Pergamon

Journal of Atmospheric and Terrestrial Physics, Vol. 56, No. 10, pp. 1399–1419, 1994

0021–9169/94 $7.00+0.00

A spectral approach for studying middle and upper atmospheric phenomena

KWING L. CHAN,* HANS G. MAYR,† J. G. MENGEL* and I. HARRIS†

* Applied Research Corporation, Landover, Maryland, U.S.A.;
† Goddard Space Flight Center, Greenbelt, Maryland, U.S.A.

(*Received in final form* 15 *June* 1993; *accepted* 1 *July* 1993)

Abstract—We report on the application of a newly developed spectral code to the study of the middle/upper atmosphere. The spectral approach offers conceptual and practical convenience for analyzing the generation and interaction of different components of atmospheric activity through the decomposition of the dynamical fields into components with different zonal wave numbers (m). As examples and tests, we obtain solutions for the $m = 0$ (the mean circulation) and 1 (the diurnal tides) components separately (no mutual interactions), as well as the $m = 1$ component under the influence of the mean circulation. By simulating gravity wave effects with Rayleigh friction and eddy diffusion peaking near 90 km altitude, the mean circulation thus generated can reproduce the observed mesospheric temperature anomaly under solstice conditions. The computed diurnal tides are in good agreement with results obtained earlier by other authors. The large temperature gradient (associated with the $m = 0$ component) set up by the mesospheric temperature anomaly under solstice conditions creates a condition favorable for the development of baroclinic instabilities in the mesopause layer, especially near the summer pole. In our time dependent calculation, waves with approximately 4-day period are generated in the $m = 1$ component, superimposing with the 1-day period tides.

1. INTRODUCTION

With the rising interest in the Earth's climate, the influential roles of the middle and upper atmosphere (MUA; altitude range of interest 20–200 km) have attracted increasing attention. Relative to the lower atmosphere, this part of the Earth's atmosphere contains little material. Since only 5% of the atmospheric mass lies above 20 km, external perturbations can induce large responses in this region. Such perturbations can be due to the variation of the solar radiative forcing, upward propagating waves from the lower atmosphere, auroral particle precipitation from the magnetosphere, or anthropogenically created chemical changes. This is a region where signals of atmospheric changes are more readily detected than in the lower atmosphere.

Because this upper region of the atmosphere is so tenuous, it is often believed that its perturbations cannot propagate downward to the lower atmosphere with significant effects. However, such an argument can only be applied to energy conserving interactions. The upper atmosphere, on the other hand, can act as a control valve between the Sun, the true energy source, and the lower atmosphere and thus exert its influence with leverage. Another way to view this influence is to consider the lower atmosphere as a dynamical system with upper boundary conditions controlled by the MUA. It is conceivable that the state of the dynamical system can be quite sensitive to these boundary conditions.

The MUA region referred to here is roughly defined to be between 20 and 200 km altitude. The middle atmosphere includes the upper stratosphere (20–50 km altitude) and the mesosphere (50–90 km altitude). The primary energy source of this region is the solar UV radiation absorbed by ozone, through which a multitude of phenomena (e.g. global circulation, tides) is induced. Most of these dynamical phenomena penetrate or propagate to the region directly above the lower thermosphere (roughly limited to be the 90–200 km altitude region) whose energetics is primarily controlled by the solar EUV radiation and magnetospheric sources. Transport of latent energy in the lower thermosphere contributes to the energy balance of the circulation that traverses the 90 km altitude boundary (KELLOG, 1961). The dynamics of the middle atmosphere and the lower thermosphere is therefore tightly coupled.

The lower atmosphere, in which water vapor plays a dominant role, has been an area of intense research for a long time due to its direct relevance to weather and climate near the Earth surface. Elaborate general circulation models (GCMs) have been constructed to study this lower region in much detail (see reviews by HALTINER and WILLIAMS, 1980; SIMMONS and BENGTSSON, 1988; ARAKAWA, 1988). Similarly, GCMs for the thermosphere have been developed

(FULLER-ROWELL and REES, 1980; DICKINSON *et al.*, 1984; ROBLE and RIDLEY, 1987 [TGCM]). Simulations of the middle atmospheric region have been carried out (e.g. MATSUNO, 1982; GARCIA and SOLOMON, 1985; AKMAEV *et al.*, 1992). So far, however, there is a lack of models that describe the coupled middle and upper atmosphere (although an effort in extending the NCAR TGCM to include the mesosphere has been initiated).

While GCMs are comprehensive in the inclusion of input physics and parameterizations, they are computationally expensive and inconvenient for performing parametric studies. However, such studies are essential for analyzing the cause and effects of the different processes and for developing understanding of the complex dynamical system. Our objective is to develop a versatile model which allows us to study the interplay of the dynamical components (e.g. mean circulation, planetary waves, gravity waves) and processes (e.g. radiation, wave breaking, chemistry) in the MUA region. Other parts of the atmosphere also influence this region, and at present, we treat such influences as boundary conditions. The present paper provides a progress report on our effort in the development of such a model.

Responding to the temporally and spatially varying external energy and momentum sources, a variety of atmospheric phenomena is generated in the MUA region. In the order of decreasing spatial and temporal scales (which are positively correlated), the major components of atmospheric activities can be categorized as: global circulation (zonally averaged; annual and semi-annual variations), planetary waves (zonal wave number <6; period one day to one month); thermal tides (zonal wave numbers = 1, 2 or 3 for diurnal, semi-diurnal, and ter-diurnal tides, respectively); acoustic/gravity waves (horizontal wavelength of a few km to a few thousand km; periods from a few minutes to a few hours); and three-dimensional turbulence (spatial scale $<$a few km; temporal scale $<$a few minutes). The dynamical system is extremely complex, as each of these components can interact with the others.

Our long term goal is to understand the generation and propagation of these components and to sort out their mutual interactions. Our strategy is to work with a family of complementary models that have a common root, each of which can be tailored to describe a specific phenomenon. The common root ensures that, if needed, some or all of these models can be consistently reassembled for more comprehensive simulations. We choose the 'root' of this family to be a newly developed three-dimensional, non-hydrostatic, non-linear, time-dependent, semi-implicit, transform spectral code that has few limitations so that, in principle, it can describe many of the phenomena observed. Besides having the well-known advantage of being more efficient in low and moderate resolutions, the spectral approach also offers the very important convenience of naturally decomposing the variable fields into separate zonal components. This decomposition makes it very easy to study the behavior of individual components of atmospheric motions and to analyze their interactions. For example, the interaction between the mean circulation and the diurnal tide is described by the non-linear coupling of the $m = 0$ and 1 components where m is the zonal wave number (see the later discussion).

It is fairly straightforward to generate specialized versions of this code and this convenience has been exploited for studying the mean circulation under widely different conditions. An axisymmetric version of the code has been applied to compute the spin-up of the Venusian atmosphere and gave reasonable results (CHAN and MAYR, 1990). Furthermore, a linear axisymmetric, fully coupled (Coriolis terms included in the Jacobian matrix; see Section 2) version of the code has been applied to study and reproduce the observed internal angular velocity distribution in the Sun (CHAN and MAYR, 1991).

In this paper, we describe the initial application of this modeling approach to a study of the MUA region. First, several versions (using different solution techniques) of the original spectral code are generated to provide solutions describing separately the mean circulation and the diurnal tide. The properties of the solutions are explored under different physical conditions. Then, the behavior of the diurnal tide under the influence of the mean circulation is studied with a code that couples the two components. At this stage, we are primarily interested in understanding the performance of the codes/models and their numerical/parametric properties. We have not tried to implement the most advanced input parameterizations for the various processes (e.g. radiative exchange, wave momentum deposition). Such improvements are left for later studies.

In Section 2, we provide a description of our spectral model. In Section 3, results are presented for the zonally averaged circulation. In Section 4, we present results for the thermally driven diurnal tides. We conclude with a short summary in Section 5.

2. THE NUMERICAL MODEL

The objective of our numerical model is to provide an efficient, dynamically rigorous analytical tool for studying the MUA region. A rigorous formulation of

the equations of motion allows for a flexible, consistent application of the numerical model, which is important because a wide variety of phenomena occurs in the MUA region. Efficiency is an important aspect of our approach because we are interested in the roles of the individual processes and the cause–effect relationships. It is necessary to obtain a large number of solutions for different parameters and conditions with affordable computing resources.

Details concerning the formulation, numerical implementation and testing of our model have been given in another paper (CHAN *et al.*, 1993). Here we provide a general outline which emphasizes the salient features and describe some specific modifications used for the present applications.

Our model uses a combination of the transform spectral (horizontal direction) and finite-difference (vertical direction) approaches to solve the equations of motion in spherical coordinates (r, θ, ϕ). Its prognostic variables are the density ρ, pressure p, vertical mass flux M_r, and the divergence δ $(=\nabla_H \cdot M_H)$ and curl ζ $(=\nabla_H \times M_H)$ of the horizontal mass flux M_H. The prognostic equations are derived from the Navier Stokes equations for a compressible fluid. Neither the hydrostatic nor the anelastic approximations (used to filter sound waves; see OGURA and CHARNEY, 1962; OGURA and PHILLIPS, 1962) are invoked. A 'stratified' approximation, however, is introduced which makes the following simplifications: (i) in the momentum equations, the horizontal variation of density is ignored in the non-linear advection terms; (ii) in the energy equation, terms containing products of two or more of the horizontal variations (perturbations from the mean) of the thermodynamical variables are ignored. This approximation serves several purposes: (i) it reduces the non-linearity of all terms to quadratic so that alias-free spectral transforms can be performed easily (MACHENHAUER, 1979); (ii) the conservation of total angular momentum is enforced both formally and numerically (due to alias-free transformations); (iii) the advection terms can be expressed in simple forms similar to those derived for spectral GCM models (e.g. BOURKE, 1974; WILLIAMSON, 1983).

The equations are:

$$\partial_t \rho = -\mathscr{D}_r M_r - \delta \tag{1}$$

$$\partial_t M_r = -\partial_r p - \rho g + C_r + D_r + N_r \tag{2}$$

$$\partial_t \delta = -\nabla_H^2 p + C_\delta + D_\delta + N_\delta \tag{3}$$

$$\partial_t \zeta = C_\zeta + D_\zeta + N_\zeta \tag{4}$$

$$\partial_t p = -\mathscr{D}_r(\langle p\rangle M_r/\langle p\rangle) - (\langle\Gamma\rangle - 1)\langle p\rangle \mathscr{D}_r(M_r/\langle\rho\rangle) - \langle\Gamma\rangle\langle p\rangle\delta/\rho + D_p + N_p + \nabla_{ad}\Gamma\varepsilon, \tag{5}$$

where ∂_t and ∂_r are the partial derivatives with respect to t and r, respectively; $\mathscr{D}_r = r^{-2}\partial_r r^2$; g is the magnitude of the gravitational acceleration; ∇_H^2 is the horizontal Laplacian operator; Γ is the adiabatic exponent defined by $(\mathrm{d}\ln p/\mathrm{d}\ln\rho)_{ad}$; $\nabla_{ad} = (\mathrm{d}\ln T/\mathrm{d}\ln p)_{ad}$ is the adiabatic temperature gradient; ε is the heating rate per unit volume. The brackets $\langle\ \rangle$ denote the horizontal average. The symbol C represents the Coriolis terms:

$$C_r = 2\Omega\sigma M_\phi \tag{6}$$

$$C_\delta = 2\Omega[\chi\zeta - r^{-1}(\sigma M_\phi + \partial_\phi M_r)] \tag{7}$$

$$C_\zeta = -2\Omega[\chi\delta - r^{-1}(\sigma M_\theta - \sigma\partial_\theta M_r - 2\chi M_r)] \tag{8}$$

where Ω is the angular velocity, and σ and χ stand for $\sin\theta$ and $\cos\theta$, respectively. The symbol D represents the dissipative terms:

$$D_r = 2r^{-2}\partial_r(\mu_r r^2\partial_r V_r) + \mu_H\nabla_H^2 V_r + \mu_r\partial_r\delta_V - 2\mu_H r^{-1}(\delta_V + 2r^{-1}V_r) + \partial_r[\lambda(\mathscr{D}_r V_r + \delta_V)] \tag{9}$$

$$D_\delta = r^{-4}\partial_r[r^4(\mu_r\partial_r\delta_V + \mu_H\nabla_H^2 V_r)] + 2\mu_H(r^{-1}\nabla_H^2 V_r + \nabla_H^2\delta_V + r^{-2}\delta_V) + \lambda[r^{-4}\partial_r(r^4\nabla_H^2 V_r) + \nabla_H^2\delta_V] \tag{10}$$

$$D_\zeta = r^{-4}\partial_r(r^4\mu_r\partial_r\zeta_V) + \mu_H(\nabla_H^2\zeta_V + 2r^{-2}\zeta_V) \tag{11}$$

$$D_p = \nabla_{ad}\Gamma_0[r^{-2}\partial_r r^2(k_{pr}\partial_r p + k_{\rho r}\partial_r\rho) + k_{\rho H}\nabla_H^2 p + k_{\rho H}\nabla_H^2\rho] \tag{12}$$

where $\delta_V = \nabla_H \cdot \mathbf{V}_H$ $(\approx\delta/\langle\rho\rangle)$ and $\zeta_V = \nabla_H \times \mathbf{V}_H$ $(\approx\zeta/\langle\rho\rangle)$. The radial and horizontal viscosities, μ_r and μ_H, are allowed to be different. λ is the second coefficient of viscosity. The energy flux is modeled as the negative gradient of the temperature which is a function of ρ and p; the radial and horizontal conductivities (k) are again allowed to be different. In the derivation of equations (9)–(12), the horizontal variations of the viscosities and conductivities are neglected (but can be included as additions to the non-linear terms as discussed later). In the present formulation, the dissipative terms are linear in the prognostic variables and can then be treated implicitly. Implicit treatment of these terms is important for the thermosphere where the molecular diffusivity is large and quite uniform. Otherwise, the numerical stability condition associated with the explicit treatment of the diffusion terms would impose a very severe time step restriction on the integration of equations (1)–(5) (time step <1 s).

Finally, the symbol N represents the non-linear

terms:

$$N_r = -\mathscr{D}_r(M_r V_r) - \nabla_H \cdot (M_r \mathbf{V}_H) + r^{-1} M_H V_H \quad (13)$$

$$N_\delta = -r^{-4}\partial_r[r^4 \nabla_H \cdot (M_r \mathbf{V}_H)] + \nabla_H(\zeta \mathbf{V}_H) - \nabla_H \cdot (\delta \mathbf{V}_H) - \nabla_H^2 (M_H V_H)/2 \quad (14)$$

$$N_\zeta = -r^{-4}\partial_r[r^4 \nabla_H \times (M_r \mathbf{V}_H)] - \nabla_H \cdot (\zeta \mathbf{V}_H) - \nabla_H \times (\delta \mathbf{V}_H) \quad (15)$$

$$N_p = -\nabla \cdot (p' \mathbf{V}) - (\langle \Gamma \rangle - 1) p' \nabla \cdot \mathbf{V} + \nabla \cdot (\langle p \rangle \rho' \mathbf{V} / \langle \rho \rangle) + (\langle \Gamma \rangle - 1) \langle p \rangle \nabla \cdot (\rho' \mathbf{V} / \langle \rho \rangle) - (\langle \Gamma \rangle - 1) \langle p \rangle \nabla \cdot \mathbf{V} \quad (16)$$

where $\mathbf{V}$ is approximated by $\mathbf{M}/\langle \rho \rangle$. The superscript $'$ denotes horizontal variations.

The linear terms preceding the C terms in equations (1)–(5) contain the description of acoustic and gravity waves, and for convenience, we shall identify them with W. In the present formulation, all of the linear terms that appear in the original set of the Navier Stokes equations remain intact; therefore all the waves generated by these terms (including Rossby waves) can be described exactly.

Simplifying the notation, we use u to represent the set of prognostic variables. The dynamical equations can then be written in the following compact form:

$$\partial_t u = (W + D + C)u + N[u] + Q \quad (17)$$

where Q represents an external driver which includes the periodic solar radiative forcing. The spectral approach employs a finite series expansion of u in spherical harmonics:

$$u(r, \theta, \phi) = \sum u_n^m(r)\, Y_n^m(\theta, \phi). \quad (18)$$

The variables M_θ and M_ϕ used in equations 13–16 can be obtained by:

$$(\sigma M_\theta)_n^m = r\{-c(n,m)\delta_{n-1}^m + c(n+1,m)\delta_{n+1}^m + im[n(n+1)]^{-1}\zeta_n^m\} \quad (19)$$

$$(\sigma M_\phi)_n^m = r\{-c(n,m)\zeta_{n-1}^m + c(n+1,m)\zeta_{n+1}^m - im[n(n+1)]^{-1}\delta_n^m\} \quad (20)$$

where $c(n, m) = n^{-1}[(n^2 - m^2)/(4n^2 - 1)]^{1/2}$. These diagnostic variables, together with all the prognostic variables, are transformed to physical space for the evaluation of N.

Since the spherical harmonic functions Y_n^m are the eigenfunctions of W and D, a semi-implicit numerical treatment of equation (17) is possible, such that a time discretized approximation of this equation can be written as

$$[1 - \beta \Delta t(W + D)](u^{q+1} - u^q) = \Delta t\{(W + D + C)u^q + N[u^q] + Q(t^q)\} \quad (21)$$

where the upper index q denotes the time level; Δt is the time step; and β is a factor that determines the implicitness of the scheme. When $\beta = 0$, the time stepping algorithm corresponds to the explicit Euler scheme. When $\beta > 0.5$, the scheme is stable even when the time steps are larger than those imposed by the Courant conditions associated with the gravity and acoustic waves (which would limit Δt to less than a few seconds in our application). The Jacobian matrix $[1 - \beta \Delta t(W + D)]$ does not couple the different spherical harmonic components (since Y_n^m are eigenfunctions); it only couples variables of neighboring altitude levels. Therefore the matrix has a narrow block-tri-diagonal structure and can be inverted in a straightforward manner (LAPIDUS and SEINFELD, 1971). The large time steps and fast inversions are the reasons why the solution of equation (21) is computationally efficient, although most of the computer time is used in computing the non-linear terms N which require forward and backward transformations between the spectral and physical spaces.

The Coriolis terms C are not included in the implicit treatment of equation (21) because they couple spherical harmonics with neighboring degrees $(n \pm 1)$. Their inclusion would enlarge the tri-diagonal blocks of the Jacobian matrix by a large factor and drastically increase the computer time for inverting the matrix. For some applications, however, an implicit treatment of the Coriolis terms is desirable. In the linear study of a thermal tide, the longitudinal variation can be described by $e^{im\phi}$ and the temporal variation by $e^{im\omega t}$ where m is the zonal wave number ($=1, 2$ for diurnal and semi-diurnal tides, respectively), ϕ is the longitude, and $\omega = 2\pi$ day^{-1}. Replacing u by $u_m e^{im\omega t}$ in equation (17) and neglecting the non-linearities, one obtains

$$[im\omega - (W + D + C)]u_m = Q_m \quad (22)$$

where Q_m is the complex amplitude of the thermal heat source. In this case, the solution is obtained by inverting the Jacobian matrix only once, which can be more efficient than marching in time. By taking $m = 0$, this method allows us also to compute the zonally averaged circulation generated linearly by a mean heat source. In the next two sections, we use this method to compute the mean circulation and the diurnal tide (without background winds). For comparison, the mean circulation is also computed by the time marching scheme [equation (21)] with and without

the non-linear terms. The influence of the mean circulation on the diurnal tides is then studied. Contrasting these different solutions can illustrate the relative importance of different processes.

3. MEAN CIRCULATION

The mean (zonally averaged) global circulation generally takes on the form of meridional cells with differentially heated fluid rising. The pattern of this circulation changes with annual and semi-annual cycles and is affected by internal energy and momentum sources due to tides and waves. In the mesosphere, zonal winds are generated more or less according to geostrophic balance; and the amplitude in the annual variation has a peak of about 70 m/s near 45° latitude and 60 km altitude. In the thermosphere, the zonal winds become ageostrophic due to the large molecular viscosity and ion drag.

The global circulation dramatically influences the energy balance and temperature distribution of the MUA. An important effect is seen in the 'mesospheric temperature anomaly' at the mesopause level (~90 km) where the polar temperature in summer is 80 K lower than that in winter, contrary to the expectation based purely on solar radiative heating. The energy budget for this temperature anomaly is balanced by a diabatic circulation (MURGATROYD and GOODY, 1958; DUNKERTON, 1978). It is generally believed that the momentum balance of this circulation is maintained by breaking gravity waves from the lower atmosphere (LINDZEN, 1981; HOLTON, 1982; CHAO and SCHOEBERL, 1984; SCHOEBERL, 1988).

A great deal of work has been done to model the mean circulation of the stratosphere and mesosphere. To account for the generation of the mesospheric temperature anomaly, the effects of gravity wave breaking and wave mean-flow interaction have been considered (e.g. MATSUNO, 1982; GARCIA and SOLOMON, 1985) or parameterized in terms of eddy diffusion and Rayleigh friction (e.g. HOLTON and WEHRBEIN, 1980; GARCIA and SOLOMON, 1983). The calculations we perform are not intended for extending the results of previous studies. We are mainly interested in knowing to what extent our simplified spectral model can reproduce these results.

For the purpose of this study, a two-dimensional spectral model is derived by restricting the harmonic

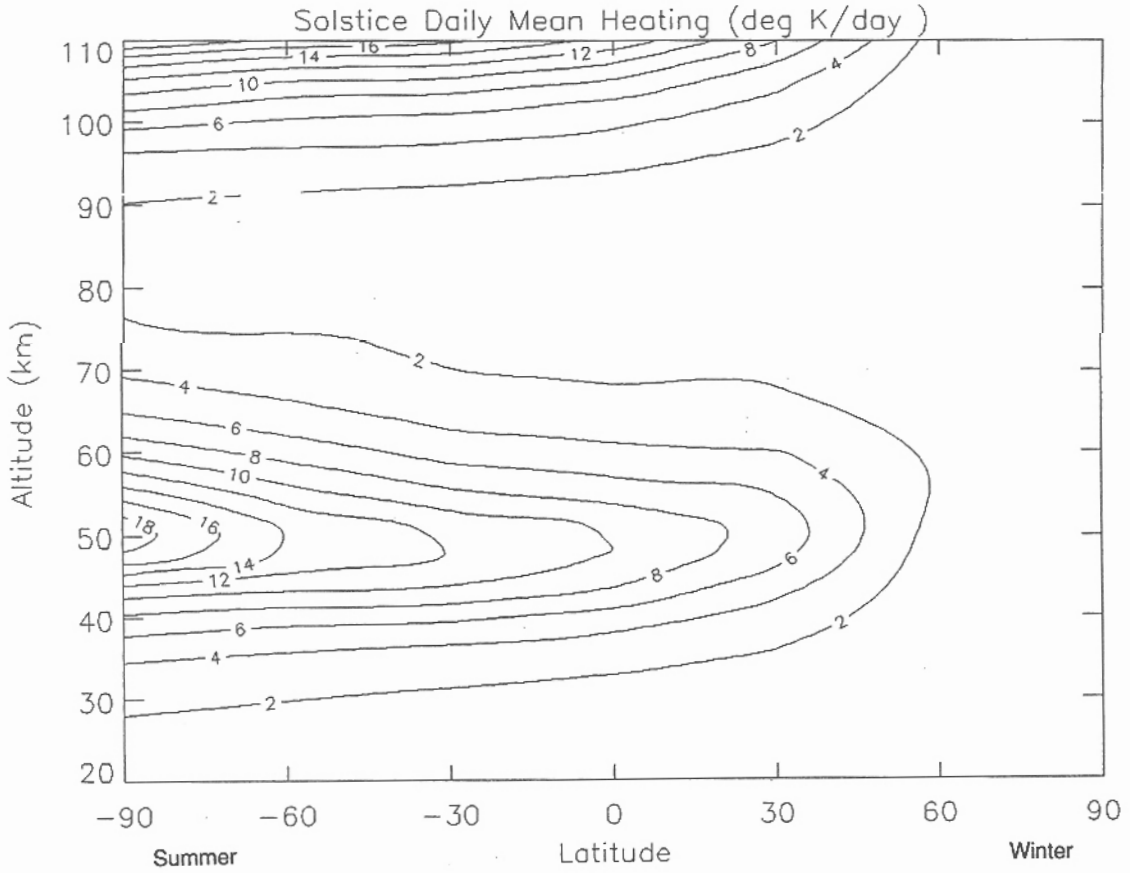

Fig. 1. Distribution of the heat source used in computing the zonally averaged circulation under perpetual solstice conditions (southern summer). Note that the negative latitudes denote southern hemisphere.

expansion to the zonal wave number 0. The globally averaged temperature and density are taken as functions of height and treated as fixed background. The background temperature asymptotically reaches 1000 K in the upper thermosphere, corresponding to a medium solar activity condition. The circulation is considered as a perturbation of the background. Our focus is on the mesospheric temperature anomaly generated under solstice conditions. Since the important time constants are generally shorter than a season, we assume perpetual solstice. The circulation is primarily driven in the upper stratosphere and mesosphere. However, the altitude domain of our model is from the surface to 600 km, allowing for sponge layers that reduce the artificial constraints from boundary conditions. At the surface, the vertical and horizontal winds are set to zero; at the top, the vertical gradients of temperature and horizontal winds are taken to be zero, due to the large molecular conductivity and viscosity in the upper thermosphere. Ten spherical harmonics are included in the spectral expansion [n in equation (18) $\leqslant 10$], and 226 vertical levels are used (eight levels per scale height); they turn out to be adequate for resolving the large-scale atmospheric response to the differential solar heating.

The heat source driving the circulation is obtained from a formulation derived by STROBEL (1978) which includes components associated with the O_3 and O_2 absorption of the UV radiation. Above 75 km, only the net heating is included, assuming that the dissociation energy is effectively removed through transport. The O_3 distribution given by Strobel is slightly modified by removing the small concentration peak near 80 km. In addition, the rates are modified by applying a scaling factor chosen to bring the heating at 50 km into agreement with the work of ROSENFIELD *et al.* (1987). The latitudinal and vertical distribution of the heating is shown in Fig. 1. Above 120 km, EUV heating and cooling due to O_2, NO, and CO_2 are included from the work of PESNELL (1993).

Following the earlier work of GARCIA and SOLOMON (1983), we first use eddy diffusion and Rayleigh friction rates that vary monotonically with altitude [Fig. 2(a), 2(b)]. The eddy diffusion coefficient is assumed to vary inversely with the square root of density and is limited by a value of 3×10^6. The turbulent Prandtl number is taken to be 3. Radiative loss is parameterized in the form of Newtonian cooling, with a distribution [see Fig. 2(c)] based on the work of DICKINSON (1973) and WEHRBEIN and LEOVY

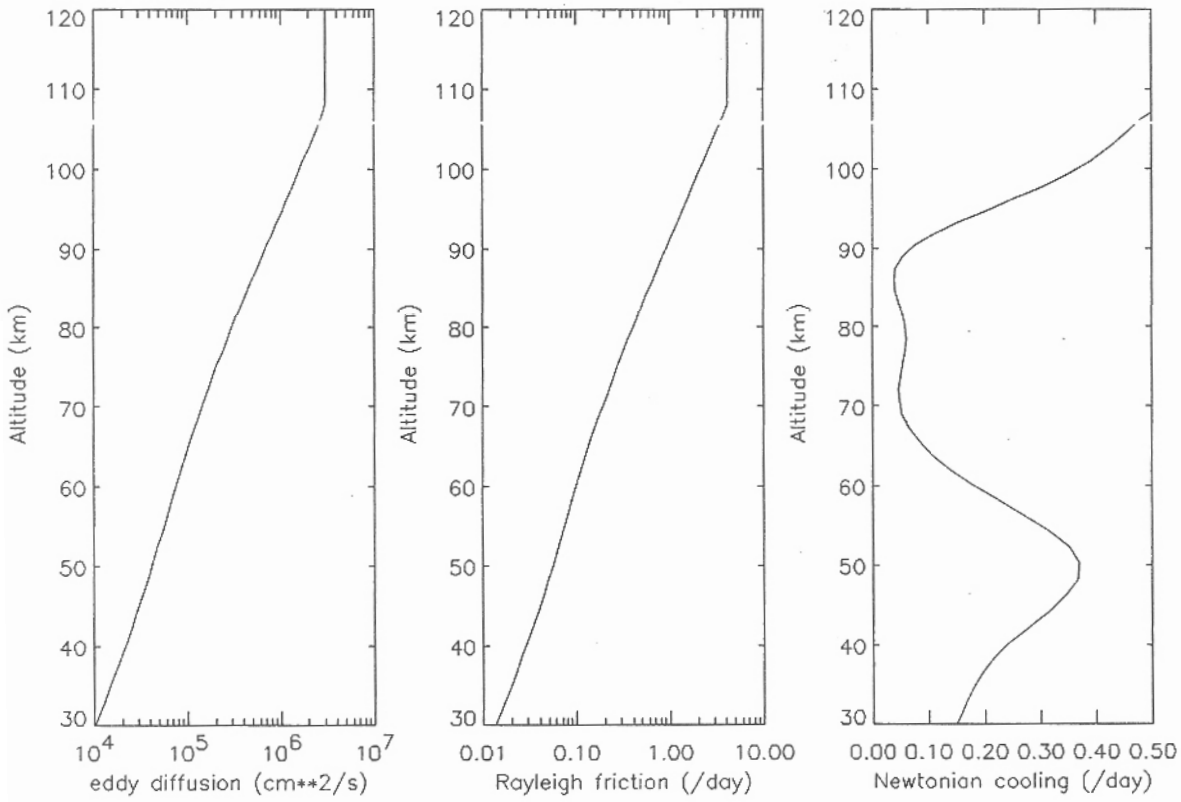

Fig. 2. Vertical distributions of (a) eddy diffusivity, (b) Rayleigh friction, and (c) Newtonian cooling for computing the circulation as depicted in Figs 3 and 4. The eddy diffusivity and Rayleigh friction are monotonic functions of altitude.

(1982). The molecular viscosity and conductivity are computed by the empirical fits given by BANKS and KOCKARTS (1973).

The temperature, the zonal, meridional and vertical velocities computed by equation (22) are presented in Figs 3(a), (b) and 4(a), (b), respectively. They show a close resemblance to the results obtained by GARCIA and SOLOMON (1983). The polar contrast associated with the mesospheric temperature anomaly in the 80–90 km region is only about 30 K. The zonal velocities are zero at the equator, due to the lack of horizontal advection, Coriolis force, and diffusion (relatively small in our case) needed for acceleration (HIDE, 1969).

In a later model, GARCIA and SOLOMON (1985) accounted for the effects of gravity wave momentum deposition and produced a larger temperature reversal (anomaly). There is no guarantee that one could simulate the gravity wave processes with Rayleigh friction and eddy diffusion, yet our numerical experiments indicate that Rayleigh friction distributions that have a hump near 90 km, as shown in Fig. 5(b), are favorable for producing a large temperature anomaly. This is compatible with the distributions of wave momentum deposition computed by gravity wave breaking models (e.g. see HOLTON, 1982; GARCIA and SOLOMON, 1985). The computed temperature and velocity fields obtained with parameterizations described by Fig. 5 are presented in Figs 6 and 7. The mesospheric temperature contrast now reaches 80 K, much closer to that observed. This is caused by the sharper meridional jet from the summer to winter hemisphere and the stronger vertical flows at the poles. Compared to observations, the zonal wind amplitudes are slightly lower and the peaks locate at higher altitudes. Improvements could be made by fine tuning the distribution of the Rayleigh friction. However, that is probably not worth the effort since the Rayleigh friction parameterization is not a faithful representation of wave effects. At present, we are content with reproducing the major features qualitatively and leave a detailed matching with observation to later modeling efforts which implement more realistic description of gravity wave breaking effects.

Up to this point, we have discussed only linear and stationary solutions obtained from a direct matrix

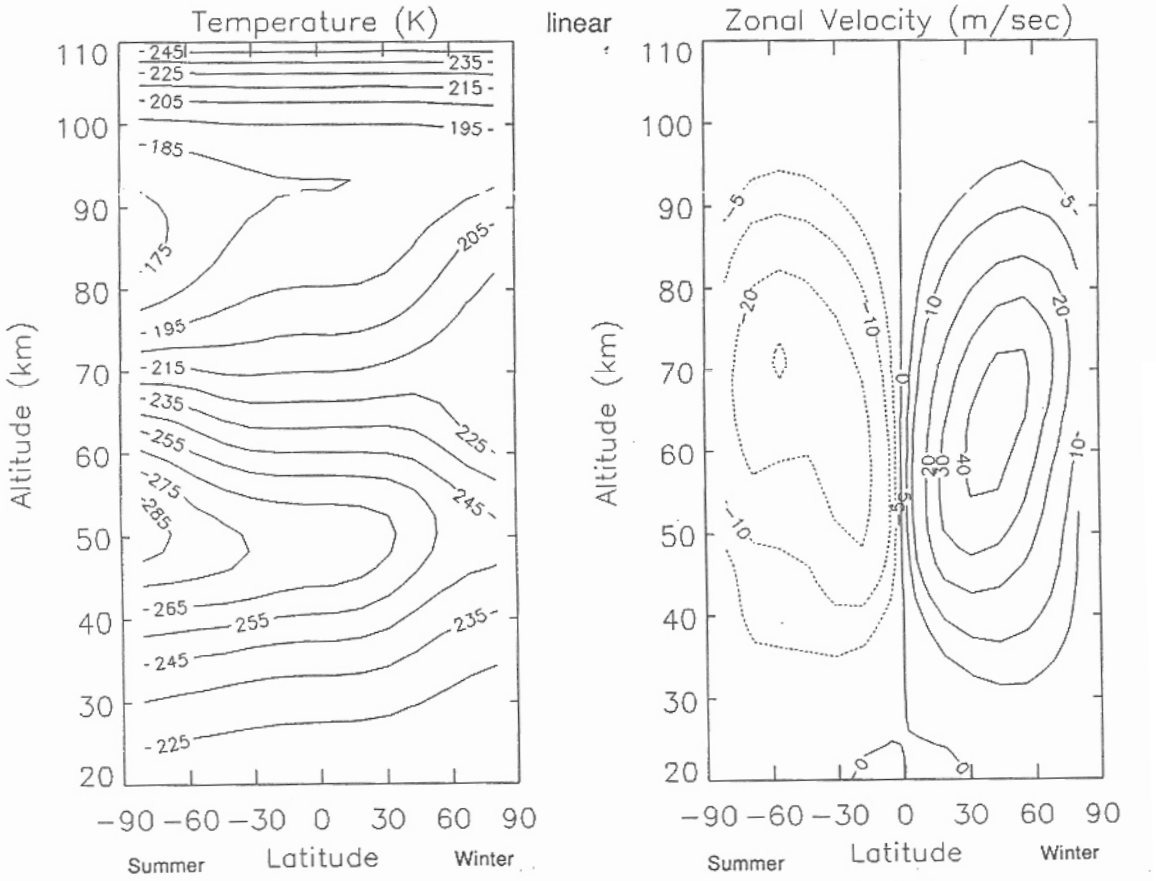

Fig. 3. Distributions of (a) temperatures and (b) zonal velocity of the mean circulation, computed with the parameterizations shown in Fig. 2. The polar temperature differential in the mesopause region is only 30 K.

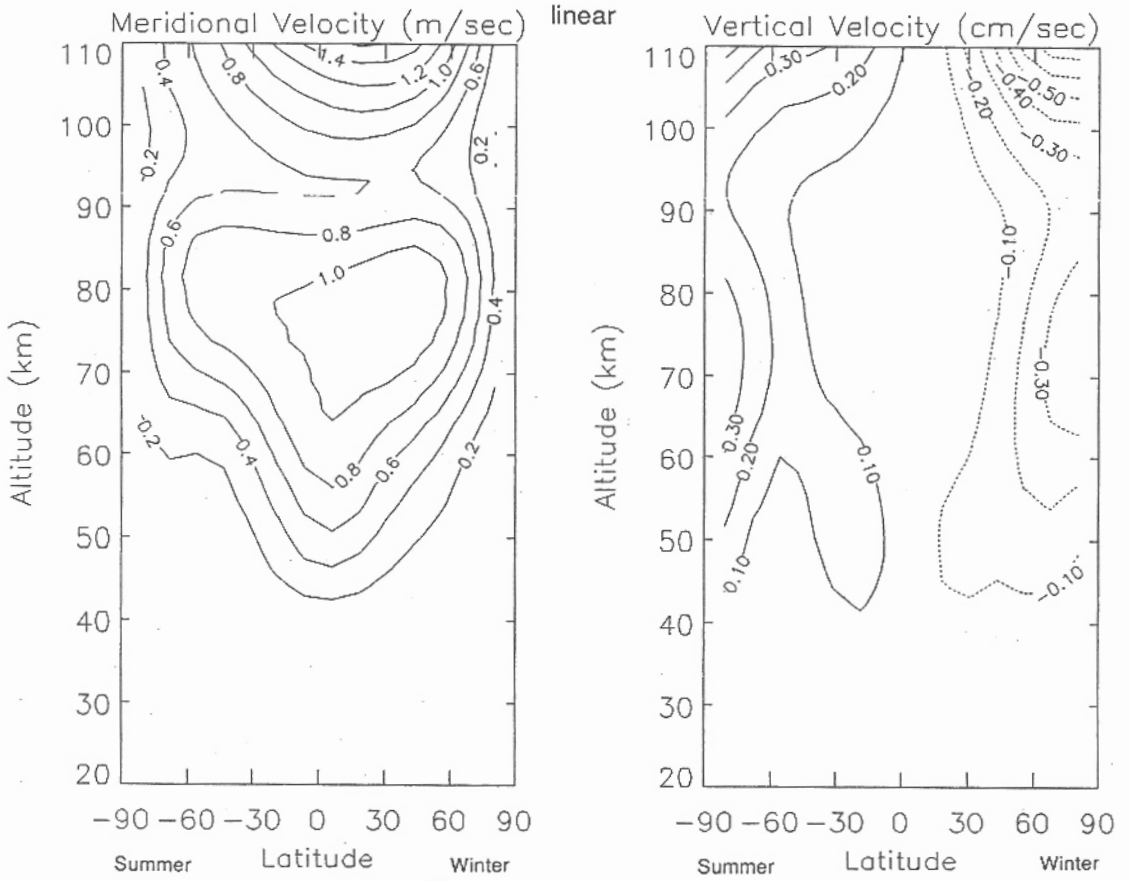

Fig. 4. Distributions of (a) meridional velocity and (b) vertical velocity of the mean circulation, computed with the parameterizations shown in Fig. 2.

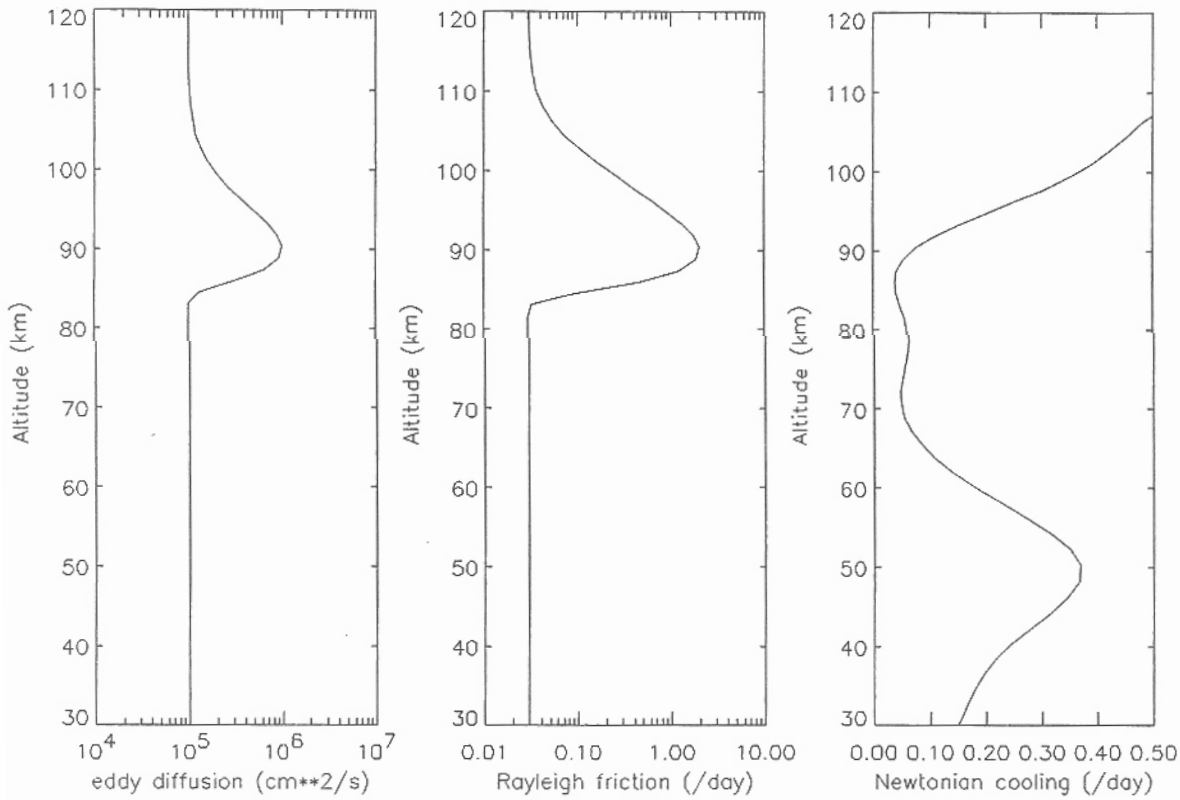

Fig. 5. Vertical distributions of (a) eddy diffusivity, (b) Rayleigh friction, and (c) Newtonian cooling for computing the circulation as depicted in Figs 6 and 7. The eddy diffusivity and Rayleigh friction are characterized by humps near 90 km to mimic the steep deposition of momentum and generation of turbulence due to gravity wave breaking.

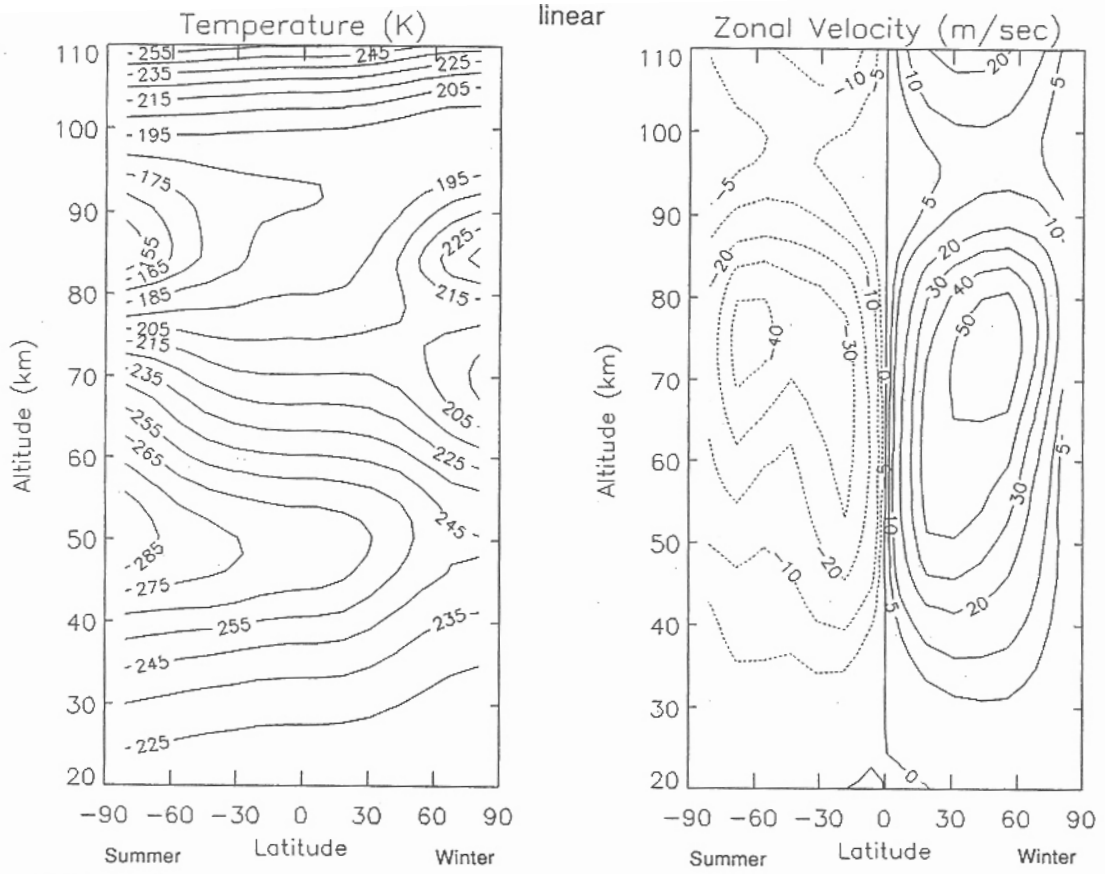

Fig. 6. Distributions of (a) temperature and (b) zonal velocity of the mean circulation, computed with the parameterizations shown in Fig. 5. The amplitude of the temperature anomaly at the mesopause region reaches 80 K.

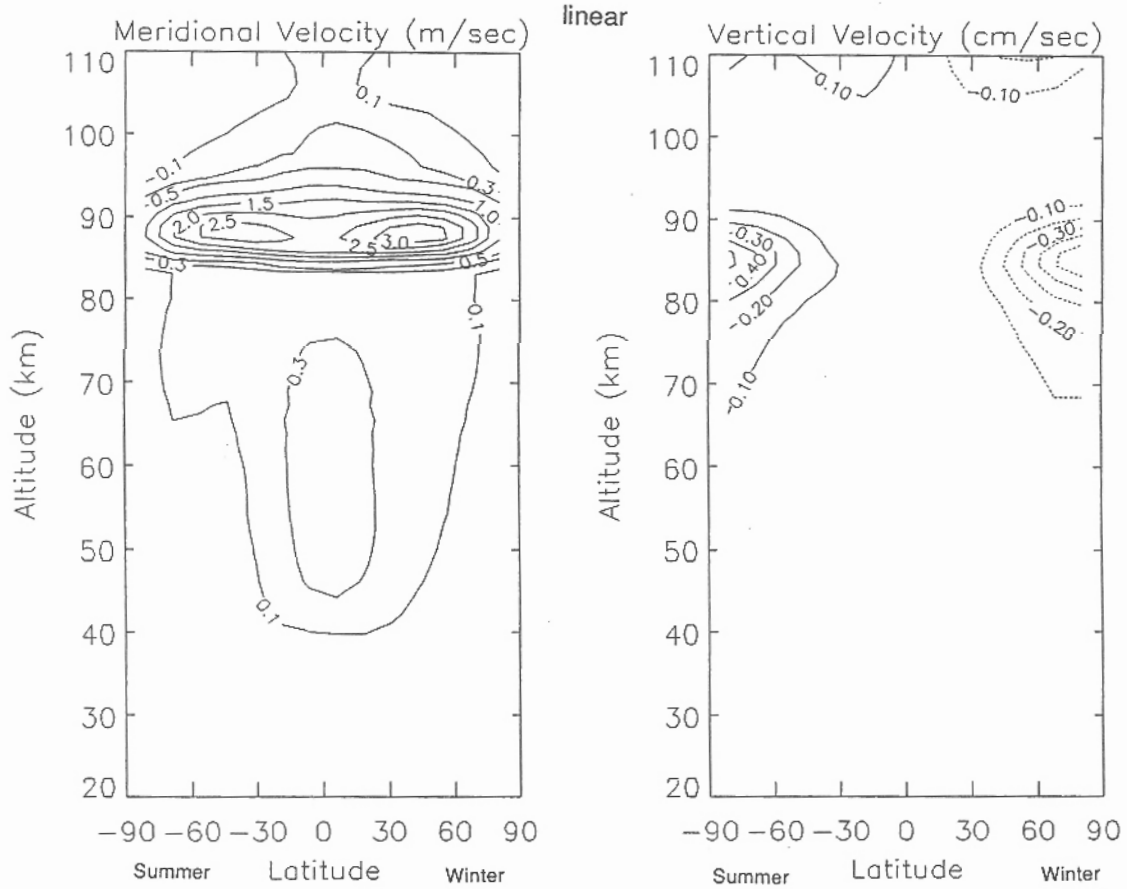

Fig. 7. Distributions of (a) meridional velocity and (b) vertical velocity of the mean circulation, computed with the parameterizations shown in Fig. 5. The meridional circulation forms a thin jet flowing from the summer to winter pole.

inversion according to equation (22). If we want to include non-linearities (though still axisymmetric), we need to solve equation (21). With the parameterizations specified in Fig. 5, we integrate equation (21) in time both with and without the non-linear terms N. The time-dependent solutions converge to stationary states in about 30 days. While the numerical procedures associated with equations (21) and (22) differ significantly, the results without the non-linearity are essentially identical to those obtained earlier, thus adding confidence in the integrity of the computer codes.

The non-linear results show some interesting differences. They are presented in Figs 8 and 9. First, one sees that the zero line of the zonal velocity is pushed by the non-linear advection to bend away from the equator, angular momentum being delivered across the equator. Second, the temperature contrasts between summer and winter are slightly reduced due to the non-linear redistribution of energy. This is better illustrated in Fig. 10 which shows the temperature gradients at the poles for the different models: (i) linear, with solid line for summer pole, dot–dashed line for winter pole; and (ii) non-linear, with long dashed line for summer pole, short dashed line for winter pole; the background temperature gradient is shown by the dotted line. The relative shifts in the positions of the maxima and minima are compatible with the directions of the vertical winds. The non-linearities produce smaller deviations from the background, but in general, the linear and non-linear results are qualitatively similar.

In a Micro-Vax 3100-80 machine, the computer times needed for finding converged solutions using equation (21) (30 days, 1 h time step), with and without the non-linear terms, are 70 and 25 min, respectively. Only 8 min are required for a linear solution of equation (22).

4. TIDES

Thermal tides are the subharmonic oscillations of a solar day and are primarily excited by the diurnal variation of the insolation absorption. The primary

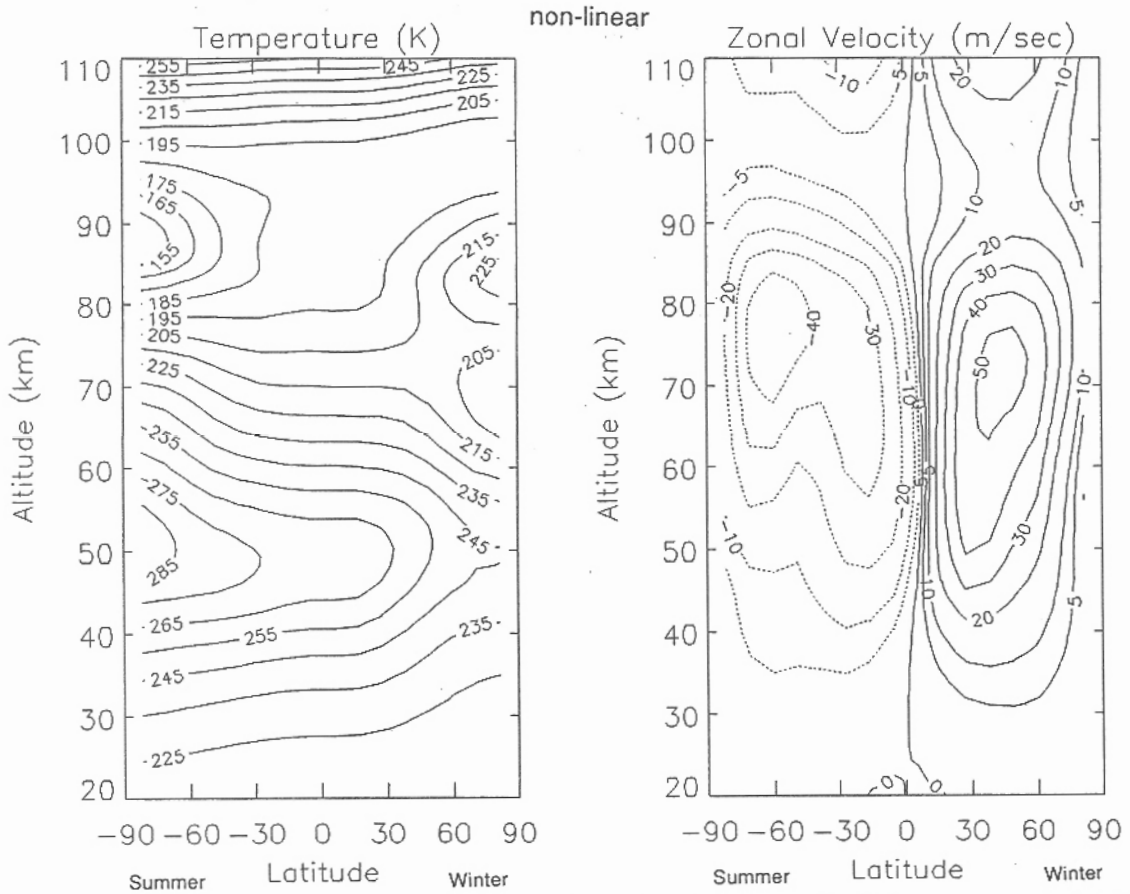

Fig. 8. Distributions of (a) temperature and (b) zonal velocity of the mean circulation, computed with the inclusion of non-linear terms. Parameterizations are the same as in Fig. 5. The magnitude of the temperature anomaly at the mesopause is slightly reduced. A zero line of the zonal velocity that used to lie on the equator is distorted by the horizontal advection.

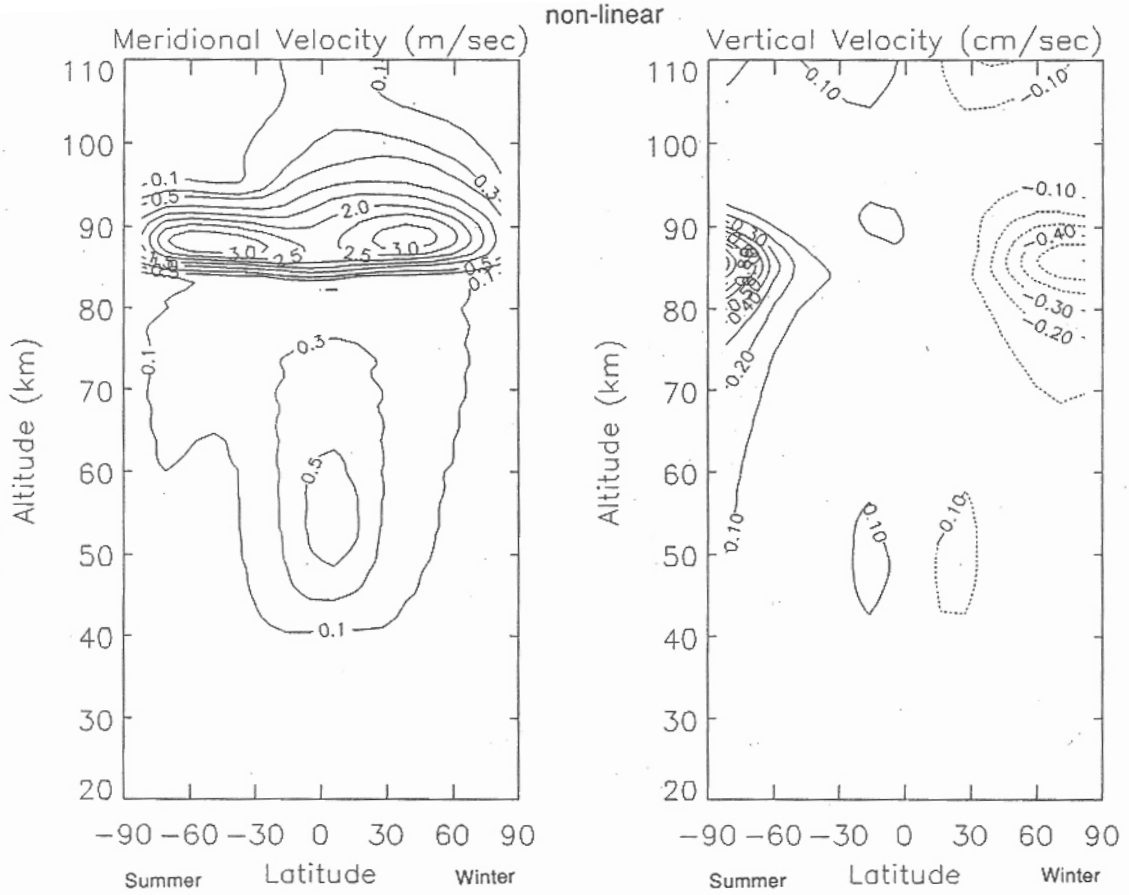

Fig. 9. Distributions of (a) meridional velocity and (b) vertical velocity of the mean circulation computed with the inclusion of non-linear terms.

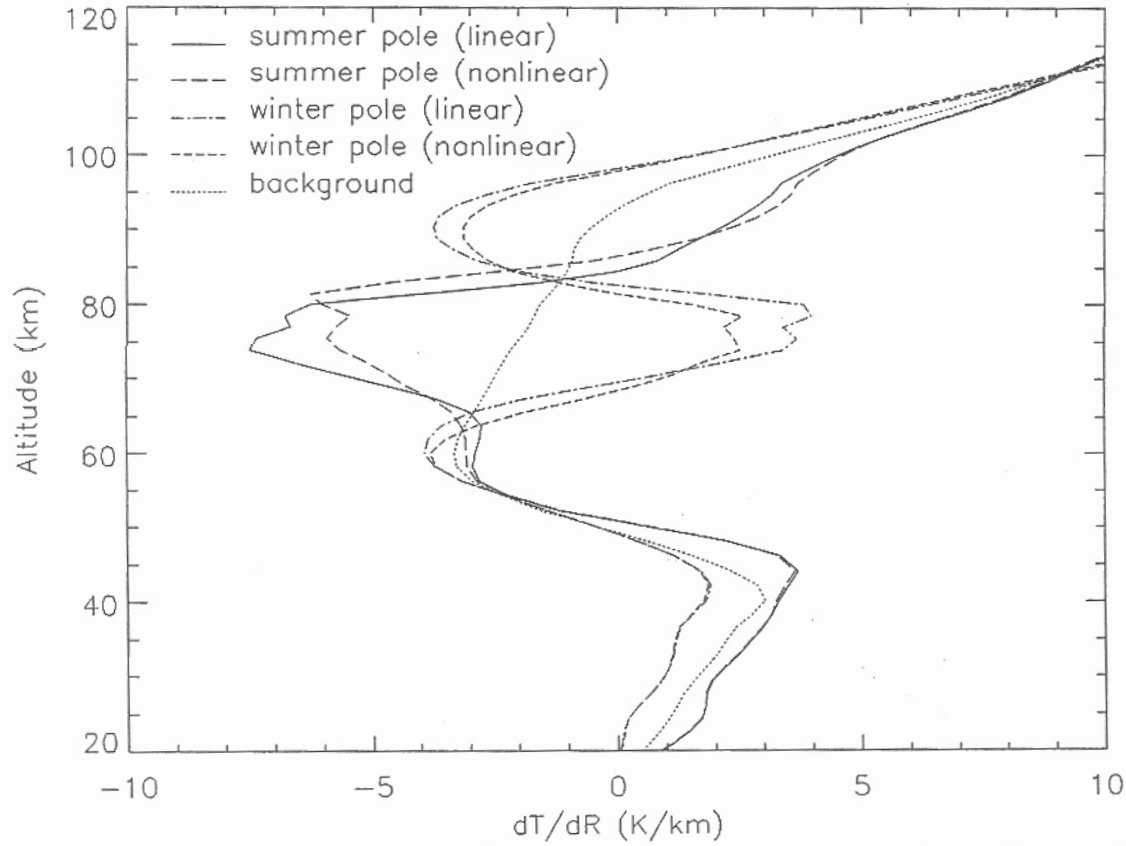

Fig. 10. Comparison of the temperature gradients at the summer and winter poles. The non-linear results show a tendency to limit the amplitudes of the perturbation. The vertical advection slightly displaces the locations of the maxima and minima.

drivers of the diurnal and semi-diurnal tides are the water vapor absorption of the solar near-IR radiation below 15 km and the ozone absorption of UV radiation between 20 and 60 km altitude. Tides can be decomposed into trapped and propagating eigenmodes identified as Hough modes (CHAPMAN and LINDZEN, 1970). When the propagating modes travel upward, their amplitudes experience exponential growth with height due to the decrease in the ambient density, and they become the most prominent dynamical feature of the upper mesosphere and thermosphere. When the propagating tides arrive at the lower thermosphere, dissipation by molecular diffusion and turbulence becomes important. The tides generated above 150 km tend to be evanescent due to the importance of molecular viscosity and ion drag, and they are predominantly driven by EUV absorption (see review by VOLLAND, 1988).

Based on classical tidal theory, tides can be described in terms of Hough modes which are eigenfunctions of the tidal equation. But that is only valid under simplifying assumptions, mainly: no viscosity, background winds, or horizontal temperature variations. Under more realistic conditions, multidimensional numerical computation is required. For this purpose, finite difference models have been employed by FORBES (1982), VIAL (1986), FESEN *et al.* (1986), MIYAHARA and WU (1989), and others. Spectral models have been employed in different forms, using spherical harmonics (e.g. MAYR and HARRIS, 1978) or Hough modes (e.g. WALTERSCHEID and VENKATESWARAN, 1979).

An important test of the numerical model is to demonstrate that, under the simplifying assumptions mentioned above, our model can reproduce the results of classical tidal theory. Minimizing the dissipation associated with eddy diffusion and eliminating the Rayleigh friction as well as the vertical component of the Coriolis force, the results for a propagating tide excited with an S_1^1 source in the water vapor layer near the surface are shown in Fig. 11. (A very small kinematic viscosity in the background, $\nu = 3 \times 10^2\,\mathrm{cm^2\,s^{-1}}$, is needed to provide pivoting for inverting the matrix associated with equation (22).

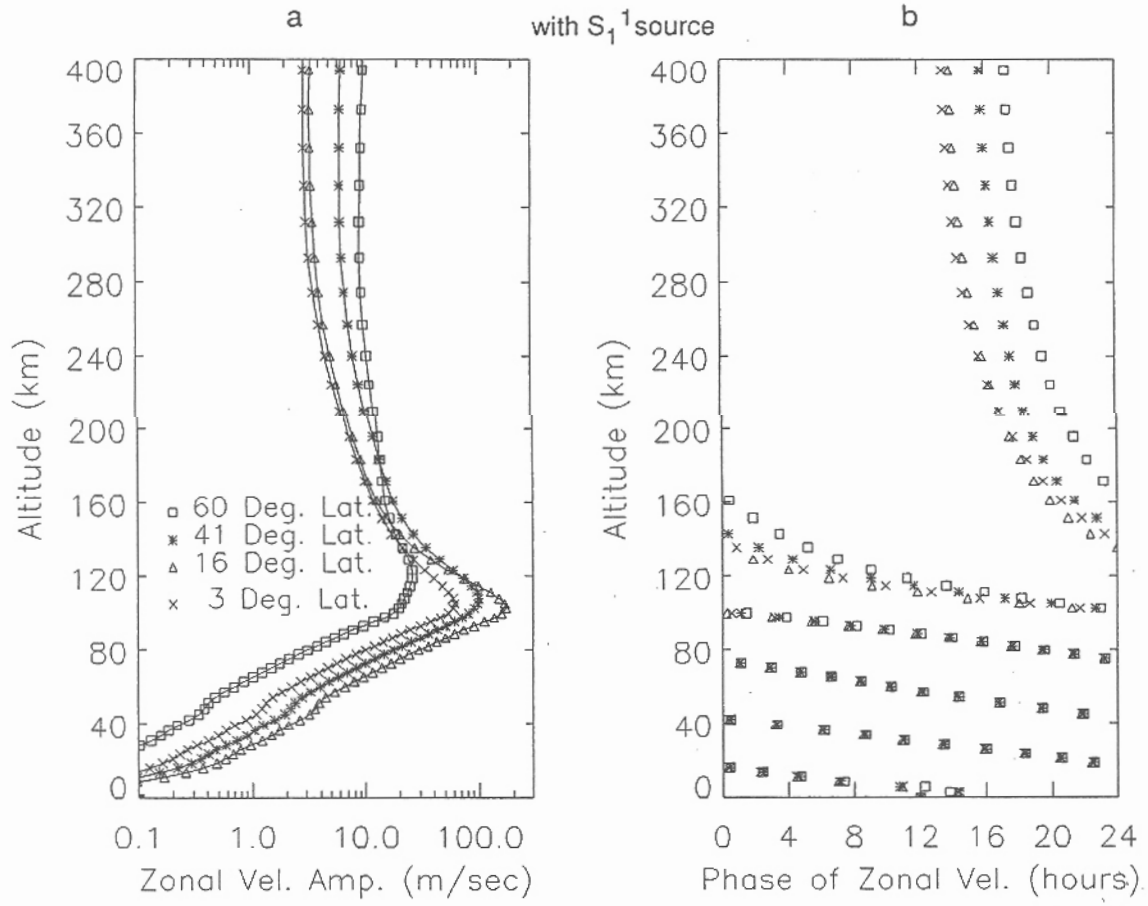

Fig. 11. A propagating tide excited with only the S_1^1 source in the water vapor layer near the surface. The squares, stars, triangles and crosses are for quantities at 60, 41, 16, and 3° latitude, respectively (the same hereafter). (a) Amplitudes; (b) phases.

The errors thus introduced are about 0.25% in amplitude and 4 s in phase.) The amplitudes and phases of the zonal velocity at 3, 16, 41 and 60° latitudes are plotted vs. altitude. The phase propagation reveals a vertical wavelength of about 35 km. Up to 100 km, while the amplitudes grow exponentially, their ratios for different latitudes remain constant and the phases coincide, showing that the horizontal structure of the solution remains the same, that of the Hough function. This demonstrates that only the S_1^1 mode is excited, as expected from classical tidal theory, and that the effects of the background viscosity are indeed negligible. At higher altitudes, however, molecular diffusion and ion drag become important, thus exciting other modes as seen from the phase digression and the changing amplitude ratios. At altitudes above 200 km, dissipation becomes so important that the tide becomes evanescent. The number of spherical harmonics used here is 10, and the vertical resolution is 8 levels per pressure scale height. Comparing the results to those with higher resolutions (up to 16 spherical harmonics, 12 levels per scale height), we find that the truncation errors are less than 5% in amplitude and half an hour in phase.

A similar analysis is also carried out for an S_{-2}^1 source in the water vapor layer. As shown in Fig. 12, a single trapped mode is excited in the lower atmosphere; the amplitude ratios and phases remain locked up to 40 km. Above this altitude, effects of mode mixing due to the non-zero dissipation (though small) become apparent. As indicated by the phase progression, propagating modes are excited and turn dominant at higher altitudes. The asymptotic amplitudes in the upper thermosphere increase with higher background viscosity. Fortunately, the amplitudes of these extraneous modes are small compared to those of the S_1^1 mode excited by the thermal forcing (see Fig. 11).

In classical tidal theory, the contributions from the vertical velocity to the Coriolis force and the vertical component of the Coriolis force are neglected. When we include these in the model, mode mixing is noticeably enhanced in the S_{-2}^1 calculation. As shown in Fig. 13, the mixed modes are excited with larger amplitude,

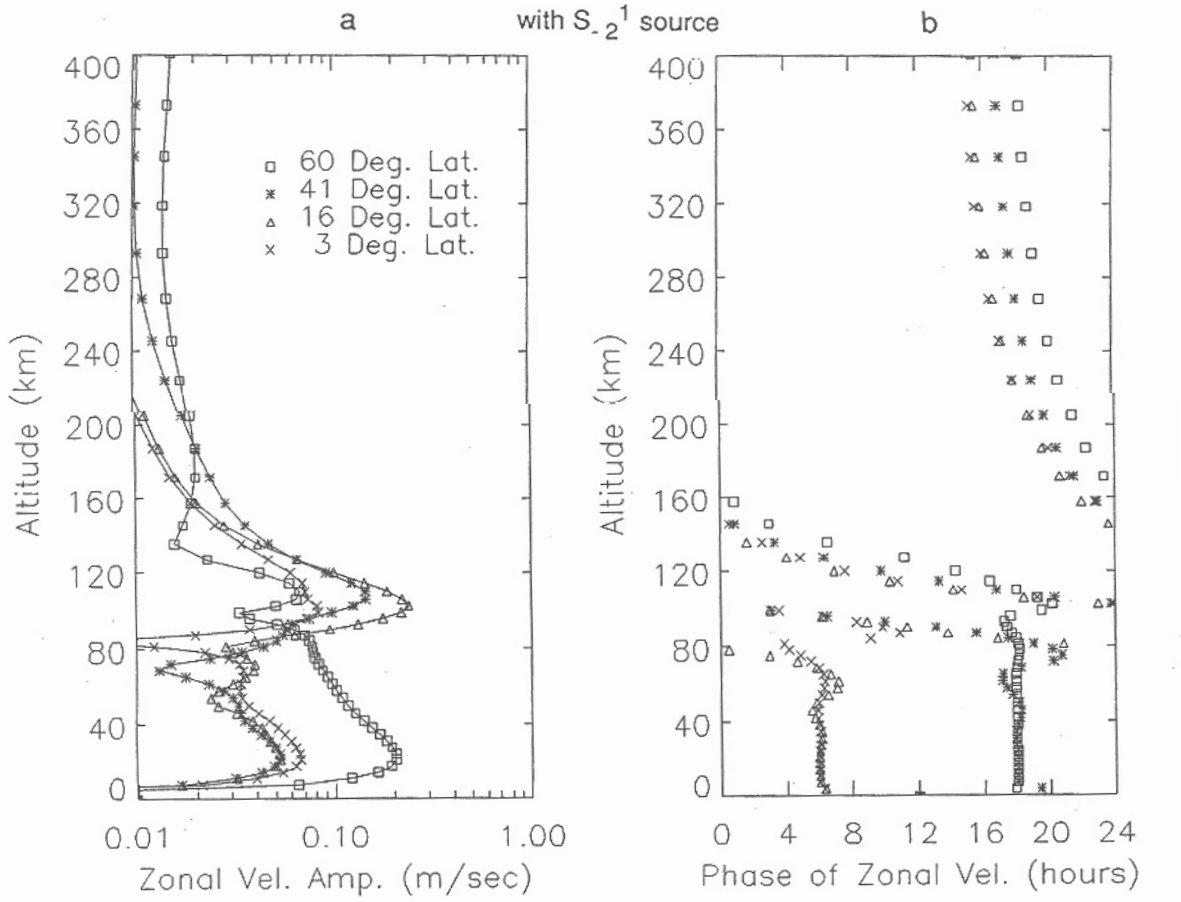

Fig. 12. A trapped mode excited with only the S_{-2}^1 source in the water vapor layer near the surface. The vertical component of the Coriolis force as well as the contributions from the vertical velocity to the Coriolis force are neglected. (a) Amplitudes; (b) phases.

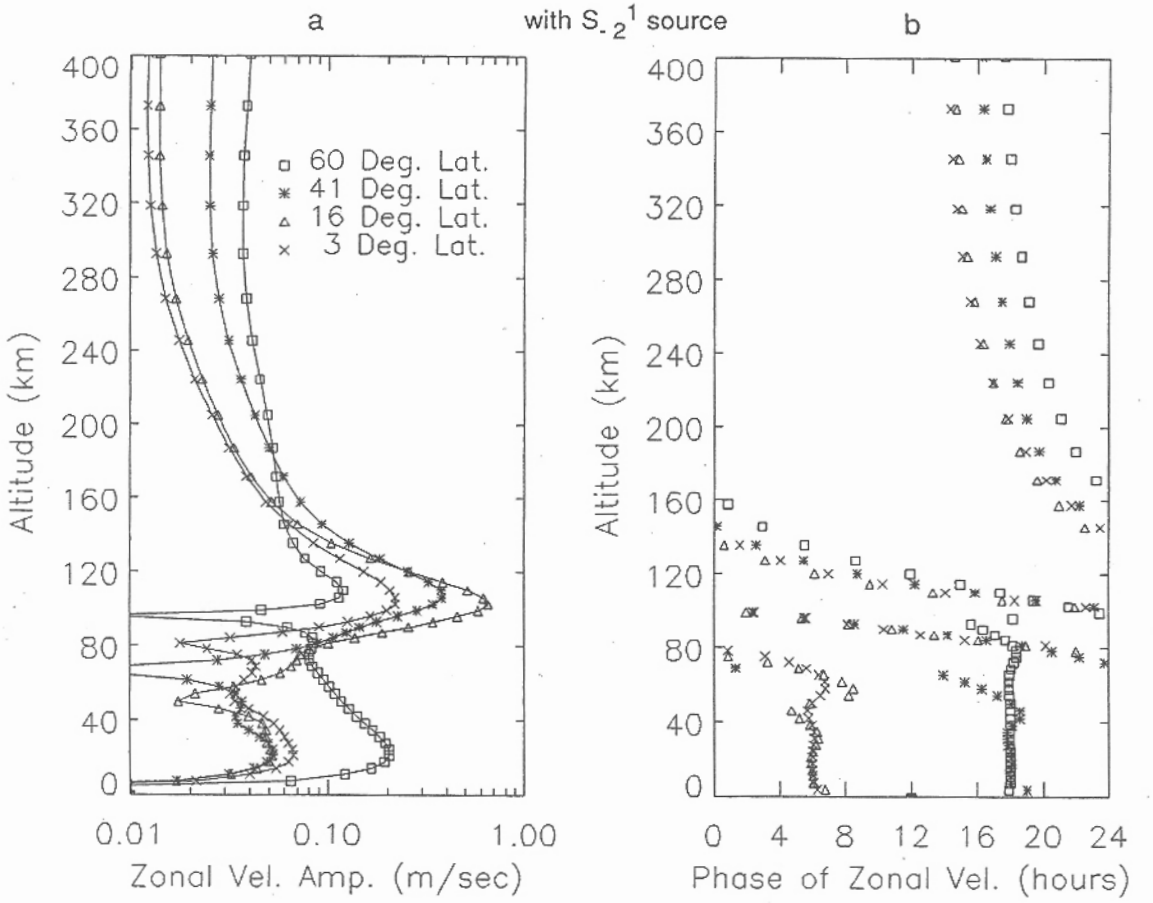

Fig. 13. The trapped mode excited the same way as for Fig. 12, but all the Coriolis terms are included in the calculation. (a) Amplitudes; (b) phases. Note that the 'residual' amplitudes excited in the thermosphere are larger than those of Fig. 2.

and the asymptotic amplitudes in the upper thermosphere are about 2.5 times larger than those of Fig. 12. Even so, they are still negligible.

Applying our model to a more realistic situation, we compute the diurnal component of thermal tides under equinox conditions, but without the consideration of background winds. The driver includes excitation by heat absorption in the water vapor layer near the surface and in the ozone layer at 50 km (FORBES and GARRETT, 1976), as well as EUV heating in the lower thermosphere. Using the matrix inversion technique described by equation (22), the solution is computed by solving simultaneously for 10 spherical harmonics. In the formulation described by equation (22), all the coefficients of the matrix on the left-hand side have to be latitude-independent. For this reason, we treat the ion drag in a way similar to the Rayleigh friction; the drag coefficient is approximated by an 'averaged' value independent of latitude (though dependent on altitude). This introduces errors in the relative amplitudes and phases in the tidal winds at higher altitudes, but only above 120 km. The latitudinal dependence of the ion drag as well as those of other processes can be taken care of in real space through the N operator in equation (21), but that is postponed to a later simulation.

Figure 14 shows the computed zonal winds associated with the diurnal tide. The same parameterizations for the eddy diffusion and Newtonian cooling are used as for the calculation of the zonally averaged circulation [Fig. 5(a),(c)]. However, no Rayleigh friction is included here as the effects of gravity wave breaking on tides are considerably complicated by the tidal winds (WALTERSCHEID, 1981; FRITTS and VINCENT, 1987). Qualitatively, our results are similar to those obtained by other workers in the field (e.g. FORBES, 1982; VIAL, 1986; TEITELBAUM *et al.*, 1989; MIYAHARA and WU, 1989). In particular, the phase progression shows very good agreement with that obtained by FORBES (1982). The magnitude of the zonal wind at 90 km, where observations concentrate, is around 35 m s^{-1}, very close to those obtained recently by VIAL (1986) and FORBES and HAGAN (1988).

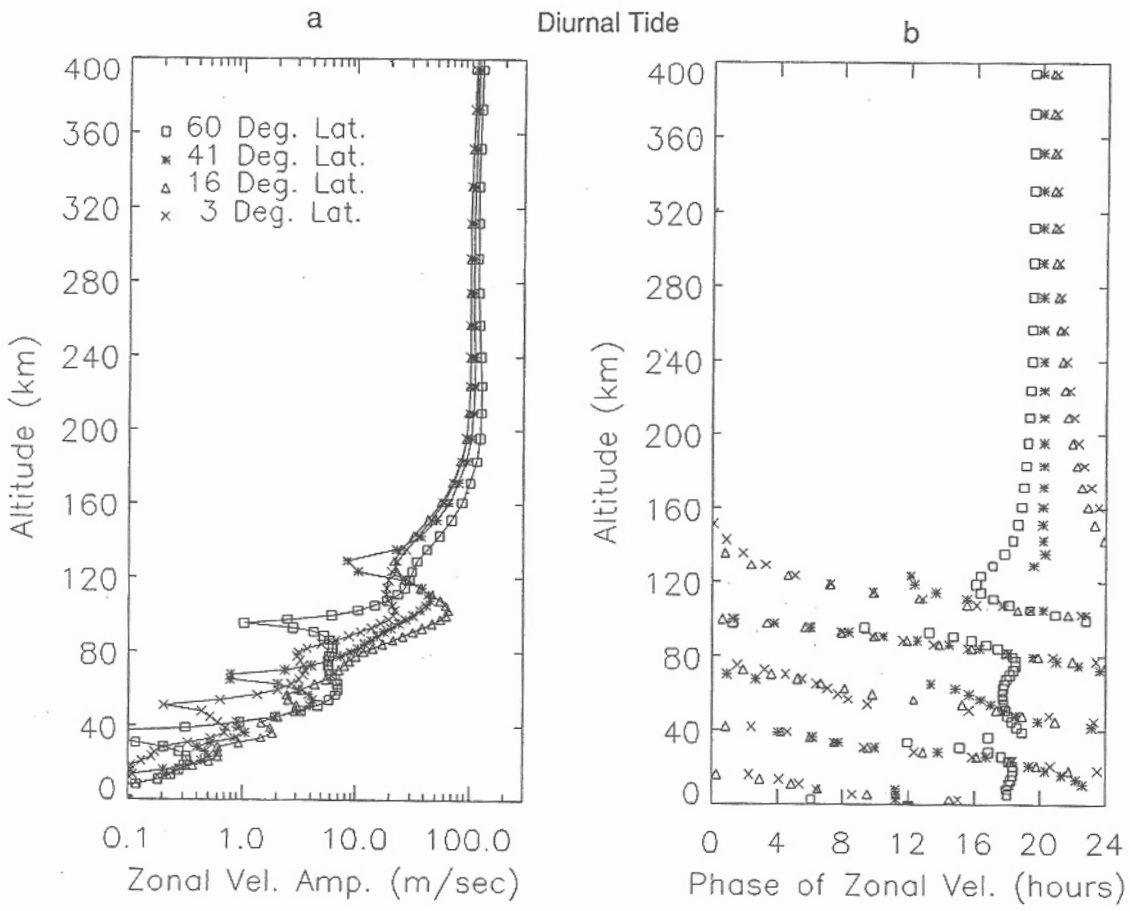

Fig. 14. Diurnal component of thermal tides under equinox conditions, without the consideration of background winds. (a) Amplitudes; (b) phases.

It is also possible to obtain the tidal solution by integrating equation (21) in time. However, to expedite convergence, it is necessary to implement a 'Rayleigh friction' type damping in the momentum equations. For a damping rate of 0.03 day^{-1}, the convergence can be achieved in about 30 days. The results are shown in Fig. 15. The differences between Fig. 14 and Fig. 15 are about 20% in amplitude and 1/2 h in phase. Better agreement can be obtained by reducing the artificial damping successively.

Background winds and temperature variations can affect the propagation of tides and induce coupling between different Hough modes (WALTERSCHEID and VENKATESWARAN, 1979). The background winds introduce extra advection terms and temperature variations introduce extra temperature gradient terms in the tidal equations. Many authors (e.g. FORBES, 1982; VIAL, 1986) solve the system of equations by finite difference approximations, with winds limited to the zonal component. In our present approach, the mean circulation (including the zonal mean component of temperature perturbation) acts on the diurnal tides through the interaction term N in equation (21) which can be specialized for the $m = 1$ case as:

$$[1-\beta\Delta t(W+D)](u_1^{q+1}-u_1^q) = \Delta t\{(W+D+C)u_1^q + N_1[u_0^q, u_1^q] + Q_1(t^q)\} \quad (23)$$

(the subscript denotes the zonal wave number). We can include the effects of all components of background winds and temperature gradients. This is realized through calculating each of the dynamical variables in equations (13)–(16) with two components: (i) the fixed $m = 0$ component taken from the solution for the mean circulation obtained previously (Figs 8 and 9), and (ii) the $m = 1$ component that changes with time. Since the non-linearity is quadratic and the transforms are alias-free, the self-interaction of the $m = 1$ components (creating $m = 0, 2$ components) is automatically filtered out. Only products of the form u_0u_1 contribute to the right hand side of equation (21). In spectral space, N_1 is in fact a linear function of u_1.

The spectral transforms required for marching

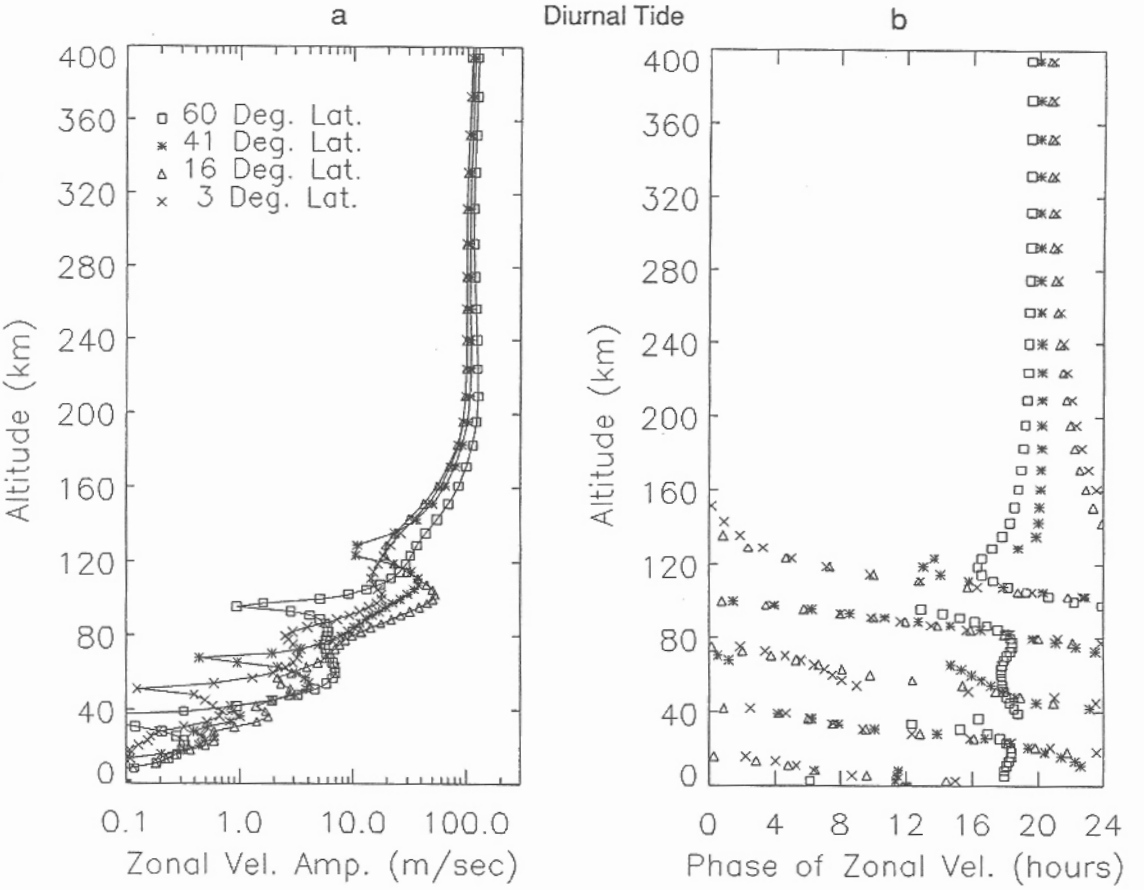

Fig. 15. Diurnal tides computed [by equation (22)] with the addition of a small 'background' Rayleigh friction ($0.03\,\text{day}^{-1}$) which is useful for accelerating the convergence of time-dependent calculations [equation (23)]. (a) Amplitudes; (b) phases.

equation (23) in time consume a considerable amount of computer time (for 10 spherical harmonics, 226 vertical levels and 15 min time steps, it typically requires 8 Micro-Vax hours to complete a 30 day run). However, the time-marching introduces extra flexibility to the system and makes it possible to study transient phenomena associated with the variability of the tides. In our study, a baroclinic type instability is discovered near the summer pole of the mesopause region. This is due to the existence of a very large horizontal temperature gradient created by the mesospheric temperature anomaly. Time series of the zonal winds under solstice condition (southern summer) are plotted for the latitudes 85°S, 60°S, 41°S, and 16°S (at 85 km altitude and longitude 0) in Fig. 16. Starting with zero amplitude at day 0, the diurnal tide is already well developed at day 20, but at the higher latitudes, the winds show extra four-day oscillations with growing amplitudes. The e-folding time for the growth of this oscillation is about 10 days. The horizontal pattern of this growing mode is a two-cell vortical structure around the summer pole and the vertical wavelength of this structure is about 33 km. A contour plot of the $m = 1$ amplitude of the zonal wind at day 50 is shown in Fig. 17. For comparison, Figs 18 and 19 present results without influence from the mean circulation. Here, all the time series are periodic (1 day period) and the amplitudes are stationary. The contour plot shows no vortex in the south pole region. Except for the south pole region, Figs 17 and 19 are quite similar.

The instability at this summer pole can be controlled by a locally enhanced diffusivity ($\sim 3 \times 10^6\,\text{cm}^2\,\text{s}^{-1}$ in the region 75–95 km altitude and 90–60°S latitude) which would occur as stronger turbulence is generated. At later times (> two months), the winter pole also shows a tendency to develop a baroclinic type instability. However, since the solstice condition is not truly perpetual, there may not be enough time for this instability to develop. We plan to study this problem by implementing a seasonally varying driver for the tides.

5. SUMMARY

We have developed a spectral model to study the dynamics of the middle and upper atmosphere. Based

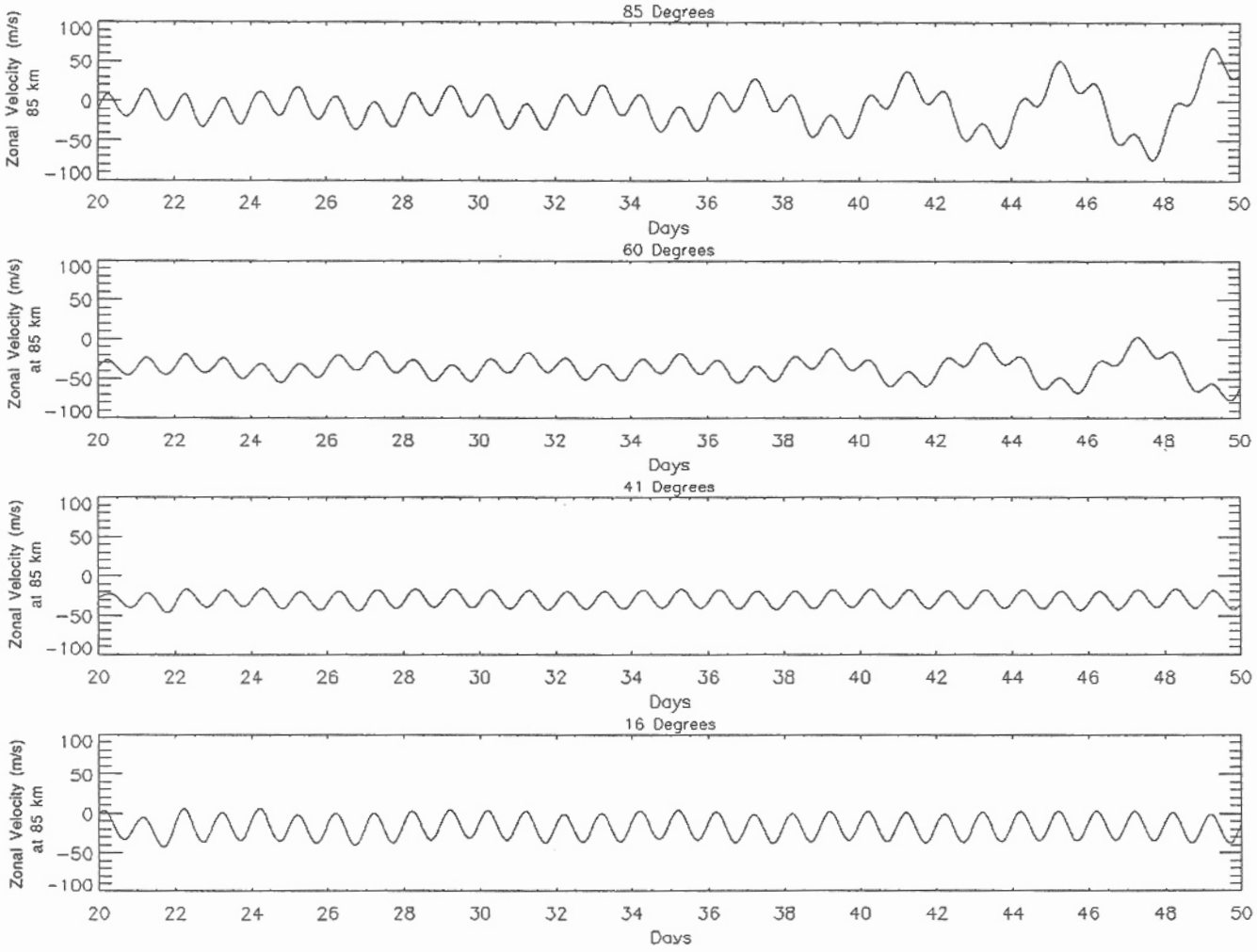

Fig. 16. Time series of zonal velocities at an altitude of 85 km and 85°S, 60°S, 41°S, 16°S (from top to bottom) for a diurnally excited tide under the influence of a background circulation. At lower latitudes, the velocities are purely periodic (1-day period). At higher latitudes, however, 4-day waves with growing amplitudes are superimposed.

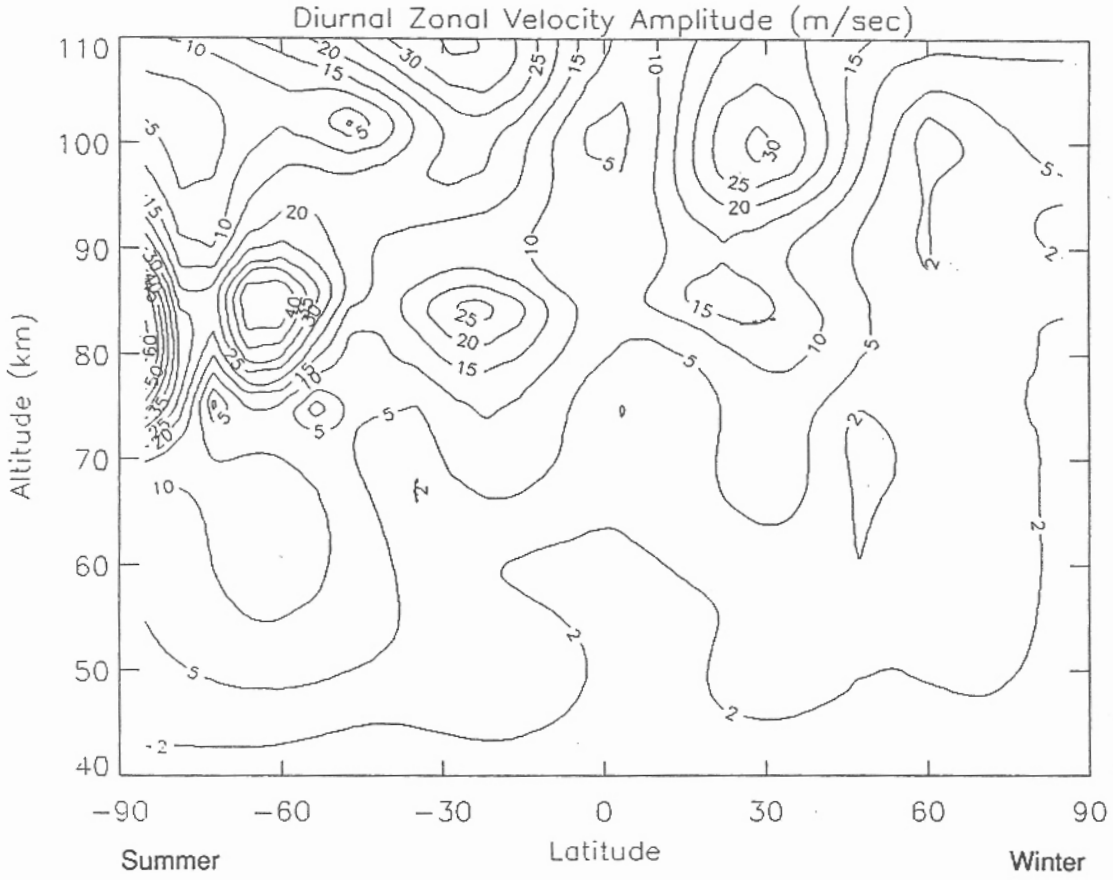

Fig. 17. Distribution of zonal wind amplitude for the tide of Fig. 16 at day 50.

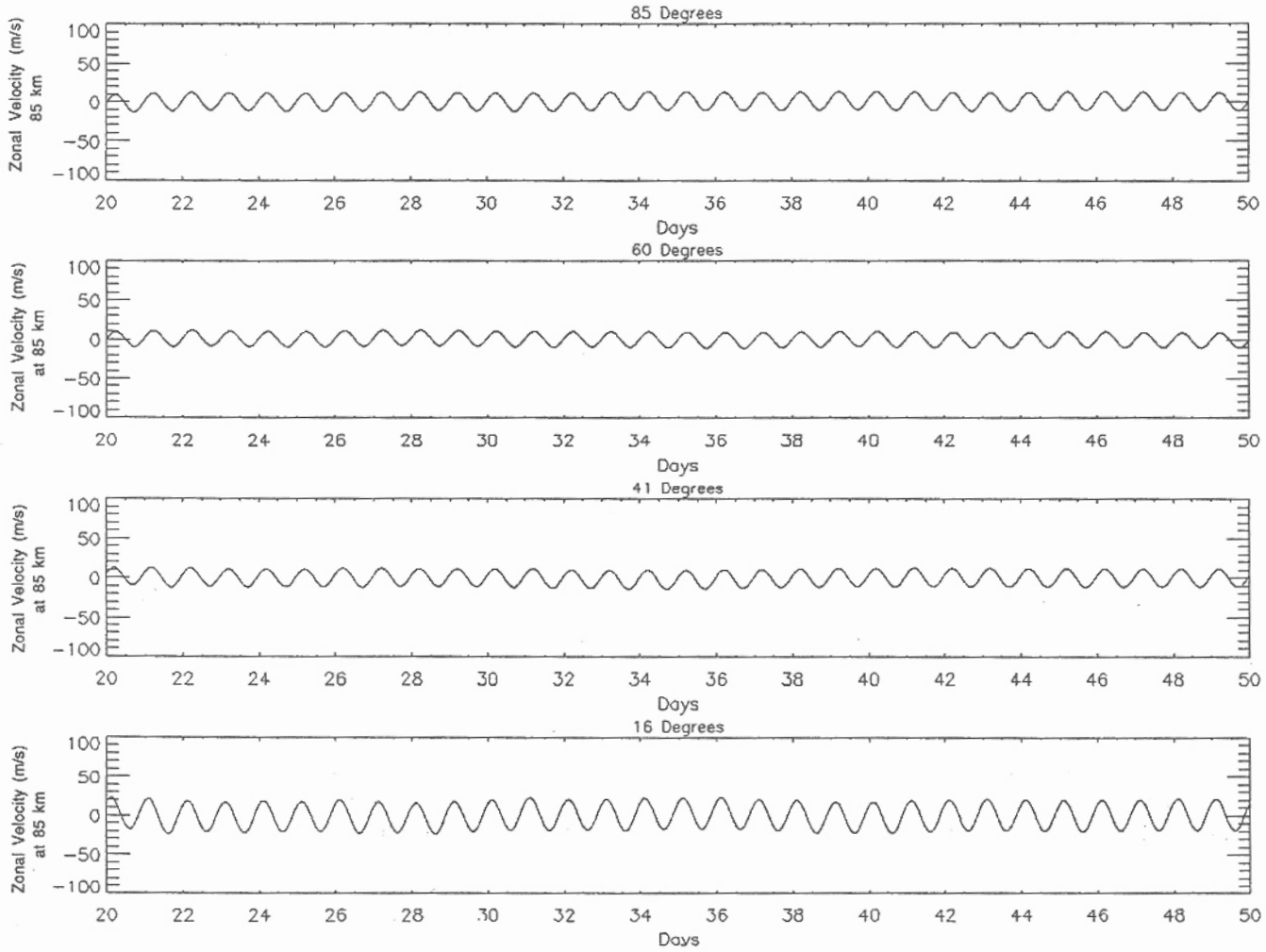

Fig. 18. Time series for a diurnal tide without background circulation.

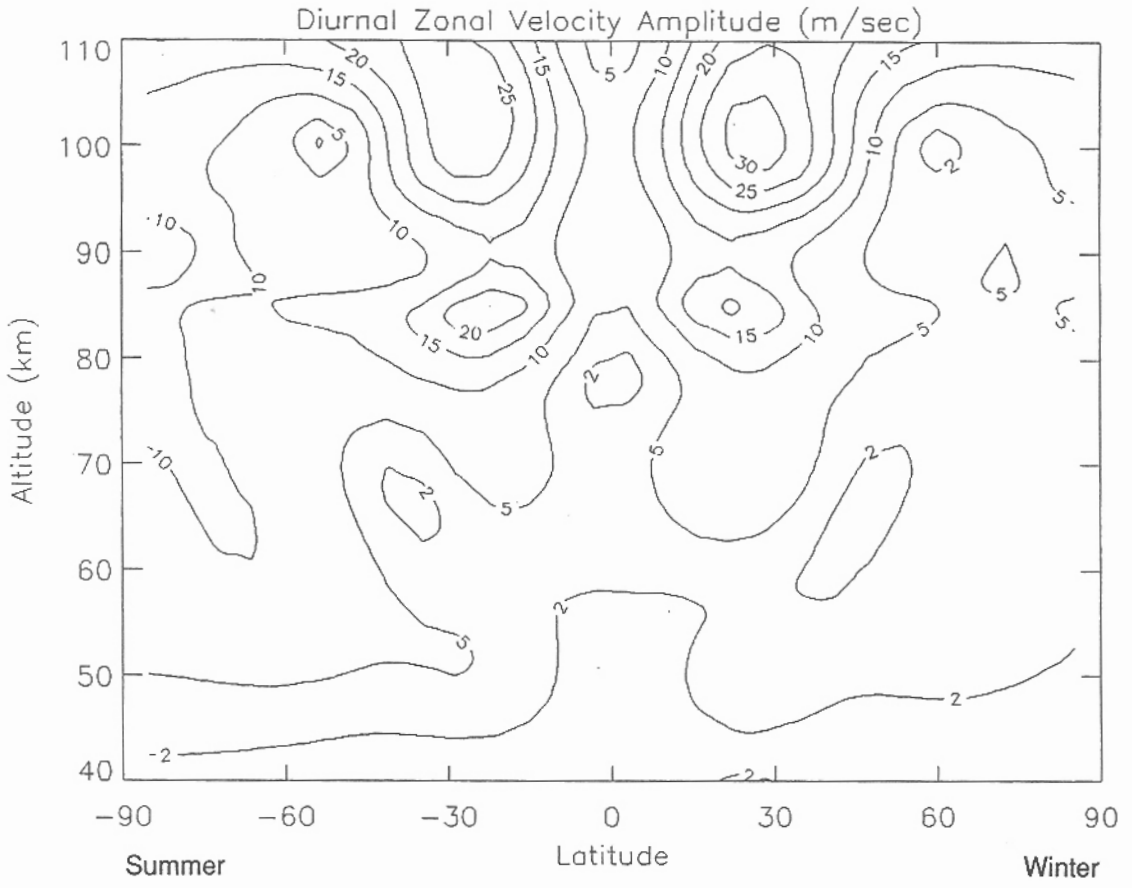

Fig. 19. Distribution of the zonal wind amplitude for the tide of Fig. 18 at day 50.

on a three-dimensional spectral numerical model, we derive 'submodels' which are specially suited to study features with particular zonal wave numbers. Having the same root, these submodels can be assembled consistently to form more complex models in which the interactions and cause–effect relationships of the different components can be examined. This methodology has been applied to study the mean circulation, the diurnal tides and the influence of the mean circulation on the diurnal tides. First, the mean circulation and the tide are treated separately using both a single step matrix inversion technique and a semi-implicit time marching technique. The results are in good agreement with those obtained previously by other authors. Both numerical techniques yield essentially identical results, showing the dynamical consistency of the two approaches. The time-marching technique generally uses more computer time, but it allows for the inclusion of non-linear interactions and provides the flexibility to account for complications associated with the latitudinal (as well as longitudinal) variations of parameters and parameterizations. This capability is essential for our second exercise—simulating the effects of latitude dependent background winds and temperature gradients on diurnal tides at solstice. Our time-dependent model shows that the large horizontal temperature gradient at the summer mesopause region can drive a baroclinic instability, generating waves with approximately 4 day period. The observational implications of these waves are intriguing, but we shall postpone that to a more comprehensive later study.

Acknowledgement—We would like to express our gratitude to Dr C. A. Reddy for helpful discussions.

REFERENCES

AKMAEV R. A., FOMICHEV V. I., GAVRILOV N. M. and SHUED G. M. 1992 *J. atmos. terr. Phys.* **54**, 119.

Arakawa A.	1988	*Physically-Based Modeling and Simulation of Climate and Climate Change*, M. E. Schlesinger (ed.), p. 79. Kluwer, Dordrecht.
Banks P. M. and Kockarts G.	1973	*Aeronomy*, part B. Academic, New York.
Bourke W.	1974	*Mon. Weath. Rev.* **102,** 687.
Chan K. L. and Mayr H. G.	1990	*EOS* **71,** 549.
Chan K. L. and Mayr H. G.	1991	*The Sun and Cool Stars: Activities, Magnetism, Dynamos*, I. Tuominen, D. Moss and G. Rudiger (eds), p. 178. Springer, Berlin.
Chan K. L., Mayr H. G., Mengel J. G. and Harris I.	1993	*J. Comput. Phys.* (in press).
Chao W. C. and Schoeberl M. R.	1984	*J. atmos. Sci.* **41,** 1893.
Chapman S. and Lindzen R. S.	1970	*Atmospheric Tides*. D. Reidel, Dordrecht.
Dickinson R. E.	1973	*J. geophys. Res.* **78,** 4451.
Dickinson R. E., Ridley E. C. and Roble R. G.	1984	*J. atmos. Sci.* **41,** 205.
Dunkerton T. J.	1978	*J. atmos. Sci.* **35,** 2325.
Fesen C. G., Dickinson R. E. and Roble R. G.	1986	*J. geophys. Res.* **91,** 4471.
Forbes J. M.	1982	*J. geophys. Res.* **87,** 5222.
Forbes J. M. and Garrett H. B.	1976	*J. atmos. Sci.* **33,** 2266.
Forbes J. M. and Hagan M. E.	1988	*Planet Space Sci.* **36,** 579.
Fritts D. C. and Vincent R. A.	1987	*J. atmos. Sci.* **44,** 605.
Fuller-Rowell T. J. and Rees D.	1980	*J. atmos. Sci.* **37,** 2545.
Garcia R. R. and Solomon S.	1983	*J. geophys. Res.* **88,** 1379.
Garcia R. R. and Solomon S.	1985	*J. geophys. Res.* **90,** 3850.
Haltiner G. J. and Williams R. T.	1980	*Numerical Prediction and Dynamical Meteorology*. John Wiley and Sons, New York.
Hide R.	1969	*J. atmos. Sci.* **26,** 841.
Holton J. R.	1982	*J. atmos. Sci.* **39,** 791.
Holton J. R. and Wehrbein W. M.	1980	*Pure appl. Geophys.* **118,** 284.
Kellogg W. W.	1961	*J. Met.* **18,** 373.
Lapidus L. and Seinfeld J. H.	1971	*Numerical Solution of Ordinary Differential Equations*. Academic Press, New York.
Lindzen R. S.	1981	*J. geophys. Res.* **86,** 9707.
Machenhauer B.	1979	*The Spectral Method*, GARP Publication Series No. 17, Vol. II, pp. 121–275.
Matsuno T.	1982	*J. met. Soc. Japan* **60,** 215.
Mayr H. G. and Harris I.	1978	*J. geophys. Res.* **83,** 3327.
Miyahara S. and Wu D. H.	1989	*J. atmos. terr. Phys.* **51,** 635.
Murgatroyd R. J. and Goody R. M.	1958	*Q. Jl R. Met. Soc.* **84,** 224.
Ogura Y. and Charney J. G.	1962	Proceedings of the International Symposium Numerical Weather Prediction, p. 431, Tokyo.
Ogura Y. and Phillips N. A.	1962	*J. atmos. Sci.* **19,** 173.
Roble R. G. and Ridley E. C.	1987	*Ann. Geophys.* **5A,** 369.
Rosenfield J., Schoeberl M. R. and Geller M. A.	1987	*J. atmos. Sci.* **44,** 859.
Schoeberl M. R.	1988	*J. atmos. Sci.* **45,** 980.
Simmons A. J. and Bengtsson L.	1988	*Physically-Based Modeling and Simulation of Climate and Climate Change*, M. E. Schlesinger (ed.), p. 23. Kluwer, Dordrecht.
Strobel D. F.	1978	*J. geophys. Res.* **83,** 6225.
Teitelbaum H., Vial F., Manson A. H., Giraldez R. and Massebeuf M.	1989	*J. atmos. terr. Phys.* **51,** 627.
Vial F.	1986	*J. geophys. Res.* **91,** 8955.
Volland H.	1988	*Atmospheric Tidal and Planetary Waves*. Kluwer, Dordrecht.
Walterscheid R. L.	1981	*J. geophys. Res.* **86,** 9698.
Walterscheid R. L. and Venkateswaran S. V.	1979	*J. atmos. Sci.* **36,** 1623.
Wehrbein W. M. and Leovy C. B.	1982	*J. atmos. Sci.* **39,** 1532.
Williamson D. L.	1983	Description of NCAR Community Climate Model (CCMOB), NCAR/TN-210+STR.