Baroclinic-wave life cycles, climate simulations and cross-isentrope mass flow in a hybrid isentrope coordinate GCM

BY JOHN THUBURN

UK Universities Global Atmospheric Modelling Programme, Department of Meteorology, University of Reading, RG6 2AU

(Received 29 April 1992; revised 14 October 1992)

SUMMARY

Results are presented from a number of simulations using a global circulation model to compare the use of a $\sigma$ or hybrid $\sigma$-$\theta$ vertical coordinate with a hybrid $\sigma$-$\theta$-$p$ vertical coordinate. In idealized baroclinic instability experiments, without diabatic processes, $\sigma$ coordinates and $\sigma$-$\theta$-$p$ coordinates give very similar results, with some suggestion of slightly improved Lagrangian potential-vorticity conservation in the $\sigma$-$\theta$-$p$ coordinate case. In climate simulations it has been necessary to include extra vertical diffusion in the presence of noisy temperature profiles in order to achieve stable integrations using $\sigma$-$\theta$-$p$ coordinates (though this may be unnecessary if improved schemes for vertical advection and radiation are used). The extra diffusion leads to a noticeably less noisy temperature profile around the tropopause. Large differences between $\sigma$-$p$ coordinate integrations and $\sigma$-$\theta$-$p$ coordinate integrations appear in the moisture field because the diabatic terms and vertical-advective terms are computed in a different order, though the net moisture budget is not affected. Other diagnostics show that the two types of vertical coordinate give very similar results. Some examples are presented of a novel diagnostic: the accumulated cross-isentrope mass flow. Global tracer conservation may be poor when hybrid isentropic coordinates are used, unless care is taken to put the horizontal scale-selective dissipation into the form of a flux divergence.

1. INTRODUCTION

Several theoretical reasons have been proposed for using potential temperature as a vertical coordinate in numerical modelling of the atmosphere. The potential temperature of an air parcel can only change when there is diabatic heating; when diabatic heating is small the vertical velocity relative to isentropes is small, and air motions become quasi-two-dimensional. Numerical errors associated with vertical-advective terms might be reduced in this case by using isentropic coordinates, leading to improved model performance (Bleck 1973; Uccellini et al. 1979; Black 1987). For example, the ability of the model to conserve Rossby–Ertel potential vorticity (PV), $\zeta$, $(\nabla \theta)/\rho$, may be improved. (Here $\zeta$ is the three-dimensional vorticity vector, $\nabla \theta$ is the three-dimensional gradient of potential temperature and $\rho$ is the air density.) Finally, the relationship between diabatic processes and vertical motions is clearer in isentropic coordinates, giving a more nearly Lagrangian view of the mean circulation (Townsend and Johnson 1985).

One of the difficulties of using a purely isentropic coordinate model is that model levels intersect the ground. This may lead to problems in calculating horizontal derivatives at the ground, and in initializing new grid points as they emerge above the ground (Bleck 1974; Hsu and Arakawa 1990). The introduction of a small number of terrain-following $\sigma$-levels (Phillips 1957) close to the ground goes some way towards solving the problems (Deavin 1976; Uccellini et al. 1979; Black 1987) though some noise is still generated because of imbalance between mass and wind fields as grid points emerge above the interface between the $\sigma$-coordinate and $\theta$-coordinate domains.

Zhu et al. (1992) describe a vertical finite-difference scheme based on a hybrid $\sigma$-$\theta$-$p$ coordinate that permits a numerical model of the atmosphere to use $\sigma$-levels near the ground and isentropic levels higher up, with a smooth transition region in between. This scheme has been implemented in the UK Universities Global Atmospheric Modelling Programme (UGAMP) global circulation model (GCM), which is based on the cycle 27 forecast model of the European Centre for Medium-range Weather Forecasts (ECMWF)
(Tiedtke et al. 1988; Palmer et al. 1990). In this paper results are presented of GCM simulations using hybrid $\sigma$-$\theta$-$p$ levels and compared with simulations using the more usual $\sigma$ or hybrid $\sigma$-$p$ levels.

Some details of the implementation are reviewed in section 2. In section 3 it is shown that models using $\sigma$ or $\sigma$-$\theta$-$p$ coordinates give very similar simulations of baroclinic-instability life cycles. In some instances there is a suggestion of improved material conservation of PV when hybrid isentropic levels are used.

In section 4 two pairs of climate simulations are discussed. Many diagnostics show that simulations using $\sigma$-$\theta$-$p$ coordinates are very similar to those using $\sigma$-$p$ coordinates. The problems of model levels crossing, and of supersaturation of the moisture field in the hybrid isentropic coordinate simulations, are addressed. Some examples are shown of a novel diagnostic, the accumulated cross-isentrope flow. This is a precise measure of the three-dimensional 'diabatic' circulation, which is important for problems of atmospheric chemistry such as the lifetimes of chlorofluorocarbons (CFCs) and the rate of supply of ozone from the stratosphere to the troposphere.

This study has brought to light a problem regarding the global conservation of tracers, which the use of $\sigma$-$\theta$-$p$ coordinate makes especially conspicuous. The cause of the problem and a solution are discussed in section 5.

2. IMPLEMENTATION

The details of the hybrid isentropic vertical finite-difference scheme are given by Zhu et al. (1992), but some important points are summarized here:

(i) The scheme is an extension of the hybrid $\sigma$-$p$ coordinate scheme due to Simmons and Burridge (1981) in which the half-level pressure, $p_{k+1/2}$, may depend on the half-level temperature, $T_{k+1/2} = \frac{1}{2}(T_k + T_{k+1})$, as well as the surface pressure, $p_s$. Temperature remains a prognostic variable and half-level pressures are diagnosed from temperature and surface pressure at each time step.

(ii) The levels used by the model are specified as part of its initial data in terms of a set of constants referred to as the vertical-coordinate table, allowing flexibility in the choice of levels. For an example, see the appendix.

(iii) The vertical velocity relative to model levels, \( \tilde{\omega} = \frac{\partial \tilde{p}}{\partial \eta} \), which is used in the vertical-advection terms, can only be diagnosed from the thermodynamic and mass-continuity equations after the diabatic heating rates are known. Parametrizations of some processes in the UGAMP GCM require a preliminary estimate of the temperature and moisture tendencies due to adiabatic processes. Therefore the computations during one time step proceed in the following order:

(a) all adiabatic terms except vertical-advection terms are calculated,

(b) \( \tilde{\omega} \) is diagnosed assuming zero diabatic heating,

(c) vertical-advection terms for temperature and moisture are calculated using this estimate of \( \tilde{\omega} \),

(d) diabatic terms are calculated using the preliminary estimates of the tendencies given by (a)-(c),

(e) \( \tilde{\omega} \) is recalculated taking into account the diabatic heating, and

(f) the vertical-advection terms for temperature and moisture are recalculated and the remaining vertical-advection terms computed using the full \( \tilde{\omega} \).

(iv) The semi-implicit corrections to the divergence, temperature and surface pressure tendencies for gravity waves (Robert et al. 1972; Simmons and Burridge 1981) are unaffected by the new vertical scheme when an isothermal reference profile is used.
(v) The subroutine of the UGAMP GCM that computes the adiabatic terms in grid-point space has been completely rewritten to enable both hybrid σ-p coordinates and hybrid σ-θ-p coordinates to be used. This includes the reordering of the diabatic and adiabatic computations described in note (iii) above. Also a new subroutine was needed to diagnose half-level pressures from temperature and surface pressure. No other major changes were necessary, for example, to the physical parametrization schemes, in order to implement the new vertical coordinate.

Because the vertical velocity and some vertical advection terms are calculated twice at each time step, simulations using σ-θ-p coordinates require more computer time than those using σ-p coordinates. This increase depends on the resolution used; it is about 10% for a T21 climate simulation. A very small amount of additional memory is also required.

3. BAROCLINIC-INSTABILITY LIFE CYCLES

Baroclinic instability has been extensively studied in the past, in particular using a model that could be considered to be an ancestor of the spectral models developed at the ECMWF and hence of the UGAMP GCM (Hoskins and Simmons 1975; Simmons and Hoskins 1978, 1980), including studies at relatively high horizontal and vertical resolutions (Thornicroft and Hoskins 1990). Therefore, experiments of this kind provide an ideal test-bed for the new vertical scheme. The absence of any parametrizations of diabatic processes means that the effects on the dynamics of changing the numerical scheme can be seen most clearly.

Two simulations were compared; both used T42 horizontal resolution (that is, an expansion in spherical harmonics, truncated at total wave number 42), and 15 levels in the vertical; the control experiment (experiment 1S) used σ-levels while the second experiment (experiment 1I) used hybrid σ-θ-p levels with six pure isentropic levels between 310 K and 370 K and a transition to σ-levels at the ground and a transition to pressure levels at the top of the model. Exact details are given in Table 1 of Zhu et al. (1992). A \( \nabla^6 \) horizontal scale-selective dissipation with a time-scale of 4 hours at the shortest resolved scale was applied to vorticity, divergence and temperature, but no other diabatic processes were included.

The initial zonally averaged zonal velocity and potential temperature for the experiment are shown in Fig. 1; they are the same as those used by Simmons and Hoskins (1978). A small-amplitude normal-mode perturbation of zonal wave number 6, symmetric about the equator, is superimposed on this, and the model integrated for 15 days. A baroclinic wave grows rapidly, reaching a maximum eddy kinetic energy around day 7, before decaying barotropically, that is, through horizontal eddy fluxes of momentum into the jet. The life cycle is essentially over by day 15, leaving a much stronger, narrower jet.

In adiabatic frictionless flow, PV is conserved following fluid parcels. However, material conservation of PV will hold only approximately in a numerical model, and how well this holds is one test of a model’s performance. Isentropic maps of PV were examined at 350 K, 330 K and 310 K on days 7, 8, 9 and 10 of the two integrations. Generally the two integrations give very similar results. Inspection of the maximum PV values on successive days suggests that, at the 350 K level, PV conservation is rather better in experiment 1I. Figure 2 shows 350 K PV from the two experiments at day 9. In experiment 1S the maximum at around 65°N is more than 0.5 PVU (1 PVU = 10^{-6} m^2 s^{-1} K kg^{-1}) greater than it is on either day 8 or day 10, while in experiment 1I the largest day-to-day
Figure 1. Zonal mean initial zonal velocity (contour interval 5 m s⁻¹) and potential temperature minus 300 K (contour interval 5 K) for the baroclinic instability life-cycle experiments 1S and 1I.

Figure 2. Potential vorticity on the 350 K isentrope at day 9 of a baroclinic-instability life cycle; (a) experiment 1S and (b) experiment 1I, contour interval 10⁻⁶ m² s⁻¹ K kg⁻¹. Only the region from the equator to the North Pole and from 0°E to 180°E is shown.
change is about 0.22 PVU. At the 330 K level both experiments show day-to-day changes of more than 0.5 PVU with the $\sigma-\theta-p$ case being slightly worse.

These errors on both the 350 K and 330 K levels are associated with the spurious increase of a PV maximum. In adiabatic flow, vertical velocities, and hence errors in vertical advection, should be reduced when isentropic coordinates are used, and this is consistent with the improved PV conservation at the 350 K level in experiment II compared with experiment IS. However, the errors are not necessarily reduced when experiment IS is repeated with the centred-difference vertical-advection scheme replaced by monotonicity-preserving schemes of varying complexity (Thuburn 1993). An alternative explanation is that gradients and changes of gradient of temperature and vorticity along model levels are likely to be sharper in experiment IS than in experiment II, and so the spectral horizontal representation may be less accurate and, in particular, more prone to Gibbs phenomena.

On the 310 K surface (not shown) the maximum PV values decrease monotonically in time between days 7 and 10 in both experiments, with slightly larger changes in the $\sigma-\theta-p$ coordinate case. The relatively small horizontal scale of the PV maxima at this level suggests that the decrease of the PV maxima may be associated with the horizontal scale-selective dissipation rather than with other model errors.

The horizontal eddy-momentum flux, $u'u'$, and heat flux, $\overline{v'T'}$, for the two experiments, time averaged from days 1 to 15, are shown in Figs. 3 and 4. The differences are very small. Above 200 hPa the momentum flux is reduced slightly and the poleward heat flux is increased slightly in experiment II, while the poleward heat flux is reduced slightly at low levels. The peak values of both fluxes are reduced by about 7% in the $\sigma-\theta-p$ coordinate case. The zonally averaged zonal wind and temperature at day 15 are almost identical in the two experiments.

Differences in these diagnostics may be due to the following reasons:

(i) the fact that the model levels are in different places is bound to lead to numerical differences in the interpolation of the initial data onto the model levels, in the integration itself and in the interpolation of diagnostics onto pressure levels or isentropic levels,

(ii) errors associated with the vertical-advection scheme may be reduced in near-adiabatic conditions when isentropic coordinates are used,

(iii) gradients, and changes of gradient, of prognostic variables along model levels may be reduced in isentropic coordinates so that the horizontal spectral representation is more accurate, and

(iv) the horizontal scale-selective dissipation acts along more steeply sloping model levels when hybrid isentropic coordinates are used. Systematic differences in the temperature gradient along model levels, for example, might lead to systematic differences in heat transport by the scale-selective dissipation. Also, in hybrid isentropic coordinates, the horizontal gradient of level thickness, $\nabla \Delta p$, may be significant so that the failure of the model's scale-selective dissipation to conserve angular momentum, potential energy or mass of tracer may become correspondingly significant (see section 5).

Two analogous experiments were carried out, from the same initial state, with horizontal resolution increased to T85 and using 2-hour $\nabla^6$ horizontal scale-selective dissipation. The conclusions are very similar to those of the T42 experiments: Lagrangian conservation of PV appears to be improved slightly on the 330 K and 350 K isentropes, and the peak eddy heat and momentum fluxes are reduced slightly in the hybrid isentropic coordinate case. The final zonal mean temperature and zonal wind fields are almost identical in the two experiments.

Clearly the new vertical-coordinate scheme performs as well in these experiments as the $\sigma$-coordinate scheme, though not necessarily any better. To put things in perspec-
Figure 3. (a) Zonal mean momentum flux, $\bar{u}'\bar{v}'$, time averaged over days 1 to 15, for experiment 1S, contour interval 20 m s$^{-1}$, and (b) zonal mean heat flux, $\bar{v}'\bar{T}'$, time averaged over days 1 to 15, for experiment 1S, contour interval 5 K m s$^{-1}$.

The differences between experiments 1S and II are far smaller than the differences obtained by increasing the horizontal resolution to T85.

4. CLIMATE SIMULATIONS

Two pairs of climate simulations were performed to compare the use of hybrid $\sigma-p$ vertical coordinates with hybrid $\sigma-\theta-p$ vertical coordinates. The first pair of experiments were at T21 horizontal resolution with 24 vertical levels, and included parametrizations of radiation, convection (a Kuo-type scheme), condensation, vertical mixing, orographic gravity-wave drag and surface processes, as in the ECMWF cycle 27 forecast model. The model was integrated for 360 days from initial data for 15 January 1987, with seasonally
Figure 4. As in Fig. 3 but for experiment II.

varying solar forcing, and sea-surface-temperature and deep-soil boundary conditions. The hybrid $\sigma$-$\theta$-$p$ coordinate experiment (experiment 21) used vertical levels defined in Table 2 of Zhu et al. (1992). These include nine pure isentropic half-levels, the lowest being at approximately 344 K. The height of the lowest pure isentropic level that may be used is limited by the height (and temperature) of the model orography. The hybrid $\sigma$-$p$ coordinate experiment (experiment 2S) used the corresponding $\sigma$-$p$ levels defined by the a's and b's in the same table, with c's identically 0.0 and d's identically 1.0.

The second pair of experiments were at T42 horizontal resolution and used a new radiation scheme due to Morcrette (1990) and an experimental convection scheme due to Betts and Miller (Betts 1986; Betts and Miller 1986 and personal communication). The model was run for 120 days with ‘perpetual January’ solar forcing, and sea-surface
and deep-soil boundary conditions. The $\sigma-\theta-p$ coordinate experiment (experiment 3I) used vertical levels defined in Table 1 of the appendix. There are five pure isentropic half-levels from approximately 380 K to 500 K. The corresponding $\sigma-p$ coordinate experiment (experiment 3S) used levels defined by the $a$'s and $b$'s in that table.

(a) The problem of levels crossing

Zhu et al. (1992) describe in detail the problem of adjacent model levels crossing each other, i.e. of layer thickness becoming negative. There are two types of situation where this can occur. The first is when the vertical-coordinate table is inappropriate so that the transition from one type of level to another is made over too small a height range. For example, the envelope orography in the UGAMP GCM is significantly higher at T42 resolution than at T21. When the vertical-coordinate table of experiment 2I was used in a T42 experiment, the height and warmth of the Himalayan plateau during the northern summer caused model levels there to cross in the $\sigma$ to $\theta$ transition region. In another experiment, with 'perpetual July' solar forcing and lower boundary conditions, the southern hemisphere stratosphere became extremely cold so that the hybrid $\sigma-p$ level, level $2\frac{1}{2}$, rose until it crossed the fixed pressure level, level $1\frac{3}{4}$. For these reasons the vertical-coordinate table of experiment 3I uses more gradual transitions from $\sigma$ to $\theta$ levels and from $\theta$ to $\sigma$ levels than that of experiment 2I.

The second situation in which model levels can cross is when strong, grid-scale features arise in the vertical profile of temperature. For example, levels crossed during one integration in which an experimental version of a convection scheme gave unreasonably large cooling rates at one level. In another case a strong temperature inversion at around 200 hPa over Antarctica caused levels to cross after about 70 days of integration. In fact grid-scale oscillations in the vertical profile of temperature are seen above the tropopause at high latitudes in both hemispheres in climate simulations using the UGAMP GCM. These are now believed to be caused by using a centred-difference scheme for the vertical-advection terms, since similar features are seen in passive tracer fields. Certainly the grid-scale noise no longer arises in tracer fields when a monotonicity-preserving advection scheme is used for vertical advection of tracers, and in hybrid $\sigma-p$ coordinate integrations use of the monotonicity-preserving scheme for vertical advection of temperature eliminates the noise from the temperature profile (Thuburn 1993).

Unfortunately the current implementation of the hybrid $\sigma-\theta-p$ coordinate scheme may only use the centred-difference scheme for vertical advection of temperature. It is possible to extend equation (2.24) of Zhu et al. (1992) to allow other vertical-advection schemes to be used on temperature (provided any new scheme is linear in $\tilde{\omega}$, which not all 'flux limited' schemes are), but this has not yet been implemented in the hybrid isentropic coordinate version of the UGAMP GCM.

In the real atmosphere one would expect radiation to damp these grid-scale temperature oscillations in the upper troposphere and lower stratosphere over a time-scale of a few days. However, the particular numerical implementations of both the original ECMWF cycle 27 radiation scheme and of the Morcrette scheme used in the UGAMP GCM are unable to damp two-grid oscillations. A modified implementation of the Morcrette scheme which is able to damp two-grid oscillations in temperature is currently being tested (D. Li, personal communication).

Other workers developing isentropic coordinate models have found similar problems, though not necessarily for the same reasons. For example, when density is used as a prognostic variable the use of an advection scheme that does not preserve positive values can lead to negative densities (Bleck and Boudra 1986; Hsu and Arakawa 1990; M. Fisher, personal communication).
HYBRID ISENTROPIC COORDINATE SIMULATIONS

For the experiments described here, an ad hoc solution to the problem has been adopted in which vertical diffusion is enhanced locally when the vertical spacing of model levels becomes very irregular, as it does when the vertical profile of temperature becomes noisy. (Details are given by Zhu et al. (1992).) This is sufficient to enable stable integrations of the UGAMP GCM to be carried out using hybrid isentropic vertical levels. It is hoped that the improvements in the vertical-advection scheme for temperature, and in the radiation scheme mentioned here, will make the enhanced vertical diffusion redundant.

(b) Results of climate simulations

The zonally averaged temperature and zonal wind give simple measures of the performance of a climate model. In experiments 2I and 3I the temperature profiles above the tropopause at high latitudes are considerably less noisy than in experiments 2S and 3S. This is because the enhanced vertical diffusion in the hybrid $\sigma-\theta-p$ coordinate experiments removes features that are probably generated by the numerical scheme for vertical advection. (Radiation acting on grid-scale features in the vertical profile of moisture, which themselves are caused by the vertical-advection scheme, may also contribute to the noisy temperature profile.) Figure 5 shows the difference in the zonally averaged temperature between experiments 2I and 2S for the month of June.

The upper troposphere, roughly 100 hPa to 300 hPa, tends to be warmer, typically by 2–4 K though sometimes more, in the hybrid isentropic coordinate integrations. This temperature difference is greatest at high latitudes, especially in experiment 2I, but is also present in the tropics. The increased temperature at high latitudes may be partly due to stronger meridional temperature gradients along the sloping model levels, leading to increased heat transport by the horizontal scale-selective dissipation (3-hour $\nabla^8$ in experiments 2S and 2I). In the UGAMP GCM the horizontal scale-selective dissipation is applied by modifying the spectral coefficients, so it is not possible to compute the corresponding fluxes directly to test this conjecture. An estimate of the temperature tendency due to scale-selective dissipation, calculated as a residual of known contributions, is inconclusive.

![Figure 5. Zonal mean temperature averaged over June for experiment 2I minus zonal mean temperature averaged over June for experiment 2S, contour interval 2 K.](image-url)
The largest differences in the zonally averaged zonal wind, $\bar{u}$, between experiments 2S and 2I occur in the winter stratosphere, where the model variability is greatest, and in the tropical upper troposphere. (Incidentally, this version of the UGAMP GCM has an easterly bias in the upper tropical tropospheric winds, which is sensitive to the parametrizations of deep convection and surface latent-heat flux.) In experiments 3S and 3I the zonal mean zonal winds are very similar, again with the largest differences in the winter stratosphere and upper tropical troposphere. Figure 6 shows the zonal mean zonal wind for the two experiments averaged over the last 60 days of the integrations. In experiment 3I the relatively warm upper tropical troposphere is accompanied by slightly reduced upper tropical easterlies, consistent with thermal-wind balance.

In experiments 2S and 2I the horizontal eddy-heat and momentum fluxes show considerable variability from month to month. The momentum fluxes into the tropospheric jets in experiment 2I seem to be generally weaker, by up to 20%, than in

Figure 6. Zonal mean zonal wind averaged over the last 60 days of (a) experiment 3S and (b) experiment 3I, contour interval 5 m s$^{-1}$. 
experiment 2S. This is consistent, at least in sign, with the results of the baroclinic
instability experiments described in section 3. However, any systematic differences in
the horizontal eddy-heat fluxes are less obvious. Experiments 3S and 3I give nearly
identical results for a wide range of zonal mean second-order diagnostics such as
horizontal and vertical fluxes of heat and momentum, and eddy kinetic energy.

(c) Moisture

The largest differences between $\sigma-p$ coordinate simulations and $\sigma-\theta-p$ coordinate
simulations are seen in the moisture field, especially when this is displayed as relative
humidity. Figure 7 shows zonal mean relative humidity from experiments 2S and 2I for
February. Two major differences are apparent. Experiment 2I has extremely large
relative humidity, in fact supersaturation, in the upper tropical troposphere. Experiment
2S has a very sharp tropical/subtropical temperature inversion at a constant height, about

![Figure 7. Zonal mean relative humidity averaged over February for (a) experiment 2S and (b) experiment 2I, contour interval 10%.](image)
700 hPa, coincident with a sharp decrease in the moisture mixing ratio with height. This feature is unrealistically sharp. In experiment 2I the height of the largest vertical moisture gradient appears to be lowered to the north and to the south of the equator.

Examination of the moisture values and moisture tendencies at various stages of the calculations for one time step reveals that the supersaturation arises because of the reordering of the calculations described in section 2, note (iii). In the usual $\sigma-p$ coordinate case, the adiabatic processes, such as vertical advection of moisture and cooling due to the energy-conversion term, $k\omega T/p$, can lead to supersaturation. The diabatic terms are subsequently calculated and any supersaturation is removed by condensation and precipitation. Model fields written out at the end of a time step would show no supersaturation. In the hybrid $\sigma-\theta-p$ coordinate scheme the vertical-advection terms are necessarily calculated after the diabatic terms. Vertical advection of moisture may lead to supersaturation that is not removed before the model fields are written out. The supersaturation is removed during the next time step, so that the net moisture budget is not affected, to a first approximation. Average cloudiness and precipitation rates are very similar in experiments 2S and 2I. Iterating steps (d)-(f) of note (iii), section 2, should reduce the apparent supersaturation in hybrid isentropic coordinate simulations, though this would be computationally very expensive.

The very large supersaturations in the hybrid isentropic coordinate integration occur in a region where there is strong ascent and where the moisture saturation mixing ratio falls very rapidly with height, typically by a factor of 10 per model level, so that large supersaturations can occur immediately after the vertical-advection term is computed. Furthermore, the actual amounts of moisture and latent energy involved are very small indeed. Therefore the supersaturation might be regarded as a problem in interpreting the model output, because of the ordering of the computations during each time step, rather than a serious flaw in the hybrid $\sigma-\theta-p$ coordinate scheme. However, it is noteworthy that very high vertical gradients of moisture mixing ratio occur in the middle and upper tropical troposphere, a region where hybrid isentropic coordinates tend to give relatively low vertical resolution because of the small (dry) static stability.

The unrealistic sharpness of the boundary-layer inversion in experiment 2S is due partly to having an artificial top at (the first level above) 1600 m in the parametrization of shallow convection. Shallow convection, when active, is represented as a strong vertical diffusion. The UGAMP GCM also includes a vertical-diffusion scheme to represent vertical mixing in the boundary layer and, with generally weaker diffusion coefficients, in the free atmosphere. Comparison of the moisture tendencies due to different processes shows that the flux of moisture due to vertical diffusion (not shallow convection) across the inversion is much greater in experiment 2I than in experiment 2S. The extra moistening above the inversion is compensated for, mainly by extra drying there in the deep convection parametrization. Statistics that monitor the enhancement of vertical diffusion discussed in section 4(b) show that this operates mainly at high latitudes and cannot account for these changes in the moisture fluxes at the inversion. The parametrization of vertical diffusion uses diffusion coefficients that depend on a Richardson number calculated from differences in geopotential, dry static energy and wind between adjacent model levels. Differences in the exact locations of model levels between experiments 2S and 2I, especially in relation to the artificial top of the shallow convection scheme, may lead to differences in the computed Richardson number, and hence in the diffusion coefficients.

Experiment 3I also shows supersaturation in the upper tropical troposphere, for the same reason as it appears in experiment 2I. However, experiments 3S and 3I used a different scheme for deep and shallow convection, and, in particular, have much less
sharp boundary-layer inversions in the tropics and subtropics. Possibly as a result of this, there are no great differences in the moisture in this region between experiments 3I and 3S.

In summary, the differences between these $\sigma-\theta-p$ coordinate climate simulations and $\sigma-p$ coordinate simulations may be due to the reasons listed in section 3 and also due to:
(i) enhanced vertical diffusion in the hybrid isentropic coordinate simulations when model levels are irregularly spaced,
(ii) reordering of the adiabatic computations, and
(iii) sensitivity of some physical parametrizations to the exact locations of model levels.

\(d\) Cross-isentrope mass flow

The rate of flow of air across an isentrope is proportional to the diabatic heating rate. Thus it provides a clear relationship between ascent and diabatic heating, and between descent and diabatic cooling. The accumulated mass flux across an isentrope may be a useful diagnostic in studies of the atmospheric transport of chemicals, especially in connection with the exchange of chemical species between the troposphere and the photolytic regions of the stratosphere. In the middle atmosphere the cross-isentrope flow may be related to eddy-induced zonal forces at higher levels using the 'downward control' ideas of Haynes et al. (1991, see also McIntyre (1992)). Pawson and Harwood (1989) have used satellite observations plus a radiative-transfer model to calculate heating rates on isentropic levels in the stratosphere, as well as the zonal mean cross-isentropic transport of ozone by the zonal mean flow and by eddies. Use of the cross-isentrope mass flow as a diagnostic may also lend a more nearly Lagrangian perspective on the atmospheric circulation.

In order to calculate the cross-isentrope mass flow in a $\sigma-p$ coordinate model, information on the three-dimensional diabatic heating field is required. (This is unlikely to be written out at every time step so the accumulated cross-isentrope flow would be especially difficult to calculate.) The quantity $\bar{\omega}$, which is $g$ times the downward mass flux across a model level, is calculated at every step in the UGAMP GCM. When hybrid isentropic levels are used, $\bar{\omega} = \partial \bar{\omega}/\partial \theta$ in the region of pure isentropic levels, it is a trivial matter to accumulate this quantity and write out the result at regular intervals.

In experiment 2I the accumulated mass flow across the lower isentropic levels in the model (342 K to 401 K) shows the largest values of both diabatic ascent and diabatic descent to be in the tropics, with the regions of ascent well correlated with the main regions of tropical convection: South America, Africa and Indonesia. There is a broad, almost featureless, area of descent covering the high latitudes of both hemispheres. At higher levels (428 K to 622 K) the features in the tropics become progressively weaker and uncorrelated with the convection below. For example, Fig. 8 shows the flow across the 358 K and 622 K levels accumulated over the month of February. At 358 K there is significant descent over subtropical desert regions such as the Sahara and Gobi, presumably associated with radiative cooling. At 622 K the largest cross-isentrope flow is descent over the North Pole, while the smaller-scale features in the tropics are reminiscent of a wave train. Interpolating between the 498 K level (not shown) and the 622 K level gives an estimate of 0.8–1.0 hPa d$^{-1}$ (approximately 0.9–1.2 mm s$^{-1}$) for the largest diabatic descent rate at high latitudes at an altitude of 18 km. This is similar to, though a fraction larger than, the estimate (approximately 0.8 mm s$^{-1}$) shown in Fig. 12 of Haynes et al. (1991).

Figure 9 (taken from Hoskins et al. 1989) shows the average diabatic heating for December, January and February, integrated between 700 hPa and 50 hPa, calculated,
as a residual in the thermodynamic equation, from 6 years of ECMWF analyses. Encouragingly, the major features in the tropical heating are reflected in the 358 K cross-isentrope flow in simulation 2I, though the latent heating associated with the northern hemisphere storm tracks occurs at too low levels to be seen in Fig. 8.

During the northern winter there is descent at high latitudes across all the model's isentropic levels in both hemispheres. However, in June and July there is ascent across the 622 K level over a large part of the northern hemisphere, with descent mainly in the southern hemisphere, as we begin to see the transition in the stratosphere from a meridional circulation cell in each hemisphere (Brewer 1949; Dobson 1956) to a single global circulation cell (Murgatroyd and Singleton 1961). It is interesting, and noteworthy, that the model's extratropical three-dimensional diabatic circulation should be so close to the standard zonally averaged residual or diabatic circulations used in two-dimensional chemical modelling. The tropics again are less simple; in June and July the flow across the 377 K to 401 K levels is dominated by the strong convective heating associated with the Indian monsoon.
Figure 9. Diabatic heating integrated from 700 hPa to 50 hPa averaged over December, January and February, calculated as a residual in the thermodynamic equation from 6 years of ECMWF analyses. Contour interval is 50 W m$^{-2}$. Taken from Hoskins et al. (1989).

Comparison of experiment 2I with experiment 3I suggests the pattern of cross-isentrope flow at certain levels of a GCM may be sensitive to model details such as the resolution and parametrization scheme for deep convection.

The cross-isentrope mass flow has obvious applications in the study of the transport of chemical species in the stratosphere. For example, maps of moisture mixing ratio and ozone mixing ratio at about 60 hPa obtained from the Limb Infrared Monitor of the Stratosphere (LIMS) instrument on board the NIMBUS 7 satellite (J. A. Pyle, personal communication) show interesting spatial variations in the tropics which correspond approximately to the patterns seen in maps of cross-isentrope mass flow. Although the diagnostic presented here shows the accumulated flow of air across an isentrope, by multiplying by the local mass mixing ratio the flow of moisture or any chemical tracer across an isentrope could also be calculated. But note that this diagnostic would capture the transport by resolved motions only. If the model includes parametrizations of vertical transport of tracers by sub-grid-scale motions, such as deep convection or turbulence due to breaking gravity waves, that transport would not be captured by this diagnostic.

5. TRACER CONSERVATION

Experiments 2S and 2I each included ten passive tracers having idealized initial distributions. The tracer mixing ratio, $\chi$, is predicted using the equation,

$$\frac{\partial \chi}{\partial t} + \mathbf{v} \cdot \nabla \chi + \frac{\partial \chi}{\partial p} = \kappa \nabla^n \chi$$

(5.1)

where $\nabla$ is the gradient operator along model levels and $n$ is a positive integer. The term on the right-hand side is a horizontal scale-selective dissipation (which is computed in spectral space in the UGAMP GCM). There is no parametrization of tracer transport by sub-grid-scale processes such as convection.

In experiment 2I the global conservation of the tracers was very poor, with the total mass of one tracer changing by as much as 40% over the year-long integration. It was
found, by removing the scale-selective dissipation term in various equations, that scale-selective dissipation acting on tracer mixing ratio and on temperature were the dominant sources of error, with comparable contributions from each. Because the model-level thickness, $\Delta p$, varies greatly in the horizontal (for example by a factor of 4 between mid latitudes and the tropics in the zonal mean at some levels) the scale-selective dissipation acting on mixing ratio changes the tracer distribution in a way that does not conserve the total mass of tracer. Also, scale-selective dissipation acting on temperature modifies the level thicknesses (since model-level pressure depends on temperature) in a way that does not conserve mass of tracer.

Horizontal gradients of model-level thickness can also occur when $\sigma-p$ coordinates are used, for example near steep orography. The impact on tracer conservation is much smaller than in the hybrid isentropic coordinate case; in experiment 2S the worst violation of conservation seen was a 4% reduction in total mass of tracer in 60 days. However, this is still large enough to be of concern to chemical modellers.

One practical solution is to use an alternative tracer variable, $X = \chi \partial p/\partial \eta$, which is the mixing ratio weighted by the density of air in model coordinates. Equation (5.1) is replaced by

$$\frac{\partial X}{\partial t} + \nabla \cdot (\mathbf{v}X) + \frac{\partial}{\partial \eta} (\eta X) = \kappa \nabla^{2n} X$$

(5.2)

where now the scale-selective dissipation is applied to $X$ rather than $\chi$. The tendency due to scale-selective dissipation is in the form of a divergence of a flux and so conserves mass of tracer. Furthermore, scale-selective dissipation acting on temperature no longer affects the total mass of tracer. In fact the adoption of this method also removes several smaller numerical contributions to non-conservation, for example, contributions associated with semi-implicit modifications to the time-stepping scheme.

Experiment 2I was repeated using (5.2) instead of (5.1) to model tracer transport. After some very small initial oscillations died down (associated with initializing the tracer distributions in a model with a leap-frog time step and time filtering), the tracers were conserved very accurately, with total masses unchanged to seven significant figures over the year-long integration.

A disadvantage of this approach is that an initial tracer distribution defined by a constant mixing ratio will not remain constant, as it should. There will be a small amount of spurious transport associated with the scale-selective dissipation and with numerical errors in the advection terms. In practice the spurious transport is small compared with the 'true' transport and is less of a problem than the poor conservation of mass. A scale-selective dissipation of the form,

$$\frac{\partial X}{\partial t} + \nabla \cdot (\mathbf{v}X) + \frac{\partial}{\partial \eta} (\eta X) = \kappa \nabla^{2n-k} \left( \frac{\partial p}{\partial n} \nabla^k \chi \right)$$

(5.3)

with integers $n$ and $k$, $2n > k > 0$, would both conserve mass and give no spurious transport when $\nabla \chi = 0$. However, this is not so easy to implement in a spectral model.

6. DISCUSSION

We have shown that, in idealized baroclinic instability life-cycle experiments and in climate simulations, a GCM using a hybrid $\sigma-\theta-p$ vertical coordinate gives broadly similar results to one using the more usual hybrid $\sigma-p$ coordinate. The anticipated benefits of using isentropic coordinates, discussed in section 1, appear to have been realized only
in a small way. In climate simulations the height of the envelope orography limits the region where isentropic coordinates may be used to the upper troposphere and above. The benefits of using this hybrid isentropic vertical scheme may be more fully realized in simulations of the middle atmosphere, where it should be possible to use isentropic coordinates over a large fraction of the model domain.

In baroclinic wave life-cycle experiments there appears to be a small improvement in the simulation of small-scale PV features at some levels when $\sigma-\theta-p$ coordinates are used. If this is due to the mechanism conjectured in note (iii) of section 3, the interaction of the horizontal spectral representation and the horizontal scale-selective dissipation is probably important. In this respect, the Anticipated Potential Vorticity method of Sadourny and Basdevant (1985) appears to offer a promising alternative to the more usual $\nabla^2$ type of dissipation, especially in isentropic coordinates, and there are plans to test this in the UGAMP GCM.

One anticipated benefit of using isentropic model levels was that numerical errors associated with vertical-advection terms might be reduced where vertical motions are predominantly adiabatic. In fact, in climate simulations using the UGAMP GCM, the largest systematic errors due to vertical advection occur in the temperature field around the high-latitude tropopause and lower stratosphere and in the moisture field above the subtropical inversion (depending on the shallow convection scheme used) where negative moisture values can arise. Both of these errors are associated with diabatic cooling and descent, so there is no benefit in changing to isentropic model levels. In this case the best remedy is to use an alternative scheme for vertical advection.

The hybrid isentropic vertical scheme is sensitive to the presence of small vertical-scale features in the temperature profile, which may be caused by flaws in the vertical-advection scheme for temperature or in parametrization schemes. Stable integrations can be achieved by enhancing the vertical diffusion when and where model-level spacing becomes irregular. The effect of the enhanced vertical diffusion on the zonal mean temperature above the tropopause is clearly visible in climate simulations. A preferred solution would be to use improved numerical schemes for vertical advection of temperature and for radiation, which are being developed, which prevent the occurrence of noisy temperature profiles.

When hybrid isentropic coordinates are used the vertical-advection terms must be calculated after the diabatic terms. In regions where the moisture amount is determined by a strong balance between vertical advection and other, parametrized, physical processes (where there is a sharp vertical gradient of moisture and a large vertical velocity) this reordering of the computations relative to the time at which model fields are written out can lead to apparent differences in the moisture field. For example, supersaturation appears in the upper tropical troposphere. The net moisture budget is unaffected, to a first approximation.

In this study the major advantage of using a hybrid isentropic coordinate GCM appears to be the availability of a new diagnostic of potential importance for stratospheric chemistry studies, the accumulated cross-isentrope mass flux. This is a precise measure of the three-dimensional diabatic circulation. Some examples of this diagnostic have been presented, and have suggested that in the extratropical lower stratosphere the diabatic circulation is remarkably like the standard two-dimensional zonally averaged picture, but is more complicated in the tropics.

When hybrid isentropic coordinates are used, horizontal scale-selective dissipation acting on tracer mixing ratio and on temperature can lead to very poor global tracer conservation. By allowing scale-selective dissipation to act on density-weighted tracer values very good global conservation can be achieved.
In the near future it is planned to test the hybrid isentropic coordinate version of the UGAMP GCM with an improved implementation of its radiation scheme and with an improved vertical-advection scheme for temperature. It is hoped that the model's temperature fields will then be sufficiently free from noise to allow stable integrations without the need for ad hoc enhancement of vertical diffusion.

Acknowledgements

I am grateful to Drs M. J. Miller and J-J. Morcrette of the ECMWF for supplying parametrization schemes to the UGAMP in the form of computer code. Discussions with Dr M. Blackburn, Prof. B. J. Hoskins, Dr M. E. McIntyre and Dr K. P. Shine were helpful during the course of this work. The UK Universities Global Atmospheric Modelling Programme is funded by the UK Natural Environment Research Council.

Appendix

In the hybrid $\sigma-\theta-p$ vertical scheme the pressure on each model half-level is defined by four constants, $a$, $b$, $c$ and $d$ (the vertical-coordinate table), the local temperature and the surface pressure:

$$d_{k+1/2} \cdot p_{k+1/2} = a_{k+1/2} + b_{k+1/2} \cdot p_\alpha - c_{k+1/2} \cdot p_{k+1/2} \cdot T_{k+1/2}^{1/2}.$$  \hfill (A.1)

When $d = 1$ and $c = 0$ this reduces to the hybrid $\sigma-p$ coordinate of Simmons and Burridge (1981). Table A.1 gives the vertical-coordinate table used for experiment 31. The half-levels correspond to approximate $\sigma$ and $\theta$ values given in columns 2 and 3. Levels $5 \frac{1}{2}$ to $9 \frac{1}{2}$ are pure isentropic levels.

<table>
<thead>
<tr>
<th>Level</th>
<th>$\sigma$</th>
<th>$\theta$</th>
<th>$a \times 10^{-5}$</th>
<th>$b$</th>
<th>$c \times 10^{-6}$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{2}$</td>
<td>0.0</td>
<td>—</td>
<td>0.0</td>
<td>0.0</td>
<td>0.183</td>
<td>1.00</td>
</tr>
<tr>
<td>$\frac{1}{4}$</td>
<td>0.0093</td>
<td>850</td>
<td>0.0093</td>
<td>0.0</td>
<td>0.0183</td>
<td>0.890</td>
</tr>
<tr>
<td>$\frac{1}{8}$</td>
<td>0.0186</td>
<td>705</td>
<td>0.0186</td>
<td>0.0</td>
<td>0.0571</td>
<td>0.670</td>
</tr>
<tr>
<td>$\frac{1}{16}$</td>
<td>0.0286</td>
<td>625</td>
<td>0.0286</td>
<td>0.0</td>
<td>0.0978</td>
<td>0.440</td>
</tr>
<tr>
<td>$\frac{1}{32}$</td>
<td>0.0406</td>
<td>550</td>
<td>0.0406</td>
<td>0.0</td>
<td>0.1236</td>
<td>0.220</td>
</tr>
<tr>
<td>$\frac{1}{64}$</td>
<td>0.0546</td>
<td>500</td>
<td>0.0546</td>
<td>0.0</td>
<td>0.1526</td>
<td>0.000</td>
</tr>
<tr>
<td>$\frac{1}{128}$</td>
<td>0.0706</td>
<td>460</td>
<td>0.0706</td>
<td>0.0</td>
<td>0.1747</td>
<td>0.000</td>
</tr>
<tr>
<td>$\frac{1}{256}$</td>
<td>0.0886</td>
<td>430</td>
<td>0.0886</td>
<td>0.0</td>
<td>0.1940</td>
<td>0.000</td>
</tr>
<tr>
<td>$\frac{1}{512}$</td>
<td>0.1096</td>
<td>403</td>
<td>0.1096</td>
<td>0.0</td>
<td>0.1940</td>
<td>0.000</td>
</tr>
<tr>
<td>$\frac{1}{1024}$</td>
<td>0.1346</td>
<td>379</td>
<td>0.1346</td>
<td>0.0</td>
<td>0.1427</td>
<td>0.000</td>
</tr>
<tr>
<td>$\frac{1}{2048}$</td>
<td>0.1636</td>
<td>360</td>
<td>0.1541</td>
<td>0.0095</td>
<td>0.1264</td>
<td>0.058</td>
</tr>
<tr>
<td>$\frac{1}{4096}$</td>
<td>0.1986</td>
<td>344</td>
<td>0.1711</td>
<td>0.0275</td>
<td>0.1292</td>
<td>0.138</td>
</tr>
<tr>
<td>$\frac{1}{8192}$</td>
<td>0.2456</td>
<td>329</td>
<td>0.1981</td>
<td>0.0565</td>
<td>0.1222</td>
<td>0.230</td>
</tr>
<tr>
<td>$\frac{1}{16384}$</td>
<td>0.3026</td>
<td>315</td>
<td>0.2045</td>
<td>0.0981</td>
<td>0.1134</td>
<td>0.324</td>
</tr>
<tr>
<td>$\frac{1}{32768}$</td>
<td>0.3716</td>
<td>303</td>
<td>0.2147</td>
<td>0.1669</td>
<td>0.1040</td>
<td>0.422</td>
</tr>
<tr>
<td>$\frac{1}{65536}$</td>
<td>0.4516</td>
<td>295</td>
<td>0.2166</td>
<td>0.2350</td>
<td>0.0955</td>
<td>0.520</td>
</tr>
<tr>
<td>$\frac{1}{131072}$</td>
<td>0.5416</td>
<td>289</td>
<td>0.2073</td>
<td>0.3343</td>
<td>0.0851</td>
<td>0.617</td>
</tr>
<tr>
<td>$\frac{1}{262144}$</td>
<td>0.6359</td>
<td>286</td>
<td>0.1857</td>
<td>0.4502</td>
<td>0.0735</td>
<td>0.708</td>
</tr>
<tr>
<td>$\frac{1}{524288}$</td>
<td>0.7274</td>
<td>285</td>
<td>0.1538</td>
<td>0.5736</td>
<td>0.0601</td>
<td>0.789</td>
</tr>
<tr>
<td>$\frac{1}{1048576}$</td>
<td>0.8109</td>
<td>285</td>
<td>0.1158</td>
<td>0.6951</td>
<td>0.0452</td>
<td>0.857</td>
</tr>
<tr>
<td>$\frac{1}{2097152}$</td>
<td>0.8819</td>
<td>285</td>
<td>0.0769</td>
<td>0.8050</td>
<td>0.0301</td>
<td>0.913</td>
</tr>
<tr>
<td>$\frac{1}{4194304}$</td>
<td>0.9366</td>
<td>285</td>
<td>0.0430</td>
<td>0.8936</td>
<td>0.0168</td>
<td>0.954</td>
</tr>
<tr>
<td>$\frac{1}{8388608}$</td>
<td>0.9730</td>
<td>285</td>
<td>0.0185</td>
<td>0.9545</td>
<td>0.0072</td>
<td>0.981</td>
</tr>
<tr>
<td>$\frac{1}{16777216}$</td>
<td>0.9923</td>
<td>285</td>
<td>0.0048</td>
<td>0.9875</td>
<td>0.0019</td>
<td>0.995</td>
</tr>
<tr>
<td>$\frac{1}{33554432}$</td>
<td>1.0000</td>
<td>—</td>
<td>0.0</td>
<td>1.0000</td>
<td>0.0</td>
<td>1.000</td>
</tr>
</tbody>
</table>
REFERENCES

Phillips, N. A. 1957 A coordinate system having some special advantages for numerical forecasting. J. Meteorol., 14, 184-185


