Structure and Evolution of Baroclinic Waves as Inferred from Regression Analysis

GYU HO LIM* AND JOHN M. WALLACE

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

(Manuscript received 2 May 1990, in final form 27 February 1991)

ABSTRACT

The structure of transient disturbances with periods shorter than a week is documented on the basis of one-point regression maps and longitude–height sections derived from the National Meteorological Center (NMC) operational analyses for nine winter seasons. The reference time series for the regression coefficients is highpass filtered, normalized 500-mb height at a grid point in the Pacific stormtrack. The field variables that were regressed on the reference time series were not filtered. The regression statistics yield quantitative estimates of the amplitudes of the high-frequency fluctuations in the geopotential height, temperature, wind, ageostrophic wind, vorticity, divergence, vertical velocity, and outgoing longwave radiation fields.

The observed high-frequency fluctuations exhibit a vertical and meridional structure suggestive of baroclinic waves evolving through the mature phase of their life cycle. Most of the results are in qualitative agreement with previous studies, but the ageostrophic wind field in the lower troposphere exhibits systematic fluctuations in the meridional component that had not been anticipated.

An analysis of the temperature and vorticity tendencies in the waves yielded dynamically plausible results, despite the deficiencies in the NMC divergence and vertical velocity fields. The vertical velocity field at the 500-mb level was found to be consistent with the forcing in the quasi-geostrophic omega equation, which exhibits a simple wavelike pattern in the Q-vector field.

1. Introduction

High-frequency, eastward-propagating transient disturbances extend over much of the extratropics at any given time. In terms of their wavelength, period, phase speed, and vertical structure, they resemble the most unstable normal mode in analyses of baroclinic instability of the zonal-mean basic state, and, hence, it is appropriate to refer to them as baroclinic waves. The extensive theoretical literature on this subject, which dates back to the seminal works of Eady (1949) and Charney (1947), contains lively debates as to how the waves originate, what determines their wavelength and growth rate, and how they feed back upon the basic state. However, relatively few observational studies have been devoted to the structure, dynamics, and life cycle of the everyday "generic" baroclinic wave. The main emphasis in the synoptic meteorology literature has tended to be placed on intense (or rapidly intensifying) cyclones and their attendant fronts and bands of heavy precipitation, rather than on the wave trains in which they are embedded.

The present study is in the tradition of the works of Hartmann (1974), Hayashi and Golder (1977), Lau (1978, 1979), Blackmon et al. (1977, 1984a,b), Böttger and Fraedrich (1980), Trenberth (1981), Wallace et al. (1988), Randel (1988), and Buzzi and Tosi (1989), which have emphasized the structure and evolution of baroclinic waves, as revealed by time-series analysis of hemispheric gridded datasets extending over many years. The waves are isolated through the use of bandpass or highpass digital filters that emphasize fluctuations with periods shorter than a week (a technique pioneered by Klein 1951; Sawyer 1970; Blackmon 1976). As in some of the aforementioned more recent studies, we will make use of correlation analysis based on reference time series for selected grid points to define the properties of the waves. We will attempt to go further than previous studies in describing the vertical structure of the waves and in documenting the dynamical relationships that exist within the waves in the vertical velocity field, the ageostrophic motions, and the vorticity and heat balances.

2. Data and analysis techniques

The primary datasets examined in this study are twice-daily geopotential height, temperature, and wind fields at standard pressure levels derived from the NMC twice-daily final analyses for the Northern Hemisphere. The fields are gridded on a 2.5° longitude × 1° latitude grid, and the analyses are interpolated to 0.5° × 0.5° resolution for the purposes of this study.

* Present affiliation: Department of Atmospheric Sciences, Seoul National University, Seoul, Korea.

Corresponding author address: Professor J. M. Wallace, Dept. of Atmospheric Sciences, AK-40, University of Washington, Seattle, WA 98195.

© 1991 American Meteorological Society
Hemisphere winters 1975/76 through 1983/84 obtained from the NCAR data library. In this study winter is defined as three consecutive months, December through February. Six-hour forecast fields of vertical velocity at the 200-, 500-, and 850-mb levels, and outgoing longwave radiation (OLR) data for the same period, were also used. During this period there were a number of modifications and improvements in the numerical prediction model and the methods of data assimilation at NMC. For example, in September 1978 the Global Data Assimilation System was implemented and the Flattery Hough function analysis scheme was replaced by optimum interpolation (Bergman 1979), and in May 1980 a nonlinear normal mode initialization procedure was introduced (Dey and Morone 1985). A comprehensive discussion of these and other changes in operational procedures at NMC and their impacts upon the quality of the analyses is given by Trenberth and Olson (1988).

To assess the homogeneity of our dataset, hemispheric maps of standard deviations of various meteorological parameters were generated for the data subsets consisting of the first five winters (1975/76 through 1979/80) versus the remaining winters (1980/81 through 1983/84), and for odd numbered winters (1975/76, 1977/78 . . .) versus the even numbered winters (1976/77, 1978/79 . . .). Similar patterns were found in all four data subsets. Variances of 500-mb height in the stormtracks were found to be slightly larger during the later period, but no larger than for the subset based on the even numbered winters. Differences in the ageostrophic component of the wind field were more substantial: the patterns described in section 3c of this paper and in the companion paper (Lim et al. 1991) are dominated by the later period starting with the 1980/81 winter, after the introduction of the nonlinear normal-mode initialization scheme. For documentation, the reader is referred to Figs. 4.8 and 4.9 of Lim (1989), which show the 300- and 850-mb ageostrophic wind fields associated with baroclinic waves for each individual winter season.

Missing grids, which comprised about 3.5% of the dataset, were supplied by linear interpolation in the time domain. Only a few of the data gaps involved more than two consecutive grids; the largest one extended for a full week. From the interpolated time series of the primary fields (geopotential height, temperature, and wind), secondary fields such as 24-h tendencies, the ageostrophic component of wind, relative vorticity, and horizontal divergence were derived.

The OLR data used in this study cover the same period, apart from the 1978/79 winter, which was missing. The fields are asynchronous mosaics, derived from sequential passes of polar-orbiting satellites. They are defined on a latitude–longitude grid with 5° × 5° resolution. Daily values were obtained by averaging twice-daily observations at each grid point. Over the region of interest for the present study roughly 10% of the observations were missing at typical grid points in the region of interest. No attempt was made to interpolate in space or time across these gaps, and no time filtering was performed on the OLR data.

Highpass-filtered, normalized 500-mb height time series at selected grid points were used as reference time series to construct one-point regression maps of various meteorological parameters on the NMC octagonal grid and on the OLR latitude–longitude grid. The highpass filter, which has a half-power point near 6.5 days, is a 61-point optimal (minimax) finite-impulse response (FIR) filter. The weights were computed using the algorithm of Rabiner and Gold (1975). The frequency response curve is shown in appendix A of Lim (1989): the filter has a low-frequency cutoff at about the same frequency as the bandpass filter used by Blackmon (1976) but the cutoff is sharper. Wallace et al. (1988; section 7) have shown that the characteristics of baroclinic waves inferred from this type of analysis are not very sensitive to the details of the filter.

Our analysis approach is similar to that used in Blackmon et al. (1984a,b) and Wallace et al. (1988), but with two distinctions. First, instead of one-point correlation maps based on selected reference time series, we have constructed one-point regression maps that may be viewed as composite fields consisting of a linear combination of all available data, where each datum is weighted by the concurrent value of the reference time series. Second, only the reference time series is highpass filtered. Formally, the regression coefficient $b(i)$, where $i$ refers to a particular variable such as 500-mb height at a particular grid point, is given by

$$b(i) = \frac{1}{N} \sum_{t} y(t) x_{t}(i),$$

where $t$ refers to time, $x_{t}(i)$ is the highpass-filtered reference time series, $y(t)$ is the unfiltered time series of the $i$th variable, the highpass-filtered time series are denoted by primes ($'$), and $N$ is the number of observation times for which concurrent values of $x_{t}$ and $y(t)$ are available. Lim (1989) has demonstrated that (1) yields fields of regression coefficients that are almost identical to those obtained when both $x_{t}(i)$ and $y(t)$ are highpass filtered. By ignoring the distinction, we are, in effect, assuming that the low-frequency component of $y(t)$ removed by the highpass filter is temporally uncorrelated with $x_{t}$. The use of $y(t)$ as opposed to $y(t)'$ in (1) greatly reduces the amount of computation required and makes it possible to obtain regression coefficients from time series that contain many gaps such as our OLR time series. Since the reference time series have been normalized, the regression coefficients $b(i)$ carry the units of the regressed variables. They may be viewed as the perturbations in the variable in question at the $i$th grid point that are observed in association with a positive high-frequency perturbation in the reference variable $x_{t}'$, with an amplitude of one standard deviation.
With the exception of Fig. 5, the reference time series used in the analysis presented in this paper is the 500-mb height at the grid point near 41°N, 178°E in the most active region of the Pacific stormtrack. A selection of corresponding results for a grid point in the Atlantic stormtrack is presented in appendix B of Lim (1989).

The geostrophic wind statistics presented here are based on centered finite differences of the geopotential height field on the NMC octagonal grid, taking into account the latitudinal variation of the Coriolis parameter and the map factor. The computed geostrophic wind was rotated into polar coordinates and subtracted from the analyzed wind to calculate the ageostrophic wind component at each grid point on the octagonal grid.

3. Observed structure of baroclinic waves

a. Geopotential height

One-point regression maps for the 250- and 850-mb height fields are shown in Fig. 1. The patterns are qualitatively similar to those presented in Figs. 3 and 5 of Wallace et al. (1988). They are characterized by alternating areas of negative and positive regression coefficients that form a zonally oriented wavetrain. Such wavetrains are characteristic of the patterns derived from linear correlation analysis of the bandpass (2.5–6 day) or highpass (shorter than 5.5 day)-filtered geopotential heights (e.g., Blackmon et al. 1984a,b; Wallace et al. 1988). The magnitude of the regression coefficients at the centers of action decreases with distance from the base grid point somewhat more rapidly than on the corresponding correlation maps shown in Fig. 2. This difference reflects the fact that the reference grid point is located in the stormtrack, where the temporal variance of highpass-filtered geopotential is larger than in the surrounding regions. The westward tilt of the axes of the waves can be inferred from the relative position of the centers of action, and an increase in amplitude with height is also evident.

The orientation of the wavetrain at the two levels is somewhat different: at 850 mb it is almost perfectly zonal, whereas at 250 mb it exhibits an equatorward component (more distinct in the correlation map), consistent with the differences in the variance statistics discussed in Wallace et al. (1988). This distinction is consistent with the notion of equatorward Rossby-wave radiation out of the stormtracks at the upper levels, as discussed in Held and Hoskins (1985).

In order to document the vertical structure of the waves in more detail, a longitude–height cross section taken along the 40°N latitude circle, assembled from one-point regression maps at ten standard pressure levels, is shown in Fig. 3. The cross in the middle of the figure denotes the location of the reference time series. The westward tilt of the waves with height decreases as one moves across the section from west to east. The average phase difference between the upper troposphere and the surface (1000 mb) is about one-sixth of a cycle, which is in agreement with results of Lau (1979) and Hartmann (1974). Successive maxima in this cross section appear at progressively higher levels as one moves eastward, indicating upward energy dispersion. There is some suggestion of secondary amplitude maximum near the surface (i.e., the negative extremum near 160°E and
the weaker positive extremum just to the east of the date line). The westward tilt of the axes of the waves with increasing height is much more pronounced in the western half of the section than in the eastern half.

In order to examine the evolution of the waves, Fig. 4 shows lag-regression cross sections, based on the same reference time series as the simultaneous regression map shown in Fig. 3. Aside from the obvious eastward phase propagation, these sections reveal little concerning the evolution and life cycle of the waves that is not evident in the simultaneous correlations (i.e., the middle panel) alone.

We observe similar east–west contrasts in longitude–height sections based on reference grid points well upstream and downstream of the stormtracks (not shown), and Randel (1988) obtained similar vertical structures in composite sections for all grid points on the 51°N and 51°S latitude circles. Hence, the evolution from a more baroclinic structure at the western end of the sections to a more barotropic structure at the eastern end has little to do with the precise location of the reference grid point. It seems more likely that it is a reflection of the evolution of individual waves as they propagate eastward. For the sake of argument, let us imagine an idealized climatology, independent of longitude, in which all grid points on a latitude circle experience the same population of waves with regard to amplitude and “level of maturity” in their life cycle. Sometimes a particular reference grid point on the latitude circle lies in a “stormtrack”; sometimes it is upstream or downstream of a stormtrack; and sometimes it is not anywhere close to a stormtrack. The regression statistics tend to be dominated by the periods when the waves at the reference grid point are relatively strong, i.e., when the stormtrack is near it. At these times, the waves upstream of the reference grid point will be growing and the waves downstream of it will be decaying, which is consistent with the structures that we observe in the cross sections.

The choice of reference grid point does have a modest influence on the appearance of the sections. Waves observed at grid points near the upstream end of the stormtrack exhibit somewhat larger vertical tilts than those observed at the downstream end, as shown in Fig. 5. However, the differences are rather subtle in comparison to the east–west contrast in the vertical tilt in individual sections. Likewise, the choice of reference variable (e.g., 1000-mb height, 500-mb height, or 100-mb temperature) has relatively little impact on the appearance of the sections (Lim 1989).

b. Temperature

Figure 6 shows regression coefficients for the 850-mb temperature field based on the same reference time series as was used in constructing Figs. 1–4. The overall pattern is similar to that of the corresponding geopotential height-regression map in Fig. 1, but the waves extend somewhat farther into the subtropics than in the 850-mb height field, and the larger amplitudes are
Fig. 5. As in Fig. 3, but based on highpass filtered, normalized 500-mb height at the reference grid points (a) (35°N, 130°E) and (b) (42°N, 140°W). The approximate position of the base grid point is marked with a cross on each section. Negative contours are dashed. Contour interval: 10 m.

located somewhat farther to the west, consistent with the larger vertical tilt of the waves observed in the vicinity of the upstream grid point in Fig. 5. A longitude–height cross section along 40°N latitude, assembled from the temperature regression maps at all pressure levels, is shown in Fig. 7. Consistent with results based on bandpass-filtered data by Lau (1978) and Lau et al. (1981), the largest temperature perturbations in the waves are observed near the 700-mb and 200-mb levels. The eastward tilt of the axes of the waves with height

Fig. 6. The 850-mb temperature regression map, for the same reference time series as in Fig. 1. Negative contours are dashed and the zero contour is omitted. Contour interval: 0.2°C.

Fig. 7. Temperature regression cross section along 40°N for the same reference time series as in Fig. 1. Contour interval: 0.2°C.

in the tropospheric temperature field, which is evident at all longitudes in the section, is consistent with the structure simulated in the numerical integrations of Hoskins and Heckley (1981), who emphasized its role in the formation of fronts. Upward energy dispersion is clearly evident in the lower stratosphere to the west of the date line. Coherent temperature perturbations are detectable as high as 50 mb, but the amplitude at that level is much smaller than at 100 mb.

c. Wind

Regression maps of the $u$- and $v$-components of the total wind field at the 300-mb level for the same reference time series are shown in Fig. 8. The wind perturbations associated with baroclinic waves are on the
order of 5 m s\(^{-1}\) for the zonal component and 8 m s\(^{-1}\) for the meridional component. The meridionally oriented dipole structure in the zonal wind field, together with the simple wavelike structure of the meridional wind field, is indicative of a series of vortices in the perturbation wind field, elongated in the meridional direction. At the 850-mb level (not shown) the wind perturbations are roughly half as large as those at 300 mb.

Corresponding results for the ageostrophic wind component (based on a latitudinally varying Coriolis parameter) are shown in Fig. 9. In these maps a few of the contours from the corresponding geopotential height regression fields have been superimposed on the wind field for reference. The solid contour encircles a high in the geopotential height field and the dashed contour encircles a low. The maximum ageostrophic wind speeds at the 300-mb level are on the order of 3–4 m s\(^{-1}\), which corresponds to approximately half the geostrophic wind speed associated with the waves (but only one-tenth of the total wind speed, including the basic-state flow) at the upper level. At the 850-mb level it is slightly less than 1 m s\(^{-1}\), which corresponds to roughly one-fifth of the corresponding geostrophic wind perturbations at that level. The ageostrophic wind vectors are oriented zonally at the 300-mb level and meridionally at the 850-mb level. These features are apparent in the ageostrophic regression maps for individual winters from 1980 (the time of the introduction of the nonlinear normal-mode initialization scheme in the NMC analyses) onward, and there is a suggestion of them in the earlier winters (Lim 1989). The ageostrophic flow at the 700-mb level, which is close to the steering level for baroclinic waves, appears to be a hybrid of the 300-mb and 850-mb patterns and is relatively weak (Lim et al. 1991). The longitude–height cross section of the zonal component of the ageostrophic wind field, shown in Fig. 10, is at variance with the notion that the ageostrophic flow in baroclinic waves is dominated by a simple closed circulation cell in the zonal plane, as implied in many schematic diagrams that have appeared in the literature. The implications and theoretical interpretation of these results are discussed in Lim et al. (1991).

d. Relative vorticity and the horizontal divergence

Figure 11 shows the regression field of relative vorticity at the 300-mb level. The structure is similar to that of 250-mb geopotential height regressions shown in Fig. 1 except for the sign reversal. However, to the west of the base grid point the features in the vorticity field appear to be somewhat more isotropic and to the east of the base grid point they show an evidence of dispersion of the waves into the subtropics. Mote et al. (1991) have shown that the high-frequency fluctuations exhibit a very similar signature in vertically integrated ozone as sensed by the total ozone mapping spectrometer (TOMS) instrument.
The corresponding regression fields for horizontal divergence at the 250- and 850-mb levels are shown in Fig. 12. The divergence patterns at the two levels exhibit the expected out-of-phase relationship, and the corresponding centers are comparable in magnitude. At the 250-mb level, there is a stronger suggestion of an equatorward propagation of wave activity, as evidenced by the northwest–southeast orientation of the wavetrain. The corresponding zonal cross section shown in Fig. 13 was slightly smoothed in the pressure domain before contouring. The node in the vertical profile is located near the 500-mb level. In the upper troposphere the maximum divergence is located about one-eighth of a wavelength to the west of the upper-level ridge. Hence, the divergent motions are giving rise to kinetic energy generation. In the lower troposphere the wave in the divergence field is almost in quadrature with the wave in the geopotential height field, with divergence occurring to the east (i.e., in advance) of the ridge and vice versa. The perturbations in the divergence field at the jetstream level are about one-tenth as large as the corresponding perturbations in the vorticity field.

In view of the much stronger ageostrophic flows at the jetstream level than in the lower troposphere as documented in the previous subsection, one might question how the divergence in the lower troposphere can be comparable to that at the higher levels. Lim et al. (1991) show that at the upper levels the divergence associated with the geostrophic wind field due to the \( \beta \) effect partially cancels the divergence associated with the ageostrophic wind field (mainly the subgeostrophic flow in the troughs and supergeostrophic flow in the ridges), while in the lower troposphere, the geostrophic and ageostrophic components of the divergence field tend to reinforce one another.

e. Vertical velocity and OLR

Figure 14 shows one-point regression maps of the 200-, 500-, and 850-mb vertical velocities (in pressure
coordinates) based on the NMC 6-h forecasts. The various features in the pattern are nearly in phase at the three levels. Peak vertical velocities on the order of 100 mb per day (2 cm s⁻¹) are observed at the 500-mb level. At all three levels the waves in the vertical velocity field are almost in quadrature with those in the geopotential height field, with ascent to the east of the negative height perturbations. At the 850-mb level the waves in the temperature and vertical velocity fields are nearly in phase (compare Fig. 6 with 14c), which is indicative of a thermally direct circulation, while at the 200-mb level they are nearly in quadrature (compare Fig. 7 with 14a). The corresponding OLR regression field, shown in Fig. 15, agrees well with the vertical velocity pattern: ascent (negative ω in pressure coordinates) corresponds to enhanced cloudiness (negative OLR perturbations). In view of the uncertainties in the timing of the OLR observations, we do not attach any importance to the small eastward displacement of the wave in the OLR field relative to the wave in the vertical velocity field.

f. Synthesis

The observed structure of baroclinic waves as inferred from the linear regression analysis agrees well with our expectations based on previous observational studies (e.g., Fleagle 1947; Palmén and Newton 1969), linear theory (Charney 1947; Eady 1949), and numerical simulations incorporating nonlinear effects (Simmons and Hoskins 1978; Hoskins and Heckley 1981), except for the dominance of the meridional component of the ageostrophic flow in the lower troposphere, which is discussed in more detail by Lim et al. (1991). The amplitudes of the waves in the various meteorological parameters are summarized in Table 1. They tend to be much weaker than those based on composite studies that have appeared in the literature because our statistical treatment of the data takes into account all high-frequency fluctuations, regardless of their intensity.

Figure 16 shows a summary of the phase relationships in baroclinic waves as deduced from our regression analysis. The distinctions between the growing, mature, and decaying phases are based on comparisons between the structures observed upstream and downstream of the reference grid point. It is perhaps worth noting that in none of our sections did we see a structure that strongly resembles the linear stage of the life cycle of baroclinic waves, with significant amplitudes confined to the lower troposphere.

4. Further analysis of the temperature and vorticity fields

In this section we will consider the adiabatic, inviscid processes that contribute to changes in temperature and vorticity, using the thermodynamic energy and vorticity equations as they appear in the conventional quasi-geostrophic formulation. The terms in these equations are estimated on the basis of regression analysis, using the same reference time series as in the previous section.

a. Diagnosis of the thermodynamic energy equation

We may write the thermodynamic energy equation without the diabatic heating term in the form

$$\frac{\partial T}{\partial t} = -V_g \cdot \nabla T + \frac{p \sigma}{R} \omega,$$  \hspace{1cm} (2)

where $T$ is the temperature, $R$ the gas constant for dry air, $\omega$ the vertical velocity, $V_g$ the geostrophic wind, and $\sigma$ is the static stability defined as $-\partial \ln \rho / \partial p$ in pressure coordinates. Here $\alpha$ is the specific volume of dry air. The values of the static stability parameter used in this study are based on the climatological mean temperatures at the standard pressure levels for the nine winters. They were averaged on pressure surfaces over the entire data grid. They are comparable at all pressure

![Fig. 15. Regression map of the daily OLR field based on the same reference time series as in Fig. 1. Contour interval: 2 W m⁻²; negative contours are dashed and the zero contour is omitted.](image-url)
levels to the values computed from upper-air observations over the United States for the winter season by Gates (1961).

Figures 17 and 18 show regression maps of (a) the observed 24-h temperature tendency, as evaluated by a centered difference, (c) the horizontal advection term, and (d) the vertical velocity term in Eq. (2) for the 850- and 200-mb levels, respectively. Panel (b) in these figures is the sum of (c) plus (d). All of the terms are expressed in units of $10^{-3}$ °C s$^{-1}$, which corresponds to roughly 1°C per day. The 850-mb temperature advection (c) is in qualitative agreement with Figs. 1b and 6; the strongest warm advection, with peak values of 5°C per day, is occurring near and slightly to the east of the region of strongest southerly flow. Adiabatic cooling is taking place in the regions of southerly geostrophic flow in Fig. 1b and vice versa. Hence, the temperature changes forced by the advection term are partially offset by the adiabatic temperature change induced by the vertical motion. The fact that the computed temperature changes in Fig. 17b are about 50% larger than the observed changes at all three major centers of action in Fig. 17a suggests that the compensation is not strong enough. Since the estimates of the

![Fig. 16](image)

Fig. 16. Schematic illustrations for the different stages of the life cycle of baroclinic waves. The italicized characters C and W refer to the cold and warm axes of temperature waves, and T and R refer to the trough and ridge lines of the geopotential height waves, respectively. The secondary circulations in the waves are indicated by arrows. The associated vertical velocity is represented by arrowheads at the 200-mb and 500-mb levels, as viewed from the upper right corner of each panel.

![Fig. 17](image)

Fig. 17. Regression maps of (a) the local time rate of change of temperature based on 24-h changes; (b) the sum of panels (c) plus (d); (c) the geostrophic temperature advection; and (d) the vertical velocity term in the simplified thermodynamic energy equation (2), as applied to the 850-mb level. The static stability parameter $a$ was computed from the climatological mean temperatures at the standard pressure levels for the nine winters and averaged over all grid points. The reference time series is the same as in Fig. 1. Contour interval: $1.0 \times 10^{-3}$ °C s$^{-1}$; negative contours are dashed and the zero contour is omitted.
time derivative and geostrophic temperature advection term should be quite reliable, and since we have no reason to expect that the geostrophic temperature advection should be systematically stronger than the advection produced by the actual winds, we suspect that the vertical velocities at this level in the NMC model may have been too weak by \( \sim 30\% \). If diabatic heating had been included in these calculations it would have tended to be in phase with the advection term, making the discrepancy somewhat larger.

At the 200-mb level (Fig. 18), the temperature advection and vertical velocity terms (c) and (d) nearly cancel, leaving a weak, disorganized residual pattern (b), which bears little relation to the observed temperature changes (a). At this level the basic-state zonal flow is \( \sim 40 \text{ m s}^{-1} \) over the reference grid point (Lau et al. 1981), almost four times as large as the phase velocity of the waves, and the basic-state meridional temperature gradient is very weak. Troughs in the geopotential height field tend to be warm and ridges tend to be cold. Therefore, air parcels must warm by adiabatic compression as they are swept eastward from ridge to trough and cool by adiabatic expansion as they go from trough to ridge. From a comparison of Figs. 1a and 18d it is evident that this is, in fact, the case. Since the phase speed is eastward, the temperature advection term must be larger than the vertical motion term. Comparison of panels (a) and (b) in Fig. 18 suggests that the advection term is in fact a bit larger, but not by a wide enough margin. Hence, either the NMC analyses are underestimating the basic-state zonal wind speed in the region of the reference grid point, or the vertical velocities in the lower stratosphere may have been too strong by \( \sim 20\% \).

b. Diagnosis of the vorticity equation

The rate change of relative vorticity can be approximated by

\[
\frac{\partial \xi}{\partial t} = -\nabla \cdot (\xi \mathbf{f}) - f \nabla \cdot \mathbf{V},
\]

(3)

where \( \xi \) is the geostrophic relative vorticity, and \( f \) is the Coriolis parameter.

Figure 19 shows regression maps for the local time rate of change of vorticity, the advection of absolute vorticity,\(^2\) and the convergence term in (3), evaluated at the 300-mb level. The panels are arranged as in the two previous figures. It is evident from comparing panels (c) and (d) that the horizontal vorticity advection term is much larger than the horizontal divergence term at this level. The sum of these two terms shown in (b) is about twice as large as the observed vorticity tendency shown in (a). Hence, either the basic-state zonal flow is much too strong, producing excessively strong vorticity advection (which seems highly unlikely), or the wave component of the divergence field at this level is too weak, by more than a factor of two. The discrepancy in our calculations is larger than in the case of the heat balance discussed in the previous subsection.

Figure 20 shows longitude–height regression cross sections along 40°N for the two terms on the right-hand side of (3). The advection is almost in phase at all levels, while the divergence term exhibits a node near the 500-mb level. In the upper troposphere the

---

\(^2\) The advection of relative and planetary vorticity were computed separately, and the former was found to be about four times as large as the latter.
5. Diagnosis of the vertical velocity field based on the quasi-geostrophic formulation

The results presented in the previous section indicate that the NMC operational analyses for the record period considered in this study seriously underestimate the divergent component of the wind. We suspect that this problem was most serious during the early years of the record, while the Hough analysis was still in use, but that it has probably persisted to some degree at least into the early 1980s. The vertical velocity field should not be as seriously impacted because it is based, not upon the NMC analyses, but upon the 6-h forecast fields. Hence, it should be possible to test the dynamical consistency of the analysis scheme used in this paper by comparing the vertical velocity fields derived from the model with the quasi-geostrophic vertical velocity field derived from the three-dimensional distribution of temperature and vorticity advection by the geostrophic wind.

The quasi-geostrophic omega equation may be written in the conventional form (e.g., see Holton 1979)

\[
\left( \nabla^2 + \frac{f^2}{\sigma} \frac{\partial^2}{\partial p^2} \right) \omega
\]

\[
= \frac{\sigma}{\partial p} \left[ V_g \cdot \nabla (f_g + f) \right] + \frac{\sigma}{\partial p} \mathcal{V}^2 (V_g \cdot \nabla T), \quad (4)
\]

Fig. 19. Regression maps of (a) the local time rate of change of relative vorticity; (b) the sum of panels (c) plus (d); (c) the advection of absolute vorticity by the geostrophic wind; and (d) the divergence term in the simplified vorticity equation. The vorticity tendencies are based on 24-h changes of vorticity (3), as applied to the 300-mb level, based on the same reference time series as in Fig. 1. Contour interval: $1.0 \times 10^{-16}$ s$^{-2}$; negative contours are dashed and the zero contour is omitted.

divergence opposes the vorticity advection, while in the lower troposphere it contributes in the same sense to the observed tendencies of relative vorticity. The divergence term serves to keep the waves in the vorticity field from being torn apart by the vertical shear in the basic state: it retards the eastward propagation of the waves in the upper troposphere and enhances it in the lower troposphere.

Fig. 20. Vertical cross sections of (a) the vorticity advection term and (b) the divergence term in (3) along 40°N, based on the same reference time series as Fig. 1. The position of the reference time series is marked with a cross. Contour interval: $1.0 \times 10^{-16}$ s$^{-2}$. 
Fig. 21. Regression maps of (a) the elliptic function of the omega velocity field on the left-hand side of (4), estimated from the regression maps of vertical velocity shown in Fig. 14, and (b) the net forcing of vertical velocity as defined by the sum of the two terms on the right-hand side of (4), evaluated at the 500-mb level, based on 850-, 500-, and 200-mb regressions. Contour interval: $1.0 \times 10^{-15}$ mb m$^{-2}$ s$^{-1}$; negative contours are dashed and the zero contour is omitted.

where we have replaced $-\partial \Phi / \partial p$ by $RT / p$. Figure 21 shows the sum of terms on the left-hand side of (4) at the 500-mb level at the reference grid point, as estimated from the NMC 6-h forecast $\omega$ fields for the 200-, 500-, and 850-mb levels, together with the sum of the two forcing terms on the right-hand side. This agrees remarkably well, considering the crudeness of our treatment of the vertical structure.

Figure 22 shows regression maps for the two terms on the left-hand side of (4). The two distributions are similar but the horizontal Laplacian term is somewhat larger. Figure 23 shows the regression maps of the two forcing terms on the right-hand side of (4). The temperature advection term (a) is almost twice as large as the term associated with the vertical derivative of the vorticity advection (b). The axis of the wavetrain in (a) exhibits a northwest–southeast orientation, whereas in the wavetrain in (b) it has a west to east orientation.

In order to assess the extent of the cancellation or reinforcement between the two forcing terms, we have shown in Fig. 24 a simplified version of the fields in the previous figure, superimposed upon the corresponding 500-mb geopotential height regression map. Along 40°N, the forcing of upward motion by the differential vorticity advection (the thin solid loops) is located roughly 90° to the west of the ridges of the baroclinic waves in the geopotential height field (denoted by the fine shading), while the thermal forcing (denoted by the thick solid loop) favors upward motion coincident with the ridge itself. Along 30°N both terms favor upward motion in the southerly flow ahead of the troughs. At no latitude do we observe what appears to be a strong compensation between the two terms.

Fig. 22. Regression maps of (a) the horizontal Laplacian of the vertical velocity and (b) the $\partial^2 \omega / \partial p^2$ term on the left-hand side of (4), evaluated at the 500-mb level. Negative contours are dashed and zero contours omitted. Contour interval: $1.0 \times 10^{-15}$ mb m$^{-2}$ s$^{-1}$.

Fig. 23. Regression maps of the forcing of vertical velocity as deduced from (4): (a) the contribution from the Laplacian of the temperature advection and (b) the contribution from the vertical derivative of the vorticity advection. Contour interval: $1.0 \times 10^{-15}$ mb m$^{-2}$ s$^{-1}$; negative contours are dashed and the zero contour is omitted.
Following Hoskins et al. (1978), the forcing terms of vertical velocity in the quasi-geostrophic formulation may be expressed in the alternative form

\[
\left( \nabla^2 + \frac{f^2}{\sigma} \frac{\partial^2}{\partial \sigma^2} \right) \omega = \frac{2}{\sigma} \nabla \cdot \mathbf{Q} + \frac{f \beta}{\sigma} v_p, \tag{5}
\]

where the so-called "\( Q \) vector" is defined as

\[
\mathbf{Q} = -\left[ (u_x \alpha_x + v_x \alpha_y), (u_y \alpha_x + v_y \alpha_y) \right], \tag{6}
\]

where the subscripts refer to spatial derivatives, \( \alpha \) is specific volume, and all the terms that involve wind refer to geostrophic wind. In this derivation, the variation of \( f \) with latitude is taken into account, but the extra terms arising from the variation of \( f \) with latitude are not included on the right-hand side in (5) because of their negligible contribution to the forcing of vertical velocity.

Figure 25 shows the \( Q \)-vector field derived from our regression scheme for the 500-mb level at the reference grid point. In the upper panel, contours of the net forcing \([i.e., \text{the sum of the two terms on the right-hand side of (5)}]\) are superimposed on the \( Q \) vectors. Since the first term is almost an order of magnitude larger than the second, there is a strong correspondence between the forcing of upward motion and the divergence of the \( Q \) vectors. In the lower panel, several 500-mb height contours have been superimposed on the \( Q \) vectors in order to show the relation between the \( Q \) field and the ridges and troughs in the geopotential height field. The pattern is dominated by westward (eastward) oriented \( Q \) vectors in the ridges (troughs) of the waves in the geopotential field, which reflect the intensification of the east-west thermal contrasts due to the advection of the basic-state temperature field by the meridional wind component in the waves. The northward (southward) directed \( Q \) vectors to the east of the lows (highs) reflect changes in the meridional gradients induced by differential advection of the disturbances in the temperature field due to the meridional shear in the basic-state flow \( (u_x \alpha_x') \) and the diffusivity of the meridional wind component in the waves acting upon the basic-state temperature gradient \( (v_x \alpha_x') \).

From a comparison of Figs. 21b and 25a it is evident that the results based on \( Q \)-vector formulation bear some resemblance to those based on the conventional omega equation, though the two patterns are not as similar as we had expected. In general, the \( Q \)-vector formulation yields a stronger forcing along 40°N, which falls off more rapidly toward the subtropics. From a comparison of the same two figures with Fig. 21a, it is apparent that the conventional and \( Q \)-vector formulations replicate the elliptic function of vertical velocity on the left-hand side of (4) and (5) with comparable fidelity. Since the total forcing fields derived from the terms on the right-hand sides of (4) and (5) should be identical, the differences between the patterns in Figs. 21b and 25a must be a reflection of inaccuracies in the finite-difference algorithms used in computing the terms. In view of the apparent sensitivity of the shape of the forcing to the method of computation, the schematic Fig. 24 cannot be regarded as definitive.

6. Concluding remarks

The methodology used in the present study combines elements of what might be called the "general circulation approach," characterized by statistical analysis...
of comprehensive, hemispheric (or global) datasets extending over many seasons and a more traditional synoptic analysis that aims at the description and diagnosis of discrete dynamical phenomena that appear on weather maps. It goes further than previous statistical studies in documenting the essential characteristics of the “generic” baroclinic wave, but it falls short of a fully satisfying description in at least three respects.

(i) The diagnostic calculations performed in the previous two sections are not as accurate as we had hoped. A likely source of error is the suppression and distortion of the secondary circulations in the waves by the gradually improving data assimilation schemes that have been in use. If the same analysis could be performed on a complete 10-winter dataset reanalyzed with a current state-of-the-art data assimilation system that includes diabatic initialization, the agreement in the various budgets should be considerably better.

(ii) As demonstrated by the studies of Mullen (1987), Lau (1988), and others, baroclinic waves assume a variety of tracks and configurations in response to the changing background flow; for example, they may be interrupted during periods of blocking; they might propagate northeastward during periods when transient longwave troughs are situated to the west of the reference grid point, etc. Only by resorting to some type of conditional sampling could one hope to recover the full sharpness inherent in the baroclinic wavetrains observed on synoptic maps.

(iii) The analysis scheme that we have employed in this study is effective for defining the linear properties of baroclinic waves, but it cannot resolve significant nonlinear features such as fronts and tropopause folds, circulations associated with comma clouds and tight vorticity maxima, and secondary cyclone developments in the polar airstream. Information on such features is inherent in the higher-order statistical moments. For example, Fig. 26 shows the hemispheric distribution of the moment coefficient of skewness of the 24-h tendency of 850-mb temperature: i.e.,

\[ m_3 = \frac{(\Delta T)^3}{(\Delta T)^2} \]

where \( \Delta T \) represents the 24-h 850-mb temperature difference computed by subtracting the temperature on the current day from the temperature on the next day, an overbar (\( \bar{\cdot} \)) denotes a mean over all days in the dataset, and a prime (\( \prime \)) denotes a deviation from the mean over all days. Fields such as highpass-filtered 500-mb height and sea level pressure and their 24-h tendencies do not exhibit strong skewness. The field that we have displayed in Fig. 26 is of much more interest. The prevalence of negative skewness (i.e., the absence of dashed contours except over the Tibetan plateau) confirms the well-known fact that cold frontal zones (in which local temperatures are decreasing) tend to be more concentrated than warm frontal zones. Note that the strong negative skewness tends to be concentrated within certain geographical regions such as the southeastern United States. If such information could be incorporated into one-point regression analyses, it might be possible to reconstruct more realistic looking baroclinic wave fields than would be possible in a strictly linear analysis such as the one that we have performed.

Despite these limitations, this analysis technique has yielded dynamically plausible baroclinic wave signatures in all the fields that we have examined, including ageostrophic wind components, divergence, model-generated and quasi-geostrophic vertical velocity, \( Q \) vector, OLR, total column ozone (see Mote et al. 1991), and stratospheric temperatures at least as high as the 50-mb level. Our results, together with those of Randel (1988), suggest that the notion of a “life cycle” of baroclinic waves characterized by baroclinic growth, barotropic decay, and upward and equatorward dispersion of wave activity is applicable, not only in the region of the climatological mean stormtracks, but wherever baroclinic waves are regularly observed.

Acknowledgments. We would like to thank James R. Holton for his encouragement and helpful suggestions. This work was supported by the Climate Dynamics Program Office of the National Science Foundation under Grants ATM 8318853 and ATM 8822872.

REFERENCES


