lower stratosphere associated with the ozone hole (discussed below), although a relative minimum in this region is a climatological feature related to the mean circulation (as discussed previously and observed in the NH winter in Fig. 6.31; see also Bojkov and Fioletov 1995, their Fig. 15).

The vertical profile of ozone and its seasonal evolution is shown in Fig. 6.32, based on HALOE data, combined with polar measurements from the microwave limb sounder (MLS) instrument on the UARS (Froidevaux et al. 1994; Manney et al. 1995). The quantity shown is ozone mixing ratio, in parts per million by volume. Note that column ozone (as shown above) is the vertical integral of mixing ratio times atmospheric density, which decreases exponentially upward, and hence the contribution to column ozone is heavily weighted from mixing ratios in the lower stratosphere in Fig. 6.32 [which is why column ozone is largest at high latitudes, even though the mixing ratio is largest in the middle stratosphere in the Tropics (the photochemical source region)]. There is a clear seasonal movement of the tropical mixing ratio maximum, following the position of the sun—over the SH sub tropics in January and over NH sub tropics in July. Relatively low values of ozone are found at high southern latitudes (inside the vortex) throughout SH winter. This is due to relatively weak poleward transport from tropical source latitudes across the vortex boundary; note the much-stronger high-latitude gradients of ozone in the SH versus the NH in winter (similar to the methane measurements seen in Fig. 6.8). Very low values are also observed in the polar lower stratosphere in October in Fig. 6.32, resulting from chemical ozone depletion (discussed below).

b. Ozone trends

Predictions of global ozone loss due to human activities were put forward by Johnston (1971) and Molina and Roland (1974), due to supersonic aircraft emissions and the release of chlorofluorocarbons, respectively. As nitrogen and chlorine compounds increase in the stratosphere, catalytic loss would also increase, and the new photochemical balance would shift ozone levels downward. The search for evidence of downward trends was inconclusive until the discovery of the Antarctic ozone hole by Farman et al. (1985). Since that time, numerous observational analyses, based on both ground and satellite measurements, have established statistically significant decreases in total ozone over much of the globe (see WMO 1992; Stolarski et al. 1991; Stolarski et al. 1992; and the most recent updates in WMO 1995; Bojkov and Fioletov 1995; Harris et al. 1997; and McPeters et al. 1996). Figure 6.33 shows the latitudeseason pattern of total ozone decreases during 1979–94 derived from TOMS data [similar to the results of Stolarski et al. (1991), but extended using data through 1994]. Similar results are derived using solar backscatter ultraviolet (BUV) data (Hollandsworth et al. 1995). Statistically significant decreases are observed in the SH at middle and high latitudes during spring (associated with the ozone hole) and extending into summer; relatively small losses are seen during autumn and winter. Large areas of ozone loss are observed in the NH midlatitudes during winter and spring; significant trends are not found in the Tropics. The majority of the column-ozone losses occur in the lower stratosphere. The SH ozone losses over midlatitudes during summer may be attributable to transport and persistence of ozone-poor air following breakup of the polar vortex (Atkinson et al. 1989; Sze et al. 1989; Frather et al. 1990; Cariolle et al. 1990; Mahlman et al. 1994; Brasseur et al. 1997).

Figure 6.34 shows a long record (1956–94) of monthly mean column-ozone anomalies at Mel-
bourne, Australia (38°S), derived from ground-based Dobson spectrophotometer measurements (Atkinson and Eason 1989). The data in Fig. 6.34 have been deseasonalized by removing the mean annual cycle to highlight interannual variations and in turn divided by the long-term mean (317 DU) to obtain percentage variations. Additionally, interannual anomalies from TOMS data (over 1979–94) are superimposed in Fig. 6.34, and these show excellent agreement with the ground-based data for this overlap period. The data in Fig. 6.34 show month-to-month variations in column ozone at Melbourne of order ±4%, associated with natural meteorological variability inherent to the atmosphere. Additionally, there is an overall downward trend apparent in Fig. 6.34 (highlighted by the heavy solid line, which is a smoothed version of the monthly mean data), such that mean ozone values are 6%–8% lower in the 1990s than in the 1950s (approximately −2% per decade decrease). The trends are somewhat larger for the period 1979–94 (near −5% per decade). The enhanced ozone loss since approximately 1980 is a feature observed over extratropics of both hemispheres; hemispheric and global averages prior to this time show near constant values (Bojkov and Fioletov 1995), unlike the decreases seen during 1955–75 at Melbourne in Fig. 6.34. The SH contributed approximately 65% of the global 5% decline during 1980–95.

c. The ozone hole

1) Observations

The basic source, sink, and transport mechanisms affecting ozone were thought to have been well understood in the early 1980s, as evidenced by reasonable
simulation of observed data (e.g., Garcia and Solomon 1983). This understanding was shaken by the unanticipated observations by Farman et al. (1985) of an approximate 50% decrease in column ozone over Antarctica during the decade 1975–85, but only during spring season (September–October). Figure 6.35 shows an updated version of the data presented in Farman et al. (see Jones and Shanklin 1995), showing October average column ozone measured by ground-based Dobson spectrophotometer over Halley Bay, Antarctica, over the years 1956–95. Values near 300 DU were observed prior to the mid-1970s, steadily declining to values below 150 DU in the 1990s.

Analysis of TOMS satellite data by Stolarski et al. (1986) showed that the depletion of ozone during SH spring occurred over a large area centered over the South Pole, comparable to the size of Antarctica. Figure 6.36 shows SH October monthly mean data from TOMS for the years 1979 and 1993–97, together with Nimbus 5 BUV measurements from 1970 to 1972 (data from these instruments are included in Fig. 6.35, showing reasonable agreement with the ground-based measurements). Figure 6.36 illustrates the large ozone losses over polar regions between the 1970s and 1990s. The most dramatic losses occur inside the SH polar vortex (cf., the PV structure in Fig. 6.10), although midlatitude values also decrease (note the change in the collar region near 60°S in Fig. 6.36; see also Fig. 6.33). Note that a relative minimum in ozone over the pole was observed in the early 1970s (prior to the recent large depletions); this is the climatological minimum discussed above related to circulation and transport effects on ozone. The “ozone hole” refers not to this weak minimum, but to the relatively large depletions observed in comparing the earlier and later periods. Column-ozone values below 100 DU have been observed in the ozone hole since 1993 (Hofmann et al. 1994; Herman et al. 1995; Hofmann et al. 1997), and these are the lowest values ever observed anywhere on earth. Figure 6.37 shows the area of the ozone hole measured by TOMS, defined as the area inside the 220-DU contour during SH spring, illustrating rapid growth during the 1980s and approximately constant size during the 1990s (note that over 1979–
there was effectively zero area, with ozone levels below 220 DU). Jiang et al. (1996) present further details of ozone-hole interannual variability based on TOMS data.

The vertical profile of ozone depletion associated with the ozone hole is shown in Fig. 6.38, based on balloon-borne ozonesonde observation at the South Pole during 1992 and 1993 (Hofmann et al. 1994); the quantity shown is ozone partial pressure, proportional to mixing ratio times atmospheric density. The data in late August 1993 shows a column ozone amount near 270 DU, with a profile maximum over 15–20 km (approximately 100–50 mb). By the middle of October 1993, the ozone has substantially vanished over 13–20 km, and the column amount is 91 DU. Similar results are found for other years, demonstrating that the ozone hole is attributable to some process depleting ozone over this 13–20-km layer. Analyses of recent South Pole ozonesonde data compared with historical records (see WMO 1995 and Hofmann et al. 1997) clearly show that anomalous ozone loss begins in spring (September) and extends into early summer (December–January). Profiles show that this depletion occurs in the lower stratosphere, with strongest ozone losses occurring during the time period September–December.

2) **Theory of Chemical Ozone Depletion in the Ozone Hole**

The primary characteristics of the ozone hole (reviewed above) show strong ozone losses within the SH polar vortex in spring, primarily in the lower stratosphere. Departures of the column amount from climatological values near 300 DU are first evident in the mid-1970s (see Fig. 6.35); dramatic changes were first noted in the mid-1980s, and the ozone hole is now a regular feature in the SH stratosphere, with a nearly complete loss of ozone in the lower stratosphere each year (Fig. 6.38). Although the causes of the ozone hole were strongly debated in the 1980s, a series of ground-based and aircraft measurement campaigns of the ozone-hole chemistry and dynamics have led to a scientific consensus that chemical ozone depletion associated with chlorine and bromine are responsible for the ozone hole (WMO 1992). Stratospheric chlorine primarily results from photochemical breakdown of chlorofluorocarbons released in the troposphere from human activities, and the ozone hole is a remarkable signature of human climate influence on a planetary scale. Subsequent observations from aircraft and satellites, together with laboratory and modeling studies, have given better understanding to details of chemical ozone losses inside the ozone hole, as discussed in WMO (1995) and summarized briefly here.

The gas-phase catalytic cycles that destroy ozone over most latitudes are not effective over winter polar regions (particularly in the lower stratosphere, where the ozone hole occurs), because there is not enough sunlight to provide atomic oxygen [needed to complete the catalytic cycle in Eq. (6.10b)]. Rather, photochemical loss of ozone here is attributed to catalytic cycles involving enhanced levels of chlorine monoxide (ClO). The principal reactions involve the formation of the ClO dimer Cl₂O₂, followed by photolysis and thermal decomposition (Molina and Molina 1987):
Fig. 6.37. Size of the Antarctic ozone hole during 1979–97, as calculated by the area of column ozone with values under 220 DU (from TOMS data).

\[
\begin{align*}
\text{ClO} + \text{ClO} + M & \rightarrow \text{Cl}_2\text{O}_2 + M \\
\text{Cl}_2\text{O}_2 + h\nu & \rightarrow \text{ClO}_2 + \text{Cl} \\
\text{ClO}_2 + M & \rightarrow \text{Cl} + \text{O}_2 + M \\
2(\text{Cl} + \text{O}_3 & \rightarrow \text{ClO} + \text{O}_2) \\
\text{net: } 2\text{O}_3 & \rightarrow 3\text{O}_2
\end{align*}
\] (6.11a-c)

Other reactions involve the interaction of ClO and bromine species (McElroy et al. 1986; see also Anderson et al. 1989; Solomon 1990; and WMO 1995). The Cl\_2\_O\_2 dimer reactions [Eq. (6.11)] account for approximately 80% of polar ozone losses (Solomon 1990). The rate-limiting step of these reactions is Eq. (6.11a), and the rate of chemical ozone loss is hence proportional to the amount of ClO squared (which helps explain the rapid increase in Antarctic ozone depletion during the 1980s in response to the approximate linear increase in anthropogenic chlorine).

In order for these reactions to proceed, two important components must occur: 1) a change of chlorine from inactive, so-called "reservoir" species (namely ClONO\_2 and HCl) into reactive chlorine (Cl\_2, ClO, and Cl\_2\_O\_2); and 2) exposure to sunlight. These processes, which determine the location and timing of the ozone hole, are illustrated schematically in Fig. 6.39 [from WMO (1995), adapted from Webster et al. (1993)]. The key factor that occurs inside the SH polar vortex is the transformation from reservoir to reactive chlorine species. This occurs by chemical reactions occurring on the surfaces of particles (heterogeneous chemistry), rather than the more normal reaction of well-mixed gases (gas-phase or homogeneous chemistry) (Solomon et al. 1986a; Solomon 1988). One important heterogeneous reaction is

\[
\text{ClONO}_2(\text{g}) + \text{HCl}(\text{s}) \rightarrow \text{Cl}_2(\text{g}) + \text{HNO}_3(\text{s}),
\]

which converts the relatively benign reservoir species ClONO\_2 and HCl into Cl\_2, which can then be photolyzed by visible wavelength light (g, gas; s, solid).

The particles on which this and similar heterogeneous reactions occur are sulfate aerosols or polar stratospheric clouds (PSCs), which in turn form only at the very cold temperatures of the vortex interior (typically at temperatures below 195 K). It had long been recognized that such clouds form in the polar winter stratosphere (Stanford and Davis 1974), and that they occur with greater frequency and thickness in the Antarctic than in the Arctic (McCormick et al. 1982). The threshold formation and growth of PSC particles have been observed by balloon-borne measurements over Antarctica (Hofmann et al. 1989; Hofmann and Deshler 1989) and by aircraft measurements (Dye et al. 1990). A climatology of PSC observations derived from satellite measurements is shown in Fig. 6.40 (from Poole and Pitts 1994), indicating the frequent occurrence of PSCs inside the SH vortex in winter, primarily in the intensely cold lower stratosphere. Note the correspondence between the PSC probability in Fig. 6.40 and the polar temperature evolution in Fig. 6.5, and the much higher frequency of PSCs in the SH than in the NH due to the colder SH temperatures. The CLAES instrument on UARS was the first to obtain synoptic measurements (maps) of enhanced aerosol (PSC) levels over Antarctica, and

Fig. 6.38. Comparison of pre-depletion South Pole ozone profile in 1993, with the profile observed when total ozone reached a minimum in 1992 and 1993. (Adapted from Hofmann et al. 1994.)
Fig. 6.39. Schematic of the photochemical and dynamical features of the polar regions related to ozone depletion. The upper panel represents the conversion of chlorine from inactive to active forms in winter in the lower stratosphere and the reformation of inactive forms in spring. The partitioning between the active chlorine species Cl₂, ClO, and Cl₂O₃ depends on exposure to sunlight after PSC processing. The corresponding stages of the polar vortex are indicated in the lower panel, where the temperature scale represents changes in the minimum polar temperatures in the lower stratosphere. (From WMO 1995; adapted from Webster et al. 1993.)

Fig. 6.41 shows examples of these data during SH winter 1992 (from Mergenthaler et al. 1997). Also shown in Fig. 6.41 are contours representing regions of temperature below 195 and 187 K (PSC threshold temperatures) from UKMO analyses. These CLAES data clearly show enhanced PSC amounts in polar regions throughout winter, generally coincident with regions of temperature below 195 K.

An additional important process in the SH involves the freezing out and sedimentation of nitric acid (HNO₃) in low-temperature PSCs. This results in an irreversible loss of reactive nitrogen inside the vortex (so-called denitrification), which in turn inhibits the reformation of the reservoir CNIO₂ from active chlorine (and hence maintains high levels of active chlorine). [A similar freezing out of water vapor inside the vortex leads to marked dehydration (Harwood et al. 1993; Tuck et al. 1993; Pierce et al. 1994), although this is not directly related to ozone chemistry.] By middle winter, most of the chlorine inside the vortex in the SH lower stratosphere is in the form of Cl₂ or Cl₂O₃, requiring only weak sunlight to initiate catalytic ozone loss.

When sunlight returns to the Antarctic in SH spring, ozone is rapidly destroyed in the lower stratosphere by catalytic cycles involving ClO and BrO (at the rate of approximately 1%–2% per day). This is why the ozone hole occurs in springtime. Some of the most compelling evidence for this sequence of events comes from direct observations of the ClO molecule (which occurs in very small amounts in the lower stratosphere away from aerosol- or PSC-processed air) and its anticorrelation with ozone. Figure 6.42 shows aircraft-based ClO and O₃ measurements covering a latitudinal sweep across the edge of the polar vortex (from Anderson et al. 1989). This shows a strong increase of ClO and decrease of O₃ inside the vortex (poleward of ~69°S) and strong anticorrelation for small-scale features near the vortex edge (over 67°–69°S). Figure 6.43 shows simultaneous observations of ClO and O₃ during September 1991 and 1992 from the MLS on UARS (Waters et al. 1993). These show strongly enhanced levels of ClO inside the SH polar vortex, with spatial patterns that are clearly mirrored in the low ozone values. Observations of other key chlorine and nitrogen species confirm the basic understanding
of the SH polar ozone budget (Solomon 1990; Santee et al. 1995; Douglass et al. 1995; Santee et al. 1996; Chipperfield et al. 1996). The most recent comprehensive modeling studies of Antarctic ozone depletion, which incorporate comprehensive chemistry and aerosol microphysical parameterizations in realistic dynamical models (Brasseur et al. 1997; Portmann et al. 1996; Schoeberl et al. 1996), produce accurate simulations of observed ozone loss rates and reasonable estimates of observed interannual variability.

6.7. Looking to the future

Our study and understanding of the dynamics, radiation, and chemistry of the Southern Hemisphere stratosphere has increased considerably over the last decade, primarily motivated by the discovery of the Antarctic ozone hole. Much of this rapid progress, however, resulted from the large body of observations compiled prior to the ozone hole discovery. These pre-ozone hole observations through the 1950s, 1960s, and 1970s in regions that appear to be pristine and that seemingly had little impact on human affairs justified the need for long-term monitoring and validated the foresight of the scientists who implemented these long-term observational programs.

There are a variety of open questions concerning the Southern Hemisphere stratosphere that require further research. The foremost question concerns the recovery of springtime Antarctic ozone levels as stratospheric chlorine levels decrease in response to the regulations on chlorine-containing compounds, via the Montreal protocol and the amendments to that protocol. Continuous global monitoring of these substances now indicate that the total ozone-destroying potential peaked in the troposphere in 1994 and will peak in the stratosphere between 1997 and 1999 (Montzka et al. 1996). A slow healing process would then begin, with stratospheric chlorine levels reaching \( \sim 1980 \) values (associ-
ated with emergence of the ozone hole) in about the year 2050 (WMO 1995). Hofmann et al. (1997) propose that early healing may be detectable from ozonesonde data as early as 2004. Additional questions concern climate-change impacts on the southern stratosphere in response to increased greenhouse gas loading and ozone depletion. Ramaswamy et al. (1996) have recently shown that much of the long-term cooling of the lower stratosphere (see Figs. 6.25–6.26) is consistent with a radiative response to observed ozone depletion; moreover, the effect on lower-stratosphere temperatures of ozone losses during the last two decades outweighs the radiative effects of changes in other greenhouse gases over the past two centuries! The details of planetary-wave variability in the southern stratosphere require continued analysis and study for understanding the characteristics of the SH polar vortex and the polar-vortex breakup. Additionally, many aspects of trace constituent transport and variability remain to be explored in the SH stratosphere; one intriguing question concerns why the SH midlatitude lower stratosphere is substantially drier than that in the NH (Kelly et al. 1990; Tuck et al. 1993; Rosenlof et al. 1997). These questions point toward a continuation and improvement of observations in the SH, particularly at high latitudes.

6.8. Summary

The stratospheric polar vortex in the SH is colder, with more intense zonal winds and potential vorticity gradients, than that in the NH. The SH vortex is also more stable on a daily basis (with an absence of midwinter breakdowns) and persists longer into springtime than its NH counterpart. The fundamental reason for this difference is the smaller amplitude of large-scale wave forcing from the troposphere in the SH (due to absence of topography and continental heat sources found in the NH).

The intense polar vortex in the SH autumn, winter, and spring exerts a strong influence on transport and mixing of stratospheric trace constituents. There is minimal horizontal transport across the vortex edge, so that the vortex interior is effectively isolated from midlatitudes; this is clearly demonstrated in observations of long-lived constituents (see Figs. 6.8–6.9). The vortex structure is also a dominant factor in the seasonal evolution of ozone, both from a transport and chemical perspective (see also Schoeberl and Hartmann 1991).

The ozone hole is now a climatological feature of the SH stratosphere. There is a near-complete loss of ozone in the layer ~13–20 km throughout the polar vortex in spring. Observations of key chemical species (namely ClO) show conclusively that the Antarctic ozone loss is due to chemical depletion, linked to the buildup of anthropogenic chlorine emissions. Chemical ozone loss in the Antarctic ozone hole is tied to heterogeneous reactions occurring on the surface of PSCs. These PSCs occur only in the intense cold of the Antarctic lower stratosphere; the return of sunlight in spring (completing the photochemical catalytic cycles) explains why the ozone hole occurs over Antarctica in spring. Analyses of long-term records show a substantial cooling of the SH lower stratosphere in spring (of order 6–10 K), and this is likely a radiative response to the depleted ozone levels. The Antarctic stratosphere thus provides the most conclusive evidence of the effect of humans on the climate system.

Acknowledgments. A significant portion of this work was completed while WJR was on sabbatical leave at the Cooperative Research Centre for Southern Hemisphere Meteorology at Monash University, Melbourne; thanks to Prof. David Karoly and the CRC staff for hospitality and a professional work environment. Annita Allman (CRC) and Marilena Stone (NCAR) expertly prepared the manuscript. The National Center for Atmospheric Research is sponsored by the National Science Foundation; partial support was also provided to WJR by NASA grants W-16215 and W-18181. The authors thank Rolando Garcia, Kevin Hamilton, Jim Holton, Bill Mankin, Alan O’Neill, Mark Schoeberl, John Stanford, and Darryn Waugh for discussions throughout the course of the work and for comments on the manuscript, and Prof. Donald Johnson for a careful review of the text.
APPENDIX
Notation

$(u, v, w)$ — three-dimensional velocity components
$(\bar{v}_h, \bar{w}_h)$ — residual mean meridional circulation
$T$ — temperature
$\theta$ — potential temperature [Eq. (6.6)]
$(\phi, \lambda)$ — latitude, longitude
$z$ — log-pressure vertical coordinate [$z = -H \ln(p/p_0)$]
$p_0$ — background atmospheric density ($p_0 = p_{surface} e^{-\phi H}$)
$p$ — pressure
$PV$ — potential vorticity [Eq. (6.7)]
$N^2$ — static stability parameter (Brunt–Väisälä frequency) $N^2 = (R/H)((\partial T/\partial z) + (2/\theta) (\partial H))$
$G$ — zonal momentum forcing [Eq. (6.1)]
$F$ — Eliassen–Palm flux vector [Eq. (6.5)]
$p_0 \nabla \cdot F$ — EP flux divergence (large-scale zonal momentum forcing)

$X$ — zonal momentum forcing from unresolved scales
$Q$ — radiative heating/cooling rate
$\Omega$ — earth rotation rate ($7.3 \times 10^{-5}$ s$^{-1}$)
$H$ — scale height (7 km)
$a$ — earth radius ($6.37 \times 10^6$ m)
$R$ — gas constant ($287$ K$^{-1}$ m$^2$ s$^{-2}$)
$\bar{A}$ — zonal mean of $A$
$A'$ — deviation of $A$ from zonal mean

REFERENCES
Andrews, D. G., and M. E. McIntyre, 1978: Generalized Eliassen–Palm and Charney–Drazin theorems for waves on axisymmetr-


