Chapter 6 Stratospheric Sudden Warmings

6.1 Introduction

It has been shown in Section 5.2.1 that the climatological zonal-mean zonal winds in the winter stratosphere are generally westerly and increase with height, peaking in the "polar night jet" vortex (see, for example, Figs. 5.2a,c). The zonal-mean climatological temperature fields (Figs. 5.1a,c) decrease towards the winter pole on each pressure surface in the stratosphere.

During some winters, however, this zonal-mean configuration is dramatically disrupted, with polar stratospheric temperatures increasing rapidly with time, leading to a poleward increase of zonal-mean temperature and, on occasion, a reversal of zonal-mean winds to an easterly direction. Such an event is called a *stratospheric sudden warming.* It is defined, somewhat arbitrarily, to be a *major warming* if at 10 mb or below the zonal-mean temperature increases poleward from 60° latitude and the zonal-mean zonal wind reverses. If the temperature gradient reverses there but the circulation does not, it is defined to be a *minor warming.* Major warmings occur on average about once every other winter in the Northern Hemisphere, but have not been observed in the Southern Hemisphere. Minor warmings in the upper stratosphere occur more frequently, and in both hemispheres. Table 6.1 gives a list of recent occurrences of major warmings in the Northern Hemisphere. The sudden and dramatic nature of stratospheric warmings is illustrated by the fact that temperatures near the pole may increase by 40–60 K in 1 week at 10 mb.

Although the definitions of major and minor stratospheric warmings are given in zonal-mean terms, these phenomena are far from zonally symmetric

Table 6.1

Occurrences of Major Stratospheric Warmings in Recent Northern-Hemisphere Winters[a]

Year	Number	Month
1964–1965	0	
1965–1966	1	F
1966–1967	0	
1967–1968	1	J
1968–1969	0	
1969–1970	1	J
1970–1971	1	J
1971–1972	0	
1972–1973	1	J–F
1973–1974	0	
1974–1975	0	
1975–1976	0	
1976–1977	1	J
1977–1978	0	
1978–1979	1	F
1979–1980	0	
1980–1981	1	F
1981–1982	0	
1982–1983	0	
1983–1984	0	
1984–1985	1	J
1985–1986	0	

[a] Note: This table differs slightly from those given by Labitzke (1982) and McInturff (1978), owing to differing definitions of final warmings.

in form. It has been known for many years that the stratospheric circulation is strongly zonally asymmetric before and during sudden warmings, and satellite data are now providing complex three-dimensional pictures of the time-development of these events and the variety of forms that they can assume. Current theories, primarily stremming from the paper of Matsuno (1971), suggest that tropospherically forced planetary waves play a crucial role in the dynamics of sudden warmings. However, although the combination of satellite data with a hierarchy of models has done much in recent years to elucidate the mechanisms of warmings, our understanding of all the observed details of these events and the necessary conditions for their occurrence is still by no means complete.

6.2. Observed Features of Sudden Warmings

The first observation of a stratospheric sudden warming was made by Scherhag in 1952, using radiosonde measurements over Berlin, and the radiosonde network has continued to provide much information on sudden warmings in the lower and middle stratosphere. This information was originally supplemented by rocket observations up to the high mesosphere at selected times and places. More recently, satellite measurements with good horizontal and temporal coverage and satisfactory vertical resolution at various levels in the stratosphere and mesosphere have given new insights into the three-dimensional morphology of sudden warmings.

As an example of the type of behavior that can be observed during a major stratospheric warming, we describe some basic features of the warming of February 1979, using a variety of diagnostics. This particular event has received much attention, since it occurred during a period of intensive observation of the atmosphere (the First GARP Global Experiment) and shortly after the launch of the *TIROS-N* and *Nimbus* 7 satellites. Although perhaps not "typical" in some respects, it nevertheless displays most of the features that are generally associated with major sudden warmings.

6.2.1 The Zonal-Mean Picture

During January and February 1979, three warming episodes took place in the northern polar stratosphere, with zonal-mean temperature maxima at 10 mb and 80°N, for example, occurring around January 25, February 6, and February 26 (see Fig. 6.1). The first two of these were minor warmings, but the third and largest was associated with a reversal of the zonal-mean wind $\bar{u}$ at and below 10 mb, corresponding to a major stratospheric warming. Latitude–height plots of $\bar{u}$ for December 8, 1978 (well before the warming

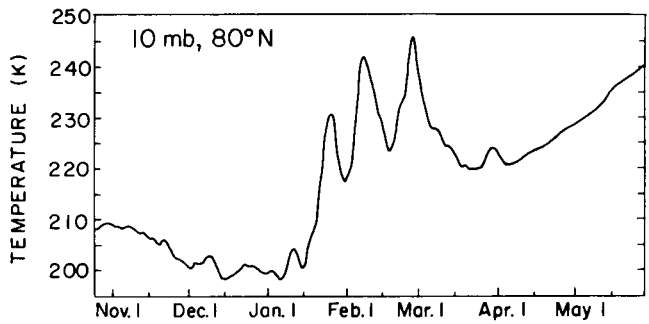

Fig. 6.1. Variation of zonal-mean temperature at 10 mb, 80°N, from October 1978 through May 1979, derived from LIMS data. [After Gille and Lyjak (1984).]

events), and January 25, and February 6 and 26, 1979 are shown in Figs. 6.2a-d, and give some idea of the time-variation of the mean zonal wind before and during the warmings. The broad region of easterlies that appears in the polar stratosphere and mesosphere at the time of the major warming is clearly visible in Fig. 6.2d. In early March the temperature falls again in the polar stratosphere (Fig. 6.1) before resuming its increase toward summer values, and winds revert to westerlies for a while (Fig. 6.2e) before the spring transition to summer easterlies.

6.2.2 Synoptic Description

As mentioned in Section 6.1, zonal-mean diagnostics give only a partial view of the rich three-dimensional structure of sudden warmings. Further observational details have traditionally been supplied by synoptic maps of temperature and geopotential at various levels in the stratosphere and by rocket soundings from isolated points on the globe. As an example, Fig. 6.3 presents polar stereographic charts of the 10-mb height and temperature fields for the days leading up to, and following, the major warming of late February 1979. On February 17 (Fig. 6.3a), the basic cyclonic polar vortex is elongated (as shown by the roughly elliptical height contours) and is

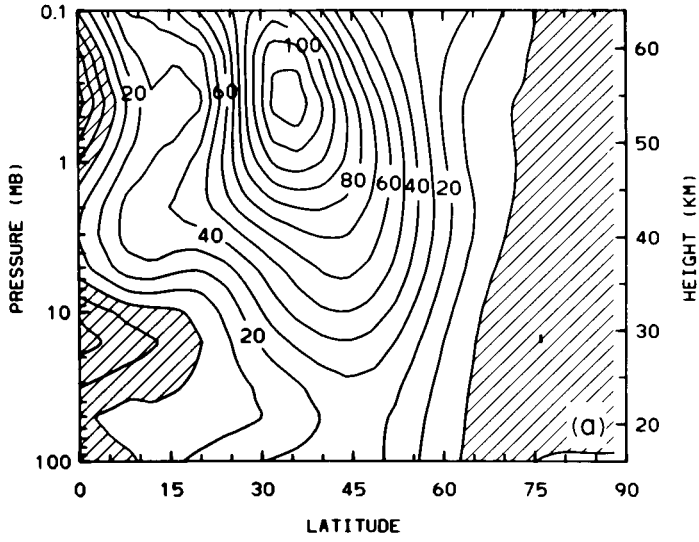

Fig. 6.2. Zonal-mean geostrophic wind (m s^{-1}), derived from LIMS data for (a) December 8, 1978, (b) January 25, 1979, (c) February 6, 1979, (d) February 26, 1979, and (e) March 3, 1979. Easterlies are shaded. [Adapted from Gille and Lyjak (1984); courtesy of Dr. John C. Gille.]

centered just off the pole, with an Aleutian High being present at about 180°W, 60°N. The temperature contours show an elongated cold region centered at about 20°E, 75°N and warm regions at about 100°E, 60°N and 80°W, 60°N. On February 19 (Fig. 6.3b), these structures have rotated slightly eastward and elongated further. By February 21 (Fig. 6.3c), a dramatic

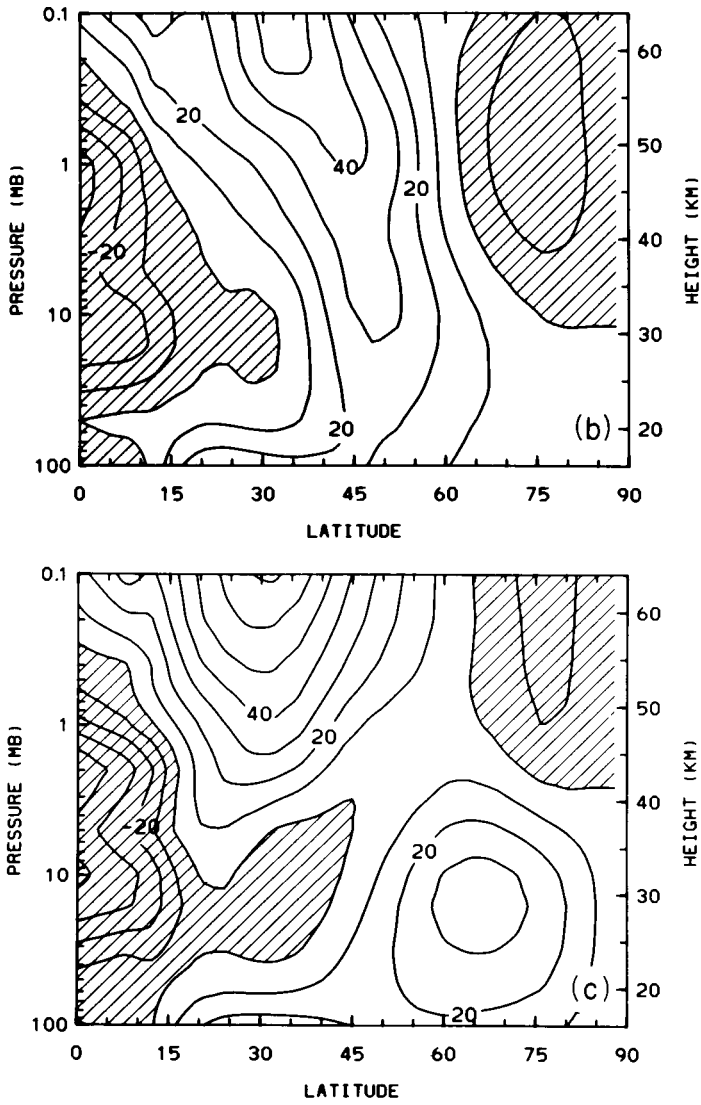

Fig. 6.2 (*figure continues*)

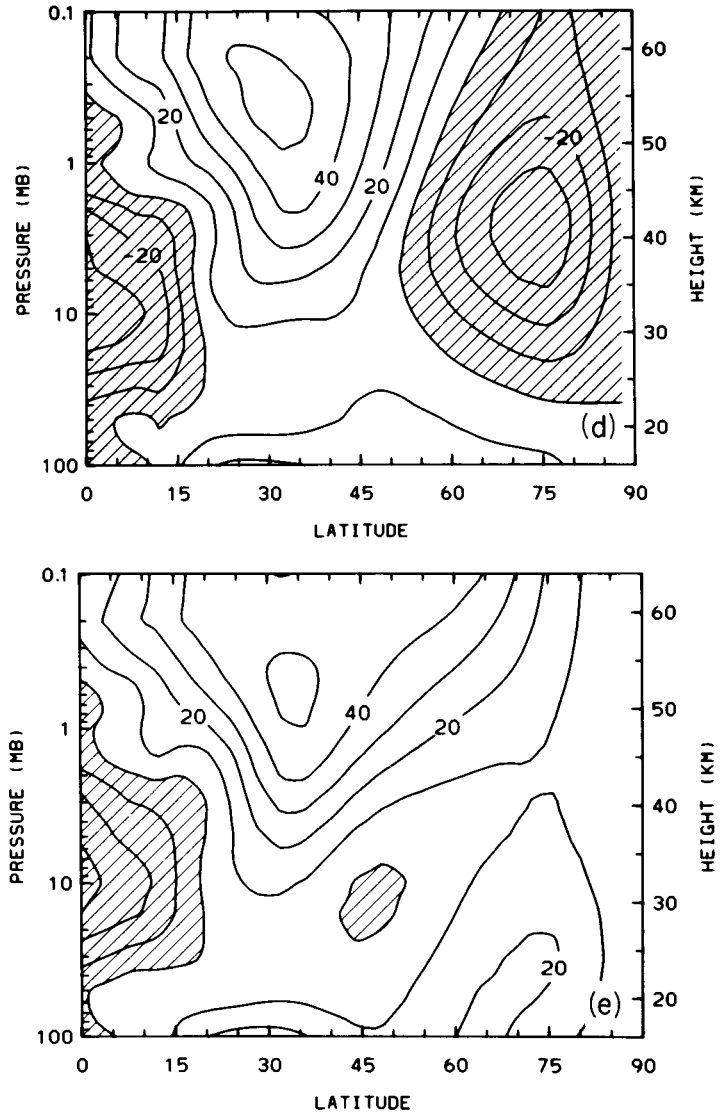

Fig. 6.2 (*continued*)

change has occurred, with the vortex splitting into two cyclonic centers as the Aleutian High extends northward and a ridge develops over the British Isles. The coldest air at 10 mb is now found at about 50°E, 55°N, with another cold patch occurring at 100°W, 55°N. Relatively warm air extending over the pole from 110°E to 60°W indicates that the polar stratospheric

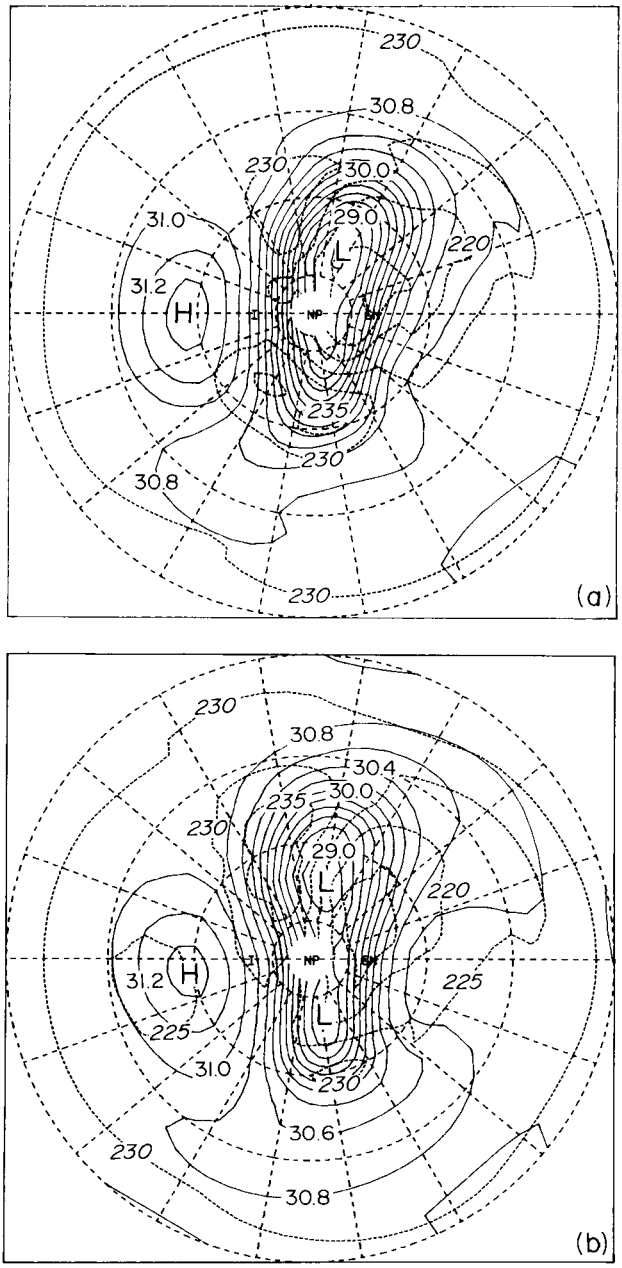

Fig. 6.3. Polar stereographic charts of 10 mb height (solid curves; contour interval 0.2 km) and temperature (dashed curves; contour interval 5 K) from LIMS data for the following days in 1979; (a) February 17, (b) February 19, (c) February 21, (d) February 26, (e) March 1, and (f) March 5. "GM" designates the Greenwich Meridian, which extends horizontally toward the right from the North Pole (NP). Latitude circles are shown at 20° intervals, with the outermost circle at 20°N. (*Figure continues.*)

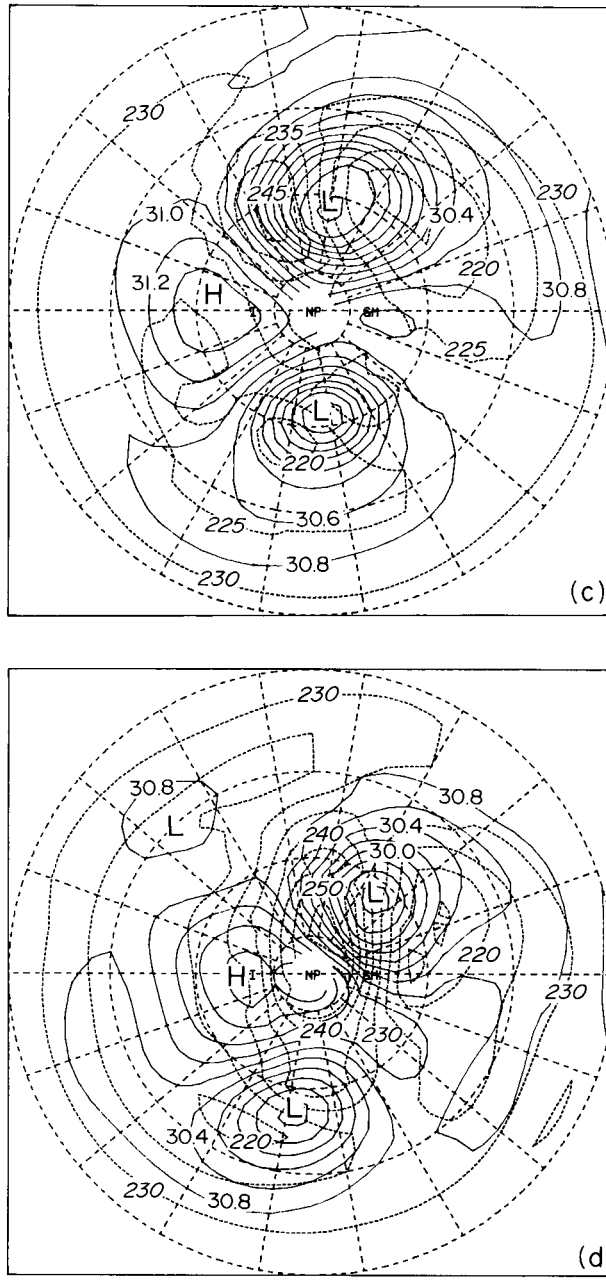

Fig. 6.3 (*continued*)

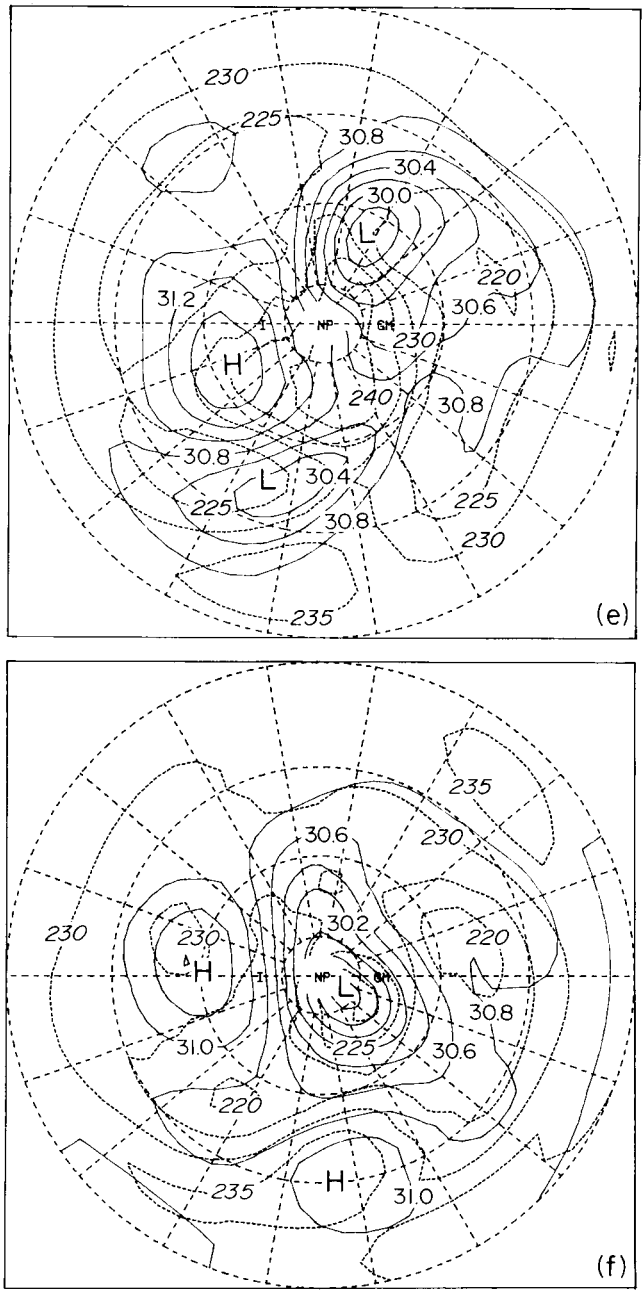

Fig. 6.3. (*continued*)

warming is taking place. This pattern rotates slightly westward, and then weakens, over the next few days (Figs. 6.3d,e), and by March 5 (Fig. 6.3f) only a single cyclonic vortex (albeit distorted, and centered slightly off the pole), accompanied by an Aleutian High, is visible.

6.2.3 *Description in Terms of Wave, Zonal-Mean Flow Interaction*

One way of quantifying the time-variations in the isobaric height and temperature fields described in the previous section is to Fourier-analyze them [as in Eq. (5.2.1)] and plot the amplitudes and phases of the various Fourier components as functions of time. Figure 6.4 shows the amplitudes of the zonal wave-number 1 and 2 components of the height fields (often called "height waves 1 and 2") as functions of time and latitude at 10 mb for the 1978–1979 winter, together with a similar plot for the zonal-mean wind $\bar{u}$. The wave-number 1 height amplitude (Fig. 6.4a) peaks in late January and early February, at roughly the times of the minor warmings shown in Fig. 6.1, corresponding to an elongation and off-polar displacement of the main vortex. The wave-number 2 amplitude (Fig. 6.4b) peaks at about February 21, as the vortex splits into two circulations (see Fig. 6.3c): the reversal of $\bar{u}$ at this time is evident in Fig. 6.4c. Diagnostics of this general type have been used for characterizing the observed behavior of many sudden warmings.

An extension of this method is to attempt dynamical interpretation using diagnostics suggested by wave, mean-flow interaction theory. For example, one may plot meridional cross sections of the Eliassen–Palm flux **F** and its divergence (see Section 3.5) at various times during the sudden warming period. These quantities are quadratic functions of disturbance fields u', v', θ', etc.; to the extent that the disturbances can be regarded as planetary waves propagating on the zonal-mean flow $\bar{u}(\phi, z, t)$, **F** may be interpreted as giving their direction of propagation in the meridional plane (see also Sections 4.5.5 and 5.2.2). Furthermore, under quasi-geostrophic scaling, but with no restriction to linear wave theory, $\nabla \cdot \mathbf{F}$ represents the sole eddy-forcing of the mean flow in the transformed Eulerian-mean equations [Eqs. (3.5.5)]. In spherical geometry the latter set become

$$\bar{u}_t - f\bar{v}^* - \bar{X} = (\rho_0 a \cos\phi)^{-1} \nabla \cdot \boldsymbol{F} \equiv D_F, \tag{6.2.1a}$$

$$\bar{\theta}_t + \bar{w}^* \theta_{0z} - \bar{Q} = 0, \tag{6.2.1b}$$

$$(a \cos\phi)^{-1} (\bar{v}^* \cos\phi)_\phi + \rho_0^{-1} (\rho_0 \bar{w}^*)_z = 0, \tag{6.2.1c}$$

$$f\bar{u}_z + \frac{R}{aH} e^{-\kappa z/H} \bar{\theta}_\phi = 0, \tag{6.2.1d}$$

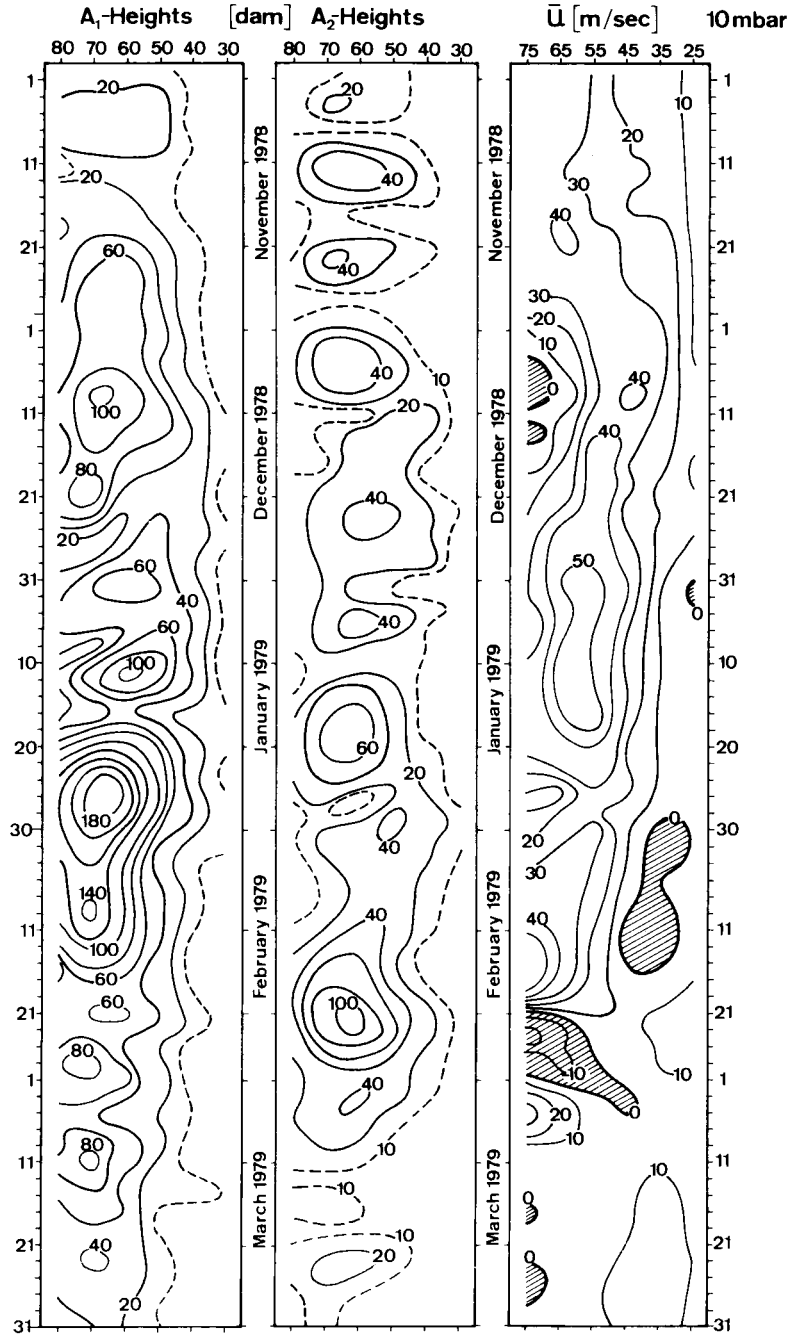

Fig. 6.4. Latitude–time sections at 10 mb from November 1978 through March 1979 for amplitudes (decameters) of the (a) wave-number 1 and (b) wave-number 2 components of the geopotential height field, together with (c) the mean zonal wind ($m\,s^{-1}$). [From Labitzke (1981a), with permission.]

where $f = 2\Omega \sin \phi$ [cf. Eqs. (3.5.2a,d) and (5.2.2)] and

$$\mathbf{F} = [0, -\rho_0 a \cos \phi \, \overline{v'u'}, \rho_0 a \cos \phi f \overline{v'\theta'}/\theta_{0z}] \tag{6.2.2}$$

[see Eq. (5.2.3)]. These equations should provide a reasonable description of slow, large-scale motions in the extratropics.

It was noted in Section 5.2.2, and illustrated in Fig. 5.5, that climatological stationary waves in winter are associated with a field of **F** vectors that indicate propagation out of the troposphere into the stratosphere, followed by an equatorward propagation. (Climatological transient planetary-wave statistics indicate similar features.) Prior to sudden warmings, however, a different behavior may be observed, as illustrated in Fig. 6.5; this shows "integral curves" of **F** (i.e. curves that are everywhere parallel to the local direction of **F**) for 6 days in late February 1979. Here **F** arrows tilt towards the polar upper stratosphere on February 19, 21, and 26 (Figs. 6.5b,c,e), suggesting that planetary waves are being diverted from their climatological equatorward propagation and "focused" at those times into the high-altitude polar cap. (This interpretation is confirmed to some extent by calculations of the refractive index during some sudden warmings: see Section 6.3.2.)

Figure 6.5 also shows contours of $D_F \equiv (\rho_0 a \cos \phi)^{-1} \nabla \cdot \mathbf{F}$, the mean zonal force per unit mass appearing on the right of Eq. (6.2.1a). The poleward focusing of the waves is accompanied by large negative values of this quantity, on the order of several tens of meters per second per day. Owing to the Coriolis term $f\bar{v}^*$ in (6.2.1a) this negative force cannot be equated exactly to the zonal-mean deceleration $\bar{u}_t$, even if $\bar{X}$ is negligible: as explained in Section 3.5, an elliptic equation like Eq. (3.5.7) must generally be solved for $\bar{u}_t$, given suitable boundary conditions. Nevertheless, the deceleration is still found to be large in the neighborhood of large negative values of D_F. (The contributions from $\bar{Q}$ and $\bar{X}$ are likely to be small for these sudden warming events.) By the thermal wind equation [Eq. (6.2.1d)], a rapid temperature rise in the stratosphere must be associated with this rapid deceleration. The observed values of D_F are thus consistent with the observed rapid deceleration and sudden stratospheric warming, and this lends support to the notion that the planetary waves are responsible for bringing about the dramatic mean-flow changes observed during sudden warmings. Further discussion of this idea and more details of the wave and mean-flow dynamics are given in Section 6.3.

6.2.4 *Isentropic Potential Vorticity Maps and Other Quasi-Lagrangian Diagnostics*

Despite the fairly self-consistent dynamical picture offered by the wave mean-flow diagnostics discussed in the previous section, the separation of

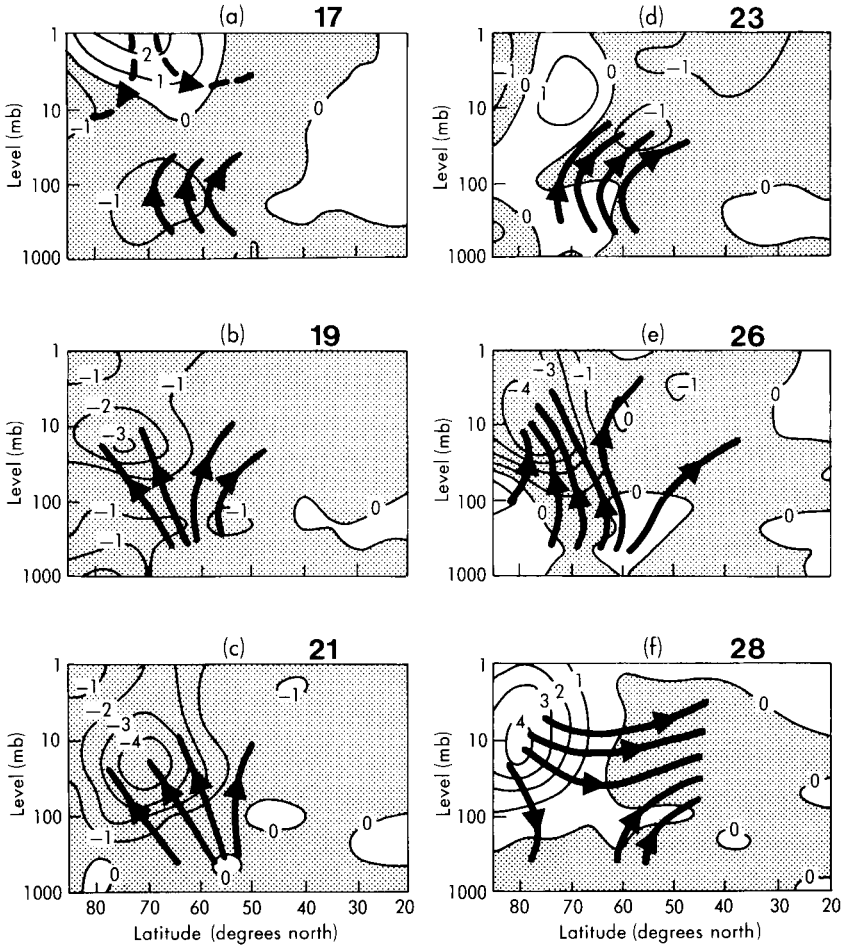

Fig. 6.5. Some integral curves of **F** and contours of D_F in units of 10^{-4} m s^{-2} (negative values stippled) for (a) February 17, (b) February 19, (c) February 21, (d) February 23, (e) February 26, and (f) February 28, 1979. Dashed integral curves are dominated by wave-number 1 contributions to **F**, full curves by wave-number 2 contributions. [From Palmer (1981a), with permission.]

flow quantities into zonal-mean and wave contributions is a rather arbitrary process and may be an unnecessarily complicated (and even misleading) way of viewing the dynamics, especially when wave amplitudes are large. The synoptic maps described in Section 6.2.2 do not make this separation but are, however, purely descriptive, and difficult to interpret in terms of a simple, coherent, dynamical framework.

An alternative approach is to use isentropic maps of Ertel's potential vorticity P (a quasi-Lagrangian tracer) in the hope of obtaining deeper insights into the dynamical processes occurring during sudden warmings. This approach was discussed in Section 5.2.3 in connection with breaking planetary waves, and applied to the diagnosis of the strong cyclonic vortex that was centered off the pole in late January 1979 (in fact at the time of the first minor warming of that year). For comparison with the preceding sections, it is thus of interest to study a sequence of isentropic potential vorticity (IPV) maps for the major warming of late February 1979.

Figure 6.6 shows such a sequence for the 850-K isentropic surface (near 10 mb). It depicts the eastward rotation of the main vortex (here distinguished by high P values) on the days preceding February 17, with material apparently being shed by the breaking process discussed in Section 5.2.3; thereafter, the vortex elongates and splits into two. Similar information can be inferred from the height maps of Fig. 6.3, but the IPV maps may give a more sharply focused view of the dynamics, to the extent that small-scale features in these maps are to be believed. Maps like these have already proved useful in the investigation of numerical models of sudden warmings (see Sections 6.3.2 and 6.3.3); further work will need to be done to establish how much they can tell us about the dynamics of observed warmings and to use them for meaningful quantitative analysis. A step in this direction has been taken by Butchart and Remsberg (1986), who studied the time-evolution of the horizontal areas enclosed by the isopleths of P on the 850-K surface for the 1978–1979 winter.

Long-lived chemical species also supply quasi-Lagrangian information on sudden warmings. For example, Leovy *et al.* (1985), in a study of the three-dimensional transport of ozone during the 1978–1979 winter as given by the LIMS data, found that each minor and major warming in January and February 1979 is associated with a surge of ozone-rich air from the tropical source region into the region of low ozone mixing ratio near the North Pole (see Section 9.5).

6.2.5 Other Observational Aspects

We conclude this brief observational account by mentioning a few further features of observed sudden warmings.

Labitzke (1981a) has noted that a typical "precondition" of many warmings is that a pulse of wave number 1 geopotential amplitude precedes the pulse of wave number 2 (corresponding to a splitting of the main vortex) that occurs at the time of the mean zonal wind reversal. This wave-number 1 pulse may be accompanied by a poleward shifting of the zonal-mean polar night jet. Such a wave-one precursor did indeed occur in 1979, at the

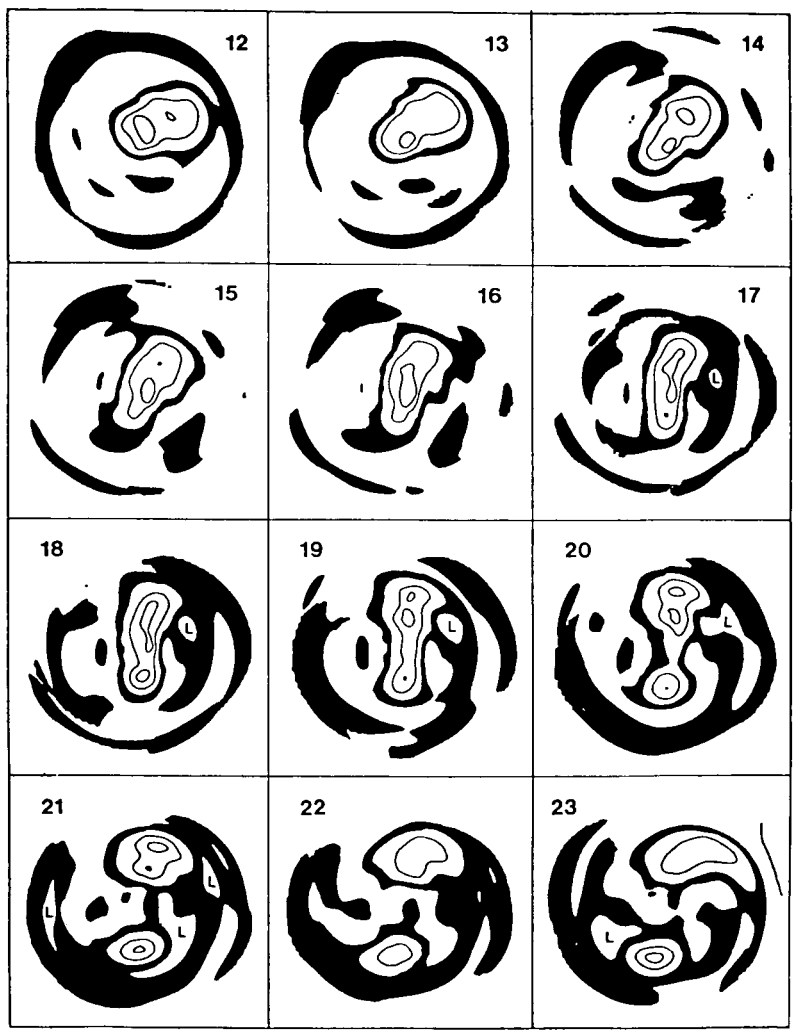

Fig. 6.6. Evolution of Ertel's potential vorticity P on the 850-K isentropic surface for part of February 1979. The north pole is at the center of each square; the sides of each square are tangent to 20°N, and no data are plotted south of that latitude. The contour values are given by $1.3 \times 10^{-4} \times n$ K kg^{-1} m^2 s^{-1} (corresponding approximately to $2 \times 10^{-4} \times n$ K m^{-1} s^{-1} in the scaled units of Figs. 5.6b and 6.12, where P is divided by gH/p_0), with $n = 0$, 2 (outer boundary of black strip), 3 (inner boundary of black strip), 5 and 7. Values increase monotonically towards the center of each plot unless an "L" is present, indicating a local minimum. [From Dunkerton and Delisi (1986), with permission.]

time of the minor warming of January 25 (see Fig. 6.4a and also Fig. 5.6), and was followed by a northward movement and tightening of the polar night jet (see also Section 6.3.3). However, these events took place an unusually long time before the major warming of late February. Other major warmings (e.g., January 1970 and January 1977) take yet a different form, in that no significant development of wave number 2 occurs. On the other hand, that of January 1985 is unusual in involving wave number 2 from the outset, with no significant wave-number 1 precursor.

Another feature that is often noted at the time of major warmings is a simultaneous cooling of the polar mesosphere and low-latitude stratosphere. Dynamical reasons for this behavior will be discussed in Section 6.3.2.

It has long been known that sudden stratospheric warmings tend to occur roughly concurrently with "blocking events," which involve strong, long-lasting, quasi-stationary distortions of the tropospheric flow. The causes of blocking are not well understood, still less their connections with sudden warmings: some theoretical suggestions for cause-and-effect relationships between the two phenomena are mentioned in Section 6.3.4.

A further possible correlation has been noted between the incidence of major stratospheric warmings and the phase of the equatorial quasi-biennial oscillation (see Section 8.2), in which the tropical stratospheric winds reverse from easterly to westerly and back, with a period of about 27 months. It was mentioned above that major warmings only occur about once in every two northern winters; more often than not, these are the winters when equatorial stratospheric winds are easterly. A speculative mechanism for this relationship is noted in Section 6.3.4.

Although the preceding sections have mostly concentrated on major warmings, minor stratospheric warmings (in which a zonal-mean wind reversal below 10 mb does not occur) are also of considerable interest. Examples include those of late January and early February 1979, and a Southern Hemisphere case that has received much attention is that of July 1974, in which considerable *midlatitude* temperature increases and wind decelerations (from an initially very strong westerly flow) were observed in the upper stratosphere. Labitzke (1981b) defines *Canadian warmings* as events that result from an anomalous strengthening of the Aleutian anticyclone and may possibly reverse the mean poleward temperature gradient north of 60°N. *Final warmings* are those events that are followed not by a reversion of stratospheric conditions to the usual winter pattern but by a transition to the summer structure of warm temperatures and easterly winds. Late winter in the Southern Hemisphere stratosphere is notable for a downward and poleward shifting of the polar night jet; the spring transition tends to occur later than in the Northern Hemisphere, and is accompanied by fluctuating wave-number 1 activity.

6.3 Theoretical Modeling of Sudden Warmings

6.3.1 Matsuno's Model

Early attempts to explain sudden warmings investigated the possibility that they may be primarily due to the baroclinic instability of the polar night jet in the zonally averaged winter stratospheric flow. A further development of this idea was the examination of the possible barotropic instability of the large-scale zonally asymmetric polar vortex. Although some unstable modes were found in these studies, their growth rates and spatial scales were too small to explain the rapid development and large horizontal and vertical scales of observed warmings.

It is now generally accepted that the essential dynamical mechanism reponsible for sudden warmings involves the upward propagation from the troposphere of planetary (Rossby) waves and their interaction with the mean stratospheric flow. This hypothesis was first proposed by Matsuno (1971), who tested it using a numerical model of the stratosphere. Most subsequent "mechanistic" models of sudden warmings have been generalizations of Matsuno's.

The dynamical reasoning on which Matsuno's model is based can be summarized as follows. He supposed that the mean-flow changes observed during sudden warmings, including the mean zonal flow deceleration and the mean temperature rise near the pole, are attributable to the nonlinear rectified effects of vertically propagating planetary waves forced in the troposphere by large-scale disturbances there. In one of the first applications of the Charney–Drazin nonacceleration theorem (see Section 3.6), he noted that steady, conservative, linear planetary waves are incapable of inducing such mean-flow changes, and therefore sought conditions under which the nonacceleration theorem is violated. One of these occurs when a transient packet of planetary waves first starts to propagate upward: at the leading edge of such a packet, wave amplitudes are growing in time. The effects of this wave growth can conveniently be described using the ideas presented in Section 3.6 (although these were not explicitly used by Matsuno); a simple schematic picture of the process is given in Fig. 6.7. If the wave-activity density

$$A \equiv \tfrac{1}{2}\rho_0 \overline{\eta'^2}\,\bar{q}_y$$

[a quadratic measure of wave amplitude: see Eq. (3.6.10)] is introduced, where η' is the northward parcel displacement and $\bar{q}_y$ is the northward mean potential vorticity gradient (assumed positive), then $\partial A/\partial t > 0$ at the leading edge of the packet. If the waves are considered to be linear and conservative, then by the generalized Eliassen–Palm theorem [Eq. (3.6.2)],

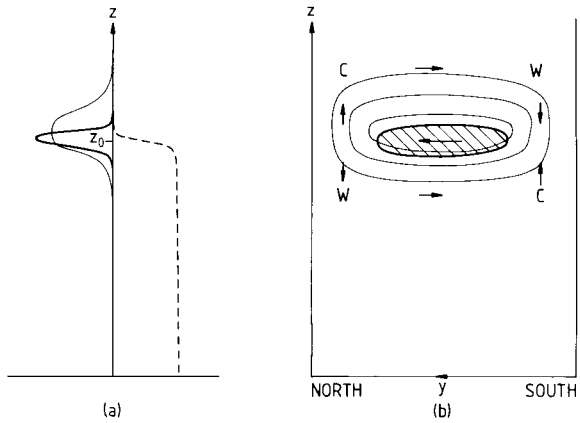

Fig. 6.7. Schematic description of the interaction with the zonal-mean flow of a transient, vertically propagating planetary-wave packet that is maintained by continual forcing at $z = 0$ in a zonal channel, to illustrate aspects of Matsuno's model of sudden warmings. (a) Midchannel height profiles of wave activity A (dashed), EP flux divergence $\nabla \cdot \mathbf{F} = -\partial A/\partial t$ (heavy line), and acceleration $\bar{u}_t$ (thin line); z_0 is the height reached by the leading edge of the packet, and is of order (vertical group velocity) × (time since packet started at $z = 0$). (b) Latitude-height cross section showing region where $\nabla \cdot \mathbf{F} < 0$ (hatched), contours of induced acceleration $\bar{u}_t$ (thin lines), and induced residual circulation $(\bar{v}^*, \bar{w}^*)$ (arrows). Regions of warming (W) and cooling (C) are also indicated.

$\nabla \cdot \mathbf{F} = -\partial A/\partial t$, where $\mathbf{F}$ is the Eliassen-Palm flux, and a patch of negative $\nabla \cdot \mathbf{F}$ is therefore to be expected near the front of the packet. As noted in Section 6.2.3, this implies a negative force per unit mass on the zonal-mean flow, the response to which will include a westward mean-flow acceleration in the vicinity of the patch of negative $\nabla \cdot \mathbf{F}$. A mean temperature increase will occur on the lower poleward flank of this patch. Thus, the mean-flow changes associated with the leading edge of a transient planetary-wave packet are qualitatively similar to those observed during a sudden warming.[1]

Matsuno tested this qualitative mechanism with a quasi-geostrophic numerical model based on a "quasi-linear" separation of the waves and the mean flow. Thus, to represent the waves he used the linearized spherical-geometry potential vorticity equation [Eq. (5.3.1)] [see also Eqs. (5.3.2)-(5.3.4)], including the time-derivative terms, together with a suitable time-

[1] Matsuno also considered the mean-flow changes associated with vertically propagating planetary waves encountering a descending horizontal critical level (see Section 5.6), with all the transience being concentrated in a thin critical layer. However, it is now known that such horizontal critical levels do not normally appear in the stratosphere during sudden warmings. On the other hand, approximately vertical critical surfaces may encroach from low latitudes and play a somewhat different role in the dynamics of some sudden warmings: see Section 6.3.2.

dependent lower boundary condition at $z = 10$ km to represent the switch-on of the tropospheric forcing associated with some unspecified large-scale transient disturbance. The geopotential disturbance Φ' given by solving this system under a specified mean-flow configuration is then substituted into the quadratic wave-forcing of the mean flow. In terms of the formulation of Section 6.2.3, this amounts to evaluating **F** by means of Eq. (6.2.2) and substituting into Eq. (6.2.1a). The mean-flow equations [Eqs. (6.2.1)] can then be solved for the tendencies $\bar{u}_t$ and $\bar{\theta}_t$, and the mean flow can be updated after one time step. The process is then repeated by solving Eq. (5.3.1), given the new mean-flow configuration and the forcing at the new time step, and so on. This quasi-linear approach excludes the possibility of interactions between waves of different zonal wave numbers.

Matsuno performed a series of experiments in which transient disturbances of a single zonal wave number, $s = 1$, 2, or 3, were forced at the lower boundary and propagated into a model stratosphere whose initial zonal wind structure $\bar{u}$ resembled the observed climatology. Interaction between the waves and mean flow led to an evolution of the mean wind and temperature fields, which in some cases was reminiscent of that occurring during sudden warmings. For example, with an $s = 2$ forcing, the climatological westerlies gave way to easterlies after about 20 days of integration, followed by a rapid polar temperature increase (of over 80 K) below about 50 km and a decrease above that level: see Fig. 6.8. The corresponding geopotential height charts (Fig. 6.9) showed an elongation and splitting of the polar cyclonic vortex in the midstratosphere, somewhat similar to the February 1979 observations given in Figs. 6.3a–d. This was followed by an amalgamation of the model's subsidiary anticyclones to form a weak anticyclonic polar vortex that was not apparent in the 1979 observations of Figs. 6.3e,f.

6.3.2 The Use of Lagrangian and Eliassen–Palm Diagnostics in Models of Sudden Warmings

In recent years there have been numerous model studies of the sudden warming phenomenon, mostly based on quasi-linear models of the type discussed in the previous section. A variety of diagnostic techniques have been used for describing and interpreting the model dynamics. Some of these techniques have also been used in observational studies of warmings and have been mentioned briefly in Section 6.2; others are less easily applied to observational data and have mainly been restricted to model experiments.

An example of the latter type of diagnostic is the air-parcel trajectory. Accurate Lagrangian calculations of this kind are difficult to achieve for

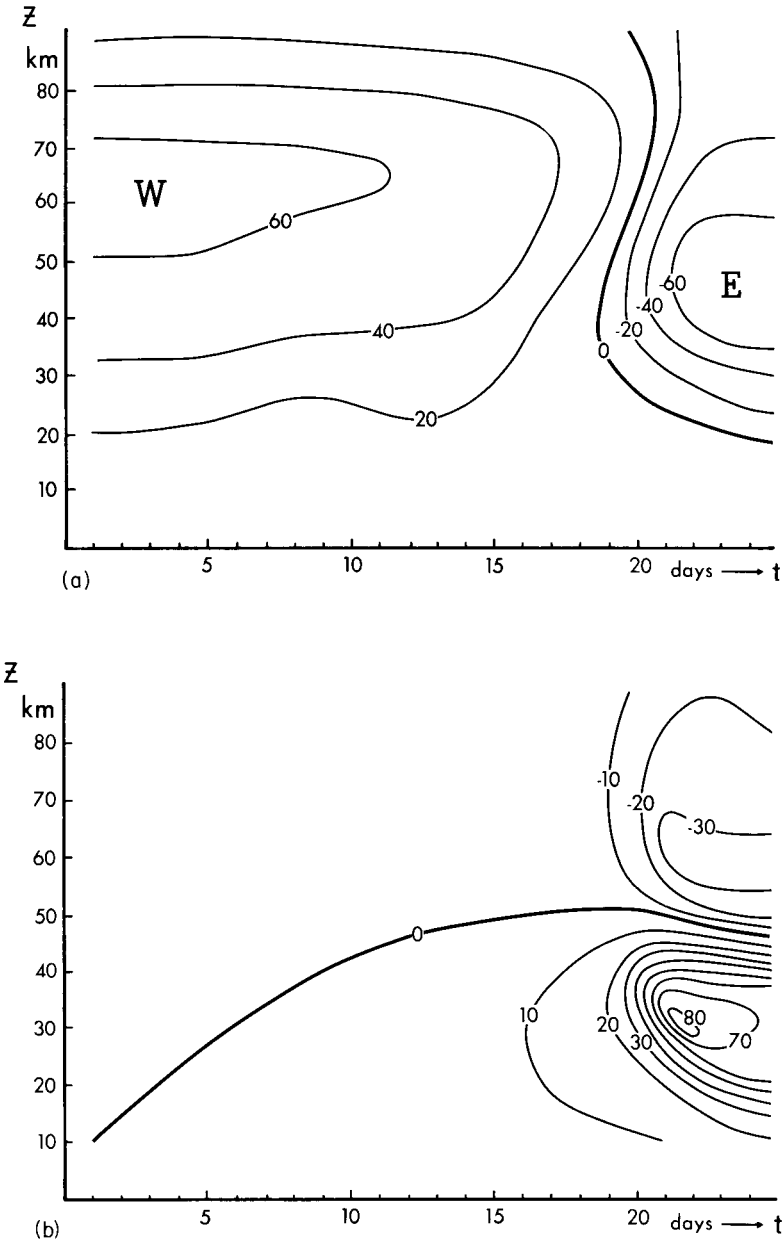

Fig. 6.8. Time–height sections for (a) zonal-mean zonal wind at 60°N and (b) temperature at the north pole, from the model wave-number 2 warming in Matsuno's (1971) model. [From Matsuno (1971). American Meteorological Society.]

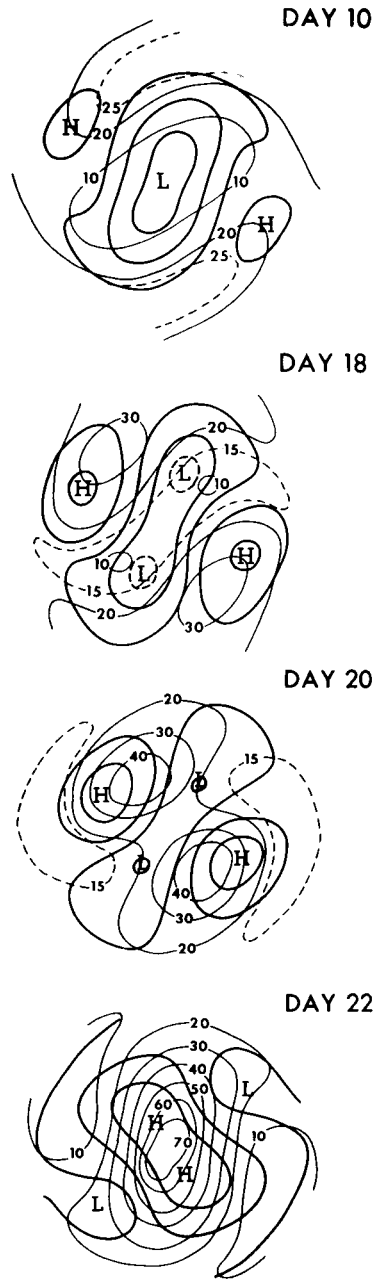

Fig. 6.9. Polar stereographic plots of the time evolution of the temperature (thin curves, marked in degrees Celsius) and height (thick curves, 500-m contour intervals) for the isobaric surface $z = 30$ km ($p \approx 13$ mb) for the same model warming as in Fig. 6.8. The contours cover the area north of 30°N. [From Matsuno (1971). American Meteorological Society.]

extended periods from stratospheric observational data, but can be performed routinely in models. Using an adaptation of the model of Holton (1976), Hsu (1980) performed an extensive study of the motion of air parcels during model sudden warmings (see also Section 9.7.4). In her "wavenumber 2" experiment the basic cyclonic vortex elongated, split into two at the peak of the warming (around days 22–24), and was replaced by a polar anticyclone, rather as in Matsuno's calculation (Fig. 6.9). These developments were reflected in the motion of air parcels, as depicted in Fig. 6.10, which shows the horizontal projections of a set of air parcels that were originally distributed uniformly around the 60°N latitude circle at $z = 32$ km. Figure 6.11 shows the meridional projections of the same set of parcels and demonstrates the general descent of midstratospheric air parcels near the pole at the peak of the warming, in agreement with observational evidence (Mahlman, 1969). Similar calculations for parcels originally located at 30°N, $z = 30.8$ km, indicated more complicated behavior; these low-latitude parcels became wrapped around the cut-off anticyclones, with some eventually penetrating to the polar region.

Hsu also presented maps of Ertel's potential vorticity P on the $\theta = 850$ K isentropic surface, near 30 km altitude: see Fig. 6.12. The isentropic isopleths of P differ from material contours in this model since θ and P are not conserved, not only because of the explicit inclusion of diabatic and frictional processes but also because of the quasi linearity, which neglects terms quadratic in wave amplitude in the disturbance equations. Nevertheless, qualitative similarities between the behavior of the strings of air parcels and the isentropic P contours provide useful cross-checks on the calculations and some guidance for the interpretation of IPV maps derived from observational data (cf. Section 6.2.4).

Further diagnosis of the same model sudden warming was provided by Dunkerton *et al.* (1981). They showed that the descent of midstratospheric air parcels near the pole during the simulation (Fig. 6.11) was closely followed by the downward motion of the 850-K isentrope relative to the isobars (or z surfaces) near the pole, thus verifying that the process was approximately adiabatic. They also noted that this polar descent was accompanied by ascent of the isentrope in the tropics, since the atmospheric mass beneath the isentrope remained roughly constant. Such a "see-saw" effect presumably accounts for the cooling of the equatorial stratosphere that is commonly observed at the time of polar warmings (see Section 6.2.5).

Dunkerton *et al.* also used information like that shown in Fig. 6.11 to calculate the meridional and vertical components $(\bar{v}^L, \bar{w}^L)$ of the Lagrangian-mean velocity during the warming. These are given by the motion of the center of mass, as viewed in the meridional plane, of the initially zonal ring of parcels (see Section 3.7.1). As expected from the

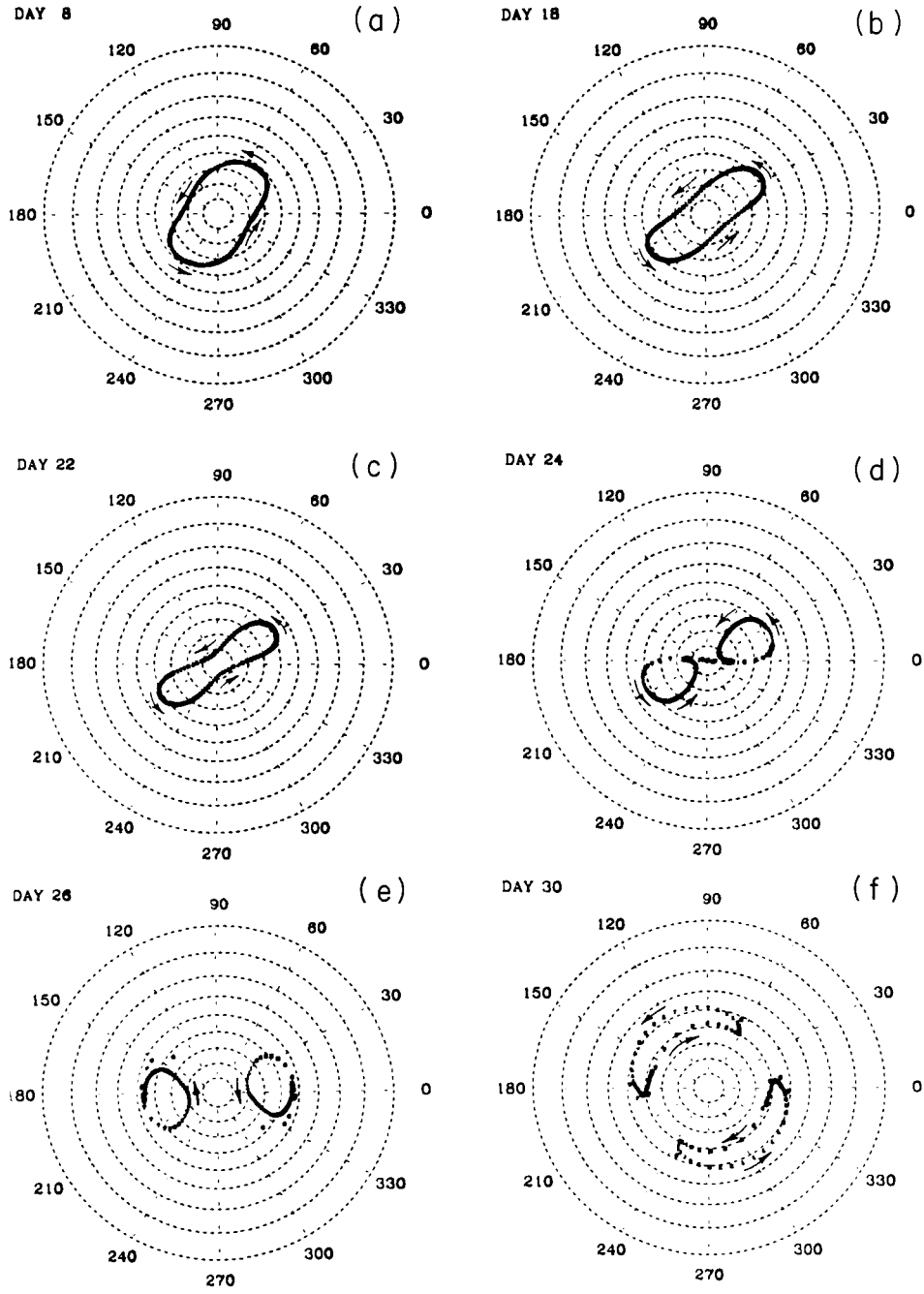

Fig. 6.10. Polar stereographic plots of the horizontal projections of a set of air parcels for 6 days during the model warming described by Hsu (1980). The parcels were distributed uniformly around the 60°N latitude circle at $z = 32$ km on day 0. Latitude circles, at 10° intervals, and longitudes, at 30° intervals, are shown dashed. [From Hsu (1980). American Meteorological Society.]

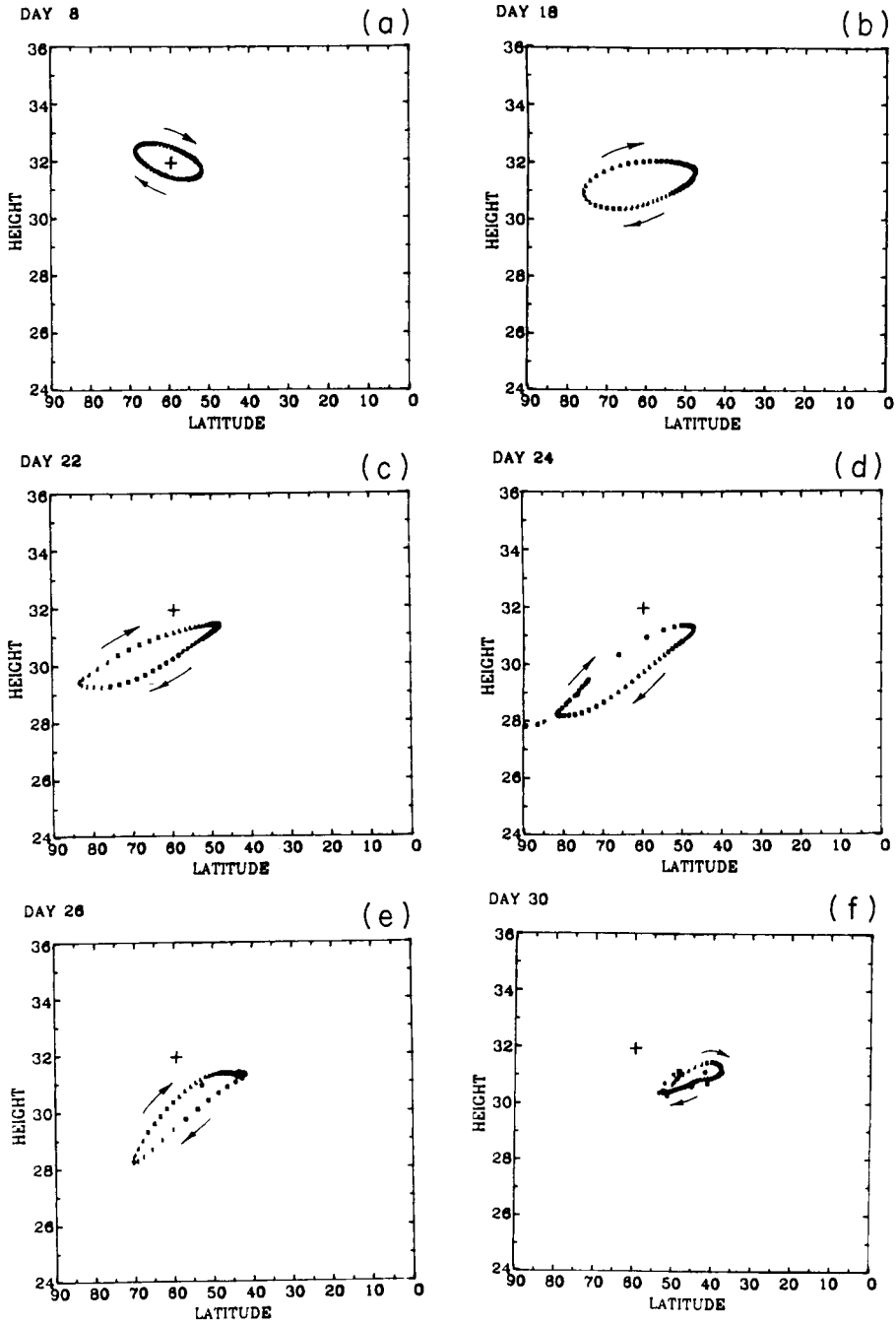

Fig. 6.11. As in Fig. 6.10, but for projections in the meridional (ϕ, z) plane. The initial position of the line of parcels is marked by a plus sign. [From Hsu (1980). American Meteorological Society.]

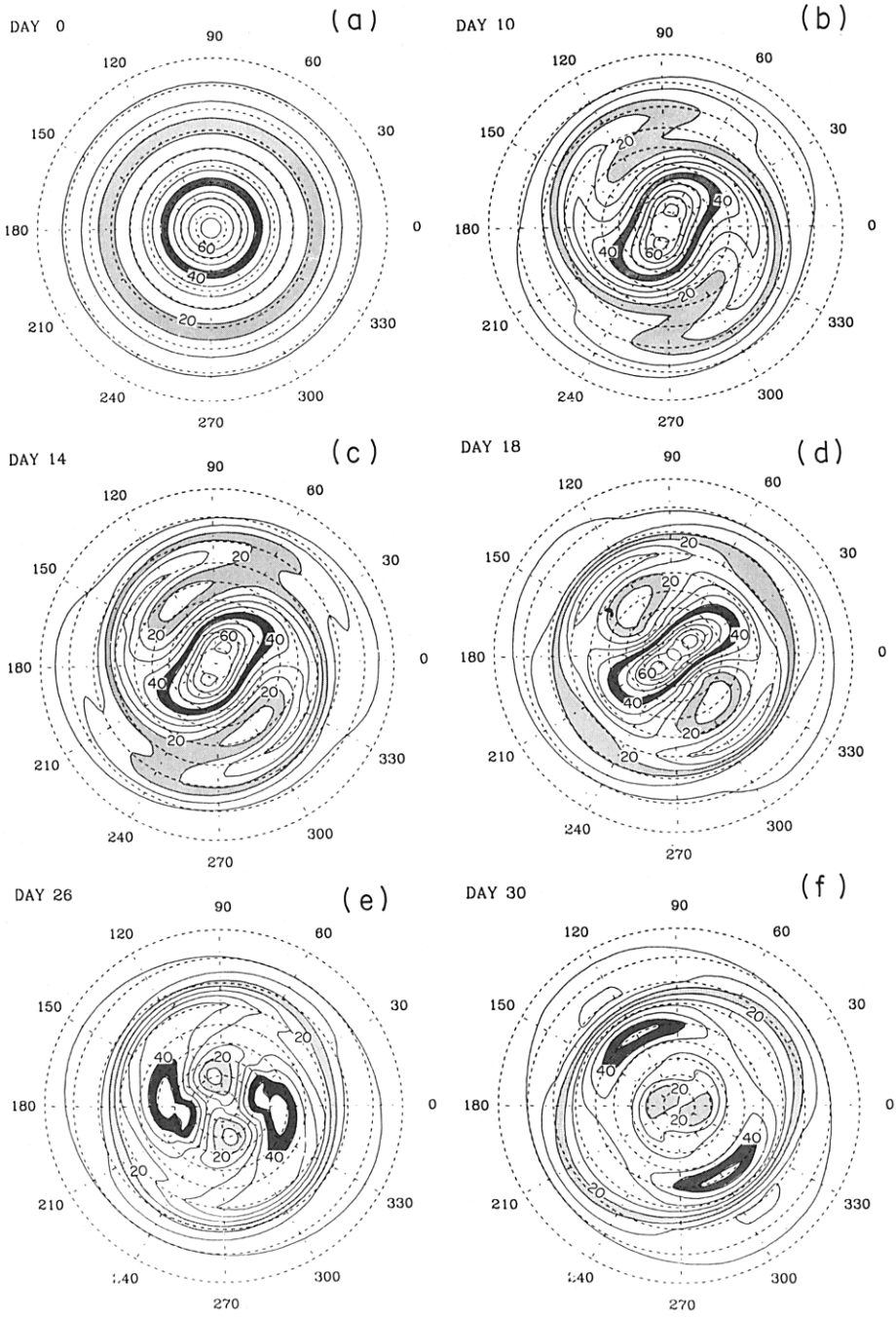

Fig. 6.12. Isopleths of Ertel's potential vorticity divided by gH/p_0 (where $H = 7$ km and $p_0 = 1000$ mb) on the 850-K isentropic surface for 6 days during the model warming described by Hsu (1980). The units are 10^{-5} K m^{-1} s^{-1} and the contour interval is 5 units. The area between the contours for 15 and 20 units is shaded lightly, and that between the contours for 40 and 45 units is shaded heavily. [From Hsu (1980). American Meteorological Society.]

previous discussion, Lagrangian-mean descent was found in high latitudes and ascent in low latitudes in the middle stratosphere. (This contrasts with an Eulerian-mean *ascent* in high latitudes.) Another feature of these calculations was a strong mass divergence of the Lagrangian-mean velocity over a deep layer, associated primarily with the north–south dispersion of air parcels. The authors also indicated ways in which the generalized Lagrangian-mean theory may fail at large wave amplitude, and they explored the possible use of an alternative "modified Lagrangian-mean" theory, based on averages around the quasi-Lagrangian isentropic P contours.

Calculations of Eliassen–Palm diagnostics (see Section 6.2.3) for the model warming were also performed by Dunkerton *et al.* Figure 6.13 shows "EP cross sections," with arrows representing the EP flux $\mathbf{F}$, and contours of its divergence, for several days. The arrows split into two separate branches by day 22, one going equatorward and one more vertically. Dunkerton *et al.* attributed this bifurcation and polar focusing to the presence of a zero-wind line (a critical surface for stationary waves, indicated by the heavy curve in the plot for day 15) encroaching from low latitudes. They argued that the model may have produced a qualitatively correct reflection of the waves from such a surface, as expected from nonlinear Rossby-wave critical layer theory (Section 5.6). This reflection would tend to deflect EP arrows upward, as seen on day 22.

Dunkerton *et al.* examined the wave, mean-flow interaction in the model using the transformed Eulerian-mean equations [Eqs. (6.2.1)]. The vertically focused planetary waves exhibited significant negative values of $\nabla \cdot \mathbf{F}$ (Fig. 6.13), due primarily to wave transience (see Section 6.3.1); the associated zonal force per unit mass $D_F \equiv (\rho_0 a \cos\phi)^{-1} \nabla \cdot \mathbf{F}$ was particularly large in magnitude in the high-altitude polar cap where both ρ_0 and $\cos\phi$ were small. The effects of this force were redistributed to some extent by the residual circulation $(\bar{v}^*, \bar{w}^*)$, but nevertheless produced decelerations (leading to the poleward migration of low-latitude easterlies) and the associated polar temperature increases. Diabatic effects were small on the short timescales involved here, and the net acceleration $\bar{u}_t$ tended to resemble D_F as a function of latitude: see Fig. 6.14. (A quite different balance of terms occurs when long-time averages are taken: see Section 7.2.) The mass stream function for the residual circulation on day 27 (Fig. 6.15) showed a descent in stratospheric polar regions roughly similar to the Lagrangian-mean descent there, and weak ascent in the polar mesosphere. The primary forcing of this circulation was D_F again, as can be seen by constructing the spherical analog of Eq. (3.5.8) from Eqs. (6.2.1) and noting that $\bar{Q}$ and $\bar{X}$ were small. The weak mesospheric ascent implies cooling there: this may be analogous to the observed mesospheric cooling noted in Section 6.2.5.

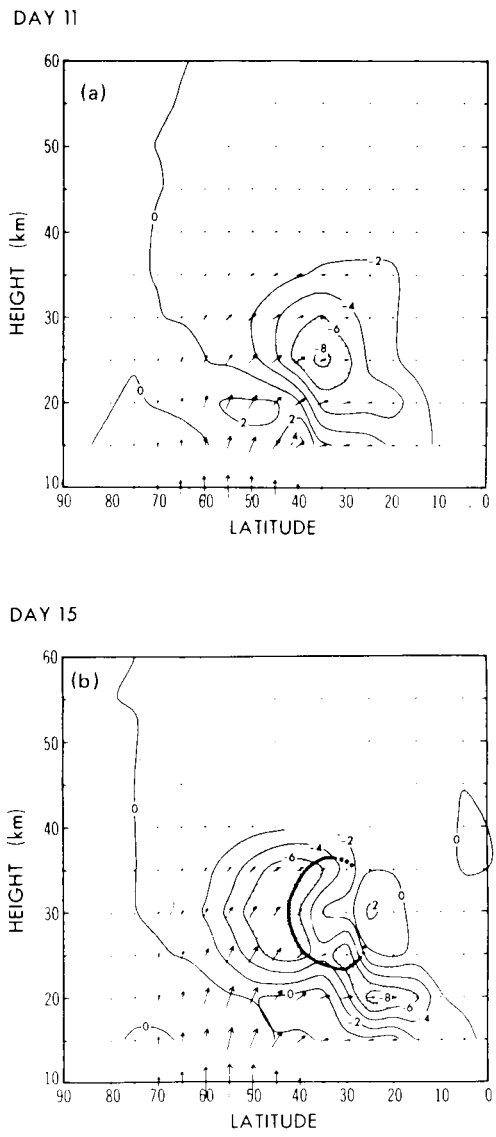

Fig. 6.13. Eliassen–Palm cross sections (in the ϕ, z plane) for (a) day 11, (b) day 15, and (c) day 22 of the evolution of Hsu's (1980) model. The heavy curve in (b) shows part of the zero-wind line for day 15. The contours represent values of $(\cos\phi)(\nabla \cdot \mathbf{F})$, in units of $a\rho_0(0) \times 10^{-7}$ m s^{-2}. The arrows represent the vector $(\cos\phi)\mathbf{F}$, scaled such that the distance occupied by 10° in ϕ represents a value $5.56\rho_0(0)a$ m^2 s^{-2} of $(\cos\phi)F^{(\phi)}$ and that occupied by 10 km in z represents a value $0.05\rho_0(0)a$ m^2 s^{-2} of $(\cos\phi)F^{(z)}$. [From Dunkerton *et al.* (1981). American Meteorological Society.] (*Figure continues.*)

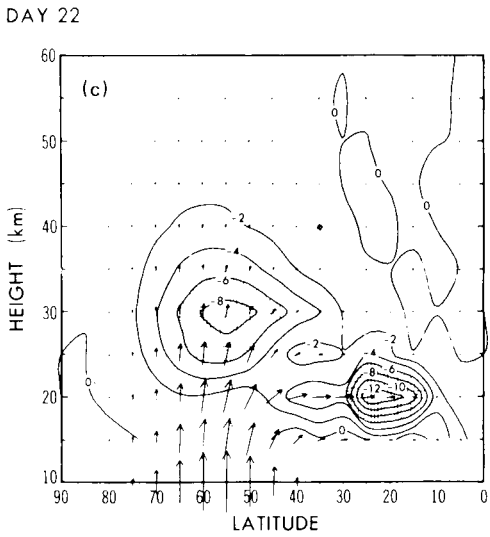

Fig. 6.13 (*continued*)

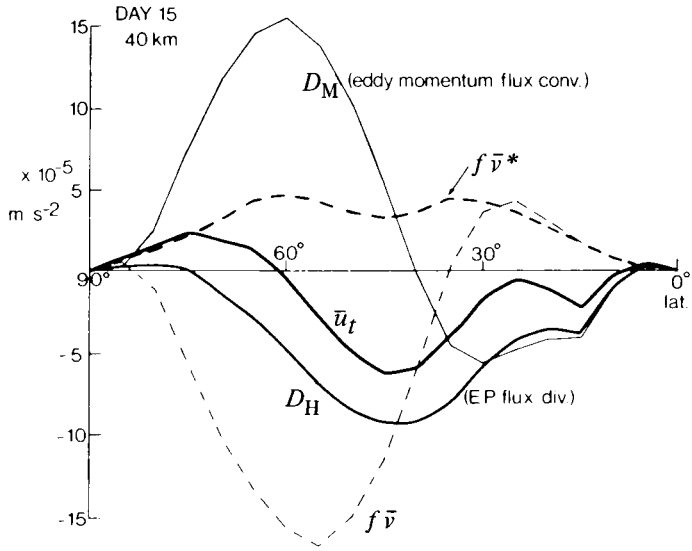

Fig. 6.14. Heavy curves: typical balance of terms in Eq. (6.2.1a) from the model warming described by Hsu (1980). Light curves: principal terms in the conventional momentum equation, Eq. (3.3.2a), where $D_M \equiv -(a\cos^2\phi)^{-1}(\overline{v'u'}\cos^2\phi)_\phi$. [From Dunkerton *et al.* (1981). American Meteorological Society.]

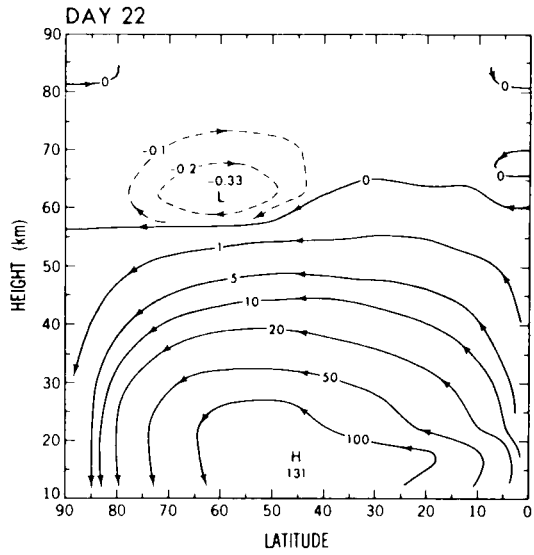

Fig. 6.15. Mass stream function $\bar{\chi}^*$ associated with the residual circulation on day 22 of the model warming described by Hsu (1980); $\bar{\chi}^*$ is defined such that $\partial\bar{\chi}^*/\partial\phi = \rho_0 a(\cos\phi)\bar{w}^*$ and $\partial\bar{\chi}^*/\partial z = -\rho_0(\cos\phi)\bar{v}^*$. Contour values are to be multiplied by $\rho_0(0)$ times 1 $m^2\ s^{-1}$. The contour interval is nonuniform. [From Dunkerton *et al.* (1981). American Meteorological Society.]

A more complex series of experiments was performed by Butchart *et al.* (1982). They first used a fully nonlinear model of the stratosphere, starting with the observed zonal-mean wind field for February 16, 1979, and forced at 100 mb by the observed height field, to simulate the observed warming of late February. Most of the main features of the observed warming were quite well reproduced. The model was then simplified in various ways, to try to isolate the essential dynamical behavior of the simulation, and hence of the observed warming. For example, runs were carried out using only the wave-number 2 Fourier component of the lower boundary forcing, or using idealized stationary or steadily progressing forcing there, and replacing the observed initial winds by their climatological counterparts. EP fluxes were used as diagnostics, as was the planetary-wave refractive index squared, n_k^2 (see Section 4.5.4). Although the strict derivation of the latter quantity assumes the waves to be steady, it nevertheless suggests interpretations of the bifurcation and polar focusing of transient-wave EP fluxes during sudden warmings, by showing the locations of possible barriers to propagation, where n_k^2 is small or negative.

Butchart *et al.* concluded that an important condition for the warming to occur was that the initial mean zonal wind should take its observed,

"preconditioned," nonclimatological form, with the polar night jet at 10 mb at about 75°N, instead of the usual 60°N. They also found it necessary that the zonal wave-number 2 component of the forcing should have the observed eastward phase speed of about 10° of longitude per day. On the other hand, they concluded that wave–wave interactions (between wave numbers 1 and 2, say) were not crucial to the development of the warming after February 16.

Several other studies have been made of sudden warmings using global circulation models of the troposphere and stratosphere, in both "forecast" and "general circulation" modes; some examples are discussed in Section 11.3.

6.3.3 Wave–Wave Interactions

The model introduced by Matsuno (1971) and most of the others discussed above were quasi-linear: they included the interaction between a single zonal wave component and the zonal-mean flow, but not the interactions between waves of different zonal wave number. The basic model used by Butchart *et al.*, on the other hand, was fully nonlinear but, as mentioned above, they found that wave–wave interactions were not essential to the simulation of the major warming of late February 1979, given the observed nonclimatological flow as an initial condition on February 16.

Several model studies have sought to elucidate the role of wave–wave interactions in the overall development of sudden warmings. For example, Hsu (1981) found that inclusion of such interactions tended to result in increased wave activity, stronger polar warmings, and more rapid mean-flow decelerations. As expected, the Lagrangian behavior of air parcels was more complex than in the quasilinear cases depicted in Figs. 6.10 and 6.11.

Palmer and Hsu (1983) used the same nonlinear model to study aspects of the stratospheric behavior in late January and early February 1979. After the minor warming associated with the wave-number 1 "precursor" in late January (see Section 6.2.5), the polar temperatures dropped once more (see Fig. 6.1) in a "sudden cooling," and high-latitude easterlies in the upper stratosphere accelerated rapidly to form a tight westerly polar night jet near 75°N by February 3. (This jet later relaxed, then formed again by February 16, just before the major warming.) Palmer and Hsu concluded that nonlinear wave–wave interactions were important during the cooling period, a period that included the essentially nonlinear planetary wave breaking event studied by McIntyre and Palmer (1983, 1984; see also Section 5.2.3). Synoptically, the sudden cooling corresponds to the movement of the cyclonic vortex in Fig. 5.6a back to a more pole-centered position. The fact that the polar night jet was further poleward after this event than before may be a

result of the erosion of potential vorticity while the vortex was in a displaced position.

Matsuno (1984) added wave-wave interactions to his earlier quasi-geostrophic model and carried out some experiments to study the differences between minor and major warmings. With lower boundary forcing in wave-number 1 of geopotential height amplitude 300 m at $z = 10$ km, on an already distorted flow, a wave-number 1 type minor warming ensued, while if the forcing amplitude was increased to 400 m, a major warming followed. The differences were diagnosed in terms of potential vorticity, among other things. In each case, low, tropical potential vorticity was advected into the polar stratosphere and high potential vorticity contours were displaced off the pole, as shown in Fig. 6.16 (cf. also Fig. 6.12). However, in the major warming irreversible mixing of potential vorticity occurred on a hemispheric scale, leading to polar easterlies, whereas in the minor warming irreversible mixing was restricted to low and middle latitudes, leading to a tight westerly polar jet. (In the latter case any mixing near the pole was reversible, being due to the temporary juxtaposition of air masses of differing potential vorticities that later returned toward their original latitudes.) This behavior was anticipated on theoretical grounds by McIntyre (1982), and is illustrated schematically in Fig. 6.17. In the model the irreversible mixing took place by the stretching-out of potential vorticity contours to such an extent that dissipation, whether explicitly represented or implicit in the model's zonal truncation, could take over: the inclusion of wave number 2 in addition to wave number 1 apparently expedited this process.

6.3.4 The Role of Instabilities

The theories discussed above have prescribed some form of planetary-wave forcing near the tropopause as a lower boundary condition. The temporal growth of this forcing leads to amplification of the planetary waves propagating into the stratosphere; these waves in turn interact with the mean stratospheric flow, and the warming follows. It has been assumed that the increased lower-boundary forcing is related in some way to large-scale tropospheric disturbances, such as blocking events, which have been taken to develop more or less independently of happenings in the stratosphere.

An alternative approach, which seeks to include an explanation of the planetary-wave amplification, rather than regarding it as imposed externally, was considered by Plumb (1981). Using both an asymptotic analytical theory and a simple quasi-linear model, which included a highly idealized troposphere and vertical sidewalls parallel to the x axis and no meridional wave propagation, he found that wave amplification and stratospheric warmings

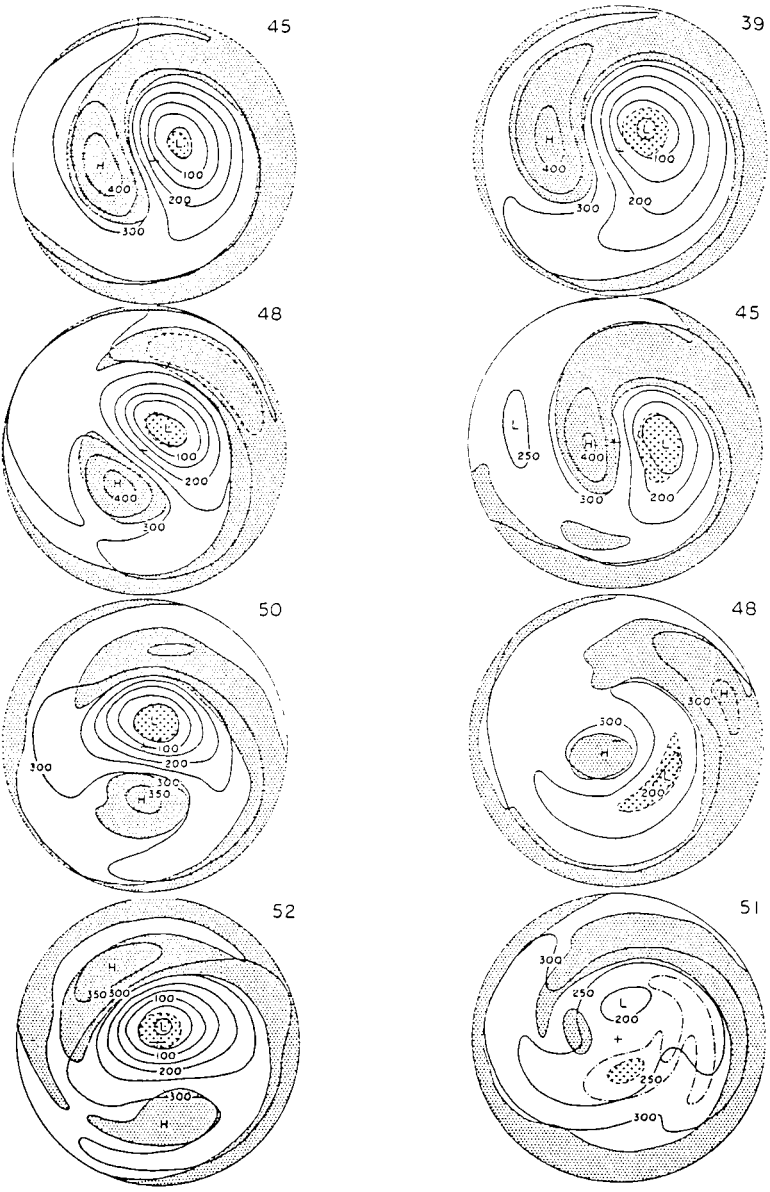

Fig. 6.16. Polar stereographic plots of the evolution of the isobaric height (solid curves, at 500-m intervals) and quasi-geostrophic potential vorticity $q'_{(M)}$ [see, e.g., Eq. (5.3.3)] at $z = 40$ km, for several days from the model experiments of Matsuno (1984). Left-hand diagrams, minor warming; right-hand diagrams, major warming. Small values of $q'_{(M)}$ ($< 2 \times 10^{-4}\,s^{-1}$) are shaded. Stippled areas indicate large values of $q'_{(M)}$ ($> 2.5 \times 10^{-4}\,s^{-1}$ in minor warming and $> 2 \times 10^{-4}\,s^{-1}$ in major warming). Day numbers marked next to each frame.

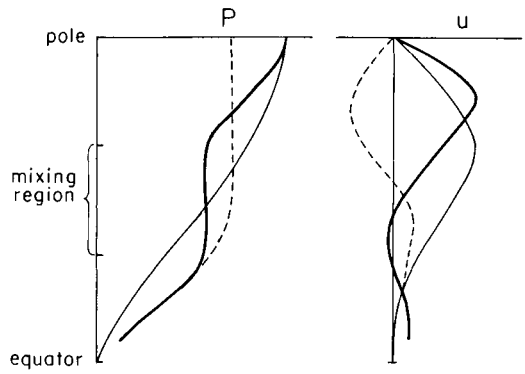

Fig. 6.17. Schematic latitudinal distributions of Ertel's potential vorticity P on an isentropic surface and corresponding polar night jet profiles u in the model of Matsuno (1984). Thin solid curves represent the initial states, and thick solid curves and dashed curves show distributions after minor and major warmings, respectively. [From McIntyre (1982) and Matsuno (1984).]

could occur through a nonlinear near-resonance between a topographically forced stationary wave and a free traveling wave in a zonal-mean flow $\bar{u}$ that depends on height.

Plumb's mechanism works essentially as follows. Suppose that under given mean-flow conditions a free normal mode has a small horizontal phase speed, so that it is nearly in resonance with the stationary, topographically forced wave. If this forced wave grows a little in amplitude, it can under certain conditions bring about mean-flow changes that make the free mode's phase speed even smaller and thus even nearer to resonance. This "self-tuning" leads to further stationary-wave growth, and a positive feedback is set up; as a result, large wave-amplification can occur, and large mean-flow changes resembling those associated with warmings. It is interesting that some observational studies, including those of the minor warming of January 1979, note a slowing-down of a traveling wave prior to the warming.

Plumb's "self-tuned resonant cavity instability" is analogous to the "topographic instability" studied by Charney and DeVore (1979) in another context. It has been discussed in detail by McIntyre (1982), who considered the likelihood of suitable resonant cavities occurring in the real atmosphere, perhaps aided by the presence of reflecting nonlinear critical layers (see Section 5.6) at low latitudes. For quasi-stationary waves, these critical layers surround zero-wind lines where $\bar{u} = 0$: the modulation of the positions of such lines by the equatorial quasi-biennial oscillation might affect the properties of any such cavity, and hence account for the roughly biennial occurrence of major warmings noted in Section 6.2.5 (but see Section 12.5.1).

McIntyre also discussed the possibility that tropospheric blocking may be part of the same nonlinear resonance phenomenon. Further observational and theoretical work will be needed to test these ideas (see also Section 12.4).

Another, quite different, type of instability that may be significant during certain phases of the development of sudden warmings is shear instability, either barotropic or baroclinic. It was noted in Section 6.3.1 that early theories of the shear instability of the global-scale stratospheric circulation were unable to find unstable planetary-scale disturbances resembling the wave-growth that occurs during warmings. However, once the stratospheric flow has become strongly disturbed by a planetary-wave, mean-flow interaction, the drawing-out of tongues of anomalous potential vorticity (as for example in several frames in Fig. 6.6, where tongues of high P become sandwiched between regions of low P) implies sign changes in $\partial P/\partial y$ on isentropic surfaces. If these occur over a sufficient spread of longitudes, for a roughly zonally aligned tongue, it is plausible that the flow may support *local* shear instabilities, owing to the local breakdown of the Charney-Stern criterion (Section 5.5.1).

6.4 Conclusions

This chapter has presented a variety of observational findings and theoretical ideas relating to stratospheric warmings. An attempt will now be made to synthesize these into an overall picture of the phenomenon.

A major sudden warming is initiated by an anomalous growth of a planetary-wave disturbance (mainly comprising wave-number 1 and 2 components) that propagates from the troposphere into the stratosphere and interacts strongly with the preexisting circulation there. The growth of the wave may be due to an independently generated forcing mechanism in the troposphere or some process (perhaps a self-tuned resonance) that disturbs the stratosphere and troposphere together. The relevant tropospheric disturbance usually manifests itself as a blocking event.

In either case, the stratosphere probably has to be in some suitable state if a warming is to follow. Thus, the zonal-mean polar night jet may need to be further poleward than its climatological position or, in synoptic terms, the basic winter polar cyclonic vortex may need to be tighter than its usual form and perhaps displaced off the pole. This "preconditioned" state can itself be time-varying, and may be the result of an earlier wave event that disrupts the climatological flow: such a disruption may be explicable in terms of wave mean-flow interaction theory or, equivalently, in terms of planetary-wave breaking mechanisms.

The major warming itself may likewise be diagnosed in terms of wave mean-flow concepts or in terms of synoptic maps. The former suggest and

quantify a dynamical scenario in which the planetary waves are diverted from their climatological equatorward propagation by the anomalous refractive properties of the preconditioned flow and are focused into the high polar cap. Once there, they lead to mean-flow deceleration and temperature rises in the stratosphere, and the associated mean meridional circulations. The chain of cause and effect provided by this picture is borne out to some extent by experiments with quasi-linear models. It is, however, complicated by the presence of wave–wave interactions, which seem to be significant in some observational and modeling studies. Moreover, it is difficult to justify for large-amplitude disturbances, when the separation into zonal-mean and wave components may be especially artificial.

Synoptic maps, on the other hand, do not make this separation, and isentropic potential vorticity maps and maps of chemical tracers such as ozone give particularly vivid descriptions of the development of observed and modeled warmings. For example, they show how the main cyclonic vortex is distorted and displaced by the growth of the Aleutian High, and how it may split and move off the pole at the height of the warming. They also give strong qualitative hints of some of the dynamical processes involved. As yet, however, few methods have been devised for the quantitative dynamical analysis of such maps. In the long run, a judicious combination of these approaches and others (such as the consideration of the propagation of localized waves on strong cross-polar flows, for example) will probably be needed if the maximum insight is to be gained from observations and models of sudden warmings. Account will also need to be taken of aspects of warmings (such as the sharp "front-like" vertical temperature gradients that have long been observed in rocket data) that can be obscured by wave mean diagnostics and IPV maps at only a few levels.

There is much interannual variability between one stratospheric winter and another: major warmings only occur in about one in two Northern-Hemisphere winters, and when they do occur they take a variety of forms. Theoretical suggestions of possible explanations for some of this variability are now beginning to emerge and will require careful testing. The total absence of major warmings in the Southern-Hemisphere winter may be partly due to the fact that the winter vortex is stronger there than in the Northern Hemisphere, and thus more deceleration (and hence more wave forcing) would be required to reverse it. Another, related, factor may be the comparative lack of large-amplitude quasi-stationary waves there: see Sections 5.2.2 and 7.4.

The stratospheric sudden warming is perhaps the most dramatic large-scale dynamical event to occur in the middle atmosphere, and the elucidation of the processes involved remains a major challenge to meteorologists. Much further work will need to be done before a full understanding of the

phenomenon is attained. This will call for further interaction between observational studies using data from satellites and other sources on the one hand, and experiments with a hierarchy of theoretical models on the other.

References

6.1. Review articles on sudden warmings include those by Quiroz *et al.* (1975), Schoeberl (1978), McInturff (1978), Holton (1980), Labitzke (1981b), and McIntyre (1982).

6.2 Synoptic descriptions of the evolution of several individual warmings are given in some of the articles cited for Section 6.1 and other papers referenced therein. Labitzke (1981b) summarizes the behavior of the wave-number 1 and 2 height field components and zonal-mean temperature field at 30 mb for all winters from 1964 to 1980. Eliassen–Palm diagnostics and refractive indices have been applied to sudden warmings by Palmer (1981a,b), O'Neill and Youngblut (1982), and Kanzawa (1982, 1984), among others. The use of potential vorticity for interpreting sudden warmings was suggested by Davies (1981).

Low-latitude stratospheric cooling at the time of a polar warming has been noted, for example, by Fritz and Soules (1972), and upper-mesospheric cooling by Hirota and Barnett (1977).

A recent statistical study of the observed association between sudden warmings and tropospheric blocking was performed by Quiroz (1986).

A possible correlation of sudden warmings with the equatorial quasi-biennial oscillation was noted by Labitzke (1982).

A major study of the Southern Hemispheric warming of July 1974 was performed by Al-Ajmi *et al.* (1985). Hartmann *et al.* (1984) and Shiotani and Hirota (1985) give detailed descriptions of wave mean-flow interaction in the southern winter, including some minor warming events.

6.3. References to numerous other developments of Matsuno's model of the sudden warming can be found in the reviews cited for Section 6.1.

The slowing-down of a traveling wave prior to a stratospheric warming was noted, for example, by Quiroz (1975) and Madden and Labitzke (1981). A recent theoretical study of the role of baroclinic instability of distorted stratospheric flows was carried out by Frederiksen (1982).