The 4-Day Wave as Observed from the Upper Atmosphere Research Satellite Microwave Limb Sounder

D. R. Allen and J. L. Stanford

Department of Physics and Astronomy, Iowa State University, Ames, Iowa

L. S. Elson, E. F. Fishbein, L. Froidevaux, and J. W. Waters

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

(Manuscript received 16 February 1996, in final form 22 July 1996)

ABSTRACT

The “4-day wave” is an eastward moving quasi-nondispersive feature with period near 4 days occurring near the winter polar stratopause. This paper presents evidence of the 4-day feature in Microwave Limb Sounder (MLS) temperature, geopotential height, and ozone data from the late southern winters of 1992 and 1993. Space–time spectral analyses reveal a double-peaked temperature structure consisting of one peak near the stratopause and another in the lower mesosphere, with an out-of-phase relationship between the two peaks. This double-peaked structure is reminiscent of recent three-dimensional barotropic/baroclinic instability model predictions and is observed here for the first time. The height variation of the 4-day ozone signal is shown to compare well with a linear advective–photochemical tracer model. Negative regions of quasigeostrophic potential vorticity (PV) gradient and positive Eliassen–Palm flux divergence are shown to occur, consistent with instability dynamics playing a role in wave forcing. Spectral analyses of PV derived from MLS geopotential height fields reveal a 4-day signal peaking near the polar stratopause. The three-dimensional structure of the 4-day wave resembles the potential vorticity “charge” concept, wherein a PV anomaly in the atmosphere (analogous to an electrical charge in a dielectric material) induces a geopotential field, a vertically oriented temperature dipole, and circulation about the vertical axis.

1. Introduction

The 4-day wave is a ubiquitous feature in the polar winter upper stratosphere. It was first observed in temperature data as a strong 4-day signal in zonal wave-number 1 (Venne and Stanford 1979). Later studies showed that it consists of waves 1 through at least 4 all moving with the same phase speed, such that the period of wave 1 is near 4 days. Synoptic plots of temperature reveal a quasi-nondispersive “warm pool” of air, which rotates eastward around the winter pole near the stratopause with a 4-day rotation period (Prata 1984; Lait and Stanford 1988b; Lawrence et al. 1995). Numerous observational (Venne and Stanford 1979, 1982; Prata 1984; Lait and Stanford 1988b; Lawrence et al. 1995) observations (Venne and Stanford 1979, 1982; Prata 1984; Lait and Stanford 1988b; Randel and Lait 1991; Manney 1991; Fraser et al. 1993; Lawrence et al. 1995; Lawrence and Randel 1996) and theoretical (Hartmann 1983; Manney et al. 1988; Manney 1991; Manney and Randel 1993; Bowman and Chen 1994; Orsolini and Simon 1995) studies have examined the 4-day phenomenon.

The wave has been observed in both hemispheres but is stronger in the Southern Hemisphere (Venne and Stanford 1982). The horizontal phase structure is variable, with both poleward and equatorward momentum flux events (Venne and Stanford 1982; Prata 1984; Manney 1991). The wave is generally barotropic, although some episodes show strong equatorward heat flux (Randel and Lait 1991). The episodes studied by Randel and Lait (1991) and Lawrence and Randel (1996) reveal a dipole structure of positive and negative Eliassen–Palm flux divergence associated with the 4-day wave, with the positive region overlapping negative quasigeostrophic potential vorticity gradient, a signature of instability processes.

Theoretical studies have focused on instability of the jet structure as the mechanism for 4-day wave growth. Hartmann (1983) and Manney et al. (1988) showed that periods of 4-day wave growth are consistent with barotropic instability of the stratospheric polar night jet. Manney and Randel (1993) used a 3D instability model to show that realistic 4-day wave growth rates exist in monthly mean background wind states only if both barotropic and baroclinic components are included. Their
model also predicted a double-lobed structure in the perturbation temperature field with one lobe in the upper stratosphere and another in the lower mesosphere, with rapid phase variation in between.

The present paper takes advantage of the coverage and vertical resolution of the Microwave Limb Sounder (MLS) onboard the Upper Atmosphere Research Satellite (UARS) to examine the 4-day wave during the late austral winters of 1992 and 1993. As will be shown below, strong 4-day signals exist in MLS temperature, geopotential height, and ozone. The vertical coverage provided by MLS allows resolution of a double-peaked structure in perturbation temperature with an out-of-phase relationship between the upper and lower peaks, as predicted by Manney and Randel (1993). Spectral analyses of MLS ozone reveal a 4-day signal in zonal wavenumber 1 near the polar stratosopause. A linear advective–photochemical model is used to calculate the ozone response to temperature and wind perturbations associated with the 4-day wave. The vertical variation of model ozone amplitude and phase is in good agreement with observations of MLS ozone. Negative regions of quasigeostrophic potential vorticity (PV) gradient and positive Eliassen–Palm flux divergence are shown to occur, consistent with wave forcing by instability dynamics.

The 4-day feature is also shown to exist in quasigeostrophic PV derived from MLS geopotential heights. The three-dimensional structure of this signal resembles the PV “charge” discussed by Hoskins et al. (1985) and more recently by Bishop and Thorpe (1994). The latter authors have also extended the concept to Ertel PV (Thorpe and Bishop 1995), although the present paper focuses on quasigeostrophic PV. In this electrostatics analogy, a PV anomaly in the atmosphere acts like an electrical charge in a dielectric material, with an associated vector field (independent of static stability, density, and boundary conditions) that produces “action at a distance.” Isolated PV charges induce geopotential anomalies, circulation about the vertical axis, and a vertically oriented temperature dipole similar to observations of the 4-day wave presented in this study.

2. Microwave Limb Sounder data

a. The MLS instrument

The Microwave Limb Sounder is 1 of 10 instruments onboard the Upper Atmosphere Research Satellite. MLS uses a 1.6-m scanning antenna to observe the atmospheric limb emission simultaneously in spectral bands at 63, 183, and 205 GHz (Barath et al. 1993). These observations allow the determination of chlorine monoxide, ozone, water vapor, nitric acid, sulfur dioxide, temperature, and geopotential height in the stratosphere and mesosphere during both day and night and in the presence of stratospheric clouds and aerosols. This study analyzes temperature, geopotential height, and ozone data from two austral late-winter time periods, 14 August–20 September 1992 and 9 August–16 September 1993, when MLS was viewing southward (coverage from 80°S to 34°N).

b. Temperature data

MLS temperature is obtained by viewing limb emission of O$_3$ at 63 GHz. The version 3 data (used here) are retrieved on every other UARS pressure level $n$ (the UARS level-3 pressure grid is $1000 \times 10^{-n}$ hPa, where $n = 0, 1, 2, \ldots$) and have precisions ranging from 1.4 K at 22 hPa to 3.1 K at 0.46 hPa. MLS temperatures are used at eight pressure levels from 22 hPa to 0.1 hPa (22, 10, 4.6, 2.2, 1.0, 0.46, 0.22, 0.1 hPa). The last two levels are contaminated by large systematic errors and are biased by a zonal mean climatology. Wave structures estimated at these levels are attenuated and should be viewed with some caution. Validation information on version 3 MLS temperature is available in Fishbein et al. (1996).

c. Ozone data

MLS ozone retrievals are obtained from limb emission in two bands centered near 205 and 183 GHz. The version 3 ozone data are retrieved on every other UARS pressure level. The 205-GHz data have smaller uncertainties throughout the stratosphere, while the 183-GHz data are less noise limited in the mesosphere. Both 205- and 183-GHz ozone data are used in this paper from 22 to 0.22 hPa. Validation information on version 3 MLS ozone is available in Froidevaux et al. (1996) and Ricaud et al. (1996). Single-profile precision for ozone retrievals is in the 0.1–0.4 ppmv range, depending on altitude.

d. Geopotential height data

The MLS geopotential height dataset used in this study is a preliminary dataset, which is not available in version 3 but is planned in future versions. The geopotential height is obtained by measuring pressure as a function of tangent-point height. Pressure is obtained from the same radiances as temperature, but the radiances are sensitive to pressure from 10 hPa to 0.046 hPa. Tangent-point altitude is determined from the orientation of the MLS antenna, measured by an angle encoder, the UARS attitude control system, and the ground-tracked UARS orbit determination. The pressure (in log pressure coordinates with atmospheric scale height 7 km) precision ranges from 5 m at 10 hPa to 10 m at 0.1 hPa, while the altitude precision is altitude independent and is estimated to be around 100 m. Errors in geopotential height (excluding precision) arising from the tangent-point altitude are synchronized to the UARS orbit and contaminate the zonal mean and high frequencies near the diurnal tide. However, the analyses
Fig. 1. (a) East–west power spectrum for wave 1 MLS temperature ($T_{\text{retrieved}}$) at 72°S for 18 August–6 September 1992. Dashed contours are at 10, 20, and 30 K day, and solid contours are from 40 to 100 K day by 10 K day. (b) Same as (a) but for temperature derived from MLS geopotential height ($T_{\text{geoht}}$). (c),(d) Same as (a),(b) but for 2–16 September 1993.

performed in this paper (zonal wave 1, 3–5-day periods) have sufficiently long periods to avoid significant effects from these errors. Good agreement is found on these spatial and temporal scales between MLS temperature (coined $T_{\text{retrieved}}$ in this paper) and temperature derived from MLS geopotential height data ($T_{\text{geoht}}$), as discussed in section 4a.

3. Analysis procedure

MLS data used in this analysis were produced by the asynoptic mapping method (Salby 1982a,b; Lait and Stanford 1988a). The specific technique used for this study is presented in detail in Elson and Froidevaux (1993). Briefly, data are binned by latitude and separated into ascending and descending time series. The time and space axes are rotated to a new coordinate system where a fast Fourier transform (FFT) is applied to calculate the spectral coefficients. An inverse transformation and interpolation allows data to be mapped onto synoptic maps, which are processed every 12 hours. Satellite orbital period drift generally limits the procedure to approximately 1-week analysis periods. The procedure here is essentially a remapping in time so that periods longer than a week can be analyzed. Data are interpolated onto horizontal grids of 4° latitude by 5° longitude and onto the standard UARS pressure levels.

The twice-daily synoptic maps described above were decomposed in space and time by applying a two-dimensional sinusoidal transform [see appendix of Ziemke and Stanford (1990)]. Power spectral density is defined in this study as $(2\Delta f)^{-1}[A(k, \omega) + B(k, \omega)]$, and amplitude is defined as $[A^2(k, \omega) + B^2(k, \omega)]^{1/2}$, where $\Delta f$ is the unit bandwidth (NDAT $\Delta t$)$^{-1}$. NDAT is the number of data points in the time series (NDAT=38 for 1992, NDAT=30 for 1993) and $\Delta t$ is the temporal sampling interval (0.5 day); $A(k, \omega)$ and $B(k, \omega)$ are the cosine and sine coefficients of the discrete Fourier expansions.
were calculated using the spherical y
y
p
zonal grid interval of 2
radius; and number;
t components; half-amplitude points at 3 and 5 day periods) given by
waves.
focal traveling
struments. In the EP
FIG. 2. Time vs longitude plot of $T_{\text{zonal}}$ at 64øS and 1 hPa (48 km) filtered for wave 1 and 3–5 day periods (eastward and westward) for 14 August–20 September 1992. Dashed contours are from −4 to 0 K by 1 K. Solid contours are from 1 to 4 K by 1 K.

\begin{align}
\{T_{z}, \Phi_{z}, \mu_{z}\}(x, t) &= \sum_{k=0}^{m/\Delta k} \sum_{\omega=0}^{n/\Delta \omega} \left[ A_{z}(k, \omega) \cos(kx \pm \omega t) \right. \\
&\quad \left. + B_{z}(k, \omega) \sin(kx \pm \omega t) \right].
\end{align}

Here $T$, $\Phi$, and $\mu$ are temperature, geopotential, and ozone mixing ratio; + (−) refers to westward (eastward) components; $t$ is time; $x$ is zonal distance; $\Delta k$ is the zonal grid interval of $2\pi a \cos \phi / 72$; $k$ is the zonal wave-number; $\omega$ is the angular frequency; $a$ is the earth’s radius; and $\phi$ is latitude. A 1–2–1 running mean was applied in frequency space to power spectral density, but not to amplitude. The time mean was removed before applying the FFT to focus the analysis on traveling waves.

Filtered time series in this study were produced by first applying the bandpass filter response function (with half-amplitude points at 3 and 5 day periods) given by Murakami (1979) to the temporal FFT coefficients and then recombining the coefficients through an inverse FFT.

As a comparison to the MLS temperature measurements, temperature was also calculated from MLS geopotential height using the hydrostatic approximation $T = R^{-1} \phi \bar{\rho}(\ln \rho)$, with $R = 287$ J kg$^{-1}$ K$^{-1}$ (dry air gas constant) and constant vertical spacing $d(\ln \rho) = 0.387$ (determined from UARS standard pressure surfaces). Quasigeostrophic potential vorticity $\mathbf{q}'$ and potential vorticity gradient $\bar{\mathbf{q}}$, were calculated using the spherical formulation of Matsuno (1970):

\begin{align}
\mathbf{q}' &= \frac{1}{fa^2} \left[ \Phi_{\phi\phi} + f^2 \cos \phi \left( \frac{\cos \phi}{F} \Phi_{\phi} \right)_{\phi} \right. \\
&\quad \left. + \frac{f^2 a^2}{\rho_0} \left( \frac{\rho_0 \Phi'}{N^2} \right) \right] z, \\
\bar{\mathbf{q}} &= \frac{2\Omega \cos \phi}{a} - \frac{ \left( \bar{u} \cos \phi \right)_{\phi} }{a^2 \cos \phi} - \frac{1}{\rho_0} \left( \frac{\rho_0 F^2}{N^2} \bar{u} \right),
\end{align}

Here $u$ is the zonal wind, $\rho_0 = \rho_0 e^{-\omega t}$, $\rho_0$ is the density at the surface (1000 hPa), $z$ is the log-pressure coordinate, $\lambda$ is longitude, $f$ is the Coriolis parameter, and $\Omega$ is the earth’s angular frequency. Overbars denote zonal means, primes denote deviation from the zonal mean, and subscripts denote partial derivatives. The Brunt–Väisälä parameter and atmospheric scale height were given the constant values $N^2 = 4 \times 10^{-4}$ s$^{-2}$ and $H = 7$ km; $\mathbf{q}'$ was calculated from MLS geopotential height data, while $\bar{\mathbf{q}}$ was calculated from U.K. Meteorological Office (UKMO) zonal winds. MLS data were not used in the calculation of $\bar{\mathbf{q}}$ due to the known inaccuracy in the zonal mean geopotential height (see section 2d).

UKMO horizontal winds and temperatures on UARS pressure surfaces, obtained from the Goddard Space Flight Center Distributed Active Archive Center (DAAC), were interpolated with a cubic spline onto horizontal grids of 4ø latitude by 5ø longitude. These winds and temperatures were used in the calculation of $\bar{\mathbf{q}}$, and also in the tracer model discussed in section 4c.

The quasigeostrophic Eliassen–Palm (EP) flux vector and the EP flux divergence were calculated with the following spherical formulation:

\begin{align}
\mathbf{F} &= (F_{\phi'}, F_{\phi}) = \rho_0 c \cos \phi \left( -u' \hat{u}', \frac{R}{H} \frac{\nabla T}{N^2} \right), \\
\nabla \cdot \mathbf{F} &= \frac{1}{a} \frac{\partial}{\partial \phi} (\cos \phi F_{\phi}) + \frac{\partial F_{\phi}}{\partial \phi}.
\end{align}

Here $u'$ and $\nu'$ (meridional wind) are calculated from MLS geopotential heights (using the geostrophic approximation) and $T'$ is from MLS-retrieved temperatures. In the EP flux diagram (Fig. 11), the vector components are multiplied by $2\pi a \cos \phi$. This factor accounts for the spherical geometry of the earth to make the arrow pattern look nondivergent if and only if $\nabla \cdot \mathbf{F} = 0$ (see Edmon et al. 1980).
4. Results

a. Temperature

Space–time spectral decomposition is a useful tool for isolating the 4-day wave from other atmospheric signals. Figure 1a presents a periodogram for zonal wavenumber 1 power spectral density (defined in section 3 and hereafter called power) for MLS-retrieved temperature (hereafter $T_{\text{retrieved}}$) at 72°S latitude for 18 August–6 September 1992. A clear 4-day eastward period signal exists near 1 hPa (48 km), and a somewhat weaker signal occurs near 0.22 hPa (60 km). These signals are also evident in temperature derived from MLS geopotential height (hereafter $T_{\text{geoht}}$) in Fig. 1b, although the lower signal maximizes near 2.2 hPa (42 km). An even stronger 4-day signal is found during 2–16 September 1993 in both $T_{\text{retrieved}}$ (Fig. 1c) and $T_{\text{geoht}}$ (Fig. 1d). The 4-day signal dominates the traveling wave spectra for this time period with two lobes centered near 2.2 hPa (42 km) and 0.22 hPa (60 km). Sections 4d and 4f further discuss the validity of the 4-day signal in terms of uncertainty and statistical significance.

The 1992 4-day wave event was further isolated by filtering a 38-day series of $T_{\text{retrieved}}$ for zonal wavenumber 1 and periods (both eastward and westward) of 3–5 days and plotting on a Hovmöller (time vs longitude) diagram for 64°S and 1 hPa (Fig. 2). The 4-day wave amplitude increases over the first 10 days and remains constant for another 10 days before dissipating near 1 September 1992. The mode makes at least five revolutions around the pole with a rotation period near 3.8 days. Spectral plots of $T_{\text{retrieved}}$ for zonal wavenumbers 2 and 3 (not shown here) for the 1992 event show signals at 1 hPa with eastward periods of 2.0 and 1.3 days, respectively. These sinusoidal components move at the same phase speed of the wave 1, 4-day signal and combine with wave 1 to form a “warm pool” of air at 1 hPa (coupled with a slightly less prominent “cold pool”) that rotates around the pole with a period near 4 days. This feature is shown in Fig. 3, where $T_{\text{retrieved}}$ data have been filtered over 14 August–20 September 1992 for waves 1–3 and eastward rotation periods of 3–5 days. (Wave 4 and higher were not analyzed because the frequency of interest, 4 days divided by integer wavenumber, is near or outside the Nyquist limit of approximately 1 day$^{-1}$.) The warm pool rotates around the pole as an identifiable entity throughout the record shown displaying the quasi-nondispersive (soliton-like) nature of the feature.

The perturbation temperature ($T_{\text{retrieved}}$) amplitude for wave 1 and 3.8-day eastward period is plotted in Fig. 4a for the 1992 4-day wave event. A large lobe exists in the upper stratosphere, peaking near 48 km and 72°S with maximum amplitude of 4.2 K. A smaller upper lobe is centered at 60 km and 68°S. This double-lobed structure is similar to that predicted by the instability study of Manney and Randel (1993). In their results, the upper lobe is larger and extends well into the meso-
February 1997

**Figure 4.**
(a) $T_{\text{retrieved}}$ amplitude for wave 1 and 3.8-day eastward period for 18 August–6 September 1992. Contours are every 10% interval of the maximum (4.2 K), with values higher than 50% shaded. (b) $T_{\text{retrieved}}$ amplitude for wave 1 and 3.75-day eastward period for 2–16 September 1993. Contours are every 10% of the maximum (5.8 K), with values higher than 50% shaded.

...sphere. However, $T_{\text{retrieved}}$ is largely climatology at 66 km, forcing zonally asymmetric wave amplitudes to zero; therefore, the abrupt fall-off with height above 60 km in Fig. 4a may not be real. The 1993 series also displays a double-lobed structure (Fig. 4b), with maxima near 42 km, 76°S (amplitude of 5.8 K) and 60 km, 80°S.

Manney and Randel (1993) also predict rapid vertical phase variation between the two lobes. Such behavior can be explicitly seen in height versus longitude plots of $T_{\text{retrieved}}$ at 72°S, filtered for waves 1–3 and eastward rotation periods of 3–5 days. Figure 5a presents 8 days during the peak amplitude of the 1992 time series. The plots reveal a warm anomaly near 48 km bounded above by a cold anomaly near 60 km. This dipole pattern continues throughout the 4-day wave life cycle, suggesting that the two peaks constitute one entity. The 1993 series also reveals this vertical temperature dipole (Fig. 5b). Here the lower peak has a strong wave 1 structure that begins near 48 km and descends to 42 km over this 8-day period. The upper feature, which maximizes near 60 km, is approximately 180° out of phase with the lower feature. The quadrupole pattern produced by the two out-of-phase lobes remains stable for at least two weeks, through the end of the UARS month (16 September 1993), when MLS changed to northward viewing (coverage from 80°N to 34°S).

**b. Geopotential**

The vertical temperature dipole structure seen in Fig. 5 suggests that the perturbation geopotential amplitude should increase with height throughout the region of the lower temperature lobe and decrease with height in the region of the upper lobe. This is confirmed in Fig. 6a, which shows the geopotential amplitude for the 1992 series for eastward moving wave 1 with a period of 3.8 days. The signal maximizes near 57 km, at the boundary between the warm and cold anomalies (see Fig. 5a), and decreases above it. The 1993 geopotential amplitude (Fig. 6b) peaks near 48 km, consistent with the perturbation temperature structure for that series (Fig. 5b). The geopotential signal decreases with altitude more rapidly in the 1993 time series, due to a larger-amplitude upper temperature lobe in 1993 than in 1992. The meridional structure shown here is similar to previous observations of the 4-day wave in geopotential height (Manney 1991; Lawrence and Randel 1996).

**c. Ozone**

Other than the unpublished conference report by Randel et al. (1992), the present paper is the first to present observations of the 4-day wave in ozone. Periodograms for wave 1 MLS ozone (both 205- and 183-GHz data) at 72°S during 18 August–6 September 1992 are presented in Fig. 7. A 4-day peak exists near the stratopause (1.0 hPa), coincident with the $T_{\text{retrieved}}$ maximum at the same altitude, see Fig. 1a. A 2-day eastward period signal exists at 2.2 hPa; the source of this signal is unknown and is not thought to be related to the 4-day wave. The 2-day feature does not exist in spectra from 5 to 20 September 1992.

To ascertain the relationship between the 4-day feature in ozone and the other atmospheric variables, a linear advective–photochemical model was used to calculate the ozone response to temperature and meridional wind perturbations associated with the 4-day wave. The procedure combines the linearized thermodynamic and ozone continuity equations to obtain an expression for the ozone mixing ratio (Hartmann and Garcia 1979; Rood and Douglass 1985; Randel 1990, 1993). Because the 4-day wave amplitudes in the various components ($T$, $u$, $v$, and $\mu$) are observed to be much less than their...
Fig. 5. (a) $T_{\text{retrieved}}$ on a height vs longitude grid at 72°S filtered for waves 1–3 and 3–5-day eastward rotation period for 22–29 August 1992. (b) Same as (a) but for 4–11 September 1993. Contours are every 1 K with solid (dashed) line having positive (negative) values. Here, W (C) refers to warm (cold) anomalies.

respective zonal mean values ($T' \ll \bar{T}$, etc.), we use a linearized thermodynamic equation,

$$T' + \bar{u} T'_{\text{e}} + \omega v' T_{\text{e}} + w' S = 0$$

$$S = \frac{1}{H} [\kappa T + HT_{\text{e}}],$$

(6)

where $w'$ is the vertical wind component, $\kappa = R/c_p \approx 2/7$, and $H = 7$ km; and a linearized ozone continuity equation,

$$\mu' + \bar{u} \mu'_{\text{e}} + \omega \bar{\mu} + w' \bar{\mu} = -\Gamma \mu' - \Theta T'.$$

(7)

The photochemical relaxation rate, $\Gamma$, and the linearized ozone response to small temperature perturbations, $\Theta$, are shown in Fig. 8a, adapted from Hartmann and Garcia (1979), who studied ozone transport by planetary waves in the winter stratosphere using a linearized $\beta$-plane model centered at 60°. Radiative damping of the temperature perturbations is neglected, since the radiative timescale is of order 10 days in the region of interest (26–57 km) [see Hartmann and Garcia (1979)], while the advective timescale $[k(\bar{u} - c)]^{-1}$ is of order 1–2 days from 26 to 42 km, and the photochemical timescale $\Gamma^{-1}$ is less than 1 day from 42 to 57 km. Radiative heating
from ozone perturbations has also been neglected, as in Randel (1990, 1993), since it is likely small compared to changes due to advection. Solving Eq.(6) for \(w'\) and substituting into Eq.(7) gives the following differential equation for \(\mu'\):

\[
\left(\frac{\partial}{\partial t} + \bar{u}\frac{\partial}{\partial x} + \Gamma\right)\mu' + \left(\bar{\mu}_z - \frac{\bar{\mu}_x}{S} T_z\right)v' = 0.
\]  

To relate \(\mu'\) to \(T'\) and \(v'\), we write \((\mu', T', v') = (M, T, V)e^{i(kx - \omega t)}\), where \(M, T,\) and \(V\) denote complex-valued perturbation amplitudes. From this,

\[
M = -\frac{(\Theta - ib)T + \xi V}{\Gamma + ia}
\]

\[
(9)
\]

\[
\xi = \bar{\mu}_z - \frac{\bar{\mu}_x}{S} T_z, \quad a = k(\bar{u} - c), \quad b = k(\bar{u} - c)\bar{\mu}_z.
\]

Figure 8b presents the resulting model ozone amplitude for the wave 1, 3.8-day period signal at 60°S (thick solid line). Also plotted are the observed ozone amplitudes at 60°S for 205-GHz (thin solid) and 183-GHz (dashed) data. Error bars on the observed data were calculated using the procedure described in section 4d. Reasonable agreement is found between the observed and model amplitudes, with low values in the middle stratosphere and lower mesosphere, and a maximum in the upper stratosphere (near 42 km). Results from a purely advective model (\(\Gamma = \Theta = 0\)) and a purely photochemical model \(\left[\mu_{amp} = (\Theta/\Gamma) T_{amp}\right]\), see Randel (1990)) show a transition region between dynamical and photochemical control from approximately 37 to 49 km, where both dynamics and photochemistry play an important role.

The phase of the model and observed 4-day feature is provided in Fig. 8c. Good agreement exists between the 205-GHz observed phase and the model throughout the stratosphere, while the 183-GHz observed phase agrees with the model in the lower mesosphere but diverges somewhat in the stratosphere. As mentioned in section 2c, the 205-GHz data have smaller uncertainties in the stratosphere, while the 183-GHz data are less noise limited in the mesosphere, so the phase results shown here are consistent with the data quality.

The good agreement (at least to first order) between model and observations suggests that the observed 4-day signal in ozone is produced largely by the coupled effects of constituent advection by the 4-day wave perturbation winds in the presence of zonal-mean constituent gradients and photochemical processes associated with the 4-day wave perturbation temperature and mixing ratio anomalies.

d. Uncertainty in 4-day wave amplitude

In order to assess the uncertainty in the 4-day wave amplitude calculation, a random number generator was used to make a noise dataset with level 3AT format (along-track, evenly spaced in time) with values between \(-0.5\) and \(+0.5\) (arbitrary units), giving a root mean square (rms) variation of 0.29. This represents the random noise in a given MLS measured quantity with an rms error (defined here as uncertainty) of 0.29. These data were passed through the asynoptic mapping routine described in section 2 and synoptic maps were produced. The rms variation in the synoptic maps is 0.19, which is 65% of the input value. The decrease is mainly due to interpolation to a regular horizontal grid. Fourier decomposition was then applied to isolate the amplitude of the wave 1, 3.8-day component from 500 different time series, each randomly generated. This Monte Carlo procedure yielded rms variations in amplitude of 0.018
for the 1992 simulated series (38 time components used, 72 longitudes) and 0.020 for the 1993 simulated series (30 time components used, 72 longitudes). The uncertainty in amplitude is 0.018/0.29 = 1/16 times the uncertainty in the level 3AT data for 1992 and 1/14 for 1993. The difference between 1/16 and 1/14 is due to the different bandwidth used in each year. This method can now be used to obtain estimated uncertainties in 4-day wave amplitudes observed in MLS temperature and ozone.

The estimated uncertainty in version 3 MLS temperature at 1 and 2.2 hPa is 3.4 and 2.2 K, respectively (Fishbein et al. 1996). The estimated wave 1, 3.8-day eastward period amplitude uncertainty in the lower peak of Fig. 4a (1992) is 3.4/16 = 0.21 K, which is much smaller than the observed amplitude of 4.2 K. In 1993 the uncertainty in the lower peak (Fig. 4b) is 2.2/14 = 0.16 K, which is much less than the observed amplitude of 5.8 K. This Monte Carlo analysis shows that the observed amplitudes are not significantly affected by random noise in the data.

This method was also used to obtain estimated errors for 205- and 183-GHz MLS ozone data. The error bars shown in Fig. 8b were computed using the above formulation along with data uncertainties at different pressure levels as given by Froidevaux et al. (1996).

### e. Statistical assessments

Because the 4-day wave is a well-established atmospheric phenomenon (see references in the introduction), statistical assessment of the 4-day features found here can be made on the basis of a priori statistics (Madden and Julian 1971). Briefly, before analysis it is decided to examine spectral features near 4-day periods. If spectral peaks are examined at previously unreported frequencies, a posteriori statistics must be used (Madden and Julian 1971).] The null hypothesis is made that any spectral peak at 4 days is due to random sampling fluctuations. Under this hypothesis, the ratio of spectral peak to nearby spectral background can be estimated statistically. For the Fourier transform spectra used here, chi-squared statistics are employed to estimate the maximum peak to background ratio that will occur in 95% of the cases for random sampling fluctuations. If the observed peak exceeds this ratio, with 95% confidence the spectral feature can be said to be statistically significant.

Figures 9a,b show the $T_{\text{retrieved}}$ spectra at 2 and 0.2 hPa during 2–16 September 1993. These are plotted without smoothing in Figs. 9a,b and correspond to the smoothed plot shown in Fig. 1c. Also indicated by dashed lines in Fig. 9 are the estimated background and 95% confidence limit (for two degrees of freedom, appropriate to this case of unsmoothed Fourier transform spectral estimates). Both features easily satisfy a priori statistical significance.

The 1992 temperature spectral 4-day peak, shown smoothed in Fig. 1a, reveals a strong peak centered at 1 hPa that also easily exceeds the a priori 95% confidence limit (not shown here). The weaker temperature peak at 0.22 hPa is close to a stronger lower-frequency feature. By examination of unsmoothed spectral results (not shown here), it can be shown that the 0.22-hPa 4-day peak also meets the criterion for a priori statistical significance.

The 183-GHz ozone spectrum at 1 hPa for 1992 (shown smoothed in Fig. 7b) is plotted unsmoothed in Fig. 9c. The 4-day peak exceeds the 95% confidence limit and is judged statistically significant. The 205-GHz ozone 4-day peak at 1 hPa (Fig. 7b) also exceeds the 95% confidence limit (not shown here).
f. Zonal wind, $\bar{q}_k$, and EP flux structure

Theoretical studies show that barotropic and/or baroclinic instability of the polar night jet can produce quasi-nondispersive modes similar to the observed 4-day wave. The instability episodes are generally accompanied by regions of negative quasigeostrophic potential vorticity gradient $\bar{q}_k$ and nonzero EP flux divergence. Figure 10a shows the zonal mean UKMO east-
west wind averaged over the life cycle of the 1992 4-day wave event (18 August–6 September 1992). A strong stratospheric westerly jet is centered near 34 km, 60°S. The associated wind curvature produces a region of negative $q_y$ near 50°S at altitudes above 50 km (Fig. 10b). Daily wind maps reveal a “double jet” structure near the beginning of the time series, which is often present during 4-day wave events (Manney 1991; Randel and Lait 1991). Manney et al. (1996) showed that daily maps of Ertel PV (on isentropic surfaces) calculated from UKMO winds sometimes exhibit unrealistic structure in the Southern Hemisphere winter middle and upper stratosphere poleward of about 70°S. Our analyses here, however, focus on quasigeostrophic PV averaged over 20 days. Comparison of Fig. 10a with National Centers for Environmental Prediction (NCEP, formerly NMC) zonal winds (not shown here) suggests good overall agreement. However, $q_y$ calculated from the NCEP data (which extends only to 54 km) does not show the negative region near 50°S for this time-averaged period, although negative $q_y$ does exist near 50°S, 54 km for several of the individual NCEP-derived daily maps during the first half of this time series. (Note that time-averaged fields are more smoothed than the wind field on any given day, so they often underestimate instability.)

Linear theory requires that wave growth by barotropic instability be accompanied by negative PV gradients. [The integral over latitude of $q_y$ times a positive definite quantity must be zero for wave growth to occur by barotropic instability. However, baroclinic instability is possible with positive $q_y$, given suitable variation with latitude of the vertical gradient of $u$ at the lower boundary; see Holton (1992).] The presence here of negative $q_y$ regions is consistent with the fact that negative $q_y$ is a necessary, although not sufficient, condition for wave growth by barotropic instability.

Figure 11 displays the EP flux diagram for this wave episode, giving details about the net effect on the zonal mean state of eddy and momentum fluxes due to wave events. The EP flux was calculated using MLS temperature and geopotential height from 18 August to 6 September 1992, filtered for the wave 1 and 3.8-day eastward period. The geostrophic approximation was used to calculate the winds from the 4-day wave perturbation geopotential field. The EP vectors are plotted over contours of $DF = \nabla \mathbf{F} / (\rho a \cos \phi)$, which is a force per unit mass on the zonal-mean wind (in m s$^{-1}$ day$^{-1}$) due to the eddy effects associated with the wave 1 component of the 4-day wave. Positive (negative) regions of $DF$ act to accelerate (decelerate) the zonal
mean westerly wind. Positive $DF$ exists between 50° and 75°S at the top of the plot, near 57 km, acting to accelerate the lower-mesospheric winds, while a negative $DF$ region near 54 km, 72°S acts to decelerate the stronger stratopause winds (see Fig. 10a). Figure 11 predicts eddy-induced reduction of vertical shear in the 60°–75°S lower-mesosphere region, consistent with baroclinic instability mechanisms driving the 4-day wave in the lower mesosphere. The EP flux structure above 48 km is similar to the model results from Manney and Randel (1993, their Fig. 2d), with a region of negative $DF$ bounded above and to the north by positive $DF$.

The EP flux vectors indicate strong equatorward heat flux (downward pointing arrows) and equatorward momentum flux (poleward pointing arrows) above 48 km and poleward of 65°S (also similar to Manney and Randel 1993, Fig. 2d) and largely poleward momentum flux elsewhere. The vectors are predominantly vertical (indicating baroclinic processes) in the lower mesosphere, while strong momentum flux (indicated by horizontal components of EP flux vectors) near 42 km, 60°S suggests that barotropic processes play a sizable role there. This is confirmed by the positive (70°S) and negative (60°S) $DF$ regions at 42 km, which act to reduce the latitudinal shear associated with the stratospheric jet (a signature of barotropic instability).

Manney and Randel (1993) and Lawrence and Randel (1996) show overlapping regions of negative $\bar{q}_r$ and positive $\nabla \cdot \mathbf{F}$ during 4-day wave episodes. This spatial coincidence is recognized as a source region of wave activity $A = (1/2)q''^2/\bar{q}_r$, where $q''$ represents the perturbation quasigeostrophic potential vorticity. Neglecting diabatic and frictional effects, a conservation equation can be written (Andrews and McIntyre 1976, 1978):

$$\frac{\partial A}{\partial t} + \nabla \cdot \mathbf{F} = 0. \quad (10)$$

Clearly, when $\bar{q}_r$ is negative, the wave activity will increase in magnitude when $\nabla \cdot \mathbf{F} > 0$. One region of negative $\bar{q}_r$, occurs in the vicinity of positive $DF$ near 57 km, 52°S. This is consistent with instability dynamics driving the 4-day wave in the lower mesosphere. According to Eq. (10), wave activity will also increase if $\bar{q}_r > 0$ and $\nabla \cdot \mathbf{F} < 0$. The regions of negative $DF$ near 54 km, 80°S–65°S and 42 km, 60°S overlap positive $\bar{q}_r$, suggesting instability forcing may also exist in these regions.

g. Quasigeostrophic potential vorticity

Spectral analyses of quasigeostrophic PV calculated from MLS geopotential heights reveal 4-day signals in both the 1992 and 1993 series. Figure 12 shows periodograms for 72°S during 18 August–6 September 1992 (Fig. 12a) and 2–16 September 1993 (Fig. 12b). Clear 4-day eastward period signals in PV are seen near the stratopause (48 km or 1 hPa) during both time periods. (In 1992, a significant longer-period feature is seen in the upper stratosphere. This eastward moving power with periods of several weeks is not directly related to the 4-day wave considered here.) A height versus latitude plot of PV amplitude for wave 1 and 3.8-day eastward period is provided in Fig. 13a for the 1992 series. A region of large PV amplitude is centered near 50 km, 68°S. Figure 13b shows the wave 1, 3.75-day signal in 1993, which has a similar structure to the 1992 feature, although centered slightly lower and with contours somewhat more compact.

The three-dimensional structure of the perturbation PV resembles the “PV charge” concept discussed by Hoskins et al. (1985) and more recently by Bishop and Thorpe (1994), who considered the effects on the atmosphere of a uniform spherical quasigeostrophic PV
anomaly in a basic state with constant static stability and density and in a two-layer vertical static stability basic state representing the effects of the tropopause. In this analogy, the atmosphere represents an isotropic dielectric material, static stability represents dielectric properties, geopotential (or streamfunction) represents electric potential, and an isolated PV anomaly represents an electrostatic charge, with an accompanying vector field that produces action-at-a-distance (a fundamental concept in PV thinking). This “PV field” (which is independent of static stability, density, and boundary conditions) induces perturbation temperature, geopotential, and wind fields as shown in Fig. 14a (for the case of constant static stability), adapted from Figs. 1 and 2 of Bishop and Thorpe (1994). Here a negative spherical PV charge induces a vertical temperature dipole with cold temperature perturbation on top, spherical geopotential anomaly, and anticyclonic circulation about the vertical axis. Note that circulation shown here is for the Southern Hemisphere ($f < 0$); the direction should be reversed for the Northern Hemisphere ($f > 0$).

In comparison, Fig. 14b shows the amplitude of the wave 1, 3.8-day period mode during the 1992 series as a function of height and latitude. The hatched region includes PV amplitude greater than 90% of the maximum. Solid lines are perturbation temperature amplitude of 60%–90%, incremented by 10% of the maximum. Dashed lines are perturbation geopotential amplitude at every 10% of the maximum. Adapted from Figs. 4a, 6a, and 13a of this paper.
Solid lines are perturbation temperature amplitude of 60%–90% of the maximum, incremented by 10%, and the dashed lines are perturbation geopotential amplitude at every 10% of the maximum. Here we see the PV anomaly bounded above and below by two temperature lobes, which are out of phase, resembling the cold and warm perturbations in Fig. 14a. The geopotential anomaly surrounds the region of large PV amplitude, although centered slightly higher. There would also be an induced circulation about the vertical axis associated with the geopotential anomaly, similar to that shown in Fig. 14a.

This analogy suggests that the 4-day wave may be more fundamentally viewed as a PV anomaly than a temperature anomaly. The wave life cycle begins with an unstable atmosphere that allows growth in PV anomaly amplitude, which in this paradigm induces anomalies in temperature, geopotential, and wind. This view also appeals to the “wave activity” interpretation of eddy features discussed in section 4f, since PV anomalies relate directly to the EP flux divergence forcing of the zonal-mean state. The PV charge paradigm, coupled with its generation by instability mechanisms, appears to provide a satisfying physical explanation for the 4-day wave.

5. Summary

This study takes advantage of the vertical coverage and resolution of the Microwave Limb Sounder instrument to analyze 4-day wave features during August–September 1992 and 1993. Strong 4-day wave events are found during both time series. A double-peaked structure is found in temperature with maxima in the upper stratosphere and lower mesosphere and strong vertical phase variation in between, as predicted by the model of Manney and Randel (1993). Four-day eastward period signals are also evident in MLS ozone, geopotential height, and quasigeostrophic potential vorticity derived from MLS geopotential height. The ozone signal compares well with results from a linear advective–photochemical model that calculates the ozone response to temperature and wind perturbations associated with the 4-day wave. A negative region of quasigeostrophic potential vorticity gradient coexists with positive EP flux divergence, suggesting that instability dynamics plays a role in the wave forcing. The three-dimensional wave structure is shown to resemble the potential vorticity charge concept, which includes a vertical temperature dipole and anomalies in PV, geopotential, and horizontal wind. The PV paradigm, in which temperature, geopotential, and wind anomalies are induced by a potential vorticity charge, itself the result of instability growth processes, offers an elegant and conceptually satisfying explanation of the physics of the 4-day wave.

Acknowledgments. We want to thank Dr. Jerald Ziemke for assistance in spectral computations and graphics. The paper benefited from conversations with Dr. Gloria Manney (who also provided the NCEP data) and Dr. William Randel, and helpful comments by two anonymous reviewers. Drs. R. Swinbank and A. O’Neill developed the UKMO data that we obtained from the EOS Distributed Active Archive Center (DAAC, code 902.2) at the Goddard Space Flight Center. The activities of the EOS DAAC and the Upper Atmosphere Research Satellite Project (code 916) are sponsored by NASA’s Mission to Planet Earth Program. ISU coauthors are sponsored in part by National Aeronautics and Space Administration Grant NAG 5-2787. Work at the Jet Propulsion Laboratory, California Institute of Technology, was performed under contract with the National Aeronautics and Space Administration. Part of this work was done by D. Allen as a guest graduate student at Argonne National Laboratory.

REFERENCES


