AN ESSAY ON THE PARAMETERIZATION OF OROGRAPHIC GRAVITY WAVE DRAG

R. T. Pierrehumbert
Geophysical Fluid Dynamics Laboratory/N.O.A.A.
Princeton University
Princeton, NJ 08542

1. Introduction

For a long time, there has been an awareness that distributed drag associated with orographically produced internal gravity waves might have a significant effect on the general circulation of the atmosphere. Perhaps the first quantitative estimates of the importance of wave drag may be found in Bretherton (1969). Lilly (1972) explicitly suggested that it might be necessary to parameterize mountain wave drag in general circulation models (GCMs), and carried out a crude experiment along the indicated lines. The first attempt to incorporate a complete drag parameterization in a GCM was made by the Canadian climate group (see, e.g. Boer, *et al.* 1984), who have carried out an extensive series of simulations with a fairly low resolution model. These calculations were important in providing quantitative evidence that mountain wave drag could produce a significant slowing of the westerlies in a realistic model; however, the first convincing evidence that the drag produces a *beneficial* effect on forecast skill and systematic error patterns came only with the high resolution calculations reported in Palmer, *et al.* (1986). It is becoming generally accepted that mountain wave drag has a more dramatic impact on the systematic error than any other improvement in model physics that has been tried to date, though the physical reality of the gravity wave drag is still a matter of considerable dispute; the possibility remains that the missing drag in the models arises from some momentum transport process other than orographic gravity wave drag.

As of this writing, parameterizations of mountain wave drag

have been for the most part based on linear theory. However, it will be seen that under typical atmospheric conditions, flow over even modest mountains lies in the strongly nonlinear range. Nonlinear effects can lead to a radical change in the character of the flow and cause amplifications of the drag by a factor of five or so. Understanding the implications of the nonlinear behavior for the drag parameterization problem will be the central concern of this paper.

We shall also discuss the effects of wave drag in simplified and realistic general circulation models; although an extensive body of simulations of the effect of wave drag have by now become available, there has been relatively little attention paid to the subtle dynamical mechanisms determining the response of the planetary scale circulations. The results with an idealized model help to elucidate these effects, while comparison with results from a realistic GCM suggest that the idealized model yields insights that are useful in a practical context; the realistic run also shows that earlier beneficial results on the effects of wave drag can be reproduced with the scheme suggested herein, which differs in some particulars from the parameterization described in Palmer, *et al.* (1986).

In order to concentrate on nonlinear aspects of the problem, we shall assume a familiarity with the basics of linear mountain wave theory. Readers requiring a review of this material may find it profitable to consult the article by Smith (1979).

In §2 we provide an overview of some aspects of nonlinear, stratified flow over mountains crucial to drag parameterization. This is followed in §3 by the specification of a parameterization based on these considerations. The effects of an idealized drag in a simplified GCM are discussed in §4, while some tentative results on the effects of realistic mountain wave drag in a full prediction model are given in §5. Our principal conclusions are rehearsed in §6.

2. Nonlinear, stratified flow over mountains: an overview

Drag and momentum flux exhibit a rather startling behavior in the nonlinear regime. In order to highlight the implied challenges for

the parameterization problem, we shall in this section focus attention on the simplest configuration which admits an archetype of the nonlinear behavior. Specifically, we consider two-dimensional nonrotating, stratified flow over a ridge with height profile h(x). Let z be the vertical coordinate, w be the vertical velocity and u the velocity in the x direction. We assume the flow impinging on the mountain to be stably stratified, with Brunt-Väisala frequency N(z). In this situation, the mountain will set up motions ("gravity waves") which propagate vertically and horizontally away from the mountain. Before proceeding further, we shall need a few elementary definitions and properties.

The *drag force* is the force exerted by the mountain on the atmosphere (which is equal and opposite to the force exerted by the atmosphere on the mountain), and is given by

$$D=-\int \hat{n} p\, ds \qquad (2.1)$$

in which p is the pressure, s is the arc length along the mountain and $\hat{n}$ is the unit normal to the mountain; for mountains of small slope, $D=-\int p(dh/dx)dx$. The *vertical momentum flux* is defined by

$$\tau(z)= \int \rho u w\, dx \qquad (2.2)$$

in which ρ is the density. The importance of the momentum flux resides in its role in determining changes in the mean flow. In the absence of other body forces acting on the fluid, the mean flow change is given by

$$\rho d[\int u dx]/dt=-d\tau/dz \qquad (2.3)$$

So far we have merely engaged in making definitions. Now we come to an important and profound properties of mountain waves. *In the absence of transience or dissipation, τ is independent of height and equal to the drag D.* Thus, according to (2.3), steady, undissipated gravity waves exert no force on the atmosphere; the drag force exerted by the mountain shows up in the vertical momentum flux of the gravity waves which is carried upward out of the domain without being deposited anywhere. Therefore, the influence of gravity waves on the atmosphere is intimately connected with the dissipation of the waves.

Since molecular viscosity and radiative damping have only a weak effect on the waves throughout most of the troposphere and stratosphere, the main dissipation is mediated by turbulent mixing of

heat and momentum. Typically the turbulent dissipation arises primarily through turbulence induced by the wave itself in regions where it becomes unstable to small scale instability. The process whereby a wave becomes unstable to instabilities of a scale substantially smaller than the wave itself is an instance of what is generally known as *wave breaking* (though we hasten to add that the existence of small scale instability should not be taken as the definition of breaking; any sort of irreversible wave-induced mixing will do). There are many types of wave breaking, but the simplest to understand is *convective wave breaking* , in which the wave attains at some height sufficient amplitude to overturn potential temperature contours, advecting a tongue of heavy fluid over a body of lighter fluid. Such a configuration is unstable to convection; since inviscid convection sets in preferentially at arbitrarily short wavelengths, there is no uncertainty in the application of a local criterion to ascertain stability, as there is a good separation of scales. As will be shown shortly, much of the momentum flux is deposited within the breaking region. The breaking process and associated flux profile is illustrated schematically in Figure 2.1, for flow impinging on the mountain from the left.

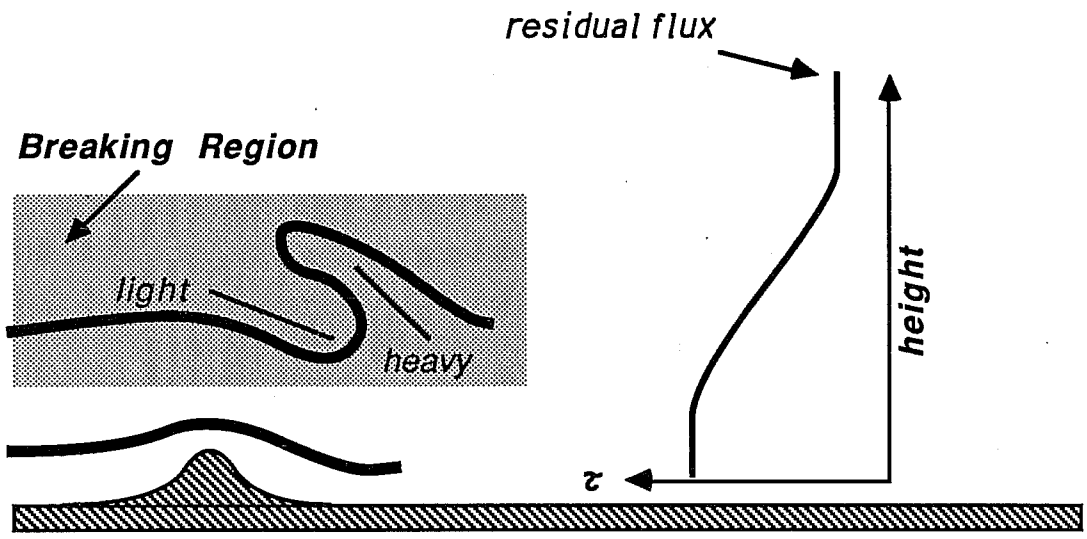

Figure 2.1

The flux is approximately constant above and below the breaking region, but decreases in magnitude as the breaking region is traversed. The flux is shown as being negative because the drag force in most situations is opposite to the direction of the oncoming flow (as a well-behaved drag should be).

Under what circumstances do waves break convectively, and at what level do they break? Consider first the case in which the flow impinging on the mountain has constant velocity U and stability frequency N far upstream of the obstacle, and in which the effects of compressibility can be neglected. Since this simple flow will play an important role in the subsequent discussion, we digress briefly to consider the nondimensionalization and scaling of the problem. If horizontal distances are nondimensionalized with the mountain half-width L, vertical distances with U/N, and times with L/U, one obtains a problem which is characterized (for fixed mountain shape) solely by the two numbers $Fr \equiv Nh_0/U$ and $b=NL/U$, where h_0 is the mountain height. If $b \gg 1$, the hydrostatic approximation is valid, and the problem is characterized by Fr alone. For smooth, symmetric mountain profiles, a range of nonlinear steady and initial value calculations show that in the hydrostatic range convective breaking occurs when Fr exceeds approximately .8 (see Pierrehumbert and Wyman 1985, Bacmeister 1986, and references therein). The breaking level is approximately 3/4 of a hydrostatic gravity wavelength above the ground (i.e. $3\pi U/4N$), though for values of Fr above the threshold for breaking, a convectively unstable region extends both above and below the breaking level. As the mountain is made narrower and the flow becomes less hydrostatic, the value of Fr required for breaking becomes larger (see, e.g. Bacmeister 1986); in the limit $b \to 0$, the mountain induces only potential flow, and hence breaking never occurs. With U=10 m/s and N=.01 sec^{-1}, breaking occurs for relatively broad (>5km) mountains whenever they exceed 800m in height; thus, low level wave breaking and the attendant strong nonlinearity are the rule rather than the exception.

The dramatic alteration of the character of the flow that occurs in conjunction with wave breaking may be seen by comparing Figures

2.2a,b. In these figures we show contours of potential temperature obtained from integrations of a 2D numerical model from shock start-up initial conditions with uniform N and U. In both cases the incident flow is from the left, and the mountain was made broad enough that the flow may be considered essentially hydrostatic. Further details can be found in Bacmeister (1986).

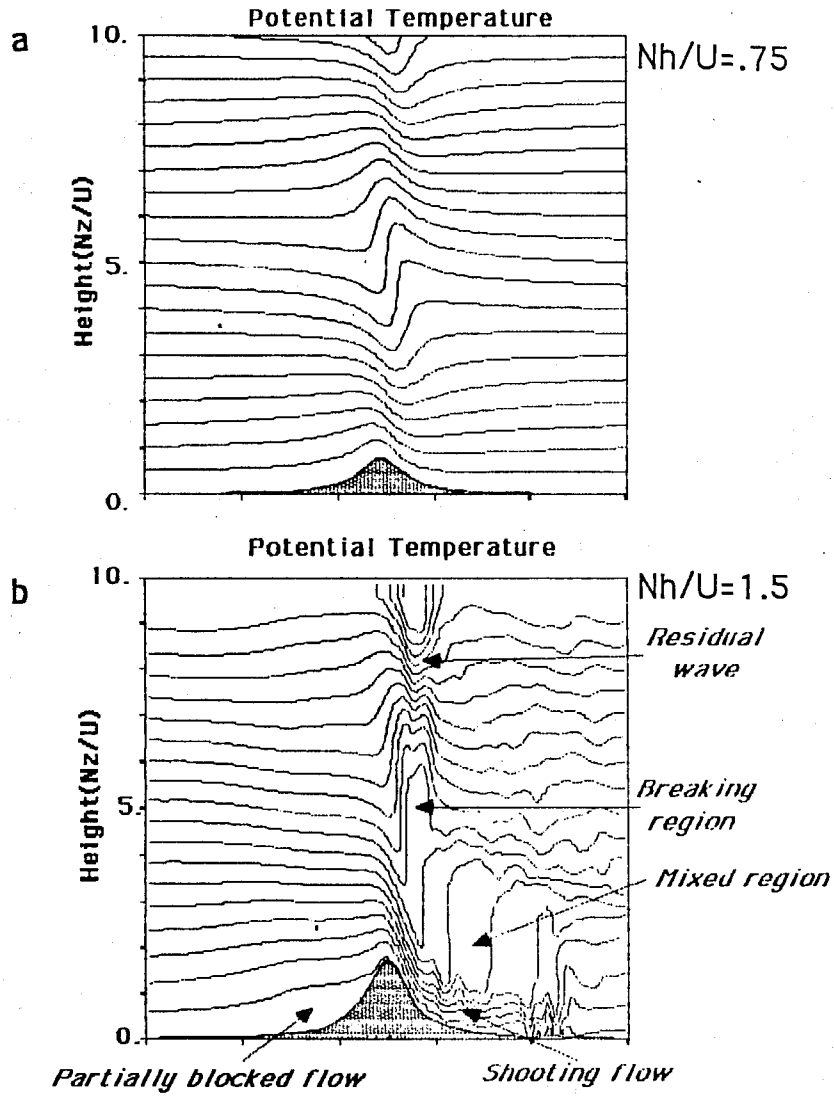

Figure 2.2

The calculation in Figure 2.2a was carried out with Nh/U=.75, just below the wave-breaking threshold, and the results are essentially the same as the classical Long's Model steady, nonlinear solutions; they also qualitatively resemble the familiar hydrostatic linear mountain wave. Far upstream and downstream of the mountain, the flow is undisturbed, though the wind is locally retarded on the upslope and locally enhanced on the downslope. Note that near z=5, the isentropes become nearly vertical, but the mountain is not high enough to cause overturning.

Results for a higher mountain (Nh/U=1.5) are shown in Figure 2.2b. Wave breaking has created a deep, stagnant mixed region in the lee of the mountain. Below this mixed layer lies a strongly stratified, swiftly moving current, the "downslope windstorm." As time goes on, this structure extends *arbitrarily far* downstream of the mountain, within the limitations of inviscid, 2D theory. In addition, the flow exhibits upstream influence and partial blocking: the low level flow is below the ambient speed within a stretch that extends arbitrarily far upstream as time progresses. This partial blocking visible in Figure 2.2b as an elevation of the low-level isentropes upstream of the mountain, terminated by a wave front which is propagating to the left. In contrast to the non-breaking case, the flow does not return to its ambient value far upstream and downstream (letting $t \to \infty$ first); the fluid experiences *permanent acceleration* as it crosses the ridge. In all, the low level flow configuration resembles that appearing in one-layer hydraulic theory. The reasons for this are by no means settled, but some useful insights may be found in Smith (1985). A residual vertically propagating wave may be seen above the breaking region; its amplitude is rather comparable to the marginally non-breaking case shown in Figure 2.2a.

The transition to the asymmetric hydraulic state described in the preceding paragraph is accompanied by a marked enhancement of the surface drag and low level momentum flux. These effects are summarized in Figure 2.3, where we plot the surface drag, the momentum flux at mountain top level, and the momentum flux aloft

(above the breaking level, if there is one) as a function of Nh/U. The estimates were obtained using the same model utilized for calculating

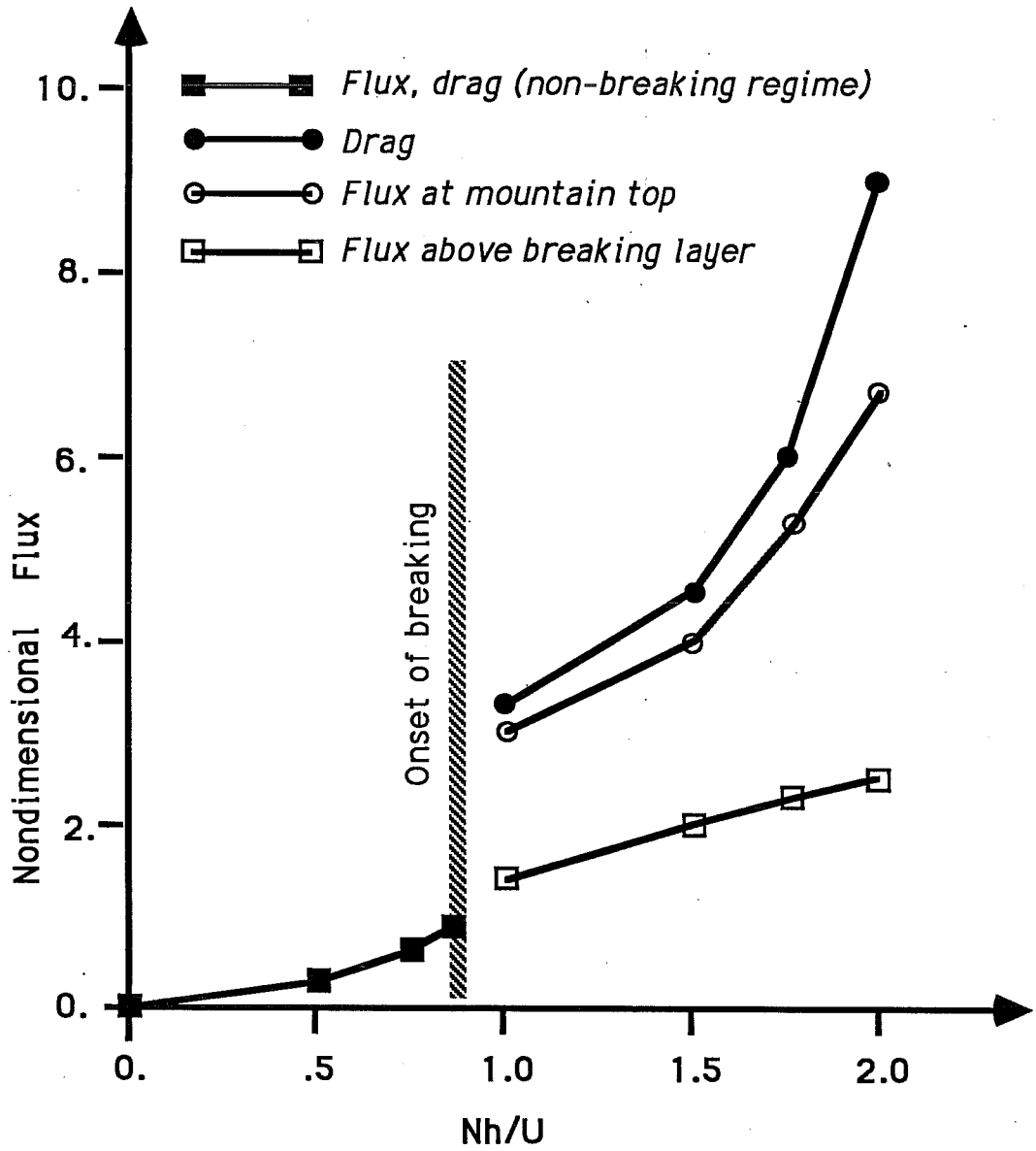

Figure 2.3
Drag and momentum flux as a function of Nh/U

Figure 2.2, integrated to a quasi-steady state. The drag D and flux τ are nondimensionalized with $\rho_0 U^3/N$. All three quantities plotted in Figure

2.3 are identical in the absence of wave-breaking. As the breaking threshold is crossed, the drag and low level flux jump by nearly a factor of 4. Nonlinear drag enhancement of this type (though not the connection with the asymmetric hydraulic states) was first noted by Peltier and Clark (1979), in a study of severe downslope wind storms. The residual flux remains roughly continuous across the breaking boundary, suggesting the operation of a "saturation hypothesis" in which the flux is limited to approximately the value for the marginally breaking case. As Nh/U is increased further, the residual flux increases only weakly, though the drag and low level flux continue to increase faster than linearly. This situation leads to large low level momentum flux convergences that are completely missing in linear theory. In fact, it is this low level convergence that drives the upstream influence and the downstream-propagating bore (Bacmeister, 1986).

We have described the breaking criterion for the case of constant N and U; now we turn to the general case in which N and U vary with height and also allow for compressibility effects. The primary interest of this problem lies in locating upper atmospheric breaking levels in cases where the mountain is too small to permit low level breaking, or to isolate "secondary breaking regions" aloft where the residual flux escaping the low layer breaking zone is deposited. The breaking criterion is most conveniently expressed in terms of a *saturation flux*, which is defined as the largest momentum flux which can be carried by the fluid without the occurrence of breaking. When the saturation flux is exceeded, breaking occurs and some momentum must be deposited. A saturation flux formulation of this sort was employed by Lindzen (1981) in his study of mesospheric gravity wave drag.

Since $\tau=\int\rho u'w'dx$, the maximum momentum flux for a wave with typical velocity magnitudes $[u']$ and $[w']$ is estimated as $\tau\approx\rho[u'][w']L$, where L is the length scale of the mountain (taken also to be the characteristic length scale of the wave). Now, if D is the characteristic vertical scale of the wave, $[w']\approx(D/L)[u']$. Further, for *convective breaking* $[u']=U$, the speed of the ambient wind. Thus, the saturation flux is

$$\tau_s = (\rho U^2 D)\alpha \qquad (2.4)$$

where α is a dimensionless order unity constant dependent on the geometry. In general, D is a complicated function of height which depends nonlocally on the flow configuration, and we can proceed no further without explicit computation of the wave pattern. However, in the WKB approximation (valid for U,N and ρ varying slowly with z, compared to a vertical wavelength), the vertical scale is determined locally. In the hydrostatic case, this characteristic vertical scale is D=U(z)/N(z) and the saturation flux becomes

$$\tau_s = (\rho U^3/N)\alpha \qquad (2.4')$$

Since τ is constant below the breaking level, it is easy to use (2.4') to provide an estimate of the breaking level. Increasing U inhibits breaking by increasing τ_s, while the decrease of ρ with height favors breaking. Because ρ decreases exponentially while U(z) remains bounded, breaking inevitably occurs at some level; however, the cubic dependence of τ_s on U makes it quite difficult for the density decay to win out over the general increase of U with height within the troposphere and lower stratosphere; this is particularly true in the winter hemisphere, where U does not fall to zero above the tropospheric jet maximum. Let us assume that τ is just slightly below its saturation value at a low level z_0 (corresponding to a low mountain, or the residual flux above the low level breaking zone) and compute where the secondary breaking level occurs if U increases with height. At one scale height, breaking can occur if $U/U(z_0)<e^{1/3}\approx 1.4$; at two scale heights we need $U/U(z_0)<e^{2/3}\approx 1.9$. Even with a fairly generous estimate of $U(z_0)=10$m/s, these conditions are unlikely to be met in the wintertime troposphere or lower stratosphere. This would suggest that convective breaking occurs predominantly at (a) low levels, and (b) levels which are higher than generally considered important in tropospheric GCM's.

It must be admitted however that the use of the WKB approximation in (2.4') is highly suspect. In the upper troposphere, with typical wind speeds of 40m/s, the vertical wavelength in WKB theory is about 40km, which is many times greater than the vertical scale of

variation of the wind field. It thus remains possible that the strong curvature above the jet maximum allows wave-breaking by limiting D to something comparable to the scale of the jet, which is much less than $2\pi U/N$. This issue cannot be addressed rigorously by means of any local argument. Nonlinear numerical simulations are required to properly determine whether breaking occurs, though steady linear theory may provide some useful qualitative insights. On the other hand, nonhydrostatic, rotation, and 3D effects all inhibit breaking by increasing dispersion, and it is not at all clear that a secondary convective breaking level can exist in the troposphere or lower stratosphere.

We note in passing that other kinds of breaking besides convective breaking may be possible. The alternate mechanism that has received the most attention is *shear breaking* , in which the wave induces a wind shear so great that the wave succumbs to local Kelvin-Helmholz instabilities. Shear breaking is usually ascertained by means of a local Richardson number criterion. It can easily be shown that when WKB is valid, convective breaking always occurs before the criterion for shear breaking is satisfied. Thus, shear breaking can occur only in the context of rapidly varying ambient flow, and in consequence its presence cannot be rigorously characterized in terms of a local criterion. Furthermore, when the local Richardson number criterion is satisfied without flow reversal (and hence convective breaking) the net velocity is everywhere positive, and it is likely that perturbations will be swept out of the limited unstable region before they have much chance to grow and do much damage to the wave structure. Despite these problematic aspects, shear breaking has been incorporated in some wave drag parameterizations (T. N. Palmer, personal communication) where it in fact plays a key role in permitting secondary breaking levels in the lower stratosphere. Much additional work is required to assess the physical reality of this process.

Concerning parameterization, the most unpleasant implication of the results described in this section is that the surface drag and low level flux cannot be determined independently of the existence and

location of the breaking level; the two are inextricably tied up via nonlinear processes of considerable subtlety. In the case of general U(z), N(z), very little has been firmly established about this problem. Certainly, for relatively uniform U,N, there is nonlinear "low level" drag enhancement, which is quantified in Figure 2.3. Although we refer to this flux absorption as a "low level" phenomenon, it should be noted that in some cases "low level" could extend right to the stratosphere; since the vertical scale in Fig. 2.2b is U/N, this could occur for a high mountain when there is a deep layer of large U (but not so large as to prevent breaking).

The nonlinearly enhanced low level drag has been left out of all parameterizations that have been attempted to date - what would its effect be in a GCM? Note that the physical effect of low level drag is to create an upstream blocked region and downstream stagnant mixed region; to the extent that envelope orography attempts to represent the same features, the employment of low level drag in conjunction with envelope terrain may be redundant.

3. Formulation of a parameterization

We shall now formulate a simple drag parameterization based on the considerations outlined in §2. Any parameterization of mountain wave drag must address two issues: how much momentum flux to launch into the atmosphere at the lower boundary, and how to distribute its absorption. As we have seen, these problems cannot really be addressed independently, because the presence of absorption (wave-breaking), and its location, greatly affects the net drag. This difficult problem will be finessed in the following treatment. The low-level wave breaking regime is accompanied by very large low level decelerations - amounting to perhaps 100 (m/s)/da; it is uncertain whether a large scale model can reproduce the correct physical response to such large localized momentum sinks. Hence, in the parameterization to follow, we attempt to represent only the residual momentum flux that escapes the low level breaking region and can affect the upper atmosphere. It will be seen that earlier parameterizations (Palmer *et al.* 1986) have in effect

accomplished the same thing, though based on very different physical reasoning. In the spirit of a sensitivity study, this decision is safe, as it almost inevitably results in an *underestimate* of wave drag effects on the atmosphere.

The wave drag due to subgrid scale momentum flux appears as an additional body force on the rhs of the momentum equations. The momentum equations become:

$$\partial_t u + \ldots = -\rho^{-1}\partial_z \tau_x = g\partial_p \tau_x \qquad (3.1a)$$

$$\partial_t v + \ldots = -\rho^{-1}\partial_z \tau_y = g\partial_p \tau_y \qquad (3.1b)$$

where the symbols represent:

- u zonal wind
- v meridional wind
- ρ density
- g acceleration of gravity
- z height
- p pressure
- τ_x flux of zonal momentum
- τ_y flux of meridional momentum

In terms of the subgrid scale velocities u',v' and w', the momentum fluxes are

$$\tau_x = (\Delta x\, \Delta y)^{-1} \iint dx\, dy\, \rho u'w' \qquad (3.2a)$$

$$\tau_y = (\Delta x\, \Delta y)^{-1} \iint dx\, dy\, \rho v'w' \qquad (3.2b)$$

where Δx and Δy are the grid increments and the integrals are to be taken over an area of a grid box.

Suppose that the waves are due to a two-dimensional ridge $h(x)$; we think of the ridge as being isolated (i.e. $h \to 0$ as $x \to \pm\infty$). The net vertical momentum transport over a region of length Δy, expressed in *dimensional* terms is

$$S_x = [\int_{-\infty}^{\infty} dx\ \rho u'w']\Delta y \qquad (3.3)$$

Note that this differs from the stress definition (3.2) in that it is not averaged over the grid area. Also, the stress is proportional to the length of the region in question because h is independent of y. The actual stress is given by

$$\tau_x = S_x/(\Delta x \Delta y) \qquad (3.4)$$

The flux is a function of the mountain height, mountain shape, the ambient flow profile, and altitude; clearly, some drastic approximations must be made if we are to make any progress. We begin by writing the flux as

$$\tau_x=(\tau_x/\tau_0)\tau_0\ , \qquad (3.5)$$

in which τ_0 is the momentum flux evaluated at a low level; when there is low level breaking, it is the residual flux measured just above the primary breaking layer. We shall refer to τ_0 as the *base level* flux. τ_0 is no longer a function of height, but its value could well be sensitive to the profile of U and N aloft. The nondimensionalized flux (τ_x/τ_0) contains the information about how the momentum is to be distributed in the vertical. Assuming no internal sources, its magnitude must be a monotonically decreasing function of height, and everywhere less than unity. Our ansatz in connecting the parameterization with the results of §2 is that the base level flux has the same behavior as in a model with constant N and U, where N and U are set equal to the average "low level" values prevailing in the atmosphere at the grid point in question. The precise definition of "low level" is, at this point, a matter of engineering.

Nondimensionalizing lengths by the mountain length L, depths by D=U/N, horizontal velocities by U and vertical velocities by UD/L, then the resulting *nondimensional* base flux is a function of Fr only (in the hydrostatic approximation). Thus the dimensional base flux due to a *single ridge* of length L and height h_0 can be written

$$\tau_0 = [\rho U^3/(N\Delta x)]\ G(Fr) \qquad (3.6)$$

where G(Fr) is some monotonically increasing function of Fr. Equation

(3.6) gives the stress in a grid box due to a single mountain of length L. If there are m such mountains in a grid box, then

$$\tau_x = m[\rho U^3/(N\Delta x)]\, G(Fr) \qquad (3.7)$$

where m is the number of mountains ($m<\Delta x/L$).
In fact, we can define the effective mountain length to be $\ell^*=\Delta x/m$.

Next we must determine G(Fr). For small Fr the wave field is linear and both u' and w' are proportional to the mountain height; hence, we must have

$$G=a\,(Fr)^2 \text{ for } Fr<<1 \qquad (3.8)$$

where a is an order unity constant determined by the mountain shape. On the other hand, the nonlinear results given in Figure 2.3 suggests that G saturates to an upper bound at large Fr. A plausible form for G is then

$$G= \tilde{\tau}Fr^2/(1+Fr^2) \qquad (3.9)$$

where $\tilde{\tau}$ is an order unity nondimensional saturation flux.

In the present formulation, the base flux is applied as if it were at the ground, and the flux convergence due to low level breaking is ignored. We note in passing, though, that a similar scaling argument can be used to parameterize the low level drag in the breaking case. The nondimensional flux above the breaking layer is still given by (3.9), but the true ground level flux based on Figure 2.3 would be

$$G'(Fr)= \begin{cases} G(Fr), & \text{for } Fr<Fr_c \\ 3.+5.(Fr-Fr_c)^2, & \text{for } Fr>Fr_c \end{cases} \qquad (3.10)$$

where $Fr_c\approx.8$ is the critical value for breaking. Between the ground and the top of the breaking layer, the flux would be relaxed from the value indicated by (3.10) to that indicated by (3.9); based on numerical simulations, a reasonable estimate for the top of the low level breaking zone is $(3/5)\pi U/N$.

The use of (3.9) has the effect of not allowing the height of the mountains contributing to the drag to exceed some (flow dependent) critical value. In this regard, it is similar to one of the schemes employed by Palmer, *et al.* (1986). Palmer, *et al.* argue for a cutoff height on the grounds that high mountains block the low layer flow so

that the atmosphere sees an effectively smaller mountain. This argument does not imply that there is any low-level drag that has been left out of the parameterization. However, for ridgelike mountains, there is little blocking in the range Fr<2, and total blocking does not appear at the ground until Fr≈2 (Pierrehumbert and Wyman 1985). For ridgelike mountains, then, the wavebreaking picture is on a firmer physical footing and moreover suggests a possible refinement of the parameterization. For axisymmetric or conelike mountains, though, flow does tend to be diverted around the obstacle for Fr>1 without any need for wavebreaking, and a blocking formulation may be more appropriate.

Table 1

Source	Nh/U	τ (Nt m^{-2})	τ (nondim)
Lilly 1982 (Rockies) (WAMFLEX avg., at 6 km)	2.[1]	.2	.7[1]
Cox 1986 (Pyrenees) (3/23/82, at 300 mb)	2. [2]	.4	1.[2]
Hoinka 1985 (Alps, 11/8/82)	2.5[3]	.3	1.2[3]
Linear *theory*, h=exp(-x^2)	2.	--	*3.8*
Steady, nonlinear hydrostatic *theory* (Long's model) h=exp(-x^2)	.5	--	*.25*
	.75	--	*.61*
	1.0	--	*1.2*
	2.0	--	*8.3*

[1] Based on N=1.25 x 10^{-2}, U=10m/s (similar to 600mb composite WAMFLEX upstream sounding) and h=2km.

[2] Based on values of N and U at 1km height in Toulouse sounding.

[3] Based on N=1.5 x 10^{-2}, U=15 m/s (similar to Milan sounding at 2km.)

Based on U=10m/s, N=.01sec^{-1}, ℓ^*=100km, and $\tilde{\tau}$=-1 the saturation flux indicated by (3.7) and (3.9) is 1 Pascal; if this flux is made to go to zero linearly in pressure over the depth of the atmosphere, the corresponding deceleration would be about 10(m/s)/da. How well do these numbers accord with observations? In Table 1 we compile some observed fluxes aloft in dimensional and nondimensional terms. The data is scanty; it does not represent a climatology as the observations were for the most part taken on days when there was some reason to suspect appreciable wave activity. Nevertheless, it is interesting that the fluxes are generally consistent with the (rather small) fluxes predicted by the parameterization. For comparison, we also give some fluxes predicted by steady linear and nonlinear theory neglecting the effects of wave-breaking. It is clear that considerable overestimates are obtained if the effects of low level breaking are not taken into account.

The choice of the vertical profile τ/τ_0 is problematic. In §2 we have seen that a local convective saturation criterion is unlikely to yield much breaking, and that the physical importance of shear breaking is questionable. In the experiments to be discussed in §5, we adopt the simple expedient of making the flux go to zero linearly in pressure. Insofar as we have a physical justification for doing this, it is that the strong curvature above the tropospheric jet may cause breaking aloft, and that the atmospheric response may not be fatally sensitive to the details of the vertical distribution. Its chief virtue, though, is as a base case with which other experiments can be compared; in the spirit of a sensitivity study, it also tells us whether fluxes of the indicated magnitude can conceivably yield an appreciable change in the circulation *supposing that* they are absorbed within the troposphere and lower stratosphere.

The above discussion has assumed that the mountain height h is a function of x alone, whence τ_y=0. We have little to say about the general case in which h=h(x,y) and the wind direction may turn with height, as nonlinear, three-dimensional mountain waves represent largely unexplored territory. In the experiments described in §5, we have assumed that all fluxes behave as in the unidirectional case, provided we

transform to a coordinate system aligned in the direction of the low level wind. This at least has the effect of imposing some retardation of largely meridional flow over largely zonal ridges, such as occur in the Himalayas. No attempt has been made to take into account the anisotropy of the terrain distribution.

4. Effect of momentum sink in idealized GCM

In order to assess the accuracy of wave drag parameterizations and identify their deficiencies, it is necessary to understand the manner in which a general circulation model responds to wave drag. In essence, wave drag parameterizations introduce momentum sinks above the rugged mountains of the Earth; intuitively, one expects the dominant effect of such sinks to be a deceleration of the westerlies. However, the factors determining the strength and distribution of the deceleration are subtle and various. The westerlies are maintained by a delicate balance between surface stress, mean meridional circulations, thermal forcing, and stationary and transient eddy fluxes. The introduction of a new momentum sink disrupts the balance, and the concomitant change in the zonal mean flow depends on the shifts in the terms in the balance.

Understanding the response of a realistic GCM to realistic wave drag is a formidable task, owing both to the complicated spatial distribution of the drag and the nonlinear interplay of the various processes in the model. For example, if the wave drag leads to an alteration in the zonal mean flow, then the Rossby wave trains emanating from the mountains will change; it would be difficult to distinguish anomalous wavetrains forced by this mechanism from wavetrains directly forced by the momentum sink. Thus, instead of attempting to diagnose the effects of realistic drag in a full GCM, we will look at the effects of an idealized momentum sink in an idealized GCM (IGCM).

The model used for these experiments is a simplified version of the GFDL Climate Group spectral GCM, with 15-wave rhomboidal truncation. Orography and all other zonally asymmetric boundary forcings (sea-surface temperature, inhomogeneous clouds, land-sea

contrast, etc.) were eliminated, and swamp boundary conditions were employed at the ground (inexhaustible moisture source, with surface temperature determined by local heat budget). The model was driven by annual mean solar forcing. A 1400 day control run was conducted in order to establish the (zonally symmetric) climatology of the base model. Next, a 1400 day "drag run" was conducted, in which a localized momentum sink was introduced in the vicinity of 33^0N, 90^0E. The momentum sink was chosen to have a Gaussian profile in the horizontal with 15^0 half-width in longitude and 5^0 half-width in latitude; the vertical profile of the drag force was chosen so as to correspond to a fixed deceleration of 10 (m/s)/day *independent of height.* A fixed momentum sink of this sort was employed in preference to a true parameterized drag dependent on surface wind in order to minimize the number of feedbacks in the model that need to be considered. The response of the model to the drag was evaluated by taking the difference between the drag run and the control run.

The response to the drag may be divided into a zonal-mean and zonally asymmetric component. Figure 4.1 shows the zonally asymmetric response of the (a)geopotential height and (b)meridional wind component, at 205mb. The response takes the form of a wave train emanating from the drag region, with a maximum amplitude of $Z'=-100$m and $v'=-6$ m/s in the immediate lee of the drag region. The response appears predominantly to the north of the drag maximum; the flow bends anticyclonically around the drag region, and there is a pronounced cyclone about 45^0 to the east.

Some aspects of this pattern can be explained in terms of the following barotropic β-plane argument. If $F(x,y)$ is the (negative) drag force appearing in the zonal momentum equation, the vorticity equation linearized about a constant wind U is

$$\partial_x(U\nabla^2\psi + \beta\psi) = -\partial_y F \qquad (4.1)$$

where ψ is the streamfunction. A localized drag region thus leads to a negative vorticity source to the north of the drag maximum and a positive vorticity source to the south; if F is symmetric in y, ψ will be antisymmetric. Supposing that the length scale of the drag region is small enough that the β term does not overwhelm the relative vorticity term in (4.1), to the north of the drag maximum the fluid will pick up

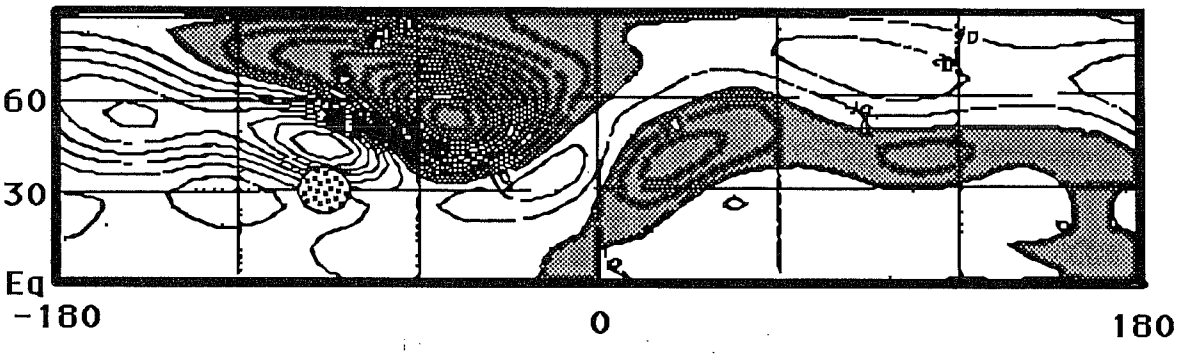

Figure 4.1a

IGCM Geopotential Height Response (zonal mean removed)

Contour interval is 10 meters. Shaded area is negative. Stippled disk gives location of drag region.

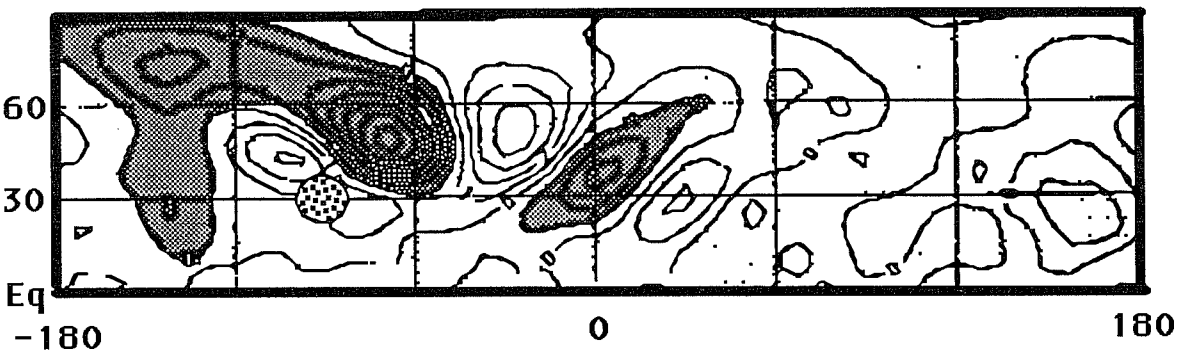

Figure 4.1b

IGCM Meridional Wind response (zonal mean removed)

Contour interval is 1m/s. Shaded area is negative. Stippled disk gives location of drag region.

negative vorticity and curve anticyclonically as it enters the drag region; the flow splits, and the streamlines suffer a net meridional displacement in traversing the drag region. This excites a free Rossby wave in the lee, giving rise to a trough roughly a half wavelength downstream. The resulting pattern is sketched in Figure 4.2.

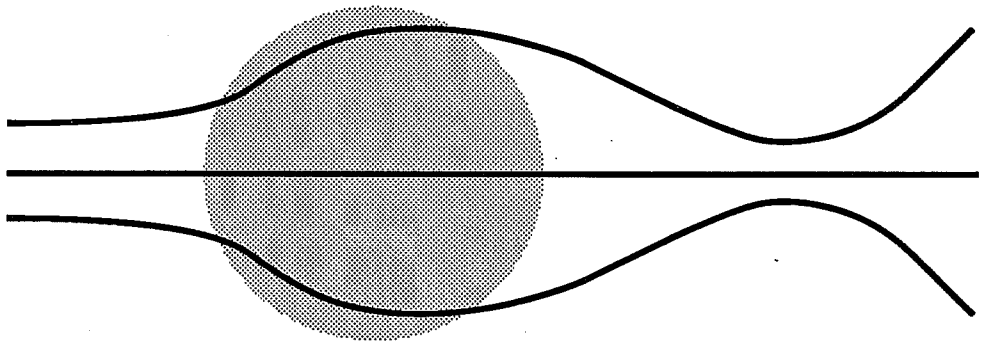

Figure 4.2
Schematic of Rossby Wave Driving by Localized Drag

(Actually, this description is greatly oversimplified; the linear equation (4.1) exhibits surprisingly subtle behavior when F has a nonzero zonal average, which we do not intend to discuss here). There are clear similarities between the northern parts of the patterns seen in Figure 4.2 and 4.1; however, the IGCM shows little response to the south of the drag maximum. This is perhaps connected with the fact that the zonal mean wind is *easterly* at low levels equatorward of 30^0, so that only the drag on the northern flank sees a westerly current throughout the depth of the model. Absorptive or other critical level effects may inhibit the response to the south.

The properties of wavetrains such as described above can be understood in great detail in terms of linear Rossby wave theory, and so we will not dwell further on this part of the response. In fact, some preliminary calculations with a linear primitive equation model on the sphere indicate that the patterns can be reproduced quite accurately (both qualitatively and quantitatively) within linear theory.

Understanding the response of the zonal mean winds is a far more difficult problem. Yet, the structure of the zonal mean wind is

critical in a full GCM, as Rossby wave paths are very sensitive to the wind field. Hence, an accurate zonal wind climatology is essential if the model is to correctly simulate response to forcing from mountains or sea surface temperature anomalies. In addition, the strength of the westerlies affects the propagation and steering of synoptic eddies. The zonal mean zonal wind response (drag run minus control run) is shown in Figure 4.3. Within a broad region in the vicinity of the drag maximum, the shift is negative, corresponding to a weakening of the westerlies. It is highly significant that the effect becomes more pronounced with height, even though the imposed deceleration was independent of height; we will see in §5 that this is precisely the vertical structure needed to

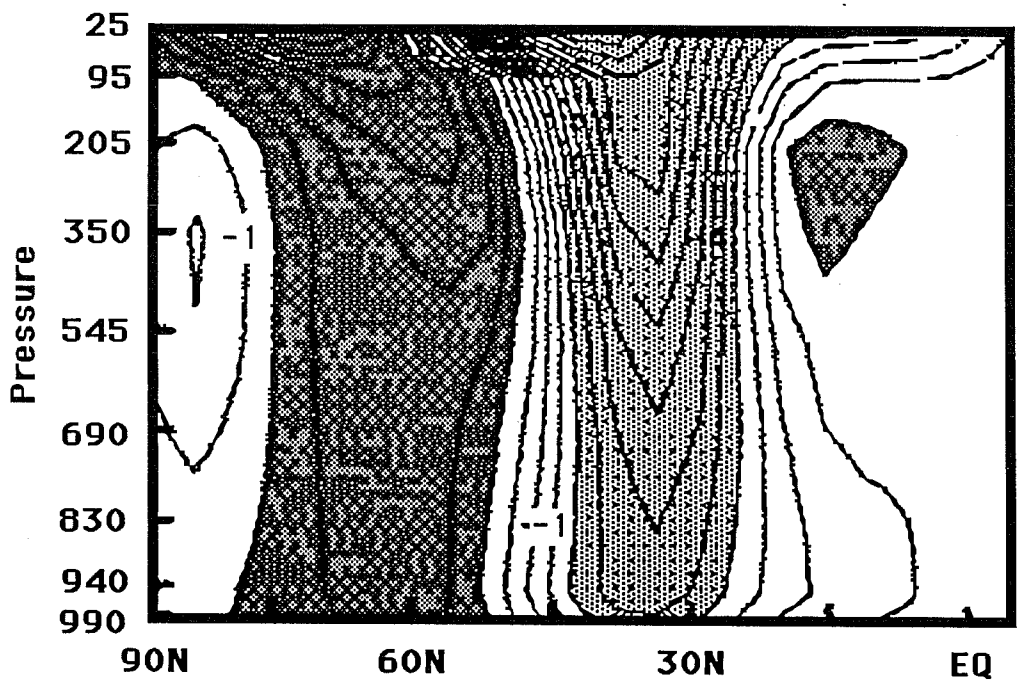

Figure 4.3
Response of Zonal Mean Zonal Wind to Drag

Contours of zonal mean zonal wind for (drag run)-(control run). Contour interval is .5m/s. Lightly shaded regions are <-2m/s, and darkly shaded regions are >0 (i.e. accelerated).

cancel the systematic error in a full GCM. The decelerated region extends well beyond the range of latitudes where appreciable drag is applied; this phenomenon will be addressed shortly. To the north of the decelerated region, there is a region of *accelerated* westerlies. There is very little response in the Southern Hemisphere.

Some insight into the zonal mean response can be obtained from an examination of the vertically integrated angular momentum budget. Schematically, the budget may be written

(Drag force)-(Convergence of transports)=(Surface stress)

Since the surface stress is a monotonic function of surface wind speed, the imbalance on the left hand side determines the surface wind. The transport term may be further subdivided into transports due to (a) the mean meridional circulation, (b) the stationary eddies, and (c) the transient eddies. We examined the *changes in* all these terms upon introduction of drag (defined as drag run minus control run). It was found that the change in the transport term was strongly dominated by the transient eddy angular momentum flux. The resulting balance, as a function of latitude, is shown in Figure 4.4. Near the drag region the surface stress change is positive, reflecting the fact that the drag weakens the westerlies in this region. The transient eddy term is comparable in magnitude to the explicit drag, and in fact cancels out more than half of the peak value of the drag; on the flanks of the drag region, the transient eddy term reverses sign, tending to enhance the local deceleration. The net effect of the eddies is thus *diffusive*, reducing the peak of the effective drag while spreading out the region which experiences drag. Linearizing the surface stress, the surface wind response U' can then be represented by an equation of the form

$$\kappa d^2U'/dy^2 + CU' = F \qquad (4.2)$$

where F is the vertically integrated drag force, C is the surface drag coefficient, and κ is a (model-dependent) effective diffusion.

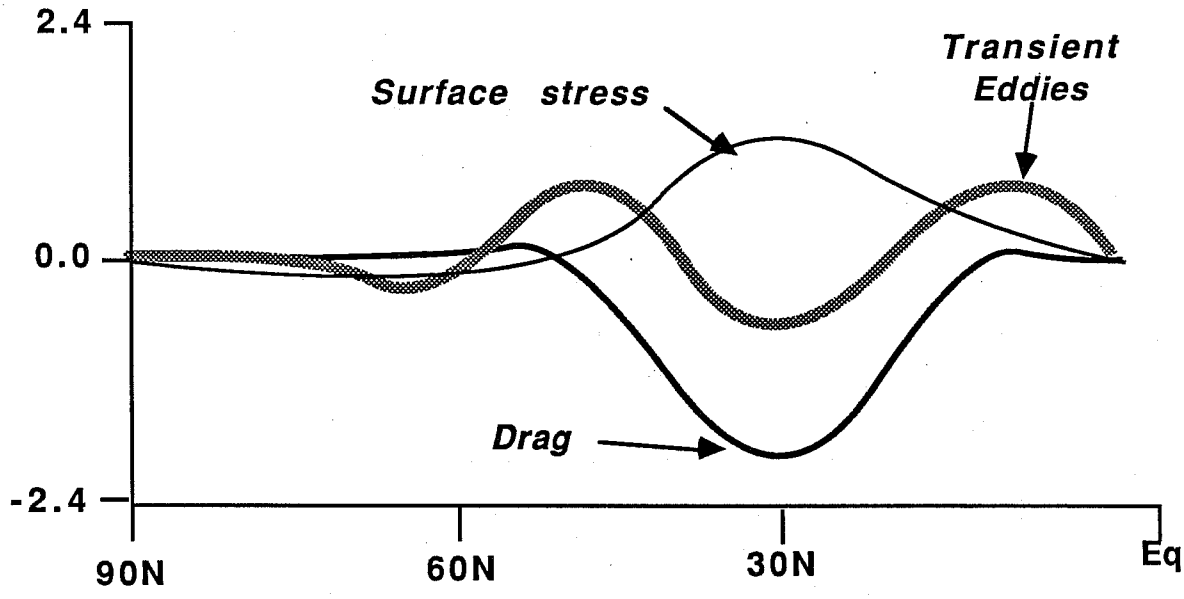

Figure 4.4
Anomaly Angular momentum budget for IGCM drag run

This figure gives the changes (Drag run - Control run) in the dominant terms in the angular momentum budget. Curve marked "transient eddies" gives change in transient eddy angular momentum flux convergence; Curve marked "Drag" is the imposed momentum sink. The sum of the drag and the surface stress term approximately balances the transient eddy term. Units are 10^{-6} Nt/m.

This observation by no means accounts for all the important features of the zonal wind response. In particular, the weak acceleration to the north of the drag region, the vertical structure of the response, and especially its sharp peak near the top of the model, have not yet been explained. Nonetheless, the angular momentum budget suggests that the nonlinear eddy fluxes are likely to be critical to the understanding of these features.

5. Tests in a prediction model

In this section we provide an indication of the effect of the drag parameterization of §3 in the GFDL global spectral model described in Gordon & Stern (1982). In implementing the scheme, N and U were taken from an average of the lowest three model layers, the mountain height was estimated from subgrid orographic variance (based on $(1/6)^0$ data,

and we set ℓ^*=100km and $\tilde{\tau}$=-1. Wave drag and control-model 30-day forecasts were carried out for four January cases; each case was run with R30L9 and R42L18 resolution. The systematic error was estimated by taking the ensemble average of the time-averaged day 20-30 forecast error over the four cases. This is a rather modest series of experiments, but given the strength of the signal there is some hope that useful insights can be obtained by analyzing the results.

Anomaly correlation coefficients for the 10-day mean 500mb height forecasts (averaged over all four cases) in the R30 model are shown in Figure 5.1. It is clear that the wave drag has a substantial beneficial impact on the accuracy of long-range forecasts. The improvement persists throughout the time period, and in the wave drag run the correlation coefficient asymptotes at a time when that of the control run is still dropping precipitously. This probably reflects the arrest of the climate drift by the wave drag, about which more will be said shortly. Insofar as the results presented in Figure 5.1 are representative, they confirm the beneficial impact of wave drag found at medium-range and climate time scales in Palmer *et al.* (1986). The results also indicate that the miscellaneous differences between our wave drag scheme and that of Palmer *et al.* have not destroyed the favorable effects. As a further comparison, we also include in Figure 5.1 results for a run including both wave drag and $\sqrt{2}\ \sigma$ envelope orography. For the medium range (roughly the first 10 days) the combination produces a clearly superior forecast; after that, the forecast deteriorates precipitously, until the combination becomes far worse than wave drag alone. To some extent, the degradation comes from an excessive buildup of planetary wave amplitude in the Pacific, suggesting that deep wave drag and envelope orography may duplicate the same physical influence in the forcing of planetary waves.

In Figure 5.2 we show the systematic error of the zonal-mean zonal wind for the R42L18 model. The control case is shown in Figure 5.2a, while the wave drag case is shown in Figure 5.2b. The error pattern for the control case is dominated by excessive strength of the midlatitude westerlies, with the magnitude of the error increasing with height. The deep wave drag response from the idealized GCM (Figure 4.3)

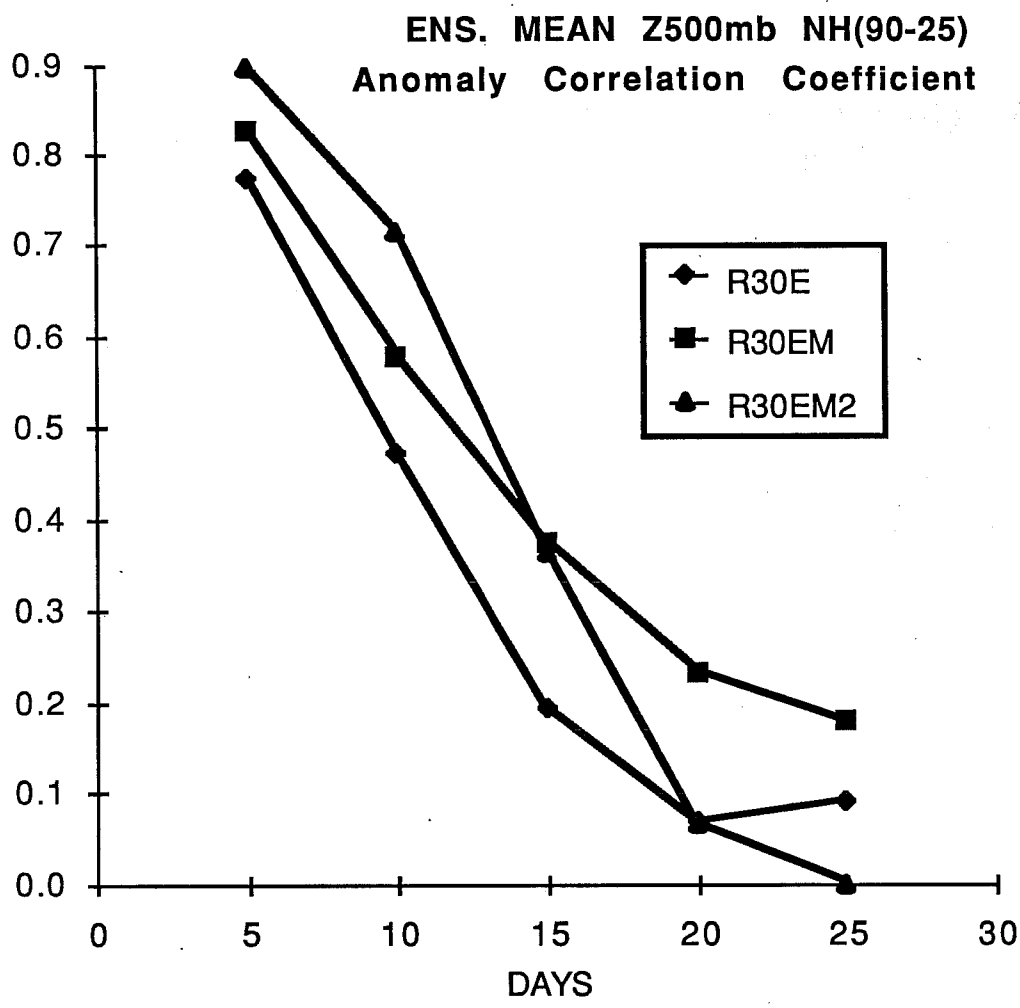

Figure 5.1

Anomaly correlation coefficients for R30 forecasts.

Correlation coefficients are for 10-day mean fields. Models used are: Control model (R30E), Wave drag model (R30EM), Wave drag + Envelope orography (R30EM2).

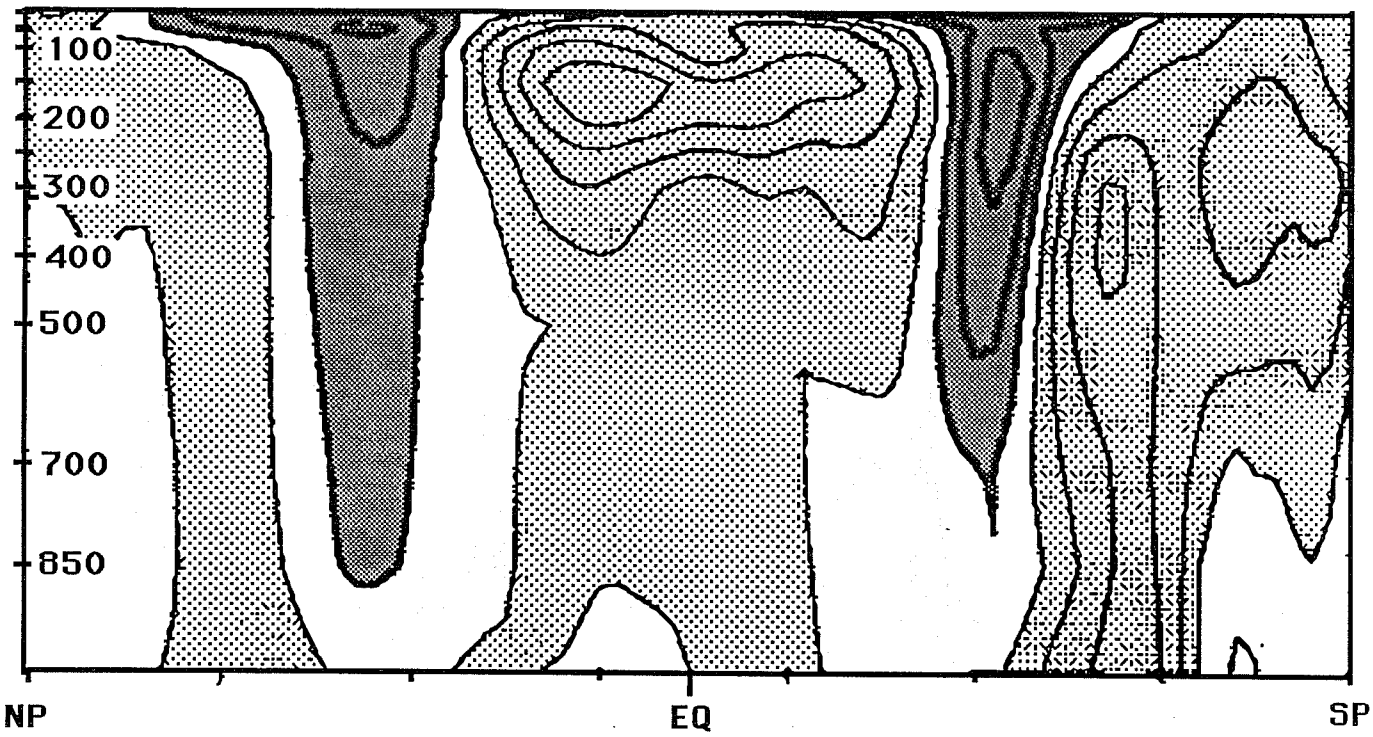

Figure 5.2a
R42 Control Run Zonal Mean Wind Systematic Error
Contour interval 2m/s. Lightly shaded areas are negative, darkly shaded>2m/s.

has precisely the structure needed to remove this error. Indeed, the northern hemisphere zonal wind error has practically vanished in the wave drag run (Figure 5.2b).

However, the systematic error analysis also casts considerable doubt as to the physical reality of the orographic wave drag effect. Specifically, the signature of the midlatitude wind error is very similar in the Northern and Southern hemispheres, even though there is radically less orography in the latter. It is evident from Figure 5.2 and from the IGCM results that orographic wave drag has little impact on the Southern Hemisphere circulation. Thus, the orographic wave drag is very good at removing in the Northern Hemisphere a component of the error that is relatively symmetrical between the hemispheres, and which therefore is unlikely to be due in reality to orographic effects.

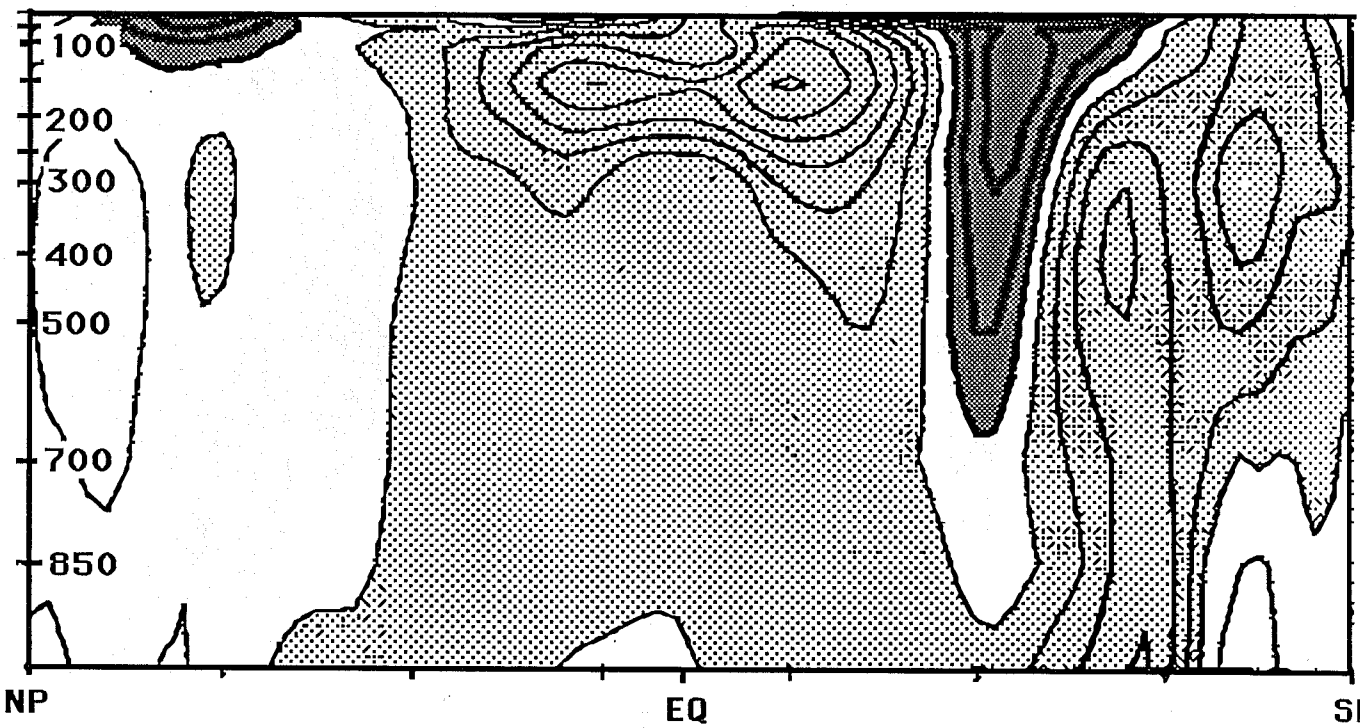

Figure 5.2b
R42 Wave Drag Run Zonal Mean Wind Systematic Error
Plotting conventions as for Figure 5.2a

This result is based on a small number of cases, and at this point provides no more than a cautionary note. Some confidence is lent to the results by the fact that Tibaldi (1986) found a similarly symmetric zonal wind systematic error for a large ensemble of 10-day forecasts. If the result holds up in the face of further investigation, it would suggest that the missing drag is due to some internal momentum transport mechanism other than orographic gravity wave drag, and that deep orographic wave drag exerts a spurious (though beneficial) effect on the westerlies aloft which should in reality be attributed to other physical causes.

6. Conclusions

Elementary considerations regarding the stratified flow of a

fluid over an obstacle imply that in typical circumstances gravity waves generated by air flow over rugged terrain will be in the strongly nonlinear range. One manifestation of this nonlinearity is the presence of a low level wave breaking zone; the onset of low level breaking is accompanied by greatly amplified drag, which is typically felt at low levels. Some momentum flux escapes the low level breaking zone and may be deposited elsewhere in the atmosphere, though it is not at this point clear that there can be a secondary breaking level in the lower stratosphere. The nonlinear phenomena known to occur in the simplest stratified flow problems pose the following questions for those of us seeking to parameterize wave drag effects: (1) Should low level nonlinear drag be parameterized? What are its effects in models? To what extent does it reproduce envelope orography? (2) How can the nonlinear amplification of drag associated with wave breaking be parameterized? What is the dependence of this effect on the height of the breaking level? (3) Can shear breaking occur? How efficient is it at removing momentum? How can the shear breaking level be determined? (4) Is there in reality a secondary breaking level in the lower stratosphere?

Tests with an idealized GCM have shown that momentum flux of a magnitude consistent with that escaping from the low level breaking layer (and also consistent with observed fluxes aloft) can exert a substantial influence on the zonal mean wind, provided that the momentum is distributed throughout the troposphere and lower stratosphere. The dominant effect of the drag is to slow down the westerlies within a broad band centered on the drag region. An examination of the angular momentum balance in the drag run shows that the shifts in surface wind are determined primarily by a balance between surface stress, wave drag, and transient eddy momentum flux; the transient eddy fluxes are predominantly diffusive in nature.

Besides its effects on the zonal mean wind, localized wave drag excites a pronounced planetary wave. The wave train resembles that found in linear theory and its dominant feature is a trough just downstream of the drag region. Thus, orographic wave drag can act to

directly enhance planetary wave amplitude, in addition to any effects mediated by the changes in the zonal mean wind. It has not yet been established whether this direct forcing effect has a beneficial impact on full GCM's.

Tests of upper-level residual wave drag in a series of 30-day forecasts with a full GCM confirm earlier results of beneficial impact of such drag, particularly with regard to systematic error. A careful examination of zonal mean zonal wind errors, however, indicates that wave drag eliminates *in the Northern Hemisphere alone* a component of error that is quite symmetric between the hemispheres. These admittedly tentative results cast some doubt on the reality of orographic wave drag as the source of the missing drag in the models.

Although incorporation of orographic gravity wave drag has yielded remarkable amelioration of the formerly resistant problem of climate drift in general circulation models, a vast number of mathematical and physical problems need to be surmounted before it can be said that the drag represents a real effect. Considering where to go from where we are now leads us into very deep questions concerning the fundamental nature of stratified flow over obstacles, and the effect of internal drag on large scale circulations. Whatever the eventual outcome, an exciting new chapter has been opened in the study of large scale effects of mesoscale phenomena.

Acknowledgements

Many people have contributed generously to the results presented above, and their assistance is gratefully acknowledged. Julio Bacmeister carried out the computations of nonlinear mountain waves and their fluxes discussed in §2. The IGCM simulations described in §4 were carried out by Thomas Knutsen. William Stern implemented the wave drag scheme in the GFDL prediction model, and carried out the experiments described in §5. We had the benefit of numerous interactions with Bertrand Carissimo on the subject of observed momentum fluxes, and with Bruce Wyman on the subject of the ALPEX data and of behavior of momentum flux in nonlinear models. I alone am,

of course, responsible for all errors and dubious speculations that remain in this manuscript.

References

Bacmeister, J. 1986: Wave breaking, upstream influence and downslope winds. *PhD Thesis,* Geophysical Fluid Dynamics Program, Princeton University.

Boer, G.J., McFarlane, N.A., Laprise, R., Henderson, J.D. and Blanchet, J.-P., 1984: The Canadian Climate Centre Spectral Atmospheric General Circulation Model. *Atmosphere-Ocean,* **22**,397-429.

Bretherton, F.P., 1969: Momentum transport by gravity waves. *Quart. J. R. Meteorol Soc* **95**, 785-803.

Cox, K. 1986: Analysis of the Pyrenees lee wave event of 23 March 1982. *Mon. Wea. Rev.* **114**,1146-1166.

Gordon, C.T. and Stern, W.F. 1982: A description of the GFDL global spectral model. *Mon. Wea. Rev.* **110**,625-644

Hoinka, K. P. 1985: Observation of the airflow over the Alps during a foehn event. *Quart. J. R. Meteorol. Soc.* **111**, 199-224.

Lilly, D.K. 1972: Wave momentum flux. A GARP problem. *Bull Am Meteorol Soc* **20**, 17.

Lilly, D.K. Nicholls, J.M, Chervin, R.M., Kennedy, P.J. and Klemp, J.B., 1982: Aircraft measurements of wave momentum flux over the Colorado Rocky Mountains. *Quart. J. R. Meteorol Soc,* **108**, 625-641.

Lindzen, R.S. 1981: Turbulence and stress owing to gravity wave and

tidal breakdown. *J. Geophys. Res.* **86**, 9707-9714.

Palmer, T.N., Shutts, G.J. and Swinbank, R.,1986: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parameterization. *Q J R Meteorol Soc*, in press.

Peltier, W.R. and Clark, T.L., 1979: The evolution and stability of finite amplitude mountain waves-II: Surface drag and severe downslope windstorms. *J Atmos Sci* **36**, 1498-1529.

Pierrehumbert, R.T. and Wyman, B., 1985: Upstream effects of mesoscale mountains. *J. Atmos. Sci.*, **42**,977-1003..

Smith, R.B.1979: The influence of mountains on the atmosphere. *Advances in Geophysics*, **21**, Academic Press, 87-230.

Smith, R.B. 1985: On severe downslope winds. *J. Atmos. Sci.* **42**, 2597-2603.

Tibaldi, S. 1986: Envelope orography and maintenance of the quasi-stationary circulation in the ECMWF global models. *Advances in Geophysics*, **29**, Academic Press, 339-372.