Numerical Weather Prediction

Convective-dynamical feedbacks and their implications for process modelling and NWP

Forecasting Research Technical Report No. 397

S.H.Derbyshire and S.J.Woolnough[1]

1; now at the University of Reading, Reading, UK

email: nwp_publications@metoffice.com

Met Office NWP Technical Report no.397

Convective-dynamical feedbacks and their implications for process modelling and NWP

S.H.Derbyshire and S.J.Woolnough*
Atmospheric Processes and Parametrizations
Met Office

14th January 2003

*now at the University of Reading

Abstract

The significance for convection modelling of systematic feedbacks between convection and larger-scale dynamics is increasingly recognized. Here we consider the representation of such feedbacks in idealized process studies, such as comparisons of convection parametrizations with cloud-resolving models. We seek to formulate a testbed which can include a large-scale response, thus allowing convection greater freedom to change the precipitation than in conventional single-column studies.

Various authors have proposed simple convective-dynamical feedback models, but there remains a need to benchmark these models more thoroughly against specific dynamical problems. Here therefore we review the dynamical solutions to a number of dynamical test problems, and evaluate (i) the fraction of diabatic heating compensated by large-scale advection and (ii) the effective timescale for dynamical relaxation of buoyancy anomalies.

Dynamical feedback parametrizations can be formulated based on either of these quantities. Which one is more nearly constant depends on whether the problem is balanced or unbalanced, and whether it is two- or three-dimensional.

As applications of this theory, we consider some specific process-issues in convection where allowance for large-scale feedback gives additional insight. These issues include the sensitivity of convection to humidity, the development of convection over tropical islands and the application of latent-heat nudging in operational mesoscale data assimilation. We offer a theoretical explanation for the problem of gridpoint storms in large-scale models, and why it matters whether convection is resolved or parametrized.

Contents

1 Introduction

This note concerns the feedbacks between moist atmospheric convection and large-scale flows, and their significance for atmospheric modelling on small or large scales.

Convective cloud systems interact strongly with many other processes (including surface, boundary layer, upper-air and large-scale phenomena). Such systems and their statistical parametrization are now being studied in increasing detail using finescale process-research models. E.g. Swann (2001) shows how cloud-resolving models (CRMs) can be used to look in detail at the assumptions of a mass-flux convection scheme. In other stud-

ies, cloud- resolving models are compared with single-column versions of NWP or GCM models as a test of the physical parametrizations.

Single-column (SCM) studies have become a standard tool for assessment of parametrizations (Randall, Xu, Somerville & Iacobellis 1996), but a SCM inherently lacks the dynamical response of a full GCM or NWP.

Even with current supercomputers, no single model can fully capture all the scales which are relevant to convection and its dynamical issues. Although it is now sometimes possible with massively-parallel computation to run coarse cloud-resolving resolution (e.g. 2km) on a planetary scale, Grabowski (2001), many current process issues in convection call for resolution refinement towards 100m (Petch & Gray 2001). Motivated by the wide scale separation, our aim here is to develop a method which allows computational resources to be targeted at the small-scale process issues, but still captures some of the essential large-scale dynamical context, insofar as it affects those process issues.

James (1994), p.93, states that the aim of scientific modelling is to distinguish between 'crucial' and 'incidental' mechanisms. Convective-dynamical feedback is potentially a crucial mechanism, even in idealized modelling, in regimes where the feedback can fundamentally change the behaviour of the system. More specifically, such feedback allows various physical effects to have long-term impact on convective precipitation, whereas if large-scale temperature forcing is prescribed then convective precipitation is effectively controlled through the latent heat equation. So the presence of feedback can test aspects of the convection scheme's behaviour which are not fully tested in conventional SCM process studies.

In the present note we seek therefore to outline a convective-dynamical testbed suitable for convective process modelling and idealized SCM/CRM comparisons, taking account of the existing convective-dynamical literature and looking critically at the underlying assumptions. We shall also comment on some points arising directly in NWP. First, though, we describe what we mean by process studies and why the issue of dynamical feedback arises.

2 Process studies and 'forcing' specification

Here a process study is defined as an investigation (by observation, theory and/or modelling) of atmospheric processes which are subgrid relative to a large-scale model, of resolution O(100km). For example the ARM-SGP site (Zhang, Lin, Cederwall, Yio & Xie 2001) has an intensive array giving cloud and other information surrounded by other sites which give information required for lateral BCs on the O(100km) scale. A process study may include SCMs or CRMs, and indeed may involve explicit intercomparison of SCMs and CRMs under the same setup. We shall assume further that in such process studies the results from the CRMs are to be interpreted statistically rather than deterministically.

Process modelling studies may be either idealized (tackling the simplest, most generic problems) or 'case studies'. The case studies provide a 'reality check' on the process models, whilst the idealized studies can evaluate specific parameter dependences. In practice there are also intermediate possibilities, such as 'semi-idealized' case studies in which a case study is progressively idealized and/or perturbed in order to test sensitivities.

As already hinted, such process studies require boundary conditions (BCs), both surface and lateral. There is a choice not only of the values but of the *type* of BC, i.e. which variables are prescribed, and which variables may adjust during the run.

Consider first, for illustration, the specification of surface BCs. Surface BCs may be specified either via prescribed surface scalar-values (SSV), or via prescribed surface scalar fluxes (SSF). With suitable choice of those values or fluxes, the two specifications might be considered almost equivalent. Nevertheless any model-model or model-observational comparisons will in general show different sensitivities because different quantities are held fixed. E.g. if a model is sensitive to a quantity λ (say, surface heat flux) then running a process model with fixed λ may fail to anticipate behaviour in a more complete model where λ can change. Idealizations are legitimate but need to be complemented by running alternative scenarios.

Turning now to the choice of lateral BC for process studies, similar general considerations apply. In convective process studies the lateral BCs are commonly idealized as horizontally cyclic with the addition of horizontally averaged forcing terms derived from a large-scale budget. This approach (which may be called a pseudo-lateral BC) fits naturally with a parametrization of the convective process as a statistical ensemble. (There are other approaches, based on open BC or nesting, but these seem more relevant for deterministic prediction of a convective system.)

The limitation of 'prescribed temperature forcing' as a pseudo-lateral BC for convective process studies is simply that it forces the convective heating (closely linked to other measures of convective activity, such as mass flux) to balance the temperature forcing, and thus makes quasi-steady convection insensitive to other mechanisms. The role of those other mechanisms can then only be inferred in a very indirect and limited way, leading to a risk that their importance to large-scale modelling may be hidden by our methodology.

2.1 Forcing or feedback in large-scale and process models

Convective-dynamical forcing or feedback is handled in a variety of ways in the existing literature, e.g.:

(i) Dynamical methods for computing the response to prescribed 'convective forcing' using linear analysis or idealized GCMs. E.g. Rodwell & Hoskins (1996) show that latent heating in the South Asian monsoon drives a very broad-scale circulation (including significant compensating descent over the Mediterranean and the Sahara desert), and that a linearized model can capture much of this effect.

(ii) A further literature where simple convection parametrization is incorporated into an essentially dynamical model of large-scale waves or instabilities, e.g. the model of the El Nino-Southern Oscillation phenomenon due to Zebiak & Cane (1987).

(iii) Process studies which usually treat 'large-scale forcing' as fixed for comparing parametrizations, process models and observations. E.g. Guichard, Redelsperger & Lafore (1996) show significant impacts of large-scale forcing on cloud-resolving model, but do not attempt to model interactively the response of large-scale ascent to convection.

(iv) An increasing literature on 'two-column' or related methods which assume some simple feedback between local convection and its environment (e.g. Raymond & Zeng (2000), Sobel & Bretherton (2000))

Here we wish to strengthen the connections between different approaches by selecting appropriate feedback methods for process studies and tracing their connection with the underlying dynamical equations using benchmark problems.

3 Benchmark problems

3.1 Scalar advection

We consider scalar advection, principally to establish to what extent it can be approximated using vertical motion and vertical gradients. In particular we consider the implications for (i) buoyancy (ii) humidity and (iii) hydrometeors.

Without loss of generality we restrict the analysis in this section to a 2-D problem, suppressing one of the horizontal dimensions (the generalization to 3-D is trivial). We also use the incompressible Boussinesq equations (rather than pressure coordinates as elsewhere in this note), as the scaling is slightly easier to understand when vertical scales are expressed in the same units as horizontal scales, and the scaling results in fact carry over straightforwardly.

We therefore write the advective term for a general scalar ϕ as

$$\mathbf{u}.\nabla\phi = u\partial_x\phi + w\partial_z\phi \tag{1}$$

and evaluate the relative contributions of the two terms.

Note first the special case of cancellation when $\partial_x\phi/\partial_z\phi = -w/u$, i.e. when isolines of ϕ are parallel to streamlines. It is easily seen then that the 'vertical gradient term' $w\partial_z\phi$ will dominate when streamlines are steeper than scalar isolines (or more generally isosurfaces).

We refer to $w\partial_z\phi$ as 'vertical gradient term' rather than 'vertical advection' because (at least in its impact on the column budget) it may be related to horizontal convergence via

$$\int w\,\partial_z\phi\,dz = -\int \phi\,\partial_z w\,dz = \int \phi\,\partial_x u\,dz \tag{2}$$

assuming $w = 0$ at the top and bottom of the integration.

Whatever the nomenclature, it is very useful in process studies to be able to approximate the large-scale advection (or at least that part of it which participates in feedback with the local convection) by $w\partial_z\phi$ since we then need to know only one 'large-scale' variable w (averaged over some appropriate scale), which we can then estimate using buoyancy arguments.

3.1.1 Cases with horizontal symmetry

A special case of scalar advection with horizontal symmetry is worth considering separately, e.g. as a model for dynamical response to localized convection around (say) $x = 0$. This case is not straightforwardly covered by the above, because both u and $\partial_x \phi$ are small.

Consider e.g. a streamfunction $\psi = \sin mz \sin kx$ so that $u = -m \cos mz \sin kx$ and $w = k \sin mz \cos kx$ (over a layer from $z = 0$ to π/m) and suppose that near $x = 0$ we can approximate

$$\phi \sim az + bx^2 \tag{3}$$

Then (for small kx)

$$(u\partial_x \phi)/(w\partial_z \phi) \sim (mbx^2/ka) \cot mz \tan kx \sim (mbx^2/a) \cot mz \tag{4}$$

whence if we treat $\cot mz = O(1)$ we see that the ratio of the horizontal gradient term to the vertical gradient term scales with the ratio of the variation in ϕ over the characteristic vertical scale m^{-1} to the variation in ϕ over the characteristic horizontal scale for x.

The factor $\cot mz$ indicates that horizontal terms will become relatively more important at the top and bottom of the convecting layer. This may alert us to the possibility of special treatment for the boundary layer. In fact the horizontal terms are not formally large there, but the vertical terms are small.

3.1.2 Buoyancy, humidity and hydrometeors

Let us now assume further that the horizontal variation arises from a convective source in a region of horizontal dimensions L_x, with a vertical scale π/m $\sim$10km. We estimate the ratio of vertical to horizontal terms for three important scalars. For present scaling purposes we may regard buoyancy as measured by potential temperature θ, although strictly we should look at θ_v, with correction terms due to water vapour and hydrometeors. For convenience we again use the model with horizontal symmetry.

For θ, the vertical variation scale a/m $\sim$10K, a value which is very large by the standards of horizontal temperature variation in the tropics. We shall show later in §3 that the large-scale dynamics strongly resist θ-variation on scales less than an outer dynamical scale (k_r^{-1}, for future reference). So for current purposes we can assume that the vertical gradient term dominates for θ.

For humidity q the calculation is slightly more complicated because q profiles have a more nonlinear shape. Typically (at least in the lower troposphere, say from 1-5km, including most of the column vapour) there is a quasi-exponential decay with height over a vertical scale $L_q \simeq$3km. This implies that $a/m \simeq q$ and so the ratio of horizontal to vertical gradient terms will scale simply on the horizontal fractional variation in q over the scale L_x. Here the dynamics do not significantly resist q-variation and so substantial variations are possible.

For hydrometeors e.g. cloud ice q_i, the picture is different again. We may expect cloud ice in particular (or that part of it which persists for long enough to interact with large-scale advection) to be concentrated in relatively thin layers, say 1km thick. We may then have $a/m \gg 1$, in which case the vertical term will dominate even if there is O(1) variation in q_i over the scale L_x.

3.1.3 Significance of scalar advection

In summary, for present purposes we can relate buoyancy advection straightforwardly to vertical velocity (a very useful simplification).

For humidity the horizontal term may be significant quantitatively, but in many problems this corresponds to a 'nonlinear stage' in which the convecting region has greatly distorted the initial basic state. It may still be acceptable to neglect this term in process model comparisons, provided that the neglect is carried consistently through the comparison (e.g. for both CRMs and SCMs).

For hydrometeors in thin layers the vertical term is again likely to dominate. But in fact the net hydrometeor removal (implied outflow) is best computed from the convergence form of the budget shown above, from which the characteristic removal rate is $\partial w/\partial z$. In effect the hydrometeors are ventilated by horizontal outflow on the same timescale on which the column is 'vertically ventilated' by upward motion (e.g. for $w = 0.1\text{ms}^{-1}$ over 10km depth this gives a timescale 10^4s), because continuity requires the mass transport timescales to match. As discussed above, this ventilation can be represented via the vertical term.

These arguments show that the inclusion of hydrometeor advection in process study setups is formally desirable, and that it can be done relatively straightforwardly without additional adjustable parameters. The significance of such hydrometeor advection de-

pends on their characteristic lifetimes within the cloud system. Probably only cloud ice lives long enough for such advection to be significant. Support for the significance of such terms may be drawn from the tendency of some 1D climate models of tropical convection to generate excessive cirrus in a manner that is unrepresentative of the 3D climate model (J.Mitchell, personal communication).

In fact the 'symmetric' (self-generated) ventilation discussed here is not the only type. There may be additional horizontal advection by the background wind (if any), i.e. by modes superimposed on any self-generated component. These background winds may not couple dynamically but may lead to further exchange with the background at a rate scaling on U/L_x.

That the vertical gradient term follows directly from knowledge of w (without further dependence on L_x) is perhaps initially surprising, but a consequence of continuity. Of course the dynamics which determine w will be influenced by L_x, as we now discuss.

3.2 Dynamical benchmarks

Our main task now is to determine large-scale w from quantities which we can evaluate in a process study. Dynamical studies such as Rodwell & Hoskins (1996) broadly support the treatment of dynamical response as a linear function of convective forcing, at least to the level of accuracy needed for our purposes. We can therefore characterize the key feedback on the buoyancy equation using either a dynamical adjustment timescale t_{dyn} or an effective fractional area r. Either of these quantities could potentially be used to parametrize ω.

The variable $\partial_p \Phi'$ (where Φ' is the deviation of Φ from some constant reference state) is a measure of negative buoyancy anomaly or negative thickness anomaly. Using the notation and derivations of the Annex we have the following equation:

$$\partial_t \partial_p \Phi' = -s^2 \omega - B \tag{5}$$

whence we may define a dynamical timescale for buoyancy adjustment as

$$t_{dyn} \equiv \partial_p \Phi' / s^2 \omega \tag{6}$$

and an effective fractional area for the convecting region as

$$r \equiv (B + s^2 \omega)/B \tag{7}$$

i.e. the fraction of the convective heating which is *not* compensated by the dynamical feedback. For if the dynamics acted effectively to distribute convective heating evenly over a box, from a convective source with fractional area r, then the net warming in the source-region would be a fraction r of the source.

The parameters r and t_{dyn} are related by the equation

$$t_{dyn} = (-B/s^2\omega - 1)/\sigma = r/\sigma(1-r) \tag{8}$$

where σ^{-1} is the forcing timescale. Although we have written this in terms of an e-folding timescale, a general relationship of this kind follows directly from time-integration of (5) without any assumptions about either the convection or the dynamics.

Note that for long timescales σ^{-1}, either r must become small or t_{dyn} must become large. The case where r tends to a nonzero limiting value as $\sigma \to 0$ may be referred to as the 'balanced' case, as it is consistent with the existence of an asymptotically quasi-steady spatial structure. In contrast, problems where $r \to 0$ as $\sigma \to 0$ may be described as 'unbalanced', as this suggests that on long timescales the influence of a given region of convection is propagating without limit.

In fact a number of forcing timescales are of interest here, including the diurnal development of convection and also the quasi-steady limit which corresponds to the 'DC component'. E.g. the essentially single-signed nature of convective heating might be modelled by a dependence of the form $(1+\sin\nu t)(1+\sin kx)$. The component which is homogeneous in space does not, however, drive any large-scale circulation and may therefore be assumed to have been removed from our problem.

We now seek to evaluate the timescale parameter t_{dyn} and the fractional-area parameter r in some simplified dynamical test problems.

3.3 Prescribed sinusoidal buoyancy forcing

This test is based on our f-plane $\omega-$equation (79). Suppose forcing is sinusoidal in space, and sinusoidal or exponential in time with the complex form $\exp(ikx + imp + \sigma t)$. Then from (79) we have

$$1 - r \equiv -s^2\omega/B = s^2k^2/[(f^2+\sigma^2)m^2 + s^2k^2] \tag{9}$$

Let us define a characteristic wavenumber $k_r = (f^2+\sigma^2)^{1/2}m/s$. In the steady rotating case k_r is an inverse Rossby radius corresponding to the prescribed vertical scale. However in the unsteady case k_r is affected by effects related to the time taken for gravity waves to propagate.

We then have

$$1 - r = -s^2\omega/B = k^2/(k^2 + k_r^2) \tag{10}$$

or

$$t_{dyn} = (-B/s^2\omega - 1)/\sigma = \sigma^{-1}k_r^2/k^2 \tag{11}$$

The assumed sinusoidal time-dependence implies that σ^2 is real, but more generally σ itself could be real or (pure) imaginary.

It can also be shown that if Rayleigh damping at a rate c_R is applied to all fluctuating variables (represented collectively by a fluctuating state vector X) then a transformed variable $X^* = X\exp(c_R t)$ obeys the same equations as X obeys in the undamped problem. In this sense a damped problem is formally equivalent to an undamped problem with exponentially growing forcing (i.e. real, positive σ).

(i) In the balanced (strongly rotating or slowly-evolving) limit with real positive σ ($\ll |f|$), t_{dyn} is real, positive and varies as $\sigma^{-1}k^{-2}$

(ii) In the unbalanced (weakly rotating or rapidly-evolving) limit $t_{dyn} \sim \sigma(m/sk)^2$. Note that $t_{GW} = m/sk$ corresponds to the timescale for gravity-waves to travel one horizontal wavelength. Note also that the dependence on σ is opposite to the dependence found in the balanced case.

(iii) There is a resonance at $\sigma = \pm if$ (e.g. diurnal forcing at 30° latitude), which formally gives $t_{dyn} = 0$, i.e. very strong dynamical response.

(iv) If σ is imaginary then t_{dyn} is also imaginary, i.e. the dynamical feedback is out of phase with the buoyancy anomaly. It may then be more difficult to base a t_{dyn} parametrization strictly on the benchmark problem, although for some purposes use of the timescale might still be legitimate.

As noted above, time-varying convection usually contains a steady component, so both 'slow' and 'fast' limits are relevant.

Note finally that our equations for r and t_{dyn} in the sinusoidally forced problem can be written more compactly using a modified Burger number Bu_1

$$Bu_1 = k/k_r \tag{12}$$

$$r = 1/(1 + Bu_1^2) \tag{13}$$

$$t_{dyn} = \sigma^{-1} Bu_1^{-2} \tag{14}$$

For forcing at wavelengths small compared to the modified Rossby radius k_r^{-1} we have high values of Bu_1, small area parameter r and small t_{dyn}, i.e. the large-scale dynamics respond strongly and rapidly to oppose the forcing. Conversely, for forcing at wavelengths large compared to k_r^{-1} we have low Bu_1 and $r \to 1$ whilst $t_{\rm dyn} \to \infty$, so the large-scale dynamical feedback can be neglected.

3.4 Localized buoyancy forcing

We now consider the case when convection is forced over a finite region, but dynamical response occurs within a large domain. The relevant forcing might come from a local sea-surface temperature anomaly or island. Localized forcing might also arise from a gridpoint storm in a large-scale model.

It might be thought that the localized problem is essentially the same as the previous case, only with the wavelength k replaced by the reciprocal linear dimension of the forcing region. However the localized non-sinusoidal problem differs in containing a range of wavelengths and an asymmetry between ascent and descent.

We specify the 2D problem as follows. The inner forcing region extends between $x = \pm x_1$, with buoyancy effectively prescribed at the boundaries. The outer domain is unbounded. At the inner boundaries u and p, and therefore b are continuous (from continuity and the momentum equation). The anomalies of u and b are again assumed to vary sinusoidally or exponentially in time, and sinusoidally in pressure. Furthermore the inner region is assumed to be horizontally homogeneous.

Outside the forcing region we therefore assume a dependence $\exp(ik|x| + imp + \sigma t)$ which satisfies the ω-equation

$$[(f^2 + \sigma^2)\partial_p^2 + s^2 \nabla_H^2]\omega = 0 \tag{15}$$

if

$$(f^2+\sigma^2)m^2+s^2k^2=0 \tag{16}$$

i.e. $k=[-(f^2+\sigma^2)]^{1/2}m/s$. For admissible solutions (bounded as $|x|\to\infty$), $Re(ik)\le 0$.

Averaging the continuity equation $\partial_x u+\partial_p\omega=0$ horizontally across the inner domain (average denoted by $\langle\rangle$) gives

$$\partial_p\langle\omega\rangle=-u(x=x_1)/x_1=\partial_p\omega(x=x_1^+)/ikx_1 \tag{17}$$

so that

$$\langle\omega\rangle=\omega(x=x_1^+)/ikx_1 \tag{18}$$

We use the notation $(x=x_1^+)$ to denote the one-sided limit because ω is not in general continuous at the inner boundaries and may change suddenly to compensate for sudden changes in B. Thus using continuity of buoyancy at $|x|=x_1$ we obtain

$$t_{dyn}=\frac{\partial_p\Phi'}{s^2\langle\omega\rangle}=\frac{-\omega(x=x_1^+)}{\sigma\langle\omega\rangle}=-\sigma^{-1}ikx_1 \tag{19}$$

since $\sigma\partial_p\Phi'=-s^2\omega$ in the unforced outer region.

Thus

(i) In the balanced (slowly-varying or rapidly rotating) limit $|\sigma|\ll|f|$ we have $ik\sim-|f|m/s$ and $t_{dyn}\sim(|f|/\sigma)(mx_1/s)$ so that t_{dyn} exceeds the timescale mx_1/s for gravity-waves to cross the inner domain.

(ii) In the unbalanced limit ($|\sigma|\gg|f|$) we have $ik\sim-\sigma m/s$ so $t_{dyn}\sim mx_1/s$ (scaling on the gravity-wave crossing time). In this case parametrization based on t_{dyn} is straightforward because t_{dyn} is real, positive and independent of σ.

(iii) Forcing at the frequency $\sigma=\pm if$ again induces a strong resonant response with formally $t_{dyn}=0$

In §3.5 we will check this *ad hoc* derivation against a more formal solution for localized forcing.

3.4.1 Comparison between localized and non-localized forcing

There is an obvious difference in scaling of t_{dyn} between the localized and the spatially sinusoidal forcing cases. For instance in the nonrotating (unbalanced) problem:

- in the localized case $t_{dyn} \sim mx_1/s$ (the gravity-wave travel-time across the inner region)
- in the sinusoidal case $t_{dyn} \sim \sigma(m/sk)^2$

Since the forcing timescale σ^{-1} is independent of the gravity-wave timescales, these are quite different scalings, suggesting that t_{dyn} becomes particularly short (i.e. dynamical feedback particularly strong) for quasi-steady sinusoidal forcing.

The difference between the scalings for the localized and sinusoidal problems respectively can be explained by considering the horizontal length-scale of the buoyancy anomaly. In the sinusoidal case this length-scale is forced to match the forcing length-scale, but in the localized case the buoyancy anomaly (and hence the pressure anomaly) decays (or varies) away from the inner region on the natural length-scale of free oscillations.

In the localized forcing case the horizontal momentum balance

$$\partial_t u = -\partial_x \Phi \tag{20}$$

can be represented rather simply by the relation $u' = \pm\Phi'/c$ for gravity waves propagating at speed $\pm c$. Applying this at the inner-region boundaries we obtain $t_{dyn} \sim x_1/c$.

Another (equivalent) explanation is by interpreting the localized forcing as a spectrum of sinusoidal modes, in which the small wavelengths dominate the response.

Both localized and sinusoidal problems are valid in some sense but localized forcing is arguably more relevant to most issues in convection parametrization.

Our characterization (on timescales $\ll |f|^{-1}$) of the localized forcing problem in terms of the relatively fast timescale t_{GW} might suggest that f and the outer-region response are unimportant for our purposes. However, as noted above, in practice the buoyancy-forcing will normally contain a quasi-steady (or time-mean) component. As shown in the

dispersion relation above, in the limit of slow forcing t_{dyn} becomes large and the 'area' parametrization based on $r/(1-r) = \sigma t_{dyn} x_1(|f|m/s)$ becomes potentially more relevant. For instance, this limit may be viewed as providing model for the behaviour of convection within the Hadley circulation (discussed further below).

The 'area' parametrization can be implemented in a two-column modelling procedure, where the two columns have areas in the ratios $r : (1-r)$ and we force the dynamical circulation to remove a fraction $1-r$ of the difference in diabatic heating between the two columns. In the limit $r \to 0$ we recover the model of Sobel & Bretherton (2000). Note that the implied vertical velocity should be applied to all scalar fields, and not just to buoyancy.

3.5 Case with top-hat forcing

Our specification of localized forcing assumed that the buoyancy in the inner region was essentially homogeneous. However one might expect in reality the core of the inner region to be somewhat more buoyant than the periphery. To show that this does not fundamentally change the answers we now solve a problem where the buoyancy forcing is homogeneous but the buoyancy distribution is not.

We therefore solve (again in 2D)

$$[-(f^2+\sigma^2)m^2 + s^2\partial_x^2]\omega = -\partial_x^2 B \tag{21}$$

$$B = \chi\{|x| < x_1\} \tag{22}$$

where the indicator function χ takes the value 1 if its argument is true, and 0 otherwise. This equation becomes easier to manipulate if we define the residual buoyancy forcing (after allowing for dynamical compensation)

$$R_B = B + s^2\omega \tag{23}$$

since then

$$(\partial_x^2 - |k|^2)R_B = -|k|^2 B \tag{24}$$

where, as before, $|k|^2 = (f^2+\sigma^2)m^2/s^2$. Since the right-hand side of (24) is finite, we know that R_B and $\partial_x R_B$ must be continuous even at the inner boundary.

The solutions are

$$R_B = \begin{cases} 1 + c_1 \cosh|kx| & \text{(inner region)} \\ c_2 \exp(-|kx|) & \text{(outer region)} \end{cases} \tag{25}$$

where continuity of $\partial_x R_B$ requires

$$c_1 \sinh|kx_1| = -c_2 \exp(-|kx_1|) \tag{26}$$

and then continuity of R_B requires

$$1 + c_1 \cosh|kx_1| = -c_1 \sinh|kx_1| \tag{27}$$

so that $c_1 = -(\cosh|kx_1| + \sinh|kx_1|)^{-1}$.

The area parameter is then given simply by

$$r/(1-r) = R_B/B = 1 - \cosh|kx|/(\cosh|kx_1| + \sinh|kx_1|) \sim |kx_1| \text{ as } kx_1 \to 0 \tag{28}$$

where, as normal, this evaluation is conducted within the convecting region. This is the same result as we obtained before, and confirms our previous *ad hoc* solution of the localized-forcing problem.

3.6 Extension to 3D and illustration with gridpoint storms

We now sketch the extension of these 2D arguments to 3D, in order to indicate the nature of the mathematical solutions and the scales which emerge.

In 3D the governing equation becomes

$$[-(f^2+\sigma^2)m^2 + s^2\nabla_H^2]\omega = -\nabla_H^2 B \tag{29}$$

so we have to solve a Helmholtz equation with circular symmetry. Away from sources B we have

$$(\nabla_H^2 - k^2)\omega = 0 \tag{30}$$

where $k^2 = (f^2+\sigma^2)m^2/s^2$. Imposing circular symmetry gives

$$r^{-1}\partial_r(r\partial_r\omega) - k^2\omega = 0 \tag{31}$$

which is a modified Bessel equation of degree 0. The relevant solution, decaying at infinity, is $\omega = K_0(kr)$ in the notation of Abramowitz & Stegun (1970), up to an arbitrary factor which we shall define via a constant c such that

$$\partial_p\omega = cK_0(kr) \tag{32}$$

since we have already assumed that the problem is separable in p.

The asymptotic properties are

$$K_0(z) \sim (\pi/2z)^{1/2} e^{-z} \tag{33}$$

for large positive real z, and

$$K_0(z) \sim -\ln z - \gamma + \ln 2 + ... \tag{34}$$

for small positive real z, where $\gamma \simeq 0.577$ is Euler's constant.

The expression for large z is accurate to 10% at $z = 1$ and to about 30% at $z = 0.1$. Roughly speaking the solutions remain exponentially decaying on the same scale as for 2D.

Following the approach given at the beginning of §3.4, we use this solution for ω to compute the inflow to an inner region of radius r_c associated with a given buoyancy anomaly. In this circularly symmetric geometry the continuity equation is

$$r\partial_p \omega + \partial_r (ru) = 0 \tag{35}$$

where u is taken as the radial velocity, we have

$$r_c u(r_c) = \int_{r_c}^{\infty} r' \partial_p \omega(r') dr' \tag{36}$$

from which, on using the outer approximation (33) we estimate

$$r_c u(r_c) \simeq c(\pi/2)^{1/2} \int_{r_c}^{\infty} r'(kr')^{-1/2} e^{-kr'} dr' \tag{37}$$

The use of this approximation is reasonable because the weighting of the integrand with r' prevents substantial contributions to this integral from regions where kr' is small. Hence for our estimates it suffices (assuming $r_c \ll k^{-1}$) to evaluate just the complete integral. By transforming to a variable ξ where $kr' = \xi^2$, and using $\int_{-\infty}^{\infty} \xi^2 e^{-\xi^2} = \pi^{1/2}/2$ we obtain

$$r_c u(r_c) \sim c(\pi/2^{3/2} k^2) \tag{38}$$

as $k'r' \to 0$, where $c = \partial_p \omega(r_c^+)/K_0(kr_c)$.

The equation for $\langle \omega \rangle$, i.e. the horizontal-mean ω over the inner disc, is then

$$\pi r_c^2 \partial_p \langle \omega \rangle = -2\pi r_c u(r_c) \tag{39}$$

whence

$$\partial_p \langle\omega\rangle = -2r_c^{-1}u(r_c) \sim -(\pi/2^{1/2})(kr_c)^{-2}\partial_p\omega(r_c^+)/K_0(kr_c) \tag{40}$$

from (38), and on further using $\sigma\partial_p\Phi' = -s^2\omega(r_c^+)$ as above, we obtain

$$\langle\omega\rangle \sim (\pi/2^{1/2})(kr_c)^{-2}[\sigma/s^2K_0(kr_c)]\partial_p\Phi' \tag{41}$$

and finally

$$t_{\rm dyn} \equiv \partial_p\Phi'/s^2\langle\omega\rangle \sim 2^{1/2}(\pi\sigma)^{-1}(kr_c)^2K_0(kr_c) \tag{42}$$

with the slight caveat that here t_{dyn} is based on the value of $\partial_p\Phi'$ at the *boundary*. We can also write from (8)

$$r/(1-r) = \sigma t_{\rm dyn} \sim 2^{1/2}(kr_c)^2\pi^{-1}K_0(kr_c) \tag{43}$$

We see then that the effective subsidence area scales on $k^{-2}/K_0(kr_c)$. This represents primarily a horizontal scale set (as in the 2D problem) by the gravity-wave travel distance, modified slightly by a weak logarithmic dependence via $K_0(kr_c)$ (e.g. $K_0(0.1) \simeq 2.4$).

Whilst in terms of the effective fractional area r the 3D scaling seems a straightforward generalization of the 2D problem, the scaling for $t_{\rm dyn}$ is less obvious. Eq. (43) implies for instance that in the non-rotating case $t_{\rm dyn} \sim \sigma m^2 r_c^2 s^2$, which is smaller by a factor $\sigma m r_c/s$ than its 2D counterpart, and shorter than the gravity-wave crossing time for the inner domain (if kr_c is small).

With hindsight one can see that the 3D problem cannot be expected to scale in the same way as 2D. If the horizontal lengthscale of influence (and compensating subsidence) is essentially the same as in 2D, as follows straightforwardly from the Helmholtz equation for both balanced and unbalanced cases, then (for small kr_c) the fractional area parameter must be much smaller in 3D. Consequently the large-scale dynamics constrain the inner-region buoyancy much more strongly and so (for given Q_1 in the inner region) the buoyancy anomaly in 3D will be much weaker, whereas the inner-region ω (balancing Q_1) will be about the same. Hence $t_{\rm dyn}$ in 3D must be much smaller than in 2D and cannot scale similarly. In 3D it is much easier for a 'point source' to attain small values of r, i.e. the situation where the convective heating Q_1 is almost entirely compensated by large-scale advection, consistent with the assumptions of Sobel & Bretherton (2000).

These results are consistent with the initial-value problem solutions of Bretherton (1993). Bretherton considers the time-development of convection after a convective point source of buoyancy is switched on at time $t = 0$, whereas here we have concentrated on normal-mode solutions, whose scaling is in some sense independent of time.

For the 3D problem, in contrast to the 2D problem, Bretherton finds that the local buoyancy anomaly at the source *decreases* with time (like t^{-1}), despite the steady buoyancy forcing (after $t = 0$). This is consistent with our arguments, which suggest this buoyancy anomaly should be governed by the total time-integrated heat-input (growing as t) divided by the effective area (growing as t^2). This again implies small values of $t_{\rm dyn}$ on long timescales in the 3D problem.

3.6.1 'Gridpoint' (column) storms

As an illustration of the 3D problem, consider the following simple description of a gridpoint storm (GPS), which may occur in large-scale models. The GPS in a column of horizontal dimensions 100km may grow roughly at a rate $\sigma \simeq 10^{-4}\mathrm{s}^{-1}$, often at locations within a few degrees latitude of the Equator, so that $|f| \sim 10^{5}\mathrm{s}^{-1}$. With these values the Coriolis effect will be negligible in determining k. Assuming further a vertical structure of the form $\omega \sim \sin \pi p/p_r$, we then obtain $k \sim \sigma m/s \sim (300\mathrm{km})^{-1}$. For an approximate comparison, let $r_c = 56$km, which gives a disc of the same area as a 100km$\times$100km square.

For these values, using the above approximations, and using $K_0(56/300 \simeq 0.19) \simeq 1.7$, we derive $r \sim 0.03$ and $t_{\rm dyn} \sim 300$s. These small values suggests that the Sobel-Bretherton model might be well suited to describe GPS behaviour.

A full analysis of the gridpoint storms problem is obviously beyond the scope of this note. Clearly one would need to consider the discretizations in space and time, and the parametrization of subgrid convection, as well as the convective-dynamical feedback. But an understanding of relevant scales, and of the extent to which dynamics compensate for local physical increments, should contribute to such an analysis.

3.7 The Gill tropical heating problem

We now evaluate our dynamical adjustment-time t_{dyn} within the analytic solutions of Gill (1980) for idealized tropical heating on a β-plane.

Gill (1980), following earlier work by Matsuno, presents a linear analytical solution of an idealized quasi-steady representation of heating over Indonesia. He uses the linearized shallow-water equations, which are isomorphic to our linearized hydrostatic equations for

given vertical wavenumber m, except that he retains the β term and introduces a small damping coefficient ϵ. He nondimensionalizes his equations with the lengthscale $(c/2\beta)^{1/2}$ and the timescale $(1/2\beta c)^{1/2}$, where the characteristic gravity-wave speed c $(= s/m$ in our set) becomes 1 in the nondimensional set.

In Gill's set pressure corresponds to negative buoyancy in our set, and obeys a nondimensional equation

$$\partial_t p = w - Q \tag{44}$$

where Q is the nondimensional heating. We may therefore write

$$t_{dyn} = -p'/w \tag{45}$$

within the same nondimensionalization.

We concentrate here on Gill's 'symmetric case', where forcing is symmetric about the equator. In this case he sets

$$Q(x, y) = F(x) \exp(-\tfrac{1}{4} y^2) \tag{46}$$

where $F(x) = \cos kx$ for $|kx| < \pi/2$ and 0 otherwise. This latitude dependence is chosen to correspond to the zeroth parabolic cylinder function.

For the symmetrically forced case Gill identifies two components of the response, namely a 'Kelvin-wave' part and a 'planetary-wave' (or Rossby-wave) part.

(a) For the Kelvin-wave part, which Gill interprets as representing the Walker circulation over the Pacific, his solutions give (as $\epsilon \to 0$)

$$t_{dyn} = -p'/w \sim q_0(x)/F(x) \tag{47}$$

where

$$q_0(x) \sim -k^{-1}(1 + \sin kx) \tag{48}$$

so that (as nondimensionalized)

$$t_{dyn} \sim (1 + \sin kx)/k \cos kx \tag{49}$$

Taking into account the normalization, this result implies that the dynamical buoyancy-adjustment time scales (as above) on the gravity-wave 'crossing time' c/k. Note also that t_{dyn} is independent of latitude, although there is some variation with x.

(b) For the Rossby-wave component (again as $\epsilon \to 0$)

$$t_{dyn} = -p'/w' \sim -q_2(x)/F(x) \sim -\sin kx/k \cos kx \tag{50}$$

The scaling is again consistent with the 'crossing-time' c/k but with the significant difference that the dynamical timescale for this component is *negative*. This can be traced to the property that, despite positive prescribed heating and a positive w component, the pressure anomaly of this component is *positive* in the heated region.

Note that the overall solution involves both components. Since both satisfy the crossing-time scaling separately, so does their sum. Gill shows that despite some positive contributions, the total pressure anomaly is everywhere negative (at least with $\epsilon = 0.1$). In other words, the net effect of prescribed heating plus dynamical response acts as expected to warm all parts of the domain, albeit in a more complex spatial pattern than seen in our simple f-plane test problems.

In summary, even though Gill's problem is dominated by β effects which are neglected in our other dynamical test problems, it obeys the same t_{dyn} scaling and in most respects the same qualitative behaviour as before. The indirect meridional circulation in the forcing region (this can be interpreted as a Sverdrup PV response, see Annex) is one significant difference, but the net dynamical feedback on the buoyancy anomaly is negative and reasonably well described by a t_{dyn} parametrization.

The Gill problem can be viewed as a stepping stone from our highly idealized benchmark problems to convective-dynamical interactions as they occur in current global models. Jin & Hoskins (1995) use an idealized configuration of a GCM, to compare the Gill-Matsuno predictions with a range of more realistic time-integrations. Depending on location and flow-field, localized heating can set off a pattern of Rossby waves affecting remote parts of the globe. However our present concern is only the feedback loop from convection to dynamics and then back to convection at the original location, rather than with remote one-way impacts which can be handled via other methodologies.

3.8 Summary of benchmark dynamical tests

We have shown that parametrizations of the dynamical buoyancy-adjustment timescale t_{dyn} or alternatively of the fractional-area parameter r emerge naturally from our dynamical test problems. These parameters are related through (8), but because this relation explicitly involves the forcing timescale σ^{-1} they cannot both be independent of σ. In

particular, in the limit of slow forcing (which approaches the quasi-steady problem as $\sigma \to 0$) we must have either $t_{dyn} \to \infty$ or $r \to 0$.

If we want the same dynamical feedback model to be valid over a range of timescales σ^{-1}, then we need to consider two possible types of behaviour in the slow limit:

(i) *balanced* cases where buoyancy anomaly can be balanced by rotational dynamics without vertical acceleration: r parameter may be independent of σ.

(ii) *unbalanced* cases where any buoyancy anomaly implies vertical acceleration: t_{dyn} parameter may be independent of σ

In case (i) only *changes* in the buoyancy anomaly can induce vertical circulation, so ω becomes linked to the convective heating Q_1, whereas in (ii) the dynamics 'drives towards' a state of flat isentropes. Our tests with idealized sinusoidal forcing encompass both cases, according to whether or not f is negligible.

The Hadley problem, although centred on the Equator, includes significant rotational forces so that thermal wind balance is a key ingredient in the Held-Hou model of the Hadley circulation. In contrast the Gill problem, also centred on the Equator, involves a significant zonal flow component. Because equatorial buoyancy anomalies cannot be sustained without vertical circulation (e.g. Walker cells), this is essentially a type-(ii) problem consistent with our derivation of a finite t_{dyn} value for the slow limit.

3.9 Interpretation of published schemes

Woolnough (2001) discusses published dynamical-feedback schemes due respectively to Nilsson & Emanuel (1999), Raymond and Zeng (2000) and Sobel and Bretherton (2000). In the light of our benchmark dynamical problems we now briefly revisit that discussion.

For all these schemes the momentum balance is key, as the scalar tendency due to advection by known wind over a prescribed length-scale is straightforward to estimate.

The Nilsson-Emanuel scheme is based on an equation

$$\partial_t u = -\partial_x \Phi + \mu \partial_p^2 u \tag{51}$$

for the large-scale flow, where μ is an artificial diffusion coefficient. In the 'unbalanced' limit of small timescales or small μ, this reduces to simple gravity-wave propagation. On longer timescales, where one might expect some kind of balance to be attained, the physical interpretation is much less clear. Presumably an essentially artificial boundary layer will be set up, governed by μ and the other parameters.

The Raymond-Zeng model is also difficult to interpret literally. It uses an equation

$$\rho u = -\tau_{RZ} \Delta p / \Delta x \tag{52}$$

where τ_{RZ} is a timescale, and $\Delta p / \Delta x$ approximates the large-scale pressure gradient. This might be viewed as approximating a fundamentally unbalanced system starting from rest and running for a time τ_{RZ}.

Owing to their limitations in relation to quasi-steady dynamics, it seems unlikely that either the Nilsson-Emanuel or Raymond-Zeng schemes can shed great light on, say, the Gill-Matsuno or Jin-Hoskins problems.

The Sobel-Bretherton model is perhaps conceptually the clearest of the published schemes, although it is also extreme in the strength of the assumed large-scale feedback. Essentially this model assumes that the convective heating Q_1 is entirely compensated for by vertical motion, and that vertical motion is then used to advect the moisture variable. In terms of our dynamical feedback analysis, Sobel and Bretherton set both r and t_{dyn} to zero.

The scheme developed by Woolnough (2001) may be viewed as an extension of the Sobel-Bretherton model in the light of the present arguments, and with implementation in a Cloud-Resolving Model. In particular the feedback parameters r and t_{dyn} are not assumed to vanish. Tests are shown with values e.g. $r = 0.1$ and $t_{dyn} = 6$ hrs in the CRM coupled to the dynamical feedback scheme. A 'test column' for CRM simulation is set up based on a sea surface temperature elevated by 1K relative to the background. With these values, significant nontrivial feedbacks on convection are found. The vertical velocity profiles derived are considered broadly realistic.

These ambiguities in some of the existing literature illustrate the need to make stronger connections with specific benchmark problems. As noted above, the feedback models for present purposes do not need to be precise, but should be capable of a steady state which can be realistically interpreted.

4 Coupling to simple representations of convection

From general considerations (§2) we expect large-scale dynamical feedback to have implications for process studies if such coupling *substantially changes the sensitivities* of the problem to process issues. We show first how coupling can affect sensitivities using a highly simplified representation of convection, parametrized simply with an adjustment timescale. The sensitivity to this timescale under dynamical coupling will illustrate how other process-sensitivities may be affected.

Consider the case of convection over a sea of prescribed surface temperature which varies horizontally over a lengthscale of order 1000km. The equilibrium state of this system will depend on both vertical transport (primarily by convection) and horizontal transport (primarily by large-scale dynamics). If there were no vertical transport then a horizontally uniform atmospheric state could exist irrespective of the SST boundary condition. On the other hand, if there were no horizontal transport then the problem would decompose into separate columns. In general, though, the equilibrium state is influenced by both vertical and horizontal transport.

4.1 An electrical analogy

It is convenient to express this problem in terms of an electrical analogy widely used in boundary-layer hydrology, e.g. Blyth, Dolman & Wood (1993). A conventional electrical circuit diagram specifies the ratio of potential difference to current as a *resistance* between one node and another. Here, by analogy, we may generically define the resistance to be the spatial differences in temperature (or some temperature-like variable) divided by the associated total flux of that quantity across the relevant surface (or by a term playing the role of such a flux). The dimensions of such resistance are then s kg^{-1}.

To analyze this problem in the simplest terms, let us consider only *two* columns of the atmosphere, each tending to adjust to its own characteristic equilibrium with the local surface. Following our analysis above, we assume these columns are dynamically coupled using a specified t_{dyn}. We further simplify the columns so that each contains only one atmospheric gridpoint, and simplify the physics such that the single-column response (including convection, radiation and cloud-radiative feedbacks) adjusts the column temperature towards a prescribed value T_{SST} (a function of the underlying sea surface temperature) on some prescribed timescale.

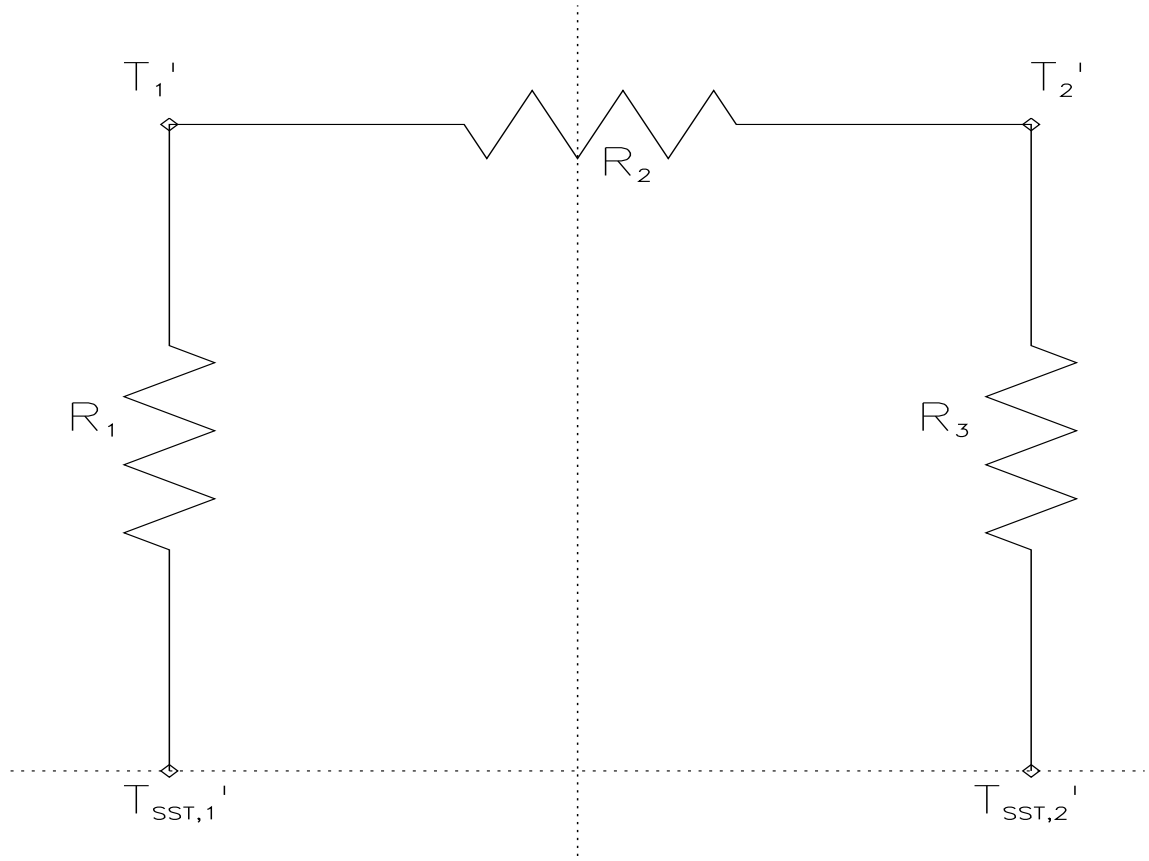

Figure 1: Sketch of electrical resistance analogy, for the problem of convective and dynamical equilibrium of two atmospheric columns over different sea surface temperatures. Here T_1' and T_2' are bulk measures of the respective atmospheric temperature anomalies, and $T'_{SST,1}$ and $T'_{SST,2}$ the corresponding values which would be in vertical equilibrium with the underlying sea surfaces. As discussed in the text, R_1, R_2, R_3 denote 'resistances' which describe the change in temperature difference needed to induce a given change in flux.

It is convenient to prescribe each temperature as an 'anomaly' relative to a full equilibrium state with uniform SST and each column in equilibrium with its SST. These anomalies will be denoted by primes, and we shall assume that we can linearize the processes. In effect we thus construct a tangent linear model.

Specifically we set

$$\partial_t T_1' = (T_{\mathrm{SST},1}' - T_1')/t_{phys1} + (T_2' - T_1')/t_{dyn1} \tag{53}$$

$$\partial_t T_2' = (T_{\mathrm{SST},2}' - T_2')/t_{phys2} + (T_1' - T_2')/t_{dyn2} \tag{54}$$

Heat conservation requires that $M_1/t_{dyn1} = M_2/t_{dyn2}$, where M_1, M_2 are the masses associated with each column. Together these equations imply that in equilibrium

$$(T_{\mathrm{SST},1}' - T_1') : (T_1' - T_2') : (T_2' - T_{\mathrm{SST},2}') = R_1 : R_2 : R_3 \tag{55}$$

where the resistances R_i have the values

$$R_1 = t_{phys1}/M_1 \tag{56}$$

$$R_2 = t_{dyn1}/M_1 = t_{dyn2}/M_2 \tag{57}$$

$$R_3 = t_{phys2}/M_2 \tag{58}$$

The 'current' (i.e. heat-flux) through this circuit is then $(T_{\mathrm{SST},2}' - T_{\mathrm{SST},1}')/(R_1+R_2+R_3)$. The vertical current $(T_{\mathrm{SST},1}' - T_1')/t_{phys1}$ corresponds then to an integrated $Q_1 + Q_R$ (in the usual convective notation) rather than to a flux of either dry or moist static energy. The non-radiative part corresponds to the surface sensible-heat flux plus the surface latent-heat flux multiplied by the precipitation efficiency.

The electrical analogy shows immediately how (in this simple problem) the convection in column 1 would change if any of the parametrized terms were changed. Essentially the current is controlled by the largest resistance, so that (cf. the three methodological approaches listed in §2)

(i) if $R_1 \gg R_2 + R_3$ then vertical transport in column 1 is the rate-limiting process, and convective heating may be regarded as a 'prescribed forcing' to the large-scale dynamics.

(ii) if if $R_1 \ll R_2 + R_3$ then the convection in column 1 may be viewed as a response to 'prescribed large-scale forcing'

(iii) if the resistances are comparable then neither large-scale forcing nor local convection can be independently prescribed.

Both cases (i) and (iii) imply that changes to convection parametrization affecting R_1 will systematically change the amount of precipitation in column 1. Only in case (ii) can precipitation be viewed as straightforwardly controlled by the large scales.

We should be able (at least in principle) to estimate these resistances within a large-scale model, or indeed in the real atmosphere. The greatest difficulty may lie in assigning appropriate values for the single-column equilibrium temperatures $T'_{\mathrm{SST},1}, T'_{\mathrm{SST},2}$ or the physical adjustment timescales. Within a given model, such parameters can be estimated by applying forcings of different magnitudes and comparing the mean buoyancy perturbations resulting from those forcings.

4.2 Evaluation of physical adjustment timescales

Convective adjustment timescales of 1-2 hours are often used in NWP, and may perhaps be used as an estimate for t_{phys} during active convection. However note that convection even in the ITCZ will not be continuously active and is influenced by diurnal cycles, intra-seasonal oscillations and geographical effects, so that an effective overall adjustment timescale might be somewhat longer. Also, as seen in §5, we are not only concerned with intense deep convection but potentially with convection of shallow or moderate depth, whose adjustment timescales might be longer.

If we base the dynamical timescale on the gravity-wave crossing-time then we may assume that gravity-waves cross 100km (cf. a typical CRM domain size, approximately 1 degree of latitude) in about 1 hour. James (1994) estimates the meridional scale of the ITCZ as only 100km, in which case its dynamical adjustment should be governed by a 1-hour timescale.

If column 2 is not convecting then its adjustment timescale may be much longer, and determined essentially by radiative adjustment. But this does not necessarily mean that radiation controls the overall 'current', because the vertical resistance of a column depends inversely on its mass. Hence the process issues have greater freedom to affect the intensity of convection in a small column than in a large column. For instance, in the descending branch of the Hadley circulation, slow radiative adjustment can effectively

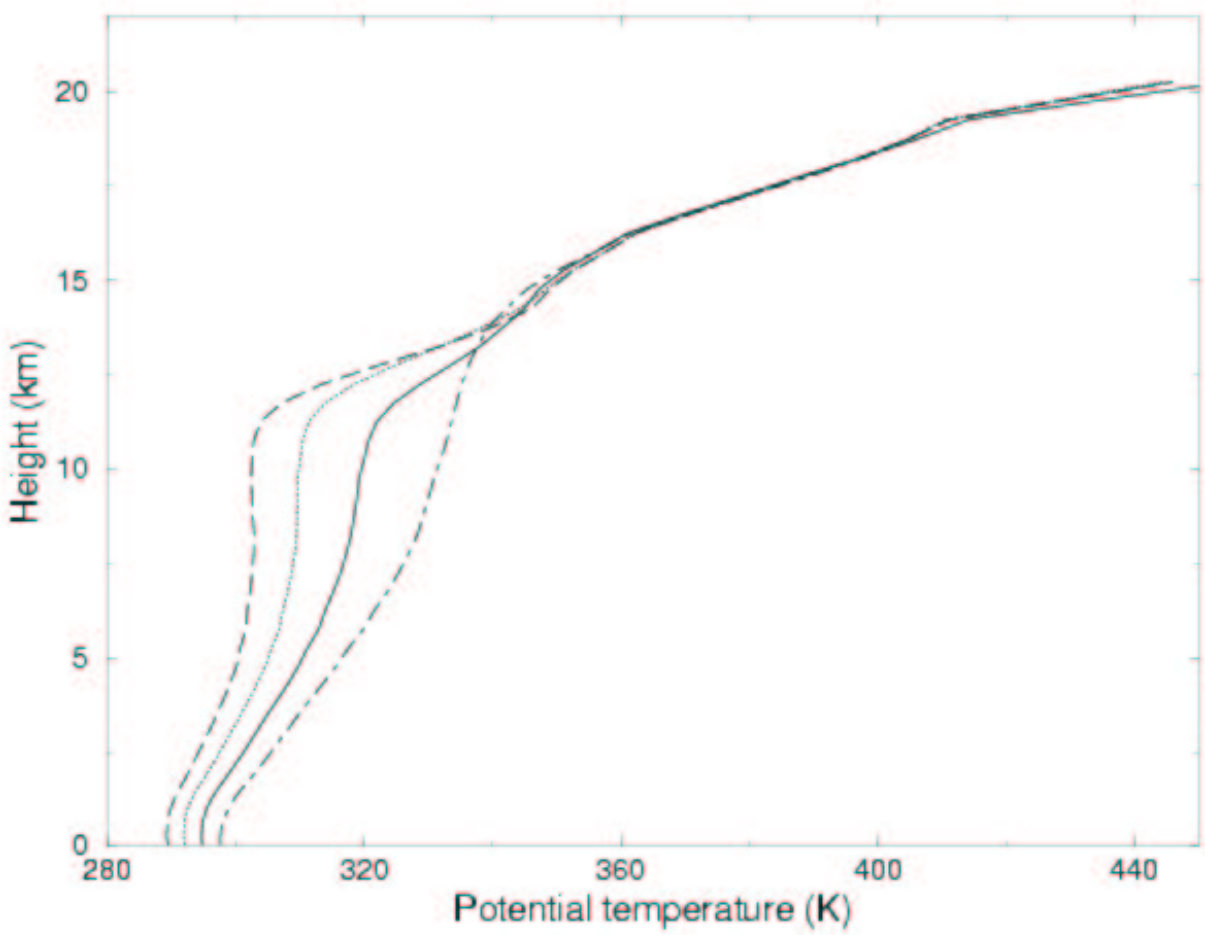

Figure 2: Equilibrium temperature profiles for radiative-convective equilibrium simulations from Cohen (2001). The radiative temperature-forcing varies between -4K/day (dot-dashed) and -16K/day (dashed). These forcing values are applied up to 7.5km, then tapered linearly to zero over the layer 7.5–12.5km. Figure courtesy of Brenda Cohen.

remove temperature increments transported from convection in the ITCZ, owing to the larger spatial scale of the descending region.

Fig. 2 is taken from the statistical-equilibrium radiatively forced Cloud-Resolving Model studies of Cohen (2001). This shows the impact on profiles of temperature and moist static energy of changing the magnitude of a prescribed radiative forcing.

Two changes in the θ-profile are evident as the radiative cooling is increased. First, the temperature profile appears to cool right down to the surface, or as close to the surface as the model can resolve. Under standard boundary layer theory, this finding is consistent with an increasing sensible heat-flux. Secondly, the lapse rate also increases significantly,

so the temperature change in the upper troposphere is much greater than in the boundary layer.

For present purposes it is not essential to distinguish how much of the change in lapse rate is an adiabatic consequence of the change in boundary layer values (rather than a non-adiabatic effect). We need only to evaluate the ratio of forcing to equilibrium impacts, which here suggests that the 'physical adjustment' t_{phys} reaches values around 10 hours in the upper troposphere.

James (1994), discussing the Held-Hou model of the Hadley circulation, comments that '..the basic mechanism in the Held-Hou model, namely angular momentum mixing and thermal wind balance, gives a general form of Hadley circulation which is insensitive to the details of how the heat enters the system. The actual location of the ITCZ and its strength will depend upon details of the tropical boundary layer, and the fluxes of moisture out of the surface...'. In a large-scale model, one might add the parametrization of convection (which cannot really be incorporated in the Held-Hou model) as another detailed factor affecting the ITCZ.

We may summarize the impact of coupling on process-sensitivities as follows. Where the transfer-resistance to convection is low compared to the dynamical transfer-resistance, then the amount of convection may be relatively insensitive to process-issues, being effectively constrained to balance an independent forcing. But in many weather and climate problems, the transfer-resistance to convection is significant relative to the dynamical transfer-resistance. In those situations, we expect convective process issues to be able to change fluxes and precipitation distributions.

4.3 Other convection schemes in the 'resistance' problem

We now briefly comment on the generalization of our simple convection scheme based on temperature-adjustment to schemes based either on a more complex combination of local parameters or alternatively on direct link to large-scale forcing.

In most convection schemes the convective heating depends not only on temperature but also on humidity q, and in many cases there is also an explicit dependence on large scale w (via moisture convergence, trigger functions or other assumptions).

Within our resistance model we can include a dependence of t_{phys} on q and/or w, or

indeed T. Such dependences would make the circuit nonlinear, but still fairly straightforward to solve numerically. A case of particular note is where convection is assigned a sharp 'threshold' dependence on some combination of these variables. In that case (depending on the numerical coefficients) the possibility arises that convection regulates itself to lie at the threshold. For example, a sufficiently strong dependence on relative humidity might in principle override other forms of regulation.

Another class of schemes is based on a direct attempt to balance large-scale forcing of moisture or instability, associated with the names of Kuo and Arakawa-Schubert respectively. Such attempts have often proved controversial, especially in regard to the conflict (at least on short timescales) between balancing moisture and balancing instability.

Raymond & Emanuel (1993) give a critique of the moisture-convergence approach, noting in particular its unconvincing properties under Galilean transformation, implying a largely spurious response to uniform horizontal advection. It is not absolutely clear that 'instability'-balancing closures entirely avoid such advection problems, but on short timescales the instability balance may be more relevant. One possibility is that the usefulness of the moisture-convergence approach stems largely from qualitative 'triggering' effects in confining deep convection to regions of ascent.

However we can show with our resistance model that *any* parametrization based on 'balancing large-scale forcing' has fundamental limitations. The reason is that *any* quasi-steady state with *any parametrization* (including all versions of our 'convective adjustment') must satisfy this condition. Balance (whether of moisture or instability) therefore *fails to constrain* our convective adjustment timescale.

The problem is perhaps even worse if the balance condition is evaluated only approximately for purposes of the convective closure. For example if (as in the most basic version of the Kuo scheme) we set convective precipitation to some fraction c_q (say 70%) of the moisture convergence, then such a scheme may have *no quasi-steady non-zero solution* to our problem of the response to prescribed SST anomalies.

If we allow time-dependence in our test problem, relaxing the steady state assumption, allowing unbalanced effects, but retaining the prescribed SST anomalies, we obtain

$$M(t) = \int_{-\infty}^{t} G(t-t')M(t')dt' \tag{59}$$

where the Green's function $G(t-t')$ accounts for the convergence caused by a given amount of convection, translated back (via the balance assumption) into current convection. This

argument suggests that in our problem a scheme based on 'balancing large-scale forcing' acts fundamentally as a persistence-based forecast resembling certain types of nowcasting. The scheme may be saved from grossly excessive convection by an overriding switch which prevents convection if the profile is overstabilized and limits sustained convective heating to $(T'_{SST,2} - T'_{SST,1})/M_2$ (essentially the rate at which the heating can be compensated by the dynamics), but is not saved from grossly under-responding in other cases.

As discussed by Raymond & Emanuel (1993), a 'moisture-balancing' closure can also be 'tweaked' by making the precipitation efficiency depend on the actual relative humidity. Although not unreasonable in itself, this adaptation does not solve all the problems, and shows the need to consider the 'actual state' as well as the 'forcing'.

Our conclusion is that whilst acknowledgment of profile balances is useful theoretically, and may also be useful numerically (if the equilibration is fast relative to a model timestep), any 'hardwiring' of such a balance in the convective closure is unreliable. We need to recognize the component of the flow which is driven by convective response to process-issues (including SST anomalies and various forms of convective inhibition), and is not controlled in any simple way by large scale dynamics.

5 Relationship to other systematic feedbacks

We now set our account of convective-dynamical feedback through buoyancy adjustment in the context of other systematic feedback loops which could affect the regulation of convection. Strong systematic positive or negative feedbacks may have implications for process studies, whereas e.g. a tendency for phase propagation may be regarded as a matter for the large-scale model.

The feedbacks discussed involve water vapour and moist-static-energy (from either horizontal transport or surface fluxes), radiative feedback and convective momentum transport. We shall see that although some of these are quantitatively significant, they do not destroy our basic picture, or the rationale for such a 'buoyancy-feedback' parametrization.

5.1 Water vapour and moist-static-energy feedbacks

In analyzing the temperature loop via resistance formulation earlier in §5 we ignored the corresponding loop for q. Actually the moisture loop does not invalidate the temperature loop if the physics timescale t_{phys} is known, but any changes to column moisture are likely to affect t_{phys}.

There are two key questions about moisture:

(i) how does moisture change in response to the combination of convection and dynamical feedback?

(ii) how do any moisture changes affect convection (or other moist physics)?

Question (ii) will be considered in §6. Insight into question (i) may be gained by thinking about moist static energy.

Moist static energy h is defined by

$$h = c_p T + \lambda q + gz \tag{60}$$

Loosely, h is a measure of 'sensible plus latent heat' and is especially relevant in situations of high precipitation efficiency, where added latent heat can be assumed to convert into sensible heat, rather than being stored in the vapour phase. Approximately related to the equivalent potential temperature θ_e, h is essentially conserved by advection, condensation and evaporation with source terms only from radiation or boundary fluxes.

A schematic plot of typical h-profiles is given in Fig. 3. Note that h varies with height more strongly than its saturated counterpart h_{sat} (defined as for h but with q replaced by q_{sat}). The latter quantity is more closely related to adiabatic moist stability arguments.

Neelin (1997) gives a theoretical analysis of feedbacks via h. He defines a 'gross moist stability' (GMS), essentially the change in total column h for a given vertical motion (roughly analogous to a bulk N^2 but measuring the impact on h rather than buoyancy). He concludes that GMS is small but positive, leading to the conclusion that sustained convective systems must draw their h from the surface, not from the ambient atmosphere.

Neelin's GMS is a measure of the 'vertical' integral $\int \omega(\partial h/\partial p)dp$, normalized by ω. Its value depends on the relative shape of the ω and h profiles. In particular, the conclusion

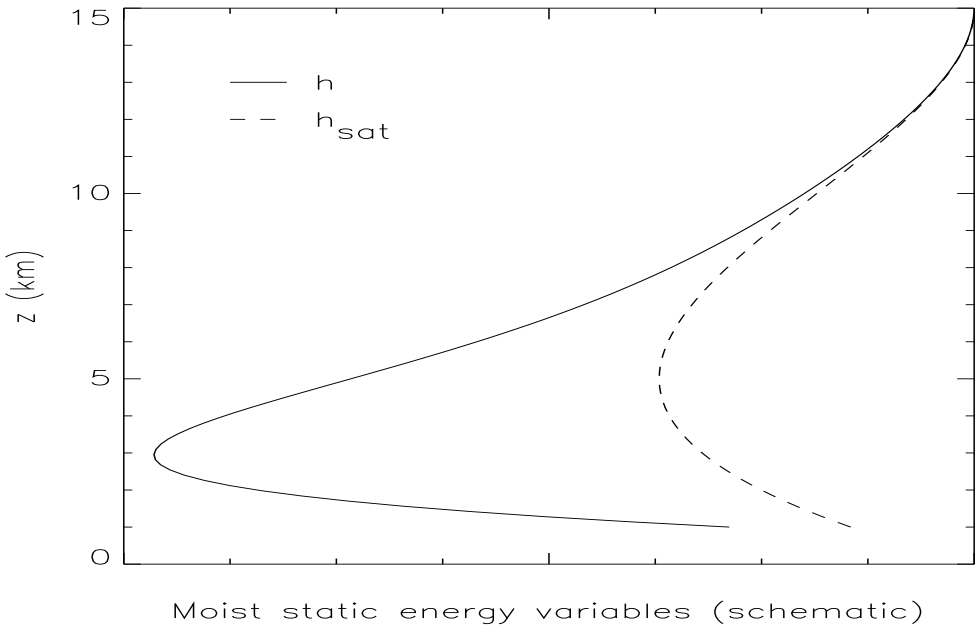

Figure 3: Schematic profiles of moist static energy h and saturated moist static energy $h_{sat} = c_p T + gz + \lambda q_{sat}$. The minimum in the h_{sat} curve contributes to conditional instability, whereas the much larger dip in the h curve has no direct significance for adiabatic parcel theory. Basic parcel theory can be expressed as a comparison between the (adiabatically conserved) parcel value of h and the 'environmental' h_{sat} profile.

that GMS is normally positive depends on the expectation that (roughly speaking) $|\omega|$ peaks higher in the troposphere (say at 5km) than the level of minimum h (say 2-3 km). It is easily seen that this finding is not completely universal without qualification. E.g. on the cloud scale, parcel ascent of high-h from the boundary layer may lead, at least temporarily, to an increase in column h as the cloud grows. In this case the 'normal GMS' argument fails because ω peaks in the layer of high h. However we may assume that GMS is positive for 'normal' large-scale atmospheric motions.

A particularly simple interpretation of GMS can be made in the Sobel-Bretherton limit, as we derived for $r \to 0$, in which vertical motion exactly compensates for temperature increments. For if temperature is fixed, then changes in h correspond to changes in q. Hence if gross moist stability is positive then S-B dynamics implies that convection plus dynamical feedback should give a net *drying* to the column.

5.1.1 Implications of moist static energy for gridpoint storms

Gridpoint storms (better called 'column storms') may be regarded essentially as a form of undesirable 'resolved convection' in a large-scale model, when the resolved dynamics and large-scale condensation and precipitation schemes take over from subgrid convection in a manner that is damaging and unphysical. The column moistens dramatically from the top of the boundary layer upwards, in a manner resembling a spurious large-scale parcel ascent.

We have seen that 'gross moist stability' arguments do not usefully apply to cloud parcel ascents because the assumptions about 'normal' shapes of ω-profiles are violated. Surface fluxes may also play a role but no clear evidence has yet been found that they drive GPSs.

However moist static energy thinking can also be applied to some extent locally, and without using GMS. Let us again apply S-B dynamics to a GPS (for which we argued earlier that the relevant r value is indeed small). This implies that the temperature profile is pegged to that of the large-scale environment, and hence that changes in h correspond to changes in q. The moistening is then controlled by the tendency terms for h:

(i) parametrized convection (vertical transport only)

(ii) large-scale dynamics $-\omega(\partial h/\partial p)$

(iii) large-scale condensation and precipitation: no effect on h

(iv) surface terms (thought not to be crucial; e.g. surface incoming radiation becomes blocked by thick large-scale cloud)

Consider first the *resolved-only* case where parametrized convection is not used. We see then that (in this S-B limit) any upward motion in a layer where h decreases with height will moisten the layer. If there is even partial condensation of this moisture increment (e.g. via a partial cloud scheme) then we have a *positive feedback*. This property may recall the concept of 'potential instability', whose relevance to real atmospheric convection may be questioned (Emanuel 1994). It is however clear from our analysis that on scales where Sobel-Bretherton dynamics apply, condensation is *counter-productive* where $dh/dz < 0$. That is, the net effect of condensation after dynamical compensation for the latent heating is actually to *increase* water vapour. This scenario is obviously unstable if increased q implies larger condensation. We have then not 'gross moist stability' but local moist instability.

If the T-profile (fixed, by the S-B assumption) is conditionally stable (the usual criterion, equivalent to $dh_{sat}/dz < 0$) then this instability is probably self-limiting: if the h profile reaches $h_{sat}(T)$ then the instability will stop because $dh_{sat}/dz < 0$. If however the T-profile is conditionally unstable, then the instability may not stop at saturation. In fact conditional instability may not be necessary for a disturbance to become a gridpoint storm. For an increase in h in layers where $dh/dz < 0$ will lead to layers of negative dh/dz developing higher up. If our dynamical scenario is maintained, we predict that the moistening will gradually progress from below until the whole convecting layer reaches the boundary-layer value of h. Based on typically observed h-profiles, e.g. those of Riehl & Malkus (1958), this would imply essentially the whole troposphere (up to say 15km) approaching large-scale saturation in the GPS column. The latent heat released would be so large as to damage severely the integrity of the large-scale model.

Can parametrized convection help the model escape from this moist instability, given that it uses the same basic thermodynamics of latent heat? Indeed there is at least one very significant difference, namely the vertical transport property of parametrized (or indeed real) convection. In particular a mass flux scheme will normally allow some penetration into more stable regions where $\partial h/\partial p$ is unfavourable for positive feedback. In this way the build-up of h (and hence q) in the lower troposphere can be relieved, whilst the greater moist stability aloft controls the impact at higher levels.

So in summary a partial cloud scheme can tap into potential instability if $dh/dz <$

0, and progress to full-scale resolved convection if there is conditional instability, i.e. $dh_{sat}/dz < 0$. The Sobel-Bretherton limit in which this may happen is extreme, but relevant to gridpoint storms by arguments of §3. Some localized trigger is probably also required so that the resolved convection can occur at very small area fraction r (as convection over a large area would stabilize the profiles too much to allow this phenomenon).

Parametrized convection is predicted to suppress or reduce GPSs even though it obeys the same fundamental thermodynamics and certainly contributes to latent heating. The key difference between parametrized and resolved convection lies in the nature of vertical transport of h. Parametrized convection using mass-flux or other nonlocal scheme, like real convection, can transport h penetratively to levels where it is less 'damaging'. This theoretical prediction is consistent with the finding in practice that speeding up parametrized convection can suppress the GPSs. It also suggests that 'the more penetrative the better' in such circumstances.

Of course GPSs will be affected by various other model issues, including the accuracy of the advection of h, details of the timestepping, and the possible need for horizontal diffusion (perhaps viewed as 'horizontally penetrating convection'). However these moist static energy arguments combined with strong dynamical feedback give a plausible explanation of their basic dynamics and thermodynamics.

5.1.2 Resistance formulation for h – a radiative-convective equilibrium?

We note here briefly that a resistance loop can be written for moist static energy h, analogous to that written earlier for T. Neelin's argument however implies an important difference, i.e. that the horizontal coupling (the counterpart for h of $t_{\rm dyn}$) is much weaker than for temperature. (In fact one might use the assumption of negligible horizontal h-flux to extend our simple coupling model of §4, but we do not pursue that exercise here.)

In bulk column terms h is much closer to being in radiative-convective equilibrium than either T or q separately. At this point though 'moist static energy thinking' risks misleading us. Convection is not determined by h alone – for the same h we could have a mid-troposphere that was cold and moist (and hence unstable relative to the surface) or warm and dry (and hence much more stable).

Even if surface h-fluxes are constrained to balance radiative cooling, this does not necessarily regulate the convective mass flux. Raymond (1995) shows that $\overline{w'h'}$ scales on

$M_v \Delta h$ but Shutts & Gray (1999) show that Δh is an elastic quantity, which can change in response to other convective parameters. As with other tracer quantities, one expects vigorous convection to drive tropospheric h towards the surface value.

5.2 Wind-induced surface heat exchange (WISHE)

The feedback on convection via wind-induced surface heat-exchange (WISHE) can be estimated as follows, using moist static energy h as a measure of heat.

First we note two simple flow patterns satisfying continuity, (i) barotropic circulations essentially in the (x, y) plane and (ii) baroclinic circulations essentially in the (x, z) plane. Although in time mixed modes are likely to develop, these geometries may be treated as initial conditions for purposes of scaling estimates.

In the baroclinic geometry the gravest tropospheric modes, spanning the troposphere depth H, scale as $u \sim w(L_x/H)$. The extra 'wind-induced' flux of h may be estimated as

$$\rho c_h u (\Delta h)_{\mathrm{S-BL}} \sim \rho c_h w (\Delta h) L_x / H \tag{61}$$

where c_h is the surface-to-BL exchange coefficient and $(\Delta h)_{\mathrm{S-BL}}$ the surface-to-BL difference in h. We take typical values for scaling to be $c_h = 10^{-3}$, $H \sim 10$km.

If we compare the flux of h into the boundary layer by WISHE and out of the boundary layer by convection (cf. Raymond (1995)), and scale the latter on $\rho M (\Delta h)_{\mathrm{BL-FE}}$ (where $(\Delta h)_{\mathrm{BL-FE}}$ is the difference in h between boundary layer and free atmosphere) then the relative magnitude of the WISHE term is $c_h (L_x/H)(\rho w/M)(\Delta h)_{\mathrm{S-BL}}/(\Delta h)_{\mathrm{BL-FE}}$.

For convection on short horizontal scales (e.g. as discussed in §4) we expect $\rho w/M = O(1)$ to satisfy buoyancy compensation. The importance of the WISHE term is then seen to be linear in L_x. If further $(\Delta h)_{\mathrm{S-BL}}/(\Delta h)_{\mathrm{BL-FE}} = O(1)$ then we can expect WISHE term to become dominant for lengthscales $\gg H/c_h \sim 1000$km.

For a number of reasons we do not recommend incorporating WISHE in our convective-dynamical model for process-modelling applications. First, the scales where WISHE becomes dominant are somewhat larger than arising in buoyancy compensation scaling. On those larger scales, moreover, the dynamics of planetary waves come into their own. WISHE effects may well interact significantly with flow components on the planetary scale,

but those are beyond the scope of the present study (e.g. we have generally neglected β effects).

Secondly the surface transfer processes will normally form an integral component of a conventional process model and it would be questionable to couple such a process model to a large-scale feedback parametrization that made specific assumptions about surface transfer.

Thirdly, the WISHE effects depend very much on detailed geometry and location, and do not suggest simple systematic feedbacks that could be used in process models. If rotation can be neglected, the WISHE term has the effect of spatial propagation rather than amplifying or damping. For a small baroclinic zonal perturbation to a zonal basic flow, the WISHE effect will enhance moistening, convection and low-level convergence *upwind* of the initial convergence point, leading to upwind propagation. For an initially barotropic mode, the WISHE effect will tend to promote convection and ascent in the region of strongest surface wind, tending to introduce a baroclinic element to the mode.

Different behaviour could arise in a system with strong rotation. In an initially barotropic mode, convection could induce convergence in the cross-wind direction, which through Coriolis forces could be converted into a long-wind component. Such an atmospheric mode would then develop a mixed barotropic-baroclinic character. Some WISHE-convection-convergence-rotational feedback appears to play a role in the development of tropical cyclones. Indeed after significant cyclonic circulation develops, the rotational feedback is enhanced by the positive relative vorticity (in addition to planetary vorticity). However tropical cyclones are a special case in their geometry, intensity and other respects, and we do not advocate here any extension of the present methodology to 'convection in hurricanes'.

More detailed analysis of tropical cyclone intensification is given in Craig & Gray (1996), who find in their numerical study a strong positive dependence on heat and moisture transfer coefficients, but little sensitivity to frictional drag.

We conclude that WISHE is a significant feedback process in the large-scale atmosphere, but that from a process modelling standpoint it cannot be treated as systematic in the same way as the buoyancy feedback of §3 and is not appropriate for inclusion in convective-scale process studies as a parametrization of large-scale feedback.

5.3 Cloud-radiative feedbacks on convection

We now give a basic scaling for the impact of cloud-radiative feedbacks on convection, again focusing on the regulation of convection in process models. The nature of the underlying surface is important to this evaluation. We discuss two limiting cases (i) a surface of small heat capacity (e.g. land) and (ii) a surface of high heat capacity, e.g. the ocean. We focus here on short-wave radiation; the argument can be extended to long-wave. Furthermore we do not consider radiative interactions with individual clouds, regarding that as a detailed 'process study' issue rather than a question about convective regulation in the sense of the current study. (Petch & Gray 2001) give a detailed evaluation of interactive radiative impacts in a CRM case-study.

As the simplest relevant calculation (and probably the greatest possible impact) suppose that a thick cloud-shield reflects almost all of the incoming solar radiation $S_\downarrow$, of order $1000\mathrm{Wm}^{-2}$ to space. This component of cloud radiative forcing tends to lower the temperature of the atmosphere or surface or both. (We do not claim that short-wave radiative effects dominate over long-wave effects, but merely that this calculation indicates in principle how radiative effects may be scaled.)

Compare this with the column impact of the non-radiative convective heating Q_1, which can be approximated as the total latent heat of the precipitation P. We find then that the non-radiative component of the total heat budget exceeds the radiative component if $P > S_\downarrow/\lambda \simeq 0.7\mathrm{mm/hr}$ (a value reported as 'light rain' if observed locally). For this comparison the computation of P should be carried out over the area of optically thick cloud (or equivalent if there is a substantial radiative contribution from large-scale diffuse cirrus etc.).

In the land case (i) the surface $S_\downarrow$ can be assumed to translate into boundary layer sensible heat-flux and the above ratio $\lambda P/S_\downarrow$ gives a measure of the relative contributions to convection via the atmospheric heat budget.

In the oceanic case (ii) however the surface is well buffered against temperature changes and only the atmospheric absorption matters. The relevant parameter then becomes $\lambda P/\epsilon S_\downarrow$ where ϵ is a measure of bulk short-wave absorptivity for the atmosphere.

We suggest therefore that on process modelling timescales cloud-radiation is significant as a feedback over land, but not over ocean.

The argument changes of course on climate timescales, where the ocean no longer functions as a heat sink, but instead recycles heat to the atmosphere. The relevant timescale for recycling is then the time for surface fluxes to change ocean temperature sufficiently to feedback significantly on surface fluxes (via convection or otherwise). E.g. a surface flux of $10^3 \mathrm{Wm}^{-2}$ would take about 4days to change the temperature of a 100m ocean 'mixed layer' by 1K.

5.4 Feedbacks via convective momentum transport

The last feedback we consider between convection and large-scale flow is that of convective momentum transport.

Our estimates for convective momentum transport (CMT) are based on Gregory, Kershaw & Inness (1997). This scheme is an extension to the UM mass-flux convection scheme.

First of all we choose a typical value of cloud-base mass-flux M over a grid square of side 100km to be $\sim 0.1 \mathrm{kg\ m}^{-2}\mathrm{s}^{-2}$. Taking a typical boundary-layer humidity q_{BL} around ~ 10g/kg, and the density of liquid water $\rho_w = 10^3 \mathrm{kg\ m}^{-3}$, this implies that the moisture extraction from the boundary layer corresponds in rough scaling terms to a rain-rate $Mq_{BL}/\rho_w \sim 3.6$mm/hr.

The Gregory et al. (1997) scheme treats CMT in a similar manner to convective scalar transport, except for a term describing the environmental pressure-reaction on the cloudy updraughts. The pressure-reaction is characterized by a nondimensional coefficient $C^u \sim 0.7$ (as evaluated using a Cloud-Resolving Model). Roughly speaking we can approximate the momentum flux τ by

$$\tau \sim M \Delta U (1 - C^u) \tag{62}$$

where ΔU is the velocity difference across the relevant layer. Taking $\Delta U = 10 \mathrm{ms}^{-1}$ and the other numbers as above, gives the estimate $\tau \sim 0.3 \mathrm{Nm}^{-2}$, a value comparable with typical boundary-layer stresses.

The response to given convective or boundary-layer stresses may be estimated from (82). The estimate is simplest on the largest scales ($L_H > L_R$) where the static stability does not significantly resist the 'Ekman-pumped' vertical motion ω_{Ek}.

In this large-scale regime

$$\omega_{\rm Ek} \sim \tau_p / f r_c \tag{63}$$

where r_c is the radius of curvature of the synoptic flow. For scaling purposes a Boussinesq or anelastic approximation suffices, whence we can replace the pressure-coordinate form by its near-equivalent form

$$\rho w_{\rm Ek} \sim \tau / f r_c \tag{64}$$

Substituting $\tau \sim M \Delta U (1 - C^u)$ as above gives

$$\rho w_{\rm Ek} \sim Ro (1 - C^u) M \tag{65}$$

Here $Ro = \Delta U / f r_c$ is essentially a standard synoptic Rossby-number, but with a velocity scale based on vertical differences across the convective layer. It follows from (65 that in the low-Ro regime which governs most midlatitude weather systems, the Ekman mass flux will be small compared to the convective mass flux.

On smaller scales $L_H \leq L_R$, i.e. $Bu = L_R / L_H \geq O(1)$, the static stability will reduce the Ekman-pumping from the simple rotational balance of (83). Hence we can use the large-scale approximation (65) as a general upper bound to the Ekman impact of CMT. In fact from our analysis in the Annex, the Ekman-pumping is strongly reduced at high Bu, owing to a static-stability correction factor $\sim Bu^{-2}$ in that limit.

Finally we consider the case of high Rossby-number, e.g. in the deep tropics where f is small. Here we can use the algebraic relation

$$Ro = Bu \, Ri^{-1/2} \tag{66}$$

where the bulk Richardson-number $Ri \equiv N^2 H^2 / (\Delta U)^2$. We can assume $Ri \geq O(1)$, otherwise (if nothing else) shear turbulence would act to raise Ri. The high-Bu case can be discounted owing to the static-stability correction factor. Hence, for our present purposes in setting upper limits to CMT impact via Ekman pumping, we can effectively discount the high-Ro case. Even in a tropical cyclone, where low Ro may occur, this argument predicts that convection feeds back on the dynamics primarily via buoyancy compensation rather than by Ekman pumping. This prediction is consistent with the detailed numerical study of Craig & Gray (1996), who find little sensitivity of tropical cyclone intensification to frictional drag.

Taking into account the factor $(1 - C^u)$ in (65), this implies that in all relevant atmospheric regimes the Ekman mass flux pumped by CMT will be significantly smaller

than the convective mass flux, and will play only a minor role in the type of feedback discussed in this paper. This finding does not contradict the importance of CMT in large-scale modelling, but suggests that we do not need to model explicitly the large-scale feedback via CMT in process models. Sensitivity tests in Met Office global models confirm that the impact of CMT is subtle and does not dramatically change circulation patterns (S.Milton, personal communication).

6 Application to further process-modelling issues

Having shown that dynamical feedback can make convection more sensitive to adjustment timescales, we now consider some more complex process issues. Specifically we discuss the impact of mid-tropospheric humidity on convection, the convective diurnal development over tropical islands and the impact of latent heat assimilation in NWP models.

6.1 Impact of mid-tropospheric humidity

There is observational evidence, e.g. Johnson (1997), for associations between mid-tropospheric humidity and convection. In the TOGA-COARE tropical experiment, composite humidity profiles for the stronger, more organized systems are systematically moister than for the weaker convection cases.

However in order to develop parametrizations based on such associations, we need to understand the causal mechanisms. The observations could be interpreted in terms of the suppression of convection by entrainment of dry air, but clear-air radiation may also play a role and in some cases moistening could be seen as a consequence of convection.

Full analysis of these mechanisms may require an idealized cloud-resolving model study. However the most idealized model, that of radiative-convective equilibrium, does not allow us to control the moisture profiles. Instead the equilibrium profiles emerge diagnostically from the balance between radiation and convection. In the real tropical atmosphere, however, as argued above, the large scale dynamics play a key role and the mean T, q profiles of a convecting column may be controlled not by the local convection but by the larger-scale atmosphere.

Motivated by these considerations, a group within the international EUROCS project

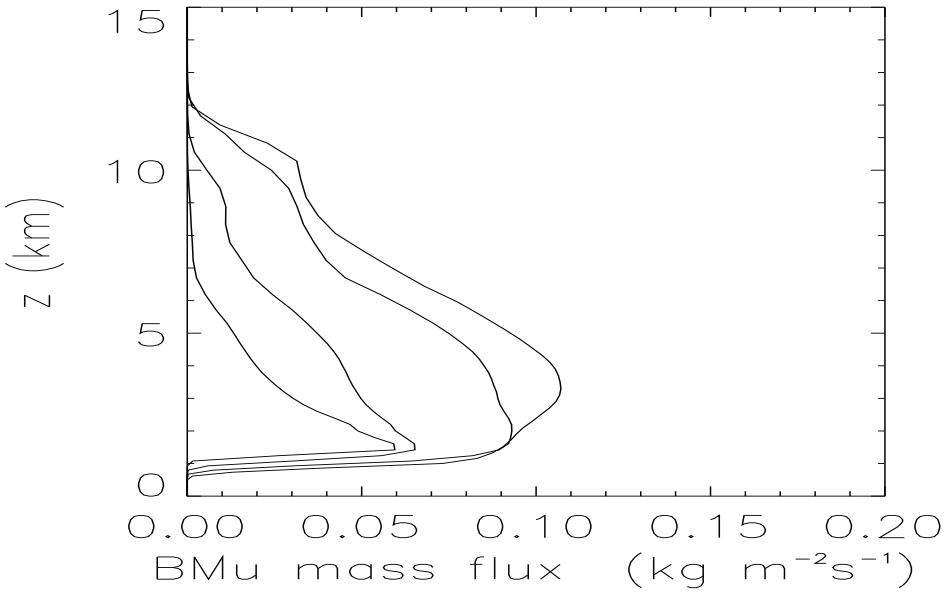

Figure 4: Profiles of quasi-steady buoyant moist updraught mass-flux for Met Office Cloud-Resolving Model (for four prescribed values of the target relative humidity RHt).

(see http://www.cnrm.meteo.fr/gcss/EUROCS/EUROCS.html) is currently studying the response of convection to mid-tropospheric humidity. The methodology of this 'idealized humidity' project is based on a simple representation of large-scale feedback via nudging of mean temperature and humidity on a 1-hour timescale. Four quasi-steady runs were made with target profiles based respectively on relative humidity values (RHt) of 25%,50%,70% and 90% at heights above 2km.

This prescribed nudging does not of course constitute a complete dynamical feedback model, as outlined in the present paper, since the large-scale moisture forcing is not derived from a model of vertical motion. Nevertheless the impacts shown in Fig. 4 show that even a partial version of dynamical feedback can enable a quasi-steady convecting ensemble to show sensitivities that do not otherwise appear.

6.2 Convection over tropical islands

From §4, we can expect dynamical feedback to play an important role in determining the amplitude of convection over sea surface temperature anomalies, whenever the horizontal scale of the anomaly is small compared to a Rossby radius. In particular such feedback enables convective activity to circumvent the limit set by 'external' forcing in the local column.

Similar arguments carry over to the problem of convection over tropical islands. However the island problem involves stronger time dependence, being driven by diurnal surface heating. In the inner tropics the diurnal insolation cycle is a faster process than inertial oscillations, and the intrinsic dynamical lengthscale for diurnal convection then becomes the distance which gravity waves can travel over the diurnal timescale.

Consider for instance the island of New Guinea, with horizontal dimensions of order 1000km × 500km (comparable in fact to the British Isles). The New Guinea orography (rising up to 2km above sea level) is known to be able to generate spurious 'gridpoint storms' in some large-scale model formulations. Whatever physical, dynamical or numerical model weaknesses contribute to this phenomenon, one may perhaps interpret such gridpoint storms as tapping into the capacity of tropical dynamics to support sustained convective events.

Close to the Equator the appropriate dynamical scale for response to *steady* forcing would be several thousand kilometres. This scale is comparable to the Hadley and Walker circulations, and also to the range of influence of the S.Asian monsoon heating (Rodwell and Hoskins).

From our analysis in §3 it can be seen that the impact of unsteady forcing is effectively to change the Coriolis parameter to an equivalent value $(f^2 + \sigma^2)^{1/2}$. Here σ may be imaginary (sinusoidal forcing) but using real σ in the same equations we can also compute the far field of exponentially growing localized disturbances.

The primary component of diurnal forcing corresponds to $\sigma = 2\pi i/24\text{hr}$, although there will also be higher frequencies (harmonics) and an important steady component. In fact the f-plane dynamical analysis shows a resonance at latitude 30°, which in principle would give a particularly large dynamical feedback to diurnal forcing. However a more complete dynamical model, with latitude variation of f, is likely to show less strong resonance, as there is then no unique inertial resonant frequency and Rossby wave effects

will also complicate the problem.

At low latitudes, if rotating effects can be neglected, the diurnal value of σ allows gravity-waves to propagate a distance of order 500km, comparable with the linear dimensions New Guinea. In fact many gridpoint storms seem to grow with an e-folding time of the same magnitude (i.e. $\sigma \sim 10^{-4}\mathrm{s}^{-1}$) so their domain of significant dynamical influence may also be of order 500km.

In summary, over tropical islands we can expect large-scale dynamical feedback to play a key role both in real convection and in any spurious convection present in a model. Some of the gridpoint storms seen in models can give local precipitation accumulations far greater than the local column moisture, showing that they have drawn in moisture from many columns. Our simple feedback model obviously cannot capture in detail the subtleties of sea-breeze triggering and other detailed interactions between topography and convection, but it can simulate the basic ability of large-scale ascent to overcome single-column stabilization.

6.3 Impact of latent heat assimilation

Jones & Macpherson (1997) show that estimates of latent heating from weather-radar observations can be assimilated usefully into a mesoscale NWP model, whilst Pullen & Butterworth (2001) have sought to extend this approach to assimilate satellite estimates of tropical convective rainfall. In each case 'usefully assimilated' implies that such latent heating can at least help maintain precipitation in a location where it is observed.

Whilst detailed consideration of these techniques lies outside the scope of the present note, some aspects of the principle are highly relevant to our argument. For traditional SCM thinking would suggest that assimilation of latent heating would act purely to *suppress* model convective precipitation. This is because in a classical SCM scenario such precipitation and latent heating respond passively to profile destabilization controlled by larger-scale dynamics. In such a scenario latent heat assimilation could not work as intended.

Our resistance model offers some pointers to what is happening in the NWP implementation of latent heat nudging (LHN), as opposed to a SCM test. We assume that latent heat nudging is applied to the bulk atmospheric temperature in 'column 1' in the terminology of §4.

First, in the simple case of prescribed resistances (including fixed temperature adjustment timescales) the electrical analogy predicts that the additional 'input current' from LHN is partitioned in the ratio $R_2 + R_3 : R_1$ between vertical:horizontal transfer. Thus

(i) if R_1 is small we recover the SCM case where convective Q_1 should adjust negatively to compensate (the opposite of what is desired)

(ii) if $R_2 + R_3$ is small or O(1) then the impact on local Q_1 is reduced but still positive

So the simple electrical analogy model still predicts that LHN should not work.

However we can understand why the simple (fixed-resistance) model does not fully capture the behaviour in NWP. In §4.3 we noted that convective 'resistance' could depend on w or q, indeed perhaps triggering quite sharply at threshold values. The dynamical compensation for positive LHN will imply positive w and positive q-forcing and shows that LHN could work via a triggering effect if $(R_2 + R_3)/R_1$ is sufficiently small, i.e. if column 1 is small compared to the area of dynamical compensation.

So our theoretical analysis predicts that the usefulness of LHN depends on at least two conditions. First, the nudging should be on relatively small scales. Secondly the convection scheme should be susceptible to triggering, either directly by w or by q. A third condition may be added, namely that the scales assimilated should not be too small for model numerics and dynamics to cope at least qualitatively (e.g. there is little point in inserting detail which merely contributes to gridscale noise or gridpoint storms).

This analysis seems to help explain the findings of Pullen & Butterworth (2001). In assimilating tropical measurements, they found that whilst nudging seemed to be helpful for the location of convection, it led to an overall reduction of model convective activity in the assimilation region. Based on our analysis, some band-pass filter should ideally be applied to LHN so that it acts only on scales which are small compared to the dynamical scales, but also avoids forcing the gridscale.

Clearly there are other physical questions about LHN which are beyond the scope of this note.

7 Conclusions

Systematic buoyancy feedbacks between convection and larger-scale dynamics, as discussed in this note, can significantly affect our approach to convective-scale process studies, whether these involve idealized problems or observed cases.

Such process studies normally involve choices about how to specify boundary conditions and related quantities. Here we are concerned with lateral boundary conditions, or the domain-averaged representations of their gross effects (large-scale forcings as 'pseudo-BCs') which are applied in convective process studies. In observed cases model specification may be influenced by an assessment of measurement accuracy, but such specification also has important implications for the nature and sensitivity of model behaviour.

The conventional specification of large-scale forcing via prescribed large-scale temperature and other tendencies can force the convection over time to balance that forcing. Such a prescription systematically diminishes the sensitivity of the process model to certain process issues (including various forms of convective inhibition), relative to the behaviour in a large-scale model. Conventional 'large-scale forcing' strategy therefore should be complemented by an alternative strategy for including in process studies some measure of large-scale feedback, insofar as this is systematic and predictable from the most basic parameters.

The importance of convective-dynamical feedback depends on the numerical magnitudes of these effects. We showed that the feedback could be conveniently described either by an effective area-fraction r (the fraction of convective heating not compensated by the large scales) or by a buoyancy-relaxation timescale t_{dyn}. We showed further that $t_{dyn} \sim \sigma^{-1} r/(1-r)$, where σ^{-1} is the characteristic overall timescale. So for instance on long timescales σ^{-1} we will have large t_{dyn} unless r is correspondingly small.

In §3 we discussed both balanced and unbalanced problems. The balanced problems can approach a quasi-steady state, in which pressure forces are balanced by Coriolis terms, leading to a range of dynamical influence scaling on the Rossby radius L_R. This type of problem can yield non-zero values of r and t_{dyn}. A scenario in which friction balances the pressure terms is also conceivable, although less realistic. In an unbalanced type of problem the time-development continues without approaching steadiness, typically through long-range propagation of gravity-waves.

We solved test problems in both 2D and 3D to determine the feedback parameters

r and t_{dyn}. In both cases r is determined by the ratio of the area of convective forcing to the area associated with the dynamical response range L_R. Taking into account the relationship between r and t_{dyn}, we found that in 2D the feedback timescale t_{dyn} scales with the gravity-wave crossing time but in 3D it does not (being significantly smaller).

We showed further that when t_{dyn} is small compared to the timescale governing the strength of feedbacks by vertical transport, or when these timescales are of the same order, the feedback can significantly affect the physical behaviour in ways not captured by conventional SCM forcing.

The recognition of dynamical feedback as *systematic* has conceptual implications for process modelling. It may be better to regard a SCM (despite the well established name) as fundamentally a 1D version of the large-scale model rather than strictly a 'single column' of the latter at some specified horizontal resolution. We make this subtle, perhaps pedantic distinction because SCMs (as normally run) cannot capture this systematic interaction with the larger scales. Numerical details of discretization in the large-scale model also militate against precise identification of the SCM with a large-scale 'single column'.

Our arguments do not invalidate observational case comparisons between SCMs, CRMs and observations based on prescribed measured forcings. Assuming the measurement is sufficiently accurate, these comparisons remain a valid and important check on our models. However we should note that systematic dynamical feedback may lead to systematic differences in the sensitivity to physics between the SCM and the parent GCM. The sensitivity of a given testbed is important especially if we go beyond 'validating' the physics against observed cases to comparing different schemes in more idealized problems.

Conventional prescribed forcing in SCMs is most applicable to a convective ensemble that is coherent over at least the length-scale of dynamical response. When buoyancy adjustment is balanced by Coriolis terms, this scale is given by a Rossby radius L_R, of order 1000km in midlatitudes and larger in the tropics. In the unbalanced case the scale is given by the gravity-wave propagation distance. On scales shorter than this, feedback can play a significant role depending on the values of t_{dyn}. Separately we may also stipulate that some of the physics (e.g. convection) are deemed to represent averages over certain scales (e.g. some cloud-system scale) and cannot be sensibly compared on shorter scales.

Similarly a CRM process study cannot be precisely identified with a gridbox-average of the large-scale model. Instead, we argue that process studies justifiably make idealizations, which are valid when carried through consistently, but sometimes require alterna-

tive assumptions to be tested. We suggest that a simple parametrization of dynamical feedback is a useful alternative to the neglect of feedback, such neglect being itself an idealization. Woolnough (2001) shows how this parametrization can be implemented in a CRM and how it affects convection.

We gave further illustrations in §6 of how consideration of large-scale feedback can bridge the gap between convective process studies and the behaviour within an NWP model. For instance a pragmatic procedure to assimilate convective latent heating into a large-scale model has been developed and used in a large-scale model. Under conventional SCM thinking this is paradoxical: would not such assimilated heating act to suppress convection, rather than enhancing it as desired? But here our feedback analysis gives us an outline explanation of how, and on what scales, this assimilation procedure can work.

In §3 and §5 we discussed gridpoint storms (or rather column storms) as found in many large-scale models, and gave a theoretical explanation which seems to fit with current practical experience.

Dynamical analyis confirmed that these storms fall into the regime $r \to 0$, and can reasonably be considered in terms of Sobel-Bretherton dynamics. Using S-B dynamics combined with moist static energy arguments adapted from Neelin, we showed how an unstable feedback between model dynamics and large-scale condensation scheme can arise.

Essentially a partial cloud scheme can tap locally into potential instability (if present). This mechanism only works on very isolated gridpoints and therefore probably requires some local trigger. The mechanism does not necessarily require conditional instability but a different (and weaker condition) sometimes called 'potential instability'.

We suggest that parametrized deep convection acts to suppress GPSs not so much through precipitation (which, after all, implies further latent heating), but through penetrative vertical transport, carrying excess moist static energy to higher, more stable levels where it is less damaging to the model. The behaviour of shallow convection, if active in this context, is not yet clear and deserves further study.

8 Annex: derivation and interpretation of ω-equation

In this Annex we review the forcing of vertical motion by convection using a fully time-dependent 'ω-equation', i.e. an equation for diagnosing 'pressure velocity' $\omega \equiv Dp/Dt$ from other forcings or fields. Quasi-geostrophic versions of the ω-equation are derived in many textbooks but here we derive for reference a fully time-dependent version with convective forcings of heat and momentum, which enables us to assess convective-dynamical interactions at all latitudes. In our pressure coordinates we use the notation ∇_H to denote 'horizontal' vector gradients at constant p.

For simplicity we linearize about a basic state at rest (thereby losing[1] some advection terms) but we include the effects of the planetary vorticity gradient β. The standard β-plane approximation effectively replaces the Earth's angular velocity $\mathbf{\Omega}$ by its vertical component, written here as $\mathbf{f}/2 = (0, 0, f/2)$. The primitive equations then become

$$\text{horizontal momentum: } \partial_t \mathbf{u} = -\mathbf{f} \wedge \mathbf{u} - \nabla_H \Phi \tag{67}$$

$$\text{buoyancy: } \partial_p \Phi = -1/\rho = -RT_v/p = -R\Pi\theta_v/p \tag{68}$$

(see Notation). Then $\partial_t(68) \rightarrow$

$$\partial_t \partial_p \Phi = -(R\Pi/p)\partial_t \theta_v = (R\Pi/p)(\partial_p \theta_v)\omega = -s^2\omega \tag{69}$$

where

$$s^2 \equiv -(R\Pi/p)\partial_p \theta_v = N^2/(\partial_z p)^2 \tag{70}$$

We now modify the primitive equations by adding convective source terms $\mathbf{F}$ and $B = (R\Pi/p)D\theta_v/Dt$, so that

$$\partial_t \mathbf{u} = -\mathbf{f} \wedge \mathbf{u} - \nabla_H \Phi + \mathbf{F} \tag{71}$$

and

$$\partial_t \partial_p \Phi = -s^2\omega - B \tag{72}$$

Then $\nabla_H.(71) \rightarrow$

$$\partial_t \nabla_H.\mathbf{u} = -\nabla_H.(\mathbf{f} \wedge \mathbf{u}) - \nabla_H^2 \Phi + \nabla_H.\mathbf{F} \tag{73}$$

$$= -\mathbf{f}.\nabla_H \wedge \mathbf{u} + \beta u - \nabla_H^2 \Phi + \nabla_H.\mathbf{F} \tag{74}$$

since

$$\nabla_H.(\mathbf{f} \wedge \mathbf{u}) = -\mathbf{f}.\nabla_H \wedge \mathbf{u} + \mathbf{u}.\nabla_H \wedge \mathbf{f}$$

[1] actually the advection terms are easily reinstated by reclassifying them as forcings

and

$$\mathbf{u}.\nabla_H \wedge \mathbf{f} = \beta u \tag{75}$$

Differentiating further, $\partial_p \partial_t (74) \rightarrow$

$$\partial_t^2 \partial_p \nabla_H.\mathbf{u} = \partial_p[\mathbf{f}.\nabla_H \wedge \partial_t \mathbf{u}] - \nabla_H^2 \partial_t \partial_t \Phi + \partial_t \partial_p (\nabla_H.\mathbf{F} + \beta u) \tag{76}$$

and substituting $\nabla_H.\mathbf{u} = -\partial_p \omega$ (continuity) on the left-hand side and $\partial_t \mathbf{u}$=(71) on the right-hand side we obtain

$$-\partial_t^2 \partial_p^2 \omega = \mathbf{f}.\nabla_H \wedge \partial_p[-\mathbf{f} \wedge \mathbf{u} - \nabla_H \Phi + \mathbf{F}] + \nabla_H^2 (s^2 \omega + B) + \partial_t \partial_p (\nabla_H.\mathbf{F} + \beta u) \tag{77}$$

which by using the vector identities $\nabla_H \wedge \nabla_H \Phi \equiv 0$ and

$$\nabla_H \wedge (\mathbf{a} \wedge \mathbf{b}) = \mathbf{a}(\nabla_H.\mathbf{b}) + (\mathbf{b}.\nabla_H)\mathbf{a} - \mathbf{b}(\nabla_H.\mathbf{a}) - (\mathbf{a}.\nabla_H)\mathbf{b}$$

(with $\mathbf{a} = -\mathbf{f}$ and $\mathbf{b} = \partial_p \mathbf{u}$) we can write as

$$\begin{aligned} -\partial_t^2 \partial_p^2 \omega &= \mathbf{f}.[-\mathbf{f} \nabla_H.(\partial_p \mathbf{u}) + \partial_p \nabla_H \wedge \mathbf{F}] - f\beta \partial_p v + \nabla_H^2 (s^2 \omega + B) + \partial_t \partial_p (\nabla.\mathbf{F} + \beta u) \\ &= \mathbf{f}.(\mathbf{f} \partial_p^2 \omega + \partial_p \nabla_H \wedge \mathbf{F}) - f\beta \partial_p v + \nabla_H^2 (s^2 \omega + B) + \partial_t \partial_p (\nabla_H.\mathbf{F} + \beta u) \end{aligned} \tag{78}$$

which we can finally write, collecting source terms on the right, as

$$[(f^2 + \partial_t^2)\partial_p^2 + s^2 \nabla_H^2]\omega = -\nabla_H^2 B - \mathbf{f}.\partial_p (\nabla_H \wedge \mathbf{F}) + f\beta \partial_p v - \partial_t \partial_p (\nabla_H.\mathbf{F} + \beta u) \tag{79}$$

8.1 Interpretation and scale analysis of the ω-equation

As compared to the quasi-geostrophic (QG) version (James 1994) our ω-equation (79) includes some ∂_t terms, which strictly make this a prognostic rather than diagnostic equation. Here though we consider those terms only for scaling purposes. Essentially these additional terms are negligible compared to the Coriolis terms if $\partial_t \ll |f|$. By contrast in the opposite limit, where rotation is slow compared to the fluid timescales, the ∂_t terms dominate the ω-equation.

In the deep tropics the Coriolis parameter f is not usually 'large' in the sense required for QG theory. For instance at 5° latitude the inertial period is about 5 days, so the QG approximation is unreliable for waves on sub-weekly timescales.

In the absence of 'forcing', i.e. setting the RHS of (79) to zero, we recover the standard dispersion relation for inertia-gravity waves. In addition to the Coriolis and time-derivative terms, the static-stability term on the left-hand side of (79) is also important at shorter wavelengths.

In order to describe the scales on which the static-stability is important, let us write

$$f^2\omega/s^2\nabla_H^2\omega \sim L_H^2/L_R^2 \tag{80}$$

where L_H is the horizontal scale, Δp the pressure-depth scale of the disturbance (usually related to the troposphere depth) and $L_R = s\Delta p/|f| \sim NH/|f|$ is the Rossby-radius of deformation. The ratio $Bu = L_R/L_H$ is called the Burger number, and the low-Bu regime may be viewed as applying to broadscale barotropic waves. Typically though, most tropical and midlatitude weather systems involve scales shorter than L_R, in which static-stability is as important as the Coriolis term, i.e. $Bu \geq O(1)$.

Amongst other mechanisms our ω-equation includes the 'Ekman pumping' effect. The Ekman effect, in its simplest terms, corresponds to a balance between the Coriolis force acting on the ageostrophic wind and the frictional acceleration $\mathbf{F}$ (due to boundary layer turbulence or convection). This leads in (79) to a balance between $f^2\partial_p^2\omega$ and $-\mathbf{f}.\partial_p(\nabla_H \wedge \mathbf{F})$. More generally let $\omega_{\rm Ek}$ be the forced response to a steady friction $\mathbf{F} = \partial_p\tau_p$ where

$$\tau_p = -\overline{(u'\omega', v'\omega')} \tag{81}$$

Then when the forcing scale is large compared to the Rossby radius ($Bu = NH/L_H|f| \ll 1$) we have

$$\omega_{\rm Ek} \sim f^{-1}(\nabla_H \wedge \tau_p)_3 + D \tag{82}$$

Here D is a compensating term linear in p which represents Dines compensation, and enables us to satisfy $\omega = 0$ at the top and bottom (or some alternative condition) even if $\mathbf{F}$ does not vanish at the bottom. As a linear, boundary-compensating term, D does not affect the scaling.

However where the forcing scale is smaller (moderate or high Bu) then static stability terms enter as we saw earlier for buoyancy forcing. The scaling then changes to

$$\omega_{\rm Ek} \sim f^{-1}(1 + Bu^2)^{-1}(\nabla_H \wedge \tau_p)_3 + D \tag{83}$$

It follows that for a given frictional forcing and given NH, the Ekman pumping $\omega_{\rm Ek}$ is maximized at a value of f such that $L_H = O(L_R)$, and that actually $\omega_{\rm Ek} \to 0$ when $f \to 0$.

The terms in the planetary vorticity gradient β come into their own when the meridional scale is 'planetary', or at least a few degrees of latitude. These terms arise from two stages in our derivation, but can be written together as $-f\partial_p(\beta u) + \partial_t\partial_p(\beta v)$.

The term $-f\partial_p(\beta u)$ reflects a tendency for Westerly flow to induce convergence and Easterly flow divergence, owing to the variation in f. Consider for instance Westerly flow near the Equator. To the North $f > 0$ and the flow turns to the right, whilst to the South the flow turns to the left, hence convergence results. Such effects militate against simple barotropic flow solutions of the kind seen in f-plane dynamics. Consequently on scales where the β-effect is significant, there is a tendency for mixed barotropic-baroclinic behaviour, especially when the characteristic lengthscale for β effects becomes smaller than the Rossby radius. See Gill (1982), p.588, and references.

The term $f\beta\partial_p v$ in (79) corresponds to the forcing of vertical motion by vertically non-uniform advection of the planetary vorticity gradient. In the quasi-geostrophic limit the vertical motion acts to balance differential vorticity advection together with other terms that tend to disturb thermal wind balance. The tendency of a cyclonic vorticity anomaly in the upper troposphere to induce ascent in the mid-troposphere (in the presence of planetary vorticity f) may be termed the 'vacuum cleaner effect'.

The case where βv balances potential vorticity forcing is known as a Sverdrup balance, particularly when the forcing is 'frictional', i.e. $\beta v = \nabla_H \wedge \mathbf{F}$. Equation (79) incorporates these terms differentiated with respect to p and multiplied by f. The factor f implies that the horizontal 'restoring tendency' of potential vorticity is slow at low latitudes, even though β is at its maximum there.

The Sverdrup response to diabatic heating in the mid-troposphere is slightly counter-intuitive. Such heating tends to reduce upper-tropospheric static stability (and hence PV), inducing equatorward flow, whilst low-level PV is increased and equilibrates by poleward flow. Hence diabatic heating close to the Equator can induce a thermally indirect flow component in the meridional plane. However this does not mean that the induced circulation as a whole is thermally indirect. An example of this behaviour occurs in the Gill problem, as discussed above in §3.

In (79) we wrote the β terms as 'forcings', although evidently a full solution involves relationships with the other terms through flow geometry and continuity. Detailed analysis shows a complex set of wave solutions, including Rossby- and Kelvin-waves as discussed in Gill (1982), p.438. Kelvin waves may be viewed as a form of equatorially trapped gravity waves with flow in the zonal plane, whereas Rossby waves primarily involve meridional

advection of planetary vorticity. Both types of waves permit ascent in convecting regions to be compensated by remote subsidence.

Gill (1982), p.454, says that the adjustment under gravity near the Equator can be thought of as a two-stage process, with an initial fast adjustment due to gravity waves followed by a slow adjustment due to planetary waves. Although planetary waves undoubtedly play an important role in propagating the influence of convection, we shall assume that their detailed properties are not critical to the feedbacks of present interest.

Notation

x, y, z	standard Cartesian coordinates (righthanded with z vertical)
u, v, w	corresponding velocity components
t	time
p	pressure
ρ	density
T	absolute temperature
Φ	geopotential
∇_H	Horizontal derivative on p-surfaces
∂_t, ∂_p etc.	partial derivatives with respect to t, p etc.
D/Dt	Lagrangian (material) time derivative
Π	Exner function $(p/1000\text{hPa})^{R/c_p}$
R	gas constant
c_p	specific heat capacity of air at constant pressure
ω	'pressure velocity' Dp/Dt
f	Coriolis parameter
β	df/dy where y is taken as the latitudinal direction
N	Brunt-Vaisala frequency (typically 10^{-2}s^{-1}) in the free troposphere
B	buoyancy forcing in pressure units
R_B	the part of B not compensated by vertical motion
$s = N/(-\partial_z p)$	parameter for static stability in pressure coordinates (typically $10^{-3}\text{mPa}^{-1}\text{s}^{-1}$)
t_{dyn}	effective dynamical response timescale for our feedback model
r	effective fractional area parameter for our feedback model
k, l, m	wavenumbers in x, y, z directions respectively
k_r	characteristic wavenumber for dynamic response to local forcing
σ	growth rate (implying time-dependence $e^{\sigma t}$)
ϕ	general scalar quantity
H	height scale for convecting layer
L_H	horizontal scale
$q_{sat}(T, p)$	saturation value of specific humidity
L_R	$NH/\lvert f\rvert$, Rossby radius of deformation
$Bu = L_R/L_H$	Burger number
$Bu_1 = k/k_r$	modified Burger number for §3.1
Q	nondimensional heating in the Gill problem
Q_1	apparent convective temperature-source (contribution to $\partial_t T$)
Q_R	radiative temperature forcing
M	convective mass flux
T_{SST}	atmospheric temperature in equilibrium with given sea surface temperature

$\mathbf{F}$	frictional force
τ	frictional stress
Ro	Rossby number U/fL_H where U is a velocity scale
Ri	gradient Richardson number
ω_{Ek}	Ekman-pumped contribution to pressure-velocity ω
w_{Ek}	Ekman-pumped contribution to vertical velocity w
C^u	coefficient governing momentum transport in Gregory-Kershaw scheme

Acknowledgments

We are grateful to Brenda Cohen for supplying some of her results, and to Glenn Shutts, Mike Gray, Alison Stirling, Olaf Stiller, Jon Petch, Gill Martin, Stuart Webster, Mark Rodwell and Damian Wilson for relevant discussions. We thank Roy Kershaw for his careful review of the draft.

References

Abramowitz, M. & Stegun, I. 1970 , Handbook of mathematical functions, Dover, New York.

Blyth, E. M., Dolman, A. J. & Wood, N. 1993 , Effective resistance to sensible- and latent-heat flux in heterogeneous terrain., *Quart. J. R. Met. Soc.* **119**, 423–442.

Bretherton, C. S. 1993 , The nature of adjustment in cumulus cloud fields, *in* K.Emanuel & D.Raymond, eds, 'The representation of cumulus convection in numerical models', Amer. Met. Soc., Boston, Mass., pp. 63–74.

Cohen, B. G. 2001 , Fluctuations in an ensemble of cumulus clouds, PhD Dissertation, University of Reading.

Craig, G. C. & Gray, S. L. 1996 , CISK or WISHE as the mechanism for tropical cyclone intensification, *J. Atmos. Sci* **23**, 3528–3540.

Emanuel, K. A. 1994 , Atmospheric convection, Oxford University Press, Oxford.

Gill, A. E. 1980 , Some simple solutions for heat-induced tropical circulations, *Quart. J. R. Met. Soc.* **106**, 447–462.

Gill, A. E. 1982 , Atmosphere-ocean dynamics, Academic Press, London.

Grabowski, W. W. 2001 , Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP), *J. Atmos. Sci* **58**, 978–997.

Gregory, D., Kershaw, R. & Inness, P. M. 1997 , Parametrization of momentum transport by convection. ii: Tests in single-column and general circulation models., *Quart. J. R. Met. Soc.* **123**, 1153–1183.

Guichard, F., Redelsperger, J.-L. & Lafore, J.-P. 1996 , The behaviour of a cloud ensemble in response to external forcings, *Quart. J. R. Met. Soc.* **122**, 1043–1073.

James, I. N. 1994 , Introduction to Circulating Atmospheres, Cambridge University Press (Cambridge), 422pp.

Jin, F. & Hoskins, B. J. 1995 , The direct response to tropical heating in a baroclinic atmosphere, *J. Atmos. Sci* **52**, 307–319.

Johnson, R. H. 1997 , Recent observations of deep convection: TOGA-COARE, *in* ECMWF, ed., 'New insights and approaches to convective parametrization', ECMWF, pp. 1–24.

Jones, C. & Macpherson, B. 1997 , A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model, *Meteorol. Appl.* **4**, 269–277.

Neelin, J. D. 1997 , Implications of convective quasi–equilibrium for the large-scale flow, *in* R. K. Smith, ed., 'The Physics and Parameterization of Moist Atmospheric Convection', Kluwer Academic Publishers, pp. 413–446.

Nilsson, J. & Emanuel, K. A. 1999 , Equilibrium atmospheres of a two-column radiative convective model, *Quart. J. R. Met. Soc.* **125**, 2239–2264.

Petch, J. C. & Gray, M. E. B. 2001 , Sensitivity studies using a cloud-resolving model simulation of the tropical west pacific, *Quart. J. R. Met. Soc.* **127**, 2287–2306.

Pullen, S. & Butterworth, P. 2001 , Assimilation of satellite-derived estimates of tropical convective rainfall, NWP Division Technical Note No. 360.

Randall, D. A., Xu, K.-M., Somerville, R. J. C. & Iacobellis, S. 1996 , Single-column models and cloud ensemble models as links between observations and climate models, *J. Climate* **9**, 1683–1697.

Raymond, D. J. 1995 , Regulation of moist convection over the West Pacific warm pool, *J. Atmos. Sci* **52**, 3945–3959.

Raymond, D. J. & Emanuel, K. A. 1993 , The Kuo cumulus parametrization, *in* K.Emanuel & D.Raymond, eds, 'The representation of cumulus convection in numerical models', Amer. Met. Soc., Boston, Mass., pp. 145–147.

Raymond, D. J. & Zeng, X. 2000 , Instability and large-scale circulations in a two-column model of the tropical troposphere, *Quart. J. R. Met. Soc.* **126**, 3117–3135.

Riehl, H. & Malkus, J. S. 1958 , On the heat balance in the equatorial trough zone, *Geophysica* **6**, 503–558.

Rodwell, M. & Hoskins, B. 1996 , Monsoons and the dynamics of deserts, *Quart. J. R. Met. Soc.* **122**, 1385–1404.

Shutts, G. J. & Gray, M. E. B. 1999 , Numerical simulations of convective equilibrium under prescribed forcing, *Quart. J. R. Met. Soc.* **125**, 2767–2787.

Sobel, A. H. & Bretherton, C. S. 2000 , Modeling tropical precipitation in a single column, *J. Climate* **13**, 4378–4392.

Swann, H. 2001 , Evaluation of the mass flux approach to parametrizing deep convection., *Quart. J. R. Met. Soc.* **127**, 1239–1260.

Woolnough, S. J. 2001 , Modelling the effects of large-scale feedbacks on convection in the Met Office Cloud-Resolving Model, *Met O (APR) Turbulence and Diffusion Note No. 278.*

Zebiak, S. E. & Cane, M. 1987 , A model El Nino-Southern Oscillation, *Mon. Wea. Rev.* **115**, 2262–78.

Zhang, M. H., Lin, J. L., Cederwall, R. T., Yio, J. J. & Xie, S. C. 2001 , Objective analysis of the ARM IOP data: Method and sensitivity, *Mon. Wea. Rev.* **129**, 295–311.