Dynamical processes in the atmosphere and the use of models

By B. J. HOSKINS

Department of Meteorology, University of Reading

[Symons Memorial Lecture, delivered 19 May 1982]

1. INTRODUCTION

The pressure for the production of a forecast and the advent of the electronic computer have had a tremendous influence which has generally been for the benefit of meteorology. However, they have also conspired to push the organization of research into the situation depicted in Fig. 1. Observations of the atmosphere and output from sophisticated numerical models are diagnosed and conveyed to others using conceptual models that have not changed in decades. Theoretical modelling frequently proceeds with little contact with data from the real atmosphere or with the complex numerical model and contributes little to diagnosis. An example of the static conceptual model is provided by the Norwegian frontal model. Sixty years and many thousands of baroclinic instability papers later this is still the manner in which atmospheric data and the output from forecast models are conveyed not only to the public but also between scientists.

That such a situation is not unique to the computer age is highlighted in the fascinating historical account by Bergeron (1959), one of the great men of the Bergen School.

![Diagram showing the relationship between observations, static conceptual models, complex numerical models, and theoretical models.]

Figure 1. A schematic illustration of the unhealthy situation which sometimes exists in meteorological research.
He also produced a challenge to computer modellers: 'Moreover, the computers have not yet disclosed any new specialized (structural) models to us'.

The purpose of this Lecture is to stress the potential advantages of the system described in Fig. 2. Here, the sophisticated numerical model is at one end of a spectrum of dynamical models of different complexity. Some of these might be laboratory models;

![Diagram](image)

Figure 2. A schematic illustration of the optimum situation for meteorological research.

some will be amenable to analytic techniques; others will require electronic computers for their study. However, all levels of modelling interact well with each other and with observations of the atmosphere to produce an evolving conceptual background. In turn this is used to aid in diagnosis of atmospheric and complex model behaviour.

Two different meteorological subjects will be discussed with these ideas in mind, stressing the desirability of application of a hierarchy of models. The first topic, teleconnections, is one in which the situation envisaged in Fig. 2 is now close to being achieved. The second topic is that of mid-latitude synoptic systems where the current situation is further from the ideal but the opportunity now exists for real progress.

2. ATMOSPHERIC TELECONNECTIONS

(a) Basic β-plane theory

The word 'teleconnection' means connection at a distance. Thus one may say that the dry subtropics situated in the descending arm of the Hadley cell have a teleconnection with the Intertropical Convergence Zone with its convection and rising arm of the Hadley cell. However, the word is usually reserved for the sort of phenomena more associated with the rotational portion of the flow and the possibility of steady wave systems.

When we wish to produce theory and a hierarchy of models for such waves we have to proceed backwards from today's primitive equation models on the sphere to the earlier quasi-geostrophic β-plane models and to the barotropic vorticity equation, β-plane channel models. Historically the barotropic vorticity equation models were superseded as the best forecast models because of phase speed problems and the non-development of new systems. In their turn, quasi-geostrophic models were not applicable to the tropics and the inclusion of parametrizations of physical processes in them was difficult, so that they too were replaced. However, despite their limitations, these simpler models can be useful for
elucidating fundamental dynamical processes at work in the atmosphere and in more complicated models.

The basic theory of wave motion in the barotropic $\beta$-plane model is due to Rossby (1939, 1945). A thorough review of Rossby wave theory is given by Platzman (1968). For our purposes we note that the equation for the streamfunction, $\psi$, of a steady perturbation to a uniform basic flow $\bar{u}$ may be written:

$$\bar{u} \frac{\partial}{\partial x} \nabla^2 \psi + \beta \frac{\partial \psi}{\partial x} = 0,$$

(1)

or

$$\frac{\partial}{\partial x} \left( \nabla^2 \psi + \frac{\beta}{\bar{u}} \psi \right) = 0.$$

(2)

Thus wave-like solutions of the form $\cos(kx + \lambda y)$ are possible provided that the total wavenumber, $K$, is given by

$$K^2 = k^2 + l^2 = \beta/\bar{u} \equiv K_s^2.$$

(3)

Rossby and his collaborators concentrated on the 'channel' problem with $l$ fixed and usually zero. Rossby (1945) and Yeh (1949) produced solutions for a source at $x = 0$ switched on at $t = 0$, or almost equivalently for a steady source at $x = 0$ in an atmosphere with wave damping (see Fig. 3). In both cases, the wavelength of the stationary train produced is $2\pi/K_s$. Although the phase speed of the pattern is zero relative to the ground,

$$\left\langle \begin{array}{c} \bar{u}T \cr \end{array} \right\rangle \quad \left\langle \begin{array}{c} \text{K} \cr \end{array} \right\rangle$$

Figure 3. A schematic illustration of the steady vorticity or height-field waves downstream from a source region at $x = 0$ on a $\beta$-plane channel in an atmosphere with a westerly flow $\bar{u}$ and wave damping on a time-scale $T$. The wavelength is $2\pi/K_s$ and the $x$-decay scale for the amplitude is $2\bar{u}T$.

... information proceeds eastwards with the group velocity $c_g (= \partial(\text{frequency})/\partial k)$ which for stationary waves is twice the basic velocity $\bar{u}$. Thus, in the initial value problem, the steady wave pattern is set up at time $T$ for $x$ less than $2\bar{u}T$. In the steady-state damped problem the pattern exists for $x$ less than $2\bar{u}T$ where $T$ is now the timescale of the damping. This is consistent with the conceptual model widely used by long-range forecasters that tele-connections proceed from west to east.

On an infinite $\beta$-plane stationary wave-crests oriented in any direction are possible, but the separation of these crests must still be $2\pi/K_s$ (see Fig. 4(a)). For stationary waves, the group velocity relative to the ground is perpendicular to the wave-crests with a component in the eastward direction (Longuet-Higgins 1964). Its magnitude is now $2\bar{u}\cos \alpha$ where $\alpha$ is the angle subtended with the $x$ axis. As highlighted by Lighthill (1966), from a point source stationary wave information may travel in all eastward directions but the distance of travel is proportional to the cosine of the angle subtended with the flow. Thus, as shown in Fig. 4(b), the stationary wave-train with circular crests separated by a distance $2\pi/K_s$ fills a circle of radius $\bar{u}T$ centred at a distance $\bar{u}T$ eastward of the source. A source of finite size which is elongated zonally has its largest Fourier components with $l$ larger than $k$, so that itfavours wave-crests quite zonally oriented and a large meridional...
component in the group velocity. Thus the amplitudes along the axis in Fig. 4(b) would be small and one would see almost separate wave-trains to the north-east and south-east of the source. Finally, we note that any point is influenced by stationary sources in a circle of radius \( \bar{u}T \) centred a distance \( \bar{u}T \) to the west. Thus the long-range forecaster on an infinite \( \beta \)-plane should look in such a region!

(b) Rossby–Haurwitz waves

The spherical equivalent of Rossby's \( \beta \)-plane waves was produced by Haurwitz (1940). For stationary perturbations to a uniform relative angular velocity (super-rotation) zonal flow \( \bar{u} = a\omega \cos \phi \), the equivalent to Eq. (1) is

\[
a\omega \frac{\partial}{\partial \lambda} \nabla^2 \psi + \frac{2(\Omega + \omega)}{a} \frac{\partial \psi}{\partial \lambda} = 0, \tag{4}
\]

where \( a \) and \( \Omega \) are the radius and rotation rate of the planet, \( \lambda \) is longitude and \( \phi \) latitude. Proceeding as before, Eq. (4) may be rewritten

\[
\frac{\partial}{\partial \lambda} \left( \nabla^2 \psi + \frac{2}{a^2} \frac{\Omega + \omega}{\omega} \psi \right) = 0. \tag{5}
\]

The eigenfunctions of the Laplacian operator on the sphere are the spherical harmonics \( Y_n^m \) (\( m \) and \( n \) normally integers, \( n \geq |m| \)) for which \( \nabla^2 Y_n^m = -n(n+1)a^{-2}Y_n^m \). Thus spherical-harmonic-type solutions of Eq. (5) are possible provided that the ‘total wave-number’ \( n \), is given by

\[
n(n + 1) = 2(1 + \Omega/\omega) \equiv n_s(n_s + 1). \tag{6}
\]

A complete description of possible stationary wave fields is given by a sum of such modes:
\[ \psi = \sum_{|m| \leq n} \psi^m Y_n^m(\lambda, \phi), \tag{7} \]

where \( n = n_z \).

It is worth pausing to consider the structure of these spherical harmonic modes \( Y_n^m \). \( m \) is the zonal wavenumber and \((n - |m|)\) the number of zeros between the north and south poles. They are symmetric or antisymmetric about the equator according as \((n - |m|)\) is even or odd. For \( n_z = 8 \), the fundamental mode \((Y_8^0)\) and the first symmetric mode \((Y_8^8)\) are shown in Fig. 5. These modes are stationary for a super-rotation zonal flow with angular velocity \( \Omega/35 \) and an earth equatorial speed of about 13-3 m s\(^{-1}\).

![Figure 5](image)

(c) More complete theory for the spherical domain

The theory so far described is of little direct relevance to the observer of the atmosphere or user of a general circulation model. However, in the last few years the gap between theory and its use has been closed to a large extent by filling in the missing elements in the hierarchy of dynamical models (Fig. 2). It may be noticed that the Haurwitz theory does not answer the question of what will happen if one places a stationary vorticity source in a spherical atmosphere which is rotating with constant angular velocity, i.e. we have no equivalent of Figs. 3 and 4(b). However, such an experiment is easily performed analytically, considering each spherical harmonic component separately. Figure 6 shows the vorticity perturbation generated by the switch-on of a circular source at 30°N in an atmosphere with angular rotation which would correspond to \( n_z = 7.5 \) and has an earth equatorial velocity of about 15-0 m s\(^{-1}\). A downstream pattern develops which is much as expected from the infinite \( \beta \)-plane studies except that there is a bias in the wave-train towards the equatorial side and towards the antipodean point of the source. This behaviour is confirmed also by the damped, steady-state perturbation forced by a mountain in this same flow illustrated in Fig. 3 of Grose and Hoskins (1979) where it is also seen that there is a tendency for the wave-train to propagate from the antipodean point eastwards and northwards back to the source region.

Inspired by the \( \beta \)-plane analysis we may expect meridional propagation to be enhanced by elongating the source zonally. Fig. 9(a) shows a super-rotation flow with earth equatorial zonal flow of about 16-2 m s\(^{-1}\) and Fig. 9(c) the response ten days after the
switch-on of a forcing centred at 5°N with a ratio of zonal to meridional scales of 3 to 1. There is now an almost isolated wave-train proceeding north and east from the source.

The solution shown in Fig. 6 was generated as a sum of standard Rossby-Haurwitz waves but for conceptual purposes this view is not very useful. However, if we return to the Haurwitz theory with these experiments in mind it is possible to make a more incisive analysis. Analysing the behaviour of waves on the spherical domain in the standard manner of Eq. (7) is equivalent to using the separable representation in terms of functions like $\cos kx \cos ly$ on the infinite $\beta$-plane. This is mathematically possible, but it was found much more convenient to analyse in terms of the modes $\cos K_x x'$ where $x'$ is the distance along an axis rotated from the zonal direction. As is easily seen from Eq. (5) and was implicit in the work of Barrett (1958) and Longuet-Higgins (1964), the spherical harmonic $Y^n_n(\lambda_i, \phi_i)$ with $n = n_x$ is a stationary solution for any rotated coordinate system on the sphere: there is no need for the mathematical 'pole' to coincide with the geographical pole. As on the $\beta$-plane, we may expect it to be most convenient to analyse in terms of the gravest mode in $\phi_i$, that is in terms of modes $Y^n_n(\lambda_i, \phi_i)$. The $n_x = 8$ mode for one coordinate orientation is shown in Fig. 7(a). The path of the wave-train is a great circle. One can now envisage any stationary wave field as being described by a sum of modes on different great circles:

$$\psi = \sum_i \psi_i Y^n_n(\lambda_i, \phi_i),$$

(8)

where $n = n_x$. For example, half the sum of the mode shown in Fig. 7(a) and the one rotated in the other direction gives the stationary wave exhibited in Fig. 7(b). A representation of $\psi$ as in Eq. (8) is clearly a much more revealing and fundamental way of analysing the results shown in Figs. 6 and 9(e). Both pictures show the great circle signature and the latter in particular looks very similar to part of a single rotated spherical harmonic. When transience or dissipation are present one must expect only a portion of the rotated mode to be present and it is not then necessary for $n_x$ to be an integer.

One can proceed directly to group velocity ideas but it appears to be best, following Hoskins and Karoly (1981), to use a Mercator projection of the sphere and then apply the ray-tracing notions developed in the field of geometrical optics (Lighthill 1978). One finds that, for a super-rotation flow, the rays are indeed great circles and that the group velocity is $2\omega \cos \alpha$ along a great circle tilted by an angle $\alpha$ from the equator. This result is entirely analogous to that for the infinite $\beta$-plane and one may now produce a figure for the sphere corresponding to Fig. 4(b). Figure 8(a) shows such a figure for a source at
Figure 7. Two perturbations that would be stationary on an \( n_s = 8 \) super-rotation basic flow. (a) A rotated fundamental spherical harmonic obtained by moving the south pole of a \( Y_8^0 \) mode (Fig. 5(a)) to 45°E 45°S where 0°E 0°S is the point in the centre of the plot. (b) The result of adding half the rotated mode shown in (a) and half that obtained by moving the north pole of a \( Y_8^0 \) mode to 45°E 45°N.

Figure 8. Steady wave propagation from sources at 15°N for two different basic zonal flows. Successive crests and troughs are marked assuming a ten-day time-scale for wave damping. (a) An \( n_s = 7.5 \) super-rotation. (b) A northern hemisphere winter climatological zonal flow. The projections are polar stereographic and lines of latitude and longitude are drawn every 30°. (Adapted from Hoskins and Karoly (1981))

15°N in an \( n_s = 7.5 \) super-rotation, assuming a ten-day dissipation time-scale.

Both the numerical model and ray-tracing can be used in more general situations. A zonal flow based on the observed northern hemisphere 300 mb zonally averaged zonal flow in winter is shown in Fig. 9(b) and the ten-day response in Fig. 9(f). The super-rotation flow in Fig. 9(a) had the same angular momentum as the flow in Fig. 9(b). Comparing Figs. 9(e) and (f) shows that the great circle propagation idea is not greatly modified though the wavelength is rather larger in the stronger mid-latitude westerlies in Fig. 9(f). Assuming that the zonal flow varies little in a wavelength, ray-tracing in the observed zonally averaged flow gives the propagation picture sketched in Fig. 8(b). There is now a split into a poleward propagating train and one trapped equatorward of 40°N. The poleward train corresponds well with that in Fig. 9(f). The split is not noticeable
there because of the bias to polar propagation given by the zonally elongated shape of the forcing. It does agree extremely well with the steady mountain solutions shown in Fig. 4 of Grose and Hoskins (1979). The larger amplitudes in Fig. 9(f) compared with 9(e) may also be understood in terms of ray theory. It should be noted that, according to ray theory, stationary waves are absorbed at a critical line at which $\tilde{u} = 0$. In this region, the linearization employed breaks down and there is much discussion as to the real processes occurring there (see, for example, Held (1983); and Killworth and McIntyre (personal communication)).

(d) *Longitudinal variation in the basic flow*

A vorticity forcing may easily be determined that makes the observed northern hemisphere 300 mb longitudinally varying winter averaged flow a solution of the barotropic vorticity equation. This flow may then be perturbed in the same manner as before
Figure 9. (e)-(h) show the vorticity perturbations at day 10 corresponding to the basic flows (a)-(d), respectively. (Courtesy of A. J. Simmons, and similar to the results in Simmons (1982))

(Simmons 1982). Figs. 9(c) and (d) show two different longitudes for the forcing with respect to the time-averaged flow, rotated so that the forcing is in the same position throughout Fig. 9. The solution for forcing at 135°W (Fig. 9(g)) still shows very similar propagation from the source region. However, that from a source at 135°E (Fig. 9(h)) gives an extremely large response at the difluence of the Asian jet, with signs of a great circle propagation downstream from this region. As discussed in Simmons (1982), the model picks out the Aleutian and Icelandic regions as the two regions for large response. This is in good agreement with the areas of maximum low-frequency (10- to 90-day) variability found in winter data by Blackmon (1976). Ray theory has been used by Karoly (personal communication) on the longitudinally varying flow and there is some slight evidence of ray convergence in the sensitive areas. However, the striking result of Fig. 9(h) is not reproduced, which is scarcely surprising considering the lack of separation in scales between the propagating wave-trains and the basic climatological flow. More recent work
by Simmons et al. (personal communication) suggests that the forced response in the Aleutian region may be enhanced by a local barotropic instability, with wave propagation towards and away from this region much as given by the theory described above.

(e) Baroclinic theory

Similar experiments can be performed with baroclinic primitive equation models though any time-dependent models will tend to include the initiation of baroclinic instability. Figure 10 shows the steady-state result of including a longitudinally elongated heating centred on 15°N in a model linearized about a northern hemisphere winter zonal flow and including some wave damping. A map of height field tends to accentuate events

![Diagram](image)

Figure 10. Steady-state, damped, linear 300 mb height field response to a heating in the stippled region perturbing a northern hemisphere winter zonal flow. The contour interval is 2 dam if the depth-integrated heating maximum is 2.5 K d$^{-1}$. For comparison, a great circle has been drawn with a heavy continuous line. (Adapted from Hoskins and Karoly (1981))

in higher latitudes but the poleward 'great circle' mode of propagation is again very apparent. Away from the source, the height field response has a simple vertical structure: in phase at all levels with a maximum amplitude at the tropopause. That barotropic theory should work so well for Rossby wave propagation in a baroclinic model has been
explained elegantly by Held (1983) in terms of the generation of this equivalent barotropic mode.

The inclusion of longitudinal variation in the basic flow in simple forcing experiments using baroclinic models is again made more difficult by the presence of baroclinic instability. However, investigations are proceeding.

(f) Recent observational studies

Since the study by Lorenz (1956), one way of looking at horizontal structures in the geopotential height field in the atmosphere has been to use an empirical orthogonal function or principal component analysis. The first component maximizes the normalized variance of a height–time series. Subsequent components explain a maximum of the residual normalized variance and are temporally uncorrelated. However, the components are generally global in character and not dissimilar from the \((\cos kx \cos ly)\)-type representation which was not useful on the \(\beta\)-plane, or the \(Y_n^m(\lambda, \phi)\) on the sphere.

Another type of observational study has been to pick a particular phenomenon at a particular point and to correlate with it other features elsewhere. This procedure was

Figure 11. The western Atlantic pattern at 500mb found by Wallace and Gutzler (1981) (adapted from their Fig. 21). Shown is the difference between the 10 months with the largest positive values of an index of the pattern and the 10 months with the largest negative values, in a set of 45 winter months. The contour interval is 4 dam. A great circle has been superimposed for comparison.
recently carried out exhaustively by Wallace and Gutzler (1981) for northern hemisphere 500 mb height analyses for 45 winter months. A calculation of the temporal correlation coefficients between data at every pair of gridpoints was made. The highest correlation coefficients then led to a choice of points for which the correlation maps will be most informative and indices of various patterns constructed. Figure 11 shows the western Atlantic pattern obtained by taking the difference in 500 mb heights between the ten months with the most positive and the ten months with the most negative values of an index of this pattern based on the gridpoints 30°N 55°W and 55°N 55°W. Also marked in Fig. 11 is a great circle which suggests that atmospheric teleconnections proceed along paths not too dissimilar from great circles. Other patterns found by Wallace and Gutzler have a similar nature or emphasize a north–south dipole structure. Horel (1981) showed that similar spatially localized patterns can be obtained by ‘rotating’ principal components and relaxing the necessity for spatial orthogonality, which is somewhat analogous to using the rotated $Y_n^m$ modes for the theoretical analysis.

The study of Wallace and Gutzler showed monthly mean patterns but gave no idea of cause and effect, analogous to the group velocity concepts introduced in the theory. However, in a recent study, Blackmon and Wallace (personal communication) have taken 5-day-mean data and produced lagged correlations for various gridpoints. For the point at 55°N 55°W, which was a key point in the western Atlantic pattern, Fig. 12(a) shows correlation with data five days previously and Fig. 12(b) that for five days afterwards. The dominant pattern in Fig. 12(a) is near a great circle mostly to the west of the point. That in Fig. 12(b) is near a different great circle mostly to the east. This strongly supports the relevance of linear wave theory with its attendant concept of group velocity.

![Figure 12. Lagged correlations of 500 mb height field with that at the point 55°N 55°W using 5-day-mean northern hemisphere winter data. The contour interval is 0·1. (a) Correlation with mean 500 mb data five days previously; (b) correlation with mean 500 mb data five days afterwards. (Courtesy of Blackmon and Wallace)](image)

**Discussion**

The subject of teleconnections now seems to fit in well with the scheme shown in Fig. 2. The notion of propagation of an equivalent barotropic wave-train along almost great circle rays is one that is proving useful in observational and general circulation
studies. Barotropic theory and models and steady-state primitive equation models can now be used in diagnosis and in giving mechanisms and a priori hypotheses that may be tested. Previously, general circulation model studies such as that of Rowntree (1972) had proved stimulating, but convincing results of the effects of changing boundary conditions were difficult to obtain. The recent work of Shukla and Wallace (personal communication) is an example of a study of the global response to tropical sea-surface temperature anomalies which has benefited from being able to test a priori hypotheses. The theory should be directly useful to the long-term forecaster in giving him a more rational basis for cause and effect.

Many diagnoses of atmospheric behaviour have proceeded on the basis of zonal Fourier analysis and it is worth considering why a 'great circle' description is more useful in a description of low-frequency variability in the northern hemisphere winter troposphere. Compelling evidence that this is the case has been given by Wallace and Hsu (personal communication). The first reason must be that the forcing of these waves, associated with mountains, heat sources and storm tracks, is essentially local as are the sensitive unstable regions (e.g. the Aleutians) which can themselves act as a source of Rossby waves. If the forcing was predominantly in zonal wavenumber three, for example, then zonal Fourier analysis would be the best tool. The second reason is that the tropospheric scales implied by Eq. (6) are definitely sub-planetary and not larger than that of the mean flow itself. Thus wave behaviour can be in some senses local. In the stratosphere, on the other hand, there is mostly global wave-like forcing and Eq. (6) suggests that quasi-stationary waves must be global. Thus although ray-tracing and wave-packet ideas can be applied to vertical propagation (Hayashi 1981, Karoly and Hoskins 1982), zonal Fourier analysis remains the predominant diagnostic tool.

At this point one may return to the rhetorical question asked in a previous Symons Memorial Lecture by Platzman (1968): 'Has a Rossby wave ever been observed in the atmosphere?' Fourteen years later, even for the troposphere, we can go beyond his own cautious reply and give a rather definite 'Yes'!

3. MIDDLE LATITUDE SYNOPTIC SYSTEMS AND THEIR FRONTS

(a) Introduction

Having quoted Bergeron's challenge on the development of structural models earlier, we now turn to the most famous of conceptual models in whose development he played a large role, the Norwegian cyclone model. Figure 13 contains a summary of the events postulated at the surface, starting from a pre-existing discontinuity surface, developing into an open wave with a cold front and a warm front, and finally the occlusion process in

![Diagram of Norwegian cyclone model](Image of Norwegian cyclone model)

(a) (b) (c)

Figure 13. The classic Norwegian cyclone model showing three stages in the life of a polar front wave.
which the cold front overtakes the warm front. This model and its application was resisted strongly for many years in many countries, but having gained acceptance it now tends to be clung to with religious tenacity. It provides the basic way of reducing and communicating the enormous quantity of numbers produced by a forecast model. However, one must surely ask whether the theoretical work of the last sixty years is able to influence the way in which such data are studied, both on output from the numerical model and also on input to the model, and even the construction of the model itself.

With the baroclinic wave theories of Charney (1947) and Eady (1949) the theoretical emphasis shifted to the idea that synoptic systems develop as an instability of a broad baroclinic region. Much has been written on this subject since the work of Charney and Eady though the gap between observed systems and theory remains wider than it should be. From the theoretical viewpoint, fronts develop in a growing baroclinic wave and the process of frontogenesis has received some attention in the literature. Though less has been written on this problem, it is also considered that shallow, smaller-scale disturbances can develop on the fronts of the major systems.

In section 3(b) we will give a very brief review of the results of two-dimensional frontogenesis theory followed in section 3(c) by a discussion of the synoptic structures shown in models of baroclinic instability.

(b) Two-dimensional frontogenesis theory summary

A review of the mathematical theory of frontogenesis is given in Hoskins (1982), and reference may be made to that article for details of the theory justifying the qualitative statements which will now be made.

Two-dimensional frontogenesis theory starts from a situation sketched in Fig. 14(a) with a horizontal temperature gradient and a tendency for large-scale quasi-geostrophic motion to increase this temperature gradient moving with the air. This tends to destroy

![Figure 14. Two-dimensional frontogenesis theory in a dry atmosphere. (a) The basic geostrophic flow. A temperature gradient between warm air (W) and cold air (C) is in thermal wind balance with the wind into the section (⊙) and out of the section (⊙). The large-scale geostrophic flow is tending to increase the temperature gradient, possibly by the convergence shown. (b) The ageostrophic circulation induced. (c) The distortion of the ageostrophic circulation as the vorticity becomes larger near the surface on the warm side of the contrast. (d) The distribution of low-level temperature (T) across the frontal region.](image-url)
thermal wind balance which can only be retained by means of a direct cross-frontal circulation with warm air rising and cold air descending, as in Fig. 14(b). The consequent adiabatic cooling in the warm air and warming in the cold air act to inhibit the formation of very large temperature gradients in the mid-troposphere, but at low levels on the warm side of the contrast the ageostrophic convergence leads to larger values of horizontal temperature gradient and vorticity. As these increase, so the ageostrophic circulation becomes distorted (Fig. 14(c)) so that the ageostrophic convergence in this region becomes larger. There is therefore a feedback between the magnitude of the ageostrophic convergence and the magnitudes of the gradients in both temperature and velocity. If the large-scale motion is such that air stays in a frontogenesis region then the magnitudes of gradients in velocity and temperature are limited only by frictional processes. The approach to a discontinuity occurs on the time-scale for which pure geostrophic advection would predict a growth in the temperature gradient by a factor of \( e \), typically of order one day. Note that the nonlinear frontogenesis occurs on the warm side of the surface temperature contrast, so that the surface temperature distribution tends to look as shown in Fig. 14(d). Since the mid-level temperature distribution tends to be more symmetric, the position of maximum temperature gradient tilts towards the cold air with height.

When comparing this simple picture with observed surface fronts, one finds many points of agreement but some of disagreement. That the strongest surface gradients occur on the warm side of the contrast and the gradients are most intense near the surface accord with reality. However, fronts can be more intense in the middle troposphere than simple models suggest. The picture of warm air rising and cold air descending is consistent with high humidity and rainfall on the warm side and low humidity and clearance of cloud on the cold. However, the model upward vertical motion maximum is typically of order 1 cm s\(^{-1}\), which is much smaller than observed.

The inclusion of other physical processes such as boundary layer friction and the release of latent heat does not drastically alter the model structures, though, as shown by Hoskins (1974) and more definitively by Williams et al. (1981), the realism of the model front increases: the intensity of the middle tropospheric front is increased and the maximum ascent rises by a factor of order five. Blumen (1980) has performed a detailed comparison of an observed front with that in a simple two-dimensional model.

It is not the intention here to discuss upper air frontogenesis in detail, but it is clear that as far as the tropopause acts as a lid on the troposphere, then the region below the tropopause on the cold side of the transition (Fig. 14(b)) is another frontogenesis region. Simple models show the folding of the tropopause that occurs in this region and exhibit a frontal structure quite similar to those observed by Reed and Danielsen (1959) and others but of less intensity than documented rather extreme events.

One of the features of two-dimensional frontogenesis theory that may be of great importance is that frontal-scale phenomena are slaves to larger-scale quasi-geostrophic flow. Frontogenesis and direct cross-frontal circulation are present only as long as the large-scale motion is tending to increase horizontal temperature gradients. If that motion reverses its tendency, then frontogenesis becomes frontolysis. Quasi-geostrophic arguments are therefore very useful in elucidating frontal behaviour. This idea will be returned to below. The prospect for mesoscale models, given the lack of data on the small scale, is also much brighter if the large scale plays such a dominant forcing role.

(c) **Model baroclinic wave structures near the surface**

The surface synoptic structures of nonlinear baroclinic waves have not received a great deal of attention in the literature and so we shall now endeavour to present a rather limited review starting from developments of the Eady model and working towards primitive equation solutions on the sphere including some parametrizations of physical processes.

Quasi-geostrophic theory gives a good indication of nonlinear baroclinic wave structure, but to exhibit realistic frontogenesis one needs to proceed to the level of the semi-
geostrophic equations (Eliassen 1948, Hoskins 1975). In two-dimensional regions the equations are identical with those used in the above frontogenesis models, and in simple cases such as those discussed here the solutions are merely distortions of the quasi-geostrophic solutions.

Arguably the simplest linear baroclinic wave is that determined by Eady (1949) as the unstable mode independent of y on a basic flow \( \bar{u} = Az \), with uniform stability and Coriolis parameter between two rigid boundaries at \( z = 0 \) and \( H \). The nonlinear two-dimensional Eady wave is a simple analytic solution of the semi-geostrophic equations, and the surface pressure and temperature fields at one particular day for a basic flow with maximum velocity 29.4 m s\(^{-1}\) are shown in Fig. 15(a). The pressure contours are straight and parallel to the y axis (beware the optical illusion!). The relative vorticity is a maximum of 2-1\( f \) in the centre of the low and the temperature gradients are enhanced in the region of the low.

Figure 15. Surface maps for nonlinear, modified Eady waves. Surface pressure contours are shown using continuous lines, temperature contours with dashed lines. The contour intervals are 12 mb and 8 K in (c) and 18 mb and 12 K otherwise. The regions in which the relative vorticity is greater than \( f/2 \) are stippled. The domain shown is approximately 5600 km\(^2\). (a) Day 5 with initial zonal flow having no horizontal shear (and latitude-independent disturbance). (b) Day 5-5 with initial zonal flow having small upper-level shear. (c) Day 6-3 with initial zonal flow at upper levels falling to zero on the edges. (d) Day 5-5 with initial zonal flow having low-level easterlies in the centre and westerlies on the edges. (For more detail see Hoskins and West (1979).)
The parts of Fig. 15 are surface maps for the nonlinear stages of baroclinic instability modes growing on various modifications of the basic Eady profile, though all maintaining the uniform potential vorticity and Coriolis parameter of the basic case. Fig. 15(b) corresponds to a basic flow which is still zero at the surface but rises to a sinusoidal profile at \( z = H \) varying from 29.4 m s\(^{-1}\) in the centre to 20.6 m s\(^{-1}\) on the edges. The surface wave tilts in sympathy with the imposed upper-level horizontal shear and has largest amplitude in the centre of the channel in the region of maximum baroclinity. A 'cold front' region is evident in the temperature and vorticity fields, the latter having an isolated maximum of 2.2\( f \) there. The positive vorticity does not pick out any 'warm front' region though there are increased temperature gradients to the north-east of the low centre. The basic flow for the Fig. 15(c) surface map is the same as that for Fig. 15(b) except that the upper flow variation is from 29.4 m s\(^{-1}\) in the centre to zero on the edge. The velocity and temperature gradients in the cold front region are intense, with relative vorticities up to 5\( f \) and temperature contrasts of 4 K in 40 km. Again there are large temperature gradients also to the east and north of the low and a late development in the wave is the increased vorticity eastwards along this region.

This progression from the Eady wave shows how nonlinear baroclinic waves can form intense cold fronts and, as a rather secondary feature, a region which may perhaps be called a warm front. This warm front is very definitely not ahead of the 'warm sector' to the south-east of the low pressure centre. However, such a warm front, more in accord with the textbook, may be obtained in a nonlinear baroclinic wave development with this model. Fig. 15(d) shows a nonlinear baroclinic wave on a basic flow which is modified so that there are surface easterlies in the centre of the channel and westerlies on the edges, both with maximum speed 4.4 m s\(^{-1}\). The upper-level flow is a uniform 25.4 m s\(^{-1}\) so that the mid-channel baroclinity is identical with the other cases. The wave tilts in sympathy with the low-level flow, and the strongest temperature gradients and maximum vorticity are ahead of the warm sector. This is a primary warm front of a very different character. Now the picture departs from the textbook in that there is no cold front to the south of the low pressure centre, though there is a feature that could be described as a cold front to the north-west of the low.

In all the solutions illustrated in Fig. 15, part of the occlusion process is present if one considers it to be the shrinking of the region of warm air at the surface. With the conversion of potential to kinetic energy that occurs, this shrinkage is inevitable. However, the textbook story of the overtaking of a warm front by a cold front is not seen.

A possible next step in a hierarchy of models applied to this problem is to use the primitive equations in a true spherical domain. They contain less approximations than the semi-geostrophic equations used above, but the numerical resolution of positive vorticity frontal regions and the production of simple physical explanations are less easy. Fig. 16(a) shows a zonal flow that has been used in a high-resolution model restricted to six-fold symmetry on a hemisphere. A baroclinic wave of zonal wavenumber 6 develops a near-surface structure which is exhibited in Fig. 17(a). Surface pressure and low-level temperature, vertical velocity and the positive relative vorticity maximum are indicated. The positive vorticity has very much a linear distribution along the line indicated and there are intense thermal gradients of order 8 K in 300 km in what is clearly a cold front region. There is strong warm advection only a small distance ahead of this front and marked ascent only on its northern end, but strong descent exists behind. To the north of the low there are again strong thermal gradients and large vorticity. The ascent is a maximum near the most eastward part of the vorticity maximum.

The structure described for the primitive equation, spherical, non-linear baroclinic wave is very similar to that in Fig. 15(c) apart from the expected rather smaller intensity in the cold front region and the lack of the positive vorticity maximum along the possible warm front region ahead of the low. However, a day later (Fig. 17(b)) such a structure is apparent, along with a general development of the system.

Various modifications to the basic flow in Fig. 16(a) and a study of zonal wavenum-
Figure 16. Basic zonal flows for the primitive equation integrations described in the text and in Simmons and Hoskins (1980). Light contours are those of zonal velocity drawn every 5 m s\(^{-1}\) with the zero contour in (b) very heavy. Medium contours are those of potential temperature drawn every 5 K.

ber 9 instead of 6 have been tried, but all except one have yielded essentially the same synoptic structure, lacking the textbook warm front. This one case was inspired by the semi-geostrophic model results. A barotropic modification to the basic zonal flow was imposed, giving 10 m s\(^{-1}\) surface westerlies at 20°N and 10 m s\(^{-1}\) surface easterlies at 50°N. This flow is shown in Fig. 16(b). As anticipated, the low-level mode tilts in sympathy with the shear and, as shown in Fig. 18, produces a primary warm front structure very similar to that in Fig. 15(d). This warm front has large positive relative vorticity along its length and ascent ahead of it. Only a small distance behind it there is cold advection, but no real cold front.

The consistency of the theoretical model surface structures shown in Figs. 15, 17 and 18 and their disagreement with the classic conceptual model sketched in Fig. 13, leads one to question whether the theory is missing a crucial ingredient or whether the conceptual model, if not invariably wrong, is by no means ubiquitous. The textbook cold front is easily obtained in models, but these cases lack a real warm front ahead of the warm sector. Conversely, the rarer cases with such a warm front lack a real cold front. The occlusion process in the models is not a simple overtaking of the warm front by the cold front. Discussions with those well versed in synoptic data suggest that 'textbook cases' are indeed difficult to find and that the model behaviour may be quite realistic.

The obvious processes missing in the theoretical models are representations of latent heat and boundary layer fluxes. Inclusion of such parametrizations in baroclinic wave simulations like that in Fig. 17 tends to produce convective rain at the cold front and large-scale precipitation in the warm air rising ahead of the low pressure system. Again, vertical velocities are increased, but the low-level synoptic structure is not fundamentally altered. Golding (1981), using an operational forecast model, has compared pure baroclinic wave growth without and with the model's parametrization package. He found more rapid development with rainfall included but little difference in surface structures. By placing tracers in the flow he was able to see that, in the moist case, warm air rises ahead of the low and becomes part of the strong north-westerly flow ahead of the upper ridge.
Figure 17. The low-level structure of a most unstable zonal wavenumber 6 on the basic flow given in Fig. 16(a). Days 5 and 6 are shown in (a) and (b) respectively. The maps show $1\frac{1}{2}$ wavelengths (90° of longitude) and are polar stereographic, cutting off at 20°N. Surface pressure contours, every 8 mb, are shown using continuous lines; 967 mb temperature contours, every 8 K, are dashed; the positive vorticity maximum at 967 mb is indicated by a heavy dashed line, ascent and descent at 967 mb greater than 1.5 mb h$^{-1}$ are shown by heavy dots and light dots, respectively.

(d) Discussion

The present situation, with disagreement between the conceptual frontal model used for forecast model and atmospheric diagnosis, and the theoretical models of the low-level structure of baroclinic waves, is unsatisfactory. However, all the tools are available to remedy this situation. An hierarchy of numerical models applicable to the problem is available, ranging from semi-geostrophic to operational forecast models. Quasi-geostrophic theory has been shown to give useful information on frontogenesis, and advances in the theory (Hoskins and Pedder 1980, Hoskins 1982) allow quite simple
diagnosis of synoptic structures. Satellite pictures are providing a wealth of information on cloud structures associated with synoptic systems, and radars (e.g. Browning 1978) are providing routine detailed views of frontal characteristics. Special observational programmes have also been mounted (e.g. Browning et al. 1973, Hobbs and Locatelli 1978).

Given the will to change, useful improved models of mid-latitude synoptic systems and their fronts can be developed over the next few years.

4. Conclusion

This lecture has been an attempt to describe two areas of current meteorological research and to use them to illustrate the advantages of the approach sketched in Fig. 2 in which atmospheric observations and models of all complexities are used to probe phenomena of interest and to produce evolving conceptual models of them. This approach is promising in many areas. One further example which is of much interest at this time is the study of climate and climate change. General circulation models are crucial in this programme but they must be seen as only part of the weaponry at our disposal. Equally, very simple low-order climate models are only one end of the complete and balanced attack that is needed.

References

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeh, T.</td>
<td>1949</td>
<td>On energy dispersion in the atmosphere. J. Met., 6, 1–16.</td>
</tr>
</tbody>
</table>