Local Dynamics of Baroclinic Waves in the Martian Atmosphere

MICHAEL J. KAVULICH JR., ISTVAN SZUNYOGH, AND GYORGYI GYARMATI

Texas A&M University, College Station, Texas

R. JOHN WILSON

Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 25 September 2012, in final form 6 June 2013)

ABSTRACT

The paper investigates the processes that drive the spatiotemporal evolution of baroclinic transient waves in the Martian atmosphere by a simulation experiment with the Geophysical Fluid Dynamics Laboratory (GFDL) Mars general circulation model (GCM). The main diagnostic tool of the study is the (local) eddy kinetic energy equation. Results are shown for a prewinter season of the Northern Hemisphere, in which a deep baroclinic wave of zonal wavenumber 2 circles the planet at an eastward phase speed of about 70° Sol\(^{-1}\) (Sol is a Martian day). The regular structure of the wave gives the impression that the classical models of baroclinic instability, which describe the underlying process by a temporally unstable global wave (e.g., Eady model and Charney model), may have a direct relevance for the description of the Martian baroclinic waves. The results of the diagnostic calculations show, however, that while the Martian waves remain zonally global features at all times, there are large spatiotemporal changes in their amplitude. The most intense episodes of baroclinic energy conversion, which take place in the two great plain regions (Acidalia Planitia and Utopia Planitia), are strongly localized in both space and time. In addition, similar to the situation for terrestrial baroclinic waves, geopotential flux convergence plays an important role in the dynamics of the downstream-propagating unstable waves.

1. Introduction

Baroclinic instability converts available potential energy of the atmosphere into kinetic energy of synoptic-scale transient waves. While baroclinic instability has been a well-established central concept of midlatitude atmospheric dynamics since it was introduced more than 60 years ago by Charney (1947) and Eady (1949), the local processes that govern the spatiotemporal evolution of the baroclinic transient waves have remained the subject of intense research (e.g., Simmons and Hoskins 1978; Farrell 1982; Swanson and Pierrehumbert 1994; Wallace et al. 1988; Chang 2005).

Terrestrial baroclinic waves are generated in spatio-temporally localized episodes of baroclinic energy conversion. Part of the kinetic energy generated in the process propagates downstream in the form of an eastward-expanding packet of upper-tropospheric Rossby waves. The carrier wavenumber of the wave packet is between 5 and 8 depending on the latitude (the lower values typically occur at higher latitudes). The wave packet, propagating at a group velocity faster than the phase speed of the baroclinic waves, can trigger new downstream events of baroclinic energy conversion in regions where the lower-tropospheric conditions are favorable for baroclinic instability. This process maintains and connects the storm tracks of the extratropics.

Earth is not the only planet in our solar system whose atmosphere can maintain large-scale transient waves generated by baroclinic energy conversion: observations and the results of numerical experiments show that baroclinic transient waves also exist in the Martian atmosphere (e.g., Read and Lewis 2004). The Martian waves provide a unique opportunity to test our understanding of the dynamics of baroclinic waves, as some of the physical parameters that play an important role in the life cycle of the waves are similar for Mars and Earth, while others are very different.

GCM simulations and observational evidence suggest that baroclinic transient waves develop in both hemispheres...
of Mars, especially in the cold season (e.g., Barnes et al. 1993; Hinson and Wilson 2002; Banfield et al. 2004; Wilson et al. 2006; Kuroda et al. 2007; Greybush et al. 2013). In this paper, we focus on the more intense and much more often studied waves of the Northern Hemisphere. These waves are different from their terrestrial cousins in a number of important ways. In particular, their zonal wavenumber is lower, trifurcating between the values of 1, 2, and 3; their vertical structure remains deep, as their kinetic energy does not propagate at the jet level faster than their phase speed; and their amplitude never becomes localized in the zonal direction, such preserving the global character of the waves while they circle around the planet from late autumn to early spring (e.g., Read and Lewis 2004). Although the amplitude of the waves is not localized zonally, it changes in both space and time, indicating that the intensity of the processes that generate and destroy the kinetic energy of the waves also changes in both space and time. This property of the waves motivates us to study the local dynamics of the baroclinic transient waves with the help of the eddy kinetic energy equation of Orlanski and Katzfey (1991), a diagnostic tool that, to the best of our knowledge, has not been applied to Martian waves before.

We expect to find that the amplitude of the waves is modulated by the spatiotemporal variability of the large-scale forcing. For instance, the amplitude of the waves tends to grow in particular geographical regions, which indicates the existence of preferred regions for the development of baroclinic instability on Mars (e.g., Hollingsworth et al. 1996). The existence of similar preferred regions on Earth is usually explained by the contrast between the thermal properties of the oceans and the continents, and the effects of orography (e.g., Hoskins and Valdes 1990; Brayshow et al. 2009, 2011; Chang 2009). In the absence of oceans on Mars, the zonal variability in the strength of the baroclinic instability is usually attributed to the effect of orography (Hollingsworth et al. 1996), although zonal variability of the surface thermal inertia and albedo may also contribute to a lesser extent.

Spatial variability of the large-scale forcing is not the only source of the spatiotemporal variability of the properties of the transient waves, as it has been shown (e.g., Newman et al. 2004; Rogberg et al. 2010) that the dynamics of the Martian atmosphere in the Northern Hemisphere in the cold season is chaotic; that is, internal variability due to nonlinear interactions also plays an important role in the spatiotemporal evolution of the transient waves. An analysis based on the eddy kinetic energy equation allows for an investigation of the combined effects of the variability in the large-scale forcing and the effects of nonlinearities in the dynamics.

One particularly challenging aspect of studying the local dynamics of baroclinic waves for Mars is the limited amount of information available from observations about its atmosphere. The only in situ atmospheric observations available for Mars are time series of surface observations of pressure, wind, and temperature from the two Viking landers at fixed locations for a few Martian years (e.g., Barnes 1980, 1981). In addition, temperature profiles retrieved (e.g., Conrath et al. 2000) from the Mars Global Surveyor (MGS) Thermal Emission Spectrometer (TES) (Christensen et al. 2001) provide global coverage of retrievals for almost three Martian years (e.g., Conrath et al. 2000). Information about surface temperature, dust and water ice aerosol optical depth, and water vapor column abundance has also been retrieved from TES radiances (e.g., Smith et al. 2004). Similar information can be retrieved from the Mars Reconnaissance Orbiter’s Mars Climate Sounder (MCS) (Kleinböhler et al. 2009) observations. Finally, a limited number of temperature and surface pressure vertical profiles have been retrieved from MGS radio occultation experiments (Hinson and Wang 2010). None of these types of observations can be used to retrieve information about the vertical profiles of wind.

Recent studies of baroclinic transient waves in the terrestrial atmosphere have been carried out by examining analyses from the different reanalysis projects (e.g., Kalnay et al. 1996; Uppala et al. 2005). While there are multiple ongoing reanalysis projects for Mars that use modern data assimilation techniques to assimilate TES and MCS observations (e.g., Montabone et al. 2006, 2011; Lewis et al. 2007; Greybush 2011; Greybush et al. 2012; Lee et al. 2011), and data from one of them, the Mars Analysis Correction Data Assimilation (MACDA), has already been released for public use (Montabone et al. 2011a); however, we do not use Martian reanalysis data in our diagnostic calculations. Instead, we apply the local diagnostic techniques to output from the Geophysical Fluid Dynamics Laboratory Mars GCM (GFDL MGCM) after an examination of the similarities and the differences between the behavior of the baroclinic waves in the model simulations and in the observations and reanalyses.

It is not obvious that studying the transient waves in one of the reanalyses would provide more reliable results than our approach. First, most Martian reanalysis systems assimilate temperature profiles retrieved from the TES and MCS observations (e.g., Montabone et al. 2006, 2011a; Lewis et al. 2007; Greybush 2011; Greybush et al. 2012) and experience accumulated with terrestrial observations suggests that retrieved temperature profiles are usually heavily influenced by the a priori assumptions made about the profiles in the retrieval algorithms (e.g., Rodgers 2000). Second, even if the temperature retrievals provided a highly accurate representation of...
the temperature profiles, the analyses of the wind profiles, which were obtained by a correction of the model-predicted wind profiles without any observational constraints from in situ wind observations, would be heavily influenced by the model dynamics. Assimilating the TES radiance observations directly, as done by Lee et al. (2011) for Mars, eliminates the effects of the retrieval algorithm, but it makes the temperature analyses even more dependent on the dynamics of the model. In addition, correcting the observation bias, which has been found to be essential for the gainful assimilation of radiances in terrestrial applications (e.g., Eyre 1992; Derber and Wu 1998), cannot be performed for Mars because of in situ observations of the temperature, which are necessary to “anchor” the bias estimates.

Finally, we note that the motivation to study the local dynamics of baroclinic waves in the Martian atmosphere is more than scientific curiosity or the desire to improve our understanding of baroclinic instability with the aim of advancing terrestrial forecast applications. First, we hope that a better understanding of the spatiotemporal evolution of the transient waves can help design efficient observing strategies, which can provide the necessary observed information for practical numerical model-based techniques to predict Martian dust storms. These storms are the main hazards to surface operations on Mars. They have been responsible for at least one probe failure, the Russian Mars 3 Lander (Lorenz 2008), and can seriously interfere with science missions; for example, they necessitated the shutdown of both Mars exploration rovers in 2007 (Seibert et al. 2009). The storms also present a significant hazard to a potential manned surface mission in the future (Sharma et al. 2009). Therefore, the ability to forecast Martian weather is essential for the success of missions to Mars.

The outline of the paper is as follows: section 2 provides background information on the Martian atmosphere and a brief description of the model we use in our simulations; section 3 describes the temporal and zonal mean properties of the atmospheric flow and the transient waves in our simulation, pointing out the most important similarities and differences between the simulated and the observed (retrieved) temperature fields; section 4 describes the time-mean behavior of the transient waves, while section 5 investigates their spatiotemporal evolution; and section 6 offers a summary and a discussion of our most important findings.

2. Background

a. Martian atmosphere

Some of the most important physical constants for Earth and Mars are listed in Table 1. The planetary–atmospheric system of Mars can arguably be described as the closest known such system to that of Earth (Leovy 2001; Haberle 2002). In particular, Mars’s length of day (24 h 37 min 22.663 s in terrestrial time), and therefore the Coriolis parameter, are similar to those for Earth. (A Martian solar day is referred to as a Sol.) Also similar are the obliquity of Mars’s rotational axis, the typical scale height, and the Rossby deformation radius in the atmosphere. There are also some important differences between the atmospheres of the two planets; in particular, the radius of Mars is approximately half that of Earth, the average atmospheric pressure of the mean geoid is less than 1% of average terrestrial sea level values, the main constituent of the Martian atmosphere is carbon dioxide, the atmospheric gas constant is significantly lower for the Martian atmosphere. There is some difference between the atmospheres of the two planets; in particular, the radius of Mars is approximately half that of Earth, the average atmospheric pressure of the mean geoid is less than 1% of average terrestrial sea level values, the main constituent of the Martian atmosphere is carbon dioxide, the atmospheric gas constant is significantly lower for the Martian atmosphere, Mars’s solar constant is less than 50% of Earth’s, and this solar constant varies more strongly than Earth’s because of Mars’s much higher orbital eccentricity. The Martian atmosphere features only trace amounts of water vapor, and its surface has no permanent liquid water. Its terrain exhibits a striking dichotomy of highlands in the Southern Hemisphere and lowlands in the Northern Hemisphere, punctuated by the large, volcanic Tharsis Plateau, which straddles the equator (Fig. 1).

b. Martian calendar

Terrestrial atmospheric studies use the date to refer to the time of the year. In Mars research, the date is replaced by the aero-centric longitude of the Sun, denoted by $L_s$; $L_s = 0^\circ$ is the Northern Hemisphere spring equinox; $L_s = 90^\circ$ is the Northern Hemisphere summer solstice; $L_s = 180^\circ$ is the Northern Hemisphere fall equinox; $L_s = 270^\circ$ is the Northern Hemisphere winter solstice; aphelion is at $L_s = 71^\circ$; and perihelion is at $L_s = 251^\circ$.

The length of a Martian year is 668.6 Sols. Because of the eccentricity of the orbit of Mars, it takes a different

<table>
<thead>
<tr>
<th>Property</th>
<th>Earth</th>
<th>Mars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of solar day</td>
<td>24 h 00 min 00 s</td>
<td>24 h 39 min 35 s</td>
</tr>
<tr>
<td>Obliquity (°)</td>
<td>23.5</td>
<td>25</td>
</tr>
<tr>
<td>Orbital eccentricity</td>
<td>0.0167</td>
<td>0.0934</td>
</tr>
<tr>
<td>Scale height (km)</td>
<td>7.6</td>
<td>10.2</td>
</tr>
<tr>
<td>Rossby deformation radius (km)</td>
<td>1150</td>
<td>920</td>
</tr>
<tr>
<td>Mean temperature of lowest scale height (K)</td>
<td>260</td>
<td>200</td>
</tr>
<tr>
<td>Mean surface pressure (Pa)</td>
<td>$1.013 \times 10^5$</td>
<td>$-600$</td>
</tr>
<tr>
<td>Atmospheric gas constant (J kg$^{-1}$ K$^{-1}$)</td>
<td>287.10</td>
<td>188.92</td>
</tr>
<tr>
<td>Average solar constant (W m$^{-2}$)</td>
<td>1366</td>
<td>590</td>
</tr>
<tr>
<td>Typical Brunt–Väisälä frequency</td>
<td>$-6 \times 10^{-4}$</td>
<td>$-1 \times 10^{-2}$</td>
</tr>
</tbody>
</table>
number of Sols for the planet to complete the same aero-centric longitude segment at the different parts of the orbit. For instance, completing the 30° segment starting at $L_s = 240°$ takes 46.1 Sols, while completing the 30° segment starting at $L_s = 60°$ takes 66.7 Sols, making the Northern Hemisphere winter significantly shorter than the Northern Hemisphere summer. In this paper, we refer to the time of the year by $L_s$ but describe the period and the frequency of the waves in Sols.

c. The GFDL MGCM

The GFDL MGCM has been used in a large number of studies of the Martian atmosphere. These studies have included investigations of tides and planetary waves (Wilson and Hamilton 1996; Hinson and Wilson 2002; Wilson 2000; Hinson et al. 2003), the water cycle (Richardson and Wilson 2002b; Richardson et al. 2002), the dust cycle (Basu et al. 2004, 2006; Wilson 2011), the influence of topography (Richardson and Wilson 2002a), and cloud radiative effects (Hinson and Wilson 2004; Wilson et al. 2007, 2008; Wilson 2011). The reanalysis project of the University of Maryland, College Park, (Greybush 2011) also uses the GFDL MGCM. The model has also been used before to study the dynamics of transient waves (Wilson et al. 2002; Hinson and Wilson 2002; Wang et al. 2003; Basu et al. 2006).

The latest detailed description of the model can be found in Wilson (2011). Here, we provide only a brief description of the model to explain its main features and the particular choices of the parameters we made for our simulation experiments. The model is based on the finite volume (FV) numerical solver (Lin 2004) included in the GFDL Flexible Modeling System (FMS). For our simulations, we chose the particular version of the dynamical core that is based on a cube–sphere geometry (Putman and Lin 2007; Donner et al. 2011). We integrate the model at horizontal resolution C22—that is, each cube face has $22 \times 22$ volume elements, yielding a nearly uniform resolution of $4° \times 4°$. The vertical discretization uses 36 vertical levels in a hybrid sigma–pressure vertical coordinate, which transitions from a terrain-following coordinate near the surface to a pressure coordinate at the higher altitudes. The top layer of the model atmosphere is centered at about 0.004 Pa, which corresponds to about 95 km above the mean geoid.

d. Calculation of the diagnostics

Our diagnostic calculations are based on four daily outputs at regular (1/4 Sol) intervals from a representative year of a multiyear simulation experiment with the
For these calculations, which are all carried out using pressure $p$ as the vertical coordinate, the model fields are interpolated to a uniform, $4\degree \times 4\degree$ resolution latitude–longitude grid. While all diagnostics use $\omega = dp/dt$, which is the vertical coordinate of the wind vector for the pressure vertical coordinate, to describe vertical motions, the results are visualized using log pressure as the vertical coordinate of the vector of position. We use the log-pressure vertical coordinate in the figures, because this way we can show results for the entire depth of the model atmosphere. Two important properties of the log-pressure coordinate, however, should be kept in mind while analyzing the results shown in the figures. First, log pressure is essentially a pressure-based height coordinate; thus, the mass associated with a unit distance along the vertical coordinate decreases exponentially with the distance from the origin. Therefore, the contribution of the flow in the upper atmosphere to the total kinetic energy of the atmospheric column is much smaller than it may seem based on values of the wind speed (kinetic energy per unit mass). Second, the vertical coordinate of the wind vector in log-pressure coordinate is $-\omega p/g$, where $\rho$ is the density and $g$ is the standard gravity. Hence, the same value of $\omega$ indicates a shorter distance traveled by the “air” parcel in the vertical direction per unit time at a higher altitude.

3. The atmospheric flow

In what follows, we start with a brief description of the seasonal cycle on Mars; then, we describe the zonal-mean structure of the flow for the period investigated in this paper; finally, we describe the statistical properties of the transient waves. Whenever there exists observation-based information for comparison, we point out the similarities and the differences between the simulated and observed properties.

a. Seasonal cycle

The main source of diabatic heating in the Martian atmosphere is radiative forcing. Because of the short time scales of the radiative processes and the absence of oceans, the surface temperature closely follows the seasonal changes in the incoming solar radiation. In the cold season, low temperatures lead to a deposition of gaseous CO$_2$ into the polar ice cap by condensation, while in the warm season, CO$_2$ is released back into the atmosphere by sublimation. Up to 30% of the atmospheric mass can be trapped in the polar ice caps (e.g., James et al. 1992; Read and Lewis 2004). The position of the southern edge of the northern polar ice cap, which disappears in the warm season and expands to 50°N in the cold season, has an important influence on baroclinic instability by being the primary factor in determining
the near-surface vertical temperature profile and the position of the polar front. The time evolution of the position of the edge of the polar ice caps throughout the Martian year in our simulation is shown in Fig. 2.

b. Zonal-mean temperature field

We show the vertical cross section of the zonal-mean fields for both $L_s = 200^\circ$ and $L_s = 230^\circ$ to illustrate the seasonal changes in the zonal-mean conditions (Fig. 3). The zonal mean of the simulated temperature field (Fig. 3, top) features a strong meridional temperature gradient (polar front) at the southern edge of the northern polar ice cap. While the strong meridional temperature gradient provides favorable conditions for baroclinic instability at the earlier time, $L_s = 200^\circ$, the polar front becomes more strongly tilted toward the pole with height at the later times (e.g., at $L_s = 230^\circ$), leading to an inverted temperature profile in the lower atmosphere at high latitudes. This temperature profile is unfavorable for baroclinic instability near the surface as an air parcel traveling poleward cannot rise because of the high static stability of the atmosphere, even in the presence of a strong meridional temperature gradient. These properties are in good qualitative agreement with those shown by the temperature fields retrieved from TES observations (e.g., Fig. 3.18 of Read and Lewis 2004).

c. Zonal-mean circulation

The most important features in the zonal mean of the zonal component of the wind are the strong westerly jet, with a vertically extended and poleward-tipped core centered at about 5-Pa pressure level in the Northern Hemisphere midlatitudes, and the easterly jet centered at the equator with a core at about 0.5 Pa. The zonal mean of the vertical velocity indicates a cross-equatorial Hadley cell, which is more intense at $L_s = 230^\circ$ than at $L_s = 200^\circ$ (Fig. 3, bottom). The features of the wind field shown here are consistent with those from other Mars GCM simulations.

Figure 4 shows the time mean of the zonal and the vertical components of the wind at the 400- and 100-Pa pressure levels for the interval $L_s = 200^\circ$–230$^\circ$. The orography has a strong effect on the meridional component of the wind: the zonal wavenumber–2 orographic forcing leads to a wavenumber-2 pattern in the meridional wind field, which extends into the layers well above the top of the mountains. The effect of the orography on the vertical component of the wind is less obvious, except for the immediate vicinity of the mountains in the lower atmosphere; in particular, the vertical motion is downward almost everywhere at the 100-Pa level. The latter feature of the flow is due to the dominance of the sinking motion in the northern branch of the Hadley cell circulation.

It is important to note that the stationary wave response of the atmosphere to orography on Mars is different from that on Earth, primarily because of the differences in the nature of the diabatic heating (e.g., Nayvelt et al. 1997). The existence of the stationary zonal wavenumber–2 pattern, which is present in our simulations, is supported by TES retrievals (Banfield et al. 2003). [Banfield et al. (2003) also identified a zonal wavenumber–1 stationary wave, whose phase shifts 90° in longitude between the beginning and the middle of the cold season and then back to its original phase by the end of the cold season.]

d. Space–time spectrum of the surface pressure

We start our analysis of the transient waves with a space–time spectral analysis of the time series of surface pressure at latitude circle 60°N. The temporal component of the analysis is based on 32-Sol running time windows. The seasonal trend is removed from the time series by discarding the frequency-zero component of the temporal spectrum, which is equal to the 32-Sol running mean.

The results of the spectral analysis (Fig. 5) show that initially, for the segment $L_s = 200^\circ$–230$^\circ$, the wavenumber-2 component dominates the spatial variability. The dominant frequency during this period is 2.6 Sols. Beyond about $L_s = 230^\circ$, the wavenumber-1 component and the related 7-Sol temporal variability becomes dominant. (The overall dominance of the 7-Sol disturbance for the entire winter is at odds with the second Viking year [Martian year (MY) 13] observations but not inconsistent with the first Viking year (MY12) observations (Barnes 1980, 1981).) While the amplitude of the wavenumber-3 component tends to be significantly smaller than the amplitude of the wavenumber-1 and -2 components, it still plays a prominent role in the spatial variability for the segment $L_s = 200^\circ$–230$^\circ$. The dominant frequency associated with the wavenumber-3 component is about the same, 2.6 Sols, as the dominant frequency for the wavenumber-2 component.

e. Eddy component of the state variables

To isolate the signal associated with the transient synoptic waves, we apply a multiband digital filter to the time series of surface pressure, removing the variability related to diurnal and semidiurnal tidal waves. To be precise, we remove the signal in the frequency band 0.95–1.05 Sol$^{-1}$ and at frequencies equal to or higher than 1.82 Sol$^{-1}$ with the help of a Hamming window digital filter. [The description of the filter can be found in IEEE (1979).] In addition, we subtract the 30-Sol
running mean to remove the seasonal trend. We will refer to the filtered state variables as the eddy components of the variables and distinguish them from the original variables in notation by a superscribed prime. For instance, we denote the eddy component of the temperature, zonal component of the wind vector, and the vertical component of the wind vector in pressure coordinates by $T'$, $u'$, and $\omega'$, respectively.

**FIG. 3.** Vertical cross section of the zonal mean of some state variables for (left) $L_s = 200^\circ$ and (right) $L_s = 230^\circ$. The state variables shown are (top) the temperature with a contour interval of 10 K (shading indicates values less than 180 K), (middle) the zonal component of the wind vector with a contour interval of 20 m s$^{-1}$ (shading indicates negative values), and (bottom) the vertical velocity in pressure coordinates with a contour interval of $0.2 \times 10^{-3}$ Pa s$^{-1}$ (shading indicates negative values).
f. Transient waves in the temperature field

We start the description of the behavior of the eddy component of the temperature by a spectral analysis of the zonal variability of the eddy component of the temperature $T'$ at a lower level (570 Pa) of the atmosphere (Fig. 6). The latitude belt where the wave amplitude is the largest follows the southward-expanding, and later retreating, edge of the polar ice cap, with a distinct minimum in midwinter. For comparison, we show the same field for MACDA (Montabone et al. 2011a) in Fig. 7. This figure shows results for the same time period as the one from which we use the dust retrievals to constrain the zonal average of the aerosol fields in our simulations. Because the reanalysis is constrained by temperature retrievals, we expect it to provide a more realistic depiction of the temperature fields in the lower atmosphere than our model simulation. Since the dynamics are chaotic, which leads to significant internal variability of the system, we can only expect a broad qualitative agreement between the simulation and the reanalysis. The qualitative differences between the two figures, however, are larger than what could be reasonably explained by internal variability. Most importantly, the amplitude of the wavenumber-3 component is too small, while the amplitude of the wavenumber-2 component is too large in the model simulation. We also note that while MACDA shows significant interannual variability of the characteristics of the waves (results not shown), the properties of the GFDL MGCM illustrated by Fig. 6 show relatively little interannual variability.

Next, we show the vertical cross section of the zonal-mean magnitude of the eddy component of the temperature for $L_s = 200^\circ$–$230^\circ$ (Fig. 8) and $L_s = 230^\circ$–$260^\circ$ (Fig. 9). These figures also show the contribution of the different zonal-wavenumber components to the amplitude of the transient waves in the temperature field. The behavior of these components can be compared to that described in detail for the TES retrievals by Banfield et al. (2004) and Hinson and Wang (2010).
factor that should be taken into account when comparing the properties of the simulated and observed temperature fields is the limited vertical resolution and limited vertical coverage of the TES retrievals. For instance, the vertical resolution of the TES retrievals at the 50-Pa level is only about two scale heights. Because of these limitations of the retrievals, they cannot be used for the assessment of the quality of the simulations in the upper atmosphere: our comments about the similarities and the differences between the simulated and observed properties of the waves in the following paragraph refer to the behavior of the waves below 50 Pa.

In good agreement with its observed behavior, the wavenumber-1 component dominates the variability of the temperature field at high altitudes, especially at the later times (for \( L_s = 230^\circ - 260^\circ \)). The behavior of the wavenumber-2 component in our simulation also closely resembles its observed properties: it dominates the variability of the temperature field in the lower atmosphere (below 100 Pa) and is present at all pressure levels.
levels for $L_s = 200^\circ-230^\circ$, while its contribution, especially near the surface, weakens at the later times ($L_s = 230^\circ-260^\circ$). Similar to its observed behavior, the wavenumber-3 component plays an important role only near the surface and only for the earlier period ($L_s = 200^\circ-230^\circ$). The amplitude of this wave component, however, is clearly smaller in the simulation than in the retrievals. This result corroborates the earlier conclusion drawn based on the comparison to the reanalysis data, that the wavenumber-3 component plays a less prominent role in the simulation than it should. While this deficiency of the simulated waves is an obvious and important limitation of our study, we do not believe that it has a major effect on the particular conclusions we will draw about the qualitative dynamics of the waves. We also note that the underrepresentation of the wavenumber-3 component is a relatively new problem in the GFDL MGCM, as past studies with the model showed a prominent role of the wavenumber-3 component (Hinson and Wilson 2002; Wang et al. 2003; Basu et al. 2006; Wilson et al. 2006).

g. Transient waves in the wind field

The propagation of the transient waves along the westerly jet is illustrated by Fig. 10, which shows a
Hovmöller diagram for the eddy component of the meridional component of the wind vector $\nu'$ at the 100-Pa level for $L_s = 200^\circ$–$230^\circ$. (In this figure, the dependence of $\nu'$ on the latitude is removed by taking its average over the latitude band $60^\circ$–$85^\circ$N.) The Hovmöller diagram shows a wavenumber-2 wave with a spatiotemporally varying amplitude that travels around the planet in about 5 Sols. (This time is consistent with the wavenumber-2 traveling wave of period 2.6 Sols that we have detected.) While the wave never becomes fully localized in the zonal direction, its amplitude typically increases in the main regions of baroclinic energy conversion and decreases elsewhere. Next, we will study the structure of the wave in detail for the 4-Sol period starting at $L_s = 212^\circ$, the period in which the amplitude of the wave reaches its maximum.

4. Time-mean properties of the transient waves

In this section, we investigate the time-mean properties of the transient waves. We focus on investigating the waves for $L_s = 200^\circ$–$230^\circ$, the time interval in which the waves are the most active in the lower atmosphere. An additional motivation to focus on this time interval is that “flushing” dust storms, which have been found to be closely related to baroclinic waves, have the peak of their activity in $L_s = 200^\circ$–$230^\circ$, while they become inactive approaching midwinter (Wang et al. 2003, 2005; Hinson and Wang 2010). We start our investigation with an analysis of the time mean of the terms of the eddy kinetic energy equation. Since the eddy kinetic energy equation does not distinguish between the contributions of the different processes in the different atmospheric
layers, we also present diagnostics to investigate the vertical structure of the processes described by the terms of the eddy kinetic energy. In what follows, as well as in the figure captions, an overbar indicates a time mean for $L_s = 200^\circ - 230^\circ$.

**a. The eddy kinetic energy equation**

The Eulerian variable $K_e$ is the eddy kinetic energy per unit mass. The eddy kinetic energy equation of Orlanski and Katzfey (1991), given as

\[
\frac{\partial}{\partial t} (K_e) = -\langle \mathbf{v} \cdot \nabla K_e \rangle - \langle \mathbf{v} \cdot \mathbf{v}' \mathbf{\phi}' \rangle - \langle \mathbf{v}' \cdot (\mathbf{v}_3' \cdot \mathbf{V}_3) \mathbf{v}_m \rangle - \mathbf{v} \cdot \mathbf{v}' - \langle (\text{Residue}) \rangle,
\]

describes the evolution of the vertical mean of the kinetic energy, which is defined by the mass-weighted vertical mean of $K_e$, in the two horizontal directions and time for an atmosphere in hydrostatic balance. In Eq. (2), $\mathbf{v}' = (u', \mathbf{v}')$, $\mathbf{v}_3 = (u', \mathbf{v}', \omega')$, while $\mathbf{v}_m$ is the 30-day-running-mean (basic flow) component of $\mathbf{v}$—that is, $\mathbf{v} = \mathbf{v}_m + \mathbf{v}'$; $\mathbf{V} = (\partial / \partial x, \partial / \partial y)$ is the horizontal del operator, while $\mathbf{V}_3 = (\partial / \partial x, \partial / \partial y, \partial / \partial p)$ is the three-dimensional del operator; the overbar indicates time averaging; the angle brackets indicate a mass-weighted mean.

\[
K_e = \frac{1}{2} (u^2 + v^2)
\]
where \( f \) is an arbitrary scalar field; \( \rho_t (=0) \) is the pressure at the top of the model atmosphere; and square brackets indicate the surface integral, normalized by \( \rho_s - \rho_t \), across the surface of the planet (subscript \( s \)) or the top of the model atmosphere (subscript \( t \)).

Each term of the right-hand side of Eq. (2) represents a different dynamical process that can change the eddy kinetic energy locally. Term 1 is the eddy kinetic energy transport term, while term 2 is the geopotential flux convergence. The spatial integral of these two terms over the globe is zero, which indicates that the processes they represent play a role solely in redistributing the eddy kinetic energy and cannot serve as its ultimate source or sink. More precisely, combining terms 1 and 2 into a single term, we would obtain a term to describe the (total) transport of the eddy kinetic energy due to advection and the work done by the pressure gradient force in the eddies.

We note that term 2 is often called the ageostrophic flux convergence term, because if \( \mathbf{v} \) is written as the sum of a “geostrophic” component and an “ageostrophic” component, where the former is defined by the geostrophic wind for a constant reference value of the Coriolis parameter, then only the ageostrophic component contributes to term 2. We do not follow this terminology, because we examine the terms of Eq. (2) for the entire hemisphere; hence, the geostrophic component would be drastically different from the actual geostrophic wind at most latitudes.

Although the processes represented by terms 1 and 2 cannot be ultimate sources or sinks of the eddy kinetic energy, they can play an important role in the local changes of the eddy kinetic energy. Most importantly, term 2 describes a process that can trigger baroclinic energy conversion in regions of the terrestrial atmosphere, where the conditions are favorable for the development of baroclinic instability (Orlanski and Katzfey 1991; Chang and Orlanski 1993; Orlanski and Sheldon 1995).

The rate of baroclinic energy conversion, which is also due to the work of the pressure gradient force in the eddies, is described by term 3. This term is positive when eddy available potential energy is converted into eddy kinetic energy by rising warm air or sinking cold air in the wave. Term 4 is the barotropic energy conversion term, which is positive when eddy kinetic energy is generated
and negative when eddy kinetic energy is destroyed. The first part of this term represents the transfer of kinetic energy between the mean flow and the eddy, while the second part of this term represents the transfer of kinetic energy between the eddy and the mixed kinetic energy \( v_m v' \). Since the time mean of the latter term is zero, it cannot be the ultimate source of barotropic energy conversion, but it can play a role in the instantaneous transfer of kinetic energy.

Terms 5 and 6, whose contribution is negligible in both the terrestrial and the Martian atmosphere, represent, respectively, the surface fluxes of the eddy kinetic energy and the potential energy at the bottom and the top of the model atmosphere. Term 7, called the “residue” term, represents the combined effect of all unaccounted sources and sinks of the eddy kinetic energy, which include interpolation errors, frictional effects, other subgrid processes, and the errors introduced by the temporal filtering of the state variables. This term is computed by first obtaining the left-hand side of Eq. (2) by

\[
\frac{\partial}{\partial t} \langle K_e \rangle = v' \cdot \frac{\partial v'}{\partial t}
\]

and then subtracting the sum of terms 1–4 from the result.

The time mean of the eddy kinetic energy and terms 1–4 and 7 is shown in Fig. 11. This figure shows that the eddy kinetic energy (top row) rapidly increases poleward from about latitude 50°N and reaches its maximum at about 75°N. In the zonal direction, the local maxima of the eddy kinetic energy are located in the regions of Utopia Planitia and Acidalia Planitia, while the eddy kinetic energy is significantly lower north of the Tharsis Plateau than elsewhere at the same latitudes. The main source of the eddy kinetic energy is baroclinic energy conversion (second row), with the local maxima of the rate of baroclinic energy conversion located upstream of the maxima of the eddy kinetic energy. (This relationship between the locations of the maxima of the rate of baroclinic energy conversion and the eddy kinetic energy is similar to that in a terrestrial storm track.) According to Fig. 11, two processes play important roles in shifting the location of the maxima of the eddy kinetic energy downstream from the primary regions of baroclinic energy conversion in the latitude band 65°–85°N; the downstream transport of the eddy kinetic energy by the main flow (fourth row) and barotropic energy conversion (fifth row), which acts as a sink of the eddy kinetic energy in the main regions of baroclinic energy conversion and as a source downstream of those regions. While ageostrophic geopotential flux convergence (third row) also transports eddy kinetic energy downstream, that transport is also toward the southern and the northern flanks of the high-eddy-kinetic-energy zone rather than a straight downstream transfer in the latitude channel between 65° and 85°N. Thus, ageostrophic geopotential flux convergence is an important sink of the eddy kinetic energy in the zonal channel of high eddy kinetic energy. [Orlanski and Katzfe (1991) drew a similar conclusion for Earth.] It is important to point out that divergence of the ageostrophic geopotential fluxes is not the only source of the eddy kinetic energy on the flanks of the high-eddy-kinetic-energy zone, as those are also the main regions of eddy kinetic energy generation by barotropic energy conversion.

The main sink of the eddy kinetic energy is an unknown combination of the processes contributing to the residue term (bottom row), which has its maxima (in absolute value) just downstream of the regions of most intense baroclinic energy conversion. This term tends to be negative for the terrestrial atmosphere as well, but with a smaller relative magnitude compared to the other terms. The negative values for Earth are usually explained by frictional effects. While frictional effects certainly play a role for Mars, the larger relative magnitude of the residue term indicates that additional processes are also likely to contribute. For instance, filtering the tidal waves from the eddy components may have a similar effect to that of friction. In particular, diurnal forcing has been found to have an important effect on the frequency of
changes between wave regimes characterized by a dominant period of 2–4 and 5–7 Sols (Collins et al. 1996). There are two possible explanations for the influence of the diurnal and semidiurnal cycles on the baroclinic waves: the associated periodic perturbations can influence the evolution of the baroclinic waves either because the chaotic nature of the wave dynamics makes the dominant period of the waves sensitive to any small perturbation or because the tidal waves and the baroclinic waves can exchange energy. The tidal waves “shaking the table” argument of Read and Lewis (2004) suggests the first explanation, while the observation of Hinson and Wang (2010), that the transition among baroclinic wave modes can strongly modulate the intensity of the meridional winds near the surface, points to the second possibility. In situations where the second explanation applies, filtering the signal associated with tidal waves can have an effect on the energy conversion terms, even though a period 2.6-Sol wave mode dominates the flow in the time interval we investigate. We examined the latter possibility by reproducing Fig. 11 without filtering the diurnal and the semidiurnal signal from the eddies (results are not shown). While the eddy kinetic energy was clearly larger when filtering was not applied, the magnitude and the patterns of the energy conversion terms remained very similar to those in Fig. 11. Most importantly, the residue term remained negative and its magnitude was not reduced. This result suggests that the large negative values of the residue term are not artifacts of the temporal filtering.

b. Zonal mean of the eddy kinetic energy

The zonal mean of the eddy kinetic energy per unit mass $K_e$ is shown in Fig. 12. The values shown in this figure should be interpreted as a measure of the local amplitude of the waves rather than a characterization of the contribution of the different regions to the total eddy kinetic energy of the atmosphere, as the density of the atmosphere, which determines the weight of the contribution of the eddy kinetic energy per unit mass to the total column integral, decreases exponentially along the $y$ axis.

The absolute maximum of $K_e$ is at about 0.04 Pa between 75° and 80°N. The local maximum in the lower atmosphere (below 10 Pa) is in the same latitude band as the absolute maximum but in the westerly jet region, centered at about 100 Pa. The contribution of $K_e$ to the total column eddy kinetic energy from the region of the lower-atmospheric maximum is overwhelming, because the density of the atmosphere is orders of magnitude higher at 100 Pa than at 0.04 Pa. The position of the eddy kinetic energy maximum in the westerly jet region suggests that baroclinic energy conversion may contribute to the maintenance of the jet. [We recall that for the terrestrial atmosphere, the convergence of the momentum fluxes associated with the waves generated by baroclinic energy conversion are thought to play an important role in the maintenance of the westerly jet (e.g., Pedlosky 1987; James 1994).]

A similar spatial distribution of the eddy kinetic energy in a GCM simulation was first reported by Barnes et al. (1993) in a model with a top at 10 Pa. That paper provided evidence that baroclinic energy conversion
played an important role in the generation of the eddy kinetic energy. It also showed that the structure of the eddies in the upper part of the model atmosphere was close to equivalent barotropic. A later study (Forget et al. 1999) found that simulations in the layer above 10 Pa were highly model sensitive. In addition, a more recent study (Angelats i Coll et al. 2005) found that the layer between about 10 and 0.01 Pa was highly sensitive to the parameterization of orographic gravity wave effects. Based on these results, we expect that gravity wave dynamics and ageostrophic effects play an important role above 10 Pa. The main focus of our analysis is on the dynamics of the waves below that level.

c. Eady index

We start the detailed investigation of the baroclinic energy conversion processes by examining how conducive the basic flow is to baroclinic instability. For this purpose, we use the Eady index (Hoskins and Valdes 1990), which is defined by the ratio of the vertical shear in the zonal wind component and the Brunt–Väisälä frequency. (The latter measures the static stability in a stably stratified atmosphere.) A larger Eady index means a larger potential for intense baroclinic energy conversion. The top panel of Fig. 13 shows the vertical cross section of the Eady index in the basic flow, while the bottom panel shows the meridional mean of the index in the latitude band 60°–80°N. For the meridional averaging, we chose this latitude band, because Fig. 11 (second row) indicates that this is the zone of the most intense baroclinic energy conversion.

The primary maximum of the index in the zonal mean (Fig. 13, top) is in a narrow layer at the surface, spreading from 60° to 80°N, while the secondary maximum is centered at 60°N in the layer between 700 and 100 Pa. As it will turn out, the much broader region of the secondary maximum plays the key role in the baroclinic energy conversion. In this layer, there are two regions of enhanced baroclinic instability in the zonal direction (Fig. 13, bottom): one extends from about 30° to 130°, while the other extends from about 200° to 340°. These regions are shifted upstream compared to the regions of main baroclinic energy conversion shown in Fig. 11. This relationship between the zonal positions of the region of high Eady indices and regions of most intense baroclinic conversion are similar to that on Earth: because it requires time for the developing unstable wave to attain a structure that can maximize its efficiency in extracting energy from the basic flow, the most intense energy conversion always takes place downstream of the location of the strongest instability in the basic flow (e.g., James 1994).

The bottom panel of Fig. 13 also shows that from about 10 Pa, the Eady index gradually decreases with
height to zero at about 1 Pa. Above that level, the Eady index is typically negative. These results indicate that the layer most favorable for the development of baroclinic instability is between the surface and the 10-Pa level, while the high static stability of the vertical stratification of the upper atmosphere makes the development of baroclinic instabilities impossible there.

d. Baroclinic energy conversion

An unstable baroclinic wave is characterized by a simultaneous meridional and vertical transport of heat. In particular, the integrand in the baroclinic energy conversion term of Eq. (2) \( \omega' \alpha' \) is directly proportional to the vertical component of the instantaneous heat flux \( T' \omega' \) because \( \alpha' \sim T'/p \). A term related to the meridional component of the instantaneous heat flux \( T' \nu' \) does not appear in the eddy kinetic energy equation because the role of the meridional heat flux is to convert available potential energy of the ambient flow into the eddy available potential energy of the waves. (The role of the upward temperature fluxes is to convert the eddy available potential energy into the eddy kinetic energy.)

Figure 14 shows the vertical cross section of the zonal mean of \( T' \omega' \) and \( T' \nu' \). This figure indicates a region of strong upward and poleward heat fluxes in the westerly jet region. The temperature fluxes are the strongest in a narrow layer near the surface in a latitude band where the zonal-mean orography is the lowest and the Eady index is the highest. In the layer between 500 and 100 Pa, however, the strongest temperature fluxes are shifted by about 10° to the north compared to the position of the secondary maximum in the Eady index.

While the vertical heat flux is the most widely used indicator of baroclinic energy conversion, it can provide a somewhat misleading picture about the vertical structure of the intensity of the energy conversion in a deep atmospheric layer: since \( \alpha' \sim T'/p \), the magnitude of the temporal mean of the energy conversion is larger than indicated by \( T' \omega' \) at higher altitudes, where \( p \) is smaller. To avoid this potential problem, we also show the zonal–vertical cross section of \( -\alpha' \omega' \) (Fig. 15, top). As expected, this figure, at high altitudes, indicates more intense energy conversion than the one that shows the vertical heat flux. A peculiar feature of the figure, however, is the pattern of negative baroclinic energy conversion south of the pattern of positive energy conversion above 40 Pa. Such a pattern is rarely observed in an atmosphere, or in a computer simulation of an atmosphere, because it indicates locations where lighter “air” is sinking and heavier air is rising. (This feature of the flow will be further discussed shortly.)
We have found no evidence for positive baroclinic energy conversion above the 10-Pa level, which suggests that the meridional heat fluxes in the upper atmosphere are not related to baroclinic instability. It is important to point out, however, that the upper-atmospheric heat fluxes shown in Fig. 14 also transport warm air into a colder environment (and cold air into a warmer environment) at all latitudes: where the fluxes change sign at around 65°N, the meridional temperature gradient in the ambient flow also changes sign. These meridional temperature fluxes do not lead to baroclinic energy conversion, because the static stability is high in that region due to the weak vertical temperature gradient (see top row of Fig. 3 and bottom panel of Fig. 13). Since we can rule out baroclinic instability as a major source of the eddy kinetic energy above 10 Pa, other processes must provide the eddy kinetic energy for the development and the maintenance of the waves above 10 Pa.

Figure 16 shows the vertical cross section of the meridional mean of $-\bar{\omega}'$ (top panel) and $\bar{T}'\bar{\omega}'$ (bottom panel). This figure shows, in agreement with Fig. 11 (second row), that the preferred region of baroclinic energy conversion extends eastward from about longitude 300° to 200° (from the western edge of Acidalia Planitia to the eastern edge of Utopia Planitia), with the most intense energy conversion regions located between 80° and 130° (in Utopia Planitia), and at around 330° (in Acidalia Planitia). In the region between longitudes 220° and 280° (north of the Tharsis Plateau), the intensity of baroclinic energy conversion is weak and mainly occurs in the layer between 100 and 20 Pa.

e. Geopotential flux

The bottom panel of Fig. 15 shows the vertical cross section of the zonal mean of the convergence of the three-dimensional geopotential flux, defined as

$$-\bar{\nabla} \cdot \bar{\nabla}' \phi' = -\bar{\nabla} \cdot \bar{\nabla}' \phi' - \frac{\partial \omega' \phi'}{\partial p}. \tag{5}$$

[The vertical mean of the first term on the right-hand side of Eq. (5) is the geopotential flux convergence term (term 3) of Eq. (2), while the vertical mean of the second term on the right-hand side of Eq. (5) is term 6 of Eq. (2).] A comparison of the two panels of Fig. 15 shows that in the time-mean sense, the convergence of the three-dimensional geopotential flux is a sink of the eddy kinetic energy in the main region of positive baroclinic energy conversion and a source of the eddy kinetic energy in the main region of negative baroclinic energy conversion. In the latter region, the magnitude of the
positive values of the geopotential flux convergence is larger than the magnitude of the negative values of the baroclinic energy conversion. This result indicates that the total work done by the pressure gradient force is positive in that region. In addition, it also suggests that the negative baroclinic energy conversion is the result of a strongly convergent flow, which can more than compensate the buoyancy force acting on the lighter air parcels.

f. Barotropic energy conversion

The top panel of Fig. 17 shows the vertical cross section of the zonal mean of $2v_0^3/C_1^3$ for $L_s = 200^\circ–230^\circ$. Contours show (top) the eddy kinetic energy (m$^2$s$^{-2}$, same as the field shown in Fig. 12), and (bottom) the variance of the geopotential height (gpm).

This term describes the contribution of barotropic energy conversion to changes in the time-mean eddy kinetic energy. The bottom panel of Fig. 17 shows the meridional mean of the same term for $60^\circ–80^\circ$N. The two panels together indicate that in the main regions of baroclinic energy conversion above the 100-Pa level, the barotropic energy conversion is negative, suggesting that the eddies generated by the baroclinic energy conversion transfer a part of their kinetic energy to the basic flow. This result suggests that, similar to the situation for the terrestrial atmosphere, the eddy kinetic energy generated by baroclinic energy conversion plays a role in the maintenance of the westerly polar jet.

The main region of positive barotropic energy conversion, the region of transfer of kinetic energy from the basic flow to the eddies, is on the southern side of the westerly jet, with a maximum in the layer between 100 and 20 Pa (Fig. 17, top). Thus, the barotropic energy conversion broadens the zone where eddies can exist in the southerly direction. In the upper atmosphere (bottom panel), the barotropic energy conversion acts as either a local source or a local sink of the eddy kinetic energy. The zonal mean of the barotropic energy conversion (results are not shown), however, is negative at all latitudes and pressure levels, which suggests that it is unlikely to be the ultimate source of the large amount of eddy kinetic energy per unit mass in the upper atmosphere. We note that this result is in contrast to the results of earlier studies (e.g., Barnes et al. 1993), which found that barotropic energy conversion can be a net source of the eddy kinetic energy in the upper atmosphere. A possible explanation for this discrepancy is that our study focuses on an early period of the cold season and that extending the time period of the study would lead to a different result. The investigation of this possibility will be the subject of future research.
5. Spatiotemporal evolution of the waves

We now turn our attention to the spatiotemporal evolution of the transient waves. In particular, we first study the zonal–temporal propagation of the waves, then the evolution of the horizontal and the vertical structures of the waves, and finally the local energy conversion processes.

a. Evolution of the amplitude of the waves

In what follows, we study the structure of the wave for the particular 4-Sol period that starts at \( L_s = 212^\circ \), the time interval in which the amplitude of the wave reaches its maximum. We first show the eddy component of the meridional wind at 100 Pa (Fig. 18, left) and the surface pressure (Fig. 18, right) for the selected time period. This figure shows that the phase speed of the wave is about 70 Sol\(^{-1}\) at both the jet level and the surface. While the wave has a wavenumber-2 structure for the entire 4-Sol-long period, its amplitude changes in both space and time. The changes in the amplitude of the wave become even more transparent by extracting and plotting the amplitude (packet envelope) of the wave (Fig. 19). Formally, the extraction of the wave packet is carried out by searching for the function \( A(x) \) that satisfies

\[
w(x) = A(x) \cos 2x, \tag{6}
\]

where \( w(x) \) is the function that describes the selected eddy variable along a latitude circle. We find \( A(x) \) by applying the Hilbert transform–based method of Zimin et al. (2003) to \( w(x) = v'(x) \) and \( w(x) = p'_s(x) \). The two-dimensional pictures of the packet envelope shown in Fig. 19 are obtained by extracting \( A(x) \) for each latitude where data are available. We note that the wave-packet extraction algorithm requires the selection of a wavenumber range around the carrier wavenumber, since a wave packet is composed of Fourier modes with wave-numbers that are similar to the carrier wavenumber. [A broader function \( A(x) \) in the \( x \) direction means that fewer Fourier components are required to recover \( w(x) \)]. We found that selecting the wavenumber range to be from 1 to 4 gave the best results, which is in good agreement with the finding of the observational and modeling studies that the Martian transient waves are composed of waves of that wavenumber range.

The wave packet analysis (Fig. 19) shows that, in addition to the zonal variability of the wave amplitude observed in Fig. 18, there is also a meridional variability in the wave amplitude. In particular, there is a significant variability in the position of the southern boundary of the region where the amplitude is notably different from

![Fig. 16. Vertical cross section of the meridional mean of (top) \(-\overline{\alpha^2w}\) and (bottom) \(\overline{T'v'}\) for \( L_s = 200^\circ - 230^\circ \) and \( 60^\circ - 80^\circ N \).](image-url)
zero. Since this boundary is set by the intensity of the barotropic energy conversion and the transport of eddy kinetic energy by the ageostrophic flux convergence downstream and southward of the main regions of baroclinic energy conversion, the spatiotemporal variability of the boundary indicates a significant variability in the intensity of those two processes.

b. Vertical structure of the wave

We start the investigation of the vertical structure of the wave by following its evolution in vertical–zonal cross sections (Fig. 20). As can be expected from a wave actively converting available potential energy into eddy kinetic energy, it has a vertical structure that is tilted westward with height below about 10 Pa at most times. The angle of the tilt, however, changes with location and time, suggesting that the intensity of the baroclinic energy conversion also changes with location and time. This conclusion is also supported by Fig. 21, which shows the instantaneous heat fluxes, $T'v'$ and $T'w'$, associated with the wave; while the spatial structure of the heat flux fields indicates a global wave of wavenumber 2, which is actively converting available potential energy into eddy kinetic energy, there are considerable differences in the intensity of the temperature fluxes in the active regions. There is also a significant spatiotemporal variability in the depth of the layer penetrated by the plumes of the heat fluxes.

The differences between the properties of the waves below and above about 10 Pa suggest that they are different types of waves. Most importantly, the structure of the waves above about 10 Pa is not tilted westward and while they actively transport heat in the meridional direction, they do not transport heat in the vertical direction. These properties further confirm that the upper-tropospheric waves are not baroclinic waves. In addition, the wavenumber and the phase velocity of the waves below and above about 10 Pa are often different. The dynamics of the upper- and lower-atmospheric waves, however, is not completely independent, as parts of the two wave trains becomes phase locked occasionally, leading to the development of troughs and ridges that extend from the bottom to the top of the model atmosphere (Fig. 20).

Figure 22 shows the evolution of the vertical structure of the convergence of the geopotential fluxes and the barotropic energy conversion. The fields above 10 Pa are not shown because these two fields have locally large
FIG. 18. Illustration of the eastward-traveling baroclinic wave for the 4-Sol period starting at \( L_s = 212^\circ \). Color shades show the eddy components of (left) the meridional component of the wind vector at the 100-Pa level and (right) the surface pressure. The fields are shown with a 0.5-Sol interval. Contours show the geopotential height at the (left) 100- and (right) 800-Pa pressure levels (time increases from top to bottom.)
FIG. 19. Temporal evolution of the wave packet envelope $A(x)$, for the period shown in Fig. 18, for $v'$ at (left) the 100-Pa pressure level and (right) $v''$. 
FIG. 20. Vertical–zonal cross section of the wave shown in Figs. 18 and 19 at latitude 68°N. Shown are (left) \( v' \) and (right) \( \phi' \).
Fig. 21. Vertical–zonal cross section of the meridional mean of the temperature fluxes associated with the wave shown in Fig. 18 for 60°–80°N. Shown are (left) $T'v'$ and (right) $T'\omega'$. (See text in section 5c for an explanation of the meaning of the black arrows and the capital letters A–D.)
Fig. 22. Vertical–zonal cross section of the meridional mean of (left) $-V_z \cdot \nabla'_z \phi'$ and (right) $\nabla'_z \cdot (V'_z \cdot \nabla'_z) - \nabla'_z \cdot (V'_z \cdot \nabla'_z)v_m$ for the wave shown in Fig. 18 for 60°–80°N.
values near the top of the model atmosphere and showing those values would lead to a complete loss of details in the figures below 10 Pa. The dominant patterns are eastward-propagating wavelike structures in both fields. The amplitude of these wavelike structures, however, shows significant vertical and temporal variability in both fields. The largest amplitudes tend to occur above 100 hPa in both fields, but the convergence of the geopotential fluxes more often penetrates deep into the lower layers.

Comparing the left panels of Figs. 21 and 22, we find that a pattern of positive geopotential flux convergence is a precursor of an event of baroclinic energy conversion at the same location about ¼ Sol later. For instance, each black arrow and capital letter (A, B, C, or D) in Figs. 21 and 22 mark a pattern of positive geopotential flux convergence and the pattern of temperature fluxes generated by the related baroclinic energy conversion event. Because each major pattern of positive geopotential flux convergence is followed by a major pattern of negative geopotential flux convergence, at the time of baroclinic energy conversion, the geopotential flux convergence plays an important role in the downstream transport of the eddy kinetic energy from the region of the energy conversion.

A comparison of the left and right panels of Fig. 21 indicates that the geopotential flux convergence and the barotropic energy conversion are usually in sync. More precisely, when the eddies gain kinetic energy by the geopotential flux convergence, they also gain energy by barotropic energy conversion (from the basic flow), and when the eddies lose kinetic energy because of the downstream transport by the geopotential flux convergence, they also lose kinetic energy to the basic flow.

c. Regression analysis of the vertical structure of the waves

To investigate the typical vertical structure of the wave for the entire segment $L_\phi = 200^\circ - 230^\circ$, we carry out a regression analysis, using a technique introduced by Lim and Wallace (1991) and Chang (1993). The regression is constructed by starting with a reference time series $X(i), i = 1, 2, \ldots, N$, which is obtained by taking a time series of a selected variable at a given point for $N$ times, dividing each data point in the time series by the standard deviation of the data in the time series, and then computing the regression coefficients $b_j$ by

\[ b_j = \frac{1}{N} \sum_{i=1}^{N} Y(j)X(i) \]

for the locations indexed by $j$, $j = 1, \ldots, M$. In Eq. (8), $Y(j), i = 1, 2, \ldots, N, j = 1, \ldots, M$, is a time series of the eddy component of a state variable for the $M$ different locations and the $N$ times used for the construction of $X(i), i = 1, 2, \ldots, N$. The two variables used for the construction of $X(i)$ and $Y(j)$ can be the same or different state variables. A large (small) absolute value of $b_j$ indicates that a unit magnitude change in the state variable that defines $X(i), i = 1, 2, \ldots, N$, is associated with a large (small) change in the state variable that defines $Y(j)$.

Pressure–longitude cross sections of the regression coefficients can be obtained by choosing the locations $j = 1, \ldots, M$, to be the grid points at the different pressure levels and longitudes for a fixed latitude. We define $X(i), i = 1, 2, \ldots, N$ by $\phi'$ at the 100-Pa level at 68°N, 110°E, the location where the variance of the eddy kinetic energy at the 100-Pa level for $L_\phi = 200^\circ - 230^\circ$ is the largest. We show results for different choices of the state variables in $Y(j)$ (Fig. 23). The regression coefficients indicate a strong relationship between $\phi'$ at the selected location and the eddy components of the selected variables along the entire latitude circle, which is not unexpected considering the global nature of the wave. The relationship is also strong in the vertical direction throughout the entire depth of the atmosphere for $\phi'$ and $\phi'_0$, which suggests that phase locking between the upper- and lower-atmospheric waves near the location 68°N, 110°E occurs regularly. The regression coefficients also show, however, that a pronounced westward tilt of the structure of the wave with height is present only below the 10-Pa level. This result supports our earlier conclusion that baroclinic energy conversion takes place only below about 10 Pa.

d. Energy conversion processes

In our final analysis, we study the spatiotemporal evolution of the terms of the eddy kinetic energy equation for the time interval already studied in connection with Figs. 18–22. The episodes of intense baroclinic energy conversion are highly localized in both space and time despite the global character of the wave (Fig. 24). There are large changes in the maximum intensity of the baroclinic energy conversion as the patterns of intense energy conversion travel eastward with the wave. The magnitude of the geopotential flux convergence (right panels) is typically larger than the magnitude of the baroclinic energy conversion (left panels), which shows that the geopotential flux convergence can be a temporarily more important source of the eddy kinetic energy than baroclinic energy conversion. Of course, the generation of eddy kinetic energy by the geopotential flux convergence must be offset by the destruction of the same amount of eddy kinetic energy in the neighboring regions (Fig. 24). This process leads to the southeasterly...
The net transfer of the eddy kinetic energy by the geopotential flux convergence, which was shown in Fig. 11 and the bottom panel of Fig. 15. Figure 24 refines the picture that emerged from the earlier figures by showing that most of the eastward transport of the eddy kinetic energy by the geopotential flux convergence occurs in intense episodes.

As mentioned before, the eddy kinetic energy transported eastward by the geostrophic flux convergence can act as the trigger of a baroclinic energy conversion event. A comparison of the left and right columns of Fig. 24 suggests that each development of a local maximum of the baroclinic energy conversion is preceded by a local maximum of the geopotential flux convergence at the same location 0.5 Sol earlier. We recall that this behavior has already been discussed in connection with Figs. 21 and 22. To help identify the related features in Figs. 21, 22, and 24, the same capital letters are used to mark them in all three figures. (The 0.5-Sol lag between the local maxima of the two terms should not be considered a precise estimate, as it is obviously influenced by the model output frequency.) The timeline of events depicted by the three figures is indicative of baroclinic downstream development. While further studies based on longer time periods and constrained more strongly by observations will be necessary to confirm that baroclinic downstream development regularly occurs in the Martian atmosphere, our results clearly show that geopotential flux convergence plays an important role in the dynamics of the Martian baroclinic waves.

Similar to the magnitude of the geopotential flux convergence, the magnitude of the energy transport (left panels of Fig. 25) is often larger than the magnitude of the baroclinic energy conversion, which indicates that the contribution of the two transport processes to the local changes in the eddy kinetic energy is typically larger than the contribution of the two energy conversion processes. (The barotropic energy conversion is shown by the right panels of Fig. 25.)

The most intense events of negative barotropic energy conversion tend to occur northwest of the locations of the intense events of baroclinic energy conversion. The most intense events of positive barotropic energy conversion tend to occur north of 80°N between 200° and 280°E. In particular, strong positive barotropic energy conversion occurs at times when a significant amount of eddy kinetic energy is released from an intense event of upstream baroclinic energy conversion. (This relationship between the energy transport and the barotropic energy conversion can be observed by comparing the left and right panels of Fig. 25.)
6. Conclusions

We investigated the local dynamics of transient waves in the Northern Hemisphere of the Martian atmosphere in the pre-winter solstice season. While our investigation was based on a model simulation of the Martian atmosphere, we demonstrated that the general characteristics of the transient waves in our simulation were consistent with some important observed properties of the transient waves. There are, however, some properties of the simulated waves that are inconsistent with the observed properties of the waves that can affect our results. Most importantly, the wavenumber-3 component plays a less prominent role in the dynamics of the simulated...
waves than in the dynamics of the observed waves. In addition, the temporal variability of the characteristics of the simulated waves is more regular.

For the investigated time period, a wavenumber-2 wave with a spatiotemporally varying amplitude dominated the structure of the transient waves in the lower part of the atmosphere (below about 10 Pa). While the amplitude of the wave changed in both space and time, it never became localized in the zonal direction, thus preserving the global character of the wave. Based on the analysis of the different terms in the eddy kinetic energy equation for the wave, we drew the following conclusions:

- The transient waves of the lower atmosphere derive their kinetic energy from baroclinic energy conversion.
Although the role of baroclinic energy conversion in the Martian atmosphere is usually to maintain an existing wave rather than to start a new localized baroclinic development, the intense episodes of baroclinic energy conversion are spatiotemporally localized events.

- Barotropic energy conversion acts as both a source and a sink of the eddy kinetic energy: kinetic energy is transferred from the waves to the main flow in the regions of intense baroclinic energy conversion, while the waves are enhanced by a transfer of the kinetic energy from the main flow in a narrow zonal channel north of the Tharsis Plateau and in large regions southeast of the main regions of baroclinic energy conversion.

- Similar to the situation for the terrestrial atmosphere, geopotential flux convergence plays an important role in the dynamics of the downstream-propagating baroclinic waves. This result suggests that the classical theoretical models of baroclinic energy conversion (e.g., Eady model, Charney model, two-layer geostrophic model), which describe the spontaneous development of global waves in a baroclinically unstable environment, may have similar limitations for both Mars and Earth.

In the future, we hope to carry out a more rigorous analysis of the waves in the upper atmosphere (above 10 Pa). Since we found that the main source of eddy kinetic energy in that layer remained hidden in the residue term of the eddy kinetic energy equation and expect that gravity wave dynamics and parameterized physical processes play an important role, we believe that the investigation of those waves requires a major refinement of our diagnostic tools. We also hope to utilize our diagnostic tools to investigate the dynamical roots of the weaker-than-expected wavenumber-3 component in the simulations with the GFDL Mars GCM.

Acknowledgments. This work was supported by Grant NNX09AT57G from the Mars Fundamental Research Program of NASA and in part by Award KUS-C1-016-04 made by King Abdullah University of Science and Technology (KAUST). The contribution by R. J. W. was made possible by funding from the NASA Planetary Atmospheres Program and the Mars Data Assimilation Program. The many helpful comments by Jeffrey Barnes and anonymous reviewers on earlier manuscript versions of the paper helped us significantly improve the design of the numerical experiments and the analysis of the results.

REFERENCES


